Model-based On-line Handwritten Digit Recognition

Xiaolin Li, Réjean Plamondon
Laboratoire Scribens
École Polytechnique de Montréal
Montréal (PQ), Canada H3C 3A7
{xiaolin, rejean}@scribens.polymtl.ca

Marc Parizeau
Département de Génie Electrique
Université Laval
Ste-Foy (PQ), Canada G1K 7P4
parizeau@gel.ulaval.ca

Abstract

This paper presents a hidden Markov model (HMM) based approach to on-line handwritten digit recognition using stroke sequences. In this approach, a character instance is represented by a sequence of symbolic strokes, and the representation is obtained by component segmentation and stroke classification. The component segmentation is based on the delta lognormal model of handwriting generation. The symbolic strokes are used for HMM multiple observation training or recognition. A training and recognition experiment has been conducted using the above techniques.

1. Introduction

On-line handwriting recognition is an area of active research that aims at providing a dynamic means of communication with computers through a digitizing tablet and an electronic pen [9]. On-line handwritten scripts captured by the digitizing tablet consist of sequences of components that are pen-tip traces from pen-down to pen-up positions. As the shape of a script depends on pen-tip traces, many researchers use equally distributed points along components to characterize a script [1, 9], while others use characteristic points to segment a script into basic elements [4, 5]. This paper presents an HMM based approach to on-line handwritten digit recognition using stroke sequences. In this approach, a character instance is represented by a sequence of symbolic strokes, and the representation is obtained by component segmentation and stroke classification. The component segmentation is based on the delta lognormal model of handwriting generation [6, 7]. And the symbolic strokes are used for HMM multiple observation training or recognition. The details of these techniques and the experimental results on handwritten digits are given in the following sections.

2. Model based segmentation and stroke classification

2.1. Model based segmentation

The kinematic theory is a powerful method for analyzing rapid human movements [6, 7]. It describes a neuromuscular synergy involved in the production of these movements in terms of the agonist and antagonist systems [6]. With respect to handwriting generation, a simple stroke is controlled by a velocity vector:

$$\overline{v}(t-t_0) = < v_\xi(t), \theta(t) >, \quad t_0 \leq t \tag{1}$$

where $v_\xi(t)$ is the magnitude and $\theta(t)$ is the angular direction of the velocity such that

$$\begin{aligned} v_\xi(t) &= D_1\Lambda(t; t_0, \mu_1, \sigma_1^2) - D_2\Lambda(t; t_0, \mu_2, \sigma_2^2) \\ \theta(t) &= \theta_0 + c_0 \int_{t_0}^{t} v_\xi(u)du \end{aligned} \tag{2}$$

where θ_0 is the initial angular direction, c_0 is a constant representing the curvature of the stroke, and

$$\begin{aligned} \Lambda(t; t_0, \mu_i, \sigma_i^2) \quad = \quad & \frac{1}{\sigma_i\sqrt{2\pi}(t-t_0)} \\ & \exp\{-\frac{[\log(t-t_0)-\mu_i]^2}{2\sigma_i^2}\} \end{aligned} \tag{3}$$

is a lognormal function. In the above equations, t_0 represents the activation time, D_1 and D_2 are the amplitude of the impulse commands, μ_1 and μ_2 are the log-time delays, σ_1 and σ_2 are the log-response time of the agonist and antagonist systems, respectively.

From the above definition, it readily follows that given $v_\xi(t)$, the shape of a stroke image is an arc. Furthermore, let $\overline{v}^{(1)}(t-t_0^{(1)})\overline{v}^{(2)}(t-t_0^{(2)})\cdots\overline{v}^{(n)}(t-t_0^{(n)})$ be an n-stroke sequence, where the ith stroke is represented as $\overline{v}^{(i)}(t-t_0^{(i)}) = < v_\xi^{(i)}(t), \theta^{(i)}(t) >$ with $t_0^{(i)} \leq t$. Then the movement of a component is the vectorial summation of the n strokes in the time domain:

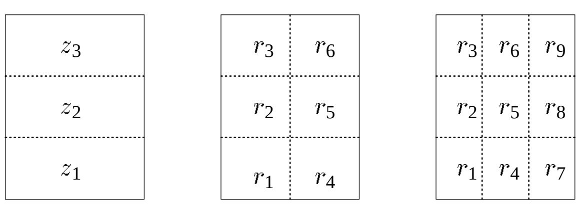
$$\overline{v}(t) = \sum_{i=1}^{n} \overline{v}^{(i)}(t - t_0^{(i)}) \tag{4}$$

Theoretically, in the above summation, each stroke in the sequence has an effect on all subsequent strokes. Since the n strokes are superimposed on one another, it is very difficult to recover them given only $\overline{v}(t)$. However, in practice, using the extrema of the static curvature $c = d\theta/d\xi$, we can segment the trace of a component into a sequence of static strokes (arcs) in the spatial domain instead of in the time domain. The details of the above technique can be found in [4]. Figure 1 shows examples of segmentation using this technique.

Figure 1. Extrema of curvature and inflection point

After component segmentation, strokes contained in a character can be classified into symbol categories.

2.2. Stroke classification



(a) 3 zones (b) 6 regions (c) 9 regions

Figure 2. Using zones and regions for symbol generation

Although handwriting signals are sequential signals, any handwritten character has a 2D shape which depends on the positions of individual components on the X-Y plane. To incorporate the positional information and produce appropriate symbols for HMM training and recognition, we center a character instance within a box and divide the box into different zones or regions (see Figure 2). Then we associate 66 symbols with 3 zones, 96 symbols with 6 regions, and 225 symbols with 9 regions using not only the location but also the curvature information of a stroke. In this manner, we can classify a stroke into a symbol category.

3. HMM Multiple Observation Training

Hidden Markov models have been successfully used in speech recognition since the 1980s [3, 8]. In recent years they have also been applied to on-line handwriting recognition [1, 2]. In this application, we use the left-right models to represent on-line handwritten character classes.

Let $S = \{S_1, S_2, \cdots, S_N\}$ be a set of hidden states and $V = \{v_1, v_2, \cdots, v_M\}$ be a set of observation symbols. Then a hidden state sequence $Q = q_1 q_2 \cdots q_T$ has its elements in S and a corresponding observation $O = o_1 o_2 \cdots o_T$ has its elements in V. Denote an HMM as $\lambda = (A, B, \pi)$ where $A = \{a_{ij} | 1 \leq i, j \leq N\}$ is the state transition distribution matrix, $B = \{b_j(k) | 1 \leq j \leq N, 1 \leq k \leq M\}$ is the symbol emission distribution matrix, and $\pi = \{\pi_i | 1 \leq i \leq N\}$ is the initial state distribution matrix. Given a set of observations $\mathbf{O} = \{O^{(1)}, O^{(2)}, \cdots, O^{(K)}\}$ from a character class, we use Levinson's equations [3] for HMM multiple observation training:

$$\overline{a}_{mn} = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T_k - 1} \xi_t^{(k)}(m, n)}{\sum_{k=1}^{K} \sum_{t=1}^{T_k - 1} \gamma_t^{(k)}(m)} \tag{5}$$

$$\overline{b}_n(m) = \frac{\sum_{k=1}^{K} \sum_{t=1, o_t^{(k)} = v_m}^{T_k} \gamma_t^{(k)}(n)}{\sum_{k=1}^{K} \sum_{t=1}^{T_k} \gamma_t^{(k)}(n)} \tag{6}$$

where $\xi_t^{(k)}(i, j)$ is the joint event and $\gamma_t^{(k)}(i)$ is the the state variable [3, 8] associated with the kth observation.

4. Training and Recognition Experiment

4.1. Data set

The data set involved here was unipen1a, which was from train_r01_v0, the fifth release of the international UNIPEN training data sets. This data set contained 6519 digits (0-9) and there were different writing styles in each digit class. In this experiment, about half of the data of each character class was randomly chosen for model training, and the other half was used for model testing.

4.2. Data preprocessing and observation extraction

We used the techniques described in [4] to segment components of character instances into strokes. Strokes were then classified into corresponding symbol categories.

4.3. Model training

In this experiment, with respect to each zone-region representation method (see Figure 2), we used just one left-right HMM per digit class. The number of training samples per class was 300, and the number of hidden states in all HMMs was fixed at $N = 6$. The convergence of the training process was judged by a small threshold $\epsilon = 0.00001$.

4.4. Digit recognition

We performed character recognition using the following steps and rules:

1. Given an observation O from the testing data, $P(O|\lambda_i)$, $i = 0, 1, \cdots, 9$ are evaluated, where λ_i are well trained HMMs corresponding to digit classes.

2. If $P_i \leq P(O|\lambda_i)$ where P_i is a threshold, then λ_i will enter the competition; otherwise, it will be eliminated. All $P(O|\lambda_i)$ that pass their thresholds then are sorted in the order of highest probability first.

3. Suppose O is from class j. If $P(O|\lambda_i) < P_i$, $i = 0, 1, \cdots, 9$, then we say O is rejected. In other cases, if $P(O|\lambda_j) = \max_i\{P(O|\lambda_i)\}$, then we say O is recognized; otherwise, we say it is substituted (confused). Furthermore, if O is not rejected and $P(O|\lambda_j)$ is among the top-three probabilities in $\{P(O|\lambda_i)|1 \leq i \leq n\}$, then we say O is top-three hitted.

The digit recognition results using different number of symbol categories are summarized in Table 1. The total number of testing digits was 3126. From Table 1, one can see that using different number of symbol categories, the recognition rate varied from 89.1% to 92.0%, the substitution rate varied from 2.8% to 8.5%, the rejection rate varied from 2.4% to 5.3%, and the top-three hit rate varied from 94.4% to 97.0%.

cat.	recognition	substitution	rejection	top-three hit
66	89.1%	8.5%	2.4%	97.0%
96	92.0%	4.1%	3.9%	95.8%
225	91.8%	2.8%	5.3%	94.4%

Table 1. (OK+GOOD) Digit recognition: 3126 testing digits

5. Conclusions

A hidden Markov model based approach to on-line handwritten digit recognition has been presented. In this approach, a character instance is represented by a sequence of symbolic strokes, and the representation is obtained by component segmentation and stroke classification. The component segmentation is based on the delta lognormal model of handwritting generation [6], and the resulted strokes are classified into symbol categories for HMM multiple observation training or recognition. An on-line recognition experiment on digit data has been conducted using the above techniques. The results have shown that this HMM based approach using stroke sequences is robust in that that a single HMM can adapt to different writing styles through multiple observation training. This important feature is very useful in on-line handwritten character recognition.

Acknowledgements

This research work was supported by NSERC Grant OGP0155389 to M. Parizeau, and NSERC Grant OGP00915 and FCAR Grant ER-1220 to R. Plamondon.

References

[1] E.J. Bellegarda, J.R. Bellegarda, D. Nahamoo and K.S. Nathan, "A fast statistical mixture algorithm for on-line handwriting recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.16, No.12, 1227–1233 (1994)

[2] J. Hu, M.K. Brown and W.Turin, "HMM based on-line handwriting recognition". *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol.18, No.10, 1039-1045 (1996)

[3] S.E. Levinson, L.R. Rabiner and M.M. Sondhi, "An introduction to the application of the theory of probabilistic functions of Markov process to automatic speech recognition". *Bell System Technical Journal*, Vol.62, No.4, 1035-1074 (1983)

[4] X. Li, M. Parizeau and R. Plamondon, "Segmentation and reconstruction of on-line handwritten scripts". In press, *Pattern Recognition*, 31/6, (1998)

[5] M. Parizeau and R. Plamondon, "A fuzzy-syntactic approach to allograph modeling for cursive script recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.17, No.7, 702–712 (1995)

[6] R. Plamondon, "A kinematic theory of rapid human movements Part I. Movement representation and generation". *Biological Cybernetics*, No.72, 295–307 (1995)

[7] R. Plamondon, "A kinematic theory of rapid human movements Part II. movement time and control". *Biological Cybernetics*, No.72, 309-320 (1995)

[8] L. Rabiner and B.H. Juang, "Fundamentals of Speech Recognition". Prentice Hall, Englewood Cliffs, N.J., 1993

[9] C.C. Tappert, C.Y. Suen and T. Wakahara, "The state of the art in on-line handwriting recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.PAMI-12, No.8, 787–808 (1990)