A Metamorphosis-Based Shape Recognition Method

Rahul Singh Ioannis Pavlidis Nikolaos P. Papanikolopoulos

Artificial Intelligence, Robotics, and Vision Laboratory
Department of Computer Science, University of Minnesota, Minneapolis, MN 55455
E-mail: {singh, pavlidis, npapas}@cs.umn.edu

Abstract: *We present a framework for matching and recognition of planar shapes based on a method from computer graphics based animation, called "shape metamorphosis." In our approach, the "degree of morphing" between two shapes is employed as a dissimilarity measure. A physics-based energy minimization approach is used for optimally computing the "degree of morphing." This measure is shown to have metric properties and invariance to translation, rotation, scaling and mirror symmetry. Experimentations in the recognition of planar shapes, hand-drawn figures, and on-line cursive words indicate the robustness of the recognition paradigm.*

1 Introduction

Shape matching and recognition both in 2D and 3D are some of the most fundamental problems in Computer Vision. In spite of the advancements in object recognition [5], recognition of 2D shapes has remained an area which evokes much interest. The problem of 2D shape matching appears amongst others in the identification and handling of planar industrial parts, in Optical Character Recognition (OCR), and in the analysis of medical imagery. Furthermore, in many cases the analysis of 3D shapes can be decomposed into 2D shape analysis. The goal of this work is to recognize planar objects, and object contours which differ not only in terms of translation, rotation and scale, but also undergo contour perturbations or have features whose arrangement may alter due to shape deformations. Typical examples of such objects include hand-drawn shapes and cursive words (see Figs. 4 and 5).

Our approach is to use physics-based contour metamorphosis for shape recognition. We use this technique in a model-based framework. The contours of the planar object or the stroke patterns of a cursive handwritten word are modeled by a piece of deformable wire. Contour metamorphosis occurs through appropriate stretching and bending of the artificial wire. The *degree of morphing* required to transform an input wire-form shape to a reference wire-form shape is used as a dissimilarity measure. The *degree of morphing* is an abstract quantity, which in our system is substantiated through a minimum energy approach to contour metamorphosis proposed by Sederberg *et al.* [7] for the purpose of computer-based animation. The proposed approach is based on the intuitive fact that similar shapes need to undergo much smaller metamorphosis to assume each other's shape than do dissimilar ones (see Fig. 1). Some of the other methods proposed by researchers for this problem include the use of joint probability density over shape to define model flexibility, Hidden Markov Model (HMM) [2, 3], multidimensional co-occurrence matrices [4], elastic matching, and eigen-mode based representation. Our approach differs from the research [2, 3, 4] in that it does not require extensive shape statistics. In real-world applications, like user-dependent pen-based interfaces it allows the system to perform with minimal training. Based on [7], we use a computationally inexpensive dynamic programming approach to find a globally optimal solution to the energy minimization problem. This is in contrast to the approaches based on calculus of variation (e.g. elastic matching) or finite elements.

This paper is organized as follows: In Section 2 we explain the matching of shapes using contour metamorphosis. Section 3 includes experimental results. Finally, in Section 4 we present the conclusions and outline the future work.

2 Shape Matching by Contour Metamorphosis

Contour metamorphosis is defined as the transformation of one shape (as represented by its contour) to another. Let $\mathbf{S}^I = [S_0^I, \ldots, S_n^I]$ and $\mathbf{S}^T = [S_0^T, \ldots, S_n^T]$ be the point sets representing the input and the target contours, respectively. Metamorphosis of $\mathbf{S}^I$ to $\mathbf{S}^T$ is defined by a sequence of intermediate

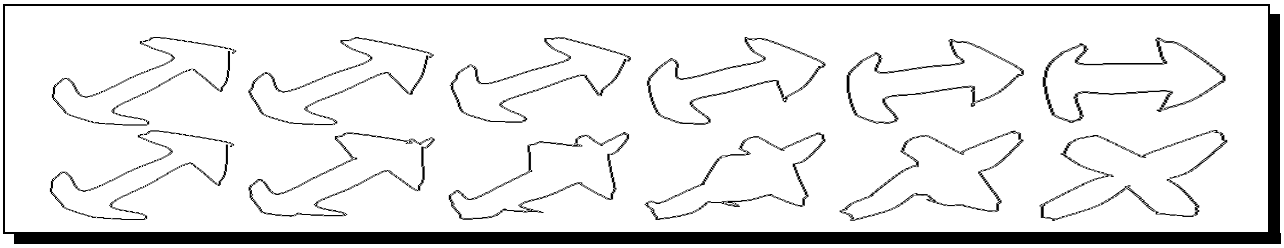

Figure 1: Physics-based metamorphosis of a shape to a similar shape (top row) and to a dissimilar shape (bottom row).

shapes as:

$$\begin{aligned}
\mathbf{S}(t) &= u\mathbf{S}^I + t\mathbf{S}^T \\
&= [uS_0^I + tS_0^T, uS_1^I + tS_1^T, \ldots, uS_n^I + tS_n^T] \\
&= [S_0(t), S_1(t), \ldots, S_n(t)] \qquad (1)
\end{aligned}$$

where $u = 1 - t$. $S_i(t)$ is the ith contour point in the intermediate shape, formed at time t. The time parameter t is normalized to the interval $[0, 1]$.

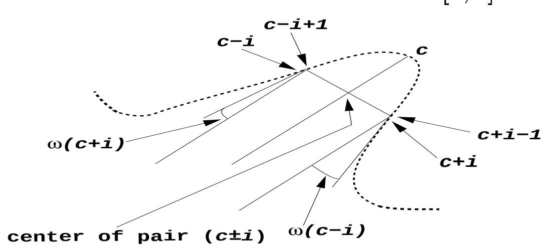

Figure 2: Geometric model for corner determination as proposed by Brault *et al.* [1].

We approximate the contour of an object by identifying certain salient points on it. Our segmentation strategy is based upon the algorithm of Brault *et al.* [1]. Specifically, the geometric parameters shown in Figs. 2 and 3 are calculated for each contour point. Following Brault *et al.* [1], a high curvature domain (corner) is defined for every contour point c as $c \pm i$ ($i = 1, 2 \ldots$). The point pair belongs to this domain *iff* the following inequalities are satisfied (see Fig. 2):

$$\omega(c+i) < \frac{\pi}{2} \quad and \quad \omega(c-i) < \frac{\pi}{2}. \qquad (2)$$

The *cornerness* of point c is computed as

$$TCF(c) = \sum_{i=1}^{M(c)} \cos(\omega(c+i)) * \cos(\omega(c-i)), \qquad (3)$$

where $c \pm 1, \ldots, c \pm n$ belong to the high curvature domain of point c. Innovatively, for each point c, we also define a low curvature domain in a manner conjugate to the method proposed for the determination of corners in [1]. The point pair $c \pm i$ belongs to the low curvature domain of c (see Fig. 3) *iff*

$$\omega(c+i) > \frac{\pi}{2} \quad or \quad \omega(c-i) > \frac{\pi}{2}. \qquad (4)$$

The *flatness* of point c is then given by

$$flatness(c) = \sum_{i=1}^{M(c)} \mid \cos(\omega(c+i)) \mid * \mid \cos(\omega(c-i)) \mid. \qquad (5)$$

where $c \pm 1, \ldots, c \pm n$ belong to the low curvature domain of point c. The points with the highest *cornerness* or *flatness* from each domain are selected. These points constitute the segmentation points for the contours. The algorithm is described in detail in [6].

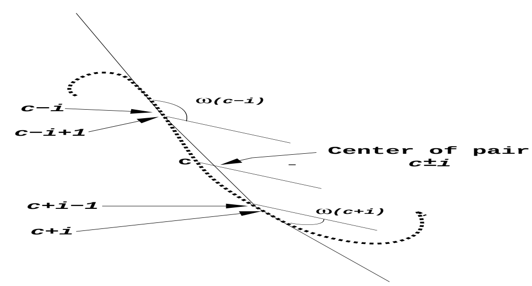

Figure 3: Geometric model for key low curvature point determination.

In general, the input and the target contours participating in the metamorphosis have different number of segmentation points. In order for the process of metamorphosis to occur, as detailed in Eq. (1), a point correspondence between the segmentation points in the input and the target needs to be established, wherein any point in the input shape corresponds to at least one point in the target shape and vice versa.

Such a correspondence can be one to one (point morphed to a point), one to many (point morphed

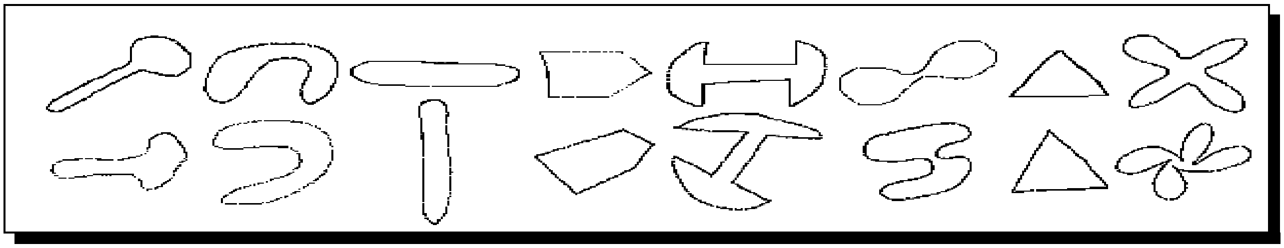

Figure 4: Templates (first row) and some of the test shapes (second row) used in experiment 1.

to segment(s)), many to one (segment(s) morphed to a point) or many to many (segment(s) morphed to segment(s)). Following the physics based formulation of metamorphosis [7], we define the cost of a point correspondence as the sum of stretching and bending energies required to bring about the correspondence. The stretching energy is computed for a pair of points and is defined as:

$$\begin{aligned} E_s &= \mid (f_s((L_T - L_O)^2 - (L_I - L_O)^2)) \div ((1 - c_s) \\ &\quad min(L_O, \ldots, L_I, L_T) + c_s \\ &\quad max(L_O, \ldots, L_I, L_T)) \mid . \end{aligned} \tag{6}$$

L_O, L_I, and L_T denote the segment lengths at the beginning, before the current deformation, and after the current deformation, respectively. The term c_s corresponds to the penalty for segments collapsing to points and f_s is the stretching stiffness parameter. The bending energy is computed for point triplets and denotes the cost of angular deformation. The bending energy is computed as:

$$E_b = \mid f_b[(\phi_T - \phi_O)^2 - (\phi_I - \phi_O)^2] \mid \tag{7}$$

where f_b indicates bending stiffness, ϕ_O represents the original angle, and ϕ_I and ϕ_T denote the angle before the current deformation and the angle after the curennt deformation, respectively.

The optimal point correspondence between two shapes is defined as the correspondence which requires the least energy in terms of Eqs. (6) and (7). If the metamorphosis is constrained to the segmentation points of the input and target shapes only (i.e., the energy minimization is considered only at the existing points), then the following optimal substructure property can be stated: The optimum cost of the point correspondence (S_i^I, S_j^T) equals the optimum cost of the previous point correspondence (S_{i-1}^I, S_j^T) or (S_{i-1}^I, S_{j-1}^T) or (S_i^I, S_{j-1}^T) and the cost of establishing the correspondence (S_i^I, S_j^T). Since the energy measures (Eqs. (6) and (7)) used in the metamorphosis depend on the relative position of the contour points of an object and not their absolute position, they are invariant to translation and rotation. Invariance to scale changes is achieved by mapping the shapes to a unit square.

Formally, we define the *degree of morphing* between two shapes A and B as:

$$D_{morph}(A, B) = \min_{\Omega} \min_{\Lambda} E(C_A, C_B), \tag{8}$$

where C_A and C_B are the contour points of shapes A and B and $E(C_A, C_B)$ is the cumulative energy spent in morphing shape A to shape B. Ω denotes the set of all starting point correspondences between shapes A and B. Λ denotes the four possible ways of traversing contours C_A and C_B. Since the contours of planar objects are simple curves and therefore can be traversed either clockwise or anticlockwise, invariance to mirror symmetry is obtained by minimizing $E(\cdot, \cdot)$ over Λ. $D_{morph}(A, B)$ satisfies all the properties of a metric. To prove the triangle inequality $D_{morph}(A, B) + D_{morph}(B, C) \geq D_{morph}(A, C)$, we may observe that the energy terms E_s and E_b (Eqs. (6) and (7)) are always positive. Furthermore, in the case of E_b any corresponding angle change between shapes A, B, and C can be either monotonic ($\Theta^A > \Theta^B > \Theta^C$) or non-monotonic ($\Theta^B > \Theta^A > \Theta^C$). In the former case, we have equality while in the latter case the inequality holds true. A similar argument can be given for E_s (length changes) and hence for the sum of E_b and E_s. The other three properties of a metric are straightforward from the formulation of the *degree of morphing*.

3 Experimental Results

The proposed method was evaluated with three data sets. The first set consisted of eight shapes taken from [2]. For each shape, 22 samples were created by different people using a WACOM tablet. Participants were shown each one of the eight shapes in turn and then asked to generate the samples without looking at the shapes. A single carefully hand drawn shape from each class was used as a template. Of the 176 test

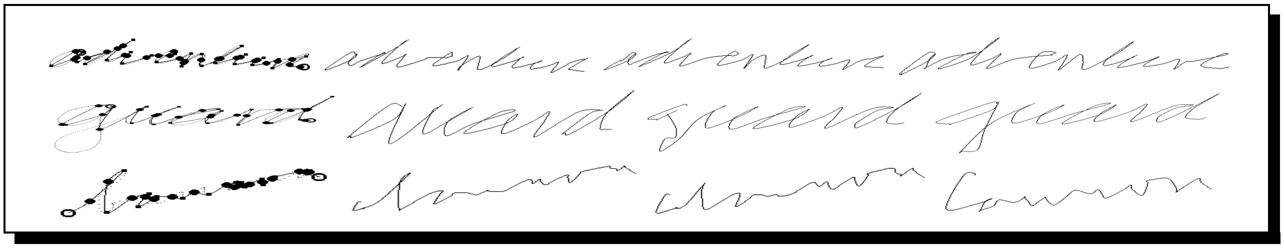

Figure 5: Metamorphosis of test samples of the words *adventure*, *guard*, and *banana*. The input words are shown with the segmentation points superimposed.

shapes, the proposed system correctly recognized 161 shapes (91.48% recognition rate). Fig. 4 shows the templates and some test samples which were correctly recognized.

In the second experiment, the method was used for the recognition of unconstrained cursive handwriting in a user dependent, on-line setting.

Four users participated in the experiment. A vocabulary of the following ten randomly selected words was used: *adventure, banana, bookshelf, cannon, flywheel, guard, landmark, mad, peach, tooth.* Each user provided one template sample and four test samples per word. The data was collected using the same setup as in the previous experiment. The recognition rates for each of the participants were: user 1 (100%), user 2 (90%), user 3 (92.5%), and user 4 (97.5%). Fig. 5 shows examples of recognition experiments for the words *adventure*, *guard* and *banana* taken from user 4. The last case shows a misrecognition caused due to loss of important feature points throughout the word. The last experiment concerns the recognition of planar industrial parts and objects. The objects were captured at different orientations and locations. An arbitrary example from each object class was chosen as a template. For a database of 21 templates and with 4 test cases per template a recognition rate of 97.62% was obtained.

4 Conclusions and Future Work

In this paper we discuss the use of shape metamorphosis as a paradigm for recognition of 2D shapes. The shape contours are modeled as a piece of wire and the amount of energy spent in converting one wire-form shape to another is used to quantify the shape difference. The proposed recognition measure can be used for both convex and concave shapes, furthermore it is invariant to translation, rotation, scale changes, mirror symmetry and satisfies the properties of a metric. Our future work is directed towards developing the paradigm to deal with contour occlusions and incorporating shape statistics to define the shape representation.

Acknowledgements

This research was supported by the National Science Foundation through Grants #IRI-9410003 and #IRI-9502245, the McKnight Land-Grant Professorship Program of the University of Minnesota, and the Department of Energy (Sandia Labs) through Contracts #AC-3752D and #AL-3021.

References

[1] J.J. Brault and R. Plamondon. "Segmenting Handwritten Signatures at Their Perceptually Important Points". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):953–957, 1993.

[2] Y. He and A. Kundu. "2D Shape Classification Using Hidden Markov Model". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11):1172–1184, 1991.

[3] J. Hu and M.K. Brown. "HMM-Based On-Line Handwriting Recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1039–1045, 1996.

[4] V. Kovalev and M. Petrou. "Multidimensional Cooccurrence Matrices for Object Recognition and Matching". *Graphical Models and Image Processing*, 58(3):187–197, 1996.

[5] S.K. Nayar, S.A. Nene, and H. Murase. "Real-Time 100 Object Recognition System". In *Proceedings 1996 IEEE International Conference on Robotics and Automation*, volume 3, pages 2321–2325, 1996.

[6] I. Pavlidis and N.P. Papanikolopoulos. "A Curve Segmentation Algorithm That Automates Deformable-Model Based Target Tracking". Technical Report TR 96-041, University of Minnesota, 1996.

[7] T. W. Sederberg and E. Greenwood. "A Physically Based Approach to 2D Shape Blending". *Computer Graphics*, 26(2):25–34, 1992.