Shape Feature Representation In Partial Object Matching

K. L. Chiew
Faculty of Information Technology, University of Malaysia Sarawak,
Jalan Dato’ Musa, 93400 Kota Samarahan, Sarawak, Malaysia
klchiew@fit.unimas.my

Y. C. Wang
Faculty of Information Technology, University of Malaysia Sarawak,
Jalan Dato’ Musa, 93400 Kota Samarahan, Sarawak, Malaysia
ycwang@fit.unimas.my

Abstract—The main objective of this research is to develop an image matching prototype that can retrieve images based on the geometry feature representation of a complete object or a partial object. Consequently, a shape decomposition method is introduced to preserve the shape feature for the remaining portion of a partial object and to simplify the shape complexity by breaking down into segments.

I. INTRODUCTION

The emergence of digital library and multimedia database require an efficient and effective maintenance technique. As stated in [1], a lot of difficulties and limitations arise from the rapid-growing size of these databases. Therefore, content-based image retrieval techniques have been developed to overcome some of the limitations associated with the conventional method. The main objective of this research is to develop new feature extraction and representation techniques to overcome the shortcomings that have been encountered especially in partial object matching.

This paper is organized as follows. In Section II, the state of the art will be briefly discussed. Whereas, the proposed methodology will be highlighted in detail in Section III. It is followed by the experimental result in Section IV and finally, Section V will offer a concluding remark.

II. STATE OF THE ART

Curvature receives a lot of attention from researchers because of its local shape features and it is believed that local shape features are more suitable for representing partial object shape as described in [2] because local shape features can describe object properties locally. A lot of attempts on using curvature can be seen in current research trends such as [3, 4, 5, 6, 7]. Among these works, they share one thing in common that is the use of curvature of the shape though used differently in each of their works.

In [11], curvature is integrated in their shape similarity retrieval system. The authors utilized multiple scales instead of single scale because they can avoid of using too small single scale, which may lead to local distortion and noise tolerance. Furthermore, too big a single scale may cause the loss of some important shape structure. In their research, the curvature of every point is obtained in the scale space form. From this curvature scale space, a graph known as CSS image is plotted from where the curvature zero crossing happens in the relative arc length position. After that, the local maximum of the CSS image peaks will be extracted as the final shape feature representation. However, CSS feature is inadequate to represent an occluded or partial object. Particularly, when the major concavities of the shape are occluded or missed then the contour in the CSS will disappear and result in a very different CSS representation.

Another similar approach can be found in [7], the authors developed an image retrieval system mainly based on the curvature of the shape. In this system, a shape is partitioned into a few tokens and the partitioning process is based on the curvature value. The authors have claimed that curvature plays an important role in shape feature. Hence, these partitioned tokens will be represented by local maximum curvature and the tokens’ angle. The performance of their system has proven to perform well and applicable to partial object retrieval, but unfortunately it is rotation variant, which means the rotation will change the spatial relationships of the object. The main reason of this drawback lies in the use of an absolute reference system for token orientation.

The next review is on decomposition, several techniques have been applied differently such as [4, 5, 7]. Among them, the technique which decomposes a shape at minimum curvature [7] is considered as the most significant technique. This is because a decomposed segment is bounded by the pair of minimum curvature points and therefore the high curvature points will be preserved. Consequently, since dominant information regarding shape is usually available at high curvature points and it is preserved, so it is sensible to conclude that this decomposition will yield an efficient feature representation which will be beneficial in later use.

III. METHODOLOGY

A. Image Interpretation and Image Smoothing

The input data is a binary image containing a single object. The file format for this input data is Portable Graymap format with the file extension named PGM. The header information of the image such as file type, the width and height of the image will be used to load pixel values of the image data bits and store these image pixel values into a two dimensional array. This array will form the object boundary.
In mathematical point of view, the object boundary is a planar continuous curve and can be parameterised as shown in (1).

\[ L(t) = (x(t), y(t)) \quad (1) \]

\( u \) in the equation is an arbitrary parameter. Since \( u \) is an arbitrary parameter, the curve will be parameterised according to its arc length, \( t \) as expressed in (2)

\[ c(t) = [(x(t), y(t)) ; t \in [0, \text{total length of curve} - 1]] \quad (2) \]

The input image may contain some noises due to the digitising process or other factor such as low resolution. This noise can be reduced by smoothing the image at a suitable scale. However, choosing the right scale is a crucial step because if the scale is too large, some important shape structure may be lost. On the other hand, if the scale used is too small, the noise still remains. In addition, using a fixed scale for all the images with different sizes is also not suitable.

Therefore, in order to overcome this problem, this research proposes the incorporation of a dynamic scale smoothing that is independent of image size. The dynamic scale, \( \sigma_0 \) is obtained by using the Equation (3). As determined by experiment, the scale ratio of 1:6 is used to produce the dynamic scale. The optimum scale is the maximum scale needed to evolve an object from the original shape to a fully convex shape with no curvature zero crossing.

\[ \text{dynamic scale, } \sigma_0 = \text{scale ratio} \times \text{optimum scale} \quad (3) \]

Curvature zero crossing is a point where the sign of curvature changes as shown in Fig. 1.

![Curvature Zero Crossings](image)

**Fig. 1. Curvature zero crossing**

The shape evolution in this research is the process of smoothing a shape using Gaussian convolution until the shape becomes fully convex. Since the shape has been parameterised in parametric vector as shown in (2), so the convolution can be carried out by convoluting each component in \( c(t) \) with the one dimensional Gaussian kernel of width \( \sigma \). Then \( X(t, \sigma) \) and \( Y(t, \sigma) \) will represent the resulting components of the \( c(t) \) after convolution as shown in (4) and (5).

\[ X(t, \sigma) = x(t) \Theta g(t, \sigma) \quad (4) \]

\[ Y(t, \sigma) = y(t) \Theta g(t, \sigma) \quad (5) \]

\( \Theta \) is the convolution process, \( g(t, \sigma) \) is one dimensional Gaussian kernel and \( \sigma \) is the kernel width (\( \sigma \) also refer to smoothing scale). The formula of Gaussian is shown in (6).

\[ g(t, \sigma) = \frac{1}{\sqrt{2\pi\sigma^3}} e^{-\frac{t^2}{2\sigma^2}} \quad (6) \]

After the dynamic scale (\( \sigma_0 \)) has been obtained, the final smoothing of an object will be performed based on this dynamic scale. The equivalent mathematical description for the resulting components of the \( c(t) \) is shown in (7) and (8).

\[ X(t, \sigma_0) = x(t) \Theta g(t, \sigma_0) \quad (7) \]

\[ Y(t, \sigma_0) = y(t) \Theta g(t, \sigma_0) \quad (8) \]

### B. Shape Feature Extraction and Representation

After smoothing the object, the next step is to extract the shape feature representation. Before the shape feature representation can be obtained, the curvature of the shape must be calculated. Since the shape of the object has been parameterised as parametric curve \( c(t) \) shown in (2), the curvature at every point on the curve can be calculated. The curvature function \( k(t) \) of \( c(t) \) at the point \( (x(t), y(t)) \) can be expressed as (9).

\[ k(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{(x'^2(t) + y'^2(t))^{3/2}} \quad (9) \]

\( x'(t) \) in (9) is the first derivative of component \( x(t) \) and \( y'(t) \) is the second derivative of component \( x(t) \). Whereas \( y'(t) \) and \( y''(t) \) is the first and second derivative of component \( y(t) \) respectively.

According to the properties of convolution, the derivatives of component \( x(t) \) is equivalent to the derivatives of Gaussian kernel convolute to component \( x(t) \). The first derivative and the second derivative of component \( x(t) \) are shown in (10) and (11) respectively. The same notation is applied to the component \( y(t) \) where its first derivative and the second derivative are shown in (12) and (13) respectively.

\[ X'(t, \sigma_0) = x(t) \Theta g'(t, \sigma_0) \quad (10) \]

\[ X''(t, \sigma_0) = x(t) \Theta g''(t, \sigma_0) \quad (11) \]

\[ Y'(t, \sigma_0) = y(t) \Theta g'(t, \sigma_0) \quad (12) \]

\[ Y''(t, \sigma_0) = y(t) \Theta g''(t, \sigma_0) \quad (13) \]

g'(t, \sigma) and g''(t, \sigma) is the first and second derivative of Gaussian kernel respectively. Since the exact forms of g'(t, \sigma) and g''(t, \sigma) are known and as shown in (14) and (15) respectively, the curvature function \( k(t, \sigma) \) can be calculated as in (16).

\[ g'(t, \sigma) = -\frac{-t}{\sqrt{2\pi}\sigma^3} e^{-\frac{t^2}{2\sigma^2}} \quad (14) \]

\[ g''(t, \sigma) = -\frac{1}{\sqrt{2\pi}\sigma^3} e^{-\frac{t^2}{2\sigma^2}} \left[ 1 - \frac{t^2}{\sigma^2} \right] \quad (15) \]

\[ k(t, \sigma) = \frac{X'(t, \sigma)Y''(t, \sigma) - X''(t, \sigma)Y'(t, \sigma)}{(X'(t, \sigma)^2 + Y'(t, \sigma)^2)^{3/2}} \quad (16) \]

The output of this process is a list of curvature that will be used to decompose the shape. The term “Shape Decomposition” used in this research means to break down or partition the shape of an object into smaller parts. This smaller part is called “token” in this research. Each of these tokens contains some significant shape features. Currently
there are a few methods that can be used to decompose a
shape, the method that is used in this research is similar to
the method used by [7], but with some modifications.

The decomposition criteria used in [7] is the minimum
curvature. Based on the minimum curvature, the process of
decomposition happens at the point where its curvature
value is minimum. The shape after decomposition will be
in the form of basic shape such as convex or concave.
However, this method cannot be used when the shape does
not have an obvious minimum curvature as illustrated in
Fig. 2.

Besides, this method also has the problem to determine
how minimum a curvature value that is considered as
minimum (Fig. 3). Sometime, in an image the curvature
value of -0.90 is considered as a significant minimum
curvature and the curvature value of -0.01 is not. However,
in some images the lowest minimum curvature may be
around -0.01, so in this case the curvature value of -0.01
may consider as a significant minimum curvature.
Therefore, using minimum curvature value as the shape
decomposition criteria will cause a lot of ambiguity and this
method is very subjective.

In order to overcome the insufficiencies described above,
this research proposes another shape decomposition criteria
that is based on the curvature zero crossing. Since the
curvature zero crossing (Fig. 1) is the pixel where the sign
of the curvature changes either from positive to negative or
negative to positive, so it can effectively solve the problem
as depicted in Fig. 2 and Fig. 3.

After the shape has been decomposed into tokens, all
these tokens will be represented by the proposed features.
The proposed features are token eccentricity, token local
maximum, token relative arc length and basic visual
components identification.

The concept of eccentricity in a shape is to measure the
ratio between its major axis and its minor axis. The
simplicity in calculation makes eccentricity a popular
method. The main reason the eccentricity is suitable for
representing token instead of representing the whole shape
is because a token is a basic or primitive shape which
consists relatively less detail alone the token boundary.
Therefore, by representing a token using its major and
minor axis is sufficient to describe its structure. On the
other hand, the whole shape of an object consists of
relatively more detail along its boundary since it is actually
formed by a list of tokens as depicted in Fig. 4.

As a result, by using token’s eccentricity, the degree of
the token elongation can be determined and the uniqueness
each token can be preserved. However, this research
found that in order to optimize the token eccentricity, the
identification of the major and minor axis has to be changed
to the identification of the height and width of the token.
This has ensured that narrow token is distinguished from
wide token. Equation (17) shows the calculation for token
eccentricity and Fig. 5 shows the example of a token which
is represented by the token eccentricity

\[
\text{Token Eccentricity} = \frac{\text{Height}}{\text{Width}}
\]  

Fig. 5. Token Eccentricity

Token local maximum is the second proposed shape
feature and is based on the sign of the local maximum
curvature. It is an important feature, since it helps to
represent the shape more uniquely and more meaningfully.
For example, given two tokens with the same eccentricity,
one with a convex shape and the other with a concave
shape, it is difficult to differentiate them without the sign.
A good shape feature should be invariant to the object's size. For instance, when an object is enlarged by 50 percent of its original size, their shape should still look alike. However, unlike object's size, the token's size is very important since the relative size of each token in an object is interrelated and describes the whole shape structure of an object. Fig. 6 shows two different objects, object A and object B. If the relative size in each token is ignored, then all the tokens in object A are the same as all the tokens in object B, and this will produce a result that object A is the same as object B.

![Tokens](image1)

**Fig. 6.** Without considering relative token size, the extracted tokens from object A and object B will be the same.

Therefore, this research shows that the incorporation of relative token size will ensure unique representation for every token in the object. Due to the ease of feature extraction and without loss of any important information in shape feature, the relative arc length is used to represent the token’s relative size. Arc length is the sum of all the distance between every adjacent pixel in a token and is calculated using the Euclidean distance as shown in (18). Later, among these arc lengths, the approximate mod arc length is determined. This approximate mod arc length is used as the denominator to compute the relative arc length for every token. The formula for computing the relative arc length is shown in (19).

\[
d_{mn} = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2}
\]

(18)

\[
m and n in (18) are two consecutive pixel, \(x_m and y_m\) are the x and y coordinate of pixel \(m\) and \(x_n and y_n\) are the x and y coordinate of pixel \(n\) respectively.

\[
\text{relative } i^{th} \text{ arc length} = \frac{\text{approximate mod arc length}}{i^{th} \text{ arc length}}
\]

(19)

The last proposed shape feature representation is basic visual components identification and this feature is based on the tokens’ properties such as token eccentricity and sign of token curvature.

**Fig. 7.** Comparison between symbolic and numeric representation

The main reason of choosing symbolic method is because it provides approximate description of a shape that is suitable for retrieving similar image. Whereby numeric method provides exact description of a shape that is only suitable for retrieving exact image. By using symbolic method, it has the advantage of simplifying the similarity measurement and still maintaining the efficiency. In Fig. 7, token A and token B are represented by both symbolic and numeric method. It can be noticed that, by using symbolic method the two tokens are 100 percent similar, but it is only 50 percent similar when the numeric method is used. Therefore, by using symbolic representation the false retrieval can be avoided.

In this research, based on the token eccentricity and the sign of token curvature, eight types of basic visual component have been identified. Fig. 8 shows these basic visual components. These components cover two extreme cases and two moderate cases. The extreme cases include the component that is near to flat curve (component A and A') and the other is near to a very sharp and narrow curve (component D and D'). Whereas the moderate cases include the component that is wide curve but not flat (component B and B') and the other is near to the curve that has aspect ratio of one (component C and C').

**Fig. 8.** Eight types of basic visual component

### C. Similarity Measurement

In the similarity measurement, image distance between the query image and the image from database is calculated by using dynamic programming [8, 9, 10]. It will calculate the edit distance between the query image and the image from database and to fill the matrix cost. After the edit distance and matrix cost have obtained, then the optimal dynamic programming (DP) path will be able to determine. Since the matching pairs can be obtained from the optimal DP path, so this research found that by calculating the shape features distance from these pairs will produce a better result.

The shape features used in this matching are the token eccentricity, token local maximum curvature and token relative arc length. The sum of the shape features distance for each pair and the edit distance will form the image distance which will be used in the similarity measurement.

### IV. EXPERIMENTAL RESULT

Since shape similarity retrieval involved the notion of similarity that cannot be measured and it is very subjective [11]. Therefore, this section presents an alternative method which involves objective evaluation. This evaluation uses a small classified subset as the image database. There are 14 groups of images in the test database and each group contains about 9 images. The object in the image is carefully
selected so that the object similarity within the group is reasonably high. In addition, there are particularly different shape characteristic in each group that distinguish from other groups. Since every group only has around 9 images, so the short-listed of 15 images is chosen in this evaluation.

**TABLE I**

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of images</th>
<th>Result (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>Group B</td>
<td>6</td>
<td>97</td>
</tr>
<tr>
<td>Group C</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Group D</td>
<td>10</td>
<td>99</td>
</tr>
<tr>
<td>Group E</td>
<td>9</td>
<td>85</td>
</tr>
<tr>
<td>Group F</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Group G</td>
<td>10</td>
<td>79</td>
</tr>
<tr>
<td>Group H</td>
<td>10</td>
<td>99</td>
</tr>
<tr>
<td>Group I</td>
<td>11</td>
<td>96</td>
</tr>
<tr>
<td>Group J</td>
<td>10</td>
<td>78</td>
</tr>
<tr>
<td>Group K</td>
<td>7</td>
<td>71</td>
</tr>
<tr>
<td>Group L</td>
<td>8</td>
<td>79</td>
</tr>
<tr>
<td>Group M</td>
<td>9</td>
<td>98</td>
</tr>
<tr>
<td>Group N</td>
<td>9</td>
<td>83</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>9</strong></td>
<td><strong>86</strong></td>
</tr>
</tbody>
</table>

The result of the retrieval is shown in TABLE 1. From the experiment result, the proposed system produces an excellent result for Group A, B, D, H, I and M, which is above 90% accuracy. Among these groups, take Group A as an example since it has the accuracy of 100%. This means that, whenever one of the members in this group is used as the query, all other members will appear in the first 15 retrieved images as the output. Apart from these groups, Group E, G, J, L and N also show a good result which is about 80% of the retrieval accuracy. The results for the remaining groups are still within the acceptable percentage which is not less than 70%. Altogether for the overall performance, the developed system can produce an average result of 86%.

Besides, the research also carried out an experiment to evaluate the partial object retrieval. 611 images were derived to form the test images. The result obtained from the experiment is depicted in Fig. 9. The graph shows that the developed system can achieve about 80% accuracy for partial object retrieval at 10% occlusion, which is almost close to the retrieval accuracy for complete object. The developed system can still achieve an average retrieval accuracy of 70% for a 20% to 30% of occlusion, and about 60% retrieval accuracy for 40% and 50% of occlusion. The retrieval accuracy falls below 30% when the degree of occlusion increases to 60% and above.

The result of the experiment shows that the developed system is able to handle partial object retrieval. The decrease in the retrieval accuracy with the increase in the degree of occlusion does not mean that the system is unable to handle partial object retrieval. This is because the retrieval accuracy is significantly affected by the degree of occlusion. Therefore, it is impossible to avoid the decrease in the retrieval accuracy for the increasing degrees of occlusion. Instead, the main concern is to have a system that can reduce the rate at which the retrieval accuracy decreases with the increase in the degree of occlusion.

Fig. 9. Partial object retrieval accuracy for the proposed system

**V. CONCLUSION**

The most profound contribution is the success in developing the algorithm to handle for partial object retrieval. The key behind this success is the use of an effective shape decomposition technique that has been developed by this research. The design and implementation of the proposed dynamic image smoothing method has been completed. Hence, it had provided the capability to handle different size of images. In addition, the feature extraction and representation methods introduced ensure that the system is similarity transformation invariance and most importantly are the uniqueness of the feature. However, there are several research issues that need to be addressed as future works. Firstly, in order to make the system more useful and applicable in the wide range, it should allow the user to express and specify the query by sketching. Secondly, some artificial intelligent methodology such as the learning capability should be incorporated to enable the system to improve its performance by time.

**ACKNOWLEDGMENT**

Sincere thanks to F. Mokhtarian’s datasets for the system testing.

**REFERENCES**


IEEE Transactions on Image Processing, Vol. 9


