AFFINE-INVARIANT SHAPE MATCHING AND RECOGNITION UNDER PARTIAL OCCLUSION

F. Mai*, C.Q. Chang and Y.S. Hung

Department of Electrical and Electronic Engineering
The University of Hong Kong

ABSTRACT

In this paper, a new approach for matching and recognizing affine-distorted planar shapes is proposed, which allows for partial occlusions. In our approach, each shape is divided into a sequence of ordered affine-invariant segments based on the properties of curvature scale space (CSS) shape descriptor. Then Smith-Waterman algorithm is applied for matching the pair of affine-invariant segment sequences. Finally the point correspondences along the two shapes are obtained and the affine transform is estimated accordingly. Experiments show that our algorithm is effective for matching and recognizing shapes which are distorted by affine transforms, including translation, rotation, scaling, and shearing. Furthermore, it is capable of dealing with the case where one of the shapes or both shapes are partially occluded.

Index Terms— Affine-invariant matching, partial occlusion, planar curves, Smith-waterman algorithm

1. INTRODUCTION

The boundaries of planar shapes contain considerable amount of information which enables human beings to recognize objects without further information. The problem of shape matching and recognition is of central importance in computer vision, and has wide applications in shape similarity retrieval, object recognition and 3D curve reconstruction. However, despite the great efforts for solving this problem, it is still a very difficult task for two main reasons: 1) the images of an object taken from different viewpoints suffer from perspective distortions, and 2) sometimes the planar shapes to be recognized are partially occluded, which makes the problem even more challenging.

It is well known that a perspective transformation between two images of an object can be approximated by an affine transformation if the object is planar and far away from the image plane [1]. Therefore, various affine invariant shape matching methods have been developed [2]. They are capable of addressing complex issues such as matching under excessive noise condition, matching under affine transformations, etc. However, they work only for closed shapes, assuming

---

*Corresponding author. Email address:feimai@eee.hku.hk

that the whole shape is always visible in images. In real applications, it is possible that the object to be recognized is only partially visible, which makes the recognition problem far more difficult than that for closed shapes. Only very few approaches for shape matching under partial occlusion have been proposed [3] [4]. However, they work only for shapes up to a similarity transform.

In this paper, we present a new method for affine-invariant shape matching and recognition under partial occlusion, aiming to solve the above mentioned problems. The rest of this paper is organized as follows. Section 2 gives an overview of the proposed method. In Sections 3 and 4, the details of the proposed method are presented. Experimental results are shown in Section 5, and concluding remarks are given in Section 6.

2. OVERVIEW OF THE METHOD

For two given planar shapes A and B, suppose the boundary of each shape has been extracted and is represented by a sequence of ordered points respectively as: $A = [x_a(t_a)]$ $(t_a = 1, 2, \cdots, N_a)$ and $B = [x_b(t_b)]$ $(t_b = 1, 2, \cdots, N_b)$, where $x_a(t_a)$ (or $x_b(t_b)$) is a 2×1 vector, giving the u and v coordinates of the t_a-th (or t_b-th) point along A's (or B's) boundary. For a shape without occlusion, its boundary is a closed curve. If a shape is partially visible, its boundary is represented by an open curve. The goal of affine-invariant shape matching is to find out if A and B are related by an affine transform. That is, to find if A and B share a maximum common part $(A^c \in A, B^c \in B)$, such that $B^c = TA^c + t$, where T is the affine transformation including rotation, scaling and shearing, t is the translation.

Let us denote the i-th segment of A as $a_i = [x_a(s_{a_i})]$, where $s_{a_i} = p_i, p_i + 1, \cdots, p_{i+1}$ (when $p_i < p_{i+1}$); and $s_{a_i} = p_i, p_i + 1, \cdots, N_a, 1, \cdots, p_{i+1}$ (when $p_i > p_{i+1}$ for the last segment of a closed curve). In both cases, $x_a(p_i)$ and $x_a(p_{i+1})$ are the starting point and end point of segment a_i respectively. Similarly, a segment of B between point $x_b(q_j)$ and $x_b(q_{j+1})$ is denoted as $b_j = [x_b(s_{b_j})]$.

In matching the two shapes A and B, the process consists of three steps.

1. The affine-invariant points along each shape are de-

tected. Every two consecutive affine-invariant points determine an affine-invariant segment along the shape, and thus a sequence of ordered affine-invariant segments are created for each shape.

2. The two sequences of affine-invariant segments are aligned by the Smith-Waterman algorithm, and thus an initial matched part is obtained.

3. If at least one shape is partially visible, the initial matched part is extended using the affine invariant affine-length, such that the maximum matched part is obtained. Accordingly, the point correspondences are obtained and the affine transform is estimated.

In the following, we will present our algorithm in details.

3. AFFINE-INVARIANT SEGMENTS CONSTRUCTION

To deal with the occlusion case, we need to use local information as well as global information of the shapes. Therefore, we divide the shape into affine-invariant segments, where an affine-invariant segment means the segment is preserved under affine transforms. In [5], it is proved that the locations of the contour local maxima in the curvature scale space (CSS) image are preserved under an affine transform. Furthermore, they are robust to noise. Therefore, we use this property to construct the affine-invariant segments of the shape.

The details for how to construct the CSS images can be found in [5], and we will not repeat it here due to the space limitation. An example is given in Figure 1, where (a) is the CSS image of the shape in (b), and the red asterisks in (b) are the affine-invariant points $x_a(p_i)$, which are extracted by detecting the local maxima of the contours in the CSS image (a). For two shapes A and B to be matched, we use this method to extract the affine-invariant points $x_a(p_i)$ in shape A, and those $x_b(q_j)$ in shape B. Figure 2 shows the segmentation result of a pair of affine related camel shapes, where one of the shapes is partially occluded.
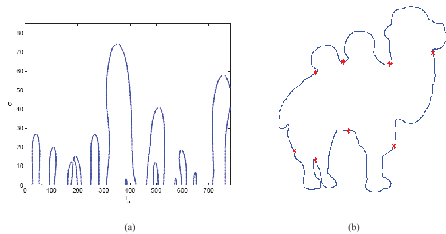

Fig. 1. Constructing affine-invariant segments using CSS image. (a): the CSS image of the shape shown in (b). (b): the affine-invariant points (denoted by red asterisks) corresponding to the local maxima of the contours in the CSS image (a).

4. AFFINE-INVARIANT SEGMENTS MATCHING

Suppose there are M affine-invariant points of shape A. Then A is represented by a sequence of M affine-invariant
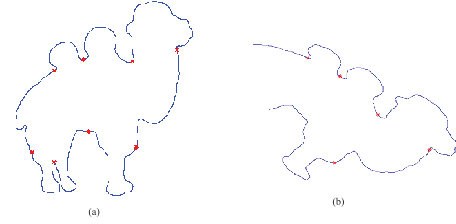

Fig. 2. Affine-invariant segmentation. (a) and (b) are related by an affine transform and (b) is partially occluded. The affine-invariant points are shown by red asterisks and it is clear that they are preserved under the affine transform.

segments if A is a closed shape, and a sequence of $M - 1$ affine-invariant segments if A is an open shape. Similarly, shape B is represented by a sequence of N segments if B is a closed shape, and $N - 1$ segments if B is partially visible. To find the common part of the two shapes, we apply the Smith-Waterman algorithm [6] to find the local alignment of the two sequences. The Smith-Waterman algorithm was proposed for identification of common molecular subsequences. More generally, it is a dynamic programming algorithm that finds the local alignment between two sequences of elements with respect to a scoring system.

4.1. Similarity Score

To establish a scoring system for the Smith-Waterman algorithm, we need to measure the similarity between two segments up to an affine transform. For a pair of elements (a_i and b_j) which are related by an affine-transform, the translation effect can be removed by moving origins of their coordinate systems to the curve segment centroids. For scaling and shearing, it is easy to prove that they can be eliminated by a whitening transformation [7]. Let us denote the shape elements after centering as $\overline{a}_i$ and $\overline{b}_j$ accordingly. We apply the singular value decomposition (SVD) to them respectively, for whitening.

$$\begin{aligned} \overline{a}_i &= U_a S_a V_a^T \\ \overline{b}_j &= U_b S_b V_b^T \end{aligned} \tag{1}$$

where V_a^T and V_b^T are the whitened sequences corresponding to a_i and b_j, and they represent the canonical forms of the two curve segments. It is easy to show that the two canonical forms of a pair of affine-related curve segments are related only by a rotation. To measure the similarity of the two canonical forms, we use the centroid distance (the distance between a boundary point to the shape centroid) as the shape signature to represent the shape, which is rotation invariant. Then the similarity of the two canonical forms for a_i and b_j is measured by computing the cross-correlation score of the corresponding 1-D shape signatures. One example is given in Figure 3, where the correlation score of the shape elements pair (in red) is 0.9994.
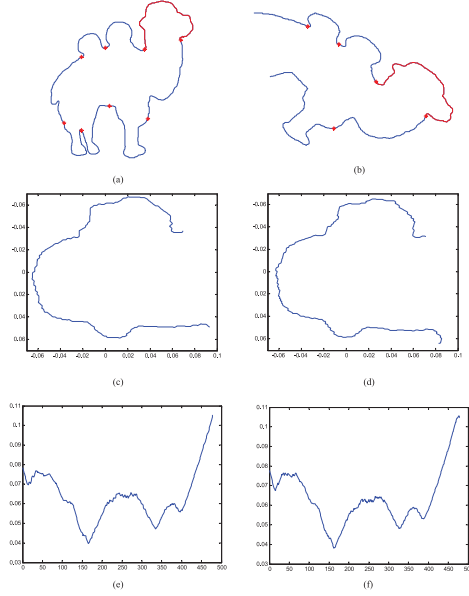
Fig. 3. Similarity measurement: the red segments in (a) and (b) are the shape elements to be compared; (c) and (d) are the canonical forms for the two shape elements; (e) and (f) are the 1-D signatures of the canonical forms.

4.2. Smith-Waterman Alignment

Totally there are three possible cases: whole-to-part matching (one shape is closed and the other is open), whole-to-whole matching (both are closed), and part-to-part matching (both are open). Let us first discuss the whole-to-part matching case. Without loss of generality, we assume A is closed and B is open. Then the two sequences can be denoted by $A = \{a_1, a_2, \cdots, a_M\}$, and the $B = \{b_1, b_2, \cdots, b_{N-1}\}$. A similarity score for a_i and b_j, denoted by $w(i,j)$, can be computed using the method presented in Section 4.1. Furthermore, a gap penalty is defined as:

$$w(a_i, -) = w(b_j, -) = -1 \qquad (2)$$

where $-$ is an insertion of a gap into the aligned sequences. In the Smith-Waterman algorithm, a matrix H is set up, whose preliminary values have the interpretation that $H(i,j)$ is the maximum similarity of two elements ending in a_i and b_j, respectively [6]. H is built as follows:

$$H(i,0) = 0, \ 0 \leq i \leq M$$
$$H(0,j) = 0, \ 0 \leq j \leq N-1$$

$$H(i,j) = \max \left\{ \begin{array}{c} 0 \\ H(i-1, j-1) + w(a_i, b_j) \\ H(i-1, j) + w(a_i, -) \\ H(i, j-1) + w(-, b_j) \end{array} \right\}, \qquad (3)$$

$$1 \leq i \leq M, 1 \leq j \leq N-1$$

The pair of elements with the maximum similarity is found by first locating the highest value in H. The other matrix entries leading to this maximum value are then sequentially determined with a trace-back procedure ending with an entry of H equal to zero. This procedure identifies the corresponding element alignment. Details of the Smith-Waterman algorithm can be found in [6]. Since A is a closed shape, a circular shift is applied to A, denoted by $A_i = \{a_{i \mod M}, a_{(i+1) \mod M}, \cdots, a_{(i+M-1) \mod M}\}, (i = 1, 2, \cdots, M)$, where $\mod$ is the circular shift operator. Every A_i is aligned with B and the best alignment is selected as the initial matched part.

For the whole-to-whole matching case, we use the same method to get the initial matched part. For the part-to-part matching case, we can still use the above method except that no circular shift is needed.

As far as the initial matched part is obtained (see the red part in Figure 4 (a) and (b)), we extend the matched part by computing the affine-length [8] (see the green part in Figure 4 (c) and (d)). Thus the point correspondences along the matched part are determined accordingly. Therefore, the affine transform between shape A and B can be computed from the point correspondences (see Figure 4 (e) and (f), where the blue shape is the original shape, and the red dashed line describes the "registered" shape, which is defined as the shape converted from its corresponding shape by the obtained affine transform).
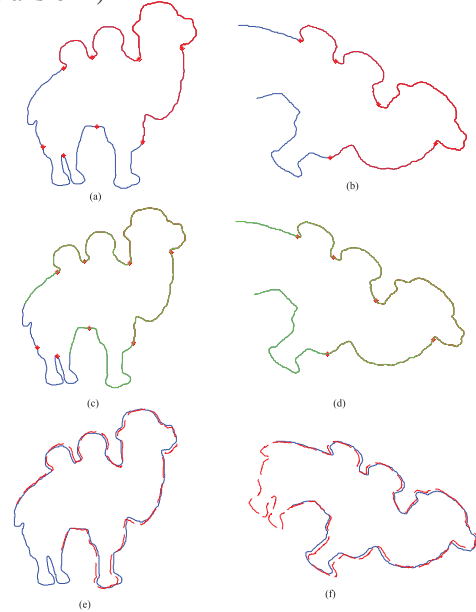

Fig. 4. Final Matching: the red segments in (a) and (b) are the initial matched part; the green segments in (c) and (d) are the final matched part; (e) and (f) : the original shape (in blue line) overlaid with the "registered" shape (in red dashed line)

5. EXPERIMENTAL RESULTS

To test our algorithm, we carry out the experiments on the Multiview Curve Database (MCD) [9]. In the original database, all shapes are closed curves. In our experiments,

4607

we randomly remove some part of the shape to make it partially visible. Due to the space limitation, apart from the whole-to-part example shown in Figure 4, two more examples (Figure 5 for whole-to-whole matching and Figure 6 for part-to-part matching) are shown in this section. It is clear that our algorithm works well for all the three cases.

In the following we focus more on the whole-to-part case to test the robustness of our algorithm to partial occlusion. For quantitative analysis, we randomly select one affine-related shape pair from each of the 40 shape classes in MCD, and remove part of one shape (randomly ranging about 10% to 50%) in each pair, resulting 40 part-to-whole affine-related pairs in total. The registration error is measured by the percentage of non-overlapping area between the common part of the original shape and its "registered" shape. We compare our method with [4], as a representative of the state-of-the-art methods for shape matching under partial occlusion. We take the average registration error over the 40 shape pairs for a statistal quantitative comparison. The average error of our proposed method is 13.12%, which is much smaller than 49.41%, the average error of the reference method [4]. It shows that our method is robust to both partial occlusions and affine distorsions, while the reference methods is sensitive to affine-transformation although it works well for partial occlusion under similarity transformations.
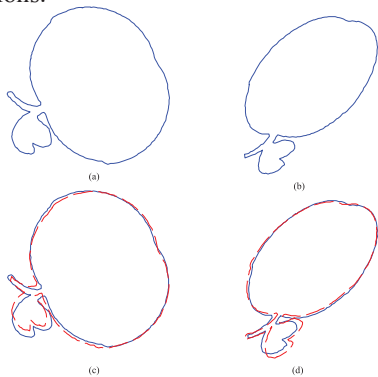

Fig. 5. Apple pair: (a) and (b) are the initial matched part; (c) and (d) show the original shape (in blue line) overlaid with the "registered" shape (in red dashed line)
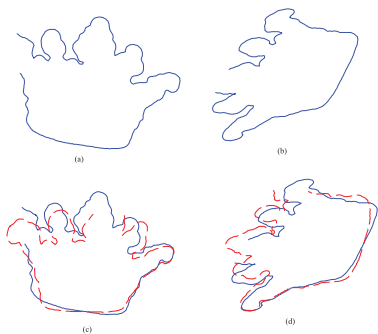

Fig. 6. Crown pair: (a) and (b) are the initial matched part; (c) and (d) show the original shape (in blue line) overlaid with the "registered" shape (in red dashed line)

6. CONCLUSIONS

We propose an approach for affine-invariant shape matching and recognition, which allows for partial occlusion. We apply the Smith-Waterman algorithm for optimal local alignment of the affine-invariant segments constructed based on the properties of CSS images. All three possible cases are tested: whole-to-whole matching, whole-to-part matching, and part-to-part matching. The experimental results show that our algorithm is independent of transforms including translation, rotation, scaling and shearing. Future work includes the experimentation with more data sets and the application of our algorithm for shape retrieval and classification.

Acknowledgment

7. REFERENCES

[1] D. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant descriptors for 3d object recognition and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 10, pp. 971–991, Oct. 1991.

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

[3] M. Cui, J. Femiani, J.X. Hu, P. Wonka, and A. Razdan, "Curve matching for open 2d curves," *PRL*, vol. 30, no. 1, pp. 1–10, January 2009.

[4] E. G.M. Petrakis, A.s Diplaros, and E. Milios, "Matching and retrieval of distorted and occluded shapes using dynamic programming," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 1501–1516, 2002.

[5] F. Mokhtarian and S. Abbasi, "Shape similarity retrieval under affine transforms," *Pattern Recognition*, vol. 35, no. 1, pp. 31–41, 2002.

[6] M. S. Waterman and T. F. Smith, "Identification of common molecular subsequences.," *J. Mol. Biol.*, vol. 147, pp. 195–197, 1981.

[7] J. Sprinzak and M. Werman, "Affine point matching," *PRL*, vol. 15, pp. 337–339, 1994.

[8] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 8, pp. 789–805, Aug. 1992.

[9] M. Zuliani, S. Bhagavathy, B. S. Manjunath, and C. S. Kenney, "Affine-invariant curve matching," in *IEEE International Conference on Image Processing*, Oct 2004.