

Chapter 2

Theoretical aspects of digital holography

2.1 Introduction

Holography is used to record sufficient information for three-dimensional imaging of an object. Conventional holography is generally regarded as a practical technique for imaging a 3D object [2.1-2.3]. The modern invention of digital megapixel sensors and a digital computer made digital holography possible [2.4-2.5]. Digital holography is thus referred to as the technique that uses a CCD camera to record holographic patterns and performing the reconstruction numerically using a computer by means of diffraction theory [2.6]. These recorded intensity patterns or holograms contain amplitude and the phase information of the object. Moreover, these digitally recorded holograms have many advantages over conventional holography such as ease of transmission, faster processing and accuracy in analyzing. In recent years digital holography technique has been demonstrated to be a valuable method in different fields of optics like measurement of refractive index, particle size, digital holographic microscopy, deformation analysis and object contouring among others [2.7, 2.8]. The most interesting aspect of digital holography is that numerical reconstruction provides whole field of the recorded wavefront. That is it yields the intensity distribution and the phase distribution of the wave field at any arbitrary plane located between the object and the recording plane.

The present chapter deals with the principles of the digital holography technique for numerical reconstruction of the complex wave field. The basic concept and procedure of wave field reconstruction using digital holography is discussed. In recent era of digital imaging, digital holography appears to be a stronger competitor in the field of optical metrology for measurement [2.9]. Since digital holography is non-destructive and non-contact testing, it is preferred for measuring refractive index distributions and

temperature mapping within transparent media, wavefront sensing, particle size measurement, applications in deformation analysis and profile measurement and for many other applications [2.10]. Thus, by using digital holography, it is possible to achieve the spatial three dimensional profile measurements of many objects.

2.2 Interference

If light originating from a source get divided into two beams, the inherent fluctuations in those two beams are in general correlated, and the beams are said to be completely or partially coherent depending on whether full or partial correlation exists. In light beams from two independent sources, the phase functions are usually uncorrelated and such beams are called incoherent beams. When coherent waves superpose, they produce visible interference effects because their amplitudes combine constructively and destructively [2.11, 2.12]. Interference produced by incoherent waves varies too rapidly in time to be practically observed.

When two mutually coherent beams pass through a point, we can observe the phenomena of interference between the wavefronts. The medium at that point is subjected to the total effect of the superposition of the two vibrations, and under certain conditions, this superposition results in stationary waves, which can be observed as interference fringes. Consider the superposition of two monochromatic plane waves of complex amplitudes U_1 and U_2 of the same frequency and with different amplitudes. The result is a monochromatic wave of the same frequency and the complex amplitude is the sum of the individual complex amplitudes, i.e. $U=U_1+U_2$. The plane waves can be expressed in terms of their intensities $I_1=|U_1|^2$ and $I_2=|U_2|^2$.

Therefore the resultant intensity I of the interference pattern is

$$I=|U|^2=|U_1+U_2|^2=|U_1|^2+|U_2|^2+U_1U_2^*+U_1^*U_2 \quad (1)$$

where the asterisk denotes complex conjugation.

The individual waves can be also represented by

$$U_1 = \sqrt{I_1}e^{i\phi_1} \quad \text{and} \quad U_2 = \sqrt{I_2}e^{i\phi_2}$$

and the resultant wave by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \varphi \quad (2)$$

Eq. (1) is known as the interference equation and the term $2\sqrt{I_1 I_2} \cos \varphi$ is known as the interference term. It contains information of the phases of the two beams. At different points in space, the resultant irradiance can be greater, less than or equal to $I_1 + I_2$ depending on the value of the interference term, which in turn depends on phase difference (φ). Irradiance maxima and minima occur for $\varphi = 2m\pi$ and $\varphi = (2m+1)\pi$ respectively. The dark and bright zones that would be seen on a screen placed in the region of interference are known as interference fringes.

An interferometer is, in the broadest sense, a device that generates interference fringes. Interferometers can basically be classified into two types: wavefront splitting interferometers and amplitude splitting interferometers. Wavefront splitting interferometers recombine two different parts of a wavefront to produce fringes. Amplitude-splitting interferometers, on the other hand, divide the intensity of the beam (splitting the amplitude), which propagate through separate paths and are then recombined. Holograms are nothing but the interference pattern between the wavefront scattered by the object and a known background wavefront usually called the reference wavefront and follows Eq (2).

2.3 Coherence

Coherence is the most important property light should exhibit to be useful for holography purposes. It is very important in relation to the interference which has been already described in the previous section. It is therefore appropriate to explain what coherence is.

Waves are said to be coherent when they are either in same phase or have constant phase difference. For the interference to take place, it is essential that the superposing waves are coherent [2.13].

Coherence quantizes the ability of the light to form a visible diffraction pattern. It directly influences the quality and the visibility of the interference pattern which consists of areas with different degrees of constructive or destructive interference. The areas are usually referred to as fringes. The fringes are more distinct if two interfering

waves are more coherent and they are less distinct if waves are less coherent. This visibility is quantified by the degree of coherence [2.14]. In other words, coherence determines the ability of two interfering waves to create total constructive or destructive interference. While perfectly coherent waves create clearly visible interference pattern the incoherent ones won't create visible interference fringe at all.

2.4 Hologram recording

Recording of hologram is an important step in the process of holography, as it provides three dimensional information of object [2.15]. Recording of holograms can be done on different photo-sensitive mediums. In conventional holography, recording is done on photosensitive materials like photographic plates. Electronic devices for capturing light intensity e.g. photo diodes have been also known for long time but the introduction of digital CCD cameras revolutionized this field of recording. In digital holography recording is done with the help of CCD cameras [2.7, 2.8].

Experimental setup for recording digital holograms using off-axis method is discussed and shown in Fig. 2.1. This setup is known as a Mach-Zehnder interferometer [2.16]. This is the most commonly used hologram recording setup in the course of this work. The beam from a coherent source (Laser) is split into two. One of these beams passes through or scattered from the object under investigation and reaches the recording medium. This constitutes the object beam. The other wave called the reference beam is allowed to illuminate the recording medium directly. These two waves interfere at the recording medium to produce the hologram. The hologram is recorded on semiconductor arrays like CCD or CMOS [2.7-2.9].

For an object wavefront with real amplitude distribution of $o(\xi, \eta)$ and phase distribution $\phi_o(\xi, \eta)$ the complex amplitude can be written as

$$O(\xi, \eta) = o(\xi, \eta) e^{i\phi_o(\xi, \eta)} \quad (3)$$

The plane (x, y) is transverse to the probe beam. Similarly the complex amplitude of the reference with real amplitude $r(\xi, \eta)$ and phase $\phi_r(\xi, \eta)$ can be written as

$$R(\xi, \eta) = r(\xi, \eta) e^{i\phi_r(\xi, \eta)} \quad (4)$$

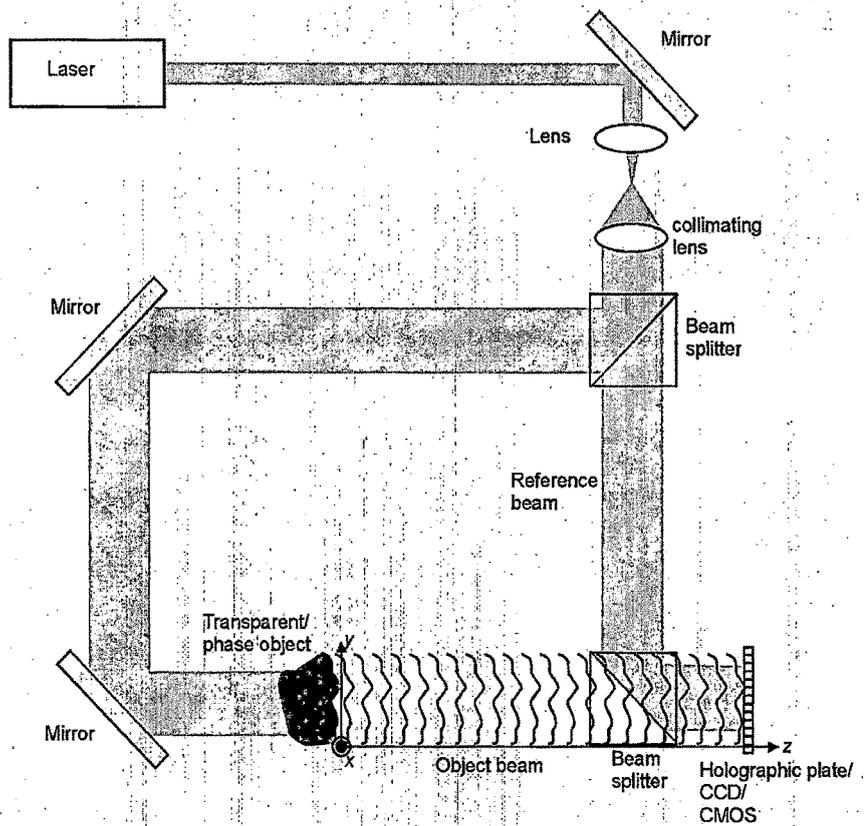


Fig. 2.1: Mach-Zehnder hologram recording setup for transparent objects

The intensity at the hologram is the square of the complex amplitude [16], which is now the addition of the object wave and reference wave complex amplitudes

$$\begin{aligned}
 I(\xi, \eta) &= |O(\xi, \eta) + R(\xi, \eta)|^2 = (O(\xi, \eta) + R(\xi, \eta))(O(\xi, \eta) + R(\xi, \eta))^* \quad (5) \\
 &= R(\xi, \eta)R^*(\xi, \eta) + O(\xi, \eta)O^*(\xi, \eta) + O(\xi, \eta)R^*(\xi, \eta) + R(\xi, \eta)O^*(\xi, \eta)
 \end{aligned}$$

Here * denotes complex conjugation.

2.5 Hologram reconstruction

In a digital holography setup, a digital camera replaces the film [2.17]. This not only offers a convenient and versatile setup, but also provides a detector that has much more linear response with respect to the incident intensity than a photographic film.

The recording setup is practically unchanged compared to the film-based version, but the reconstruction process is quite different. The digital hologram captured by the digital camera is processed by a computer program to extract the intensity and phase distributions corresponding to the object [2.18].

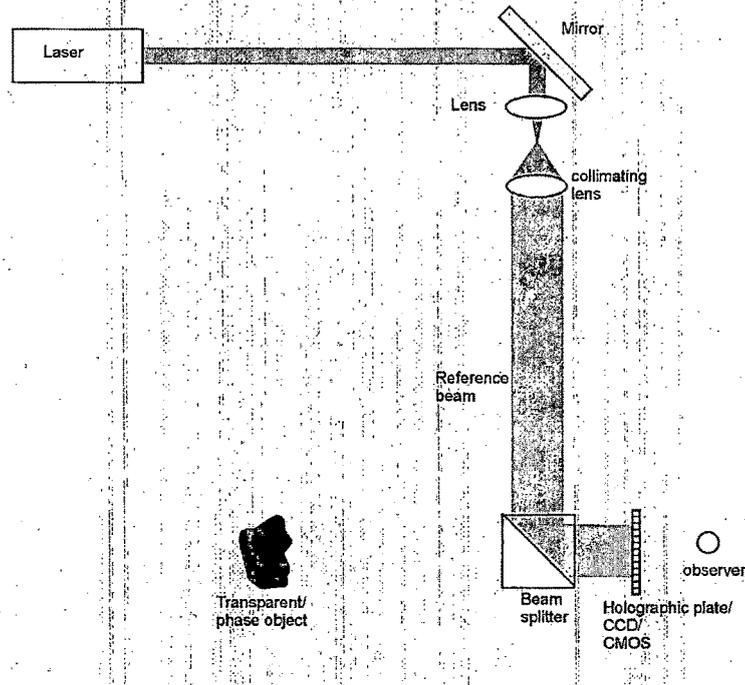


Fig. 2.2: Hologram reconstruction setup for transparent objects

The amplitude transmittance of the developed hologram is proportional to the recorded intensity of the fringe pattern and is

$$h(\xi, \eta) = h_0 + \beta \tau I(\xi, \eta) \quad (6)$$

where β is a constant, τ is the exposure time and h_0 is the amplitude transmission of the unexposed plate. $h(\xi, \eta)$ is also called the hologram function. In digital holography using CCDs as recording medium h_0 can be neglected.

Holograms are reconstructed by illuminating it with the reference wave (Fig. 2.2). For hologram reconstruction the amplitude transmission has to be multiplied with the complex amplitude of the reconstruction (reference) wave giving rise to

$$R(\xi, \eta)h(\xi, \eta) = R(\xi, \eta)[h_0 + \beta\tau I(\xi, \eta)] \quad (7)$$

Using Eq. (3), Eq. (4) and Eq. (5). Eq. (7) can be modified to

$$\begin{aligned} R(\xi, \eta)h(\xi, \eta) &= \\ R(\xi, \eta) & \left[h_0 + \beta\tau \{ R(\xi, \eta)R^*(\xi, \eta) + O(\xi, \eta)O^*(\xi, \eta) + O(\xi, \eta)R^*(\xi, \eta) + R(\xi, \eta)O^*(\xi, \eta) \} \right] \\ &= [h_0 + \beta\tau(r^2 + o^2)]R(\xi, \eta) + \beta\tau r^2 O(\xi, \eta) + \beta\tau R^2(\xi, \eta)O^*(\xi, \eta) \end{aligned} \quad (8)$$

This equation represents the amplitude of the diffracted reference beam from the micro-interference structures of the hologram. There are three main terms. The first term is the reference wave multiplied by a constant. This represents nothing but the un-diffracted reference wave transmitted from the hologram structures and does not contain any information about the object. The second term does contain object information. In fact it reconstructs the object exactly at the same position where the original object was placed. So this provides the virtual image of the object. The multiplicative constant just changes the brightness of the reconstructed virtual image. Also it can be seen that since $O(\xi, \eta)$ is the complex amplitude of the object wavefront, it contains phase information about the object. Since phase is dependent on the path length, the reconstructed virtual image of the object contains the depth or 3D profile of the original object. This reconstructed wave is indistinguishable from the original object wave. An observer sees a three-dimensional image which exhibits all the effects of perspective and depth of focus. The third term produces a distorted real image of the object. For off-axis holography the virtual image, the real image and the undiffracted wave are spatially separated.

In digital holography, the digital hologram should be multiplied by the digital reference wave, which must be a replica of the experimental reference wave [2.19]. When the distance of propagation equals the optical path distance d between the object and the hologram plane in the recording process, the real image comes into focus.

Fig.2.3 illustrates the relationship between visual image and real image in the reconstruction process. In optical reconstructions, the virtual image appears at the position of the original object and real image is formed at an equal distance d but in the opposite direction.

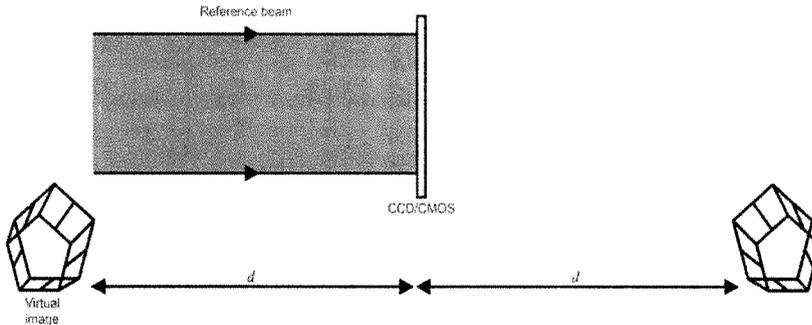
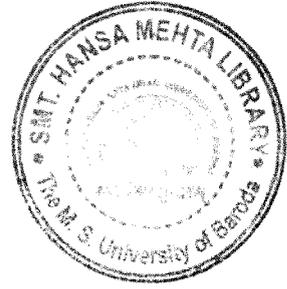


Fig. 2.3: Holographic reconstruction showing the virtual and real images

2.6 Holographic Interferometry

Using holography the coherent wave fields reflected from the two states (or two time instances) of the object can be compared to produce an image of the object on which interference fringes are superposed. This is called holographic interferometry. This method finds application in many areas like finding object deformations, mapping of refractive index variations, mapping temperature and density changes etc. This is a non-contact method with very high sensitivity that maps optical path length changes with resolutions of several nanometers.

To achieve holographic interferometry, the wavefront interacting with the two states of the object should be interfered. So the use of coherent radiation is necessary and lasers with high coherence length are the most commonly used as light sources in holography. In conventional holography, two methods can be used for achieving holographic interferometry. These are illustrated in Fig. 2.4. The first of the holographic interferometric technique involves recording the holograms of the two object states in the same photo-sensitive medium. This is called double exposure holographic interferometry. When the coherent reference beam impinges, the hologram, it gets diffracted from the microstructures. These microstructures can be imagined as randomly oriented diffraction grating. The reference beam so gets diffracted in the direction of the object location according to Eq. (8). But since the recorded double exposure hologram contains fringes due to two states of the object, the two diffracted object wavefront will result. One diffracted in the direction of the

original object state and the second one propagating in the direction of the final state of the object. Since both were reconstructed using the same coherent source, these two will interfere to produce a set of interference fringes, superposed on the image of the object. The deformation or change in the optical path length could be determined from this interference fringes.

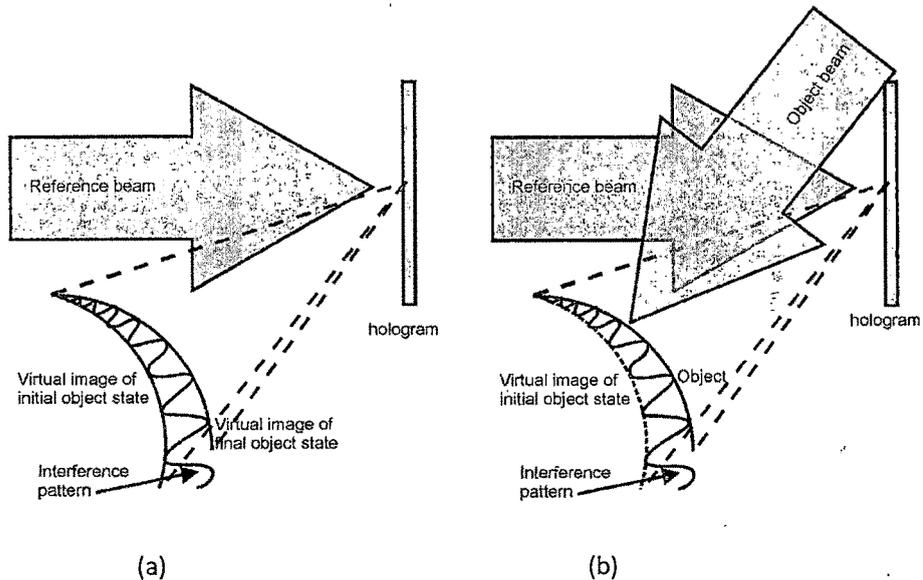


Fig. 2.4: Holographic interferometry (a) Double exposure, (b) Real-time

The second technique for holographic interferometry is called real-time method [2.21]. Here the hologram of the initial state of the object is recorded on the photo-sensitive medium. If it is a photographic plate, it is then developed and fixed at the original position. The original object and the object beam are not removed in this technique. The reference beam is now made to impinge the hologram. This produces an object wavefront propagating in the direction of the original object. This will interfere with the light from the object. This will produce a set of interference fringes on the original object itself, depending upon the amount of optical path length change between the holographic image and the original object.

In digital holography since the reconstruction is achieved numerically, the double exposure technique is applicable. Advantage of digital holographic interferometry is that due to availability of phase information due to numerical reconstruction, the

phase change of the object wavefront between exposures is directly available, making the deformation analysis simple without the need for additional optics for phase stepping.

Let us find the mathematical relationship between the initial and final object states for holographic interferometry. Assuming the complex amplitude of the object wavefront in the initial state be

$$O_i(x, y) = A_0 \exp[i\phi_i(x, y)] \quad (9)$$

where $\phi_i(x, y)$ is the phase of the object wavefront in the initial state. Here it is assumed that the reflectivity of the object is same across its surface. Due to change in the object (which could be due to deformation, change in refractive index, change in temperature etc) the phase of the object wavefront changes from $\phi_i(x, y)$ to $\phi_i(x, y) + \Delta\phi(x, y)$. Here $\Delta\phi(x, y)$ represents the change in phase due to the optical path length change and is called the interference phase. The new complex amplitude of the object wavefront can therefore be written as

$$O_f(x, y) = A_0 \exp[i\{\phi_i(x, y) + \Delta\phi(x, y)\}] \quad (10)$$

The intensity of the holographic interference pattern is the absolute square of the sum of the complex amplitudes in Eq. (9) and Eq. (10).

$$I(x, y) = |O_i(x, y) + O_f(x, y)|^2 = [O_i(x, y) + O_f(x, y)] \times [O_i(x, y) + O_f(x, y)]^* \quad (11)$$

Substituting the exponential representation of O_i and O_f in Eq. (11) one gets

$$I(x, y) = 2A_0^2 [1 + \cos \Delta\phi(x, y)] \quad (12)$$

This provides the relationship between the intensity and the interference phase distribution [23]. As mentioned in conventional holography it is not possible to compute the phase distribution without phase stepping since 'cos' is an even function. But digital holography overcomes this problem by numerical reconstruction.

The relation between the optical path length and the obtained interference phase distribution for near normal illumination directions is

$$\Delta\phi(x, y) = \frac{2\pi}{\lambda} \Delta(x, y) \quad (13)$$

It is assumed here that the propagation direction of the object beam along the z -axis. $\Delta(x, y)$ is the optical path length change along the z direction. So by knowing the phase change or the interference phase the optical path length change can be computed.

2.7 Numerical reconstruction of digital holograms

Conventional holograms recorded on photographic plates are reconstructed by illuminating it by the reference wave or they can be said as to be reconstructed optically. But in digital holography, the holograms are stored as digital images in a computer. So they have to be reconstructed numerically [24]. Numerical reconstruction of digital holograms is achieved by assumption of scalar diffraction theory [4, 5]. Here the object wavefront is reconstructed by simulating the diffraction of the reference wavefront occurring at the microstructures of the hologram. This involves the use of diffraction integrals. In the course of this work, digital holograms were reconstructed using either Fresnel-Krichoff diffraction integral or Angular spectrum propagation diffraction integral. Both are based on the scalar diffraction theory [6].

2.7.1 Wave propagation

Wave propagation has always been a topic of common interest especially in acoustics and optics. Study of wave propagation in these fields requires sharing of common wave equation for scalar fields. Since light is considered to be a wave, one needs to simulate its propagation and understand the interaction mechanism with media. In most cases experiments need to be performed to verify the theory.

The mathematical derivations in these sections are important for presenting theoretical optical analysis in digital holography. It also permits approximation of the image plane by propagating the wavefront from the hologram to various distances. Further approximations of Kirchhoff's integral then lead to the classical Fresnel and Fraunhofer diffraction integrals [2.6, 2.7]. Although the intensity of the diffracted

field is the primary concern of many digital holography applications, emphasis is on both the amplitude and the phase of the diffracted field that are important for this work. The next section deals with the meaning of a scalar wave equation and the mathematical solution of the scalar wave equation by Green's function that can lead to the well known Kirchhoff diffraction integral solution [2.6, 2.11].

2.7.2 The scalar wave equation

The first step to understand the object wave field propagation is to create a model of the processes involved, make reasonable assumptions and approximations and making it appropriate for a numerical propagation.

The explanation of propagation of light begins with the fundamentals of electromagnetism, Maxwell's equations [2.11, 2.12]. Light being an electromagnetic, its behavior can be derived from Maxwell's fundamental electromagnetic equations. Hence, if one attempt to take into account everything that takes place during the interaction of an object with a high frequency electromagnetic wave, one has to solve the Maxwell's equations.

The electromagnetic wave consists of the time varying electric and magnetic fields. In a homogeneous isotropic medium like free space or a lens with constant refractive index, the electric and magnetic field vectors form a right-handed orthogonal triad with the direction of propagation. The Maxwell's equations bring together electric and magnetic fields. The simplified Maxwell's equations in vacuum can be written as

$$\nabla \cdot \epsilon \vec{E} = 0 \quad (\text{Gauss' law for electric field}) \quad (14)$$

$$\nabla \cdot \mu \vec{H} = 0 \quad (\text{Gauss' law for magnetic field}) \quad (15)$$

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \quad (\text{Faraday's law}) \quad (16)$$

$$\nabla \times \vec{H} = \epsilon \frac{\partial \vec{E}}{\partial t} \quad (\text{Amperes' circuital law with Maxwell's correction}) \quad (17)$$

where \vec{E} represents electric field and \vec{H} represents magnetic field vector, μ and ϵ are the permeability and permittivity respectively of the medium (in vacuum, they are

represented by μ_0 and ϵ_0). The permittivity and permeability of free space (vacuum)

is linked to the speed of light in a vacuum by the equation $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$

These equations explain the electric and magnetic field properties [2.24]. Although the field's travel perpendicular to each other, they are in phase, travelling at the speed of light in vacuum.

The easiest approach to understand a scalar wave equation is to consider Maxwell's vector wave equation in a linear, homogeneous, isotropic and non-dispersive medium like vacuum [2.25]. Maxwell's equations can then be used to derive a wave equation for both E and H fields.

$$\nabla^2 \vec{E} - \frac{n^2}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad (18)$$

$$\nabla^2 \vec{H} - \frac{n^2}{c^2} \frac{\partial^2 \vec{H}}{\partial t^2} = 0 \quad (19)$$

where n is the refractive index. All the three components of \vec{E} , i.e. E_x , E_y and E_z as well as the three components of \vec{H} , i.e. H_x , H_y and H_z obey the wave equations. The relationship between the wave vectors and their components is

$$\vec{E} = E_x \hat{e}_x + E_y \hat{e}_y + E_z \hat{e}_z \quad \text{and} \quad \vec{H} = H_x \hat{e}_x + H_y \hat{e}_y + H_z \hat{e}_z \quad (20)$$

where \hat{e} 's represent the unit vector. Since a linear, isotropic, homogenous and non-dispersive medium is considered, all of the scalar field components satisfy the wave equation and so the vector wave equation becomes a scalar wave equation [2.26]. Therefore for example E_x satisfies the scalar wave equation,

$$\nabla^2 E_x - \frac{n^2}{c^2} \frac{\partial^2 E_x}{\partial t^2} = 0 \quad (21)$$

All the other components obey such wave equation. Therefore the behavior all the components can be expressed through a generalized scalar wave equation

$$\nabla^2 u(P, t) - \frac{n^2}{c^2} \frac{\partial^2 u(P, t)}{\partial t^2} = 0 \quad (22)$$

where $u(P, t)$ represents any of the scalar fields [2.27]. For a linear, homogenous, isotropic, non-dispersive medium all the field components behave in similar manner and can be described by a single scalar wave equation. But this will not be case when a inhomogeneous, anisotropic, non-linear or dispersive medium is considered. But for our case we are going to consider propagation of light (electromagnetic radiation) in air and for large objects, so scalar wave equation will suffice.

If the disturbance (light) at position P and time t can be represented by a scalar function since we know that it satisfies the scalar wave equation. This can be written as

$$u(P, t) = \text{Re} \left[A(P) e^{i(\phi(P) - \omega t)} \right] \quad (23)$$

where $A(P)$ and $\phi(P)$ are the amplitude and phase of the wave at position P , ω is the angular frequency of the optical radiation and 'Re' represents the real part. This can further be written as

$$u(P, t) = \text{Re} \left[U(P) e^{-i\omega t} \right] \quad (24)$$

$$\text{where } U(P) = A(P) e^{i\phi(P)} \quad (25)$$

is the complex amplitude. This optical disturbance should satisfy the scalar wave equation given by Eq. (22) at each source free point. Since we assume that the frequency of the radiation is known, $U(P)$ can will provide information about the disturbance at point P . So $U(P)$ must then satisfy the time independent equation (Helmholtz equation)

$$\nabla^2 U + k^2 U = 0 \quad (26)$$

where k is the wave vector which depends upon the wavelength of the radiation.

2.7.3 Fresnel-Kirchoff diffraction integral

The solution of this equation gives the value of the amplitude and phase of the field anywhere in the space at a given moment of time, provided that there are well-defined boundary conditions [2.28]. This equation is applicable to a linearly polarized monochromatic EM wave, which is being used in the system under consideration.

Howsoever simple the Helmholtz scalar equation might be compared to the Maxwell's equations; it is still difficult to find a solution, because this case deals with a surface EM distribution rather than volume distribution [2.29]. Making one more approximation and using Green's theorem the EM field anywhere in space can be expressed as a function of the EM field distribution on a finite surface as [2.30]

$$U(P_0) = \frac{1}{4\pi} \iint_S \left(\frac{\partial U}{\partial n} G - \frac{\partial G}{\partial n} U \right) ds \quad (27)$$

where $U(P_0)$ is the EM field at a particular point in space that has to be calculated provided the complex amplitude U at another plane and G is the corresponding Green's function, $\frac{\partial}{\partial n}$ denotes a directional derivative on the surface taken in the normal outward direction to the surface.

When light strikes an object, the scattered EM field is completely characterized by the illuminated surface. The Kirchhoff's formula of diffraction by a planar screen can be used to find the field at the point P_0 [2.31]. The planar screen is especially of importance in digital holography, since holograms can be considered as apertures on a planar screen. The diffraction from a planar aperture or screen is shown in Fig. 2.5. The closed surface S is chosen to consist of two parts. A plane surface S_1 , lying directly behind the diffracting screen, joined and closed by a large spherical cap, S_2 , of radius R centered at the observation point P_0 .

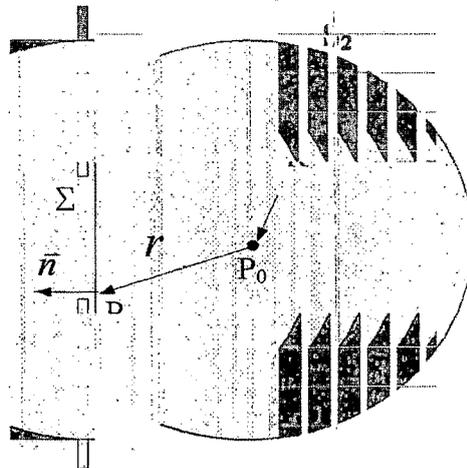


Fig. 2.5: illustration of Kirchhoff formulation of diffraction by a plane screen

The total closed surface S is simply the sum of S_1 and S_2 . The problem of diffraction of light can be illustrated by having an aperture in an infinite opaque screen. In Fig. 2.5, a wave disturbance is assumed to impinge on the screen and the aperture from the left, and the field at the point P_0 behind the aperture is to be calculated. Using as a unit amplitude expanding spherical wave as the Green's function of the problem, assuming Sommerfeld radiation condition and adopting the assumptions that (i) the across the opening U and its normal derivative are same behind after the aperture and (ii) for the screen U and its normal derivative are zero, Eq. (27) reduces to

$$U(P_0) = \frac{1}{4\pi} \iint_S \left(\frac{\partial U}{\partial n} G - \frac{\partial G}{\partial n} U \right) ds \quad (28)$$

For a point illumination source and P_0 large distance away from the planar screen, this formula reduces to

$$U(P_0) = \frac{1}{i\lambda} \iint_S U(P_1) \frac{e^{-ikr}}{r} \frac{1}{2} (1 + \cos \theta) ds \quad (29)$$

where θ is the angle between the normal to the aperture plane and r . This equation is known as the integral theorem of Helmholtz and Kirchhoff, which plays an important role in the development of the scalar theory of diffraction [33].

2.7.4 Numerical reconstruction of digital holograms by Fresnel Transform

The diffraction of a light wave at an aperture perpendicular to the incoming beam (or a hologram perpendicular to the incoming beam) as shown in Fig. 2.6 can be described by the Fresnel–Kirchhoff integral [2.7, 2.8].

$$U(x, y) = \frac{1}{i\lambda} \iint_H h(\xi, \eta) R(\xi, \eta) \frac{e^{-ikr}}{r} \frac{1}{2} (1 + \cos\theta) d\xi d\eta \quad (30)$$

$h(\xi, \eta)$ is the hologram function or just the hologram and $R(\xi, \eta)$ is the complex amplitude of the reference wave and r is the distance between a point in the hologram plane and point in the reconstruction plane, which not necessarily be the image plane. For plane reference beam the R is nothing but an array of real numbers. The distance r can be written in terms of the co-ordinates of the two planes as

$$r = \sqrt{(x - \xi)^2 + (y - \eta)^2 + z^2} \quad (31)$$

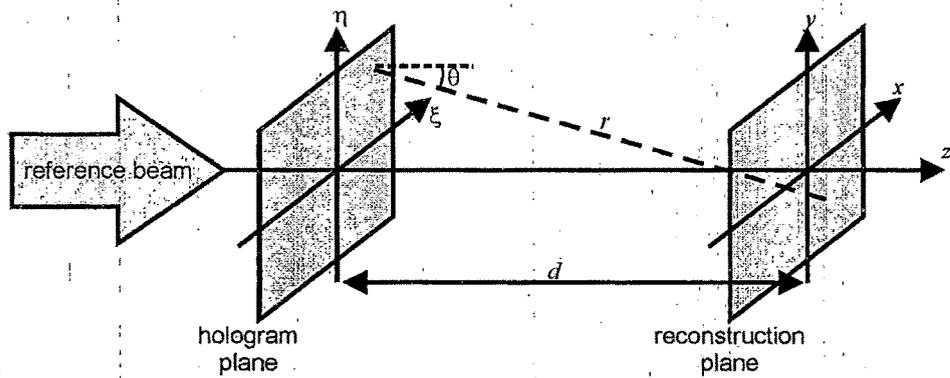


Fig. 2.6: Coordinate system for digital hologram reconstruction using Fresnel-Kirchoff diffraction integral.

The diffracted wavefront from the hologram plane is propagated to a plane which is situated at distance d from the hologram plane but parallel to it. This plane could be the image plane. It can be seen that propagation using diffraction integral involves a complex function. Numerical reconstruction yields the complex amplitude at the desired plane. It can then be used to determine the intensity or phase of the object wavefront. For large propagation distances compared to the hologram size, r can be replaced by the first two terms of its Taylor series expansion.

$$r = d + \frac{(x-\xi)^2}{2d} + \frac{(y-\eta)^2}{2d} \quad (32)$$

Also when d is large the angle θ between the normal and r becomes very small so that $\cos(\theta)$ can be approximated by 1. Putting the value of r and $\cos(\theta)$ into Eq. (30) and re-arranging the terms yield

$$U(x, y) = \frac{1}{i\lambda d} \exp(-ikd) \exp\left[-i\frac{k}{2d}(x^2 + y^2)\right] \iint_{\#} h(\xi, \eta) R(\xi, \eta) \exp\left[-i\frac{k}{2d}(\xi^2 + \eta^2)\right] \exp\left[i\frac{k}{d}(\xi x + \eta y)\right] d\xi d\eta \quad (33)$$

This is called the Fresnel approximation or Fresnel transformation [2.33]. From this equation it can be seen that it enables the reconstruction of the wave field at any plane behind the hologram by changing the reconstruction distance d . If one looks closely at the Eq. (33), it can be seen that it is nothing but the Fourier transform of the first three terms in the integral. This means that the numerical reconstruction using Fresnel transform could be implemented by the use of Fourier Transforms. Since we have propagated through a distance d behind the hologram, it basically reconstructs the real image of the object [2.34]. In conventional holography also the real image can be directly seen by placing a screen on the path of the diffracted beam. Virtual image that lies behind the hologram plane is a source of diverging wavefronts. Reconstruction of virtual image requires the use of a lens, like the eye of the observer. In digital holography also the reconstruction of virtual image can be achieved by using a digital lens with focal length equal to half the propagation distance placed just after the hologram to reconstruct the virtual image with unit magnification. This is represented in Fig. 2.7.

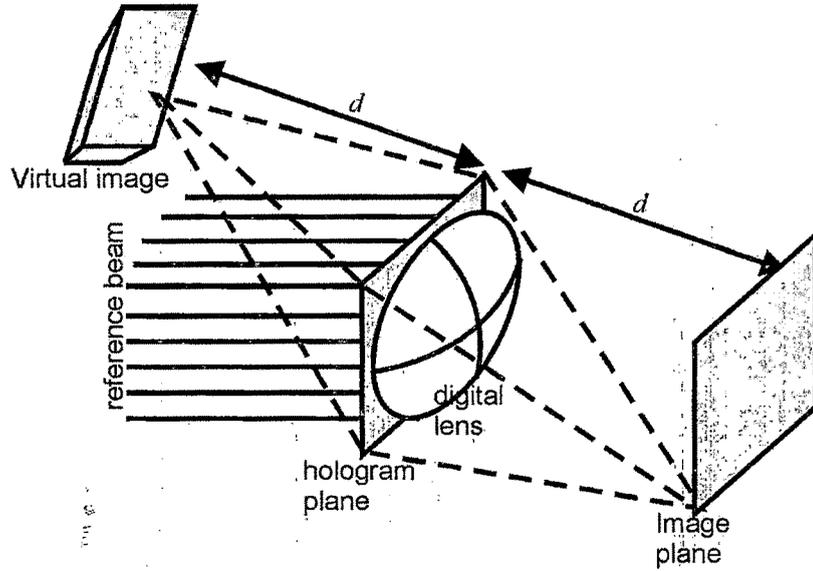


Fig. 2.7: Reconstruction of virtual image in digital holography using Fresnel-Kirchoff integral and a digital lens

A converging lens which produces a unit magnification virtual image ($d=2f$, where f is focal length of the digital lens) can be mathematically represented as

$$L(\xi, \eta) = \exp\left[i\frac{k}{d}(\xi^2 + \eta^2)\right] \quad (34)$$

So now at the complex amplitude at the imaging for the reconstructed virtual image is given by,

$$U(x, y) = \frac{1}{i\lambda d} \exp(-ikd) \exp\left[-i\frac{k}{2d}(x^2 + y^2)\right] \iint_{\#} h(\xi, \eta) R(\xi, \eta) \exp\left[i\frac{k}{2d}(\xi^2 + \eta^2)\right] \exp\left[i\frac{k}{d}(\xi x + \eta y)\right] d\xi d\eta \quad (35)$$

Now digital imaging devices has discrete sized pixels along both ξ and η directions. This requires a digitization of the Fresnel Transform. Noting that $k=2\pi/\lambda$ and using the following substitutions

$$\nu = \frac{x}{\lambda d} \text{ and } \mu = \frac{y}{\lambda d} \quad (36)$$

Eq. (33) becomes

$$U(\nu, \mu) = \frac{1}{i\lambda d} \exp\left(-i\frac{2\pi}{\lambda}d\right) \exp\left[-i\pi\lambda d(\nu^2 + \mu^2)\right] \iint_H h(\xi, \eta) R(\xi, \eta) \exp\left[-i\frac{\pi}{\lambda d}(\xi^2 + \eta^2)\right] \exp\left[i2\pi(\xi\nu + \eta\mu)\right] d\xi d\eta \quad (37)$$

The Fresnel transform can be digitized since the hologram function given by Eq. (6) in the case of digital holography is matrix of size $N \times N$ sampled at regular steps of $\Delta\xi$ and $\Delta\eta$ in the ξ and η axes. This makes the pixel sizes in the horizontal and vertical directions of the recording device $\Delta\xi$ and $\Delta\eta$ respectively. Using the discrete values at the hologram plane, the integral equation can be converted into a sum

$$U(m, n) = \frac{1}{i\lambda d} \exp\left(-i\frac{2\pi}{\lambda}d\right) \exp\left[-i\pi\lambda d(m^2\Delta\nu^2 + n^2\Delta\mu^2)\right] \sum_{k=1}^N \sum_{l=1}^N h(k, l) R(k, l) \exp\left[-i\frac{\pi}{\lambda d}(k^2\Delta\xi^2 + l^2\Delta\eta^2)\right] \exp\left[i2\pi(k\Delta\xi m\Delta\nu + l\Delta\eta n\Delta\mu)\right] \quad (38)$$

The relationship between the sensor pixel size and the reconstructed image pixel size can be written as

$$\Delta x = \frac{\lambda d}{N\Delta\xi} \text{ and } \Delta y = \frac{\lambda d}{N\Delta\eta} \quad (39)$$

Using these in Eq. (38)

$$U(m, n) = \frac{1}{i\lambda d} \exp\left(-i\frac{2\pi}{\lambda}d\right) \exp\left[-i\pi\lambda d\left(\frac{m^2}{N^2\Delta\xi^2} + \frac{n^2}{N^2\Delta\eta^2}\right)\right] \sum_{k=1}^N \sum_{l=1}^N h(k, l) R(k, l) \exp\left[-i\frac{\pi}{\lambda d}(k^2\Delta\xi^2 + l^2\Delta\eta^2)\right] \exp\left[i2\pi\left(\frac{km}{N} + \frac{ln}{N}\right)\right] \quad (40)$$

Eq. (40) represents the discrete Fresnel transform. The complex amplitude at the reconstruction plane is calculated by taking the Fourier transform of the product of the digitized hologram function, the matrix representing the reference beam and the exponential function. This can be implemented numerically using FFT algorithm [2.35].

2.7.5 Reconstruction using Angular Spectrum Propagation (ASP) integral

Angular spectrum method has many applications, but one of utmost importance is in the optical holography. Based on angular spectrum descriptions of the scattered field, E. Wolf provided the relationship between the detected scattered signal and the object, which later became the foundation of optical diffraction tomography [2.36, 2.37]. Angular spectrum propagation assumption of scalar diffraction theory used in this method is more appropriate for short distance non-paraxial propagation which occurs especially near the object. Since the angular spectrum of a complex field is evaluated through 2-D Fourier transform, it is natural to take advantage of the computation efficiency of Fast Fourier Transform (FFT) [2.35].

Provided the complex disturbance at any other point, the calculation of the complex disturbance U at any observation point in space is the core of the scalar diffraction theory. This could be accomplished in three ways, out of which two employ the use of Green's functions. These methods yield the Fresnel-Kirchoff diffraction formula and the Rayleigh-Sommerfeld diffraction formula [2.31-2.33]. These method also assumes that the field is observed at a distance much larger than the wavelength ($k \gg 1/r$), where r is the distance between the aperture and observation point) from the aperture.

Another approach is to formulate the scalar diffraction theory from the framework of angular spectrum. The angular-spectrum representation of plane waves has been

widely used to study the propagation properties of the wave fields in homogeneous media. If the complex field distribution of a monochromatic disturbance is Fourier-analyzed across any plane, the various spatial Fourier components can be identified as plane waves traveling in different directions away from that plane [2.37]. The field amplitude at any other point (or across any other parallel plane) can be calculated by adding the contributions of these plane waves, taking due account of the phase shifts they have undergone during propagation [2.38].

Suppose a wave is propagating along the positive z direction. Here (x,y) is the transverse plane. Let the complex amplitude be given by $U(x, y, 0)$ at $z = 0$. The aim of the diffraction theory is to compute the complex amplitude $U(x, y, z)$ at a point (x, y, z) at distance z to the right of the original plane (see Fig. 2.8).

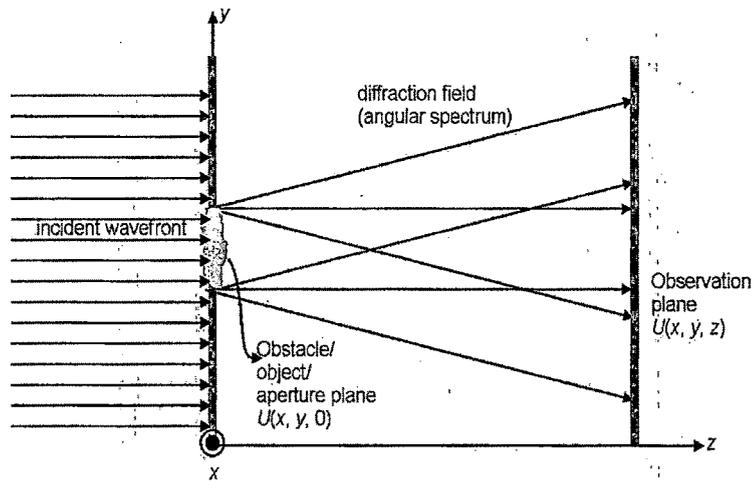


Fig. 2.8: Propagation of wavefront using ASP diffraction integral

The Fourier transform of $U(x, y, 0)$, i.e., its spectrum at $z = 0$ is given by

$$\hat{U}(f_x, f_y; 0) = \iint_{-\infty}^{\infty} U(x, y, 0) e^{-[i2\pi(f_x x + f_y y)]} dx dy \quad (41)$$

where f_x, f_y are the spatial frequencies in the x and y directions respectively. Basically, Fourier transform decomposes [40] the complicated function into a series

of simple complex functions. So, the inverse Fourier transform of Eq. (41) provides the complex amplitude at $(x, y, 0)$.

$$U(x, y, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{U}(f_x, f_y; 0) e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (42)$$

If the complex amplitude at a plane parallel to the (x, y) plane but a distance d from the $(x, y, 0)$ plane is known, then the angular spectrum at this plane is

$$\hat{U}(f_x, f_y; d) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, z) e^{-[i2\pi(f_x x + f_y y)]} dx dy \quad (43)$$

If the relationship between the angular spectrums at $(x, y, 0)$ and (x, y, d) can be found, the effect of wave propagation on the angular spectrum can be determined. Now the complex amplitude at (x, y, d) can be written as

$$U(x, y, d) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{U}(f_x, f_y; d) e^{i2\pi(f_x x + f_y y)} df_x df_y \quad (44)$$

This complex amplitude should satisfy the Helmholtz equation (Eq. (15)). Using U given by Eq. (44) in Helmholtz equation give by Eq. (26)

$$\frac{d^2}{dz^2} \hat{U}(f_x, f_y; d) + k^2 [1 - \lambda^2 f_x^2 - \lambda^2 f_y^2] \hat{U}(f_x, f_y; d) = 0 \quad (45)$$

where λf_x and λf_y are the direction cosines in the x and y directions respectively. A solution to the above equation may be written as

$$\hat{U}(f_x, f_y; d) = \hat{U}(f_x, f_y; 0) e^{ik\sqrt{1 - \lambda^2 f_x^2 - \lambda^2 f_y^2} d} \quad (46)$$

When the direction cosines have the relationship $\lambda^2 f_x^2 + \lambda^2 f_y^2 \leq 1$, the effect of propagation over a distance d is simply a change in the relative phases of the various components of the angular spectrum. Each plane wave component travels at a different angle and hence travels different distances between the parallel planes in introducing relative phase delays. When the direction cosines follow the

relationship $\lambda^2 f_X^2 + \lambda^2 f_Y^2 > 1$, λf_X and λf_Y are no longer interpretable as direction cosines and Eq. (46) can then be written as

$$\hat{U}(f_X, f_Y, d) = \hat{U}(f_X, f_Y, 0) e^{-\mu d} \quad (47)$$

where μ is a real number. So these wave components are rapidly attenuated by propagation phenomenon. These components are called evanescent waves. When the distance d is larger than several wavelengths, the evanescent components will not be present as they are attenuated rapidly. By inverse Fourier transforming Eq. (46) one gets

$$U(x, y, d) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{U}(f_X, f_Y, 0) e^{ik\sqrt{1-\lambda^2 f_X^2 - \lambda^2 f_Y^2} d} e^{i2\pi(f_X x + f_Y y)} df_X df_Y \quad (48)$$

Therefore the disturbance observed at (x, y, d) can be written in terms of the initial angular spectrum and a phase factor called the free space propagation function. So by knowing the complex amplitude of the wavefront at any plane, the wavefront at any other plane can be obtained by propagating the angular spectrum using Eq. (48). This method has some advantages over Fresnel-Kirchoff integral such that there is no assumption on the propagation distance between the planes under consideration other than that the field should not be observed very near the aperture producing the diffraction [2.41]. Whereas the Fresnel or paraxial approximation used in the Fresnel-Kirchoff diffraction integral needs the observation planes be appreciable distance apart. So the angular spectrum propagation (ASP) will be useful in the case where the distance between the hologram and image plane are small. For reconstruction of digital holograms using ASP integral the hologram the product of the matrix of hologram transfer function and the matrix representing the reference beam is Fourier transformed [2.42, 2.43]. This will provide the frequency spectrum of the scattered reference beam from the hologram micro-structures. As we have seen from Eq. (8), this will contain three terms corresponding to the un-diffracted reference beam, the virtual object and the real object. Since the frequency spectrum is available, it is possible to separate out the different beams and perform a filtering in the frequency space. The filtered spectrum containing the information about the object only can be obtained and propagated to the image plane. So in the reconstructions there will not be any overlap between the three beams. Numerical reconstructions in digital

holography using ASP integral can be done by first Fourier transforming the product of the hologram function and the reference beam and then applying a frequency domain filter to extract information due to object along and then propagating it to the image plane. This can be written as

$$U(x, y, z) = \mathfrak{F}^{-1} \left\{ \text{filt}[\mathfrak{F}\{U(x, y, 0)\}] e^{ik\sqrt{1-\lambda^2 f_x^2 - \lambda^2 f_y^2} z} \right\} \quad (49)$$

$$\text{Here } U(x, y, 0) = R(x, y, 0) \times h(x, y, 0) \quad (50)$$

2.8 Computing intensity and phase of the object wavefront

The complex amplitude obtained from the either Fresnel transform or angular spectrum integral, is an array of complex numbers. This can be used to determine either the intensity or the phase of the object wavefront [2.44]. The intensity is calculated from the absolute square of the Complex amplitude distribution of the object wavefront and can be written as

$$I(x, y) = |U(x, y)|^2 \quad (51)$$

Phase distribution of the object wavefront is calculated from the angle that complex amplitude makes with the real axis

$$\varphi(x, y) = \arctan \frac{\text{Im}[U(x, y)]}{\text{Re}[U(x, y)]} \quad (52)$$

The value of the computed phase ranges from $-\pi$ to π and is called wrapped phase distribution. Unwrapping has to be applied to obtain the continuous phase distribution.

As described holographic interferometry is able to compare wavefront existing at two time instances. But digital holography [45] has an added advantage. It can directly compare the phase of the object wavefronts existing at two time instances. This is done by reconstructing the object wavefront for the two states separately and computing the phase using Eq. (52).

$$\varphi_1(x, y) = \arctan \frac{\text{Im}[U_1(x, y)]}{\text{Re}[U_1(x, y)]} \quad (53)$$

$$\phi_2(x, y) = \arctan \frac{\text{Im}[U_2(x, y)]}{\text{Re}[U_2(x, y)]} \quad (54)$$

where U_1 is the complex amplitude of the object wavefront in the first state and the U_2 is its complex amplitude distribution in the second state and θ_1 and θ_2 are the corresponding phases computed. The phase difference or the interference phase can then be computed using

$$\Delta\phi(x, y) = \begin{cases} \phi_1(x, y) - \phi_2(x, y) & \text{if } \phi_1(x, y) \geq \phi_2(x, y) \\ \phi_1(x, y) - \phi_2(x, y) + 2\pi & \text{if } \phi_1(x, y) < \phi_2(x, y) \end{cases} \quad (55)$$

From this equation, we can calculate interference phase directly from the digital holograms. The relationship of the phase change to the change in optical path length is given by

$$\Delta\phi(x, y) = \frac{2\pi}{\lambda} \Delta l(x, y) \quad (56)$$

This optical path length change could be brought about by object deformation $\Delta l(x, y)$, refractive index change $\Delta n(x, y)$, temperature change $\Delta T(x, y)$ etc. So from the interference phase information on these parameters can be obtained.

2.9 Codes for numerical reconstruction of holograms

Digital holograms can be reconstructed using either the Fresnel transform or the Angular spectrum integral. Both involve taking Fourier transforms. Computer codes for hologram reconstruction using these approaches were written in MATLAB. These were tested on several holograms for the reconstruction of object intensity and phase.

Other than this MATLAB code for computation of phase difference was also developed. Other developed codes included tomographic inversion by Abel inversion.

References

- [2.1] D. Gabor, "A new microscopic principle", *Nature*, **161**, 777 (1948).
- [2.2] P. Hariharan, *Basics of Holography*, Cambridge University Press, Cambridge (2002).
- [2.3] T. Kreis, *Handbook of Holographic Interferometry*, Wiley-VCH, Weinheim (2005).
- [2.4] U. Schnars and W. Juptner, "Direct recording of holograms by a CCD-target and numerical reconstruction" *Appl. Opt.* **33**, 179 (1994).
- [2.5] U. Schnars and W. Juptner, "Digital recording and reconstruction of holograms in hologram interferometry and shearography", *Appl. Opt.* **33**, 4373 (1994).
- [2.6] Eugene Hecht, *Optics*, Fourth Edition. San Francisco: Addison Wesley, 2002.
- [2.7] U. Schnars, "Direct phase determination in hologram interferometry with use of digitally recorded holograms," *J. Opt. Soc. Am. A*, **11**, 2011 (1994).
- [2.8] U. Schnars and W. P. Jueptner, "Digital recording and numerical reconstruction of holograms," *Meas. Sci. Technol.* **13**, R85 (2002).
- [2.9] E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," *Opt. Lett.* **24**, 291 (1999).
- [2.10] M. Centurion, Y. Pu, Z. Liu, D. Psaltis, T. W. Hänsch, "Holographic recording of laser-induced plasma", *Opt. Lett.*, **29**, 7, (2004).
- [2.11] J Goodman, *Introduction to Fourier Optics*, McGraw-Hill companies Inc, New York (1996)
- [2.12] P. Hariharan, *Basics of Interferometry*, Academic Press, New York (2007).
- [2.13] James E. Harvey, "Nonparaxial Scalar Diffraction Theory", SPIE Press
- [2.14] Rastogi, P. K, "Holographic Interferometry Principles and Methods", Germany: Springer-Verlag. (1995).
- [2.15] Anand Asundi and Vijay Raj Singh, "Amplitude and Phase Analysis in Digital Dynamic Holography", *Opt. Lett.*, **31**, 2420, (2006).

- [2.16] U. Schnars and W. P. Jueptner, *Digital Holography: Digital Recording, Numerical Reconstruction and Related Techniques*, Springer, Berlin (2005).
- [2.17] P. Hariharan, "Basics of Holography", (2004).
- [2.18] K. Creath. "Phase measurement interferometry techniques", in *Progress in Optics*, volume 26. E. Wolf, ed. (Elsevier, 1988), 1988.
- [2.19] O. Coquoz, C. Depeursinge, R. Conde, and F. Taleblou, "Numerical reconstruction of images from endoscopic holograms," in *14th Annual International Conference of the IEEE-EMBS*, pp. 338-339, IEEE, (Paris), 1992.
- [2.20] H.I. Bjelkhagen and E. Mirlis. Color holography to produce highly realistic three-dimensional images. *Appl. Opt.*, **47**, 123, 2008.
- [2.21] Vest, C. M., and D. W. Sweeney, "Application of Holographic Interferometry to Nondestructive Testing", *International Advances in Nondestructive Testing*, **5**: 17-29, (1977).
- [2.22] Caussignac, J. M, "Application of Holographic Interferometry to the Study of Structural Deformations In Civil Engineering", *First European Conference on Optics Applied to Metrology*. 136: 136-142.(1978).
- [2.23] J. H. Massig, "Digital off-axis holography with a synthetic aperture," *Opt. Lett.*, vol. 27, no. 24, pp. 2179-2181, 2002.
- [2.24] J. Durnin, "Exact solutions for non diffracting beams, the scalar theory," *J. Opt. Soc. Am. A* **4**, 651-654 (1987).
- [2.25] J. Durnin, J. J. Miceli Jr. and J. H. Eberly, "Diffraction free beams," *Phys. Rev. Lett.* **58**, 1499 -1501 (1987).
- [2.26] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast Fourier approach," *J. Opt. Soc. A*, **15**, 857-867 (1998).
- [2.27] G. B. Esmer and L. Onural, "Simulation of scalar optical diffraction between arbitrarily oriented planes", In *First International Symposium on Control, Communications and Signal Processing*, pages 225-228, (2004).
- [2.28] L. Yu, Y. An and L. Cai, "Numerical reconstruction of digital holograms with variable viewing angles," *Opt. Express* **10**, 1250-1257 (2002).

- [2.29] David J. Griffiths. Introduction to Electrodynamics. Prentice-Hall, Inc., 1989.
- [2.30] M. T. Heideman, D. H. Johnson, and C. S. Burrus, "Gauss and the history of the fast fourier transform", *Archive for History of Exact Sciences*, 34(3):265–277, September 1985.
- [2.31] M. Hain, W. von Spiegel, M. Schmiedchen, T. Tschudi, and B. Javidi, "3D integral imaging using diffractive fresnel lens arrays", *Optics Express*, 13:315–326, 2005.
- [2.32] L. Onural and H.M. Ozaktas. Signal processing issues in diffraction and holographic 3DTV. In Proc. EURASIP 13th European Signal Processing Conference, 2005.
- [2.33] E. CuChe, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," *Opt. Lett.* **24**, 291-293 (1999).
- [2.34] P. Ferraro, S. De Nicola, A. Finizio, G. Pierattini, G. Coppola, "Recovering image resolution in reconstructing digital off-axis holograms by Fresnel-transform method," *Appl. Phys. Lett.* **85**, 2709-2711 (2004).
- [2.35] K. Matsushima, H. Schimmel, F. Wyrowski, "Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves," *J. Opt. Soc. A.* **20**, 1755-1762 (2003).
- [2.36] S. De Nicola, A. Finizio, G. Pierattini, P. Ferraro, and D. Alfieri, "Angular spectrum method with correction of anamorphism for numerical reconstruction of digital holograms on tilted planes", *Optics Express*, 13(24):9935–9940, November 2005.
- [2.37] P. Cloetens, W. Ludwig, J. Baruchel, D. Van Dyke, J. Van Landuyt, J. P. Guigay, and M. Schlenker, "Holotomography quantitative phase tomography with micrometer resolution using hard synchrotron radiation x rays," *Appl. Phys. Lett.* **75** 2912-2914 (1999).
- [2.38] A.V. Bronnikov, "Reconstruction formulas in phase-contrast tomography," *Optics Communications* 171 239-244 (1999).
- [2.40] A.V. Bronnikov, "Theory of quantitative phase-contrast computed tomography," *J. Opt. Soc. Am. A* 19 472-480 (2002).

- [2.41] M.A. Anastasio, D. Shi, F. De Carlo, and X. Pan, "Analytic image reconstruction in local phase-contrast tomography", *Phys. Med. Biol.* 49 121-144 (2004).
- [2.42] I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," *Opt. Lett.* **22**, 1268-1270 (1997)
- [2.43]. A. Groso, R. Abela, M. Stampanoni, "Implementation of a fast method for high resolution phase contrast tomography", *Optics Express*, to be published (2006).
- [2.44] C. Wagner, S. Seebacher, W. Osten, and W. Juptner. Digital recording and numerical reconstruction of lensless Fourier holograms in optical metrology. *Appl. Opt.*, **38**(22), 4812–4820, August 1999.
- [2.45] Y. Frauel, B. Javidi, "Neural network for three-dimensional object recognition based on digital holography" *Opt. Lett.*, **26**, 19-21 (2001).