Shape Descriptor Based on Centroid with Chord Lengths for Image Retrieval

Jayadevi Karur[✉] and Jagadeesh Pujari

S D M College of Engineering and Technology, Dharawad, India
jck1965@gmail.com, jaggudp@yahoo.com

Abstract. Shape is considered as a significant characteristic in object detection. In applications such as recognizing the shapes and classification, retrieval of shapes, extraction of shape features and their depiction plays an important role. We propose a shape descriptor which is computationally economical, compact and fast. In this work, we present three methods of feature computation based on chords. Among three, good result is obtained by the third method. Since the proposed work is based on centroid involving a global histogram, shape descriptor is invariant to all transformations. When compared to other shape signatures the proposed shape descriptor resulted in bull's eye score of 84.56% on MPEG-7 dataset.

Keywords: Computationally economical · Compact · Chords · New centroid

1   Introduction

Speedy progress in technologies of imaging and web has made it feasible for persons to refer and access an enormous number of images. Present day's applications, for example, matching and image retrieval are most natural phenomena. Retrieval of images from a huge database is the most important and thought-provoking problem in CVPR-Computer Vision and Pattern Recognition. Earlier methods of textual annotation of images are insufficient and at times, difficult in large databases due to requirement of intensive human labor. To overcome this, techniques based on CBIR (Content Based Image Retrieval), was presented. Image retrieval, in these methods, depends on visual contents instead of textual explanations. For representation and alphabetical listing of images, CBIR methods use visual fillings, for example, texture, shape and color. In object recognition, other than color and texture, shape plays an important role. Usually the shape plays a critical role in identifying regular objects [3]. Hence, in CBIR techniques, use of shape features of objects is increasing.

Two important operations of shape retrieval are shape depiction and matching. The first operation abstracts useful and perceptually significant shape features. These features are organized using data structure like string, tree, vector and so on. In the second operation, shape descriptors obtained in the first

operation are used to measure closeness (or difference) among the two shapes. Numerous shape depiction and evaluation techniques have been projected all through the previous decades. Generally, the present shape depiction and portrayal techniques are divided into two categories based on contour and region. Each of these techniques is further classified into two classes as global and local.

Some renowned and latest shape descriptors based on region comprise Zernike moments, multi-scale Fourier descriptor, generic Fourier descriptor and so on. There have been many works, based on contour, for shape representation and matching. Both local and global information are contained in many contour–based shape descriptors. The descriptors belonging to this category are CSS-Curvature Scale Space, PMEM-Polygonal Multi-resolution and Elastic Matching, SC-Shape Context, IDSC-Inner Distance Shape Context, CPDH-Contour Points Distribution Histogram, HPM-Hierarchical Procrustes Matching, ST-Shape Tree, CF-Contour Flexibility and HF-Height Functions and so on. MPEG-7 community has recommended CSS [7], as one of the standards of global shape depiction. In CSS, calculation of a closeness degree is centered on the greatest potential correspondence among utmost concave/convex curvatures that are present in basic forms of the boundary outlines. In [1] PMEM three primitives of every contour section are mined at various scales. After that, the sum of absolute differences, enhanced by the elastic matching, is calculated. This is used to quantify the comparability among shapes. In SC [2], spatial dispersal of all other sample points relative to every sample point is computed. This information is represented by a histogram in log-polar space. The bins considered in histogram are uniform bins. Shape descriptor in SC is more profound to adjacent points (sample) compared to distant points (sample). Ling et al. [5] proposed a new distance named inner distance, which they defined as "length of the shortest path between two points within the shape profile". It is used as an alternative for Euclidean distance. This makes shape descriptors, based on contour, not sensitive to articulation. They have defined new shape descriptor called Inner Distance Shape Context (IDSC), which is an extension of shape context with inner distance. CPDH [13] is a global method of shape descriptor. The shape is represented by spatial dispersal of boundary points in the polar co-ordinate space. Spatial dispersals of two shapes are equated using Earth Mover's Distance. In [6] HPM, shape information is captured across different hierarchies. The ordered geometric characteristics of a shape are extracted in ST [4]. Tree matching technique is used for shape matching. The shape descriptor proposed in CF [15], defines every contour point by its deformable potential and shape matching is done using DP-Dynamic Programming. In HF [14], height function is used to define a shape descriptor. For shape matching DP-Dynamic Programming is used in HF.

In [10], to contest corresponding pairs of random features, a method, called dynamic time wrapping (DTW) is presented. The method proposed by authors in [9], DP is used for matching the random features. In this method compacting pattern into a single vector is avoided. Usually medical images are complex and contain various regions. These regions are marked using arrows in biomed-

ical documents/publications. Simple and proficient method for recognition of arrows in images is presented in [8]. In this method, first, fuzzy binarization technique is used to convert images as binary images. Based on principle of connected component a set of candidates is calculated and convexity defect-based filtering is used to choose arrow candidates. This is followed by template matching via DTW-Dynamic Time Warping. Sequential classifier, encompassing of BLSTM-bidirectional long short-term memory classifier, trailed by convexity defect based arrowhead detection is proposed in [11]. For pattern recognition, a ranked method to extract subsets of soft output classifiers (decision rules) automatically is proposed in [12].

The above explained techniques produce outstanding retrieval precisions in the existence of diversity of distortions. These approaches utilize multifaceted distance measures, computationally complex correspondence techniques and advanced shape descriptors to attain great retrieval precisions. As a result, substantial quantity of time is spent for shape equivalent and retrieval. The existing approaches are not proper for retrieval of shapes from huge databases and online retrieval, even though they result in high retrieval accuracies. This results in dilemma to users with respect to speed-accuracy disparity. Hence we are motivated to design a new shape descriptor and matching procedure that takes into account, both accuracy and speed in its design. According to [13], a good shape depiction should be compact and should contain vital characteristics of the shape. Also it should be invariant to all types of transformations like translation, scaling and rotation, just as in the reliable human vision system. The added significant desires for favorable shape descriptor are broad-spectrum of applications, computational proficiency and compactness [14].

In this paper, we propose a new shape descriptor centered on centroid (not normal centroid) with chord lengths. It is simple and easy in computation, compact, fast retrieval rate and invariant to all transformations. MPEG-7 data set is used for experimentation purpose. The paper is organized as follows: The detailed discussion on the proposed shape descriptor is given in Sect. 2. Experimental results are presented in Sect. 3 and Sect. 4 concludes the paper.

2   Proposed Work

In this paper three methods, which are evolutionary in nature, for feature extraction and their performances are presented. The flow diagram of the work is given in Fig. 1.

MPEG-7 dataset consists of both binary and grey images. Therefore the images read from data set are binarized. The technique used in this work for determining edge pixels is called zero-crossing technique which is based on 8-neighborhood pixels only. The idea behind this technique is, all the pixels in the image are traced using 3X3 window for their 8-neighborhood connectivity. The sum of pixel values in that window is calculated. If this sum is less than 9 then it is edge pixel otherwise not edge pixel. By applying this technique over a whole image we obtained all edge pixels. The technique determines both
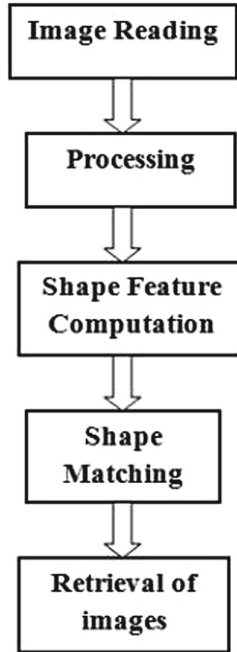
Fig. 1. Flow diagram

outer boundary pixels of the image as well as inner contour formed in the image object. Figure 2 represents result of this technique to determine edge pixels for the horse image in the dataset.

By removing the extra rows and columns that do not contain any of the object pixels we can limit the image by using minimum bounding box. The center of the bounding box is considered as a centroid (new centroid-NC) which is different from the normal definition of a centroid. NC is calculated as half of the length and breadth of minimum bounding box as shown in Fig. 3.

The circle is drawn with NC as a center and the distance of farthest edge pixel from NC as a radius using Bresenhams-circle algorithm. The algorithm returns coordinates of all points on the circle. The proposed work considers bounding box centroid rather than centroid of the image. The above procedure is followed in all three methods presented in this work as a first step. The remaining part of the work is explained below:

In the first method (M1) eight equidistance reference points on the circle are chosen. These points are set as viewing points of the object. The procedure for the same is as following:

Diameters are drawn using points on the circle. These diameters cut the edge pixels of the image. To fix first reference point on the circle, the procedure followed is given below:
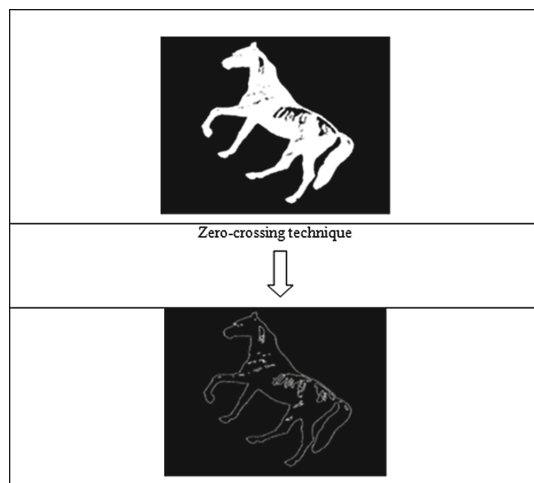
Fig. 2. Edge pixel image

Fig. 3. Bounding box with centroid

– First distances between extreme edge pixels on every diagonal are calculated.
– The average of highest ten distances is computed. The diameter yielding the closest distance to this average establishes the first pair of reference points.

This process will choose a reference line that divides the image into symmetrical loops along its length. The remaining three pairs of reference diameters are located on the circle at offsets of $\frac{1}{8}^{th}$, $\frac{2}{8}^{th}$ and $\frac{3}{8}^{th}$ portions of the circle with reference to the first reference diameter. The eight points chosen remain relative to the structure of the shape with respect to all the transformations. Lines from each reference point (v41) to all points (like c1) in the opposite half circle are considered for feature computation as shown in Fig. 4. Each line will intersect multiple edge pixels (consider e1, e2, e3, e4 on line v41 to c1). Distances between every pair of edge pixels on every line are computed and recorded in a histogram. This is repeated for all eight points on the circle. All the features are recorded in a single histogram which serves as a transformation invariant feature vector. Closeness between two descriptors is measured by comparing feature vectors of both query image and each shape of data set using L1 norm.
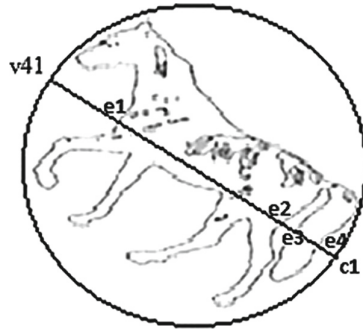
Fig. 4. Feature computation in M1

In second method (M2), feature computation involves both edge and region pixels. Fixing up reference diameter continues as explained in M1. The structure of the shape is evaluated along diameters. The diameter corresponding to this pair of reference points divides the circle into two halves. Lines from each point of one half of the circle to the corresponding point in the other half of the circle traced in the same direction are considered for feature computation (all the diameters within the half circle). These lines will intersect with edge pixels. Fraction of line lying in the original shape image is computed by counting the number of on pixels (with value 1) on a line. The ratio, named as OR (Occupancy Ratio), of this value to the length (digital length) between the extreme edge pixels on the line is calculated which represents one feature of the image as shown in Fig. 5. The Euclidean distance between extreme edge pixels, say p and q, is calculated as $distpq = \sqrt{(px - qx)^2 + (py - qy)^2}$. All the distances between extreme edge pixels are normalized by the largest distance to avoid scaling effect. All these distances and ORs are recorded in 2D histogram (64×32) where x-axis represents Euclidean distance and y axis OR values. Histograms of both query image and each shape of data set are compared using L1 norm for similarity of the two descriptors.

In third method (M3), the procedure is same as in M1, up to setting up of first pair of reference points on the circle. The diameter is drawn using these pair of points which divides the circle into two halves. Here the structure of the shape is determined in both horizontal and vertical directions of the shape along chords. All the parallel and perpendicular chords to this diameter are drawn using points on the circle and are considered for feature computation. These chords intersect with edge pixels as given in Fig. 6. Distances between extreme edge pixels on all the chords are computed. All these features are recorded in a single histogram. These distances are normalized by the largest distance in the image. The normalized histogram serves as the feature vector for the image. Closeness between two descriptors is measured by comparing feature vectors of both query image and each shape of data set using L1 norm.
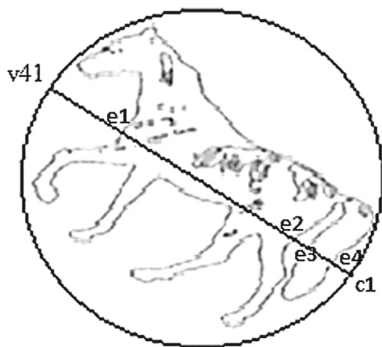
Fig. 5. Feature computation in M2 OR = on pixel(with value 1) on line e1e4/Length (digital length) of line between edge pixels e1 and e4
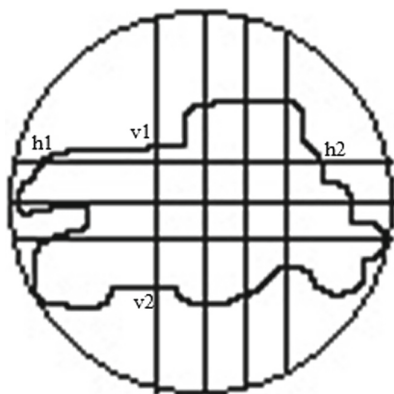

Fig. 6. Feature computation in M3

3   Experimental Results

The data set considered for the work is MPEG-7 which contains 1400 images (70 classes of objects, each class with 20 different shapes). Shapes within the same class are with some multifaceted distortions and some shapes from different classes are similar. Hence it is very challenging. The dataset is publicly available for testing.

A set of 70 arbitrarily chosen images, one from each class, forms the query set. For each query, full depth retrieval is done i.e. all 20 images of the query image class have been retrieved. To measure the retrieval performance precision and recall are used and are defined as follows:

$$\text{precision} = \frac{x}{y}$$

$$\text{recall} = \frac{x}{z}$$

where, x = number of correct images retrieved

y = total number of images retrieved

z = total number of correct images

Precisions are recorded for a standard set of eleven recall levels, in steps of 0.1 from 0 to 1 (inclusive of end values). The average precision–recall graph for the work in given in Fig. 7 and from this it is clear that method 3 of feature computation is efficient. We obtained bull's eye score of 84.56% on MPEG-7.
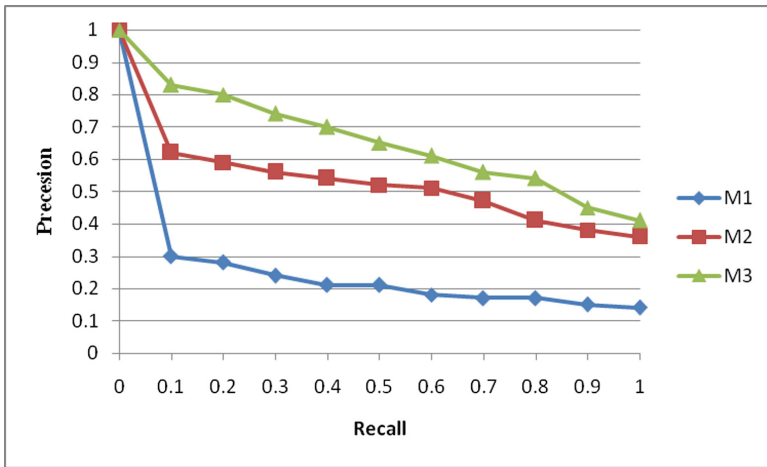

Fig. 7. Average precision – recall graph

4   Conclusion

A new shape descriptor is proposed which is centered on centroid (not normal centroid) with chord lengths. It is simple, compact, fast and computationally economical (points on circle alone are considered). Since the proposed work is based on centroid involving a global histogram, shape descriptor is invariant to all transformations.

References

1. Attalla, E., Siy, P.: Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. Pattern Recogn. **38**(12), 2229–2241 (2005). https://doi.org/10.1016/j.patcog.2005.02.009. http://www.sciencedirect.com/science/article/pii/S0031320305000919

2. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002). https://doi.org/10.1109/34.993558

3. El-ghazal, A., Basir, O., Belkasim, S.: Farthest point distance: a new shape signature for Fourier descriptors. Sig. Process.: Image Commun. **24**(7), 572–586 (2009). https://doi.org/10.1016/j.image.2009.04.001. http://www.sciencedirect.com/science/article/pii/S0923596509000393

4. Felzenszwalb, P.F., Schwartz, J.D.: Hierarchical matching of deformable shapes. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007). https://doi.org/10.1109/CVPR.2007.383018

5. Ling, H., Jacobs, D.W.: Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 286–299 (2007). https://doi.org/10.1109/TPAMI.2007.41

6. McNeill, G., Vijayakumar, S.: Hierarchical procrustes matching for shape retrieval. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 885–894. IEEE (2006)

7. Mokhtarian, F., Abbasi, S., Kittler, J.: Efficient and robust retrieval by shape content through curvature scale space. In: Image Databases and Multi-Media Search, pp. 51–58. World Scientific (1997)

8. Santosh, K.C., Alam, N., Roy, P.P., Wendling, L., Antani, S., Thoma, G.R.: A simple and efficient arrowhead detection technique in biomedical images. Int. J. Pattern Recognit. Artif. Intell. **30**(05), 1657002 (2016)

9. Santosh, K.C., Lamiroy, B., Wendling, L.: DTW for matching radon features: a pattern recognition and retrieval method. In: Blanc-Talon, J., Kleihorst, R., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2011. LNCS, vol. 6915, pp. 249–260. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23687-7_23

10. Santosh, K.C., Lamiroy, B., Wendling, L.: DTW-radon-based shape descriptor for pattern recognition. Int. J. Pattern Recognit. Artif. Intell. **27**(03), 1350008 (2013)

11. Santosh, K.C., Roy, P.P.: Arrow detection in biomedical images using sequential classifier. Int. J. Mach. Learn. Cybern. **9**(6), 993–1006 (2018)

12. Santosh, K.C., Wendling, L.: Pattern recognition based on hierarchical description of decision rules using choquet integral. In: Santosh, K.C., Hangarge, M., Bevilacqua, V., Negi, A. (eds.) RTIP2R 2016. CCIS, vol. 709, pp. 146–158. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-4859-3_14

13. Shu, X., Wu, X.J.: A novel contour descriptor for 2D shape matching and its application to image retrieval. Image Vis. Comput. **29**(4), 286–294 (2011)

14. Wang, J., Bai, X., You, X., Liu, W., Latecki, L.J.: Shape matching and classification using height functions. Pattern Recogn. Lett. **33**(2), 134–143 (2012)

15. Xu, C., Liu, J., Tang, X.: 2D shape matching by contour flexibility. IEEE Trans. Pattern Anal. Mach. Intell. **31**(1), 180–186 (2009). https://doi.org/10.1109/TPAMI.2008.199