An On-Line Handwritten Note Recognition Method Using Shape Metamorphosis

Ioannis Pavlidis   Rahul Singh   Nikolaos P. Papanikolopoulos

Honeywell Technology Center
Honeywell Inc.
Minneapolis, MN 55418
pavlidis@htc.honeywell.com

Dept. of Computer Science
University of Minnesota
Minneapolis, MN 55455
singh@cs.umn.edu

Dept. of Computer Science
University of Minnesota
Minneapolis, MN 55455
npapas@cs.umn.edu

Abstract

We propose a novel user-dependent method for the recognition of on-line handwritten notes. The method employs as a dissimilarity measure the "degree of morphing" between an input curve and a template curve. A physics-based approach substantiates the "degree of morphing" as a deformation energy and casts the problem as an energy minimization problem. The method operates upon key segmentation points that are provided by an appropriate segmentation algorithm. The segmentation objective is not to locate letters, but instead to locate corners and some key low curvature points (an easier task). This is part of the method's strategy to see the word as a generic on-line curve. Due to this strategy, the proposed method can handle collectively both, cursive words and hand-drawn line figures, the two key ingredients of handwritten notes. Most importantly, the proposed system achieves high recognition rates without ever resorting to statistical models.

1   Introduction

In this paper we introduce a framework for handling collectively the recognition of on-line handwritten note patterns in a user-dependent setting. A preliminary version of this work was reported in [9]. We propose the use of shape metamorphosis as a potentially powerful way of dealing with the recognition of complex hand-drawn, on-line patterns. Shape metamorphosis is a well established graphics technique [11] that refers to the problem of computing a continuous shape transformation from an initial shape to a target shape. The proposed system is based on the fact that two similar shapes do not undergo extensive metamorphosis if one of them is morphed to the other. Thus, the *degree of morphing* between a test pattern and a reference pattern may serve as the matching

criterion. We compute the *degree of morphing* using a physics-based approach to shape metamorphosis first proposed by T. W. Sederberg *et al.* [11], for computer based animation. The framework proposed by us includes modifications to Sederberg's method [11], so as to deal with issues specific to the problem of on-line hand-drawn shape recognition.

Most of the current research in handwritten note recognition deals with cursive script recognition [3, 7, 10, 12] and on-line hand-drawn line figure recognition [5, 6], in an independent manner. In our approach, both cursive words and line figures are considered in the unifying framework of an on-line (possibly discontinuous) curve. The recognition task is thus reduced to the computation of the similarity of these curves. Furthermore the use of a dynamic programming approach to compute the degree of morphing allows for the incorporation of local deformations in computing a global measure of dissimilarity.

The metamorphosis-based shape recognition system consists of three modules: Shape Sampling and Preprocessing, Shape Segmentation, and Shape Metamorphosis. Preprocessing is performed according to the work in [1] and [4]. The segmentation algorithm is described in Section 2. Section 3 unveils the shape metamorphosis method and its use in the pattern recognition context. In Section 4, the results from experimental tests are presented and discussed. Finally, in Section 5 the paper is summarized, conclusions are drawn, and future work is outlined.

2   Shape Segmentation

Morphing a curve by involving all its constituent pixels is both computationally expensive and prone to pixel level variations caused by quantization errors. We use an alternate representation wherein, a curve is modeled by identifying its perceptually important

points (*corner points*), along with points identifying regions of slowly varying curvature (*key low curvature points*). The detection of *corner points* is done by using the algorithm suggested by Brault *et al.* [2]. By using criteria conjugate to that of locating corners, the regions of the curve with a slowly varying curvature are identified. These regions are represented by placing a *key low curvature point* in their middle. A detailed account of this segmentation approach can be found in [8].
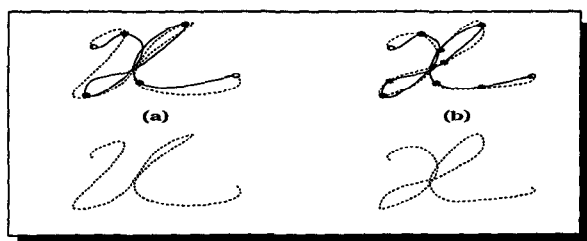

Figure 1: (a) Reconstruction of the handwritten character "x" using corner points only. (b) Reconstruction of the handwritten character "x" using corner and key low curvature points.

The representation of hand-drawn curves, by interleaving points of high curvature with points of low curvature not only facilitates their reconstruction using a small number of points (see Fig. 1), but also provides a consistent representation pattern (a *key low curvature point* between every two *corner points*), which facilitates the metamorphosis process.

3 Shape Metamorphosis

Shape metamorphosis is defined as the transformation of one shape (initial) to another shape (target). If P_i ($i = 0, 1, \ldots, n$) denote the segmentation points of the shape $\mathbf{S}$, then the shape $\mathbf{S}$ could be represented in vector form as

$$\mathbf{S} = [P_0, P_1, \ldots, P_n]. \tag{1}$$

In our work, metamorphosis of an initial shape $\mathbf{S}^I$ to a target shape $\mathbf{S}^T$ involves a linear interpolation between the corresponding segmentation points of the two shapes. This interpolation could be discretely timed to give a series of intermediate shapes.

Since most initial and target shapes do not have the same number of segmentation points, establishing a point correspondence between the shapes is a central issue in shape metamorphosis. An important constraint in obtaining the necessary correspondences is the avoidance of intermediate *chaotic shapes*. Such

shapes are obtained if, similar parts between the initial and the target shapes are not maintained during metamorphosis. As our approach is based on relating the *degree of morphing* between two on-line curves to the similarity between their shapes, maintaining similar parts between the initial and final shapes is essential. An intuitive consequence of this strategy would be the lack of any metamorphosis, if the initial and target shapes are identical.

Our approach to establish the point correspondences is based on the work of Sederberg *et al.* [11]. All the possible correspondences between the n segmentation points of the initial shape and m segmentation points of the target shape are represented by an $m \times n$ matrix. Each shape is considered to be made of virtual wires. The metamorphosis of one shape to another takes place through stretching and bending of the respective wireform shapes. The cost of a point correspondence is equal to the virtual energy spent in deforming the wire to realize that correspondence. The computation of the optimal point correspondence can, therefore, be regarded as an optimization problem which involves associating each point of the initial curve to at least one point of the target curve and vice-versa, by using minimal energy. A quadratic time (in the number of segmentation points) solution to this problem can be found using dynamic programming.

The main goal in [11], was to produce visually pleasing metamorphoses with the least possible human (animator) intervention. A visually pleasing metamorphosis is achieved by maintaining the similar parts between the initial and the target shapes and smoothly deforming the dissimilar parts. At the intuitive level, the requirements for a visually pleasing metamorphosis appear to facilitate also the production of a *non-chaotic* metamorphosis. In other words, Sederberg's method [11] tends to link the *degree of morphing* with the shape dissimilarity, which is our main goal (pattern recognition).

3.1 Metamorphosis-based shape matching

A direct application of Sederberg's approach [11] to quantify the differences between hand-drawn shapes is precluded due to the following reasons:

1. It does not elaborate on how to automatically select points for representing a shape.

2. The computation of the deformation energy (stretching and bending) depends on a number of parameters, the values of which need to be specified appropriately if we are to obtain the desired

915

non-chaotic behavior. We have found, in particular, that special care should be given to the fine tuning of the bending energy parameters.

3. It does not include any knowledge specific to the handwriting domain. Specifically, we have found that the response of the deformation measure becomes more robust if we take into account the following fact: two different versions of the same shape (word or figure), if they have been written by the same user, don't have radically different numbers of segmentation points. This is emphasized in our method by the introduction of an additional term in the energy formula.
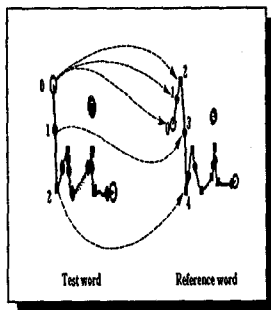

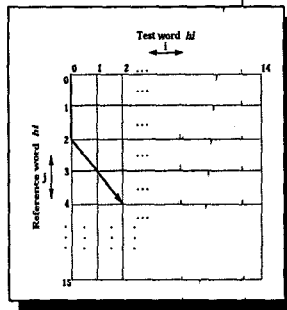**Figure 2:** Optimal point correspondence for the handwritten word *hi*.

Figure 3: Matrix representation of all possible point correspondences.

The automatic selection of shape points has been solved by the introduction of the segmentation algorithm [8] described in Section 2. The other two issues have been addressed by the introduction of two appropriate modifications in the original physics-based metamorphosis method [11]. Specifically, the proposed metamorphosis-based matching works as follows: It employs the dynamic programming technique suggested in [11] to find the optimal point correspondence set. It is the by-product of this process (i.e., the cost of the optimal correspondence) which is really used for the shape matching task. Fig. 2 shows partially the optimal point correspondence set between a test sample and the prototype of the handwritten word *hi*. The beginning of the reference word differs from the beginning of the test word in that it features an extra stroke. Fig. 2 shows how the dynamic programming algorithm optimally solves the local imbalance in points by corresponding point 0 from the test word to points 0, 1, and 2 in the reference word. Fig. 3 shows a part of the the optimal point correspondence set (bold arrow) for the words in Fig. 2.

The method responds predictably as the shape dissimilarity increases by producing proportionally greater energy deformation values. We have found, however, that the response of the system becomes more robust by augmenting the cost (deformation energy E_{total}) of the optimal correspondence set in the following way:

$$E_a = E_{total} + 5* \mid n - m \mid \qquad (2)$$

where n is the number of segmentation points of the test shape, m is the number of segmentation points of the reference shape, and E_a is the augmented energy requirement. The weight factor 5 in Eq. (2) has been chosen for its adequate experimental behavior. The additional term added to the original energy requirement emphasizes the fact that different versions of the same shape (either cursive word or line figure) don't have radically different numbers of segmentation points.

The point correspondence cost is materialized in terms of stretching and bending energy. Following the formulation by Sederberg *et al.* [11], the stretching energy for a piece of virtual wire is given by the equation:

$$E_s = f_s \frac{\mid L_T - L_I \mid^2}{(1 - c_s)min(L_I, L_T) + c_s max(L_I, L_T)}, \qquad (3)$$

where L_I is the initial length of the wire and L_T is its length after the deformation. The term c_s (set to 0.5) is a user-definable parameter which controls the penalty for segments of the wire that collapse to points during the metamorphosis. Finally, f_s (set to 100.0) is the user-definable stretching stiffness parameter. Different values in the user-definable parameters give different interpretations in the meaning of minimum energy. Since we are interested in linking minimum energy metamorphosis to shape dissimilarity, a good way to test our minimum energy model is to see how it behaves in the case of two very similar shapes (see Fig. 4). The correct model should give almost imperceptibly different in-between shapes and it should yield almost zero energy expenditure. We have experimentally determined that this is always the behavior of the model only when the wire is made stiffer with respect to bending and looser with respect to stretching.

The bending energy equation is formulated analogously to the stretching energy equation. The notable difference is that bending deforms angles and not wire segments and hence applies to consecutive triples of segmentation points instead of pairs of segmentation
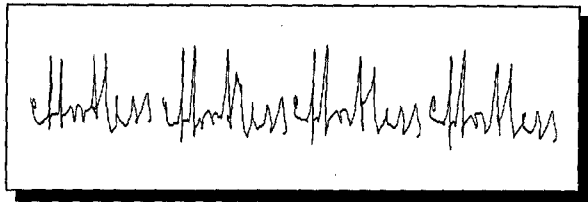
Figure 4: Metamorphosis of a test sample of the word *effortless* to its prototype sample kept in the database (user 1). The metamorphosis is an example of a minimal deformation between similar handwritten words and produces the correct matching.

points. More specifically, the bending energy is given by the equation

$$
E_b = \begin{cases}
f_b(\Delta\phi + m_b\Delta\phi^*)^2 + p_b \\
\quad \text{if } \phi(t) \quad \text{does go to zero,} \\
f_b(\Delta\phi + m_b\Delta\phi^*)^2 \\
\quad \text{if } \phi(t) \quad \text{never goes to zero}
\end{cases} \tag{4}
$$

where f_b indicates bending stiffness (set to 0.02), m_b penalizes bending angles which deform non-monotonically (set to 100.0), and p_b penalizes bending angles from going to zero (set to 1000.0). The term $\Delta\phi$ represents the bending angle change in radians due to the metamorphosis of the particular triplet of segmentation points. $\Delta\phi^*$ is a non-negative quantity that expresses in radians how much the bending angle change deviates from monotonicity. The severe penalties expressed by f_b and m_b account mostly for the restrictions placed in the amount of bending involved in each metamorphosis.

The method as it is described applies to singly connected shapes only (no pen-up movements). A writer can break a word or a figure in variable points from writing to writing. The problem is, to have the test and reference shapes have the same number and arrangement of segments. Only then, similar shapes will appear similar to the metamorphosis method too.

Because for cursive words and typical line figures, the number of segments is very small (usually not more than 6 − 8 segments per shape), we opted to solve the problem in the following way: we compute the degree of morphing as the minimum energy value over all possible fusion combinations of consecutive segments. Delayed strokes are handled in a special manner. They are connected from left to right and they constitute always a single segment labeled as *delayed*.

4 Experimental Results

The experimental setup, in which the above method has been implemented, consists of a graphics work-

station (SGI Indigo^{TM}R4000) and a graphics tablet (WACOM $UD - 0608R$). Twelve users have participated in the experiments. A reference database of one hundred cursive words, and seven hand-drawn line figures has been established from each user. For each shape in the reference databases, four test shapes have been collected at different days and times over a period of four months. Fig. 5 shows handwritten samples from four different users. The test shapes from each user were matched against the corresponding reference database only (user-dependent system). The results of the experiment are presented in Table 1.

Users	Success Rate
1	99.0%
2	90.1%
3	91.8%
4	91.3%
5	93.4%
6	90.2%
7	89.9%
8	92.0%
9	86.2%
10	91.6%
11	89.7%
12	93.2%

Table 1: Test results.

The method appears to be ideal for interpreting online notes that include cursive words and simple line figures. Fig. 6(a) for example, shows the handwritten block diagram of a typical low-level vision system as it was produced by user 1. Fig. 6(b) shows the interpretation our system gave to Fig. 6(a) after having successfully matched each word and line figure in the sketch.
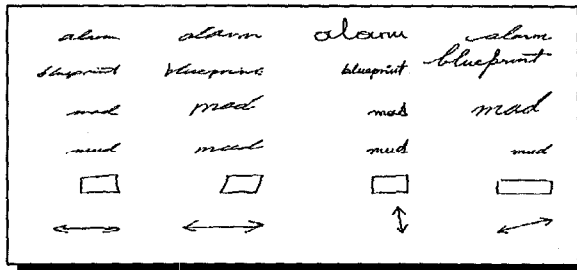

Figure 5: Handwritten samples from users 9, 12, 1, and 5.

Two major sources of failure were identified: One source of failure was poor preprocessing performance,

917

especially in words which are very look alike (e.g., *mad* and *mud*). The other source of failure was poor postprocessing performance of the segment fusion algorithm in the case of line figures. The time sequence of the various segments in certain line figures can vary from drawing to drawing. For example, the double arrow was found to be drawn by a user one time as *arrow tip - line - arrow tip* and another time as *arrow tip - arrow tip - line*. The current segment fusion algorithm always preserves the time sequence of the segments (except of delayed strokes). Consequently, it is unable to handle the change in the time sequence observed sometimes in the case of the double arrow.
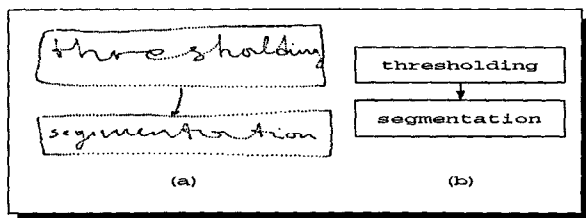

Figure 6: (a) Handwritten block-diagram from user 1. (b) Interpretation by the metamorphosis-based system.

5 Conclusions and Future Work

A novel method for the recognition of on-line handwritten notes (cursive words and line figures) has been described. The strong points of the method are:

1. robust behavior verified experimentally,
2. real-time performance,
3. single reference samples from each shape,
4. collective handling of words and figures.

Currently, the weakest point of the method appears to be the unidirectional flow of information from the preprocessing stage towards the metamorphosis stage. For example, the presence of multiple close matching scores may indicate the need to adjust the preprocessing to capture more discriminant features. The metamorphosis-based matching proved to be as good as the quality of information that is provided.

The segment fusion algorithm currently in use, is better suited for cursive words than 2-D on-line figures. Further research is needed towards more effective ways of handling the segment correspondence problem in the case of line figures.

Acknowledgements

References

[1] H. Arakawa. "On-Line Recognition of Handwritten Characters - Alphanumerics, Hiragana, Katakana, Kanji". *Pattern Recognition*, 16(1):9–16, 1983.

[2] 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.

[3] M. K. Brown and S. Ganapathy. "Cursive Script Recognition". In *Proceedings of the International Conference on Cybernetics and Society*, pages 47–51, 1980.

[4] W. Guerfali and R. Plamondon. "Normalizing and Restoring On-Line Handwriting". *Pattern Recognition*, 26(3):419–431, 1993.

[5] W.C. Lin and J.H. Pun. "Machine Recognition and Plotting of Hand-Sketched Line Figures". *IEEE Transactions on Systems, Man, and Cybernetics*, pages 52–57, 1978.

[6] H. Murase and T. Wakahara. "On-Line Hand-Sketched Figure Recognition". *Pattern Recognition*, 19:147–160, 1986.

[7] K.S. Nathan, H.S.M. Beigi, S. Jayashree, G.J. Clary, and H. Maruyama. "Real-Time On-Line Unconstrained Handwriting Recognition Using Statistical Methods". In *Proceedings 1995 International Conference on Acoustics, Speech, and Signal Processing*, pages 2619–2622, 1995.

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

[9] I. Pavlidis, R. Singh, and N. Papanikolopoulos. "Recognition of On-Line Handwritten Patterns Through Shape Metamorphosis". In *Proceedings 13th International Conference on Pattern Recognition*, pages 18–22, 1996.

[10] L. Schomaker. "Using Stroke- or Character-Based Self-Organizing Maps in the Recognition of On-Line, Connected Cursive Script". *Pattern Recognition*, 26(3):443–450, 1993.

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

[12] G. Seni, R.K. Srihari, and N. Nasrabadi. "Large Vocabulary Recognition of On-Line Handwritten Cursive Words". *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 18(7):757–762, 1996.