Identifying Features via Homotopy on Handwritten Mathematical Symbols

Rui Hu, Stephen M. Watt

Department of Computer Science
Western University
London Ontario, Canada
Outline

- Introduction
- Digital Ink Approximation
- Determining Points
- Determining Point Algorithm
- Experimental Results
- Conclusion and Future Work
Handwriting Mathematics

- Handwriting mathematics recognition is different:
  - Symbols are taken from many alphabets.
  - Symbols are written in a two dimensional layout in various sizes.
  - Symbol size and layout carry meaning.
  - No fixed dictionary of words to help with disambiguation.

\[ e^x = \int e^x \, dx = \sum_{i=0}^{\infty} \frac{x^i}{i!} \]
Baseline Estimation

- Mathematical handwriting does not follow simple baselines.

\[ C_1 x^2 + C_2 x \]

- We can locate some important features by identifying special points.

We refer to a point such as this, that determines the height of a metric line, as a *determining point*. 
Motivation

- Juxtaposition ambiguity

- Handwriting neatening

\[ a_1 x^2 + a_2 \rightarrow a_1 x^2 + a_2 \]
Earlier Work

- We have previously addressed the problem of how to find determining points automatically. [CICM’13]

- The algorithm derived determining points on new samples from a reference symbol of the same type.

- For samples that are significantly different from the reference symbol, one can use a numerical homotopy between the reference symbol and the target.
Objectives

- We wish to investigate the choice of the homotopy method and the efficiency implications of the choices.

- We examine two strategies: *average symbol* vs *nearest neighbor*.

- We ask these questions:
  1) What are the differences between starting the homotopy from average symbols vs nearest labelled neighbor?
  2) What is the relationship between the distance from the sample to the reference symbol and the number of homotopy steps?
  3) How can we best use the results to find determining points in new samples?
Digital Handwriting

- Represented as a sequence of points 
  \((x_0,y_0), (x_1,y_1), (x_2,y_2)\)...
Decomposition of channels

- We consider an ink trace as a segment of plane curve \((X(s), Y(s))\), parameterized by arc length.
- A function can be approximated with orthogonal polynomials
  \[ f(s) \approx \sum_{i=0}^{d} c_i P_i(s) \]
- We approximate \(X(s)\) and \(Y(s)\) and obtain
  \[ C_0^X, C_1^X, \ldots, C_d^X, C_0^Y, C_1^Y, \ldots, C_d^Y \]
Determining Points

- We focus on European alphabets.
- We consider 6 types of determining points.

\[ ax + by = c \]
Average Symbol

- The average symbol of a set of known samples for a class can be computed as the average point in the functional space,

\[
\bar{C} = \sum_{i=1}^{n} \frac{C_i}{n}
\]
Distance Between Curves

\( \bar{x}(t) = x(t) + \xi(t) \quad \xi(t) = \sum_{i=0}^{\infty} \xi_i \phi_i(t), \quad \phi_i \text{ ortho on } [a, b] \text{ with } w(t) = 1. \)

\( \bar{y}(t) = y(t) + \eta(t) \quad \eta(t) = \sum_{i=0}^{\infty} \eta_i \Phi_i(t) \)

\[
\rho^2(C, \bar{C}) = \int_{a}^{b} \left[ (x(t) - \bar{x}(t))^2 + (y(t) - \bar{y}(t))^2 \right] dt
\]

\[
= \int_{a}^{b} \left[ \xi(t)^2 + \eta(t)^2 \right] dt
\]

\[
\approx \int_{a}^{b} \left[ \sum_{i=0}^{d} \xi_i^2 \phi_i^2(t) + \text{cross terms} + \sum_{i=0}^{d} \eta_i^2 \phi_i^2(t) + \text{cross terms} \right] dt
\]

\[
= \sum_{i=0}^{d} \xi_i^2 + \sum_{i=0}^{d} \eta_i^2
\]
Nearest Neighbor

• Similar samples tend to have a shorter Euclidean distance in the functional space.

• Deriving determining points from a similar reference symbol leads to a lower measured error.
Deriving from a Reference Symbol [c.f. CICM]

Reference Symbol

Sample-1-Initial  Sample-2-Initial

Arc-length guess

Sample-1-Derived  Sample-2-Derived

Optimization
A Homotopy Method [SYNASC]

- Some samples are far away from the reference symbol.

- We use a homotopy between the reference symbol and the target sample in a multi-step method.
Dataset

- 240 symbols
- 388 classes
- 45637 classified samples
Evaluation I: Building Ground Truth

- We first computed the average symbol of each class.
- On each average symbol the determining points were manually identified.
- We then used this information to find corresponding determining points in all samples using a 4 step homotopy.
- We then visually inspected the determining points in each of the samples, and manually adjusted the few incorrect ones.
- This collection of samples with corrected determining points then served as the ground truth.
Evaluation II: Measuring no. steps needed

- We went back and tested the data by forgetting the marked points and seeing how many steps it took to recover them.
- For each sample, we tried 1 step, 2 steps, etc up to 30, until a correct result was achieved.
- Samples that did not converge with a 30 step homotopy were considered to fail.
- We examined the relationship between the distance to the reference symbols and the number of steps the homotopy required.
Average Symbol $\Delta s = 0.02, \Delta y = 1\%$

The overall success rate is 99.57%. Each distance interval in the dense area ($x \leq 0.59$) contains 2800 samples while each interval in the sparse area ($x > 0.59$) contains 1500 samples except the last one contains 1340 samples.
The overall success rate is 99.63%. Each distance interval in the dense area (x≤0.59) contains 2800 samples while each interval in the sparse area (x>0.59) contains 1500 samples except the last one contains 1370 samples.
The overall success rate is 99.69%. Each distance interval in the dense area (x≤0.59) contains 2800 samples while each interval in the sparse area (x>0.59) contains 1500 samples except the last one contains 1396 samples.
Nearest Neighbor $\Delta s = 0.02$, $\Delta y = 1\%$

The overall success rate is 98.86%. Each distance interval in the dense area ($x \leq 0.32$) contains 2800 samples while each interval in the sparse area ($x > 0.32$) contains 1500 samples except the last one contains 1016 samples.
The overall success rate is 99.45%. Each distance interval in the dense area (x≤0.32) contains 2800 samples while each interval in the sparse area (x>0.32) contains 1500 samples except the last one contains 1284 samples.
Nearest Neighbor $\Delta s = 0.02, \Delta y = 7\%$

The overall success rate is 99.68%. Each distance interval in the dense area ($x \leq 0.32$) contains 2800 samples while each interval in the sparse area ($x > 0.32$) contains 1500 samples except the last one contains 1390 samples.
Conclusion

- We have examined two strategies: *average symbol* vs *nearest neighbor*.
- Both achieved very high success rate (> 99%) by choosing $\Delta y > 1\%$.
- We have found a clear relation between the distance to reference symbols and the required number of homotopy steps.

- The correlation of distance to the number of steps was greater using the average symbol strategy.
- The fewer steps were needed with the nearest neighbor strategy. Nearest neighbor strategy has a slightly better predictability.

- Overall, the nearest neighbor strategy with suitable error tolerances gives the most efficient performance with no significant loss of accuracy.