

A geometric invariant shape descriptor based on the Radon, Fourier, and Mellin transforms

Thai V. Hoang, Salvatore Tabbone

► **To cite this version:**

Thai V. Hoang, Salvatore Tabbone. A geometric invariant shape descriptor based on the Radon, Fourier, and Mellin transforms. International Conference on Pattern Recognition - ICPR'2010, Aug 2010, Istanbul, Turkey. IEEE, pp.2085-2088, 2010, ICPR 2010 - 20th International Conference on Pattern Recognition. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5597270>. <10.1109/ICPR.2010.512>. <inria-00512312>

HAL Id: inria-00512312

<https://hal.inria.fr/inria-00512312>

Submitted on 13 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Geometric Invariant Shape Descriptor Based on the Radon, Fourier, and Mellin Transforms

Thai V. Hoang^{1,2}, Salvatore Tabbone²

¹MICA, UMI 2954, Hanoi University of Technology, Hanoi, Vietnam

²LORIA, UMR 7503, Université Nancy 2, 54506 Vandœuvre-lès-Nancy, France

Abstract—A new shape descriptor invariant to geometric transformation based on the Radon, Fourier, and Mellin transforms is proposed. The Radon transform converts the geometric transformation applied on a shape image into transformation in the columns and rows of the Radon image. Invariances to translation, rotation, and scaling are obtained by applying 1D Fourier-Mellin and Fourier transforms on the columns and rows of the shape’s Radon image respectively. Experimental results on different datasets show the usefulness of the proposed shape descriptor.

Keywords—Geometric invariant regions-based shape descriptor, Radon transform, Fourier transform, Mellin transform

I. INTRODUCTION

Shape description is an important research topic in pattern recognition as it is one of the basic features used to describe image contents. Describing shape means looking for effective and perceptually important shape features based on its boundary information and/or internal structure. A desirable shape descriptor should have the following properties: good retrieval accuracy, geometric invariances (translation, rotation, scaling), application independence, low computation complexity, etc.

Many descriptors have been proposed [1] for describing shape and they can be roughly classified into two main classes: *contour-based* descriptors and *region-based* descriptors. Contour-based shape descriptors are extracted from the shape contour exploiting its boundary information. Important extraction techniques include 1D Fourier transform [2], curvature scale-space [3], 2D histogram of neighboring contour pixels [4], inner-distance in the shape silhouette [5], and rotation invariant kernel [6]. In spite of their popularity, contour-based shape descriptors are applicable only to certain kinds of application due to several limitations. Firstly, they are generally sensitive to noise that exists in shape contours. Secondly, they cannot capture the internal structure of a shape. Thirdly, they are not suitable for disjoint shapes or shapes with holes inside.

The limitations of contour-based shape descriptors can be overcome by region-based shape descriptors which are extracted from the whole shape region. Common extraction techniques are based on the theory of moments [7], 2D Fourier-Mellin transform [8], 2D Fourier transform [9]. Although region-based shape descriptors are more suitable

for general applications, they are more computationally intensive and most methods need normalization steps (centroid position, re-sampling, re-quantization) in order to achieve common geometric invariances. These normalizations introduce errors, are sensitive to noise, and thus induce inaccuracy in the later recognition/matching process.

Shape descriptors defined on the Radon transform are region-based shape descriptors. Geometric transformation parameters are encoded in the columns (for translation and scaling) and rows (for rotation) of the Radon image. Current techniques thus usually exploit this encoded information to define invariant descriptors. Notable work in this direction is the \mathcal{R} -signature proposed in [10]. This approach uses an integral function for the columns and the 1D Fourier transform for the rows of the Radon image respectively to get an 1D signature of the shape image that is invariant to translation, rotation, and scaling. However, even if an extension to 2D signature has been proposed, the obtained signature has low discriminatory power as there is a loss of information in the compression process from the Radon image to the 1D signature. Recently, there was an effort to apply the 2D Fourier-Mellin transform on the Radon image [11]. Similarly, Mellin and Fourier transforms are applied on the columns and rows of the Radon image respectively to get a shape descriptor that is invariant to scaling and rotation. The main weakness of this approach is the lack of translation invariance. Moving the origin to the shape’s centroid is a common solution to have translation invariance. However, this normalization step may introduce errors.

This paper presents a new region-based geometric invariant shape descriptor, called the \mathcal{RFM} shape descriptor, based on the Radon, Fourier, and Mellin transforms that is invariant not only to rotation, and scaling but also to translation. Geometric invariances are obtained by applying 1D Fourier-Mellin and Fourier transforms on the columns and rows of the shape’s Radon image respectively. Experimental results on different databases show the usefulness of the proposed \mathcal{RFM} shape descriptor.

The remainder of this paper is organized as follows. Section II gives some background on the Radon, Fourier, and Mellin transforms. The proposed \mathcal{RFM} shape descriptor is defined in Section III. Experimental results are given in Section IV, and finally conclusions are drawn in Section V.

II. BASIC MATERIAL

A. The Radon transform

Let $f(x, y) \in \mathbb{R}^2$ be a two-dimensional function, $L(\theta, \rho)$ be a straight line in \mathbb{R}^2 represented by:

$$L = \{(x, y) \in \mathbb{R}^2 : x \cos \theta + y \sin \theta = \rho\}, \quad (1)$$

where θ is the angle L makes with the y axis and ρ is the distance from the origin to L . Concretely, any straight line L can be parameterized by a parameter t as follows:

$$(x(t), y(t)) = t(\sin \theta, -\cos \theta) + \rho(\cos \theta, \sin \theta). \quad (2)$$

The Radon transform [12] of f , denoted by \mathcal{R}_f , is a function defined on the space of lines L by the line integral along each line:

$$\mathcal{R}_f(L) = \mathcal{R}_f(\theta, \rho) = \int_{-\infty}^{\infty} f(x(t), y(t)) dt. \quad (3)$$

The Radon transform has some useful properties on translation, rotation, and scaling as outlined below:

- *P1*: A translation of f by a vector $\vec{u} = (x_0, y_0)$ results in a shift of its transform in the variable ρ by a distance $d = x_0 \cos \theta + y_0 \sin \theta$ equal to the projection of \vec{u} on the line $x \cos \theta + y \sin \theta = \rho$.
- *P2*: A rotation of the image by an angle θ_0 implies a shift θ_0 of the transform in the variable θ .
- *P3*: A scaling of f by a factor α results in a scaling of the ρ coordinate and the amplitude by a factor α and $\frac{1}{\alpha}$ of the transform respectively.

B. The 1D Fourier-Mellin transform for one dimensional signal

Consider the Fourier transform of a function $g(x) = f(\alpha x - x_0)$, a scaled and translated version of $f(x)$ (α is supposed to be positive):

$$\mathcal{F}_g(\xi) = \int_{-\infty}^{\infty} f(\alpha x - x_0) e^{-i2\pi\xi x} dx, \quad (4)$$

where ξ is a real number. Letting $y = \alpha x - x_0$, then:

$$\begin{aligned} \mathcal{F}_g(\xi) &= \frac{1}{\alpha} e^{-i2\pi\frac{\xi}{\alpha}x_0} \int_{-\infty}^{\infty} f(y) e^{-i2\pi\frac{\xi}{\alpha}y} dy \\ &= \frac{1}{\alpha} e^{-i2\pi\frac{\xi}{\alpha}x_0} \mathcal{F}_f\left(\frac{\xi}{\alpha}\right). \end{aligned} \quad (5)$$

$$|\mathcal{F}_g(\xi)| = \frac{1}{\alpha} \left| \mathcal{F}_f\left(\frac{\xi}{\alpha}\right) \right|. \quad (6)$$

The translation parameter x_0 has disappeared in Eq. (6), this agrees with the shift or translation property of the Fourier transform. The remaining scaling parameter α could be removed by using the Mellin transform [13] \mathcal{M}_f of a function f :

$$\mathcal{M}_f(s) = \int_0^{\infty} f(x) x^{s-1} dx. \quad (7)$$

Thus, the Mellin transform of $|\mathcal{F}_g(\xi)|$ is:

$$\begin{aligned} \mathcal{M}_{|\mathcal{F}_g|}(s) &= \int_0^{\infty} \frac{1}{\alpha} \left| \mathcal{F}_f\left(\frac{\xi}{\alpha}\right) \right| \xi^{s-1} d\xi \\ &= \alpha^{s-1} \mathcal{M}_{|\mathcal{F}_f|}(s). \end{aligned} \quad (8)$$

where $s = \sigma + i\tau$ (σ is a constant chosen such that the integral in Eq. (8) converges and τ is the transform variable). Taking the absolute magnitude of the two sides of Eq. (8) results in:

$$|\mathcal{M}_{|\mathcal{F}_g|}(s)| = \alpha^{\sigma-1} |\mathcal{M}_{|\mathcal{F}_f|}(s)|. \quad (9)$$

Defining the above calculating steps from Eq. (4) to Eq. (9) as the 1D Fourier-Mellin transform of the function $g(x)$, denoted by $\mathcal{MF}_g(s)$. Thus, this 1D Fourier-Mellin transform is, except for a constant multiplicative factor $\alpha^{\sigma-1}$, independent of the translation and scaling parameters x_0 and α of f .

C. Mellin transform implementation

Although the Mellin transform, as defined in Eq. (7), has a very attractive property of scaling invariance, there are reported problems with its implementation [8] and its use with Fourier transform for feature extraction [14]. The first problem comes from the FFT-based implementation of the Mellin transform which requires exponential sampling at $x = 0$. The second problem is the obscurity of the discriminatory information in the input function by the 1D Fourier-Mellin transform.

To avoid these problems, an alternative to the Mellin transform proposed in [15], which is called the direct Mellin transform, is adopted for this work. Assuming $f(x)$ is in the form of sampled data with sampling period T (the value of $f(x)$ is assumed to be piecewise constant), expanding Eq. (7) gives:

$$\begin{aligned} \mathcal{M}_f(s) &= \int_0^T f(x) x^{s-1} dx + \int_T^{2T} f(x) x^{s-1} dx \\ &\quad + \dots + \int_{(N-1)T}^{NT} f(x) x^{s-1} dx. \end{aligned} \quad (10)$$

Denoting $f(iT) = f_{i+1}$ and without loss of generality assuming $T = 1$ and $f_N = 0$, Eq. (10) becomes:

$$\begin{aligned} s\mathcal{M}_f(s) &= f_1 x^s \Big|_0^1 + f_2 x^s \Big|_1^2 + \dots + f_N x^s \Big|_{N-1}^N \\ &= \sum_{k=1}^{N-1} k^s (f_k - f_{k+1}). \end{aligned} \quad (11)$$

The direct Mellin transform, as defined in Eq. (11), is an exact implementation of the Mellin transform for sampled data. It can be proven to maintains the scaling invariance property of the Mellin transform.

III. THE PROPOSED \mathcal{RFM} SHAPE DESCRIPTOR

Let I_2 be the shape image obtained after scaling, rotating, and translating an shape image I_1 using transformation parameters α , θ_0 , and $\vec{u} = (x_0, y_0)$. Properties *P1*–*P3* of the

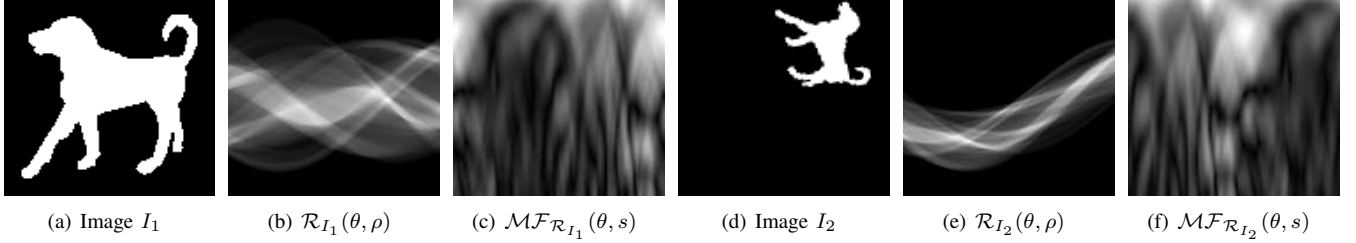


Figure 1. Radon and 1D Fourier-Mellin transforms performed on shape images. The image I_2 in (d) is a scaled, rotated, and translated version of the image I_1 in (a). Correspondingly, $\mathcal{MF}_{\mathcal{R}_{I_2}}(\theta, s)$ in (f) is a horizontally shifted version of $\mathcal{MF}_{\mathcal{R}_{I_1}}(\theta, s)$ in (c).

Radon transform imply $\mathcal{R}_{I_2}(\theta, \rho) = \frac{1}{\alpha} \mathcal{R}_{I_1}(\theta + \theta_0, \alpha\rho - d)$, where $d = x_0 \cos(\theta + \theta_0) + y_0 \sin(\theta + \theta_0)$. It is clear that, except for a constant multiplicative factor $\frac{1}{\alpha}$, $\mathcal{R}_{I_2}(\theta, \cdot)$ can be obtained by scaling and translating $\mathcal{R}_{I_1}(\theta + \theta_0, \cdot)$ by a factor α and a distance d . This observation is illustrated in Fig. 1 where shape images are given in Fig. 1(a) and 1(d) (one of which is the scaled, rotated, and translated version of the other) and their corresponding Radon images are given in Fig. 1(b) and 1(e).

The invariant property of the 1D Fourier-Mellin transform thus guarantees the same transformed data when applying it on $\mathcal{R}_{I_2}(\theta, \cdot)$ and $\mathcal{R}_{I_1}(\theta + \theta_0, \cdot)$. Fig. 1(c) and 1(f) provide the image data obtained after performing the 1D Fourier-Mellin transform on Radon images in Fig. 1(b) and 1(e) respectively using 169 values of τ ranging from 2.0 to 18.8 with increment of 0.1. The two images in Fig. 1(c) and 1(f) demonstrate clearly the scaling and translation invariant property of the 1D Fourier-Mellin transform, they have the same pattern except for a horizontal shift by θ_0 as a result of the rotation in I_2 . In order to have a geometric invariant shape descriptor, this shifting phenomena can be overcome by applying Fourier transform on the rows of the Fourier-Mellin image $\mathcal{MF}_{\mathcal{R}_I}(\theta, s)$ and then ignoring the phase information in the coefficients. And finally, by normalizing these coefficients by the magnitude of the DC component ($\xi = 0$), the effect of the multiplicative factor, which is a by-product of the 1D Fourier-Mellin transform due to scaling, will be eliminated.

To sum up, the proposed descriptor of a shape image I that is invariant to scaling, rotation, and translation is calculated by $\mathcal{RFM}(I) = \left| \mathcal{F}_{\mathcal{MF}_{\mathcal{R}_I}}(\xi, s) \right|$, namely:

- The Radon transform on the shape image I .
- The 1D Fourier-Mellin transform performed on the columns of the obtained Radon image.
- The magnitude of Fourier transform performed on the rows of the obtained Fourier-Mellin image normalized by the DC component.

Similarity measure: For any two shape images I_1 and I_2 , their measure of similarity is defined as the ℓ_2 -norm distance between their \mathcal{RFM} descriptors:

$$\text{sim}(I_1, I_2) = \|\mathcal{RFM}(I_1) - \mathcal{RFM}(I_2)\|_2. \quad (12)$$

IV. EXPERIMENTAL RESULTS

Performance of the proposed \mathcal{RFM} shape descriptor has been evaluated on two different datasets and compared to other commonly used shape descriptors. The first dataset is the Shapes216 [16] which contains 18 classes of shape with 12 samples per class. Shapes216 is used to evaluate the robustness of the proposed descriptor to occlusion and elastic deformation. The second dataset is the Logos275 which contains 25 classes corresponding to the first 25 logo images of the UMD Logo dataset [17]. Each class has 11 samples obtained after scaling, rotating, and adding salt-and-pepper noise to the original logo image. This dataset is used to evaluate the robustness of the proposed descriptor to noise. Fig. 2 provides some images in the Shapes216 and Logos275 datasets.

The \mathcal{RFM} shape descriptor is compared with shape context (SC) [4], generic Fourier descriptor (GFD) [9], Zernike moments [18], angular radial transform (ART) [19], \mathcal{R} -signature [10], and Radon 2D Fourier-Mellin transforms (R2DFM) [11]. Except for the contour-based SC descriptor, all other descriptors are region-based and additionally \mathcal{R} -signature and R2DFM descriptors are also defined on the Radon transform. These descriptors are selected because they are commonly used and have good reported performance. The average relevant rank (ARR) is used to measure the performance of all the comparing descriptors. Each of the images in the dataset is used as a query to which all images in the dataset are compared with. Thus, 46656 and 75625 comparisons are performed for the Shapes216 and Logos275 datasets respectively. ARR is then computed as the average of retrieval correctness for the first k^{th} nearest match ($k = 12$ for Shapes216 and 11 for Logos275).

Comparison results are given in Fig. 3 for the two datasets. For the Shapes216 dataset, the performance of the \mathcal{RFM} descriptor outperforms the performance of \mathcal{R} -signature and is comparable to the performance of SC, GFD, ART, Zernike, and R2DFM descriptors. However, for the Logos275 dataset, the \mathcal{RFM} descriptor outstrips all other descriptors and has nearly perfect performance. The insensitivity of the \mathcal{RFM} descriptor to noise can be attributed to the use of the Radon transform. In continuous domain, the effect of random noise will be reduced when

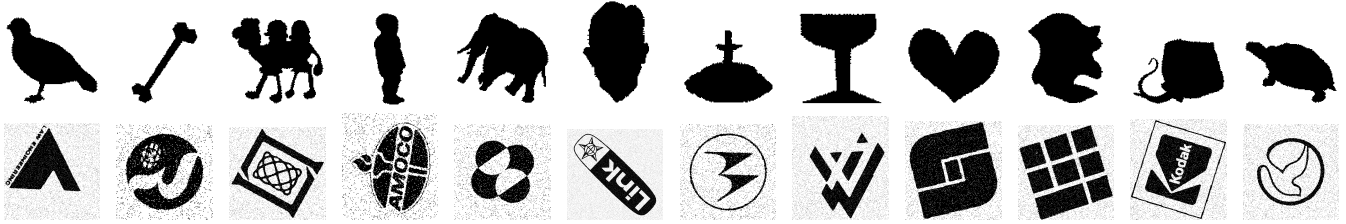


Figure 2. Some sample images taken from the Shapes216 dataset (top row) and the Logos275 dataset (second row).

an integral along a line L is to be taken. In addition, the R2DFM descriptor, even it also uses the Radon transform, performs less than the $\mathcal{R}\mathcal{F}\mathcal{M}$ descriptor. This is due to the incorrectness of the calculated centroid position in the presence of noise.

It should also be noted from the comparison results that although the contour-based SC descriptor provides good performance for the Shapes216 dataset, it performs poorly with the Logos275 dataset. This demonstrates clearly the inappropriateness of the contour-based descriptors for datasets of noisy shape images.

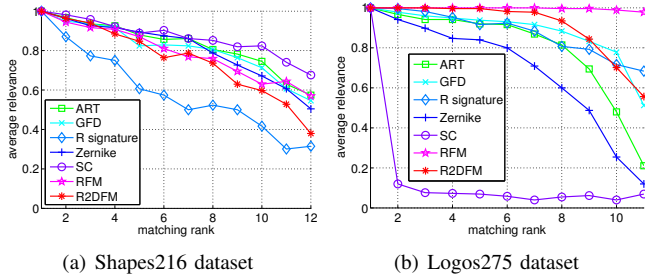


Figure 3. Comparison results in term of average relevant rank.

V. CONCLUSION

This paper presents a new region-based shape descriptor that is invariant to geometric transformation based on the Radon, Fourier, and Mellin transforms. Invariances to translation, rotation, and scaling are obtained by applying 1D Fourier-Mellin and Fourier transforms on the columns and rows of the shape's Radon image respectively. Experimental results show that, when compared to commonly used shape descriptors, the proposed $\mathcal{R}\mathcal{F}\mathcal{M}$ shape descriptor has comparable performance on elastic deformation dataset and outperforms on noisy dataset.

REFERENCES

- [1] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, 2004.
- [2] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Transactions on Computers*, vol. 21, no. 3, pp. 269–281, 1972.
- [3] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. PAMI*, vol. 14, no. 8, pp. 789–805, 1992.

- [4] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. PAMI*, vol. 24, no. 4, pp. 509–522, 2002.
- [5] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. PAMI*, vol. 29, no. 2, pp. 286–299, 2007.
- [6] O. C. Hamsici and A. M. Martínez, "Rotation invariant kernels and their application to shape analysis," *IEEE Trans. PAMI*, vol. 31, no. 11, pp. 1985–1999, 2009.
- [7] M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Am.*, vol. 70, no. 8, pp. 920–930, 1980.
- [8] S. Derrode and F. Ghorbel, "Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description," *Comp. Vis. and Image Under.*, vol. 83, no. 1, pp. 57–78, 2001.
- [9] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825–848, 2002.
- [10] S. Tabbone, L. Wendling, and J.-P. Salmon, "A new shape descriptor defined on the Radon transform," *Comp. Vis. and Image Under.*, vol. 102, no. 1, pp. 42–51, 2006.
- [11] X. Wang, B. Xiao, J.-F. Ma, and X.-L. Bi, "Scaling and rotation invariant analysis approach to object recognition based on Radon and Fourier-Mellin transforms," *Pattern Recognition*, vol. 40, no. 12, pp. 3503–3508, 2007.
- [12] S. R. Deans, *The Radon Transform and Some of Its Applications*. Krieger Publishing Company, 1993.
- [13] J. Bertrand, P. Bertrand, and J. P. Ovarlez, "The Mellin transform," in *The Transforms and Applications Handbook*, A. D. Poularikas, Ed. CRC & IEEE Presses, 2000, ch. 11.
- [14] L. H. Johnson, "The shift and scale invariant Fourier-Mellin transform for radar applications," Massachusetts Institute of Technology, Tech. Rep., 1980.
- [15] P. E. Zwicke and I. Kiss, "A new implementation of the Mellin transform and its application to radar classification of ships," *IEEE Trans. PAMI*, vol. 5, no. 2, pp. 191–199, 1983.
- [16] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Trans. PAMI*, vol. 26, no. 5, pp. 550–571, 2004.
- [17] D. S. Doermann, E. Rivlin, and I. Weiss, "Applying algebraic and differential invariants for logo recognition," *Mach. Vis. Appl.*, vol. 9, no. 2, pp. 73–86, 1996.
- [18] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. PAMI*, vol. 12, no. 5, pp. 489–497, 1990.
- [19] M. Bober, F. Preteux, and W.-Y. Y. Kim, "Shape descriptors," in *Introduction to MPEG 7: Multimedia Content Description Language*, B. S. Manjunat, P. Salembier, and T. Sikora, Eds. Wiley, 2002, pp. 231–260.