J. Vis. Commun. Image R. 23 (2012) 622–633

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Invariant curvature-based Fourier shape descriptors

A. El-ghazal [a,*], O. Basir [a], S. Belkasim [b]

[a] Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1
[b] Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA

ARTICLE INFO

Article history:
Received 8 February 2011
Accepted 16 January 2012
Available online 02 February 2012

Keywords:
Fourier descriptors
Feature extraction
Curvature-scale space
Content-based image retrieval
Shape matching
Invariant descriptors
Shape retrieval
Shape signature

ABSTRACT

Shape descriptors have demonstrated encouraging potential for retrieving images based on image content, and a number of them have been reported in the literature. Nevertheless, most of the reported descriptors are still face accuracy and computational challenges. Fourier descriptors are considered to be promising descriptors as they are based on a sound theoretical foundation and also have the advantages of computational efficiency and attractive invariance properties. This paper proposes a new curvature-based Fourier descriptor (CBFD) for shape retrieval. The proposed descriptor takes an unconventional view of the curvature-scale-space representation of a shape contour as it treats it as a 2-D binary image (hence referred to as curvature-scale image, or CSI). The invariant descriptor is derived from the 2-D Fourier transform of the curvature-scale image. This method allows the descriptor to capture the detailed dynamics of the shape curvature and enhance the efficiency of the shape-matching process. Experiments using the widely known MPEG-7 databases in conjunction with a created noisy database have been conducted in order to compare the performance of the proposed descriptor with six commonly used shape-retrieval descriptors: curvature-scale-space descriptor (CSSD), angular radial transform descriptors (ARTD), Zernike moment descriptors (ZMD), radial Tchebichef moment descriptors (RTMD), generic Fourier descriptor (GFD), and the 1-D Fourier descriptor (1-FD). The performance of the proposed descriptor has surpassed that of many of these notable descriptors.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

With the increasing number of images in digital databases in last decades, content-based image retrieval (CBIR) has become an active research area. These databases cover a wide range of applications, including, expert medical systems, military use, crime prevention, remote sensing systems, and digital libraries, to name a few. The storage format of such image data is relatively standardized; however, the effective retrieval of images from such databases remains a significant challenge.

Images in a database are retrieved based on either textual information or content information. Text-based techniques have many limitations due to their reliance on manual annotation – a tedious and error-prone process, especially for large data sets. Furthermore, the rich content typically found in images and the subjectivity of human perception make using words to describe images a difficult, if not impossible, task. Due to these difficulties, Content Based Image Retrieval was proposed [1]. This approach to image retrieval relies on the visual content of images, rather than on textual annotations, to search for images, and therefore has the potential to respond to more specific user queries. CBIR techniques use such visual contents as color, texture, and shape to represent and index images. Visual contents such as color and texture have been explored more thoroughly than shape. The increasing interest in using shape features of objects for CBIR is not surprising, since shape is a more intrinsic property of objects than color or texture, given the considerable evidence that natural objects are recognized based primarily on their shape [2,3]. A survey of users with respect to the cognition aspects of image retrieval indicates that users are more interested in retrieval based on shape than on color or texture [4]. However, retrieval based on shape remains a more difficult task than that based on other visual features [2].

Despite the fact that there are some generic and effective approaches such as SIFT (scale-invariant feature transform) and SURF (Speeded Up Robust Feature) that can be used to retrieve shapes, the use of shape-based image retrieval is still more efficient and effective for content based image retrieval applications that include either figurative images or objects within an image such as trademarks, logos, clip art, and icons [5–7].

Significant progress has been made in the development of shape-based image retrieval systems [8,9]. Typically, there are two approaches to shape representation: the region-based approach and the boundary-based approach. Region-based techniques often use moment descriptors to describe shapes. These descriptors include geometrical moments [10,11], Zernike moments [12,13], pseudo-Zernike moments [14], Legendre moments [12], Tchebichef

* Corresponding author.
E-mail addresses: aelghaza@engmail.uwaterloo.ca, akrem@pami.uwaterloo.ca (A. El-ghazal), obasir@uwaterloo.ca (O. Basir), sbelkasim@cs.gsu.edu (S. Belkasim).

1047-3203/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2012.01.011

moments [15], and the radial Tchebichef moments [16]. Other notable region-based techniques include the generic Fourier descriptor (GFD) [17], the compound image descriptor [18], the shape matrix [19], and the grid technique [20]. Although region-based approaches are global in nature and can be applied to generic shapes, they often involve intensive computation and fail to distinguish between similar objects [21].

In many applications, the internal content of the shape is not as important as its boundary. Boundary-based techniques tend to be more efficient for handling shapes that are describable by their object contours [21]. Many boundary-based techniques have been proposed in the literature, including Fourier descriptors (FDs) [22–24], curvature-scale-space descriptors (CSSD) [25–27], wavelet descriptors [28,29], contour displacement [30], Chain codes [31], autoregressive models [32], and the Delaunay triangulation technique [33]. Recently, Dynamic Programming (DP) has been adopted in order to achieve a high accuracy rate using shape boundaries [30,34–38]. Although the DP-based techniques offer generally better performance than other techniques that do not use DP, the DP-based techniques have the disadvantage of being computationally expensive, making them impractical for large databases.

In general, Fourier descriptors are a promising boundary-based approach for shape-based image retrieval, as the FDs are based on the well-known Fourier theory, making them easy to compute and simple to normalize and interpret [22–24,39,40]. In addition, the computational efficiency and compactness of FDs make them well suited to online image retrieval.

Fourier descriptors are often based on applying the Fourier transform to a shape signature. Several shape signatures have been proposed to derive Fourier descriptors. However, in the literature, the complex coordinate is the most frequently used signature. Recent work shows that for shape-based image retrieval, the radial distance signature outperforms complex coordinates and other signatures [41]. To increase the ability of the Fourier-based technique to capture local features, Eichmann et al. [42] have used the Short Fourier Transform (SFT). The SFT is not suitable for image retrieval because the matching process using SFT is computationally more expensive than traditional FDs. Invariance to affine transforms allows considerable robustness in the case of rotating shapes in all three dimensions. Arbter et al. [43,44] have used a complex mathematical analysis and proposed a set of normalized descriptors which are invariant under any affine transformation. Oirrak et al. [45] have also used one-dimensional Fourier series coefficients to derive affine invariant descriptors. Zhang and Lu [41] have shown that although the affine Fourier descriptor [43] was proposed as a way to target affined shape distortion, it does not perform well on the standard affine invariance retrieval set of the MPEG-7 database because the affine Fourier descriptors are designed to work on a polygonal shape under affine transformation and are not designed for a non-rigid shape [41]. Most of the Fourier-based techniques utilize the magnitude of the Fourier transform and ignore the phase information in order to achieve rotation invariance as well as to make the descriptors independent from the starting point. However, Bartolini et al. [36] have described a technique which exploits the phase information. Recently, Kunttu et al. [24] have introduced a multiscale Fourier descriptor for shape-based image retrieval. These descriptors are presented in multiple scales through the adoption of the wavelet and Fourier transform, which improves the shape-retrieval accuracy of the traditional Fourier descriptors.

To generate the FDs of a shape, the 2-D shape is converted to a 1-D signature. A good number of shape signatures have been reported in the literature [22,41,46]. The complex coordinate (CC), the radial distance (RD), and the triangular centroid area (TCA) are some of the notable signatures available to drive FDs. Fourier descriptors derived from different signatures can have significantly different effects on the results of the retrieval [41]. Zhang et al. have conducted a comprehensive study and evaluation of a variety of Fourier methods [41]. Recently, El-ghazal et al., have proposed a shape signature that includes corner information to enhance the performance of shape retrieval using Fourier descriptors [47,48].

This paper proposes a new curvature-based Fourier descriptor and compares its performance with that of other frequently used descriptors. This descriptor is an extension to the descriptor we previously presented [49]; however, in this paper a new version of the descriptor is introduced to handle shape convexities. Furthermore, more comparative analysis and new experimental results are reported in this paper based on four additional databases and three other descriptors to make our study more comprehensive and to strengthen the quality of the comparative analysis.

The paper is organized as follows: Section 2 provides a brief introduction of the curvature-scale-space technique. Section 3 introduces the proposed curvature-based Fourier descriptor (CBFD) as well as its normalization scheme. In Section 4, experimental results using all the sets of the well-known MPEG-7 databases and a created noisy database are presented as means of comparing the proposed descriptor with six notable descriptors: curvature-scale-space descriptor (CSSD), angular radial transform descriptors (ARTD), Zernike moment descriptors (ZMD), radial Tchebichef moment descriptors (RTMD), generic Fourier descriptor (GFD), and the 1-D Fourier descriptor (1-FD). Section 5 provides concluding remarks drawn from this study.

2. Related work

The curvature-space-scale (CSS) technique is based on the multi-scale curvature information obtained from the inflection points of the contour as it evolves. To obtain the curvature-scale-space map, first an image contour is parameterized using an arc-length [25]. This step is accomplished by sampling the contour at equal intervals and recording the 2-D coordinates of each sampled point $(x(s),y(s))$, $s = 0,1,\ldots,L$. Then, the curvature is derived from the shape boundary as follows [21]:

$$k(s,\sigma) = \frac{\dot{X}(s,\sigma)\ddot{Y}(s,\sigma) - \ddot{X}(s,\sigma)\dot{Y}(s,\sigma)}{(\dot{X}(s,\sigma)^2 + \dot{Y}(s,\sigma)^2)^{3/2}}, \quad (1)$$

where $\dot{X}(s,\sigma) = x(s) * \dot{g}(s,\sigma)$, $\ddot{X}(s,\sigma) = x(s) * \ddot{g}(s,\sigma)$, $\dot{Y}(s,\sigma) = y(s) * \dot{g}(s,\sigma)$, $\ddot{Y}(s,\sigma) = y(s) * \ddot{g}(s,\sigma)$, "$*$" is the convolution operator, and $g(s,\sigma)$ is a Gaussian function of standard deviation σ, while $\dot{g}(s,\sigma)$ and $\ddot{g}(s,\sigma)$ are the first and second derivative of $g(s,\sigma)$, respectively.

After each convolution, the curvature zero-crossing points are recovered and mapped to the curvature-scale space. This process continues, and new zero-crossing points are recovered and mapped until no curvature zero-crossing points are found. The result of the mapping is usually an interval tree, called a CSS map, consisting of inflection points. The CSS technique is based on extracting only the peaks that are more than 1/6 of the highest peak of the CSS map. These peaks are used as descriptors to index the shape. A shape, a sample of the shape contours during the evolution process, the CSS map, and the peaks of the CSS map are shown in Fig. 1.

The CSS descriptor is a translation invariant. The scale invariant can be achieved by normalizing all the shapes into a fixed number of boundary points [50]. Rotation causes a circular shifting on the s axis, as shown in Fig. 2. To achieve the rotation invariance, the highest peak is shifted to the origin of the CSS map. The similarities of the two shapes A and B are then, measured by summing the differences between all the matched peaks (peaks within a threshold value are considered to be matched) and the peak values of the

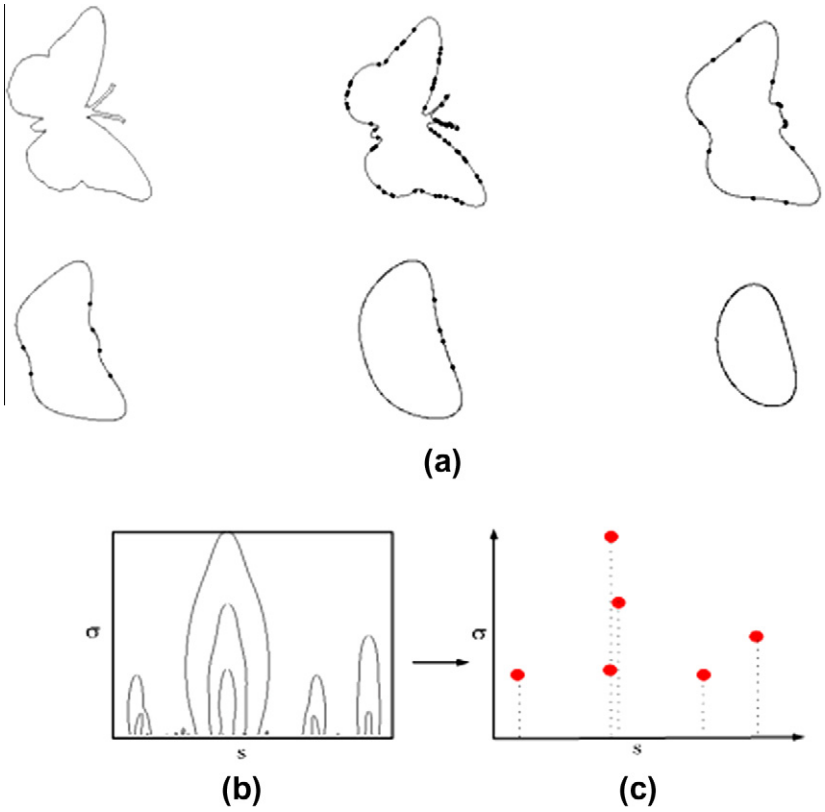


Fig. 1. (a) A shape and samples of its contour during the evolution processes with zero-crossings. (b) The CSS map. (c) The peaks of the CSS map.

unmatched peaks. To increase the accuracy of the result, other schemes of circular shifting matching are used. For example, rather than shifting the primary peak of A to match the primary peak of B, the primary peak of A can be shifted to match the secondary peak of B, or the secondary peak of A can be shifted to match the primary peak of B. The mirror shape has a CSS that differs from the original one, and matching must include the mirrored shape as well. The details of the implementation of the CSS descriptor are described in [21].

Although the CSS technique has many advantages, including its ability to capture local shape features and its robustness with respect to noise, it also has shortcomings. The technique uses only the peaks of the CSS map and ignores other important features, such as the dynamic change in the curvature during the process of extracting the zero-crossing points. This kind of feature might be a key factor in distinguishing between different shapes. Moreover, the CSS descriptor is not a rotation invariance unless a circle shift is applied in the matching stage, and hence, the online matching of the CSS involves many schemes of circular shift in order to align the peaks [17]. Another drawback of the CSS technique is that the number of peaks varies by shape, and these peaks are often mismatched and can be ordered quite differently. This fact adds to the complexity of the matching stage and increase the processing time required. These shortcomings affect the overall performance of the CSS descriptor for shape retrieval. The following section proposes a new curvature-based Fourier descriptor that will overcome the shortcomings of the CSS descriptors.

3. A novel curvature-based Fourier descriptor (CBFD)

This section proposes a new Fourier descriptor based on mapping the curvature-scale space of a shape contour into the 2-D Fourier domain in order to overcome the shortcomings of the well-known CSS descriptors described in the previous section.

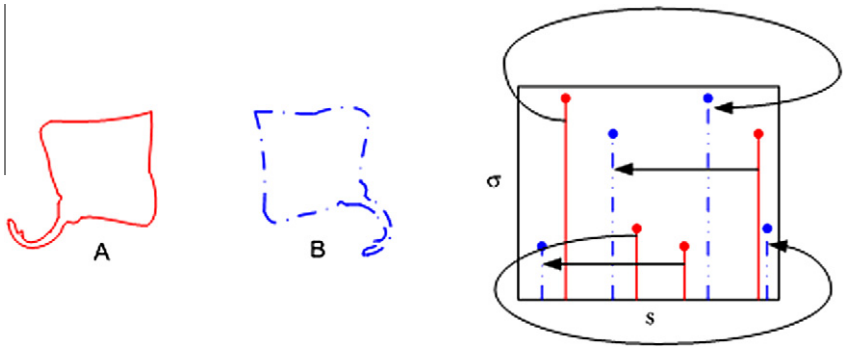


Fig. 2. A shape before and after rotation and the peaks of the CSS map before and after rotation.

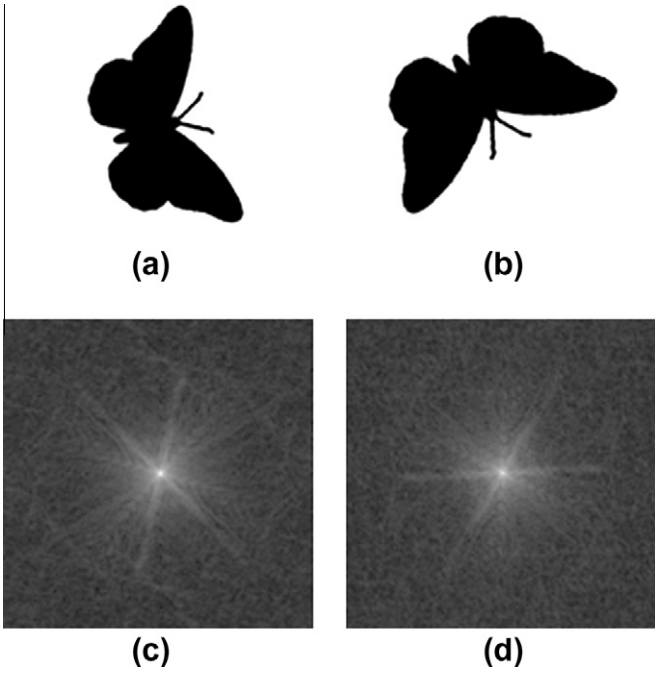


Fig. 3. The magnitude of the Fourier transform for a shape and its rotated version.

The 2-D Fourier transform can be applied to any binary image $f(x,y)$ in the normal Cartesian coordinate system in order to extract Fourier descriptors, as follows:

$$F(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cdot\exp[-j2\pi(ux/M+vy/N)] \tag{2}$$

However, the resulting descriptor is not a rotation invariant even though the phase of the Fourier transform is ignored. In other words, the 2-D Fourier transform produces different descriptors if a 2-D shape $f(x,y)$ is rotated by θ in the normal Cartesian space [51]:

$$f(x\cos\theta+y\sin\theta,-x\sin\theta+y\cos\theta)\leftrightarrow F(u\cos\theta+v\sin\theta,-u\sin\theta+v\cos\theta) \tag{3}$$

Fig. 3 shows the Fourier transform of a shape and its rotated version. According to Eq. (3) the magnitudes of the Fourier transform for the original shape and for its rotated version are different.

It is difficult to obtain the rotation invariance of the features when the images are represented in the normal Cartesian coordinate system [17]. On the other hand, the 2-D Fourier transform has many properties that make it a useful tool for image processing and analysis. One of these properties is that a shift in the normal Cartesian coordinate system affects only the phase of the Fourier transform:

$$f(x-x_0,y-y_0)\leftrightarrow F(u,v)\cdot\exp[-j2\pi(ux_0+vy_0)] \tag{4}$$

A practical application of the 2-D Fourier transform requires a representation of the image that affects the phase of the 2-D Fourier transform when the original shape is rotated [51]. To apply the 2-D Fourier transform, the procedure explained in Section 2 is used, and the contour of the original shape is represented by a binary image that has boundaries similar to those of the CSS map. This image will be referred to as a curvature-scale image (CSI), which can be obtained as follows:

$$\mathrm{CSI}(s,\sigma)=\begin{cases}0 & \text{if } k(s,\sigma)<0\\ 1 & \text{if } k(s,\sigma)\geqslant 0\end{cases} \tag{5}$$

where $k(s,\sigma)$ is the contour curvature given by Eq. (1).

Fig. 4a shows the CSI of the shape presented Fig. 3a. It can be seen that the boundaries of the CSI are exactly the same as those of the CSS map shown in Fig. 1b. Fig. 4b shows the CSI of a rotated version of the shape from Fig. 1a. It can be seen in Fig. 4 that the effect of the rotation in the original shape corresponds to that of a circular shift of the CSI image. This property makes it possible to derive rotation invariance descriptors using the 2-D Fourier transform and the property given in Eq. (4). To apply the 2-D Fourier transform, the CSI(s,σ) is treated as a two-dimensional rectangular binary image in the normal Cartesian coordinate system. Therefore, if the 2-D Fourier transform is applied to the CSI(s,σ), the effect of the phase on the resulting descriptor can be eliminated by taking only the magnitude and ignoring the phase of the 2-D Fourier transform of the CSI(s,σ). If the CSI(s,σ) is constructed from a contour of a binary image, then the 2-D Fourier transform of the CSI(s,σ) is defined as

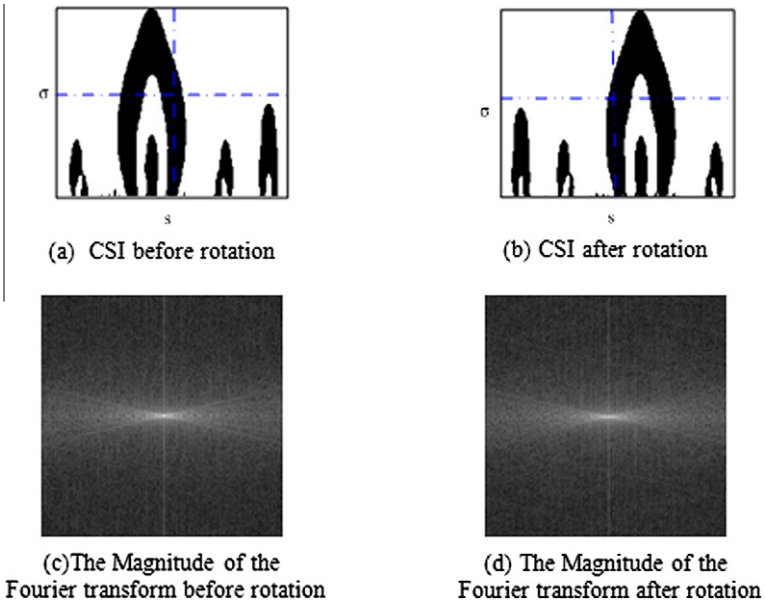


Fig. 4. The CSI images for the image shown in Fig. 3a and its rotated version, and the magnitude of Fourier transform for each one.

$$\mathrm{FCSI}(u,v)=\sum_{t=0}^{m-1}\sum_{\sigma=0}^{n-1} CSI(s,\sigma)\cdot\exp[-j2\pi(us/m+v\sigma/n)] \tag{6}$$

where m is the maximum number of contour points and n is the maximum number of iterations needed to construct the CSI image using different standard deviations σ.

The Fourier descriptors derived from FCSI(u,v) coefficients have to be invariant with respect to translation, rotation, and scaling. Given that the CSI(s,σ) image is translation invariant, it is obvious that the FCSI(u,v) coefficients also possess translation invariance. The rotation invariance of the FCSI(u,v) coefficients is achieved by ignoring the phase of the FCSI(u,v) coefficients. To achieve scale invariance, all the FCSI(u,v) coefficients are divided by the magnitude of the first coefficient [17]. The reason for choosing the first coefficient as a factor in scale normalization is that it represents the average energy of the CSI image. Moreover, the |FCSI(0,0)| is usually the largest coefficient, and consequently, the normalized descriptors are within [0 1] [17]. The proposed translation, scaling, and rotation invariance descriptors are as follows:

$$\mathrm{FD}=\frac{|\mathrm{FCSI}(0,1)|}{|\mathrm{FCSI}(0,0)|},\frac{|\mathrm{FCSI}(0,2)|}{|\mathrm{FCSI}(0,0)|},\ldots,\frac{|\mathrm{FCSI}(0,n)|}{|\mathrm{FCSI}(0,0)|},\ldots,\frac{|\mathrm{FCSI}(m,0)|}{|\mathrm{FCSI}(0,0)|},\ldots,\frac{|\mathrm{FCSI}(m,n)|}{|\mathrm{FCSI}(0,0)|} \tag{7}$$

The first descriptor is ignored because it is used for scale normalization. For efficient retrieval, only a few descriptors are used to index the shape.

Zhang and Lu described and evaluated commonly used similarity measurements [52]. Their results shows that Euclidean distance, city block distance and χ^2 statistics are the more desirable distance measures in terms of both retrieval effectiveness and retrieval efficiency. In this paper, the similarity measure between two shapes indexed with M normalized Fourier descriptors is the Euclidian distance D_{FCSI} between the proposed normalized Fourier descriptors of the query image F^q and the normalized Fourier descriptors of an image from the database F^d, as follows:

$$D_{\mathrm{FCSI}}(F^q,F^d)=\sqrt{\sum_{i=1}^{M}(f_i^q-f_i^d)^2} \tag{8}$$

The proposed descriptor has many advantages over the CSS descriptor. First, the proposed descriptor uses all of the information from the curvature-scale image while the CSS technique uses only the peaks of the CSS map. Second, the number of features of the proposed descriptor can be determined in advance, and hence, the number of features for different shapes can be equal. With the CSS technique, on the other hand, the number of features is not equal for different shapes. Third, the proposed descriptor possesses rotation invariance, so there is no need for the circular shift used in the matching stage of the CSS technique, which translates into considerable computational savings. With all of these advantages, the proposed descriptor clearly outperforms the CSS descriptor.

4. Combined curvature scale image

The proposed descriptors cannot represent a convex shape, because the CSI image is constructed for a concave shape only. To overcome this shortcoming, the procedure for mapping a convex shape to a concave shape proposed in [53] is applied:

$$D_{x(s),y(s)}=\sqrt{(x_c-x(s))^2+(y_c-y(s))^2} \tag{9}$$

$$x_m(s)=(x(s)-x_c)\cdot\frac{2\cdot R-D_{x(s),y(s)}}{D_{x(s),y(s)}}+x_c \tag{10}$$

$$y_m(s)=(y(s)-y_c)\cdot\frac{2\cdot R-D_{x(s),y(s)}}{D_{x(s),y(s)}}+y_c \tag{11}$$

where R: the circle that encloses the shape; (x_c,x_y): the center of the circle; $(x_m(s),y_m(s))$: the coordinates of the mapped shape

The CSI of the mapped shape and the original shape are combined into one image, referred to as the combined CSI image. The upper side of the combined CSI image represents the shape's concavities and the lower side represents the shape's convexities. The proposed features are derived from the combined CSI image using the procedure explained in Section 3.

Fig. 5 shows a convex shape (a square) and its mapped version along with the CSI images for both cases. It is clear that the combined CSI image makes it possible to distinguish between convex shapes such as a triangle and a square. Furthermore, it is noteworthy that for any convex shape, the upper side of the combined CSI image is represented by a blank image that has the same size as the lower side. The CSI for the shape shown in Fig. 4a, its CSI for the mapped version, and the combined version are shown in Fig. 6.

5. Comparative study

The theoretical properties of the proposed invariant descriptors are confirmed by the experimental results presented in this section. To obtain comprehensive results, the proposed descriptors were compared with six commonly used descriptors, the first two of which are the curvature scale space (CSS) and the radial angular transform descriptor (RATD) techniques. These techniques were used for the comparison because in the MPEG-7 standard, the curvature scale space (CSS) and radial angular transform descriptors have been adopted as the contour-based shape descriptor and the region-based shape descriptor, respectively [21]. The other comparative descriptors are the 1-D descriptor, Zernike moment descriptor (ZMD), radial Tchebichef moment descriptor (RTMD), 1-D Fourier descriptor (1-FD), and generic Fourier descriptor (GFD). These descriptors were adopted for the comparison because it has recently been shown that they are efficient and effective techniques for shape retrieval [9,17,40,41,54,55]. Furthermore, both 1-FD and GFD techniques are based on a Fourier transform as are the newly proposed descriptors. In the case of the 1-FD, the centroid signature was utilized because recent work has proven that the centroid distance signature outperforms other signatures for shape retrieval [41].

5.1. MPEG-7 databases

Due to the lack of a standard database, the evaluation of the different shape descriptors is not an easy task. Researchers in this field tend to develop their own databases, which are often limited in size or scope of application, or both. The MPEG-7 developers have set up a database of a reasonable size and generality [21]. It consists of three main sets: set A, set B, and set C. Set A consists of two subsets, A_1 and A_2, which are used to test robustness in scaling and rotation, respectively. Subset A_1 includes 420 shapes: 70 primary shapes and 5 shapes derived from each primary shape, with scale factors ranging from 0.1 to 2. Subset A_2 includes 420 shapes: 70 primary shapes and 5 subset shapes generated by rotating the primary shape with angles ranging from 9° to 150°. Sample shapes from set A of the MPEG-7 database are shown in Fig. 7.

Set B consists of 1400 images that are classified into 70 classes, each class having 20 images. Set B is used to test for similarity-based retrieval performance, and to test the shape descriptors for robustness to a variety of arbitrary shape distortions that include rotation, scaling, arbitrary skew, stretching, deflection, and indentation. Samples of shapes from set B of the MPEG-7 database are shown in Fig. 8.

Set C consists of 200 affine transformed Bream fish and 1100 marine fish that are unclassified. The 200 Bream fish are frames extracted from a short video clip of a bream fish swimming. This set

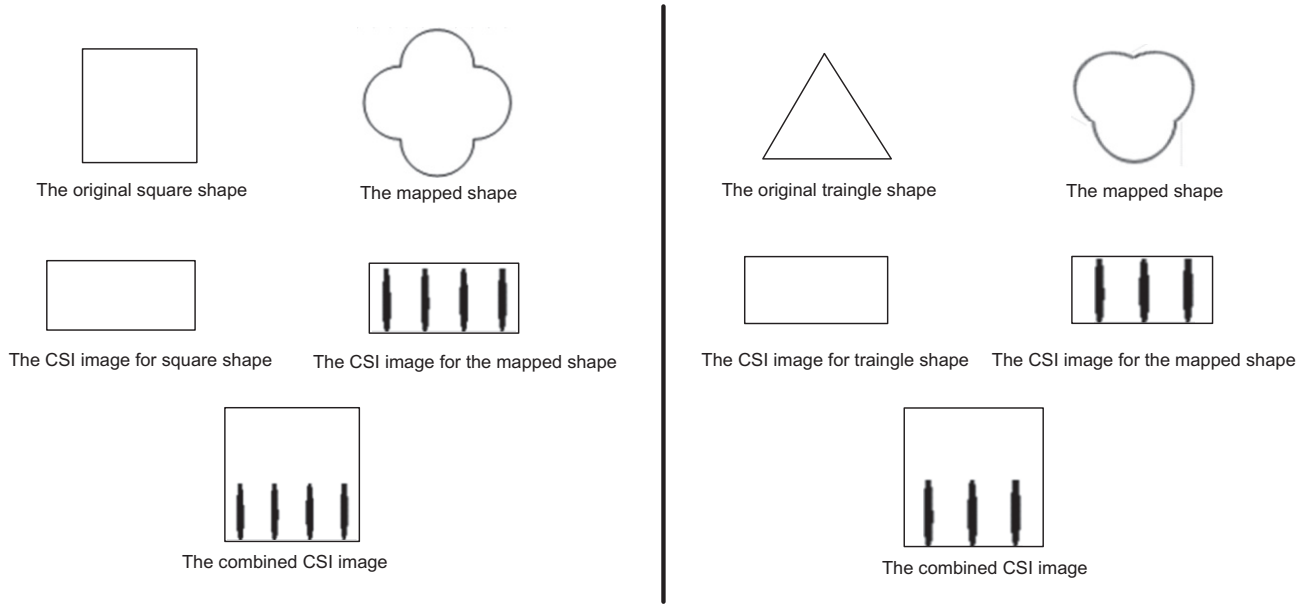
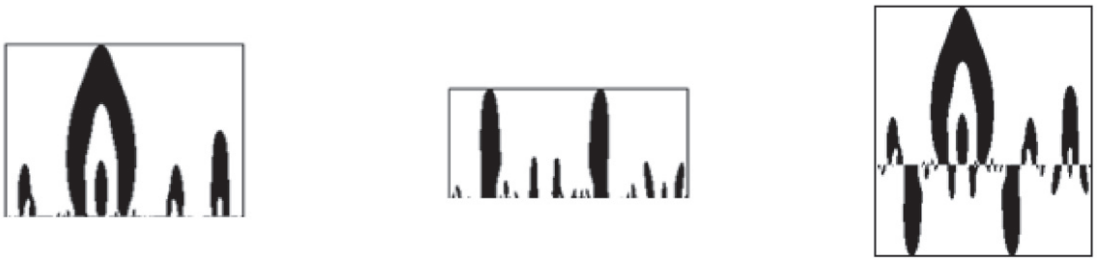


Fig. 5. The square and triangle shapes, their mapped shapes, the CSI images for the shapes, the CSI images for the mapped shapes, and the combined CSI images.



(a) CSI Image for the shape shown in Figure 3.a

(b) CSI Image for the mapped version of the shape shown in Figure 3.a.

(c) Combined CSI Image

Fig. 6. CSI for the shape shown in Fig. 3a, CSI for the transformed shape, and the combined CSI image.

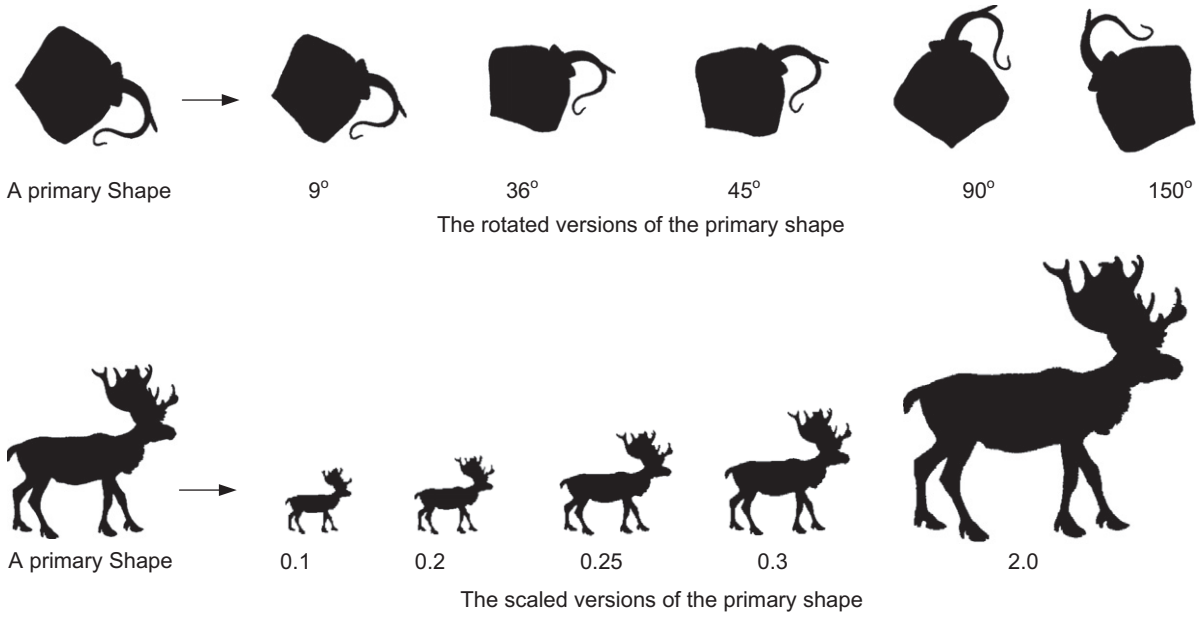


Fig. 7. Primary shapes from set *A* of the MPEG-7 database and their rotated and scaled versions.

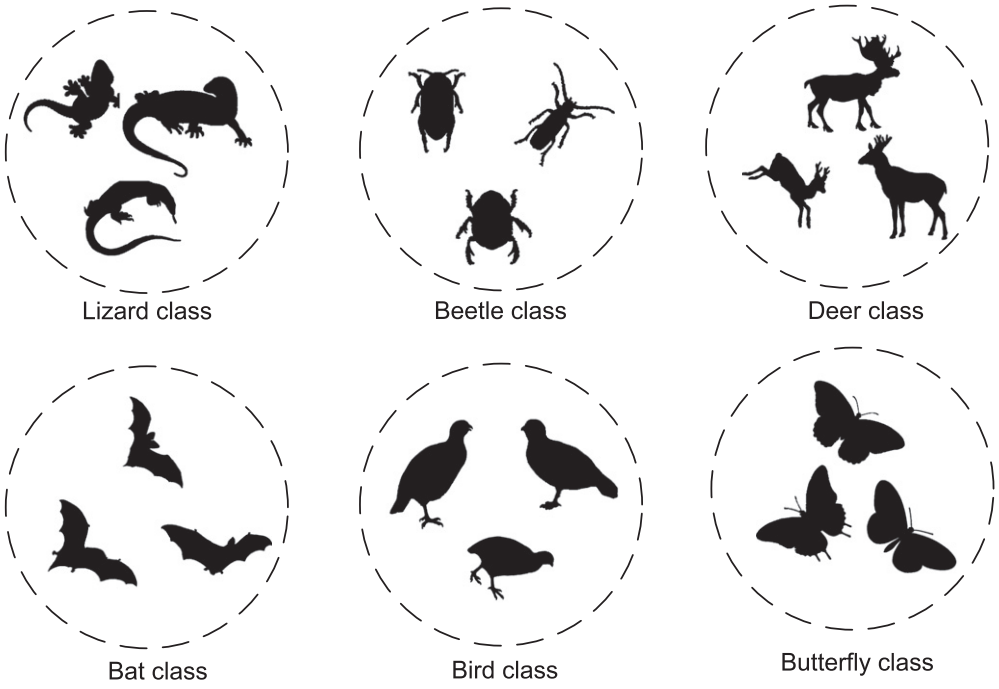


Fig. 8. Samples of shapes from set *B* of the MPEG-7 database.

is used to test shape descriptors for robustness to non-rigid object distortions. Usually the first frame of the video is used as a query, and the number of Bream shapes in the top 200 retrieved shapes is counted. However, in the experiment for this research, the 200 Bream fish are designated as queries in order to obtain a comprehensive comparison. Samples of the images from this database are depicted in Fig. 9.

In [48] a noisy database (set *D*) has been created to test the performance of the shape descriptors in presence of noise. This set consists of 420 shapes: 70 primary shapes and 5 shapes derived from each primary shape by adding random Gaussian noise to the boundary of the primary shape. The signal-to-noise ratios for the distorted shapes are 40, 35, 30, 25, and 20 dB. Sample shapes from set *D* database are shown in Fig. 10.

5.2. Evaluation measures

To evaluate the performance of the different techniques with respect to image retrieval, a performance measure is required.

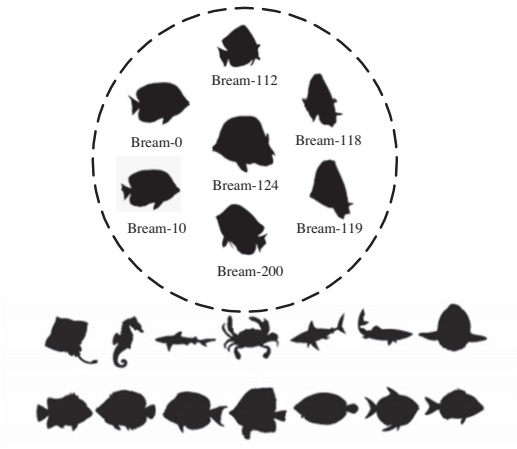


Fig. 9. Samples of shapes from set *C* of the MPEG-7 database.

The precision and recall measures are the most commonly used measures and were deemed appropriate for measuring retrieval performance for classified datasets. If *A* is the number of relevant retrieved shapes, *B* the total number of retrieved shapes, and *C* the number of relevant shapes in the whole database used, then precision and recall are defined as A/B and A/C, respectively. Precision measures the accuracy of the retrieval, whereas recall measures the ability to retrieve relevant items from the database [41]. The precision value of a specific recall is the average of the precision values of all the database shapes included in that recall.

5.3. Comparison of the proposed descriptors and other descriptors

The proposed descriptor (CBFD) and the CSS descriptors are not robust in a global sense. Consequently, the proposed descriptor and the CSS descriptors were combined with four simple global descriptors (SGD): eccentricity (E), aspect ratio (AR), circularity (C), and solidity (S) [21]. These simple global descriptors enhance the ability of the proposed descriptors to capture global shape information. The distance of the proposed descriptors (D_{CBFD}) is directly added to the average distance of the simple global descriptors (D_{SGD}). The total distance between a query shape (q) and a shape from the database (d) is expressed as follows:

$$D(q,d) = D_{CBFD}(q,d) + D_{SGD}(q,d) \tag{12}$$

$$D_{SGD}(q,d) = \frac{D_S(q,d) + D_C(q,d) + D_E(q,d) + D_{AR}(q,d)}{4} \tag{13}$$

$D_S(q,d) = \frac{|S^q - S^d|}{masx(S^d)}$ is the solidity distance, $D_C(q,d) = \frac{|C^q - C^d|}{masx(C^d)}$ is the circularity distance, $D_S(q,d) = \frac{|E^q - E^d|}{masx(E^d)}$ is the eccentricity distance, and $D_{AR}(q,d) = \frac{|AR^q - AR^d|}{masx(AR^d)}$ is the aspect ratio distance.

The number of features for the CSS is not constant, while the number of features for the proposed descriptors can be specified in advance. However, selecting a small number of features affects the retrieval performance of the proposed descriptors. On the other hand, if the selected number of features is too large, the system requires more processing time and more storage space. Since the

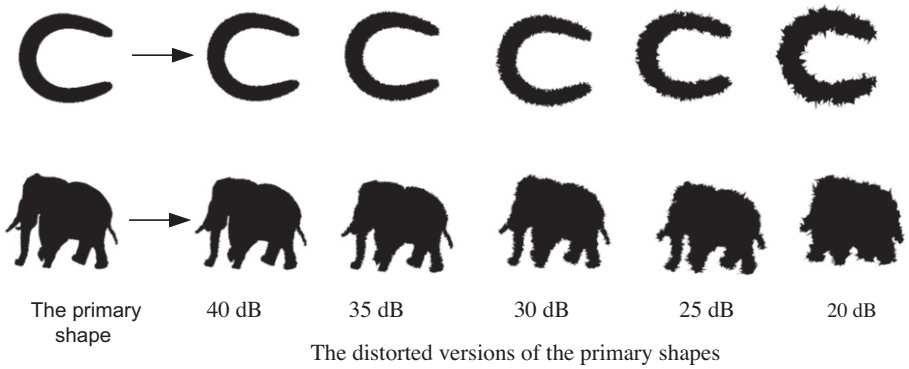


Fig. 10. Samples of shapes from set *D* of the proposed distorted database.

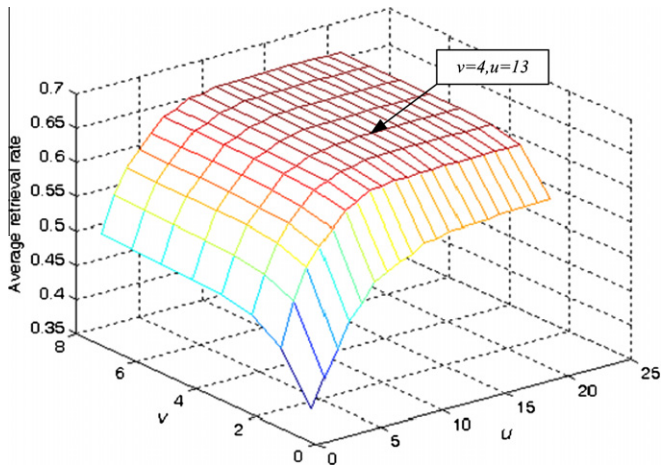


Fig. 11. The average retrieval rate for set B at different frequencies: u and v.

proposed descriptors derived from the 2-D Fourier transform there are two different parameters (frequencies) (u, v) that should be selected. To select theses parameters the average retrieval rates of the first top 20 shapes retrieved from set *B* of the MPEG-7 database were calculated at different frequencies and the result is shown in Fig. 11. The experiments revealed that the 52 descriptors ($v = 4$, $u = 13$ according to Eq. (6)) of the 2-D Fourier transform of the CSI result in a performance superior to that of other descriptors. Moreover, the experiments revealed that no significant improvement in the performance of the CBFD technique results using $v > 4$ and $u > 13$. The first descriptor is not included because it is used for scale normalization. Accordingly, the total number of descriptors used for the proposed technique is 55 (51 Fourier descriptors plus 4 simple global descriptors). For consistency and fair comparison, the number of descriptors is limited to 55 for all other descriptors.

The recall and precision curves using the four sets of the MPEG-7 database and the created noisy database are plotted in Figs. 12–16, and the average precision rates for the low and high recall are shown in Tables 1–5. The proposed descriptors outperform the CSSD, as can be seen in Figs. 12–16, and they produce results comparable to those of the CSSD in the case of the database of non-rigid object distortions. The reason for the improvement is that the proposed descriptor uses all of the information from the curvature scale image (CSI) while the CSS descriptor uses only the peaks of the CSS map. Fig. 12 shows that the proposed descriptors outperform all other descriptors with respect to both high and low recall. The CSSD, GFD and RTMD yield comparable results in cases of low and high recall; however, their results are not as accurate as that of the ZMD. The ARTD, ZMD, RTMD, GFD, and the 1-FD descriptors often misclassify shapes that are globally similar and have different local characteristics. The proposed descriptors, on the other hand,

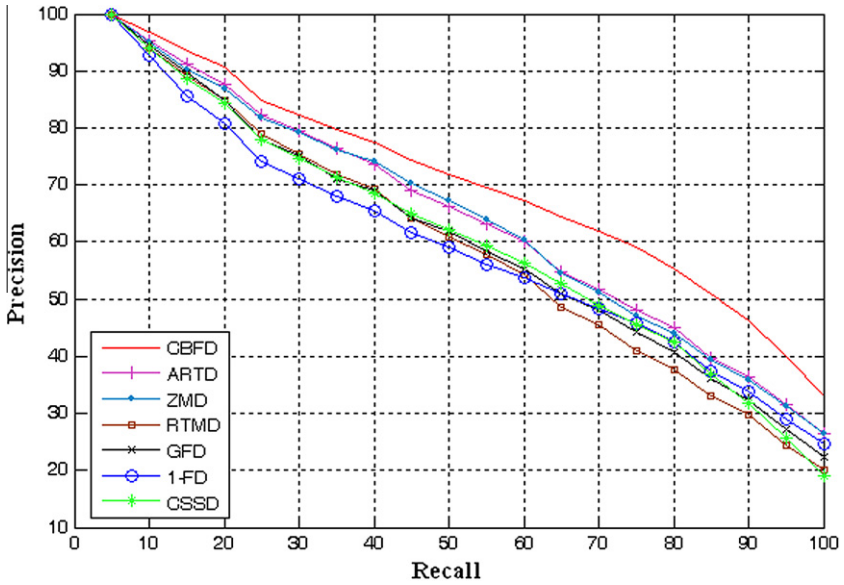


Fig. 12. Precision–recall curves of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *B*.

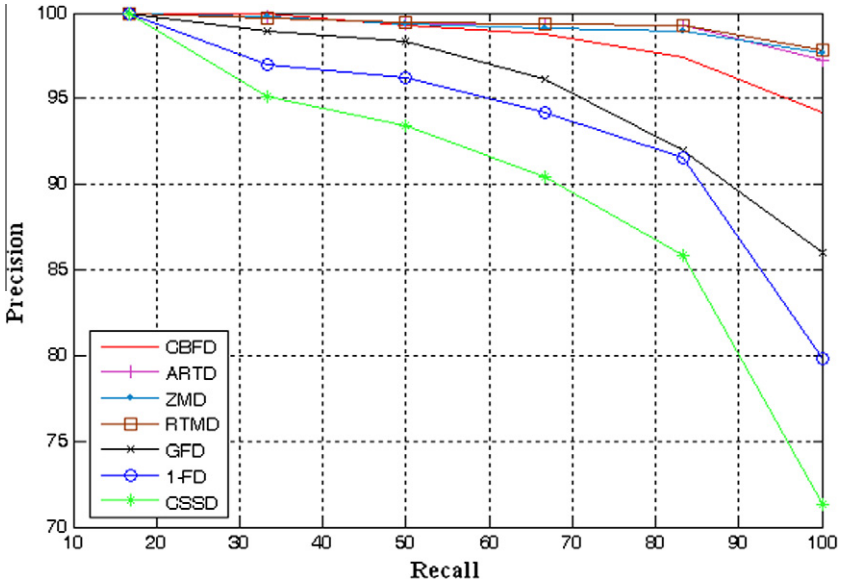


Fig. 13. Precision-recall curves of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set A_1.

have a strong structure for capturing both local and global features using the curvature of the shape contour and the simple global descriptors, respectively.

The performance of the proposed descriptors under scale invariance was tested using the database from subset A_1, as shown in Fig. 13. The RTMD give the highest accuracy for low and high recalls. In the case of low recall the proposed descriptors gives a result comparable to that of the ARTD, ZMD, and GFD. The GFD and the 1-FD give comparable results, which are considered very low in the case of high recall compared to the other descriptors. In this test, the region-based descriptors are better than the boundary-based descriptors because scaling of a shape changes the local characteristics of the shape and keep the global ones which are captured by the region-based descriptors.

In the case of the rotation invariance test, as shown in Fig. 14, all descriptors give a perfect result except the CSSD and 1-FD. The CSS technique has compact features; however, its matching algorithm is very complex and fails to distinguish between objects within the same class but having different rotations.

In the case of the database of non-rigid object distortions (set *C*), as shown in Figs. 15 and 16, all descriptors produce high accuracy for low recall. However, in the case of high recall, the results of the CBFD are not as accurate as those of the ARTD, ZMD, RTMD, GFD, and 1-FD because set C has many relevant shapes that are not visually close to the query shapes. Furthermore, it contains irrelevant shapes that are visually closer to the query shapes than the relevant shapes. For example, the first three shapes in the last row of Fig. 9 are more similar to the Bream-0 fish than to the Bream-118 and Bream-119 fish (in the first row) shown in the same figure. This observation has also been mentioned by Latecki et al. [56].

In the case of noisy shapes it is clear that the performance each technique decreases by reducing the signal-to-noise ratio as shown

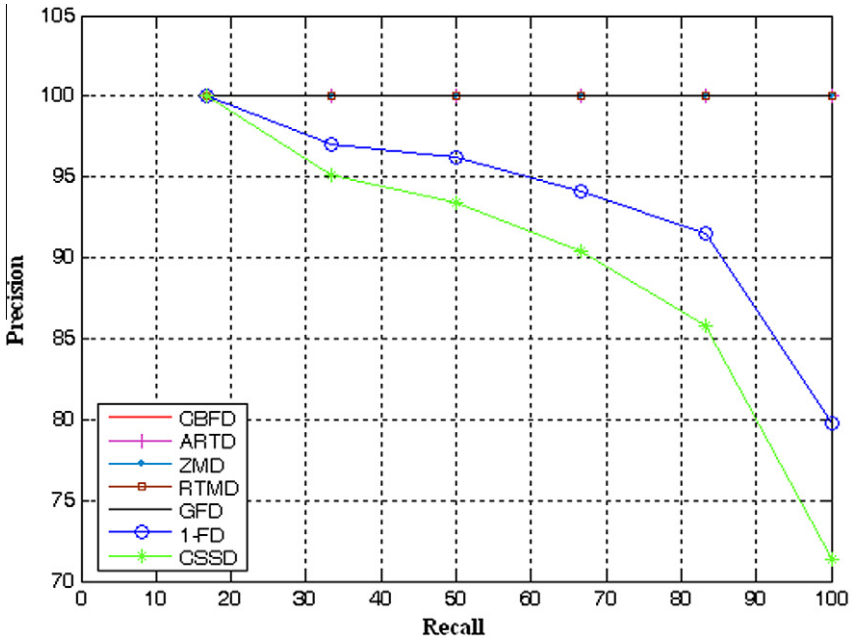


Fig. 14. Precision-recall curves of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set A_2.

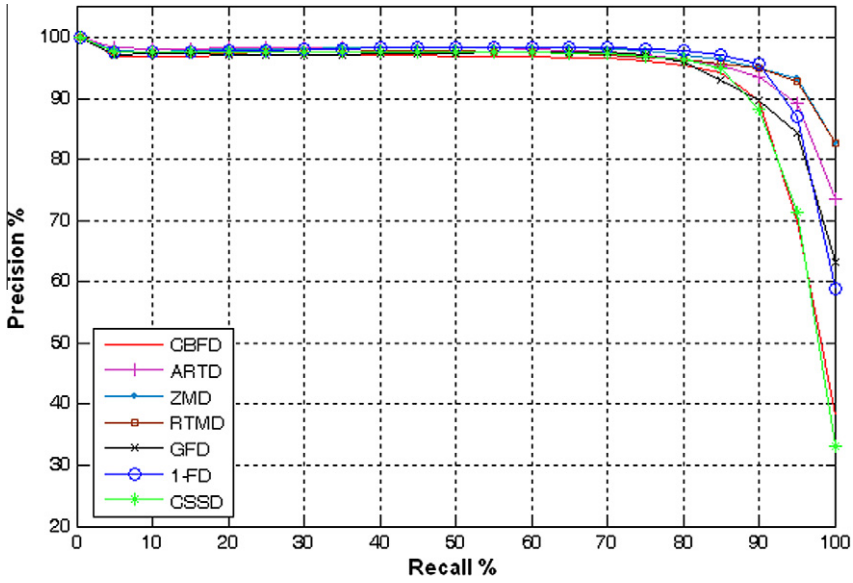


Fig. 15. Precision-recall curves of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *C*.

in Fig. 16. The proposed descriptors give the best performance in all cases. It is noteworthy that in the proposed technique the boundary noise is filtered because of the smoothing process during the construction of the CSI map. Moreover, the noise is filtered due to ignoring the higher order of the Fourier descriptors that are more sensitive to noise. These two processes of filtering render the proposed descriptors to outperform the other descriptors.

5.4. Overall performance using the MPEG-7 database

The overall average performance of the proposed descriptors was compared with the other descriptors based on the following weighted average:

$$\begin{aligned} \textit{Weighted Average} = {} & \frac{420}{2860} * A_1 + \frac{420}{2860} * A_2 + \frac{1400}{2860} * B \\ & + \frac{200}{2860} * C + \frac{420}{2860} * D \end{aligned} \tag{14}$$

The weights are calculated based on the number of query images used to obtain the recall–precision curve for each set. Table 6 shows the overall average performance of the CBFD descriptors compared to that of the ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD. Based on Table 6, it is clear that, compared with other descriptors; the proposed descriptors provide the highest performance in both low and high recalls.

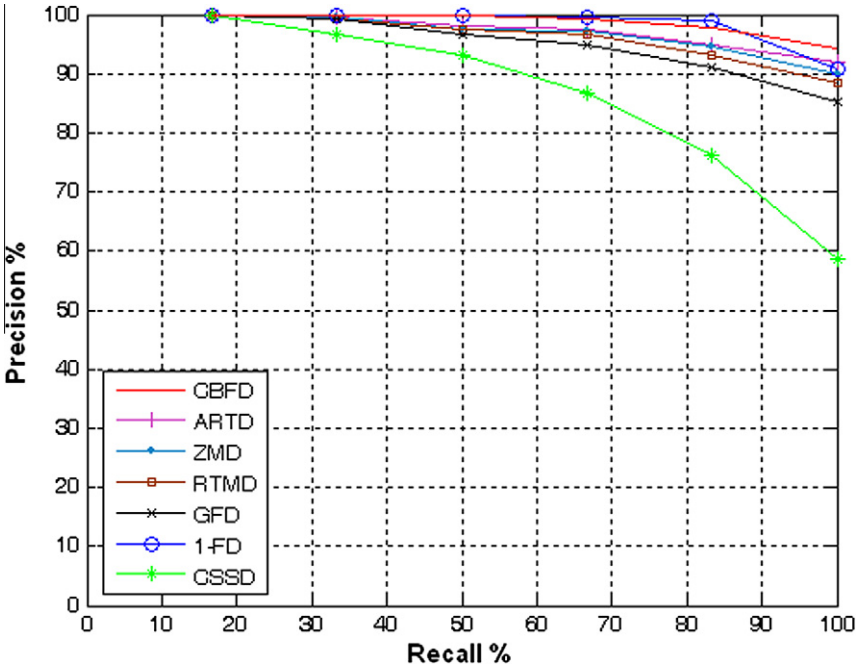


Fig. 16. Precision-recall curves of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *D*.

Table 1
The average precision rates of low and high recalls for the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *B*.

Method	Low recall (%) The average precision for recall rates ⩽ 50%	High recall (%) The average precision for recall rates > 50%
CBFD	85.15	54.77
ARTD	82.10	45.69
ZMD	82.11	45.42
RTMD	78.87	39.23
GFD	78.80	41.56
1-FD	75.83	42.25
CSSD	78.61	41.81

Table 2
The average precision rates of low and high recalls for the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using A_1.

Method	Low recall (%) The average precision for recall rates ⩽ 50%	High recall (%) The average precision for recall rates > 50%
CBFD	99.74	96.78
ARTD	99.71	98.62
ZMD	99.71	98.59
RTMD	99.77	99.38
GFD	99.11	91.39
1-FD	97.76	88.50
CSSD	96.18	82.52

Table 3
The average precision rates of low and high recalls for the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set A_2.

Method	Low recall (%) The average precision for recall rates ⩽ 50%	High recall (%) The average precision for recall rates > 50%
CBFD	100.00	100.00
ARTD	100.00	100.00
ZMD	100.00	100.00
RTMD	100.00	100.00
GFD	100.00	100.00
1-FD	100.00	98.35
CSSD	99.25	95.55

Table 4
The average precision rates of low and high recalls for the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *C*.

Method	Low recall (%) The average precision for recall rates ⩽ 50%	High recall (%) The average precision for recall rates > 50%
CBFD	97.38	88.03
ARTD	98.48	94.16
ZMD	98.32	95.85
RTMD	97.79	95.22
GFD	97.54	91.98
1-FD	98.15	93.34
CSSD	97.95	88.10

5.5. Complexity analysis

To compare the computational efficiency of the five descriptors, the processing time in the matching stage for set B of the MPEG-7 database was computed using the same processor and software for all descriptors. Matlab (version 7.0), running on a Pentium IV CPU 2.6 GHZ PC with 1.5 GB of memory was used as a testing platform. Table 7 shows the average processing time (the test was repeated 40 times) for each query in the matching stage for the proposed descriptors, as well as the ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD. The data displayed in Table 7 reveal that the CSSD require the highest time whereas the other descriptors have the same processing time because they all use the same matching procedure with the same number of features. The CSSD time is high compared to the proposed and the other descriptors because many factors must be considered in order to align two peaks of the CSSD features.

Table 5
The average precision rates of low and high recalls for the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD using set *D*.

Method	Low recall (%) The average precision for recall rates ⩽ 50%	High recall (%) The average precision for recall rates > 50%
CBFD	99.96	97.19
ARTD	99.75	94.75
ZMD	99.24	94.92
RTMD	98.99	92.83
GFD	98.66	90.38
1-FD	99.91	96.5
CSS D	96.65	73.77

Table 6
The overall average precision rates of the low and high recall of the CBFD, ARTD, ZMD, RTMD, GFD, 1-FD and the CSSD.

Method	Low recall The average precision for recall rates ⩽ 50%	High recall The average precision for recall rates > 50%
CBFD	92.50	76.14%
ARTD	91.05%	72.03
ZMD	90.97%	72.04
RTMD	89.32	68.77
GFD	89.12	68.16
1-FD	87.70	68.82
CSSD	88.22	63.61

Table 7
The average time required for each query in the matching stage using set *B* of the MPEG-7 database.

Descriptor	The average time required for each query
CBFD, ARTD, ZMD, RTMD, GFD, and 1-FD	1.73 (ms)
CSSD	2.16 (s)

6. Conclusions

This paper presents a new curvature-based Fourier descriptor for shape retrieval. The proposed descriptors have been evaluated against six commonly used descriptors. Several experiments based on the MPEG-7 database were conducted in order to confirm the theoretical properties of the proposed descriptors.

The experimental results demonstrate that the proposed descriptors (CBFD) outperform the CSS and 1-D Fourier descriptors for three sets of the MPEG-7 database, and they produce results comparable to those of the CSSD in the case of the database of non-rigid object distortions. Moreover, the CBFD descriptors outperform six notable descriptors in one of the most challenging databases (set *B*) for both high and low recall. The reason for this

superior performance is that the proposed descriptor is able to capture both the local and the global properties of an object, while the ARTD, ZMD, RTMD and the GFD tend to capture the global properties. The experimental results also show that the overall average performance of the proposed descriptors is better than that of all other descriptors.

The proposed descriptors overcome several shortcomings of the well-known CSS descriptors. The overall performance of the proposed descriptors is 12.53% higher than that of the CSSD for high recall. Furthermore, the proposed descriptors use a simple matching procedure that makes them efficient for retrieving images from huge databases.

The overall results reveal the effectiveness and efficiency of the proposed descriptors for image retrieval applications.

References

[1] K. Hirata, T. Kato, Query by visual example – content-based image retrieval, in: Third International Conference on Extending Database Technology, 1992, pp. 56–71.
[2] D. Mumford, Mathematical theories of shape: Do they model perception? in: SPIE Conference on Geometric Methods in Computer Vision, San Diego, California, vol. 1570, 1991, pp. 2–10.
[3] I. Biederman, Recognition-by-components: a theory of human image understanding, Psychological Review 94 (1987) 115–147.
[4] L. Schomaker, E.d. Leau, L. Vuurpijl, Using pen-based outlines for object-based annotation and image-based queries, Visual Information and Information Systems (1999) 585–592.
[5] R.H. Van Leuken, M.F. Demirci, V.J. Hodge, J. Austin, R.C. Veltkamp, Layout indexing of trademark images, in: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, 2007, pp. 525–532.
[6] S. Jan, P.E. John, C.V. Remco, Practice and challenges in trademark image retrieval, in: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, ACM, Amsterdam, The Netherlands, 2007.
[7] R.H. van Leuken, O. Symonova, R.C. Veltkamp, Topological and directional logo layout indexing using Hermitian spectra, in: Proceedings of the 5th Conference on Signal Processing, Pattern Recognition, and Applications 2008, pp. 93–98.
[8] Z. Wang, Z. Chi, D. Feng, Shape based leaf image retrieval, in: IEE Proceedings, Vision Image and Signal Processing, vol. 150, 2003, pp. 34–43.
[9] M.E. Celebi, Y.A. Aslandogan, A comparative study of three moment-based shape descriptors, in: International Conference on Information Technology: Coding and Computing, vol. 1, 2005, pp. 788–793.
[10] M. Hu, Visual pattern recognition by moment invariants, IRE Transactions on Information Theory IT-8 (1962) 179–187.
[11] J. Flusser, On the independence of rotation moment invariants, Pattern Recognition 33 (2000) 1405–1410.
[12] M. Teague, Image analysis via the general theory of moments, Journal of the Optical Society of America 70 (1980) 920–930.
[13] A. Khotanzad, Invariant image recognition by Zernike moments, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1990) 489–497.
[14] S.O. Belkasim, M. Shridhar, M. Ahmadi, Pattern recognition with moment invariants: a comparative study and new results, Pattern Recognition 24 (1991) 1117–1138.
[15] R. Mukundan, S.H. Ong, P.A. Lee, Image analysis by Tchebichef moments, IEEE Transactions on Image Processing 10 (2001) 1357–1364.
[16] R. Mukundan, Radial Tchebichef invariants for pattern recognition, in: IEEE Tencon Conference, 2005, pp. 2098–2103.
[17] D. Zhang, G. Lu, Shape-based image retrieval using generic Fourier descriptor, Signal Processing: Image Communication 17 (2002) 825–848.
[18] S. Li, M.-C. Lee, Effective invariant features for shape-based image retrieval, Journal of the American Society for Information Science and Technology 56 (2005) 729–740.
[19] A. Goshtasby, Description and discrimination of planar shapes using shape matrices, IEEE Transactions on Pattern Analysis and Machine Intelligence 7 (1985) 738–743.
[20] G. Lu, A. Sajjanhar, Region-based shape representation and similarity measure suitable for content based image retrieval, Multimedia Systems 7 (1999) 165–174.
[21] F. Mokhtarian, M. Bober, Curvature scale space representation: Theory application and MPEG-7 standardization, first ed., Kluwer Academic Publishers, 2003.
[22] T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, IEEE Transactions on Computers 21 (1972) 269–281.
[23] T.P. Wallace, P A Wintz, An efficient three dimensional aircraft recognition algorithm using normalized Fourier descriptors, Computer Graphics and Image Processing 13 (1980) 99–126.
[24] I. Kunttu, L. Lepisto, J. Rauhamaa, A. Visa, Multiscale Fourier descriptors for defect image retrieval, Pattern Recognition Letters 27 (2006) 123–132.
[25] F. Mokhtarian, A. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (1986) 34–43.
[26] S. Abbasi, F. Mokhtarian, J. Kittler, Curvature scale space image in shape similarity retrieval, Multimedia Systems 7 (1999) 467–476.
[27] S. Abbasi, F. Mokhtarian, J. Kittler, Enhancing CSS-based shape retrieval for objects with shallow concavities, Image and Vision Computing 18 (2000) 199–211.
[28] G. Chauang, C. Kuo, Wavelet descriptor of planar curves: theory and applications, IEEE Transactions on Image Processing 5 (1996) 56–70.
[29] R.B. Yadav, N.K. Nishchal, A.K. Gupta, V.K. Rastogi, Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: a comparative study, Optics and Lasers in Engineering 45 (2007) 695–708.
[30] T. Adamek, N.E. O'Connor, A multiscale representation method for nonrigid shapes with a single closed contour, IEEE Transactions on Circuits and Systems for Video Technology 14 (2004) 742–753.
[31] S. Junding, W. Xiaosheng, Chain code distribution-based image retrieval, in: International Conference on Intelligent Information Hiding and Multimedia Signal Processing, China, 2006, pp. 139–142.
[32] S.R. Dubois, F.H. Glanz, An autoregressive model approach to two-dimensional shape classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (1986) 55–65.
[33] Y. Tao, G. WI, Delaunay triangulation for image object indexing: a novel method for shape representation, in: 7th SPIE Symposium on Storage and Retrieval for Image and Video Databases, San Jose, CA, 1999, pp. 631–642.
[34] E.G.M. Petrakis, A. Diplaros, E. Milios, Matching and retrieval of distorted and occluded shapes using dynamic programming, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 1501–1516.
[35] N. Arica, F.T.Y. Vural, BAS: a perceptual shade descriptor based on the beam angle statistics, Pattern Recognition Letters 24 (2003) 1627–1639.
[36] I. Bartolini, P. Ciaccia, M. Patella, WARP: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2005) 142–147.
[37] L.J. Lateckia, R. Lakaempera, D. Wolter, Optimal partial shape similarity, Image and Vision Computing 23 (2005) 227–236.
[38] N. Alajlan, I.E. Rube, M.S. Kamel, G. Freeman, Shape retrieval using triangle-area representation and dynamic space warping, Pattern Recognition 40 (2007) 1911–1920.
[39] E. Persoon, K.S. Fu, Shape discrimination using Fourier descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (1986) 388–397.
[40] D.S. Zhang, G. Lu, A comparative study of curvature scale space and Fourier descriptors, Journal of Visual Communication and Image Representation 14 (2003) 41–60.
[41] D.S. Zhang, G. Lu, Study and evaluation of different fourier methods for image retrieval, Image and Vision Computing 23 (2005) 33–49.
[42] G. Eichmann, C. Lu, M. Jankowski, R. Tolimieri, Shape representation by Gabor expansion, in: Hybrid Image and Signal Processing II, vol. 1297, 1990, pp. 86–94.
[43] K. Arbter, Affine-invariant Fourier descriptors, in: From Pixels to Feature, Elsevier Science, Amsterdam, The Netherlands, 1989.
[44] K. Arbter, W.E. Snyder, H. Burkhardt, G. Hirzinger, Application of affine-invariant Fourier descriptors to recognition of 3-D objects, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1990) 640–647.
[45] A.E. Oirrak, M. Daoudi, D. Aboutajdin, Affine invariant descriptors using Fourier series, Pattern Recognition Letters 23 (2002) 1109–1118.
[46] I. Kunttu, L. Lepistö, Shape-based retrieval of industrial surface defects using angular radius Fourier descriptor, IET Image Processing 1 (2007) 231–236.
[47] A. El-Ghazal, O. Basir, S. Belkasim, A new shape signature for Fourier descriptors, in: The 14th IEEE International Conference on Image Processing San Antonio, TX, USA, 2007, pp. 161–164.
[48] A. El-ghazal, O. Basir, S. Belkasim, Farthest point distance: a new shape signature for Fourier descriptors, Signal Processing: Image Communication 24 (2009) 572–586.
[49] A. El-ghazal, O. Basir, S. Belkasim, A novel curvature-based Fourier shape descriptor, in: 15th IEEE International Conference on Image Processing (ICIP), 2008, pp. 953–956.
[50] F. Mokhtarian, S. Abbasi, J. Kittler, Efficient and robust retrieval by shape content through curvature scale space, in: International Workshop on Image DataBases and MultiMedia Search, 1996, pp. 35–42.
[51] D. Casasent, D. Psaltis, Position, rotation, and scale invariant optical correlation, Applied Optics 15 (1976) 1795–1799.
[52] D. Zhang, G. Lu, Evaluation of similarity measurement for image retrieval, in: Proceedings of the IEEE 2003 International Conference on Neural Networks and Signal Processing, vol. 2, 2003, pp. 928–931.
[53] S. Kopf, T. Haenselmann, W. Effelsberg, Enhancing curvature scale space features for robust shape classification, in: Proceedings of IEEE International Conference on Multimedia and Expo (ICME), Amsterdam, Netherlands, 2005, pp. 1–4.
[54] D. Zhang, G. Lu, Evaluation of MPEG-7 shape descriptors against other shape descriptors, Multimedia Systems 9 (2003) 15–30.
[55] S. Li, M.C. Lee, C.M. Pun, Complex Zernike moments features for shape-based image retrieval, IEEE Transactions on Systems, Man and Cybernetics, Part A 39 (2009) 227–237.
[56] L. J. Latecki, R. Lakamper, T. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, in: IEEE Conference Computer Vision and Pattern Recognition, 2000.