192400-13-F

Final Report

GENERALIZED BEAMFORMING WITH MULTIPLE COMPUTER ORIGINATED HOLOGRAMS

October 1985-October 1988

J.N. CEDERQUIST
A.M. TAI
M.T. EISMANN
M.A. STUFF
J.R. FIENUP
Advanced Concepts Division

DECEMBER 1988

Sponsored by:
U.S. Army Research Office
P.C. Box 12211
Research Triangle Park, NC 27709-2211
Contract: DAAL03-86-K-0006

The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 192400-13-F	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Environmental Research Institute of Michigan 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U. S. Army Research Office
6c. ADDRESS (City, State, and ZIP Code) P.O. Box 8618 Ann Arbor, MI 48107-8618	7b. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U. S. Army Research Office 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAL03-86-K-0006
8c. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Generalized Beamforming with Multiple Computer Originated Holograms

12. PERSONAL AUTHOR(S)
Cederquist, J.N.; Tai, A.M.; Eismann, M.T.; Stuff, M.A.; and Fienup, J.R.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 10/85 TO 10/88	1988 December	vi + 97 = 103

16. SUPPLEMENTARY NOTATION
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
20	05	
20	06	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Computer generated holograms (CGH)
Holographic optical elements (HOE)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Generalized beamforming with optical systems using multiple computer-originated holographic optical elements (HOEs) was studied theoretically, through computer simulations, and through fabrication and experiments. A two HOE system was built for polar formatting of spotlight synthetic aperture radar data and was used with a spatial light modulator and a third phase-correcting HOE for real-time image formation. A two HOE system was designed using a novel iterative algorithm and successfully used for energy-efficient conversion of a Gaussian beam to a uniform far-field beam. A necessary condition was derived that a one-to-one coordinate transformation must satisfy to enable it to be performed with only two optical coordinate transformations, each using the stationary phase approximation. Because of the importance of computer-generated holograms to this work, a brief study was made of computer-generated hologram fabrication facilities.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL B.D. Guenther	22b. TELEPHONE (Include Area Code) 919-549-0641 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

PREFACE

The work reported here was performed by the Optical Science Laboratory of the Advanced Concepts Division, Environmental Research Institute of Michigan (ERIM). The work was sponsored by the U.S. Army Research Office (ARO) under Contract DAAL03-86-K-0006. At ARO, the technical monitors were Dr. Bob Guenther and Dr. David Skatrud.

This final technical report covers work performed from October 1985 to October 1988. The principal investigators at ERIM were Jack Cederquist and Anthony M. Tai. Major contributors to this work were Jack Cederquist, Anthony M. Tai, Michael T. Eismann, James R. Fienup, and Mark A. Stuff.

TABLE OF CONTENTS

1.0 PROBLEM STATEMENT

Optical systems employing a single holographic optical element (HOE) have proven useful for many applications. Consequently, many theoretical results and computer simulation techniques are available for the study and design of single HOE systems. One of the results from this work has been an understanding of the limitations, both in applications and in performance, of single HOE systems. Recently, a few researchers have studied systems using two or more HOEs. The initial effect has been to increase awareness of the potential new applications and performance improvements possible with such systems. Of particular interest have been applications involving the redirection of light for beamforming, optical coordinate transformation, and optical data interconnection. The HOEs used have generally been of the computer-generated type to give the needed flexibility.

With the above results as background, the goal of this research program was to explore the possibilities of multiple HOE systems when used for generalized beamforming. Generalized beamforming includes both the redistribution of light intensities and the more demanding one-to-one mapping of optical data coordinate transformation. Specific goals were to increase the theoretical understanding of generalized beam forming with multiple HOEs, to develop practical methods, whether theoretical or algorithmic, for determining the phase functions which specify the individual HOEs, to investigate applications for such systems, and to fabricate and experimentally investigate the performance of multiple HOE systems for generalized beamforming. It is understood that the required HOEs will need to be computer-originated, that is, at some stage of their fabrication, computer-generation techniques will be necessary to realize the complicated phase functions.

2.0 SUMMARY OF RESULTS

2.1 HOLOGRAPHIC POLAR FORMATTING

A two holographic optical element (HOE) system for polar formatting of spotlight mode synthetic aperture radar (SAR) data was designed, fabricated, and successfully tested. With the addition of a spatial light modulator, a third phase-compensating HOE, and a Fourier transform lens, the real-time polar formatting of SAR data and SAR image formation was experimentally demonstrated.

SAR image formation is an excellent application for coherent, analog optical processing. However, for spotlight mode SAR, polar formatting of the SAR data is required. One method for real-time optical polar formatting would be to use coordinate transform HOEs designed using the stationary phase approximation. Unfortunately, a single HOE cannot perform the polar coordinate transformation for the SAR data processing application. However, it was found that polar formatting can be performed optically by the successive use of two holographic coordinate transformations. The approach investigated and demonstrated uses a rectilinearly addressed SLM to input the SAR data and a two HOE system to optically polar format the data. By using a rectilinearly addressed SLM, this approach will benefit from the current intensive development of electronically addressed SLMs.

Briefly, the system operates as follows: The SAR data was input either by recording it on film or by using an electronically addressed, liquid crystal television spatial light modulator. In either case, the data was imaged onto the optical polar formatter consisting of two HOEs and two Fourier transform lenses. A third HOE is necessary to correct the output phase so that a final lens can perform a Fourier transformation to produce the desired SAR image. The phase functions for the HOEs were implemented as computer-generated holograms (CGHs) by encoding them as real, non-negative amplitude transmittance functions.

To improve their diffraction efficiency, the CGHs were copied onto a volume phase material and placed on a high carrier frequency. The result is a computer-originated holographic optical element (COHOE).

The multiple HOE system was used experimentally to polar format and process SAR data from multiple point scatterers with both the film and the liquid crystal television input. The experimental results demonstrated high SNR, low aberrations, and acceptable efficiency. The theoretical space-bandwidth product (SBWP) of the system (given by the square root of the CGH SBWP) was achieved experimentally. Further details are given in Appendix A.

2.2 HOLOGRAPHIC BEAMFORMER

A design method was developed for computing the phase functions of an energy-efficient system using two holographic optical elements (HOEs) for converting a Gaussian beam into a uniform beam with rectangular support in the far-field of the source. The method is based on a modification of the Gerchberg-Saxton algorithm which includes an x-y separability constraint on the phase of one of the HOEs. A beamforming system was fabricated using this method, and experimental results were obtained which support the design approach.

In applications such as laser radar and active coherent imaging, it is often desired to uniformly illuminate a rectangular area which is in the far-field of the source. We investigated the use of a holographic beamforming system to provide the energy redistribution into a near-field $\text{sinc}^2(x)\text{sinc}^2(y)$ intensity profile with the correct phase. In order to provide both the energy redistribution and phase correction, two separated HOEs are required. The beamformer concept which was developed therefore uses a first HOE to redistribute the energy of the input beam into the necessary near-field intensity profile, and then uses a second HOE to place the correct phase onto this beam. By relying solely on energy redistribution and phase correction, the

system is limited in efficiency only by the maximum achievable diffraction efficiency of the HOEs, which can be very high.

After determining a near-field amplitude distribution (a weighted sinc) which yields the desired far-field characteristics, the design of the beamforming system consists of computing the required phases for the two HOEs to produce the desired near-field beam. The first design method considered was adapted from that of Bryngdahl for producing geometric transformations with a single HOE. This stationary phase design for the first HOE was tested by computer simulation and found to be inadequate for accurately producing the zeros of the distribution which, in turn, resulted in unacceptable ringing and high sidelobes in the far-field.

Because of the Fourier transform relationship between the two planes, the design problem is equivalent to that of computing an object phase for a known object modulus (truncated Gaussian) and a known Fourier modulus (weighted sinc). This synthesis problem is equivalent to that of phase retrieval from a pair of intensity measurements, in which the complex fields in two domains related by a Fourier transform are reconstructed from incomplete information in each domain. A 2-D iterative synthesis of the first HOE phase using the Gerchberg-Saxton algorithm was therefore attempted. The algorithm was modified to enforce a separability constraint on the object phase. The stationary phase design was used as an initial phase estimate to start the Gerchberg-Saxton algorithm, and a specified number of iterations were performed. A 1-D slice of the current phase estimate was then extracted, and an outer product performed to produce a 2-D separable phase estimate. This was used to start another set of iterations, and the loop repeated until an acceptable solution was found by examining the simulated far-field intensity. After each separability constraint is imposed, the phase estimate improves, so that overall convergence to a good separable phase solution is obtained.

To verify the results of the computer simulation, the first HOE was produced as a COHOE using a CGH having the phase synthesized by the modified Gerchberg-Saxton algorithm. The second HOE was produced by interfering the output of the first element with a beam carrying the desired near-field phase of the sinc distribution. This desired phase (alternating π phase shifts) was encoded and recorded as a CGH. The experimental results agreed with the simulation results sufficiently well to support the design method. Further details are given in Appendix B.

2.3 REALIZABLE COORDINATE TRANSFORMATIONS

Optical processors often operate on data that are presented in two spatial dimensions. In some applications, the data need to be spatially rearranged at some point in the processing. These data rearrangements may be thought of as one-to-one coordinate transformations. Conventional optics cannot easily perform complicated transformations, so holographic optical elements (HOEs) are often used for coordinate transformations. Unfortunately, many coordinate transformations of interest in optical processing are not realizable with a single HOE. An example is polar formatting for image formation from spotlight mode SAR data. However, two HOEs in series can be used to perform this coordinate transformation. In this case, the desired transformation is decomposed into two serial transformations, each of which is realizable with a single HOE.

Given this example, two questions naturally arise: First, can all one-to-one coordinate transformations satisfying some modest assumptions be achieved by the use of multiple HOEs in series? Second, if so, what is the minimum number of HOEs required? It was shown that all coordinate transformations which are sufficiently differentiable can be decomposed into two transformations in series and therefore performed using two holographic elements. Although the method developed for designing the two required holographic elements is not

explicit for the general case, a useful necessary condition for the decomposition was found. A simple explicit solution based on separability of the necessary equations for the first of the two serial transformations was developed and is useful in some cases. Further details are given in Appendix C.

2.4 CGH FABRICATION FACILITIES

Because of the importance of computer-generated holograms (CGHs) to this research, a small part of the total effort was devoted to a consideration of existing CGH fabrication facilities and techniques. Although many facilities are currently fabricating CGHs of various types, all of these facilities suffer from one or more of the following difficiencies: (1) performance limitations (e.g., resolution, space-bandwidth product, or recording speed), (2) cost, (3) CGH encoding or materials restrictions (e.g., restriction to binary amplitude or phase, or restriction to recording fringes), (4) data format inefficiencies, and (5) lack of technical support in CGH design and optics in general. These deficiencies raise the issue of the need for a national facility for the fabrication of CGHs. Further details are given in Appendix D.

The SPIE Symposium in Los Angeles in January 1988 included conferences entitled Holographic Optics: Design and Applications and Computer-Generated Holography II. To benefit those attending these conferences, an evening panel discussion entitled "The Need for a National CGH Facility" was organized and held. The consensus of the panel members was that software difficulties (e.g., data format inefficiency and incompatibility) are the most important limitations on the development and use of CGHs and that, for the near term, electron beam writers are the preferred hardware for the fabrication of high performance CGHs.

3.0 LIST OF PUBLICATIONS

1. J.N. Cederquist and A.M. Tai, "Coordinate Transformation with Multiple Computer-Originated Holograms," J. Opt. Soc. Am. A 4, P121-P122 (1986).

2. J.N. Cederquist, A.M. Tai, and M.T. Eismann, "Design and Applications of Multiple Hologram Systems," in Holographic Optics: Design and Applications, Proceedings of the SPIE, Volume 883, I. Cindrich, ed. (Los Angeles, January 1988), pp. 177-184.

3. J.N. Cederquist, J.R. Fienup, and A.M. Tai, "CGH Fabrication Techniques and Facilities," in Computer-Generated Holography II, Proceedings of the SPIE, Volume 884, S.H. Lee, ed. (Los Angeles, January 1988), pp. 40-45.

4. M.T. Eismann, A.M. Tai, and J.N. Cederquist, "Holographic Beamformer Designed by an Iterative Technique," in Holographic Optics: Optically and Computer Generated, Proceedings of the SPIE, Volume 1052, I. Cindrich and S.H. Lee, eds. (Los Angeles, January 1989).

5. M.T. Eismann, A.M. Tai, and J.N. Cederquist, "Iterative Design of a Holographic Beamformer," Appl. Opt. (accepted 1989).

6. J.N. Cederquist, A.M. Tai, and M.T. Eismann, "Holographic Polar Formatting and Real-Time Optical Processing of Synthetic Aperture Radar Data," Appl. Opt. (submitted 1989).

7. M.A. Stuff and J.N. Cederquist, "Coordinate Transformations Realizable with Multiple Holographic Optical Elements," J. Opt. Soc. Am. A (submitted 1989).

8. M.T. Eismann, A.M. Tai, and J.N. Cederquist, "Synthesis of a Holographic Beamformer Using a Separability Constraint," in Digest of Topical Meeting on Signal Recovery and Synthesis III (Optical Society of America, Washington, D.C., 1989), accepted for publication.

4.0 PARTICIPATING SCIENTIFIC PERSONNEL

Participating scientific personnel were Jack Cederquist, Michael T. Eismann, James R. Fienup, Mark A. Stuff, and Anthony M. Tai. No advanced degrees were earned by participating personnel while they were employed on this project.

APPENDIX A

HOLOGRAPHIC POLAR FORMATTING AND REAL-TIME OPTICAL PROCESSING OF SYNTHETIC APERTURE RADAR DATA

(submitted to Applied Optics, 1989)

Abstract

A two holographic optical element (HOE) system for polar formatting of spotlight mode synthetic aperture radar (SAR) data was designed, fabricated, and successfully tested. With the addition of a spatial light modulator, a third phase-compensating HOE, and a Fourier transform lens, the real-time polar formatting of SAR data and SAR image formation was experimentally demonstrated.

I. Introduction

Synthetic aperture radar (SAR) image formation is an excellent application for coherent, analog optical processing. However, for spotlight mode SAR, polar formatting of the SAR data is required before the image can be formed via 2-D Fourier transformation. Previous work has investigated the use of mechanical rotation and, in one case, spatial light modulators with electron beam writing to perform polar formatting.[1,2,3] We view mechanical motion as undesirable and wish to take advantage of the current intensive development of SLMs with rectilinearly placed pixels and electronic addressing. We therefore seek optical methods for polar formatting. Single holographic optical elements (HOEs) to perform coordinate transformations were first designed using the stationary phase approximation by Bryngdahl[4,5] in 1974. Unfortunately, a single HOE cannot perform the polar coordinate transformation for the SAR data processing application. However, the use of two coordinate transform HOEs in series has been investigated[6,7] to overcome the limitations of single HOE systems.

In this paper, the design, fabrication, and testing of a three HOE system for the real-time polar formatting of, and image formation from, spotlight SAR data is described. In Sec. II, background information on SAR data processing is given. The design of the required HOEs for polar coordinate transformation is described in Sec. III. In Sec. IV, the fabrication of the HOEs and the construction of the optical system is discussed. Experimental results for both polar formatting and real-time SAR image formation are given in Sec. V. Conclusions and suggestions for further research are presented in Sec. VI.

II. Synthetic Aperture Radar Data Processing

In synthetic aperture radar, an appropriate waveform is transmitted and the return signal scattered from a scene is processed to obtain range and Doppler information about the scene.[8] Range resolution is achieved in inverse proportion to the frequency bandwidth of the transmitted signal. The Doppler information corresponds to azimuthal, or cross-range, spatial information. Azimuthal resolution is achieved in inverse proportion to the time interval over which return signals are collected as the SAR flies by the scene (i.e., in inverse proportion to the extent of the synthetic aperture). Together, the range and Doppler information give a 2-D image of the scene. For a SAR which transmits a sequence of linear FM (chirp) pulses, the required data processing consists of four steps. The first mixes the return signals with a reference signal to demodulate the signal to an intermediate frequency. The return signal from each pulse thus gives a 1-D data set. The second step forms a 2-D data set by placing the 1-D data from each of many pulses side by side. In the important cases where the scene consists of a rotating object or the SAR is used in the spotlight mode, Fig. 1(a) shows the resulting data for a single point scatterer. In these cases, the third step polar formats the data as shown in Fig. 1(b) to compensate for the continuously changing viewing angle during data collection. The fourth step performs a 2-D Fourier transformation of the data to produce an image of the scene.

The necessary polar formatting operation is a difficult data processing step. In any sampled data processor (such as an electronic digital computer), polar formatting requires interpolation which often introduces error and is computationally intensive, about twice that of the FFT operation. Since a spatial light modulator (SLM) must be used to input the data to an optical processor, one possibility that has been investigated is to polar format the data by writing it onto a rotating SLM with an electron beam.[1] An optically addressed SLM could also be used, but it must have a storage time as long as the time taken to write all the 2-D data coupled with a short erasure time to allow real-time input of the next 2-D data set. Other possibilities such as rotating the output detector of the data processor and processing each 1-D data set separately have also been investigated.[2,3]

The approach investigated and demonstrated in this paper, however, uses a rectilinearly addressed 2-D SLM to input the data in the form shown in Fig. 1(a) and a two HOE system to optically polar format the data as shown in Fig. 1(b). It does not suffer from the frame rate limitations of optically or electron beam addressed SLMs and does not involve mechanical motion. By using a rectilinearly addressed SLM, this approach will benefit from the current intensive development of electronically addressed SLMs, such as pixel addressed liquid crystal light valves[9] and the Si-PLZT device.[10] These devices are compact, have controllable storage and erasure rates, and are relatively fast. Together with optical polar formatting and optical processing, they have great potential for application to real-time processing of SAR data.

III. Design of Polar Coordinate Transformation HOEs

As an introduction to the design method, consider the optical system shown in Fig. 2 where the input field at P_1 in the front focal plane of a lens L is uniphase and has an amplitude distribution $a_1(x, y)$. A HOE with phase $\phi(x, y)$ is placed in plane P_1. The complex

amplitude distribution $a_2(u, v)$ found in the back focal plane P_2 of the lens is given approximately by the Fourier transform:

$$a_2(u,v) = -\frac{i}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_1(x,y)\,\exp\left[i\phi(x,y) - \frac{i2\pi}{\lambda f}(xu + yv)\right] dxdy \qquad (1)$$

where λ is the wavelength and f is the lens focal length. Following Bryngdahl,[4,5] this integral can be approximated by the method of stationary phase[11] to show that a point (x, y) in the input is mapped to a point (u, v) in the output where

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{\lambda f}{2\pi}\frac{\partial\phi(x,y)}{\partial x} \\ \frac{\lambda f}{2\pi}\frac{\partial\phi(x,y)}{\partial y} \end{bmatrix} . \qquad (2)$$

If a desired coordinate transformation is

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} g(x, y) \\ h(x, y) \end{bmatrix} , \qquad (3)$$

then the required HOE phase $\phi(x,y)$ is found by solving the set of partial differential equations:

$$\begin{aligned} \frac{\lambda f}{2\pi}\frac{\partial\phi}{\partial x} &= g(x, y) \\ \frac{\lambda f}{2\pi}\frac{\partial\phi}{\partial y} &= h(x, y) \quad . \end{aligned} \qquad (4)$$

Equation (4) has a solution if and only if[12]

$$\frac{\partial g(x, y)}{\partial y} = \frac{\partial h(x, y)}{\partial x} \quad . \qquad (5)$$

Unfortunately, for many coordinate transformations of interest, Eq. (5) is not satisfied. Specifically, for polar formatting, the desired coordinate transformation is

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \cos y \\ x \sin y \end{bmatrix} \tag{6}$$

and the transformation functions are

$$\begin{aligned} g(x, y) &= x \cos y \\ h(x, y) &= x \sin y \quad . \end{aligned} \tag{7}$$

The partial derivatives,

$$\begin{aligned} \frac{\partial g(x, y)}{\partial y} &= -x \sin y \\ \frac{\partial h(x, y)}{\partial x} &= \sin y \quad , \end{aligned} \tag{8}$$

do not satisfy Eq. (5) and, hence, the transformation is not realizable with a single coordinate transform HOE.

However, the polar formatting transformation can be separated into two cascaded transformations, each of which does satisfy Eq. (5).[13,14] The transformations are:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \ln x \\ y \end{bmatrix} \tag{9}$$

and

$$\begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} \exp u \cos v \\ -\exp u \sin v \end{bmatrix} \tag{10}$$

where s and t are Cartesian coordinates. Eq. (5) is satisfied for both transformations since

$$\frac{\partial(\ln x)}{\partial y} = 0$$
$$\frac{\partial y}{\partial x} = 0 \qquad (11)$$

and

$$\frac{\partial(\exp u \cos v)}{\partial v} = -\exp u \sin v$$
$$\frac{\partial(-\exp u \sin v)}{\partial u} = -\exp u \sin v \quad . \qquad (12)$$

Therefore, both transformations are realizable optically. Note that the minus sign in Eq. (10) is necessary to make the second transformation realizable. It corresponds to an inversion in one of the output coordinates. Similar use of a minus sign in Eq. (6) does not make that transformation realizable, however.

Polar formatting can therefore be performed optically by the successive use of two holographic coordinate transformations as shown in Fig. 3. HOEs H1 and H2 perform coordinate transformations which, when cascaded, produce the polar coordinate transformation. HOEs H1' and H2' correct the phase at the output of each transformation such that the output phase is uniform for a uniphase input. It should be noted that the phase correction for the output of the first coordinate transformation and the phase for the second transformation can be combined and implemented as a single HOE, so the complete system, including the final phase correction, requires only three HOEs.

The next step is to solve for the phase functions of the two coordinate transformation HOEs. To do this, we need to define coordinate systems in each of the three optical planes in Fig. 3 and to establish the relationships of these systems to coordinate systems for the SAR data. As explained in Sec. II, after the second step in the data processing, the input SAR data lie in a rectangular region from r_1 to r_2 and $-\theta_1$ to θ_1 in a coordinate system in which r and θ are orthogonal [see Fig. 4(a)]. This SAR data will be loaded into an electronically addressed SLM with rectilinear pixel placement such that the return signal data from each pulse will reside in one row of SLM pixels. In the optical plane located at the HOE H1, this data can be placed in a region defined to be from $-x_1$ to x_1 and $-y_1$ to y_1 [again, see Fig. 4(a)]. The relationship between these two coordinate systems (one optical and one for the input SAR data) is:

$$x = \frac{2x_1}{r_2 - r_1}\left[r - \frac{r_2 + r_1}{2}\right]$$
$$y = \frac{y_1\theta}{\theta_1} \tag{13}$$

and

$$r = \frac{x}{x_1}\left[\frac{r_2 - r_1}{2}\right] + \frac{r_2 + r_1}{2}$$
$$\theta = \frac{\theta_1 y}{y_1} \quad . \tag{14}$$

The first HOE H1 will transform [see Eq. (9)] the SAR data into a region from $\ln r_1$ to $\ln r_2$ and $-\theta_1$ to θ_1 in a coordinate system in which $\ln r$ and θ are orthogonal. In the optical system, this data can be placed in a region from $-u_1$ to u_1 and $-v_1$ to v_1 [see Fig. 4(b)]. The relationship between these data and optical coordinate systems at

the output plane of the first coordinate transformation (at the plane of the HOE incorporating H1' and H2) is:

$$u = \frac{2u_1}{\ln (r_2/r_1)} \left[\ln r - \frac{\ln (r_2 r_1)}{2}\right]$$
$$v = \frac{v_1 \theta}{\theta_1} \tag{15}$$

and

$$r = \sqrt{r_2 r_1}\, \exp\left[\frac{u \ln (r_2/r_1)}{2u_1}\right]$$
$$\theta = \frac{\theta_1 v}{v_1} \quad . \tag{16}$$

Combining Eqs. (14) and (15) gives the required form of the first optical coordinate transformation:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} \tag{17a}$$

where

$$u = \frac{2u_1}{\ln (r_2/r_1)} \left\{\ln \left[\frac{x}{x_1}\left[\frac{r_2 - r_1}{2}\right] + \frac{r_2 + r_1}{2}\right] - \frac{\ln (r_2 r_1)}{2}\right\}$$
$$v = \frac{v_1 y}{y_1} \quad . \tag{17b}$$

Note that Eq. (17) is simply a more general form of Eq. (9) which still satisfies Eq. (5). Solving the partial differential Eq. (4) by using the expressions in Eq. (17) and integration gives the required holographic phase function for the HOE H1:

$$\phi_1(x, y) = \frac{2\pi}{\lambda f}\left[\frac{2u_1}{\ln (r_2/r_1)}\left[\frac{r_2 - r_1}{2x_1}\right]^{-1}\left[\frac{x}{x_1}\left[\frac{r_2 - r_1}{2}\right] + \frac{r_2 + r_1}{2}\right]\right.$$

$$\left.\cdot\left\{\ln\left[\frac{x}{x_1}\left[\frac{r_2 - r_1}{2}\right] + \frac{r_2 + r_1}{2}\right] - \frac{\ln (r_2 r_1)}{2} - 1\right\} + \frac{v_1 y^2}{2y_1}\right] . \tag{18}$$

The second transformation [see Eq. (10)] will put the SAR data into a region from r_1 to r_2 and $-\theta_1$ to θ_1 in a coordinate system in which r and θ are polar coordinates. In the optical system, the data will be placed in a region from $-s_1$ to s_1 on the s-axis and $-t_1$ to t_1 on the arc corresponding to $r = (r_1 + r_2)/2$ [see Fig. 4(c)]. The relationship between these data and optical coordinate systems is:

$$s = \frac{2s_1}{r_2 - r_1}\left[r \cos\theta - \frac{r_2 + r_1}{2}\right]$$

$$t = \frac{-2t_1 r \sin\theta}{\left(r_2 + r_1\right) \sin\theta_1} . \tag{19}$$

It is convenient to require that

$$\frac{t_1}{s_1} = \frac{\left(r_2 + r_1\right) \sin\theta_1}{r_2 - r_1} \tag{20}$$

so that SAR image produced is scaled identically in both coordinates.

Combining Eqs. (16) and (19) gives the required second optical coordinate transformation:

$$\begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} \tag{21a}$$

where

$$\begin{aligned} s &= \frac{2s_1}{r_2 - r_1}\left[\sqrt{r_2 r_1}\,\exp\left[\frac{u \ln (r_2/r_1)}{2u_1}\right]\cos\left[\frac{\theta_1 v}{v_1}\right] - \frac{r_2 + r_1}{2}\right] \\ t &= \frac{-2s_1}{r_2 - r_1}\,\sqrt{r_2 r_1}\,\exp\left[\frac{u \ln (r_2/r_1)}{2u_1}\right]\sin\left[\frac{\theta_1 v}{v_1}\right] \quad . \end{aligned} \tag{21b}$$

Note that Eq. (21) is simply a more general form of Eq. (10) which still satisfies Eq. (5) (when written in terms of u-v coordinates) provided that

$$\frac{u_1}{v_1} = \frac{\ln (r_2/r_1)}{2\theta_1} \quad . \tag{22}$$

Eq. (22) expresses the requirement that the SAR data be scaled identically in both coordinates in the u-v plane. In these experiments, u_1 was chosen equal to v_1 and $r_2/r_1 = 5/3$, so $\theta_1 = 14.6°$. Solving the partial differential Eq. (4) (in u-v coordinates) by using the expressions given in Eqs. (21) and (22) and integration gives the required holographic phase function for the HOE H2:

$$\phi_2(u, v) = \frac{2\pi}{\lambda f}\left[\frac{2s_1}{r_2 - r_1}\right]\left[\frac{2u_1\sqrt{r_2 r_1}}{\ln (r_2/r_1)}\exp\left[\frac{u \ln (r_2/r_1)}{2u_1}\right]\cos\left[\frac{v \ln (r_2/r_1)}{2u_1}\right] - \left[\frac{r_2 + r_1}{2}\right]u\right] \quad . \tag{23}$$

IV. HOE Fabrication and Optical Setup

The phase functions of Eqs. (18) and (23) were implemented as computer-generated holograms (CGHs) by encoding them as real, non-negative amplitude transmittance functions t(x, y) of the form:

$$t(x, y) = b + a \cos [\phi(x, y) + \alpha x + \beta y] \qquad (24)$$

where a and b are the amplitude and bias of the cosinusoidal fringe pattern. For these experiments, a = 0.4 and b = 0.5. The term of the form $\alpha x + \beta y$ was added to each phase to give a spatial carrier frequency. α and β were chosen so that, in a Fourier plane with respect to the CGH, the desired first diffraction order (the coordinate transform region) did not overlap with the zero order. The amplitude transmittance functions were recorded on high resolution film using a rotating drum laser beam recorder. A CGH size of $\sim(1\ \text{cm})^2$ was used with a 5 μm pixel size to produce a CGH space-bandwidth product of 2048 by 2048. The film was mounted between 1/4 inch microflat plates using a UV-curing optical cement in order to eliminate any phase errors due to thickness variations in the film.

These CGHs are thin absorption holograms and are therefore too low in diffraction efficiency (~1%) to be used in a multiple HOE system. To improve their diffraction efficiency, the CGHs were copied onto a volume phase material and placed on a high carrier frequency. The result is a computer-originated holographic optical element (COHOE).[15,16] The optical setup for producing a COHOE is shown in Fig. 5. A laser beam is split into two parts. One part is collimated and used to read out the CGH. The desired diffracted order of the CGH is passed by a Fourier plane spatial filter and imaged to the COHOE recording plane. There it is interfered with a reference plane wave at an offset angle suitable for forming a volume hologram. When illuminated by the reference plane wave, the phase distribution encoded in the CGH is reconstructed with high efficiency by the COHOE.

As shown in Fig. 3, two coordinate transform HOEs and a third phase-compensating HOE are used in the polar formatting system. The first HOE is a COHOE of the phase function for the first coordinate transformation [Eq. (18)]. The second HOE corrects the output phase of the first coordinate transformation and produces the phase for the second transformation [Eq. (23)]. This second HOE was fabricated by interfering the output wavefront of the first HOE (for a plane wave input) with the desired phase produced by the second CGH. It is therefore a COHOE made with a nonplanar reference wavefront. The third HOE corrects the output phase of the second coordinate transformation. It was made by interfering the output wavefront from the the second HOE with a uniform reference beam. All three HOEs were made on Kodak 649F spectroscopic plates using a HeNe laser (λ = 632.8 nm). The HOEs were processed by the silver halide (sensitized) gelatin process[17] and then covered with microflats to avoid degradation of diffraction efficiency when exposed to humidity. The resulting diffraction efficiencies for the three HOEs were 37%, 14%, and 36%, respectively. Because the second and third HOEs are made by interfering beams of non-uniform intensity, the recording must be confined to the linear exposure region to avoid intermodulation noise. This, of course, decreases the maximum diffraction efficiency that can be achieved.

The experimental multiple HOE system for polar formatting and SAR image formation is shown in Fig. 6. The SAR data was input either by recording it on film and placing it in a liquid gate or by an electronically addressed, liquid crystal television spatial light modulator. In either case, the data was coherently illuminated and imaged onto the first HOE (H1). Fourier transformation with a lens then produced the log transformed data in the first diffracted order where it illuminated the second HOE (H1' + H2). The second HOE and a second lens then performed the remainder of the polar formatting operation. The third HOE (H2') corrected the output phase, and a final lens was used to perform a Fourier transformation and produce the desired SAR image.

V. Experimental Results

The multiple HOE system was used experimentally to polar format and form images from SAR data of multiple point scatterers. Both film and a liquid crystal television (LCTV) were used as input devices. SAR data was written on the input devices by computer generation techniques and the output image was magnified and detected by a COHU CCD camera. The overall efficiency of the system was about 1.6%; however, the output intensity using a 50 mW laser source was more than sufficient for the CCD detector.

Figure 7 shows the operation of the system at each stage of the formatting and processing of SAR data for a single point scatterer. For this experiment, in which the input was recorded on film and placed in a liquid gate, the ability of the multiple HOE system to properly format the data and provide both azimuth and range compression can clearly be seen.

To measure system performance quantitatively, SAR data representing eight point scatterers in a single line was generated and recorded on film. Because the data was recorded on a bias [as in Eq. (24)], an intense DC reference point and high sidelobes along the x' and y' axes are present in the output image plane. The eight point test image was therefore intentionally placed along a line at 45° to these axes to reduce the effect of these sidelobes. In Fig. 8(a), the image is shown for an experiment in which the data was processed without polar formatting (i.e., optical Fourier transformation alone). The output of the multiple HOE system is shown in Fig. 8(b) with a line scan through the points shown in Fig. 8(c). Detector noise was reduced by temporal averaging. As expected, the image obtained with polar formatting shows far superior azimuth compression.

Several performance characteristics of the multiple HOE system can be determined from the line scan in Fig. 8(c). First, the space-bandwidth product (SBWP) can be measured. Theoretically, the SBWP is given approximately by the square root of the hologram SBWP.[6] For the $(1\ cm)^2$ hologram used with 2048 by 2048 pixels, a SBWP of approximately 45 by 45 would be predicted. The SAR test data was computed such that the image points are spaced by 10 diffraction-limited impulse response widths. The system modulation transfer function (MTF), specified by the envelope of the line scan, falls to -3 dB of the peak value at 62 diffraction-limited impulse response widths along the diagonal from the DC reference point, giving an experimental SBWP of 44 by 44.

Second, to evaluate the system phase errors, the impulse response widths of the images of the eight points were measured and found to be approximately 1.25 times the diffraction-limited width predicted for an ideal system. Not all of this increase in width is due to phase errors, however. Aperture weighting due to the MTF of the CGH recorder is partially responsible for this increase. Another part of the increase is due to the coordinate transformation itself. This aperture weighting is the result of the change in the differential element of area from $drd\theta$ (rectilinear r-θ coordinates) to $rdrd\theta$ (polar coordinates)[18] and can also be predicted by evaluating Eq. (1) using the method of stationary phase.[11] Both of these effects can be corrected by appropriate inverse weighting in the input plane of the polar formatting system, but this correction was not implemented in these experiments.

The final system performance characteristic which can be computed is the signal-to-noise ratio (SNR). The system SNR is signal dependent for two reasons. First, as the number of points in the image increases, the available system dynamic range is shared by more points and the SNR decreases. Second, SNR variations over the image roughly follow the system MTF and therefore the SNR decreases as the location of the point moves farther from the DC reference point (i.e., SNR

decreases as carrier frequency increases). For the point of the eight point image which was nearest to the DC point, a SNR of 34 dB was measured.

Real-time operation of the SAR processor was demonstrated by using a low cost Radio Shack Model 16-151 LCD Pocketvision TV as an input spatial light modulator. The LCTV was modified at ERIM for this application by removing the polarizers and cementing the device between optical flats to obtain higher optical quality.[19] In order to fully utilize the SBWP of the LCTV, the input imaging optics were altered to provide 3X demagnification of the LCTV onto the first HOE H1. The LCTV was driven by the video output of a Data Translation DT2861 frame grabber in an IBM PC-AT computer. This allowed SAR test data to be placed on a standard video signal and sent to the LCTV. A polarizer oriented at 90° to the input laser polarization was placed just before the detector.

The output of this real-time, multiple HOE, SAR polar formatter and processor for an input representing SAR data from three point scatterers is shown in Fig. 9. The input data is shown in Fig. 9(a) and the output image in Fig. 9(b). Note that Fig. 9(b) is magnified with respect to Fig. 8(b). Because the SBWP and dynamic range of the LCTV are less than that of the film input, the processor performance is reduced somewhat. The smaller SBWP of the LCTV can be seen by comparing the MTF of the real-time processor to that of the film-input processor (see Fig. 10). The real-time processor MTF falls off much more rapidly, limiting the processor SBWP to approximately 20 by 20. The limited dynamic range reduces the signal-to-bias ratio and lowers the SNR. For the image of the test point nearest the DC reference point, a SNR of approximately 18 dB was measured.

VI. Conclusion

A multiple HOE system was designed, constructed, and shown experimentally to perform polar formatting for the real-time optical processing of SAR data. An electronically addressed SLM was used for data input. The experimental system demonstrated high SNR, low aberrations, acceptable efficiency, and a SBWP at the output image of 44 by 44. In order to increase the processor SBWP, the SBWP of the CGHs used must be increased by reducing the pixel size and increasing the CGH size. Using a $(6.25\ cm)^2$ CGH with a 1 μm spot size, a 250 x 250 coordinate transform SBWP could be achieved which would also match the performance of current real-time input and output devices.

The promising results obtained in this research provide a useful foundation for future development of compact, real-time spotlight SAR processors. Further work in this area should first be aimed at improving the system performance by increasing the SBWP and reducing the scattering and intermodulation noise in the COHOEs. The next step would be the design of a more compact system with the Fourier transform lenses incorporated into the coordinate transform HOEs.

This research was supported by the U.S. Army Research Office. Portions of this research were presented at the Optical Society of America Annual Meeting in Seattle, Washington, October 20-24, 1986[13] and at the Society of Photo-Optical Instrumentation Engineers Meeting on Holographic Optics: Design and Applications in Los Angeles, California, January 13-14, 1988.[14]

References

1. G.D. Currie, I. Cindrich, and C.D. Leonard, "The ERIM TOPR in Optical Data Processing," Proc. SPIE 83, 8-13 (1976).

2. C.C. Aleksoff, I.J. LaHaie, and A.M. Tai, 453 Optical-Hybrid Backprojection Processing," Proc. SPIE 422, 89-95 (1983).

3. M. Haney and D. Psaltis, "Real-Time Programmable Acoustooptic Synthetic Aperture Radar Processor," Appl. Opt. 27, 1786-1796 (1988).

4. O. Bryngdahl, "Optical Map Transformations," Opt. Commun. 10, 164-168 (1974).

5. O. Bryngdahl, "Geometrical Transformations in Optics," J. Opt. Soc. Am. 64, 1092-1099 (1974).

6. J. Cederquist and A.M. Tai, "Computer-Generated Holograms for Geometric Transformations," Appl. Opt. 23, 3099-3104 (1984).

7. W.J. Hossack, A.M. Darling, and A. Dahdouh, "Coordinate Transformations with Multiple Computer-Generated Optical Elements," J. Modern Optics 34, 1235-1250 (1987).

8. J.L. Walker, "Range-Doppler Imaging of Rotating Objects," IEEE Trans. AES-16, 23-51 (1980).

9. M.S. Welkowsky, U. Efron, W. Byles, and N.W. Goodwin, "Status of the Hughes Charge-Coupled-Device-Addressed Liquid Crystal Light Valve," Opt. Eng. 26, 414-417 (1987).

10. S.H. Lee, S.C. Esener, M.A. Title, and T.J. Drabik, "Two-Dimensional Silicon/PLZT Spatial Light Modulators: Design Considerations and Technology," Opt. Eng. 25, 250-260 (1986).

11. M. Born and E. Wolf, Principles of Optics (Pergamon, New York, 1975), pp. 752-754.

12. W. Kaplan, Advanced Calculus (Addison-Wesley, Reading, Mass., 1959), pp. 243-249.

13. J.N. Cederquist and A.M. Tai, "Coordinate Transformation with Multiple Computer-Originated Holograms," J. Opt. Soc. Am. A 4, P121-P122 (1986).

14. J.N. Cederquist, A.M. Tai, and M.T. Eismann, "Design and Applications of Multiple Hologram Systems," Proc. SPIE 883, 177-184 (1988).

15. R.C. Fairchild and J.R. Fienup, "Computer-Originated Aspheric Holographic Optical Elements," Opt. Eng. 21, 133-140 (1982).

16. H. Bartelt and S.K. Case, "High Efficiency Hybrid Computer Generated Holograms," Appl. Opt. 21, 2886-2890 (1982).

17. D.K. Angell, "Improved Diffraction Efficiency of Silver Halide (Sensitized) Gelatin," Appl. Opt. 26, 4692-4702 (1987).

18. G. Arfken, Mathematical Methods for Physicists (Academic, New York, 1966), pp. 65-66.

19. A.M. Tai, "Low-cost LCD Spatial Light Modulator with High Optical Quality," Appl. Opt. 25, 1380-1382 (1986).

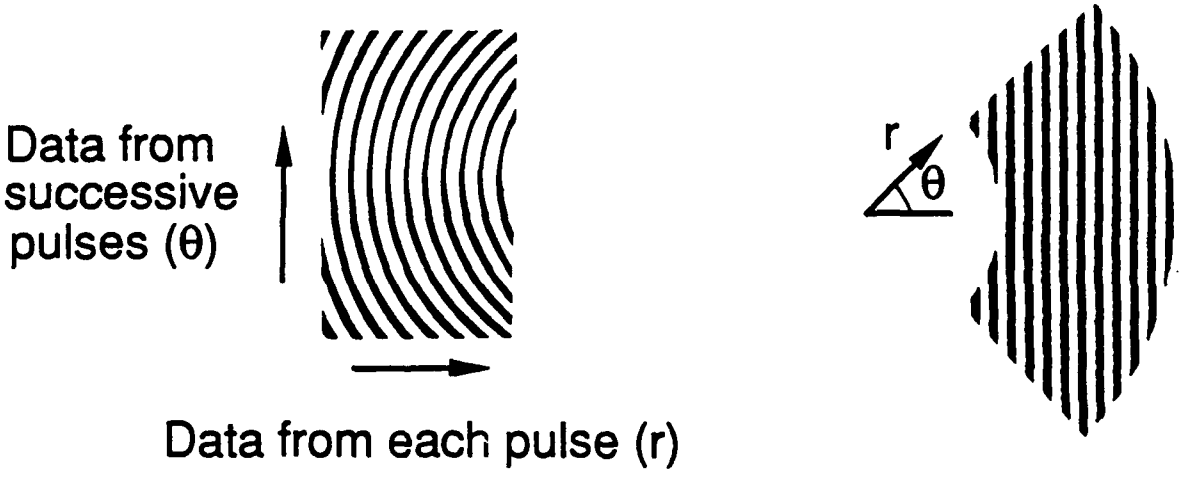


Fig. 1. SAR data for a point scatterer.

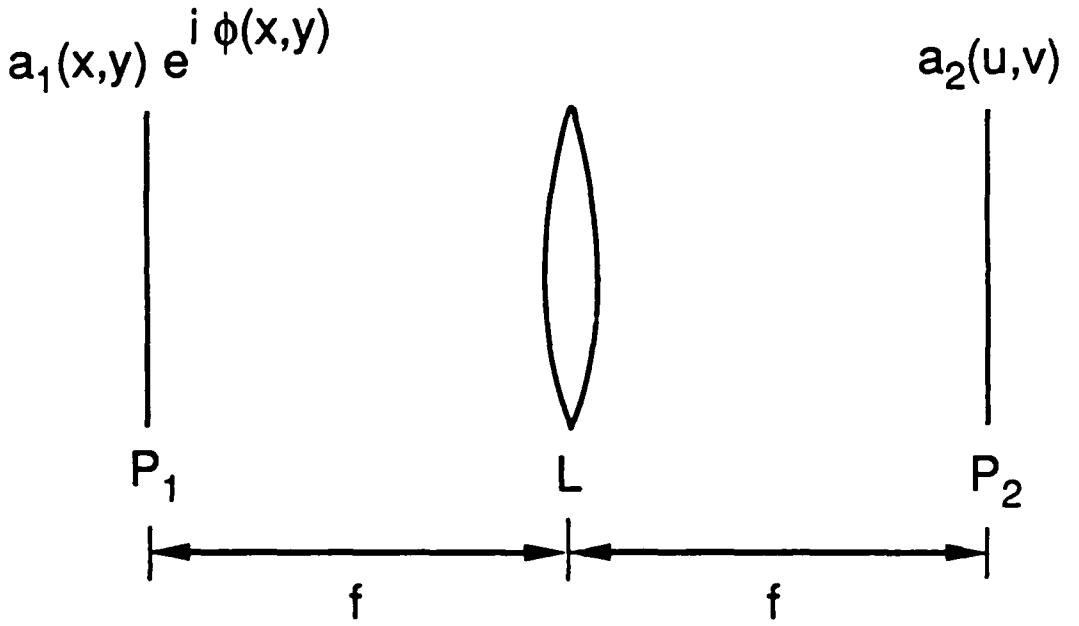


Fig. 2. Optical system for coordinate transformation. The lens, L, has focal length, f, and Fourier transforms plane P_1 to plane P_2.

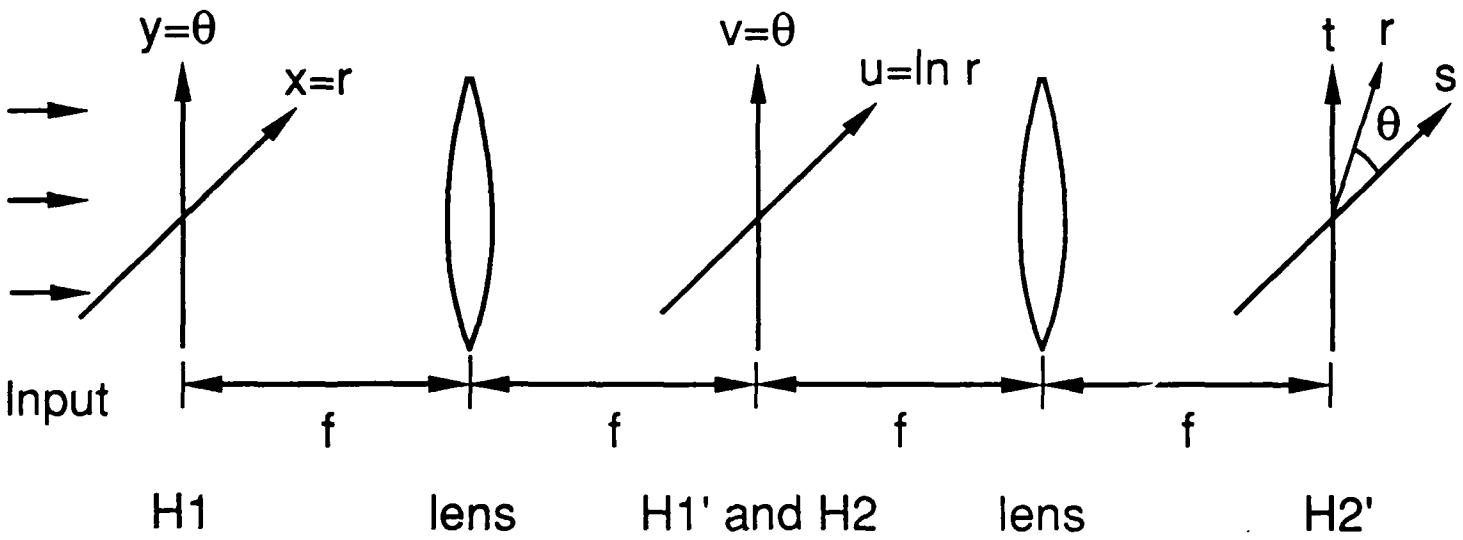


Fig. 3. Optical system for polar formatting.

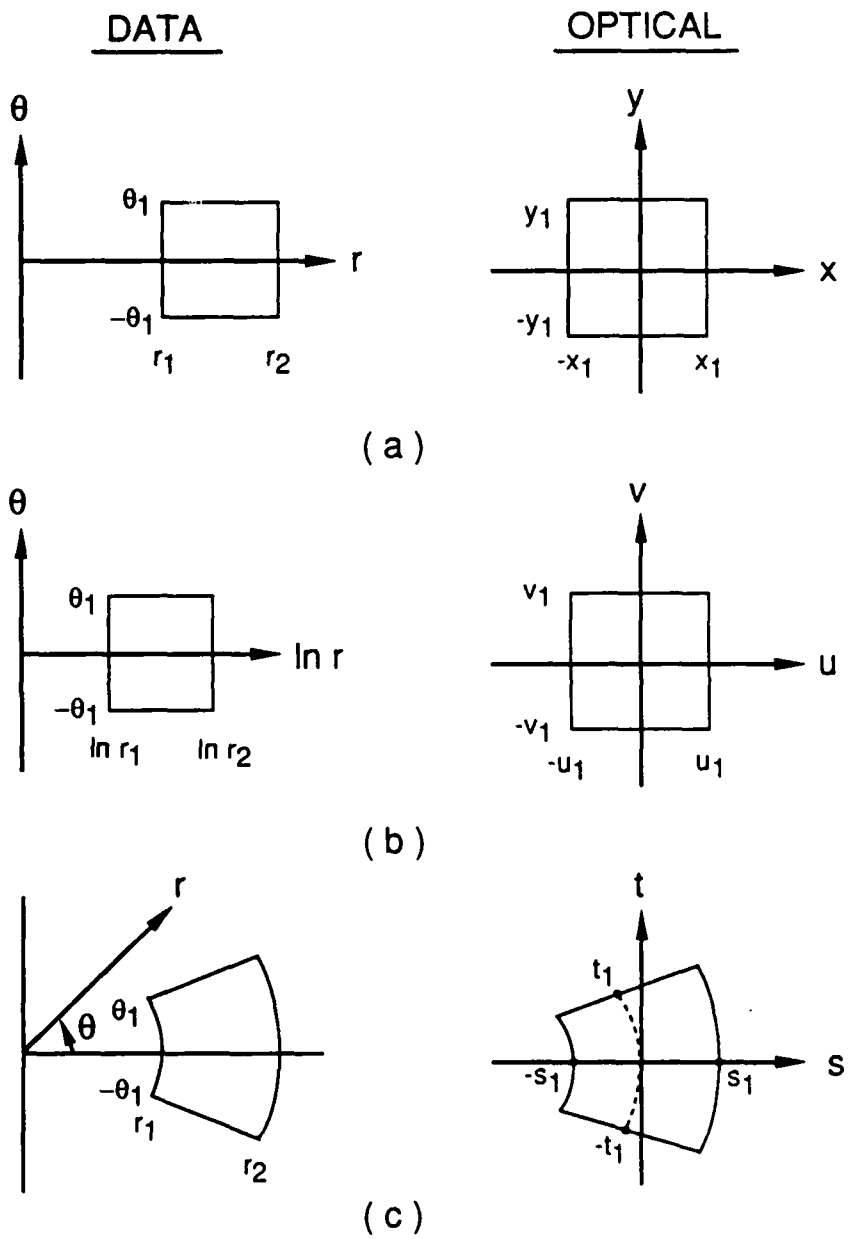


Fig. 4. SAR data and optical coordinate systems. (a) Input, (b) intermediate, and (c) output planes of polar formatting system.

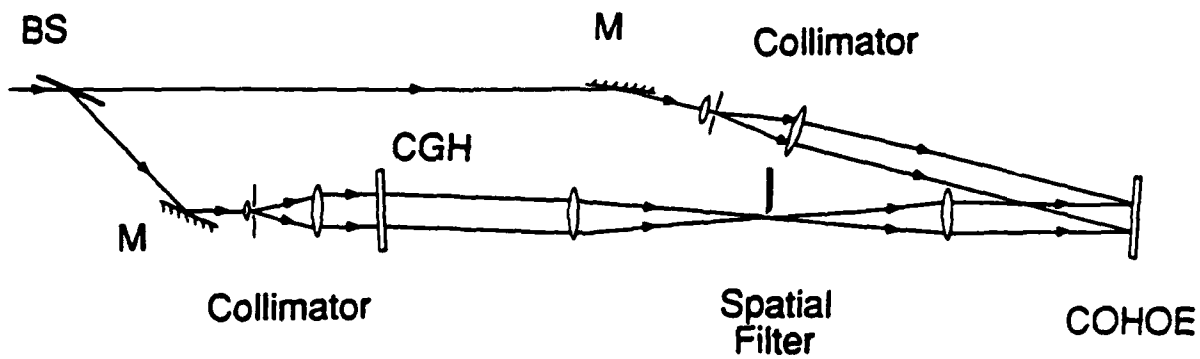


Fig. 5. COHOE recording method.

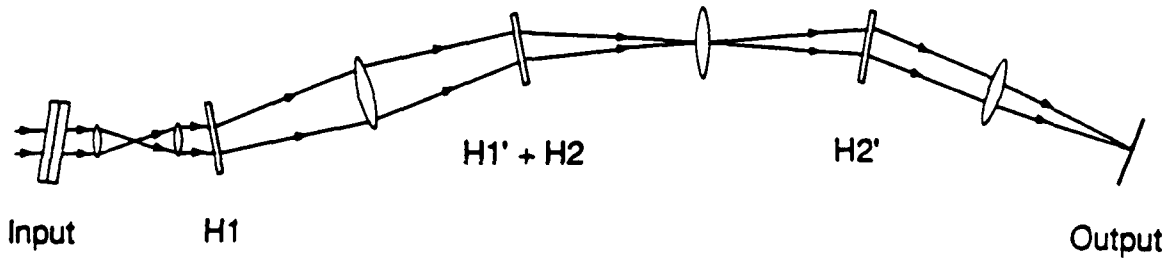


Fig. 6. Experimental multiple HOE system for polar formatting and SAR image formation.

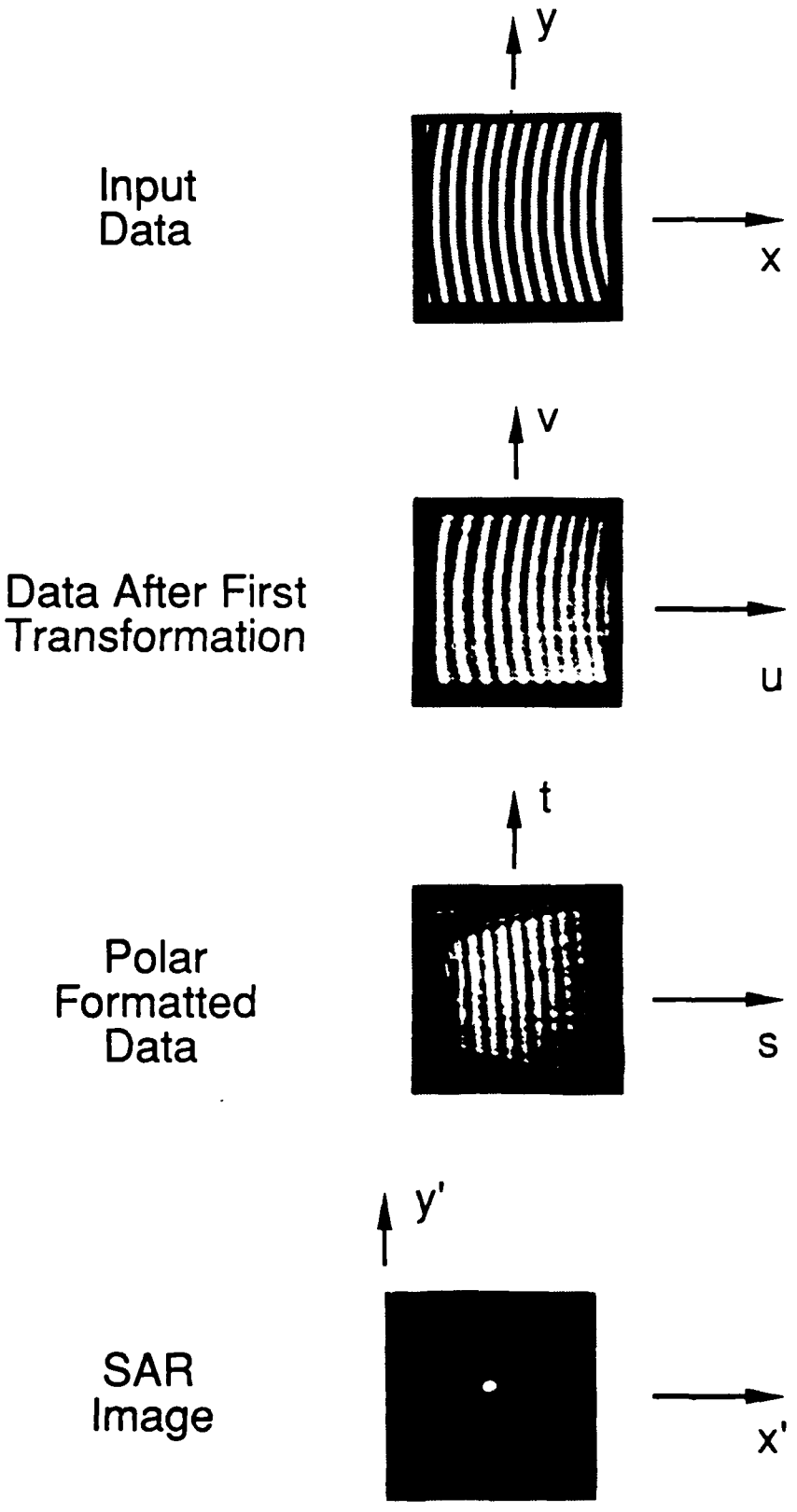


Fig. 7. Experimental results showing polar formatting and SAR image formation.

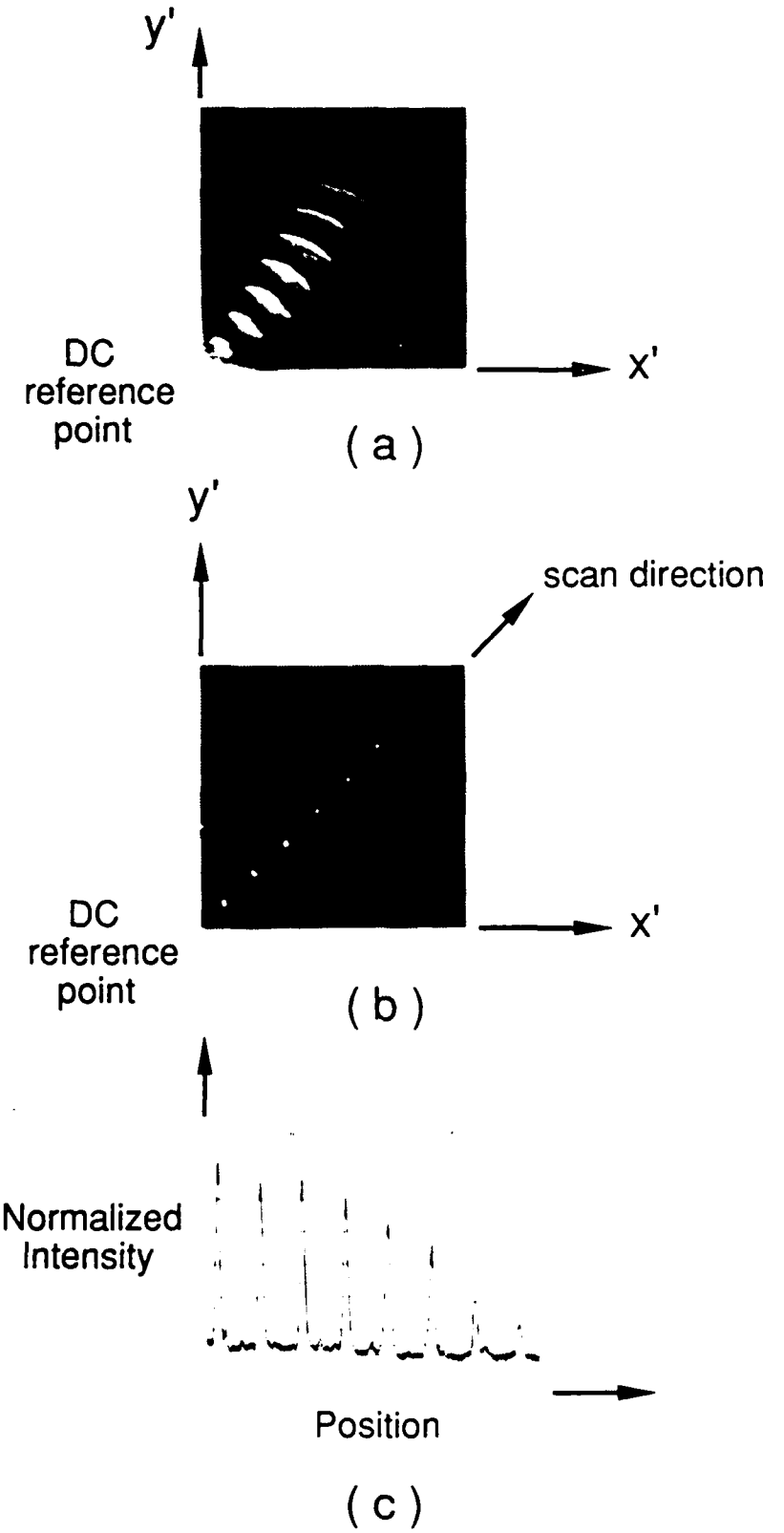


Fig. 8. Experimental results for SAR data for eight point scatterers. 2-D image produced (a) without polar formatting and (b) with polar formatting. (c) Line scan through point images in (b).

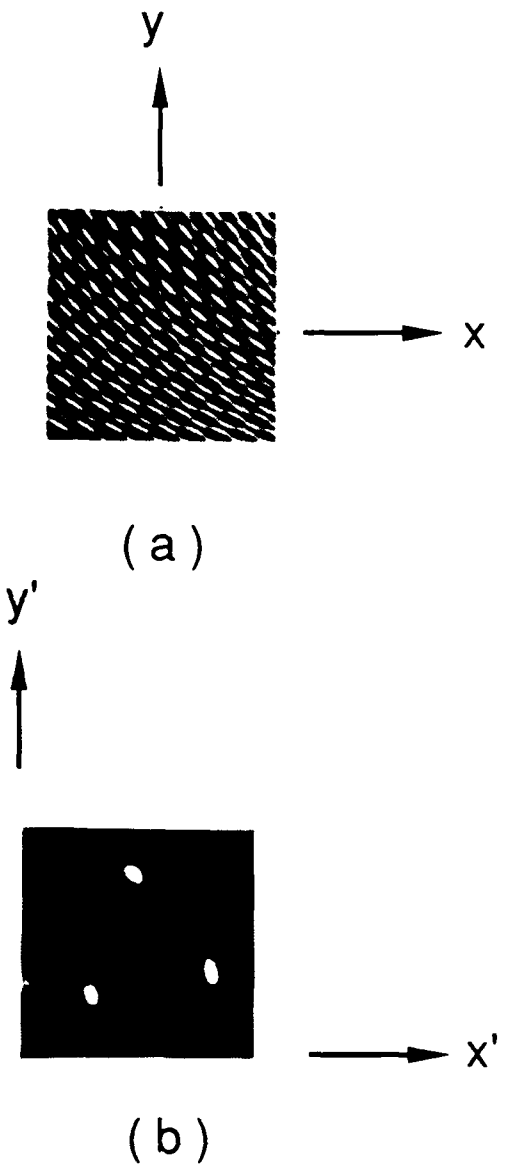


Fig. 9. Experimental results with LCTV input. (a) Data for three point scatterers which was input to LCTV. (b) Output SAR image.

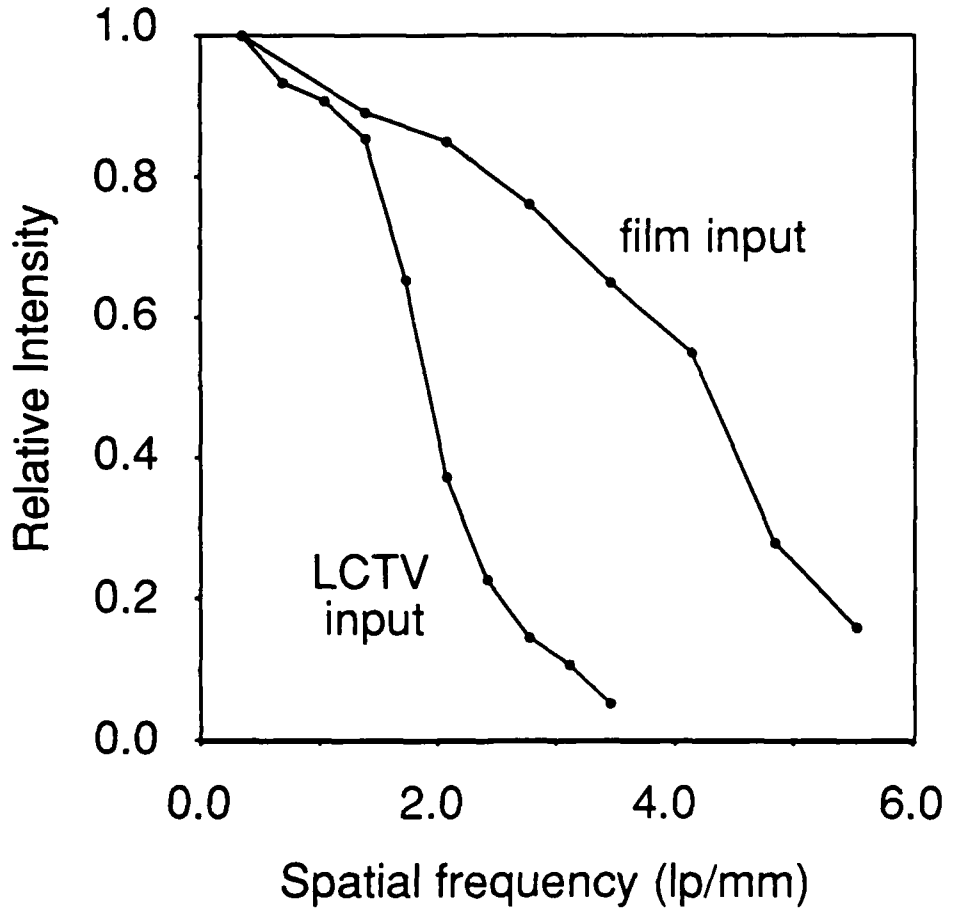


Fig. 10. Relative output image intensity versus spatial frequency of input data for film and LCTV input devices.

APPENDIX B

ITERATIVE DESIGN OF A HOLOGRAPHIC BEAMFORMER

(accepted by Applied Optics, 1989)

Abstract

A design method is presented for computing the phase functions of an energy-efficient system using two holographic elements for converting a Gaussian beam into a uniform beam with rectangular support in the far-field of the source. The method is based on a modification of the Gerchberg-Saxton algorithm which includes an x - y separability constraint on the phase of one of the holographic elements. A beamforming system was fabricated using this method, and experimental results were obtained which support the design approach.

I. Introduction

In applications such as laser radar and active coherent imaging, it is often desired to uniformly illuminate a rectangular area which is in the far-field of the source. When using a single mode laser with a Gaussian (TEM_{00}) intensity distribution as a source, an optical device is required to modify the near-field amplitude profile of the beam such that its phase and intensity become uniform upon propagation to the far-field.

Several methods have been investigated for beamforming in the near-field. The most straightforward approach is through the use of apodization and truncation; however, this approach is fundamentally energy inefficient. For energy efficiency, reflective[1] or refractive[2,3] surfaces have been used to redistribute the energy of a Gaussian beam to a near-field uniform beam of circular support. Unfortunately, these methods generally require complicated aspheric

surfaces which are difficult to fabricate, and are also applicable only in a limited set of situations. A more general implementation approach is to use diffractive optics in the form of computer-generated holograms.[4] The conversion of a Gaussian beam into a uniform beam of rectangular support, for example, has been demonstrated using this approach.[5]

Diffractive techniques have also been considered to obtain a practical beamformer which is useful in shaping a far-field beam. One system uses binary phase, interlaced gratings to provide shaping in one dimension.[6] In this system, the grating pattern provides both intensity weighting and the required phase shifts to produce a 1-D $\text{sinc}(x) = (\pi x)^{-1} \sin(\pi x)$ distribution in the near field. Upon propagation to the far-field, a 1-D rect(x) beam profile is formed. A similar system could be used to obtain a uniform far-field distribution with a circular region of support by producing a 2-D $J_1(\pi r)/(\pi r)$ distribution in the near-field. The system is fairly efficient in producing a $J_1(\pi r)/(\pi r)$ distribution even though weighting is used to generate the desired profile due to the similarity between the Gaussian and $J_1(\pi r)/(\pi r)$ functions. System efficiency, however, will be very poor if the approach is extended to producing a uniform far-field beam of 2-D rectangular support since the necessary near-field sinc(x)sinc(y) distribution has its energy concentrated near the axes. As shown in the Appendix, the efficiency achievable with this apodization approach in generating a $\text{sinc}^2(x)\text{sinc}^2(y)$ intensity distribution is only 5.8%. This efficiency does not include losses from truncation of the Gaussian beam.

To obtain greater energy efficiency, we investigated the use of a holographic beamforming system to provide the energy redistribution into a near-field $\text{sinc}^2(x)\text{sinc}^2(y)$ intensity profile with the correct phase without the use of intensity weighting. The beamformer concept is described in Section II. In Section III, several design procedures are described for synthesizing the required phases for the holographic

elements. Experimental procedures and results are then discussed in Section IV. Section V gives conclusions and suggestions for further research.

II. Beamformer Concept Using Two Separated Holographic Elements

A single holographic optical element does not provide enough degrees of freedom to produce a far-field beam with both uniform intensity and phase without using intensity weighting. In order to provide both the energy redistribution and phase correction, two separated holographic elements are required. The beamformer concept therefore uses a holographic element to redistribute the energy of the input beam into the necessary near-field intensity profile, and then uses a second holographic element to place the correct phase onto this beam. By relying solely on energy redistribution and phase correction, the system is limited in efficiency only by the maximum achievable diffraction efficiency of the holographic elements, which can be very high.

The basic beamformer system is illustrated in Fig. 1. The truncated Gaussian beam from a single transverse mode laser is incident on a holographic element (H1) which redistributes the energy into a desired near-field intensity distribution. Then, a second holographic element (H2) places the desired near-field phase distribution onto the beam such that, upon propagation to the far-field, a uniform intensity and phase distribution of approximately rectangular support results. After determining a near-field amplitude distribution which yields the desired far-field characteristics, the design of the beamforming system consists of computing the required phases for H1 and H2 to produce the desired near-field beam.

Since the illuminating beam is coherent, a Fourier transform relationship exists between the near-field and far-field amplitude distributions.[7] To produce a uniform far-field distribution of

rectangular support, denoted as rect(x) rect(y), a near-field sinc(x) sinc(y) distribution is required. This distribution, however, is of infinite extent, and is therefore unrealizable. To obtain a realizable near-field solution, the ideal solution must be windowed by some function which is zero outside the area of the beamformer elements.

In choosing the optimum window function, one inevitably encounters a tradeoff between uniformity and the ability to obtain sharp edges in the far-field due to the finite aperture size available. A rectangular window (i.e., truncation) in the near-field results in the sharpest far-field edges, but with very significant ringing. In order to reduce the ringing, the window function must roll off more uniformly to zero, but, in doing this, the far-field edges become less sharp. For the purpose of this investigation, a linear-weighted window with a cutoff at the sixth zero of the sinc function was arbitrarily chosen. The desired near-field amplitude distribution, $A_N(x,y)$, is given by

$$A_N(x,y) = \mathrm{sinc}\left[6x/x_o,\ 6y/y_o\right]\left[1 - |x|/x_o\right]\left[1 - |y|/y_o\right] \quad , \qquad (1)$$

where the second holographic element (H2) size is $2x_o$ by $2y_o$. The truncated Gaussian input beam was chosen to have a 10% truncation loss. The modulus of this beam is shown in Fig. 2(a). The magnitude of the distribution from Eq. (1) is shown in Fig. 2(b), and its corresponding far-field beam modulus in Fig. 2(c).

III. Holographic Element Design

Now that the desired near-field distribution has been chosen, a holographic system must be designed which can produce this distribution from a uniform phase, truncated Gaussian input. Referring to Fig. 1, holographic element H1 provides the energy redistribution from the truncated Gaussian beam into a beam with a weighted and truncated sinc^2 intensity profile. In this section, several methods are discussed for determining the phase of this hologram. The second holographic element

H2 corrects the phase of the output beam of H1, and is discussed at the end of this section.

A. Design Based on Stationary Phase Approximation

The first design method to be discussed is based on the method of Bryngdahl[8,9] for producing geometric transformations with holographic elements. The basic geometry is shown in Fig. 3(a), where the input field at P_1 has an amplitude distribution $a_1(x,y)$ with uniform phase. It is desired to add a holographic element with phase $\phi(x,y)$ in plane P_1 such that a given output amplitude distribution $a_2(u,v)$ results in the back focal plane P_2 of lens L. The approximate relationship between the two planes is given by the Fourier transform:

$$a_2(u,v) = -\frac{i}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_1(x,y)\,\exp\left[i\phi(x,y) - \frac{i2\pi}{\lambda f}(xu + yv)\right]dxdy \qquad (2)$$

where λ is the wavelength, f is the focal length of lens L, and (x,y) and (u,v) are the coordinates in planes P_1 and P_2, respectively. To simplify this system, note that by placing the input plane P_1 in contact with the lens L, as in Fig. 3(b), the output distribution only changes by a quadratic phase:

$$a_2'(u,v) = \exp\left[\frac{i\pi}{\lambda f}\left(u^2 + v^2\right)\right] a_2(u,v). \qquad (3)$$

Further, as shown in Fig. 3(c), the quadratic phase of the lens may be included in the holographic element by defining a new phase ϕ' where

$$\phi'(x,y) = \phi(x,y) - \frac{\pi}{\lambda f}\left(x^2 + y^2\right). \qquad (4)$$

Therefore, the necessary phase of the first beamforming hologram (H1) in Fig. 1 can be derived from $\phi(x,y)$.

By approximating the Fourier transform integral in Eq. (2) by the method of stationary phase,[10] it can be shown that a point (x,y) in the object plane is mapped to a point (u,v) in the Fourier plane given by the relation:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} (\lambda f/2\pi)\ \partial\phi(x,y)/\partial x \\ (\lambda f/2\pi)\ \partial\phi(x,y)/\partial y \end{bmatrix} . \tag{5}$$

If the desired mapping is defined by

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} g(x,y) \\ h(x,y) \end{bmatrix} , \tag{6}$$

the necessary phase to produce this mapping is the solution of the partial differential equations:

$$\begin{aligned} \frac{\partial\phi}{\partial x} &= \frac{2\pi}{\lambda f}\, g(x,y), \\ \frac{\partial\phi}{\partial y} &= \frac{2\pi}{\lambda f}\, h(x,y) . \end{aligned} \tag{7}$$

For conversion of a truncated Gaussian beam into a weighted sinc distribution, both distributions are separable functions. Because of this, the partial differential equations in Eq. (7) become ordinary differential equations in x and y. If the mapping functions g(x) and h(y) are bounded, the required coordinate transforms can be found independently in the x and y directions. Once the mapping functions are determined, the phase of the hologram is given by

$$\phi'(x,y) = \frac{2\pi}{\lambda f}\left[\int_0^x g(x')dx' + \int_0^y h(y')dy'\right] - \frac{\pi}{\lambda f}\left[x^2 + y^2\right]. \tag{8}$$

Note that the quadratic phase of the lens has been added.

The method of conservation of energy[11] can be used in both the x and y directions to determine the mapping functions. Let one dimension of the incident Gaussian intensity be represented by

$$I_g(x) = \begin{cases} I_o \exp(-x^2/2\alpha^2) & |x| \leq x_o \\ 0 & |x| > x_o, \end{cases} \tag{9}$$

where I_o is the peak intensity, x_o is the half-width of the truncated beam, and α is a constant. From Eq. (1), the desired near-field 1-D intensity distribution is

$$I_N(u) = \begin{cases} I_{No} \operatorname{sinc}^2\left(6u/u_o\right)\left[1 - |u|/u_o\right]^2 & |u| \leq u_o \\ 0 & |u| > u_o \, . \end{cases} \tag{10}$$

The mapping can be determined with the aid of Fig. 4. All of the energy bounded by $0 \leq x' \leq x$ from the Gaussian intensity profile is mapped into a region bounded by $0 \leq u' \leq u$ in the sinc^2 intensity distribution. This can be described mathematically by

$$I_o \int_o^x \exp(-x'^2/2\alpha^2)\, dx' = I_{No} \int_o^u \operatorname{sinc}^2(6u'/u_o)\,(1 - |u'|/u_o)^2\, du' \tag{11}$$

for $|x| < x_o$ and $|u| < u_o$. The peak intensity in the two planes is related by conservation of the total energy in both distributions, or

$$\int_{-\infty}^{\infty} I_g(x)\, dx = \int_{-\infty}^{\infty} I_N(u)\, du \, . \tag{12}$$

Inserting the functions given in Eqs. (9) and (10) and solving for I_{No} gives

$$I_{No} = I_o \frac{\int_{-x_o}^{x_o} \exp(-x^2/2\alpha^2)\, dx}{\int_{-u_o}^{u_o} \mathrm{sinc}^2(6u/u_o)\,(1 - |u|/u_o)^2\, du} \quad . \tag{13}$$

which can be evaluated numerically.

The mapping function g(x) can be determined by numerically solving Eq. (11) to determine a set of points u as a function of x. This is implemented by computing the left-hand-side of Eq. (11) for a given value of x, and then computing the right-hand-side for increasing values of u until the equality is satisfied. Due to the separability of the Gaussian and sinc functions, h(y) can be similarly obtained. With the mapping functions defined, Eq. (8) can then be numerically integrated to determine the needed phase function for the light redistribution hologram.

For this investigation, a symmetric Gaussian beam truncated to contain 90% of its energy within a $(10.24\text{mm})^2$ area was used as an input. The beamformer was designed to produce the weighted sinc^2 intensity of Eq. (10) with x_o and y_o equal to 5.12mm. The separation between the holographic elements was 375mm. After evaluating Eq. (13) numerically for I_{No}, the mapping function was determined by balancing the integrals of Eq. (11) for 100 (x,u) pairs. The resulting function was interpolated to 1024 points using a cubic spline interpolation and numerically integrated using Eq. (8) to compute the required holographic element phase.

The stationary phase design method was tested by computer simulation. This was done by creating a 256 x 256 array with Gaussian intensity and the numerically determined phase (without the quadratic phase), padding it with zeros to a 512 x 512 size, and taking a 2D FFT to give the near-field distribution. The simulated far-field distribution was computed by replacing the simulated near-field phase with the ideal phase (alternating π phase shifts of the sinc function), and taking a 2D FFT. The resulting near-field and simulated far-field moduli are shown in Fig. 5. It can be seen by comparing to Fig. 2 that, while the stationary phase approximation was sufficient to provide the general shape of the near-field distribution, it was clearly insufficient for accurately producing the zeros of the distribution, and results in unacceptable ringing and high sidelobes in the far-field.

B. Design Using Iterative Fourier Transform Algorithms

To improve the design of the previous section, a second design method was investigated. Because of the Fourier transform relationship between the two planes in Fig. 3(a), the design problem is equivalent to that of computing an object phase (CGH phase not including quadratic phase) for a known object modulus (truncated Gaussian) and a known Fourier modulus (weighted sinc). This synthesis problem is equivalent to that of phase retrieval from a pair of intensity measurements,[12] in which the complex fields in two domains related by a Fourier transform are reconstructed from incomplete information in each domain. This problem has gained much attention in the past, in applications such as electron microscopy, and is typically solved using an iterative Fourier transform algorithm. For this specific case, in which the object and Fourier moduli are known, the Gerchberg-Saxton algorithm[13] allows reconstruction of the phase.

A block diagram of the Gerchberg-Saxton algorithm is shown in Fig. 6. The algorithm consists of an iterative loop having four steps: (1) the known object modulus with an initial phase estimate ϕ_0 is

transformed to the Fourier domain, (2) the Fourier modulus |G| is replaced with the known or desired modulus, (3) the resulting Fourier data is transformed back to the image domain, and (4) the object modulus |g'| is replaced with the known object modulus. The resulting object data is then input to step (1), and the cycle is repeated until convergence is achieved. This algorithm has been used to design holographic beamformers[14,15] and computer-generated Fourier holograms.[16-18]

A 2-D iterative synthesis of the first holographic element phase using this algorithm was attempted. The results were found to be very sensitive to the initial phase estimate used to start the algorithm. It was found that by starting with a random phase, the resulting redistributed modulus is approximately a weighted sinc, but with randomly distributed error, as shown in Fig. 7. The resulting CGH phase, though an approximate solution, has significant random features which vary for each different random input. This lack of repeatability was considered unacceptable for a synthesis procedure.

Better results were obtained by using the stationary phase solution as an initial phase estimate. After 300 iterations, the algorithm converged to the solution whose simulated near-field and far-field moduli are shown in Fig. 8. Although the algorithm was initialized with a separable object modulus, object phase, and Fourier modulus, the resulting object phase became less separable as the number of iterations increased. This is attributed to finite word length effects in the digital computation.

This lack of separability in the phase solution is an undesirable feature of the Gerchberg-Saxton result. The solution would not be expected to be optimal, since the coordinate transform problem is exactly separable. Also, a separable solution would be easier to encode as a CGH, because it would make the interpolation to the 2048 x 2048 holographic element size less computationally intensive. For these reasons, it was desired to find a separable phase solution.

The iterative algorithm was therefore modified to enforce a separability constraint on the object phase. This hybrid algorithm is shown in Fig. 9. An initial phase estimate is used to start the Gerchberg-Saxton algorithm, and a specified number of iterations are performed. Then a 1-D slice of the current phase estimate is extracted, and an outer product is performed to produce a 2-D separable phase estimate. This is used to start another set of iterations, and the loop is repeated until an acceptable solution is found by examining the simulated far-field intensity. After each separability constraint is imposed, the phase estimate improves, so that the iterative algorithm converges with fewer iterations. Each successive phase solution is closer to being separable, so that overall convergence to a good separable phase solution would be expected.

A phase solution was generated by performing 5 loops of the hybrid algorithm with a successively decreasing number of Gerchberg-Saxton iterations per hybrid loop (100, 50, 25, 10, and 5). The results are shown in Fig. 10. When compared to the nonseparable phase solution generated by the Gerchberg-Saxton algorithm alone [Fig. 9(a)], the hybrid solution has slightly better uniformity and lower sidelobes in the simulated far-field beam for the same total number of iterations. Since it is separable, it is also easier to encode as a CGH.

C. Design of Second Holographic Element

To complete the beamformer system design, the phase of the second holographic element H2 must be specified. The purpose of H2 is to place on the beam the desired near-field phase which is a simple, known function consisting of alternating π phase shifts at the zeroes of the sinc(x) sinc(y) distribution. H2 must also compensate for the non-uniform phase of the beam arriving from H1. It is therefore necessary to determine this phase function, which can be done by numerically propagating the complex amplitude at the plane of H1 to that of H2, and extracting the phase. Since this may require some approximations and

would ignore any aberrations in the system, it may be limited in its accuracy. For this reason, it may be necessary to measure the phase at the plane of H2 with an interferometric system. Once this is done, the required phase of H2 can also be encoded in a CGH. These two methods may be very difficult to perform experimentally. First, it may be difficult to compute or measure the phase at the plane of H2 within the needed accuracy since the beam phase changes very rapidly in this region. Second, even if the phase can be determined, the tolerances for aligning H2 to H1 would be very tight. Small errors in alignment or phase compensation would lead to unacceptable results in the beamforming.

A third method for producing H2 is by interfering the output of H1 with a beam carrying the desired near-field phase. Upon reconstruction, the desired phase is imposed on the beam after it is diffracted by H2. Ideally, the diffraction efficiency across the entire phase correction hologram should be constant in order for the diffracted output to maintain the desired profile. To achieve uniform diffraction efficiency with a recording beam (coming from H1) having a nonuniform intensity distribution requires a highly non-linear recording process. Unfortunately, such a non-linear recording will result in high intermodulation noise. To avoid the problem with intermodulation noise, the recording has to be kept linear. A linear recording, however, produces a diffraction efficiency that is proportional to the input intensity distribution. The result is a squaring of the input intensity profile. This squaring effect can be compensated by designing the light redistribution hologram to produce an intensity distribution equal to the square root of the desired intensity distribution. Another drawback of restricting the recording to the linear region is the low average diffraction efficiency that can be obtained.

IV. Experiment

To verify the results of the computer simulation, H1 was produced using the phases synthesized by both the stationary phase and the hybrid algorithm design methods. The conclusions which were drawn earlier by simulation of the stationary phase design method are supported by experimental results. A CGH was produced by adding the phase solution for H1 from this method to an appropriate quadratic phase, encoding the phase as a real, non-negative amplitude transmittance function on a carrier, and recording the distribution on film using a rotating drum laser beam recorder.[4] A hologram size of $(1.024\ cm)^2$ was used with a 5μm pixel size to produce a hologram space-bandwidth product of 2048 x 2048. The hologram was mounted between 1/4 inch microflats using UV curing cement in order to eliminate phase errors due to thickness variations in the film. When illuminated with the Gaussian beam, the redistributed near-field intensity profile occurs in the first diffracted order in a plane 375mm away, and is shown in Fig. 11. Because the diffraction efficiency of the CGH was very low (~ 1%), the intensity distribution was measured with a scanning fiber optic probe which was coupled into a photomultiplier tube detector. The slight asymmetry in the distribution is due to the MTF of the laser writer (i.e., higher spatial frequencies are diffracted less strongly).

Significantly better results were obtained with the hybrid algorithm design method. A CGH with the designed phase function was produced as above using the laser writer. The intensity distribution of the diffracted Gaussian beam at plane H2 is shown in Fig. 12(a). Since the result of the hybrid algorithm design method is clearly better than that of the stationary phase method, the CGH designed using the former approach was used for H1. The CGH recorded on film as an absorption hologram, however, is too low in diffraction efficiency (~ 1%) to be practical. In order to improve the diffraction efficiency, the CGH was copied onto a volume phase material with a high carrier frequency, which then becomes known as a computer-originated holographic optical element

(COHOE).[19] The COHOE was recorded as shown in Fig. 13 by exposing a Kodak 649F spectroscopic plate using a HeNe laser (λ = 0.6328 μm), and processing using the silver halide (sensitized) gelatin process.[20] A diffraction efficiency of approximately 43% was achieved, which could be increased by using the dichromated gelatin process[21] and an argon laser. When reconstructed in the same manner as the CGH, the resulting near-field intensity of the COHOE is shown in Fig. 12(b).

A note should be made here about the methods used to obtain the intensity line scans. The intensity distributions of the output of the CGHs [Figures 11(b) and 12(e)] were obtained with the scanning fiber optic probe since the diffraction efficiency was so low. For the output of the COHOE [Figure 12(f)], however, more light was available for detection, so the line scans were obtained by using a CCD camera and grabbing video lines with a digital oscilloscope. The differing methods of detection account for the different characteristics of the line scans. In particular, the pixel size of the fiber optic probe is larger than that of the CCD camera, resulting in more spatial integration which smooths the output of the probe.

The second hologram H2 was produced using the third method described at the end of Section III by interfering the output of the first hologram H1 with a beam carrying the desired near-field phase of the sinc distribution. This desired phase (alternating π phase shifts) was encoded and recorded as a CGH on film. This step was very sensitive to alignment errors, since phase compensation has to occur at the right plane and the desired π phase shifts must line up almost exactly with the zeroes of the sinc distribution. Very slight misalignment errors were found to result in poor beamforming.

An added problem involves the linearity of the recording. Since the sidelobes of the near-field distribution are much lower than the mainlobe, a linear recording over a very wide dynamic range is required. Relatively small nonlinearities result in incorrect sidelobe-to-mainlobe

intensity ratios, which correspond to either excessive ringing or edge blurring in the far-field beam. Keeping the recording very linear severely limits the overall diffraction efficiency that can be obtained. The resulting holographic element in our experiment had a diffraction efficiency of only 5.2%.

The results from the complete beamforming system are shown in Fig. 12(c) and (d). In the near-field, the squaring effect of the second hologram can be readily seen by the relative sidelobe intensities. In the far-field, the beam is relatively uniform with very low sidelobes and sharp edges, although there is still some ringing in the edges and corners. The entire distribution also seems to be weighted slightly higher in one corner. This is due to the presence of nonlinearities in the recording and small misalignment errors. The fringes seen in the far-field beam are due to interference in the faceplate of the CCD camera, and are not present in the actual output beam of the beamforming system. The line scans [Figures 12(g) and 12(h)] were obtained from the video output of the CCD camera.

V. Discussion

In order to obtain a highly efficient, uniform far-field beamforming device, the system must rely solely on redistribution of energy. Several methods were investigated for computing the necessary phase functions of a holographic system for this application. It was found that a method based on using a stationary phase approximation to the Fourier transform integral was inadequate for producing good zeroes in the near-field distribution. Significantly better results, however, were achieved using a hybrid algorithm which combined an x-y separability constraint with an iterative Fourier transform algorithm. Further improvements may also be made by using alternate algorithms such as the input-output algorithm that speeds convergence.[12,17] Another possibility would be to use a two-step design procedure which would optimize the beam in the far-field plane rather than specifying the desired beam at H2.

Experimental results demonstrated the ability of the holographic system to provide accurate beamshaping. To obtain uniformly high efficiency across the second holographic element, it must be fabricated as a COHOE by computing or measuring the necessary phase as described in Section III. While this puts very high demands both on alignment and the ability to produce precise phase compensation, it is necessary in order to make an efficient far-field beamforming device. By doing this, the overall efficiency becomes limited only by that of the holographic optical elements which would be expected to be >80% with existing materials (i.e., dichromated gelatin). For comparison, using intensity weighting to obtain the near-field weighted sinc^2 intensity distribution of Eq. (10), the maximum efficiency that can be achieved is only 8% (see Addendum). Neither of these efficiencies include losses due to truncation of the input Gaussian beam.

This research was sponsored by the U.S. Army Research Office. The authors would like to thank J.R. Fienup for helpful advice on the use of the Gerchberg-Saxton algorithm.

ADDENDUM

Since both the truncated Gaussian and weighted sinc^2 intensity distributions are separable functions, the efficiency in apodizing a 2-D Gaussian into a weighted sinc^2 beam profile is the square of the 1-D efficiency. Let the 1-D truncated Gaussian intensity be described by

$$I_G(x) = \begin{cases} I_o \exp(-x^2/2\alpha^2) & |x| \leq x_o \\ 0 & |x| > x_o \end{cases} \tag{A1}$$

and the 1-D weighted sinc^2 intensity by

$$I_N(x) = \begin{cases} I_{No} \, \text{sinc}^2(Nx/x_o)(1 - |x|/x_o)^2 & |x| \leq x_o \\ 0 & |x| > x_o \end{cases} \tag{A2}$$

where $2x_o$ is the window width and $(N-1)$ is the number of sidelobes on each side of the sinc function inside the window.

For efficient apodization, the intensities are equalized at $x = 0$:

$$I_{No} = I_o. \tag{A3}$$

In order to exactly produce the distribution in Eq. (A2) without energy redistribution, the α parameter of the Gaussian must be chosen such that

$$I_G(x) \geq I_N(x) \quad \text{for} \quad |x| \leq x_o. \tag{A4}$$

The greatest efficiency is obtained when the intensities are equal at the peak of the $(N-1)$th sidelobe, which is located approximately at $x = [(N - 1/2)/N] \, x_o$. The condition is therefore given by

$$I_o \exp\left[-\left(\frac{N - 1/2}{N}\right)^2 \frac{x_o^2}{2\alpha^2}\right] = I_o \operatorname{sinc}^2\left[(N - 1/2)\right]\left[1 - \left(\frac{N - 1/2}{N}\right)\right]^2. \quad \text{(A5)}$$

This can be solved for the required α parameter:

$$\alpha = \left[\left(\frac{N - 1/2}{N}\right)^2 \frac{x_o^2}{2\ln\left[4N^2\left(N - 1/2\right)^2\pi^2\right]}\right]^{1/2}. \quad \text{(A6)}$$

The 1-D apodization efficiency, η, is equal to the ratio of the energy in the apodized beam to that in the incident beam. For this case,

$$\eta = \frac{\int_{-x_o}^{x_o} I_o \operatorname{sinc}^2\left[Nx/x_o\right]\left[1 - |x|/x_o\right]^2 dx}{\int_{-x_o}^{x_o} I_o \exp\left[-x^2/2\alpha^2\right] dx} \quad \text{(A7)}$$

where α is given by Eq. (A6). The 2-D efficiency is η^2. For the case where N = 6, the integrals in Eq. (A7) can be computed numerically to obtain $\eta^2 = 8.1\%$. This efficiency can be improved by using a smaller α and clipping the higher sidelobes of the sinc^2 intensity, but this will result in some degradation in the far-field beam.[6] We have also calculated the efficiency of converting a Gaussian beam into a non-weighted sinc^2 intensity distribution. This was done by eliminating the weighting term in Eq. (A2) and computing the efficiency in the same manner as shown above. In this case, the efficiency is only 5.8%.

References

1. J.W. Ogland, "Mirror System for Uniform Beam Transformation in High Power Annular Lasers," Appl. Opt. 17, 2917 (1978).

2. P.W. Rhodes and D.L. Shealy, "Refractive Optical Systems for Irradiance Redistribution of Collimated Radiation: Their Design and Analysis," Appl. Opt. 19, 3545 (1980).

3. D. Shafer, "Gaussian to Flat-top Intensity Distributing Lens," Opt. and Laser Tech., p. 159 (June 1982).

4. J. Cederquist and A.M. Tai, "Computer-generated Holograms for Geometric Transformation," Appl. Opt. 23, 3099 (1984).

5. C.C. Aleksoff, K.K. Ellis, and B.D. Neagle, "Holographic Conversion of a Gaussian Beam to a Uniform Beam," Proc. SPIE 883, 220 (1988).

6. W.B. Veldkamp, "Laser Beam Profile Shaping with Interlaced Binary Diffraction Gratings," Appl. Opt. 21, 3209 (1982).

7. J.W. Goodman, Introduction to Fourier Optics (McGraw-Hill, San Francisco, CA, 1968), p. 57.

8. O. Bryndgdahl, "Optical Map Transformations," Opt. Commun. 10, 164 (1974).

9. O. Bryndahl, "Geometrical Transformations in Optics," J. Opt. Soc. Am. 64, 1092 (1974).

10. M. Born and E. Wolf, Principles of Optics (Pergamon, New York, 1959), p. 749.

11. C.-Y. Han, Y. I. Shii, and K. Murata, "Reshaping Collimated Laser Beams with Gaussian Profile to Uniform Beams," Appl. Opt. 22, 3644 (1983).

12. J.R. Fienup, "Reconstruction and Synthesis Applications of an Iterative Algorithm," Proc. SPIE 373, 147 (1981).

13. R.W. Gerchberg and W.O. Saxton, "A Practical Algorithm for Determination of Phase from Image and Diffraction Plane Pictures," Optik 35, 237 (1972).

14. W.-H. Lee, "Method for Converting a Gaussian Laser Beam into a Uniform Beam," Opt. Commun. 36, 469 (1981).

15. H.O. Bartelt, "Applications of the Tandem Component: an Element with Optimum Light Efficiency," Appl. Opt. 24, 3811 (1985).

16. N.C. Gallagher and B. Liu, "Method for Computing Kinoforms that Reduce Image Reconstruction Error," Appl. Opt. 12, 2328 (1973).

17. J.R. Fienup, "Iterative Method Applied to Image Reconstruction and to Computer-Generated Holograms," Opt. Eng. 19, 297 (1980).

18. F. Wyrowski and O. Bryngdahl, "Iterative Fourier-transform Algorithm Applied to Computer Holography," J. Opt. Soc. Am. A 5, 1058 (1988).

19. R.C. Fairchild and J.R. Fienup, "Computer Originated Aspheric Holographic Optical Elements," Opt. Eng. 21, 133 (1982).

20. D.K. Angell, "Improved Diffraction Efficiency of Silver Halide (Sensitized) Gelatin," Appl. Opt. 26, 4692 (1987).

21. B.J. Chang and C.D. Leonard, "Dichromated Gelatin for the Fabrication of Holographic Optical Elements," Appl. Opt. 18, 2407 (1979).

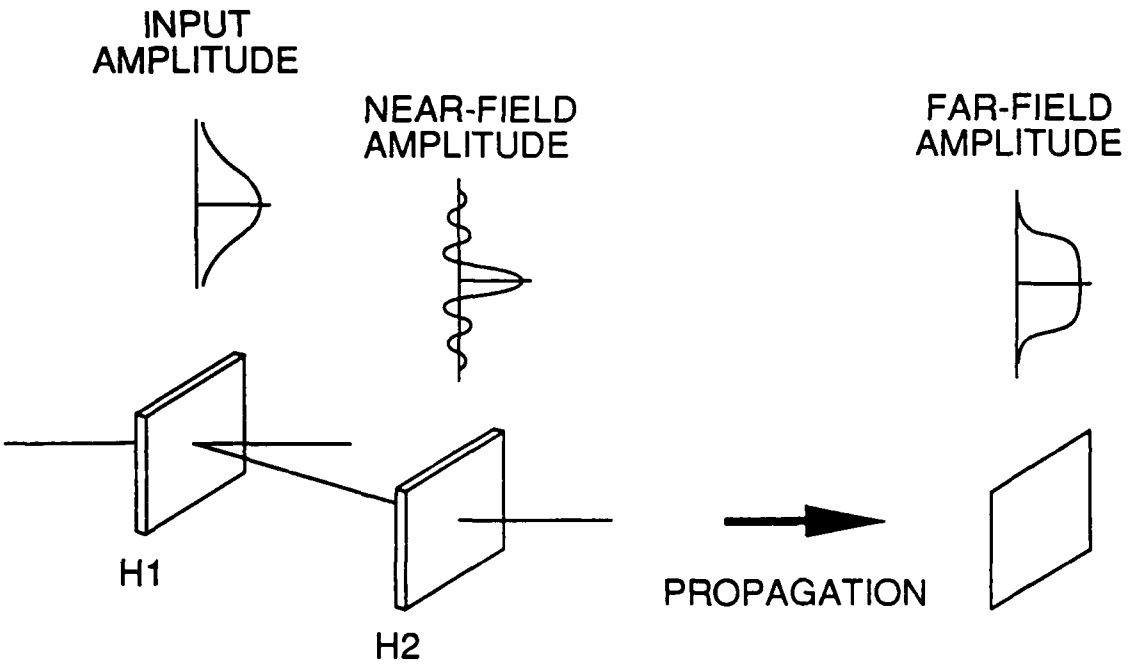


Fig. 1. Holographic beamformer. Holographic element H1 redistributes the energy of the incident beam into the desired near-field intensity distribution. Holographic element H2 places the correct phase on the beam. Upon propagation to the far-field, a uniform intensity and phase results.

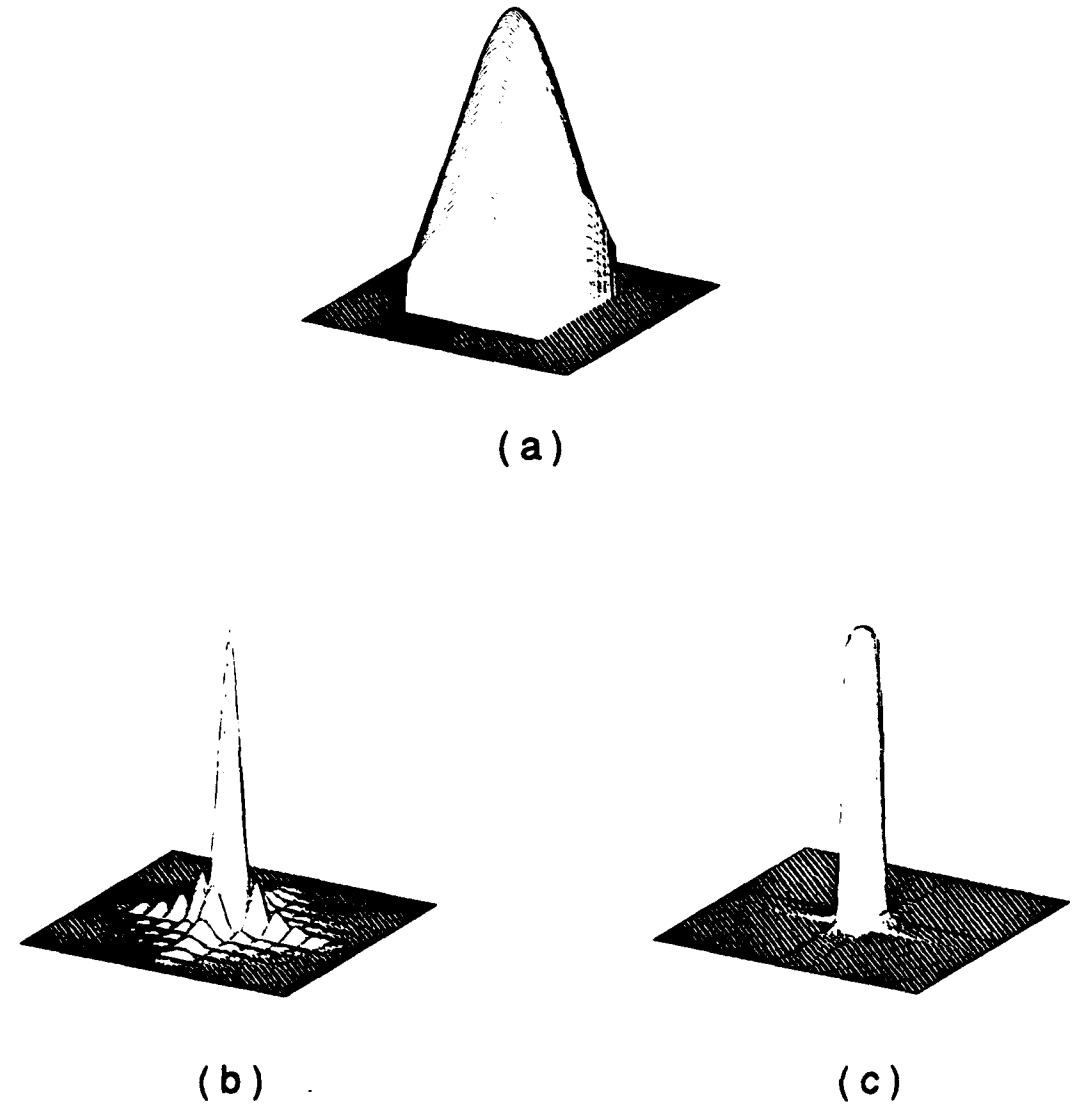


Fig. 2. Beam profiles. (a) Truncated Gaussian profile; (b) desired near-field modulus; (c) desired far-field modulus.

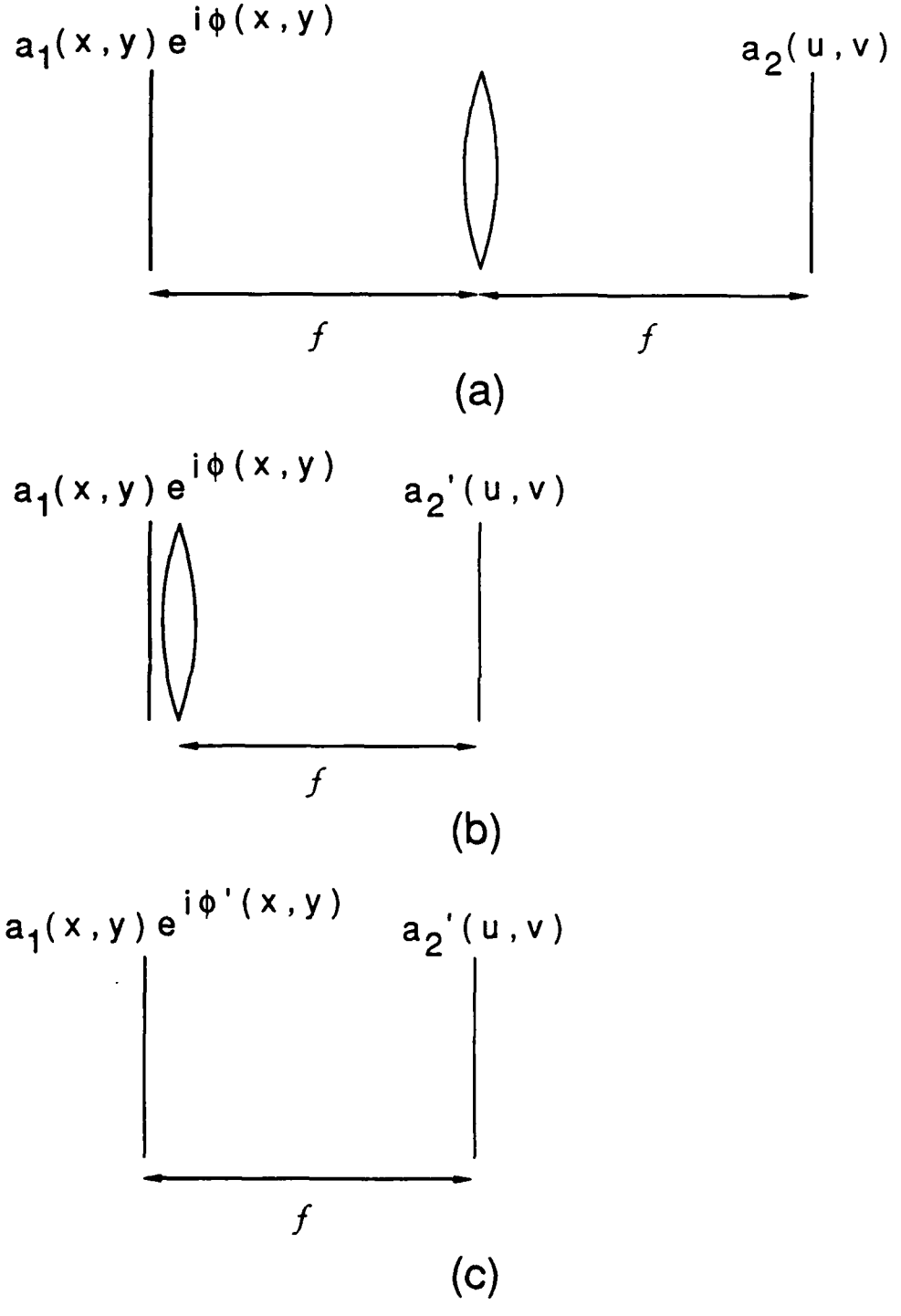


Fig. 3. Equivalent systems for holographic redistribution of energy. In (a), the complex amplitude distributions are related by a Fourier transform. If the holographic element is placed in contact with the lens, as in (b), the intensity distribution at the output plane is unchanged. In (c), the quadratic phase of the lens is added to the holographic element.

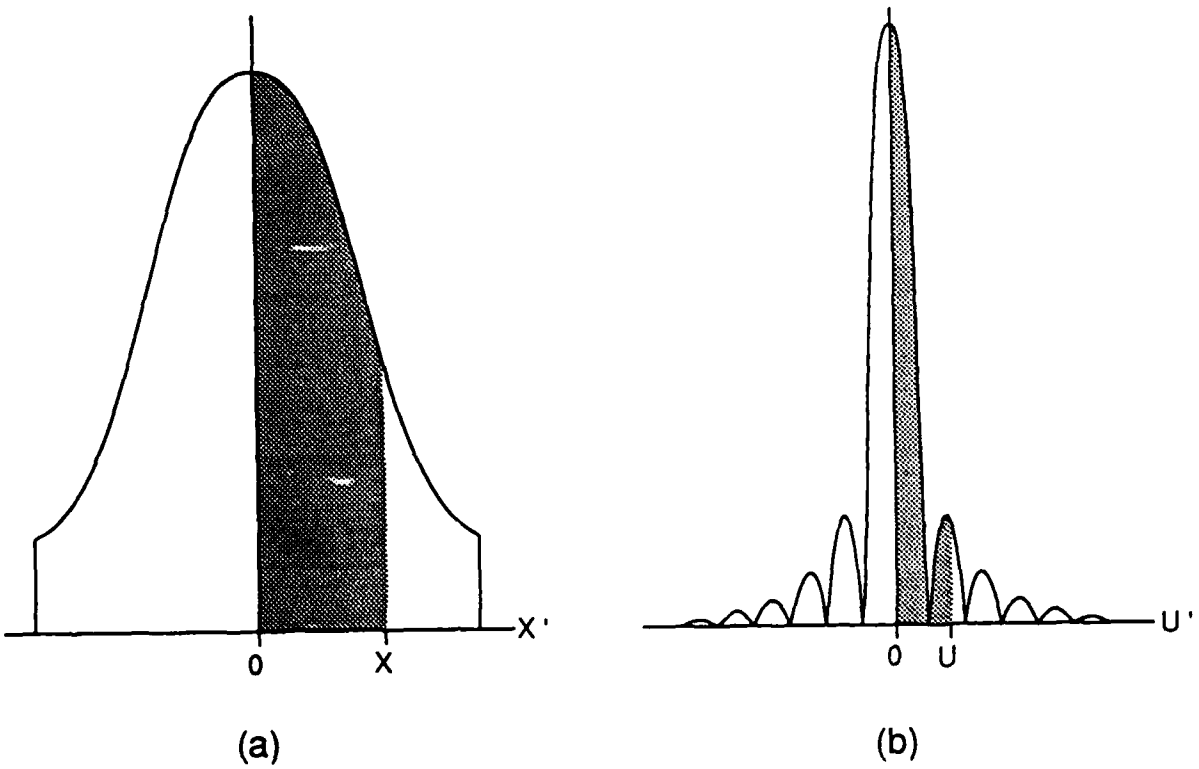
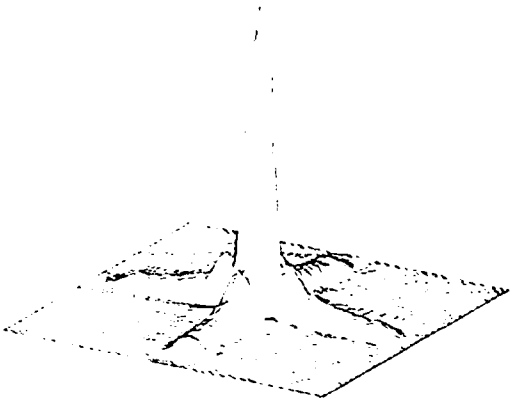
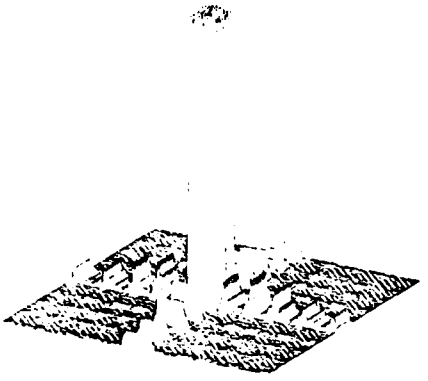


Fig. 4. Determination of the mapping function through conservation of energy. The energy bounded by $0 \leq x' \leq x$ in (a) must equal that bounded by $0 \leq u' \leq u$ in (b).



(a)



(b)

Fig. 5. Simulated results for beamformer designed by stationary phase method. (a) Near-field modulus; (b) far-field modulus.

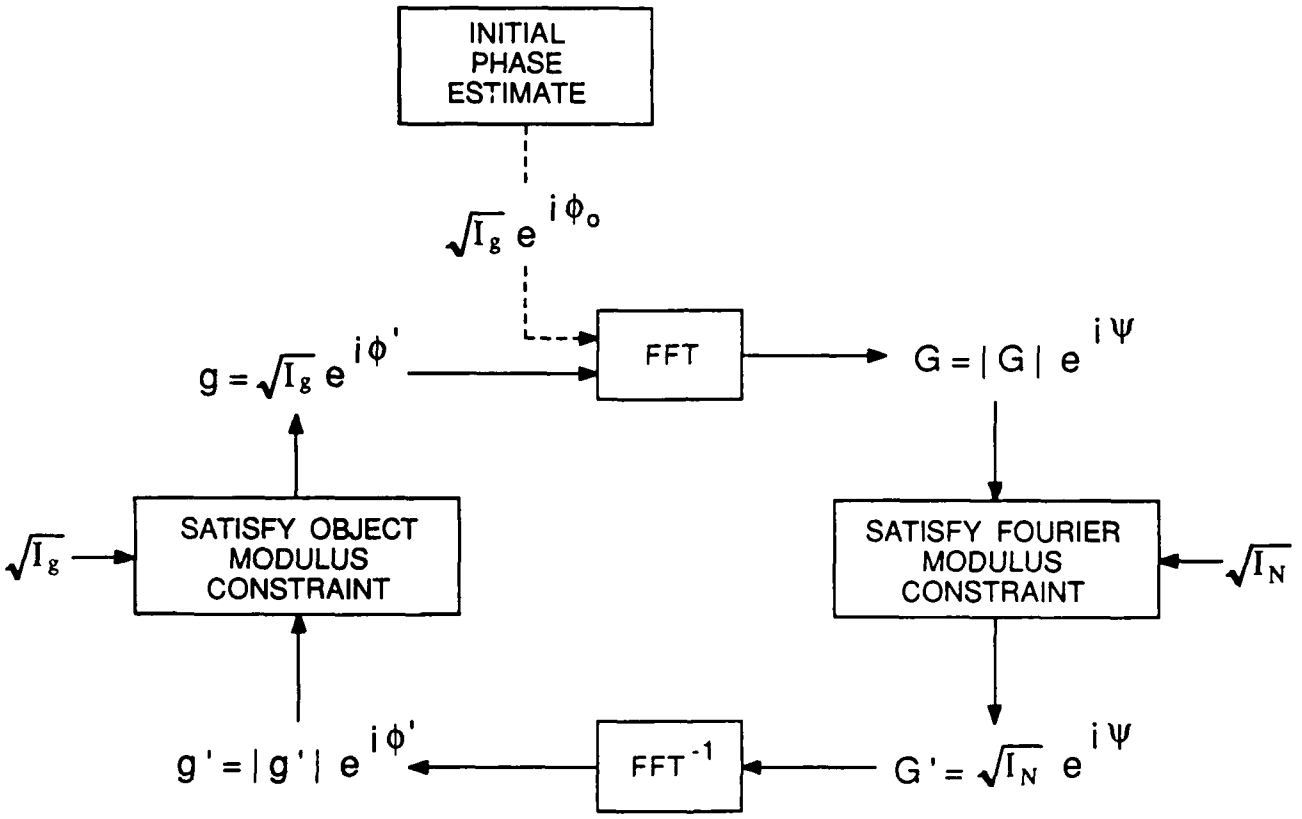


Fig. 6. The Gerchberg-Saxton algorithm.

88-11487

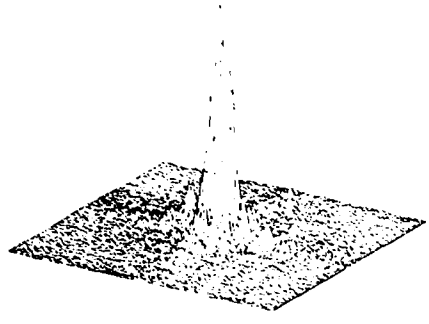
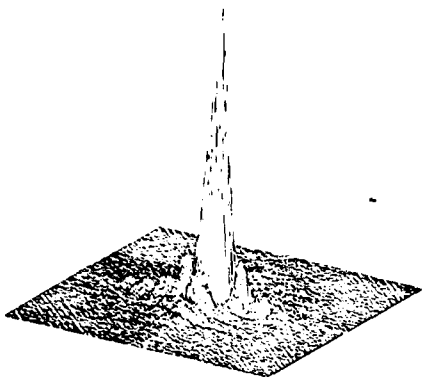
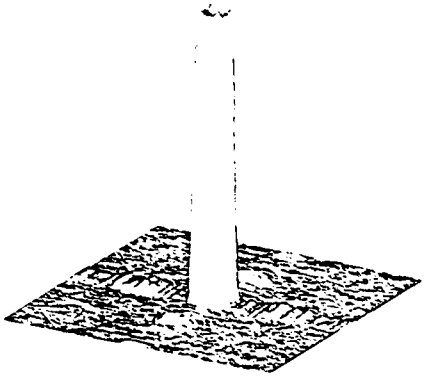


Fig. 7. Near-field modulus obtained using Gerchberg-Saxton algorithm with a random initial phase estimate.



(a)



(b)

Fig. 8. Synthesis results using Gerchberg-Saxton algorithm with initial phase estimate from stationary phase method. (a) Near-field modulus; (b) far-field modulus.

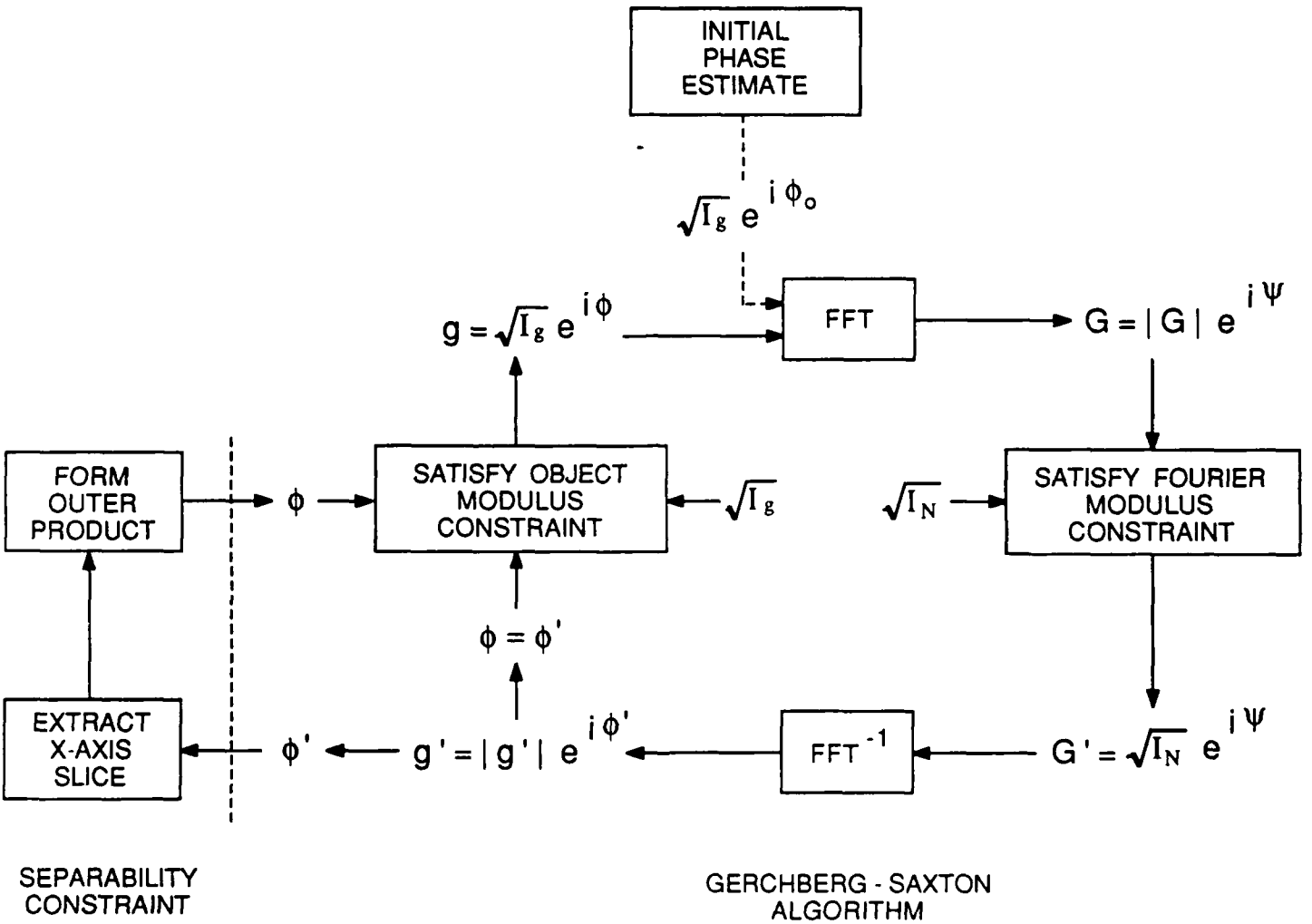
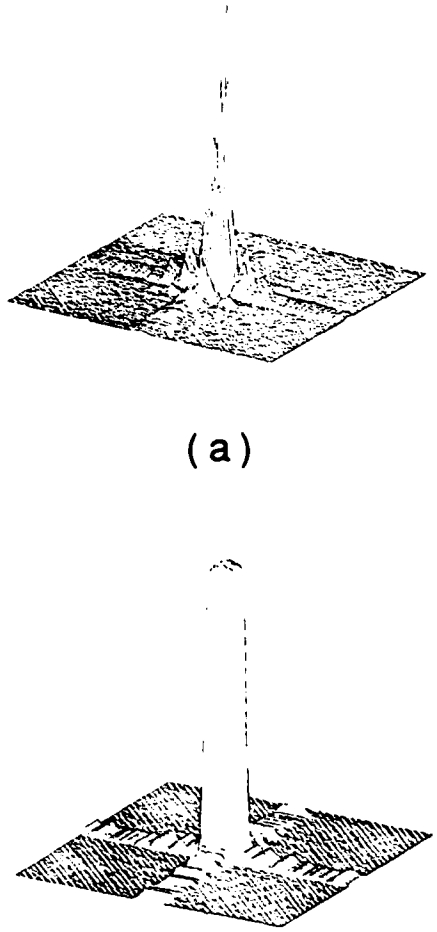


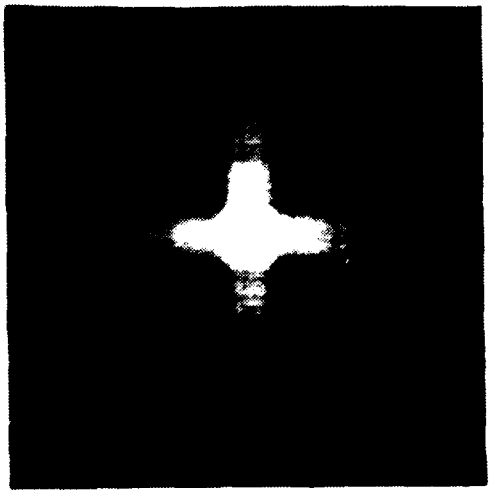
Fig. 9. The hybrid iterative algorithm. The separability constraint is applied after each set of Gerchberg-Saxton iterations.



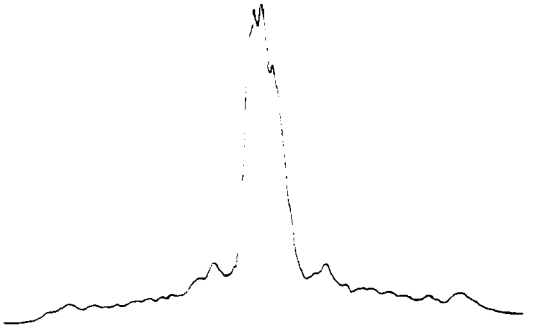
(a)

(b)

Fig. 10. Synthesis results for beamformer designed using hybrid algorithm. (a) Near-field modulus; (b) far-field modulus.



(a)



(b)

Fig. 11. Experimental results for CGH computed by the stationary phase method. (a) Intensity distribution at a plane 375mm from the CGH; (b) line scan through (a).

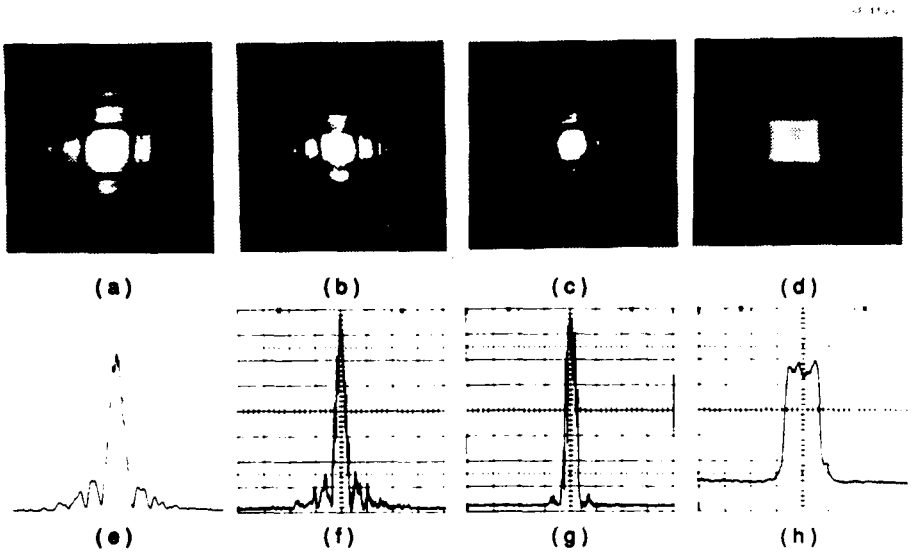


Fig. 12. Beamformer experimental results. (a) Intensity distribution at a plane 375 mm from CGH; (b) same result from COHOE; (c) near-field intensity distribution directly after H2; (d) far-field intensity distribution; (e)-(h) are line scans through the distributions in (a)-(d), respectively.

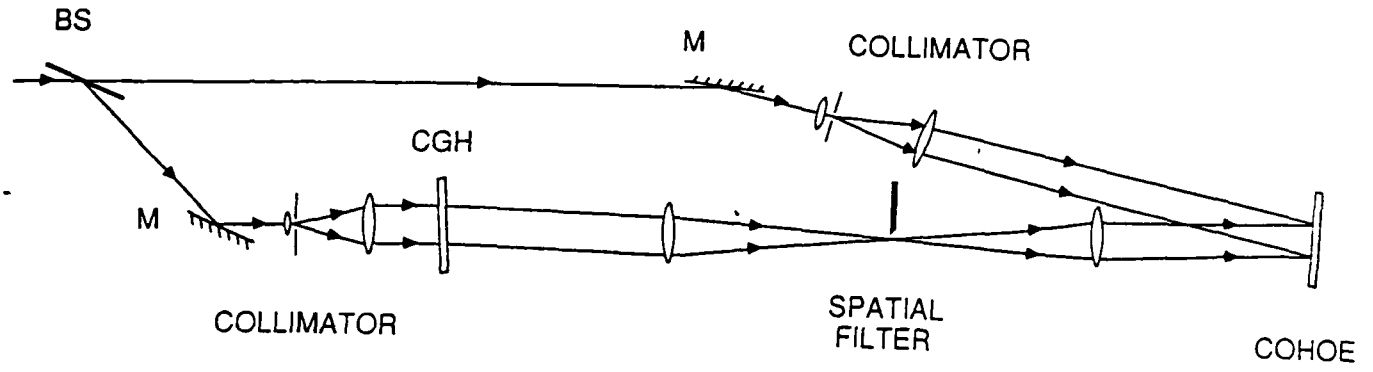


Fig. 13. COHOE fabrication. The desired first diffraction order of the CGH is imaged onto the COHOE plane and interfered with a reference beam.

APPENDIX C

COORDINATE TRANSFORMATIONS REALIZABLE WITH MULTIPLE HOLOGRAPHIC OPTICAL ELEMENTS

(submitted to Journal of the Optical Society of America A, 1989)

Abstract

Using the stationary phase approximation, holographic optical elements can be designed which perform one-to-one coordinate transformations. All coordinate transformations which are sufficiently differentiable can be decomposed into two transformations in series and therefore performed using two holographic elements. Although the method presented here for designing the two required holographic elements is not explicit for the general case, a useful necessary condition for the decomposition is given. A simple explicit solution based on separability of the necessary equations for the first of the two serial transformations is useful in some cases.

1. Introduction

Optical processors often operate on data that are presented in two spatial dimensions. In some applications, the data need to be spatially rearranged at some point in the processing. These data rearrangements may be thought of as one-to-one coordinate transformations (also called geometric distortions or mappings). Conventional optics cannot easily perform complicated transformations, so holographic optical elements (HOEs) are often used for coordinate transformations. In particular, computer-originated HOEs are used because of the greater ease with which complicated phase functions may be fabricated.[1,2]

One method for designing the HOEs required for a given coordinate transformation uses the stationary phase approximation and was developed by Bryngdahl.[3,4] To review this method, consider the optical system shown in Fig. 1 where the input field at P_1 in the front focal plane of a lens L is uniphase and has an amplitude distribution $a_1(x, y)$. A HOE with phase $\phi(x, y)$ is placed in plane P_1. The complex amplitude distribution $a_2(s, t)$ found in the back focal plane P_2 of the lens is given approximately by the Fourier transform:

$$a_2(s, t) = -\frac{i}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_1(x, y)\exp\left[i\phi(x, y) - \frac{i2\pi}{\lambda f}(xs + yt)\right] dxdy \quad (1)$$

where λ is the wavelength and f is the lens focal length. Following Bryngdahl,[3,4] this integral can be approximated by the method of stationary phase[5] to show that a point (x, y) in the input is mapped to a point (s, t) in the output where

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} \frac{\lambda f}{2\pi}\frac{\partial\phi(x, y)}{\partial x} \\ \frac{\lambda f}{2\pi}\frac{\partial\phi(x, y)}{\partial y} \end{bmatrix} . \quad (2)$$

The transformation is one-to-one if no two points (x_1, y_1) and (x_2, y_2) are mapped to the same point (s, t). We will consider only one-to-one transformations.

If a desired coordinate transformation is

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} g(x, y) \\ h(x, y) \end{bmatrix} , \quad (3)$$

then the required HOE phase $\phi(x, y)$ is found by solving the system of partial differential equations:

$$\frac{\lambda f}{2\pi}\frac{\partial \phi}{\partial x} = g(x, y)$$
$$\frac{\lambda f}{2\pi}\frac{\partial \phi}{\partial y} = h(x, y) \quad . \tag{4}$$

It will be assumed throughout that the region in the x-y plane to be transformed is simply connected (i.e., that it has no holes) and that the functions g and h are continuously differentiable (i.e., that the first partial derivatives of g and h are continuous). Many coordinate transformations of interest satisfy these assumptions.[1,2,3,4] Under these assumptions, the system of Eq. (4) has a solution if and only if[6]

$$\frac{\partial \phi(x, y)}{\partial y \partial x} = \frac{\partial \phi(x, y)}{\partial x \partial y} \tag{5}$$

which is equivalent to

$$\frac{\partial g(x, y)}{\partial y} = \frac{\partial h(x, y)}{\partial x} \quad . \tag{6}$$

A coordinate transformation is said to be realizable with a single HOE designed using the stationary phase approximation if and only if Eq. (6) is satisfied.

Unfortunately, many coordinate transformations of interest in optical processing[1,2] are not realizable with a single HOE. An example is polar formatting for image formation from spotlight mode SAR data. However, two HOEs in series can be used to perform this coordinate transformation.[2] In this case, the desired transformation is decomposed into two serial transformations, each of which is realizable with a single HOE. [Note that, as described in Ref. 2, phase correction is performed after the first transformation so that the input to the next

transformation is uniphase for a uniphase input. The phase correction of the output of the first transformation can be included in the HOE which performs the second transformation. If the output of the second transformation must be uniphase (for example, for further coherent optical processing), then a third HOE must be used to correct the output phase.]

Given this example, two questions naturally arise: First, can all one-to-one coordinate transformations satisfying the above assumptions of a simply connected region and continuous differentiability be achieved by the use of multiple HOEs in series? Second, if so, what is the minimum number of HOEs required? In this paper, these two questions are investigated and answered. In Section 2, the questions are developed mathematically and expressed as a system of partial differential equations. In Section 3, it is shown that all such transformations can be achieved with at most two HOEs in series. Section 4 discusses orientation preserving and reversing transformations and their relation to the earlier results. Section 5 applies these results to an example. Conclusions and suggestions for further research are presented in Section 6.

2. Mathematical Development

We begin by investigating the conditions under which a transformation:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} s(x, y) \\ t(x, y) \end{bmatrix} \tag{7}$$

for which

$$\frac{\partial s(x, y)}{\partial y} \neq \frac{\partial t(x, y)}{\partial x} , \tag{8}$$

i.e., which is not realizable with a single HOE, can be performed by two transformations,

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix} \tag{9}$$

and

$$\begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} p(u, v) \\ q(u, v) \end{bmatrix} \tag{10}$$

which satisfy

$$\frac{\partial u(x, y)}{\partial y} = \frac{\partial v(x, y)}{\partial x} \quad , \tag{11}$$

$$\frac{\partial p(u, v)}{\partial v} = \frac{\partial q(u, v)}{\partial u} \quad , \tag{12}$$

i.e., each of which is realizable, and for which, of course,

$$\begin{aligned} s(x, y) &= p[u(x, y), v(x, y)] \\ t(x, y) &= q[u(x, y), v(x, y)] \quad . \end{aligned} \tag{13}$$

If this can be done, then the transformation can be performed optically by two HOEs in series.

We begin by observing that a one-to-one, continuously differentiable first transformation [given by Eq. (9)] onto a simply connected region can always be chosen that is realizable [satisfies Eq. (11)] and has a differentiable inverse. (The differentiability of the inverse is used later.) Since the transformations given by Eqs. (7) and (9) are one-to-one, then a second one-to-one transformation [given by Eq. (10)] which satisfies Eq. (13) will always be well-defined. The question then

becomes: under what conditions will this second transformation be realizable [satisfy Eq. (12)]? Clearly, for some choices of the first transformation, the required second transformation will not be realizable. However, we now derive a system of partial differential equations which, if satisfied by the first transformation, guarantees that the second transformation will be realizable.

By applying the chain rule for partial derivatives to Eq. (13), we arrive at a system of equations which can be written as:

$$\begin{bmatrix} \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial p}{\partial u} & \frac{\partial p}{\partial v} \\ \frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} . \tag{14}$$

The individual matrices have the same elements as the Jacobian determinants for each transformation[7] and the equation expresses the fact that the Jacobian matrix of the total transformation is the product of the Jacobian matrices of the individual transformations. It is interesting to note that the realizability conditions given by Eqs. (11) and (12) are equivalent to requiring that the matrices arising from the transformations defined in Eqs. (9) and (10) be symmetric.

Our next step is to solve Eq. (14) for $\partial p/\partial v$ and $\partial q/\partial u$ for substitution into Eq. (12). This can be done by inverting the matrix Eq. (14), i.e., multiplying Eq. (14) from the right by the inverse of the right-hand matrix, performing the matrix multiplications, and choosing the desired terms. The result is:

$$\begin{aligned} D \frac{\partial p}{\partial v} &= \frac{\partial s}{\partial y}\frac{\partial u}{\partial x} - \frac{\partial s}{\partial x}\frac{\partial u}{\partial y} \\ D \frac{\partial q}{\partial u} &= \frac{\partial t}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial t}{\partial y}\frac{\partial v}{\partial x} \end{aligned} \tag{15}$$

where

$$D = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} \quad . \tag{16}$$

D is the Jacobian determinant for the first transformation. We now use the assumed differentiability of the inverse of the first transformation. Since D is never equal to zero for a transformation with a differentiable inverse,[8] Eq. (15) and similar expressions for $\partial p/\partial u$ and $\partial q/\partial v$ show that the second transformation is continuously differentiable.

We can now state our first important result: For the total transformation to be performed as a series of two realizable transformations [i.e., for Eqs. (11), (12), and (13) to be satisfied], a necessary condition is that the first transformation in the series satisfies the system of partial differential equations:

$$\begin{aligned} &\frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} \\ &\frac{\partial t}{\partial x}\frac{\partial v}{\partial y} = \frac{\partial t}{\partial y}\frac{\partial v}{\partial x} + \frac{\partial s}{\partial y}\frac{\partial u}{\partial x} - \frac{\partial s}{\partial x}\frac{\partial u}{\partial y} \quad . \end{aligned} \tag{17}$$

The first equation is just Eq. (11) and the second is obtained from Eq. (12) by using Eq. (15) and the fact that D is never equal to zero. In addition, the first transformation must be one-to-one, continuously differentiable, onto a simply connected region, and have a differentiable inverse. Together, the two transformations in series make it possible to perform with two HOEs, under the stationary phase approximation, the total transformation that was not realizable with a single HOE.

3. Existence of Solutions to the Partial Differential Equation System

The usefulness of the result of Section 2 depends on the existence, for general transformations as specified by Eq. (7), of invertible solutions to the system of partial differential equations (PDEs) given in Eq. (17). Previous work has shown that solutions always exist to linear hyperbolic PDE systems in two independent variables.[9] We therefore find a condition for which the PDE system is hyperbolic. We do not explicitly find a general solution to the system of Eq. (17).

In canonical form, Eq. (17) becomes:[10]

$$\begin{aligned} &\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} = 0 \\ &\frac{\partial v}{\partial y} + \left[\frac{\partial t}{\partial x}\right]^{-1}\left\{-\frac{\partial s}{\partial y}\frac{\partial u}{\partial x} + \left[\frac{\partial s}{\partial x} - \frac{\partial t}{\partial y}\right]\frac{\partial v}{\partial x}\right\} = 0 \quad . \end{aligned} \tag{18}$$

The classification of the PDE system is determined by the eigenvalues $\lambda(x, y)$ which are the roots of the characteristic equation:

$$\begin{vmatrix} -\lambda & -1 \\ -\left[\frac{\partial t}{\partial x}\right]^{-1}\frac{\partial s}{\partial y} & \left[\frac{\partial t}{\partial x}\right]^{-1}\left[\frac{\partial s}{\partial x} - \frac{\partial t}{\partial y}\right] - \lambda \end{vmatrix} = 0 \quad . \tag{19}$$

Solving Eq. (19) gives:

$$\lambda(x, y) = \left[2\frac{\partial t}{\partial x}\right]^{-1}\left\{\frac{\partial s}{\partial x} - \frac{\partial t}{\partial y} \pm \left[\left(\frac{\partial s}{\partial x} - \frac{\partial t}{\partial y}\right)^2 + 4\frac{\partial s}{\partial y}\frac{\partial t}{\partial x}\right]^{1/2}\right\} \quad . \tag{20}$$

The PDE system will be hyperbolic if the eigenvalues are real and distinct for all x and y in the region of the coordinate transformation. This is the case when the expression within the square root in Eq. (20) is greater than zero. This expression can be manipulated to give:

$$\left(\frac{\partial s}{\partial x} - \frac{\partial t}{\partial y}\right)^2 + 4\,\frac{\partial s}{\partial y}\,\frac{\partial t}{\partial x} = \left(\frac{\partial s}{\partial x} + \frac{\partial t}{\partial y}\right)^2 + 4\left(\frac{\partial s}{\partial y}\,\frac{\partial t}{\partial x} - \frac{\partial s}{\partial x}\,\frac{\partial t}{\partial y}\right) \quad . \tag{21}$$

We now need to require only that the second term be greater than zero. A sufficient condition for real and distinct eigenvalues is therefore:

$$\begin{vmatrix} \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{vmatrix} < 0 \quad . \tag{22}$$

The condition of Eq. (22) is very important because, if it is satisfied, then the PDE system is hyperbolic. For a hyperbolic system in two independent variables, it is known that an infinite family of solutions always exists.[9] This is proven by developing a series solution and showing that it converges uniformly. We assume that, for the one-to-one, continuously differentiable coordinate transformations onto simply connected regions which are of interest here, at least one member of this family will have an inverse that is also differentiable.

In summary, we have shown that the total transformation can be decomposed into two transformations in series if a solution to a related PDE system [Eq. (18)] exists. This PDE system will be hyperbolic, and therefore an infinite family of solutions will exist, if the total transformation [defined in Eq. (7)] has a Jacobian determinant which is negative everywhere in the region to be transformed. (It will be shown

in the next section that this result can be trivially extended to total transformations for which the determinant is always positive.)

It should be noted that, if $\partial t/\partial x = 0$, then this approach cannot be used. (This means that t is a function of y only and that data on any line parallel to the x-axis will, after transformation, lie on a line parallel to the s-axis.) However, if $\partial s/\partial y \neq 0$, then a second approach is possible which begins by dividing Eq. (17) by $\partial s/\partial y$. If both $\partial t/\partial x$ and $\partial s/\partial y = 0$, then the desired transformation satisfies the realizability condition and there is no need to decompose it into two serial transformations.

4. Orientation Preserving and Reversing Transformations

To understand the results of Section 3, we need to define orientation preserving and orientation reversing transformations. A transformation is said to be orientation preserving if a point traversing a closed curve clockwise in the original coordinate system also traverses the transformed curve clockwise.[11] A transformation is orientation reversing if the point traverses the transformed curve counterclockwise. Orientation preserving transformations are like ordinary images while orientation reversing transformations are like mirror images.

It is also true that a transformation is orientation preserving if the Jacobian determinant is positive and orientation reversing when the determinant is negative.[11] By Eq. (22), we require the determinant to be negative for the resulting PDE system to be hyperbolic and therefore have a solution, so we are in effect requiring that the transformation be orientation reversing.

This seeming restriction to transformations which are orientation reversing can easily be relaxed. An orientation preserving transformation can be made to be orientation reversing simply by reversing the

direction of one of the output coordinates (thus making it a mirror image). In an optical processor, this usually means just redefining the output data coordinates. If, for some reason, the coordinates cannot be redefined, then a third coordinate transformation which turns the orientation preserving transformation into an orientation reversing one is necessary. One possible transform is:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ -y \end{bmatrix} \tag{23}$$

which, of course, is realizable with a single HOE. It should also be noted that the requirement for an orientation reversing transformation has only been shown to be a sufficient, not a necessary, condition for the existence of a solution to the PDE system.

5. Illustrative Example

To better understand the above results, it may be helpful to consider a specific example. For polar formatting, the desired coordinate transformation is:[2]

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} x \cos y \\ x \sin y \end{bmatrix} \tag{24}$$

where x-y and s-t are both orthogonal coordinates. This transformation performs polar formatting because, if we identify the r and θ of polar coordinates with x and y, respectively, then r and θ are rectilinear coordinates in the x-y system, but polar coordinates in the s-t system. The transformation functions are:

$$\begin{aligned} s(x, y) &= x \cos y \\ t(x, y) &= x \sin y \quad . \end{aligned} \tag{25}$$

The region of transformation has to be restricted to $x > 0$ or $x < 0$ for the transformation to be one-to-one.

Upon presentation of a coordinate transformation such as this, the first step is to determine whether or not Eq. (6) is satisfied. The partial derivatives,

$$\begin{aligned} \frac{\partial s(x, y)}{\partial y} &= -x \sin y \\ \frac{\partial t(x, y)}{\partial x} &= \sin y \quad , \end{aligned} \tag{26}$$

are not equal and, hence, the transformation is not realizable with a single HOE.

The second step is to determine if it is orientation reversing (i.e., if Eq. (22) is satisfied). Again, it is easily shown that it is not. Therefore, to make use of the results of Sections 2 and 3, the desired transformation functions can be redefined as:

$$\begin{aligned} s(x, y) &= x \cos y \\ t(x, y) &= -x \sin y \quad . \end{aligned} \tag{27}$$

(Other choices are also possible.) The minus sign in Eq. (27) is used so that Eq. (22) is satisfied, and the transformation is orientation reversing, for $x > 0$. The only physical effect of the minus sign is to invert one of the axes in the output plane.

Using Eq. (25), the canonical form [Eq. (18)] for this example is:

$$\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} = 0 \tag{28a}$$

$$\frac{\partial v}{\partial y} + \left[- \sin y\right]^{-1} \left\{x \sin y \frac{\partial u}{\partial x} + \left[(x + 1) \cos y\right] \frac{\partial v}{\partial x}\right\} = 0 \quad . \tag{28b}$$

A solution of Eq. (28) can be found by assuming that $u = g(x)$ and $v = h(y)$. Then Eq. (28a) is satisfied trivially and Eq. (28b) becomes:

$$\frac{dh}{dy} - x\,\frac{dg}{dx} = 0 \quad . \tag{29}$$

Eq. (29) is easily solved and the solutions give a first transformation of the form:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} c \ln x + d \\ cy + e \end{bmatrix} \tag{30}$$

where c, d, and e are constants. This transformation is one-to-one, continuously differentiable, onto a simply connected region, and has a differentiable inverse. Using Eqs. (27) and (30), the second transformation is:

$$\begin{bmatrix} u \\ v \end{bmatrix} \rightarrow \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} \exp u \cos v \\ -\exp u \sin v \end{bmatrix} \tag{31}$$

where the constants have been determined to have the values $c = 1$, $d = e = 0$. Note the minus sign that comes from Eq. (27) and makes the total transformation orientation reversing. The polar formatting transformation has therefore been separated into two transformations performed in series that are realizable [i.e., that satisfy Eqs. (11), (12), and (13)]. These transformations have been implemented with HOEs and performed as predicted.[2]

Note that this example was solved by trying a first coordinate transformation which was separable, i.e., u was a function of x only and v a function of y only. This method may be useful in designing HOEs for other transformations. As another aid to design, it is important to note that a system, designed to perform a specific coordinate

transformation, will perform the inverse transformation if the propagation direction of the light is reversed. In this case, the phase correction HOEs discussed in Section 1 and the designed coordinate transformation HOEs simply reverse their roles.

6. Conclusion

Because of the need to perform data transformations optically, the general problem of one-to-one coordinate transformations realizable with multiple HOEs was investigated. Subject to some requirements on the transformation that can be met in most, perhaps all, practical cases, any one-to-one, continuously differentiable transformation over a simply connected region can be performed with at most two HOEs designed using the stationary phase approximation. One design method was described that was useful for the polar coordinate transformation and may be useful in other cases as well. However, further research is necessary to develop a method for designing the two HOEs that would be applicable in all cases. Another problem of interest is the general theory of many-to-one and one-to-many optical transformations. Of possible mathematical interest would be the development of a proof that at least one member of the infinite family of solutions to a hyperbolic system of PDEs has a differentiable inverse.

This research was supported by the U.S. Army Research Office.

References

1. J.N. Cederquist and A.M. Tai, "Computer-Generated Holograms for Geometric Transformations," Appl. Opt. 23, 3099-3104 (1984).

2. J.N. Cederquist, M.T. Eismann, and A.M. Tai, "Holographic Polar Formatting and Real-Time Optical Processing of Synthetic Aperture Radar Data," Appl. Opt. (submitted 1989).

3. O. Bryngdahl, "Optical Map Transformations," Opt. Commun. 10, 164-168 (1974).

4. O. Bryngdahl, "Geometrical Transformations in Optics," J. Opt. Soc. Am. 64, 1092-1099 (1974).

5. M. Born and E. Wolf, Principles of Optics (Pergamon, New York, 1959), pp. 749-750.

6. W. Kaplan, Advanced Calculus (Addison-Wesley, Reading, Mass., 1959), pp. 243-249.

7. W. Kaplan, op. cit., pp. 90-97.

8. M. Spivak, Calculus on Manifolds (Benjamin, Menlo Park, Calif., 1965), pp. 34-39.

9. I.G. Petrovsky, Lectures on Partial Differential Equations (Interscience, New York, 1954), pp. 67-72.

10. E.C. Zachmanoglou and D.W. Thoe, Introduction to Partial Differential Equations with Applications (Dover, New York, 1986), pp. 361-362.

11. A.E. Taylor, Advanced Calculus (Ginn, Boston, 1955), pp. 267-271, 428-431.

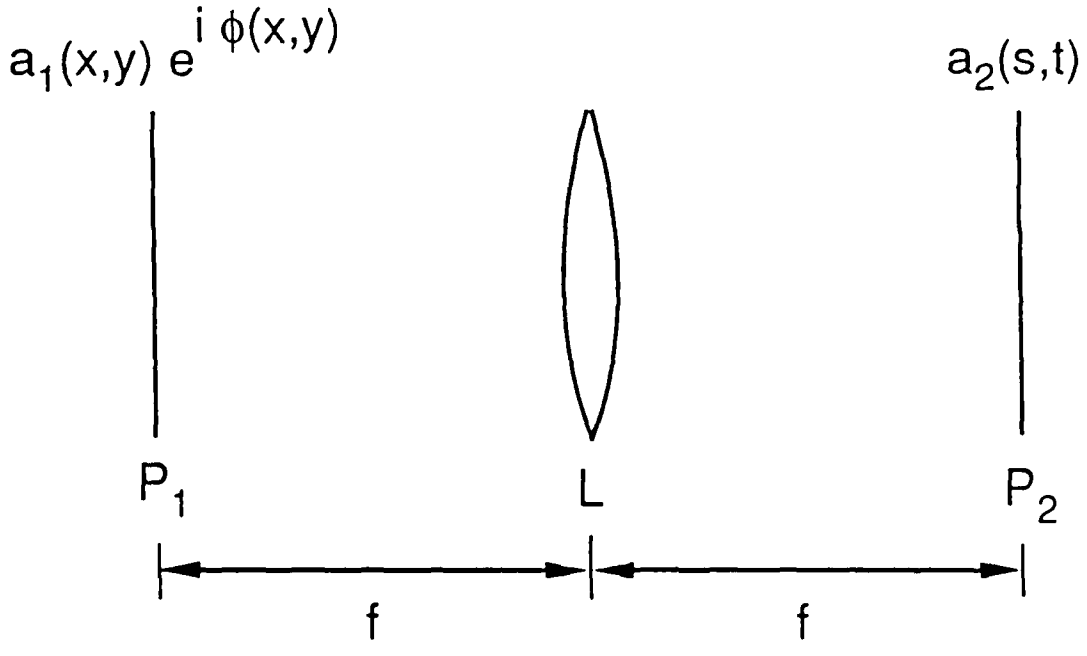


Fig. 1. Optical system for coordinate transformation. The lens L has focal length f and Fourier transforms plane P_1 to plane P_2.

APPENDIX D

CGH FABRICATION TECHNIQUES AND FACILITIES

(in Computer-Generated Holography II,
Proceedings of the SPIE, Volume 884,
S.H. Lee, ed., Los Angeles, January 1988, pp. 40-45)

CGH fabrication techniques and facilities

J.N. Cederquist, J.R. Fienup, and A.M. Tai

Optical Science Laboratory, Advanced Concepts Division
Environmental Research Institute of Michigan
P.O. Box 8618, Ann Arbor, Michigan 48107

ABSTRACT

Computer-generated holograms (CGHs) are used in a number of important optical technology application areas such as holographic optical elements, optical processing and computing, optical testing, image and information display, beam forming, and beam scanning. Many different CGH fabrication devices (e.g., laser beam scanners, electron beam writers) and facilities have been developed and are in use. However, none of these devices ideally suit the requirements (e.g., resolution, space-bandwidth product, recording material) of many CGH applications. Furthermore, the access of many researchers to these facilities is limited and the technical support available is often poor. New facilities specifically designed for CGH fabrication would better serve the needs of the CGH research and development community. The requirements for a successful CGH fabrication facility including appropriate technical support for users are established.

1. INTRODUCTION

The possibility of a national facility for the fabrication of computer-generated holograms (CGHs) has been considered for the past decade. Some of the possible benefits of a national CGH facility are: (1) it would fabricate high quality CGHs for use by the research, development, and manufacturing communities, (2) it would avoid further expensive, needless duplication of facilities, (3) it would speed the development and application of diffractive optics techniques and devices, and (4) it would improve system performance in several important application areas (e.g., holographic optical elements, optical interconnect devices for high speed parallel processors, and invariant correlation filters for object detection and recognition).

Although many facilities are currently fabricating CGHs of various types, all of these facilities suffer from one or more of the following difficiencies: (1) performance limitations (e.g., resolution, space-bandwidth product, or recording speed), (2) cost, (3) CGH encoding or materials restrictions (e.g., restriction to binary amplitude or phase, or restriction to recording fringes), (4) data format inefficiencies, and (5) lack of technical support in CGH design and optics in general. Clearly, no existing facility can yet be considered to be acceptable as a national CGH fabrication facility.

This paper makes an effort to take the next step toward a better, and possibly national, CGH fabrication facility. A broad overview of CGH applications and technology is given in Section 2. CGH fabrication methods and devices are discussed in Section 3. Finally, in Section 4, the requirements which a successful CGH fabrication facility must meet are given and an in-depth study of user requirements is recommended.

2. REVIEW OF CGH TECHNOLOGY

The range of applications which can potentially take advantage of CGH is very wide. Some of the applications are shown in Figure 1. They have been grouped according to whether the CGH is used to modify only the phase of a wavefront or to modify both amplitude and phase.

CGH can be classified in many ways. Two possible methods are by material type and by type of diffraction effect utilized. The different types of CGH under each of these categories are shown in Figure 2.

CGH APPLICATIONS

Phase-only wavefront modification
- Conventional optics
  - HOE - lens, mirrors, phase correctors
  - Testing of aspherics
  - Wavefront creation for COHOE fabrication
  - Scanners
  - Heads-up/helmet displays
- Light redistribution
  - Beam intensity profile shaping
  - Data formatting - coordinate transformation
  - Digital data processor interconnects
  - Beam combiners, dividers

Amplitude and phase wavefront modification
- Displays
  - 2-D, 3-D images
  - Sighting devices, reticles
- Spatial filters for pattern recognition
- Data memory, storage

Figure 1. The wide range of CGH applications.

CGH CLASSIFICATION

Material types
- Amplitude-only
  - Continuous
  - Binary
- Phase-only
  - Continuous
  - Binary
- Amplitude and phase

Diffraction effects
- Scalar
  - Planar media and surface relief
  - Volume Bragg diffraction
- Vector

Figure 2. Examples of CGH classification methods.

CGH can be copied in several ways. A contact copy can be made, usually for the purpose of transferring the CGH onto a new, more suitable material. Another possibility is the COHOE or Computer-Originated Holographic Optical Element. In this copying method, the wavefront reconstructed by the CGH is interfered with a second wavefront. For example, if this second wavefront is planar, then the effect is simply to copy the phase of the CGH while adding a high carrier frequency to achieve higher diffraction efficiency or a larger diffraction angle. Another example is for the second wavefront to be spherical and of high numerical aperture and the CGH wavefront to represent a desired phase correction; then the resulting COHOE can be a high numerical aperture aspheric. In the latter example, the space-bandwidth product required of the CGH is reduced by using a nonplanar wavefront generated by conventional optics.

Because of the wide range of applications in which CGHs may be used, there is also a wide range in the performance required of CGHs (see Figure 3). The specifications shown in the low range might apply to a CGH for a Fourier-transform-domain matched filter in a coherent optical correlator and those in the high range to an aspheric holographic combiner for a head-up display. For many applications, the required signal-to-noise ratio has not been adequately analyzed, hence the question marks in the table rather than numerical values. In addition, the diffraction angle required of a CGH can vary from zero to nearly 180° depending on the application. Although CGHs have been most commonly used in the wavelength region from the ultraviolet to the infrared, CGHs have also been made and used at x-ray and microwave wavelengths. It is neither expected or required that any one CGH fabrication device would meet all the specifications shown in Figure 3.

RANGE OF CGH PERFORMANCE SPECIFICATIONS

	Range	
	Low	High
Resolution	100 μm	0.05 μm
Space-bandwidth product	$10^2 \times 10^2$	$10^6 \times 10^6$
Size	$(1\ mm)^2$	$(50\ cm)^2$
Diffraction efficiency	<1%	~100%
Phase accuracy	λ	$\lambda/50$
Signal-to-noise ratio	?	?

Figure 3. The wide range of CGH specifications for different applications.

There are also a wide variety of methods for encoding the desired two-dimensional distribution of amplitude and phase in a CGH. The major encoding methods are shown in Figure 4. Each method has its own advantages and disadvantages. For the case of CGH used for two-dimensional image reconstruction, the inaccuracies inherent in most of these methods have been analyzed.

MAJOR CGH ENCODING METHODS

- Amplitude and phase directly - ROACH
- Phase directly - Kinoform
  - Continuous
  - Binary
- Phase on carrier
  - Amplitude transmittance
    - Fringes
      - Continuous - Burch
      - Binary - Lee
    - Detour-phase
      - Continuous - Burckhardt and Lee
      - Binary - Lohmann
  - Phase transmittance
    - Phase versions of above

Figure 4. A classification of the major CGH encoding methods.

As has been the case in each CGH characteristic examined in this section, the range of CGH recording media is also very wide. Some of the principal types are listed in Figure 5. It is important to note that, even when the CGH can not be recorded in a material suitable for use in the application, it can usually be replicated in a suitable material. Not included in this list are the various real-time spatial light modulators that can be used to record CGHs. They are ordinarily restricted to uses in which the hologram must be recorded in near real-time, and are therefore not appropriate for a CGH facility. (They also typically have modest performance specifications).

CGH RECORDING MEDIA

- Photographic emulsions
  - Amplitude
  - Bleached
- Gelatin
  - Silver halide
  - Dichromated
- Photoresist, electron resist
- Photopolymers
- Multiemulsion - ROACH
- Machinable materials

- Replication materials
  - Above plus:
    - Glass, quarts
    - Si, GaAs, Ge
    - Metal

Figure 5. Examples of CGH recording and replication media.

3. CGH FABRICATION DEVICES

A large number of devices have been used to fabricate CGH. The principle methods and devices are listed in Figure 6. In addition to the categories listed in Figure 6, the devices can be divided into two classes: (1) those which were built for another purpose and were adapted for CGH fabrication and (2) those which have been designed and built specifically for CGH fabrication. Plotters, printers, rotating drum scanners, some translating flat bed scanners, optical pattern generators for microelectronic mask making, electron beam writers, and computer-controlled milling machines are members of the first class. Although useful CGH have been made with all these devices, better performance (e.g., smaller writing spot size, greater space-bandwidth product, more appropriate data formatting and CGH encoding) could be obtained if the device were redesigned specifically for CGH fabrication. Galvanometer scanners, devices which image a CRT to the recording medium, and devices which interfere two computer-controlled wavefronts belong to the second class. Although some of these devices are excellent for fabricating specific types of holograms, they all have performance limitations which make them unsuitable for fabricating a wide variety of CGH types.

CGH FABRICATION METHODS AND DEVICES

- Drawing
  - Plotter
  - Printer
- Optical writing
  - With one beam
    - Galvanometer
    - Rotating drum
    - Translating flat bed - microdensitometer
    - Optical pattern generator
  - With two beams
    - Interference of two wavefronts
  - With CRT image
- Electron beam writing
- Machining of surface profile

Figure 6. The major CGH fabrication methods and representative device types.

For example, consider devices which use a laser beam to write on a photosensitive film placed on a rotating drum. These devices have the advantages of (1) continuous modulation of the intensity of the laser beam allowing a continuous amplitude recording, (2) being relatively inexpensive to purchase and maintain, (3) fast recording speed, (4) potentially high space-bandwidth product, and (5) moderate resolution. Disadvantages are (1) positional inaccuracy can cause phase errors, (2) current use with photographic film produces a CGH with (a) phase errors due to thickness variations (correctable by mounting in an index matching fluid) and (b) low diffraction efficiency (correctable by copying onto a high efficiency phase material), and (3) restriction to a raster exposure pattern which is not optimum for writing CGH which consist of curved fringe patterns.

The currently popular electron beam recorder which is used for mask generation for microelectronics is another example. It has the advantages of (1) high (submicron) resolution, (2) high space-bandwidth product, and (3) allowing optics to build on microelectronics technology. However, current use of this device suffers from several disadvantages: (1) the recording is principally binary, (2) the devices and their maintenance are expensive, (3) recording speed is only moderate to low, and (4) much software development is need to allow easier user access, to permit efficient data formating, and to make use of the random-access ability to write curved CGH fringe patterns.

4. CGH FACILITY REQUIREMENTS

The brief review of CGH applications and technology given in Section 2 indicates that multiple issues are involved in the choice of a CGH fabrication method and in the design of a CGH fabrication facility. As indicated by the examples of fabrication devices given in Section 3, all current facilities have some difficiencies. Designers and builders of future facililites should consider the CGH facility requirements summarized in Figure 7. First, the projected facility must offer high performance levels. High performance is not sufficient for success by itself, however. The facility must also have the flexiblility to accomodate the various encoding methods and recording materials required by different applications. Finally, the facility hardware, software, and technical and management staff must be user-oriented. Only if all three of these requirements are met will it be asked to produce a sufficient volume of CGHs to justify the financial investment in a new CGH fabrication facility.

REQUIREMENTS FOR A SUCCESSFUL CGH FACILITY

- High performance
  - High space-bandwidth product
  - High resolution
  - Low errors
  - Fast recording
- Flexibility to accomodate various
  - Encoding methods
  - Recording media
- User-oriented
  - Flexible CGH description
  - Low cost
  - Reasonable delivery time
  - Technical support

Figure 7. The three-fold requirements for a successful CGH facility.

It is not possible at this time to determine which fabrication method (or methods), hardware, software, and support structure should be chosen for a new CGH fabrication facility. The CGH performance requirements of potential users are complex and require more study. When this has been accomplished, a trade-off study can be performed in an effort to select a CGH fabrication method and facility that best meets user community needs.

5. ACKNOWLEDGEMENT

This work was supported in part by the U.S. Army Research Office.