BASICS OF HOLOGRAPHY

P. HARIHARAN
School of Physics, University of Sydney, Australia

CAMBRIDGE
UNIVERSITY PRESS

1

Holographic imaging

A hologram is usually recorded on a photographic plate or a flat piece of film, but produces a three-dimensional image. In addition, making a hologram does not involve recording an image in the conventional sense. To resolve this apparent paradox and understand how holography works, we have to start from first principles.

In conventional imaging techniques, such as photography, what is recorded is merely the intensity distribution in the original scene. As a result, all information about the optical paths to different parts of the scene is lost.

The unique characteristic of holography is the idea of recording both the phase and the amplitude of the light waves from an object. Since all recording materials respond only to the intensity in the image, it is necessary to convert the phase information into variations of intensity. Holography does this by using coherent illumination and introducing, as shown in fig. 1.1, a reference beam derived from the same source. The photographic film records the interference pattern produced by this reference beam and the light waves scattered by the object.

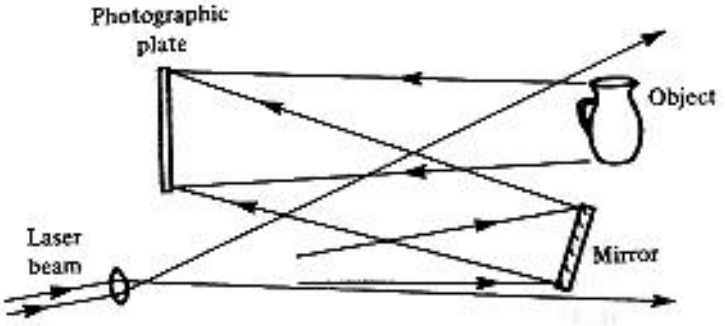


Fig. 1.1. Hologram recording: the interference pattern produced by the reference wave and the object wave is recorded.

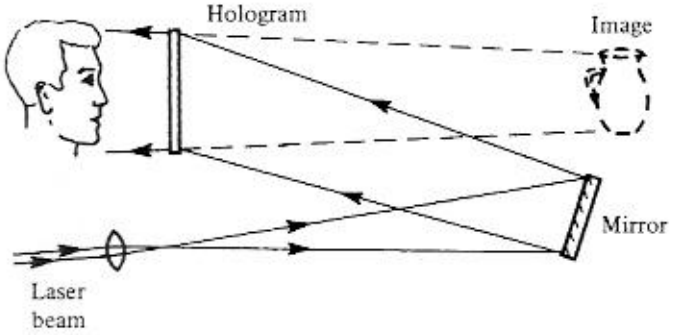


Fig. 1.2. Image reconstruction: light diffracted by the hologram reconstructs the object wave.

Since the intensity at any point in this interference pattern also depends on the phase of the object wave, the resulting recording (the hologram) contains information on the phase as well as the amplitude of the object wave. If the hologram is illuminated once again with the original reference wave, as shown in fig. 1.2, it reconstructs the original object wave.

An observer looking through the hologram sees a perfect three-dimensional image. This image exhibits, as shown in figs. 1.3 and 1.4, all the effects of perspective, and depth of focus when photographed, that characterized the original object.

1.1 Early development

In Gabor's historical demonstration of holographic imaging [Gabor, 1948], a transparency consisting of opaque lines on a clear background was illuminated with a collimated beam of monochromatic light, and the interference pattern produced by the directly transmitted beam (the reference wave) and the light scattered by the lines on the transparency was recorded on a photographic plate. When the hologram (a positive transparency made from this photographic negative) was illuminated with the original collimated beam, it produced two diffracted waves, one reconstructing an image of the object in its original location, and the other, with the same amplitude but the opposite phase, forming a second, *conjugate* image.

A major drawback of this technique was the poor quality of the reconstructed image, because it was degraded by the conjugate image, which was superimposed on it, as well as by scattered light from the directly transmitted beam.

The twin-image problem was finally solved when Leith and Upatnieks

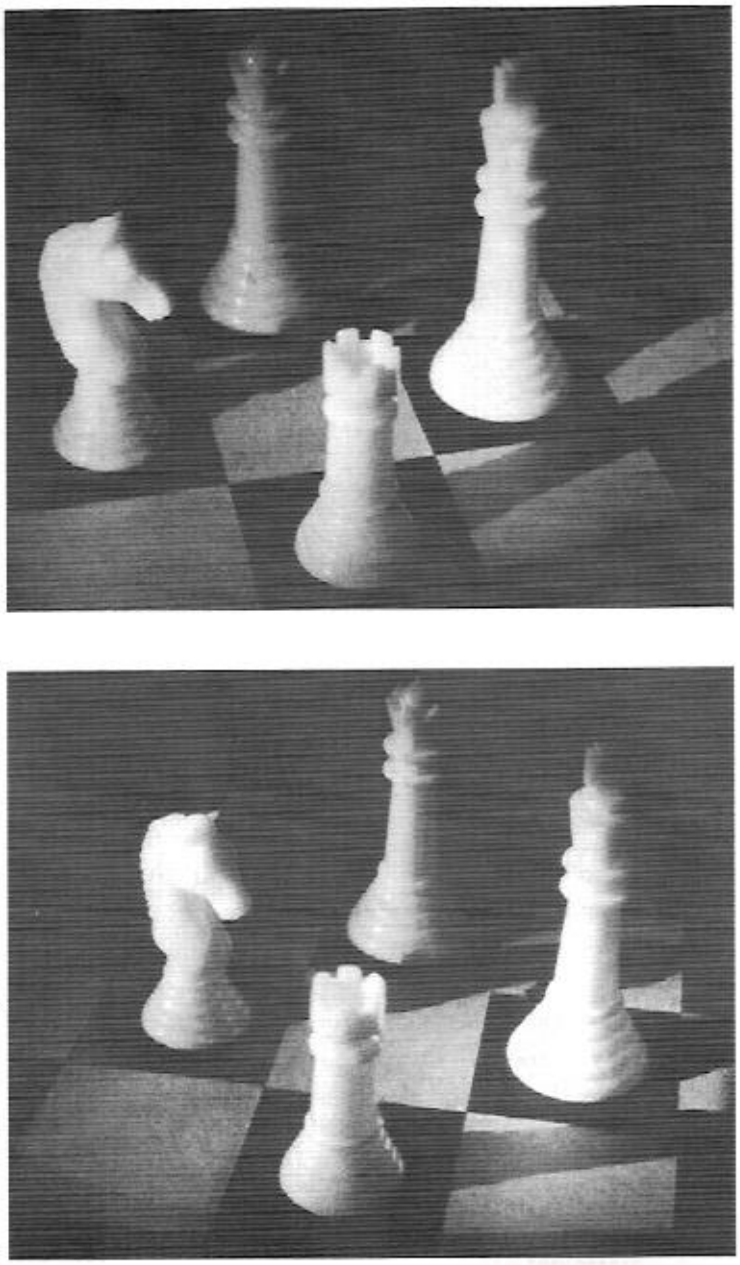


Fig. 1.3. Views from different angles of the image reconstructed by a hologram, showing changes in perspective.

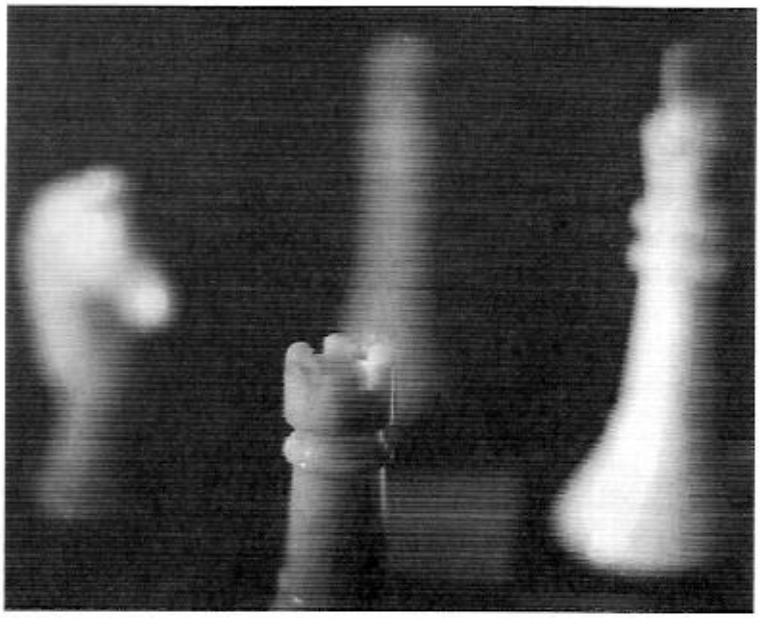


Fig. 1.4. Picture of the reconstructed image taken with the camera lens wide open (f/1.8), showing the effect of limited depth of focus.

[1962, 1963, 1964] developed the off-axis reference beam technique shown schematically in figs. 1.1 and 1.2. They used a separate reference wave incident on the photographic plate at an appreciable angle to the object wave. As a result, when the hologram was illuminated with the original reference beam, the two images were separated by large enough angles from the directly transmitted beam, and from each other, to ensure that they did not overlap.

The development of the off-axis technique, followed by the invention of the laser, which provided a powerful source of coherent light, resulted in a surge of activity in holography that led to several important applications.

1.2 The in-line hologram

We consider the optical system shown in fig. 1.5 in which the object (a transparency containing small opaque details on a clear background) is illuminated by a collimated beam of monochromatic light along an axis normal to the photographic plate.

The light incident on the photographic plate then contains two components. The first is the directly transmitted wave, which is a plane wave whose amplitude and phase do not vary across the photographic plate. Its complex amplitude (see Appendix A) can, therefore, be written as a real constant r. The

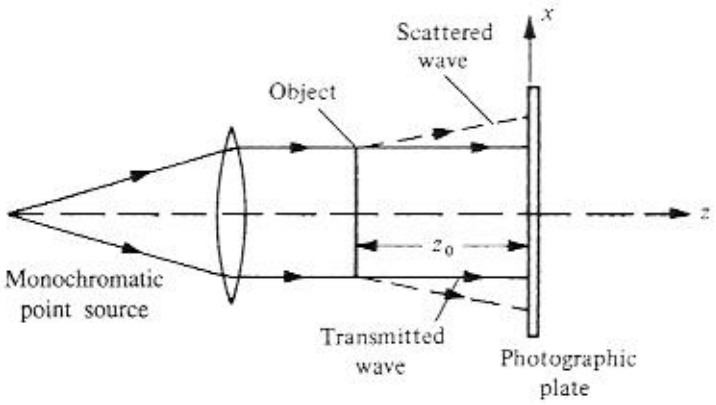


Fig. 1.5. Optical system used to record an in-line hologram.

second is a weak scattered wave whose complex amplitude at any point (x, y) on the photographic plate can be written as $o(x, y)$, where $|o(x, y)| \ll r$.

Since the resultant complex amplitude is the sum of these two complex amplitudes, the intensity at this point is

$$\begin{aligned} I(x, y) &= |r + o(x, y)|^2, \\ &= r^2 + |o(x, y)|^2 + ro(x, y) + ro^*(x, y), \end{aligned} \tag{1.1}$$

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

A 'positive' transparency (the hologram) is then made by contact printing from this recording. If we assume that this transparency is processed so that its amplitude transmittance (the ratio of the transmitted amplitude to that incident on it) can be written as

$$\mathbf{t} = \mathbf{t}_0 + \beta T I, \tag{1.2}$$

where $\mathbf{t}_0$ is a constant background transmittance, T is the exposure time and β is a parameter determined by the photographic material used and the processing conditions, the amplitude transmittance of the hologram is

$$\mathbf{t}(x, y) = \mathbf{t}_0 + \beta T[r^2 + |o(x, y)|^2 + ro(x, y) + ro^*(x, y)]. \tag{1.3}$$

Finally, the hologram is illuminated, as shown in fig. 1.6, with the same collimated beam of monochromatic light used to make the original recording. Since the complex amplitude at any point in this beam is, apart from a constant factor, the same as that in the original reference beam, the complex amplitude transmitted by the hologram can be written as

$$\begin{aligned} u(x, y) &= r\mathbf{t}(x, y) \\ &= r(\mathbf{t}_0 + \beta T r^2) + \beta T r|o(x, y)|^2 \\ &\quad + \beta T r^2 o(x, y) + \beta T r^2 o^*(x, y). \end{aligned} \tag{1.4}$$

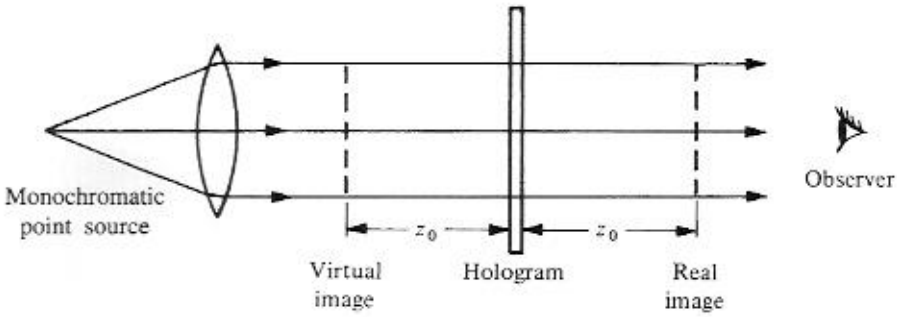


Fig. 1.6. Optical system used to reconstruct the image with an in-line hologram, showing the formation of the twin images.

The right-hand side of (1.4) contains four terms. The first of these, $r(\mathbf{t}_0 + \beta Tr^2)$, which represents a uniformly attenuated plane wave, corresponds to the directly transmitted beam.

The second term, $\beta Tr\,|o(x, y)|^2$, is extremely small, compared to the other terms, and can be neglected.

The third term, $\beta Tr^2 o(x, y)$, is, except for a constant factor, identical with the complex amplitude of the scattered wave from the object and reconstructs an image of the object in its original position. Since this image is formed behind the hologram, and the reconstructed wave appears to diverge from it, it is a virtual image.

The fourth term, $\beta Tr^2 o^*(x, y)$, represents a wave similar to the object wave, but with the opposite curvature. This wave converges to form a real image (the conjugate image) at the same distance in front of the hologram.

With an in-line hologram, an observer viewing one image sees it superimposed on the out-of-focus twin image as well as a strong coherent background. Another drawback is that the object must have a high average transmittance for the second term on the right-hand side of (1.4) to be negligible. As a result, it is possible to form images of fine opaque lines on a transparent background, but not *vice versa*. Finally, the hologram must be a 'positive' transparency. If the initial recording is used directly, β in (1.2) is negative, and the reconstructed image resembles a photographic negative of the object.

1.3 Off-axis holograms

To understand the formation of an image by an off-axis hologram, we consider the recording arrangement shown in fig. 1.7, in which (for simplicity) the reference beam is a collimated beam of uniform intensity, derived from the same source as that used to illuminate the object.

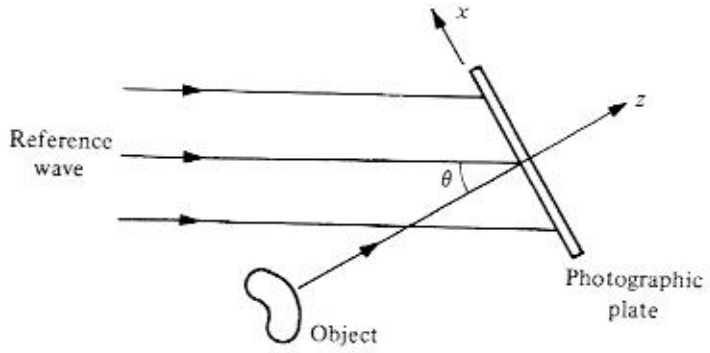


Fig. 1.7. The off-axis hologram: recording.

The complex amplitude at any point (x, y) on the photographic plate due to the reference beam can then be written (see Appendix A) as

$$r(x, y) = r \exp(\mathrm{i}2\pi\xi x), \tag{1.5}$$

where $\xi = (\sin\theta)/\lambda$, since only the phase of the reference beam varies across the photographic plate, while that due to the object beam, for which both the amplitude and phase vary, can be written as

$$o(x, y) = |o(x, y)| \exp[-\mathrm{i}\phi(x, y)]. \tag{1.6}$$

The resultant intensity is, therefore,

$$\begin{aligned} I(x, y) &= |r(x, y) + o(x, y)|^2 \\ &= |r(x, y)|^2 + |o(x, y)|^2 \\ &\quad + r|o(x, y)| \exp[-\mathrm{i}\phi(x, y)] \exp(-\mathrm{i}2\pi\xi x) \\ &\quad + r|o(x, y)| \exp[\mathrm{i}\phi(x, y)] \exp(\mathrm{i}2\pi\xi x) \\ &= r^2 + |o(x, y)|^2 + 2r|o(x, y)| \cos[2\pi\xi x + \phi(x, y)]. \end{aligned} \tag{1.7}$$

The amplitude and phase of the object wave are encoded as amplitude and phase modulation, respectively, of a set of interference fringes equivalent to a carrier with a spatial frequency of ξ.

If, as in (1.2), we assume that the amplitude transmittance of the processed photographic plate is a linear function of the intensity, the resultant amplitude transmittance of the hologram is

$$\begin{aligned} \mathbf{t}(x, y) &= \mathbf{t}_0' + \beta T |o(x, y)|^2 \\ &\quad + \beta T r |o(x, y)| \exp[-\mathrm{i}\phi(x, y)] \exp(-\mathrm{i}2\pi\xi x) \\ &\quad + \beta T r |o(x, y)| \exp[\mathrm{i}\phi(x, y)] \exp(\mathrm{i}2\pi\xi x), \end{aligned} \tag{1.8}$$

where $\mathbf{t}_0' = \mathbf{t}_0 + \beta T r^2$ is a constant background transmittance.

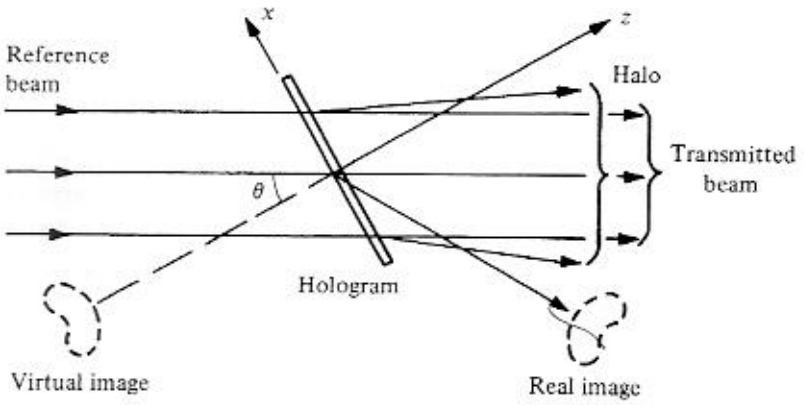


Fig. 1.8. The off-axis hologram: image reconstruction.

When the hologram is illuminated once again with the original reference beam, as shown in fig. 1.8, the complex amplitude of the transmitted wave can be written as

$$\begin{aligned} u(x, y) &= r(x, y)\mathbf{t}(x, y) \\ &= \mathbf{t}_0' r \exp(\mathrm{i}2\pi\xi x) + \beta T r |o(x, y)|^2 \exp(\mathrm{i}2\pi\xi x) \\ &\quad + \beta T r^2 o(x, y) + \beta T r^2 o^*(x, y) \exp(\mathrm{i}4\pi\xi x). \end{aligned} \tag{1.9}$$

The first term on the right-hand side of (1.9) corresponds to the directly transmitted beam, while the second term yields a halo surrounding it, with approximately twice the angular spread of the object. The third term is identical to the original object wave, except for a constant factor $\beta T r^2$, and produces a virtual image of the object in its original position. The fourth term corresponds to the conjugate image which, in this case, is a real image. If the offset angle of the reference beam is made large enough, the virtual image can be separated from the directly transmitted beam and the conjugate image.

In this arrangement, corresponding points on the real and virtual images are located at equal distances from the hologram, but on opposite sides of it. Since the depth of the real image is inverted, it is called a pseudoscopic image, as opposed to the normal, or orthoscopic, virtual image. It should also be noted that the sign of β only affects the phase of the reconstructed image, so that a 'positive' image is always obtained, even if the hologram recording is a photographic negative.

1.4 Fourier holograms

An interesting hologram recording configuration is one in which the complex amplitudes of the waves that interfere at the hologram are the Fourier

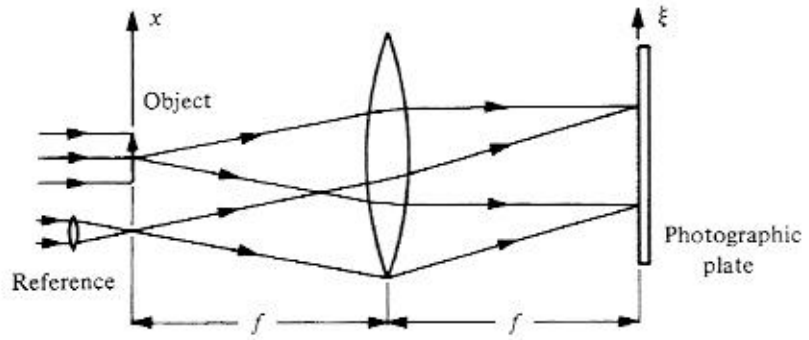


Fig. 1.9. Optical system used to record a Fourier hologram.

transforms (see Appendix B) of the complex amplitudes of the original object and reference waves. Normally, this implies an object of limited thickness, such as a transparency.

To record a Fourier hologram, the object transparency is placed in the front focal plane of a lens, as shown in fig. 1.9, and illuminated with a collimated beam of monochromatic light. The reference beam is derived from a point source also located in the front focal plane of the lens. The hologram is recorded on a photographic plate placed in the back focal plane of the lens [Vander Lugt, 1964].

If the complex amplitude of the wave leaving the object plane is $o(x, y)$, its complex amplitude at the photographic plate located in the back focal plane of the lens is

$$O(\xi,\eta)=\mathcal{F}\{o(x, y)\}. \tag{1.10}$$

The reference beam is derived from a point source also located in the front focal plane of the lens. If $\delta(x+b, y)$ is the complex amplitude of the wave leaving this point source, the complex amplitude of the reference wave at the photographic plate can be written as

$$R(\xi,\eta)=\exp(-\mathrm{i}2\pi\xi b). \tag{1.11}$$

The intensity in the interference pattern produced by these two waves is, therefore,

$$\begin{aligned} I(\xi,\eta)=1+|O(\xi,\eta)|^2+O(\xi,\eta)\exp(\mathrm{i}2\pi\xi b)\\ +O^*(\xi,\eta)\exp(-\mathrm{i}2\pi\xi b). \end{aligned} \tag{1.12}$$

To reconstruct the image, the processed hologram is replaced in the front focal plane of the lens, as shown in fig. 1.10, and illuminated with a collimated beam of monochromatic light.

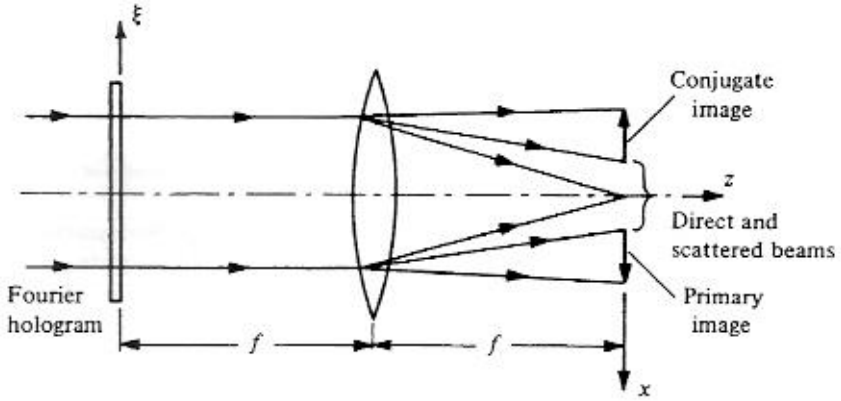


Fig. 1.10. Image reconstruction by a Fourier hologram.

If the incident wave has unit amplitude, and the amplitude transmittance of the processed hologram is a linear function of the intensity, the complex amplitude of the transmitted wave is

$$U(\xi,\eta) = \mathbf{t}_0 + \beta T I(\xi,\eta). \tag{1.13}$$

The complex amplitude in the back focal plane of the lens is then the Fourier transform of $U(\xi,\eta)$. We have

$$\begin{aligned} u(x, y) &= \mathcal{F}\{U(\xi,\eta)\} \\ &= (\mathbf{t}_0 + \beta T)\delta(x, y) + \beta T o(x, y) \bigstar o(x, y) \\ &\quad + \beta T o(x - b, y) + \beta T o^*(-x + b, -y). \end{aligned} \tag{1.14}$$

As shown in fig. 1.10, the wave corresponding to the first term on the right-hand side of (1.14) comes to a focus on the axis, while that corresponding to the second term forms a halo around it. The third term produces an image of the original object, shifted downwards by a distance b, while the fourth term gives rise to a conjugate image, rotated by 180° and shifted upwards by the same distance b.

Fourier holograms have the useful property that the reconstructed image does not move when the hologram is translated in its own plane. This is because a shift of a function in the spatial domain only results in its Fourier transform being multiplied by a phase factor which has no effect on the intensity distribution.

1.5 Lensless Fourier holograms

A hologram with the same properties as a Fourier hologram can be produced, without a lens, with the arrangement shown in fig. 1.11 in which the object is illuminated with a plane wave, and the reference wave comes from a point source in the plane of the object [Stroke, 1965].

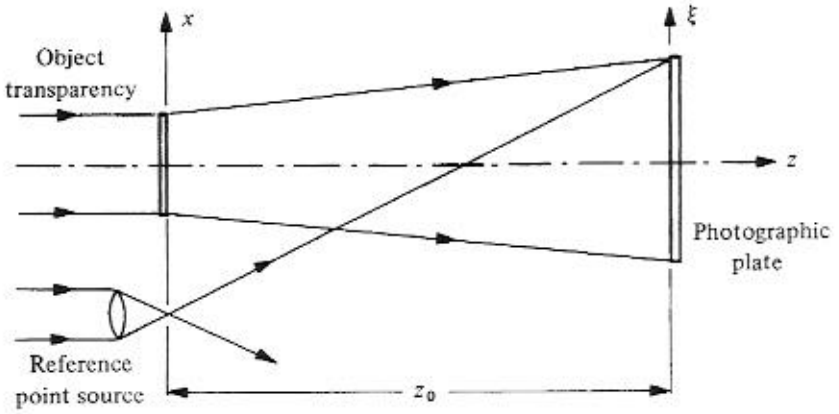


Fig. 1.11. Optical system used to record a Fourier hologram without a lens.

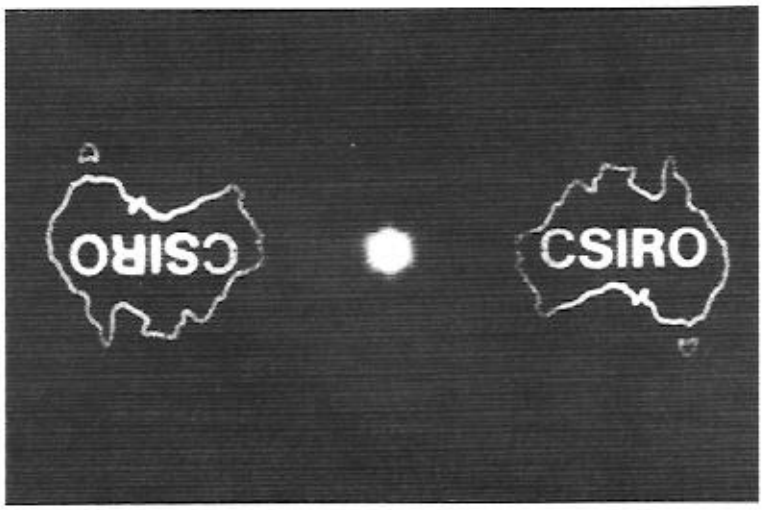


Fig. 1.12. Images reconstructed by a lensless Fourier hologram.

In this recording arrangement, the effect of the spherical phase factor associated with the near-field (or Fresnel) diffraction pattern of the object transparency (see Appendix C) is eliminated by using a spherical reference wave with the same average curvature.

Figure 1.12 shows the images reconstructed by a lensless Fourier hologram.

1.6 Image holograms

For some applications, there are advantages in recording a hologram of a real image of the object formed by a lens. As shown in fig. 1.13, the hologram plate

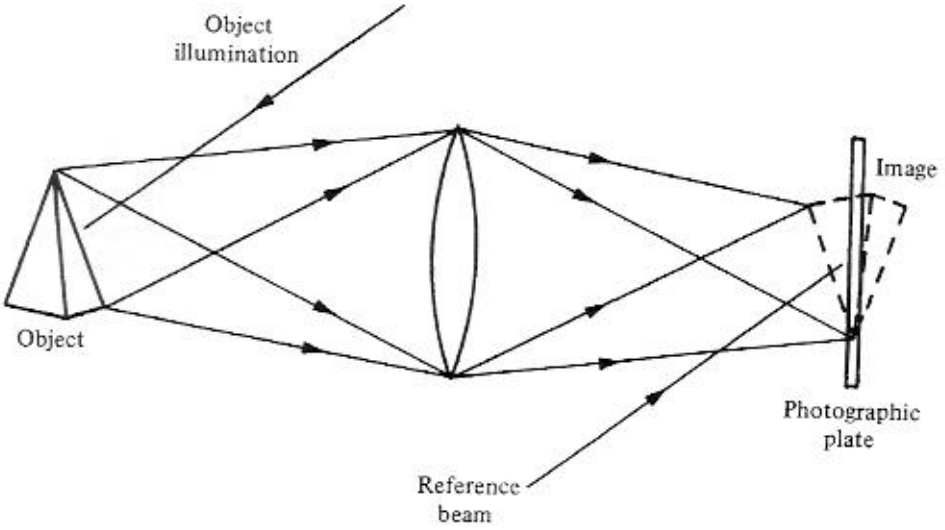


Fig. 1.13. Recording an image hologram.

is set in the central plane of the image, and a hologram is recorded in the normal fashion with an off-axis reference beam.

When the hologram is illuminated with the original reference beam, part of the reconstructed image lies in front of the hologram, and part of the image lies behind it. Since the image is very close to the hologram plane, the spatial and temporal coherence requirements for the illumination used to reconstruct the image are much less critical, and it is even possible, with an object of limited depth, to use an extended white-light source [Rosen, 1966].

1.7 Reflection holograms

It is also possible to record a hologram with the object beam and the reference beam incident on the photographic emulsion from opposite sides. The interference fringes then form a series of planes within the emulsion layer, at a small angle to its surface and about half a wavelength apart. Such holograms, when illuminated with a point source of white light, reflect a sufficiently narrow band of wavelengths to reconstruct an image of acceptable quality.

A simpler way of recording such a hologram, with an object of limited depth, is to use the portion of the reference beam transmitted by the photographic plate to illuminate the object [Denisyuk, 1962].

References

Denisyuk, Yu. N. (1962). Photographic reconstruction of the optical properties of an object in its own scattered radiation field. *Soviet Physics – Doklady*, 7, 543–5.

Gabor, D. (1948). A new microscopic principle. *Nature*, **161**, 777–8.
Leith, E.N. & Upatnieks, J. (1962). Reconstructed wavefronts and communication theory. *Journal of the Optical Society of America*, **52**, 1123–30.
Leith, E. N. & Upatnieks, J. (1963). Wavefront reconstruction with continuous-tone objects. *Journal of the Optical Society of America*, **53**, 1377–81.
Leith, E. N. & Upatnieks, J. (1964). Wavefront reconstruction with diffused illumination and three-dimensional objects. *Journal of the Optical Society of America*, **54**, 1295–301.
Rosen, L. (1966). Focused-image holography with extended sources. *Applied Physics Letters*, **9**, 337–9.
Stroke, G. W. (1965). Lensless Fourier-transform method for optical holography. *Applied Physics Letters*, **6**, 201–3.
Vander Lugt, A. (1964). Signal detection by complex spatial filtering. *IEEE Transactions on Information Theory*, **IT-10**, 139–45.

Problems

1.1. A transmission hologram is recorded using a He–Ne laser ($\lambda = 633$ nm) with the object and reference beams making angles of $+30°$ and $-30°$, respectively, with the normal to the photographic plate. What is the average spatial frequency of the hologram fringes?

The average spatial frequency of the hologram fringes is

$$\xi = \frac{2 \sin 30°}{633 \times 10^{-9}} \text{ m}^{-1}$$
$$= 1579 \text{ lines/mm}. \tag{1.15}$$

1.2. What would be the spatial frequency of the fringes in a reflection hologram recorded with a He–Ne laser ($\lambda = 633$ nm) if the object beam is normal to the photographic plate, and the reference beam makes an angle of 45° with the normal?

The refractive index of the unexposed photographic emulsion is about 1.6.

As a result, after refraction, the reference beam makes an angle of 26.3° with the normal. The angle between the two beams in the photographic emulsion is 153.7°, and the two beams make angles of $\pm 76.85°$ with the fringe planes, which are inclined at 13.15° to the surface of the emulsion.

In addition, the wavelength of the laser in the unexposed photographic emulsion is $633/1.6 = 395.6$ nm. The spatial frequency of the fringes is, therefore,

$$\xi = \frac{2 \sin 76.85°}{395.6 \times 10^{-9}} \text{ m}^{-1}$$
$$= 4923 \text{ lines/mm}. \tag{1.16}$$

1.3. A lensless Fourier hologram is recorded with a He–Ne laser ($\lambda = 633$ nm) using an arrangement similar to that shown in fig. 1.11. The object is a

transparency with a width of 2 cm placed with its inner edge at a distance of 0.5 cm from the point source providing the reference wave. The hologram is recorded on a photographic plate placed at a distance of 25 cm from the reference source. What is the range of spatial frequencies of the hologram fringes?

At any point in the hologram, the angle between the object and reference beams ranges from a maximum of approximately 0.1 radian (5.73°) to a minimum of approximately 0.02 radian (1.15°). The spatial frequency of the hologram fringes ranges, therefore, from

$$\xi_{\max} \approx \frac{\sin 5.73^\circ}{633 \times 10^{-9}}\,\mathrm{m}^{-1} \approx 158\ \mathrm{lines/mm} \tag{1.17}$$

to

$$\xi_{\min} \approx \frac{\sin 1.15^\circ}{633 \times 10^{-9}}\,\mathrm{m}^{-1} \approx 31.6\ \mathrm{lines/mm}. \tag{1.18}$$

2

The reconstructed image

2.1 Images of a point

To study the characteristics of the reconstructed image and their dependence on the optical system, we consider, as shown in fig. 2.1(*a*), the hologram of a point object $O(x_O, y_O, z_O)$, recorded with a reference wave from a point source $R(x_R, y_R, z_R)$, using light of wavelength λ_1.

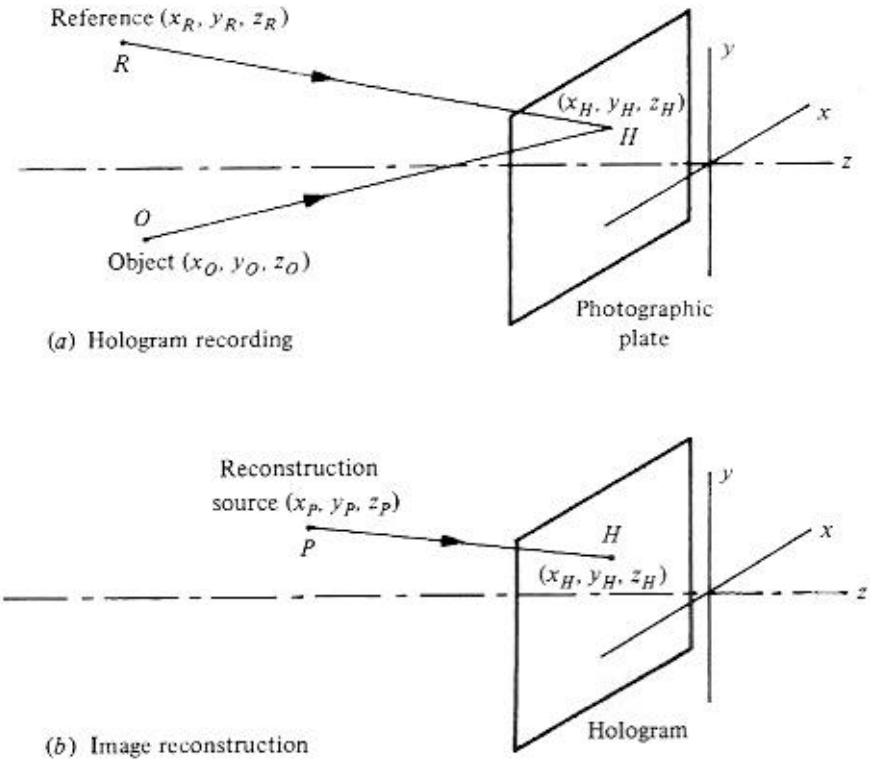


Fig. 2.1. Formation of the image of a point object by a hologram.

If, then, this hologram is illuminated with monochromatic light with a wavelength λ_2 from a point source $P(x_P, y_P, z_P)$, as shown in fig. 2.1(*b*), it can be shown that the coordinates of the primary image of O are [Meier, 1965]

$$x_I = \frac{x_P z_O z_R + \mu x_O z_P z_R - \mu x_R z_P z_O}{z_O z_R + \mu z_P z_R - \mu z_P z_O}, \tag{2.1}$$

$$y_I = \frac{y_P z_O z_R + \mu y_O z_P z_R - \mu y_R z_P z_O}{z_O z_R + \mu z_P z_R - \mu z_P z_O}, \tag{2.2}$$

$$z_I = \frac{z_P z_O z_R}{z_O z_R + \mu z_P z_R - \mu z_P z_O}, \tag{2.3}$$

where $\mu = (\lambda_2/\lambda_1)$, while those of the conjugate image of O are

$$x_C = \frac{x_P z_O z_R - \mu x_O z_P z_R + \mu x_R z_P z_O}{z_O z_R - \mu z_P z_R + \mu z_P z_O}, \tag{2.4}$$

$$y_C = \frac{y_P z_O z_R - \mu y_O z_P z_R + \mu y_R z_P z_O}{z_O z_R - \mu z_P z_R + \mu z_P z_O}, \tag{2.5}$$

$$z_C = \frac{z_P z_O z_R}{z_O z_R - \mu z_P z_R + \mu z_P z_O}. \tag{2.6}$$

The lateral magnification of the primary image can be defined as

$$M_{\text{lat},I} = \mathrm{d}x_I/\mathrm{d}x_O = \mathrm{d}y_I/\mathrm{d}y_O = \left[1 + z_O\left(\frac{1}{\mu z_P} - \frac{1}{z_R}\right)\right]^{-1}, \tag{2.7}$$

while that of the conjugate image is

$$M_{\text{lat},C} = \left[1 - z_O\left(\frac{1}{\mu z_P} - \frac{1}{z_R}\right)\right]^{-1}. \tag{2.8}$$

Similarly, the longitudinal magnification of the primary image is

$$M_{\text{long},I} = \mathrm{d}z_I/\mathrm{d}z_O = \frac{1}{\mu}(M_{\text{lat},I})^2, \tag{2.9}$$

while that of the conjugate image is

$$M_{\text{long},C} = -\frac{1}{\mu}(M_{\text{lat},C})^2, \tag{2.10}$$

implying that it has the opposite sign.

If the hologram is illuminated with the reference wave originally used to record it, the primary image has the same size as the original object and coincides with it. However, any change in the position, or the wavelength, of the point source used for reconstruction results in a change in the position and magnification of the reconstructed images and can introduce aberrations.

2.2 Orthoscopic and pseudoscopic images

To understand the implications of the opposite signs of $M_{\text{long},I}$ and $M_{\text{long},C}$, we consider an off-axis hologram recorded with a collimated reference beam incident normal to the photographic plate, as shown in fig. 2.2(*a*).

When the hologram is illuminated once again with the same collimated reference beam, as shown in fig. 2.2(*b*), it reconstructs two images, one virtual and the other real. While the virtual image is located behind the hologram, in the same position as the object, the real image is located at the same distance from the hologram, but in front of it. Since corresponding points on the virtual and real images are located at equal distance from the plane of the hologram, the real image has the curious property that its depth appears inverted. Such an image is called a pseudoscopic image.

A hologram that produces an orthoscopic real image of an object can be produced in two steps [Rotz & Friesem, 1966]. In the first step, as shown in fig. 2.3, a hologram (H1) is recorded of the object with a collimated reference beam. H1 is then illuminated once again with the same reference beam, and a second hologram (H2) is recorded of the real image formed by H1 using another collimated reference beam.

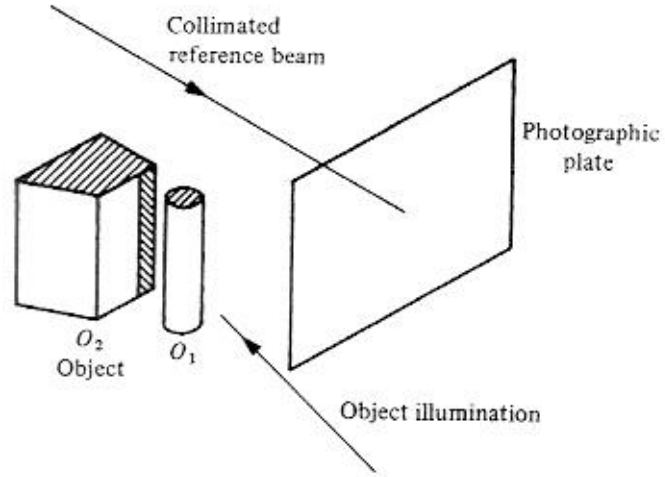
When H2 is illuminated with a collimated beam, it reconstructs a pseudoscopic virtual image, located in the same position as the real image formed by H1, and an orthoscopic real image.

A simpler method is to record a hologram of an orthoscopic real image of the object formed by a large concave mirror or lens. When this hologram is illuminated with the original reference beam, it reconstructs the original object wave, producing an orthoscopic real image.

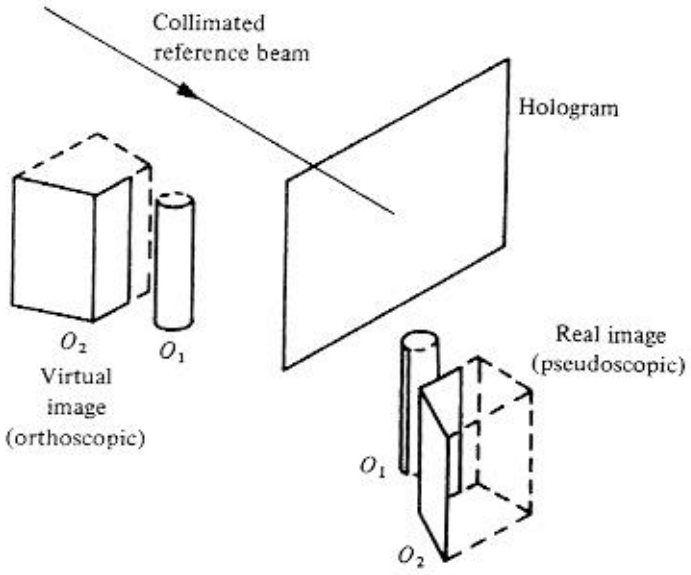
2.3 Image aberrations

If the hologram is replaced in the position in which it was recorded and illuminated with the original reference beam, $z_p = z_R$, and $\mu = 1$. The primary image then coincides with the object. In any other case, the image may exhibit aberrations.

These aberrations can be defined as the phase difference between the reference spheres centered on the image points defined in Section 2.1 and the



(*a*) Hologram recording



(*b*) Image reconstruction

Fig. 2.2. Formation of orthoscopic and pseudoscopic images by a hologram.

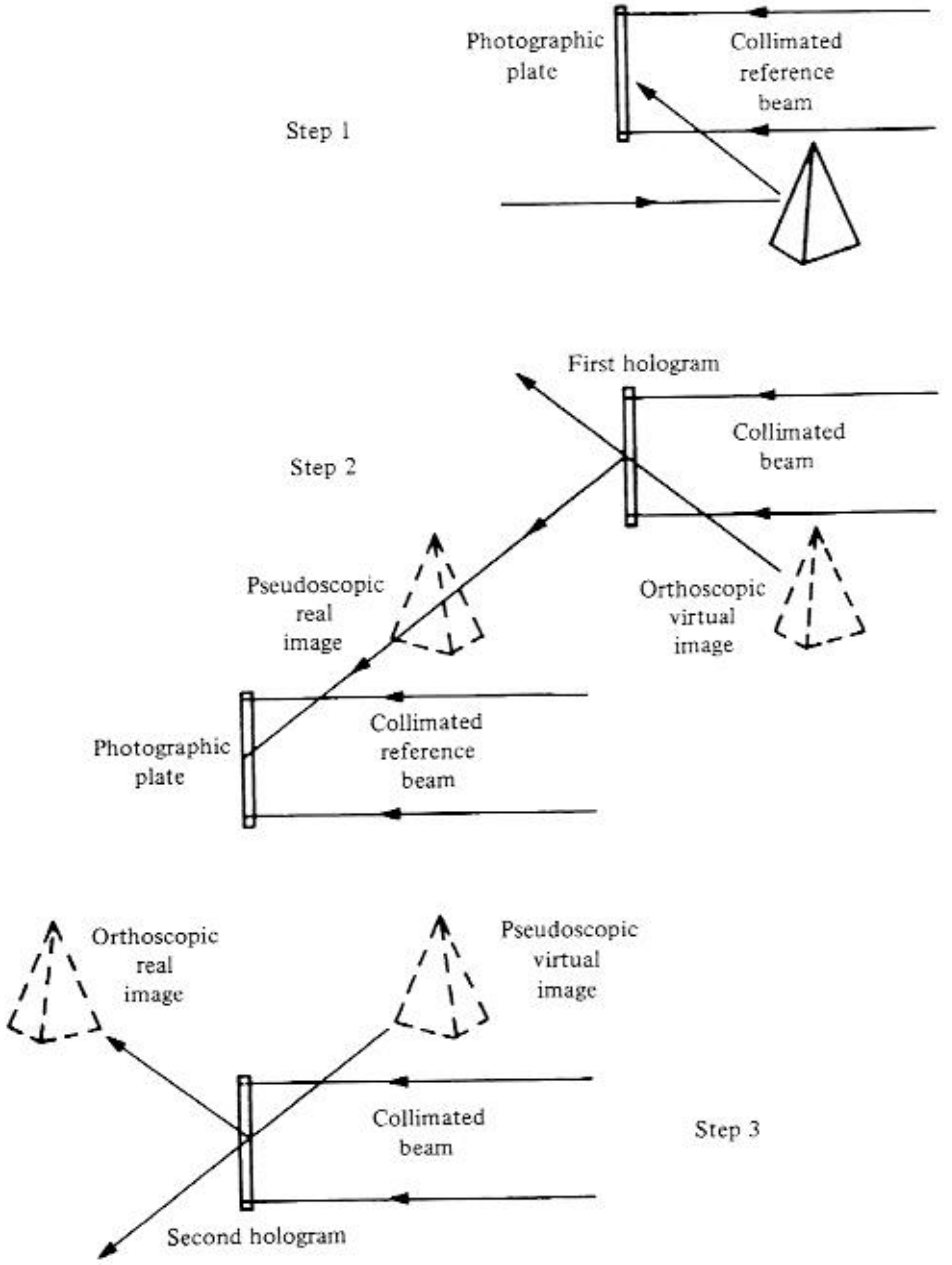


Fig. 2.3. Production of an orthoscopic real image by recording two holograms in succession [Rotz & Friesem, 1966].

actual reconstructed wavefronts. They can be evaluated if we retain the third-degree terms in the expansion for the phase of a spherical wavefront [Meier, 1965].

2.3.1 Classification of aberrations

Hologram aberrations can be classified in the same manner as lens aberrations [Hopkins, 1950]. If we use polar coordinates (ρ,θ) in the hologram plane, the third-order aberration can be written as

$$\begin{aligned}\Delta\phi_I=(2\pi/\lambda_2)\,[&-(1/8)\,\rho^4 S+(1/2)\rho^3(C_x\cos\theta+C_y\sin\theta)\\&-(1/2)\rho^2(A_x\cos^2\theta+A_y\sin^2\theta+2A_xA_y\cos\theta\sin\theta)\\&-(1/4)\rho^2F+(1/2)\rho(D_x\cos\theta+D_y\sin\theta],\end{aligned}\tag{2.11}$$

where S is the coefficient of spherical aberration, C_x, C_y are the coefficients of coma, A_x, A_y are the coefficients of astigmatism, F is the coefficient of field curvature and D_x, D_y are the coefficients of distortion.

For simplicity, we will assume that the object lies on the x axis ($y_O=0$) and consider only the conjugate (real) image. The aberration coefficients for the primary (virtual) image can be obtained by changing the signs of z_O and z_R in the expressions obtained for the conjugate image.

It can then be shown that the coefficient of spherical aberration is

$$S=(1/z_P^3)-(\mu/z_O^3)+(\mu/z_R^3)-(1/z_C^3).\tag{2.12}$$

When $z_R=z_O$, $z_C=z_P$, and the spherical aberration is zero.

The coefficient of coma is

$$C_x=(x_P/z_P^3)-(\mu x_O/z_O^3)+(\mu x_R/z_R^3)-(x_C/z_C^3).\tag{2.13}$$

Coma can be eliminated only if $z_R=z_O$ and $z_P=\pm z_O$.

The coefficient of astigmatism is

$$A=(x_P^2/z_P^3)-(\mu x_O^2/z_O^3)+(\mu x_R^2/z_R^3)-(x_C^2/z_C^3).\tag{2.14}$$

Astigmatism is eliminated when $z_R=z_O$, $(x_P/z_P)=-(\mu x_R/z_R)$ and $z_P=\mu z_O$. Coma and astigmatism can, therefore, be eliminated simultaneously only when $\mu=1$.

The coefficient for curvature of field is

$$\begin{aligned}F=&[(x_P^2+y_P^2)/z_P^3]-[\mu(x_O^2+y_O^2)/z_O^3]\\&+[\mu(x_R^2+y_R^2)/z_R^3]-[(x_C^2+y_C^2)/z_C^3]\end{aligned}\tag{2.15}$$

which disappears when the astigmatism is reduced to zero.

Finally, the coefficient of distortion is

$$D_x = [(x_P^3 + x_P y_P^2)/z_P^3] - [\mu(x_O^3 + x_O y_O^2)/z_O^3] + [\mu(x_R^3 + x_R y_R^2)/z_R^3] - [(x_C^3 + x_C y_C^2)/z_C^3]. \quad (2.16)$$

Distortion cannot normally be eliminated when $\mu \neq 1$.

In particular, unless $M_{lat} = 1$ and $\mu = 1$, the longitudinal magnification is not the same as the lateral magnification, resulting in a distortion in depth. However, longitudinal distortion due to the recording and reconstruction wavelengths not being the same ($\mu \neq 1$) can be minimized by a proper choice of the recording and reconstruction geometry [Hariharan, 1976].

2.4 Image blur

A source other than a laser (one with a finite size and spectral bandwidth) is often used to illuminate the hologram and view the reconstructed image. However, the use of such a source affects the sharpness of the reconstructed image.

From (2.1), it can be shown that if the source used to illuminate the hologram occupies very nearly the same position as the reference source used to record it and has very nearly the same wavelength, the blur of the reconstructed image along the x axis for a source size Δx_P is

$$\Delta x_I = (z_O/z_P)\,\Delta x_P. \quad (2.17)$$

Similarly, if the source used to illuminate the hologram has a mean wavelength λ_2 very nearly equal to λ_1, the wavelength used to record the hologram, and a spectral bandwidth $\Delta\lambda_2$, the transverse image blur due to the finite spectral bandwidth of the source can be shown to be

$$|\Delta x_I| = (x_P/z_P) z_O (\Delta\lambda_2/\lambda_2). \quad (2.18)$$

The image blur increases with the depth of the image and the interbeam angle.

It follows from (2.17) and (2.18) that if the central plane of the image lies in the hologram plane, as in an image hologram (see Section 1.6), the restrictions on the size and the spectral bandwidth of the source used to illuminate the hologram can be relaxed. In fact, if the interbeam angle and the depth of the object are small, it is possible to use an extended white-light source to view the image.

2.5 Image speckle

When a diffusely reflecting object is illuminated with light from a laser, each element on its surface produces a diffracted wave. Since the differences in their

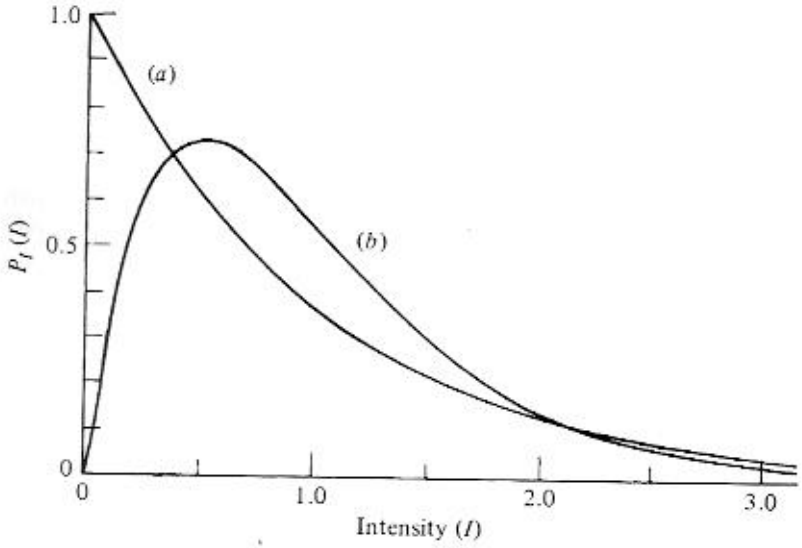


Fig. 2.4. Intensity distribution in (*a*) a single speckle pattern and (*b*) the incoherent sum of two different speckle patterns.

optical paths have a random distribution, these diffracted waves interfere with each other to produce a speckle pattern in the far field (see Appendix D). As a result, the image reconstructed by a hologram exhibits a speckled appearance. With polarized light, the intensity in the speckle pattern has, as shown in fig. 2.4(*a*), the negative exponential distribution

$$p(I) = \frac{1}{2\sigma^2} \exp\left(\frac{-I}{2\sigma^2}\right), \tag{2.19}$$

where $\langle I \rangle = 2\sigma^2$ is the mean intensity [Goodman, 1975].

The contrast of the speckle pattern is unity, and its appearance is almost independent of the nature of the surface, but the size of the speckles increases with the viewing distance and the *f*-number of the imaging system. With a circular pupil of radius ρ, the average dimensions of the speckles in the image are

$$\Delta x = \Delta y = 0.61 \lambda f / \rho. \tag{2.20}$$

Speckle is a serious problem in holographic imaging. While several techniques have been described to reduce speckle in the reconstructed image [McKechnie, 1975], the most common method is to record a number of holograms with the object illuminated from slightly different directions. Each of these holograms reconstructs the same image, but a different speckle pattern.

The intensity distribution in the speckle pattern obtained by summing the intensities in two such images with average intensities of $\langle I/2 \rangle$ is, then

$$p(I) = \frac{4I}{\langle I \rangle^2} \exp\left(\frac{-2I}{\langle I \rangle}\right). \tag{2.21}$$

As can be seen from fig. 2.4(*b*), most of the dark areas are eliminated, and the contrast of the pattern is only $1/\sqrt{2}$. If we sum the intensities in the images produced by N holograms, the intensity fluctuations due to speckle are reduced by a factor equal to $\sqrt{N}$.

2.6 Signal-to-noise ratio

Random spatial variations in the intensity of the reconstructed image, commonly caused by scattered light, are referred to as noise. In the case of holographic imaging, when calculating the signal-to-noise ratio, the amplitudes of the signal and the noise must be added, since they are both encoded on a common carrier [Goodman, 1967].

We consider the reconstructed image of a uniform bright patch on a dark background, and assume that the intensity due to the nominally uniform signal is I_S, while that of the randomly varying background is I_N. The noise N in the bright area is given by the variance of the resulting fluctuations of the intensity. It can then be shown that when, as is usually the case, $I_S \gg \langle I_N \rangle$, (where $\langle I_N \rangle$ is the average intensity of the dark background), the signal-to-noise ratio is

$$\frac{I_S}{N} = \left(\frac{I_S}{2\langle I_N \rangle}\right)^{1/2}. \tag{2.22}$$

It follows that even a small amount of scattered light can result in relatively large fluctuations in intensity in the bright areas of the image.

2.7 Image luminance

The luminance of the reconstructed image is directly proportional to the diffraction efficiency of the hologram (defined as the ratio of the energy diffracted into the image to that incident on the hologram) and inversely proportional to the solid angle over which the image can be viewed [Hariharan, 1978].

A hologram recorded of a real image projected either by an optical system or another hologram (see fig. 2.3) also reconstructs an image of the optical system, including any aperture (or pupil) that limits the angular spread of the

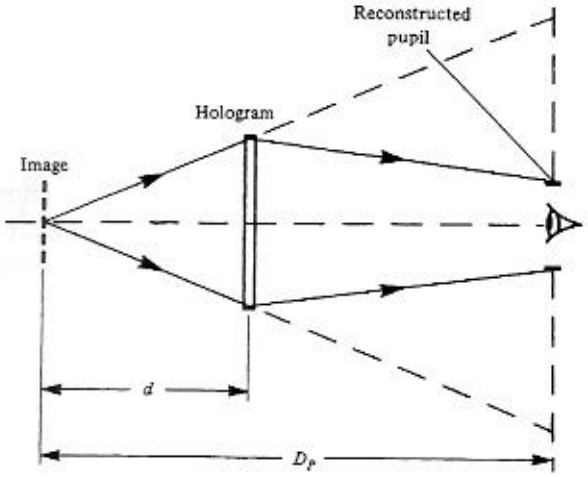


Fig. 2.5. Image reconstruction with an external reconstructed pupil [Hariharan, 1978].

object beam. As shown in fig. 2.5, the flux from the reconstructed image is confined within this external reconstructed pupil.

As a result, the solid angle over which the image can be viewed is reduced, but the luminance of the image is increased. A substantial improvement in image luminance can be obtained if the aperture of the imaging system is masked so that its shape and size match the range of angles over which the hologram is actually to be viewed.

References

Goodman, J. W. (1967). Film grain noise in wavefront reconstruction imaging. *Journal of the Optical Society of America*, **57**, 493–502.

Goodman, J. W. (1975) Statistical properties of laser speckle patterns. In *Laser Speckle & Related Phenomena*, Topics in Applied Physics, vol. 9, ed. J. C. Dainty, pp. 9–75. Berlin: Springer-Verlag.

Hariharan, P. (1976). Longitudinal distortion in images reconstructed by reflection holograms. *Optics Communications*, **17**, 52–4.

Hariharan, P. (1978). Hologram recording geometry: its influence on image luminance. *Optica Acta*, **25**, 527–30.

Hopkins, H. H. (1950). *Wave Theory of Aberrations*. Oxford: The Clarendon Press.

McKechnie, T. S. (1975). Speckle reduction. In *Laser Speckle & Related Phenomena*, Topics in Applied Physics, vol. 9, ed. J. C. Dainty, pp. 123–70. Berlin: Springer-Verlag.

Meier, R. W. (1965). Magnification and third-order aberrations in holography. *Journal of the Optical Society of America*, **55**, 987–92.

Rotz, F. B. & Friesem, A. A. (1966). Holograms with non-pseudoscopic real images. *Applied Physics Letters*, **8**, 146–8.

Problems

2.1. A hologram is recorded using a pulsed ruby laser ($\lambda = 694$ nm), and illuminated with a He–Ne laser ($\lambda = 633$ nm) to view the image. The reference beam in the recording system appears to diverge from a point at a distance of 1 m from the hologram. How far from the hologram should the beam from the He–Ne laser be brought to a focus to ensure that the image is reconstructed with unit lateral magnification?

From (2.7), it follows that the condition for the image to be reconstructed with unit lateral magnification is

$$\mu z_P = z_R, \tag{2.23}$$

where z_P and z_R are the distances from the hologram of the sources used for reconstruction and recording, respectively, and μ is the ratio of their wavelengths. Accordingly, the beam from the He–Ne laser should be brought to a focus at a distance from the hologram

$$\begin{aligned} z_P &= 1 \times \frac{694}{633} \text{ m} \\ &= 1.096 \text{ m}. \end{aligned} \tag{2.24}$$

2.2. A hologram recorded with an Ar^+ laser ($\lambda_1 = 514$ nm) and an interbeam angle of 30° is to be illuminated with green light from a high-pressure mercury vapor lamp which has a mean wavelength λ_2 of 546 nm and a spectral bandwidth $\Delta\lambda_2$ of about 5 nm. The lamp can be regarded as a source with an effective diameter of 5 mm and is placed at a distance of 1 m from the hologram. What is the maximum depth of the image that can be reconstructed with an acceptable value of the image blur?

The acceptable value of the image blur for a display is determined by the resolution of the eye, which is about 0.5 mrad, or 0.5 mm at a viewing distance of 1 m.

The total image blur is the sum of the contributions from the finite size of the source and its spectral bandwidth and, from (2.17) and (2.18), can be written as

$$\Delta x_{I,\text{Total}} = z_O \left(\frac{\Delta x_P}{z_P} \right) + z_O \left(\frac{x_P}{z_P} \right) \left(\frac{\Delta\lambda_2}{\lambda_2} \right). \tag{2.25}$$

In the present case, $\Delta x_p = 5$ mm, $z_p = 1$ m and $(x_p/z_p) = \tan 30° = 1/\sqrt{3}$, so that we have

$$0.5 \times 10^{-3} = z_O \left[\frac{5}{1000} + \frac{1}{\sqrt{3}} \frac{5}{546} \right],$$

$$= z_O [5 \times 10^{-3} + 5.293 \times 10^{-3}], \tag{2.26}$$

and

$$z_O = 48.6 \text{ mm}. \tag{2.27}$$

It follows that the distance of any point in the image from the hologram should not exceed this value.

Note that, in this case, the contributions of the size of the source and its spectral bandwidth to the image blur are almost equal.

2.3. The image reconstructed by a hologram illuminated with a He Ne laser ($\lambda = 633$ nm) is to be photographed with a camera having a lens with a focal length of 100 mm. We need to stop the lens down to get the maximum depth of focus in the image. What is the highest f-number that can be used for the speckle size in the image to be less than 0.01 mm?

From (2.20), the minimum radius of the lens aperture is

$$\rho_{\min} = \frac{0.61 \times 633 \times 10^{-9} \times 100 \times 10^{-3}}{0.01 \times 10^{-3}} \text{ m}$$

$$= 3.86 \text{ mm}, \tag{2.28}$$

which would correspond to an f-number of $100/(2 \times 3.86) = 12.95$.

3

Thin and thick holograms

So far, we have treated a hologram recorded on a photographic film as equivalent, to a first approximation, to a grating of negligible thickness with a spatially varying transmittance. However, if the thickness of the recording medium is larger than the average spacing of the fringes, volume effects cannot be neglected. It is even possible, as mentioned in Section 1.7, to produce holograms in which the interference pattern that is recorded consists of planes running almost parallel to the surface of the recording material; such holograms reconstruct an image in reflected light.

In addition, with modified processing techniques, or with other recording materials, it is possible to reproduce the variations in the intensity in the interference pattern produced by the object and reference beams as variations in the refractive index, or the thickness, of the hologram. Accordingly, holograms recorded in a medium whose thickness is much less than the spacing of the interference fringes (thin holograms) can be classified as amplitude holograms and phase holograms.

Similarly, holograms recorded in thick media (volume holograms) can be subdivided into transmission amplitude holograms, transmission phase holograms, reflection amplitude holograms and reflection phase holograms.

In the next few sections we review the characteristics of these six types of holograms. For simplicity, we consider only gratings produced by the interference of two plane wavefronts.

3.1 Thin gratings

3.1.1 Thin amplitude gratings

The amplitude transmittance of a thin amplitude grating can be written as

$$t(x) = t_0 + \Delta t \cos Kx, \tag{3.1}$$

where t_0 is the average amplitude transmittance, Δt is the amplitude of the spatial variations of $t(x)$, and $K = 2\pi/\Lambda$, where Λ is the average spacing of the fringes. The maximum amplitude in each of the two diffracted orders is a fourth of that in the incident wave, so that the maximum diffraction efficiency is

$$\eta_{\max} = 0.0625. \tag{3.2}$$

3.1.2 *Thin phase gratings*

If the phase shift produced by the recording medium is proportional to the intensity in the interference pattern, the resulting recording can be regarded as a thin phase grating whose complex amplitude transmittance can be written as

$$\mathbf{t}(x) = \exp(-\mathrm{i}\phi_0) \exp[-\mathrm{i}\Delta\phi \cos(Kx)], \tag{3.3}$$

where ϕ_0 is a constant phase factor, and $\Delta\phi$ is the amplitude of the phase variations. If we neglect this constant phase factor, the right-hand side of (3.3) can be expanded to obtain the relation

$$\mathbf{t}(x) = \sum_{n=-\infty}^{\infty} \mathrm{i}^n J_n(\Delta\phi) \exp(\mathrm{i}nKx), \tag{3.4}$$

where J_n is the Bessel function of the first kind, of order n. The incident beam is diffracted into a number of orders, with the diffracted amplitude in the nth order proportional to the value of the Bessel function $J_n(\Delta\phi)$, but only the first diffracted order contributes to the primary image. The diffraction efficiency of the phase grating can therefore be written as

$$\eta = J_1^2(\Delta\phi). \tag{3.5}$$

As shown in fig. 3.1, the amplitude diffracted into the first order increases initially with the phase modulation and then decreases; the maximum value of the diffraction efficiency is

$$\eta_{\max} = 0.339. \tag{3.6}$$

3.2 Volume gratings

With a thick recording medium, the hologram is made up of layers corresponding to a periodic variation of transmittance or refractive index. If the two interfering wavefronts are incident on the recording medium from the same side, these layers are approximately perpendicular to its surface, and the hologram produces an image by transmission. However, it is also possible to have the two

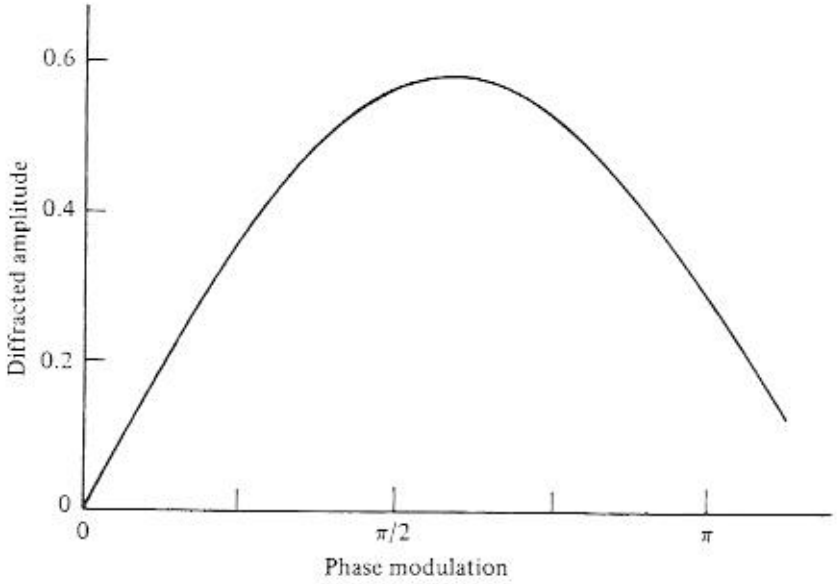


Fig. 3.1. Amplitude diffracted, as a function of the phase modulation, for a thin phase grating [Kogelnik, 1967].

interfering wavefronts incident on the recording medium from opposite sides (see Section 1.7), in which case the interference surfaces run approximately parallel to the surface of the recording medium. In this case, the reconstructed image is produced by the light reflected from the hologram.

We will use a coordinate system in which, as shown in fig. 3.2, the z axis is perpendicular to the surfaces of the recording medium, and the x axis is in the plane of incidence. For simplicity, we will assume that the interference surfaces are either perpendicular or parallel to the surfaces of the recording medium. The grating vector $\mathbf{K}$, which is perpendicular to the interference surfaces, is of length $K = 2\pi/\Lambda$, where Λ is the grating period, and makes an angle ψ ($\psi = 90°$ or 0° in this case) with the z axis.

In both cases, the diffracted amplitude is a maximum only when the angle of incidence is equal to the Bragg angle θ_B, satisfying the Bragg condition

$$\cos(\psi - \theta_B) = \frac{K}{2n_0 k_0}, \tag{3.7}$$

where n_0 is the average refractive index of the recording medium, and $k_0 = 2\pi/\lambda$. As a result, only two waves need be considered; they are the incoming reference wave R and the outgoing signal wave S. Since the other diffracted orders violate the Bragg condition strongly, they are severely attenuated and can be neglected. With volume reflection holograms, the angular and wavelength selectivity can

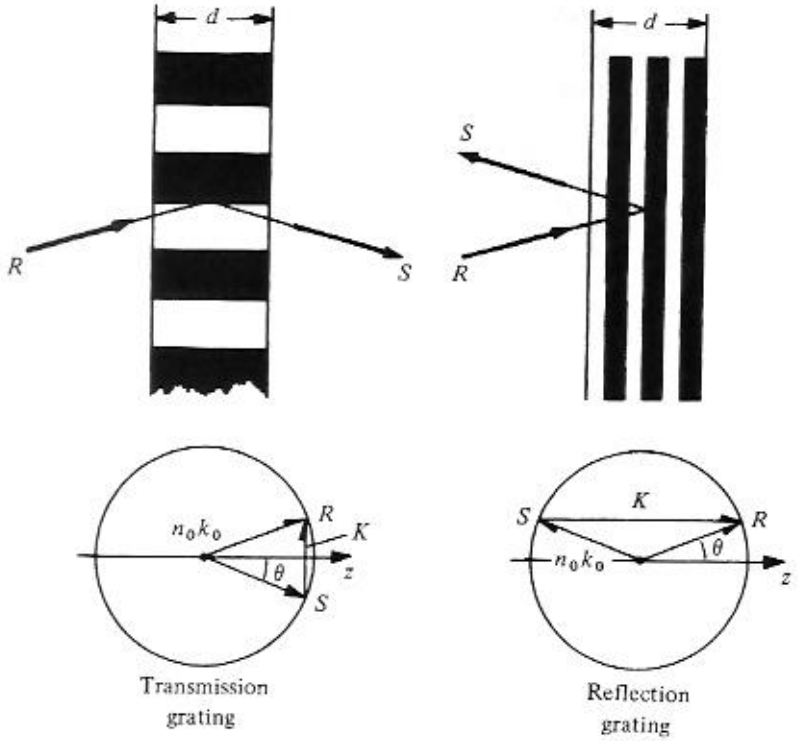


Fig. 3.2. Volume transmission and reflection gratings and their associated vector diagrams for Bragg incidence [Kogelnik, 1967].

be high enough to produce a sharp, monochromatic image when the hologram is illuminated with white light.

When analyzing the diffraction of light by volume gratings, it is necessary to take into account the fact that the amplitude of the diffracted wave increases progressively, while that of the incident wave decreases, as they propagate through the grating. This problem was solved by the development of a coupled wave theory [Kogelnik, 1969; Solymar & Cooke, 1981]. Some of the most important results for volume gratings are summarized below.

3.2.1 Volume transmission gratings

We consider, in the first instance, a lossless, volume transmission phase grating of thickness d, with the grating planes running normal to its surface. If we assume that the refractive index varies sinusoidally with an amplitude Δn about a mean value n_0, the diffraction efficiency of the grating at the Bragg angle θ_B is

$$\eta_B = \sin^2 \Phi, \tag{3.8}$$

where $\Phi = \pi \Delta n\, d/\lambda \cos \theta_B$ is known as the modulation parameter. The diffraction efficiency increases initially as the modulation parameter Φ is increased, until, when $\Phi = \pi/2$, $\eta_B = 1$. Beyond this point, the diffraction efficiency decreases.

For a deviation $\Delta\theta$ in the angle of incidence, or a deviation $\Delta\lambda$ in the wavelength of the incident beam, from the values required to satisfy the Bragg condition, the diffraction efficiency drops to

$$\eta = \frac{\sin^2(\Phi^2 + \chi^2)^{1/2}}{(1 + \chi^2/\Phi^2)}, \tag{3.9}$$

where

$$\chi = \Delta\theta \frac{Kd}{2}, \tag{3.10}$$

or, alternatively,

$$\chi = \frac{\Delta\lambda\, K^2 d}{8\pi n_0 \cos\theta_B}. \tag{3.11}$$

Figure 3.3 shows the normalized diffraction efficiency (η/η_B) of a volume transmission phase grating, as a function of the parameter χ, for three values of the modulation parameter Φ.

The other case we shall consider is that of a volume transmission grating in which the refractive index does not vary, but the absorption constant varies with an amplitude $\Delta\alpha$ about its mean value α. In this case, the diffraction efficiency, for incidence at the Bragg angle, is given by the expression

$$\eta = \exp\left(\frac{-2\alpha d}{\cos\theta_B}\right) \sinh^2\left(\frac{\Delta\alpha\, d}{2\cos\theta_B}\right). \tag{3.12}$$

The maximum diffraction efficiency is obtained when

$$\Delta\alpha = \alpha = \frac{\ln 3}{d\cos\theta_B} \tag{3.13}$$

and has a value $\eta_{max} = 0.037$.

3.2.2 *Volume reflection gratings*

The diffraction efficiency of a volume reflection phase grating at the Bragg angle is given by the relation

$$\eta_B = \tanh^2 \Phi_r, \tag{3.14}$$

where $\Phi_r = \pi \Delta n\, d/\lambda \cos\theta_B$ and Δn is the amplitude of the variation in the refractive index. As the value of Φ_r increases, the diffraction efficiency increases steadily, as shown in fig. 3.4, to a limiting value of 1.00.

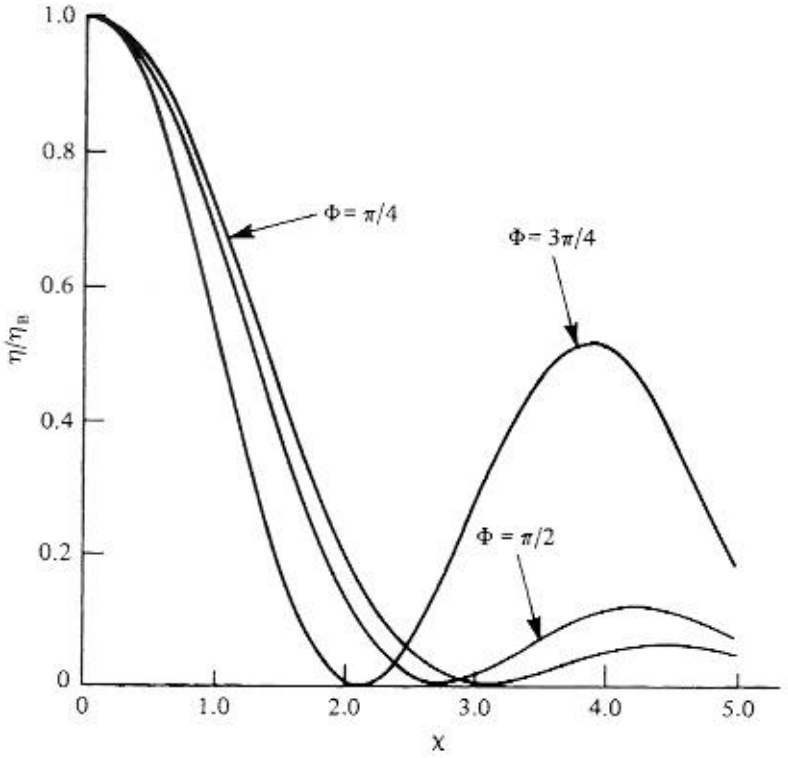


Fig. 3.3. Normalized diffraction efficiency of a volume transmission phase grating as a function of the parameter χ, which is a measure of the deviation from the Bragg condition, for three values of the modulation parameter Φ [Kogelnik, 1969].

For a deviation from the Bragg condition, the normalized diffraction efficiency decreases, as shown in fig. 3.5, as a function of the parameter χ_r specified by the relations

$$\chi_r = \Delta\theta\left(\frac{2\pi n_0 d}{\lambda}\right)\sin\theta_B \tag{3.15}$$

$$= \left(\frac{\Delta\lambda}{\lambda}\right)\left(\frac{2\pi n_0 d}{\lambda}\right)\cos\theta_B. \tag{3.16}$$

The diffraction efficiency drops to zero for a value of $\chi_r \approx 3.5$.

3.3 Imaging properties

With volume holograms, which exhibit a high degree of angular and wavelength selectivity, the amplitude of the diffracted wavefront is affected by any changes of wavelength or geometry between recording and reconstruction.

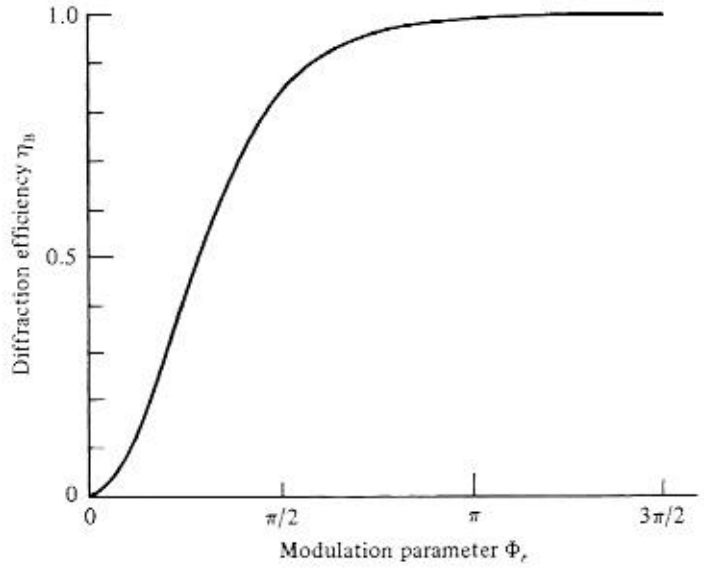


Fig. 3.4. Diffraction efficiency of a volume reflection phase grating as a function of the modulation parameter Φ_r.

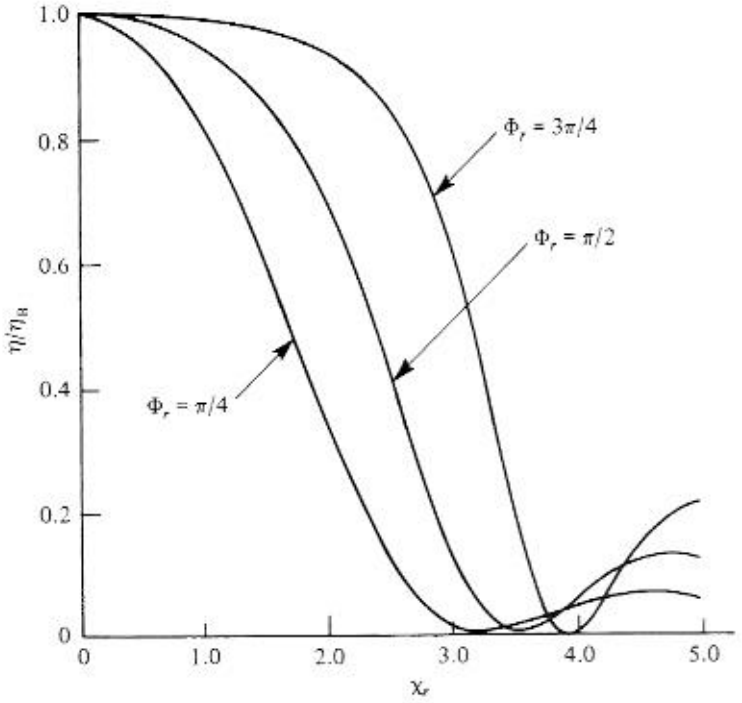


Fig. 3.5. Normalized diffraction efficiency (η/η_B) of a volume reflection phase grating as a function of the parameter χ_r, which is a measure of the deviation from the Bragg condition, for different values of the modulation parameter Φ_r [Kogelnik, 1969].

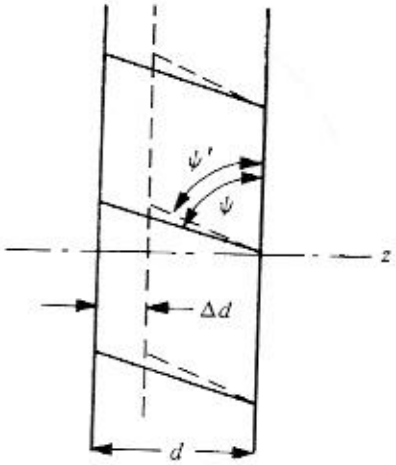


Fig. 3.6. Changes in the orientation and spacing of the fringe planes in a volume transmission hologram due to a change in the thickness of the photographic emulsion.

Such effects are particularly noticeable with holograms recorded on photographic materials, where processing usually results in a change in the thickness of the photographic emulsion.

In the case of a reflection hologram, such a change in thickness results in a change in the spacing of the interference surfaces and, consequently, a change in the color of the image formed when the hologram is illuminated with white light.

With transmission holograms, a change in thickness can result, as shown in fig. 3.6, in a rotation of the fringe planes as well as a change in their spacing [Vilkomerson & Bostwick, 1967; Hariharan, 1999]. For gratings recorded with plane wavefronts, it is possible to compensate for these changes by changing either the angle of illumination or the wavelength. However, complete compensation is not possible with a hologram of a point at a finite distance or a hologram of an extended object.

3.4 Thick and thin gratings

The distinction between thin gratings and volume gratings is commonly made [Klein & Cook, 1967] on the basis of a parameter Q defined by the relation

$$Q = \frac{2\pi\lambda d}{n_0 \Lambda^2}. \tag{3.17}$$

Small values of Q ($Q < 1$) correspond to thin gratings, while large values of Q ($Q > 1$) correspond to volume gratings. However, more detailed studies have

Table 3.1. *Theoretical maximum diffraction efficiencies for transmission phase holograms*

Hologram type	Thin		Volume	
Object beam	Collimated	Diffuse	Collimated	Diffuse
η_{max}	0.33	0.22	1.00	0.64

shown that the transition between the two regimes is not completely defined by (3.17) and that, as the modulation amplitude increases, an intermediate regime appears and widens [Moharam, Gaylord & Magnusson, 1980*a*,*b*].

3.5 Diffusely reflecting objects

The values of diffraction efficiency obtained with a hologram of a diffusely reflecting object are always much lower than those for a grating, because it is not possible to maintain optimum modulation over the entire area, due to the local variations in the amplitude of the object wave. The maximum diffraction efficiencies of transmission phase holograms recorded with a diffuse object beam have been calculated [Upatnieks & Leonard, 1970] on the assumption that the amplitude of the object wave has statistics similar to those of a speckle pattern and are presented in Table 3.1.

3.6 Multiply exposed holograms

With a thick recording medium, it is possible to record two or more holograms in the same medium and read them out separately. In order to do this, their Bragg angles should be sufficiently far apart that the maximum of the angular selectivity curve for one hologram coincides with the first minimum for the other. However, with N amplitude transmission holograms, the diffraction efficiency of each hologram drops to $1/N^2$ of that for a single hologram, since the available dynamic range is divided equally between the N holograms.

On the other hand, with volume phase holograms, if the Bragg angles are far enough apart for coupling between the gratings to be negligible, each hologram diffracts independently of the others [Case, 1975], and there is no loss in diffraction efficiency. However, if the recording medium is nearing saturation, the consequent reduction in modulation can result in a decrease in the diffraction efficiencies of the individual holograms.

References

Case, S. K. (1975). Coupled wave theory for multiply exposed thick holograms. *Journal of the Optical Society of America*, **65**, 724–9.
Hariharan, P. (1999). Real-time holographic interferometry: effects of emulsion shrinkage. *Optics & Lasers in Engineering*, **31**, 339–44.
Klein, W. R. & Cook, B. D. (1967). Unified approach to ultrasonic light diffraction. *IEEE Transactions on Sonics & Ultrasonics*, **SU-14**, 123–34.
Kogelnik, H. (1967). Reconstructing response and efficiency of hologram gratings. In *Proceedings of the Symposium on Modern Optics*, pp. 605–17. Brooklyn Polytechnic Press.
Kogelnik, H. (1969). Coupled wave theory for thick hologram gratings. *Bell System Technical Journal*, **48**, 2909–47.
Moharam, M., Gaylord, T. K. & Magnusson, R. (1980a). Criteria for Bragg regime diffraction by phase gratings. *Optics Communications*, **32**, 14–8.
Moharam, M., Gaylord, T. K. & Magnusson, R. (1980b). Criteria for Raman-Nath regime diffraction by phase gratings. *Optics Communications*, **32**, 19–23.
Solymar, L. & Cooke, D. J. (1981). *Volume Holography & Volume Gratings*. New York: Academic Press.
Upatnieks, J. & Leonard, C. (1970). Efficiency and image contrast of dielectric holograms. *Journal of the Optical Society of America*, **60**, 297–305.
Vilkomerson, D. H. R. & Bostwick, D. (1967). Some effects of emulsion shrinkage on a hologram's image space. *Applied Optics*, **6**, 1270–2.

Problems

3.1. A volume transmission phase grating with a spatial frequency of 1579 lines/mm (see Problem 1.1), recorded in a gelatin layer with a thickness of 15 μm, has a diffraction efficiency of 100 percent at the Bragg angle. What would be the angular selectivity of the grating at a wavelength of 633 nm?

The spacing of the grating fringes is $\Lambda = (1/1579)$ mm $= 0.633$ μm, so that the value of the grating vector $K = 2\pi/\Lambda = 9.926 \times 10^6$ m^{-1}. Since the diffraction efficiency of the grating at the Bragg angle is 100 percent, the modulation parameter $\Phi = \pi/2$. From the curve in fig. 3.3 corresponding to this value of Φ, the diffraction efficiency of the grating drops to zero when the dephasing parameter $\chi = 2.7$. Accordingly, from (3.10), the deviation from the Bragg angle at which the diffraction efficiency drops to zero is

$$\begin{aligned}\Delta\theta &= 2\chi/Kd \\ &= \frac{2 \times 2.7}{9.926 \times 10^6 \times 15 \times 10^{-6}} \text{ radian} \\ &= 3.63 \times 10^{-2} \text{ radian} \\ &= 2.08°. \end{aligned} \tag{3.18}$$

3.2. A volume reflection phase grating (see Problem 1.2) is recorded in a photographic emulsion layer with a thickness of 15 μm. What would be the wavelength selectivity of this grating at a mean wavelength of 633 nm?

Within the emulsion layer, the effective wavelength is 395.6 nm and the Bragg angle is $90° - 76.85° = 13.15°$. From the curves in fig. 3.5, the diffraction efficiency drops to zero when $\chi_r = 3.5$. From (3.16), this would correspond to a change in wavelength

$$\Delta\lambda = \frac{3.5(395.6\times10^{-9})^2}{2\pi\times1.6\times15\times10^{-6}\cos 13.15°}\ \text{m}$$
$$= 3.73\ \text{nm}, \tag{3.19}$$

in the emulsion, or a change in wavelength of 5.97 nm in air.

4

Light sources

In order to maximize the visibility of the interference fringes formed by the object and reference beams, while recording a hologram, it is essential to use coherent illumination (see Appendix A). In addition to being spatially coherent, the coherence length of the light must be much greater than the maximum value of the optical path difference between the object and the reference beams in the recording system. Lasers are therefore employed almost universally as light sources for recording holograms.

4.1 Lasers

The characteristics of some of the lasers used for holography are listed in Table 4.1.

For a simple holographic system, the helium–neon (He–Ne) laser is the usual choice. It is inexpensive and operates on a single spectral line at 633 nm which is well matched to the peak sensitivity of many photographic emulsions. In addition, it does not require water cooling and has a long life.

Table 4.1. *Lasers used for holography*

Laser	Output	Wavelength (nm)	Power
Ar^+	cw	514, 488	1 W
He–Cd	cw	442	25 mW
He–Ne	cw	633	2–50 mW
Kr^+	cw	647	500 mW
Diode	cw	670–650	5 mW
Diode–YAG	cw	532	100 mW
Dye	cw	tunable	200 mW
Ruby	pulsed	694	1–10 J

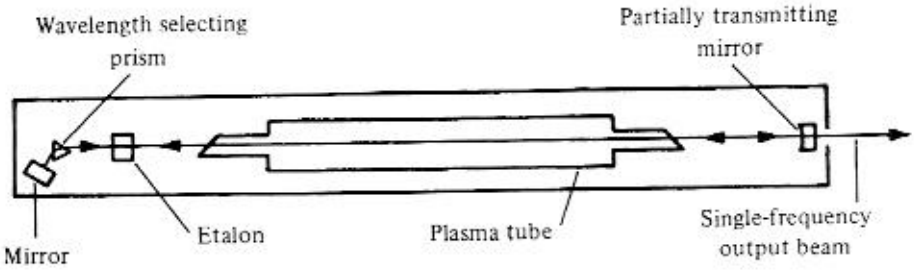


Fig. 4.1. Optical system of an argon-ion laser.

Argon-ion (Ar^+) lasers are more expensive and complex, but give much higher outputs in the green (514 nm) and blue (488 nm) regions of the spectrum. The Ar^+ laser normally has a multiline output but can be made to operate at a single wavelength by replacing the reflecting end mirror by a prism and mirror assembly, as shown in fig. 4.1. The krypton-ion (Kr^+) laser is useful where high output power is required at the red end of the spectrum (647 nm).

The helium–cadmium (He–Cd) laser provides a stable output at a relatively short wavelength (442 nm) and is useful with recording materials such as photoresists.

Diode lasers can be used as a source of red light, and are now available with output wavelengths that are fairly well matched to the sensitivity of commercial photographic materials.

Diode lasers can also be used to pump a Nd:YAG laser with a frequency doubler. Such a system provides a compact cw source of green light ($\lambda = 532$ nm) with output powers up to 100 mW.

Dye lasers are relatively expensive and complex, but have the advantage that they can be operated over a wide range of wavelengths by switching dyes; in addition, for applications such as contouring (see Section 14.3), the output can be tuned over a range of wavelengths (50–80 nm) by incorporating a wavelength selector, such as a diffraction grating or a birefringent filter, in the laser cavity.

While pulsed Nd:YAG lasers (with a frequency doubler) have been used for holography, the most commonly used type of pulsed laser is the ruby laser, mainly because of the large output energy available (up to 10 J per pulse) and the wavelength of the light emitted (694 nm), which is fairly well matched to the peak sensitivity of available photographic materials for holography.

The active medium in a ruby laser is a rod, typically 5–10 mm in diameter and 75–100 mm in length, made of synthetic sapphire (Al_2O_3) doped with 0.05 percent of Cr_2O_3. When this rod is mounted in an optical resonator, as shown

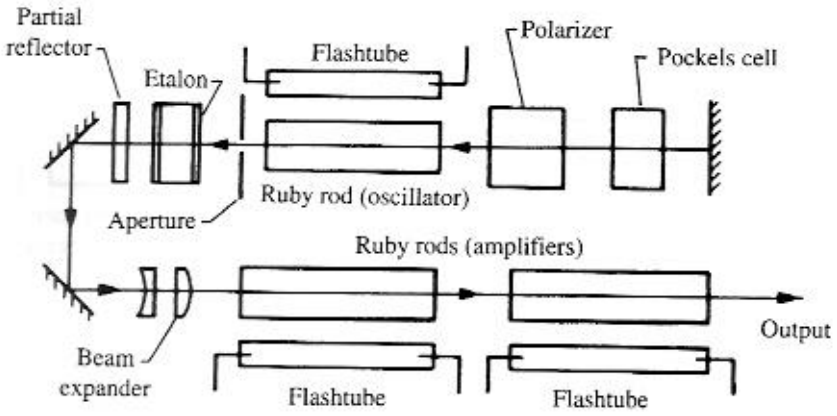


Fig. 4.2. Optical system of a Q-switched ruby laser.

in fig. 4.2, and pumped with a xenon flash lamp, it emits a series of pulses [Lengyel, 1971]. Single pulses are obtained by using a Q-switch (a Pockels cell) in the cavity.

Since the output from a ruby oscillator is limited, high-power ruby lasers use one or more additional ruby rods as amplifiers. With two stages of amplification, an output of 10 J can be obtained.

4.2 Coherence requirements

As mentioned earlier, to obtain a satisfactory hologram, the light used must be spatially coherent and its coherence length (see Appendix A) must be much greater than the maximum optical path difference between the object and reference beams in the recording system.

Spatial coherence is automatically ensured if the laser oscillates in the lowest order transverse mode (the TEM_{00} mode); this mode also gives the most uniform illumination.

Operation of Ar^+ and Kr^+ lasers on a single spectral line can be obtained, where necessary, by means of a wavelength selector prism. However, this may not ensure adequate temporal coherence since many lasers, even when operating on a single spectral line, will oscillate, as shown in the upper part of fig. 4.3, in a number of longitudinal modes lying within the profile of this spectral line, at which the gain of the laser medium is adequate to overcome the cavity losses. These modes correspond to the resonant frequencies of the laser cavity and are separated by a frequency interval

$$\Delta\nu = c/2L, \tag{4.1}$$

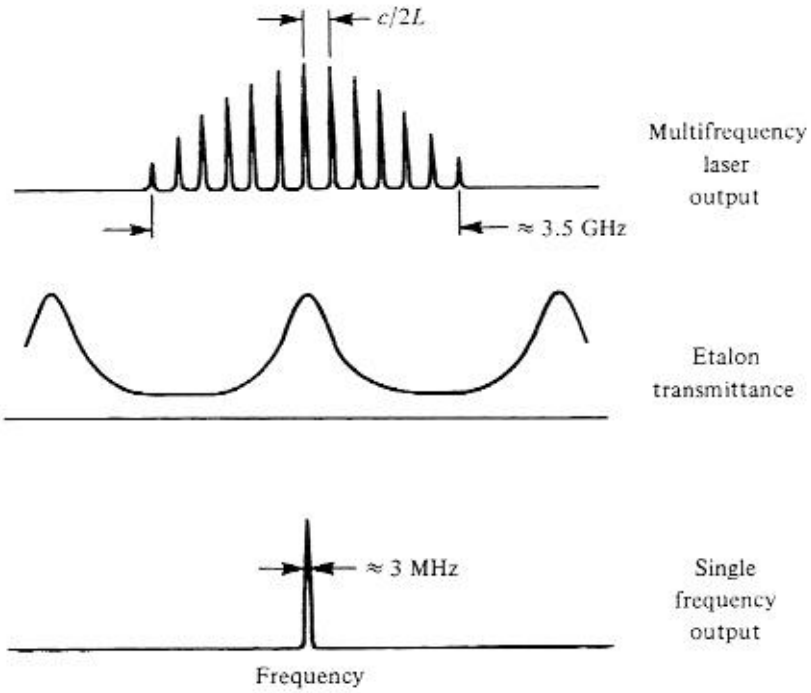


Fig. 4.3. Laser modes, without and with an intracavity etalon.

where c is the speed of light, and L is the length of the laser cavity. If we assume that the output power is divided equally between N longitudinal modes, the effective coherence length of the output is

$$\Delta l = 2L/N. \tag{4.2}$$

Equation (4.2) shows that the existence of more than one longitudinal mode in the output reduces the coherence length severely. Even if the mean optical paths of the object and reference beams are equalized carefully, severe restrictions are placed on the maximum depth of the object.

Depending on their power, commercial He–Ne lasers oscillate, typically, in two to five longitudinal modes, and the coherence length of the output is limited to a few centimeters. With high-power Ar^+ and Kr^+ lasers, it is possible to obtain operation in a single longitudinal mode, and coherence lengths in excess of a meter, by using an intracavity etalon. This etalon is tuned to obtain maximum power output, and stabilize the output wavelength, by mounting it in a temperature-controlled oven.

Diode lasers can be made to operate in a single longitudinal mode by increasing the drive current to a figure just below their maximum rating [Henry, 1991].

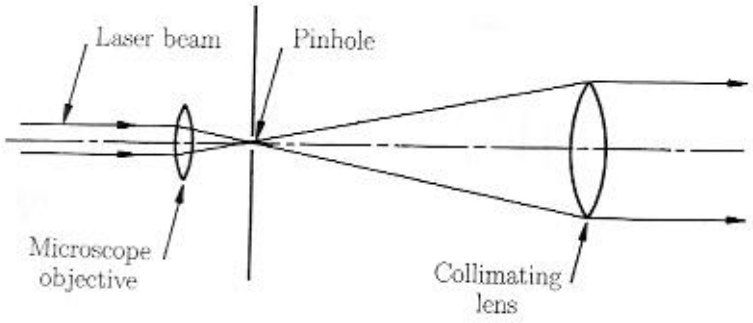


Fig. 4.4. Optical system used to expand and spatially filter a laser beam.

4.3 Laser beam expansion

Since the beam from most lasers has a diameter of only 1–2 mm and a very low divergence, it is usually necessary to use low-power microscope objectives to expand the laser beam to illuminate the object as well as the hologram. A problem is that, due to the high coherence of laser light, the expanded beam exhibits random diffraction patterns (spatial noise) produced by defects and dust on the optical surfaces in the path of the beam. These diffraction patterns can be eliminated, and a clean beam obtained, by placing a pinhole at the focus of the microscope objective, as shown in fig. 4.4.

If the laser is oscillating in the TEM_{00} mode, the beam has a Gaussian intensity profile given by the relation

$$I(r) = I(0)\exp\left(\frac{-2r^2}{w^2}\right), \tag{4.3}$$

where r is the radial distance from the center of the beam and w is the distance at which the intensity drops to $(1/e^2)$ of that at the center of the beam. Since the aperture of the microscope objective is usually greater than $2w$, the diameter of the focal spot is

$$d = \frac{2\lambda f}{\pi w}, \tag{4.4}$$

where f is the focal length of the microscope objective. If the diameter of the pinhole is less than d, randomly diffracted light is blocked, and the transmitted beam has a smooth profile.

An exceptional case, where it is possible to dispense with beam expansion, is with a diode laser. Because of the extremely small emitting area, a diode laser produces a clean, divergent beam similar to that obtained from a He–Ne laser after expansion and spatial filtration. Diode lasers can therefore be used

to record some types of holograms (see Section 7.2) without any external optics.

4.4 Beam splitters

If we define the beam ratio $R=(r/o)^2$ as the ratio of the irradiances of the reference and object beams, we would expect the visibility of the interference fringes forming the hologram to be a maximum when $R=1$. However, the wave scattered by a diffusely reflecting object exhibits very strong local variations in amplitude (see Section 2.5). As a result it is usually necessary to work with a value of $R \gg 1$ (typically $R \approx 3$), to avoid nonlinear effects.

A convenient way to vary the ratio of the powers in the two beams and optimize the visibility of the hologram fringes is to use a beam splitter consisting of a glass disc, which can be rotated, coated with a thin aluminum film whose reflectivity is a linear function of the azimuth. Such a variable-ratio beam splitter must be used in the unexpanded laser beam to minimize the variation in reflectivity across the beam.

4.5 Beam polarization

Most gas lasers have Brewster-angle windows on the plasma tube so that the output is linearly polarized. The visibility of the interference fringes forming the hologram is then a maximum when the electric vectors of the object and reference beams are parallel (see Appendix A). This condition is always satisfied if the two beams are linearly polarized with their electric vectors normal to the plane containing the beams. If, on the other hand, they are polarized with their electric vectors in the plane containing the beams, the visibility of the hologram fringes can drop to zero when the beams intersect at right angles.

It should be noted that in the case of an object with a rough surface, a substantial fraction of the reflected light may be depolarized. The resulting decrease in the visibility of the interference fringes can be minimized, where necessary, by using a sheet polarizer in front of the hologram to eliminate the cross-polarized component.

4.6 Pulsed laser holography

Very short light pulses (<20 nanoseconds) can be obtained with a pulsed laser, if a Pockels cell is used as a Q-switch in the laser cavity. As a result, problems of vibration and air currents are largely eliminated. Because their output

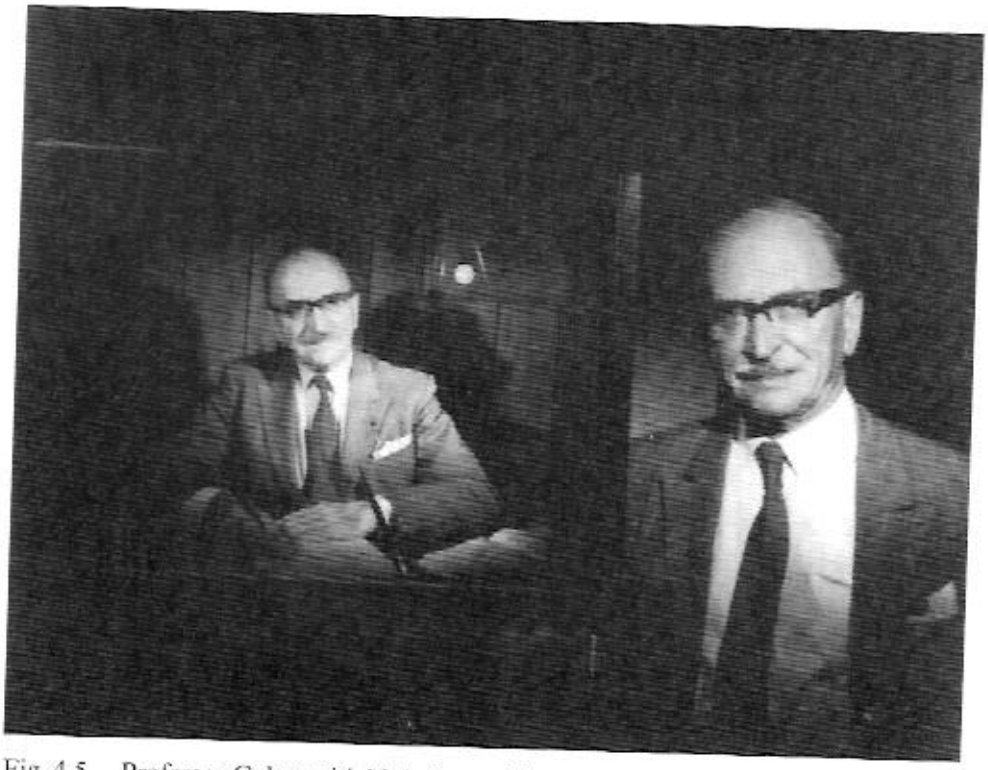


Fig. 4.5. Professor Gabor with his holographic portrait: this hologram was produced by R. Rinehart at the McDonnell Douglas Electronics company in 1971, using a pulsed ruby laser.

wavelength is well matched to the peak sensitivity of available photographic materials, and their output energy is fairly large, pulsed ruby lasers are used widely to record holograms in an industrial environment [Koechner, 1979].

A Q-switched ruby laser can also be used to record holograms of living human subjects [Siebert, 1968] as shown in fig. 4.5. However, to avoid eye damage (see Section 4.7), instead of illuminating the subject directly, the expanded laser beam should be allowed to fall on a large diffuser which constitutes an extended source illuminating the subject.

A studio setup for making holographic portraits has been described by Bjelkhagen [1992].

4.7 Laser safety

Since the beam from a laser is focused by the lens of the eye to a very small spot on the retina, direct exposure to low power lasers can cause eye damage. With pulsed lasers, even stray reflections can be dangerous. It is essential to take all

due precautions to avoid accidental exposure and, where required, to use appropriate eye protection [Sliney & Wolbarsht, 1980].

References

Bjelkhagen, H. I. (1992). Holographic portraits made by pulse lasers. *Leonardo*, **25**, 443–8.

Henry, C. H. (1991). Theory of the linewidth of semiconductor lasers. In *Semiconductor Diode Lasers*, Vol. 1, pp. 36–41. New York: IEEE.

Koechner, W. (1979). Solid state lasers. In *Handbook of Optical Holography*, ed. H. J. Caulfield, pp. 257–67. New York: Academic Press.

Lengyel, B. A. (1971). *Lasers*. New York: Wiley-Interscience.

Siebert, L. D. (1968). Large scene front-lighted hologram of a human subject. *Proceedings of the IEEE*, **56**, 1242–3.

Sliney, D. & Wolbarsht, M. (1980). *Safety with Lasers and Other Optical Sources: a Comprehensive Handbook*. New York: Plenum Press.

Problems

4.1. A He Ne laser with a 200 mm long resonant cavity oscillates in two longitudinal modes. What is the coherence length of the radiation?

From (4.2) the coherence length of the radiation is

$$\begin{aligned}\Delta l &= 2L/N \\ &= (2\times 0.2/2)\ \text{m} \\ &= 0.2\ \text{m}.\end{aligned} \tag{4.5}$$

4.2. In the arrangement shown in fig. 4.4, the central part of the beam from a He–Ne laser is isolated by an aperture with a diameter of 2.0 mm and brought to a focus by a microscope objective with a focal length of 32 mm. What would be a suitable size for the pinhole?

The diameter of the focal spot is equal to the diameter of the Airy disc which, in this case, is

$$\begin{aligned}d &= \frac{2.44\times 0.633\times 10^{-6}\times 32\times 10^{-3}}{2\times 10^{-3}}\ \text{m} \\ &= 24.4\ \mu\text{m}.\end{aligned} \tag{4.6}$$

A pinhole with a diameter of 20 μm would ensure a clean beam with only a marginal loss of light.

5

The recording medium

5.1 Amplitude and phase holograms

The complex amplitude transmittance of the medium used to record a hologram can be written as

$$\begin{aligned} \mathbf{t} &= \exp(-\alpha d)\exp[-\mathrm{i}(2\pi n d/\lambda)], \\ &= |\mathbf{t}|\exp(-\mathrm{i}\phi), \end{aligned} \tag{5.1}$$

where α is the absorption constant of the medium, d is its thickness and n is its refractive index. A change in α with the exposure produces an amplitude hologram, while a change in n or d will produce a phase hologram.

The response of a recording material used for amplitude holograms can be characterized on a macroscopic scale, by a curve of the amplitude transmittance as a function of the exposure (a $|\mathbf{t}| - E$ curve), as shown in fig. 5.1.

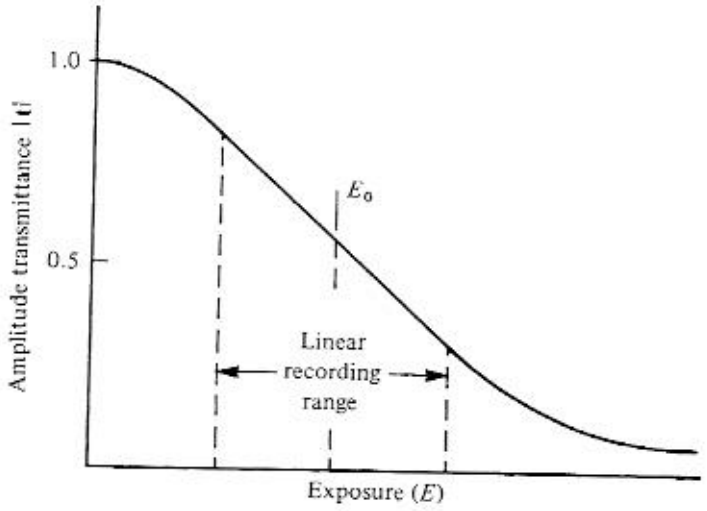


Fig. 5.1. Typical amplitude transmittance vs. exposure curve for a recording material.

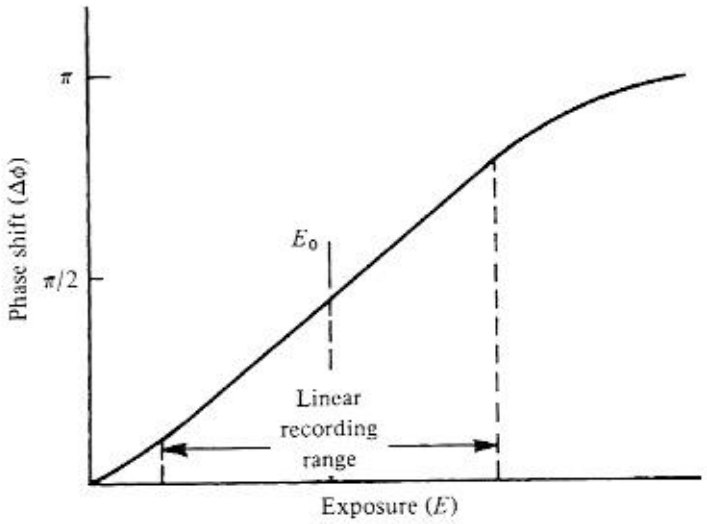


Fig. 5.2. Typical phase shift vs. exposure curve for a recording material.

Similarly, the response of a recording material used for phase holograms can be described by a curve of the effective phase shift as a function of the exposure (a $\Delta\phi - E$ curve), as shown in fig. 5.2.

5.2 The modulation transfer function

Curves such as those shown in figs. 5.1 and 5.2 may not describe the response of the recording medium on a microscopic scale because the actual intensity distribution to which the recording medium is exposed always differs from that incident on it, due to scattering and absorption in the medium. As a result, the actual modulation of the intensity within the recording medium is always less than that in the original interference pattern. In addition, the response of the medium at different spatial frequencies may be affected by the processing technique that is used.

For a given modulation of the input intensity, the ratio of the response of the recording medium at any spatial frequency s, relative to that at low spatial frequencies ($s \rightarrow 0$), is specified by a parameter $M(s)$, termed the modulation transfer function.

5.3 Effects of nonlinearity

For simplicity, we have assumed so far that the amplitude transmittance of the hologram is a linear function of the intensity, as described by (1.2). However,

in practice, this assumption is not always valid. The amplitude transmittance of the processed recording material can then be represented by a polynomial [Bryngdahl & Lohmann, 1968]

$$\begin{aligned}\mathbf{t} &= \mathbf{t}_0 + \beta_1 TI + \beta_2 T^2 I^2 + \cdots \\ &= \mathbf{t}_0 + \beta_1 T[rr^* + oo^* + r^* o + ro^*] \\ &\quad + \beta_2 T^2 [rr^* + oo^* + r^* o + ro^*]^2 \\ &\quad + \cdots.\end{aligned} \tag{5.2}$$

If the hologram is illuminated once again with a plane wave of unit amplitude, the complex amplitude of the wave transmitted by the hologram can be written in the form

$$\begin{aligned}u &= \text{linear terms} \\ &\quad + \beta_2 T^2 [(oo^*)^2 + o^2 + o^{*2} + 2o^2 o^* + 2oo^{*2}] \\ &\quad + \cdots.\end{aligned} \tag{5.3}$$

As can be seen, nonlinearity leads to the production of additional spurious terms. An examination of (5.3) shows that the term involving $(oo^*)^2$ results in a doubling of the width of the halo surrounding the directly transmitted beam, while the terms involving o^2 and o^{*2} correspond to higher-order diffracted images, and the terms involving $2o^2 o^*$ and $2oo^{*2}$ are intermodulation terms, producing false images.

The effects of nonlinearity are particularly noticeable with phase holograms. Even if we assume that the phase shift produced by the recording medium is proportional to the exposure, the complex amplitude transmittance is given by the expression

$$\begin{aligned}\mathbf{t} &= \exp(-\mathrm{i}\phi) \\ &= 1 - \mathrm{i}\phi - (1/2)\phi^2 + (1/6)\mathrm{i}\phi^3 + \cdots.\end{aligned} \tag{5.4}$$

If the phase modulation is increased, to obtain better diffraction efficiency, the effects of the higher-order terms cannot be neglected.

However, with volume holograms, the effects of nonlinearity are reduced significantly by the angular selectivity of the hologram. If the angle between the beams in the recording setup is large enough that the diffracted beams corresponding to different orders do not overlap, a simple analysis [Hariharan, 1979] shows that the signal-to-noise ratio should improve by a factor approximately equal to $(\psi/\Delta\theta)$, where $2\Delta\theta$ is the width of the passband of the angular selectivity function, and ψ is the angle subtended by the object at the hologram.

References

Bryngdahl, O. & Lohmann, A. (1968). Nonlinear effects in holography. *Journal of the Optical Society of America*, **58**, 1325–34.
Hariharan, P. (1979). Intermodulation noise in amplitude holograms: the effect of hologram thickness. *Optica Acta*, **26**, 211–5.

Problems

5.1. A hologram is recorded in an optical system in which the object and reference beams make angles of $+30°$ and $-30°$, respectively, with the normal to the photographic film. After exposure, the film is processed to produce a volume phase hologram. If the object subtends an angle of 30° at the photographic film, and the thickness of the photographic emulsion layer is 15 μm, what is the improvement in the signal-to-noise ratio due to the angular selectivity of the hologram?

Since the modulation parameter for such a hologram would be significantly lower than the optimum value of $\pi/2$ (say, $\pi/4$), it follows from fig. 3.3 that the diffraction efficiency drops to zero when the dephasing parameter $\chi \approx 3$. From Problems 1.1 and 3.1, this would correspond to a deviation from the Bragg angle $\Delta\theta = \pm 2.31°$. The signal-to-noise ratio should improve by a factor of $(30/4.62) \approx 6.5$.

6

Recording materials

Several recording materials have been used for holography [Smith, 1977]. Table 6.1 lists the principal characteristics of those that have been found most useful.

6.1 Photographic materials

High-resolution photographic plates and films were the first materials used to record holograms. They are still used widely because of their relatively high sensitivity when compared to other hologram recording materials [Bjelkhagen, 1993]. In addition, they can be dye sensitized so that their spectral sensitivity matches the most commonly used laser wavelengths.

Conventional processing produces an amplitude hologram. It also results in a reduction in the thickness of the emulsion layer of about 15 percent, due to the removal of the unexposed silver halide grains in the fixing bath. This reduction in thickness can cause a rotation of the fringe planes as well as a reduction in their spacing (see Section 3.3).

With volume reflection holograms (see Section 3.2.2), the reduction in the

Table 6.1. *Recording materials for holography*

Material	Exposure (J/m^2)	Resolution (mm^{-1})	Processing	Type	η_{max}
Photographic	≈1.5	≈5000	Normal	Amplitude	0.06
			Bleach	Phase	0.60
DCG	10^2	10000	Wet	Phase	0.90
Photoresists	10^2	3000	Wet	Phase	0.30
Photopolymers	$10–10^4$	5000	Dry	Phase	0.90
PTP	10^{-1}	500–1200	Dry	Phase	0.30
BSO	10	10000	None	Phase	0.20

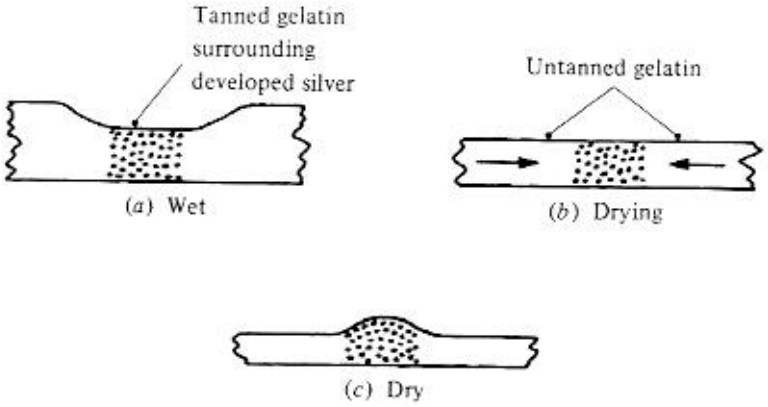


Fig. 6.1. Formation of a relief image by local tanning of the gelatin.

spacing of the fringe planes is immediately apparent as a shift in the color of the image, which is reconstructed at a shorter wavelength than that used to record it.

With volume transmission holograms, emulsion shrinkage can result in a reduction in diffraction efficiency when the hologram is replaced in the original recording setup, since the reference beam is no longer incident on it at the Bragg angle (see Section 3.2.1). This loss in diffraction efficiency can be minimized by having the object and reference beams incident at equal, but opposite, angles on the hologram, so that the fringe planes are normal to the surface of the photographic emulsion.

Amplitude holograms recorded in photographic emulsions can also exhibit phase modulation due to a surface relief image arising from local tanning (hardening) of the gelatin by the oxidation products of the developer [Smith, 1968]. After developing and fixing, as shown in fig. 6.1, the unexposed areas of the wet emulsion layer, which are not tanned, are swollen to more than five times their normal thickness and are very soft. The tanned gelatin in the exposed areas absorbs less water and, therefore, dries more quickly. Shrinkage during drying pulls some of the soft gelatin from adjacent unexposed areas into the exposed areas, so that, when the emulsion has dried, the exposed areas stand out in relief.

Since this relief image is confined to low spatial frequencies (≤ 200 mm^{-1}), it normally contributes very little to the reconstructed image, but can give rise to intermodulation terms.

Because of the low diffraction efficiency of amplitude holograms, holograms produced for displays using photographic materials are usually processed to obtain volume phase holograms (see Section 3.2) which have much higher

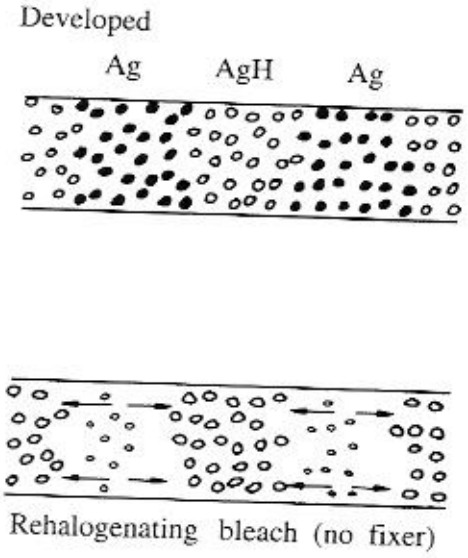


Fig. 6.2. Production of photographic volume phase holograms by a rehalogenating bleach.

diffraction efficiencies. The most common procedure is to use a rehalogenating bleach bath, without fixing, to convert the developed silver back into a transparent silver halide with a high refractive index [Hariharan, Ramanathan & Kaushik, 1971; Phillips *et al.*, 1980]. The regenerated silver halide goes into solution and is redeposited on the adjacent unexposed silver halide grains. The result is that, as shown in fig. 6.2, a volume phase hologram is produced by material transfer from the exposed areas to the adjacent unexposed areas [Hariharan, 1990].

Material transfer through diffusion is effective only over a very limited distance. As a result, the diffraction efficiency of the hologram drops off rapidly, as shown in fig. 6.3, for fringe spacings greater than a critical value corresponding to the diffusion length of the silver ion in the bleach bath [Hariharan & Chidley, 1988].

Since, with a rehalogenating bleach, little or no silver or silver halide is removed from the emulsion layer, the reduction in the thickness of the emulsion due to processing is minimal. In the case of a volume reflection hologram, the change in the color of the reconstructed image is quite small, even with a nontanning developer, while the use of a tanning developer may result in an increase in thickness, so that the image is reconstructed at a longer wavelength [Hariharan & Chidley, 1989].

Until a few years ago, the most commonly used photographic plates and films were Kodak 649F and Agfa 8E75HD and 8E56HD. After the

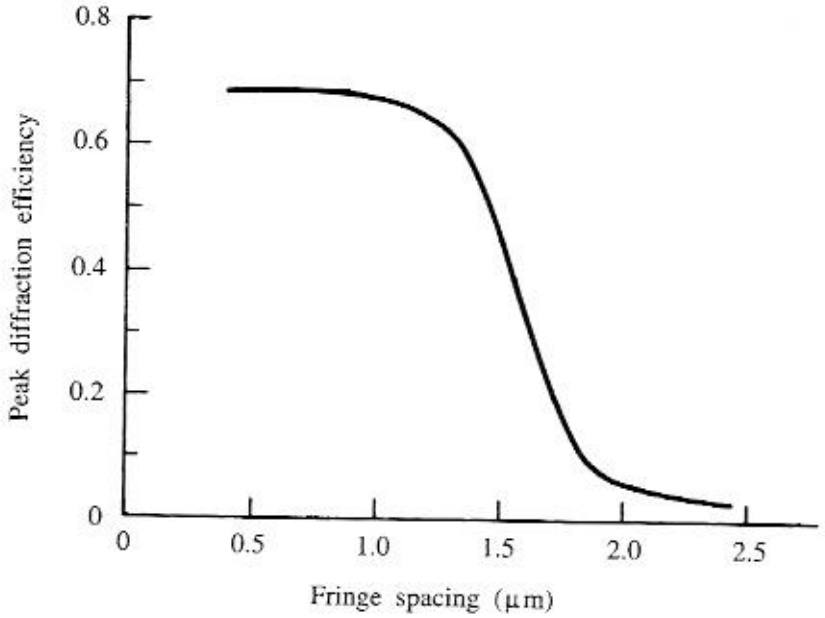


Fig. 6.3. Peak diffraction efficiency, plotted as a function of the fringe spacing, for transmission gratings produced in a photographic emulsion processed with a rehalogenating bleach [Hariharan & Chidley, 1988].

manufacture of these materials was discontinued, their place has been taken by the BB emulsions produced by Holographic Recording Technologies and the Slavich emulsions marketed by Geola.

The BB series of plates includes three emulsions (BB-640, BB-520 and BB-450) sensitized, respectively, for red, green and blue laser light. They have a grain size of 20–25 nm and require an exposure of approximately 1.5 J/m^2 to obtain a density of 2.5 when developed in a metol-ascorbate developer. An emulsion with panchromatic sensitization (BB-PAN) is also available for making full-color reflection holograms (see Section 8.1), as well as one (BB-700) for use with pulsed lasers.

The Slavich series of films and plates features two emulsions (PFG-01 and VRP-M, sensitized for red and green laser light, respectively), which are similar to Agfa 8E75HD and 8E56HD, as well as two emulsions (PFG-03M, sensitized for the red, and PFG-03C, with panchromatic sensitization) with ultra-fine grains. The latter two emulsions require exposures of 15–20 J/m^2 but can resolve more than 5000 lines/mm.

Information on the BB emulsions is available at http://www.hrt-gmbh.de, and on the Slavich plates and films at http://www.geola.com.

6.2 Dichromated gelatin

A volume phase hologram can be recorded in a gelatin layer containing a small amount of a dichromate, such as $(NH_4)_2Cr_2O_7$, by making use of the fact that dichromated gelatin (DCG) becomes locally hardened on exposure to light, due to the formation of cross-links between the carboxylate groups on neighboring gelatin chains. This effect is used to obtain a local modulation of the refractive index.

After exposure in the holographic system using blue light from an Ar^+ laser ($\lambda = 488$ nm), the gelatin layer is washed in water at 20–30 °C for 10 min, so that it absorbs water and swells. The swollen gelatin layer is then immersed in two successive baths of isopropanol, to extract the water, and dried thoroughly. With care in processing, volume phase holograms with high diffraction efficiency and low scattering can be produced [Chang & Leonard, 1979; Jeong, Song & Lee, 1991].

A method commonly used to prepare plates coated with DCG is to dissolve and remove the silver halide from the emulsion layer in a photographic plate by soaking it in a nonhardening fixing bath. The plates are sensitized by soaking them for about 5 min at 20 °C in an aqueous solution of $(NH_4)_2Cr_2O_7$ to which a small amount of a wetting agent has been added. They are then allowed to drain and dried at 25–30 °C in darkness.

Geola also make plates (PFG-04) coated with dichromated gelatin which can be used directly to record reflection holograms. These plates require an exposure of 1000 J/m^2 at 488 nm and can yield gratings with a diffraction efficiency of 0.8.

6.3 Silver-halide sensitized gelatin

This technique makes it possible to combine the high sensitivity of photographic materials with the high diffraction efficiency, low scattering and high light-stability of DCG.

In this technique [Pennington, Harper & Laming, 1971], the exposed photographic emulsion is developed in a metol-hydroquinone developer and then bleached in a bath containing $(NH_4)_2Cr_2O_7$. During the bleaching process, the developed silver is oxidized to Ag^+, while the Cr^{6+} ions in the bleach are reduced to Cr^{3+} ions. The oxidation products of the developer, as well as the Cr^{3+} ions formed by reduction of the bleach, form cross-links between the gelatin chains in the vicinity of the oxidized silver grains, causing local hardening of the gelatin [Hariharan, 1986]. The emulsion is then fixed to remove the unexposed silver halide, washed, dehydrated with isopropanol and dried exactly as for a

dichromated gelatin hologram. With this technique it is possible to obtain diffraction efficiencies up to 80 percent with transmission gratings and 55 percent with reflection gratings [Angell, 1987; Fimia, Pascual & Belendez, 1992].

Optimized processing techniques for SHSG holograms using BB-640 plates have been described by Belendez *et al.* [1998].

6.4 Photoresists

In positive photoresists, such as Shipley AZ-1350, the areas exposed to light become soluble and are washed away during development to produce a relief image [Bartolini, 1977].

The photoresist is coated on a glass substrate by spinning to form a layer 1–2 μm thick. This layer is then baked at 75 °C for 15 min to ensure complete removal of the solvent. Holograms are recorded with a He–Cd laser at a wavelength of 442 nm. The exposed plate is processed in AZ-303 developer diluted with four parts of distilled water.

Holograms recorded on a photoresist can be replicated, using a thermoplastic (see Section 9.2). Multiple copies of holographic optical elements (see Section 12.4) can also be made.

6.5 Photopolymers

Several organic materials can be activated by a photosensitizer to produce refractive index changes, due to photopolymerization, when exposed to light [Booth, 1977]. A commercial photopolymer is also available coated on a polyester film base (DuPont OmniDex) that can be used to produce volume phase holograms with high diffraction efficiency [Smothers *et al.*, 1990].

The film is supplied with a polyester cover sheet laminated on to the tacky photopolymer layer. Since the exposure required is around 300 J/m^2 at 514 nm, holograms are recorded by contact copying a master hologram (see Section 9.1). Close contact is ensured by removing the cover sheet and laminating the tacky film to the master hologram. The film is then exposed to UV light to cure the photopolymer, after which it can be separated from the master hologram. Finally the film is baked at 100–120 °C for 1–2 hours to obtain increased index modulation [Smothers *et al.*, 1990; Weber *et al.*, 1990].

This procedure yields volume reflection holograms that reconstruct an image at almost the same wavelength as that used to record them. If necessary, the reconstruction can be shifted to a longer wavelength by laminating the holographic recording film, after exposure to UV light, to a color tuning film and baking the sandwich [Zager & Weber, 1991].

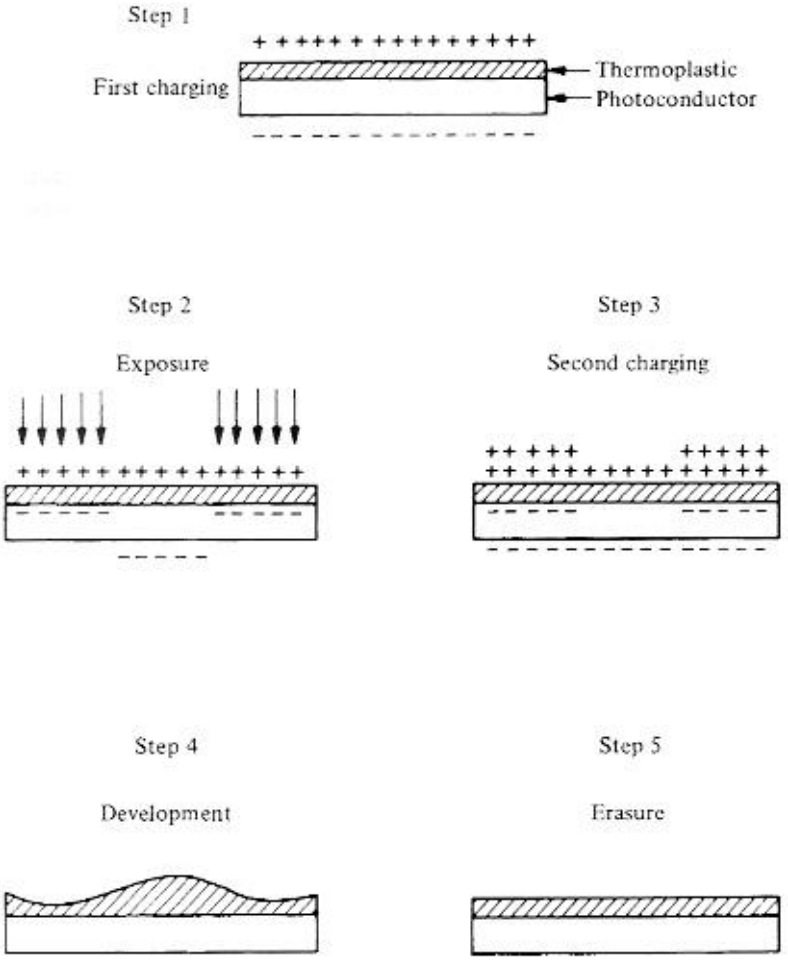


Fig. 6.4. Record erase cycle for a photothermoplastic [Lin and Beauchamp, 1970].

6.6 Photothermoplastics (PTP)

A hologram can be recorded in a multilayer structure consisting, as shown in fig. 6.4, of a glass or Mylar substrate coated with a thin, transparent, conducting layer of indium oxide, a photoconductor, and a thermoplastic [Lin & Beauchamp, 1970; Urbach, 1977].

The film is initially sensitized in darkness by applying a uniform electric charge to the top surface. On exposure and recharging, a spatially varying electrostatic field is created. The thermoplastic is then heated briefly, so that it becomes soft enough to be deformed by this field, and cooled to fix the variations in thickness.

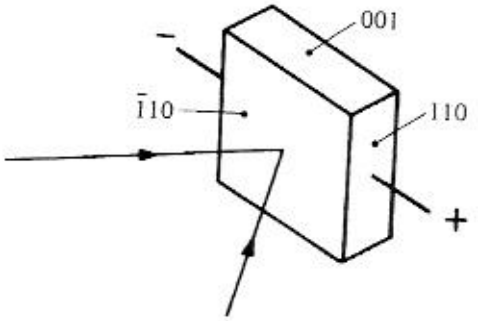


Fig. 6.5. Hologram recording configuration for BSO [Huignard, 1981].

Photothermoplastics have a reasonably high sensitivity and yield a thin phase hologram with good diffraction efficiency. They have the advantage that they can be processed rapidly *in situ*; in addition, if they are produced on a glass substrate, the hologram can be erased by heating the substrate, and the material reused.

6.7 Photorefractive crystals

When a photorefractive crystal is exposed to a spatially varying light pattern, electrons are liberated in the illuminated areas. These electrons migrate to adjacent dark regions and are trapped there. The spatially varying electric field produced by this space-charge pattern modulates the refractive index through the electro-optic effect, producing the equivalent of a phase grating. The space-charge pattern can be erased by uniformly illuminating the crystal, after which another recording can be made.

The photorefractive crystals most commonly used for recording holograms are Fe-doped $LiNbO_3$ and $Bi_{12}SiO_{20}$ (BSO),which has a higher sensitivity. The best results are obtained with BSO with the recording configuration shown in fig. 6.5, in which an electric field is applied at right angles to the hologram fringes [Huignard & Micheron, 1976; Huignard, 1981].

Typical recording erasure curves for BSO at different applied fields, at an incident power density of 2.45 W/m^2 ($\lambda = 514$ nm) are presented in fig. 6.6. A maximum diffraction efficiency of 0.25 can be obtained with a field of 900 V/mm and the hologram can be stored in darkness for about 30 h. Since readout is destructive, the reconstructed image is best recorded and stored for viewing.

Several interesting possibilities have been opened up by such photorefractive materials.

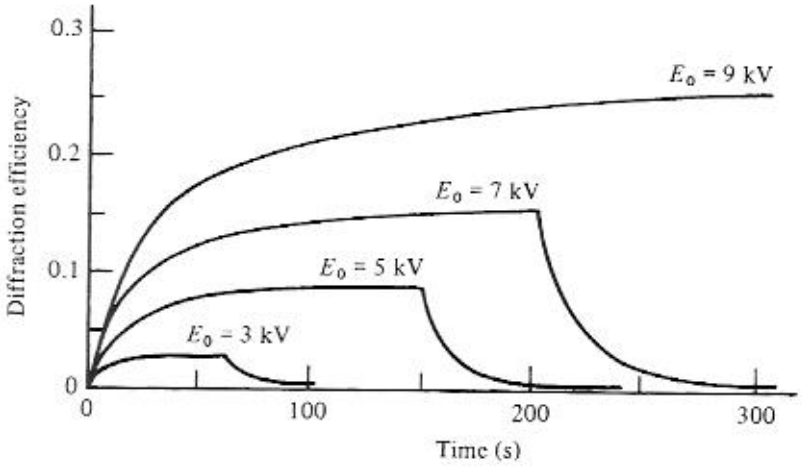


Fig. 6.6. Record–erase cycles for BSO at different applied fields (sample size $10 \times 10 \times 10$ mm) [Huignard & Micheron, 1976].

References

Angell, D. K. (1987). Improved diffraction efficiency of silver halide (sensitized) gelatin. *Applied Optics*, **26**, 4692–702.

Bartolini, R. A. (1977). Photoresists. In *Holographic Recording Materials*, Topics in Applied Physics, vol.20, ed. H. M. Smith, pp. 209–27. Berlin: Springer-Verlag.

Belendez, A., Neipp, C., Flores, M. & Pascual, I. (1998). High-efficiency silver halide sensitized gelatin holograms with low absorption and scatter. *Journal of Modern Optics*, **45**, 1985–92.

Bjelkhagen, H. I. (1993). *Silver Halide Materials for Holography & Their Processing*. Berlin: Springer-Verlag.

Booth, B. L. (1977). Photopolymer laser recording materials. *Journal of Applied Photographic Engineering*, **3**, 24–30.

Chang, B. J. & Leonard, C. D. (1979). Dichromated gelatin for the fabrication of holographic optical elements. *Applied Optics*, **18**, 2407–17.

Fimia, A., Pascual, I. & Belendez, A. (1992). Optimized spatial frequency response in silver halide sensitized gelatin. *Applied Optics*, **31**, 4625–7.

Hariharan, P. (1986). Silver-halide sensitized gelatin holograms: mechanism of hologram formation. *Applied Optics*, **25**, 2040–2.

Hariharan, P. (1990). Basic processes involved in the production of bleached holograms. *Journal of Photographic Science*, **38**, 76–81.

Hariharan, P. & Chidley, C. M. (1988). Rehalogenating bleaches for photographic phase holograms. 2: Spatial frequency effects. *Applied Optics*, **27**, 3852–4.

Hariharan, P. & Chidley, C. M. (1989). Bleached reflection holograms: a study of color shifts due to processing. *Applied Optics*, **28**, 422–4.

Hariharan, P., Ramanathan, C. S. & Kaushik, G. S. (1971). Simplified processing technique for photographic phase holograms. *Optics Communications*, **3**, 246–7.

Huignard, J. P. (1981). Phase conjugation, real time holography and degenerate four-wave mixing in photoreactive BSO crystals. In *Current Trends in Optics*, ed. F. T. Arecchi & F. R. Aussenegg, pp. 150–60. London: Taylor & Francis.

Huignard, J. P. & Micheron, F. (1976). High sensitivity read-write volume holographic storage in $Bi_{12}SiO_{20}$ and $Bi_{12}GeO_{20}$ crystals. *Applied Physics Letters*, **29**, 591–3.
Jeong, M. H., Song, J. B. & Lee, I. W. (1991). Simplified processing method of dichromated gelatin holographic recording material. *Applied Optics*, **30**, 4172–3.
Lin, L. H. & Beauchamp, H. L. (1970). Write-read-erase in situ optical memory using thermoplastic holograms. *Applied Optics*, **9**, 2088–92.
Pennington, K. S., Harper, J. S. & Laming, F. P. (1971). New phototechnology suitable for recording phase holograms and similar information in hardened gelatine. *Applied Physics Letters*, **18**, 80–4.
Phillips, N. J., Ward, A. A., Cullen, R. & Porter, D. (1980). Advances in holographic bleaches. *Photographic Science & Engineering*, **24**, 120–4.
Smith, H. M. (1968). Photographic relief images. *Journal of the Optical Society of America*, **58**, 533–9.
Smith, H. M., ed. (1977). *Holographic Recording Materials*. Berlin: Springer-Verlag.
Smothers, W. K., Monroe, B. M., Weber, A. M. & Keys, D. E. (1990). Photopolymers for holography. In *Practical Holography IV*, Proceedings of the SPIE, vol. 1212, ed. S. A. Benton, pp. 20–9, Bellingham: SPIE.
Urbach, J. C. (1977). Thermoplastic hologram recording. In *Holographic Recording Materials*, Topics in Applied Physics, vol. 20, ed. H. M. Smith, pp. 161–207, Berlin: Springer-Verlag.
Weber, A. M., Smothers, W. K., Trout, T. J. & Mickish, D. J. (1990). Hologram recording in Du Pont's new photopolymer materials. In *Practical Holography IV*, Proceedings of the SPIE, vol. 1212, ed. S. A. Benton, pp. 30–9, Bellingham: SPIE.
Zager, S. A. & Weber, A. M. (1991). Display holograms in Du Pont's OmniDex films. In *Practical Holography V*, Proceedings of the SPIE, vol. 1461, ed. S. A. Benton, pp. 58–67, Bellingham: SPIE.

Problems

6.1. A hologram is to be recorded with a He–Ne laser on a BB-640 plate. The illumination level in the hologram plane due to the object beam is 0.003 W/m^2, and that due to the reference beam is 0.009 W/m^2. The plate is to be processed to produce a phase hologram. What is the exposure time required?

To obtain good diffraction efficiency after bleaching, the exposure should result in an optical density of at least 2.5 after development. From the published data for BB-640 plates, this would correspond to an exposure of approximately 1.5 J/m^2.

Accordingly, the exposure time required would be

$$\begin{aligned} T &\approx 1.5/(0.009+0.003) \\ &\approx 125 \text{ seconds.} \end{aligned} \tag{6.1}$$

The exposure time yielding the maximum diffraction efficiency can be selected from trials with exposures ranging from 100 seconds to 200 seconds.

7

Display holograms

7.1 Transmission holograms

A typical optical system for recording transmission holograms is shown in fig.7.1.

Making a hologram involves recording a two-beam interference pattern. Any change in the phase difference between the two beams during the exposure results in a movement of the fringes and reduces modulation in the hologram. Accordingly, to avoid mechanical disturbances, all the optical components, as well as the object and the photographic film or plate, should be mounted on a rigid surface resting on shock absorbers. The system is adjusted to have a low natural frequency of vibration (<1 Hz). A concrete or granite slab resting on inflated scooter inner tubes can be used, but most experimenters prefer a

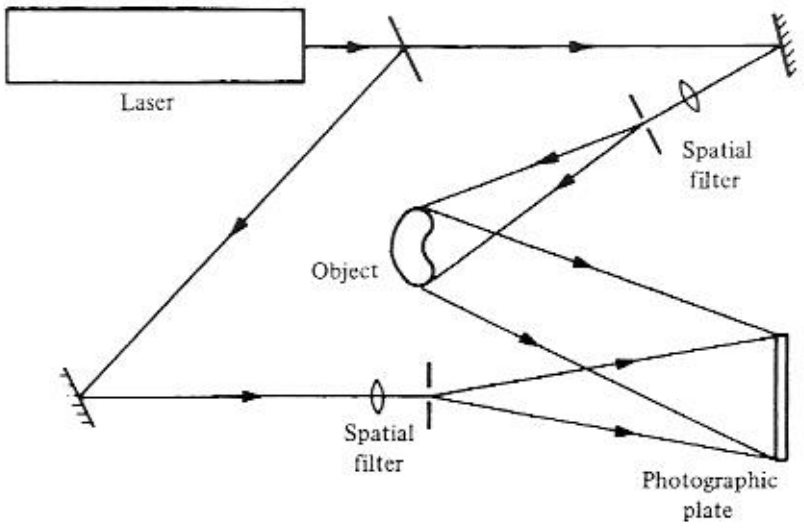


Fig. 7.1. Optical arrangement used to record a transmission hologram.

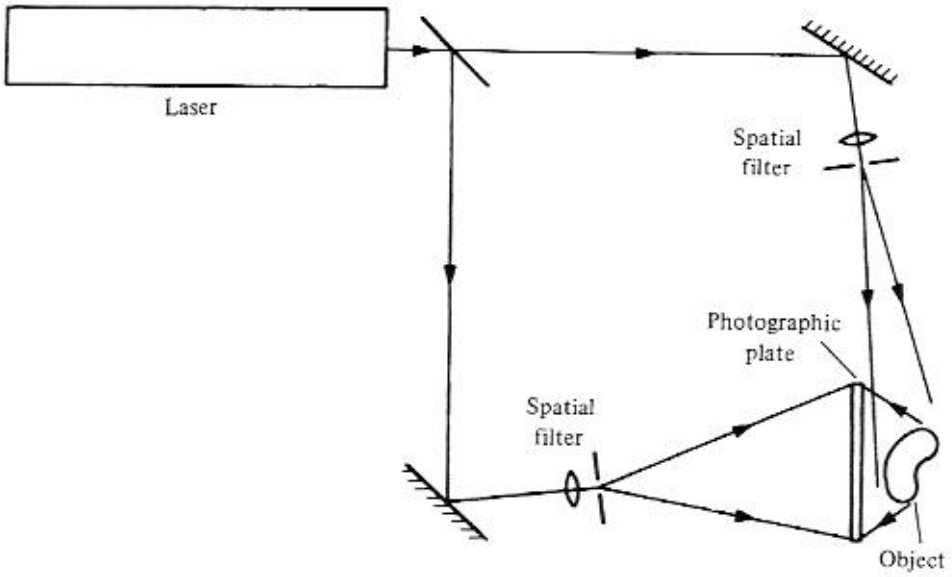


Fig. 7.2. Optical arrangement used to record a reflection hologram.

honeycomb optical table, with a steel top, supported on pneumatic legs. The steel top has the advantage that optical components can be bolted down to its surface or mounted on magnetic bases. The effects of air currents and acoustic disturbances can be minimized by enclosing the working area with heavy curtains. An accurate power meter should be used to measure the intensities of the object and reference beams in the hologram plane, in order to set the ratio of their intensities to a suitable value and calculate the required exposure.

Where very long exposures have to be made, residual disturbances can be eliminated by a feedback system which stabilizes the optical path difference between the beams [Neumann & Rose, 1967]. Any motion of the interference fringes in the hologram plane is picked up by a photodetector, and the variations in its output are amplified and applied to a piezoelectric translator (PZT) which controls the position of one of the mirrors in the beam path.

7.2 Reflection holograms

A typical optical system that can be used for recording reflection holograms is shown schematically in fig. 7.2.

As can be seen, the object and reference waves are incident on the photographic emulsion from opposite sides. Since the thickness of the photographic emulsion is typically between 6 μm and 15 μm, the interference fringes are recorded as layers within it, about half a wavelength apart. The hologram exposure should

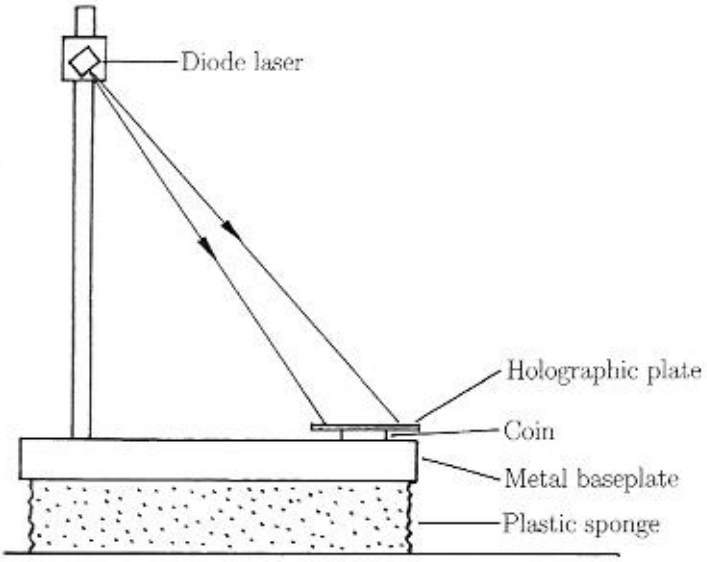


Fig. 7.3. Recording arrangement for reflection holograms using a diode laser.

be adjusted to give a density greater than 2.0, after development. The hologram is then processed in a rehalogenating bleach to control emulsion shrinkage and the resulting wavelength shift [Hariharan & Chidley, 1989].

Reflection holograms of objects with limited depth can be made with an arrangement similar to that used originally by Denisyuk [1965], in which the portion of the reference beam transmitted by the photographic plate illuminates the object. A simple overhead configuration for recording such holograms, using a diode laser, has been described by Koch & Petros [1998].

Because of the small size of the diode laser, the entire optical system can be put together on a $20 \times 25 \times 1$ cm metal baseplate, which, as shown in fig. 7.3, is placed on a thick sheet of plastic sponge to provide vibration isolation. The diode laser (see Section 4.3) produces a clean beam whose divergence (typically, about $8° \times 25°$) makes it possible to illuminate the object through the holographic plate without any external optics. A hologram of a flat, specular reflecting object, such as a coin, can be recorded by placing a high resolution photographic plate directly on the object.

7.3 Full-view holograms

A drawback of conventional holograms is the limited angle over which they can be viewed. Holograms that give a full 360° view of an object can be recorded using either four plates, or a cylinder of film, surrounding the object.

A very simple optical system for this purpose [Jeong, 1967] is shown in

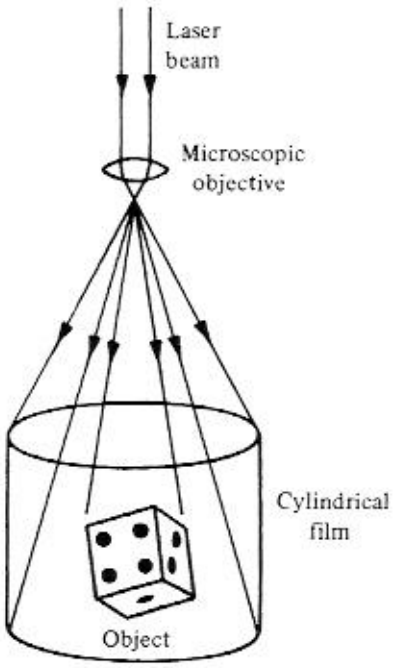


Fig. 7.4. Simple optical arrangement for making a 360° hologram [Jeong, 1967].

fig. 7.4. The object is placed at the center of a glass cylinder which has a strip of photographic film taped to its inner surface with the emulsion side facing inwards, and the expanded laser beam is incident on the object from above. The central portion of the expanded laser beam illuminates the object, while the outer portions, which fall directly on the film, constitute the reference beam.

To view the reconstructed image, the processed film is replaced in its original position and illuminated with the same laser beam.

7.4 Rainbow holograms

The rainbow hologram is a transmission hologram which reconstructs a bright, sharp, monochromatic image when illuminated with white light [Benton, 1977].

As shown schematically in fig. 7.5(*a*), the first step in making a rainbow hologram is to record a conventional transmission hologram of the object.

When this primary hologram (H_1) is illuminated, as shown in fig. 7.5(*b*), by the conjugate of the original reference wave, it reconstructs the conjugate of the original object wave and produces a real image of the object with unit magnification. A horizontal slit is then placed over H_1, as shown in figs. 7.5(*c*) and 7.5(*d*), and a second hologram (H_2) is recorded of the real image produced by H_1. The reference beam for H_2 is a convergent beam inclined in the vertical

plane, and the photographic plate used to record H_2 is placed so that the real image formed by H_1 straddles it.

When H_2 is illuminated with the conjugate of the reference beam used to make it, it forms an orthoscopic image of the object straddling the plane of the hologram, as shown in fig. 7.6(*a*). In addition, it also forms a real image of the slit placed across H_1. All the light diffracted by the hologram passes through this slit pupil, so that a very bright image is seen from this position. Since the observer can move his head from side to side, horizontal parallax is retained. However, the image disappears if the observer's eyes move outside this slit pupil, so that vertical parallax is eliminated.

With a white light source, the slit image is dispersed in the vertical plane, as shown in fig. 7.6(*b*), to form a continuous spectrum. An observer whose eyes are positioned at any part of this spectrum then sees a sharp, three-dimensional image of the object in the corresponding color.

Figure 7.7 shows an optical system that permits both steps of the recording process to be carried out with a minimum of adjustments.

In this arrangement, the plane of the figure corresponds to the vertical plane in the final viewing geometry. A collimated reference beam is used to record the primary hologram (H_1), so that it is necessary only to turn H_1 through 180° about an axis normal to the plane of the figure and replace it in the holder, for an undistorted real image to be projected into the space in front of H_1. Vertical parallax is eliminated by a slit a few millimeters wide placed over H_1 with its long dimension normal to the plane of the figure. This orientation of the slit corresponds to the horizontal in the final viewing geometry. A convergent reference beam is used to record the final hologram (H_2), which, after processing, is reversed for viewing. When H_2 is illuminated with a divergent beam from a point source of white light, an orthoscopic image of the object is formed, and a dispersed real image of the slit is projected into the viewing space.

7.4.1 Image blur

The image reconstructed by a rainbow hologram is free from speckle, because it is illuminated with incoherent light, but not from blur [Wyant, 1977].

One cause of image blur is the finite wavelength spread in the image. If we consider a rainbow hologram made with the optical system shown schematically in fig. 7.8, the wavelength spread observed when the rainbow hologram is illuminated with white light is

$$\Delta\lambda = \left(\frac{\lambda}{\sin\theta}\right)\left(\frac{b+a}{D}\right), \tag{7.1}$$

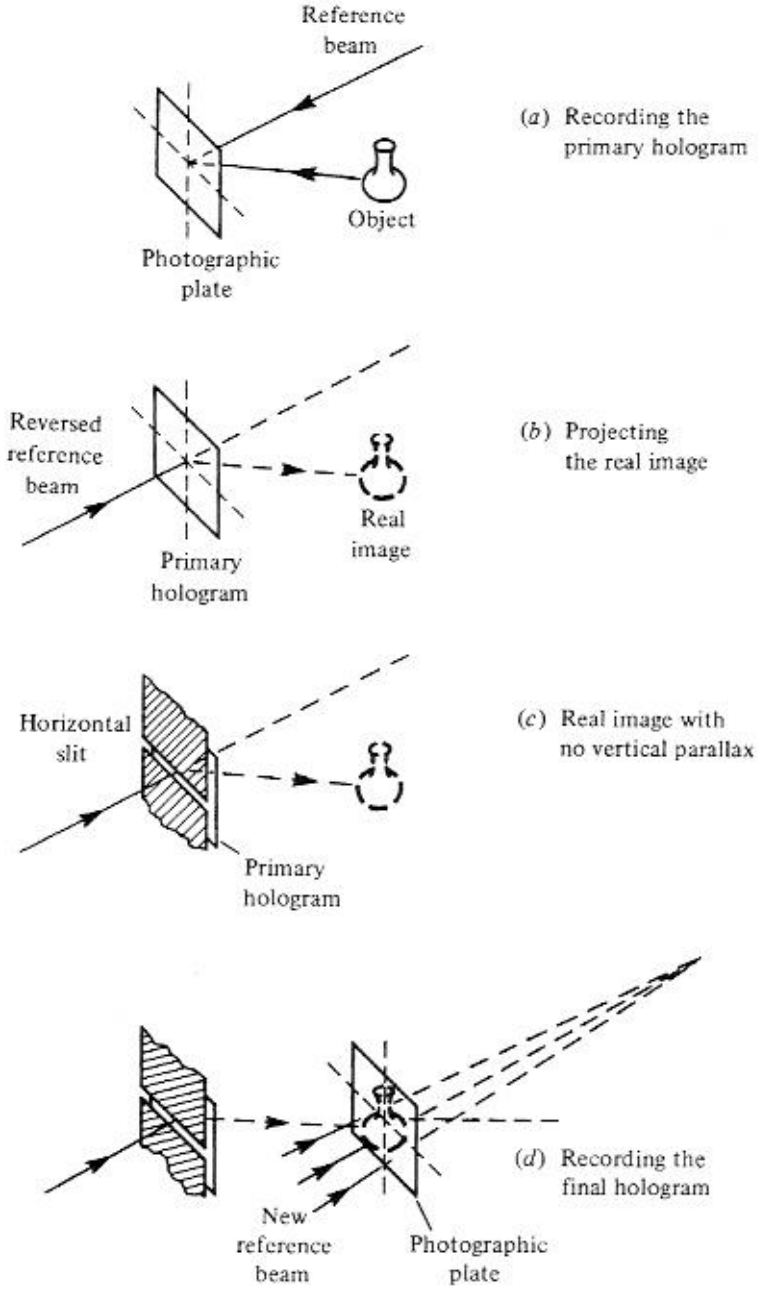


Fig. 7.5. Steps involved in the production of a rainbow hologram.

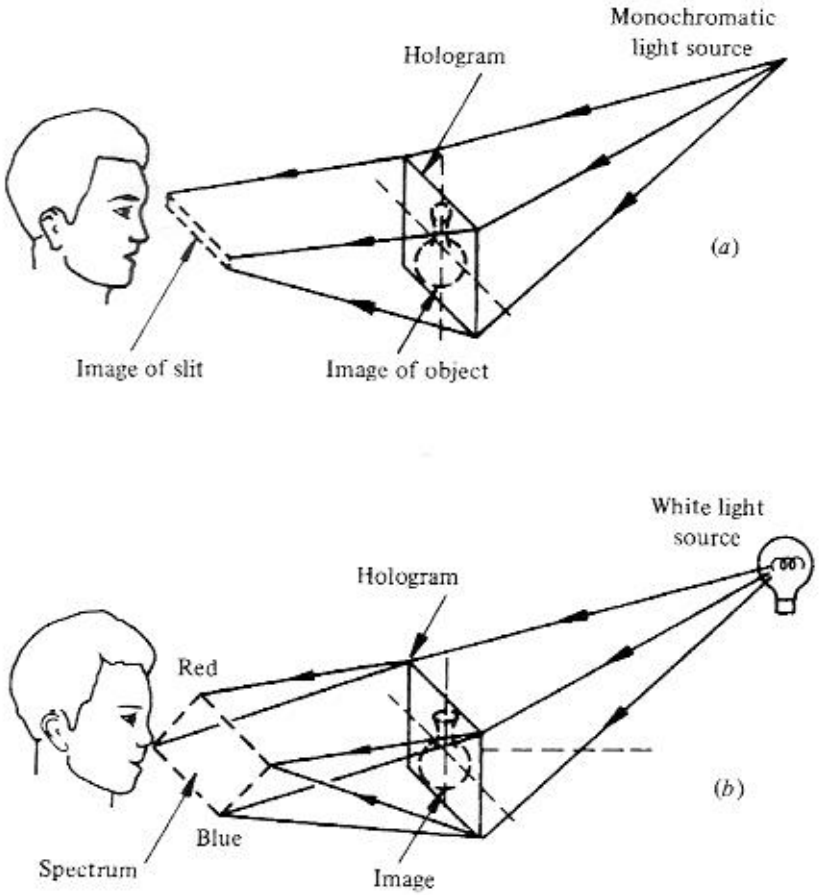


Fig. 7.6. Image reconstruction by a rainbow hologram (*a*) with monochromatic light, and (*b*) with white light.

where λ is the mean wavelength of the reconstructed image, θ is the angle made by the reference beam with the axis, b is the width of the slit, a is the diameter of the pupil of the eye and D is the distance from the primary hologram to the final rainbow hologram.

The image blur due to this wavelength spread is then

$$\Delta y_{\Delta\lambda} = z_0 \left(\frac{\Delta\lambda}{\lambda} \right) \sin \theta,$$

$$= z_0 \left(\frac{a+b}{D} \right), \tag{7.2}$$

where z_0 is the distance of the image from the hologram.

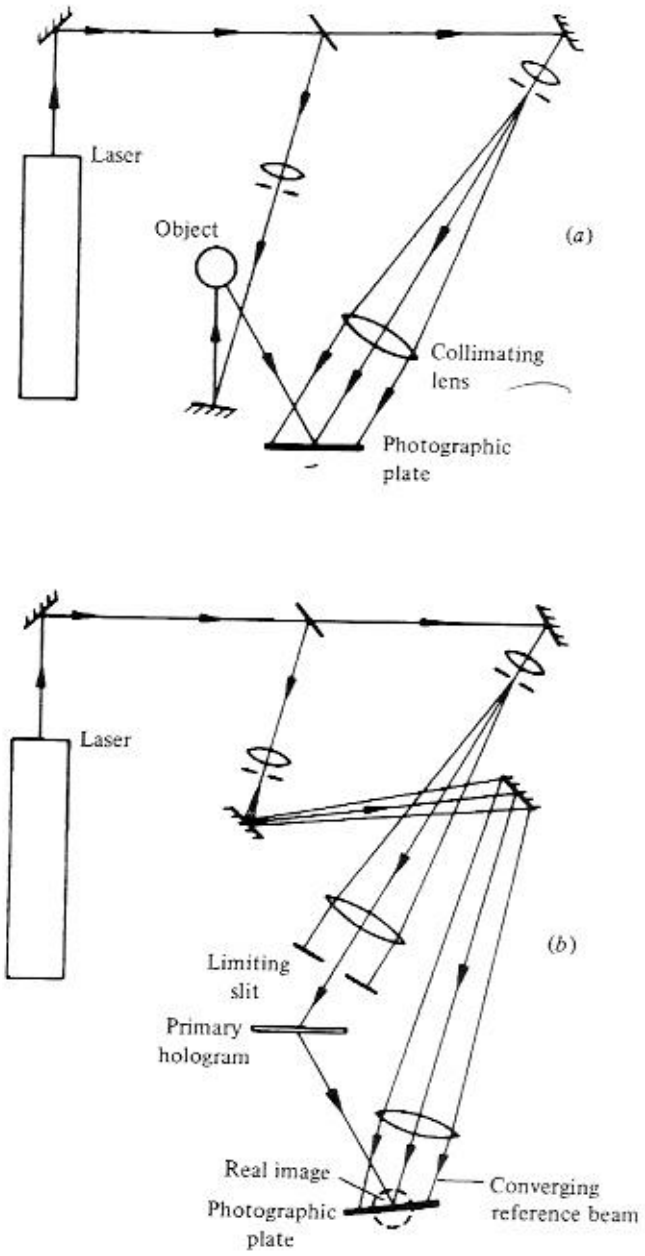


Fig. 7.7. Optical system used to record a rainbow hologram [Hariharan, Steel & Hegedus, 1977].

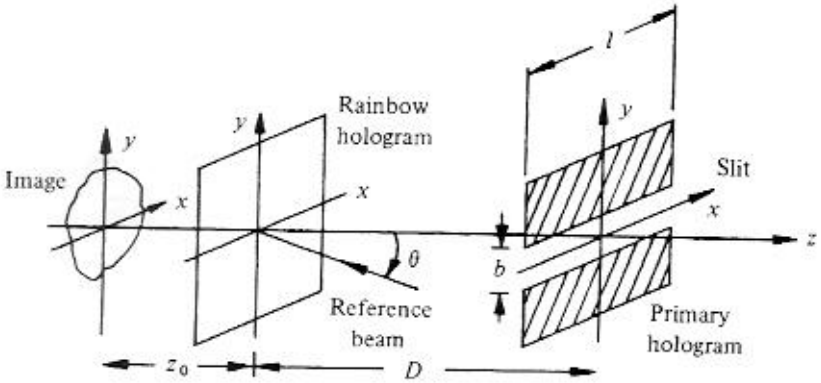


Fig. 7.8. Analysis of image blur in a rainbow hologram [Wyant, 1977].

Another cause of image blur is the finite size of the source used to illuminate the hologram. If the source has an angular spread ψ_s, as viewed from the hologram, the resultant image blur is

$$\Delta y_s = z_0 \psi_s. \tag{7.3}$$

For the blur due to source size not to exceed the blur due to the wavelength spread, the size of the source must satisfy the condition

$$\psi_s \leq \left(\frac{a+b}{D}\right). \tag{7.4}$$

A final cause of image blur is diffraction at the slit; this can be neglected unless the width of the slit is very small.

7.5 Holographic stereograms

It is also possible to synthesize a hologram that reconstructs an acceptable three-dimensional image from a series of two-dimensional views of a subject from different angles [McCrickerd & George, 1968]. To produce such a holographic stereogram, a series of photographs of the subject is taken from equally spaced positions along a horizontal line. Alternatively, the subject is placed on a slowly rotating turntable, and a movie camera is used to make a record of a 120° or 360° rotation. Typically, three movie frames are recorded for each degree of rotation [Benton, 1975].

The optical system used to produce a white-light holographic stereogram from such a movie film is shown schematically in fig. 7.9. Each frame is imaged in the vertical plane on the film used to record the hologram. However, in the horizontal plane, the cylindrical lens brings the image to a line focus. A contiguous

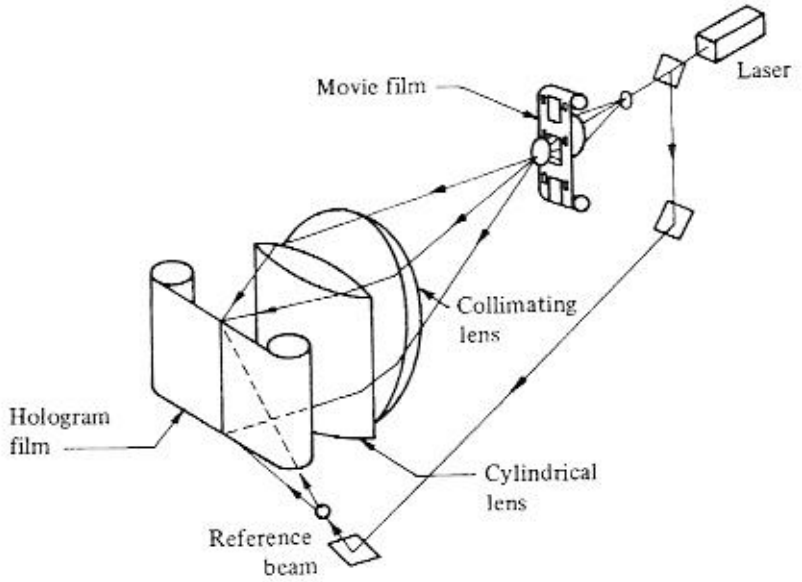


Fig. 7.9. Optical system used to produce a white-light holographic stereogram [Huff & Fusek, 1980].

sequence of vertical strip holograms is recorded of successive movie frames, covering the full range of views of the original subject, using a reference beam incident at an appropriate angle, either from above or from below. With a rotating subject, the processed film is formed into a cylinder for viewing.

When the holographic stereogram is illuminated with white light, the viewer sees a monochromatic three-dimensional image. This image changes color, as with a rainbow hologram, when the observer moves his head up or down. The image lacks vertical parallax, but it exhibits horizontal parallax over the range of angles covered by the original photographs.

Because of the difficulties involved in scaling up the system shown in fig. 7.9, large white-light holographic stereograms (up to 2 m × 1 m) are made by a two-step process [Newswanger & Outwater, 1985].

The obvious advantage of this technique, over recording a hologram directly, is that white light can be used to illuminate the subject in the first stage, so that holographic stereograms can be made of living subjects as well as large scenes. Some subject movement can also be displayed without destroying the stereoscopic effect.

7.6 Holographic movies

Moving three-dimensional images can be produced by presenting a sequence of holograms. With Fourier transform holograms (see Section 1.4),

the reconstructed image can be viewed through the film by a single observer as the film is moved continuously. This limitation is acceptable for technical studies [Heflinger, Stewart & Booth, 1978; Smigielski, Fagot & Albe, 1985] but is a serious drawback for entertainment.

This problem was overcome [Komar, 1977; Komar and Serov, 1989] by using a lens with a diameter of 200 mm to record a series of image holograms of the scene on 70 mm film. The reconstructed image was projected with an identical lens on to a special holographic screen, equivalent to several superimposed concave mirrors. Each of these holographic mirrors then formed a real image of the projection lens in front of a spectator so that, when he looked through this pupil, he saw a full-size three-dimensional image.

Two artistic movies using holographic stereograms have also been produced by Alexander [Lucie-Smith, 1992].

References

Benton, S. A. (1975). Holographic displays – a review. *Optical Engineering*, **14**, 402–7.

Benton, S. A. (1977). White light transmission/reflection holographic imaging. In *Applications of Holography & Optical Data Processing*, ed. E. Marom, A. A. Friesem & E. Wiener-Avnear, pp. 401–9. Oxford: The Pergamon Press.

Denisyuk, Yu. N. (1965). On the reproduction of the optical properties of an object by the wave field of its scattered radiation. II. *Optics & Spectroscopy*, **18**, 152–7.

Hariharan, P. & Chidley, C. M. (1989). Bleached reflection holograms: a study of color shifts due to processing. *Applied Optics*, **28**, 422–4.

Hariharan, P., Steel, W. H. & Hegedus, Z. S. (1977). Multicolor holographic imaging with a white light source. *Optics Letters*, **1**, 8–9.

Heflinger, L. O., Stewart, G. L. & Booth, C. R. (1978). Holographic motion pictures of microscopic plankton. *Applied Optics*, **17**, 951–4.

Huff, L. & Fusek, R. L. (1980). Color holographic stereograms. *Optical Engineering*, **19**, 691–5.

Jeong, T. H. (1967). Cylindrical holography and some proposed applications. *Journal of the Optical Society of America*, **57**, 1396–8.

Koch, G. J. & Petros, M. (1998). A simple overhead Denisyuk configuration for making reflection holograms with a diode laser. *American Journal of Physics*, **66**, 933–4.

Komar, V. G. (1977). Progress on the holographic movie process in the USSR. In *Three-Dimensional Imaging*, Proceedings of the SPIE, vol. 120, ed. S. A. Benton, pp. 127–4. Redondo Beach: SPIE.

Komar, V. G. & Serov, O. B. (1989). Works on the holographic cinematography in the USSR. In *Holography '89*, Proceedings of the SPIE, vol. 1183, eds. Yu. N. Denisyuk & T. H. Jeong, pp. 170–82. Bellingham: SPIE.

Lucie-Smith, E. (1992). *Alexander*. London: Art Books International.

McCrickerd, J. T. & George, N. (1968). Holographic stereogram from sequential component photographs. *Applied Physics Letters*, **12**, 10–2.

Neumann, D. B. & Rose, H. W. (1967). Improvement of recorded holographic fringes by feedback control. *Applied Optics*, **6**, 1097–104.
Newswanger, C. & Outwater, C. (1985). Large format holographic stereograms and their applications. In *Applications of Holography*, Proceedings of the SPIE, vol. 523, ed. L. Huff, pp. 26–32. Bellingham: SPIE.
Smigielski, P., Fagot, H. & Albe, F. (1985). Progress in holographic cinematography. In *Progress in Holographic Applications*, Proceedings of the SPIE, vol. 600, ed. J. Ebbeni, pp. 186–93. Bellingham: SPIE.
Wyant, J. C. (1977). Image blur for rainbow holograms. *Optics Letters*, **1**, 130–2.

Problems

7.1 A rainbow hologram is recorded with an optical system similar to that shown in fig. 7.8, in which $\theta = 45°$, $D = 300$ mm and $b = 3$ mm. If the diameter of the pupil of the observer's eye is 3 mm and the image is reconstructed at a distance $z_0 = 50$ mm from the hologram, what is the image blur due to the wavelength spread? What is the minimum distance from the hologram at which a light source with a diameter of 10 mm should be placed for the image blur due to source size not to exceed that due to the wavelength spread?

From (7.2), the image blur due to the wavelength spread is

$$\Delta y_{\Delta\lambda} = z_0\left(\frac{a+b}{D}\right)$$

$$= 50\left(\frac{3+3}{300}\right) = 1 \text{ mm}. \tag{7.5}$$

For the image blur due to source size not to exceed that due to the wavelength spread, the angular extent of the source, viewed from the hologram, should satisfy (7.4). We have

$$\psi_s \leq \frac{a+b}{D}$$

$$\leq \frac{3+3}{300} \leq 0.02 \text{ radian}. \tag{7.6}$$

A 10 mm diameter source should be placed at a minimum distance of 500 mm from the hologram.

8

Multicolor images

In principle, a multicolor image can be produced by a hologram recorded with three suitably chosen wavelengths, when it is illuminated once again with these wavelengths. However, a problem is that each hologram diffracts, in addition to the wavelength used to record it, the other two wavelengths as well. The cross-talk images produced in this fashion overlap with, and degrade, the desired multicolored image. This problem has been overcome, and several methods are now available to produce multicolor images [Hariharan, 1983].

8.1 Multicolor reflection holograms

The first technique employed to eliminate cross-talk made use of the high wavelength selectivity of volume reflection holograms. If such a hologram is recorded with three wavelengths, one set of fringe planes is produced for each wavelength. When the hologram is illuminated with white light, each set of fringe planes diffracts a narrow band of wavelengths centered on the original wavelength used to record it, giving a multicolor image free from cross-talk [Upatnieks, Marks & Federowicz, 1966].

Higher diffraction efficiency can be obtained by superimposing three bleached volume reflection holograms recorded on two plates, one with optimum characteristics for the red, and the other with optimum characteristics for the green and blue. Brighter images can also be obtained if the final holograms are produced using real images of the object projected by primary holograms whose aperture is limited by a suitably shaped stop [Hariharan, 1980*a*].

8.2 Multicolor rainbow holograms

Another method of producing multicolor images makes use of superimposed rainbow holograms [Hariharan, Steel & Hegedus, 1977]. In this technique, three primary holograms are made with red, green and blue laser light. These

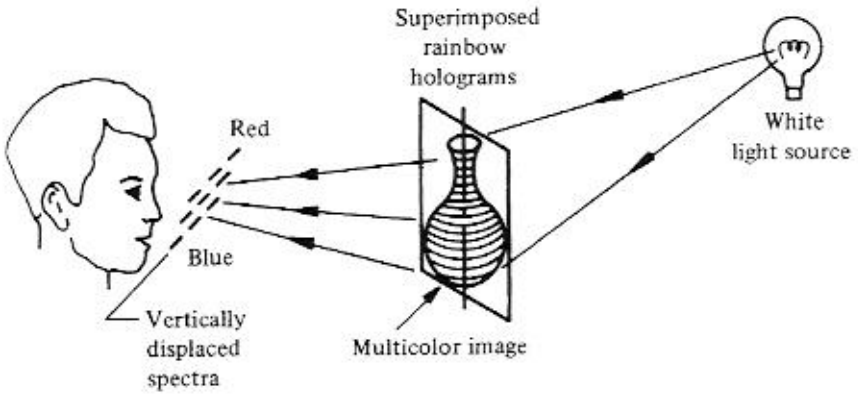


Fig. 8.1. Reconstruction of a multicolor image by superimposed rainbow holograms [Hariharan, 1983].

primary holograms are then used with the same laser sources to make a set of three rainbow holograms, which are then superimposed. When this multiplexed hologram is illuminated with a white light source, it reconstructs three images of the object. In addition, as shown in fig. 8.1, three dispersed images of the slit are formed in the viewing space. Since the corresponding three spectra are displaced by appropriate amounts with respect to each other, an observer, viewing the hologram from the original position of the slit, sees three superimposed images of the object reconstructed in the colors with which the primary holograms were made.

Rainbow holograms can be used very effectively in multicolor displays, since the reconstructed images are very bright and exhibit high color saturation, and are also free from cross-talk.

8.3 Light sources

The most commonly used lasers for color holography are the He–Ne laser which provides an output in the red ($\lambda = 633$ nm), and the Ar^+ laser which provides outputs in the green ($\lambda = 514$ nm) and the blue ($\lambda = 488$ nm). The range of wavelengths that can be reconstructed with these three wavelengths as primaries can be determined by means of the C.I.E. chromaticity diagram. As shown in fig. 8.2, points representing monochromatic light of different wavelengths are located on a horseshoe-shaped curve known as the spectrum locus; all other colors lie within this boundary. New colors obtained by mixing light of two wavelengths, such as 633 nm and 514 nm, lie on the straight line AB joining these primaries. If light with a wavelength of 488 nm is also used, any color within the triangle ABC can be obtained.

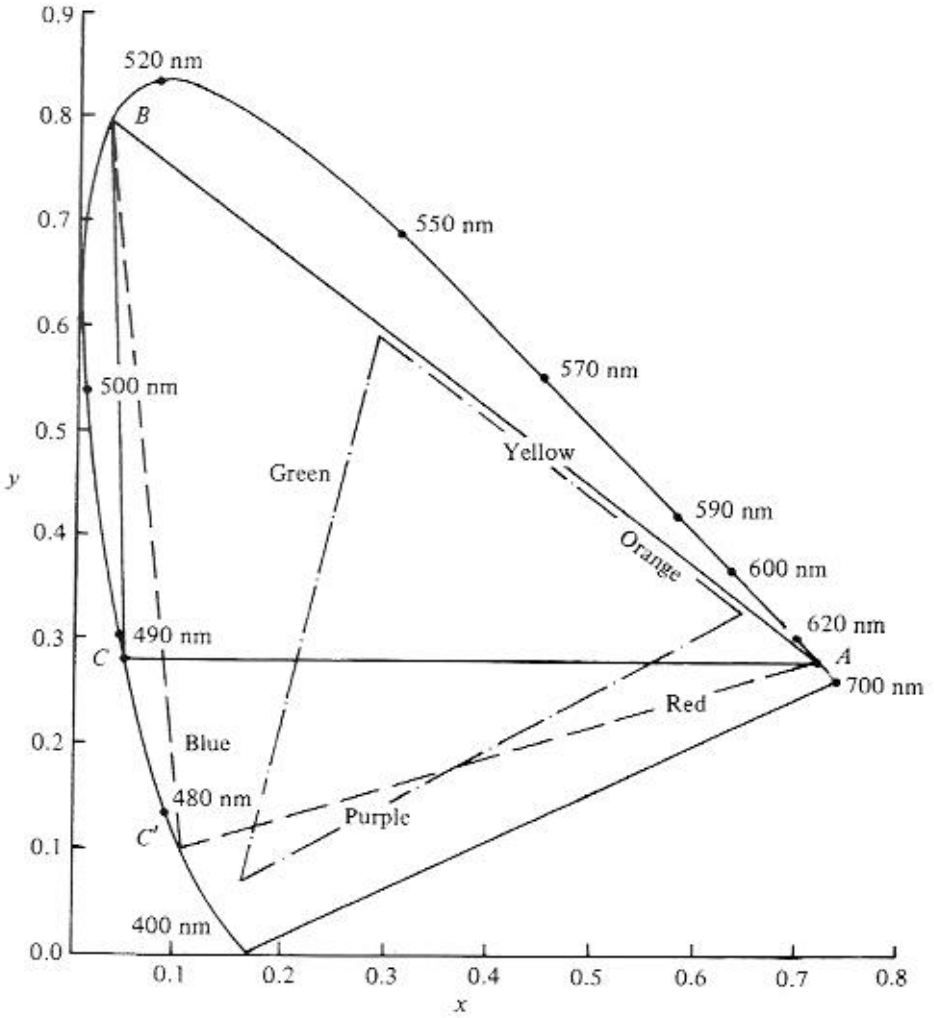


Fig. 8.2. C.I.E. chromaticity diagram. The triangle ABC shows the range of hues that can be produced by a hologram illuminated with primary wavelengths of 633 nm, 514 nm and 488 nm, while the broken lines show the extended range obtained with a blue primary at 477 nm. The chain lines enclose the range of hues that can be reproduced by a typical color-television display [Hariharan, 1983].

A wider range of colors can be obtained, at some sacrifice of power, by using the blue output from an Ar^+ laser at a wavelength of 477 nm, or a He Cd laser at 422 nm.

8.4 Pseudocolor images

The color information in a hologram is encoded in the spatial frequencies of the carrier fringes. It is therefore possible to generate different carrier fringe

frequencies and, consequently, images of different colors, even with a single laser wavelength, by different means.

8.4.1 Pseudocolor rainbow holograms

One way to produce a multicolor image is to record three superimposed rainbow holograms with different reference beam angles [Tamura, 1978]. Alternatively the position of the limiting slit can be changed between exposures.

A problem is that the images reconstructed in a different color from that used to record the hologram are displaced with respect to it. The displacement in the vertical plane is

$$\Delta y = z_0 \left(\frac{2\Delta\lambda}{\lambda} \right) \tan^3 \theta, \tag{8.1}$$

while the longitudinal displacement is

$$\Delta z = z_0 \left(\frac{2\Delta\lambda}{\lambda} \right), \tag{8.2}$$

where z_0 is the distance of the image from the hologram, θ is the interbeam angle in the recording system (see fig. 7.8), λ is the recording wavelength and $(\lambda + \Delta\lambda)$ is the wavelength at which the image is reconstructed. These displacements can be tolerated if the image is formed close to the hologram.

8.4.2 Pseudocolor reflection holograms

With volume reflection holograms, the color of the reconstructed image is affected by changes in the thickness of the recording medium, and these changes can be controlled and used to produce pseudocolor images even with a He–Ne laser [Hariharan, 1980*b*].

The red component hologram is recorded first, on a plate exposed with the emulsion side towards the reference beam. This plate is processed using a rehalogenating bleach to minimize emulsion shrinkage. The green and blue images are then recorded on another plate exposed with the emulsion side towards the object beam. After the green component is exposed, the emulsion is soaked in a 3 percent solution of triethanolamine, to swell it, and dried in darkness. The blue component is then exposed. Normal bleach processing eliminates the triethanolamine and produces the usual shrinkage. As a result, the first exposure yields a green image, while the second produces a blue image.

After drying, the plates are cemented together, with the emulsions in contact, and viewed with the hologram reconstructing the blue and green images in front.

Very high quality pseudocolor images can be produced from contact copies (see Section 9.1) of three master holograms made on DuPont photopolymer Holographic Recording Film with an Ar^+ laser (476 nm). Two of the holograms are swollen during processing by baking in contact with DuPont Color Tuning Film, so that they reconstruct green and red images, before the three films are sandwiched together [Hubel & Klug, 1992].

References

Hariharan, P. (1980*a*). Improved techniques for multicolour reflection holograms. *Journal of Optics (Paris)*, **11**, 53–5.

Hariharan, P. (1980*b*). Pseudocolour images with volume reflection holograms. *Optics Communications*, **35**, 42–4.

Hariharan, P. (1983). Colour Holography. In *Progress in Optics*, vol. 20, ed. E. Wolf, pp. 265–324. Amsterdam: North-Holland.

Hariharan, P., Steel, W. H. & Hegedus, Z. S. (1977). Multicolor holographic imaging with a white light source. *Optics Letters*, **1**, 8–9.

Hubel, P. & Klug, M. A. (1992). Color holography using multiple layers of DuPont photopolymer. In *Practical Holography V*, Proceedings of the SPIE, vol. 1667, ed. S. A. Benton, pp. 215–24. Bellingham: SPIE.

Tamura, P. N. (1978). Pseudocolor encoding of holographic images using a single wavelength. *Applied Optics*, **17**, 2532–6.

Upatnieks, J., Marks, J. & Federowicz, R. (1966). Color holograms for white light reconstruction. *Applied Physics Letters*, **8**, 286–7.

Problems

8.1. What is the theoretical improvement in diffraction efficiency possible in a multicolor rainbow hologram by recording the three component holograms on three separate plates instead of on a single plate?

If all the three holograms are recorded on a single photographic plate, the available dynamic range is divided between the three holograms (see Section 3.6). The diffraction efficiency of each component hologram is then only $(1/3)^2 = 1/9$ of that for a single hologram recorded on the same plate. The diffraction efficiency should therefore improve, theoretically, by a factor of 9 (the actual gain is about half of this, because of transmission losses).

8.2. A pseudocolor rainbow hologram is recorded with an Ar^+ laser ($\lambda = 514$ nm) with an interbeam angle of 30°. While recording the red and blue components, the slit is shifted so that these images are reconstructed at wavelengths of 630 nm and 450 nm respectively. If the lateral displacements of the red and blue images are not to exceed 1 mm, what is the maximum permissible depth of these images?

We consider, first, the displacement of the red image, since the wavelength shift for it is greater ($\Delta\lambda = 116$ nm). From (8.1), we have

$$\begin{aligned}\Delta y &= z_0\left(\frac{2\Delta\lambda}{\lambda}\right)\tan^3\theta,\\ &= z_0\left(\frac{2\times 116}{514}\right)\left(\frac{1}{\sqrt{3}}\right)^3\\ &= 0.124\times z_0.\end{aligned} \tag{8.3}$$

Accordingly, for the lateral displacement of the edges of the red image to be less than 1 mm, they must lie within ± 8 mm of the hologram.

A similar calculation shows that the edges of the blue image (for which $\Delta\lambda = 64$ nm) must lie within ± 14 mm of the hologram.

Appendix A

Interference and coherence

A.1 Interference

The time-varying electric field E at any point due to a linearly polarized monochromatic light wave propagating in a vacuum in the z direction can be represented by the relation

$$E = a\cos[2\pi\nu(t - z/c)], \tag{A.1}$$

where a is the amplitude, ν the frequency and c the speed of propagation of the light wave. Equation (A.1) can be written in the form

$$E = \mathrm{Re}\{a\exp[\mathrm{i}2\pi\nu(t - z/c)]\},$$

$$= \mathrm{Re}\{a\exp(-\mathrm{i}\phi)\exp(\mathrm{i}2\pi\nu t)\}, \tag{A.2}$$

where Re{. . .} represents the real part of the expression within the braces, $\mathrm{i} = (-1)^{1/2}$, and $\phi = 2\pi\nu z/c = 2\pi z/\lambda$.

If we assume that all operations on E are linear, we can use the complex representation

$$E = a\exp(-\mathrm{i}\phi)\exp(\mathrm{i}2\pi\nu t),$$

$$= A\exp(\mathrm{i}2\pi\nu t), \tag{A.3}$$

where $A = a\exp(-\mathrm{i}\phi)$ is known as the complex amplitude. Multiplication of the complex amplitude by its complex conjugate yields the optical intensity

$$I = AA^* = |A|^2. \tag{A.4}$$

Since the complex amplitude at any point due to a number of waves of the same frequency is the sum of the complex amplitudes of the individual waves

$$A = A_1 + A_2 + \cdots, \tag{A.5}$$

the intensity due to the interference of two waves is

$$I = |A_1 + A_2|^2,$$

$$= |A_1|^2 + |A_2|^2 + A_1A_2^* + A_1^*A_2,$$

$$= I_1 + I_2 + 2(I_1I_2)^{1/2}\cos(\phi_1 - \phi_2). \tag{A.6}$$

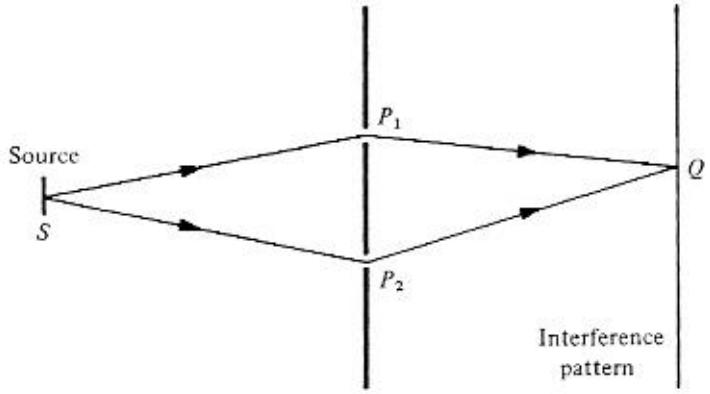


Fig. A.1. Evaluation of the degree of coherence.

The visibility of the interference fringes is defined as

$$\mathbf{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2(I_1 I_2)^{1/2}}{I_1 + I_2}. \tag{A.7}$$

Equations (A.6) and (A.7) assume that the electric vectors of the two waves are parallel. If the electric vectors make an angle θ with each other, the visibility of the interference fringes is only

$$\mathbf{V}_\theta = \mathbf{V} \cos \theta, \tag{A.8}$$

which drops to zero when $\theta = \pi/2$.

A.2 Coherence

We have assumed in the previous section that the light waves are derived from a single point source emitting an infinitely long, monochromatic wave train, in which case we say that the fields due to the two light waves are perfectly coherent. In reality, all wave fields are only partially coherent.

We can represent the time-varying electric field at any point due to a linearly polarized, quasi-monochromatic light wave from a source of finite size by the analytic signal [Born & Wolf, 1999]

$$V(t) = \int_0^\infty a(\nu) \exp[-i\phi(\nu) \exp(i2\pi\nu t)] d\nu, \tag{A.9}$$

where $a(\nu)$ is the amplitude and $\phi(\nu)$ is the phase of a component with frequency ν. To evaluate the degree of coherence of the fields at two points illuminated by such a source, we consider the optical system shown in fig. A.1. In this arrangement, if $V_1(t)$ and $V_2(t)$ are the analytic signals corresponding to the electric fields at P_1 and P_2, the complex degree of coherence $\gamma_{12}(\tau)$ of the fields, for a time delay τ (the difference in the transit times from S to P_1 and P_2), is defined as the normalized correlation of $V_1(t)$ and $V_2(t)$ (see Appendix B)

$$\gamma_{12}(\tau) = \frac{\langle V_1(t+\tau) V_2^*(t) \rangle}{[\langle V_1(t) V_1^*(t) \rangle \langle V_2(t) V_2^*(t) \rangle]^{1/2}}. \tag{A.10}$$

The physical significance of (A.10) can be understood if the light waves are allowed to emerge through pinholes at P_1 and P_2, so that they form an interference pattern on a screen. P_1 and P_2 can then be considered as two secondary sources, so that, from (A.6), the intensity at Q is

$$I = I_1 + I_2 + \langle V_1(t+\tau)V_2^*(t) + V_1^*(t+\tau)V_2(t)\rangle,$$

$$= I_1 + I_2 + 2\mathrm{Re}[\langle V_1(t+\tau)V_2^*(t)\rangle], \tag{A.11}$$

where I_1 and I_2 are the intensities at Q when P_1 and P_2 act separately, and τ is the difference in the transit times for the paths P_1Q and P_2Q. We can rewrite this relation, from (A.10), as

$$I = I_1 + I_2 + 2(I_1I_2)^{1/2}\,\mathrm{Re}[\gamma_{12}(\tau)]$$

$$= I_1 + I_2 + 2(I_1I_2)^{1/2}|\gamma_{12}(\tau)|\cos\phi_{12}(\tau), \tag{A.12}$$

where $\phi_{12}(\tau)$ is the phase of $\gamma_{12}(\tau)$.

Interference fringes are produced by the variations in $\cos\phi_{12}(\tau)$ across the screen. If $I_1 = I_2$, the visibility of these fringes is, from (A.7),

$$\mathbf{V} = |\gamma_{12}(\tau)|. \tag{A.13}$$

The coherence of the field produced by any light source can be studied from two aspects.

A.2.1 Spatial coherence

When the difference in the optical paths from the source to P_1 and P_2 is negligibly small, so that $\tau \approx 0$, effects due to the spectral bandwidth of the source can be neglected, and we are essentially concerned with what is termed the spatial coherence of the field. A special case of interest is when the dimensions of the source and the separation of P_1 and P_2 are extremely small compared to the distances between the source and P_1 and P_2. In this case, the complex degree of coherence is given by the normalized Fourier transform (see Appendix B) of the intensity distribution over the source.

A.2.2 Temporal coherence

If the source is effectively a point source, but radiates over a range of wavelengths, we are concerned with the temporal coherence of the field. In this case, the complex degree of coherence depends only on τ, the difference in the transit times. Equation (A.10) can then be transformed and written as

$$\gamma(\tau) = \frac{\mathcal{F}\{S(\nu)\}}{\int_{-\infty}^{\infty} S(\nu)\mathrm{d}\nu}, \tag{A.14}$$

where $S(\nu)$ is the frequency spectrum of the radiation. In this case, it follows, from (A.12) and (A.13), that the degree of temporal coherence can be obtained from the visibility of the interference fringes as the difference in the lengths of the optical paths from the source is varied.

This argument leads us to the concepts of the coherence time and the coherence length of the radiation. It can be shown that, with radiation having a bandwidth $\Delta\nu$, the visibility of the interference fringes drops to zero for a difference in the transit times

$$\Delta\tau \approx 1/\Delta\nu. \tag{A.15}$$

This time $\Delta\tau$ is called the coherence time of the radiation; the corresponding value of the coherence length is

$$\Delta l \approx c\Delta\tau \approx \lambda_0^2/\Delta\lambda, \tag{A.16}$$

where λ_0 is the mean wavelength and $\Delta\lambda$ the bandwidth of the radiation. To obtain interference fringes with good visibility, the optical path difference must be small compared to the coherence length of the radiation.

References

Born, M. & Wolf, E. (1999). *Principles of Optics*. Cambridge: Cambridge University Press.