ELSEVIER

Signal Processing: *Image Communication* 17 (2002) 825–848

SIGNAL PROCESSING: IMAGE COMMUNICATION

www.elsevier.com/locate/image

Shape-based image retrieval using generic Fourier descriptor

Dengsheng Zhang*, Guojun Lu

Gippsland School of Computing and Information Technology, Monash University, Churchill, Vic. 3842, Australia

Received 4 September 2001; accepted 22 July 2002

Abstract

Shape description is one of the key parts of image content description for image retrieval. Most of the existing shape descriptors are usually either application dependent or non-robust, making them undesirable for generic shape description. In this paper, a generic Fourier descriptor (GFD) is proposed to overcome the drawbacks of existing shape representation techniques. The proposed shape descriptor is derived by applying two-dimensional Fourier transform on a polar-raster sampled shape image. The acquired shape descriptor is application independent and robust. Experimental results show that the proposed GFD outperforms common contour-based and region-based shape descriptors.

Keywords: Fourier descriptor; Shape; CBIR; Retrieval

1. Introduction

Due to the rapid increase of multimedia information, there is an urgent need of multimedia content description so that automatic searching is possible. The newly emerging multimedia application MPEG-7 is to address this issue. In MPEG-7, shape is one of the key components for describing digital image along with other features such as texture and color. Six criteria have been set for shape description by MPEG-7, they are: good retrieval accuracy, compact features, general application, low computation complexity, robust retrieval performance and hierarchical representation [12].

Various shape descriptors exist in the literature, these descriptors are broadly categorized into two groups: contour-based shape descriptors and region-based shape descriptors.

Contour-based shape descriptors include Fourier descriptor (FD) [7,21,31,11,32,33], wavelet descriptors [29,30], curvature scale space [17,16,5] and shape signatures [6,4]. Since contour-based shape descriptors exploit only boundary information, they cannot capture shape interior content. Besides, these methods cannot deal with disjoint shapes where boundary information is not available. Consequently, they have limited applications.

In region-based techniques, shape descriptors are derived using all the pixel information within a shape region. Region-based shape descriptors can be applied to general applications. Common region-based methods use moment descriptors to describe shape [8,26,28,13,27,18]. These include geometric moments, Legendre moments, Zernike

*Corresponding author. Fax: +61-3-99026842.

E-mail addresses: dengsheng.zhang@infotech.monash.edu.au (D. Zhang), guojun.lu@infotech.monash.edu.au (G. Lu).

0923-5965/02/$ - see front matter
PII: S0923-5965(02)00084-X

moments and pseudo-Zernike moments. It has been shown in [28] that Zernike moments outperforms other moment methods in terms of overall performance. Recently, several researchers also use the grid method to describe shape [14,23,3]. The grid-based method attracts interest for its simplicity in representation and conformance to intuition; however, its rotation normalization does consider shape interior content. Other region-based shape descriptors are also proposed, these include bounding box descriptor [20] and multi-layer eigenvector shape descriptor [12]. Most of the region-based shape descriptors are extracted from spatial domain, as the result, they are sensitive to noise and shape variations.

In this paper, we propose a generic Fourier descriptor (GFD) which can be applied to general applications. The GFD is extracted from spectral domain by applying two-dimensional (2-D) Fourier transform on polar-raster sampled shape image. It captures shape finer features in both radial and circular directions. The rest of the paper is organized as follows. In Section 2, background of related work and GFD are described in details. In Section 3, we give detailed experimental results on the proposed methods and compare GFD with other shape descriptors. Section 4 concludes the paper.

2. Generic Fourier descriptor

In this section, we describe GFD in details. First, we give some background information on related work in Section 2.1. We then introduce two polar Fourier transforms used to derive GFD in Section 2.2. The implementation details of GFD are described in Sections 2.3 and 2.4.

2.1. Related work

2.1.1. One-dimensional Fourier descriptor

One-dimensional Fourier descriptor (FD) has been successfully applied to many shape representation applications, especially to character recognition. The nice characteristics of FD, such as simple derivation, simple normalization, simple to do matching, robust to noise, perceptually meaningful, compact and hierarchical coarse to fine representation, make it a popular shape descriptor [7,21,31,1,19,22,11,15,9,33]. Generally, one-dimensional (1-D) FD is obtained through Fourier transform (FT) on a shape signature function derived from shape boundary coordinates $\{(x(t), y(t)),\ t = 0, 1, \ldots, N-1\}$. A typical shape signature function is the centroid distance function which is given by the distance of the boundary points from the centroid (x_c, y_c) of the shape

$$r(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2}, \quad t = 0, 1, ..., N-1,$$

where

$$x_c = \frac{1}{N}\sum_{t=0}^{N-1} x(t), \quad y_c = \frac{1}{N}\sum_{t=0}^{N-1} y(t).$$

An example of centroid distance function of an apple shape is shown in Fig. 1.

One dimensional FT is then applied on $r(t)$ to obtain the Fourier transformed coefficients

$$a_n = \frac{1}{N}\sum_{t=0}^{N-1} r(t)\exp\left(\frac{-\mathrm{j}2\pi nt}{N}\right), \quad n = 0, 1, \ldots, N-1.$$

The magnitudes of the coefficients a_n $(n = 0, 1, \ldots, N-1)$ normalized by the magnitude of the first coefficient a_0 are used as shape descriptors, called Fourier descriptors. The acquired FDs are translation, rotation and scale invariant. It has been shown that shape representation using FD outperforms many other contour-shape descriptors [11,33]. However, all these methods assume the knowledge of shape boundary information which may not be available in general situations. For example, it is difficult to derive 1-D FD for the shape in Fig. 2(a), because the contour of the shape is not available. Furthermore, 1-D FD cannot capture shape interior content which is important for shape discrimination. For example, FD is not able to discriminate the shape in Fig. 2(b) from the shape in Fig. 2(c). The drawbacks limit the application of 1-D FD.

2.1.2. Zernike-moments descriptor

The application of Zernike moments on shape overcomes the aforementioned drawbacks of 1-D FD. Zernike-moment descriptor (ZMD) is

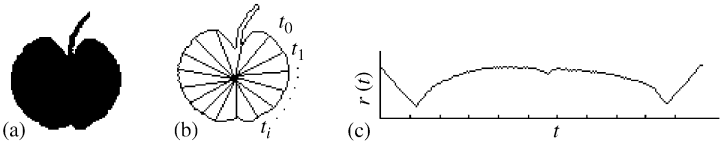


Fig. 1. (a) An apple shape; (b) the contour of (a); (c) centroid distance function of (a).

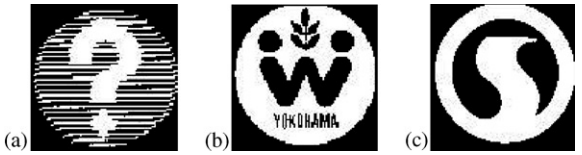


Fig. 2. (a) A shape without contour; (b), (c) two shapes with same contour but with different interior content.

obtained by using all the pixel information within a shape region. It does not assume shape boundary information. ZMD is one of the best shape descriptors among the existing shape descriptors. Many researchers report promising result of ZMD [28,12,33]. It has been tested on MPEG-7 shape databases and adopted in MPEG-7 as region-based shape descriptor. An examination into ZMD reveals that it is essentially a transform-based descriptor which is derived from two-dimensional transform of shape on polar space. The theory of ZMD is similar to FD. However, it is not a fully spectral transform. In the following, we examine ZMD in details.

The complex Zernike moments are derived from Zernike polynomials,

$$V_{nm}(x,y) = V_{nm}(r\cos\theta, r\sin\theta) = R_{nm}(r)\exp(\mathrm{j}m\theta) \tag{2.1}$$

and

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} r^{n-2s}, \tag{2.2}$$

where r is the radius from (x,y) to the shape centroid, θ is the angle between r and x-axis, n and m are integers and subject to $n-|m| = \text{even}$, $|m| \leqslant n$. Zernike polynomials are a complete set of complex-valued function orthogonal over the unit disk, i.e., $x^2+y^2=1$. Then the complex Zernike moments of order n with repetition m are defined as

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V^*_{nm}(x,y) = \frac{n+1}{\pi}\sum_r\sum_\theta f(r\cos\theta, r\sin\theta)R_{nm}(r) \times \exp(\mathrm{j}m\theta), \quad r\leqslant 1, \tag{2.3}$$

where $f(x,y)$ is a binary shape function, $V^*(x,y)$ is the complex conjugate of $V(x,y)$. A list of Zernike moments up to order 10 is given in Table 1 [12]. The magnitudes of the acquired Zernike moments normalized by the mass of the shape are used as shape descriptors.

It can be seen from (2.3) that the basis of Zernike moments $R_{nm}(\rho)\exp(\mathrm{j}m\theta)$ only reflects angular frequency in its trigonometric harmonic. This indicates that the radial spectral features of the shape are not captured in ZMD. Furthermore, the repetition of m in each order n of the basis reduces the number of angular frequencies each order of Zernike moment (or coefficient) captures. This indicates that the circular spectral features captured by ZMD are too coarse if the number of moments used is not sufficiently large. For example, the number of angular frequencies captured by the first 36 Zernike moments is 10. In other words, if 36 Zernike moments are used as shape descriptor, then the descriptor only captures 10 circular features. More circular features can be otherwise captured if there is no repetition in each order of the basis. To prove this fact, we propose a variation of ZMD (VZM) in (2.4). The variation is an extension to Zernike moments in a way by removing the repetition in each order of Zernike moment. It is given by

$$\mathrm{VF}(u) = \sum_r\sum_\theta f(r\cos\theta, r\sin\theta)r\exp(\mathrm{j}u\theta), \tag{2.4}$$

where r and θ have the same meanings as those in (2.1). Eq. (2.4) has simpler form than (2.3).

Table 1
List of Zernike moments up to order 10

Order (n)	Zernike moment of order n with repetition m (A_{nm})	Number of moments in each order n	Total number of moments up to order 10
0	$A_{0,0}$	1	
1	$A_{1,1}$	1	
2	$A_{2,0}$, $A_{2,2}$	2	
3	$A_{3,1}$, $A_{3,3}$	2	
4	$A_{4,0}$, $A_{4,2}$, $A_{4,4}$	3	
5	$A_{5,1}$, $A_{5,3}$, $A_{5,5}$	3	36
6	$A_{6,0}$, $A_{6,2}$, $A_{6,4}$, $A_{6,6}$	4	
7	$A_{7,1}$, $A_{7,3}$, $A_{7,5}$, $A_{7,7}$	4	
8	$A_{8,0}$, $A_{8,2}$, $A_{8,4}$, $A_{8,6}$, $A_{8,8}$	5	
9	$A_{9,1}$, $A_{9,3}$, $A_{9,5}$, $A_{9,7}$, $A_{9,9}$	5	
10	$A_{10,0}$, $A_{10,2}$, $A_{10,4}$, $A_{10,6}$, $A_{10,8}$, $A_{10,10}$	6	

However, the same number of transformed coefficients VF(u) capture more circular features than the same number of Zernike moments A_{nm} do. The retrieval effectiveness of $F(u)$ will be shown in Section 3. Similar to Zernike moments, VF(u) does not capture radial features.

To summarize this section, we conclude that the Zernike polynomials $R_{nm}(r)$ only add weight to the bases. They contribute little to the capturing of shape features. Furthermore, the polynomials create a number of repetitions in each order of the calculated moment. The repetitions are actually the weighted moments of the previous orders. In other words, for the number of Zernike moments calculated in each order, only one is essentially important, the others are the repetitions of the moments of the previous orders. The repetitions in each order can be saved for capturing radial features. With this idea in mind, we attempt to use 2-D polar FT (PFT) instead of Zernike moments. 2-D polar FT allows multiresolution in both radial and angular directions, and the same number of FT coefficients can capture more shape features than ZMD. In the next section, we introduce PFT that captures both radial and circular features from formal FT theory.

2.2. Polar Fourier transform

Fourier transform has been widely used for image processing and analysis. The advantage of analyzing image in spectral domain over analyzing shape in spatial domain is that it is easy to overcome the noise problem which is common to digital images. Besides, the spectral features of an image are usually more concise than the features extracted from spatial domain. One-dimensional FT has been successfully applied to contour shape to derive FD as has been mentioned in Section 2.1. The application of one-dimensional FT on shape assumes the knowledge of shape boundary information. There is no reported work on region-based FD. In this section we introduce generic FD derived from 2-D PFT.

The continuous and discrete 2-D Fourier transform of a shape image $f(x,y)$ ($0 \leqslant x < M$, $0 \leqslant y < N$) are given by (2.5) and (2.6), respectively.

$$F(u,v) = \int_x \int_y f(x,y) \times \exp[-\mathrm{j}2\pi(ux+vy)]\,\mathrm{d}x\,\mathrm{d}y, \tag{2.5}$$

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \times \exp[-\mathrm{j}2\pi(ux/M+vy/N)]. \tag{2.6}$$

The u and v in (2.6) are the uth and vth spatial frequency in horizontal and vertical direction, respectively. 2-D FT can be directly applied to any shape image without assuming the knowledge of boundary information. However, direct applying 2-D FT on a shape image in Cartesian space to derive FD is not practical because the features captured by 2-D FT are not rotation invariant. Rotation invariance of a shape is important because similar shapes can be under different

orientations. For example, the two patterns (shapes) in Fig. 3(a) and (b) are similar patterns (shapes); however, their Fourier spectra distributions (Fig. 3(c) and (d)) on frequency plane are different. The difference of feature distributions makes it impractical to match the two patterns, especially online.

Therefore, we consider shape image in polar space and applying PFT on shape image. The PFT produces rotation-invariant data particularly well suited for accurate extraction of orientation features. In the following, we study and describe two PFTs. The study is necessary, because theoretically sound method may not readily applicable for implementation.

To derive PFT, both the data $f(x,y)$ and the spectra $F(u,v)$ are put into polar space, that is, let

$$\begin{aligned} x &= r\cos\theta, \quad y = r\sin\theta, \\ u &= \rho\cos\psi, \quad v = \rho\sin\psi. \end{aligned} \tag{2.7}$$

(r,θ) is the polar coordinates in image plane and (ρ,ψ) is the polar coordinates in frequency plane. The definition of (r,θ) and (ρ,ψ) is the same as that in (2.1). The differentials of x and y are

$$\begin{aligned} \mathrm{d}x &= \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta, \\ \mathrm{d}y &= \sin\theta\,\mathrm{d}r + r\sin\theta\,\mathrm{d}\theta. \end{aligned} \tag{2.8}$$

The Jacobian of (2.8) is r. By replacing (2.7) and (2.8) into (2.5) we have the PFT1,

$$\mathrm{PF}_1(\rho,\psi) = \int_r \int_\theta rf(r,\theta) \times \exp[-\mathrm{j}2\pi r\rho\sin(\theta+\psi)]\,\mathrm{d}r\,\mathrm{d}\theta. \tag{2.9}$$

The discrete PFT1 is then obtained as

$$\mathrm{PF}_1(\rho_l,\psi_m) = \sum_p \sum_i f(r_p,\theta_i) r_p \times \exp[-\mathrm{j}2\pi r_p\rho_l\sin(\theta_i+\psi_m)], \tag{2.10}$$

where $r_p = p/R$, $\theta_i = i(2\pi/T)$ $(0 \leqslant i < T)$; $\rho_l = l$ $(0 \leqslant l < R)$ and $\psi_m = m\theta_i$. R and T are the resolution of radial frequency and angular frequency, respectively. The acquired polar Fourier coefficients $F(\rho,\psi)$ are used to derive normalized FD for shape representation.

PFT1 is the direct result from the polar transform of (2.5). However, due to the presence of ψ_m within sin function $\sin(\theta_i+\psi_m)$, the physical meaning of ψ_m is not the mth angular frequency. The features captured by the PFT1 lose physical meaning in circular direction. To overcome the problem, a modified polar FT (PFT2) is derived by treating the polar image in polar space as a normal two-dimensional rectangular image in Cartesian space. Fig. 4 demonstrate the rectangular polar images. Fig. 4(a) is the original shape image in polar space, Fig. 4(b) is the rectangular polar image plotted into Cartesian space. Fig. 4(c) is a binary trade mark image, (d) is the polar-raster sampled image plotted in Cartesian space.

The polar image of Fig. 4(b) is the normal rectangular image. Therefore, if we apply 2-D FT on this rectangular image, the polar FT has the similar form to the normal 2-D discrete FT of (2.6) in Cartesian space. Consequently, the modified polar FT is obtained as

$$\mathrm{PF}_2(\rho,\phi) = \sum_r \sum_i f(r,\theta_i) \times \exp\left[-\mathrm{j}2\pi\left(\frac{r}{R}\rho + \frac{2\pi\mathrm{i}}{T}\phi\right)\right], \tag{2.11}$$

where $0 \leqslant r < R$ and $\theta_i = i(2\pi/T)$ $(0 \leqslant i < T)$; $0 \leqslant \rho < R$, $0 \leqslant \phi < T$. R and T are the radial and angular resolutions. $f(x,y)$ is a binary function in shape application. PFT2 has a simpler form than ZMD and PFT1. There is no need to constrain the shape into a unit circle (the constraint requires

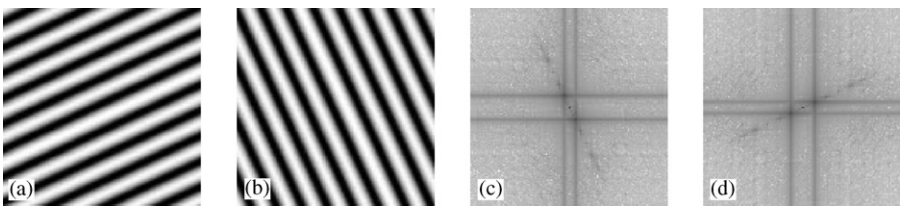


Fig. 3. (a) A pattern; (b) pattern (a) rotated by 90°; (c) Fourier spectra of (a); (d) Fourier spectra of (b).

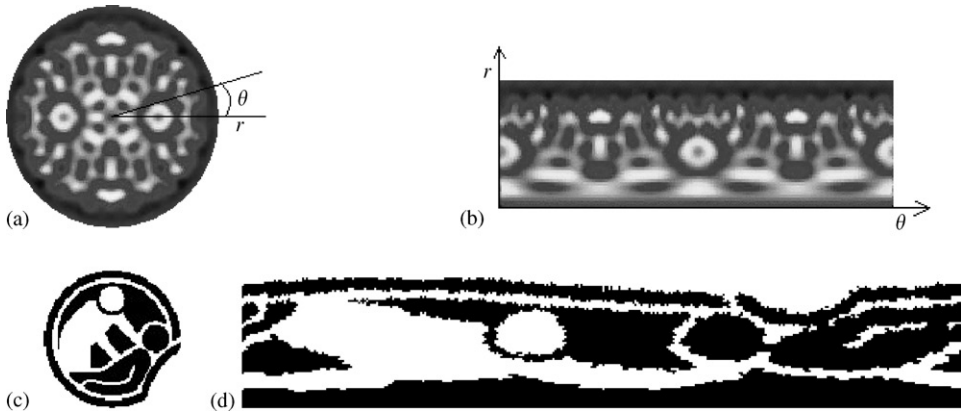


Fig. 4. (a) An original shape image in polar space; (b) the polar image of (a) plotted into Cartesian space; (c) a trade mark shape image; (d) the polar image of (c) plotted into Cartesian space.

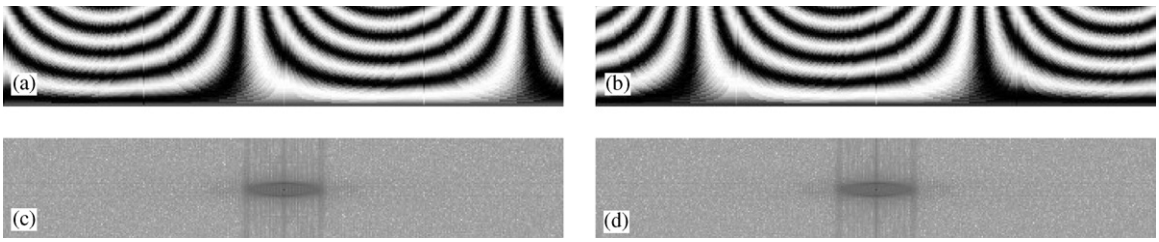


Fig. 5. (a,b) Polar images of the two patterns in Fig. 3(a) and (b); (c) Fourier spectra of (a); (d) Fourier spectra of (b).

a extra scale normalization in spatial domain) as required in the implementation of ZMD (because Zernike moment is defined in a unit circle). Moreover, the physical meaning of ρ and ϕ is similar to u and v in (2.6). The ρ and ϕ are simply the number of radial frequencies selected and the number of angular frequencies selected. The determination of ρ and ϕ is physically achievable, because shape features are usually captured by the few low frequencies.

Fig. 5(a,b) shows the polar images of the two patterns in Fig. 3(a) and (b) and their polar Fourier spectra are shown in (c) and (d). It can be observed from Fig. 5 that rotation of pattern in Cartesian space results in circularly shift in polar space. The circular shift does not change the spectra distribution on polar space. This is demonstrated in Fig. 5(c) and (d). The polar Fourier spectra is more concentrated around the origin of the polar space. This is particularly well suited for shape representation, because for efficient shape representation, the number of spectra features selected to describe the shape should not be large. Since $f(x, y)$ is a real function, the spectra is circular symmetric, only one quarter of the spectra features are needed to describe the shape. The acquired polar Fourier coefficients $F(\rho, \phi)$ are used to derive normalized FD for shape representation.

2.3. *Derivation of generic* FD

In this section, the derivation of FD using the above-described VZM and PFT is given in details. The VZM and the two polar FTs: PFT1 and PFT2 are all implemented in the experiments to derive FD in the purpose to determine which is the most appropriate for shape retrieval.

Given a shape image $I = \{f(x, y);\ 0 \leqslant x < M,\ 0 \leqslant y < N\}$. To apply VZM and PFT, the shape image is converted from Cartesian space to polar space $I_p = \{f(r, \theta);\ 0 \leqslant r < R,\ 0 \leqslant \theta < 2\pi\}$, R is the maximum radius of the shape. The origin of the polar space is set to be the centroid of the

shape, so that the shape is translation invariant. The centroid (x_c, y_c) is given by

$$x_c = \frac{1}{M}\sum_{x=0}^{N-1} x, \quad y_c = \frac{1}{N}\sum_{y=0}^{M-1} y \tag{2.12}$$

and

$$r = \sqrt{(x-x_c)^2 + (y-y_c)^2}, \quad \theta = \arctan\frac{y-y_c}{x-x_c}. \tag{2.13}$$

The VZM and PFTs are applied on I_p. The acquired coefficients of the three transform are translation invariant due to the use of centroid as polar space origin. Rotation invariance is achieved by ignoring the phase information in the coefficients and only retaining the magnitudes of the coefficients. To achieve scale invariance, the first magnitude value is normalized by the area of the circle (area) in which the polar image resides or the mass of the shape (mass), and all the other magnitude values are normalized by the magnitude of the first coefficient. The translation, rotation and scale normalized PFT coefficients are used as the shape descriptors. To summarize, the shape descriptor derived from VZM and the FD derived from PFT1 and PFT2 are **VZMD**, **FD1** and **FD2**, respectively, they are shown as follows:

$$\mathbf{VZMD} = \left\{\frac{|\mathrm{VF}(0)|}{\mathrm{mass}}, \frac{|\mathrm{VF}(1)|}{|\mathrm{VF}(0)|}, \ldots, \frac{|\mathrm{VF}(n)|}{|\mathrm{VF}(0)|}\right\},$$

$$\mathbf{FD1} = \left\{\frac{|\mathrm{PF}_1(0,0)|}{\mathrm{mass}}, \frac{|\mathrm{PF}_1(0,1)|}{|\mathrm{PF}_1(0,0)|}, \ldots, \frac{|\mathrm{PF}_1(0,n)|}{|\mathrm{PF}_1(0,0)|}, \ldots, \frac{|\mathrm{PF}_1(m,0)|}{|\mathrm{PF}_1(0,0)|}, \ldots, \frac{|\mathrm{PF}_1(m,n)|}{|\mathrm{PF}_1(0,0)|}\right\},$$

$$\mathbf{FD2} = \left\{\frac{|\mathrm{PF}_2(0,0)|}{\mathrm{area}}, \frac{|\mathrm{PF}_2(0,1)|}{|\mathrm{PF}_2(0,0)|}, \ldots, \frac{|\mathrm{PF}_2(0,n)|}{|\mathrm{PF}_2(0,0)|}, \ldots, \frac{|\mathrm{PF}_2(m,0)|}{|\mathrm{PF}_2(0,0)|}, \ldots, \frac{|\mathrm{PF}_2(m,n)|}{|\mathrm{PF}_2(0,0)|}\right\},$$

where m is the maximum number of the radial frequencies selected and n is the maximum number of angular frequencies selected. m and n can be adjusted to achieve hierarchical coarse to fine representation requirement.

For efficient shape description, only a small number of the acquired descriptors are selected for shape representation. The selected descriptors form a feature vector which is used for indexing the shape. For two shapes represented by their Fourier descriptors, the similarity between the two shapes is measured by the Euclidean distance between the two feature vectors of the shapes. Therefore, the online matching is efficient and simple.

2.4. *Implementation of* GFD

The implementation of GFD can be summarized into four steps, translation normalization, polar Fourier transform, rotation normalization and scale normalization. The algorithm of deriving GFD using PFT2 is given in Fig. 6. The algorithms of deriving VZMD and GFD using PFT1 are similar, with only difference in the basis calculation of polar Fourier transform step and scaling normalization step.

3. Test of retrieval effectiveness

In order to test retrieval effectiveness of the proposed methods, three sets of experiments are conducted. The first experiment is to compare the three proposed methods to determine which is the most suitable for shape retrieval. The other two experiments are to compare the proposed GFD with contour FD and MPEG-7 shape descriptors.

3.1. *Comparison of VZMD, FD1 and FD2*

To test the retrieval effectiveness of the VZMD and the two FDs derived from PFT1 and PFT2, a Java-based indexing and retrieval framework is implemented. The framework runs on Windows platform of a Pentium III-866 PC. The retrieval effectiveness of the VZMD and the two types of FD described in Section 2.4 is tested on the region-based shape database of MPEG-7. MPEG-7 region shape database consists of 3621 shapes of mainly trademarks. 651 shapes from 31 classes of shapes are selected as queries. The 31 classes of

1. Input binary shape image data $f(x, y)$;
2. Get centroid of the shape (x_c, y_c);
3. Set the centroid as the origion; /* translation normalization */
4. Get the maximum radius of the shape image (*maxRad*);
5. Polar Fourier transform

```
For radial frequency (rad) from zero to maximum radial frequency (m)
  For angular frequency (ang) from zero to maximum angular frequency (n)
    For x from zero to width of the shape image
      For y from zero to height of the shape image
      {
        radius = square root[(x-maxRad)^2 + (y-maxRad)^2];
        theta = arctan2[(y-maxRad)/(x-maxRad)];  /* theta falls within [–π, +π] */
        if(theta<0) theta += 2π;  /* extend theta to [0, 2π] */
        FR[rad][ang] += f(x,y)×cos[2π×rad×(radius/maxRad) + ang×theta]; /* real part of spectra */
        FI[rad][ang]  = f(x,y) ×sin[2π×rad×(radius/maxRad) + ang×theta]; /* imaginary part of spectra */
      }
```

6. Calculate FD

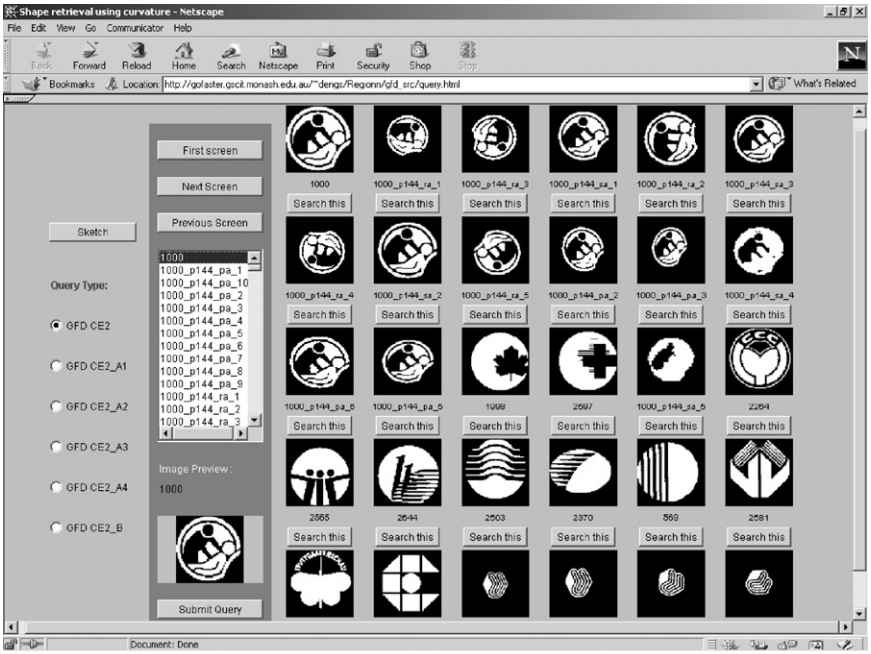
```
For rad from zero to m
  For ang from zero to n
  {
    /* rotation and scale normalization */
    If (rad=0 & ang=0)
      FD[0] = square root[(FR^2[0][0] + FR^2[0][0])/(π×maxRad^2)];
    Else
      FD[rad×n+ang] = square root[(FR^2[rad][ang] + FI^2[rad][ang])/ FD[0]];
  }
```

7. Output feature vector **FD**.

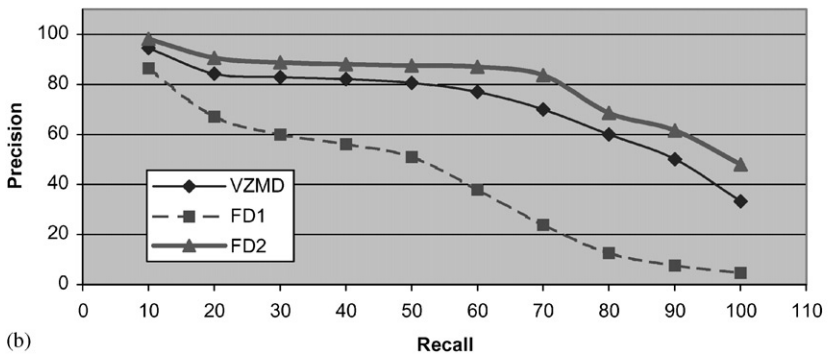
Fig. 6. Procedure of computing GFD from PFT2.

shapes reflect general variations (rotation, scaling and perspective transform) of shapes. Each class has 21 members generated through scaling, rotation and perspective transformation. Since the IDs of all the similar shapes to each query in the classes are known, the retrieval is done automatically. However, the retrieval system is also put online to test real-time retrieval. For online retrieval, the indexed data and the shape databases are put in a web server, user can do online retrieval by visiting the retrieval site using either common browsers or Java appletviewer (http://www.gscit.monash.edu.au/~dengs/) (Fig. 7(a)).

Common performance measure, i.e., precision and recall of the retrieval [2,24,25], are used as the evaluation of the query result. 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. $P = r/n$. Precision P measures the accuracy of the retrieval and the speed of the recall. Recall R is defined as the ratio of the number of retrieved relevant images r to the



(a)



(b)

Fig. 7. (a) Retrieval online using browser; (b) comparison of retrieval effectiveness of VZMD, FD1 and FD2.

total number m of relevant shapes in the whole database, i.e. $R = r/m$. Recall R measures the robustness of the retrieval.

For each query, the precision of the retrieval at each level of the recall is obtained. The result precision of retrieval using a type of shape descriptors is the average precision of all the query retrievals using the type of shape descriptors. The average precision and recall of the 651queries using the three derived shape descriptors are shown in Fig. 7(b).

It is clear from Fig. 7(b) that FD derived from PFT2 outperforms VZMD and FD derived from PFT1. The fact that FD2 outperforms VZMD and FD1 significantly indicates that FD derived from PFT2 is the most suitable for shape description. Therefore, FD derived from PFT2 is selected as the generic FD (GFD) as shape representation. Hereafter, GFD refers to FD derived using PFT2.

60 GFDs (reflecting 5 radial frequencies and 12 angular frequencies) are selected as shape descriptors. However, different number of FDs with different parameters are tested to decide which is the most appropriate number of FDs to describe the shape. The test results are given in Table 1. From Table 1, it is observed that retrieval effectiveness is improved by increasing radial resolution. However, retrieval effectiveness does not improve significantly when the radial resolution is greater than 3. It is also observed from the table that retrieval effectiveness does not improve significantly when the angular resolution is greater than 12. The observation indicates that for efficient retrieval, 36 GFDs (reflecting 3 radial frequencies and 12 angular frequencies) or 60 GFDs (reflecting 4 radial frequencies and 15 angular frequencies) is the suitable number of GFDs for shape description.

3.2. Comparison between GFD and contour-based shape descriptors

The above-derived GFD is compared with common shape contour-shape descriptors: 1-D FD and curvature scale space descriptor (CSSD) which has been adopted as contour shape descriptor in MPEG-7. The technique of 1-D FD has been briefly described in Section 2.1. The technical details of CSSD are described in [16], and the implementation details are given in [10]. Basically, CSSD applies iterative Gaussian smoothing (scale space) to the shape boundary. The curvature zero-crossing points at each scale are recorded. A CSS contour map is created consisting of zero-crossing points from all scales. The normalized peaks of the contour branches in the CSS contour map are used as shape descriptor.

The retrieval tests are conducted on MPEG-7 contour shape database (CE-1). MPEG-7 contour shape database CE-1 is composed of Set A1, A2, B and C which are for testing different types of robustness. The following explains how to use the database:

- Set A1 consists of 420 shapes which are organized into 70 groups. There are 6 similar shapes in each group. Set A1 is for test of scale invariance.
- Set A2 consists of 420 shapes which are organized into 70 groups. There are 6 similar shapes in each group. Set A2 is for test of rotation invariance.
- Set B consists of 1400 shapes of 70 groups. There are 20 similar shapes in each group. Set B is for similarity-based retrieval which tests overall robustness of the shape representations.
- Set C consists of 1300 marine fishes, 200 bream fishes are generated through affine transform with different parameters. Set C is for test of robustness of non-rigid deformations.

For Set A1, A2 and B, all the shapes in the database are used as queries. For Set C, the 200 bream fishes are used as queries. The common retrieval measurement precision–recall which is described in Section 3.1 is used for evaluation of the retrieval effectiveness. The average precision and recall of the retrieval using the three shape descriptors on each set are shown in Fig. 8(a)–(d). Some screen shots of retrieval on MPEG-7 contour shape database are shown in Fig. 9(a)–(g). In all the screen shots, the top left shape is the query shape. The retrieved shapes are ranked in descending order of similarity to the query shape, they are arranged in left to right and top to bottom order.

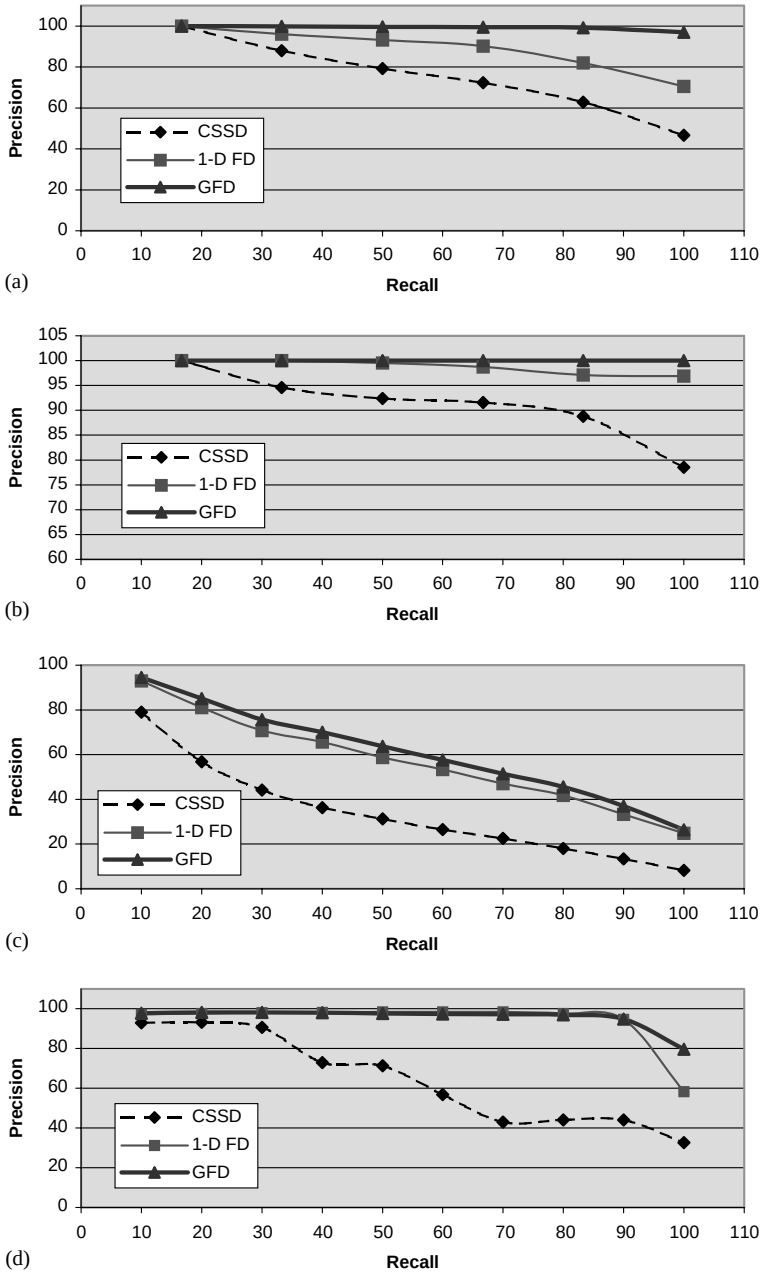


Fig. 8. Average precision–recall charts of retrieval on MPEG-7 contour shape database. (a) Average precision–recall of 420 queries in Set A1 of CE-1. (b) Average precision–recall of 420 queries in Set A2 of CE-1. (c) Average precision–recall of 1400 queries in Set B of CE-1. (d) Average precision–recall of 200 queries in Set C of CE-1.

It can be seen from Fig. 8 that the proposed GFD outperforms 1-D FD and CSSD on all the sets of MPEG-7 contour-shape database. The performance of CSSD is significantly lower than GFD and 1-D FD due to its complex normalization and matching [33]. GFD has 100% correct retrieval of rotated shapes. It has almost 100% correct retrieval of scaled shapes and non-rigid

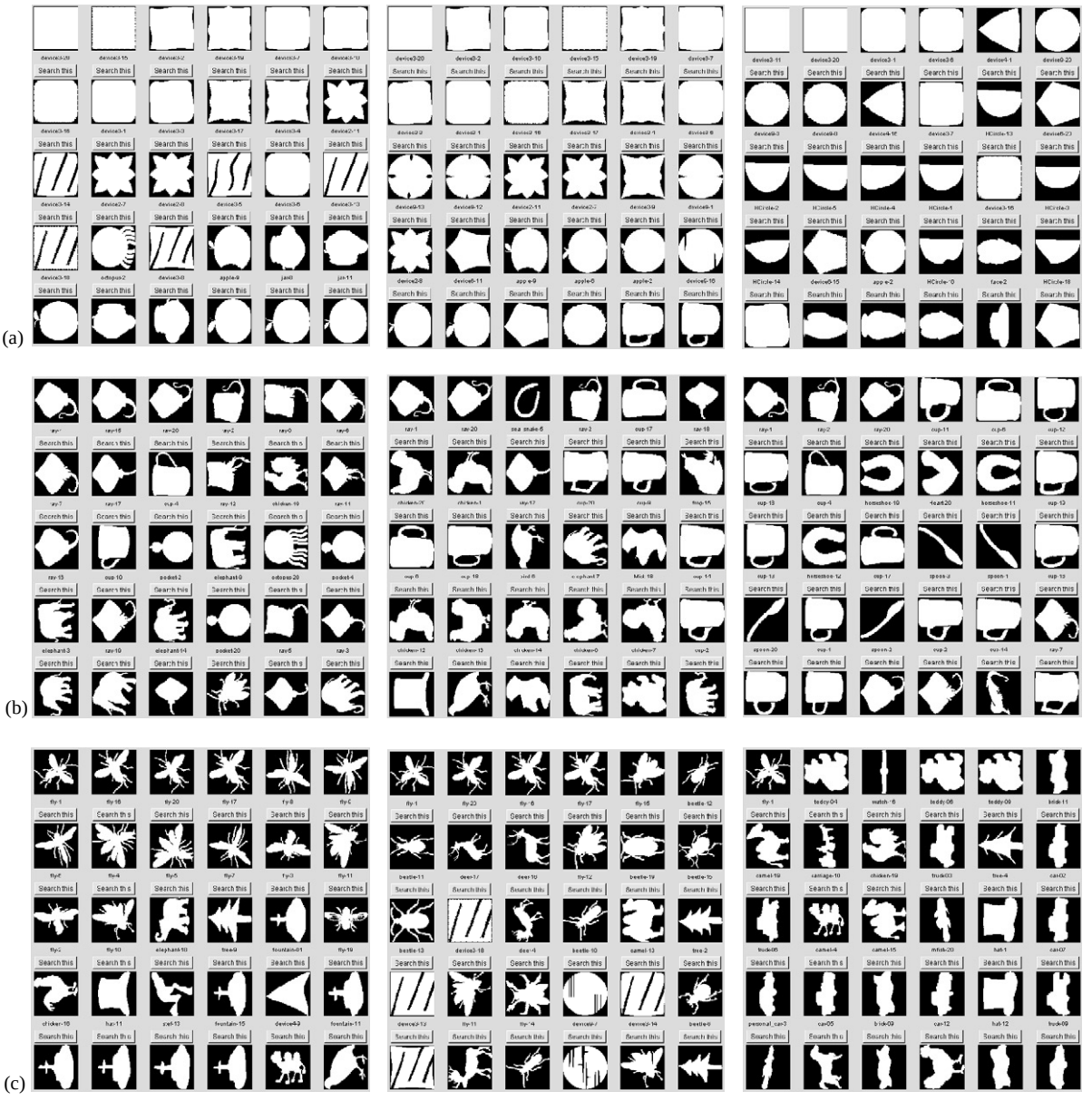


Fig. 9. Example retrievals using GFD, 1-D FD and CSSD on different sets of CE-1. (a) Retrieval of query device3-20 on Set B using (left) GFD; (middle) 1-D FD; (right) CSSD. (b) Retrieval of query ray-1 on Set B using (left) GFD; (middle) 1-D FD; (right) CSSD. (c) Retrieval of query fly-1 on Set B using (left) GFD; (middle) 1-D FD; (right) CSSD. (d) Retrieval of query fork-1 on Set B using (left) GFD; (middle) 1-D FD; (right) CSSD. (e) Retrieval of query bream-120 on Set C using (left) GFD; (middle) 1-D FD; (right) CSSD. (f) Retrievals of device6-1_SD_01, fly-1_SD_01 and butterfly-1 on Set A1 using (left) GFD; (middle) 1-D FD; (right) CSSD. (g) Retrieval of query spring-1 on Set A2 using (left) GFD; (middle) 1-D FD; (right) CSSD.

shapes. The advantage of GFD over contour-based shape descriptors is obvious in situations where severe protrusions and indentations occur. For example, in Fig. 9(a), GFD not only retrieves those distorted square shapes but also retrieves those squares with severe indentations. For the

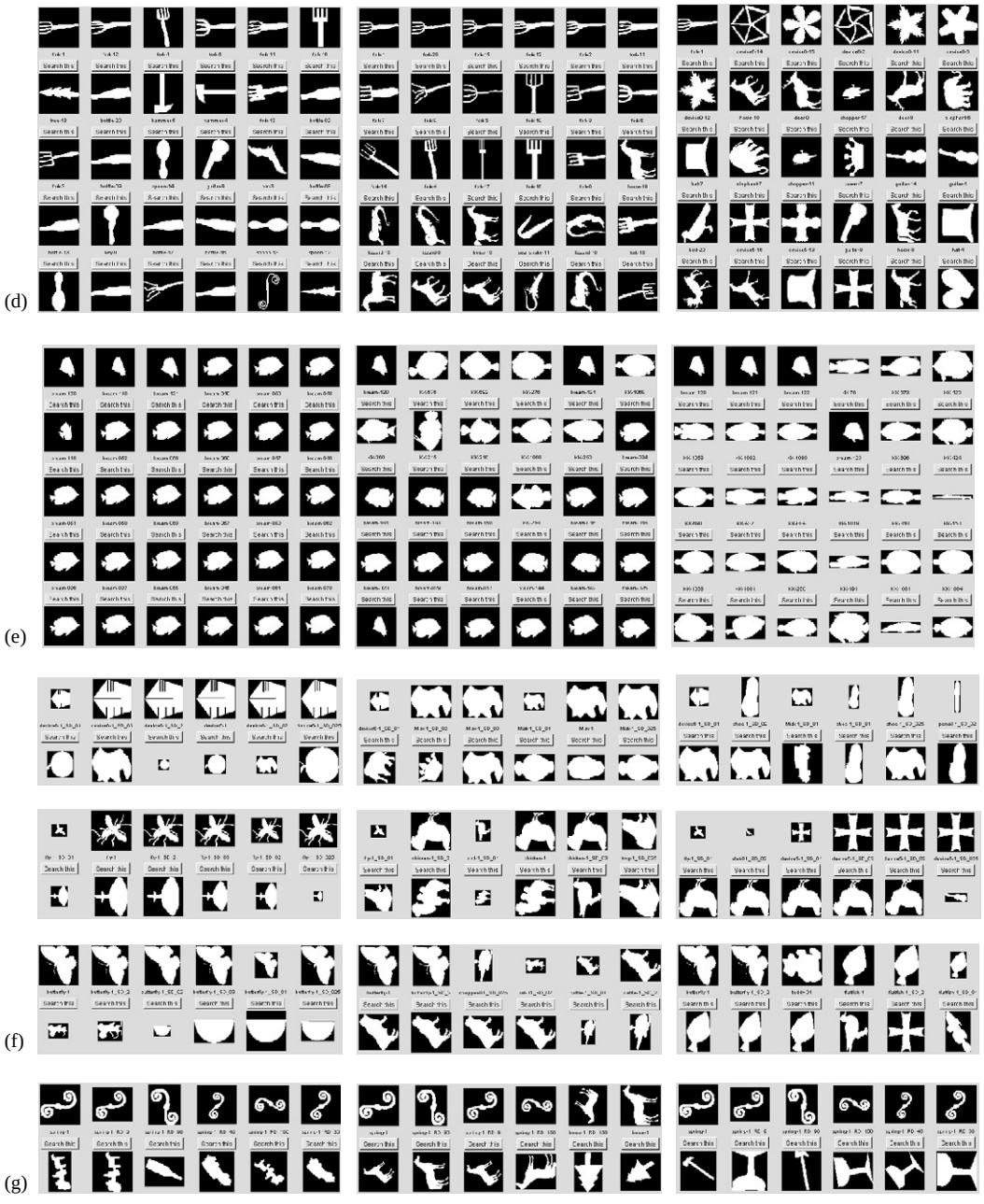


Fig. 9 (*continued*).

contour-shape descriptors, they can only retrieve square shapes without any indentations. In Fig. 9(b), GFD retrieves most of the ray fish shapes in the first screen; however, the contour shape descriptors are easily trapped by those protrusions of the shapes. They treat any shapes with hook-like parts as similar shapes to the ray fish shape which has a hook-like tail. Similar to (b), in Fig. 9(c), contour shape descriptors are easily distracted by the complex arms of the flies.

They are confused with shapes with similar protrusions, or similar number of protrusions. While GFD is able to concentrate on the main body of the fly shapes, successfully retrieve most of the fly shapes in the first screen.

When a shape is scaled, its boundary can be substantially changed. The contour shape descriptors can fail completely when large scaling (scaling factor larger than 2) occurs (Fig. 9(f)). However, GFD is not affected by large scaling.

GFD is also more robust to severe deformation of shape than the contour-shape descriptors. The fish bream-120 is a severely distorted shape; however, GFD correctly retrieve its similar shapes (Fig. 10(e)). 1-D FD only works better than GFD in situations where the protrusions and indentations constitute the main body of the shape. The fork shape in Fig. 9(d) consists only of protrusion parts. 1-D FD has very high performance on this shape.

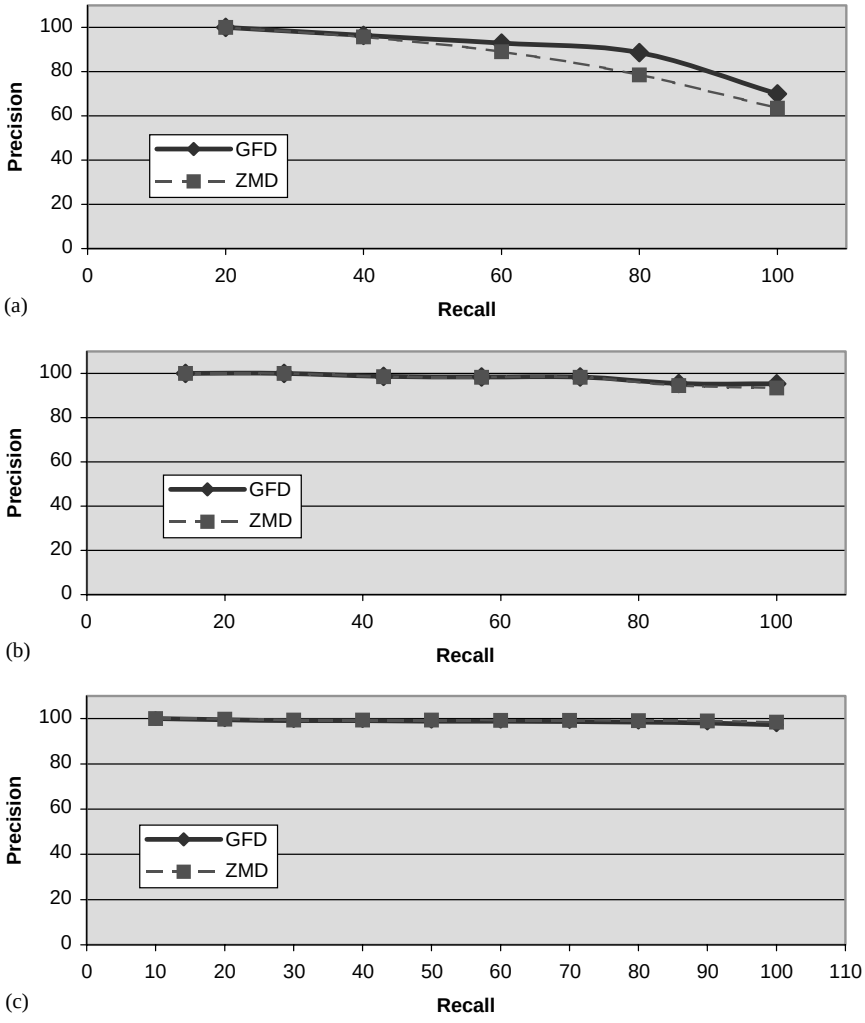
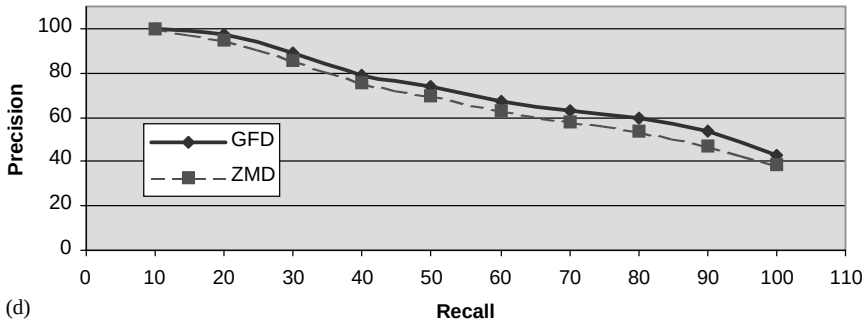
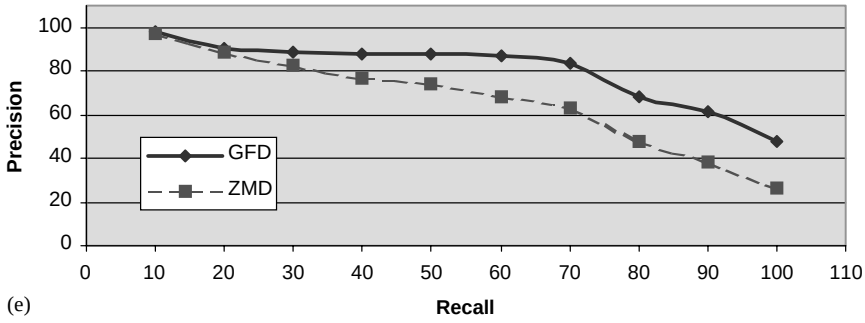


Fig. 10. Retrieval performance of GFD and ZMD on different sets of CE-2. (a) Average precision–recall of 100 queries in Set A1 of CE-2. (b) Average precision–recall of 140 queries in Set A2 of CE-2. (c) Average precision–recall of 330 queries in Set A3 of CE-2. (d) Average precision–recall of 330 queries in Set A4 of CE-2. (e) Average precision–recall of 651 queries in CE-2.



(d)



(e)

Fig. 10 (*continued*).

To summarize, because of using only boundary information, 1-D FD and CSSD are more likely affected by various types of shape variations such as scaling, protrusions, indentations and deformations. Due to using all the information within a shape region, GFD is more robust to shape variations than 1-D FD and CSSD.

Although the extraction of GFD requires more computation than the extraction of contour FD, the computation of online matching using GFD is about the same as that using contour FD. Because the number of GFDs used to index the shape is about the same as the number of FDs used to index the shape, and both matching using Euclidean distance [33,11]. For image retrieval application, low computation of online matching is essential while low computation of offline feature extraction is not as essential.

3.3. Comparison between GFD and region-based shape descriptors

The GFD is also compared with ZMD which is adopted as region-based shape descriptor in MPEG-7. The comparison is conducted on MPEG-7 region-based shape database (CE-2). MPEG-7 region shape database CE-2 consists of 3621 shapes of mainly trademarks. It is organized as 5 sets for testing different types of robustness. The use of the database is summarized as follows:

- Set A1 consists of 2881 shapes from the whole database, it is for test of scale invariance. One hundred shapes in Set A1 are organized into 20 groups (5 similar shapes in each group) which can be used as queries for test of retrieval. In our experiment, all the 100 shapes from the 20 groups are used as queries to test the retrieval.
- Set A2 consists of 2921 shapes from the whole database, it is for test of rotation invariance. 140 shapes in Set A2 are organized into 20 groups (7 similar shapes in each group) which can be used as queries for test of retrieval. In our experiment, all the 140 shapes from the 20 groups are used as queries to test the retrieval.
- Set A3 consists of 3101 shapes from the whole database, it is for test of rotation/scale invariance. 330 shapes in Set A3 are organized into

30 groups (11 similar shapes in each group) which can be used as queries for test of retrieval. In our experiment, all the 330 shapes from the 30 groups are used as queries to test the retrieval.

- Set A4 consists of 3101 from the whole database, it is for test of robustness to perspective transform. 330 shapes in Set A4 are organized into 30 groups (11 similar shapes in each group) which can be used as queries for test of retrieval. In our experiment, all the 330 shapes from the 30 groups are used as queries to test the retrieval.
- Set B consists of 2811 shapes from the whole database, it is for subjective test. 682 shapes in Set B are manually sorted out into 10 groups by MPEG-7. The grouping is rough even for a normal observer. Some of the members in groups 1 and 2 are shown in Fig. 11. In our experiment, all the 682 shapes from 10 classes are used as queries to test the retrieval.
- The whole database consists of 3621 shapes, 651 shapes of the 3621 shapes are organized into 31 groups (21 similar shapes in each groups). For the 21 similar shapes in each group, there are 10 perspective transformed shapes, 5 rotated shapes and 5 scaled shapes. The 31 groups of shapes reflect overall shape operations, and they test the overall robustness of a shape descriptor. The whole database is 17–29% larger in size than the individual sets.

For Set A1, A2, A3, A4 and the whole database, the precision–recall which described in Section 3 is used for evaluation of retrieval effectiveness. The average precision–recall of retrieval using the two shape descriptors on each set are shown in Fig. 10(a)–(e). For Set B, because the number of members in each group is different, the Bull's eye performance (BEP) is used for the evaluation of retrieval effectiveness. The BEP is measured by the correct retrievals among the top $2N$ retrievals, where N is the number of relevant (or similar shapes) shapes to the query in the database. The BEP of Set B is given in Tables 2 and 3.

It can be seen from Fig. 10 that there is only slight difference (overall precision is less than 1% different) of retrieval performance between GFD and ZMD on Sets A2 and A3. Both GFD and ZMD have very high performance on these two sets. However, the difference between GFD and ZMD on Set A1, A4 and Set B is obvious (difference of overall precision on each set is over 4%) and the difference between GFD and ZMD on the whole database is significant (difference of

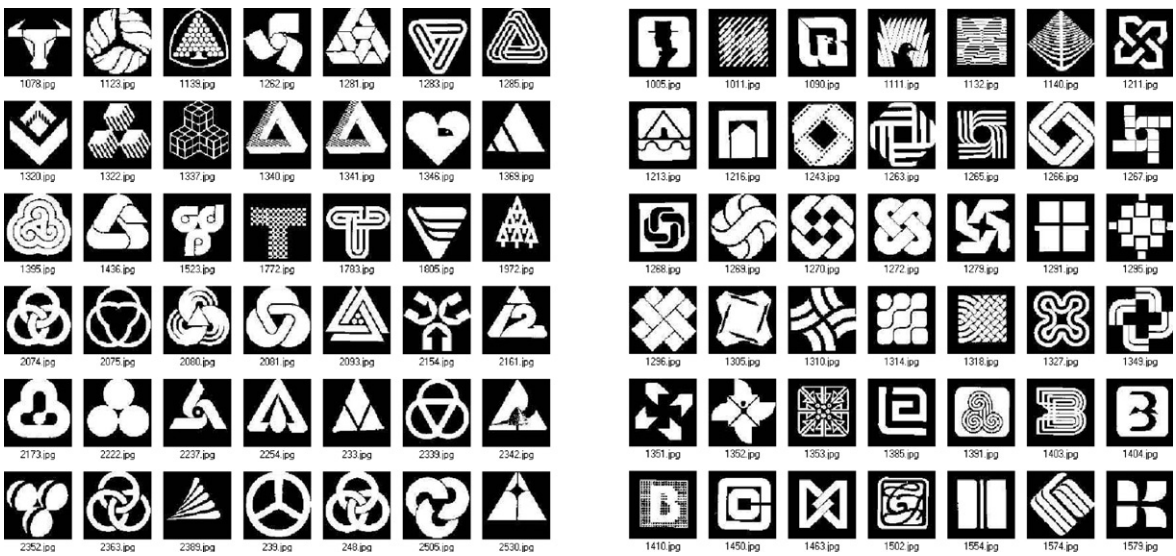


Fig. 11. Part of members in group 1 (left) and group 2 (right) of Set B of CE-2.

Table 2
Retrieval performance of FDs with different number of radial and angular frequencies. r: number of radial frequencies selected; t: number of angular frequencies selected

Recall (%) Parameters	10	20	30	40	50	60	70	80	90	100	Overall precision on full recall
$r = 1$, $t = 20$	94.4	85.5	83.4	82.3	81.0	76.4	70.5	59.7	53.0	37.6	72.4
$r = 2$, $t = 15$	96.3	89.0	87.2	86.5	86.0	85.0	82.0	66.5	60.8	46.5	78.6
$r = 3$, $t = 12$	97.6	90.6	88.9	88.0	87.6	87.0	84.0	68.5	62.0	48.8	80.3
$r = 3$, $t = 15$	97.8	90.7	89.0	88.2	87.7	87.2	84.3	68.6	62.5	48.9	80.5
$r = 4$, $t = 15$	98.2	90.8	89.2	88.4	88.0	87.4	84.1	69.0	62.7	48.3	80.6
$r = 4$, $t = 20$	98.3	91.0	89.4	88.5	88.1	87.5	84.5	69.1	63.0	48.8	80.8
$r = 5$, $t = 12$	98.3	90.8	88.9	88.0	87.7	87.1	84.0	68.8	62.0	48.2	80.4
$r = 5$, $t = 20$	98.3	91.0	89.1	88.3	87.9	87.3	84.3	68.9	62.3	48.7	80.6
$r = 6$, $t = 6$	97.4	88.6	86.8	86.0	85.7	84.7	81.0	66.4	58.0	44.0	77.9
$r = 8$, $t = 8$	97.8	89.6	87.7	87.0	86.7	85.7	82.2	68.1	60.2	46.8	79.2
$r = 10$, $t = 6$	97.4	88.7	86.7	86.0	85.6	84.6	80.8	66.7	58.3	44.7	78.0
$r = 8$, $t = 15$	98.3	90.6	88.7	87.8	87.3	86.9	83.1	68.7	62.1	48.5	80.2
$r = 10$, $t = 15$	98.3	90.7	88.7	87.8	87.3	87.0	83.2	68.7	62.1	48.5	80.2

Table 3
Bull's eye performance of the 682 queries in Set B of CE-2

Class	1	2	3	4	5	6	7	8	9	10	Average
No. of shapes	68	248	22	28	17	22	45	145	45	42	
GFD (%)	47.0	66.4	55.6	50.0	50.0	24.8	30.4	50.8	55.6	29.0	46.0
ZMD (%)	37.0	58.0	55.0	41.2	42.6	22.6	33.6	52.0	41.4	34.0	41.7

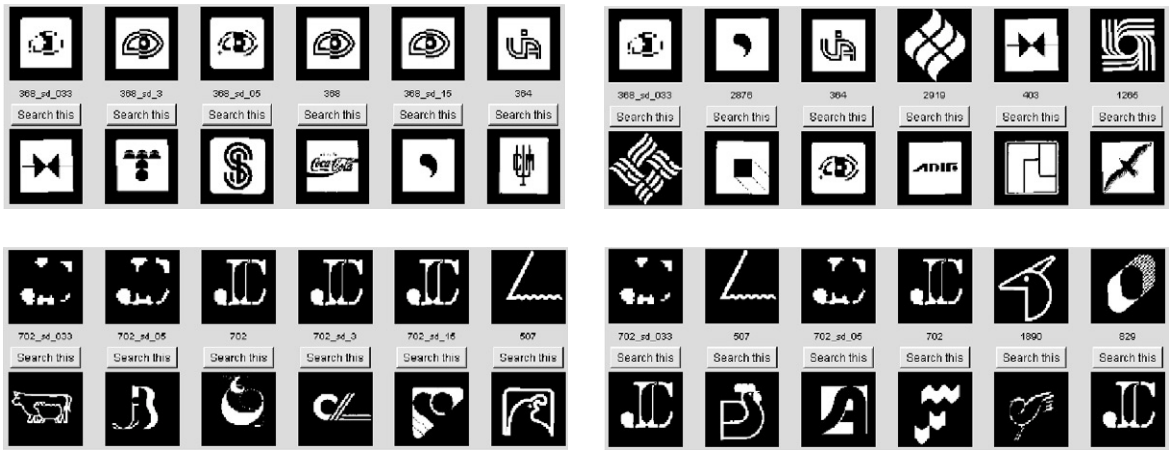


Fig. 12. Example retrievals on Set A1 of CE-2. (a) Retrieval of query 368_SD_033 using GFD (left) and using ZMD (right). (b) Retrieval of query 702_SD_033 using GFD (left) and using ZMD (right).

overall precision is over 12%). The reasons are explained as follows:

- Scaling, especially large scaling, can cause shape content or spatial distribution substantially changed. ZMD meets problem in dealing this type of situations because it is only able to examine shape in circular direction. However, GFD can successfully deal with this type of situations by examining shape more

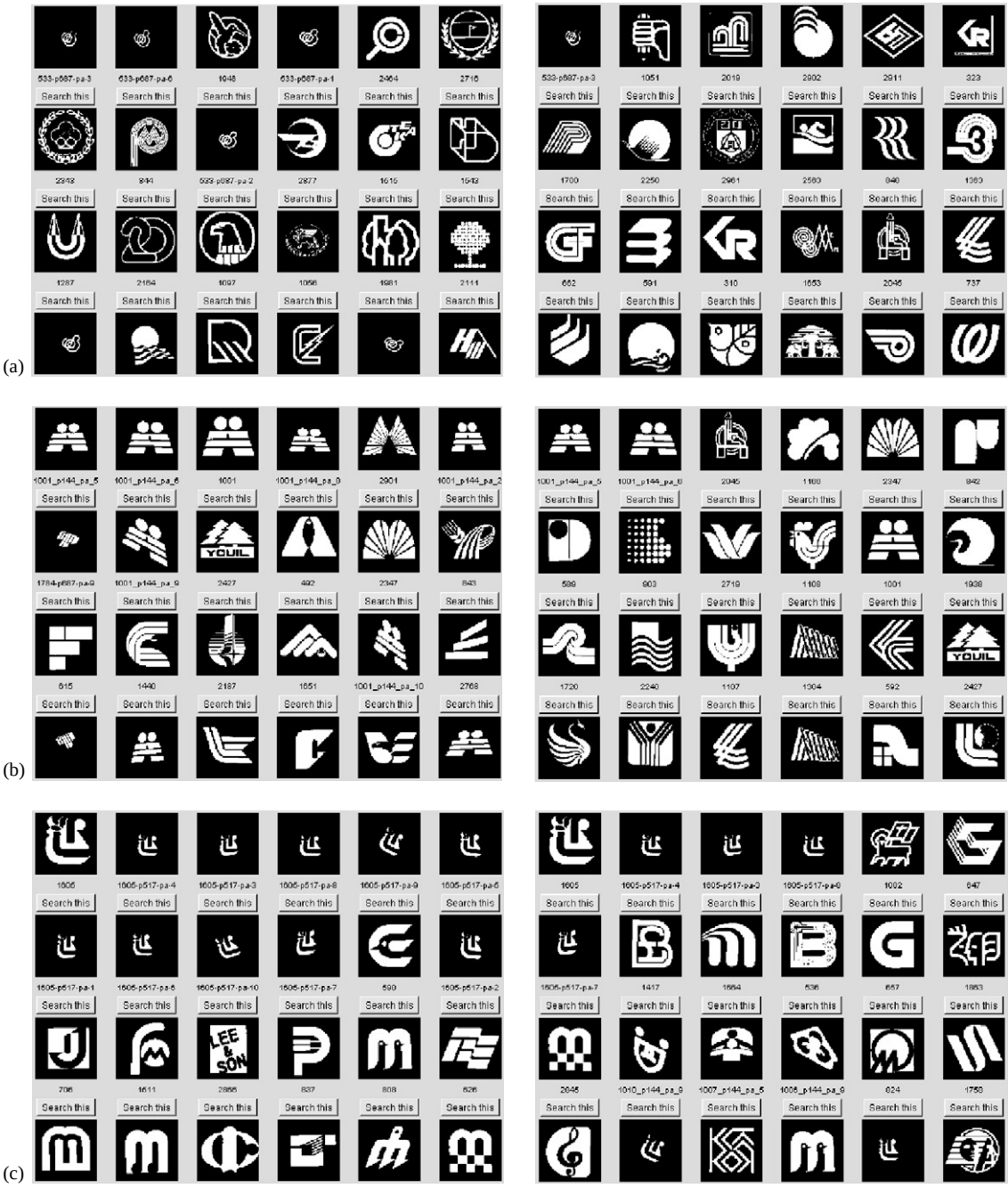


Fig. 13. Example retrievals on Set A4 of CE-2. (a) Retrieval of query 533_p687_pa_3 using GFD (left) and using ZMD (right). (b) Retrieval of query 1001_p144_pa_5 using GFD (left) and using ZMD (right). (c) Retrieval of query 1605 using GFD (left) and using ZMD (right).

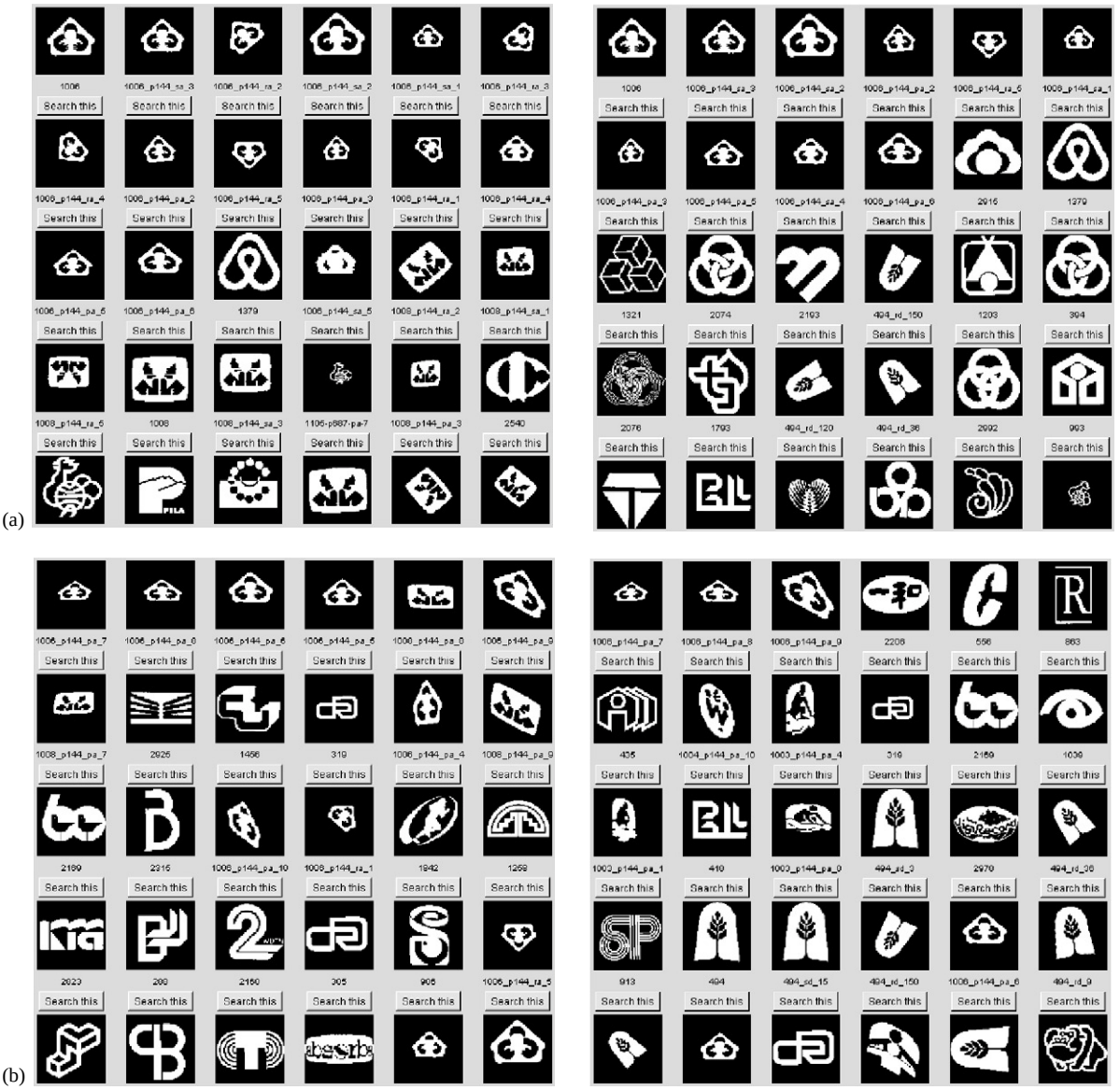


Fig. 14. Example retrievals on CE-2. (a) Retrieval of query 1006 using GFD and using ZMD. (b) Retrieval of query 1006_p144_pa_7 using GFD (left) and using ZMD (right). (c) Retrieval of query 1004_p144_pa_3 using GFD (left) and using ZMD (right).

carefully on the radial directions (Figs. 12(a,b) and 14(a)).

- Perspective deformations can also result in scaling effect, as a result, shape spatial distribution can be changed (Figs. 13(b) and 14(b,c)). Parts of shape can be lost due to the transform (Fig. 13(a)). GFD can cope with this type of situations by examine shape features in radial directions.
- Due to the capturing of shape features in both radial and circular directions, the retrieved shapes are more perceptually acceptable. For example, in Fig. 15, both GFD and ZMD retrieve all the similar shapes to the query. However, GFD not only retrieve those similar shapes, but also retrieve perceptually relevant shapes such as the members in group

(c)

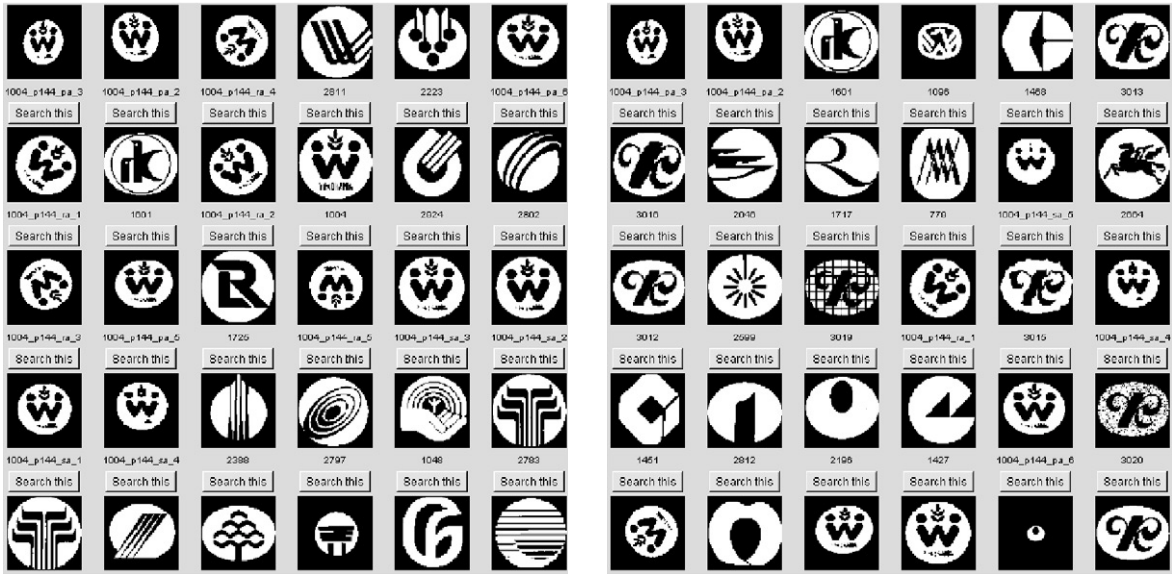
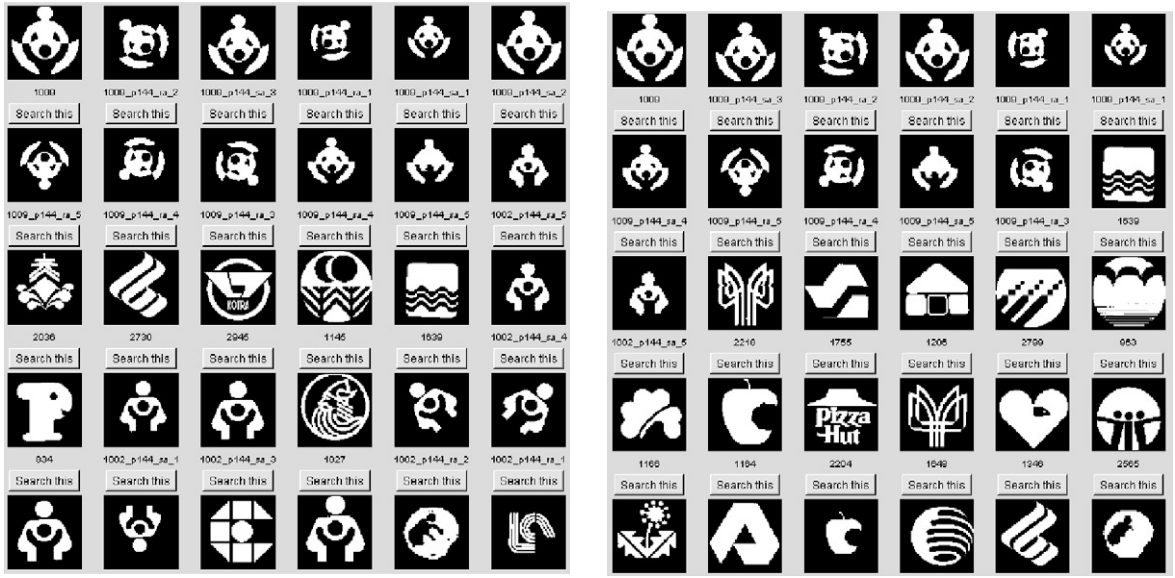


Fig. 14 (*continued*).



Retrieval of query 1009 using GFD (left) and using ZMD (right)

Fig. 15. Example retrievals on Set A3.

1002_p144. Examples from Set B (Fig. 16) and Set A4 (Fig. 13(c)) also demonstrate retrievals using GFD are more perceptually acceptable than ZMD.

- GFD is more robust than ZMD when the size of the shape database is increased. This is reflected in the retrieval performance on the whole database (Fig. 10(e)).

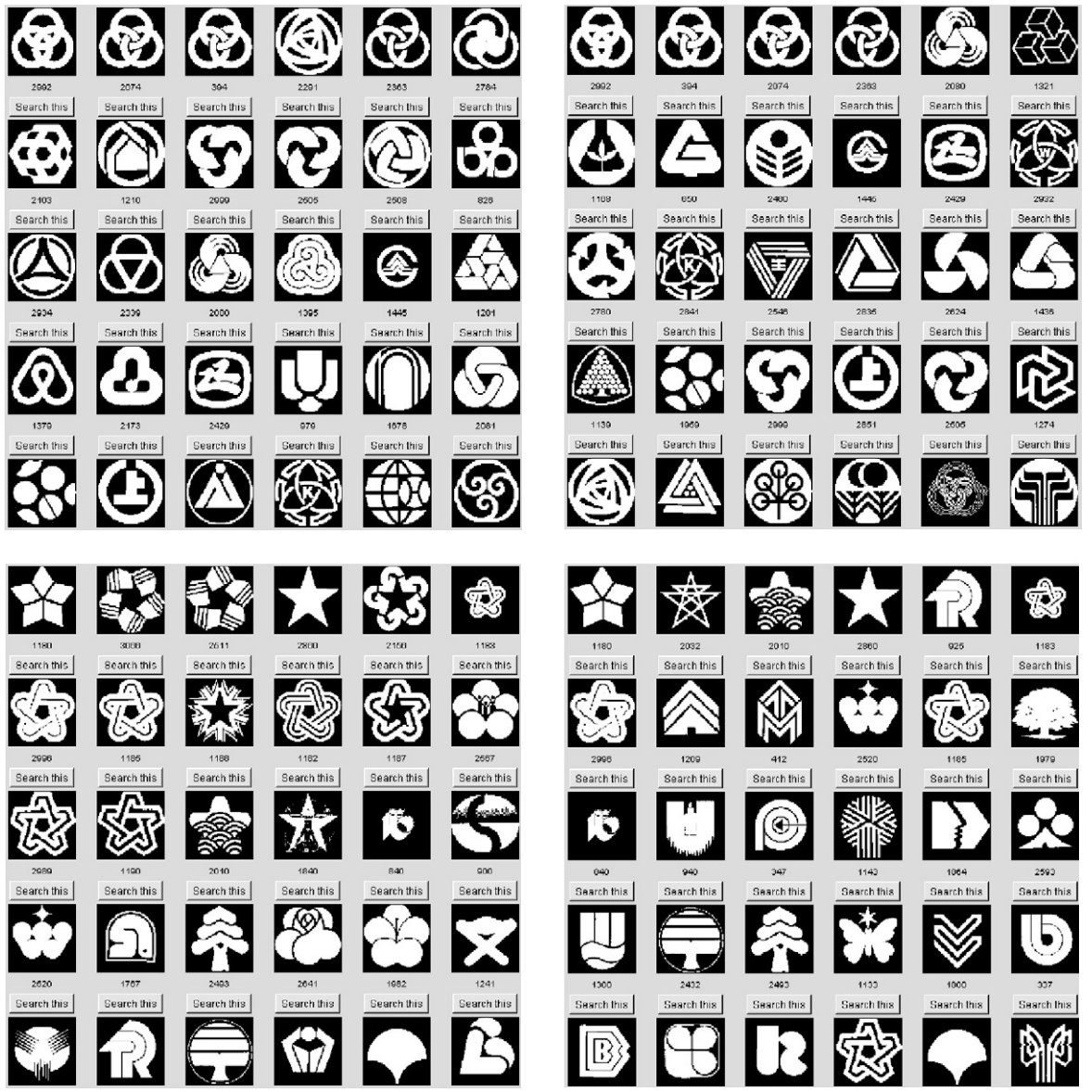
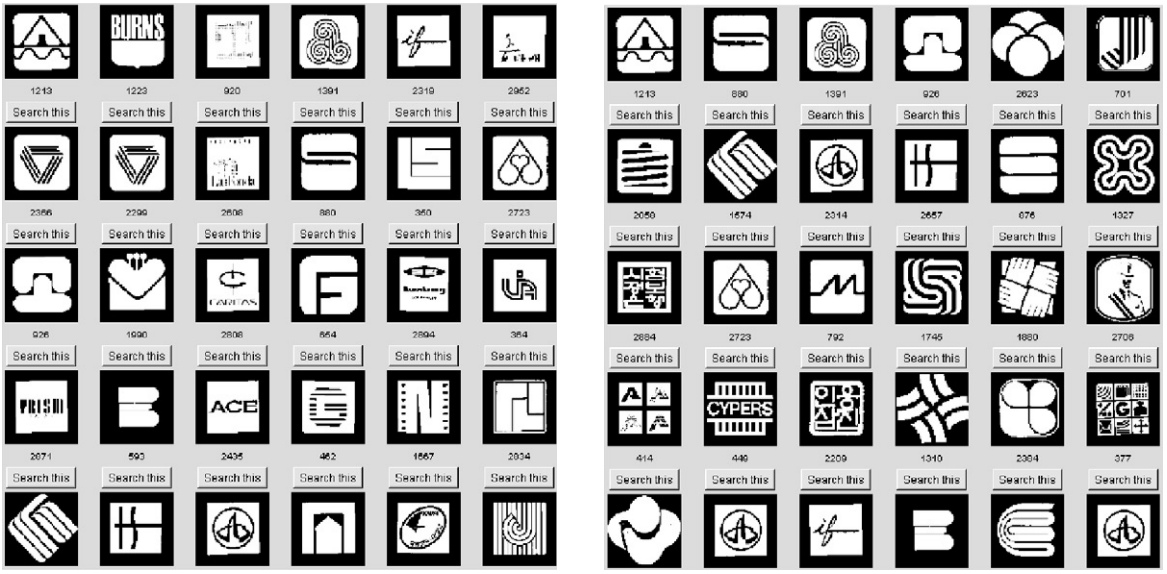


Fig. 16. Example retrievals on Set B of CE-2. (a) Retrieval of 2992 using GFD (left) and using ZMD (right). (b) Retrieval of 1180 using GFD (left) and using ZMD (right). (c) Retrieval of 1213 using GFD (left) and using ZMD (right). (d) Retrieval of 1011using GFD (left) and using ZMD (right).

The comparative low retrieval performance of GFD on Set B is due to that the grouping within the Set is too rough, as can been seen from the example shapes in Fig. 11. The comparative low retrieval performance of GFD on Set A4 is due to the circular scanning of shape with constant radius when applying PFT. This intrinsic problem will be considered in future implementation to increase GFD robustness to affine or perspective deformed shapes which are expected common in nature.

The computation of extracting GFD is simpler than ZMD. First, it does not need to normalize

(c)



(d)

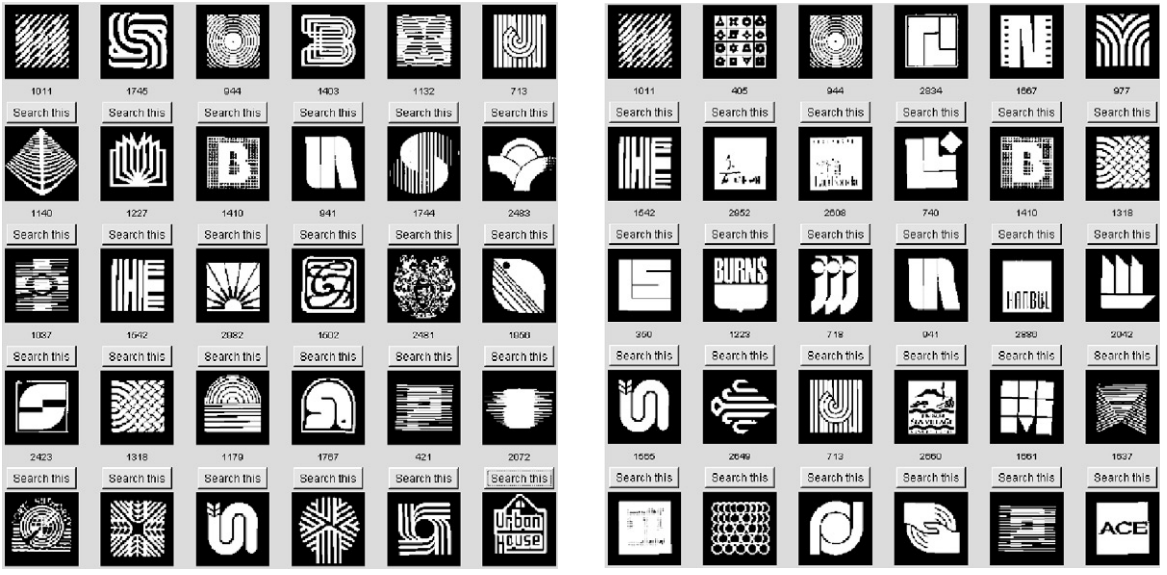


Fig. 16 (*continued*).

shape into an unit disk as is required in extracting ZMD (because Zernike moments is defined within a unit disk). Furthermore, the polar Fourier transform of (2.5) is simpler than the Zernike moments of (2.3). PFT avoids the complex computation of Zernike polynomials. The computations of online matching using GFD and ZMD are about the same, because both methods use Euclidean distance for similarity measurement and the number of GFDs used to index the shape is about the same as the number of ZMD used to index the shape [10].

4. Conclusions

In this paper, we have proposed a generic Fourier descriptor for general applications. The

main contributions of the paper are in the following two aspects.

- It improves common 1-D FD in that: (i) it does not assume shape contour information which may not be available; (ii) it captures shape interior content as well as shape boundary features; (iii) it is more robust.
- It improves ZMD in that: (i) it captures both radial and circular features of a shape; (ii) it is simpler in computation; (iii) it is more robust and perceptually meaningful.

The proposed GFD satisfies all the six requirements set by MPEG-7 for shape representation, that is, good retrieval accuracy, compact features, general application, low computation complexity, robust retrieval performance and hierarchical coarse to fine representation. It has been tested on both MPEG-7 contour shape database and MPEG-7 region shape database. Comparisons have been made between GFD, 1-D FD, and MPEG-7 shape descriptors, results show that the proposed GFD outperforms these shape descriptors. Compared with scaling and rotation retrieval, however, the retrieval accuracy for perspective transformed shapes is significantly lower. This problem will be studied in future research by using non-uniform polar-raster sampling.

References

[1] K. Arbter, W.E. Snyder, H. Burkhardt, G. Hirzinger, Application of affine-invariant Fourier descriptors to recognition of 3-D objects, IEEE Trans. PAMI 12 (7) (1990) 640–647.

[2] A.D. Bimbo, Visual Information Retrieval, Morgan Kaufmann Publishers, Inc., San Francisco, USA, 1999, pp. 56–57.

[3] K. Chakrabarti, M.O. Binderberger, K. Porkaev, S. Mehrotra, Similar shape retrieval in MARS, in: Proc. IEEE Int. Conf. on Multimedia and Expo (CD-ROM), New York, USA, 2000.

[4] E.R. Davies, Machine Vision: Theory, Algorithms, Practicalities, Academic Press, New York, 1997.

[5] G. Dudek, J.K. Tsotsos, Shape representation and recognition from multiscale curvature, Comput. Vision Image Understanding 68 (2) (November 1997) 170–189.

[6] H. Freeman, A. Saghri, Generalized chain codes for planar curves, in: Proceedings of the 4th International Joint Conference on Pattern Recognition, Kyoto, Japan, 7–10 November, 1978, pp. 701–703.

[7] G. Granlund, Fourier preprocessing for hand print character recognition, IEEE Trans. Computers 21 (1972) 195–201.

[8] M.-K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inform. Theory IT-8 (1962) 179–187.

[9] C.-L. Huang, D.-H. Huang, A content-based image retrieval system, Image Vision Comput. 16 (1998) 149–163.

[10] S. Jeannin (Ed.), MPEG-7 Visual part of experimentation Model Version 5.0. ISO/IEC JTC1/SC29/WG11/N3321, Nordwijkerhout, March 2000.

[11] H. Kauppinen, T. Seppanen, M. Pietikainen, An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification, IEEE Trans. PAMI-17 (2) (1995) 201–207.

[12] H. Kim, J. Kim, Region-based shape descriptor invariant to rotation, scale and translation, Signal Processing: Image Communication 16 (2000) 87–93.

[13] S.X. Liao, M. Pawlak, On image analysis by moments, IEEE Trans. Pattern Anal. Mach. Intell. 18 (3) (1996) 254–266.

[14] G.J. Lu, A. Sajjanhar, Region-based shape representation and similarity measure suitable for content-based image retrieval, Multimedia System 7 (1999) 165–174.

[15] B.M. Mehtre, M.S. Kankanhalli, W.F. Lee, Shape measures for content based image retrieval: a comparison, Inform. Process. Manage. 33 (3) (1997) 319–337.

[16] F. Mokhtarian, S. Abbasi, J. Kittler, Robust and efficient shape indexing through curvature scale space, in: Proc. British Machine Vision Conference, Edinburgh, UK, 1996, pp. 53–62.

[17] F. Mokhtarian, A. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, IEEE PAMI-8 (1) (1986) 34–43.

[18] W. Niblack, et al., The QBIC project: querying images by content using color, texture and shape, in: SPIE Conference on Storage and Retrieval for Image and Video Databases, Vol. 1908, San Jose, CA, 1993, pp. 173–187.

[19] P.J. van Otterloo, A Contour-Oriented Approach to Shape Analysis, Prentice Hall International Ltd., UK, C1991, pp. 90–108.

[20] E. Paquet, M. Rioux, A. Murching, T. Naveen, A. Tabatabai, Description of shape information for 2-D and 3-D objects, Signal Processing: Image Communication 16 (2000) 103–122.

[21] E. Persoon, K.-s. Fu, Shape discrimination using Fourier descriptors, IEEE Trans. Syst. Man Cybern. SMC-7 (3) (1977) 170–179.

[22] T.W. Rauber, Two-dimensional shape description, Technical Report: GR UNINOVA-RT-10-94, University Nova de Lisboa, Portugal, 1994.

[23] M. Safar, C. Shahabi, X. Sun, Image retrieval by shape: a comparative study, in: IEEE International Conference on Multimedia and Expo, New York, USA, 2000 (CD-ROM Proc.).

[24] G. Salton, The state of retrieval system evaluation, Inform. Process. Manage. 28 (4) (1992) 441–450.

[25] J. Tague-Sutcliffe, The pragmatics of information retrieval experimentation, revisited, in: K.S. Jones, P. Willett (Eds.), Readings in information retrieval, Multimedia Information and Systems, Morgan Kaufmann Publisher Inc., San Francisco, USA, 1997, Chapter 4, pp. 205–216.

[26] G. Taubin, D.B. Cooper, Recognition and positioning of rigid objects using algebraic moment invariants, in: SPIE Conference on Geometric Methods in Computer Vision, Vol. 1570, 1991, pp. 175–186.

[27] M.R. Teague, Image analysis via the general theory of moments, J. Opt. Soc. Am. 70 (8) (1980) 920–930.

[28] C.-H. Teh, R.T. Chin, On image analysis by the methods of moments, IEEE Trans. Pattern Anal. Mach. Intell. 10 (4) (1988) 496–513.

[29] Q.M. Tieng, W.W. Boles, Recognition of 2D object contours using the wavelet transform zero-crossing representation, IEEE Trans. PAMI 19 (8) (August 1997) 910–916.

[30] H.S. Yang, S.U. Lee, K.M. Lee, Recognition of 2D object contours using starting-point-independent wavelet coefficient matching, J. Visual Comm. Image Representation 9 (2) (1998) 171–181.

[31] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, IEEE Trans. Comput. c-21 (3) (1972) 269–281.

[32] D.S. Zhang, G.J. Lu, A comparative study on shape retrieval using Fourier descriptors with different shape signatures, in: Proceedings of the International Conference on Multimedia and Distance Education, Fargo, ND, USA, June 2001, pp. 1–9.

[33] D.S. Zhang, G. Lu, Content-based shape retrieval using different shape descriptors: a comparative study. in: Proceedings of IEEE International Conference on Multimedia and Expo (ICME2001), 22–25 August 2001, Tokyo, Japan, pp. 317–320.