A New Multi-resolution Affine Invariant Planar Contour Descriptor

Taha Faidi, Faten Chaieb(✉), and Faouzi Ghorbel

CRISTAL Laboratory, ENSI, La Manouba University, Manouba, Tunisia
taha.faidi@ensi-uma.tn, {faten.chaieb,Faouzi.ghorbel}@ensi.rnu.tn

Abstract. In this paper, a novel affine invariant shape descriptor for planar contours is proposed. It is based on a multi-resolution representation of the contour. For each contour resolution, a shape signature is defined from the contour points and the initial contour centroid and points. Finally, Fourier descriptors are computed for each signature. The proposed descriptor is invariant to affine transformations. Experiments carried on the MPEG-7 coutour database and the Multiview Curve Dataset (MCD) show that our proposed descriptor outperforms other contour shape descriptors proposed in the literature.

1 Introduction

The recognition of planar shapes that are subjected to certain viewing transformations has an increasing interest in many computer vision applications such as robotic vision, content-based image retrieval, registration and 3D reconstruction. Three dimensional objects could be also considered as planar when the camera is far away from the object and distances within the object are negligible.

The use of shape descriptors to deal with such problem seems to be the most efficient method. The main challenging task of shape descriptors is to provide an accurate representation of shape information that deals with two critical problems : (a) viewpoint invariance since object shapes can change as the viewpoint changes due to perspective transformation and (b) parametrization invariance.

The viewpoint is usually described by a group of geometric transformations. The main groups of interest for pattern recognition and image analysis applications are Euclidean, affine and projective. Most work have considered the affine transformations group, which is a pretty good approximation when the object is far from the camera, since the slight distortion that may result from the more general projection can be regarded as part of a deformation.

Furthermore, the shape parametrization is chosen arbitrarily and would not be necessary the same for different views. Thus, an isotropic representation invariant to the geometric transformations group should be considered.

Many shape descriptors have been proposed in the literature and could be classified into two main classes : contour-based shape descriptors and region-based shape descriptors methods [14].

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part II, LNCS 9280, pp. 494–505, 2015.
DOI: 10.1007/978-3-319-23234-8_46

Region-based descriptors are obtained from all the pixels within a shape. They include the angular radial transform (ART) descriptor [6], geometric moments [11], Zernike moments [12] and affine moment invariant [8], etc. In contour-based approach, planar shapes are generally assumed to have a piecewise smooth boundary that is represented by a bidimensional (2D) continuous contour.

Many contour-based descriptors have been proposed in the last decades [9]. Early work is based on the well-known Fourier descriptors (FD) which are derived from Fourier transform of shape signatures [5,13]. The main drawback of such descriptors is that Fourier transform does not provide local shape information. Hence, they are known to be inaccurate in the case of small contour variations and sensitive to the contour extraction method. Multi-scale theory based shape descriptors are considered as a good solution to deal with this drawback. Furthermore, these descriptors are more robust to noise since prevalent features are preserved across scales. Many multi-scale contour based descriptors have been developed in the literature [1–3]. The ISO/IEC MPEG-7 shape descriptor based on the curvature scale space (CSS) representation has been proposed in [1,2]. It is based on the multi scale space theory where the arc-length parametrized contour is convolved with increasing values of Gaussian kernel. The CSS image of each boundary is computed and then the maxima of the curvature zero-crossing are used as a shape descriptor for each object. This descriptor is very compact and quite robust with respect to noise, scale and orientation changes of objects. However, it fail to distinguish shallow concavity from deep concavity on the shape boundary. Furthermore, it is only robust to local variations and it is not robust in global sense. Many extensions of the proposed CSS descriptor have been developed in order to overcome its main drawbacks. In particular, MPEG-7 recommends combing the CSS descriptor with global shape descriptors such as eccentricity and circularity (CSS+). The triangle area representation (TAR) computes the areas of the triangles formed by the boundary points to measure the convexity/concavity of each point at different scales [4]. The area value of every triangle measures the curvature of corresponding contour point and the sign of the area indicates if a contour point is convex, concave or on a straight line. Unlike CSS descriptors, scale levels are obtained by considering different triangle side lengths.

In this work, we propose a new multi-scale planar contour based signature descriptor invariant to affine transformations. The proposed descriptor, is based on the Fourier Descriptors of a multi-scale shape signature from a progressive contour representation. In order to achieve representation isotropy, an affine arc-length based reparametrization is carried out [10].

The remainder of the paper is organized as follows. In Section 2, we briefly recall the affine arclength reparametrization method. The proposed descriptor, also called Fourier Descriptors of Affine invariant Progressive signature (*FD-APS*), is described in Section 3. Experimental results are reported in section 4. Finally, conclusions and future work are presented in Section 5.

2 Affine Arc-Length Reparametrization

In this work we focus on planar shapes represented by simple and closed 2D continuous parametric curve. A curve parametrization $\gamma(t)$ of a geometric curve $\mathbf{\Gamma}$ is an 1-periodic function of a continuous parameter t defined by:

$$\begin{array}{rl} \gamma : [0,1] & \longrightarrow \mathbb{R}^2 \\ t & \longmapsto \gamma(t) = [x(t), y(t)]^t. \end{array} \tag{1}$$

It's well known that a same parametric curve may have different parameterizations. The invariant-descriptors computed from two different parameterizations of the same geometric curve are generally different. This is due to parametrization dependance on transformations. One solution to this problem consist in performing a $\mathbb{G}$-invariant reparametrization of the curve where $\mathbb{G}$ is the geometric transformations group. A reparametrization of (t, γ), noted $(\widehat{t}, \gamma)$, is defined as follows :

$$\gamma(\tilde{t}) = \gamma(\tau(t)) = [x(\tau(t)), y(\tau(t))]^t, \; t \in [0,1]. \tag{2}$$

where τ is an increasing function defined on $[0,1]$.

Let (t_1, γ_1) and (t_2, γ_2) two parameterizations of a geometric curve Γ and its image by a geometric transformation g. After $\mathbb{G}$-invariant reparametrization, both curve parametrizations verify the following equation :

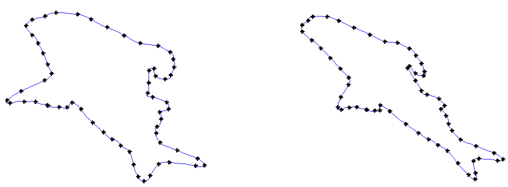
$$\gamma_2(\,\widehat{t}\,) = g(\gamma_1(\,\widehat{t} + t_0)), \;\; t_0 \in \mathbb{Z} \text{ et } g \in \mathcal{G}, \tag{3}$$

where t_0 is departure points difference between the contours.

In the case of affine transformations group, we carry out a reparametrization by the normalized affine arc-length defined as:

$$\bar{s}_a(t) = \frac{1}{L_a} \int_0^t (||\det(\gamma'(t), \gamma''(t))||)^{\frac{1}{3}} dt, \; t \in [0, T]. \tag{4}$$

where L_a is the curve affine length. In figure 1, an affine arc-length reparametrization of a contour and its image by an affine transformation is shown. It's important to notice that after affine arc-length parametrization, a point to point matching between a curve and its image under an affine transformation could be performed up to an overall shift due to different starting points.



(a) (b)

Fig. 1. Affine arc-length reparametrization of a contour (a) and its image by an affine transformation.

3 Fourier Descriptors of Affine Invariant Progressive Signature

Figure 2, shows the main steps performed to obtain the proposed descriptor, denoted *FD-APS*. First, the contour centroid is computed and translated to the origin. Then, an affine arc-length reparametrization is performed to meet parametrization invariance constraint. After, a progressive contour representation is defined and a multi-scale shape signature is generated. Finally, Fourier descriptors of the multi-scale signature are obtained.

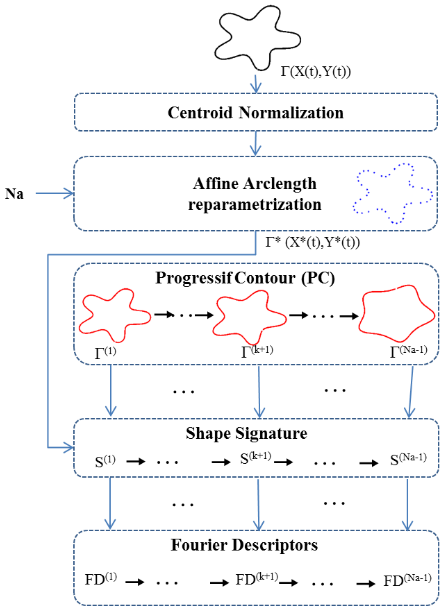


Fig. 2. Block diagram of *FD-APS* descriptor.

3.1 Progressive Contour

Let us consider the discrete parametric representation of a closed curve: $\Gamma = \{m_i = (x_i, y_i)\}_{i=1\cdots N_a}$ where N_a is a number of contour points after affine arc-length reparametrization.

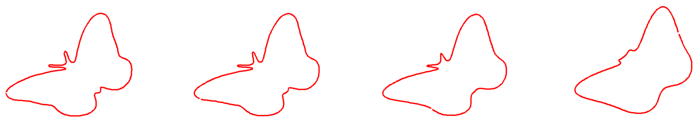
An initial contour can be transformed into a coarser contour Γ^N by applying a sequence of n successive elementary operations called *MiddleArc*. This former, replaces two consecutive points by their medium point:

$$MiddleArc(m_i, m_{i+1}) = \frac{m_i + m_{i+1}}{2}.$$

Thus, N contour resolutions $\{\Gamma^{(k)}\}_{k=1\cdots N}$ are generated as follows :

$$\begin{cases} \Gamma^{(0)} = \Gamma \\ \Gamma^{(k)} = \{MiddleArc(m_i^{(k-1)}, m_{i+1}^{(k-1)})\}_{i=1\cdots N_a-1} \bigcup , \\ \qquad\quad \{MiddleArc(m_{N_a}^{(k-1)}, m_1^{(k-1)})\} \end{cases} \tag{5}$$

An example of contour resolutions is illustrated in figure 3. It's important to note that the *MiddleArc* operation preserves the affine transformation between two contours at different resolutions.



(a) $\Gamma^{(0)}$ (b) $\Gamma^{(20)}$ (c) $\Gamma^{(70)}$ (d) $\Gamma^{(200)}$

Fig. 3. A contour with different resolutions.

3.2 Shape Signature

Basically, our proposed signature is conceived around the linearity of triangular area under affine transform. A multi-scale approach is used to compute the area between the centroid point from the original contour and two given points from respectively original contour and the contour at scale k. By this definition, the signature is expected to preserve the affine invariance and strengthen its discrimination capability.

Therefore, the shape signature $S^{(k)}(t)$ of a given contour $\Gamma^{(k)}$ at resolution k is based on the area of triangles $T(k,t) = (m_t^{(0)}\, m_t^{(k)}\, G)$ where $m_t^{(0)} = (x^{(0)}(t), y^{(0)}(t))$ is a point of $\Gamma^{(0)}$, $m_t^{(k)} = (x^{(k)}(t), y^{(k)}(t))$ is a point of $\Gamma^{(k)}$ and G the centroid of the initial contour (see figure 4). It is given by the following equation:

$$S^{(k)}(t) = \frac{|x^{(0)}(t)\; y^{(k)}(t) - y^{(0)}(t)\; x^{(k)}(t)|}{2} \tag{6}$$

where $x^{(k)}(t)$ and $y^{(k)}(t)$ are defined as:

$$x^{(k)}(t) = \frac{1}{2^{(k)}} \sum_{i=0}^{k} \binom{k}{i} x^{(0)}(r_i)$$

$$y^{(k)}(t) = \frac{1}{2^{(k)}} \sum_{i=0}^{k} \binom{k}{i} y^{(0)}(r_i)$$

Where $r_i = (t+i)\ Mod\ N_a\ and\ t = 0, \cdots, N_a - 1.$

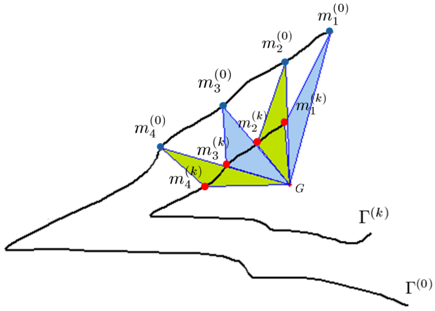


Fig. 4. Shape signature at a given scale.

3.3 Fourier Descriptors

The Fourier Descriptors are obtained by applying Fourier transform to the shape signature. The discrete Fourier transform of the signature $S^{(k)}(t)$ is given by:

$$a_n^{(k)} = \frac{1}{N_a} \sum_{t=0}^{N_a-1} S^{(k)}(t) \exp\left(\frac{-j2\pi nt}{N_a}\right), n = 0, \cdots, N_a - 1$$

The Fourier descriptors of the signature $S^{(k)}$ are derived from the Fourier coefficients $a_n^{(k)}$ as follows:

$$DF^{(k)} = DF^{(k)}(S^{(k)}) = \left\{ \frac{|a_n^{(k)}|}{|a_0^{(k)}|} \right\}_{n=1\cdots p}, \tag{7}$$

where p is the number of Fourier coefficients. Therefore, the proposed *FD-APS* descriptor is defined by $\{J_k\}_{k=1\cdots N}$:

$$\{J_k\}_{k=1\cdots N} = \{DF^{(k)}\}_{k=1\cdots N}, \tag{8}$$

where N is the number of scales. In this work, we consider N_a scales.

In order to prove the invariance property of the proposed descriptor, let us consider a contour Γ_1 and its image by an affine transformation Γ_2 characterized by its matrix A.

It's easy to verify that $S_2^{(k)}(t) = det(A)\ S^{(k)}(t)$ where $det(.)$ is the determinant. Hence, the affine transformation applied to a contour is mapped to a scaling between correspondent signatures. As the Fourier descriptor of the signature is invariant to scaling and starting point, the proposed descriptor is invariant to affine transformations and starting point.

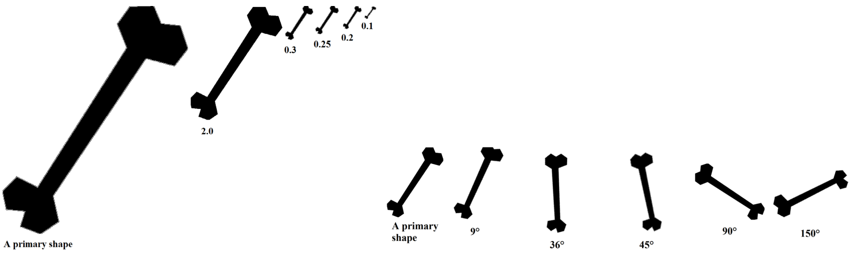
The similarity measure used to compare two shape contours $\Gamma 1$ and $\Gamma 2$ can be formalized as follows:

$$d(\Gamma 1, \Gamma 2) = \frac{1}{N} \sum_{k=1}^{N} \frac{\|J1_k - J2_k\|_2}{\max(\|J1_k\|_2, \|J2_k\|_2)} \tag{9}$$

where N is the number of contour resolutions and $\|.\|_2$ is the L_2 norm.

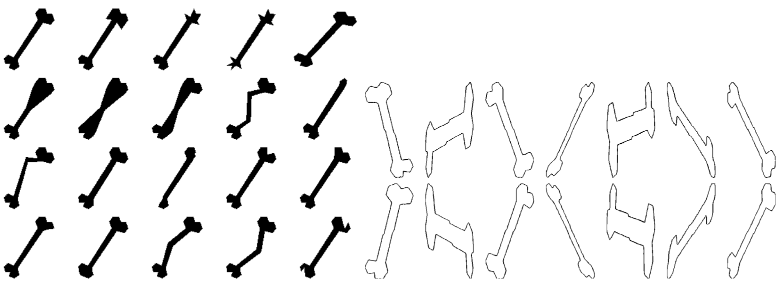
4 Experimental Results

Our experimentations were conducted on two databases: MPEG-7 shape database and the Multiview Curve dataset (MCD) [15]. The MPEG-7 database consists of three parts, Set A and Set B. Set A has two parts, Set A1 and Set A2, each contain 420 shapes of 70 classes. Set A1 is for scale invariance and Set A2 is for rotation invariance. Set B has 1400 shapes classified into 70 classes. Set B is for similarity-based retrieval and shape descriptors robustness under various arbitrary shape distortions, that include rotation, scaling, arbitrary skew, stretching, defection, and indentation. The Multiview Curve dataset is composed of 40 shape classes taken from MPEG-7 database. Each class contains 14 curve samples that correspond to different perspective distortions of the original curve.this database is used to test shape descriptors under affine transformation. Samples of shapes from the MPEG-7 (set A and B) and MCD databases are shown in figures 5 and 6.



(a) MPEG-7 SetA1 (Scale: 0.1 to 2) (b) MPEG-7 SetA2 (Rotation: 9 to 150)

Fig. 5. Samples of shapes from the MPEG-7 database (Set A1 and Set A2).



(a) MPEG-7 SetB (b) MCD

Fig. 6. Samples of shapes from the MPEG-7 database (Set B) and MCD database.

To evalute the performance of shape descriptors techniques in terms of image retrieval efficiency, a precision-recall curve are the most commonly used measure. The precision is defined as the number of relevant shapes retrieved divided by the total number of shapes retrieved and the recall is defined as the number of relevant shapes retrieved divided by the total number of relevant shapes in the class (size of a class). The precision recall curve is plotted by averaging precision and recall over all database shapes.

In the case of rigid transformations, the proposed descriptor is compared with Fourier Descriptors (FD), Curvature Scale Space (CSS) and the Affine Invariant Fourier Descriptor (AIFD) [7] on the MPEG-7 Set A and MCD databases. The precision-recall curves of the above mentioned descriptors conducted respectively on MPEG-7 parts A1, A2 and MCD are respectively shown in the figures 7, 8 and 9. The results for MPEG-7 database part A, used to evaluate descriptors under scale and rotation invariance, show that our descriptor outperforms FD, CSS and AIFD descriptors.

Regarding robustness to arbitrary shape distortions described in the MPEG-7 database part B, a comparison with some commonly used descriptors is performed [9]. In particular, we consider the farthest points distance (FPD), radius

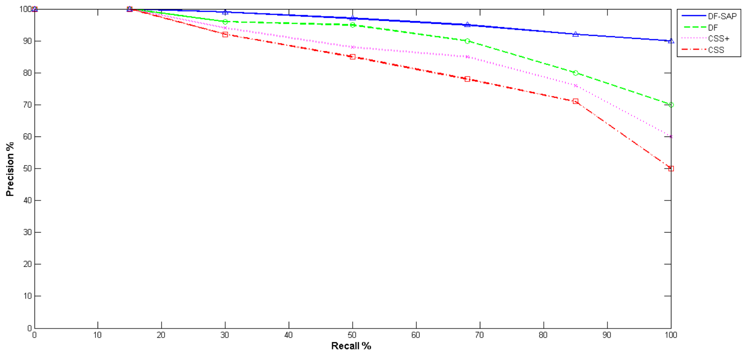


Fig. 7. Average precision and recall of retrieval on Set A1

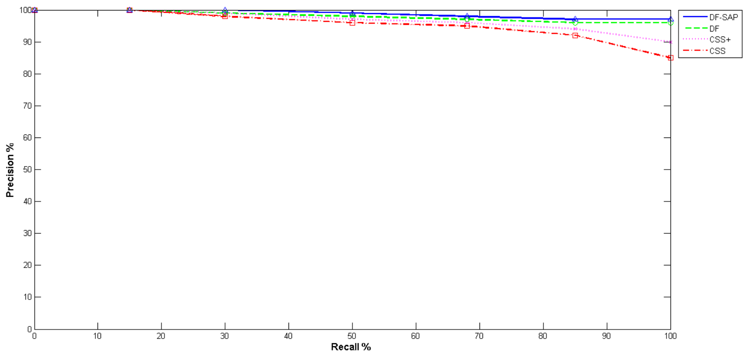


Fig. 8. Average precision and recall of retrieval on Set A2

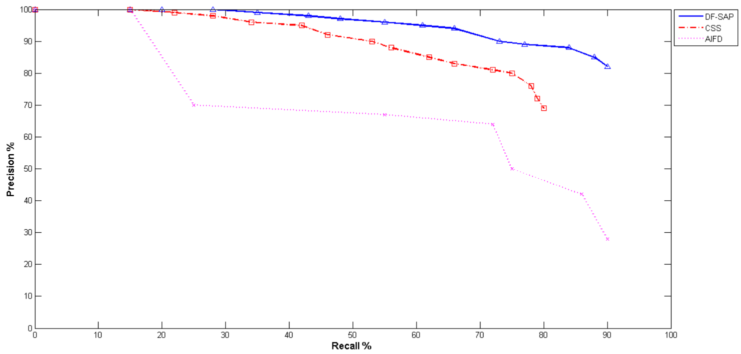


Fig. 9. Average precision and recall of retrieval on MCD

distance (RD), triangle centroid area (TCA), complex coordinates (CC), polar coordinates (PC), angular radial coordinates (ARC), triangular area representation (TAR) [4], chord length distance (CLD) and angular function (AF).

Table 1 shows the average retrieval precision scores for low and high recalls for the proposed descriptor and above mentioned descriptors from literature. The results of the other descriptors are reported from [9]. The average retrieval precision of our descriptor is respectively 78.69 and 44.46 for low and high recall rates and is higher than that in the other methods. Figures 10 and 11, show respectively the retrieval results of 10 random queries from MPEG-7 part B and MCD databases based on our proposed descriptor. It's easy to notice that the proposed descriptor performs better on the MCD database than on the MPEG-7 part B database. In fact, this is due to the theoretical invariance property of our

Table 1. The average precision for low and high recalls for the DF-SAP descriptor and other signatures using MPEG-7 database setB.

Signature	Low Recall The average precision for recall rate <= 50 %	High Recall The average precision for recall rate >= 50%
DF-SAP	78.69	44.46
FPD	75.82	42.13
RD	75.69	41.77
TCA	73.4	38.5
CC	64.76	22.59
PC	64.4	35.12
ARC	58.93	26.83
TAR	58.7	23.54
CLD	57.8	24.00
AF	57.39	27.88

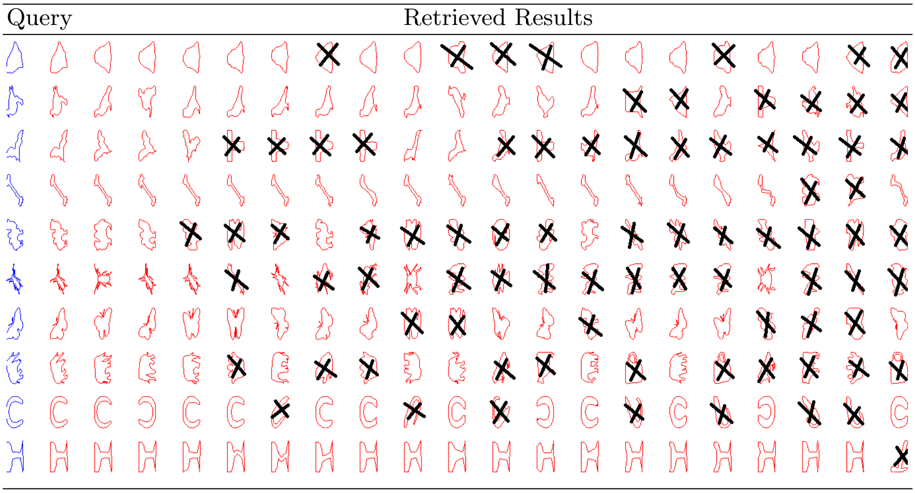


Fig. 10. 10 random retrieval results from MPEG-7 SetB database.

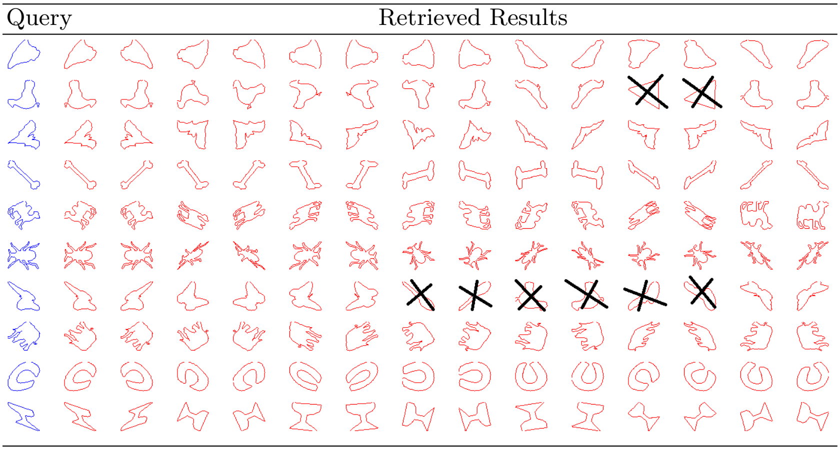


Fig. 11. 10 random retrieval results from MCD database.

descriptor under affine transformations. Furthermore, multi-scale used approach makes our descriptor robust to small distortions.

5 Conclusion

In this paper, a novel multi-scale and affine shape descriptor is proposed. The performance of this descriptor is evaluated using the MPEG-7 and MCD database and compared with other shape descriptors. Results show that in terms of recall-precision performance measures, our descriptor is efficient and competitive for shape retrieval.

Furthermore, it has been shown in [10] that completeness property of descriptors is required to ensure the existence of a distance between shapes which have a right physical mean. In fact, the completeness property allows us to recover shape from its descriptors. In our future work, we will aim to derive a complete family of descriptors from the proposed descriptor which is not complete.

Acknowledgments. This work is carried out as part of a MOBIDOC thesis financed by the European Union under the PASRI program in partnership with the Research and Studies Telecommunications Center (CERT) and CRISTAL laboratory.

References

1. Abbasi, S., Mokhtarian, F., Kittler, J.: Curvature scale space image in shape similarity retrieval. Multimedia Syst. **7**(6), 467–476 (1999)
2. Abbasi, S., Mokhtarian, F., Kittler, J.: Enhancing css-based shape retrieval for objects with shallow concavities. Image and Vision Computing **18**(3), 199–211 (2000)
3. Adamek, T., O'Connor, N.: A multiscale representation method for nonrigid shapes with a single closed contour. IEEE Transactions on Circuits and Systems for Video Technology **14**(5), 742–753 (2004)
4. Alajlan, N., Rube, I., Kamel, M.S., Freeman, G
5. Bartolini, I., Ciaccia, P., Patella, M.: Warp: accurate retrieval of shapes using phase of fourier descriptors and time warping distance. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(1), 142–147 (2005)
6. Bober, M.: Mpeg-7 visual shape descriptors. IEEE Trans. Circuits and Systems for Video Technology **11**(6), 716–719 (2001)
7. Chaker, F., Bannour, M., Ghorbel, F.: A complete and stable set of affine-invariant fourier descriptors. In: Proceedings of the 12th International Conference on Image Analysis and Processing 2003, pp. 578–581, September 2003
8. Chong, C.W., Raveendran, P., Mukundan, R.: An efficient algorithm for fast computation of pseudo-zernike moments. International Journal of Pattern Recognition and Artificial Intelligence **17**(06), 1011–1023 (2003)
9. El-ghazal, A., Basir, O., Belkasim, S.: Farthest point distance: A new shape signature for fourier descriptors. Signal Processing: Image Communication **24**(7), 572–586 (2009)
10. Ghorbel, F.: Towards a unitary formulation for invariant images description; application to images coding. Annales des télécommunications **53**(5,6), 242–260 (1998)
11. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory **8**(2), 179–187 (1962)
12. Lin, T.W., Chou, Y.F.: A comparative study of zernike moments. In: Proceedings of the IEEE/WIC International Conference on Web Intelligence WI 2003, pp. 516–519, October 2003

13. Zhang, D., Lu, G.: A comparative study of fourier descriptors for shape representation and retrieval. In: Proceedings of the Fifth Asian Conference on Computer Vision (ACCV02), pp. 646–651. Melbourne, Australia, January 2002
14. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition **37**(1), 1–19 (2004)
15. Zuliani, M., Kenney, C., Bhagavathy, S., Manjunath, B.S.: Drums and curve descriptors. In: In British Machine Vision Conference (2004)