Envelope Extraction for Composite Shapes for Shape Retrieval

Jianguo Song[1], Xiaoqing Lu[1], Haibin Ling[2], Xiao Wang[1], Zhi Tang[1]

[1]*Institute of Computer Science & Technology, Peking University, Beijing, China*
[2]*Department of Computer & Information Sciences, Temple University, Philadelphia, U.S.A.*
Email: lvxiaoqing@pku.edu.cn

Abstract

Analysis of composite shapes recently receives increasing amount of research attention. Different from a silhouette, a composite shape rarely contains a complete envelope. In the paper, we propose a novel envelope extraction algorithm based on the Delaunay triangulation for composite shapes. By analyzing the spatial relationship among individual components of contours and their concavities, we establish new models to describe the envelope edges and their corresponding local enclosed regions. These new models are then used to extract accurate envelopes for composite shapes. We then apply the extracted envelopes to improve shape classification used in shape retrieval. The experimental results show that our algorithm effectively boosts existing shape retrieval algorithms.

1. Introduction

As one of the most important visual attributes for describing image content, shape plays an essential role in content-based image retrieval (CBIR). Obtaining an accurate envelope of a composite shape is one of the significant processes of shape retrieval.

The Gestalt theory proposed by psychologists in the early 19th century indicates that human visual cognition is inferred more by the properties of an image as a whole than by the simple sum of the parts. However, accurate extraction of holistic features remains an open problem. For example, in Figure 1, the six trademarks share similar circle patterns, although their internal contents are very different from one another. The Gestalt perception enables humans to perceive complex configurations holistically with global properties that cannot be simply derived by aggregating local properties. Nevertheless, extracting a reasonable outline of a complex shape, which possesses high visual salience, is crucial for the shape to match correctly the corresponding semantic concept, and is also helpful for further understanding the image.

Figure 1. Trademark examples with similar circle outlines.

A large number of shape analysis approaches have been developed, with the main focus on shape matching and retrieval. One of the earliest work on envelope extraction appears in 1996 [1], which applies the Gestalt principles in ARTISAN, a shape retrieval system for the UK Patent Office. In [2], geometric elements, such as parallel lines, concentric circles and arcs, are grouped following the Gestalt perception. In [3, 4] an approach is introduced for retrieving the envelope of high-level object groupings in bi-level images with multiple objects. In [5] a shape-based similarity retrieval system is developed to exploit both the contour and interior region of a shape. One perceptually motivated strategy for shape classification proposed in [6] with focus on decomposing a shape into a base structure and a set of "strand" structures. This strategy is also adopted to extract some specific envelopes, which are treated as the base structure of a shape contour with inward strand structures. An algorithm is proposed in [7] for computing the opening based on the corresponding concavity in order to eliminate such irrelevant segments. In [8] a concave hull algorithm based on improved Graham scanning is proposed to recover the object boundary through scattered points.

Although the above mentioned algorithms yield useful results, to the best of our knowledge, two issues remain to be improved: (1) an increase in the distance between components (also called objects) in a

shape may distort its envelope; and (2) the complicated blank area between irregular components makes it impossible to obtain an accurate envelope using simple judgments. To overcome the problems and to extract envelopes closer to human perception, we analyze the spatial relationship between an envelope and the related contours, and adopt an iteration method of contraction based on the Delaunay triangulation. The most important contribution of this paper is to establish a model to implement the shrinking criterion with the calculation of the proximity of a complex shape and local concavities.

2. Extraction Algorithm

2.1. Definition

We define a composite shape as a shape containing one or more disconnected components, and contours of the components are provided as 2D point sequences ordered clockwise.

A set $S = \{OB_1, OB_2, \cdots, OB_m\}$ is a collection of objects OBs in one shape, where m is the number of objects. Each object $OB_i = \{p_{i1}, p_{i1}, \cdots p_{iN_i}\}, i = 1,2, \cdots m$ is an ordered sequence of length N_i containing points $p_{ij}, j = 1,2, \cdots N_i$, such that for $1 < j \leq N_i$, p_{ij+1} is the next neighbor of p_{ij} clockwise in the same contour. A combined point set $TP = OB_1 \cup OB_2 \cup \cdots OB_m$ assembling all points from OBs serves as the foundation for the Delaunay triangulation. As shown in Figure 2, the Delaunay triangulation generates a group of triangles for the given S. An edge $E = (p_s, p_e)$ consists of two points, the start point p_s and end point p_e. The extracting algorithm starts from one of the outermost boundaries and performs in a clockwise order; hence only the two edges at the right-hand side of each edge need to be considered. For example, edges E_s, E_e are on the right-hand side of E, E_s starting at p_s and E_e ending at point p_e. The middle point p_{nn} is shared by E_s and E_e.
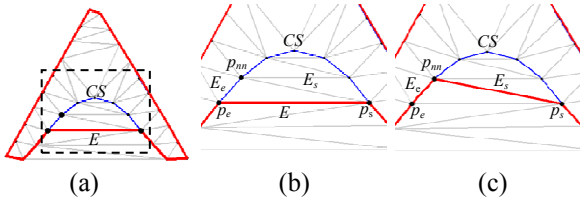
Figure 2. Example of the shrinking operation. (a) A shape with a rectangle region enlarged in (b) and (c); (b) Before shrinking at edge E; (c) After shrinking.

2.2. Workflow

As illustrated in Figure 3, the proposed algorithm contains 4 steps: (1) the Delaunay Triangulation (DT) is adopted to decompose the original boundaries in a shape; (2) a primal envelope (convex hull) $S_{envelop} = \{E_1, E_2, \cdots E_{N_{convexhull}}\}$ is obtained. E_i is one of the edges composing the envelope in a clockwise order, and $N_{convexhull}$ is the number of edges; (3) an iteration process of contraction is conducted, where each edge in the primal envelope is judged by shrinking criterion for further adjustment; (4) the satisfied envelope $S_{envelop} = \{E_1, E_2, \cdots E_{N_{envelop}}\}$ is obtained, wherein $N_{envelop}$ denotes the number of the updated edges.
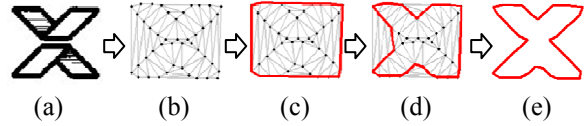
Figure 3. The workflow of the proposed algorithm. (a) Original Image; (b) The result of DT; (c) Initial envelope (in red); (d) Operation of shrinking; (e) Final envelope.

2.3. Shrinking criterion

The judgment of shrinking clearly plays a key role in the extraction of envelopes. However, three challenging problems need to be solved: First, a threshold is calculated based on the proximity of shapes to select and keep the edges shorter than it in the envelope. Second, to judge each edge larger than the threshold, a new polygon, called dynamic local enclosed region (DLER), needs to be established. Third, a method is adopted to judge the concavity of the DLER and determine whether the envelope should shrink.

2.3.1. Analysis of the proximity. In the first step, an acceptable threshold for preserving the very short bridge lines must be found. However, a fixed value can hardly be applicable to all cases. A very small threshold limits the function of the threshold, whereas a too large one may lead to the miss of some necessary adjustments for the envelope. In accordance with our observations, such a threshold can be more appropriate if calculated based on the proximity of a shape.

To describe the proximity, a representative distance for all objects in a shape is first computed by

$$Rep\_Dist = \max_{1 < i \leq m} \left\{ \min_{p_x \in OB_i, p_y \in \{S - OB_i\}} \{d(p_x, p_y)\} \right\}, \quad (1)$$

where $d(p_x, p_y)$ denotes the Euclidian distance between point p_x and point p_y. Thus, the proximity of a shape, called S_Prox, can be obtained by normalizing the corresponding Rep_Dist with formula (2)

$$S\_Prox = 1 - \frac{Rep\_Dist}{L_{Shape}}, \tag{2}$$

where L_{Shape} is the size of the whole shape. Figure 4 demonstrates the capability of S_Prox to describe the proximity of two sample shapes. Therefore, a more applicable threshold can be obtained based on S_Prox with the following formula

$$D_T = \alpha(1 - S\_Prox)Avg_{Dist}, \tag{3}$$

where Avg_{Dist} is the average of the distances and α is a weight parameter. Formula (3) implements the adjustment that if the proximity of the shape increases, D_T automatically decreases, and vice versa, as also shown in Figure 4.

$S\_Prox = 0.96$
$D_T = 4$

$S\_Prox = 0.81$
$D_T = 18$

Figure 4. Examples with different D_T.

2.3.2. Establishment of DLER. The second step focuses on edges larger than the threshold. For further analysis, a virtual polygon must be constructed to combine long edges and their related object boundaries.
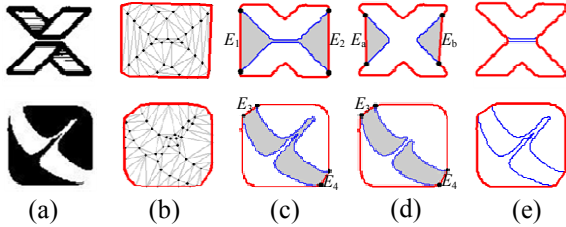

(a)   (b)   (c)   (d)   (e)

Figure 5. Examples of shrinking determination with DLERs. (a) Original images; (b) Initial envelopes(in red); (c) Space area between objects (in gray); (d) DLERs; (e) Final envelope.

Figure 5 shows two typical examples consisting of two discrete objects and four long edges E_1, E_2, E_3, and E_4, which need to be judged for shrinking. The unenclosed space areas between two irregular objects are shown in grey in the third column of Figure 5. However, the characteristics cannot be easily computed without a complete figure. To enclose the regions for E_1, E_2, E_3, and E_4, four new lines are calculated and embedded by connecting two points on the two boundaries separately, leaving the current edges as far as possible. Simultaneously, their lengths are kept no less than D_Ts. As shown in grey in the fourth column of Figure 5, new polygons surrounded by the embedded lines, the current edges, and two boundaries are then formed. These polygons are called dynamic local enclosed region (DLER).

2.3.3. Shrinking judgment based on the concavity. The judgment for shrinking mainly focuses on edges whose start point $\boldsymbol{p_s}$ and end point $\boldsymbol{p_e}$ are not adjacent to each other in the object contour, such as $\boldsymbol{E}$ illustrated in Figure 2 (a) and $\boldsymbol{E_1}$, $\boldsymbol{E_2}$, $\boldsymbol{E_3}$, $\boldsymbol{E_4}$ in Figure 5. In this case, the current edge $\boldsymbol{E}$ and the segment of the contour ($\boldsymbol{CS}$) also starting at point $\boldsymbol{p_s}$ and ending at point $\boldsymbol{p_e}$ form a temporary area called the local enclosed region (LER) or DLER. LER is treated as a special case of DLER. Figures 6(a)(b)(c)(d) show four different local regions. Undoubtedly, a strong relationship exists between the concavity of $\boldsymbol{CS}$ and the attraction to shrink $\boldsymbol{E}$. As shown in Figure 6(a), a small concavity of a $\boldsymbol{CS}$ means that its contribution to the envelope is substantial, and the envelope should shrink toward the $\boldsymbol{CS}$. On the contrary, as shown in Figures 5(c)(d), when the concavity of a $\boldsymbol{CS}$ is large, it contributes less to the envelope and the current $\boldsymbol{E}$ is more suitable to be kept in the envelope. The possibility of shrinking at the current edge is then transformed into calculating the concavity of the $\boldsymbol{CS}$ corresponding to the current edge.

Instead of a complicated calculation for the accurate concavity, the ratio of the CS relative depth to the CS opening width L_E is adopted as the bending degree of the CS to simulate its concavity. The present study only concerns with the depth of CS along the inward normal direction to E; hence, the rough value of depth can be adopted and calculated by $Area_{\mathrm{DLER}}/L_E$, where $Area_{\mathrm{DLER}}$ is the area of the current DLER and L_E is the length of the current E. In the proposed model, the definition of the rough concavity of a CS is briefly described as
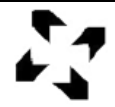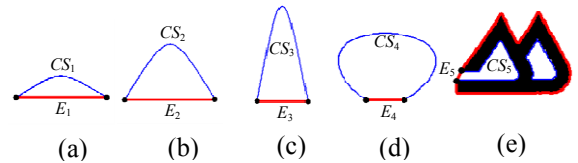
$$RC_{CS} = \frac{Area_{\mathrm{DLER}}}{(L_E)^2}, \tag{4}$$

(a)   (b)   (c)   (d)   (e)

Figure 6. Five situations for the analysis of concavity of CS.

Although the shapes of DLERs have diverse appearances, RC_{CS} correspondingly reflects the possibility of the envelope shrinking. A low value of RC_{CS} means that the envelope should be adjusted toward the original contour to reflect clear boundaries. A high value resists the shrinking from the current edge of envelope.

3. Experiments

This section demonstrates the validity of the proposed approach and how it improves the accuracy of shape retrieval. For all shapes, which is used to show its envelope in the literature [4-6], we obtain similar or smoother envelopes with our algorithm. Next, the whole MPEG7 CE-2 Set B database [9] is used for envelop extraction. Although the database consists of 2811 shapes, the proposed method can extract all the satisfied envelopes of them in one process by the aid of the adaptive proximity and the local concavity. We emphasize that our resulting envelopes of the complex shapes, shown in Table 1, match the overall impression of human properly, which, to the best of our knowledge, have not been achieved in previous literature.

Table 1. Envelopes extracted from samples in MPEG7 CE-2.

Initial Image	Envelope	Initial Image	Envelope

Table 2. Bull's eye performance of the 682 queries in MPEG-7 CE-2 SET B.

	ZMD	ART	GFD	IDSC
Without envelope	52.67%	63.45%	58.64%	N.A.
With envelope	72.50%	70.53%	71.58%	60.13%

Furthermore, we combine the algorithm with four methods of shape classification and retrieval to show that the envelopes are helpful. In our experiment, 682 shapes have been manually sorted out into 10 classes from MPEG-7 CE-2 SET B. In Table.2, the enhanced ZMD [10], GMD [9] and ART [5] methods which take envelops as additional features improve their precisions of retrieval obviously. Especially for those methods, such as the Inner Distance Shape Context [11] (IDSC), which are very effective on the shapes with single outline but cannot be used directly on the shapes with composite, the envelops provide a bridge. The last column in Table 2 shows that the IDSC not only can be applied on the composite shapes with the help of envelops, but also achieves parallel result on MPEG7 CE-2 Set B database. In fact, envelops can be adopted by many other kinds of methods, such as [12], and the combination of them can obtain more promising results.

4. Conclusions

In this paper we presented an effective envelope extraction method for composite shapes. The proposed method has two main advantages compared with existing techniques. First, the concavity analysis of the bridge edges between different objects is distinguished from the analysis of a single contour. Second, different kinds of spaces among objects can be processed in a new model by individually calculating their proximities. The experiment results show that the proposed method possesses a strong adaptability for composite shapes, such as trademarks shapes.

Acknowledgment

References

[1] Eakins, J.P., Boardman, J.M., and Shields, K. Retrieval of trade mark images by shape feature-the ARTISAN project. In *IEE Intelligent Image Databases*, p. 9/1-9/6, 1996.

[2] Jiang, H., Ngo, C., and Tan, H. Gestalt-based feature similarity measure in trademark database. *Pattern Recognition*. 39,5, 988-1001, 2006.

[3] Alajlan, N., Badawy, O.E., Kamel, M.S., and Freeman, G.H. Envelope Detection of Multi-object Shapes. *ICIAR*, p. 399-406, 2005.

[4] Alajlan, N. Retrieval of Hand-Sketched Envelopes in Logo Images. *ICIAR*, p. 436-446, 2007.

[5] Hung, M.H., Hsieh, C.H., and Kuo, C.M. Similarity retrieval of shape images based on database classification. *J. of Visual Communication & Image Representation*, 17(5): 970-985, 2006.

[6] Temlyakov, A., et al., Two perceptually motivated strategies for shape classification, *Computer Vision and Pattern Recognition*, p. 2289-2296, 2010.

[7] Torres., D.C.G.P. Shape retrieval using contour features and distance optmization. In *VISAPP*, p. 197-202, 2010.

[8] Junyi, X., et al., A Concave Hull algorithm for scattered data and its applications. *Image and Signal Processing*, p. 2430-2433, 2010.

[9] Zhang, D. and G. Lu, Shape-based image retrieval using generic Fourier descriptor. *Signal Processing: Image Communication*. 17(10): p. 825-848, 2002.

[10] Kim, W. and Y. Kim, A region-based shape descriptor using Zernike moments. *Signal Processing: Image Communication*. 16(1–2): p. 95-102, 2000.

[11] Ling, H. and Jacobs, D. W. Shape Classification Using the Inner-Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 29(2):p. 286-299, 2007.

[12] Cerri, A., Ferri, M., and Giorgi, D. Retrieval of trademark images by means of size functions, *Graphical Models*. 68: p. 451–471, 2006.