MOPM
MEXICAN OPTICS AND PHOTONICS MEETING
September 9 -11, 2015
Leon, Guanajuato Mexico

A digital image pattern recognition system invariant to rotation, scale and translation for color images

Carolina Barajas-García (1), Selene Solorza-Calderón (1), Josué Álvarez-Borrego (2)

1. Facultad de Ciencias, Universidad Autónoma de Baja California, México.
2. Div. de Física Aplicada, Centro de Investigación Científica y de Educación Superior de Ensenada, México.

Corresponding author email: selene.solorza@gmail.com

ABSTRACT:

This work presents a new pattern recognition system for color images invariant to rotation, scale and translation (RST). The digital system is based on the Fourier transform, the normalized analytic Fourier-Mellin transform and Bessel binary rings masks to generate 1D RST invariant signatures for each channel in the RGB color space. Using the instantaneous amplitudes of those 1D signatures a classifier cuboids space with confidence level of 95.4% is constructed.

Key words: Pattern recognition, feature extraction, color image, 1D signature, Bessel masks.

REFERENCES AND LINKS

[1] B.L. Boese, P.J. Clinton, D. Dennis, R.C. Golden y B. Kim, “Digital image analysis of Zostera marina leaf injury”, Aquat. Bot. **88**, 87-90 (2008).

[2] D.G. Lowe, “Distinctive image features from scale-invariant key points”, IJCV **60**, 91-110 (2004).

[3] H. Bay, A. Essa, T. Tuytelaars y L. Van Gool, “Speeded-Up Robust Features (SURF)”, CVIU **110**, 346-359 (2008).

[4] Y. Ke y R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors”, CVPR, 506-513 (2004).

[5] E.N. Mortensen, H. Deng y L. Shapiro, “A SIFT descriptor with global context”, CVPR **1**, 184-190 (2005).

[6] D. Su, J. Wu, Z. Cui, V.S. Sheng y S. Gong, “CGCI-SIFT: A more efficient and compact representation of local descriptor” Meas. Sci. Rev. **13**, 132-141 (2013).

[7] A.E. Abdel-Hakin y A.A. Farag, “CSIFT: A SIFT descriptor with color invariant characteristics”, CVPR, 1978-1983 (2006).

[8] R. Javanmard-Alitappeh, K. Jeddi-Saravi y F. Mahmoudi, “A new illumination invariant feature based on SIFT descriptor in color space”, Procedia Eng. **41**, 305-311 (2012).

[9] C. Ancuti y P. Bekaert, “SIFT-CCH: Increasing the SIFT distinctness by color co-occurrence histograms”, ISPA, 130-135 (2007).

[10] S. Solorza y J. Álvarez-Borrego, “Translation and rotation invariant pattern recognition by binary rings masks”, J. Mod. Opt. **62**, 851-864 (2015).

[11] S. Derrode y F. Ghorbel, “Robust and efficient Fourier-Mellin transform approximations for Gray-level image reconstruction and complete invariant description”, CVIU **83**, 57-78 (2001).

---

1. Introduction

Reproduce the pattern recognition human functions are a great challenge and a very difficult task. The research community has been employed a lot effort to create robots and automation systems to this purpose. Color is a very important feature to human pattern recognition process, if this information is neglected very important characteristic could be lost. For example, the color is used to study the Zostera marina leaf injury [1], but the processing of the images is done by hand-operated through multiple imaging programs (Adobe PhotoShop, Canon Photostitch and ERMapper) although local feature descriptors are used in a variety of pattern recognition real-world applications due to the identification efficiency [2-5]. The color-SIFT descriptors are developed to take into account the color feature, but the complexity and the calculation increase considerably for the training and the testing phases [6-9].

This work presents a rotation, scale and translation (RST) invariant color image descriptor based on Bessel binary rings masks methodology developed in [10]. This RT invariant methodology is robust and efficient in the pattern recognition for gray-level images regardless the position and rotation the object presents. To introduce the scale invariance, here is proposed the use of the amplitude spectrum of the normalized analytic Fourier-Mellin transform (AFMT). This spectrum is filtered by a Bessel binary rings mask in order to obtain a RST invariant 1D signature for each channel in the RGB color space. The instantaneous amplitudes of the signatures for the training color images are used to construct cuboids with 95.4% confidence level (based on the statistical boxplot technique); those cuboids are used to build the classifier space; in this manner, the classification step reduces considerably the computational time investment. The rest of the work is organized as follows: Section 2 describes the procedure to develop the RST invariant color image pattern recognition system based on Bessel masks. Section 3 exposes the methodology to construct the classifier cuboids space. Finally, conclusions are given in section 4.

2. The RST Invariant Pattern Recognition System for Color Images

2.a. The Bessel binary rings masks

The binary rings masks are obtained using the ratio of Bessel function of first kind and first order by its argument, given by

$$y(x) = \begin{cases} \frac{J_1(x-c_x)}{x-c_x}, & x \neq c_x, \\ 1, & x = c_x, \end{cases} \tag{1}$$

where $x = 1, \dots, n$, $n \times n$ is the size of the image I and (c_x, c_x) is the central pixel of I. Fig. 1 shows the graph of $y(x)$ with $n = 203$ and $c_x = 102$.

Base on Eq. (1), it is built the following binary functions

$$Z_P(x) = \begin{cases} 1, & y(x) > 0, \\ 0, & y(x) \leq 0, \end{cases} \tag{2}$$

and

$$Z_N(x) = \begin{cases} 1, & y(x) > 0, \\ 0, & y(x) \leq 0, \end{cases} \tag{3}$$

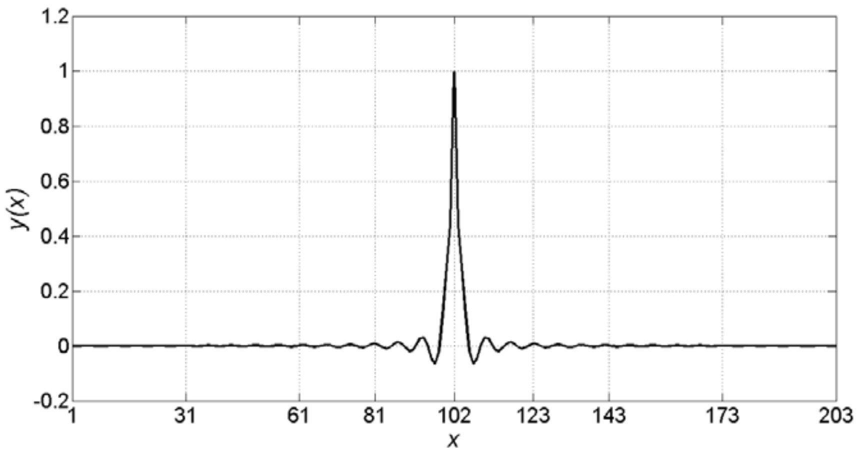

Fig.1. Graph of Eq. (1).

Finally, taking the vertical axis $x = c_x$ as the rotation axis, the Z_P function is rotated 180 degrees to obtain concentric cylinders of height one, different widths and centred in (c_x, c_x) pixel. Taking a cross-section, it is built the Bessel binary rings mask B_P [10]. Analogously, the binary ring mask B_N is generated using the Z_N function. Fig. 2 shows the two Bessel binary rings masks for 203×203 images.

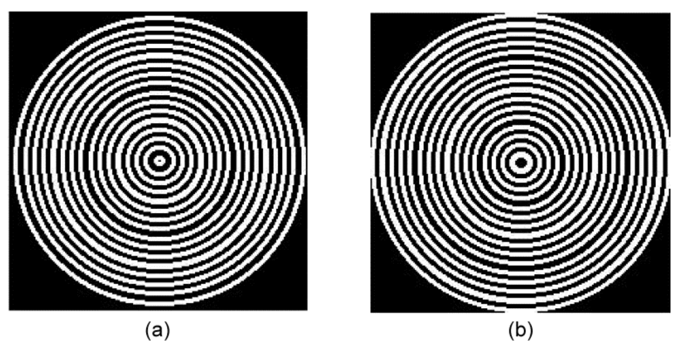

Fig.2. Bessel binary rings masks. (a) Mask B_P. (b) Mask B_N.

2.b. The signature of the image

To obtain the signatures of the color image I, it is split in the three monochromatic images I^C, where C represents the color channel R, G or B. Then, the amplitude spectrums A^C of I^C are introduced in the analytical Fourier-Mellin transform

$$M^C(k,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{0}^{2\pi} A^C(e^\rho,\theta)\, e^{\sigma\rho}\, e^{-i\,(k\theta+\omega\rho)}\, d\theta\, d\rho\,, \qquad (4)$$

where $\rho = ln(r)$ and $\sigma > 0$ [11]. Fig. 3(b) yields the amplitude spectrum for the red channel of Fig. 3(a), because in Eq. (4) this spectrum is set in log-polar coordinates, Fig. 3(c) displays it. Eq. (4) is not invariant to scale yet, but normalizing the AFMT (analytical Fourier-Mellin transform) by its value in the central pixel, the amplitude spectrum of the normalized AFMT is scale invariance,

$$G^C = \left|\frac{M^C}{M^C(c_x,c_x)}\right|. \qquad (5)$$

Fig. 3(d) shows the normalized AFMT amplitude spectrum of Fig. 3(c). The next step is filters the G^C images by a Bessel mask, that is

$$H_P^C = G^C \otimes B_P,$$
$$H_N^C = G^C \otimes B_N, \tag{6}$$

where $\otimes$ means an element-wise product or Hadamard product. An example of this is given in Fig. 3(f). The rings in the normalized AFMT amplitude spectrum (H_P^C or H_N^C) are numbered from inside to outside (without considering the black rings, because they represents zero-intensity values). Then, the intensity values in each ring are added and assigned to the corresponding ring index to construct the function S_P^C called signature of the image [10]. The signature S_P^R obtained from the red channel of Fig. 3(a) using the Bessel mask B_P is shown in Fig. 3(g). Fig. 4(b) shows the three signatures S_P^R, S_P^G and S_P^B associated to Fig. 4(a) using the Bessel mask B_P. It is observes in Fig. 4(b), the three signatures are quite similar, that is because the intensity values in the monochromatic images are similar. If the intensity values are different (e.g. Fig. 4(c)) the signatures differ, as it is exemplified in Fig. 4(d).

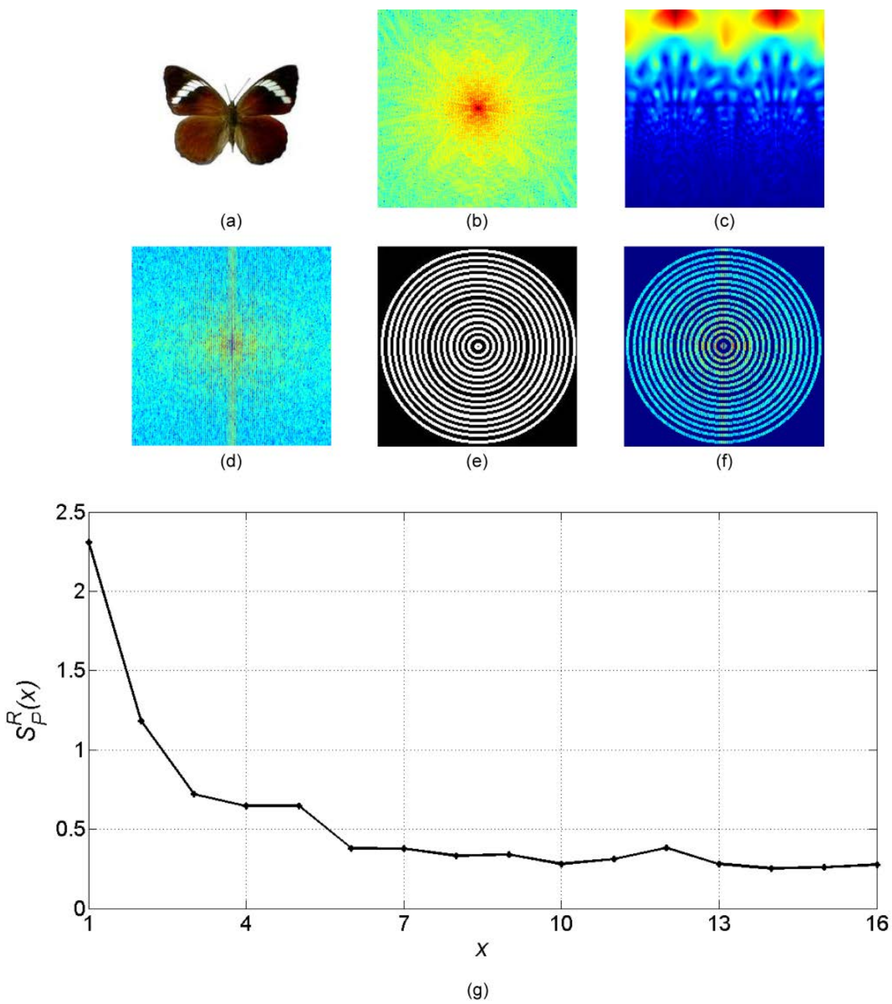

Fig.3. Signature procedure example. (a) Image I. (b) Amplitude spectrum $A^R(x,y)$. (c) $A^R(e^\rho,\theta)\ e^{\sigma\rho}$, with $\sigma = 0.5$. (d) Normalized AFMT amplitude spectrum G^R. (e) Bessel mask B_P. (f) $H_P^R = G^R \otimes B_P$. (g) The signature S_P^R generated using the red channel of Fig. 3(a). For visualization purposes the spectrums are given in logarithmic scale.

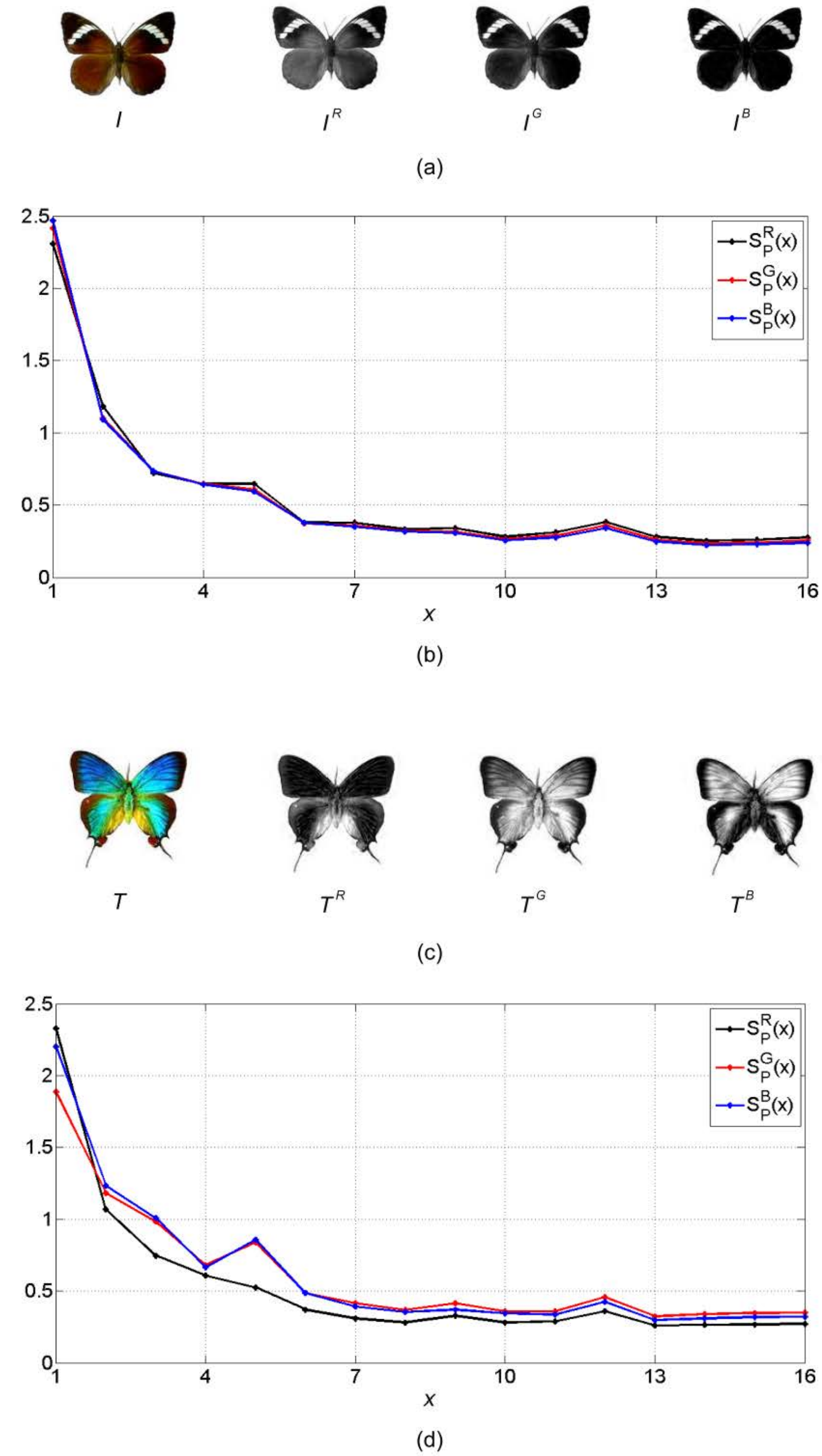

Fig.4. The signatures of the images. (a) The color image I and the monochromatic images I^R, I^G and I^B. (b) The signatures of I^R, I^G and I^B using the binary rings mask B_P. (c) The color image T and the monochromatic images T^R, T^G and T^B. (d) The signatures of T^R, T^G and T^B using the binary rings mask B_P.

The features selected to characterize the image I are the instantaneous amplitude of the signatures, given by

$$A_k^C = \sqrt{\sum\left(S_k^C(x)\right)^2}, \tag{7}$$

where $k = P$ or N.

3. The classifier cuboids space

To train the RST invariant pattern recognition system, each image in the reference image database (e.g. Fig. 5) was rotated 360° using $\Delta\theta = 1°$. Thereafter, those images were scaled ±20% with a scale step $\Delta h = 1\%$. Next, the three RST invariant 1D signatures of all those images were obtained. Finally, the instantaneous amplitude of the signatures were determined in the following form: lets R_k be the k-reference image in the database (e.g. using the B_P mask), from the instantaneous amplitude values of their rotated and scaled sample images for the red channel a 95.4% confidence interval (CI) was built using the statistical method of box-plots with $\mu_R \pm 2EE$, here μ_R represents the mean of those instantaneous amplitude values and EE the standard error. Analogously, the confidence intervals for B and G channels are set. Next, it is constructed a cuboid with edges being the confidence intervals $\mu_R \pm 2EE$, $\mu_G \pm 2EE$ and $\mu_B \pm 2EE$; and the vertices are set in the coordinates: $(\mu_R - 2EE, \mu_G - 2EE, \mu_B - 2EE)$, $(\mu_R + 2EE, \mu_G - 2EE, \mu_B - 2EE)$, $(\mu_R + 2EE, \mu_G + 2EE, \mu_B - 2EE)$, $(\mu_R - 2EE, \mu_G + 2EE, \mu_B - 2EE)$, $(\mu_R - 2EE, \mu_G - 2EE, \mu_B + 2EE)$, $(\mu_R + 2EE, \mu_G - 2EE, \mu_B + 2EE)$, $(\mu_R - 2EE, \mu_G + 2EE, \mu_B + 2EE)$ and $(\mu_R + 2EE, \mu_G + 2EE, \mu_B + 2EE)$.

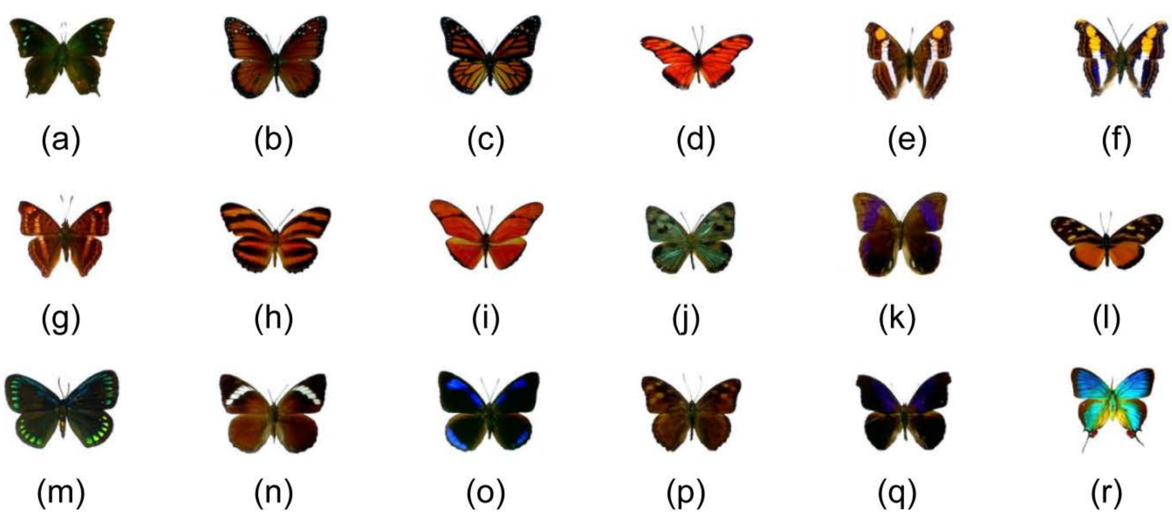

Fig.5. Butterflies database. (a) *Danaus plexippus plexippus*. (b) *Dione juno huascuma*. (c) *Doxocopa laure acca h*. (d) *Doxocopa laure laure m*. (e) *Doxocopa pavon m*. (f) *Dryadula phaetusa*. (g) *Dryas julia moderata*. (h) *Dynamine mylitta m*. (i) *Eryphanis aesacus*. (j) *Eueides procula asidia*. (k) *Eumaeus debora*. (l) *Eunica alcmena h*. (m) *Eunica alcmena m*. (n) *Eunica caresa h*. (o) *Eunica caresa m*. (p) *Evenus regalis*. (q) *Cymatogramma arginussa eubaena*. (r) *Danaus eresimus montezuma*.

Fig. 6 and Fig. 8 show the classifier cuboids space for the database in Fig. 5 using B_P and B_N masks, respectively. Hence, a volume space could be assigned to each image without overlapping (Fig. 7 shows an amplification zone of the output space to observe the cuboids assigned to some butterflies); in both cases, the RST invariant pattern recognition color image presents a confidence level at least of 95.4%. Contrary at correlator pattern recognition systems in [10] (where multiple correlation output planes are generated: one for each image in the database) here is established one classifier space, achieving in this form reduces the computational time investment.

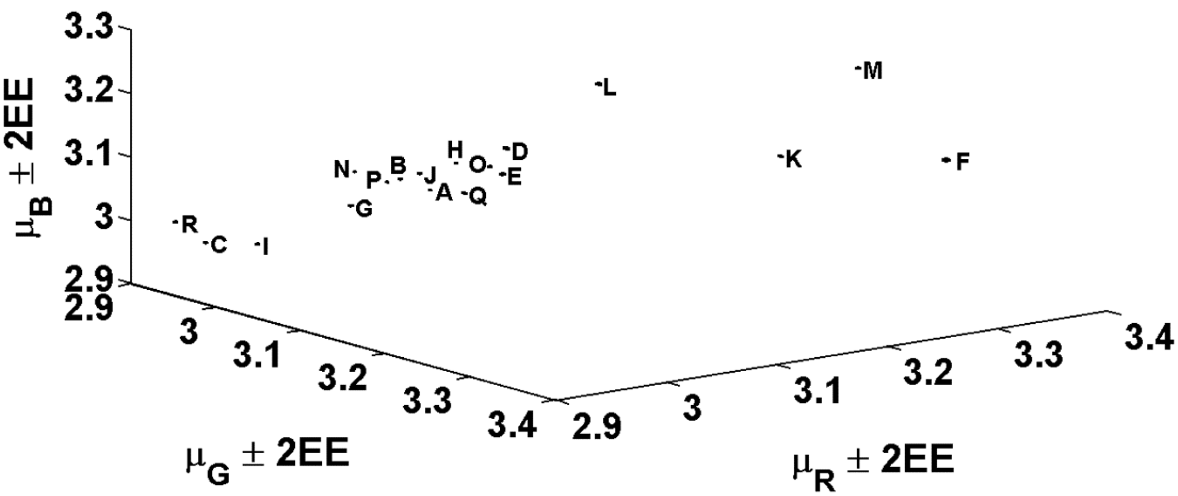

Fig.6. Classifier cuboids space using the Bessel mask B_P.

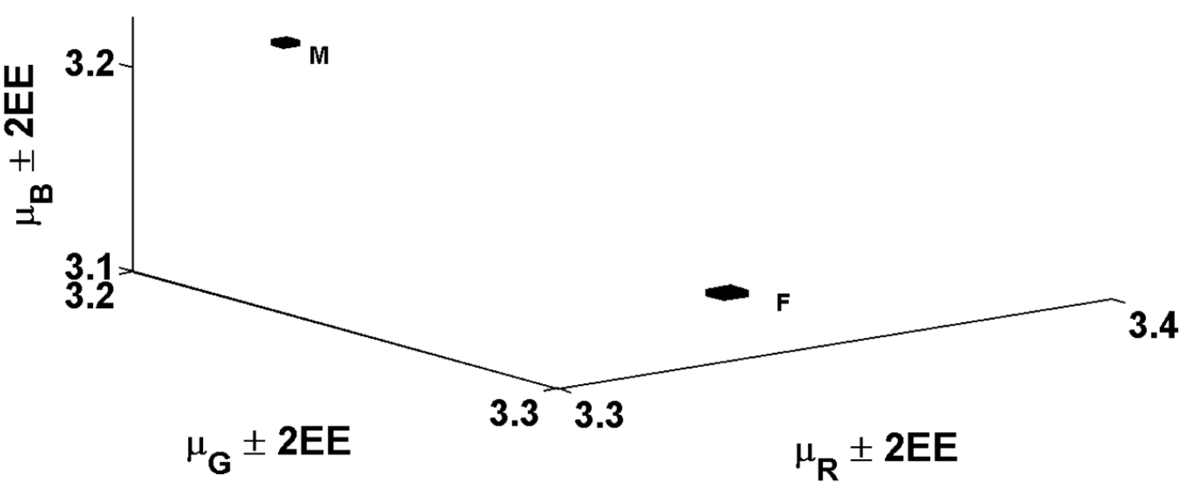

Fig.7. Amplification zone of the classifier cuboids space using the Bessel mask B_P.

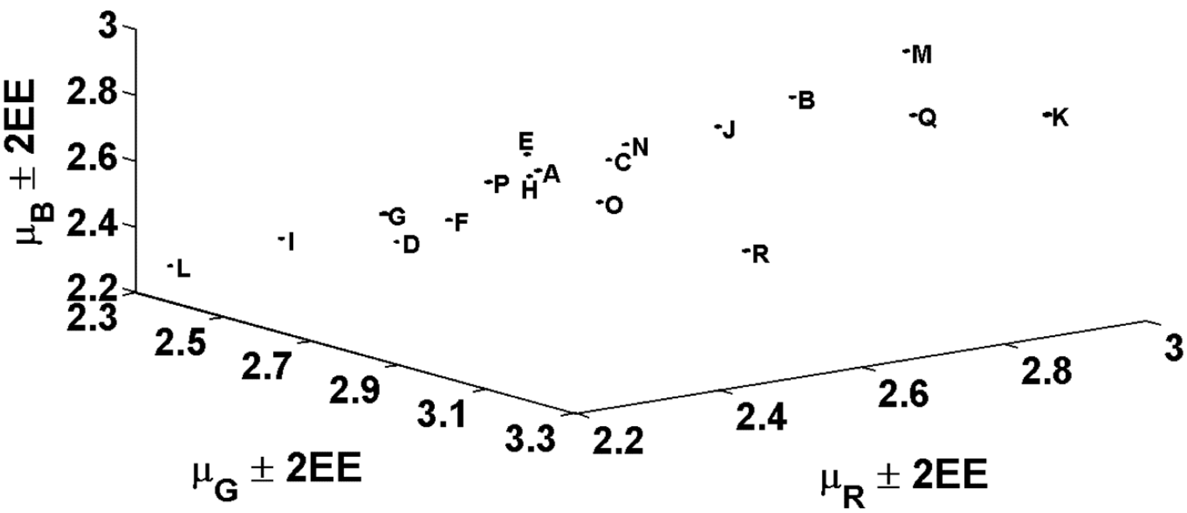

Fig.8. Classifier cuboids space using the Bessel mask B_N.

4. Conclusions

This work presents a new rotation, scale and translation invariant 1D signatures pattern recognition system specialized for color images. The RST system is based on Fourier transform, the analytic Fourier-Mellin transform and Bessel binary rings masks. The 1D RST invariant pattern recognition systems present confidence levels at least of 95.4% using the classifier cuboids space methodology. Moreover, the proposal of 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. Carolina Barajas-García is a student in the PhD program MyDCI offered by Universidad Autónoma de Baja California and she is supported by CONACyT's scholarship.