Modeling of Pen-Coordinate Information in SCPR-based HMM for On-line Recognition of Handwritten Japanese Characters

Junko Tokuno, Yiping Yang, Gleidson Pegoretti da Silva, Akihito Kitadai, Masaki Nakagawa

Modeling of Pen-Coordinate Information in SCPR-based HMM for On-line Recognition of Handwritten Japanese Characters

Junko Tokuno, Yiping Yang, Gleidson Pegoretti da Silva, Akihito Kitadai and Masaki Nakagawa†
Tokyo University of Agriculture and Technology, 2-24-16, Naka-cho, Koganei, Tokyo, Japan
{j-tokuno, yangyiping, pegoretti, ak}@hands.ei.tuat.ac.jp, †nakagawa@cc.tuat.ac.jp

Abstract

This paper describes stochastic modeling of pen-coordinate information in HMMs with structured character pattern representation (SCPR) for on-line recognition of handwritten Japanese characters. SCPR allows HMMs for Kanji character patterns to share common subpatterns. Although SCPR-based HMMs have been successfully applied to Kanji character recognition, the pen-coordinate feature has not been modeled since it is an unique feature for each character pattern. In this paper, we employ mapping from a common subpattern to each occurrence in Kanji patterns in the estimation step of SCPR-based HMMs. We also employ adaptation of state parameters to each character pattern in generating character HMMs by composing SCPR-based HMMs. Experimental results showed that the correct recognition rate improved from 83.6% to 92.3% for Japanese Kanji character patterns by modeling the pen-coordinate information in the SCPR-based HMMs.

Keywords: On-line handwriting recognition, Structured character pattern representation, HMM, Pen-coordinate feature

1. Introduction

As the popularity of PDA, tablet PC, and other pen-based or paper-based systems, the demands for improvement of on-line handwriting recognition accuracy have increased. The hidden Markov model (HMM) has been successfully applied not only to Western handwriting recognition [1] but also to on-line handwriting recognition of Chinese [2], Kanji characters of Chinese origin [3]-[6] and Hangul characters in Korea [7] because of its promising ability to model deformations of strokes and variations of the number of sample points. In case of alphanumeric on-line handwriting recognition, "character HMM" has been widely employed, where each whole letter of the alphabet is modeled typically by one HMM and all words are represented by employing only some dozens of HMMs at most. On the other hand, there are thousands of characters in Oriental characters of Chinese origin, so that modeling each character by an HMM leads to an infeasible character recognition system requiring huge amount of memory and training data.

To tackle this problem, structured character pattern representation (SCPR)-based HMM [2] has been proposed where each Kanji character is represented as a composite of constituent subpatterns, which are shared among several Kanji characters. The SCPR-based HMM provides such advantages as reducing the total size of the models, making the recognition system robust against deformation of common subpatterns and so on. In the SCPR-based HMMs, the pen-direction feature extracted from consecutive pen-tip positions has almost always been employed [2][6]. In contrast, the pen-coordinate feature has not been employed, though it is no less important than the pen-direction feature. This is due to that the pen-coordinate information is more subject to change when subpatterns are composed into each character pattern.

The proposed SCPR-based HMMs model both the pen-direction feature and pen-coordinate feature by employing mapping from a common subpattern to each occurrence in Kanji character pattern in the estimation step of SCPR-based HMMs. Moreover, we adapt SCPR-based HMM parameters to each character pattern according to the information of the size and the position of raw patterns in the recognition step.

This paper is organized as follows. Section 2 describes our recognition system which employs HMM and SCPR as key components. Section 3 presents the proposed approach for stochastic modeling of the pen-coordinate feature by SCPR-based HMMs. Section 4 shows experimental results and Section 5 concludes this work.

2. Recognition System

Since HMM-based on-line handwriting recognition for thousands of Kanji characters requires an enormous computation time, it is difficult for computers with low performance such as PDAs to work practically. Therefore, we first carry out coarse classification and reduce recognition candidates [8] and then apply HMM-based recognition for those candidates. In our research, we set the number of maximum recognition candidates to 200.

The proposed HMM-based recognition system basically consists of a feature extraction module, SCPR, SCPR-based HMMs, and a decoder. In this section, we show the outline of our recognition system.

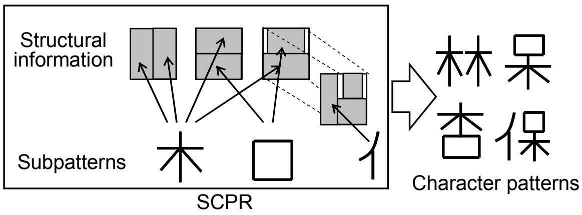

Figure 1. Structured character pattern representation.

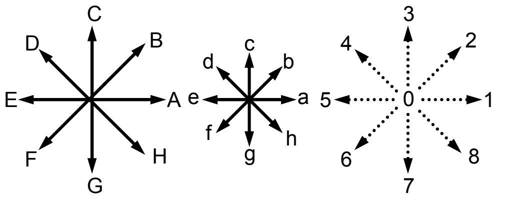

Figure 2. Sub-stroke categories: A-H (a-h) are long (short) sub-strokes and 0-8 are the direction of off-strokes.

2.1. Feature Extraction

We use a sequence of pen-tip positions (x_t,y_t), $t=1,\cdots,T$, sampled at a certain interval from a pen tablet as the pen-coordinate features. Moreover, we extract the displacement $(\Delta x_t,\Delta y_t)=(x_t-x_{t-1},y_t-y_{t-1})$ from two consecutive pen-tip positions and use (r_t,θ_t) as the pen-direction feature, where $r_t=\sqrt{\Delta x_t^2+\Delta y_t^2}$ and θ_t denote the velocity and the direction of the pen movement, respectively. We then denote by $O=o_1,o_2,\cdots,o_T$, $o_t=(x_t,y_t,r_t,\theta_t)$ the feature sequence representing each character pattern in case that we employ both the pen-coordinate features and the pen-direction features.

2.2. Structured Character Pattern Representation (SCPR)

Kanji character patterns, ideographic character patterns of Chinese origin, are mostly composed of multiple subpatterns. Very often subpatterns are shared among several Kanji character patterns as shown in Figure 1. Each subpattern is composed by a sequence of strokes and each stroke is made of a sequence of sub-strokes. In this paper, a *stroke* denotes a sequence of pen-tip coordinates sampled from pen-down to pen-up and an *off-stroke* denotes a vector from pen-up to the next pen-down. A stroke is divided into a sequence of straight-line *sub-strokes* [6]. Sub-strokes are classified into eight directions of two-quantized length while off-strokes are classified into eight directions and another of very short distance with arbitrary direction as shown in Figure 2.

Structured character pattern representation (SCPR) represents a character pattern as a composite of subpatterns and their structures, where common subpatterns are shared among several character patterns that include the subpatterns is their shapes. SCPR is suitable for patterns that have structures like Chinese characters and provides advantages such as the size reduction of a dictionary (a set of prototype patterns) and greater recognition robustness against deformation of common subpatterns.

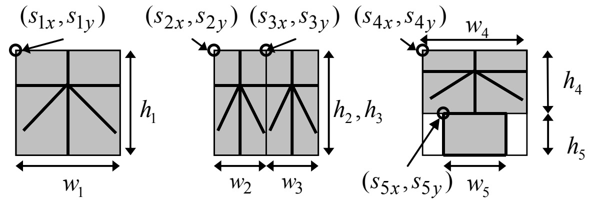

Figure 3. Structural information of SCPR.

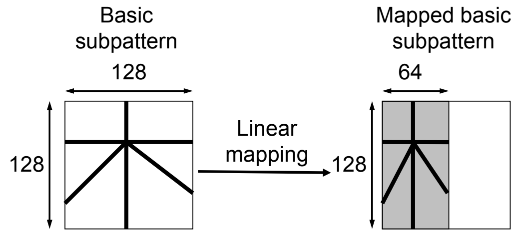

Figure 4. Generating mapped basic subpattern by linear mapping.

Japanese Kanji characters are of Chinese origin and are almost the same as Chinese characters. Korean Hangul characters, although they are phonetic characters, also have a structural composition similar to that used for Chinese characters, so structured approaches have also been applied.

In our on-line recognition system, each character pattern prototype is represented by a SCPR that is a composite of subpatterns and off-strokes. Each SCPR has constituent subpatterns and the structural information on where in the character pattern each subpattern is placed in terms of the bounding box (s_x,s_y) $<w,h>$ for the subpattern, where (s_x,s_y) denotes the top-left corner and $<w,h>$ denotes <width, height> of the bounding box as shown in Figure 3.

All the basic subpatterns (subpatterns which are not decomposed further into smaller subpatterns) as well as the character patterns are represented by a square shape of 128×128 resolution, and are reduced to bounding boxes in structural information through linear mapping when they are included in larger subpatterns or character patterns (Figure 4). In this paper, we call a result of the linear mapping a "mapped basic subpattern", even if the mapping is identical.

Since the recognition method is sensitive to stroke order variations, there are multiple prototypes for each subpattern and multiple SCPRs for each character pattern. Before the start of this study, all the subpattern prototypes had already been clustered by employing a simple clustering algorithm based on the LBG algorithm [11]. Though the number of definitions is different among subpatterns, the average is 4.53 for each subpattern and the standard deviation is 5.86.

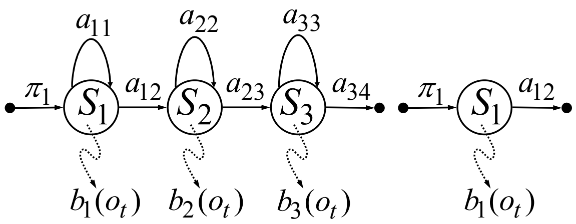

Figure 5. Sub-stroke HMMs : (Left) pen down model, (Right) pen up model.

2.3. SCPR-based HMMs

We model each subpattern by HMMs. In this paper, we call our subpattern HMMs as SCPR-based HMMs since they are SCPRs. In the previous study [6], each sub-stroke is represented by a left-to-right HMM of three states while each off-stroke is represented by a single state (Figure 5). On the other hand, each stroke is represented by concatenation of sub-stroke HMMs and each subpattern is represented by that of stroke HMMs interleaved with off-stroke HMMs in this work. Therefore, the total number of states in each subpattern HMM is determined by connecting three states sub-stroke HMMs and single-state off-stroke HMMs.

Here, let $\lambda^{(k)}=(A^{(k)},B^{(k)},\pi^{(k)})$ be the set of HMM parameters of a subpattern or off-stroke k, with the following notations:

$A^{(k)}=\{a_{ij}^{(k)}\}$: set of all the state-transition probability distributions from state S_i to S_j,

$B^{(k)}=\{b_i^{(k)}(o_t)\}$: set of all the probability distributions of observing symbols o_t at state S_i,

$\pi^{(k)}=\{\pi_i^{(k)}\}$: initial state probability distributions.

The observation probability distributions are represented by mixtures of M Gaussian distributions given by

$$b_i(o_t)=\sum_{m=1}^{M} c_{im}\frac{\exp(-\frac{1}{2}(o_t-\mu_{im})^t\Sigma_{im}^{-1}(o_t-\mu_{im}))}{\sqrt{(2\pi)^n|\Sigma_{im}|}} \tag{1}$$

with the mean vectors μ_i, the covariance matrices Σ_i (n is the dimension of the observation feature vector o_t.) and weighting coefficients c_{im}. Here, the direction feature θ has a continuous probability distribution with 2π cycle. In this study, each state of HMM has a single Gaussian distribution with diagonal covariance. These parameters can be trained through the Viterbi training or the Baum-Welch algorithm.

2.4. Decoder

According to a SCPR and SCPR-based HMMs, the decoder concatenates the SCPR-based HMMs to generate an HMM of each candidate character pattern, and then calculates the probability that the input pattern is produced from the HMM. This operation is effectively done by the Viterbi search algorithm.

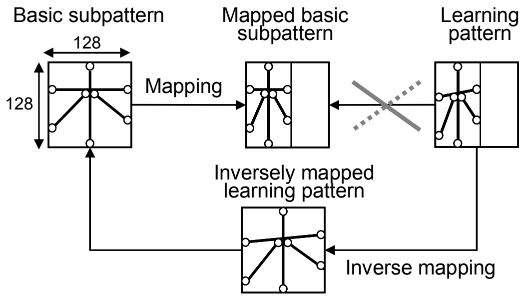

Figure 6. Mapping and inverse mapping.

3. Modeling of Pen-coordinate Features

The basic idea of modeling pen-coordinate features by SCPR-based HMMs is based on the following mapping and inverse mapping. As mentioned before, all subpatterns are represented by a square shape of 128×128 resolution in our recognition system. In this study, we call the square size of 128×128 normalization size. When we generate character patterns by combining subpatterns, we apply a linear mapping from each subpattern to its area in a character pattern. In contrast, we apply the inverse of the above mapping when we learn subpatterns (Figure 6). A simple idea is to enlarge the size of the bounding box of a mapped basic subpattern in a learning pattern to the normalization size. By applying the inverse mapping, we can exclude character dependency of each subpattern (difference in size and position when it appears in different character patterns) to model pen-coordinate features of the subpattern by SCPR-based HMM.

3.1. Initial Parameters of SCPR-based HMMs

Before estimating SCPR-based HMM parameters by the Viterbi training or the Baum-Welch method, we set initial parameters in each SCPR-based HMM by applying the inverse mapping. However, handwriting usually has noises due to a hand vibration etc. and these noises have no or little correlation with the bounding box size of a subpattern, so that the inverse mapping may magnify these noises and reflect them into the subpattern. Therefore, we set the initial parameters from character patterns which are composed of only a single subpattern. First of all, we extract the features: $O=o_1,o_2,\cdots,o_T$, $o_t=(x_t,y_t,r_t,\theta_t)$ from each character pattern in training data and then we assign those features to each state of concatenated SCPR-based HMMs equally by employing pen-up information. We set the initial parameters for each state by taking the average of those features in the training data.

Exceptionally, there are some subpatterns which do not appear as character patterns as them alone. In this case, we first generate character HMMs and carry out the Viterbi segmentation for the HMMs. According to the result of the Viterbi segmentation, we extract those subpatterns and apply inverse mapping to them. Finally, we assign the features: $O=o_1,o_2,\cdots,o_T$, $o_t=(x_t,y_t,r_t,\theta_t)$ to each state of SCPR-based HMMs according to the result of the Viterbi segmentation and set the initial parameters based on the above procedure.

3.2. Adaptation of SCPR-based HMM Parameters

In both the estimation step and the recognition step, we generate HMMs for character patterns by connecting one or more than one SCPR-based HMM according to SCPRs described in Section 2.4. Each SCPR-based HMM which is a composition of generated character HMMs corresponds to a subpattern whose square size is smaller than normalization size, though each SCPR-based HMM parameters are estimated by normalizing each training patterns. In other words, the parameters for the pen-coordinate feature (o_t=(x_t,y_t)) of a common subpattern are mapped to distinct values when it is composed into different character patterns. Then, we need to adapt the parameters of SCPR-based HMM to each character.

Here, let $\vec{\mu}_i = (\mu_i(x), \mu_i(y))$ be the mean vector of the Gaussian distribution at a state S_i of SCPR-based HMMs, an adapted mean vector is given by

$$\hat{\mu}_i(x) = \mu_i(x) \times \frac{w}{128} + s_x \qquad (2)$$

$$\hat{\mu}_i(y) = \mu_i(y) \times \frac{h}{128} + s_y \qquad (3)$$

where $\mu_i(x)$ and $\mu_i(y)$ denote the mean vectors of the pen-coordinate feature x and y, and w,h,s_x,s_y are noted in Section 2.2. Moreover, we assume that the diagonal covariance matrix $\Sigma_i=(\sigma^2_{ixx},\sigma^2_{iyy})$ of the Gaussian distribution at each state S_i of SCPR-based HMMs is correlated to the bounding box sizes of subpatterns and adapt Σ_i to each character according to the correlations. Then, in order to obtain the correlations between the average of the covariance and the bounding box size we analyzed them using the database HANDS_kuchibue_d_97_06 [9]. The result is shown in Figure 7. According to the result, we convert the diagonal covariance matrix $\Sigma_i=(\sigma^2_{ixx},\sigma^2_{iyy})$ to $\hat{\Sigma}_i = (\hat{\sigma}^2_{ixx}, \hat{\sigma}^2_{iyy})$ as follows:

$$\hat{\sigma}_{ixx} = \begin{cases} \sigma_{ixx} - 0.085 \times (128 - w) & (\sigma_{ixx} \geq 9.4707) \\ 0.085w + 0.005 & (\sigma_{ixx} < 9.4707) \end{cases} \qquad (4)$$

$$\hat{\sigma}_{iyy} = \begin{cases} \sigma_{iyy} - 0.080 \times (128 - h) & (\sigma_{iyy} \geq 8.99614) \\ 0.080h + 0.007 & (\sigma_{iyy} < 8.99614) \end{cases} \qquad (5)$$

3.3. Estimation of SCPR-based HMM

In the estimation, in order to avoid noise influence, we employ displacement normalization proposed in our previous study [10] instead of the inverse mapping for each subpattern. This approach is based on the assumption that there are some correlations between the bounding box size of a subpattern and the freedom of movement of each feature point in the subpattern because each feature point can move in larger area if the bounding box size of the subpattern is larger as shown in Figure 8. From the analysis of our previous study, the relationship between the bounding box size: <w,h> and displacement: (D_{Ax},D_{Ay}) is as follows:

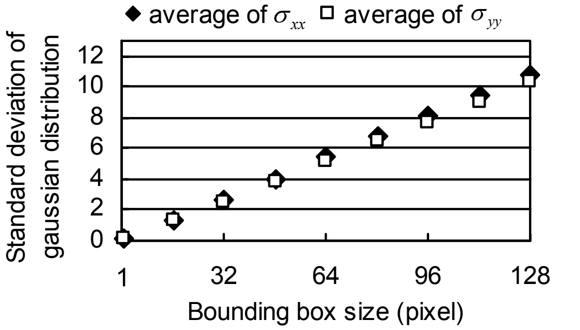

Figure 7. The relations between the standard deviation and the bounding box size.

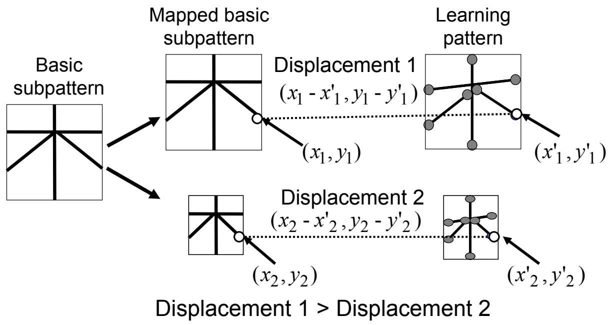

Figure 8. Displacement between the basic subpattern and the learning pattern.

$$D_{Ax}(w) = 0.0846w + 1.7 \qquad (6)$$

$$D_{Ay}(h) = 0.0539h + 3.5 \qquad (7)$$

The detailed procedure of estimating parameters for SCPR-based HMM based on this formulation is as follows:

1. We first extract both the pen-direction features and the pen-coordinate features: O=o_1,o_2,$\cdots$,o_T, o_t=(x_t,y_t,r_t,θ_t) from each character pattern in training data and carry out the Viterbi training for concatenated initial SCPR-based HMMs. According to the result of the Viterbi segmentation, we extract subpatterns which do not appear as character patterns.
2. We convert pen-direction features of each subpattern as follows:

$$x_t' = \mu_i(x) + (x_t - \hat{\mu}_i(x)) \times \frac{D_{Ax}(128)}{D_{Ax}(w)} \qquad (8)$$

$$y_t' = \mu_i(y) + (y_t - \hat{\mu}_i(y)) \times \frac{D_{Ay}(128)}{D_{Ay}(h)} \quad (9)$$

where (x_t, y_t) and (x'_t, y'_t) are feature points of a pre-normalized subpattern and a normalized subpattern, respectively. And $\vec{\hat{\mu}}_i = (\hat{\mu}_i(x), \hat{\mu}_i(y))$ is the mean vector of a state S_i where a feature point (x_t, y_t) is observed.

3. We update parameters for each state by taking the average of those converted features in the training data.

4. Experiments

We carried out experiments to show the effect of the pen-coordinate feature and that of SCPR-based HMMs for Japanese Kanji recognition. The database HANDS_kuchibue_d_97_06 [9] contains 1,435,440 characters written by 120 writers. In the experiments, we used 675,840 patterns (only JIS1 Kanji patterns). Patterns from 60 writers were used for estimating the HMM parameters by the Viterbi training, and those from the remaining 60 writers were used for test**.**

4.1. Experiment 1: Comparison of Features

In this experiment, we compared the conventional pen-direction features with the pen-coordinate features added to them as modeling features for SCPR-based HMMs to show the effect of the pen-coordinate features. The recognition rates according to the feature sets are shown in Table 1. We can see from the result that the recognition rate with pen-coordinate features is higher than that with only pen-direction features about 9 point.

Some samples of correctly recognized character patterns by modeling the pen-coordinate features are shown in Table 2. We can see from the result that misclassifications between character patterns which have the same direction of stroke movement are mostly corrected by modeling the pen-coordinate features. Moreover, HANDS_kuchibue_d_97_06 is a collection of character patterns of usual text. As shown in Table 3, it includes many patterns composed of a few stroke and they are often misrecognized by only the pen-direction features but recognized by both the pen-direction features and the pen-coordinate features. These facts are the reasons of the large improvement of recognition rate.

Conversely, the character patterns that have been turned to incorrect recognition are shown in Table 4. These character patterns have the similar SCPR such as "頴" and "穎" though the total number of those misclassified character patterns is small compared to the improved character patterns.

From these results, it can be concluded that the pen-coordinate features are effective for the SCPR-based HMMs.

Table 1. Comparison of features.

Features	*N*-best cumulative recognition rate [%]			
	1	~2	~3	~10
(r,θ)	83.6	88.4	90.0	92.7
(x,y,r,θ)	92.3	95.1	95.9	96.9

Table 2. Examples of character patterns correctly recognized by employing pen-coordinate features (x,y,r,θ)

Input	Correct recognition rate		Misrecognized character patterns of (r,θ)
	(r,θ)	(x,y,r,θ)	
治	0.2%	90%	活，沿，冶
何	1%	84%	佑
偵	2%	95%	順，悼，須
床	2%	97%	庇，居，呆
操	3%	97%	燥

Table 3. Correct recognition rate of character patterns which often appear in the HANDS_kuchibue_d_97_06 database.

Input	Appea-rance rate	Correct recognition rate		Misrecognized character patterns of (r,θ)
		(r,θ)	(x,y,r,θ)	
人	8%	88%	96%	八，入
日	6%	6%	65%	白，月，用
大	5%	51%	80%	丈
子	4%	53%	84%	与，土
十	4%	87%	99%	七，土

Table 4. Examples of character patterns that have been turned to incorrect recognition by employing pen-coordinate features (x,y,r,θ)

Input	Correct recognition rate		Misrecognized character patterns of (x,y,r,θ)
	(x,y,r,θ)	(r,θ)	
頴	70%	92%	穎
未	72%	82%	朱
埋	77%	98%	浬，哩
卿	78%	90%	郷
鉛	78%	97%	鈷

4.2. Experiment 2: SCPR-based HMMs vs. Character HMMs

We compared the conventional character HMMs and the proposed SCPR-based HMMs with respect to the amount of training patterns to model deformation of strokes. Figure 9 shows the recognition rates when varying the amount of training patterns. Note that there are 5,632 character patterns per writer.

As the result, the SCPR-based HMMs achieved better recognition performance with a smaller amount of training patterns than the character HMMs. This is because a larger number of training patterns are

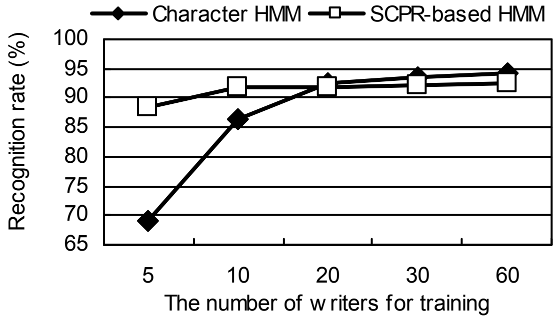

Figure 9. Comparison of character HMM and SCPR-based HMM about the amount of training patterns.

employed for each SCPR-based HMM than for each character HMM even when the amount of training patterns are limited. In this study, there are multiple prototypes for each subpattern and multiple SCPRs for each character pattern to recognize character patterns with wrong stroke order as mentioned in Section 2.2. SCPR-based HMMs can be trained from a lot of subpatterns with various stroke orders because subpatterns are shared among several character patterns. On the other hand, when the stroke orders of the same character class are different, each character HMM is trained separately, even if the stroke order of constituent subpatterns are the same. Therefore, there is not an enough amount of training patterns for the character HMMs to model deformations of strokes.

This result also shows that there is no appreciable difference between the recognition rate of the SCPR-based HMMs and the character HMMs in case that there is enough amount of training data.

5. Conclusion

In this paper, we have proposed the stochastic modeling of pen-coordinate information by SCPR-based HMMs. Through the experiments, it has been shown that the recognition accuracy is vastly improved by modeling the pen-coordinate information for SCPR-based HMMs. We also showed that the proposed SCPR-based HMMs are superior to conventional character HMMs in case that training patterns are limited. Even if there are a large amount of training patterns, the recognition accuracy of the SCPR-based HMMs is no less than that of the character HMMs.

In the near future, we will consider the stability of pen-coordinate features observed at each state of SCPR-based HMMs [5] in order to improve recognition accuracy.

Acknowledgment

The HMM based approach on which this work is extended is due to the research of the first author with Prof. Shigeki Sagayama currently at Tokyo University and Hiroshi Shimodaira currently at Edinburgh University and Mitsuru Nakai at Toyama Prefectural University. The authors would like to thank them.

This work is partially supported by Grant-in-Aid for Scientific Research under the contract number B:17300031.

References

[1] J. Hu, M. K. Brown and W. Turin: HMM based on-line handwriting recognition, *IEEE Trans. PAMI*, **18**, 10, pp.1039-1045, Oct.1996.

[2] H. J. Kim, K. H. Kim, S. K. Kim and F. T-P. Lee: On-line recognition of handwritten Chinese characters based on hidden Markov models, *Pattern Recognition*, **30**, 9, pp.1489-1499, 1997.

[3] H. Ito and M. Nakagawa: An on-line handwritten character recognition method based on hidden Markov model (in Japanese), Technical report of IEICE, PRMU97-85 pp.95-100, July.1997.

[4] K. Takahashi, H. Yasuda and T. Matsumoto: On-line handwritten character recognition using hidden Markov model (in Japanese). *Technical report of IEICE, PRMU96-211,* pp.143-150, Mar.1997.

[5] D. Okumura, S. Uchida and H. Sakoe: An HMM implementation for on-line handwriting recognition based on pen-coordinate feature and pen-direction feature, *Proc. ICDAR'05*, pp.26-30, Aug.2005.

[6] M. Nakai, N. Akira, H. Shimodaira and S. Sagayama: Sub-stroke approach to HMM-based on-line Kanji handwriting recognition, *Proc. ICDAR'01*, pp.491-495, Sept.2001.

[7] S-J. Cho and J. H. Kim: Bayesian Network Modeling of Hangul Characters for On-line Handwriting Recognition, *Proc. ICDAR'03*, pp.207-211, Aug. 2003.

[8] K. Matsumoto and M. Nakagawa: Improvement of a coarse classification method for on-line recognition of handwritten Japanese characters (in Japanese), *Technical Report of IEICE,* PRMU2001-272, pp.9-16, Mar. 2002.

[9] M. Nakagawa and K. Matsumoto: Collection of on-line handwritten Japanese character pattern databases and their analysis, *IJDAR*, **7**, 1, pp.69-81, 2004.

[10] A. Kitadai and M. Nakagawa: A learning algorithm for structured character pattern representation used in on-line recognition of handwritten Japanese characters, *Proc. IWFHR'02*, pp.163-168, Aug. 2002.

[11] Y. Linde, A. Buzo and R. M. Gray: An algorithm for vector quantizer design, *IEEE Trans.* on Communications COM-28(1): 84-95. 1980.