Pen-Coordinate Information Modeling by SCPR-based HMM for On-line Japanese Handwriting Recognition

Junko Tokuno, Yiping Yang, Gleidson Pegoretti da Silva, Akihito Kitadai, 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 Japanese handwriting recognition. 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 unique feature in each character pattern. In this paper, we employ mapping from a common subpattern to each occurrence in Kanji patterns and adaptation of state parameters to each character pattern in generating character HMMs by composing SCPR-based HMMs. Experimental results show that the pen-coordinate feature modeled in the SCPR-based HMMs effects significantly.

1. Introduction

The hidden Markov model (HMM) has been successfully applied to not only Western handwriting recognition[1] but also on-line handwriting recognition of Chinese[2], Kanji characters of Chinese origin[3] and Korean Hangul characters[4] 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, 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 been almost always employed[2][3]. In contrast, 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 the pen-coordinate feature by employing mapping from a common subpattern to each occurrence in Kanji character patterns in the estimation step of SCPR-based HMMs. Moreover, we adapt SCPR-based HMM parameters to each character pattern according to the structural information.

After describing our recognition system in Sec. 2, we propose stochastic modeling of the pen-coordinate feature by SCPR-based HMMs in Sec. 3. We show experimental results in Sec. 4 and conclude this work in Sec. 5.

2. Recognition system

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

2.1. Feature extraction

We use a sequence of pen-tip positions (x_t,y_t), $t=1,\ldots,T$, sampled at a certain interval from a pen tablet as the pen-coordinate feature. 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 employ (r_t,θ_t) as

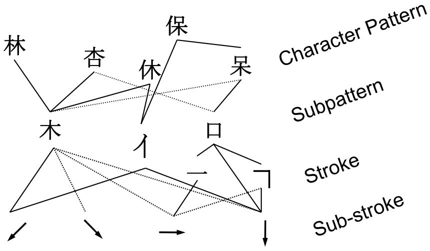

Figure 1. SCPR.

the pen-direction feature, where $r_t=\sqrt{\Delta x_t^2+\Delta y_t^2}$ and θ_t denote velocity and direction of the pen movement, respectively. We then denote by $\boldsymbol{O}=\boldsymbol{o}_1,\boldsymbol{o}_2,\ \ldots,\ \boldsymbol{o}_T$, $\boldsymbol{o}_t=(x_t,y_t,r_t,\theta_t)$ the feature sequence representing each character pattern in case that we employ both pen-coordinate features and pen-direction features.

2.2. SCPR

Kanji character patterns are mostly composed of multiple subpatterns. Very often subpatterns are shared among several Kanji character patterns as shown in Fig. 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.

Each character pattern is represented by structured character pattern representation (SCPR) that is a composite of subpattern and off-stroke models like BNF definitions. We employ 675 subpatterns and eight directional off-strokes. Each SCPR has the structural information on where in the character pattern each subpattern is placed in terms of the bounding box (s_x,s_y) $\langle w,h\rangle$ for the subpattern where (s_x,s_y) denotes the top-left corner and $\langle w,h\rangle$ denotes $\langle$*width, height*$\rangle$ of the bounding box as shown in Fig. 2.

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. Though the number of definitions is different among subpatterns, the average is 4.53 for each subpattern and the standard deviation is 5.86.

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[3], each

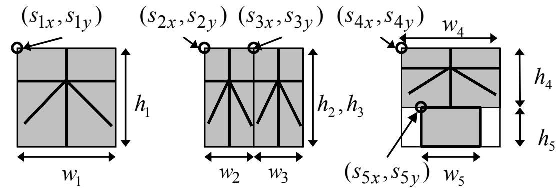

Figure 2. Structural Information of SCPR.

sub-stroke is represented by left-to-right HMM of three states while each off-stroke is represented by a single state. 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. 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, in which

$A^{(k)}=\{a_{ij}^{(k)}\}$: state-transition probability distribution from state S_i to S_j,

$B^{(k)}=\{b_i^{(k)}(\boldsymbol{o}_t)\}$: probability distribution of observation symbols $\boldsymbol{o}_t$ at state S_i,

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

The observation probability distribution is represented by a M mixtures of Gaussian distribution given by

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

with the mean vector μ, the covariance matrix Σ and the weighting coefficient c. Here, the direction feature (θ) has a continuous probability distribution with 2π cycle. These model parameters can be trained by the Viterbi training or the Baum-Welch method.

2.4. Decoder

We first carry out coarse classification and reduce recognition candidates until 200 and then apply an HMM decoder for those candidates. The decoder generates connection models for each character pattern from SCPR and SCPR-based HMMs, and then calculates probability that an input pattern is produced from the models by the Viterbi search algorithm.

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. In our system, all subpatterns (basic subpatterns) are represented by a square shape of 128×128 resolution which is called normalization size. The basic subpatterns are reduced

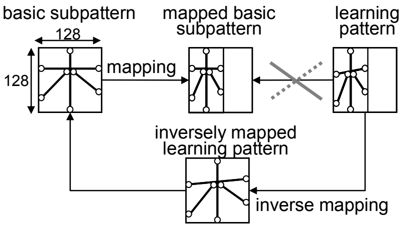

Figure 3. Mapping and inverse mapping.

to bounding boxes in structural information through linear mapping when they are included in larger subpatterns or character patterns (Fig. 3). In this paper, we call a result of the linear mapping a "mapped basic subpattern", even if the mapping is identical.

In contrast, we apply the inverse of the above mapping when we learn subpatterns. A simple idea is to enlarge the size of the bounding box of a mapped basic subpattern in a learning pattern to 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

First of all, 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., 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. After extracting the features: $\boldsymbol{O}=\boldsymbol{o}_1,\ldots,\boldsymbol{o}_T$, $\boldsymbol{o}_t=(x_t,y_t,r_t,\theta_t)$ from each character pattern in training data, we assign them to each state of the SCPR-based HMMs equally. We set the initial parameters for each state by taking the average of those features. Exceptionally, there are some subpatterns which do not appear as character patterns as them alone. In that case, we extract those subpatterns by the result of the Viterbi segmentation and apply inverse mapping to them. We then set the initial parameters according the above procedure.

3.2 Adaptation of SCPR-based HMM parameters

As mentioned in Sec. 2.4, we generate HMMs for character patterns by connecting one or more than one SCPR-based HMM according to SCPRs. 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

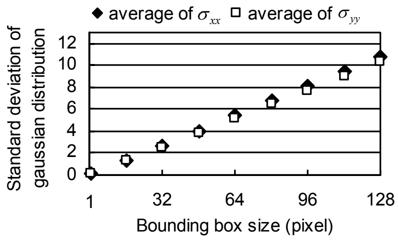

Figure 4. Relations between the standard deviation and the bounding box size.

each SCPR-based HMM parameters are estimated by normalizing each training patterns. In other words, the parameters for the pen-coordinate feature $(\boldsymbol{o}_t=(x_t,y_t))$ 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

$$\hat{\mu}_i(x)=\mu_i(x)\times\frac{w}{128}+s_x,\quad \hat{\mu}_i(y)=\mu_i(y)\times\frac{h}{128}+s_y$$

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 Sec. 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. In order to analyze the correlations we use the database HANS_kuchibue_d_97_06[6]. From the result shown in Fig. 4, 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:

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

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

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[5] 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. The relationship between the bounding box size: $<w,h>$ and displacement: (D_{Ax},D_{Ay}) is as follows:

$$D_{Ax}(w)=0.0846w+1.7,\quad D_{Ay}(h)=0.0539h+3.5$$

Based on this formulation, we estimate SCPR-based HMM. After extracting the features: $\boldsymbol{O}=\boldsymbol{o}_1,\ldots,\boldsymbol{o}_T$, $\boldsymbol{o}_t=(x_t,y_t,r_t,\theta_t)$ from each character pattern, we apply the Viterbi training for the concatenated initial SCPR-based HMMs and extract subpatterns which do not appear as character patterns. Then, we convert the pen-coordinate features of each subpattern as follow:

$$x_t' = \mu_i(x) + (x_t - \hat{\mu}_i(x)) \times \frac{D_{Ax}(128)}{D_{Ax}(w)}, y_t' = \mu_i(y) + (y_t - \hat{\mu}_i(y)) \times \frac{D_{Ay}(128)}{D_{Ay}(h)}$$

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. Finally, we update parameters for each state by taking the average of those converted features.

4. Experiments

We made 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[6] 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 training, and those from the remaining 60 writers were used for test.

4.1. Comparison of features

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 Tab. 1. We can see from the result that the recognition accuracy is vastly improved by adding the pen-coordinate features to the pen-direction features.

Table 1. Comparison of features.

Features	*N*-best cumulative recognition rate [%]			
	1	~2	~3	~10
(r_t,θ_t)	83.6	88.4	90.0	92.7
(x_t,y_t,r_t,θ_t)	92.3	95.1	95.9	96.9

4.2. 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. Fig. 5 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 employed for each SCPR-based HMM than for each character HMM even when the amount of training patterns are limited. SCPR-based HMMs can be

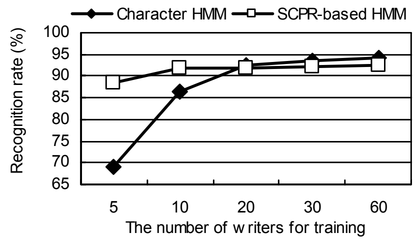

Figure 5. Comparison of character HMM and SCPR-based HMM.

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.

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

5. Conclusion

In this paper, we have proposed the stochastic modeling of pen-coordinate information by SCPR-based HMMs. Experimental results showed that the recognition accuracy is vastly improved by modeling the pen-coordinate features for SCPR-based HMMs.

6. References

[1] J. Hu et al.: HMM based on-line handwriting recognition, *IEEE Trans. PAMI*, **18**, 10, pp.1039-1045, 1996.

[2] H. J. Kim et al.: On-line recognition of handwritten Chinese characters based on hidden Markov models, *Pattern Recognition*, **30**, 9, pp.1489-1499, 1997.

[3] M. Nakai et al.: Sub-stroke approach to HMM-based on-line Kanji handwriting recognition, *Proc. ICDAR'01*, pp.491-495, 2001.

[4] K. Kang, et al.: Handwritten Hangul Character Recognition with Hierarchical Stochastic Character Representation, *Proc. ICDAR'03*, pp.212-216, 2003.

[5] 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, 2002.

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