PROCEEDINGS OF SPIE

Applications of Digital Image Processing XXXVIII

Andrew G. Tescher
Editor

10–13 August 2015
San Diego, California, United States

Sponsored and Published by
SPIE

Volume 9599
Part One of Two Parts

Proceedings of SPIE 0277-786X, V. 9599

Applications of Digital Image Processing XXXVIII, edited by Andrew G. Tescher,
Proc. of SPIE Vol. 9599, 959901 · © 2015 SPIE · CCC code: 0277-786X/15/$18
doi: 10.1117/12.2218079

A POSITION, ROTATION AND SCALE INVARIANT IMAGE DESCRIPTOR BASED ON RAYS AND CIRCULAR PATHS

Selene Solorza-Calderón

Facultad de Ciencias, UABC, Km. 103, Carretera Tijuana-Ensenada, C.P. 22860, Ensenada, B.C., México

ABSTRACT

In this paper a rotation, scale and translation (RST) invariant image descriptor based on 1D signatures is presented. The position invariant is obtained using the amplitude spectrum of the Fourier transform of the image. That spectrum is introduced in the analytical Fourier-Mellin transform (AFMT) to obtain the scale invariance. From the normalized AFMT amplitude spectrum two 1D signatures are constructed. To build a 1D circular signature, circular path binary masks are used to filter the spectrum image. On the other hand, ray path binary filters are utilized in the construction of the 1D ray signature. These 1D signatures are RST invariant image descriptors. The latin alphabet letters in arial font style were used to test the descriptor efficiency. According with the statistical analysis of bootstrap with a constant replacement $B = 1000$ and normal distribution, the descriptor has a confidence level at least of 95%.

Keywords: Pattern recognition, feature extraction, RST invariant image descriptor, 1D signatures, analytical Fourier-Melling transform.

1. INTRODUCTION

Nowadays, with the great advance in technology, the pattern recognition via digital images is a very productive area. However, the feature extraction process to generate a descriptor invariant to geometric transformations of the object (translation, rotation, scale, noise, illumination and others) is not a trivial problem. A lot of methodologies in digital image features extraction based on joint transforms correlators are developed. In Ref. 1, the Radon transform and the 1D Fourier-Mellin transform are used to build a 2D RST invariant classifier. The classification step is realized by the use of the 2D cross-correlation of the target and the problem image. Because the Radon transform generates a circular shift in the angular variable, 180 2D cross-correlation values are calculated for each pair of images, employing a lot of computation time in the classification process. On the other hand, the 2D Fourier-Mellin transform (FMT) is utilized to design 2D RST invariant classifiers. Because of the factor $\frac{1}{r}$ in this transform, generally the translation invariance is done in the spatial domain via the centroid or the center of mass of the objects,[2] but removing a small disk around the centroid or the center of mass to reduce the large effect of the factor $\frac{1}{r}$. To eliminate the influence of $\frac{1}{r}$, in Ref. 3 is proposes the analytical Fourier-Mellin transform (AFMT) where the images are weighted by the factor r^{σ} with $\sigma > 0$. However, this transform not preserves the rotation and scale invariance, but weighting it by two of the AFMT harmonics a RST invariant descriptor is obtained.[2]

Recently, pattern recognition systems based on binary rings masks[4–7] are developed. These methodologies are robust and efficient in the pattern recognition for gray-level images regardless the position, rotation and scale the object presents. Also, the response of these systems are great under non-homogenous illumination and noise. In Refs. 4–6 the pattern recognition systems are position and rotation invariance only. In Ref. 7 the scale invariance is achieved via the 2D non-separable scale transform. This 2D transform is not invariant to translation. Therefore, the center of mass of the object is used as the center of the image to solve it. Those methodologies work with 1D signatures building through binary rings masks. The use of 1D signatures reduces

E-mail: selene.solorza@uabc.edu.mx. Telephone: +52 (646) 1744560.

Applications of Digital Image Processing XXXVIII, edited by Andrew G. Tescher,
Proc. of SPIE Vol. 9599, 95991V · © 2015 SPIE · CCC code: 0277-786X/15/$18
doi: 10.1117/12.2187258

considerably the computational time investment in comparison step by correlation functions.

This work proposes two 1D signatures invariant to rotation, scale and translation. Also, it presents a new methodology to construct a single output plane (without the use of correlation functions) instead of the multiple output planes generated in the correlation descriptors. The pattern recognition digital system utilizes the amplitude spectrum of the Fourier transform to obtain the translation invariance. Next, it is used in the normalized analytic Fourier-Mellin transform (AFMT) to achieve an amplitude spectrum invariant to position and scale. Then, this is filtered with circular masks to build the 1D RST circular signature. Additionally, the AFMT spectrum is filtered with ray masks to build the 1D RST ray signature. The output plane to classify images is constructed using the 95% confidence interval (CI) for the amplitude values of the circular and ray signatures of each training images. The CI is generated by the statistical method of bootstrap using a replacement value of $B = 1000$ and normal distribution.[8,9] The plot of the CI for the amplitude values for the circular signature versus the CI for the amplitude values for the radial signatures generates a single output plane in which the images are located into a specific region, reducing in this form the computational cost time in the classification step.

This work is organized as follows: Section 2 explains the mathematical foundations of the RST pattern recognition system. Section 3 exposes the procedure to construct the classifier output plane of 95% confidence level. Finally, conclusions are given in Sec. 4.

2. ROTATION, SCALE AND TRANSLATION (RST) INVARIANT PATTERN RECOGNITION SYSTEM

The amplitude spectrum of the Fourier transform of a given image $I(x, y)$ is invariant to translation.[10] Mathematically, it is

$$A(u, v) = |F(u, v)| = \sqrt{R_{_F}^2(u, v) + I_{_F}^2(u, v)}, \tag{1}$$

where

$$F(u, v) \quad = \quad R_{_F}(u, v) + iI_{_F}(u, v) = \mathcal{F}\{I(x, y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x, y)e^{-i2\pi(ux+vy)}dxdy. \tag{2}$$

Then, $A(u, v)$ is introduced in the fast analytical Fourier-Mellin transform (AFMT),

$$M(k, \omega) \quad = \quad \mathcal{M}\{A(e^{\rho}, \theta)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{0}^{2\pi} A(e^{\rho}, \theta)e^{\rho\sigma}e^{-i(k\theta+\rho\omega)}d\theta d\rho, \tag{3}$$

where $\rho = \ln(r)$ and $\sigma > 0$. Equation (3) is not an invariant to scale and rotation, but normalizing the $M(k, \omega)$ by its value in the central pixel (c_x, c_y), the amplitude spectrum is a scale invariance,[2] that is

$$S(k, \omega) = \left| \frac{M(k, \omega)}{M(c_x, c_y)} \right|. \tag{4}$$

Figures 1(j) to 1(l) show $S_1(k, \omega)$, $S_2(k, \omega)$ and $S_3(k, \omega)$, the normalized analytical Fourier-Mellin amplitude spectrum associated to Figs. 1(a) to 1(c), respectively.

2.1 The 1D circular signature

To achieve the RST invariant 1D circular signature, the normalized AFMT spectrum S, given by Eq.(4) with $\sigma = \frac{1}{2}$, is filtered by binary masks with a circle of radius r. The circle is centred in the (c_x, c_y) pixel, which is the central pixel of the image. Fig.2 shows some circle mask examples. Mathematically, the filter procedure is

$$H_r = S \otimes C_r, \tag{5}$$

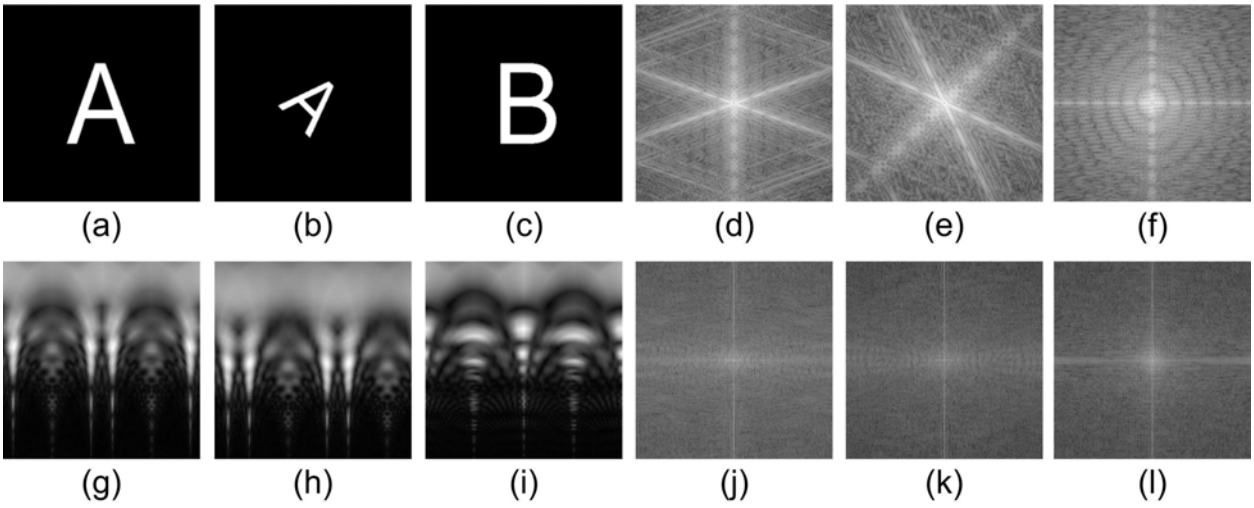


Figure 1. AFMT amplitude spectrum examples. (a) Image I_1: The A arial font letter without geometric transformations. (b) Image I_2: The A arial font letter with a rotation angle of 315° and scaled −25%. (c) Image I_3: The B arial font letter without geometric transformations. (d) $A_1(u,v) = |\mathcal{F}\{I_1(x,y)\}|$. (e) $A_2(u,v) = |\mathcal{F}\{I_2(x,y)\}|$. (f) $A_3(u,v) = |\mathcal{F}\{I_3(x,y)\}|$. (g) $A_1(e^{\rho},\theta)e^{\rho\sigma}$. (h) $A_2(e^{\rho},\theta)e^{\rho\sigma}$. (i) $A_3(e^{\rho},\theta)e^{\rho\sigma}$. (j) The $S_1(k,\omega)$ of Fig. 1(g). (k) The $S_2(k,\omega)$ of Fig. 1(h). (l) The $S_3(k,\omega)$ of Fig. 1(i).

where C_r is the circle binary mask of the same size of S, $r = 1,\ldots,N$, $N = min\{c_x, c_y\}$ and $\otimes$ represents the element-wise product (Hadamard multiplication). Next, the addition of the intensity values in H_r is calculated like

$$k_r = \sum_{x=1}^{n}\sum_{y=1}^{m} H_r(x,y). \tag{6}$$

Finally, the 1D circular signature S_c is built just as

$$\begin{aligned} S_c &: \{1,\ldots,N\} \to \mathbb{R}, \\ S_c(r) &= k_r. \end{aligned} \tag{7}$$

Figure 3(c) shows the circular signature of the images in Figs. 3(a) and 3(b). The length of the signature is the number of circles that sampling the entire disk of radius $r = N$, central pixel (c_x, c_y) and $\Delta r = 1$.

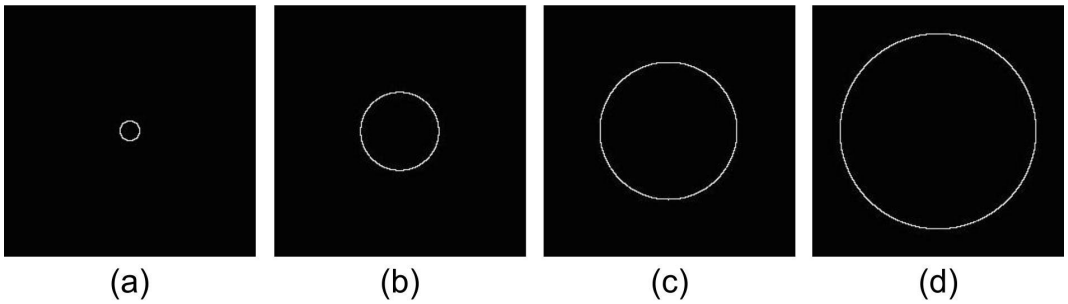


Figure 2. Circle binary mask examples. (a) C_{10}. (b) C_{40}. (c) C_{70}. (d) C_{100}.

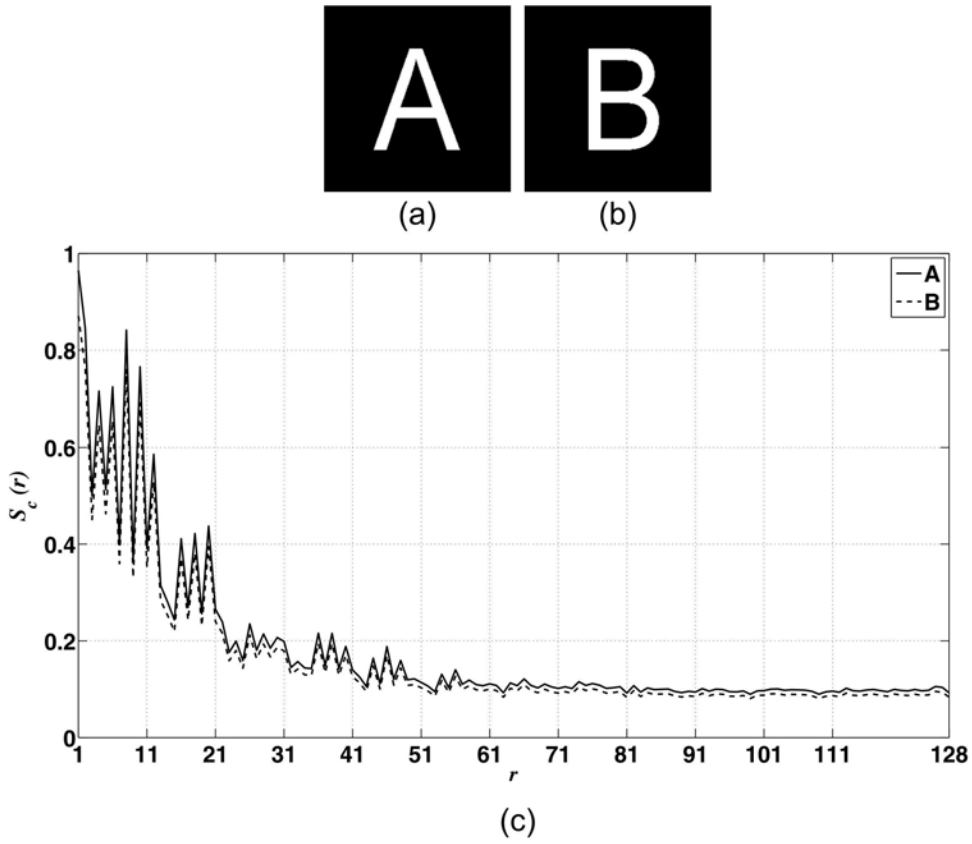


Figure 3. Circular signature examples. (a) Black and white image with the A arial font letter. (b) Black and white image with the B arial font letter. (c) Circular signatures of Figs. 3(a) and 3(b).

2.2 The 1D ray signature

On the other hand, to get the RST invariant 1D ray signature, the normalized AFMT spectrum S, given by Eq.(4) with $\sigma = \frac{1}{2}$, is filtered by binary ray images R_θ of the same size of S. Figure 4 shows some binary ray filters. Mathematically, this procedure is

$$G_\theta = S \otimes R_\theta, \tag{8}$$

where $\theta = 0, \ldots, 359$. The ray has length $N = min\{c_x, c_y\}$ and it departures from the central pixel (c_x, c_y) of the image. Then, the addition of the intensity values in G_θ is calculated by

$$\tilde{k}_\theta = \sum_{x=1}^{n} \sum_{y=1}^{m} G_\theta(x, y), \tag{9}$$

next, a 1D ray signature is built just as

$$\begin{aligned} \tilde{S}_r &: \{0, \ldots, 359\} \to \mathbb{R}, \\ \tilde{S}_r(\theta) &= \tilde{k}_\theta. \end{aligned} \tag{10}$$

Because of the circular shift in the AFMT spectrum which appears from the log-polar transformation (for example Figs. 1(g) and 1(h)), the $\tilde{S}_r(\theta)$ values are ordered in descendent manner to obtain the RST invariant 1D ray signature S_r, that is

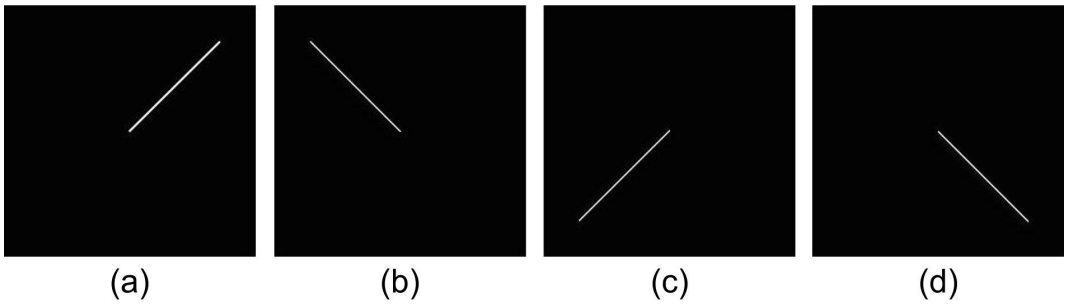


Figure 4. Ray binary mask examples. (a) R_{45}. (b) R_{135}. (c) R_{225}. (d) R_{315}.

$$S_r(\alpha) \quad = \quad \tilde{k}_\alpha, \tag{11}$$

where $\tilde{k}_0 \leq \tilde{k}_2 \leq \ldots \leq \tilde{k}_{359}$. In Fig. 5(c) is shown the circular signatures for Figs. 5(a) and 5(b). The length of the signature is the number of rays that sampling the entire circle with $\Delta\theta = 1°$.

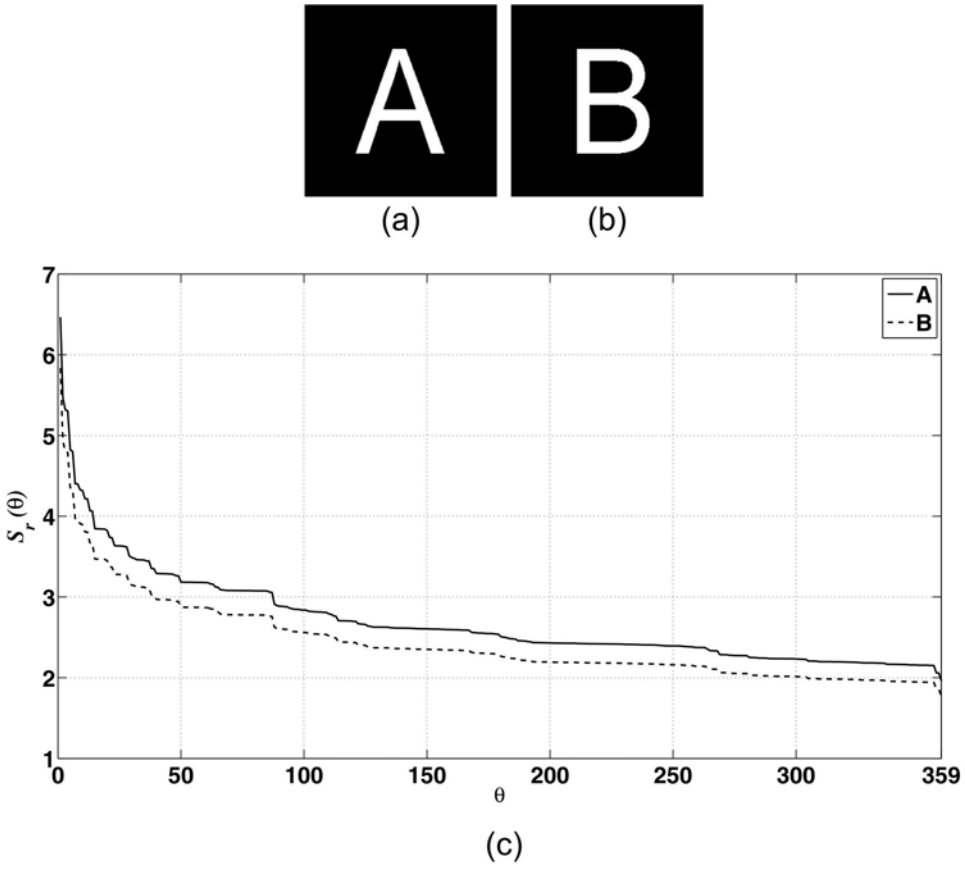


Figure 5. Ray signature examples. (a) Black and white image with the A Arial font letter. (b) Black and white image with the B Arial font letter. (c) Ray signatures of Figs. 5(a) and 5(b).

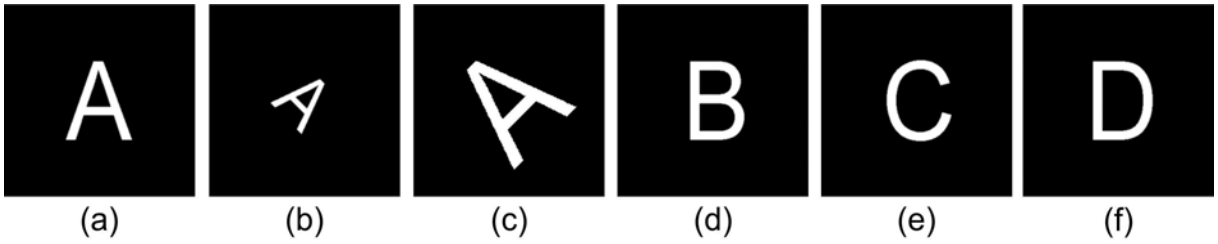


Figure 6. Some target images in the database. (a) The A arial font letter without geometric transformations. (b) The A arial font letter with a rotation angle of 315° and scaled −25%. (c) The A arial font letter with a rotation angle of 45° and scaled 25%. (d) The B arial font letter without geometric transformations. (e) The C arial font letter without geometric transformations. (f) The D arial font letter without geometric transformations.

3. CONFIDENCE INTERVAL FOR THE RST INVARIANT DESCRIPTOR

To train the pattern recognition RST invariant image descriptor, black and white 257×257pixel digital images with the latin alphabet letters in arial font style were used. The images were rotated 360° with $\Delta\theta = 1°$. Then, each of those images were scaled ±25%. Figure 6 shows some examples. Next, the amplitude values of the 1D signatures S_c and S_r of each image are calculated like

$$A_c = \sqrt{\sum_{k=1}^{N} S_c(k)^2}, \tag{12}$$

$$A_r = \sqrt{\sum_{k=0}^{359} S_r(k)^2}. \tag{13}$$

For example, to Fig. 6(a) 18,360 images were obtained, from those images 18,360 A_c values were generated and based on that values a 95% confidence interval (CI) was constructed by the statistical method of bootstrap using a replacement constant $B = 1,000$ and normal distribution.[8,9] Analogously, the 95% confidence interval (CI) corresponding to the A_r values was constructed. Figure 7(a) shows the output plane for the latin alphabet letter in arial font style. The horizontal and vertical axes represent the CI of the A_r and A_c values, respectively. A rectangle zone is assigned to classify each type of image (Fig. 7(b) displays an amplification of the output plane to observe the rectangle zone assigned to some letters). Because those rectangles are not overlapped, the descriptor has a confidence level at least of 95%.

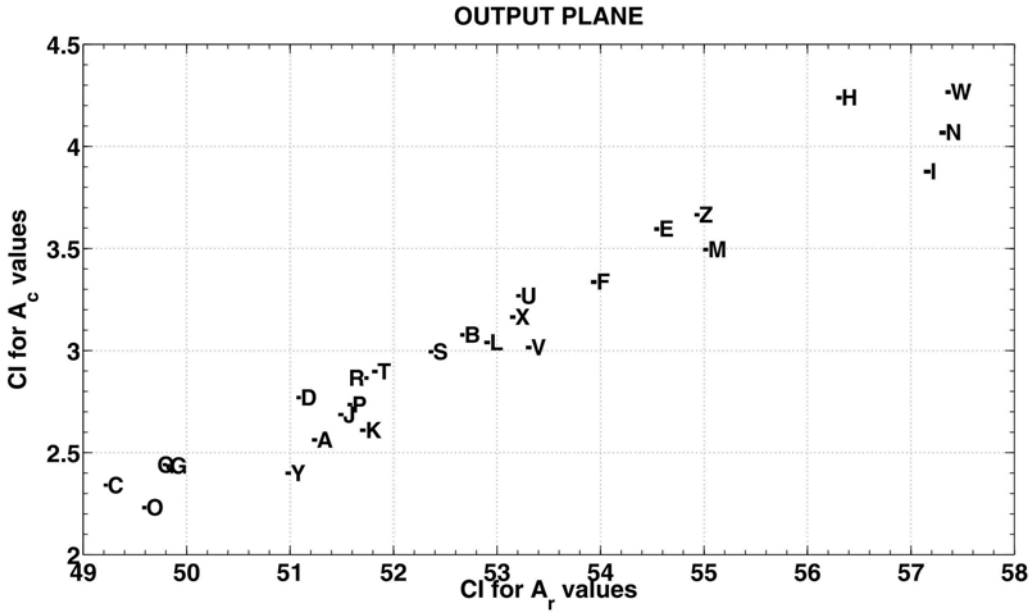
The descriptors based on the correlation function[1,5,6] or distance functions[2,11] generate an output plane for each target image. Therefore, in that works 26 output planes will be constructed to identify the digital images. This work propose a new methodology to reduce the computational cost time in the classification step by the use of only one output plane.

4. CONCLUSIONS

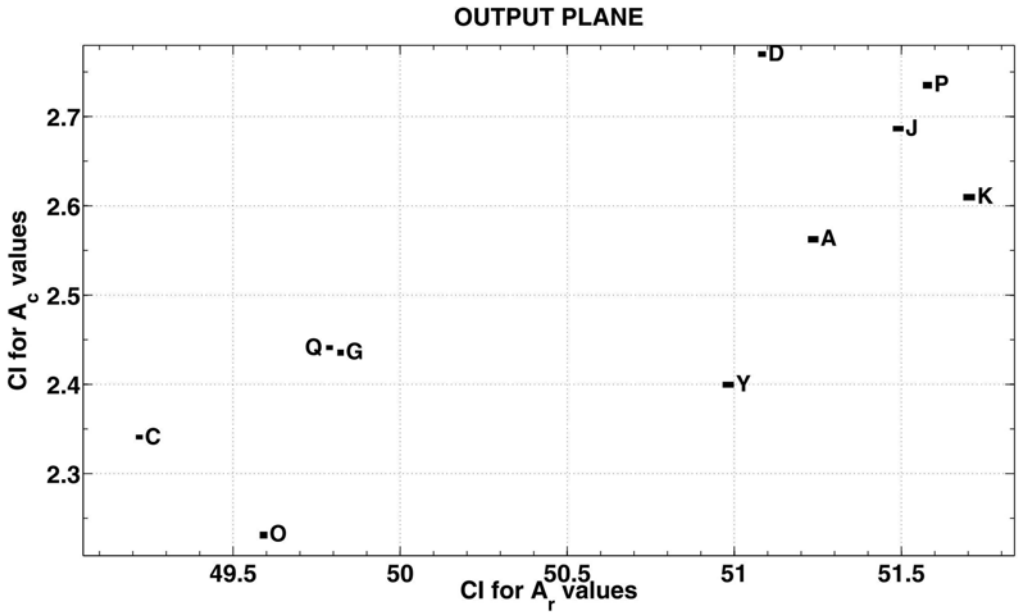
This work presents a new digital image descriptor invariant to rotation, scale and translation based 1D signatures obtained from binary masks of circular and ray paths, and the amplitude spectrum of the normalized analytical Fourier-Mellin transform. The RST descriptor presents a confidence level of 95% in the pattern recognition of translated, rotated and scaled black and white images with the latin alphabet letters in arial font style. Moreover, in this RST image descriptor is proposed a new methodology to generate a single classifier output plane, reducing in this manner the comparison cost time in the classification step.

Acknowledgments

This work was partially supported by CONACyT under grant No. 169174.



(a)



(b)

Figure 7. (a) The classifier output plane. (b) Amplification zone of the classifier output plane.

REFERENCES

[1] T. Hoang and S. Tabbone, "Invariant pattern recognition using the rfm descriptor," *Pattern Recognition* **45**(1), pp. 271–284, 2012.

[2] S. Derrode and F. Ghorbel, "Robust and efficient fourier-mellin transform approximations for gray-level image reconstruction and complete invariant description," *Computer Vision and Image Understanding* **83**(1), pp. 57–78, 2001.

[3] F. Ghorbel, "Towards a unitary formulation for invariant image description ; application to towards a unitary formualtion for invariant image description; application to image coding," *Annals of Telecommunication* **53**, pp. 242–260, 1998.

[4] S. Solorza and J. Álvarez-Borrego, "Digital system of invariant correlation to position and rotation," *Optics Communications* **283**, pp. 3613–3630, 2010.

[5] S. Solorza and J. Álvarez-Borrego, "Translation and rotation invariant pattern recognition by binary rings masks," *Journal of Modern Optics* **62**(10), pp. 851–864, 2015.

[6] J. Álvarez-Borrego, S. Solorza, and M. Bueno-Ibarra, "Invariant correlation to position and rotation using a binary mask applied to binary and gray images," *Optics Communications* **294**, pp. 105–117, 2013.

[7] A. Solís-Ventura, J. Álvarez-Borrego, and S. Solorza, "An adaptive nonlinear correlation with a binary mask invariant to rotation and scale applied to identify phytoplankton," *Optics Communications* **339**, pp. 185–193, 2015.

[8] A. Davison and D. Hinkley, *Bootstrap methods and their application*, Cambridge University Press, 1997.

[9] B. Efron and R. Tibshirani, *An introduction to the bootstrap*, Chapman and Hall, 1993.

[10] R. Gonzalez, R. Woods, and S. Eddins, *Digital image processing using Matlab*, Tata McGraw Hill Education Private Limited, 2010.

[11] J. Lerma-Aragón and J. Álvarez-Borrego, "Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition," *Journal of Modern Optics* **56**(14), pp. 1598–1606, 2009.