

座標特徴と方向特徴の選択的利用に基づくオンライン文字認識HMM

片山, 喜規

Faculty of Information Science and Electrical Engineering, Kyushu University

内田, 誠一

Faculty of Information Science and Electrical Engineering, Kyushu University

迫江, 博昭

Faculty of Information Science and Electrical Engineering, Kyushu University

<https://hdl.handle.net/2324/7402678>

出版情報：電子情報通信学会論文誌 D. J91-D (8), pp.2112-2120, 2008-08-01. The Institute of Electronics, Information and Communication Engineers

バージョン：

権利関係：© (社) 電子情報通信学会2008



KYUSHU UNIVERSITY

座標特徴と方向特徴の選択的利用に基づくオンライン文字認識 HMM

片山 喜規^{†a)} 内田 誠一^{†b)} 迫江 博昭[†]

HMM for On-Line Handwriting Recognition by Selective Use of Pen-Coordinate Feature and Pen-Direction Feature

Yoshinori KATAYAMA^{†a)}, Seiichi UCHIDA^{†b)}, and Hiroaki SAKOE[†]

あらまし 本論文では、高精度なオンライン文字認識のために、方向特徴並びに座標特徴を適切に使い分け可能な隠れマルコフモデル（HMM）を提案する。両特徴はいずれもオンライン文字認識の基本的な特徴量でありますから、全く異なる性質を示す。すなわち、線分内で方向特徴が定常的なのに対し、座標特徴は常に非定常である。したがって、HMM の枠組みにおいて両特徴を同等に扱うのは問題が多い。実際従来法では、座標特徴を用いずに方向特徴だけが用いられることが多かった。本論文で提案する HMM では、方向特徴を状態内自己遷移における出力シンボルとして使用し、座標特徴を状態間遷移における出力シンボルとして使用する。このようにすることで、線分方向が一定した定常的な部分においては方向特徴が、線分の方向が変化する過渡的な部分においては座標特徴が評価されることになる。このように特徴を使い分けることで、従来法に比べ認識精度を大幅に向上できることを、多画文字（漢字）の筆順フリー認識実験並びにその詳細な考察を通して示す。

キーワード オンライン文字認識、筆順フリー、隠れマルコフモデル、特徴の選択的利用

1. まえがき

オンライン文字認識において、各ストローク（画）の形状や標本点数などの変動を吸収しつつマッチング処理を行う手段の一つとして隠れマルコフモデル（HMM）が利用されている。HMM は状態をノードとするネットワークモデルであり、状態遷移の際に既定の確率分布に従ってシンボル（文字の局所的性質を表す特徴量）を出力する。HMM を用いる際は、ネットワークのトポロジー及び出力シンボルなどの設計が必要となる。トポロジーについては、オンライン文字認識の場合、一般的に自己遷移を含んだ left-to-right 型が適用される。

ストロークを表現する HMM（以下、必要に応じてストローク HMM と呼ぶ）の出力シンボルについては運筆方向特徴（以下、方向特徴）が採用されることが多い。各状態の確率分布は定常的であるため、自己遷

移の回数だけほぼ一定の方向特徴が出力されることになる。したがって、ストロークを折れ線近似したときの一線分を HMM の一状態に対応づければ、文字形状の局所線形伸縮を自然に表現できることになる。別の言い方をすれば、方向特徴の線分内定常性を自己遷移により適切に表現していることになる。

これに対し、座標特徴（二次元平面内におけるペン先の位置）は、方向特徴のもととなる重要な特徴であるにもかかわらず、あまり利用されていない。この理由としては座標特徴の非定常性が挙げられる。すなわち、折れ線近似した一線分内においても座標特徴は常に変化するため、一状態の確率分布による定常的な表現が困難になる。対応策としてより多くの状態に再分割することも考えられるが、パラメータ数の増加のためにより多くの学習データが必要となる。また、そもそも定常性を表現可能な HMM を用いる意義自体が失われることになる。

しかし、座標特徴の無視は、ストロークの位置や長さを直接表現できなくなるという副作用をもたらす。一画文字や筆順変動を考えなくてよい文字（すなわちペンアップ部分を接続することで実質一画文字とみなせる文字）では、これが大きな問題にならない場合も

[†]九州大学大学院システム情報科学研究院、福岡市

Faculty of Information Science and Electrical Engineering,
Kyushu University, 744 Motooka, Nishi-ku, Fukuoka-shi,
819-0395 Japan

a) E-mail: yosinori@is.kyushu-u.ac.jp

b) E-mail: uchida@is.kyushu-u.ac.jp

ある。しかし、多画文字を筆順フリー条件下で認識する場合、座標特徴の無視は深刻な問題となる。例えば、“四”と“兄”は、方向特徴だけで考えれば、ほとんど同一の文字として認識されるであろう。

そこで本論文では、方向特徴の線分内定常性と座標特徴の非定常性という性質を生かし、それらを適切に使い分けた HMM の構成手法を提案する。具体的には、方向特徴を状態内自己遷移における出力シンボルとして使用し、座標特徴を状態間遷移における出力シンボルとして使用する。このようにすることで、線分方向が一定した定常的な部分においては方向特徴が、線分の方向が変化する過渡的な部分（すなわちストロークの始端・終端・屈曲点）においては座標特徴が評価されることになる。

以下では、まず従来の HMM における方向及び座標特徴の使われ方について再考し、問題点及び限界について指摘する。次に、それらを踏まえた上で、本論文で提案する HMM について説明する。その後、従来法に対する本手法の優位性について、大規模かつ詳細な実験を通して、定量的並びに定性的に検証する。

2. 従来のストローク HMM

以下、 $xy_t = (x_t, y_t)$ を時刻 t における座標特徴とし、その時系列 $xy_1, \dots, xy_t, \dots, xy_T$ により单一のストロークを表現する。また $\theta_t = \arg(xy_t - xy_{t-1})$ により、時刻 t での方向特徴を表す。

2.1 方向特徴のみを用いた HMM

図 1 に典型的な従来法の一つである方向特徴 θ_t を出力シンボルとするストローク HMM (以下、 θ -HMM) [1] ~ [4] を示す。1. でも述べたように、方向特徴 θ_t のみを出力シンボルとし、座標特徴 xy_t を無視することで、特徴量の線分内定常性を HMM の自己遷移により適切に表現できる。

この θ -HMM の学習、すなわち状態 i から j への遷移確率 $a_{i,j}$ 及び状態 i から j への遷移で方向特徴 θ_t を

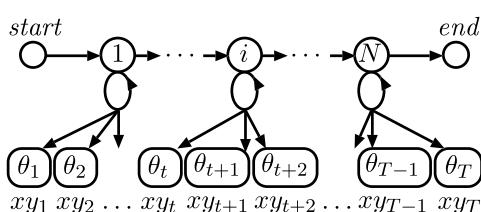


図 1 従来の HMM (θ -HMM)

Fig. 1 A conventional HMM (θ -HMM).

観測する確率 $b_{i,j}(\theta_t)$ の学習は、極めて一般的な HMM の枠組みで行われる。例えば、図 1 の left-to-right 型 HMM の学習では、まず前向き変数 $\alpha_t(i)$ (時刻 1 から始まって時刻 t で状態 i にある確率) を、広く知られた以下の漸化式を用いて効率的に計算する。

$$\begin{aligned}\alpha_t(i) &= \alpha_{t-1}(i-1) a_{i-1,i} b_{i-1,i}(\theta_t) \\ &\quad + \alpha_{t-1}(i) a_{i,i} b_{i,i}(\theta_t)\end{aligned}\quad (1)$$

なお θ -HMM においては $b_{i-1,i}(\theta_t) = b_{i,i}(\theta_t)$ であるが、後の議論との整合性のためにここでは区別している。後向き変数 $\beta_t(i)$ についても同様に計算する。

HMM パラメータは、以下のように通常の Baum-Welch アルゴリズムを適用することにより再推定される。出力シンボル確率 $b_{i,j}(\theta_t)$ を正規分布で表現し、 $b_{i,j}(\theta_t | \mu_{i,j}, \Sigma_{i,j})$ とする。ただし、 $\mu_{i,j}, \Sigma_{i,j}$ は出力シンボル θ_t の平均、分散を表す。学習用データの引数を n とするとき、 $a_{i,j}, b_{i,j}(\theta_t | \mu_{i,j}, \Sigma_{i,j})$ の再推定式は以下のとおりである。

$$\begin{aligned}\gamma(i, j, t) &= \alpha_{t-1}(i) a_{i,j} b_{i,j}(\theta_t | \mu_{i,j}, \Sigma_{i,j}) \beta_t(j) \\ \hat{a}_{i,j} &= \frac{\sum_n \sum_t \gamma(i, j, t)}{\sum_n \sum_t \alpha_t(i) \beta_t(j)} \\ \hat{\mu}_{i,j} &= \frac{\sum_n \sum_t \gamma(i, j, t) \theta_t}{\sum_n \sum_t \gamma(i, j, t)} \\ \hat{\Sigma}_{i,j} &= \frac{\sum_n \sum_t \gamma(i, j, t) (\theta_t - \mu_{i,j})^2}{\sum_n \sum_t \gamma(i, j, t)}\end{aligned}$$

ここで左辺の $\hat{a}_{i,j}, \hat{\mu}_{i,j}, \hat{\Sigma}_{i,j}$ はパラメータの再推定値である。HMM パラメータの値を再推定値に置き換え、パラメータの値が収束するまで再推定を繰り返す。

1. で述べたように、方向特徴だけではストロークの位置や長さが明示的に制御できないという問題がある。この問題は、多画文字における筆順フリー文字認識において特に深刻になる。例えば、筆順フリー条件のもと、座標特徴を無視して方向特徴だけを用いると、“四”は“兄”に誤認識される可能性が高くなる。文献 [1] ~ [3] では、起こり得るすべての筆順変動についてそれぞれ標準パターンを準備し、更に必要に応じてペンアップ部分を接続することにより、事実上筆順固定問題とすることで対処している。若しくは文献 [4] のように部首単位の位置制御などに座標特徴を補助的に利用する形態がとられる。

2.2 座標特徴を用いた HMM

図 2 及び図 3 は、出力シンボルに座標特徴を用い

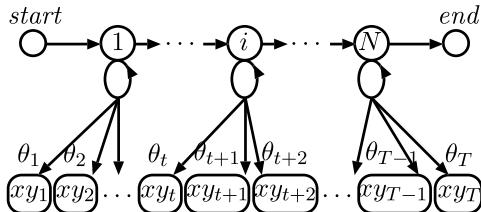


図 2 座標特徴を用いた HMM (1) (xy-HMM)
Fig. 2 HMM with coordinate feature (xy-HMM).

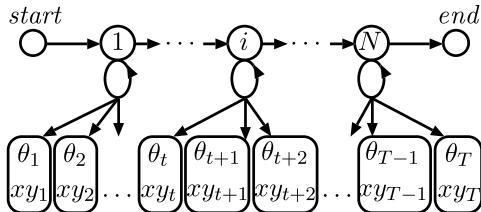


図 3 座標特徴を用いた HMM (2) (xytheta-HMM)
Fig. 3 Another HMM with coordinate feature (xytheta-HMM).

た HMM である。このうち前者は、座標特徴のみを出力シンボルとするもの [3] で、以下 *xy-HMM* と呼ぶ。一方後者は各状態で座標特徴と方向特徴を同時に出力するもの [5] で、以下 *xyθ-HMM* と呼ぶ。

先に述べたように、座標特徴 xy_t を用いた HMM は従来あまり用いられていない。これは座標特徴がストロークの一線分内で定常的とはみなされず、一線分内での座標特徴の変化を表現するには一状態では不十分であるためと予想される。すなわち座標特徴は時間とともに変化する特徴であり、したがって基本的に非定常的である。座標特徴を使うために、状態数を大幅に増やし、サンプル点数 T と同程度にすることも考えられるが、それでは HMM を使う意義が失われる。実際、状態数を増やすならば、Stochastic DP [6] や Bahlmann ら [7] の検討に見られるように、HMM による確率過程としての枠組みではなく、決定論的な枠組みにする方が妥当であろう。加えて、学習すべきパラメータももとの HMM 状態数 N に対して T/N 倍（一般に $N \ll T$ ）に増えるため、学習パターンの不足も問題となる。

3. 座標/方向出力型 HMM

2. で述べたように、方向特徴量はそのストローク形記述能力に限界があり、座標特徴は HMM との相性が良くない。そこで両者を適切に使い分ける工夫が必

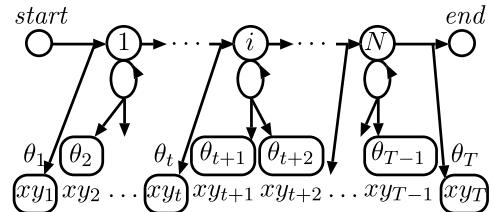


図 4 本手法の HMM ((xy/θ)-HMM)
Fig. 4 The proposed HMM ((xy/θ)-HMM).

要である。これまでに θ -HMM と *xy-HMM* を加重結合した事例 [3] もあるが、本質的な解決にはなっていないと思われる。実際、その *xy-HMM* には 2.2 の問題点がそのまま残っている。また方向特徴と座標特徴が HMM トレリス上で独立に非線形時間伸縮することになり、整合性の面でも問題がある。

本論文で提案する (xy/θ) -HMM は、方向特徴と座標特徴を単一の HMM 上で適切に使い分けたものである。図 4 にこの (xy/θ) -HMM を示す。この“適切な使い分け”とは、各特徴と定常・非定常性を考慮し、ストロークを線分近似した各線分の始端・終端では座標特徴、始端以降の各線分内では方向特徴を使用する方式を指す。HMM 上では、過渡的な座標特徴 xy は状態間遷移で 1 回だけ出力され、定常的な方向特徴 θ は状態内遷移（自己遷移）で繰り返し出力される。言い換えると、状態間遷移によりストロークの始端・屈曲点・終端の座標分布が規定され、状態内遷移においてストロークを構成する各線分の方向分布が規定される。

この (xy/θ) -HMM の学習に使用される前向き変数 $\alpha_t(i)$ は、以下のように式 (1) を若干変更した漸化式によって計算される。

$$\begin{aligned} \alpha_t(i) = & \alpha_{t-1}(i-1) a_{i-1,i} b_{i-1,i}(xy_t) \\ & + \alpha_{t-1}(i) a_{i,i} b_{i,i}(\theta_t) \end{aligned} \quad (2)$$

ここで $b_{i-1,i}(xy_t)$ は二次元正規分布である。後向き変数も同様な方法で計算される。式 (1) に対し、式 (2) は、状態遷移の違いに応じて出力シンボルを使い分けていることが分かる。 (xy/θ) -HMM の学習は、こうして計算される前向き変数及び後向き変数を用いて、従来どおりの Baum-Welch アルゴリズムで計算される。

4. 筆順フリーなオンライン文字認識におけるストローク HMM の利用

本論文では主として筆順フリーなオンライン文字認識実験を通して、各種ストローク HMM の性能を比

較評価する。筆順フリー認識のためには、 K ストロークからなる入力パターン $I = I_1, I_2, \dots, I_K$ と標準パターン $R = R_1, R_2, \dots, R_K$ の間に最適な 1 対 1 ストローク対応を求める必要がある。これら最適対応するストローク間の類似度を文字全体で総和したものが、筆順によらない I と R 間の類似度となる。ストローク対応の最適化手法はこれまでに様々な手法が提案されている（例えば [8] ~ [12]）。いずれの手法も比較評価実験に利用可能であるが、本論文では大局的に最適なストローク対応を求めるべく、キューブサーチ法 [11], [12] を採用した。

この筆順フリー認識の枠組みにおいて、ストローク HMM はストローク間の類似度として確率ゆう度の計算に利用される。具体的には、 R_l を表現するストローク HMM に I_k を入力して得られるゆう度が R_l と I_k 間の類似度となる。ストローク HMM により、ストロークの局所的な変動を吸収しながら、類似ストロークの位置関係の差異を適切に評価できれば、認識に適した類似度が与えられ、結果的に高い認識率が得られる。

画数変動問題は、画数 K の入力文字と画数 L の標準文字の間の最適な K 対 L 対応を求めるにあり、本論文の主題である画の表現方法に対して比較的独立した問題である。対策としては、画数変動を標準モデルの分岐で表現したり、主要な画数変動に標準モデルを準備する手法が挙げられる。一方、上述のキューブサーチ法を画数変動に対応するべく発展させた多層キューブサーチ法 [12] の利用も考えられる。

5. 実験

5.1 データセット

実験用データセットは東京農工大のオンライン文字データベース “HANDS-kuchibue.d-97-06-10” [13] 中の教育漢字を使用した。4. で述べた画数変動問題の解釈から、本論文では画数固定条件による実験としたため、正規画数で書かれたものだけを対象とした。画数別の傾向を観察するべく、5, 10, 15, 20 画のカテゴリーのみを用いた。本手法では画単位の HMM モデルを適切に学習させるために、画対応が既知で HMM モデル学習に十分な量の文字データ集合を用意する必要がある。本論文では画対応既知の条件として正規筆順を重視した。更に後述する三重交差確認法においてストローク HMM 学習用に 20 パターンの正規画数・正規筆順データを確保するために、正規画数・正規筆

表 1 実験で用いたデータセット

Table 1 Statistics of dataset.

#strokes	5	10	15	20
#categories	67	76	21	2
#samples	28,525	17,605	3,336	624

順パターン数が 30 未満のカテゴリーについては除外した。表 1 に実験対象となったパターン数などを示す。

各パターンについて前処理及び特徴抽出を行った。具体的には、まず 128×128 への線形スケーリング、平滑化、リサンプリングを適用した。次にこうして得られた座標特徴量系列から方向特徴量を求めた。方向特徴量については、その分布が 2.1 で述べたように一次元正規分布すなわち連続分布として表現されるため、8 方向や 16 方向などの量子化は特段行っていない。

5.2 ストローク HMM の学習

各カテゴリーより正規筆順パターンのみを抽出し、それらの $2/3$ を用いてストローク HMM を学習した。残る $1/3$ は 5.3 の認識実験の評価用パターンに加えた。ストローク HMM としては、2. で挙げた θ -HMM, xy -HMM, $xy\theta$ -HMM, 及び本手法 (xy/θ)-HMM を用いた。各ストローク HMM の状態数（開始・終了状態を除く）は、対応するストロークの標準形を折れ線近似した際の線分数とした。例えば “|” “ノ” “—” 状のストロークならば 1 状態，“フ” 状ならば 2 状態とした。折れ線近似における最大線分数は 4 (“乙” 状ストローク及び “しんじょう” 2 画目）とした。

カテゴリー “右” の全 5 ストロークについて、 xy -HMM, $xy\theta$ -HMM, (xy/θ) -HMM の各手法が学習によって獲得した出力シンボルの分布を図 5 に示す。それぞれの機能円の中心と形状は座標特徴の平均と分散 (2σ 等分散面) を表す^(注1)。ストローク HMM 内の状態の時系列は、機能円の線種を実線、破線、一点鎖線の順に変えて表現している。また機能円中心から出ている 3 本の線分は、方向特徴の平均と分散 ($\pm 2\sigma$ 方向) を表す。

同図より、 xy -HMM と $xy\theta$ -HMM の座標特徴の分布は、線分全体を無理に定常的なものとみなしたために線分の伸長方向に細長く、かつ線分を包含する広範囲な分布となっている。この場合、想定外のストローク形状であっても、その分布内に含まれてさえいれば不適に高いゆう度を与えてしまう点が問題になる。例

(注1): $xy\theta$ -HMM の機能円形状は、座標特徴の分散 (2σ 等分散面 (機能円球)) の筆記方向平均値による断面で表している。

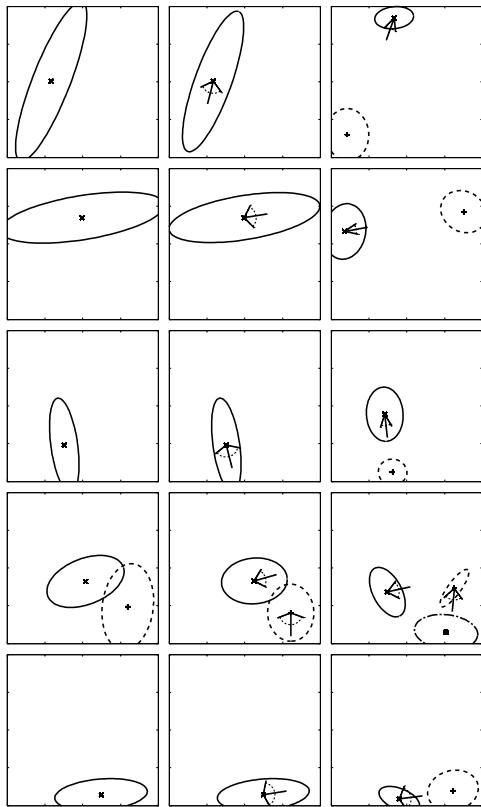


図 5 HMM の出力シンボル分布。上より “右” の第 1 (ノ) , 2 (一) , 3 ($|$) , 4 (ヽ) , 5 (ヽ) ストロークの HMM。左列 : xy -HMM , 中央列 : $xy\theta$ -HMM , 右列 : (xy/θ) -HMM

Fig. 5 Symbol distributions of trained stroke HMMs.

えば画像の左端から右端まで広がる長い水平線分を学習した HMM は、分布内にある半分の長さの線分や、分布内を左右に大きく行き戻りするようなストロークにも高いゆう度を与えててしまう。

一方、本手法である (xy/θ) -HMM では、ストローク始端・終端・屈曲点の座標特徴を、適度な局所性を保ったまま獲得することに成功している。特に、4 画目 (“ ヽ ”)について注目すると、学習時に屈曲点を示的に教示していないにもかかわらず、学習結果として屈曲点の座標特徴の分布が極めて妥当に自動獲得されていることが分かる。また、座標特徴平均から方向特徴平均方向へのほぼ延長線上に次の座標特徴平均が存在することから、HMM の学習によって得られた始端・終端・屈曲点の座標特徴と線分区間内の方向特徴との関連性は高く、かつ破綻を来していないことも分かる。

表 2 筆順フリー条件における本手法及び他のストローク HMM による誤認識率 (%)

Table 2 Error rates by the proposed HMM and the other HMMs at stroke order-free recognition experiment.

#strokes, K	5	10	15	20
θ -HMM	closed	68.56	58.63	62.13
	open	69.42	59.73	62.13
xy -HMM	closed	9.49	4.21	2.38
	open	10.04	4.58	2.25
$xy\theta$ -HMM	closed	7.18	1.62	1.00
	open	7.81	1.91	1.08
(xy/θ) -HMM (proposed)	closed	3.43	0.37	0.15
	open	3.80	0.47	0.39
				0.0

5.3 筆順フリー条件での認識実験

4. で述べたように、主として筆順フリーなオンライン文字認識の枠組みにおいて各種ストローク HMM の性能を比較評価した。なお、多様な条件下で本手法と従来法の比較を行うべく、筆順固定の条件下での評価も行っており、これについては 5.4.3 にて後述する。

誤認識率は三重交差確認法により求めた。すなわち、全データを 3 分割し、そのうち二つをストローク HMM の学習用、残る一つを評価用として誤認識率を求める処理を 3 回繰り返し、その平均値を誤認識とした。以下、表中の “open” がそうして求めた誤認識率である。一方の “closed” は、参考のために示した学習用パターンに対する誤認識率である。

本手法及び他の HMM による誤認識率を表 2 に示す。全般的に座標特徴に方向特徴を併用した手法が低い誤認識率を得ている。更にその中でも本手法 (xy/θ) -HMM が最も低い誤認識率を与えた。実際、 $xy\theta$ -HMM と (xy/θ) -HMM を評価用パターン (open) について比べると、5 画で 51%，10 画で 75%，15 画で 64% の誤認識低減を達成した。20 画ではすべての評価用パターンが正しく認識された。

一方、同表からは θ -HMM の精度が極めて低いことも分かる。筆順フリー条件は、ストロークの位置情報が使えない θ -HMM にとって極めて不利な条件である。実際 θ -HMM は筆順固定条件下で利用されることが多い [1] ~ [4]。このため 5.4.3 では筆順固定条件下において θ -HMM と本手法を比較する。

ところで、図 4、図 5 から分かるように、本手法は座標特徴出力の機会が独立に存在する分、他のストローク HMM に比べて出力回数が多い。このため、上述の本手法の優位性が出力回数の差すなわちモデルの自由度の差によると解釈される可能性もある。そこ

表 3 状態数を 2 倍としたときの誤認識率 (%)

Table 3 Error rates by the conventional HMMs with double states.

#strokes, K	5	10	15	20	
xy -HMM	closed	6.83	3.03	1.93	0.3
w/ double states	open	7.23	3.37	1.98	0.2
$xy\theta$ -HMM	closed	6.17	1.45	1.21	0.1
w/ double states	open	7.00	1.89	1.14	0.2

表 4 本手法による改善 (improved), 改悪 (degraded) パターン数, 及び本手法・従来法ともに誤認識となったパターン数 (failed). 括弧内はその全データに占める割合 (%)

Table 4 The numbers of improved samples and degraded samples.

 $xy\theta$ -HMM \rightarrow (xy/θ)-HMM

#strokes, K	5	10	15	20	
improved	closed	2,978 (5.22)	496 (4.41)	58 (0.87)	0 (0.0)
	open	1,618 (5.67)	285 (1.62)	29 (0.87)	0 (0.0)
degraded	closed	838 (1.47)	57 (0.16)	1 (0.01)	1 (0.1)
	open	473 (1.66)	31 (0.18)	6 (0.18)	0 (0.0)
failed	closed	1,120 (1.96)	75 (0.21)	9 (0.09)	0 (0.0)
	open	611 (2.14)	52 (0.30)	7 (0.21)	0 (0.0)

で、本手法の優位性があくまで特徴量の選択的利用にあることを示すために、従来の HMM の状態数を 2 倍にして出力回数の差異をなくした上で、同様の実験を行った。結果を表 3 に示す。 xy -HMM では認識性能が向上したが、 $xy\theta$ -HMM では一部に認識性能の低下すら見られた。いずれにせよ、状態数を 2 倍にしても (xy/θ)-HMM の認識性能には及ばないことが分かる。

5.4 本手法による改善及び改悪

表 4 に、従来法の中で最も高い認識性能を与えた $xy\theta$ -HMM に対する本手法の改善パターン数 (improved) 及び改悪パターン数 (degraded) を示す。なお同表中の failed は $xy\theta$ -HMM でも本手法でも誤認識されたパターン数である。

同表から、いずれの画数 K においても、改善パターンの方が改悪パターンや誤認識のままのパターンに比べて優勢であることが分かる。すなわち、多少の副作用はあるものの本手法によって大多数の誤認識が改善されている。以下では、具体的な改善例及び改悪例を挙げ、それらが生じた理由について考察する。

5.4.1 改善例

表 5 に、本手法による改善パターンのうち主要なもの (5 画) を示す。また、ここで述べる改善例の実例を図 6 に示す。例えば、 $xy\theta$ -HMM では “末” に誤認識していた 249 サンプルが (xy/θ)-HMM では “末” に正しく認識されたことを表す。図 6 上段に示すように、この例では (xy/θ)-HMM の適用によってストローク

表 5 本手法による改善例 (5 画)

Table 5 Major improved results.

#samples	$xy\theta$ -HMM \rightarrow (xy/θ)-HMM	xy -HMM \rightarrow (xy/θ)-HMM	θ -HMM \rightarrow (xy/θ)-HMM
249	未 \rightarrow 未	由 \rightarrow 生	司 \rightarrow 目
228	布 \rightarrow 市	主 \rightarrow 立	正 \rightarrow 主
227	由 \rightarrow 生	布 \rightarrow 市	功 \rightarrow 用
165	皿 \rightarrow 四	田 \rightarrow 四	司 \rightarrow 田
145	皿 \rightarrow 立	失 \rightarrow 矢	犯 \rightarrow 他

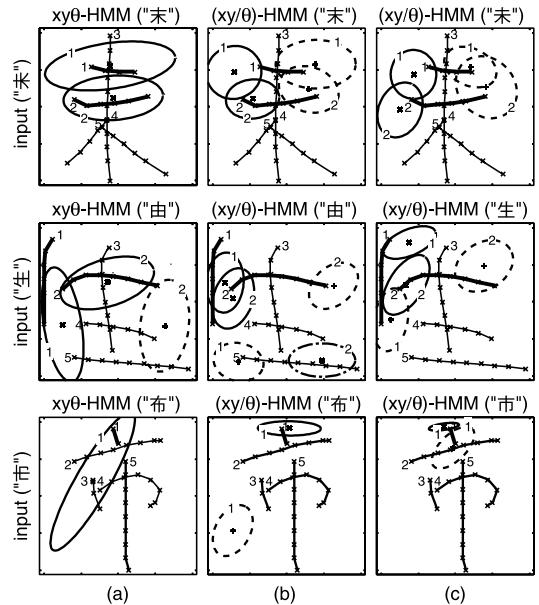


図 6 $xy\theta$ -HMM での誤認識が (xy/θ)-HMM で改善された例。(a) $xy\theta$ -HMM による誤認識,(b) 誤認識カテゴリーの (xy/θ)-HMM 表現,(c) 正解カテゴリーの (xy/θ)-HMM 表現

Fig. 6 Samples improved by (xy/θ)-HMM.

の端点の位置関係がより明確に表現された結果，“未”と“末”的各 HMM の 1 画目と 2 画目の端点の到達位置の違いが評価され、正答に転じたと判断される。

次に、 $xy\theta$ -HMM によって “由” に誤認識していた 227 サンプルが (xy/θ)-HMM では “生” に正しく認識された事例に着目する。図 6 中段に示すように、この改善は本手法の二つの効果による。第一は、“由” 1 画目の位置評価における改善効果である。すなわち、“由” 1 画目 (左端の “|”) が文字の下方に分布するのに対して、“生” 1 画目 (“ノ”) は文字の上方に分布する。この 1 画目の始端・終端座標分布の違いが何度も差を生じさせ、それぞれを正しいカテゴリーに分類させたと考えられる。第二は、 $xy\theta$ -HMM では座標特徴の分布が広範囲となつたために起つっていた問題

が、 (xy/θ) -HMM で解消されたという効果である。具体的には、 $xy\theta$ -HMM では“由”2画目(“ヽ”)の屈曲点から下方向に延びて終端に至る部分が大幅に省略され、“生”2画目(“-”)に不当に高い類似度で対応してしまう問題があった。一方、 (xy/θ) -HMM では、終端範囲が明確になり、屈曲点から終端までがある程度の長さをとるようになった結果、上記の合わせすぎの問題を回避できた。

“布”の228サンプルが“市”に正しく認識された事例も上の効果による。これら2カテゴリーは、1画目の後半以外はほぼ同一な構造をもつ。図6下段に示すように、 $xy\theta$ -HMM では“布”1画目(“ノ”)のストロークの後半部分が省略され、“市”1画目(“|”)に高い類似度で対応していた。この省略が (xy/θ) -HMM では回避され、正解に転じたものと考えられる。部分ストローク欠損の有無による改善例は、これら以外にも“皿”→“四”(4画)、“皿”→“立”(2画)等が挙げられる。

最後に θ -HMMとの比較において特徴的な例として、“功”に誤認識していた769サンプルが (xy/θ) -HMMでは“用”に正しく認識された事例を挙げる。 θ -HMMでは位置や長さ情報が欠損しているため、筆順フリー条件下では方向特徴上の類似ストロークがすべてそろっている“功”と正解の“用”との区別がつかなかつた。一方 (xy/θ) -HMMでは座標特徴の利用によりゆう度に差が生じ、正しく認識されるようになったと解釈できる。

5.4.2 改悪例

5.2で述べた本手法の長所は、そのまま短所にもなり得る。すなわち、ストローク端点における大きな位置ずれや各種ノイズ要因に対して、本手法は xy -HMM や $xy\theta$ -HMMよりも敏感であり、影響を受けやすい。

表6に、筆順フリー条件下での認識実験において本手法による主な改悪パターン(5画)を示す。“本”が“末”や“未”に誤認識したパターンを観察したこと、 (xy/θ) -HMM でストローク端点座標ゆう度分布

表6 本手法による改悪例(5画)
Table 6 Major degraded results.

#samples	$xy\theta$ -HMM → (xy/θ) -HMM	xy -HMM → (xy/θ) -HMM	θ -HMM → (xy/θ) -HMM
206	本 → 末	188 未 → 末	22 用 → 田
119	本 → 未	185 本 → 未	14 左 → 壓
98	右 → 石	159 右 → 石	8 本 → 未
90	未 → 末	111 本 → 未	7 矢 → 失
79	未 → 未	80 失 → 矢	7 代 → 他

から外れる要因として、長さが極端に長い/短いストローク、標準的な位置からはずれた場所に書かれたストロークが見られた。図7に、以上の要因が観測されたパターン事例を示す。これらは (xy/θ) -HMM でストローク端点座標ゆう度分布から外れる要因となり、誤認識を起こしたものと考えられる。

5.4.3 筆順固定下の認識結果

筆順フリー条件下では認識性能が極端に低かった θ -HMM に対して、通常用いられている筆順固定条件下での認識実験を行い、同一条件下での (xy/θ) -HMM の性能と比較した。筆順固定とするため、本実験ではデータセットから正規筆順のパターンだけを利用した。

更に、ペンアップ時の取扱い方の異なる次の2通りのゆう度評価を行った。一つ目は、ペンドウン時の各ストロークのゆう度のみ使用しペンアップ時の情報を使用しない方法である。二つ目は、文献[1]を模した形態である。すなわち、ペンアップ時の画間移動方向を一つの方向出力シンボルとして表現してペンドウン時の方向特徴とは区別する。ストローク HMM 学習時

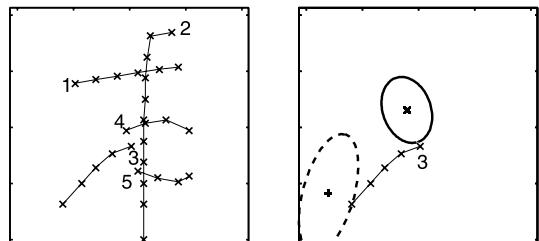


図7 改悪例のパターン事例(数字はストローク番号、右は出力シンボル分布との比較(3画目))

Fig. 7 Degraded patterns.

表7 筆順固定条件下における本手法及び従来法の誤認識率(%)

Table 7 Error rates by the proposed HMM and the other HMMs at stroke order-fixed recognition experiment.

without pen-up information

#strokes, K	5	10	15	20
θ -HMM	closed	30.28	1.10	0.0
	open	31.04	1.46	0.1
(xy/θ) -HMM (proposed)	closed	1.58	0.09	0.0
	open	1.78	0.12	0.0

with pen-up 1 information

#strokes, K	5	10	15	20
θ -HMM	closed	8.17	0.14	0.0
	open	8.62	0.14	0.0
(xy/θ) -HMM (proposed)	closed	1.24	0.03	0.0
	open	1.38	0.04	0.0

にペンアップ時の画間移動方向分布を学習することにより、各ストロークのゆう度とペンアップ時の画間移動方向のゆう度を連結して評価する方法である。

表 7 に本手法及び従来法の θ -HMM の認識率を示す。いずれの場合でも θ -HMM より (xy/θ) -HMM の方が低い誤認識率を示す。このことから、方向特徴だけでは識別が難しい類似ストロークに対して、 (xy/θ) -HMM の座標特徴が識別の一助となることが分かる。この結果は、本手法が英筆記体文字など 1 画からなる他の文字の認識にも有効であることを示唆している。

6. む す び

方向特徴の線分内定常性と座標特徴の非定常性を生かして両特徴量を適切に使い分けたストローク HMM の表現方法として (xy/θ) -HMM を提案した。HMM の状態内の自己遷移では方向特徴、状態間遷移では座標特徴というように、状態遷移に応じて出力シンボルを使い分けながらも、従来法と同等の Baum-Welch アルゴリズムの枠組みで HMM の学習・認識が可能であることを実験により示した。特に、学習の結果、極めて妥当な出力シンボル分布が獲得されたことも確認した。筆順フリー及び筆順固定条件下での認識実験における誤認識率の改善によって従来法に対する本手法の有効性を確認した。

今後の課題は、HMM のシンボル出力を状況に応じて使い分けるという本手法の基本構想を更に発展させることや、改悪例への対策が挙げられる。また、オンライン文字認識が一種の行動認識であることをかんがみ、他の行動パターン認識への適用も考えられる。

実験規模の拡大も一つの課題であろう。すなわち、今回は画数固定の条件のもと、5, 10, 15, 20 画のカテゴリーのみを用いて手法の有効性を示したが、例えば HANDS-kuchibue のすべてのデータを認識対象とした実験も必要であろう。この場合、画数変動と 20 画以上の多画文字への対応が主たる問題となる。前者については、4. で挙げたように、主要な画接続が施された文字を標準パターンに登録するといった最も単純な対処法に加え、本論文で用いたキューブサーチ法の画接続対応型である多層キューブサーチ法 [12] を用いることも考えられる。また前者及び後者に共通して、計算量の低減が重要となる。これについても種々の対策が考えられる。例えば大分類の導入やキューブサーチ過程への枝刈りの導入 [11] はその一例である。いずれにおいても、本論文で提案した高精度なストローク

HMM の利用によって正解カテゴリーと不正解カテゴリーのゆう度差を大きくできる可能性が高く、こうした計算量低減手法の効果を強めることができると期待できる。

謝辞 当研究室在籍時に、ストローク HMM の構成に関する基本的研究に協力して頂いた菅沢昭善氏並びに奥村大樹氏に感謝する。

文 献

- [1] 中井 满, 嵐嶋山茂樹, 下平 博, “サブストローク HMM を用いたオンライン手書き文字認識,” 信学論 (D-II), vol.J88-D-II, no.9, pp.1825–1835, Sept. 2005.
- [2] 高橋賢一郎, 安田英史, 松本 隆, “Hidden Markov Model を用いたオンライン手書き文字認識,” 信学技報, PRMU96-211, 1997.
- [3] 伊藤 等, 中川正樹, “Hidden Markov Model に基づくオンライン手書き文字認識,” 信学技報, IE97-54, 1997.
- [4] 徳野淳子, 楊 一平, Gleidson Pegoretti da Silva, 末代誠仁, 中川正樹, “構造字体表現型 HMM に基づくオンライン手書き文字認識における座標特徴の利用法と効果,” 信学技報, PRMU2005-222, 2006.
- [5] K.C. Jung, S.M. Yoon, and H.J. Kim, “Continuous HMM applied to quantization of on-line Korean character spaces,” Pattern Recognit. Lett., vol.21, no.4, pp.303–310, 2000.
- [6] 中川聖一, “ストキャスティック DP 法および統計的手法による不特定話者の英語子音の認識,” 信学論 (D), vol.J70-D, no.1, pp.155–163, Jan. 1987.
- [7] C. Bahlmann and H. Burkhardt, “The writer independent online handwriting recognition system *frog on hand* and cluster generative statistical dynamic time warping,” IEEE Trans. Pattern Anal. Mach. Intell., vol.26, no.2, pp.299–310, 2004.
- [8] 小高和己, 若原 徹, 増田 功, “筆順に依存しないオンライン手書き文字認識アルゴリズム,” 信学論 (D), vol.J65-D, no.6, pp.679–686, June 1982.
- [9] C.K. Lin, K.C. Fan, and F.T.P. Lee, “On-line recognition by deviation-expansion model and dynamic programming matching,” Pattern Recognit., vol.26, no.2, pp.259–268, 1993.
- [10] A.J. Hsieh, K.C. Fan, and T.I. Fan, “Bipartite weighted matching for on-line handwritten Chinese character recognition,” Pattern Recognit., vol.28, no.2, pp.143–151, 1995.
- [11] 迫江博昭, 懇 重弼, “筆順フリーなオンライン文字認識のための画対応サーチアルゴリズム,” 九州大学大学院システム情報科学研究科報告, vol.2, no.1, pp.99–104, 1997.
- [12] 懇 重弼, 迫江博昭, “筆順・画数自由オンライン文字認識のための画対応決定法—多層キューブサーチ,” 信学論 (D-II), vol.J82-D-II, no.2, pp.230–239, Feb. 1999.
- [13] 松本 騒, 中川正樹, “オンライン手書き日本語文字パターンの解析,” 信学論 (D-II), vol.J87-D-II, no.5, pp.1112–1119, May 2004.

(平成 19 年 10 月 24 日受付, 20 年 2 月 19 日再受付)



片山 喜規（正員）

昭 62 九大・工・電子卒。平元同大大学院修士課程（電子）了。同年九大・工・情報工学科助手。現在同大システム情報科学研究院知能システム学部門准助教。修士（工学）。時系列パターン解析・認識に関する研究に従事。日本音響学会会員。



内田 誠一（正員）

平 2 九大・工・電子卒。平 4 同大大学院修士課程（情報）了。セコム（株）勤務を経て、現在、同大システム情報科学研究院知能システム学部門教授。博士（工学）。画像パターン及び時系列パターンの解析・認識に関する研究に従事。平 15 本会 PRMU 研究奨励賞、平 18 MIRU 長尾賞、平 19 ICDAR2007 Best Paper Award 各受賞。IEEE、情報処理学会、映像情報メディア学会各会員。



迫江 博昭（正員：フェロー）

昭 41 九大工・制御卒。昭 43 九大大学院修士課程（電子）了。同年日本電気（株）入社。平元、九大に移り、システム情報科学研究院知能システム学部門教授を経て、現在同大名誉教授。工博。音声認識及び画像パターン解析の研究に従事。IEEE ASSP Senior Award、本会業績賞、米沢ファウンダーズメダル受賞記念特別賞及び論文賞、（社）発明協会・内閣総理大臣発明賞、九大嘉村記念賞各受賞。