

Computer-simulated Fresnel holography

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2000 Eur. J. Phys. 21 317

(<http://iopscience.iop.org/0143-0807/21/4/305>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 93.180.53.211

This content was downloaded on 10/10/2013 at 06:31

Please note that [terms and conditions apply](#).

Computer-simulated Fresnel holography

Seymour Trester

Physics Department, C W Post Campus, Long Island University, Brookville, NY 11548, USA

E-mail: strestre@liu.edu

Received 10 November 1999, in final form 22 February 2000

Abstract. This paper presents a method of producing a computer-generated hologram for an object that has depth. It is a continuation of the author's previously published work on computer-simulated holography, which dealt with the generation of Fraunhofer holograms for planar objects. The technique is now extended to the case of Fresnel holograms for simple objects whose points do not lie in a single plane. The method utilizes the computer's ability, using packaged mathematical software, to compute the Fresnel diffraction pattern of a planar aperture by taking the Fourier transform of a modified aperture function. By adding a plane wave reference beam to the diffraction pattern, a hologram for each plane of the object is constructed, and the superposition of these holograms yields a resultant hologram whose reconstruction exhibits the depth of the object. Since computer-simulated holography is based on the interference and diffraction of light waves, the material presented in this paper could serve to enhance the understanding of these topics in a university course in optics.

1. Introduction

The use of the personal computer to calculate and display diffraction patterns for various apertures is well documented in the literature [1–3]. For the most part, these calculations are based on the ability of packaged mathematical software programs to compute the Fourier transform of a function. The method used for the calculation of Fraunhofer diffraction has been extended by the author [4] to computer generation of Fraunhofer holograms of planar objects. Similarly, the present paper extends the treatment of Fresnel diffraction as a Fourier transform of an aperture function to computer generation of Fresnel holograms, that is, holograms which have the ability to focus light and hence, for non-planar objects, show depth in their reconstruction.

In general, the evaluation of the Fresnel integral used to obtain the diffraction pattern for an aperture is rather difficult. The theory presented below demonstrates how the Fresnel integral can be evaluated using the Fourier transform. The work of Mas *et al* [5] discusses this method and some of the accuracy and aliasing problems it presents. Recently, the author [6] used the fast Fourier transform algorithm to calculate and display Fresnel diffraction patterns for various apertures.

2. Theory

The foundations of the scalar theory of diffraction can be found in most texts on optics, such as that of Hecht (see chapter 10 of [7]). Using this theory, Goodman (see chapters 3 and 4 of [8])

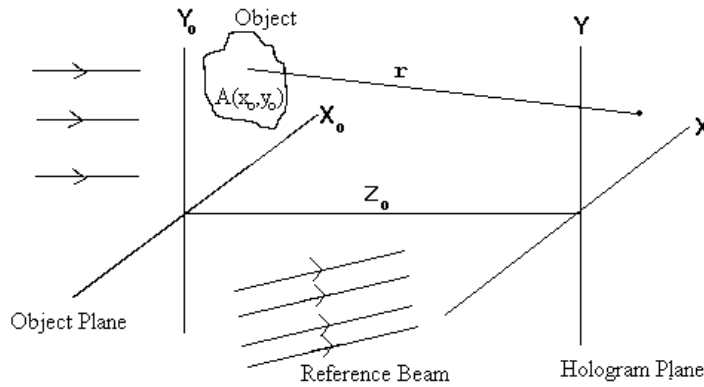


Figure 1. Relevant geometry and illumination for the construction of a Fresnel hologram.

shows how the Fresnel diffraction due to a planar aperture can be treated as a Fourier transform. Consider the case shown in figure 1, where a planar transparent object in an otherwise opaque plane is illuminated from the left by a beam of monochromatic light of wavelength λ travelling parallel to the z -axis. For the geometry shown in this figure, the electric field $E(x, y)$ of the diffracted light at any point (x, y) in the hologram plane, a plane parallel to and at a distance z_0 to the right of the object plane, is given by the mathematical expression of the Huygens–Fresnel principle (see p 60, equation (4-10) of [8]). That is,

$$E(x, y) = (-i/\lambda z_0) \exp(ikz_0) \exp[(ik/2z_0)(x^2 + y^2)] \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x_0, y_0) \exp[(ik/2z_0)(x_0^2 + y_0^2)] \exp[(-ik/z_0)(xx_0 + yy_0)] dx_0 dy_0 \quad (1)$$

where $k = 2\pi/\lambda$ and x_0, y_0 and x, y are the x and y coordinates in the object and hologram planes, respectively. The function $A(x_0, y_0)$ is the amplitude of the light transmitted by the object and for the unit-amplitude incident plane wave assumed in this paper, is numerically equal to the amplitude transmission function of the object. Equation (1) was derived using the paraxial ray approximation, that is, $z_0^2 \gg [(x_0 - x)^2 + (y_0 - y)^2]$, which allows one to expand the distance r of figure 1 as $r \approx z_0 + (1/2z_0)\{(x_0 - x)^2 + (y_0 - y)^2\}$. Hence, as Goodman points out, the spherical wavelets of Huygens are replaced by quadratic surfaces. Equation (1) shows that the resulting Fresnel diffraction field in the hologram plane is related to the Fourier transform of the object's transmission function modified by $\exp[(ik/2z_0)(x_0^2 + y_0^2)]$ and is to be evaluated at the spatial frequencies $f_x = x/\lambda z_0$ and $f_y = y/\lambda z_0$. It will be shown later that when the computer is used to take the Fourier transform, a method is used in which the spatial frequencies will automatically satisfy these relations and thus have the correct spacing in the hologram plane.

The construction of a Fresnel hologram is achieved by letting the field of a plane wave reference beam, whose wavelength is the same as that of the light used to illuminate the aperture, interfere in the hologram plane with the diffraction field of equation (1). This is also shown in figure 1. In general, the field of the off-axis reference beam at the hologram plane is given by $E_r = \tilde{A}_r \exp[ik(x \cos \alpha + y \cos \beta)]$ where \tilde{A}_r is the complex amplitude of the reference beam and the angles α and β are between the propagation vector \mathbf{k} and the x - and y -axes, respectively. In this paper, the optical Gabor in-line Fresnel hologram (see [8], section 8-3 and [9], section 2.5) will be simulated by computer. For this type of hologram the reference beam travels along the z -axis so that the angles α and β are both 90° . One way to achieve this, as Gabor did, is to have the object lie in an object plane which is highly transmissive, rather than opaque, and let the reference beam be formed by that part of the incident plane wave

illumination which goes through undiffracted (see [8], section 8-3). In this case, the reference beam field in the hologram plane can be written as $E_r = A_r \exp(ikz_0)$, where the amplitude A_r is real. Hence, to emulate the construction of a Gabor hologram for the object in the opaque plane shown in figure 1, we use the above expression for E_r and add it to the diffraction field given by equation (1). The resulting light intensity in the hologram plane is then

$$I(x, y) = |E_r + E(x, y)|^2 = A_r^2 + |E(x, y)|^2 + A_r \exp(ikz_0) E^*(x, y) + A_r \exp(-ikz_0) E(x, y) \quad (2)$$

where $E^*(x, y)$ is the complex conjugate of $E(x, y)$.

If the intensity pattern $I(x, y)$ is recorded on film, the developed transparency is the hologram. If, as is usual for Gabor holograms, the amplitude of the reference beam, A_r , is much larger than the amplitude of the light diffracted by the object, the second term in $I(x, y)$ can be neglected. In any case, in the computer-generated hologram discussed in the next section, the second term will be subtracted out by spatial filtering. In addition, we assume that the development process is such that this transparency has an amplitude transmission function, $T(x, y)$, which is linear in the intensity $I(x, y)$. Under these conditions, it is shown in section 8-3 of [8] and section 2.5 of [9], that the transmission function can be written as $T(x, y) = a + b[E_r E^*(x, y) + E_r^* E(x, y)]$ where a and b are constants.

We now reconstruct or play back the hologram by illuminating it with light from a monochromatic point source of wavelength λ located on the z -axis at a distance z_p to the left of the hologram. For the case in which $z_p^2 \gg x^2 + y^2$, one can approximate the field of this incident spherical wave, at the plane of the hologram, by $A_p \exp(ikz_p) \exp\{(ik/2z_p)(x^2 + y^2)\}$ where A_p is the amplitude of the reconstruction light. The field of the diffracted light in an image plane a distance z_i to the right of the hologram, for which the approximation $z_i^2 \gg x_i^2 + y_i^2$ is valid, will then be given by the Huygens–Fresnel integral

$$E(x_i, y_i) = C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[(ik/2z_p)(x^2 + y^2)] T(x, y) \exp[(ik/2z_i)(x^2 + y^2)] \times \exp[(-ik/2z_i)(xx_i + yy_i)] dx dy \quad (3)$$

where $C = (-iA_p/\lambda z_i) \exp[ik(z_p + z_i)] \exp[(ik/2z_i)(x_i^2 + y_i^2)]$. In the reconstruction, the constant term in $T(x, y)$ gives rise to an attenuated spherical wave which serves as a background field, and the fields due to the terms $bE_r E^*(x, y)$ and $bE_r^* E(x, y)$ of $T(x, y)$ then add to this background field. It is the fields of these two terms which produce the images of the object in the reconstruction process. To see this explicitly, we first calculate $E_2(x_i, y_i)$, the contribution made to the reconstruction field by the term $bE_r E^*(x, y)$. By inserting $bA_r \exp(ikz_0) E^*(x, y)$ for the $T(x, y)$ in equation (3), since $E_r = A_r \exp(ikz_0)$ for Gabor holograms, and using the expression for $E^*(x, y)$ from equation (1) one finds that, after rearranging terms, $E_2(x_i, y_i)$ can be written as

$$E_2(x_i, y_i) = C_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A^*(x_0, y_0) \exp[(ik/2z_0)(x_0^2 + y_0^2)] \times \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left[\frac{ik(x^2 + y^2)}{2} \left(\frac{1}{z_i} + \frac{1}{z_p} - \frac{1}{z_0} \right) \right] \exp \left[-ikx \left(\frac{x_i}{z_i} - \frac{x_0}{z_0} \right) \right] \times \exp \left[-iky \left(\frac{y_i}{z_i} - \frac{y_0}{z_0} \right) \right] dx dy \right\} dx_0 dy_0 \quad (4a)$$

where

$$C_2 = \left(\frac{bA_r A_p}{\lambda^2 z_i z_0} \right) \exp[ik(z_p + z_i)] \exp[(ik/2z_i)(x_i^2 + y_i^2)].$$

For the case in which z_i , the distance from the hologram to the image plane, is such that $1/z_i = 1/z_0 - 1/z_p$, the exponential containing the quadratic term in the inner integral is 1 and the inner integral is then proportional to the two-dimensional delta function, $\delta(x_i/z_i - x_0/z_0, y_i/z_i -$

y_0/z_0). Using this expression in equation (4a) and evaluating the outer integral we see that the reconstructed field, $E_2(x_i, y_i)$, is proportional to $\exp[i\phi(x_i, y_i)]A^*(z_0x_i/z_i, z_0y_i/z_i)$, a phase factor times the complex conjugate of the object's transmission function. Hence, in a plane located at this distance z_i from the hologram, the light intensity forms an image of the object. Recall that in the construction of the hologram, the hologram plane was to the right of the illuminated object plane and the object distance z_0 was taken to be positive. Likewise, the distance z_p is positive when the hologram plane is to the right of the reconstruction point source. Consistent with this sign convention, for the reconstruction process, the image distance z_i is taken as positive when the image plane is to the right of the illuminated hologram plane, that is, for real images. Thus, we see that a given object point whose coordinates are x_0, y_0 in the object plane at z_0 will be imaged at x_i, y_i in the image plane at z_i , where $1/z_i = 1/z_0 - 1/z_p$, $x_i = (z_i/z_0)x_0$ and $y_i = (z_i/z_0)y_0$. In effect, the hologram acts as a thin positive lens of focal length z_0 . Furthermore, if z_p is infinite, that is, if the reconstruction beam is a plane wave, then $z_i = z_0$, $x_i = x_0$, $y_i = y_0$, and $E_2(x_i, y_i)$ equals a constant times $A^*(x_0, y_0)$. This means that a real image, with the same size and orientation as the object, is formed in a plane to the right of the hologram plane.

In a similar fashion to the above, if the term $A_r \exp(-ikz_0)E(x, y)$ from equation (2) is inserted for $T(x, y)$ in equation (3), then the contribution of this term to the reconstructed field is given by

$$E_3(x_i, y_i) = C_3 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x_0, y_0) \exp[(ik/2z_0)(x_0^2 + y_0^2)] \\ \times \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left[\frac{ik(x^2 + y^2)}{2} \left(\frac{1}{z_i} + \frac{1}{z_p} + \frac{1}{z_0} \right) \right] \exp \left[-ikx \left(\frac{x_i}{z_i} + \frac{x_0}{z_0} \right) \right] \right. \\ \left. \times \exp \left[-iky \left(\frac{y_i}{z_i} + \frac{y_0}{z_0} \right) \right] dx dy \right\} dx_0 dy_0 \quad (4b)$$

where

$$C_3 = \left(\frac{-bA_rA_p}{\lambda^2 z_i z_0} \right) \exp[ik(z_p + z_i)] \exp[(ik/2z_i)(x_i^2 + y_i^2)].$$

In this case, when the z_i for the image plane satisfies $1/z_i = -1/z_0 - 1/z_p$, the inner integral reduces to the two-dimensional delta function, $\delta(x_i/z_i + x_0/z_0, y_i/z_i + y_0/z_0)$ and the field $E_3(x_i, y_i)$ is proportional to $A(-z_0x_i/z_i, -z_0y_i/z_i)$ times a phase factor. Thus, a given object point whose coordinates are x_0, y_0 in the object plane at z_0 will be imaged at $x_i = -z_i x_0/z_0$ and $y_i = -z_i y_0/z_0$ in the plane at z_i . The relation $1/z_i = -1/z_0 - 1/z_p$ shows that the hologram acts as a negative lens of focal length $-z_0$. For z_p positive, z_i will be negative, meaning that a virtual image is formed in a plane to the left of the hologram. If z_p is infinite, then $z_i = -z_0$, $x_i = x_0$, $y_i = y_0$, and a virtual image, with the same size and orientation as the object, will be formed behind the hologram at the original position of the object.

To summarize, the two images formed upon reconstruction of a Gabor type hologram are found at image distances given by

$$1/z_i = 1/z_0 - 1/z_p \quad \text{with } x_i = z_i x_0/z_0 \text{ and } y_i = z_i y_0/z_0 \quad (5a)$$

with a lateral magnification

$$m = \Delta x_i / \Delta x_0 = z_i / z_0 = 1 + z_i / z_p \quad (5b)$$

and

$$1/z_i = -1/z_0 - 1/z_p \quad \text{with } x_i = -z_i x_0/z_0 \text{ and } y_i = -z_i y_0/z_0 \quad (6a)$$

with a lateral magnification

$$m = \Delta x_i / \Delta x_0 = -z_i / z_0 = 1 + z_i / z_p. \quad (6b)$$

From the above derivation for a planar object, one can see that the diffracted light from a given point P on the object interferes with the plane wave reference beam to form an interference pattern which, when recorded on film, acts in the reconstruction process as both a positive and a negative lens of focal lengths $+z_0$ and $-z_0$, respectively, for P . These two lenses are due to the sinusoidal Fresnel zone plate [10] formed by the interference pattern. Hence, the imaging ability of a hologram of an extended object can be explained by interpreting the hologram as a superposition of zone plates (see [7], p 595), one plate for each point of the object. For holograms of objects that are non-planar, that is, that have depth, z_0 will be different for each plane of the object and thus zone plates of different focal lengths will be superposed. This superposition will be illustrated in the next section, which deals with the computer construction of a hologram.

For completeness, it should be noted that if the reference beam, shown in figure 1, is not a plane wave but, rather, is due to a point source located at x_r, y_r and z_r where $z_r^2 \gg [(x_r - x)^2 + (y_r - y)^2]$, then the quadratic-surface approximation for the spherical waves is valid, and one has $E_r(x, y) = C_r \exp\{ik/2z_r(x^2 + y^2)\} \exp\{(-ik/z_r)(xx_r + yy_r)\}$ for the field of the reference wave at the hologram plane. Using this expression for the reference field in equations (2), (3), (4a) and (4b), the above derivation again yields two images in the reconstruction. One image will be at z_i given by

$$1/z_i = 1/z_0 - 1/z_r - 1/z_p \quad (7a)$$

and the other at

$$1/z_i = -1/z_0 + 1/z_r - 1/z_p. \quad (7b)$$

Furthermore, if the reconstruction point source is not on the z -axis but at x_p, y_p , where $z_p^2 \gg [(x_p - x)^2 + (y_p - y)^2]$, then a point on the object whose x and y coordinates are x_0 and y_0 will image at

$$x_i = z_i(x_0/z_0 - x_r/z_r - x_p/z_p) \quad y_i = z_i(y_0/z_0 - y_r/z_r - y_p/z_p) \quad (7c)$$

for the image whose z_i is given by equation (7a), and at

$$x_i = z_i(-x_0/z_0 + x_r/z_r - x_p/z_p) \quad y_i = z_i(-y_0/z_0 + y_r/z_r - y_p/z_p) \quad (7d)$$

for the image whose z_i is given by equation (7b). Except for the sign convention used, equations (7a)–(7d) are equivalent to those obtained for the off-axis hologram of point objects by Goodman (see [8], pp 217–8, equations (8-41) and (8-44)–(8.46)) and Collier (see [9], section 3-3, equations (3-27)–(3.29)), using a method which does not involve the Fourier transform. The author is not familiar with any prior derivation of these imaging relations using the Fourier transform approach presented here.

3. Construction of the computer-generated hologram

The construction of the computer-generated Fresnel hologram was accomplished using the Mathcad 6.0† mathematical software package for Windows‡ on a 486 PC with 16 MB RAM. The first step is to form an $N \times N$ matrix to represent the transmission function of an object in an x_0, y_0 plane which is a distance z_0 from the hologram plane, that is, the transmission function $A(x_0, y_0)$ discussed in the theory section. Similar to the procedure used in [4], the elements A_{y_0, x_0} of this transmission function matrix are made equal to 1 for points in the object plane that are to be transparent and 0 for points that are to be opaque. The coordinate indices x_0, y_0 of the matrix elements both range from 0 to $N - 1$ in integer steps where x_0 is the column index and y_0 the row index. Thus, when the matrix is displayed graphically the x_0 - and y_0 -axes will have their normal orientations. To evaluate the Fresnel diffraction field for this transmission function, we follow the expression in equation (1) and multiply the matrix

† Mathcad is a registered trademark of Mathsoft, Inc., 201 Broadway, Cambridge, MA 02139, USA.

‡ Windows 3.1 is a trademark of Microsoft Corporation.

elements A_{y_0, x_0} by $\exp[(ik/2z_0)(x_0^2 + y_0^2)]$ and, choosing the ratio $k/z_0 = 2\pi/N$, take the discrete Fourier transform of the resulting matrix. In Mathcad this is achieved by using the 'icfft' operator which, in turn, yields an $N \times N$ matrix. The reason for choosing the value of k/z_0 to be $2\pi/N$ is that for any k/z_0 the discrete Fourier transform of the matrix M whose elements are $A_{y_0, x_0} \exp[(ik/2z_0)(x_0^2 + y_0^2)]$ yields a matrix whose elements are

$$[\text{icfft}(M)]_{y,x} = \frac{1}{N} \sum_{x_0=0}^{N-1} \sum_{y_0=0}^{N-1} A_{y_0, x_0} \exp[(ik/2z_0)(x_0^2 + y_0^2)] \exp[(-i2\pi/N)(xx_0 + yy_0)] \quad (8)$$

where the coordinates x and y in the hologram plane also take the values $0, 1, \dots, N-1$. By choosing $k/z_0 = 2\pi/N$, or $\lambda z_0 = N$, one sees that the discrete transform of equation (8) and the optical transform contained in equation (1) will both have the same frequency spectrum as a function of the x, y coordinates of the hologram plane. This means that in the hologram plane, the computer-calculated Fresnel diffraction pattern for an object will replicate the optically obtained pattern with no scaling of the x - and y -axes required. The equating of λz_0 to N is a crucial step in this paper since the matrix size, as given by N , will determine z_0 , the planar object's distance from the hologram plane. Hence, for an object, which will be considered later, consisting of planar sections at different distances from the hologram plane, a transmission function matrix for each section will be formed using a different N for each matrix. It should be noted that, unlike some fast Fourier transform algorithms, that used by Mathcad does not require N to be a power of 2.

The last step in obtaining the matrix representing the diffraction field given by equation (1) is to multiply the elements of the Fourier transform matrix of equation (8) by $-i \exp[(ik/2z_0)(x^2 + y^2)]$, so that the elements of the diffraction field matrix become

$$E_{y,x} = -i \exp[(i\pi/N)(x^2 + y^2)][\text{icfft}\{M\}]_{y,x} \quad (9)$$

where the matrix elements of M are now $M_{y_0, x_0} = A_{y_0, x_0} \exp[(i\pi/N)(x_0^2 + y_0^2)]$ since we have substituted $2\pi/N$ for k/z_0 . The constant phase factor $\exp(ikz_0)$ which appears in the factor multiplying the integral in equation (1) does not have to be explicitly stated since, for Gabor type holograms, the calculation of the light intensity forming the hologram given by equation (2), shows that this phase factor, regardless of its value, will always be cancelled by the conjugate phase factor of the reference beam. Note also that the term $1/\lambda z_0$ which appears as a factor in equation (1) is naturally accounted for by the $1/N$ factor of the discrete Fourier transform of equation (8).

As seen from equation (2), to calculate the elements of the matrix I that represents the light intensity forming the hologram, we add the reference beam matrix elements $E_{r,y,x}$ to the $E_{y,x}$ of equation (9) and take the absolute value squared of this sum. In this case we let $E_{r,y,x} = A_r$, a constant, since as explained above, the phase factor $\exp(ikz_0)$ will be cancelled in the intensity calculation. Thus, $I_{y,x} = |E_{r,y,x} + E_{y,x}|^2 = A_r^2 + |E_{y,x}|^2 + A_r E_{y,x}^* + A_r E_{y,x}$. As discussed in the theory section, it is the last two terms of $I_{y,x}$ which produce the images of the object in the reconstruction process. By subtracting the elements $|E_{y,x}|^2$ from $I_{y,x}$ we obtain an intensity matrix I^H which, when displayed graphically, will serve as the hologram. That is,

$$I_{y,x}^H = |E_{r,y,x} + E_{y,x}|^2 - |E_{y,x}|^2. \quad (10)$$

The reconstruction of I^H will lack the contribution of the $|E_{y,x}|^2$ term which is detrimental to the quality of the reconstructed image of the object.

As an example of the above process let us consider an object which is composed of three points each at a different distance from the hologram plane. The first point is taken to have the coordinates (25, 25) in a 128×128 matrix; thus $\lambda z_{01} = N_1 = 128$. The elements $A1_{y_0, x_0}$ of the transmission matrix for the plane in which this point lies are defined as 0 for all y_0, x_0 except $A1_{25,25}$ which is made equal to 1. This corresponds to an opaque plane with a transparent point at (25, 25). The diffraction field matrix elements, $E1_{y,x}$, are then calculated using the matrix $A1$ in equation (9). The hologram intensity matrix elements $I^H(1)_{y,x}$ are then calculated using

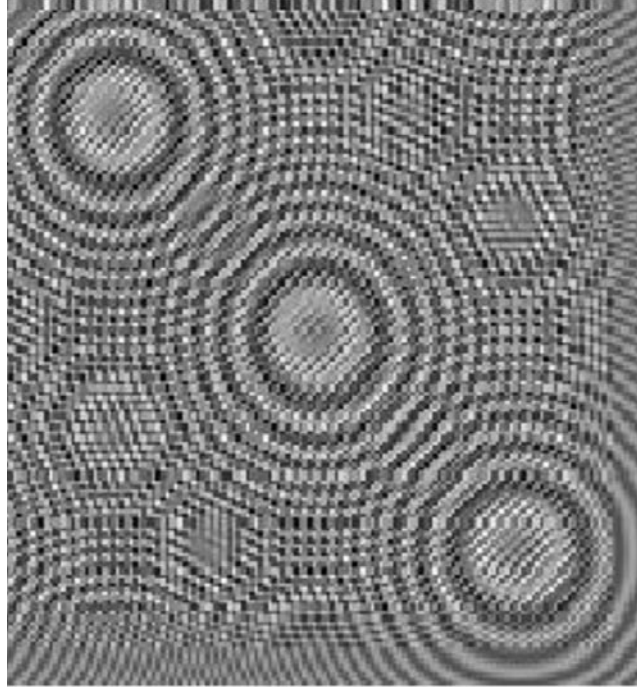


Figure 2. Graphic of the computer-generated Fresnel hologram matrix I^H for an object which consists of three points at different distances from the hologram plane.

equation (10) where the value for the constant A_r is taken to be 6. This emulates optically constructed holograms where the amplitude of the reference beam is made significantly larger than the amplitude of the diffracted field. The second point has the coordinates (70, 70) in a 136×136 matrix, so that $\lambda z_{02} = N_2 = 136$, and, in a fashion similar to the above, the intensity matrix elements $I^H(2)_{y,x}$ are computed for the transmission matrix A_2 . Lastly, the third point has the coordinates 115,115 in a 144×144 matrix and the elements $I^H(3)_{y,x}$ for A_3 are computed. The final hologram intensity matrix I^H is formed by superimposing the three computed intensities. In order to do this addition, the three matrices must, of course, be of the same size. To achieve this, we define the elements $I^H(1)_{128+j,128+k}$ to be zero for $j, k = 0, 1, \dots, 15$, and the elements $I^H(2)_{136+m,136+n}$ to be zero for $m, n = 0, 1, \dots, 7$. We now form the 144×144 matrix I^H , whose elements are

$$I^H_{y,x} = I^H(1)_{y,x} + I^H(2)_{y,x} + I^H(3)_{y,x} \quad (11)$$

which is the hologram matrix for the three point sources. As will be seen later, this device of adding zero-valued matrix elements to extend the size of the $I^H(1)$ and $I^H(2)$ matrices results in little distortion of the reconstructed images. A graphic of the I^H matrix is shown in figure 2 which was made using the three-dimensional surface plot capability of the Mathcad software. The plane of the page is the x - y plane with the origin in the upper left-hand corner. In this graphic the value of each matrix element $I^H_{y,x}$ is plotted at its x, y position as a shade of grey with the maximum value of I^H appearing as white and the minimum value as black. In the reconstruction process discussed below, it is convenient, although not necessary, to normalize the I^H matrix so that the values of its elements lie between 0 and 1, that is,

$$I^H(\text{normalized}) = [I^H - I^H_{\text{minimum}}] / [I^H_{\text{maximum}} - I^H_{\text{minimum}}].$$

Since $I^H(\text{normalized})$ is a linear function of I^H , Mathcad's surface plot software produces a graphic of $I^H(\text{normalized})$ which is identical to the graphic of I^H . For the rest of this paper

I^H will refer to I^H (normalized). This and all other graphics shown in this paper were printed on an HP LaserJet 5L printer. Figure 2 is a superposition of three Fresnel zone plates whose centres are at (25, 25) and (70, 70), and (115, 115) respectively in the hologram plane. That is, using the relation $\lambda z_0 = N$ effectively results in a one-to-one correspondence between the x_0, y_0 coordinates of the object plane and the x, y coordinates of the hologram plane and, thus, for these Fresnel diffraction computations, no shifting or scaling of the displayed pattern is required.

In figure 2 one notices what appear to be spurious circular fringe patterns. The author surmises that they are due to the fact that when one takes the discrete Fourier transform of a function, the frequencies present in the transform are such *as if* the object function is periodic with period N (see theorems 8.2 and 8.3, pp 242–3 of [11]). Of course, if the hologram were constructed optically the x and y coordinates would be continuous variables and not the discrete integer values used in the computer construction and, hence, these patterns would not appear. In any case, as will be demonstrated below, the presence of these fringes does not preclude the hologram from producing a satisfactory image when reconstructed optically.

Before discussing the reconstruction process, it is instructive to present a hologram for a more complicated object, since the hologram shown in figure 2 for the simple three-point object could have been obtained just as easily by direct calculation [12] without using the Fourier transform. For the new hologram, the object in the $N_1 = 128$ matrix is composed of the set of transparent points, (i.e. points at which the matrix elements $A1_{y_0, x_0}$ have the value 1), which form the letters HUYGENS in the upper third of the matrix. Likewise, the transparent letters YOUNG are placed in the middle of the $N_2 = 136$ matrix, and the letters FRESNEL are placed in the lower third of the $N_3 = 144$ matrix. These transmission matrices $A1, A2$ and $A3$ are shown in figures 3(a)–(c), respectively. It should be noted that to form these matrices, the three names were first printed out in the Paintbrush application of Windows and saved as a bitmap file. This graphic was then imported into Mathcad 6.0 where it was converted into the three matrices $A1, A2$ and $A3$, which consist of 0's and 1's. Those readers who are familiar with the history of holography will recognize that the hologram of these three names was one of the first made optically by Gabor [13]; a reprint of it can be found in Stroke's book on holography, page 216 [14]. However, in Gabor's work, unlike the situation presented here, the three names were all in a single plane.

The hologram for each matrix is computed as described above and their superposition forms the hologram I^H shown in figure 4. To make the hologram into a transparency suitable for optical reconstruction, the graphic of I^H was photographed and reduced to a 3×3 mm² square. This was achieved using a Polaroid MP-3 Land camera with type 665 positive/negative

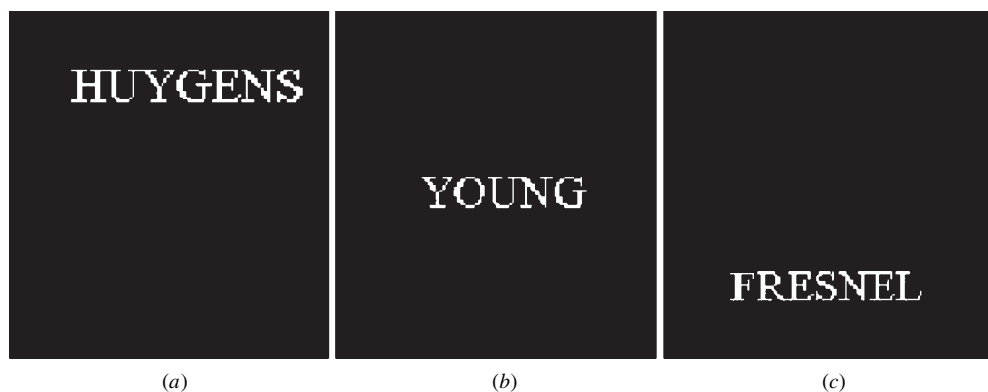


Figure 3. Parts (a), (b) and (c) are the object transmission matrices $A1, A2$ and $A3$, respectively, used in the construction of the hologram shown in figure 4.

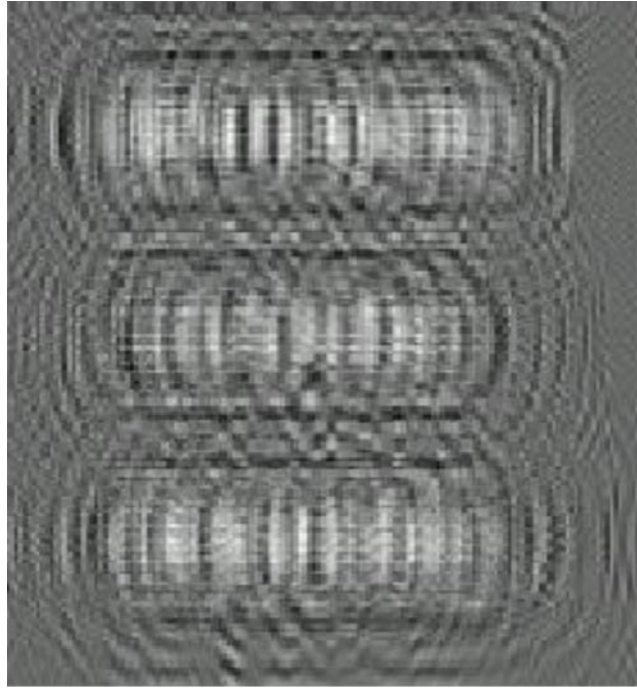


Figure 4. The computer-generated hologram matrix I^H for an object consisting of the three names shown in figure 3.

film. The negative is the hologram transparency. Hence, in those cases in which a positive transparency of the I^H is desired, the graphic of the negative hologram matrix, $I^{H(\text{neg})} = 1 - I^H$ was photographed. Since the matrix I^H is normalized, the matrix elements $I_{y,x}^{H(\text{neg})}$ will have values which lie between 0 and 1 and hence, the matrix $I^{H(\text{neg})}$ is normalized.

4. Replay of the computer-generated hologram

4.1. Computer reconstruction

Equation (3) of the reconstruction theory presented in section 2 is an expression for the field of the diffracted light, $E(x_i, y_i)$, in a plane at distance z_i from the hologram. It shows that this field is proportional to the Fourier transform of the product of the hologram's amplitude transmission function, $T(x, y)$, and the term $\exp[(ik/2)(x^2 + y^2)(1/z_p + 1/z_i)]$. As discussed in that section, the function $T(x, y)$ is taken to be $T(x, y) = a + bI^H(y, x)$. For a computer simulation of the reconstruction process, all continuous functions contained in equation (3) are replaced by their corresponding matrices. Thus, the matrix T has the elements $T_{y,x} = a + bI_{y,x}^H$. In order to have a meaningful comparison between the computer-reconstructed images presented here and the optical reconstructed images presented in the next section, the constants a and b appearing in the expression for T must be evaluated. This was accomplished using a test transparency made by photographing a pattern that consisted of a black and a white area. The fractional intensity transmitted by this transparency in the two regions was measured. By taking the square root of these measurements the maximum and minimum values of the amplitude transmission function T were found. Using these values and the corresponding maximum and minimum values of I^H (i.e. 1 and 0) in the relation for maximum T and minimum T respectively, the values of the constants were determined to be $a = 0.45$ and $b = 0.46$.

The first case of computer-simulated reconstruction we consider is that in which the hologram is illuminated from the left by a plane wave travelling to the right along the z -axis. In this case, z_p is infinite and, as can be seen from equation (5a) of section 2, the reconstructed real image of any point in the object plane at z_0 will lie in the image plane at $z_i = z_0$, to the right of the hologram, with x_i and y_i equal to x_0 and y_0 , respectively. Thus, the magnification m as given by equation (5b) is $m = 1$, and the image is erect and of the same size as the object. Likewise, equation (6a) shows that the virtual image will lie in the plane at $z_i = -z_0$, to the left of the hologram, again with $x_i = x_0$, and $y_i = y_0$. To realize the computer reconstruction of the real image, we recall that when taking the discrete Fourier transform of a matrix, it is the matrix size, as given by N , which determines the z coordinate of the plane through the relation $\lambda z = N$, in which the transform's spectrum equals that of the optical transform. For the hologram shown in figure 4, the real image of points which lie in the z_{01} object plane is formed in the image plane for which $z_{i1} = z_{01}$. For computer calculation of the elements $E_{y_i, x_i}^{\text{rec}}$ of the reconstruction field matrix in this image plane we use, in equation (3), $\lambda z_{i1} = \lambda z_{01} = N_1$, where, for this hologram, $N_1 = 128$. To compute the integral of equation (3) for this case of z_p infinite, we first define the range of the coordinate indices x, y as $0, 1, \dots, N_1 - 1$ and then form the matrix $M(1)$ whose elements are $M(1)_{y,x} = T_{y,x} \exp[(ik/2z_i)(x^2 + y^2)] = (0.45 + 0.46I_{y,x}^H) \exp[(i\pi/128)(x^2 + y^2)]$. Then, as seen from equation (3), the reconstruction field matrix, $E^{\text{rec}}(1)$, is obtained by taking the discrete Fourier transform of $M(1)$ and multiplying each element of the resulting matrix by a corresponding matrix element C_{y_i, x_i} where $C_{y_i, x_i} = (-A_p/128) \exp(ikz_p) \exp\{(ik/2z_i)(x_i^2 + y_i^2)\}$. That is, $E^{\text{rec}}(1)_{y_i, x_i} = C_{y_i, x_i} \text{icfft}(M(1))_{y_i, x_i}$. Thus, in the calculation of the reconstruction intensity matrix whose elements are $I^{\text{rec}}(1)_{y_i, x_i} = |E^{\text{rec}}(1)_{y_i, x_i}|^2$ we find that $I^{\text{rec}}(1)_{y_i, x_i}$ equals a constant times $|\text{icfft}(M(1))_{y_i, x_i}|^2$. As stated earlier, the $\text{icfft}(M(1))$ operation returns a matrix of the same size as $M(1)$. Thus, the reconstruction matrix $E^{\text{rec}}(1)$ is 128×128 and the x_i, y_i indices go from $0, 1, \dots, 127$. The graphic of $I^{\text{rec}}(1)$ is shown in figure 5(a), and one clearly sees a sharp image of the letters HUYGENS, which was the object in the z_{01} plane. The position, size and orientation of this image are those predicted by equations (5a) and (5b) for this case of z_p infinite.

In figure 5(a) we also see the out-of-focus images of the letters YOUNG and FRESNEL. This is because, for z_p infinite, the sharp reconstructed real images of these letters lie in the image planes $z_{i2} = z_{02}$ and $z_{i3} = z_{03}$, respectively, and since $z_{03} > z_{02} > z_{01}$, they lie further to the right of the hologram than the z_{i1} plane. For computer reconstruction of these two sharp images we use, in the above procedure, $\lambda z_{i2} = \lambda z_{02} = N_2$ with $x, y = 0, 1, \dots, N_2 - 1$ and $N_2 = 136$ for the letters YOUNG, and $\lambda z_{i3} = \lambda z_{03} = N_3$ with $x, y = 0, 1, \dots, N_3 - 1$ and

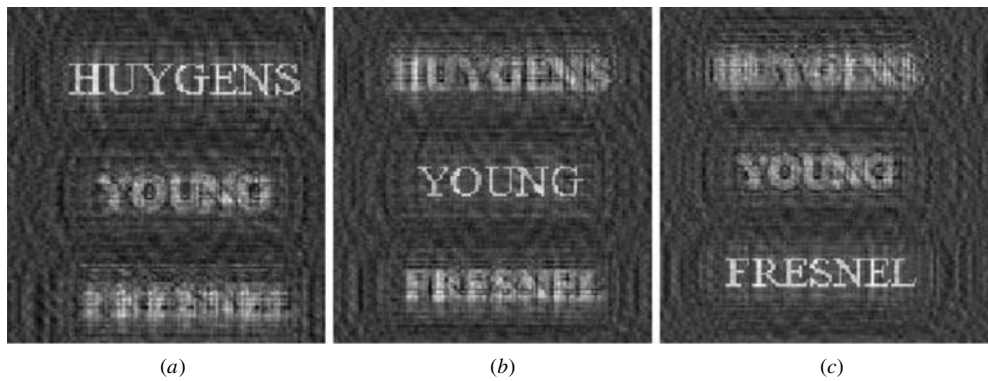


Figure 5. Parts (a), (b) and (c) are the computer-reconstructed image matrices $I^{\text{rec}}(1)$, $I^{\text{rec}}(2)$ and $I^{\text{rec}}(3)$, respectively, for the hologram shown in figure 4 for the case of z_p infinite.

$N_3 = 144$ for the letters FRESNEL, and calculate the reconstruction intensity matrices $I^{\text{rec}}(2)$ and $I^{\text{rec}}(3)$, respectively. The graphics of these matrices are shown in figures 5(b) and (c). In the next section, these results will be compared to those obtained by optical reconstruction.

Thus far, the computer reconstruction of real images, that is, images for which z_i is positive, has been discussed. Recall that for a virtual image z_i is negative and the image is formed in a plane to the left of the hologram plane. For the case of z_p infinite considered above, equation (6a) of the theory predicts that a reconstructed virtual image will lie in the image plane at $z_i = -z_0$. Thus, we now use $\lambda z_i = -\lambda z_0 = -N$ in equation (3) and the subsequent relations of the reconstruction procedure discussed above. This has the effect of changing i to $-i$ in these relations. In turn, this means that the `icfft` Mathcad operator is replaced by the `cfft` operator when taking the Fourier transform. Since we are concerned with the reconstructed intensity, $I^{\text{rec}} = |E^{\text{rec}}(1)|^2$, the matrices $I^{\text{rec}}(1)$, $I^{\text{rec}}(2)$, and $I^{\text{rec}}(3)$ obtained for the virtual images for this case of z_p infinite, are identical to the corresponding matrices for the real images shown in figures 5(a)–(c).

An interesting case which shows that the hologram behaves as a negative lens of focal length $-z_0$ (see equation (6a)), is given by computer simulation of the reconstruction for $z_p = -z_0/2$, that is, the case of illuminating the hologram from the left with converging light. Using this value in equation (6a) we find that $z_i = z_0$, which is positive. Hence, once again, a real image is formed at this distance to the right of the hologram. However, for this case of $x_i = -x_0$ and $y_i = -y_0$, the magnification m as given by equation (6b) is $m = -1$ and the image is of the same size as the object but inverted. For computer reconstruction of this real image of the z_{01} object plane for the hologram I^H shown in figure 4, we use $z_p = -z_{01}/2$ and $\lambda z_{i1} = \lambda z_{01} = N_1 = 128$ in equation (3). For this case, the term $\exp[(ik/2z_p)(x^2 + y^2)]$ in the integral of equation (3) is not equal to 1 as it was for z_p infinite. Consequently, the elements of the matrix $M(1)$ discussed above are now $M(1)_{y,x} = \exp[(ik/2z_p)(x^2 + y^2)]T_{y,x} \exp[(ik/2z_i)(x^2 + y^2)]$ which becomes $M(1)_{y,x} = (0.45 + 0.46I_{y,x}^H) \exp[-(i\pi/128)(x^2 + y^2)]$ when the expression for $T_{y,x}$ in terms of $I_{y,x}^H$ and the values $\lambda z_p = -N_1/2$ and $\lambda z_i = N_1$ are substituted. As before, to calculate the reconstruction field E^{rec} we first define the range of the indices x and y as $0, 1, \dots, N_1 - 1$ and then take the Fourier transform of the matrix $M(1)$. That is, $E_{y_1,x_1}^{\text{rec}} = C_{y_1,x_1} \text{icfft}(M(1))_{y_1,x_1}$ and the elements C_{y_1,x_1} are such that the reconstruction intensity matrix $I^{\text{rec}}(1)_{y_1,x_1} = |E^{\text{rec}}(1)_{y_1,x_1}|^2$ can again be written as a constant times $|\text{icfft}(M(1))_{y_1,x_1}|^2$. The graphic of I^{rec} for this case of $z_p = -z_{01}/2$ is shown in figure 6(a) and the inverted image predicted by equation (6b) of the theory is obtained. For $z_p = -z_{01}/2$ equation (6a) also predicts that the in-focus image

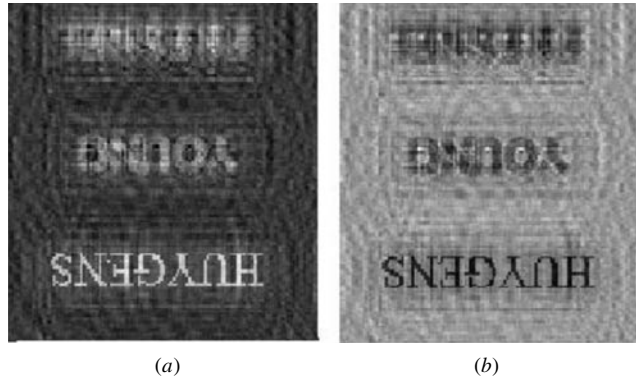


Figure 6. Computer-reconstructed image matrices, I^{rec} , for the case $z_p = -z_{01}$. The reconstructed image shown in (a) was obtained using the positive hologram matrix I^H shown in figure 4. The image shown in (b) was obtained using the negative hologram matrix, $I^{H(\text{neg})}$.

distances z_{i2} and z_{i3} for the object planes at z_{02} and z_{03} , respectively, are such that $z_{i3} < z_{i2} < z_{i1}$ and hence lie in planes to the left of the z_{i1} plane, which accounts for the out-of-focus images of YOUNG and FRESNEL seen in figure 6(a).

For completeness, it should be noted that, for the above computer reconstructions, if one uses the normalized negative hologram matrix, $I^{H(\text{neg})}$, in the expression for the transmission matrix T , the calculated I^{rec} yields a negative image of the object. Since the matrix elements of $I^{H(\text{neg})}$ (normalized) are given by $I_{y,x}^{H(\text{neg})} = 1 - I_{y,x}^H$ they have values between 0 and 1. Thus the previously determined values for the constants a and b which appear in the equation for T are also applicable in this case. An example of the I^{rec} which results when the negative hologram matrix is used is shown in figure 6(b) for which the reconstruction calculations described above for figure 6(a) were repeated using $I^{H(\text{neg})}$. The occurrence of the negative image in these in-line holograms is due to the interference of the image field with the undiffracted background field (see section 8-3 of [8] and section 2-5 of [9]). For holograms made with an off-axis reference beam, not presented in this paper, the interference between the image and background fields does not occur in the reconstruction. Thus, one always obtains, in off-axis holography, a positive image using either a positive or negative transparency.

In the next section, the optically reconstructed images corresponding to all of the above computer-simulated reconstructions will be presented for comparison.

4.2. Optical reconstruction

The photographic reduction of the graphic of the matrix I^H into a small square hologram transparency suitable for optical reconstruction was discussed at the end of section 3. Recall that in order to produce a positive transparency, which will result in positive images upon reconstruction, the graphic of the matrix $I^{H(\text{neg})}$ was photographed. The first case of optical reconstruction we consider is that in which the hologram transparency is illuminated from the left by diverging laser light and forms a real image. This was achieved by passing the beam from a He-Ne laser, of wavelength $\lambda = 632.8$ nm, through a short focal length positive lens as shown in figure 7. The general experimental procedure was as follows. For a given reconstruction distance z_p , a screen was placed to the right of the hologram and the real image plane for each of the objects used in the construction of the hologram was determined. The image distances z_i were then measured. A specific experiment, using a 3 mm positive transparency of the hologram matrix I^H shown in figure 4, was to adjust the distance z_p so that the size of the image formed of one of the names was equal to the size of the corresponding computer-reconstructed image shown in figure 5. The optically reconstructed image was recorded directly on film, without the use of a camera, by replacing the screen with the camera back film holder containing the Polaroid P/N 655 film pack. This process was repeated for the other two names and optically scanned copies of the Polaroid photographs of the three images are shown in figures 8(a)–(c). A close examination of the optically reconstructed images shown in figure 8 and the corresponding computer-reconstructed images of figure 5 reveals that, in addition to the legible letters forming the names, much of the substructure appearing

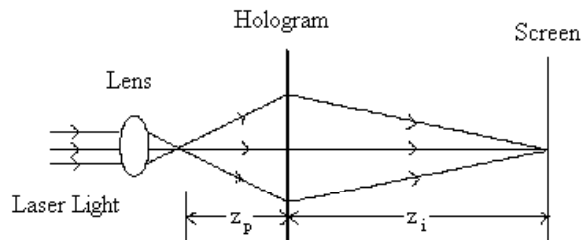


Figure 7. Apparatus configuration used in the optical reconstruction of a hologram.

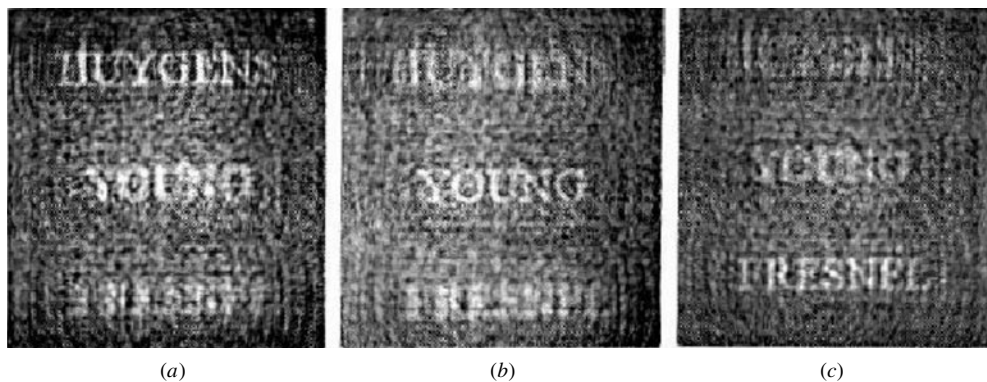


Figure 8. Optically reconstructed real images using a positive transparency of the hologram shown in figure 4. The value of z_p was adjusted for each image so that the image size is that of the images shown in figure 5.

in figure 5 is also present in figure 8 and thus is not an artifact of the photographic process. It should be noted that for these small holograms, the entire image cannot be seen, all at once, when directly viewed by eye. This is because the angle subtended by the light rays forming each point of the image is also small resulting in little divergence.

When the three pairs of measured values of z_p and z_i used to form the images shown in figure 8 are inserted in equation (5b), the magnification for each figure is obtained. The magnifications so computed for the names HUYGENS and YOUNG are, respectively, about 143/127 and 143/135 times the magnification for the name FRESNEL. This arises from the fact that the linear sizes of the matrices for the names HUYGENS, YOUNG, and FRESNEL, which together comprise the hologram matrix shown in figure 4, are 127, 135, and 143, respectively.

By varying the value of z_p and measuring the resulting image distances z_{i1} , z_{i2} and z_{i3} for each z_p , a graph of $1/z_i$ versus $1/z_p$ can be constructed for each of the three images. The resulting graphs, shown in figure 9, are straight lines as predicted by equation (5a), and their slopes are each within a few per cent of the predicted value of -1 . Furthermore, for each graph the intercept on the $1/z_i$ axis should equal $1/z_0$. The values for these effective object distances obtained from the intercepts are $z_{01} = 8.68$ cm, $z_{02} = 9.29$ cm, and $z_{03} = 9.81$ cm. These values represent the distances of the three object planes to the hologram plane had this hologram transparency been constructed optically using coherent light of 632.8 nm wavelength. These distances can be computed from theory by considering the number of circular fringes contained in a sinusoidal zone plate. The details of this calculation are available from the author and the computed values of the z_0 's so obtained are within 1% of their respective values given above.

Equation (6a) of the theory predicts that the hologram should exhibit the properties of a negative lens and hence, in the reconstruction process with z_p positive, produce a virtual image to the left of the hologram. For these holograms this virtual image is difficult to see and photograph due to its small size and the intensity of the transmitted undiffracted light. However, if one illuminates the hologram with converging light, that is z_p negative, equation (6a) predicts that if the absolute value of z_p is less than z_0 , a real image will result. This corresponds to the computer reconstruction shown in figure 6(a). To optically reproduce this reconstruction, we illuminated the hologram transparency with converging light by interposing a positive lens between the lens and the hologram shown in figure 7. For this case of z_p negative, the image distance z_i is positive and the magnification m is negative so that a real, inverted image is formed. Figure 10(a) shows this optically reconstructed real image of the z_{01} object plane. Once again, the image was formed directly on the film in its holder without the use of a camera. The value of z_p was adjusted so that the size of the image is comparable to that obtained in the computer reconstruction shown in figure 6(a). Inserting the measured values of $z_p = -8.4$ cm

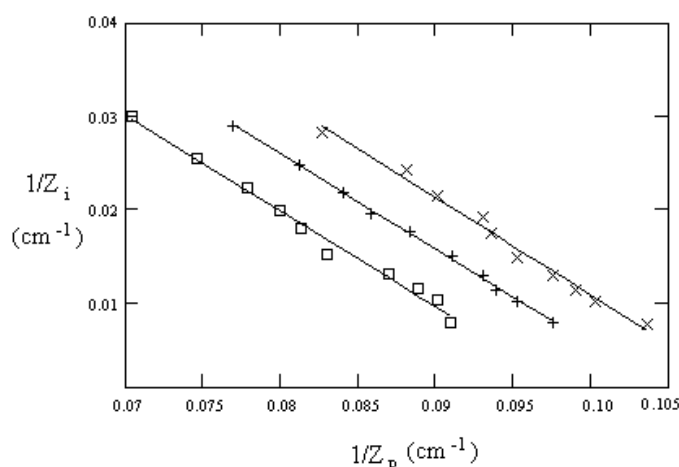


Figure 9. Graphs of the reciprocal image distances $1/z_{i1}$, $1/z_{i2}$, and $1/z_{i3}$ for the three object planes at distances z_{01} , z_{02} , and z_{03} , respectively, versus the reciprocal of the reconstruction distance $1/z_p$. Data points labelled \times , $+$ and \square are for the reciprocal image distances $1/z_{i1}$, $1/z_{i2}$, and $1/z_{i3}$, respectively. The lines are drawn using a least-squares fit.

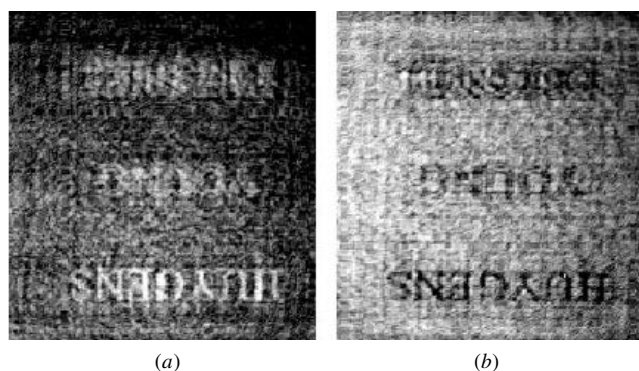


Figure 10. Optically reconstructed inverted real images of the object plane at z_{01} for the case of negative z_p . The reconstructions shown were achieved using a positive transparency of the hologram in (a) and a negative hologram transparency in (b).

and $z_{i1} = 189.5$ cm, used to obtain the image shown in figure 10a, into equation (6a) one obtains a value for $z_{01} = 8.79$ cm which again is in excellent agreement with the z_{01} value predicted by the theory. Furthermore, it was experimentally verified that, for $|z_p| < z_{01}$, the image distances z_{i3} and z_{i2} are also positive with $z_{i3} < z_{i2} < z_{i1}$. This result is predicted by equation (6a) since $z_{03} > z_{02} > z_{01}$. Thus, these real inverted images appear in the reverse order, from the hologram plane, to the reconstructed erect real images obtained using z_p positive and greater than z_{03} . (See figure 9 for the relative z_i values for a given z_p .) The effect of the images reversing their order and orientation as the hologram's reconstruction illumination is changed from converging to diverging light is striking and easily observed.

The image shown in figure 10(a) was achieved using a positive transparency. In figure 10(b) we show the reconstructed image obtained using the negative hologram transparency for the same z_p and, as expected from the theory, a negative image results. By comparing the images shown in figures 6(a) and 6(b) with those shown in figures 10(a) and 10(b), respectively, one again sees the general agreement in the characteristics of the

corresponding computer and optically reconstructed images.

Lastly, it should be mentioned that when the negative hologram transparency is illuminated from the left using converging light, a simple experiment can be performed which dramatically demonstrates that the negative image shown in figure 10(b) is the result of interference between the image field and the undiffracted background field transmitted by the hologram. The experiment consists of carefully positioning the tip of a fine wire at the point, a distance z_p to the right of the hologram, where the undiffracted converging light comes to focus before diverging. With the background field obscured in this manner, a positive image is now observed in the place of the negative image of figure 10(b).

5. Summary and conclusions

This paper presents the theory of Fresnel holography in terms of the Fourier transform of a modified aperture function, which allows for the simulation of this type of holography on a personal computer. The computer generation of a hologram for a simple object having depth is achieved by computing a hologram in matrix form for each plane of the object using the prescription $\lambda z_0 = N$, which relates the position of the object plane at z_0 to the matrix size N . The computer-generated hologram for the entire object is obtained by superimposing these matrices using a method which makes them equal in size. The reconstruction of the hologram is performed both on the computer and optically, and the predictions of the theory concerning the position, size, orientation, and nature (real or virtual) of the image are quantitatively verified.

This work should be of interest to researchers in holography, as well as to teachers of modern optics. In addition, it seems that a natural extension of this work would be to use the techniques presented for computer generation of Fresnel holograms for the case of an off-axis reference beam. Moreover, it might be possible, using an optical scanner, to generate holograms of more complicated objects. That is, by importing the scanned image of an object in bitmap form into the mathematical software program, the resulting matrix could serve as the object's transmission function in the hologram formation procedures presented in this paper.

Acknowledgments

It is with great appreciation that I thank my colleague, Dr Donald Gelman, for his thorough reading of this manuscript and his many useful comments and suggestions. I should also like to thank the Research Committee of the C W Post Campus of Long Island University for the released time made available to me for the completion of this work.

References

- [1] Wilson R G, McCreary S M and Thompson J L 1992 Optical transforms in three-space: simulations with a PC *Am. J. Phys.* **60**, 49–56
- [2] Chen X, Huang J and Loh E 1988 Two-dimensional fast Fourier transform and pattern processing with IBM PC *Am. J. Phys.* **56** 747–9
- [3] Dodds S A 1990 An optical diffraction experiment for the advanced undergraduate *Am. J. Phys.* **58** 663–8
- [4] Trester S 1996 Computer-simulated holography and computer-generated holograms *Am. J. Phys.* **64** 472–8
- [5] Mas D, Garcia J, Ferreira C, Bernardo L M and Marinho F 1999 Fast algorithms for the free space diffraction pattern calculation *Opt. Commun.* **164** 233–45
- [6] Trester S 1999 Computer-simulated Fresnel diffraction using the Fourier transform *Comput. Sci. Engng* **1** 77–83
- [7] Hecht E 1987 *Optics* 2nd edn (Reading, MA: Addison-Wesley)
- [8] Goodman J W 1968 *Introduction to Fourier Optics* (New York: McGraw-Hill)
- [9] Collier R J, Burckhardt C B and Lin L H 1971 *Optical Holography* (New York: Academic)
- [10] Klein M V 1970 *Optics* (Chichester: Wiley) pp 386–8
- [11] Weaver H J 1989 *Theory of Discrete and Continuous Fourier Analysis* (New York: Wiley)
- [12] Dittmann H and Schneider W B 1992 Simulated holograms *Phys. Teacher* **30**, 244–8
- [13] Gabor D 1949 Microscopy by reconstructed wave-fronts *Proc. R. Soc. A* **197** 454–87
- [14] Stroke G W 1966 *An Introduction to Coherent Optics and Holography* (New York: Academic)