This is the published version


Available from Deakin Research Online

http://hdl.handle.net/10536/DRO/DU:30005398

Reproduced with the kind permission of the copyright owner

Copyright: 2005, The Author
Coherence Based Histograms for Shape Retrieval

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

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

Abstract

Histograms have been used for Shape Representation and Retrieval. The drawback of the histograms method is that histograms can be same for dissimilar shapes, which renders the method less effective for retrieval of shapes. In this paper, we describe the concept of coherence. We show how coherence can be used with distance and angular histograms. We perform experiments to test the effectiveness of the proposed method. It is found that coherence improves accuracy of retrieval significantly.

1 Introduction

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

Approaches for shape representation and retrieval can be broadly classified into contour based and region based. Some of the region based methods are moments invariants [3], and grid based method [4]. Some of the contour based methods are polygonal approximation [5], autoregressive model [6], Fourier Descrptors [7][8][9] and chain code [10].

Recently, a contour based method for shape representation and retrieval was used by Fan [1]. The method used by Fan is based on distance histograms. Angular histograms have also been used for shape representation and retrieval [13]. We propose the concept of coherence which can be used to modify distance histograms and angular histograms. Experiments are performed to test the effectiveness of the proposed method.

Distance histograms and angular histograms for shape representation and retrieval are described in Sections 2 and 3. The proposed methods which are based on spatial coherence are described in Sections 4 and 5. Experimental Setup and Results are presented in Section 6. Section 7 is the Conclusion.

2 Distance Histograms

Fan [1] has proposed a method of using distance histograms for shape representation and retrieval. In this method, points are sampled along the shape boundary and their distances are computed from the centroid. The centroidal distances obtained are discretised into buckets. The resulting histograms are used for shape representation and retrieval. A histogram is represented as below.

\[ D : (d_0, d_1 \ldots d_{N-1}) \] (1)

where, \( N \) is the number of buckets in the histograms and \( d_i \) is the number of centroidal distances, which were discretised into bucket \( i \).

The distance between two shapes is measured as the Euclidean distance below.

\[ \text{Dist} (D_1, D_2) = \sqrt{\sum_{i=0}^{N-1} (d_{1i} - d_{2i})^2} \] (2)
where, N is the number of buckets for each shape, \(d_i\) is the count in bucket \(i\) for histogram \(D_1\) and \(d_{ij}\) is the count in bucket \(i\) for histogram \(D_2\).

### 3 Angular Histograms

Angular histograms have been used for shape representation [13]. Points are sampled along the shape boundary; the centroid is computed from the sample points as shown below.

\[
x_c = \frac{\sum_{i=0}^{N-1} x_i}{N}, \quad y_c = \frac{\sum_{i=0}^{N-1} y_i}{N}
\]

(3)

The major-axis is defined as the line obtained by joining the centroid to the sample point on the shape boundary which is farthest from the centroid. Gaussian smoothing of the contour is used for robustness of the major-axis.

Angle for each sample point is computed as the angle between the x-axis and the line joining the centroid and the sample point. The angles thus obtained are invariant to scale and translation. However, the angles need to be normalised for rotation. To normalise the angles for rotation, we rotate the shape by angle \(\theta\) so that the major-axis aligns with the x-axis, where, \(\theta\) is computed as below.

\[
\theta = \arctan \left( \frac{y_m - y_c}{x_m - x_c} \right)
\]

(4)

where, \((x_m, y_m)\) is an extremity of the major axis on the shape boundary and \((x_c, y_c)\) is the centroid. Normalisation of a shape for rotation is illustrated in the Fig. 1.

![Fig. 1. Rotation normalization](image)

The angles made by sample points on the normalized shape boundary are discretised into buckets. Rotating the shape by angle \(\theta\) about the centroid will move all points \((x, y)\) on the shape boundary to \((x', y')\) as shown below.

\[
x' = x \cos \theta - y \sin \theta
\]

(5)

\[
y' = x \sin \theta + y \cos \theta
\]

(6)

### 4 Distance Coherence

There is no spatial information about the centroidal distances in the method described in Section 2. Hence, two shapes, which are entirely different, may have the same distance histograms.

We propose to incorporate spatial information when discretising the centroidal distances into a histogram. An analogy is drawn from Color Coherence Vectors (CCV) proposed by Pass and Zabih [10]. CCV is used for image retrieval based on colour. Pass et al [10] defined colour coherence of pixels as the degree to which pixels of that colour are members of a large similarly coloured region. Pixels are classified as coherent or incoherent. Coherent pixels are part of a sizable contiguous region of similar colour while incoherent pixels are not.

In the case of shape representation, we define “distance coherence” of pixels on the shape boundary. If the centroidal distance of a sample point on the shape boundary is part of a segment in which all the sample points have the same discretised centroidal distance then it is coherent. On the other hand, if the discretised centroidal distance of a sample point is different from the discretised centroidal distances of it’s immediate neighbours then it is incoherent. In other words, if the centroidal distance of a sample point belongs to the same bucket as it’s neighbours then it is coherent otherwise it is incoherent. The number of consecutive sample points, which should belong to the same bucket for the comprising sample points to be coherent is given by a variable \(\tau\). Hence, if \(\tau\) consecutive sample points have centroidal distances belonging to the same bucket then these pixels will be coherent.

In the method described in Section 2, each bucket in the histogram has one count. This count represents the number of pixels, which had their centroidal distances discretised into that bucket. In the proposed method, pixels in each bucket are classified into coherent and incoherent. Hence, there will be two counts for each bucket. One count will represent the
coherent pixels and the other will represent the incoherent pixels.
During the indexing process, N equi-spaced points are sampled along the shape boundary. The coordinates of the centroid \((x_c, y_c)\) for a shape are computed from Eqn. 3. The distance between a sample point \(s(x_v, y_v)\) and the centroid \((x_c, y_c)\) is computed as below.

\[
d(s_c, c) = \sqrt{(x_c - x_v)^2 + (y_c - y_v)^2} \tag{7}
\]

The distances thus obtained are invariant to rotation and translation. However, the distance needs to be normalised for scale. To normalise the distance for scale, we scale them from 0 to 100. Thus, the minimum distance will be 0 and the maximum will be 100. The normalisation is done as shown in Eqn 8 below.

\[
norm\_dist = \frac{\text{dist} - \text{dist}_{\text{min}}}{\text{dist}_{\text{max}} - \text{dist}_{\text{min}}} \times 100 \tag{8}
\]

where, \(\text{norm\_dist}\) is the normalised centroidal distance for \(\text{dist}\). \(\text{dist}_{\text{max}}\) is the maximum centroidal distance and \(\text{dist}_{\text{min}}\) is the minimum centroidal distance.

The normalised centroidal distances of the pixels are discretised into buckets. The pixels in each bucket are divided into coherent and incoherent. For instance, bucket “i” will have \(\alpha_i\) coherent pixels and \(\beta_i\) incoherent pixels. The total number of pixels in the bucket is \((\alpha_i + \beta_i)\). A histogram for a shape is represented as follows.

\[
D: [(\alpha_0, \beta_0), \ldots, (\alpha_{N-1}, \beta_{N-1})] \tag{9}
\]

where, \(N\) is the number of buckets in the histogram. Fig. 2 shows distance histograms for a shape contour with 100 equi-spaced sample points along the boundary. Fig. 2b shows the breakdown of the count in each bucket into coherent and incoherent categories. The distance between the two shapes represented by histograms \(D_1, D_2\) is computed as below.

\[
\text{Dist}(D_1, D_2) = \sum_{i=0}^{N-1} [\alpha_i - \alpha_i] + |\beta_i - \beta_i| \tag{10}
\]

The distance measure in Eqn. 10 computes the distance between “like-buckets”. Quadratic distance is a better distance measure because it takes into account the similarity across buckets [15]. The quadratic-form distance between the two feature vectors Q and T is given by Eqn. 11.

\[
d_{\text{quad}}(Q, T) = \left\| (Q - T)^T A(Q - T) \right\|^{1/2} \tag{11}
\]

where, \(A = [a_{ij}]\) is an \(N \times N\) matrix, and \(a_{ij}\) is the similarity coefficient between indexes (dimensions) \(i\) and \(j\). \(a_{ij}\) is given by Eqn. 12.

\[
a_{ij} = 1 - d_{ij} / d_{\text{max}} \quad \text{and} \quad d_{ij} = |Q_i - T_j| \tag{12}
\]

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

\[
d_{\text{quad}}(Q, T) = \left( \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{ij}Q_iQ_j + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{ij}T_iT_j + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} QT_j \right)
\]

**5 Angular Coherence**

There is no spatial information about the angles made by sample points on the shape boundary in the method described in Section 3. Hence, two shapes, which are entirely different, may have similar angular histograms.

We propose to incorporate spatial information when discretising the angles into a histogram. An
analogy is drawn from “distance coherence” which is described in Section 4. We define “angular coherence” of pixels on the shape boundary. If the angle of a sample point on the shape boundary is part of a segment in which all the sample points have the same discretised angle then it is coherent. On the other hand, if the discretised angle of a sample point is different from the discretised angle of its immediate neighbours then it is incoherent. The number of consecutive sample points, which should belong to the same bucket for the comprising sample points to be coherent is given by a variable \( \tau \). Hence, if \( \tau \) consecutive sample points have angles belonging to the same bucket then these pixels will be coherent.

In the method described in Section 3, each bucket in the histogram has one count. This count represents the number of pixels, which had their angles, discretised into that bucket. In the proposed method, we classify the pixels in each bucket into coherent and incoherent. Hence, there will be two counts for each bucket. One count will represent the coherent pixels and the other will represent the incoherent pixels.

During the indexing process, \( N \) equi-spaced points are sampled along the shape boundary. The angle made by the line joining the sample point \( s(x, y) \) and the centroid \( c(x, y) \) with the \( x \)-axis is computed from Eqn. 4. The count in each bucket is divided into coherent and incoherent. For instance, bucket “1” will have \( \theta_1 \) coherent pixels and \( \phi_1 \) incoherent pixels. The total number of pixels in the bucket is \( (\theta_1 + \phi_1) \). A histogram for a shape is represented as follows.

\[
D : [(\theta_0, \phi_0), (\theta_1, \phi_1), \ldots (\theta_{N-1}, \phi_{N-1})] \quad (13)
\]

where, \( N \) is the number of buckets in the histogram. The distance between two shapes represented by histograms \( D_1, D_2 \) is computed as the quadratic distance between the histograms as shown in Section 4.

6 Experimental Setup and Results

The SQUID database [11] is used to perform experiments and test the proposed method. This database consists of 1100 fish shapes and has been extensively used by researchers for testing. A system is built, using Java for indexing the shapes and performing queries. Indexing and Retrieval experiments are performed on the database for the traditional method and the proposed method.

Queries are performed using the traditional method and the proposed method. The shapes in the database, which are perceptually similar to the query shapes are noted. Ranks of relevant shapes for each query are obtained using the traditional approach and the proposed approach. The change in the rank of the relevant shapes retrieved is used to compare the effectiveness of the proposed method with the traditional method.

For distance histograms, 150 sample points are used along the shape boundary. Centroidal distances are computed for the sample points. The distances are then scaled for each shape such that all distances are in the range 0 to 100. 10 buckets are created which are populated for distances 0 to 10, 10 to 20, 20 to 30 and so on.

![Fig. 3. Query Shapes](image)

During the process of indexing, histograms are constructed for each shape. The buckets in the histograms consist of a single count for the traditional method. On the other hand, each bucket in the histograms consists of two counts for the proposed method. Queries are performed for \( \tau = 1, \tau = 2, \text{and} \tau = 3 \). \( \tau = 1 \) corresponds to the traditional method because all the pixels will be regarded as coherent for this case.

The shapes in the database, which are perceptually similar to the query shapes are noted. Six queries are made for each value of \( \tau \). The six query shapes are shown in Fig. 3. Recall vs. Precision values obtained for the six queries are averaged. The average values are shown in Fig. 4.

Another set of experiments are performed to test the angular histograms method. 150 points are sampled along the shape boundary. The angles are discretised into 20 buckets. \( \tau \) is defined to be 4 pixels. The average values of recall vs. precision are shown in Fig. 5.
The experimental results of distance histograms are shown in Fig. 4. We see from the results that the effectiveness for $\tau = 2$ is almost the same as for $\tau = 1$. However, the effectiveness improves significantly for $\tau = 3$. This is explained by the fact that $\tau = 2$ does not contribute significantly to the spatial information in the distance histograms. $\tau = 3$ contributes significantly to the spatial information, thus improving the effectiveness of retrieval.

The experimental results from angular histograms are shown in Fig. 5. There is significant improvement when $\tau$ is increased from 1 to 4. $\tau = 1$ represents the traditional method. $\tau = 4$ represents the proposed method based on angular histograms.

It is observed from experimental results that using (spatial and angular) coherence in histograms improves the accuracy of retrieval significantly. The computation expense for obtaining coherence is $O(n)$ which may be outweighed by the improvement in accuracy.

7 Conclusion

The experimental results of distance histograms are shown in Fig. 4. We see from the results that the effectiveness for $\tau = 2$ is almost the same as for $\tau = 1$. However, the effectiveness improves significantly for $\tau = 3$. This is explained by the fact that $\tau = 2$ does not contribute significantly to the spatial information in the distance histograms. $\tau = 3$ contributes significantly to the spatial information, thus improving the effectiveness of retrieval.

The experimental results from angular histograms are shown in Fig. 5. There is significant improvement when $\tau$ is increased from 1 to 4. $\tau = 1$ represents the traditional method. $\tau = 4$ represents the proposed method based on angular histograms.

It is observed from experimental results that using (spatial and angular) coherence in histograms improves the accuracy of retrieval significantly. The computation expense for obtaining coherence is $O(n)$ which may be outweighed by the improvement in accuracy.

References

12. SQUID: www.ee.surrey.ac.uk/Research/VSSP/imagedb