Detection of three-dimensional objects under arbitrary rotations based on range images

Javier García, Jose J. Valles, and Carlos Ferreira

Departamento de Optica. Universitat deValencia. Calle Dr. Moliner 50 Burjassot Spain javier.garcia.monreal@uv.es

Abstract: In this paper a unique map or signature of three dimensional objects is defined. The map is obtained locally, for every possible rotation of the object, by the Fourier transform of the phase-encoded range-image at each specific rotation. From these local maps, a global map of orientations is built that contains the information about the surface normals of the object. The map is defined on a unit radius sphere and permits, by correlation techniques, the detection and orientation evaluation of three dimensional objects with three axis translation invariance from a single range image.

©2003 Optical Society of America

OCIS codes: (100.6890) Three-dimensional image processing; (070.5010) Pattern recognition and feature extraction; (070.2590) Fourier transforms

References and links

- 1. R. Campbell and P. Flynn, "A survey of free-form object representation and recognition techniques," Computer Vision and Image Understanding **81** 2 (2001), pp. 166–210.
- 2. M. Rioux, "Laser range finder based on synchronized scanners," Appl. Opt. 23, 3837-3844 (1984).
- M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," Appl. Opt. 22, 3977-3982 (1983).
- 4. J. Rosen, "Three-dimensional electro-optical correlation," J. Opt. Soc. Am. A 15, 430-436 (1998).
- 5. T. Poon and T. Kim, "Optical image recognition of three-dimensional objects," Appl. Opt. 38, 370-381 (1999).
- E. Tajahuerce, O. Matoba, and B. Javidi, "Shift-Invariant Three-Dimensional Object Recognition by Means of Digital Holography," App. Opt. 40, 3877-3886 (2001)
- E Paquet, H H Arsenault and M Rioux, "Recognition of faces from range images by means of the phase Fourier transform," Pure Appl. Opt. 4, 709-721 (1995).
- P. Parrein, J. Taboury, P. Chavel, "Evaluation of the shape conformity using correlation of range images," Opt. Commun. 195 (5-6), 393-397 (2001).
- Huber DF, Hebert M, "Fully automatic registration of multiple 3D data sets," Image And Vision Computing 21 (7): 637-650 (2003)
- M. Rioux, P. Boulanger and T. Kasvand, "Segmentation of range images using sine wave coding and Fourier transformation," App. Opt. 26, 287-292 (1987).
- E. Paquet, M. Rioux and H. H. Arsenault, "Range image segmentation using the Fourier transform," Opt. Eng. 32, 2173-2180 (1994)
- 12. J. J. Esteve-Taboada, D. Mas, and J. García, "Three-dimensional object recognition by Fourier transform profilometry," App. Opt. **38**, 4760-4765 (1999).
- Esteve-Taboada JJ, Garcia J, Ferreira C, "Rotation-invariant optical recognition of three-dimensional objects," App. Opt. 39, 5998-6005 (2000)
- Hsu Y., Arsenault HH, "Optical-pattern recognition using circular harmonic expansion," App. Opt. 21, 4016-4019 (1982)
- Chang S, Rioux M, Grover CP, "Range face recognition based on the phase Fourier transform," Opt. Commun. 222, 143-153 (2003)
- Hassebrook LG, Lhamon ME, Wang M, Chatterjee JP, "Postprocessing of correlation for orientation estimation," Opt. Eng. 36, 2710-2718 (1997)
- J. J. Esteve-Taboada and J. García, "Detection and orientation evaluation for three-dimensional objects," Opt. Com. 217, 123-131 (2002).
- 18. http://scienceworld.wolfram.com/astronomy/EquatorialCoordinates.html
- 19. B. D. Wandelt and K. M. Górski, "Fast convolution on the sphere," Phys. Rev. D 63, 123002 (2001).
- J. R. Driscoll and D. M. Healy, "Computing Fourier transforms and convolutions on the 2-Sphere," Adv. In App. Math. 15, 202-250 (1994).

#3304 - \$15.00 US		
(C) 2003 OSA		

Received November 07, 2003; Revised November 24, 2003 15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3352

1. Introduction

A considerable effort has been devoted in the last years to optical digitizing, processing and recognition of three dimensional (3D) objects [1-8]. In general, the information of a 3D object is contained in a set of triplets with the (x, y, z) coordinates of the object surface. Detecting an object means scanning over the 6 degrees of freedom (3 rotations and 3 translations) of the object for the best matching with the reference set of points. A significant reduction of the complexity of the problem can be achieved storing a few views of the object as planar images. For instance, a range image can contain the full 3D geometrical information, although only from a given point of view. With a sufficient amount of overlapping between views the full 3D object information can be stored in a few images [9]. Also, range images can be obtained from most of the 3D digitizing techniques [2, 3]. A main limitation of most 3D matching methods is the limited tolerance to rotation in the object. This problem can only be addressed by adding a significant simplification to the problem or by very heavy calculation (see reference [1] for a survey of some digital techniques).

In this paper we propose and demonstrate a method for detection and orientation evaluation of three dimensional objects that have undergone an arbitrary rotation in space. The method is based on computing a map or signature of the three dimensional object based on the local orientation of its surfaces, by means of range images. A matching between this map an the information extracted from a range image from an arbitrary point of view provides a means for detection and, also, for estimation of the object orientation in space.

2. Fourier transforms of phase encoded range images

A range image, z=f(x, y), contains the depth information of an object from a given view line, that defines the z axis. Note that the range image is a single valued function, therefore only the part of the surface closer to the positive z axis is contained on it. The encoding of the depth information has been used in the literature to extend the possibilities of range images [7, 10, 11]. Following this approach, we encode the range image as phase as follows [7]:

$$P(x, y) = exp[iwz(x, y)].$$
(1)

where w is a constant that permits the adjustment of the phase slope of the object.

,

A way to deal with range images keeping translation invariance is to use their Fourier transform. The Fourier transform of the phase encoded range image (PhFT) is then:

$$PhFT(u,v) = F_{2D} \{ exp[iwz(x, y)] \}.$$
⁽²⁾

where F_{2D} stands for two dimensional Fourier transform. Note that w determines the scaling of the Fourier transform frequencies. From now, we will assume w=1, that correspond to the PhFT mapping detailed in the following.

A planar surface of the object, after phase encoding, will become a linear phase factor. Thus its Fourier transform will be peaked around a well defined location. Referring to figure 1(a), the z axis defines the view line, the polar axis is the y axis and the azimuth angle is measured with respect to the z axis. A facet of the object is determined by the angles (α_{y}, α_{y}) , that are the angles of the z axis with the projection of the normal to the planar surface. The tangents of these angles are proportional to the location of the peak in the Fourier domain:

$$(u, v) \equiv (tan(\alpha_x)/2\pi, tan(\alpha_y)/2\pi).$$
(3)

Based on this expression we can relate the peak location with the orientation of the normal in conventional spherical angles, (θ, ϕ) as:

#3304 - \$15.00 US	Received November 07, 2003; Revised November 24, 2003
(C) 2003 OSA	15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3353

$$(u,v) = \left(\frac{\tan(\varphi)}{2\pi} \ \frac{\tan(\theta)}{2\pi\cos(\varphi)}\right). \tag{4}$$

These relations correspond to a value of w=1. A lower value will concentrate the information in low frequencies, while a higher one will expand the scale of the PhFT. From Eq. (4), the PhFT can be coordinate transformed to obtain a distribution $PhFT_{sph}(\theta, \varphi)$, expressed in spherical coordinates, where the location of a peak gives directly the orientation of the corresponding facet. In general, for non-planar surfaces, the PhFT will contain the information of the orientations of the surfaces in the object. Figs. 1(b) and (c) show a sample object and the corresponding PhFT expressed in angular coordinates. It is worth noting that the angular variation for θ and φ has maximum span of ±45 degrees. For higher angles the energy contents of a surface will be small because the apparent size of the facet will be reduced.



Fig. 1. (a) Definition of angular coordinates. (b) Range image of a pyramid shaped object. (c) PhFT intensity.

The intensity of the PhFT exhibit a crucial property: it is invariant to arbitrary translations of the object. This property is obvious for translations in the (x,y) plane, as they will produce just a linear phase factor in the PhFT. On the other hand, from the definition of the phase encoded range image (Eq. 1), a shift along the view line (*z* axis) will influence just as a constant phase factor. In both cases the PhFT is just altered by a phase that is irrelevant in intensity. This property makes the PhFT advantageous for 3D object matching, because the translation invariance is automatically obtained, when performing the correlation between the intensities of the PhFTs [8, 12].

3. Correlation under limited rotations

The problem of rotations is by far more complex than the shifts. Two particular cases have been addressed in the bibliography. In [13] circular harmonic components [14] are used to achieve full rotation invariance under rotations around the view line (z axis). This result is extended in [15] to partial rotations that permits the increase of the information content of the impulse response. The second case involves rotations around an axis perpendicular to the view line [i.e., contained in the (x,y) plane].

In order to analyze this situation we consider two coordinates system. The first one (x,y,z) is attached to the object and rotates with it, whilst the second one (x',y',z') is fixed in space, the z' axis defined by the fixed view line. Assuming a rotation of the object (and its attached coordinate system) around the y' axis of angle ω , from the definition of angular coordinates, it is obvious that a normal defined by (θ, φ) in the rotated system will be represented by the angles $(\theta', \varphi') = (\theta, \varphi + \omega)$ in the fixed system [See Fig. 1(a)]. Therefore, $PhFT(\theta', \varphi')$, expressed in angular coordinates, will undergo just a displacement. Even in spatial frequencies, (u,v), at first order approximation of ω the variations are linear with the rotation angle ω , as can be checked using Eq. (4). A rotation around any axis contained in the (x,y) plane can be reduced to this case simply by choosing the y axis as the rotation axis. This property has been used to obtain detection of 3D objects by correlation of the PhFT intensity with limited tolerance to rotations with respect to an axis perpendicular to the view line [15].

 #3304 - \$15.00 US
 Received November 07, 2003; Revised November 24, 2003

 (C) 2003 OSA
 15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3354

If the full 3D information of the object is known, it is possible to calculate in advance the intensity of the PhFT for any possible orientation. This makes it possible to build a map (3D object orientation map or 3DOOM) that contains the information of all view points by displacing and pasting on a large image the PhFT expressed in spherical coordinates. This 3DOOM contains the information about the normals for all possible angles in a single image, not only around a given view line. This approach is similar to that of a bank of filters [16], although there is a partial overlapping between the impulse responses at different angles. This approach was followed in [17] to achive detection and orientation estimation with rotations around an axis perpendicular to the view line. In this work the 3DOOM was limited to a region around the equator (θ smaller than 45 degrees), because larger nodding of the object was not permitted and thus no information of the facet pointing to the poles was available.

3. Object orientations map on the unit sphere

A first step to extend the detection capability to any arbitrary rotation is to obtain a 3DOOM that covers the full angular range. This step involves the rotation of the model object scanning the possible orientations in such a way that any facet is facing the view line for at least one orientation. Alternatively we can consider the scan on the view line keeping the object fixed. An arbitrary rotation is usually described by the Euler angles. Figure 2(a) shows the convention that we are assuming in the following. With respect to conventional definition, the axes have been permuted cyclically, for a better correspondence with the previous spherical coordinates definition, which assumes that the y axis is the polar one. The arbitrary rotation is decomposed in (a) a rotation of angle α around the y axis (b) a rotation of angle β around the rotated x axis (c) a rotation of angle γ around the new z axis.

In matrix notation this decomposition is expressed as

$$R(\alpha, \beta, \gamma) \equiv R_{Y'}(\gamma)R_{X'}(\beta)R_Y(\alpha).$$
⁽⁵⁾

Note that the last rotation (of angle γ) is not needed for scanning over all (θ , φ) range. The procedure for building the 3DOOM starts by preparing an image for the full (θ , φ) range. Then we scan the view line for all angles (θ , φ), by means of the two rotations around the y and x' axes. For every angle of the view line the following procedure is performed: (a) the range image is computed (b) the PhFT is obtained in the coordinates of the rotated system (θ' , φ') (c) the Intensity of the PhFT is coordinate transformed into the object axes (θ , φ) (d) The resulting image is pasted (by averaging) on the full 3DOOM image. Figure 2 shows a sample 3D image and the corresponding map in spherical angles.



Fig. 2. (a) Definition of Euler angles. (b) Range image of an object. (c) 3DOOM on $\theta {-} \varphi$ Coordinates.

Once the 3DOOM is obtained the task of detection of an object at an arbitrary angular orientation consist on a matching of the intensity of the PhFT for this orientation with the 3DOOM. Unfortunately, although the 3DOOM has been up to now represented as a planar image, it correspond to angular variables, and problems of representation, well known in cartography, arise. In particular, the choice of the polar axis makes that the flat map will be

#3304 - \$15.00 US	Received November 07, 2003; Revised November 24, 2003
(C) 2003 OSA	15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3355

greatly distorted for orientations close to it. The problem derives from the coordinate transformation from local spherical coordinates (θ', φ'), where the PhFT is obtained, into object spherical coordinates (θ, φ), that are the variables for the 3DOOM. This coordinate transformation is highly non linear [18], except for the case of rotations around y axis, as previously discussed. The distortion is indeed an artifact due to the representation of the 3DOOM as a planar map. As in the case of cartography there is no singular point in the representation if the 3DOOM is represented on a unit radius sphere.

3. Correlations on the unit sphere

In a conventional two dimensional correlation, the images to be correlated and the resulting correlation are 2D distributions. The correlation is performed by scanning the spatial coordinates for the matching of one of the two distributions on the other. The scanning is 2D, what defines the dimensionality of the output correlation. The problem is quite different for correlations on the unit sphere. The scan for matching has to be performed on the group of rotations on the sphere (SO3), as described by Euler angles, defining a 3D correlation output. Therefore, the problem is to obtain the correlation between two functions expressed in spherical coordinates. This is equivalent to define the matching between two functions defined on a unit radius sphere, namely $PhFT(\theta', \phi')$ and the 3DOOM. Figure 3 shows an example of this statement of the problem of correlations on the unit sphere



Fig. 3. (a) PhFT of the target shown in figure 2 (b). (b) The same PhFT depicted on the unit sphere. (c) 3DOOM on the unit sphere. The correlation consist on the matching of distributions on figures (b) and (c)

A mathematical definition of the correlation on the unit sphere is [19]:

$$T(\alpha,\beta,\gamma)_{\in SO(3)} = f(\theta,\varphi) * g(\theta,\varphi) \equiv \int [D(\alpha,\beta,\gamma)f](\theta,\varphi)g(\theta,\varphi)sin(\theta)d\theta d\varphi.$$
(6)

where $D(\alpha, \beta, \gamma)$ is the rotation characterized by the proper Euler angles. Note that is not the only definition in the literature. In [20] the correlation is defined for objects with symetry around the view line (or equivalently integrating on last rotation angle).

As in the planar case, it is possible to perform the correlation on the frequency domain, using Fourier transforms. On the unit sphere the Fourier expansion involves the use of spherical harmonics, instead of linear exponential functions. The Fourier transform of an object in terms of spherical harmonics expansion is defined as [19]:

$$f_{lm} = \int f(\theta, \varphi) Y_{lm}^*(\theta, \varphi) \sin(\theta) d\theta d\varphi.$$
⁽⁷⁾

where Y_{lm} are the spherical harmonics functions. Following Wendelt *et al.* [19], the correlation will be defined as:

$$T(\alpha,\beta,\gamma) = \sum_{m,m',m''} T_{mm'm''} e^{im\alpha + im'\beta + m''\gamma}.$$
(8)

 #3304 - \$15.00 US
 Received November 07, 2003; Revised November 24, 2003

 (C) 2003 OSA
 15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3356

$$T_{mm'm''} = \sum_{l} g_{lm} d_{mm'}^{l} \left(\frac{\pi}{2} \right) d_{m'm''}^{l} \left(\frac{\pi}{2} \right) f_{lm''}^{*}.$$
(9)

Where d_{mn}^{l} are the reduced rotation matrices [21] and f_{lm} and g_{lm} are the spherical Fourier transforms. In our case *f* and *g* are the PhFT from a given (unknown) point of view and the 3DOOM, respectively. Therefore, following this definition, the output of the correlation will be three dimensional, and the location of the correlation peak will give the (α , β , γ) triplet with the object orientation. This correlation admits a rapid implementation using the fast Fourier transform algorithm [20].

To test the performance of this method of 3D object matching we have implemented it digitally. In order to simplify the interpretation of the results, without losing generality, the rotations are limited to the two first rotations (the rotation around the final polar axis is not permitted). This way the correlation output is bidimensional and can be directly represented.

The target in our tests is the one depicted in Figure 2(b). Figures 4 (a-c) show a range image corresponding to an arbitrary orientation of the target, the corresponding intensity of the PhFT and its correlation with the 3DOOM [see Fig. 3(c)]. The correlation has been represented as a planar image. The correlation shows a distinct correlation peak from which the orientation of the object ($\alpha = -120^\circ$, $\beta = 53^\circ$) can be estimated. It is clear that the method provide a correlation that can sharply determine the presence and the orientation of the object.

The discrimination capabilities of the method have been tested by using another object shown on Figure 4(d). The PhFT for an arbitrary orientation [Fig. 4(e)] along with the correlation the correlation, with the same intensity normalization as the correlations for the true target, [Fig. 4(f)] are shown. The low output correlation value discriminates this object from the target. Additional non-shown simulations demonstrate that this discrimination happens for all values of rotations of the false target.



Fig. 4. (851KB) (a) Range image of the 3-D object rotated with α =-120°, β =53.4°. (b) PhFT of the range image given in (a). (c) Output of the correlation. (d) Range image of another object. (e) PhFT of the range image given in (d). (e) Output of the correlation.

4. Conclusions

We have proposed and demonstrated a method for detection of three dimensional objects. The method is based on constructing a map of orientations of the object on a unit sphere using the intensity of the PhFT of range images. The results prove the capability for detection as well as to estimate the orientation of the detected object.

#3304 - \$15.00 US (C) 2003 OSA Received November 07, 2003; Revised November 24, 2003 15 December 2003 / Vol. 11, No. 25 / OPTICS EXPRESS 3357

Acknowledgments

This work was supported by FEDER funds and the Spanish Ministerio de Ciencia y Tecnología under the project BFM2001-3004. Jose J. Vallés acknowledges the grant from the Spanish Ministerio de Educación, Cultura y Deporte.