Optics Communications 283 (2010) 3613–3630

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Digital system of invariant correlation to position and rotation

Selene Solorza [a], Josué Álvarez-Borrego [b,*]

[a] Facultad de Ciencias, UABC, Ensenada, B.C., México, Km. 103 Carretera Tijuana-Ensenada, Ensenada, B.C., México, C.P. 22860, Mexico
[b] División de Física Aplicada, Departamento de Óptica, CICESE, Ensenada, B.C., México, Carretera Ensenada-Tijuana No. 3918, Fraccionamiento Zona Playitas, Ensenada, B.C., México, C.P. 22860, Mexico

ARTICLE INFO

Article history:
Received 27 August 2009
Received in revised form 18 May 2010
Accepted 18 May 2010

Keywords:
Fourier optics
Pattern recognition
Image forming and processing
Computer vision
Robotic vision

ABSTRACT

A new correlation digital system invariant to position and rotation is presented. This new algorithm requires low computational cost, because it uses uni-dimensional signatures (vectors). The signature of the target so like the signature of the object to be recognized in the problem image is obtained using a binary ring mask constructed based on the real positive values of the Fourier transform of the corresponding image. In this manner, each image will have one unique binary ring mask, avoiding in this form the relevant information leak. Using linear and non-linear correlations, this methodology is applied first in the identification of the alphabet letters in Arial font style and then in the classification of fossil diatoms images. Also, this system is tested using the diatom images with additive Gaussian noise. The non-linear correlation results were excellent, obtaining in this way a simple but efficient method to achieve rotation and translation invariance pattern recognition.

1. Introduction

Since few decades ago, the design of new filters for pattern recognition based on correlation has attracted considerable attention [1–8]. The correlation digital system applications are several and different. For example, the optical diffraction pattern sampling was used to identify two alphabetic characters [9]. The digital systems of correlation invariant to position and rotation have been used to classify by sex the *Acartia* species [10], because the male and female size are different. Moreover, they have been used to detect the *Vibrio cholerae* 01 (cholera bacterium) and tuberculosis bacteria [11–15]. Also, the clasification of fossil diatoms are achieved by digital systems [16]. Then, digital systems invariant to position and rotation are a useful tool to identify micro- and macro-structures regardless the shift and orientation of the object.

There are many works about digital systems invariant to position and rotation but they have a big computational cost. The composite non-linear filters [17] like adaptive SDF filters [18] can be used for obtaining a digital invariant correlation to rotation, but they must have the information of all angles of rotation of the object to be recognized. In order to have a new correlation digital system invariant to position and rotation with low computational cost, in this work a new methodology based on uni-dimensional signatures of the images is presented.

An important aspect in the study of digital system is the information leak. Here, the frequencies are filtered using binary ring masks. Then, the objective is to neglect only the irrelevant information when the mask applies. Thus, a methodology to build a mask associated to the image is proposed. In this form, the information leak is avoided because the frequencies filtered are not the same always. Other relevant issue is the correlation process in the discrimination between objects, thus the complete analyses for the linear and non-linear correlations involved in the proposed system are presented.

The material of this work is organized as follows: in Section 2, the digital system invariance to position and rotation process is explained. Section 3 presents the results of the correlation method applied to the alphabetical letters in Arial font. In Section 4, the method is applied to identify different fossil diatoms. Section 5 shows the analysis of the diatoms images with additive Gaussian noise. Finally, in Section 6 the conclusions are given.

2. Position and rotational invariance correlation method

The invariant correlation digital system to position and rotation presented in this work uses a binary ring mask of the image chosen.

2.1. Binary ring mask

The mask of a selected image, named I, is built by taking the real part of its 2D-Fourier transform (FT), given by

$$f(x,y) = Re(FT(I(x,y))) \quad (1)$$

* Corresponding author. Tel.: +52 646 1750500; fax: +52 646 1750553.
E-mail addresses: selene.solorza@gmail.com (S. Solorza), josue@cicese.mx (J. Álvarez-Borrego).

0030-4018/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2010.05.035

where (x,y) represent a pixel of the image. For example, in Fig. 1b the real part of the Fourier transform (the f function in Eq. (1)) for a black and white image containing the contour of T letter in Arial font style (Fig. 1a) is shown.

Based on Eq. (1), we obtained the one-variable function,

$$Z(y) = \begin{cases} 1, & \text{if } f(c_x, y) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where (c_x,c_y) is the centred-pixel of I and y goes from zero to wide of I. In the example presented in Fig. 2 (from Fig. 1), $c_x = 129$ and $0 \le y \le 256$.

Fig. 2a presents the graph of f in Eq. (1) setting $x = 129$. It is a cross-section cut of the graph in Fig. 1b along the 129-axes in the y-direction. In Fig. 2b the graph of Z (Eq. (2)) is shown.

Next, taking the vertical axis as the rotation axis, the graph of Z is rotated 360° to obtain concentric cylinders of height one, different widths and centred in (c_x,c_y). Finally, mapping those cylinders in two dimensions we built the binary ring mask associated to image I. Fig. 3 presents the mask for the images of T and X Arial font letters.

2.2. The signature of the image

The objective is identifying a specific target (the object to recognize) no matter the angle of rotation presented on the vertical-axis. In order to have an invariance rotation we did the binary mask (step 1 in Fig. 4) beginning from the frequency content of the image (step 2 in Fig. 4). So, the mask is applied in the Fourier plane for sampling the frequencies pattern of the object (step 3 in Fig. 4). Finally, the modulus of the Fourier transform for each ring is summed and then assigned to the corresponding ring index to obtain the signature of the image (step 4 in Fig. 4). Because the signature of the image depends upon the number of rings in the mask, and the mask changes with the target, the length of the signature also changes (as shown in Fig. 4). In this way each signature is unique for each image. Contrary in [19,20] where the mask used is the same for all images, hence the number of rings in the mask is constant.

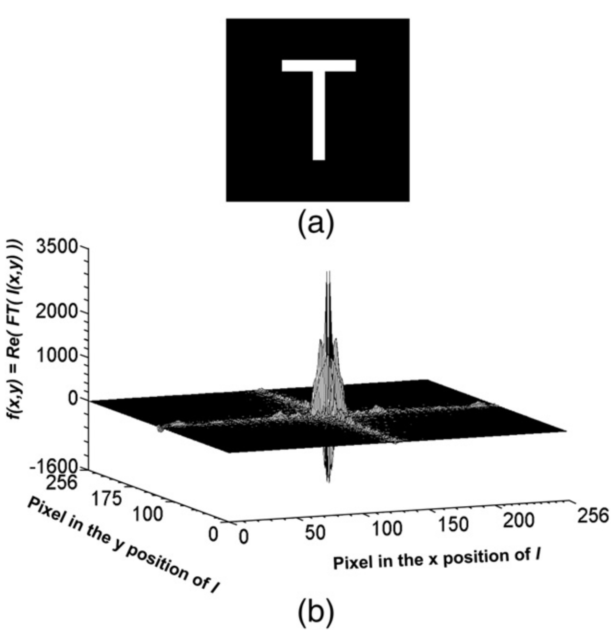

Fig. 1. (a) An image, called I, of 256×256 pixels with a centred Arial T letter. (b) The real part of the Fourier transform of the image I.

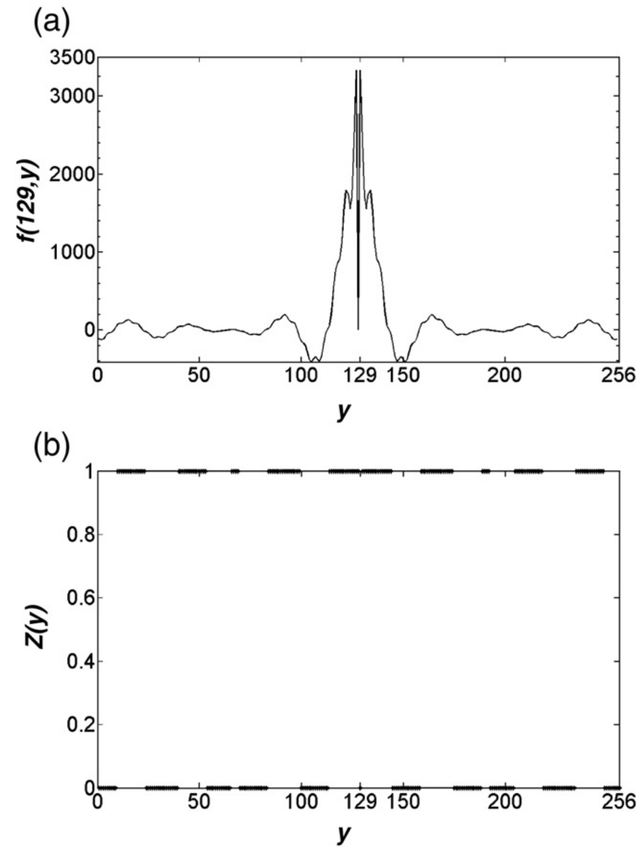

Fig. 2. (a) The graph of the real part of the Fourier transform of $I(129,y)$ obtained from Fig. 1b. (b) The graph of Z function associated to panel a.

2.3. The filter

For making the average signature filter of a given target, a complete rotation of 360° is performed to it with a rotation angle of $\Delta\theta$°; hence we generate $360/\Delta\theta$ images. Then, the signatures of those images are obtained to average them to build the average filter of the target. Fig. 5a and c shows, respectively, the average filters for the images with Arial font letters T and X (those shown in Fig. 3) with $\Delta\theta = 1$°, thus the filters are constructed using 360 images. In this methodology, we can choose the number of images in the filter, making this work more flexible, as

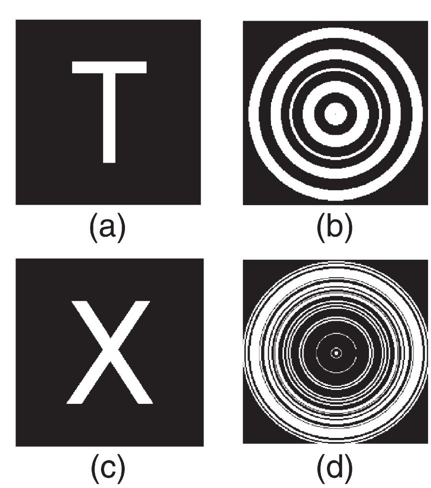

Fig. 3. Binary ring mask examples. (a) An image of 256×256 pixels with a centred Arial T letter. (b) The mask corresponding to T letter. (c) An image of 256×256 pixels with a centred Arial X font letter. (d) The mask corresponding to X font letter.