

Discriminating Shape Descriptors Based on Connectivity

Atul Sajjanhar
School of Info. Tech.
Deakin University
221 Burwood Highway
Burwood, VIC 3125
Australia
atuls@deakin.edu.au

Guojun Lu Dengsheng Zhang
Gippsland School of Computing & Info. Tech.
Monash University
Northways Road
Churchill, VIC 3842
Australia
{guojun.lu, dengsheng.zhang}@infotech.monash.edu.au

Abstract

In this paper, we have proposed a method for enhancing the accuracy of shape descriptors. The concept of connectivity, to obtain discriminating shape descriptors, has been introduced. We have shown how connectivity can be applied to two popular shape descriptors. Experiments were performed to test the effect of using connectivity with Generic Fourier Descriptors and distance histograms. Item S8 within the MPEG-7 Still Images Content Set [5] was used for performing experiments. This dataset consists of 3621 still images. The experimental results show that connectivity enhances the performance of the methods significantly.

Keywords: Histograms, Fourier descriptors, shape descriptors, CBIR.

1. Introduction

Much research is being done to develop tools for analyzing images based on their content and then representing them in a manner that the images can then be searched based on these representations. Content based image retrieval (CBIR) allows users to retrieve images using queries based on sketches, user constructed query images, color and texture patterns, layout or structural descriptions, and other example images or iconic and graphical information. Retrieval of images based on the shape of objects in images is an important part of CBIR.

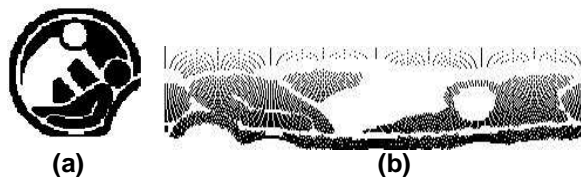
Approaches for shape representation and retrieval can be broadly classified into contour based and region based. Some of the region based methods are geometric moments [3], moments constructed from orthogonal functions [11] and grid based method [4]. Recently, Generic Fourier Descriptors (GFD) method was proposed

by Zhang and Lu [9] for region based matching of shapes. Some of the contour based methods are polygonal approximation [10], autoregressive model [6], Fourier Descriptors [7], distance histograms [1]. In this paper, we propose the concept of *connectivity* to obtain discriminating shape descriptors. We have shown how to apply *connectivity* to GFD method and histograms method. Experiments have been performed to show the effectiveness of using *connectivity*.

In Section 2, GFD and the proposed enhancement to GFD has been described. In Section 3, histograms method and the proposed enhancement to the histograms method has been described. Experimental Setup and Results are presented in Section 4. We provide the conclusion in Section 5.

2. Generic Fourier Descriptors

Generic Fourier Descriptors (GFD) have been used for image retrieval based on region-based shape matching [9]. In GFD, the feature vectors are created by extracting spectral information in the frequency domain. Fourier transform is applied 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 [7][9].

We draw an analogy from Color Coherence Vectors (CCV) proposed by Pass and Zabih [8]. CCV is used for image retrieval based on color. Pass et al [8] defined color coherence of pixels as the degree to which pixels of that color are members of a large similarly colored region. Pixels are classified as coherent or incoherent. Coherent pixels are part of a sizable contiguous region of similar color while incoherent pixels are not.

In the case of shape representation, we define “connectivity” of pixels in an image. The state of the nearest 8-neighbours is computed for each *set* pixel. Connectivity of a *set* pixel is obtained as the number of *set* pixels amongst the nearest 8-neighbours. Figure 2(a) provides additional information for the image in Figure 1(a). Connectivity information is added in Cartesian coordinates. Hence, the z-axis is obtained which provides information regarding connectivity of pixels. For each *set* pixel within the image, the connectivity can take values 0 through 8. A connectivity of 0 indicates that none of the nearest 8-neighbours are *set*. A connectivity of 8 indicates that all of the nearest 8-neighbours are *set*. Figure 2(b), represents the image in cylindrical coordinates. Cylindrical coordinates (r, θ, ϕ) are obtained from the 3D Cartesian co-ordinates (x, y, z) as shown below.

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

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

$$\phi = z \quad (3)$$

where, (x_c, y_c) is the centroid of the 2D Cartesian image and z represents the connectivity of pixel (x, y) .

Feature vectors are constructed from the cylindrical coordinates by computing the 2D DFT for each value of ϕ in Eqn. 3. 2D DFT of the polar coordinates for each value of ϕ is defined as below.

$$PF_\phi(\rho, \tau) = \sum_r \sum_\theta f(r, \theta) e^{-j2\pi\left(\frac{r}{R}\rho + \frac{\theta}{T}\tau\right)} \quad (4)$$

where, R and T is the radial and angular resolution. r, θ is obtained from Eqn. 1 and Eqn. 2.

A set of nine feature vectors is obtained for each image. Feature vectors are represented as shown below.

$$F: \left(\frac{PF_\phi(0,1) \quad PF_\phi(0,2) \quad PF_\phi(0,T-1) \quad PF_\phi(1,0) \quad PF_\phi(R-1,T-1)}{PF_\phi(0,0) \quad PF_\phi(0,0) \quad \dots \quad PF_\phi(0,0) \quad \dots \quad PF_\phi(0,0)} \right)$$

where, ϕ is the connectivity and has values 0 to 8. R and T is the radial and angular resolution as used in Eqn. 4.

The difference between two images is computed as the sum of the Euclidean distances between feature vectors as shown in Eqn. 5.

$$Dist(F_1, F_2) = \sqrt{\sum_{i=0}^8 \sum_{j=0}^{RT-1} (f_{1,i,j} - f_{2,i,j})^2} \quad (5)$$

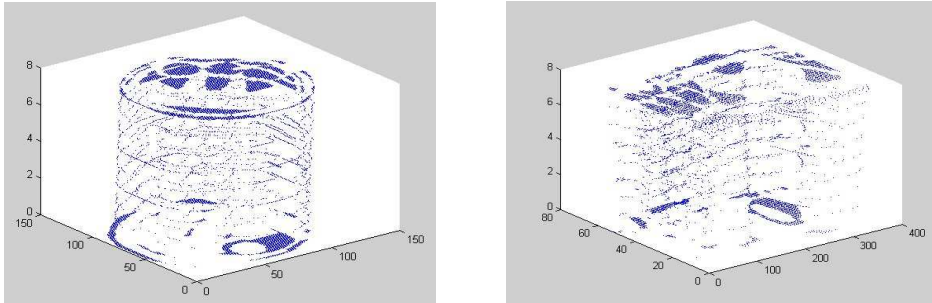
where, $f_{x,i,j}$ is a descriptor within the feature vector of image x. $0 \leq i \leq 8$ is the connectivity. $0 < j < RT$, where R, T is the radial and angular resolution.

3. Distance Histograms

The distance histograms method treats an image as a point cloud. The centroid is computed for the image as shown in Eqn. 6.

$$x_c = \frac{\sum_{i=0}^{N-1} x_i}{N} \quad y_c = \frac{\sum_{i=0}^{N-1} y_i}{N} \quad (6)$$

where, N is the number of points in the image.



**Figure 2. (a) Connectivity Information for Image in Figure 1
(b) Connectivity Information in Cylindrical Coordinates**

After the centroid has been identified, the centroidal distance is computed for each point as shown in Eqn. 7.

$$d(s_i, c) = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} \quad (7)$$

where $s_i(x_i, y_i)$ is a point on the image for $i=0$ to $N-1$ and $c(x_c, y_c)$ is the centroid.

The distances thus obtained are invariant to rotation and translation. However, the distance needs to be normalised for scale. To normalise the distance for scale, we scale them from 0 to 100.

Normalised centroidal distances of the points are discretised into buckets. A histogram for an image is represented as below.

$$D: (d_0, d_1 \dots d_{N-1}) \quad (8)$$

where N is the number of buckets in the histogram.

Quadratic distance is used to compute the distance between histograms because it takes into account the similarity across buckets [12][13]. Euclidean distance on the other hand only considers the distance between “like-buckets”. The quadratic-form distance between the two feature vectors Q and T is given by Eqn. 9.

$$d_{quad}(Q, T) = \left[(Q - T)^t A (Q - T) \right]^{1/2} \quad (9)$$

where $A = [a_{ij}]$ is an $N \times N$ matrix, and a_{ij} is the similarity coefficient between indexes (dimensions) i and j . a_{ij} is given by Eqn. 10.

$$a_{ij} = 1 - d_{ij} / d_{\max} \text{ and } d_{ij} = |Q_i - T_j| \quad (10)$$

For calculation, the quadratic-form distance is written as shown below [14].

$$d_{quad}(Q, T) = \left(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{ij} Q_i Q_j + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{ij} T_i T_j + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} Q_i T_j \right)$$

The concept of connectivity was introduced in Section 2. Connectivity can be applied to the histograms method. Images are decomposed according to connectivities as shown in Figure 3(a). Histograms are obtained for each value of connectivity. This gives a set of nine histograms for connectivity 0 through 8. The nine histograms form the feature vector for each image as shown in Eqn. 11.

$$D: (d_{0,\phi}, d_{1,\phi} \dots d_{N-1,\phi}) \quad (11)$$

where, ϕ is the connectivity and has values 0 to 8, N is the number of buckets into which the centroidal distances have been discretised.

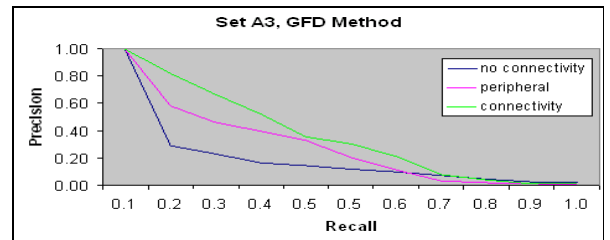
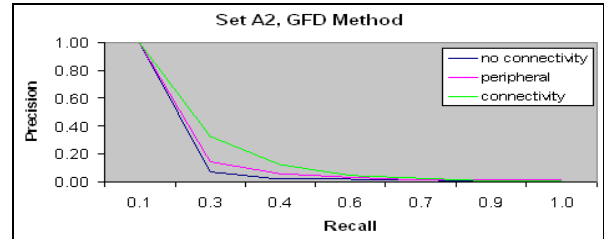
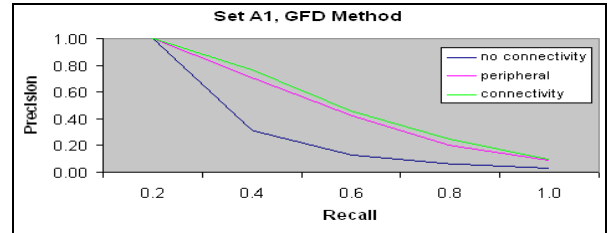
The distance between two images is computed as the sum of quadratic distances between corresponding histograms.

4. Experimental Results

Experiments were conducted on Item number S8 within the MPEG-7 Still Images Content Set [5]. This is a collection of trademark images and 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.

Queries were performed using the basic GFD method and the proposed enhancement to the GFD method. Another set of queries were performed using the histograms method and the proposed enhancement to the histograms method. In Figure 3, average recall-precision has been plotted for each method. The basic method is represented by “no connectivity” within the legends. The proposed enhancement to the basic methods is represented by “connectivity” within the legends.

In Figure 2, we see that the pixel density is high for connectivity=0 and connectivity=8. This observation led us to perform experiments while ignoring intermediate values of connectivity. This is a special case of the proposed method. In this method, images are indexed for connectivity=0 and connectivity=8 only. This method is represented by “peripheral” within the legends.



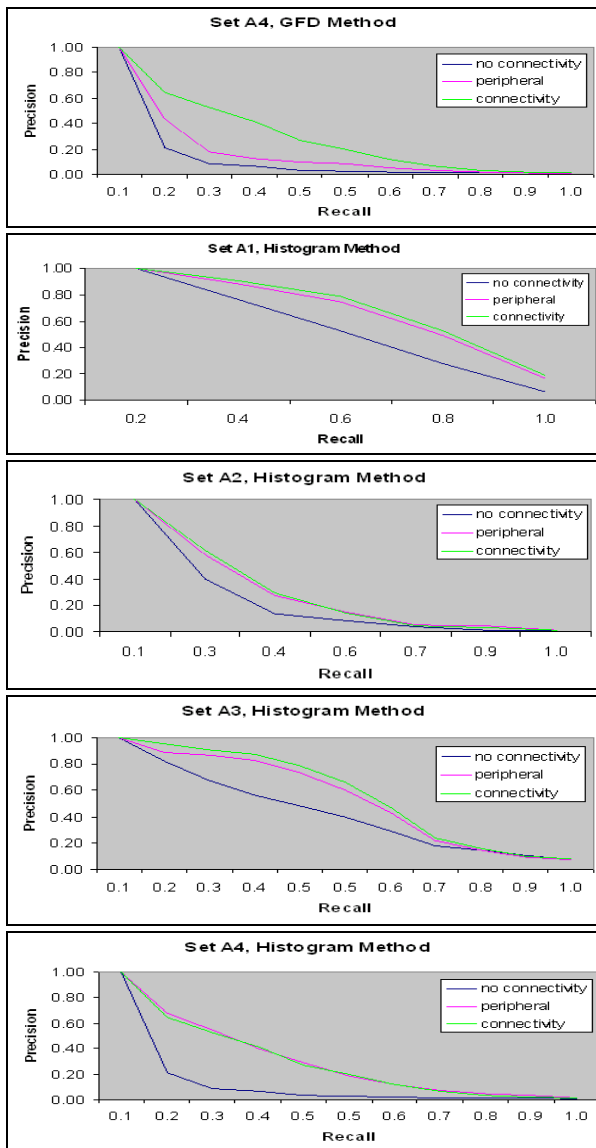


Figure 3. Average Recall-Precision Plots

5. Conclusion

There is a drop in the effectiveness of retrieval, when intermediate values of connectivity are excluded, from indexing and retrieval.

We note that the data set does not contain intricate images. In Figure 2, we see that the pixel density is high for connectivity=0 and connectivity=8. We believe that the relative improvement in the effectiveness of the proposed method will be higher with an increase in pixel densities for intermediate values of connectivity.

In this paper, a method has been proposed to obtain highly discriminating shape descriptors. The proposed method has been tested for two common shape based

image retrieval techniques. Experiments were performed on the MPEG-7 Still Images Content Set. Experimental results show that the proposed method improves the performance of shape based image retrieval techniques.

References

- [1] S. Fan, "Shape Representation and Retrieval using Distance Histograms", Technical Report – University of Alberta, 2001.
- [2] A. Sajjanhar, G. Lu, "Effect of Spatial Coherence on Shape Retrieval", International Conference on Imaging Science, Systems, and Technology, Las Vegas, pp. 118-122, June 2003.
- [3] R. Gonzalez. R. E. Woods, "Digital Image Processing", Prentice Hall Inc., New Jersey, USA, 2002.
- [4] A. Sajjanhar, G. Lu, "A Grid Based Shape Indexing and Retrieval Method", The Australian Computer Journal, Vol. 29, No. 4, pp. 131-140, November 1997.
- [5] <http://ipsi.fraunhofer.de/delite/Projects/MPEG7/>
- [6] R. L. Kashyap, R. Chellappa, "Stochastic models for closed boundary analysis: Representation and reconstruction", IEEE transactions on Information Theory, IT-27(5), pp. 627-637, September 1981.
- [7] H. Kauppinen, T. Seppanen, M. Pietikainen, "An Experimental Comparison of Autoregressive Fourier-Based Descriptors in 2D Shape Classification", IEEE transactions on Pattern Analysis and Machine Intelligence, 17(2), pp. 201-207, February 1995.
- [8] G. Pass, R. Zabih, "Histogram refinement for content-based image retrieval", IEEE Workshop on Applications of Computer Vision, pp. 96-102, December 1996.
- [9] D. S. Zhang, G. Lu, "Generic Fourier Descriptors for Shape-based Image Retrieval", IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland, August 26-29, 2002.
- [10] T. P. Pavlidis, "Polygonal approximation by Newton's method", IEEE transactions on Computers, C-26(8), pp. 800-807, August 1977.
- [11] M. Teague, "Image analysis via the general theory of moments", Journal Optical Society of America, 70(8), pp. 920-930, 1980.
- [12] J. R. Smith, "Integrated Spatial and Feature Image System: Retrieval, Analysis and Compression", PhD thesis, Columbia University, 1997.
- [13] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Image By Content Using Color, Texture and Shape", SPIE Storage and Retrieval for Image and Video Databases, Vol.1908, pp. 173-187, 1993.
- [14] Y. Deng, "A Region Based Representation for Image and Video Retrieval", PhD thesis, University of California, 1999.