Optical Society of America. Ting-Chung Poon and Taegeun Kim, "Optical Image Recognition of Three-Dimensional Objects," Appl. Opt. 38, 370-381 (1999). doi: 10.1364/ao.38.000370

Optical image recognition of three-dimensional objects

Ting-Chung Poon and Taegeun Kim

A three-dimensional (3-D) optical image-recognition technique is proposed and studied. The proposed technique is based on two-pupil optical heterodyne scanning and is capable of performing 3-D image recognition. A hologram of the 3-D reference object is first created and then is used to modulate spatially one of the pupils of the optical system; the other pupil is a point source. A 3-D target object to be recognized is then scanned in two dimensions by optical beams modulated by the two pupils. The result of the two-dimensional scan pattern effectively displays the correlation of the holographic information of the 3-D reference object and that of the 3-D target object. A strong correlation peak results if the two pieces of the holographic information are matched. We analyze the proposed technique and thereby lay a theoretical foundation for optical implementations of the idea. Finally, computer simulations are performed to verify the proposed idea. © 1999 Optical Society of America

OCIS codes: 100.6890, 100.5010, 100.4550.

1. Introduction

The main objective of this paper is to investigate optical techniques for optical image recognition of three-dimensional (3-D) objects. Three-dimensional optical image recognition finds applications in the areas of 3-D microscopy, medical imaging and recognition, robotic vision, 3-D data acquisition and processing, and optical remote sensing. A key operation to the achievement of 3-D optical image recognition is 3-D optical correlation, i.e., an optical technique capable of performing 3-D correlation of a 3-D reference object and a 3-D target object to be recognized. This is a formidable task.

Research in two-dimensional (2-D) pattern recognition has its root in the 1960's,[1,2] and it has been rejuvenated in past decades because of the advancement of the development of spatial light modulators.[3,4] Optical image recognition of 3-D objects, on the other hand, is rarely tackled, mainly because of the lack of optical systems capable of performing 3-D correlation. Some studies involving transdimensional mapping have been done in optics, but they dealt exclusively with one-dimensional (1-D) to 2-D or 2-D to 1-D transformation.[5–7] A relevant 3-D to 2-D mapping has been proposed that involves sampling along the 3-D object's depth, hence representing the object by 2-D sectional depth images.[8] Recently, a 3-D joint transform correlator was demonstrated in that 2-D images can be recognized even when the images are translated along the z direction. The technique involves capturing the 3-D information by transverse displacement (i.e., x and y scanning) of a CCD camera.[9] However, these two proposed techniques suffer a major drawback of dimensional reduction in that the number of these many 2-D sampling images should be high for good resolution performance. For alleviating the need for many 2-D sampling images, range images have been used, but coding schemes must be applied to the input scene and the reference image, which makes the proposed system lack real-time capability.[10] Most recently, a planar encoding theory of 3-D images was proposed that does not require the many 2-D images.[11]

In this paper we propose a real-time optical technique for the recognition of 3-D objects. Our goal is to recognize a 3-D object, i.e., an object of three spatial coordinates. In essence, our proposed optical system performs the correlation of two pieces of holographic information pertaining to the two 3-D objects—a 3-D reference object and a 3-D target object—to be matched, hence 3-D recognition is possible. In the technique there is no need to record 2-D images for 3-D object representation, as a holographic technique is employed. The proposed 3-D recognition system is based on a two-pupil optical heterodyne scanning technique,[12,13] which is briefly

The authors are with the Optical Image Processing Laboratory, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. T.-C. Poon's e-mail address is tcpoon@vt.edu.

Received 20 May 1998; revised manuscript received 20 October 1998.

0003-6935/99/020370-12$15.00/0

discussed and summarized in Section 2. Section 3 develops 3-D optical image-recognition theory based on two-pupil optical heterodyne scanning, whereby the technique for correlating two sets of holographic information is developed. Section 4 discusses the pupils necessary to implement the idea for 3-D optical image recognition. Section 5 presents and discusses computer-simulation results. Section 6 addresses the important 3-D invariance of the proposed idea, and, finally, in Section 7 we present some concluding remarks.

2. Two-Pupil Optical Heterodyne Scanning: Introduction and Background

In incoherent image processing, intensity quantities are manipulated, hence the most severe limitation of incoherent image processing has been its inability to synthesize bipolar point-spread functions (PSF's), which precludes important image operations such as high-pass spatial filtering and the derivative. The most important development of incoherent image processing has therefore been the introduction of dual-channel processing, commonly known as two-pupil processing.[12–15] The methods of two-pupil synthesis utilize two parallel channels and realize the negative part of the PSF either by a direct difference of illumination or by an interferometric interaction of two pupils. The latter method involves encoding and decoding steps, and some possible schemes based on temporal frequency modulation[12,16,17] and spatial encoding have been proposed.[18–20]

Among the temporal frequency-modulation techniques, the use of active optical heterodyne scanning has been studied extensively by Poon and Korpel,[12] Poon,[13] and Indebetouw and Poon.[15] In the optical heterodyne scanning system two laser beams of different temporal frequency that carry the information of two different pupils are combined and used to scan an object in two dimensions. The heterodyne current from a photodetector is obtained serially as an output of the optical system. When the heterodyne current is demodulated and displayed on a 2-D display device, we have the processed image. The main result of the technique is that the synthesized optical transfer function (OTF) is the cross correlation of the two pupils, hence complex PSF's can be generated in incoherent image processing. An extension of this research includes an important solution in incoherent holography that has the drawback of a large background buildup. Indeed, Poon proposed the use of a two-pupil optical heterodyne scanning system to achieve real-time incoherent holographic recording without the drawback of background buildup. In the proposed system, one pupil is a spherical wave, and the other is a point source of different temporal frequency, whereby the synthesized PSF becomes a Fresnel zone plate (FZP) with rings running in and out as time advances.[13] Effectively, the system, which utilizes 2-D scanning, performs correlation of the input object with a FZP. Each point of the object is encoded as a FZP, hence the recording has become a hologram of the object. This real-time holographic recording technique is now known as optical scanning holography.[21]

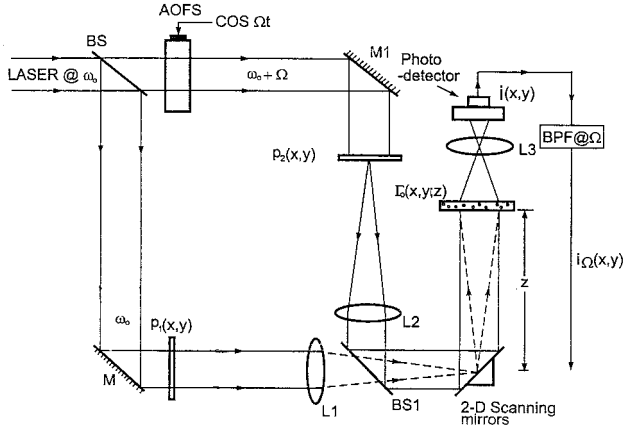

Fig. 1. Two-pupil optical heterodyne scanning system. BPF@Ω, bandpass filter tuned at the frequency Ω; AOFS, acousto-optic frequency shifter.

Holography has been used for 2-D pattern recognition. A hologram of the reference object is inserted into the Fourier plane of a 4*f* system for 2-D image recognition.[1] In principle, 3-D recognition is possible when the 3-D object to be recognized is presented as sectional 2-D images of sequential inputs to the 4*f* system, but again this technique has the drawback of dimensional reduction, as discussed in Section 1. In this paper we propose the use of optical scanning holography for real-time 3-D object recognition. The proposed idea of 3-D object recognition is to perform a 2-D correlation of two pieces of holographic data, one piece corresponding to the 3-D reference and the other to the 3-D target object to be recognized. This scheme is possible for 3-D recognition because holographic data contain the 3-D information of the reference and the target. The scheme of the idea is as follows: (1) A hologram of the 3-D reference is first generated by the two-pupil heterodyne scanning system in which one of the pupils is a plane wave and the other a point source. (2) The heterodyne scanning system is used again, but one of the pupils carries the holographic information of the 3-D reference object, and the other remains the same as a point source. A 3-D target to be recognized is then scanned in two dimensions. If the two 3-D objects are the same, we have a strong correlation peak at the output of a 2-D display.

We now summarize and briefly describe the two-pupil heterodyne scanning system. A detailed mathematical description of the system is presented in Appendix A. A typical two-pupil heterodyne optical scanning system is shown in Fig. 1. Beam splitters BS and BS_1 and mirrors M and M_1 form a Mach–Zehnder interferometer. A collimated laser at a temporal frequency ω_0 is used to illuminate the pupil $p_1(x, y)$. The other pupil, $p_2(x, y)$, is illuminated by a laser of temporal frequency $\omega_0 + \Omega$. The laser's temporal frequency offset of Ω is introduced by an acousto-optic frequency shifter, as shown in Fig.

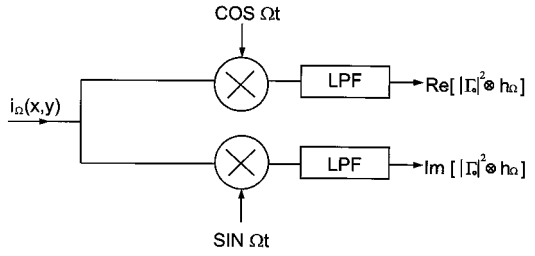

Fig. 2. Parallel processing for obtaining the in-phase and the quadrature-phase information of the scanned signal simultaneously. LPF, low-pass filter.

1.[22] The two pupils are located at the front focal planes of lenses L_1 and L_2, both with a focal length f. The two pupils are then combined by beam splitter BS_1 to focus the light onto the 2-D x–y scanning mirrors, which are located on the back focal plane of lenses L_1 and L_2. The combined optical scanning field is used to two-dimensionally raster scan over an object of amplitude distribution $\Gamma_o(x, y; z)$, located a distance z away from the focal plane of the two lenses. Lens L_3 is used to collect all the transmitted light (or scattered light if the object is diffusely reflecting) onto the photodetector. An electronic bandpass filter tuned at the heterodyne frequency Ω gives a scanned and processed current i_Ω as output. When i_Ω is mixed electronically with a sine and a cosine wave, as shown in Fig. 2, we have two outputs given by

$$\begin{aligned} i_c(x, y; z) &= \mathrm{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o|^2\} \times \mathrm{OTF}_\Omega\}] \\ &= \mathrm{Re}[|\Gamma_o(x, y; z)|^2 \otimes h_\Omega(x, y; z)], \quad (1a) \\ i_s(x, y; z) &= \mathrm{Im}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o|^2\} \times \mathrm{OTF}_\Omega\}] \\ &= \mathrm{Im}[|\Gamma_o(x, y; z)|^2 \otimes h_\Omega(x, y; z)], \quad (1b) \end{aligned}$$

where the symbol $\otimes$ denotes a convolution operation [see Eq. (A5)] and $h_\Omega(x, y; z) = \mathcal{F}^{-1}\{\mathrm{OTF}_\Omega\}$ is the PSF of the optical heterodyne scanning system with $\mathcal{F}$ and $\mathcal{F}^{-1}$ denoting the Fourier transform operation and its inverse, respectively, as defined in Appendix A. Finally, OTF_Ω is the OTF of the system, given by

$$\begin{aligned} &\mathrm{OTF}_\Omega(k_x, k_y; z) \\ &= \exp\left[j \frac{z}{2k_0}(k_x^2 + k_y^2)\right] \\ &\times \iint p_1^*(x', y') p_2\left(x' + \frac{f}{k_0} k_x, y' + \frac{f}{k_0} k_y\right) \\ &\times \exp\left[j \frac{z}{f}(x'k_x + y'k_y)\right] \mathrm{d}x'\mathrm{d}y', \quad (2) \end{aligned}$$

where k_0 is the wave number of the light. k_x and k_y are the spatial frequencies associated with the coordinates x and y, respectively. $p_1(x, y)$ and $p_2(x, y)$ are the two pupils located in the front focal planes of lenses L_1 and L_2, respectively, with focal length f. Note that the two pupils completely define the spatial-filtering characteristic of the scanning system and that only the intensity distribution $|\Gamma_o|^2$ is processed; hence the optical system is incoherent. A detailed derivation of Eqs. (1) is shown in Appendix A.

In Eqs. (1), the input object $|\Gamma_o(x, y; z)|^2$ has been assumed to be an infinitely thin 2-D object located a distance z away from the focus of the spherical wave, which is on the 2-D scanning mirrors as shown in Fig. 1. To generalize Eqs. (1) for 3-D objects, we need to integrate them over the depth, i.e., over z, of the 3-D object. Equations (1) become

$$\begin{aligned} i_c(x, y) &= \mathrm{Re}\left[\int \mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\} \times \mathrm{OTF}_\Omega\}\mathrm{d}z\right] \\ &= \mathrm{Re}\left[\int |\Gamma_o(x, y; z)|^2 \otimes h_\Omega(x, y; z)\mathrm{d}z\right], \quad (3a) \\ i_s(x, y) &= \mathrm{Im}\left[\int \mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\} \times \mathrm{OTF}_\Omega\}\mathrm{d}z\right] \\ &= \mathrm{Im}\left[\int |\Gamma_o(x, y; z)|^2 \otimes h_\Omega(x, y; z)\mathrm{d}z\right]. \quad (3b) \end{aligned}$$

Note that we have left the z dependence out of the left-hand side of Eqs. (3) to emphasize that the recorded information is strictly two dimensional, even for 3-D objects. $i_c(x, y)$ or $i_s(x, y)$ represents the scanned and the processed current or the information of $|\Gamma_o|^2$ and can be displayed as a 2-D record if these currents are stored in synchronization with the signals used to drive the x–y scanning mirrors.

3. Recognition of a Three-Dimensional Object

The proposed 3-D recognition technique is divided into two steps: (a) construction of a reference hologram and (b) correlation of two sets of holographic information by 2-D scanning of the 3-D target object. The technique is based on optical heterodyne scanning, as described in Section 2. We now describe the two steps separately.

A. Construction of Reference Holograms

To construct the hologram of a 3-D reference object, we choose the two pupils such that $p_1(x, y) = 1$ and $p_2(x, y) = \delta(x, y)$, as illustrated in Fig. 1. With this choice of pupils the OTF of the scanning system, according to Eq. (2), becomes

$$\begin{aligned} \mathrm{OTF}_\Omega(k_x, k_y; z) &= \exp\left[-j \frac{z}{2k_0}(k_x^2 + k_y^2)\right] \\ &= \mathrm{OTF}_{\mathrm{osh}}(k_x, k_y; z), \quad (4) \end{aligned}$$

where the subscript osh denotes that the particular OTF achieved is for optical scanning holography.[21] What is now being recorded in two dimensions is a hologram $H(x, y)$. We use Eq. (3a) as an example.

The $i_c(x, y)$ now becomes a hologram and is given by

$$H(x, y) = \mathrm{Re}\left[\int \mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\}\mathrm{OTF}_{\mathrm{osh}}(k_x, k_y; z)\}\mathrm{d}z\right]$$
$$= \mathrm{Re}\left[\int \mathcal{F}^{-1}\left\{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]\right\}\mathrm{d}z\right] \quad (5)$$

when using Eq. (4). From Eq. (5) it is clear that the scanning holographic recording process in the frequency domain can be interpreted as the object's spectrum along its depth (z) being processed by the OTF of the form given by Eq. (4). To see clearly why this interpretation corresponds to holographic recording, we rewrite Eq. (5) in terms of convolution; we then have

$$H(x, y) = \mathrm{Re}\left[\int |\Gamma_o|^2 \otimes h^*(x, y; z)\mathrm{d}z\right], \quad (6)$$

where, interestingly, Eq. (5) can be written in terms of $h(x, y; z)$, which is the free-space impulse response[23] [also see Eq. (A4) in Appendix A]. Now Eq. (6) can be rewritten in terms of the correlation operation [see Eq. (A9)]:

$$H_{\sin}(x, y) = \mathrm{Re}\left[\int h(x, y; z) \odot |\Gamma_o|^2\mathrm{d}z\right]$$
$$= \int \frac{k_0}{2\pi z}\sin\left[\frac{k_0}{2z}(x^2 + y^2)\right] \odot |\Gamma_o|^2\mathrm{d}z, \quad (7)$$

where the symbol $\odot$ denotes correlation. In writing the last step of Eq. (7), because $|\Gamma_o|^2$ represents the intensity distribution, which is strictly positive, the Re operation has been distributed to the function h. The subscript sin on the left-hand side of Eq. (7) denotes that the object has been correlated with a sine-type zone pattern.

If we now let $|\Gamma_o|^2 = \delta(x - x_0, y - y_0, z - z_0)$, we have $H_{\sin}(x, y) = (k_0/2\pi z)\sin\{(k_0/2z_0)[(x - x_0)^2 + (y - y_0)^2]\}$, which is the hologram of an offset delta function located at a distance $z = z_0$ away from the scanning mirror. Equation (7) is a very important result. In optical scanning holography the 3-D holographic recording process can be thought of as a 2-D transverse correlation between the real part of the free-space impulse response, $k_0/2\pi z \sin[(k_0/2z)(x^2 + y^2)]$, and the 3-D object, $|\Gamma_o(x, y; z)|^2$. The resulting correlation is then integrated along the depth of the object to obtained the hologram. To put Eq. (7) into a wider context, the Re operation can be replaced by the Im operation. This corresponds to the use of Eq. (3b) as output:

$$H_{\cos}(x, y) = \mathrm{Im}\left[\int h(x, y; z) \odot |\Gamma_o|^2\mathrm{d}z\right]$$
$$= \int \frac{k_0}{2\pi z}\cos\left[\frac{k_0}{2z}(x^2 + y^2)\right] \odot |\Gamma_o|^2\mathrm{d}z$$
$$= \frac{k_0}{2\pi z_0}\cos\left\{\frac{k_0}{2z_0}[(x - x_0)^2 + (y - y_0)^2]\right\}, \quad (8)$$

for $|\Gamma_o|^2 = \delta(x - x_0, y - y_0, z - z_0)$. We call $H_{\sin}(x, y)$ and $H_{\cos}(x, y)$ the sine FZP hologram and the cosine FZP hologram, respectively, of the object $|\Gamma_o|^2$. Indeed, because of the fact that holographic information is available in electronic form, as given by $i_\Omega(x, y)$ shown in Fig. 1, we can have a sine FZP hologram and a cosine FZP hologram simultaneously through parallel processing, as shown in Fig. 2. It turns out that these holograms are important for the implementation of complex pupil functions for 3-D optical image recognition, which is discussed in Section 4.

B. Correlation of Two Sets of Holographic Information

From Subsection 3.A we can see that holographic recording can be analyzed by using an OTF approach and that the OTF_Ω is expressed in terms of the two-pupil functions, as shown in Eq. (2), in general. By particularly choosing values of $p_1 = 1$ and $p_2 = \delta(x, y)$, we have the type of OTF_Ω [see Eq. (4)] that is capable of performing holographic recording.

To be more general, we now choose a value of $p_2 = \delta(x, y)$ and keep $p_1(x, y)$ as is. The OTF_Ω of the system then becomes, according to Eq. (2),

$$\mathrm{OTF}_\Omega(k_x, k_y; z) = \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right] \times p_1^*\left(-\frac{f}{k_0}k_x, -\frac{f}{k_0}k_y\right). \quad (9)$$

The 2-D record is then, according to Eq. (3a), i.e., by use of the Re operation as an example first,

$$i_c(x, y) = \mathrm{Re}\left[\int \mathcal{F}^{-1}\left\{\mathcal{F}\{O(x, y; z)\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right] \times p_1^*\left(-\frac{f}{k_0}k_x, -\frac{f}{k_0}k_y\right)\right\}\mathrm{d}z\right], \quad (10)$$

where we assume that $|\Gamma_o|^2 = O(x, y; z)$ is the 3-D target object to be recognized. Now we see that $\int\mathcal{F}\{O(x, y; z)\}\exp[-j(z/2k_0)(k_x^2 + k_y^2)]\mathrm{d}z$ contains a holographic recording of $O(x, y; z)$ if we conveniently interchange the integral and $\mathcal{F}^{-1}$ in Eq. (10) [also see Eq. (5)]. If we now assume that p_1 carries the holo-

graphic information of a 3-D reference object $R(x, y; z)$, we have

$$p_1\left(-\frac{f}{k_0}k_x, -\frac{f}{k_0}k_y\right) = \int \mathcal{F}\{R(x, y; z)\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]dz. \quad (11)$$

Substituting Eq. (11) into Eq. (10), we have a 2-D recording of the form

$$i_c(x, y) = \mathrm{Re}\left[\mathcal{F}^{-1}\left\{\left(\int \mathcal{F}\{O(x, y; z)\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]dz\right) \times \left(\int \mathcal{F}\{R(x, y; z)\}\exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]dz\right)^*\right\}\right]. \quad (12)$$

By use of the correlation property, i.e., $\mathcal{F}^{-1}\{F^*G\} = \mathcal{F}^{-1}\{F\} \odot \mathcal{F}^{-1}\{G\} = f \odot g$, Eq. (12) becomes

$$\begin{aligned} i_c(x, y) &= \mathrm{Re}\left[\left(\int \mathcal{F}^{-1}\left\{\mathcal{F}\{R(x, y; z)\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]\right\}dz\right) \odot \left(\int \mathcal{F}^{-1}\left\{\mathcal{F}\{O(x, y; z)\} \times \exp\left[-j\frac{z}{2k_0}(k_x^2 + k_y^2)\right]\right\}dz\right)\right] \\ &= \mathrm{Re}[H_R(x, y) \odot H_O(x, y)], \end{aligned} \quad (13a)$$

where $H_R(x, y) = \int h(x, y; z) \odot R(x, y; z)dz$ and $H_O(x, y) = \int h(x, y; z) \odot O(x, y; z)dz$. From inspection of Eq. (7), $H_R(x, y)$ and $H_O(x, y)$ represent the holographic information of $R(x, y; z)$ and $O(x, y; z)$, respectively, obtained by optical scanning holography. By the same token, when Eq. (3b) is used, we have

$$i_s(x, y) = \mathrm{Im}[H_R(x, y) \odot H_O(x, y)]. \quad (13b)$$

Equations (13) are the major result of the paper: $i_c(x, y)$ or $i_s(x, y)$ given by Eqs. (13) is the display of a 2-D pattern that basically represents the correlation of the holograms of two 3-D objects, R and O. Only when the two sets of holographic data are identical, i.e., $H_R(x, y) = H_O(x, y)$, do we have a strong peak on the pattern, hence a match of the two 3-D objects. The 3-D target object has to match with the 3-D reference object throughout the depth to produce a sharp correlation peak, hence a true 3-D matching scheme. We do not need to record 2-D images of the 3-D objects to be matched, but a single hologram of the reference 3-D object is required to perform 3-D matching.

4. Pupil Implementation for Three-Dimensional Optical Image Recognition

In Section 3 we discussed the object $O(x, y; z)$ to be recognized at a distance z away from the scanner mirror as being scanned with the two pupils set as $p_2 = \delta(x, y)$ and $p_1(x, y)$ and as having a functional form such as that given by Eq. (11), which contains the holographic information of R. If $O = R$, we have a 3-D match. Let us now discuss how p_1 can be implemented. We rewrite Eq. (11) as

$$p_1\left(-\frac{f}{k_0}k_x, -\frac{f}{k_0}k_y\right) = \int [\mathcal{F}\{h(x, y; z) \odot R(x, y; z)\}_{k_x, k_y}]dz. \quad (14)$$

Substituting $x = -(f/k_0)k_x$ and $y = -(f/k_0)k_x$ into Eq. (14), we have

$$\begin{aligned} p_1(x, y) &= \int [\mathcal{F}\{h(x, y; z) \odot R(x, y; z)\}_{k_x=-k_0x/f, k_y=-k_0y/f}]dz \\ &= \mathcal{F}\left\{\int h(x, y; z) \odot R(x, y; z)dz\right\}_{k_x=-k_0x/f, k_y=-k_0y/f}. \end{aligned} \quad (15)$$

Note that $\int h(x, y; z) \odot R(x, y; z)dz = H_R(x, y)$ is the complex hologram of the reference object R. This result can be achieved conveniently by optical scanning holography, as discussed in Subsection 3.A. To be more precise, Eqs. (7) and (8) can be added to give a complex hologram $H_R(x, y)$ of the reference object R as follows:

$$H_R(x, y) = H_{\cos}(x, y) + jH_{\sin}(x, y) \sim \int h(x, y; z) \odot R(x, y; z)dz, \quad (16)$$

for $|\Gamma_o|^2 = R(x, y; z)$, where $H_{\cos}(x, y)$ and $H_{\sin}(x, y)$ in this case are the cosine FZP and the sine FZP holograms, respectively, of the reference object R. In summary, for 3-D optical image recognition, the scanning system shown in Fig. 1 will have one pupil of $p_2 = \delta(x, y)$ with the other pupil p_1 specified by Eq. (15), a Fourier transform of the hologram of the 3-D reference object R. The target object $O(x, y; z)$ to be

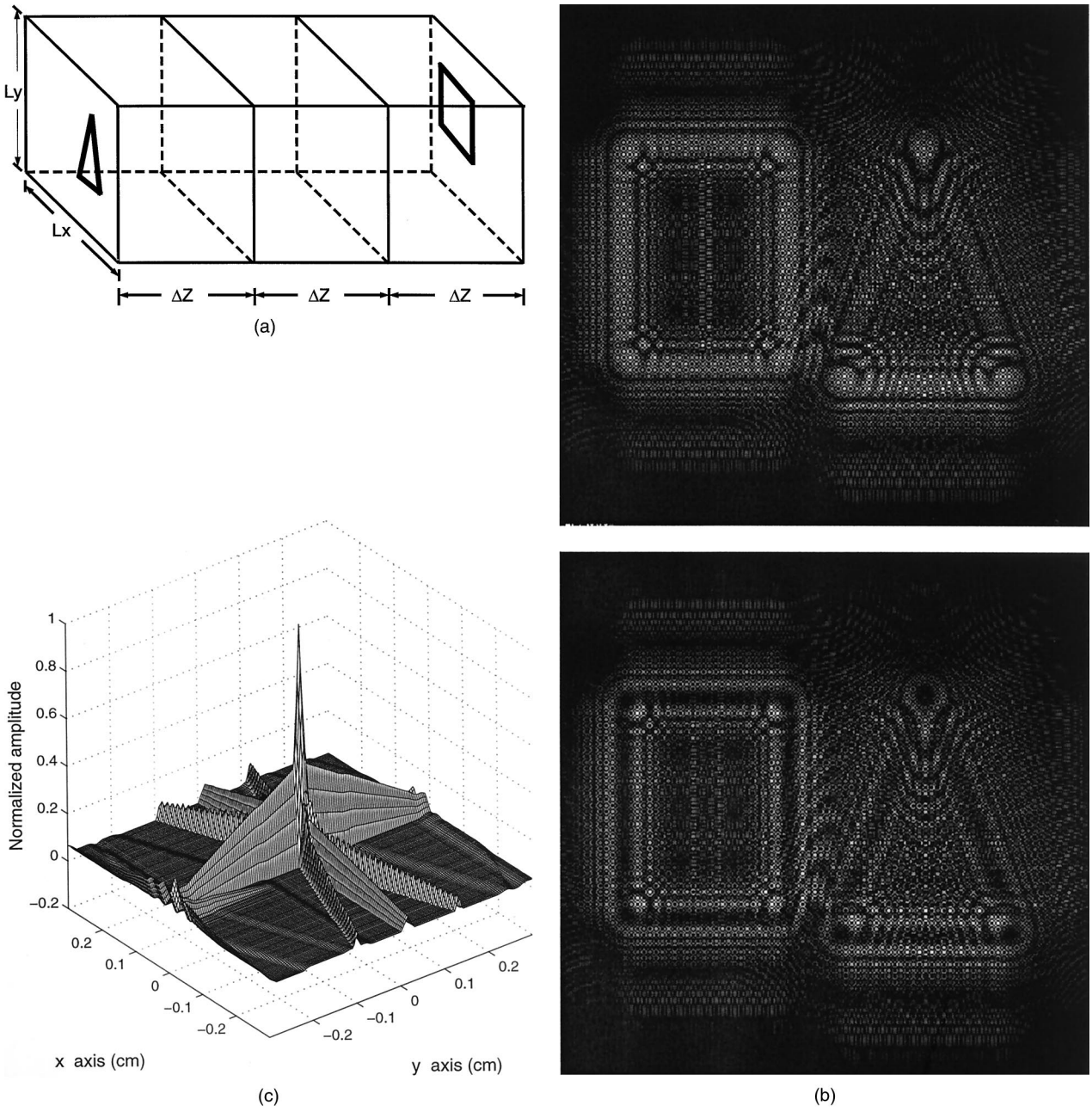

Fig. 3. (a) Three-dimensional reference object R, with $L_x = L_y = 1$ cm and $\Delta z = 1$ cm. (b) Holograms of the 3-D reference object of (a) (the scanned area is 1 cm × 1 cm): Shown are the cosine FZP hologram (upper image) and the sine FZP hologram (lower image). A bias has been added to the holograms to avoid plotting negative values. (c) Correlation output (0.4 cm × 0.4 cm) when the target object O is matched with the reference object R.

recognized will be two-dimensionally scanned, and a strong correlation peak will occur when $O = R$.

5. Simulation Results

Figure 3(a) shows a 1 cm × 1 cm × 3 cm 3-D reference object R, consisting of a triangle and a square that are separated by 3 cm along the depth of the object. Figure 3(b) plots the cosine FZP hologram (upper graph) and the sine FZP hologram (lower graph) of the reference object R, as calculated by Eqs. (8) and (7), respectively, with $z = 26$ cm for the triangle and $z = 29$ cm for the square of the 3-D reference object. The wavelength of light used is 0.6 μm for h. The scanned area of the hologram is 1 cm × 1 cm. Physically these holograms would correspond to the scanning of R, with its front face located 26 cm away from the scanning mirrors and the outputs given as shown in Fig. 2. Again, the pupils have be chosen such that $p_1(x, y) = 1$ and $p_2(x, y) = \delta(x, y)$ for holographic recording, as described in Subsection 3.A. The holograms shown in Fig. 3(b) are now used for the pupil specified by Eq. (15), which results in the output calculated according to Eq. (13a) [alternatively, one can

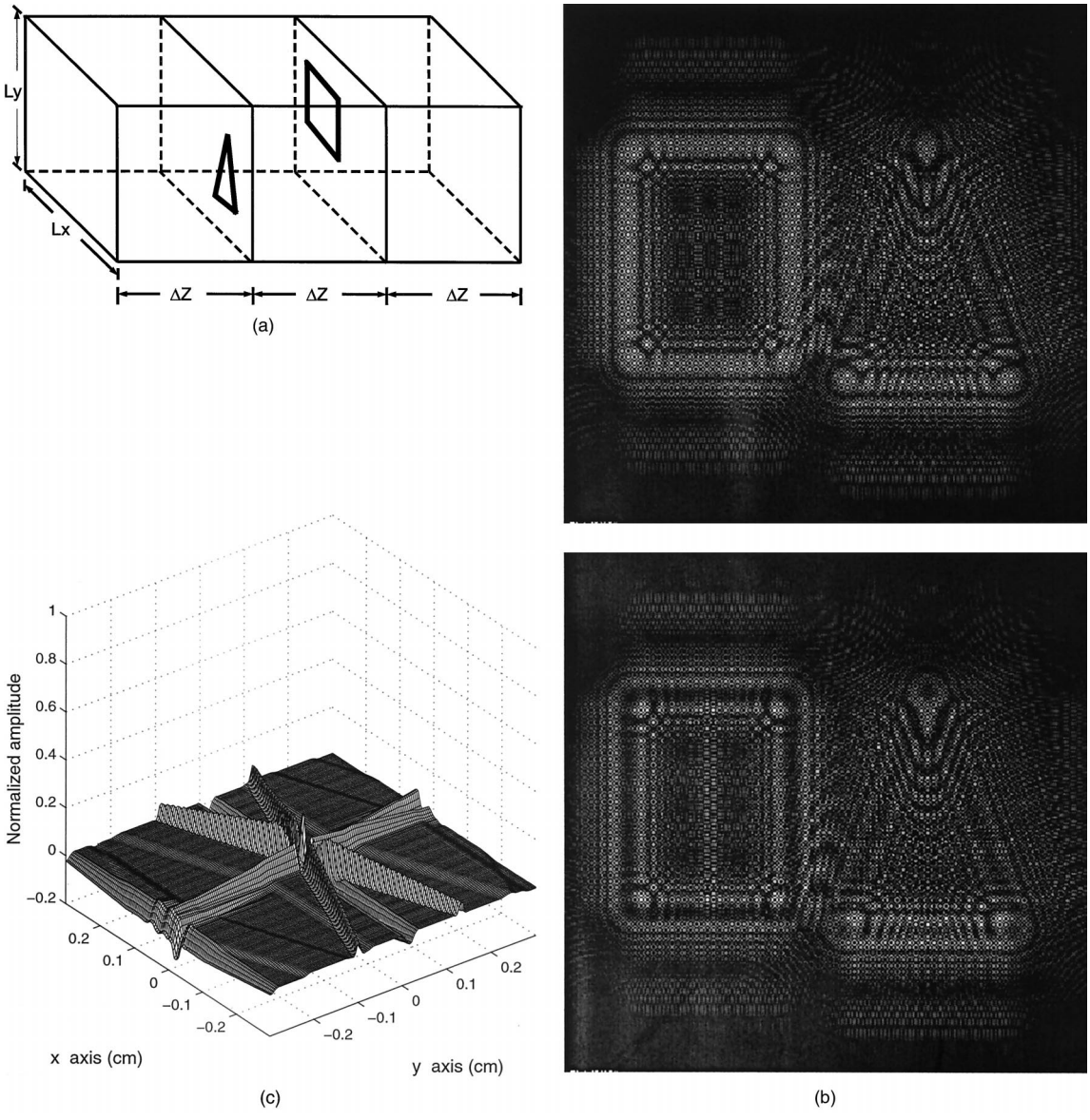

Fig. 4. (a) Three-dimensional target object O. (b) Holograms of the 3-D target object of (a): Shown are the cosine FZP (upper image) and the sine FZP (lower image) holograms. (c) Correlation output when the target object of (a) is scanned. The reference object is shown in Fig. 3(a).

use Eq. (13b) for the output (see Fig. 2 for parallel-processing discussion)]:

$$t(x,y) = \mathrm{Re}\left\{\left[\int h(x,y;z) \odot R(x,y;z)\mathrm{d}z\right] \odot \left[\int h(x,y;z) \odot O(x,y;z)\mathrm{d}z\right]\right\}. \quad (17)$$

The output represented by Eq. (17) corresponds to the 2-D scanning of a 3-D target object O with the hologram of a 3-D reference object R used to form the pupil. Figure 3(c) shows Eq. (17) plotted for $R = O$, where O is the object shown in Fig. 3(a).

With the same reference object shown in Fig. 3(a), Fig. 4 shows the results of another 3-D target object to be scanned. Figure 4(a) shows the target object, and its front face is again 26 cm away from the scanning mirrors in the simulations. Note that the target has the same 2-D patterns as those in the reference object but located at different depths. Figure 4(b) shows the cosine FZP hologram (upper image) and the sine FZP hologram (lower image) of the target object O. Figure 4(c) shows the correlation output according to Eq. (17). Note that the amplitude axis of Fig. 4(c) has been normalized by the peak value of Fig. 3(c). The lack of a strong correlation peak in Fig. 4(c) clearly indicates a mismatch of the two 3-D objects. It is important to point out that the holograms shown in Figs. 3(b) and 4(b) are similar, and yet, when they are correlated with the reference hologram, as our proposed system effectively performs, we do not find a match for the two 3-D objects shown in Figs. 3(a) and 4(a). The reason is that,

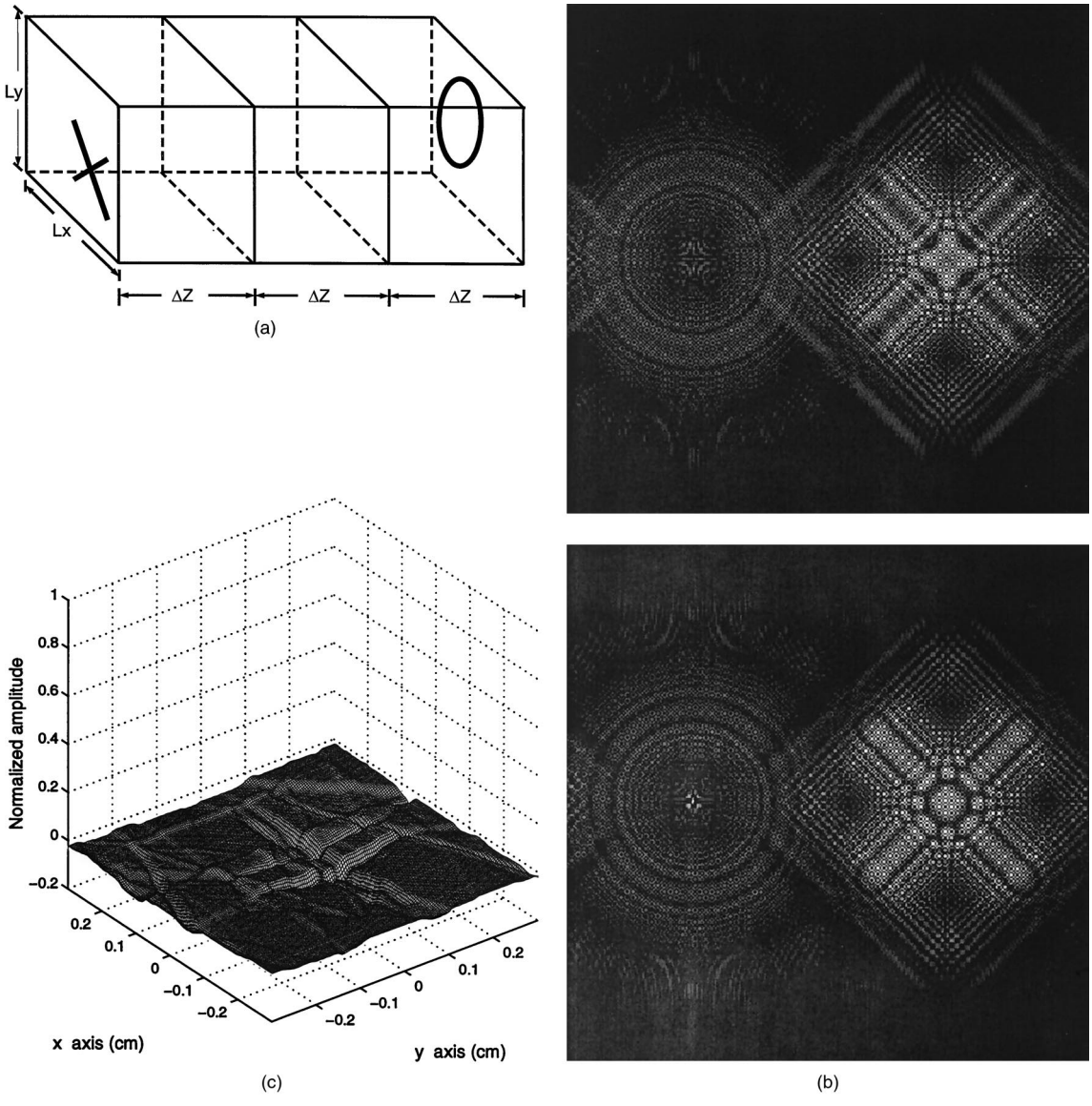

Fig. 5. (a) Three-dimensional target object O. (b) Holograms of the 3-D target object in (a): Shown are the cosine FZP (upper image) and the sine FZP (lower image) holograms. (c) Correlation output when the target object in (a) is scanned. The reference object is shown in Fig. 3(a).

because we are really correlating two holographic data sets, the depth information is as important as the planar distributions. To have a strong correlation peak, we must match the 3-D objects precisely throughout the whole 3-D volume. Finally, in Figs. 5(a) and 5(b) we show another target object and its holograms. The correlation outputs between Fig. 3(a), the reference, and Fig. 5(a), the target, are shown in Fig. 5(c). Clearly, no correlation peak is observed in Fig. 5(c). Note that in this case the X and the O of the 3-D object are located at the same distance as the triangle and the square, respectively.

6. Three-Dimensional Shift Invariance

Three-dimensional shift invariance in object recognition is of practical importance. The system proposed so far is not invariant to the target-object shift along the z axis. In this section we analyze z invariance and provide a solution. Let the target 3-D object be shifted along x, y, and z by amounts Δx, Δy, and Δz, respectively, i.e., we let $O(x, y; z) = O(x - \Delta x, y - \Delta y; z - \Delta z)$. Equation (13a) becomes

$$
\begin{aligned}
i_c(x, y) &= \mathrm{Re}\Bigg(H_R(x, y) \odot \Bigg\{\int h(x, y; z) \\
&\qquad \odot O(x - \Delta x, y - \Delta y; z - \Delta z)dz\Bigg\}\Bigg) \\
&= \mathrm{Re}\Bigg(H_R(x, y) \odot \Bigg\{\int h(x, y; z + \Delta z) \\
&\qquad \odot O(x - \Delta x, y - \Delta y; z)dz\Bigg\}\Bigg). \qquad (18)
\end{aligned}
$$

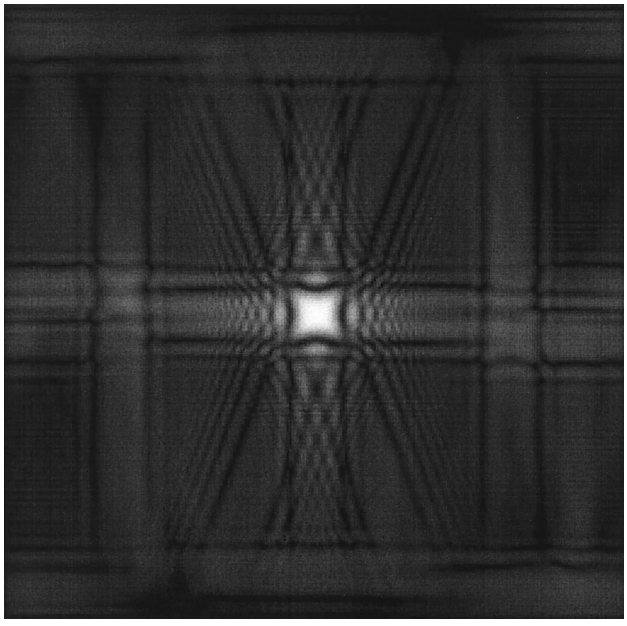

Fig. 6. Correlation output when the 3-D target object and the 3-D reference object are displaced along the depth (z) direction; otherwise the object and the reference are the same.

After some manipulation Eq. (18) can be cast into the following form involving convolution:

$$i_c(x,y) = \mathrm{Re}\Bigg(\Bigg\{H_R(x,y) \odot \Bigg[\int h(x,y;z) \odot O(x-\Delta x, y-\Delta y; z)dz\Bigg]\Bigg\} \otimes h^*(x,y;\Delta z)\Bigg)$$
$$= \mathrm{Re}\{[H_R(x,y) \odot H_O(x-\Delta x, y-\Delta y)] \otimes h^*(x,y;\Delta z)\}. \quad (19a)$$

Similarly, Eq. (13b) becomes

$$i_s(x,y) = \mathrm{Im}\{[H_R(x,y) \odot H_O(x-\Delta x, y-\Delta y)] \otimes h^*(x,y;\Delta z)\}. \quad (19b)$$

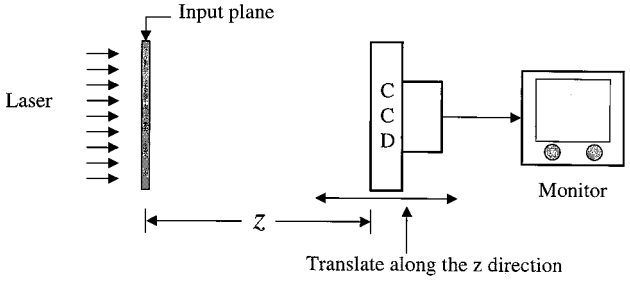

Fig. 7. Optical system capable of extracting the depth difference between the 3-D target object and the 3-D reference object. Otherwise the two objects are the same.

Now, if we add the correlation results from Eqs. (19) to form a complex correlation $C(x, y)$, we have

$$C(x,y) = i_c(x,y) + j i_s(x,y) = [H_R(x,y) \odot H_O(x-\Delta x, y-\Delta y)] \otimes h^*(x,y;\Delta z). \quad (20)$$

Equation (20) states that, if the 3-D target object has a value of $O(x, y; z) = R(x - \Delta x, y - \Delta y; z - \Delta z)$ (i.e., it is the same as the 3-D reference object but shifted along Δx, Δy, and Δz), the correlation peak is located at Δx and Δy, and the peak is broadened by the convolution of $h^*(x, y; \Delta z)$ because of the longitudinal shift of the target object. Figure (6) plots the absolute value of Eq. (20) and shows a broadened correlation output when the 3-D target object shown in Fig. 3(a) is now located 51 cm away from the scanning mirror and where the 3-D reference object is the same as the target object but is located 26 cm away. To extract the information of Δz optically, we propose the system shown in Fig. 7. $C(x, y)$, in the form of a complex transparency, would be placed on the input plane and illuminated uniformly by a laser. A z-translated CCD camera detects the diffraction pattern and produces 2-D output. When the distance from the input plane to the CCD camera is $z = \Delta z$, the output of the shifted CCD camera gives the 3-D coordinates of the shifted correlation peak. Mathematically, it is clear from Eq. (20) that $C(x, y) \otimes h(x, y; \Delta z)$, resulting from diffraction, gives $H_R(x, y) \odot H_O(x - \Delta x, y - \Delta y)$ as $h(x, y; \Delta z) \otimes h^*(x, y; \Delta z) \sim \delta(x, y)$. Figure 8 shows the intensity distribution of the correlation output as the CCD camera moves from $z = 10$ cm to $z = 35$ cm. Note that, when $z = 25$ cm (i.e., the difference between 51 cm and 26 cm), we have a sharp correlation peak, as expected.

7. Concluding Remarks

We have developed a method for 3-D image matching. Optical implementations of the idea have been discussed, and systems have been analyzed. Real-time implementation is possible through 2-D optical scanning and the use of spatial light modulators for hologram display for the 3-D reference object. We have also addressed the important and practical z invariance of the system and showed that the overall scheme of the system is 3-D shift invariant. Computer simulations have been performed for clarification and confirmation of the idea. Inasmuch as our proposed system effectively performs the correlation of two sets of holographic data pertaining to the two 3-D objects to be matched, it is conceivable that currently available 2-D optical joint transform correlation systems can provide 3-D optical image recognition if two holographic data sets are presented as inputs to the 2-D joint transform correlation system. This aspect deserves some future investigation.

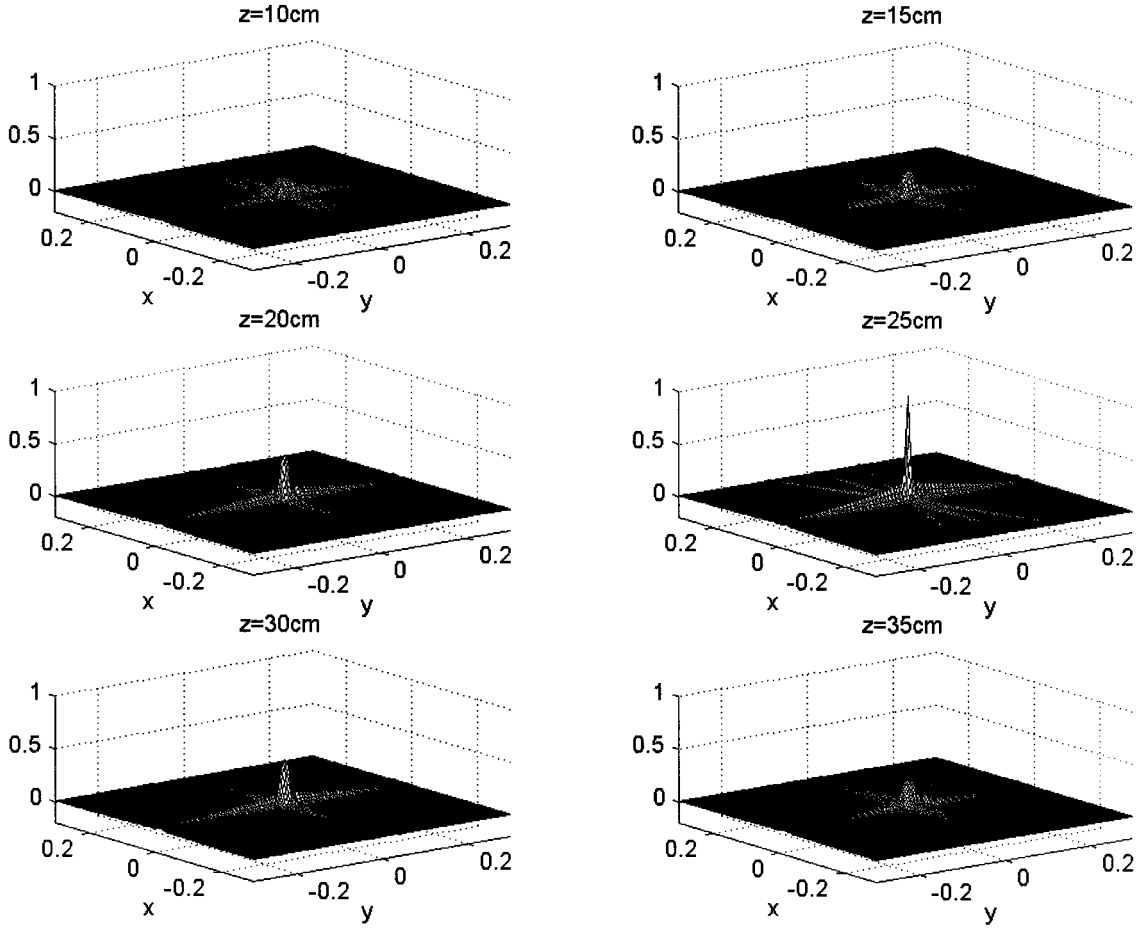

Fig. 8. Correlation outputs as observed by the CCD camera when the camera is translating along the z direction. The two 3-D objects are different in depth by 25 cm, but otherwise they are the same. Note the strong correlation peak at $z = 25$ cm.

Appendix A

The beam splitter BS_1 (see Fig. 1) combines the fields from the two arms of the interferometer to give the total field distribution at a distance z away from the 2-D scanning mirrors:

$$P_{1z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right)\exp(j\omega_0 t) + P_{2z}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \times \exp[j(\omega_0 + \Omega)t], \quad \text{(A1)}$$

where $P_{iz}[(k_0 x/f), (k_0 y/f)]$ is the field distribution z away from the scanning mirrors and is given through Fresnel diffraction by

$$P_{iz}\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) = P_i\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) \otimes h(x, y; z), \quad i = 1, 2. \quad \text{(A2)}$$

In Eq. (A2), $P_i[(k_0 x/f), (k_0 y/f)]$ is the field distribution in the back focal plane of lenses L_1 and L_2 and is given, aside from some nonessential constants and a phase factor, by

$$P_i\left(\frac{k_0 x}{f}, \frac{k_0 y}{f}\right) = \iint p_i(x', y') \times \exp\left[j\frac{k_0}{f}(xx' + yy')\right]dx'dy' = \mathcal{F}\{p_i(x, y)\}_{k_0 x/f, k_0 y/f}, \quad \text{(A3)}$$

where $\mathcal{F}$ denotes the optical Fourier transform operation and is defined as $\mathcal{F}\{u(x, y)\}_{k_x,k_y} = \iint u(x, y)\exp(jk_x x + jk_y y)dxdy = U(k_x, k_y)$, with k_x and k_y denoting the spatial frequencies and with the uppercase function U denoting the transform of the lowercase function u. The term $h(x, y; z)$ is the free-space impulse response and, aside from some phase constants, is given by

$$h(x, y; z) = \frac{jk_0}{2\pi z}\exp\left[-j\frac{k_0}{2z}(x^2 + y^2)\right], \quad \text{(A4)}$$

where k_0 denotes the wave number of the light and, finally, the symbol $\otimes$ in Eq. (A2) denotes the 2-D convolution operation, defined as

$$g_1(x, y) \otimes g_2(x, y) = \iint g_1(x', y')g_2(x - x', y - y')\mathrm{d}x'\mathrm{d}y'. \quad \text{(A5)}$$

Let us return to Eq. (A1); the combined total optical field, or scanning pattern, is used to scan, in two dimensions, an object amplitude transparency $\Gamma_o(x, y; z)$ located a distance z away from the scanning mirrors, as shown in Fig. 1. The photodetector, which responds to the incident intensity of the optical transmitted field or scattered field, generates a current that is given by

$$i(x, y; z) = \int\int_A \left| \left\{ P_{1z}\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right)\exp(j\omega_0 t) + P_{2z}\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right)\exp[j(\omega_0 + \Omega)t] \right\} \times \Gamma_o(x + x', y + y'; z) \right|^2 \mathrm{d}x'\mathrm{d}y'. \quad \text{(A6)}$$

Note that the integration is over the area A of the photodetector, that $x = x(t)$ and $y = y(t)$ represent the instantaneous position of the scanning pattern, and that the shifted coordinates of Γ_o represent the action of scanning. The heterodyne current at the temporal frequency Ω of Eq. (A6), after a bandpass filter tuned at the frequency Ω, becomes

$$i_\Omega(x, y; z) = \mathrm{Re}\left[\int\int_A P_{1z}^*\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right) P_{2z}\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right) \times |\Gamma_o(x + x', y + y'; z)|^2 \mathrm{d}x'\mathrm{d}y' \exp(j\Omega t) \right], \quad \text{(A7)}$$

where we adopted the convention for phasors ψ_p as $\psi(x, y, t) = \mathrm{Re}[\psi_p(x, y, t)\exp(j\Omega t)]$ and $\mathrm{Re}[\cdots]$ denotes the real part of Eq. (A7), which can be written as

$$i_\Omega(x, y) = \mathrm{Re}[i_{\Omega_p}(x, y; z)\exp(j\Omega t)], \quad \text{(A8)}$$

where

$$i_{\Omega_p}(x, y; z) = \int\int_A P_{1z}^*\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right) P_{2z}\left(\frac{k_0x'}{f}, \frac{k_0y'}{f}\right) \times |\Gamma_o(x + x', y + y'; z)|^2 \mathrm{d}x'\mathrm{d}y'$$

is the output phasor and denotes the amplitude and the phase information of the heterodyne current. The amplitude and the phase information constitute the scanned and the processed version of the object $|\Gamma_o|^2$. Defining the correlation operation as

$$g(x, y) \odot h(x, y) = \iint g^*(x', y')h(x + x', y + y')\mathrm{d}x'\mathrm{d}y', \quad \text{(A9)}$$

we can now rewrite the output phasor as

$$i_{\Omega_p}(x, y; z) = P_{1z}\left(\frac{k_0x}{f}, \frac{k_0y}{f}\right) P_{2z}^*\left(\frac{k_0x}{f}, \frac{k_0y}{f}\right) \odot |\Gamma_o(x, y; z)|^2. \quad \text{(A10)}$$

Note that, as in conventional optical scanning systems, only the intensity distribution, i.e., $|\Gamma_o|^2$, is processed, hence the optical system is incoherent, as discussed in the main text.

We now define the OTF of the system as

$$\mathrm{OTF}_\Omega(k_x, k_y; z) = \frac{\mathcal{F}\{i_{\Omega_p}(x, y; z)\}}{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\}}. \quad \text{(A11)}$$

Substituting Eq. (A10) into Eq. (A11), we have

$$\mathrm{OTF}_\Omega(k_x, k_y; z) = \mathcal{F}^*\left\{ P_{1z}\left(\frac{k_0x}{f}, \frac{k_0y}{f}\right) P_{2z}^*\left(\frac{k_0x}{f}, \frac{k_0y}{f}\right) \right\}. \quad \text{(A12)}$$

In terms of the pupils p_1 and p_2, we substitute Eqs. (A2) and (A3) into Eq. (A12) to obtain

$$\mathrm{OTF}_\Omega(k_x, k_y; z) = \exp\left[j\frac{z}{2k_0}(k_x^2 + k_y^2) \right] \times \iint p_1^*(x', y')p_2\left(x' + \frac{f}{k_0}k_x, y' + \frac{f}{k_0}k_y\right) \times \exp\left[j\frac{z}{f}(x'k_x + y'k_y) \right]\mathrm{d}x'\mathrm{d}y'. \quad \text{(A13)}$$

Equation (A12) states that the OTF_Ω of the system can be modified according to the selection of the two pupils. Now, by using Eq. (A11) and rewriting Eq. (A8) in terms of OTF_Ω, we have

$$i_\Omega(x, y; z) = \mathrm{Re}[i_{\Omega_p}(x, y; z)\exp(j\Omega t)] = \mathrm{Re}[\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o(x, y; z)|^2\}\mathrm{OTF}_\Omega(k_x, k_y; z)\} \times \exp(j\Omega t)]. \quad \text{(A14)}$$

By defining the PSF of the optical heterodyne scanning system as $h_\Omega(x, y; z) = \mathcal{F}^{-1}\{\mathrm{OTF}_\Omega\}$, we can now rewrite Eq. (A14) in the spatial domain as

$$i_\Omega(x, y; z) = \mathrm{Re}[|\Gamma_o(x, y; z)|^2 \otimes h_\Omega(x, y; z)\exp(j\Omega t)]. \quad \text{(A15)}$$

Equation (A14) or (A15) represents the scanned and the processed output current modulated by a temporal carrier at frequency Ω. Note that this processed

information is modulated on a carrier at temporal frequency Ω. We can demodulate and extract the in-phase and the quadrature components by mixing it with $\cos(\Omega t)$ and $\sin(\Omega t)$, hence the idea of parallel processing. The demodulation system is shown in Fig. 2, and the two outputs can be shown and given as

$$\begin{aligned} i_c(x, y; z) &= \mathrm{Re}\{\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o|^2\}\mathrm{OTF}_\Omega\}\} \\ &= \mathrm{Re}[|\Gamma_o|^2 \otimes h_\Omega(x, y; z)], \quad \text{(A16a)} \end{aligned}$$

$$\begin{aligned} i_s(x, y; z) &= \mathrm{Im}\{\mathcal{F}^{-1}\{\mathcal{F}\{|\Gamma_o|^2\}\mathrm{OTF}_\Omega\}\} \\ &= \mathrm{Im}[|\Gamma_o|^2 \otimes h_\Omega(x, y; z)], \quad \text{(A16b)} \end{aligned}$$

where $\mathrm{Im}[\cdots]$ denotes the imaginary part of and the subscripts c and s represent the use of $\cos \Omega t$ and $\sin \Omega t$ for mixing, respectively, to extract the information from i_Ω.

We acknowledge the support of the National Science Foundation (grant NSF-ECS-9319211).

References

1. A. B. VanderLugt, "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory **IT-10,** 139–145 (1964).
2. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. **5,** 1248–1249 (1966).
3. Special issue on Advances in Recognition Technique, Opt. Eng. **37**(1), 1998.
4. Special issue on Spatial Light Modulators: Research, Development, and Application, Appl. Opt. **37**(32), 1998.
5. W. W. Stoner, W. J. Miceli, and F. A. Horrigan, "One-dimensional to two-dimensional transformations in signal correlation," in *Transformations in Optical Signal Processing*, W. T. Rhodes, J. R. Fineup, and B. E. A. Saleh, eds., Proc. SPIE **373,** 21–30 (1981).
6. W. T. Rhodes, "The falling raster in optical signal processing," in *Transformations in Optical Signal Processing*, W. T. Rhodes, J. R. Fineup, and B. E. A. Saleh, eds., Proc. SPIE **373,** 11–20 (1981).
7. A. E. Siegman, "Two-dimensional calculations using one-dimensional arrays, or 'Life on the Skew,' Comput. Phys. Nov./Dec. 74–75 (1988).
8. J. Hofer-Alfeis and R. Bamler, "Three- and four-dimensional convolution by coherent optical filtering," in *Transformations in Optical Signal Processing*, W. T. Rhodes, J. R. Fienup, and B. E. A. Saleh, eds., Proc. SPIE **373,** 77–87 (1981).
9. J. Rosen, "Three-dimensional optical Fourier transform and correlation," Opt. Lett. **22,** 964–966 (1997).
10. E. Paquet, P. Garcia-Martinez, and J. Garcia, "Tridimensional invariant correlation based on phase-coded and sine-coded range images," J. Opt. **29,** 35–39 (1998).
11. Y. B. Karasik, "Evaluation of three-dimensional convolution by use of two-dimensional filtering," Appl. Opt. **36,** 7397–7401 (1997).
12. T.-C. Poon and A. Korpel, "Optical transfer function of an acousto-optic heterodyning image processor," Opt. Lett. **4,** 317–319 (1979).
13. T.-C. Poon, "Scanning holography and two-dimensional image processing by acousto-optic two-pupil synthesis," J. Opt. Soc. Am. A **2,** 621–627 (1985).
14. A. W. Lohmann and W. T. Rhodes, "Two-pupil synthesis of optical transfer functions," Appl. Opt. **17,** 11451–1151 (1978).
15. G. Indebetouw and T.-C. Poon, "Novel approaches of incoherent image processing with emphasis on scanning methods," Opt. Eng. **31,** 2159–2167 (1992).
16. D. Gorlitz and F. Lanzl, "Method of zero-order noncoherent filtering," Opt. Commun. **20,** 68–72 (1977).
17. W. T. Rhodes, "Bipolar point spread function synthesis by phase switching," Appl. Opt. **16,** 265–267 (1977).
18. A. W. Lohmann, "Incoherent optical processing of complex data," Appl. Opt. **16,** 261–263 (1977).
19. W. Stoner, "Edge enhancement with incoherent optics," Appl. Opt. **16,** 1451–1453 (1997).
20. E. N. Leith and D. K. Angell, "Generalization of some incoherent spatial filtering techniques," Appl. Opt. **25,** 499–502 (1986).
21. T.-C. Poon, M. Wu, K. Shinoda, and Y. Suzuki, "Optical scanning holography," Proc. IEEE **84,** 753–764 (1996).
22. A. Korpel, *Acousto-Optics* (Marcel Dekker, New York, 1997).
23. P. P. Banerjee and T.-C. Poon, *Principles of Applied Optics* (Richard D. Irwin, Inc., Homewood, Ill., 1991).