2006

Quantiative biological microsocopy by digital holography

Christopher J. Mann
University of South Florida

Quantitative Biological Microscopy by Digital Holography

by

Christopher J. Mann

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Physics
College of Arts and Sciences
University of South Florida

Major Professor: Myung K. Kim, Ph.D.
Dennis K. Killinger, Ph.D.
Chun-Min Lo, Ph.D.
George S. Nolas, Ph.D.

Date of Approval:
July 14, 2006

Keywords: computer holography, holographic interferometry , interference microscopy , numerical reconstruction, phase-unwrapping, phase-contrast microscopy

Dedications

To all my family and friends who have supported me on my journey.

Acknowledgments

I would like to first express my sincere appreciation to Dr. Kim for his support and guidance during this project and for his continued patience and encouragement throughout my study. He is an outstanding advisor.

I would also like to thank all my committee members Dr. Dennis Killinger, Dr. Chun-Min Lo, and Dr. George Nolas for providing technical advice and serving on my dissertation and candidacy exam committee. I would also like to thank Dr. Shekhar Bhansali for chairing my defense examination.

I would also like to convey my thanks to my fellow lab members Leo Krzewina, Alex Khmaladze, Nilanthi Warnasooriya, and Mariana Potcoava as well as Dr. Lingfeng Yu who helped me to build this work.

Also I must say a huge thank you to Anali Makoui who has provided me which much impetus and gave so much valuable support and encouragement throughout this project. Also I acknowledge the numerous helpful discussions with Avis Pal, Ed Navarro, and Yaroslav Shtogun.

I would like to thank finally the staff of the Physics department for their help, Sue Wolfe, Evelyn Keeton-Williams, Sam Valenti and Phil Bergeron.

Table of Contents

List of Tables

List of Figures

Quantitative Biological Microscopy by Digital Holography

Christopher J. Mann

ABSTRACT

In this dissertation, improved techniques in digital holography, that have produced high-resolution, high-fidelity images, are discussed. In particular, the angular spectrum method of calculating holographic optical field is noted to have several advantages over the more commonly used Fresnel transformation or Huygens convolution method. It is observed that spurious noise and interference components can be tightly controlled through the analysis and filtering of the angular spectrum. In the angular spectrum method, the reconstruction distance does not have a lower limit, and the off-axis angle between the object and reference waves can be lower than that of the Fresnel requirement, while still allowing the zero-order background to be cleanly separated. Holographic phase images are largely immune from the coherent noise commonly found in amplitude images. With the use of a miniature pulsed laser, the resulting images have 0.5μm diffraction-limited lateral resolution and the phase profile is accurate to about several nanometers of optical path length. Samples such as ovarian cancer cells (SKOV-3) and mouse-embryo fibroblast cells have been imaged. These images display intra-cellular and intra-nuclear organelles with clarity and quantitative accuracy. This technique clearly exceeds currently available methods in phase-contrast optical

microscopy in both resolution and detail and provides a new modality for imaging morphology of cellular and intracellular structures that is not currently available.

Furthermore, we also demonstrate that phase imaging digital holographic movies provide a novel method of non-invasive quantitative viewing of living cells and other objects. This technique is shown to have significant advantages over conventional microscopy.

Chapter 1

Introduction

Holography, formed from the Greek words *Holo* (whole), and *graphe* (write), is a method of recording both the intensity and phase information of a wave-field. It was first suggested by Dennis Gabor in 1948 while working to improve upon the resolution of electron microscopy [1], which suffers from limiting aberrations. Since its discovery, holography has become an active field of research in modern optics and is now an established tool for scientific and engineering studies.

This chapter presents an introduction to digital holography and the motivations behind our research, in the following order: Section 1.1 discusses the foundations of holography and its development in a historical context. Section 1.2 describes the advent of digital holography, which has seen rapid progression due to advances in computers and technology. Its applications to microscopy are investigated in Section 1.3. Section 1.4 summarizes research contributions and motivations. Finally Section 1.5 outlines the organisation of this thesis.

1.1 Foundations of holography

An object illuminated with a coherent light source reflects or transmits the light wave, which consequently carries information about the object's physical and optical properties. The intensity or amplitude information of the wave is associated with the reflectance or attenuation property of the object while the phase information is related to

the topography or thickness. Recording devices such as Charged-Coupled Device (CCD) cameras and photographic plates are able to capture the intensity of the incoming light, but not the phase information. As an inherent three-dimensional technique, holography presents a means of encoding the phase information as an intensity variation termed a "hologram", which can then be recorded by devices sensitive only to the intensity component of light.

A hologram is defined as the recorded interference pattern between a coherent object wave transmitted or reflected by an object and a reference wave which interferes directly with this wave at the recording medium. The hologram contains information about the entire three-dimensional wave-field (amplitude and phase) which is encoded in the form of interference lines at high spatial frequencies. In conventional holography the recorded object wave is reconstructed optically by illuminating the chemically processed hologram with an optical replica of the reference wave. This produces a reconstructed object which is indistinguishable from the original object and contains the 3D information, both amplitude and phase.

In the original experimental set-up of Gabor, both the object wave and reference wave are located along the optical axis normal to the recording medium. This leads to a reconstructed image superimposed by the bright reconstruction wave called the *dc* term or zero order and a second component, the so-called 'twin or conjugate image'. These unwanted components obscure the true amplitude and phase of the reconstructed object wave.

While the basic foundations of holography were laid down by Gabor, it was not until the 1960s, that holography gained widespread interest and practical use. This was

due to the availability of the laser and the introduction of the off-axis technique by Leith and Uptanieks [2], in which the twin images and the *dc* term are spatially separated by the introduction of an angle between the object and reference waves. In the late 1960`s the first paper on Holographic Interferometry (HI) was published by Stetson, Powell *et al* [3]. Holographic Interferometry, was designed as a non-contact optical test tool for deformation studies of strains, vibrations and heating, and enabled non-destructive measurement of surfaces with an accuracy of less than a micron.

Other significant advances in holography made around the same time were the creation of computer generated holograms using Fraunhofer diffraction. Numerical methods for reconstruction of the hologram were initially proposed by Goodman *et al* [4] and Kronrod *et al* [5]. These methods included digitizing holograms recorded on photographic plates which were then reconstructed numerically.

1.2 Digital holography

The conventional process of performing holography using photographic plates is time consuming and cumbersome. This has precluded the widespread use of conventional holography except in the case where measurements are considered important enough to justify the time and expense involved. Also real time processing of a conventional hologram is not feasible unless one uses photo-refractives and other nonlinear optical materials. The advances in digital imaging and computation technologies have now made it feasible and advantageous to replace the photochemical processing of conventional holography with CCD arrays and numerical computation [6, 7]. With the development of higher performance CCD and computational techniques, digital holography is fast

becoming an increasingly attractive alternative to conventional film-based holography. It offers a number of significant advantages, such as simple, fast image acquisition and the availability of many powerful digital processing algorithms. By calculating the complex optical field of an image volume, the amplitude and phase of the optical field are simultaneously available for direct manipulation [8-12].

There are numerous digital processing techniques for manipulating the optical field information in ways that are difficult or impossible in real space processing. For example, optical system aberration can be numerically corrected [13,14] and holographic interferometry can be performed between remotely situated objects through telecommunication links [15].

Figure 1.1 illustrates the digital holography process. A digital hologram is created by the interference between a coherent object and a reference beam which is recorded by a CCD camera and digitally transferred to a computer, where it is processed by computational methods to obtain the holographic images. The digital hologram contains not only amplitude information of the object, but also phase.

Moreover, the ability of the CCD camera to quantify the recorded light gives rise to a number of post processing methods that can for instance be used to obtain quantitative three-dimensional topography of an object surface from the phase information. The fact that quantities such as the phase of the wave cannot be measured directly with conventional detectors, which are only sensitive to the wave-field's intensity, makes digital holography a particularly attractive technique for measuring morphological properties of an object.

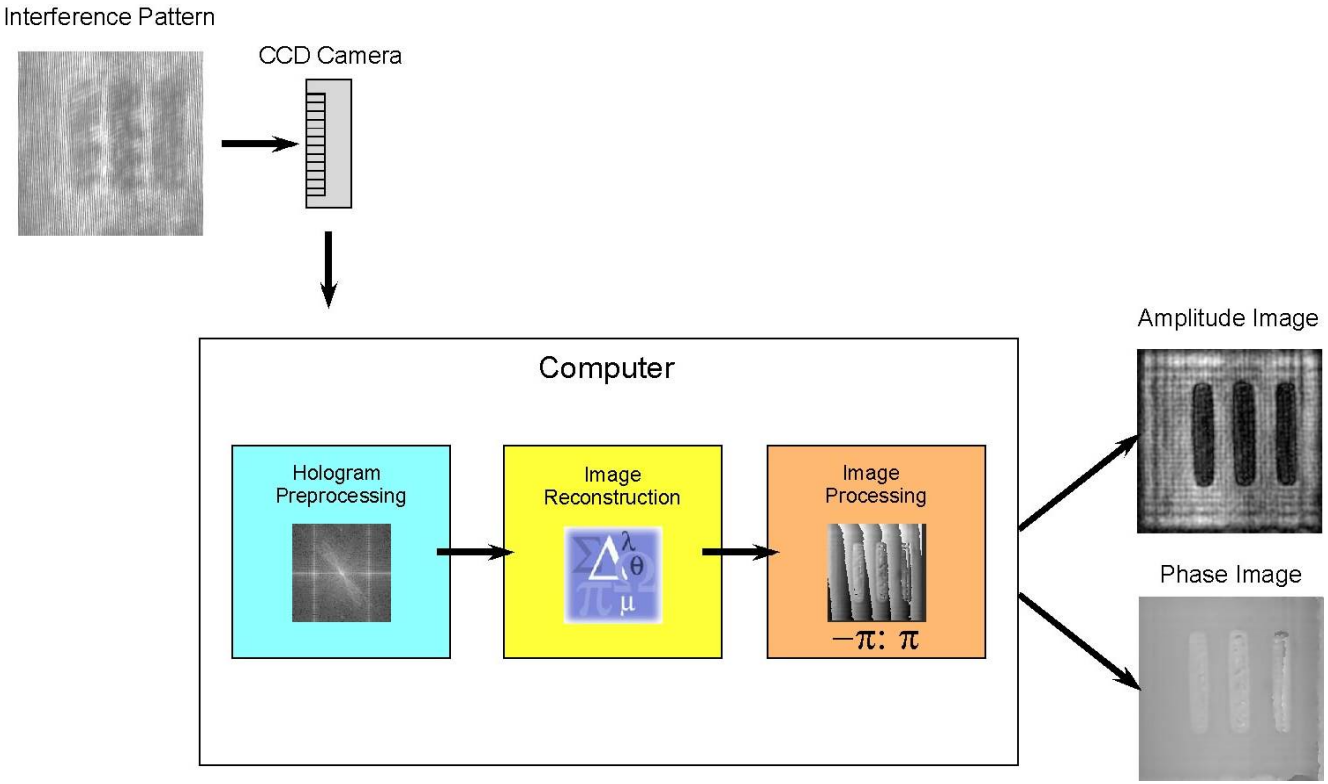


Figure 1.1: The process of digital holography

1.3. Applications to microscopy

Digital holography has been applied in diverse fields, including metrology [16], deformation measurement [17], vibrational analysis [18], art conservation [19], microstructures [20, 21] and more recently biological microscopy [22-27].

The applications to microscopy are particularly appealing due to a number of technical advantages.

1.3.1 Three-Dimensional Microscopy

One of the main drawbacks in conventional optical microscopy is the small depth of focus at higher spatial resolution, due to the objective's large numerical aperture and magnification. Only two dimensional information of an object can be obtained. More specifically the diffraction-limited depth of focus is given by

$$\Delta z_{focus} = \frac{n\lambda}{NA^2} \tag{1.1}$$

where n is the index of refraction of the medium surrounding the object, λ is the illumination wavelength, and NA is the numerical aperture of the focusing lens or the microscope objective.

Common ways to circumvent the problem of the low depth of focus in conventional microscopy include sectioning of the sample or a mechanical motion along the optical axis to scan the complete experimental volume. However, physical sectioning of a biological sample is invasive and the mechanical operation of scanning increases the time for acquisition of three-dimensional images. These are constraints for three-dimensional imaging of a live sample.

This problem can be resolved using digital holographic microscopy. The information on a complete volume can be recorded in just a single hologram and optically reconstructed with visual 3-D observation.

1.3.2 Phase Contrast Microscopy

As well as the limited depth of focus, conventional bright-field microscopes have difficulty in observing transparent samples such as living biological cells which exhibit little intensity contrast. The conventionally applied phase contrast techniques of Zernike and differential interference contrast (DIC) microscopy of Nomarksi that have been developed for visualization of unstained transparent specimens do not offer direct quantitative evaluation of the phase information.

The unavailability of quantitative phase information in these techniques presents a difficulty in observing and interpreting morphological changes and properties of a sample. Quantitative phase imaging is particularly important in microscopy because it enables determination of either the refractive index or optical thickness variations from the measured optical path length with an accuracy of a few nanometers.

Unlike the above-mentioned conventional phase-contrast microscopic techniques, interferometric techniques allows for quantitative measurement of the phase information produced by the object. In addition, digital holography as an interferometric technique offers quantitative phase information and high fidelity and high resolution images with a precision of optical thickness in the order of tens of nanometers [28]. While other interferometric configurations such as interference microscopy are commonly used in metrology, very few applications have been reported in biological microscopy. These

techniques often require the phase to be measured through the process of multiple image acquisition and phase modulation. Digital holography on the other hand requires only a single image (hologram) and no phase modifying devices in order to obtain phase information. This is a significant advantage for real-time sample analysis.

A common presumption is that coherent imaging suffers from the image degrading effect of coherent noise; however, through careful control of laser beam and other optics quality, remarkably clean images can be obtained. This is especially true with phase imaging in digital holography because of its relative immunity to the amplitude or phase noise of the laser profile.

1.3.3 Numerical Focusing

Another unique capability of digital holography is numerical focusing, which emulates the focusing control of conventional microscopes. As the focusing can be adjusted in the reconstruction process, digital holography is free from the process of mechanical focusing and can be used to monitor the dynamic change of objects.

A number of algorithms have been developed for use in the reconstruction of a hologram such as the Fresnel transform, convolution and angular spectrum methods [28-31]. In particular, the use of the angular spectrum reconstruction algorithm provides a significant advantage in focusing and reconstruction [32]. It has no minimum distance requirement from the object plane to the hologram plane, and allows for flexible and effective filtering and control of the *dc* term and spurious noise components from sources such as stray reflections within the experiment.

1.3.4 Phase Unwrapping

From a computational point of view, phase images are directly determined from the computed field. However, the phase images are required to be unwrapped in order to simplify their interpretation. Of particular significance is the 2π-discontunuity problem in phase-imaging. A conventional approach is to apply one of many phase-unwrapping algorithms [33, 34], but often these require substantial user intervention and strict requirements on the level of phase noise and phase discontinuity.

It has long been recognized that the range of unambiguous phase measurement can be extended beyond a single wavelength by synthesizing a beat wavelength between two wavelengths [35, 36]. The 2π-discontunuities inherent in the wrapped phase image can be effectively removed without the addition of any noise, by using a multi-wavelength digital holographic approach in a microscope configuration [37]. This technique, unlike the mathematically challenging and often time consuming processes of conventional phase unwrapping algorithms, makes real time processing much simpler to perform because any calibration for removing 2π-discontinuities can be made prior to any image capture. The drawback with this method lies in the increased technical effort required for optical alignment of more than one laser (unless the same laser produces the required wavelengths) along the same optical path.

1.4 Research Contributions

As a recently developed technique, digital holography has yet to be firmly established in the microscopy world and many of its technical advantages are yet to be fully explored and examined. The motivation behind this work has been to improve and

develop techniques of digital holography in order to obtain images which effectively provide more information and clarity than images produced by other comparable microscopy techniques.

In this thesis, techniques of digital holography have been vastly improved to obtain high-resolution, high-fidelity, quantitative phase-contrast images of microscopic samples. The level of resolution and details in these images clearly exceed currently available techniques in phase-contrast optical microscopy and provides a new modality for imaging the morphology of cellular, intracellular and intra-nuclear structures that is not currently available with non-invasive optical methods. Furthermore for the first time to our knowledge, quantitative, dynamic effects have been measured on living biological samples using digital holography [38].

This research will help with precise measurements of such properties as the physical thickness of a cell, to a vertical resolution of just a few nanometers. This might be of significance to biologists who can then accurately investigate the physical morphology of specimens like cancer cells, which are known to have different thickness and other morphological properties as compared to healthy cells [39].

We also demonstrate the use of the angular spectrum reconstruction method for digital holography, which is seen to solve some of the significant problems that have so far prevented the wider use of digital holography in biomedical imaging applications. Spurious noise components can be easily controlled and there is no minimum reconstruction distance as with other methods for numerical reconstruction. Also the *dc* term of diffraction and the conjugate image can be eliminated by an appropriate selection of spatial frequencies in the angular spectrum of an off-axis hologram.

Finally we show that multi-wavelength phase imaging digital holography is an effective and efficient technique for removal of 2π-discontunuities in the phase image by the extension of the axial range and also overcomes many of the limitations imposed by conventional, mathematical based phase unwrapping techniques.

1.5 Thesis organization

This thesis is organized into seven chapters and four appendices. The main body of content in this dissertation is contained in Chapters 2, 3, 4, 5, and 6. Chapter 2 presents some of the basic optical concepts and principles of holography, beginning with an analysis of the wave theory of light and later describing the properties and roles of effects such as interference, coherence, diffraction, and speckle in holographic recording and reconstruction. Chapter 3 presents the technique of recording digital holograms and discusses the reconstruction of the recorded optical wave field by the Fresnel transform, the convolution approach and the angular spectrum. In Chapter 4, digital Gabor holography is described and experimental results are presented which show the capabilities for microscopy. Chapter 5 describes methods of determining quantitative phase information and demonstrates the high quality results achievable by digital off-axis holography. Chapter 6 describes phase imaging digital holography using two or more wavelengths that achieves unambiguous phase unwrapping by an optical method. Future works and conclusions are presented in Chapter 7.

Chapter 2

Principles of Holography

This chapter introduces a number of fundamental optical principles which together form the conceptual and mathematical foundations of holography. Section 2.1 discusses the wave theory of light, in which both the intensity and phase information form the three-dimensional nature of the light wave. Section 2.2 describes the phenomena of interference which stems directly from the wave theory of light. Section 2.3 examines different aspects of coherence. Section 2.4 includes a detailed analysis of the development of models of diffraction, which provide the foundations of wave-field propagation in holography. Finally Section 2.5 introduces the theoretical background and underlying principles of holography.

2.1 Wave Theory of Light

A light wave can produce both diffraction and interference effects which form the basis of holographic recording and reconstruction. Light is an electromagnetic wave and the theory used to explain its motion is described by the wave model. The propagation of a light wave in a vacuum can be described by

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \qquad (2.1)$$

where $\vec{E}$ is the electric field and ∇^2 is the Laplace operator defined as

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.2}$$

The electric field $\vec{E}$ is a vector quantity and therefore has direction associated with it. It can oscillate in any direction perpendicular to the light waves' propagation direction. However, it frequently oscillates only within a single plane. In such cases we say that the light is linearly polarized and so the scalar wave equation can be considered where the light propagates in one direction. If the light is propagating in the *z*-direction, then

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = 0 \tag{2.3}$$

The electromagnetic wave can be considered in terms of two components, namely the real part and the imaginary part of a complex quantity.

A sinusoidal time-varying signal, ψ, at a time *t* and at a point *z* of a wave propagating in the *z* direction can be represented by

$$\psi(z,t) = Ae^{i(kz-\omega t)} \tag{2.4}$$

where *A* is the amplitude, ω is the angular frequency and *k* is the wave number. The $(kz-\omega t)$ term is the absolute phase and is dimensionless. The real part of Equation (2.4) is taken to be that of the measurable signal *E*,

$$E = \Re[Ae^{i(kz-\omega t)}] = A\cos(kz-\omega t) \tag{2.5}$$

Equation (2.4) can be rewritten as

$$\psi(z,t) = Ae^{i\phi(z)}e^{-i\omega t} = u(z)e^{-i\omega t} \tag{2.6}$$

where $u(z) = Ae^{i\phi(z)}$ is the complex amplitude of the signal and $\phi(z) = kz$ is the phase.

The only directly measurable quantity is intensity, which is proportional to the time average of the square of the electric field.

$$I = \varepsilon_0 c \left\langle E^2 \right\rangle_t = \varepsilon_0 c \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E^2 dt \tag{2.7}$$

The intensity is calculated by taking the square of the modulus of the complex amplitude. The intensity at a point *z* is then defined as

$$I(z) = u^*(z)u(z) = \left|u(z)\right|^2 = A^2 \tag{2.8}$$

2.2 Interference

The concept of interference was first introduced by Thomas Young in 1804. He noted that under appropriate conditions, light could be combined with light to create darkness.

Interference is, in essence, the superposition of two or more waves in space. The result of adding two waves of the same frequency depends on the value of the phase of the wave at the point in which the waves are added. For two sources of electromagnetic waves to interfere they must be coherent, have the same frequency and polarization, and the superposition principle must apply.

If we consider each individual wave of the form $\vec{E}_j(\vec{r},t)$, which is a solution to the wave equation, the superposition of the waves is found by

$$\vec{E}(\vec{r},t) = \sum_j \vec{E}_j(\vec{r},t) \qquad j=1, 2 \tag{2.9}$$

For two monochromatic waves with the same polarization direction and frequency, the complex amplitude of the waves are

$$A_1(x,y,z) = a_1 e^{(i\phi_1)} \tag{2.10}$$

$$A_2(x,y,z) = a_2 e^{(i\phi_2)}$$

The scalar approach can be considered in this case because the waves have the same polarization. The addition of each single wave in Equation (2.10) creates a new resultant wave with complex amplitude *A*,

$$A = A_1 + A_2 \tag{2.11}$$

From Equation (2.8), then

$$I = |A_1 + A_2|^2 = (A_1 + A_2)(A_1 + A_2)^* \tag{2.12}$$

$$= a_1^{\,2} + a_2^{\,2} + 2a_1 a_2 \cos(\phi_1 - \phi_2)$$

$$= I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\Delta\phi$$

where I_1 and I_2 are the individual intensities and the phase difference between the sources of the two waves is

$$\Delta\phi = \phi_1 - \phi_2 \tag{2.13}$$

If the phase difference $\phi_1 - \phi_2$ is constant, the two sources are mutually coherent. If the sources are mutually incoherent, then $\phi_1 - \phi_2$ varies randomly with time. If this latter condition occurs, it signifies that the average value of the $\cos\Delta\phi$ term is zero and therefore no interference exists. In effect, this explains why no interference occurs between two incoherent sources.

In Equation (2.12) the total intensity is calculated by the sum of the individual intensities I_1 and I_2 plus the addition of the last term $2\sqrt{I_1 I_2}\cos\Delta\phi$.

This last term depends on the phase difference between the waves as seen from Equation (2.13). The intensity reaches its maximum at points for which

$$\Delta\phi = 2n\pi \qquad \text{for } n = 0,1,2\ldots\ldots \qquad (2.14)$$

which represents the condition for constructive interference.

The intensity is a minimum where

$$\Delta\phi = (2n+1)\pi \qquad \text{for } n = 0,1,2\ldots.. \qquad (2.15)$$

This is the condition for destructive interference. *n* represents the interference order and is an integer. On examination of an interference pattern, one sees bright and dark fringes as a consequence of the constructive and destructive interference of the light waves.

Consider the superposition of two plane waves E_1 (dashed lines) and E_2 (solid lines) which creates an interference pattern with equidistant spacing *d* as illustrated in Figure 2.1.

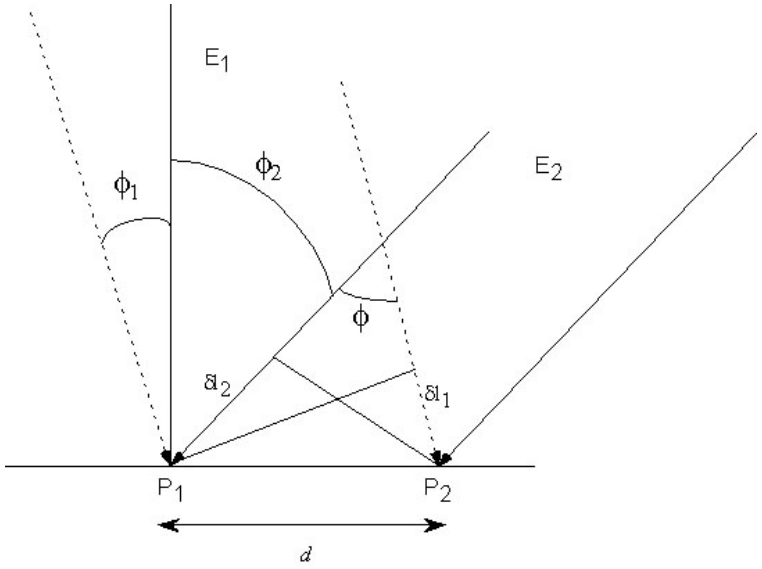


Figure 2.1: The Interference between two plane waves. (Adapted from "Digital Holography" by U. Schnars, W. Jueptner).

The distance between the interference maxima formed at points P_1 and P_2 is defined by the fringe spacing *d.* By geometry the following relations are true

$$\sin\phi_1 = \frac{\delta l_1}{d} \text{ and } \sin\phi_2 = \frac{\delta l_2}{d} \tag{2.16}$$

where ϕ_1 and ϕ_2 represent the angles between the vertical plane and the propagation direction of each individual wave-front. δl_2 is the path difference between the wave-front E_2 with respect to that of the second wave-front E_1, which is located at the position of the interference maximum P_1. In addition $-\delta l_1$ is the path difference between the wave-front E_2 with respect to E_1.

One notices in the Figure that E_2 travels a longer distance to P_1 than E_1 and conversely at P_2, the opposite is also true with E_1 traveling a further distance than E_2. It follow that the path differences between the maxima at points P_1 and P_2 is

$$\delta l_1 + \delta l_2 \tag{2.17}$$

This path difference is equal to one wavelength so that

$$\delta l_1 + \delta l_2 = \lambda \tag{2.18}$$

By combination of equations and substitution then

$$d = \frac{\lambda}{\sin\phi_1 + \sin\phi_2} = \frac{\lambda}{2\sin\frac{\phi_1+\phi_2}{2}\cos\frac{\phi_1-\phi_2}{2}} \tag{2.19}$$

By substitution of the approximation of $\cos(\phi_1 - \phi_2)/2 \approx 1$ and $\phi = \phi_1 + \phi_2$ we obtain the fringe spacing

$$d = \frac{\lambda}{2\sin\frac{\phi}{2}} \tag{2.20}$$

2.3 Coherence

Coherence is the measure of the ability of light to interfere. The coherence of two waves follows from how well correlated the waves are as quantified by the cross-correlation function and is derived from the phase relationship between two points, separated in either time or space, along a wave-train.

Consider two points along the train that are spatially separated and moving with the train. If the phase relationship between the waves at these points remains constant in time, then the waves between these points are coherent. On the other hand, if the phase relationship is random or rapidly changing, the waves at these two points are incoherent.

2.3.1 Temporal Coherence

Temporal coherence describes the correlation of a wave with itself at different instances in time and is related to the finite bandwidth of the source $\Delta\nu$. If light were strictly monochromatic then the following conditions would apply: $\Delta\nu = 0$ and $\Delta t_c = \infty$. This is not practically attainable. However, over a shorter period of time Δt_c, a wave can behave like it's monochromatic. The temporal coherence length is the greatest distance between two points for which there is a phase difference that remains constant in time. For points separated by a distance greater than the coherence length, there is no phase correlation. The coherence time Δt_c is defined as the temporal interval required for the train to traverse the coherence length when traveling in vacuum. In optics, temporal coherence is measured by combining beams from the same source but having a known path length difference, and observing the interference pattern produced.

2.3.2 Spatial Coherence

Like temporal coherence where a source is considered to be never strictly monochromatic, spatial coherence derives from the fact that a source is never truly a point source. Spatial coherence depicts the mutual correlation of different parts of the same wave-front. Thus, for two laterally displaced points residing on the same wave-front at a given time, the fields are spatially coherent.

2.4 Diffraction Theory

A mathematical description of the propagation of light waves must recognize diffraction effects. Diffraction is broadly defined as the spreading out of light from its geometrically defined path. Figure 2.2 illustrates the problem posed by diffraction. Light emitted by a source at a location *Q* falls on an opaque plate containing an aperture. The problem is then to describe the light field that will be received at point *R* to the right of the aperture. One of the earliest solutions was put forward by Huygens, whose principle underlies the idea that each point on a wave-front acts as a source of secondary wavelets. In terms of Figure 2.2 this means that that the illuminated aperture can be replaced by an array of point sources.

Huygens theory was mathematically described by Fresnel, who integrated the appropriate contributions and Kirchhoff, who put in place all the correct multiplying terms. The problem was later solved by Fresnel and Fraunhofer in the form established by Huygens, and was later solved in more detail by Köttler and Sommerfield. Köttler considered the vectorial nature of light, and Sommerfield presented a rigorous scalar derivation.

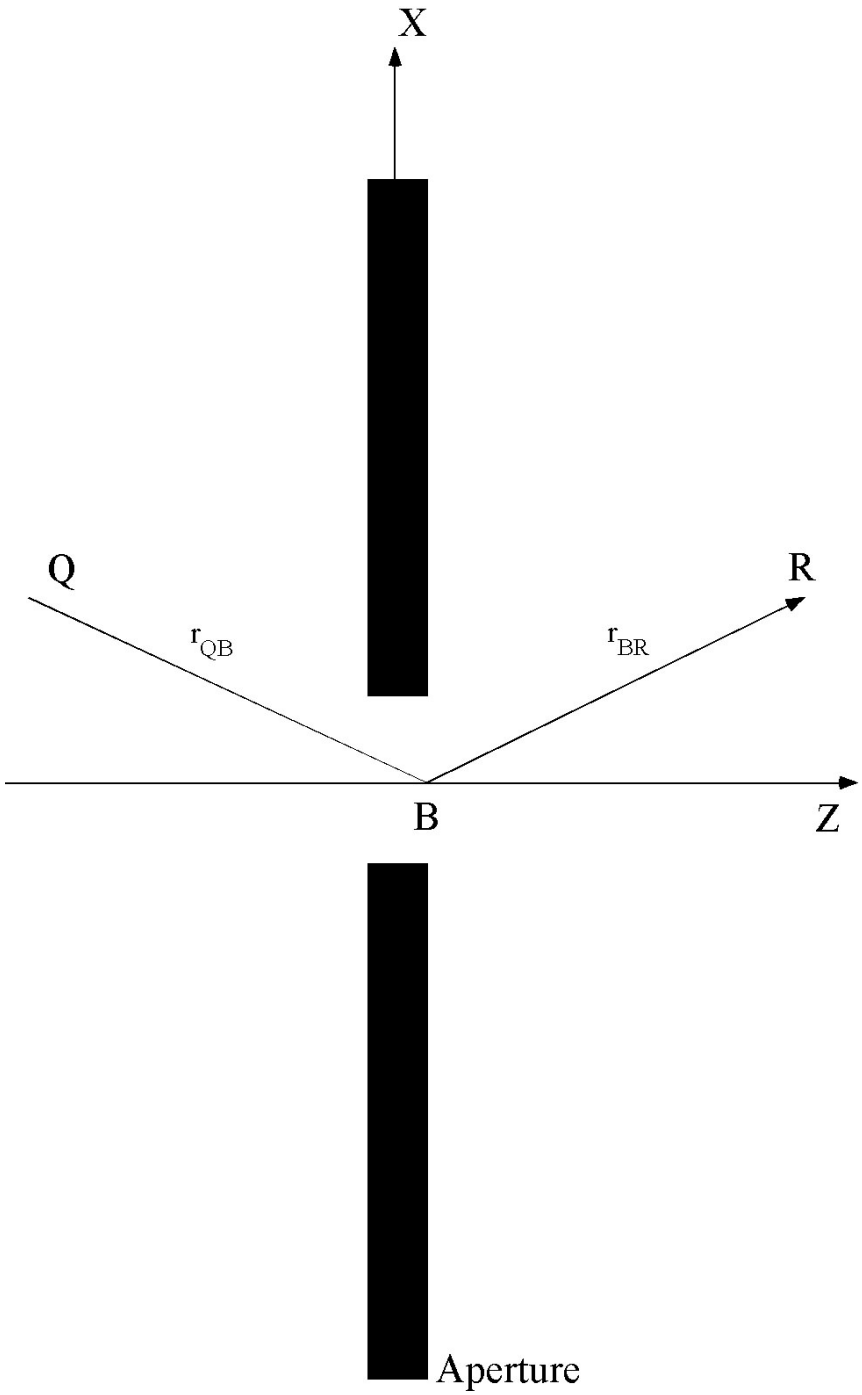


Figure 2.2: Diffraction from an aperture

2.4.1 Huygens-Fresnel Principle

Diffraction is a phenomenon by which wave-fronts of propagating waves bend in the neighborhood of obstacles. It can be described qualitatively by Huygens principle as depicted in Figure 2.3. Every point on a primary wave front serves as the source of spherical secondary wavelets, such that the primary wave front at some later time is the 'envelope' of these wavelets. Moreover, the wavelets advance with a speed and frequency equal to that of the primary wave at each point in space. By simple assumptions about the amplitude and phase of the secondary waves that form the envelopes, Huygens principle was able to accurately determine the light distribution of diffraction patterns.

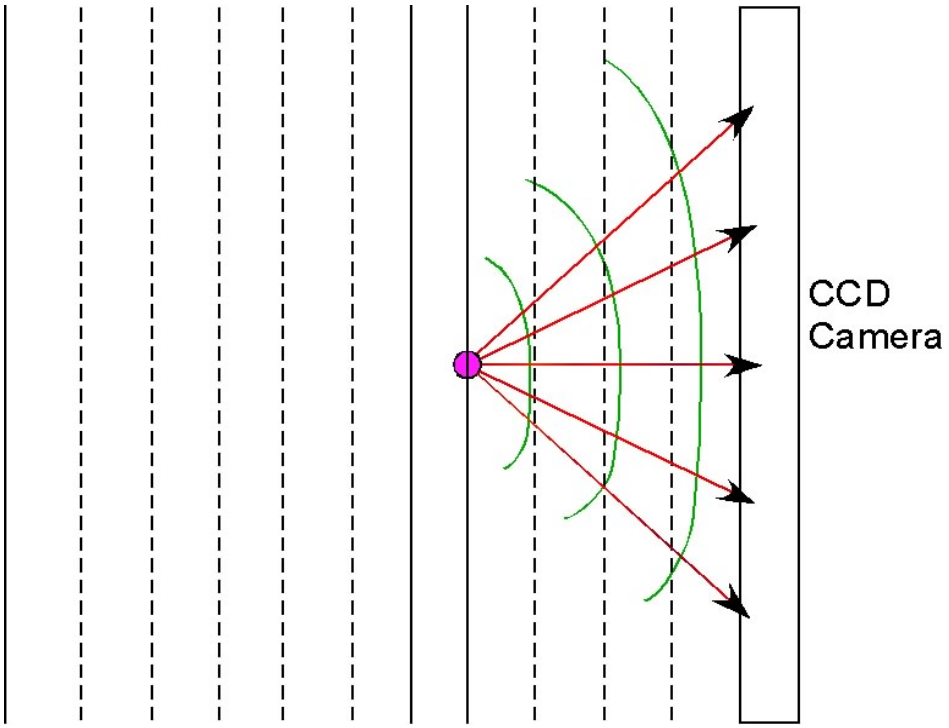


Figure 2.3: Huygens's principle. The construction of envelopes, propagating from a point source.

Huygens principle accounted satisfactorily for the reflection and refraction of light waves. However, his theory did not explain why there is no backward propagation of the waves. Furthermore, even though it predicts diffraction, it does not explain quantitatively the results obtained from diffraction experiments.

Fresnel made changes to Huygens's principle to account for the deficiency in the back wave propagation. He introduced a new term, $K(\chi)$ known as the obliquity or inclination factor, in order to solve this problem.

Figure 2.4 illustrates a geometrical representation of a spherical wave-front, arising from a point source P_0. In the figure, S represents the instantaneous position of the wave-front with a radius r_0 and a frequency ω. P is a point at which the light disturbance is to be determined. Omitting the periodic factor $e^{-i\omega t}$, the disturbance at point T on the wave-front may be represented by Ae^{ikr_0}/r_0, where A is the amplitude at a unit distance from the source.

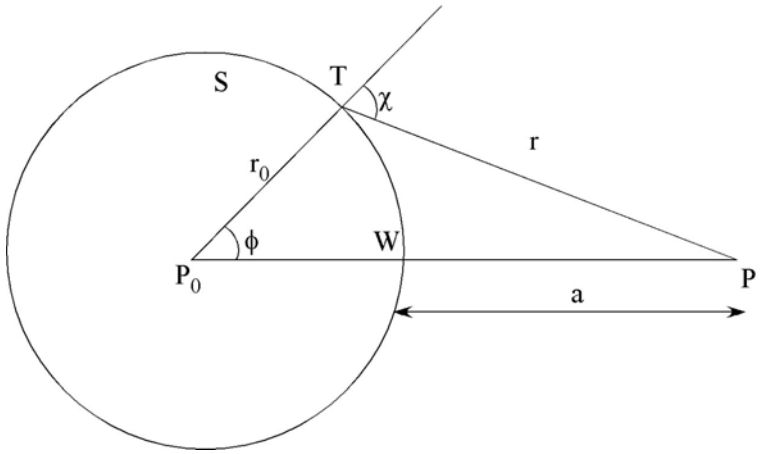


Figure 2.4: Huygens-Fresnel Principle. (Adapted from M. Born, E. Wolf, "Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light").

From Huygens-Fresnel principle, each point on the wave-front is the source for a secondary disturbance which propagates as a spherical wave. For an element of *dS* at *T* the contribution to *du(P)* is found by

$$du(P) = K(\chi)\frac{Ae^{ikr_0}}{r_0}\frac{e^{ikr}}{r}dS \tag{2.21}$$

where $r = TP$, and the obliquity factor $K(\chi)$ is introduced to describe the variation in the propagation direction of the amplitudes of the secondary waves. χ is the angle between the normal to *S* at *T* and the direction *TP*.

Fresnel made the assumption that $K(\chi)$ was a maximum in the original direction of propagation, i.e. for $\chi = 0$ and that it rapidly decreases with increasing χ, eventually being zero when *TP* is tangential to the wave-front, i.e. when $\chi = \pi/2$. The total disturbance at *P* is then found by

$$u(P) = \frac{Ae^{ikr_0}}{r_0}\iint_S K(\chi)\frac{e^{ikr}}{r}dS \tag{2.22}$$

The Huygens-Fresnel principle is a superposition integral, or alternatively it can be regarded as a convolution integral.

2.4.2 Integral Theorem of Kirchhoff and Helmholtz

The Huygens-Fresnel principle contained certain inconsistencies which were remedied by Kirchhoff. He established that the results and theory could be directly derived from the wave equation. The scalar wave equation in three-dimensions is

$$\frac{\partial^2\psi(x,t)}{\partial t^2} = c^2\nabla^2\psi(x,t) \tag{2.23}$$

where $\psi(x,t)$ is the wave-field at the point x = (x; y; z)^T at time t and c is the speed of light in vacuum.

For a monochromatic scalar wave then

$$\psi(x,t) = u(x)e^{-i\omega t} \tag{2.24}$$

Using separation of variables, then *u(x)* satisfies the time-independent wave equation

$$(\nabla^2 + k^2)u = 0 \tag{2.25}$$

Equation (2.25) is known as the Helmholtz equation.

Referring to Figure 2.5, if *V* is the volume bounded by a closed surface *S*, and *P* is a point within it, we can assume that *u(x)* possesses continuous first and second order partial derivatives within and on *S*.

If *f* and *g* are defined as scalar functions, are continuous and integrable in *V*, and bounded by *S*, then we have the general form of Greens theorem

$$\oint_S \left(g\nabla f - f\nabla g\right) \bullet d\mathbf{S} = \int_V \left(g\nabla^2 f - f\nabla^2 g\right) dV \tag{2.26}$$

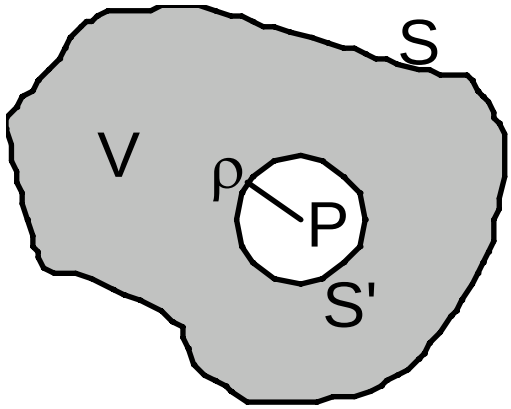


Figure 2.5: Integration surface in Green's Theorem.

If *v(x)* is another function which satisfies the same continuity requirements as *u(x)*, we have by substitution into Green's theorem in Equation (2.26)

$$\iiint_V d^3x'[u(x')\nabla^2 v(x') - v(x')\nabla^2 u(x')] = \iint_S d^2x'[u(x')\frac{\partial v(x')}{\partial n} - v(x')\frac{\partial u(x')}{\partial n}] \qquad (2.27)$$

where $\frac{\partial}{\partial n} = \hat{n}.\nabla$ is the directional derivative along the inward normal $\hat{n}$ to S.

One now chooses a Green function of the Helmholtz Equation in (2.25) so that $v(x) = G(x.x')$. This function describes the disturbance at a point x' on the surface *S* due to a spherical point source at a position *x*.

Within the volume the disturbance, $G(x.x')$ satisfies

$$(\nabla^2 + k^2)G(x, x') = \frac{\delta(|x - x'|)}{4\pi|x - x'|^2} \qquad (2.28)$$

The equation in (2.28) contains spherical symmetry, and $G(x, x') = G(r)$,

where $r = |x - x'|$. By spherical coordinates relations

$$\nabla^2 = \frac{1}{r^2}\frac{\partial^2}{\partial r^2}(r) \qquad (2.29)$$

and subsequent substitution into Equation (2.28) yields

$$\frac{1}{r}(\frac{\partial^2}{\partial r^2}(rG) + k^2(rG)) = \frac{\delta(r)}{4\pi r^2} \qquad (2.30)$$

For $r > 0$, then

$$\frac{\partial^2}{\partial r^2}(rG) + k^2(rG) = 0 \qquad (2.31)$$

This implies the condition that $rG = A^{'}e^{ikr} + B^{'}e^{-ikr}$ where $A^{'}$ and $B^{'}$ are arbitrary constants. From the Helmholtz equation

$$G = \frac{A}{4\pi r}e^{ikr} + \frac{B}{4\pi r}e^{-ikr} \tag{2.32}$$

with $A^{'} = A/4\pi$ and $B^{'} = B/4\pi$. To avoid back waves, when $B = 0$, then

$$G = \frac{A}{4\pi r}e^{ikr} \qquad \text{for } r > 0 \tag{2.33}$$

Now, extending *G* to all values of *r* then

$$G = \lim_{\varepsilon \to 0}\left(\frac{A\Theta(r-\varepsilon)}{4\pi r}e^{ikr}\right) \tag{2.34}$$

Where $\Theta(x) = 1$ if *x>0*, and $\Theta(x) = 0$ for all other x. $\Theta(x)$ is the Heaviside function. Then

$$\nabla^2 G = \lim_{\varepsilon \to 0}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\left(\frac{A\Theta(r-\varepsilon)}{4\pi r}e^{ikr}\right)\right)\right] \tag{2.35}$$

$$= \lim_{\varepsilon \to 0}\left[\frac{1}{4\pi r^2}\frac{\partial}{\partial r}\left(rAe^{ikr}\delta(r-\varepsilon) + (rAike^{ikr} - Ae^{ikr})\Theta(r-\varepsilon)\right)\right]$$

$$= \lim_{\varepsilon \to 0}\left[\frac{1}{4\pi r^2}\frac{\partial}{\partial r}\left(\varepsilon Ae^{ik\varepsilon}\delta(r-\varepsilon) + (rAike^{ikr} - Ae^{ikr})\Theta(r-\varepsilon)\right)\right]$$

$$= \lim_{\varepsilon \to 0}\left[\frac{1}{4\pi r^2}\left(\varepsilon Ae^{ik\varepsilon}\delta^{'}(r-\varepsilon) + (\varepsilon Aike^{ik\varepsilon} - Ae^{ik\varepsilon})\cdot\right.\right.$$

$$\left.\left.\delta(r-\varepsilon) + (-rk^2Ae^{ikx} + ikAe^{ikx} - ikAe^{ikx})\Theta(r-\varepsilon\right]\right]$$

$$= \frac{-Ak^2e^{ikr}}{4\pi r} - A\delta^3(x-x^{'})$$

and so

$$(\nabla^2 + k^2)G(x,x^{'}) = -A\delta^3(x-x^{'}) \tag{2.36}$$

Choosing *A=-1* to agree with Greens theorem, then

$$G(x,x') = \frac{-e^{ik|x-x'|}}{4\pi|x-x'|} \tag{2.37}$$

Substituting the Helmholtz equations for *u(x)* and $G(x,x')$, ((2.25) and (2.36)) into the volume integral in Green's theorem (2.27), we obtain

$$u(x) = \frac{1}{4\pi}\iint_S d^2x'[u(x')\frac{\partial}{\partial n}(\frac{e^{ik|x-x'|}}{|x-x'|} - \frac{e^{ik|x-x'|}}{|x-x'|}\frac{\partial u(x')}{\partial n}] \tag{2.38}$$

This is the integral theorem of Helmholtz and Kirchhoff.

2.4.3 Fresnel-Kirchhoff Diffraction

Consider a monochromatic wave, propagating from a point source, *Q*, to an opening in a plane opaque screen as depicted in Figure 2.6. The problem is to find the light disturbance at a point *R*. The disturbance can be found by taking Kirchhoff's integral over a surface S formed by the opening *A*, a portion of the non-illuminated side of the screen, *B* , and a portion of a large sphere of radius R_0, centered at *R*, *C,* which, together with *A* and *B*, form a closed surface.

Using Equation (2.38), we obtain the total disturbance at point $R, u(R)$, as

$$u(R) = \frac{1}{4\pi}[\iint_A + \iint_B + \iint_C]d^2x'[u(x')\frac{\partial}{\partial n}(\frac{e^{ik|x-x'|}}{|x-x'|}) - \frac{e^{ik|x-x'|}}{|x-x'|}\frac{\partial u(x')}{\partial n}] \tag{2.39}$$

where, $r = |x-x'|$ is the distance of the element *dS* from *R* and $\partial/\partial n = \hat{n}.\nabla$ denotes differentiation along the inward normal to the surface of integration.

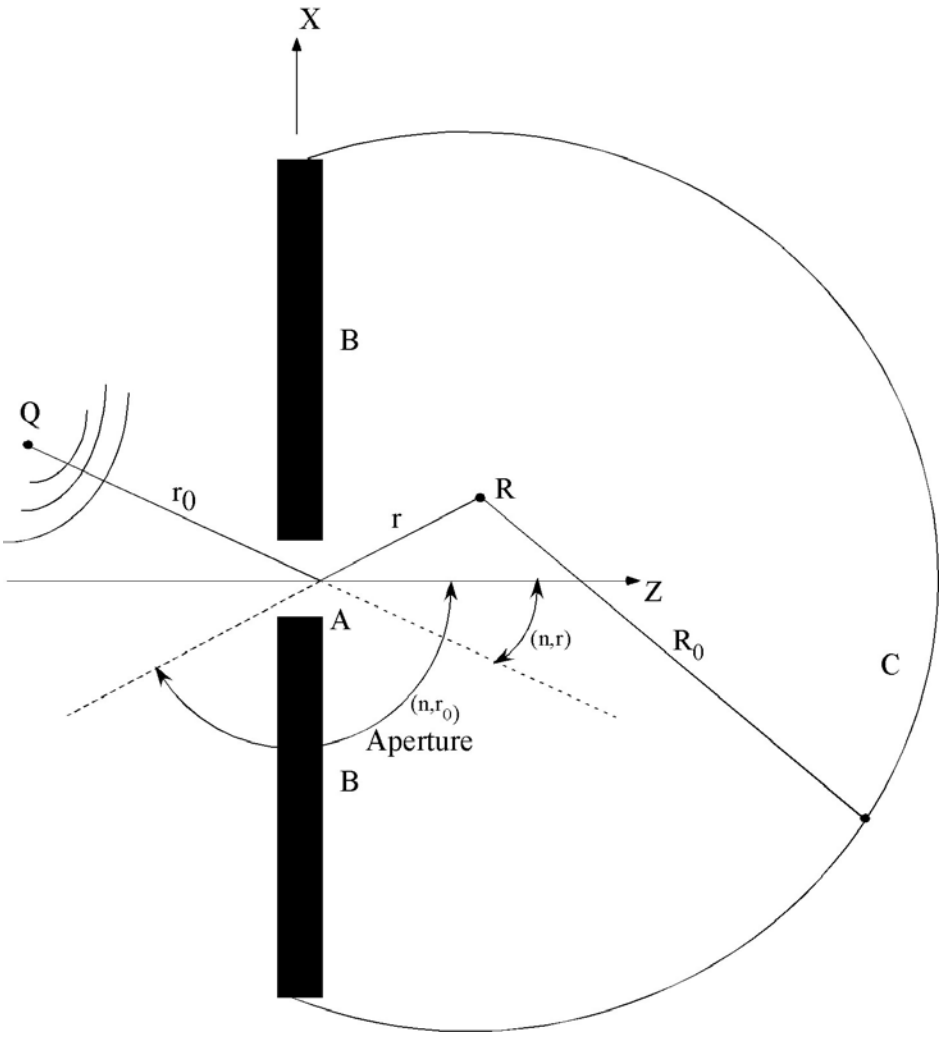


Figure 2.6: Fresnel-Kirchoff Diffraction. (Adapted from M. Born, E. Wolf, "Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light").

The values of the function of u and $\partial u / \partial n$ on *A*, *B*, and *C* are not known accurately, which presents a difficulty. On the other hand, by the assumption that at *A*, u and $\partial u / \partial n$ will not differ significantly from the values obtained in the absence of the screen, and that at *B* these quantities are approximately zero, then the boundary conditions are

$$\text{on } A: \qquad u = u^{(i)}, \qquad \frac{\partial u}{\partial n} = \frac{\partial u^{(i)}}{\partial n} \tag{2.40}$$

$$\text{on } B: \qquad u = 0, \qquad \frac{\partial u}{\partial n} = 0$$

The conditions in Equation (2.40) are known Kirchhoff's boundary conditions and these allow for substantial simplification of the results. If the aperture is large compared to the wavelength, then the boundary conditions can yield very accurate results.

For the incident field then

$$u^{(i)} = \frac{A}{4\pi r_0} e^{ikr_0}, \qquad \frac{\partial u^{(i)}}{\partial n} = \frac{A}{4\pi r_0} e^{ikr_0}[ik - \frac{1}{r_0}]\cos(n, r_0)$$

where A is a constant and $\cos(n, r_0)$ is the cosine of the angle between n and r_0 in Figure 2.6.

Finally, one considers the point *C*. The field at this point is produced by a source that emits at a time $t = t_i$. At a later time, $t > t_i$ the wave propagates outward from the source. However, the field's outer boundary is at distance not greater than $c(t - t_i)$ from *Q*. Therefore, if one chooses the radius R_0 to be large enough, so that at the time when the disturbance at *R* is considered no contributions from *C* have reached *R*, the integral over *C* will vanish.

Thus, on substituting the above conditions into Equation (2.39), we obtain

$$u(R) = -\frac{iA}{2\lambda} \iint_A d^2x' \, \frac{e^{ik(r+r_0)}}{rr_0} [\cos(n, r_0) - \cos(n, r)] \tag{2.41}$$

Equation (2.41) is the Fresnel-Kirchhoff diffraction formula.

2.4.4 Rayleigh-Somerfield Diffraction

The Fresnel-Kirchhoff diffraction theory has limitations; it has attached boundary conditions which make it mathematically inconsistent. The difference between the Rayleigh-Sommerfield theorem and the Fresnel-Kirchhoff theorem is in the inclination factor.

For Kirchhoff, the obliquity factor can be written as

$$K(\chi) = \frac{\cos(\vec{n},\vec{r}) - \cos(\vec{n},\vec{r}_0)}{2} \tag{2.42}$$

In the Rayleigh-Sommerfield approach the obliquity factor is

$$K(\chi) = \cos(\vec{n},\vec{r}) \tag{2.43}$$

A drawback of Kirchhoff's theory lies in the attached boundary conditions which are required for the field and its derivative. If a two-dimensional potential function and its normal derivative vanish together along any finite curve segment, then that potential function must vanish over the entire plane. Similarly, if a solution of the three dimensional wave equation vanishes on any finite surface element; it must vanish in all space. The two Kirchhoff boundary conditions together imply that the field is identically zero everywhere behind the aperture, a result which contradicts the known physical situation. From the integral theorem of Helmholtz and Kirchhoff, Equation (2.38), then

$$u(x) = \frac{1}{4\pi}\iint_S d^2x'[u(x')\frac{\partial G(x,x')}{\partial n} - G(x,x')\frac{\partial u(x')}{\partial n}] \tag{2.44}$$

By making adjustments to the Greens function in Equation (2.44), *G*, or its derivative, $\partial G/\partial n$, disappears at the aperture A and the non-illuminated portion B. In both cases, the necessity of imposing simultaneous boundary conditions on *u* and $\partial u/\partial n$ is

removed. Using the former case in which *G* vanished over the aperture, the integral theorem is

$$u(x) = -\frac{i}{\lambda}\iint_A d^2x'[u(x')\frac{e^{ikr}}{r}\cos(n,r)] \tag{2.45}$$

and the Rayleigh-Somerfield diffraction formula is

$$u(x) = -\frac{iA}{\lambda}\iint_A d^2x'[\frac{e^{ik(r+r_0)}}{rr_0}\cos(n,r)] \tag{2.46}$$

The derivation is more self consistent than Kirchhoff's derivation.

2.5 Speckles

Holography uses coherent light sources to form the interference pattern which creates the hologram. However, a significant issue with using a coherent source is the degradation in image quality due to a phenomenon called speckle**.**

A rough, uneven surface illuminated with a coherent light source produces a coarse indeterminate pattern as displayed in Figure 2.7. The coherent light which illuminates the surface is scattered by the surface and produces randomly oriented bright and dark patches.

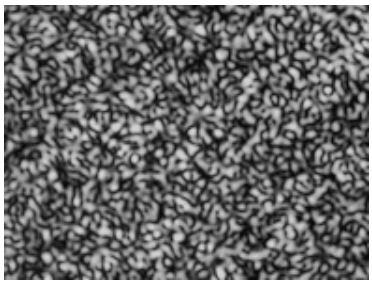


Figure 2.7: A typical speckle pattern formed by coherent illumination of a rough surface. (Obtained from http://www.polytec.com/)

The occurrence of speckle is a result of the interference between different scattered beams from the surface. This kind of speckle pattern is created if the height variations of the rough surface are larger than the wavelength of the light. The coherent noise of speckle effectively disturbs the image quality to a particularly high degree and makes identification of features in highly scattering structures difficult. This has been a major barrier for widespread coherent imaging and in particular, microscopy.

The probability density function for the intensity in a speckle pattern is given by

$$P(I)dI = \frac{1}{\langle I \rangle} e^{\left(-\frac{I}{\langle I \rangle}\right)} \tag{2.47}$$

where $P(I)dI$ is the probability that the intensity at a certain point is lying between I and dI. $\langle I \rangle$ is the mean intensity of the entire speckle field. The most probable intensity value of a speckle is therefore zero, which is why most speckles, in fact, appear black. The standard deviation is found by

$$\sigma_I = \langle I \rangle \tag{2.48}$$

That means the intensity variations are in the same order as the mean value. The contrast is given by

$$V = \frac{\sigma_I}{\langle I \rangle} \tag{2.49}$$

The contrast of the speckle pattern is unity.

2.6 Holography Recording and Reconstruction

In conventional imaging techniques, such as photography, what is recorded is simply the intensity distribution of the original object. As a result all information about

the relative optical paths to the different parts of the object is lost. The unique characteristic of holography is the recording of both intensity and phase information of the light wave reflected or transmitted by an object. As recording media respond only to the intensity, it is necessary to convert the phase information into variations in intensity. This is performed by using coherent illumination.

The general principle of conventional holography is shown in Figure 2.8, one wave illuminates the object and is scattered onto the recording medium. This is called the object beam. The second wave, the reference beam, illuminates the recording medium directly. Both waves interfere and the pattern is recorded onto the photographic plate. This recorded interference pattern is termed a 'Hologram'.

Let the complex amplitude object and reference waves at the photographic plate be denoted by U_o and U_r, respectively.

The complex amplitude of the object wave is described by

$$U_0(x,y) = a_o(x,y)\exp(\phi_o(x,y)) \tag{2.50}$$

with real amplitude a_o and phase ϕ_o

The reference wave is described by

$$U_r(x,y) = a_r(x,y)\exp(\phi_r(x,y)) \tag{2.51}$$

with real amplitude a_r and phase ϕ_r.

Both waves interfere at the recording medium and the intensity may be written as

$$I(x,y) = \left|U_o(x,y)+U_r(x,y)\right|^2 = (U_o(x,y)+U_r(x,y))(U_o(x,y)+U_r(x,y))^* \tag{2.52}$$

$$= U_o(x,y)U_o{}^*(x,y)+U_r(x,y)U_r{}^*(x,y)+U_o(x,y)U_r{}^*(x,y)+U_o{}^*(x,y)U_r(x,y)$$

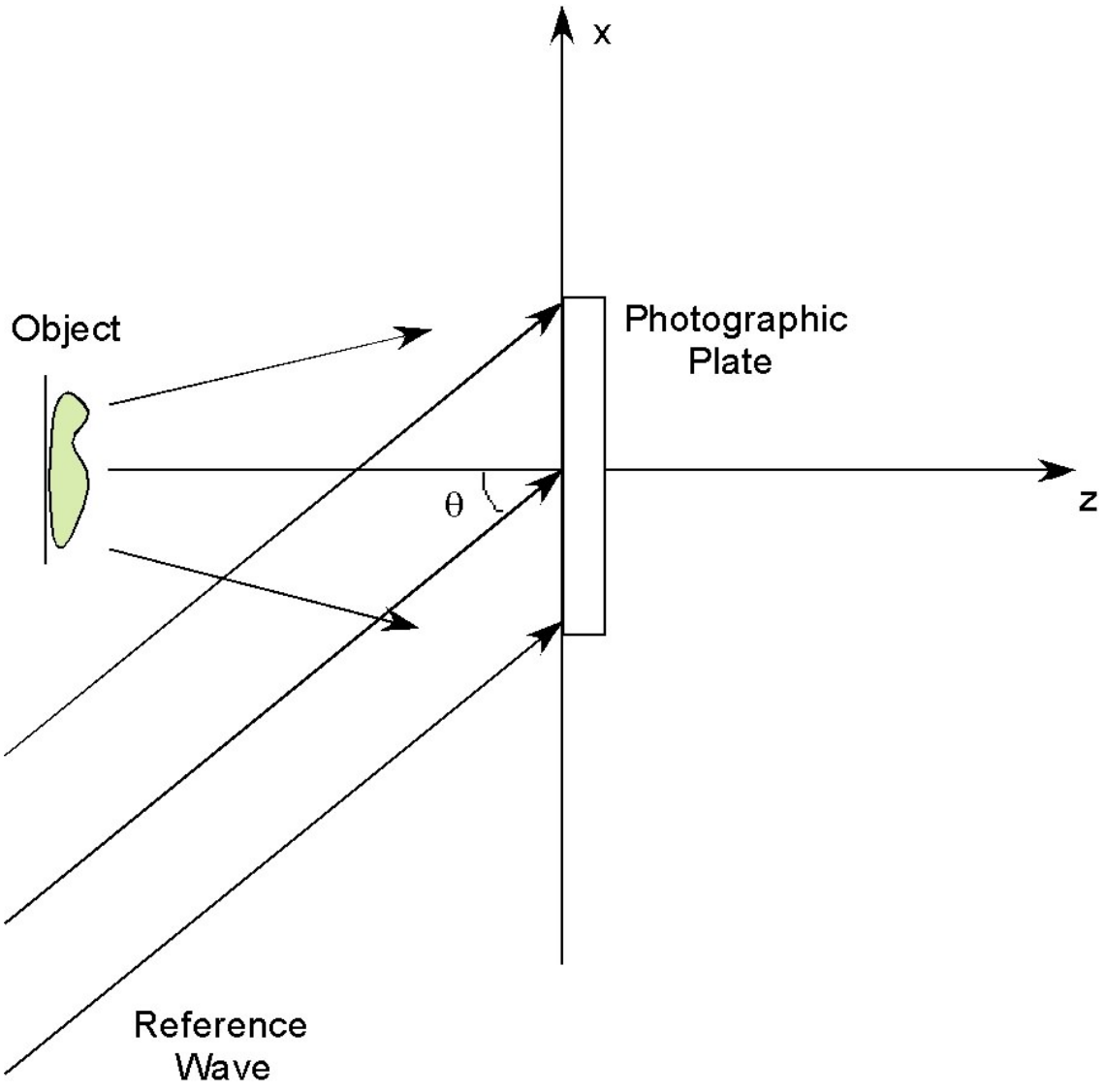


Figure 2.8: Recording a hologram

The amplitude transmission of the developed photographic plate is proportional to the intensity and can be represented by

$$t = \alpha_0 + \alpha I(x,y) = \tag{2.53}$$

$$= \alpha_0 + \alpha[U_o(x,y)U_o^*(x,y) + U_r(x,y)U_r^*(x,y) + U_o(x,y)U_r^*(x,y) + U_o^*(x,y)U_r(x,y)]$$

where α_0 and α are constants determined by the recording medium type.

The original object wave is reconstructed by illuminating the recorded hologram at the photographic plate with the reference beam as shown in Figure 2.9. The observer sees the virtual image, which is indistinguishable from the image of the original object.

When the plate is illuminated by the reference beam for reconstruction

$$U_r\, t = U_r\alpha_0 + \alpha[U_o(x,y)U_o^*(x,y)U_r + U_r(x,y)U_r^*(x,y)U_r + U_o(x,y)U_r^*(x,y)U_r(x,y) + U_o^*(x,y)U_r(x,y)U_r(x,y)]$$

$$= U_r\alpha_0 + \alpha[|U_0|^2 U_r + |U_r|^2 U_r + U_0 |U_r|^2 + U_0^* U_r^2] \tag{2.54}$$

The first term on the right hand side of Equation (2.54) is usually called the ambiguity term, and is the product between the object wave's intensity and the reference wave. Since $|U_o|^2$ is generally non-uniform, it produces a wave that propagates along the direction of the illuminating plane wave within a cone, the angular aperture of which depends on the spatial spectrum of the object wave. The second term on the right-hand side is the product of the intensity of the reference wave the illumination wave. Since the reference wave is a plane wave, $|U_r|^2$ is uniform, this term represents a plane wave propagating along the direction of the illuminating plane wave. Together, these first two terms on the right-hand side of Equation (2.54) form the zero order of diffraction also known as the *dc* term.

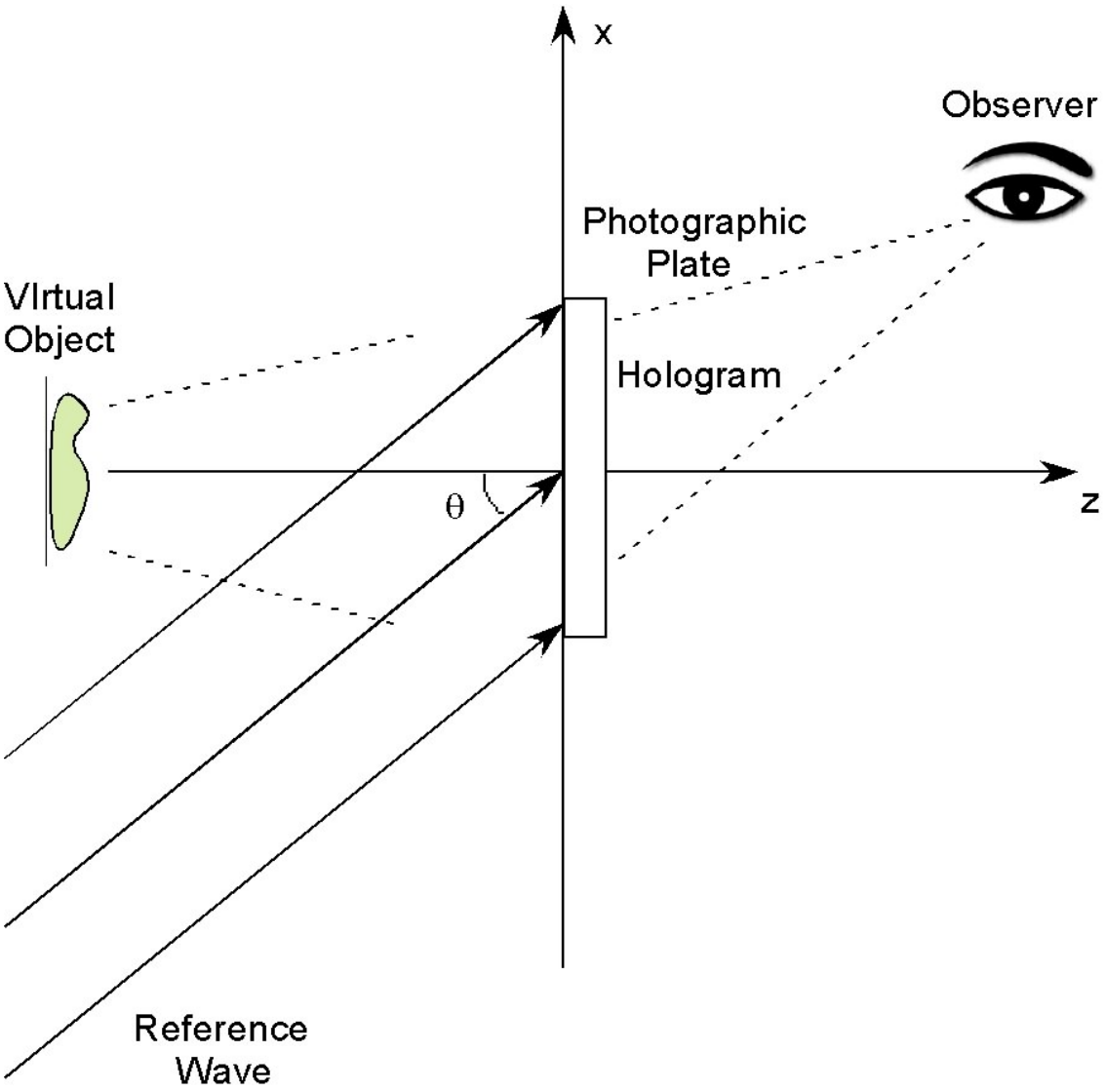


Figure 2.9: Reconstruction of a hologram

The third and the fourth terms in Equation (2.54) are produced by the holographic interference terms, and each of them generates an image of the object. The third term is the reconstructed object wave forming the virtual image located at the position initially occupied by the object. The fourth term produces a real image located on the other side of the hologram. Since this term is formed by the conjugate of the object wave U_o^* the depth is inverted. Points on the virtual image and real image are located at equal distances from the hologram plane but in opposite directions. The real image, therefore, appears inside out and is called a pseudoscopic image.

Chapter 3

Digital Holography

This chapter discusses principles of digital holography. Section 3.1 examines the physical and mathematical aspects of recording holograms onto a CCD camera. Section 3.2 analyses numerical diffraction methods for reconstructing the digital hologram in the diffractive planes of the holographic terms. Section 3.3 describes methods of separating the *dc* and holographic terms. In Section 3.4, the use of CCD cameras for holographic recording is discussed. Finally, Section 3.5 examines resolution and calibration issues in the holographic optical system.

3.1 Digital Recording of Holograms

The method of recording digital holograms is illustrated in Figure 3.1. A coherent object wave reflected or transmitted from an object and a reference wave interfere at the surface of a CCD camera. The holographic interference pattern formed at the CCD is electronically recorded and stored.

In digital holography this recorded hologram is subsequently reconstructed computationally by using numerical algorithms to provide direct access to amplitude and phase information of the object.

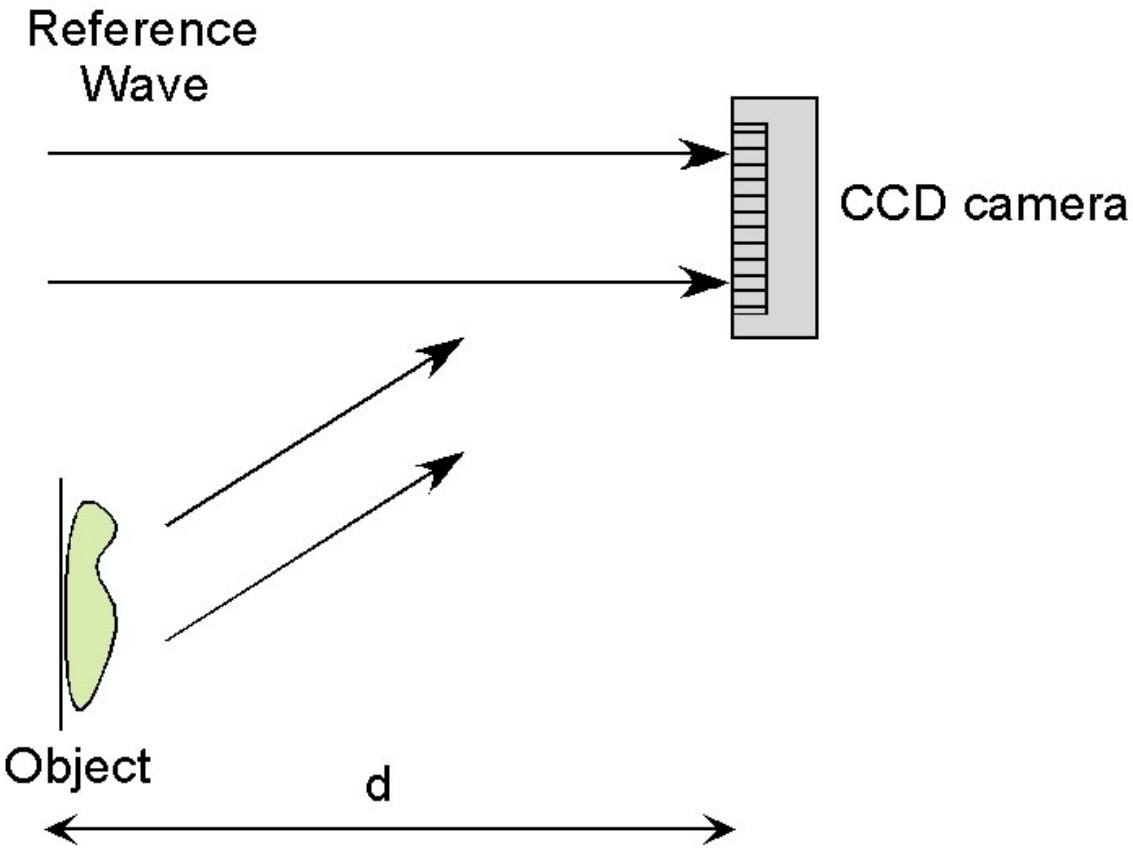


Figure 3.1: Recording of a digital hologram.

In the numerical reconstruction of the hologram, the holographic terms, the virtual image and real image are formed at distances, *d and d'* respectively from the hologram plane which is assumed to be in the $d' = d = 0$ plane as displayed in Figure 3.2, which illustrates the coordinate system for hologram reconstruction. In classical holography, the hologram is reconstructed by illumination with an optical replica of the wave-field. In digital holography reconstruction takes place by multiplication with the numerical model of the reference wavefield. The interference between the object wave $o(\xi,\eta)$ transmitted by the object located at the object plane and the reference wave $r(\xi,\eta)$ is recorded in the hologram plane (ξ,η). The complex object wave at the hologram plane is

$$o(\xi,\eta) = o_0 \exp\{\frac{ik}{2d}(\xi - x)^2 + (\eta - y)^2\} \tag{3.1}$$

where o_0 is the amplitude of the object wave. The complex amplitude of the reference wave diverging from a point (x_r, y_r, d_r) is

$$r(\xi,\eta) = r_0 \exp\{\frac{ik}{2d_r}(\xi - x_r)^2 + (\eta - y_r)^2\} \tag{3.2}$$

where r_0 is the amplitude of the reference wave. The complex amplitude of the interference pattern at the hologram plane is

$$U(\xi,\eta) = o(\xi,\eta) + r(\xi,\eta) \tag{3.3}$$

A CCD sensor records this interference pattern in the form of the intensity distribution

$$h(\xi,\eta) = U(\xi,\eta).U^*(\xi,\eta) \tag{3.4}$$

which is stored in the computer as the digital hologram.

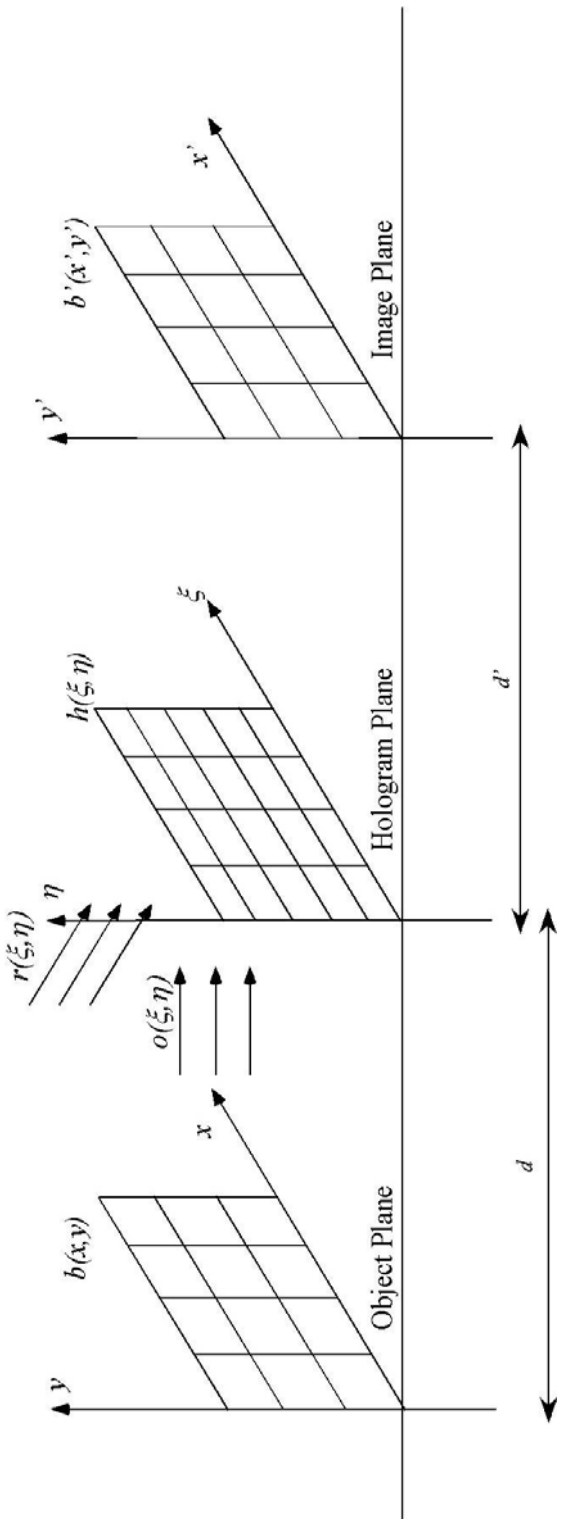


Figure 3.2: Coordinate system for digital holography

The interference pattern is then mathematically described by

$$h(\xi,\eta) = \left|o(\xi,\eta)\right|^2 + \left|r(\xi,\eta)\right|^2 + o^*(\xi,\eta)r(\xi,\eta) + o(\xi,\eta)r^*(\xi,\eta) \qquad (3.5)$$

3.2 Numerical diffraction

Once the hologram has been captured, the optically diffracted field is numerically propagated by the use of reconstruction algorithms. The diffraction pattern is calculated at a distance *d'* behind the CCD focal plane (hologram plane), which means, it reconstructs the complex amplitude in the plane of the real image. Note however, that one can focus on either the real or virtual image located a respective distances *d'* and *d* from the hologram plane. The magnitude of the distances is the same but the signs are opposite. In this thesis we define the general expression for the reconstruction distance for either of the holographic images as *z*.

Mathematically, the reconstruction process reduces to the calculation of the Rayleigh-Sommerfeld diffraction integral. If one treats the digital holographic system as a coherent imaging system, then the reconstruction algorithm functions like the lens in an optical system. All computer calculations can be performed based on the Rayleigh-Sommerfeld diffraction formula, however, using this formula directly is often time consuming. There are a number of computationally efficient numerical methods available to calculate the holographic diffraction[40]. The most commonly applied method has been the discrete Fresnel transform where the approximation of spherical Huygens wavelet by a parabolic surface allows the calculation of diffraction integral using a single Fourier transform[6].

In the convolution method, the diffraction integral is calculated without such approximation using three Fourier transforms through the convolution theorem [41]. On the other hand, the angular spectrum method involves two Fourier transforms, plus simple filtering of the angular spectrum[32]. Here we review these methods.

3.2.1 Reconstruction by Fresnel Transform method.

If the wave-field is known in one plane, it can be calculated in any other plane in the 3D space. The Rayleigh–Sommerfeld diffraction theory is a precise mathematical form of this statement. Referring to Figure 3.2, the Rayleigh-Sommefield diffraction formula in this coordinate system for the reconstruction of the holographic image in the image plane is

$$b'(x',y') = \frac{1}{i\lambda}\iint h(\xi,\eta) r(\xi,\eta) \frac{\exp(ik\rho)}{\rho} \cos\Theta d\xi d\eta \tag{3.6}$$

where $b'(x',y')$ is the reconstructed image in the image plane, $h(\xi,\eta)$ is the hologram recorded by the CCD sensor array, $r(\xi,\eta)$ is the reference wave-field and ρ is the distance between a point in the hologram plane and a point in the reconstruction plane defined as

$$\rho = \sqrt{d'^2 + \left(\xi - x'\right)^2 + \left(\eta - y'\right)^2} \tag{3.7}$$

d' is the reconstruction distance i.e. the distance between the hologram and image plane. Due to the small angles between the hologram normal and the rays from the hologram to the image points, the obliquity factor can be set to $\cos\Theta = 1$ in Equation (3.6).

The expression in Equation (3.7) can be expanded to a Taylor series so that

$$\rho = d' + \frac{(\xi - x')^2}{2d'} + \frac{(\eta - y')^2}{2d'} - \frac{1}{8}\frac{\left[\left(\xi - x'\right)^2 + \left(\eta - y'\right)^2\right]^2}{d'^3} + \ldots \tag{3.8}$$

The fourth term can be neglected if it is small compared to the wavelength and so

$$\rho \approx d' + \frac{(\xi - x')^2}{2d'} + \frac{(\eta - y')^2}{2d'} \tag{3.9}$$

With the additional approximation of replacing the factor ρ by the reconstruction distance d' in the denominator of Equation (3.6) then

$$b'(x', y') = \frac{\exp(ikd')}{i\lambda d'} \iint h(\xi,\eta) r(\xi,\eta) \exp[\frac{i\pi}{\lambda d'}\left[\left(\xi - x\right)^2 + \left(\eta - y\right)^2\right]] d\xi d\eta \tag{3.10}$$

Further multiplication in the argument of the exponential yields

$$b'(x', y') = \frac{\exp(ikd')}{i\lambda d'} \iint h(\xi,\eta) r(\xi,\eta) \exp[\frac{i\pi}{\lambda d'}\left[\left(\xi\right)^2 + \left(\eta\right)^2\right]] \cdot$$

$$\exp[\frac{i\pi}{\lambda d'}\left[\left(x'\right)^2 + \left(y'\right)^2\right]] \exp[-\frac{2i\pi}{\lambda d'}\left[\left(x'\xi\right) + \left(y'\eta\right)\right]] d\xi d\eta \tag{3.11}$$

Defining the impulse response function of the optical system as

$$g(\xi,\eta) = \frac{\exp(ikd')}{i\lambda d'} \exp[\frac{i\pi}{\lambda d'}(\xi^2 + \eta^2)] \tag{3.12}$$

the reconstructed wavefield is then

$$b'(x', y') = \iint h(\xi,\eta) r(\xi,\eta) g(\xi,\eta) \exp[\frac{i\pi}{\lambda d'}\left[\left(x'\right)^2 + \left(y'\right)^2\right]] \exp[-\frac{2i\pi}{\lambda d'}\left[\left(x'\xi\right) + \left(y'\eta\right)\right]] d\xi d\eta \tag{3.13}$$

This equation is known as the Fresnel approximation. This approximation states that the size of the aperture is small in comparison to the distance to the viewing plane.

Fresnel diffraction is essentially the Fourier transform of the aperture multiplied by the phase factor of the light at the aperture. When the reconstruction distance z (*d* or *d'*) is sufficiently large for this approximation to be an accurate one, the observer is said to be in the region of Fresnel diffraction. For the Fresnel approximation to remain valid, it is only necessary that the higher order terms of the expansion do not change the value of the integral in Equation (3.13).

The intensity is calculated by

$$I(x',y') = \left|b'(x',y')\right|^2 \tag{3.14}$$

and the phase by

$$\phi(x',y') = \arctan\frac{\mathrm{Im}\left|b'(x',y')\right|}{\mathrm{Re}\left|b'(x',y')\right|} \tag{3.15}$$

For digitization of the Fresnel transform then the following spatial frequency terms are introduced

$$\nu = \frac{x'}{\lambda d'} \text{ and } \mu = \frac{y'}{\lambda d'} \tag{3.16}$$

Equation (3.13) then becomes

$$b'(\nu,\mu) = \iint h(\xi,\eta)r(\xi,\eta)g(\xi,\eta)\exp[i\pi\lambda d'\left[\nu^2+\mu^2\right]]\exp[-2i\pi\left[(\nu\xi)+(\mu\eta)\right]]d\xi d\eta \tag{3.17}$$

Comparison of Equation (3.17) with the definition of the two dimensional Fourier transform

$$\Im\{(f(\xi,\eta)\} = \iint f(\xi,\eta)\exp[-2i\pi(\nu\xi+\mu\eta)]d\xi d\eta \tag{3.18}$$

with $f(\xi,\eta) = h(\xi,\eta)r(\xi,\eta)g(\xi,\eta)$,

then Equation (3.17) becomes

$$b'(\nu,\mu) = \exp[i\pi\lambda d'\left[\nu^2+\mu^2\right]]\Im\{f(\xi,\eta)\} \tag{3.19}$$

The function $b'(\nu,\mu)$ can be digitized if the hologram, $h(\xi,\eta)$ is sampled on a $M \times N$ pixel array on the CCD sensor, with steps $\Delta\xi$ and $\Delta\eta$ along the coordinates. Therefore $\Delta\xi$ and $\Delta\eta$ represent the distances between neighboring pixels in the horizontal and vertical directions and so

$$b'(m,n) = \frac{\exp(ikd')}{i\lambda d'}\exp[i\pi\lambda d'\left[m^2\Delta\nu^2+n^2\Delta\mu^2\right]]\Im[h(k,l)r(k,l)\exp[\frac{i\pi}{\lambda d'}(k^2\Delta\xi^2+l^2\Delta\eta^2)] \tag{3.20}$$

where *k, l, m, n* are integers defined in the range from $(-M/2 \le k,l \le M/2)$ and $(-N/2 \le m,n \le N/2)$.

The pixel sizes of the CCD array $(\Delta\xi,\Delta\eta)$ located at the hologram plane are different to those in the image plane $(\Delta x', \Delta y')$. The discrete Fourier transform theorem tells us the maximum frequency is determined by the sampling interval in the spatial domain and so

$$\Delta\nu = \frac{1}{M\Delta\xi} \text{ and } \Delta\mu = \frac{1}{N\Delta\eta} \tag{3.21}$$

substitution of these terms provides the relations between the hologram and image plane

$$\Delta\xi = \frac{d'\lambda}{M\Delta x'} \text{ and } \Delta\eta = \frac{d'\lambda}{N\Delta y'} \tag{3.22}$$

this gives the reconstructed wave-field as

$$b'(m,n) = \frac{\exp(ikd')}{i\lambda d'}\exp[\frac{i\pi}{\lambda d'}\left[m^2\Delta x'^2+n^2\Delta y'^2\right]]\Im[h(k,l)r(k,l)\exp[\frac{i\pi}{\lambda d'}(k^2\Delta\xi^2+l^2\Delta\eta^2)]_{m,n} \tag{3.23}$$

This is the discrete Fresnel transform. The calculation in Equation (3.23) can be made more efficient by applying the Fast Fourier Transform (FFT) so that Equation (3.23) becomes

$$b'(m,n) = \frac{\exp(ikd')}{i\lambda d'} \exp[\frac{i\pi}{\lambda d'}\left[m^2\Delta x'^2 + n^2\Delta y'^2\right]]FFT[h(k,l)r(k,l)\exp[\frac{i\pi}{\lambda d'}(k^2\Delta\xi^2 + l^2\Delta\eta^2)]_{m,n} \quad (3.24)$$

The pixel resolution $\Delta x'$ and $\Delta y'$ of the reconstructed images which is determined directly from the Fresnel transform will vary as a function of the reconstruction distance *d'* as

$$\Delta x' = \frac{\lambda d'}{M\Delta\xi} \; and \; \Delta y' = \frac{\lambda d'}{N\Delta\eta} \quad (3.25)$$

Therefore the reconstructed image changes in size according to the reconstruction distance. This is an inconvenient limitation of this method. Equation (3.25) is also the horizontal and vertical diameters of the Airy disk in the image plane and sets the diffraction limited resolution of the optical system.

Although the Fresnel diffraction formula can still give an accurate reconstruction for smooth and slowly varying objects where the Fresnel approximation is not strictly satisfied, it cannot correctly reconstruct near wave-fields for more diffractive objects, where the higher-order terms in the Taylor expansion of the Fresnel approximation in Equation (3.8) are more significant.

The Fresnel transform also requires that the distance between the object and the hologram plane be sufficiently large in comparison to the size of the object or the hologram. This sets the minimum reconstruction distance *z* requirement so that

$$z_{\min} = \frac{a_x{}^2}{N\lambda} \quad (3.26)$$

where a_x is the image size and N is the number of pixels. At too close a distance, the spatial frequency of the pixelated hologram is not high enough to reproduce a large angular size of the object without aliasing.

3.2.2 Reconstruction by Fresnel Convolution method

The Rayleigh-Sommerfield diffraction formula in Equation (3.4) can be rewritten as a superposition integral so that

$$b'(x', y') = \iint g(\xi,\eta)h(\xi,\eta)r(\xi,\eta)d\xi d\eta \tag{3.27}$$

where the Fresnel impulse response function is

$$g(\xi,\eta) = \frac{\exp(ikd')}{i\lambda d'}\exp[\frac{i\pi}{\lambda d'}(\xi^2+\eta^2)] \tag{3.28}$$

The linear system is space invariant so that for the impulse response function

$$g(x', y', \xi, \eta) = g(x'-\xi, y'-\eta) \tag{3.29}$$

The superposition integral in (3.27) can be expressed as a convolution

$$b(x', y') = \Im^{-1}\{\Im[h(\xi,\eta)r(\xi,\eta)]\Im[g((\xi,\eta)]\} \tag{3.30}$$

$\Im[g((\xi,\eta)]$ is the Fourier transform of the impulse response function. From Equation 3.16 the definition for the Fourier transform of the impulse response function is

$$\begin{aligned}\Im[g(\xi,\eta)] &= \iint g(\xi,\eta)\exp[-i2\pi(\nu\xi+\mu\eta)]d\xi d\eta \\ &= \exp(ikd')\exp[-i\pi\lambda d'(\nu^2+\mu^2)]\end{aligned} \tag{3.31}$$

The reconstructed wave-field is then

$$b(x',y') = \Im^{-1}\{\Im[h(\xi,\eta)r(\xi,\eta)]\exp(ikd')\exp[-i\pi\lambda d'(\nu^2+\mu^2)]\} \qquad (3.32)$$

The convolution approach gives an image with constant pixel dimensions between the hologram and image plane.

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

However, due to the more complex algorithm, the convolution approach is slower than the Fresnel transform approach. The whole process requires three Fourier transforms, which are carried out using the FFT algorithm for more effective calculation.

The sampling requirements in the Fresnel and convolution algorithm are the same and therefore this method is also subject to the minimum reconstruction distance set by Equation (3.26). At too close a distance, the spatial frequency of the hologram is too low and aliasing occurs. Normally the object is placed just outside this minimum distance found from Equation (3.26).

3.2.3 Reconstruction by Huygens Convolution method

The Fresnel approximation in convolution is not wholly justified for all *z*-values. The Huygens convolution method often yields better quality images, especially when the hologram and image plane are close [31]. The Rayleigh-Sommerfield diffraction formula in Equation 3.4 is again rewritten as a superposition integral`

$$b'(x',y') = \iint \frac{1}{i\lambda}\frac{\exp(ik\rho)}{\rho}h(\xi,\eta)r(\xi,\eta)d\xi d\eta \qquad (3.34)$$

where the impulse response function is

$$g(\xi,\eta,x',y') = \frac{1}{i\lambda}\frac{\exp(ik\rho)}{\rho} \qquad (3.35)$$

The Huygens impulse response function is approximated by

$$g(x^{'},y^{'},\xi,\eta)=\frac{1}{i\lambda}\frac{\exp\left(ik\sqrt{d^{'2}+\left(\xi-x^{'}\right)^{2}+\left(\eta-y^{'}\right)^{2}}\right)}{\sqrt{d^{'2}+\left(\xi-x^{'}\right)^{2}+\left(\eta-y^{'}\right)^{2}}} \tag{3.36}$$

$$=\frac{1}{i\lambda d^{'}}\exp\left(ik\sqrt{d^{'2}+x^{'2}+y^{'2}}\right) \tag{3.37}$$

Finally

$$b(x^{'},y^{'})=\Im^{-1}\{\Im[h(\xi,\eta)r(\xi,\eta)]\Im[g((\xi,\eta)]\} \tag{3.38}$$

The final inverse Fourier transform brings the convolution result back to the spatial domain, whereby the reconstructed pixel size becomes independent of *z* and one can conveniently focus the reconstruction at different reconstruction distances without re-scaling issues as in the Fresnel transform.

The whole process requires three Fourier transforms, which are carried out using the FFT algorithm for a more efficient calculation. The pixel sizes of the images reconstructed by the convolution approach are equal to that of the hologram as in the Fresnel convolution. The minimum distance for reconstruction of the image by this method is given by Equation (3.26).

3.2.4 Reconstruction by Angular Spectrum method

Suppose that, due to some unspecified system of monochromatic sources, a plane wave is incident on the x_0y_0 plane of Figure 3.3, traveling in the positive *z* direction. Let the complex field across that plane be represented by $E_0(x; y; 0)$. The objective is to

calculate the consequent field E(x; y; z) that appears at a second point with coordinates (x; y; z). By use of the generalized form of the two-dimensional Fourier transform

$$\Im\{f(x,y) = F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\exp\{-i2\pi(ux+vy)dxdy \tag{3.39}$$

Across the x_0y_0 plane, the function E_0 has a two-dimensional Fourier transform which is then given by

$$A(f_x, f_y, 0) = \int\int_{-\infty}^{\infty} E_0(x_0, y_0, 0)\exp[-i2\pi(f_x x + f_y y)]dx_0 dy_0 \tag{3.40}$$

where $A(f_x, f_y; 0)$ is the angular spectrum. With the substitution of

$$f_x = k_x / 2\pi \quad \text{and} \quad f_y = k_y / 2\pi \tag{3.41}$$

then we obtain

$$A(k_x, k_y; 0) = \int \int_{-\infty}^{\infty} E_0(x_0, y_0, 0)\exp[-i(k_x x_0 + k_y y_0)]dx_0 dy_0 \tag{3.42}$$

where k_x and k_y are corresponding spatial frequencies of x and y.

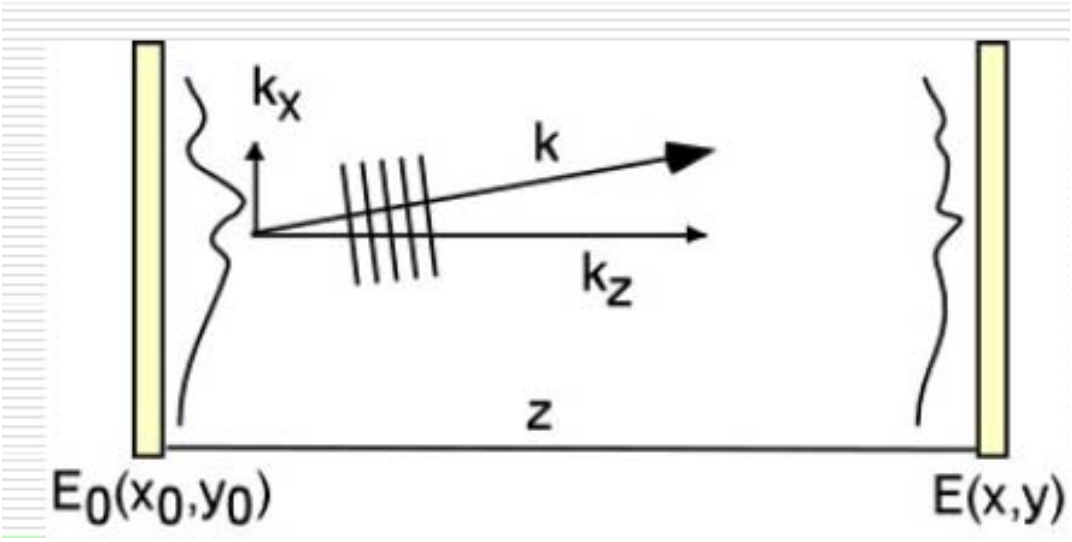


Figure 3.3: Coordinate system for reconstruction of the hologram in the angular spectrum method.

Fourier-domain filtering can be applied to the spectrum to block unwanted spectral terms in the hologram and select a region of interest corresponding only to the object spectrum. A modified wave-field $E_0(x_0, y_0; 0)$ can be written as the inverse Fourier transform of the filtered angular spectrum $A(k_x, k_y; 0)$. The operation of a Fourier transform may be regarded as a decomposition of a complicated function into a collection of more simple complex-exponential functions. So we can write E_0 as an inverse transform of its spectrum

$$E_0(x_0, y_0, 0) = \int\int_{-\infty}^{\infty} A(k_x, k_y; 0) e[i(k_x x_0 + k_y y_0)] dx_0 dy_0 \tag{3.43}$$

Propagating plane waves are of the form, (writing the dot product of vectors k and r out fully, r being position vector)

$$b(x, y, z) = \exp[i(k_x x + k_y y + k_z z) \tag{3.44}$$

Where

$$k^2 = k_x{}^2 + k_y{}^2 + k_z{}^2 \tag{3.45}$$

The new angular spectrum at plane z, $A(k_x, k_y; z)$ is calculated from $A(k_x, k_y; 0)$ as

$$A(k_x, k_y; z) = A(k_x, k_y; 0) \exp\left[ik_z z\right] \tag{3.46}$$

The spatial frequency diffraction kernel is the *z*-portion from Equation (3.44) and substitution of Equation (3.45) gives

$$\exp[ik_z z] = \exp[i\sqrt{k^2 - k_x{}^2 - k_y{}^2}\, z] \tag{3.47}$$

This can be written by substitution of terms in Equation (3.41) as

$$\exp[ik_z z] = \exp[i\sqrt{\left(\frac{2\pi}{\lambda}\right)^2 - \left(2\pi f_x\right)^2 - \left(2\pi f_y\right)^2}\, z] \tag{3.48}$$

Removing the 2π term from square root bracket in Equation (3.48) then the angular spectrum diffraction kernel or optical transfer function is

$$H(f_x, f_y) = \exp[i2\pi\sqrt{\left(\frac{1}{\lambda}\right)^2 - (f_x)^2 - (f_y)^2}] \tag{3.49}$$

The reconstructed complex wave-field of any plane perpendicular to the propagating *z* axis is found by

$$E(x, y; z) = \iint A(k_x, k_y; z)\exp[i(k_x x + k_y y)]dk_x dk_y \tag{3.50}$$

Therefore by substitution of Equations (3.48) and (3.49) then

$$E(x, y; z) = \iint A(k_x, k_y; 0)\exp[i2\pi\sqrt{\left(\frac{1}{\lambda}\right)^2 - (f_x)^2 - (f_y)^2}]\exp[i(k_x x + k_y y)]dk_x dk_y \tag{3.51}$$

$$E(x, y; z) = \Im^{-1}\{filter[\Im\{U(x_0, y_0, 0)\}]\exp[i2\pi\sqrt{\left(\frac{1}{\lambda}\right)^2 - (f_x)^2 - (f_y)^2}]\} \tag{3.52}$$

Here 'filter' represents filtering in the spectral domain.

The effect of propagation over a distance *z* is simply a change in the relative phases of the various components of the angular spectrum. Since each plane wave component propagates at a different angle, each travels a different distance to reach a given observation point and relative phase delays are thus introduced.

In the calculation in Equation (3.52) two Fourier transforms are needed for the calculation in comparison to the one needed by the Fresnel transform. However once the field is known at any one plane, only one additional Fourier transform is needed to calculate the field at different values of *z*. This method allows frequency-domain spectrum filtering to be applied, which for example can be used to block or remove the

disturbance of the *dc* term and twin image components. The angular spectrum method of calculating the holographic optical field is seen to have several advantages over the more commonly used Fresnel transformation or Huygens convolution method. Spurious noise and interference components can be tightly controlled through the analysis and filtering of the angular spectrum. The reconstruction distance does not have a lower limit and the off-axis angle between the object and reference can be lower that the Fresnel requirement and still be able to cleanly separate out the zero-order background.

3.3 Separation of the Virtual Image, Real Image and *dc* term

In Gabor (in-line) holography, the virtual image, real image and *dc* term are superimposed thus creating difficulties for the acquisition of the object wave information. There have been numerous ideas and suggestions proposed as solutions to this problem, including some by Gabor himself, however most of these require extra technical effort and are not efficient procedures [42].

Off-axis holography, as first devised by Leith and Upatnieks provides both an effective and simple means of laterally separating the *dc* and holographic terms by the introduction of a sufficiently large angle between the object and reference beams. The increase in the angle between the beams increases the spatial distance between the *dc* and holographic terms which is desired so that one can easily extract the relevant frequency information for the holographic image. However there is a limitation associated with this increased spatial distance due to the low spatial resolution of CCD cameras which means that often these terms may still partially overlap. In this section we explore the ‘anatomy’ of a digital hologram.

3.3.1 Suppression of the DC term

Figure 3.4 displays a digital hologram, Fig.3.4 (a), recorded in an off-axis configuration and its associated Fourier transform Fig.3.4 (b). The bright area at the centre of the image is the un-diffracted reconstruction wave known as the *dc* term or *zero order*. This term holds most of the energy in the image. The holographic terms, the virtual image and real image are located slightly shifted to the top left and bottom right of the *dc* term. The spatial frequencies of the interference terms are located symmetrically with respect to the center of the image. This shift is caused by the angle introduced between the object and reference beams. Higher order holographic terms are also present to the extreme top left and bottom right of the figure. Also one can see spurious noise components whose source is from the CCD camera. The energy is concentrated around three main frequencies: (0,0) for the *dc* term, $(-kx, -ky)$ for the image and (kx, ky) for the conjugate image.

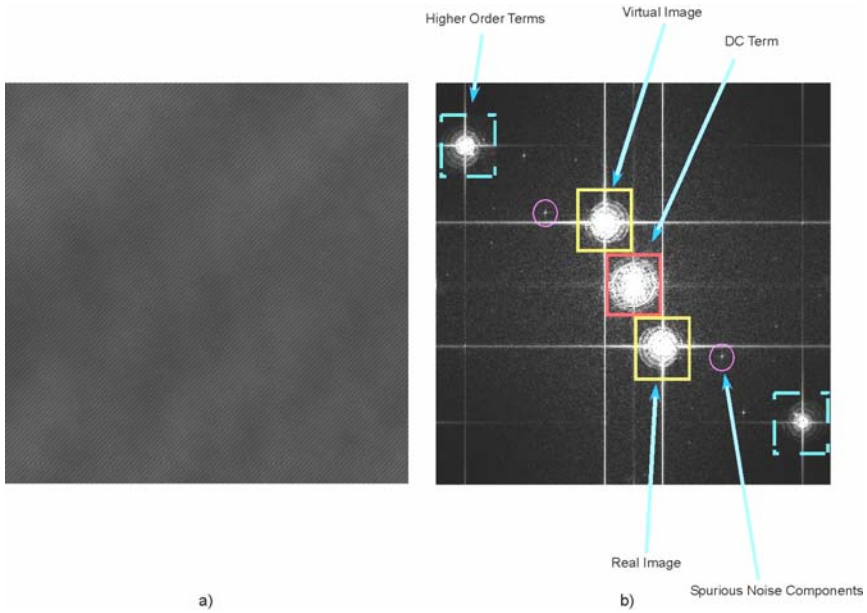


Figure 3.4: Absolute value of the Fourier transform of the hologram. The frequency origin is in the center.

As the reconstruction distance increases, the *dc* term stays located in the central part of the image, while the holographic terms move away from the center. Either of the holographic terms contains the entire information of the object wave and so in the numerical reconstruction one focuses either on the virtual or real image located at the object and image planes respectively. The objective is then to remove the *dc* term, conjugate image and any residual noise in the hologram or alternatively one can simply 'cut' out the relevant frequency information of the holographic image and discard the rest via a band-pass filtering procedure. This procedure can be implemented in either the spatial or frequency domain. A combination of windows of various shapes and sizes can be used to filter the undesired spatial frequencies while keeping intact the contributions of the interference terms. Inserting, a mask into the image spectrum in Fourier space allows the holographic term to be isolated and cut out. On the other hand, one notices in Fig. 3.4b) that if either holographic term were to be cut out then there would still be weak frequency components of the *dc* term that still would disturb either of the selected holographic images and as a consequence would degrade the quality of the reconstructed image. These components cannot be removed by filtering in either domain, unlike the easily recognizable spurious noise components.

A simple method for suppression of the *dc* term is to subtract the average intensity from the hologram before reconstruction. A similar effect can be achieved by the subtraction of two holograms. This provides a particular effective way of *dc* suppression by experiment. However a disadvantage of either of these methods is the increased experimental effort.

3.4 Recording of Digital Holograms

3.4.1 CCD Cameras for Digital Holography

Digital holography is very different to conventional holography in terms of the recording medium and the way in which the hologram is stored. In conventional holography the hologram is recorded on a photographic plate, typically made of silver halide which is then chemically processed. The original object wave is subsequently obtained by illuminating the recorded hologram with original reference wave.

In digital holography, the hologram is directly recorded onto the CCD and digitally stored in a computer where the object wave is reconstructed numerically on the computer. The exposure time of a CCD camera recording a hologram is significantly shorter than that of the photographic plate, so it relaxes the stability requirement on the recording system, and as a result, moving objects can be recorded with a short pulse of light.

The elimination of the need for chemical processes, quantitative information, easy data storage, and the fast recording of holograms are just some of the important advantages of using a CCD camera in digital holography. CCD cameras are an essential component of exploring the many advantages of digital holography and should be carefully selected dependant upon the specific application. There are a number of CCD cameras which can be used effectively in digital holography with differing architectures.

In a full frame CCD, the entire detector surface is first exposed to light and then the light is externally blocked off so that readout can occur without further charge accumulation during the readout process and in an interline transfer CCD, alternate columns of the detector array are masked off with opaque material. To read out the

image, the charge in each column of "light" pixels is simultaneously shifted into the adjacent "dark" column, which is blocked off from light. The drawback of full frame CCDs is that they require an external means of controlling exposure, typically a mechanical shutter. This increases camera cost, size and weight. Readout speed, which influences image noise, is also slower than for interline transfer CCDs.

3.4.2 Spatial Frequency Requirements for Hologram Recording

For a hologram to be effectively recorded, the CCD camera must be able to record all of the interference fringes in the hologram. From signal sampling theory, the Shannon criterion requires the sampling frequency be at least twice the highest signal frequency. This means that every interferometric fringe of the hologram has to be sampled by at least two pixels of the CCD array to resolve the fringe frequency. This requires that the angle between the object and reference waves must be sufficiently small.

The low spatial resolution of a CCD camera means that the maximum angle between the reference wave and the object wave is limited to just a few degrees. The maximum spatial frequency $f_{\max}$ which needs to be resolved, for the wavelength λ, is determined by the maximum angle $\phi_{\max}$ between the reference and the object beams as

$$f_{\max} = \frac{2}{\lambda} \sin \frac{\phi_{\max}}{2} \tag{3.53}$$

or

$$f_{\max} = \frac{1}{2\Delta x} \tag{3.54}$$

where Δx is the spacing between the pixel elements on the CCD sensor.

As the maximum spatial frequency $f_{\max}$ recorded on the CCD is limited by the pixel size of CCD, the angle between the reference and the object wave at any point of CCD must not exceed the maximum value $f_{\max}$. If however the angle between the reference and the object wave at some region of the CCD exceeds the maximum value $f_{\max}$, the interference fringe becomes under sampled, this often results in a decrease in the resolution of the reconstructed and therefore image aliasing appears in the recorded image of the fringe. Hence the distance between neighboring pixels is what limits the maximum angle between the object and reference wave. In digital holography, the angle between the reference and the object wave must be less than the maximum value $\phi_{\max}$.

The poor resolution of CCD cameras available today compares rather poorly to that of photographic plates. This creates a limitation on the maximum angle between object and reference beams and hence the spatial distance that can occur between the real and virtual image as shown earlier in Figure 3.4. As a consequence, the three diffracted waves often partially overlap during reconstruction. With the introduction of better CCD cameras this restriction will be less of a problem.

3.4.3 Lateral Resolution of the Optical System

The resolution of the reconstructed image depends on the information recorded in the hologram. However, this information is influenced by the size and the spatial resolution of the recording material. For digital holography the information recorded in the hologram is determined by the number of pixels and the pixel size and dynamic range of CCD [43,44].

In particular, the lateral resolution in digital holography is restricted by the pixel size of the CCD sensor and therefore the resolution of the reconstructed image is determined by the highest spatial frequency maximum of the object wave recorded by the CCD camera. For this reason, microscope lenses have to be applied for magnification of the object.

3.4.4 Calibration of Optical System in Digital Holography

The USAF 1951 Test Target shown in Figure 3.5 is one of the most commonly used resolution targets in optics. It allows one to determine the lateral resolution of the optical system. Table 1 displays the chart for the 1951 resolution target that specifies the frequency and line width for each group and element. This standard test object contains horizontal and vertical three-bar patterns in the form of a reflecting chromium coating set on a glass substrate. The vertical bars are used to calculate horizontal resolution and the horizontal bars are used to calculate vertical resolution.

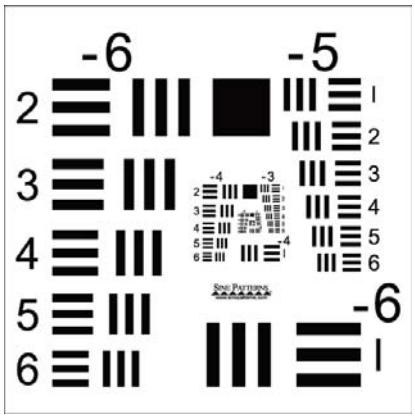


Figure 3.5: USAF1951 resolution test target pattern (see www.sinepatterns.com).

1951 USAF Resolution Target Data

Group	Element	Frequency	Line Width (um)	Group	Element	Frequency	Line Width (um)
-2	1	0.2500	2000	4	1	16.00	31.3
	2	0.2806	1782		2	17.96	27.8
	3	0.3150	1587		3	20.16	24.8
	4	0.3536	1414		4	22.63	22.1
	5	0.3969	1260		5	25.40	19.7
	6	0.4454	1122		6	28.51	.17.5
-1	1	0.5000	100	5	1	32.00	15.6
	2	0.5612	891		2	35.92	13.9
	3	0.6300	794		3	40.32	12.4
	4	0.7071	707		4	45.25	11.0
	5	0.7937	630		5	50.80	9.84
	6	0.8909	561		6	57.02	8.77
0	1	1.00	500	6	1	64.00	7.81
	2	1.122	445		2	71.84	6.96
	3	1.260	397		3	80.63	6.20
	4	1.414	354		4	90.51	5.52
	5	1.587	315		5	101.6	4.92
	6	1.782	281		6	114.0	4.38
1	1	2.000	250	7	1	128.0	3.91
	2	2.245	223		2	143.7	3.48
	3	2.520	198		3	161.3	3.10
	4	2.828	177		4	181.0	2.76
	5	3.174	157		5	203.2	2.46
	6	3.564	140		6	228.1	2.19
2	1	4.00	125	8	1	256.0	1.95

Table 3.1: Specification data for the groups and elements of a 1951 USAF resolution test target.

The resolution target is a vital instrument in the calibration of the holographic optical system. It allows determination of parameters such as the image size, number of pixels and the lateral resolution of the hologram and reconstructed image. The methodology to obtain the correct image size and number of pixels is conducted as follows: One calculates the number of pixels on the CCD camera required to draw a square box around a specific element made up of 2.5 line pairs. From Table 1 we are then able to calculate the physical dimensions of the square box corresponding to the number of pixels on the camera. One then increases the required image size to use the highest number of square pixels on the camera.

Chapter 4

Digital Gabor Holography for Microscopy

The structure of this chapter is as follows: Section 4.1 introduces the background of Gabor holography and the theoretical basis for recording and reconstruction of the hologram. Section 4.2 presents a simulation of digital Gabor holography using a theoretical object. Section 4.3 examines some of the main applications and limitations of the Gabor holographic technique. In Section 4.4 technical details about the experiment are discussed. Section 4.5 presents a host of applications of digital Gabor holography in microscopy. Section 4.6 looks at the use of digital Gabor holographic movies for microscopy and demonstrates some of the technical advantages over conventional microscopy. Finally Section 4.7 summarizes and concludes the chapter.

4.1 Introduction to Digital Gabor Holography

Gabor or 'in-line' holography as it is commonly known provides a simple method of high-resolution imaging that routinely achieves both micron level, lateral and depth resolution in three-dimensional imaging. The same beam serves as the reference wave and illuminates the object; no mirrors or beam splitters are strictly needed, but can add more capabilities to the holographic optical system.

In Gabor holography, the object is required to be mostly transparent with small thin obstructions, which diffract and scatter the illuminating beam. The light wave

arriving at the detector plane consists of the un-diffracted plane wave, the reference (R) and the diffracted object wave (O), which interfere to give the holographic interference pattern, $|H|^2$ at the CCD camera

$$|H|^2 = |R+O|^2 = |R|^2 + |O|^2 + R^*O + RO^* \tag{4.1}$$

The reference term $|R|^2$ on the right side of Equation (4.1) is removed by exposing another frame with the object removed, whereas the $|O|^2$ term cannot be removed. However, by assuming that the object is mostly transparent then this term contributes only a small amount of background noise. The two remaining terms are the virtual image R^*O and the real image RO^*. These terms are spatially superimposed in the reconstructed field, which results in the loss of phase information. Therefore when performing numerical focusing in the reconstruction of the hologram, the defocused blur from the conjugate image although weak, is problematic as the superposition of these terms effectively blurs the details and structures of the reconstructed object wave. This effect combined with the need for a mostly transparent object to minimize the $|O|^2$ term is the main limitation of Gabor holography.

4.2 Gabor Holography Simulation

Figure 4.1 shows a simulation of Gabor holography using a cross and circle as the theoretical object. For a mostly transparent object (Fig.4.1a), this results in a positive image against dark background (Fig.4.1b). On the other hand, if one uses a Michelson interferometer set up, with a separate reference wave, as in off-axis, then both of the first

two terms on the right side of Equation (4.1) can be subtracted, and the object – and the holographic image – is dark against light background (Fig.4.1c). For a mostly opaque object (Fig.4.1d), there is not enough un-diffracted reference in Gabor holography to generate image correctly (Fig.4.1e), where as the Michelson holography generates a correct image (Fig.4.1f).

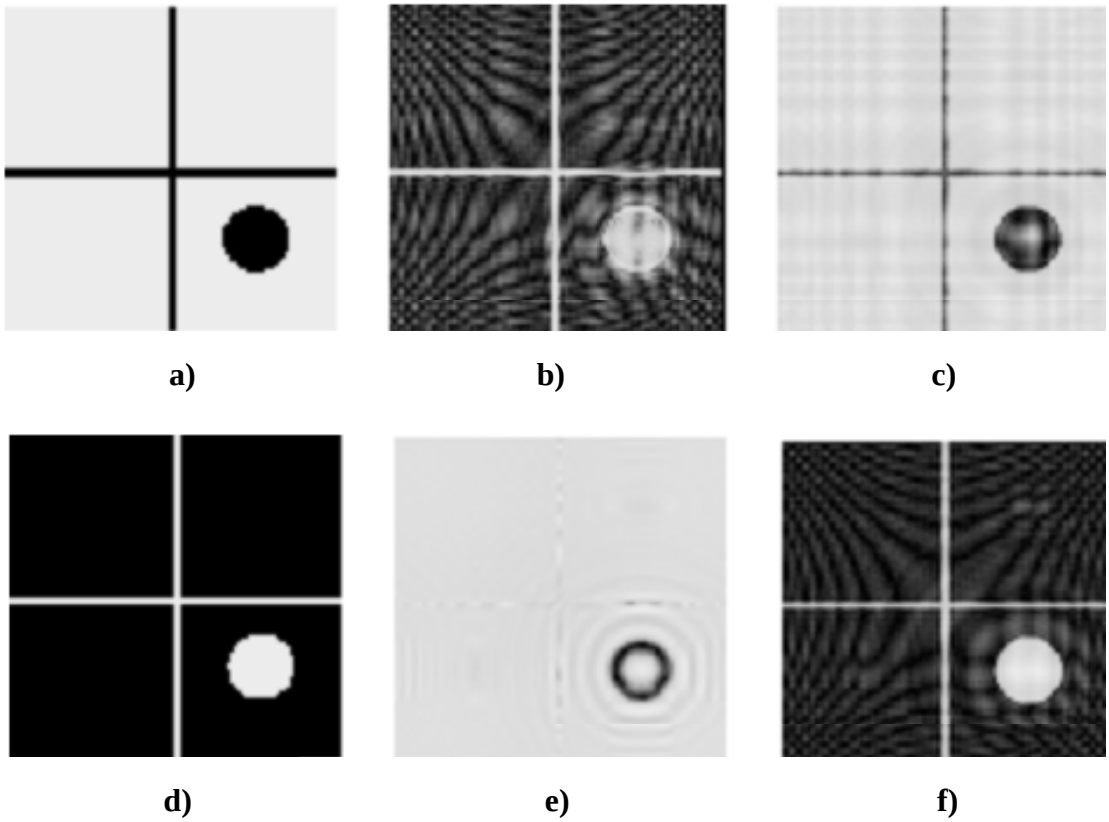


Figure 4.1: Gabor and Michelson holography of mostly transparent object: a) mostly transparent object; b) Gabor holographic image; c) Michelson holographic image.;
Gabor and Michelson holography of mostly opaque object: d) mostly opaque object; e) Gabor holographic image; f) Michelson holographic image.

4.3 Limitations of Gabor Holography

Gabor holography is seen to suffer from certain limitations which restrict the extent of its applicability. From the simulation in Figure 4.1 one can see that Gabor holography is best suited towards the imaging of low density, small discrete opaque objects such as particle fields or larger but highly transparent objects such as biological cells. This type of object requirement minimizes the amount of noise from the disturbing $|O|^2$ term. Besides being limited to specific types of objects, a second limitation of Gabor holography is the overlapping conjugate image as depicted in Figure 4.2. The problem lies with the fact that they are not spatially separable as is the case in off-axis holography. When the real image is brought into focus it is accompanied by the out of focus virtual image. Alternately, an observer looking at a focused virtual image also sees a defocused real image. The conjugate image however is generally not considered to be a significant problem in Gabor holography because its effect at large enough recording distances appears as the background signal in the reconstructed image which is often negligibly small. The in-focus image amplitude to out-of-focus conjugate image amplitude ratio is proportional to the inverse of the recording distance *z*. At small recording distances, the contrast between the image and its conjugate image is low and both images merge. Therefore, the recording distance must be sufficiently large for adequate suppression of the disturbing conjugate image.

A number of methods have been developed for removal of the conjugate image and recovery of the phase information [45,46]. If the phase distribution in the hologram plane is recovered then it is possible to reconstruct the whole wave field exactly.

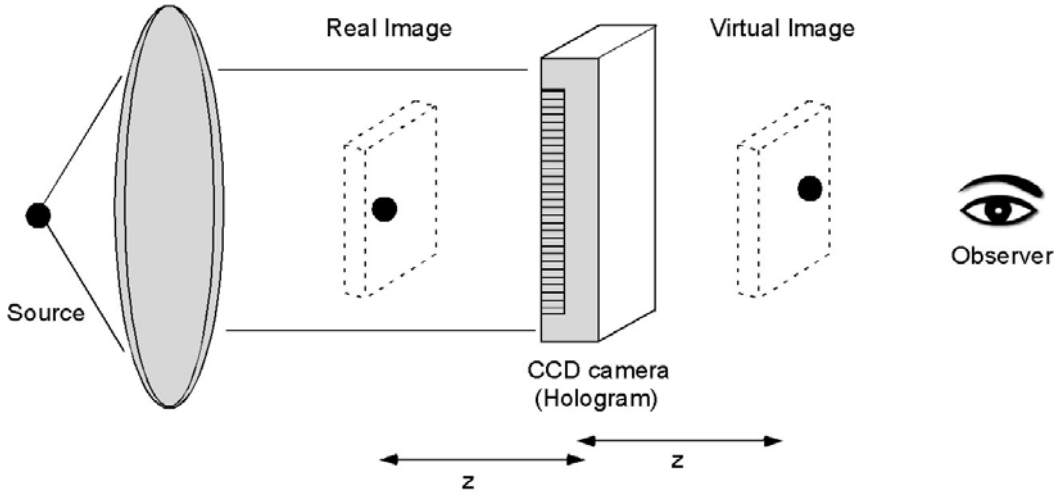


Figure 4.2: Formation of the conjugate image in a Gabor hologram.

Phase-shifting digital holography is one technique in particular which has attracted widespread interest [47]. In the phase-shifting technique, the *dc* term and conjugate image are removed through multi-exposure holographic recording while shifting the phase of the reference field by an integer fraction of 2π. While this method has the disadvantage of increasing the experimental effort; a more serious limitation for practical use in microscopy is the requirement of multiple exposures, therefore hindering the investigation of dynamic objects.

Although Gabor holography does not record the phase, it is still very useful for the study of suitable objects at high resolution. By taking advantage of the large depth of field and the plane-to-plane numerical reconstruction capability of digital holography, one can produce 4-D representations of the paths followed by micron-sized objects such as biological samples and particles [48].

4.4 Recording of Digital Gabor Holograms

Figures 4.3 and 4.4 depict the optical setup for hologram recording in the digital Gabor holographic optical system in transmission geometry. The 532nm coherent light from a frequency doubled Q-switched Nd:YAG laser (Continuum Minilite II, shown schematically with specifications in Figures 4.5 and 4.6), operating at a power of ~2mJ/pulse is used for hologram recording.

The main advantage of using a pulsed laser is that it can record an object field at selected times and so is able to serve as a useful tool in studying the dynamics of fast phenomena. To ensure proper timing in the experiment, a digital delay/pulse generator (Stanford DG535) is used to synchronize the pulsed laser and digital camera.

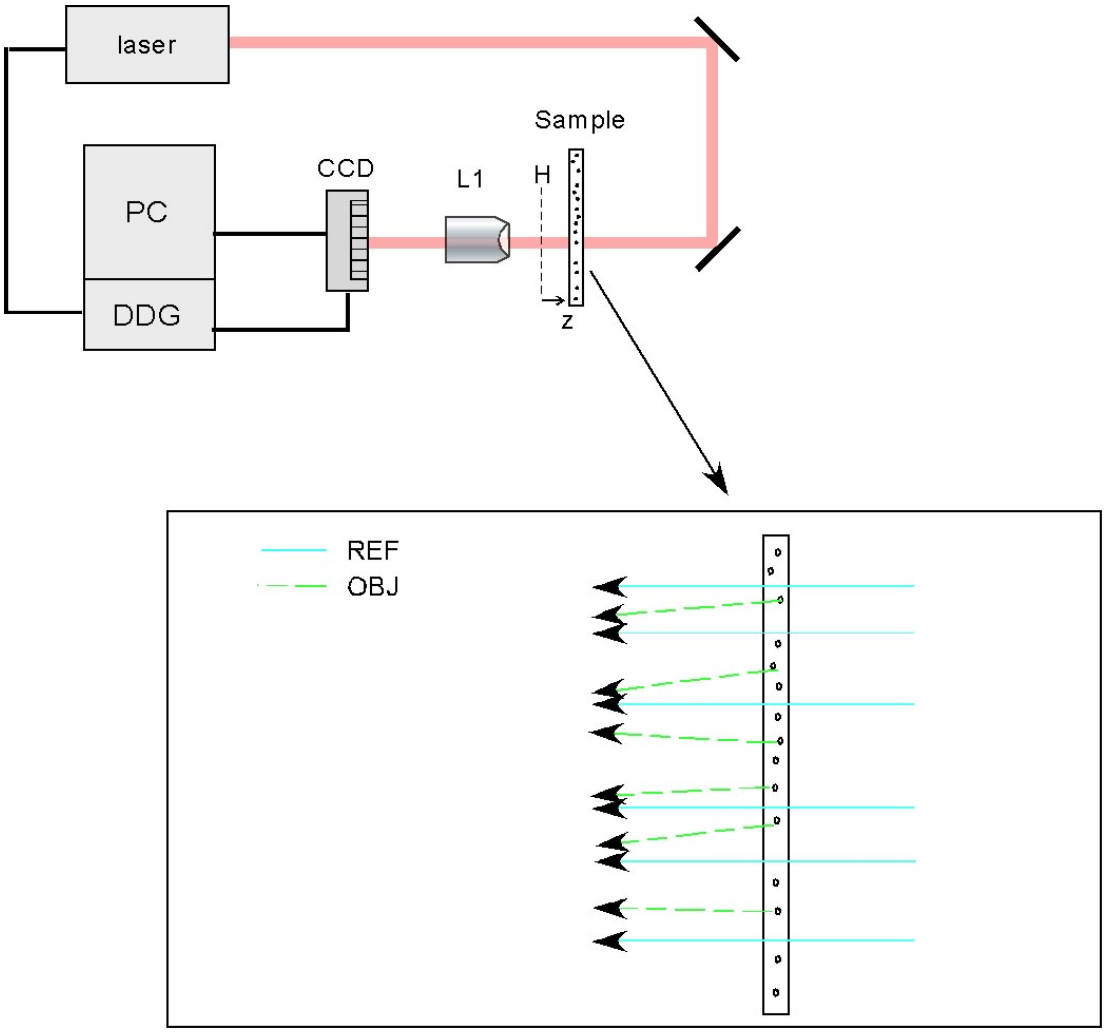


Figure 4.3: Experiment setup for recording of digital Gabor holograms.

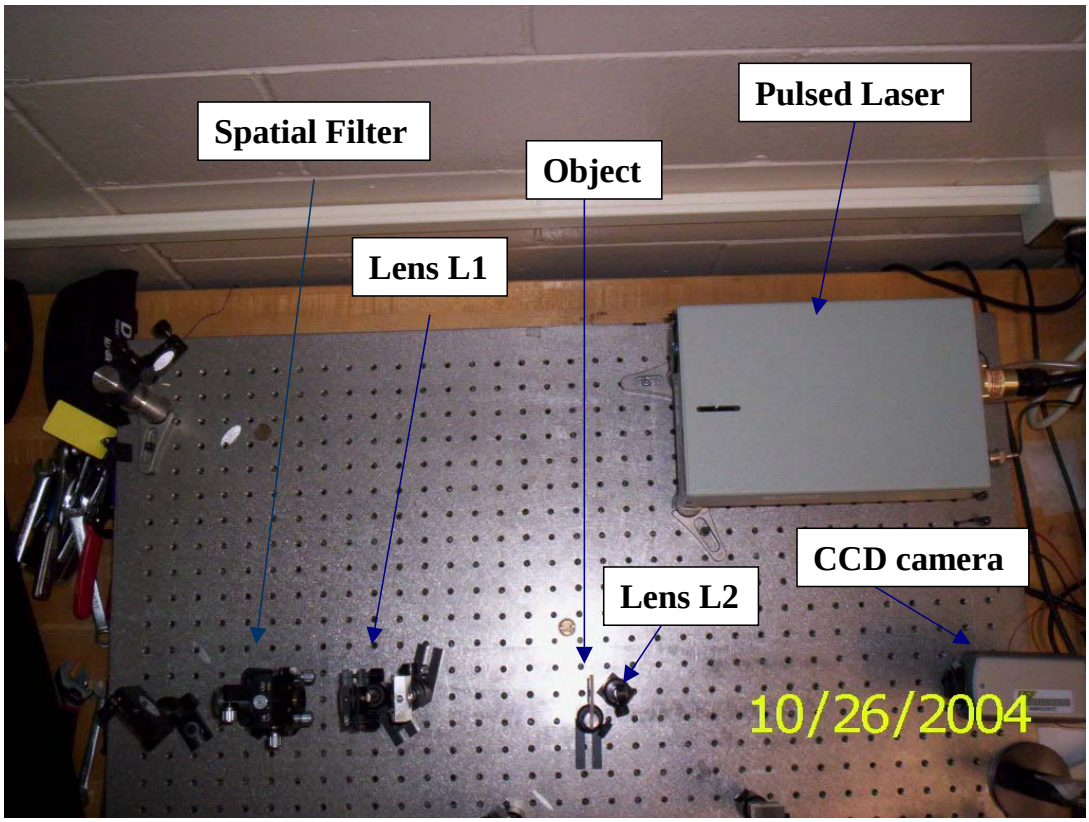
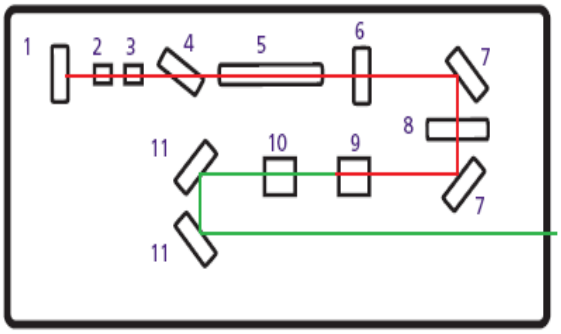


Figure 4.4: Birds-eye photograph of experimental setup for digital Gabor holography.

MINILITE™ OPTICAL LAYOUT



1. Rear Mirror
2. Pockels Cell
3. 1/4 Wave Plate
4. Dielectric Polarizer
5. Rod
6. Output Coupler
7. IR Mirror
8. Attenuator
9. Second Harmonic Generator
10. Third or Fourth Harmonic Generators
11. 532, 355 or 266 nm Mirrors

Figure 4.5: Optical layout of the Contiuum Nd:Yag Minilite pulsed laser.

DESCRIPTION	MINILITE™ I	MINILITE™ II
Repetition Rate (Hz)	1-15	1-15
Energy (mJ)		
1064 nm	28	50
532 nm	12	25
355 nm	4	8
266 nm	2	4
Pulsewidth[1] (nsec)		
1064 nm	5-7	5-7
532 nm	3-5	3-5
355 nm	3-5	3-5
266 nm	3-5	3-5
Linewidth (cm^{-1})	1	1
Divergence [2] (mrad)	< 3	< 3
Rod Diameter (mm)	3	3
Jitter [3] (±ns)	0.5	0.5
Energy Stability [4] (±%)		
1064 nm	2; 0.6	2; 0.6
532 nm	3; 1.0	3; 1.0
355 nm	4; 1.3	4; 1.3
266 nm	8; 2.6	8; 2.6
Polarization		
1064 nm	HORIZONTAL	HORIZONTAL
532 nm	VERTICAL	VERTICAL
355 nm	HORIZONTAL	HORIZONTAL
266 nm	HORIZONTAL	HORIZONTAL

Figure 4.6: Continuum Minilite laser optical layout and specifications.

Three channels on the delay/pulse generator are used for triggering, channels T_0, A and B as displayed in Figure 4.7. Channel T_0 is used for triggering of the flash-lamp and Channel A is used to trigger the Q-switch on the pulsed laser. Channel B is used to trigger the CCD camera to capture a frame. In triggering each channel a TTL pulse is used. A delay of 250μs between Channels T_0 and A fires the pulsed laser at maximum output. Channel B is delayed for 80ns with respect to Channel A. The main purpose of this type of triggering is that it allows the speed of frame capture on the CCD to synchronize with the dynamics of the object. The CCD camera is a Sony DFW-V500 YUV/Monochrome, 640x480 pixels with pixel size 7.4μm x 7.4μm. The camera can switch between frame rates of 3.75, 7.5, 15, 30 frames per second and also one shot.

A pulse fired from the laser is reflected from mirrors M1 and M2 of quality $\lambda/10$. The spatial filter effectively brings the beam into sharp focus by means of a 20x microscope objective with focal length 9.0mm. A 5μm pinhole placed at the focal plane of the microscope objective constitutes the filter. This acts to remove the unwanted multiple-order energy peaks and pass only the central maximum of the diffraction pattern. This is a diffraction plane in itself. A lens of focal length f = 3.0cm is used to collimate the beam from the spatial filter. The coherent plane wave then propagates through the object in the setup which is located a distance z from the focal plane of the CCD detector, *H*. The light scattered by the object serves as the object wave and the un-scattered light serves as the reference wave. The interference between these two waves forms a geometrically magnified diffraction pattern on the CCD sensor using a 20x lens (L1). Only a small amount of spherical aberration is introduced to the image since both reference and object beam angles are normal to the recording plane. The role of

reconstruction is to obtain the 3-D structure of the object from the 2-D hologram on the screen or, in physical terms, to reconstruct the wave front at the object. In the reconstruction stage the same plane wave is diffracted numerically. A number of LabView and Matlab programs are used for control of the experiment for numerical computations and for presentation of the resulting images.

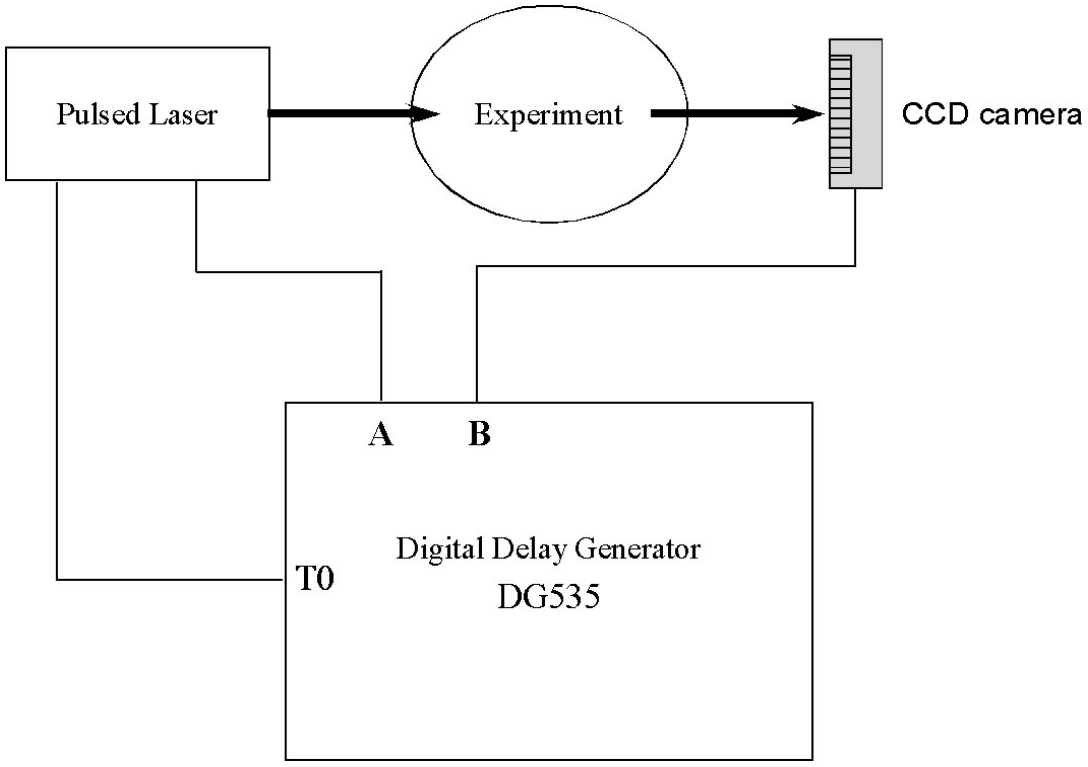


Figure 4.7: Triggering the system using the pulsed laser in order to capture dynamic events on the CCD.

4.5 Experimental Results of Gabor holography

Figures 4.8(a) and 4.8(b) are the hologram and the reconstructed amplitude image of a U.S. Air Force (USAF) 1951 positive resolution test target recorded with the experimental setup in Figure 4.3. The image area is 250 x 250 μm^2, recorded with 464 x 464 pixels and the image is reconstructed using the Huygens convolution algorithm at z = 1.30mm from the hologram. The determination of the best focus for the reconstructed image is performed by observation alone.

After image capture of the hologram and digital transfer to the computer, the object is removed from the experiment setup and twenty consecutive images of the reference beam are captured. The reference beam images are then digitally transferred to the computer and averaged to facilitate the removal of the $|R|^2$ term from the holographic interference in Equation (4.1). This in turn leaves

$$|H|_{-r}^{\ 2} = |O|^2 + R^* O + R O^* \tag{4.2}$$

The *dc* term is mainly made up by the contribution of the reference wave, and therefore this term is largely eliminated by this subtraction of the average intensity of the hologram. Note however, that this method does not fully remove the *dc* term. The contribution of the $|O|^2$ term remains and cannot be removed.

The smallest observable reconstructed vertical three-bar pattern of the resolution target in the reconstructed amplitude image is that of the third element of group 7 which corresponds to a spatial frequency of 161 line pairs per mm, a line width of 3.1μm and a line length of 0.0155mm (see Table 3.1 in Chapter 3). The details of this element are

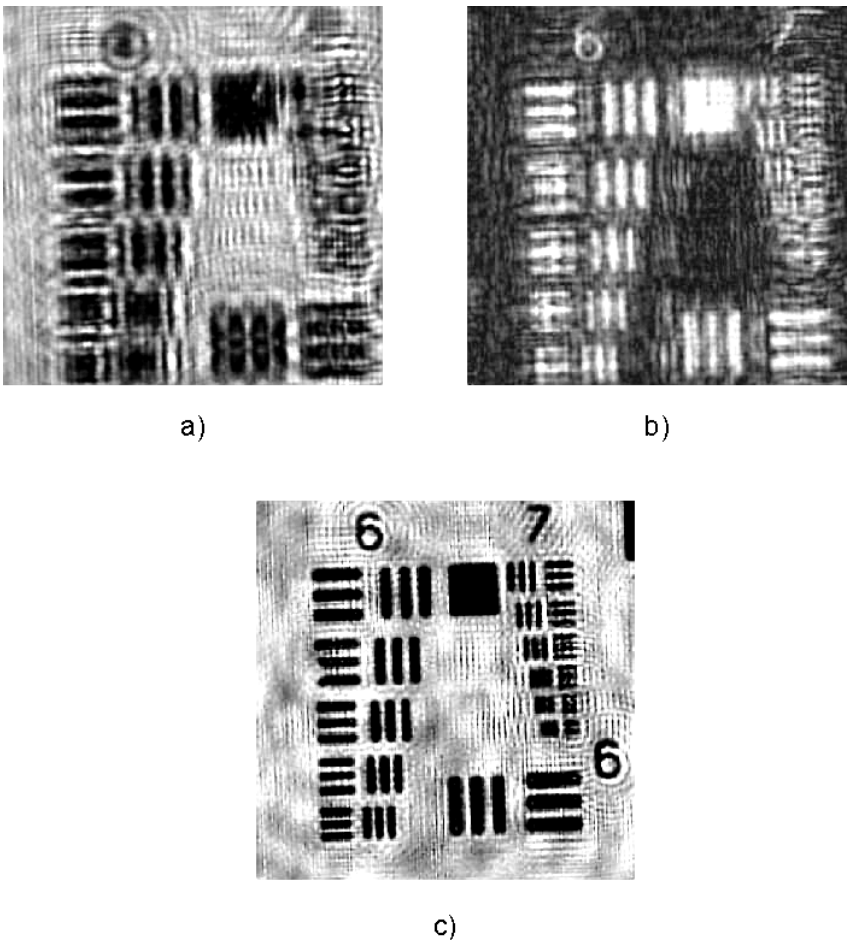


Figure 4.8: The hologram (a), the reconstructed amplitude image (b), and the direct image (c), of groups 6 and 7 of USAF 1951 positive resolution Target (Area =250μm x 250μm^2, Pixels = 464 x 464, z = 1.30mm)

clearly observable in the reconstructed amplitude image. This agrees well with the predicted resolution limit of around 2μm using;

$$\Delta\xi = \frac{\lambda z}{N\Delta x} \sim 2\mu m \tag{4.3}$$

where $\Delta\xi$ is the lateral resolution, λ is the wavelength =532nm, z is the distance between the object and the CCD array and Δx is the CCD pixel size $= 7.4\mu m$.

Comparing the reconstructed amplitude image to the direct image as displayed in Fig. 4.8(c), one sees that the image quality of the reconstructed amplitude image is degraded. The opaque bars of the resolution target occupy a significant portion of the image area thus reducing the overall image transparency and as a result there is not enough un-diffracted light to reconstruct the image correctly. In addition, the localized presence of the defocused conjugate image in the reconstructed image plane creates a blur around the bars in the image.

Figures 4.9(a) and 4.9(b) are the hologram and reconstructed amplitude image of groups 6 and 7 of the resolution target. The image area is 250 x 250 μm^2, recorded with 464 x 464 pixels and the image is reconstructed using the Huygens convolution algorithm at z = 2.20mm from the hologram. Note the relatively higher image quality of the amplitude reconstruction compared to that shown in Figure 4.8(b). This could be because of the larger recording distance used in this example. However, the image is still largely degraded due to the reduction in overall transparency of the image.

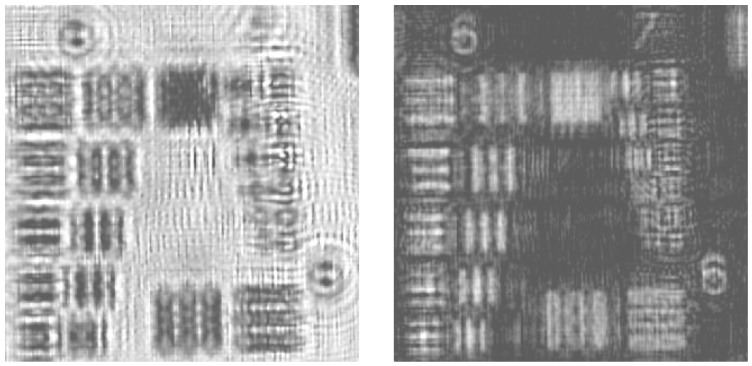


Figure 4.9: The hologram (a) and reconstructed amplitude image (b), of groups 6 and 7 of the USAF 1951 positive resolution Target.
(Area =250μm x 250μm^2, Pixels = 464 x 464, z = 2.20mm)

Figure 4.10 displays the Fourier transform of the hologram in Fig.4.9a). The overlap of the *dc* and holographic terms are seen. Most of the energy in the image is concentrated at low frequencies from the *dc* term, most of which can be removed by subtracting the average of the reference beam.

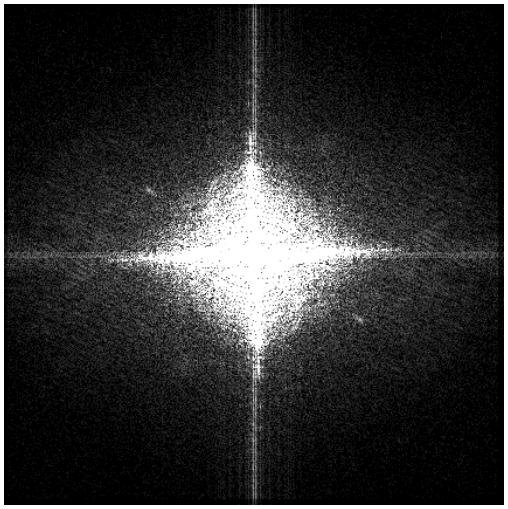


Figure 4.10: Fourier Transform of the digital Gabor hologram in Fig.4.9 (a).

Figures 4.11(a) and 4.11(b) are the hologram and the reconstructed amplitude image of an assembly of onion cells displaying an area of 400μm x 400μm^2, recorded at 456 x 456 pixels, and reconstructed using the Huygens convolution algorithm at a distance of $z = 2.50mm$. The onion cells size and structure are well established, easily identifiable and range from 0.25 to 0.4 millimeters in length. The samples were mounted by cutting into the surface of the fresh onion material, and peeling off a layer of the outermost cells with a razor blade. The sample was then moistened, and applied directly between the cover-slip and microscope slide.

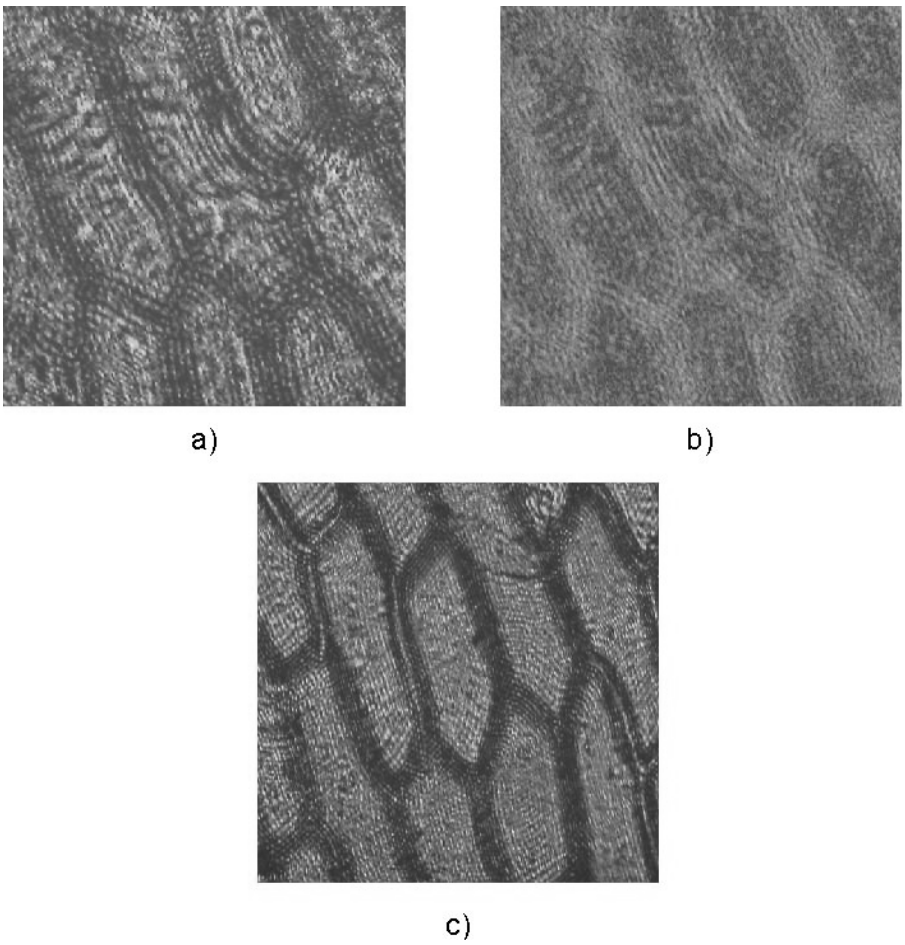


Figure 4.11: The hologram (a), the reconstructed amplitude image (b), and the direct image (c) of a collection of onion cells
(Area =400μm x 400μm^2, Pixels = 456 x 456, z = 2.50mm)

The amplitude reconstruction is able to partially reconstruct the conspicuous cell walls of the cells. Nonetheless, when compared to the direct image in Figure 4.11(c) one can see that the image quality is again rather poor. There is the presence of large scale residual features in the interior of the cell, which suggest that the cells may be dead or dying and this introduces a significant amount of coherent noise into the image. These residual features also reduce the overall transparency of the cell, and in turn generate

adverse effects on the Gabor reconstruction due to the disturbing $|O|^2$ term.

The large depth of field of digital Gabor holographic microscopy with numerical reconstruction provides an ideal tool for the study of small discrete objects at different focal planes. The hologram and the reconstructed amplitude image of copolymer microspheres with a mean diameter of 9.6μm are shown in Figures 4.12(a) and (b) respectively, each displaying an area of 300μm x 300μm^2, recorded at 364 x 364 pixels, and reconstructed at a distance of z = 5.20mm using the Huygens convolution method. The microspheres were prepared by mounting in a thin layer of deionized water between a microscope slide and a glass cover-slip. The diffraction rings of each particle are recognizable from the hologram. When the distance between the particles and the CCD sensor is increased, the spacing between concentric rings belonging to an individual particle also increases. Figure 4.12(c) shows the direct image of the microspheres.

In comparison with the direct image the reconstructed amplitude image clearly resolved all isolated spheres as well as the two spheres that appear to be in direct contact. The fact that the overall transparency of the hologram is increased minimizes the degrading effect of the $|O|^2$ term and results in a higher quality image reconstruction when compared to that of the previous examples of the resolution target and onion cells.

Figures 4.13(a) and Figure 4.13(b) are the hologram and the reconstructed amplitude image of a particle field containing differing densities of copolymer microspheres within the image. The image area is 300 x 300 μm^2, recorded with 364 x 364 pixels and the image is reconstructed using the Huygens convolution algorithm at z =

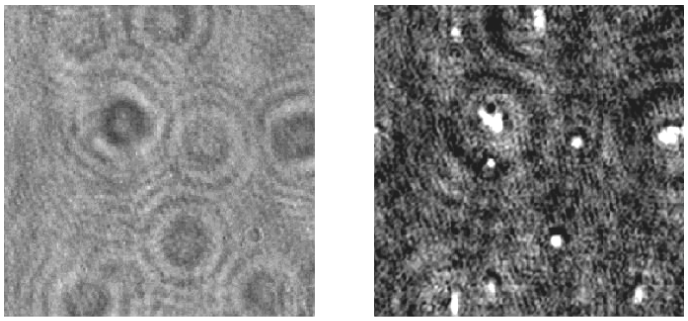


Figure 4.12: The hologram (a), the reconstructed amplitude image (b), and the direct image (c) of copolymer microspheres.
(Area =300μm x 300μm^2, Pixels = 364 x 364, z = 5.20mm)

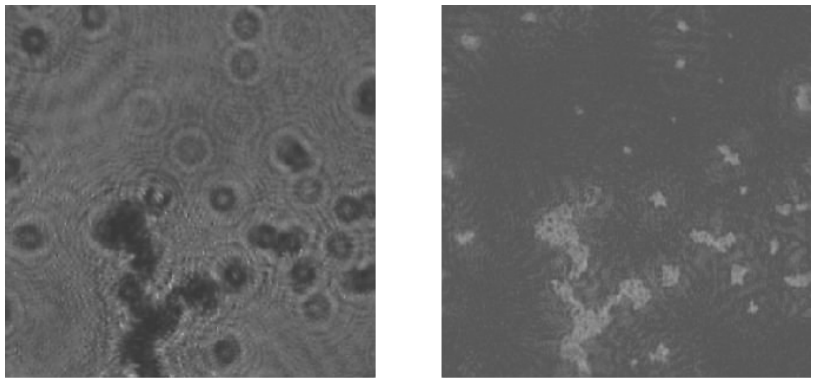


Figure 4.13: The hologram (a) and the reconstructed amplitude image (b) showing differing densities of copolymer microspheres
(Area =300μm x 300μm^2, Pixels = 364 x 364, z = 5.60mm)

5.60mm from the hologram. The speckle noise in the image increases with the particle density. Therefore for a large group of small objects such as the microscopheres clustered

together as in the bottom part of the images in Fig. 4.13 the speckle intensity may actually be comparable to the focused particle intensity and this makes individual particle detection complicated. One may see this effect in Fig.4.13 in the amplitude reconstruction

Figures 4.14(a) and Figure 4.14(b) are the hologram and the reconstructed amplitude image of a paramecium, a small unicellular organism that is found in freshwater ponds. The image area is 250μm x 250μm^2, recorded with 464 x 464 pixels, and reconstructed at a distance of $z = 0.90\text{mm}$ using the Huygens convolution method.

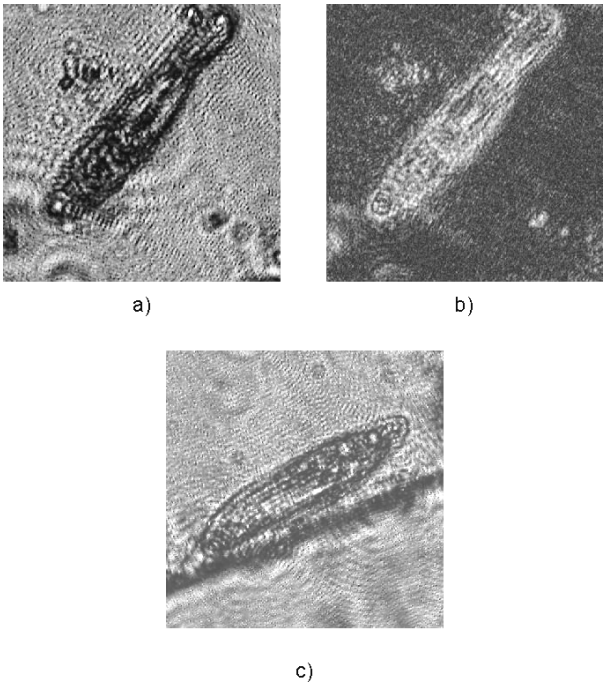


Figure 4.14: The hologram (a), the reconstructed amplitude image (b), and the direct image (c) of a paramecium
(Area =250μm x 250μm, Pixels = 464 x 464, z = 0.90mm)

Paramecium belong to the protozoa family and range in size from approximately 100-300μm in length. These organisms contain many complex and interesting features within a single cell as shown in the anatomical diagram presented in Figure 4.15 [49].

In the reconstructed amplitude image it is difficult to identify specific intra-cellular features as depicted in the anatomical diagram. However, when one makes a comparison with the direct image in Fig. 4.14(c) it becomes clear that the paramecium contains a number of unresolved intracellular particles with a large range of structural scales which effectively increase the speckle noise and simultaneously reduces the transparency of the paramecium.

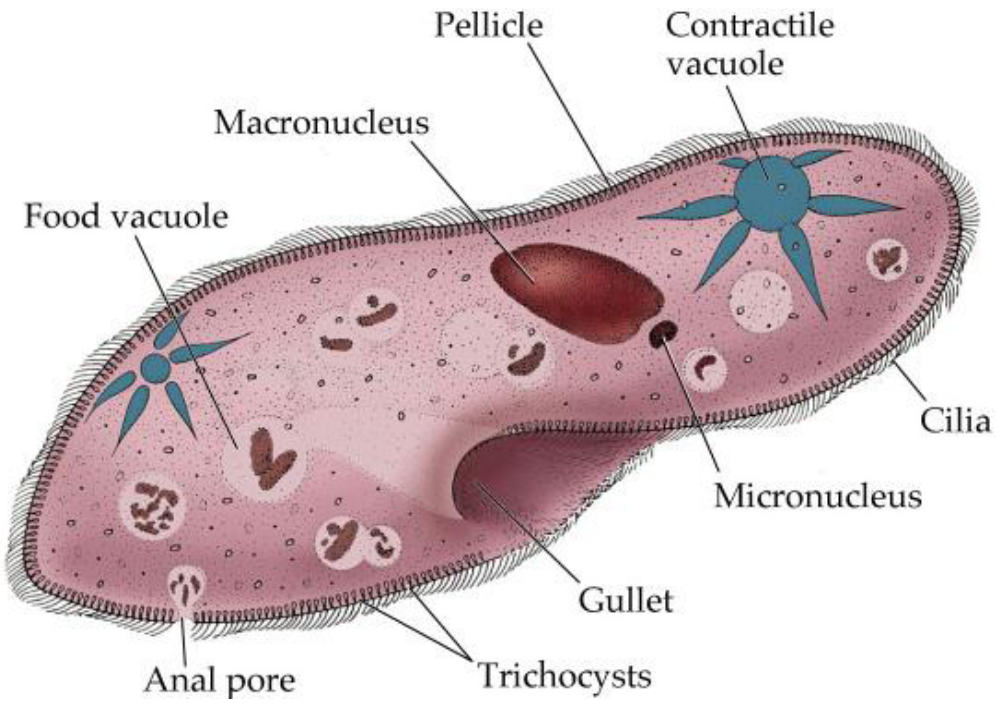


Figure 4.15: Anatomy of Paramecium.

4.6 Digital Gabor Holographic Movies

Gabor holography offers both a rapid and efficient approach in recording digital holograms. This property is particular useful in capturing the motion of moving objects in depth. A digital hologram contains information of the whole optical field in a three-dimensional image volume, which allows for calculation of the image field in any focal plane. A time series of digital holograms then provides complete four-dimensional information of the object's three-dimensional spatial images as well as the time evolution of those images. Once a movie of digital holograms are recorded, one can reconstruct the images of the object and be able to focus on any focal plane as a specimen under observation moves up and down in the image space. This is the most unique capability of digital holographic movie in contrast to conventional video microscopy, where only the images of the specific focal plane that was used in recording are preserved and the information of all the other planes is lost. This is a critical advantage especially in the microscopic recording of animated microbes that are constantly swimming in and out of a large range of focal distances. In conventional video microscopy, it is not feasible to track the focal distances of rapidly moving microbes, whereas in digital holography a series of holograms can be recorded at a fixed distance and the images are later numerically focused at leisure and one is able to track the rapidly moving microbes as needed.

To demonstrate these technical advantages for the creation of holographic movies, Figure 4.16 displays a time series of twelve holograms showing a microbe moving progressively through pond water between each of the hologram frames. The image area of each frame is $100\text{x}100\mu\text{m}^2$, recorded at 428 x 428 pixels and each frame is captured with an interval of $t = 0.40s$ apart.

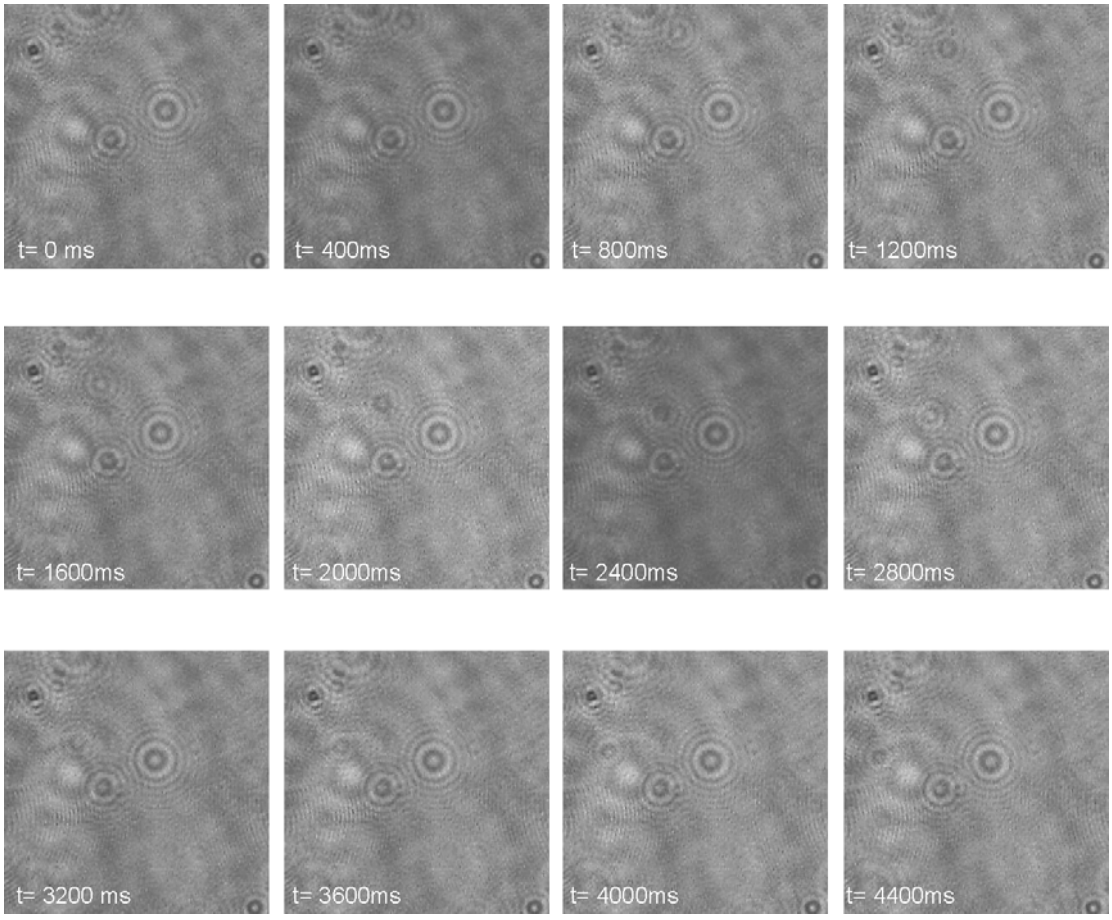


Figure 4.16: Time-series of holograms of a microbe moving progressively in pond water. (Area =100μm x 100μm^2, Pixels = 428 x 428)

The holograms are recorded by synchronizing the camera with the laser pulses under the control of the digital delay generator. The time-series amplitude images in Figure 4.17 corresponding to the holograms in Figure 4.15 are reconstructed using the Huygens convolution method adjusting image distances in the range z = 0.352-0.354mm for best focus of the microbe (highlighted in circle) under observation. The images show the track of the specimen to demonstrate

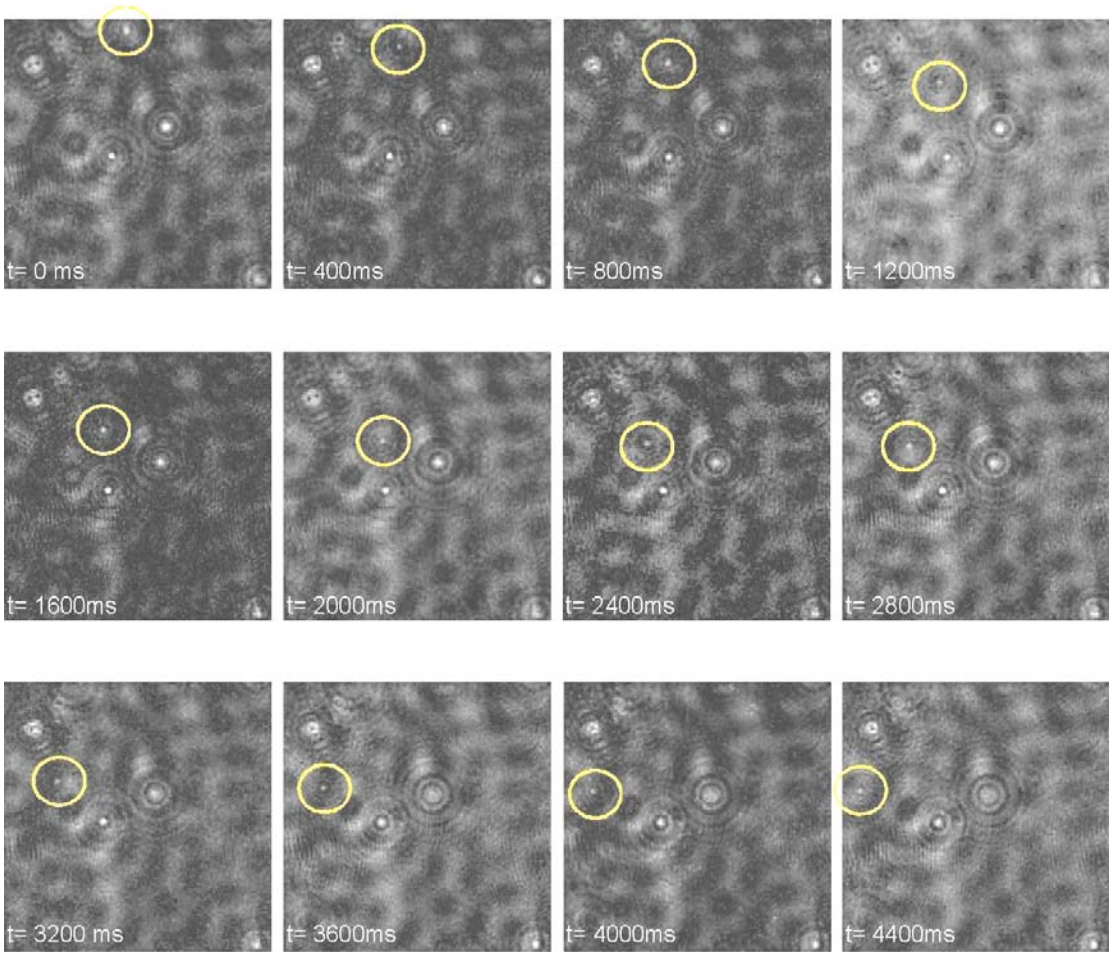


Figure 4.17: Time-series of reconstructed amplitude images of a microbe moving in pond water.
(Area =100μm x 100μm^2, Pixels = 428 x 428, z = 0.352-0.354mm)

that time resolution is achieved. The reconstructed amplitude images are then composed into AVI files to produce the holographic movie in with the microbe is consistently in focus even though it is moving through focal planes in depth. This is not possible with conventional microscopy.

The next biological example concerns euglena, a single-celled organism with spindle-shaped bodies ranging in size from 0.025 to 0.050mm in length. Figure 4.18 (a-d) shows a time series of holograms recorded 0.40s apart, displaying a number of euglenas at different focal depths swimming in pond water. Figures 4.18 (e-h) are the corresponding amplitude images, reconstructed using the Huygens convolution method adjusting image distances in the range z = 0.464-0.471mm for best focus of the individual euglenas. Reconstruction at different depths throughout the volume brings different individuals of euglena into focus. One can see how we have focused on one particular euglena moving between frames. Note how the other euglenas one of which appears to be in focus in the holograms now appear out of focus in the reconstructed amplitude frames as they are located at different depths to the plane of focus. In addition, one may observe smaller objects swimming about. Closer examination reveals it to be algae of about 10μm in diameter.

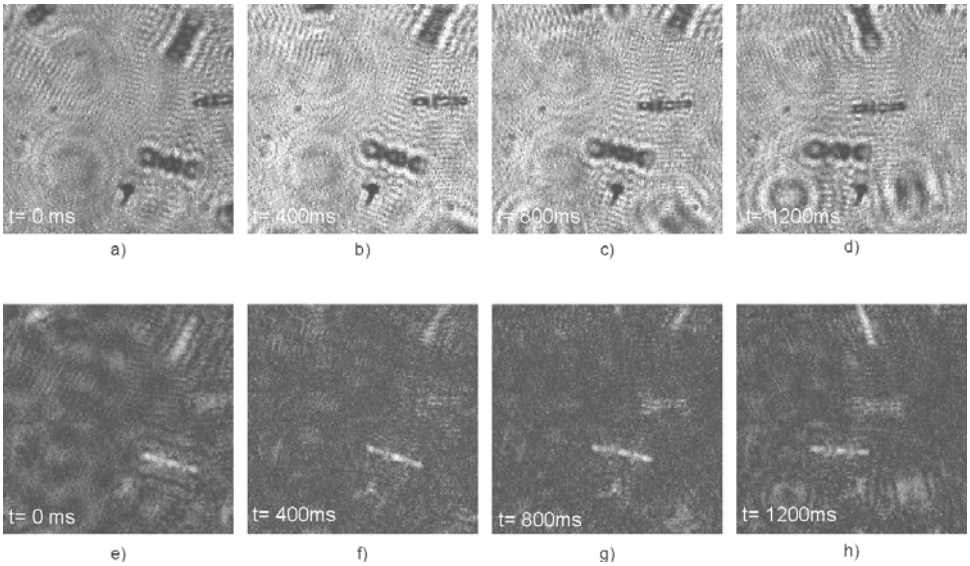


Figure 4.18: Time-series of holograms (a-d) and their corresponding reconstructed amplitude images (e-h) of a number of euglenas swimming in different focal planes. (Area =200μm x 200μm^2, Pixels = 400 x 400, z = 0.464-0.471mm)

Any of the individual euglenas can focused throughout the entire movie. This is a significant advantage for a biologist interested in a number of objects moving in depth during a time sequence. In comparison to a frame captured from conventional video microscopy in Figure 4.19, the information of euglenas at different focal planes to the one viewed in the figure are effectively lost.

Figure 4.19: A single frame of euglenas in pond water captured by a conventional video microscope

4.6.1 Automatic Focusing for Digital Gabor Movies

After a number of holograms have been recorded, it remains to reconstruct each of the holograms. If an object recorded in the holograms is moving in depth, then the reconstruction distance is required to be changed between each holographic frame. For a user to reconstruct for example hundreds of holograms for the objects best focus by observation alone, then this becomes a time consuming and tedious procedure.

An interesting property about Gabor holograms is that in the reconstructed amplitude image the object appears bright against a mostly dark background. Coupled with the fact that only specific objects with certain characteristics are suitable for imaging with Gabor holography it is possible to construct auto-focusing algorithms that use the contrast and other properties of the image to find the best focus of the object. Upon examining the frequency components of an unfocused image, one will find that there are relatively few high frequency components. As the image comes into focus, high frequency components increase.

An example of the use of an auto-focus program is illustrated in Figure 4.20. The figure displays the hologram and the reconstructed amplitude image of the resolution target. The program uses two approaches. The first one uses frequency content to determine edge sharpness. The other method uses an edge detector and measures the standard deviation of the intensity of the edges, the larger the value, the better the focus. The program works by iteratively changing the reconstruction distance starting from a user specified value to find the largest numerical values of contrast and therefore find the best focus.

The program algorithm is not able to find the correct best focus in the figure. The starting z value was already fairly close to the actual reconstruction plane for best focus. The methodology is also probably made more difficult because of how the noise effects different parts of the image and the sheer number of objects involved may also cause problems.

There are a number of other approaches that require investigation and other

contrast methods such as squared gradient, absolute variation, or Laplacian methods which may yield better results. The successful integration of a program that automatically identifies the correct focus of an object will create a significant improvement for the application of this technique in holographic movies. This is a subject of future study.

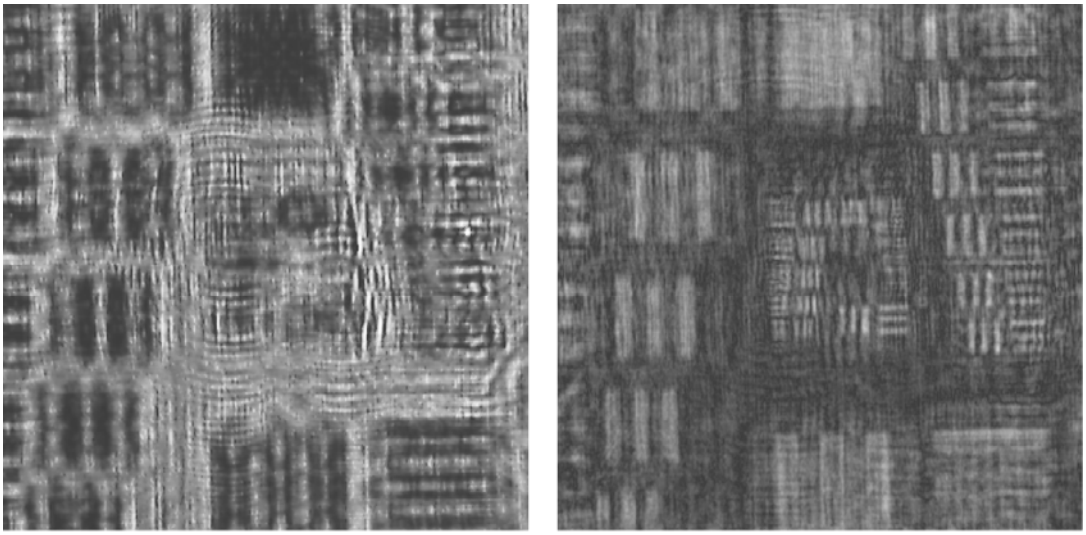


Figure 4.20: Auto-focusing of the resolution target.

4.7 Conclusions of Digital Gabor Holography

In this chapter, we have described digital Gabor holography for applications to biological microscopy. The holographic optical system records Gabor holograms directly from an object onto a CCD camera and then reconstructs the object wave-field numerically by use of the Huygens convolution method. Although there is a minimum reconstruction distance requirement associated with the Huygens convolution method, this is insignificant since large recording distances are usually required in order to

minimize the merging of the holographic terms. The superposition between the focused holographic image and the unfocused conjugate image cannot be avoided and there is, inevitably, some noise introduced into the image that cannot be avoided. However, by increasing the recording distance, it is possible to partially suppress the conjugate image so that it does not disturb the image to a significant degree.

The microscopic capability of the Gabor holographic system is demonstrated through the experimental results of reconstructed amplitude images of diverse objects such as a resolution target, copolymer microspheres, onion cells, and protozoa. The simulation and experimental results clearly show that only specific types of objects with certain physical characteristics are suitable for imaging with Gabor holography. These objects are required to be either discrete, small opaque objects, or larger but mostly transparent objects. In either case the overall transparency of the image must be significant enough so that the $|O|^2$ term approximates a small amount of background noise.

Biological samples such as paramecium and onion cells, which are considered by microscopy terms to be mostly transparent, actually contain a large range of intracellular features which create speckle noise in the coherent imaging system. Also, these features, in turn, reduce the transparency and cause further degradation in image quality.

Digital Gabor holographic movies are seen to offer a rapid and efficient approach for image capture of a time-series of holograms. In the experimental results, we have demonstrated that high-resolution tracking of many objects in 4-D from a single hologram data set can be performed. High lateral resolution is achievable, and, thus,

tracking of organisms as small as bacteria may be possible. By the use of numerical focusing, each single hologram in the data set can be reconstructed at selected depths so that the trajectory and speed of a collection of objects can be captured as 3D data sets by stacking of the 2D reconstructions. The results show that the digital Gabor holographic system outperforms traditional conventional microscopic methods and can accurately calculate the object's position in the *z*-direction. Other advantages of digital Gabor holography are the simplicity of sample preparation, particularly for biological samples where neither sectioning nor staining is required, which means that living cells which are largely transparent can also be viewed.

Outside of biological microcopy, there are many other applications for 4-D Gabor holographic movies, including particle tracking.

Chapter 5

Quantitative Digital Holographic Microscopy

This chapter presents the results of digital holography experiments conducted in an off-axis configuration. Section 5.1 begins the chapter with an introduction to phase-contrast microscopy and reviews some of the commonly used techniques in this field. Section 5.2 examines how phase information from an object is produced and how one may employ this information in order to obtain quantitative evaluation. Section 5.3 compares the use of the different reconstruction methods to a standard resolution test target. Section 5.4 studies the differences between phase and amplitude information in biological objects. Section 5.5 presents examples of quantitative phase images obtained from digital holography experiments that examine the resolution of the optical system and demonstrate the quantitative and qualitative capabilities of digital holography for biological microscopy. Section 5.6 discusses the use of quantitative digital holographic movies in microscopy and demonstrates the advantages over conventional techniques. Finally Section 5.7 ends this chapter with a conclusion and discussion.

5.1 Introduction

Many microscopic biological specimens, such as living cells and their intracellular constituents, are mostly transparent, and therefore are problematic for conventional bright-field microscopy. There have been developed a number of techniques

for rendering transparent phase objects visible that have played very important roles in the development of modern biology and medicine[50], and these include dark field, Zernike phase-contrast, and Nomarski differential interference contrast (DIC) microscopies.

In dark field microscopy, only the scattering centers and boundaries contribute to the image signal against a zero background. In the phase contrast microscope, the phase variation is converted into amplitude variation, and in DIC, the interference of two sheared polarization components result in images that have shadow effect and thus give a three-dimensional perception of the object. In any of these techniques, the phase to amplitude conversion is nonlinear and there are significant artifacts in the images, such as the halo in phase contrast and the disappearance of contrast along the direction perpendicular to shear in DIC. Quantitative phase imaging is not feasible with these techniques. Quantitative phase imaging is important because it allows the determination of the optical thickness profile of a transparent object with sub-wavelength accuracy. The optical thickness profile depends on the physical thickness as well as the optical index variation, and thus one can extract this information with great accuracy [51].

White-light interference microscopy [22] and optical coherence microscopy [52] have been used to generate quantitative phase images but these require multiple exposures or mechanical scanning.

Digital holography offers an excellent approach for quantitative phase imaging. A hologram that consists of the interference between the object and the reference beams is recorded by a CCD camera and the holographic image is numerically reconstructed inside a computer using the results of diffraction theory. Calculation of the complex optical field

allows direct access to both the amplitude and the phase information of the optical field, and by numerical focusing the images can be obtained at any distance from a single recorded hologram.

5.2 Quantitative Evaluation by Digital Holography

Amplitude objects change only the amplitude component of the light, leaving the phase unchanged. On the other hand, phase objects change the phase of the light passing through them, leaving the amplitude unaltered as illustrated in Figure 5.1. As each wave-front passes through the specimen, it is deformed and retarded (or advanced) according to the geometry, refractive index differential, and the thickness of the specimen. An incident wave-front present divides into two different components upon passing through a phase specimen. The first component is the un-diffracted, zero-order wave-front that passes through and around the specimen, but does not interact with it. In addition, a second component is the deviated or diffracted wave-front, which becomes scattered in many directions and increases with specimen size. This scattering is essentially caused by discontinuities in the specimen's refractive index. A non-absorbing substance with a homogeneous refractive index is transparent.

Biological tissue, however, often contains many refractive index discontinuities. A refractive index change in a transparent medium leads to a change of the optical path length, which is proportional to the product of the refractive index and the optical thickness. This optical path length change can be formulated mathematically in a phase change in the wave-front.

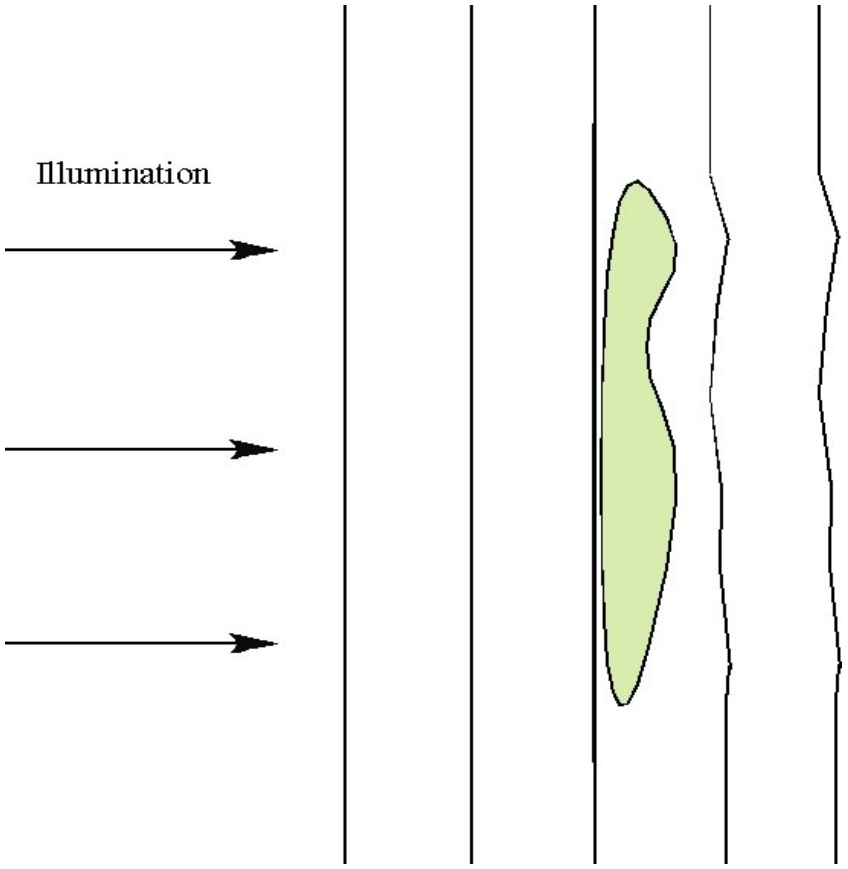


Figure 5.1: A phase shift is produced when a wave-field impinges on an object with a refractive index different to the surrounding medium. The phase shift is a product of the variation of the relative index of refraction and the optical thickness of the object.

This corresponding change in the interference phase is

$$\Delta\phi(x,y) = \frac{2\pi}{\lambda}\int[n(x,y,z) - n_0]dz \tag{5.1}$$

where n_0 is the refractive index of the medium under observation in its initial, unperturbed state and *n(x, y, z)* is the final refractive index distribution.

From Equation (5.1) the physical thickness of the object can be expressed as

$$d_{obj} = \frac{\lambda\Delta\phi_{obj}}{2\pi}.\frac{1}{n_{obj} - n_0} \tag{5.2}$$

where d_{obj} represents the physical thickness of the object, $\Delta\phi_{obj}$ is the optical path length change of the object with refractive index n_{obj} and n_0 is the refractive index of the surrounding medium. Once the phase information $\Delta\phi_{obj}$ is obtained from the reconstruction of the hologram, knowledge of the physical thickness of the sample allows calculation of the refractive index or vice versa. This method can be applied for the shape measurement of transparent microscopic samples, such as living cells.

5.3 Experimental Setup for Off-Axis Digital Holography

The digital holography experiments are performed using the apparatus schematically depicted in Figure 5.2. In Figure 5.3, we show a birds-eye view of the setup by a photograph. A miniature pulsed Nd:YAG laser (Continuum Minilite) operates at 532nm, with an energy of 2 mJ per 10ns Q-switched pulse. The pulse repetition rate is 15Hz. The laser output, which is spatially-filtered and collimated, is split into a reference and object beam in a transmissive interferometer setup based on the Mach-

Zehnder configuration. The benefit of this type of interferometer is that the path length in both arms is equal due to the symmetry of the setup. Therefore it can be used to measure the amount of phase modulation caused by an object placed in one arm of the interferometer.

The object specimen, mounted on a xyz-translation stage, is placed at a distance *z* from the hologram plane *H*, whose magnified image is projected on the CCD camera, along with the reference beam. A pair of similar microscope objectives, either 20X, 0.4NA, or 40X, 0.65NA depending on the desired lateral magnification, is used in the two optical branches to match the curvatures of the two wave-fronts. A slight angle is introduced between the object and the reference beams by tilting the beam splitter BS2 for off-axis holography. The camera (Sony DFW-V500) has an array of 640 x 480 pixels on a 4.7 x 3.6 mm^2 active area, with 8-bit gray scale output. A digital delay generator (Stanford Research DG535) triggers both the laser and the camera at a repetition rate of 15 Hz.

The microscope objectives L1 and L2 and the beam splitter BS2 are adjusted in the experimental setup so that the interference fringes are straight and so avoiding the need to perform any digital correction due to spherical aberration introduced by the microscope objectives. Another parameter that characterizes the interference fringes is the spacing between them.

An IEEE1394 cable connects the camera to the desktop computer, which processes the acquired images and calculates the holographic diffraction using a number of programs based on LabVIEW® and MatLab®.

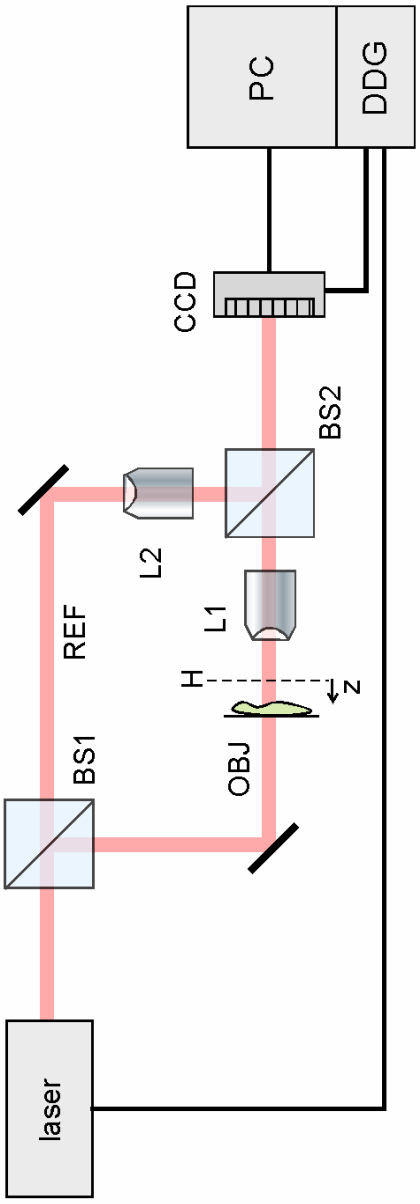


Figure 5.2: Digital holographic setup for Off-Axis holography.

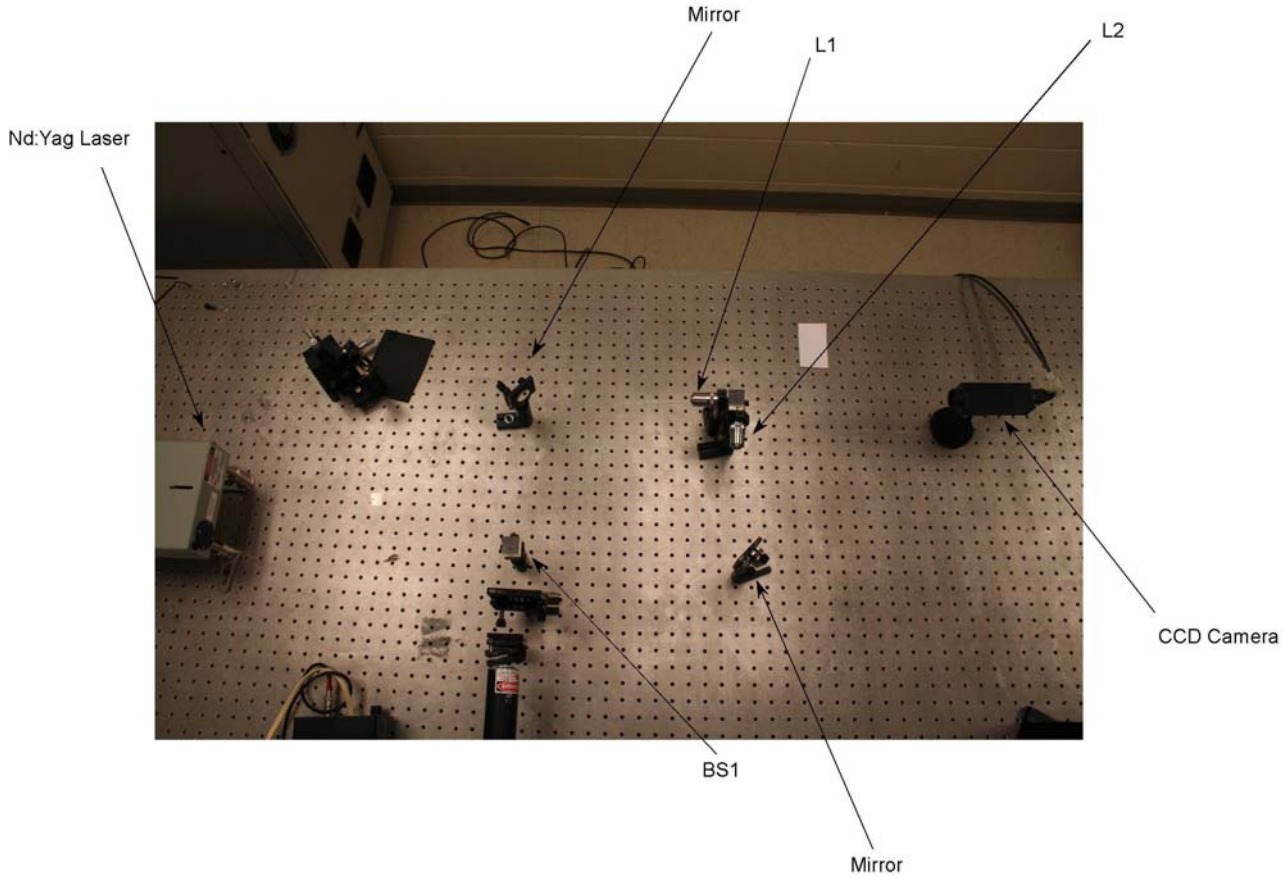


Figure 5.3: Birds-eye view of the digital holographic setup for Off-Axis holography.

The Labview programs allow the user to view the amplitude, the phase, and the real and imaginary parts of the reproduced object. The program can also change the angle of the plane at which the object is recreated. This is equivalent to viewing the object from a different angle. This is a very useful feature for flat objects because the off-axis hologram reproduces the image as having a horizontal tilt equal to the angle of the reference mirror. This would generate phase maps that exceed the repeat distance and cause wrap around. The angle is selected so that the phase maps have the least number of fringes. This can also compensate for any tilt the object may have originally had.

5.3.1 Optical alignment

An important factor in the experimental methodology is to quickly observe whether the hologram produced at the CCD camera is of a sufficiently good quality. It is very difficult to know whether hologram will be able to provide a high quality reconstruction of the amplitude and phase image just by simply looking at the holographic image. Furthermore the high sensitivity of the hologram to minute changes in alignment in the experiment also makes a substantial difference to the quality of the image reconstruction.

An efficient methodology that allows for efficient and accurate alignment of the holographic optical system is to look at the angular spectrum (Fourier transform) of the hologram recorded by the CCD camera in real time. Any changes that are made in the experiment are registered in the angular spectrum in real time. This allows one to make subtle changes in the experiment to obtain the best possible spectrum that will be able to

produce a good reconstruction. Factors to consider are sufficient angular separation of the *dc* and holographic terms and any spurious noise.

Particular attention must be paid to the adjustment of the angle between object and reference beams, which must not exceed a maximum value.

5.4 Phase Unwrapping

Once the hologram has been reconstructed, both the amplitude and the phase information of the object are available. The phase information, however, contains 2π discontinuities wherever the extreme value of $\Delta\phi$ corresponding to either $-\pi$ or π are reached. In general, the true phase may range over an interval greater than 2π, in which case the phase contains artificial discontinuities. To improve interpretation of the reconstructed phase image the process of *phase unwrapping* ensures that all appropriate multiples of 2π phase errors are removed from the image. Unwrapping these discontinuities is a matter of adding an appropriate integer multiple of 2π to each pixel element of the wrapped phase map. In practice, however, the presence of noise and residues complicates effective phase unwrapping.

Figure 5.4 displays the problem of the 2π discontinuity. The theoretical object is a tilted ramp of arbitrary height. Wherever a phase variation of 2π exists, a phase discontinuity occurs. The true unwrapped phase image can be expressed as

$$\phi_{uwap}(r) = \phi_{wrap}(r) + 2\pi n(r) \tag{5.3}$$

where $\phi_{uwap}(r)$ is the unwrapped phase, $\phi_{wrap}(r)$ is the wrapped phase in the phase range of $-\pi$ to π, r represents the pixel position and *n(r)*is an integer factor to be determined.

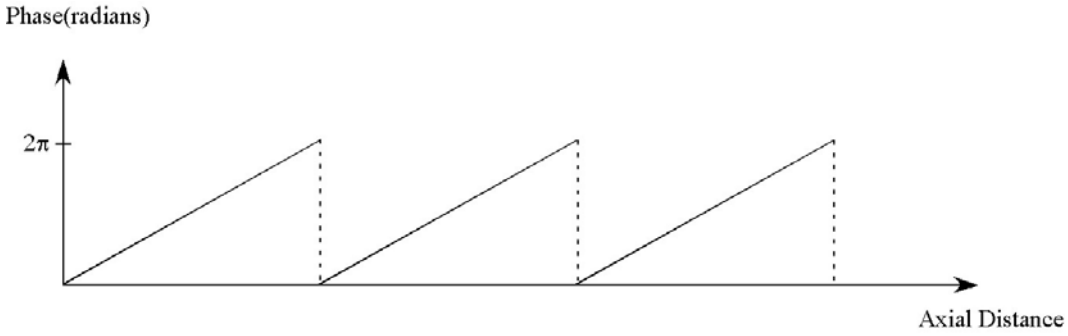


Figure 5.4: The phase variation of a tilted ramp displaying discontinuities wherever the object height is greater than the phase variation of 2π .

Figure 5.5 presents the unwrapped phase data for the tilted object in Figure 5.4 with the implementation of Equation (5.3). The discontinuities are removed by the addition of integer values of 2π .

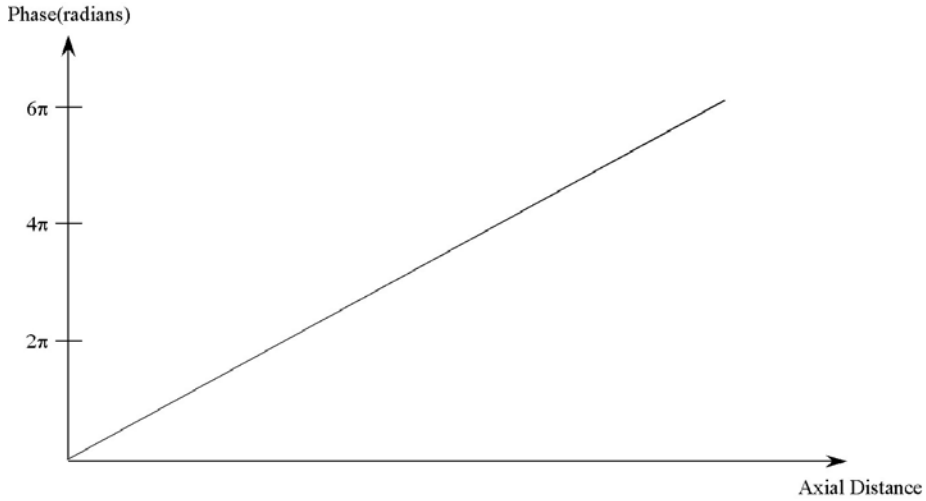


Figure 5.5: Removal of 2π discontinuities for the tilted object in Figure 5.3 by addition of integer multiples of 2π .

A number of phase unwrapping algorithms have been developed, each with varying complexity and requirements [53,54]. The simplest algorithm checks adjacent pixels in the wrapped phase and, when it exceeds some threshold value, takes this difference as the phase jump and adds $2\pi\pm$. However depending on the initial start position and thereby the unwrapping path chosen, one may get different unwrapped phase fields, depending on the unwrapping path chosen. As phase unwrapping is an integrated process, an error on the previous point can propagate along the unwrapping path. Therefore, determining an optimized path is a critical step toward completely reliable phase unwrapping. Other issues in phase unwrapping include the presence of speckle noise and also complex surfaces which may cause an acute change in the fringe distribution, or even fringe discontinuity. Therefore the process is not as simplistic as may first seem.

In this thesis the phase images are all processed by the use of a flood-fill algorithm to perform computational phase unwrapping as outlined in Appendix C. Like many other phase unwrapping algorithms, it starts from a defined position and calculates the difference of neighboring pixels in the wrapped phase and, when it exceeds some threshold value, takes this difference as the phase jump. The algorithm also makes use of edge detection before unwrapping to determine features in the image which may cause errors.

The main drawback of this algorithm is the rather long time it requires for processing. The average time to unwrap an image is approximately three minutes and this creates a serious limitation for obtaining real time processing of dynamic effects.

5.5 Comparison of numerical reconstruction methods

Figure 5.6 shows the application of the Fresnel, convolution and angular spectrum methods to the reconstruction of the image of a resolution target. A 25 x 25μm^2 area of the USAF resolution target is imaged containing the group 7 element 6 using 452 x 452 pixels. Figure 5.6(a) shows the holographic interference pattern recorded by the CCD camera, and its Fourier transform in Fig. 5.6(b) is the angular spectrum. It contains three main peaks that correspond to the spectra of the zero-order and the two twin-images. One of these peaks – the highlighted rectangular area – is selected, a propagation phase factor (z = 7μm) is multiplied, and inverse-Fourier transformed to obtain the amplitude image in Fig. 5.6(c) and the phase image in Fig. 5.6(d).

Any further filtering may be carried out in the selected rectangular area highlighted in Fig. 5.6(b) to remove spurious noise components. The noise contributions are clearly observable in the hologram spectrum, because they often arrive on the CCD with incidence angles that are distinct from the incidence angle of the object wave. In addition noise associated with the camera can also be removed.

The individual bars displayed in the figure are 2.2μm wide, and are clearly resolved and consistent with the diffraction-limited resolution of 0.5μm for a 0.65 NA lens. The resolution target is a positive mask with opaque chrome film pattern on transparent glass plate, used in transmission. From the analysis of the phase map, the phase step of the metal strip is ~ 52° . The physical thickness of film is given by Equation (5.2) where λ is the wavelength, $\Delta\phi_{obj}$ is the phase step, and $n_{obj} - n_0$ is the index difference between the film and air.

Using a known estimate of the film thickness of 50nm, the optical index (real part) of the metal film is ~2.5. The noise level of the flat area is ~ 10° , which corresponds to glass thickness variation of ~30 nm. (On the film-coated bar areas, the lack of light causes larger uncertainty in phase.) The phase map is rendered in pseudo-colored 3D perspective in Fig. 5.6(i). Especially notable in the phase map is the lack of the coherent noise conspicuous in the amplitude image and prevalent in most other holographic imaging methods. The amplitude and phase images obtained from the Huygens convolution method are shown in Fig. 5.6(e) and 5.6(f), as well as those obtained from the Fresnel method in Fig. 5.6(g) and 5.6(h). The main reason for the obvious degradation of these images is the insufficient off-axis angle to separate out the zero-order component. The effect is most detrimental in the Fresnel images, where part of the holographic image is buried in the zero order background, and its phase image is completely scrambled. The effect shows up as the spurious interference patterns in the convolution images. While the minimum off-axis angle must be strictly satisfied in order to avoid the zero-order intrusion in Fresnel or convolution methods, the control and removal of zero-order component is straightforward and flexible in the angular spectrum method. Another potential problem is that the Fresnel and convolution methods require minimum hologram distance to avoid aliasing, whereas the angular spectrum method does not have such minimum and the image can be calculated even at zero distance[15]. For the particular example shown here the minimum distance happens to be 2.6μm and therefore is not an issue, but the images shown below in Fig. 5.6 are obtained at

$$z = 5\mu m < z_{min} = 18\mu m,$$

the Fresnel or convolution methods would not have worked.

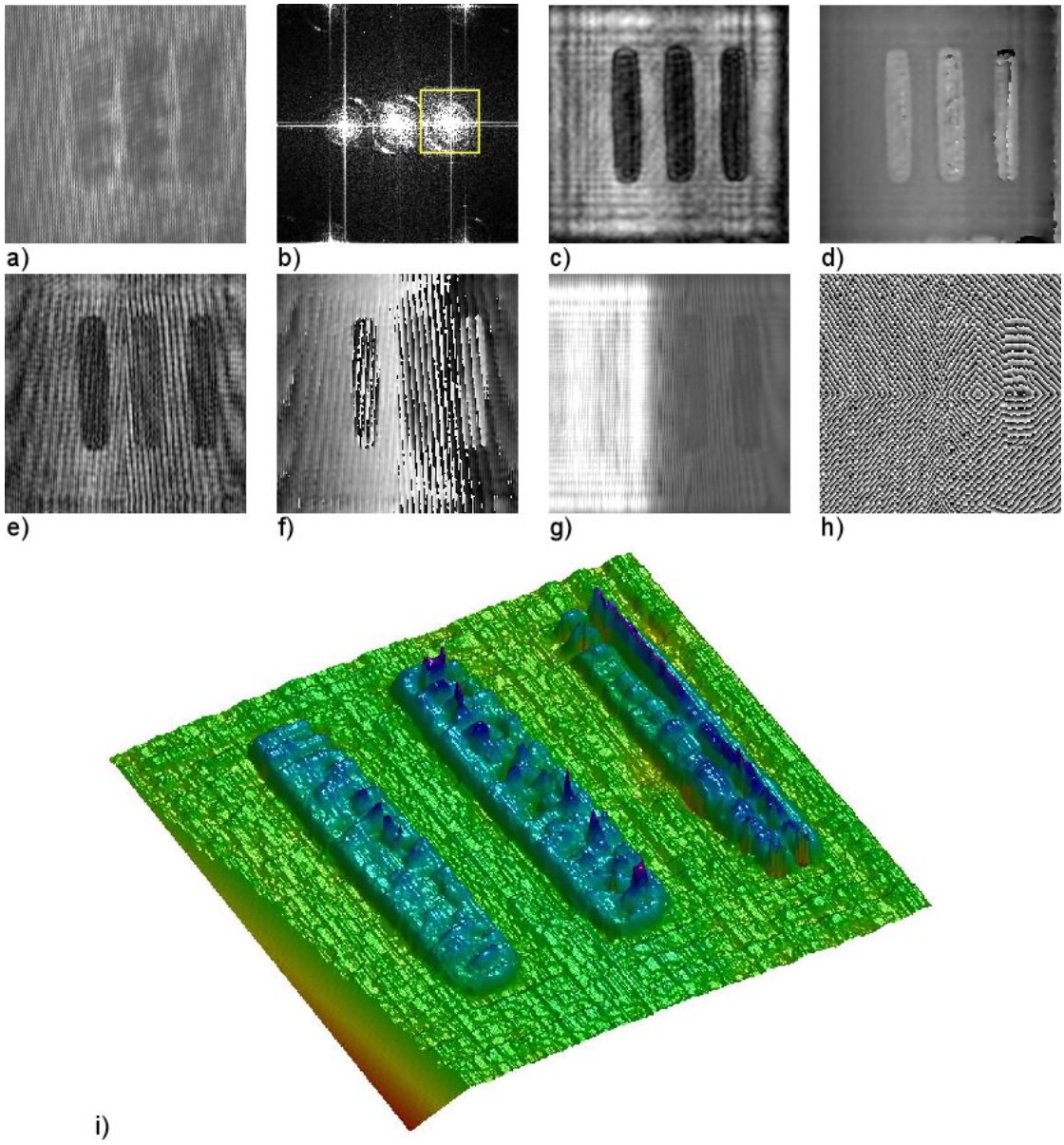


Figure 5.6: Holography of a resolution target. The image area is 25 x 25 μm^2 (452 x 452 pixels) and the image is at z = 7μm from the hologram: (a) hologram; (b) angular spectrum; (c) amplitude and (d) phase images by angular spectrum method; (e) amplitude and (f) phase images by Huygens convolution method; (g) amplitude and (h) phase images by Fresnel transform method; (i) 3D pseudo-color rendering of (d). The individual bars are 2.2μm wide.

The optical field can be calculated at any number of image planes from a single hologram. Figure 5.7 illustrates the use of the angular spectrum method in numerical focusing of a group 6 element of the resolution target. The panel shows a sequence of eight images calculated in the range of z = 1–15μm in steps of 2μm. Each image is a 30×30 μm^2 area of a resolution target. As the focus is scanned, one observes the bars move into focus as it passes through the various image planes. Note that the object comes into focus at around 10μm. For Huygens and Fresnel methods this distance would have been smaller than z_{min}.

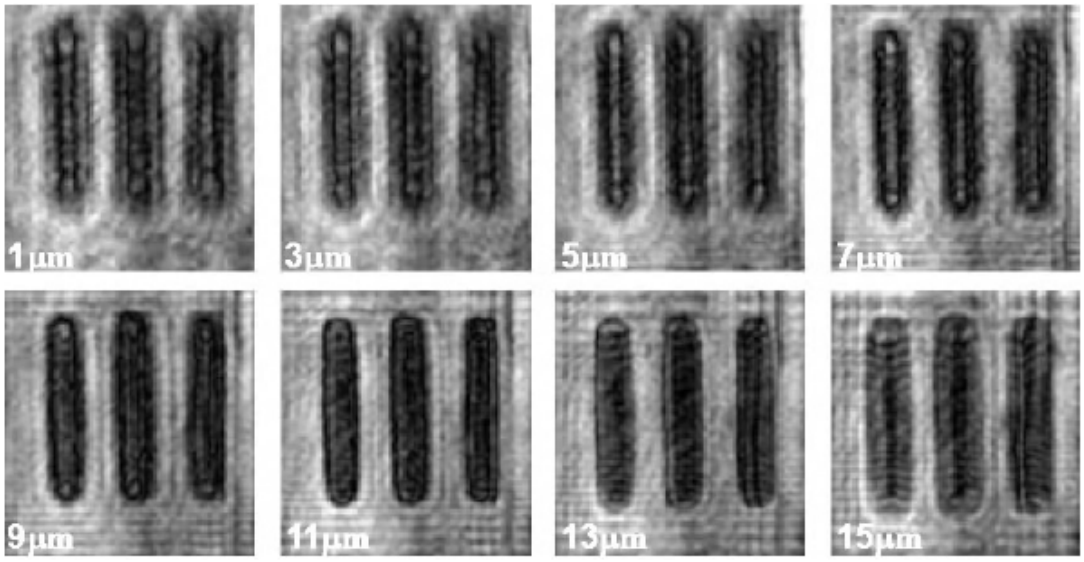


Figure 5.7: Numerical focusing in digital holography of an element of a USAF 1951 resolution target from a single hologram. Images are of a 30 × 30μm^2 area (360 × 360 pixels) with z scanned from 1 to 15μm in steps of 2μm.

5.6 Phase Imaging Digital Holography

Coherent noise is a major problem for imaging with lasers and often leads to degradation of the image quality. In digital holography, the phase images tend to suffer from the coherent noise to a significantly lesser degree compared to the amplitude images. This is because the amplitude image reflects the intensity variations in the reference wave, whereas the phase noise mostly comes from the quality of the optical surfaces in the imaging system. Furthermore while amplitude measurements are equivalent in resolution and quality to classical optical microscopy, phase measurements can lead to images with an optical thickness resolution of a few nanometers.

Figure 5.8 displays digital holography of a paramecium, which is a partially transparent. unicellular protozoan with considerable intracellular complexity. The image area is 250 x 250μm^2, recorded at 464 x 464 pixels and the images are reconstructed at z = 770μm from the hologram. The most conspicuous feature in the amplitude image in Fig.5.8 (b) is one of the two contractile water vacuoles, which is shown as a bright spot. The other water vacuole and the macronucleus are not readily visible in the amplitude image, while there is a hint of them in the phase images. Slightly different directions of reconstruction reference waves are used in the phase images figures 5.8(c) and 5.8(d), which may be useful for emphasizing different aspects of the microscopic images. One may notice from these images that the paramecium is not highly transparent as demonstrated by the amplitude image and there exists a large number of features with varying indexes of refraction that produce speckle and somewhat degrade the image quality.

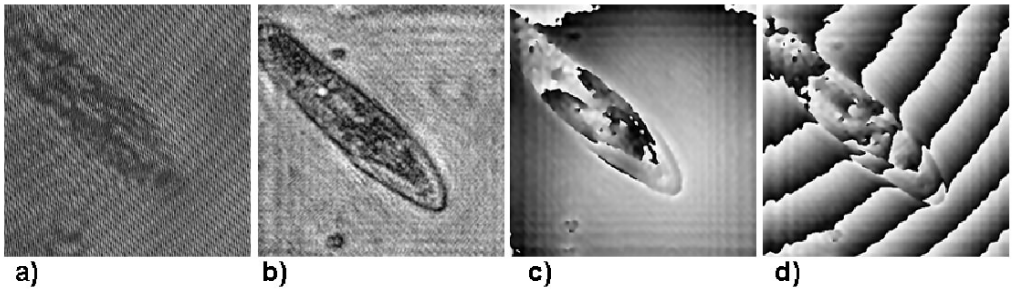


Figure 5.8: Holography of a paramecium, showing a) hologram, b) amplitude, c) and d) phase images of a 250 x 250 μm^2 (464 x 464 pixels) area with z = 770μm**.**

Figure 5.9 presents another example of digital holography of a paramecium. Note the phase variations that occur on the edge of and within the paramecium itself in the phase image in Fig. 5.9(c). Compared to the amplitude image in Fig.5.9 (b), the phase image reveals more information of the intracellular features.

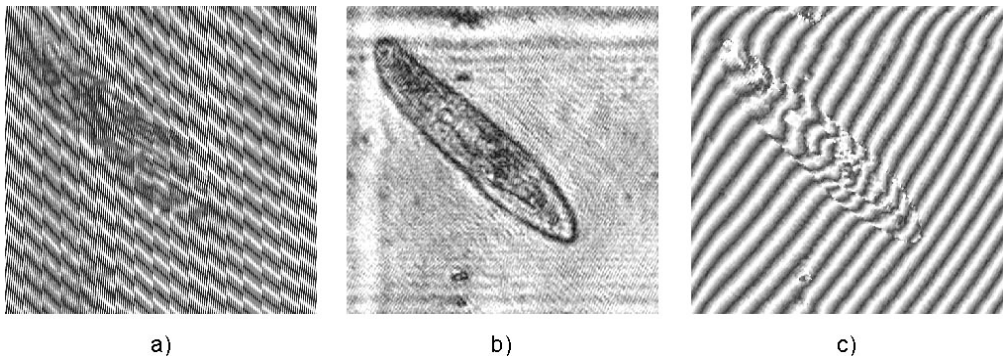


Figure 5.9: Holography of a paramecium, showing a) hologram, b) amplitude, and c) phase image (Area =250μm x 250μm^2, Pixels = 464 x 464, z = 0.561mm).

Examples of digital holography of onion cells are shown in Figure 5.10. In the upper row, a 100 x 100 μm^2 area of a layer of onion cells is shown. The cell walls are sharply focused in the amplitude image and the phase image shows an accurate representation of the optical thickness, modulo wavelength, of the cell bodies. The

images of 70 x 70 μm^2 area in the lower row focus on the nucleus of a cell. The phase image is a clear view of the optical thickness variation of the nucleus in the middle of the bulged body of the cell. A simple quantitative analysis of the cell's index of refraction is possible. By counting the number of fringes, the optical thickness of the nucleus is easily determined to be $3.5\lambda = 1.86\mu m$ thicker than the cell body. If we assume that the shape of the nucleus is spherical so that its thickness is the same as the 19μm diameter of the circular image, then the refractive index difference between the nucleus and the cell body is 0.093.

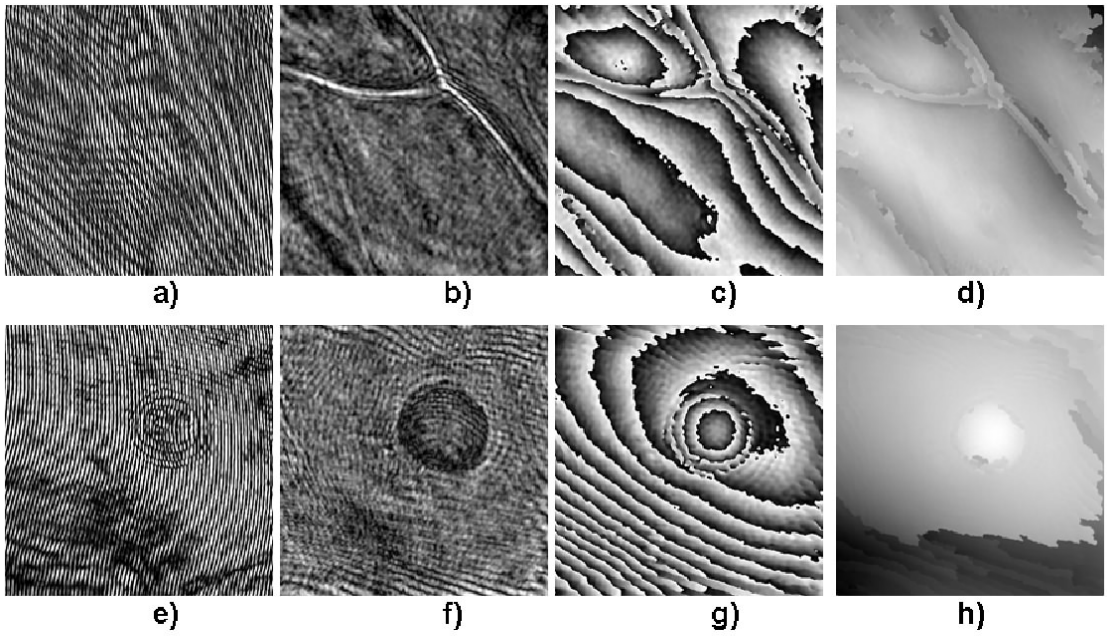


Figure 5.10: Holography of a layer of onion cells. Images of a) hologram, b) amplitude, c) wrapped phase, and d) unwrapped phase in the upper row are a 100 x 100μm^2 (416 x 416 pixels) area with z = 174μm, while the images e), f), g), and h) in the lower row are a 70 x 70μm^2 (464 x 464 pixels) area with z = 6μm.

5.7 Quantitative Phase-Contrast microscopy by Digital holography

Quantitative phase imaging is particularly effective in digital holography – one only needs to plot the phase of the calculated complex optical field. Figure 5.11 displays digital holography of the smallest elements of group 7 on the resolution target. The panel displays the (a) hologram, (b) amplitude image, (c) phase image, (d) phase image unwrapped by software algorithm, (e) z-profile of the optical thickness cross section and (f) is the pseudo-color 3D rendering of the phase image unwrapped by software algorithm.

The hologram is recorded at a distance z = 3μm from the focal plane of the CCD camera and thus is not recorded in focus. The reconstruction is performed by application of the angular spectrum algorithm which has no set minimum reconstruction distance and the determination of best focus of the image is done by observation. The smallest resolvable bars are that of group 7, element 6 corresponding to a line width of 2.2μm. The optical thickness resolution was determined quantitatively by analysis of the noise level in the flat area in the unwrapped phase image, which corresponds to glass thickness variation and is calculated to be around 8nm by the use of Equation 5.2 (On the film-coated bar areas, the lack of light causes a larger uncertainty in phase).

The low noise characteristics of phase imaging are further demonstrated in Figure 5.12, which shows digital holography of group 6, element 2 of the resolution target. The panel displays the (a) unwrapped phase image, (b) the z-profile of the optical thickness cross section, and (c) the pseudo-color 3D rendering of the unwrapped phase image. The quantified noise in the flat areas of the unwrapped phase image is determined to be around 3nm as displayed in the R.M.S cross section profile.

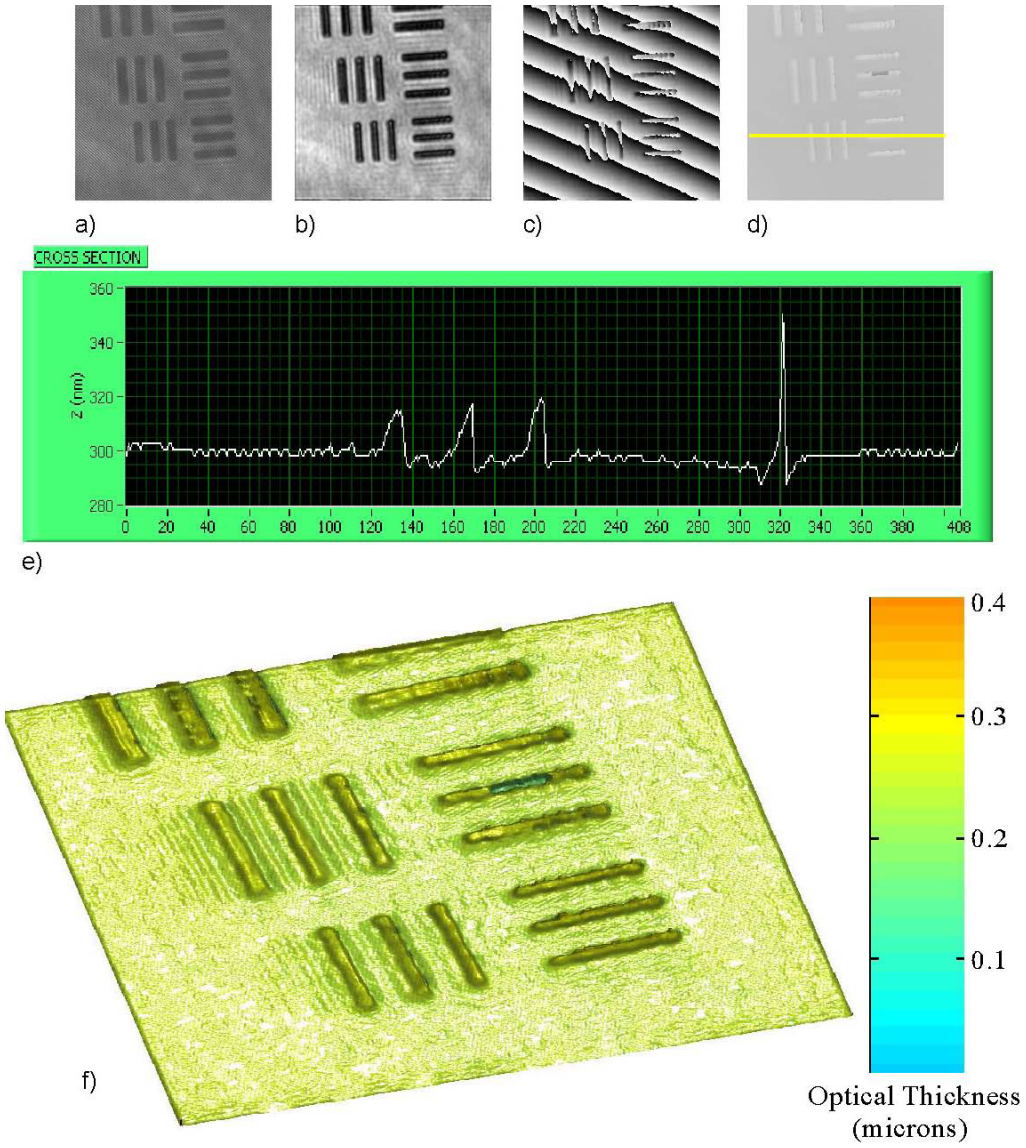


Figure 5.11: Digital holographic images of group 6 of a resolution target with area 50 x 50 μm^2. The image is at z = 3μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image, (e) z-profile of cross section as displayed in (d) and (f) 3D pseudocolor rendering of (d).

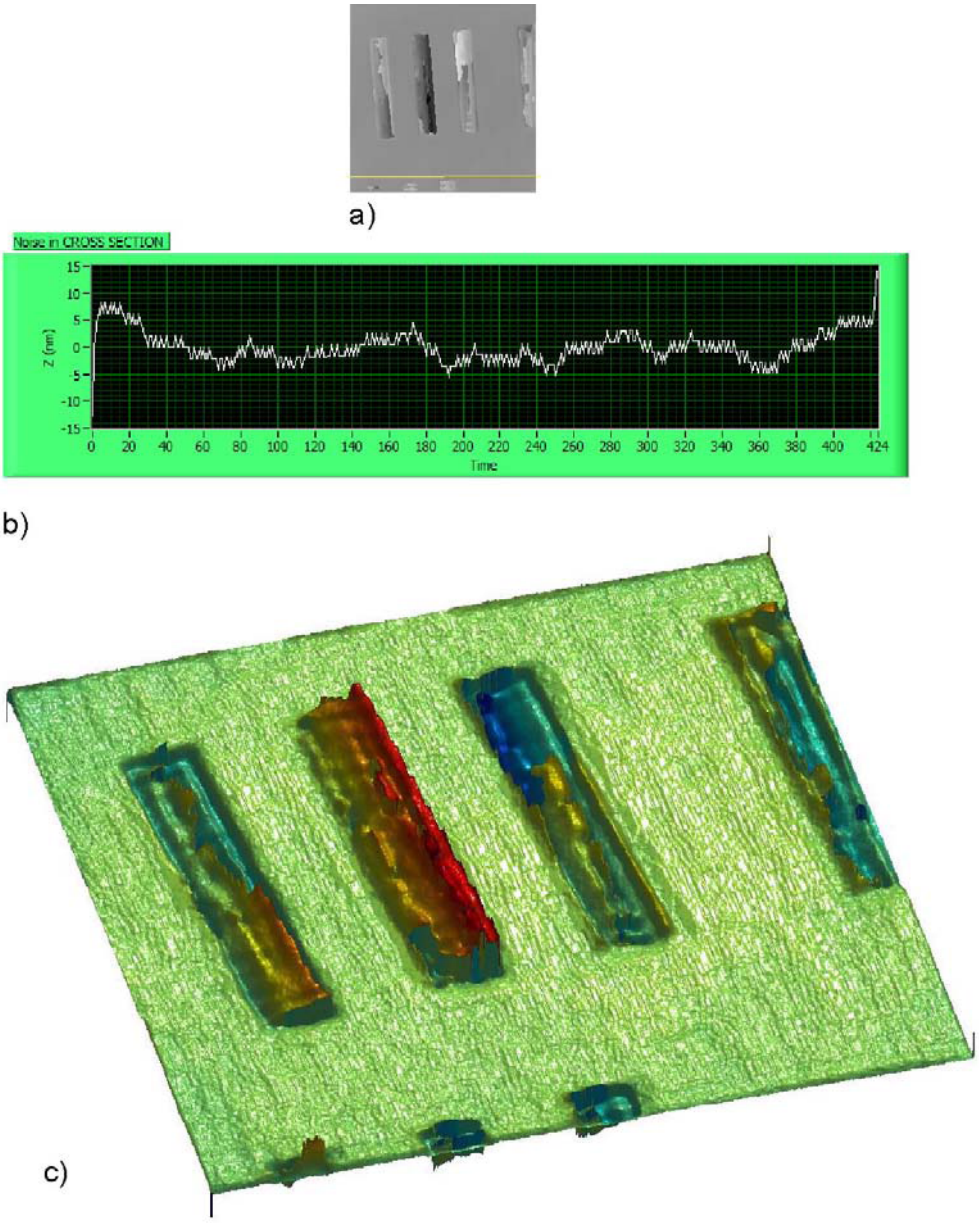


Figure 5.12: Digital holography of group 6 element 2 of the resolution target. The images are the (a) unwrapped phase image, (b) z-profile of cross section as displayed in (a) and (c) the 3D pseudocolor rendering of (a).

Figure 5.13 shows another example of digital holography of a group 6 element of the resolution target. The panel displays the (a) hologram, (b) amplitude image, (c) phase image, (d) phase image unwrapped by software algorithm, and the pseudo-color 3D rendering of (d) is shown in (e). In Fig. 5.13(e), one may notice how there is a large amount of noise produced in the bars due to a lack of signal from an opaque object in a transmission setup. However, this noise is reduced considerably at the edges of the bars suggesting that there may indeed be some signal from these areas. Further evidence of this phenomenon is also observed when looking at the amplitude image reconstruction, where there is a small hole in one of the bars. In the phase image the immediate area around this hole appears to be fairly noiseless with respect to that of the inner parts of the bars. The area around the bars is again demonstrated to contain low noise as can be seen in the 3D profile or unwrapped phase image.

Figure 5.14 displays higher magnification digital holographic images of groups 8 and 9 of the resolution target, where the smallest resolvable bars are that of group 9, element 2, with line width of 0.87μm. This value is consistent with that of the Abbe criterion

$$\Delta\xi = 0.61\lambda / N.A \sim 0.81\mu m \tag{5.4}$$

for diffraction limited imaging of the optical system.

Thus the lateral resolution is only restricted by the diffraction limit. In general, one may note that the lateral resolution of the system by Equation (5.4) can be improved either by use of a higher-N.A. objective or by an increase in the size of the beam incident onto the objective lens.

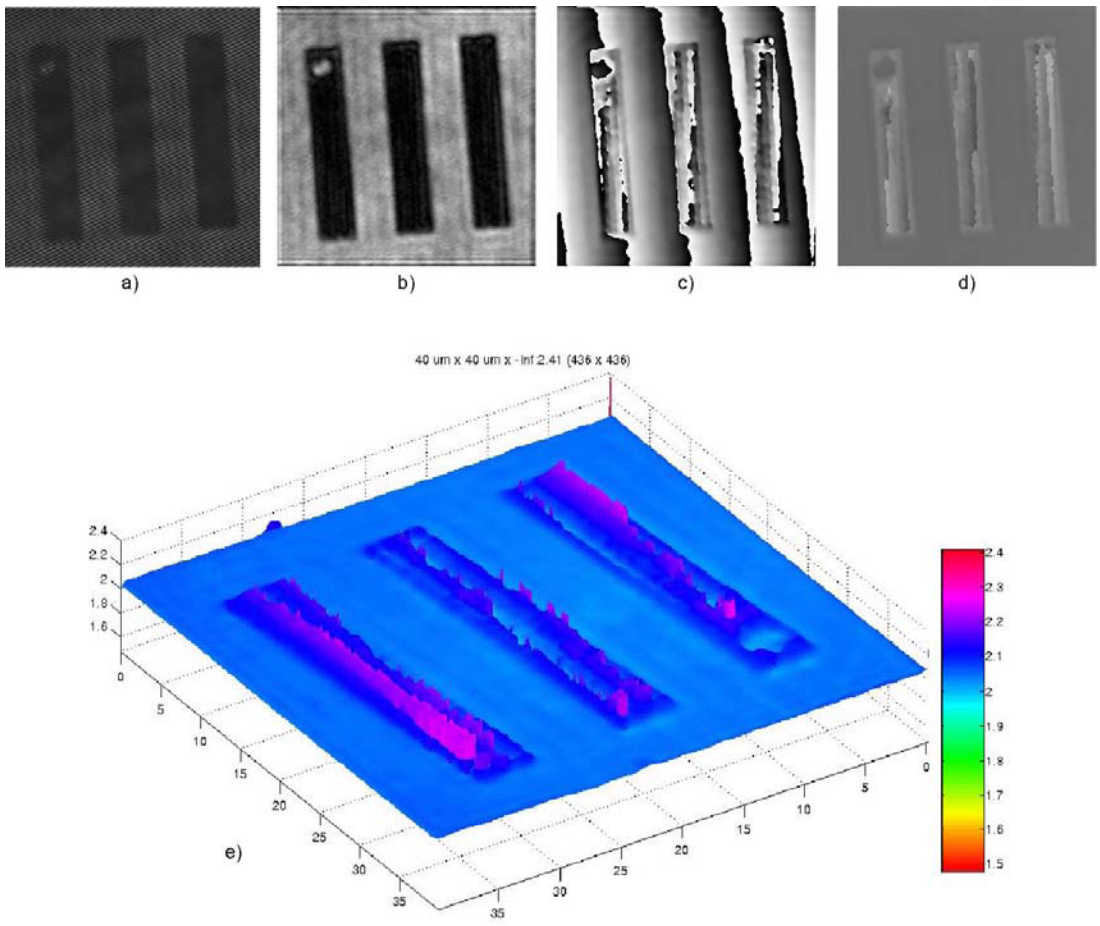


Figure 5.13: Digital holographic images of group 6 element of the resolution target with area 40 x 40 μm^2. The image is at z = 3μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image, and (e) 3D pseudo-color rendering of (d).

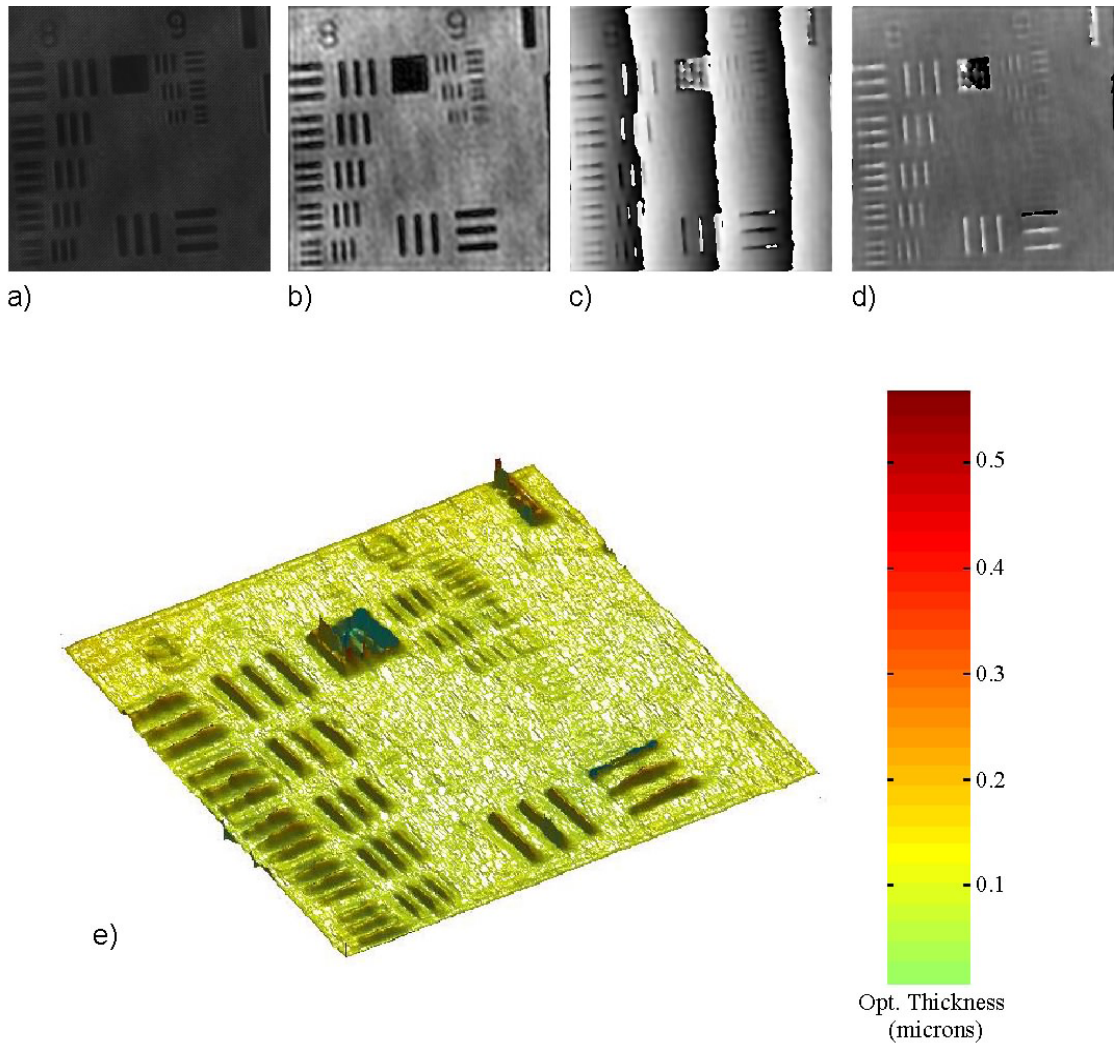


Figure 5.14: Digital holographic images of group 8 and 9 of a resolution target with area 25 x 25 μm^2 (424 x 424pixels). The image is at z = 2μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image and (e) 3D pseudo-color rendering of (d). The smallest bar is that of group 9, element 2 and is 0.78μm wide.

A problem with the form of the experimental setup as depicted in Figure 5.2 is that the sample orientation is required to be mounted vertically instead of horizontal, as is the case in most conventional microscopic systems. This is fine for most samples that require no specific preparation and can be carried out with some sticky tape and a lens holder. However, it causes an inconvenience when performing imaging analysis of living biological cells. Furthermore, another problem is that cells require both the correct amount of CO_2 and a temperature of around $37.2^{\circ}C$ in order to be sustained over long periods of time. The usual methodology in conventional microscopy is to place the living cells horizontally in a chamber maintained at the correct temperature and CO_2 gas concentration. With these factors in mind, adaptations were made to both the sample preparation and the mounting in order to satisfy the requirements for the experiments.

The sample preparation for live cell imaging is shown schematically in Figure 5.15. The cells were grown directly on a cover slip in a culture solution. The cover slip was then put inside a rectangular shaped quartz cell (STARNA Rectangular cell with path length 1mm) as illustrated in Figure 5.15 and DPBS, 1X w/Ca and Mg solution was simultaneously added into the quartz cell along with the cover slip containing the cells in order to keep the cells alive for periods of up to approximately 4-6 hours. The object beam entered the sample from the slide glass side. As the cell was attached to the cover slip, the focus was adjusted to the inner side of the cover slip. Although we were able to keep the cells alive for sustained periods of time using the DPBS, the cells were not in optimum conditions. This, we believe, severely reduced both the mobility and also the quality of the cells. Long periods of cell analysis require a special preparation of the biological sample. This preparation mainly refers to the fixation of the biological sample.

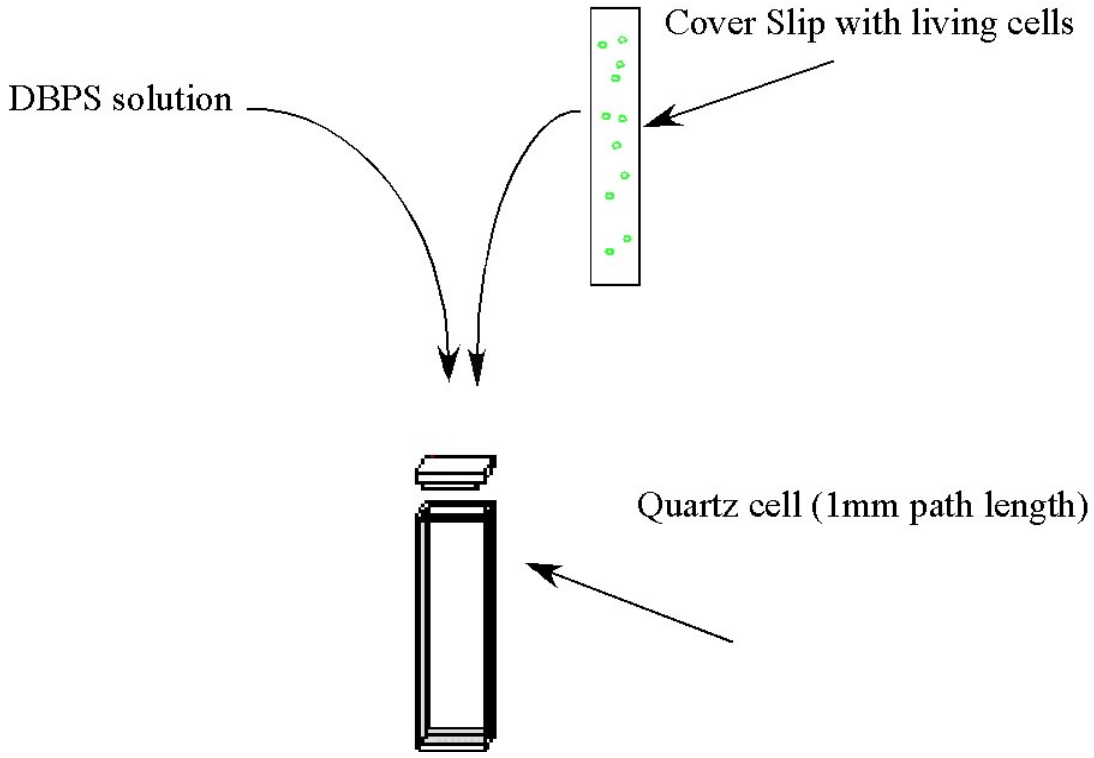
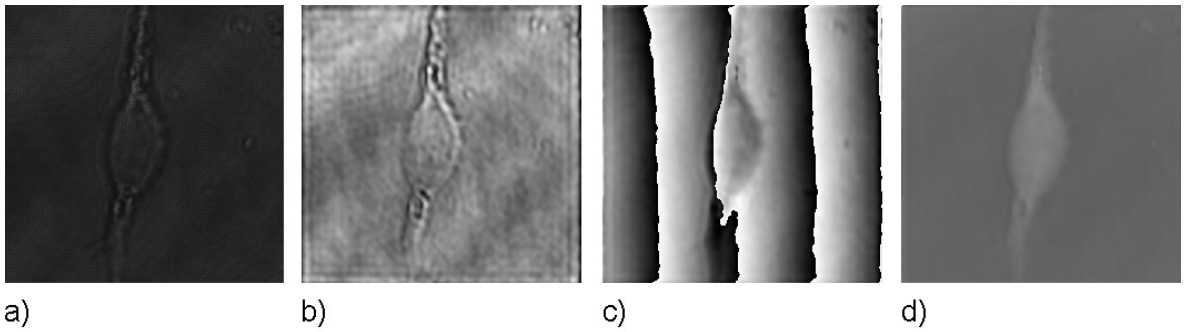


Figure 5.15: Schematic of live sample preparation.

The fixant is applied to kill the cell and to keep its form at the certain stage of development or change by external influence. This is particularly useful for studies of cells in an essentially 'frozen' state.

Figure 5.16 show images of a living mouse embryo fibroblast cell. Figure 5.16(e) which is the pseudo-color 3D rendering of Fig. 5.16(d) shows the high quality and contrast of surface detail of the cell obtained with phase imaging and provides an accurate and qualitative profile of optical thickness. To produce the optical thickness profile by Equation (5.2), a refractive index estimate of 1.375 was used for the cell [55].



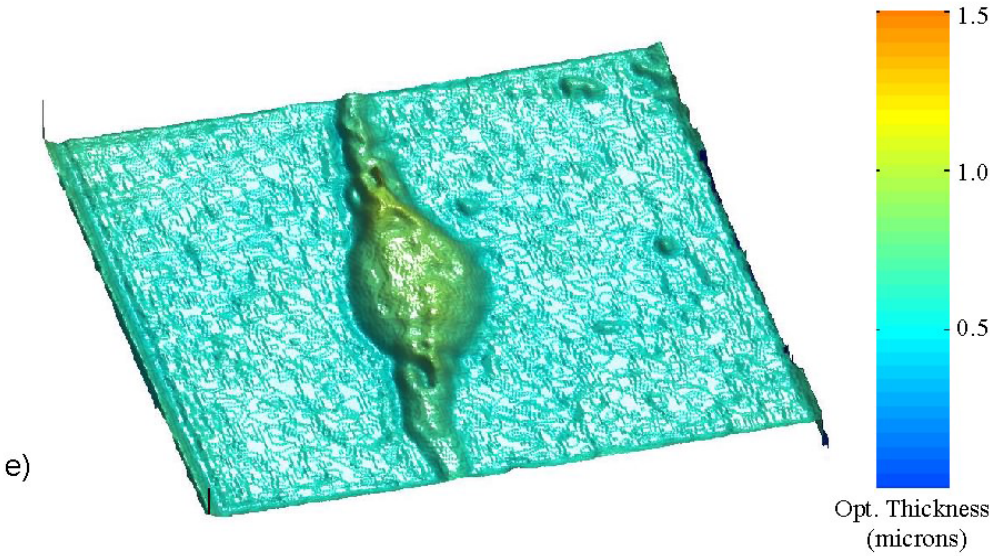


Figure 5.16: Holography of mouse embryo fibroblast cell. The image area is 60 x 60 μm^2 (424 x 424 pixels) and the image is at z = 14μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image and (e) 3D pseudocolor rendering of (d).

Figure 5.17 shows digital holography of an onion nucleus. The panels display the (a) hologram, (b) amplitude image, (c) phase image, and (d) phase image unwrapped by a software algorithm. Pseudo-colour 3D rendering of (d) is shown in (e).The image size is $30 \times 30\mu m^2$ with 436×436 pixels and is reconstructed at $z = 22\mu m$ from the hologram. The phase image is a clear view of the optical thickness variation of the nucleus in the middle of the body of the cell.

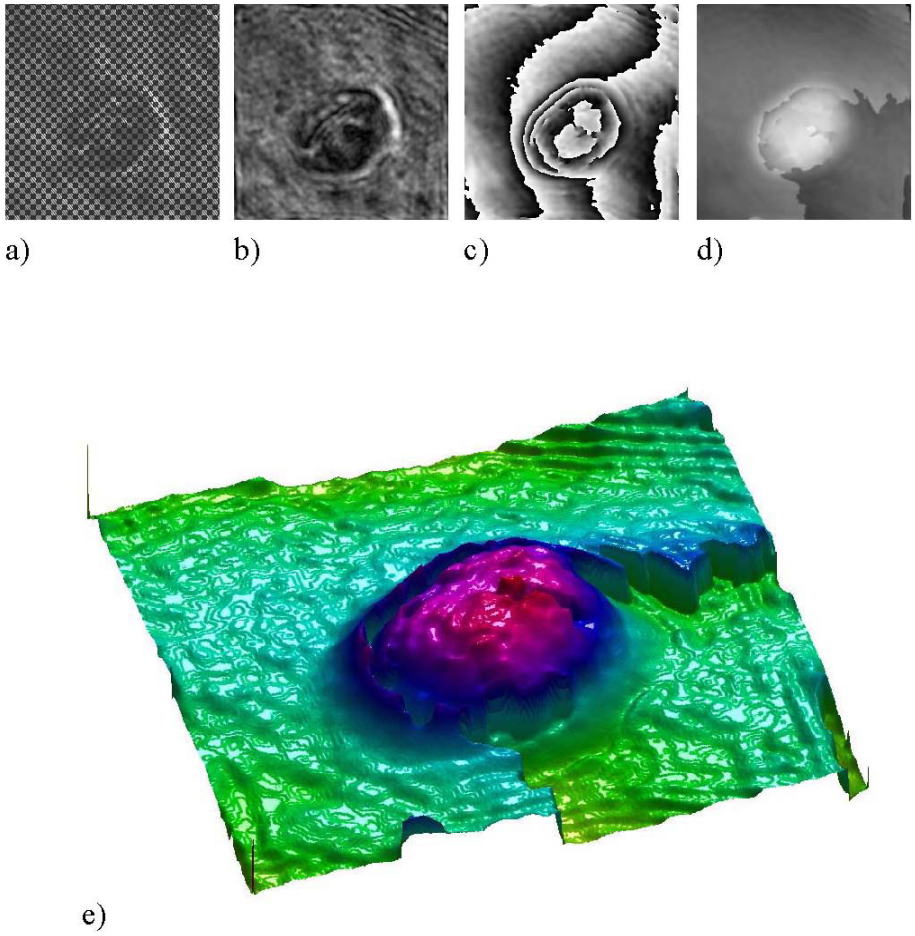


Figure 5.17: Holography of an onion nucleus. The image area is 30 x $30\mu m^2$ (436 x 436 pixels) and the image is at $z = 22\mu m$ from the hologram: a) hologram, b) holographic amplitude and c) phase images; d) unwrapped phase image; e) 3D pseudocolour rendering of d).

Figure 5.18 shows another example of digital holography of an onion nucleus. The panels display the (a) hologram, (b) amplitude image, (c) phase image, and (d) phase image unwrapped by a software algorithm. Pseudo-colour 3D rendering of (d) is shown in (e).The image size is 30 × 30μm^2 with 452 × 452 pixels and is reconstructed at z = 13μm from the hologram. The nucleus appears close to the cell wall and appears to be enlarged to a significant degree.

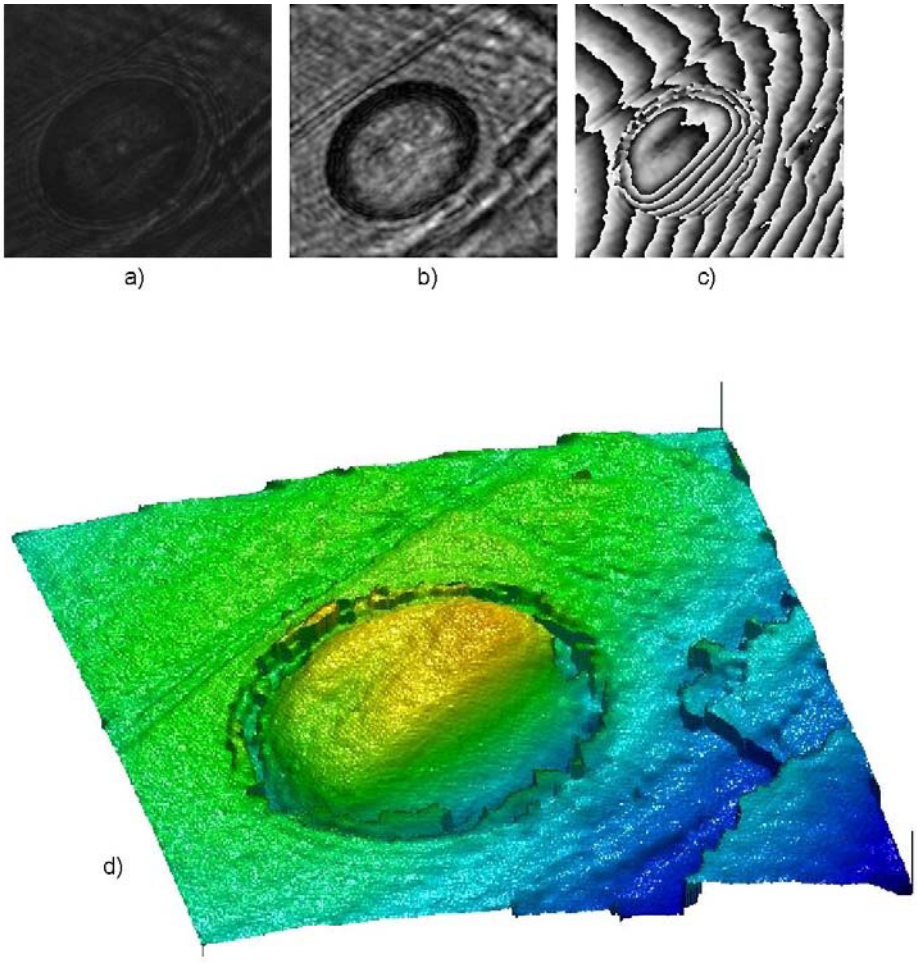


Figure 5.18: Holography of an onion nucleus. The image area is 40 x 40μm^2 (452 x 452 pixels) and the image is at z = 13μm from the hologram: a) hologram, b) holographic amplitude and c) phase images; d) unwrapped phase image; e) 3D pseudocolour rendering of d).

Figure 5.19 displays digital holography of red blood cells. The panels display the (a) hologram, (b) amplitude image, (c) phase image, and (d) phase image unwrapped by a software algorithm. Pseudo-colour 3D rendering of (d) is shown in (e).The surrounding medium is air and we apply a constant index of refraction estimate of $n = 1.375$ for the blood cell to produce the optical thickness scaling as seen in the pseudo-color 3D rendering. From the analysis of the phase map we infer the average optical thickness of the blood cells to be around 0.6μm. Figure 5.20 shows the 3D rendering from Fig.5.19 (e) where the color-map has been adjusted for better color visualization of the red blood cells.

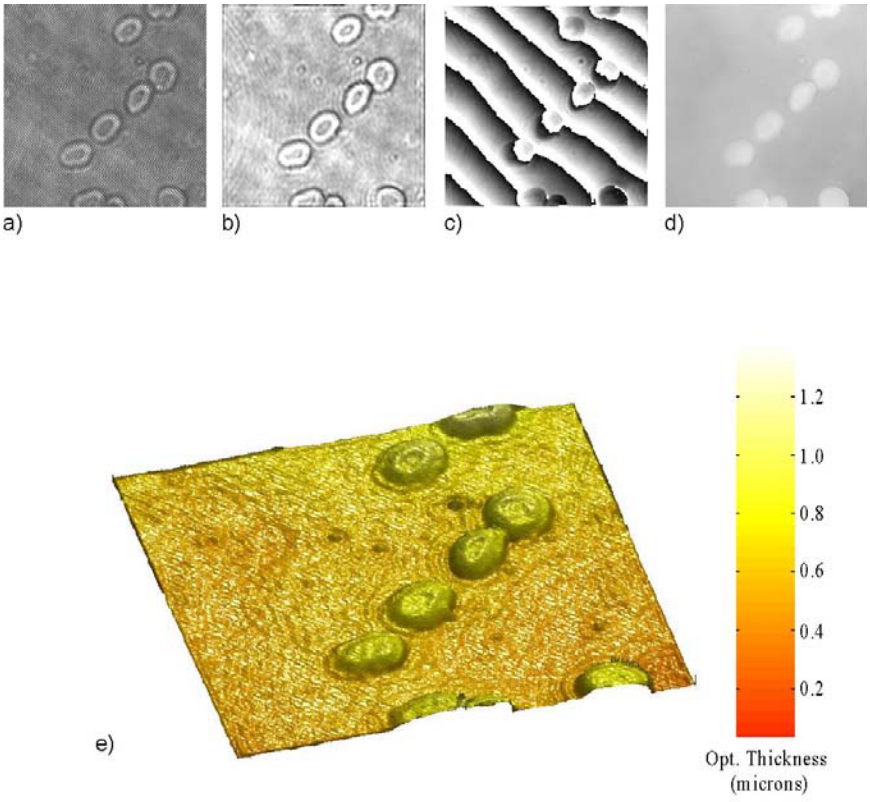


Figure 5.19: Digital holographic images of red blood cells of area 60 x 60 μm^2 (408 x 408pixels). The image is at z = 4μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image and (e) 3D pseudocolor rendering of (d).

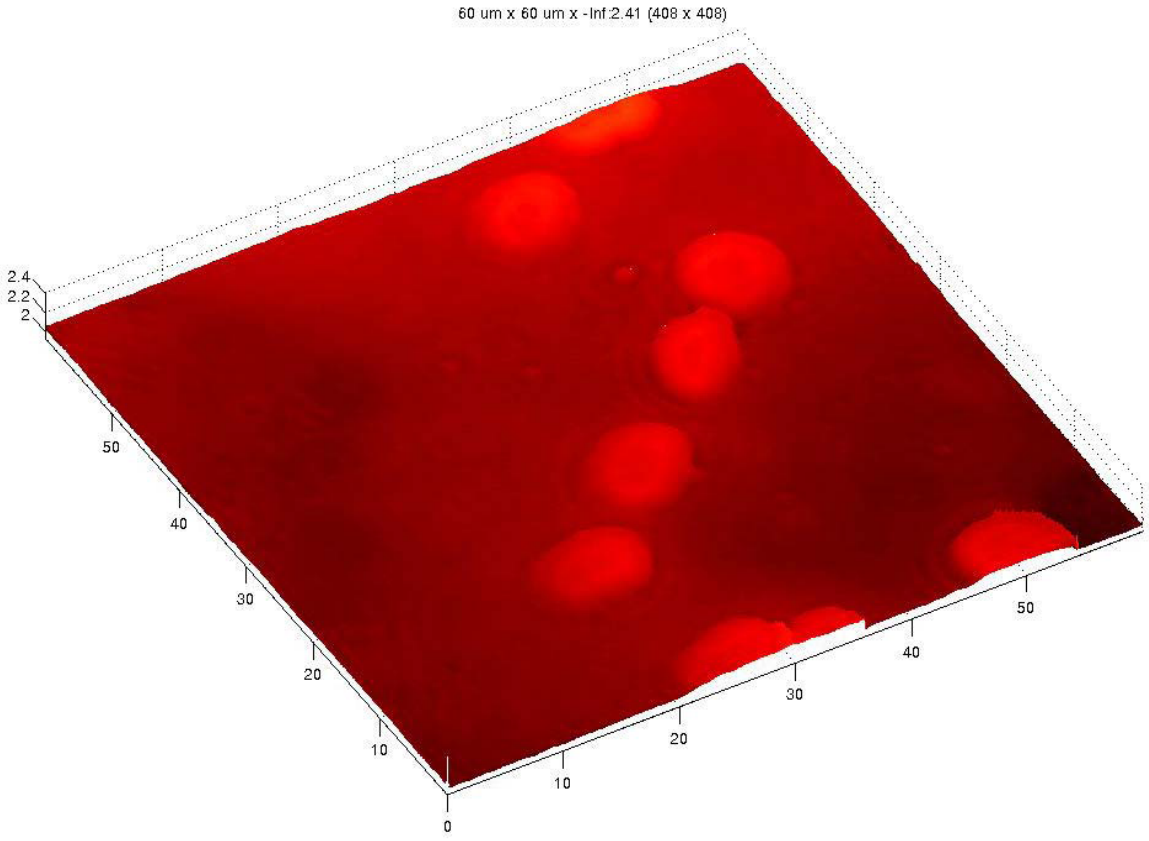


Figure 5.20: 3D pseudo-color rendering with color-map adjustment of the unwrapped phase image in Fig. 5.18 (d).

Figure 5.21 displays digital holography of a single red blood cell. The panels display the (a) hologram, (b) reconstructed amplitude image, (c) reconstructed phase image, and (d) phase image unwrapped by a software algorithm. (e) The Fourier Transform of (d) is shown in Fig. 5.21 (e) and the pseudo-color 3D rendering of (d) is shown in (f). Figure 5.21 (g) is an SEM image of a red blood cell at 14,000x. The morphological shape of the cell in the 3D rendering clearly defines the cell as a flat disk with the center pushed in which provides a larger surface area for absorption of oxygen. The shape in the 3D rendering corresponds well to that of the SEM image. The thickness of the cell is again measured to be approximately 0.6μm and the diameter is approximately 8μm. The angular spectrum shows the clear separation of the holographic terms and the *dc* component by careful control of the optical elements in combination with the use of the angular spectrum in real time.

Figure 5.22 shows digital holography of a number of red blood cells. Some of which are seen to be clumped together. The panels display the (a) hologram, (b) reconstructed amplitude image, (c) reconstructed phase image, and (d) the phase image unwrapped by a software algorithm. A pseudo-color 3D rendering of figure 5.22(d) is shown in figure 5.22(e). The red blood cells show the well defined central indentation in the form of the familiar doughnut shape. Note here, the uniformity in the shape of the cells but slight variations in the size. The variation in size could be due to drying over time, which would undoubtedly affect the morphological properties of the cell. The red blood cells are seen to have no internal organelles and are internally uniform and amorphous.

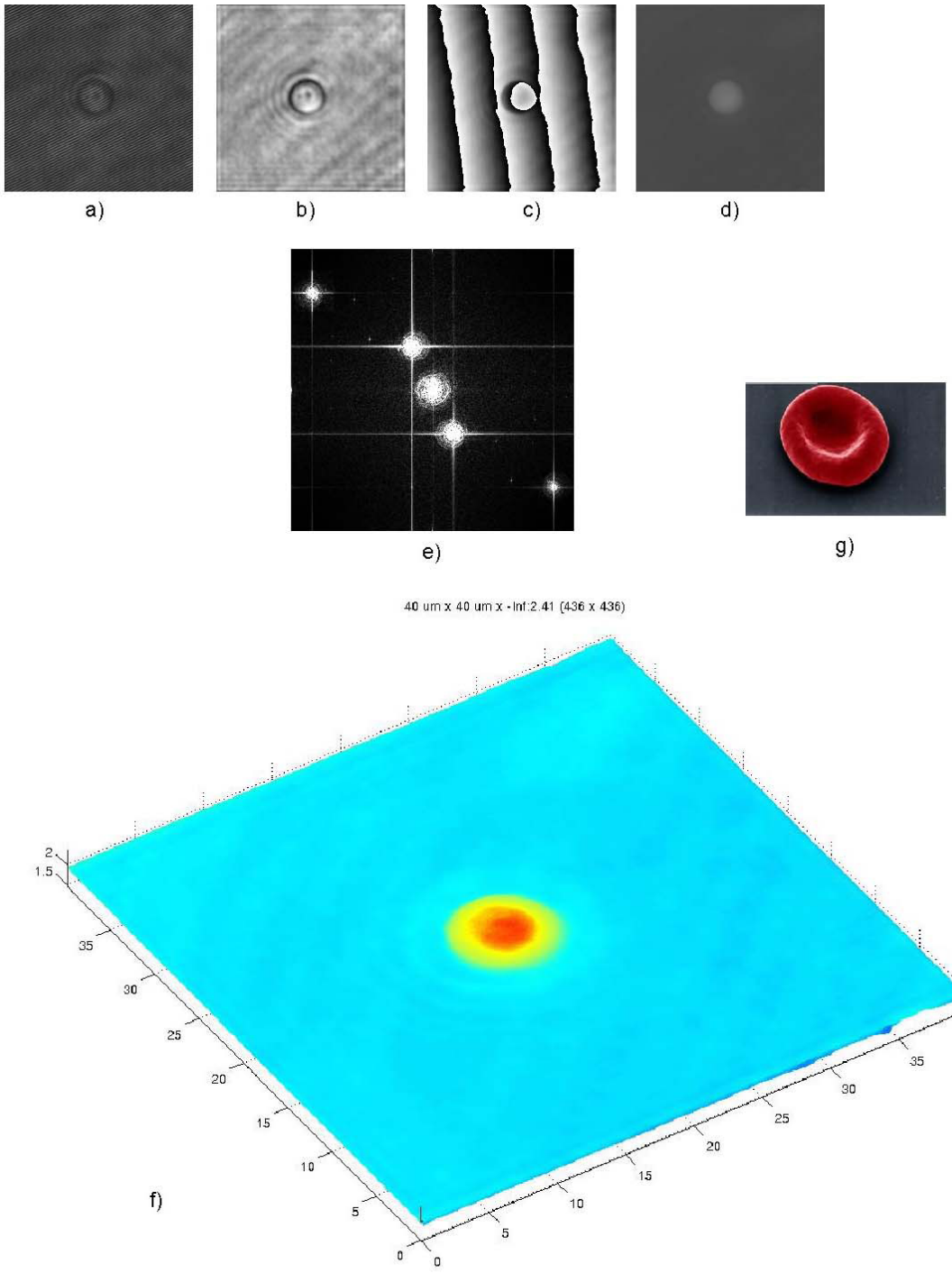


Figure 5.21: Digital holographic images of a single red blood cell of area 60 x 60 μm^2 (408 x 408pixels). The image is at z = 4μm from the hologram: (a) hologram; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image, (e) angular spectrum, (f) 3D pseudo-color rendering of (d) and (g) SEM image at 14,000x.

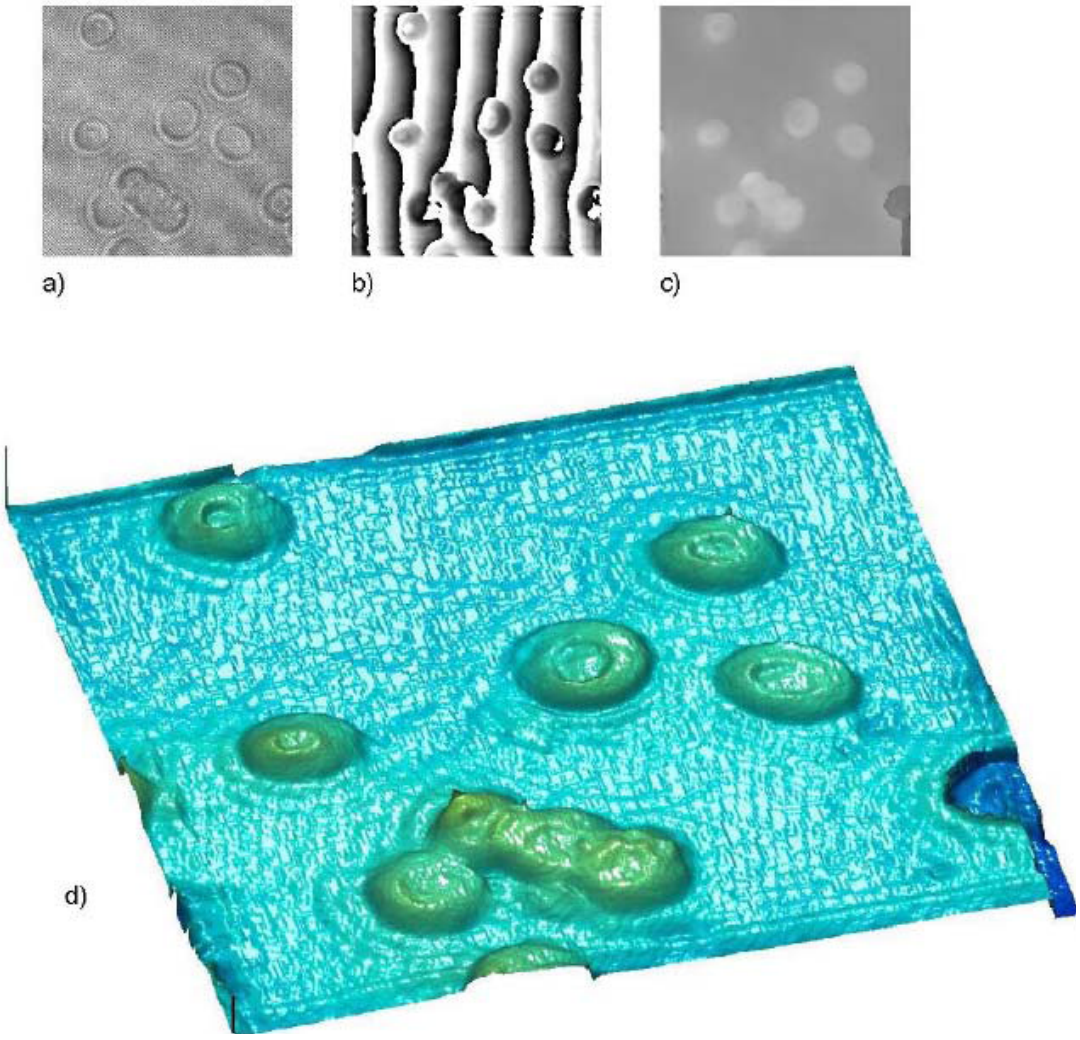


Figure 5.22: Digital holographic images of red blood cells with area 50 x 50μm^2 . The image is at z = 6μm from the hologram: (a) Hologram; (b) wrapped phase image; (c) unwrapped phase image and (d) 3D pseudocolor rendering of (c).

Figure 5.23 displays the reconstruction of a SKOV-3 ovarian cancer cell. The area in the image is $60\times60\mu m^2$, recorded with 424×424 pixels. Figure 5.23(a) is the holographic interference pattern recorded by the CCD camera, and its Fourier transform in figure 5.23(b) is the angular spectrum. It contains both the zero order and twin images, as well as an artifact due to stray interference components. The virtual image component, the highlighted circular area, is selected. A propagation phase factor (z=1.0μm) is multiplied, and finally inverse-Fourier transformed to obtain the amplitude image in figure 5.23(c) and the phase image in figure 5.23(d).

The layer of lamellipodia around the edge of the cell is found to be about 110 nm, assuming n = 1.375 for the cell. The phase map is rendered in pseudo-coloured 3D perspective in figure 5.23(h). Especially notable in the phase map is the lack of the coherent noise conspicuous in the amplitude image and prevalent in most other holographic imaging methods. The amplitude and phase images obtained from the Huygens convolution method are shown in figures 5.23(f) and 5.23(g), while those obtained from the Fresnel method are omitted because they are completely scrambled. The main reason for the obvious degradation of these images is the insufficient off-axis angle at such short z distance to separate out the zero-order component.

In Figure 5.24 we show digital holography of a living HUVEC (Human Umbilical Vein Endothelial Cell). The image demonstrates high clarity and quality, displaying the nuclear membranes morphological structure in detail. The lamellipodia of the cell are seen to extend out in order to occupy a large area as it attempts to migrate. The improvements in the digital holographic process is achieved in part by the use of the

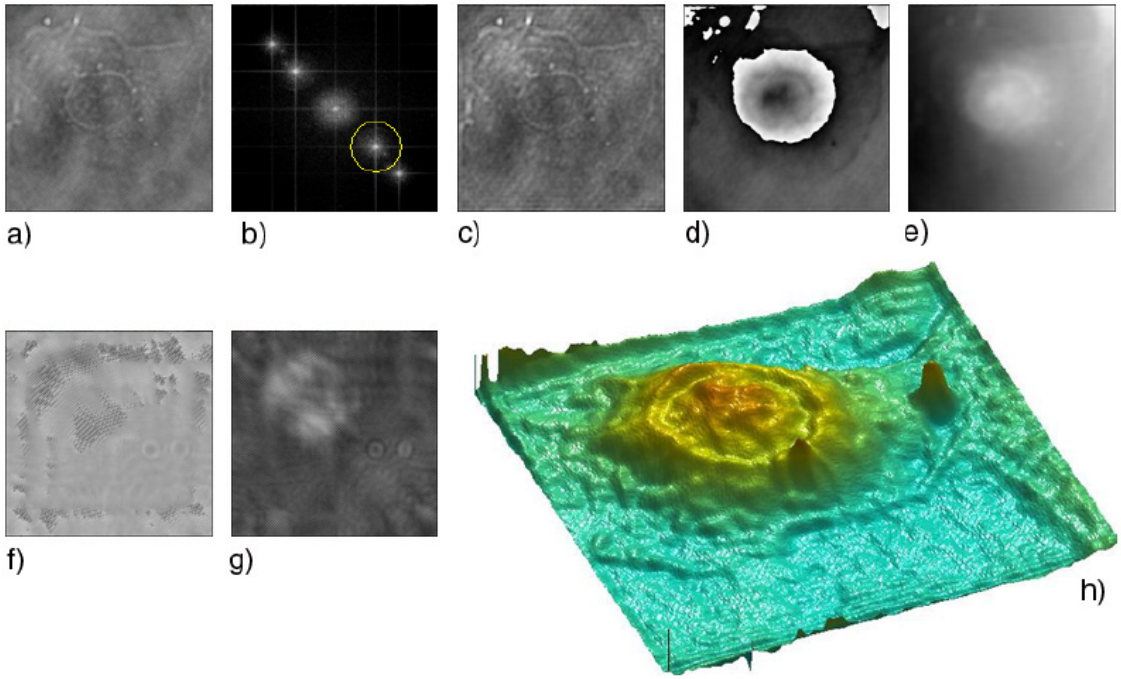
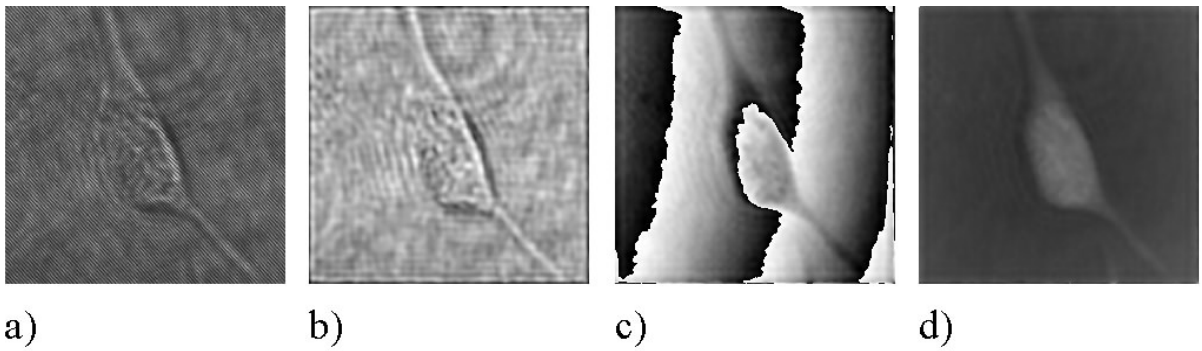
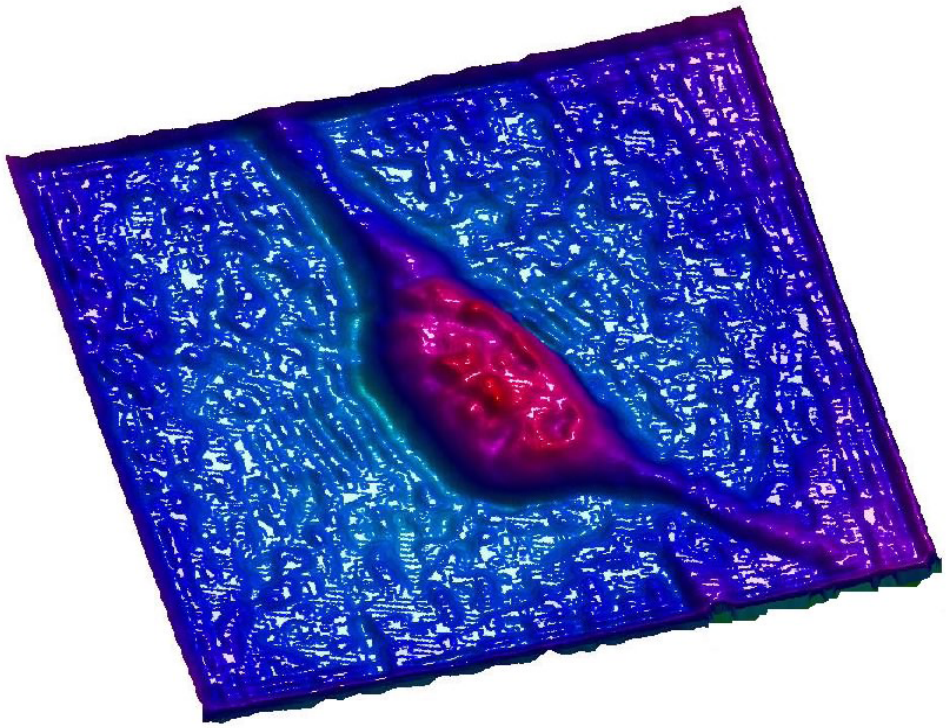


Figure 5.23: Digital holography of a SKOV ovarian cancer cell. The image area is 60 × 60μm (424 × 424 pixels) and the image is at $z = 1.0\mu$m from the hologram: (a) hologram; (b) angular spectrum; (c) amplitude and (d) phase images by the angular spectrum method; (e) unwrapped phase image of (d); (f) phase and (g) amplitude images by the Huygens convolution method. (h) 3D perspective rendering of (e).



a) b) c) d)



e)

Figure 5.24: Holography of a HUVEC cell. The image area is 70 x 70 μm^2 (448 x 448 pixels) and the image is at z = 12μm from the hologram: a) hologram; b) amplitude and c) phase images; d) unwrapped phase image; e) 3D perspective rendering of d).

angular spectrum method for diffraction calculation, which has several advantages over the more commonly used Fresnel transformation or Huygens convolution methods. Spurious noise and interference components can be tightly controlled through the analysis and filtering of the angular spectrum. The reconstruction distance does not have a lower limit and the off-axis angle between the object and reference can be lower than the Fresnel requirement and still be able to cleanly separate out the zero-order background.

Figure 5.25 presents holographic images of a confluent group of SKOV-3 ovarian cancer cells by angular spectrum digital holography. The panels display (a) Zernike phase contrast image (of similar but different portions of the sample than the holography images), (b) holographic amplitude and (c) phase images, and (d) phase image unwrapped by a software algorithm. Pseudocolor 3D rendering of (d) is shown in (e). The image size is 60 x 60 μm^2 with 404 x 404 pixels.

In the figure we see the phenomenon of cuboidal cells connecting together into an epithelial sheet and producing the grooves between cells. A gap in the confluence is also accurately imaged in Fig. 5.25(e), except for a few spikes due to a defect in the phase-unwrapping algorithm.

The comparison between the amplitude and phase images highlights the main differences in image quality and noise dependence. In the amplitude image many features of the cells are hard to distinguish from the background and significant intensity variation may indicate the presence of coherence noise which is not visible in the phase images.

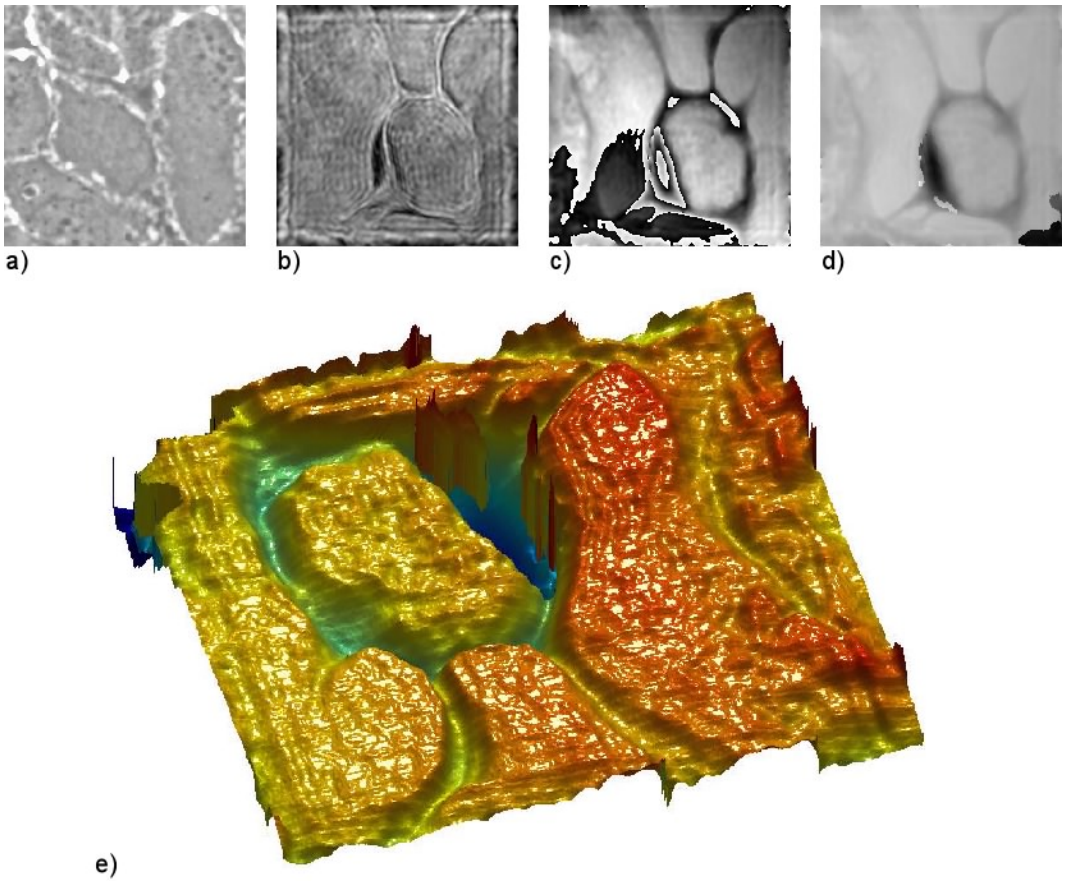


Figure 5.25: Holography of confluent SKOV-3 ovarian cancer cells. The image area is 60 x 60μm^2 (404 x 404 pixels) and the image is at z = 10μm from the hologram: (a) Zernike phase contrast image; (b) holographic amplitude and (c) phase images; (d) unwrapped phase image; (e) 3D pseudocolor rendering of (d).

Figure 5.26 is a particularly unambiguous demonstration of the level of image resolution and fidelity that can be obtained by the present technique, displaying the nuclear membranes and chromosomes. The panels display (a) Zernike phase contrast image (of similar but different portions of the sample than the holography images), (b) holographic amplitude and (c) phase images, and (d) phase image unwrapped by a software algorithm. Pseudocolor 3D rendering of (d) is shown in (e). The image size is 60 x 60 μm^2 with 404 x 404 pixels. The overall height of the cell is calculated to be about 2.8μm, with the assumption of the average index of the cell to be 1.375. Thickness of the lamelipodium around the edge of the cell is determined to be about 320 nm. The noise level in the substrate area is 60 nm, which may be partly due to the residues from fixing of the cells. From this we infer that the corrugated texture of cellular surfaces in Fig. 5.26(e) is not likely to be noise or artifact of holographic process. Such texture is evident in the Zernike phase contrast image of Fig. 5.26(a) and is known to exist in such confluent cells. The depth of the texture is estimated to be about 120 nm.

Figure 5.27 presents digital holography of a single SKOV-3 ovarian cancer cell. The panel displays the (a) hologram, (b) holographic amplitude and (c) phase images, and (d) phase image unwrapped by a software algorithm. Pseudocolor 3D rendering of (d) is shown in (e). The image size is 60 x 60 μm^2 with 404 x 404 pixels. The image shows a clear view of the thin but broad projections of the lamellipodia at the edge of the mobile cell which pushes out in front of the cell in the direction of movement as it migrates. The mechanical tension generated by the lamellipodia appears to stretch the cell into an elongated form when compared to the SKOV-3 cells as shown in Fig. 5.26. Also of interest here is the rather irregular shape of the lamellipodia.

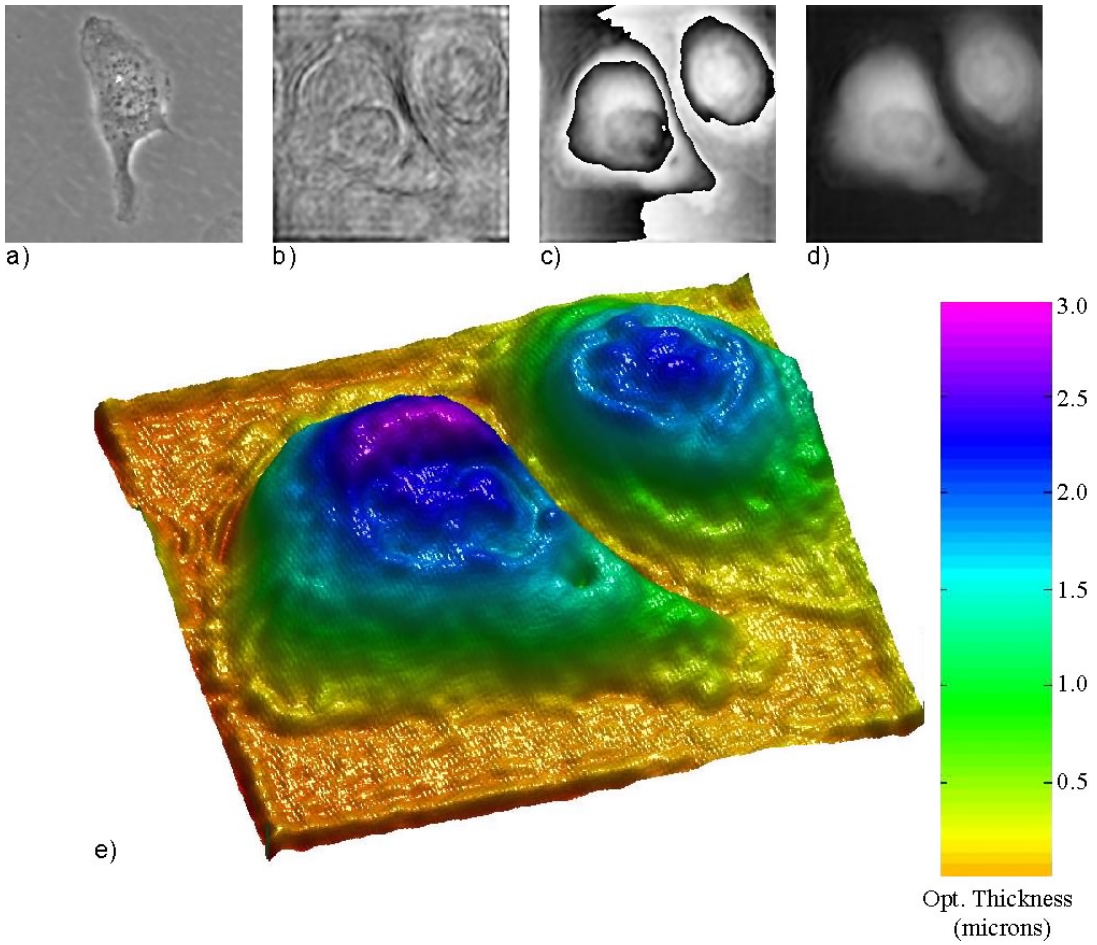


Figure 5.26: Holography of non-confluent SKOV-3 cells. The image area is 60 x 60 μm2 (404 x 404 pixels) and the image is at z = 5μm from the hologram: (a) Zernike phase contrast image; (b) holographic amplitude and (c) phase images; (d) unwrapped phase image; (e) 3D pseudocolor rendering of (d).

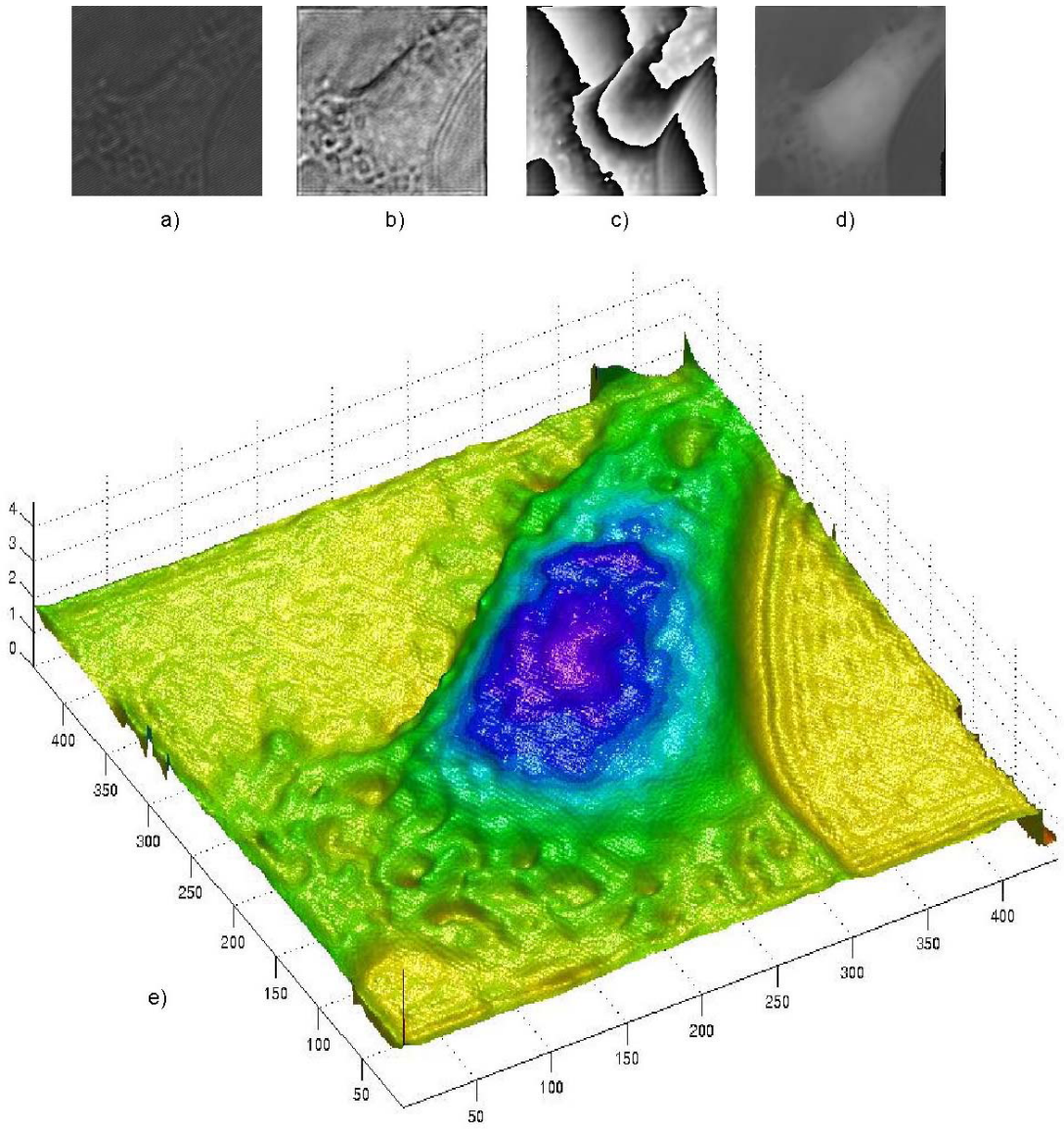


Figure 5.27: Holography of a single SKOV-3 cell. The image area is 60 x 60 μm2 (404 x 404 pixels) and the image is at z = 4μm from the hologram: (a) hologram; (b) amplitude and (c) phase image; (d) unwrapped phase image; (e) 3D pseudocolor rendering of (d).

Figure 5.28 displays images of a single human epithelial cheek cell obtained by quantitative phase contrast microscopy. The image size is 60 x 60 μm^2 with 404 x 404 pixels. The cheek cell is about 50μm in diameter and again corresponds well with literature. The nucleus and cell membrane are clearly visible as is the mitochondria, distributed in the proximity of the nucleus. Also of note here is the irregular shape and contours of the cell membrane, with some sections of the cells showing folding.

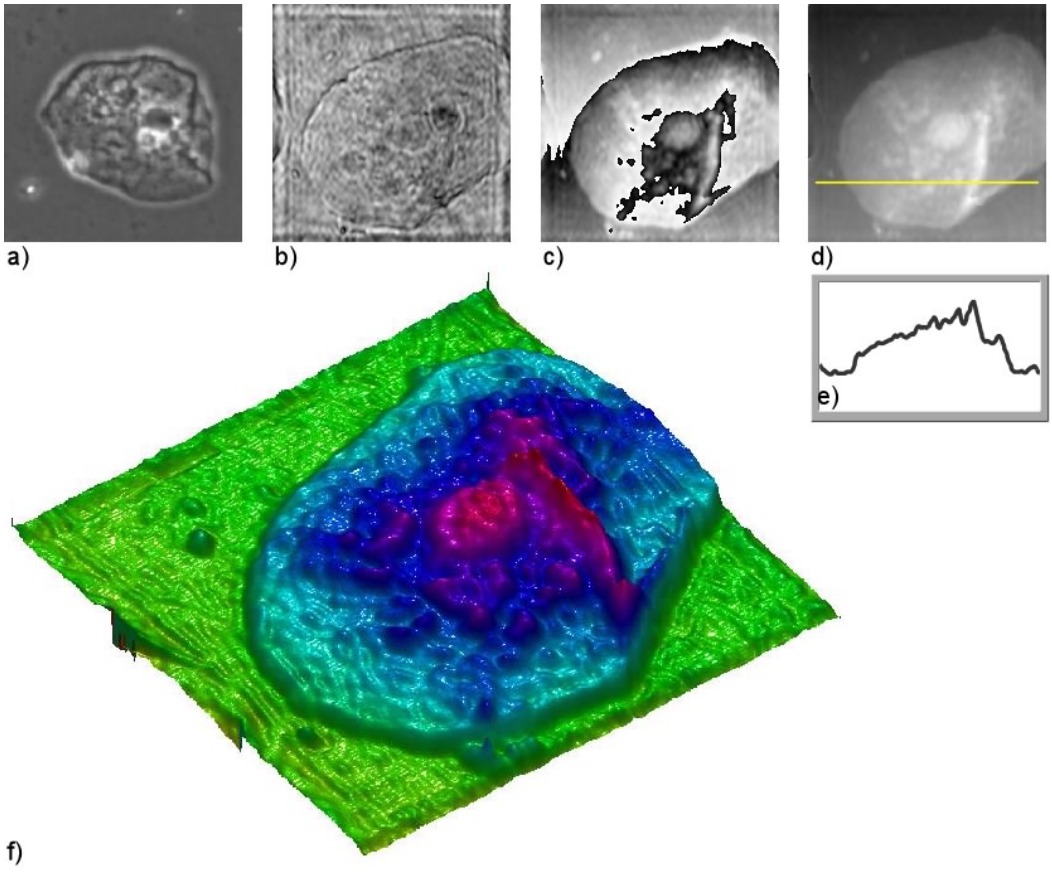
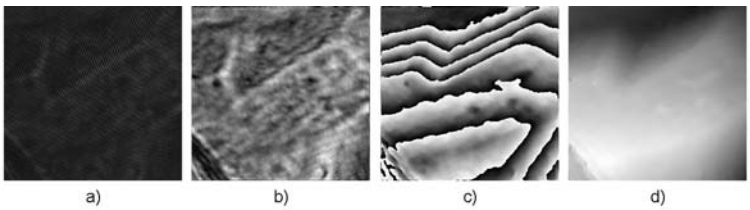


Figure 5.28: Digital holographic images of a Human cheek cell with area 60 x 60μm^2. The image is at z = 9μm from the hologram:
(a) Zernike phase contrast image; (b) amplitude; (c) wrapped phase image; (d) unwrapped phase image; (e) z-profile cross section and(f) 3D pseudo-color rendering of (d).

As well as biological applications for microscopy, digital holography can also be applied for accurate and high resolution shape measurement in a number of other fields including materials characterization.

Figure 5.29 presents an example of quantitative digital holography from a fragment piece of quartz. The panel displays the (a) hologram, (b) holographic amplitude and (c) phase images, and (d) phase image unwrapped by a software algorithm. Pseudocolor 3D rendering of (d) is shown in (e). The image size is 60 x 60 μm^2 with 404 x 404 pixels. The contour of the quartz is seen to be somewhat graded and jagged in its thickness profile.



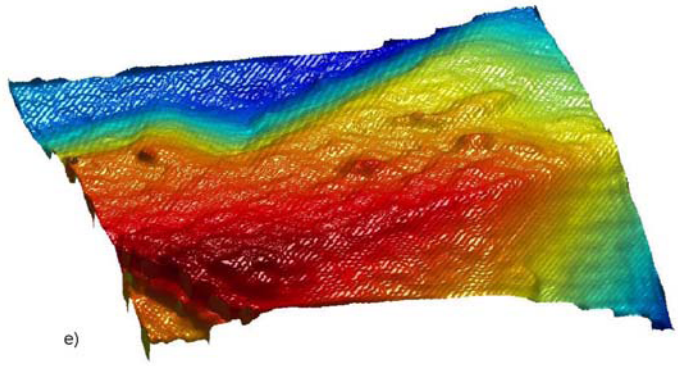


Figure 5.29: Holography of a fragment of quartz material. The image area is 60 x 60 μm^2 (404 x 404 pixels) and the image is at z = 7μm from the hologram: (a) Zernike phase contrast image; (b) holographic amplitude and (c) phase images; (d) unwrapped phase image; (e) 3D pseudocolor rendering of (d).

Figure 5.30 displays digital holography of a water droplet. The shape is spherical and smoothed with the thickest part of the droplet occurring at the center.

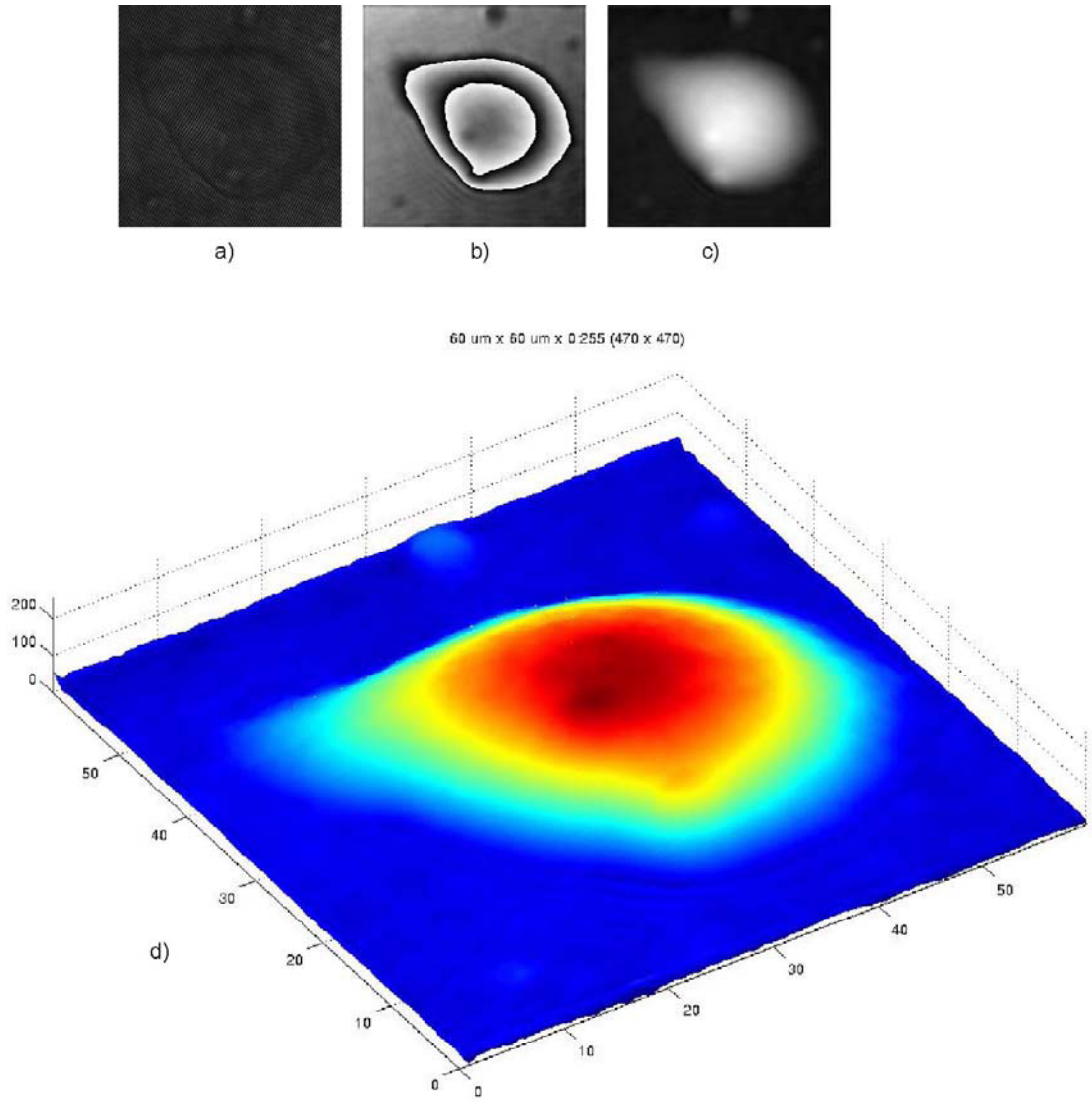


Figure 5.30: Holography of a water droplet. The image area is 60 x 60 μm^2 (404 x 404 pixels) and the image is at z = 12μm from the hologram: (a) hologram; (b) phase images; (c) unwrapped phase image; (d) 3D pseudocolor rendering of (c).

5.8 Off-Axis Digital Holographic Movies

In digital holography, a series of holograms can be recorded at a fixed distance and the reconstructed images later numerically focused at leisure. In this way, one is able to track the rapidly moving object(s) quantitatively, as needed. We have recorded a series of holograms by synchronizing the camera with the laser pulses under the control of the digital delay generator. The amplitude and phase images are calculated by the angular spectrum method while adjusting image distances for best focus of the object under observation. The reconstructed images are subsequently composed into AVI files. The phase movie generated is quantitative, and therefore, it is possible to make dynamic measurements of physical and morphological changes in the sample over time. In particular, this has many applications in microscopy where tracking of changes in events, like mitosis in cells, can be analyzed in detail.

Figure 5.31 shows a time series of selected frames from the reconstructed amplitude movie of a paramecium swimming around with a number of euglenas. The image area of each frame is 250 x 250μm^2, recorded with 464 x 464 pixels. The amplitude images are reconstructed from a series of holograms while adjusting the image distances over a 20μm range in the range of z = 850 ~ 870 μm in order to maintain the paramecium in focus. The time interval between each reconstructed frame is t= 800ms and the measured reconstruction rate of each recorded hologram is achieved in around 300 ms.

Figure 5.32 shows the reconstructed phase images in a time series corresponding to the amplitude images in Figure 5.31. Note the phase information of the paramecium's intracellular structure is much more definitive than that of the amplitude information.

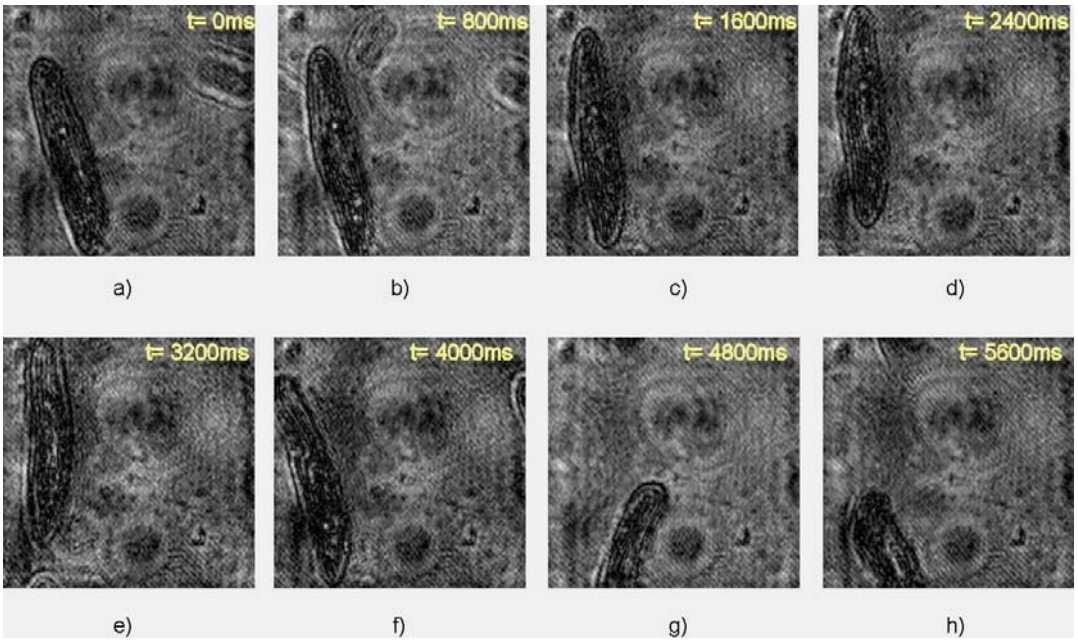


Figure 5.31: Time series of reconstructed amplitude images of a moving paramecium. The image area is 250 x 250μm^2 (464 x 464 pixels). The numerical focus was adjusted in the range of z = 850 ~ 870 μm in order to track the paramecium in focus.

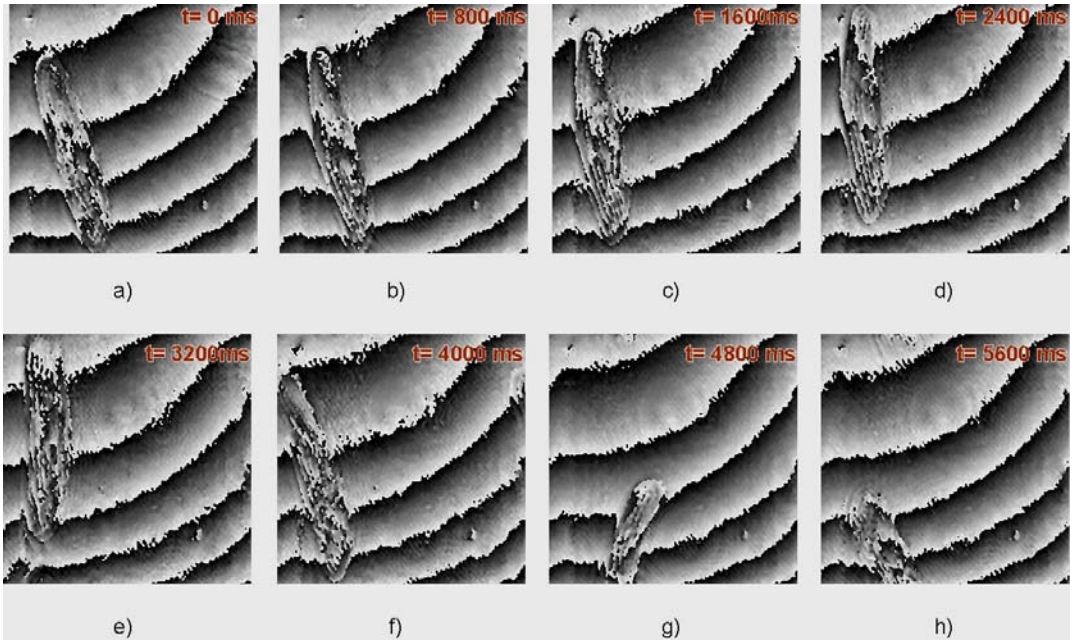


Figure 5.32: Time series of reconstructed phase images of a moving paramecium corresponding to the amplitude images in Fig.5.29.

Higher magnification time-series amplitude images, selected from an amplitude movie are illustrated in Figure 5.33. The image area of each frame is 80 × 80 μm^2, recorded with 452 x 452 pixels and reconstructed at z = 123μm. The time interval between each reconstructed frame is *t*= 800ms.

In order to maintain the paramecium within the field of view for long enough, the paramecium is slowed down using a drop of the thickening agent methylcellulose. In the amplitude images, one observes the flickering due to celial motion in the oral groove that forces food (bacteria) into the food vacuoles. Note the large range of intracellular structures and particles which cause speckle and image degradation.

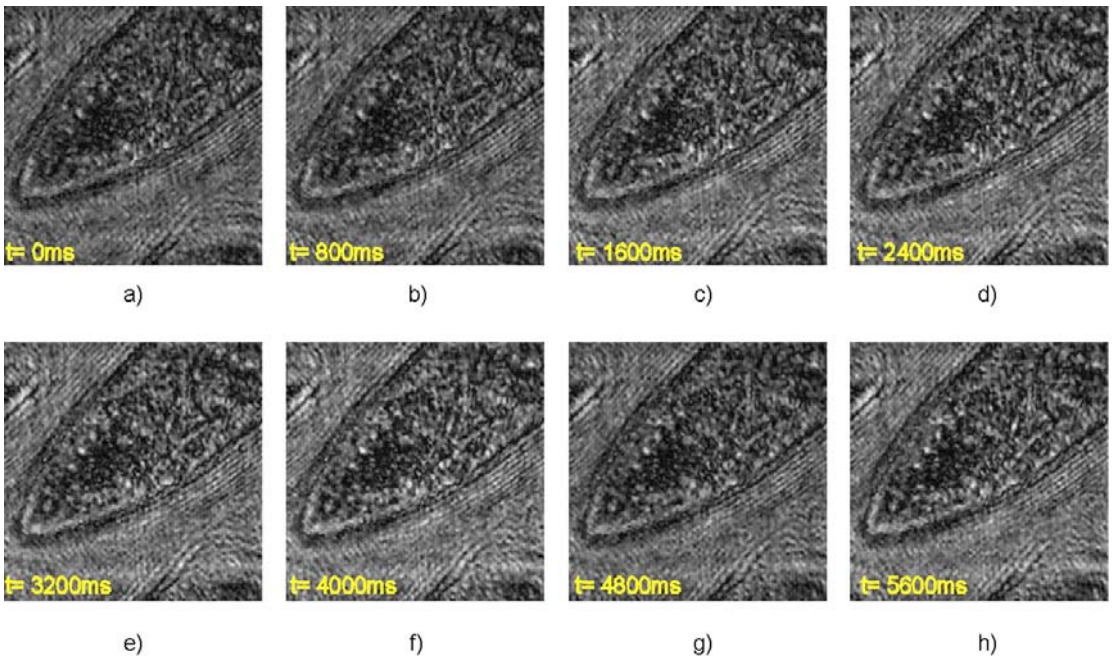


Figure 5.33: Time series of reconstructed amplitude images of a paramecium with image area 80 x 80 μm^2 (452 x 452 pixels) with z = 123μm.

Figure 5.34 shows the time series of reconstructed phase images corresponding to the amplitude images in Figure 5.33. The phase images are reconstructed simultaneously to the amplitude images and either movie is available to the user at the same time.

A notable phenomenon in the phase movie is the shrinking of the contractile vacuole, highlighted, as it pumps out water from the paramecium. This feature is not easily observed in the amplitude frames in Fig. 5.33. The speckle noise in the paramecium does not appear to be as much of a problem in the phase images as it is in the amplitude images, in areas where the paramecium contains a large number of intracellular features.

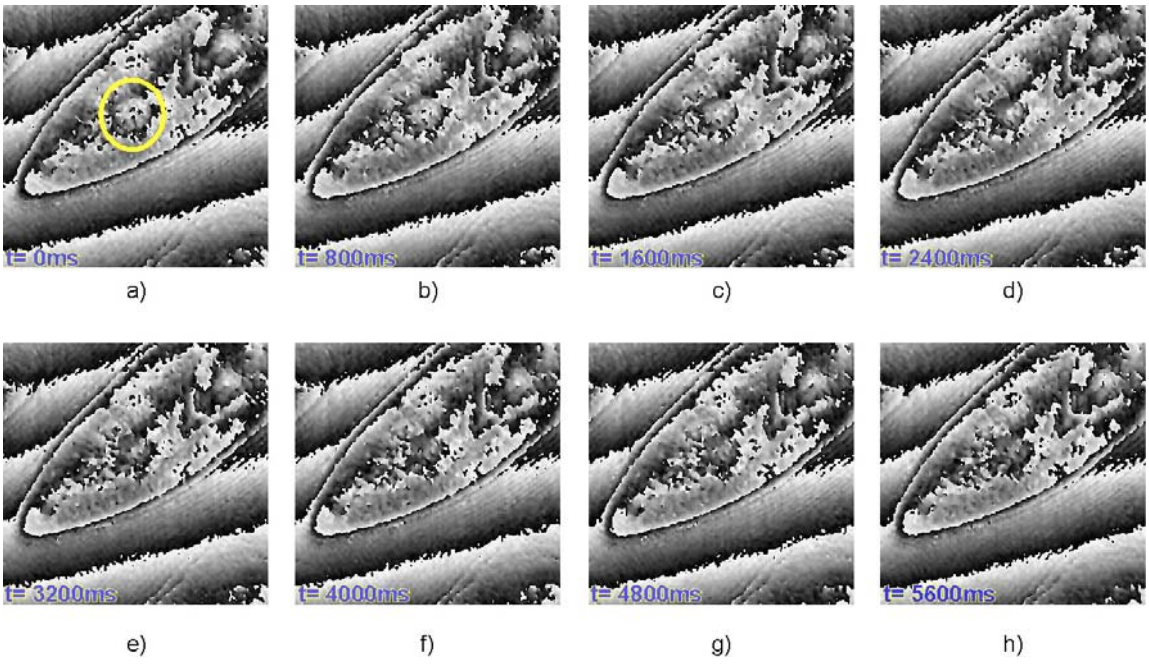


Figure 5.34: Time series of reconstructed phase images of paramecium corresponding to the amplitude images in Fig.5.31.

Figure 5.35 displays a time series of selected frames from the amplitude and phase movies of the migration process by living mouse-embryo fibroblast cells (3T3). The image area is 170 × 170 μm^2, recorded with 460 x 460 pixels and the reconstruction distance is z = 6μm. In the upper row, Figures 5.35 (a-d) show selected frames from the amplitude movie and in the bottom row Figures 5.35 (e-h) are the simultaneously produced phase images from the phase movie. The time interval between each frame is one hour. Cell migration is a fundamental function of normal cellular processes and its understanding is very much an issue for scientific researchers.

The cells are not very mobile suggesting that the conditions for migration are not ideal. This is almost certainly because of the low temperature and CO_2 concentration which inhibit the cell motion.

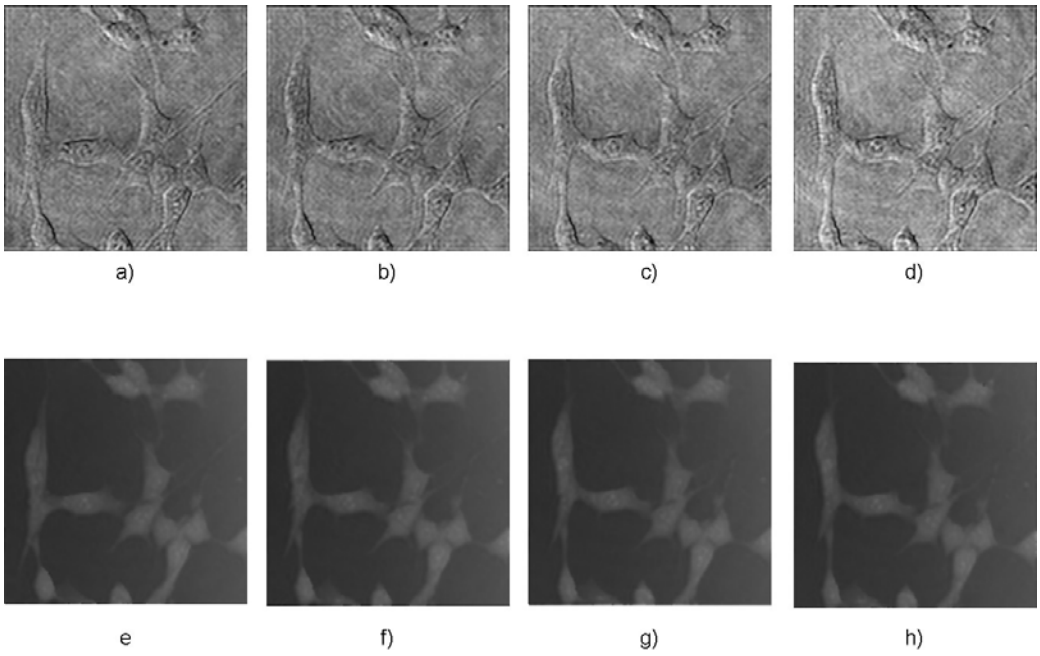


Figure 5.35: Selected frames from the amplitude movie (a-d) and from the corresponding phase unwrapped movie (e-h) of mouse-embryo fibroblast cells in the process of migration.

Figure 5.36 presents selected frames from the 3-D optical thickness profile phase movie corresponding to the reconstructed amplitude and phase images presented in Fig. 5.35. The nucleus is clearly seen in the cells as are smaller intra-nuclear characteristics. The cells exhibit directed migration by formation of lamellipodia extensions towards a specific target cell. Connections between some of the cells are also displayed.

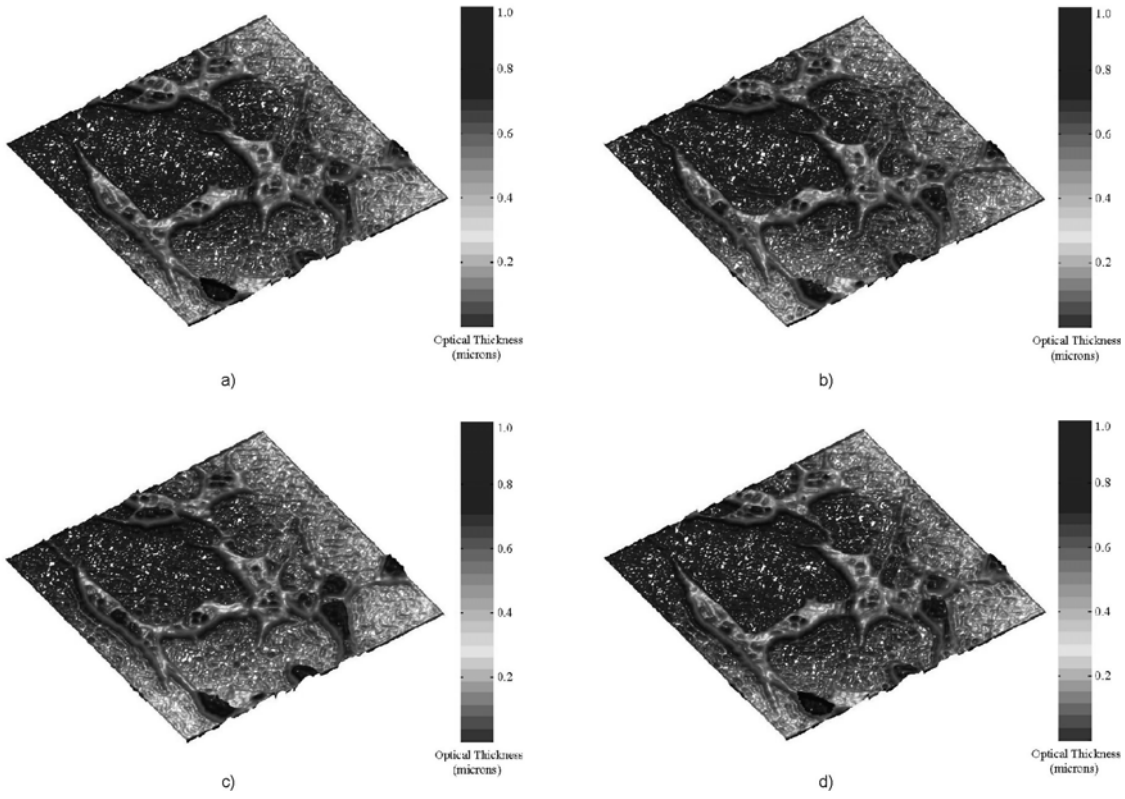


Figure 5.36: Selected frames from the 3-D optical thickness profile phase movie of mouse-embryo fibroblast cells in the process of migration

Figure 5.37 displays selected frames from the 3-D optical thickness profile phase movie showing the process of fibroblast cell mitosis. The study demonstrates that time lapse video microscopy is a simple but very useful approach to monitor the dynamics of movements which vary in speed and frequency during migration and mitosis of living cells.

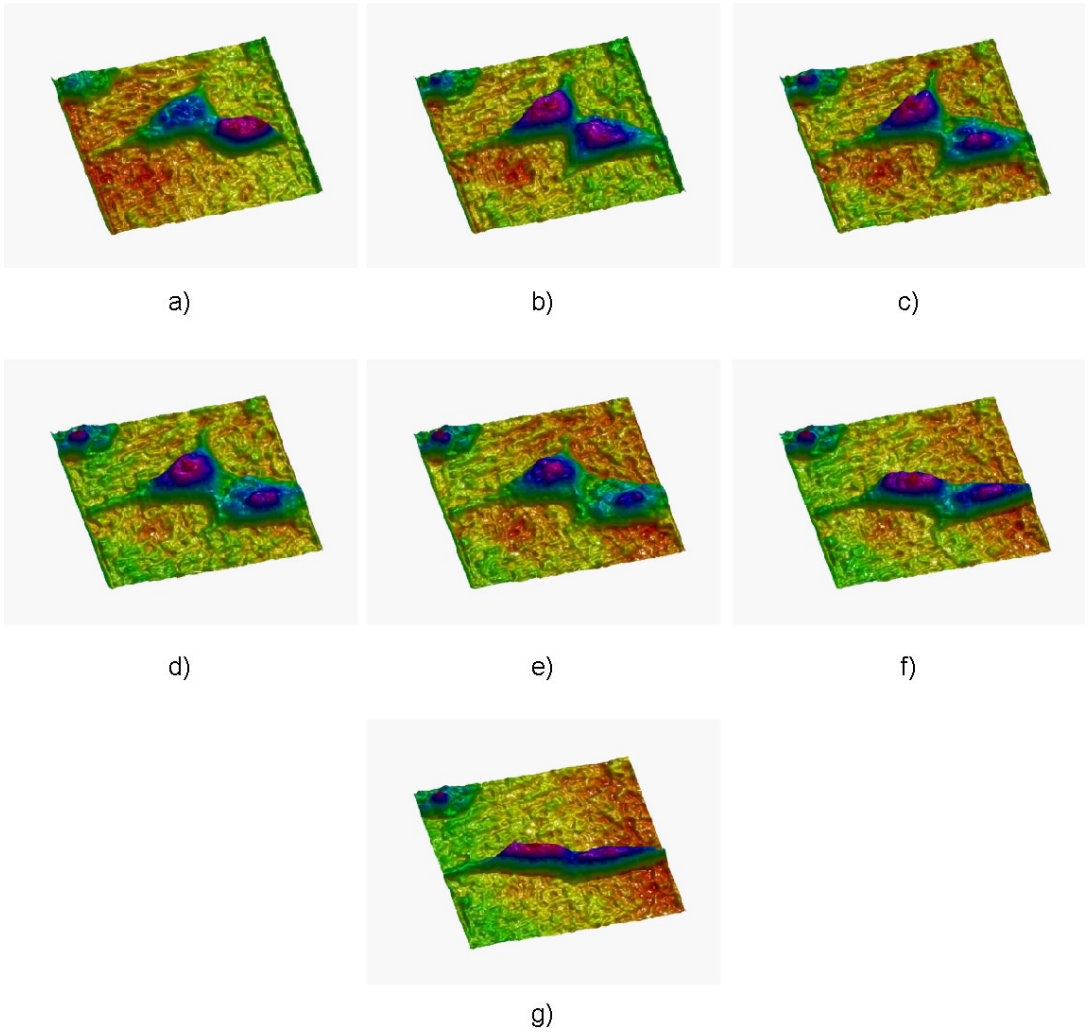


Figure 5.37: Twelve-hour time lapse movie of fibroblast cells undergoing mitosis, obtained by digital holographic microscopy. The pseudo-color 3D rendering represents the optical thickness profile of the cells.

5.9 Discussion and Conclusion

Digital holography offers a highly sensitive and versatile means of measuring and monitoring optical path variations. From the recorded hologram, both amplitude and phase information are simultaneously available, therefore enabling digital holography to act essentially as both a conventional and a phase microscope. The holographic reconstruction is carried out by the angular spectrum method in an off-axis configuration so that the twin images and *dc* term background can be subtracted out using frequency domain analysis. The use of angular spectrum method is also seen to solve some of the significant problems that have prevented wider use of digital holography in biomedical imaging applications. These advantages include tight control of spurious noise components and no minimum recording distance. Also a combination of filtering windows of various shapes and sizes can be used to filter the undesired spatial frequencies while keeping intact the contributions of the interference terms.

Phase imaging digital holography is particularly promising in terms of the lack of coherent noise and the precision with which the optical thickness can be profiled, which leads to images with high optical thickness resolution. Experimental results have been presented which demonstrate the effectiveness of digital holography for quantitative phase-contrast microscopic imaging of transparent objects. Quantitative knowledge of the structures and dynamics of a transparent specimen is of great importance in microscopy, particularly in cell biology. Conventional phase-contrast microscopies, such as Zernike and DIC, produce high-contrast qualitative images of transparent objects, but they cannot be directly applied to quantitative studies.

The experimental results display 0.5μm diffraction-limited resolution, with the noise level in the phase profile corresponding to several nm of optical thickness. The lateral and longitudinal resolutions obtained are consistent with diffraction limited imaging. Reconstructed images of objects such as living biological cells display intracellular and intra-nuclear organelles with sufficient clarity and quantitative accuracy for applications in biomedical research. The level of resolution and details in these images clearly exceeds currently available techniques in phase-contrast optical microscopy, and provides a new modality for imaging morphology of cellular, intracellular, and intranuclear structures that is not currently available with non-invasive optical methods.

We have also presented experimental results that demonstrate the capabilities of digital holographic movies for biological microscopy. A series of holograms are recorded and the images are reconstructed with numerically adjustable focus so that the moving objects can be accurately tracked.

On the other hand, the speckle noise of a coherent imaging system is a significant issue for biological microscopy where there is often a large range of structural scales. For example, in the presented holographic amplitude images of paramecium the intracellular regions contain various assortments of unresolved particles, which may be food particles or other organelles. These degrade the quality of images to various degrees. However, one also notices that the phase images tend to suffer from the coherent noise to a significantly lesser degree compared to the amplitude images.

A disadvantage of the current method of producing digital holographic movies is the time-consuming procedure of phase unwrapping which prevents the user from

observing changes in an object in real time. There are other phase unwrapping procedures available that will require testing and this will be a subject of future study. In the next chapter we describe and present results of multi-wavelength digital holography which allows for real time phase imaging.

We note that in this study we have presented the images with minimal post-processing, the goal being the demonstration of digital holography processes. Other than the overall brightness and contrast adjustments, we have not applied any of the numerous image enhancing techniques that are available which can significantly improve the perceived image quality for biological applications. This will be the subject of a future study.

Digital holography and digital holographic movies are seen to be a useful new tool for biological microscopy, with noteworthy advantages over traditional microscopic techniques for biological imaging. As a non invasive and high resolution measurement technique, digital holography offers enormous advantages to the field of microscopy. The number of potentially useful applications is vast. Such examples include profiling of 3-D morphological structures of diseased cells, such as cancer cells where the nuclei are often enlarged and elasticity is different to healthy cells. Cell changes over time can also be quantitatively assessed; therefore, one may obtain high contrast 4-D movies.

A limitation of the current experimental setup is that the sample is required to be vertical. For observation of living cells over long periods, under favorable conditions for migration and other processes, changes in the setup for axial illumination of the object are necessary. We are currently investigating the possibility of modifying an interference microscope.

Chapter 6

Multi-Wavelength Phase Imaging Digital Holography (MWPIDH)

This chapter introduces a novel optical technique for removal of 2π discontinuities from the phase image. Section 6.1 reviews some of the motivations and advantages of using an optical based approach for phase unwrapping as opposed to that of computational algorithms as demonstrated in the previous chapter. Section 6.2 introduces the mathematical and theoretical principles of using multiple-wavelengths for phase unwrapping. Section 6.3 discusses the experimental setup and procedure for MWPIDH. Section 6.4 presents experimental results. Section 6.5 displays selected frames from MWPIDH movies, which demonstrate the capabilities and future potential for quantitative analysis of dynamic objects or samples such as living cells in real time. Finally section 6.6 concludes this chapter.

6.1 Introduction

Digital holography has been shown to be an effective method for performing high resolution quantitative phase microscopy. However the phase images contain 2π-discontinuities for objects of an optical depth greater than the wavelength, thus providing a difficulty in both interferometry and phase imaging. A conventional approach is to apply any one of a number of phase unwrapping algorithms that have been developed to remove these discontinuities and improve the quality and interpretation of the image. On the other hand, these algorithms require both substantial user intervention and the level of phase noise and phase discontinuity to lie

within strict limits. Furthermore, complex object shapes are often particularly difficult to unwrap correctly and may create phase errors. Another significant disadvantage is that often the unwrapping procedure may be time consuming and hence the capability for real time imaging is lost.

It has been recognized that the phase measurement range can be extended to that of a longer synthetic or beat wavelength by the generation and combination of two phase maps using two or more separate wavelengths. The phase discontinuities which exist in the image are subsequently removed in this process by the effective extension of the axial range. The noise in the final phase profile from the combination of the phase images produced by the multiple wavelengths is equal to that of the single wavelength. With the further extension to more wavelengths it should be possible to achieve even longer axial range with undiminished resolution. The technique can be applied to incoherent interference imaging as well as holographic imaging.

6.2 Theory of Multi-Wavelength Phase Imaging Digital Holography

The principle of multi-wavelength phase imaging is described by referring to Figures 6.1 and 6.2, with numerical values that were used in generating the simulation plots. Suppose the object is a tilted plane of height $h = 5.0\mu m$. Figures 6.1(a) and 6(b) display the phase maps ϕ_1 and ϕ_2 of the tilted object using wavelengths of $\lambda_1 = 532nm$ and $\lambda_2 = 633nm$ respectively. The phase maps contain a 2π discontinuity wherever the height is a multiple of the wavelength. Subtraction of the two phase maps ϕ_1 and ϕ_2 in Figures 6.1(a) and 6(b), results in a new phase map $\phi_{12} = \phi_1 - \phi_2$ as shown in Figure 6.1(c) which has numerous discontinuities of 2π.

By the addition of 2π to the phase map in Figure 6.1(c) wherever $\phi_{12} < 0$ produces a new phase map $\phi_{12}(x) = \phi'_{12} + 2\pi \cdot (\phi'_{12} < 0)$ with a longer range free of discontinuities and extended axial range. The new phase map is equivalent to that of a longer 'beat wavelength'. Using wavelengths of $\lambda_1 = 532nm$ and $\lambda_2 = 633nm$ the new axial range is defined as

$$\Lambda_{12} = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} = 3.33\mu m \tag{6.1}$$

The phase map $\phi_m(x)$ can be converted to the corresponding surface profile Z_m using

$$Z_m = \frac{\lambda_m \phi_m}{2\pi} \tag{6.2}$$

The corresponding surface profile of the object is the coarse map shown in Figure 6.1.(d) which is defined by this new axial range as

$$Z'_{12} = \frac{\Lambda_{12} \phi_{12}(x)}{2\pi} \tag{6.3}$$

By proper choice of two-wavelengths the axial range Λ_{12} can be adjusted to any value that would fit the axial size of the object being imaged.

This technique provides a straightforward and efficient phase imaging method in a wide range of applications. On the other hand a limitation is that any phase noise in each single-wavelength phase map is amplified by a factor equal to the magnification of the wavelengths.

Suppose the single-wavelength phase maps $\phi_m(x)$ contain phase noise of $2\pi\varepsilon_m$, or the corresponding surface profiles $Z_m(x)$ contain a noise level of

$$\varepsilon_m \lambda_m \sim 12nm \tag{6.4}$$

where we have used $\varepsilon_m \sim 2\%$ to simulate the noise. The noise in the difference phase map $\phi_{12}(x)$ is

$$2\pi\varepsilon_{12} = 2\pi(\varepsilon_1 + \varepsilon_2) \tag{6.5}$$

and that in the surface profile $Z'_{12}(x)$ is

$$\varepsilon_{12}\Lambda_{12} \sim 130nm \tag{6.6}$$

The noise has in effect been amplified approximately by a factor of $2\Lambda_{12} / \lambda_m$ as one can see in the coarse map $Z'_{12}(x)$ in Fig.6.1.d) when compared to the single-wavelength phase maps in Fig.6.1 a) or Fig.6.1.b).

The other half of the phase imaging method consists of an algorithm to reduce the noise in $Z'_{12}(x)$ back to the level of the single-wavelength phase maps. This method is described by referring to Figure 6.2. The coarse profile, $Z'_{12}(x)$ shown in Fig.6.1d) can be divided into integer multiples of either of the wavelengths to produce a new coarse profile $Z''_{12}(x)$ as illustrated in Fig. 6.2a). Using for example λ_1, the new coarse profile is defined as

$$Z''_{12}(x) = \text{int}(\frac{Z'_{12}(x)}{\lambda_1}) \cdot \lambda_1 \tag{6.7}$$

Pasting on the single single-wavelength coarse map $Z_1(x)$ to the newly created coarse profile $Z''_{12}(x)$ in Fig 6.2a), produces another coarse profile $Z'''_{12}(x)$ as shown in Fig.6.2b) such that

$$Z'''_{12}(x) = Z''_{12}(x) + Z_1 . \tag{6.8}$$

This almost recovers the surface profile with reduced noise, except at the boundaries of wavelength intervals, where the noise in the single-wavelength phase map causes numerous jumps of size, $\pm\lambda_1$ as displayed in Figure 6.2c). If the noise level is not

excessive, most of the spikes in the can be removed simply by the comparison of $Z_{12}'''(x)$ with the coarse map $Z_{12}'(x)$ to produce Z_b as shown in Figure 6.2d). If the difference is more than λ_1, then λ_1 is either added or subtracted depending on the sign of the difference. Finally Fig.6.2.e) shows the final result, as the fine map Z_{12} where the noise level is approximately the same as that of $Z_1(x)$, the single wavelength profile, at around 12nm. The remaining spikes in the map are due to places where the coarse map is more than one half of λ_1. That is, the maximum noise level for the method to work properly is given approximately by

$$\varepsilon_m < \frac{\lambda_m}{4\Lambda_{12}} \sim 4\% \tag{6.9}$$

The phase-unwrapping technique can be further extended to an iterative procedure of three or more wavelengths as displayed in Figure 6.3. The object simulated this time is a tilted plane of height $h = 10.0\mu m$ and the noise level is approximated to be $\varepsilon_m = 5\%$.

First note that the axial range Λ_{12} can be increased by using closer values of λ_1 and λ_2 as one can see from Equation (6.1). Figures 6.3.a) and 6.3.b), are the surface profiles, Z_1 and Z_2 of the object using wavelengths of $\lambda_1 = 620nm$ and $\lambda_2 = 580nm$ respectively. The axial range from Equation (6.1) is then $\Lambda_{12} = 8.99\mu m$. From Equation (6.9) as described above, for two-wavelength phase imaging the noise limit is found to be

$$\varepsilon_m < \frac{\lambda_m}{4\Lambda_{12}} \sim 1.7\% \tag{6.10}$$

The noise limit has been effectively reduced because of the larger Λ_{12}. Figs.6.3.c) and 6.3.d) show the coarse map, $Z'_{12}(x)$ and the fine map, $Z_{12}(x)$ generated from the phase maps of λ_1 and λ_2. The noise in the coarse map $Z'_{12}(x)$ is displayed in Fig.6.3 j) and corresponds to

$$(\varepsilon_1 + \varepsilon_2)\Lambda_{12} \sim 900nm \tag{6.11}$$

which is much larger than half of λ_1, and therefore the fine map $Z_{12}(x)$ has too many λ_1 spikes in it, as one can see from Fig.6.3k). With the addition of a third wavelength at $\lambda_3 = 500nm$, whose surface profile is displayed in Fig.6.3e), new coarse maps $Z'_{13}(x)$ and $Z'_{23}(x)$ are generated as shown in Figures 6.3f) and 6.3g) with beat wavelengths of $\Lambda_{13} = 2.58\mu m$ and $\Lambda_{23} = 3.63\mu m$. Combining two coarse maps coarse maps $Z'_{13}(x)$ and $Z'_{23}(x)$ using the coarse map procedure, produces a new coarse map Z'_{13-23} as seen in Figure 6.3h). Instead of pasting Z_1 to Z'_{12} which is too noisy to produce a useful result as shown in Fig.6.3k), we go through the following two steps. First paste Z'_{13} onto $Z'_{13-23}(\equiv Z'_{12})$, to obtain an 'intermediate fine map', Z''_{13-23}, as shown in Fig.6.3.h). The noise level in Z''_{13-23} as shown in Fig.6.3 l) is

$$Z'_{13} = (\varepsilon_1 + \varepsilon_2)\Lambda_{13} = (0.05 + 0.05)2.58\mu m \sim 260nm \tag{6.11}$$

This is now smaller than one half of λ_1 and we can paste Z_1 to obtain the final fine map, Fig.6.3.i). The noise in this map is approximately the same as that of Z_1, $\varepsilon_1\lambda_1 \sim 31nm$. The maximum noise level ε_m in the single-wavelength phase map for the three-wavelength phase imaging to work is given by the smaller of

$$\frac{\Lambda_{13}}{4\Lambda_{12}} \sim 7\% \text{ or } \frac{\lambda_1}{4\Lambda_{13}} \sim 6\% \tag{6.12}$$

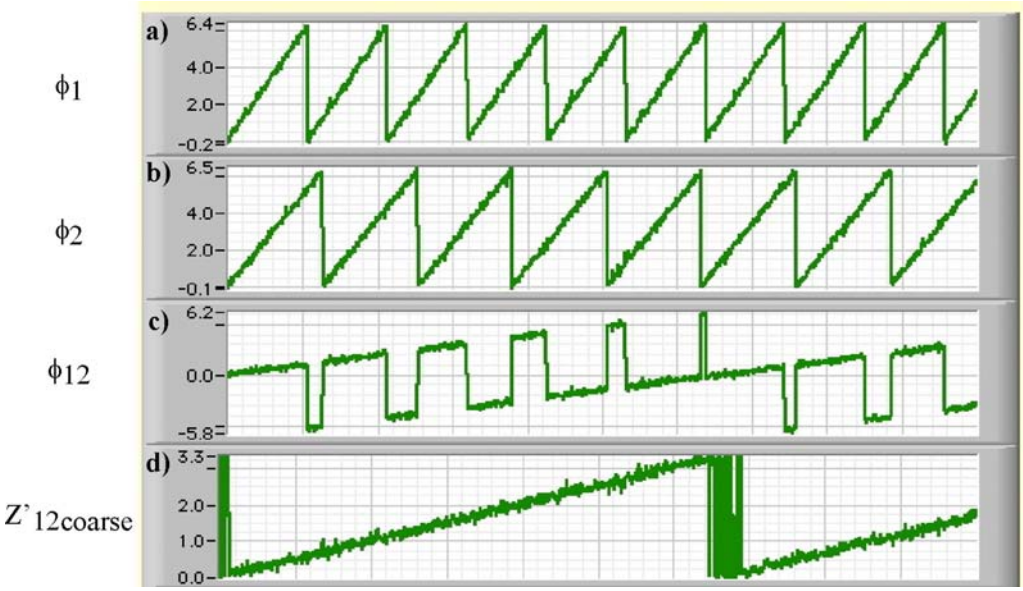


Figure 6.1: Simulation of two-wavelength phase imaging digital holography. **a)** Actual height profile of object with a tilted plane of axial height $h = 5.0\mu m$; **b)** height profile $Z_1(x)$ of a $5\mu m$ high incline derived from phase $\phi_1(x)$ of $\lambda_1 = 532nm$; **c)** $Z_2(x)$ derived from phase $\phi_2(x)$ of $\lambda_2 = 633nm$; **d)** phase map $\phi_1(x)$ of $\lambda_1 = 532nm$; **e)** phase map $\phi_2(x)$ of $\lambda_2 = 633nm$; **f)** difference phase map $\phi_{12}'(x) = \phi_1 - \phi_2$; **g)** coarse map $Z_{12}'(x)$, with beat wavelength $\Lambda_{12} = 3.33\mu m$.

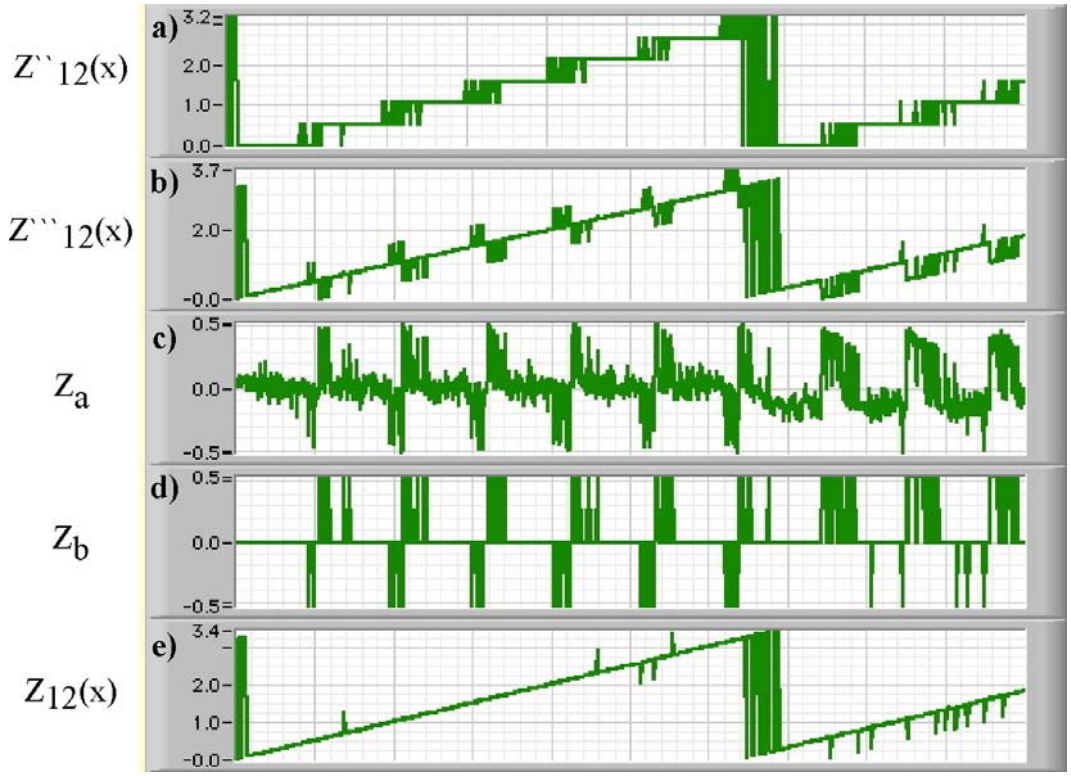


Figure 6.2: Simulation of two-wavelength phase imaging digital holography for noise removal. **a)** $Z''_{12}(x)$, where $Z'_{12}(x)$ is divided into integer multiples of λ_1; **b)** $Z'''_{12}(x)$, where $Z_1(x)$ is pasted on $Z''_{12}(x)$; **c)** Z_a, derived from comparing $Z'''_{12}(x)$ with the coarse map $Z'_{12}(x)$; **d)** Z_b ,resulting from addition or subtraction of λ_1 in Z_a to remove spikes; **e)** the fine map, $Z_{12}(x)$

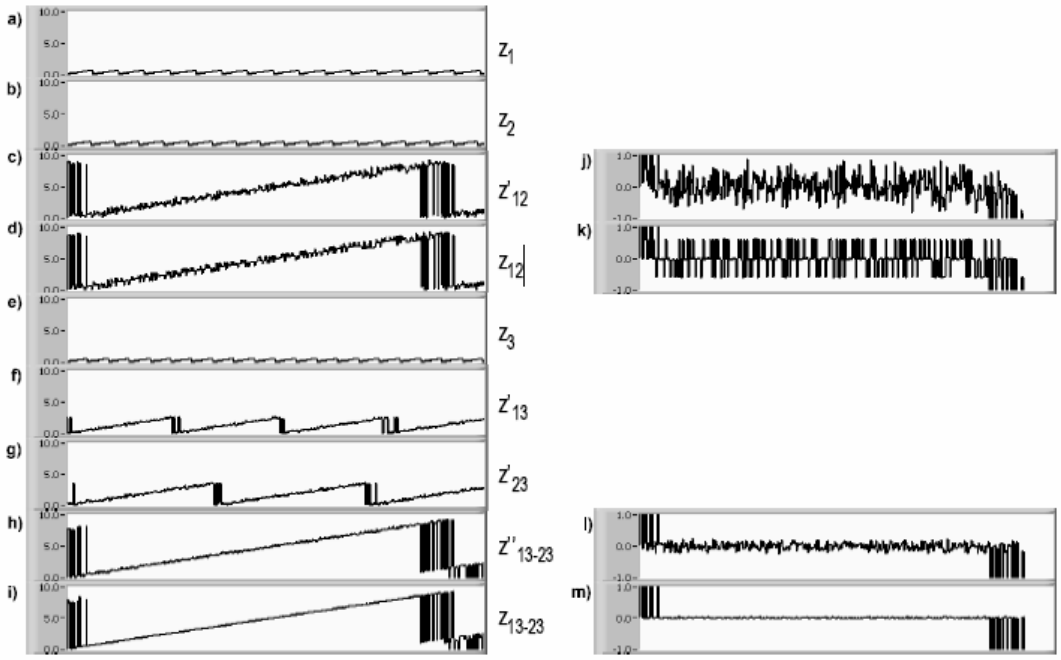


Figure 6.3: Simulation of three-wavelength phase imaging digital holography. **a)** height profile $Z_1(x)$ of a $10\mu m$ high incline derived from phase $\phi_1(x)$ of $\lambda_1 = 620nm$; **b)** $Z_2(x)$ derived from phase $\phi_2(x)$ of $\lambda_2 = 580nm$; **c)** coarse map $Z'_{12}(x)$, with beat wavelength $\Lambda_{12} = 8.99\mu m$; **d)** fine map, $Z_{12}(x)$; **e)** $Z_3(x)$ derived from phase $\phi_3(x)$ of $\lambda_2 = 500nm$; **f)** coarse map $Z'_{13}(x)$, with beat wavelength $\Lambda_{12} = 2.58\mu m$; **g)** coarse map $Z'_{23}(x)$, with beat wavelength $\Lambda_{12} = 3.63\mu m$; **h)** 'intermediate fine map Z''_{13-23}, where $Z'_{13}(x)$ is pasted on $Z'_{13-23}(x) = Z'_{12}(x)$; **i)** final fine map $Z_{13-23}(x)$, where $Z_1(x)$ is pasted onto Z''_{13-23}; **j)** Noise in c), $Z'_{12}(x) - Z(x)$; **k)** Noise in d), $Z_{12}(x) - Z(x)$; **l)** Noise in h), $Z''_{13-23}(x) - Z(x)$; **m)** Noise in i), $Z_{13-23}(x) - Z(x)$

6.3 Experimental Setup for MWPIDH

Figure 6.4 depicts the experimental setup for MWPIDH. A Continuous Wave (CW) HeNe laser (633nm) is added to the experimental setup in Figure 5.2 to obtain the second wavelength for the MWPIDH process. The beam from the HeNe laser is combined with the beam from the Nd:YAG laser (532nm) using BS1 to form a two color beam which is then split into an object and a reference beam. The beams require highly accurate optical alignment in order for them to traverse the same exact optical path.

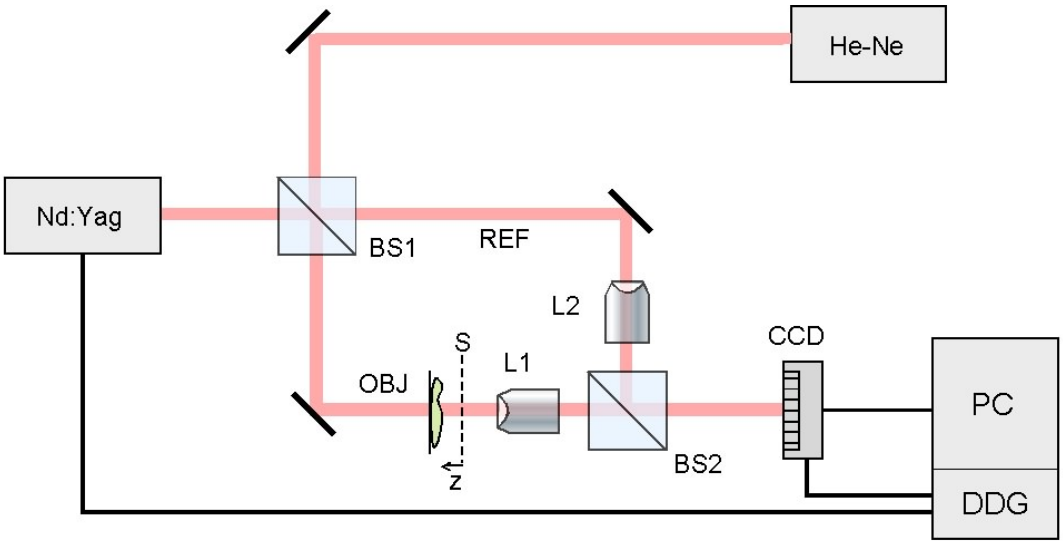


Figure 6.4: Experiment setup for MWPIDH

6.4 Two-Wavelength Phase Imaging Digital Holography Results

Figure 6.5 presents holographic images of a group 6 element of the resolution target. The procedure for MWPIDH image capture and processing is as follows. First a hologram is captured using only the beam emitted from the Nd:YAG laser at the green wavelength, $\lambda_1 = 532nm$, with the HeNe laser beam simultaneously blocked off before BS1, so it does not enter into the holographic optical system. Once this hologram is captured, the Nd:YAG beam is then blocked off and a second hologram is captured using the beam from the HeNe laser at the red wavelength, $\lambda_2 = 633nm$. The two holograms are then processed separately to obtain their respective reconstructed amplitude and phase images.

The top row of Figure 6.5 displays the hologram (a), amplitude (b) and phase (c) images by the green wavelength. The bottom row of Figure 6.5 shows the hologram (d), amplitude (e) and phase (f) images by the red wavelength. The image size is 60x60μm^2, recorded at 452 x 452 pixels. The 2π discontinuities are present in both of the phase images. One also notices that the quality of both the hologram and the reconstructed amplitude and phase images in Fig.6.5 (d-f) by the red wavelength is significantly poorer than that those by the green wavelength in Fig.6.5 (a-c). In the amplitude image in Fig.6.5 (e) for example, one observes faint horizontal lines suggesting a disturbing interference source.

Using the technique of combining the two wavelength phase maps as outlined in Section 6.2, creates a longer range free of 2π discontinuities with extended axial range. The fine map is displayed in Fig. 6.5(h). Some discontinuities still remain in the image as

the bars contain a large fluctuation of phase due to the small amount of signal that is obtained from these areas.

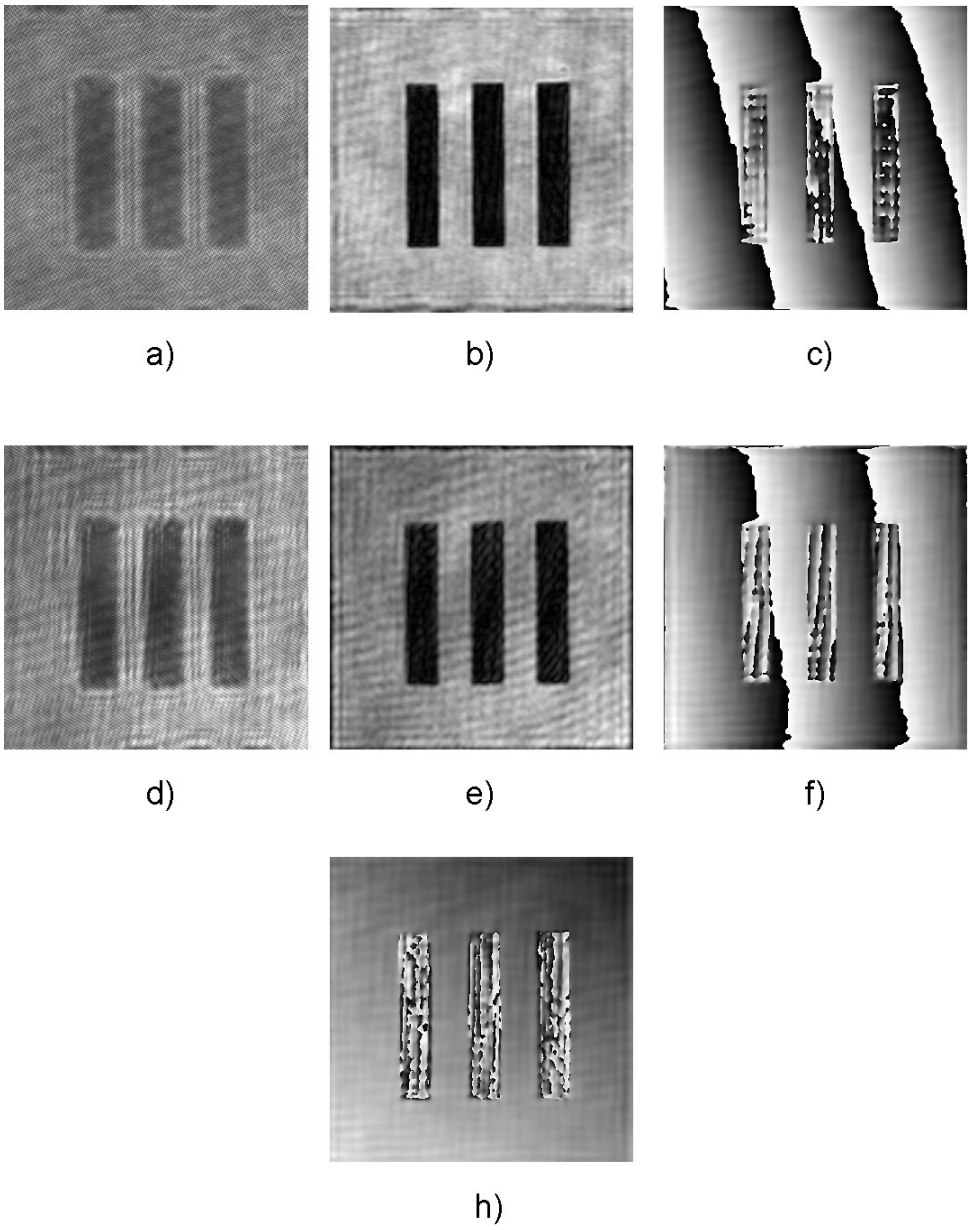


Figure 6.5: (a) The hologram, (b) amplitude, (c) and wrapped phase map by the green wavelength λ_1 = 0.532μm and (d) the hologram, (e) amplitude, (f) wrapped phase map by the red wavelength at λ_2 = 0.633μm; (h) is the fine map obtained by the phase maps shown in (c) and (f).
(Area =60μm x 60μm^2, Pixels = 452 x 452)

Another example of two-wavelength phase imaging digital holography is illustrated in Figure 6.6. The combination of the phase maps by the green wavelength, Fig. 6.6(a), and the red wavelength, Fig.6.6(b), produces a fine map with a new larger beat wavelength $\Lambda_{12} = 3.33\mu$m, Fig.6.6(c). The software implemented phase unwrapping algorithm in Fig.6.6(d) has a defect that propagates beyond the noisy regions.

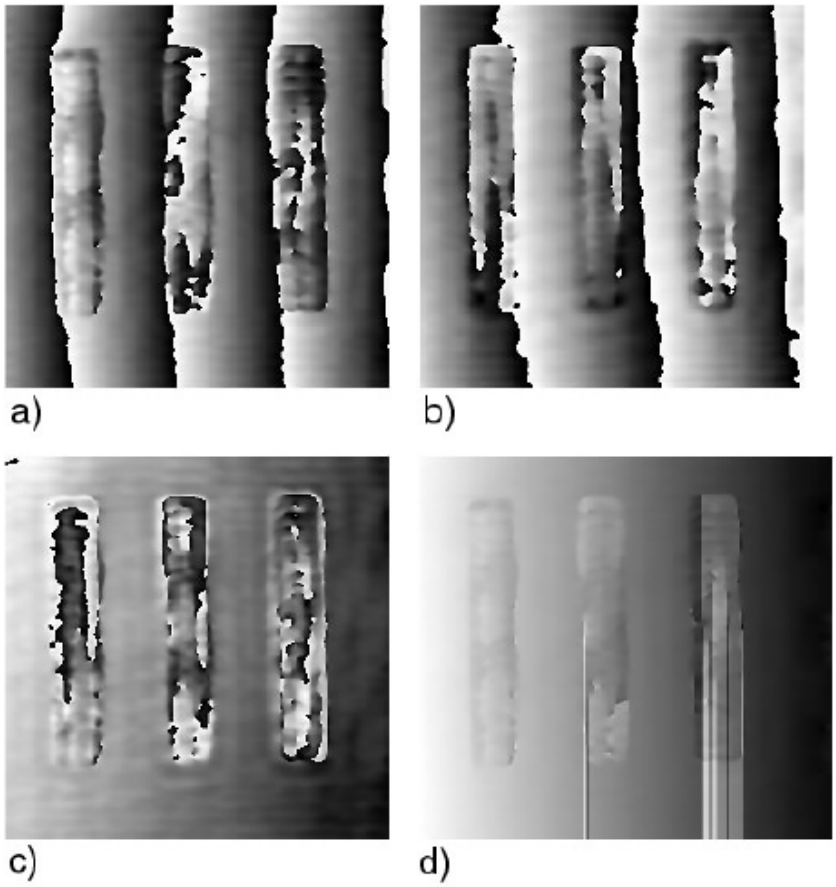


Figure 6.6: (a) The wrapped phase map reconstructed from the hologram at the green wavelength $\lambda_1 = 0.532\mu$m and (b) the red wavelength at $\lambda_2 = 0.633\mu$m; (c) the fine map obtained by the phase maps shown in (a) and (b); (d) is the unwrapped phase map by a software program.

Figure 6.7 displays two-wavelength phase imaging digital holography of onion cells. The image size is 250 x 250μm^2, recorded at 452 x 452 pixels. The panels display the (a) hologram, (b) reconstructed amplitude image, (c) and reconstructed phase image by the green wavelength and the (d) hologram, (e) reconstructed amplitude image, (f) and reconstructed phase image, by the red wavelength. (g) is the fine map produced by combination of the phase maps. The onion cell walls are easily visible in the reconstruction and the fine phase map is of a good quality. Of note again here is the red wavelength reconstruction is significantly degraded compared to the green wavelength reconstruction. This consequently causes some image degradation in the fine map.

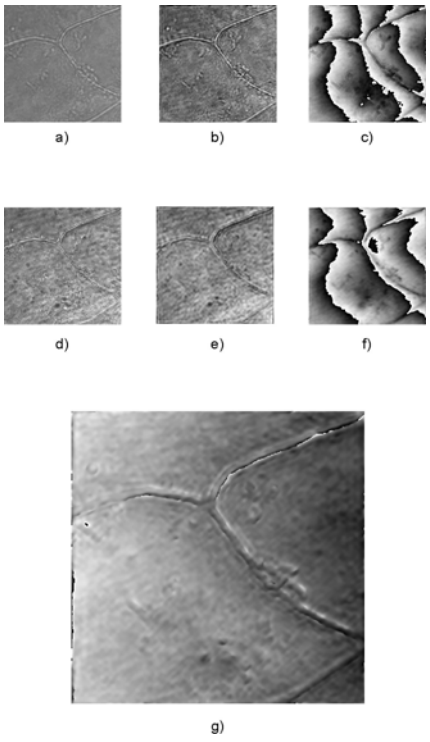


Figure 6.7: (a) The hologram, (b) amplitude, (c) and wrapped phase map by the green wavelength λ_1 = 0.532μm and (d) the hologram, (e) amplitude, (f) wrapped phase map by the red wavelength at λ_2 = 0.633μm; (g) is the fine map obtained by the phase maps shown in (c) and (f).

6.5 MWPIDH Movies

Besides the numerous difficulties involved in phase unwrapping with computation algorithms, a significant disadvantage of their use is that most of them are time consuming and therefore the capability to perform real time or close to real time processing is lost. On the other hand the procedure of MWPIDH for phase unwrapping can be performed in real time. Phase images reconstructed from holograms recorded at two different wavelengths are simply combined in a simple way to produce a fine phase map without the usual 2π discontinuities.

For a static object we capture holograms successively for each individual wavelength. However to produce multi-wavelength holographic movies of dynamic changes requires that the holograms corresponding to each wavelength be captured simultaneously on the CCD camera. The problem then becomes how to separate the individual holograms corresponding to each wavelength from the two color beam.

An effective method of undertaking this is to use the spectral property of the CCD camera. The CCD is able to capture images in color, i.e. 24-bit, Red, Green and Blue (RGB), each color corresponding to 8-bit and the separate planes can be extracted. Figure 6.8 shows the relative response of the CCD with wavelength. By choosing wavelengths near the CCD's peak response for each RGB color, one effectively maximizes the response to that particular color plane while simultaneously minimizing the response to the other colors. For our specific example at approximately $\lambda_1 = 532nm$, one can see that the relative response to the green color plane is approximately the highest, while it is fairly low for red and blue and the same is true for our second selected wavelength at $\lambda_2 = 633nm$.

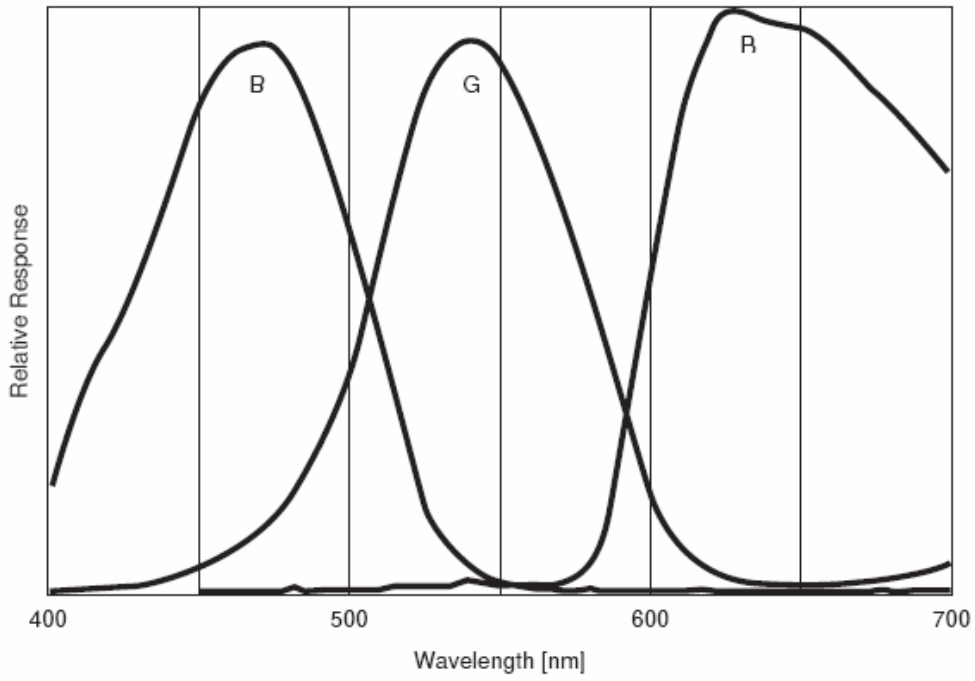


Figure 6.8: Spectral characteristics of the color CCD camera.

After digital capture of the two color beam holographic image corresponding to the selected wavelengths $\lambda_1 = 532nm$ and $\lambda_2 = 633nm$ onto the CCD camera, the red, green and blue planes of the captured image are extracted. One then obtains the holograms specific to each wavelength. The blue extraction is subtracted as noise from both the red and green holograms. The two holograms are then processed separately to obtain the phase images. These are subsequently combined as earlier outlined in the theory in order to remove 2π discontinuities and extend the axial range.

Figure 6.9 shows a time series of selected amplitude images from an amplitude movie illustrating the motion of a rotifer through water. The amplitude images are reconstructed from holograms created by the green wavelength and the image area is 70μm x 70μm^2, recorded with 460 x 460 pixels. The time interval between each image is t = 800ms. The frame rate on the camera for image capture is 25 frames/sec and the holograms are reconstructed and combined in 300ms.

Rotifers generally have a thick cylindrical body. So, when looking at them from above, there will always be a problem with depth of focus with conventional light microscopes. Furthermore they cannot be fixed and flattened by the pressure of a cover-glass because they would disintegrate and die at once. Digital holography as a non-invasive technique with three-dimensional information overcomes these limitations.

The rotifer's cilia are used to trap food and to move around through the water. A feature of interest in the series of images in the figure is the view of the stomach which appears somewhat indented in the images. Some food particles are also visible moving throughout the body.

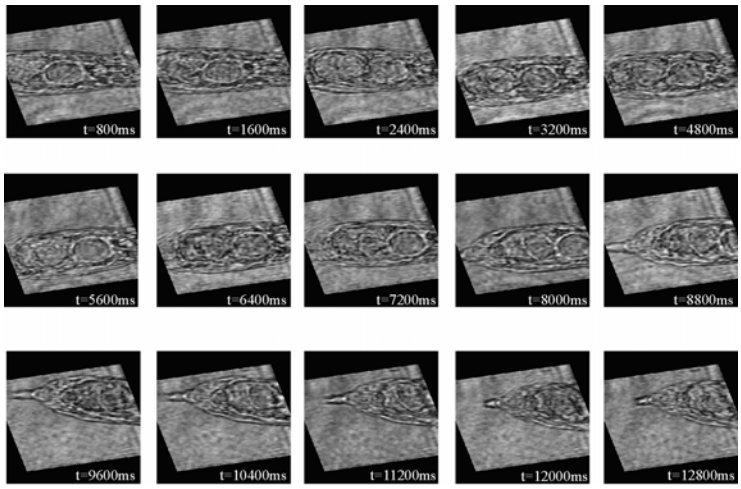


Figure 6.9: Time-Series of reconstructed amplitude images showing a rotifer moving across the field of view.
(Area =70μm x 70μm^2, Pixels = 460 x 460, z = 0.253mm)

Figure 6.10 presents the reconstructed time series of phase images reconstructed simultaneously along with the amplitude images in Figure 6.9 by the green wavelength. The 2π discontinuities are seen to be present in the image.

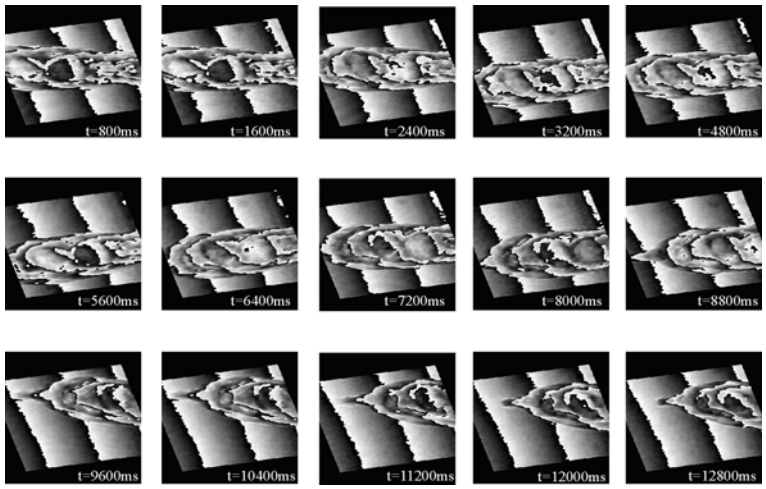


Figure 6.10: Time-Series of reconstructed phase images showing a rotifer moving across the field of view
(Area =70μm x 70μm^2, Pixels = 460 x 460, z = 0.253mm)

Figure 6.11 shows the corresponding time series of unwrapped fine map phase images created by the combination of phase maps reconstructed from the hologram produced by the green and red wavelengths. There are some discontinuities still present in the images where the object thickness is greater than the new axial range at 3.3μm

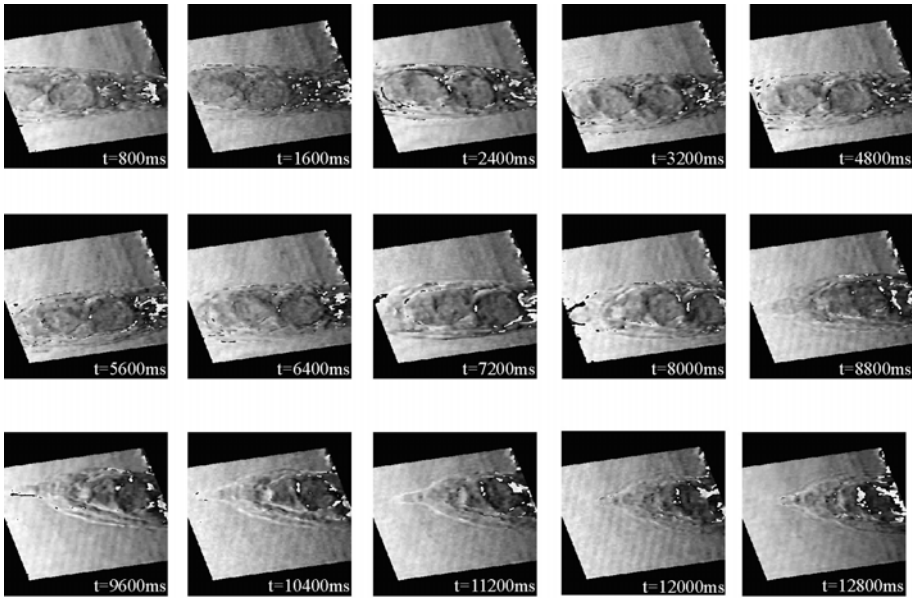


Figure 6.11: Time-Series of unwrapped phase images by MWPIDH showing a rotifer moving across the field of view.
(Area =70μm x 70μm^2, Pixels = 460 x 460)

Figure 6.12 shows a time series of selected amplitude images from an amplitude movie showing a high magnification view of a paramecium feeding in water. The amplitude images are reconstructed from holograms created by the green wavelength and the image area is 60μm x 60μm^2, recorded with 442 x 442 pixels. The time interval between each image is t = 800ms. In the images there are some slight variations between the frames but nothing particularly striking, in fact the interior of the paramecium looks pretty noisy and dense.

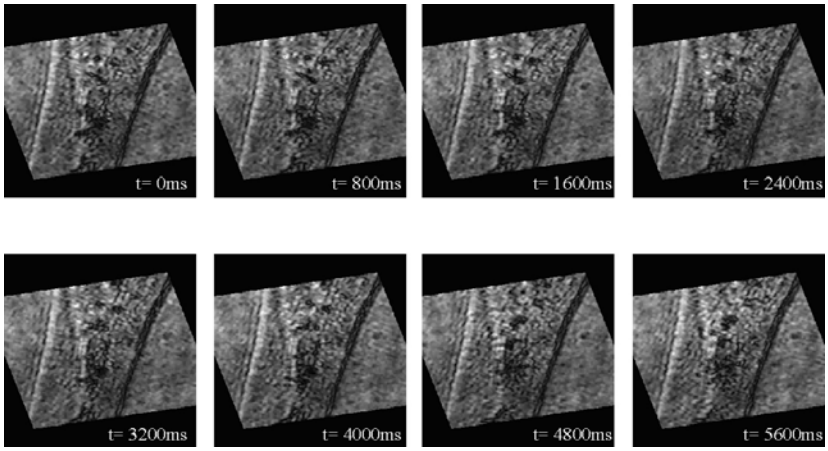


Figure 6.12: Time-Series of reconstructed amplitude images of a paramecium feeding in pond water.
(Area =60μm x 60μm^2, Pixels = 442 x 442)

Figure 6.13 presents the reconstructed time series of phase images reconstructed simultaneously along with the amplitude images in Figure 6.12 by the green wavelength. The 2π discontinuities are seen to be present in the image.

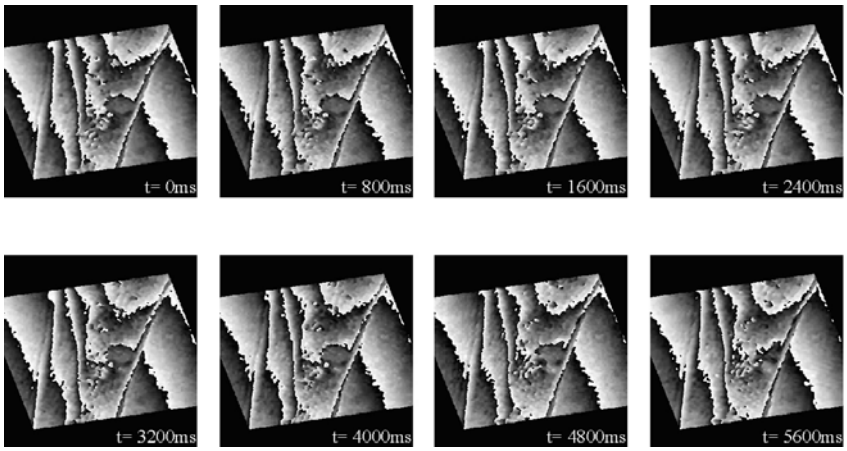


Figure 6.13: Time-Series of reconstructed phase images of a paramecium feeding in pond water.
(Area =60μm x 60μm^2, Pixels = 442 x 442)

Figure 6.14 presents the images created by the combination of phase maps by the green and red wavelengths. The contractile vacuole (highlighted in the second frame) is shown to contract as it pumps out water. This phenomenon cannot be observed in the amplitude images in Fig. 6.12. There are some discontinuities still present in the images where the paramecium's physical thickness is greater than the axial range.

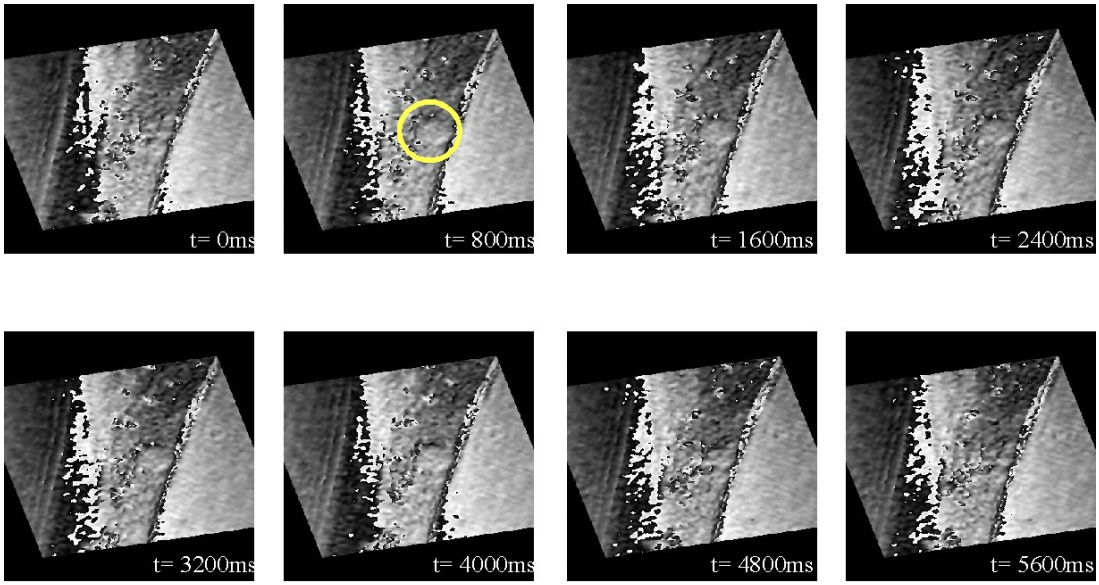


Figure 6.14: Time-Series of unwrapped phase images by MWPIDH showing a rotifer moving across the field of view.

6.5 Conclusions and Discussion

MWPIDH offers a simple and effective way of removing the discontinuities in the phase image. The axial range can be changed to a pre-determined value depending on the wavelengths selected for the procedure. Therefore one can pre-determine the axial range required for a specific object. If both a short wavelength and a long wavelength are used, it is possible to examine a larger object and still have high resolution. The object

would first be imaged with the long wavelength; this would provide a range of the object's position in space. Then the object would be imaged with the shorter wavelength to obtain higher resolution. The idea is to use the longer wavelength to resolve the degeneracy's of the shorter wavelength. The longer wavelength is used to give information about which 2π discontinuity the object point is on. This requires that the uncertainty in distance given by the longer wavelength be less than the shorter wavelength. One must take care when choosing the beat wavelength. If it is too long, then the uncertainty in distance will be greater than the single frequency wavelength, causing the system to inaccurately determine the fringe number.

Conventional computational algorithms for phase unwrapping are limited by a number of factors such as the object shape, noise requirements, significant user intervention and as a fairly time consuming process, real time imaging is a difficulty. This is a serious limitation for the continued growth of digital holography for microscopy. MWPIDH is seen to overcome these problems. The technique can be performed in close to real time and is quantitative; therefore it is an ideal method for performing quantitative phase-contrast microscopy.

The presented results of test and biological samples show that the methodology for MWPIDH works well. However, the results also confirm that the red wavelength reconstruction is of an inferior quality when compared to that of the green wavelength reconstruction and this causes image degradation of the final fine map phase image.

Analysis of other recording wavelengths and CCD cameras will require further investigation and the use of the angular spectrum should be invaluable for this process.

Chapter 7

Conclusion and Future Work

In this research we have successfully demonstrated the improvements made in techniques of digital holography. Experiments conducted in both digital Gabor and off-axis holographic configurations have revealed both of these methods to be highly effective for applications in biological microscopy.

In Gabor holography there is a limitation in the method due to specific requirements of the object's physical characteristics. It is required to be either highly transparent, or small and opaque. This condition is necessary in order to maximize the degree of transparency in the overall image. If this object requirement is not satisfied, then degradation in the reconstructed image occurs. Furthermore, phase information is not available in Gabor holography due to the superposition of the holographic terms.

While these limitations tend to produce reconstructed images which are generally not of a high quality, the advantages of this technique when compared to that of conventional bright-field microscopic techniques are observed in the 3-D reconstruction capability. From a single hologram, numerical reconstruction can be performed in numerous axial planes. By contrast, in conventional bright-field microscopy the information is only available for one specific axial plane. Therefore, digital Gabor holography is able to track any object moving in depth over time within a certain field of view, so that it remains in focus throughout the entire time-frame. If one is interested

specifically in the trajectory, speed, motion, pattern, or any dynamic analysis of the object, digital Gabor holography is a particularly attractive technique.

One may also produce three dimensional optical profiles of motion by stacking numerous 2D-reconstructions of selected depths. For this purpose, the use of algorithms for auto-focusing will also be of enormous advantage for the application of this method in the creation of holographic movies.

Off-Axis digital holography enables simultaneous reconstruction of both amplitude and phase information of the object. The *dc* and holographic terms can be spatially separated due to the introduction of an angle between the object and reference beams. However, due to the low spatial resolution of CCD cameras, this angle is limited. This in turn limits the distance of spatial separation between the terms. There are a number of numerical reconstruction algorithms available for use in digital holography. In this thesis we have reviewed each of the methods, and their application to the reconstruction of a resolution test target and a biological sample. The angular spectrum algorithm is shown to be a particularly advantageous method for holographic reconstruction, since it makes it possible to filter the undesired spatial frequencies while keeping the contributions from the interference terms intact. Also, a significant advantage of the angular spectrum algorithm is the absence of a minimum reconstruction distance, which is a strict requirement in the more commonly used Huygens convolution and Fresnel transform reconstruction methods. Furthermore, in this thesis we have demonstrated the application of the angular spectrum method for obtaining high quality images of biological objects with quantitative phase analysis. The use of the angular

spectrum is seen to solve the problems that have so far prevented the widespread use of digital holography for microscopy.

Digital holography provides amplitude images, as obtained by a conventional optical microscope. The advantage over conventional bright-field microscopes is in the 3D reconstruction, which overcomes the limited depth of focus. However, the significant advantage of off-axis digital holography is actually seen in the phase information, which is simultaneously reconstructed along with the amplitude information.

Conventional phase-contrast microscopy techniques are unable to yield quantitative phase information and have other disadvantages associated with the image quality, such as the halo effect. Digital holography, on the other hand, offers quantitative phase information, and, therefore, reveals the 3D topography of an object's optical profile with a resolution at the nanometer scale along the optical axis.

Quantitative holographic movies can also be created for analysis of moving objects such as living cells. We have demonstrated, as shown in the experimental results, that digital holography is well suitable for living cell analysis, and contains many advantages over traditionally applied methods. However, the current experimental setup is required to be modified in order for horizontal mounting of the sample.

The phase image needs to be unwrapped to improve interpretation of the image. A conventional approach is to apply any one of a number of computational phase unwrapping techniques. However, there are a number of disadvantages associated with their use. The main constraint for live sample imaging is the time taken for the unwrapping of each phase image, which essentially takes away the ability for a biologist to view the motion in real time or close to real time. Extensive investigation of other

phase unwrapping techniques is required which may produce good quality unwrapping in less time than the present flood-fill algorithm.

We show the use of a novel, optically based approach for unwrapping the phase by use of multiple wavelengths. This technique offers a convenient and attractive alternative to using a software-based phase unwrapping algorithm. The advantage of the multi-wavelength imaging technique is clearly demonstrated when un-wrapping an object that does not fulfill the strict requirements of the conventional unwrapping algorithms. As shown in the experimental results, this method is effective and able to perform unwrapping in close to real time. Although the reconstructed image is shown to be of a fairly good quality with the current set up, it could be significantly enhanced by improving the quality of the red reconstructed image from the red hologram. We are currently performing investigations to pinpoint the cause of the poor outcome of the reconstructed images using the red wavelength. The angular spectrum has been invaluable in so far eliminating the possible causes. However, the reason for this degradation is as yet still unknown, although we suspect that the camera itself may be the main cause. This will be the subject of future investigations.

References

[1] Gabor D, “A new microscope principle”, *Nature* 1948 **161**: 777-8.

[2] Leith E and Upatnieks J, “Reconstructed wavefronts and communication theory”, J. *Opt. Soc. Am*. 1962, **52**: 1123.

[3] Powell R.L, Stetson K.A, "Interferometric vibration analysis by wavefront reconstruction," *J. Opt. Soc. Am.* 1965, **55**: 1593-1598.

[4] Goodman J.W and Lawrence R.W, “Digital image formation from electronically detected holograms,” *Applied Physics Letters* 1967, **11**: 77-79.

[5] Kronrod M.A, Merzlyakov N.S, Yaroslavsky L.P, “Reconstruction of a Hologram. with a Computer,” *Soviet Physics-Technical Physics* 1972, **17(2)**: 333-334.

[6] Schnars U, Jueptner WP, “Direct recording of holograms by a CCD target and numerical reconstruction,” *Appl Opt.* 1994, **33**: 179-181.

[7] Schnars U, Jueptner WP, “Digital recording and numerical reconstruction of holograms,” *Meas Sci Technol* 2002, **13**: R85-R101.

[8] Cuche E, Bevilacqua F, Depeursinge C, “Digital holography for quantitative phase-contrast imaging,” *Opt Lett* 1999, **24**: 291-293.

[9] Ferraro P, De Nicola S, Coppola G, Finizio A, Alfieri D, Pierattini G, “Controlling image size as a function of distance and wavelength in Fresnel-transform reconstruction of digital holograms,” *Opt Lett* 2004, **29**: 854-856.

[10] Ferraro P, Grilli S, Alfieri D, De Nicola S, Finizio A, Pierattini G, Javidi B, Coppola G, Striano V, “Extended focused image in microscopy by digital holography,” *Opt Exp* 2005, **13**: 6738-6749.

[11] Palacios F, Ricardo J, Palacios D, Gonalves E, Valin J, De Souza R, “3D image reconstruction of transparent microscopic objects using digital holography,” *Opt Commun* 2005, **248**: 41-50.

[12] Cuche E, Marquet P, Depeursinge C: “Simultaneous amplitude contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms,” *Appl Opt* 1999, **38**: 6994-7001.

[13] Ferraro P, De Nicola S, Finizio A, Coppola G, Grilli S, Magro C, Pierattini G, “Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging,” *Appl Opt* 2003, **42**: 1938-1946.

[14] Stadelmaier A, Massig J, “Compensation of lens aberrations in digital holography,” *Opt Lett* 2000, **25:** 1630-1632.

[15] Osten W, Baumbach T, Jueptner WP, “Comparative digital holography,” *Opt Lett* 2002, **27:** 1764-1766.

[16] Xu M L, Peng X, Miao J and Asundi A, “Studies of digital microscopic holography with applications to microstructure testing,” *Appl. Opt.* 2001, **40:** 5046–51.

[17] Pedrini G, Tiziani H J, “Quantitative evaluation of two-dimensional dynamic deformations using digital holography,” *Opt. Laser Technol.* 1997, **29:** 249–56.

[18] Picart P, Leval J, Mounier D and Gougeon S, “Some opportunities for vibration analysis with time averaging in digital Fresnel holography,” *Appl. Opt.* 2005, **44:** 337–343.

[19]. Alexander, "Development of integral holographic motion pictures," *SPIE* Art and design, 1995, **2333**: 187-197.

[20] Jacquot M, Sandoz P, Tribillon G, “High resolution digital holography,” *Opt Commun* 2001, **190**: 87-94.

[21] Xu W, Jericho M H, Meinertzhagen I A and Kreuzer H J, "Digital in-line holography for biological applications," *Proc. Natl Acad. Sci. USA* 2001, **98:** 11301–5.

[22] Popescu G, Delflores L P, Vaughan J C, Badizadegan K, Iwai H, Dasari R and Feld M S, "Fourier phase microscopy for investigation of biological structures and Dynamics," *Opt. Lett.* 2004, **29:** 2503–5.

[23] Marquet P, Rappaz B, Magistretti P J, Cuche E, Emery Y, Colomb T and Depeursinge C, "Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwevelength axial accuracy," *Opt. Lett.* 2005, **30:** 468–70.

[24] Carl D, Kemper B, Wernicke G and Von Bally G, "Parameter-optimized digital holographic microscope for high-resolution living-cell analysis," *Appl. Opt.* 2004, **43:** 6536–44.

[25] Haddad W.S, Cullen D, Solem J.C, Longworth J.W, McPherson A, Boyer K, Rhodes C.K, "Fourier-transform holographic microscopy," *Appl Opt* 1992, **31**: 4973-4978.

[26] Yamaguchi I, Kato J, Ohta S, Mizuno J, "Image formation in phase-shifting digital holography and applications to microscopy," *Appl Opt* 2001, **40**: 6177-6186.

[27] Sun H, Player M, Watson J, Hendry D, Perkins R, Gust G, Paterson D, "The use of digital/electronic holography for biological applications," *J Opt A-Pure Appl Opt* 2005, **7**: S399-S407.

[28] Mann C, Yu L, Lo C, Kim MK, "High-resolution quantitative phase-contrast microscopy by digital holography," *Opt Exp* 2005, **13**: 8693-8698.

[29] Ma L, Wang H, Li Y, Jin H, "Numerical reconstruction of digital holograms for three-dimensional shape measurement," *J Opt A-Pure Appl Opt* 2004, **6**: 396-400.

[30] Grilli S, Ferraro P, De Nicola S, Finizio A, Pierattini G, Meucci R, "Whole optical wavefields reconstruction by digital holography," *Opt Exp* 2001, **9**: 294-302.

[31] Dakoff A, Gass J, Kim M.K: "Microscopic three-dimensional imaging by digital interference," *Journal of Electronic Imaging* 2003, **12**: 643-647.

[32] Yu L, Kim M.K: "Wavelength-scanning digital interference holography for tomographic 3D imaging using the angular spectrum method," *Opt Lett* 2005, **30**: 2092.

[33] Servin M., Marroquin J.L, Malacara D, and Cuevas F.J, "Phase unwrapping with a regularized phase-tracking system," *Appl. Opt.* 1998, **37**: 1917-23.

[34] Schofield M.A. and Zhu Y, "Fast phase unwrapping algorithm for interferometric

Applications," *Opt. Lett.* 2003, **28**: 1194.

[35] Cheng Y.Y and Wyant J.C, "Two-wavelength phase shifting interferometry," *Appl. Opt.* 1984, **23**: 4539-43.

[36] Creath K, "Step height measurement using two-wavelength phase-shifting interferometry," *Appl. Opt.* 1987, **26**: 2810-6.

[37] Gass J, Dakoff A, Kim M.K, "Phase imaging without 2π ambiguity by multiple-wavelength digital holography," *Opt Lett* 2003, **28**: 1141-1143.

[38] Mann C. J, Yu L, and Kim M.K**,** "Movies of cellular and sub-cellular motion by digital holographic microscopy," *BioMedical Engineering OnLine* 2006, **5**:21-33.

[39] Backman V. *et al*, "Detection of pre-invasive cancer cells," *Nature* 2000, **406**: 35-36.

[40] T. M. Kreis, M. Adams and W. P. O. Jueptner, "Methods of digital holography: a comparison," *Proc. SPIE* 1997, **3098**: 224-33.

[41] Demetrakopoulos TH, Mittra R, "Digital and optical reconstruction of images from sub-optical diffraction patterns," *Appl Opt* 1974, **13**: 665-670.

[42] Nugent K. A., "Twin-image elimination in Gabor holography," *Opt. Commun.* 1990, **78** 293-299.

[43] Gustafsson M *et al*., "High-resolution digital transmission microscopy: a Fourier holography approach," *Opt. and Lasers in Eng*. 2004, **41**: 553-563.

[44] Liu C., Liu Z.G, and Bo F. *et al.*, "Super-resolution digital holographic imaging method," *Appl. Phys. Lett*. 2002, **81**: 3143-3145.

[45] Liu G and Scott P. D, "Phase retrieval and twin-image elimination for in-line Fresnel holograms," *J. Opt. Soc. Am. A* 1987, **4**: 159-165.

[46] S. Lai, B. Kemper, G. v. Bally, "Off-axis reconstructions of in-line holograms for twin-image elimination," *Opt. Commun.* 1999. **169**: 37-43 (1999).

[47] Yamaguchi I, Kato J., Matsuzaki H., Measurement of surface shape and deformation by phase-shifting image digital holography," Proceedings of SPIE 2002, **4778**: 251-256.

[48] Xu W, Jericho M.H, Meinertzhagen I.A, Kreuzer H.J, "Digital in-line holography of microspheres". Applied Optics 41, 5367-5375, 2002.

[49] Carolina Biological Supply Handbook. (2002).

[50]. Torok P and Kao F.J, *Optical imaging and microscopy* (Springer-Verlag, 2003).

[51]. Barty A, Nugent K. A, Paganin D. and Roberts A., "Quantitative optical phase microscopy," *Opt. Lett.* 1998, **23**: 817-9.

[52]. Rylander C. G, Dave D, Akkin T, Milner T. E, Diller K. R. and Welch A. J., "Quantitative phase-contrast imaging of cells with phase-sensitive optical coherence microscopy," *Opt. Lett.* 2004, **29**: 1509-11.

[53] Van den Doel L.R and Van Vliet L.J, **"**Temporal phase-unwrapping algorithm for dynamic interference pattern analysis in interference-contrast microscopy," *App. Opt.* 2001, **40**: 4487-4500.

[54] Strand J., Taxt T, and Jain A. K, "Two-dimensional phase unwrapping using a block least-squares method," *IEEE Trans. Image Process*. 1999, 8**:** 375–386.

[55] Farinas J, Verkman A.S, "Cell volume and plasma membrane osmotic water permeability in epithelial cell layers measured by interferemetry," *Biophys J* 1996, **71**: 3511-3522.

Bibliography

R. E. Fischer and B. Tadic-Galeb, *Optical System Design*, SPIE Press; McGraw-Hill, ISBN 0-07-134916-2 (2000).

J.W. Goodman, *Introduction to Fourier Optics*, *Second Edition*, McGraw-Hill, Inc., New York, NY (1996).

E. Hecht, *Optics*, *Fourth Edition*, Pearson Education, Inc., ISBN 81-7808-617-4 (2002).

C. S. Williams and O. A. Becklund, *Introduction to the Optical Transfer Function*, Wiley-Interscience, ISBN 0-471-94770-9 (1989).

U. Schnars, W Jueptner, *Digital Holography: Digital Hologram Recording, Numerical Reconstruction, and Related Techniques,* Springer; ISBN: 354021934X (2004).

P. Hariharan, *Basics of Holography***,** Cambridge University Press ISBN 0521002001 (2002).

Dennis C. Ghiglia, Mark D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, Wiley-Interscience ISBN: 0471249351 (1998).

M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light,* Cambridge University Press ISBN: 0521642221 (1999)

Appendices

Appendix A

Theorems

A.1. The Fourier Transform

The one-dimensional Fourier transform of a reasonably well-behaved function *f(x)* is

$$\Im\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\exp[-i2\pi ux]dx$$

The inverse is then found by

$$\Im^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u)\exp[i2\pi ux]du$$

where $f(x)$ and $F(u)$ are Fourier transform pairs.

The two-dimensional Fourier transform of *f(x)* is

$$\Im\{f(x)\} = F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\exp[-i2\pi(ux+vy)]dxdy$$

The inverse is then

$$\Im^{-1}\{F(u)\} = f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)\exp[i2\pi(ux+vy)]dudv$$

(x, y) are the spatial coordinates and (u, v) are the corresponding spatial frequencies.

A.2. Theorems of Fourier Transforms

A.2.1 Convolution:

$$(f \oplus g)(x,y) \equiv \int dx'dy' f(x',y')\, g(x-x', y-y')$$

where the convolution operation is defined by $\oplus$

A.2.2 Correlation:

$$f \otimes g(x) \equiv \int dx' f(x') g^*(x' - x)$$
$$= [g \otimes f]^*(-x)$$
$$= f(x) \oplus g^*(-x)$$

A.2.3 Auto-Correlation:

$$f \otimes f(x) \equiv \int dx' f(x') f^*(x' - x)$$
$$= f(x) \oplus f^*(-x)$$

A.2.4 Other useful convolution theorems:

$$\text{F} \quad \{f \oplus g\} = \sqrt{2\pi} F \cdot G \qquad \text{F} \quad \{f \cdot g\} = \frac{1}{\sqrt{2\pi}} F \oplus G$$

$$\text{F} \quad \{f \otimes g\} = \sqrt{2\pi} F \cdot G^* \qquad \text{F} \quad \{f \cdot g^*\} = \frac{1}{\sqrt{2\pi}} F \otimes G$$

$$\text{F} \quad \{f \otimes f\} = \sqrt{2\pi} |F|^2 \qquad \text{F} \quad \{|f|^2\} = \frac{1}{\sqrt{2\pi}} F \otimes F$$

$$f \oplus (g \otimes h) = (f \oplus g) \otimes h \equiv f \oplus g \otimes h$$
$$f \oplus g \otimes h = \bar{h}^* \oplus f \otimes \bar{g}^* = g \oplus \bar{h}^* \otimes \bar{f}^*$$
$$\text{F} \quad \{f \oplus g \otimes h\} = 2\pi F \cdot G \cdot H^*$$

$$\left[\bar{f}(x) \equiv f(-x)\right]$$

$$f \oplus g(x) = g \oplus f(x)$$
$$f \otimes g(x) = g^* \otimes \bar{f}^*(x)$$
$$f \otimes g(x) = f \oplus \bar{g}^*(x)$$

$$f(x) \oplus \delta(x-a) = \delta(x-a) \oplus f(x) = f(x-a)$$
$$f(x) \otimes \delta(x-a) = f(x+a)$$
$$\delta(x-a) \otimes f(x) = f^{*}(a-x)$$

$$f(x-a) \otimes f(x-b) = \delta'(x-a+b)$$

A.3. Imaging with a coherent light source

A.3.1 Impulse Response Function, *h:*

$$\delta(x_0 - X_0)\delta(y_0 - Y_0) \xrightarrow[\text{System}]{\text{Optical}} h(x_1 - X_1, y_1 - Y_1)$$
$$X_1 = mX_0; \quad Y_1 = mY_0; \quad [\text{Let m} = 1]$$

Then

$$E_1(x_1, y_1) = \iint dx_0 dy_0 E_0(x_0, y_0) h(x_1 - x_0, y_1 - y_0)$$

$$\boxed{E_1(x_1, y_1) = E_0(x_1, y_1) \oplus h(x_1, y_1)}$$

A coherent imaging system is linear in complex amplitude.

A.3.2 Coherent Transfer Function (CTF), *H:*

$$\text{F}\quad \{E_0\} \equiv F_0; \ \text{F}\quad \{E_1\} \equiv F_1; \ \text{F}\quad \{h\} \equiv H$$
$$\kappa_x \equiv k\frac{x_1}{z_1}; \qquad \kappa_y \equiv k\frac{y_1}{z_1}$$

Then

$$\boxed{F_1(\kappa_x, \kappa_y) = F_0 \cdot H(\kappa_x, \kappa_y)}$$

For a coherently illuminated object, the image intensity pattern is:

$$I_1(x_1, y_1) = |E_1(x_1, y_1)|^2 = |E_0 \oplus h(x_1, y_1)|^2$$

A.3.3 Transfer function, *H:*

Fourier decomposition of the fields, with **k** = (k_x, k_y) = (kx/z, ky/z):

$$\text{F} \quad \{f_0\} \equiv F_0(\mathbf{k}); \qquad \text{F} \quad \{f_1\} \equiv F_1(\mathbf{k}); \qquad \text{F} \quad \{h\} \equiv H(\mathbf{k})$$

Then

$$\boxed{F_1(\mathbf{k}) = F_0 \cdot H(\mathbf{k})}$$

For an invariant system, the effect of the imaging system to a sinusoidal input is limited to amplitude change and phase shift. The transfer function approach is valid only in an invariant system.

A.4. Discrete Fourier Transform

A.4.1 Fourier Series

If $f(x)$ is a periodic function of period L, then

$$\boxed{\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} F_n \exp(inKx) \\ F_n &= \frac{1}{\Lambda}\int_0^{\Lambda} f(x)\exp(-inKx)dx \end{aligned}}$$

where $K \equiv \dfrac{2\pi}{\Lambda}$ is the fundamental frequency.

A.4.2 Cosine and Sine series:

$$f(x) = \tfrac{1}{2} C_0 + \sum_{n=1}^{\infty} C_n \cos(nKx) + \sum_{n=1}^{\infty} S_n \sin(nKx)$$

$$C_n = \frac{2}{\Lambda} \int_0^{\Lambda} f(x) \cos(nKx) dx$$

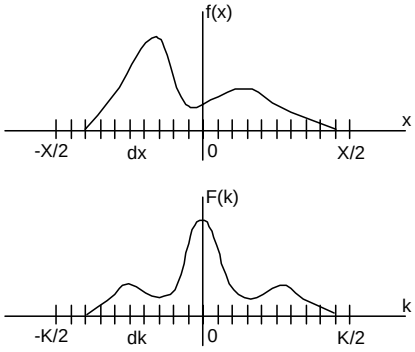
$$S_n = \frac{2}{\Lambda} \int_0^{\Lambda} f(x) \sin(nKx) dx$$

$$F_n = \tfrac{1}{2} [C_n - iS_n]$$

A.4.3 Discrete Fourier Transform

f(x) defined in [-X/2, X/2]:

$$f(x) = 0 \text{ for } x \notin \left[-\frac{X}{2}, \frac{X}{2} \right]$$



$f_s(x)$ sampled at dx intervals:

$$f_s(x) = f(x) \cdot \text{comb}\left(\frac{x}{\delta x} \right)$$

i.e., $f_s(x_i) = f(x_i)$ for $x_i \in \left[-\frac{X}{2} : \delta x : \frac{X}{2} \right]$

Fourier transform $F_s(k)$:

$$F_s(k) = \text{F} \quad \{f_s(x)\}[k] = \text{F} \quad \left\{ f(x) \cdot \text{comb}\left(\frac{x}{\delta x} \right) \right\}$$

$$= F(k) \oplus \frac{\sqrt{2\pi}}{\delta x} \text{comb}\left(\frac{k}{2\pi / \delta x} \right) = \frac{\sqrt{2\pi}}{\delta x} \sum_{n=-\infty}^{\infty} F\left(k - n \frac{2\pi}{\delta x} \right)$$

Therefore, if $F(k) = 0$ for $k \notin \left[-\frac{K}{2}, \frac{K}{2}\right]$, where $K = \frac{2\pi}{\delta x}$, then

$$F_s(k) = F(k) \text{ for } k \in \left[-\frac{K}{2}, \frac{K}{2}\right]$$

Conversely:

F(k) defined in [-K/2, K/2]: $F(k) = 0$ for $k \notin \left[-\frac{K}{2}, \frac{K}{2}\right]$

F_s(k) sampled at dk intervals: $F_s(k) = F(k) \cdot \text{comb}\left(\frac{k}{\delta k}\right)$

Fourier transform f_s(x): $f_s(x) = f(x) \oplus \frac{\sqrt{2\pi}}{\delta k} \text{comb}\left(\frac{x}{2\pi / \delta k}\right)$

Therefore, if $f(x) = 0$ for $x \notin \left[-\frac{X}{2}, \frac{X}{2}\right]$, where $X = \frac{2\pi}{\delta k}$, then

$$f_s(x) = f(x) \text{ for } x \in \left[-\frac{X}{2}, \frac{X}{2}\right]$$

Therefore, if both f(x) and F(k) are discretized with N+1 points, then:

$$\begin{cases} K = N\delta k = N\frac{2\pi}{X} = \frac{2\pi}{\delta x} & \delta k = \frac{K}{N} = \frac{2\pi}{N\delta x} = \frac{2\pi}{X} \\ X = N\delta x = N\frac{2\pi}{K} = \frac{2\pi}{\delta k} & \delta x = \frac{X}{N} = \frac{2\pi}{N\delta k} = \frac{2\pi}{K} \end{cases}$$

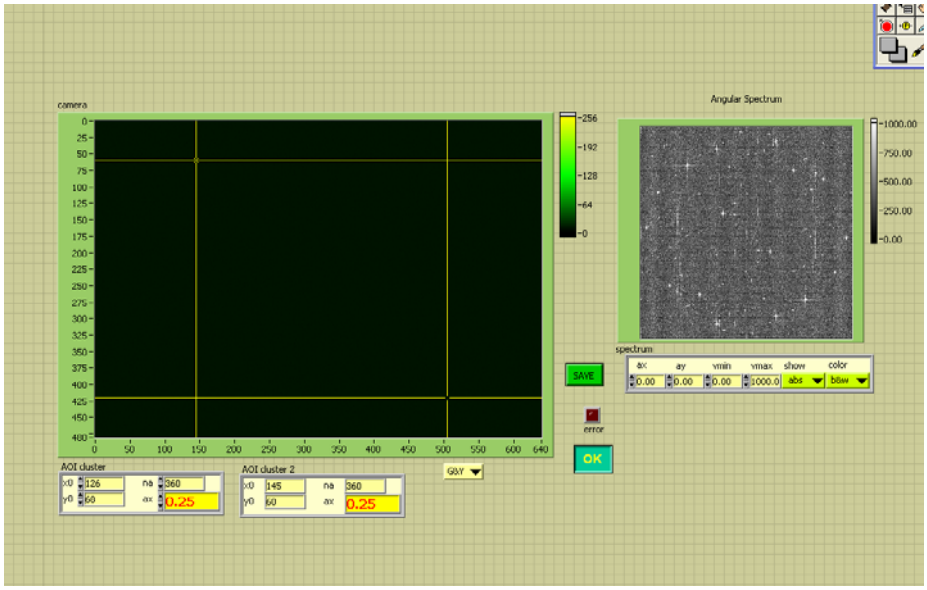
Appendix B

Listing of Labview Programs

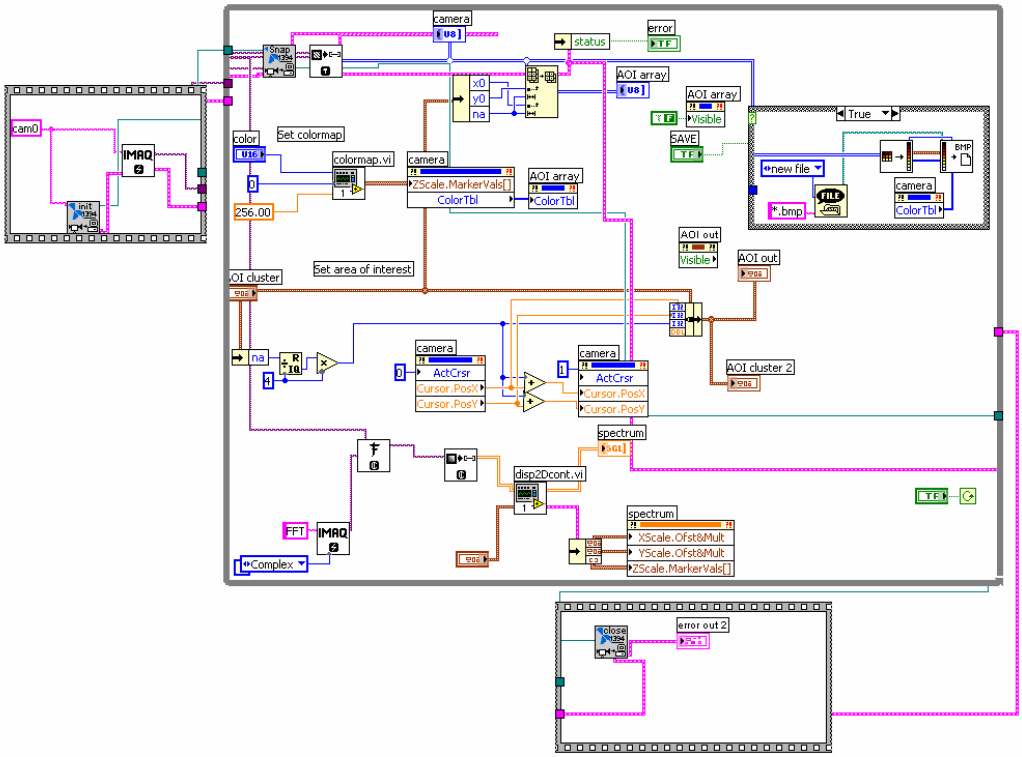
This appendix contains a listing of the computer programs written in Labview which are used to acquire holograms from the firewire CCD camera and perform numerical reconstruction of the hologram.

B.1 Calibration of data parameters and optical alignment via the angular spectrum (Firecamera.vi)

B.1.1 Front Panel

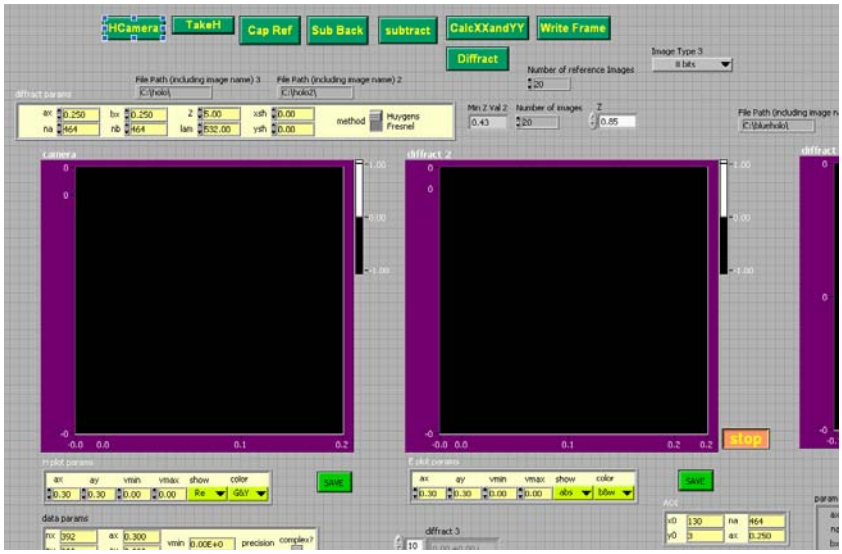


B.1.2 Block Diagram

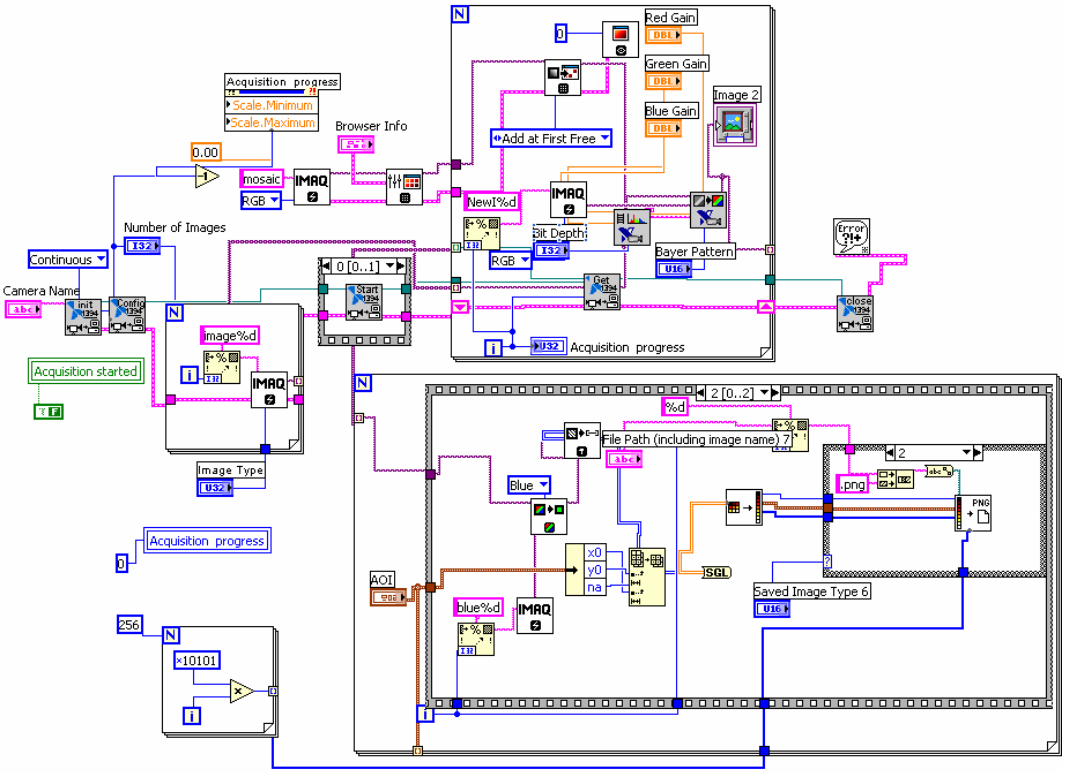


B.2 Image capture of holograms (ExperimentalHolo.vi)

B.2.1 Front Panel

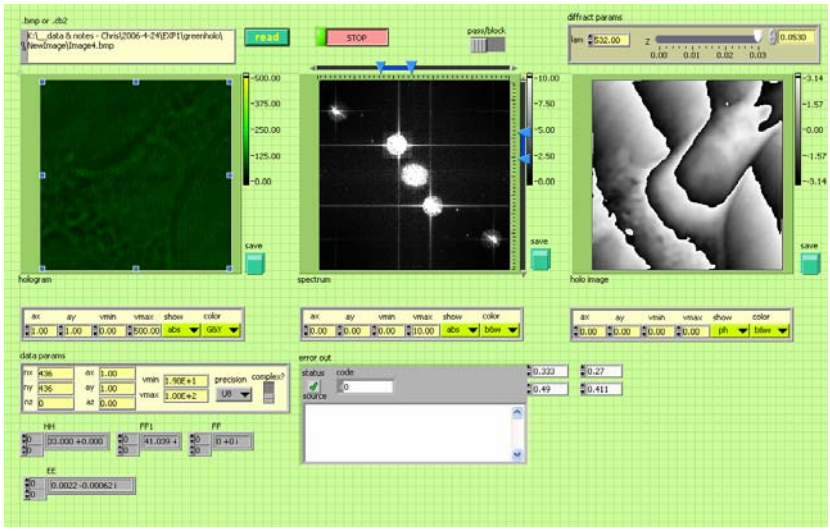


B.2.2 Block Diagram

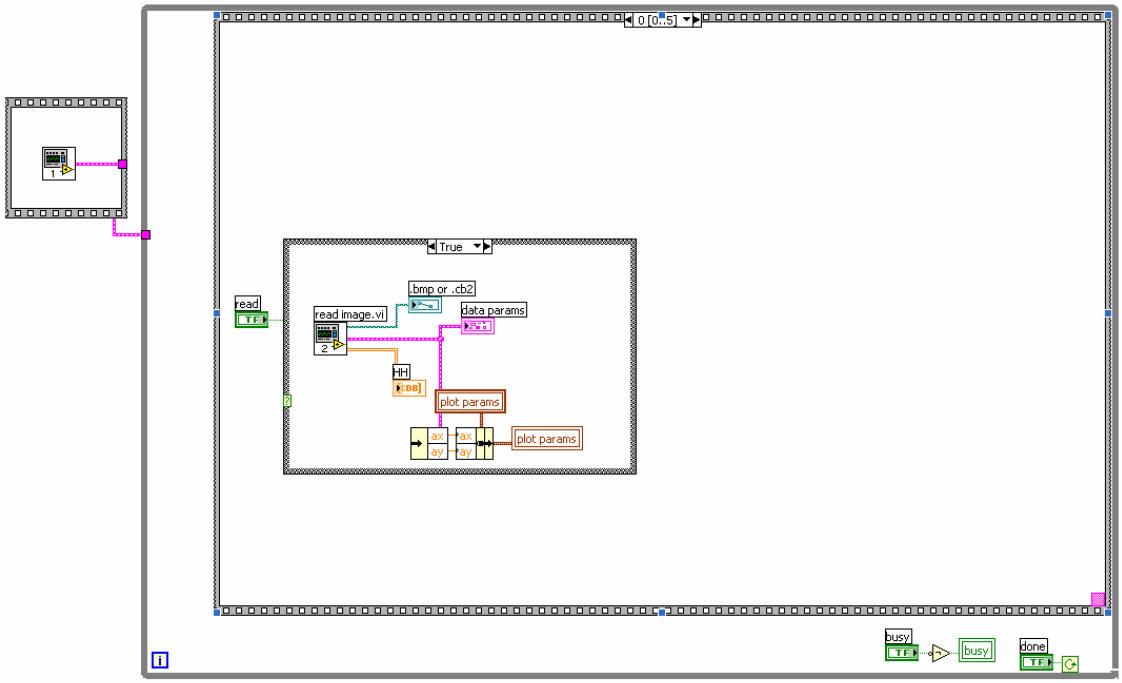


B.3 Reconstruction of the hologram (Holodiffract.vi)

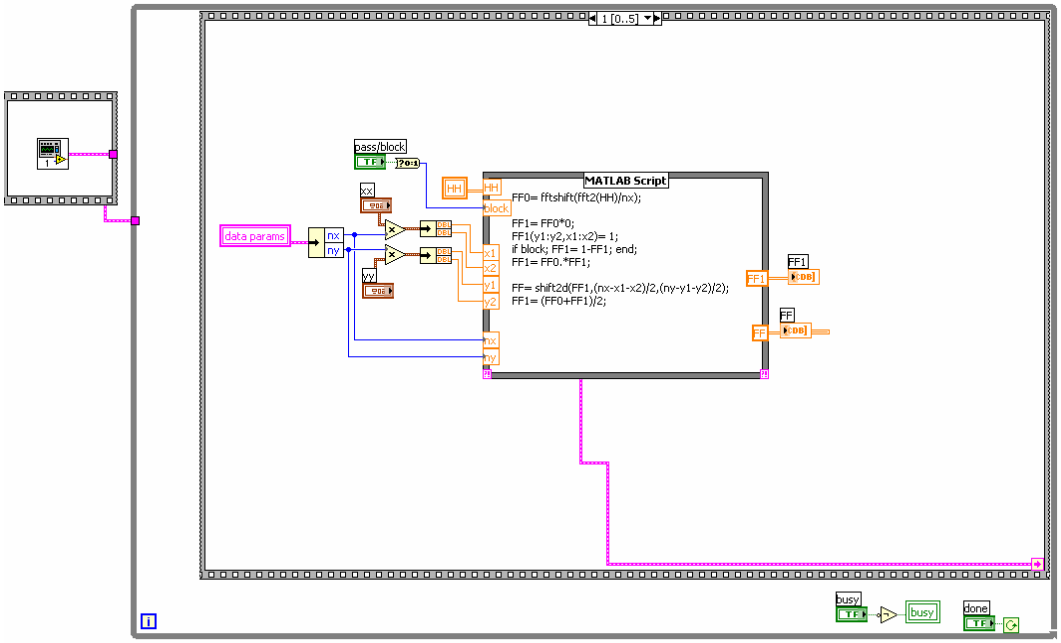
B.3.1 Front Panel



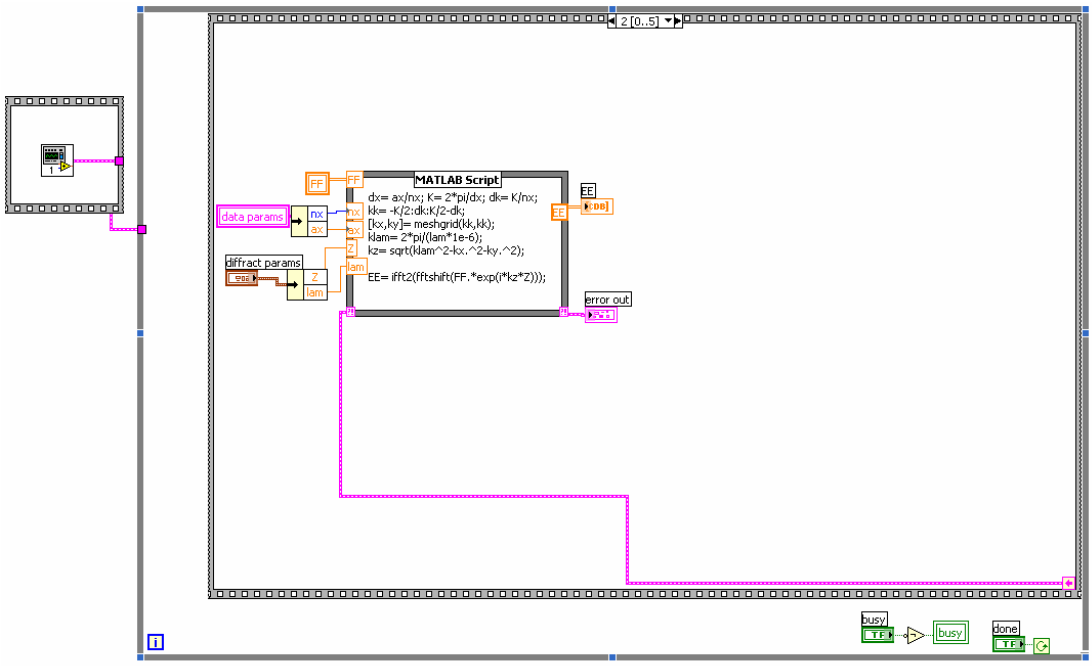
B.3.2 Block Diagram- Reading hologram into reconstruction program



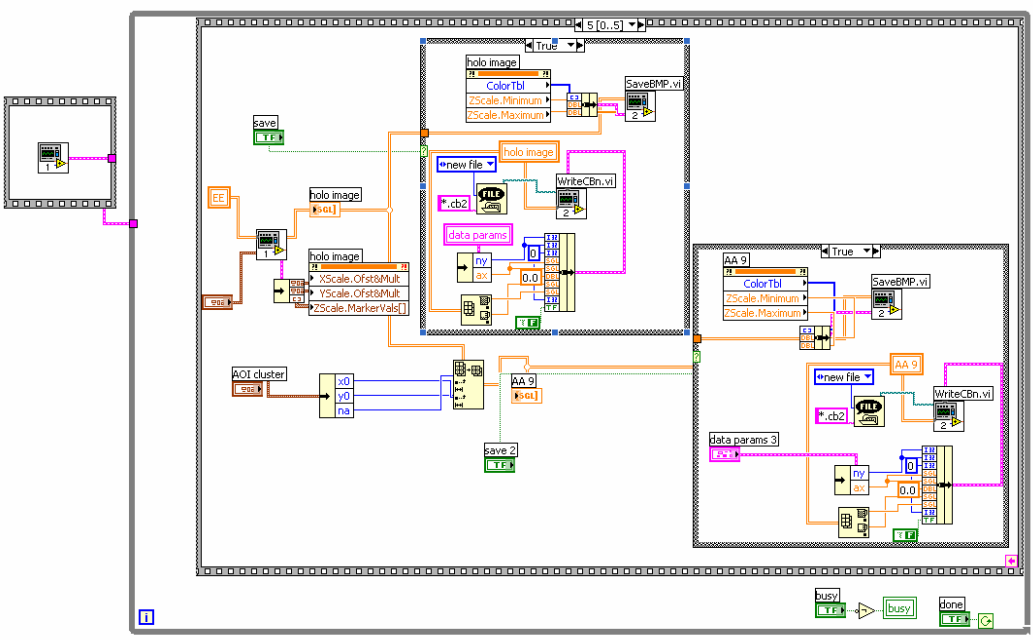
B.3.3 Block Diagram- Performing frequency domain analysis



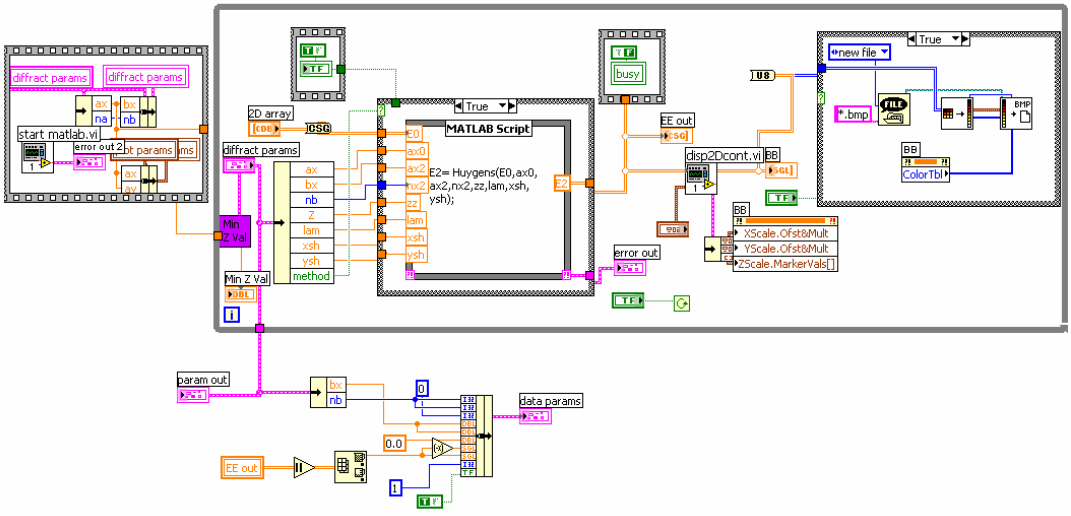
B.3.4 Block Diagram- Calculating Angular Spectrum



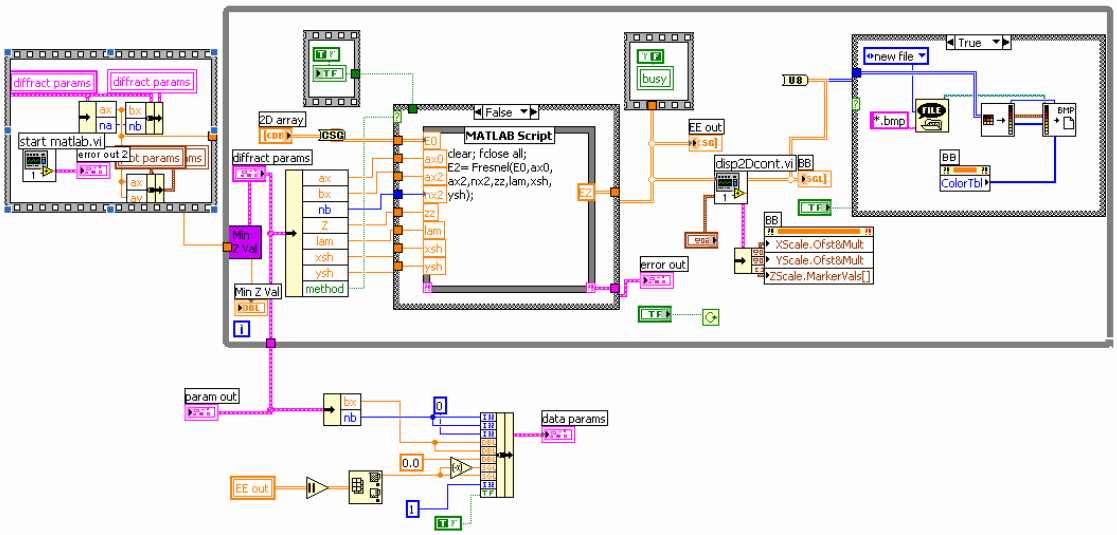
B.3.5 Block Diagram- Saving images



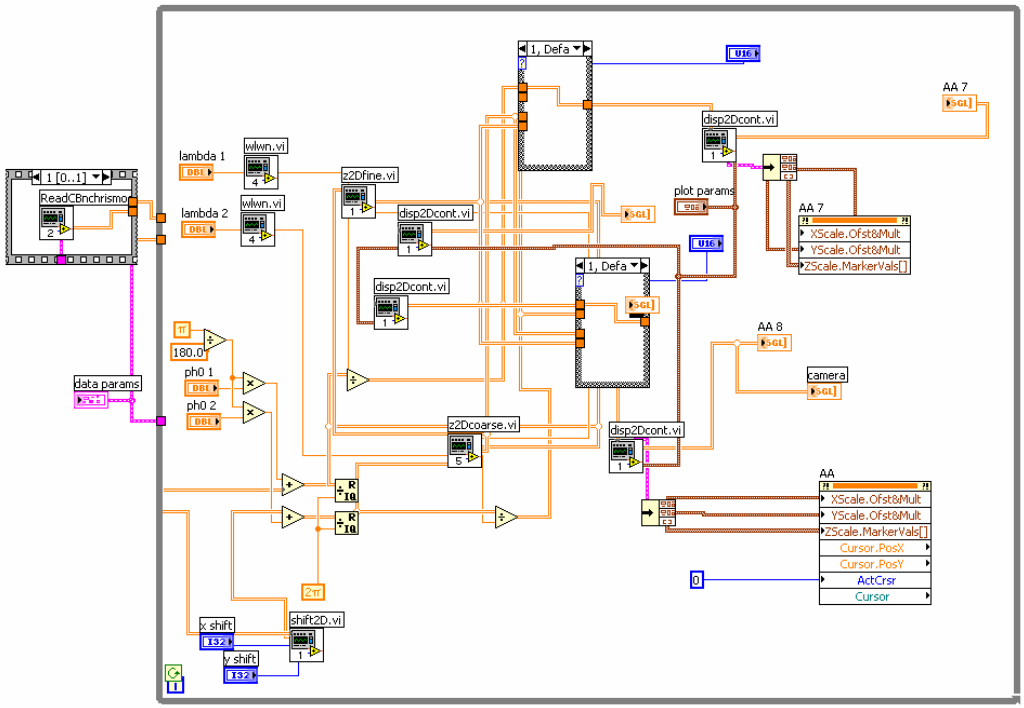
B.4 Calculating Huygens diffraction



B.5 Calculating Fresnel diffraction



B.6 Multi-Wavelength Phase imaging



Appendix C

Listing of Matlab programs

This appendix contains a listing of the computer programs written in Matlab which are used to numerically reconstruct the hologram and un-wrap the reconstructed phase image. These algorithms can be integrated into the Labview environment one may observe in the block diagrams displayed in Appendix B.

C.1 Numerical Reconstruction in Matlab

C.1.1 Huygens Algorithm

```
function E1= Huygens(E0,ax0,ax1,nx1,zz,lam,xth,yth)
% function E2= Huygens(E0,ax0,ax2,nx2,zz,lam,xth,yth)
% xth and yth are horizontal and vertical angular shift of object
% in degrees relative to reference for off-axis hologram.
%
% 3/10/2005 mkk
%

if zz==0; E1= E0; return; end

pi2= pi*2;
kk= pi2/(lam*1e-6);

nx0= size(E0,1);
dx0= ax0/nx0;
dx1= ax1/nx1;

XX0= [-ax0/2:dx0:ax0/2-dx0];
[X0,Y0]= meshgrid(XX0,XX0);
XX1= [-ax1/2:dx1:ax1/2-dx1];
[X1,Y1]= meshgrid(XX1,XX1);

xth= pi*xth/180; yth= pi*yth/180;
E0= E0.*exp(i*kk*(xth*X0+yth*Y0));

Eo= reframe(E0,'n',ax0,nx0,ax1,nx1);
```

```
%SS= pi2*(-i*kk/(pi2*zz))*exp(i*kk*sqrt(X1.^2+Y1.^2+zz^2));
SS= exp(i*kk*sqrt(X1.^2+Y1.^2+zz^2));

Eo= fft2(Eo)/nx1;
SS= fft2(SS)/nx1;
E1= fft2(Eo.*SS)/nx1;
E1= fftshift(E1);
E1= flipud(fliplr(E1));
```

C.1.2 Fresnel Algorithm

```
function E2= Fresnel(E0,ax0,ax2,nx2,zz,lam,xsh,ysh)
% function E2= Fresnel(E0,ax0,ax2,nx2,zz,lam,xsh,ysh)
% xsh and ysh are horizontal and vertical shift of object
% relative to reference for off-axis hologram.
%
% 6/24/01 mkk
%

if zz==0; E2= E0; return; end

pi2= pi*2;
kk= pi2/(lam*1e-6);

nx0= size(E0,1);
dx0= ax0/nx0;
ax1= axbx(ax0,nx0,zz,kk);
nx1= nx0;
dx1= ax1/nx1;

XX0= [-ax0/2:dx0:ax0/2-dx0];
[X0,Y0]= meshgrid(XX0,XX0);
XX1= [-ax1/2:dx1:ax1/2-dx1];
[X1,Y1]= meshgrid(XX1,XX1);

thx= xsh/zz; thy= ysh/zz;
E0= E0.*exp(i*kk*(thx*X0+thy*Y0));

S0= exp((i*kk/(2*zz))*(X0.^2+Y0.^2));
%SS= pi2*(-i*kk/(pi2*zz))*exp(i*kk*(zz+(X1.^2+Y1.^2)/(2*zz)));
SS= exp(i*kk*(zz+(X1.^2+Y1.^2)/(2*zz)));

E1= fft2(E0.*S0)/nx0;
E1= SS.*E1;
E1= fftshift(E1);

E2= reframe(E1,'n',ax1,nx1,ax2,nx2);
```

C.1.3 Angular Spectrum Algorithm

```
function [E2,SO]= Angspc(E0,ax0,ax2,nx2,zz,lam,xsh,ysh);
% function E2= Angspc(E0,ax0,ax2,nx2,zz,lam,xsh,ysh);
% xsh and ysh are horizontal and vertical shift of object
% relative to reference for off-axis hologram.
%

[M,N]=size(E0);

I1=fftshift(fft2(double(E0)));
%figure(1); imshow(mat2gray(abs(I1))/0.003);

%=========================================================Fresnel
Diffraction
wlength=lam*1e-9;%594e-9;
x=ax0*1e-3;
y=ax0*1e-3;
d=1e-3*zz;

%=============================================filtering

I2=zeros(M,N);
%I2(:,1:70)=I1(:,1:70);
I2(1:M,1:100)=I1(1:M,1:100);
%layer102=shiftxy(I2,xsh,ysh);
nx=xsh*100;
ny=ysh*100;
F00=zeros(M,N);
F00(1:M-nx,1:N-ny)=I2(nx+1:M,ny+1:N);
F00(M-nx+1:M,1:N-ny)=I2(1:nx,ny+1:N);
F00(1:M-nx,N-ny+1:N)=I2(nx+1:M,1:ny);
F00(M-nx+1:M,N-ny+1:N)=I2(1:nx,1:ny);
SO=F00;
%=========================================================Angular
Spectrum


F01=(conj(F00));
MM=M/2;
NN=N/2;
%----------------------------------------

%d is the position of reconstruction
%d=1.2;
c1=j*2*pi*d;
c2=1/wlength^2;

dk=1/x;
dl=1/y;

for k=1:M
   for l=1:N
```

```
            T3(k,l)=F01(k,l)*exp(c1*sqrt(c2-((MM-k)*dk)^2-((NN-
l)*dl)^2));
   end
end


E2=ifft2(fftshift(T3));%/(j*wlength*d);
```

C.2 Flood Fill Phase Unwrapping Programs

C.2.1 *Main.m* - Loads a wrapped phase image

```
a = imread('Phase1.bmp','bmp');
c=double(a);
size_of_b=size(c);
for i=1:size_of_b(1,1)
   for j=1:size_of_b(1,1)
      b(i,j)=c(i,j,1); %^2+c(i,j,2)^2+c(i,j,3)^2)^0.5;
   end
end

arraydisp(b,0,255);              %display image
theta=b*2*pi/255-pi;
quality_map=qualitymap(b,size_of_b(1,1));
arraydisp(quality_map,0,0.4);   %display quality map
ture_quality_map=rmd(quality_map,size_of_b(1,1),0.01);
arraydisp(ture_quality_map,0,0.4);   %diplay quality map without
discontinuity

function [a1,a2]=add2list(adjoin1,adjoin2,x,y,q)
```

C.2.2 *add2list.m* –Adds a position and quality value to adjoining x,y

```
% x,y is the postion, q is the quality

[h,long]=size(adjoin2);
for i=1:long
   if adjoin2(i)==0
      adjoin1(i,1)=x;
      adjoin1(i,2)=y;
      adjoin2(i)=q;
      break;
   end
end
if i==long
   adjoin1(long+1,1)=x;
   adjoin1(long+1,2)=y;
   adjoin2(long+1)=q;
end
```

C.2.3 *arraydisp.m* – Displays the array image

```
a1=adjoin1;
a2=adjoin2;
% ARRAYDISP() display the array image


function arraydisp(array,minv,maxv)

% ARRAYDISP() display the array image

figure
imshow(mat2gray(array,[minv,maxv]))
axis on
```

C.2.4 *main2.m* – Central program for phase unwrapping

```
%for i=1:232
%   for j=1:232
%      test(i,j)=theta(i,j);
%      test_map(i,j)=ture_quality_map(i,j);
%   end
%end
uwphase=qgunwrap(theta,ture_quality_map,size_of_b(1,1),190,190);
mdisp(uwphase,min(min(uwphase)),max(max(uwphase)));

cmax=max(max(uwphase));
cmin=min(min(uwphase));
%uwph=uwphase-cmin;
uwph=cmax-cmin-(uwphase-cmin);
max(max(uwph))
value=80-uwph;
newvalue=max(max(value));
uwphnew=newvalue+uwph;

figure(2)
imshow(mat2gray(uwphnew))
imwrite(uint8(255*mat2gray(uwphnew)/1),colormap,'Unwrapphaserr123234new
cellnewTa0.bmp')
```

C.2.5 *qmdisp.m* – display for array

```
function mdisp(map,min,max)
%QMDISP

%t=map.*mask0+mask1*max;
arraydisp(map,min,max);                          % FIGURE
```

```
function uwphase=qgunwrap(theta,qmap,N,s_x,s_y)
```

C.2.6 *qgunwrap.m* – quality guided path following method

```
%QGUNWRAP.m
%N=100;

start_time=datestr(now);
disp(['File generating begins on ' start_time ', please wait ...'])
mphase=theta;
uwphase=zeros(N);
ifunwrap=zeros(N);

% start to unwrap phase using quality guided plus flood-fill algorithm
%--------------------------------------------------------------------------
start_x=s_x; %starting point
start_y=s_y;
uwphase(start_x,start_y)=mphase(start_x,start_y);
ifunwrap(start_x,start_y)=1;
long=1;
ifempty=1;
adjoin1(1,1)=start_x;
adjoin1(1,2)=start_y;
adjoin2(1)=qmap(start_x,start_y);
tt=0;

% perform flood-fill algorithm until the ajoint list is empty
%--------------------------------------------------------------
while ifempty>0
   [t,position]=max(adjoin2);       % First find the maximun value from
adjoin list
   i=adjoin1(position,1);
   j=adjoin1(position,2);
   adjoin1(position,1)=0;           % remove the pixel being operated
from the adjoin list
   adjoin1(position,2)=0;
   adjoin2(position)=0;
   if i>1 & i<N & j>1 &j<N
      if ifunwrap(i,j-1)==0   %if not border and unwrap pixel, perform
phase unwrapping
         d=unwrp(mphase(i,j),mphase(i,j-1));
         uwphase(i,j-1)=uwphase(i,j)+d;
         ifunwrap(i,j-1)=1;
         [adjoin1,adjoin2]=add2list(adjoin1,adjoin2,i,j-1,qmap(i,j-1));
      end
      if ifunwrap(i-1,j)==0 %& mask0(i-1,j)==1  %if not border and
unwrap pixel, perform phase unwrapping
         d=unwrp(mphase(i,j),mphase(i-1,j));
         uwphase(i-1,j)=uwphase(i,j)+d;
         ifunwrap(i-1,j)=1;
         [adjoin1,adjoin2]=add2list(adjoin1,adjoin2,i-1,j,qmap(i-1,j));
      end
      if ifunwrap(i,j+1)==0 %& mask0(i,j+1)==1  %if not border and
unwrap pixel, perform phase unwrapping
```

```
            d=unwrp(mphase(i,j),mphase(i,j+1));
            uwphase(i,j+1)=uwphase(i,j)+d;
            ifunwrap(i,j+1)=1;
            [adjoin1,adjoin2]=add2list(adjoin1,adjoin2,i,j+1,qmap(i,j+1));
        end
        if ifunwrap(i+1,j)==0 %& mask0(i+1,j)==1  %if not border and
unwrap pixel, perform phase unwrapping
            d=unwrp(mphase(i,j),mphase(i+1,j));
            uwphase(i+1,j)=uwphase(i,j)+d;
            ifunwrap(i+1,j)=1;
            [adjoin1,adjoin2]=add2list(adjoin1,adjoin2,i+1,j,qmap(i+1,j));
        end
    end
    ifempty=sum(adjoin2);
    tt=tt+1
end
tt
end_time=datestr(now);
disp(['File generating ends on ' end_time '.']);



function map=qualitymap(theta,s)
```

C.2.7 *qualitymap.m* –

```
%QUALITYMAP.m calculate the quality map of input image

start_time=datestr(now);
disp(['File generating begins on ' start_time ', please wait ...'])

theta_temp=(flipud(theta))';

for i=1:2:(s-3)
    for j=1:(s-1)
        ma=theta_temp(i+2,j)-theta_temp(i,j);
        mb=theta_temp(i,j+1)-theta_temp(i,j);
        mc=theta_temp(i,j+1)-theta_temp(i+2,j+1);
        md=theta_temp(i+2,j)-theta_temp(i+2,j+1);
        c(i,j)=2/sqrt(ma^2+4*mb^2+4);
        c(i+1,j)=2/sqrt(mc^2+4*md^2+4);
    end
end
map=flipud(c');
map(:,s-1:s)=1;
map(s,:)=1;

end_time=datestr(now);
disp(['Programe runnig ends on ' end_time '.'])
function mapt=rmd(map,s,t)
```

C.2.8 *rmd.m* – remove the discontinuity, s is image size. t is theshold

```
%RMD.m
```

```
disp('Precessing ...')
mapt=map;
if mapt(1,1)<t            % point (1,1)
   mapt(1,1)=t;
end
for j=2:s                 % first line
   if mapt(1,j)<t
      mapt(1,j)=mapt(1,j-1);
   end
end

for i=2:s        % second line to end
   if mapt(i,1)<t
      mapt(i,1)=mapt(i-1,1);
   end
   for j=2:s
      if mapt(i,j)<t
         mapt(i,j)=max(mapt(i,j-1),mapt(i-1,j));
      end
   end
end
```

C.2.9 *unwrap.m* - not is unfinished point; finish is finished point % d is the difference

```
function d=unwrp(finish,not)

t=not-finish;
if t>pi
   d=not-2*pi-finish;
elseif t<-pi
   d=not+2*pi-finish;
else
   d=not-finish;
end
```

C.2.10 *view2D.m* – view unwrapped phase image

```
A = IMREAD('UnwraNewPhase','bmp');
%A= double(A)/255;
%A=double(A)*2*pi/255;
cmax=max(max(A));
cmin=min(min(A));
%A=(A*0.532)/(2*pi)/(0.375);
%A=(A*0.532)/(2*pi);
A=(double(A)/255)*(20.8700);
A=(A*0.532)/(2*pi*0.375);
cmax=max(max(A));
cmin=min(min(A));
A= flipud(A);
```

```
surf(A);
shading flat;
axis tight;
rotate3d on;
colormap hsv;
lighting phong;
light;
```

Appendix D

List of Accomplishments

Journal Publications

1. C. J. Mann, L. Yu, C. Lo, and M. K. Kim, "High-resolution quantitative phase-contrast microscopy by digital holography," *Opt. Express* **13,** 8693-8698 (2005) http://www.opticsexpress.org/issue.cfm?volume=13&issue=22

2. C.J. Mann, L. Yu, & M.K. Kim, "Movies of cellular and sub-cellular motion by digital holographic microscopy", *J. Biomedical Engineering* **5,** 21 (2006) http://www.biomedical-engineering-online.com/content/5/1/21/abstract.

3. M.K. Kim, L. Yu, and C.J. Mann, "Interference techniques in digital holographic microscopy", *J. Opt. A: Pure Appl. Opt.* **8:** S518–S523 (2006)

Conference Papers

1. C.J. Mann & M.K Kim, "Digital Gabor holography for particle field imaging", *OSA Biomedical Topical Meetings*, Miami, FL (April 14-17, 2004) poster presentation

2. C.J. Mann & M.K. Kim, “Digital Gabor holographic movie of animal microbes”, *OSA. Annual Meeting,* Rochester, NY, (October 10-14, 2004) oral presentation

3. C.J. Mann & M.K. Kim “Movies of amplitude and phase images of paramecium by digital holographic microscopy”, poster presentation at *USF Interdisciplinary Graduate Research Symposium* (April 2005)

4. C.J. Mann & M.K. Kim, “Phase Imaging Digital holographic movies of animal cells”, *OSA, CLEO/QELS*, Baltimore, MA, (May 24-26, 2005) oral presentation

5. M.K. Kim, L. Yu, and C.J. Mann, “Interference techniques in digital holographic microscopy”, *OMS 05 Optical Microsystems,* Capri, Italy (Sep.2005) oral presentation

6. C.J. Mann & M.K Kim, “Quantitative phase-contrast microscopy by angular spectrum digital holography”, *SPIE BIOS*, San Jose, CA (Jan. 21-26, 2006) oral presentation.

7. C.J. Mann & M.K Kim, “Quantitative biological microscopy of cells by digital holography”, *OSA Biomedical Topical Meetings,* Fort Lauderdale, FL (March 19-22, 2006): oral presentation.

8. C.J. Mann & M.K. Kim “Microscopy by digital holography”, poster presentation at *USF Interdisciplinary Graduate Research Symposium* (April 2006)

Book Chapters

M.K. Kim, L. Yu, and C.J. Mann, “Digital holography and multi-wavelength interference techniques”, Chap. 2 in *Digital holography and three-dimensional display*, T.C. Poon, ed. (Springer 2006)

About the Author

Christopher Mann currently resides in the Tampa Bay area. He grew up in Willenhall, West Midlands, England and received a Bachelor of Science degree in Physics from the University of Birmingham, England in 2001 and a Master of Science degree in Physics with Astrophysics from the University of Birmingham in 2002. He entered the Ph.D program in Applied Physics at the University of South Florida in Fall 2002.

He has completed an Industrial Practicum at MediVision, Tel Aviv Israel as part of the Applied Physics practical training. He has two journal publications in Optics Express and the Journal of Biomedical Engineering and has presented at numerous technical conferences including CLEO, SPIE and OSA topical meetings.