



Shape classification via image-based multiscale description

Cem Direkoğlu*, Mark S. Nixon

School of Electronics and Computer Science, University of Southampton, SO17 1BJ, UK

ARTICLE INFO

Article history:

Received 12 June 2010

Received in revised form

23 December 2010

Accepted 16 February 2011

Available online 23 February 2011

Keywords:

Shape classification

Fourier-based description

Multiscale representation

Gaussian filter

Feature extraction

Computer vision

ABSTRACT

We introduce a new multiscale Fourier-based object description in 2-D space using a low-pass Gaussian filter (LPGF) and a high-pass Gaussian filter (HPGF), separately. Using the LPGF at different scales (standard deviation) represents the inner and central part of an object more than the boundary. On the other hand using the HPGF at different scales represents the boundary and exterior parts of an object more than the central part. Our algorithms are also organized to achieve size, translation and rotation invariance. Evaluation indicates that representing the boundary and exterior parts more than the central part using the HPGF performs better than the LPGF-based multiscale representation, and in comparison to Zernike moments and elliptic Fourier descriptors with respect to increasing noise. Multiscale description using HPGF in 2-D also outperforms wavelet transform-based multiscale contour Fourier descriptors and performs similar to the perimeter descriptors without any noise.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Silhouette-based object description and recognition is an important task in computer vision. The descriptor must be invariant to size, translation and rotation, and it must be effective in adverse conditions such as noise and occlusion. There are two main types of shape description methods: boundary-based methods and region-based methods.

1.1. Boundary-based shape descriptors

In boundary-based methods only the boundary pixels of a shape are taken into account to obtain the shape representation. Boundary-based techniques have some limitations. First, they are generally sensitive to noise and variations of shape, since they only use boundary information. Second, in many cases, the object boundary is not complete with disjoint regions or holes. Region-based methods can overcome these limitations. The most common boundary-based shape descriptors are Fourier descriptors [1–4], wavelet descriptors [5], wavelet–Fourier descriptors [6–8] and curvature scale space (CSS) [9].

Shape representation using Fourier descriptors is easy to compute and robust. Fourier descriptors are obtained from the Fourier transform on a shape signature. The shape signature is a 1-D function that represents the shape derived from the boundary points of a 2-D binary image. Many shape signatures exist such as,

centroid distance, complex coordinates (position function), curvature and cumulative angle [10,11]. Geometric invariance can be achieved at the shape signature extraction stage or after the Fourier transform by normalizing Fourier coefficients appropriately, which depends on the choice of shape signature type. The lower frequency descriptors contain information about the general features of the shape and the higher frequency descriptors contain finer details of the shape.

Wavelet descriptors are derived from wavelet transform on a 1-D shape signature. The wavelet transform can be considered as a signal decomposition onto a set of basis functions. It has multiresolution, denoising and feature extraction capabilities. Chang and Kuo [5] used 1-D discrete periodized wavelet transform to describe shapes. However, the matching schema was more complicated than for Fourier descriptors. Kunttu et al. [6–8] introduce multiscale Fourier descriptors using wavelet and Fourier transforms. The multiscale contour Fourier descriptors are obtained by applying the Fourier transform to the coefficients of the multiscale complex wavelet transform.

McNeil and Vijayakumar [12] introduced perimeter and radial descriptors. In this work, shapes are represented by a large number of points from their boundaries. These points are selected at fixed intervals in terms of distance along the boundary (perimeter distance) or radial angle. Then, a probabilistic correspondence-based algorithm, which also incorporates with scale, translation and rotation invariance, is applied for shape matching. Later on, McNeil and Vijayakumar [13] improved their algorithm by segment-based shape matching, which can overcome limitations of global shape matching such as independent movement of parts or smooth deformations. There are also some recent

* Corresponding author. Tel.: +44 7785990782; fax: +44 2380594498.

E-mail address: cdirekoglu@googlemail.com (C. Direkoğlu).

advances in boundary-based shape description and classification techniques such as using the inner-distance [14], based on the contour flexibility [15] and based on the two perceptually motivated strategies [16].

Multiscale shape description is the most promising approach for recognition. Different features of the shape can be obtained at different scales and the combination of these features can increase discrimination power, so increasing the correct classification rate. In addition, it is more robust to noise since the dominant features are those that persist across scales. There are many boundary-based multiscale description techniques [6,9,17,18,19].

One of the most influential techniques is curvature scale space (CSS) introduced by Mokhtarian and Mackworth [9]. This method uses the scale space framework in 1-D space [20]. The boundary of a shape is filtered by LPGF of varying scales (standard deviation). For each specific scale, the locations of those curvatures zero crossings designated as one and otherwise as zero. The binary CSS image, which is generated by the location and scale in the horizontal and vertical axes, is used for matching.

Adamek and O'Connor [17] proposed a multiscale representation for a single closed contour that makes use of both concavities and convexities of all contour points. It is called multiscale convexity concavity (MCC) representation, where different scales are obtained by smoothing the boundary with LPGF of different scales. There are also other boundary-based multiscale description techniques such as graph-based approach [18] and triangle-area-based approach [19].

1.2. Region-based shape descriptors

In region-based methods, all the pixels within a shape are used to obtain the shape representation. Popular region-based shape descriptors include moments [4,21] and generic Fourier descriptors (GFDs) [22]. There are different types of moments and they can be classified as non-orthogonal and orthogonal moments depending on the basis function used. Geometric moments [23] are the first and simplest type of moments, which has been used for character recognition. They use non-orthogonal basis functions called a monomial. Low-order moments capture global description, whereas as the order increases, more detail is captured. The main problem with geometric moments is the high degree of information redundancy, because of a non-orthogonal basis function (monomials) is used. If the basis functions are orthogonal then each moment should highlight independent features. Teague [24] proposed Legendre moments that use Legendre polynomials as basis functions. These polynomials are orthogonal and cause Legendre moments to extract independent features within the image, with no information redundancy. This property also provides good reconstruction capability. These moments are based on Cartesian coordinates but the image function has to be mapped to a specific range of values. Zernike moments were also first proposed by Teague [24] and are based on the complex valued Zernike polynomials. These polynomials are defined in polar coordinates, which help to achieve rotation invariance. Zernike moments were found to be the best performing type of moment in image analysis and description task in terms of noise resilience, information redundancy and reconstruction capabilities [25].

Generic Fourier descriptors (GFD) [22] are other popular region-based shape descriptors. A 2-D Fourier transform is applied on a polar raster sampled shape image. The translational invariance is achieved due to using the shape centroid as origin in polar transform. The obtained polar Fourier coefficients represent translation and rotation invariant features. The scale invariance is

achieved by normalizing the polar Fourier coefficients. GFDs capture features of the shape in both polar and radial directions. GFDs are simple to compute and efficient.

Although many boundary-based multiscale description techniques exist, there is no region-based multiscale description technique in the image space. It is important to note that moments and GFDs are multiscale approach in the feature space, but not in the image space. In our work, we introduce image-based multiscale description using LPGF and HPGF, separately. The LPGF applies smoothing to the object and as scale (standard deviation) decreases, it causes loss of boundary and exterior regions. Therefore using the LPGF at different scales focuses on the inner and central part more than on the boundary of an object. On the other hand, using the HPGF at different scales emphasizes the boundary and exterior parts of an object more than the central part. Our algorithm is organized to achieve size, translation and rotation invariance. By classifying objects with the HPGF-based multiscale description, increase immunity to noise as well as increase correct classification rate is observed. Evaluation indicates that the HPGF-based multiscale representation performs better than the LPGF-based multiscale representation, and in comparison to Zernike moments and elliptic Fourier descriptors with respect to increasing noise. Multiscale description using HPGF in 2-D also outperforms wavelet transform-based multiscale contour Fourier descriptors and performs similar to the perimeter descriptors without any noise. Note that part of this work and the preliminary version were presented in [26,27], respectively. In this paper, we extend the basis and evaluation of the new multiscale shape description technique. We investigate and compare single scales (filtering at different scales) and average distance results of LPGF- and HPGF-based representation with respect to increasing noise in the MPEG-7 dataset [28]. We also compare the proposed descriptors with wavelet transform-based multiscale contour Fourier descriptors [6] and with the perimeter descriptors [12]. In addition, we experiment the proposed multiscale shape description on Swedish leaf dataset [29], which is a real and challenging dataset. We also have time evaluations for our model and for other models on two different databases.

2. Fourier-based description with multiscale representation in 2-D space

We produce multiscale Fourier-based object descriptors in 2-D space. For this purpose, we investigate the LPGF and the HPGF, separately. The new algorithm starts with size normalization of an object using bilinear interpolation in an image. We choose bilinear interpolation, since it scales better than nearest neighbor interpolation and it is faster than bicubic interpolation. The object size (the sum of intensities over the image) and the image size are determined experimentally depending on the database to locate each object in the image without any occlusion with image edges.

We also note that it is optional to centralize object in the image, since the next step is 2-D Fourier transform, as given in Eq. (1), which provides translation invariance:

$$\mathbf{FT}(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \mathbf{I}(x,y) e^{[-j2\pi((ux/M) + (vy/N))]} \quad (1)$$

where $\mathbf{FT}(u, v)$ is Fourier transform of the silhouette image $\mathbf{I}(x, y)$. $M \times N$ is the size of the silhouette image.

We also note that there is no “windowing” operation before the Fourier transform. The Fourier transform treats an image as it is part of a periodically repeated set of images extending horizontally and vertically to infinity, which can cause strong edges

between the neighbors of the periodic image. Therefore, the Fourier transform is the combination of the actual Fourier transform of the given image and that caused by the edge effects at image neighbors. These edge effects can be significantly reduced by using “windowing” operations, which in general makes image values zero towards edges. In our application, the given image is a pre-segmented object on a zero-valued background. Since the object does not occlude image edges, the image values are already zero towards image edges, and there is no need for a “windowing” operation.

In general, the result of the Fourier transform is a complex number and the transform can be represented in terms of its magnitude and phase. The magnitude describes the amount of each frequency component and the phase describes timing, when the frequency components occur. Here, we choose to use the Fourier magnitude image, which is translation invariant. However, the phase also carries considerable information that is discarded here. Oppenheim and Lim [30] showed that if we construct synthetic images from the magnitude information of one image and the phase information of another, we perceive mostly the image corresponding to the phase data. We leave investigation of the phase information as future work and continue with the magnitude information.

The computed Fourier magnitude image, $|\mathbf{FT}(u, v)|$, is translation invariant, however it retains rotation. Given the shift operation (the zero-frequency components are at the center), multiscale generation is achieved at this stage. To represent the inner and central part of an object more than the boundary, an LPGF with a selection of scale parameters (standard deviation) is applied to the Fourier magnitude image as shown below:

$$|\mathbf{FT}(u, v)|^s = |\mathbf{FT}(u, v)| (e^{-(u^2 + v^2)/2\sigma_s^2}) \quad (2)$$

where $|\mathbf{FT}(u, v)|^s$ and σ_s are Fourier magnitude and scale parameter of scale index s , respectively. This method is generating the scale space [20] of the object in 2-D as shown in Fig. 1. It is observed that the LPGF smooths the object and as scale decreases, it causes loss of the boundary and exterior regions. The LPGF emphasizes lower frequency components, but retains some contribution of higher frequency components.

On the other hand, to represent the boundary and exterior parts of an object more than the central part, an HPGF with a selection of scale parameters (standard deviation) is similarly applied to the Fourier magnitude image as shown below:

$$|\mathbf{FT}(u, v)|^s = |\mathbf{FT}(u, v)| (1 - e^{-(u^2 + v^2)/2\sigma_s^2}) \quad (3)$$

Filtering with the HPGF at different scales is illustrated in Fig. 2. It is observed that the HPGF detects the object boundary and as scale decreases, it represents exterior regions. The HPGF emphasizes higher frequency components, but retains a slight contribution of lower frequency components.

The obtained Fourier magnitude images are not convenient for matching at this stage, since they still vary with rotation. To remove rotation variance, the coordinates of each Fourier magnitude image are polar mapped to make rotations appear as translations in the new image.

Consider the polar coordinate system (r, θ) , where $r \in \mathbb{R}$ denotes radial distance from the center of the Fourier magnitude image (x_c, y_c) and $0 \leq \theta \leq 2\pi$ denotes angle. Any point $(x, y) \in \mathbb{R}^2$ can be represented in polar coordinates as follows:

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \\ \theta = \tan^{-1} \left(\frac{y - y_c}{x - x_c} \right) \quad (4)$$

Eq. (4) describes conversion from Cartesian to polar coordinates. The reverse process, which is the polar to Cartesian coordinates transform, is defined below:

$$x = r \cos(\theta) \\ y = r \sin(\theta) \quad (5)$$

For every point (x, y) , there is a unique point (r, θ) . Rotating the Cartesian coordinate system about an origin, while preserving position and size, can be written with the following matrix notation:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad (6)$$

where, (x_1, y_1) is the point before rotation and (x_2, y_2) is the point after rotation by the angle φ . Assuming that $(x_1, y_1) = (r \cos(\theta), r \sin(\theta))$ and after substitution to Eq. (6), we obtain the new

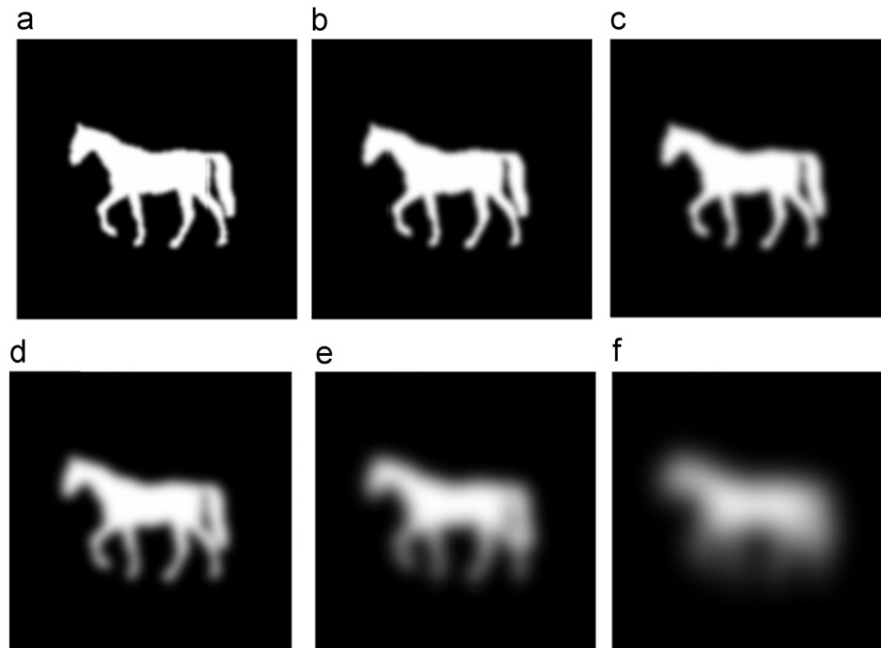


Fig. 1. Horse object filtered by LPGF with respect to decreasing scale: (a) $\sigma_1=20$, (b) $\sigma_2=15$, (c) $\sigma_3=11$, (d) $\sigma_4=8$, (e) $\sigma_5=5$ and (f) $\sigma_6=3$.

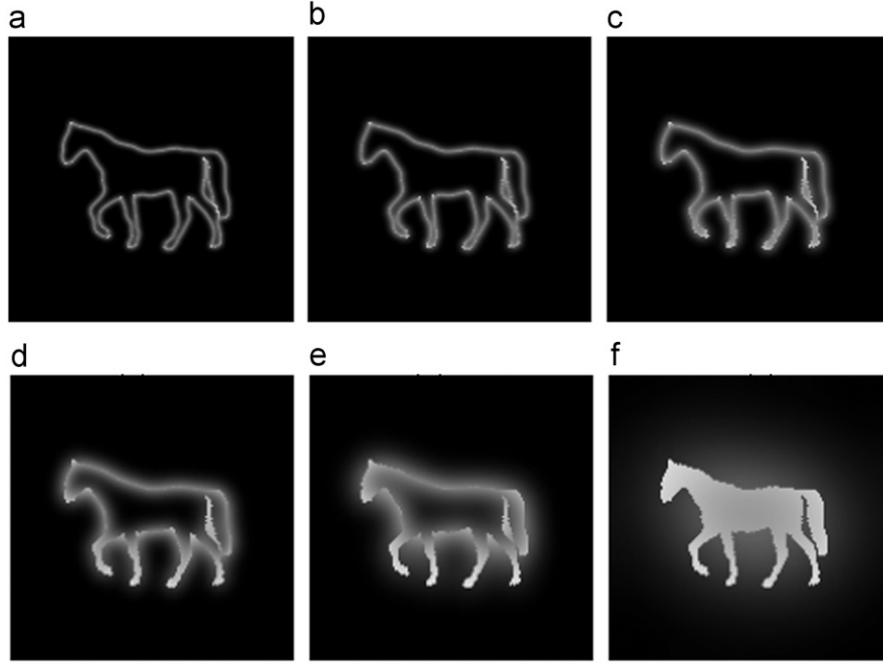


Fig. 2. Horse object filtered by HPGF with respect to decreasing scale: (a) $\sigma_1=15$, (b) $\sigma_2=11$, (c) $\sigma_3=8$, (d) $\sigma_4=5$, (e) $\sigma_5=3$ and (f) $\sigma_6=1$.

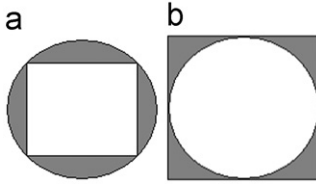


Fig. 3. Alternative approaches for mapping a square image to the circle: (a) Fitting the image into the circle, where the shaded area shows parts of the circle ignored in the mapping process and (b) fitting the circle to the square image, where shaded areas represent parts of the image lost in mapping.

coordinates as

$$\begin{aligned} x_2 &= r \cos(\theta + \varphi) \\ y_2 &= r \sin(\theta + \varphi) \end{aligned} \quad (7)$$

Here, we can observe that rotation in Cartesian coordinates causes translation in polar coordinates,

$$\begin{aligned} (x_1, y_1) &\leftrightarrow (r, \theta) \\ (x_2, y_2) &\leftrightarrow (r, \theta + \varphi) \end{aligned} \quad (8)$$

There are two principal methods for mapping a rectangular image to a circle in polar transform. The image can either be fitted within the circle as shown in Fig. 3(a) or the circle can be fitted within the boundaries of the image as shown in Fig. 3(b). The main problem with fitting circle within the boundaries of the image is losing the information in the corners. Since we want to use all information in the Fourier magnitude image, we use the method that fits the image within a circle. In this method, all pixels will be taken into account but some invalid pixels will also be included, which fall inside the circle but outside the image. In our algorithm these invalid pixel values are set to zero. Fig. 4 shows the polar transform of a Fourier magnitude image.

Finally, another 2-D Fourier transform is applied, as given in Eq. (9), to compute Fourier magnitude, which removes these translations:

$$\mathbf{FPT}^s(k, l) = \frac{1}{EF} \sum_{r=0}^{C-1} \sum_{\theta=0}^{E-1} \mathbf{P}^s(r, \theta) e^{-j2\pi((kr/C) + (l\theta/E))} \quad (9)$$

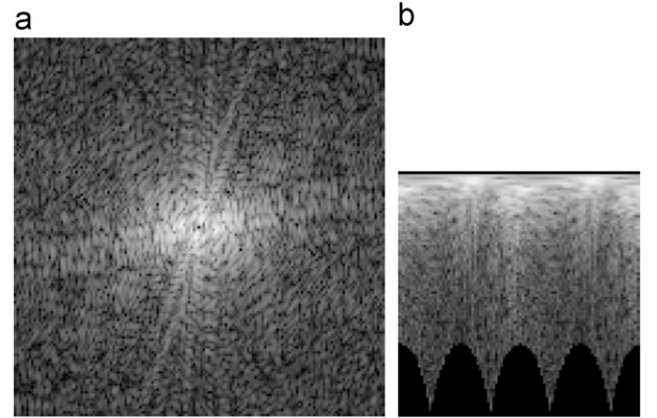


Fig. 4. Cartesian to polar transform with fitting the image into the circle: (a) Fourier magnitude image of the horse object filtered by HPGF ($\sigma=3$) and the image size is 151×151 and (b) polar transformed Fourier magnitude image of size 90×90 , the invalid pixels are zero.

where $\mathbf{FPT}^s(k, l)$ is the Fourier transform of the polar mapped image $\mathbf{P}^s(r, \theta)$ of size $C \times E$ and at scale index s . Note that there is no “windowing” operation before the Fourier transform. Although it can remove edge effect between the neighbors of the periodic image, it may also cause losses of some important information in the polar mapped image.

The resultant Fourier magnitude image, $|\mathbf{FPT}(k, l)|^s$, is translation, size and rotation invariant and represents object descriptors \mathbf{OD}^s of a shape at scale index s . Fig. 5 shows the proposed algorithm to obtain multiscale Fourier-based object descriptors.

The Fourier–Mellin transform is similar to our algorithm in terms of achieving rotation, size and translation invariance. The Fourier–Mellin transform is a method for rotation, size and translation invariant image feature extraction in 2-D space [31]. The first stage is a 2-D Fourier transform to calculate the Fourier magnitude image ($|\mathbf{FT}|$), which removes translation variance while keeping scale and rotation variances, then the coordinates are log-polar transformed (LPT) to make scaling and rotation

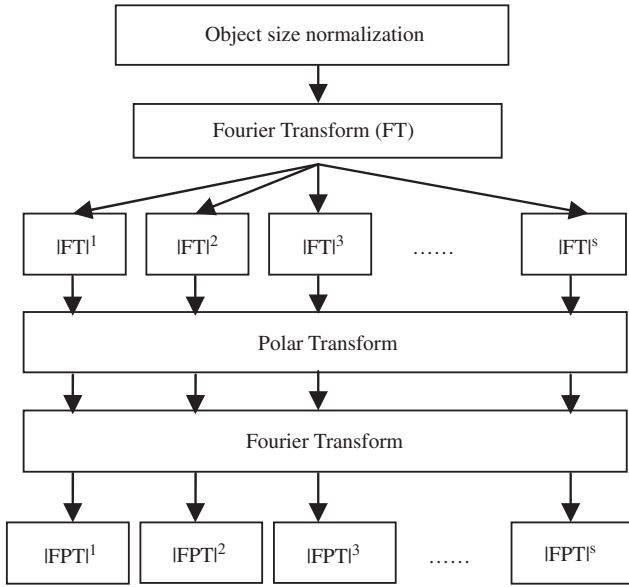


Fig. 5. Producing the proposed multiscale Fourier-based object descriptors.

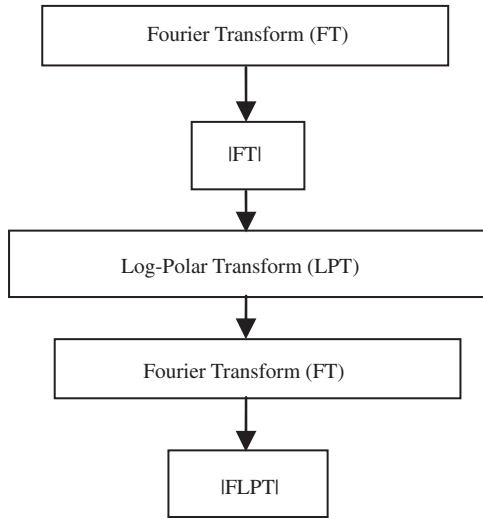


Fig. 6. Fourier–Mellin transform to produce rotation, size and translation invariant image features.

appear as translations, and finally another 2-D Fourier transform is applied to compute Fourier magnitude image ($|FLPT|$), which remove these translations. Fig. 6 shows the Fourier–Mellin transform to obtain rotation, size and translation invariant image features. In the log-polar transform, converting scale change to translation is achieved by logarithmic scaling the radius coordinate of the polar map image [32]. The difference from our new approach is that we now have a filtering approach to create a multiscale representation, which must be applied to the objects of the same size. Because of this, object size is normalized in the first step and we do not apply logarithmic scaling to the radius coordinate of the polar transformed image.

3. Classification with multiscale Fourier-based description

Classification is achieved using the nearest neighbor algorithm. We are using a standard approach to allow comparison, though other classifiers are equally appropriate. Euclidean distance (Ed) is used to measure similarity between objects and is

computed separately in each scale as given below:

$$Ed^s(\mathbf{T}, \mathbf{D}) = \sqrt{\sum_{x=1}^C \sum_{y=1}^E (\mathbf{OD}_T^s(x, y) - \mathbf{OD}_D^s(x, y))^2} \quad (10)$$

where $Ed^s(\mathbf{T}, \mathbf{D})$ is the Euclidean distance between the object descriptors, \mathbf{OD}_T^s , of the test image \mathbf{T} and object descriptors \mathbf{OD}_D^s of an image from database \mathbf{D} , at scale index s . Then average distance (Ad) is computed for each object:

$$Ad = \frac{1}{Y} \sum_{s=1}^Y Ed^s \quad (11)$$

where Ad represents average distance and Y is the number of scales. Classifying with average distance, instead of single scale distance, increase correct classification as well as increase immunity to noise.

4. Evaluations and experimental results

For evaluation, we use MPEG-7 CE-Shape-1 Part B database [28] and Swedish leaf database [29]. MPEG-7 CE-Shape-1 Part B is a commonly used dataset in shape classification experiments, which consists of shapes acquired from real world objects. On the other hand, the Swedish leaf database is a real and challenging dataset due to their high between class similarity and large inner class deformations. Computational time evaluations are also conducted on these two datasets, which is presented in this section.

4.1. Evaluation on MPEG-7 shape database

There are 1400 images in MPEG-7 CE-Shape-1 Part B dataset [28], which are pre-segmented and are in binary form. The objects are divided into 70 classes with 20 images in each class. The object classes are shown in Fig. 7. The appearance of these silhouettes changes due to

- viewpoint with respect to objects (size, translation and rotation variance),
- non-rigid object motion (e.g. people walking and fish swimming),
- noise inside shape (e.g. digitization and segmentation noise).

Some object variations are shown in Fig. 8. Leave-one-out cross-validation is applied to validate classification. The correct classification rate (CCR%) is measured as follows:

$$CCR(\%) = \frac{c_o}{t_o} \times 100 \quad (12)$$

where c_o is the total number of correctly classified objects and t_o is the total number of classified objects.

In evaluation, first we investigate and compare single scales (filtering at different scales) and average distance (with the method given in Section 3) results of LPGF- and HPGF-based representation without any noise in silhouette images. Single scales and average distance results are also compared with the original result, where the original result represents the classification result without any filtering operation. Second, we experiment with the original, single scales and average distance performances with respect to increasing noise in the dataset. Finally, LPGF- and HPGF-based multiscale description (average distance performances) are compared with other object description techniques.



Fig. 7. A sample from each object class in the database.

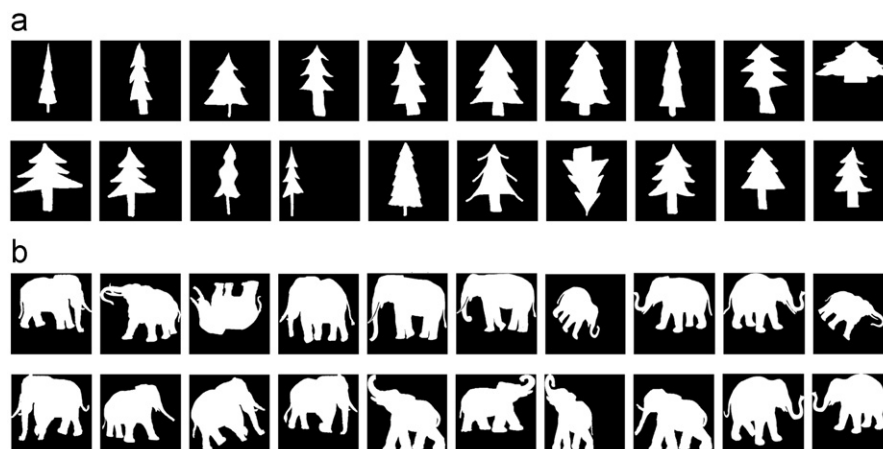


Fig. 8. Some objects variations: (a) tree and (b) elephant.

4.1.1. Original, single scales and average distance results without any noise in the database

We analyze the original (without any filtering operation), single scales and average distance performances of LPGF- and HPGF-based multiscale description without adding any noise to the database. We also remove existing noise in the database by filling object region (using morphological flood-fill operation), since there is noise only inside shapes.

In a multiscale description using LPGF, the object size is normalized to be 2500, which is the sum of intensities over the image, in a 151×151 size image. Five different scales are selected for multiscale representation. The selected scales are: $\sigma_1=20$, $\sigma_2=15$, $\sigma_3=11$, $\sigma_4=8$ and $\sigma_5=5$. The size of the object descriptor matrix is 90×90 at each scale. These five-scale values are determined experimentally to achieve the best performance of the proposed algorithm with LPGF. Note that as the number of scales increases, the computational complexity increases.

In a multiscale description using HPGF, the object size is similarly normalized to be 2500, which is the sum of intensities over the image, in a 151×151 size image. Five different scales are selected for multiscale representation. The selected scales are: $\sigma_1=11$, $\sigma_2=8$, $\sigma_3=5$, $\sigma_4=3$ and $\sigma_5=1$. The size of the object descriptor matrix is 90×90 at each scale. These five-scale values are also determined experimentally to achieve the best performance of the proposed algorithm with HPGF. As the number of scales increases, the computational complexity increases.

Table 1 shows the CCR% of the original, selected single scales using LPGF and average distance of selected scales. It is observed that the highest CCR% is achieved with the original that is without applying any LPGF. The CCR% of the original is 92.6% and as we apply LPGF with decreasing scales, which means as the objects become smoother, CCR% decreases. Taking average distances from these selected scales, with the method given in Section 3, results with 91.1%. This is not higher than the original result and some single scale results. Therefore using LPGF is not effective, when there is no noise in the database.

Table 2 similarly shows the CCR% of the original, selected single scales using HPGF and average distance of selected scales. It is observed that applying HPGF with scales $\sigma_3=5$, $\sigma_4=3$ and $\sigma_5=1$ perform better than the original (92.6%). The highest CCR% is 95% among the single scale results and is achieved at scale $\sigma_4=3$. This is the scale that represents the exterior parts of the

object more than the boundary and the central part. The scales $\sigma_5=1$ and $\sigma_3=5$ give exactly the same result (93.9%). After $\sigma_3=5$, as scale increases, the CCR% decreases. This is because we start to focus more on the boundary alone, which is more sensitive to shape variations. Averaging the distances of these five scales, which represents the boundary and exterior parts of an object more than the central part, even increase CCR% more and makes it 95.7%.

4.1.2. Original, single scales and average distance results with added noise in the database

We experiment with the original (without any filtering operation), single scales and average distance performances with respect to increasing salt and pepper noise in the database. Fig. 9 illustrates salt and pepper noise corrupted binary images with increasing density. In this evaluation, we do not remove the existing noise in the database as well (no region filling). Although some objects in the dataset contain noise inside the shape, adding salt- and pepper-type noise cause noise outside the shape as well. Salt and pepper noise is added to all objects in the database; therefore noisy test image is matched with the noisy images from database. It is also important to note that the noise is added after the object size normalization stage.

Table 3 and Fig. 10 show the CCR% of the original images, LPGF-filtered images at different scales and the average distance of these scales. The results represent mean values obtained over four applications of each scale at each noise level. In Fig. 10, the error bar represents minimum and maximum values at the data points. It is simpler to follow our explanations from the table, since the obtained results are very close to each other and cannot be seen well in the figure. It is observed that when there is no noise or small amounts of noise such as $D=0.1$ and $D=0.2$, applying LPGF at selected scales does not increase CCR% in comparison to the original. Even averaging the distances with selected scales does not become effective. When there is noise more than $D=0.2$, applying LPGF at higher selected scales ($\sigma_1=20$, $\sigma_2=15$ and $\sigma_3=11$) increase CCR% slightly. Averaging the distances from these selected scales, at noise levels $D=0.3$ and $D=0.4$, increases the original as well as the single scales performances slightly. However, at noise levels $D=0.5$ and $D=0.6$, we do not observe any increased performance by average

Table 1
CCR% of the original, single scales using LPGF and average distance using LPGF.













LPGF	Original	$\sigma_1=20$	$\sigma_2=15$	$\sigma_3=11$	$\sigma_4=8$	$\sigma_5=5$	Average distance
CCR%	92.6%	92.2%	91.7%	91.4%	90.2%	88.5%	91.1%
							

Table 2
CCR% of the original, single scales using HPGF and average distance using HPGF.

HPGF	Original	$\sigma_1=11$	$\sigma_2=8$	$\sigma_3=5$	$\sigma_4=3$	$\sigma_5=1$	Average distance
CCR%	92.6%	92.2%	92.4%	93.9%	95%	93.9%	95.7%
							

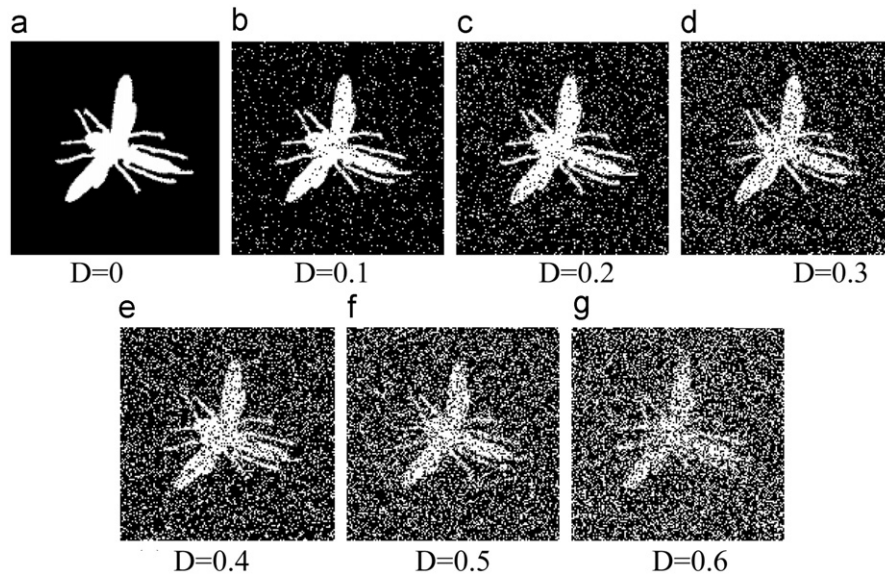


Fig. 9. Fly object with increasing density (D) of salt and pepper noise.

Table 3

CCR% of the original, single scales using LPGF and average distance using LPGF with respect to the increasing density of salt and pepper noise.

LPGF	Salt and pepper noise density (D)						
	0	0.1	0.2	0.3	0.4	0.5	0.6
Original	92.6%	89.5%	86.1%	77.4%	69.2%	57.7%	41.5%
$\sigma_1=20$	91.7%	88.1%	84.0%	78.7%	69.2%	59.4%	44.2%
$\sigma_2=15$	91.2%	89.4%	83.5%	78.2%	69.7%	58.7%	44.5%
$\sigma_3=11$	90.7%	89.0%	82.5%	79.0%	69.6%	57.8%	43.7%
$\sigma_4=8$	89.4%	86.1%	81.1%	76.5%	65.5%	56.1%	39.7%
$\sigma_5=5$	88.1%	84.3%	77.2%	69.2%	58.7%	48.1%	34.4%
Average distance	90.4%	89.5%	85.2%	80.5%	70.3%	55.0%	41.5%

Table 4

CCR% of the original, the single scales using HPGF and the average distance using HPGF with respect to the increasing density of salt and pepper noise.

HPGF	Salt and pepper noise density (D)						
	0	0.1	0.2	0.3	0.4	0.5	0.6
Original	92.6%	89.5%	86.1%	77.4%	69.2%	57.7%	41.5%
$\sigma_1=11$	91.7%	59.1%	25.4%	10.2%	3.7%	2.6%	2.4%
$\sigma_2=8$	92.5%	78.1%	48.4%	23.0%	11.2%	5.2%	2.5%
$\sigma_3=5$	93.8%	89.3%	77.4%	56.7%	34.7%	15.2%	4.7%
$\sigma_4=3$	94.7%	91.7%	88.8%	80.7%	67.5%	45.7%	21.8%
$\sigma_5=1$	93.7%	92.1%	90.9%	86.6%	80.3%	69.0%	52.3%
Average distance	95.5%	93.6%	92.2%	88.5%	82.0%	71.3%	52.0%

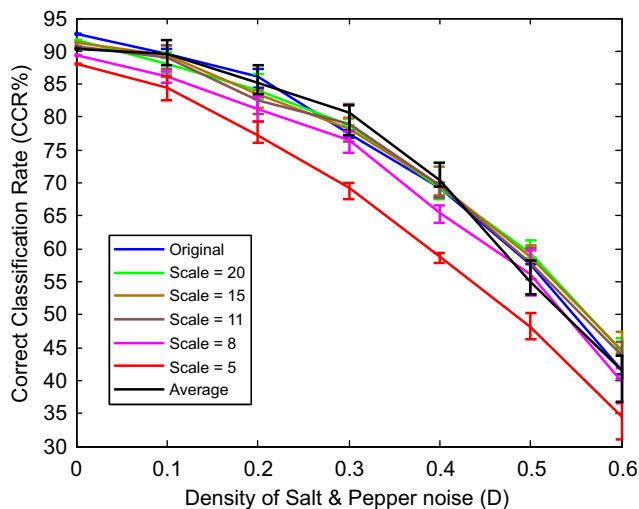


Fig. 10. Classification performance of the original, the single scales and the average distance by LPGF-based representation with respect to increasing salt and pepper noise in the database. CCR% is plotted with minimum and maximum values using error bars.

distance in comparison to the original and some single sales (higher scales).

Table 4 and Fig. 11 show the CCR% of the original image, the single scales using HPGF and the average distance using HPGF-based representation, with respect to increasing density of salt

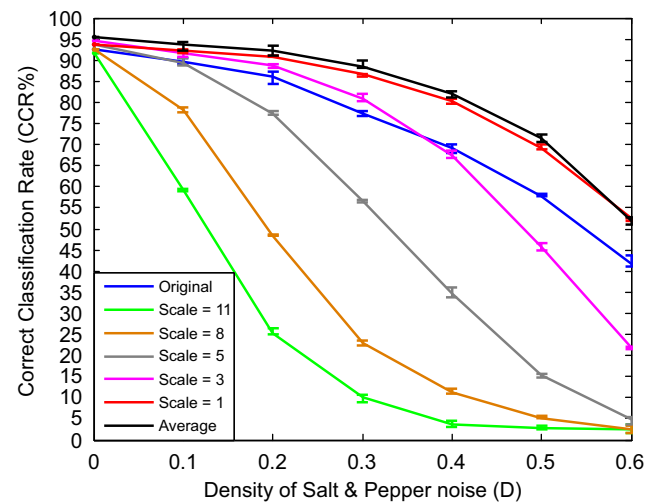


Fig. 11. Classification performance of the original, the single scales and the average distance by HPGF-based description with respect to increasing salt and pepper noise in the database. CCR% is plotted with minimum and maximum values using error bars.

and pepper noise. The results represent mean values obtained over four applications of each scale at each noise level. In Fig. 11, the error bar represents minimum and maximum values at the data points. It is observed that when $D=0$, lower scales ($\sigma_3=5$, $\sigma_4=3$ and $\sigma_5=1$) perform better than the original (92.6%).

The best single scale result is achieved at $\sigma_4=3$, which is 94.7%. This scale represents the exterior regions of an object more than the boundary and the central part. Averaging the distances of the selected scales also improves the CCR% (95.5%). When we add salt and pepper noise with increasing density, average distance always performs better than the original and the single scales. Only at $D=0.6$, which is very noisy and objects are not visible, the scale $\sigma_5=1$ performs slightly better than average distance. The scale $\sigma_5=1$ also performs better than the original at all noise levels. The scale $\sigma_4=3$, which achieves the best result without any added noise, performs better than original until $D=0.3$. The performances of the higher selected scales goes down faster than lower selected scales, since the higher selected scales represent the boundary more than the exterior parts and the central part, and more sensitive to noise and shape variations.

Applying HPGF at selected scales and computing average distance improve CCR% in the dataset. This result occurs because of representing more the boundary and the exterior parts, which are more discriminative, than the central part.

4.1.3. Comparison with other techniques

Performance evaluation is also employed by comparing the multiscale description using LPGF (average distance) and multiscale description using HPGF (average distance) with each other as well as with elliptic Fourier descriptors (EFD) and Zernike moments (ZM). The evaluation is again achieved with respect to increasing salt and pepper noise in the database, and the noisy test image is matched with the noisy images from the database.

EFD are fast and robust boundary-based shape descriptors. The contour is represented with complex coordinates (position function) and then the Fourier expansion is performed to obtain the

EFD, where the number of descriptors is 80 in this evaluation. To evaluate EFD, we use a Matlab implementation given in [33], and note that this is a non-optimal Matlab framework. We describe the boundary of the biggest region in the image, since there will be many regions after noise has been added.

Zernike moments (ZM) are region-based shape descriptors. They are an orthogonal moment set, which makes optimal utilization of shape information and allows accurate recognition. It is a potent moment technique for shape description [34]. To evaluate ZM, we use the algorithm given in [34], which uses 36 moments for description. We use a Matlab implementation given in [35] that is also a non-optimal framework.

Fig. 12 and Table 5 show the correct classification rate (CCR%) of the multiscale description in 2-D using LPGF, of multiscale description in 2-D using HPGF, of EFD and of ZM, with respect to increasing salt and pepper noise. The results represent mean values obtained over four applications of each algorithm at each noise level. In Fig. 12, the rectangle on graphs represents standard deviation from the mean value and error bar represents minimum and maximum values at data point. It is observed that HPGF-based multiscale description performs better than LPGF-based multiscale description, EFD and ZM. HPGF-based multiscale description achieves 95.5% correct classification rate, whereas LPGF-based multiscale description achieves 90.4%, ZM achieves 90% and EFD achieves 82% without adding noise to the database. As noise increases, the performance of all algorithms decreases and their performances degrade similarly. It is also observed that LPGF-based multiscale description and ZM have very close performances. The success of HPGF-based multiscale description in 2-D appears due to emphasizing the boundary and exterior parts of objects and also allowing the central part contribute slightly to classification.

There are also other techniques that used the same database (MPEG-7 CE-Shape-1 Part B) for classification purpose. A subset of this shape database was used by Kunttu et al. [6]. Their descriptors are wavelet transform-based multiscale contour Fourier descriptors, which is obtained by applying the Fourier transform to the coefficients of the multiscale complex wavelet transform. They applied classification for 30 classes without any noise in the dataset. The selected classes are: bone, bottle, brick, car, cellular phone, children, chopper, comma, deer, device0, device1, device2, device7, device8, face, fish, fountain, frog, glass, heart, key, lmfish, misk, octopus, pencil, personal car, pocket, shoe, teddy and truck. Using the leave-one-out classification with a nearest neighbor classifier, they achieve 94.2–96.3% with respect to the length of descriptors.

Same subset was also recently used by McNeil and Vijayakumar [12] for classification without any noise in the dataset. In their work, shape boundary is represented with a large number of equally spaced points either defined by perimeter distance (perimeter descriptors) or radial angle (radial descriptors). Then, a probabilistic correspondence-based algorithm, which also incorporates with scale, translation and rotation invariance, is applied for shape matching. They note that suitability of the perimeter distance or radial angle for description depends on the classes in

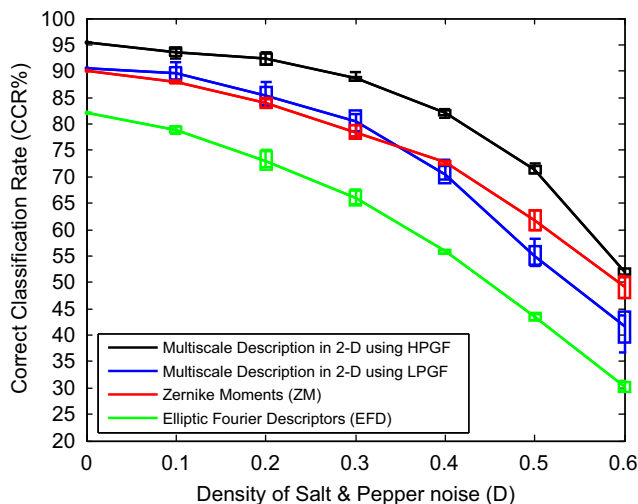


Fig. 12. Classification performance of HPGF- and LPGF-based multiscale description in 2-D, ZM and EFD, with respect to increasing salt and pepper noise in the database. In graphs, the rectangle represents standard deviation from the mean value and error bar represents minimum and maximum values of the CCR%.

Table 5

CCR% of HPGF- and LPGF-based multiscale description in 2-D, ZM and EFD, with respect to increasing salt and pepper noise in the database.

Descriptions	Salt and pepper noise density (D)						
	0	0.1	0.2	0.3	0.4	0.5	0.6
Multiscale description using HPGF	95.5%	93.6%	92.2%	88.5%	82.0%	71.3%	52.0%
Multiscale description using LPGF	90.4%	89.5%	85.2%	80.5%	70.3%	55.0%	41.5%
Zernike moments (ZM)	90.0%	87.9%	83.9%	78.4%	72.6%	61.7%	49.0%
Elliptic Fourier descriptors (EFD)	82.0%	78.9%	73.0%	65.9%	55.8%	43.5%	30.2%

the dataset and these two descriptions can also be combined to improve classification in some datasets. They used the same testing procedure, leave-one-out classification with a nearest neighbor classifier, to compare with the wavelet-based multiscale contour Fourier descriptors described above. They only show the results of the perimeter descriptors, which performs 95.6–98.0% with respect to the number of points selected on the boundary.

Table 6

CCR% of multiscale contour Fourier descriptors [6], perimeter descriptors [12], HPGF- and LPGF-based multiscale description in 2-D, ZM and EFD on the subset (30 classes) without any noise.

Descriptors	CCR %
Multiscale description using HPGF in 2-D	99.2%
Perimeter descriptors [12]	95.6–98.0%
Wavelet-based multiscale contour Fourier descriptors [6]	94.2–96.3%
Multiscale description using LPGF in 2-D	95.8%
Zernike moments (ZM)	92.6%
Elliptic Fourier descriptors (EFD)	87.8%

Table 7

CCR% of perimeter descriptors, radial descriptors, combined perimeter and radial descriptors [12], HPGF- and LPGF-based multiscale description in 2-D, ZM and EFD on the full dataset (70 classes) without any noise.

Descriptors	CCR %
Combined perimeter-radial descriptors [12]	96.2%
Multiscale description using HPGF in 2-D	95.7%
Perimeter descriptors [12]	95.7%
Multiscale description using LPGF in 2-D	91.1%
Radial descriptors [12]	91.0%
Zernike moments (ZM)	90.2%
Elliptic Fourier descriptors (EFD)	82.0%

They also evaluated their descriptors on the full dataset, which includes 70 classes, without any noise in the dataset. They achieved 95.7% and 91.0% with perimeter descriptors and radial descriptors, respectively. They also combined perimeter and radial descriptors and achieved 96.2% classification accuracy on the full dataset.

On the other hand, HPGF-based multiscale description in 2-D achieves 99.2% on the same subset using the leave-one-out classification with a nearest neighbor classifier. This result shows that our algorithm, with HPGF-based multiscale description, outperforms both perimeter descriptors and multiscale contour Fourier descriptors on the subset. We also evaluated other algorithms on the same subset and observe that LPGF-based multiscale description in 2-D achieves 95.8%, ZM achieves 92.6% and EFD achieves 87.8%. Table 6 shows CCR% of the algorithms on the subset without any noise.

On the full dataset, HPGF-based multiscale description in 2-D achieves 95.7%, which is better than radial descriptors, same as perimeter descriptors, and slightly less than the combined perimeter and radial descriptors. LPGF-based multiscale description in 2-D achieves 91.1%, ZM achieves 90.2% and EFD achieves 82%. Table 7 shows CCR% of the algorithms on the full dataset without any noise.

4.2. Comparison on Swedish leaf database

We also experiment the proposed multiscale shape description technique on Swedish leaf database [29]. This is a real and very challenging database for shape classification experiments. It contains isolated leaves from 15 different Swedish tree species, with 75 leaves per species. In this dataset, there is a high between class similarity and large inner class deformations. Fig. 13 (a) shows examples from each leaf species in gray-scale image form.

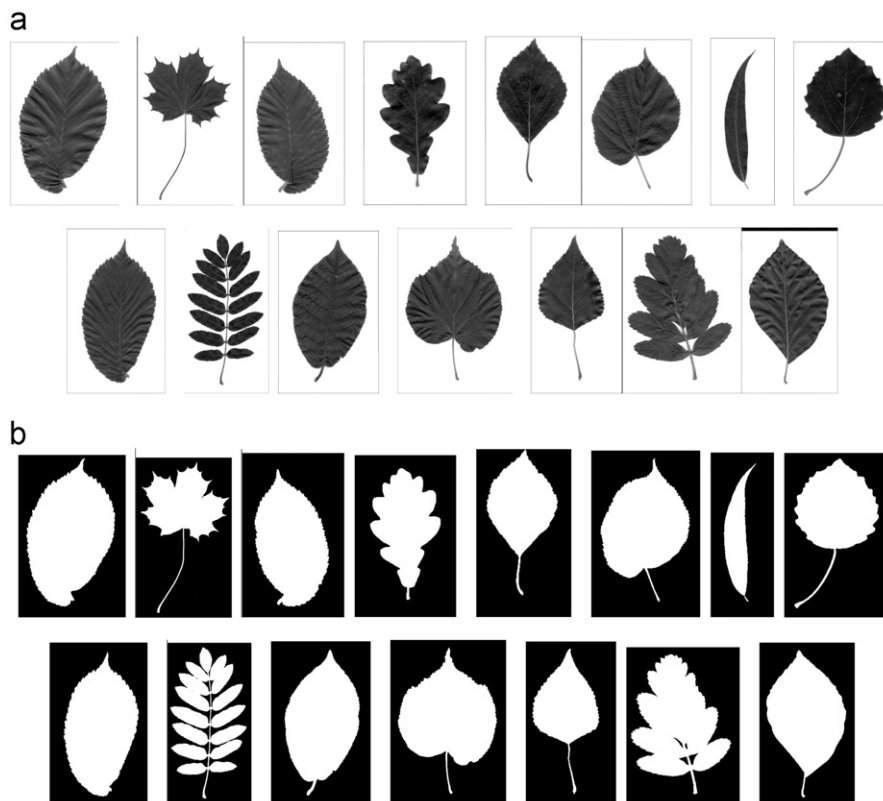


Fig. 13. A sample from each leaf species in the database: (a) in gray-scale image form and (b) in binary image form.

Some species have very similar shapes but different texture, which makes the combination of shape and texture features more suitable for this dataset. However, in our evaluation we use only the shape features which are obtained using the proposed multiscale shape description technique. For this purpose, first the gray-scale images are thresholded to obtain binary segmentation (shape of the leaves), and then the shape features are computed from the binary images. Fig. 13 (b) shows examples from each leaf species in binary image form.

Performance evaluation is employed by comparing the multiscale description using LPGF (average distance) and multiscale description using HPGF (average distance) with each other as well as with EFD and ZM. We use the same testing procedure as we use in MPEG-7 dataset, which is leave-one-out cross-validation with nearest neighbor classifier and the CCR% is measured with Eq. (12). The number of descriptors for EFD is 80 and the number of descriptors for ZM is 36.

In a multiscale description using LPGF, the object size is normalized to be 3500, which is the sum of intensities over the image, in a 211×211 size image. Five different scales are also selected for multiscale representation. The selected scales are: $\sigma_1=20$, $\sigma_2=15$, $\sigma_3=11$, $\sigma_4=8$ and $\sigma_5=5$. The size of the object descriptor matrix is 90×90 at each scale.

In a multiscale description using HPGF, the object size is similarly normalized to be 3500, which is the sum of intensities over the image, in a 211×211 size image. Five different scales are selected for multiscale representation. The selected scales are: $\sigma_1=11$, $\sigma_2=8$, $\sigma_3=5$, $\sigma_4=3$ and $\sigma_5=1$. The size of the object descriptor matrix is 90×90 at each scale.

Table 8 shows the CCR% of the multiscale description in 2-D using LPGF, of multiscale description in 2-D using HPGF, of EFD and of ZM on the Swedish leaf dataset. It is observed that HPGF-based multiscale description performs better than LPGF-based multiscale description, ZM and EFD. HPGF-based multiscale description achieves 90.2% correct classification rate, while LPGF-based multiscale description achieves 84.6%, ZM achieves 85.0% and EFD achieves 80.8%.

4.3. Comparison of computational efficiency

We evaluate the computational efficiency of the HPGF- and LPGF-based multiscale description, on both MPEG-7 dataset and Swedish leaf dataset, by comparing with EFD and ZM techniques. Feature extraction and classification times are measured by using

Table 8
CCR% of HPGF- and LPGF-based multiscale description in 2-D, ZM and EFD on the Swedish leaf database [29] without any noise.

Descriptors	CCR %
Multiscale description using HPGF in 2-D	90.2%
Zernike moments (ZM)	85.0%
Multiscale description using LPGF in 2-D	84.6%
Elliptic Fourier descriptors (EFD)	80.8%

Table 9
The computation time of feature extraction and classification for 1400 shapes in MPEG-7 CE-Shape-1 Part B database.

Descriptors	Total time of feature extraction of 1400 shapes (ms)	Average time of feature extraction of each shape (ms)	Total time of classification of 1400 shapes (ms)	Average time of classification of each shape (ms)
Multiscale description using HPGF in 2-D	683970	488.5	864120	617.2
Multiscale description Using LPGF in 2-D	684953	489.2	871203	622.2
Zernike moments (ZM)	965693	689.7	3369	2.4
Elliptic Fourier descriptors (EFD)	163941	117.1	4602	3.2

MATLAB 7.0 on an Intel Core 2 Quad (processor) computer, which runs Windows Vista operating system with 2.66 GHz CPU and 4GB RAM.

4.3.1. Computational efficiency on MPEG-7 dataset

Table 9 shows the total time of feature extraction of 1400 shapes in the MPEG-7 CE-Shape-1 Part B dataset and average time of feature extraction of each shape. Table 9 also shows the total time for classification of 1400 shapes and the average time of classification of each shape. In feature extraction, it is observed that EFD is more efficient than ZM, HPGF- and LPGF-based multiscale description (five scales for LPGF and five scales for HPGF). Because the shape signature, in EFD, is a 1-D function that represents the shape derived from the boundary points of a 2-D binary image. Geometric invariance, in EFD, is also achieved after the Fourier transform by normalizing Fourier coefficients appropriately. HPGF- and LPGF-based multiscale description are more efficient than ZM in feature extraction. Although we use a multiscale representation (LPGF or HPGF), the proposed algorithm is easier to compute in comparison to Zernike moments. To obtain the Zernike moments, Zernike polynomials are computed, which are difficult and complex. On the other hand, in our algorithm, we rely on a polar transform and two Fourier transforms that are computed by the Fast Fourier Transform. HPGF- and LPGF-based multiscale description have very similar computational performances, since they are the same algorithms with different filtering models.

In classification, ZM is more efficient than EFD, HPGF- and LPGF-based multiscale description. The reason is ZM uses only 36 descriptors to measure similarity between objects, which is lower number of descriptors in comparison to other methods. EDF uses 80 descriptors to measure similarity and it is slightly less efficient than ZM. HPGF- and LPGF-based multiscale description are computationally more complex than ZM and EFD. Since the size of the object descriptor matrix is 90×90 and we have five scales in this evaluation. Due to the higher number of descriptors in HPGF- and LPGF-based multiscale description, the similarity measurement between objects takes more time than ZM and EFD. Although HPGF-based multiscale description has higher number of descriptors and takes more time in classification, it has significantly better classification performance in comparison to ZM and EFD. We also strongly believe that we can reduce the number of descriptors by analyzing the features. We can find efficient and persistent features over the selected scales to increase discrimination and also reduce the number of features.

4.3.2. Computational efficiency on Swedish leaf dataset

Table 10 illustrates the total time of feature extraction of 1125 shapes in the Swedish leaf dataset and average time of feature extraction of each shape. Table 10 also shows the total time for classification of 1125 shapes and the average time of classification of each leaf shape.

In feature extraction, EFD is the most efficient model, since it is computed from a 1-D function that represents the shape

Table 10

The computation time of feature extraction and classification for 1125 shapes in Swedish leaf database.

Descriptors	Total time of feature extraction of 1125 shapes (ms)	Average time of feature extraction of each shape (ms)	Total time of classification of 1125 shapes (ms)	Average time of classification of each shape (ms)
Multiscale description using HPGF in 2-D	538967	479	557469	495.5
Multiscale description using LPGF in 2-D	547594	486.7	559435	497.2
Zernike moments (ZM)	727744	646.8	2230	1.9
Elliptic Fourier descriptors (EFD)	105643	93.9	2667	2.3

boundary points and the geometric invariance is also achieved after the Fourier transform by normalizing Fourier coefficients. HPGF- and LPGF-based multiscale description (five scales) are more efficient than ZM in feature extraction. In ZM, Zernike polynomials are computed, which are difficult and complex. On the other hand, in HPGF- and LPGF-based multiscale description, we rely on a polar transform and two Fourier transforms that are computed by the Fast Fourier Transform. HPGF- and LPGF-based multiscale description have very similar computational performances, since they have the same algorithm but different filtering models.

In classification, ZM is the most efficient model, since it uses only 36 descriptors to measure similarity between objects. ZM has the lowest number of descriptors in comparison to other methods. EDF uses 80 descriptors to measure similarity and it is slightly less efficient than ZM. HPGF- and LPGF-based multiscale description are computationally more complex than ZM and EFD. Since we are using five scales in this evaluation and the size of the object descriptor matrix is 90×90 at each scale. Due to the higher number of descriptors in HPGF- and LPGF-based multiscale description, the similarity measurement between objects takes more time than ZM and EFD. Although HPGF-based multiscale description has higher number of descriptors and computationally less efficient, it has significantly better classification accuracy in comparison to ZM and EFD. We also strongly believe that we can reduce the number of descriptors by analyzing the features. We can find persistent and effective features over the selected scales to increase discrimination and also reduce the number of features.

5. Conclusions and future work

Mutiscale description is a promising approach for shape recognition. Different features can be obtained at different scales and combining these features can increase the discrimination power between objects and therefore increase the correct classification rate. Although many boundary-based multiscale description techniques exist, there is no region-based multiscale description technique in the image space. We have presented a novel image-based multiscale description using a low-pass Gaussian filter (LPGF) and a high-pass Gaussian filter (HPGF), separately. Using the LPGF at different scales represents the inner and central part of an object more than the boundary. On the other hand using the HPGF at different scales represents the boundary and exterior parts of an object more than the central part. In addition, most of the existing multiscale description techniques are based on low-pass filtering (such as LPGF). In our work, we also show that the HPGF-based multiscale description in 2-D space can perform better than the well-known techniques even in noisy conditions.

Our algorithm starts with object size normalization and we then compute a Fourier magnitude image that is translation invariant. At this stage, a LPGF or a HPGF with a selection of scale parameters is used to obtain multiscale Fourier magnitude

images. To give rotation invariance, each image of different scale is polar mapped and then another Fourier magnitude image is computed to obtain the proposed object descriptors. For classification, the Euclidean distance is measured separately at each scale, and then the average distance is computed for each object. Multiscale description using HPGF, which represents the boundary and exterior parts of an object more than the central part, outperforms multiscale description using LPGF, elliptic Fourier descriptors (EFD) and Zernike moments (ZM) with respect to increasing salt and pepper noise in the database. Multiscale description using HPGF in 2-D also performs better than wavelet transform-based multiscale contour Fourier descriptors and performs similar to the perimeter descriptors without any noise in the dataset. Classifying objects with this new multiscale Fourier-based object description using the HPGF in 2-D space increases immunity to noise and discrimination power.

In the future work, we can find effective and persistent features over the selected scales to increase discrimination and also reduce the number of features. A new classifier can be used instead of nearest neighbor classifier to increase correct classification rate. In addition, we can also investigate phase information of the Fourier transforms, which is currently discarded in our algorithm. The phase has significant information about the image and it could be beneficial to include it in object description.

References

- [1] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane close curves, *IEEE Transactions on Computers* C-21 (1972) 269–281.
- [2] G.H. Granlund, Fourier preprocessing for hand print character recognition, *IEEE Transactions on Computers* C-21 (2) (1972) 195–201.
- [3] E. Persoon, K. Fu, Shape discrimination using Fourier descriptors, *IEEE Transactions on Systems, Man, and Cybernetics* 7 (1977) 170–179.
- [4] M.S. Nixon, A. Aguado, *Feature Extraction and Image Processing*, 2nd ed., Elsevier, 2007.
- [5] G.C.H. Chang, C.C.J. Kuo, Wavelet descriptor of planer curves: theory and applications, *IEEE Transactions on Image Processing* 5 (1996) 56–70.
- [6] I. Kunttu, L. Lepisto, J. Rauhamaa, A. Visa, Multiscale Fourier descriptor for shape classification, *IEEE International Conference on Image Analysis and Processing* (2003) 536–541.
- [7] I. Kunttu, L. Lepisto, J. Rauhamaa, A. Visa, Multiscale Fourier descriptor for shape-based image retrieval, *IEEE International Conference on Pattern Recognition* 2 (2004) 765–768.
- [8] I. Kunttu, L. Lepisto, J. Rauhamaa, A. Visa, Multiscale Fourier descriptors for defect image retrieval, *Pattern Recognition Letters* 27 (2) (2006) 123–132.
- [9] F. Mokhtarian, A.K. Mackworth, A theory of multiscale, curvature-based shape representation for planer curves, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (8) (1992) 789–805.
- [10] D.S. Zhang, G. Lu, A comparative study on shape retrieval using Fourier descriptors with different shape signatures, *International Conference on Intelligent Multimedia and Distance Education* (2001) 1–9.
- [11] D.S. Zhang, G. Lu, A comparative study of Fourier descriptors for shape representation and retrieval, *Asian Conference on Computer Vision* (2002) 646–651.
- [12] G. McNeil, S. Vijayakumar, 2D shape classification and retrieval, *International Joint Conference on Artificial Intelligence* (2005) 1483–1488.
- [13] G. McNeil, S. Vijayakumar, Hierarchical procrustes matching for shape retrieval, *IEEE International Conference on Computer Vision and Pattern Recognition* 1 (2006) 885–894.
- [14] H. Ling, D.W. Jacobs, Shape classification using the inner-distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (2) (2007) 286–299.

- [15] C. Xu, J. Liu, X. Tang, 2D shape matching by contour flexibility, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (1) (2009) 180–186.
- [16] A. Temlyakov, B.C. Munsell, J.W. Waggoner, S. Wang, Two perceptually motivated strategies for shape classification, *IEEE International Conference on Computer Vision and Pattern Recognition* (2010) 2289–2296.
- [17] T. Adamek, N.E. O'Connor, A multiscale representation method for nonrigid shapes with a single closed contour, *IEEE Transactions on Circuit and Systems for Video Technology* 14 (5) (2004) 742–753.
- [18] R.S. Torres, A.X. Falcao, L.F. Costa, A graph-based approach for multiscale shape analysis, *Pattern Recognition* 37 (6) (2004) 1163–1174.
- [19] I.E. Rube, N. Alajlan, M. Kamel, M. Ahmed, G. Freeman, Robust multiscale triangle-area representation for 2D shapes, *IEEE International Conference on Image Processing* 1 (2005) 545–548.
- [20] A. Witkin, Scale-space filtering, *International Joint Conference on Artificial Intelligence* (1983) 1019–1021.
- [21] R.J. Prokop, A.P. Reeves, A survey of moment-based techniques for unoccluded object representation and recognition, *CVGIP: Graphical Models and Image Processing* 54 (5) (1992) 438–460.
- [22] D.S. Zhang, G. Lu, Generic Fourier descriptor for shape-based image retrieval, *IEEE International Conference on Multimedia and Expo* 1 (2002) 425–428.
- [23] M.K. Hu, Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory IT-8* (1962) 179–187.
- [24] M.R. Teague, Image analysis via the general theory of moments, *Journal of the Optical Society of America* 70 (8) (1979) 920–930.
- [25] C.-H. Teh, R.T. Chin, On image analysis by the method of moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10 (4) (1988) 496–513.
- [26] C. Direkoğlu, M.S. Nixon, Shape classification using multiscale Fourier-based description in 2-D space, *IEEE International Conference on Signal Processing* 1 (2008) 820–823.
- [27] C. Direkoğlu, M.S. Nixon, Image-based multiscale shape description using Gaussian filter, *IEEE Indian Conference on Computer Vision, Graphics and Image Processing* (2008) 673–678.
- [28] L.J. Latecki, R. Lakamper, U. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, *IEEE International Conference on Computer Vision and Pattern Recognition* (2000) 424–429.
- [29] O. Soderkvist, Computer vision classification of leaves from Swedish trees, Master Thesis, Linköping University, 2001.
- [30] A.V. Oppenheim, J.S. Lim, The importance of phase in signals, *Proceedings of the IEEE* 69 (5) (1981) 529–541.
- [31] J. Wood, Invariant pattern recognition: a review, *Pattern Recognition* 29 (1) (1996) 1–17.
- [32] G. Wolberg, S. Zokai, Robust image registration using log-polar transform, *IEEE International Conference on Image Processing* 1 (2000) 493–496.
- [33] M. Nixon, A. Aquado, *Elliptic Fourier Descriptors Matlab Code, Feature Extraction and Image Processing (Book)*, 2nd ed., 2008, pp. 307–308.
- [34] W.Y. Kim, Y.S. Kim, A region-based shape descriptor using Zernike moments, *Signal Processing: Image Communication* 16 (1) (2000) 95–102.
- [35] K. Chang, Zernike moments Matlab code, LANS Pattern Recognition Toolbox (2005) <<http://www.mathworks.com/matlabcentral/fileexchange/7972>>.

Cem Direkoğlu is a research fellow in Center for Digital Video Processing (School of Electronics Engineering) at Dublin City University since December 2010. Before joining to Dublin City University, he was a research fellow in Graphics, Vision and Visualization group (School of Computer Science and Statistics) at Trinity College Dublin, from October 2009 to September 2010. He has his Ph.D. degree in Electrical and Electronics Engineering from University of Southampton (in June 2009). He also has his M.S. (in February 2005) and B.S. (in February 2003) degrees in Electrical and Electronics Engineering from Eastern Mediterranean University, North Cyprus. He has expertise in image processing and computer vision. Mainly, he has interests in physics- and signal processing-based analogies in image processing and computer vision. He has published peer-reviewed papers in international conferences and journals.

Mark S. Nixon is the Professor in Computer Vision at the University of Southampton UK. His research interests are in image processing and computer vision. His team develops new techniques for static and moving shape extraction, which have found application in automatic face and automatic gait recognition and in medical image analysis. His team were early workers in face recognition, later came to pioneer gait recognition and more recently joined the pioneers of ear biometrics. Amongst research contracts, he was the Principal Investigator with John Carter on the DARPA supported project Automatic Gait Recognition for Human ID at a Distance. He has chaired or had major involvement in many conferences (BMVC, AVBPA, IEEE Face and Gesture, ICPR, ICB, IEEE BTAS) and given many invited talks. His vision textbook, co-written with Alberto Aguado, *Feature Extraction and Image Processing* (Academic Press) reached 2nd Edition in 2008. With Tieniu Tan and Rama Chellappa, their book *Human ID based on Gait*, which is part of the new Springer Series on Biometrics was published in 2005. Dr. Nixon is a member of the IEEE and Fellow IET and FIAPR.