Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Optics and Lasers in Engineering 45 (2007) 695–708
OPTICS and LASERS in ENGINEERING

Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study

Raj Bahadur Yadav, Naveen K. Nishchal, Arun K. Gupta*, Vinod K. Rastogi[1]

Photonics Division, Instrument Research and Development Establishment, Dehradun, India

Received 5 October 2006; received in revised form 31 October 2006; accepted 15 November 2006
Available online 24 January 2007

Abstract

In this paper, we report retrieval and classification of shape-based objects employing three techniques-conventional Fourier descriptors (FD), generic Fourier descriptors (GFD) and wavelet-Fourier descriptors (WFD) techniques. All the three techniques have been applied to a database of seven different types of shapes. The centroid distance based shape signatures have been used for the derivation of descriptors. The Euclidean distance has been calculated as a similarity measure parameter for shape classification. For WFD technique, a Mexican-hat wavelet function was used. Classification results from all the three techniques were compared and it was observed that WFD performs better than FD and GFD technique. To study the effect of the noise on the retrieval and classification of shapes of different objects, additive and multiplicative noise of various variances were applied to the database. Precision and recall were also measured as parameters of performance metric.

Keywords: Shape classification; Fourier descriptors; Wavelet-Fourier descriptors; Generic Fourier descriptor; Euclidean distance; Image retrieval

1. Introduction

The current age, which is often known as information age, more and more images are generated in digital form that leads to a growing interest of finding out images/shapes of interest from large collections/database/scene. In order to find an image from a database, the image has to be described by certain features. Shape is an important visual feature for describing an object within the scene. The primary purpose of shape description was only shape classification, but now it is also used for image retrieval [1–4]. There are several techniques in literature for shape description in image processing and pattern recognition [1–31]. Shape description techniques are divided into two categories—boundary-based and region-based. The region-based techniques consider whole area of the object while boundary based techniques concentrate merely on the boundary lines. The boundary-based methods are more popular than region-based methods, because the shape classifications are mainly attentive towards the contour features [4]. The most common boundary based shape descriptors are chain codes [5], autoregressive models [6], wavelet descriptors (WD) [7–10], curvature scale space (CSS) [11], moment [12,13] and Fourier descriptors (FD) [14–20]. The FD technique is one of the most promising techniques for classification of various kinds of objects. The technique finds application in contour coding [18], invariant 2D shape recognition [1], aircraft identification [19], classification of teeth [20], and scene analysis etc.

Zahn and Roskies [21] developed a method for the analysis and synthesis of plane closed curves using FD and showed that it described the shape on the basis of similarity between curves. Persoon and Fu [22] presented some experimental results for character recognition and machine

*Corresponding author. Tel.: +91 135 2787089; fax: +91 135 2787128.
E-mail address: akgupta@irde.res.in (A.K. Gupta).
[1]Present address: Department of Physics, C.C.S. University, Meerut, India.

0143-8166/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.optlaseng.2006.11.001

part recognition. A number of researchers have reported comparison among autoregressive modeling, chain code, CSS, and moment descriptors and showed that FD technique is the best technique for shape retrieval and classification [23–26]. Although, FD is 40-year-old technique, it is still considered as a valid description tool. The shape description and classification using FD either in 1D or 2D space is simple to compute, robust to noise, and compact and it has many applications in different area, but it does not have a multiresolution property.

Zhang and Lu [27] proposed an improved technique based on the interior content with boundary information, which can be applied to general applications, is known as generic FD (GFD). The technique overcomes the drawbacks of existing shape representation techniques. The proposed shape descriptor is divided by applying 2D Fourier transform on a polar-raster sampled shape image. GFD captures finer features of the shape in both radial and circular directions. The GFD and contour-based FD technique have some similarity for hierarchical representation of the shape with same retrieval efficiency for simple shape. The GFD technique cannot be applied for partial matching, while contour-based FD can be applied for partial and occluded shape matching. GFD uses all the interior content information of the shape.

Wavelet transform (WT) can be seen as signal decomposition onto a set of basis functions, which can be obtained from mother wavelet by dilation and shift [32,33]. It has been drawing increasing attention in pattern recognition community because of its attractive multi-resolution, denoising and feature extraction capabilities. WT has been widely used in area of image processing and pattern classification but only few applications in shape description [28,29]. In case of complicated shapes we require effectively describing the shape at multiple resolutions [1]. Yang et al. [28] reported that WD are not rotation invariant and also the matching scheme is complicated and time consuming than that of FD. Kunttu et al. [30] reported efficient Fourier shape descriptors technique for industrial defect images using wavelets. They reported Fourier-based multiresolution shape descriptors combing Fourier shape description with multiple resolutions. They used WT and boundary smoothing to produce the multi-resolution property to FD [31].

In this paper, we used WT to achieve multiscale representation of the shape considering a database of different object shapes using wavelet-Fourier descriptor (WFD). Seven different types of military targets such as aircrafts, helicopters, missiles, tractors, trucks, cars, and biplanes have been used for this study. WFDs were derived from centroid distance shape signature and was found more suitable for image retrieval. To obtain the similarity between the shapes, minimum Euclidean distance (MED) method has been used [3]. The WFD is obtained by applying the Fourier transform to the coefficients of WT to the boundary of shapes. After approximating shapes of objects by boundaries, a set of features is computed for each object in the form of a set of WFD. We used only 20 WFD, all achieved invariant to scale, translation, and rotation for shape retrieval. The classification results using WFD technique has also been compared with GFD and conventional FD techniques. The results have been compared also with the noisy database. Precision and recall have been computed as a performance metric.

2. Shape signature

Any 1D function that represents 2D areas or boundaries to uniquely describe a shape, is defined as a shape signature $u(t)$. The $u(t)$ usually captures the perceptual feature of the shape. Generally boundaries are represented in complex form. Therefore, to represent the position function we consider two shape signatures viz—complex coordinates and centroid distance for test and comparison of object shapes. These shape signatures are mostly used in FDs implementations and for shape representation of various objects. In this study, we used two types of shape signatures; boundary-based method for WFD and FD, and region-based method for GFD.

2.1. Complex function

The complex function is used for the derivation of FD [3]. A complex coordinate function is simply the complex number generated from the boundary coordinates. Assuming $(x(t), y(t)), t = 0, 1, \ldots, L-1)$ the shape boundary coordinates of an object, the complex coordinates $z(t)$ of the target is written as

$$z(t) = x(t) + \mathrm{j}y(t), \tag{1}$$

For removing the effect of the bias, we used the shifted coordinate function e.g.

$$z(t) = [x(t) - x_c] + \mathrm{j}[y(t) - y_c], \tag{2}$$

where (x_c, y_c) is the centroid of the shape. Here x_c and y_c are averages of the boundary coordinates, and are given as

$$x_c = \frac{1}{L\sum_{t=0}^{L-1} x(t)} \quad \text{and} \quad y_c = \frac{1}{L\sum_{t=0}^{L-1} y(t)}, \tag{3}$$

$z(t)$ represents the shape boundary and is a translation invariant signature. Rotation causes $z(t)$ circular shift, and scaling of shape only introduces linearly change in $z(t)$.

2.2. Centroid distance

The distance of the boundary points from the centroid (x_c, y_c) of shape is defined as the *centroid distance* function [3]. The centroid distance is given as

$$r(t) = \sqrt{([x(t) - x_c]^2 + [y(t) - y_c]^2)}. \tag{4}$$

Due to the subtraction of centroid, which represents the position of the shape from boundary coordinates, the centroid distance is invariant to translation. Rotation causes $r(t)$ circular shift, and scaling of shape only changes $r(t)$ linearly. Recently, Zhang and Lu [3] showed that FDs derived from centroid distance perform better than FD derived from other shape signatures. Therefore, in this study we used centroid distance based shape signatures. One image from each class of the database and its centroid distance has been shown in Fig. 1. Fig. 1(a,c,e,g,i,k,m) shows the original images and Fig. 1(b,d,f,h,j,l,n) shows the obtained corresponding centroid distances, respectively. It shows that each class image has a unique shape signature, which causes more tolerant power for the derivation of feature descriptors.

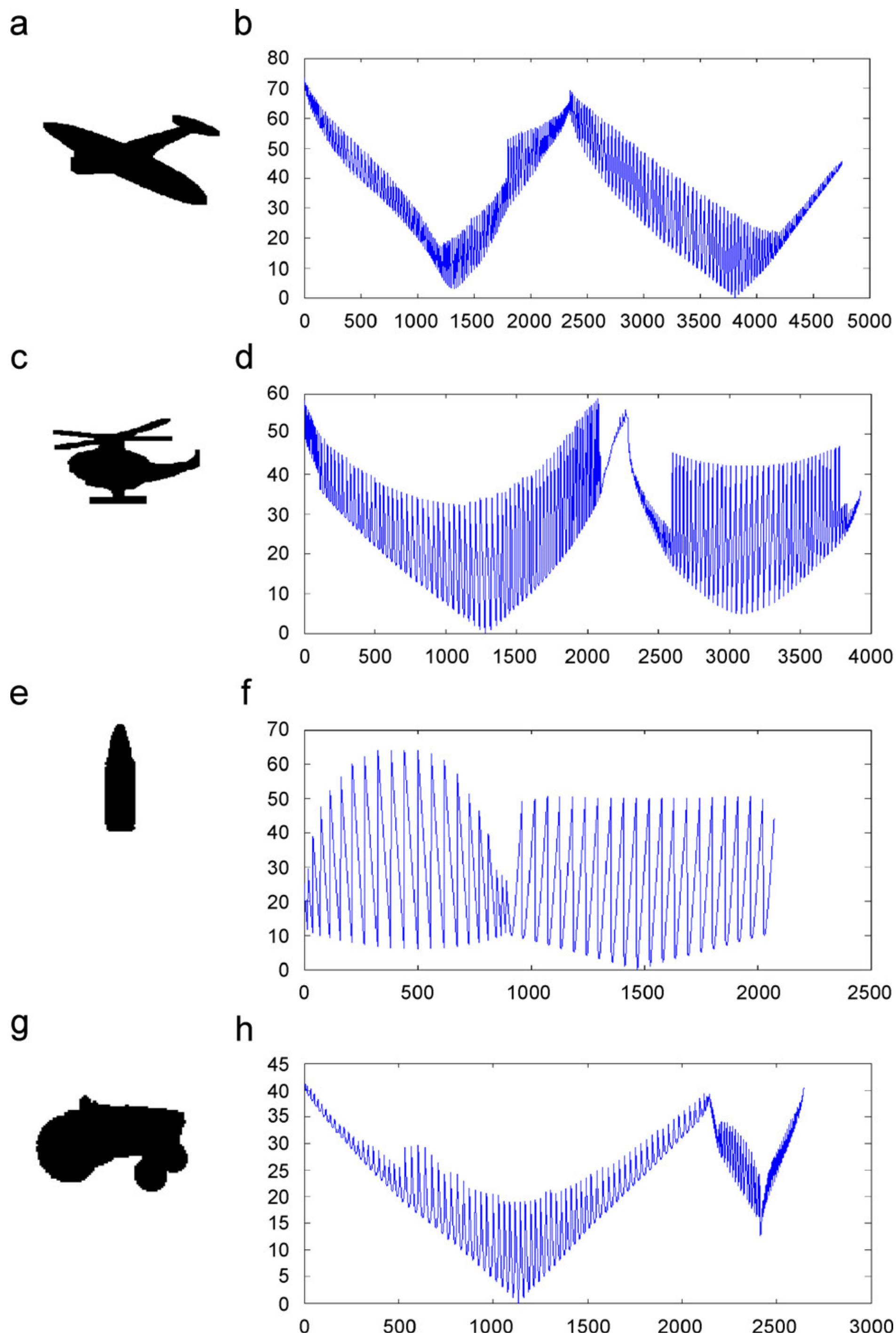

Fig. 1. Images and their centroid distance of each class of database (a) aircraft, (b) its centroid distance, (c) helicopter, (d) its centroid distance, (e) missile, (f) its centroid distance, (g) tractor, (h) its centroid distance, (i) truck, (j) its centroid distance, (k) car, (l) its centroid distance, (m) biplane, and (n) its centroid distance.

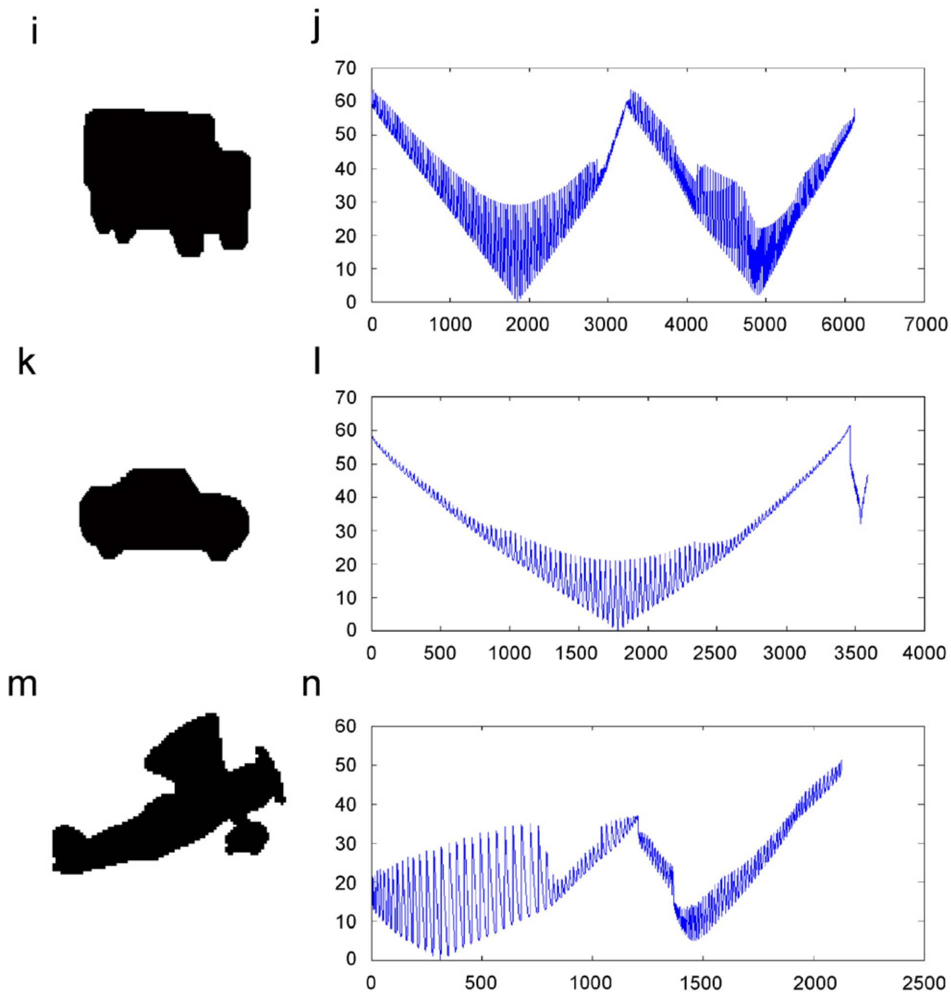

Fig. 1. (*Continued*)

3. Shape descriptors

Shape descriptors are basically feature vectors, which describe the shape of an object [25]. For good retrieval accuracy a shape descriptor should be able to find perceptually similar shapes from a database and also invariant to different geometrical transformation like translation, scale, and rotation etc. for image retrieval and classification. We used centroid distance and complex coordinates as shape signatures, which are translation invariant. Manipulating the amplitude and phase information, the shape descriptors were made invariant to rotation and scaling.

3.1. Contour-based Fourier descriptor

One-dimensional FD also referred to as ordinary FD has been successfully applied to a number of shape representation problems. It has several nice characteristics such as simple derivation, simple normalization, simple to do matching, and robust to noise [27]. Discrete Fourier transform (DFT) and inverse DFT are defined as $a(k)$ and $z(t)$, respectively:

$$a(k) = 1/N \sum_{t=0}^{N-1} r(t) \exp(-\mathrm{j}2\pi kt/N), \tag{5}$$

where $k = -N/2, \ldots, N/2-1$.

$$r(t) = 1/N \sum_{k=0}^{N-1} a(k) \exp(\mathrm{j}2\pi kt/N), \tag{6}$$

where $t = -N/2, \ldots, N/2-1$.

The coefficients $a(k)$ are called FDs, which represent the discrete contour of a shape in the spectral domain. The magnitudes of the coefficients of Fourier transform $a(k)$ are normalized by dividing the first coefficient for achieving the scale invariance and for rotation invariance, the phase information is discarded. In this way, the FDs become translation, rotation, and scale invariant. The performance of FD is better than other contour-based shape descriptors in terms of the effectiveness and efficiency [25]. However, all the contour-based techniques depend upon the boundary information, which may not be available in general situations. Specifically, for complex shapes that consist of several disjoint regions e.g. emblems, trademarks or logos, contour-based shape descriptors are not suitable.

3.2. Region-based Fourier descriptor

In region-based techniques, all the pixels within shape region are taken into account for shape representation, rather than only using boundary information as in case of the contour-based techniques. Most common region-based techniques are Zernike moments, Legendre moments, GFDs, and WDs. These are used for shape representation and classification. The region-based FD is referred to as generic FD [27], which can be used for general applications. We used the GFD for reconstruction and classification of object shapes. The process of deriving GFD is similar to the ordinary FD as has been mentioned in the previous section. Since most of the region-based descriptors are not invariant to translation and scaling therefore, normalization is done before feature extraction. The following steps are used:

(i) the length in x and y direction of the target image is found out,
(ii) the image region (window) which contains the complete object is cropped,
(iii) the cropped image is placed at the center of a rectangular window whose aspect ratio is same as that of the input image with breadth equal to the root mean square value of object's lengths in x and y direction, and
(iv) the image obtained in the step (iii) is resized to the original image.

It is difficult to obtain rotation invariance when images are represented in Cartesian coordinate system. Therefore, two approaches are usually employed to make them rotation-invariant. First method is to represent the image in polar co-ordinate system and then expand the image function in terms of appropriate radial and azimuthal basis functions. The second method is to expand the function in terms of basis functions in x and y direction and then apply the coordinate transformation [1].

$$x = \rho * \cos\theta \quad \text{and} \quad y = \rho * \sin\theta. \tag{7}$$

We used first method, which involves representing an image in polar coordinates. Zhang and Lu [27] modified the polar FT by treating the polar image in polar space as a normal 2D rectangular image in Cartesian space. It was stated that the features captured by the polar FT loses physical meaning in circular direction. Therefore, in this study we used the modified polar FT as derived by Zhang and Lu [27]. The steps followed are

(i) the approximated normalized image is rotated anticlockwise by an angular step sufficiently small,
(ii) the pixel values along positive x-direction starting from the image center are copied and pasted into a new matrix as row elements,
(iii) the steps (i) and (ii) are repeated until the image is rotated by 360°. The result of these steps is an image matrix in polar coordinates (ρ, θ).

The GFD is acquired by applying a discrete 2D Fourier transform on a polar-raster sampled shape image [27].

$$\mathrm{GFD}(\rho,\theta) = \sum_r \sum_i f(r,\theta_i)\exp\left[-\mathrm{j}2\pi\left(\frac{r}{R}\rho + \frac{2\pi\mathrm{i}}{T}\phi\right)\right], \tag{8}$$

where $0 \leqslant r < R$ and $\theta_i = \mathrm{i}(2\pi/T)$ $(0 \leqslant i < T)$; $0 \leqslant \rho < R$, $0 \leqslant \phi < T$ Here R and T are the radial and angular resolutions and $f(x,y)$ is a binary function in shape application.

3.3. Wavelet-Fourier descriptor

WDs are coefficients of WT to the boundary function of an object. WT belongs to the multiresolution transformation, performing the decomposition of the signal object on different levels [7–10,28–33]. Wavelet-based methods are ideally suited for highlighting local features in the decomposed subimages. WDs are formed on the basis of wavelet representation of the original sequence, which describes the boundary of the shape. WT of an image boundary $z(t)$ is defined [31] as

$$C_\mathrm{a}(b) = \frac{1}{\sqrt{|a|}} \int_R z(t)\psi\left(\frac{t-b}{a}\right)\mathrm{d}t. \tag{9}$$

This expression gives the wavelet coefficients of the boundary $z(t)$ at a scale a and position b. For an optimal scale a, a set of wavelet coefficients $C_\mathrm{a}(b)$ is obtained. WFDs are the selected coefficients obtained after applying the Fourier transform to the wavelet coefficients $C_\mathrm{a}(b)$ and is defined as

$$F^\mathrm{a}(k) = \frac{1}{N}\sum_{b=0}^{N-1} C_\mathrm{a}(b)\exp(-\mathrm{j}2\pi b/N). \tag{10}$$

Applying Fourier transform to the wavelet coefficients the benefits of multiscale representation and Fourier shape representation are combined that yields better performance. Due to use of centroid distance shape signature, the WFD are invariant to the translation effect. For the rotation invariance, we discard the phase part and use only the magnitude part of the feature vectors. We normalize the feature vectors to achieve the scale invariance. In this way, we made the WFD as invariant to translation, rotation, and scale.

Generally, WFD of a shape contour represents the object in frequency domain. There are two levels of frequency descriptors—low and high frequency descriptors. The low frequency descriptors have the details of the general features, while high frequency descriptors contain finer details of the object. After applying Fourier transform operation to the wavelet coefficients of an image, we obtain a large number of WFD coefficients. Out of these coefficients, some are significant for image retrieval, which are also known as the representative of the contour. To acquire the knowledge of the number of coefficients necessary for shape retrieval and computation power, we carried out an experiment. We used 10 and 20 descriptors, respectively, to retrieve the object shape. With 10 descriptors we could obtain the global form of the object but finer details were missing, which are essentially required for better classification. We achieved optimal global form and optimal finer details of the object with 20 descriptors. Therefore, we used 20 descriptors for object shape retrieval. The selected 20 normalized descriptors, form feature vectors used for shape indexing.

We used several wavelet functions; such as Harr, Biorthogonal spline, Complex Gaussian, Daubechies, Mexican-hat, Meyer, and Symlet at various scales and observed their comparative performance. The complex Gaussian, Symlet, and Mexican-hat wavelets performed better than others for shape classification. In this study, we used a Mexican-hat wavelet function, which is the second derivative of the Gaussian function and is represented as

$$h(t) = (1 - |t|^2)\exp\left(\frac{t^2}{2}\right). \tag{11}$$

The Fourier transform of the Mexican-hat wavelet function, which is real valued and symmetrical, is given as

$$H(f) = 4\pi f\ \exp(-2\pi f^2). \tag{12}$$

4. Effect of noise

We used additive (Gaussian) and multiplicative (speckle) noise to the database to study the effect on retrieval and classification performance. Since the main parameters for noise study are mean and standard deviation, we studied the effect after applying different values of variance. Considering shape function composed of complex coordinate function $z(t)$ and noise function $N'(t)$ the shape function can be written as

$$z'(t) = z(t) + N'(t), \tag{13}$$

where $N'(t)$ is zero mean, uncorrelated random error with a finite variance σ^2. Following the procedure mentioned in the previous section, we computed and retrieved the shape with the WFD coefficients from noisy data. For the retrieval of noisy shapes also we used only 20 normalized descriptors.

5. Experiment

We created a database of 280 shapes (1–40 aircrafts [Fig. 2(a)], 41–80 helicopters [Fig. 2(b)], 81–120 missiles [Fig. 2(c)], 121–160 tractors [Fig. 2(d)], 161–200 trucks [Fig. 2(e)], 201–240 cars [Fig. 2(f)], and 241–280 biplanes [Fig. 2(g)]) consisting 40 images of seven different types of object shapes. Within the class, there is no difference in the size of the images over the whole database. In this study, we used the shape information of the image for retrieval and classification. After preprocessing, we obtained the contour of the shape. Each class has a standard set of shape descriptors. To compute the similarity between features of two images a metric function was used, which gives the degree of match or similarity for a given pairs of images. We computed ED as a similarity measure, which is defined as

$$\mathrm{ED}\left(F^I, F^Q\right) = \sqrt{\sum_{j=1}^{n} \left\{F_j^Q - F_j^I\right\}^2}, \tag{14}$$

where n is the number of shape descriptors, $F^Q = (f_1^Q, f_2^Q, \ldots, f_n^Q)$ for the query image, and $F^I = (f_1^I, f_2^I, \ldots, f_n^I)$ for a database image.

The class corresponding to the minimum ED value was chosen for the shape classification. The discrimination ability between the shapes of different classes depends upon the ED difference. It means that the ED should be minimum for native class and maximum for other class. If this margin can be increased, we have better discrimination and classification performance. Therefore, a metric called MED between native and other class of images was used, which is defined as

$$\begin{aligned}\mathrm{MED} = {} & \min(\text{ED w.r.t. other classes}) \\ & - \max(\text{ED w.r.t. native class}),\end{aligned} \tag{15}$$

In the experiment, query images were selected randomly from each class of the database. The ED has been calculated for the query object within class and MED has been calculated with respect to the whole database.

5.1. Invariance nature

With the database, as shown in Fig. 2, we created a subdatabase consisting of 42 images obtained after rotating the original object by 90°, 180°, and 270° and scaled upto 0.5, 0.7, and 0.9 of the original image from each class. A number of researchers have shown FD, GFD, and WFD as invariant to translation, rotation, and scaling [8,26]. We also carried out the experiment to check the invariance properties. Chen and Bui [8] reported WFDs performance on rotation and scaling. They demonstrated the effectiveness of the feature extraction algorithm against geometric distortion. With our database having distorted shapes, we applied a relatively simple algorithm to check the invariant nature of WFD technique. In the WFD methodology, Fourier transform is applied to the coefficients of WT of the boundary function. Therefore, the shape information in form of feature vectors is found in frequency domain. Hence, the derivations of shape descriptors become simple and the matching process is straightforward.

Considering feature vectors of second object (180° rotated shape of aircraft) as query object, we compared it with the feature vectors of all objects of the database. Fig. 3(a) shows the computed values with other distorted

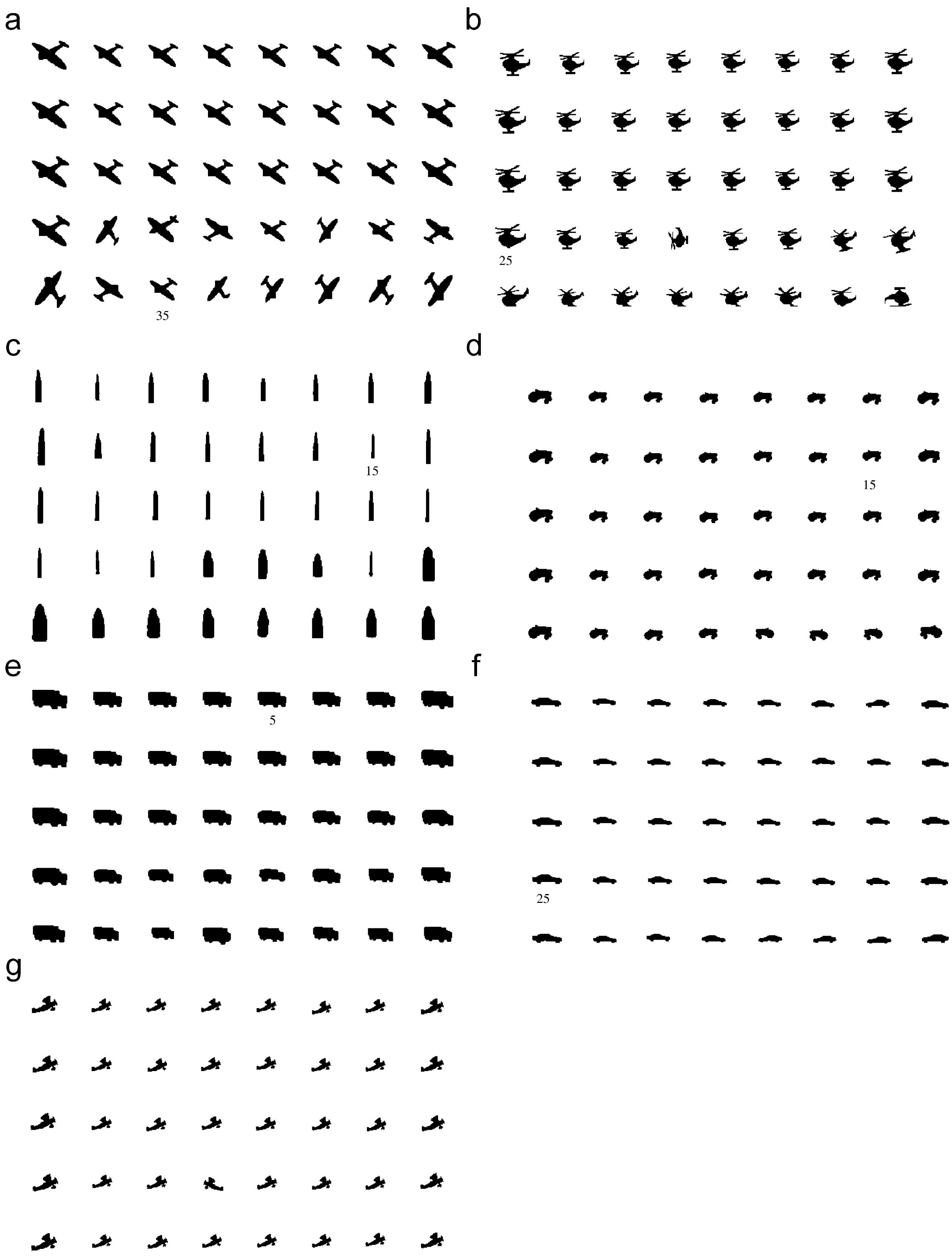

Fig. 2. Image database: (a) aircraft, (b) helicopter, (c) missile, (d) tractor, (e) truck, (f) car, and (g) biplane.

shape of every class with respect to the query shape for WFD technique. With the computed values of ED 0.7, all the shapes belonged to the aircraft class. After a careful observation, we chose this value as the threshold. Each object's ED is corresponding to their classes; hence WFD is invariant to 180° rotations. Similarly we repeated the experiment for other rotations and found the technique rotation-invariant.

Scaled versions of the original image were used to the check the scale invariance property of WFD. We classified the whole database with respect to a scaled query image. As a representative example, results of a single class military

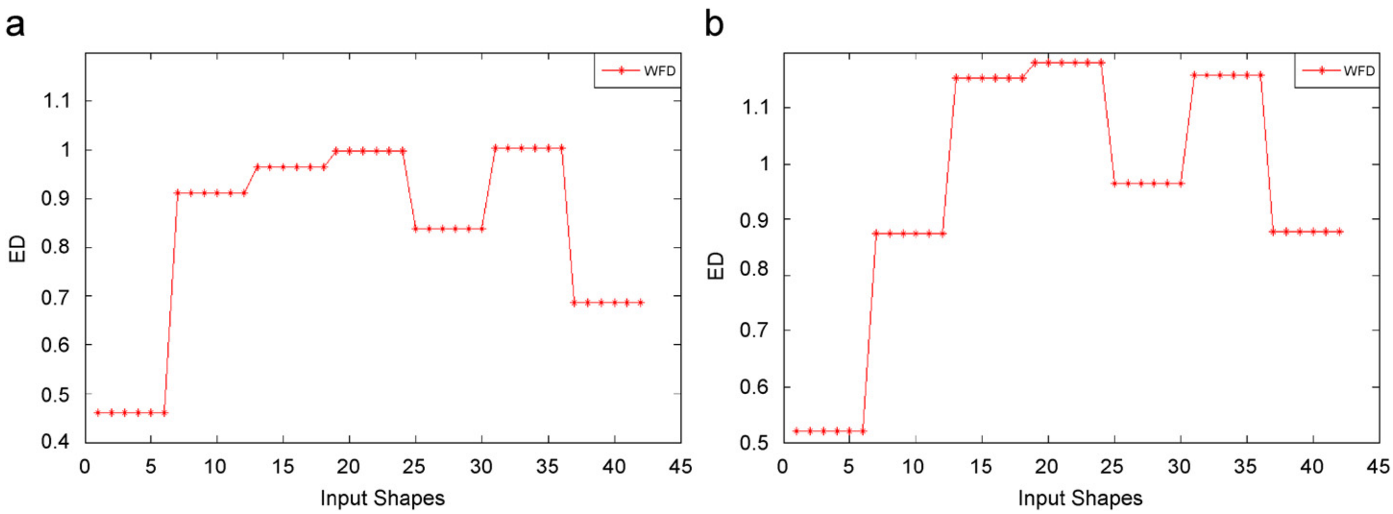

Fig. 3. (a) Invariant WFD: 180° rotated shape of an aircraft image (second shape of database as query), and (b) 50% scaled shape of aircraft image (second shape of database as query).

target have been shown. We used the feature vectors of 10th object (50% scale of original helicopter shape) as query object and compared it with the feature vectors of all objects of the database. Fig. 3(b) shows the computed values with distorted objects of all class with respect to query shape using WFD technique. With the computed value of ED 1.01, all the shapes belonged to the helicopter class. Hence we chose this value as the threshold. Objects with MED values are less than the chosen threshold. No object was found misclassified, hence WFD is said to be invariant to 50% scaling. Similarly we repeated the experiment with 70% and 90% scaled shape of the helicopter image and found WFD scale invariant. Results showed that scaled objects belonged to their native class. Hence it was inferred that the WFD technique is invariant to the geometrical transformation like translation, rotation, and scale. It was also observed that the WFD threshold values were always greater than the FD and GFD threshold values. Therefore, WFD provides additional margin as compared to the FD and GFD, which provided additional accuracy for invariant shape classification.

5.2. Classification of the whole database

The interclass discrimination is the ability of shape descriptors to differentiate shapes belonging to different classes. Shape descriptors of each shape of every class is derived and stored in the database, which are then used for finding out the maximum shape variation within the class. The maximum of these values is used as the marginal threshold for decision-making. The different steps involved in the derivation of WFD and decision making for classification is given below.

i. binarisation of the object,
ii. boundarisation of object and derivation of the contour function,
iii. normalization for translation invariance,
iv. continuous WT of the contour function,
v. Fourier transform to the coefficient of the continuous WT,
vi. selection of the wavelet-Fourier coefficients, and
vii. ED-based comparison for shape classification.

In average shape classification, we first take the average of the feature vectors of the entire image within class, and then compare the query image to each average shape of the class instead of each image of whole database. While in leave-one-out shape classification method, we compare the feature vectors of query shape to each image of the whole database. For unsupervised shape classification we do not take average shape of each class. In this study, we carried out classification using leave-one-out shape method. The obtained results with all the seven different types of object shapes [Fig. 4] employing all the three descriptors FD, GFD, and WFD are discussed below.

(i) The first set of testing database consists 40 images of aircraft shapes [Fig. 2(a)]. We take the feature vectors of first object of this class as query object and compare it with the feature vectors of all objects of the database. Fig. 4(a) shows the computed values of ED with each object of all class with respect to aircraft shape for all techniques. With the computed value of ED = 0.5 for WFD technique, all the shapes belonged to the aircraft class. Hence we chose this value as the threshold. Objects with ED values greater than the threshold do not classify. There are very small variations in ED values within class.

The GFD technique was applied to the same database and EDs were computed. It was found that below the value of 0.1845, all the shapes belong to the aircraft class. Objects with ED values greater than the chosen threshold do not classify. The variation of ED in GFD comparatively is less, which is not suitable for multiscale shape description.

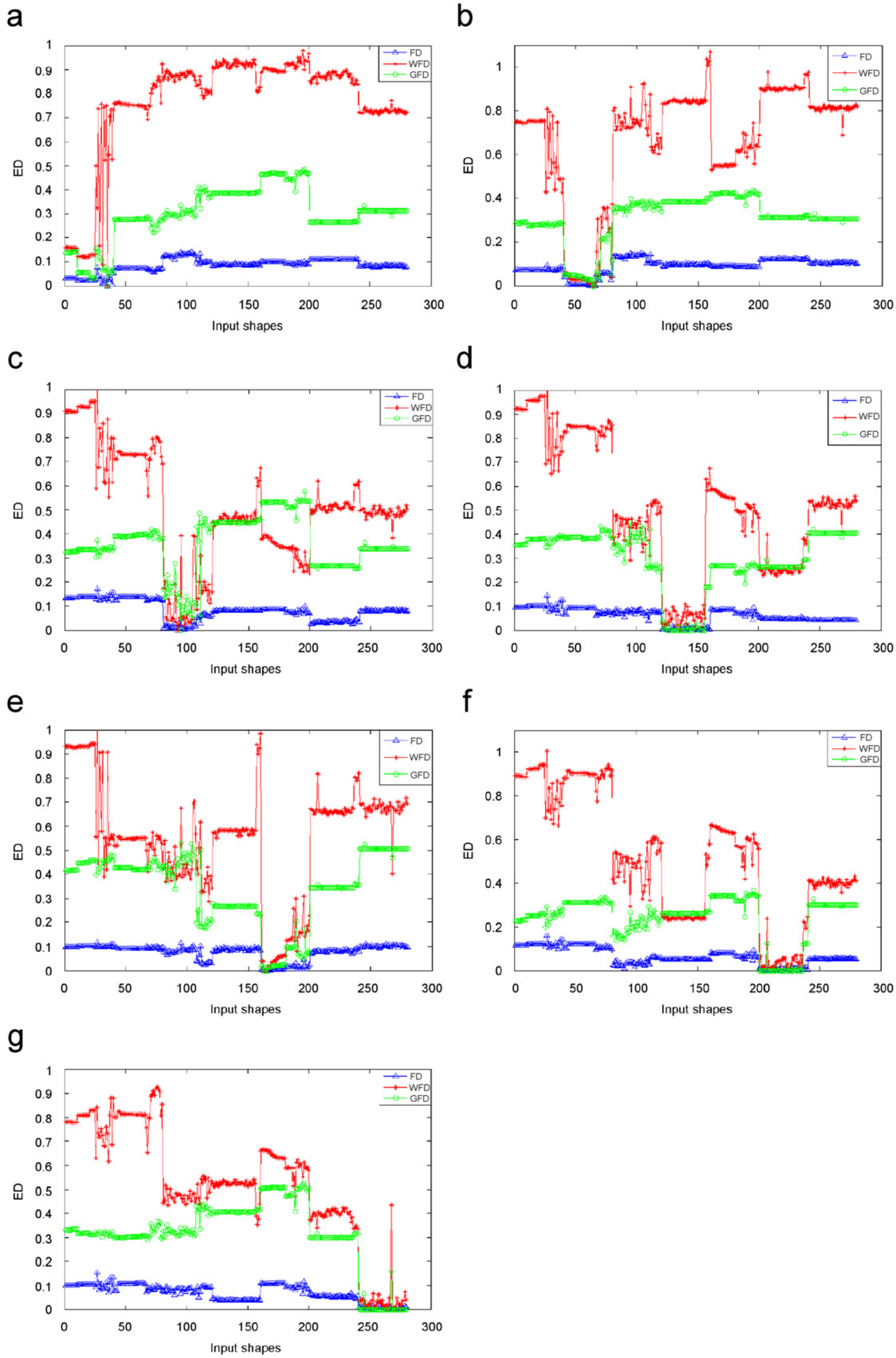

Fig. 4. Classification of images of every class: (a) variation of ED of each object of the database with respect to query object (35th image of the aircraft class), (b) variation of ED of each object of the database with respect to query object (25th image of the helicopter class), (c) variation of ED of each object of the database with respect to query object (15th image of the missile class), (d) variation of ED of each object of the database with respect to query object (15th image of the tractor class), (e) variation of ED of each object of the database with respect to query object (5th image of the truck class), (f) variation of ED of each object of the database with respect to query object (25th image of the car class), and (g) variation of ED of each object of the database with respect to query object (35th image of the biplane class).

The FD technique was also applied to the same database and EDs were computed. It was found that below the value of 0.0303, all the shapes belong to the aircraft class. Objects with ED values greater than the chosen threshold do not classify. The variation of ED in FD is comparatively very less, which is not suitable for multiscale shape description. There are also very small variations in the values of ED within class.

(ii) The second set of testing database consists 40 images of helicopter shapes [Fig. 2(b)]. Randomly selecting the feature vectors of 12th object of this class as query object, we compared it with the feature vectors of all objects from each class. Fig. 4(b) shows the computed values with each object of all class for all the three techniques. We observed that with the computed values of ED 0.4, all the shapes belong to helicopter class. Hence we considered this value as the threshold. Objects with ED values greater than this do not classify.

The GFD technique was also applied to the same database and EDs were computed. It was found that below the value of 0.1746, all the shapes belong to the helicopter class. Objects with ED values greater than the chosen threshold do not classify.

FD was also applied to this database and EDs were computed. It was found that below the value of 0.0524, all the shapes belong to the helicopter class. Objects with ED values greater than the chosen threshold do not classify.

Similar to the above, we selected randomly a query shape from all the seven different types of the military shapes and compared with the feature vectors of all the objects. The values of ED were computed employing all the three FD, GFD, and WFD techniques [Figs. 4(c–g)]. Note that with computed value of the ED, all shapes belonged to one class. Hence that particular value was taken as the threshold. The threshold value was found different for the different class of shapes. Also the values were different for different techniques. Objects with ED values greater than the threshold did not classify. Values of ED for all the shapes and FD, GFD, and WFD techniques were tabulated [Table 1]. In all the above seven class of target shapes it has been observed that there are very small variations in ED within class in all techniques. We found two objects, one from helicopter class [Fig. 2(b)] and another one from fighter class [Fig. 2(g)] not properly classified. Both the objects are having different features. It is also inferred that ED values are more in case of GFD than simple FD by at least a factor of 10. In case of WFD the ED values are very high as compared to GFD, which help classify the shapes better.

From experiments it was observed that the feature enhancement step increases the values of ED of the query shape belonging to all classes. The WFD method has better performance for a range of wavelet scale values, where it enhances the shape features. Therefore, the method can be employed for classification performance by choosing an appropriate scale of optimum performance. In the database used in this study, it was observed that there are very small variations in ED values within class and large for inter class for shape classification using all techniques.

5.3. Effect of noise

To study the effect of noise on shape retrieval and classification, additive and multiplicative noise of various magnitudes were applied to the whole database. We used *imnoise* command of MATLAB to add Gaussian white noise of mean 'M' and variance 'V' to all the images. Fig. 5(a) shows images with additive Gaussian noise of $M = 0$ and $V = 0.9$ applied to one shape of the tank database. After removing the salt and pepper noise from the Fig. 5(a) and then binarised the shape of image. Thereafter we obtained the boundary of the shape using 8-connectivity algorithm and we applied the wavelet-Fourier transform to the boundary function of the noisy shape and we selected 20 wavelet-Fourier coefficients and normalized them. With the help of these coefficients we retrieved the shape of object as shown in Fig. 5(b). The experiment was repeated using constant value of mean 0 in each case and varied the values of variance from 0.1 to 5.0. After iterating the experiment with different values of variance it was observed that the retrieval of shape was possible up to $V = 0.9$. For higher values of variance e.g. $V = 1.5$ the noisy image and retrieved shape have been shown in the Figs. 5(c,d) respectively. We found that for higher values of variance most of the relevant information about the shape were lost e.g. boundary information and hence shape were not classified. Also with any other value of mean the zero, shapes were neither retrieved nor classified.

Similarly, we applied other types of noise i.e. speckle on shape of different categories in database. Fig. 5(e) shows the image of tank after applying speckle noise of $V = 0.7$. Then we removed the salt and pepper noise from the Fig. 5(e) and get the binarised shape after selecting a threshold value. Thereafter we used 8-connectivity

Table 1
ED values used as threshold for the whole database using WFD, GFD and FD techniques

Technique	Aircraft	Helicopter	Missile	Tractor	Trucks	Cars	Biplane
FD	0.0303	0.0524	0.00323	0.0243	0.02130	0.0114	0. 0479
GFD	0.1845	0.1746	0.1158	0.1952	0.0694	0.1401	0.2863
WFD	0.5	0.4	0.3	0.28	0.27	0.24	0.35

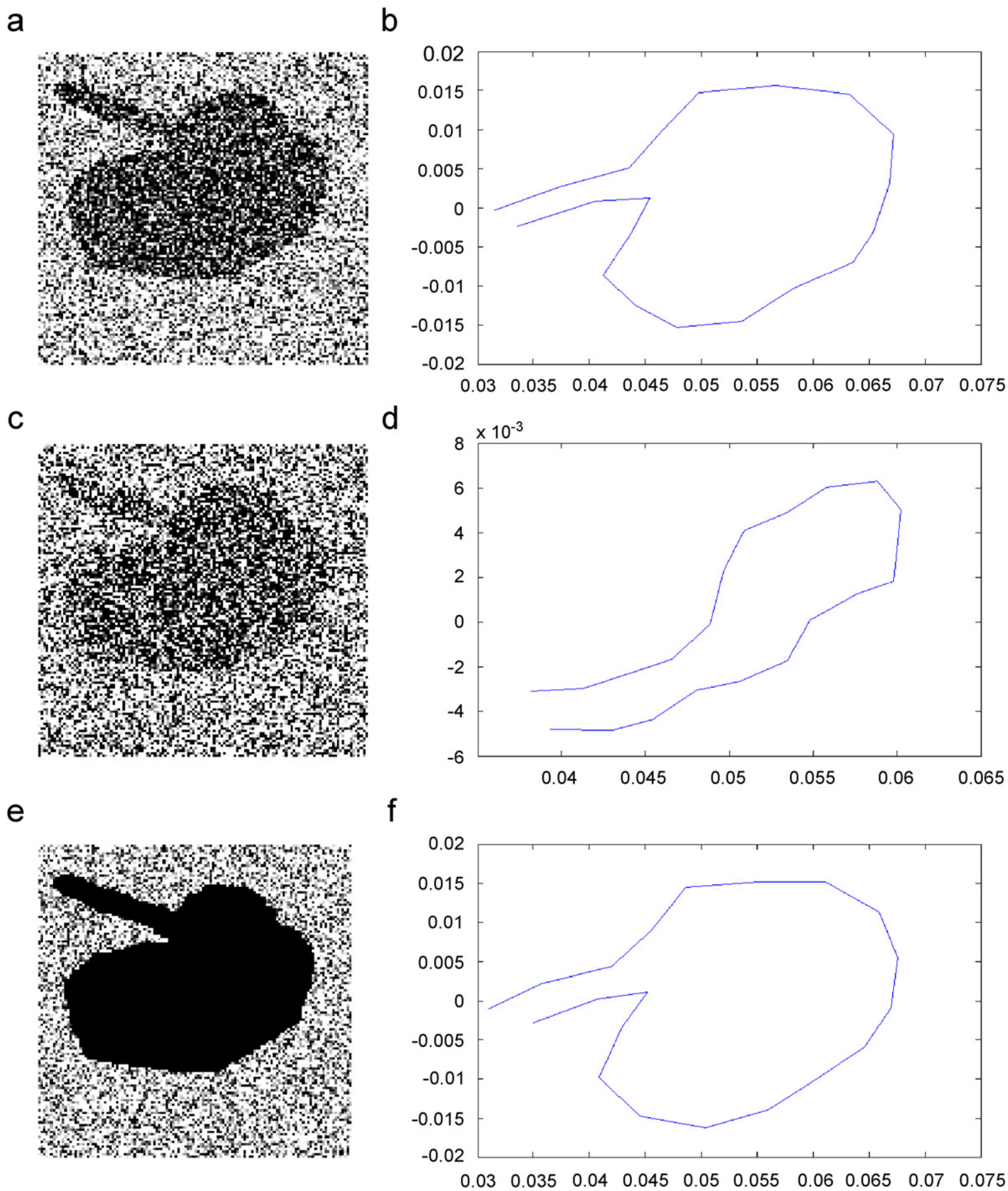

Fig. 5. Retrieval of noise affected tank image: (a) Gaussian noise with $M = 0$ and $V = 0.9$, (b) retrieved image of (a), (c) Gaussian noise with $M = 0$ and $V = 1.5$, (d) retrieved image of (c), (e) speckle noise with $V = 0.7$, and (f) retrieved image of (e).

algorithm [1] to obtain the boundary of the shape and then we applied the Fourier transform to the wavelet coefficients of the boundary function of the noisy shape and selected the normalized 20 WFD for shape retrieval as shown in Fig. 5(f) similar to Gaussian noise case. For the values of more than 0.7 the shapes neither retrieved nor classified.

We also studied the effect of noise (Gaussian and speckle) on the classification of the whole database using WFD, GFD and FD techniques. The experiment was repeated using constant value of mean 0 and different values of variance. After iterating the experiment with different values of variance it was observed that the noisy target images were classified up to $V = 0.9$ with Gaussian noise and 0.7 with speckle noise. With higher values of variance, the shapes neither retrieved nor classified. The experiment was carried out for each class of the database. Results for Gaussian noise ($V = 0.9$) with respect to missile shape query have been shown in Fig. 6(a). From this plot, it is observed that there are very small variations in ED in the intraclass and large variations with interclass. Similarly,

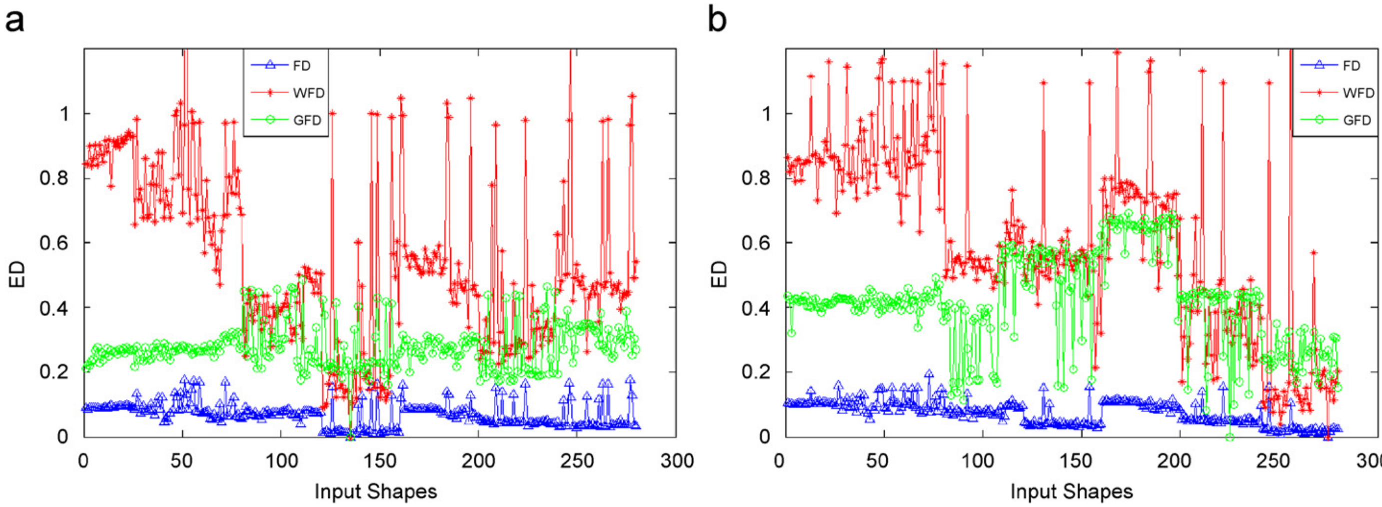

Fig. 6. Classification of noise affected image (a) variation of ED of each object of whole database with respect to query object (randomly selected missile image), and (b) variation of ED of each object of whole database with respect to query object (randomly selected tank image).

results for the classification of speckle noise affected database with respect to missile shape query have been shown in Fig. 6(b). With $V = 0.7$, WFD technique produced small variations in intra-class and large variations in inter-class than GFD and FD. Using higher variance values than 0.7, the shape of object was not retrieved and hence shapes were not classified. Hence, WFD also performed better than FD for noisy database. It is inferred that WFD have more tolerance power than GFD and FD.

6. Classification efficiency

The classification performance ($\eta\%$) is a metric, which we used to evaluate the classification experiment conducted for shape-based objects. It is defined as

$$\eta\% = \frac{m-n}{m} \times 100, \quad (16)$$

where m is the total number of classified image and n the total number of misclassified images. We carried out the classification performance using FD, GFD, and WFD techniques in three ways; original, Gaussian noise affected, speckle noise affected of whole database and has been shown in Fig. 7(a). It is inferred that the WFD technique performs better than FD and GFD for whole database classification. Also, similar performances have been achieved using WFD technique when additive and multiplicative noises are added to the database in Figs. 7(b) and (c) respectively. The classification performances are above 90% in all cases as aircrafts, helicopters, missiles, tanks, military trucks, cars and fighters. In case of fighter class, slight degradation of performance is attributed to inside variation in orientation. The performance of noisy database is slightly degraded than the noise free database classification.

6.1. Retrieval performance

For the measurement of retrieval performance, precision and recall [26] have been calculated as the evaluation measure of the query results. Precision P is defined as the ratio of the number of retrieved relevant shapes 'r' to the total number of retrieved shapes 'n', i.e., $P = r/n$. Recall R is defined as ratio of the number of retrieved relevant shapes 'r' to the total number 'm' of relevant shapes in whole database i.e., $R = r/m$. The physical significance of precision is the measurement of accuracy and recall is the measurement of robustness of retrieval. The performances in terms of precision and recall for original database have been shown in Fig. 7(d) using WFD, GFD and FD techniques. With obtained results we infer that WFD performs better than GFD and FD techniques.

7. Conclusion

The FD has been an effective shape description tool for classifying different kinds of shapes for more than four decades. Exploiting the multiresolution property of wavelets, we applied WFD technique to the created database of 280 shape-based objects. In WFD technique, we utilize the benefits of Fourier transform and WT. One object from each class from the database consisting seven classes, was chosen and it was matched with all objects from the whole database. We observed that when shape similarity is above 90% then class of the object is well recognized. The ED was used as a metric for similarity measure. The results show that there is little variation in ED within class and large variation for other classes. In WFD technique, the computation time becomes less than the GFD technique, because it concentrates only on the boundary of the target. Therefore, it is computationally more efficient for large database. The WFD method has also much tolerance

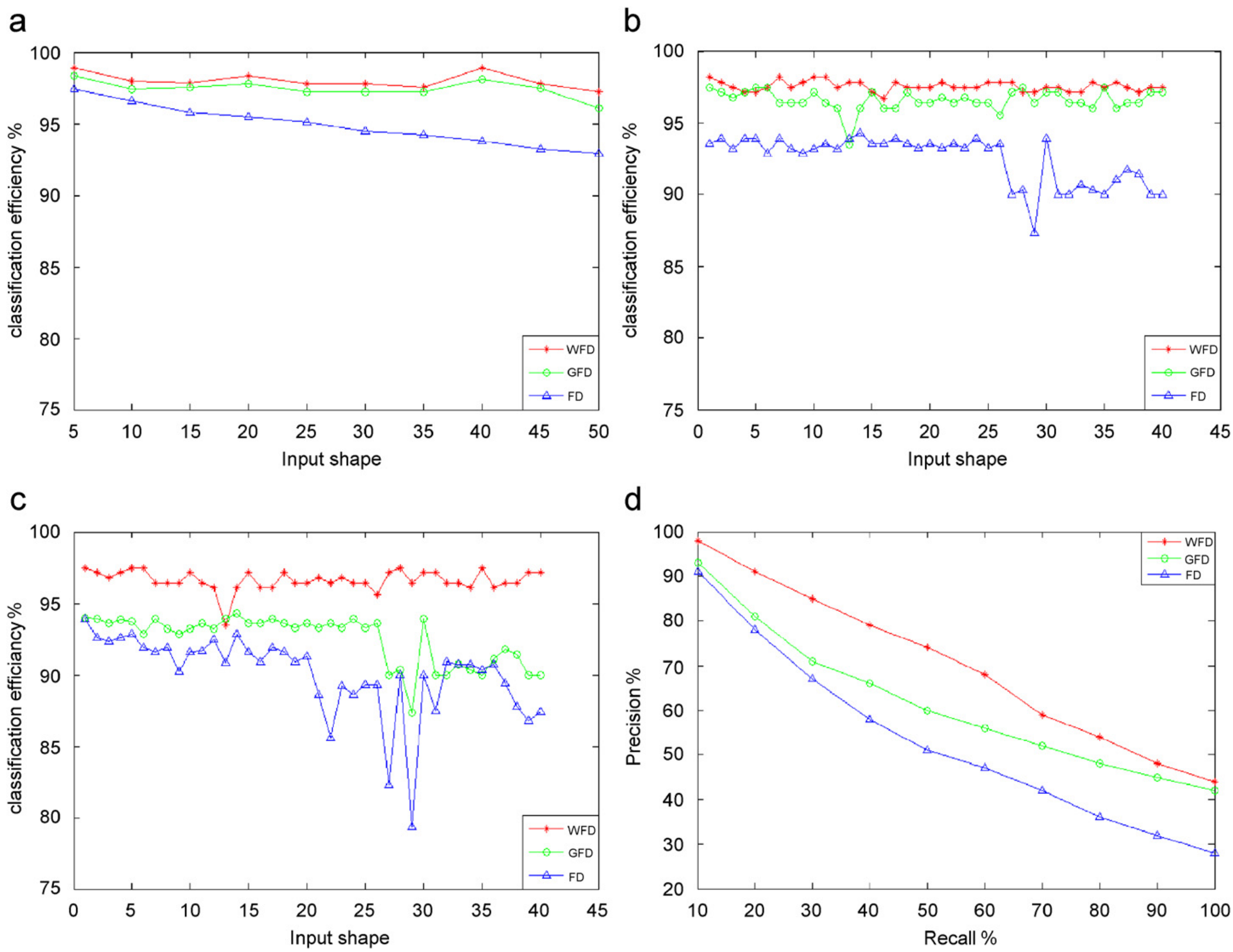

Fig. 7. Performance of the WFD, GFD and FD techniques: (a) classification performance of database, (b) classification performance of Gaussian noise affected database, (c) classification performance of speckle noise affected database, and (d) retrieval performance of database.

power than FD and GFD. The performance of the WFD technique is above than 95% in every class, which is better than FD and GFD. The influence is applicable to noisy database. With WFD technique, additional accuracy could be achieved without increasing the computational cost. The technique could also be applied in other areas such as document analysis, computer vision etc.

Acknowledgment

The authors are also grateful to Shri J.A.R. Krishna Moorty, Director, IRDE for his encouragement and permission to publish this work.

References

[1] Costa LF, Cesar RM. Shape analysis and classification: theory and practice. Boca Raton, Florida: CRC Press; 2001.
[2] Duda RO, Hart PE, Stork DG. Pattern classification, second ed. New York: Willey; 2000.
[3] Zhang D, Lu G. Study and evaluation of different Fourier methods for image retrieval. Image Vision Comput 2005;23:33–49.
[4] Loncaric S. A survey of shape analysis techniques. Pattern Recogn 1998;31:983–1001.
[5] Liu H-C, Srinath MD. Corner detection from chain code. Pattern Recogn 1990;23:51–68.
[6] Dubois SR, Glanz FH. An autoregressive model approach to two-dimensional shape classification. IEEE Trans Pattern Anal Machine Intell 1986;8:55–65.
[7] Chang GC-H, Kuo C-CJ. Wavelet descriptor of planer curves: theory and applications. IEEE Trans Image Process 1996;5:56–70.
[8] Chen G, Bui TD. Invariant Fourier-wavelet descriptors for pattern recognition. Pattern Recogn 1999;32:1083–8.
[9] Osowski S, Nghia DD. Fourier and wavelet descriptors for shape recognition using neural networks—a comparison study. Pattern Recogn 2002;35:1949–57.
[10] Shen D, Ip HHS. Discriminative wavelet shapes descriptors for recognition of 2-D patterns. Pattern Recogn 1999;32:151–65.
[11] Mokhtarian F, Mackworth AK. A theory of multiscale, curvature based shape representation for planer curves. IEEE Trans Pattern Anal Machine Intell 1992;14:789–805.
[12] Dudani SA, Breeding KJ, McGhee RB. Aircraft identification by moments invariants. IEEE Trans Comput 1977;26:39–46.

[13] Liao SX, Pawlak M. On image analysis by moments. IEEE Trans Pattern Anal Machine Intell 1996;18:254–66.
[14] de Leon RD, Sucar LE. Human silhouette recognition with Fourier descriptors. International Conference on Pattern Recognition 2000;3:709–12.
[15] Wimmer A, Ruppert GS, Sidla O, Konard H, Gretzmacher F. FFT-descriptors for shape recognition of military vehicles. Proc SPIE 2000;4029:81–7.
[16] Lam KP. Contour map registration using Fourier descriptors of gradient codes. IEEE Trans Pattern Anal Machine Intell 1985;7:332–8.
[17] Wallace TP, Mitchell OR. Analysis of three-dimensional movement using Fourier descriptors. IEEE Trans Pattern Anal Machine Intell 1980;2:583–8.
[18] Granlund GH. Fourier preprocessing for hand printed character recognition. IEEE Trans Comput 1972;21:195–201.
[19] Lin CC, Chellappa R. Classification of partial 2-D shapes using Fourier descriptors. IEEE Trans Pattern Anal Machine Intell 1987;9:686–90.
[20] Mahoor MH, Abdel-Mottaleb M. Classification of numbering of teeth in dental bitewing images. Pattern Recogn 2005;38:577–86.
[21] Zahn T, Roskies RZ. Fourier descriptors for plane closed curves. IEEE Trans Comput 1972;21:269–81.
[22] Persoon E, Fu K. Shape discrimination using Fourier descriptors. IEEE Trans Syst Man Cybernet 1977;7:170–9.
[23] Kauppinen H, Seppänen T, Pietikäinen M. An experimental comparison of autoregressive and Fourier-based descriptors in 2-D shape classification. IEEE Trans Pattern Anal Machine Intell 1995;17:201–7.
[24] Mehtre BM, Kankanhalli MS, Lee WF. Shape measures for content-based image retrieval: a comparison. Inform Process Manage 1997;33:319–37.
[25] Zhang D, Lu G. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval. J Visual Commun Image Representation 2003;14:41–60.
[26] Golden JP. Terrain contour matching (TECOM): a cruise missile guidance aid. Proc SPIE 1980;238:10–8.
[27] Zhang D, Lu G. Shape-based image retrieval using generic Fourier descriptor. Signal Process Image Commun 2002;17:825–48.
[28] Yang HS, Lee SU, Lee KM. Recognition of 2-D object contours using starting-point independent wavelet coefficient matching. J Visual Commun Image Representation 1998;9:171–81.
[29] Tsai D-M, Chiang C-H. Rotation-invariant pattern matching using wavelet decomposition. Pattern Recognition Lett 2002;23: 191–201.
[30] Kunttu I, Lepisto L, Visa A. Efficient Fourier shape descriptor for industrial defect images using wavelets. Opt Eng 2005;44:080503.
[31] Kunttu I, Lepisto L, Rauhamaa J, Visa A. Multiscale Fourier descriptors for defect image retrieval. Pattern Recognition Lett 2006;27:123–32.
[32] Mallat S. A wavelet tour of signal processing. New York: Academic Press; 1999.
[33] Rao RM, Bopardikar AS. Wavelet transforms: introduction to theory and applications. Reading, MA: Addison-Wesley; 1998.