CURVES VS SKELETONS IN OBJECT RECOGNITION

Thomas B. Sebastian *Benjamin B. Kimia*

Division of Engineering, Brown University
Providence, RI 02912, USA

ABSTRACT

The type of representation used in describing shape can have a significant impact on the effectiveness of a recognition strategy. Shape has been represented by its bounding curve as well as by the medial axis representation which captures the regional interaction of the boundaries. Shape matching with the former representation is achieved by curve matching, while the latter is achieved by matching skeletal graphs. In this paper, we compare the effectiveness of these two methods using approaches which we have developed recently for each. The results indicate that skeletal matching involves a higher degree of computational complexity, but is better than curve matching in the presence of articulation or rearrangement of parts. However, when these variations are not present, curve matching is a better strategy due to its lower complexity and roughly equivalent recognition rate.

1. INTRODUCTION

The type of shape representation used can have a significant impact on the effectiveness of a recognition strategy. A successful recognition technique has to be robust to visual transformations like articulation and deformation of parts, viewpoint variation, occlusion. Thus, the shape representation has to effectively capture the variations in shape due to these transformations. In previous recognition applications, shapes have been represented as curves [1, 2], point sets or feature sets, and by medial axis [3, 4, 5, 6, 7], among others. This paper compares two techniques for matching shapes, one based on matching their outline curves [8] and the second based on matching their shock graphs [9].

In many object recognition and content-based image indexing applications, the objects are represented by their outline curves and matched. Outline curves typically do not represent a notion of the interior of the shapes. Despite this well-known drawback, it has been effectively used in certain applications [1, 10, 2]. Matching typically involves finding a mapping from one curve to the other that minimizes an "elastic" performance functional, which penalizes "stretching" and "bending" [2]. The minimization problem in the discrete domain is transformed into one of matching shape signatures with curvature, bending angle, or orientation as attributes [1, 10]. The curve-based methods in general typically suffer from one or more of the following drawbacks: asymmetric treatment of the two curves, lack of rotation and scaling invariance, and sensitivity to articulations and deformations of parts.

Shapes have also been represented by the medial axis or its variants, which represent both the interior and the outline of the shape. Medial axis has been effectively used for matching shapes [3, 7]. The shock graph is the medial axis endowed with geometric and dynamics information, and is a richer descriptor of shape than the medial axis graph, since its graph topology is more in accord with our perceptual notions of shape. Shock graph matching has been used in [4, 5, 6] for object recognition and image indexing tasks. A recent approach [9] uses an edit-distance between shock graph to match shapes effectively. However, the use of graph matching techniques in general is computationally more intensive than curve matching. This gives rise to the question of whether the additional effort required in skeletal matching is justified by the improvements in recognition rates for particular applications. This paper compares the matching of shapes based on curves and skeletal graphs in general and specifically for the recognition task.

2. CURVE MATCHING

In this section, we briefly review the outline-based recognition method, which is based on finding the minimum-cost deformation of one curve to the other. The cost of the deformation is defined as the sum of "stretching" and "bending" energies [2]. The basic premise of the approach is that the cost of the deformation can be expressed as the sum of the cost of deforming infinitesimal subsegments, which is defined by length and curvature differences as $\mu = |d\hat{s} - ds| + R|d\hat{\theta} - d\theta|$, where R is a scale-dependent constant. The problem is then cast as minimizing an energy functional over all possible *alignments* between the two curves. The notion of an *alignment curve* is introduced to ensure the symmetric treatment of the two curves, and the optimal alignment is found using an efficient dynamic-programming

The support of NSF grants IRI-9700497 and BCS-9980091 is gratefully acknowledged.

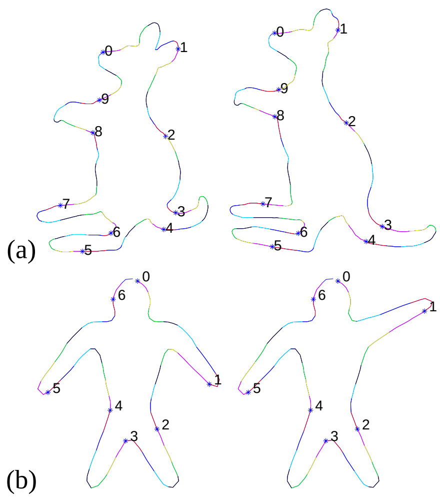

Fig. 1. This figure illustrates that the curve matching algorithm works well in presence of view-point variation (a) and articulation and deformation of parts (b). The alignment is indicated by arbitrarily coloring portions of the aligned curves by identical colors with a number indicating the each portion's end point. Observe that the matches are intuitive, *e.g.*, hands, legs and head of the dolls correspond.

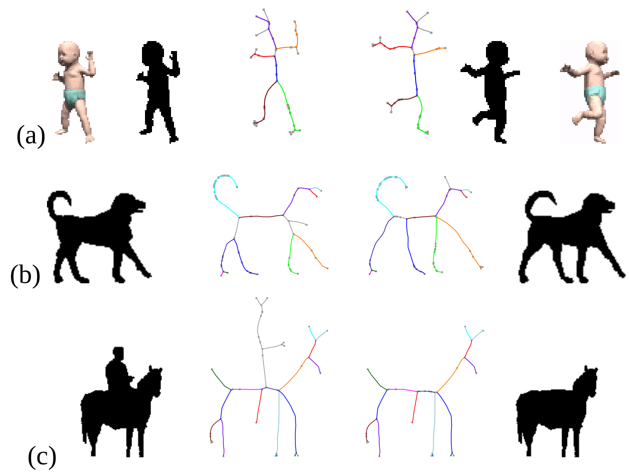

Fig. 2. This figure illustrates that the shock graph matching algorithm works well in presence of view-point variation (a) and articulation and deformation of parts (b) and occlusion (c). Same colors indicate matching shock branches, and grey colored branches in the shock graphs have been spliced. The matching is intuitive in all cases. Observe, in particular, how the rider on the horse is pruned when matched against horse by itself.

algorithm.

This technique works well in the presence of commonly occurring visual transformations, modest amounts of view-point variation, affine transformations, and under some articulation and deformations like stretching and bending of parts Figure 1. It has been applied to several applications including hand-written character recognition, prototype formation, and morphing [8].

3. SHOCK GRAPH MATCHING

In this section we briefly review the matching of shapes which are represented by the shock graphs. Shapes are viewed as points in a shape space and the distance between shapes is defined as the minimum cost of the deformation path connecting one shape to another. To make this search practical, an equivalence class on shapes is defined, where all shapes with the same shock graph topology are equivalent. In addition, all the deformation paths having the same set of transition points (boundaries between shape equivalence classes) are defined as equivalent. These shock transitions are points where the shock graph topology changes, and have been formally classified [11]. In the graph domain, each shock transition is represented by an "edit" operation on the shock graph, and there are four types of edit operations: (i) the *splice* operation deletes a shock branch and merges the remaining two; (ii) the *contract* operations deletes a shock branch connecting two degree-three nodes; (iii) the *merge* operations combines two branches at a degree-two node; (iv) we also define a *deform* edit to relate two shapes in the same shape cell, $i.e.$, shapes with the same shock graph topology but with different attributes. We associate a cost to each edit operation, and then find the minimum-cost sequence of edits by an efficient polynomial-time graph edit distance algorithm developed in [12].

The shock graph matching technique works well in the presence of articulation, deformation of parts, occlusion, presence of shadow and highlights, boundary noise, and viewpoint variation, Figure 2.

4. CURVES VS SKELETONS

This section discusses our experience with the use of curve and shock-graph based representations for matching shapes. The major advantage of directly matching the outline curves of shapes is the computational efficiency of curve matching, which is an order of magnitude faster than shock graph matching. Curve matching is the natural choice in applications where the item to be matched is inherently one-dimensional, $e.g.$, handwritten character recognition, signature verification $etc.$. However, we have identified some fundamental limitations in using a curve-based representation for general purpose object recognition. Specifically, a well-known shortcoming of curve-based representation is that it does not represent the interior of a shape. Hence, curve matching cannot easily distinguish between those per-

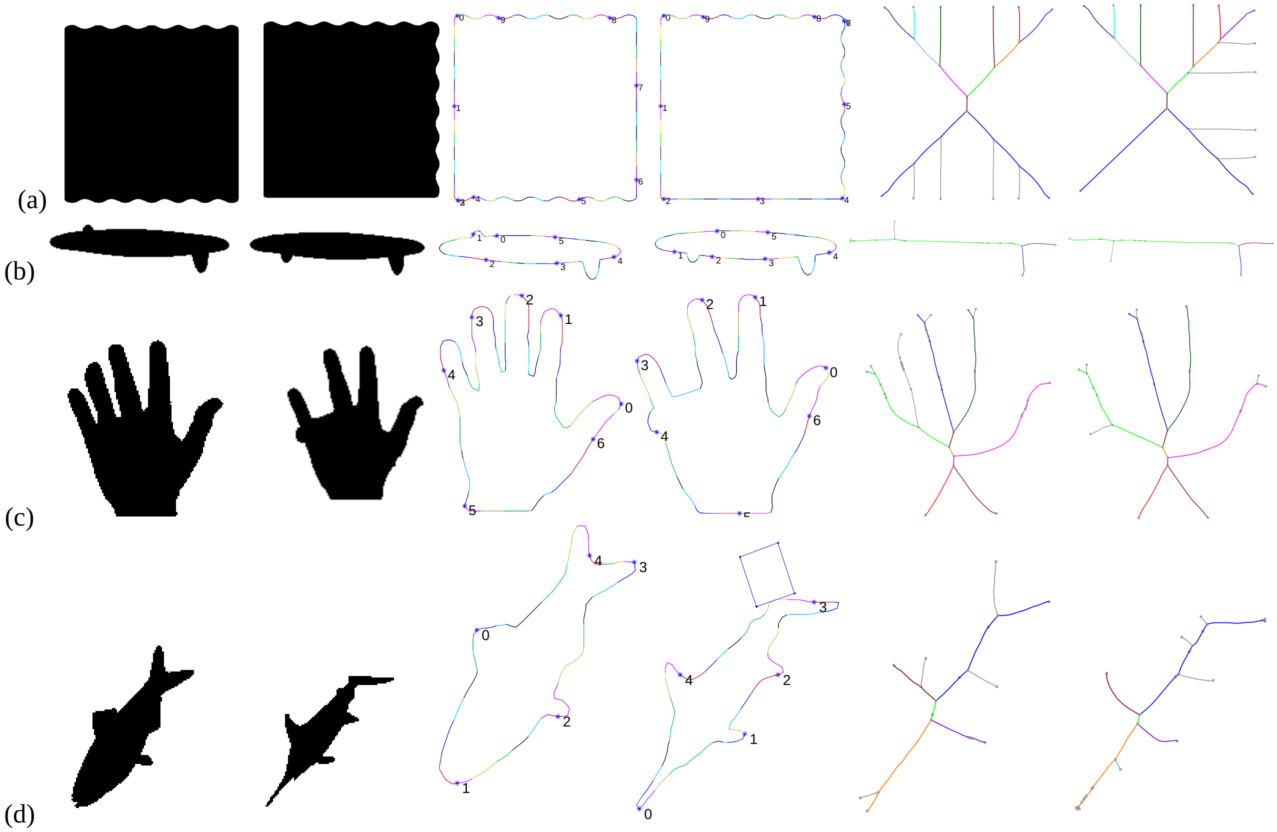

Fig. 3. Curves do not have a notion of the interior of the shape, and hence curve matching cannot find the intuitive correspondence in some cases. (a) The optimal matching is found for a pair of noisy squares, where two of the sides are replaced by a "wavy" lines. Curve matching aligns the wavy sides, ignoring the spatial information that shows a square. Shock-graph matching gives intuitive correspondence. (b)-(d) illustrates the sensitivity of the curve matching to spatial arrangement of parts and how shock-graph matching avoids the problem. (b) Two ellipses with protrusions. Curve matching correctly matches the larger protrusion, as it lies on the same side of the ellipse, but fails in the presence of the smaller protrusions that on opposite sides of the ellipse. Shock-graph matching splices out the small protrusions on different sides and matches the ellipses intuitively. (c) The missing finger on the hand on the left and the small bump of the hand on the right causes curve matching to give an un-intuitive match, where the bump is matched to a finger. On the other hand, shock-graph matching gives the intuitive correspondence. (d) Occlusion of part of the tail of the fish on the right affects the overall part structure, and curve matching gives the wrong correspondence, as a fin on the fish on the left is matched to the head of the fish on the right. Shock-graph matching splices the part of the tail on left fish corresponding to the occluded part of right fish is spliced, thus giving an intuitive correspondence, *i.e.*, the heads, tails and the fins of the two fishes matching intuitively.

ceptually distinct shapes whose local curve-based features are in conflict with the global shape percept, as shown in Figure 3a.

Another drawback of the curve representation and hence curve matching is the sensitivity to the presence and spatial arrangement of parts. Figure 3b-d shows examples where curve matching gives the un-intuitive correspondence when the parts around the shape are arranged differently. Note that the curve matching approach [8] is robust to the presence of occlusion, *if* it does not affect the overall part structure of the object. However, when the occlusion adds or deletes a part, curve matching can fail, as shown in Figure 3c. The shock graph of shape, on the other hand, inherently induces a part-based representation of shape based on regional aspects, and the spatial and hierarchical relationships among these parts. Shock-graph matching not only matches pairs of individual parts, but also the overall part hierarchy. These fundamental drawbacks favor shock-graph matching for the generic recognition problem where the space of variation is enormously large including the visual transformations such as those highlighted above. Tables 1 and 2 show recognition results for a database of 99 shapes.

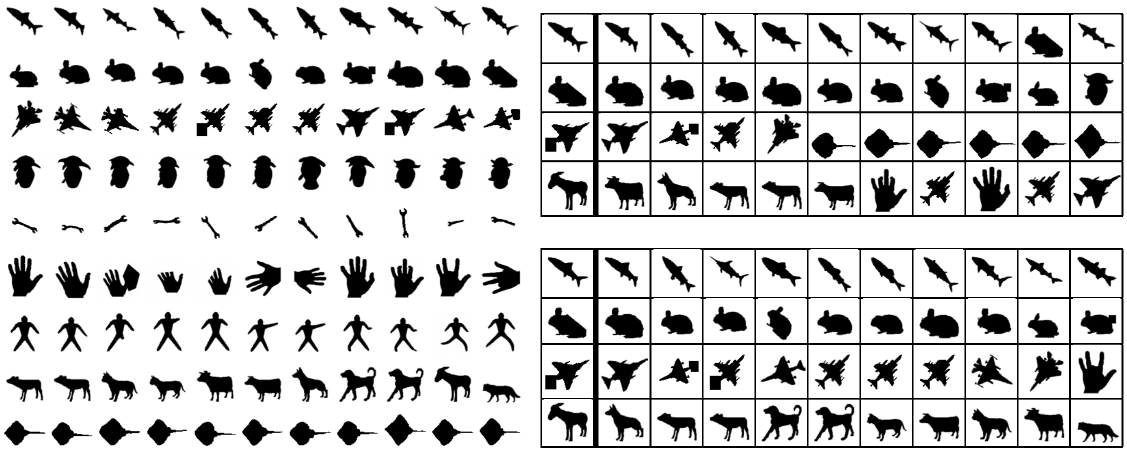

Table 1. Left: A database of 99 shapes with 9 categories and 11 shapes in each category. Right: Each shape in the database is matched against every other shape and the 10 nearest neighbors for a few representative shapes are shown for curve matching (top) and shock graph matching (bottom). Table 2 summarizes the results.

	1	2	3	4	5	6	7	8	9	10
Curve	100	98	98	99	95	95	91	85	75	51
Shocks	100	100	100	99	99	99	97	96	95	87

Table 2. Summary of results of curve and shock graph matching for classifying shapes from Table 1. Each entry corresponds to the number of times (in percentages) the n^{th} best match belongs to the same category as the query.

5. CONCLUSION

We conclude that curve matching is adequate for matching shapes whose space of variations does not include rearrangement and articulation of parts and where the regional interaction is not significant, $e.g.$, handwritten character recognition, signature verification, $etc.$ On the other hand, in the general object recognition problem due to the enormity of the space of variations, matching based on shock graphs gives superior results.

6. REFERENCES

[1] Yoram Gdalyahu and Daphna Weinshall, "Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes," *PAMI*, vol. 21, no. 12, pp. 1312–1328, 1999.

[2] Laurent Younes, "Computable elastic distance between shapes," *SIAM J. Appl. Math.*, vol. 58, pp. 565–586, 1998.

[3] S. C. Zhu and A. L. Yuille, "FORMS: A flexible object recognition and modeling system," *IJCV*, vol. 20, no. 3, 1996.

[4] Daniel Sharvit, Jacky Chan, Hüseyin Tek, and Benjamin B. Kimia, "Symmetry-based indexing of image databases," *JVCIR*, vol. 9, no. 4, pp. 366–380, 1998.

[5] K. Siddiqi, A. Shokoufandeh, S.J. Dickinson, and S.W. Zucker, "Shock graphs and shape matching," *IJCV*, vol. 35, no. 1, pp. 13–32, 1999.

[6] M. Pelillo, K. Siddiqi, and S.W. Zucker, "Matching hierarchical structures using association graphs," *PAMI*, vol. 21, no. 11, pp. 1105–1120, 1999.

[7] T. Liu and D. Geiger, "Approximate tree matching and shape similarity," *ICCV*, pp. 456–462, 1999.

[8] Thomas B. Sebastian, Philip N. Klein, and Benjamin B. Kimia, "Alignment-based recognition of shape outlines," *IWVF*, pp. 606-618, 2001.

[9] Thomas B. Sebastian, Philip N. Klein, and Benjamin B. Kimia, "Recognition of shapes by editing shock graphs," *ICCV*, Accepted – To appear, 2001.

[10] E. Milios and E.G.M. Petrakis, "Shape retrieval based on dynamic programming," *Trans. Image Proc.*, vol. 9, no. 1, pp. 141–146, 2000.

[11] P. J. Giblin and B. B. Kimia, "On the local form and transitions of symmetry sets, and medial axes, and shocks in 2D," *ICCV*, pp. 385–391, 1999.

[12] Philip Klein, Srikanta Tirthapura, Daniel Sharvit, and Benjamin Kimia, "A tree-edit distance algorithm for comparing simple, closed shapes," *SODA*, pp. 696–704, 2000.