Shape Retrieval Based on Distance Ratio Distribution

Dalong Li[1], Steven Simske
Intelligent Enterprise Technologies Laboratory
HP Laboratories Palo Alto
HPL-2002-251
September 9th , 2002*

image database, shape retrieval, geometric transformation, standard deviation, down-sampling

We propose a shape-matching algorithm based on a distance ratio distribution (DRD), which is used to measure the regularity of the transformation between the shapes to be compared. The standard deviation of the DRD is used to represent the distribution. Our algorithm is robust to translation, scaling and rotation. We illustrate the effectiveness of our algorithm in the content-based retrieval of shapes using subsets of the MPEG-7 standard database SQUID [31]. The experimental results show the competitiveness of our approach.

* Internal Accession Date Only
Approved for External Publication
[1] Georgia Institute of Technology

Shape Retrieval Based on Distance Ratio Distribution

Dalong Li[+], Steven Simske
Intelligent Enterprise Technologies Lab
HP Laboratories Palo Alto
([+]Georgia Institute of Technology / HP Labs)
July 30th, 2002

ABSTRACT

We propose a shape-matching algorithm based on distance ratio distribution, which is used to measure the regularity of the transformation between the shapes to be compared. The standard deviation is used to represent the distribution. Our algorithm is robust to translation, scaling and rotation. We illustrate the effectiveness of our algorithm in the retrieval of shapes by content using the subsets of the MPEG-7 standard database SQUID[31]. The experimental results show the competitiveness of our approach.

Keywords: image database, shape retrieval, geometric transformation, standard deviation, sampling theorem.

1 Introduction

Recently, visual information retrieval has become a major research area due to the increasing rate at which images are generated in many application fields. Visual information retrieval systems support retrieval by visual content by directly addressing image perceptual features such as color [1]–[3], shape [4]–[8], texture [9]–[11] and spatial relationship [12]–[14]. In this paper, our concern is shape.

Shape-matching has been approached in a number of ways including tree-pruning[15], the generalized Hough transformation[16], or pose clustering[17], geometric hashing[18], the alignment method[19], deformable template[20], relaxation labeling[21], Fourier descriptors[22], wavelet transformation[23], curvature scale space[24], neural networks[25], dynamic programming[32] and shape context [33].

Persoon and Fu [26] first proposed the technique of using the Fourier descriptor as the representation of shapes. A large amount of research had been done following their idea. It is easy to make the Fourier descriptor invariant to translation and scaling. To make it invariant to rotation, the optimal rotation angle must be obtained. Like in many other techniques, rotation is always something hard to handle compared with translation and scaling.

The grid-based method was proposed by Sajjanhar and Lu [27] for shape representation and similarity measurement. The shape is mapped onto a grid of fixed cell size, and is justified to the top left corner. Then, this grid is scanned from left to right, and also from top to bottom. "1" is assigned to each cell of the grid which is partially or wholly covered by the shape, and "0" to the cell which is not covered by the shape, obtaining a sequence of 0's or 1's. This sequence can be used for shape representation, which is robust to translation but not to scaling and rotation. The "major axis" is introduced to make this representation invariant to rotation and scaling. The major axis of a shape is the straight-line segment joining the two points on the shape boundary farthest away from each other [27]. It is expensive in computing to locate the major axis from a shape. For most shapes, the major axis is unique. Denoting the major axis as AB, we start with a random point named C_0 on the boundary. This point is fixed and the distance to all other points on the boundary is computed. The longer the contour, the more computing required. The maximum distance from point C_0 will then have to be compared to the maximum distance from $C_1, \ldots, C_n$ to obtain a global maximum.

The related research that we are most interested in is the Centroid-Radii Model. The paper is organized in the following way: related work is reviewed in section 2 and our proposal follows in section 3. Experimental results are also reported in section 3, followed by the conclusion.

2 Related works: The Centroid-Radii Model

Tan and Thiang proposed a novel method for shape representation in their paper [28]. The basic idea of their proposal is to use the centroid-radii model to represent shapes. Figure 1[30] illustrates an example. In this method, lengths of the shape's radii from centroid to boundary are used to represent the shape. The interval between radii, measured in degrees, is fixed.

Without loss of generality, the intervals are taken counter-clockwise, starting from the X-axis direction. The shape can then be represented by a vector $\{L_0, L_1, \ldots, L_i, \ldots L_{n-1}\}$: where L_i is the i th radius from the centroid to the boundary of shape. To make this approach invariant to scale, all radius lengths can be normalized by dividing by the longest radius among them: In this way, the shape can be represented in detail with sufficient radii used. In Tan and Thiang's method, two shapes are considered similar to each other if and only if the lengths of their radii at the respective angles differ by a trivial value [28].

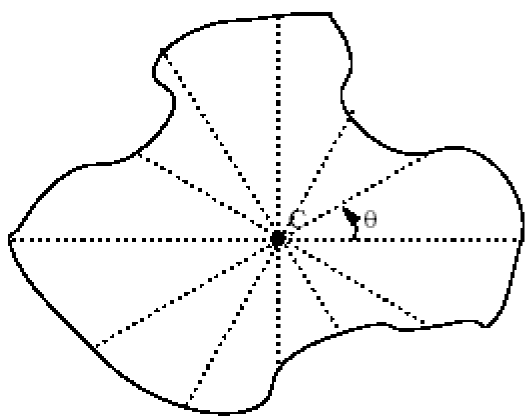

Figure 1 Centroid-Radii Model [30]

In most cases, the representation itself is not robust to rotation. When a different starting point is chosen, a different radius vector is obtained. In some cases, this representation is rotation invariant; for example, for any circle. Another thing is all the boundary points are used in computing the centroid. Once the centroid is located, the sample points corresponding to each angle can be located. Here the computing cost is quite high. Another potential problem is that the radius may cross the curve more than once…i.e. for non-convex contours. One possible solution is to choose the radius with minimum distance to the centroid.

Green's Theorem in a plane is used in locating the centroid from a polygon [29]. In order to compute the x and y coordinate of the centroid, we first need to calculate the area of the polygon. The following formula can be used to calculate the area of a polygon in a plane: Given points (x_i, y_i), i = 0, ..., n, with $x_0 = x_n$ and $y_0 = y_n$

$$\overline{x} = \frac{\mu_x}{A} \qquad \text{Eq 1}$$

$$\overline{y} = \frac{\mu_y}{A} \qquad \text{Eq 2}$$

$$\mu_x = \frac{1}{6}\sum_{i=0}^{n-1}(x_{i+1} + x_i)\alpha_i \qquad \text{Eq 3}$$

$$\mu_y = \frac{1}{6}\sum_{i=0}^{n-1}(x_{i+1} + x_i)\alpha_i \qquad \text{Eq 4}$$

$$A = \frac{1}{2}\sum_{i=0}^{n-1} a_i \qquad a_i = x_i y_{i+1} - x_{i+1} y_i \qquad \text{Eq 5}$$

The area computed in this way is a signed value, where a negative sign indicates that the vertices are in clockwise order, and a positive sign indicates that the vertices are in a counter-clockwise order.

Another similar work based on the model is reported by Fan [30]. To make the comparison robust to rotation, a histogram is computed from the radii and used to compare shapes. The approach contains four parts:

1. Given a polygon that approximates the boundary of a shape, calculate its centroid.
2. A set of sample points in the boundary of the polygon is selected and the radii are computed.
3. A histogram based on the distances computed in the previous step is constructed.
4. Finally, the distance histogram is normalized.

As with all histograms, the distance histogram discards spatial information to achieve the robustness to rotation. It is not surprising, therefore, for it to fail to distinguish different shapes which have the same histograms as shown in Figure 2.

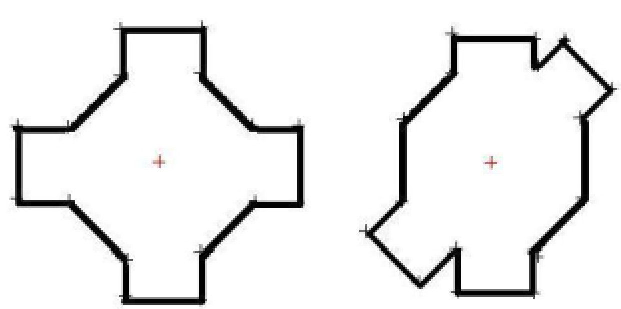

Figure 2 Two different shapes with same distance histogram

3 Distance Ratio Distribution (DRD)

Our algorithm also uses the Centroid-radii model, but in a different way. To explain our algorithm, let us look at the following triangle examples as an analogy for understanding. The basic idea is illustrated in Figures 3 and 4. In both of the figures, o is the centroid of the triangles, which are similar to each other. And we can see that the ratios {oa/oA} = {ob/oB} = {oc/oC}; hence the standard deviation of the ratio is zero. On the other hand if the two triangles are not similar to each other, for example, a right-angled triangle (one of the inner angles is 90 degrees) and an equilateral triangle (all inner angles are 60 degrees), the corresponding standard deviations will be larger than 0. Notice here in the example, no rotation is necessary. But it is necessary to rotate the distance vectors before a general comparison is made. For a given distance vector {a, b, c}, there are six possible distance vectors with circle shifts. They are: {a, b, c}; {b, c, a}; {c, a, b}; {c, b, a}; {b, a, c}; {a, c, b}. Generally, there are 2*N different vectors after a N-point vector is circular-shifted.

To have a better understanding of the basic idea, suppose contour is moving outwards. If the speeds are proportional to their radius, the position of the centroid remains and more importantly, the ratio of the distance is still **uniform**. Otherwise if their speed is not proportional to their radius, a different shape is formed. For any two shapes, we can assume that there is a transform between them. If the transform is linear and uniform, the transformed shape is regarded as similar as the original one. If the transform is geometrical such as rotation, translation, scaling, then the shapes are actually geometrically identical. The key idea in our work is to use the distance ratio standard deviation as a measurement of the transform, thus to measure the similarity of the shapes.

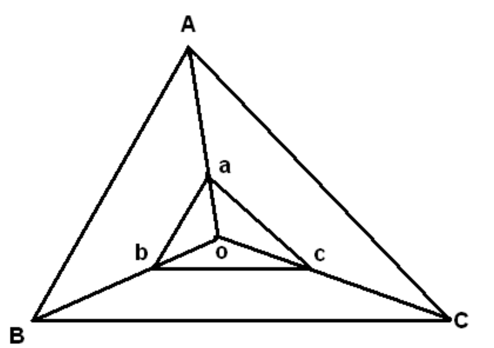

Figure 3 Geometric similarity(static view)

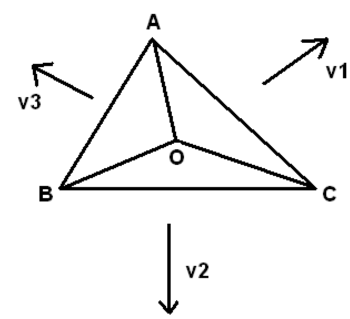

Figure 4 Geometric similarity(dynamic view)

What we discussed in the examples so far is exact matching. There is approximation error when general shapes are compared. However these errors are not always a bad thing for retrieval. For the purpose of retrieval, similar shapes besides same shapes shall be retrieved. When an exact matching algorithm is used in a retrieval system, you usually cannot return similar shapes. The recall rate will be too low, though you will get a very high precision rate.

In general our algorithm takes the following steps:

- Uniformly sample N points along the contour of the shapes
- Compute the centroid from the N sample points
- Get the radii vectors (rv1, rv2)
- Circularly shift one of the radii vectors and compute the distance ratios
- From the ratios, compute the standard deviations and choose the minimal one from the 2N standard deviations

- Compare the minimal with the trained threshold to decide whether the two shapes are similar

The shape is sampled evenly along the contour. It does not matter where you start since we actually circularly shift the vectors when compared. From the above, it is clear that the parameters are N, the number of sample Points, and T, the threshold. The optimal threshold is automatically selected from the training sets where we manually classify the shapes so that we already know whether the compared shapes are similar or not. The threshold is related to the number of the sample points.

From sampling theory, we know that the minimum sample rate to guarantee the exact recovery of original signal is $2*f_{max}$, where the f_{max} is the maximum frequency of the signal. However, in our work, the sample frequency is much lower than the required minimum sample rate (down-sampling). Down-sample is effectively a low-pass filter. Therefore, what we get after the down-sampling is the low frequency component of the shapes. Consequently the low frequency component of the shape is compared. It is very important for the low-frequency components to be compared since the low frequency components of two similar shapes are similar though they usually differ in high frequency. If high frequency component is included when two shapes are compared, many similar shapes cannot be retrieved. This will lead to a low recall rate in a retrieval system.

To determine N, the optimal number of sample points and T, the optimal threshold, we conduct experiments on a training set. In the training set, we select a certain number of pairs of fish from the SQUID [31] database. There are two groups of fish, one group is made up of individually different fish; another group of fish is made up of similar fishes. The following figures display the distribution of the comparisons under different sample points. Each blue(#) point is the minimum standard deviation of all the 2*N distance ratio vectors between the similar fish and the red(*) one is for the dissimilar fishes.

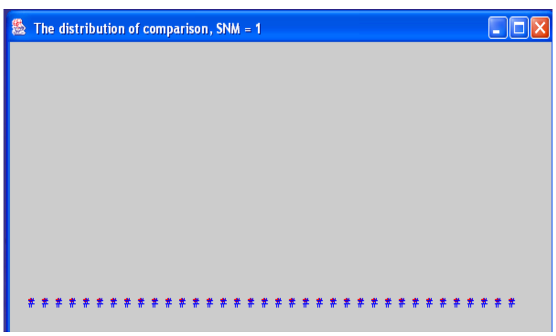

Figure 5-1 the comparison when N is 1

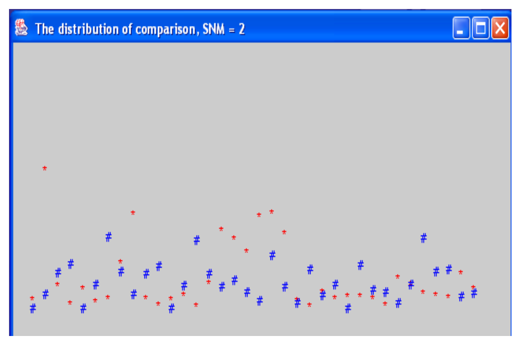

Figure 5-2 the comparison when N is 2

In Figure 5-1 N = 1 and the standard deviation of a sequence with one element is not defined. We cannot distinguish anything if N = 1. When N = 2, It is better than in Figure 5-1. But the similar shapes and the dissimilar shapes are still mixed. For N = 8 (Figure 5-3), an improved capacity to distinguish like images is evidenced. When N = 200 (Figure 5-4), we can see that the performance is worse. The reason lies in sampling theory. When N is higher, high frequency components are included and compared; however similar shapes usually differ in high frequency.

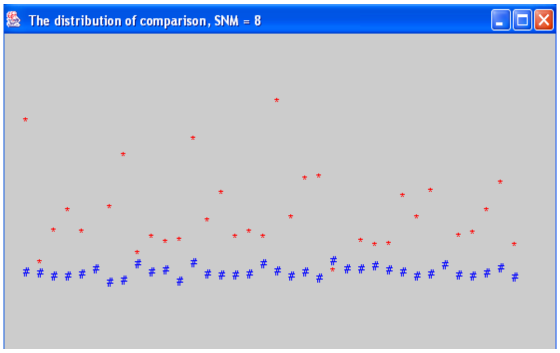

Figure 5-3 the comparison when N is 8

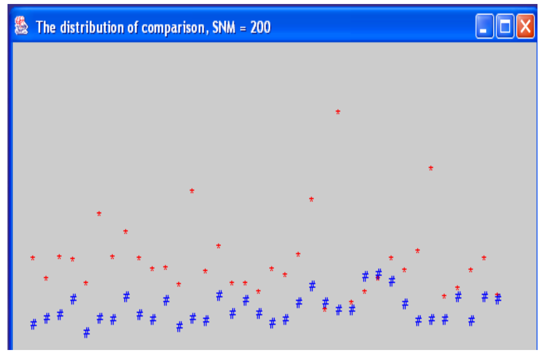

Figure 5-4 the comparison when N is 200

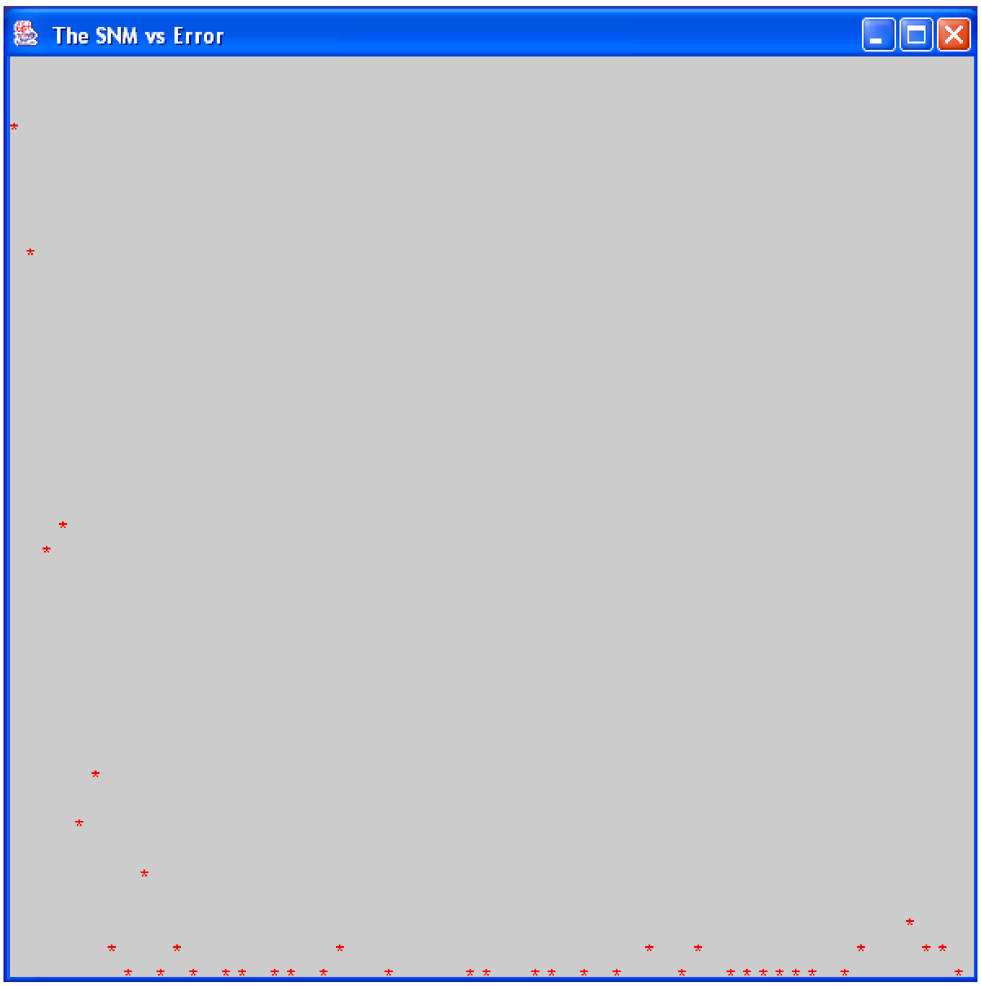

Figure 6 Errors under different sample points

The plot in Figure 6 summarizes the relationship between N and the minimum error, which is the sum of both recall and precision errors. The optimal threshold is the boundary line which best distinguishes the blue # points and the red * points in the plots such as Figure 5-3. Under the optimal threshold, the smallest number of total mistakes is made. At the beginning the error drastically drops with the increase of the sampling number. At some point (optimal N), it become stable and later on it becomes even worse when N is further increased. The optimal N is the smallest number of sample points where the shapes can be distinguished correctly: the standard deviation of the similar shapes is below T and that of the unlike shapes is above T. The optimal N is determined by the frequency components. Usually the optimal N is obtained by experiment.

From the training set, we can obtain the optimal sample points and the associated threshold we need to build a retrieval system. Table 1 is one group of them.

Table 1 a sample of trained T and N

N	13	16	19	22	23	25	26
T	0.107	0.099	0.088	0.105	0.082	0.099	0.107

Next, the retrieval system is tested on the SQUID database. There are 1100 unclassified fish. If you have 20 sample fish then you need to make 20*1100 (22000) comparisons! To evaluate the retrieval performance, these 22000 comparisons need to be done manually since the database itself is not classified. In our preliminary experiments, a representative set of fish is chosen from SQUID.

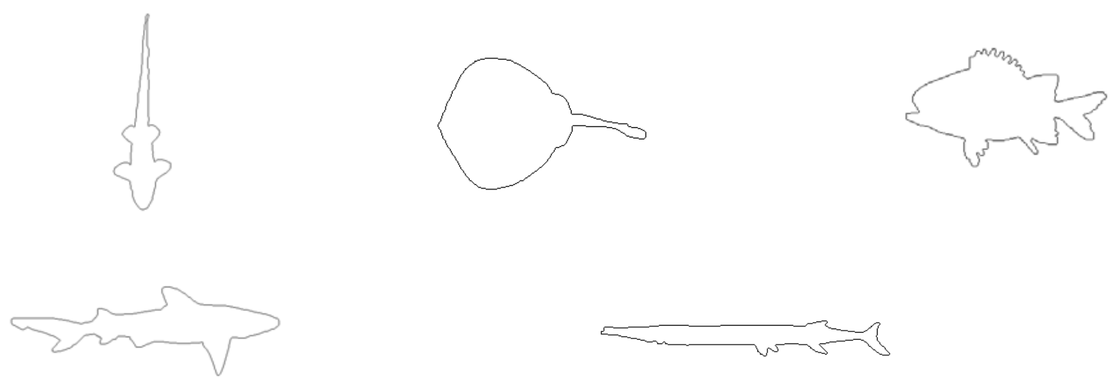

Figure 7 The sample index fish in test 1

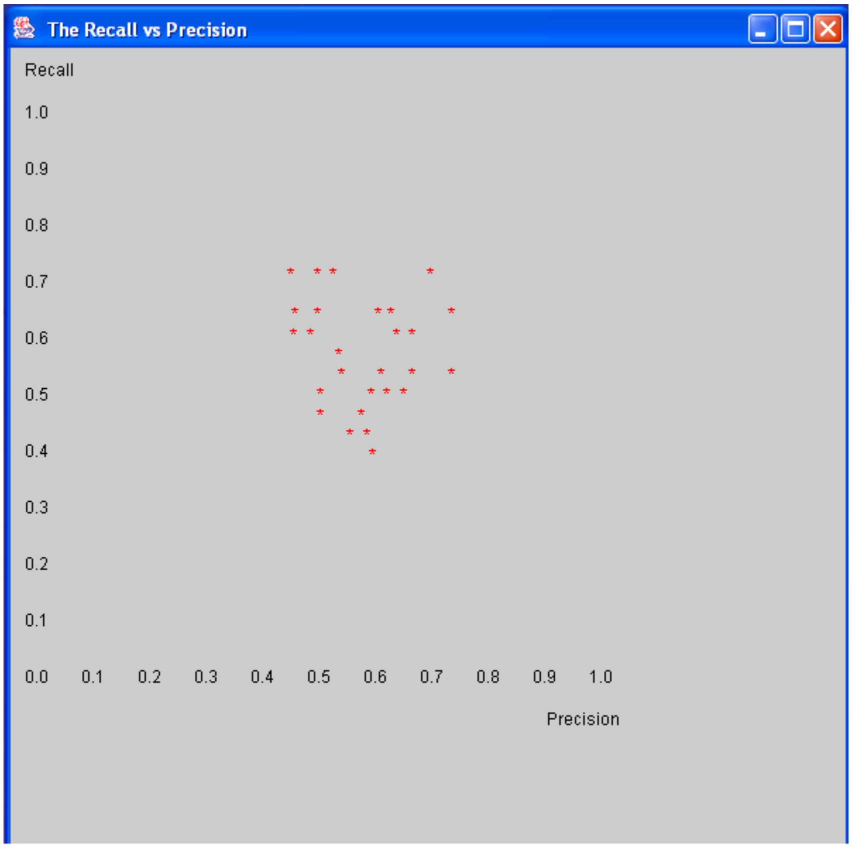

Figure 8 Precision-recall plots for test 1

Like all retrieval systems, we compute the recall-precision. Precision and recall are standard measures for Information Retrieval systems. Given a query Q: R is the answer set the system has retrieved; C is the right answers in R; A is the set of the right answers for the query. Precision and recall measures are defined as follows:

$$\text{Precision} = C/R$$

$$\text{Recall} = C/A$$

In one of our tests, we use sample fish shown in Figure 7. For each index fish, there are a certain number of fish which are similar to it. The number of the similar fish for each sample fish are different. For example, there are 6 fish which are similar to the fist index fish while there are 4 for the second index fish. Also we tried different (N,T), which was obtained from the training set while we test the algorithm. Each time, we compute the average recall and precision under a certain (N, T), that is to say, for the group of the 5 index fish, the total number of the fish retrieved is used in computing the recall rate and precision rate. What appears in the plot of Figure 8 is the average recall and precision of the 5 index fish set. Ideally, we want the curve/point in the recall-precision plot to be as "high" as possible. In Figure 8, both the recall and the precision are between 0.4-0.8.

Figure 9 shows the sample fish in another experiment. The set is made up of 110 fish. The performance is worse than the above one since the index fish looks quite similar and it becomes harder to distinguish them.

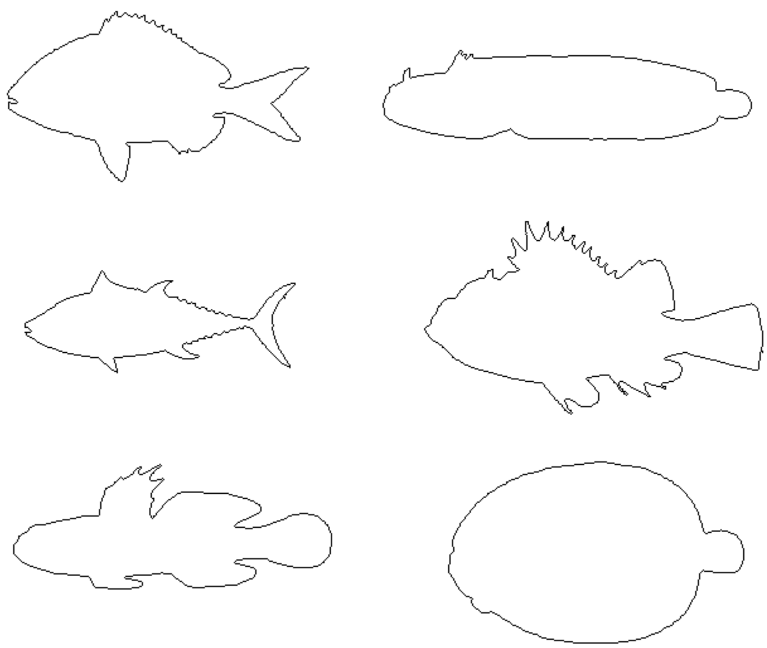

Figure 9 The index fish used to retrieve in test 2

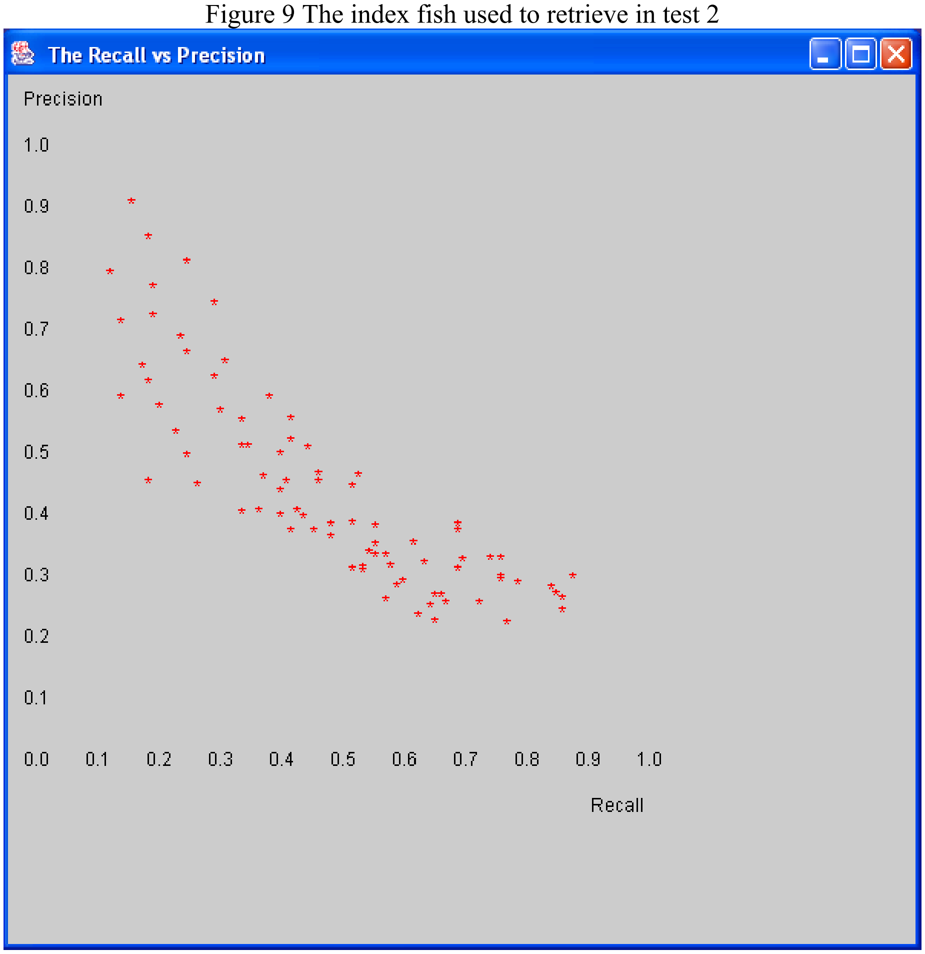

Figure 10 The recall-precision for test 2

4 Conclusion

In this paper, a shape retrieval algorithm is proposed. The algorithm is based on the geometry similarity theorem and Centroid-radii model. The distance ratio distribution is used to measure the similarity of the shapes. From another important point of view, a transformation is assumed to exist between any two shapes. If the transformation is uniform, the shapes can be regarded as similar and they have same origin. Otherwise if the transform is non-uniform, they shall not be classified in a same catalog. Further work is to test the algorithm with more shapes and with other algorithms to give comparison results.

Acknowledgement

The authors would like to thank Prof. Mokhtarian, Centre for Vision, Speech, and Signal Processing Laboratory, University of Surrey, U.K., for providing the SQUID database for our experiments.

References

[1] E. Binaghi, I. Gagliardi, and R. Schettini, "Image retrieval using fuzzy evaluation of color similarity," *Int. J. Pattern Recognit. Artif. Intell.*, vol.8, no. 4, pp. 945–968, May 1994.

[2] A. Del Bimbo, M. Mugnaini, P. Pala, and F. Turco, "Visual querying by color perceptive regions," *Pattern Recognit.*, vol. 31, pp. 1241–1253, 1998.

[3] C. Faloutsos, M. Flickner, W. Niblack, D. Petkovic, W. Equitz, and R.Barber, "The Qbic Project: Efficient and Effective Querying by Image Content," IBM Res. Div. Almaden Res. Center, Res. Rep. 9453, Aug. 1993.

[4] A. Del Bimbo and P. Pala, "Visual image retrieval by elastic matching of user sketches," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 121–132, Feb. 1997.

[5] I.-J. Lin, A. Vetro, H. Sun, and S.-Y. Kung, "Shape representation and comparison using voronoi order skeletons and dag-ordered trees," presented at the Int. Workshop on Very Low Bit-Rate Video Coding, Oct. 1999.

[6] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Tools for content-based manipulation of image databases," presented at the SPIE Storage and Retrieval for Image and Video Database, Feb. 1994.

[7] L. G. Shapiro and R. M. Haralick, "Decomposition of two-dimensional shapes by graph theoretic clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, pp. 10–20, 1979.

[8] L. G. Shapiro and R. M. Haralick, "Structural descriptions and inexact matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, no. 5, pp. 504–519, Oct. 1981.

[9] F. Liu and R. W. Picard, "Periodicity, directionality, and randomness —Wold features for image modeling and retrieval," *IEEE Trans.Pattern Anal. Machine Intell.*, vol. 18, pp. 722–733, July 1996.

[10] R. Picard, A society of models for video and image libraries, in MIT Media Lab Perceptual Computing Section Tech. Rep., Cambridge, MA, no. 360, 1995.

[11] H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 460–473, Apr. 1976. BERRETTI *et al.*: RETRIEVAL BY SHAPE SIMILARITY 239

[12] S. K. Chang, Q. Y. Shi, and C.W. Yan, "Iconic indexing by 2-d strings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 413–427, July 1987.

[13] M. J. Egenhofer and R. Franzosa, "Point-set topological spatial relations," *Int. J. Geogr. Inform. Syst.*, vol. 9, no. 2, 1992.

[14] H. Samet, "Hierarchical representations of collections of small rectangles," *ACM Comput. Surv.*, vol. 20, no. 4, pp. 271–309, Dec. 1988.

[15] S. Umeyama, "Parameterized point pattern matching and its application to recognition of object families," *IEEE Transaction on pattern Analysis and Machine Intelligence*, 15(1):136-144,1993.

[16] D.H.Ballard, "Generalized Hough transform to detect arbitrary patterns", *IEEE Transaction on pattern Analysis and Machine Intelligence*, 13(2): 111-122,1981.

[17] G.Stockman, "Object recognition and localization via pose clustering", *Computer vision, Graphics, and Image Processing*, 40(3):361-387,1987.

[18] H.Wolfson and I. Rigoutsos, "Geometric hashing; an overview", *IEEE Computational Sciences & Engineering* , pages 10-21, October-December 1997.

[19] D. Huttenlocher and S Ullman, “Object recognition using alignment”, *Proceeding of international conference on computer vision*, London, pages 102-111,1987
[20] S. Sclaroff and A. P. Pentland, “Modal matching for correspondence and recognition”, *IEEE Transaction on pattern Analysis and Machine Intelligence*, 17(6) June 1995
[21] S. Ranade and A Rosenfeld, “Point pattern matching by relaxation”, *Pattern recognition*, 12:269-275,1980.
[22] S. Loncaric, “A survey of shape analysis techniques”, *Pattern recognition*, 31(8):983-1001,1998.
[23] C. Jacobs, A. Finkelstein, and D. Salesin, “Fast multi-resolution image querying”, *Computer Graphics Proceeding SIGGRAPH*, page 277-286,1995.
[24] F. Mokhtarian, S. Abbasi, and J. Kittler, “Efficient and robust retrieval by shape content through curvature scale space”, *In image databases and multi media search, proceeding of the first international workshop IDB-MMS’96*, Amsterdam, the Netherlands, pages 35-42,1996.
[25] S. Gold, “Matching and learning structural and spatial representation with neural networks”, *PHD thesis*, Yale University, 1995
[26] 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.
[27] 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.
[28] K.L. Tan, B.C. Ooi and L.F. Thiang, "Indexing Shapes in Image Databases Using the Centriod-Radii Model", *Data and Knowledge Engineering*, Elsevier Science Publishing Inc, North-Holland, Vol 32, pp. 271-289, 2000.
[29] http://www.efg2.com/Lab/Graphics/PolygonArea.htm, June 2001.
[30] Shuang Fan, “Shape Representation and Retrieval Using Distance Histograms”, *Technical Report TR 01-14*,department of computer science, University of Alberta, Edmonton, Alberta, Canada, October 2001.
[31] http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html
[32] Evangelos Milios and Euripides G. M. Petrakis “Shape Retrieval Based on Dynamic Programming”, *IEEE Transaction on image processing*, Vol 9, No 1, pp 141-147, Jan 2001.
[33] S. Belongie, J. Malik, J. Puzicha, “Shape matching and object recognition using shape context”, *IEEE Trans. PAMI* April 2002, 24(4), pp. 509-522.