Tridimensional invariant correlation based on phase-coded and sine-coded range images

Eric Paquet†, Pasquala García-Martínez‡ and Javier García‡

† Visual Information Technology, Institute for Information Technology, Building M-50, Montreal Road. Ottawa, Ontario, K1A 0R6 Canada
‡ Departament Interuniversitari d'Òptica, Universitat de València, C/Dr Moliner, 50, 46100 Burjassot, Spain

Received 1 September 1997, accepted 25 November 1997

Abstract. New methods for recognition of range images based on correlation are presented. The techniques allow the detection of a range image of a tridimensional pattern whatever its position along the x, y and z axes. The first method uses the phase-coding of the range image in order to obtain invariance along the z-axis in the correlation process. The second method uses the sine-coding of the range image in order to obtain better discrimination than that obtained with phase-coding. The third method is a hybrid one that makes use of sine-coding for the input scene and phase-coding for the reference image, so both the full translation invariance and high discrimination abilities are obtained. Computer simulations are presented. Also, an experimental implementation of the method has been carried out showing good agreement with theoretical predictions.

Keywords: Pattern recognition, image coding, range images

Corrélation invariante à trois dimensions basée sur des images en distance codées en phase et sinusoidalement

Résumé. On présente de nouvelles méthodes pour reconnaître des images en distance basées sur la corrélation. Ces techniques permettent la détection d'une image en distance d'une structure tridimensionnelle quelque soit sa position le long de axes x, y et z. La première méthode utilise le codage de phase d'une telle image pour obtenir l'invariance selon l'axe z dans le processus de corrélation. La seconde méthode utilise un codage sinusoïdal de l'image pour obtenir une meilleure discrimination qu'avec la méthode précédente. La troisième est une méthode hybride qui utilise le codage sinusoïdal pour la scène d'entée et le codage de phase par l'image de référence, ainsi on obtient à la fois l'invariance par translation et de fortes possibilités de discrimination. On montre des simulations numériques, et aussi un montage expérimental traduisant cette méthode et indiquant un bon accord avec les prévisions théoriques.

Mots clés: Reconnaissance de forme, codage d'images, images en distance

1. Introduction

Much work has been devoted to the study of bidimensional image recognition based on correlation methods [1–3]. Correlation is shift invariant if the translation is performed in the xy plane [2]. But it suffers from two problems. Firstly, if there is a translation along the z-axis, in most practical cases the apparent scale of the image will change: therefore bidimensional image correlation is not invariant under a translation along the z-axis [2]. The second problem comes from the nature of the information present in the bidimensional image. In most cases, a bidimensional image is an intensity image. This means that if the surrounding lighting changes, the intensity of the image will change as well. The only information that does not change with the illumination is the contour of the object.

Practically speaking this is, by far, the most interesting information in a bidimensional image and this is why the methods that use only the contours of the objects are so effective [4]

With recent developments in active range cameras [5] it is possible to obtain tridimensional images. The main advantage of this type of image is that it contains only geometrical information. Thus, the image remains the same whatever the surrounding lighting. This geometrical information can be used to increase the discrimination in the correlation process.

Several methods can be used for tridimensional image recognition. Some of them are based on some kind of segmentation, others on the normals or on the curvatures [6]. A common problem for these methods is the need to analyse each object in the scene separately. In contrast, methods based on bidimensional correlations can deal with an entire image at once in real-time. In this paper we propose to codify the geometrical pattern recognition scheme to obtain translation detection invariance along the x, y and z axes.

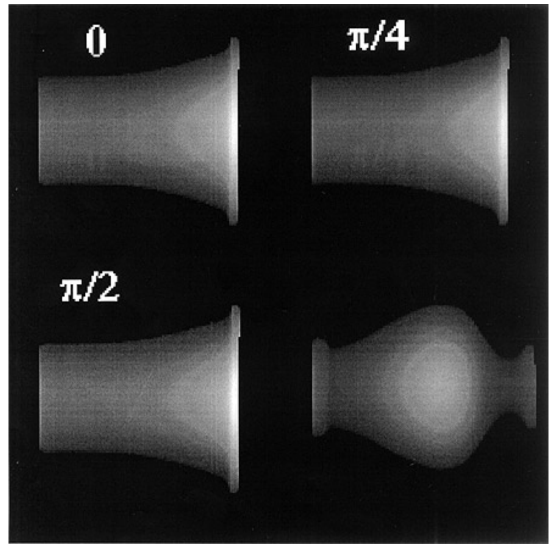

Figure 1. Input scene representing vase 1 at three different positions on the z-axis, given by a shift of $z_0 = 0$, $z_0 = \pi/4$ and $z_0 = \pi/2$ and vase 2.

2. Codification of range images

Let us consider a tridimensional object. From this object, two types of images may be obtained: a classical intensity image (projection through a point as in a standard camera) and a range image obtained from a range camera, which scans a laser beam over the object, obtaining the 3D spatial coordinates of every point in the valid range of the camera [5]. The intensity image is defined as $I(x, y)$ and the range image as $z(x, y)$. Each point (x, y) has a corresponding intensity I and range z. Since $I(x, y)$ contains both photometric and geometrical structure information, the range image contains only geometrical information on the object. If the lighting changes, $I(x, y)$ will change but $z(x, y)$ does not depend on the surrounding lighting.

If there is a translation along the z-axis, $I(x, y)$ and $z(x, y)$ will not behave in the same way. In the intensity image, the scale of the object changes. The change of scale is only apparent because the original object always maintains the same geometrical scale (if the object is rigid). Since $z(x, y)$ represents the geometrical structure of the object, the only effect of a translation along the z-axis is to add a bias to the corresponding image of the same amount [6], i.e. the origin of the range variable will change but the shape of the object will remain the same:

$$T_{z_0} z(x, y) = z(x, y) + z_0. \qquad (1)$$

A correlation procedure to recognize a tridimensional object is considered. This method should fulfil three requirements: it should be invariant under a translation in the xy plane, invariant under translation along the z-axis and, finally the discrimination capability of the system should be good. The correlation of an input scene (z_s) and a reference image (z_r) is defined in terms of the Fourier transform as

$$z_s * z_r = \mathcal{F}^{-1}\{\mathcal{F}[z_s]\,\mathcal{F}^*[z_r]\} \qquad (2)$$

where $\mathcal{F}$ is the Fourier transform and $*$ denotes correlation. For every instance of the reference object in the input scene, a correlation peak is obtained at the location of the target. However, if the input and the reference objects do not have the same position on the z-axis, equation (2) will not work properly because the reference and the input object will not be the same. In the following, a way to overcome this problem by codifying the tridimensional information is proposed. Two different codifications are used for the range: phase and sine codification.

2.1. Phase-coding

Phase-coding consists of codifying the range $z(x, y)$ as a phase as follows:

$$P(z(x, y)) = \exp[\mathrm{i}kz(x, y)] \qquad (3)$$

where k is a constant factor [6]. The influence of this will be explained in the following subsection. We consider that the input scene (z_s) is the target (z_r) translated an amount z_0 on the z-axis. Then, the result of the correlation of a z_s and z_r using the phase-coding method is

$$\begin{aligned} P(z_s(x, y)) * P(z_r(x, y)) \\ &= P(T_{z_0} z_r(x, y)) * P(z_r(x, y)) \\ &= \exp[\mathrm{i}kz_0][P(z_r(x, y)) * P(z_r(x, y))]. \end{aligned} \qquad (4)$$

If the correlation image is registered by a conventional intensity detector, the phase factor does not influence the result and an autocorrelation peak is obtained. So, the output is fully translation invariant. Let us consider the scene shown in figure 1†, which represents a vase (vase 1) at different z locations ($z_0 = 0$, $z_0 = \pi/4$ and $z_0 = \pi/2$) and another different vase to be rejected (vase 2).

† The 3D images are from M Rioux of the Institute for Information Technology, National Research Council of Canada.

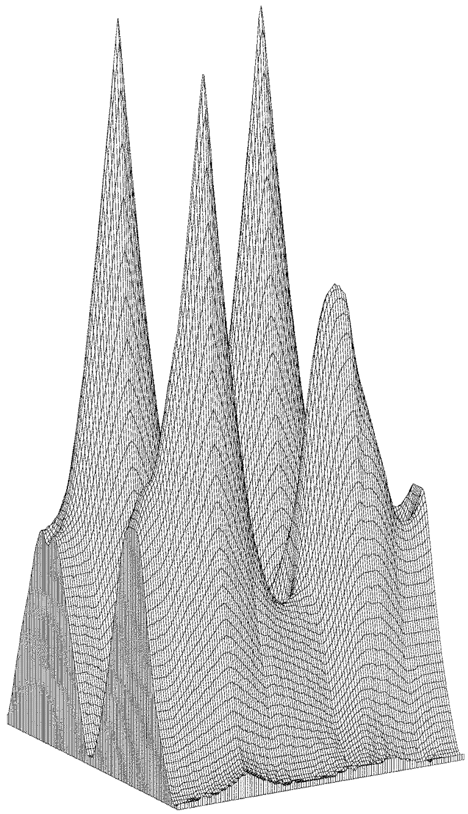

Figure 2. Computer simulations of figure 1 using the phase codification correlation method.

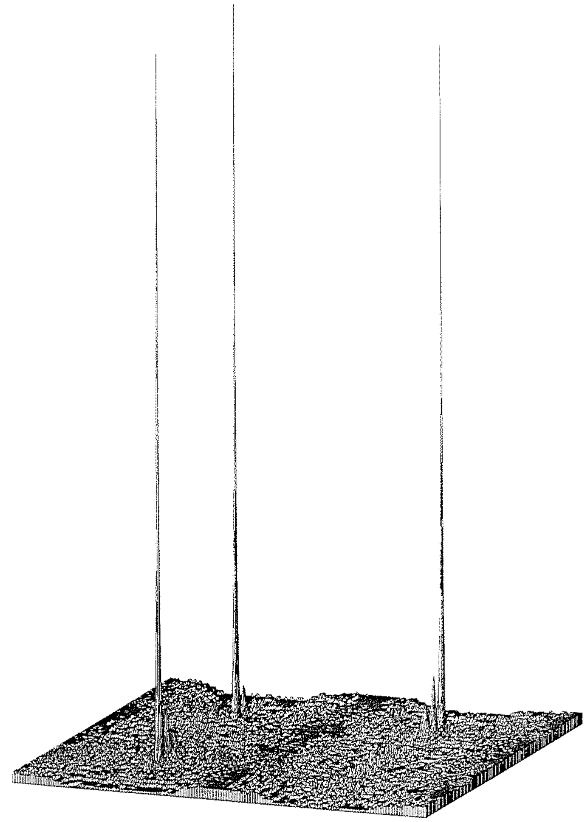

Figure 3. Computer simulations of figure 1 using the hybrid codification correlation method.

The correlation using phase-coding method is shown in figure 2. The threshold detection is 60%. The k value in all the numerical simulations has been $k = 1$. As can be seen in that figure, full translation invariance is obtained. Nevertheless, the discrimination abilities are not as quite good as one would desire. Moreover, bearing in mind the possibility of optical implementation we must consider the difficulty of introducing phase images in optical systems. So, a real function coding could be more feasible. Also, high discrimination is desired. Because of this we propose another codification.

2.2. Sine-coding

Sine-coding consists in modulating the range with a sine function [7]. It is defined as

$$S(z(x, y)) = \sin[kz(x, y)]. \tag{5}$$

Unfortunately, this codification produces correlation output which is not invariant under translation along the z-axis. This can be seen from the correlation output at the origin:

$$\gamma(0, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sin[kz(x, y) + kz_0] \sin[kz(x, y)] \, \mathrm{d}x \, \mathrm{d}y. \tag{6}$$

Nevertheless, this codification can serve to explain the influence of the parameter k in the recognition process. The sine function is an oscillating function made up of positive and negative values. The parameter k is related to the discrimination and the tolerance of the process. If k is small, the sine function oscillates slowly producing a relatively smooth image so that discrimination is low and tolerance is high. In contrast, when k has a significant content at high frequencies the discrimination is high and the tolerance is low. Thus, k can be set to a specific value depending on the specific applications, providing a way to control the robustness of the systems against distortions.

In the next section it is shown that it is possible to obtain invariance under translation along the z-axis by combining the phase-coding and the sine-coding methods.

3. Hybrid coding of the range images

We propose a hybrid-coding method consisting of sine-coding for the input scene and phase-coding for the reference image. Thus, the correlation can be written as

$$\begin{aligned} \mathrm{HC}(z_\mathrm{s}, z_\mathrm{r}) &= S(z_\mathrm{s}) * P(z_\mathrm{r}) \\ &= \mathcal{F}^{-1}\{\mathcal{F}[\sin(kz_\mathrm{s})]\, \mathcal{F}^*[\exp(\mathrm{i}kz_\mathrm{r})]\} \end{aligned} \tag{7}$$

where HC is the hybrid correlation. Let us consider the correlation of two identical images: the second being at a different position on the z-axis:

$$\mathrm{HC}(z, z + z_0) = \exp[-\mathrm{i}kz_0]\mathcal{F}^{-1}[\mathcal{F}[\sin(kz)]\, \mathcal{F}^*[\exp[\mathrm{i}kz]]]. \tag{8}$$

The constant phase factor can be eliminated by computing the intensity spectrum of equation (8). It is normal if we detect the optical correlation output with a quadratic detector. From equation (8), the autocorrelation can be written as

$$\gamma(0,0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sin[kz(x,y)][\cos[kz(x,y)] + \mathrm{i}\sin[kz(x,y)]]\,\mathrm{d}x\,\mathrm{d}y. \tag{9}$$

Note that the integral in equation (9) reduces to the autocorrelation of two sine-coded images. In the first term of the right-hand side of equation (9) the sine-coded image and the real part of the phase image are in quadrature. Now, the main difference with sine-coding is that full translation invariance is obtained with the hybrid coding method.

Some computer simulations are performed to demonstrate the hybrid codification correlation method.

Let us consider the same scene as we have considered in the phase-coding method (see figure 1). The correlation using the hybrid method is shown in figure 3. The result shows a good discrimination capability, and a threshold of 4.3% is needed to reject vase 2. Note that the different z-shifted versions of vase 1 are detected. It can be concluded that hybrid modulation is, from the point of view of correlation, a sine-coding correlation and the introduction of phase-coding for the reference image provides translation invariance along the z-axis. Note that equation (8) has a periodicity of 2π, so any translation can be folded on this interval.

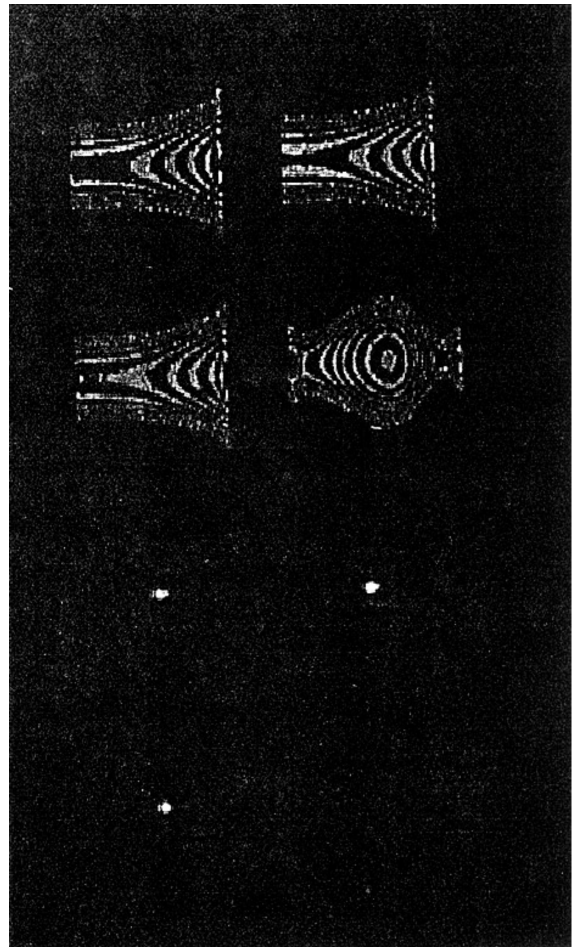

Figure 4. Optical correlation of figure 1 with vase 1 using binary modulation and a phase-only filter.

4. Optical implementation of the hybrid codification

In the following we consider the optical implementation with a classical 4-f correlator where a complex filter is introduced in the Fourier plane of the input transparency. The hybrid-modulation method presents some problems in this set-up: the input scene is made up of positive and negative amplitude values and the matched filter is complex (see equation (7)). The method described in the previous section has been modified to the usual technical requirements of an optical correlator.

For the input scene, binary modulation can be used. It is easy to demonstrate that this modulation also produces full translation invariance, the main drawback being an increasing level of noise and the appearance of a bias in the object. This non-discriminating bias will also contribute to the input. Let us consider the reference image. In order to obtain good discrimination, the contribution of the background in the phase function is eliminated, i.e. the intensity is made equal to zero every time there is zero at the corresponding position in the reference image.

Another problem comes from the matched filter. In an optical implementation, one can use a computer-generated hologram in order to introduce the filter in the Fourier plane. To improve the discrimination of the method, a phase-only filter can be used. Such a filter is obtained by extracting the phase for each point of the original complex filter [8]. When all these modifications are introduced the results are similar to those obtained with a regular hybrid coding. The translation invariance along the z-axis is also preserved in this case.

An optical implementation has also been performed. An optical correlator using a He–Ne laser and spherical beam illumination has been prepared. This allows the scale between the target and the filter to be matched. The input scene is binarized using the previously described binary modulation. The experimental correlation of figure 1 with vase 1 is shown in figure 4. As expected, the targets are detected independently of their position on the z-axis.

5. Conclusions

There are many advantages in using tridimensional images for correlation: one of the most significant being that the scale-invariance problem is transformed into a translation-invariance problem. Translation invariance in the xy plane is a problem that can be solved using correlation techniques. Modulation of the range image has been introduced in order to obtain translation invariance along the z-axis. The method presents a good correlation detection threshold. This comes in part from the modulations used which give rise to contour–contour correlation and the use of range images which carry useful information that is not present in bidimensional images. This tridimensional information leads to an improvement in the discrimination between images. Also, the optical implementation of this codification provides significant advantages in the whole recognition process.

Acknowledgments

EP would like to thank the National Research Council of Canada and especially Mr Marc Rioux for the range images and the Research Council in Natural Sciences and Engineering of Canada for funding this project. The work was also carried out with the financial support of CICYT Project TAP96-1015-C03-03 of Spain.

References

[1] Vander Lugt A 1964 Signal detection by complex spatial filtering *IEEE Trans. Inform. Theory* **IT-10** 139–45
[2] Feiltelson D G 1988 *Optical Computing: a Survey for Computer Scientists* (Cambridge, MA: MIT Press)
[3] Maze S and Refregier Ph 1994 Optical correlation: influence of the coding of the input image *Appl. Opt.* **33** 6788–96
[4] Chang S and Arsenault H H 1994 Invariant optical-pattern recognition based on a contour bank *Appl. Opt.* **33** 3076–85
[5] Rioux M, Beraldin J A, O'Sullivan M and Cournoyer L 1991 Eye-safe laser scanner for range imaging *Appl. Opt.* **30** 2219–23
[6] Paquet E, Rioux M and Arsenault H H 1994 Range image segmentation using the Fourier transform *Opt. Eng.* **32** 2173–80
[7] Rioux M, Boulanger P and Kasvand T 1987 Segmentation of range images using sine wave coding and Fourier transformation *Appl. Opt.* **26** 287–93
[8] Campos J, Turon F, Yzuel M J and Vallmitjana S 1990 Phase-only filter codified with Buckhardt's method *Appl. Opt.* **29** 5232–4