Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms

Nafaa Nacereddine, Salvatore Tabbone
LORIA-UMR 7503
Campus Scientifique, BP 239
Vandœuvre-lès-Nancy, France
{naceredn,tabbone}@loria.fr

Djemel Ziou
D^{pt} Informatique
Université de Sherbrooke
Sherbrooke (QC), Canada
djemel.ziou@usherbrooke.ca

Latifa Hamami
Dpt Electronique
Ecole Nationale Polytechnique
Alger, Algérie
latifa.hamami@enp.edu.dz

Abstract–**In this paper, the Radon transform is used to design a new descriptor called Phi-signature invariant to usual geometric transformations. Experiments show the effectiveness of the multilevel representation of the descriptor built from Phi-signature and R-signature, compared to the powerful generic Fourier descriptor.**

Keywords- Radon transform; Phi-signature; R-signature; wavelet transform

I. INTRODUCTION

Shape is recognized as one of the most fundamental characteristics that describe image content [1]. Therefore, effective shape description is essential in object retrieval systems. These objects may consist especially in graphical symbols, which have been widely used for technical document recognition. In general, a graphical symbol can represent a meaning through the use of shape only [2] and it may be complex and may contain unfilled parts.

In this paper, the proposed descriptor is built so that it will be effective for a complex shape in general and a graphic symbol in particular.

In general, a good descriptor should satisfy the following requirements. Firstly, it should have high discrimination ability, i.e. the descriptor derived from different samples of the same class should be close and descriptors derived from samples of different classes should differ significantly [1]. Secondly, descriptors should be independent of the size, orientation and the location of an object.

A lot of region-based shape descriptors have been proposed in the literature. Common methods are based on moment theory including among others, Legendre and Zernike moments [3]. About the generic Fourier descriptors (GFD) proposed by [4], although the experimental results show the superiority of this descriptor compared to a lot of descriptors, whilst its disadvantage is its time-consuming. Tabbone *et al.* [5] use the Radon transform (RT) [6] to define the $\Re$-signature, which consists to do the sum of the elements squared along each Radon matrix column. It reveals its particular suitability to the detection of the rectilinear shapes and conics arc. In this paper, we propose a new shape descriptor, called Φ-signature, which is complementary to the $\Re$-signature transformation by considering the line elements in the Radon matrix. It will be thus expected that the descriptor resulting from the concatenation of the two signatures has a larger discriminating powerfulness for complex shapes, like graphical symbols. All morphological properties and geometrical transformations necessary to define the Φ-signature are calculated directly. The multilevel representation of the signature using the wavelet transform is used for possible descriptor performance improvement. The paper is organized as follows. Sect. 2 gives a summary on Radon transform and the $\Re$-signature. Φ-signature and the wavelet-based descriptors are detailed in Sect. 3. Experiments and a conclusion are given in Sections 4 and 5.

II. RADON TRANSFORM AND $\Re$-SIGNATURE

Let L be a straight line in the x-y plane and ds be the arc length along L, the Radon transform, denoted RT, of a real valued function f is defined by its integral as

$$f^{\vee}(p,\phi) = RT\{f(x,y)\} = \int_{L} f(x,y)\,ds \qquad (1)$$

$RT\{f(x,y)\}$ is then determined by integration of all lines $L_{p,\phi}$ in the x-y plane, $p \in \mathbb{R}$, $\phi \in [0\ \pi]$. From Fig. 1, the equation of L is written as $p = x\cos\phi + y\sin\phi$. If we rotate the coordinate system by an angle ϕ, and label the new axes p and s, $x = p\cos\phi - s\sin\phi$, $y = p\sin\phi + s\cos\phi$. Then, RT can be defined by:

$$f^{\vee}(p,\phi) = \int_{-\infty}^{\infty} f(p\cos\phi - s\sin\phi, p\sin\phi + s\cos\phi)\,ds \quad (2)$$

The properties of linearity, periodicity and symmetry of the RT can be easily obtained [1]. The Radon transform is known mostly for its role in computed tomography [6]. However, the RT possesses many others properties [5,6] and some are relevant for shape representation. But, the RT can not represent the shape structures because it is not invariant to geometric transformations and using it to measure similarity between two shapes is not practical.

The authors in [5] propose a first adaptation of the Radon transform called $\mathfrak{R}$-transform. Through the properties enumerated by the authors, it results that $\mathfrak{R}$-signature is invariant to translation and scaling however a rotation implies a shifting of the signature modulo π. To make this signature invariant also to rotation, we have centered the $\mathfrak{R}$-signature using the Fourier transform so that the zero-frequency component is shifted to center of $\mathfrak{R}^f$ spectrum. Another original RT adaptation will be detailed in the next section.

III. Φ-SIGNATURE

Before we define the Φ-signature, recall that the translation affects considerably the elements of the Radon matrix. Then, f is first translated to place its centre of mass at the x-y coordinate system origin taken at the centre of the image. The coordinates of the centre of mass (x^f, y^f) can be expressed by exploiting solely the Radon transform $f^\vee$. Indeed, by using (2) and the fact that $f^\vee(p,\phi)=g^\vee(p,\phi+\pi/2)$ where $g(x,y)=R_{\pi/2}(x,y)[f]$ and $R_{\pi/2}$ denotes a rotation transformation by an angle of $\pi/2$ (see Fig. 1), the mass centre coordinates can be expressed, for any ϕ, by

$$x^f = \cos\phi \frac{\int_{-\infty}^{\infty} p f^\vee(p,\phi)dp}{\int_{-\infty}^{\infty} f^\vee(p,\phi)dp} - \sin\phi \frac{\int_{-\infty}^{\infty} p f^\vee(p,\phi+\pi/2)dp}{\int_{-\infty}^{\infty} f^\vee(p,\phi)dp} \quad (3)$$

$$y^f = \sin\phi \frac{\int_{-\infty}^{\infty} p f^\vee(p,\phi)dp}{\int_{-\infty}^{\infty} f^\vee(p,\phi)dp} + \cos\phi \frac{\int_{-\infty}^{\infty} p f^\vee(p,\phi+\pi/2)dp}{\int_{-\infty}^{\infty} f^\vee(p,\phi)dp}$$

x^f and y^f can be computed for several values of ϕ and the results are averaged.

Then, the translated function f_c is given by

$$f_c(x,y) = f\left(x+x^f, y+y^f\right) \quad (4)$$

By using the translation property, a direct determination of a centered object RT can be obtained without using the spatial centering operation.

$$f_c^\vee(p,\phi) = f^\vee\left(p+x^f\cos\phi+y^f\sin\phi,\phi\right) \quad (5)$$

Now, let us provide a first definition of the Φ-signature as the integral of $RT\{f_c\}$ according to ϕ

$$\Phi_1^f(p) = \int_0^\pi f_c^\vee(p,\phi)d\phi \quad (6)$$

It can be shown that Φ_1^f can not be invariant to rotation since $f_c^\vee$ is defined on $[0\ \pi]$ and is not π-periodic. By using the symmetry property, an RT extended version on $[0\ 2\pi]$ is derived
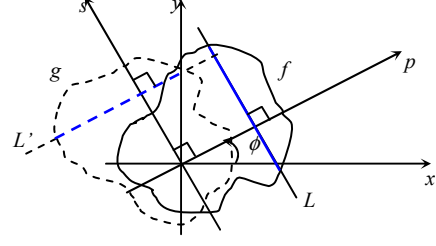

Figure 1. The object and its $\pi/2$ rotated version in the (x,y) and (p,s) axes systems

$$f_{c2\pi}^\vee(p,\phi)=\begin{cases} f_c^\vee(p,\phi) & \text{if } \phi\in[0\ \pi] \\ f_c^\vee(-p,\phi) & \text{if } \phi\in[\pi\ 2\pi] \end{cases} \quad (7)$$

Let us state the second definition of Φ-signature

$$\Phi_2^f(p) = \int_0^{2\pi} f_{c2\pi}^\vee(p,\phi)d\phi \quad (8)$$

Thus, by using this definition and RT periodicity, the invariance of Φ_2^f to rotation can be demonstrated. Indeed, by taking $\phi+\phi_0=\phi'$, we have

$$\Phi_2^{f_r}(p) = \int_0^{2\pi} f_{c2\pi}^\vee(p,\phi+\phi_0)d\phi = \int_{\phi_0}^{2\pi} f_{c2\pi}^\vee(p,\phi')d\phi'$$
$$+ \int_{2\pi}^{2\pi+\phi_0} f_{c2\pi}^\vee(p,\phi')d\phi' = \Phi_2^f(p) \quad (9)$$

where f_r is the rotated centered object with an angle ϕ_0 ($\phi_0\in[-\pi\ \pi]$) i.e. $\forall\ x,y\in\mathfrak{R}^2$, $f_r(x,y)=R_{\phi_0}(x,y)[f_c]$.

Before giving the final formulation of Φ-signature, some transformations are needed in order to make the mentioned signature invariant to scaling.

Let f_S be the centered object scaling transform with a factor α, then $\forall\ x,y\in\mathfrak{R}^2$, $f_S(x,y)=f_c(\alpha x,\alpha y)$, and

$$\Phi_2^{f_S}(p) = \int_0^{2\pi} \frac{1}{\alpha} f_{c2\pi}^\vee(\alpha p,\phi)d\phi = \frac{1}{\alpha}\Phi_2^f(\alpha p) \quad (10)$$

Consequently, the invariance of Φ-signature to scaling can be achieved by normalization regarding the amplitude and range of p. The amplitude normalization can be obtained by dividing Φ_2^f by its maximum i.e. $\Phi_2^{f-amp\_norm}(p)=\Phi_2^f(p)/\max_p(\Phi_2^f)$; whilst p can be normalized by interpolation of Φ_2^f so that, the nonzero part of this latter lies on all the range of p. In other words and as shown in Fig. 2, the p-range normalization can be expressed as $\Phi_2^{f_S}(p')=\Phi_2^{f-amp\_norm}(p)$ where $p'=p(D/2R)$ with D the image diagonal and $R=\min_{\{p\ge 0;\Phi_2^f(p)=0\}}\{p\}$.
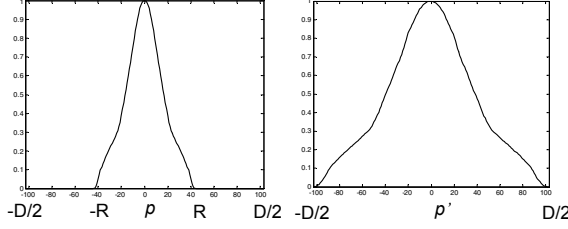
Figure 2.   Scale normalization of Φ-signature regarding the range of p
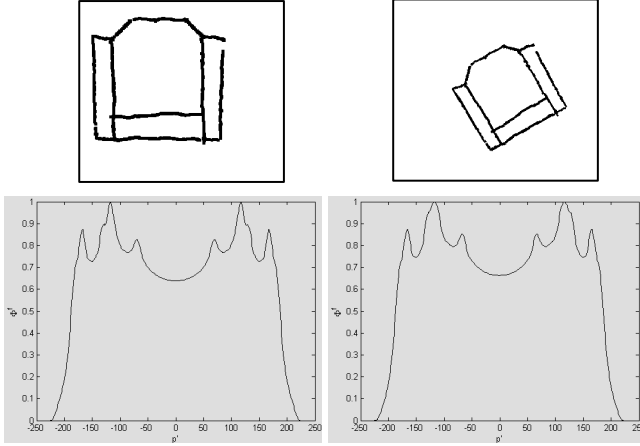

Figure 3.   Φ-signature of an object and that of the same object which is transformed by RST operations

In the following, Φ^f will designate the final version of Φ-signature invariant to translation (T), rotation (R) and scaling (S). Fig. 3 shows Φ^f for an object and the same object rotated, scaled and translated.

Wavelet transforms have been found useful in a variety of applications. This is because they provide the analyst with an approximation of the signal and a detail of the signal as well. In this paper, continuous wavelet transforms (CWT) [7] using some wavelet ψ is applied to represent Φ^f and $\Re^f$ on many levels, differing with the number of details. Since these signatures are already invariant to RST operations, their CWT is exploited without transformation. The CWT of a signal $s(t)$ at a scale a and position b is defined as

$$C_a(b) = \frac{1}{\sqrt{a}} \int_{D_f} s(t)\psi\left(\frac{t-b}{a}\right)dt \qquad (11)$$

where s represents :
Φ-signature with $t = p'$ and $D_f = [-D/2 \ D/2]$, or $\Re$-signature with $t = \phi$ and $D_f = [0 \ \pi]$.

Thus, the multilevel versions of Φ^f and $\Re^f$ are obtained, respectively, as

$$w\Phi^f = \left[C^{\Phi}_{a_1}(1), \cdots, C^{\Phi}_{a_1}(T), \cdots, C^{\Phi}_{a_r}(1), \cdots, C^{\Phi}_{a_r}(T)\right] \qquad (12)$$

$$w\Re^f = \left[C^{\Re}_{a_1}(1), \cdots, C^{\Re}_{a_1}(T), \cdots, C^{\Re}_{a_r}(1), \cdots, C^{\Re}_{a_r}(T)\right] \qquad (13)$$

where $\{a_1, a_2, \ldots, a_r\}$ is the set scales and T is the descriptor resolution for each scale.
Another descriptor is deduced by concatenation as

$$w\Phi\Re^f = \left[w\Phi^f \ \ w\Re^f\right] \qquad (14)$$

IV.   EXPERIMENTS

To show the discrimination ability of the studied descriptors, let see the results in Fig. 4 with simple geometric shapes. It appears from the curves provided by Φ^f and $\Re^f$ that the former gives good discrimination between filled and unfilled objects, whilst curves of the latter are almost similar for the disk and the ring, and just the width of the peaks differentiate the filled rectangle from the unfilled one. Conversely, $\Re^f$ separates better the rectilinear shapes from the round ones. This is why; if the information brought by each signature is considered then the resulting descriptor can be efficient particularly, for unfilled shapes.

In order to evaluate the proposed descriptors, several experiments are carried out on two shape databases. The first one, called *CVC-db*, consists of graphical symbols drawn manually[1]. A part of this database representing 1380 symbols distributed in 10 classes is used. A 2nd database called *Kimia-db* [8] is also tested. It is composed of 216 shapes, grouped in 18 categories with 12 shapes per category.

For comparison purposes, the tests are conducted on the following descriptors: GFD, Φ^f, $w\Phi^f$, $\Re^f$, $w\Re^f$, $\Phi\Re^f = [\Phi^f \ \Re^f]$ and $w\Phi\Re^f$. It is noticed after using several wavelet functions at various scales and resolutions that the Haar wavelet has the best performance and the values of $L = 40$ and $r = 5$ reveals to be more suitable for our databases. The scales are chosen to be dyadic, i.e. $a_i = 2^i$, $i = 1, \ldots, r$.

Relevance feedback (RF), which introduces human visual perception into the retrieval process gradually, is an efficient improvement for narrowing down the gap between low-level visual feature representation of an image and its semantic meaning in content-based image retrieval (CBIR) [9]. In these retrieval experiments, a CBIR system with relevance feedback based on the MindReader approach [10] is used. In our work, since the number of training examples is far less than the dimensionality (54~400) of the tested descriptors, a weighted Euclidean distance is taken instead of a generalized Euclidean one in order to avoid the problem of the covariance matrix inversion needed for the computation of the optimal weight matrix. The retrieval performance is determined by the average *precision versus recall curve* for databases (see Fig. 5).
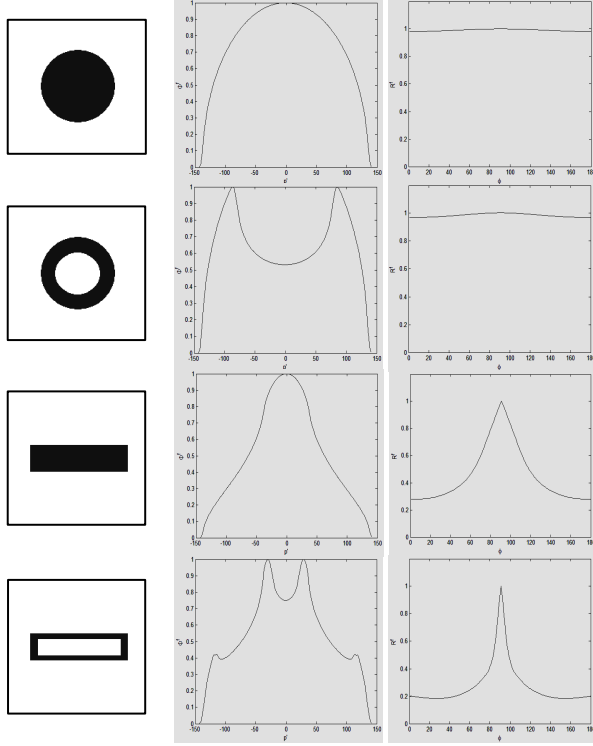
---

[1]   http://www.cvc.uab.es/iapr-tc10/grec2007/

Figure 4. Φ^f (column 2) and $\Re^f$ (column 3) of some simple shapes

For *CVC-db*, the retrieval results without RF (*Rf*0) show that $w\Phi\Re^f$ descriptor outperforms all the other descriptors, especially GFD. Concerning *Kimia-db*, $w\Phi\Re^f$ is slightly better than GFD although this latter is known for its powerfulness for this kind of shapes. In retrieval without RF, one shape for each database class is taken as a query. Some relevant shapes are chosen from the returned images and fed back to the system as training examples with chosen relevance degrees to launch the 1st RF sequence (*Rf*1), and same process for *Rf*2. The improved retrieval results for GFD and $w\Phi\Re^f$ are shown in Fig. 5 (b,d) where the latter still remains better than the former. In addition to its satisfactory discrimination ability, the wavelet version of the concatenated Φ and $\Re$ signatures is not time consuming.

Indeed, it is 10 and 20 times faster than the GFD, for *Kimia-db* and *CVC-db*, respectively. For indication, if the scope used in the precision/scope curve is taken equal to the class cardinality, the recognition rates for $w\Phi\Re^f$ after *Rf*2 reach 85.7% and 93.5% for *CVC-db* and *Kimia-db*, respectively.

V. CONCLUSION

In this paper, a new approach on the design of RT-based descriptors has been developed. The proposed descriptor $\Phi\Re^f$ exploits in a complementary way the rows and columns of Radon matrix. Its multilevel representation increases the discrimination capacity. It is much faster than GFD and outperforms it in retrieval of complex shapes particularly, graphical symbols.

REFERENCES

[1] Y. W. Chen, Y. Q. Chen, "Invariant description and retrieval of planar shapes using Radon composite features," IEEE Trans. on Signal Processing, vol. 56, pp. 4762–4771, 2008.

[2] M.-K. Kim, "A contour shape description method via transformation to rotation and scale invariant coordinates systems," GREC, LNCS, vol. 3926, pp. 312–322, 2006.

[3] R. Mukundan, K. R. Ramakrishnan, "Fast computation of Legendre and Zernike moments," Pattern Recognition, vol. 28, pp. 1433–1442, 1995.

[4] D. Zhang, G. Lu, "Shape-based image retrieval using generic Fourier descriptor," Signal Processing, vol. 17, pp. 825–848, 2002.

[5] S. Tabbone, L. Wendling, J. -P. Salmon, "A new shape descriptor defined on the Radon transform," Computer Vision and Image Understanding, vol. 102, pp. 42–51, 2006.

[6] S. R. Deans, The Radon Transform and some of its Applications, Krieger Publish. Comp., Florida, 1993.

[7] C. K. Chui, An introduction to Wavelets. Academic Press, San Diego, 1992.

[8] T. B. Sebastian, P. N. Klein, B. B. Kimia, "Recognition of shapes by editing shock graphs," ICCV2001, Vancouver, Canada, pp. 755–762, July 2001.

[9] M. L. Kkerfi, D. Ziou, "Image retrieval from the World Wide Web: Issues, techniques and systems," ACM Computing Surveys, vol.36, pp. 35–67, 2004.

[10] Y. Ishikawa, R. Subramanya, C. Faloutsos, "MindReader: Querying databases through multiple examples," 24th Int. Conf. on Very Large DBs, New York, pp. 24–27, 1998.
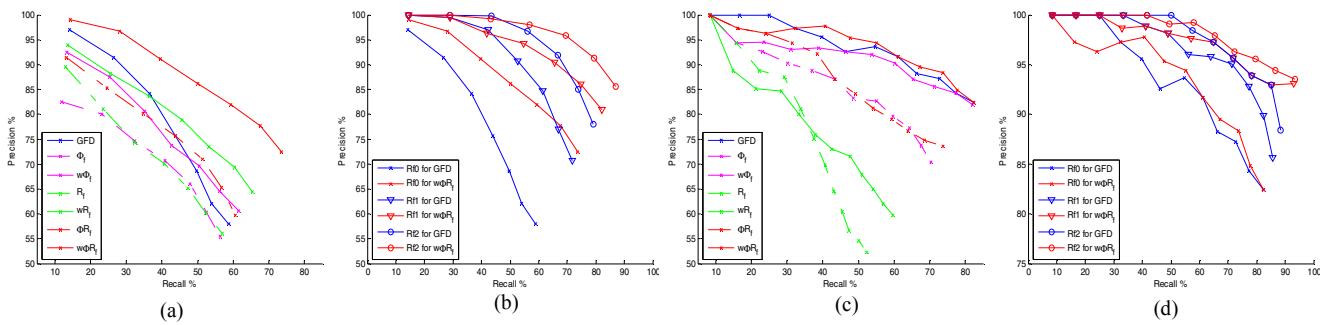
Figure 5. Average precision/recall curves for *CVC-db* (a,b) and *Kimia-db* (c,d) without (a,c) / with (b,d) RF.