Reprinted from APPLIED OPTICS, Vol. 10, page 619. March 1971

Computer Generated Holograms for Testing Optical Elements

A. J. MacGovern and J. C. Wyant

The application of computer generated holograms to the interferometric testing of aspheric optical elements has been investigated, and it has been shown that they provide a convenient and practical method of producing an aspheric reference wavefront.

Introduction

The measurement of the surface shape of an optical element, or the wavefront produced by an optical element, is usually a relative and not an absolute measurement. The surface is determined with respect to a reference surface, which is supposedly of known shape to an order of magnitude better accuracy than the surface under test. For example, the Twyman-Green interferometer is a common arrangement for testing spherical optical elements, the reference surface being a flat mirror. If the element under test is aspheric, this arrangement is limited in accuracy because the difference between the reference surface and the test surface becomes too large. This paper discusses the use of computer generated holograms to modify predictably either the reference wavefront or the object wavefront so that the wavefronts exiting from the interferometer are more closely matched.

Figure 1 shows one method of using a computer generated hologram with a Twyman-Green interferometer. If the hologram is placed in the image plane of the exit pupil of the surface (or system) to be tested, the wavefront that the hologram is required to reconstruct is simply the departure of the desired wavefront from a best fitting spherical wave. For many cases, this is an analytic function that can be evaluated at a large number of points with very little computing time.

In general, several wavefronts (different diffraction orders) emerge from the hologram. The order of particular interest is selected by spatial filtering, an indicated in Fig. 1.

The authors are with Itek Corporation. Lexington, Massachusetts 02173.

Received 10 June 1970.

Binary Holograms

With available graphic devices, binary holograms are the most convenient to generate by computer. In 1967, Lohmann and Paris[1] showed that wavefronts could be produced with a binary hologram divided into equispaced resolution cells. The number of resolution cells needed depends on the complexity of the wavefront that is to be produced. A rectangular aperture is placed in each cell. The area of this aperture is made proportional to the amplitude, as calculated at the hologram plane, and it is shifted from the center of the resolution cell by an amount proportional to the phase of the wavefront at the hologram plane. It is inconsequential whether the hologram consists of opaque rectangles and clear background or vice versa.

Another method for encoding the complex amplitude information into the hologram was deduced by Lee.[2] Lee considered the wavefront in terms of real and imaginary part's, both positive and negative, rather than in terms of its amplitude and phase. Lee also considered equispaced resolution cells. Figure 2 shows the construction of a typical cell. The cell is divided into four equal sections. The positive real part of the wavefront is calculated at the center of the first section, and a rectangle proportional to this value is placed there. Rectangles having areas proportional to the positive imaginary, negative real, and negative imaginary are placed in the second, third, and fourth sections, respectively. In general, only two rectangles need to be placed in each cell.

The Lee approach is of particular interest if several gray levels, rather than strictly binary apertures, are available.

Theoretical Considerations

In Fig. 1, the hologram is shown in the reference arm of the interferometer. It is positioned so that its image and the image of the piece under test coincide. The imaging and spatial filtering of the wavefront produced by the hologram are shown in more detail in Fig. 3. This situation is slightly different from that described and analyzed by Lohmann and Paris,[1] and it is **neces-**

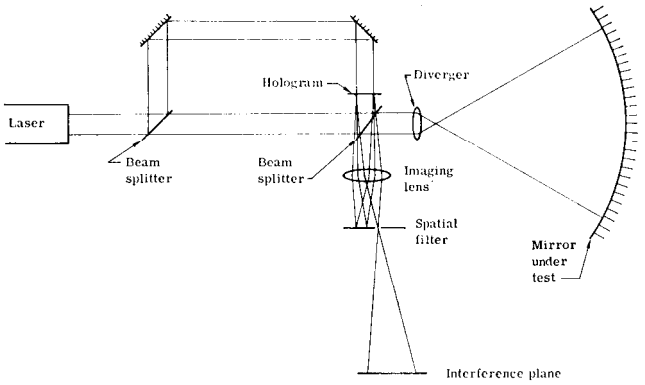


Fig. 1. Modified Twyman-Green interferometer.

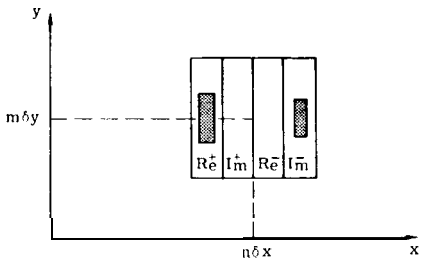


Fig. 2. Structure of Lee type resolution cell.

sary for completeness to show that this arrangement gives a close approximation to the required reference wavefront.

The derivation of the expression for the reconstructed wavefront is given in the Appendix. It is of value here in assessing the use of computer generated holograms for providing accurate reference wavefronts to clarify the physical process by which they give a reconstruction. This is made especially simple in the present application since the required wavefronts have uniform amplitude.

Consider a Lohmann and Paris type hologram. It consists of a set of small apertures (or point sources if the apertures are small enough) positioned in such a way that the phase of some illuminating wavefront at these apertures is precisely that of the required wavefront at the same points. Therefore, a sampled version of the required wavefront exists from the hologram. It is clear from Fig. 3 that if the filtering aperture in the focal plane of the imaging lens is removed, an image of this sampled wavefront occurs in the image plane, i.e., the required wavefront is obtained at a set of discrete points but not between the points. This is true no matter where or how many sampling points are taken. The spatial filter introduced into the focal plane of the imaging lens has the effect of distributing the light from each sampling point to the area between the corresponding points in the image plane. If the hologram is divided into sufficient equispaced resolution cells and one sample is taken per cell, then this distribution of light results in a reasonable approximation to the required wavefront at all points in the image plane. It should be noted that the wavefront can never be exact even for functions that may be considered band limited, because the hologram is a nonequispaced sampled version of the required wavefront and as such is not subject to the Whittaker-Shannon sampling theorem.[3]

A Lee type hologram is slightly different in that it is an equispaced sampled version of the required wavefront, or at least of four separate wavefronts that combine to give the correct wavefront. If these four wavefronts are band-limited, they are subject to the sampling theorem and may give an exact result.

The relationship of a Lohmann and Paris type hologram to a real hologram becomes evident by again realizing that the former consists of a set of apertures positioned so that the phase of the illuminating wavefront is equal to the required phase of the reconstructed wavefront (± integer x 2π). A real hologram, made by interfering the reconstructing wavefront with the required wavefront (if it existed), would consist of a set of interference fringes whose maxima are also the loci of points where the phases of the two wavefronts are equal (to within an integer X 2π). Thus, the apertures of a Lohmann and Paris type hologram simply lie along the maxima of the interference fringes of a conventional hologram. Normally, the reconstructing wavefront is a plane wave tilted so that there is one wave optical path difference between the centers of adjacent resolution cells (of side δx). The apertures

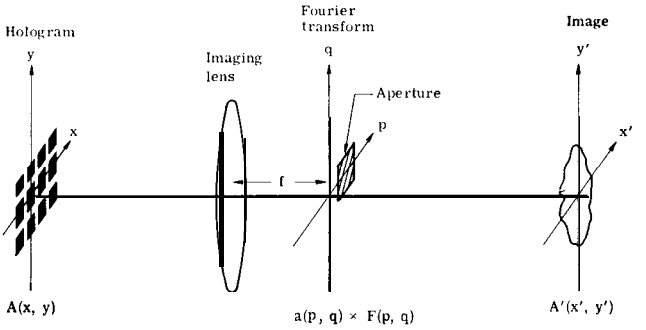


Fig. 3. Spatial filtering of hologram reconstruction.

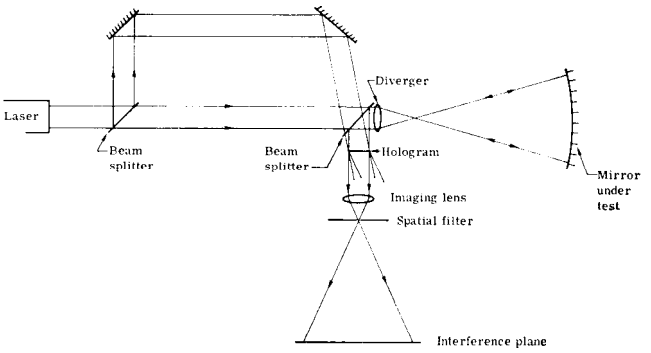


Fig. 4. Arrangement that reduces errors due to sampling.

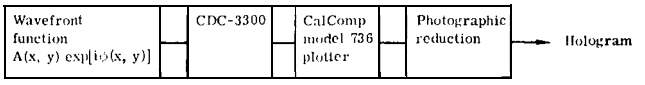


Fig. 5. Procedure for generating synthetic holograms.

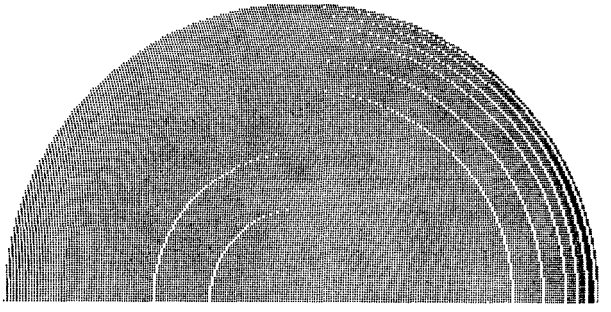


Fig. 6 Half of Lohmann and Paris hologram of defocused parabolic wavefront.

must then be at x positions, which satisfy the following equation

$$(2\pi/\delta x)\cdot x - j2\pi = \phi(x,y), \quad (1)$$

where ø(x, y) is the required phase function and j is an integer.

In practice, it may be difficult to solve this equation if ø(x, y) is complicated. It is easier to find the x values that satisfy

$$(2\pi/\delta x)\cdot x - j2\pi = \phi(n\delta x, m\delta y), \quad (2)$$

where m, n, and j are integers and δ y is the y extension of a resolution cell. j is selected such that x lies within the nth resolution cell, i.e., x is between $x + n\,\delta x/2$ and $x - n\,\delta x/2$. In the terminology of Lohmann and Paris, this corresponds to shifting the apertures from the centers of the resolution cells by an amount proportional to the phase (modulo 2 π) at the center of the cell rather than the phase at the shifted position. Most of the holograms described in this paper were made this way.

Error Analysis

The following sources of error are in the phase of the wavefront produced by computer generated holograms: (1) finite size of apertures making up hologram, (2) unequal spacing of sampling points, (3) quantization errors, and (4) graphic recorder nonlinearity and photographic distortion.

The effect of these sources of error depends on the number of sampling points taken and on the complexity of the wavefront produced. The first three errors can be reduced to any desired value, if the graphic recorder has a sufficient number of resolution points.

The finite aperture size has the effect of convolving the distribution in the image plane with a rectangular function having the width of the image of the aperture. In other words, the phase at any point is the result of averaging the continuous phase variation over an area the size of the aperture. Theoretically, this phase error can be made as small as desired by either decreasing the size of the apertures or by increasing the number of sampling points.

Experimental results will be presented later in this report for a hologram made to test an f/5 18-in. (46-cm) diam parabolic mirror. For this hologram, there were 256 X 256 resolution cells and the rectangle placed inside each resolution cell had half the width of the cell. For this situation, the finite aperture size theoretically causes a peak-to-peak error of approximately 0.001 wave.

Because of the unequal spacing of sampling points, points in the image plane receive erroneous contributions from neighboring apertures. This error can also be reduced by increasing the number of sampling points. For the hologram described for testing the f/5 parabolic mirror, the peak-to-peak error caused by the unequal spacing of sampling points was calculated to be approximately 0.001 wave.

It is interesting to note that by placing the hologram in the same plane as the image of the exit pupil of the element under test, as shown in Fig. 4, the effects of the error due to the spreading out of the image points are essentially removed. In this arrangement the same error is introduced into both interfering wavefronts. Any fringe pattern observed is simply the moiré between the hologram and the interference of the plane wave and the wavefront from the aspheric element. However, the small error estimates and the practical difficulties of this arrangement make it less desirable for general testing.

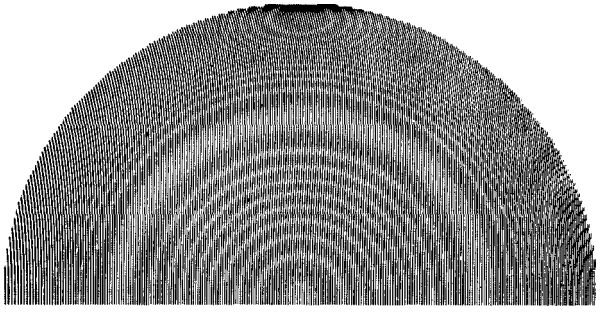


Fig. 7. Half of Lee type hologram of defocused parabolic wavefront.

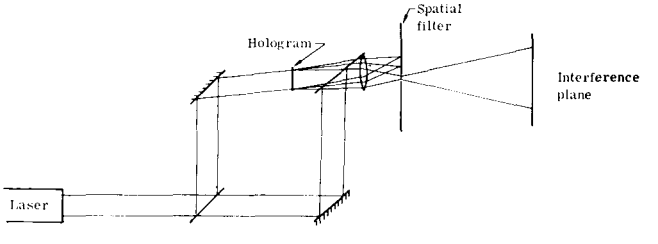
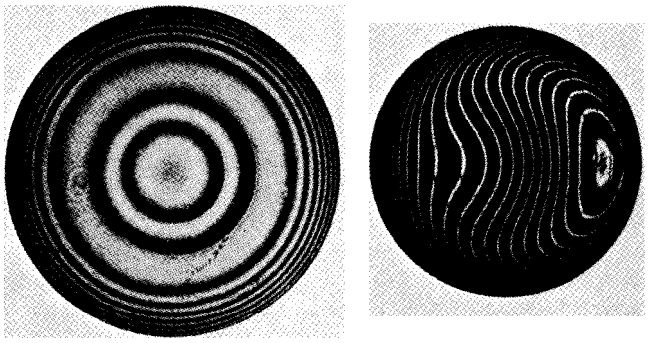


Fig. 8. Arrangement for testing quality of synthetic hologram.



(a) No tilt (b) Tilt introduced

Fig. 9. Interferogram resulting from interfering plane wavefront with parabolic wavefront produced by Lohmann and Paris type hologram.

A simple argument gives an estimate of the quantization error. Suppose the graphic device has P resolution points. Let N be the chosen number of resolution cells (in the x direction). Thus, there are P/N aperture positions per cell. The maximum error due to quantization of position of an aperture is $N/2P$ of a wave. The rms error will be approximately half this or $N/4P$ of a wave. The graphic device used to make the hologram for testing the f/5 parabolic mirror had 6000 resolution points. Since the hologram had 256 X 256 resolution cells, there were 23 aperture positions per cell. Thus, the theoretical rms error resulting from phase quantization was approximately 0.01 wave.

Any graphic recorder nonlinearity or photographic distortion causes some of the apertures making up the hologram to be positioned incorrectly. The error due to nonlinearity and distortion can be found empirically by analyzing the reconstruction of a hologram made to produce a plane wave. Experiments showed that for holograms included in this paper the nonlinearity and distortion errors were less than l/15 wave peak to peak.

As mentioned previously, it is generally easier to solve Eq. (2) for the approximate hologram aperture positions rather than to solve Eq. (1) for the correct positions. For the hologram made to test the f/5 parabola, the theoretical peak-to-peak error that results from using Eq. (2) to determine aperture position is approximately l/20 wave. If this error were too large, the magnitude of the error could be reduced by increasing the number of sampling points, or the error could be eliminated completely by using Eq. (1) to determine aperture position.

Experimental Work

Figure 5 shows the general procedure for generating the synthetic holograms. The mathematical expression describing the wavefront to be reconstructed is fed into a CDC-3300 computer, and the output data are delivered to a CalComp plotter. The output plot is photographically reduced to give the required hologram.

Figure 6. shows half of a hologram made, using the Lohmann and Paris technique, for reconstructing a defocused parabolic wavefront suitable for testing an f/5 parabola. Because the amplitude is constant, the rectangles do not vary in size. They are adjusted in position according to the phase varration across the wavefront.

The function was sampled at 256 X 256 points. A plotter width of 30 in. (76 cm) was used with 6000 possible pen positions available. Thus, with 256 resolution cells across the plotter there are about 23 pen positions available in each resolution cell. This results in a phase quantization of about 15.5°. The other half of this hologram is identical to that shown in Fig. 6, so rather than plot the whole function, two photographically reduced copies were made and attached together. Figure 7 shows a hologram representing the same function sampled at the same rate but formed using the Lee scheme. The Lohmann and Paris type hologram required about 5 min of computer time and 10 h of plotting time, whereas the Lee type required 15 min of computing time and 14 h of plotting time. The output plots were photographically reduced about 100 times to give the required holograms.

To test the fidelity of the wavefronts produced by these holograms, their reconstructions were interfered with a plane wave, using the arrangement shown in Fig. 8. Figure 9(a) shows the interference pattern obtained from the hologram, using the Lohmann and Paris scheme. As expected, circular fringes were obtained. Analysis of the fringe positions shows that the hologram produced the required wavefront to the rms accuracy of about l/15 wave. Figure 9(b) shows the same pattern, but with some tilt introduced. The fringe pattern from the Lee type hologram is shown in Fig. 10. Analysis of these fringes also shows an rms error of about l/15 wave. It appears that for both types of holograms the graphic recorder nonlinearity is the limiting factor in the fidelity of the wavefront produced by the hologram.

For the plotter used, the Lee scheme appears to have no advantage over the Lohmann and Paris scheme since it takes longer to calculate and plot and does not give better results. However, the Lee scheme may have considerable advantages with gray level plotters

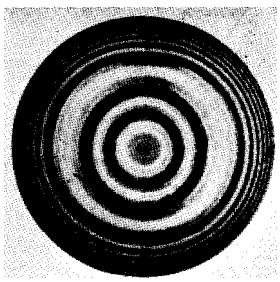


Fig. 10. Interferogram resulting from interfering plane wavefront with parabolic wavefront produced by Lee type hologram.

Fig. 11. Interferogram resulting from interfering parabolic wavefront produced by hologram with wavefront produced by parabolic mirror.

since it does not require as many plotter resolution points as the Lohmann and Paris scheme.

The Lohmann and Paris type hologram of Fig. 6 was used to test an actual f/5 parabola, using the modified Twyman-Green interferometer shown in Fig. 1. Figure 11 shows the fringes resulting from this test with some tilt introduced. Since the hologram was already tested to l/15 wave, most of the departure from straightness of these fringes may be directly attributed to errors in the parabola. An optician looking at the fringe pattern could tell what work must be done to correct, the mirror surface.

The parabola that was tested had a departure of only sixteen waves from the base sphere. A hologram made to test more severe aspherics must contain more resolution cells. This means that either a plotter having a greater number of resolution points must be used or the phase must be quantized in larger intervals. Some investigation was carried out to see how few times the phase need be quantized while maintaining adequate reconstruction. Figure 12 shows the interference patterns obtained with 36° and 90° quantization for holograms of the same parabolic function, sampled at the same rate. The rms deviation of the reconstruction is about l/10 wave for the 36° quantization and about l/8 wave for 90° quantization. This may be adequate for many applications, and it shows that reasonably good reconstructions can be obtained even with coarse phase quantizations.

A problem associated with testing more severe aspherics, using the interferometer of Fig. 1, is that the diverger may contribute aberrations to the distorted wavefront returning through it. One way to eliminate this difficulty is to use the interferometer of Fig. 13 in which the hologram is placed in the object beam. The incident spherical wavefront on the hologram will reconstruct a wavefront which, if it matches the test piece surface, will return to the hologram unchanged. This wavefront returning through the hologram will reconstruct a spherical wavefront that will be collimated by the diverger and interfered with the reference wavefront via the beam splitter. Since there are two sequential reconstructions involved, the intensity of the returning object beam will be low. However, initial experiments using this arrangement show it to be a feasible method with adequate light returned.

Conclusion

The application of some forms of synthetic holograms to the testing of optical elements has been investigated and it has been shown, at least for simple aspherics, that the method provides a convenient and practical means of producing an aspheric reference wavefront.

Results have been shown only for the testing of a reflective parabolic element, but there is no reason why the same technique cannot be applied to the testing of nonrotationally symmetric elements, both reflective and refractive. The effect of the various plotting parameters on the accuracy to be expected and how this accuracy is affected by the severity of the aspheric have been determined. It is expected that with more sophisticated graphic devices, aspheric wavefronts that are several times more extreme than have been considered up to now may be generated.

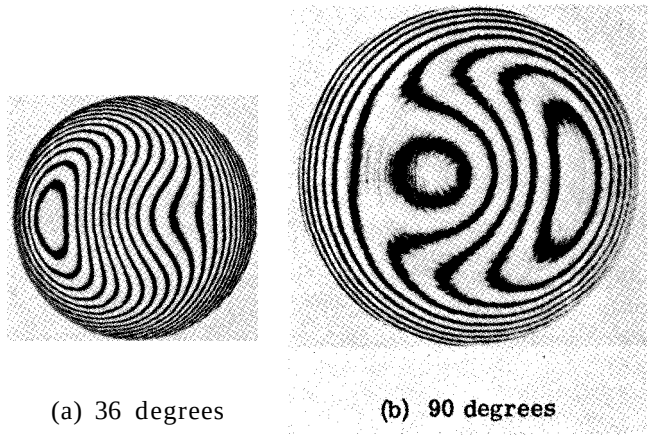


Fig. 12. Interferograms obtained from holograms having 36° and 90° phase quantization.

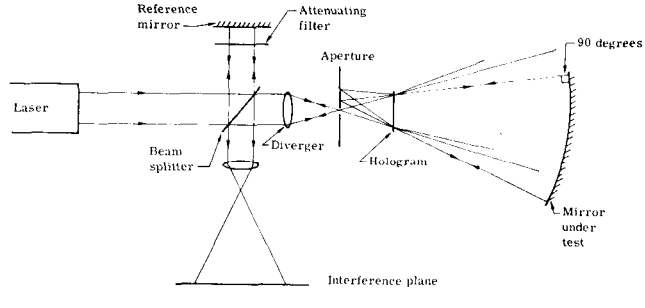


Fig. 13. Arrangement for testing aspherics with hologram acting as a null lens.

The authors wish to acknowledge the support and technical assistance of W. P. Barnes, I. P. Adachi, and J. C. Chastang and the invaluable contribution of M. G. Hurwitz in the development of the computer programs.

Appendix

Figure 3 shows the physical situation to be theoretically analyzed. A Lohmann and Paris hologram consisting of a set of shifted apertures is fourier transformed by the imaging lens, and the result is filtered by the aperture function and effectively retransformed (aside from a quadratic phase factor) to give the wavefront in the image plane.

Let $A(x, y)$ represent the hologram transmission function, a(p, q) its transform in normalized coordinates, $F(p, q)$ the filter function, and $A'(x', y')$ the result in the image plane. Equation (Al) shows $a(p, q)$ written as the transform of $A(x, y)$:

$$a(p,q) = \int_{-\infty}^{\infty}\int A(x,y)\exp[i(px+qy)]\,dx\,dy. \quad (A1)$$

Since $A(x, y)$ is the sum of a set of rectangles, the integral with infinite limits may be written as the summation of integrals with finite limits. The limits contain the position and the boundaries of each rectangle. Thus,

$$a(p,q) = \sum_{(n,m)}\sum \int_{n\delta x-b_{nm/2}+S_{nm}}^{n\delta x+b_{nm/2}+S_{nm}} \exp(ipx)\,dx \times \int_{m\delta y-c_{nm/2}}^{m\delta y+c_{nm/2}} \exp(iqy)\,dy, \quad (A2)$$

where x and y are the sides of a resolution cell, n and m are integers, b_{nm} and c_{nm} are the sides of each rectangle, and S_{nm} is the shift of each rectangle away from the center of the corresponding resolution cell. When the integrals are solved, we obtain

$$a(p,q) = \sum_{(n,m)}\sum b_{nm}c_{nm}\,\mathrm{sinc}\left(\frac{pb_{mn}}{2}\right)\mathrm{sinc}\left(\frac{qc_{nm}}{2}\right) \times \exp[ip(n\delta x+S_{nm})]\exp[iq(m\delta y)], \quad (A3)$$

where $\mathrm{sinc}\,\alpha = \sin\alpha/\alpha$. $a(p, q)$ is filtered by the function $F(p, q)$ and retransformed to give the image distribution $A(x', y')$. This process is indicated by Eq. (A4) :

$$A'(x',y') = \int_{-\infty}^{\infty}\int a(p,q)F(p,q)\exp[i(px'+qy')]\,dp\,dq. \quad (A4)$$

Because the filter function, F, is just a aperture of unity transmission, the limits of integration become the limits of the aperture. Usually, the aperture is chosen to select the first-order reconstruction in the p direction and the zero-order reconstruction in the q direction. The limits would then be $3\pi/\delta$ x and π/δ x for p and $\pm\pi/\delta$ y for q. Generally, the Kth-order reconstruction can be chosen in the p direction, so that Eq. (A4) may be rewritten as

$$A'(x',y') = \int_{(K-1/2)2\pi/\delta x}^{(K+1/2)2\pi/\delta x}\int_{-\pi/\delta y}^{\pi/\delta y} a(p,q)\exp[i(px'+qy')]\,dp\,dq. \quad (A5)$$

If Eq. (A3) is substituted into Eq. (A5) and both b_{nm} and c_{nm}, the rectangular aperture sides, are small so that the sinc functions both approximate to unity, Eq. (A5) may be solved to give

$$A'(x',y') = \text{constant}\exp\left(iK\frac{2\pi}{\delta x}\cdot x'\right)\sum_{(n,m)}^{N}\sum^{M} b_{nm}c_{nm} \times \mathrm{sinc}\left[\frac{\pi}{\delta x}(n\delta x+S_{nm}+x')\right]\exp\left(iKS_{nm}\frac{2\pi}{\delta x}\right) \times \mathrm{sinc}\frac{\pi}{\delta y}(m\delta y+y'). \quad (A6)$$

If the amplitude of the wavefront that the hologram is to produce is constant across the hologram plane, as it was for all the Lohmann and Paris type holograms described in this paper, b_{nm} and c_{nm} can be removed from the summation signs. If n and m are large, the summations approach integration and the sinc functions approximate to delta functions which sift the exponential function. The factor exp [iK(2 π/δ x). x'] represents a linear phase factor corresponding to the tilt of the wavefront from the axis of the system.

By choosing

$$S_{nm} = \phi(n\delta x+S_{nm}, m\delta y)(\delta x/2\pi),$$

where Ø(x', y') is the phase of the required wavefront, it is clear that the exponential function has the required form, and except for slight contributions from neighboring apertures, Eq. (A6) shows that the reconstruction is as required.

References

1. A. W. Lohmann and D. P. Paris, Appl. Opt. 6, 739 (1967).
2. W. H. Lee, Appl. Opt. 9, 639 (1970).
3. J. W. Goodman, Introduction to Fourier Optics (McGraw-Hill, New York, 1968).