Scale-invariant recognition of three-dimensional objects by use of a quasi-correlator

Youzhi Li and Joseph Rosen

A method of scale-invariant recognition of three-dimensional (3-D) objects is presented. Several images of the observed scene are recorded under white-light illumination from several different points of view and compressed into a single complex two-dimensional matrix. After filtering with a single scale-invariant filter, the resultant function is then coded into a computer-generated hologram (CGH). When this CGH is coherently illuminated, a correlation space is reconstructed in which light peaks indicate the existence and location of true targets in the tested 3-D scene. The light peaks are detectable for different sizes of the true objects, as long as they are within the invariance range of the filter. Experimental results in a complete electro-optical system are presented, and comparisons with other systems are investigated by use of computer simulation. © 2003 Optical Society of America

OCIS codes: 100.6890, 100.5090, 100.5010, 100.4550, 110.6880, 110.4190, 090.1760, 090.2870.

1. Introduction

Several different methods of extending two-dimensional (2-D) optical pattern recognition toward operating in three-dimensional (3-D) space have been proposed recently.[1–7] In 1997 the optical correlator was extended to process a 3-D scene containing scattered targets.[1] The basic idea behind this technique is to extend the optical Fourier transform (the correlator's building block) from two to three dimensions. The 3-D correlation was first demonstrated by a 3-D joint-transform correlator (JTC).[2,3] As an alternative, we later proposed a hybrid correlator with general complex filters,[4] which can be regarded as a 3-D extension of the 2-D Vander Lugt correlator.[8] Recently, we proposed a 3-D optical quasi-correlator (OQC).[5] In this method, the spatial spectrum of several images from the 3-D tested scene is compressed into a single computer-generated hologram (CGH). This CGH is then accordingly filtered and coherently illuminated. As a result, a correlation space between the tested scene and the reference is reconstructed from this CGH. The OQC is actually a degenerate version of the 3-D hybrid correlator.[4] In comparison with previous 3-D correlators,[1–4] some of the OQC's performances, such as signal-to-noise ratio (SNR) are degraded, but the 3-D space invariance is maintained and other aspects such as simplicity and operation speed are superior.[5]

All these methods[1–5] involve fusing images of a 3-D tested scene from different points of view and allow targets to be recognized and located in their 3-D surroundings. However, all these schemes[1–7] suffer from similar limitations of the conventional 2-D optical correlators,[8,9] namely, sensitivity to geometrical distortions. A preliminary solution to the problem of sensitivity to distortions was presented in Ref. 10. By this method, a single 2-D synthetic reference function is employed to achieve out-of-plane rotation-invariant 3-D object recognition. Esteve-Taboada *et al*. adopted a different attitude to the problem of distortion invariance in 3-D object recognition.[11,12] They have introduced the Fourier transform profilometry technique[13] into a 2-D optical correlator equipped with rotation-invariant[11] and later scale-invariant filters.[12] Although their system is relatively simple, this technique works well only when all of the objects are located at almost the same transverse plane and in more or less the same distance from the viewer.

In the present study we employ the scale invariant filter in the recently invented OQC.[5] By this technique, three goals are simultaneously achieved. First, the system is scale invariant for a limited range of the object's size. Second, targets can be scattered in the entire observed 3-D scene at different distances

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel.

Received 5 June 2002; revised manuscript received 3 September 2002.

0003-6935/03/050811-09$15.00/0

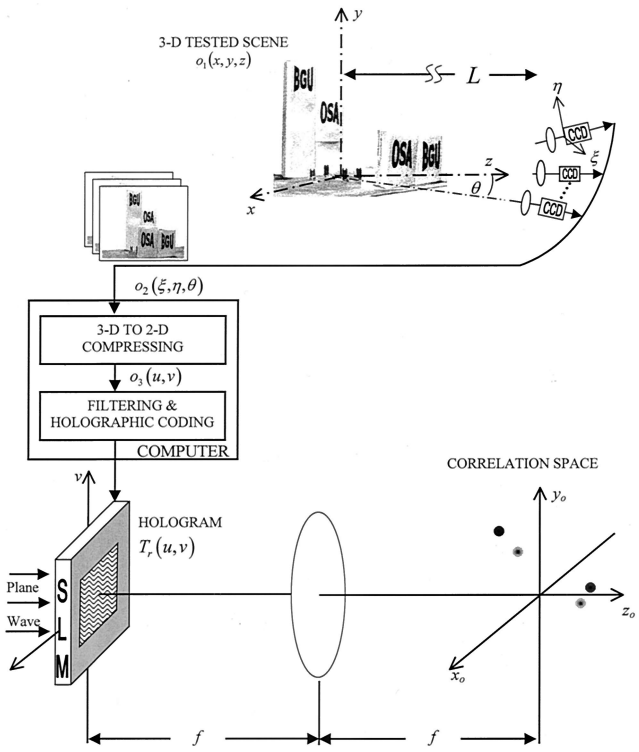

Fig. 1. Schematic of the proposed system.

from the observer, and can still be recognized and located properly. Finally, in comparison to the other 3-D optical correlators,[1–4,10] the inherent simplicity and the reduced computation of the OQC is maintained.

2. Three-Dimensional Scale-Invariant Quasi-Correlation

Before describing the scale-invariant recognition, we briefly summarize the description of the OQC first presented in Ref. 5, and shown in Fig. 1. The tested scene containing several objects is denoted by $o_1(x, y, z)$ in the coordinate system (x, y, z). A digital camera located by a distance L from the origin of (x, y, z) captures the scene function $o_1(x, y, z)$ from several different points of view. The camera is shifted in constant angular steps along a horizontal arc centered on the origin, whereas the camera is always directed toward the origin. The angle between the optical axis of the digital camera and the z axis is denoted by θ. For each θ, the projected image $o_2(\xi, \eta, \theta)$ is recorded into a computer, where (ξ, η) are the coordinates of the camera's image plane for every point of view. Based on simple geometrical considerations, the relation between a point (ξ, η) and a point (x, y, z), for some value of θ, is given by

$$(\xi, \eta) = M_z(x \cos \theta + z \sin \theta, y), \tag{1}$$

where M_z is the transverse magnification factor, and the subscript z indicates that this factor is slowly dependent on the object location along z. This dependence is given by the relation $M_z = d_i/(d_o - z)$, where d_o and d_i are the distances from the origin to the imaging lens and from the imaging lens to the image plane, respectively. The maximum view angle θ is chosen to be small (in the current example, it is 16° from each side of z axis), thus substituting the small-angle approximation: $\cos \theta \cong 1$, $\sin \theta \cong \theta$ to Eq. (1) yields

$$(\xi, \eta) \cong M_z(x + z\theta, y). \tag{2}$$

Inside the computer, the set of captured images are compressed in a particular way to a single spectral matrix as described next. The digital operations along the horizontal and the vertical axes are different because the camera collects the projections of the various viewpoints only along the horizontal axis. Along the horizontal axis there is no need for complete information on the spatial spectrum of every projection. For each kth projection from the view angle θ_k we just calculate, and store, the spectral content of the u_kth frequency element, with a scale factor of $1/(\lambda f)$. λ is the wavelength of a plane wave illuminating the output hologram, and f is the focal length of the spherical Fourier lens used later in the system's output. θ_k and u_k are related by the linear relation $\theta_k = au_k$, and a is some chosen parameter. In other words, from the spatial spectrum of each projected image of some angle θ_k, a different column u_kth is picked up according to the rule $\theta_k = au_k$. But instead of calculating the entire spectrum, we calculate only the value of the spectrum in the relevant column u_kth by the following way: Each kth projection is multiplied by $\exp(-j2\pi\xi u_k/\lambda f)$, and the product is summed along the rows (i.e., along ξ) to a single column. In the following we show that the linear relation $\theta_k = au_k$ is responsible for the desired holographic reconstruction of the correlation space. The parameter a can be an arbitrary real value, and it eventually affects the magnification of the correlation space along the longitudinal axis. Of course, there is no need to calculate the complete horizontal spectrum, but only the columns that are used in the new obtained matrix. The obtained single 2-D matrix in this stage is compressed version of the entire collection of the angular projections. Next, this matrix is Fourier transformed along the vertical axis only, with the above-mentioned scale factor of $1/(\lambda f)$. Thus, in continuous formalism, the dimension of the captured data is reduced from three to two in the following way,

$$o_3(u, v) \propto \iint o_2(\xi, \eta, au)\exp[-i2\pi(u\xi + v\eta)/\lambda f]d\xi d\eta. \tag{3}$$

Note that along ξ the operation is not a Fourier transform because for each frequency value u, the integration is done on a different projection according to the relation $\theta = au$. Now let us look on a single infinitesimal element of the size $(\Delta x, \Delta y, \Delta z)$ at (x, y, z) space, and with brightness $o_1(x, y, z)$ from the entire 3-D tested scene. The element appears as a

single element on every projection plane. Thus, according to relation (3), the point (x, y, z) is transformed into a phase function $o_3(u, v)$ given by

$$o_3(u, v) \propto o_1(x, y, z)\exp[-i2\pi(u\xi + v\eta)/\lambda f]\Delta x\Delta y\Delta z. \quad (4)$$

Substituting Eq. (2) and the relation $\theta = au$ into relation (4) yields

$$o_3(u, v) \propto o_1(x, y, z)\exp[-i2\pi M_z(ux + vy + au^2 z)/\lambda f]\Delta x\Delta y\Delta z. \quad (5)$$

Next we consider the influence of all points of the observed scene $o_1(x, y, z)$ on the function $o_3(u, v)$. Because the scene is 3-D, the overall distribution of $o_3(u, v)$ is obtained as a 3-D integral on all of the points of the input scene, as follows:

$$o_3(u, v) \propto \iiint o_1(x, y, z)\exp[-i2\pi M_z(ux + vy + au^2 z)/\lambda f]\mathrm{d}x\mathrm{d}y\mathrm{d}z. \quad (6)$$

Thus we obtain a 2-D function, which contains 3-D information from the tested scene similarly to the way that a 2-D optical hologram contains the 3-D information of the recorded 3-D scene.[14] First $o_3(u, v)$ is multiplied by a proper filter, and then the obtained product is encoded into a CGH. The CGH, when illuminated by a plane wave, yields a holographic reconstruction of the correlation space, as shown in the lower part of Fig. 1. Thus we can detect the targets and also locate them in their 3-D environment.

Implementing spatial filtering is the next stage. The filter function is synthesized in a similar procedure given by relation (6),

$$H(u, v) \propto \iiint f(x, y, z)\exp[-i2\pi M_0(ux + vy + au^2 z)/\lambda f]\mathrm{d}x\mathrm{d}y\mathrm{d}z. \quad (7)$$

Where $f(x, y, z)$ is a reference function located at the origin of the coordinate system (x, y, z), M_0 is the magnification factor of the imaging system for an object in the origin. Multiplying relation (6) by the complex conjugate of the filter $H(u, v)$, given by relation (7), yields

$$\begin{aligned} T(u, v) = o_3(u, v)H^*(u, v) &\propto \iiint o_1(x, y, z) \\ &\times \exp[-i2\pi M_z(ux + vy + au^2 z)/\lambda f]\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &\times \iiint f^*(-x', -y', -z') \\ &\times \exp[-i2\pi M_0(ux' + vy' + au^2 z')/\lambda f]\mathrm{d}x'\mathrm{d}y'\mathrm{d}z' \\ &= \iiint \iiint o_1(x, y, z)f^*(-x', -y', -z') \\ &\times \exp\left\{-i\frac{2\pi M_0}{\lambda f}\left[u\left(\frac{M_z}{M_0}x + x'\right) + v\left(\frac{M_z}{M_0}y + y'\right) + au^2\left(\frac{M_z}{M_0}z + z'\right)\right]\right\} \\ &\times \mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}x'\mathrm{d}y'\mathrm{d}z' \\ &= \iiint \iiint o_1(x, y, z)f^*\left(\frac{M_z}{M_0}x - x_c, \frac{M_z}{M_0}y - y_c, \frac{M_z}{M_0}z - z_c\right) \\ &\times \exp\left[-i\frac{2\pi M_0}{\lambda f}(ux_c + vy_c + au^2 z_c)\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}x_c\mathrm{d}y_c\mathrm{d}z_c \\ &= \iiint g(x_c, y_c, z_c)\left[-i\frac{2\pi M_0}{\lambda f}(ux_c + vy_c + au^2 z_c)\right]\mathrm{d}x_c\mathrm{d}y_c\mathrm{d}z_c, \end{aligned} \quad (8)$$

where

$$x_c = \frac{M_z}{M_0}x + x', \; y_c = \frac{M_z}{M_0}y + y', \; z_c = \frac{M_z}{M_0}z + z'$$

and

$$g(x_c, y_c, z_c) = \iiint o_1(x, y, z)f^* \times \left(\frac{M_z}{M_0}x - x_c, \frac{M_z}{M_0}y - y_c, \frac{M_z}{M_0}z - z_c\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z. \quad (9)$$

The obtained complex function $T(u, v)$ contains 3-D information of $g(x_c, y_c, z_c)$, which is the correlation function between the tested and the reference objects. The 3-D information exists inside $T(u, v)$ in a similar sense that a 2-D hologram of an object contains the 3-D information of this object.[14] Thus the correlation space is recorded inside the function $T(u, v)$ in a holographic format and can be reconstructed when the resultant hologram is properly illuminated.

The light spots distributed in the reconstructed correlation space indicate the existence and the location of the targets as shown in the lower part of Fig. 1. However, because of the scanning direction along the horizontal axis only, the function $T(u, v)$ is not symmetrical by means of the absence of the term av^2z in relation (8). As a result, the correlation peaks of the true targets obtained out of the back focal plane of the lens are sharp only along the x axis, but slightly spread out along the y axis. This astigmatism is

avoidable if the scanning of the input scene is performed along all the directions of a transversal plane with improved scanning technology. As mentioned, the function $g(x_c, y_c, z_c)$ given by Eq. (9) is the correlation between the tested object function and the scaled reference function. To maintain the same intensity correlation peak for a wide range of object's scale, one needs to introduce a scale-invariant reference function. In the next section we propose using a scale-invariant filter for this purpose.

3. Synthesis of the Scale Invariant Filter

Two different scale-invariant filters for 2-D pattern recognition were proposed almost simultaneously in the end of the 1980s.[15,16] Recently, in a comparative test under similar conditions, Esteve-Taboada *et al.*[12] have shown that the logarithmic radial harmonic (LRH) filter[16] has a longer range of scale invariance than the Mellin harmonic filters have.[15] We follow the recommendation implied from this comparison, and implement the LRH filters in our 3-D OQC. The main points describing the LRH filter are summarized next.

To use the technique of the 2-D LRH, it is assumed that the significant changes of scale from object to object in the scene are in their transverse 2-D shape. This assumption enables us to use the spectral function $H(u, v)$, given by relation (7), of the reference function, as the basis for synthesizing the LRH filter. In case the depth of objects is significant relative to their transversal dimensions, the filter design should be modified accordingly. Such filter design may be a topic for a future research, but besides a different filter distribution, we do not expect any other change in the proposed system or in the present recognition procedure. Using the LRH filter, we intend that the value of the correlation peak intensity will be independent on the value of the scale factor. This goal cannot be completely achieved but only approximated. This is because the change of the scale factor is not a cycle process such as, for instance, changing the object's rotation angle.[16] The LRH filter affects the correlation peak to be approximately related to the scale factor by a phase function, as the following:

$$c_0{}^{\alpha} \cong c_0{}^{1} \exp[i\sigma(\alpha)], \tag{10}$$

where $\sigma(\alpha)$ is a real function of the scale factor α, and $c_o{}^{\alpha}$ is the correlation peak given by

$$\begin{aligned} c_0{}^{\alpha} &= \int_0^{2\pi} \int_d^R o_3\left(\frac{\rho}{\alpha}, \phi\right) \tilde{H}^*(\rho, \phi)\rho d\phi d\rho \\ &= \int_0^{2\pi} \int_{d/\alpha}^{R/\alpha} o_3(\tau, \phi)\tilde{H}^*(\alpha\tau, \phi)\tau d\phi d\tau. \end{aligned} \tag{11}$$

$o_3(\rho/\alpha, \phi)$ is the compressed spectral function obtained from the scaled object according to relation (6) in polar frequency coordinates (ρ, ϕ). $\tilde{H}(\rho, \phi)$ is the filter function defined below. d and R are the internal and external radiuses of the filter, respectively. Thus with power detection in mind, as is shown in Ref. 16, the partial scale invariance in the sense of Eq. (10) is obtained by using a phase only LRH filter defined as

$$\tilde{H}(\rho, \phi) = \exp[i\Omega(\phi)]\left(\frac{\rho}{d}\right)^{i\frac{p}{w}}, \tag{12}$$

where p is the LRH frequency, and w is a normalization constant defined by

$$w = \frac{1}{2\pi} \ln\left(\frac{R}{d}\right). \tag{13}$$

The LRH frequency p has to be properly chosen to optimize the performance of the filter. $\Omega(\phi)$ is an angular phase function that carries all the angular information contained in the phase of $H(\rho, \phi)$ given in relation (7). Thus $\Omega(\phi)$ is computed according to

$$\Omega(\phi) = \arg\left[\int_d^R H(\rho, \phi)\left(\frac{\rho}{d}\right)^{-i\frac{p}{w}} \rho d\rho\right]. \tag{14}$$

The obtained LRH filter given in Eq. (12) is expected to be scale-invariant at least within a limited range of the scale factor. As explained in Ref. 16, the invariance is accomplished approximately for a certain scale range because of α dependence of the integral limits of Eq. (11).

4. Experimental Results

To reconstruct the correlation space from the complex function, the computer should modulate some transparency medium with the hologram values. The spatial light modulator (SLM) that we use in this study can modulate an incident light intensity with positive continuous gray tones. Therefore the complex function $T(u, v)$ is coded into a positive real transparency as follows:

$$\begin{aligned} T_r(u, v) = 0.5\Bigg(1 + \mathrm{Re}\Bigg\{T(u, v)\exp\Bigg[-\frac{i2\pi}{\lambda f}(d_x u \\ + d_y v)\Bigg]\Bigg\}\Bigg), \end{aligned} \tag{15}$$

where (d_x, d_y) is the new origin point of the reconstruction space, and $|T(u, v)|$ is normalized between 0 and 1. In practice the reconstruction space should be stretched far from the zero and the first diffraction orders created from the inherent grating of the SLM. Given that (u, v) represents the indexes of the computer arrays from 1 to N pixels, this goal can be achieved if the factor $(d_x/\lambda f, d_y/\lambda f)$ in Eq. (15) is equal to (1/4, 1/4). Under these conditions, Eq. (15) indicates on the exact computation algorithm of the desired CGH displayed on the SLM.

The holographic reconstruction setup is shown in the lower part of Fig. 1. The SLM is illuminated by a plane wave, which propagates through the SLM and the Fourier spherical lens. $T_r(u, v)$ given in Eq. (15) contains the term $T^*(u, v)$. From the Fourier

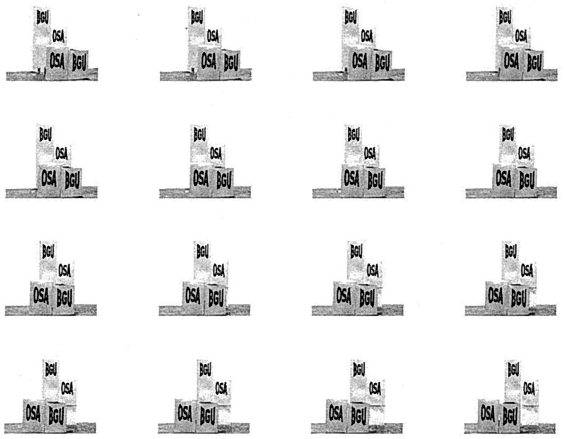

Fig. 2. Sixteen out of sixty-five projections of the tested scene, imaged from −16° to 16°.

transform of this term one can get the output correlation peaks with the same orientation as the original objects in the vicinity of the back focal plane of the Fourier lens around the origin point $(d_x, d_y, 0)$ of the output space.

Two electro-optical experiments were carried out for demonstrating our proposed system. As shown in Fig. 1, four objects are located in the 3-D tested scene. The two with the acronym BGU on their faces are identical. One of them is located in the front of the scene at the point (−2.5, 3, 10.5) cm and the other is in the back at (0, 12, −5.5) cm. The other two with OSA on their faces are also identical and located at (1.5, 3, 10.5) cm and at (−4.5, 8, −5.5) cm. A digital camera observes the 3-D scene from a distance of $L = 68$ cm, and shifted around the origin from −16° to 16° in a 0.5° step between every two successive points of view. At each point of view, an image is recorded into the computer. 16, out of the total 65 projected images of the tested scene are depicted in Fig. 2.

In the first experiment, the two objects with OSA on their faces are regarded as true targets, and should be recognized and located at their 3-D space. The other two with BGU are the false targets, which should be ignored. The LRH filter in this case is computed based on the object of OSA with the size of the front object located at the origin. The optimal frequency of the LRH for this object, in the sense of maximum SNR defined by Eq. (16) given below, is $p = 3.2$. For finding the optimal frequency we calculated the SNR values in the range of p between 0–3.5, with step size of 0.1. The projections of the reference were compressed into a 2-D complex function according to relation (7), and the LRH filter is then computed according to Eq. (12). The phase angle of the resultant LRH filter is shown in Fig. 3(a). Figs. 3(b) and 3(c) are the amplitude and the phase of the spectral function $T(u, v)$ of the correlation space, respectively. The obtained complex function $T(u, v)$ is then encoded into a CGH. The central part of the CGH is

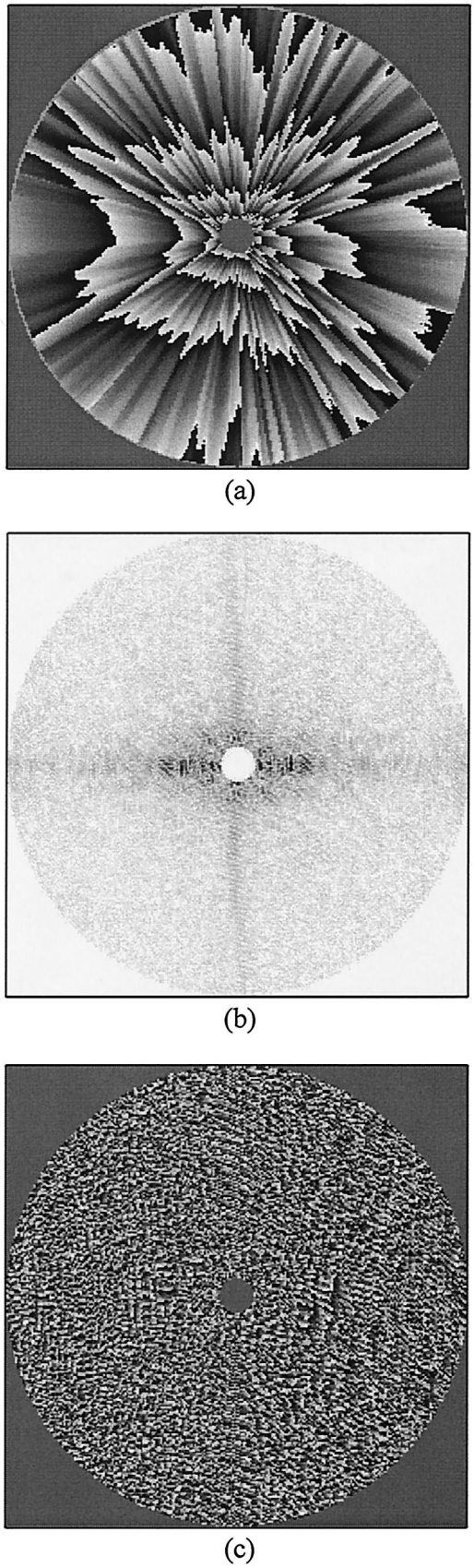

Fig. 3. (a) Phase angle of the LRH filter when OSA is on the reference object, (b) amplitude and (c) the phase angle of the generated spatial spectrum of the correlation space.

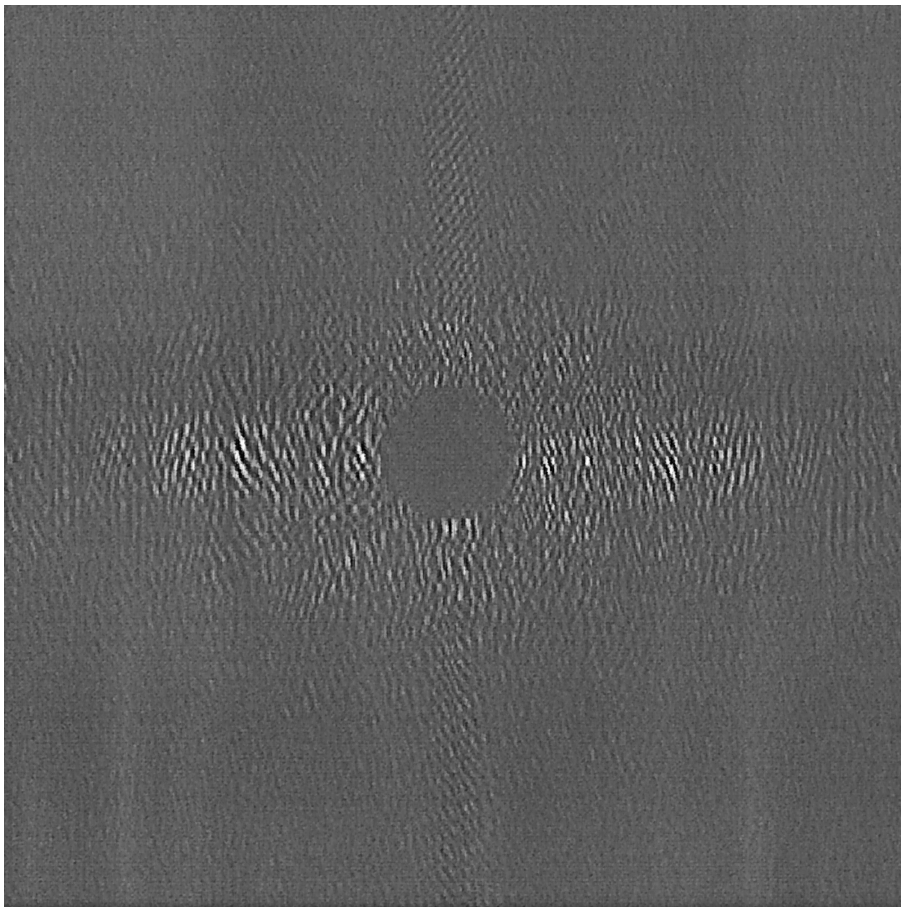

Fig. 4. Central part of the CGH generated from $T(u, v)$ by use of the holographic coding method. The reference is the object with OSA on its face.

depicted in Fig. 4. The CGH was displayed on the SLM in the optical setup shown in the lower part of Fig. 1. The SLM was then illuminated with a plane wave, and the reconstructed correlation planes were recorded in the vicinity of the back focal plane of the Fourier lens along the longitudinal axis. The results are plotted in Fig. 5. Fig. 5(a) was recorded at $z_0 = 7.1$ cm, and Fig. 5(b) at $z_0 = -3.7$ cm. Figs. 5(c) and 5(d) are the 3-D plots of the transverse plane shown in Figs. 5(a) and 5(b), respectively. As shown in Fig. 5, the two objects with OSA on their faces have been successfully detected and located in their 3-D space with longitudinal demagnification of approximately 1.5. The relatively strong light spots indicate the existence and the location of the true targets. Knowing in advance the system's magnification or the demagnification enables one to locate the various targets in the scene, exactly as is commonly done in every imaging or pattern recognition system. In other words, the system's demagnification is the ratio between the real and the measured object's location. The objects with BGU on their faces have not raised meaningful correlation peaks above the noise level in any transverse plane.

In the second electro-optical experiment, the reference object was replaced to be the BGU object. The two BGU objects should be detected now whereas the other two OSA objects should be ignored. As previously, the LRH was computed based on the front BGU object when it is located at the origin of the scene. The optimal LRHs frequency this time is $p = 2.4$. The optically reconstructed planes, on which the correlation peaks are located, are depicted in Fig. 6. Figs. 6(a) and 6(b) are the correlation planes recorded at $z_0 = 7.1$ cm and at $z_0 = -3.7$ cm, respectively. Figs. 6(c) and 6(d) are the 3-D plots of Figs. 6(a) and 6(b), respectively. As shown in Fig. 6, again the task of recognizing the two BGU objects, and ignoring the two OSA objects by the proposed system is successfully achieved. It should be noted that the high noise level on the correlation plane is due to the relatively low quality of the SLM as a holographic display medium. It is expected that improving the SLMs contrast ratio and increasing the display resolution will reduce this noise.

To test the performances of the proposed system, different filters were employed on the same captured data mentioned above. These data were checked by computer simulation of the hologram reconstruction stage. As a measure of comparison we chose the SNR as a degree of discrimination

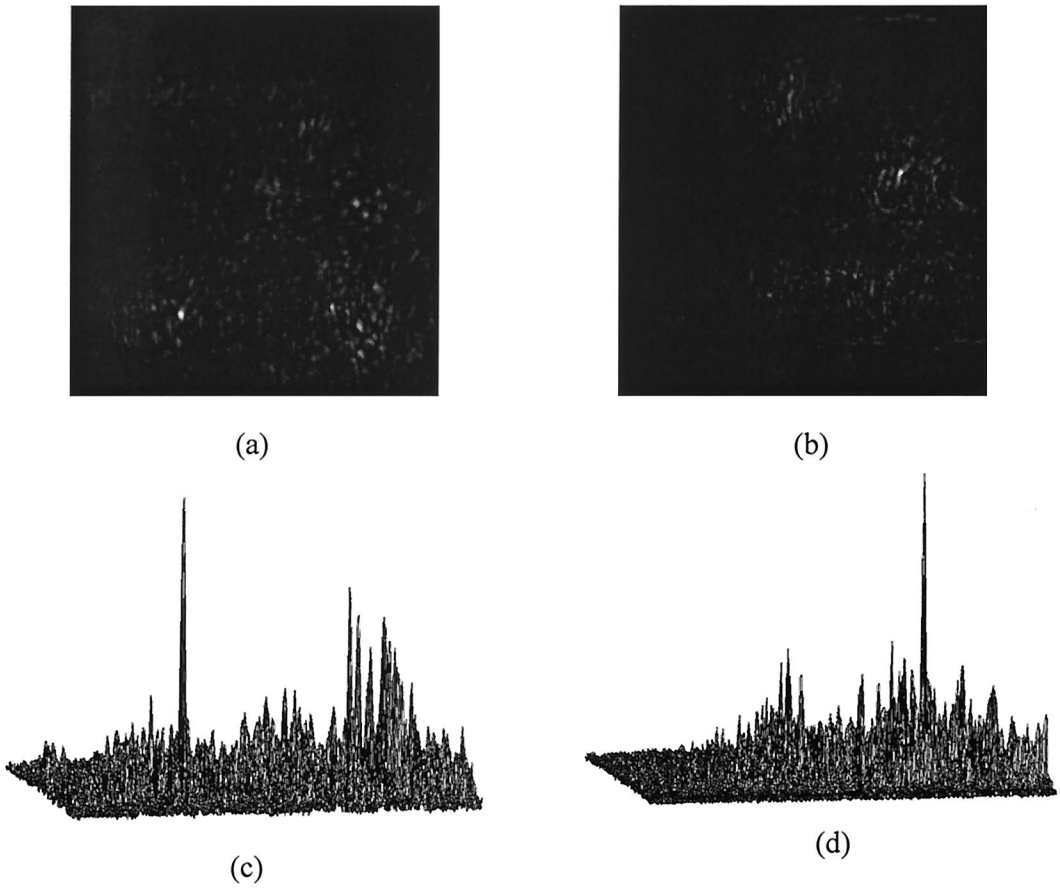

Fig. 5. Optically reconstructed correlation results when the OSA is on the reference object. (a) The pattern recorded at $z_0 = 7.1$ cm, (b) at $z_0 = -3.7$ cm, where the back focal point is at $z_0 = 0$, (c) and (d) are the 3-D plots of (a) and (b), respectively.

between the true and the false targets. The SNR is defined as

$$\mathrm{SNR} = \frac{\text{maximum correlation peak intensity of the true target}}{\text{maximum noise intensity}}. \tag{16}$$

The results of SNR for three kinds of filters are summarized in Table 1. A proper recognition is achieved only if the SNR is greater than 1. All the correlation results in this table are measured in the 3-D OQC, whereas the various filters are synthesized from one of the front objects, each time. As shown in the table, when the front OSA object is taken as the reference, the correlation peak of the back OSA is lost below the noise level for the matched filter (MF) and the phase-only filter (POF). Only the front OSA is recognized when MF and POF are employed, because only these front objects are matched with the reference. However, the two OSA objects are both successfully detected in their 3-D scene when the LRH filter is utilized, as demonstrated by the fifth line of the table. When the reference is switched to be BGU, similar results are obtained. Explicitly, only when the LRH filter is employed, both true targets are identified. However, one of the targets will be lost below the noise level when MF and POF are employed. Thus a 3-D scale-invariant object recognition system is thereby demonstrated.

The scale dependence of the correlation peak for various filters was checked by computer simulation, where a single object was positioned in a virtual scene. The results are depicted in Fig. 7(a) for the three kinds of filters. The asterisks, the rectangles, and the triangles indicate the results of the MF, the POF, and the LRH filter, respectively. The object from which the filter's computation is initiated is the OSA object in a middle-sized scale designated as $\alpha = 1$. When MF and POF are employed and the scale factor α is greater than 1 (i.e., the object appears smaller than the reference) the values of the correlation peaks sharply decrease. When α is smaller than 1, the correlation peak caused by the POF de-

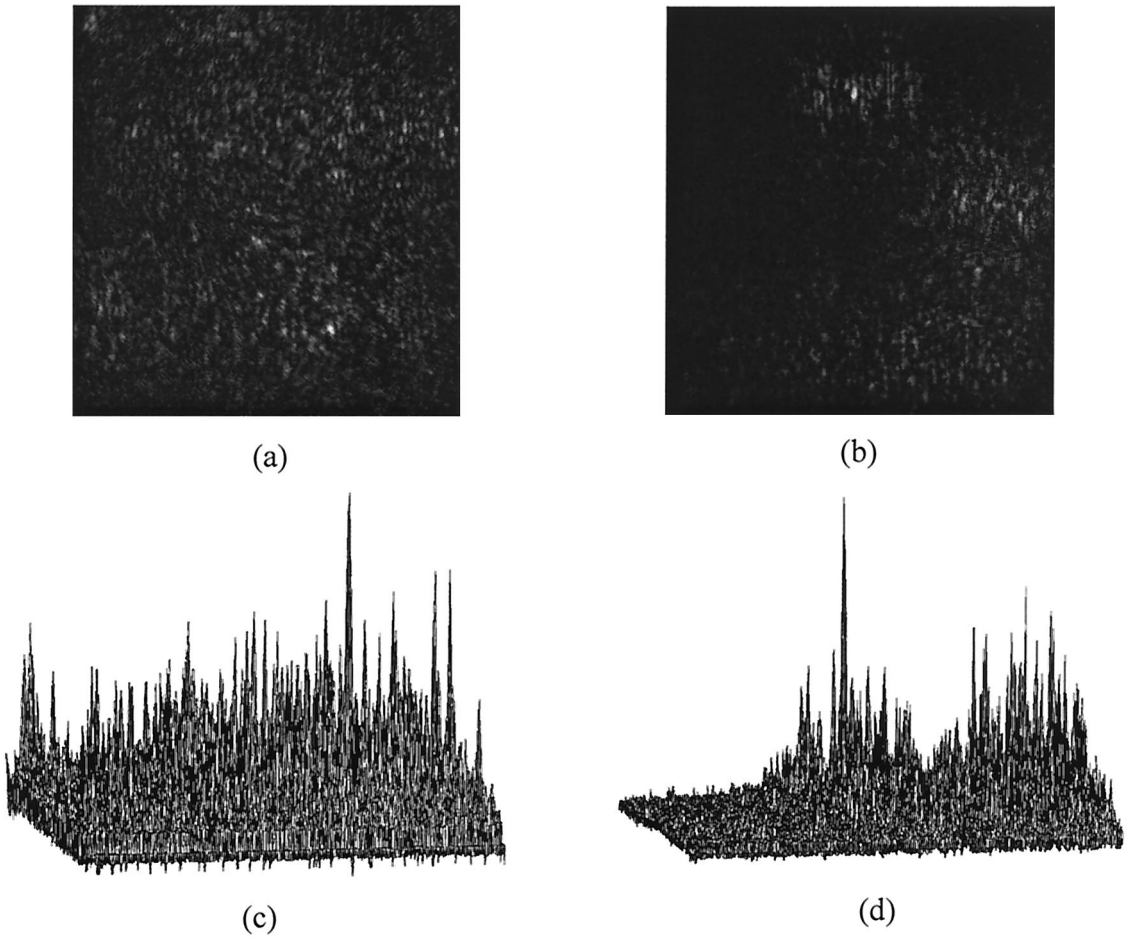

Fig. 6. Same as Fig. 5, but for the reference object with BGU on its face.

creases sharply, but the values of the peaks from the MF are scattered around half of their initial value at $\alpha = 1$. For both directions of α, the peak values caused by the LRH decrease in a relatively moderate curve. The SNR of the various filters versus the scaling factor α are plotted in Fig. 7(b). From this figure, one can see that the SNR of the LRH versus α maintains acceptable and stable SNR values along a wide-scale range, whereas the SNRs of the MF and the POF rapidly decrease to a level close to one and occasionally even below one. We have found that the range in which the SNR is greater than 2 is in the range of $\alpha = 0.4–1.9$ for the OSA sample and $\alpha = 0.4–2.2$ for the BGU sample. Therefore the proposed system equipped with the LRH filters performs better in the sense of scale invariance than the same system with the MF and the POF.

Table 1. SNR Performances of the Proposed System when Matched Filter, Phase-Only Filter, and the LRH Filter Were Employed

Filter	Reference	Back Target	Front Target
Matched filter	OSA	0.3143	1.6384
	BGU	0.3571	1.8599
Phase-only filter	OSA	0.2936	7.9933
	BGU	0.3975	6.2883
Logarithmic radial harmonic filter	OSA	2.3114	6.0064
	BGU	3.3626	4.3107

5. Conclusions

We have presented a technique of scale-invariant 3-D object recognition. This method is based on the recently proposed 3-D OQC.[5] By fusing several images of the tested 3-D scene, and filtering with a single 2-D LRH filter, a 2-D hologram of the 3-D correlation space is generated. When the hologram is coherently illuminated, the 3-D correlation space is reconstructed in front of the observer's eyes. The correlation peaks in the reconstructed space indicate the existences and locations of all true targets in their 3-D scene for a relatively wide range of size of the objects. To compute the 2-D LRH filter, the projections of the reference are first compressed into a 2-D complex matrix. Then, the LRH values are calculated from the compressed matrix. Using a single filter simplifies the system, saves computational time, and yet enables us to achieve a wide degree of scale invariance in the 3-D pattern-recognition system. By using LRH filters this system cannot only detect the targets but also locates them in their 3-D space in a single reconstruction step.

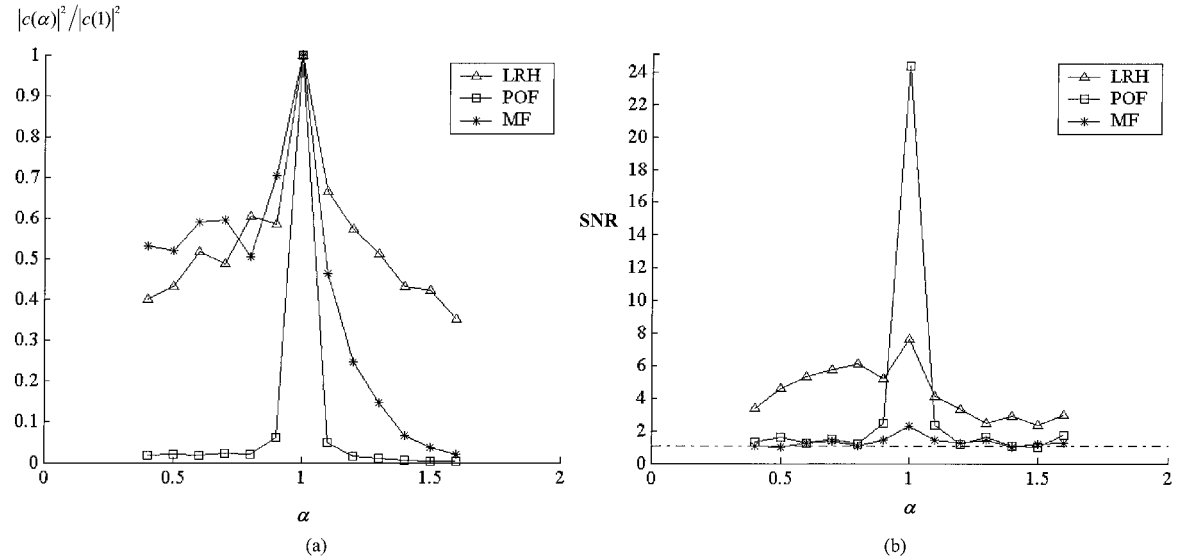

Fig. 7. (a) Scale dependence of correlation peaks normalized to unity. MF, matched filter; POF, phase-only matched filter; LRH, logarithmic radial harmonic filter. (b) Scale dependence of the discrimination ability measured by SNR for different filters.

This study is an additional stage in a long-term effort to create an automatic electro-optical pattern-recognition system that can identify a memory-stored object invariantly to its position in the 3-D space. This time we have achieved and demonstrated shift invariance in three spatial dimensions together with a considerable amount of scale invariance. In addition, the proposed system is capable of identifying the exact locations of the observed objects in a 3-D scene, a feature that rarely exists currently in electro-optical pattern-recognition systems.

References

1. J. Rosen, "Three-dimensional optical Fourier transform and correlation," Opt. Lett. **22,** 964–966 (1997).
2. J. Rosen, "Three-dimensional electro-optical correlator," J. Opt. Soc. Am. A **15,** 430–436 (1998).
3. J. Rosen, "Three-dimensional joint transform correlator," Appl. Opt. **37,** 7538–7544 (1998).
4. Y. Li and J. Rosen, "Three-dimensional correlator with general complex filters," Appl. Opt. **39,** 6561–6572 (2000).
5. Y. Li and J. Rosen, "Object recognition using three-dimensional optical quasi-correlation," J. Opt. Soc. Am. A **19,** 1755–1762 (2002).
6. J. J. Esteve-Taboada, D. Mas, and J. Garcia, "Three-dimensional object recognition by Fourier transform profilometry," Appl. Opt. **38,** 4760–4765 (1999).
7. B. Javidi and E. Tajahuerce, "Three-dimensional object recognition by use of digital holography," Opt. Lett. **25,** 610–612 (2000).
8. A. Vander Lugt, "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory **IT-10,** 139–145 (1964).
9. J. Goodman, *Introduction to Fourier Optics*, 2nd ed., (McGraw-Hill, New York, 1996) Chap. 8.
10. Y. Li and J. Rosen, "Three-dimensional pattern recognition with a single two-dimensional synthetic reference function," Appl. Opt. **39,** 1251–1259 (2000).
11. J. J. Esteve-Taboada, J. Garcia, and C. Ferreira, "Rotation-invariant optical recognition of three-dimensional objects," Appl. Opt. **39,** 5998–6005 (2000).
12. J. J. Esteve-Taboada, J. Garcia, and C. Ferreira, "Optical recognition of three-dimensional objects with scale invariance using a classical convergent correlator," Opt. Eng. **41,** 1324–1330 (2002).
13. M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," Appl. Opt. **22,** 3977–3982 (1983).
14. Y. Li, D. Abookasis, and J. Rosen, "Computer-generated holograms of three-dimensional realistic objects recorded without wave interference," Appl. Opt. **40,** 2864–2870 (2001).
15. D. Mendlovic, E. Marom, and N. Konforti, "Shift and scale invariant pattern recognition using Mellin radial harmonic," Opt. Commun. **67,** 172–176 (1988).
16. J. Rosen and J. Shamir, "Scale invariant pattern recognition with logarithmic radial harmonic filters," Appl. Opt. **28,** 240–244 (1989).