IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 7, Issue 5, Oct - Nov, 2019
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

SHAPE BASED IMAGE RETRIEVAL USING STRUCTURAL AND GLOBAL DESCRIPTORS

Shatajbegum Nadaf[1], Dr. J C Karur[2], Dr. Jagadeesh Pujari[3]

[1] Student, S D M College of Engineering and Technology, Dharwad, India,
[2] Professor, Dept of CSE, S D M College of Engineering and Technology, Dharwad, India
[3] Professor & HOD, Dept of CSE, S D M College of Engineering and Technology, Dharwad, India,

Abstract

Content-based Image Retrieval (CBIR) systems often rely on the basic features of an image such as color, texture, shape and other features to search and retrieve similar images from a database. Among these features, shape plays an important role as objects are mainly perceived by their shape. In this paper we have proposed shape based image retrieval using structural and global shape descriptors. The structural shape descriptors are based on the radial distance ratio. A novel edge pixel sampling method is proposed for uniform sampling of edge pixels to be used by a global descriptor based on distance interior ratio. Shape descriptor matching is done using bin-to-bin matching and Most Similar Highest Priority matching algorithms. MPEG-7 shape dataset is used to conduct the experiments. The experimental results show that the proposed method is effective and efficient for image matching and retrieval with a Bull's-eye score of 80.275%. The proposed method is simple and computationally effective which makes it suitable for practical applications with large databases.

I. INTRODUCTION

"A picture is worth a thousand words", which means any complex concept can be easily described and communicated using an image. Due to the increasing number of digital assets such as photographs and videos, an efficient computerized retrieval system is always in demand. Earlier image retrieval methods used to describe an image using keywords. Image retrieval based on image content is more effective than text-based retrieval in most of the applications. Hence there is a need to automatically extract visual features from an image and using these features to retrieve similar images. The Content-based Image Retrieval (CBIR) systems were designed to solve this problem by using the visual features of an image as opposed to its textual description. Basic features such as color, texture, shape and other distinctive features are extracted from the image. These features are used to retrieve similar images from the database.

Shape is considered as the most important intrinsic feature of an object compared to color and texture. Shape feature gives appearance and outline of an image. It is a known fact that objects are perceived mainly by their shape. Thus use of shape features of an object in CBIR techniques is increasing. Shape descriptor refers to a numeric descriptor of the shape of an object, organized using data structures like character array, tree, one dimensional array, or a matrix. These simplified shape descriptors must carry most of the important information, while being easier to work with and to compare with other descriptors than comparing the shapes directly. For a shape descriptor to be considered robust it must be invariant to rotation, scaling and translation.

Shape analysis is a well-explored research area with many shape depiction and matching techniques. Shape extraction techniques are usually grouped into two categories: contour (boundary) based and region based methods. Moment descriptor is a method normally employed in region-based techniques such as Zernike moments [1], Legendre moments [2], Tchebichef moments [3] etc. Other works include generic Fourier descriptor [4], shape matrix, multi-scale Fourier-based description [5], grid technique etc. Despite the fact that region based descriptors are robust in handling generic shapes, they often involve time consuming and intensive computations.

In many applications object's internal content is not as relevant as its boundary. Boundary based techniques mainly rely on the contour of the object. Computations using silhouette of an object makes the shape description efficient compared to region based methods. There has been a lot of work on contour based shape representation and matching. These methods include curvature scale space (CSS) [6], triangle area representation (TAR) [7], shape context (SC) [8], inner-distance shape context (IDSC) [9], contour points distribution histogram (CPDH) [10] and so on.

These techniques can be further classified into structural and global methods based on whether the object's shape is considered as whole or as a set of segments. In this paper we have concentrated on contour based method to define structural and global shape descriptors. Two types of shape descriptor matching methods are defined. A novel method has been proposed for edge pixel sampling. The proposed sampling technique can be combined with any contour based shape description method.

The paper is organized as follows: Section II provides the methodology used in proposed approach. Section III gives the experimentation results and Section IV states the conclusion and future scope.

II. PROPOSED WORK

The proposed work consists of structural and global methods for shape feature extraction along with shape descriptor matching methods based on bin-to-bin matching and MSHP matching. The flow diagram of the proposed work is presented in Fig.1.

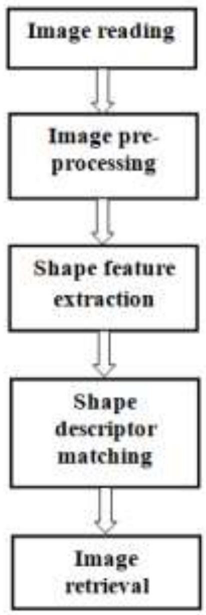

Figure 1. Flow Diagram

A. *Image Pre-processing*

Image datasets may contain both binary and grey images. Hence images read from the dataset are binarized and edge pixels are extracted before extracting features. Zero crossing technique is used to extract the edge pixels which are based on the value of 8 neighbouring pixels. A 3x3 window is used to trace the image for the 8-neighborhood connectivity. If the sum of pixels in the window is less than 9, it is considered as an edge pixel. This method is used to extract both outer and inner boundary pixels of the image.

Most of the images in the dataset contain areas which do not contain any useful information about the object under consideration. These extra pixels are removed by a minimum bounding box around the object. Lines will be drawn to connect the opposite sides of the bounding box at their mid points. The point of intersection of these lines is taken as the centroid of the image. This centroid will be different from the conventional image centroid. A circle is drawn using Bresenham's circle drawing algorithm with the distance from the centroid to farthest edge pixel as radius [11].

B. *Feature extraction*

Feature extraction refers to using a shape description method to convert the given shape into a numerical descriptor. Two types of shape descriptors are proposed: Structural and Global.

1) Structural Shape Descriptors:

The structural shape descriptor proposed is a contour based method. In this method the circle drawn around the image in the pre-processing stage is used as a guiding circle for feature computation. The circle is divided into 36 bins. The number of points on the circle lying in each bin is computed as the ratio of number of pixels on the circle to the total number or bins.

For each point on the circle in a bin, a radius r is drawn. These radii drawn intersect with the edge pixels. Fig.2 shows two such radii intersecting the edge pixels. Radius r1 intersects the image at e1, e2 and e3 whereas r2 intersects at e4.

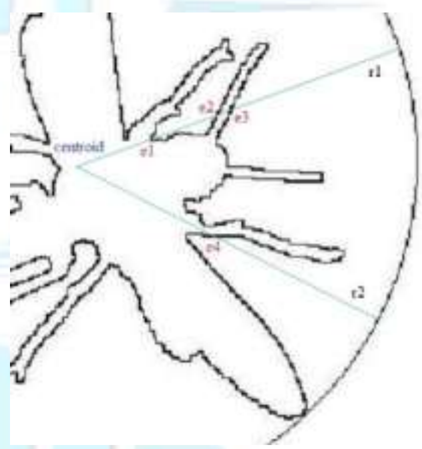

Figure 2. Intersection of radius with edge pixels

a) Radial distance ratio based shape descriptor

In this method, Euclidean distance from the centroid to the edge pixel on the radius drawn is calculated using the formula $\text{Dist} = \sqrt{(cx - ex)^2 + (cy - ey)^2}$ where cx, ex and cy, ey are the x and y coordinates of the centroid C and edge pixel E respectively. The ratio of Euclidean distance of edge pixels to radius is calculated and is called Distance Ratio (DR). Distance Ratios are calculated for every radius in each bin.

The feature set is represented as a 2D histogram. Each bin is divided into 64 blocks, where each block represents a distance-ratio interval. The distance-ratio interval is calculated by dividing the difference between the maximum and minimum distance ratios by 64. The number of bins and intervals is decided empirically. All distance ratios computed in each bin are mapped to respective block in the feature space. The entire feature set is then normalized to avoid the scaling effect.

The histogram of an individual bin is shown in Fig.3.

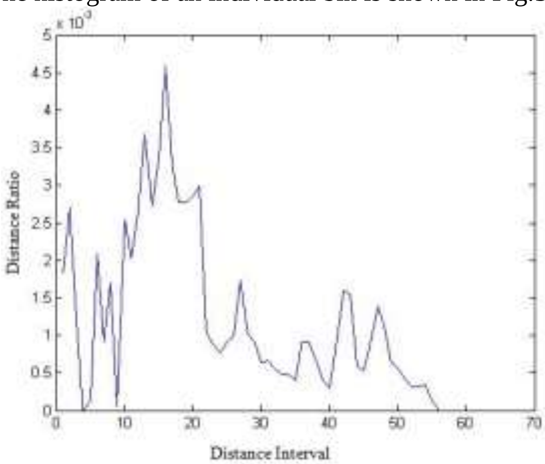

Figure 3. Histogram of individual bin

b) Shape description using Line Fitting

In this method a line will be fitted on the cumulative histogram of each bin, obtained using the DR method. Three features are computed for each bin as follows:

i. Let H be the histogram of bin b_i.
ii. Obtain the cumulative histogram (CH) of H as shown in fig.4

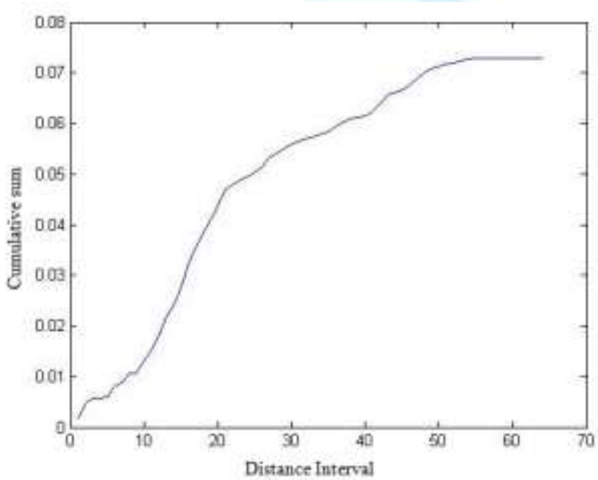

Figure 4. Cumulative histogram (CH) of a single bin

iii. The points on the CH are considered for fitting a regression line as shown in fig.5.

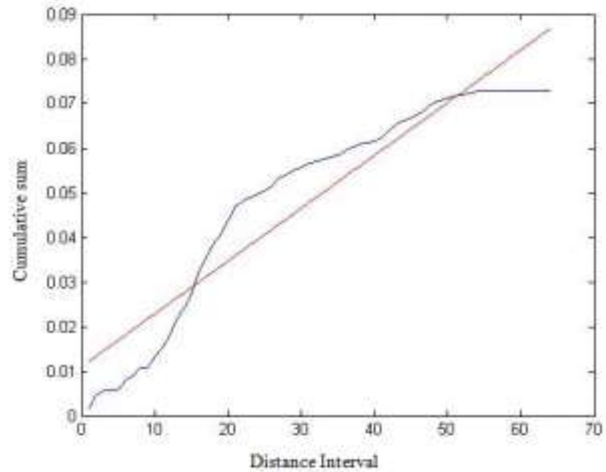

Figure 5. Line fitted across CH

iv. From the regression line fitted, following three features are computed:

Slope, s = slope of the regression line fitted across CH

Mean, μ = mean of all the points on CH using (1).

$$\mu = \frac{\sum_{i=1}^{64} CH_i}{64} \qquad (1)$$

Mean Deviation is calculated using (2),

$$D_{CH} = \frac{\sum_{i=1}^{64} |CH_i - \mu_{CH}|}{64} \qquad (2)$$

Slope, mean and mean deviation are used to represent feature space of the object. The size of the shape descriptor is reduced drastically using this method.

2) Global Shape Descriptor:

A global shape descriptor considers entire image as a single entity for feature extraction. Shape of an object is usually described by its edge pixels. But the region within the edge pixels of an image can also provide significant information about the object. Hence a combination of region and contour based shape description method is used to define a global shape descriptor which is computed in two steps.

a) Edge pixel sampling

The proposed algorithm is based on using the contour points to obtain the shape descriptor. But the number of edge pixels varies for each image ranging from few hundreds to few thousands. Randomly selecting a fixed number of edge pixels may compromise the shape of the object. Hence a novel approach is proposed to restrict the number of edge pixels selected for feature computation while maintaining its shape. The procedure is as follows:

i. Bresenham's circle drawing algorithm returns two set of points, one for each semi circle.

ii. *'n'* sample points on each semi circle are selected which are uniformly spread across the perimeter.
iii. Lines are drawn from every sample point on one semi circle to all the points on the other semi circle.
iv. Lines drawn may intersect with the edge pixels of the image as shown in Fig.6. The line *l* intersects the edges at points *a*, *b*, *c* and *d*. Distances between all pair of points are calculated and the points that give maximum distance are considered. In this case points *a* and *d* are considered. For each such line a pair of farthest edge pixels on the line will be obtained. This method gives only the outermost edge pixels.

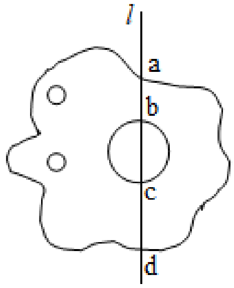

Figure 6. Edge pixel selection method

v. Each line returns two edge pixels. Hence $2n^2$ pixels are obtained through this sampling method which also maintains the original shape.

Fig. 7 shows the result of edge sampling method applied on a boundary image.

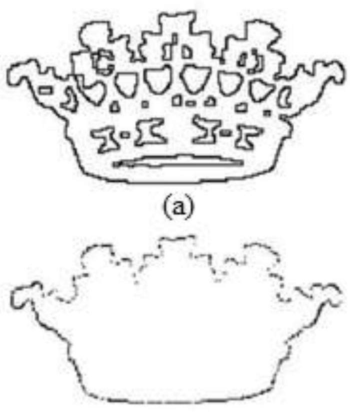

Figure 7. Edge pixel sampling (a) Original edge pixels (b) Sampled edge pixels

b) Descriptor

A histogram based descriptor called Distance Interior Ratio (DIR) is used to describe the shape [12]. When a line is drawn between two edge pixels, the line segment can be one among three types as shown in Fig.8. DIR of a line segment is the ratio of part of line segment lying within the object to the length of the line segment.

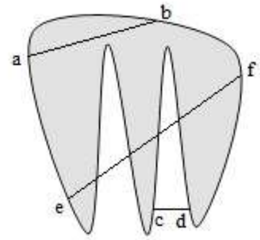

Figure 8. Line segment types (a) line [ab] with DIR 1 (b) line [cd] with DIR 0 (c) line [ef] with DIR between 1 and 0

Let *a*, *b* be two points belonging to a set of edge pixels of an object. Bresenham's line drawing algorithm is used to draw a digital line between two points. The Euclidean distance between a and b can be computed as $d_{ab} = \sqrt{(ax - bx)^2 + (ay - by)^2}$ which represents the total number of pixels on the digital line drawn between the pixels *a* and *b*.

Let *n* be the number of pixels on the line. Since the algorithm considers only binary images, the intensity of the pixels can be either 1 or 0. Let *k* be the count of the number of pixels with value 1 on the line. DIR of the line is calculated as $dir_{ab} = k / n$ where dir_{ab} is the DIR of line *ab*.

A structural histogram is used to represent the feature space. The x-axis corresponds to the Euclidean distance between the points with range d_{min} and d_{max}. Y-axis corresponds to the DIR with range dir_{min} and dir_{max}. The feature space size is 64x32 blocks. The x-axis is divided into 64 blocks with interval $(d_{max} - d_{min}) / 64$ and y-axis is divided into 32 blocks with interval $(dir_{max} - dir_{min}) / 32$. Each DIR is mapped to a block in the feature space. All the entries in the feature space are normalized by dividing them by the total number of points in the feature space to avoid scaling effect.

C. Matching and Retrieval

Two methods have been proposed for shape descriptor matching: Bin-to-Bin (B2B) matching and Most Similar Highest Priority (MSHP) Matching.

1) Bin-to-Bin Matching:

The set of histograms of query image Q is compared with the set of histograms of each image D in the dataset with relative positions of the bins maintained. Difference between query image Q and database image D is calculated using (3).

$$\mathrm{Diff}(Q, D) = \sum_{i=0}^{r-1} \sum_{j=0}^{c-1} |PQ(i, j) - PD(i, j)| \qquad (3)$$

where r=36, c=64 and PQ, PD are the values in the corresponding blocks of query image and image in the dataset.

Bin wise histogram difference is calculated with the bins of each image in dataset rotated, in both clock-wise and anti-clockwise direction, one bin at a time. Rotating the histogram in clockwise direction provides invariance to rotation. The anticlockwise rotation helps to provide invariance to mirroring of the image, which is an added advantage. Among all the bin wise differences, the one with minimum difference is taken as the distance between the images.

2) Most Similar Highest Priority Matching

In this descriptor matching method, one bin from query image is compared with all the bins in the database image. A bipartite graph is built using the bins of query image and database image. The edges of the bipartite graph refer to the distances between each bin, which are saved in an adjacency matrix.

Minimum cost matching is done for the matrix using the following procedure. The minimum value in the adjacency matrix is obtained in the first step. The row and column giving that minimum distance are ignored in further iterations. This process is repeated until every bin finds a match. A particular bin participates in the matching process only once. All these distances are added to get the difference between the images. The complexity of the matching procedure will be linear as only *n* minimum values are obtained where *n* is the number of bins.

Distances between all the images in the dataset and query image are calculated for both the matching methods. All these calculated distances are sorted and top 'n' images are retrieved. Retrieval efficiency is computed by counting the number of correct matches among the images retrieved.

III. Experimentation and Results

MPEG-7 dataset is used to evaluate the proposed shape descriptors. The dataset considered consists of 70 classes with different types of objects and each class contains 20 diverse shapes. This dataset comprises of variety of shapes with a lot of similarity between shapes of different classes which makes it hard to distinguish objects of one class from other. We have considered 70 images, one image from each class to test the retrieval efficiency of the descriptor.

Shape features of all the images in the database are extracted offline. One random image from each class is taken as a query image and the shape descriptor of the query image is compared with all the descriptors in the database using the matching algorithms. Most similar '*n*' matches are retrieved and displayed. The retrieval accuracy is calculated as the ratio of number of correct matches to the number of images retrieved.

Retrieval efficiency is calculated for different number of image retrievals i.e. 5, 10, 15 and 20. The retrieval efficiencies for different number of image retrievals for all the methods are listed in Table.1.

TABLE I.
Retrieval Accuracy for Different No of Images Retrieved

Descriptor + Matching method	**No of images retrieved**			
	5	**10**	**15**	**20**
DR + B2B	96.29	91.57	84.29	76.21
DR + MSHP	**96.57**	**93.14**	**87.81**	**81.43**
Line Fitting + B2B	92.57	82.29	74.48	67.36
Line Fitting + MSHP	89.14	79.43	72.38	65.57
DIR with novel pixel sampling method	85.14	72.71	63.43	55.79

The structural descriptor based on the radial Distance Ratio (DR) combined with MSHP matching algorithm provided better retrieval accuracy compared to the other methods. Bulls-eye test was carried out on this combination which gave a retrieval accuracy of 80.275%.

IV. Conclusion and Future Work

Shape is considered as an important low level feature in Content Based Image Retrieval. In this paper we have presented one global and two structural shape description methods along with shape matching methods using bin-to-bin matching and Most Similar Highest Priority matching. A novel method for edge pixel sampling has been proposed which gives uniformly sampled outer boundary edge pixels of an image. Computations in all the methods are simple and flexible. Since the techniques proposed are based on a guiding circle, the descriptors are invariant to rotation and translation. The distance ratios make the technique invariant to scaling. Experiments were conducted on the MPEG-7 benchmark dataset. The results of the experiments show that the proposed method is effective and efficient for image matching and retrieval with a Bull's-eye score of 80.275%. The simplicity and flexibility of the descriptor makes it feasible to be used in an online environment with large image databases. The future work will be focused on building an even more efficient shape descriptor by combining other visual features of an image along with shape and to experiment with 3D shape matching and retrieval.

V. Bibliography

[1] Yong-Sung Kim Whoi-Yul Kim, "Region-based shape descriptor using Zernike moments," *Signal Processing: Image Communication*, pp. 95-102, 2000.

[2] Jianzhong Zhou, Yuhong Li, Guojun Zhang, Cheng Wang Bo Fu, "Image analysis by modified Legendre moments," *Pattern Recognition*, vol. 40, no. 2, pp. 691-704, February 2007.

[3] S. H. Ong and P. A. Lee R. Mukundan, "Image analysis by Tchebichef moments," *IEEE Transactions on Image Processing*, vol. 10, pp. pp. 1357-1364, Sept 2001.

[4] Guojun Lu Dengsheng Zhang, "Generic Fourier descriptor for shape-based image retrieval," *Proceedings. IEEE International Conference on Multimedia and Expo, Lausanne, Switzerland*, pp. pp. 425-428 vol.1, 2002.

[5] Mark S. Nixon Cem Direkoglu, "Shape classification using multiscale Fourier-based description in 2-D space," *9th International Conference on Signal Processing, Beijing*, pp. 820-823, 2008.

[6] Farzin, Sadegh Abbasi and Josef Kittler Mokhtarian, "Efficient and Robust Retrieval by Shape Content through Curvature Scale Space," *Image Databases and Multi-Media Search*, 1998.

[7] Ibrahim El rube, Mohamed S. Kamel, George Freeman Naif Alajlan, "Shape retrieval using triangle-area representation and dynamic space warping," *Pattern Recognition*, pp. 1911-1920, 2007.

[8] J. Malik and J. Puzicha S. Belongie, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no.4*, pp. pp. 509-522, 2002.

[9] Haibin Ling and David W. Jacobs, "Shape Classification Using the Inner-Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. pp 286–299, 2007.

[10] Xiao-Jun Wu Xin Shu, "A novel contour descriptor for 2D shape matching and its application to image retrieval," *Elsevier*, 2011.

[11] Jack Bresenham, "A Linear Algorithm for Incremental Display of circular Arcs," *Communications of the ACM*, vol. 20, no. 2, pp. 100-106, February 1977.

[12] Jinhee Chun, Takeshi Tokuyama Natsuda Kaothanthong, "Distance Interior Ratio: a new shape signature for 2D shape retrieval," *Pattern Recognition Letters*, 2016.