

筆順変動を表現するHMMとそのオンライン文字認識への応用

片山, 喜規
九州大学大学院システム情報科学研究院

内田, 誠一
九州大学大学院システム情報科学研究院

迫江, 博昭
九州大学大学院システム情報科学研究院

<https://hdl.handle.net/2324/7402679>

出版情報：電子情報通信学会論文誌 D. J91-D (5), pp.1434-1441, 2008-05-01. 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



筆順変動を表現する HMM とそのオンライン文字認識への応用

片山 喜規^{†a)} 内田 誠一^{†b)} 迫江 博昭[†]

An HMM Representing Stroke Order Variations and Its Application to Online Character Recognition

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

あらまし 本論文では、筆順フリーなオンライン文字認識の高精度化を目指し、(i) 筆順変動の統計的モデルの構築、及び(ii) その認識における利用、の2点について検討する。一般に筆順フリー化には不自然な画対応の許容による誤認識の問題があるが、提案する筆順変動モデルを用いることでそれらを抑制できる。この筆順変動モデルは、筆順フリー認識のためのグラフモデル（キューブグラフ）の確率的拡張として定式化され、結果的に文字形状に関するゆう度と筆順のゆう度を同時に扱うことが可能な隠れマルコフモデル（HMM）の一種となる。公開されているオンライン文字データベース“HANDS-kuchibue-d-97-06-10”を用いた認識実験により、筆順変動モデル導入の有効性及び妥当性を明らかにした。

キーワード オンライン文字認識、筆順フリー、筆順変動モデル、隠れマルコフモデル、キューブサーチ

1. まえがき

漢字など多画文字のオンライン認識においては、筆順変動が問題となる。すなわち、筆順変動を伴った入力パターンは、正規筆順で登録されている標準パターンと部分的若しくは全体的に大きく異なったデータとなり、誤認識される可能性が高くなる。そこで、筆順変動があっても正しく認識できる手法（筆順フリーなオンライン文字認識手法）が多数検討されてきた[1], [2]。

後述するように、筆順フリーなオンライン文字認識手法にはいくつかの実現方式がある。これらの中に、入力パターンと標準パターン間に最適な画対応を決定することで、筆順フリー性を実現するという方式がある。適切な画対応が求まれば、その対応下での両パターン間の類似度は筆順によらず一定となるために、筆順変動に頑強な認識を実現できる。

しかしこの最適画対応方式には副作用もある。すなわち、画の形状や位置に基づいて画対応を定めると、異なるカテゴリーの文字パターンであっても強引な画

対応付けにより類似度が過大評価される場合があり、結果的に誤認識を起こすという問題もあった。例えば、正規筆順で書かれた入力パターン“本”的各画を標準パターン“末”的第1, 3, 4, 5, 2画にそれぞれ対応づけることで誤認識してしまう可能性が高くなる。

本論文では、筆順フリーなオンライン文字認識の高精度化を図るべく、筆順変動の統計的モデル（キューブHMM）の構築及びその認識における利用について検討する。認識の際に、この筆順変動モデルを用いながら画対応並びに類似度を求めて、起こりやすい画対応のみを許容し、強引な画対応を排除することを可能にする。先の“本”と“末”的例では、“末”において第5画 → 第2画の順に書かれることはほとんどないために、それをモデルに反映できれば、誤認識を回避できると考えられる。

本手法は、従来の筆順フリー認識手法であるキューブサーチ法[3], [4]を発展させたものである。後述するように、キューブサーチ法では、すべての筆順を表現するグラフ（キューブグラフ）を用い、その始端ノードから終端ノードまでの可能なすべての経路により、すべての筆順を表現する。そして最適画対応決定問題をこのグラフ上の最適経路探索問題として表現する。本手法ではキューブグラフを一種の隠れマルコフモデル（HMM）と見て、グラフの各エッジの遷移確率とし

[†]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: yoshinori@is.kyushu-u.ac.jp

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

て筆順変動の統計モデルすなわちキューブ HMM を構築する。簡単にいえば、上述の各経路についてその確率が計算されることになり、これがその経路で表現される筆順の生起確率となる。遷移確率は通常の HMM と同様、学習により自動的に設定できる。このようにグラフに基づくキューブサーチ法を出発点として、HMM という非常に自然な形式で筆順変動モデルを構築できる。

2. 関連技術

2.1 従来の筆順フリー認識手法

筆順フリーなオンライン文字認識には、大きく 2 種類の方式がある。

- 第 1 は、予想される筆順変動をすべて標準パターンとして登録しておくという方式である [5] ~ [7]。例えば「右」「女」「必」といった筆順変動が起こりやすい文字カテゴリーについては、複数の標準パターンを準備する。極めて実際的な戦略であるものの、当然ながら予想外の筆順変動には対応できない。更に続け書きによる画数変動が起きた場合まで考えると、相当に大量の標準パターンを準備しておかなくてはならない。
- 第 2 は、入力及び標準パターン間に最適な画対応を求め、その対応下で両パターンの類似度を求めるという方式である [3], [4], [8] ~ [11]。画対応を探索するため、第 1 の戦略に比べて計算量を要することが多いが、上述の問題を解消できるため、より本質的な筆順フリー化戦略であるといえよう。加えて多くの場合続け書きに対応するための拡張が可能であることとは大きな特長である。

このように第 2 の方式すなわち最適画対応決定に基づく方式は、その拡張性を含めて重要なアプローチであるといえる。本手法もこの第 2 の方式の一種である。

従来、この最適画対応決定問題は、画間類似度の文字全体での総和を 1 対 1 対応制約条件のもとで最小化する問題として定式化されてきた。この類似度は、画の形状や画の位置といった幾何的情報をもとに設計されることが多い。その結果、入力パターンを表す画の集合と、標準パターンを表す画の集合間に、画の幾何的類似度評価に基づいて 1 対 1 対応を定めることになる。筆順情報を用いることなく、集合の要素間の単純な対応付け問題としているので、対応付けの自由度が非常に高い。

しかし、この自由度の高さは第 2 の方式の問題点にもなっている。すなわち、1. の“本”と“末”的例のように、実際には起こり得ないような対応付けをも許

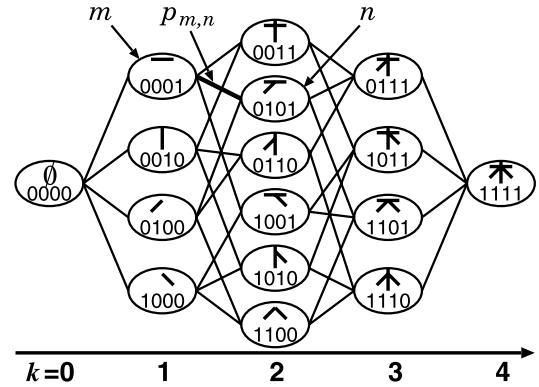


Fig. 1 キューブグラフ
Cube graph for stroke order-free recognition.

容してしまうということにより、異なるカテゴリーのパターン間の類似度が過大評価される危険性がある。実際、後述する実験結果から明らかになるように、この問題による誤認識は非常に多い。

本手法では、筆順変動モデルを導入することでこの対応付けの自由度を制御し、出現頻度の低い筆順についてはその頻度に応じて抑制することを目指す。この導入により、非常にまれな筆順で書かれた入力パターンについては誤認識となる可能性があるが、上記第 1 の方式のように完全に未登録の筆順変動を排除する枠組みとは異なるため、こうした副作用は少ないと考える。

なお、以上第 1、第 2 の方式とは別に、最近では、オンライン文字パターンを画像化することで筆順情報を含めた動的情報を消去し、オフライン文字認識の技法を用いて認識するという方式も検討されている [12], [13]。ただし、実際には動的情報を完全に消去することは少なく、多くはオンライン文字認識手法との相補的なハイブリッド構成をとる。したがって本論文の結果もこのハイブリッド方式の性能向上に資することができる。

2.2 キューブサーチ [3], [4]

本節では、本手法の出発点となるキューブサーチ法について概観する。既述のように、キューブサーチ法は最適画対応を決定する方式の一つであり、キューブグラフを用いて各文字のすべての筆順を表現する。図 1 は 4 画文字“木”に関するキューブグラフである。このグラフの始端ノードから終端ノードまでの可能なすべての経路により、“木”的すべての筆順が表現されていることが分かる。なお各ノード内の 4 けたの 2 進数は、そのノードに到達した時点で対応済みの画番号を表現している。例えば 0101 は、入力パターンの

1画と2画が標準パターンの1画(一)と3画(ノ)若しくは3画と1画に対応済みであることを意味している。

キューブグラフの各エッジは特定の画対応を表現する。図1の例でいえば、ノード $m = 0001$ から $n = 0101$ へのエッジには標準パターンの第3画(ノ)と入力パターンの第2画の類似度(若しくは相違度)が計算される。このように、隣接ノード m, n により標準パターンと入力パターンのどの画を対応づけるかが一意に定まるが、 m や n は画番号を直接表しているのではない点には注意されたい。

最適画対応決定問題はキューブグラフ上の最適経路探索問題として表現される。最適化の基準は、キューブグラフの各エッジで計算される入力画と標準画の画間類似度である。この類似度の累積値が最大となる経路が最適経路であり、最適画対応、すなわち入力パターンの筆順を表現している。

他の手法[8]～[11]と同様、このキューブサーチも画の位置関係・形状といった幾何学的評価のみを考慮しながら画対応が最適化される。したがって、筆順変動の起こりやすさといった評価はなされていなかった。

3. キューブ HMM

本章では、提案するキューブ HMM の構造、学習法、並び文字認識における利用について述べる。以下では $k = 1, \dots, K$ を入力パターンの画番号とする。

3.1 キューブ HMM の構造

キューブ HMM は、図1のキューブグラフの各ノードを状態とみなし、エッジで結ばれた2状態 m, n 間に遷移確率 $p_{m,n}$ を付与したものである(ただし $\sum_n p_{m,n} = 1$)。したがってキューブ HMM は、キューブグラフをトポロジーとし、自己状態への遷移(セルフループ)をもたない left-to-right HMM である。

2.2 のキューブサーチと同様、キューブ HMM の各エッジも特定の画対応を表現する。例えば図1の状態 $m = 0001$ から $n = 0101$ へのエッジは、標準パターンの第3画(ノ)と入力パターンの第2画の対応付けを表現する。したがってそのエッジの通りやすさを示す $p_{m,n}$ は、その特定の画対応が起こりやすいか否かを表現している。遷移確率 $p_{m,n}$ は3.2で述べるように学習により決定される。

各エッジでは、確率 $p_{m,n}$ に加え、そのエッジが表現する対応画間の類似度 $q_{m,n}$ が計算される。通常の HMM の用語にたとえれば、 $p_{m,n}$ が状態遷移確率、

$q_{m,n}$ がシンボル出力確率に相当する。この場合のシンボルは、入力画を表す特徴ベクトル系列である。本論文では、時刻 t における座標特徴 (x_t, y_t) と方向特徴 θ_t による特徴ベクトル (x_t, y_t, θ_t) の時系列を用いた。

画間類似度 $q_{m,n}$ には様々な形態が考えられるが、本論文では、その計算にも HMM を利用することとする。すなわち標準パターンの各画についてその形状を表現する HMM(いわゆるストローク HMM)を準備しておく。入力画との画間類似度 $q_{m,n}$ はその HMM による入力画のゆう度として求められる。この場合、キューブ HMM の各エッジにストローク HMM が埋め込まれた構造となり、全体として一つの大規模な HMM となる。なおストローク HMM 自体にも様々な形態が考えられる。本論文では文献[15]の座標特徴と方向特徴を併用する手法を用いた^(注1)。ストローク HMM の状態数は画の屈曲点数に応じておよそ 1～4 状態とした。

3.2 キューブ HMM の学習

キューブ HMM の状態遷移確率 $p_{m,n}$ は、通常の HMM と同様、EM アルゴリズムによる反復学習(再推定)が可能である。以下、再推定に必要な前向き確率と後向き確率、及びそれらを用いた再推定の方式について述べる。なお、画間ゆう度 $q_{m,n}$ 計算のためのストローク HMM は既に学習済みであり、そのパラメータは固定とする(ただし、4.5においてストローク HMM の追加学習について実験的に検討する。)

3.2.1 前向き確率

まず入力画番号が 1 から順に k まできたときに状態 $m \in S(k)$ に存在し得る確率、いわゆる前向き確率 α_m を計算する。ここで $S(k)$ は入力の第 1 画から第 k 画までの画対応が終了したことを示す状態の集合である。4 画の例ならば、図2のように $S(2) = \{0011, 0101, 0110, 1001, 1010, 1100\}$ である。

具体的には α_m は以下のように計算される。

(1) 初期状態 $\alpha_0 = 1.0$

(2) $k = 1, \dots, K$ に対して以下を実行

(3) $\forall m \in S(k)$ に対して以下を実行

$$(4) \quad \alpha_m = \sum_{l \in \text{Pre}(m)} \alpha_l p_{l,m} q_{l,m} \quad (1)$$

ここで、 $\text{Pre}(m)$ は、状態 m に先行可能な状態の集合で

(注1): 具体的には、 N 本の線分で近似できる画について、 N 状態の left-to-right 型 HMM により表現し、ゆう度計算を行った。線分の始端では座標特徴を出し、線分内では方向特徴を出力する工夫を施し、各特徴の観測確率密度関数としてはガウス分布を用いた。

ある。同じく 4 画の場合、 $m = 0011$ ならば $\text{Pre}(m) = \{0001, 0010\}$ である。

3.2.2 後向き確率

前向き確率 α_m と同様、後向き確率 β_m についても計算する。後向き確率 β_m は入力画が第 k 画から第 K 画まで筆記されたとき、それが状態 $m \in S(k)$ から開始され得る確率であり、次式で計算される。

- (1) 状態 $m \in S(K)$ について $\beta_m = 1.0$
- (2) $k = K-1, \dots, 1, 0$ に対して以下を実行
- (3) $\forall m \in S(k)$ に対して以下を実行
- (4) $\beta_m = \sum_{n \in \text{Post}(m)} \beta_n p_{m,n} q_{m,n}$

ここで $\text{Post}(m)$ は、状態 m に後続可能な状態の集合である。4 画の場合、 $m = 0011$ ならば $\text{Post}(m) = \{0111, 1011\}$ である。

3.2.3 遷移確率 $p_{m,n}$ の再推定

上で得られた α_m, β_m から、以下のように EM アルゴリズムに基づく $p_{m,n}$ の再推定式が得られる。

$$\hat{p}_{m,n} = \frac{\alpha_m p_{m,n} q_{m,n} \beta_n}{\sum_{n' \in \text{Post}(m)} \alpha_m p_{m,n'} q_{m,n'} \beta_{n'}} \quad (2)$$

ただし再推定の初期値はすべて等確率すなわち $p_{m,n} = 1/|\text{Post}(m)|$ とした。なお、以下の実験では学習パターンセット全体での一括学習を行った。すなわち、一つの学習パターンごとに $p_{m,n}$ を修正するのではなく、学習パターンセット全体での総和を式 (2) の分子分母ともに適用した再推定式により、ゆう度の最大化を図った。

3.3 キューブ HMM を用いた認識

入力パターンの認識は、以上で準備された各文字カ

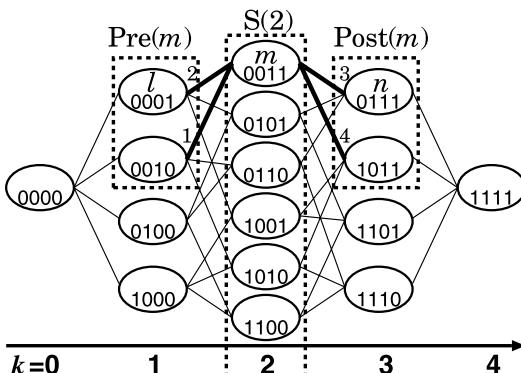


図 2 キューブ HMM における $S(k)$, $\text{Pre}(m)$, $\text{Post}(m)$

Fig. 2 Sets of states, $S(k)$, $\text{Pre}(m)$, and $\text{Post}(m)$.

テgorie のキューブ HMM について入力パターンのゆう度を計算すればよい。具体的には、まず入力パターンと標準パターンの各画について、ストローク HMM により画間ゆう度 $q_{m,n}$ を算出する。その後、以下に述べる計算法により、状態遷移確率 $p_{m,n}$ を考慮しながら、キューブ HMM によって最ゆう画対応及び文字全体のゆう度が決定される。最も高いゆう度を示したキューブ HMM のカテゴリーを認識結果とする。このようにキューブ HMM は、 $q_{m,n}$ による文字形状のゆう度と $p_{m,n}$ による筆順のゆう度の同時評価を実現している。

文字全体のゆう度の具体的計算法には 2 種類が考えられる。すなわち、最終状態 $m \in S(K)$ における前向き確率 α_m を用いる場合と、ビタビアルゴリズムによるゆう度を用いる場合である。後者はキューブ HMM 上の開始状態から最終状態までの経路で求まり得る最大確率であり、3.2.1 の漸化式 (1) を $\alpha_m = \max_{l \in \text{Pre}(m)} \alpha_l p_{l,m} q_{l,m}$ に置き換えた手続きで導出される。予備実験の結果、両計算法の認識率にほとんど違いが見られなかった。以下本論文では、ビタビアルゴリズムによるゆう度計算を行った結果を示している。

4. 実験

4.1 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” [14] 中の教育漢字でかつ正規画数で書かれたものを使用した。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。ただし、ストローク HMM の学習に必要となる正規画数・正規筆順パターン数が 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を行った。具体的には、 128×128 への線形スケーリング、平滑化、リサンプリング、及び方向量子化を適用した。

4.2 ストローク HMM の学習

各カテゴリーの正規筆順・画数パターン（前述のよ

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

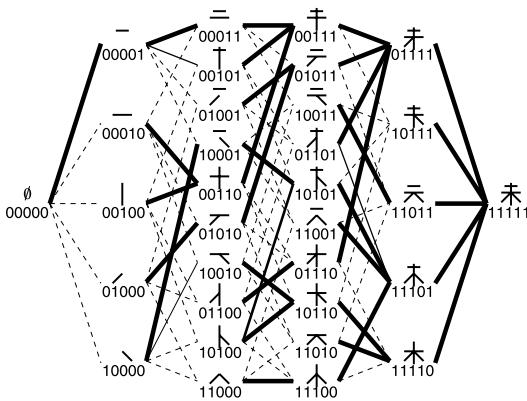


図 3 “未”の筆順変動の学習結果
Fig. 3 Trained stroke order variation of “未”.

うに 30 パターン以上存在) のうち 2/3 を用いてストローク HMM を学習した。残る 1/3 は評価用パターンとして 4.4 の認識実験の評価用パターンに含めた。ストローク HMM としては 3.1 で述べた手法を用いた。

4.3 筆順変動モデル学習結果

図 3 にカテゴリー “未”について、学習された状態遷移確率 $p_{m,n}$ を可視化したものを示す。実線は $p_{m,n} \geq 10^{-3}$ となった状態遷移であり、 $p_{m,n}$ が大きくなるほど太く示している。逆に点線は $p_{m,n} < 10^{-3}$ となった状態遷移を示す。

同図より、キューブグラフ上縁の経路に相当する正規筆順については、大きなゆう度をもつことが分かる。実際、“未”的正規画数パターン 343 個の筆順を調査した結果、1-3-2-4-5 の筆順で書かれた 1 パターンを除いてすべて正規筆順で書かれており、 $p_{m,n}$ は実際の筆順情報を反映していることが分かる。なお、1-3-2-4-5 という表現は正規筆順の画順番に対して第 1, 3, 2, 4, 5 画の筆順 (“未”では → | → — → / → \ の順) を意味する。

図 4 に、より筆順変動の種類の多かったカテゴリー “用”的学習結果を示す。“用”的正規画数パターン 1496 個の筆順を調査した結果、正規筆順が 1315 個、1-2-5-3-4 の筆順が 178 個、4-2-1-3-5, 1-3-5-2-4, 2-3-1-4-5 の筆順が各 1 個であった。これら実際の筆順に対応する経路が学習によって獲得されているのが確認できる。(例えば、2 番目に多かった筆順 1-2-5-3-4 は、経路 00000 → 00001 → 00011 → 10011 → 10111 → 11111 に対応。)

なお、図 3、図 4において、主要な筆順以外に、高いゆう度をもった状態遷移が散在している。これらは

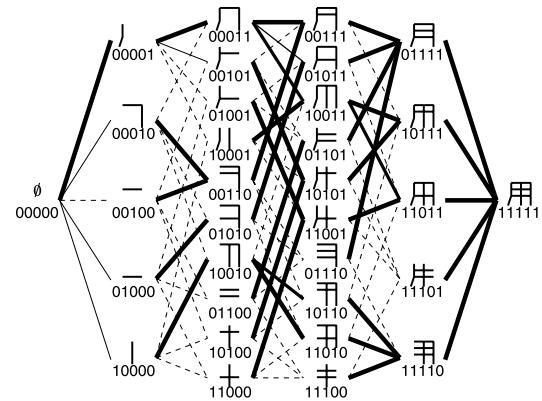


図 4 “用”的筆順変動の学習結果
Fig. 4 Trained stroke order variation of “用”.

表 2 本手法及び従来法の誤認識率(%)

Table 2 Error rates by the proposed method and the conventional method.

#strokes, K	5	10	15	20
conventional	closed	9.06	2.13	1.14
	open	9.55	2.26	1.26
proposed	closed	5.98	0.22	0.06
	open	6.47	0.70	0.54

同一文字内の画形状の類似性(例えば“用”的場合は 1 画と 5 画)と実際の筆順の組合せにより発生したものである。一見すると筆順変動モデルとして適切でないようにも見えるが、散在的すなわち主要な経路に組み込まれていない遷移なので、筆順情報としてほとんど機能しない。実際、後述する認識実験の考察結果からも、特に悪影響を及ぼさないことが判明している。

4.4 認識結果

4.4.1 認識率

表 2 に本手法による誤認識率を示す。これらの誤認識率は 3 重交差確認法によった。すなわち、同一筆者データが極力分割されないように全データを 3 分割し、そのうち二つを確率 $p_{m,n}$ 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。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。

同表中の従来法とは、筆順変動モデルを用いない 2.2 の手法のことであり、具体的には本手法において $p_{m,n}$ をすべて 1.0 にした場合に相当する。したがってこの従来法と比較することで、筆順変動を学習することの効果を検証できる。ストローク HMM の学習用パターンと画間ゆう度 $q_{m,n}$ 、及び評価用パターンは、

表 3 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つたパターン数 (failed). 括弧内はその全データに占める割合 (%)

Table 3 The numbers of improved samples and degraded samples.

#strokes, K	5	10	15	20	
improved	closed open	1,821 (3.19) 981 (3.44)	674 (1.91) 349 (1.98)	72 (1.08) 34 (1.02)	2 (0.2) 1 (0.2)
	closed open	63 (0.11) 103 (0.36)	3 (0.01) 75 (0.43)	0 (0.00) 10 (0.30)	0 (0.0) 0 (0.0)
failed	closed open	3,347 (5.87) 1,742 (6.11)	76 (0.22) 48 (0.27)	4 (0.06) 8 (0.24)	0 (0.0) 0 (0.0)

本手法のそれらと同一とした。

表 2 より, 筆順変動モデルを利用して誤認識率を大幅に低減できることが分かる。評価用パターン (open) について誤認識が低減された割合を計算すると, 5 画では誤認識が 32%, 10 画では 69%, 15 画では 57% 減っている。20 画については,すべての評価用パターンが正しく認識された。なお,他の 10 画以上の文字に対して 5 画における誤認識の低減効果が高いことに関しては 4.5 にて後述する。

表 3 に従来法を基準とした本手法による改善パターン数及び改悪パターン数を示す。いずれの画数 K においても, 改悪パターン数は少ない。すなわち, 前述の誤認識率低減は, 従来法で誤認識されていたものがそのまま正解に転じたことによるものであり, 筆順変動モデル導入による副作用は少ないと分かる。

4.4.2 改善例

表 4 に, 5 画における従来法に対する本手法の主要な改善結果を示す。例えば, 正規筆順で書かれた“本”的うち, 従来法では誤った筆順同定の結果“未”と誤認識されていた 151 パターンが, 本手法では正しく筆順同定されたことで正解に転じたことを示している。また, 同表では従来法では誤認識した“未”モデルでの筆順同定結果には, 1-3-4-2-5 などが多かったことも併せて示している。これらの筆順同定結果は, 4.3 で述べたように“未”的学習パターン中には存在せず, 図 3 に示す“未”的筆順変動モデルにおいても極めてまれな筆順である。本手法で獲得した統計的な筆順情報によって誤認識カテゴリーの過度な画対応が抑制された結果, 正しいカテゴリーに認識されたことが分かる。

次に, 従来法では“用”と誤認識されていた“田”106 パターンが本手法によって正解となった事例に着目する。図 5 にここで言及される各事例の物理ストロークを示す。表 4 の (*1) は, 非正規筆順で書かれた“田”

表 4 従来法に対する本手法の改善例 (5 画)

Table 4 Major improved samples.

#samples	input cat.	order	recognition result	
			conventional cat.	proposed order
151	本	1-2-3-4-5	未	1-3-4-2-5, 1-3-4-5-2, 2-3-4-1-5, 2-3-4-5-1
139	本	1-2-3-4-5	未	1-3-4-2-5, 1-3-4-5-2, 2-3-4-5-1, 2-3-4-1-5
107	目	1-2-3-4-5	司	3-1-2-4-5, 3-1-4-2-5, 3-1-4-5-2
106	田	1-2-4-3-5 1-2-3-4-5 1-2-4-5-3 1-3-5-2-4	用 (*1) (*2)	1-2-4-5-3, 1-2-3-5-4 1-2-5-4-3, 1-2-5-3-4 1-2-4-3-5 1-5-4-2-3
103	市	1-2-3-4-5	布	2-1-3-4-5, 2-1-5-4-3
				田 1-2-4-3-5 用 1-2-4-5-3 用 1-2-3-5-4 田 1-2-3-4-5 用 1-2-5-4-3 用 1-2-5-3-4 田 1-2-4-5-3 用 1-2-5-3-4 用 1-2-5-2-4



図 5 改善例 (*1), (*2) の物理ストローク対応
Fig. 5 Physical stroke order of improved samples.

に対して“用”モデルの過度な画対応が抑制されたことで正答に転じた例である。一方, 同表の (*2) は, 正規筆順で書かれた入力パターン“田”的改善例である。4.3 で見たように, カテゴリー“用”には“田”的正規筆順に対して形状的に類似した非正規筆順 (1-2-5-3-4) が学習パターン中に有意数存在するので“用”モデルにとって過度な画対応ではない。したがって筆順変動モデルを用いてもこれら入力パターンは“用”に誤認識される可能性はある。しかし, 筆順変動モデル適用によるゆう度低下は, “田”的正規筆順に対して“用”的非正規筆順の方が大きい。その結果, “田”的正規筆順モデルの方が“用”的非正規筆順モデルよりもゆう度的に上回るため, 正答に転じている。

以上から, 本手法では筆順変動モデルに学習された実際の筆順情報に基づいて現実には起こらない筆順や, 起こり得るが頻度の低い筆順のゆう度を下げることによって認識率を改善している。なお, 表 4 から, 筆順変動の多い“田”においても, 入力パターンの筆順が本手法により正しく同定できていることが分かる。

表 5 従来法に対する本手法の改悪例(5画)

Table 5 Major degraded samples.

samples	input	recognition result	
		conventional	proposed
		cat. order	cat. order
15	用 1-2-3-4-5	用 1-2-4-3-5, 1-2-3-4-5	申 1-2-3-4-5
12	用 1-2-5-3-4	用 1-2-5-4-3, 1-2-5-3-4	田 1-2-3-4-5
10	生 1-2-4-3-5 1-2-3-4-5	生 1-2-4-3-5, 3-2-4-1-5 3-2-1-4-5	主 1-2-4-3-5 1-2-3-4-5
10	四 1-2-3-4-5	四 1-2-4-3-5	皿 1-2-3-4-5
9	右 2-1-3-4-5	右 2-1-3-4-5, 5-1-3-4-2	石 1-2-3-4-5

4.4.3 改悪例

4.4.1 でも述べたように、改善パターンに比べ改悪パターンの数は非常に少なく、1/10程度である。表5に、本手法による5画文字での主な改悪パターンを示す。誤認識に転じた要因は様々あるが、それらの中でも一つの特徴的な例を同表2番目の“用”の“田”への誤認識に見ることができる。これは、非正規筆順のゆう度に対して正規筆順のゆう度が相対的に高くなるという、4.4.2の“本”的改善例でも見た本手法の全般的な傾向が逆効果となった例である。すなわちこの入力パターン“用”は非正規筆順(1-2-5-3-4)で書かれたものであり、したがってそのゆう度は正規筆順の場合に比べて多少低くなった。一方、不正解となった文字カテゴリー“田”との画対応付けの結果は正規筆順(1-2-3-4-5)であり、そのゆう度は非正規筆順に対して相対的に高くなっている。この結果“田”とのゆう度が“用”とのゆう度に勝り、誤認識に転じた。こうした改悪パターンはあまり多くないものの、筆順変動の統計的モデルを使う場合に必ず起り得る問題といえよう。

4.5 ストロークHMMの追加学習

筆順変動モデルには、不自然な画対応による誤認識の抑制という上述の効果に加え、利用可能な学習パターンの拡大によるストロークHMMの高精度化という二次的効果をもたらす。すなわち、筆順変動モデルにより良好な画対応が得られるようになれば、言い換えると、筆順変動を伴う学習パターンについてもその第 k 画が正規筆順の第 k' 画であることが分かれれば、前者を第 k' 画用ストロークHMMの学習に利用できることになる。結果的に、ストロークHMM用学習パターンとして、少数の正規筆順パターンだけでなく、任意筆順のパターンを利用できることになる。

表 6 追加学習を導入した場合の誤認識率(%)

Table 6 Error rate by the proposed method with incremental training.

#strokes, K		5	10	15	20
proposed with	closed	3.38	0.10	0.00	0.0
inc. learn.	open	3.94	0.60	0.45	0.0

このことを踏まえ、正規筆順でないパターンをもストロークHMMの学習に利用すべく、 $p_{m,n}$ の学習と同時にストロークHMMを追加学習する方式について実験的に検討した。具体的な追加学習の方式は以下のとおりである。まず3.2での $p_{m,n}$ 学習の際に、各学習パターンについてキューブHMMグラフ上で最大ゆう度を与える経路をViterbiアルゴリズムによって求める。次に、この経路により学習パターンの筆順が正しく表現されているとして、学習パターンの筆順を正規筆順に変換し、対応するストロークHMMの再学習に適用する。各ストロークHMMの追加学習が収束した後に、 $p_{m,n}$ の学習を継続する。これによりストロークHMMの学習パターンが増加し、ストロークHMMの精度が向上すると考えられる。

表6に追加学習を導入した場合の誤認識率を示す。表2と比べると、追加学習の導入により、更に誤認識率を低減できていることが分かる。

より細かく見ると、低画数文字ほど誤認識率低減効果が高いことが分かる。すなわち、5画文字について9.55%→6.47%と32%であった誤認識低減率(表2参照)が、追加学習導入によって59%となり、他の画数における誤認識低減率と比肩するに至った。これは、追加学習の導入によってストロークHMM用の学習パターンを増やすことができ、変動の大きな低画数文字のストローク形状を追加学習導入前よりも適切に表現できるようになったためと考えられる。

5. むすび

キューブサーチのもつ高い画対応能力を抑制し適正化するためにノード間遷移確率を導入したキューブHMMを提案した。学習パターンから実際の筆順変動を学習するために、従来のトレリスとは異なるキューブHMM上でありながら従来のEMアルゴリズムの枠内で機能するノード間遷移確率の再推定手法を提案した。本手法の学習実験により実際の筆順の獲得が観察され、認識率の向上によってその有効性を確認した。更にキューブHMMの有する筆順判定機能を用いて学習データの筆順を正規筆順に変換する追加学習手法を

提案し、認識実験によりその有効性を確認した。

今後の課題は、ストローク HMM の表現能力に関する検討が挙げられる。これはキューブ HMM で獲得された筆順情報や認識結果に、ストローク HMM の表現能力に起因した問題が見られたためである。また、ストローク HMM を内包したキューブ HMM 全体が一つの HMM であることを生かして、一括同時学習に関する検討も挙げられる。4.5 でのストローク HMM の追加学習が効果を挙げたように、それを更に一般化した形式となる同時学習にも一定の効果が期待される。

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] C.-L. Liu, S. Jaeger, and M. Nakagawa, "Online recognition of Chinese characters: The state-of-the-art," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol.26, no.2, pp.198–213, Feb. 2004.
- [2] 東代誠仁, 徳野順子, 朱 碧蘭, 中川正樹, 田中 宏, "ペンインタフェースのための手書き文字認識技術," *ヒューマンインタフェース学会誌*, vol.8, no.3, pp.161–166, 2006.
- [3] 追江博昭, 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.
- [4] 懇 重弼, 追江博昭, "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.2300–239, Feb. 1999.
- [5] 小林 充, 真崎晋哉, 宮本 修, 中川洋一, 小宮義光, 松本 隆, "オンライン手書き文字認識アルゴリズム RAV (Reparameterized Angle Variations)," *情処学論*, vol.41, no.9, pp.2536–2544, 2000.
- [6] 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*, PRMU 96-211, 1997.
- [7] 中川正樹, 池田裕治, 相澤 正, 茂田千冬, 高橋延匠, "構造解析的オンライン手書き文字認識における字体表現辞書の拡張と統計的相違度の評価," *信学論(D)*, vol.J71-D, no.5, pp.800–811, May 1988.
- [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] 横田登志美, 葛貴壯四郎, 郡司圭子, 桂 晃洋, 浜田長晴, 福永 泰, "筆順画数同時フリーを実現する楔文字モデル

によるオンライン日本文字認識方式," *情処学論*, vol.44, no.3, pp.980–990, 2003.

- [12] M. Hamanaka, K. Yamada, and J. Tsukumo, "Online Japanese character recognition based on flexible pattern matching method using normalization-cooperative feature extraction," *IEICE Trans. Inf. & Syst.*, vol.E77-D, no.7, pp.825–831, July 1994.
- [13] M. Okamoto and K. Yamamoto, "On-line handwriting character recognition using direction-change features that consider imaginary strokes," *Pattern Recognit.*, vol.32, no.7, pp.1115–1128, July 1999.
- [14] 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.
- [15] 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. 8th Int. Conf. Doc. Anal. Recog.*, vol.1 of 2, pp.26–30, 2005.

(平成 19 年 7 月 30 日受付, 11 月 29 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准教授。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・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)日本発明協会・内閣総理大臣発明賞、九大嘉村記念賞各受賞。