Distortion invariant 3D object recognition using digital holography

Bahram Javidi[*a], Yann Frauel[a], Enrique Tajahuerce[b], Maria-Albertina Castro[a]

[a]Electrical & Computer Engineering Dept., University of Connecticut, USA
[b]Dept. de Ciencies Experimentals, Universitat Jaume I, SPAIN

ABSTRACT

We present a method to recognize three-dimensional objects from phase-shift digital holograms. The holograms are used to reconstruct various views of the objects. These views are combined to create nonlinear composite filters in order to achieve distortion invariance. We present experiments to illustrate the recognition of a 3D object in the presence of out-of-plane rotation and longitudinal shift along the *z*-axis.

Keywords: three-dimensional object recognition, digital holography, distortion-invariant pattern recognition, nonlinear composite filters.

1. INTRODUCTION

Thanks to its massive parallelism and its high space-bandwidth product, free-space optics has proven to be particularly suitable for information processing.[1-15] One of the most widely studied applications is the recognition of objects or images. In particular, the properties of diffraction of light allow the computation of Fourier transforms with a single lens.[2] This explains why optical correlation techniques have been extensively studied for pattern recognition.[1-8] However, although these techniques are easy to use with two-dimensional images, they cannot be applied directly to three-dimensional objects. Yet, it would be very useful to recognize this kind of objects because we actually live in a 3D environment and the use of the three dimensions would improve the accuracy of classification.

Some proposals have been made in order to generalize correlation techniques to 3D objects. A simple method consists in processing several 2D views of both the 3D test scene and the 3D reference object.[16] Another approach is to merge all the views of the reference object into a composite filter which is compared to a 2D picture of the scene.[17] A different possibility is to map 3D functions into 2D images.[18,19] All of these three techniques actually use conventional 2D correlators to perform the computations. More elaborate methods have also been reported in order to use real 3D information on the objects. This information can be contained in a range image obtained from a rangefinder system.[20] In this case, the 3D correlation is performed digitally. The 3D spectrum of an object can also be obtained from a collection of various 2D perspectives, and can then be used in a digital 3D correlation.[21]

Holographic techniques provide a very good way of acquiring 3D information because holograms record the phase of an optical beam as well as its magnitude.[22] A promising approach consists in using digital holography to acquire 3D information about an object. Since this information is then available in digital form, it can be used to compute digital correlations.[23-25] It has thus been demonstrated that it is possible to recognize a 3D object by directly computing correlations between two digital Fresnel holograms.[23] However this method is limited because it provides no distortion tolerance. It can be improved by using the holograms to reconstruct views of the objects thanks to a simulated propagation. These reconstructed views are compared to reconstructed views of the reference object through correlations. The recognition thus becomes translation invariant and it is possible to find the 3D location of the object.[24] However this technique does not allow to recognize distorted objects. In this paper, we show how it is possible to make the 3D recognition really robust to distortions by using nonlinear composite filters.[25] This kind of filters has been used previously for 2D pattern recognition.[4,5,26] We propose here to generalize the principle for recognizing 3D objects recorded with digital holography. The proposed filter is a combination of several matched filters corresponding to

* bahram@engr.uconn.edu; phone +1 860-486-2867; fax +1 860-486-2447; ECE Dept. U-157, 260 Glenbrook Road, Storrs CT 06269-2157, USA.

various distortions of the 3D reference object. As an example, we demonstrate tolerance to out-of-plane rotations and longitudinal shifts. The same approach could be generalized, for instance to achieve tolerance to in-plane rotation or scaling.

2. EXPERIMENTAL SETUP

In Fig. 1 is described the system to record digital holograms. The beam of an Ar laser is divided into one reference arm and one object arm. On both arms, the beams are expanded and spatially filtered. The object beam illuminates a rough 3D object which scatters some light in the direction of the CCD camera. The reference beam crosses one half-wave plate RP_1 and one quarter-wave plate RP_2. The beam is linearly polarized and can be phase-modulated by rotating the two retardation plates. Depending on the position of the fast and slow axes of both plates, we can achieve a phase retardation of $0, \pi/2, \pi$ or $3\pi/2$. We can thus record four phase-shifted interferograms with which we can compute the complex field $U_0(x,y)$ of the wave in the plane of the sensor.[27,28] This function contains the phase of the wave as well as its amplitude and it allows to reconstruct the complex field $U_d(x,y)$ in any plane at distance d from the camera.[23,28] This is done using the Fresnel-Kirchhoff formula:

$$U_d(x,y) = U_0(x,y) * h_d(x,y), \tag{1}$$

where $h_d(x,y) = -\frac{i}{\lambda d}\exp\left(i\frac{2\pi}{\lambda}d\right)\exp\left(i\pi\frac{x^2+y^2}{\lambda d}\right)$ represents the point spread function of the free space, λ denotes the wavelength of the beam, and the symbol $*$ stands for the 2D convolution.

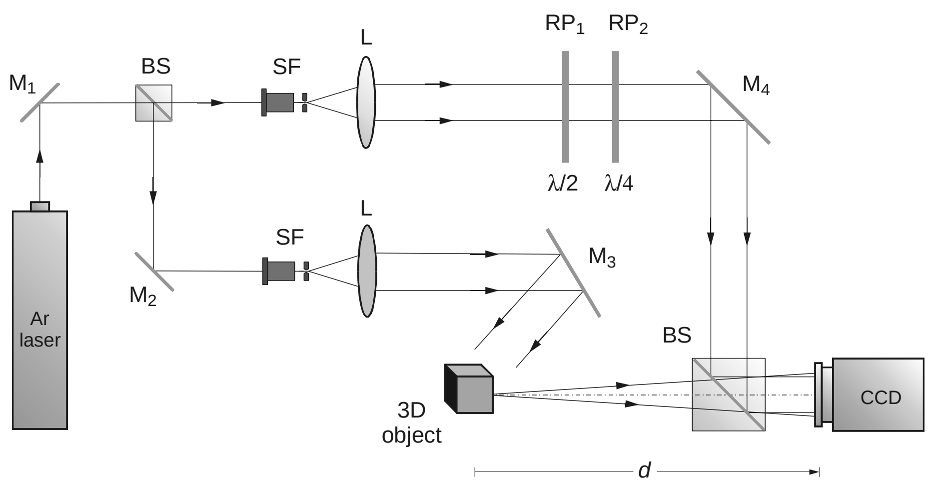
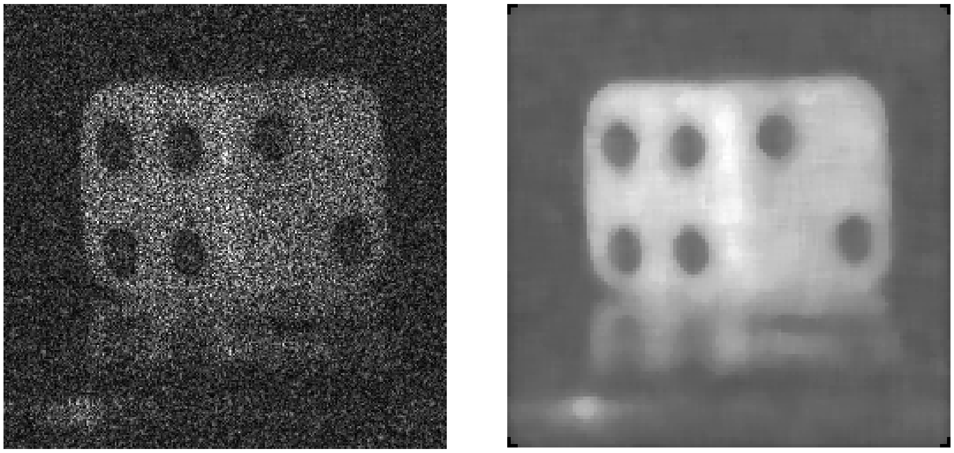


Figure 1. Experimental setup -M mirror, BS beamsplitter, SF spatial filter, L lens, RP retardation plate.

3. MAKING OF THE COMPOSITE FILTER

First of all, we need to improve the quality of the reconstructed views of the object by ridding them of the speckle noise due to the roughness of the surfaces. We can start with dismissing the phase information of the complex amplitude distribution in the reconstruction plane. However, the magnitude of this reconstructed distribution also contains a speckle pattern featuring high frequency spatial variations. We remove this pattern in two steps.[25] First, we average each block of 8×8 pixels of the image. We thus divide the resolution of the image by 8 in both direction which also allows to reduce the computation time. However some magnitude variations remain after this process and we remove them by performing a median filtering over 7×7 pixels. Namely for each pixel we classify the values of the 49 closest neighbours in an ascending order and we keep the 25th value. Fig. 2 presents reconstructed images of a die both before and after filtering the speckle pattern. The levels have been scaled to enhance the contrast.



(a) (b)

Figure 2. Image of a die reconstructed from a digital hologram

a) Before and b) after filtering the speckle pattern.

In order to obtain a high discrimination, we use a *k*th-law nonlinear correlation filter.[5,25] Specifically we compute the Fourier transforms of the filtered images and we raise their Fourier amplitudes to the *k*-th power while retaining their Fourier phase. The case $k = 1$ corresponds to a linear correlation; the case $k = 0$ corresponds to a phase-only filter. For the rest of this paper, we will use $k = 0.1$, which is not a phase-only filter but is still a strongly nonlinear correlation and is therefore highly discriminant.

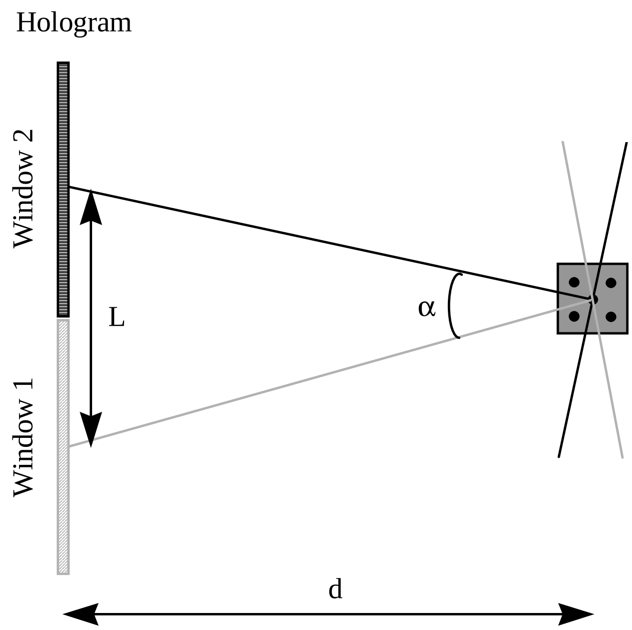


Figure 3. Shifting the reconstruction window inside the hologram results in changing the angle of view of the object.

Our holograms have 2028×2044 pixels but we only use 1024×1024 pixel windows in order to reconstruct the views of the objects. Depending on the location of the window in the hologram, it is possible to reconstruct different perspectives of the object (Fig. 3). Thus, from one or several holograms, we reconstruct a collection of views of the object with various distortions (for instance out-of-plane rotation or longitudinal shift along the *z*-axis). These views are then used to construct a synthetic discriminant function filter.[4,5,25] Specifically, we normalize the energy of these views and we combine them linearly in the Fourier domain in such a way that the resulting composite filter produces the same output peak for each training image.

4. EXPERIMENTAL RESULTS

4.1. Out-of-plane rotation tolerance

Our reference object is a die. We record 19 holograms of this die with several out-of plane rotations. For each new hologram, the die is rotated by roughly 0.5° around the axis which is orthogonal to Fig. 1. The overall rotation angle is around 9°. Although we only use out-of-plane rotation around the vertical axis, our approach is easy to generalize to any axis of rotation. For every hologram, we reconstruct the corresponding image in the plane of the object. These 19 images are our nontraining true targets (Images #1-19). In order to study the robustness of the object recognition, we also record holograms of the die with a very different illumination (Image #20) and in a different 3D position (Image #21). For this latter case, we will present the correlation in the best focus plane. Finally, we need several false targets to test the discrimination of our filter. Hence we use seven various objects (Images #22-28) which are different from the die.

First, we synthetize a nonlinear filter with only one view of the reference die (the "training" image). This view is reconstructed from the same hologram we used for reconstructing Image #10, but with a different window. Fig. 4 presents the values of the correlation peaks for all the images. It can be seen that - except for the one that is very close to the training image - the true targets are barely distinguished from the false targets.

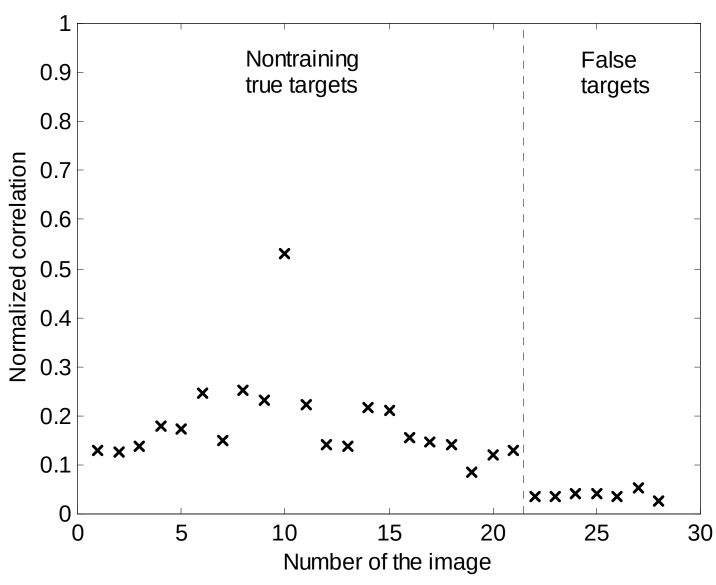


Figure 4. Correlation results obtained with various test objects.

Filter made from one single view.

In order to improve the recognition range, we construct a nonlinear composite filter with several reconstructed views of our reference object.[25] Since we want to achieve rotation tolerance, we have to include in the training images different perspectives of the die. We therefore use three different holograms for reconstructing the views. These holograms are the ones corresponding to Images #3, #10 and #16. They were recorded with the die rotated by -3°, 0° and +3° respectively compared to the orientation for the previous filter. With each of these holograms, we use three different windows: one centered window and one laterally shifted window in both direction. These windows allow to reconstruct views of the die with a regular angle change of 0.6°. We thus have $3\times 3 = 9$ different views of the reference object with which we make our composite filter. Fig. 5 shows the results of the correlations with the test images. The filter has been designed in order to obtain an output peak of 1 for the training targets. For the nontraining images, we obtain obviously an output lower than 1. However, it can be seen that it is easy to discriminate true targets from false targets by using a threshold, for instance at 0.2. When comparing Figs. 4 and 5, it is clear that using a composite filter enhances the recognition of the out-of-plane rotated object.

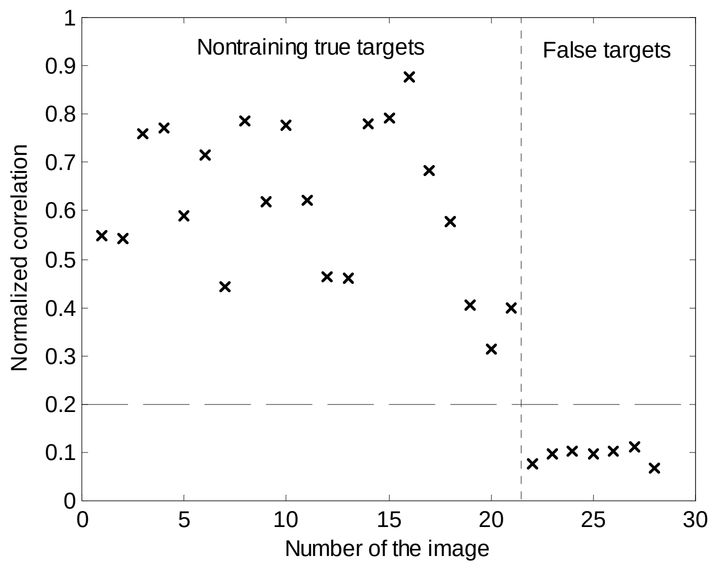


Figure 5. Correlation results obtained with various test objects.

Filter made from 9 views based on 3 different holograms.

4.2. Longitudinal shift tolerance

Even when testing an object which is similar to the reference, we only get a high output peak when the reconstruction plane is the plane of the object. In some cases, it is useful to lower the sensitivity to longitudinal shifts along the *z*-axis. This allows to reduce the number of reconstructions that have to be computed in order to recognize the object. Therefore we design a new composite filter which includes defocused images of the reference object.[25] Namely, we use again the hologram corresponding to Image #10 and also the same three reconstruction windows as above. However, for each window, in addition to reconstructing the image of the object in the focus plane, we also reconstruct views with a defocus of -20 mm, -10 mm, 10 mm and 20 mm. We finally obtain $3\times 5 = 15$ images with which we make the composite filter. We then test the filter with the hologram corresponding to Image #21. The evolution of the output peak value versus longitudinal shift along the *z*-axis can be seen in Fig. 6 for both this new filter and a filter made with the 3 focused images only. It appears that the new filter is less sensitive to longitudinal shift in the reconstruction of the image. Indeed, the recognition of the die is achieved over a wider range of longitudinal shifts along the *z*-axis. Moreover, Fig. 7 gives the output values for all the test images. It appears that the performance of this filter is comparable to the one made with 3 different holograms (Fig. 5), but this time we only need one single hologram to construct the filter.

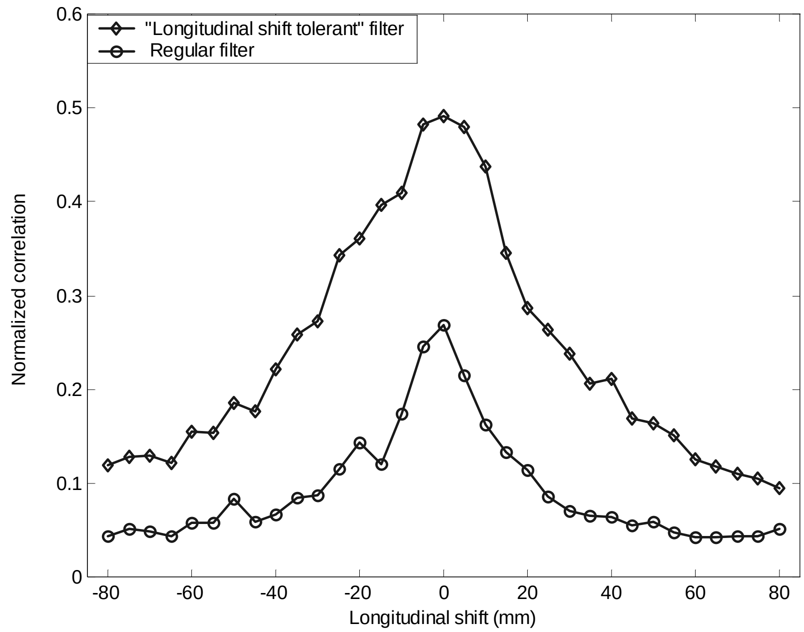


Figure 6. Value of the output peak versus longitudinal shift along the z-axis. Comparison between a filter made with focused images ("regular filter") and a filter including defocused images ("longitudinal shift tolerant filter").

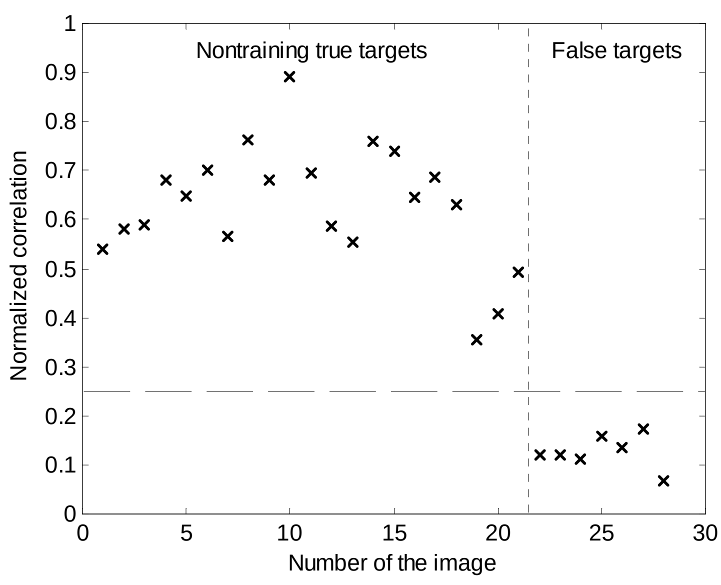


Figure 7. Correlation results obtained with various test objects. Filter made from 15 focused and defocused views based on one single hologram.

5. CONCLUSION

We have described a method to perform pattern recognition of 3D objects based on the ability of a digital hologram to record 3D information on a reference and an input objects. We have demonstrated a technique to achieve distortion tolerance in the recognition process. First, we have shown that the quality of the reconstructed distributions can be improved by filtering the irradiance speckle pattern. Then we have described how to construct nonlinear composite filters to take into account distortions of the reference object. These composite filters are made with several views of the object obtained from one or several digital holograms. As an example, we have demonstrated some tolerance to out-of-plane rotation and to longitudinal shift along the *z*-axis. This technique can be applied to other kinds of distortions.

ACKNOWLEDGEMENTS

M.-A. Castro acknowledges the support of the Consejo Nacional de Ciencia y Tecnología of Mexico.

REFERENCES

1. A. B. VanderLugt, *Optical Signal Processing*, John-Wiley (1992).
2. J. W. Goodman, *Introduction to Fourier optics*, McGraw-Hill, New-York (1968).
3. J. L. Horner and P. D. Gianino, "Phase-only matched filtering," *Appl. Opt.* **23**, 812-816 (1984).
4. D. Casasent, "Unified synthetic discriminant function computational formulation," *Appl. Opt.* **23**, 1620-1627 (1984).
5. B. Javidi and D. Painchaud, "Distortion-invariant pattern recognition with Fourier-plane nonlinear filters," *Appl. Opt.* **35**, 318-331 (1999).
6. B. Javidi and J. Wang, "Optimum Distortion Invariant Filters for Detecting a Noisy Distorted Target in Background Noise, " *JOSA A* **12**, 2604-2614 (1995).
7. A. D. McAulay, *Optical computers architectures*, John Wiley and sons, New York (1991).
8. B. Javidi, J.L. Horner, *Real-time optical information processing*, Academic Press, Orlando, FL (1994).
9. Ph. Réfrégier and B. Javidi, "Optical image encryption based on input plane and Fourier plane random encoding," *Opt. Lett.* **20,** 767-769 (1995).
10. Y. Owechko, "Optical implementation of back-propagation neural networks using cascaded-grating holography," *Int. J. of Opt. Comp.* **2,** 201-231 (1991).
11. F. H. Mok, H. M. Stoll, "Holographic inner-product processor for pattern recognition", *Proc. of SPIE* **1701**, 312-322 (1992).
12. H.-Y. S. Li, Y. Qiao, D. Psaltis, "Optical network for real-time face recognition," *Appl. Opt.* **32**, 5026-5035 (1993).
13. Y. Frauel, T. Galstyan, G. Pauliat, A. Villing, G. Roosen, "Topological map from a photorefractive self-organizing neural network," *Opt. Comm.* **135**, 179-188 (1997).
14. C. Berger, N. Collings, R. Völke, M. T. Gale, T. Hessler, "A microlens-array-based optical neural Network Application," *Pure Appl. Opt.* **6**, 683-689 (1997).
15. B. Javidi ed., *Smart Imaging Systems*, SPIE Press, Bellingham, Washington (2001).
16. A. Pu, R. Denkewalter, and D. Psaltis, "Real-time vehicle navigation using a holographic memory," *Opt. Eng.* **36**, 2737-2746 (1997).
17. S. Deschênes and Y. Sheng, "Three-dimensional object recognition from two-dimensional images using wavelet transforms and neural networks," *Opt. Eng.* **37**, 763-770 (1998).
18. R. Bamler, J. Hofer-Alfeis, "Three- and four-dimensional filter operations by coherent optics," *Opt. Acta* **29**, 747-757 (1982).
19. Y. Karasik, "Evaluation of three-dimensional convolutions by two-dimensional filtering," *Appl. Opt.* **36**, 7397-7401 (1997).
20. J. Guerrero-Bermúdez, J. Meneses, and O. Gualdrón, "Object recognition using three-dimensional correlation of range images," *Opt. Eng.* **39**, 2828-2831 (2000).
21. J. Rosen, "Three-dimensional joint transform correlator," *Appl. Opt.* **37**, 7538-7544 (1998).
22. H. J. Caulfield, ed., *Handbook of Optical Holography,* Academic Press, London (1979).

23. B. Javidi and E. Tajahuerce, “Three dimensional object recognition using digital holography,” *Opt. Lett.* **25**, 610-612 (2000).
24. E. Tajahuerce, O. Matoba, B. Javidi, “Shift-invariant three-dimensional object recognition by means of digital holography,” *Appl. Opt.* (accepted).
25. Y. Frauel, E. Tajahuerce, M. A. Castro, B. Javidi, “Distortion-tolerant 3D object recognition using digital holography,” *Appl. Opt.* (accepted).
26. A. Mahalanobis, “Review of correlation filters and their application for scene matching,” *Optoelectronic Devices and Systems for Processing*, *Critical Reviews of Optical Science Technology*, eds. B. Javidi and K. Johnson, Vol. CR 65, pp. 240-260, SPIE Press (1996).
27. J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, D. J. Brangaccio, “Digital wavefront measuring interferometer for testing optical surfaces and lenses,” *Appl. Opt.* **13**, 2693-2703 (1974).
28. I. Yamaguchi, T. Zhang, “Phase-shifting digital holography,” *Opt. Lett.* **22**, 1268-1270 (1997).