

Corners-Based Composite Descriptor for Shapes

Atul Sajjanhar^{1†}, Guojun Lu², Dengsheng Zhang², Wanlei Zhou¹

¹*School of Engineering and Information Technology
Deakin University
221 Burwood Highway
Burwood, VIC 3125
Australia
{atuls, wanlei}@deakin.edu.au*

²*Gippsland School of Information Technology
Monash University
Northways Road
Churchill, VIC 3842
Australia
{guojun.lu, dengsheng.zhang}@infotech.monash.edu.au*

Abstract

In this paper, a composite descriptor for shape retrieval is proposed. The composite descriptor is obtained based upon corner-points and shape region. In an earlier paper, we proposed a composite descriptor based on shape region and shape contour, however, the descriptor was not effective for all perspective and geometric transformations. Hence, we modify the composite descriptor by replacing contour features with corner-points features. The proposed descriptor is obtained from Generic Fourier Descriptors (GFD) of the shape region and the GFD of the corner-points. We study the performance of the proposed composite descriptor. The proposed method is evaluated using Item S8 within the MPEG-7 Still Images Content Set. Experimental results show that the proposed descriptor is effective.

1. Introduction

Approaches for shape representation and retrieval can be broadly classified into contour based and region based. Some of the region based methods are grid based method, geometric moments and moments constructed

orthogonal functions [1]. Recently, Generic Fourier Descriptors (GFD) was proposed by Zhang and Lu [2] for region based matching of shapes. Some of the contour based methods are polygonal approximation, autoregressive model, Fourier Descriptors and distance histograms [1].

In this paper, a composite descriptor based on a region based method and a corner-points based method is proposed for 2d shape retrieval. Corners are obtained using the Harris-Plessey operator. GFD is used for representation of the corners obtained a priori. Region-based GFD is also computed for the 2d shape [2]. A composite descriptor is constructed from the corners descriptor and the region descriptor.

It has been shown that the performance of GFD is comparable with other contemporary techniques [2]. Hence, we evaluate the proposed descriptor by comparing it with GFD. GFD is described in Section 2. The proposed method is described in Section 3. Experimental Setup and Results are presented in Section 4. Finally, Discussion and Conclusion are presented in Sections 5 and 6 respectively.

2. Generic Fourier Descriptors

Generic Fourier Descriptors (GFD) is a region-based method for image retrieval [1][2]. In GFD, feature vectors are created by extracting spectral information in the frequency domain. Fourier transform is applied

[†] PhD candidate, GSIT, Monash University, Australia

to the polar raster sampled shape image. Consider the image shown in Figure 1. To obtain the GFD for the image, the image is first plotted in polar space. The polar image of Figure 1(a), is shown in Figure 1(b).

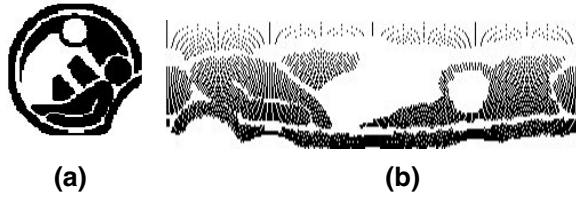


Figure 1. (a) An Image in Cartesian Coordinates (b) Polar Image

Before obtaining the polar image, the image is normalized for scale. 2d DFT is applied to the rectangular region in polar coordinates to obtain Fourier coefficients which are used to construct feature vectors for shape representation and similarity measure [1][2].

Polar coordinates (r, θ) are obtained from the 3d Cartesian co-ordinates (x, y) as shown below.

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \quad (1)$$

$$\theta = \arctan\left(\frac{y - y_c}{x - x_c}\right) \quad (2)$$

where, (x_c, y_c) is the centroid of the 2d Cartesian image.

Feature vectors are constructed from the polar coordinates by computing the 2d DFT. 2d DFT of the polar coordinates is defined as below.

$$PF(\rho, \tau) = \sum_r \sum_\theta f(r, \theta) e^{-j2\pi\left(\frac{r}{R}\rho + \frac{\theta}{T}\tau\right)} \quad (3)$$

where, R and T is the radial and angular resolution. r, θ is obtained from Eqn. 1 and Eqn. 2. Feature vectors are represented as shown below.

$$F: \left(\frac{F(0,1)}{F(0,0)}, \dots, \frac{F(0,T-1)}{F(0,0)}, \dots, \frac{F(1,0)}{F(0,0)}, \dots, \frac{F(R-1,T-1)}{F(0,0)} \right)$$

where, R and T is the radial and angular resolution as used in Eqn. 3.

The difference between two images is computed as the Euclidean distance between their feature vectors as shown in Eqn. 4.

$$Dist(F_1, F_2) = \sqrt{\sum_{i=0}^{RT-1} (f_{1,i} - f_{2,i})^2} \quad (4)$$

where, $f_{x,i}$ is a descriptor within the feature vector of image x . $0 < i < RT$, where R, T is the radial and angular resolution.

3. Proposed Method

In this section, the drawback of the method proposed in an earlier paper is highlighted. A shape representation method based on corner points is proposed.

In an earlier paper, we proposed a method to obtain the shape contour from connectivity information [4][5]. Connectivity method is based on pixel neighbourhood information. Advantage of capturing the contour using connectivity is that the contour continuity is non-essential. Hence, contour propagation techniques are not required. The contour for the shape in Figure 1(a) is shown in Figure 2.

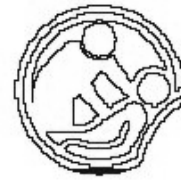


Figure 2. Contour Information obtained from Connectivity

We showed that a composite descriptor obtained from the GFD of the shape region and the GFD of the shape contour performed well [12]. The drawback of the method is that it does not perform well when the images are scaled down. The drop in performance for scaled down images is attributed to the fact that there is loss of contour information when the contour extraction method is applied to the scaled down images.

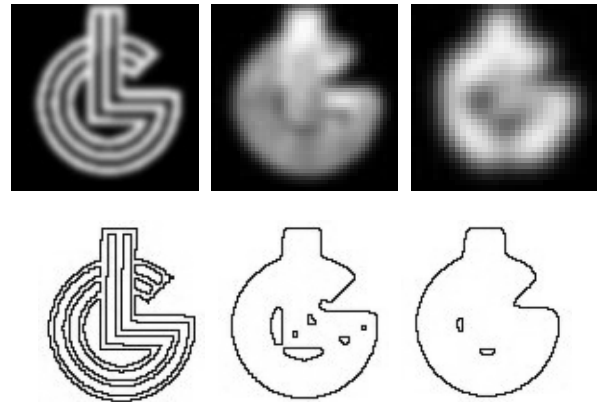


Figure 3. Contour Extraction after Smoothing the Scaled-down Images

In this paper, the composite descriptor is built based on the GFD of the shape region and the GFD of the corner-points. Harris-Plessey corner detector is used to extract corner-points. Harris-Plessey corner

detector has been shown to be effective in detecting corners. It addresses the limitations of the Moravec operator which has an anisotropic response as the intensity variation is only computed at discrete set of shifts. The result of applying the corner-points detector to scaled-down, Gaussian smoothed images is shown below.

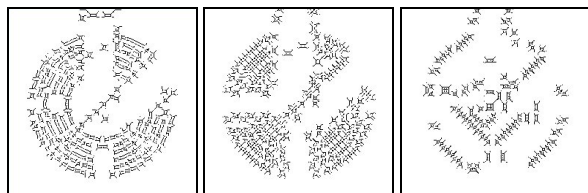


Figure 4. Corner-Points Extraction after Smoothing the Scaled-down Images

During retrieval, images are ranked based on the two features. Ranking of images based on multiple features is a two step process. First, distances are computed for each feature thus giving a ranked list for each feature. Second, distances are combined to obtain an overall measure of distance, namely, global distance. When combining distances for different features, some features may contribute disproportionately to the distance measure because the scale of the distance will be different for different features. Celentano et al. [6] have proposed two methods to mitigate the disproportionate contribution by different features. The methods are *norm1* and *norm2*.

norm1: For each element in a ranked list of k elements having distance d_i from the query image, the normalized distance is:

$$d_i' = \frac{d_i}{d_k} \quad (5)$$

where d_k is the distance for the lowest ranked element.

norm2: The normalised distance is:

$$d_i' = \frac{d_i - d_1}{d_1 - d_k} \quad (6)$$

where d_1 and d_k are the distances for the highest and the lowest ranked elements.

After normalising the distances the global distance for an image which appears in n ranked lists may be computed from *mean1* or *mean2* which are computed as shown in Eqn. 7 and Eqn. 8.

$$d = \frac{d_1 + d_2 \dots d_n}{n} \quad (7)$$

$$d = \frac{d_1 + d_2 \dots d_n}{n^2} \quad (8)$$

Experimental results provided by Celentano et al. suggest that *norm1* and *mean1* perform better than *norm2* and *mean2*. A method similar to *norm1* has also been used by Jain et al [8].

Another method for integration of disparate feature vectors is used in the QBIC system [9][10][11]. The distance between two images is computed using a weighted Euclidean distance with the inverse of feature variances used for normalization. The distance between two images is computed as,

$$d_{ij} = \sum_k \frac{(k_i - k_j)^2}{\sigma^2} \quad (9)$$

where k is a feature and i and j are the two images.

4. Experimental Results

Two sets of experiments are conducted. In the first set of experiments, effectiveness of two methods for the integration of disparate feature vectors (namely, *Corner-points GFD* and *Region GFD*) are compared. In the second set of experiments, the effectiveness of the composite descriptor for GFD is compared with traditional GFD.

Experiments are performed on Item number S8 within the MPEG-7 Still Images Content Set; this is a collection of trademark images which was originally provided by the Korean Industrial Property Office. S8 consists of 3621 still images. It is divided into sets A1, A2, A3, A4 to test the robustness of methods to geometric and perspective transformations.

Set A1 consists of 2881 shapes from the whole database; it is for test of scale invariance. 100 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.

The effectiveness of two methods to integrate Corner-points GFD and Region GFD are compared. The result for recall and precision for *norm1* and *QBIC method (variance based)*, as described in Section 3, are shown below. The results are averaged over 220 queries from the dataset described above. The queries are divided into 20 classes of shapes with each class containing 11 member shapes which are generated through geometric and perspective transformations. The results are shown in Table 1 where Method A refers to *norm1* method and Method B refers to *variance based* method. Weight refers to the weight assigned to the Corner-points GFD when computing the distance between two images. *norm1* method clearly outperforms the other method.

Table 1. Retrieval Performance of Two Methods to Integrate Feature Vectors

Method, Weight	Recall (%)				
	10	20	30	40	50
A, 0.1	100	100	100	100	100
B, 0.1	100	100	100	99.2	98.6
A, 0.4	100	100	100	100	99.5
B, 0.4	100	100	99.1	98.4	94.3
A, 0.8	100	100	96.7	95.2	94.4
B, 0.8	100	99.0	96.3	87.3	81.3

Method, Weight	Recall (%)				
	60	70	80	90	100
A, 0.1	100	97.9	94.1	86.7	55.0
B, 0.1	95.2	90.7	84.2	68.5	58.3
A, 0.4	98.6	97.6	95.9	73.7	52.4
B, 0.4	88.7	86.2	74.1	53.7	24.8
A, 0.8	93.7	92.2	88.9	71.3	47.2
B, 0.8	72.6	63.9	46.3	36.1	14.0

In the proposed method, normalised distance between two images is computed for *Region GFD* and *Corner-points GFD* using the *norm1* method described in Section 3. The total distance between two images is computed as the weighted sum of *normalised* distances. Queries used are those supplied by MPEG-7. 60 features (reflecting 5 radial frequencies and 12 angular frequencies) each are selected for the Corner-points GFD and the Region GFD. Queries are performed using GFD; these are represented by

'*traditional*' within the legends. Another set of queries are performed using the proposed method; these are represented by '*composite*' within the legends. Average recall-precision plots for queries in Sets A1, A2, A3 and A4 are shown in Figure 5. Precision along the vertical axis is plotted for recall along the horizontal axis. The weight assigned to *Corner-points GFD* is 0.1.

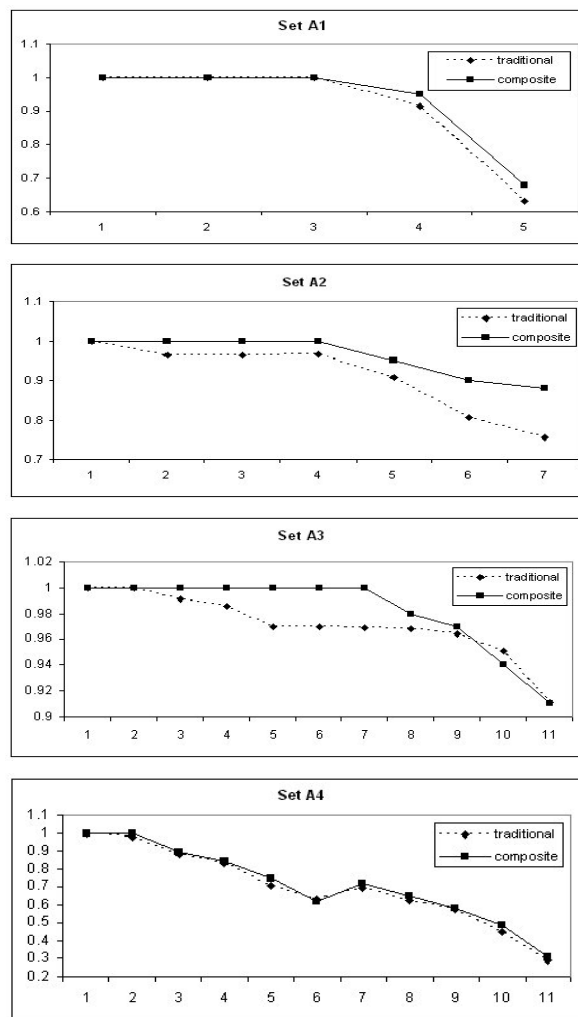


Figure 5. Recall-Precision Plots for Queries in Set A1, A2, A3, A4

From the results, we observe that there is no significant difference in performance for Set A1. The proposed method shows significant improvement for Sets A2 and A3. The proposed method shows a small improvement for Set A4. Different datasets respond differently to the weight assigned to *Corner-points GFD*. The effectiveness will drop when the weight of *Corner-points GFD* is increased beyond a threshold. The optimal value of the weight assigned to *Corner-*

points GFD needs to be determined empirically and it will depend on the dataset.

5. Discussion

There are two factors which contribute to the relative improvement of the proposed method when compared with traditional GFD. First, additional information is captured by the corners; descriptors which encode corners are able to discriminate better between shapes.

Improvement of the proposed method is also attributed to the use of spectral analysis. Spectral analysis of images has been widely used for image retrieval. There are two advantages of spectral features. First, they are robust compared with spatial features. Second, spectral features are inherently multi-resolutional and this property can be leveraged to determine the degree of detail encoded during indexing.

6. Conclusion

In this paper, we have proposed a new index which complements GFD by capturing corners information. Contour information is captured based on connectivity. Experiments have been performed on the MPEG-7 Still Images Content Set. Experimental results prove that the proposed method is promising. The proposed method may be modified to incorporate other shape retrieval techniques such as geometric moments, however, this will need further investigation.

7. References

- [1] D. Zhang, Image Retrieval Based on Shape, *PhD Thesis*, Monash University, Australia, 2002.
- [2] D. Zhang and G. Lu, Generic Fourier Descriptors for Shape-based Image Retrieval, *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland, August 26-29, 2002.
- [3] G. Pass and R. Zabih, Histogram refinement for content-based image retrieval, *IEEE Workshop on Applications of Computer Vision*, pp. 96-102, December 1996.
- [4] A. Sajjanhar, G. Lu and D. Zhang, Discriminating Shape Descriptors Based on Connectivity, *IEEE International Conference on Multimedia and Expo*, Taipei, Taiwan, 2004.
- [5] A. Sajjanhar, G. Lu and D. Zhang, Spherical Harmonics Descriptors for 2D Image Retrieval, *IEEE International Conference on Multimedia and Expo*, Amsterdam, The Netherlands, 2005.
- [6] A. Celentano and S. Sabbadin, Multiple Feature Indexing in Image Retrieval Systems, *First European Workshop on Content-Based Multimedia Indexing*, Toulouse, France, October 25-27, 1999.
- [7] R. Gonzalez and R. Woods, *Digital Image Processing*, Prentice Hall, 2002.
- [8] A. K. Jain; A. Vailaya, Shape-based retrieval: a case study with trademark image databases, *Pattern Recognition*, Volume 31, Number 9, September 1998, pp. 1369-1390
- [9] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, Efficient and effective querying by image content, *Journal of Intelligent Information Systems*, 3(3/4):231-262, July 1994.
- [10] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic, Automatic and semi-automatic methods for image annotation and retrieval in QBIC, *Storage and Retrieval for Image and Video Databases III*, Volume 2420, pages 24-35, February 1995.
- [11] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, The QBIC project: querying images by content using color, texture and shape, *Storage and Retrieval for Image and Video Databases*, Volume 1908, pages 173-187, 1993.
- [12] A. Sajjanhar, G. Lu, D. Zhang, W. Zhou, A Composite Descriptor for Shape Retrieval, *IEEE International Conference on Computer and Information Science*, 11-13 July 2007, Melbourne, Australia.