5

Chord Context Algorithm for Shape Feature Extraction

Yang Mingqiang[1], Kpalma Kidiyo[2] and Ronsin Joseph[2]
[1]*ISE, Shandong University, 250100, Jinan*
[2]*Université Européenne de Bretagne,*
France - INSA, IETR, UMR 6164, F-35708 RENNES
[1]*China*
[2]*France*

1. Introduction

The emergence of new technologies makes it easy to generate information in visual forms, leading everyday to an increasing number of generated digital images. At the same time, the rapid advances in imaging technologies and the widespread availability of Internet access motivate data browsing into these data bases. For image description and retrieval, manual annotation of these images becomes impractical and inefficient. Image retrieval is based on observation of an ordering of match scores obtained by searching through a database. The key challenges in building a retrieval system are the choice of attributes, their representations, query specification methods, match metrics and indexing strategies.
A large number of retrieval methods using shape descriptors has been described in literature. Compared to other features, for example, color or texture, object shape is unique. It enables us to recognize an object without further information. However, since shapes are 2D images that are projections of 3D objects, the silhouettes may change from one viewpoint to another with respect to objects and non-rigid object motion (e.g., walking people or flying bird) and segmentation errors caused by lighting variations, partial occultation, scaling, boundary distortion and corruption by noise are unavoidable. As we know, while computers can easily distinguish slight differences between similar objects, it is very difficult to estimate the similarity between two objects as perceived by human beings, even when considering very simple objects. This is because human perception is not a mere interpretation of a retinal patch, but an active interaction between the retinal patch and a representation of our knowledge about objects. Thus the problem is complicated by the fact that a shape does not have a mathematical definition that exactly matches what the user feels as a shape. Solutions proposed in the literature use various approaches and emphasize different aspects of the problem. The choice of a particular representation scheme is usually driven by the need to cope with requirements such as robustness against noise, stability with respect to minor distortions, and invariance to common geometrical transforms or tolerance to occultation, etc. For general shape representation, a recent review is given in [1] [2].
In this chapter, a shape descriptor based on chord context is proposed. The basic idea of chord context is to observe the lengths of all parallel and equidistant chords in a shape, and

to build their histogram in each direction. The sequence of vector features extracted forms the feature matrix for a shape descriptor. Because all the viewpoint directions, considered with a certain angle interval, are chosen to produce the chord length histogram, this representation is unlike conventional shape representation schemes, where a shape descriptor has to correspond to key points such as maxima of curvature or inflection points, for example, Smooth Curve Decomposition [3], Convex Hull [4], Triangle-area representation (TAR) [5] and Curvature Scale Space (CSS) [6][7] etc. The proposed method needs no special landmarks or key points. There is also no need for certain axes of a shape. The proposed descriptor scheme is able to capture the internal details, specifically holes, in addition to capturing the external boundary details. A similarity measure is defined over chord context according to its characteristics and it confirms efficiency for shape retrieval from a database. This method is shown to be invariant under image transformations, rotations, scaling and robust to non-rigid deformations, occultation and boundary perturbations by noise thus it is well-adapted to shape description and retrieval. In addition, the size of the descriptor attribute is not very great; it has low-computational complexity compared to other similar methods.

2. Chord context

This section details the proposed method, chord context, for extracting attributes from the contour or silhouette of a shape. It then proposes a method of measuring similarities between two shapes.

2.1 Feature extraction

Chord context analysis corresponds to finding the distribution of all chord lengths in different directions in a given shape. For discrete binary image data, we consider each object point as one and the background as zero. In the shape recognition field, it is common to consider the case where the general function $f(x, y)$ is

$$f(x,y)=\begin{cases}1 & \text{if } f(x,y)\in D\\ 0 & otherwise,\end{cases}$$

where D is the domain of the binary shape.

In each direction, we can find all the chords in the shape. Fig. 1 shows an example of chords in direction θ.

A set of lines $T(\rho,\theta)$ is defined by

$$\rho = x\cos(\theta-\pi/2)+y\sin(\theta-\pi/2)\ ,\ \theta\in[0,\ \pi],\ \text{and}\ \rho\in(-\infty,\ \infty).$$

The chords are defined by the parts of these lines within the domain of the binary shape. So a shape can be represented by a discrete set of chords sampled from its silhouette. Considering different angles θ, the number and length of chords obtained in different directions may not be the same, except in the case of a circle. One way to capture this information is to use the distribution of chord lengths in the same direction in a spatial histogram.

Concretely, let us assume that the set of chords in directions θ_i are represented by $C=\left\{c_{i,n}\,\middle|\, n\in[1,N_i]\right\}$, where N_i is the number of the chords in direction θ_i. Let $L(c_{i,n})$ be the length of chord $c_{i,n}$. So we can compute a histogram h_i in direction θ_i by

$$h_i^l = \#\left\{L(c_{i,n}) \in bin(l)\right\} \quad l \in [1,\ L_{i,\ \max}] \tag{1}$$

where $L_{i,max}$ is the longest chord in direction θ_i.

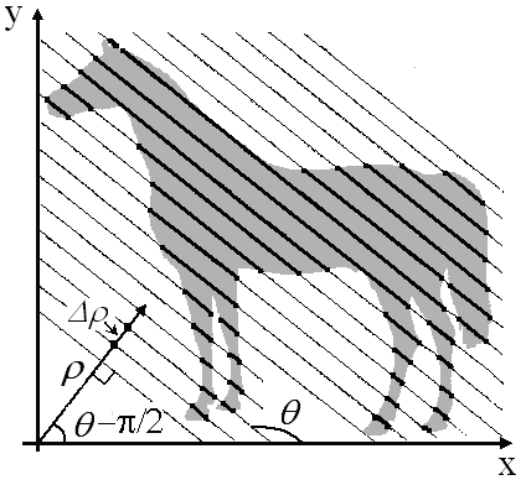


Fig. 1. Representation of chords in direction θ with the interval $\Delta\rho$. The bold lines are the chords of the shape.

In order to capture the details of a shape, the interval $\Delta\rho$ of ρ, i.e. the distance between two parallel chords, should not be great. In practice, $\Delta\rho=L_{max}/(50\sim100)$, where L_{max} is the length of the longest axis of the shape. The histogram h_i of Fig. 1 in direction θ_i is shown in Fig. 2.

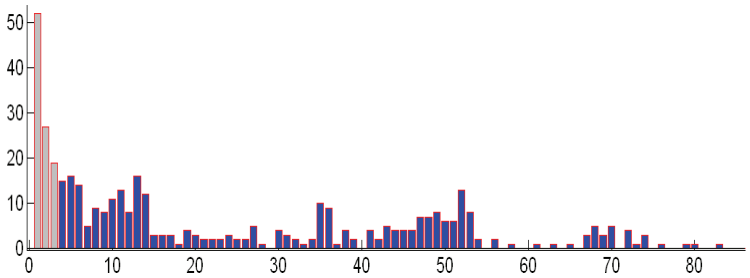


Fig. 2. Histogram of chord lengths in direction θ_i for the shape shown in Fig. 1

An excessive number of too short chords is counted when line T is close to a tangent along the edge of the shape (see Fig. 3). This is because a scraggy edge is produced by the minor disturbances resulting from digitization noise or normalization of the image to a certain size. In fact, these uncertain short chords are harmful to our shape descriptor, so we remove these too short chords directly: this could be seen as a low-pass filtering of the shape contour. Empirical tests show that, if we normalize a shape in an image with 128×128 pixels, i.e. the largest size of the shape is 128 pixels, and the shorter size transforms in proportion, then we can consider the set of chords whose length is shorter than 4 to be too short chords. So they should be discarded. In Fig. 2, the first 3 bins, plotted in gray, should be removed.

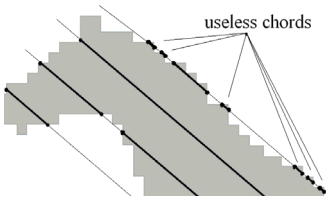


Fig. 3. Illustration of producing very short chords

With θ increasing from 0 to 179 degrees, all the chords in different directions in the silhouette can be recorded. If we divide the orientation range [0, 179] into D', then we can obtain D' histograms h_i, $i \in [0, D'-1]$. They can form a matrix M arranged by a set of histograms with column vector h_i according to the order of angles:

$$M = \left[\mathbf{h_0}, \mathbf{h_1}, \cdots \mathbf{h_{D'-1}}\right]$$

The matrix M of Fig. 1 is shown in Fig. 4.

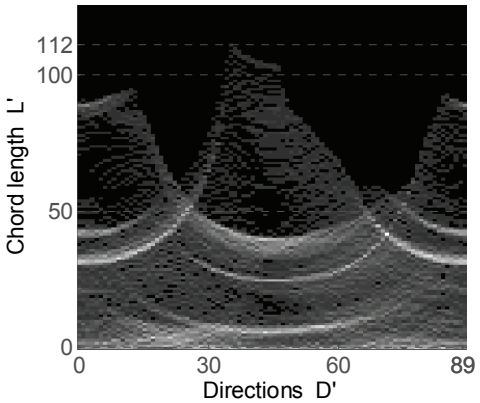


Fig. 4. The matrix M of the shape in Fig. 1 with all orientations.

The matrix element is the number of equal-length chords whose direction and length are given by the value of abscissa and y-axis, respectively. The abscissa is the orientation angle θ, and the y-axis is the length of the chords. The value in each row of the matrix M is the number of the chords with same length in different directions; and each column is the chord length histogram in the same direction.

Due to its very great size, it is unreasonable to use this matrix directly as a shape attribute. For example, in an image with 128×128 pixels, the longest possible chord in the shape is $128 \times \sqrt{2} \approx 181$. So, if D'=90, the size of the matrix will be 181×90=16290. Clearly, it is not appropriate as a direct feature of a shape.

In order to reduce the size of the matrix M and, at the same time, make the extracted feature invariant to scale transforms, we normalize this matrix M as follows:

First, find the maximum of non-zero bin L' for all the histograms in the matrix M. In Fig. 4 for example, L'_{max}=112. Then remove all the bins that are greater than L'_{max}, and form a matrix M' with dimension $L'_{max} \times D'$:

$$M' = \left[\mathbf{h}'_0, \mathbf{h}'_1, \cdots \mathbf{h}'_{D'-1}\right]$$

And then, for the next normalization, we expand the matrix M' to matrix M'', using a wrap-around effect, so that to eliminate border effects:

$$M'' = \left[\mathbf{h}'_{D'-2}, \mathbf{h}'_{D'-1}, \mathbf{h}'_0, \mathbf{h}'_1, \cdots \mathbf{h}'_{D'-1}, \mathbf{h}'_0, \mathbf{h}'_1\right]$$

Finally, the matrix M'' is subsampled down to a new matrix F with the dimension $L \times D$, after a 4×4 bicubic interpolation algorithm. For convenience, D is even. The bicubic interpolation algorithm means that the interpolated surface is continuous everywhere and also continuous in the first derivative in all directions. Thus, the rate of change in the value is continuous. Each value of matrix F contains a synthesis of its 4×4 neighbouring point values. The feature matrix F can be represented by:

$$F = \left[\mathbf{f}_0, \mathbf{f}_1, \cdots \mathbf{f}_{D-1}\right]$$

where $\mathbf{f}_i$, $i \in [0, D-1]$, is a L dimensions column vector given by $\mathbf{f}_i = [f_{i,0}, f_{i,1}, \cdots f_{i,L-1}]^T$.

The feature matrix F is the attribute of a shape. We call this feature matrix F of a shape the "chord context" descriptor.

For L=30 and D=36, the chord context of Fig. 1 is shown in Fig. 5.

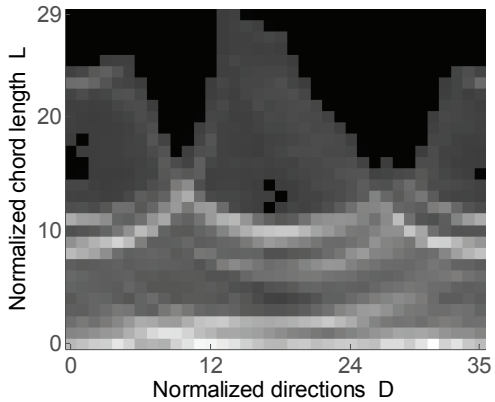


Fig. 5. Chord context of Fig. 1 with 30 rows and 36 columns.

The experiment in section 3 shows that chord context as the feature of a shape can retain the visual invariance to some extent.

2.2 Similarity measure

In determining the correspondence between shapes, we aim to meet the distance between two feature matrices. Such matching combines two criteria: one is the calculation of the minimum value of the distances between histograms of two feature matrices, *e.i.* the *Character Matrix Distance (CMD),* and the other is a comparison of the *Perpendicular Chord Length Eccentricity (PCLE).*

In the first criterion, we calculate all the distances between the query feature matrix and the model feature matrix while shifting its histograms one by one. Similar shapes have similar

histograms in a same direction, and the rearranged order of these histograms is also similar. To calculate the distance between two attributes of shapes, we first calculate the distance between each corresponding histogram, according to their arrangement orders, and, then calculate the sum of all these distance values. Regarding rotation invariance, we shift the model feature matrix by one histogram, i.e. change the direction used to obtain the histograms, and repeat the same step to calculate the sum of all the values of these distances between the two feature matrices.

We assume that the query feature matrix is F_Q and the model feature matrix is F_M. F_Q and F_M are given by

$$F_q = \left[\mathbf{fq_0}, \mathbf{fq_1}, \cdots \mathbf{fq_{D-1}}\right] \text{ and } F_m = \left[\mathbf{fm_0}, \mathbf{fm_1}, \cdots \mathbf{fm_{D-1}}\right]$$

According to subsection 2.1, $\boldsymbol{f\alpha_i}$, where α is q or m, $i \in [0, D-1]$, is an L dimensions column vector $\mathbf{f\alpha_i} = [f\alpha_{i,0}, f\alpha_{i,1}, \cdots f\alpha_{i,L-1}]^T$.

So the set of similarity distance is given by

$$DistF_{F_q,F_m}(n) = \sum_{i=0}^{D-1} DistH(\mathbf{fq_i}, \mathbf{fm_j}),\ \ j = \mathrm{mod}(i+n, D) \tag{2}$$

where $n \in [0, D-1]$ is the number of shifts applied to each histogram in the model feature matrix. The formula shows that the set of similarity distances is the sum of the distances *DistH* between the two corresponding normalized histograms in two feature matrices.

To quantify the similarity between two histograms, there are many methods being reported: Minkowski-form, Kullback-Leibler Divergence, Jeffrey Divergence, Quadratic-form, Earth Mover's Distance, χ^2 statistics, Hausdorff distance, etc. Because of the properties of the chord context histogram:

- they have the same number of bins.
- the value in each bin has great variances; some of values are even zeros, cf. Fig. 2.

We compare χ^2 statistics distance

$$DistH_{\chi^2}(\mathbf{fq_i}, \mathbf{fm_j}) = \frac{1}{2L}\sum_{k=0}^{L-1}\frac{\left(fq_{i,k} - fm_{j,k}\right)^2}{\left(fq_{i,k} + fm_{j,k}\right)} \tag{3}$$

and our proposed distance formula defined here by

$$DistH(\mathbf{fq_i}, \mathbf{fm_j}) = \begin{cases} 0, & fq_{i,k} = 0 \text{ and } fm_{i,k} = 0, \forall k \in [1, L-1] \\ \dfrac{1}{L}\displaystyle\sum_{k=0}^{L-1}\frac{\left|fq_{i,k} - fm_{j,k}\right|}{\max(fq_{i,k}, fm_{j,k})}, & otherwise \end{cases} \tag{4}$$

on the database of Kimia silhouettes with 216 shapes [8] (see section 3.4). For convenience, we consider the minimum value of the distance set as the similarity distance and call it the *Character Matrix Distance (CMD).*

$$CMD_{F_q,F_m} = \min_{n=0}^{D-1} DistF_{F_q,F_m}(n) \tag{5}$$

The comparison result of precision vs. recall is shown in Fig. 6. Precision is the ratio of the number of relevant shapes retrieved to the total number of retrieved shapes, while recall is the ratio of the number of relevant shapes retrieved to the total number of relevant shapes in the database.

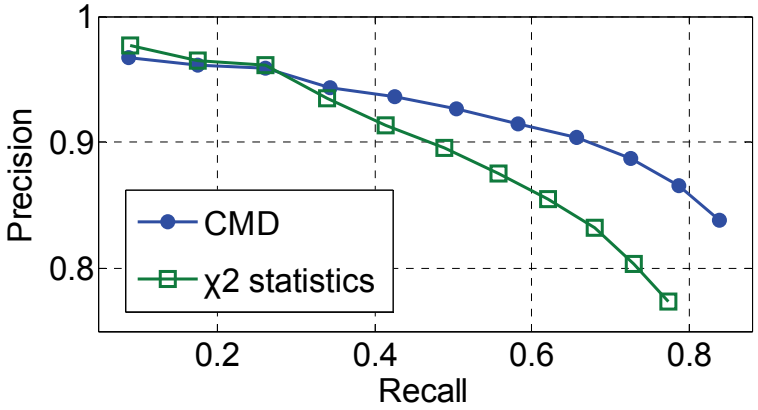


Fig. 6. The precision-recall diagrams for indexing into the database of Kimia silhouettes with 216 shapes.

It is clear from Fig. 6, that our proposed distance formula is better than χ^2 statistics on this similarity measure.

In the second criterion, we propose a new concept, the *Perpendicular Chord Length Eccentricity (PCLE)* as following:

$$PCLE(i)=\begin{cases}\left\|\mathbf{f}_i-\mathbf{f}_{i+(\mathbf{D/2})}\right\| & i\in[0,\frac{D}{2}-1]\\ \left\|\mathbf{f}_i-\mathbf{f}_{i-(\mathbf{D/2})}\right\| & i\in[\frac{D}{2},D-1]\end{cases} \tag{6}$$

Where $\|\bullet\|$ is the norm; $\mathbf{f_i}$, $i\in[0,D-1]$ (D is even) are vectors of the feature matrix. So (6) is the Euclidean distance between any two histograms of perpendicular directions of chord lengths. Since the norm is symmetric, we have

$$PCLE(i+\frac{D}{2})=PCLE(i),\quad i\in[0,\ \frac{D}{2}-1] \tag{7}$$

Clearly, *PCLE* represents the perpendicular directions chord feature in a shape. To compare the query's *PCLE* P_q and the model's *PCLE* P_m, we define the distance between P_q and P_m as follows:

$$D\_PCLE_{P_q,P_m}(n)=\sum_{i=0}^{D-1}\frac{\left|P_q(i)-P_m(j)\right|}{\max(P_q(i),P_m(j))},\ j=\mathrm{mod}(i+n,D) \tag{8}$$

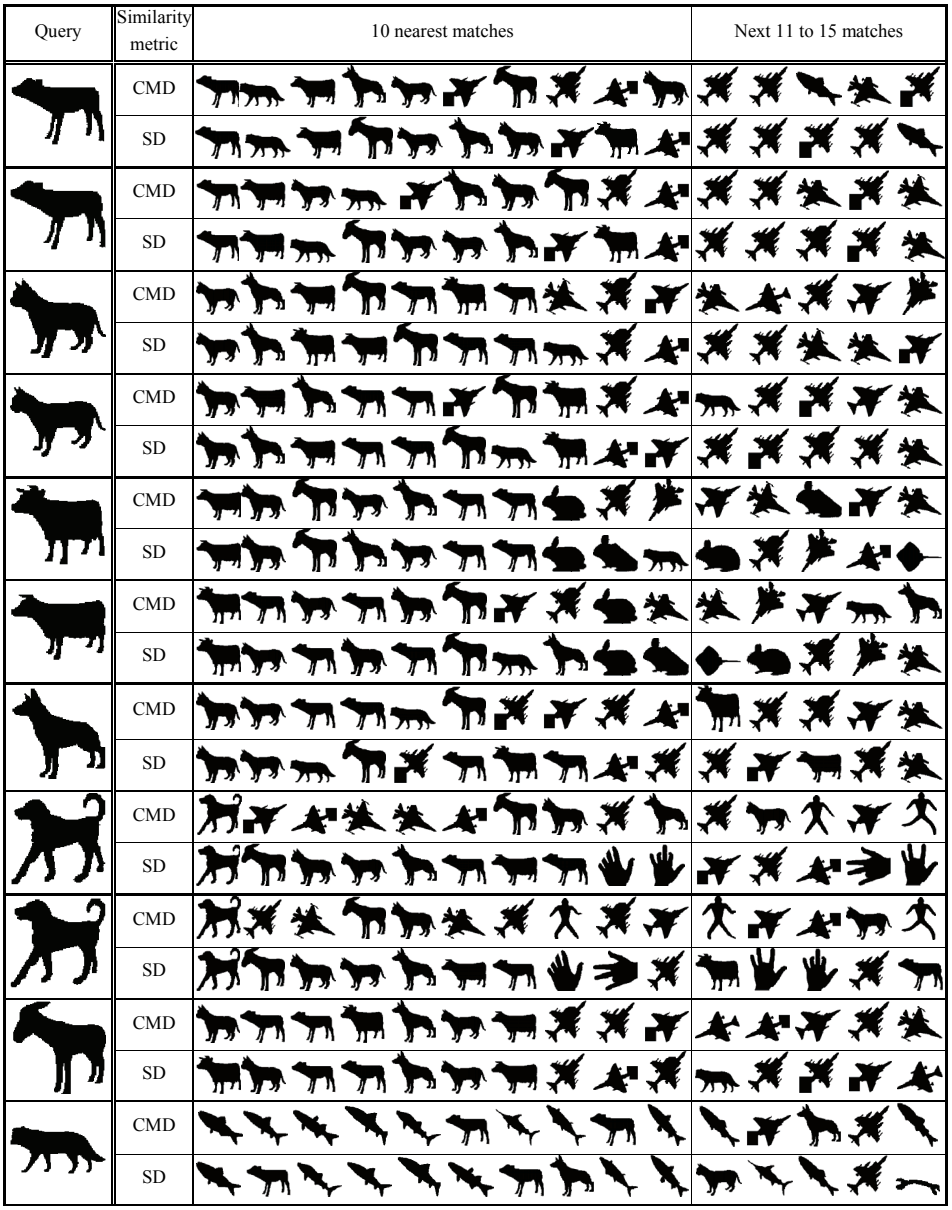


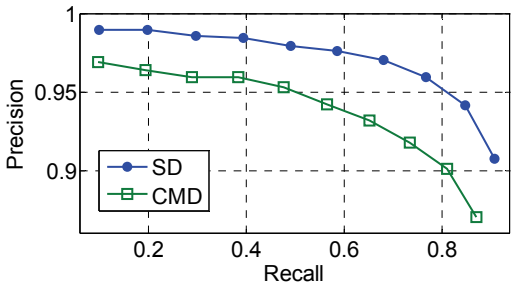
Fig. 7. Illustration of retrieval results from the 'Quadruped' category in Kimia silhouettes set of 99 shapes. The 11 quadruped shape queries in the dataset are shown in the first column. The 10 nearest retrieved shapes for each query are shown in order (from small similarity distance to large similarity distance) in the 3rd column by their similarity metric *CMD* and SD. The next five matches are shown in the 4th column for completeness.

As in (2), n is the number of shifts applied to each histogram in the *PCLE*.
Intuitively, we assume that, in general, if shape S_1 is more similar to shape S_2 than shape S_3, then the smallest value of $D\_PCLE_{P_{S_1},P_{S_2}}$ is less than the smallest value of $D\_PCLE_{P_{S_1},P_{S_3}}$.

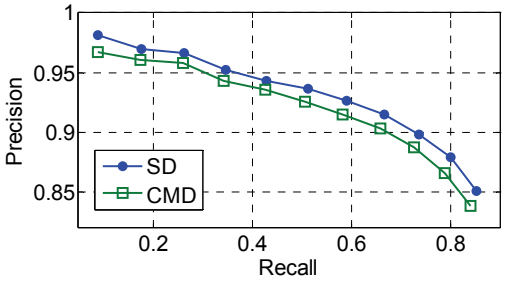
So we can use *D_PCLE* to adjust the similarity distance of the *Character Matrix Distance (CMD)* to improve retrieval precision and recall. The combined similarity metric of shape Query and shape Model is computed using a weighted sum:

$$SD_{Q,M} = \alpha CMD_{F_q,F_m} + (1-\alpha)\min_{n=0}^{D-1}(D\_PCLE_{P_q,P_m}(n)) \tag{9}$$

where $\alpha \in [0, 1]$.
Let us explain this hypothesis by the following experiments, where we have used α=0.85. First let us look at the experiment running on the Kimia silhouettes set of 99 shapes [8]. Fig. 7 shows the two retrieval results from querying the 'Quadruped' category. One shows the retrieval results only with the similarity distance of *Character Matrix Distance (CMD)* and the other with the combined similarity metric weighted sum *SD*. As there are 11 shapes in the 'Quadruped' category, each shape is matched against all the other shapes in the database. As there are 11 shapes in each category, up to 10 nearest neighbors can be retrieved from the same category. We count in the *n*th (*n* from 1 to 15) nearest neighbors the number of times that the test image is correctly classified. The best possible result is 110 matches (except the query itself) in all 10 nearest matches. With the similarity metric *CMD*, we found 63 matches



(a)



(b)

Fig. 8. The precision-recall diagrams for indexing into the database of Kimia silhouettes with (a) 99 shapes and (b) 216 shapes.

in the first ten retrieved shapes, recall is 63/110=57.3%, and there were 70 matches in the first fifteen retrieved shapes. Whereas with the similarity metric *SD*, we found 80 matches in the first ten retrieved shapes, recall is 80/110=72.7%, and there were 85 matches in the first fifteen retrieved shapes. The result shows that the recall rate in the first ten retrieved shapes was improved by 15.4 percentage points for the 'Quadruped' cluster when we used the similarity metric *SD* instead of *CMD*. It also shows a good performance rate when compared with [9], [10] and [11] which have the same retrieval results of 51 matches in the first ten retrieved shapes.

Let us look at the statistical results. We compare the retrieval results using Character Matrix Distance (*CMD*) to the retrieval results when using the similarity metric *SD* by calculating precision vs. recall in the Kimia silhouette datasets of 99 and 216 shapes [8]. The results are shown in Fig. 8. It is evident that similarity metric *SD* outperforms *CMD*.

3. Experimental evaluations of chord context matching

In this section, we present the results obtained during a more realistic shape recognition process in presence of different possible visual deformations to study the comparative performances of the proposed algorithm. We show that the chord context matching is effective in the presence of commonly occurring visual transformations like scale changes, boundary perturbations, viewpoint variation, non-rigid transform and partial occultation. We also compare its results with ten other well-known algorithms. All the experiments are conducted on the standard database: MPEG-7 CE-shape-1 database (1400 shapes) [12], Columbia University Image Library Coil-100 database (7200 images) [13], and 3 databases of Kimia silhouettes [8]. In all the experiments, the feature matrix was normalized to 30 bins and 36 directions; the similarity measure uses formula (9) with α=0.85 : these values are found to be the most efficient, during various experiments.

3.1 Scale and rotation transforms

Scale and rotation transforms are the important intuitive correspondences for a variety of shapes. They can be regarded as a necessary condition that every shape descriptor should satisfy. In order to study retrieval performance in terms of scale changes and image rotations, we use the test-sets Part A, from CE-Shape-1 database that was defined during the standardization process of MPEG-7, consisting of 1400 shapes semantically classified into 70 classes [12].

For robustness to scaling during the test, i.e. Part A-1, we created a database in which there were 70 basic shapes taken from the 70 different classes and 5 shapes derived from each basic shape by scaling digital images with factors 2, 0.3, 0.25, 0.2, and 0.1. Thus in the database, there were 420 shapes. Each of the 420 images was used as a query image. A number of correct matches were computed in the top 6 retrieved images. Thus, the best possible result was 2520 matches.

For robustness to rotation during the test, i.e. Part A-2, we again created a database including 420 shapes. The 70 basic shapes were the same as in part A-1 and 5 shapes were derived from each basic shape by rotation with angles: 9, 36, 45, 90 and 150 degrees. As in Part A-1, each of the 420 images was used as a query image. The best result was 2520 matches.

The similarity rate in each experiment was calculated by taking the ratio of correct matches to the maximum number of possible matches. Table 1 indicates the similarity rate of

comparison of the chord context descriptor with the reported results of certain studies. Note that the proposed descriptor has the best performance in all the experiments.

Data Set	Tangent Space [14]	Curvature Scale Space [15]	Zernike Moments [16]	Wavelet [17]	BAS [18]	Chord context
Part A-1	88.65	89.76	92.54	88.04	90.87	**99.37**
Part A-2	100	99.37	99.60	97.46	100	**100**
Part A	94.33	94.57	96.07	92.75	95.44	**99.69**

Table 1. Comparison of the Retrieval Results of Different Methods in the MPEG-7 CE-Shape-1 Part A Test

The results show that chord context is very invariant to scale and rotation transforms. Reviewing the extracted attribute algorithm in section 2, we are not surprised by the almost perfect results. The attribute matrix is obtained by the statistic of all the chord lengths of a shape in all directions. The rotation of the shape affects the attribute matrix only when shifting the chord length histograms. In the similarity measure, we have considered this point and compared the two attribute matrices by shifting the histograms of either matrix. The rotation of plane shapes does not affect the retrieval result. Since we normalize all the images to a certain size before extracting their features, the scale transform of a shape does not significantly affect the retrieval result.

3.2 Boundary perturbations by noise

The query shape can be perturbed by different noises. It may simply result from digitization. As a reminder, to fight perturbations resulting from shape digitization, and in order to alleviate the influence of boundary perturbation, we have removed the very short chords from attribute matrices. To evaluate the performance of chord context when boundary perturbations are present, we use noisy images with different noise powers as queries to retrieve the relevant image in a database. We generated a 20 sub-database test-set based on MPEG-7 CE-Shape-1. In each sub-database, there were 70 shapes from the 70 different classes according to their orders in the database. The query shapes were all 70 shapes in each sub-database subjected to noise with 4 different noise powers. Thus, the best possible result in each sub-database was 70 matches for each noise power. Suppose the average distance of all the points on the edge of a shape to its centroid is *D*. We then define the signal-to-noise ratio (*SNR)* as follow:

$$\mathrm{SNR} = 20\lg \frac{\mathrm{D}}{\mathrm{r}}(\mathrm{dB})$$

where *r* is the largest deviation of the points on the edge. Fig. 9 shows an example of an original shape and its contaminated shapes produced by random uniform noise with *SNR* equal to 30dB, 25dB, 20dB and 15dB.

Table 2 shows the average similarity rates of the 20 sub-databases at 4 *SNRs*.

SNR(dB)	30	25	20	15
Average Similarity Rate (%)	99.9	99.8	99.1	82.9

Table 2. Average Similarity Rates of the 20 Sub-Databases at *SNR* of 30dB, 25dB, 20dB and 15dB

Fig. 9. Examples of noisy shapes from a model in MPEG-7 CE-Shape-1. From left to right, the original and the resulting noised shapes at an *SNR* of 30dB, 25dB, 20dB and 15dB.

From the results in Table 2, we notice that using noise to impair boundaries does not produce significant differences between similar shapes. As the chord context descriptor utilizes the edge as well as the region feature of a shape, it can bear boundary disturbances due to noise to certain extent.

3.3 Partial occultation

In general, a global descriptor is not robust when a shape is partially occulted. Since the chord context descriptor has the statistical information for a shape, this drawback is alleviated to some extent. To evaluate robustness to partial occultation we ran experiments using the same sub-databases as mentioned in subsection 3.2. We occulted the shapes, applying 4 different percentages of occultation from the left, Fig. 10(a), or right, Fig. 10(b), respectively to them, in raster-scan order. The occultation percentages were 5%, 10%, 15% and 20%. Each occulted shape is retrieved in its sub-database as a query. Thus, the best result in each sub-database is 70 matches for one occultation.

(a) (b)

Fig. 10. Examples of occulted shapes from a model in MPEG-7 CE-Shape-1. (a) From left-occulted objects; they are occulted by 5%, 10%, 15% and 20%. (b) From right-occulted objects; they are occulted by 5%, 10%, 15% and 20%.

Table 3 shows the average similarity rate of the 20 sub-databases on the 4 partial examples of occultation.

Occultation (%)		5	10	15	20
Average Similarity Rate (%)	Left	99.5	91.3	78.6	58.7
	Right	98.9	94.1	80.7	64.6
	Average	99.2	92.7	79.7	61.7

Table 3. Average similarity rates of the 20 sub-databases at shape occultation of 5%, 10%, 15% and 20%

It is clear from Table 3, that small occultations do not affect chord context significantly. However, the problem of significant occultation remains to be explored. The results show that chord context is robust to minor occultation.

3.4 Similarity-based evaluation

The performance in similarity-based retrieval is perhaps the most important of all tests performed. In order to demonstrate the performance of chord context when deformed parts are present, we turn to three shape databases. All three databases were provided by Kimia's group [8] [19].

The first database is Kimia's data set 1 which contains 25 images from 6 categories (Fig. 11).

Fig. 11. Kimia's data set 1: 25 instances from six categories. Each row shows instances of a different object category.

This property has been tested by shape contexts [20], Sharvit et. al [19], Gdalyahu et. al [21] and Ling et. al [22]. The retrieval results are summarized as the number of first, second, and third closest matches that fall into the correct category. The results are listed in Table 4. It shows that the proposed method outperforms the first 3 reported methods. For the fourth approach, chord context is slightly better than it in the Top 2 closest matches.

Method	Top 1	Top 2	Top 3
Sharvit et. al [19]	23/25	21/25	20/25
Gdalyahu et. al [21]	25/25	21/25	19/25
Belongie et. al [20]	25/25	24/25	22/25
Ling et. al [22]	25/25	24/25	25/25
Chord context	**25/25**	**25/25**	**23/25**

Table 4. Comparison of the Retrieval Results of Different Methods on the Kimia Data Set 1 (Fig. 11)

The second database contains 99 images from nine categories with 11 shapes in each category [8]. It has been tested by D.S. Guru [10], Shape contexts [20], Bernier [9] and Tabbone [11]. Each shape was used as a query to which all other shapes were compared and thus 9801 shape comparisons were made. Ideal results would be that the 10 closest matches (except the query itself) belong to the same category. The results are summarized by precision-recall diagrams in Fig. 12. The proposed method shows better precision and recall rate than the other methods.

The third database contains 216 images from 18 categories with 12 shapes in each category [8]. All the shapes were selected from the MPEG-7 test database [8]. It has been tested by shape contexts [20]. As in the case of second database, a comparison of the results of our

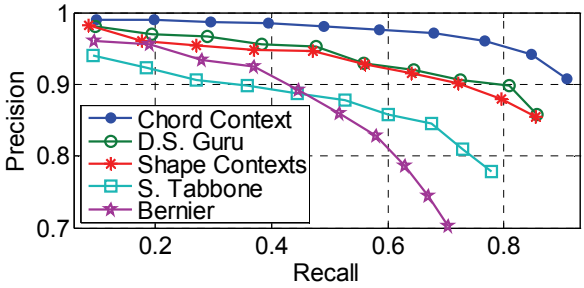


Fig. 12. Comparison of precision and recall rates of different methods on the Kimia Data Set of 99 shapes.

approach to the shape context method is given in Fig. 13. As we see in Fig. 13, the two precision/recall curves cross. This means that the shape contexts [20] method performs better for small answer sets, while our proposed method performs better for larger answer sets. According to [24], the method achieving highest precision and recall for large answer sets is considered to be the best one, so the proposed method is better than shape contexts' method.

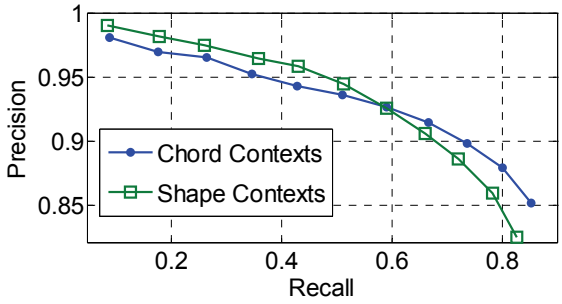


Fig. 13. Comparison of precision and recall rates of different methods on the Kimia Data Set of 216 shapes.

From the above results, it appears that chord context produces outstanding performance in the presence of non-rigid deformations.

3.5 Viewpoint variations

For a better evaluation in the realistic context of image retrieval with industrial vision where the picture of an object from real world is observed from different viewpoints, we have performed tests on shapes extracted directly from these pictures. To test the retrieval performance of the proposed method in the presence of viewpoint changes, the Columbia University Image Library Coil-100 3D object dataset is used. This dataset contains 7200 color images of 100 household objects and toys. Each object was placed on a turntable and an image was taken for every 5 degrees of rotation, resulting in 72 images per object. We converted the color images to grayscale, and then into shapes by using the same gray-value

threshold settings for the whole set (e.g., Fig. 14). Since the shapes are projections of 3D objects by a simple gray-value threshold, not only may their silhouettes change from one viewpoint to another; the lighting and object textures may also differ.

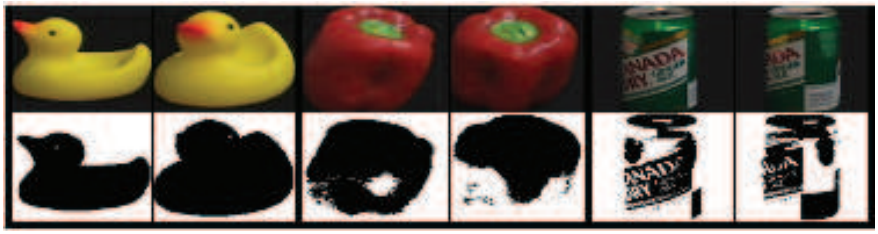


Fig. 14. 6 images taken as an example from the COIL-100 3D object dataset. The first row shows the original images and the second row shows their corresponding silhouettes as produced by gray-value thresholding. In 1st and 2nd columns, the duck's silhouettes change significantly due to a change of viewpoint. In the 3rd and 4th columns, the red pepper's silhouettes change significantly due to a difference in lighting. In 5th and 6th columns, the tin's silhouettes change significantly due to a change in textures.

In the following subsections we present two experiments showing the performance of the proposed method on these test sets. First we compare the new approach to the shape contexts [23]. We converted the color images into shapes, then selected 3 images per object with a 15° viewpoint interval, for example 0°, 15° and 30°. To measure performance, we counted the number of times the closest match was a rotated view of the same object. Our result was 285/300. The result reported in [23] is 280/300.
In the second experiment, we generated 7 sub-databases in which we selected 5 to 17 consecutive viewpoints per object for a total from 500 to 1700 images respectively (cf. Table 5). We use the middle view images in a sub-dataset as a query to retrieve in them. For each query, the number of the best possible matches is 5 to 17 in the 7 sub-databases, respectively. The results are shown in Table 5 and Fig. 15.

	Models in Sub-database (view angle)	Query (view angle)	The best retrieved number	Retrieved results	Retrieved precision
Sub-database1	0, 5, **10**, 15, 20	10	500	**490**	**98.0%**
Sub-database2	0, 5, 10, **15**, 20, 25, 30	15	700	**679**	**97.0%**
Sub-database3	0, 5, 10, 15, **20**, 25, 30, 35, 40	20	900	**827**	**91.9%**
Sub-database4	0, 5, 10, 15, 20, **25**, 30, 35, 40, 45, 50	25	1100	**958**	**87.1%**
Sub-database5	0, 5, 10, 15, 20, 25, **30**, 35, 40, 45, 50,55, 60	30	1300	**1062**	**81.7%**
Sub-database6	0, 5, 10, 15, 20, 25, 30, **35**, 40, 45, 50,55, 60, 65, 70	35	1500	**1149**	**76.6%**
Sub-database7	0,5, 10, 15, 20, 25, 30, 35, **40**, 45, 50,55,60,65, 70,75,80	40	1700	**1211**	**71.2%**

Table 5. The Components of 7 Sub-Databases with Different Viewpoints on Coil-100 3D Object Dataset and Their Retrieval Results

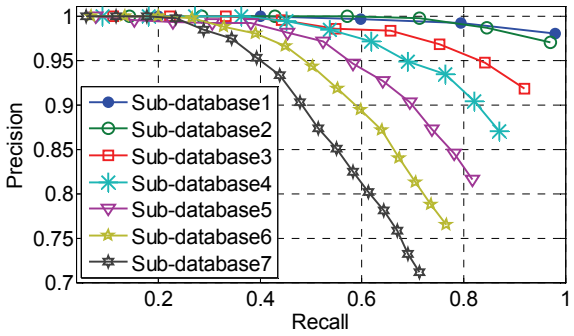


Fig. 15. Precision and recall rates in the 7 sub-databases on Coil-100 3D object dataset.

These results are very encouraging, since they indicate that we can perform satisfactory retrieval with mean average precision of more than 91% for view angle differences of under 20°: see the results of Sub-database1-3. For viewpoint difference of less than 40°, the retrieved precision is more than 71%. Note that this is done exclusively on shape images (without using any intensity information). Clearly, if other information and a more specialized feature set were used, even higher precision scores could be achieved.

4. Conclusions

We have presented a new approach which is simple and easy to apply in the context of shape recognition. This study has two major contributions: (1) defining a new algorithm which can capture the main feature of a shape, from either its contour or its region; (2) proposing an assistant similarity measure algorithm *Perpendicular Chord Length Eccentricity (PCLE)* which can help to improve retrieval precision and recall to some extent.

In various experiments we have demonstrated the invariance of the proposed approach to several common image transforms, such as scaling, rotation, boundary perturbations, minor partial occultation, non-rigid deformations and 3D rotations of real-world objects.

The particular strengths of the proposed descriptor for retrieving images are summarized in the following points:

- **Flexibility:** Chord context can handle various types of 2D queries, even if it has holes or separates itself into several parts. It is robust to noise and minor occultation. So we can use this simple method to segment objects from images.
- **Accuracy:** the proposed method has the advantage of achieving higher retrieval accuracy than the other methods in the literature based on MPEG-7 CE-1 database, Coil-100 database and the Kimia silhouettes datasets retrieval test.

Currently we have not considered the effect of an affine transform of a shape. The chord context method has no special operations that resist affine transforms. We consider this to be the main weakness of this approach and, to achieve more accuracy and have more applications; further work will be carried out on invariance to affine transforms.

5. References

[1] R.C. Veltkamp, M. Hagedoorn, "State of the Art in Shape Matching," *Principles of Visual Information Retrieval,* 2001, pp. 89-119.

[2] D. Zhang, G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.,* vol. 37, 2004, pp. 1–19.

[3] S.Berretti, A.D.Bimbo, and P.Pala, "Retrieval by shape similarity with perceptual distance and effective indexing," *IEEE Trans. Multimedia*, vol. 2, no. 4, 2000, pp. 225–239.

[4] O. El Badawy, M. Kamel, "Shape Retrieval using Concavity Trees," *Proceedings of the 17th International Conference on Pattern Recognition,* vol. 3, 2004, pp. 111-114.

[5] N. Alajlan, Mohamed S. Kamel and George Freeman, "Multi-object image retrieval based on shape and topology," *Signal Processing: Image Communication,* vol. 21, 2006, pp. 904–918.

[6] F. Mokhtarian, A. K. Mackworth, "A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 8, 1992, pp. 789–805.

[7] F. Mokhtarian, Abbasi S and Kittler J, "Robust and efficient shape indexing through curvature scale space," *Proceedings British Machine Vision Conference, Edinburgh, UK,* 1996, pp. 53-62.

[8] T.B. Sebastian, P.N. Klein and B.B. Kimia, "Recognition of Shapes by Editing Their Shock Graphs," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 5, 2004, pp. 550-571.

[9] T. Bernier, J.-A. Landry, "A new method for representing and matching shapes of natural objects," Pattern Recognit., vol. 36, issue 8, 2003, pp. 1711-1723.

[10] D.S. Guru, H.S. Nagendraswamy, "Symbolic representation of two-dimensional shapes," *Pattern Recognit. Letters,* vol. 28, 2007, pp. 144–155.

[11] S. Tabbone, L. Wendling and J.-P. Salmon, "A new shape descriptor defined on the Radon transform," *Computer Vision and Image Understanding,* vol. 102, issue 1, 2006, pp. 42-51.

[12] S. Jeannin, M. Bober, "Description of core experiments for MPEG-7 motion/shape," *MPEG-7, ISO/IEC JTC1/SC29/WG11/MPEG99/N2690,* Seoul, March 1999.

[13] Dataset available on the website: http://www1.cs.columbia.edu/CAVE/software/softlib/coil-100.php.

[14] L. J. Latecki, R. Lakamper, "Shape Similarity Measure Based on Correspondence of Visual Parts," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 10, 2000, pp. 1185-1190.

[15] 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, 1996, pp. 35-42.

[16] A. Khotanzan, Y. H. Hong, "Invariant Image Recognition By Zernike Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, 1990, pp. 489-497.

[17] G. C.-H. Chuang, C.-C. J. Kuo, "Wavelet Descriptor of Planar Curves: Theory and Applications," *IEEE Trans. Image Processing*, vol. 5, no. 1, 1996, pp. 56-70.

[18] N. Arica, F. T. Yarman-Vural, "BAS: a perceptual shape descriptor based on the beam angle statistics Source," *Pattern Recognit. Letters, v*ol. 24, issue 9-10, 2003, pp. 1627-1639.

[19] D. Sharvit, J. Chan, H. Tek and B. Kimia, "Symmetry-Based Indexing of Image Database," *J. Visual Comm. and Image Representation,* vol. 9, no. 4, 1998, pp. 366-380.

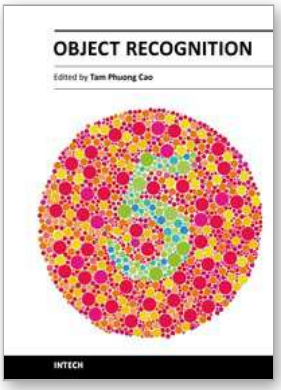
[20] S. Belongie, J. Malik and J. Puzicha, "Shape Matching and Object Recognition Using Shape Context," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 4, 2002, pp. 509-522.

[21] Y. Gdalyahu, D. Weinshall, "Flexible Syntactic Matching of Curves and Its Application to Automatic Hierarchical Classification of Silhouettes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 12, 1999, pp. 1312-1328.

[22] Haibin Ling, David W. Jacobs, "Shape Classification Using the Inner-Distance," *IEEE Trans. On Pattern Analysis and Machine Intelligence,* 2007, vol. 29, no. 2, pp. 286-299.

[23] S. Belongie, J. Malik, "Matching with shape contexts," Content-based Access of Image and Video Libraries, Proceedings IEEE, 2000, pp. 20-26.

[24] E. Petrakis, A. Diplaros, and E. Milios, "Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, 2002, pp. 1501-1516.



Object Recognition

Edited by Dr. Tam Phuong Cao

ISBN 978-953-307-222-7
Hard cover, 350 pages
Publisher InTech
Published online 01, April, 2011
Published in print edition April, 2011

Vision-based object recognition tasks are very familiar in our everyday activities, such as driving our car in the correct lane. We do these tasks effortlessly in real-time. In the last decades, with the advancement of computer technology, researchers and application developers are trying to mimic the humanâ€™s capability of visually recognising. Such capability will allow machine to free human from boring or dangerous jobs.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yang Mingqiang, Kpalma Kidiyo and Ronsin Joseph (2011). Chord Context Algorithm for Shape Feature Extraction, Object Recognition, Dr. Tam Phuong Cao (Ed.), ISBN: 978-953-307-222-7, InTech, Available from: http://www.intechopen.com/books/object-recognition/chord-context-algorithm-for-shape-feature-extraction

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821