Electro-optic hologram generation on spatial light modulators

Joseph Rosen, Lior Shiv, Jeremy Stein, and Joseph Shamir

Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel

Received November 25, 1991; revised manuscript received February 24, 1992; accepted February 27, 1992

An iterative method for generating holograms on spatial light modulators is based on measuring the reconstructed complex image and on-line correction of the hologram. Apart from recording complex amplitude distributions, the new procedure may become a useful tool for applications such as adaptive optics and reconfigurable interconnection networks.

1. INTRODUCTION

Binary computer-generated holograms[1] (BCGH's) are becoming essential components in optical signal processing schemes. Apart from image reconstruction, these holograms can serve as spatial filters, optical elements, and the basic component in sophisticated interconnection networks.[2] The present paper deals with BCGH's presented on spatial light modulators (SLM's), which provide real-time variability. The enhanced flexibility of holograms presented on SLM's makes them suitable for applications such as adaptive optical elements and reconfigurable interconnects.

Many techniques for designing a BCGH are known.[3,4] These include the iterative computing of BCGH's[5,6] that offers higher light efficiency and more accurate reconstructed images.[5] The conventional procedure is to perform all the calculations on digital computers and use the results for preparing the hologram.[4-6] A recently proposed method is to implement the iterative process directly on a SLM, employing an electro-optic hybrid system.[7] The advantage of this method is that it takes into account the actual characteristics of the system. No physical approximations are needed during the calculations in a computer, and the optical aberrations, the SLM phase distortions, and other system defects are automatically taken into account. The electro-optic system also obviates slow computer calculations, such as the Fourier transform (FT), and hence the direct way becomes efficient when a large number of pixels in the BCGH are considered.

This paper describes a new and sophisticated procedure for the iterative generation of holograms on SLM's. Whereas in our previous study[7] only the intensity distribution was reconstructed, the present paper records and reconstructs the complete complex amplitude distribution. Although the basic principles were proposed in Ref. 7, the actual implementation required some modifications and a more advanced optimization algorithm. In the study reported in Ref. 7 the BCGH was calculated by the direct binary search (DBS) algorithm.[5] The DBS bypasses the need for a FT every iteration, and thus it is efficient for digital computer algorithms. However, the FT is not the bottleneck of the electro-optic system, and thus the DBS becomes unnecessary. In this paper we introduce an alternative algorithm that is faster and more efficient than the DBS.

After a brief discussion of the background for this study in Section 2, the theoretical analysis is given in Sections 3–5, and the optimization procedure is described in Section 6. The experimental results are shown in Section 7, and concluding remarks are given in Section 8.

2. GENERAL BACKGROUND

A generalized schematic of a hybrid system is shown as Fig. 1. A BCGH is displayed on the SLM at the input plane to a linear optical system. The SLM is illuminated by a monochromatic plane wave, and the reconstructed image is measured over the output plane by using a CCD camera. Our goal is to find the BCGH that produces the desired image in a given region of the output plane with a minimal error. This BCGH has to be created directly, on line, in the same optical system. Since this concept is applicable to any linear optical system the BCGH can be a Fresnel hologram, a FT hologram, or any other kind of hologram. To keep this paper within reasonable bounds we consider mainly FT holograms.

The iterative process for generating the BCGH includes, in general, the following steps. First, an initial BCGH is displayed on the SLM, and the CCD camera detects the intensity of the reconstructed wave front at the output plane. The detected intensity distribution is compared with a pattern stored in the computer memory with the help of a cost function that quantifies the deviation from the desired pattern. The value of the cost function is computed and is used to make a decision and to change the displayed BCGH for the next iteration. This process proceeds, iteratively, until the cost function decreases below some desired minimal value that indicates an acceptable error in the reconstruction.

If we look for a BCGH that yields only a desired intensity distribution at the output, the above description is almost the entire story. There is only the need to define the nature of the linear system, to pick the desired image, and to find a reasonably efficient searching algorithm for the iterative process. The process becomes much more complicated, however, if the aim is to find a BCGH that

0740-3232/92/071159-08$05.00

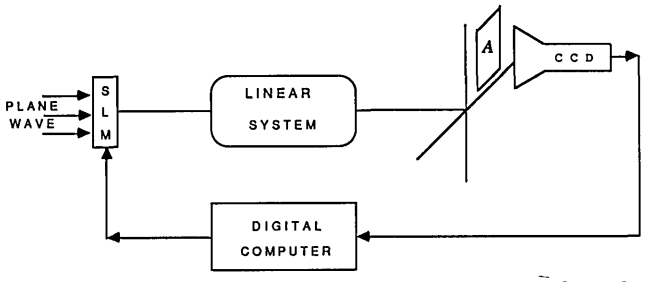

Fig. 1. Experimental system. The pattern generated on the SLM is transformed by the linear system and observed by the CCD.

produces a desired complex amplitude distribution. Since ordinary detectors are sensitive to the light intensity only, we have to use some indirect method to measure the phase distribution of the reconstructed image.

Suppose that the BCGH is designed to reconstruct the complex amplitude $f(x,y)$. We denote the SLM transmittance by the binary function $H(u,v)$ and the point-spread function of the linear system by $S(u,v,x,y)$, and the output distribution of the system is given by

$$h(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(u,v)S(u,v,x,y)\mathrm{d}u\mathrm{d}v. \tag{1}$$

The aim is to obtain, within a given region $\mathcal{A}$ of the output plane, an image $\tilde{f}(x,y)$ as close as possible to $f(x,y)$. For this purpose, we may look for the minimum of the mean-square error that serves as the cost function[5]:

$$e = \iint_{\mathcal{A}} |f(x,y) - \gamma\tilde{f}(x,y)|^2\mathrm{d}x\mathrm{d}y, \tag{2}$$

where γ is a real constant and

$$\tilde{f}(x,y) = w(x,y)h(x,y),$$

where

$$w(x,y) = \begin{cases} 1 & \text{if } x \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

Since the detector array can measure only the intensity distribution that is proportional to $|\tilde{f}(x,y)|^2$, the error of Eq. (2) cannot be directly calculated from the measurements. In Section 3 we review some modifications that make the calculation of the complex amplitude distribution possible from the intensity measurements.

3. RECONSTRUCTION OF COMPLEX AMPLITUDE DISTRIBUTIONS

A conventional procedure to obtain the complete information of a complex amplitude from the intensity distribution is to use interferometric techniques. In the present case the distribution $h(x,y)$ may be superposed by a plane wave such that the intensity distribution within the region $\mathcal{A}$ is given by

$$\tilde{I}_{\mathcal{A}}(x,y) = A^2 + |\tilde{f}(x,y)|^2 + A|\tilde{f}(x,y)| \times \cos[xk\sin\theta + \tilde{\varphi}(x,y)], \tag{4}$$

where k is the wave number of the illuminating light, θ is the angle between the **k** vector and the optic axis (it is assumed, for simplicity, that the **k** vector has no component in the y direction), A is the amplitude of the plane wave, and $\tilde{\varphi}(x,y)$ is the phase distribution of $\tilde{f}(x,y)$. $\tilde{I}_{\mathcal{A}}(x,y)$ is compared with a memory-stored function of the form

$$I_{\mathcal{A}}(x,y) = A^2 + |f(x,y)|^2 + A|f(x,y)| \times \cos[xk\sin\theta + \varphi(x,y)]. \tag{5}$$

The cost function, redefined for this case,

$$e_I = \iint_{\mathcal{A}} |I_{\mathcal{A}}(x,y) - \gamma\tilde{I}_{\mathcal{A}}(x,y)|^2\mathrm{d}x\mathrm{d}y, \tag{6}$$

can now be minimized. The complex amplitude distribution of the reconstructed image, $\tilde{f}(x,y)$, is close to the desired image, $f(x,y)$, if all the parameters, A, k, and θ, are equal in both distributions $I_{\mathcal{A}}(x,y)$ and $\tilde{I}_{\mathcal{A}}(x,y)$.

Unfortunately, this interferometric procedure is quite problematic. First, the angle θ and the constant A have to be known exactly and have to be stored in the computer together with $f(x,y)$. Second, the cost function of Eq. (6) becomes sensitive to aberrations of the plane wave in addition to the difference between $\tilde{f}(x,y)$ and $f(x,y)$. Finally, the notable drawback of the interferometric method is an increased bandwidth of the measured signal $\tilde{I}(x,y)$ as compared with $\tilde{f}(x,y)$, by a factor exceeding 2.

An alternative technique, based on two intensity measurements (TIM), was proposed in Ref. 7. This method does not suffer from the above-mentioned disadvantages. In the TIM method, in addition to the measurement of the intensity distribution, $|\tilde{f}(x,y)|^2$, the intensity of the cross correlation between the two functions $\tilde{f}(x,y)$ and $f(x,y)$ is measured at the origin. This is the inner product between the two functions. We will show that these two separate intensity measurements are sufficient for calculating a cost function that expresses the distance between a reconstructed complex amplitude distribution and a desired complex function.

The cost function of Eq. (2) can be rewritten in the form

$$e = \iint_{\mathcal{A}} |f(x,y)|^2\mathrm{d}x\mathrm{d}y - 2\gamma\cos\phi\left|\iint_{\mathcal{A}} f^*(x,y)\tilde{f}(x,y)\mathrm{d}x\mathrm{d}y\right| + \gamma^2\iint_{\mathcal{A}} |\tilde{f}(x,y)|^2\mathrm{d}x\mathrm{d}y, \tag{7}$$

where ϕ is the phase of the inner product between $f(x,y)$ and $\tilde{f}(x,y)$. If we use the γ proposed in Ref. 5 for minimizing the error in Eq. (2),

$$\gamma_0 = \frac{\cos\phi\left|\iint_{\mathcal{A}} \tilde{f}(x,y)f^*(x,y)\mathrm{d}x\mathrm{d}y\right|}{\iint_{\mathcal{A}} |\tilde{f}(x,y)|^2\mathrm{d}x\mathrm{d}y}, \tag{8}$$

we have

$$e_0 = \iint_{\mathcal{A}} |f(x,y)|^2\mathrm{d}x\mathrm{d}y - \frac{\cos^2\phi\left|\iint_{\mathcal{A}} \tilde{f}(x,y)f^*(x,y)\mathrm{d}x\mathrm{d}y\right|^2}{\iint_{\mathcal{A}} |\tilde{f}(x,y)|^2\mathrm{d}x\mathrm{d}y}. \tag{9}$$

The numerator of the second term contains the measured inner product, while the denominator is an integration

over the intensity distribution of $\tilde{f}(x, y)$. Unfortunately this cost function can be measured only for real functions when it is known that ϕ is zero or π. Nevertheless, the two intensity measurements contain all the information about the reconstructed image, and the only difficulty is to find a proper cost function that can be measured.

Consider, for instance, the function

$$e_c = \iint_{\mathcal{A}} |f(x, y)|^2 \mathrm{d}x\mathrm{d}y - \frac{\left| \iint_{\mathcal{A}} \tilde{f}(x, y) f^*(x, y) \mathrm{d}x\mathrm{d}y \right|^2}{\iint_{\mathcal{A}} |\tilde{f}(x, y)|^2 \mathrm{d}x\mathrm{d}y}. \quad (10)$$

From the Schwartz inequality we have that, for any pair of complex functions, $e_c = 0$ when and only when $\tilde{f}(x, y)$ is proportional to $f(x, y)$; otherwise $e_c > 0$. Moreover, e_c becomes identical to e_0 of Eq. (9) when both $\tilde{f}(x, y)$ and $f(x, y)$ are real. Thus this cost function is close to the mean-square error for real functions and is reasonable for complex functions too. Since the first term is always a constant, the minimization of e_c is identical to the minimization of

$$e_a = \frac{\iint_{\mathcal{A}} |\tilde{f}(x, y)|^2 \mathrm{d}x\mathrm{d}y}{\left| \iint_{\mathcal{A}} \tilde{f}(x, y) f^*(x, y) \mathrm{d}x\mathrm{d}y \right|^2}. \quad (11)$$

This is the cost function used for the present investigation.

Learning systems that implement the TIM method are depicted in Figs. 2 and 3. To generate a FT hologram, the two-channel linear system shown in Fig. 2 is suitable. The hologram is to be generated on the SLM, and one of the channels is used to reconstruct $\tilde{f}(x, y)$ in the region $\mathcal{A}$ on plane P_3. The other channel constitutes a 4-f optical correlator. If the function $f^*(-x, -y)$ is introduced in plane P_1, then in plane P_3 the correlation distribution between $\tilde{f}(x, y)$ and $f(x, y)$ is obtained. The value of their inner product can be measured, while the other channel is blocked, at the center of the subregion $\mathcal{A}$ where the peak of the correlation function is usually obtained. The integral over the reconstructed image intensity at the subregion $\mathcal{A}$ is taken in plane P_3, with the shutter (S) blocking the light coming from P_1 while leaving the plane-wave illumination over the SLM. Note that there is no problem

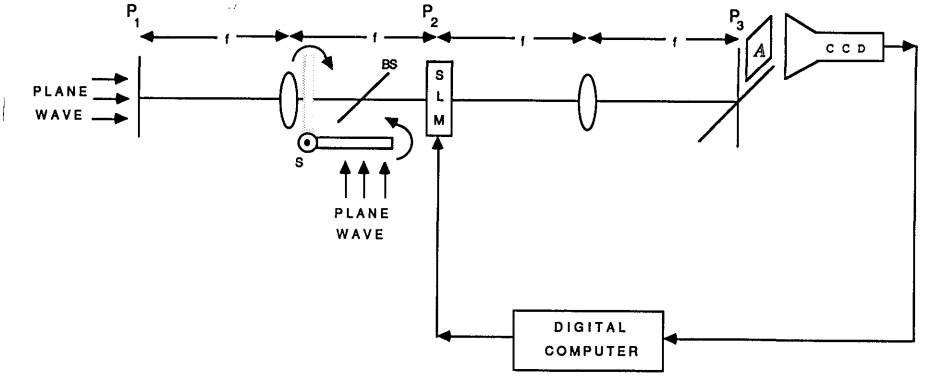

Fig. 2. Experimental system for calculating a Fourier hologram with a control on the complex distribution of the reconstructed image. BS, beam splitter.

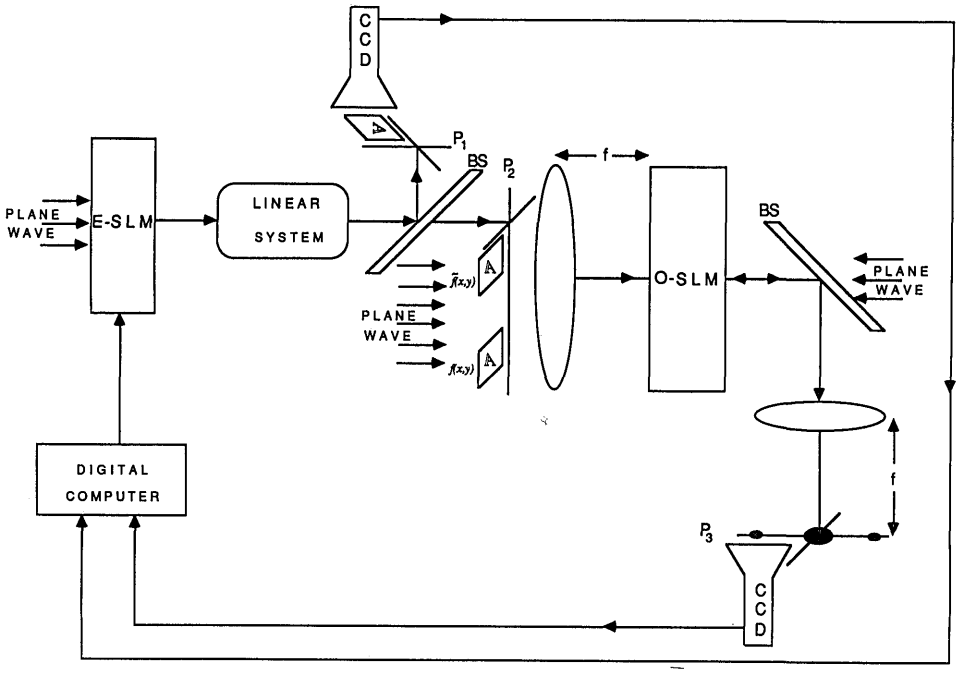

Fig. 3. Schematic illustration of the system for calculating any kind of hologram. BS's, beam splitters.

of alignment between the input spatial spectrum and the filter since the filter is adapted to the input spectrum during the iterative process. If, for some reason, it is too complicated to prepare $f^*(-x,-y)$ in the input plane, it may be replaced by a simpler function, as is shown in Section 4.

After the optimal BCGH is obtained, only one channel needs to be retained. If the hologram is used for information storage we need only the reconstruction channel for information retrieval. On the other hand, if the aim is the synthesis of a FT hologram for spatial filtering, the correlation channel should be kept for signal processing operations.

When the desired hologram is not a FT hologram but, e.g., a Fresnel hologram, we usually need a more complicated system, with that of Fig. 3 being one possibility. In this architecture too, the hologram is displayed on an electrically addressed SLM (E-SLM) and illuminated by a plane wave. The intensity distribution of the reconstructed image is measured at plane P_1 after the wave front passes from the hologram through the linear system. The reconstructed image, $\tilde{f}(x,y)$, is also projected onto plane P_2, where it is joined by $f(x,y)$ to be jointly Fourier transformed in a joint transform correlator.[8] The joint spectrum is projected onto the optically addressed SLM (O-SLM). Illuminating the O-SLM from the readout side by a plane wave yields the correlation distribution at plane P_3 around a point that is displaced from the origin by a distance equal to the gap between $\tilde{f}(x,y)$ and $f(x,y)$ in plane P_2. The value of the correlation peak and the total energy distribution over region $\mathcal{A}$ in plane P_1 are sufficient for calculating the cost function e_a of Eq. (11). To keep the extent of this study reasonable, in what follows we consider only the architecture of Fig. 2 for the generation of a FT hologram.

4. PROCEDURE GENERALIZATION

In some cases it may be inconvenient or even impossible to place the function $f^*(-x,-y)$ into the input plane of the correlation channel. For example, if $f(x,y)$ represents a positive gray-level image, it cannot easily be displayed linearly on a transparency or a SLM. To treat such a situation we may generalize the procedure by replacing $f^*(-x,-y)$ with a more convenient function $g^*(-x,-y)$. If we are to do so, $g(x,y)$ must satisfy the following constraints:

$$\arg\{g(x,y)\} = \arg\{f(x,y)\},$$
$$g(x,y) \neq 0, \qquad \text{for all } x,y \text{ where } f(x,y) \neq 0, \quad (12)$$

i.e., the phase distributions of $g(x,y)$ and $f(x,y)$ are equal. When the inner product between $g(x,y)$ and $\tilde{f}(x,y)$ is maximized and the differences between $|\tilde{f}(x,y)|^2$ and $|f(x,y)|^2$ are minimized, then $\tilde{f}(x,y)$ converges to $f(x,y)$. This statement is acceptable since the maximum value of the inner product between $\tilde{f}(x,y)$ and $g(x,y)$ is obtained if and only if the phase distributions of these two functions are equal and thus equal to that of $f(x,y)$. The convergence of the phase distribution is guaranteed from the correlation channel, while the convergence of the magnitude distribution is considered directly by the comparison of $|\tilde{f}(x,y)|^2$ with $|f(x,y)|^2$. Although the phase and the magnitude are treated separately, taking both of them guarantees the converge of $\tilde{f}(x,y)$ to $f(x,y)$. The cost function that represents the above discussion can be written in the form

$$e_g = \frac{\iint_{\mathcal{A}} \left| |f(x,y)|^2 - \gamma|\tilde{f}(x,y)|^2 \right|^2 \mathrm{d}x\mathrm{d}y}{\left| \iint_{\mathcal{A}} \tilde{f}(x,y) g^*(x,y) \mathrm{d}x\mathrm{d}y \right|^2}. \quad (13)$$

The numerator is the mean-squared error between the reconstructed intensity distribution and the desired image intensity, while the denominator expresses the intensity of the inner product. As above, this cost function contains the two measurements, and its minimization leads to the convergence of $\tilde{f}(x,y)$ toward $f(x,y)$. This constitutes a generalization of the previous procedure since the special case of $g(x,y) = f(x,y)$ satisfies the conditions of expressions (12), and therefore Eq. (13) is a suitable cost function for the case treated in Section 3.

The advantage of using this generalized procedure can be demonstrated with the example indicated above, a positive gray-level image. For this case choosing a simple function such as $g(x,y) = w(x,y)$ satisfies the conditions of expressions (12). Thus minimization of the cost function of Eq. (13) by using just this simple window function yields the desired BCGH.

5. SINGLE-CHANNEL OPERATION

In this section we address the interesting question of what happens when only one of the two channels is operated. The answer is obvious when the intensity distribution channel alone is considered. If the distribution of $|\tilde{f}(x,y)|^2$ is detected and compared with that of $|f(x,y)|^2$ by using the cost function of Eq. (6), the reconstruction of the final BCGH is similar to having the minimum mean-square error to $|f(x,y)|^2$.

Less obvious is what is obtained if only the output of the correlation channel is measured during the process. The answer depends on which cost function is used. Let $\tilde{f}(x,y)$ be the reconstructed image from the BCGH, and let $g^*(-x,-y)$ be the input function of the correlation channel. If we define a cost function by

$$e_p = \left| \iint_{\mathcal{A}} \tilde{f}(x,y) g^*(x,y) \mathrm{d}x\mathrm{d}y \right|^{-2}, \quad (14)$$

the minimum error is achieved when the inner product is in its maximal value that is obtained for

$$\tilde{F}(u,v) = G(u,v)/|G(u,v)|, \quad (15)$$

where $\tilde{F}(u,v)$ and $G(u,v)$ are the FT's of $\tilde{f}(x,y)$ and $g(x,y)$, respectively. The hologram converges such that the FT of the reconstructed image is similar to the phase distribution of the FT of the input function $g(x,y)$. This procedure can thus be used to generate a phase-only filter[9] for pattern recognition purposes.

A different cost function,

$$e_i = \frac{\iint_{\mathcal{A}} \left| \iint_{\mathcal{A}} \tilde{f}(x,y) g^*(x-x', y-y') \mathrm{d}x\mathrm{d}y \right|^2 \mathrm{d}x'\mathrm{d}y'}{\left| \iint_{\mathcal{A}} \tilde{f}(x,y) g^*(x,y) \mathrm{d}x\mathrm{d}y \right|^2}, \quad (16)$$

attains its minimum[10] when

$$\tilde{F}(u,v) = [G(u,v)]^{-1}. \tag{17}$$

The hologram in this case represents the inverse filter of $G(u,v)$. Note that in this special case, in addition to the measurement of the inner product, the denominator of e_i, the whole energy of the correlation distribution appears in the numerator.

In conclusion, operation with the correlation channel alone can be useful for calculating various kinds of spatial filters for pattern recognition.[11,12] However, the above analysis indicates that, when a discriminant filter for classification is needed, we may have to present the patterns of a training set one at a time to make the measurement of the cost function possible. Otherwise the generated filter may be matched to a specific spatial combination of objects rather than to the objects themselves.

6. SEARCHING PROCEDURE

A searching algorithm to implement the procedures discussed above must include two major aspects. First, it must have an acceptance criterion for each iteration. The second aspect is the operation to be done on the BCGH for the next iteration. In the early algorithm of the DBS,[5] a single pixel was changed in every iteration. The policy was to accept every result that decreased the error. If the last trial reduced the error, another pixel of the BCGH was changed for the next iteration. Otherwise, the changed pixel was converted back, and only then was another pixel inverted. This algorithm eliminates the need for performing the FT for every iteration. However, it is a slow searching algorithm, and it does not make any use of the information buried in the trials that fail.

In the experiment of Ref. 7 the algorithm was the DBS with a modified version of the acceptance policy. The problems there were a high noise level and a low reliability of the detected results. In the present algorithm we intend to use all the available information from the measurements, including those results that increase the cost function.

The problem here can be defined as one of binary nonlinear unconstrained optimization, which means that the minimum of the cost function $e(\mathbf{x})$ [defined in Eq. (11), for instance] has to be found with $\mathbf{x}$ being a binary vector. We would like to adapt concepts from the continuous nonlinear unconstrained optimization problem but without converting the binary vector to continuous variables at every step. This would be quite inefficient for dealing with a 128×128 binary array, as was the case in our experiments. The adapted concept in our algorithm got its inspiration from the Nelder–Mead algorithm[13] for continuous variables. It can be summarized as follows: Progress along the course of reduction in the cost function until the cost function increases, and then change course to one with the highest probability of further reduction.

To implement this principle we have to know, at every stage, how much and to where we proceed. That is, we have to define a distance between two vectors and something that indicates a direction between vectors. The distance issue can be resolved by using the Hamming distance, as it is well suited for binary variables. The problem of defining the directions among various vectors is more difficult since we have not found a definition that satisfies the Euclidean rules. Converting the binary variables to their decimal values is possible but complicated and has no physical meaning.

The algorithm starts with a group G of $M (M \geq 1)$ randomly chosen holograms stored in the computer memory. Every hologram from this group is situated at some Hamming distance from the others, and each one is displayed on the SLM in the initial steps of the procedure. As a result of every measurement, each hologram receives a mark according to the value of its cost function.

The searching procedure begins with a randomly chosen hologram denoted by x_0. Every pixel in x_0 may be changed according to a probability distribution determined by the different weights of the M holograms and another weight W_{M+1} that indicates the preference of the pixels to be unchanged. The error function is calculated for the actual hologram x_i by displaying it on the SLM, where the index i indicates the number of the measurement during the process. If the cost function is less than that of one of G, the new hologram becomes a member of G, and we assign new grades to all the holograms in G. The hologram with the highest error becomes the x_{i+1} tested hologram. The trial is repeated while x_{i+1} is modified according to the same partition of probability weights. If the error increases, group G is retained, but the partition of the weights is altered such that the probability weights increase more toward the holograms with the lowest error, and the probability of the pixel to be unchanged decreases. The fact that the x_i hologram did not improve the cost function was included in the determination of the next hologram to be considered for the subsequent iteration.

Our algorithm contained three additional parameters. If after K trials the cost function for x_i was not improved, then a new x_{i+1} was determined randomly, and the partition of the probability weights was reset to their starting point. If during the process the distance between two holograms of G decreased below d pixels, then the best graded hologram was retained while the rest were modified by a random perturbation imposed on p percent of their pixels.

In our experiments group G contained two holograms. The different parameters were tried many times with different values, and the best results were obtained with the selections of $K = 40$, $d = 5$, and $p = 2$. The algorithm was stopped after no significant improvement was observed for a long time. On the average it was stopped after a total of 10,000 measurements, which, for a matrix of 128×128, is less than a single-scan DBS process. Accordingly and compared with other results,[5–7] this algorithm is faster than the DBS by at least 1 order of magnitude.

7. EXPERIMENTAL RESULTS

All the experiments reported here were performed with the system of Fig. 1 or 2 to generate a FT hologram. The linear system was a thin lens of focal length f located between two identical free spaces of length f. Thus the linear system constituted half the correlation channel, and the spatial sharing of two channels[7] was replaced by time sharing, as shown in Fig. 2. This configuration is also

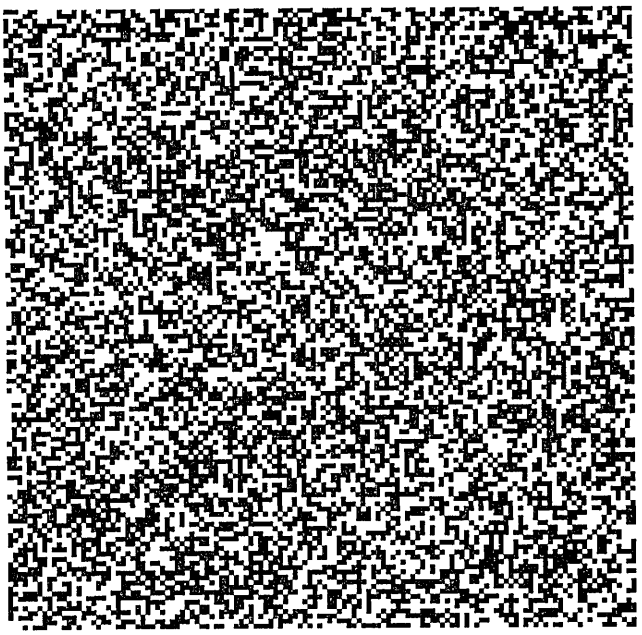

Fig. 4. Final BCGH of the reconstructed intensity distribution image shown in Fig. 5.

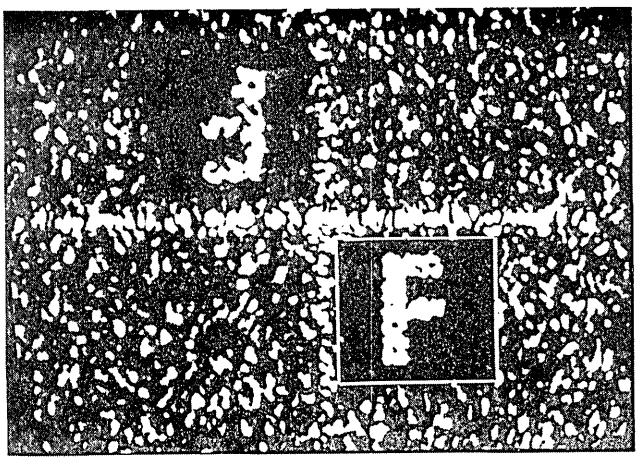

Fig. 5. Reconstructed image of the letter F, containing magnitude information only.

more economic in components. The SLM used here was a magneto-optical SLM with 128 × 128 pixels controlled by a personal computer enhanced with a CUE-2[14] image-processing package.

The first experiment used a single channel with only image intensity measurements starting with a random binary distribution. The iterative process was stopped after 10,000 measurements for the letter F and 20,000 measurements for the airplane after no significant improvements were observed for e_1, defined in Eq. (6). The final hologram is shown in Fig. 4, with its diffraction pattern over plane P_2 shown in Fig. 5. The rectangle drawn around the letter F indicates the area $\mathcal{A}$. Note that this pattern is not perfectly symmetric as would be expected from the FT of a binary function. This may be caused by slight phase distortions over the SLM and a displacement of the reconstructed plane from the exact focal plane of the FT lens. The end result, however, is not affected, since the hologram corrects any distortion or calibration mistakes because the process reduces the error in the actual subarea $\mathcal{A}$. The convergence of the process in a sequence of reconstructed images within the region $\mathcal{A}$ is shown in Fig. 6 for two different examples. The pictures were taken every 500 iterations for the letter F and every 1000 iterations for the airplane. It appears that the main

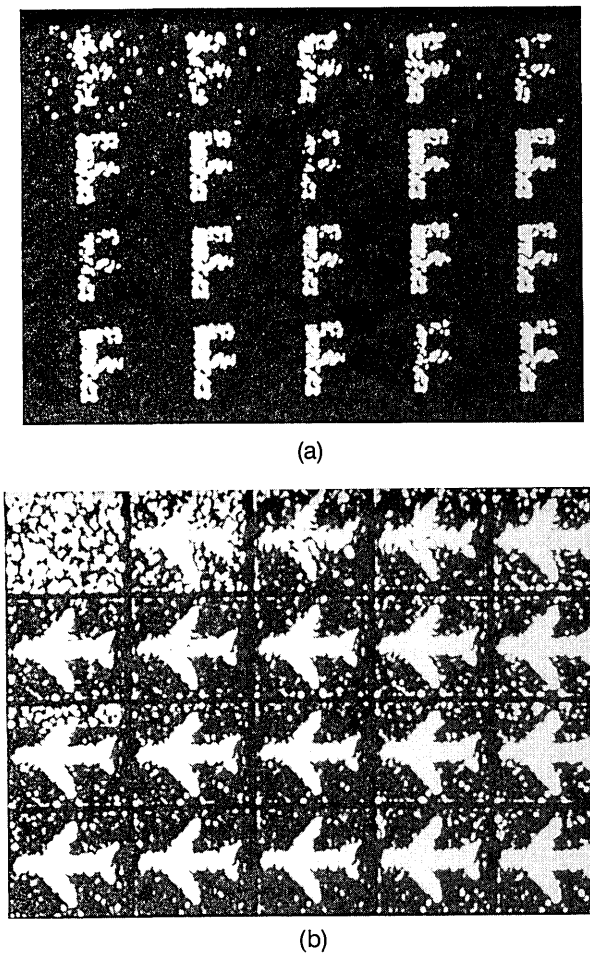

Fig. 6. Two examples of sequences of reconstructed images with random phase distribution. The numbers of iterations between every picture are (a) 500 and (b) 1000.

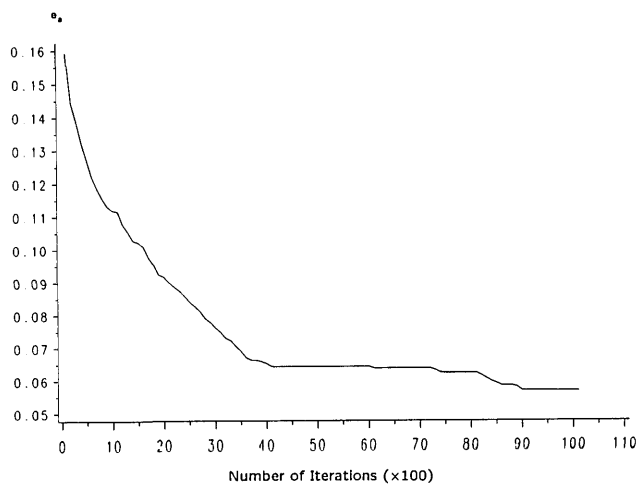

Fig. 7. Error, defined in Eq. (6), versus the number of iterations. The reconstructed images are shown in Fig. 6(a).

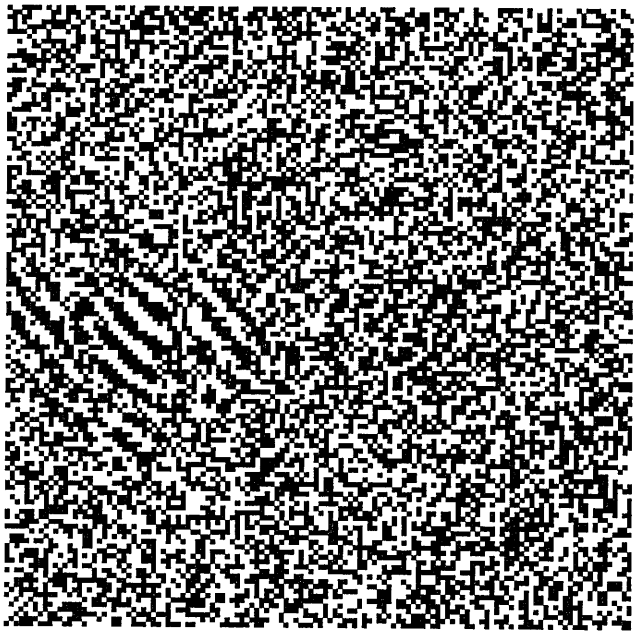

Fig. 8. Final BCGH of the reconstructed complex amplitude distribution shown in Fig. 9.

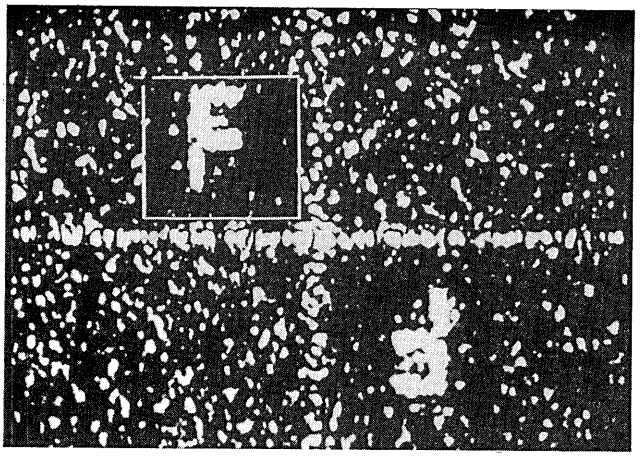

Fig. 9. Reconstructed image of the letter F containing all the complex information.

progress in the convergence was over in the first 4000 iterations. This conclusion is confirmed by the curve of the cost function versus the number of iterations, shown in Fig. 7, for the example of the letter F.

In the second experiment the complete system of two channels (Fig. 2) was utilized for generating a FT hologram of the letter F with constant phase distribution. The results are shown in Figs. 8–11. The final reconstructed image is shown in Fig. 9, while the correlation results are shown in Fig. 10. In the upper left-hand region it is the correlation distribution between $\tilde{f}(x, y)$ and $f(x, y)$, while in the lower right-hand region it is the convolution distribution between them. The convolution is in fact the correlation between the input object and the twin reconstructed image, shown at the lower right-hand side of Fig. 9.

When the reconstructed images of Figs. 9 and 5 are compared, we notice that the overall quality of the image is reduced for the two-channel experiment. This is reasonable since it is more difficult to satisfy two constraints than just one. Nevertheless, the strongly speckled character of the pattern (the image of F) produced in the single channel (Fig. 5) is significantly reduced in the reconstructed image of the double-channel case (Fig. 9). This observation indicates that the phase distribution in the double-channel process is much closer to the required uniform distribution than that obtained by a single channel.

The process was also repeated with only the correlation channel. The cost function defined in Eq. (14) was used in this experiment, and the results are shown in Figs. 12 and 13. Only the peak of the inner product, the correlation shown in the upper part of Fig. 12, was measured and maximized during the process. It is quite difficult to recognize the original letter F in the image shown in Fig. 13. This reflects the fact that for a good cross correlation the phase distribution of $F(u,v)$ is more important than the amplitude distribution.

8. DISCUSSION

The TIM method is a kind of compromise between conventional photographic optical holography and classical

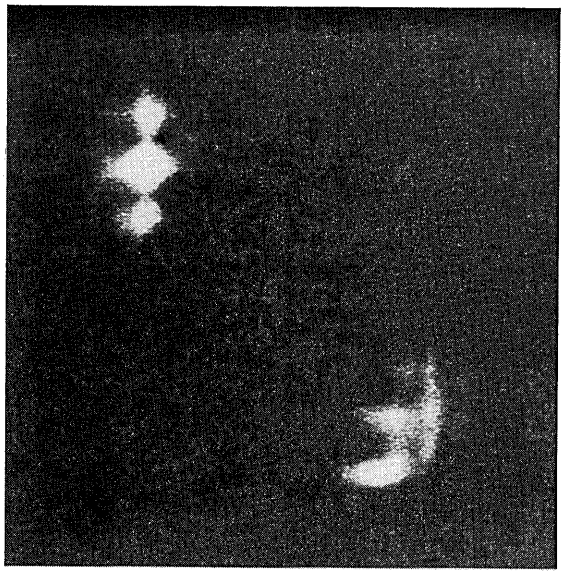

Fig. 10. Output distribution of the correlation channel.

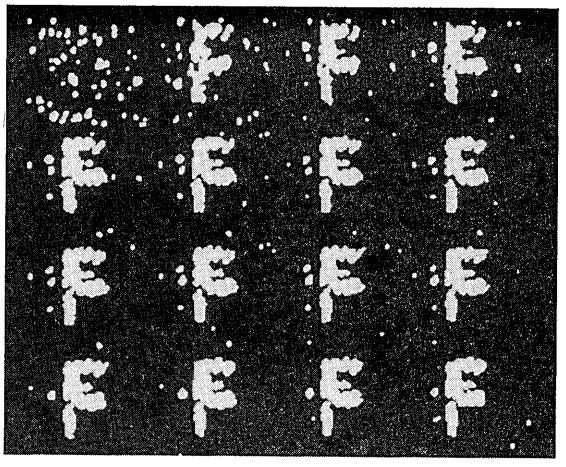

Fig. 11. Sequence of reconstructed images from two-channel process. The number of iterations between the pictures is 2000.

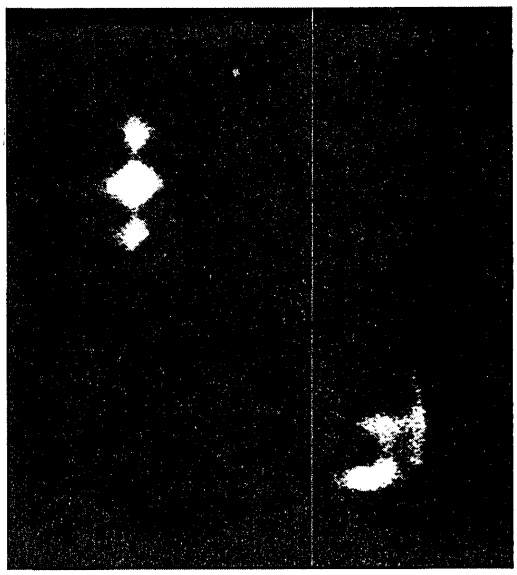

Fig. 12. Final output distribution of the correlation channel when it was measured alone, after 25,000 iterations.

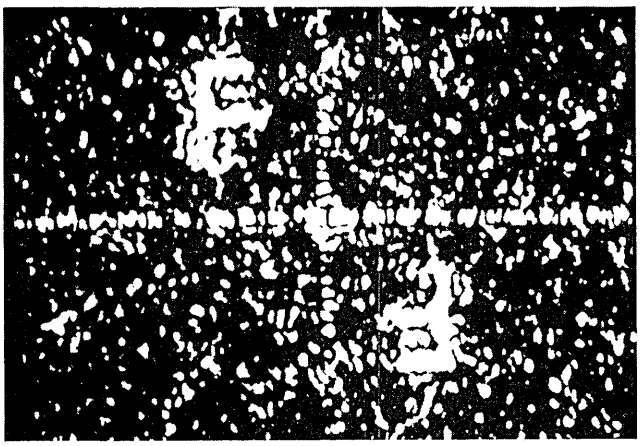

Fig. 13. Reconstructed object from the hologram that was created by maximization of the inner product value alone.

computer-generated holograms. The flexibility of digital computers enables one to record a hologram that represents any complex function that does not necessarily correspond to a real object. Optical photography, on the other hand, relates only to physical objects and is fast and direct without mathematical approximations. The TIM method tries to combine these features. The reconstructed image from the BCGH, generated by the TIM, is more accurate than the ordinary BCGH, since all the parameters of the optical system are accounted for during the fabrication. Part of the flexibility is lost since the process is connected to an actual input object. However, the generalization introduced in Section 4 and the single-correlation-channel operation extend the possibilities beyond what an input object permits in optical photography. Moreover, any object stored in the computer memory may serve as the reference object to be reconstructed. A typical and useful example of such a hypothetical object is an interconnection scheme when the hologram is used for a reconfigurable interconnection array.

The simple idea of measuring the phase distribution of an object by noninterferometric methods may have significant meaning in the field of holography beyond the BCGH issue. This idea may lead, with improvements in SLM technology, to a hologram recording process without the need for coherent interfering waves. Our initial experiments indicate that with a SLM capable of phase and amplitude modulation and a fast computer we may record a hologram on the SLM without carrier wave coding. On the one hand, the intensity of an object is recorded directly, and on the other hand, the phase transparency of the SLM is changed until the inner product between the object and the SLM transparency obtains its maximum value. The main advantage of such an arrangement is that the bandwidths of the reconstructed distributions from the hologram and the photographed object are equal. The reconstructed image is carried on the zero-order diffraction, and there are no higher diffraction orders. Although we do not argue that such a scheme can replace classic holography, it may be of significant value in various applications.

REFERENCES AND NOTES

1. A. W. Lohmann and D. P. Paris, "Binary Fraunhofer holograms, generated by computer," Appl. Opt. **6,** 1739–1748 (1967).
2. J. Shamir, H. J. Caulfield, and R. B. Johnson, "Massive holographic interconnections and their limitations," Appl. Opt. **28,** 311–324 (1989).
3. W. H. Lee, "Computer-generated holograms: techniques and applications," Prog. Opt. **16,** 119–232 (1978).
4. O. Bryngdahl and F. Wyrowski, "Digital holography—computer-generated holograms," Prog. Opt. **28,** 1–86 (1990).
5. M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, "Synthesis of digital holograms by direct binary search," Appl. Opt. **26,** 2788–2798 (1987).
6. B. K. Jennison, J. P. Allebach, and D. W. Sweeney, "Iterative approaches to computer-generated holography," Opt. Eng. **28,** 629–637 (1989).
7. U. Mahlab, J. Rosen, and J. Shamir, "Iterative generation of holograms on spatial light modulators," Opt. Lett. **15,** 556–558 (1990).
8. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. **5,** 1248–1249 (1966).
9. J. L. Horner and P. D. Gianino, "Phase-only matched filtering," Appl. Opt. **23,** 812–816 (1984).
10. B. V. K. Vijaya Kumar and L. Hassebrook, "Performance measures for correlation filters," Appl. Opt. **29,** 2997–3006 (1990).
11. M. B. Reid, P. W. Ma, J. D. Downie, and E. Ochoa, "Experimental verification of modified synthetic discriminant function filters for rotation invariance," Appl. Opt. **29,** 1209–1214 (1990).
12. U. Mahlab and J. Shamir, "Iterative optimization algorithms for filter generation in optical correlators: a comparison," Appl. Opt. **31,** 1117–1125 (1992).
13. J. A. Nelder and R. Mead, "A simpler method for function minimization," Comput. J. **8,** 308–313 (1965).
14. The CUE-2 is an image-processing system based on a personal computer manufactured by GALAI Laboratories, Industrial Zone, Migdal Haemek, Israel.