Three-dimensional object detection under arbitrary lighting conditions

José J. Vallés, Javier García, Pascuala García-Martínez, and Henri H. Arsenault

A novel method of 3D object recognition independent of lighting conditions is presented. The recognition model is based on a vector space representation using an orthonormal basis generated by the Lambertian reflectance functions obtained with distant light sources. Changing the lighting conditions corresponds to multiplying the elementary images by a constant factor and because of that, all possible lighting views will be elements that belong to that vector space. The recognition method proposed is based on the calculation of the angle between the vector associated with a certain illuminated 3D object and that subspace. We define the angle in terms of linear correlations to get shift and illumination-invariant detection. © 2006 Optical Society of America

OCIS codes: 070.0070, 070.4550, 100.5010, 100.6890.

1. Introduction

Much effort has recently been devoted to optical digitizing, processing, and recognition of three-dimensional (3D) objects.[1–8] For example, automatic measurements of 3D object shape can be accomplished using Fourier transform profilometry.[3] Such surface measurement techniques have also been used for 3D pattern recognition.[9] Other 3D recognition methods obtain a range image of the 3D object as an intermediate step of the recognition task.[7,10,11] Another approach is to use digital holographic techniques,[6,12] but such methods require a long time to calculate the 3D optical field. Some 3D recognition methods use optical correlations. Bamler and Hofer-Alfeis[13] showed a method for performing 3D optical correlations by slice-by-slice mapping the 3D observed scene along its longitudinal axis. After mapping, they used conventional 2D optical correlations between each slice and all the others. However, the algorithms to reconstruct the 3D image require intensive computer resources. Rosen has extended the correlation from 2D to 3D by introducing a 3D optical Fourier transform.[4] By fusion of several projections of the tested scene, a 3D object function is first Fourier transformed, then filtered by some 3D reference filter, and finally inversely Fourier transformed into the correlation space, so a target can be detected and located in its 3D environment. This highly complex method requires extensive use of digital computation. To reduce the difficulty, the authors proposed an improvement of the 3D correlation space.[14] Matoba *et al.*[15] captured multiple perspectives of 3D objects by a microlens array, after which all the perspectives of the reference and of the input scene were cross correlated by a 2D joint transform correlator (JTC). The authors showed that the system can recognize 3D objects with slightly out-of-plane rotations. In fact, distortion, rotation, and scaling changes of 3D objects are a challenge in many detection systems. Some of the above 3D recognition methods have been extended and new ones were proposed to deal with that automatic target recognition issue.[16–19] Recently, the addition of object color information in the 3D detection process has been considered.[20] The method is based on a multichannel correlation in various chromatic systems based on Fourier-transform profilometry.

Although previous methods dealt with important topics in correlation-based 3D pattern recognition, taking into account changes of illumination of objects has received less attention. It is true that topics related to illumination such as shape from shading, and

J. J. Vallés, J. García, and P. García-Martínez (pascuala.garcia@uv.es) are with the Departament d'Òptica, Universitat de València, C/Doctor Moliner, 50, 46100 Burjassot, Spain. H. H. Arsenault is with the Departement de Physique, Génie Physique et Optique, Université Laval, Quebec, G1K 7P4, Canada.

Received 21 November 2005; accepted 30 January 2006; posted 14 February 2006 (Doc. ID 66083).
0003-6935/06/215237-11$15.00/0

photometric stereo have been extensively studied in computer vision.[21] We are particularly interested in detection methods based on linear subspaces created by an image database with multiple illumination changes of a given object.[22,23] Those processes have been mainly applied to face recognition, where the recognition compared a new query image to each model in turn. To compare the models, a measure of similarity is required. The differences between 3D recognition methods depend on the selection of the subspace and on the measure of similarity applied. Other recognition methods are vision systems based on principal components analysis (PCA) and various Lambertian reflectance illumination models.[23] Although the above methods are considered as 3D recognition, in fact they are not, because no 3D image is utilized. Instead, a wide 2D database is created with all perspectives and changes in illumination for given objects, usually faces.

In this paper, we introduce a 3D recognition method based on the detection of 3D objects under changes of illumination conditions. The illumination model is defined in terms of Lambertian reflectance surfaces, surfaces that scatter light equally in all directions and appear equally bright from all viewing directions. We generate a subspace using an orthonormal basis. The basis is defined in terms of different directional light sources. The recognition process uses the calculation of the cosine of the angle between a target and the vector subspace. If the target belongs to that subspace, the vector that corresponds to the target will be contained in the subspace, the angle will be zero, and the cosine will be equal to unity. On the other hand, if the target does not belong to that space, the cosine will be smaller than 1. We will define the cosine measurements in terms of correlations. From the point of view of vector spaces, intensity invariant pattern recognition consists of recognizing vectors independently of their length, which can be viewed as an angle measurement between vectors in vector spaces. This angle provides a measure of the similarity between the object and the reference function.

Lefebvre *et al.*[24] defined a nonlinear filtering method called the locally adaptive contrast invariant filter (LACIF), which is invariant under any linear intensity transformation. This LACIF operation uses three correlation operations involving local statistics and nonlinearities. It was applied directly to scenes containing unsegmented targets. One of the advantages of the LACIF method is that no *a priori* information about the constant values involved in the linear illumination model is assumed. The LACIF method can be combined with synthetic discrimination filters to achieve both illumination invariance and out-of-plane rotation invariance.[25] The authors recently[26] generalized the LACIF filtering for situations where an additional linear intensity gradient across an object is present. It is interesting to consider the LACIF technique in the context of a vector space interpretation. In this paper, we have extended the LACIF idea to 3D object recognition, and we compare both illumination invariant detection methods.

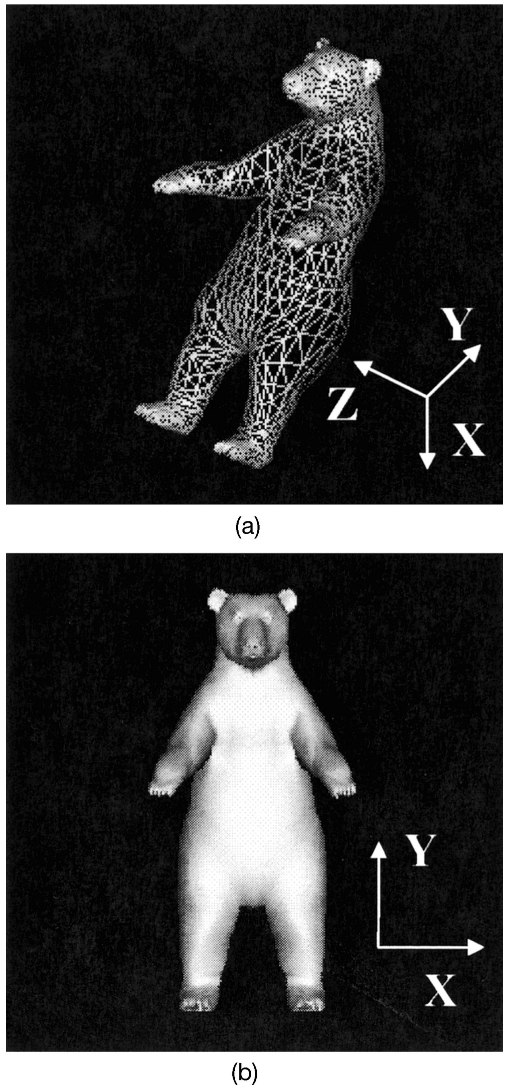


Fig. 1. (a) 3D object mesh. (b) Sample view of 3D object shaded.

The paper is organized as follows: In Section 2 the different illumination techniques for 3D objects are introduced. The 3D image representation in terms of vector spaces is in Section 3. The results comparing the new method with the LACIF are presented in Section 4, and in Section 5 we present our conclusion.

2. Illumination Model for Three-Dimensional Images

Variations in lighting can have a big effect on the appearance of 3D images. When the sources of light are far from the object, the lighting conditions may be described by specifying the intensity of light as a function of direction. Our illumination model is defined in terms of ambient illumination and Lambertian illumination. An ambient light is a diffuse, nondirectional source of light, typically the result of multiple reflections of light from many surfaces in the environment. It has a uniform intensity at any point

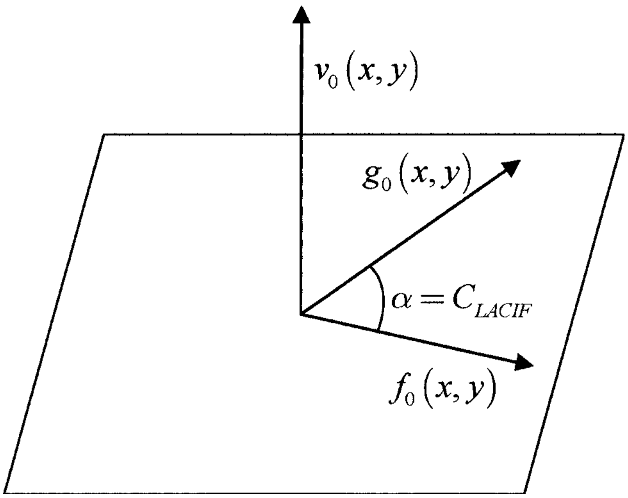


Fig. 2. Geometrical interpretation of LACIF.

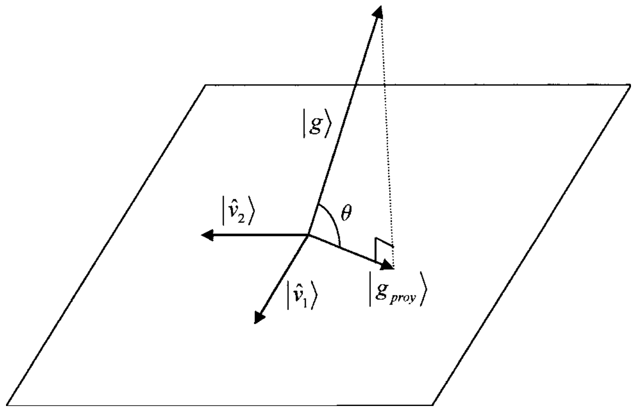


Fig. 3. Definition of the angle between a vector and the subspace.

in the environment. The influence of ambient illumination for a given surface can be expressed mathematically as

$$I_1(x, y, z) = I_a k_a(x, y, z), \quad (1)$$

Plp;&-3.77q where I_a is the intensity of ambient light, and $k_a(x, y, z)$ is the ambient-reflection coefficient of an object's surface that is determined by the material properties of the surface. Ambient light will generate a uniform intensity across the surface of an object. Note that Eq. (1) gives an intensity distribution for a 3D space. Moreover, we assume that, aside from the ambient light, the surface of the object reflects light according to Lambert's law[27] from localized light sources as

$$I_2(x, y, z) = I_p k_l(x, y, z)(\mathbf{L}(x, y, z) \cdot \mathbf{N}(x, y, z)), \quad (2)$$

where I_p is the point light source's intensity, $k_l(x, y, z)$ is the material's diffuse reflection coefficient, $\mathbf{L}(x, y, z)$ is the unit vector of light direction, $\mathbf{N}(x, y, z)$ is the normal unit vector of the surface's object for a given point, and $\cdot$ represents the inner or scalar product.

Finally our illumination model is given by the contribution of ambient illumination [Eq. (1)] and Lambertian illumination [Eq. (2)] as

$$I(x, y, z) = I_a k_a(x, y, z) + I_p k_l(x, y, z) \times (\mathbf{L}(x, y, z) \cdot \mathbf{N}(x, y, z)). \quad (3)$$

From Eq. (3) we may obtain images that are not physically realizable, because the corresponding linear combination may contain negative values. These areas define the shadows, so a point P of the surface is a shadow if the angle between the surface normal and the direction of light is obtuse, i.e., $\cos(\mathbf{L} \cdot \mathbf{N}) \leq 0$. We shall assume that when the linear combination of Eq. (3) produces negative gray values, those values can be set to zero. This criterion is widely used for purposes of display or recognition.

Although the illumination model considered in this paper deals with the 3D information of an object, from now on we will consider the information of Eq. (3) as looking at the object from a particular point of view or perspective. It is similar to taking a picture of the 3D image, so the images will be two dimensional. Figure 1(a) shows a 3D object and Fig. 1(b) the corresponding 2D object for a given illumination and perspective. Figure 1(b) has been calculated as

$$I(x, y) = \begin{cases} I_a k_a(x, y) + I_d k_l(x, y)[L_x N_x(x, y) + L_y N_y(x, y) + L_z N_z(x, y)] & \text{if } (\mathbf{L} \cdot \mathbf{N}) \geq 0 \\ I_a k_a(x, y) & \text{otherwise} \end{cases}, \quad (4)$$

where $\{N_i(x, y)\}_{i=x,y,z}$ and $\{L_i\}_{i=x,y,z}$ are the Cartesian coordinates of $\mathbf{L}$ and $\mathbf{N}(x, y)$, respectively. In the literature, there are more complex illumination models[27] that take into account only matte surfaces, etc.

3. Vector Spaces for Images

Images may be considered as vectors in a Hilbert space. Then, any vector can be expressed in terms of an given basis. Correlation is a measure of the similarity between images, because one way to understand the correlation operation is to consider it as an inner product between two functions, the object and the reference functions.

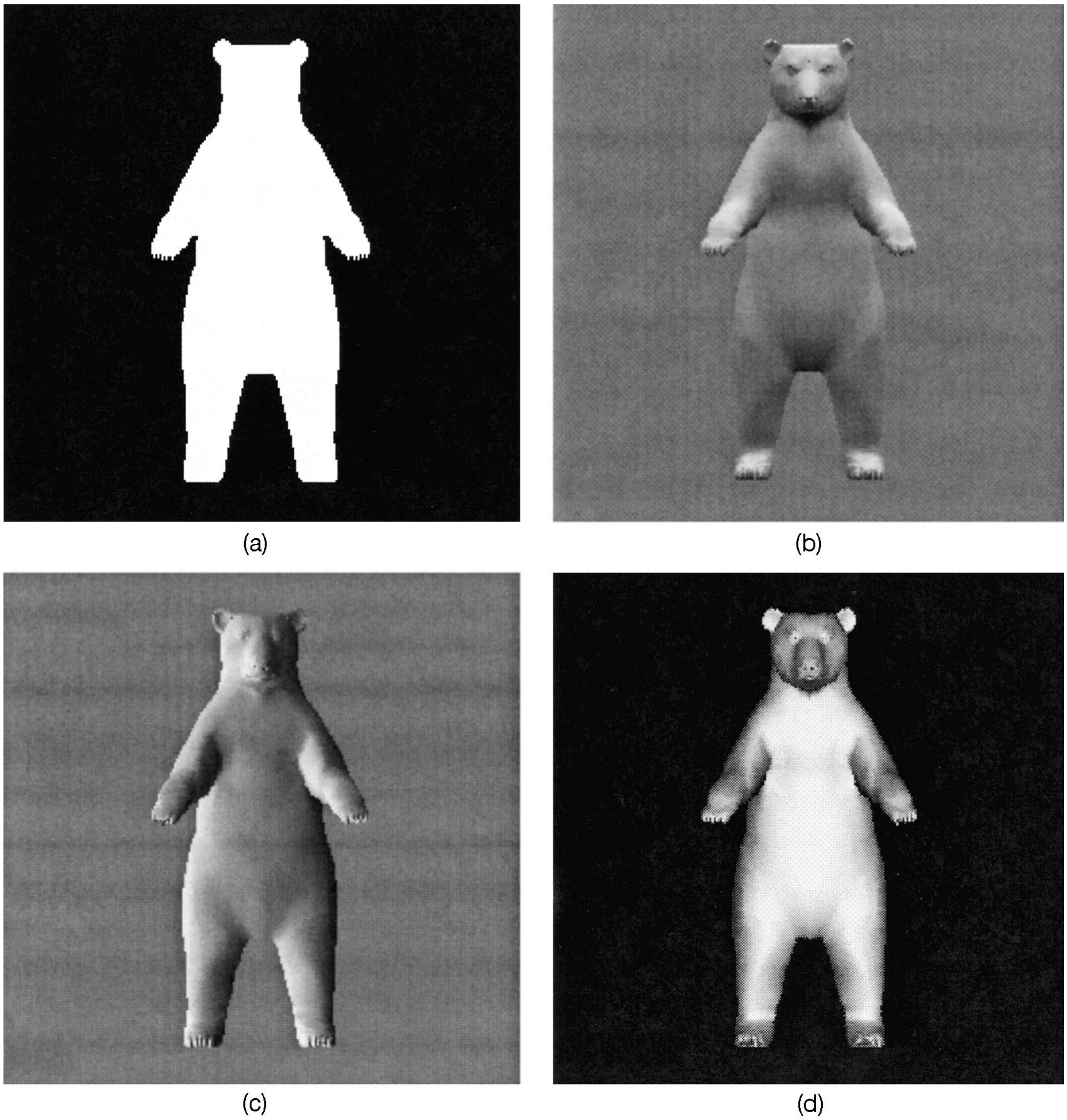
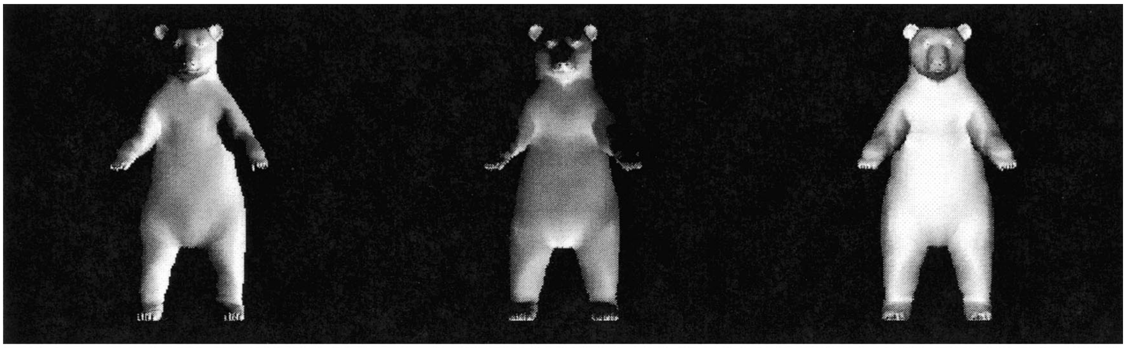


Fig. 4. Vector basis $\{k_a(x, y),[N_i(x, y)]\}$. (a) is $k_a(x, y)$ equal to ν_0 (x, y), and (b), (c), and (d) are $\{N_i(x, y)\}_i=X,Y,Z$, respectively.

From the point of view of vector spaces, intensity invariant pattern recognition consists of recognizing vectors independently of their length, which can be viewed as an angle measurement between vectors in vector spaces. This angle provides a measure of the similarity between the object and the reference function.

For 2D images, Dickey and Romero[28] defined a normalized correlation as a normalization of the inner product represented by the correlation integral. This method yields correlation peak values that are invariant under a multiplicative factor. Arsenault and Lefebvre[29] used a homomorphic transformation to change a multiplicative-intensity problem into an additive-intensity problem that can be addressed with the synthetic discriminant filter mentioned above. Lefebvre *et al.*[24] defined a nonlinear filtering method LACIF, which is invariant under any linear intensity transformation. This LACIF operation uses three correlations involving local statistics and nonlinearities. It was applied directly to scenes containing unsegmented targets. One of the advantages of the LACIF method is that no *a priori* information about the constant values involved in the linear illumination model is assumed. The LACIF method can be combined with synthetic discrimination filters to achieve both illumination invariance and out-of-plane rotation invariance.[24] Recently, in Ref. 25 the authors generalized the LACIF filtering for situations where a linear intensity gradient across an object is present.

From our illumination model defined in Eq. (4), any 3D image can be expressed as a linear combination of four images ($k_a(x, y)$, $\{N_i(x, y)\}_{i=x,y,z}$). In this paper, we define a new method based on a vector space representation in order to detect 3D objects under different lighting conditions. Our motivation is to develop an algorithm to recognize images that can be defined as a linear combination of other elementary images. That linear combination is mathematically defined as



(a) (b) (c)

Fig. 5. Three versions of the target with different illuminations.

$$f(x,\ y) = c_0 v_0(x,\ y) + c_1 v_1(x,\ y) + c_2 v_2(x,\ y) + \cdots, \tag{5}$$

where $f(x, y)$ is a certain image and $\{c_i\}_{i=0,1,\ldots}$ are arbitrary coefficients. For convenience, we assume that $v_0(x, y)$ is the silhouette of the image defined as

$$v_0(x,\ y) = \begin{cases} 1 & \text{if } f(x,\ y) \neq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

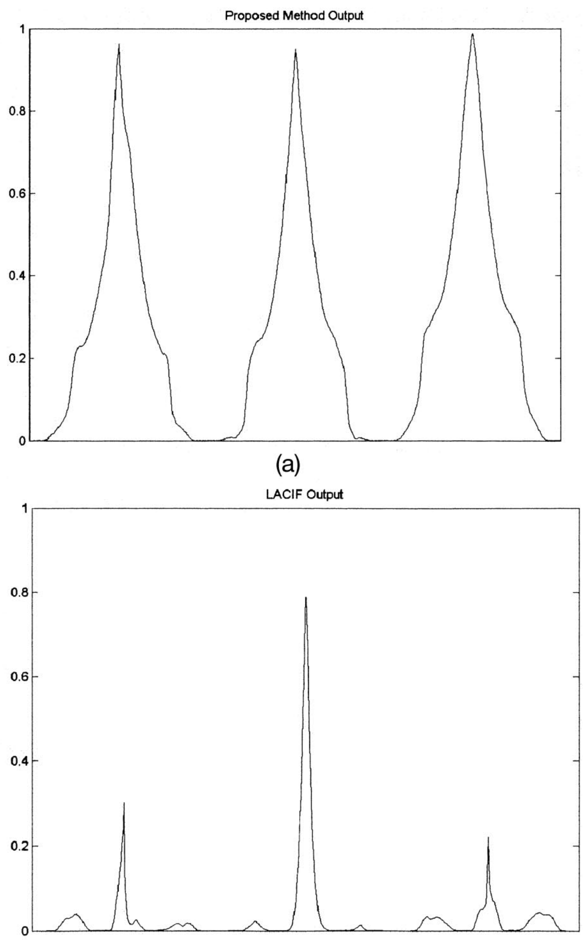


Fig. 6. (a) Correlation peak profile for Fig. 5 using the proposed method. (b) Correlation peak profile for Fig. 5 using the LACIF.

Moreover, from Eq. (5), $\{v_i(x, y)\}_{i=1,2,\ldots}$ are 2D images defined inside $v_0(x, y)$. From the point of view of vector spaces, $f(x, y)$ is a vector ($|f\rangle$) of a vector subspace generated by $\{|v_i\rangle\} = \{v_i(x, y)\}$. In addition, the inner product between 2D images, $a(x, y)$ and $b(x, y)$, is defined as

$$\langle a|b\rangle \equiv (a * b)(0,\ 0) = \int_{\mathbb{R}^2} a^*(x,\ y) b(x,\ y) \mathrm{d}x\mathrm{d}y, \tag{7}$$

where $*$ represents correlation. Note that we use as the inner product the value of the correlation between two functions at the origin. In the previous equations, we have supposed that images are continuous functions defined in $\mathbb{R}^2$. Because of that, the canonical basis of the vector space is represented by $\{\delta(x - x', y - y'), \forall x', y' \in \mathbb{R}^2\}$. Although the basis has infinite dimensions, for computing the results we have used discrete images of $N_1 \times N_2$ pixels. For a certain $N_1 = N_2 = 256$ pixel image size, the theoretical dimension of our space is $256 \times 256 = 65{,}536$, which is difficult to handle from a practical point of view unless strong restrictions apply. However, from Eq. (5), we just operate with a vector subspace whose dimensions are defined by the number of images $\{v_i(x, y)\}$, so the complexity of the dimensionality problem has been reduced from infinity to a small finite number.

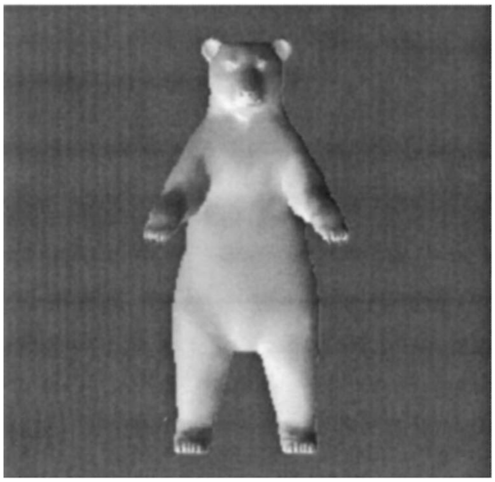
The basis defined by $\{v_i(x, y)\}$ is neither orthogonal nor orthonormal. We apply the Gram–Schmidt[30] orthonormalization method to the basis. Then the reference image, $f(x, y)$, can be defined as

$$f(x,\ y) = d_0 \cdot \hat{v}_0(x,\ y) + d_1 \cdot \hat{v}_1(x,\ y) + d_2 \cdot \hat{v}_2(x,\ y) \cdots, \tag{8}$$

where $\{\hat{v}_i(x, y)\}$ defines the orthonormal basis of the subspace, and $\{d_i\}$ are the components for the reference image. In vector space notation, the components can be defined as $\langle v_i|f\rangle$.

After those explanations, we can reformulate our original detection goal. The recognition method must

(a)



(b)

(c)

Fig. 7. (a) A 3D target. (b) Representation of (a) in the basis defined by $\{k_\alpha(x, y), [N_i(x, y)]\}$ with no consideration of shadow effects. (c) Subtraction of (a) from (b).

distinguish between images that belong or not to a certain subspace or to a certain family of images defined in Eq. (5) or Eq. (8). For 2D images, Lefebvre *et al.*[24] defined the LACIF operation as an operation able to perform this classification.

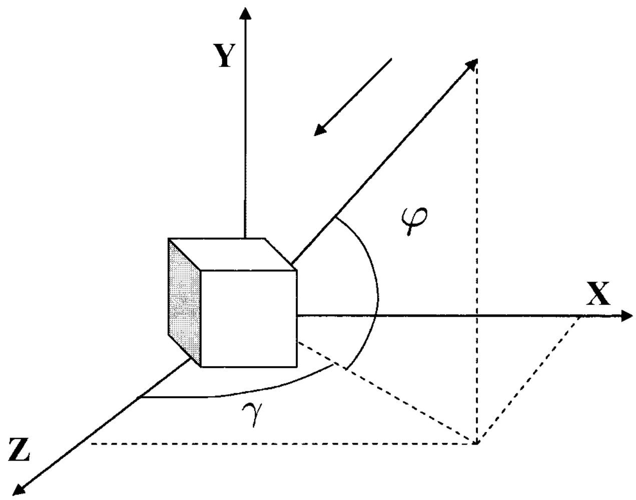


Fig. 8. Illumination variation in terms of the position of the point source.

We now review the LACIF filtering and its extension to 3D object recognition. We shall also define a novel algorithm based on a local angular distance measurement.

A. Locally Adaptive Contrast Invariant Filter

This filtering operation was used to detect images independently of intensity changes. For an arbitrary target $(g(x, y))$ and a reference object $(f(x, y))$, the LACIF method is defined in terms of the correlation as

$$C_{\mathrm{LACIF}}(g, f; x, y) = \frac{(f_0 * g)^2(x, y)}{\sqrt{N}(\hat{v}_0 * g^2)(x, y) - (\hat{v}_0 * g)^2(x, y)}, \tag{9}$$

where N is the number of pixels in the region of support, $\hat{v}_0(x, y)$ is the normalized version of $v_0(x, y)$, $f_0(x, y)$ is the zero-mean target defined as $f_0(x, y) = f(x, y) - \mu_f v_0(x, y)$, and μ_f is the mean of $f(x, y)$.

If the intensity transformation of the target is $af(x, y) + b$, where a and b are unknown parameters that are constant over the size of a single target, the LACIF output is equal to that obtained by the reference function, $f(x, y)$, i.e., the output is intensity invariant. In other words, if $g(x, y)$ is a linear combination of the orthonormal basis, $\{\hat{v}_i(x, y)\}$, the LACIF peak will be equal to 1, and be smaller than 1 if it is not. In terms of a vector interpretation, the LACIF technique is equivalent to projecting the target $(g(x, y))$ onto the subspace orthogonal to $v_0(x, y)$; that is, to project the zero-mean target onto this region of support, and then to calculating the cosine of the angle α between the target and the reference $(f(x, y))$.[24] The vector interpretation of the LACIF filter is illustrated in Fig. 2.

The LACIF method is used to determine whether an object belongs or not to the vector subspace gen-

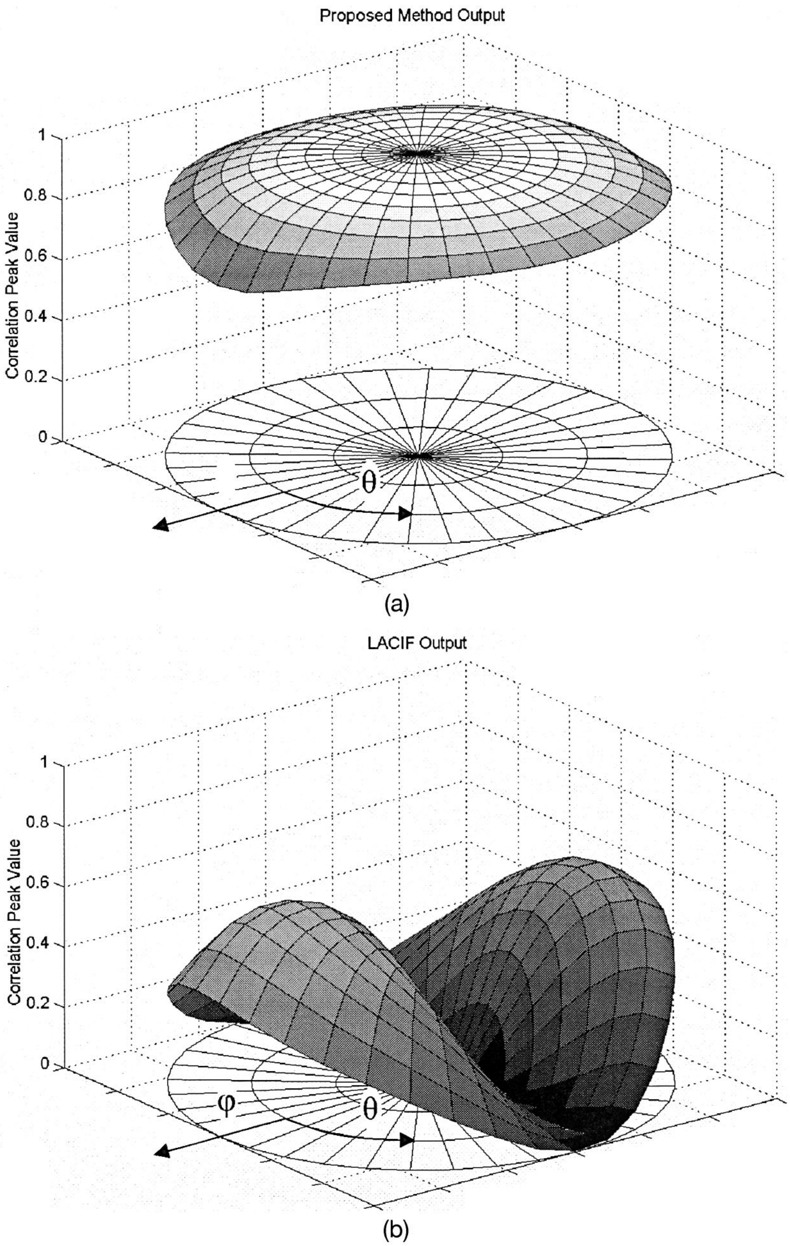


Fig. 9. Results for all illumination sampling for (a) the proposed LADC method and (b) the LACIF method.

erated by the vector basis $\{v_0(x, y), f(x, y)\}$. In addition, the LACIF can be generalized for the case of having more that two elements of the basis. Taking into account $\{\hat{v}_i(x, y)\}_{i=0,1,\ldots}$ basis, LACIF can be defined as

$$C_{\mathrm{LACIF}}(g,\ \hat{v}_i;\ x,\ y) = \frac{(\hat{v}_1 * g)^2(x,\ y)}{\sqrt{N}(\hat{v}_0 * g^2)(x,\ y) - \sum_{i \neq 1} (\hat{v}_i * g)^2(x,\ y)}. \quad (10)$$

Lefebvre *et al.*[25] applied Eq. (10) to affine transformations of intensity for 2D images.

According to our 3D illumination model, any 3D image can be expressed as a linear combination of four 2D images. Because of that, we propose the use of the LACIF to detect 3D illuminated objects. The detection method will be based on knowing if a certain 3D illuminated object is similar or not to a 3D reference object. The reference object is defined in terms of a certain basis given by the normally illuminated images $\{N_i(x, y)\}$. Because the linear combination implies multiplication by constant factors, the 3D recognition result will be invariant for darker or lighter objects.

B. Proposed Method: Local Angular Distance Correlation

We now propose a method to detect 3D objects based on the vector subspace idea. Instead of using the LACIF, we propose a simpler and more robust method based on calculating the angle between a given 3D target (e.g., a vector) and the whole subspace defined by a base. For the sake of clarity, consider the geometrical interpretation shown in Fig. 3. The vector space is generated only by two elements, $|\hat{v}_1\rangle$ and $|\hat{v}_2\rangle$. The vector $|g\rangle$ is the target. From Fig. 3, the angle, θ, between $|g\rangle$ and its projection, $|g_{\mathrm{proj}}\rangle$ can be defined as

$$\cos^2(\theta) = \frac{\|g_{\mathrm{proj}}\|^2}{\|g\|^2}, \tag{11}$$

where $\|g\|$ and $\|g_{\mathrm{proj}}\|$ are the moduli of $|g\rangle$ and $|g_{\mathrm{proj}}\rangle$. The projection of that vector onto the subspace defined by Eq. (8), can be expressed[31] as

$$|g_{\mathrm{proj}}\rangle \equiv \sum_i \langle \hat{v}_i | g\rangle |\hat{v}_i\rangle. \tag{12}$$

If vector $|g\rangle$ is contained in the subspace defined by Eq. (8), then $|g\rangle \equiv |g_{\mathrm{proj}}\rangle$. In that case, the angle between the vector and the subspace will be equal to zero, and so the cosine is equal to unity. But if the vector is orthogonal to that subspace, it means that no information about the object is contained in that subspace, so $|g_{\mathrm{proj}}\rangle = 0$ and the angle will be equal to $\pi/2$, and the cosine will vanish. Otherwise, the cosine value will be between zero and 1.

Note that Eq. (12) is a measurement of the similarity between two vectors, i.e., between the reference object (under varying illumination) and the target. However, the operation does not give the location of the target in a scene. It is not shift invariant because it is evaluated only at the origin. In order to extend the definition to the whole (x, y) space, the numerator of Eq. (11) is

$$\|g_{\mathrm{proj}}(x, y)\|^2 = \sum_i |\langle \hat{v}_i | g\rangle|^2 = \sum_i |\hat{v}_i * g|^2(x, y), \tag{13}$$

and the denominator of Eq. (11) can be written as

$$\|g(x, y)\|^2 = \int_{\mathbb{R}^2} v_0(x, y) g^2(x, y) = \sqrt{N}(\hat{v}_0 * g^2)(x, y), \tag{14}$$

where

$$\hat{v}_0(x, y) = \frac{v_0(x, y)}{\sqrt{\langle v_0 | v_0\rangle}} = \frac{v_0(x, y)}{\sqrt{N}}. \tag{15}$$

From Eq. (14), the energy of the target $(g^2(x, y))$ is evaluated inside the region of support for each point (x, y) of the image. This is the reason why our approach is a local angular distance measurement.

Taking into account previous equations, the final expression for our local angular distance correlation (LADC) is

$$\cos^2(\theta; g, \hat{v}_i; x, y) = \frac{\sum_i (\hat{v}_i * g)^2(x, y)}{\sqrt{N}(\hat{v}_0 * g^2)(x, y)}. \tag{16}$$

Equation (16) is defined in terms of correlations,

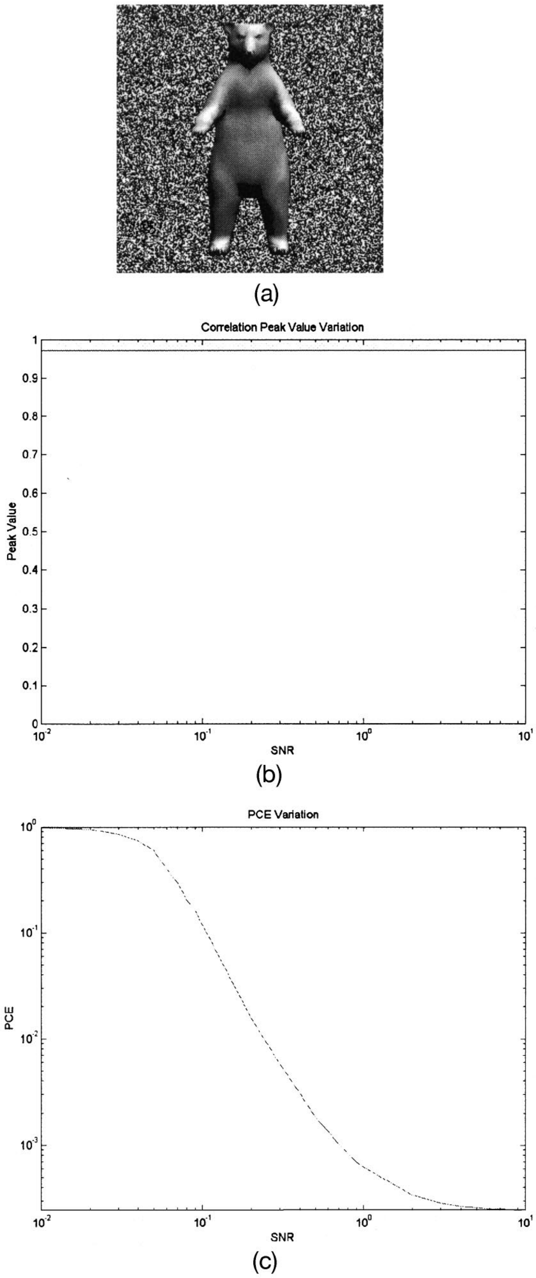


Fig. 10. (a) Sample noisy image (SNR = 5). (b) Correlation peak value variation. (c) PCE variation.

which can be an advantage if for optical implementations. Moreover, it gives a measure of the similarity between objects that belongs to a given vector subspace, and it localizes those targets in the scene because it is shift invariant. In contrast with the LACIF, which considers only the projection onto the silhouette, our method takes into account the projection onto all the subspace. This implies important differences for detecting images under different illuminations as we shall show in Section 4.

4. Illuminated Three-Dimensional Object Detection Results

In Section 2 we described our illumination model, which consists of expressing a 3D object as a linear combination of normal surface vector basis images [see Eq. (4)]. Changing the lighting conditions implies changing the constants that multiply the vector basis. Figure 4 shows the vector basis for the reference 3D object of Fig. 1(b). Figure 4(a) is the region of support and Figs. 4(b), 4(c), and 4(d) are $\{N_i(x, y)\}_{i=X,Y,Z}$, respectively.

We now compare the results of 3D detection using the LACIF and the new correlation based on angular distance. Figure 5 shows three instances of the target with changes in illumination. The corresponding correlation peaks for the proposed method, and for the LACIF, are shown in Figs. 6(a) and 6(b), respectively. Note that the LACIF procedure yields missed detections for two of the targets (the output peaks are very low). This means that the operation is not intensity invariant for this case. On the other hand, the LADC output value is approximately the same for all illumination views. To illustrate the reason for those differences between the LACIF and the LADC, we must explain the choice of the vector basis. Because we are not taking into account shadow effects in our vector basis definition, if a certain illuminated 3D target is affected by shadows, the vector basis will not completely define the object, so the basis will not be complete. This is illustrated in Fig. 7, where we see differences between a certain illuminated target [Fig. 7(a)] and the projection of the target in the four-vector basis [Fig. 7(b)]. Figure 7(c) shows the result of subtracting Fig. 7(a) from Fig. 7(b). The difference is not zero because the basis is not complete. For this reason, a target cannot be expressed exactly in terms of a linear combination of our four-vector space basis, and the LACIF will fail in the detection. Because our LADC is a measurement of the distance between the target and the four-vector subspace, and because the targets are quite similar, the distance will be small and the correlation peak value will be close to 1. As a result, our LADC is more robust to changes in illumination than the LACIF.

To show the intensity invariance of the LADC method, we represent the correlation output for all possible illuminations as an angular map representation. The changes of illumination are generated by changing the position of the point source that illuminates the 3D object. Figure 8 shows the two angles that will define a specific position of the point source that illuminates the object.

The comparison between the LADC and LACIF methods are shown in Fig. 9. Note that the LADC values of Fig. 9(a) are quite stable around the value 1, while the LACIF values of Fig. 9(b) fails for most of the illuminations.

The next step is to study the LADC for 3D images in the presence of noise. We studied the robustness of the correlation peak for the LADC detection of targets under Gaussian disjoint noise. Figure 10(a) is a target corrupted with nonoverlapping noise. Figure 10(b) shows the variation of the correlation peak as the noise increases. A low signal-to-noise (SNR) ratio value implies a highly noisy image. Figure 10(b) shows the stability of the correlation peak, which means that the LADC method is not affected by this kind of noise. Another measure of the performance of the correlation peak is given by the peak to correlation energy (PCE). Figure 10(c) shows the values of the PCE as the noise is increased. The PCE measures the sharpness of the correlation peak. A high value of PCE means a sharp correlation peak. From Fig. 10(c) we see that instead of decreasing the correlation peak sharpness, the noise improves the results. The following paragraphs explain that result.

Consider an image

$$f'(x, y) \equiv f(x, y) + R(x, y), \tag{17}$$

where $f(x, y)$ is an arbitrary image, and $R(x, y)$ is a zero-mean Gaussian noise distribution with standard deviation σ that does not overlap the arbitrary image:

$$R(x, y) = \begin{cases} 0 & \text{if } \hat{v}_i\,(x, y) = 0 \\ \text{Gaussian noise elsewhere} \end{cases}. \tag{18}$$

The PCE value after applying our LADC method is

$$\mathrm{PCE}(\sigma) = \frac{\left[\cos^2(\theta; f', \hat{v}_i, 0, 0)\right]^2}{\int_{\mathbb{R}^2} \left[\cos^2(\theta; f', \hat{v}_i, x, y)\right]^2 \mathrm{d}x\mathrm{d}y}. \tag{19}$$

Because the noise is nonoverlapping, the denominator of Eq. (19) can be split into two independent terms:

$$\mathrm{PCE}(\sigma) = \frac{\left[\cos^2(\theta; f, \hat{v}_i, 0, 0)\right]^2}{\left[\cos^2(\theta; f, \hat{v}_i, 0, 0)\right]^2 + \int_{(x, y)\neq(0, 0)} \left[\cos^2(\theta; f', \hat{v}_i, x, y)\right]^2 \mathrm{d}x\mathrm{d}y}. \tag{20}$$

The second term of the denominator can be expressed as

$$\cos^2(\theta; f', \hat{v}_i, x, y) = \frac{\sum_i \left[\int_{\mathbb{R}^2} \hat{v}_i^*(x'-x, y'-y)[f(x, y) + R(x, y)]\mathrm{d}x'\mathrm{d}y' \right]^2}{\sqrt{N} \int_{\mathbb{R}^2} \hat{v}_0^*(x'-x, y'-y)\left[f^2(x', y') + f \cdot R(x', y') + R^2(x', y')\right]\mathrm{d}x'\mathrm{d}y'}. \tag{21}$$

Considering a zero-mean Gaussian noise, Eq. (21) can be rewritten as

$$\begin{aligned} \cos^2(\theta; f', \hat{v}_i, x, y) &= \frac{K}{\sqrt{N} \int_{\mathbb{R}^2} \hat{v}_0^*(x'-x, y'-y)\left[R^2(x', y')\right]\mathrm{d}x'\mathrm{d}y'} \\ &\propto K'\sigma^{-2}, \end{aligned} \tag{22}$$

where K and K' are constant terms.
Then,

$$\mathrm{PCE}(\sigma) = \frac{\left[\cos^2(\theta; f, \hat{v}_i, 0, 0)\right]^2}{\left[\cos^2(\theta; f, \hat{v}_i, 0, 0)\right]^2 + \mathbb{O}(\sigma^{-4})}. \tag{23}$$

If $\sigma \to \infty$, the PCE value will converge to 1.

Finally, in order to show the discrimination abilities of the new method, Fig. 11(a) shows an input image made of two reference objects with different illuminations and two false objects. The result for the LADC is in Fig. 11(b). The correlation peaks for the true targets are equal to 1, whereas for the false target the output is lower, which demonstrates good discrimination even in the presence of nonoverlapping Gaussian noise. Note that we have used computer simulated objects in which 3D mesh is obtained digitally. For the case of a real 3D object, one needs to register different illumination views of the object in order to perform the vector basis.

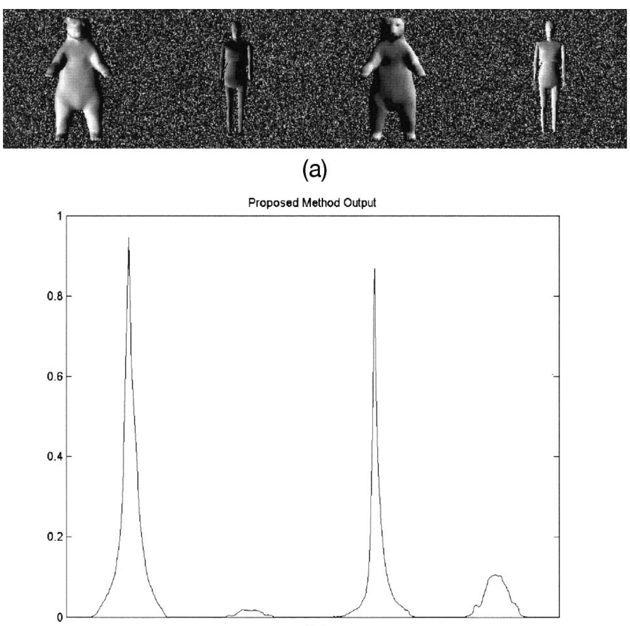


Fig. 11. (a) Input scene with true and false objects in a noisy environment (SNR = 1). (b) Output for the proposed method using (a).

5. Conclusion

We propose an illumination-invariant 3D recognition method based on vector spaces. The detection is done in terms of a local angular distance between a given target vector and a vector subspace defined by a 3D elementary image basis. We extended a previous intensity-invariant LACIF expression to 3D image recognition, and we have compared both methods. The results show the robustness of the new proposed method under changes of illumination. Experiments have validated the intensity invariant recognition of 3D images in the presence of nonoverlapping Gaussian noise. We have shown that the PCE value increases with the amount of noise in the image. We also successfully tested the method for composite images, when other false targets were in the scene.

This work was supported by FEDER funds and the Spanish "Ministerio de Educación y Ciencia" under the project FIS2004-06947-C02-01, the D.G Investigació i Transferència Tecnològica, project IIARC0/2004/217, and the Agencia Valenciana de Ciencia y Tecnología (AVCT), project GRUPOS03/117. J. J. Vallés acknowledges a FPU grant from the Ministry of Education and Science.

References

1. R. Campbell and P. Flynn, "A survey of free-form object representation and recognition techniques," Comput. Vis. Image Underst. **81,** 166–210 (2001).
2. M. Rioux, "Laser range finder based on synchronized scanners," Appl. Opt. **23,** 3837–3844 (1984).
3. M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3D 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).
6. E. Tajahuerce, O. Matoba, and B. Javidi, "Shift-invariant

three-dimensional object recognition by means of digital holography," Appl. Opt. **40,** 3877–3886 (2001).
7. 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).
8. P. Parrein, J. Taboury, and P. Chavel, "Evaluation of the shape conformity using correlation of range images," Opt. Commun. **195,** 393–397 (2001).
9. J. J. Esteve-Taboada, D. Mas, and J. Garcia, "Three-dimensional object recognition by Fourier transform profilometry," Appl. Opt. **38,** 4760–4765 (1999).
10. E. Paquet, P. García-Martínez, and J. García, "Tridimensional invariant correlation based on phase-coded and sine-coded range images," J. Opt. **29,** 35–39 (1998).
11. J. J. Esteve-Taboada and J. García, "Detection and orientation evaluation for three-dimensional objects," Opt. Commun. **217,** 123–131 (2003).
12. B. Javidi and E. Tajahuerce, "Three-dimensional object recognition by use of digital holography," Opt. Lett. **25,** 610–612 (2000).
13. R. Bamler and J. Hofer-Alfeis, "Three- and four-dimensional filter operations by coherent optics," Opt. Acta **29,** 747–757 (1982).
14. Y. Li and J. Rosen, "Object recognition using three-dimensional optical quasi-correlation," J. Opt. Soc. Am. A **19,** 1755–1762 (2002).
15. O. Matoba, E. Tajahuerce, and B. Javidi, "Real-time three-dimensional object recognition with multiple perspectives imaging," Appl. Opt. **40,** 3318–3325 (2001).
16. J. J. Esteve-Taboada, J. García, and C. Ferreira, "Rotations-invariant optical recognition of three-dimensional objects," Appl. Opt. **39,** 5998–6005 (2000).
17. J. J. Esteve-Taboada and J. García, "Detection and orientation evaluations for three-dimensional objects," Opt. Commun. **217,** 123–131 (2002).
18. J. Garcia, J. J. Valles, and C. Ferreira, "Detection of three-dimensional objects under arbitrary rotations base don range images," Opt. Express **11,** 3352–3358 (2003).
19. S. Yeom and B. Javidi, "Three-dimensional distortion-tolerant object recognition using integral imaging," Opt. Express **12,** 5795–5809 (2004).
20. J. J. Esteve-Taboada, N. Palmer, J.-C. Giannessini, J. Garcia, and C. Ferreira, "Recognition of polychromatic three-dimensional objects," Appl. Opt. **43,** 433–441 (2004).
21. B. K. Hom, *Robot Vision* (MIT Press, 1986).
22. Y. Tian, H. T. Tsui, S. Y. Yeung, and S. Ma, "Shape from shading for multiple light sources," J. Opt. Soc. Am. A. **16,** 36–52 (1999).
23. S. K. Zhou and R. Chellappa, "Image-based face recognition under illumination and pose variations," J. Opt. Soc. Am. A. **22,** 217–229 (2005).
24. D. Lefebvre, H. H. Arsenault, P. Garcia-Martinez, and C. Ferreira, "Recognition of unsegmented targets invariant under transformations of intensity," Appl. Opt. **41,** 6135–6142 (2002).
25. D. Lefebvre, H. H. Arsenault, and S. Roy, "Nonlinear filter for pattern recognition invariant to illumination and to out-of-plane rotations," Appl. Opt. **42,** 4658–4662 (2003).
26. S. Roy, D. Lefebvre, and H. H. Arsenault, "Recognition invariant under unknown affine transformations of intensity," Opt. Commun. **238,** 69–77 (2004).
27. J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics, Principle and Practics,* 2nd ed. (Addison-Wesley, 1996).
28. F. M. Dickey and L. A. Romero, "Normalized correlation for pattern recognition," Opt. Lett. **16,** 1186–1188 (1991).
29. H. H. Arsenault and D. Lefebvre, "Homomorphic cameo filter for pattern recognition that is invariant with change of illumination," Opt. Lett. **25,** 1567–1569 (2000).
30. http://mathworld.wolfram.com/Gram-SchmidtOrthonormalization.html.
31. http://mathworld.wolfram.com/VectorSpaceProjection.html.