On-line Handwritten Signature Verification Using Hidden Semi-Markov Model

Daw-Tung Lin and Yu-Chia Liao

Department of Computer Science and Information Engineering,
National Taipei University,
151, University Rd., San-Shia, New Taipei City, 237 Taiwan

Abstract. Handwritten signature has been extensively adopted as biometric for identity verification in daily life, as it is the most widely accepted personal authentication method. Automatic signature recognition technologies can definitely facilitate the verification process. Many research attempts and advances have occurred in this field, automatic signature verification still is a challenging and important issue. This work presents a novel and robust on-line signature verification approach using Hidden Semi-Markov Model (HSMM). The proposed system comprises three stages. First, dynamic features are extracted according to the local statistical information of velocity, acceleration, azimuth, altitude, and pressure. Next, the extracted features are normalized into unified observation length. To improve the verification accuracy, features with slight variation are clustered into the same class using K-means classification algorithm. Furthermore, the Forward-Backward algorithm is utilized to accelerate the computation of HSMM parameters. Finally, the system builds a unique HSMM for each identity and estimates the signature baseline in corresponding to the features. To assess the recognition performance of the proposed algorithm, experiments were conducted using SVC2004 signature database. Analytical results reveal that the proposed method is very promising.

Keywords: On-line handwritten signature verification, dynamic features, Hidden Semi-Markov Model, Forward-Backward algorithm.

1 Introduction

Handwritten signature biometrics is commonly adopted for person identity verification and authentication in daily life, such as authority of financial transaction and document approval, etc.. Automat recognition technologies can definitely facilitate the verification process. Many research attempts and advances have occurred in this field, automatic signature verification still is a challenging and important issue. Srihari *et al.* have done a comprehensive survey to online and offline handwritten signature recognition and applications [1]. Lai and Chen considered the virtual stroke feature and made a fusion of real and virtual stroke to develop Chinese signature verification systems [2]. Bovino *et al.* developed a multi-expert method which matches the position, velocity and

C. Stephanidis (Ed.): Posters, Part I, HCII 2011, CCIS 173, pp. 584–589, 2011.

acceleration of each stroke with weighting to verify handwritten signatures [3]. Martens *et al.* employed Dynamic Time Warping (DTW) for signature verification and improved the efficiency of DTW by modifyied feature extraction procedure [4]. Quan *et al.* developed an online signature verification system by fusing the HMM and Artificial Neural Network (ANN) [5]. Rigoll *et al.* utilized HMM and combined several features to classify the signature [6]. Fierrez *et al.* also made extensive discussions about feature extraction and model issue on the verification system using HMM [7]. This work presents a novel and robust on-line signature verification approach using Hidden Semi-Markov Model (HSMM).

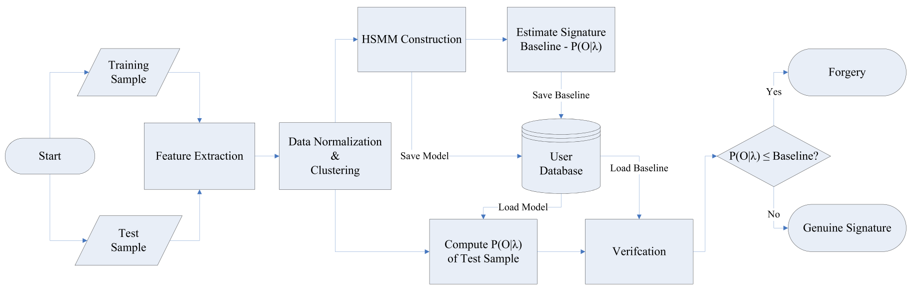

Fig. 1. System flowchart

2 Proposed System Architecture

The proposed system comprises three stages. Figure 1 depicts the flowchart of the proposed system and is described in detail as follows.

2.1 Feature Extraction

Generally, signature is captured from tablet device, on-line information such as velocity, acceleration and pressure is available. In this work, we employ various features for signature verification including velocity, acceleration, pressure, pen tilt orientation, and total signature time.

2.2 Data Normalization and Outlier Clustering

Due to slight variation of signature duration, the data captured from tablet device needs to be normalized to the same observation length or unit for further process. We divide the time sequence data into same number of frames for each signature. Then, the feature of each frame is represented by the average of all data in a single frame. Besides, although each signature has been unified to the same observation length, there are still too many outlier variations that the features could be. Thus, we adopt K-means algorithm to cluster the feature outliers for each signature. By clustering the feature value to limited classes and giving each class an index, signature feature representation can be simplified with limited types.

2.3 HSMM Construction and Verification

In this work, each signature was converted into a index sequence with the observation length T: $(I_1, I_2, ..., I_T)$. We employed the advanced version of HMM, Hidden Semi-Markov Model (HSMM), to perform signature verification [8]. Typically, signature is written starting from the first stroke. Therefore, we set $\pi_1 = 1$, while the other states are not activated, i.e. $\pi_2 = \pi_3 = \cdots = \pi_N = 0$, where π_i denote the probability of initial state i. Based on the specific signature strokes sequence of each person, the characteristics can be transferred and feeded into the HSMM. The state transition in HSMM is based on the variable duration of each state. In this work, we simplify and set each state with a constant duration, i.e. $d_1 = d_2 = d_3 = \cdots = d_N = 1$, so that the state number N is equal to the observation length T. Then, the HSMM state transition probability a_{ij} between state i and state j is determined by the time shift: $a_{ij} = 1$, if $j = i+1$; otherwise $a_{ij} = 0$. Suppose there are Q training sample sequences, the observation sequence $O^{(q)}$ for the q-th training sample is expressed as: $O^{(q)} = [I_1^{(q)}, I_2^{(q)}, ..., I_t^{(q)}, ..., I_T^{(q)}]$, where $I_t^{(q)}$ means the feature cluster index of the t-th observation point. Define $W(I_t^{(q)})$ as the weight of feature cluster index of each observation point and set $W(I_t^{(q)}) = 1$, if $I_t^{(q)} = k$; otherwise $W(I_t^{(q)}) = 0$. The probability of the feature cluster index k under state j is computed by accumulating the number of times that index k occurred under this state: $b_j(k) = \frac{\sum_{q=1}^{Q}\sum_{r=0}^{R-1} W(I_{1+r}^{(q)})}{Q \times R}$, where R is the number of observation points of state j. Since we assume each state with the constant duration $d = 1$, we set $R = 1$.

After establishing the initial HMM λ, we can compute the signature probability in model λ, i.e. finding $P(O|\lambda)$ to determine the verification result. Intuitively, $P(O|\lambda)$ can be obtained by the brute-force method, however it is time consuming. Instead, we employ Forward-Backward algorithm [9]. The forward variable is defined as $\alpha_t(i) = P(I_1, I_2, \cdots, I_t, S_t = i \mid \lambda)$, i.e, the probability of partial observation sequence, $I_1, I_2, \cdots, I_t$, (until time t and state i), given the model λ. Then we can solve the forward variable $\alpha_t(i)$ inductively as follows.

$$\text{Initialization:} \quad \alpha_1(i) = \pi_i b_i(I_1). \tag{1}$$

$$\text{Induction:} \quad \alpha_{t+1}(j) = [\sum_{i=1}^{N} \alpha_t(i) a_{ij}] b_j(I_{t+1}). \tag{2}$$

$$\text{Termination: } P(O|\lambda) = P(I_1, I_2, ..., I_T \mid \lambda) = \sum_{i=1}^{N} P(I_1, I_2, ..., I_T, S_T = i \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{3}$$

Equation (1) initializes the forward probabilities as the joint probability of state i and the initial cluster index I_1. Equation (2) shows the heart of forward process, which computes the probability that state j can be reached at time $t+1$ from N possible states with observed index sequence $I_1 I_2 I_3 \cdots I_t$. Notice that we used the concept of HSMM as the main design to deal with the index sequence of signature, so there is only one state i at time t can reach state j at time

$t+1$. Next, Eq. (3) shows $P(O|\lambda)$ can be obtained by summing terminal forward variables $\alpha_T(i)$. The backward procedure is similar to the forward procedure by computing the backward variable $\beta_t(i)$ reversely in time. Finally, by fusing two procedures, the proper $P(O|\lambda)$ can be computed as follow :

$$\begin{aligned} P(O|\lambda) &= P(I_1, I_2, ..., I_t, ..., I_T \mid \lambda) \\ &= \sum_{i=1}^{N} P(I_1, I_2, ..., I_t, S_t = i \mid \lambda) P(I_{t+1}, ..., I_T \mid S_t = i, \lambda) \\ &= \sum_{i=1}^{N} \alpha_t(i)\beta_t(i) \end{aligned} \tag{4}$$

Moreover, we can further compute $P_k(O|\lambda)$ for various types of feature (X types) and associate it with proper weights as: $P_{overall}(O|\lambda) = \sum_{k=1}^{X} P_k(O|\lambda)W_k$ Therefore, when several genuine signatures has been trained with user's HMM λ, and find an appropriate $P_{overall}(O|\lambda)$ as the baseline of user's signature, then a test sample can be verified by comparing the $P_{overall}(O|\lambda)$. The higher the $P_{overall}(O|\lambda)$ is, the more likely the signature is true.

3 Experimental Results

To assess the performance of the proposed system, experiments were conducted using 40 data sets consisting 1600 signatures of SVC2004 database with 20 genuine signatures and 20 skill forgeries in each set [10]. Five-fold cross validation was performed. To attain an ensemble accuracy, fifty complete five-fold cross-validations were performed on each database. The observation length T was set to 60, and the number of feature clusters is six. The complexity is reduced to $P(O|\lambda)$. Table 1 shows the experimental results of using single feature as the base of verification, and Table 2 shows the results when we associate multiple features with different weights.

Table 1. Experimental results using single feature

		Velocity	Acceleration	Azimuth	Altitude	Pressure
Training	EER mean	7.38%	9.25%	10.03%	15.45%	9.95%
Set	EER SD	4.75%	4.8%	6.78%	8.23%	8.33%
Test	EER mean	18.83%	23.88%	18.15%	26.8%	19.03%
Set	EER SD	11.66%	12.35%	12.33%	14.09%	14.14%

Table 2. Experimental results with multiple features

Feature	Weights							
Velocity	0.75	0.6	0.5	0.7	0.6	0.5	0.5	0.2
Acceleration	0	0	0	0	0	0	0.05	0.2
Azimuth	0.25	0.4	0.5	0.2	0.3	0.3	0.3	0.2
Altitude	0	0	0	0	0	0	0.05	0.2
Pressure	0	0	0	0.1	0.1	0.2	0.1	0.2
EER mean	15.93%	16%	16.03%	16.3%	16.28%	16.35%	17.7%	17.93%
EER SD	10.99%	11.09%	11.07%	11.44%	11.39%	11.52%	9.43%	9.48%

4 Conclusion

This work presents a novel and robust on-line signature verification approach using Hidden Semi-Markov Model (HSMM). The proposed system comprises three stages. First, dynamic features are extracted according to the local statistical information of velocity, acceleration, azimuth, altitude, and pressure. Next, the extracted features are normalized into unified observation length. To improve the verification accuracy, features with slight variation are clustered into the same class using K-means classification algorithm. Furthermore, the Forward-Backward algorithm is utilized to accelerate the computation of HSMM model. Finally, the system builds a unique HSMM for each identity and estimates the signature baseline in corresponding to the features. Analytical results reveal that the proposed method is very promising. This system is expected to further improve the performance and reliability by incorporating additional features such as virtual strokes.

References

1. Plamondon, R., Srihari, S.N.: Online and off-line handwriting recognition: a comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(1), 63–84 (2002)
2. Chen, J.H.: Chinese Signature Verification Using Combination of Real-Stroke and Virtual-Stroke Information. Master's thesis, National Central University (2005)
3. Bovino, L., Impedovo, S., Pirlo, G., Sarcinella, L.: Multi-expert verification of handwritten signatures. Document Analysis and Recognition 2, 932 (2003)
4. Martens, R., Claesen, L.: On-line signature verification by dynamic time-warping. In: Proceedings of the 13th International Conference on Pattern Recognition, vol. 3, pp. 38–42 (2002)
5. Quan, Z.H., Liu, K.H.: Online Signature Verification based on the Hybrid HMM/ANN model. IJCSNS 7(3), 313 (2007)
6. Rigoll, G., Kosmala, A.: A systematic comparison between on-line and off-line methods for signature verification with hidden Markov models. In: Fourteenth International Conference on Pattern Recognition, vol. 2, pp. 1755–1757 (2002)

7. Fierrez, J., Ortega-Garcia, J., Ramos, D., Gonzalez-Rodriguez, J.: HMM-based on-line signature verification: Feature extraction and signature modeling. Pattern Recognition Letters 28(16), 2325–2334 (2007)
8. Yu, S.Z.: Hidden semi-Markov models. Artificial Intelligence 174(2), 215–243 (2010)
9. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2), 257–286 (1989)
10. Yeung, D.Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First international signature verification competition. Biometric Authentication, 1–30 (2004)