PHYSICS 131, WINTER 1985

excerpt from "OPTICS", Eugene Hecht and Alfred Zajac, Adelphi University
4th Printing, 1979  Addison-Wesley Publishing Company
overlapping wave trains change rapidly and randomly in time, washing out the large-scale interference pattern.

A real system of fringes is formed of the scattered waves which converge in front of the screen. The fringes can be viewed by intersecting the interference pattern with a sheet of paper at a convenient location. After forming the real image in space, the rays proceed to diverge and any region of the image can therefore be viewed directly with the eye appropriately focused. In contrast, rays which initially diverge appear to the eye as if they had originated behind the scattering screen and thus form a virtual image.

It seems that as a result of chromatic aberration, normal and farsighted eyes tend to focus red light behind the screen. Contrarily, a nearsighted person observes the real field in front of the screen (regardless of wavelength). Thus if the viewer moves her head to the right, the pattern will move to the right in the first instance (where the focus is beyond the screen) and to the left in the second (focus in front). The pattern will follow the motion of your head if you’re viewing it very close to the surface. The same apparent parallax motion can be seen by looking through a window; outside objects will seem to move with your head, inside ones opposite to it. The brilliant, narrow-bandwidth, spatially coherent laser beam is ideally suited for observing the granular effect although other means are certainly possible. In unfocused sunlight the grains are minute on the surface and multicolored. The effect is easy to observe on a smooth, flat-black material (e.g., poster-painted paper) but you can see it on a fingernail or a worn coin as well.

Although it provides a marvelous demonstration, both esthetically and pedagogically, the granular effect can be a real practical nuisance in coherently illuminated systems. For example, in holographic imagery the speckle pattern corresponds to troublesome background noise.

14.3 HOLOGRAPHY

The technology of photography has been with us for a long time and we’ve all grown accustomed to seeing the three-dimensional world compressed into the flatness of a scrapbook page. The depthless TV pitchman who smiles out of a myriad of phosphorescent flashes, although inescapably there, seems no more palpable than a postcard image of the Eiffel Tower. Both share the severe limitation of being simply irradiance mappings. In other words, when the image of a scene is ordinarily reproduced, by whatever traditional means, what we ultimately see is not an accurate reproduction of the light field which once inundated the object, but rather a point-by-point record of just the square of the field’s amplitude. The light reflecting off a photograph carries with it information about the irradiance but nothing about the phase of the wave which once emanated from the object. Indeed, if both the amplitude and phase of the original wave could be reconstructed somehow, the resulting light field (presuming the frequencies are the same) would be indistinguishable from the original. This means that you would then see (and could photograph) the reformed image in perfect three-dimensionality, exactly as if the object were there before you, actually generating the wave.

14.3.1 Methods

Dennis Gabor had been thinking along these lines for a number of years prior to 1947 when he began conducting his now famous experiments in holography at the Research Laboratory of the British Thomson–Houston Company. His original set-up, depicted in Fig. 14.35, was a two-step lensless imaging process in which he first photographically recorded an interference pattern, generated by the interaction of scattered quasimonochromatic light from an object and a coherent reference wave. The resulting pattern was something he called a hologram after the Greek word holos meaning whole. The second step in the procedure was the reconstruction of the optical field or image and this was
Fig. 14.35 Holographic (in-line) recording and reconstruction of an image.
done via the diffraction of a coherent beam by a transparency, which was the developed hologram. In a way quite reminiscent of Zernike's phase-contrast technique (Section 14.1.4), the hologram was formed when the unscattered background or reference wave interfered with the diffracted wave from the small semitransparent object, \( S \) — which was, in those early days, often a piece of microfilm. The key point is that the interference pattern or hologram contains, by way of the fringe configuration, information corresponding to both the amplitude and phase of the wave scattered by the object.

Admittedly, it's not at all obvious that by now shining a plane wave through the processed hologram one could reconstruct an image of the original object. Suffice it to say for the moment that if the object were very small the scattered wave would be nearly spherical and the interference pattern a series of concentric rings (centered about an axis through the object and normal to the plane wave). Except for the fact that the circular fringes would vary gradually in irradiance from one to the next, the resulting flux-density distribution would correspond to a conventional Fresnel zone plate (Section 10.3.5). Recall that a zone plate functions something like a lens in that it diffracts collimated light into a beam converging to a real focal point, \( P_r \). In addition, it produces a diverging wave which appears to come from the point \( P_r \) and constitutes a virtual image. Thus we can imagine, albeit rather simplistically, that each point on an extended object generates its own zone plate displaced from the others and that the ensemble of all such partially overlapping zone plates forms the hologram. During the reconstruction step, each constituent zone plate forms both a real and virtual image of a single object point and in this way, point by point, the hologram regenerates the original light field. When the reconstructing beam has the same wavelength as the initial recording beam (which need not necessarily be the case, and quite often isn't), the virtual image is undistorted and appears at the location formerly occupied by the object. Thus it is the virtual image field which actually corresponds to the original object field. As such, the virtual image is sometimes spoken of as the true image while the other is the real or, perhaps more fittingly, the conjugate image.

Gabor's research, which won him the 1971 Nobel Prize for physics, had as its motivation an improvement in electron microscopy. His work initially generated some interest, but all in all it remained in a state of quasi-unnoticed oblivion for about fifteen years. In the early nineteen-sixties, there was a resurgence of interest in Gabor's wavefront reconstruction process and, in particular, in its relation to certain radar problems. Soon, aided by an abundance of the new coherent laser light and extended by a number of technological advances, holography became a subject of widespread research and tremendous promise. This rebirth had its origin in the Radar Laboratory of the University of Michigan with the work of Emmett N. Leith and Juris Uppnienks. Among other things, they introduced an improved arrangement for generating holograms which is illustrated in Fig. 14.35. Unlike Gabor's in-line configuration where the conjugate image was inconveniently located in front of the true image, the two were now satisfactorily separated off axis as shown in the diagram. Once again, the hologram is an interference pattern arising from a coherent reference wave and a wave scattered from the object (this type is sometimes referred to as a side band Fresnel hologram).

The process can be treated analytically as follows; suppose that the \( xy \)-plane is the plane of the hologram, \( \Sigma_H \). Then

\[
E_g(x, y) = E_{0g} \cos \left[ 2\pi f + \phi(x, y) \right] \tag{14.8}
\]

describes the planar background or reference wave at \( \Sigma_H \), overlooking considerations of polarization. Its amplitude, \( E_{0g} \), is constant while the phase is a function of position. This just means that the reference wavefront is tilted in some known manner with respect to \( \Sigma_H \). For example, if the wave were oriented such that it could be brought into coincidence with \( \Sigma_H \) by a single rotation through an angle of \( \theta \) about \( y \), the phase at any point on the hologram plane would depend on its value of \( x \). Thus \( \phi \) would have the form

\[
\phi = \frac{2\pi}{\lambda} x \sin \theta = k x \sin \theta
\]

being, in that particular case, independent of \( y \) and varying linearly with \( x \). For the sake of simplicity we'll just write it quite generally as \( \phi(x, y) \) and keep in mind that it's a simple known function. The wave scattered from the object can, in turn, be expressed as

\[
E_o(x, y) = E_{0o}(x, y) \cos \left[ 2\pi f + \phi_o(x, y) \right], \tag{14.9}
\]

where both the amplitude and phase are now complicated functions of position corresponding to an irregular wavefront. From the communications-theoretic point of view, this is an amplitude- and phase-modulated carrier wave bearing

---

all of the available information about the object. Note that this information is coded in spatial rather than temporal variations of the wave. The two disturbances $E_a$ and $E_0$ superimpose and interfere to form an irradiance distribution which is recorded by the photographic emulsion. The resulting irradiance, except for a multiplicative constant, is $I(x, y) = \langle (E_a + E_0)^2 \rangle$ which, from Section 9.1, is given by

$$I(x, y) = \frac{E_a^2}{2} + \frac{E_0^2}{2} + E_a E_0 \cos(\phi - \phi_0).$$ (14.10)

Observe that the phase of the object wave determines the location on $\Sigma_H$ of the irradiance maxima and minima. Moreover, the contrast or fringe visibility

$$\gamma^* = (I_{\text{max}} - I_{\text{min}}) / (I_{\text{max}} + I_{\text{min}})$$ [12.1]

across the hologram plane, which is

$$\gamma^* = 2E_a E_0 / (E_a^2 + E_0^2).$$ (14.11)

contains the appropriate information about the object wave's amplitude.

Once more, in the parlance of communications theory, we might observe that the film plate serves as both the storage device and detector or mixer. It produces, over its surface, a distribution of opaque regions corresponding to a modulated spatial waveform. Accordingly, the third or difference frequency term in Eq. (14.10) is both amplitude
and phase modulated by way of the position dependence of $E_0(x, y)$ and $\phi_0(x, y)$.

We’ve shown the configuration utilizing diffusely reflected light from an opaque object but it could equally well be rearranged a bit, as in Fig. 14.37, to get sideband Fresnel holograms from transparent objects. Figure 14.38(b) is an enlarged view of a portion of the fringe pattern which constitutes the hologram for a simple, essentially two-dimensional, semitransparent object. Were the two interfering waves perfectly planar (as in Fig. 14.38(a)), the evident variations in fringe position and irradiance, which represent the information, would be absent, yielding the traditional Young’s pattern (Section 9.3). The sinusoidal transmission-grating configuration [Fig. 14.38(a)] may be thought of as the carrier waveform which is then modulated by the signal. Furthermore, we can imagine that the coherent superposition of countless zone-plate patterns, one arising from each point on a large object, have metamorphosed into the modulated fringes of Fig. 14.38(b). When the amount of modulation is further greatly increased, as it would be for a large three-dimensional diffusely reflecting object, the fringes lose the kind of symmetry still discernible in Fig. 14.38(b) and become considerably more complicated. Incidentally, holograms are often covered with extraneous swirls and concentric ring systems that arise from diffraction by dust and the like on the optical elements.

The amplitude transmittance of the processed hologram can be made proportional to $I(x, y)$. In that case, the final

![Fig. 14.37 A side-band Fresnel holographic set-up for a transparent object.](image)

**Fig. 14.37** A side-band Fresnel holographic set-up for a transparent object.

The emerging wave, $E_r(x, y)$, is proportional to the product $I(x, y)E_r(x, y)$, where $E_r(x, y)$ is the reconstructing wave incident on the hologram. Thus if the reconstructing wave of frequency $v$ is incident obliquely on $\Sigma_h$ as was the background wave, we can write

$$E_r(x, y) = E_0 \cos [2\pi vt + \phi(x, y)].$$  \hspace{1cm} (14.12)

The final wave (except for a multiplicative constant) is the

![Fig. 14.38 Various degrees of modulation of hologram fringes.](image)

**Fig. 14.38** Various degrees of modulation of hologram fringes. [Photo courtesy Emmett N. Leith and Scientific American.]
product of Eqs. (14.10) and (14.12):
\[ E_I(x, y) = \frac{1}{2} E_0 [E_{00}^* + E_{00}^*] \cos(2\pi y + \phi(x, y)) + \frac{1}{2} E_0 E_{00} \cos(2\pi y + 2\phi - \phi_0) + \frac{1}{2} E_0 E_{00} \cos(2\pi y + \phi_0). \]  

Three terms describe the light issuing from the hologram; the first can be rewritten as
\[ \frac{1}{2} (E_{00}^* + E_{00}^*) E_I(x, y). \]
and is an amplitude-modulated version of the reconstructing wave. In effect, each portion of the hologram functions as a diffraction grating and this is the zeroth-order, undeflected, direct beam. Since it contains no information about the phase of the object wave, \( \phi_0 \), it is of little concern here.

The next two or side band waves are the sum and difference terms, respectively. These are the two first-order waves diffracted by the grating-like hologram. The first of these, i.e., the sum term, represents a wave which, except for a multiplicative constant, has the same amplitude as the object wave \( E_{00}(x, y) \). Moreover, its phase contains a \( 2\phi(x, y) \) contribution which, as you recall, arose from tilting the background and reconstructing wavefronts with respect to \( z \). It’s this phase factor which provides the angular separation between the real and virtual images. Furthermore, rather than containing the phase of the object wave, the sum term contains its negative. Thus it’s a wave carrying all of the appropriate information about the object but in a way which is not quite right. Indeed, this is the real image formed in converging light in the space beyond the hologram, i.e., between it and the viewer. The negative phase is manifest in an inside-out image something like the pseudoscopic effect occurring when the elements of a photographic stereo pair are interchanged. Bumps appear as indentations and object points which were in front of and nearer to \( \Sigma_H \) are now imaged nearer to but beyond \( \Sigma_H \). Thus a point on the original subject closest to the observer appears farthest away in the real image. The scene is turned in on itself along one axis in a way that perhaps must be seen to be appreciated. For example, imagine you are looking down the holographic conjugate image of a bowing alley. The “back” rows of pins, even though partially obscured by the “front” rows, are nonetheless imaged closer to the viewer than is the one-pin. Despite this, bear in mind that it’s not as if you were looking at the array from behind. No light from the very backs of the pins was ever recorded—you’re seeing an inside-out front view. As a consequence the conjugate image is usually of limited utility although it can be made to have a normal configuration by forming a second hologram with the real image as the object.

The difference term in Eq. (14.13), except for a multiplicative constant, has precisely the form of the object wave \( E_0(x, y) \). If you were to peer into (not at) the illuminated hologram, as if it were a window looking out onto the scene beyond, you would “see” the object exactly as if it were truly sitting there. You could move your head a bit and look around an item in the foreground in order to see the view it had previously been obstructing. In other words, in addition to complete three-dimensionality, parallax effects are apparent as they are in no other reproducing technique (Fig. 14.39). Imagine that you are viewing the holographic image of a magnifying glass focused on a page of print. As you move your eye with respect to the hologram plane, the words being magnified by the lens (which is itself just an image) actually change just as they would in “real” life with a “real” lens and “real” print. In the case of an extended scene having considerable depth, your eyes would have to refocus as you viewed different regions of it at various distances. In precisely the same way, a camera lens would have to be readjusted if you were photographing those regions of the virtual image.

There are several other extremely interesting features which holograms display. For example, if you were standing close to a window you could obscure all of it with, say, a piece of cardboard except for a tiny area through which you could then peer and still see the objects beyond. The same is true of a hologram since each small fragment of it contains information about the entire object, at least as seen from that vantage point, and can reproduce, albeit with diminishing resolution, the entire image.

The zone-plate interpretation has been applicable to the various holographic schemes we’ve considered thus far and this regardless of whether the diffracted wave was of the near- or far-field variety (i.e., whether we had Fresnel or Fraunhofer holograms, respectively). Indeed it applies generally where the interferogram results from the superpositioning of the scattered spherical wavelets from each object point and a coherent plane or even spherical reference wave (provided the latter’s curvature is different from that of the wavelets). An inherent failing, which these schemes therefore have in common, arises from the fact that the zone plate radii, \( R_m \), vary as \( m^{1/2} \) via Eq. (10.91). Thus the zone fringes are more densely packed farther from the center of each zone lens (i.e. at larger values of \( m \)). This is tantamount to an increasing spatial frequency of bright and dark rings.
which must be recorded by the photographic plate. But since film, no matter how fine grained, is limited in its spatial frequency response, there will be a cut-off beyond which it cannot record data. All of this represents a built-in limitation on resolution. In contrast, if the mean frequency of the fringes could be made constant, the limitations imposed by the photographic medium would be considerably reduced and the resolution correspondingly increased. So long as it could record the average spatial fringe frequency even coarse emulsions, such as Polaroid P/N, could be used without extensive loss of resolution. Figure 14.40 shows an arrangement which accomplishes just this by having the
Fig. 14.40 Lensless Fourier transform holography (a transparent object).
diffracted object wavelets interfere with a spherical reference wave of about the same curvature. The resulting interferogram is known as a Fourier transform hologram (in this specific instance, it's of the high-resolution lensless variety). This scheme is designed to have the reference wave cancel the quadratic (zone-lens type) dependence of the phase with position on $\Sigma_H$. But that will occur precisely only for a planar two-dimensional object. In the case of a three-dimensional object (Fig. 14.41) this only happens over one plane and the resulting hologram is therefore a composite of both types, i.e., a zone lens and Fourier transform. Unlike the other arrangements, both images generated by a Fourier-transform hologram are virtual, in the same plane and oriented as if reflected through the origin (Fig. 14.42).

The grating-like nature of all previous holograms is evident here as well. In fact, if you look through a Fourier-transform hologram at a small white-light source (a flashlight in a dark room works beautifully), you see the two mirror images but they are extremely vague and surrounded by bands of spectral colors. The similarity with white light which has passed through a grating is unmistakable.\footnote{See DeVelis and Reynolds, Theory and Applications of Holography, Stroke, An Introduction to Coherent Optics and Holography, Goodman, Introduction to Fourier Optics, Smith, Principles of Holography or perhaps The Engineering Uses of Holography, edited by E. R. Robertson and J. M. Harvey.}

\begin{figure}
\centering
\includegraphics{fig14_41}
\caption{Lensless Fourier transform holography (an opaque object).}
\end{figure}

\begin{figure}
\centering
\includegraphics{fig14_42}
\end{figure}

14.3.2 Recent Developments

For years holography was an invention in search of application, that notwithstanding certain obvious possibilities like the all too inevitable 3-D billboard. Fortunately, several significant technological developments have in recent times begun what will surely be an ongoing extension of the scope and utility of holography. The early efforts in the field were typified by countless images of toy cars and trains, chesspieces and statuettes—small objects resting on giant blocks of granite. They had to be small because of limited laser power and coherence length; while the ever-present massive granite platform served to isolate the slightest vibrations which might blur the fringes and thereby degrade or obliterate the stored data. A loud sound or gust of air could result in deterioration of the reconstructed image by causing the photo plate, object, or mirrors to shift several millionths of an inch during the exposure, which itself might last of the order of a minute or so. That was the still-life era of holography. But now, using new, more sensitive films and the short duration (~40 ns) high-power light flashes from a single-mode pulsed ruby laser, even portraiture and stop-action holography have become a reality\footnote{L. D. Siebert, Appl. Phys. Letters 11, 326 (1967) and R. G. Zech and L. D. Siebert, Appl. Phys. Letters 13, 417 (1968).} (Fig. 14.43).
Volume Holograms

Nikolayevich Denisyuk of the Soviet Union, in 1962, introduced a scheme for generating holograms which was conceptually similar to the early (1891) color photographic process of Gabriel Lippmann. In brief, the object wave is reflected from the subject and propagates backward, overlapping the incoming coherent background wave. In so doing, the two waves set up a three-dimensional pattern of standing waves. The concomitant spatial distribution of fringes is, in part, recorded by the photoemulsion throughout its entire thickness to form what has become known as a volume hologram. Several variations have since been introduced, but the basic idea are the same; rather than generating a two-dimensional grating-like structure, the volume hologram is a three-dimensional grating. In other words, it's a three-dimensional, modulated, periodic array of phase or amplitude objects which represent the data. It can be recorded in several media, for example, in thick photoemulsions wherein the amplitude objects are grains of deposited silver; in photochromic glass; with halogen crystals like KBr which respond to irradiation via color-center variations; or with a ferroelectric crystal such as lithium niobate which undergoes local alterations in its index of refraction, thus forming what might be called a phase volume hologram. In any event, one is left with a volume array of data, however stored in the medium which, in the reconstruction process, behaves very much like a crystal being irradiated by x-rays. It scatters the incident (reconstructing) wave according to Bragg’s law (Section 10.2.7). This isn’t very surprising since both the scattering centers and \( \lambda \) have simply been scaled up proportionally. One important feature of volume holograms is the interdependence \( 2d \sin \theta = m \lambda \) of the wavelength and the scattering angle, i.e. only a given color light will be diffracted at a particular angle by the hologram. Another significant property is that by successively altering the incident angle (or the wavelength), a single-volume medium can store a great many coexisting holograms at one time. This latter property makes such systems extremely appealing as densely packed memory devices. A single lithium niobate crystal is capable of easily storing thousands of holograms and any one of them could be replayed by addressing the crystal with a laser beam at the appropriate angle. Imagine a 3-D holographic motion picture: a library; or everyone’s vital statistics—beauty marks, credit cards, taxes, bad habits, income, life history, etc., all recorded on a handful of small transparent crystals. Multicolored reconstructions have been formed using (black and white) volume holographic plates. Two, three
or more different colored and mutually incoherent overlapping laser beams are used to generate separate, cohabitating, component holograms of the object and this can be done one at a time or all at once. When these are illuminated simultaneously by the various constituent beams, a multicolor image results.

Another important and highly promising scheme devised by G. W. Stroke and A. E. Labeyrie is known as white-light reflection holography. Here, the reconstructing wave is an ordinary white-light beam from say, a flashlight or projector, having a wavefront similar to the original quasimonochromatic background wave. When illuminated on the same side as the viewer, only the specific wavelength which enters the volume hologram at the proper Bragg angle is reflected off to form a reconstructed 3-D virtual image. Thus if the scene were recorded in red laser light, only red light is presumably reflected as an image. It is of pedagogical interest to point out, however, that the emulsion may shrink during the fixing process and if it is not swollen back to its original form chemically (with say triethylenolamine) the spacing of the Bragg planes, d, decreases. That means that at a given angle θ, the reflected wavelength will decrease proportionately. Hence, a scene recorded in He-Ne red might play back in orange or even green when reconstructed by a beam of white light.

If several overlapping holograms corresponding to different wavelengths are stored, a multicolor image will result. The advantages of using an ordinary source of white light to reconstruct full-color 3-D images are obvious and far reaching.

ii) Holographic Interferometry

One of the most innovative and practical of recent holographic advances is in the area of interferometry. Three distinctive approaches have proved to be quite useful in a wealth of nondestructive testing situations where, for example, one might wish to study microinch distortions in an object resulting from strain, vibration, heat etc. In the double exposure technique, one simply makes a hologram of the undisturbed object and then, before processing, exposes the hologram for a second time to the light coming from the now distorted object. The ultimate result is two overlapping reconstructed waves which proceed to form a fringe pattern indicative of the displacements suffered by the object, i.e. the changes in optical path length (Fig. 14.44). Variations in index such as those arising in wind tunnels and the like will generate the same sort of pattern.

In the real-time method, the subject is left in its original position throughout; a processed hologram is formed and the resulting virtual image is made to precisely overlap the object (Fig. 14.45). Any distortions which arise during subsequent testing show up, on looking through the hologram, as a system of fringes which can be studied as they evolve in real time. The method applies to both opaque and transparent objects. Motion pictures can be taken to form a continuous record of the response.

The third method is the time-average approach and is particularly applicable to rapid, small amplitude, oscillatory systems. Here the film plate is exposed for a relatively long duration, during which time the vibrating object has executed a number of oscillations. The resulting hologram

Fig. 14.44 Double exposure holographic interferogram. [From S. M. Zivi and G. H. Humberstone, "Chest Motion Visualized by Holographic Interferometry," Medical Research Eng. p.5 (June 1970).]
can be thought of as a superposition of a multiplicity of images with the effect that a standing-wave pattern emerges. Bright areas reveal undiffracted or stationary nodal regions while contour lines trace out areas of constant vibrational amplitude.

iii) More to Come

Two recent developments seem particularly worthy of note, if only briefly; one is *acoustical* holography, the other, *computer-generated* holography.

In acoustical holography, an ultra high-frequency sound wave (ultrasound) is used to create the hologram initially and a laser beam then serves to form a recognizable reconstructed image. In one application, the stationary ripple pattern on the surface of a water body produced by submerged coherent transducers corresponds to a hologram of the object beneath (Fig. 14.46). Photographing it creates a hologram that can be illuminated optically to form a visual image. Alternatively, the ripples can be irradiated from above with a laser beam to produce an instantaneous reconstruction in reflected light.

The advantages of acoustical techniques reside in the fact that sound waves can propagate considerable distances in dense liquids and solids where light cannot. Thus acoustical holograms can record such diverse things as underwater submarines and internal body organs. In the case of Fig. 14.46 one would see something that resembled an x-ray motion picture of the fish.

It is possible to synthesize, point by point, a hologram of a fictitious object. In other words, in the most direct approach holograms can be produced by calculating, via digital computer, the irradiance distribution which would arise were some object appropriately illuminated in a hypothetical recording session. A computer-controlled plotter drawing or cathode ray tube read out of the interferogram is then photographed, thence to serve as the actual hologram. The result upon illumination is a three-dimensional reconstructed image of an object which never had any real existence in the first place.
