

---

# OPTICAL SCANNING HOLOGRAPHY WITH MATLAB®

# OPTICAL SCANNING HOLOGRAPHY WITH MATLAB®

TING-CHUNG POON

Bradley Department of Electrical and Computer Engineering, Virginia Tech,  
Blacksburg, Virginia 24061.



Dr. Ting-Chung Poon  
Virginia Tech  
Bradley Dept. Electrical and Computer Engineering  
Blacksburg, VA 24061  
USA  
tcpoon@vt.edu

Library of Congress Control Number: 2007921127

ISBN-10: 0-387-36826-4      e-ISBN-10: 0-387-36826-4  
ISBN-13: 978-0-387-36826-9      e-ISBN-13: 978-0-387-68851-0

Printed on acid-free paper.

© 2007 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1

springer.com

## **Dedication**

*This book is dedicated to*

*Eliza (M.S., Iowa 1980),  
Christina (B.S., Cornell 2004), and  
Justine (B.S., Virginia Tech 2007).*

# Contents

Preface	ix
1. Mathematical Background and Linear System. . . . .	1
1.1 Fourier Transformation . . . . .	1
1.2 Linear and Invariant Systems . . . . .	10
1.2.1 Linearity and Invariance . . . . .	10
1.2.2 Convolution and Correlation Concept. . . . .	14
2. Wave Optics and Holography . . . . .	21
2.1 Maxwell's Equations and Homogeneous Vector Wave Equation . . . . .	21
2.2 Three-Dimensional Scalar Wave Equation . . . . .	25
2.2.1 Plane Wave Solution . . . . .	25
2.2.2 Spherical Wave Solution . . . . .	27
2.3 Scalar Diffraction Theory . . . . .	29
2.3.1 Fresnel Diffraction . . . . .	33
2.3.2 Diffraction of a Square Aperture. . . . .	36
2.4 Ideal Lens, Imaging Systems, Pupil Functions and Transfer Functions. . . . .	40
2.4.1 Ideal Lens and Optical Fourier Transformation . . . . .	40
2.4.2 Coherent Image Processing . . . . .	43
2.4.3 Incoherent Image Processing . . . . .	45
2.5 Holography . . . . .	49
2.5.1 Fresnel Zone Plate as a Point-Source Hologram . . . . .	49
2.5.2 Off-Axis Holography . . . . .	57
2.5.3 Digital Holography . . . . .	60

3. Optical Scanning Holography: Principles .....	65
3.1 Principle of Optical Scanning .....	65
3.2 Optical Heterodyning .....	69
3.3 Acousto-Optic Frequency Shifting .....	72
3.4 Two-Pupil Optical Heterodyne Scanning Image Processor .....	75
3.5 Scanning Holography .....	81
3.6 Physical Intuition to Optical Scanning Holography .....	92
4. Optical Scanning Holography: Applications .....	97
4.1 Scanning Holographic Microscopy .....	97
4.2 Three-Dimensional Holographic TV and 3-D Display .....	106
4.3 Optical Scanning Cryptography .....	117
5. Optical Scanning Holography: Advances .....	135
5.1 Coherent and Incoherent Holographic Processing .....	135
5.2 Single-Beam Scanning vs. Double-Beam Scanning .....	141
5.3 PSF Engineering .....	143
Index .....	149

## Preface

This book serves two purposes. The first is to succinctly cover the necessary mathematical background and wave optics that pertain to Fourier optics and holography. The second is to introduce optical scanning holography (OSH) - a form of electronic (or digital) holography - to the readers, and to provide them with experience in modeling the theory and applications utilizing MATLAB®.

*Optical Scanning Holography with MATLAB®* consists of tutorials (with numerous MATLAB examples throughout the text), research material, as well as new ideas and insights that are useful for engineering or physics students, scientists, and engineers working in the fields of Fourier optics, optical scanning imaging and holography. The book is self-contained and covers the basic principles of OSH. Thus, this book will be relevant for years to come. The writing style of this book is geared towards undergraduate seniors or first-year graduate-level students in the fields of engineering and physics. The material covered in this book is suitable for a one-semester course in Fourier optics, optical scanning imaging and holography.

Optical scanning holography is a highly sophisticated technology that consists of numerous facets and applications. It is a real-time (or on-the-fly) holographic recording technique that is based on active optical heterodyne scanning. It is a relatively new area in electronic holography and will potentially lead science and technology to many novel applications such as cryptography, 3-D display, scanning holographic microscopy, 3-D pattern recognition and 3-D optical remote sensing. The main purpose of this book is to introduce optical scanning holography to the readers in a manner that will allow them to feel comfortable enough to explore the technology on

their own - possibly even encourage them to begin implementing their own set-ups in order to create novel OSH applications. Optical scanning holography is generally a simple yet powerful technique for 3-D imaging, and it is my aspiration that this book will stimulate further research of optical scanning holography and its various novel applications.

I have incorporated some of the material from this book into my short course entitled “Optical Scanning Holography” at SPIE Photonics West, in lectures given at the Institute of Optical Sciences (IOS), which is now known as the Department of Optics and Photonics, National Central University (NCU), Taiwan, and also at the Department of Electronics and Computer Science, Nihon University, Japan. The book was finally completed during my time as a visiting professor at Nihon University. I want to take this opportunity to thank my host, Professor Hiroshi Yoshikawa, for his hospitality and arranging a spacious office for me that allowed me to concentrate on the last phase of this book. I would also like to thank Professor Hon-Fai Yau of NCU for providing me with some early opportunities (when the book was still in its infancy) to “rehearse” my optical scanning holography lectures at IOS.

I would like to thank my wife, Eliza, and my children, Christina and Justine, for their encouragement, patience, and love. This book is dedicated to them. In addition, I would also like to thank Christina Poon for reading the manuscript and providing comments and suggestions for improvement.



## Chapter 1

### Mathematical Background and Linear Systems

#### 1.1 Fourier Transformation

In electrical engineering, we are most concerned with a signal as a function of time,  $f(t)$ . The signal in question could be a voltage or a current. The forward temporal *Fourier transform* of  $f(t)$  is given as

$$\mathcal{F}\{f(t)\} = F(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-j\omega t) dt, \quad (1.1-1a)$$

where the transform variables are time,  $t$  [second], and temporal radian frequency,  $\omega$  [radian/second]. In Eq. (1.1a),  $j = \sqrt{-1}$ . The inverse Fourier transform is

$$\mathcal{F}^{-1}\{F(\omega)\} = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp(j\omega t) d\omega. \quad (1.1-1b)$$

In optics, we are most interested in dealing with a two-dimensional (2-D) signal. Examples include images or the transverse profile of an electromagnetic or optical field at some plane of spatial variables  $x$  and  $y$ . Hence, the two-dimensional spatial *Fourier transform* of a signal  $f(x, y)$  is given as [Banerjee and Poon (1991), Poon and Banerjee (2001)]

$$\mathcal{F}_{xy}\{f(x, y)\} = F(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp(jk_x x + jk_y y) dx dy, \quad (1.1-2a)$$

and the inverse Fourier transform is

$$\begin{aligned} & \mathcal{F}_{xy}^{-1}\{F(k_x, k_y)\} \\ &= f(x, y) \\ &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(k_x, k_y) \exp(-jk_x x - jk_y y) dk_x dk_y, \end{aligned} \quad (1.1-2b)$$

where the transform variables are spatial variables,  $x, y$  [meter], and spatial radian frequencies,  $k_x, k_y$  [radian/meter].  $f(x, y)$  and  $F(k_x, k_y)$  are a Fourier

transform pair and the statement is symbolically represented by

$$f(x, y) \Leftrightarrow F(k_x, k_y).$$

Note that the definitions for the forward and inverse transforms [see Eqs. (1.1-2a) and (1.1-2b)] are consistent with the engineering convention for a traveling wave, as explained in *Principles of Applied Optics* [Banerjee and Poon (1991)]. Common properties and examples of 2-D Fourier transform appear in the Table below.

Table 1.1 Properties and examples of some two-dimensional Fourier Transforms.

Function in $(x, y)$	Fourier transform in $(k_x, k_y)$
1. $f(x, y)$	$F(k_x, k_y)$
2. $f(x - x_0, y - y_0)$	$F(k_x, k_y) \exp(jk_x x_0 + jk_y y_0)$
3. $f(ax, by)$ ; $a, b$ complex constants	$\frac{1}{ ab } F(\frac{k_x}{a}, \frac{k_y}{b})$
4. $f^*(x, y)$	$F^*(-k_x, -k_y)$
5. $\partial f(x, y)/\partial x$	$-jk_x F(k_x, k_y)$
6. $\partial^2 f(x, y)/\partial x \partial y$	$-k_x k_y F(k_x, k_y)$
7. <b>delta function</b>	
$\delta(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\pm jk_x x \pm jk_y y} dk_x dk_y$	1
8. 1	$4\pi^2 \delta(k_x, k_y)$
9. <b>rectangle function</b>	<b>sinc function</b>
$\text{rect}(x, y) = \text{rect}(x) \text{rect}(y)$ , where $\text{rect}(x) = \begin{cases} 1, &  x  < 1/2 \\ 0, & \text{otherwise} \end{cases}$	$\text{sinc}(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}) = \text{sinc}(\frac{k_x}{2\pi}) \text{sinc}(\frac{k_y}{2\pi})$ , where $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$
10. <b>Gaussian function</b>	<b>Gaussian function</b>
$\exp[-\alpha(x^2 + y^2)]$	$\frac{\pi}{\alpha} \exp[-\frac{k_x^2 + k_y^2}{4\alpha}]$

### Example 1.1 Fourier Transform of $\text{rect}(x, y)$ plus MATLAB

The one-dimensional (1-D) rectangular function or simply *rect function*,  $\text{rect}(x/a)$ , is given by

$$\text{rect}(x/a) = \begin{cases} 1, & |x| < a/2 \\ 0, & \text{otherwise} \end{cases}, \quad (1.1-3a)$$

where  $a$  is the width of the function. The function is shown in Fig. 1.1a). The two-dimensional version of the rectangular function is given by

$$\text{rect}(x/a, y/b) = \text{rect}(x/b) \text{rect}(y/b). \quad (1.1-3b)$$

Figure 1.1b) and 1.1c) show the three-dimensional plot and the gray scale plot of the function. In the gray scale plot, we have assumed that an amplitude of 1 translates to “white” and an amplitude of zero to “black”. Therefore, from the definition of Eq. (1.1-3b), the white area is  $a \times b$ .

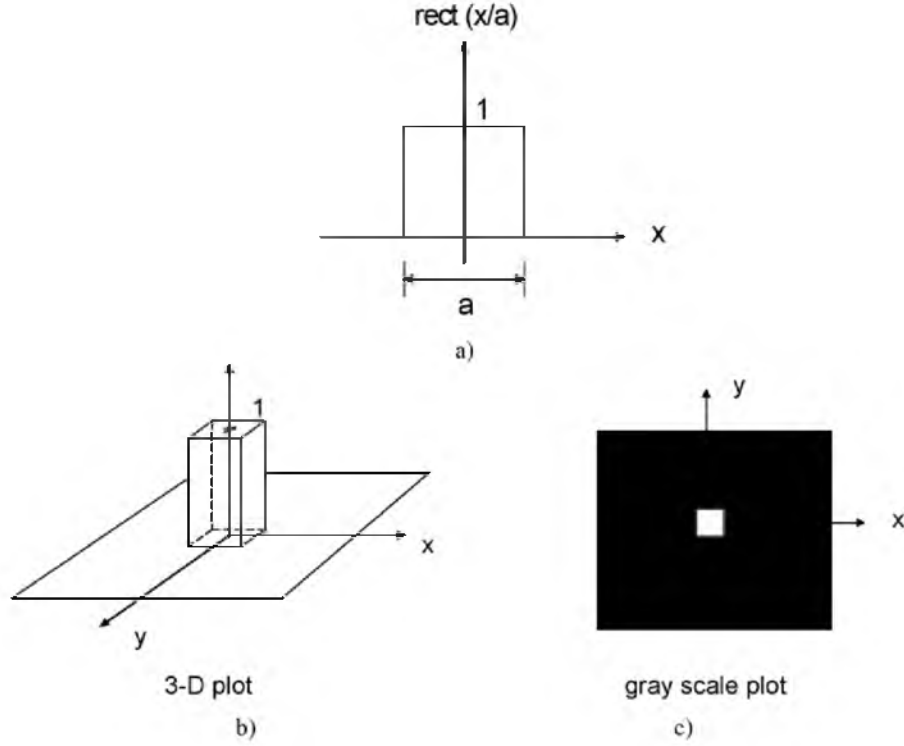


Fig. 1.1 Rect function.

To find the Fourier transform of the 2-D rectangular function, we simply evaluate the integral given by Eq. (1.1-2a) by recognizing that  $f(x, y) = \text{rect}(x/a, y/b)$ . Therefore, we write

$$\begin{aligned} \mathcal{F}_{xy}\{f(x, y)\} &= \mathcal{F}_{xy}\{\text{rect}(x/a, y/b)\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{rect}(x/a, y/b) \exp(jk_x x + jk_y y) dx dy. \end{aligned} \quad (1.1-4)$$

Since  $\text{rect}(x/a, y/b)$  is a *separable function* [see Eq. 1.1-3b)], we re-write Eq. (1.1-4) as follows:

$$\begin{aligned} &\mathcal{F}_{xy}\{\text{rect}(x/a, y/b)\} \\ &= \int_{-\infty}^{\infty} \text{rect}(x/a) \exp(jk_x x) dx \times \int_{-\infty}^{\infty} \text{rect}(y/b) \exp(jk_y y) dy \\ &= \int_{-a/2}^{a/2} 1 \exp(jk_x x) dx \times \int_{-b/2}^{b/2} 1 \exp(jk_y y) dy. \end{aligned} \quad (1.1-5)$$

By writing the last step, Eq. (1.1-5), we have used the definition of the rectangular function given by Eq. (1.1-3a). We can now evaluate Eq. (1.1-5) by using

$$\int \exp(cx) dx = \frac{1}{c} \exp(cx). \quad (1.1-6)$$

Therefore,

$$\int_{-a/2}^{a/2} 1 \exp(jk_x x) dx = a \operatorname{sinc}\left(\frac{ak_x}{2\pi}\right), \quad (1.1-7)$$

where  $\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$  is defined as the *sinc function*. Table 1.2 shows the m-file for plotting the sinc function and its output is shown in Fig. 1.2. Note that the sinc function has zeros at  $x = \pm 1, \pm 2, \pm 3, \dots$

Table 1.2 Plot\_sinc.m: m-file for plotting the sinc function.

```
-----
%Plot_sinc.m Plotting of sinc(x) function
x=-5.5:0.01:5.5;
sinc=sin(pi*x)./(pi*x);
plot(x,sinc)
axis([-5.5 5.5 -0.3 1.1])
grid on
xlabel('x')
ylabel('sinc (x)')
-----
```

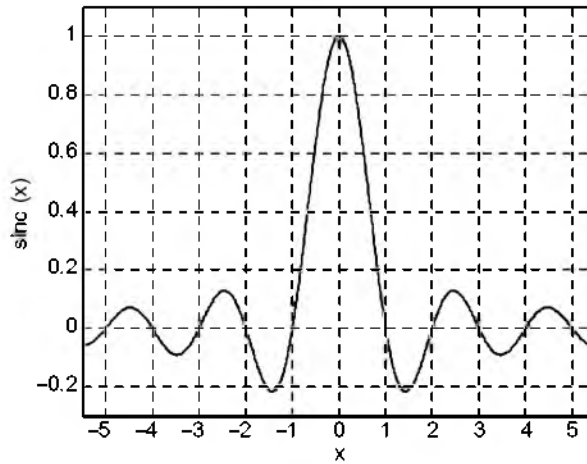


Fig. 1.2 Sinc function.

To complete the original problem of determining the Fourier transform of a rect function, we take advantage of the result of Eq. (1.1-7); Eq. (1.1-5)

becomes

$$\begin{aligned}\mathcal{F}_{xy}\{\text{rect}(x/a, y/b)\} &= a\text{sinc}\left(\frac{ak_x}{2\pi}\right)\text{sinc}\left(\frac{bk_y}{2\pi}\right) \\ &= a\text{sinc}\left(\frac{ak_x}{2\pi}, \frac{bk_y}{2\pi}\right).\end{aligned}\quad (1.1-8a)$$

Hence, we may write

$$\text{rect}(x/a, y/b) \Leftrightarrow a\text{sinc}\left(\frac{ak_x}{2\pi}, \frac{bk_y}{2\pi}\right). \quad (1.1-8b)$$

Note that when the width of the rect function along  $x$  is  $a$ , the first zero along  $k_x$  is  $k_{x,0} = 2\pi/a$ . Figure 1.3 shows the transform pair of Eq. (1.1-8b). The top figures are 2-D gray-scale plots, and the bottom figures are line traces along the horizontal axis through the center of the top figures. These figures are generated using the m-file shown in Table 1.3 where  $M = 11$ . For this value of  $M$ ,  $a = 0.0429$  units of length and the first zero  $k_{x,0} = 146.23$  radian/(unit of length). Note that the area of display in the  $x$ - $y$  plane has been scaled to 1 unit of length by 1 unit of length.

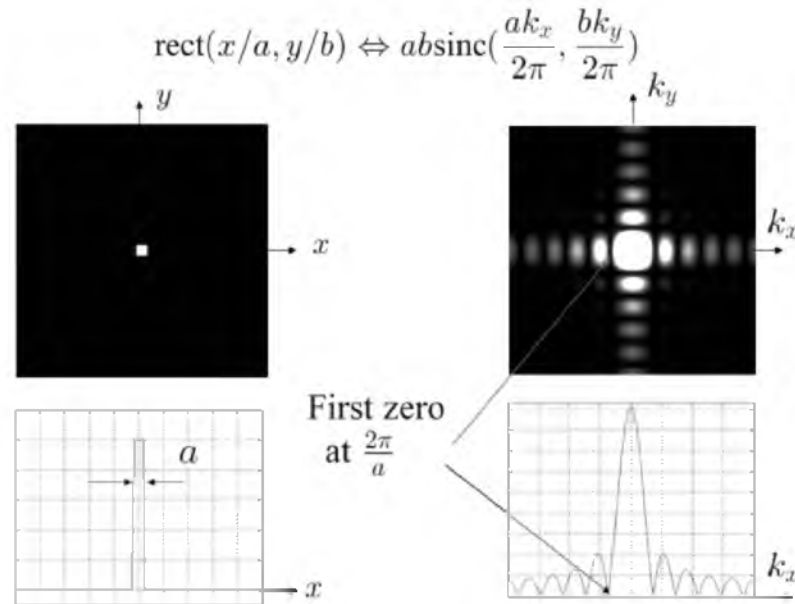


Fig. 1.3 Rect function and its Fourier transform.

Table 1.3 fft2Drect.m: m-file for 2-D Fourier transform of  $\text{rect}(x/a, y/b)$ .

---

```
%fft2Drect.m %Simulation of Fourier transformation of a 2-D rect function
%
clear
```

```

L=1; %display area is L by L, L has unit of length
N=256; % number of sampling points
dx=L/(N-1); % dx : step size

% Create square image, M by M square, rect(x/a,y/a), M=odd number
M=input('M (size of rect(x/a,y/a), enter odd numbers from 3-33)=');
a=M/256;
kx0=2*pi/a;
sprintf('a = %0.5g[unit of length]',a)
sprintf('kx0 (first zero)= %0.5g[radian/unit of length]',kx0)
R=zeros(256); %assign a matrix (256x256) of zeros
r=ones(M); % assign a matrix (MxM) of ones
n=(M-1)/2;
R(128-n:128+n,128-n:128+n)=r;
%End of creating square input image M by M

%Axis Scaling
for k=1:256
    X(k)=1/255*(k-1)-L/2;
    Y(k)=1/255*(k-1)-L/2;
    %Kx=(2*pi*k)/((N-1)*dx)
    %in our case, N=256, dx=1/255
    Kx(k)=(2*pi*(k-1))/((N-1)*dx)-((2*pi*(256-1))/((N-1)*dx))/2;
    Ky(k)=(2*pi*(k-1))/((N-1)*dx)-((2*pi*(256-1))/((N-1)*dx))/2;
end

%Image of the rect function
figure(1)
image(X+dx/2,Y+dx/2,255*R);
title('rect function: gray-scale plot')
xlabel('x')
ylabel('y')
colormap(gray(256));
axis square

%Computing Fourier transform
FR=(1/256)^2*fft2(R);
FR=fftshift(FR);

% plot of cross-section of rect function
figure(2)
plot(X+dx/2,R(:,127))
title('rect function: cross-section plot')
xlabel('x')
ylabel('rect(x/a)')
grid
axis([-0.5 0.5 -0.1 1.2])

%Centering the axis and plot of cross-section of transform along kx
figure(3)
plot(Kx-pi/(dx*(N-1)),10*abs(FR(:,127)))
title('Square-absolute value of Fourier transform of rect function: cross-section plot')
xlabel('kx')
ylabel('|a*b*sinc(a*kx/2pi)|')

```

```

axis([-800 800 0 max(max(abs(FR)))*10.1])
grid

%Mesh the Fourier transformation
figure(4);
mesh(Kx,Ky,(abs(FR)).^2)
title('Square-absolute value of Fourier transform of rect function: 3-D plot,scale
arbitrary')
xlabel('kx')
ylabel('ky')
axis square

%Image of the Fourier transformation of rectangular function
figure(5);
gain=10000;
image(Kx,Ky,gain*(abs(FR)).^2/max(max(abs(FR))).^2)
title('Square-absolute value of Fourier transform of the rect function: gray-scale plot')
xlabel('kx')
ylabel('ky')
axis square
colormap(gray(256))
-----

```

### Example 1.2 MATLAB Example: Fourier Transform of Bitmap Images

When the 2-D function or image is given with a bitmap file, we can use the m-file given in Table 1.4 to find its Fourier transform. Figure 1.4a) is the bitmap image used when the image file of the size is 256 by 256. It is easily generated with Microsoft® Paint. Figure 1.4b) is the corresponding absolute value of the transformed image.

Table 1.4 fft2Dbitmap\_image.m: m-file for 2-D Fourier transform of bitmap image.

```

-----
%fft2Dbitmap_image.m
%Simulation of Fourier transformation of bitmap images
clear

I=imread('triangle.bmp','bmp'); %Input bitmap image
I=I(:,:,1);
figure(1) %displaying input
colormap(gray(255));
image(I)
axis off
FI=fft2(I);
FI=fftshift(FI);
max1=max(FI);
max2=max(max1);
scale=1.0/max2;
FI=FI.*scale;
figure(2) %Gray scale image of the absolute value of transform

```

```
colormap(gray(255));
image(10*(abs(256*FI)));
axis off
```

---



a) Bitmap image of a triangle.



b) Displaying the absolute value of the transform of a).

Fig. 1.4 Bitmap image and its transform generated using the m-file in Table 1.4.

### Example 1.3 Delta Function and its Transform

The *delta function*,  $\delta(x)$ , is one of the most important functions in the study of systems. We can define the delta function as follows:

$$\delta(x) = \lim_{a \rightarrow 0} \left\{ \frac{1}{a} \text{rect}\left(\frac{x}{a}\right) \right\}. \quad (1.1-9)$$

The situation is shown graphically in Fig. 1.5.

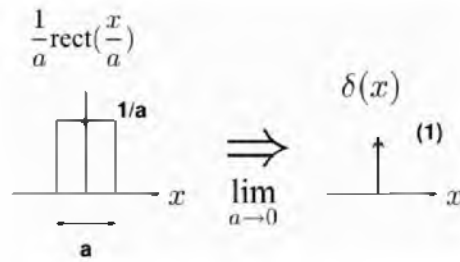


Fig. 1.5 Illustration of the definition of the delta function graphically.

The delta function has three important properties, which are listed as follows:

**Property #1: Unit Area**

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1. \quad (1.1-10a)$$



The delta function has a unit area (or strength), which is denoted by a “(1)” beside the arrow, as shown in Fig. 1.5. This unit area property is clearly demonstrated by the definition illustrated on the left hand side of Fig. 1.5. The area is always a unity regardless of the value of  $a$ .

#### Property #2: Product Property

$$f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0). \quad (1.1-10b)$$

The result of this property can be confirmed graphically by the illustration shown in Fig. 1.6 where an arbitrary function,  $f(x)$ , is shown to be overlapped with the offset delta function,  $\delta(x - x_0)$ , located at  $x = x_0$ . The product of the two functions is clearly equal to  $f(x_0)$  multiplied by  $\delta(x - x_0)$ . Therefore, the result has become an offset delta function with its strength given by  $f(x_0)$ .

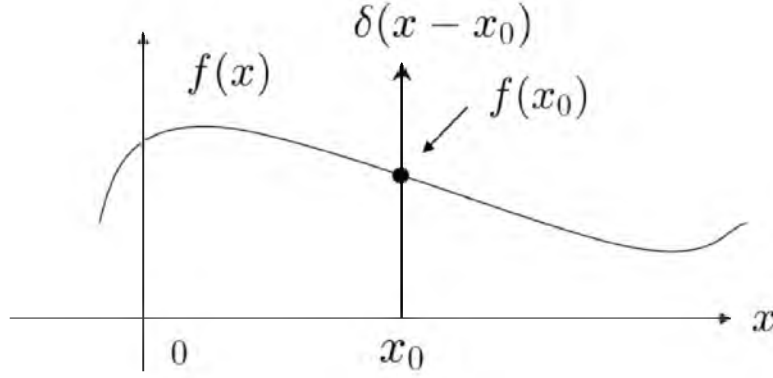


Fig. 1.6 Illustrating the result of the Product Property.

#### Property #3: Sampling Property

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0). \quad (1.1-10c)$$

To obtain the result above, we simply use Properties #1 and #2. From Eq. (1.1-10c) and by using Property #2, we have

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx &= \int_{-\infty}^{\infty} f(x_0)\delta(x - x_0)dx \\ &= f(x_0) \int_{-\infty}^{\infty} \delta(x - x_0)dx = f(x_0), \end{aligned}$$

where we have used Property #1 to obtain the last step of the result. Equation (1.1-10c) is known as the *sampling property* because the delta function selects, or samples, a particular value of the function,  $f(x)$ , at the location of the delta function (i.e.,  $x_0$ ) in the integration process.

While a 1-D delta function is called an impulse function in electrical engineering, the 2-D version of a delta function,  $\delta(x, y) = \delta(x)\delta(y)$ , represents an idealized point source of light in optics. According to Eq. (1.1-2a), the 2-D Fourier transform of  $\delta(x, y)$  is given by

$$\begin{aligned}\mathcal{F}_{xy}\{\delta(x, y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) \exp(jk_x x + jk_y y) dx dy \\ &= \int_{-\infty}^{\infty} \delta(x) \exp(jk_x x) dx \int_{-\infty}^{\infty} \delta(y) \exp(jk_y y) dy \\ &= 1,\end{aligned}$$

where we have used the sampling property of the delta function to evaluate the above integrals. Figure 1.7 shows the 2-D delta function as well as its corresponding Fourier transform.

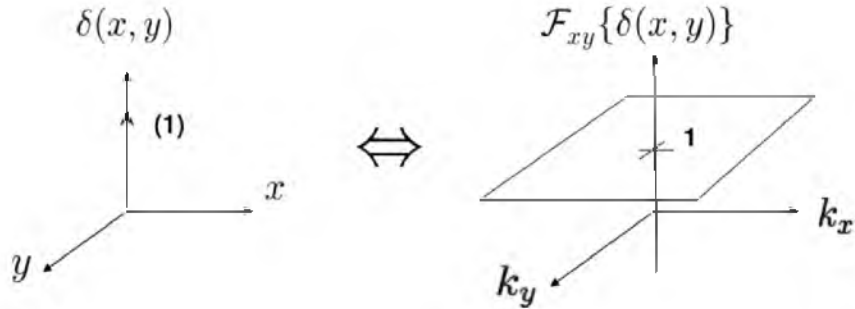


Fig. 1.7 Two-dimensional delta function and its Fourier transform.

## 1.2 Linear and Invariant Systems

### 1.2.1 Linearity and Invariance

A system is defined as the mapping of an input or set of inputs into an output or set of outputs. A system is linear if *superposition* applies. For a single-input – single-output system, if an input  $f_1(t)$  gives an output of  $g_1(t)$ , and if another input  $f_2(t)$  gives an output of  $g_2(t)$ , then superposition means if the input is given by  $af_1(t) + bf_2(t)$ , the system's output is  $ag_1(t) + bg_2(t)$ , where  $a$  and  $b$  are some constants. The situation of a *linear system* is further illustrated in Fig. 1.8.

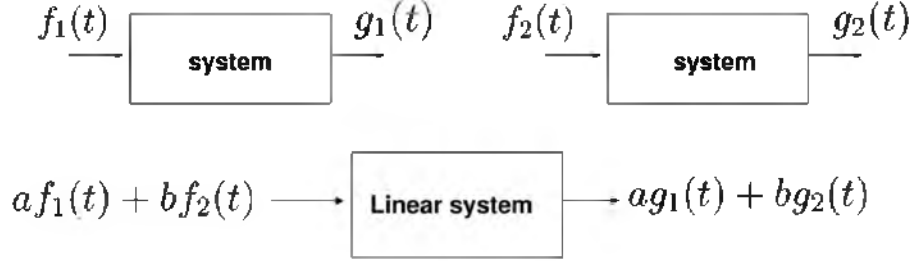


Figure 1.8 Linear system.

Systems with parameters that do not change with time are *time-invariant systems*. Consequently, a time delay in the input results in a corresponding time delay in the output. This property of the system is shown graphically in Fig. 1.9, where  $t_0$  is the time delay.

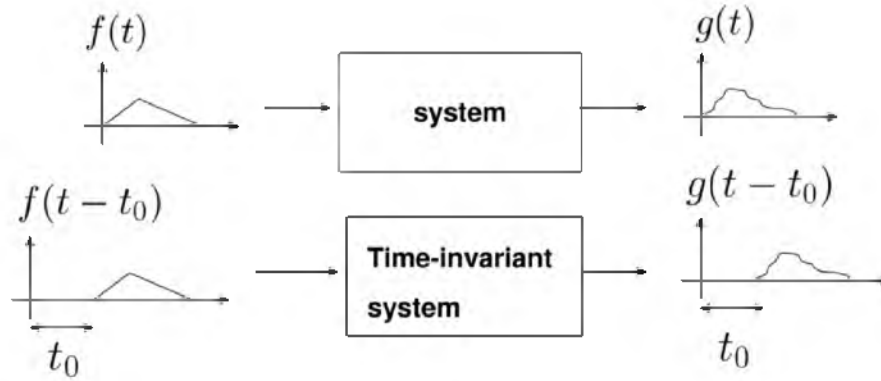


Fig. 1.9 Time-invariant system.

As it turns out, if a system is linear and time-invariant (LTI) with all initial conditions being zero, there is a definite relationship between the input and output. The relationship is given by the so-called *convolution integral*,

$$g(t) = \int_{-\infty}^{\infty} f(t')h(t-t')dt' = f(t)*h(t), \quad (1.2-1)$$

where  $h(t)$  is called the *impulse response* of the LTI system, and  $*$  is a symbol denoting the convolution of  $f(t)$  and  $h(t)$ . The expression  $f*g$  reads as  $f$  convolves with  $g$ . To see why  $h(t)$  is called the impulse response, if we let the input be a delta function,  $\delta(t)$ , then the output, according to Eq. (1.2-1), is

$$g(t) = \delta(t)*h(t) = \int_{-\infty}^{\infty} \delta(t')h(t-t')dt' = h(t),$$

where we use the sampling property of the delta function to obtain the last step of the result. Once we know  $h(t)$  of the LTI system, which can be determined experimentally by simply applying an impulse to the input of the system, we can find the response to any arbitrary input, say,  $f(t)$ , to the system through the calculation of Eq. (1.2-1).

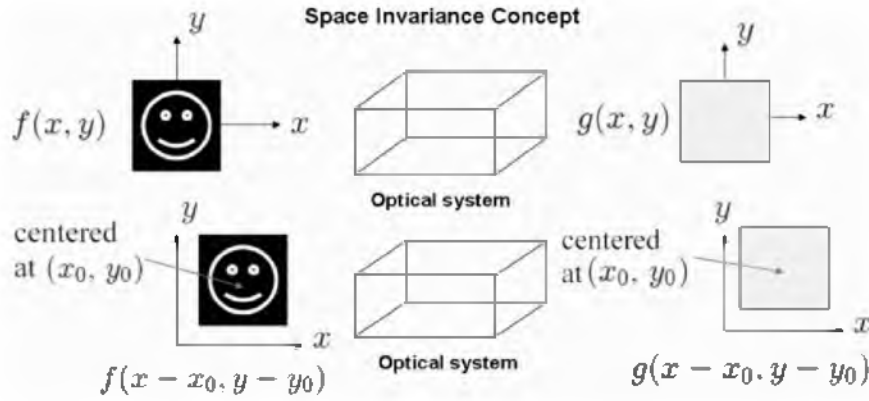


Fig. 1.10 Concept of space-invariance.

In optics, when we are dealing with signals of spatial coordinates, we can extend the concept of LTI systems to the so-called *linear space-invariant (LSI) system*. Hence we can extend the 1-D convolution integral to two dimensions as follows:

$$\begin{aligned} g(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') h(x - x', y - y') dx' dy' \\ &= f(x, y) * h(x, y), \end{aligned} \quad (1.2-2)$$

where  $f(x, y)$  is the 2-D input to the LSI system.  $h(x, y)$  and  $g(x, y)$  are the corresponding impulse response and output of the system, respectively. While the concept of time-invariance is clearly delineated by Fig. 1.9 for electrical signals, the concept of space-invariance for optical signals is not immediately clear. In Fig. 1.10, we can clarify this concept. We see that as the input image,  $f(x, y)$ , is shifted or translated to a new origin,  $(x_0, y_0)$ , its output,  $g(x, y)$ , is shifted accordingly on the  $x$ - $y$  plane. Hence, we see that the delay of an input signal in electrical systems corresponds to the translation of an output image over the output plane.

Figure 1.11 shows the block diagrams of a LSI optical system both in spatial and frequency domain. To analyze the LSI system in frequency domain, we simply take the Fourier transform of Eq. (1.2-2) to obtain

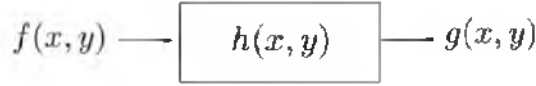
$$\mathcal{F}_{xy}\{g(x, y)\} = \mathcal{F}_{xy}\{f(x, y) * h(x, y)\}, \quad (1.2-3a)$$

which is shown to be

$$G(k_x, k_y) = F(k_x, k_y)H(k_x, k_y), \quad (1.2-3b)$$

where  $G(k_x, k_y)$  and  $H(k_x, k_y)$  are the Fourier transform of  $g(x, y)$  and  $h(x, y)$ , respectively. While  $h(x, y)$  is called the *spatial impulse response* or *point spread function (PSF)* of the LSI system, its Fourier transform,  $H(k_x, k_y)$ , is called the *spatial frequency response* or the system's *frequency transfer function*. The proof of Eq. (1.2-3b) is demonstrated in Example 1.4.

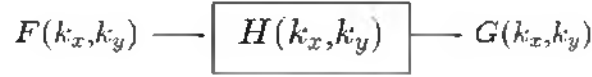
#### Block diagram of LSI optical system in spatial domain



$$g(x, y) = f(x, y) * h(x, y)$$

$h(x, y)$ : spatial impulse response or point spread function

#### Block diagram of LSI optical system in frequency domain



$$\begin{aligned} \mathcal{F}_{xy}\{g(x, y)\} &= \mathcal{F}_{xy}\{f(x, y) * h(x, y)\} \\ \text{or } G(k_x, k_y) &= F(k_x, k_y)H(k_x, k_y) \end{aligned}$$

$H(k_x, k_y)$ : spatial frequency response or frequency transfer function

Fig. 1.11 Block diagrams of LSI system.

#### Example 1.4 Fourier Transform of the Convolution of Two Functions

From Eq. (1.2-3a), we have

$$\begin{aligned} \mathcal{F}_{xy}\{g(x, y)\} &= \mathcal{F}_{xy}\{f(x, y) * h(x, y)\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f(x, y) * h(x, y)] \exp(jk_x x + jk_y y) \, dx dy \end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') h(x - x', y - y') dx' dy' \right] \\
&\quad \times \exp(jk_x x + jk_y y) dx dy,
\end{aligned}$$

where we have utilized the definition of convolution. After grouping the  $x$  and  $y$  variables together, the above equation can be written as

$$\begin{aligned}
&\mathcal{F}_{xy}\{f(x, y) * h(x, y)\} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \\
&\quad \times \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x - x', y - y') \exp(jk_x x + jk_y y) dx dy \right] dx' dy'.
\end{aligned}$$

The inner integral is the Fourier transform of  $h(x - x', y - y')$ . Using Table 1.1 (item #2), the transform is given by  $H(k_x, k_y) \exp(jk_x x' + jk_y y')$ . Hence

$$\begin{aligned}
&\mathcal{F}_{xy}\{f(x, y) * h(x, y)\} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \left[ H(k_x, k_y) \exp(jk_x x' + jk_y y') \right] dx' dy' \\
&= H(k_x, k_y) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \exp(jk_x x' + jk_y y') dx' dy' \\
&= F(k_x, k_y) H(k_x, k_y).
\end{aligned}$$

### 1.2.2 Convolution and Correlation Concept

In the last section, we have demonstrated that in the LSI system, the convolution integral is involved. In this section we will first explain the concept of convolution, and then, we will discuss another important operation called *correlation*. Finally we will make distinction between the two processes.

In Fig. 1.12, we illustrate the convolution of two images,  $f(x, y)$  and  $h(x, y)$ . According to the definition in Eq. (1.2-2), the convolution of the two images involves the calculation of the area under the product of two functions,  $f(x', y')$  and  $h(x - x', y - y')$ , for different shifts,  $(x, y)$ .

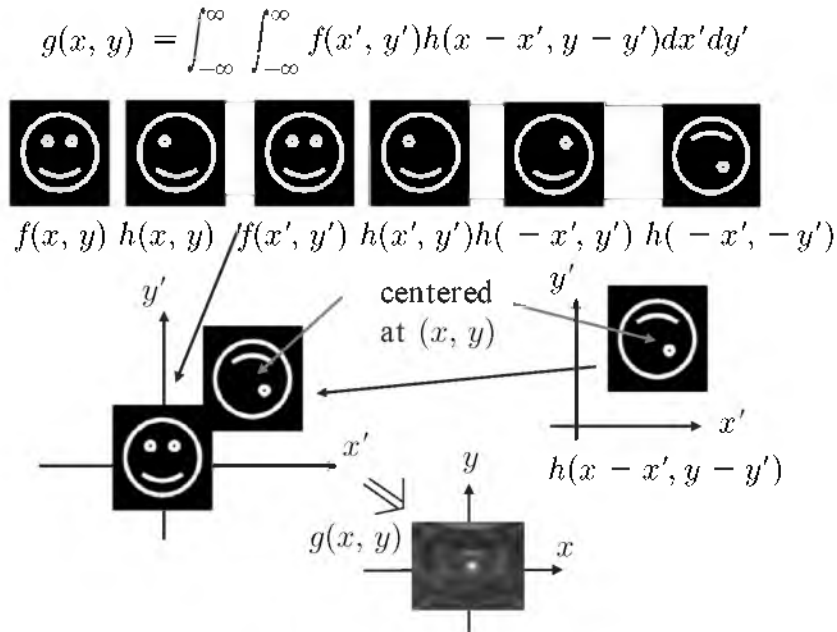
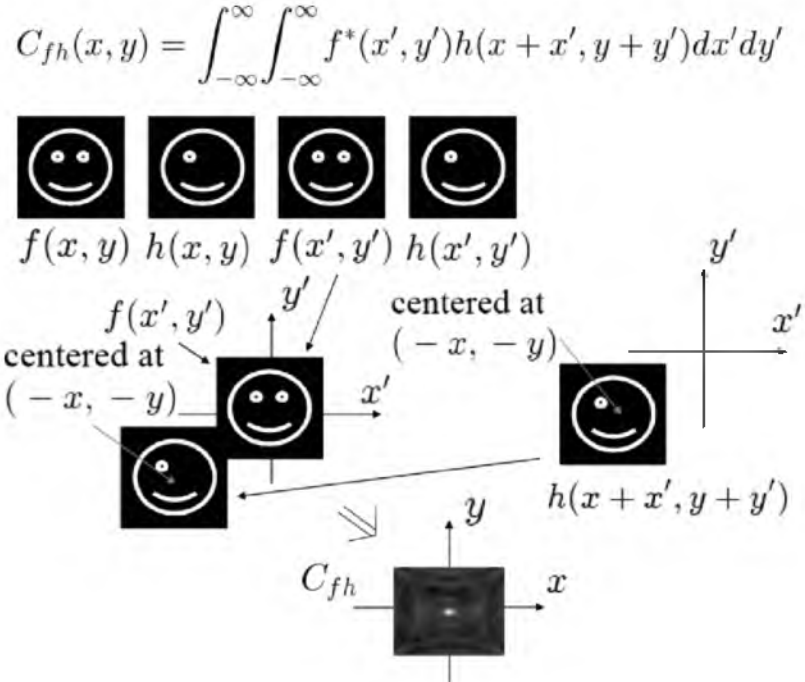


Fig. 1.12 Concept of 2-D convolution.

Fig. 1.13 Concept of 2-D correlation (assuming  $f$  is real).

The first row of figures in Fig. 1.12 shows the construction of  $f(x', y')$  and  $h(-x', -y')$  from the original images  $f(x, y)$  and  $h(x, y)$ . We then construct  $h(x - x', y - y')$  as shown in Fig. 1.12 by translating  $h(-x', -y')$  to a center at  $(x, y)$  to form  $h(x - x', y - y')$ . Once we have  $f(x', y')$  and  $h(x - x', y - y')$ , we superimpose them on the  $x'$ - $y'$  plane as illustrated in Fig. 1.12. Finally, we need to calculate the area of the product of  $f(x', y')$  and  $h(x - x', y - y')$  for different shifts  $(x, y)$  to obtain a 2-D gray-scale plot of  $g(x, y)$ .

Another important integral is called the *correlation integral*. The correlation,  $C_{fh}(x, y)$ , of two functions  $f(x, y)$  and  $h(x, y)$ , is defined as

$$\begin{aligned} C_{fh}(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x', y') h(x + x', y + y') dx' dy' \\ &= f(x, y) \otimes h(x, y). \end{aligned} \quad (1.2-4)$$

This integral is useful when comparing the similarity of two functions, and it has been knowingly used for applications in pattern recognition. For simplicity, if we assume in Fig. 1.13 that  $f(x, y)$  is real, we can illustrate the correlation of the two images,  $f(x, y)$  and  $h(x, y)$ . Similar to the convolution of the two images, the correlation involves the calculation of the area under the product of two functions,  $f(x', y')$  and  $h(x + x', y + y')$ , for different shifts,  $(x, y)$ . The first row of images in Fig. 1.13 shows the construction of  $f(x', y')$  and  $h(x', y')$  from the original images,  $f(x, y)$  and  $h(x, y)$ . Unlike convolution, to calculate the area of the product of  $f(x', y')$  and  $h(x + x', y + y')$  for different shifts,  $(x, y)$ , there is no need to flip the image,  $h(x', y')$ , upon the  $x'$ -axis and the  $y'$ -axis to obtain the 2-D plot of  $C_{fh}(x, y)$ .

### Example 1.5 Relationship between Convolution and Correlation

In this example, we will show that correlation can be expressed in terms of convolution through the following relationship:

$$f(x, y) \otimes h(x, y) = f^*(-x, -y) * h(x, y). \quad (1.2-5)$$

According to the definition of convolution [see Eq. (1.2-2)], we write

$$\begin{aligned} f^*(-x, -y) * h(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(-x', -y') h(x - x', y - y') dx' dy' \end{aligned}$$



$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x'' - x, y'' - y) h(x'', y'') (-dx'') (-dy''),$$

where we have made the substitutions  $x - x' = x''$  and  $y - y' = y''$  to obtain the last step of the equation. By re-arranging the last step and substituting the equivalents for  $x'' - x = \tilde{x}$  and  $y'' - y = \tilde{y}$ , we obtain

$$\begin{aligned} & f^*(-x, -y) * h(x, y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(\tilde{x}, \tilde{y}) h(\tilde{x} + x, \tilde{y} + y) d\tilde{x} d\tilde{y}, \\ &= f(x, y) \otimes h(x, y) \end{aligned}$$

by the definition of correlation. Therefore, we have proven Eq. (1.2-5).

With reference to Eq. (1.2-4), when  $f \neq h$ , the result is known as *cross-correlation*,  $C_{fh}$ . When  $f = h$ , the result is known as *auto-correlation*,  $C_{ff}$ , of the function  $f$ . As it turns out, we can show that

$$|C_{ff}(0, 0)| \geq |C_{ff}(x, y)|, \quad (1.2-6)$$

i.e., autocorrelation always has a central maximum. The use of this fact has been employed by *pattern recognition*. Pioneering schemes of optical pattern recognition, implementing Eq. (1.2-5), are due to Vander Lugt [1964], and Weaver and Goodman [1966]. The book, *Optical Pattern Recognition*, provides a comprehensive review of optical pattern recognition, covering theoretical aspects and details of some practical implementations [Yu and Jutamulia (1998)]. For some of the most novel approaches to optical pattern recognition, the reader is encouraged to refer to the article by Poon and Qi [2003].

### Example 1.6 MATLAB Example: Pattern Recognition

For pattern recognition applications, one implements correlation given by Eq. (1.2-4). In this example, we implement the equation in the frequency domain. To do this, we realize that

$$\mathcal{F}_{xy}\{f(x, y) \otimes h(x, y)\} = F^*(k_x, k_y) H(k_x, k_y), \quad (1.2-7)$$

which can be shown using the procedure similar to Example 1.4. For the given images  $f$  and  $h$ , we first find their corresponding 2-D Fourier

transforms, and then the correlation is evident when we take the inverse transform of Eq. (1.2-7):

$$f(x, y) \otimes h(x, y) = \mathcal{F}_{xy}^{-1} \{ F^*(k_x, k_y) H(k_x, k_y) \}. \quad (1.2-8)$$

Figure 1.14 shows the result of auto-correlation for two identical images, while Fig. 1.15 shows the cross-correlation result for two different images. These figures are generated using the m-file shown in Table 1.5. Two 256 by 256 smiley.bmp files have been used for the auto-correlation calculation. Note that in auto-correlation, shown in Fig. 1.14, a bright spot in the center of the correlation output represents the “match” of the two patterns, as suggested by Eq. (1.2-6), whereas in Fig. 1.15, there is no discernible bright spot in the center.



Fig. 1.14 Auto-correlation.



Fig. 1.15 Cross-correlation.

Table 1.5 correlation.m: m-file for performing 2-D correlation.

```
-----
%correlation.m
clear

I1=imread('smiley.bmp','bmp'); %Input image 1 (reference image)
I1=I1(:,:,1);
figure(1) %displaying input image 1
colormap(gray(255));
image(I1)
axis off

F11=fft2(I1);
max1=max(F11);
max2=max(max1);
scale=1.0/max2;
F11=F11.*scale;

I2=imread('smiley.bmp','bmp'); %Input image 2 (image to be recognized)
```

```

I2=I2(:,:,1);
figure(2) %displaying input image 2
colormap(gray(255));
image(I2)
axis off

FI2=fft2(I2);
max1=max(FI2);
max2=max(max1);
scale=1.0/max2;
FI2=FI2.*scale;

FPR=FI1.*conj(FI2);%calculating correlation
PR=ifft2(FPR);
PR=fftshift(PR);

max1=max(PR);
max2=max(max1);
scale=1.0/max2;
PR=PR.*scale;

figure(3)%display of correlation in spatial domain
colormap(gray(255));
image(abs(256*PR));
axis off
-----

```

## References

- 1.1 Banerjee, P.P. and T.-C. Poon (1991). *Principles of Applied Optics*. Irwin, Illinois.
- 1.2 Poon T.-C. and P. P. Banerjee (2001). *Contemporary Optical Image Processing with MATLAB®*. Elsevier, Oxford, UK.
- 1.3 Poon, T.-C. and Y. Qi (2003). "Novel real-time joint-transform correlation by use of acousto-optic heterodyning," *Applied Optics*, 42, 4663-4669.
- 1.4 VanderLugt, A. (1964). "Signal detection by complex spatial filter," *IEEE Trans. Inf. Theory* IT-10, 139-146.
- 1.5 Weaver, C.S. and J. W. Goodman (1969). "A technique for optical convolving two functions," *Applied Optics*, 5, 1248-1249.
- 1.6 Yu, F.T.S. and S. Jutamulia, ed. (1998). *Optical Pattern Recognition*. Cambridge University Press, Cambridge, UK.

## Chapter 2

### Wave Optics and Holography

In Chapter 1, we presented some mathematical background of Fourier optics as well as some important systems properties including linearity and space invariance. In this chapter, we present some fundamentals of wave optics by starting from Maxwell's equations and deriving the vector wave equation. We will then discuss some simple solutions of the scalar wave equation. Next, we will develop diffraction theory by using the Fresnel diffraction formula, which is uniquely derived by using Fourier transforms. In the process, we will define the spatial frequency transfer function and the spatial impulse response in Fourier optics. In the context of diffraction, we will also develop wavefront transformation by using a lens, show the Fourier transforming properties of the lens, and discuss how spatial filtering is obtained by using a standard two-lens system, leading to the distinction between coherent and incoherent image processing. In the last section of this chapter, we will discuss the basics of holography and show that a Fresnel zone plate is the hologram of a point source object, leading to the concept that the hologram of an arbitrary 3-D object can be considered as a collection of Fresnel zone plates. Finally, we will discuss electronic holography (often called digital holography in literature). This will culminate with the next chapter, which we will discuss a unique holographic recording technique called optical scanning holography.

#### 2.1 Maxwell's Equations and Homogenous Vector Wave Equation

Generally, in the study of optics, we are concerned with four vector quantities called electromagnetic (EM) fields: the electric field strength  $\mathcal{E}$  (V/m), the electric flux density  $\mathcal{D}$  (C/m<sup>2</sup>), the magnetic field strength  $\mathcal{H}$  (A/m), and the magnetic flux density  $\mathcal{B}$  (Wb/m<sup>2</sup>). The fundamental theory of electromagnetic fields is based on *Maxwell's Equations*. In differential form, these equations are expressed as

$$\nabla \cdot \mathcal{D} = \rho_v, \quad (2.1-1)$$

$$\nabla \cdot \mathcal{B} = 0, \quad (2.1-2)$$

$$\nabla \times \mathcal{E} = - \frac{\partial \mathcal{B}}{\partial t}, \quad (2.1-3)$$

$$\nabla \times \mathcal{H} = \mathcal{J} = \mathcal{J}_c + \frac{\partial \mathcal{D}}{\partial t}, \quad (2.1-4)$$

where  $\mathcal{J}_c$  is the current density [A/m<sup>2</sup>] and  $\rho_v$  denotes the electric charge density [C/m<sup>3</sup>].  $\mathcal{J}_c$  and  $\rho_v$  are the sources generating the electromagnetic fields. Maxwell's equations express the physical laws governing the *electric fields*  $\mathcal{E}$  and  $\mathcal{D}$ , *magnetic fields*  $\mathcal{H}$  and  $\mathcal{B}$ , and the *sources*  $\mathcal{J}_c$  and  $\rho_v$ . From Eqs. (2.1-3) and (2.1-4), we see that a time-varying magnetic field produces a time-varying electric field. Conversely, a time-varying electric field produces a time-varying magnetic field. It is precisely this coupling between the electric and magnetic fields that generate electromagnetic waves capable of propagating through a medium or even in free space.

For any given current and charge density distribution, we can solve Maxwell's equations. However, we need to note that Eq. (2.1-1) is not independent of Eq. (2.1-4). Similarly, Eq. (2.1-2) is a consequence of Eq. (2.1-3). By taking the divergence on both sides of Eqs. (2.1-3) and (2.1-4) and using the *continuity equation*:

$$\nabla \cdot \mathcal{J}_c + \frac{\partial \rho_v}{\partial t} = 0, \quad (2.1-5)$$

which is the *principle of conservation of charge*, we can show that  $\nabla \cdot \mathcal{D} = \rho_v$ . Similarly, Eq. (2.1-2) is a consequence of Eq. (2.1-3). Hence, from Eqs. (2.1-1) to (2.1-4), we really have six independent scalar equations (three scalar equations for each curl equation) and twelve unknowns. The unknowns are the  $x$ ,  $y$ , and  $z$  components of  $\mathcal{E}$ ,  $\mathcal{D}$ ,  $\mathcal{H}$ , and  $\mathcal{B}$ . The six more scalar equations required are provided by the *constitutive relations*:

$$\mathcal{D} = \epsilon \mathcal{E}, \quad (2.1-6a)$$

and

$$\mathcal{B} = \mu \mathcal{H}, \quad (2.1-6b)$$

where  $\epsilon$  denotes the permittivity [F/m] and  $\mu$  denotes the permeability [H/m] of the medium. In this book, we take  $\epsilon$  and  $\mu$  to be scalar constants. Indeed, this is true for a *linear, homogeneous, and isotropic* medium. A medium is *linear* if its properties do not depend on the amplitude of the fields in the medium. It is *homogeneous* if its properties are not functions of space. And the medium is *isotropic* if its properties are the same in all direction from any given point.

Returning our focus to linear, homogeneous, and isotropic media, constants worth remembering are the values of  $\epsilon$  and  $\mu$  for free space (or vacuum):  $\epsilon_0 = (1/36\pi) \times 10^{-9}$  F/m and  $\mu_0 = 4\pi \times 10^{-7}$  H/m.

Using Maxwell's equations and the constitutive relations, we can derive the wave equation, which describes the propagation of the electric and magnetic fields. Example 2.1 shows the derivation of the wave equation for  $\mathcal{E}$ .

### Example 2.1 Derivation of Vector Wave Equation in a Linear, Homogenous, and Isotropic Medium

By taking the curl of both sides of Eq. (2.1-3), we have

$$\begin{aligned}\nabla \times \nabla \times \mathcal{E} &= -\nabla \times \frac{\partial \mathcal{B}}{\partial t} \\ &= -\frac{\partial}{\partial t}(\nabla \times \mathcal{B}) = -\mu \frac{\partial}{\partial t}(\nabla \times \mathcal{H}),\end{aligned}\quad (2.1-7)$$

where we have used the second of the constitutive relations [Eq. (2.1-6b)] and assumed  $\mu$  to be space- and time-independent. Now, by employing Eq. (2.1-4), Eq. (2.1-7) becomes

$$\nabla \times \nabla \times \mathcal{E} = -\mu\epsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} - \mu \frac{\partial \mathcal{J}_c}{\partial t}, \quad (2.1-8)$$

where we have used the first of the constitutive relations [Eq. (2.1-6a)] and assumed  $\epsilon$  to be time-independent. Then, by using the following vector identity ( $\mathcal{A}$  is some arbitrary vector)

$$\nabla \times \nabla \times \mathcal{A} = \nabla(\nabla \cdot \mathcal{A}) - \nabla^2 \mathcal{A}, \quad \nabla^2 = \nabla \cdot \nabla, \quad (2.1-9)$$

in Eq. (2.1-8), we get

$$\nabla^2 \mathcal{E} - \mu\epsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} = \mu \frac{\partial \mathcal{J}_c}{\partial t} + \nabla(\nabla \cdot \mathcal{E}). \quad (2.1-10)$$

If we also assume the permittivity,  $\epsilon$ , to be space-independent, then we can now recast the first of Maxwell's equations [Eq. (2.1-1)] in the form of

$$\nabla \cdot \mathcal{E} = \frac{\rho_v}{\epsilon}, \quad (2.1-11)$$

by using the first of the constitutive relations [Eq. (2.1-6a)]. Incorporating Eq. (2.1-11) into Eq. (2.1-10), we can finally obtain

$$\nabla^2 \mathcal{E} - \mu\epsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} = \mu \frac{\partial \mathcal{J}_c}{\partial t} + \frac{1}{\epsilon} \nabla \rho_v, \quad (2.1-12)$$

which is a *vector wave equation* having source terms on the right-hand side. This is the wave equation for  $\mathcal{E}$  in a linear, homogeneous, and isotropic medium characterized by  $\mu$  and  $\epsilon$ .

For the given  $\mathcal{J}_c$  and  $\rho_v$  in a localized region characterized by  $\mu$  and  $\epsilon$ , say,  $V'$ , we can solve for the electric field,  $\mathcal{E}$ , in the region according to Eq. (2.1-12). Once the generated field reaches the source-free region  $V$  ( $\mathcal{J}_c = 0$ ,  $\rho_v = 0$ ), the field must then satisfy the *homogenous vector wave equation*,

$$\nabla^2 \mathcal{E} - \mu\epsilon \frac{\partial^2 \mathcal{E}}{\partial t^2} = 0. \quad (2.1-13)$$

The situation is delineated in Fig. 2.1. Note that the quantity,  $\mu\epsilon$ , has the unit value of  $(1/\text{velocity})^2$ . We call this velocity  $v$  and define it as

$$v^2 = \frac{1}{\mu\epsilon}. \quad (2.1-14)$$

For free space,  $\mu = \mu_0$ ,  $\epsilon = \epsilon_0$ , and  $v = c$ . We can calculate the value of  $c$  from the values of  $\epsilon_0$  and  $\mu_0$ . This works out to be about  $3 \times 10^8$  m/s. This theoretical value, which was first calculated by Maxwell, was in remarkable agreement with Fizeau's previously measured speed of light (315,300 km/s). This led Maxwell to the conclusion that light is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field based on electromagnetic laws.

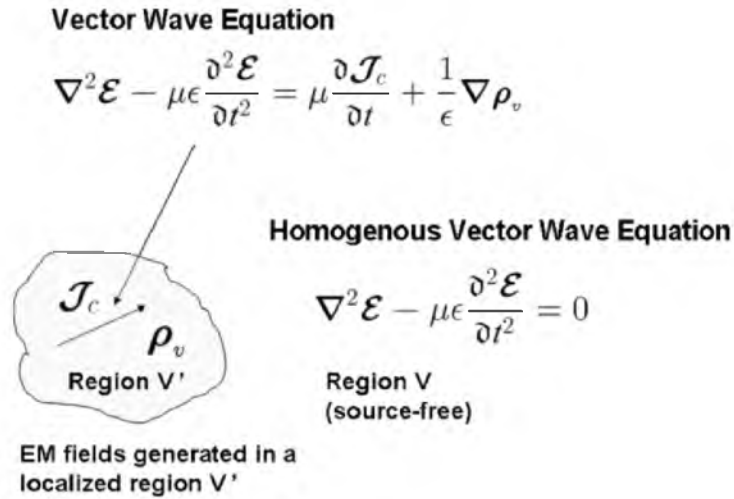


Fig. 2.1 Vector wave equations in a linear, homogeneous, and isotropic medium.

## 2.2 Three-Dimensional Scalar Wave Equation

Equation (2.1-13) is equivalent to three scalar equations - one for every component of  $\mathcal{E}$ . We shall let the field  $\mathcal{E}$  to be of the form

$$\mathcal{E} = \mathcal{E}_x \mathbf{a}_x + \mathcal{E}_y \mathbf{a}_y + \mathcal{E}_z \mathbf{a}_z, \quad (2.2-1)$$

where  $\mathbf{a}_x$ ,  $\mathbf{a}_y$ , and  $\mathbf{a}_z$  denote the unit vectors in the  $x$ ,  $y$ , and  $z$  directions, respectively. Now, the expression for the Laplacian ( $\nabla^2$ ) operator in Cartesian ( $x, y, z$ ) coordinates is given by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (2.2-2)$$

Using the above equation, Equation (2.1-13) becomes

$$\begin{aligned} & \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) (\mathcal{E}_x \mathbf{a}_x + \mathcal{E}_y \mathbf{a}_y + \mathcal{E}_z \mathbf{a}_z) \\ &= \mu \epsilon \frac{\partial^2}{\partial t^2} (\mathcal{E}_x \mathbf{a}_x + \mathcal{E}_y \mathbf{a}_y + \mathcal{E}_z \mathbf{a}_z). \end{aligned} \quad (2.2-3)$$

Comparing the  $\mathbf{a}_x$ -component on both sides of the equation, we have

$$\frac{\partial^2 \mathcal{E}_x}{\partial x^2} + \frac{\partial^2 \mathcal{E}_x}{\partial y^2} + \frac{\partial^2 \mathcal{E}_x}{\partial z^2} = \mu \epsilon \frac{\partial^2 \mathcal{E}_x}{\partial t^2}.$$

Similarly, we end up with the same type of equation shown above for the  $\mathcal{E}_y$  and  $\mathcal{E}_z$  component by comparing the other components in Eq. (2.2-3). Therefore, we can write

$$\nabla^2 \psi = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}, \quad (2.2-4)$$

where  $\psi$  may represent a component,  $\mathcal{E}_x$ ,  $\mathcal{E}_y$  or  $\mathcal{E}_z$ , of the electric field  $\mathcal{E}$ , and where  $v$  is the velocity of the wave in the medium by using Eq. (2.1-14). Equation (2.2-4) is called the *3-D scalar wave equation*. We shall look at some of its simplest solutions in the next section.

### 2.2.1 Plane Wave Solution

For waves oscillating at the *angular frequency*,  $\omega_0$  (rad/s), one of the simplest solutions to Eq. (2.2-4) is

$$\psi(x, y, z, t) = \exp[j(\omega_0 t - \mathbf{k}_0 \cdot \mathbf{R})]$$



$$= \exp[j(\omega_0 t - k_{0x}x - k_{0y}y - k_{0z}z)], \quad (2.2-5)$$

where  $\mathbf{R} = x\mathbf{a}_x + y\mathbf{a}_y + z\mathbf{a}_z$  is the position vector,  $\mathbf{k}_0 = k_{0x}\mathbf{a}_x + k_{0y}\mathbf{a}_y + k_{0z}\mathbf{a}_z$  is the *propagation vector*, and  $|\mathbf{k}_0| = k_0$  is called the *propagation constant* [rad/m]. With the condition that

$$\frac{\omega_0^2}{k_{0x}^2 + k_{0y}^2 + k_{0z}^2} = \frac{\omega_0^2}{k_0^2} = v^2, \quad (2.2-6)$$

Eq. (2.2-5) is called a plane-wave solution and the wave is called a *plane wave* of unit amplitude. Figure 2.2 shows the direction of propagation of the plane wave, which is determined from the three components  $k_{0x}$ ,  $k_{0y}$ , and  $k_{0z}$ .

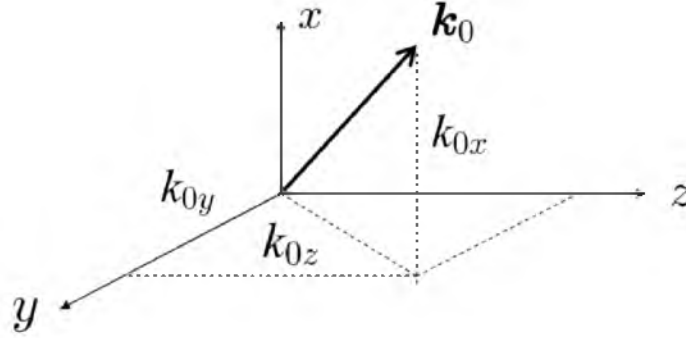


Fig. 2.2 Plane wave propagating along the  $\mathbf{k}_0$  direction.

Since the electromagnetic fields are real functions of space and time, we can define, for example, the electric field by taking the real part of  $\psi$  to obtain a real quantity,

$$\text{Re}[\psi(x, y, z, t)] = \cos(\omega_0 t - k_{0x}x - k_{0y}y - k_{0z}z). \quad (2.2-7)$$

Let us now consider a plane wave propagating along the  $z$ -direction. In one spatial dimension, i.e.,  $\psi(z, t)$ , the wave equation [Eq. (2.2-4)] reads

$$\frac{\partial^2 \psi}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} \quad (2.2-8)$$

and its plane wave solution then becomes

$$\psi(z, t) = \exp[j(\omega_0 t - k_0 z)] = \exp[j(\omega_0 t)] \exp[-j\theta(z)], \quad (2.2-9)$$

where  $\theta(z) = k_0 z = \frac{2\pi}{\lambda_0} z$  is called the *phase* of the wave with  $\lambda_0$  indicating the wavelength of the wave. Let us take the origin of the coordinates as a

zero-phase position, i.e.,  $\theta(z = 0) = 0$ . In fact, over the whole plane  $z = 0$ , the phase is zero. At  $z = \lambda_0$ , we have  $\theta(z = \lambda_0) = \frac{2\pi}{\lambda_0} \lambda_0 = 2\pi$ . So for every distance of propagation of a wavelength, the phase of the wave gains  $2\pi$ . Therefore, we have what is known as the *planar wavefronts* along the  $z$ -direction. The situation is demonstrated in Fig. 2.3.

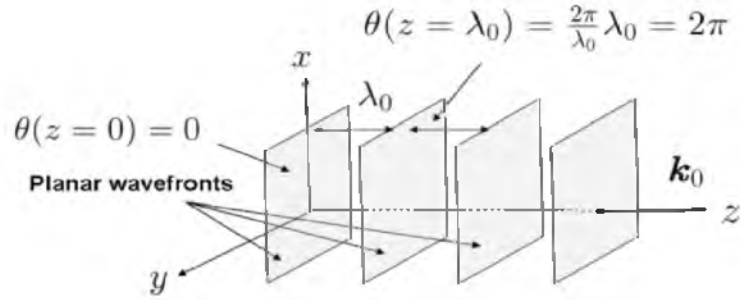


Fig. 2.3 Plane wave propagating along the  $z$ -direction exhibiting planar wavefronts.

### 2.2.2 Spherical Wave Solution

Consider now the spherical coordinates shown in Fig. 2.4.

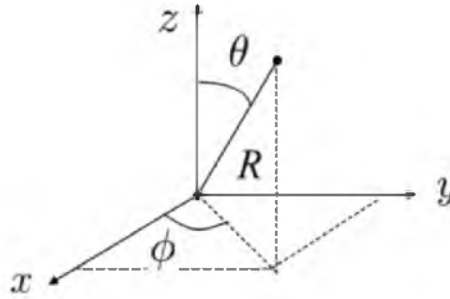


Fig. 2.4 Spherical coordinate system.

The expression for the Laplacian ( $\nabla^2$ ) operator is

$$\nabla^2 = \frac{\partial^2}{\partial R^2} + \frac{2}{R} \frac{\partial}{\partial R} + \frac{1}{R^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{R^2} \frac{\partial^2}{\partial \theta^2} + \frac{\cot \theta}{R^2} \frac{\partial}{\partial \theta}. \quad (2.2-10)$$

One of the simplest cases is called spherical symmetry, which requires that  $\psi(R, \theta, \phi, t) = \psi(R, t)$ . Therefore, for spherical symmetry ( $\partial/\partial \phi = 0 = \partial/\partial \theta$ ), the wave equation, Eq. (2.2-4), combined with Eq. (2.2-10) assumes the form

$$\left( \frac{\partial^2 \psi}{\partial R^2} + \frac{2}{R} \frac{\partial \psi}{\partial R} \right) = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}. \quad (2.2-11)$$

Since

$$R \left( \frac{\partial^2 \psi}{\partial R^2} + \frac{2}{R} \frac{\partial \psi}{\partial R} \right) = \frac{\partial^2 (R\psi)}{\partial R^2},$$

we can re-write Eq. (2.2-11) to become

$$\frac{\partial^2 (R\psi)}{\partial R^2} = \frac{1}{v^2} \frac{\partial^2 (R\psi)}{\partial t^2}. \quad (2.2-12)$$

Now, the above equation is of the same form as that of Eq. (2.2-8). Since Eq. (2.2-9) is the solution to Eq. (2.2-8), we can therefore construct a simple solution to Eq. (2.2-12) as

$$\psi(R, t) = \frac{1}{R} \exp[j(\omega_0 t - k_0 R)], \quad (2.2-13)$$

which is called a *spherical wave*. Again, we can write

$$\psi(R, t) = \frac{1}{R} \exp[j(\omega_0 t - k_0 R)] = \frac{1}{R} \exp(j\omega_0 t) \exp[-j\theta(R)],$$

where  $\theta(R) = k_0 R = \frac{2\pi}{\lambda_0} R$ . We then take the origin of the coordinates as a zero-phase position, i.e.,  $\theta(R = 0) = 0$  and  $\theta(R = \lambda_0) = \frac{2\pi}{\lambda_0} \lambda_0 = 2\pi$ . So, for every distance of propagation of a wavelength, the phase of the wave gains  $2\pi$ . We, therefore, have the so-called *spherical wavefronts* moving along the  $R$ -direction. The situation is shown in Fig. 2.5.

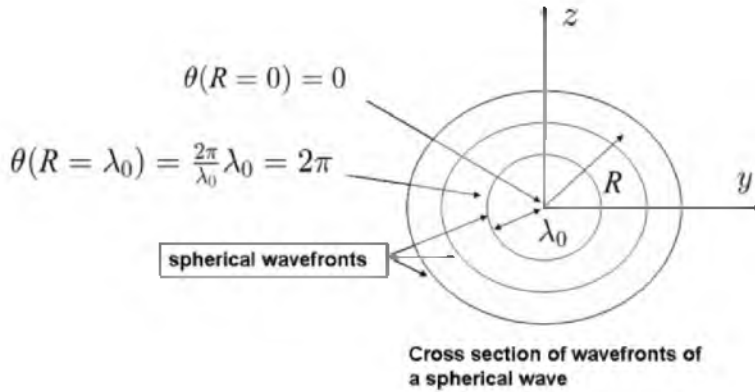


Fig. 2.5 Spherical wavefronts.

While we have previously seen that plane waves and spherical waves are some of the simplest solutions of the 3-D scalar wave equation, we can effectively generate these useful waves in the laboratory. The situation is

shown in Fig. 2.6, where the distance between the two lenses are separated by the sum of their focal lengths,  $f_1 + f_2$ , and we have assumed that the rays emitting from the laser are parallel, i.e., the wave fronts are planar. Note that the parallel rays emerging from the lens of focal length  $f_2$  have a separation of an expansion factor,  $M = f_2/f_1$ , larger than the separation of the rays originally emerging from the laser.

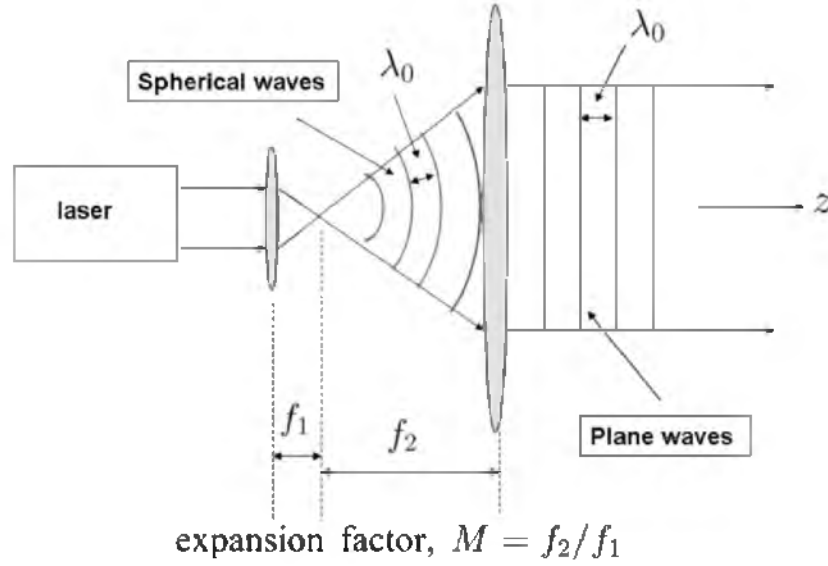


Fig. 2.6 Practical implementation of spherical waves and plane waves.

### 2.3 Scalar Diffraction Theory

Figure 2.7 shows a simple example of diffraction geometry where a plane wave oscillating at  $\omega_0$  is incident on an *aperture* or a *diffracting screen*, located on the plane  $z = 0$ . The problem is to determine the diffracted field distribution after the aperture. To tackle the problem, we will need to solve the 3-D scalar wave equation, which is subject to an initial condition. Let us now formulate the problem mathematically.

Since a plane wave of amplitude  $A$  propagating along the  $z$ -direction is given by  $\psi(z, t) = A \exp[j(\omega_0 t - k_0 z)]$  with the wave's zero-phase position defined at  $z = 0$ , we can then model the field immediately in front of the aperture as  $\psi(z = 0, t) = A \exp(j\omega_0 t)$ . The field immediately after the aperture is then given by  $\psi(x, y, z = 0, t) = \psi_p(x, y; z = 0) \exp(j\omega_0 t)$ .  $\psi_p(x, y; z = 0)$  is called the *initial condition* under consideration. For example, if the aperture has a rectangular opening of width  $x_0$  by  $y_0$ , we can then write  $\psi_p(x, y; z = 0) = A \text{rect}(x/x_0, y/y_0)$ .

It is necessary to find the field distribution at  $z$ , and to do so, we can model the solution as

$$\psi(x, y, z, t) = \psi_p(x, y, z) \exp(j\omega_0 t), \quad (2.3-1)$$

where  $\psi_p(x, y, z)$  is the unknown to be found. In optics,  $\psi_p(x, y, z)$  is called a *complex amplitude*  $\psi_p(x, y, z)$  and we see that it is riding on a *carrier* of frequency  $\omega_0$  ( $\psi_p$  is known as a *phasor* in electrical engineering).

Since the light field must satisfy the wave equation, we therefore substitute this into the 3-D scalar wave equation [Eq. (2.2-4)] to find  $\psi_p(x, y, z)$  under the given initial condition,  $\psi_p(x, y, z = 0)$ .

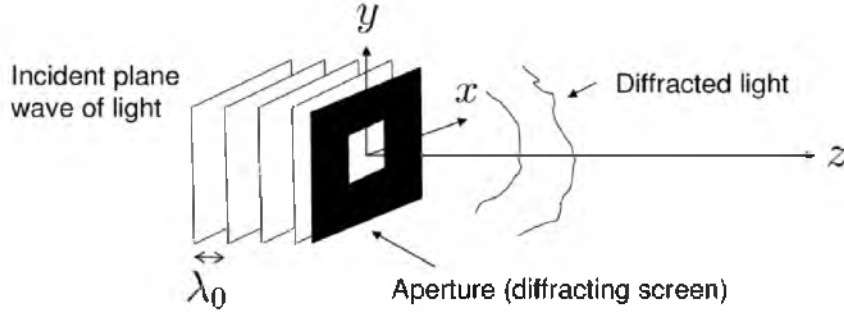


Fig. 2.7 Diffraction geometry.

After substituting Eq. (2.3-1) into Eq. (2.2-4), we get the *Helmholtz equation* for  $\psi_p$ ,

$$\frac{\partial^2 \psi_p}{\partial x^2} + \frac{\partial^2 \psi_p}{\partial y^2} + \frac{\partial^2 \psi_p}{\partial z^2} + k_0^2 \psi_p = 0, \quad k_0 = \frac{\omega_0}{v}. \quad (2.3-2)$$

By taking the 2-D Fourier transform, i.e.,  $\mathcal{F}_{xy}$ , of Eq. (2.3-2) and after further manipulations, we obtain

$$\frac{d^2 \Psi_p}{dz^2} + k_0^2 \left(1 - \frac{k_x^2}{k_0^2} - \frac{k_y^2}{k_0^2}\right) \Psi_p = 0, \quad (2.3-3)$$

where  $\Psi_p(k_x, k_y, z)$  is the Fourier transform of  $\psi_p(x, y, z)$ . We can now readily solve the above equation to get

$$\Psi_p(k_x, k_y, z) = \Psi_{p0}(k_x, k_y) \exp\left[-jk_0 \sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z\right], \quad (2.3-4)$$

where  $\Psi_{p0}(k_x, k_y) = \Psi_p(k_x, k_y, z = 0)$

$$= \mathcal{F}_{xy}\{\psi_p(x, y, z = 0)\} = \mathcal{F}_{xy}\{\psi_{p0}(x, y)\}.$$

We can interpret Eq. (2.3-4) by considering a linear system with  $\Psi_{p0}(k_x, k_y)$  as its input spectrum (i.e., at  $z = 0$ ) and where the output spectrum is  $\Psi_p(k_x, k_y; z)$ . Conclusively, the spatial frequency response of the system is given by

$$\begin{aligned} \frac{\Psi_p(k_x, k_y; z)}{\Psi_{p0}(k_x, k_y)} &= \mathcal{H}(k_x, k_y; z) \\ &= \exp\left[-jk_0\sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z\right]. \end{aligned} \quad (2.3-5)$$

We call  $\mathcal{H}(k_x, k_y; z)$  the *spatial frequency transfer function of propagation* of light through a distance  $z$  in the medium. Figure 2.8 shows the relationship between the input spectrum and the output spectrum.

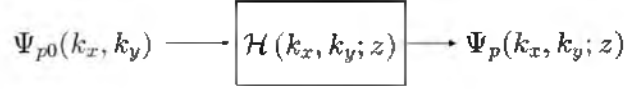


Fig. 2.8 Spatial frequency transfer function of propagation relating input spectrum to output spectrum.

### Example 2.2 Derivation of the Helmholtz Equation

When we substitute  $\psi(x, y, z, t) = \psi_p(x, y; z)\exp(j\omega_0 t)$  into the 3-D scalar wave equation given by Eq. (2.2-4), we have

$$\left[\frac{\partial^2 \psi_p}{\partial x^2} + \frac{\partial^2 \psi_p}{\partial y^2} + \frac{\partial^2 \psi_p}{\partial z^2}\right] \exp(j\omega_0 t) = \frac{(j\omega_0)^2}{v^2} \psi_p(x, y; z) \exp(j\omega_0 t)$$

or

$$\left[\frac{\partial^2 \psi_p}{\partial x^2} + \frac{\partial^2 \psi_p}{\partial y^2} + \frac{\partial^2 \psi_p}{\partial z^2}\right] = \frac{-\omega_0^2}{v^2} \psi_p(x, y; z) = -k_0^2 \psi_p$$

which is the Helmholtz equation [Eq. (2.3-2)], where we have incorporated the fact that  $k_0 = \omega_0/v$ . Note that the Helmholtz equation contains no time variable.

### Example 2.3 Derivation of Eq. (2.3-3) and its Solution

By taking the 2-D Fourier transform, i.e.,  $\mathcal{F}_{xy}$ , of Eq. (2.3-2) and by using item #5 of Table 1.1, we can obtain

$$\mathcal{F}_{xy} \left\{ \frac{\partial^2 \psi_p}{\partial x^2} + \frac{\partial^2 \psi_p}{\partial y^2} + \frac{\partial^2 \psi_p}{\partial z^2} + k_0^2 \psi_p \right\} = 0$$

or

$$- (k_x^2 + k_y^2) \Psi_p(k_x, k_y; z) + \frac{d^2 \Psi_p(k_x, k_y; z)}{dz^2} + k_0^2 \Psi_p(k_x, k_y; z) = 0,$$

which can then be re-arranged to become

$$\frac{d^2 \Psi_p}{dz^2} + k_0^2 \left( 1 - \frac{k_x^2}{k_0^2} - \frac{k_y^2}{k_0^2} \right) \Psi_p = 0. \quad (2.3-6)$$

This equation is of the form

$$\frac{d^2 y}{dz^2} + \alpha^2 y = 0,$$

which has the solution  $y(z) = y_0 \exp(-j\alpha z)$  where  $y_0 = y(z=0)$  is given as the initial condition. Using this result, the solution to Eq. (2.3-6) becomes

$$\begin{aligned} \Psi_p(k_x, k_y; z) &= \Psi(k_x, k_y; z=0) \exp \left[ -jk_0 \sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z \right] \\ &= \Psi_{p0}(k_x, k_y) \exp \left[ -jk_0 \sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z \right], \end{aligned} \quad (2.3-7)$$

which is Eq. (2.3-4).

To find the field distribution at  $z$  in the spatial domain, we take the inverse Fourier transform of Eq. (2.3-7):

$$\begin{aligned} \psi_p(x, y; z) &= \mathcal{F}_{xy}^{-1} \left\{ \Psi_p(k_x, k_y; z) \right\} \\ &= \frac{1}{4\pi^2} \iint \Psi_{p0}(k_x, k_y) \exp \left[ -jk_0 \sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z \right] \\ &\quad \times \exp(-jk_x x - jk_y y) dk_x dk_y. \end{aligned} \quad (2.3-8)$$

Now, by substituting  $\Psi_{p0}(k_x, k_y) = \mathcal{F}_{xy} \left\{ \psi_{p0}(x, y) \right\}$  into Eq. (2.3-8), we can

express  $\psi_p(x, y; z)$  as

$$\begin{aligned}\psi_p(x, y; z) &= \iint \psi_{p0}(x', y') G(x - x', y - y'; z) dx' dy' \\ &= \psi_{p0}(x, y) * G(x, y; z),\end{aligned}\quad (2.3-9)$$

where

$$\begin{aligned}G(x, y; z) &= \frac{1}{4\pi^2} \iint \exp[-jk_0 \sqrt{1 - k_x^2/k_0^2 - k_y^2/k_0^2} z] \\ &\quad \times \exp(-jk_x x - jk_y y) dk_x dk_y.\end{aligned}$$

The result of Eq. (2.3-9) indicates that  $G(x, y; z)$  can be considered as the *spatial impulse response of propagation* of light, which can be evaluated to become [Poon and Banerjee (2001)]

$$\begin{aligned}G(x, y; z) &= \frac{jk_0 \exp(-jk_0 \sqrt{x^2 + y^2 + z^2})}{2\pi \sqrt{x^2 + y^2 + z^2}} \\ &\quad \times \frac{z}{\sqrt{x^2 + y^2 + z^2}} \left(1 + \frac{1}{jk_0 \sqrt{x^2 + y^2 + z^2}}\right).\end{aligned}\quad (2.3-10)$$

### 2.3.1 Fresnel Diffraction

Equation (2.3-10) is complicated to use as is, and we shall need to make the following approximations to obtain the well-known Fresnel diffraction formula commonly used in Fourier optics:

(1) For  $z \gg \lambda_0 = 2\pi/k_0$ , i.e., we observe the field distribution many wavelengths away from the diffracting aperture, and we have

$$\left(1 + \frac{1}{jk_0 \sqrt{x^2 + y^2 + z^2}}\right) \approx 1.$$

(2) Using the binomial expansion, the factor

$$\sqrt{x^2 + y^2 + z^2} \approx z + \frac{x^2 + y^2}{2z},$$

provided that  $x^2 + y^2 \ll z^2$ . This condition is called the *paraxial approximation*. If this approximation is used in the more sensitive phase term and only used the first expansion term in the less sensitive denominators of



the first and second terms of Eq. (2.3-10), then  $G(x, y; z)$  becomes the so-called *spatial impulse response in Fourier optics*,  $h(x, y; z)$  [Poon and Banerjee (2001), Goodman (2005)]:

$$h(x, y; z) = \exp(-jk_0 z) \frac{jk_0}{2\pi z} \exp\left[\frac{-jk_0(x^2 + y^2)}{2z}\right]. \quad (2.3-11)$$

If Eq. (2.3-11) is now used in Eq. (2.3-9), we obtain

$$\begin{aligned} \psi_p(x, y; z) &= \psi_{p0}(x, y) * h(x, y; z) \\ &= \exp(-jk_0 z) \frac{jk_0}{2\pi z} \iint \psi_{p0}(x', y') \\ &\quad \times \exp\left\{\frac{-jk_0}{2z}[(x - x')^2 + (y - y')^2]\right\} dx' dy'. \end{aligned} \quad (2.3-12)$$

Equation (2.3-12) is called the *Fresnel diffraction formula* and describes the *Fresnel diffraction* of a beam during propagation and having an arbitrary initial complex profile,  $\psi_{p0}(x, y)$ . To obtain the output field distribution  $\psi_p(x, y; z)$  at a distance  $z$  away from the input (the location of the diffracting screen), we would simply convolve the input field distribution,  $\psi_{p0}(x, y)$ , with the spatial impulse response,  $h(x, y; z)$ .

By taking the 2-D Fourier transform of  $h(x, y; z)$ , we obtain

$$\begin{aligned} H(k_x, k_y; z) &= \mathcal{F}_{xy}\{h(x, y; z)\} \\ &= \exp(-jk_0 z) \exp\left[\frac{j(k_x^2 + k_y^2)z}{2k_0}\right]. \end{aligned} \quad (2.3-13)$$

$H(k_x, k_y; z)$  is known as the *spatial frequency transfer function in Fourier optics*. Indeed, we can derive Eq. (2.3-13) directly if we assume that  $k_x^2 + k_y^2 \ll k_0^2$ , meaning that the  $x$  and  $y$  components of the propagation vector of a wave are relatively small, from Eq. (2.3-5), we have

$$\begin{aligned} \frac{\Psi_p(k_x, k_y; z)}{\Psi_{p0}(k_x, k_y)} &= \mathcal{H}(k_x, k_y; z) \\ &= \exp\left[-jk_0 \sqrt{1 - (k_x^2 + k_y^2)/k_0^2} z\right] \\ &\simeq \exp(-jk_0 z) \exp\left[\frac{j(k_x^2 + k_y^2)z}{2k_0}\right] \\ &= H(k_x, k_y; z) \end{aligned}$$

or

$$\Psi_p(k_x, k_y; z) = \Psi_{p0}(k_x, k_y) H(k_x, k_y; z). \quad (2.3-14)$$

Figure 2.9 summarizes the results of Fresnel diffraction in terms of block diagrams in the spatial domain as well as in the spatial frequency domain.

#### Block diagram representation in spatial domain

$$\psi_{p0}(x, y) \longrightarrow \boxed{h(x, y; z)} \longrightarrow \psi_p(x, y; z)$$

$$\psi_p(x, y; z) = \psi_{p0}(x, y) * h(x, y; z)$$

$$h(x, y; z) = \exp(-jk_0 z) \frac{jk_0}{2\pi z} \exp\left[\frac{-jk_0(x^2 + y^2)}{2z}\right]$$

#### Block diagram representation in spatial frequency domain

$$\Psi_{p0}(k_x, k_y) \longrightarrow \boxed{H(k_x, k_y; z)} \longrightarrow \Psi_p(k_x, k_y; z)$$

$$\Psi_p(k_x, k_y; z) = \Psi_{p0}(k_x, k_y) H(k_x, k_y; z)$$

$$H(k_x, k_y; z) = \mathcal{F}_{xy}\{h(x, y; z)\} = \exp(-jk_0 z) \exp\left[\frac{j(k_x^2 + k_y^2)z}{2k_0}\right]$$

Fig. 2.9 Block diagrams to summarize Fresnel diffraction.

#### Example 2.4 Diffraction of a Point Source

A point source is represented by  $\psi_{p0}(x, y) = \delta(x, y)$ . From Eq. (2.3-12), the complex field at a distance  $z$  away is given by

$$\begin{aligned} \psi_p(x, y, z) &= \delta(x, y) * h(x, y; z) \\ &= \exp(-jk_0 z) \frac{jk_0}{2\pi z} \exp\left[-\frac{jk_0(x^2 + y^2)}{2z}\right]. \end{aligned} \quad (2.3-15)$$

This expression is the paraxial approximation to a *diverging spherical wave*. The variable  $z$  in the argument of the exponential function is called the *radius of curvature* of the spherical wave. The wavefronts are divergent

when  $z > 0$  and convergent when  $z < 0$ . We can re-write Eq. (2.3-15) as

$$\psi_p(x, y, z) = \frac{jk_0}{2\pi z} \exp\left[-jk_0\left(z + \frac{x^2 + y^2}{2z}\right)\right].$$

Now, by considering the argument of the exponential function, we see that by using the binomial expansion  $\sqrt{x^2 + y^2 + z^2} \approx z + \frac{x^2 + y^2}{2z}$ , we can write

$$\begin{aligned} \psi_p(x, y, z) &\simeq \frac{jk_0}{2\pi z} \exp\left[-jk_0(x^2 + y^2 + z^2)^{\frac{1}{2}}\right] \\ &\simeq \frac{jk_0}{2\pi R} \exp(-jk_0 R), \end{aligned} \quad (2.3-16)$$

where we have used  $z \simeq R$  in the less sensitive denominator. Eq. (2.3-16) corresponds to Eq. (2.2-13) for a diverging spherical wave.

### Example 2.5 Diffraction of a Plane Wave

For a plane wave, we can write  $\psi_{p0}(x, y) = 1$ . Then  $\Psi_{p0}(k_x, k_y) = 4\pi^2 \delta(k_x) \delta(k_y)$ . Using Eq. (2.3-14), we have

$$\begin{aligned} \Psi_p(k_x, k_y; z) &= 4\pi^2 \delta(k_x) \delta(k_y) \exp(-jk_0 z) \exp\left[\frac{j(k_x^2 + k_y^2)z}{2k_0}\right] \\ &= 4\pi^2 \delta(k_x) \delta(k_y) \exp(-jk_0 z). \end{aligned}$$

Its inverse transform gives the expression of a plane wave [see Eq. (2.2-9)],

$$\psi_p(x, y, z) = \exp(-jk_0 z).$$

As the plane wave travels, it only acquires phase shift and, as expected, is undiffracted.

### 2.3.2 Diffraction of a Square Aperture

In general, when a light field illuminates a transparency of transmission function given by  $t(x, y)$ , and if the complex amplitude of the light just in front of the transparency is  $\psi_{i,p}(x, y)$ , then the complex field immediately after the transparency is given by  $\psi_{i,p}(x, y)t(x, y)$ . In writing this product result, we assume that the transparency is infinitely thin.

Now, let us consider a simple situation where a plane wave of unit amplitude is incident normally on the transparency  $t(x, y)$ , and the field emerging from the transparency is then  $1 \times t(x, y)$  as  $\psi_{i,p}(x, y) = 1$  in the present case. We want to find the field distribution, which is a distance  $z$  away from the transparency. This corresponds to the Fresnel diffraction of an arbitrary beam profile as the transparency modifies the incident plane wave. The situation is demonstrated in Fig. 2.10.

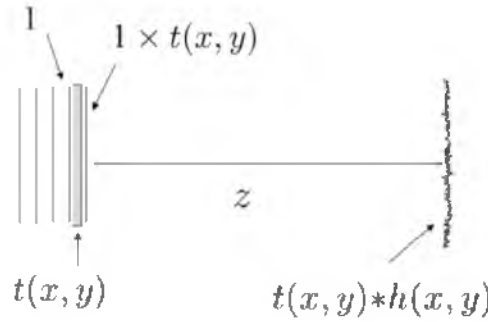
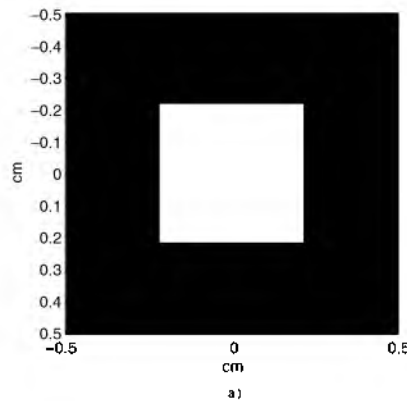


Fig. 2.10 Fresnel diffraction of an arbitrary beam profile  $t(x, y)$ .

Let us further consider a specific case where  $t(x, y) = \text{rect}(x/a, y/a)$ , a square aperture, is used for MATLAB simulations. We then implement  $\psi_p(x, y; z) = \psi_{i0}(x, y) * h(x, y; z)$  in the spatial frequency domain, i.e., using Eq. (2.3-14), where  $\psi_{i0}(x, y)$  is  $t(x, y)$  and is given by  $\text{rect}(x/a, y/a)$  with  $a = 0.4336\text{cm}$ . The m-file, `Fresnel_diffraction.m` shown in Table 2.1, generates the three figures shown below. Figure 2.11a) shows the square aperture,  $\text{rect}(x/a, y/a)$ , which is illuminated by a plane wave of red wavelength ( $\lambda_0 = 0.6328 \times 10^{-4}\text{cm}$ ). Figure 2.11b) and c) show the central cross-section of the square aperture, i.e.,  $|\psi_p(x, 0; 0)|$ , and the Fresnel diffracted magnitude, i.e.,  $|\psi_p(x, 0; z)|$ , at  $z = 5\text{cm}$ , respectively.



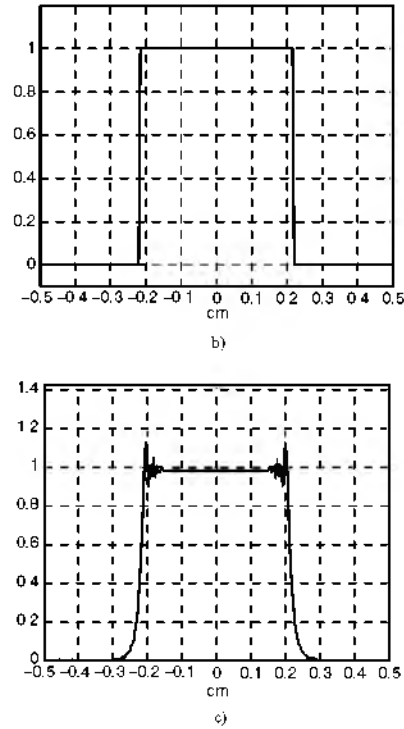


Fig. 2.11 a) Square aperture, b) Central cross-section of a),  
c) Central cross-section of diffracted amplitude at  $z = 5$  cm.

Table 2.1 Fresnel\_diffraction.m:  
m-file for calculating the Fresnel diffraction of a square aperture.

```

-----
%Fresnel_diffraction.m
%Simulation of Fresnel diffraction of a square aperture
%Adapted from "Contemporary optical image processing with MATLAB®,"
%by T.-C. Poon and P. P. Banerjee, Elsevier 2001, pp. 64-65.
clear

L=1; %L : length of display area
N=256; %N : number of sampling points
dx=L/(N-1); % dx : step size

%Create square image, M by M square, rect(x/a), M=odd number
M=111;
a=M/256
R=zeros(256); %assign a matrix (256x256) of zeros
r=ones(M); %assign a matrix (MxM) of ones
n=(M-1)/2;
R(128-n:128+n,128-n:128+n)=r;
%End of creating input image

%Axis Scaling

```

```

for k=1:256
    X(k)=1/255*(k-1)-L/2;
    Y(k)=1/255*(k-1)-L/2;

    %Kx=(2*pi*k)/((N-1)*dx)
    %in our case, N=256, dx=1/255

    Kx(k)=(2*pi*(k-1))/((N-1)*dx)-((2*pi*(256-1))/((N-1)*dx))/2;
    Ky(k)=(2*pi*(k-1))/((N-1)*dx)-((2*pi*(256-1))/((N-1)*dx))/2;
end

%Fourier transformation of R

FR=(1/256)^2*fft2(R);
FR=fftshift(FR);

%Free space impulse response function
% The constant factor exp(-jk0*z) is not calculated
%sigma=ko/(2*z)=pi/(wavelength*z)
%z=5cm, red light=0.6328*10^-4(cm)
sigma=pi/((0.6328*10^-4)*5);

for r=1:256,
    for c=1:256,
        %compute free-space impulse response with Gaussian apodization against aliasing
        h(r,c)=j*(sigma/pi)*exp(-4*200*(X(r).^2+Y(c).^2))*exp(-j*sigma*(X(r).^2+Y(c).^2));
    end
end

H=(1/256)^2*fft2(h);
H=fftshift(H);
HR=FR.*H;
H=(1/256)^2*fft2(h);
H=fftshift(H);
HR=FR.*H;

hr=ifft2(HR);
hr=(256^2)*hr;
hr=fftshift(hr);

%Image of the rectangle object
figure(1)
image(X,Y,255*HR);
colormap(gray(256));
axis square
xlabel('cm')
ylabel('cm')

% plot of cross section of square
figure(2)
plot(X+dx/2,R(:,127))
grid

```

```

axis([-0.5 0.5 -0.1 1.2])
xlabel('cm')

figure(3)
plot(X+dx/2,abs(hr(:,127)))
grid
axis([-0.5 0.5 0 max(max(abs(hr)))*1.1])
xlabel('cm')

```

---

## 2.4 Ideal Lens, Imaging Systems, Pupil Functions and Transfer Functions

### 2.4.1 Ideal Lens and Optical Fourier Transformation

In the previous section, we have discussed light diffraction by apertures. In this section, we will discuss the passage of light through an *ideal lens*. An ideal lens is a phase object. When an ideal focusing (or convex) lens has focal length  $f$ , its phase transformation function,  $t_f(x, y)$ , is given by

$$t_f(x, y) = \exp\left[j\frac{k_0}{2f}(x^2 + y^2)\right], \quad (2.4-1)$$

where we have assumed that the ideal lens is infinitely thin. For a uniform plane wave incident upon the lens, the wavefront behind the lens is a converging spherical wave (for  $f > 0$ ) that converges ideally to a point source (a distance of  $z = f$ ) behind the lens. We can see that this is the case when we apply the Fresnel diffraction formula [see Eq. (2.3-12)] :

$$\psi_p(x, y, z = f) = \psi_{p0}(x, y) * h(x, y; z = f), \quad (2.4-2)$$

where  $\psi_{p0}(x, y)$  is now given by  $1 \times t_f(x, y)$ . The constant, 1, in front of  $t_f(x, y)$  signifies that we have a plane wave (of unit amplitude) incident. For example, if we have an incident *Gaussian beam* of the profile given by  $\exp[-a(x^2 + y^2)]$ , then  $\psi_{p0}(x, y)$  will be given by  $\exp[-a(x^2 + y^2)] \times t_f(x, y)$ . Let us now return to Eq. (2.4-2) where  $\psi_{p0}(x, y) = t_f(x, y)$ , and by using Eq. (2.3-12) we have

$$\begin{aligned}
 \psi_p(x, y; f) &= \exp(-jk_0 f) \frac{jk_0}{2\pi f} \iint t_f(x', y') \\
 &\quad \times \exp\left\{-\frac{jk_0}{2f}[(x - x')^2 + (y - y')^2]\right\} dx' dy' \\
 &\propto \iint 1 \exp\left[\frac{jk_0}{2f}(x'^2 + y'^2)\right]
 \end{aligned}$$

$$\begin{aligned} & \times \exp\left[\frac{-jk_0}{2f}(x'^2 + y'^2 - 2xx' - 2yy')\right] dx' dy' \\ & = \iint 1 \exp\left[\frac{jk_0}{f}(xx' + yy')\right] dx' dy', \end{aligned}$$

which is recognized to be proportional to a 2-D Fourier transform of 1, i.e.,  $\delta(x, y)$ .

Let us now investigate the effect of placing a transparency,  $t(x, y)$ , against the ideal lens, which is shown in Fig. 2.12. In general,  $t(x, y)$  is a complex function such that if a complex field,  $\psi_{i,p}(x, y)$ , is incident on it, then the field immediately behind the transparency-lens combination is

$$\psi_{i,p}(x, y) t(x, y) t_f(x, y) = \psi_{i,p}(x, y) t(x, y) \exp\left[j\frac{k_0}{2f}(x^2 + y^2)\right].$$

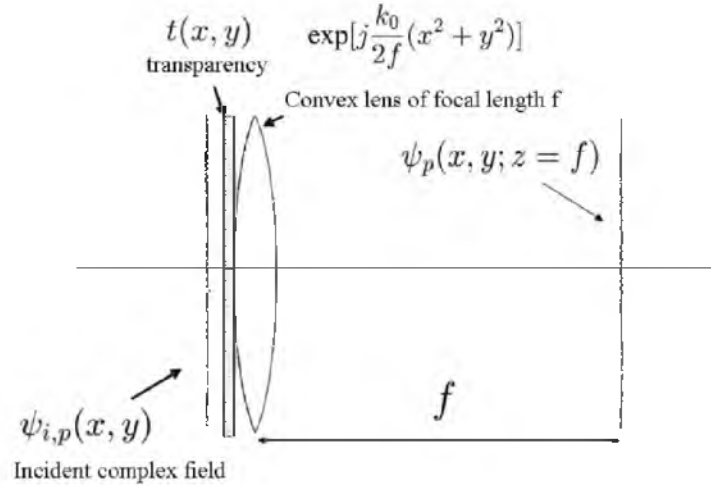


Fig. 2.12 A transparency immediately before an ideal lens under complex field illumination.

Again, for brevity, when illuminated by a unit-amplitude plane wave, the field immediately behind the combination is given by  $1 \times t(x, y) \exp[j\frac{k_0}{2f}(x^2 + y^2)]$ . We can then find the field distribution at a distance  $z = f$  by using the Fresnel diffraction formula, Eq. (2.3-12), as

$$\psi_p(x, y; z = f) = \exp(-jk_0 f) \frac{jk_0}{2\pi f} \exp\left[\frac{-jk_0}{2f}(x^2 + y^2)\right]$$



$$\begin{aligned}
& \times \int \int t(x', y') \exp \left[ j \frac{k_0}{f} (xx' + yy') \right] dx' dy' \\
& = \exp(-jk_0 f) \frac{jk_0}{2\pi f} \exp \left[ \frac{-jk_0}{2f} (x^2 + y^2) \right] \\
& \quad \times \mathcal{F}_{xy} \left\{ t(x, y) \right\} \bigg|_{\substack{k_x = k_0 x/f \\ k_y = k_0 y/f}}. \tag{2.4-3}
\end{aligned}$$

where  $x$  and  $y$  denote the transverse coordinates at  $z = f$ . Hence, the complex field on the focal plane ( $z = f$ ) is proportional to the Fourier transform of  $t(x, y)$ , but has *phase curvature* term  $\exp[\frac{-jk_0}{2f}(x^2 + y^2)]$ . Note that if  $t(x, y) = 1$ , i.e., the transparency is completely clear, then we have  $\psi_p(x, y, z = f) \propto \delta(x, y)$ , which corresponds to the focusing of a plane wave by a lens, as discussed earlier.

### Example 2.5 Transparency in front of a Lens

Suppose that a transparency,  $t(x, y)$ , is located at a distance,  $d_0$ , in front of an ideal convex lens and is illuminated by a plane wave with a unit strength shown in Fig. 2.13. The physical situation is shown in Fig. 2.13a), which can be represented by a block diagram given by Fig. 2.13b). According to the block diagram, we write

$$\psi_p(x, y; f) = \{ [t(x, y) * h(x, y; d_0)] t_f(x, y) \} * h(x, y; f), \tag{2.4-4}$$

which, apart from some constant, can be evaluated to obtain

$$\begin{aligned}
\psi_p(x, y; f) &= \frac{jk_0}{2\pi f} \exp[-jk_0(d_0 + f)] \exp \left[ -j \frac{k_0}{2f} \left(1 - \frac{d_0}{f}\right) (x^2 + y^2) \right] \\
& \quad \times \mathcal{F}_{xy} \{ t(x, y) \} \bigg|_{\substack{k_x = k_0 x/f \\ k_y = k_0 y/f}}. \tag{2.4-5}
\end{aligned}$$

As in Eq. (2.4-3), note that a phase curvature factor as a function of  $x$  and  $y$  again precedes the Fourier transform, which represents the phase error if one wishes to compute the optical Fourier transformation. However, the phase curvature vanishes for the special case of  $d_0 = f$ . That is, from Eq. (2.4-5) and by disregarding some inessential constant, we now have

$$\psi_p(x, y; f) = \mathcal{F}_{xy}\{t(x, y)\} \Big|_{\substack{k_x = k_0 x/f \\ k_y = k_0 y/f}} = T(k_0 x/f, k_0 y/f). \quad (2.4-6)$$

Therefore, when the transparency is placed in the front focal plane of the convex lens, the phase curvature disappears and we recover the exact Fourier transform on the back focal plane. *Fourier-plane processing* of an “input” transparency located on the front focal plane can now be performed on the back focal plane. This is the essence of Fourier optics to perform *coherent image processing*.

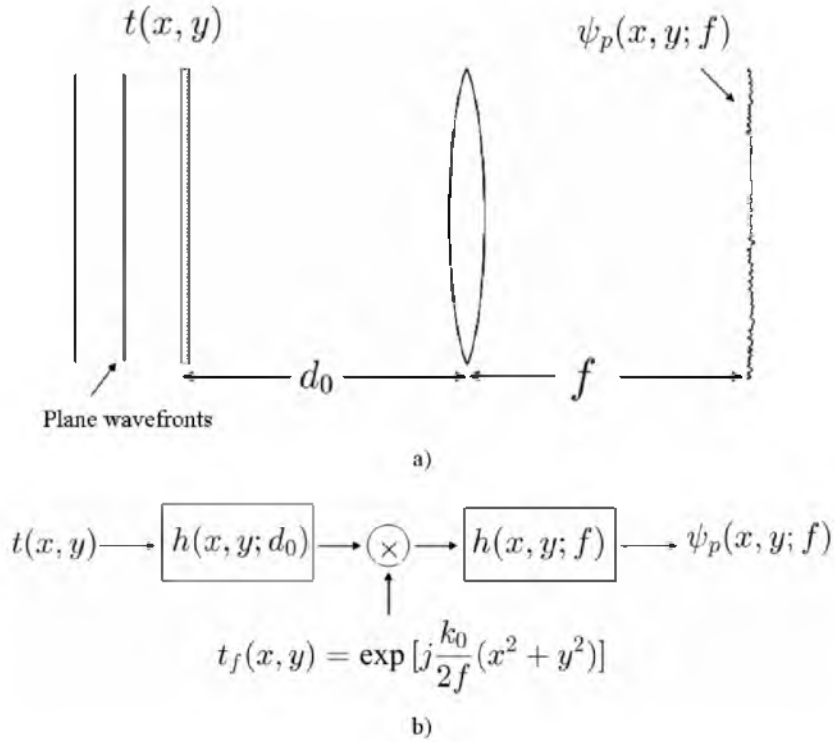


Fig. 2.13 Plane-wave illumination of a transparency  $t(x, y)$  located a distance  $d_0$  in front of a convex lens of focal length  $f$ : a) Physical situation, b) Block diagram.

### 2.4.2 Coherent Image Processing

The two-lens system is traditionally attractive for coherent image processing because, in the configuration shown in the Fig. 2.14, the Fourier transform of the input transparency,  $t(x, y)$ , appears on the common focal plane, or *Fourier plane*. In order to perform Fourier-plane processing of the input

transparency, we can insert a transparency on the Fourier plane that will suitably modify the Fourier transform of the input transparency.

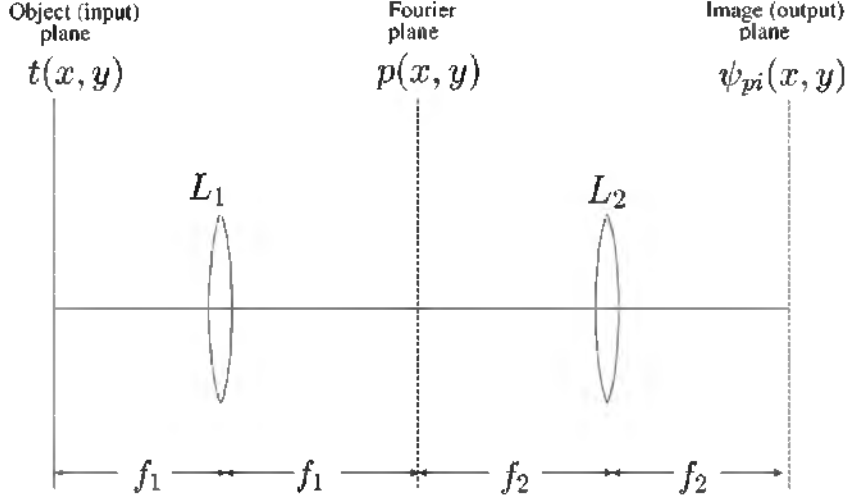


Fig. 2.14 Standard two-lens imaging processing system.

The Fourier plane transparency is commonly called a *spatial filter*,  $p(x, y)$ . According to Eq. (2.4-6), when a transparency  $t(x, y)$  is placed on the front focal plane of lens  $L_1$  as shown in Fig. 2.14, the field distribution on the common focal plane is given by  $T(k_0x/f_1, k_0y/f_1)$ , where we assume that the transparency is illuminated by a plane wave. After this field distribution is modified by the spatial filter, we can finally find the field distribution on the back focal plane of lens  $L_2$ ,  $\psi_{pi}$ , by again using Eq. (2.4-6) and neglecting some constant, as

$$\psi_{pi}(x, y) = \mathcal{F}_{xy} \left\{ T(k_0x/f_1, k_0y/f_1) p(x, y) \right\} \Big|_{\substack{k_x = k_0x/f_2 \\ k_y = k_0y/f_2}},$$

which can be evaluated, in terms of convolution, to give

$$\begin{aligned} \psi_{pi}(x, y) &= t(x/M, y/M) * \mathcal{F}_{xy} \left\{ p(x, y) \right\} \Big|_{\substack{k_x = k_0x/f_2 \\ k_y = k_0y/f_2}} \\ &= t(x/M, y/M) * P\left(\frac{k_0x}{f_2}, \frac{k_0y}{f_2}\right), \end{aligned} \quad (2.4-7)$$

where  $M = -f_2/f_1$  is the magnification factor and  $P$  is the Fourier transform of  $p$ . By comparing Eq. (2.4-7) with Eq. (1.2-2), we can describe the impulse response of the two-lens system, or the *coherent point spread function* (CPSF), as

$$h_c(x, y) = \mathcal{F}_{xy} \left\{ p(x, y) \right\} \Big|_{\substack{k_x = k_0 x / f_2 \\ k_y = k_0 y / f_2}} = P\left(\frac{k_0 x}{f_2}, \frac{k_0 y}{f_2}\right). \quad (2.4-8)$$

$p(x, y)$  is often called the *pupil function* of the system. We can see that the coherent PSF is given by the Fourier transform of the pupil function as shown in Eq. (2.4-8). By definition, the corresponding *coherent transfer function* is the Fourier transform of the coherent PSF:

$$\begin{aligned} H_c(k_x, k_y) &= \mathcal{F}_{xy} \left\{ h_c(x, y) \right\} \\ &= \mathcal{F}_{xy} \left\{ P\left(\frac{k_0 x}{f_2}, \frac{k_0 y}{f_2}\right) \right\} = p\left(\frac{-f_2 k_x}{k_0}, \frac{-f_2 k_y}{k_0}\right). \end{aligned} \quad (2.4-9)$$

We observe that *spatial filtering* is directly proportional to the functional form of the pupil function in coherent image processing.

The complex field on the image plane can then be written as

$$\psi_{pi}(x, y) \propto t(x/M, y/M) * h_c(x, y), \quad (2.4-10)$$

and hence the corresponding *image intensity* is

$$I_i(x, y) = |\psi_{pi}(x, y)|^2 \propto |t(x/M, y/M) * h_c(x, y)|^2. \quad (2.4-11)$$

### 2.4.3 Incoherent Image Processing

So far, we have discussed that the illumination of an object is *spatially coherent* - an example being the use of a laser. This means that the complex amplitudes of light falling on all parts of an object vary in unison, meaning that any two points on an object receive light that has a fixed relative phase and does not vary with time. On the other hand, an object may be illuminated with light having the property that the complex amplitudes on all parts of the object vary randomly, so that any two points on the object receive light of illumination is termed *spatially incoherent*. Light from extended sources, such as fluorescent tube lights, is incoherent. As it turns out, a *coherent system is linear with respect to the complex fields* and hence Eqs. (2.4-10) and (2.4-11) hold for *coherent optical systems*. On the other hand, an *incoherent optical system is linear with respect to the intensities*. To find the image intensity, we perform convolution with the given intensity quantities as follows:

$$I_i(x, y) \propto |t(x/M, y/M)|^2 * |h_c(x, y)|^2. \quad (2.4-12)$$

This equation is the basis for incoherent image processing.  $|h_c(x, y)|^2$  is the impulse response of the incoherent optical system and is often called the *intensity point spread function (IPSF)* of the optical system. Note that the IPSF is real and non-negative, which particularly means that it is not possible to directly implement even the simplest enhancement and restoration algorithms (e.g., highpass, derivative, etc.), which requires a *bipolar PSF* [Lohmann and Rhodes (1978)].

As usual, the Fourier transform of an impulse response will give a transfer function known as the *optical transfer function (OTF)* of the incoherent imaging system. For this case, it is given by

$$OTF(k_x, k_y) = \mathcal{F}_{xy}\{|h_c(x, y)|^2\} = H_c(k_x, k_y) \otimes H_c(k_x, k_y), \quad (2.4-13)$$

which can be explicitly written in terms of the coherent transfer function  $H_c$ :

$$OTF(k_x, k_y) = \iint H_c^*(k'_x, k'_y) H_c(k'_x + k_x, k'_y + k_y) dk'_x dk'_y. \quad (2.4-14)$$

Note that one of the most important properties of the *OTF*, which follows a property of correlation, is that

$$|OTF(k_x, k_y)| \leq |OTF(0, 0)|. \quad (2.4-15)$$

This property states that the OTF always has a central maximum, which always signifies *lowpass filtering* regardless of the pupil function used in the system [Lukosz (1962)].

### Example 2.6 Coherent Transfer Functions and OTFs

Consider a two-lens system as shown in Fig. 2.14 with  $f_1 = f_2 = f$  and  $p(x, y) = \text{rect}(x/X)$ , i.e., a slit of width  $X$  along the  $y$ -direction. Using Eq. (2.4-9), the coherent transfer function becomes

$$\begin{aligned} H_c(k_x, k_y) &= \text{rect}\left(\frac{x}{X}\right) \Big|_{x=-fk_x/k_0} \\ &= \text{rect}\left(\frac{k_x}{Xk_0/f}\right), \end{aligned} \quad (2.4-16)$$

which is plotted in Fig. 2.15a). Now, the OTF is the autocorrelation of  $H_c$  as calculated by Eq. (2.4-13) and is plotted in Fig. 2.15b). Observe that both

situation perform *lowpass filtering* of spatial frequencies on an input image. Under incoherent illumination, it is possible to transmit twice the range of spatial frequency of an image as compared to the use of coherent illumination. However, the spectrum of an image transmitted through the pass-band is modified by the shape of the OTF.

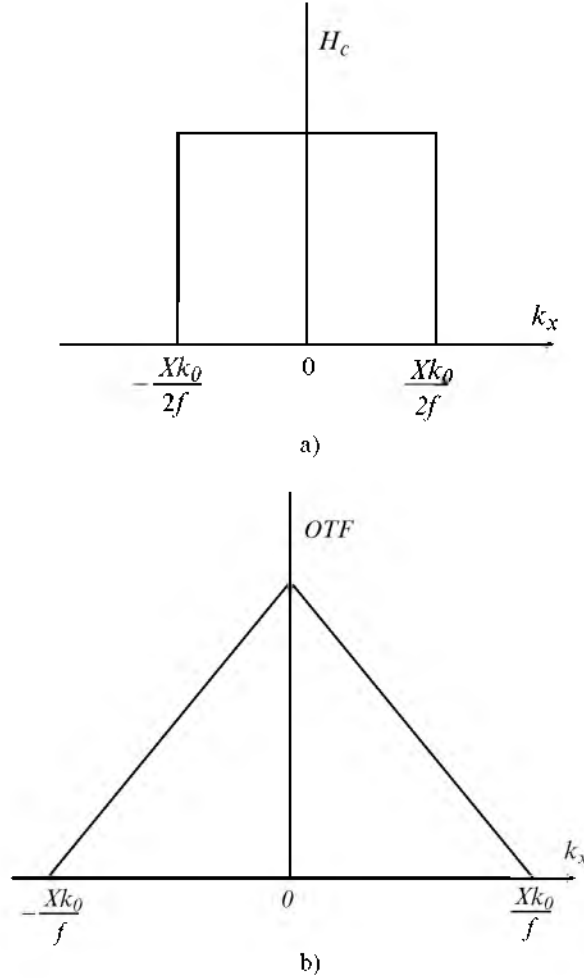


Fig. 2.15 a) The coherent transfer function, and b) the OTF for the pupil function  $p(x, y) = \text{rect}(x/X)$ .

Now, let us consider

$$p(x, y) = \left[ \text{rect}\left(\frac{x - x_0}{X}\right) + \text{rect}\left(\frac{x + x_0}{X}\right) \right], \quad x_0 > \frac{X}{2},$$

which is a two-slit object aligned along the  $y$ -direction. The coherent transfer function is

$$\begin{aligned}
 H_c(k_x, k_y) &= \left[ \text{rect}\left(\frac{x - x_0}{X}\right) + \text{rect}\left(\frac{x + x_0}{X}\right) \right] \Big|_{x = -f k_z / k_0} \\
 &= \left[ \text{rect}\left(\frac{-k_x - x_0 k_0 / f}{X k_0 / f}\right) + \text{rect}\left(\frac{-k_x + x_0 k_0 / f}{X k_0 / f}\right) \right].
 \end{aligned}$$

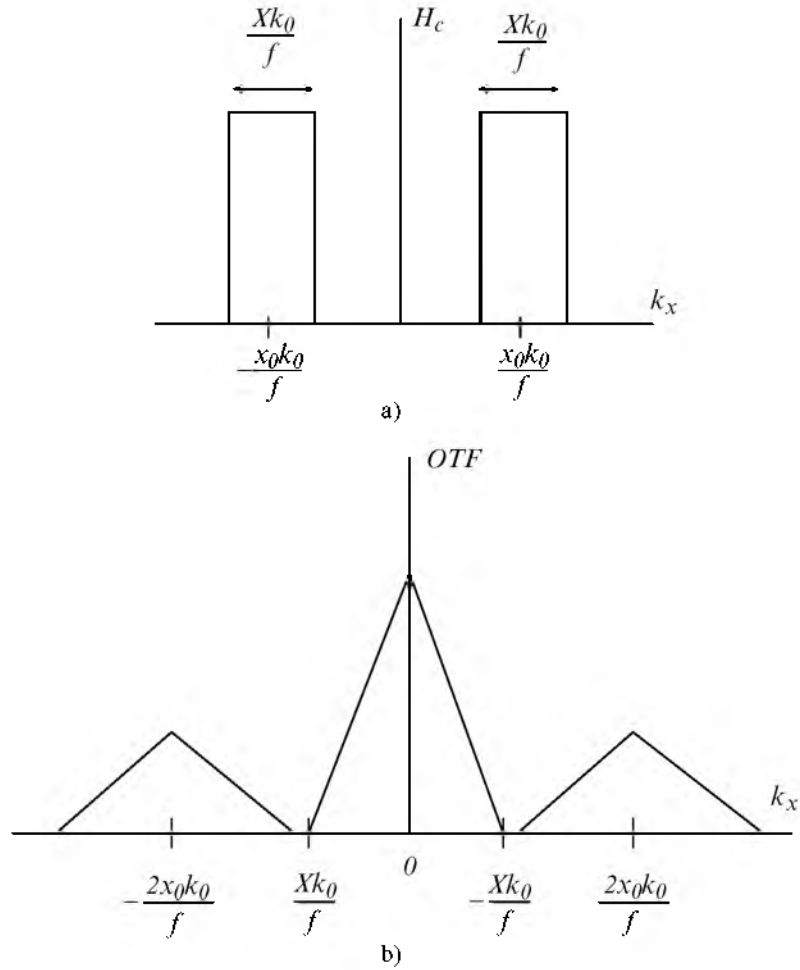


Fig. 2.16 a) The coherent transfer function, and b) the OTF for the pupil function  $p(x, y) = \text{rect}[(x - x_0)/X] + \text{rect}[(x + x_0)/X]$ .

We plot  $H_c(k_x, k_y)$  in Fig. 2.16a) along with the OTF in Fig. 2.16b). Note that even though it may be possible to achieve *band-pass filtering* with

coherent illumination, incoherent processing always gives rise to inherently low-pass characteristics because its point spread function is real and positive [see Eq. (2.4-13)]. A large amount of attention has been focused on devising methods to realize band-pass characteristics by using novel incoherent image processing techniques [see, e.g., Lohmann and Rhodes (1978), Stoner (1978), Poon and Korpel (1979), Mait (1987)], where the synthesis of bipolar or even complex point spread functions (PSFs) in incoherent optical systems is possible. Such techniques are called *bipolar incoherent image processing*. The article by Indebetouw and Poon [1992] provides a comprehensive review of bipolar incoherent image processing.

## 2.5 Holography

### 2.5.1 Fresnel Zone Plate as a Point-Source Hologram

A photograph is a 2-D recording of a 3-D scene. What is actually recorded is the light intensity at the plane of the photographic recording film - the film being light sensitive only to the intensity variations. Hence, the developed film's amplitude transparency is  $t(x, y) \propto I(x, y) = |\psi_p|^2$ , where  $\psi_p$  is the complex field on the film. As a result of this intensity recording, all the information on the relative phases of light waves from the original 3-D scene is lost. This loss of phase information on the light field destroys the 3-D character of the scene, i.e., we cannot change the perspective of the image in the photograph by viewing it from a different angle (i.e., *parallax*) and we cannot interpret the depth of the original 3-D scene.

As an example, let us take the photographic recording of a point source located at the origin, but with a distance of  $z_0$  away from the film. The situation is shown in Fig. 2.17a). Now, according to Eq. (2.3-15), the complex field just before the film is given by

$$\begin{aligned}\psi_p(x, y; z_0) &= \delta(x, y) * h(x, y; z_0) \\ &= \exp(-jk_0 z_0) \frac{jk_0}{2\pi z_0} \exp\left[-\frac{jk_0(x^2 + y^2)}{2z_0}\right].\end{aligned}$$

Hence, the developed film's amplitude transparency is

$$t(x, y) \propto I(x, y) = |\psi_p(x, y; z_0)|^2 = \left(\frac{k_0}{2\pi z_0}\right)^2. \quad (2.5-1)$$

Note that the phase information of  $\psi_p(x, y; z_0)$  is completely lost. Now, for a point source located at  $(x_0, y_0)$ , as shown in Fig. 2.17b), the complex field just before the film is given by



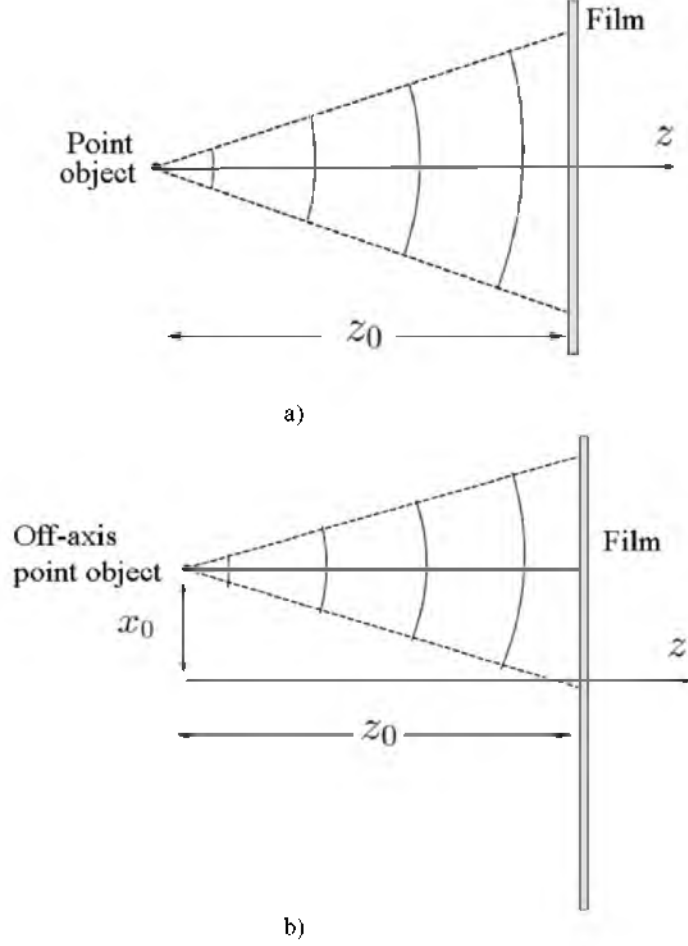


Fig. 2.17 Photographic recording of a point source: a) located at the origin  $(0, 0)$ , and b) located at  $(x_0, y_0)$ , both  $z_0$  away from the film.

$$\begin{aligned}\psi_p(x, y; x_0, y_0, z_0) &= \delta(x - x_0, y - y_0) * h(x, y; z_0) \\ &= \exp(-jk_0 z_0) \frac{jk_0}{2\pi z_0} \exp\left[-\frac{jk_0[(x - x_0)^2 + (y - y_0)^2]}{2z_0}\right],\end{aligned}$$

and what is recorded is

$$t(x, y) \propto I(x, y) = |\psi_p(x, y; x_0, y_0, z_0)|^2 = \left(\frac{k_0}{2\pi z_0}\right)^2, \quad (2.5-2)$$

which is identical to the result given by Eq. (2.5-1). Again the phase information of  $\psi_p(x, y; x_0, y_0, z_0)$  is lost, and we also notice that the 3-D location of the point source, i.e.,  $x_0, y_0$ , and  $z_0$ , is mostly lost.

Holography is an extraordinary technique that was invented by Gabor [1948], where not only the amplitude, but also the phase of a light field can be recorded. The word “holography” combines parts of two Greek words: *holos*, meaning “complete,” and *graphein*, meaning “to record.” Thus, holography means the recording of complete information. Hence, in the holographic process, the film records both the amplitude and phase of a light field. The resulting recorded film is called a “*hologram*.” When a hologram is properly illuminated, an exact replica of the original 3-D wave field is reconstructed. We shall discuss the *holographic recording* of a point object as an example. Once we know how a single point is recorded, the recording of a complicated object can be regarded as the recording of a collection of points.

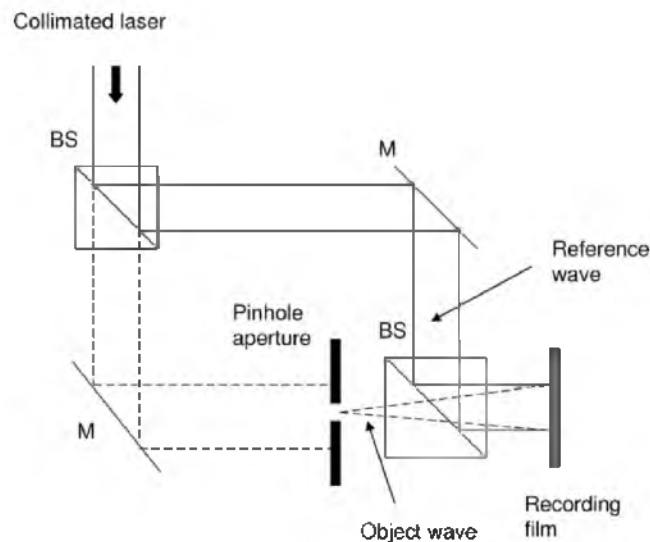


Fig. 2.18 Holographic recording of a point source object.

Figure 2.18 shows a collimated laser which is split into two plane waves and then recombined by using two mirrors (M) and two beam splitters (BS). One plane wave is used to illuminate the pinhole aperture (our point object), and the other is used to illuminate directly the recording film. The plane wave that is diffracted by the pinhole aperture generates a diverging spherical wave. In holography, this diverging wave is known as an *object wave*. The plane wave that directly illuminates the photographic plate is known as a *reference wave*. Let  $\psi_o$  represent the field distribution of the object wave on the plane of the recording film, and similarly, let  $\psi_r$  represent the field distribution of the reference wave on the plane of the recording film. The film now records the interference of the reference wave and the object wave, i.e., what is recorded is given by  $|\psi_r + \psi_o|^2$ , provided that the reference wave and the object wave are mutually coherent over the film. The

coherency of the light waves is guaranteed by the use of a laser source and ensures that the difference between the two paths is less than the coherent length of the laser. This kind of recording is known as *holographic recording*, and is distinct from a photographic recording where the reference wave does not exist and, therefore, only the object wave is recorded.

Let us now consider the recording of an off-axis point object at a distance of  $z_0$  from the recording film. The pinhole aperture is then modeled as  $\delta(x - x_0, y - y_0)$ . According to Fresnel diffraction, the object wave arises from the point object on the film and is given by

$$\begin{aligned}\psi_o &= \delta(x - x_0, y - y_0) * h(x, y; z_0) \\ &= \exp(-jk_0 z_0) \frac{jk_0}{2\pi z_0} \exp\{-jk_0[(x - x_0)^2 + (y - y_0)^2]/2z_0\}.\end{aligned}$$

This object wave is a *spherical wave*.

For the reference plane wave, we assume that the plane wave has the same initial phase as the point object at a distance of  $z_0$  away from the film. Therefore, its field distribution on the film is  $\psi_r = a \exp(-jk_0 z_0)$ , where  $a$  is the amplitude of the plane wave. Hence, the intensity distribution that is being recorded on the film, or the transmittance of the hologram, is given by

$$\begin{aligned}t(x, y) &\propto |\psi_r + \psi_o|^2 \\ &= \left| a + \frac{jk_0}{2\pi z_0} \exp\{-jk_0[(x - x_0)^2 + (y - y_0)^2]/2z_0\} \right|^2 \\ &= A + B \sin\left\{ \frac{k_0}{2z_0} [(x - x_0)^2 + (y - y_0)^2] \right\} \\ &= FZP(x - x_0, y - y_0; z_0)\end{aligned}\tag{2.5-3}$$

where  $A = a^2 + \left(\frac{k_0}{2\pi z_0}\right)^2$ ,  $B = \frac{k_0}{\pi z_0}$  and  $k_0 = 2\pi/\lambda_0$ .

The expression in Eq. (2.5-3) is called the sinusoidal *Fresnel zone plate (FZP)*, which is a hologram of the point source object. Note that the center of the zone plate specifies the location,  $x_0$  and  $y_0$ , of the point object, and the spatial variation of the zone plate is governed by a sine function with a quadratic spatial dependence. For an on-axis point source, i.e.,  $x_0 = y_0 = 0$  in Eq. (2.5-3), located a distance of  $z_0$  away from the film, we have an on-axis Fresnel zone plate as

$$\begin{aligned}
t(x, y) &\propto |\psi_r + \psi_o|^2 \\
&= \left| a + \frac{jk_0}{2\pi z_0} \exp\left[-\frac{jk_0}{2z_0}(x^2 + y^2)\right] \right|^2 \\
&= A + B \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right] \\
&= FZP(x, y; z_0).
\end{aligned} \tag{2.5-4}$$

Let us now investigate the quadratic spatial dependence of  $FZP(x, y; z_0)$ . The spatial rate of change of the phase on the zone plate along the  $x$ -direction is

$$f_{local} = \frac{1}{2\pi} \frac{d}{dx} \left( \frac{k_0}{2z_0} x^2 \right) = \frac{x}{\lambda_0 z_0}. \tag{2.5-5}$$

This is a local fringe frequency that increases linearly with the spatial coordinate,  $x$ . The farther it is away from the origin of the zone, the higher the frequency will be. So, for a fixed point (local) on the hologram, we can deduce the depth information,  $z_0$ , by finding the local fringe frequency for a given wavelength of light,  $\lambda_0$ . Therefore, we see that the depth information is encoded within the phase of the FZP. Figure 2.19 shows the dependence of the Fresnel zone plate characteristic as a function of the depth parameter  $z$  (for  $z = z_0$  and  $2z_0$ ). As the point source becomes further away from the recording film, the recorded FZP has a lower local fringe frequency.

Figure 2.20 shows us that as the point source moves to a new location  $x_0, y_0$ , the center of the zone plate translates accordingly. Hence, we see that the zone contains the complete 3-D information of the point source. The center of the zone,  $x_0$  and  $y_0$ , defines the transverse location of the point object, and the fringe variation defines the depth location,  $z_0$ . Table 2.2 shows the MATLAB code used to generate Fresnel zone plates that are presented in Figs. 2.19 and 2.20. For an arbitrary 3-D object, we can think of the object as a collection of points, and therefore, we can envision that we have a collection of zones on the hologram, where each zone carries the transverse location as well as the depth information of each individual point. In fact, a hologram has been considered as a type of Fresnel zone plate, and the holographic imaging process has been discussed previously in terms of zone plates [Rogers (1950), Siemens-Wapniarski and Parker Givens (1968)].

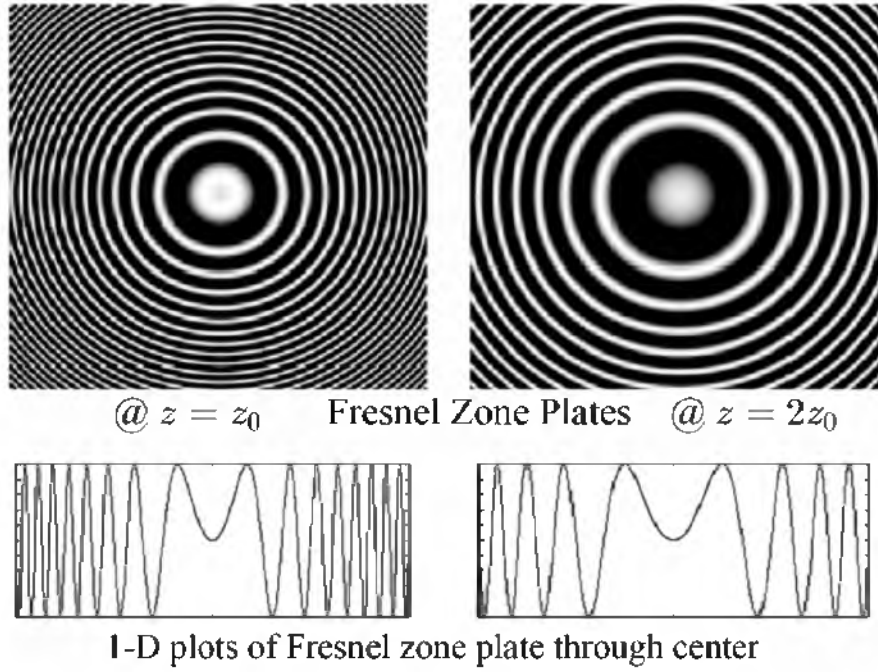


Fig. 2.19 On-axis Fresnel zone plate as a function of depth,  $z$ .

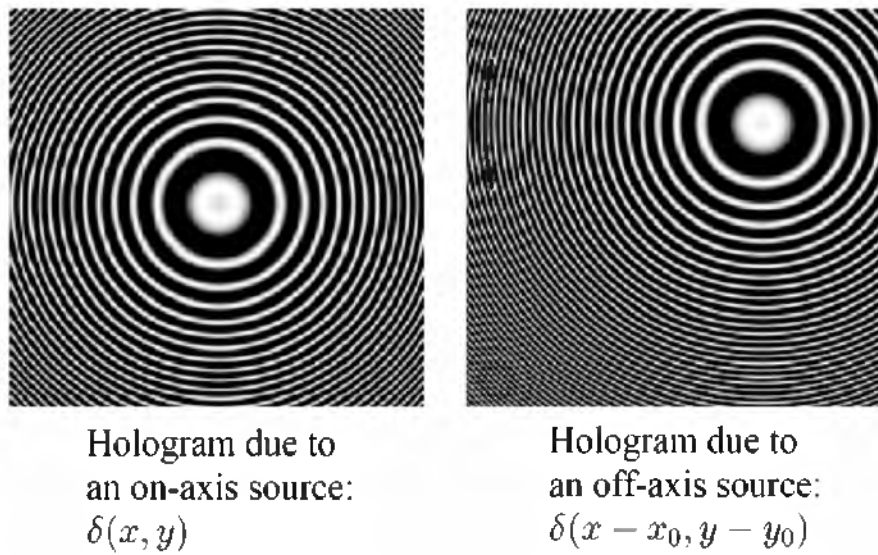


Fig. 2.20 Fresnel zone plate due to point sources at different transverse locations but otherwise located at the same depth  $z_0$ .

Table 2.2 Fresnel\_zone\_plate.m:  
m-file for calculating FZPs illustrated in Figs. 2.19 and 2.20.

```
-----
%Fresnel_zone_plate.m
%Adapted from "Contemporary optical image processing with MATLAB®,"
%by T.-C. Poon and P. P. Banerjee, Elsevier 2001, pp.177-178.
%
%display function is 1+sin(sigma*((x-x0)^2+(y-y0)^2)). All scales are arbitrary.
%sigma=pi/(wavelength*z)
clear;

z0=input('z0, distance from the point object to film, enter z0 (from 2 to 10)=');
x0=input('Inputting the location of the center of the FZP x0=y0,enter x0 (from -8 to 8) =');

ROWS=256;
COLS=256;
colormap(gray(255))
sigma=1/z0;
y0=-x0;
y=-12.8;

for r=1:COLS,
    x=-12.8;
    for c=1:ROWS, %compute Fresnel zone plate
        fFZP(r,c)=exp(j*sigma*(x-x0)*(x-x0)+j*sigma*(y-y0)*(y-y0));
        x=x+.1;
    end
    y=y+.1;
end

%normalization
max1=max(fFZP);
max2=max(max1);
scale=1.0/max2;
fFZP=fFZP.*scale;
R=127*(1+imag(fFZP));
figure(1)
image(R);
axis square on
axis off
-----
```

So far, we have discussed the transformation of a point object to a zone plate on the hologram, which corresponds to a *recording* or *coding process*. In order to retrieve the point object from the hologram, we need a *reconstruction* or *decoding process*. This can be done by simply illuminating the hologram with a *reconstruction wave*. Figure 2.21 corresponds to the reconstruction of a hologram of the point object located on-axis, i.e., the reconstruction of the hologram given by Eq. (2.5-4).

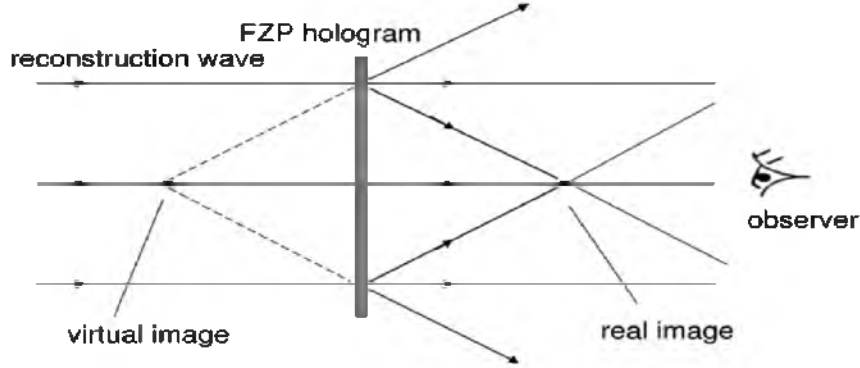


Fig. 2.21 Holographic reconstruction of a point source object.

Note that in practice as is shown in Fig. 2.21, the reconstruction wave is usually identical to the reference wave. Therefore, we take the reconstruction wave to have a field distribution on the plane of the hologram given by  $\psi_{\text{re}}(x, y) = a$ . Hence, the field distribution of the transmitted wave immediately after the hologram is  $\psi_{\text{re}}t(x, y) = at(x, y)$  and the field at arbitrary distance of  $z$  away, according to Fresnel diffraction, is given by the evaluation of

$$at(x, y) * h(x, y; z).$$

For the point-object hologram given by (2.5-4), after we expand the sine term of the hologram  $t(x, y)$ , we obtain

$$t(x, y) = A + \frac{B}{2j} \left\{ \exp\left[j \frac{k_0}{2z_0} (x^2 + y^2)\right] - \exp\left[-j \frac{k_0}{2z_0} (x^2 + y^2)\right] \right\}.$$

Therefore, as a result of the illumination of the hologram by the reconstruction wave, we have three waves. These waves, according to the convolution operation,  $at(x, y) * h(x, y; z)$ , are as follows:

*Zero-order beam:*

$$aA * h(x, y; z = z_0) = aA. \quad (2.5-6a)$$

*Real image (or the twin image):*

$$\sim \exp\left[j \frac{k_0}{2z_0} (x^2 + y^2)\right] * h(x, y; z = z_0) \sim \delta(x, y). \quad (2.5-6b)$$

*Virtual image:*

$$\sim \exp\left[-j\frac{k_0}{2z_0}(x^2 + y^2)\right] * h(x, y; z = -z_0) \sim \delta(x, y). \quad (2.5-6c)$$

By writing Eq. (2.5-6c), we have back-propagated the field immediately behind the hologram by a distance of  $z_0$  to demonstrate that a virtual image will form behind the hologram. As illustrated in Fig. 2.21, optical fields from this virtual image correspond to a diverging wave behind the hologram. We notice that the zero-order beam is caused by the bias in the hologram and the virtual image is the reconstructed original point object. The real image is located at a distance of  $z_0$  in front of the hologram, which is known as the *twin image*.

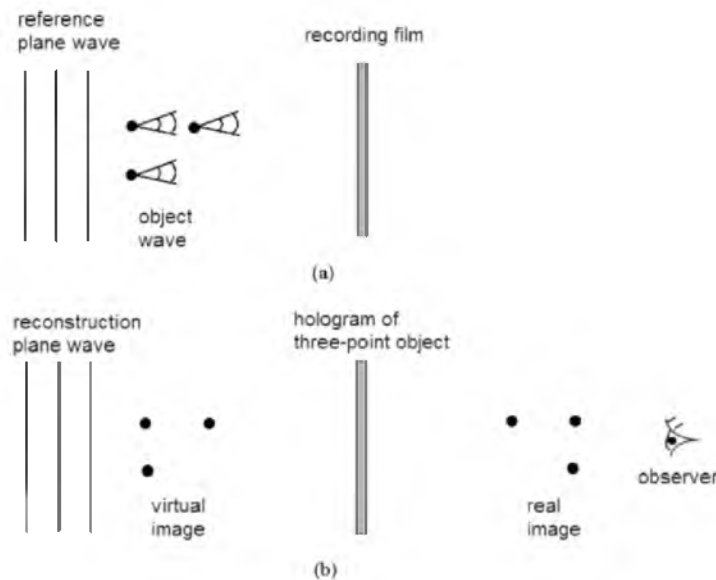


Fig. 2.22 Holographic recording and reconstruction of a three-point object.

Figure 2.22 shows the holographic recording of a 3-point object and its reconstruction. Note that the virtual image appears at the correct 3-D location as the original object, while the real image (the twin image) is the mirror-image of the original object, with the axis of reflection on the plane of the hologram.

### 2.5.2 Off-Axis Holography

In the last section, we discussed the so-called *on-axis holography*. The term “on-axis” refers to the use of a reference wave that is coaxially illuminating the hologram with the object wave. Although this technique can record 3-D



information of an object, it also create an annoying effect when we view the reconstructed virtual image. The real image (or the twin image) will also be reconstructed along the viewing direction [see Figs. 2.21 and 2.22]. In holography, this is infamously known as the “*twin-image problem*.”

*Off-axis holography* is a method that was devised by Leith and Upatnieks [1964] to separate the twin-image and the zero-order beam from the desired image. To achieve off-axis recording, the reference plane wave will need to be incident on the recording film off-axis. Referring back to Fig. 2.18, this can be done by simply, for example, rotating the beamsplitter (BS) between the pinhole aperture and the film in a clockwise direction so that the reference plane wave is incident on the film at an angle. The situation is shown in Fig. 2.23, where the plane reference wave is incident at an angle,  $\theta$ .  $\theta$  is called the *recording angle* in off-axis holographic recording.

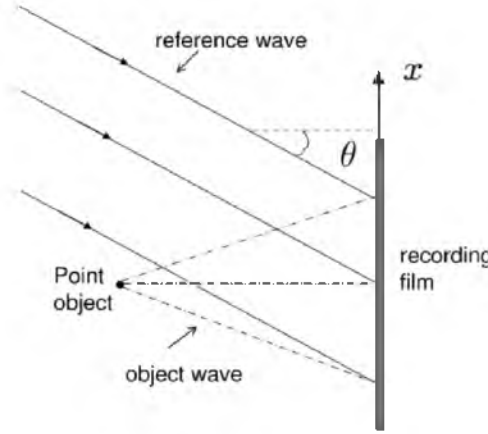


Fig. 2.23 Recording with off-axis reference plane wave.  
The point object is  $z_0$  away from the film.

For off-axis recording, we have  $t(x, y) = |\psi_r + \psi_o|^2$ , where the reference plane wave,  $\psi_r$ , is now an off-axis plane wave given by  $a \exp(jk_0 x \sin \theta)$ . The object wave,  $\psi_o$ , is the spherical wave generated by the on-axis point source. Similar to Eq. (2.5-3), where  $x_0 = y_0 = 0$  for an on-axis point object, we now have

$$\begin{aligned}
 t(x, y) &= \left| a \exp(jk_0 x \sin \theta) + \frac{jk_0}{2\pi z_0} \exp[-jk_0(x^2 + y^2)/2z_0] \right|^2 \\
 &= A + B \sin\left[\frac{k_0}{2z_0}(x^2 + y^2) + k_0 x \sin \theta\right], \quad (2.5-7)
 \end{aligned}$$

where  $A = a^2 + \left(\frac{k_0}{2\pi z_0}\right)^2$  and  $B = \frac{ak_0}{\pi z_0}$ .  $t(x, y)$  given by Eq. (2.5-7) is called

an *off-axis hologram*. Eq. (2.5-7) can be expanded into three terms as

$$t(x, y) = A + \frac{B}{2j} \left\{ \exp \left[ j \left[ \frac{k_0}{2z_0} (x^2 + y^2) + k_0 x \sin \theta \right] \right] \right. \\ \left. - \exp \left[ -j \left[ \frac{k_0}{2z_0} (x^2 + y^2) + k_0 x \sin \theta \right] \right] \right\}.$$

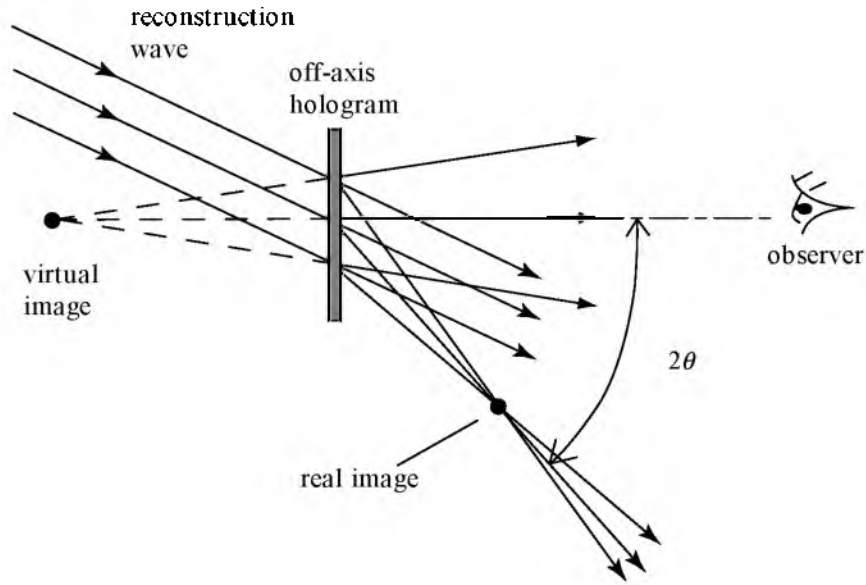


Fig. 2.24 Holographic reconstruction of off-axis hologram.  
The twin image (or the real image) is not observed if  $\theta$  is large enough.

By illuminating the hologram with a reconstruction wave identical to the reference wave, we have  $\psi_{rc} t(x, y)$  immediately after the hologram, where  $\psi_{rc} = a \exp(jk_0 x \sin \theta) = \psi_r$ . As in the case of on-axis holography, by performing Fresnel diffraction, we have  $\psi_{rc} t(x, y) * h(x, y; z)$ , thereby creating three waves as follows:

*Zero-order beam:*

$$A a \exp(jk_0 x \sin \theta) * h(x, y; z = z_0) \\ \sim \exp(jk_0 x \sin \theta). \quad (2.5-8a)$$

*Real image (or the twin image):*

$$\begin{aligned} & a \exp(jk_0 x \sin \theta) \exp \left\{ j \left[ \frac{k_0}{2z_0} (x^2 + y^2) + k_0 x \sin \theta \right] \right\} * h(x, y; z = z_0) \\ & \sim \delta(x + 2z_0 \sin \theta, y) . \end{aligned} \quad (2.5-8b)$$

*Virtual image:*

$$\begin{aligned} & a \exp(jk_0 x \sin \theta) \exp \left\{ -j \left[ \frac{k_0}{2z_0} (x^2 + y^2) + k_0 x \sin \theta \right] \right\} * h(x, y; z = -z_0) \\ & \sim \delta(x, y) . \end{aligned} \quad (2.5-8c)$$

The situation is depicted in Fig. 2.24.

### 2.5.3 Digital Holography

As discussed in the last section, in regards to off-axis holographic reconstruction, the three reconstructed beams propagate along different directions, and if the recording angle is sufficiently large, the virtual image can be viewed without any disturbances from the zero-order beam and the real image. This technique of off-axis recording is also known as *carrier-frequency holography*. We can re-write Eq. (2.5-7) as

$$t(x, y) = A + B \sin \left[ \frac{k_0}{2z_0} (x^2 + y^2) + 2\pi f_c x \right], \quad (2.5-9)$$

where  $f_c = k_0 \sin \theta / 2\pi = \sin \theta / \lambda_0$  is the *spatial carrier*. For realistic parameter values,  $\theta = 45^\circ$  and  $\lambda_0 = 0.6 \mu m$  for red laser light, we have  $\sin \theta / \lambda_0 \sim 1,000$  cycle/mm. This technique translates to a film resolution of at least 1000 lp/mm [or line-pair/mm] in order to employ this technique for holographic recording. Common holographic films have a resolution of about 5000 lp/mm. For comparison, standard black and white film resolution is about 80-100 lp/mm and color film is about 40-60 lp/mm. But can we use electronic devices such as CCD cameras for holographic recording? If we can do it, we can bypass the darkroom preparation of films and, therefore, we can perform real-time or electronic recording of holographic information. Some of the best CCD camera in the market, such as Canon D60 (3072x2048 pixels, 67.7 lp/mm, 7.4  $\mu m$  pixel size), could not record off-axis holograms efficiently because its resolution is about a couple of orders of magnitude worse than the resolution of holographic films. We can see that off-axis recording places a stringent resolution requirement on electronic recording

media. We can relax the resolution requirement by making the recording angle smaller, but this approach requires a very small recording angle that often makes it impractical. Because of this reason, on-axis holography seems to be prevalent in digital holography [Piestun, Shamir, Wekamp, and Bryngdahl (1997)]. On the other hand, twin-image problems need to be tackled when on-axis holography is employed. Indeed twin-image elimination is an important research topic [Poon et al. (2000)].

While we have discussed holographic recording electronically or digitally, for reconstruction, we can also perform it digitally. Once the holographic information is in the electronic or the digital domain, we can digitally evaluate Fresnel diffraction by performing the convolution, which is  $at(x, y) * h(x, y; z)$ , where  $a$  is some constant amplitude of the reconstruction beam,  $t(x, y)$  is the recorded hologram, and  $h(x, y; z)$  is the spatial impulse response in Fourier optics. For various values of  $z = z_1, z_2$ , etc., we can reconstruct different planes normal to the hologram. The whole 3-D volume of the object is then constructed plane by plane. The situation is shown in Fig. 2.25.

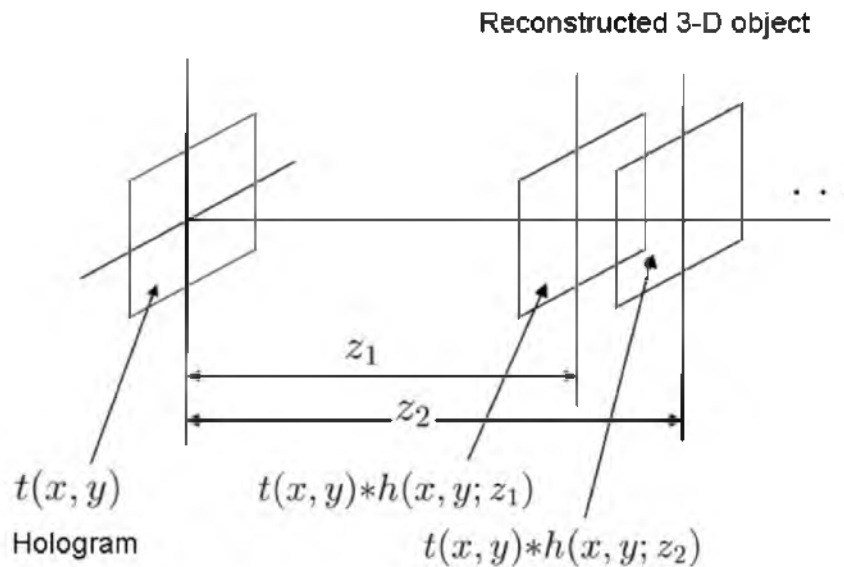


Fig. 2.25 Digital holographic reconstruction.

An alternative way to utilize electronically or digitally recorded hologram is to have it displayed on some sort of *spatial light modulator* (SLM) for real-time coherent reconstruction. A 2-D spatial light modulator is a device with which one can imprint a 2-D pattern on a laser beam by passing the laser beam through it (or by reflecting the laser beam off the device). A *liquid*

*crystal television* (LCTV) (upon suitably modification) is a good example of spatial light modulators. In fact, we can think of a spatial light modulator as a real-time transparency because one can update 2-D images or holograms upon the spatial light modulator in real time without developing films into transparencies. Again, off-axis holographic recording places stringent resolution requirement on SLMs and we will come back to this later when discussing 3-D display applications in Chapter 4.

All in all, in this section we mention electronic or digital recording and manipulation of holographic information. This type of research is commonly known as *digital* (or *electronic*) *holography*. The reader may find a pioneering contribution in the work of Goodman and Lawrence [1967]. Ever since, digital holography has become a practical tool with an increasing number of applications [Schnars and Juptner (2002)]. Most recently, an edited book on the subject organizes a collection of key chapters that covers digital holography and 3-D display techniques so as to provide the reader with the state-of-the-art developments in these important areas around the world [Poon (2006)]. Starting in the next chapter, we will discuss a unique electronic holographic recording technique called *optical scanning holography*.

## References

- 2.1 Banerjee, P.P. and T.-C. Poon (1991). *Principles of Applied Optics*. Irwin, Illinois.
- 2.2 Gabor, D. (1948). "A new microscopic principle," *Nature*
- 2.3 Goodman, J. W. and R.W. Lawrence (1967). "Digital image formation from electronically detected holograms," *Applied Physics Letters* 11, 77-79.
- 2.4 Goodman, J. W. (2005). *Introduction to Fourier Optics*. 3rd. ed., Roberts and Company Publishers, Englewood, Colorado.
- 2.5 Indebetouw, G. and T.-C. Poon (1992). "Novel approaches of incoherent image processing with emphasis on scanning methods," *Optical Engineering* 31, 2159–2167.
- 2.6 E. N. Leith, E.N and J. Upatnieks (1962). "Reconstructed wavefronts and communication theory," *Journal of the Optical Society of America* 52, 1123–1130.
- 2.7 Lohmann A.W. and W. T. Rhodes (1978). "Two-pupil synthesis of optical transfer functions," *Applied Optics* 17, 1141-1150.
- 2.8 Lukosz, W. (1962). "Properties of linear low pass filters for non-negative signals," *Journal of the Optical Society of America* 52, 827-829.
- 2.9 Mait, J. N. (1987). "Pupil-function design for complex incoherent spatial filtering," *Journal of the Optical Society of America A* 4, 1185- 1193.
- 2.10 Piestun R., J. Shamir, B. Wekamp, and O. Bryngdahl (1997). "On-axis computer-generated holograms for three-dimensional display," *Optics Letters* 22, 922-924.
- 2.11 Poon, T.-C. and A. Korpel (1979). "Optical transfer function of an acousto-optic heterodyning image processor," *Optics Letters* 4, 317-319.
- 2.12 Poon, T.-C., T. Kim, G. Indebetouw, B. W. Schilling, M. H. Wu, K. Shinoda, and Y. Suzuki (2000). "Twin-image elimination experiments for three-dimensional images in optical scanning holography," *Optics Letters* 25, 215-217.

- 2.13** Poon T.-C. and P. P. Banerjee (2001). *Contemporary Optical Image Processing with MATLAB®*. Elsevier, Oxford, UK.
- 2.14** Poon, T.-C., ed., (2006). *Digital Holography and Three-Dimensional Display: Principles and Applications*. Springer, New York, USA.
- 2.15** Rogers, G.L. (1950). "The black and white holograms," *Nature* 166, 1027.
- 2.16** Siemens-Wapniarski, W.J. and M. Parker Givens (1968). "The experimental production of synthetic holograms," *Applied Optics* 7, 535-538.
- 2.17** Stoner, W. (1978). "Incoherent optical processing via spatially offset pupil masks," *Applied Optics* 17, 2454-2466.
- 2.18** Schnars, U. and W. P. O. Juptner (2002). "Digital recording and numerical reconstruction of holograms," *Meas. Sci. Technol.* 13, R85–R101.

## Chapter 3

### Optical Scanning Holography: Principles

*Optical scanning holography (OSH)* is a form of electronic (or digital) holography. It is a unique, real-time technique where holographic information of a three-dimensional (3-D) object can be acquired by using a single 2-D optical scan. OSH was first implicated by Poon and Korpel when they investigated bipolar incoherent image processing on their acousto-optic heterodyning image processor [1979]. The original idea is later formulated and becomes known as *scanning holography* [Poon (1985)]. The first experimental results were then demonstrated and the technique was eventually called optical scanning holography in order to emphasize the novel fact that holographic recording can be achieved by active optical scanning [Duncan and Poon (1992)]. Thus far, applications of OSH include scanning holographic microscopy [(Poon, Doh, Schilling, Wu, Shinoda, and Suzuki (1995))], 3-D image recognition [(Poon and Kim (1999))], 3-D optical remote sensing [Kim and Poon (1999)], 3-D TV and display [Poon (2002a)], and 3-D cryptography [Poon, Kim, and Doh (2003)]. Scanning holographic microscopy is, by far, the most developed technique that utilizes OSH. Unlike any other holographic microscopes, scanning holographic microscope has a unique property that allows it to take the holographic information of fluorescent specimens in three dimensions. Recently, scientists have been able to achieve better than one-micron resolution in holographic fluorescence microscopy [Indebetouw and Zhong (2006)]. While in chapters 1 and 2, we have covered the necessary backgrounds in mathematics and optics to better understand OSH, in this chapter, we discuss the basic principles of OSH. In chapter 4, we will then discuss some of the previously mentioned applications of OSH in detail. Finally in chapter 5, we will discuss some recent advances in OSH.

#### 3.1 Principle of Optical Scanning

Optical scanning holography involves active *optical scanning* and *optical heterodyning*. In this section, we will discuss the basics of optical scanning. An optical scanner or optical processor scans out a transparency, i.e., information, with an optical beam by moving either the beam or the transparency. A photodetector accepts all light and gives an electrical output

that can either be stored or displayed by some means or another. Hence, optical information will have been converted into electrical information.

Figure 3.1 shows a standard, active optical scanning image processing system. A plane wave (such as the use of a laser in practice) of frequency  $\omega_0$ , illuminates a pupil function,  $p(x, y)$ . The complex field emerging from the pupil is then projected through the x-y optical scanner in order to scan over the input object specified by a transparency of  $\Gamma_0(x, y)$ . The photodetector (PD) then accepts all the light to give out an electrical signal, which contains the processed information for the scanned object. If the scanned electrical signal is digitally stored (i.e., in a computer) in synchronization with the 2-D scan signals of the scanning mechanism (such as the x-y scanning mirrors), what is stored as a 2-D record is then a processed image of the scanned object.

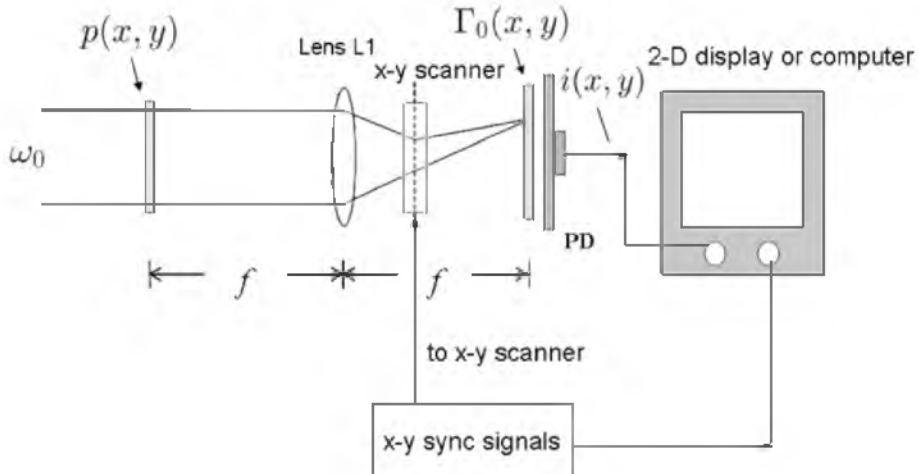


Fig. 3.1 An active optical scanning image processing system.

Let us now discuss photodetection and see how light information can be converted into electrical information. Assume that the photodetector's surface is on the  $z = 0$  plane and that the incident complex field on the detector's surface is given by  $\psi_p(x, y)\exp(j\omega_0 t)$  as shown in Fig. 3.2. Since the photodetector only responds to intensity, i.e.,  $|\psi_p\exp(j\omega_0 t)|^2$ , it gives the current,  $i$ , as an output by spatially integrating the intensity over the active area,  $D$ , of the detector:

$$i \propto \int_D |\psi_p \exp(j\omega_0 t)|^2 dx dy = \int_D |\psi_p|^2 dx dy. \quad (3.1-1)$$



For example, if the incident field is a plane wave of amplitude  $A$ , i.e.,  $\psi_p = A$ , the current output is given by

$$i \propto \int_D |A|^2 dx dy = A^2 D, \quad (3.1-2)$$

which is a constant. However, take for instance that if the light has been intensity-modulated, i.e.,  $|\psi_p|^2 = m(t)$ , where  $m(t)$  is the modulating signal, the current will then give an output that varies with the modulation. This is useful for laser communications systems [Pratt (1969)].

Note that since  $\psi_p(x, y) = |\psi_p(x, y)| \exp[j\phi(x, y)]$ , the output current can only contain the magnitude information, i.e.,  $|\psi_p|$ , and the phase information is completely lost. This type of photodetection is known as *optical direct detection* (or *optical incoherent detection*).

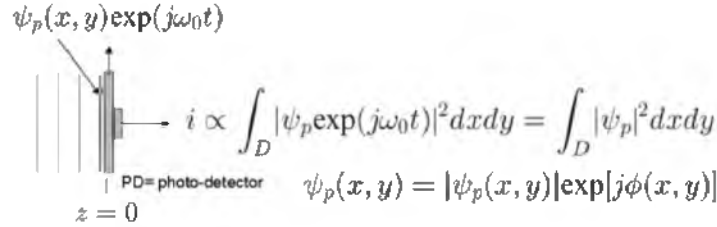


Figure 3.2 Optical direct detection.

Once we comprehend photodetection, we can return to Fig. 3.1 to calculate the current output given after scanning the transparency,  $\Gamma_0(x, y)$ . Instead of modeling the transparency that is being scanned by an optical beam, as shown by Fig. 3.3, we assume that the transparency,  $\Gamma_0$ , is moving through the optical beam. In Fig. 3.3, the plane of the photodetector is on the  $x' - y'$  plane and the optical scanning beam specified by a complex field,  $b(x', y')$ , is stationary at the origin of the  $x' - y'$  plane. By scanning or sampling we mean that successive points  $(x, y)$  in transparency coordinates) of  $\Gamma_0$  are brought into coincidence with the center  $(x', y' = 0)$  of the optical beam in the  $x' - y'$  plane.

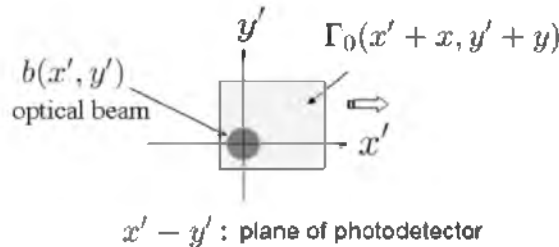


Fig. 3.3 Scanning situation.

In Fig. 3.3, the arguments,  $x$  and  $y$ , of  $\Gamma_0$  signify that the transparency is moving or translating with respect to the optical beam. Therefore, the total complex field reaching the photodetector is  $b(x', y')\Gamma_0(x' + x, y' + y)$ . The photodetector collects all the transmitted light and delivers a current,  $i$ . According to Eq. (3.1-1),  $i$  is given by

$$i(x, y) \propto \int_D |b(x', y') \Gamma_0(x' + x, y' + y)|^2 dx' dy', \quad (3.1-3)$$

where  $x = x(t)$  and  $y = y(t)$  represent the instantaneous position of the transparency. Alternatively, scanning imaging can be modeled by moving the optical beam across the transparency, which results in the following equation:

$$i(x, y) \propto \int_D |\Gamma_0(x', y') b(x' - x, y' - y)|^2 dx' dy'.$$

If we let  $x' - x = x''$  and  $y' - y = y''$  and then substitute them into the above equation, we have

$$i(x, y) \propto \int_D |\Gamma_0(x'' + x, y'' + y) b(x'', y'')|^2 dx'' dy'',$$

which is identical to Eq. (3.1-3). We shall use the formulation shown in Eq. (3.1-3) to represent optical scanning throughout the book. Note that for uniform scan speed  $V$ , we have  $x(t) = Vt$  and  $y(t) = Vt$ . When we rearrange Eq. (3.1-3), we have

$$\begin{aligned} i(x, y) &\propto \int_D |b(x', y')|^2 |\Gamma_0(x' + x, y' + y)|^2 dx' dy' \\ &= |b(x, y)|^2 \otimes |\Gamma_0(x, y)|^2. \end{aligned} \quad (3.1-4)$$

Note that this result is interesting because it is an incoherent optical system where only the intensities are processed, i.e.,  $|\Gamma_0(x, y)|^2$  is processed by  $|b(x, y)|^2$  even though the object,  $\Gamma_0(x, y)$ , originally may be complex in nature. Since the beam complex field,  $b(x, y)$ , and the pupil,  $p(x, y)$ , are in the back and the front focal plane of lens L1, respectively, as shown in Fig. 3.1, they are related by a Fourier transformation where [see Eq. (2.4-6)]

$$b(x, y) = \mathcal{F}_{xy}\{p(x, y)\} \Big|_{\substack{k_x = k_0 x/f \\ k_y = k_0 y/f}} \quad (3.1-5)$$

Figure 3.4 shows a commercially available x-y scanning system from General Scanning<sup>TM</sup>. The mirrors are driven by galvanometers. The figure on

the right is a close-up of the x-y scanning mirrors positioned orthogonally to each other (one direction for the x-scanning and the other for the y-scanning).

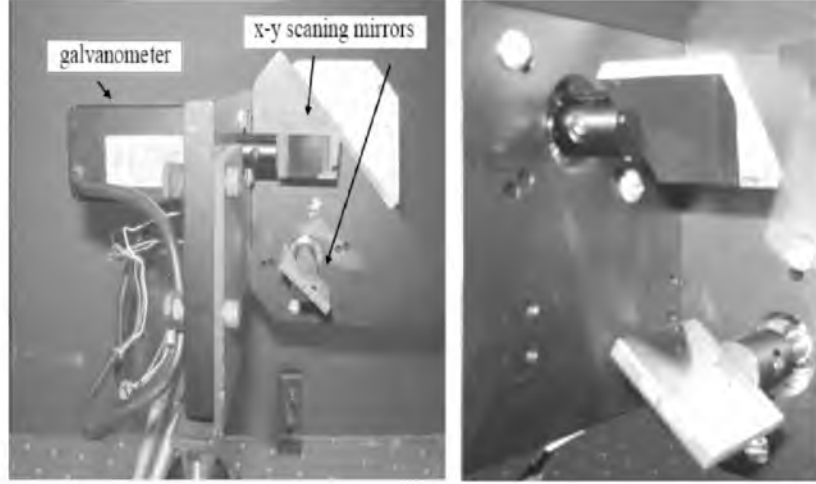


Fig. 3.4 x-y optical scanning system.

### 3.2 Optical Heterodyning

In the last section, we have shown that a simple optical scanning system that employs optical direct detection cannot extract any phase information of the incident complex field. While holography requires the preservation of the phase information, we, therefore, need to find a way to preserve the phase information during photodetection if we are expected to use optical scanning to record holographic information. The solution to this problem is *optical heterodyning*.

Figure 3.5 shows an optical heterodyne detection. The half-silvered mirror combines two mutually coherent laser beams; the information optical signal beam and the reference optical signal beam, having temporal frequencies of  $\omega_0$  and  $\omega_0 + \Omega$ , respectively [in the next section, we will show how laser beams of different temporal frequencies can be achieved by using acousto-optics]. For simplicity, we will consider the information signal and the reference signal, both as plane waves expressed by  $\psi_p \exp(j\omega_0 t)$  and  $B \exp[j(\omega_0 + \Omega)t]$  on the surface of the photodetector, respectively. Hence, the total field on the surface of the photodetector is  $\psi_t = \psi_p \exp(j\omega_0 t) + B \exp[j(\omega_0 + \Omega)t]$ . Since the photodetector only detects intensity, the current output is then given by

$$i \propto \int_D |\psi_t|^2 dx dy$$

$$\begin{aligned}
&= \int_D |\psi_p \exp(j\omega_0 t) + B \exp[j(\omega_0 + \Omega)t]|^2 dx dy \\
&= D[A^2 + B^2 + 2AB \cos(\Omega t - \phi)], \quad (3.2-1)
\end{aligned}$$

where we assume that the information signal is  $\psi_p = Ae^{j\phi}$ , which has an amplitude of  $A$  and phase information,  $\phi$ . Also, for simplicity, we assume that  $B$  is real in the above equation. The term,  $A^2 + B^2$ , is the DC current (or the *baseband current*), whereas the term  $AB \cos(\Omega t - \phi)$  is the AC current (or the *heterodyne current*) due to the *mixing* or *heterodyning* of the two optical signals at different frequencies [Poon (2002b), Poon and Kim (2006)]. Also note that the amplitude and the phase of the information signal both have been preserved in the current as it is clearly indicated in the last term of Eq. (3.2-1). Hence, optical heterodyning can preserve the amplitude and phase of the information signal. This type of photodetection is known as *optical heterodyne detection* (or *optical coherent detection*).

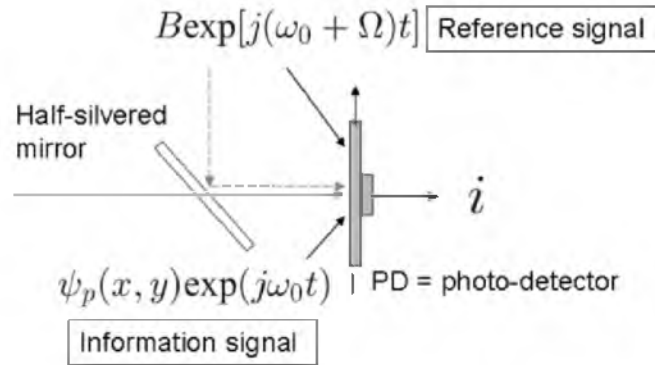


Fig. 3.5 Optical heterodyne detection.

Now that we have shown how the current,  $i$ , contains the amplitude and phase information through heterodyning, we will discuss how to extract this information electronically. Figure 3.6 shows an *electronic multiplexing detection*.

The current,  $i$ , first passes through a bandpass filter that is tuned to the heterodyne frequency,  $\Omega$ , in order to reject the baseband current and to extract the heterodyne current,  $i_\Omega \propto A \cos(\Omega t - \phi)$ . The heterodyne current splits into two channels to obtain two outputs,  $i_c$  and  $i_s$ , as shown in Fig. 3.6. Each channel actually performs *lock-in detection*, which consists of electronically multiplying the incoming signal with the cosine or sine of the heterodyne frequency, and then using lowpass filtering to extract the phase of the heterodyne current. Let us now see how it is mathematically performed.

First, consider the upper channel where the electronic multiplier gives

$$\begin{aligned} i_{\Omega} \times \cos(\Omega t) &= A \cos(\Omega t - \phi) \cos(\Omega t) \\ &= \frac{1}{2} A \cos(-\phi) + \frac{1}{2} A \cos(2\Omega t - \phi) \end{aligned} \quad (3.2-2)$$

as output, and where we have used the following trigonometric identity:

$$\cos \alpha \cos \beta = \frac{1}{2} \cos(\alpha - \beta) + \frac{1}{2} \cos(\alpha + \beta).$$

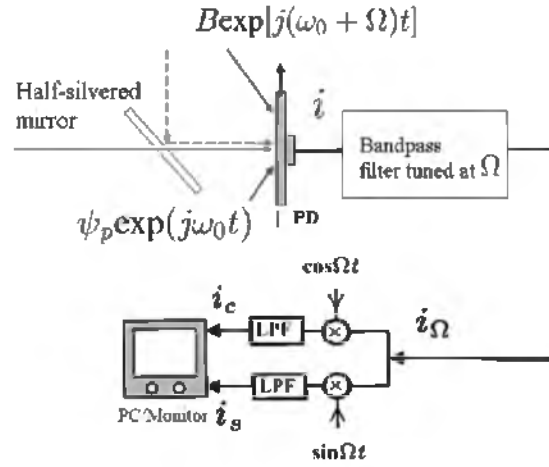


Fig. 3.6 Electronic multiplexing detection.

By using a lowpass filter on the output of the multiplier (which means we are rejecting the frequency of  $2\Omega$ ), we can obtain the *in-phase component* of the heterodyne current,  $i_{\Omega}$ , which is given by

$$i_c = A \cos(\phi). \quad (3.2-3a)$$

Apart from some constant, Eq. (3.2-3a) is really the first term of Eq. (3.2-2). Similarly, the lower channel of Fig. 3.6 gives the *quadrature component* of the heterodyne current,  $i_{\Omega}$ , which is given by

$$i_s = A \sin(\phi), \quad (3.2-3b)$$

where we can use the identity,

$$\cos \alpha \sin \beta = \frac{1}{2} \sin(\alpha + \beta) - \frac{1}{2} \sin(\alpha - \beta),$$

to obtain this result. Now, once  $i_c$  and  $i_s$  have been extracted and stored in a computer, we can perform the following complex addition:

$$i_c + j i_s = A \cos(\phi) + j A \sin(\phi) = A \exp(j\phi). \quad (3.2-4)$$

Note that this result is the full recovery of the information signal,  $\psi_p = A e^{j\phi}$ , from the photodetector's current given by Eq. (3.2-1). In fact, we will take advantage of optical heterodyning and electronic multiplexing detection by obtaining holographic information without the twin-image noise. We will return to this topic later.

### 3.3 Acousto-Optic Frequency Shifting

When we employ optical heterodyning as shown in Fig. 3.5, we need to create two laser beams of different temporal frequencies. In this section, we will discuss a common device used for shifting light frequency known as the *acousto-optic frequency shifter* (AOFS) or *acousto-optic modulator* (AOM) [Korpel (1981)].

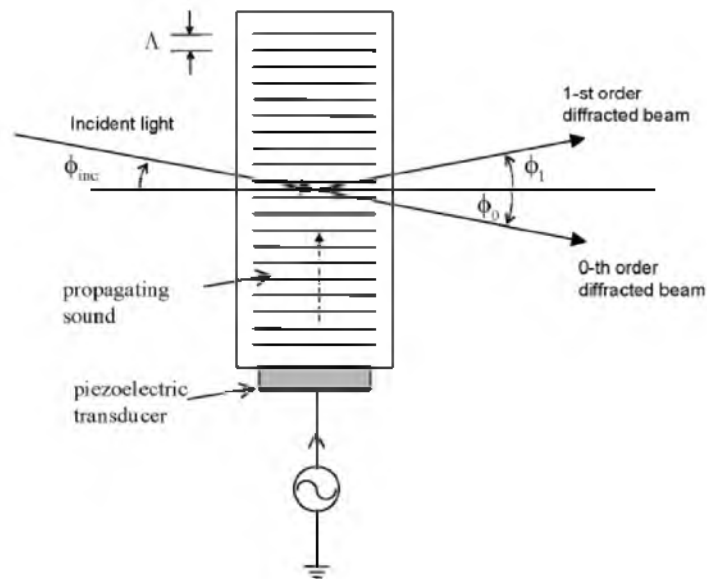


Fig. 3.7 Acousto-optic modulator.

An acousto-optic modulator is a spatial light modulator that consists of an acoustic medium, such as glass, that is bonded to a piezoelectric transducer. When an electrical signal is applied to the transducer, a sound wave propagates through the acoustic medium causing perturbations in the index of refraction, which is proportional to the electrical excitation. This in turn modulates the laser beam that traverses the acoustic medium. Thus, the

acousto-optic modulator, as shown in Figure 3.7, may be thought to act similar to *phase grating* with an effective grating line separation equal to the wavelength,  $\Lambda$ , of the sound in the acoustic medium. As it turns out, for a very specific incident angle,  $\phi_{\text{inc}}$ , the sound grating splits incident light into two diffracted beams, namely the 1-st order diffracted beam and the 0-th order diffracted beam at angles  $\phi_1$  and  $\phi_0$ , respectively. This is shown in Fig. 3.7. We shall identify these angles next.

One of the simplest explanations used to describe the interaction between sound and a laser is to treat the interaction as a collision of particles, namely *photons* and *phonons*. In order for these particles to have well-defined momenta and energies, we must assume that we have an interaction of plane waves of light and sound. In other words, we assume that the width of the transducer is sufficiently wide enough to produce plane wave fronts at a single frequency.

We will now consider two conservation laws that exist during the collision: the *conservation of energy* and the *conservation of momentum*. The condition for conservation of momentum is

$$\hbar \vec{k}_{+1} = \hbar \vec{k}_0 + \hbar \vec{K}, \quad (3.3-1)$$

where  $\hbar \vec{k}_0$  and  $\hbar \vec{K}$  are the momenta of the incident photon and phonon, respectively, and  $\hbar \vec{k}_{+1}$  is the momentum of the scattered photon.  $\vec{k}_{+1}$ ,  $\vec{k}_0$ , and  $\vec{K}$  are the corresponding wavevectors of the particles, and  $\hbar = h/2\pi$  where  $h$  is *Planck's constant*. Now, from the conservation of energy, we have

$$\hbar \omega_{+1} = \hbar \omega_0 + \hbar \Omega, \quad (3.3-2)$$

where  $\hbar \omega_{+1}$ ,  $\hbar \omega_0$ , and  $\hbar \Omega$  are the energies of the scattered photon, incident photon, and phonon, respectively.  $\omega_{+1}$ ,  $\omega_0$ , and  $\Omega$  are the corresponding radian frequencies of the particles.

After dividing Eq. (3.3-1) by  $\hbar$ , we have

$$\vec{k}_{+1} = \vec{k}_0 + \vec{K}. \quad (3.3-3)$$

Also, from Eq. (3.3-2), the corresponding conservation of energy takes the form

$$\omega_{+1} = \omega_0 + \Omega. \quad (3.3-4)$$

Figure 3.8a) shows the wave-vector interaction diagram and is constructed based on Eq. (3.3-3). For all practical cases,  $|\vec{K}| \ll |\vec{k}_0|$ , the magnitude of  $\vec{k}_{+1}$  is essentially equal to that of  $\vec{k}_0$ , and the wave-vector triangle shown in Fig. 3.8a) is, therefore, nearly isosceles. Note that the closed triangle in Figure 3.8a) stipulates that there are certain critical angles of incidence for the interaction of plane waves of light and sound. The stipulated incident angle,  $\phi_{\text{inc}}$ , is called the *Bragg angle*, which is given by

$$\sin(\phi_B) = \frac{K}{2k_0} = \frac{\lambda_0}{2\Lambda}, \quad (3.3-5)$$

where  $k_0 = |\vec{k}_0| = 2\pi/\lambda_0$  is the wavenumber of light inside the acoustic medium and  $\lambda_0$  is the wavelength of light.  $K = |\vec{K}| = 2\pi/\Lambda$  is the wavenumber of sound and  $\Lambda$  is the wavelength of sound. Note that the diffracted beams differ in direction by an angle equal to  $2\phi_B$  as shown in Fig. 3.8a).  $\phi_1$  and  $\phi_0$  in Fig. 3.7 must then be equal to  $\phi_B$ .

Figure 3.8b) shows that the 1-st diffracted beam is being up-shifted in frequency as is required by Eq. (3.3-4). The interaction described by Eq. (3.3-4) is called an *upshifted interaction* because the frequency of the diffracted beam,  $\omega_{+1}$ , has been upshifted by the amount of the sound frequency,  $\Omega$ . It is also clear that since we really do have a traveling sound wave, the frequency of the diffracted light is Doppler shifted.

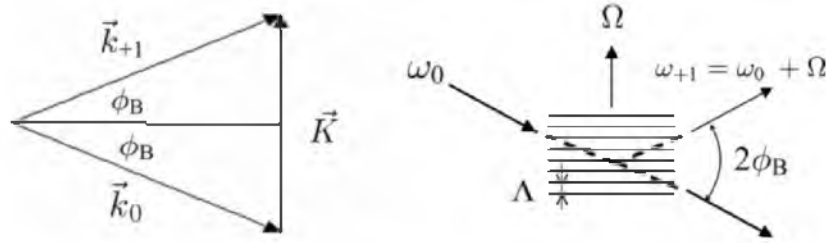


Fig. 3.8 Acousto-optic interaction:  
a) wave-vector diagram, b) experimental configuration.

The frequencies of sound waves produced in laboratories range from about 100 KHz to 3 GHz. Figure 3.9 shows a commercially available acousto-optic modulator, Model AOM-40, from IntraAction Corporation. It uses dense flint glass as an acoustic medium (refractive index  $n_0 \sim 1.65$ ) and operates at a sound center frequency of  $f_s = 40$  MHz. Therefore, in Fig. 3.9, the sound wave travels in the glass from the left to the right at a velocity of  $V_s \sim 4000$  m/s with a sound wavelength of  $\Lambda = V_s/f_s \sim 0.1$  mm. If a He-Ne laser is used (its wavelength is about  $0.6328 \mu\text{m}$  in air), its wavelength inside the glass is  $\lambda_0 \sim 0.6328 \mu\text{m}/n_0 \sim 0.3743 \mu\text{m}$ . Hence, according to Eq. (3.3-5), the Bragg angle, inside the acoustic medium is  $\sim 1.9 \times 10^{-3}$  radian or about 0.1 degrees. In Fig. 3.9, we have identified the two diffracted laser spots at the far background. The incident laser beam (not visible to the eye as it traverses across a transparent medium of glass) is traveling through the glass along the long dimension of the piezoelectric transducer.



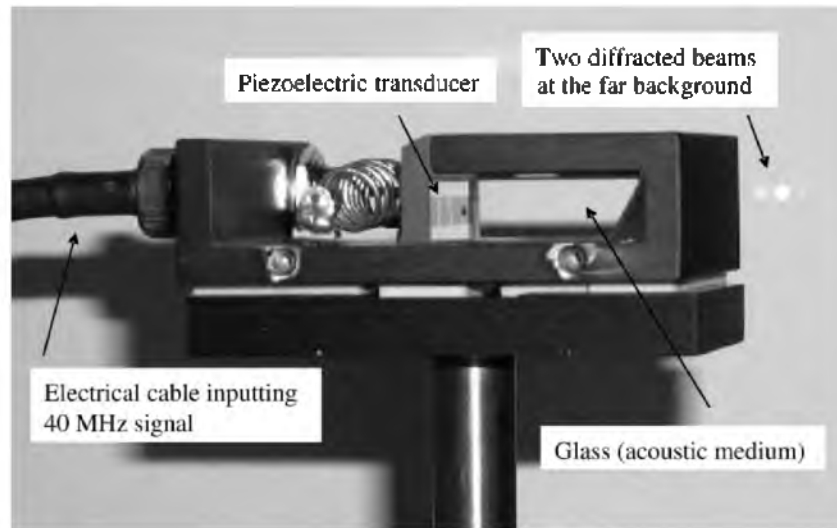


Fig. 3.9 Typical acousto-optic modulator operating at 40 MHz  
[Adapted from Poon (2002b)].

### 3.4 Two-Pupil Optical Heterodyne Scanning Image Processor

We have previously discussed optical scanning and optical heterodyning in prior sections. We have also discussed pupil function in an optical system, and we have shown that the pupil function can modify the characteristics of spatial filtering in an optical system [see Fig. 2.14]. In this section, we are in a position to discuss optical heterodyne scanning from which optical scanning holography is based. Since optical heterodyne scanning requires two optical beams to mix or heterodyne, we must, therefore, have two pupils in the optical system. One can envision that an optical system that has two pupils will have greater processing power because spatial filtering will now be controlled not only by a single pupil, as in conventional optical systems, but by two pupils. These systems are called *two-pupil systems* [Lohmann and Rhodes (1978), Poon and Korpel (1979)]. The article by Indebetouw and Poon provides a review of two-pupil approaches on incoherent image processing [1992].

Figure 3.10 shows a typical two-pupil optical heterodyne scanning image processor, which was originally developed and analyzed by Poon [1985]. We shall develop some mathematical descriptions of this system, which will eventually lead to the concept of optical scanning holography.

Beamsplitters BS and BS<sub>1</sub>, and mirrors M and M<sub>1</sub> form the Mach-Zehnder interferometer. The pupil,  $p_1(x, y)$ , is illuminated by a collimated laser at temporal frequency  $\omega_0$ . The other pupil,  $p_2(x, y)$ , is illuminated by the laser of temporal frequency  $\omega_0 + \Omega$ . The laser's temporal frequency offset by  $\Omega$  is introduced by an acousto-optic frequency shifter (AOFS) as shown in

the figure. Note that the figure shown in Fig. 3.10 is highly schematic because the details on selecting the first-order diffracted beam, which is the frequency-shifted beam at frequency  $\omega_0 + \Omega$ , emerging from the AOFS are not shown. The two pupils are located at the front focal planes of lens  $L_1$  and  $L_2$ , both with a focal length of  $f$ . The two pupils are then combined by the beamsplitter,  $BS_1$ , in order to focus the light onto the 2-D,  $x - y$  scanning mirrors, which are located on the back focal plane of lenses  $L_1$  and  $L_2$ . The combined optical beams are then used to 2-D raster scan over an object of amplitude distribution,  $\Gamma_0(x, y; z)$ , which is located at a distance of  $z$  away from the focal plane of the two lenses. Lens  $L_3$  is used to collect all the transmitted light (or scattered light if the object is diffusely reflecting) onto the photodetector (PD), which gives  $i(x, y)$  as its current output. An electronic bandpass filter (BPF) tuned at the heterodyne frequency of  $\Omega$  provides an output of a scanned and processed current  $i_\Omega(x, y)$ . We shall further develop the mathematical expression of  $i_\Omega(x, y)$ .

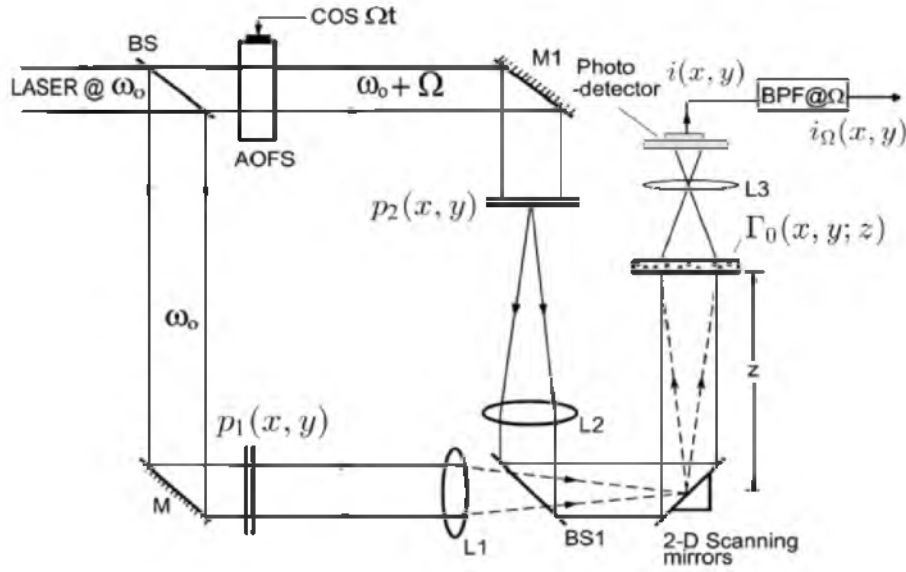


Fig. 3.10 A typical two-pupil optical heterodyne scanning system.

The combined optical scanning complex field,  $S(x, y; z)$ , located at a distance of  $z$  away from the focal plane of the two lenses, is given by

$$S(x, y; z) = P_{1z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \exp(j\omega_0 t) + P_{2z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \exp[j(\omega_0 + \Omega)t] , \quad (3.4-1)$$

where  $P_{iz}(\frac{k_0x}{f}, \frac{k_0y}{f})$  is the field distribution  $z$  away from the scanning mirrors and through Fresnel diffraction, is given by

$$P_{iz}(\frac{k_0x}{f}, \frac{k_0y}{f}) = P_i(\frac{k_0x}{f}, \frac{k_0y}{f}) * h(x, y; z), i = 1, 2. \quad (3.4-2)$$

In Eq. (3.4-2),  $P_i(\frac{k_0x}{f}, \frac{k_0y}{f})$  is the field distribution in the back focal plane of lenses  $L_1$  and  $L_2$ , and apart from some inessential constant and a constant phase factor, is given by [see Eq. (2.4-6)]

$$P_i(\frac{k_0x}{f}, \frac{k_0y}{f}) = \mathcal{F}\{p_i(x, y)\} \Big|_{\substack{k_x=k_0x/f \\ k_y=k_0y/f}}. \quad (3.4-3)$$

Now, as previously mentioned, the combined optical field or the scanning pattern, given by Eq. (3.4-1), is used to two-dimensionally scan an object with an amplitude transparency of  $\Gamma_0(x, y, z)$  located at a distance  $z$  from the scanning mirrors. According to the principle established by Eq. (3.1-3) for optical scanning, the photodetector, which responds to the incident intensity of the optical transmitted field or scattered field, generates a current given by

$$\begin{aligned} i(x, y, z) &\propto \int_D |S(x', y'; z) \Gamma_0(x' + x, y' + y; z)|^2 dx' dy' \\ &= \int_D \left| \left[ P_{1z}(\frac{k_0x'}{f}, \frac{k_0y'}{f}) \exp(j\omega_0 t) + P_{2z}(\frac{k_0x'}{f}, \frac{k_0y'}{f}) \exp[j(\omega_0 + \Omega)t] \right] \right. \\ &\quad \left. \times \Gamma_0(x + x', y + y'; z) \right|^2 dx' dy'. \end{aligned} \quad (3.4-4)$$

After a bandpass filter (BPF) is tuned to a frequency of  $\Omega$ , the heterodyne current from Eq. (3.4-4) becomes

$$\begin{aligned} i_\Omega(x, y, z) &= \text{Re} \left[ \int_D P_{1z}^*(\frac{k_0x'}{f}, \frac{k_0y'}{f}) P_{2z}(\frac{k_0x'}{f}, \frac{k_0y'}{f}) \right. \\ &\quad \left. \times |\Gamma_0(x + x', y + y'; z)|^2 dx' dy' \exp(j\Omega t) \right], \end{aligned} \quad (3.4-5)$$

where we have adopted the convention for the phasor  $\psi_p$  as  $\psi(x, y, t) = \text{Re}[\psi_p(x, y, t) \exp(j\Omega t)]$ , where  $\text{Re}[\cdot]$  denotes the real part of the content inside the bracket. Equation (3.4-5) can be written as

$$i_{\Omega}(x, y; z) = \text{Re}[i_{\Omega_p}(x, y; z)\exp(j\Omega t)], \quad (3.4-6a)$$

where

$$i_{\Omega_p}(x, y; z) = \iint_D P_{1z}^*\left(\frac{k_0 x'}{f}, \frac{k_0 y'}{f}\right) P_{2z}\left(\frac{k_0 x'}{f}, \frac{k_0 y'}{f}\right) \times |\Gamma_0(x + x', y + y'; z)|^2 dx' dy' \quad (3.4-6b)$$

is the output phasor containing the amplitude and the phase information of the heterodyne current. The amplitude and the phase information of the current constitute the scanned and the processed version of the object  $|\Gamma_0|^2$  and from Eq. (3.4-6), we can write

$$i_{\Omega}(x, y; z) = |i_{\Omega_p}(x, y; z)| \cos[(\Omega t + \phi_p(x, y; z))],$$

where  $i_{\Omega_p} = |i_{\Omega_p}| \exp(j\phi_p)$ . Note that we can re-write Eq. (3.4-6b) in terms of the following correlation:

$$i_{\Omega_p}(x, y; z) = P_{1z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) P_{2z}^*\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \otimes |\Gamma_0(x, y; z)|^2. \quad (3.4-7)$$

Similar to conventional optical scanning systems (or incoherent optical systems), only the intensity distribution, i.e.,  $|\Gamma_0|^2$  will be processed and the optical system is therefore incoherent. However,  $|\Gamma_0|^2$  is not strictly processed by an intensity quantity, and as indicated by Eq. (3.4-7) the processing element,  $P_{1z}P_{2z}^*$ , can be bipolar or even complex, thereby leading to the concept of *complex incoherent image processing*.

Equation (3.4-7) relates the input quantity to the output quantity and from this we can now define the optical transfer function (OTF) of the system to be

$$OTF_{\Omega}(k_x, k_y; z) = \mathcal{F}\{i_{\Omega_p}(x, y; z)\} / \mathcal{F}\{|\Gamma_0(x, y; z)|^2\}. \quad (3.4-8)$$

By taking the Fourier transform of Eq. (3.4-7) and combining its result with Eq. (3.4-8), we obtain the equivalent,

$$OTF_{\Omega}(k_x, k_y; z) = \mathcal{F}^*\{P_{1z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) P_{2z}^*\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right)\}. \quad (3.4-9)$$

In terms of the pupils  $p_1$  and  $p_2$ , we substitute Eqs. (3.4-2) and (3.4-3) into Eq. (3.4-9) to get

$$\begin{aligned}
OTF_{\Omega}(k_x, k_y; z) &= \exp\left[j\frac{z}{2k_0}(k_x^2 + k_y^2)\right] \\
&\times \iint p_1^*(x', y') p_2\left(x' + \frac{f}{k_0}k_x, y' + \frac{f}{k_0}k_y\right) \exp\left[j\frac{z}{f}(x'k_x + y'k_y)\right] dx' dy'.
\end{aligned}
\tag{3.4-10}$$

This equation was first derived by Poon [1985], and it states that the optical transfer function of the system,  $OTF_{\Omega}$ , can be modified based on the selection of the two pupils. Now, by using Eq. (3.4-8) and by re-writing Eq. (3.4-6a) in terms of  $OTF_{\Omega}$ , we obtain

$$\begin{aligned}
i_{\Omega}(x, y; z) &= \text{Re}[i_{\Omega_p}(x, y; z) \exp(j\Omega t)] \\
&= \text{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z)|^2\} OTF_{\Omega}(k_x, k_y; z)\} \exp(j\Omega t)].
\end{aligned}
\tag{3.4-11}$$

By defining the spatial impulse response (or the point spread function) of the optical heterodyne scanning system as

$$h_{\Omega}(x, y; z) = \mathcal{F}^{-1}\{OTF_{\Omega}\}, \tag{3.4-12}$$

we can now re-write Eq. (3.4-11) in the spatial domain as

$$i_{\Omega}(x, y; z) = \text{Re}\left\{[|\Gamma_0(x, y; z)|^2 * h_{\Omega}(x, y; z)] \exp(j\Omega t)\right\}. \tag{3.4-13}$$

Equation (3.4-11) or (3.4-13) represents the scanned and processed output current, which is modulated by a temporal carrier at a frequency of  $\Omega$ . By mixing  $i_{\Omega}$  with  $\cos(\Omega t)$  or  $\sin(\Omega t)$ , we can demodulate and extract the in-phase component or the quadrature component, respectively. The demodulation system is shown in Fig. 3.6, and the two outputs are given by

$$\begin{aligned}
i_c(x, y; z) &= \text{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0|^2\} OTF_{\Omega}\}] \quad (\text{frequency domain}) \\
&= \text{Re}[|\Gamma_0|^2 * h_{\Omega}(x, y; z)] \quad (\text{spatial domain})
\end{aligned}
\tag{3.4-14a}$$

and

$$\begin{aligned}
i_s(x, y; z) &= \text{Im}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0|^2\} OTF_{\Omega}\}] \quad (\text{frequency domain}) \\
&= \text{Im}[|\Gamma_0|^2 * h_{\Omega}(x, y; z)], \quad (\text{spatial domain})
\end{aligned}
\tag{3.4-14b}$$

where  $\text{Im}[\cdot]$  denotes the imaginary part of the quantity within the bracket. The subscripts “c” and “s” represent the use of  $\cos(\Omega t)$  and  $\sin(\Omega t)$  respectively, to extract the information from  $i_{\Omega}$ .

In Eqs. (3.4-14), we have assumed that the input object,  $|\Gamma_0(x, y; z)|^2$ , is an infinitely thin 2-D object located at a distance of  $z$  away from the 2-D scanning mirrors shown in Fig. 3.10. To generalize Eqs. (3.4-14) for 3-D objects, we need to integrate the equations over the depth, i.e., over  $z$ , of the 3-D objects. Eqs. (3.4-14) then become [Poon and Kim (1999)]

$$i_c(x, y) = \text{Re} \left[ \int \mathcal{F}^{-1} \{ \mathcal{F} \{ |\Gamma_0(x, y; z)|^2 \} OTF_\Omega \} dz \right] \quad (3.4-15a)$$

$$= \text{Re} \left[ \int |\Gamma_0(x, y; z)|^2 * h_\Omega(x, y; z) dz \right] \quad (3.4-15b)$$

and

$$i_s(x, y) = \text{Im} \left[ \int \mathcal{F}^{-1} \{ \mathcal{F} \{ |\Gamma_0(x, y; z)|^2 \} OTF_\Omega \} dz \right] \quad (3.4-15c)$$

$$= \text{Im} \left[ \int |\Gamma_0(x, y; z)|^2 * h_\Omega(x, y; z) dz \right]. \quad (3.4-15d)$$

Note that we have left the  $z$ -dependence out on the left-hand side of Eqs. (3.4-15) to emphasize that the recorded information is strictly 2-D even for 3-D objects.  $i_c(x, y)$  and  $i_s(x, y)$  represent the scanned and processed current (or information) of  $|\Gamma_0|^2$  and can be stored as 2-D records if these currents are stored in synchronization with the signals used to drive the  $x - y$  scanning mirrors. Equations (3.4-15) represent the major results of the two-pupil optical heterodyne scanning of a 3-D object. Figure 3.11 shows the overall two-pupil optical heterodyne image processor. The 3-D object is  $|\Gamma_0(x, y; z)|^2$ , and the final outputs given by Eqs. (3.4-15) are  $i_c(x, y)$  and  $i_s(x, y)$ . Note that while the input object is given by the amplitude distribution,  $\Gamma_0(x, y; z)$ , the information that can be processed is the intensity distribution given by  $|\Gamma_0(x, y; z)|^2$ . As it turns out, this is the incoherent mode of operation for the processor, which has so far been inclusively used for various applications such as 3-D fluorescence holographic microscopy, 3-D pattern recognition, optical remote sensing, and 3-D cryptography. In chapter 4, we will further elaborate on some of these applications. In chapter 5, when we consider the advancements towards optical scanning holography, we will describe a coherent mode where the complex distribution of the object can be processed [Indebetouw, Klysubun, Kim, and Poon (2000)]. This can be important when we deal with the phase specimens in biological applications.

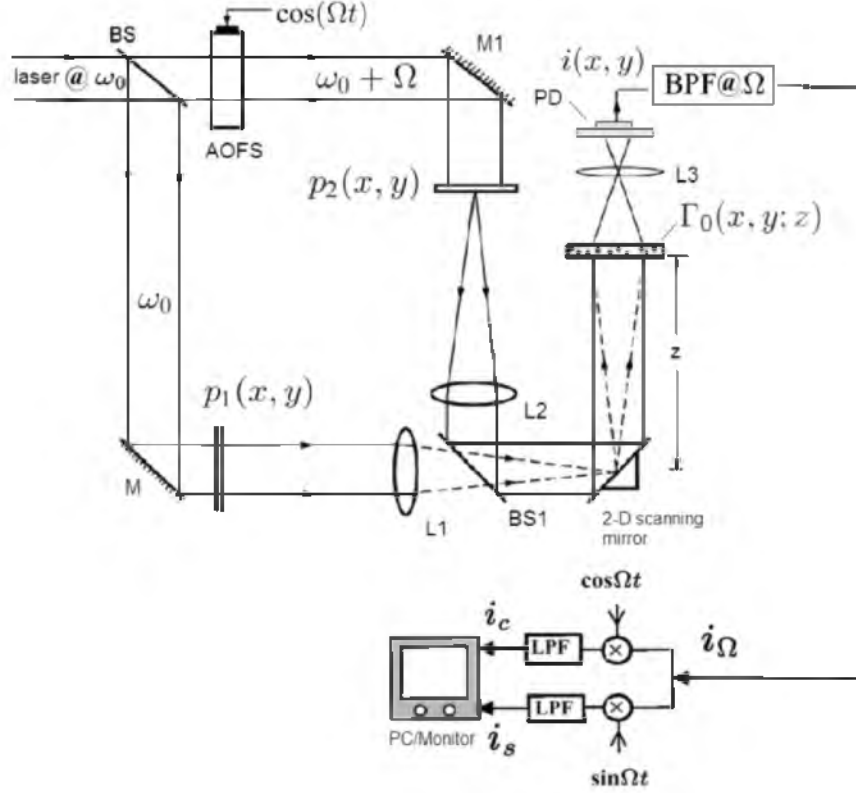


Fig. 3.11 The complete two-pupil optical scanning image processor.

### 3.5 Scanning Holography

In this section, we will discuss how holographic recording can be accomplished by using the two-pupil optical heterodyne scanning image processor discussed in the last section. The idea was first implicated by Poon and Korpel [1979]. They realized that interesting OTF's can be obtained by drastically modifying one of the pupils in relation to the other. In this context, it is intriguing to realize that there exists a possibility of creating a Fresnel-zone-plate-type impulse response (i.e., its phase is a quadratic function of  $x$  and  $y$ ) in an out-of-focus plane near the focal plane of lenses L1 and L2, i.e.,  $z$  away from the scanning mirrors shown in Fig. 3.11, by making  $p_1(x, y)$  uniform and  $p_2(x, y)$  a delta function. When investigating the chirp property, we can determine how far  $z$  is from the scanning mirrors, and thus, it carries obvious implication to holographic recording. The original idea, which was later analyzed and called *scanning holography* [Poon (1985)], is to scan the 3-D object in a 2-D raster with a complex

Fresnel-zone-plate-type impulse response created by the interference of a point source and a plane wave emerging from each pupil. A temporal frequency offset is introduced between the two pupils, and the desired signal from a spatially integrating detector is obtained using a heterodyne detection.

Hence, for scanning holography we mathematically let  $p_1(x, y) = 1$  and  $p_2(x, y) = \delta(x, y)$ , which are both clearly pictured in Fig. 3.11. With this choice of pupils, according to Eq. (3.4-10), the OTF of the heterodyne scanning system becomes

$$\begin{aligned} OTF_{\Omega}(k_x, k_y; z) \Big|_{\text{osh}} &= \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right] \\ &= OTF_{\text{osh}}(k_x, k_y; z), \end{aligned} \quad (3.5-1a)$$

and according to Eq. (3.4-12), the corresponding spatial impulse response is

$$h_{\Omega}(x, y; z) \Big|_{\text{osh}} = \frac{-jk_0}{2\pi z} \exp\left[\frac{jk_0(x^2 + y^2)}{2z}\right]. \quad (3.5-1b)$$

Apart from the constant phase factor, it is interesting to point out that by comparing the spatial frequency transfer function in Fourier optics [see Eq. (2.3-13)] to Eq. (3.5-1a), we have

$$OTF_{\text{osh}}(k_x, k_y; z) = H^*(k_x, k_y; z), \quad (3.5-2a)$$

and similarly in reference to Eq. (2.3-11), we have

$$h_{\Omega}(x, y; z) \Big|_{\text{osh}} = h^*(x, y; z). \quad (3.5-2b)$$

From the result of Eq. (3.5-1b) for scanning holography, apart from some constant, the spatial domain equations, Eqs. (3.4-15b) and (3.4-15d), become

$$\begin{aligned} i_c(x, y) &= \int \left\{ |\Gamma_0(x, y; z)|^2 * \frac{k_0}{2\pi z} \sin\left[\frac{k_0}{2z}(x^2 + y^2)\right] \right\} dz \\ &= H_{\sin}(x, y) \end{aligned} \quad (3.5-3a)$$

and

$$\begin{aligned} i_s(x, y) &= \int \left\{ |\Gamma_0(x, y; z)|^2 * \frac{k_0}{2\pi z} \cos\left[\frac{k_0}{2z}(x^2 + y^2)\right] \right\} dz \\ &= H_{\cos}(x, y), \end{aligned} \quad (3.5-3b)$$



respectively. What is being two-dimensionally recorded is a hologram.  $H_{\sin}(x, y)$  is called the *sine-FZP hologram*, and  $H_{\cos}(x, y)$  is the *cosine-FZP hologram* of  $|\Gamma_0(x, y; z)|^2$ .

To see why Eqs. (3.5-3) correspond to holographic recordings, we will let  $|\Gamma_0(x, y; z)|^2 = \delta(x, y)\delta(z - z_0)$ , which is a point source located  $z_0$  away from the scanning mirrors. Then Eq. (3.5-3a) becomes

$$\begin{aligned} H_{\sin}(x, y) &= \int \left\{ \delta(x, y)\delta(z - z_0) * \frac{k_0}{2\pi z} \sin\left[\frac{k_0}{2z}(x^2 + y^2)\right] \right\} dz \\ &= \int \left\{ \delta(z - z_0) \frac{k_0}{2\pi z} \sin\left[\frac{k_0}{2z}(x^2 + y^2)\right] \right\} dz \end{aligned}$$

after the 2-D convolution involving  $x$  and  $y$ . And finally, after the integration along  $z$ , the above equation becomes

$$H_{\sin}(x, y) = \frac{k_0}{2\pi z_0} \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right]. \quad (3.5-4a)$$

Note that this is basically the hologram of a point source without the constant bias of  $A$ , which appears in Eq. (2.5-4). The constant bias simply gives a zero-order beam upon optical reconstruction. Similarly, Eq. (3.5-3b) gives

$$H_{\cos}(x, y) = \frac{k_0}{2\pi z_0} \cos\left[\frac{k_0}{2z_0}(x^2 + y^2)\right]. \quad (3.5-4b)$$

In summary, in scanning holography for a single 2-D raster-scan we have two records of the holograms, due to electronic multiplexing detection. Both of the holograms given by Eq. (3.5-3) contain holographic information, but they are not redundant as we will later see that with the two holograms, we can obtain a *twin-image-free* hologram even the recording is made on-axis.

Figure 3.12 shows the very first hologram using scanning holography [Duncan and Poon (1992)]. The hologram is a slit with the size of 50  $\mu\text{m}$ . The term “*optical scanning holography*” was first coined in this article to emphasize that this was the first electronic hologram created by using the active optical scanning technique. To put optical scanning holography into perspective, holograms obtained by scanning techniques at long wavelengths have long been achieved. This has been possible because there is no need to supply a physical reference beam in order to extract holographic information because detectors that are capable of measuring the oscillation of low-frequency radiation (such as acoustic waves or microwaves) are commonly available, permitting amplitude and phase information to be directly extracted from long wavelength signals.

In general, optical scanning holography (OSH) can be applied to other shorter and longer wavelength systems as long as we can find devices

of that particular wavelength that can generate a collimated beam and a focused beam so that the two beams interfere on the object. In addition, a frequency shifter for that wavelength must be available. In a more futuristic vision, *CO<sub>2</sub> scanning holography* for active optical remote sensing can be possible as  $10.6\text{ }\mu\text{m}$  can penetrate the atmosphere with little absorption. At the other end of the light spectrum, *X-ray scanning holography* is becoming a reality because of the increasing existence of X-ray lasers, which should be important if atomic resolution for 3-D specimens is required. In the remainder of the book, we shall use the term *optical scanning holography* or OSH instead of scanning holography.

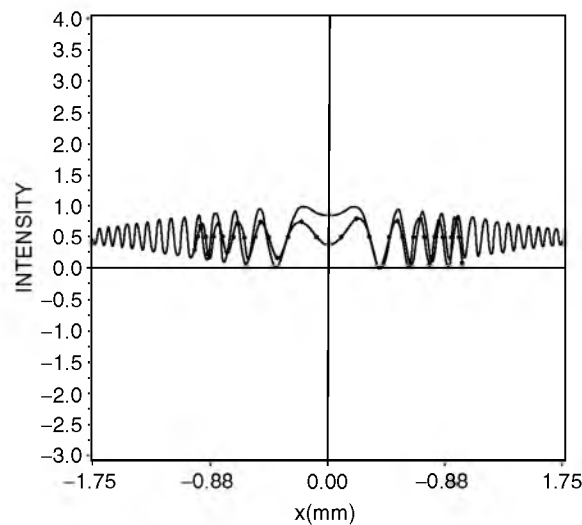


Fig. 3.12 The first hologram obtained using optical scanning holography: solid line: theoretical results; dotted line: experimental results. The object is a  $50\text{ }\mu\text{m}$ -slit. Reprinted from Duncan and Poon, *JOSA A* 9, 229 (1992), with permission. © OSA.

The extension of the capabilities of one-dimensional to two-dimensional imaging by using optical scanning holography was subsequently demonstrated by Poon, Doh, Schilling, Wu, Shinoda, and Suzuki [1995]. Fig. 3.13 shows a hologram of a pinhole object. A sine-hologram of the pinhole is shown (this is the well-known FZP). The pinhole is about  $50\text{ }\mu\text{m}$  in diameter and approximately 10 cm away from the 2-D scanning mirrors. The plane wave on the pinhole is about 10mm, and comes from a collimated HeNe laser. The spherical wave on the pinhole comes from a focused laser beam of a size of about  $3.5\text{ }\mu\text{m}$ . The temporal frequency difference between the plane wave and the spherical wave is 40 MHz. In the same paper mentioned, the authors reported the first 3-D imaging capability using optical scanning holography. Three-dimensional imaging was demonstrated by digitally reconstructing an acquired hologram to show various depths of an image.

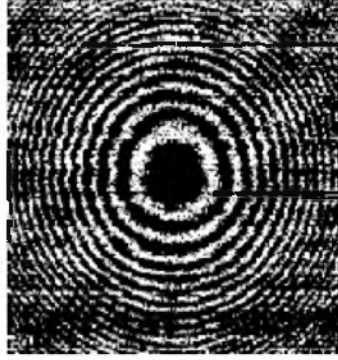


Fig. 3.13 Sine-hologram of a pinhole. The first 3-D imaging capability using OSH is also reported in the paper. Reprinted from T.-C. Poon et al., *Optical Engineering* 34, 1338, with permission. © 1995 SPIE.

We have discussed the twin-image problem in on-axis holography. And off-axis (or carrier-frequency) holography has been employed to avoid the annoying effect of the twin-image. One of the most popular electronic holographic techniques used to obtain reconstruction free of the twin image is called *phase-shifting holography* [Yamaguchi and Zhang (1997)]. This technique employs phase-shifting on the reference beam to obtain four on-axis holograms in order to calculate the phase of the complex object wave. Using optical scanning holography, we only need to perform a single 2-D scan in order to simultaneously obtain two on-axis holograms - namely the sine-hologram and the cosine-hologram. Since the two holograms can be stored digitally, we can perform a complex addition or subtraction as follows:

$$H_{c\pm}(x, y) = H_{\cos}(x, y) \pm jH_{\sin}(x, y)$$

$$= \int \left\{ |\Gamma_0(x, y; z)|^2 * \frac{k_0}{2\pi z} \exp[\pm j \frac{k_0}{2z}(x^2 + y^2)] \right\} dz, \quad (3.5-5)$$

where we have used Eqs. (3.5-3a) and (3.5-3b).  $H_{c\pm}(x, y)$  is called a *complex Fresnel zone plate hologram*, which contains no twin-image information [Doh, Poon, Wu, Shinoda, and Suzuki (1996)]. To better understand this, we will construct a complex hologram for a point object. Substituting Eqs. (3.5-4a) and (3.5-4b) into Eq. (3.5-5), and apart from some constant, we have

$$H_{c\pm}(x, y) = \exp[\pm j \frac{k_0}{2z_0}(x^2 + y^2)]. \quad (3.5-6)$$

According to Eqs. (2.5-6b) and (2.5-6c), this hologram can construct either a real point source or a virtual point source, depending on the sign of the argument being chosen in Eq. (3.5-6). For the positive sign in the argument, we will have a real image reconstruction, and for the negative sign, we will have a virtual image reconstruction. In either case, there is no twin-image formation within the complex hologram.

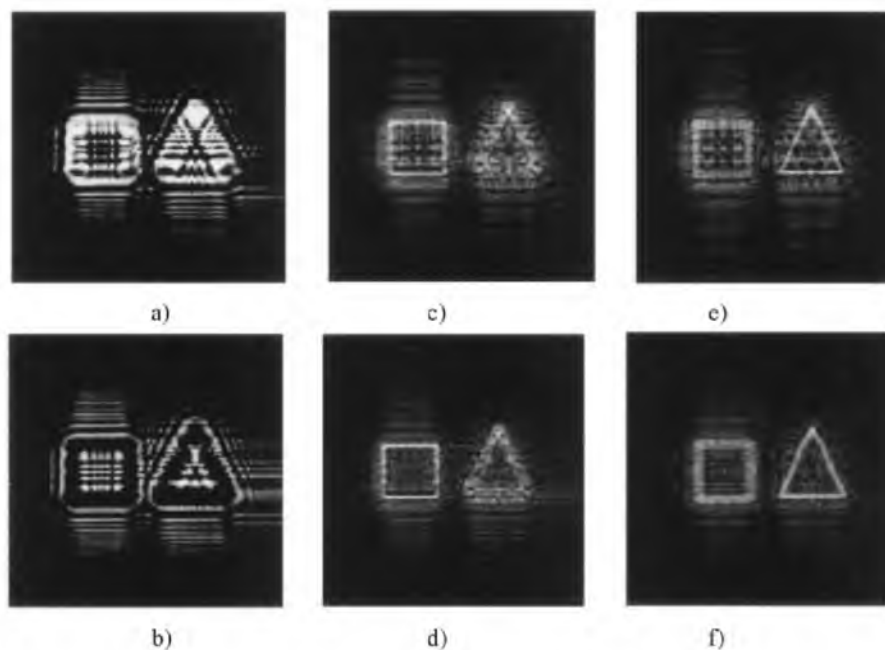


Fig. 3.14 Twin-image elimination with two holograms: a) Cosine-hologram. b) Sine-hologram. c) Reconstruction of a), the cosine-hologram, focused on the square (twin-image noise noticeable). d) Reconstruction of the complex hologram, focused on the square (no twin-image noise). e) Reconstruction of a), focused on the triangle (twin-image noise noticeable). f) Reconstruction of the complex hologram, focused on the triangle (no twin-image noise). Reprinted from Poon et al., *Optics Letters* 5, 215, (2000), with permission. © OSA.

Figure 3.14 shows the experimental results of eliminating the twin image for three-dimensional images in optical scanning holography. In the experiment, the 3D object consisted of two 35-mm slides that were set up side-by-side and at different distances from the scanning beam. One slide is a square, and the other is a triangle. The depth difference between the two slides is 15-cm. Figure 3.14a) and b) show the cosine- and sine-holograms, respectively. Figure 3.14c) shows the reconstruction of the cosine-hologram. This

reconstruction is focused on the square object but is spoiled by twin-image noise. In Fig. 3.14d), where the complex hologram is used to reconstruct and focus on the square object, twin-image noise is not present. In Fig. 3.14e), the cosine hologram reconstructs an image that is focused on the triangle but spoiled by twin-image noise. In fig. 3.14f), the complex hologram gives a reconstruction without twin-image noise. Thus, the elimination of twin-image in optical scanning holography has been demonstrated using only two holograms. In the demonstration, the use of digital reconstruction has been employed. By using two SLM's (one for the display of a sine-hologram and the other for a cosine-hologram), we can achieve all-optical reconstruction. However, this has yet to be demonstrated optically [Poon (2006)].

### Example 3.1 MATLAB Example: Optical Scanning Holography

By substituting the OTF of optical scanning holography, given by Eq. (3.5-1a), into Eqs. (3.4-15a) and (3.4-15c), the sine-hologram and the cosine-hologram will be expressed in terms of spatial frequencies. Therefore, we have

$$\begin{aligned} i_c(x, y) &= \text{Re} \left[ \int \mathcal{F}^{-1} \{ \mathcal{F} \{ |\Gamma_0(x, y; z)|^2 \} OTF_{\text{osh}}(k_x, k_y; z) \} dz \right] \\ &= H_{\sin}(x, y) \end{aligned} \quad (3.5-7a)$$

and

$$\begin{aligned} i_s(x, y) &= \text{Im} \left[ \int \mathcal{F}^{-1} \{ \mathcal{F} \{ |\Gamma_0(x, y; z)|^2 \} OTF_{\text{osh}}(k_x, k_y; z) \} dz \right] \\ &= H_{\cos}(x, y), \end{aligned} \quad (3.5-7b)$$

where  $OTF_{\text{osh}}(k_x, k_y; z) = \exp[-j\frac{z}{2k_0}(k_x^2 + k_y^2)]$ . For this example, we are assuming a planar object to be at a distance of  $z_0$  away from the x-y scanning mirrors, i.e.,  $|\Gamma_0(x, y; z)|^2 = I(x, y)\delta(z - z_0)$ , where  $I(x, y)$  is the planar intensity distribution shown in Fig. 3.15a). For the planar intensity object, after integrating over  $z$ , Eqs. (3.5-7a) and (3.5-7b) become

$$H_{\sin}(x, y) = \text{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{I(x, y)\}OTF_{\text{osh}}(k_x, k_y; z_0)\}] \quad (3.5-8a)$$

and

$$H_{\cos}(x, y) = \text{Im}[\mathcal{F}^{-1}\{\mathcal{F}\{I(x, y)\}OTF_{\text{osh}}(k_x, k_y; z_0)\}], \quad (3.5-8b)$$

respectively. The above holograms are simulated and shown in Fig. 3.15b) and Fig. 3.15c), respectively where  $\sigma = z_0/2k_0 = 2.0$  in OSH.m listed in Table 3.1. We can also construct a complex FZP hologram by using Eq. (3.5-5):

$$\begin{aligned} H_{c+}(x, y) &= H_{\cos}(x, y) + jH_{\sin}(x, y) \\ &= \mathcal{F}^{-1}\{\mathcal{F}\{I(x, y)\}OTF_{\text{osh}}(k_x, k_y; z_0)\}. \end{aligned} \quad (3.5-9)$$

For digital reconstruction, we will simply convolve the above holograms with the spatial impulse response in order to simulate Fresnel diffraction for a distance of  $z_0$ . To obtain real image reconstruction formed in front of the hologram, we will use the following equation:

$$H_{\text{any}}(x, y) * h(x, y; z_0),$$

where  $H_{\text{any}}(x, y)$  represents any one of the above holograms, i.e., the sine-hologram, the cosine-hologram or the complex hologram. In OSH.m, the above equation is implemented in the Fourier domain using the following equation [see Eqs. (1.2-3a) and (1.2-3b)]:

Reconstructed real image

$$\begin{aligned} &\propto \mathcal{F}^{-1}\{\mathcal{F}\{H_{\text{any}}(x, y)\}H(k_x, k_y; z_0)\} \\ &= \mathcal{F}^{-1}\{\mathcal{F}\{H_{\text{any}}(x, y)\}OTF_{\text{osh}}^*(k_x, k_y; z_0)\}, \end{aligned} \quad (3.5-10)$$

where we have used Eq. (3.5-2a) to relate  $OTF_{\text{osh}}$  with the spatial frequency response,  $H(k_x, k_y; z_0)$ , to obtain the last step. Figures 3.15d), e) and f) show the reconstruction of the sine-hologram, the cosine-hologram, and the complex hologram, respectively.

Note that if the complex hologram is constructed as

$$H_{c-}(x, y) = H_{\cos}(x, y) - jH_{\sin}(x, y),$$

then it will have a reconstructed virtual image that is located at a distance of  $z_0$  behind the hologram. However, if we perform

$$H_{c-}(x, y) * h(x, y; z_0)$$

for reconstruction, we will see a severely out-of-focus image formed at  $z = z_0$  as shown in Fig. (3.15g), which is equivalent to the Fresnel diffraction pattern of the original object at a distance of  $z = 2z_0$ .

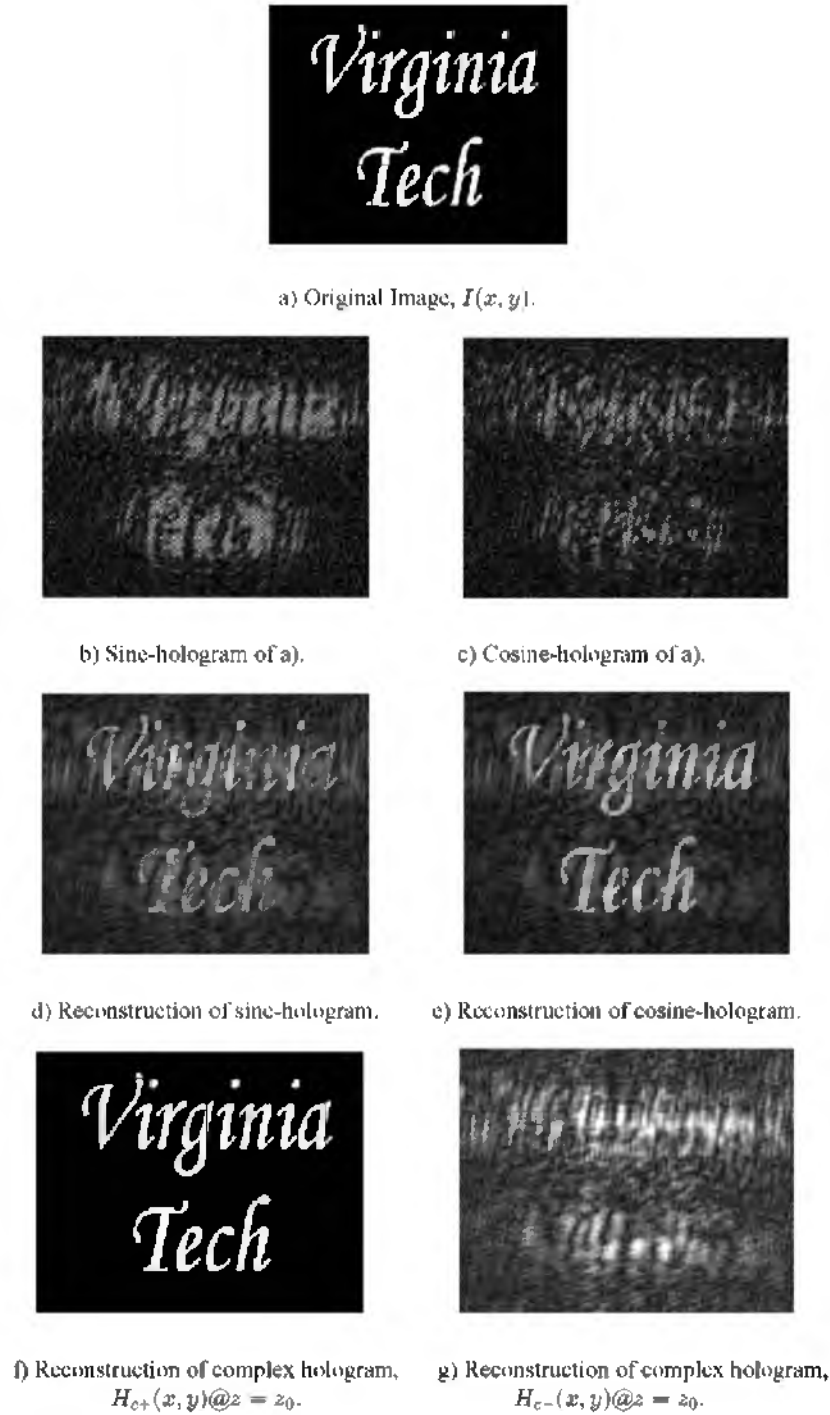


Fig. 3.15 Simulation results using OSH.m.

Table 3.1 OSH.m : m-file for illustrating optical scanning holography.

```

% OSH.m
% Adapted from "Contemporary Optical Image Processing with MATLAB,"
% by Ting-Chung Poon and Partha Banerjee, Table 7.2,
% Pages 222-223, Elsevier (2001).

clear all,
%%Reading input bitmap file
I=imread('vatech.bmp','bmp');
I=I(:,:,1);
figure(1)%displaying input
colormap(gray(255));
image(I)
title('Original image')
axis off
pause

%%Creating OTFosh with SIGMA=z/2*k0 (Eq.(3.5-1a))
ROWS=256;
COLS=256;
sigma=2.0; %not necessary to scale
%kx,ky are spatial frequencies
ky=-12.8;
for r=1:COLS,
    kx=-12.8;
    for c=1:ROWS,
        OTFosh(r,c)=exp(-j*sigma*kx*kx-j*sigma*ky*ky);
        kx=kx+.1;
    end
    ky=ky+.1;
end
max1=max(OTFosh);
max2=max(max1);
scale=1.0/max2;
OTFosh=OTFosh.*scale;

%Recording hologram
% Taking Fourier transform of I
FI=fft2(I);
FI=fftshift(FI);
max1=max(FI);
max2=max(max1);
scale=1.0/max2;
FI=FI.*scale;
% FH is the recorded hologram in Fourier domain
FH=FI.*OTFosh;
H=ifft2(FH);
max1=max(H);
max2=max(max1);
scale=1.0/max2;

```



```

H=H.*scale;
figure(1)
colormap(gray(255));
%Displaying the real part becomes sine-FZP hologram
% Eq. (3.5-8a)
image(2.5*real(256*H));
title('Sine-FZP hologram')
axis off
figure(2)
colormap(gray(255));
%Displaying the imaginary part becomes cosine-FZP hologram
% Eq. (3.5-8b)
image(2.5*imag(256*H));
title('Cosine-FZP hologram')
axis off
%Reconstructing holograms
%Reconstruction of sine-hologram,twin-image noise exists
figure(3)
colormap(gray(255))
H=ifft2(FH);
FRSINEH=fft2(real(H)).*conj(OTFosh); %Eq. (2.5-10)
RSINEH=ifft2(FRSINEH);
image(256*abs(RSINEH)/max(max(abs(RSINEH))))
title('Reconstruction of sine-FZP hologram')
axis off
%FH=FHI;
%Reconstruction with cosine-hologram, twin-image noise exists
figure(4)
colormap(gray(255))
FRCOSINEH=fft2(imag(H)).*conj(OTFosh);
RCOSINEH=ifft2(FRCOSINEH); %Eq. (3.5-10)
image(256*abs(RCOSINEH)/max(max(abs(RCOSINEH))))
title('Reconstruction of cosine-FZP hologram')
axis off

figure(5)
colormap(gray(255))
FRCOMPLEXH=fft2(real(H)+j*imag(H)).*conj(OTFosh);
RCOMPLEX=ifft2(FRCOMPLEXH);
image(1.4*256*abs(RCOMPLEX)/max(max(abs(RCOMPLEX))))
title('Real image reconstruction of complex FZP hologram, Hc+')
axis off

figure(6)
colormap(gray(255))
FRCOMPLEXH2=fft2(real(H)-j*imag(H)).*conj(OTFosh);
RCOMPLEX2=ifft2(FRCOMPLEXH2);
image(1.4*256*abs(RCOMPLEX2)/max(max(abs(RCOMPLEX2))))
title('Reconstruction of complex FZP hologram, Hc-')
axis off
-----

```

### 3.6 Physical Intuition to Optical Scanning Holography

While we have devoted the last section to developing optical scanning holography from a mathematical point of view, in this section we will be describing optical scanning holography from a physical point of view. Again, to accomplish optical scanning holography, we select  $p_1(x, y) = 1$  and  $p_2(x, y) = \delta(x, y)$  in the two-pupil heterodyne image processor (the situation is shown in Fig. 3.10 clearly) where lens L1 forms a point source and lens L2 forms a plane wave on the scanning mirrors. At a distance of  $z = z_0$  away from the scanning mirrors, where the object,  $|\Gamma_0(x, y; z)|^2 = I(x, y)\delta(z - z_0)$ , is located, we have an interference between a plane wave and a spherical wave of different temporal frequencies. Hence, the scanning beam intensity is

$$\begin{aligned}
 I_{\text{scan}}(x, y; t) &= |a \exp[j(\omega_0 + \Omega)t] + \frac{jk_0}{2\pi z_0} \exp[-\frac{jk_0(x^2 + y^2)}{2z_0}] \exp(j\omega_0 t)|^2 \\
 &= A + B \sin[\frac{k_0}{2z_0}(x^2 + y^2) + \Omega t] \\
 &= TDFZP(x, y; z_0, t), \tag{3.6-1}
 \end{aligned}$$

where we have assumed that the plane wave is of amplitude  $a$ , and  $A$  and  $B$  have been defined in Eq. (2.5-3). Note that this equation is basically the same form as that of Eq. (2.5-4) for the recording of an on-axis point source except for the time variable,  $t$ . We shall call this a *time-dependent Fresnel zone plate* (TDFZP), which is used to scan over the object in a raster fashion. For a pinhole object, i.e.,  $I(x, y) = \delta(x, y)$ , the photodetector's current,  $i(x, y)$ , is clearly given by  $TDFZP(x, y; z_0, t)$  as the pinhole samples the intensity pattern, which then gives the intensity pattern as output. We can also see mathematically that if we make use of Eq. (3.1-4), we can obtain

$$\begin{aligned}
 i(x, y) &\sim TDFZP(x, y; z_0, t) \otimes \delta(x, y) \\
 &= TDFZP(x, y; z_0, t) \\
 &= A + B \sin[\frac{k_0}{2z_0}(x^2 + y^2) + \Omega t].
 \end{aligned}$$

After bandpass filtering at  $\Omega$ , the heterodyne current becomes

$$i_{\Omega}(x, y) \sim \sin\left[\frac{k_0}{2z_0}(x^2 + y^2) + \Omega t\right], \quad (3.6-2)$$

and after electronic detection, i.e., multiplying with, say  $\cos(\Omega t)$ , and lowpass filtering, we obtain

$$i_c(x, y) \sim \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right], \quad (3.6-3)$$

which is again Eq. (3.5-4a). So we can see that optical scanning holography is simply accomplished by raster-scanning a TDFZP over a 3-D object in order to obtain two holograms. The physical situation is shown in Fig. 3.16, where in the pupil plane, we have a point source and a plane wave. In the figure, we show the pattern of the scanning beam on the object slice for a fixed time, say, at  $t = t_0 = 0$ , which becomes a “static” Fresnel zone plate (FZP). If we let the time run in Eq. (3.6-1), we will physically have running fringes that will be moving toward the center of the zone pattern. Hence, the basic principle of OSH is to simply use the TDFZP to 2-D scan a 3-D object to obtain holographic information for the scanned object. The m-file presented in Table 3.2 will allow us to generate a TDFZP and illustrate the running of fringes.

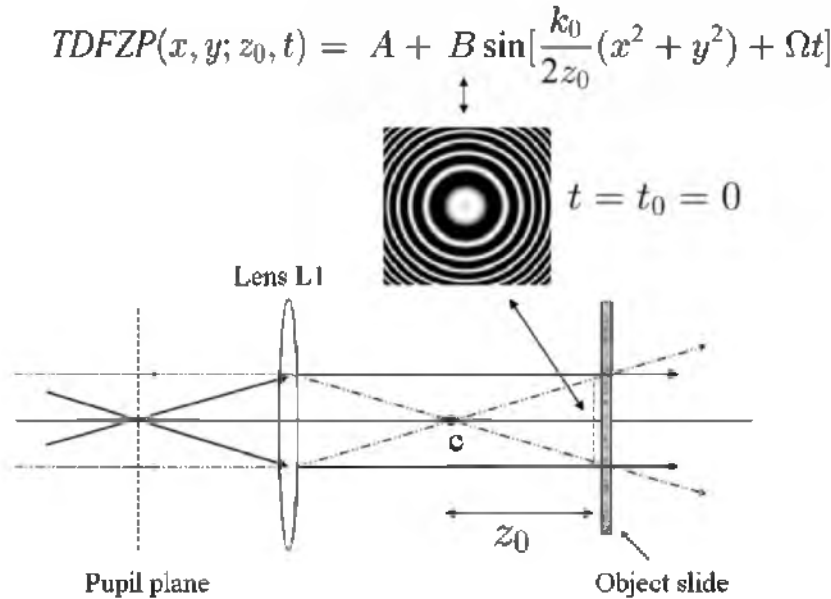


Fig. 3.16 Principle of OSH: use of time-dependent Fresnel zone plate to scan over an object. Adapted from T. -C. Poon, *Journal of Holography and Speckle* 1, 6-25, (2004).

We can also select the pupils differently to achieve optical scanning holography as long as we can create a time-dependent FZP to scan the 3-D object. For example, we can select  $p_1(x, y) = \delta(x, y)$  and  $p_2(x, y) = 1$  instead of  $p_1(x, y) = 1$  and  $p_2(x, y) = \delta(x, y)$  as we had previously discussed in the two-pupil heterodyne image processor. Then, the scanning intensity become

$$\begin{aligned}
 I_{\text{scan}}(x, y; t) &= \left| a \exp(j\omega_0 t) + \frac{jk_0}{2\pi z_0} \exp\left[-\frac{jk_0(x^2 + y^2)}{2z_0}\right] \exp[j(\omega_0 + \Omega)t] \right|^2 \\
 &= A + B \sin\left[\frac{k_0}{2z_0}(x^2 + y^2) - \Omega t\right] \\
 &= TDFZP(x, y; z_0, -t). \tag{3.6-4}
 \end{aligned}$$

This will give fringes running away from the center of the zone, and this can be verified by changing the sign in front of B in the expression of FZP(ii,jj,kk) in the TDFZP.m file. As it turns out, this scanning beam gives the same expression for the sine-hologram given by Eq. (3.5-3a). However, the expression for the cosine-hologram is different with a “negative” sign in front of it. Note that the errors and sensitivities produced by the heterodyne method, as well as the impact created by these errors, will not be discussed in this book. Readers should refer to Section 5.4 of the *Handbook of Holographic Interferometry* [Kreis, 2005], which will provide insight into the practical limitations of heterodyning holographic interferometry.

When  $\Omega$  is set to zero, we have *homodyning* and the time-dependent Fresnel zone plate becomes static. Now by introducing phase shifts between the two interfering waves that are used to generate the Fresnel zone plate, one can obtain three holograms (as a result of three different phase shifts) to alternatively solve the twin-image problem in optical scanning holography [Rosen, Indebetouw and Brooker (2006)].

Table 3.2 TDFZP.m : m-file for illustrating running fringes in TDFZP.

```

-----
% TDFZP.m
% Illustration of running fringes in TDFZP
% The author thanks Kelly Dobson for her initial programming

clear;
B=10.01*10^6; %temporal frequency, arbitrary
D=6; %Scale arbitrary

t=linspace(0,1,35);
x=linspace(-2.5,2.5,256);
y=linspace(-2.5,2.5,256);

```

```

for ii=1:length(x)
for jj=1:length(y)
for kk=1:length(t)
FZP(ii,jj,kk)=(1+sin(D*(x(ii)^2+y(jj)^2)+B*t(kk))); %TDFZP
end
end
end

for ll=1:length(t)
max1=max(FZP(:, :, ll));
max2=max(max1);
scale=1/max2;
FZP(:, :, ll)=FZP(:, :, ll).*scale;
figure(ll);
colormap(gray(256));
image(256*FZP(:, :, ll));
axis off
F(ll)=getframe;
end

movie(F,10)
-----

```

## References

- 3.1 Duncan, B.D. and T.-C. Poon (1992). "Gaussian beam analysis of optical scanning holography," *Journal of the Optical Society of America A* 9, 229-236.
- 3.2 Doh, K., T.-C. Poon, M. Wu, K. Shinoda, and Y. Suzuki (1996). "Twin-image elimination in optical scanning holography," *Laser & Optics Technology* 28, 135-141.
- 3.3 General Scanning. <http://www.gsig.com/scanners/>
- 3.4 Indebetouw, G. and T.-C. Poon (1992). "Novel approaches of incoherent image processing with emphasis on scanning methods," *Optical Engineering* 31, 2159-2167.
- 3.5 Indebetouw, G., P. Klysubun, T. Kim, and T.-C. Poon (2000). "Imaging properties of scanning holographic microscopy," *Journal of the Optical Society of America A* 17, 380-390.
- 3.6 Indebetouw, G. and W. Zhong (2006). "Scanning holographic microscopy of three-dimensional fluorescent specimens," *Journal of the Optical Society of America A* 23, 1699-1707.
- 3.7 IntraAction Corp. <http://www.intraaction.com/>
- 3.8 Kim, T. and T.-C. Poon (1999). "Extraction of 3-D location of matched 3-D object using power fringe-adjusted filtering and Wigner analysis," *Optical Engineering* 38, 2176-2183.
- 3.9 Korpel, A. (1981). "Acousto-optics-a review of fundamentals," *Proceedings of the IEEE* 69, 48-53.
- 3.10 Kreis, T. (2005). *Handbook of Holographic Interferometry*. Wiley-VCH GmbH & Co. KGaA, Weinheim.
- 3.11 Lohmann, A. W. and W. T. Rhodes (1978). "Two-pupil synthesis of optical transfer functions," *Applied Optics* 17, 1141-1150.

- 3.12 Poon, T.-C. (1985). "Scanning holography and two-dimensional image processing by acousto-optic two-pupil synthesis," *Journal of the Optical Society of America A* 4, 521-527
- 3.13 Poon, T.-C. (2002a). "Three-dimensional television using optical scanning holography," *Journal of Information Display* 3, 12-16.
- 3.14 Poon, T.-C. (2002b). "Acousto-Optics," *Encyclopedia of Physical Science and Technology*, Academic Press.
- 3.15 Poon, T.-C. (2004). "Recent progress in optical scanning holography," *Journal of Holography and Speckle* 1, 6-25.
- 3.16 Poon, T.-C. (2006). "Horizontal-parallax-only optical scanning holography," in chapter 10 of *Digital Holography and Three-Dimensional Display: Principles and Applications*, T.-C. Poon ed., Springer, New York, USA.
- 3.17 Poon, T.-C. and A. Korpel. (1979). "Optical transfer function of an acousto-optic heterodyning image processor," *Optics Letters* 4, 317-319.
- 3.18 Poon, T.-C., K. Doh, B. Schilling, M. Wu, K. Shinoda, and Y. Suzuki (1995). "Three-dimensional microscopy by optical scanning holography," *Optical Engineering* 34, 1338-1344.
- 3.19 Poon, T.-C. and T. Kim (1999). "Optical image recognition of three-dimensional objects," *Applied Optics* 38, 370-381.
- 3.20 Poon, T.-C., T. Kim, G. Indebetouw, B. W. Schilling, M. H. Wu, K. Shinoda, and Y. Suzuki (2000). "Twin-image elimination experiments for three-dimensional images in optical scanning holography," *Optics Letters* 25, 215-217.
- 3.21 Poon T.-C. and P. P. Banerjee (2001). *Contemporary Optical Image Processing with MATLAB®*. Elsevier, Oxford, UK.
- 3.22 Poon, T.-C., T. Kim and K. Doh (2003) "Optical scanning cryptography for secure wireless transmission," *Applied Optics* 42, 6496-6503.
- 3.23 Poon, T.-C. and T. Kim (2006). *Engineering Optics with MATLAB®*. World Scientific Publishing Co., Singapore.
- 3.24 Pratt, W.K. (1969). *Laser Communications Systems*, John Wiley & Sons.
- 3.25 Rosen, J., G. Indebetouw, and G. Brooker. "Homodyne scanning holography," *Optics Express* 14, 4280-4285.
- 3.26 Yamaguchi, I. and T. Zhang (1997). "Phase-shifting digital holography," *Optics Letters* 22, 1268-1270.

## Chapter 4

### Optical Scanning Holography: Applications

So far, the applications of optical scanning holography span over five diverse areas: scanning holographic microscopy [(Poon, Doh, Schilling, Wu, Shinoda, and Suzuki (1995)], 3-D image recognition [(Poon and Kim (1999)], 3-D optical remote sensing [Kim and Poon (1999)], 3-D holographic TV and 3-D display [Poon (2002a)], and 3-D cryptography [Poon, Kim, and Doh (2003)]. In this chapter, we will only cover three areas of the applications mentioned above. We will focus on scanning holographic microscopy, 3-D holographic TV and 3-D display, and 3-D cryptography in that order, as the other areas have been recently reviewed in the book chapters [Poon (2002b), Poon (2005)].

#### 4.1 Scanning Holographic Microscopy

Three-dimensional (3-D) imaging is a formidable task for optical microscopy as it is well known that the greater the *lateral resolution*  $\Delta r$  is, the shorter the *depth of focus*  $\Delta z$  will be. In other words, if we want to create a higher lateral resolution in the microscopic imaging system, say by using a high *numerical aperture* (NA) lens, we will have to compromise the system with a shorter depth of focus where only a thin section of the specimen can be imaged. To demonstrate this fact, a simple quantum mechanical argument is convenient to use here.

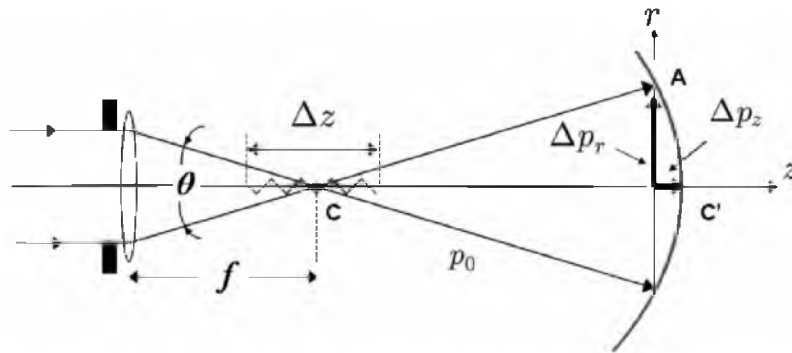


Fig. 4.1 Uncertainty principle used to find resolution and depth of focus.

Let us first find the lateral resolution,  $\Delta r$ . Quantum mechanics relates the minimum uncertainty in a position of quantum,  $\Delta r$ , to the uncertainty of its momentum,  $\Delta p_r$ , according to the relationship

$$\Delta r \Delta p_r \geq h, \quad (4.1-1)$$

where  $\Delta p_r$  is the momentum difference between rays CA and CC' along the  $r$ -direction, i.e., the transverse direction as shown in Fig. 4.1, where parallel rays are focused by a lens. Now the momentum of the CA ray and the CC' ray is  $p_0 \sin(\theta/2)$  and zero along the  $r$ -direction, respectively, where  $p_0 = h/\lambda_0$  is the momentum of the quantum. Hence,  $\Delta p_r = p_0 \sin(\theta/2)$ . By substituting this into Eq. (4.1-1), we obtain

$$\Delta r \geq \frac{h}{\Delta p_r} = \frac{h}{p_0 \sin(\theta/2)} = \frac{h}{(h/\lambda_0) \sin(\theta/2)} = \frac{\lambda_0}{\sin(\theta/2)}.$$

If the object space or the specimen space has a refractive index of  $n_0$ , we must use the fact that the wavelength in the medium is equal to  $\lambda/n_0$ , where  $\lambda$  is the wavelength in air or in vacuum. The above equation, therefore, becomes

$$\Delta r \geq \frac{\lambda}{n_0 \sin(\theta/2)} = \frac{\lambda}{NA}, \quad (4.1-2)$$

where  $NA = n_0 \sin(\theta/2)$  is called the *numerical aperture*. Similarly, to find the depth of focus,  $\Delta z$ , we have

$$\Delta z \Delta p_z \geq h, \quad (4.1-3)$$

where  $\Delta p_z$  is the momentum difference between rays CA' and CC' along the  $z$ -direction as shown in Fig. 4.1, which is given by

$$\Delta p_z = p_0 - p_0 \cos(\theta/2).$$

By substituting this expression into Eq. (4.1-3), we have

$$\Delta z \geq \frac{h}{\Delta p_z} = \frac{h}{p_0 [1 - \cos(\theta/2)]} = \frac{\lambda_0}{[1 - \cos(\theta/2)]},$$

which can be written as

$$\Delta z \geq \frac{\lambda_0}{[1 - \sqrt{1 - \sin^2(\theta/2)}]}$$



$$\approx \frac{2\lambda_0}{\sin^2(\theta/2)} = \frac{2n_0\lambda}{NA^2}, \quad (4.1-4)$$

where we have used  $\sqrt{1 - \sin^2(\theta/2)} \approx 1 - \frac{1}{2}\sin^2(\theta/2)$  by assuming  $\sin^2(\theta/2) \ll 1$  in order to obtain the last expression. Now, by combining Eqs. (4.1-2) and (4.1-4), we have

$$\frac{(\Delta r)^2}{\Delta z} \geq \frac{\lambda_0}{2}. \quad (4.1-5)$$

This “*uncertainty relationship*” tells us that, for example, by increasing the lateral resolution by a factor of two, the depth of focus must then be decreased by a factor of four. Hence, we see that the higher the lateral resolution is, the shorter the depth of focus will be. Three-dimensional imaging in microscopy therefore aims to develop techniques that can provide high lateral resolution, and at the same time maintain a large depth of focus in order to observe a thick specimen without any difficulty.

In the past decade, we have witnessed an impressive emergence of three-dimensional (3-D) imaging techniques for microscopy. *Optical sectioning microscopy* and *scanning confocal microscopy* are the two most common techniques currently available in practice.

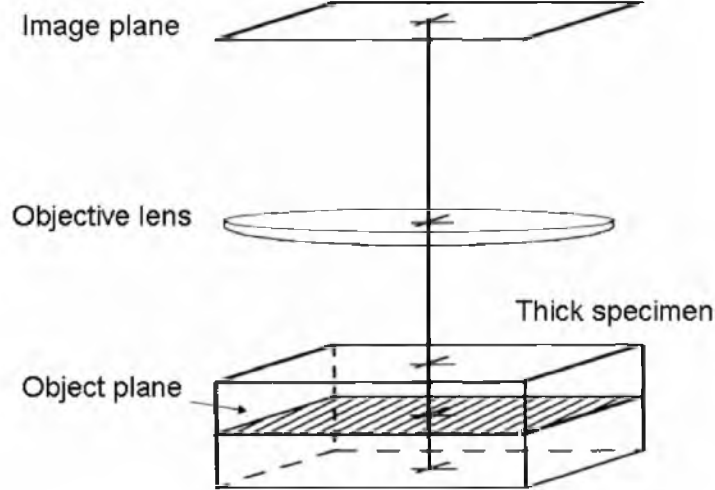


Fig. 4.2 Optical sectioning microscope (OSM).

An optical sectioning microscope (OSM), shown in Fig. 4.2, uses a wide-field microscope to sequentially record a series of images focused at different depths [Agard (1984)]. Since each 2-D image contains the in-focus as well as the out-of-focus information, reconstruction of the 3-D

information, i.e., extraction of the in-focus information from these 2-D images, is required. Many reconstruction algorithms have been developed for this purpose. However, the difficulty of optical sectioning lies in the fact that during the recording stage it is important that exact longitudinal spacing between adjacent 2-D images must be accurately controlled. Also, precise registration of the 2-D images is critical even before any computer processing can be performed.

By recognizing these problems, a radically new microscope design known as the scanning confocal microscope (SCM) has emerged [Wilson and Sheppard (1984)]. The confocal principle was first described by Minsky [Minsky, US Patent (1961)]. In scanning confocal microscopy, a doubly focused objective lens system and a pin-hole aperture in front of a photodetector are used to image only a single point within the 3-D specimen, as shown in Fig. 4.3. All the light from the point on the plane of focus (solid rays) is focused at the pin-hole aperture and passed into the photodetector. The light from the out-of-focus plane (dashed rays) is rejected by the pin-hole. Three-dimensional information is gathered by scanning the specimen in three dimensions while collecting the light transmitted through the specimen with the photodetector.

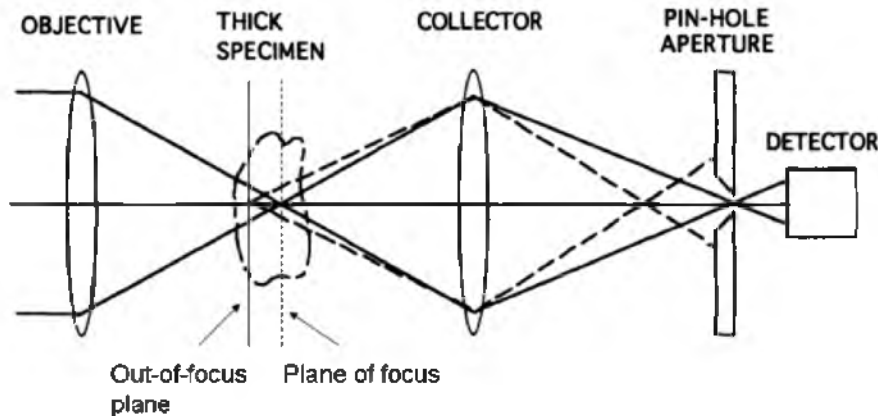


Fig. 4.3 Scanning confocal microscope (SCM).

Theoretically, SCMs provide a slightly better lateral resolution than that of OSMs. If the lateral resolution in optical sectioning microscopy is  $\Delta r = \lambda/NA$ , the lateral resolution of the scanning confocal microscope is  $0.73\Delta r$  [Corle and Kino (1996)]. In other words, one can achieve better resolution through confocal imaging. However, such a theoretical limit has never been achieved in practice. One of the main problems associated with SCMs is that the scanning instrumental tolerances required to achieve high resolution imaging and a long working *depth of field* are very difficult to obtain in practice. The term *depth of field*,  $\Delta l$ , here refers to the range of

object (or specimen) distances that are imaged within a distance,  $\Delta z$ , i.e., the depth of focus in the image space. In fact, if  $M$  is the lateral magnification of the imaging system, then  $\Delta l = \Delta z/M^2$  [see Example 4.1, page 112].

In essence, both methods (OSM and SCM) require precise 3-D positioning devices. This is particularly critical for the confocal methods whose technologically sophisticated and expensive equipment need special technical training to ensure proper use. However, for certain applications in biology, the main drawback of these instruments is that the data is sequentially acquired by a slow 3-D scan. This tedious data acquisition time is a serious drawback for *in vivo* studies. It precludes, for example, the possibility of monitoring dynamic interactions at the intercellular level. In addition, an extensive data acquisition time exacerbates the *photo-bleaching* problem in *fluorescence microscopy* [Pawley (1995)]. Briefly, damage caused by photo-bleaching refers to the fact that a specimen will not fluoresce when it is over-exposed. The severity of this problem in cell studies has prompted the development of extremely sophisticated techniques such as *two-photon scanning fluorescence imaging*, for example. We also want to point out that *optical coherent tomography* (OCT), which is based on interferometry, is another pertinent and emerging technique used for 3-D microscopy [Huang, Swanson, Lin, Shuman, Stinson, Chang, Hee, Flotte, Gregory, Puliafito, and Fujimoto (1991)]. But once again, this technique also requires scanning the object along the depth direction. Actually, all existing commercialized microscopes (OSM, SCM, and OCT) require axial-scanning in order to achieve 3-D imaging. Eliminating the need for a 3-D scan, or more specifically for a depth-scan gives us the impetus to study novel holographic methods for 3-D microscopy.

Holography can be used in applications where 3-D data sets are desirable but axial scanning is difficult or sometimes impossible to utilize. With holography, we have the ability to capture high resolution 3-D information in a large volume space. In recent years, holographic microscopy has become more and more prevalent because it represents a novel departure from conventional 3-D microscopy mentioned above [Zhang and Yamaguchi (1998), Kim (2000)]. Traditionally, holographic microscopy has been used in biology, however it is inherently insensitive to incoherent emissions, such as fluorescence, which makes its usefulness severely limited in life sciences applications. A scanning holographic microscope (SHM), which is based on the principle of optical scanning holography (OSH), can acquire 3-D information by only using single 2-D x-y scanning (no axial scanning is involved and hence there are reduced acquisition times for 3-D imaging). Most importantly, the SHM has the ability to image fluorescent samples, which is a breakthrough in holography since conventional holographic microscopes simply could not capture fluorescent specimens prior to the invention of optical scanning holography [Poon, Schilling, Indebetouw, and

Storrie, U.S. Patent (2000)]. In addition, the SHM will provide better theoretical transverse resolution even in comparison to a SCM. The resolution of a SHM is  $0.5\Delta r$ , where  $\Delta r$  is again the resolution of a wide field microscope [Indebetouw (2002)]. We will further elaborate on this topic in chapter 5. Note that OCT techniques are also not capable of fluorescent imaging. Hence, OSH is a very unique technique for 3-D biomedical applications.

In the spirit of keeping the scanning holographic microscope (SHM) at the same level of simplicity illustrated with the OSM and the SCM shown in Figs. 4.2 and 4.3, respectively, we show a SHM in Fig. 4.4. In principle, we need a time-dependent Fresnel zone plate (TDFZP) to two-dimensionally raster scan the thick specimen as shown in Fig. 4.4.

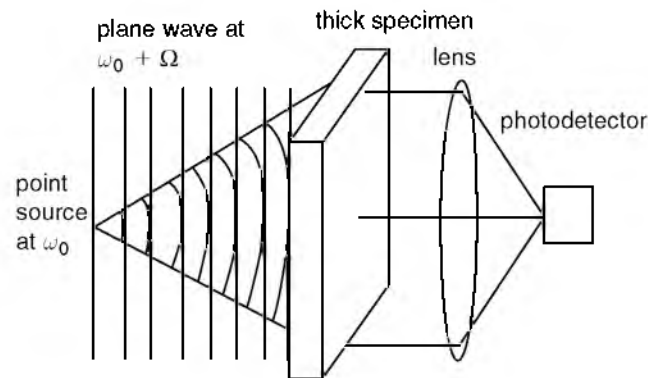


Fig. 4.4 Scanning holographic microscope (SHM).

Figure 4.5a) shows the actual setup of a scanning holographic microscope used for fluorescence applications. In this figure, two broad laser beams (originating from a 514nm-line of an argon laser), separated in temporal frequency by  $\Delta\Omega$ , are incident on the mirror and the beamsplitter (BS). The frequency shift in each beam is achieved by using an acousto-optic frequency shifter (AOFS). The AOFS is used in a configuration that allows us to split the laser into two beams separated in frequency by  $\Delta\Omega/2\pi = 10.7$  MHz. The beams are then collimated and set parallel to each other as shown in the figure. Lens  $L_1$  is placed in one of the beams to form the spherical wave, which is then combined collinearly with the other beam at the beamsplitter (BS). This will form a TDFZP on the object, which is at a distance of  $z$  beyond the focus of the spherical wave.

The dichroic beamsplitter transmits light at 514nm and reflects light at around 595nm. Therefore, the laser light is allowed to pass through the dichroic beamsplitter and excite the fluorescent sample, which fluoresces at 560nm. Also, the emission filter directly in front of the photomultiplier tube

(PMT) allows fluorescent light to pass through while rejecting the background laser light at 514nm. The sample is scanned through the TDFZP in a raster pattern using a computer-controlled mechanical x-y scanning platform. The current of the PMT, which contains holographic information of the scanned object, is electronically filtered, amplified at 10.7 MHz, demodulated and then digitized in synchronization with the x-y scanners in order to ultimately produce an electronic hologram. Note that the demodulation is performed by the usual electronic detection as we had previously discussed in Fig. 3.6. However, only one channel, specifically the in-phase component of the heterodyne current, was utilized during the experiment.

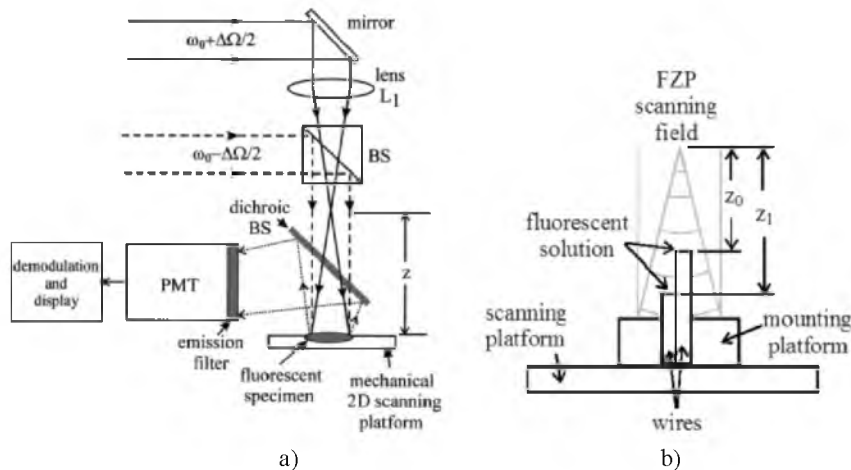


Fig. 4.5a) Experimental setup used to record the hologram of a fluorescent specimen by using OSH. PMT is a photomultiplier tube. Reprinted from B.W. Schilling et al., *Optics Letters* 22, 1506 (1997), with permission. © OSA. b) Experimental configuration of fluorescent solution on the ends of two wires. After Schilling (1997).

The fluorescent sample used in the experiment consists of a solution containing a high concentration of fluorescent latex beads. The beads are 15  $\mu\text{m}$  in diameter and characteristically reached peak excitation at 530nm and peak emission at 560nm. To demonstrate the depth-discriminating capability of the system, we use a fluorescent object that consists of two wires adjacently placed and parallel to the optical axis but with their ends at slightly different distances from the focus of lens L1. A drop of fluorescent solution is placed on the end of each of these wires, and the two drops of solution are separated in depth by approximately 2mm (the drop on the right at  $z_0 \approx 35$  mm and the drop on the left at  $z_1 \approx 37$  mm, as shown in Fig. 4.5b). A hologram of this fluorescent sample was recorded and is displayed in Fig. 4.6. The two drops are easily distinguishable in the hologram.

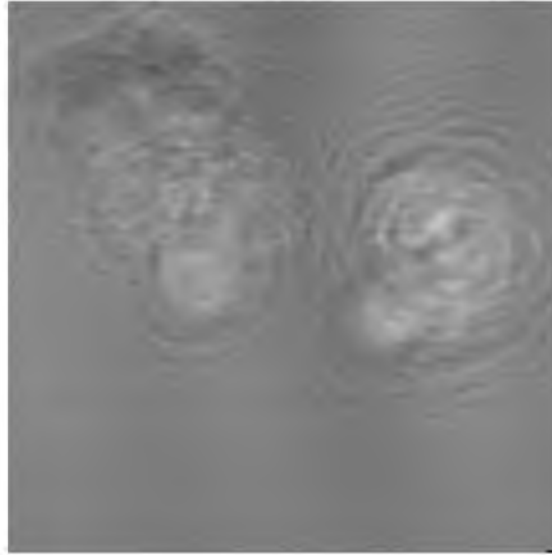


Fig. 4.6 Hologram of a fluorescent specimen recorded by using optical scanning holography. The object consists of two drops of solution containing a high concentration of fluorescent latex beads separated in depth by about 2 mm. The image is at a 256 level gray scale image consisting of 256x256 pixels. The area scanned is about 2.0 mm x 2.0 mm. Reprinted from B.W. Schilling et al., *Optics Letters* 22, 1506 (1997), with permission. © OSA.

The resolution of the OSH system is limited by the system's numerical aperture (NA), which actually depends on the focal length of lens  $L_1$  ( $f = 150\text{mm}$ ) and the diameter of the plane wave focused by lens  $L_1$  ( $D = 10\text{mm}$ ). The NA of the system is approximately 0.033 which corresponds to the diffraction-limited resolution limit of  $\Delta r \approx 18.5\mu\text{m}$  and  $\Delta z \approx 1028.4\mu\text{m}$ , according to Eqs. (4.1-2) and (4.1-4), respectively. The  $15\mu\text{m}$  bead size is very close to the limit that we can expect to resolve laterally with the setup.

Once the hologram has been recorded and stored, the 3-D image can then be reconstructed either optically or numerically. Numerical image reconstruction has been performed on the hologram at two different depths. Figure 4.7a) is a reconstructed image at  $z_0 = 35\text{mm}$ , and Fig. 4.7b) is an image reconstruction at  $z_1 = 37\text{mm}$ . Since the individual attributes of each fluorescent drop are not obvious in these figures, arrows are marked on the figures to indicate particular areas of interest. In Fig. 4.7a), the fluorescent drop on the left is in better focus than that on the right. The arrow in Fig. 4.7a) indicates particular beads that are more clearly visible when the hologram is reconstructed at a depth of  $z_0$  than in Fig. 4.7b) for a depth at  $z_1$ . Similarly, the arrow in Fig. 4.7b) points out a string of four beads that are individually distinguishable when the hologram is reconstructed at a depth of  $z_1$ , but that are blurred in the image reconstruction plane  $z_0$  in Fig. 4.7a).

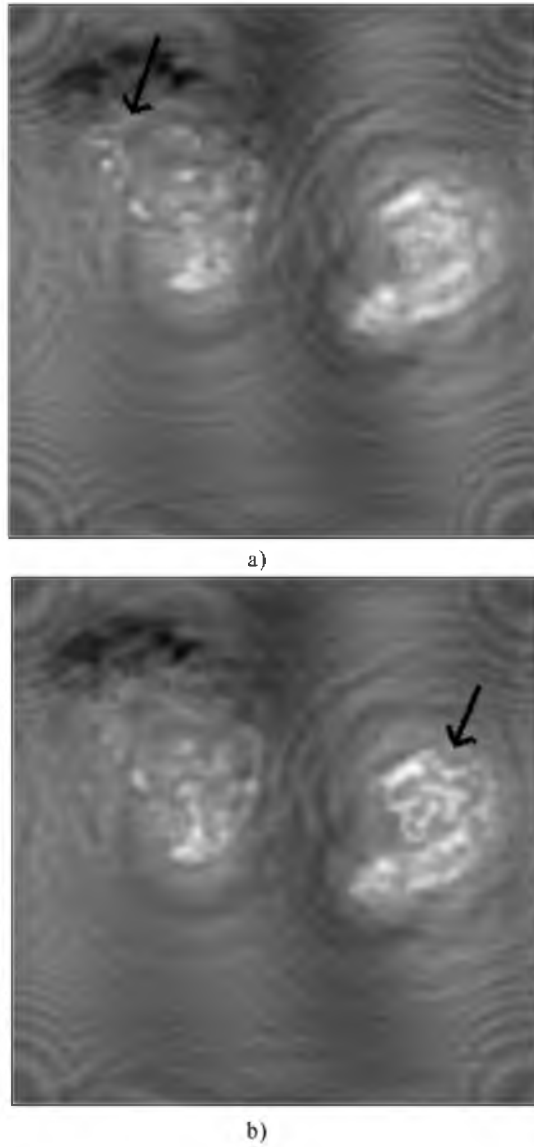


Fig. 4.7 Reconstruction of the hologram shown in Fig. 4.6. a) At a depth of  $z_0 = 35\text{mm}$ . Arrow shows individual fluorescent beads that are in focus at this depth. b) At a depth of  $z_1 = 37\text{mm}$ . The arrow shows four individual fluorescent beads that are in focus at this depth. Reprinted from B.W. Schilling et al., *Optics Letters* 22, 1506 (1997), with permission. © OSA.

Note that since only the in-phase component of the heterodyne current has been recorded to create the hologram shown in Fig. 4.6, twin-image noise (residual “fringing”) exists in these reconstructions.

An important attribute of scanning holographic imaging is that this is the first time the hologram of a fluorescent specimen has been recorded by using an optical holographic technique [Poon et al., U.S. Patent (2000)]. Holography and fluorescence imaging would never seem to make conventional sense because holographic techniques require the coherent interference of light waves and fluorescence imaging does not generate coherent light. And yet, we have been able to record holograms of fluorescent specimens because optical scanning holography makes this possible. In fact, optical scanning holographic techniques can be applied to 3-D biomedical applications as fluorescence imaging [Indebetouw, Kim, Poon, and Schilling (1998)] as well as near-infrared imaging [Sun and Xie (2004)] through turbid media have been demonstrated. Most recently a better than  $1\mu\text{m}$ -resolution also has been established with a holographic fluorescence microscope [Indebetouw and Zhong (2006)].

## 4.2 Three-Dimensional Holographic TV and 3-D Display

Figure 4.8 shows a conceptual holographic system used for 3-D display with a complete recording and reconstruction stages of a point source object. As we had mentioned in chapter 2, if the recording film is replaced by, say, some electronic device such as a CCD video camera, then we can create a 3-D display by transferring the CCD’s electronic output into some spatial light modulator. As we transfer the holographic information at video rate to a spatial light modulator, we create a holographic 3-D display system.

The first television transmission of a hologram was demonstrated by Enloe, Murphy, and Rubinstein [1966]. A television camera was used to record an off-axis hologram where the interference between the Fresnel diffraction pattern of an object transparency and an off-axis plane wave was recorded. The recorded hologram was then transmitted over a closed-circuit TV and displayed on a 2-D monitor. The displayed 2-D record was then photographed to form a hologram, which was subsequently reconstructed by a coherent optical system. Since then, much progress has been made and many novel devices have been invented [Macovski (1971), Brown, Noble and Markevitch, U.S. Patent (1983), Kirk, International Patent (1984), Benton (1991), Shinoda, Suzuki, Wu and Poon, U.S. Patent (1991), Schilling and Poon, U.S. Patent (2004)].



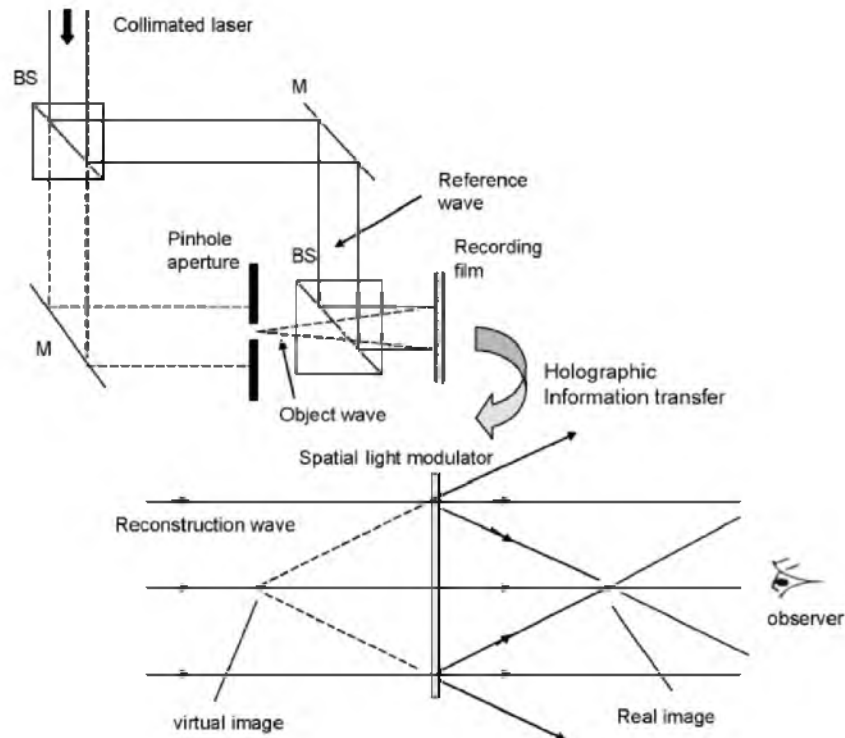


Fig. 4.8 Conceptual holographic system for 3-D display.

In this section, we will describe a recently proposed holographic TV system that uses optical scanning holography (OSH) to acquire holographic information and employs a spatial light modulator (SLM) for an eventual coherent 3-D display [Poon (2002a)]. We should be familiar with OSH by now and, therefore, we will first describe the SLM used in this system. The overall system will be subsequently discussed. The SLM, which has been experimented on the proposed TV holographic system, is called an *electron-beam-addressed spatial light modulator* (EBSLM) [Hamamatsu Photonics K.K., Japan and Hamamatsu Corp., Bridgewater, NJ]. The device is shown in Fig. 4.9.

A serial video signal is the required input to the EBSLM controller. The controller in turn provides the signal that modulates the intensity of the emission from the electron gun within the EBSLM head. This electron beam is then two-dimensionally scanned onto the surface of a  $\text{LiNbO}_3$  crystal with a deflection coil. As a result, electric charges accumulate on the surface of

the crystal. The spatially induced electric field deforms the crystal as a result of the *Pockels effect* [Poon and Kim (2006)]. A pair of crossed polarizers is used in order to read the resulting spatial distribution on the crystal by laser. Conjointly, a coherent spatial distribution of the output laser would correspond to the 2-D scanned video information on the crystal.

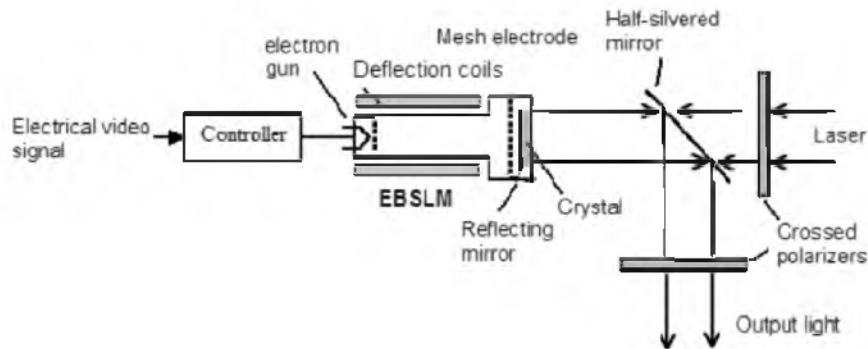


Fig. 4.9 Electron-beam-addressed spatial light modulator (EBSLM) for coherent display.

By incorporating optical scanning holography for holographic recording with EBSLM for coherent display, we can create a complete holographic TV system. This is shown in Fig. 4.10, where we have included a system for optical scanning holography on the top portion of the figure.

In the top part of the system, M1, M2, and M3 represent the mirrors, BS1 and BS2 denote the beamsplitters, AOM is an acousto-optic modulator used to shift the laser beam at a frequency of  $\Omega$ , and BE1 and BE2 are the beam expanders. Note that lens L is used to focus a point source on the BS2, that projects a spherical wave through the x-y scanner to the object, while BE2 provides a plane wave onto the object. After 2-D raster scanning of the object, the photomultiplier picks up the scattered light from the object and delivers a heterodyne current as an output current. If the heterodyne current is at the radio frequency (rf) range, then it can be directly radiated through an antenna to a remote site for demodulation. At the demodulation site, we will have the usual electronic multiplexing detection. The PC can manipulate the two holograms (sine and cosine holograms) and thus deliver its output to the controller of the EBSLM for coherent reconstruction of the holographic information in order to display the output light for the audience. Hence, we have a complete holographic TV system. This system has been proposed by Poon [2002a], and this idea of using OSH to acquire holographic information and to use SLMs for display has been tested in the system shown in Fig. 4.11. It is clear from the figure that a TDFZP is used to scan the 3-D object,

and the photodetector's output is bandpass filtered at,  $\Omega$  and then it is finally mixed with  $\cos(\Omega t)$  to get the sine-FZP hologram,  $i_c(x, y)$ , as given by Eq. (3.5-3a), for coherent reconstruction by the EBSLM.

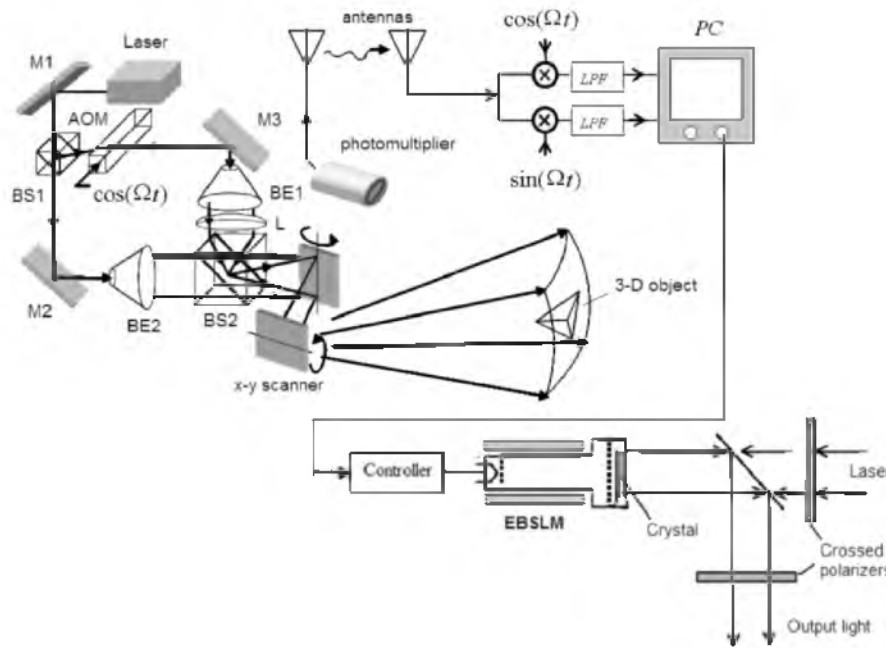


Fig. 4.10 Proposed holographic TV system. Adapted from T.-C. Poon, J. Information Display 3, 12 (2002a).

The Image Processing and Measuring System (IPMS) is a device that acts as an interface that accepts a slow-scan electrical signal and stores the information in its digital memory. The information is then converted into a NTSC video signal [Hamamatsu Photonics K.K. and Hamamatsu Corp., NJ]. When the video of the IPMS is displayed on a TV monitor, the sine-FZP hologram of a 3-D object is displayed as shown in Fig. 4.12. The 3-D object consists of two transparencies, the letters "V" and "T," located side by side but separated by a depth-distance of about 15 cm. The "V" is located closer to the 2-D scanner at a distance of about 23 cm, i.e.,  $z = 23$  cm. Both letters are printed on 35mm film, have a line width of about  $100\mu\text{m}$ , and are transmissive on an opaque background. By passing the reflected light of the EBSLM through an analyzer as shown in Fig. 4.11, a coherent image is reconstructed at a distance of  $M \times z$  away from the analyzer, where  $z$  is the distance from the scanning mirror to the object (as indicated at the holographic recording stage in the figure), and  $M$  is a magnification factor

that takes into account the longitudinal magnification of the holographic imaging system that arises because of various hologram scalings. An example of the cause of *hologram scaling* may be that the displaying area of the hologram in the EBSLM is different from the actual optical scan area of the object. Holographic magnification is discussed in Example 4.1.

The reconstruction of a hologram along depth can be observed through the movable CCD camera, which focuses on different reconstruction planes. Figure 4.13a), c), and e) show the real-time reconstruction of a hologram of different depths by using the electron-beam-addressed spatial light modulator. In the 3-D reconstruction, the  $M \times z$  for Fig. 4.13a) and Fig. 4.13e) are 23cm and 41cm, respectively. In Fig. 4.13a), we notice that the “V” is in focus, and in Fig. 4.13e) the “T” is now in focus. Also note that the reconstructed image planes have been contaminated by the twin-image noise because only one channel, namely the sine-FZP hologram, has been used. For comparison, we have shown digital reconstructions in Fig. 4.13b), d), and f) [Poon, Doh, Schilling, Wu, Shinoda, and Suzuki (1995)].

The EBSLM system is capable of displaying holograms at a video rate and, of course, some commercial  $x-y$  scanners are also capable of working at a video rate. But what we have done is really the use of a SLM to display the acquired hologram along depth for coherent reconstruction. So what is the prospect of displaying true 3-D images in holographic television?

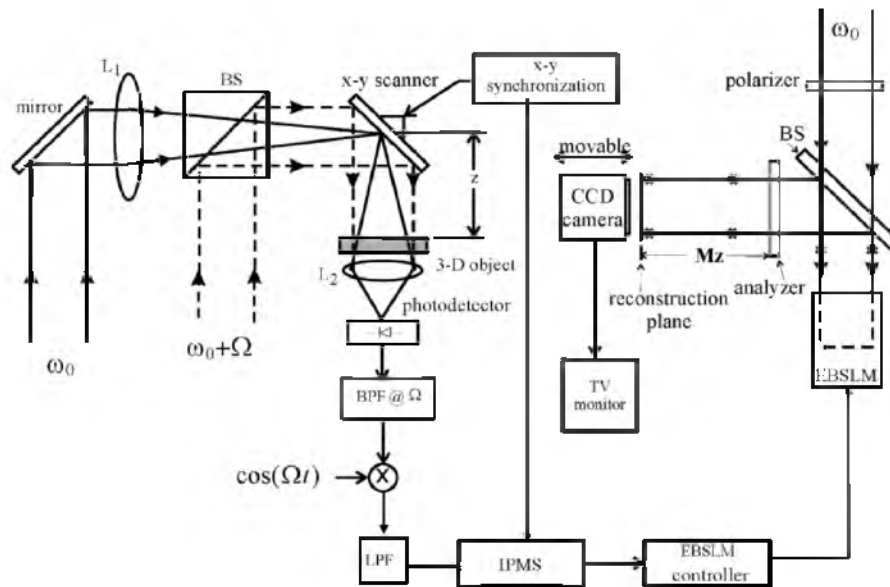
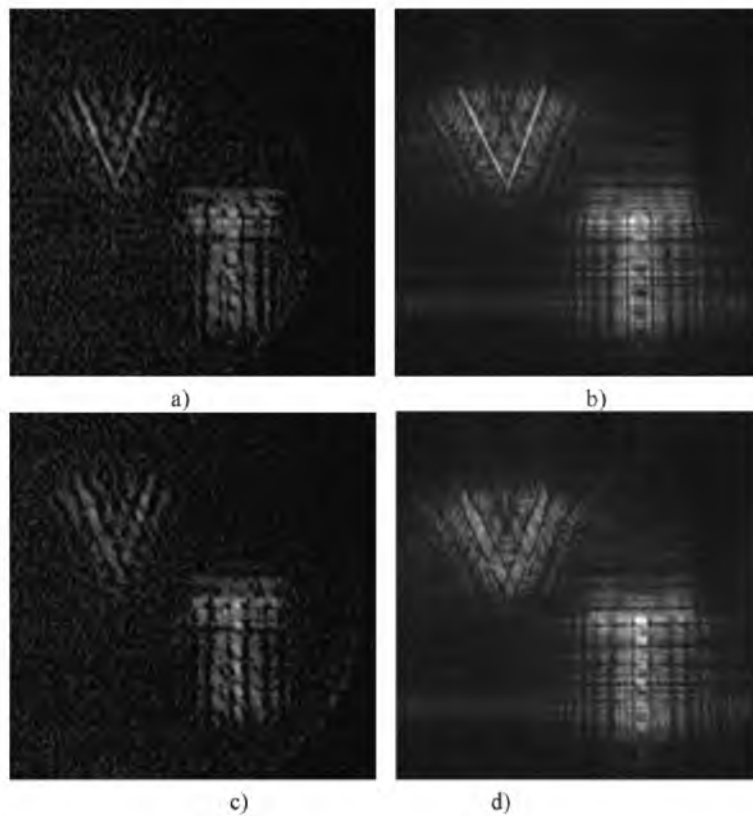


Fig. 4.11 Experimental 3-D holographic television system.  
Adapted from T.-C. Poon et al. Optical Review 4, 576 (1997).



Figure 4.12 Sine-FZP hologram of two letters "V" and "T" located at different depths.  
Reprinted from T.-C. Poon et al., *Optical Engineering* 34, 1338 (1995),  
with permission. © SPIE.



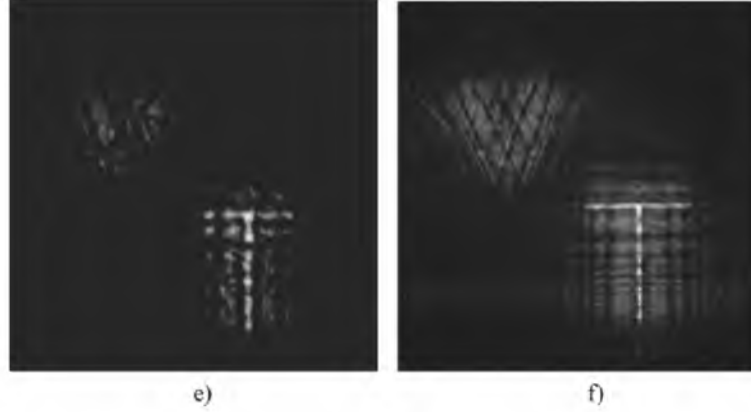


Fig. 4.13 Holographic Reconstruction. a), c), and e) EBSLM's reconstruction. Reprinted from T.-C. Poon et al., Optical Review 4, 576 (1997); b), d), and f) Digital reconstruction. Reprinted from T.-C. Poon et al., Optical Engineering 34, 1338 (1995), with permission. © SPIE.

#### Example 4.1 Holographic Magnification

We will derive holographic magnification in the context of OSH. We consider a three-point object given by

$$\delta(x, y, z - z_0) + \delta(x - x_0, y, z - z_0) + \delta(x, y, z - (z_0 + \Delta z_0)), \quad (4.2-1)$$

where the first two points are located at a distance of  $z_0$  away from the point source, which generates the spherical wave shown in Fig. 4.4. The two points have a lateral separation of  $x_0$ . The third point is located at a distance of  $z_0 + \Delta z_0$  away from the first two points. According to Eq. (3.5-3a), when this three-point object is scanned, the scanned demodulated electrical signal,  $i_c$ , gives a sine-hologram, which is given by

$$\begin{aligned} H_{3-p}(x, y) \sim & \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right] + \sin\left\{\frac{k_0}{2z_0}[(x - x_0)^2 + y^2]\right\} \\ & + \sin\left[\frac{k_0}{2(z_0 + \Delta z_0)}(x^2 + y^2)\right]. \end{aligned} \quad (4.2-2)$$

If this hologram is illuminated by a plane wave at  $\lambda_0$ , the three points will be reconstructed at their respective locations. We will now consider holographic magnification.

## a) Hologram Scaling

Magnification can be achieved by enlarging holograms, however, it is a difficult task especially when we deal with off-axis holograms where fringe densities are of the order of several thousands lp/mm. Most photographic enlargers do not have a sufficient resolution in order to handle these details. Hence, the method is not very practical. With the scanning technique, however, on-axis holograms are generated and scaling is straightforward. The hologram can be scaled by a factor of  $M$ , simply by displaying the hologram in an area that is different from the optical scan area. In this case, Eq. (4.2-2) becomes

$$\begin{aligned}
 H_{3-p}(Mx, My) = & \sin\left\{\frac{k_0}{2z_0}[(Mx)^2 + (My)^2]\right\} \\
 & + \sin\left\{\frac{k_0}{2z_0}[(Mx - x_0)^2 + (My)^2]\right\} \\
 & + \sin\left\{\frac{k_0}{2(z_0 + \Delta z_0)}[(Mx)^2 + (My)^2]\right\}. \quad (4.2-3)
 \end{aligned}$$

When  $M < 1$ , we have magnification, whereas  $M > 1$  corresponds to demagnification. By re-writing Eq. (4.2-3), we have

$$\begin{aligned}
 H_{3-p}(Mx, My) = & \sin\left[\frac{k_0}{2z_0/M^2}(x^2 + y^2)\right] \\
 & + \sin\left\{\frac{k_0}{2z_0/M^2}[(x - x_0/M)^2 + y^2]\right\} + \sin\left\{\frac{k_0}{2(z_0 + \Delta z_0)/M^2}(x^2 + y^2)\right\}. \quad (4.2-4)
 \end{aligned}$$

Now, during optical reconstruction using a wavelength  $\lambda_0$ , we see that by inspection of the first and second term in Eq. (4.2-4), the two real image points are now formed at a distance of  $z_0/M^2$  away from the hologram, and with a reconstructed lateral distance of  $x_0/M$  away from each other. By defining the *lateral magnification*,  $M_{lat}$ , as the ratio of the reconstructed lateral distance to the original lateral distance,  $x_0$ , we have  $M_{lat} = 1/M$ . In order to determine the magnification along the longitudinal direction, we must focus on the first and the third terms, and upon reconstruction, we see that the two points are reconstructed at  $z_0/M^2$  and  $(z_0 + \Delta z_0)/M^2$ , respectively. By defining the *longitudinal magnification*,  $M_{long}$ , as the ratio of the reconstructed longitudinal distance,  $\Delta z_0/M^2$ , to the original longitudinal distance,  $\Delta z_0$ , we have  $M_{long} = 1/M^2$ .

*b) Wavelength Scaling*

We could reconstruct the hologram with a different wavelength, say  $m\lambda_0$  (or  $k_0/m$ ), where  $m$  is a constant. Hence, according to Fresnel diffraction, the field distribution at  $z$  away from the hologram is now given by

$$\begin{aligned}
 H_{3-p}(x,y) * h(x,y; z, k_0/m) \\
 \propto \left\{ \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right] + \sin\left\{\frac{k_0}{2z_0}[(x - x_0)^2 + y^2]\right\} \right. \\
 \left. + \sin\left[\frac{k_0}{2(z_0 + \Delta z_0)}(x^2 + y^2)\right] \right\} * \frac{jk_0/m}{2\pi z} \exp\left[-\frac{jk_0/m}{2z}(x^2 + y^2)\right].
 \end{aligned}
 \tag{4.2-5}$$

This equation suggests that there will be no magnification in the lateral direction. Along the longitudinal direction, we can inspect the results of the first and the third terms. And again when we consider real image reconstruction, the first and the second term will form an image at  $z = z_0/m$ . The third term gives rise to a real image at  $z = (z_0 + \Delta z_0)/m$ . Hence, in this case  $M_{\text{lat}} = 1$ , and  $M_{\text{long}} = 1/m$ . The reconstructed volume is either compressed or expanded by a factor of  $1/m$  with the same lateral magnification. When  $m > 1$ , we have compression. With visible light for recording and reconstruction,  $m$  is in the range of 0.5 to 1.8. However, when using digital reconstruction,  $m$  can be arbitrarily chosen.

*c) Reconstruction combining hologram scaling and wavelength scaling*

If we change the scale of the hologram and use a different wavelength for reconstruction, then the combined magnification along the lateral and longitudinal direction will be  $M_{\text{lat}} = 1/M$  and  $M_{\text{long}} = 1/mM^2$ . Therefore, we see that the reconstructed volume,  $x_0\Delta z_0/mM^3$ , is different from the original volume  $x_0\Delta z_0$ . This creates distortion when we magnify the original 3-D object. This is a well-known result of magnification in 3-D optical imaging. In order to have a true 3-D perspective on reconstruction, when the scale change of the hologram is given by  $M$ , we let  $m = 1/M$  which gives us  $M_{\text{lat}} = m$  and  $M_{\text{long}} = m$ , and therefore, we obtain  $M_{\text{lat}} = M_{\text{long}}$ . In other words, in order to prevent distortion in 3-D imaging, we scale the hologram by a factor  $M$ , and then the reconstructing wavelength should be  $m\lambda_0$ , where  $\lambda_0$  is the recording wavelength and  $m$  is equal to  $1/M$ . This is the original idea of Gabor [1949] who first conceptualized the notion during the pre-laser era to improve upon the electron microscope. Electron microscopy was the motivational factor in the development of holography.



In order to address the prospect of a true 3-D holographic TV, we will consider some issues of 3-D holographic display.

#### A. Spatial frequency resolution issue

Let us first delve into the spatial frequency resolution of an SLM for 3-D display. For simplicity, we will take a point source hologram as our hologram displayed on a SLM. From previous chapters, we know that the expression of such a hologram is given by  $\sin[\frac{k_0}{2z_0}(x^2 + y^2)]$ . Remember that  $z_0$  is the distance of the point source away from the recording device. The local spatial frequency along the  $x$ -direction across the hologram has been given by Eq. (2.5-5), and is defined as

$$f_{local} = \frac{1}{2\pi} \frac{d}{dx} \left( \frac{k_0}{2z_0} x^2 \right) = \frac{x}{\lambda_0 z_0}. \quad (4.2-6)$$

If the size of the limiting aperture of the hologram is  $x_{max}$ , then  $f_{local}$  at  $x_{max}$  is

$$f_{max} = \frac{x_{max}}{\lambda_0 z_0}, \quad (4.2-7)$$

which is the highest spatial frequency of the hologram fringes. Now assume that the SLM has a maximum spatial resolution of  $f_0$ , and if we want to record  $f_{max}$ , then we must obey the requirement of  $f_0 = f_{max}$ . Now, according to the geometry shown in Fig. 4.14, the  $NA$  of the hologram is

$$\sin(\theta/2) = x_{max}/z_0, \quad (4.2-8)$$

where  $\theta$  is defined as the viewing angle. By using Eq. (4.2-7), Eq. (4.2-8) becomes

$$NA = \sin(\theta/2) = \lambda_0 f_{max} = \lambda_0 f_0. \quad (4.2-9)$$

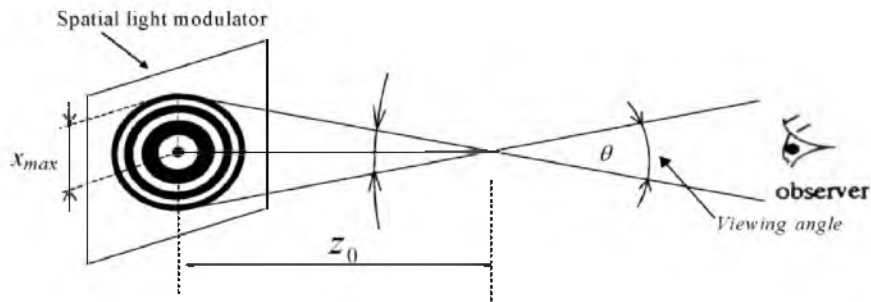


Fig. 4.14 Viewing angle.

For any given spatial resolution of the SLM, we can therefore find the viewing angle according to Eq. (4.2-9). For example, Hamamatsu's EBSLM has a spatial resolution of about  $f_0 = 8 \text{ lp/mm}$ , which gives us a viewing angle of about  $0.6^\circ$  at  $\lambda_0 = 0.6328 \mu\text{m}$ . Hence, such a device is not useful for applications in 3-D display. However, if we desire to obtain a sequential 2-D display along the depth, then the EBSLM system is adequate enough because it is capable of updating holograms at a video rate. Table 4.1 shows the viewing angle for some existing SLMs. We currently do not have SLMs that are suitable for 3-D display as their viewing angles are severely limited. The situation becomes even worse if we use off-axis holography because we will need to resolve the carrier frequency. As we recall, for 45 degrees of the recording angle, the carrier frequency is about 1,000 lp/mm, which is well beyond the capability of existing SLMs (see Table 4.1).

Table 4.1 Viewing angles for  $\lambda_0 = 0.6328 \mu\text{m}$ .

$f_0$ (SLM's resolution)	$\theta$ (viewing angle)	device/company
8 lp/mm	0.6 degree	EBSLM/Hamamatsu
100 lp/mm	6.8 degrees	PALSLM/Hamamatsu
500 lp/mm	34 degrees	not available

### B. Spatial resolution issue

Let us calculate the number of samples that are required for a hologram to be displayed by an SLM. For a given spatial frequency resolution of an SLM,  $f_0$ , and according to the *Nyquist sampling*, the minimum sampling frequency,  $f_s$ , required in order to generate the hologram is

$$f_s = 2f_0.$$

The number of samples,  $N$ , that will then be required to create a hologram of size  $l \times l$  is

$$N = (lf_s)^2 = (l \times 2f_0)^2 = (2l \frac{NA}{\lambda_0})^2, \quad (4.2-10)$$

where we have used Eq. (4.2-9). According to Eq. (4.2-10), for full parallax,  $20\text{mm} \times 20\text{mm}$  on-axis hologram to be presented on an SLM, with a viewing angle of  $60^\circ$ , the number of resolvable pixels required is about 1.1 billion. To put things into perspective, some of the best CCD cameras, such as Canon D60 (3072x2048 pixels, 67.7 lp/mm,  $7.4 \mu\text{m}$  pixel size), have just a little over 6 Mega pixels.

### C. Data transmission issue

As we had previously calculated, a single frame of a  $20\text{mm} \times 20\text{mm}$  hologram with a viewing angle of  $60^\circ$  requires about 1.1 billion pixels on the SLM. To update such a frame with an 8-bit resolution at 30 frames/s, a serial data rate of

$$1.1 \text{ billion samples/frame} \times 8 \text{ bits/sample} \times 30 \text{ frames/s} = 0.26 \text{ Tbit/s}$$

is required for full parallax.

Basically, all the issues that we have discussed illustrate the fact that content of information held within a hologram is enormous. This implies that the content of information in the hologram must be significantly reduced in order to achieve 3-D holographic TV for 3-D display. Live 3-D TV with holographic images is truly a formidable problem. However, since we are used to looking at the world with our two eyes more or less on a horizontal level, we are usually satisfied with horizontal parallax. Hence, for 512 vertical lines, the number of pixels required becomes  $512 \times (2l \times 2NA/\lambda_0)$ , which is approximately 17 million, if we are to eliminate vertical parallax. By sacrificing vertical parallax, the data rate becomes 4 Gbits per second instead of 0.26 Tbits per second, which is calculated for full parallax. And this is manageable with advanced modern optical communications systems. By using fiber optics, data rates of up to 40 Gbit/s indeed can be achieved in real-world applications. The possibility of real-time holographic TV becomes a reality if the horizontal-parallax-only (HPO)-electronic holographic recording technique becomes available. Indeed, by using computer-generated, horizontal-parallax-only holographic information, the MIT group has demonstrated a 3-D holographic display having 64 vertical lines and with viewing angle of about 15 degrees [St. Hilaire, Benton, Lucente, Jepsen, Kollin, Yoshikawa, Underkoffler (1990), St. Hilaire, Benton, and Lucente (1992)]. However, this HPO-holographic information is computer generated and no HPO-holographic information has actually been generated by or recorded from actual real objects.

Optical scanning holography with horizontal-parallax-only recording is possible if we scan the object with a 1-D TDFZP. This idea, called *HPO-optical scanning holography*, has been proposed recently while computer simulations have been performed [Poon, Akin, Indebetouw, and Kim (2005), Poon (2006)].

## 4.3 Optical Scanning Cryptography

Due to the recent progress in the development of optical components and the increased technical performance of optical systems, optical cryptography has

a significant potential in advances for security applications. Indeed there has been a plethora of articles that deal with secure systems that use optical methods [Lohmann, Stork, and Stucke (1986), Refregier and Javidi (1995), Lai and Neifeld (2000), Wang, Sun, Su, and Chiou (2000), Magensen and Gluckstad (2001)]. One of the reasons for using optical encryption is that information, such as images, that needs to be encrypted exists already in the optical domain. Another reason is that optical encryption, as opposed to electronic or digital encryption, can provide many degrees of freedom when securing sensitive information. When large volumes of information need to be encrypted, such as a 3-D object, using optical encryption methods is probably the most logical choice. Although most optical encryption techniques are typically coherent, some incoherent optical techniques for encryption have recently been proposed [Tajahuerce, Lancis, Javidi, and Andres (2001)]. In general, incoherent optical techniques have many advantages over their coherent counterparts. This includes a better S/N ratio and insensitivity to the misalignment of optical elements. In this section, we will discuss an incoherent optical method based on optical scanning holography for encryption. This method is called *optical scanning cryptography (OSC)* [Poon, Kim, and Doh (2003)]. While having the capability to utilize incoherent processing, the method also has many other advantages. These advantages include the following. 1) Since it is an optical scanning method, it can process incoherent objects, such as printed documents, without using a spatial light modulator (SLM) to convert an incoherent image into a coherent image as existing coherent techniques currently do. The proposed system can indeed perform real-time or *on-the-fly encryption*. 2) Since the output signal is a heterodyne electrical signal, and hence the encrypted information is riding on a heterodyne frequency (or a carrier frequency as used in communications), it can immediately be radiated for wireless transmission to a secure site for storage and then subsequently be encrypted. This may have important applications in *radio frequency identification (RFID)* [Radio Frequency Identification Technologies: A Workshop Summary (2004)]. 3) Because the technique is based on holography, it can be easily extended to the use of encrypting 3-D information.

Figure 4.15 shows the optical system utilized for encryption and decryption. The system contains two subsystems: an encryption stage and a decryption stage. It is noted that the two subsystems have an identical two-pupil optical heterodyne scanning image processor, which we have extensively studied in section 3.4. We will first briefly summarize the previous results of the image processor, and then we will discuss encryption and decryption.

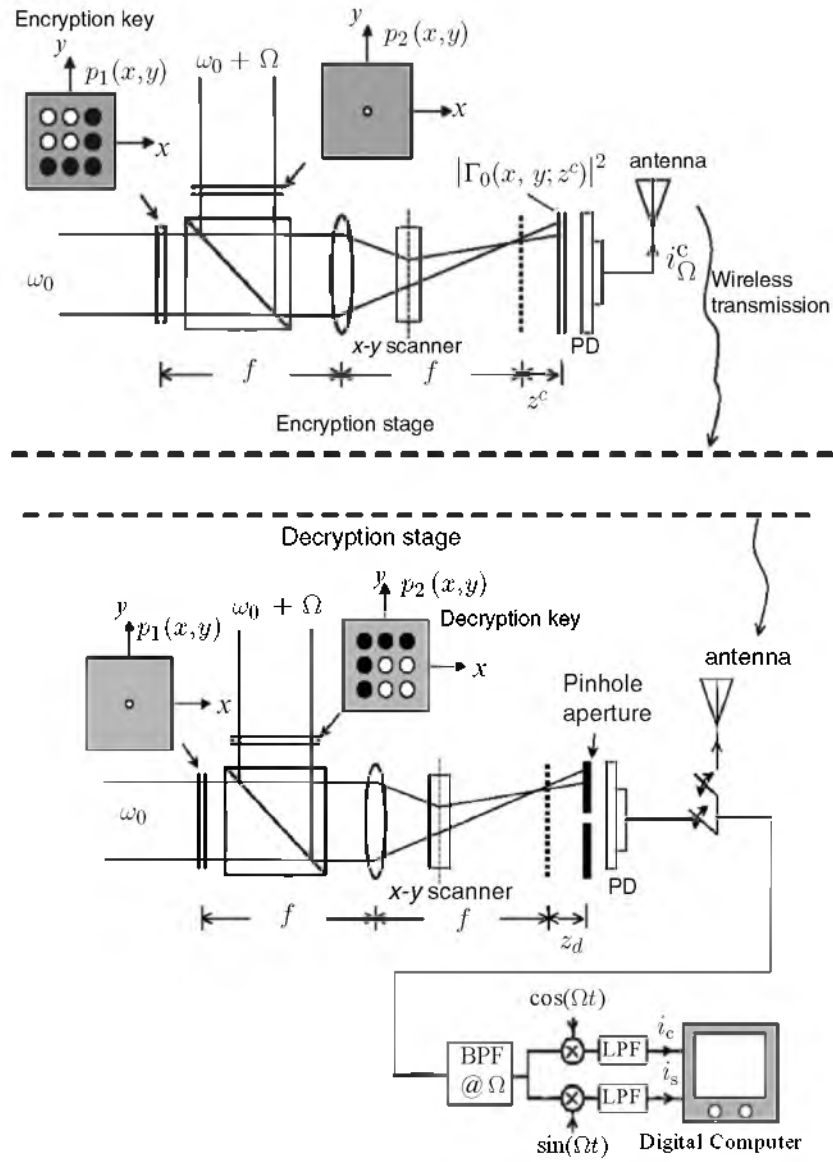


Fig. 4.15 Optical scanning cryptography.  
Adapted from T.-C. Poon et al., Applied Optics 42, 496 (2003).

If we concentrate on the optical system of the encryption stage, we can see that the two pupils,  $p_1(x, y)$  and  $p_2(x, y)$ , are located at the front focal plane of the lens and are illuminated by two broad laser beams of temporal frequencies,  $\omega_0$  and  $\omega_0 + \Omega$ , respectively. The two beams are then combined by a beamsplitter and used to 2-D scan a planar object,  $|\Gamma_0(x, y; z^c)|^2$ , located at a distance  $z^c$  away from the back focal plane of the lens. The distance of  $z^c$  is called a *coding distance*, and  $|\Gamma_0(x, y; z^c)|^2$  is the object to be encrypted. The photodetector, PD, collects all the light transmitted by the object if the object is transparent (or collects all the scattered light if the object is diffusely reflecting). The photodetector will have a heterodyning current at a frequency of  $\Omega$  as one of its outputs (where the other output is a baseband signal). After electronic tuning at  $\Omega$ , the heterodyne current,  $i_\Omega^c(x, y; z^c)$ , is given by Eq. (3.4-11) where  $z$  is replaced by the coding distance  $z^c$ ,

$$\begin{aligned} i_\Omega^c(x, y; z^c) &= \text{Re}[i_{\Omega_p}(x, y; z^c)\exp(j\Omega t)] \\ &= \text{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\}OTF_\Omega(k_x, k_y; z^c)\}\exp(j\Omega t)]. \end{aligned} \quad (4.3-1)$$

Again  $OTF_\Omega(k_x, k_y; z^c)$  is called the optical transfer function (OTF) of the heterodyne scanning system, which has been given by Eq. (3.4-10) as follows:

$$\begin{aligned} OTF_\Omega(k_x, k_y; z^c) &= \exp\left[j\frac{z^c}{2k_0}(k_x^2 + k_y^2)\right] \\ &\times \int \int p_1^*(x', y') p_2\left(x' + \frac{f}{k_0}k_x, y' + \frac{f}{k_0}k_y\right) \exp\left[j\frac{z^c}{f}(x'k_x + y'k_y)\right] dx' dy', \end{aligned} \quad (4.3-2)$$

where  $f$  is the focal length of the lens shown in the encryption stage in Fig. 4.15. The processing elements are the two pupil functions,  $p_1(x, y)$  and  $p_2(x, y)$ .  $i_\Omega^c(x, y; z^c)$  is the scanned and processed version of the input,  $|\Gamma_0(x, y; z^c)|^2$ . By manipulating the pupils, we will have a different processed output because the OTF in Eq. (4.3-2) is expressed in terms of the two pupils. Now the processed information is carried by a temporal carrier at a frequency of  $\Omega$ , and if  $\Omega$  is chosen to be in the radio frequency domain (which can be done easily through the use of acousto-optic modulators), then the processed information can be readily radiated to a secure site (or a decryption site) for further processing. The situation is shown in Fig. 4.15. After receiving the decrypted information from an antenna at the secure site (the output of the antenna in the secure site is switched to the input of the

bandpass filter for electronic processing), the information is further processed electronically as shown in Fig. 4.15. In other words, by multiplying the incoming signal by  $\cos(\Omega t)$  and  $\sin(\Omega t)$ , and through lowpass filtering, we obtain two signals,  $i_c$  and  $i_s$ , respectively, which have been given by Eqs. (3.4-14a) and (3.4-14b) as follows:

$$i_c(x, y; z^c) = \text{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\}OTF_\Omega(k_x, k_y; z^c)\}] \quad (4.3-3a)$$

and

$$i_s(x, y; z^c) = \text{Im}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\}OTF_\Omega(k_x, k_y; z^c)\}]. \quad (4.3-3b)$$

If we now apply addition to the above expressions in the following manner:  $i(x, y; z^c) = i_c(x, y; z^c) + j i_s(x, y; z^c)$ , we have a complex expression where the full amplitude and phase information of the processed object are available:

$$i(x, y; z^c) = \mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\}OTF_\Omega(k_x, k_y; z^c)\}. \quad (4.3-4)$$

### Encryption

From Eq. (4.3-4), we will now discuss encryption. To perform encryption on the input object,  $|\Gamma_0(x, y; z^c)|^2$ , located at a distance of  $z^c$  away from the back focal plane of the lens at the encryption stage, we generally can manipulate the two pupils,  $p_1(x, y)$  and  $p_2(x, y)$ . As a simple example, we will let  $p_2(x, y) = \delta(x, y)$ , a pin hole, and keep  $p_1(x, y)$  as is. The situation is shown in Fig. 4.15. We shall call  $p_1(x, y)$  an *encryption key*. Under these conditions, according to Eq. (4.3-2), the OTF of the system becomes

$$OTF_\Omega(k_x, k_y; z^c) = \exp\left[-j\frac{z^c}{2k_0}(k_x^2 + k_y^2)\right]p_1^*\left(\frac{-f}{k_0}k_x, \frac{-f}{k_0}k_y\right), \quad (4.3-5)$$

and Eq. (4.3-4) then becomes

$$\begin{aligned} i(x, y; z^c) &= \mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\} \\ &\times \exp\left[-j\frac{z^c}{2k_0}(k_x^2 + k_y^2)\right] \times p_1^*\left(\frac{-f}{k_0}k_x, \frac{-f}{k_0}k_y\right)\}. \end{aligned} \quad (4.3-6)$$

$i(x, y; z^c)$  is the coded or encrypted object and can be stored by the digital computer. Note that the spectrum of  $|\Gamma_0(x, y; z^c)|^2$  is multiplied by two terms. Since the product of the object's spectrum with the term

$\exp \left[ -j \frac{z^c}{2k_0} (k_x^2 + k_y^2) \right]$  corresponds to the spectrum of the hologram of  $|\Gamma_0(x, y; z^c)|^2$  [see Eq. (3.5-9)], where the object is recorded at a distance of  $z^c$  away from the focal plane of the lens, we can interpret Eq. (4.3-6) as the holographic information (or hologram) of the object that is encrypted or coded by  $p_1$ , i.e., we “encrypt the hologram of the object.” This idea of coding holographic information was first investigated by Schilling and Poon in the context of optical scanning holography [Schilling and Poon (1995)].

### Decryption

After the object has been coded or encrypted, it will be necessary to decode or decrypt it. To do this, we turn to the optical system at the secure site. Again, note that the optical system is the same except for the choice of the selected pupils, and the laser beams are now scanning a pin hole as an object, i.e.,  $|\Gamma_0(x, y; z^d)|^2 = \delta(x, y; z^d)$ , located at a distance  $z^d$  away from the back focal plane of the lens at the decryption stage. We shall call  $z^d$  the *decoding distance*.

However, this time the switch, as shown in Fig. 4.15, is connected to the output of the optical system at the secure site. Through electronic processing, the output of the photodetector will then be processed. The result of Eq. (4.3-4) can be applied again, but by replacing  $z^c$  with by  $z^d$ . Now we choose  $p_1(x, y) = \delta(x, y)$ , a pin hole, and keep  $p_2(x, y)$  as is. We shall call  $p_2(x, y)$  a *decryption key*. According to Eq. (4.3-2), this selection of the pupils gives the following OTF,

$$OTF_{\Omega}(k_x, k_y; z^d) = \exp \left[ j \frac{z^d}{2k_0} (k_x^2 + k_y^2) \right] p_2 \left( \frac{f}{k_0} k_x, \frac{f}{k_0} k_y \right). \quad (4.3-7)$$

By using Eq. (4.3-4) and the fact that  $\mathcal{F}\{|\Gamma_0(x, y; z^d)|^2\} = 1$ , we have

$$i(x, y; z^d) = \mathcal{F}^{-1} \left\{ \exp \left[ j \frac{z^d}{2k_0} (k_x^2 + k_y^2) \right] p_2 \left( \frac{f}{k_0} k_x, \frac{f}{k_0} k_y \right) \right\}. \quad (4.3-8)$$

This is the output generated at the decryption stage, where the decryption key has been inserted into the stage and a pin hole has been scanned. The information is now stored in the digital computer to be used to later decrypt the information coming from the encryption site via a wireless transmission. To decrypt the information represented by Eq. (4.3-6), a digital decryption unit (DDU), shown in Fig. 4.16, has been proposed.

Again,  $i(x, y; z^c)$  is the transmitted encrypted information transmitted via wireless from the encryption stage, and  $i(x, y; z^d)$  is the information generated at the decryption site. We see that by using Eq. (4.3-6)



and (4.3-8), at the output of the unit we have

$$\begin{aligned}
 \text{output of DDU} &\propto \mathcal{F}^{-1}\{\mathcal{F}\{i(x, y; z^c)\} \times \mathcal{F}\{i(x, y; z^d)\}\} \\
 &= \mathcal{F}^{-1}\left\{\mathcal{F}\{|\Gamma_0(x, y; z^c)|^2\} \exp\left[-j \frac{z^c}{2k_0}(k_x^2 + k_y^2)\right] \right. \\
 &\quad \times p_1^*\left(\frac{-f}{k_0}k_x, \frac{-f}{k_0}k_y\right) \exp\left[j \frac{z^d}{2k_0}(k_x^2 + k_y^2)\right] p_2\left(\frac{f}{k_0}k_x, \frac{f}{k_0}k_y\right)\} \\
 &= |\Gamma_0(x, y; z^c)|^2
 \end{aligned} \tag{4.3-9}$$

if the following conditions are simultaneously met. 1)  $z^d = z^c$ , i.e., the coding distance in the encryption stage and the decoding distance in the decryption stage are the same, and 2)  $p_1^*(-x, -y)p_2(x, y) = 1$ . Condition (1) simply means that the holographic reconstruction is in focus if Condition (2) has already been met. For any values of  $z^d \neq z^c$ , we have what is known as defocused image reconstruction. Condition (2) allows us to choose the functional form of the encryption key,  $p_1(x, y)$ , in the encryption stage and the decryption key,  $p_2(x, y)$ , at the decryption site. As a simple example, the choice of phase keys works well. We shall demonstrate this in the following example.

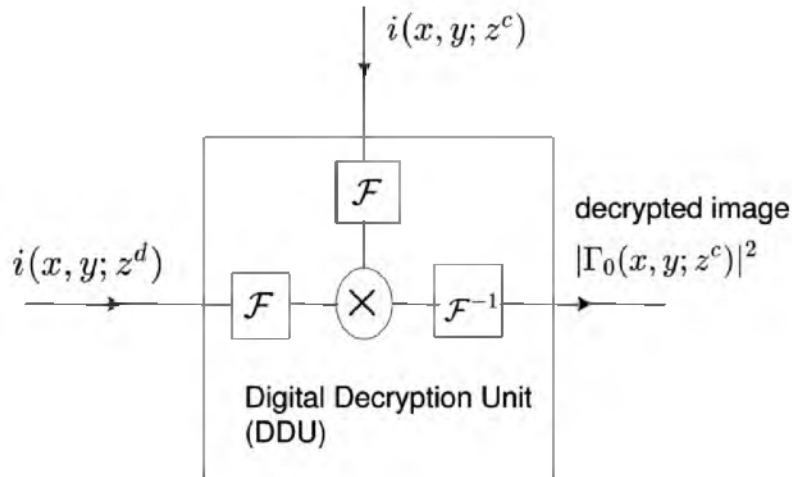


Fig. 4.16 Digital Decryption Unit (DDU) :  $i(x, y; z^c)$  is the encrypted information with encryption key,  $p_1(x, y)$ , inserted into the encryption stage, which is sent from the encryption site via wireless transmission.  $i(x, y; z^d)$  is the signal generated at the decryption site where the decryption key,  $p_2(x, y)$ , is inserted into the scanning stage to scan a pin hole aperture. Adapted from T.-C. Poon et al., Applied Optics 42, 496 (2003).

### Example 4.2 MATLAB Example on Optical Scanning Cryptography

As a simple example, we can immediately see that the choice of a random phase mask is a good encryption key, i.e., we let  $p_1(x, y) = \exp[j2\pi M(x, y)]$  where  $M(x, y)$  is a function of random numbers chosen from a uniform distribution between the interval (0.0,1.0). Again,  $p_2(x, y) = \delta(x, y)$ , i.e., a pin-hole is the other pupil in the encryption stage. For this choice of pupils, the encrypted image from Eq. (4.3-6) becomes

$$i(x, y; z^c) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ |\Gamma_0(x, y; z^c)|^2 \} \exp \left[ -j \frac{z^c}{2k_0} (k_x^2 + k_y^2) \right] \right. \\ \left. \times \exp \left[ -j2\pi M \left( \frac{-fk_x}{k_0}, \frac{-fk_y}{k_0} \right) \right] \right\}. \quad (4.3-10)$$

The above encrypted information can be made more secure if, for example, the original document is multiplied by a random phase mask,  $\exp[j2\pi r(x, y)]$ , where  $r(x, y)$  is a function of random numbers. By using Eq. (4.3-10), the overall encrypted image then becomes

$$i(x, y; z^c) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ |\Gamma_0(x, y; z^c)|^2 \exp[j2\pi r(x, y)] \} \right. \\ \left. \times \exp \left[ -j \frac{z^c}{2k_0} (k_x^2 + k_y^2) \right] \exp \left[ -j2\pi M \left( \frac{-fk_x}{k_0}, \frac{-fk_y}{k_0} \right) \right] \right\}. \quad (4.3-11)$$

This technique used to obtain the above resulting encrypted image is called *double-random phase encoding* [Refregier and Javidi (1995)]. We shall use MATLAB in order to simulate this coding. Note that  $r(x, y)$  and  $M(x, y)$  should be chosen as two independent random functions.

Table 4.2 contains the m-file for the simulations shown in this example. Figure 4.17a) shows the original document,  $|\Gamma_0(x, y; z^c)|^2$ . Figure 4.17b) and 4.17c) show the real part and imaginary part of the original document multiplied by a random phase mask,  $\exp[j2\pi r(x, y)]$ , placed immediately in front of the original document. Figure 4.17d) shows the “intensity” of the encrypted document,  $|i(x, y; z^c)|$ , which is calculated using Eq. (4.3-11) where we have used  $\text{sigma} = (zc * \text{ld}) / (4 * \text{pi}) = (30) * (0.6 * 10^{-6}) / 4\pi$  in the m-file [sigma is  $z^c / 2k_0$  in Eq. (4.3-11)].



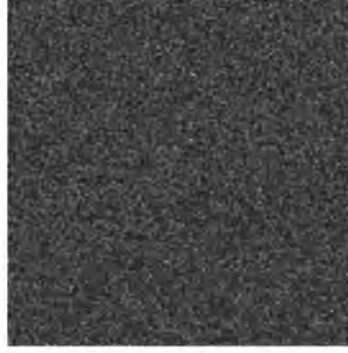
a) Original document.



b) Real part of the image multiplied by random phase mask.



c) Imaginary part of the image multiplied by random phase mask.



d) Intensity of encrypted document.

Fig. 4.17 Cryptography simulations.

For decryption, we need to gather information in the decryption stage by scanning a pin-hole object located at a distance of  $z^d$  away from the back focal plane of the lens, where the pupils are  $p_1(x, y) = \delta(x, y)$  (a pin hole) and  $p_2(x, y) = \exp[j2\pi M(-x, -y)]$ , which satisfies Condition (2) as previously discussed. According to Eq. (4.3-8), the scanned output to be stored in the digital computer then becomes

$$i(x, y; z^d) = \mathcal{F}^{-1} \left\{ \exp \left[ j \frac{z^d}{2k_0} (k_x^2 + k_y^2) \right] \exp \left[ j2\pi M \left( \frac{-fk_x}{k_0}, \frac{-fk_y}{k_0} \right) \right] \right\}. \quad (4.3-12)$$

According to Fig. 4.16, when the information from Eqs. (4.3-11) and (4.3-12) are the inputs of the DDU, the output is as follows:

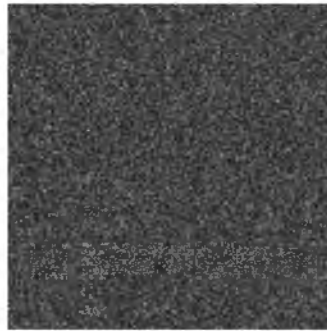
$$\begin{aligned} \text{output of DDU} &\propto \mathcal{F}^{-1} \left\{ \mathcal{F} \{ |\Gamma_0(x, y; z^c)|^2 \exp[j2\pi r(x, y)] \} \right. \\ &\quad \times \exp \left[ -j \frac{z^c}{2k_0} (k_x^2 + k_y^2) \right] \exp \left[ j \frac{z^d}{2k_0} (k_x^2 + k_y^2) \right] \left. \right\} \\ &= |\Gamma_0(x, y; z^c)|^2 \exp[j2\pi r(x, y)], \end{aligned} \quad (4.3-12)$$

when the sigma used in the m-file is the same as that used for encryption, i.e., theoretically  $z^d = z^c$ . The decrypted output intensity, i.e., the absolute value of the DDU's output, is shown in Fig. 4.18a). If the decryption key is chosen incorrectly (such as the guessing of a random phase mask), Fig. 4.18b) shows its unusable output intensity. Finally, when the decryption key is used correctly and if sigma or  $z^d$  is guessed or chosen incorrectly, say,  $z^d = 1.5 \times z^c$ , then the absolute value of the DDU's output is shown in Fig.

4.18c), which is a defocused version of Fig. 4.18a). We see that the introduction of  $z^c$  and  $z^d$  gives an extra security measure.



a) Intensity of decrypted document with matched key and  $z^d = z^c$



b) Intensity of decrypted document with mismatched key and  $z^d = z^c$



c) Intensity of decrypted document with matched key but with  $z^d = 1.5z^c$

Fig. 4.18 Decryption simulations.

Table 4.2 Cryptography.m:  
m-file for simulating optical scanning cryptography.

```

%Cryptography.m
%Simulation of Optical Encryption and Decryption
%This program was adapted from the one developed by Taegeun Kim
%of Sejong Univ., Korea

clear
%L : Normalized length of back ground (field of view)
L=1;
%Dl: Physical field of view in this simulation 20% of L
Dl=0.02;
%N : sampling number
N=255;
% dx : step size
dx=L/N;

%Unit Axis Scaling
%Normalized length and Spatial Frequency according to the Normalized length
for k=1:256
    X(k)=1/255*(k-1)-L/2;
    Y(k)=1/255*(k-1)-L/2;

    %Kx=(2*pi*k)/(N*dx)
    %k is sampling number, N is number of sample,
    %in our case, N=255, dx=1/255(unit length)

    Kx(k)=(2*pi*(k-1))/(N*dx)-((2*pi*(256-1))/(N*dx))/2;
    Ky(k)=(2*pi*(k-1))/(N*dx)-((2*pi*(256-1))/(N*dx))/2;
end
%Real length and real spatial frequency
X=Dl*X;
Y=Dl*Y;
Kx=Kx./Dl;
Ky=Ky./Dl;

%Read Input image, image size must be 575x577x3
%for the program to function properly
CH1=imread('vt.bmp','bmp');
CH1=CH1(:,:,1);
[x,y]=meshgrid(1:577,1:575);
[xi,yi]=meshgrid(1: 2.2539:577,1:2.2461:575);
CH1p=double(CH1);
I0=interp2(x,y,CH1p,xi,yi);
I0=double(I0);
I0=I0./max(max(I0)); %Image to be encrypted

M=rand(256);
M2=rand(256);
%Encryption key in frequency domain,

```

```

%the last term in Eq. (4.3-10)
P=exp(-j*2*pi*M2);

RPM=exp(j*2*pi*M); %random phase mask, exp(j*2*pi*r(x,y))
R1=I0.*RPM;%Random phase mask times the image

%OTF(kx,ky,zc)
%sigma=z/(2Ko)=(z*ld)/(4*pi)
%where Ko is the wave number, z is the distance from the source
%and ld is the wavelength of the source
ld=0.6*10^-6; % wavelength=ld=0.6*10^-6
zc=0.3; %coding distance
sigma=(zc*ld)/(4*pi);

for r=1:256,
    for c=1:256,
        OTF(r,c)=exp(-j*sigma*(Kx(r).^2+Ky(c).^2));
    end
end

for r=1:256,
    for c=1:256,
        OTF2(r,c)=exp(-j*1.5*sigma*(Kx(r).^2+Ky(c).^2));
    end
end

%Fourier transformation
FR=(1/256)^2*fft2(R1);
FR=fftshift(FR);
Ho=FR.*OTF;

%Encrypted image in the frequency domain
E=Ho.*P;

%Encrypted image in the space domain
e=ifft2(E); %Eq. (4.3-10)

%Key info for decryption key is achieved by scanning the pin hole that is
%located at z=zc
Key_info=conj(OTF.*P); % Fourier transform of Eq. (4.3-11), zd=zc
Key_info2=conj(OTF2.*P); % Fourier transform of Eq. (4.3-11), zd=1.5zc
%Different random phase
M3=rand(256);
P1=exp(j*2*pi*M3);
Key_info_mis=conj(OTF.*P1);% Fourier transform of Eq. (4.3-11) but with a wrong
phase key

%Decrypted image with matched key in the frequency domain
De=E.*Key_info;

%Decrypted image with matched key in the frequency domain
%but with twice distance of zc

```

```

De2=E.*Key_info2;

%Decrypted image with matched key in the space domain
de=ifft2(De); %Eq. (4.3-12)

%Decrypted image with matched key in the space domain
%but with zd=1.5 zc
de2=ifft2(De2); %Eq. (4.3-12)

%Decrypted image with mis_matched key in the frequency domain
De_mis=E.*Key_info_mis;

%Decrypted image with mis_matched key in the space domain
de_mis=ifft2(De_mis);

figure(1)
image(X,Y,256*I0);
colormap(gray(256));
axis square
title('image to be encrypted')
axis off

figure(2)
image(X,Y,255*real(R1));
colormap(gray(256));
axis square
title('Real part of the image multiplied by random phase mask')
axis off

figure(3)
image(X,Y,255*imag(R1));
colormap(gray(256));
axis square
title('Imaginary part of the image multiplied by random phase mask')
axis off

figure(4)
image(X,Y,255*abs(e)/max(max(abs(e))))
colormap(gray(256))
axis square
title('Intensity of encrypted image')% absolute value of Eq. (4.3-10)
axis off

figure(5)
image(X,Y,255*abs(de)/max(max(abs(de))));
colormap(gray(256));
axis square
title('Intensity of decrypted image with matched key with zd=zc')
axis off

figure(6)
image(X,Y,255*abs(de_mis)/max(max(abs(de_mis))));

```



```

colormap(gray(256));
axis square
title('Intensity of decrypted image with mismatched key with zd=zc')
axis off

figure(7)
image(X,Y,255*abs(de2)/max(max(abs(de2))));
colormap(gray(256));
axis square
title('Intensity of decrypted image with matched key with zd=1.5zc')
axis off
-----

```

## References

- 4.1 Agard, D. A. (1984). "Optical sectioning microscopy: cellular architecture in three dimensions," *Ann. Rev. Biophys. Bioeng.* 13, 191-219.
- 4.2 Benton, S.A. (1991). "Experiments in holographic video," *Proc. SPIE* IS-08, 247-267.
- 4.3 Brown, H. B., S.C. Noble and B.V. Markevitch (1983). Three-dimensional television system using holographic technique, *U.S. Patent* # 4,376,950.
- 4.4 Corle, T. R. and G. S. Kino (1996). *Confocal Scanning Optical Microscopy and Related Imaging Systems*, Academic Press, San Diego, CA.
- 4.5 Enloe, L. H., J.A. Murphy, and C. B. Rubinstein (1966). "Hologram transmission via television," *Bell Syst. Techn. J.* 45, 333-335.
- 4.6 Gabor, D. (1949). "Microscopy by reconstructed wavefronts," *Proc. Roy. Soc., ser. A*, 197, 454-487.
- 4.7 Hamamatsu Photonics K.K., Japan, and Hamamatsu Corp., Bridgewater, NJ. Image Processing and Measuring System (IPMS), Model DVS-3010/SS.
- 4.8 Hamamatsu Photonics K.K., Japan, and Hamamatsu Corp., Bridgewater, NJ. Product information sheet for EBSLM model X3636.
- 4.9 Huang, D., E.A. Swanson, C. P. Lin, J.S. Shuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, C.A. Puliafito, and J. G. Fujimoto (1991). "Optical coherent tomography," *Science*, Vol. 254, 1178-1181.
- 4.10 Indebetouw, G. (2002). "Properties of a scanning holographic microscope: improved resolution, extended depth of focus, and/or optical sectioning," *Journal of Modern Optics* 49, 1479-1500.
- 4.11 Indebetouw, G., T. Kim, T.-C. Poon, and B. Schilling (1998). "Three-dimensional location of fluorescent inhomogeneities in turbid media using scanning heterodyne holography," *Optics Letters* 23, 135-137.
- 4.12 Indebetouw, G. and W. Zhong (2006) "Scanning holographic microscopy of three-dimensional fluorescent specimens," *Journal of the Optical Society of America A* 23, 1699-1707.
- 4.13 Kim, M (2000) "Tomographic three-dimensional imaging of a biological specimen using wave-scanning digital interference holography," *Optics Express* 7, 305-310.
- 4.14 Kim, T and T.-C. Poon (1999). "Extraction of 3-D location of matched 3-D object using power fringe-adjusted filtering and Wigner analysis," *Optical Engineering* 38, 2176-2183.
- 4.15 Kirk, R. L. (1984). Electronically generated holography, *International Patent* No. WO 84/00070.

- 4.16 Lai, S. and M. A. Neifeld (2000). "Digital wavefront reconstruction and its applications to image encryption," *Optics Communications* 178, 283-289.
- 4.17 Lohmann, A.W., W. Stork, and G. Stucke (1986), "Optical perfect shuffle," *Applied Optics* 25, 1530-1531.
- 4.18 Macovski, A. (1971). "Considerations of television holography," *Optica Acta*, 18, 31-39.
- 4.19 Magensen, P.C. and J. Gluckstad (2001). "Phase-only optical decryption of a fixed mask," *Applied Optics* 8, 1226-1235.
- 4.20 Minsky, M. (1961). Microscopy Apparatus, *US Patent* # 3,013,467.
- 4.21 Pawley, J. ed. (1995). "Fundamental limits in confocal microscopy," in chapter 2 of *Handbook of Biological Confocal Microscopy*, 2<sup>nd</sup> ed., Plenum Press.
- 4.22 Poon, T.-C. (2002a). "Three-dimensional television using optical scanning holography," *Journal of Information Display* 3, 12-16.
- 4.23 Poon, T.-C. (2002b). "Optical scanning holography: principles and applications," in *Three-Dimensional Holographic Imaging*, C.J. Kuo and M. H. Tsai, ed., John Wiley & Sons, Inc.
- 4.24 Poon, T.-C. (2005). "Three-dimensional optical remote sensing by optical scanning holography," Current Research on Image Processing for 3D information displays, sponsored by SPIE Russia Chapter, V. Petrov, ed., *Proc. SPIE*, Vol. 5821, 41-59.
- 4.25 Poon, T.-C. (2006). "Horizontal-parallax-only optical scanning holography," in chapter 10 of *Digital Holography and Three-Dimensional Display: Principles and Applications*, T.-C. Poon ed., Springer, New York, USA.
- 4.26 Poon, T.-C, K. Doh, B. Schilling, M. Wu, K. Shinoda, and Y. Suzuki (1995). "Three-dimensional microscopy by optical scanning holography," *Optical Engineering* 34, 1338-1344.
- 4.27 Poon, T.-C., K. Doh, B. Schilling, K. Shinoda, Y. Suzuki, and M. Wu (1997). "Holographic three-dimensional display using an electron-beam-addressed spatial-light-modulator," *Optical Review* 567-571.
- 4.28 Poon, T.-C. and T. Kim (1999). "Optical image recognition of three-dimensional objects," *Applied Optics* 38, 370-381.
- 4.29 Poon, T.-C., B. D. Schilling, G. Indebetouw, and B. Storrie (2000). Three-dimensional holographic fluorescence microscopy, *U.S. Patent* # 6,038,041.
- 4.30 Poon, T.-C., T. Kim, and K. Doh (2003) "Optical scanning cryptography for secure wireless transmission," *Applied Optics* 42, 6496-6503.
- 4.31 Poon, T.-C., T. Akin, G. Indebetouw and T. Kim (2005). "Horizontal-parallax-only electronic holography," *Optics Express* 13, 2427-2432.
- 4.32 Poon, T.-C. and T. Kim (2006). *Engineering Optics with MATLAB®*, World Scientific, Singapore.
- 4.33 Radio Frequency Identification Technologies: A Workshop Summary (2004). *The National Academies Press*, Washington, D.C.
- 4.34 Refregier P. and B. Javidi (1995) "Optical image encryption using input and Fourier plane random phase encoding," *Optics Letters* 20, 767-769.
- 4.35 Schilling, B. W. and T.-C. Poon (1995). "Real-time pre-processing of holographic information," *Optical Engineering* 34, 3174-3180.
- 4.36 Schilling, B. W. (1997). "Three-dimensional fluorescence microscopy by optical scanning holography," Ph.D. dissertation, Virginia Tech.
- 4.37 Schilling, B., T.-C. Poon, G. Indebetouw, B. Storrie, K. Shinoda, and M. Wu (1997). "Three-dimensional holographic fluorescence microscopy," *Optics Letters* 22, 1506-1508.

- 4.38** Schilling, B.W. and T.-C. Poon (2004). Multicolor electronic holography and 3-D image projection system, *U.S. Patent # 6760134*.
- 4.39** Shinoda, K., Y. Suzuki, M. Wu and T.-C. Poon (1991). Optical heterodyne scanning type holography device, *U.S. Patent # 5064257*.
- 4.40** St. Hilaire, P., S. A. Benton, M. Lucente, M. Jepsen, J. Kollin, H. Yoshikawa, and J. Underkoffler (1990). "Electronic display system for computational holography," *Proc. SPIE*, vol. 1212, 174-182.
- 4.41** St. Hilaire, P., S. A. Benton, and M. Lucente (1992). "Synthetic aperture holography: a novel approach to three-dimensional displays," *Journal of the Optical Society of America A* 9, 1969-1977.
- 4.42** Sun, P and J. -H. Xie (2004). "Method for reduction of background artifacts of images in scanning holography with a Fresnel-zone-plate coded aperture," *Applied Optics* 43, 4214-4218.
- 4.43** Tajahuerce, E., J. Lancis, B. Javidi, and P. Andres (2001). "Optical security and encryption with totally incoherent light," *Applied Optics* 26, 678-680.
- 4.44** Wang, B, C.-C. Sun, W.-C. Su, and A. Chiou (2000). "Shift-tolerance property of an optical double-random phase-encoding encryption system," *Applied Optics* 39, 4788-4793.
- 4.45** Wilson, T. and C. Sheppard (1984). *Theory and Practice of Scanning Optical Microscopy*, Academic Press.
- 4.46** Zhang, T. and I. Yamaguchi (1998). "Three-dimensional microscopy with phase-shifting digital holography," *Optics Letters* 23, 1221-1223.

## Chapter 5

### Optical Scanning Holography: Advances

#### 5.1 Coherent and Incoherent Holographic Processing

In chapter 4, we have discussed some applications that employ optical scanning holography (OSH). OSH has been implemented by the two-pupil optical heterodyne scanning image processor that we have discussed in chapter 3 [see Fig. 3.11]. All of our discussions regarding the applications that we use OSH have so far been confined to incoherent image processing, i.e., the objects that are processed are incoherent and this leads to some of the important applications of 3-D fluorescence microscopy and remote sensing. Coherent 3-D imaging, nevertheless, is an important extension of the processor in biological imaging for the area of *quantitative phase-contrast imaging* [Cuche, Bevilacqua, and Depeursinge (1999)].

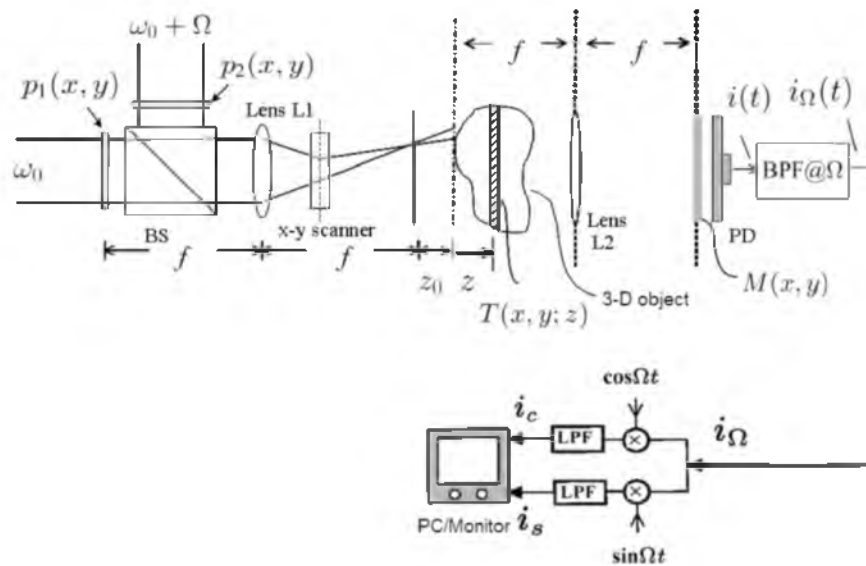


Fig. 5.1 Generalized two-pupil image processor. Adapted from T.-C. Poon, J. Holography Speckle 1, 6-25 (2004).

In this section, we will discuss how the image processor can be configured (or generalized) in order to work in a coherent mode, i.e., the amplitude instead of the intensity of the object can be processed. We shall use Fourier optics discussed in chapter 2, in order to fully analyze the processor.

The generalized processor is shown in Fig. 5.1 with its usual two-pupil set-up for optical scanning of the 3-D object. As compared to the standard setup shown in Fig. 3.11, note that the Fourier transform lens, L2, and the mask,  $M(x, y)$ , are the additional elements in the system. We will model the 3-D object as a stack of transverse slices where each slice of the object is represented by an amplitude transmittance,  $T(x, y; z)$ , which is thin and weakly scattering. We will place the 3-D object in front of the Fourier transform lens, L2.  $M(x, y)$  is a mask located at the back focal plane of lens L2. The photodetector, PD, collects all the light transmitted by the mask and delivers the processed and scanned current,  $i(t)$ , as output of the system. Finally, for the usual multiplexing electronic detection the bandpass filter (BPF) is tuned to  $\Omega$  to give the heterodyne current,  $i_\Omega(t)$ . By using Fourier optics, we shall outline the procedures used to obtain  $i(t)$  upon scanning the object.

The amplitude distribution of the light field, at position  $z$  just before the object slice, is given by

$$P_1(x, y; z + z_0)\exp(j\omega_0 t) + P_2(x, y; z + z_0)\exp[j(\omega_0 + \Omega)t], \quad (5.1-1a)$$

where, according to Fresnel diffraction,

$$P_i(x, y; z + z_0) = \mathcal{F}\{p_i(x, y)\}_{k_x=\frac{k_0 x}{f}, k_y=\frac{k_0 y}{f}} * h(x, y; z + z_0) \quad (5.1-1b)$$

with  $i = 1$  or  $2$  and  $p_i(x, y)$  is the pupil functions shown in Fig. 5.1.

According to the principle of optical scanning developed in section 3.1, the field just after the object slice is

$$\begin{aligned} \chi(x', y', x, y; z) = & \{P_1(x', y'; z + z_0)\exp(j\omega_0 t) \\ & + P_2(x', y'; z + z_0)\exp[j(\omega_0 + \Omega)t]\}T(x' + x, y' + y; z), \end{aligned} \quad (5.1-2)$$

where  $x = x(t)$  and  $y = y(t)$  represent the instantaneous 2-D position of the object with respect to the incident light amplitude distribution. This field then propagates through the Fourier transform lens, L2, and reaches the mask,  $M(x, y)$ . According to Eq. (2.4-5), the field just before the mask, apart from some inessential constant, is

$$\exp\left[-j\frac{k_0 z}{2f^2}(x_m^2 + y_m^2)\right] \\ \times \int \int \chi(x', y', x, y; z) \exp\left[j\frac{k_0}{f}(x'x_m + y'y_m)\right] dx' dy',$$

where we have set  $d_0 = f - z$  in Eq. (2.4-5) to obtain the phase factor in front of the integral.  $x_m$  and  $y_m$  are the coordinates in the plane of the mask. The above field is caused by a single object slide. For a 3-D object, we need to integrate the above field over the thickness,  $z$ , of the 3-D object to find the total field just before the mask. This becomes the following expression

$$\int \left\{ \exp\left[-j\frac{k_0 z}{2f^2}(x_m^2 + y_m^2)\right] \right. \\ \left. \times \int \int \chi(x', y', x, y; z) \exp\left[j\frac{k_0}{f}(x'x_m + y'y_m)\right] dx' dy' \right\} dz.$$

By multiplying the above field by the mask, the field, just after the mask is then given by

$$\psi(x, y; x_m, y_m) = \left\{ \int \left[ \exp\left[-j\frac{k_0 z}{2f^2}(x_m^2 + y_m^2)\right] \right. \right. \\ \left. \left. \times \int \int \chi(x', y', x, y; z) \exp\left[j\frac{k_0}{f}(x'x_m + y'y_m)\right] dx' dy' \right] dz \right\} M(x_m, y_m).$$

Finally, the photodetector, PD, which responds to intensity, gives the current output  $i(t)$  by spatially integrating the intensity:

$$i(t) \propto \int |\psi(x, y; x_m, y_m)|^2 dx_m dy_m.$$

$i(t)$  consists of a baseband current and a heterodyne current at a frequency of  $\Omega$ . After some manipulations, the heterodyne current,  $i_\Omega(t)$ , at the output of a bandpass filter [see Fig. (5.1)], becomes

$$i_\Omega(t) \propto \int [P_1(x', y'; z' + z_0) P_2^*(x'', y''; z'' + z_0) \exp(-j\Omega t) \\ + P_2(x', y'; z' + z_0) P_1^*(x'', y''; z'' + z_0) \exp(j\Omega t)] \\ \times \exp\left\{j\frac{k_0}{f}[x_m(x' - x'') + y_m(y' - y'')]\right\}$$

$$\begin{aligned}
& \times \exp\left[-j\frac{k_0(z' - z'')}{2f^2}(x_m^2 + y_m^2)\right] T(x' + x, y' + y; z') \\
& \times T^*(x'' + x, y'' + y; z'') |M(x_m, y_m)|^2 \\
& \times dx' dy' dx'' dy'' dz' dz'' dx_m dy_m.
\end{aligned} \tag{5.1-3}$$

This heterodyne current contains the scanned and processed information of the 3-D object. Different processing operations can be expected with the specified selections of the pupils,  $p_1(x, y)$  and  $p_2(x, y)$ , as well as the mask,  $M(x, y)$ , which is located at the back focal plane of lens L2. As calculated by Indebetouw, Klysubun, Kim, and Poon [2000], the coherency of the two-pupil scanning system can be modified by changing the mask,  $M(x_m, y_m)$ . We shall summarize the results of Eq. (5.1-3) for  $M(x_m, y_m) = 1$  and  $M(x_m, y_m) = \delta(x, y)$ .

For the mask being an open mask, i.e.,  $M(x, y) = 1$ , Eq. (5.1-3) becomes

$$\begin{aligned}
i_\Omega(t) & \propto \text{Re} \left[ \int P_1^*(x', y'; z + z_0) P_2(x', y'; z + z_0) \right. \\
& \quad \left. \times |T(x' + x, y' + y; z)|^2 dx' dy' dz \exp(j\Omega t) \right].
\end{aligned} \tag{5.1-4}$$

This equation is basically identical to Eq. (3.4-5), which corresponds to incoherent processing, because only the intensity values, i.e.,  $|T|^2$ , are processed. However, note that the intensity can be processed by the two pupils,  $p_1(x, y)$  and  $p_2(x, y)$ .

On the other hand, for a pinhole mask centered on the axis, i.e.,  $M(x, y) = \delta(x, y)$ . Equation (5.1-3) then becomes

$$\begin{aligned}
i_\Omega(t) & \propto \text{Re} \left[ \int P_2(x', y'; z' + z_0) T(x' + x, y' + y; z') dx' dy' dz' \right] \\
& \quad \times \left[ \int P_1^*(x'', y''; z'' + z_0) T^*(x'' + x, y'' + y; z'') dx'' dy'' dz'' \right] \exp(j\Omega t).
\end{aligned}$$

For a specific case, if we let  $p_1(x, y) = \delta(x, y)$ , i.e., one of the scanning beams is a uniform plane wave, and leave  $p_2(x, y)$  as is, then  $\int_{-\infty}^{\infty} P_1^*(x'', y''; z'' + z_0) T^*(x'' + x, y'' + y; z'') dx'' dy'' dz''$  is a constant, and the above equation becomes

$$i_\Omega(t) \propto \text{Re} \left[ \int P_2(x', y'; z + z_0) \right]$$

$$\times T(x' + x, y' + y; z) dx' dy' dz \exp(j\Omega t) \Big]. \quad (5.1-5)$$

We see that we can process the object's amplitude transmittance by pupil  $p_2(x, y)$ . Equations (5.1-4) and (5.1-5) represent the important results of the generalized two-pupil heterodyne scanning image processor. By varying the detection mode from pinhole to spatially integrating detection, we are able to change the coherence property of the imaging process of a 3-D object from linear in intensity [see Eq. (5.1-4)] to linear in amplitude [see Eq. (5.1-5)]. By varying the size of the mask, it is possible to obtain 3-D *partial coherent image processing* [Poon and Indebetouw (2003)].

By incorporating Eq. (5.1-4) and (5.1-5) into one simple important result, we have

$$i_\Omega(t) \propto \text{Re}[i_{\Omega p}(x, y) \exp(j\Omega t)], \quad (5.1-6a)$$

where, for  $M(x, y) = \delta(x, y)$  and  $p_1(x, y) = \delta(x, y)$  we have the following coherent processing equation,

$$i_{\Omega p}(x, y) = \int P_2(x', y'; z + z_0) T(x' + x, y' + y; z) dx' dy' dz. \quad (5.1-6b)$$

For  $M(x, y) = 1$ , we have the incoherent processing equation,

$$\begin{aligned} i_{\Omega p}(x, y) &= \int P_1^*(x', y'; z + z_0) P_2(x', y'; z + z_0) \\ &\times |T(x' + x, y' + y; z)|^2 dx' dy' dz. \end{aligned} \quad (5.1-6c)$$

Again  $T$  or  $|T|^2$  is the input object that is being scanned and it can be a complex amplitude object or intensity object.  $i_\Omega(t)$  is the scanned and processed heterodyne output current at a temporal frequency of  $\Omega$  from the photodetector, and  $i_{\Omega p}(x, y)$  is in general a complex function. Hence, the amplitude and the phase of the heterodyne current,  $i_\Omega(t)$ , carry the complete processed information.

The scanned and processed current can be demodulated according to Fig. 5.1 under the usual multiplexing electronic detection, and this gives the following two outputs:

$$i_c(x, y) \propto |i_{\Omega p}(x, y)| \sin(\theta), \quad (5.1-7a)$$

and

$$i_s(x, y) \propto |i_{\Omega p}(x, y)| \cos(\theta), \quad (5.1-7b)$$



where  $i_{\Omega_p}(x, y) = |i_{\Omega_p}(x, y)| \exp[j\theta(x, y)]$ . Processing operations can be manipulated by the selected choice of pupils,  $p_1(x, y)$  and/or  $p_2(x, y)$ , according to Eq. (5.1-6).

These important results open up new avenues for unconventional imaging. For one, coherent processing has been recently assessed experimentally [Indebetouw, Tada and Leacock (2006)], and the other is its ability to perform 3-D coherent image processing. In the case of incoherent processing, we can perform *3-D complex incoherent image processing*. The term “3-D” means the object being processed can be of a 3-D nature, and the term “complex” means that the processing element for the intensity object can be represented by a complex function [see Eq. (5.1-4) as the processing element is  $P_1^* P_2$ ]. Thus, we have mentioned some of the virtually unexplored topics in 3-D optical image processing [Poon and Indebetouw (2003)].

### Example 5.1 Holographic Recording in Coherent Mode

For coherent processing, we will use Eq. (5.1-6b). In general,  $T(x, y; z)$  is processed by selecting  $p_2(x, y)$ . For a simple holographic recording, we select  $p_2(x, y) = 1$ . For this selection, according to Eq. (5.1-1b) and by using Table 1.1 and Eq. (2.3-11),  $P_2(x, y; z + z_0)$  becomes

$$\begin{aligned} P_2(x, y; z + z_0) &= \mathcal{F}\{p_2(x, y)\}_{k_x = \frac{k_0 x}{f}, k_y = \frac{k_0 y}{f}} * h(x, y; z + z_0) \\ &= 4\pi^2 \delta\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \\ &\quad * \left\{ \exp[-jk_0(z + z_0)] \frac{jk_0}{2\pi(z + z_0)} \exp\left[\frac{-jk_0(x^2 + y^2)}{2(z + z_0)}\right] \right\} \\ &\propto \exp[-jk_0(z + z_0)] \frac{jk_0}{2\pi(z + z_0)} \exp\left[\frac{-jk_0(x^2 + y^2)}{2(z + z_0)}\right]. \end{aligned} \quad (5.1-8)$$

By substituting Eq. (5.1-8) into Eq. (5.1-6b), we now have

$$\begin{aligned} i_{\Omega_p}(x, y) &= \int \exp[-jk_0(z + z_0)] \frac{jk_0}{2\pi(z + z_0)} \\ &\quad \times \exp\left[\frac{-jk_0(x'^2 + y'^2)}{2(z + z_0)}\right] T(x' + x, y' + y, z) dx' dy' dz \\ &= \int \exp[-jk_0(z + z_0)] \frac{jk_0}{2\pi(z + z_0)} \end{aligned}$$

$$\times \left[ \left\{ \exp \left[ \frac{-jk_0(x^2 + y^2)}{2(z + z_0)} \right] \right\}^* \otimes T(x, y; z) \right] dz$$

according to the definition of correlation. By using Eq. (1.2-5), the above integration can be written in terms of convolution to become

$$i_{\Omega_P}(x, y) = \int \exp[-jk_0(z + z_0)] \frac{jk_0}{2\pi(z + z_0)} \\ \times \left[ \left\{ \exp \left[ \frac{-jk_0(x^2 + y^2)}{2(z + z_0)} \right] \right\} * T(x, y; z) \right] dz. \quad (5.1-9)$$

Note that this corresponds to the holographic recording of coherent information,  $T(x, y; z)$ , because it is clearly shown by comparing it to the incoherent case given by Eq. (3.5-5). Indeed, Eq. (5.1-9) tells us that we have a complex Fresnel zone plate hologram of  $T(x, y; z)$ .

## 5.2 Single-Beam Scanning vs. Double-Beam Scanning

Recently, the applicability of optical scanning holography (OSH) to 3-D microscopy has been assessed by taking into account polarization effects, high numerical apertures, and generalized illumination wavefronts [Swoger, Martinez-Corral, Huiskens, and Stelzer (2002)]. In low- $NA$  systems, polarization remains the same during light propagation. Polarization is necessarily taken into account when high- $NA$  lenses are used. Generalized illumination refers to the use of the two pupils associated with the two scanning beams. Ideally, one of the pupils is a delta function and the other is unity. This gives a plane wave and a spherical wave on the object, respectively. For arbitrary pupil functions, we have a generalized illumination wavefront on the object [Indeed, for example, Eq. (5.1-4) points to the fact that the object,  $|T(x, y; z)|^2$ , is actually illuminated by  $P_1^* P_2$ , thereby providing a generalized illumination].

The authors also use the term “reference beam” and “object beam” when referring to the plane wave and the spherical wave in the ideal case. In addition, a single-beam scanning technique has been proposed and compared to double-beam scanning. *Double-beam scanning* in OSH refers to the fact that the combination of a plane wave and a spherical wave is used to raster scan a thick specimen, while *single-beam scanning* means that only one of the waves is used to scan and the other wave remains stationary with respect to the specimen. Figure 5.2 shows the schematics of the scanning configurations for OSH. For single-beam scanning, mirror M1 is used to scan the beam and mirror M2 is fixed. For double-beam scanning, M2 is used to

scan and M1 is fixed. The specimen and the mask are placed in the front focal plane and in the back focal plane of lens L2, respectively. This corresponds to the situation shown in Fig. 5.1. In the following section, we will summarize some observations made by Swoger et al.

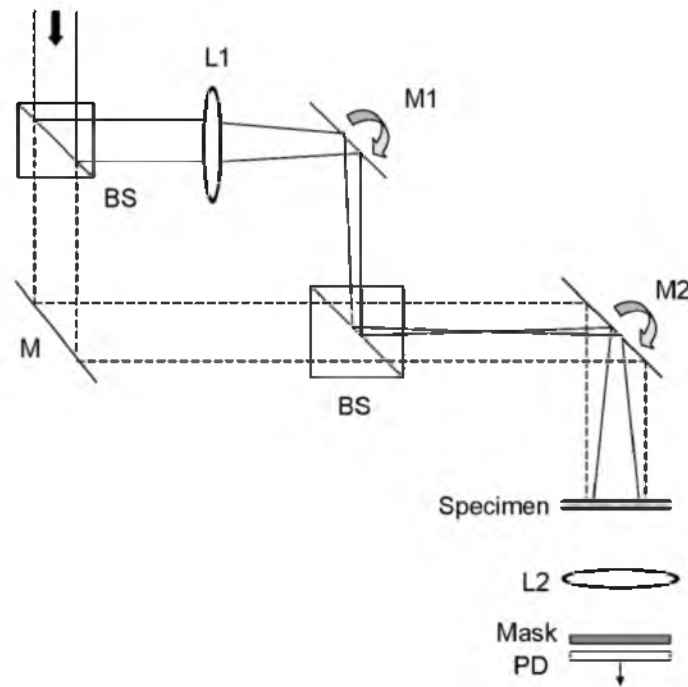


Fig. 5.2 Scanning configurations for optical scanning holography.  
(M: mirror, M1,2: scanning mirrors, BS: beamsplitter, L1 and L2: lenses, PD: photodetector).

Regardless of the polarization directions of the two scanning beams, double-beam scanning is well suited for optical scanning holography when working in the incoherent mode (open mask in front of the detector). However, in order to obtain holographic information in a coherent mode (pin-hole mask in front of the detector), it is necessary to have a uniform plane wave as one of the scanning beams on the object. In any case, if the reference beam is a uniform plane wave and the object wave has the same polarization of the plane wave, the results reduce to the conclusion developed earlier by Indebetouw, Klysubun Kim, and Poon [2000]. On the other hand, during single-beam scanning for the coherent mode, there is no need for the reference beam to be uniform and the polarization directions do not need to be constant. The reference beam here now refers specifically to the beam that is not being scanned. However, the incoherent mode is restricted by the constant polarization direction of the reference beam when we attempt to obtain holographic information. Therefore, there are several pros and cons

associated with the two scanning configurations and for future work we should include the actual implementation of a high- $NA$  optical scanning holographic system that is capable of allowing us to observe live biological specimens in both fluorescence and phase contrasts. Recently, optical scanning holographic systems operating in the coherent mode have been implemented to test its phase contrast capabilities [Indebetouw, Tada, and Leacock (2006)]. To end this section, I also want to point out that recent experiments have been demonstrated by using single-beam scanning [Chien, Dilworth, Liu and Leith (2006)]. The authors consider the term “scanning holography” and “synthetic-aperture optics” to be essentially interchangeable because both techniques imply phase-preserving scanning. Indeed, it was pointed out as early as the late 1970's in the article by Poon and Korpel [1979] that scanning holographic recording was analogous to *synthetic-aperture radar*.

### 5.3 PSF Engineering

As pointed out in section 3.6, the basic principle of optical scanning holography is that we simply create the time-dependent Fresnel zone plate (TDFZP) in order to raster scan the object that allows us to obtain holographic information. This can be implemented by the two-pupil heterodyne image processor that was discussed in section 3.4. In the processor, we let the two pupil functions be  $p_1(x, y) = 1$  and  $p_2(x, y) = \delta(x, y)$  on the pupil plane, as shown in Fig. 5.3a), so that the scanning beam intensity at a distance of  $z_0$  away from the focal point,  $c$ , of the lens is given by Eq. (3.6-1) as follows:

$$\begin{aligned} I_{\text{scan}}(x, y; t) &= \left| a \exp[j(\omega_0 + \Omega)t] + \frac{jk_0}{2\pi z_0} \exp\left[-\frac{jk_0(x^2 + y^2)}{2z_0}\right] \exp(j\omega_0 t) \right|^2, \\ &= A + B \sin\left[\frac{k_0}{2z_0}(x^2 + y^2) + \Omega t\right], \end{aligned} \quad (5.3-1)$$

where it is understood that the two pupils are of different temporal frequencies accordingly. The scanning beam intensity is the time-dependent Fresnel zone plate, and the situation is shown schematically in Fig. 5.3a).

If the object is a pin hole, a delta function mathematically, then the output after electronic detection (for example, multiplying with  $\cos\Omega t$  and lowpass filtering) is given by Eq. (3.6-3) as follows:

$$i_c(x, y) \sim \sin\left[\frac{k_0}{2z_0}(x^2 + y^2)\right]. \quad (5.3-2)$$

The result above is the hologram of a pin-hole object for the single channel. For brevity, the channel due to the multiplying with  $\sin\Omega t$  and lowpass filtering is not considered here. Upon plane-wave illumination of the hologram, a real image will focus at  $z_0$  in front of the hologram. If the hologram has a limiting aperture size of radius  $r$ , then its  $NA$  is  $r/z_0$ , which gives  $\lambda_0/NA$  as the resolution of the reconstructed point source. This reconstructed point source is the point spread function (PSF) of the system being considered. Now, by manipulating the functional form of the pupils, we can modify the PSF of the system as we see fit. Nowadays, this is known as *PSF-engineering* [Martinez-Corral (2003)].

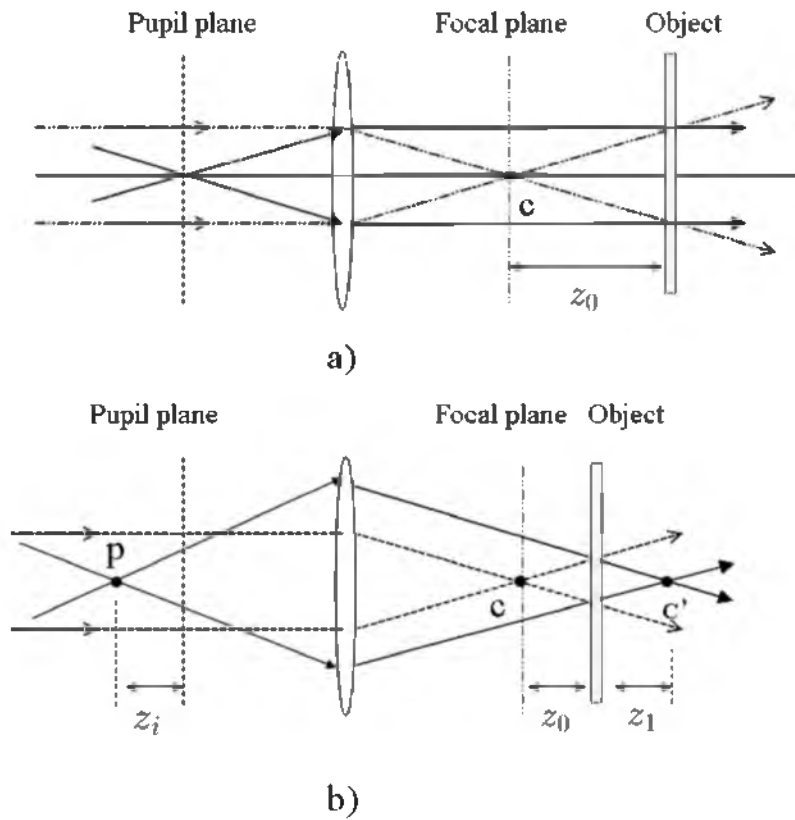


Fig. 5.3a) Optical scanning holography with time-dependent FZP as a scanning beam.  
b) Optical scanning holography with scanning spherical waves of opposite curvatures.  
Adapted from T.-C. Poon, J. Holography and Speckle I, 6 (2004).

Let us consider the following situation. We let  $p_1(x, y) = 1$  as before, but set  $p_2(x, y) = \exp[-\frac{j k_0(x^2 + y^2)}{2z_0}]$ . We shall find the PSF with these choices of the pupils. We are familiar with the choice of  $p_1(x, y) = 1$  because it gives a

spherical wave on the object at a distance  $z_0$  away from the focal plane of the lens, which corresponds to the term  $\frac{jk_0}{2\pi z_0} \exp[-\frac{jk_0(x^2+y^2)}{2z_0}] \exp(j\omega_0 t)$  in Eq. (5.3-1). The situation is shown in Fig. 5.3b). Let us find what  $a$  is in Eq. (5.3-1) for the choice of  $p_2(x, y)$  given above. Note that  $a(x, y)$  is now a function of  $x$  and  $y$ . Through the Fourier transformation of the lens [see Eq. (2.4-6)] and Fresnel diffraction at a distance of  $z_0$  [see Eq. (2.3-12)], we can write the field distribution due to  $p_2(x, y)$  on the object as

$$\begin{aligned} a(x, y) &= \mathcal{F}_{xy}\{p_2(x, y)\} \bigg|_{\substack{k_x=k_0x/f \\ k_y=k_0y/f}} * h(x, y; z_0) \\ &= \mathcal{F}_{xy}\left\{\exp\left[-\frac{jk_0(x^2+y^2)}{2z_i}\right]\right\} \bigg|_{\substack{k_x=k_0x/f \\ k_y=k_0y/f}} * h(x, y; z_0), \end{aligned} \quad (5.3-3)$$

where  $f$  is the focal length of the lens. The above Fourier transformation can be found by using Table 1.1 on page 2. We then perform the convolution and the result, apart from some constant, is given as follows:

$$a(x, y) = \frac{jk_0}{2\pi z_1} \exp\left[\frac{jk_0(x^2+y^2)}{2z_1}\right]. \quad (5.3-4)$$

Note that  $a(x, y)$  can either be converging wavefronts or divergent wavefronts on the specimen and this is dependent on the location of the focused point,  $p$ . In Fig. 5.3b), we show that point  $c'$  is the image point of the focused point,  $p$ , and hence the radius of curvature,  $z_1$ , is positive in the situation because we have converging wavefronts illuminating the specimen [see Example 2.4 on spherical wave]. The value of  $z_1$  can be designed accordingly by properly locating point  $p$  or the distance,  $z_i$ , under the imaging condition so that

$$\frac{1}{f+z_i} + \frac{1}{f+z_0+z_1} = \frac{1}{f}. \quad (5.3-5)$$

Therefore, we can design  $p_2(x, y)$  with a different radius of curvature.

For the choices of  $z_i = 2f$  and  $z_0 = z_1 = f/4$ , Eq. (5.3-5) will be satisfied. Since  $z_0 = z_1$ , we will have spherical waves of opposite curvatures illuminating the specimen. Hence, with Eq. (5.3-4) and  $z_0 = z_1$ , the scanning beam intensity becomes

$$\begin{aligned} I_{\text{scan}}(x, y; t) \\ \propto \left| \exp\left[\frac{jk_0(x^2+y^2)}{2z_0}\right] \exp[j(\omega_0 + \Omega)t] + \exp\left[-\frac{jk_0(x^2+y^2)}{2z_0}\right] \exp(j\omega_0 t) \right|^2 \end{aligned}$$

$$= A' + B' \sin\left[\frac{k_0}{2(z_0/2)}(x^2 + y^2) + \Omega t\right], \quad (5.3-6)$$

where  $A'$  and  $B'$  are some constants. The above scanning beam gives

$$i_c(x, y) \sim \sin\left[\frac{k_0}{z_0}(x^2 + y^2)\right] \quad (5.3-7)$$

as our new hologram for the selected pupils. Upon real image reconstruction of the hologram given by Eq. (5.3-7), we see that the image is formed at a distance of  $z_0/2$ . For the same limiting aperture size of the hologram,  $r$ , similar to the standard optical scanning setup, its  $NA$  now becomes  $r/(z_0/2)$ . This gives  $\lambda_0/(2NA)$  as the resolution of the reconstructed point source, which is our new PSF. The result means that for the same hologram aperture size, it is possible to synthesize the point-spread function in optical scanning holography in order to obtain holographic reconstructions with a transverse resolution exceeding the Rayleigh limit of the aperture up to a factor of 2, at least in the limit of low  $NA$ . Indeed, this has been investigated by Indebetouw (2002) and was recently confirmed by optical experiments [Indebetouw, Maghnouji, and Foster (2005)]. Other pupils examined so far include *axicons*, which have been used to achieve optical sectioning, however, thus far only simulations have been provided [Indebetouw, Zhong, and Chamberlin-Long (2006)].

In general, arbitrary complex amplitude distributions of the pupils can be synthesized by using masks, refractive or diffractive optical elements (DOEs), or spatial light modulators which allow dynamic changes. For  $p_2(x, y) = \exp\left[-\frac{j k_0(x^2 + y^2)}{2z_0}\right]$  shown in this example, we can simply implement it by placing a point source at a distance,  $z_i$ , in front of the pupil plane as shown in Fig. 5.3b). In general, I want to point out that the two-pupil method offers a broad range of possibilities when synthesizing unconventional PSFs for potential novel applications. One of the new applications is *super-resolution* [Indebetouw, Tada, Rosen and Brooker (2007)].

## References

- 5.1 Chien, W.-C., D. S. Dilworth, E. Liu, and E. N. Leith (2006). "Synthetic-aperture chirp confocal imaging," *Applied Optics* 45, 501-510.
- 5.2 Cuhe, E., F. Bevilacqua, and C. Depeursinge (1999). "Digital holography for quantitative phase-contrast imaging," *Optics Letters* 24, 291-293.

- 5.3 Indebetouw, G. (2002). "Properties of a scanning holographic microscope: improved resolution, extended depth of focus, and/or optical sectioning," *Journal of Modern Optics* 49, 1479-1500.
- 5.4 Indebetouw, G., P. Klysubun, T. Kim, and T.-C. Poon (2000). "Imaging properties of scanning holographic microscopy," *Journal of the Optical Society of America A* 17, 380-390.
- 5.5 Indebetouw, G., A. El Maghnouji, and R. Foster (2005). "Scanning holographic microscopy with transverse resolution exceeding the Rayleigh limit and extended depth of focus," *Journal of the Optical Society of America A* 22, 829-898.
- 5.6 Indebetouw, G., Y. Tada and J. Leacock (2006). "Quantitative phase imaging with scanning holographic microscopy: an experimental assessment," available at <http://www.biomedical-engineering-online.com/content/5/1/63>.
- 5.7 Indebetouw, G., W. Zhong, and D. Chamberlin-Long (2006). "Point-spread function synthesis in scanning holographic microscopy," *Journal of the Optical Society of America A* 23, 1708-1717.
- 5.8 Indebetouw, J., Y. Tada, J. Rosen, and G. Brooker (2007). "Scanning holographic microscopy with resolution exceeding the Rayleigh limit of the objective by superposition of off-axis holograms," *Applied Optics*, to appear.
- 5.9 Martinez-Corral, M. (2003). "Point spread function engineering in confocal scanning microscopy," *Proceedings of SPIE*, Vol. 5182, 112-122.
- 5.10 Poon, T.-C. and A. Korpel (1979). "Optical transfer function of an acousto-optic heterodyning image processor," *Optics Letters* 4, 317-319.
- 5.11 Poon, T.-C. (2004). "Recent progress in optical scanning holography," *Journal of Holography and Speckle* 1, 6-25.
- 5.12 Poon, T.-C. and G. Indebetouw (2003). "Three-dimensional point spread functions of an optical heterodyne scanning image processor," *Applied Optics* 42, 1485-1492.
- 5.13 Swoger, J., M. Martínez-Corral, J. Huisken, and E. H. K. Stelzer (2002). "Optical scanning holography as a technique for high-resolution three-dimensional biological microscopy," *Journal of the Optical Society of America A* 19, 1910-1918.



## Index

### A

- acousto-optic frequency shifter 72
- acousto-optic modulator 72
- angular frequency 25
- aperture 29
- axicon 146

### B

- bipolar PSF 46,49
- bitmap 7
- Bragg angle 73

### C

- carrier 30
- coding process 55
- coherent point spread function 44
- coherent transfer function 45
- complex amplitude 30
- conservation of energy 73
- conservation of momentum 73
- constitution relations 22
- continuity equation 22
- convolution 14, 16
- correlation 14, 16
  - auto- 17

- cross- 17

- current
  - baseband 70
  - heterodyne 70

### D

- decoding process 55
- decryption 121
- depth of field 100
- depth of focus 97
- detection
  - electronic multiplexing 70
  - lock-in 70
  - optical coherent 70
  - optical direct 67
  - optical incoherent 67
- diffracting screen 29
- display
  - 3-D 106
- distance
  - coding 120
  - decoding 122

### E

- electric field 22
- encoding
  - double-random phase 124

encryption 121  
on-the-fly 118

**F**

filtering  
bandpass 48  
lowpass 47  
spatial 45  
Fourier optics  
spatial frequency transfer  
function in 34  
spatial impulse response in 34  
Fourier plane 43  
frequency transfer function 13  
Fourier transform 1  
Fresnel diffraction 34  
Fresnel diffraction formula 34  
Fresnel zone plate (FZP) 49, 52  
time-dependent 92, 102  
function  
delta 2, 8  
Gaussian 2  
rect 2, 3  
separable 3  
sinc 2, 4  
transfer 13

**G**

Gaussian beam 40  
Gaussian function 2

**H**

Helmholtz equation 30  
heterodyning 69  
hologram 51  
complex Fresnel zone plate 85  
cosine-FZP 83  
off-axis 59  
sine-FZP 83  
twin-image-free 83  
hologram scaling 110, 113

holographic recording 51  
holography 49  
carrier-frequency 60  
CO<sub>2</sub> scanning 84  
digital 60  
electronic 62  
HPO-optical scanning 117  
phase-shifting 85  
scanning 65, 81  
off-axis 56  
optical scanning 65, 83, 87, 92  
X-ray scanning 84  
homodyning 94

**I**

ideal lens 40  
image intensity 45  
image processing  
3-D complex incoherent 140  
bipolar incoherent 48  
coherent 43  
complex incoherent 78  
incoherent 44  
impulse response 11  
initial condition 29  
in-phase component 71  
integral  
convolution 11  
correlation 16  
invariance 10  
issue  
spatial frequency resolution  
115  
spatial resolution 116  
data transmission 117

**K**

key  
decryption 122  
encryption 121

**L**

lateral resolution 97  
linearity 10  
liquid crystal television 61  
lock-in detection 70

**M**

magnetic field 22  
magnification  
    lateral 113  
    longitudinal 113  
    holographic 112  
Maxwell's equations 21  
medium  
    homogeneous 22  
    isotropic 22  
    linear 22  
microscopy  
    fluorescence 101  
    optical sectioning 99  
    scanning confocal 99  
    scanning holographic 97  
mixing 70  
modulator  
    acousto-optic 72  
    electron-beam-addressed  
    spatial light 107

**N**

numerical aperture 98  
Nyquist sampling 116

**O**

optical coherent detection 70  
optical coherent tomography  
    (OCT) 101  
optical direct detection 67  
optical heterodyning 65, 69  
optical incoherent detection 67  
optical pattern recognition 17  
optical scanning 65

optical scanning cryptography 117  
optical scanning holography 65,  
    97, 117  
    HPO- 117  
optical transfer function(OTF) 46,  
    78

**P**

parallax 49  
paraxial approximation 33  
pattern recognition 17  
phase 26  
phase curvature 42  
phase grating 73  
phasor 30  
phonon 73  
photo-bleaching 101  
photon 73  
planar wavefront 27  
Planck's constant 73  
Pockels effect 108  
point spread function (PSF) 13  
    bipolar 46  
    coherent 44  
    complex 49  
    intensity 46  
principle of conservation of  
    charge 22  
processing  
    coherent holographic 136  
    Fourier-plane 43  
    incoherent holographic 136  
propagation constant 26  
propagation vector 26  
PSF-engineering 143  
pupil function 45

**Q**

quadrature component 71  
quantitative phase-contrast  
    imaging 133

**R**

radio frequency identification (RFID) 118  
 radius of curvature 35  
 real image 56, 60  
 reconstruction 55  
 recording 55  
 recording angle 58

**S**

sampling property 9  
 scalar diffraction theory 29  
 scalar wave equation 25  
 scanning  
   single-beam 141  
   double-beam scanning 141  
 separable function 3  
 solution  
   planar wave 25  
   spherical wave 27  
 source 22  
 spatial carrier 60  
 spatial filter 44  
 spatial frequency response 13  
 spatial frequency transfer function  
   in Fourier optics 34  
 spatial frequency transfer function  
   of propagation 31  
 spatial impulse response  
   in Fourier optics 34  
 spatial impulse response 13  
 spatial impulse response  
   of propagation 33  
 spatial light modulator (SLM) 61  
   electron-beam addressed 107  
 spherical wavefront 28  
 square aperture 36  
 superposition 10

super-resolution 146  
 synthetic-aperture radar 143  
 system  
   coherent optical 45  
   incoherent optical 45  
   linear 10  
   linear space-invariant 12  
   time-invariant 11  
   two-pupil 75

**T**

TV  
   holographic 106  
 twin image 57  
 twin-image problem 58  
 two-pupil system 75

**U**

uncertainty relationship 99  
 upshifted interaction 74

**V**

vector wave equation 23  
   homogeneous 24  
 viewing angle 115  
 virtual image 57, 60

**W**

Wave  
   diverging spherical 35  
   object 51  
   plane 26  
   reconstruction 55, 59  
   reference 51  
   spherical 28, 52  
 wavelength scaling 114

**Z**

zero-order beam 56, 59

## Author

### **Ting-Chung Poon**

Bradley Department of Electrical and Computer Engineering  
Virginia Tech  
Blacksburg, Virginia, USA 24061  
e-mail: tcpoon@vt.edu

Dr. Ting-Chung Poon received his physics and engineering degrees from the *University of Iowa*, Iowa City. He is a professor at *Virginia Tech* in the *Bradley Department of Electrical and Computer Engineering*. His current research interests include acousto-optics, hybrid (optical/electronic/digital) 3-D image processing; optical scanning holography and its applications in 3-D cryptography, 3-D display, 3-D microscopy, 3-D optical remote sensing, and 3-D pattern recognition. Dr. Poon is the co-author of the textbooks *Engineering Optics with MATLAB®* (World Scientific 2006), *Contemporary Optical Image Processing with MATLAB®* (Elsevier 2001), and *Principles of Applied Optics* (McGraw-Hill 1991). He is also editor of the book *Digital Holography and Three-Dimensional Display* (Springer 2006) and has served as a panelist for the National Institutes of Health and the National Science Foundation, and as guest editor of, among other journals, *International Journal of Optoelectronics* and *Optical Engineering*. Dr. Poon currently serves as a topical/associate editor for *Applied Optics* and the *International Journal of Optomechatronics*. He is also on the editorial board of *Optics and Laser Technology* and the *Journal of Holography and Speckle*. Dr. Poon is a fellow of the OSA and SPIE and is a senior member of the IEEE.