The Classification of Fragments of Objects by the Fourier Mask Digital System

Carolina Barajas-García[a], Selene Solorza[a], Josué Álvarez-Borrego[b]

[a]Universidad Autónoma de Baja California, Facultad de Ciencias.
km.103, Carretera Tijuana-Ensenada, Ensenada, B.C., México, C.P. 22860.
{carolinab49,selene.solorza}@gmail.com
[b]Centro de Investigación Científica y Eduación Superior de Ensenada, División de Física Aplicada.
Carretera Ensenada-Tijuana No.3918, Frac. Zona Playitas, Ensenada, B.C., México, C.P. 22860.
josue@cicese.mx

2014 Published by $\mathcal{DIFU}_{100}ci@$ http://nautilus.uaz.edu.mx/difu100cia

Abstract

The automation process of the pattern recognition for fragments of objects is a challenge to humanity. For humans it is relatively easy to classify a fragment of some object even if it is isolate and maybe this identification could be more complicated if it is partially overlapped by other object. However, the emulation of the functions of the human eye and brain by a computer is not a trivial issue. This paper presents a pattern recognition digital system based on Fourier binary rings mask to classify fragments of objects. The digital system is invariant to position and rotation, it is robust in the classification of images that have noise and non-homogenous illumination, moreover it classifies images that present an occlusion or elimination until the 15% of the area of the object.

Keywords: Pattern recognition, Image processing, Digital systems, Binary rings masks.

1. Introduction

One of the first tasks of a human being are the identification of the objects around him or her, therefore almost through their life, him or her will be dedicated to recognize patterns whether by necessity or just for curiosity. Nowadays, with almost all the production processes automated and the fast development of technology, the necessity to improve digital systems for pattern classification is increasing [1]. One of the challenge task in the pattern recognition field is to identify the pattern of the objects into images of the real world, because into those images could be that some of the objects are incomplete, that is, a fragment of the object is given only or either because the objects present occlusions or maybe they are not fully contained into the image. Then, the problem to recognize and to classify objects by a computer becomes a non-trivial issue. Because of this, the development of robust techniques to identify fragments of objects is necessary.

In this paper, is proposed a digital system based on a Fourier mask to classify fragments of objects in digital

images. This digital system is invariant to position and rotation, it is robust in classifying images with noise and non-homogenous illumination. Moreover, the system also classifies images that present an occlusion or elimination until the 15% of the area of the object. The paper is organized as follows: In section 2, the methodology to develop the digital system is presented. The classification framework is described in section 3. Finally, conclusions are shown in section 4.

2. The pattern recognition digital system

2.1. The Fourier mask

The mask of a selected gray scale image $I(x,y), x = 1,...,n, y = 1,...,n$ can be built by taking the imaginary part of its Fourier transform [2], that is, $Im(FT(I(x,y)))$; for example, the imaginary part of the Fourier transform of Figure 1(a) is shown in Figure 1(b). Next, the image $Im(FT(I(x,y)))$ is filtered by the binary disk mask $D(x,y)$,

$$D(x,y) = \begin{cases} 1, & \text{if } d((c_x,c_x),(x,y)) \leq n\,, \\ 0, & \text{otherwise}\,, \end{cases} \tag{1}$$

where (c_x,c_x) is the center-pixel of the image and $d(p,q)$ is the euclidean-distance between p and q points, thus the $D(x,y)$ image is centered in the (c_x,c_x)-pixel. Figure 1(c) presents a disk filter of diameter n and the $Im(FT(I(x,y)))$ is filtered by $D(x,y)$, as shown in Figure 1(d), mathematically that operation is given by

$$f(x,y) = D(x,y) \circ Im(FT(I(x,y)))\,, \tag{2}$$

where $\circ$ means an element-wise product or Hadamard product [3]. For the image $f(x,y)$, 180 profiles of n pixels length that passes for (c_x,c_x) where obtained. They are separated by $\Delta\theta = 1°$, sampling in this manner the entire disk. Figure 1(d) shows the zero-degree profile $P^0(x)$. In general, the profile equations are expressed as

$$P^\theta(x) = f(x,y(x))\,, \tag{3}$$

where $x = 1,...,n$, $y(x) = m(x - x_1) + y_1$, m is the slope of y, $(x_1,y_1) = (c_x + c_x cos\theta, c_x - c_x sin\theta)$ and $(x_2,y_2) = (c_x + c_x cos(\theta+\pi), c_x + c_x sin(\theta+\pi))$ are the two distinct end points of that line segment and θ is the angle that y has respect to the horizontal axis in the Cartesian plane (considering that the origin $(0,0)$ of the Cartesian plane are set at the center pixel of the image (c_x,c_x)).

Then, the addition of the square of the intensity values in each profile is computed, that is,

$$S^\theta = \sum_{x=1}^{n} (P^\theta(x))^2\,, \tag{4}$$

and the profile whose sum has the maximum value will be selected,

$$\alpha_\beta = max_{0\leq\theta\leq179}\{S^\theta\}\,, \qquad T(x) = P^\beta(x)\,, \tag{5}$$

where β is the angle of the profile in $f(x,y)$ whose sum has the maximum value, hence, these profile is called the maximum intensity profile. For example, in Figure 1(d) is indicated the maximum intensity profile $T(x)$ as a dashed-black line and Figure 1(e) shows the maximum intensity profile $T(x)$ in the Cartesian plane. Next, based on equation (5) the $Z(x)$ binary function (Figure 1(f)) is obtained,

$$Z(x) = \begin{cases} 1, & \text{if } T(x) > 0\,, \\ 0, & \text{if } T(x) \leq 0\,. \end{cases} \tag{6}$$

for $x = c_x,...,n$. Finally, taking the vertical axis $x = c_x$ as the rotation axis, the graph of $Z(x)$ is rotated $360°$ to obtain concentric cylinders of height one, different widths and centered in (c_x,c_x) pixel. Taking a cross-section of those concentric cylinders, the binary rings mask associated to the given image is built. In Figure 1(g) the binary rings mask $M(x,y)$ corresponding to the image in Figure 1(a) is shown.

2.2. The signature

The digital system uses the modulus of the Fourier transform of the image, $|FT(I(x,y))|$, because it is invariant to translation [4], that is, $|FT(I(x,y)) = FT(I(x+\tau,y+\zeta))|$ and $\tau,\zeta \in \mathbb{R}$, hence, the system is invariant to translation in an easy manner. To obtain the rotational invariance, signatures based on binary rings mask are build. The first step is to filter the modulus of the Fourier transform of the image by the binary rings mask. For example, the amplitude spectrum (Figure 2(b)) of Figure 2(a) is filtered by the binary rings mask $M(x,y)$ (Figure 2(c)) as

$$H(x,y) = M(x,y) \circ |FT(I(x,y)|\,. \tag{7}$$

The result of equation (7) is presented in Figure 2(d). After that, the rings in $H(x,y)$ are enumerated from inside to outside to obtain the following set,

$$Index = \{ring\ \ index : ring\ \ index\ \ x \in \bar{n}\}\,, \tag{8}$$

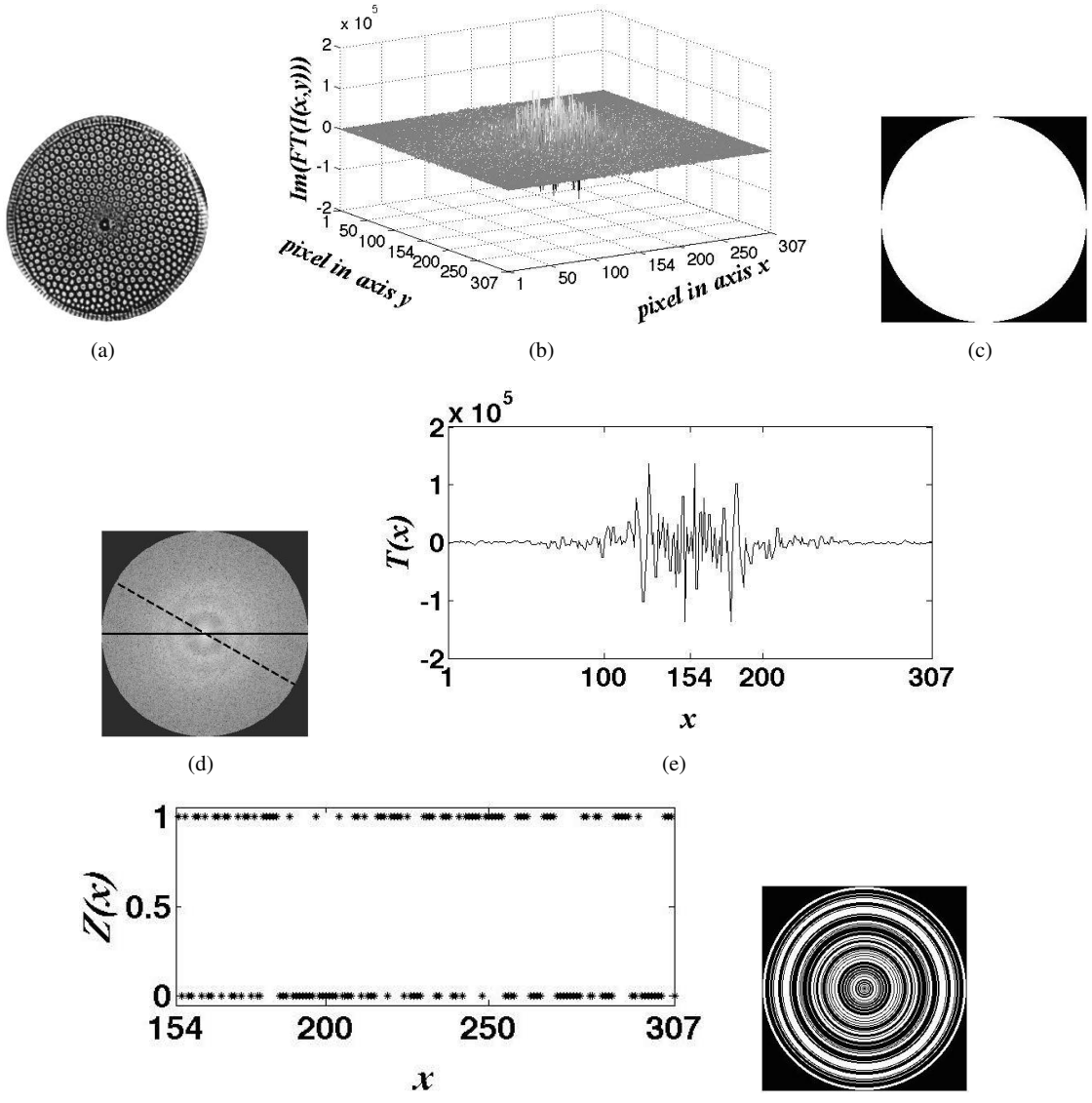

Figure 1. (a) Image $I(x,y)$. (b) Imaginary part of the Fourier transform of $I(x,y)$, that is $Im(FT(I(x,y)))$. (c) Binary disk $D(x,y)$. (d) $f(x,y) = D(x,y) \circ Im(FT(I(x,y)))$. The solid line shows the profile $P^0(x)$ and the dashed line the profile $T(x)$. (e) Graph of the maximum intensity profile $T(x)$. (f) Graph of the binary function $Z(x)$. (g) Binary mask $M(x)$.

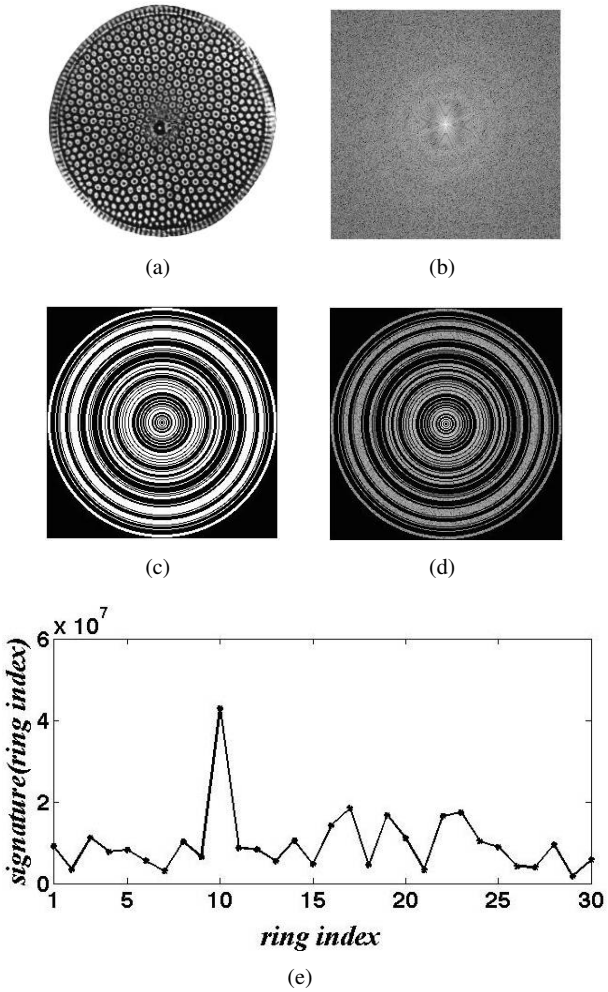

Figure 2. (a) Image $I(x,y)$. (b) $|FT(I(x,y))|$. (c) Binary mask $M(x,y)$. (d) $H(x,y) = M(x,y) \circ |FT(I(x,y))|$. (e) Signature of image $I(x,y)$.

where $\bar{n} = \{1,...,n\}$. The addition of the intensity values in each ring of $H(x,y)$ are computed to build the function

$$signature = index \rightarrow A \subset \mathbb{R}\,, \tag{9}$$

$$signature(ring\ index) = \sum H, \text{ if } H(x,y) \text{ are in the ring } ring\ index\,.$$

Figure 2(e) show the signature constructed by the binary rings mask $M(x,y)$.

3. Classification

To perform the pattern recognition, first of all it is set the signature of each image of the target images data base, $\beta_R = \left\{R_j \in M_{n\times n} :\ j = 1,\ldots,k;\ k \in \mathbb{N}\right\}$. Lets called S_{R_j} to those signatures. The pattern in the target image R_j is calculated by applying the Pearsons correlation coefficient [5]

$$r_{R_j} = \frac{\max\left\{C_L(S_{R_j})\right\}}{(N-1)\sigma^2_{S_{R_j}}}\,, \tag{10}$$

where N is the cardinality of the domain of S_{R_j} and $\sigma_{S_{R_j}}$ is the standard deviation of the signature. C_L represents the linear correlation of two signatures S_1 and S_2, given by

$$C_L(S_1,S_2) = FT^{-1}\left\{|FT(S_2)|\, e^{i\phi}\, |FT(S_1)|\, e^{-i\varphi}\right\}\,, \tag{11}$$

here φ and ϕ are the phases of the Fourier transform of the signatures S_1 and S_2, respectively. The notation $C_L(S_1)$ implies the autocorrelation function [4]. Analogously, the pattern in a problem image P, is set by [5]

$$r_P = \frac{\max\left\{C_L(S_{R_j},S_P)\right\}}{(N-1)\sigma_{S_{R_j}}\sigma_{S_P}}\,, \tag{12}$$

where S_P is the signature of the problem image and σ_{S_P} is the standard deviation of that signature. If r_P is similar to r_{R_j}, therefore P and R_j are the same, otherwise they are different.

To show that the signatures are invariant to rotation an experiment was conducted. Two pictures of the same object at different rotation angle were taken with a Panasonic camera, model Lumix DMC-FP3, the images were named $I_1(x,y)$ and $I_2(x,y)$ (Figure 3(a) and Figure 3(b) respectively). Figure 3(d) shows the signatures of $I_1(x,y)$ and $I_2(x,y)$ in dashed line and dashed line with asterisks respectively; In order to explain clarity, only the values of the outer rings are shown. As it is expected, both curves are practically the same, the small difference appears from the computers intrinsic round-off error due to the floating-point computation. On the other hand, when the image $I_1(x,y)$ is rotated by a computer the *sawtooth* noise is presented like in Figure 3(c). That noise affects the signature of the image, as it is shown in Figure 3(d), where the dashed line with squares has the same trend like the dashed line but the values in some ring index are different. Therefore, when the images samples are taken in a microscope or in a digital camera, the signatures remain the same does not matter the rotation angle of the object.

To classify fragments of objects, the system will be trained using 150 images of each reference image R_j: 50 images without randomly 5% of the area of the object, 50 without 10% and 50 without 15%. Then, their

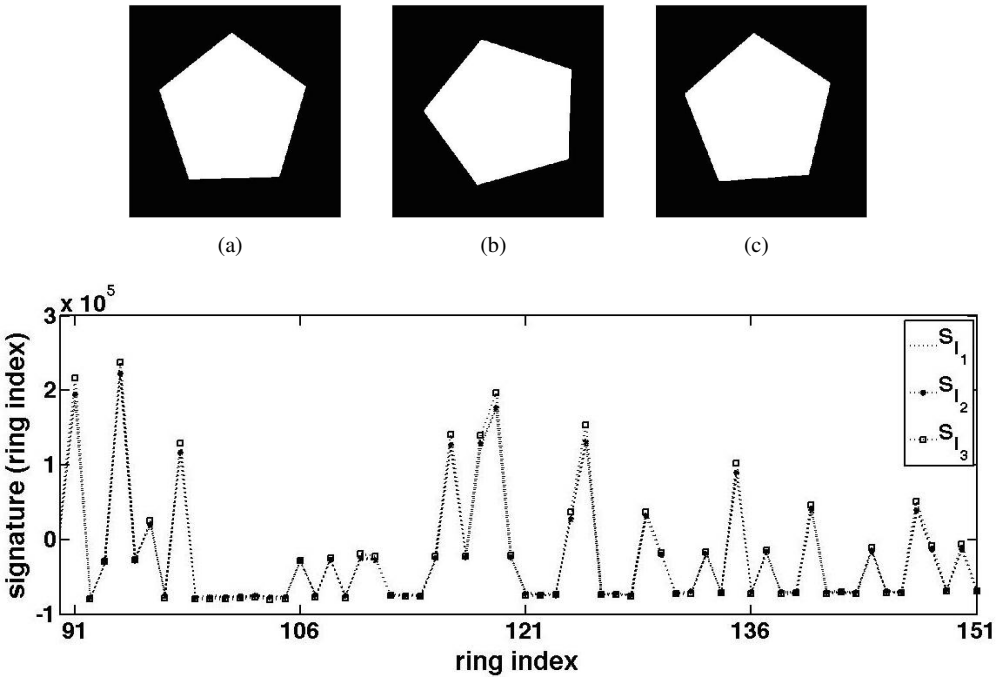

Figure 3. (a) Image $I_1(x,y)$. (b) Image $I_2(x,y)$. (c) Image $I_3(x,y)$. (d) The S_{I_1}, S_{I_2} and S_{I_3} are the signatures of the images $I_1(x,y)$, $I_2(x,y)$ and $I_3(x,y)$, respectively.

corresponding Pearsons correlation coefficient r^k, $k = 1, ..., 150$ were obtained. Because of $\{r^k, k = 1, ..., 150\}$ does not have a normal distribution, those values are normalised by the Z-Fisher transform to get the confidence interval for the correlation values [5]. The Z-Fisher value for $r^k_{R_j}$ is given by

$$Z_{r^k_{R_j}} = 1.1513\ ln\left(\frac{1 + r^k_{R_j}}{1 - r^k_{R_j}}\right). \tag{13}$$

Thus, the 95% confidence interval for $Z_{r^k_{R_j}}$ is

$$\left[Z^-_{r^k_{R_j}}, Z^+_{r^k_{R_j}}\right] = \left[Z_{r^k_{R_j}} - 1.96\,\sigma_Z,\ Z_{r^k_{R_j}} + 1.96\,\sigma_Z\right], \tag{14}$$

with a standard deviation of $\sigma_Z = \frac{1}{\sqrt{n-3}}$ and $n = 151$ the size of the sample. Hence, the confidence interval for the correlation coefficient $\rho_{r^k_{R_j}}$ is

$$\rho^-_{r^k_{R_j}} \leq \rho_{r^k_{R_j}} \leq \rho^+_{r^k_{R_j}}, \tag{15}$$

where

$$\rho^-_{r^k_{R_j}} = \frac{\exp\left(2Z^-_{r^k_{R_j}}\right) - 1}{\exp\left(2Z^-_{r^k_{R_j}}\right) + 1}, \quad \rho^+_{r^k_{R_j}} = \frac{\exp\left(2Z^+_{r^k_{R_j}}\right) - 1}{\exp\left(2Z^+_{r^k_{R_j}}\right) + 1}. \tag{16}$$

For R_j there are 151 values of $\rho^-_{r^k_{R_j}}$ and another 151 values for $\rho^+_{r^k_{R_j}}$ then the confidence interval of 95% to decide if a problem image and R_j are the same is given by

$$\left[\min_{1\leq k\leq 151}\left\{\rho^-_{r^k_{R_j}}\right\}, \max_{1\leq k\leq 151}\left\{\rho^+_{r^k_{R_j}}\right\}\right]. \tag{17}$$

The digital system by Fourier mask was tested using, as target images database, 18 gray scale digital images of diatoms fossil. Those images were selected because of their similarity in their morphology. The problem images database has 2700 images (50 images of each diatom without randomly 5%, 10% and 15% of the area of the object). The digital system in Figure 4 classifies the problem images with a confidence level of 95%.

The same procedure presented in this work was previously performed using the Bessel binary rings mask [6] obtaining that the Bessel mask system recognize objects which have 15% of occlusion, but the Fourier mask digital system is a better option, because it is robust in the classification of images that have noise (like Gaussian additive noise or salt and pepper noise) and non-homogeneous illumination. A similar digital system was presented in [2], essentially the difference between that work and this; it is that the step for the classification

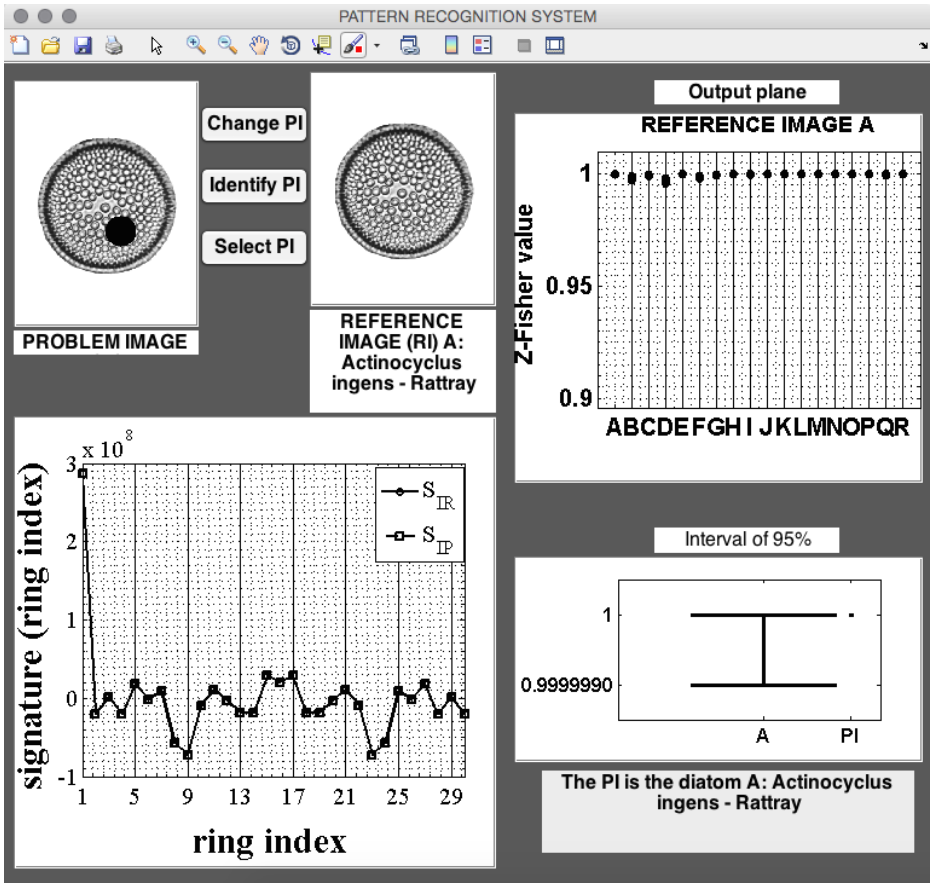

Figure 4. Pattern recognition digital system.

of the images is not achieved in [2]. In there, the confidence level of the system is determined by the boxplot technique. Here, by the Z-Fisher transform the classification step is obtained and a Matlab application was developed to classify fossil diatom digital images. Moreover, this system also classifies images that present fragments of objects.

4. Conclusions

The digital system presented in this work has a confidence level of 95% in the classification of gray scale images, which is an excellent performance. The images to be identified could have a translation and rotation of the object, noise, non-homogenous illumination and the occlusion or elimination of a fragment until 15% in the area of the object.

Acknowledgements

This work was partially supported by CONACyT under grant No.169174. Carolina Barajas-García is a student in the MSc program MyDCI offered by Universidad Autónoma de Baja California and she is supported by CONACyTs scholarship. The grammatical critical review of the manuscript by L.D.I Horacio Padilla Calderón is greatly appreciated. We also thank the unknown reviewers for excellent comments to improve this manuscript.

References

[1] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, third edition. Pearson Education, Inc. New Jersey 2008.

[2] J. Álvarez-Borrego, S. Solorza and M.A. Bueno-Ibarra. "Invariant Correlation to Position and Rotation Using a Binary Mask Applied to Binary and Gray Images", *Optics Communication*, 294, 105-107, 2013.

[3] G.H. Golub and C.F. Van loan, *Matrix Computations*, third edition. The Johns Hopkins University Press. Baltimore 1996.

[4] H.P. Hsu, *Anlisis de Fourier*, Addison Wesley Longman de México S.A. de C.V. 1998.

[5] A. Sánches-Bruno and A. Borges del Rosal. "Transformación Z de Fisher para la determinación de intervalos de confianza del coeficiente de correlación de Pearson", *Psicothema*, vol. 17, No.1, 148-153,2005.

[6] S. Solorza, C. Barajas-García and J. Álvarez-Borrego. "Digital system to classify images by Bessel masks", *To be published in the Journal of Physics: Conference Series enlightening the future*, 23rd Congress of the International Commission for Optics.