

# Three-dimensional object recognition using a quasi-correlator invariant to imaging distances

Natan T. Shaked\*, Gideon Segev, and Joseph Rosen

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

\*[natis@ee.bgu.ac.il](mailto:natis@ee.bgu.ac.il)

**Abstract:** We present a new method for performing electro-optical three-dimensional (3-D) object recognition under incoherent white-light illumination. Perspective projections of the 3-D scene are acquired from multiple points of view and then processed into a single complex two-dimensional modified Fresnel hologram of the scene. This hologram is processed with a single filter which is matched to a single object, so that all identical objects in the scene yield similar correlation peaks in the 3-D space with almost no dependency on the distances of the objects from the acquisition plane. The new method is demonstrated by experiments.

©2008 Optical Society of America

**OCIS codes:** (070.4450) Correlators; (070.5010) Pattern Recognition; (070.6110) Spatial filtering; (100.3005) Image recognition devices; (100.3008) Image recognition, algorithms and filters; (100.6890) Three-dimensional image processing; (090.1760) Computer holography.

---

## References and links

1. R. Bamler and J. Hofer-Alfeis, "Three- and four dimensional filter operations by coherent optics," *Opt. Acta.* **29**, 747–757 (1982).
2. J. Rosen, "Three-dimensional joint transform correlator," *Appl. Opt.* **37**, 7538–7544 (1998).
3. T. C. Poon and T. Kim, "Optical image recognition of three dimensional objects," *Appl. Opt.* **38**, 370–381 (1999).
4. J. J. Esteve-Taboada, D. Mas, and J. Garcia, "Three dimensional object recognition by Fourier transform profilometry," *Appl. Opt.* **38**, 4760–4765 (1999).
5. B. Javidi, R. Ponce-Díaz, and S. -H. Hong, "Three-dimensional recognition of occluded objects by using computational integral imaging," *Opt. Lett.* **31**, 1106–1108 (2006).
6. J.-S. Park, D.-C. Hwang, D.-H. Shin, and E.-S. Kim, "Resolution-enhanced three-dimensional image correlator using computationally reconstructed integral images," *Opt. Commun.* **26**, 72–79 (2007).
7. D.-H. Shin and H. Yoo, "Scale-variant magnification for computational integral imaging and its application to 3D object correlator," *Opt. Express* **16**, 8855–8867 (2008).
8. Y. Li and J. Rosen, "Object recognition using three-dimensional optical quasi-correlation," *J. Opt. Soc. Am. A* **19**, 1755–1762 (2002).
9. N. T. Shaked, J. Rosen, and A. Stern, "Integral holography: white-light single-shot hologram acquisition," *Opt. Express* **15**, 5754–5760 (2007).
10. N. T. Shaked, B. Katz, and J. Rosen, "Fluorescence multicolor hologram recorded by using a macrolens array," *Opt. Lett.* **33**, 1461–1463 (2008).
11. N. T. Shaked and J. Rosen, "Modified Fresnel computer-generated hologram directly recorded by multiple-viewpoint projections," *Appl. Opt.* **47**, D21–D27 (2008).
12. N. T. Shaked and J. Rosen, "Multiple-viewpoint projection holograms synthesized by spatially-incoherent correlation with broad functions," *J. Opt. Soc. Am. A* **25**, 2129–2138 (2008).
13. J. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996), Chap. 8.

---

## 1. Introduction

Optical three-dimensional (3-D) spatial correlation has been used for numerous applications including 3-D object recognition and target tracking. Several methods of optical 3-D spatial correlation have been suggested in the literature. For instance, Ref. [1] proposes to map the 3-D scene slices on a large two-dimensional (2-D) plane, and then to perform 2-D correlation between each slice and all other slices. More recently [2], one of us has suggested to extend the optical correlation from 2-D to 3-D using 3-D optical Fourier transform. According to this approach, multiple view-point projections (MVPs) of the 3-D scene, illuminated by

incoherent white light, are fused together. The resulting 3-D object function is first 3-D Fourier transformed, then filtered by a 3-D reference filter, and finally inversely 3-D Fourier transformed into the correlation space. Other techniques are the two-pupil optical heterodyne scanning method [3] which uses a laser to scan the objects, and the Fourier transform profilometry method, which projects grating on the objects [4]. Various methods based on a non-holographic integral imaging have also been introduced lately [5–7].

A different approach is presented in Ref. [8] according to which MVPs of the 3-D scene, illuminated by incoherent white light, are fused together into a Fourier hologram of the scene. This hologram is multiplied by a filter matched to one of the objects in the scene, producing a Fourier hologram of correlation peaks. The hologram can be reconstructed to yield the 3-D correlation space containing distinct correlation peaks at points where the object, to be recognized, is located. The system mathematical operation is a 3-D quasi-correlation since the entire correlation space is obtained from a single 2-D hologram. This is in contrast to a complete 3-D correlation [2], where each transverse slice is obtained from a different hologram, and consequently the 3-D correlation consumes much more computational power.

The MVP acquisition can be performed by shifting the camera mechanically or by using microlens [9], marcolens [10], or camera arrays. The three latter methods improve the real-time performances of the system and enable the acquisition of moving objects.

Recently, we have suggested a new type of MVP hologram called digital incoherent modified Fresnel hologram (DIMFH) [11,12]. The advantages of this hologram, compared to other MVP holograms, are a direct, efficient, and more accurate digital process, where no redundant calculations, approximations, or assumptions are needed. However, this hologram also has a 'side effect' according to which the reconstructed size of all objects of equal size in the 3-D scene is the same regardless of their distances to the acquisition plane. As shown in this study, this feature can be useful for optical 3-D object recognition, since only a single quasi-correlation operation has to be performed to detect all identical objects in the 3-D scene, no matter whether they are close to or far from the acquisition plane.

## 2. Description of the method

Figure 1 illustrates the proposed method. The first stage in this method is generating a DIMFH of the 3-D scene. This is done by first capturing the MVPs of the scene, illuminated by incoherent white light. Then, each of the MVPs is multiplied by the same quadratic phase function, and the result is summed into a single pixel in the DIMFH. Let  $P_{m,n}(x_p, y_p)$  be the  $(m,n)$ -th captured projection of the 3-D scene. The DIMFH of this scene is defined by [11,12]

$$H(m, n) = \iint P_{m,n}(x_p, y_p) \exp\left[j2\pi b(x_p^2 + y_p^2)\right] dx_p dy_p, \quad (1)$$

where  $b$  is an adjustable parameter. To reconstruct the recorded 3-D scene, we can illuminate the DIMFH by a coherent plane wave, or alternatively, compute the Fresnel propagation along the optical axis in the computer. In the latter method, the reconstructed plane located at axial distance  $d$  from the hologram is defined as follows:

$$s_d(m, n) = |H(m, n) * Q_d(m, n)|, \quad (2)$$

where  $*$  denotes 2-D convolution, and  $Q_d$  is a quadratic phase function defined as follows:

$$Q_d(m, n) = \exp\left[-j2\pi\gamma(m^2 + n^2)\Delta p^2/d^2\right], \quad (3)$$

where  $\Delta p$  is the pixel size of the camera, and  $\gamma = bf^2\alpha^2/\Delta p^2$  (where  $\alpha$  is the camera gap

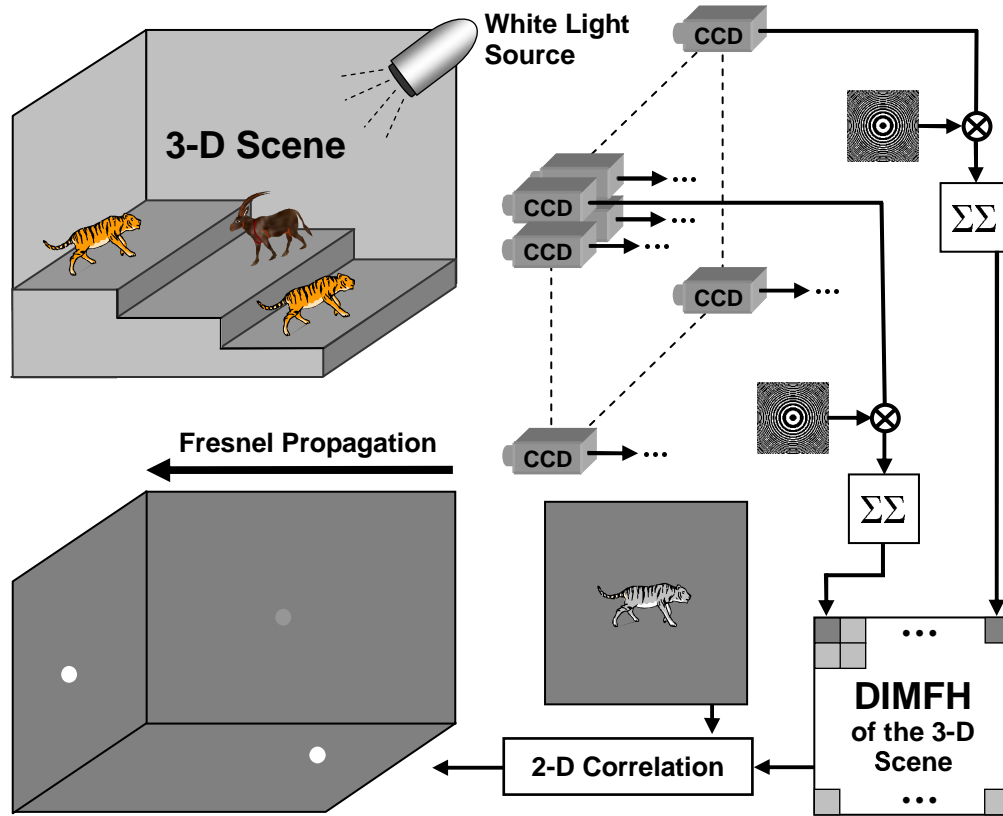


Fig. 1. Schematics of the 3-D object recognition method.

between two adjacent projections, and  $f$  is the focal length of the imaging lens of the camera). The transverse magnification of the DIMFH is constant regardless of the object axial distance, and equal to  $\Delta p/\alpha$  [12], contrary to the transverse magnification of conventional imaging systems which is dependent on the axial distance. Therefore, in case of the DIMFH, objects of equal size, located at different axial distances from the acquisition plane, appear the same size at the reconstructed space. Due to this effect, the hologram is called a modified Fresnel hologram. Another difference from the regular MVP Fresnel hologram is that the reconstruction axial distance  $d$  is in a quadratic relation with the coinciding axial distance in the 3-D scene [11,12].

The point spread function (PSF)  $f(m,n)$  of the correlator is generated beforehand by using the middle projection of the object to be recognized, magnified by  $\Delta p/\alpha$ . Alternatively, a DIMFH of the object to be recognized can be generated beforehand using Eq. (1) [but now each projection is a different perspective view of this single object, located at  $(0,0,d)$ , rather than of the entire 3-D scene], and then convolving this DIMFH with  $Q_d$  to yield  $f(m,n)$ . In both methods, the correlation plane located at axial distance  $d$  from the hologram plane is given by

$$C_d(m,n) = \left| H(m,n) \otimes f(m,n) * Q_d(m,n) \right|, \quad (4)$$

where  $\otimes$  denotes 2-D correlation. The out-of-focus objects (which are not located at axial distance  $d$ ), as well as the in-focus objects (located at axial distance  $d$ ) that do not match the PSF  $f(m,n)$ , do not yield distinct correlation peaks in correlation plane  $C_d(m,n)$ . As explained above, since the transverse magnification of the hologram is constant,  $f(m,n)$  is matched to the object to be recognized in the 3-D scene, never mind whether this object is

close to or far from the acquisition plane. Hence, the highest correlation peaks in the entire correlation space appear at all corresponding positions in which the object to be recognized is located at the 3-D scene.

Note that the correlation between  $H(m,n)$  and  $f(m,n)$  can be performed optically by various optical correlators [13]. In addition,  $H(m,n) \otimes f(m,n)$  itself is a hologram of correlation peaks that can be reconstructed optically by illuminating it with coherent plane wave to yield the 3-D correlation space, or alternatively, as suggested by Eq. (4), can be reconstructed digitally by convolving it with scaled quadratic phase functions.

### 3. Experimental results

We have implemented the electro-optical system shown in Fig. 1. The 3-D scene contained three objects, two identical tiger models and one goat model, at the average size of about  $2\text{ cm} \times 4\text{ cm}$  each. The axial distances to the CCD digital camera (PCO, Scientific 230XS), used in this experiment, were 36 cm for the close tiger, 42.5 cm for the goat, and 50.5 cm for the distant tiger. A set of  $200 \times 200$  MVPs was acquired across a transverse range of  $12.5\text{ cm} \times 12.5\text{ cm}$ . Figure 2 shows eight extreme and the central projections out of the  $200 \times 200$  projections acquired by the camera. The DIMFH  $H(m,n)$  of the 3-D scene, shown in Figs. 3(a) and 3(b), was generated by Eq. (1). Then, the correlation space was obtained by calculating  $C_d(m,n)$  according to Eq. (4) for all axial distances in the range of the 3-D scene, where we used  $f(m,n)$ , shown in Fig. 3(c), having phase-only spatial spectrum.

Figures 4(a)-4(c) show the best-in-focus reconstructed planes obtained from the DIMFH of the 3-D scene according to Eq. (2). As seen in these figures, in each plane a different object is in focus, whereas the other two objects are out of focus. Figures 5(a)-5(c) show 3-D plots of the correlation planes located at the corresponding axial distances used for Figs. 4(a)-4(c),

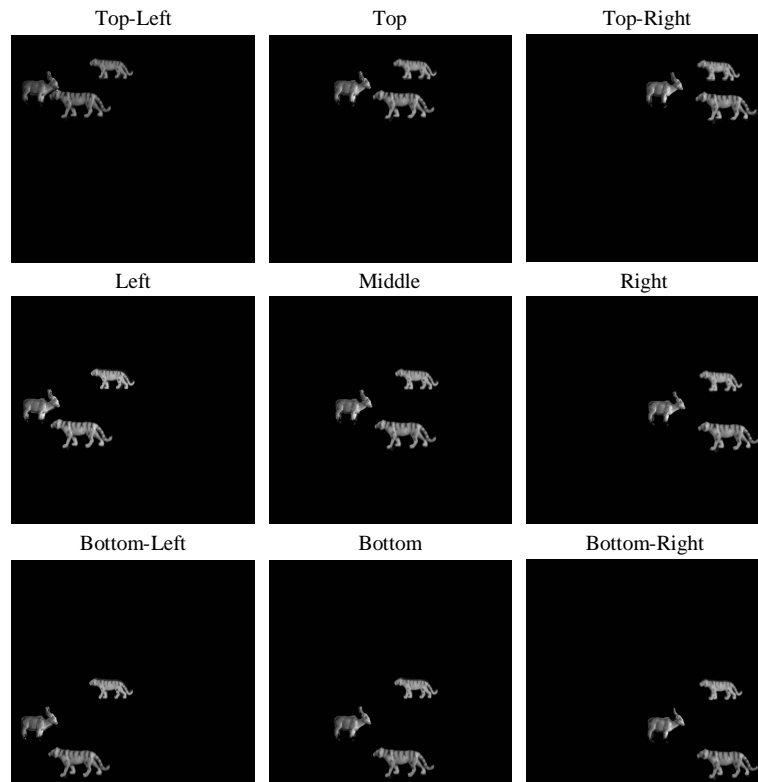


Fig. 2. Nine out of the  $200 \times 200$  projections used to generate the DIMFH of the 3-D scene.

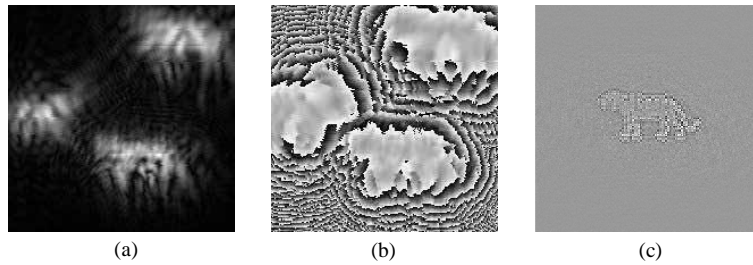


Fig. 3. (a) The resulting DIMFH magnitude; (b) The resulting DIMFH phase; (c) The inverse Fourier transform of the phase-only filter used in the correlation process (the correlator's PSF).

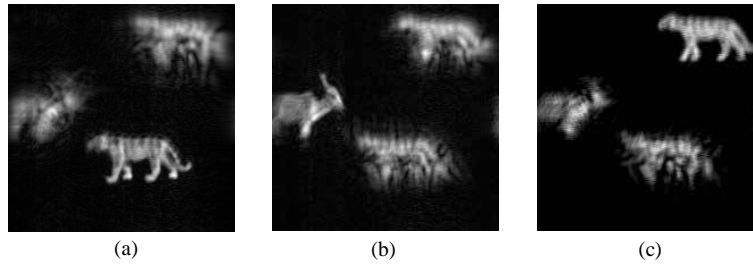


Fig. 4. Best-in-focus reconstructed planes of the hologram shown in Figs. 3(a) and (b): (a) For the close tiger; (b) For the goat; (c) For the distant tiger.

respectively. Note, that all correlation planes were normalized to the maximum value of the entire 3-D correlation space. As seen in Fig. 5(a), a distinct correlation peak appears at the close tiger transverse position, whereas in Fig. 5(c) a distinct correlation peak appears at the distant tiger transverse position. On the other hand, any of the peaks that appear in Fig. 5(b) (obtained at the best-in-focus distance of the goat) is lower than the tigers' peaks [Figs. 5(a) and (c)]. Therefore, the two tigers can be easily recognized by the correlation process, whereas the goat can be rejected by this process since its peak is low even at its best-in-focus reconstruction distance [Fig. 5(b)]. This is still valid although only a single filter was used to detect both the close tiger and the distant one. Figure 6(a) displays three correlation plots along the optical axis of the 3-D correlation space, at the transverse locations of the three objects. These graphs show that the correlation peaks of the two tigers are well-located in the 3-D correlation space, although only one filter is used in the correlation process. Thus, the method can reject the goat and detect each of the two tigers, and it does not matter whether the tigers are close to or far from the acquisition plane.

Since in the DIMFH more distant objects undergo larger magnification, they produce reconstructed images with decreased resolution [12]. Therefore, the distant tiger in Fig. 4(c) seems less clear than the close tiger in Fig. 4(a). This effect becomes smaller as the number of projections increases. Still, even in this demonstration, where this effect is noticeable, both tigers yield distinct correlation peaks as seen in Figs. 5(a) and (c).

To signify the advantage of the proposed method, we also generated an MVP Fourier hologram [8] of the 3-D scene by using the same set of projections (part of which is shown in Fig. 2) and a phase-only filter that is matched to the close tiger. Figures 5(d)-5(f) show 3-D plots of the resulting correlation planes at the same axial distances used for Figs. 5(a)-5(c), respectively. As expected from this old method, only Fig. 5(d), which is obtained at the best-in-focus axial distance of the close tiger, contains a distinct correlation peak. On the other hand, as shown in Fig. 5(f), the correlation plane, obtained at the best-in-focus axial distance of the distant tiger, contains several low correlation peaks, and thus the distant tiger might be wrongly rejected by the correlation process. Hence, in the old method, more than one filter is required to detect the same objects located at different distances from the acquisition plane,

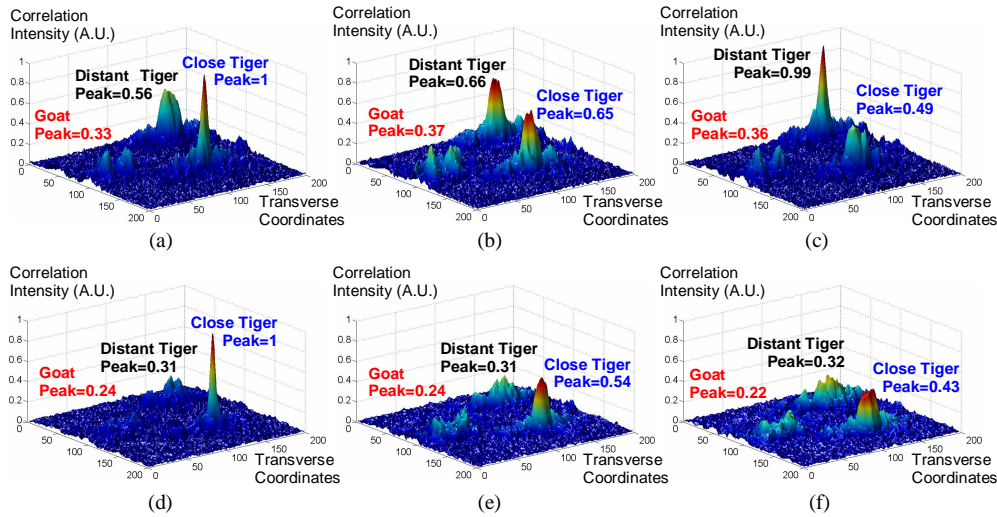


Fig. 5. (a)-(c) The correlation planes corresponding to the reconstruction distances used in Figs. 4(a)-(c), respectively, obtained by the proposed method; (d)-(f) Similar correlation planes obtained by the old method (based on a hologram which does not have a constant magnification). (A.U. = arbitrary units.)

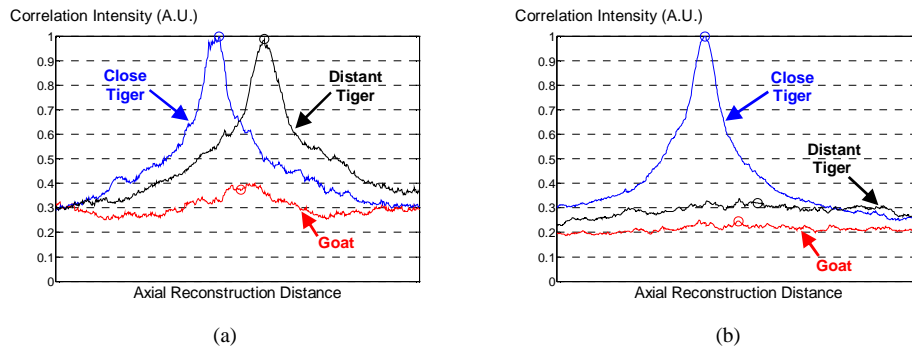


Fig. 6. The correlation peaks along the optical axis of the 3-D correlation space for each of the three objects: (a) In the proposed method (Media 1); (b) In the old method (Media 2); The best-in-focus distance points for each of the objects are circled.

which, as shown above, is not the case of the new method proposed in this paper. This outcome is also obvious by comparing Fig. 6(a) with Fig. 6(b). The latter figure reveals the correlation cross section plots of the old method along the optical axis of the 3-D correlation space at the transverse locations of each of the three objects, whereas Fig. 6(a) shows the corresponding graphs of the new method.

#### 4. Conclusions

We have suggested a new method for 3-D object recognition by a quasi-correlator which is invariant to imaging distances by using a simple digital camera working under incoherent white-light illumination. The method is able to detect, utilizing only a single 2-D filter, all identical objects in the 3-D scene, and it makes no difference whether they are close to or far from the acquisition plane. Since only one filter is used, this method is more efficient compared to the previous similar methods. Experimental results have demonstrated sound performances of the proposed method.