Wavefront Reconstruction Photography

Wavefront reconstruction, or holography, is a fascinating photographic process which is a major departure from conventional photography. In this process, discovered in 1947 by D. Gabor of Imperial College, London, the photosensitive device does not directly record an image of the subject; instead, the electromagnetic waves reflected or scattered from the subject are recorded as a standing wave pattern. The resulting photographic record is called a hologram (from the Greek word holos, meaning whole), a name given by Professor Gabor to indicate that the whole, or entirety, of the wave pattern is recorded.

The photographic record thus produced is quite unintelligible, consisting of numerous whorls, swirls, and other irrelevant patterns. However, beneath this occluding veil is an image with many fascinating and remarkable properties. When the image is uncovered, it appears projected in space in full three-dimensional form, complete with all the visual properties of the original subject, including change of perspective with shift in the observer's viewing position, and parallax between near and more distant parts. The image, in short, is an accurate re-creation of the original subject.

Since its origin by Gabor, holography has been explored by many researchers throughout the world. Recently, there has been a resurgence of interest in this area, due in part to efforts by the present authors at the University of Michigan, as well as to others elsewhere, and in part to the development of the laser, which, through the highly coherent light it produces, enables the potential inherent in holography to be realized to a degree that heretofore has simply not been possible.

Theory of holography

The fundamentals of holography have been described in many ways. The most complete treatment is contained in the various papers of Gabor.[1] Rogers[2] has described the process in a physically attractive way on the basis that each point on the subject produces on the hologram a Fresnel-zone plate. Kirkpatrick and El-Sum[3] have given a pleasing heuristic description.

The authors are members of the staff of the University of Michigan's Institute of Science and Technology. Emmett Leith is head of the optics group in the Radar Laboratory at Michigan. Juris Upatnieks is a graduate research assistant.

New interest has arisen in the wavefront reconstruction process of Gabor. With the aid of the laser, photographic imagery has been produced in which the image is, to all appearances, a three-dimensional reconstruction of the original, complete with parallax and other visual effects.

By Emmett Leith and Juris Upatnieks

We prefer to describe the process from a more communication-theory-oriented viewpoint; Gabor briefly discussed the process from this viewpoint, and Lohmann[4] carried its development further.

In Fig. 1, monochromatic, spatially coherent light illuminates the subject, which then reflects, or scatters, a portion of this light. At some plane P, the light waves can be written in the general form

$$u = a\,(x,y)\ \cos\,[2\pi ft + \phi\,(x,y)].$$

This expression represents a carrier wave that is simultaneously amplitude- and phase-modulated; f is the frequency of the light. The quantities a and ϕ are related to the reflecting surface of the subject through the Kirchhoff diffraction integral. We have assumed that scalar theory applies here, and that polarization effects can be ignored. The data a and ϕ, although functions of spatial variables, are contained on a temporal carrier wave. In holography, the objective is to transfer the data from the temporal to a spatial carrier. This is done by introducing a second beam of light, which bypasses the object but impinges also on the plane P. This beam we can write as

$$u_o = a_o\,(\cos 2\pi ft + 2\pi f_c x_o).$$

The second term of the argument indicates a linear phase shift across the plane, and the wave thus impinges on the plane at some oblique angle.

At plane P, we place a photosensitive device, for example, a photographic plate; such a device responds to the time-averaged value of the intensity of the incident light, thus acting as a square-law device. As a result, the function

$$<(u_o + u)^2> = a^2{}_o/2 + a^2/2 + a_o a \cos\,(2\pi f_c x - \phi)$$

is recorded on the plate, where the $< >$ indicates

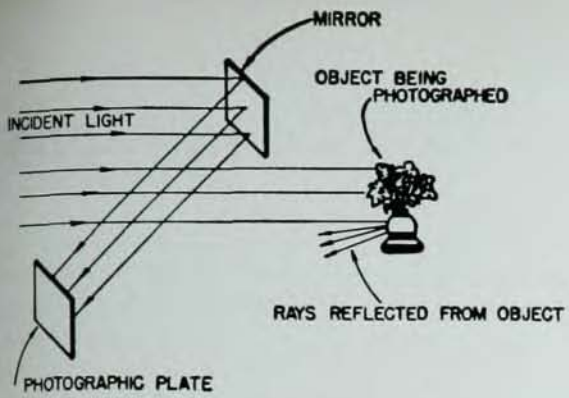


Fig. 1. Optical system for making a hologram

a time average. The auxiliary beam has functioned as a local oscillator signal, and the photographic plate, in addition to its role as a storage device, has functioned as a mixer, producing the difference frequency term $a_0 a \cos(2\pi f_c x - \phi)$. The signal carried by the light beam has been modulated onto a spatial carrier wave $\cos 2\pi f_c x$, in such a way that the information is preserved in its entirety without degradation.

The reconstruction

The reconstruction process is essentially the inverse of the hologram producing process; the interaction of a coherent light beam with the hologram record causes the modulation present on the spatial carrier of the hologram to be transferred to the light beam, in the following manner. Let $u_1 = a_1 \cos\,(2\pi f't + 2\pi f_c x)$ represent the light impinging on the hologram; the light impinges obliquely, just as did the reference beam used in making the hologram. Note that the frequency f' of the light used in the reconstruction is not necessarily that of the light used in making the hologram. The light emerging from the surface of the hologram is thus

$$[a_1\cos(2\pi f't+2\pi f_c x)]\left[\frac{a_0^2}{2}+\frac{a^2}{2}+a_0 a\cos(2\pi f_c x-\phi)\right]$$
$$=\frac{a_1}{2}(a_0^2+a^2)\cos\,(2\pi f't+2\pi f_c x)$$
$$+\frac{a_1\,a_0 a}{2}\cos\,(2\pi f't+4\pi f_c x-\phi)+\frac{a_1\,a_0 a}{2}\cos\,(2\pi f't+\phi).$$

The emerging light is seen to consist of three terms; the first of these is of no interest, since it carries no phase modulation. The other two are the usual sum and difference frequencies produced by a mixing process. The difference frequency is identical in form to the original signal-bearing wave recorded on the hologram. This term represents a reconstruction of the original waves, and, when presented to an observer, appears to emanate from a virtual image located on the source side of the plate. This image has all the visual properties of the original object; if the original object were three-dimensional, then the reconstructed image is likewise; this three-dimensional effect is produced without the need for stereo pairs of holograms and without the need for any viewing devices such as polaroid glasses. In addition, as the viewer moves his head, his perspective of the picture changes. Parallax between near and far objects is observed.

The sum term represents a real image, which forms on the side of the plate away from the source. The observer can view this image also; he finds it suspended in space between himself and the plate. The real image, however, has the property of being pseudoscopic; this is an effect that occurs in stereo viewers when the stereo pair of pictures is interchanged. Anomalous effects occur; hills become valleys, protrusions become indentations, etc. These effects give the real image an unnatural and confusing appearance.

One of the major problems in holography has been the separation of the three terms given above. Pursuing our carrier-wave viewpoint, it is apparent that a spatial filtering technique can achieve this separation, providing the spatial carrier term f_c is sufficiently great that the spatial-frequency spectra of the three terms do not overlap. This separation could be effected on the basis of ideas developed by Duffieux,[5] who pointed out that, in an imaging process carried out with coherent illumination, the Fourier spectrum of the object is displayed at the back focal plane of the imaging lens. Stops and slits placed here act like stop-band and pass-band filters. When the hologram is thus imaged through an optical system, a slit can select either the real- or the virtual-image term while rejecting the others.

The three terms can be separated on a more simple basis, which is readily explained by thinking of the hologram as a diffraction grating. The three waves emerging from the hologram can be identified with the zero order and the two first orders of the grating. The waves associated with these orders propagate in different directions and at some distance from the hologram are separated, as shown in Fig. 2. The two first-order diffractions correspond to the real and virtual image terms.

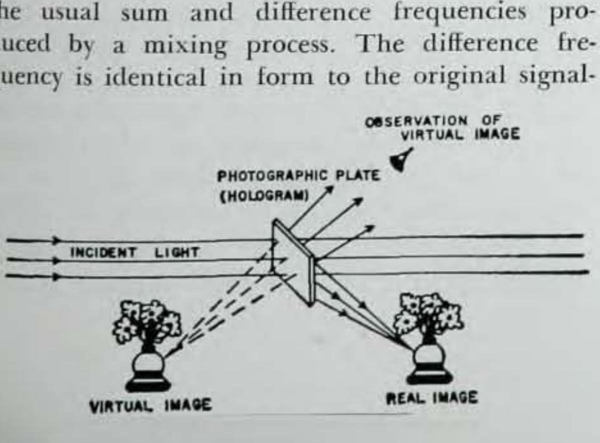
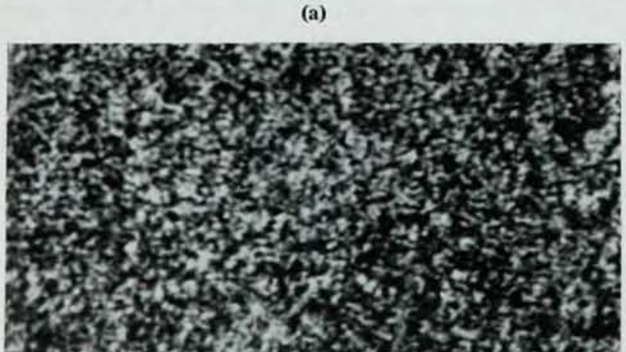


Fig. 2. The reconstruction process

Holograms have become familiar to many as patterns of lines, specks, blobs, and whorls; this is an association which, if not wrong, is misleading. Such manifestations arise from dust particles on the reference-beam mirror and other such anomalies; they can be eliminated with good experimental technique, as the pictured hologram, Fig. 3 (a), shows. The recorded signal appears as an isotropic random granular pattern, like that shown in the magnified view, Fig. 3 (b).

In Fig. 4 we have attempted to show the three-dimensional properties by means of several reconstructions of the real image, taken at different positions of the recording plate, and at different *f*-numbers (the *f*-number of the recording system can be varied by stopping the hologram itself—i.e., by restricting its aperture, just as one stops a lens). In Fig. 4 (*a*) and (*b*) the hologram was stopped considerably by illuminating only a small area (a circular area approximately 3 mm in diameter). The parallax differences between (*a*) and (*b*) were produced by using different portions of the hologram for the two pictures. A similar effect would be produced by moving one's head while viewing the reconstruction. In (*a*) we are looking up at the tank; in (*b*) we are looking from above the tank. The parallax differences are most prominent in the way the gun barrel is seen against the rest of the tank. Pictures (*c*) and (*d*)

(a)



(b)

Fig. 3. Example of a hologram; b shows a highly magnified portion of a hologram.

were produced by broadening the illuminating beam to about 40 mm; the depth of focus is small compared to the depth of the tank, and only a small piece is in focus at a time.

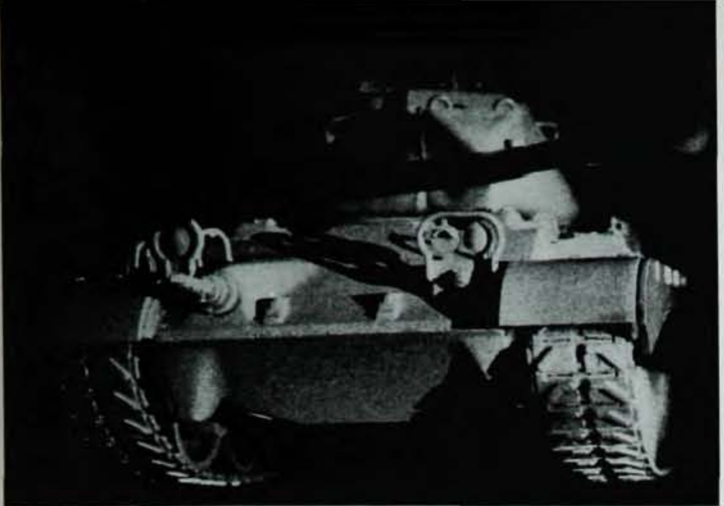
In the past year, many persons have attempted to produce holograms; some have been successful, others have encountered difficulty. Good results are readily attained if proper equipment is used; otherwise failure is likely. Since the method is a type of interferometry, stability of the various components is essential. The work is best performed on a granite block using subjects and equipment that will not vibrate during the exposure. A helium-neon cw gas laser of 1-mW output power is suitable, but a 40-mW laser is vastly better, since the exposure times are reduced from several minutes to several seconds.

The development of holography

Gabor originally conceived of the wavefront reconstruction process as a means for improving the imagery of the electron microscope. The problem was to overcome the spherical aberration inherent in the electron lenses; the solution was to produce a hologram as the output image from the electron microscope and to make the reconstruction using visible light. The spherical aberration would appear in the reconstructed waves and could be removed by using the well-developed methods of visible-light optics. The potential gains are great, since the electron microscope falls far short of approaching the theoretical resolution limits. Although Haine and others[6,7] continued the development, technical difficulties have prevented the practical realization of this goal.

El-Sum and Baez[8,9] in the early 1950's investigated the adaptation of the wavefront reconstruction technique to x-ray microscopy. The hologram would be made with x radiation, and the reconstruction would be made with visible light. Since x rays cannot be focused except crudely and with great difficulty, the resolutions achieved with x rays fall short of the theoretical limits by several orders of magnitude; the wavefront reconstruction technique, however, has the possibility of overcoming this limitation. The reconstructed waves, produced with visible light, could be sharply focused and an image of high quality could thus be produced.

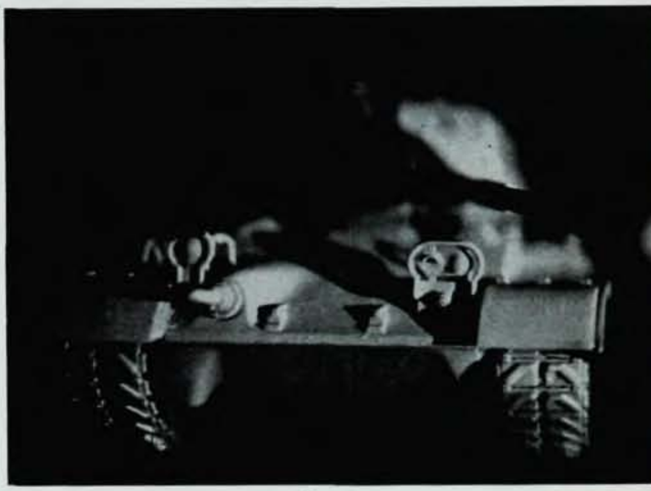
Baez and El-Sum[10] demonstrated the method, but technical difficulties prevented the realization of the full potential of wavefront reconstruction methods. The difficulty lay primarily in finding an x-ray source with sufficient spatial coherence and yet with reasonable intensity.



(a)

(b)

Fig. 4. Three-dimensional reconstruction of a toy tank; a and b show two views with different parallax. An aperture of 3 mm exists at the hologram. In c and d, the aperture is opened to 40 mm. Note the decrease in depth of field. In each case, only a small portion of the object is in focus. Pictures are further discussed in text.



(c)

(d)

Holograms made with visible light were produced by Gabor and by many others since. The resolutions achieved have been comparable to those attained by means of more conventional optical systems. However, the overall quality has been deficient, largely because of the inability to separate the three terms previously noted. This has meant generally that whichever image was used, real or virtual, the other image was present as a defocused background. In addition, the remaining term, which has noise-like properties, was also present. Consequently, most examples of wavefront reconstruction have been carried out with high-contrast, relatively simple objects.

It may be wondered why this separation has been a problem, since the separation, as we have seen, automatically occurs in the reconstruction process as we have described it. The difficulty has been that the traditional hologram-producing methods have used a spatial carrier frequency, f_c, equal to zero. Under this condition, it is seen that the three terms noted earlier do not separate.

The separation of the real and virtual image has been one of the traditional problems of holography, and many methods have been proposed for doing this. The most simple of these methods involves placing masks of some sort in the light path between source and reconstructed image. These methods gave moderate improvements, but are not among the more effective solutions.

The presence of the twin images results from the incomplete recording of the phase of the incident light. Accordingly, one might consider the production of two holograms which are complementary in the sense that the deficiencies of one are compensated by the other. The reconstruction would be made by using both holograms together. Bragg and Rogers,[11] Gabor,[12] and El-Sum[9] have proposed this kind of technique. Such techniques are sound, but they present technical problems, one of which is that, in the reconstruction process, extremely precise alignment of the two holograms is required. These methods have some advantages over the spatial-carrier method employed by the authors[13]; for example, they require less spatial coherence and monochromaticity in the light source.

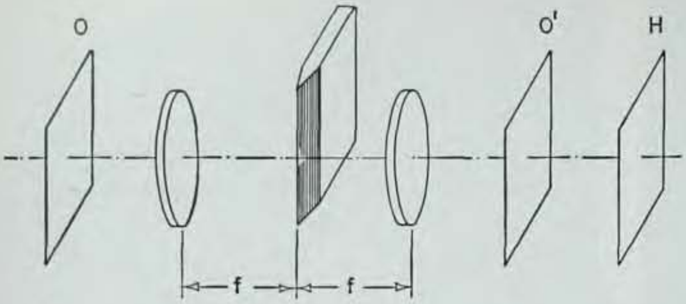


Fig. 5. Lohmann's single-sideband method for removal of the twin image; object is placed at O, image appears at O′, and the hologram is made at H. A half-plane stop is placed midway between the two lenses, each of focal length f and separated by the distance 2f.

Lohmann[4] in 1956 proposed for removal of the twin image, a method which is similar to the spatial-carrier or off-axis–reference-beam method. Lohmann's method is illustrated in Fig. 5. The hologram is made through a lens system containing a stop in the focal plane of the first lens. The object, in this case a transparency, is imaged at plane O′, after removal of half of its spatial frequencies as a consequence of the half-plane spatial filter. The hologram is made at the plane labeled *H*.

In the reconstruction process, the hologram may be placed in its original position and the light sent through the optical system in the opposite direction. The spectrum is displayed, as before, at the focal plane of the lens. The virtual image term appears on one side of the optic axis; the real term on the other; the spectra are thus separated. A half-plane stop removes either the real- or the virtual-image term, and in the reconstruction (which occurs to the left of the lens system) the interfering twin image is absent.

The holograms being produced nowadays have a quality that is considerably higher than had heretofore been published. There are two reasons for this. First is the use of the carrier-frequency or off-axis reference-beam method, which effects the removal of the desired image from its twin image and from other, noise-like terms. Second is the use of the laser, which provides a highly coherent source far more intense than previously used sources. While conventional sources, such as the mercury arc, have yielded good results with photographic transparencies for the subject, the holograms made from diffusely reflecting, three-dimensional objects, which produce the most dramatic results, almost have to be made with a laser source. The chief reason for this is that the source must have a coherence length equal to the depth of the scene. Coherence length is proportional to the monochromaticity of the source, and, if a conventional source is made monochromatic to the required degree for an object several inches in depth, the intensity becomes quite low and the required exposure time becomes quite long. If laser technology were extended into the far ultraviolet or x-ray regions, holography at these short wavelengths would presumably become an important method of x-ray and uv microscopy.

The technique of bringing the reference beam onto the recording plate from an off-axis position is an adaptation of well-known techniques used in electronic engineering, and it was because of our experience in microwave work that the idea occurred to us.

For example, in certain radars of the pulse-compression type, a linearly frequency-modulated waveform is radiated, and the received signal is passed through a network which compresses the pulse, thus providing resolution equal to that which would have been produced had a simple pulse with the same bandwidth been radiated. Such a radar system is a close analog of the wavefront reconstruction process. In each case, the signal produced by a point target is a structure resembling a Fresnel-zone plate, and in both cases the final processing consists of converting each zone-plate response into a point spot. In holography, this is done through the focal properties of the zone plates recorded on the hologram; in pulse-compression radars this is done usually by means of electrical networks that carry out an equivalent operation.

Before the radar signal is compressed, it is converted from its rf carrier onto a suitable if carrier. The carrier, however, is selected to be at least equal to the bandwidth of the pulse, so that the sidebands of the pulse spectrum are not folded into each other. The recording of the hologram is analogous to conversion of the radar pulse to if, and clearly, on the basis of the comparison drawn here, the hologram should be on a spatial carrier; the off-axis reference beam accomplishes exactly this. The off-axis reference beam, then, is but the adaptation of a technique from one discipline to another.

Current status of holography

Currently, there appears to be a high level of interest in holography; throughout the country experimenters are producing holograms, and suppliers of the Lippman-type emulsions which are so suitable to this application are reportedly doing a brisk business. At the recent Optical Society meeting in Dallas, a total of eight papers on holography were presented. Areas of application for holography are being extensively explored, including the original ones proposed by Gabor, El-Sum, and others who pioneered in this area.

One of the most promising applications is reported by Thompson and Parrent. These investigators use the technique for examination of small fog-like particles enclosed in a confined volume. A hologram is made of the entire volume, using a pulse laser; this, in effect, freezes the motions of the particles and enables one, in the reconstruction process, to examine at length the particle configuration as it existed at the instant the hologram was made. Thompson and Parrent have thus (as have the authors) made use of the excellent coherence properties of the laser, and of the three-dimensional imagery inherent in holography. Their effort, it should be noted, has paralleled and in no way stems from, the work of the present authors. Thompson and Parrent do not employ an off-axis reference beam, but the subjects they work with do not require such a technique.

Two of the authors' colleagues, Powell and and Stetson,[14] have explored the use of off-axis reference-beam holography for examination of vibratory motions. The method is based on the loss of coherence and its consequent effect on the reconstructed image, of light reflected from a vibrating object. Horman,[15] in a recent paper, has proposed several applications of wavefront reconstruction to interferometry.

The possibility of carrying out holographic techniques with incoherent light was proposed by Mertz.[16] As yet, high-quality imagery by this method has not been attained, although new attention is being directed to this end by Cochran.[17] The need for coherent light is a severe constraint, the elimination of which would be highly desirable.

The successes achieved in the visible region of the spectrum have generated new interest in the old problems of making holograms with x rays and with electrons in an electron microscope. El-Sum has previously demonstrated such holograms, but the technique has not been developed thus far to a practical stage. We hope for new advances in these areas, but the problems remain formidable.

The production of holograms in a three-dimensional storage medium was discussed in a paper by P. Van Heerden.[18] Here, the recorded fringes become surfaces within the medium, and the readout, or reconstruction, is carried out on the basis of Bragg-angle diffraction, in a manner analogous to diffraction from crystals. The reconstructed image is produced only when the orientation of the hologram relative to the illuminating beam is proper.

A photographic emulsion is generally regarded as a two-dimensional medium, but, when the recorded detail becomes greater than the emulsion thickness, the emulsion must be regarded as a three-dimensional medium. The spectrographic plate on which we have made our holograms have an emulsion thickness of about 6×10^{-3} mm; thus, when the recorded fringe pattern exceeds about 200 lm/mm, the emulsion behaves as a three-dimensional medium. Our holograms have been made at spatial frequencies ranging from a few lm/mm up to about 2000 lm/mm, and at the higher range the Bragg diffraction characteristics become prominent. Hence, our holograms certainly demonstrate the ideas of Van Heerden, although this was not our intent.

We were not motivated in our work by practical applications; however, we are quite interested in seeing such applications develop; indeed it seems to us that the dramatic imagery produced by this method must inevitably lead to worthwhile uses in addition to those already proposed or demonstrated.

References

1. D. Gabor, Nature **161**, 777 (1948); Proc. Roy. Soc. (London) **A 197**, 454 (1949); Proc. Phys. Soc. (London) **B 64**, 449 (1951).
2. G. L. Rogers, Proc. Roy. Soc. (Edinburgh) **A 63**, 193 and 313 (1952); **A 64**, 209 (1956).
3. P. Kirkpatrick and H. M. A. El-Sum, J. Opt. Soc. Am. **46**, 825 (1956).
4. A. Lohmann, Opt. Acta (Paris) **3**, 19 (1956).
5. P. M. Duffieux, *l'Intégrale de Fourier et ses Application à l'Optique*, chez l'Auteur, (Université de Besançon, Besançon, Franch, 1946).
6. M. E. Haine and J. Dyson, Nature **166**, 315 (1950).
7. M. E. Haine and T. Mulvey, J. Opt. Soc. Am. **42**, 763 (1952).
8. A. V. Baez, J. Opt. Soc. Am. **42**, 756 (1952).
9. H. M. A. El-Sum, "Reconstructed Wavefront Microscopy", PhD Thesis, Stanford University, November 1952 (Available from University Microfilm, Inc., Ann Arbor, Michigan).
10. H. M. A. El-Sum and A. V. Baez, Phys. Rev. **99**, 624 (1955).
11. W. L. Bragg and G. L. Rogers, Nature **167**, 190 (1951).
12. Papers presented at Washington conference on electron microscopy, National Bureau of Standards, Nov. 1951.
13. E. Leith and J. Upatnieks, J. Opt. Soc. Am. **53**, 1377 (1963); **54**, 1295 (1964).
14. R. Powell and K. Stetson, J. Opt. Soc. Am. **5**, 612 (1965).
15. M. H. Horman, J. Appl. Opt. **4**, 333 (1965).
16. L. Mertz and N. O. Young, Proceedings of the International Commission on Optics Conference on Optical Instruments (London 1961) p. 305.
17. G. Cochran, J. Opt. Soc. Am. **5**, 615 (1965).
18. P. Van Heerden, J. Appl. Opt. **2**, 393 (1963).