# Digital Image Pattern Recognition System Using Normalized Fourier Transform and Normalized Analytical Fourier-Mellin Transform

Rodrigo Vélez-Rábago, Selene Solorza-Calderón and Adina Jordan-Aramburo

*Universidad Autónoma de Baja California, Facultad de Ciencias, Km.103 Carretera Tijuana-Ensenada, C.P. 22860, Ensenada B.C., México.*[1]

**Abstract.** This work presents an image pattern recognition system invariant to translation, scale and rotation. The system uses the Fourier transform to achieve the invariance to translation and the analytical Forier-Mellin transform for the invariance to scale and rotation. According with the statistical theory of box-plots, the pattern recognition system has a confidence level at least of 95.4%.

**Keywords:** Pattern recognition algorithms, Fourier transform, Fourier-Mellin transform, digital image classification method.
**PACS:** 07.05.Pj; 42.30.Tz; 89.20.Ff.

## INTRODUCTION

This work presents an image descriptor invariant to translation, scale and rotation (TSR) based on the Fourier transform and the analytical Fourier-Mellin transform. The analytical Fourier-Mellin transform (AFTM) has proven to be a robust and efficient method to get the invariance to scale and rotation[1]; however, the translation invariant is often left out arguing that this can be done by using the object's centroid, rather than the actual center of each image, in the AFTM. In this work, the Fourier transform is used to accomplish the invariance to translation instead of the object's centroid[2]. Also, it is used a single classifier output space of 95.4% confidence cuboids to reduce considerably the computational time investment in the classification step[3].

The rest of the work is organized as follows: Section II describes the procedure to develop the TSR invariant image pattern recognition system based on the Fourier transform and the AFMT. Section III presents the methodology to construct the 95.4% cuboids output space. Finally, conclusions are given in section IV.

## THE TSR INVARIANT PATTERN RECOGNITION SYSTEM

### The 1D TSR signatures of the image

The pattern recognition system works with gray-scale square images. Let $I_2(x,y)$ be a translated, rotated and scaled version of the image $I_1(x,y)$, that is

$$I_2(x,y) = I_1(\alpha x \cos\phi + \alpha y \sin\phi - \hat{x}, -\alpha x \sin\phi + \alpha y \cos\phi - \hat{y}), \tag{1}$$

where $x,y = 1,2\ldots,n$; $\hat{x}$ and $\hat{y}$ are the values of the object's translation in the horizontal and vertical axis, respectively; $\phi$ and $\alpha \neq 0$ are the rotation and scale factors of the object, respectively. Then, the Fourier transforms of the images $I_1$ and $I_2$ are related by

$$F_2(u,v) = e^{-i[u(\frac{\hat{x}}{\alpha}\cos\phi - \frac{\hat{y}}{\alpha}\sin\phi)+v(\frac{\hat{y}}{\alpha}\cos\phi+\frac{\hat{x}}{\alpha}\sin\phi)]} F_1(\alpha u \cos\phi + \alpha v \sin\phi, -\alpha u \sin\phi + \alpha v \cos\phi), \tag{2}$$

[1] Corresponding author: rvelez@uabc.edu.mx

where $u, v = 1, 2, \ldots, n$; $F_1$ and $F_2$ represent the Fourier transform of the images $I_1$ and $I_2$, respectively[4]. Let's denote by $A_1$ and $A_2$ the amplitude spectra of $F_1$ and $F_2$, therefore the spectra are related by

$$A_2(u,v) = A_1(\alpha u \cos\phi + \alpha v \sin\phi, -\alpha u \sin\phi + \alpha v \cos\phi). \tag{3}$$

As it is observed in Eq. (3), $A_2(u,v)$ is a rotated and scaled version of $A_1(u,v)$, furthermore $\hat{x}$ and $\hat{y}$ are not part of this relation, hence the invariance to translation is achieved. The amplitude spectrum $A(u,v)$ of a given image $I(x,y)$ is used in the analytical Fourier-Mellin transform, but $A(u,v)$ is expressed in log-polar coordinates to obtain the AFMT via the Fourier transform[1], that is

$$M(k,\omega) = \mathscr{M}\{A(e^\rho,\theta)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_0^{2\pi} A(e^\rho,\theta) e^{\rho\sigma} e^{-i(k\theta+\rho\omega)} d\theta d\rho, \tag{4}$$

where $\rho = \ln(r)$ and $\sigma > 0$; in this work $\sigma = 0.5$. Equation (4) is not invariant to scale and rotation, but normalizing the AFMT by its value in the central pixel $(c_x, c_y)$ the AFMT amplitude spectrum is scale invariant, that is

$$S(k,\omega) = \left|\frac{M(k,\omega)}{M(c_x,c_y)}\right|. \tag{5}$$

In order to achieve the invariance to rotation and also, to reduce the computational time investment in the pattern recognition system, the row and column marginals of $S(k,\omega)$ are computed to obtain two 1D TSR invariant signatures. The signature built from the row marginals is called $S_r$ and the signature based on the column marginals is named $S_c$. A third 1D TSR invariant signature is constructed by $S_p = Arg\{\mathscr{F}\{S_c\}\}$, where $\mathscr{F}$ is the 1D Fourier transform and $Arg$ is the phase of the 1D Fourier transform of $S_p$.

For the classification step, the features assigned to the image are the power of the signatures, given by

$$P_q = \frac{\Sigma(S_q)^2}{N_q}, \tag{6}$$

where $q = r, c, p$ and $N_q$ is the length of the signature $S_q$.

## THE OUTPUT SPACE

Two set of images were used to test the confidence of the TSR invariant pattern recognition system. The first set contains Black and White (BW) $369 \times 369$ Latin alphabet letters. The second set is formed with $318 \times 318$ gray-scale images from the Coil-20 object image library: http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php.

### BW images

To train the TSR invariant pattern recognition system, each image in the reference image database in Fig. 1 was rotated 360° using $\Delta\theta = 1°$. Afterwards, those images were scaled $\pm 20\%$ with a scale step $\Delta h = 1\%$. Hence, 14,760 samples are generated from each image in Fig. 1. Next, the three TSR invariant signatures of each one of all those images were obtained. Finally, once the power values of those signatures were determined, the confidence intervals are built in the following form: let $R_k$ be the $k$-th reference image in the database, using the 14,760 power values of $S_c$, the 95.4% confidence interval (CI) was constructed by the statistical method of box-plots with $\mu_c \pm 2EE_c$, where $\mu_c$ is the mean of the power values and $EE_c$ is the standard error. Analogously, the confidence intervals for $S_r$ and $S_p$ are obtained. Then a cuboid was settled with edges being the confidence intervals $\mu_c \pm 2EE_c$, $\mu_r \pm 2EE_r$, $\mu_p \pm 2EE_p$ and the vertices are set in the coordinates: $(\mu_c - 2EE_c, \mu_r - 2EE_r, \mu_p - 2EE_p)$, $(\mu_c + 2EE_c, \mu_r - 2EE_r, \mu_p - 2EE_p)$, $(\mu_c + 2EE_c, \mu_r + 2EE_r, \mu_p - 2EE_p)$, $(\mu_c - 2EE_c, \mu_r + 2EE_r, \mu_p - 2EE_p)$, $(\mu_c - 2EE_c, \mu_r - 2EE_r, \mu_p + 2EE_p)$, $(\mu_c + 2EE_c, \mu_r - 2EE_r, \mu_p + 2EE_p)$, $(\mu_c - 2EE_c, \mu_r + 2EE_r, \mu_p + 2EE_p)$ and $(\mu_c + 2EE_c, \mu_r + 2EE_r, \mu_p + 2EE_p)$ [3]. Figure 2 shows the cuboids classifier output space for the reference images database in Fig. 1. A volume space was assigned to each image without overlaps, hence the TSR invariant pattern recognition system presents a confidence level at least of 95.4%.


**FIGURE 1.** BW reference image database.

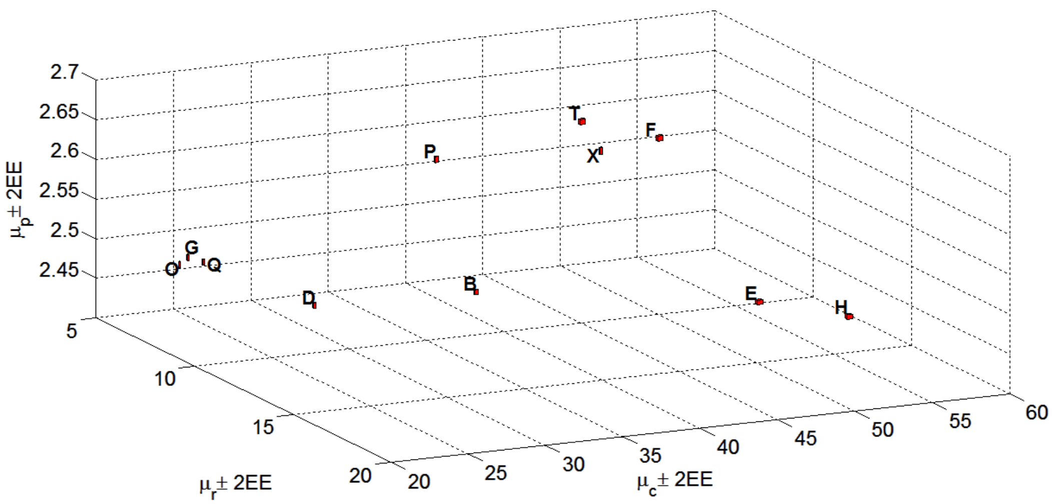


**FIGURE 2.** Cuboids classifier output space for BW images.

## Gray-scale images

To train the TSR invariant pattern recognition system with gray-scale images, the images in the Coil-20 object image library in Fig. 3 were rotated 360° using $\Delta\theta = 10°$. Afterwards, those images were scaled ±16% with a scale step $\Delta h = 2\%$, hence 612 samples are generated from each image. Next, the three TSR invariant signatures are obtained and after that the power values of those signatures were determined to construct the cuboids. Figure 4 shows the cuboids classifier output space for the reference image database in Fig. 3. Because a volume space was assigned to each image without overlaps, the TSR invariant pattern recognition system presents a confidence level at least of 95.4%.


**FIGURE 3.** Coil-20 object image library: http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php.

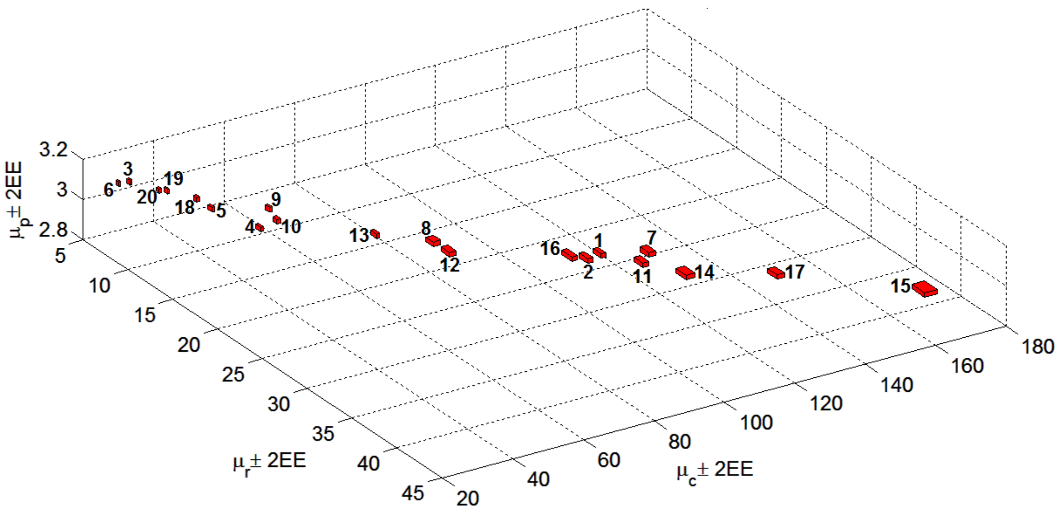


**FIGURE 4.** Cuboids classifier output space for the gray-scale image database in Fig. 3.

## CONCLUSIONS

This work presents a new 1D signatures pattern recognition system invariant to translation, scale and rotation specialized for BW and gray-scale images. The TSR pattern recognition system is based on the Fourier transform and the analytic Fourier-Mellin transform. The pattern recognition descriptor presents a confidence level at least of 95.4% using the cuboids classifier output space. Moreover, the use of the single output space reduces the computation time investment in the classification step.

## ACKNOWLEDGMENTS

This work was partially supported by CONACyT under grant No.169174. Rodrigo Vélez-Rábago is a physics undergraduate student at Universidad Autónoma de Baja California and he is supported by CONACyT's scholarship.

## REFERENCES

1. S. Derrode, and F. Ghorbel, *Computer Vision and Image Understanding* **1**, 57–78 (2001).
2. S. Solorza-Calderón, "A position, rotation and scale invariant image descriptor based on rays and circular paths," in *Proceedings of SPIE: Applications of Digital Image Processing XXXVIII*, edited by A. G. Tescher, SPIE, 2015, vol. 9599, pp. 95991V–1.
3. J. Verdugo-Olachea, *Reconocimiento de patrones en imágenes digitales usando la transformada de Radon*, Tesis de Licenciatura, Universidad Autónoma de Baja California (2015).
4. R. Gonzalez, and R. Woods, *Digital image processing*, Pearson-Prentice Hall, 2008, Third edn.