A RFM Pattern Recognition System Invariant to Rotation, Scale and Translation

Selene Solorza-Calderón(✉) and Jonathan Verdugo-Olachea

Facultad de Ciencias, Universidad Autónoma de Baja California,
Km. 103, Carretera Tijuana-Ensenada, 22860 Ensenada, B.C., México
selene.solorza@uabc.edu.mx

Abstract. In this paper a rotation, scale and translation (RST) invariant pattern recognition digital system based on 1D signatures is proposed. The rotation invariance is obtained using the Radon transform, the scale invariance is achieved by the analytical Fourier-Mellin transform and the translation invariance is realized through the Fourier's amplitude spectrum of the image. Once, the RST invariant Radon-Fourier-Mellin (RFM) image is generated (a 2D RST invariant), the marginal frequencies of that image are used to build a RST invariant 1D signature. The Latin alphabet letters in Arial font style were used to test the system. According with the statistical method of bootstrap the pattern recognition system yields a confidence level at least of 95%.

Keywords: Pattern recognition · Radon-Fourier-Mellin Images · 1D RST invariant signature · Radon transform · Analytical Fourier-Mellin transform

1 Introduction

In the pattern recognition field, 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. Since the first optical experiments in the middle of last century, the features extraction has been a subject of interest and a great progress were done since the introduction of the classical joint transform correlator by Vander Lugt [1], that is the classical matched filter (CMF). Due to the fact that the CMF filter has low response to additive Gaussian noise other filters were generated, just as the phase-only filter (POF), the synthetic discriminant function filter (SDF) and others. In general, the filters are specialized to solve specific problems, for example the filter could have an excellent performance in the discrimination step and the signal-to-noise ratio but low efficiency under non-homogeneous illumination [2]. Although composite filters are being used, the RST invariant image classification problem is an active field due to its intrinsic complexity[3–7].

Actually, with the great advance in technology, the pattern recognition via digital images is a very productive area. A lot of methodologies in digital images

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 477–484, 2015.
DOI: 10.1007/978-3-319-25751-8_57

features extraction based on joint transforms correlators are developed. T.V. Hoang and S. Tabbone [8] uses the Radon and the 1D Fourier-Mellin transform to build a 2D RST invariant classifier. However, the classification step is realized by the use of the 2D cross-correlation of the target and the problem images. On the other hand, the 2D Fourier-Mellin transform (FMT) are used 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 using the centroid or the center of mass of the objects, but removing a small disk around the centroid or the center of mass to reduce the large effect of the factor $\frac{1}{r}$ [9]. Ghorbel [10] propose the analytical Fourier-Mellin transform (AFMT), where the images are weighted by the factor r^{σ} with $\sigma > 0$ to eliminate the influence of $\frac{1}{r}$ in the FMT. However, this transform not preserves the rotation and scale invariance. Derrode [9] propose a normalization of the AFMT by two of the AFMT harmonics to obtain a RST invariant descriptor together with the Euclidean distance for the classification mechanism.

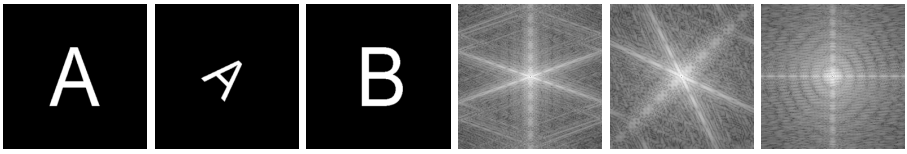
In this work a 1D RST invariant Radon-Fourier-Mellin (RFM) digital image pattern recognition system is designed. Moreover, a methodology to generate one and only one classifier output plane is proposed, instead of the multiple classifier output planes obtained with the correlator pattern recognition systems [5–8]. The work is organized as follows: Section 2 explains the mathematical foundations of the RST invariant images and the methodology to obtain the 1D signature. Section 3 exposes the procedure to construct the classifier output plane of 95% confidence level. Finally, conclusions are given in section 5.

2 Digital System Invariant to Rotation, Scale and Translation

In the Radon-Fourier-Mellin (RFM) digital image pattern recognition, the first step is obtain the shift invariance. This is achieved using the amplitude spectrum $A(u, v)$ of the Fourier transform [12]. Fig. 1 shows the Fourier's amplitude spectrums for black and white 257×257 pixel images of: the image with the A Arial font letter without geometric transformations, called I_1; the image with the A Arial font letter with a rotation angle of 315° and scaling of −25%, named I_2; the image with the B Arial font letter without geometric transformations, denominated I_3. Also, Fig. 1 shows that $A_3(x, y) = |\mathcal{F}\{I_3(x, y)\}|$ is different of $A_1(x, y) = |\mathcal{F}\{I_1(x, y)\}|$ and $A_2(x, y) = |\mathcal{F}\{I_2(x, y)\}|$. Moreover, in Fig. 1 is seen that A_2 presents the same rotation angle of I_2 and it has a stretch deformation due to the scale variation.

The next step of the RMF system is the scale invariance, which is given via the fast analytical Fourier-Mellin transform (AFMT),

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



(a) (b) (c) (d) (e) (f)

Fig. 1. Fourier's 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) $A_3(u,v) = |\mathcal{F}\{I_3(x,y)\}|$.

where $\rho = \ln(r)$ and $\sigma > 0$. Eq. (1) is not an invariant to scale and rotation, but normalizing the AFMT by its *dc*-value the amplitude spectrum is a scale invariance [9], that is

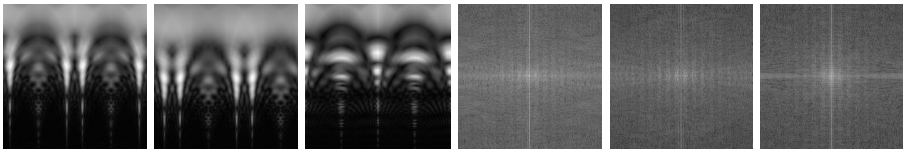
$$S(k,\omega) = \left| \frac{M(k,\omega)}{M(c_x,c_y)} \right| , \tag{2}$$

where (c_x, c_y) is the central pixel of the image. Fig. 2(d), Fig. 2(e) and Fig. 2(f) present $S_1(k,\omega)$, $S_2(k,\omega)$ and $S_3(k,\omega)$ images, respectively. These are the normalized analytical Fourier-Mellin amplitude spectrums of $A_1(e^\rho,\theta)e^{\rho\sigma}$, $A_2(e^\rho,\theta)e^{\rho\sigma}$ and $A_3(e^\rho,\theta)e^{\rho\sigma}$ (Fig. 2(a), Fig. 2(b) and Fig. 2(c)), respectively. The images are not rotation invariant yet. Fig. 2(a) and Fig. 2(b) show the circular shift in the angular variable.

Finally, the rotation invariant is obtained by the Radon transform [8] of the normalized analytical Fourier-Mellin amplitude spectrum $S(k,\omega)$, that is

$$R(r,\theta) = \mathcal{R}\{S(k,\omega)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(k,\omega)\delta(r - kcos\theta - \omega sin\theta)dkd\omega , \tag{3}$$

where $r \in (-\infty,\infty)$, $\theta \in [0,\pi)$ and δ is the Dirac delta function. Fig. 3(a), Fig. 3(b) and Fig. 3(c) show the RFM images invariant to rotation, scale and translation associated to Fig. 1(a), Fig. 1(b) and Fig. 1(c), respectively. Practically,



(a) (b) (c) (d) (e) (f)

Fig. 2. Normalized analytical Fourier-Mellin spectrum with $\sigma = 0.5$. (a) $A_1(e^\rho,\theta)e^{\rho\sigma}$. (b) $A_2(e^\rho,\theta)e^{\rho\sigma}$. (c) $A_3(e^\rho,\theta)e^{\rho\sigma}$. (d) The $S_1(k,\omega)$ of Fig. 2(a). (e) The $S_2(k,\omega)$ of Fig. 2(b). (f) The $S_3(k,\omega)$ of Fig. 2(c).

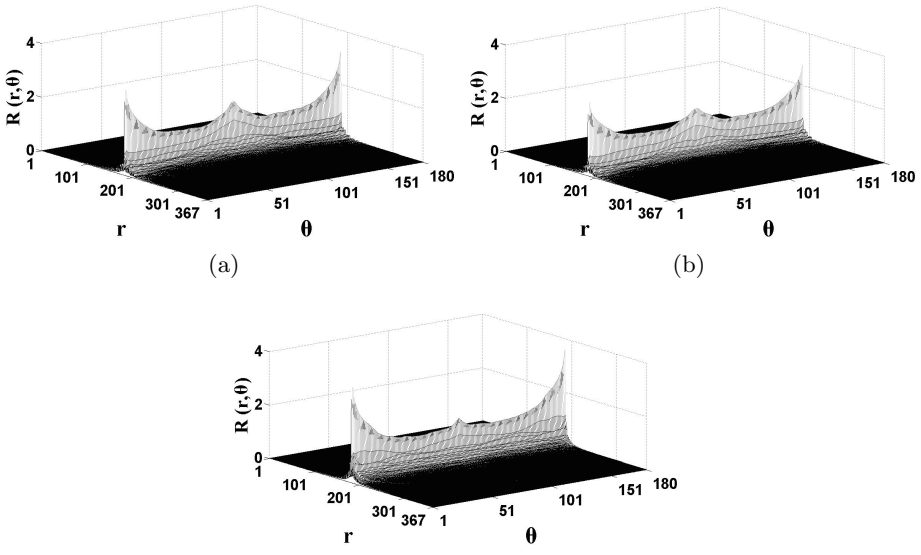


Fig. 3. The Radon transform examples. (a) The $R(r, \theta)$ of Fig. 2(d). (b) The $R(r, \theta)$ of Fig. 2(e). (c) The $R(r, \theta)$ of Fig. 2(f).

Fig. 3(a) and Fig. 3(b) are equal, the former is obtained from the image with the A Arial font letter without geometric distortions and the latter is generated with the A Arial font letter rotated and scaled. Fig. 3(c) is different to the others two images, this is associated to the B Arial font letter without distortions.

To reduce the computational time cost, the marginal frequencies are used in the 1D RST invariant signature construction, that is

$$V(x) = \sum_{y} R(x, y) \ . \tag{4}$$

3 The Confidence Level

The RST invariant Radon-Fourier-Mellin pattern recognition system was training using black and white (BW) 257×257 pixel digital images with the Latin alphabet letters in Arial font style, each image was rotated 360° with $\Delta\theta = 1°$. Thereafter, each of those images were scaled $\pm 25\%$ with a scale step size of $\Delta k = \pm 1\%$.

The real and imaginary parts of the Fourier transform of the 1D signature are obtained, just like

$$R_V(u) + iI_V(u) = F(u) = \mathcal{F}\{V(x)\} = \int_{-\infty}^{\infty} V(x) e^{-i2\pi ux} dx \ , \tag{5}$$

to determine the signature's power by

$$P_{R} = \frac{1}{N_{R}} \sum R_{V}^{2} \ , \tag{6}$$

$$P_{I} = \frac{1}{N_{I}} \sum I_{V}^{2} \ , \tag{7}$$

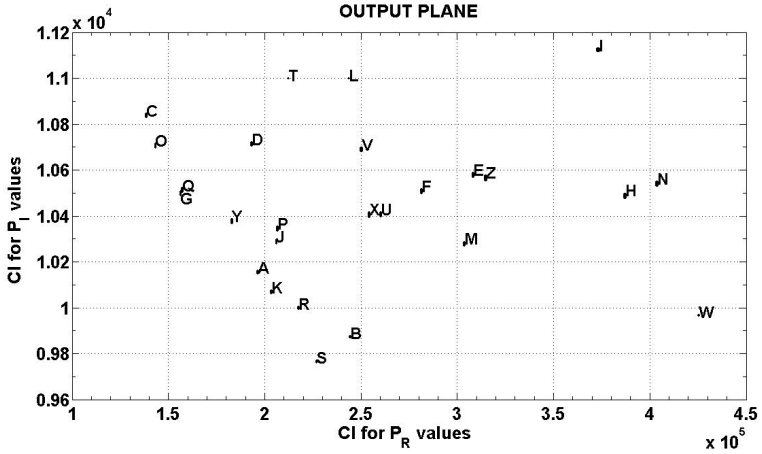
where N_{R} and N_{I} are the length of the signatures R_{V} and I_{V}, respectively. A database of target images should be established to train the RFM pattern recognition system. The P_{R} and P_{I} of those images are obtained to determine the 95% confidence interval (CI) by the statistical method of bootstrap using a replacement constant $B = 1,000$ and normal distribution [11]. Fig. 4(a) shows the output plane for the 26 Arial font letter of the Latin alphabet. Each letter image was rotated 360° ($\Delta\theta = 1°$) and then scaled ±25% ($\Delta k = \pm 1\%$), therefore 18,360 images for each letter were created. Then, each CI is built from 18,360 values. In Fig. 4(a), the horizontal and vertical axes represent the CI for the P_{R} and P_{I} values, respectively. A rectangle area is assigned to each image (Fig. 4(b) displays an amplification zone of the output plane to observe the rectangle area assigned to some letters). Because those rectangles are not overlapped, the RFM pattern recognition system has a confidence level at least of 95% in the digital image classification. Therefore, a one and only one classifier output plane was used, instead of the multiple classifier output planes (one for each reference image in the database) for correlator systems[6–8] or distance systems [5,9]. For example, in the case of the latin alphabet letters the correlator systems [6,7] uses 26 output planes. In [8], the classification step is realized by the use of the 2D cross-correlation of the target and the problem images. 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. Therefore, the single output plane methodology reduces the investment computational time considerably.

4 Noise Analysis

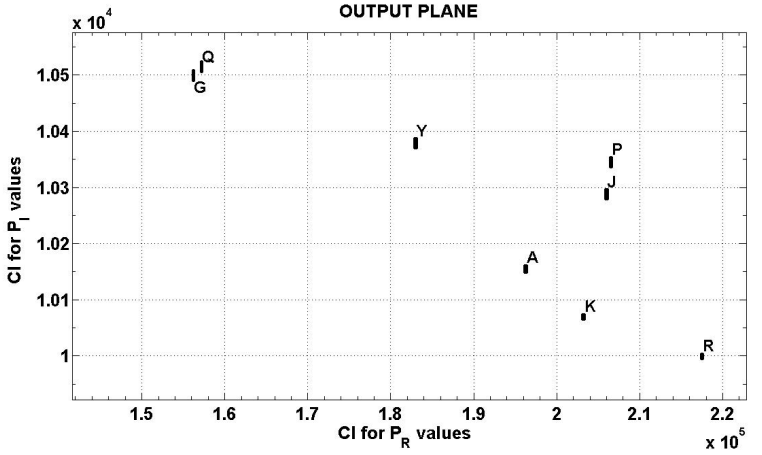
To test the performance of the system when images have additive Gaussian noise, the similarity coefficient was used, it is defined like

$$SC = 1 - \frac{||S_{T} - S_{TN}||_{\infty}}{||S_{T} - S_{FN}||_{\infty}} \ , \tag{8}$$

where $||\mathbf{x}||_{\infty} = max\{|x_1|, |x_2|, \ldots, |x_n|\}$, S_{T} is the signature of the image, S_{TN} is the signature of the image with noise and S_{FN} is the signature of the background image with noise. When S_{T} and S_{TN} are similar the $||S_{T} - S_{TN}||_{\infty} \to 0$, then $SC \approx 1$. On the other hand, when the problem image has to much noise that it looks like the background image with noise, the S_{T} and S_{FN} are similar and $||S_{T} - S_{FN}||_{\infty} \to 0$, thus $SC < 0$. For the sake of comparison, the performance of



(a)



(b)

Fig. 4. (a) The classifier output plane. (b) Amplification zone of the classifier output plane.

SURF methodology when the images have noise is included, but here the results are given in terms of the repeatability parameter r,

$$r = \frac{C\,(T, PI)}{mean(N_T, N_{PI})}\ , \tag{9}$$

where $C(T, PI)$ represents the number of the common detected points in the target T and the problem image PI; N_T and N_{PI} are the number of points detected in T and PI, respectively. Fig. 5 presents the graphs of the mean of the SC response for the RFM system and the repeatability analysis (r values) for the SURF algorithm. The images were altered with additive Gaussian noise of media zero and variance from zero to 0.3, using 40 images per sample. In Fig. 5 is shown that RFM system has a better response under this kind of noise than the SURF methodology.

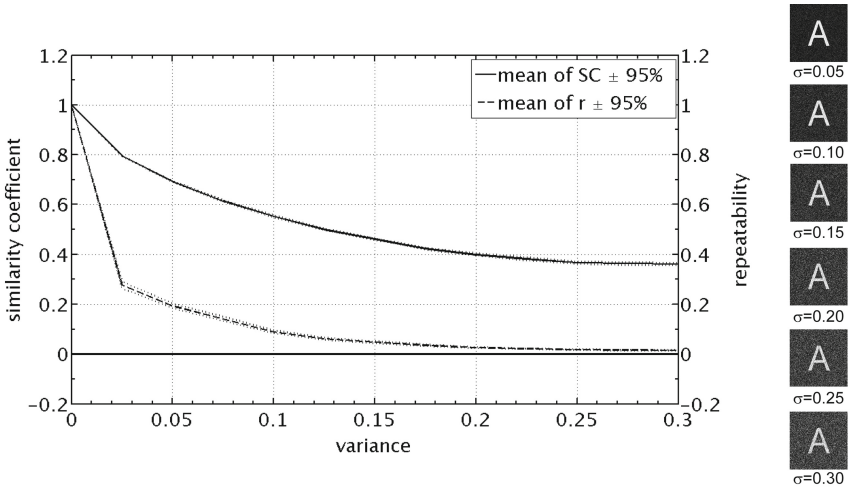


Fig. 5. The RFM and SURF pattern recognition systems performance when images have additive Gaussian noise with variance σ from 0 to 0.3 and step size of $\Delta\sigma = 0.025$.

5 Conclusions

This work presents a RFM pattern recognition system using a 1D signature invariant to rotation, scale and translation (RST) based on the analytical Fourier-Mellin and Radon transforms. The system presents a confidence level at least of 95% in the pattern recognition of rotated, scaled and translated black and white images with the Latin alphabet letters in Arial font style. Moreover, this RST invariant Radon-Fourier-Mellin methodology generates a single classifier output plane to reduce the computational cost time of the classification procedure.

Acknowledgments. This work was partially supported by CONACyT under grant No. 169174. Jonathan Verdugo-Olachea is a student in the Math program offered by the Facultad de Ciencias, Universidad Autónoma de Baja California and he is supported by a CONACyT scholarship under the research proyect No. 169174.

References

1. Van der Lugt, A.: Signal detection by complex spatial filtering. IEEE Trans. Inf. Theory **IT–10**, 139–145 (1964)
2. Vijaya Kumar, B.V.K., Hassebrook, L.: Performance measures for correlation filters. Appl. Opt. **29**, 2997–3006 (1990)
3. Lowe, D.G.: Distinctive image features from scale-invariant key points. IJCV **60**, 91–110 (2004)
4. Bay, H., Essa, A., Tuytelaars, T., Van Gool, L.: Speeded-Up Robust Features (SURF). CVIU **110**, 346–359 (2008)
5. Lerma-Aragón, J.L., Álvarez-Borrego, J.: Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition. J. Mod. Opt. **56**, 1598–1606 (2009)
6. Solorza, S., Álvarez-Borrego, J.: Translation and rotation invariant pattern recognition by binary rings masks. J. Mod. Opt. **62**, 851–864 (2015)
7. Solís-Ventura, A., Álvarez-Borrego, J., Solorza, S.: An adaptive nonlinear correlation with a binary mask invariant to rotation and scale applied to identify phytoplankton. Opt. Commun. **339**, 185–193 (2015)
8. Hoang, T.V., Tabbone, S.: Invariant pattern recognition using the RFM descriptor. Pattern Recogn. **45**, 271–284 (2012)
9. Derrode, S., Ghorbel, F.: Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description. CVIU **83**, 57–78 (2001)
10. Ghorbel, F.: Towards a unitary formulation for invariant image description; Application to Towards a unitary formualtion for invariant image description. Application to image coding. Ann. Telecommun. **53**, 242–260 (1998)
11. Davison, A.C., Hinkley, D.V.: Bootstrap methods and their application. Cambridge University Press, New York (1997)
12. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital image processing using Matlab. Gatesmark Publishing, MA (2009)