

# Shape Based Image Representation and Retrieval

Adrijit Basu\*  
CMC Limited,  
India

## Abstract—

**A**mong all the issues related to Content Based Image Retrieval systems, retrieving images based on their shapes is an important one. Many approaches exist utilizing shape representation and comparison, e.g., the methods based on Fourier descriptors. In this thesis, we propose a novel method for shape representation. In our method, we calculate the centroid of a shape and choose a set of sample points around this shape's boundary. From those, we obtain a set of radii. We then use these radii lengths to construct a Distance Histogram as the representation of shape. The natural characteristic of this method makes itself invariant to rotation and translation. Furthermore, it can be made invariant to scale by a simple normalization. To evaluate the approach, we perform a large set of experiments on a database of shapes.

**Keywords—** Image, Polygon, Shape, Content Based Retrieval

## I. INTRODUCTION

Image retrieval - how to retrieve images similar to a query - is becoming a more important problem with the rapid increase of media information on the web. Generally, users want to provide query images and use some kind of system which can present them with a set of similar images. Therefore, how to describe and model an image, how to compare different images and judge whether they are similar, how to construct an index of image databases, and how to conduct searches efficiently are some of the key problems in any so-called image retrieval system.

An image can be described by several low level features. These features, which include shape, color, texture, and spatial relationship are called the content of image. By using these features, we can not only describe and model an image, but also make comparisons among different images. Therefore, an image retrieval system which is based on these low level features is called content-based image representation and retrieval system.

## II. PROBLEM STATEMENT

Shape based image representation and retrieval is a defined problem where aim is to extract and store the low level features of images with a searchable database and the technique allows to search and retrieve similar images using those shape features.

## III. ASSUMPTIONS AND SCOPE

In this paper, we only concentrate on the issues of shape representation, shape similarity measure and shape indexing. Extracting shapes from images are not our concern. We assume that when an image is added into an image database, it is associated with a set of shapes representing the objects inside the image. This can be an automated process by incorporating any available feature extraction techniques or it may require user intervention. Therefore, we assume that the inputs to our work are the shapes extracted from images and the shapes are in the form that we described below.

Our work concentrates on 2D polygonal shapes instead of arbitrary shapes. We will use the word shape and the word polygon interchangeably hereafter. At present stage, our work only handles simple non-degenerate closed polygons as defined above.

### A. Definitions

1) *Definition 1:* A polygon is represented by an ordered list of vertices  $P = \{V_1, V_2, \dots, V_n\}$ , where  $n$  is the number of vertices of the polygon and  $V_i \in \mathbb{R}^2$ .

2) *Definition 2:* A polygon is simple if no two edges of the polygon cross each other.

*Definition 3:* A polygon is non-degenerate if  $1 \leq i \leq n$  such that  $V_i, V_{(i+1) \bmod n}$  and  $V_{(i+2) \bmod n}$  are collinear.

## IV. LITERATURE SURVEY

Shape based image representation and retrieval is a defined problem where aim is to extract and store the low level features of images with a searchable database and the technique allows to search and retrieve similar images using those shape features.

### A. Fourier Descriptors-based Approach

In 1977, Persoon and Fu [1] first proposed the technique of using Fourier descriptor as the representation of shapes. A large amount of research had been done following their idea.

### B. Grid-based Method

Sajjanhar and Lu [2] propose another method which can be used for shape representation and similarity measure. It is called the grid-based method.

### C. Turning Angle

Arkin and colleagues proposed an efficiently computable method to represent the polygonal shapes [3], called the turning angle method.

### D. Rectangular Covers

In Jagadish's paper [4], another approach for representation of shapes, called rectangular covers, is proposed.

### E. Partition Token

Berretti, Bimbo and Pala proposed an interesting representation for generic shapes [5], using local features. Each shape is partitioned into several tokens corresponding to a set of perceptually salient attributes.

### F. Centroid-Radii Model

Tan and Thiang proposed a novel method for shape representation in their paper [6]. The basic idea of their proposal is to use the centroid-radii model to represent shapes.

### G. Other Representations for Shapes

Grosky and Mehrotra's proposal [7][8] also presents a method for shape representation. This representation is based on local structural features, in which a local structural feature is defined by considering the internal angle at a vertex, the distance from the adjacent vertex, and the vertex coordinates. In each shape, a fixed number of these local features can be extracted.

Mehrotra and Gary develop a technique which is quite similar to Grosky and Mehrotra's [9], but is more robust. In their approach, they still use local structural features to represent the shape, and each structural feature is represented as a point in a multidimensional space.

Scaroff and Pentland, in their proposal for the Photobook system [10], represent the shapes as modal deformations of a prototype object. As a result, shape similarity can be computed by calculating the distances between mode vectors.

## V. SHAPE REPRESENTATION & RETRIEVAL - A METHOD BASED ON DISTANCE HISTOGRAMS

### A. Motivation

In the previous section, we reviewed several approaches to shape representation. In our opinion, for good shape representation, there are three important criteria:

- The representation should be invariant to translation, rotation, and scale. Fig. 1 illustrates an example of translating, rotating, and scaling a polygon. In this case, we assume the polygons in Fig. 1 should be considered similar by a shape representation
- It should be reasonably easy to implement
- It should match the human intuitive notion of shape similarity

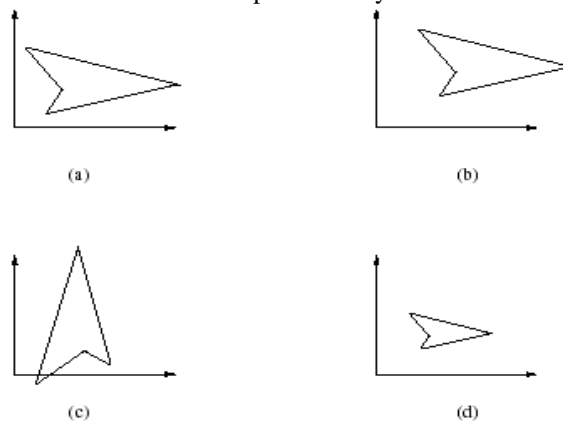


Fig. 1 (a) Polygon P (b) P after translating (c) P after rotating (d) P after scaling

When we think about the geometric feature of shapes, we find that just as with the center of a circle, the centroid of a shape is quite important (See Fig. 2). Obviously, given a shape, the position of its centroid corresponding to its boundary will not change under translation and rotation.

We know that any point in a circle's boundary has the same distance, called the radius, to its centroid. We can apply this method to an ordinary shape. However, in other shapes, only one radius may not be enough because the distances from points in the boundary to the centroid may be different. Therefore, we can use a set of radii which begin at the centroid and end at the boundary, to represent a shape. It means we can describe the shape as long as we use enough radii. Fortunately, this representation is invariant to translation and rotation.

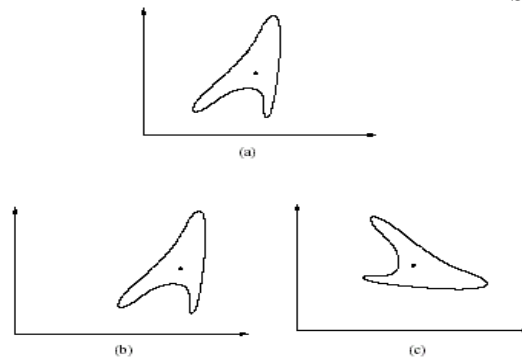


Fig. 2 (a) A shape and its centroid (b) A shape and its centroid after moving (c) A shape and its centroid after rotation

Based on the discussion above, we deal with a new methodology to describe and represent shapes. Further, we will also discuss how to use this methodology to compare and search for shapes.

Given an image which contains a single object, we assume the boundary of the object, i.e., its shape, has already been extracted. There are several types of algorithms which can be used to extract the boundary of objects. These are, however, beyond the scope of this thesis. The problem we are concerned with is how to describe the object's shape when its boundary is already available. Basically, the approach we will propose contains four parts:

1. Given a polygon which approximates the boundary of a shape, we calculate its centroid.
2. A set of sample points in the boundary of the polygon is selected and the distances between them and the centroid are computed.
3. A histogram based on the distances computed in the previous step is constructed.
4. Finally, we normalize the Distance Histogram.

### B. Centroid of Polygon

Given a polygon such as the one in Fig. 3, we need first to find its centroid. In order to compute the x and y coordinate of the centroid, we first need to calculate the area of the polygon. The method we will use is based on Green's Theorem in a plane. Given a polygon and its vertices  $(x_i, y_i)$ ,  $i=0, \dots, n$ , with  $x_0 = x_n$  and  $y_0 = y_n$ , the following formula can be used to calculate the area of a polygon in a plane:

$$A = \frac{1}{2} \sum_{i=0}^{n-1} \alpha_i$$

where

$$\alpha_i = x_i y_{i+1} - x_{i+1} y_i$$

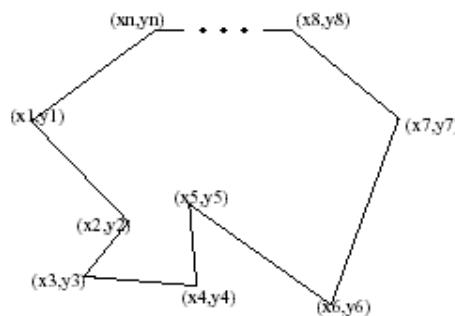


Fig. 3 A polygon which has n vertices

The area computed in this way is a signed value, where a negative sign indicates the vertices are in clockwise order, and a positive sign indicates the vertices are in a counter-clockwise order. Now we can compute the centroid based on the area. The coordinates for the centroid are:

$$\bar{x} = \frac{\mu_x}{A} \quad \bar{y} = \frac{\mu_y}{A}$$

where

$$\mu_x = \frac{1}{6} \sum_{i=0}^{n-1} (x_{i+1} + x_i) \alpha_i$$

$$\mu_y = \frac{1}{6} \sum_{i=0}^{n-1} (y_{i+1} + y_i) \alpha_i$$

The centroid of any polygon in a plane can be calculated by this method.

**C. Sample Points Selection and Distance Calculation**

In the previous section, we presented the method of computing the centroid of a polygon. Naturally, for describing the entire shape, the centroid alone is not enough. In our approach to shape representation, we need a set of sample points selected from the boundary to describe the shape. Therefore, in this section, we will discuss how to select the sample points around the boundary of a polygon, and how to calculate the distance from the centroid to the sample points.

Firstly, we should decide how many sample points will be selected around the boundary. We consider this number to be a variable, therefore we can change it for different situations. Secondly, we should decide how to select these sample points. Here, we have two problems. One is how many sample points should be in each edge; the other is how to select them within the edges.

To deal with the first problem, we can assign the number of sample points randomly to each edge. This method presents a difficulty though: sometimes a very short edge is assigned more sample points than a much longer one. Furthermore, two polygons with the same shape may have different sample points. Therefore we choose to assign the number of sample points in each edge proportionally to its length. For example, if the length of an edge is  $L_i$ , the sum of the length of all edges is  $L_{sum}$ , and the entire number of sample points is  $N$ , then the number of sample points in this edge  $N_i$  will be:

$$N_i = \frac{L_i}{L_{sum}} N$$

To deal with the second problem, we decide the location of sample points so that they are evenly spread in the edges.

Finally, we can use sample points and the centroid to calculate the distances. For example, given a sample point  $s_i = (x_i, y_i)$  and the centroid  $c = (x_c, y_c)$ , the distance between them is:

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

In this way, we can calculate the distances from the centroid to all sample points around the boundary, and we are able to use them to construct a Distance Histogram to describe the shape. Using the centroid and sample points in the boundary to describe the shape is one of the central ideas of our proposal. This is discussed in the next section.

This approach, based on distances has some apparent advantages. The distance set we use to represent the shape will not change after translating the shape. This is because the distances only have a relationship with the centroid and sample points in boundary. If we just move the shape, the location of the centroid will not change, and neither will the locations of the sample points. Therefore, the distance set will not change. If we rotate the shape, the distance set will still not change, the reason being that we assigned the sample points in each edge proportionally to its length, and the sample points are spread evenly in each edge. Therefore, the location of the sample points will not change after rotating the entire shape. As for the problem of invariance under scale, we can make a simple normalization of the distances. We will discuss this in next Section.

**D. Histogram**

A histogram is a very useful tool to represent some property of data, and is used in many fields and applications. In this section, we will present the idea of how to construct the Distance Histogram.

For shape representation, generally it is not trivial to be invariant to rotation. For example, in the methods based on Fourier descriptors [11] and turning angle [3], some kind of complicated process is required in order to solve the problem of invariant to rotation. Fortunately, in our method, solving this problem is quite easy and straightforward.

In the methods based on Fourier descriptors and turning angle, it is not only necessary to select the sample points around the boundary, but also to pick a starting point (e.g., the top-left corner) which will be used in the next computation. This starting point will be changed when the polygon is rotated, which will affect the next computation, and result in generating a different representation of a polygon before and after rotation. To avoid this problem, we use the Distance Histogram in our approach. It works as follows:

Firstly, we separate the range of all distances, which is  $[0, D_{max}]$ , into several ranges, e.g.,  $R$  ranges. Then, the ranges of distances can be represented by:  $[0, D_{max}/R]$ ,  $[D_{max}/R, 2 \times D_{max}/R]$ ,  $[2 \times D_{max}/R, 3 \times D_{max}/R]$ , ...,  $[(R - 1) \times D_{max}/R, D_{max}]$ . Secondly, we compute the number of distances in each range, e.g., if we have a sample point and its distance from the centroid is  $1.5 \times D_{max}/R$ , then we will increase the number of ranges  $[1 \times D_{max}/R, 2 \times D_{max}/R]$  by 1.

**VI. RESULTS AND INTERPRETATIONS**

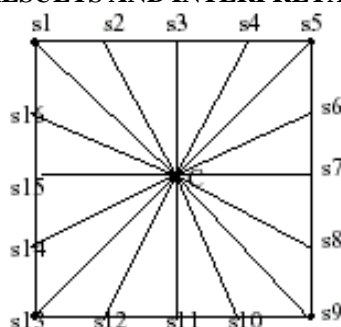


Fig. 4 A polygon and its sample points and radii

Fig. 4 illustrates an example. The polygon in Fig. 4 is a square, and we evenly select 16 sample points around its boundary. We define  $D_n$  is the original distance from centroid C to sample point  $s_n$ , which  $n \in (1, 2, 3, \dots, 16)$ . Given that the lengths of each edge of this rectangle are the same, which is 4, then we have:

$$\begin{aligned} \mathbf{D}_1 = \mathbf{D}_5 = \mathbf{D}_9 = \mathbf{D}_{13} &= \mathbf{2.828} \\ \mathbf{D}_2 = \mathbf{D}_4 = \mathbf{D}_6 = \mathbf{D}_8 = \mathbf{D}_{10} = \mathbf{D}_{12} = \mathbf{D}_{14} = \mathbf{D}_{16} &= \mathbf{2.236} \\ \mathbf{D}_3 = \mathbf{D}_7 = \mathbf{D}_{11} = \mathbf{D}_{15} &= \mathbf{2.000} \end{aligned}$$

The maximum value is 2.828. We separate the range of value  $[0, \text{maximum}]$  into 4 ranges, which are  $[0, 0.707]$ ,  $[0.707, 1.414]$ ,  $[1.414, 2.121]$ ,  $[2.121, 2.828]$ . We then know that:  $D_3, D_7, D_{11}, D_{15}$  belong to range  $[1.414, 2.121]$  and the other distances belong to range  $[2.121, 2.828]$ . Fig. 5 illustrates the histogram:

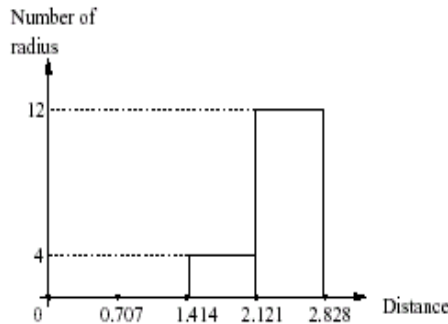


Fig. 5 Distance Histogram

This histogram can be used to represent the shape of polygons. In this sense, our approach can be referred to as a method based on Distance Histograms.

In the procedure above, the selection of sample points will not change with the rotation, nor will the radii. Therefore, the shape representation by Distance Histograms is independent of the starting point in the boundary; what matters is only the number of radii in each range. This number will not change under the rotation. Therefore, our shape representation is invariant to rotation. However, this method has a problem: two similar shapes with different sizes may have different Distance Histograms as representations. This means that our approach is still sensitive to scale. To solve this problem, we need to make some kind of normalization to the distances. This is discussed next.

**E. Normalization**

Given two polygons with different sizes, but the same shape, they should be considered similar. However, if we only use original distances and a Distance Histogram constructed from them to represent and compare them, they will be considered different.

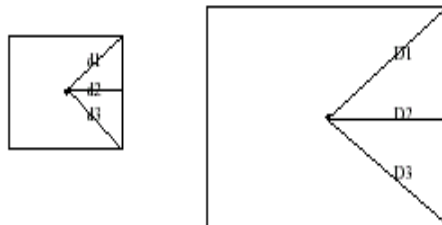


Fig. 6 Two polygons with different sizes but similar shapes

For example, in Fig. 6, two rectangles have different sizes but the same shape. If we only use  $(d_1, d_2, d_3)$  and  $(D_1, D_2, D_3)$  to represent the shapes and make the comparison, these two similar shapes will get different representations and the result of the comparison will be not similar. Therefore, we need to perform some kind of process on the distances which were obtained in previous Section.

Normalization is a very useful process for solving this problem. Firstly, we need to find the maximum distance among all the distances. Then, we divide all of them by this maximum one and get the normalized distances. After this process, the value of all normalized distances will be in the range  $[0, 1]$ . And because we assign the sample points based on the length of the edges, and evenly spread them in each edge, two polygons which have different sizes but the same shape will generate the same normalized distances. Therefore, our method is invariant to scale after normalization.

After normalization, the histogram shown in Fig. 4 will become the one illustrated in Fig. 7.

**F. Similarity Metric**

Given a shape, we can use the method discussed in the previous section to construct a Distance Histogram as its shape representation. The Distance Histogram can be represented as:

**D:**  $(d_0, d_1, d_2, \dots, d_n)$

where  $n$  is the number of ranges in the Histogram and  $d_i$  ( $i \in [0, n - 1]$ ) is the number of distances belong to this distance range. In this way, given two shapes:

**D<sub>1</sub>:**  $(d_{10}, d_{11}, d_{12}, \dots, d_{1(n-1)})$

**D<sub>2</sub>:**  $(d_{20}, d_{21}, d_{22}, \dots, d_{2(n-1)})$

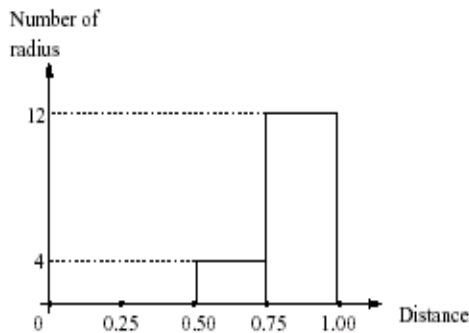


Fig. 7 Normalized Distance Histogram of polygon

Their similarity can be measured by Euclidean distance:

$$SIM(D_1, D_2) = \sqrt{\sum_{i=0}^{n-1} (d_{1i} - d_{2i})^2}$$

## VII. CONCLUSIONS

The goal of finding a simple and straightforward method for shape representation, with relatively good efficiency, was the motivation for our research. As we have discussed, the existing approaches to shape representation, (the method based on Fourier descriptors, grid-based method, etc.), have their own special advantages; however, they have limitations as well. Therefore, we proposed a novel idea for shape representation: a method based on the Distance Histograms. In our approach, we represent shapes by using their Distance Histograms. A Distance Histogram represents the distribution of radii lengths from sample boundary points to the shape's centroid, and provides the main data for us to represent, store, compare, and index shapes. Compared with other classic approaches to shape representation – like the method based on Fourier descriptors - our approach has these features:

- It is much more economical in its use of storage overhead and query processing time during the retrieval procedure
- It is highly flexible and may be adapted to be as efficient as the method based on Fourier descriptors

## REFERENCES

- [1] E. Persoon, and K. S. Fu. Shape discrimination using Fourier Descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7(2):170-179, March 1977.
- [2] A. Sajjanhar, and G. Lu. A grid based shape indexing and retrieval method. *Special Issue of Australian Computer Journal on Multimedia Storage and Archiving Systems*, 29(4):131-140, November 1997.
- [3] E. Arkin, L. P. Chew, D.P. Huttenlocher, K. Kedem, and J. Mitchell. An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3):209-215, 1991.
- [4] H. V. Jagadish. A retrieval technique for similar shapes. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '91)*, pages 208-217, Denver, May 1991.
- [5] S. Berretti, A. D. Bimbo, and P. Pala. Retrieval by shape similarity with perceptual distance and effective indexing. *IEEE Transactions on Multimedia*, 2(4):225-240, December 2000.
- [6] K-L. Tan, and L. F. Thiang. Retrieving similar shapes effectively and efficiently. *Multimedia Tools and Applications*, Kluwer Academic Publishers, the Netherlands. (Accepted for publication in 2001).
- [7] W. I. Grosky and R. Mehrotra, Index-based object recognition in pictorial data management. *Comput. Vis. Graph. Image Process.*, Vol.52, pages 416-436, 1990.
- [8] W. I. Grosdy, P. Neo, and R. Mehrotra. A pictorial index mechanism for model-based matching. *Data Knowledge Engineering*, vol. 8, pages 309-327, 1992.
- [9] R. Mehrotra, and J. E. Gary. Similar-shape retrieval in shape data management. *IEEE Computer*, 28(9):57-62, 1995.
- [10] [PPS94] A.P. Pentland, R. Picard, and S. Sclaroff. Photobook: Tools for content-based manipulation of image databases. In *Storage and Retrieval for Image and Video Databases*, pages 34-47, 1994.
- [11] D. Rafiei, and A. Mendelzon. Efficient Retrieval of Similar Shapes. *The VLDB Journal*.
- [12] I. L. Cheng. Image databases: A content-based type system and query by similarity match. Masters thesis, Department of Computing Science, University of Alberta. Available as Technical Report TR-99-03, 1999.
- [13] B. Scassellati, S. Alexopoulos, and M.D. Flickner. Retrieving images by 2D shape: a comparison of computation methods with human perceptual judgments. In *Proceeding of SPIE Storage and Retrieval for Image and Video Databases*, pages 2-9, 1994.
- [14] <http://www.ctr.columbia.edu/VisualSEEK>, 2001.
- [15] Y. Sakurai. The A-tree: an index structure for high-dimensional spaces using relative approximation. *Proceeding of 26th International Conference on Very Large Data Bases*, pages 516-526, 2000.