Rotation, Scale and Translation Invariant Pattern Recognition System for Color Images

Carolina Barajas-García*, Selene Solorza-Calderón* and Josué Álvarez-Borrego†

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

†Centro de Investigación Científica y Educación Superior de Ensenada, Departamento de Física Aplicada, Carretera Ensenada-Tijuana No.3918, Frac. Zona Playitas, C.P. 22860, Ensenada, B.C., México.

Abstract. This work presents a color image pattern recognition system invariant to rotation, scale and translation. The system works with three 1D signatures, one for each RGB color channel. The signatures are constructed based on Fourier transform, analytic Fourier-Mellin transform and Hilbert binary rings mask. According with the statistical theory of box-plots, the pattern recognition system has a confidence level at least of 95.4%.

Keywords: Image processing algorithms, pattern recognition algorithms, Hilbert masks, 1D signatures, digital image classification method.
PACS: 07.05.Pj; 42.30.Tz; 89.20.Ff.

INTRODUCTION

Color is very important feature to human pattern recognition process, if this information is discarded a very important characteristic will be lost. For example, in fishing industry color is very important to classify and count the harvested production of aquatic organisms for the proper exploitation of the marine resources. In [1] the color is used to study the Zostera marina leaf injury but the processing of the images is done by hand-operated trough multiple image processing programs: Adobe PhotoShop, Canon Photostitch and ERMapper. Therefore, color image pattern recognition systems will be a helpful option in the development of robots and automation systems specialized to preserve the natural environment and its resources.

This work presents a color image descriptor invariant to rotation, scale and translation (RST) based on binary rings masks methodologies [2, 3]. These methodologies are robust and efficient in the pattern recognition for gray-level images regardless the position, rotation and scale the object presents. Also, these pattern recognition systems have a great response under non-homogenous illumination and noises. This work proposes the use of the Hilbert binary rings mask in order to obtain a 1D RST invariant signature for each channel of the RGB color space. This methodology proposes a new single classifier output space of 95.4% confidence cuboids to reduce considerably the computational time investment in the classification step. The power of the signatures are used to get three confidence intervals, those will be used to built the confidence cuboids.

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

THE RST INVARIANT PATTERN RECOGNITION SYSTEM FOR COLOR IMAGES

The Hilbert binary rings masks

The Hilbert binary rings masks are obtained from the generalized radial Hilbert kernel given in [4]. It is defined by

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

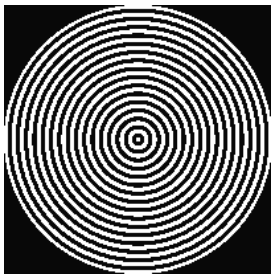

FIGURE 1. Hilbert binary rings mask.

$$H(x,y) = \Phi(\phi(x,y)), \tag{1}$$

where (x,y) is a pixel of the image and $\Phi(\phi(x,y))$ is any function. In this work it is used the specific form

$$H(x,y) = e^{iR\theta}, \tag{2}$$

with $R = \sqrt{(x-c_x)^2 + (y-c_y)^2}$, (c_x, c_y) is the central pixel of the image, $\theta = cos\big((x-c_x)/r^2\big)$, $r = \sqrt{x^2+y^2}$. Because $H(x,y)$ is a complex number, it yields the two images: $H_R(x,y)$ and $H_I(x,y)$, corresponding to the real and imaginary parts of Eq. (2), respectively. Hereafter, it will be working with $H_R(x,y)$ only; however the same procedure could be applied to $H_I(x,y)$, obtaining excellent results too. Based on $H_R(x,y)$, the binary filter $G(x,y)$ is built like

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

as it is shown in Fig. 1.

The RST invariant signatures

The RST invariant descriptor uses the amplitude spectrum of the Fourier transform of the given color image $I(x,y)$, because it is invariant to translation [5]. The RGB color space was selected in this work, hence three monochromatic images $I_q(x,y)$ are obtained from $I(x,y)$, where q represents the color channel R, G or B. The Fourier amplitude spectrum $A^{(q)}(u,v)$ is introduced in the fast analytical Fourier-Mellin transform (AFMT), just as

$$M^{(q)}(k,\omega) = \mathscr{F}\Big\{A^{(q)}(e^{\rho},\theta)e^{\rho\sigma}\Big\}, \tag{4}$$

with $\rho = ln(r)$, $\sigma > 0$ and $\mathscr{F}$ representing the Fourier transform in radial coordinates [6]. In this work $\sigma = 1/2$. Figure 2(c) displays the Fourier amplitude spectrum for the red channel image in Fig. 2(b). Figure 2(d) shows this amplitude spectrum in log-polar coordinates weighted by $e^{\rho\sigma}$. Equation (4) is not invariant to scale and rotation yet,

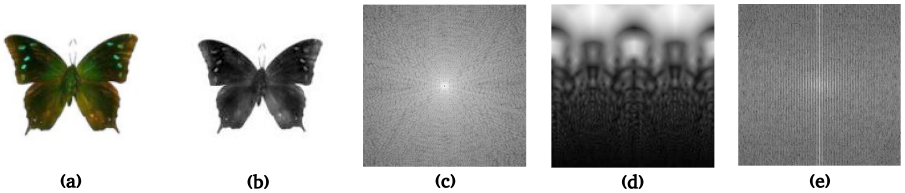

FIGURE 2. The analytic Fourier-Mellin transform example. (a) The color image $I(x,y)$. (b) The red channel image, $I_R(x,y)$. (c) The Fourier amplitude spectrum $A^{(R)}(u,v)$. (d) $A^{(R)}(e^{\rho},\theta)e^{\rho\sigma}$, with $\sigma = 1/2$. (e) $A_M^{(R)}(k,\omega)$.

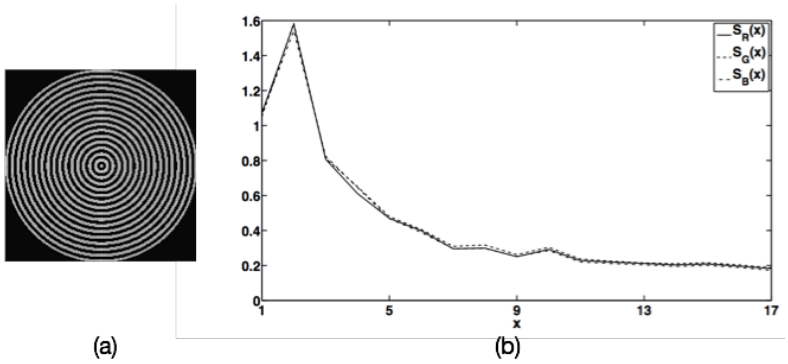

FIGURE 3. (a) $H^{(R)}$. (b) The three 1D RST signatures of the color image in Figure 2(a).

but normalizing the AFMT by its value in the central pixel (c_u, c_v), the amplitude spectrum is invariant to scale [6], that is

$$A_M^{(q)}(k,\omega) = \left| \frac{M^{(q)}(k,\omega)}{M^{(q)}(c_u,c_v)} \right|. \tag{5}$$

Figure 2(e) shows the normalized AFMT amplitude spectrum of the image in Fig. 2(b). To achieve the invariant to rotation, the image $A_M^{(q)}$ is filtered by the binary rings mask G like

$$H^{(q)} = A_M^{(q)} \odot G, \tag{6}$$

here $\odot$ indicates the element-wise product or Hadamard multiplication. Figure 3(a) displays the normalized AFMT amplitude spectrum of Fig. 2(e) filtered with the mask in Fig. 1. The rings in $H^{(q)}$ are numbered from the center towards out-side and the intensity values in each ring are added to get the 1D signature S_q. In Fig. 3(b) is shown the RST signatures of Fig. 2(a). In this image, the color contribution is balanced and it is observed in the similarity of the signatures in Fig. 3(b). On the other hand, when one of the colours is dominant the signatures will be quite different. The features assigned to color image are the power of the 1D RST signatures, which are computed by

$$P_q = \frac{\sum (S_q)^2}{N}, \tag{7}$$

where N is the length of the signature.

THE OUTPUT SPACE

To train the RST invariant pattern recognition system, each image in the reference image database (Fig. 4) was rotated 360° using $\Delta\theta = 1°$. Thereafter, those images were scaled $\pm 15\%$ with a scale step $\Delta h = 1\%$. Hence 11,160 samples are generated from each image in Fig. 4. Next, the three RST invariant 1D signatures of each one of all those images

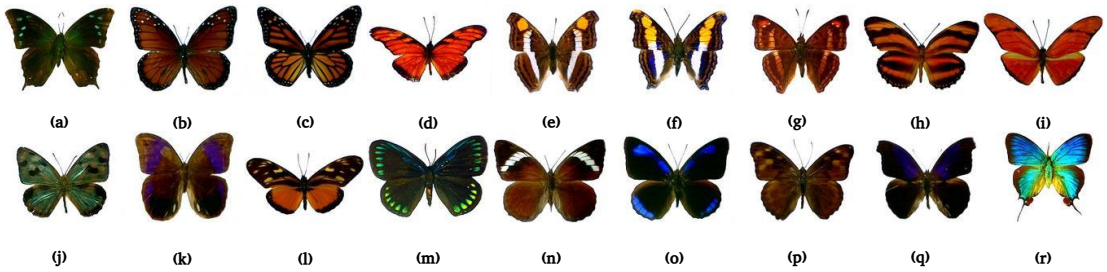

FIGURE 4. Butterflies database.

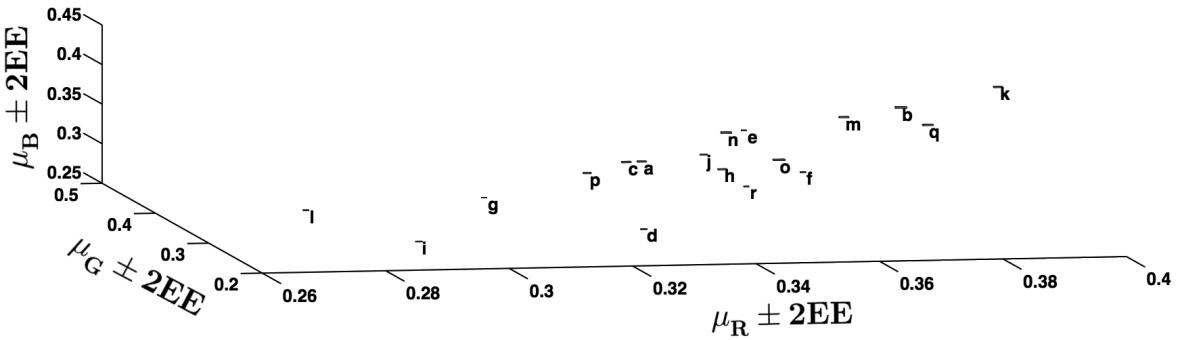

FIGURE 5. Cuboids classifier output space.

were obtained. Once the power values of those signatures were determined, the confidence intervals are built in the following form: lets R_k be the k-th reference image in the database, using the 11,160 power values of the red channel, the 95.4% confidence interval (CI) was built using the statistical method of box-plots with $\mu_R \pm 2EE$, here μ_R is the mean of the power values and EE is the standard error. Analogously, the confidence intervals for B and G channels are obtained. Next, it is get a cuboid with edges being the confidence intervals $\mu_R \pm 2EE$, $\mu_G \pm 2EE$, $\mu_B \pm 2EE$; 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 $(\mu_R + 2EE, \mu_G + 2EE, \mu_B + 2EE)$. Figure 5 shows the cuboids classifier output space for the database in Fig. 4. A volume space was assigned to each image without overlapping, hence the RST invariant pattern recognition color image presents a confidence level at least of 95.4%.

CONCLUSIONS

This work presents a new 1D signatures pattern recognition system invariant to rotation, scale and translation specialized for color images. The RST pattern recognition system is based on the Fourier transform, the analytic Fourier-Mellin transform and Hilbert binary rings masks. 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. 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.

REFERENCES

1. B. L. Boese, P. J. Clinton, D. Dennis, R. C. Golden, and B. Kim, *Aquatic Botany* **88**, 87–90 (2008).
2. S. Solorza, and J. Álvarez-Borrego, *Journal of Modern Optics* **62**, 851–864 (2015).
3. C. Barajas-García, S. Solorza-Calderón, and J. Álvarez-Borrego, "Pattern recognition digital systems using the Z-Fisher transform," in *Proceedings of SPIE: Applications of Digital Image Processing XXXVIII*, edited by A. G. Tescher, SPIE, SPIE, 2015, vol. 9599, pp. 95992M–1–8.
4. S. C. Pei, and J. J. Ding, "The generalized radial Hilbert transform and its applications to 2D edge detection (any direction or specified directions)," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, IEEE, 2003, vol. 3, pp. III–357–360.
5. R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using MATLAB*, Pearson Education India, 2004, second edn.
6. S. Derrode, and F. Ghorbel, *Computer Vision and Image Understanding* **83**, 57–78 (2001).