

Hidden Markov Model に基づくオンライン手書き文字認識システムの研究開発

1P-6

伊藤 等 中川 正樹
東京農工大学工学部電子情報工学科

1.はじめに

本研究では、オンラインで入力された手書き文字パターンの特徴ベクトルをラベル列に量子化し、従来から音声認識で広く応用されている HMM を用いて、揺らぎの大きい文字パターンについても認識できる文字認識システムについて提案する。学習や認識の基本的な手法は音声認識の分野で培われてきたものと同様であるが、我々は量子化する特徴の取り方に重点を置いている。また、我々の研究室で数年に渡って収集してきたオンライン手書き文字パターンデータベースを基に、大量のデータを統計的に学習することで、崩れ字や続け字に対処させている。今回は、平仮名 46 字種を対象としておこなった認識実験について報告する。

2.認識アルゴリズム

本認識システムの大まかな流れを図1に示す。

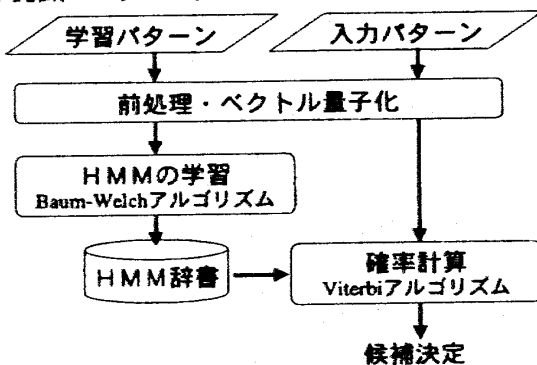


図1 認識アルゴリズムの流れ

2.1.前処理

前処理では、線形正規化と特徴点抽出をおこなう。また位置の特徴を扱う際には、入力パターンの密度を一定にしパターン間での位置のばらつきを抑えるため、非線形正規化をおこなう。

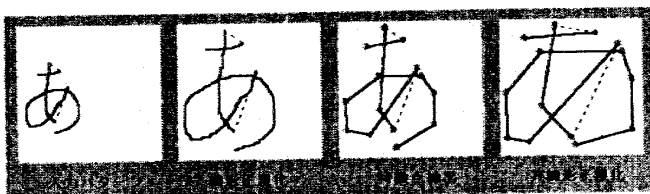


図2 前処理の例

2.2.ベクトル量子化

この処理はデータ量を削減するとともに、HMMへ適用しやすいラベル列に置換する役割をもつ。ここでは異なった2種類の特徴に対する量子化の手法について述べる。

2.2.1.方向特徴の量子化

前処理を施したパターンの特徴点から次の特徴点へのベクトルに着目し、その方向情報を量子化する。今回は8方向のラベルに、裏ストローク(pen up時のベクトル)に対する同じ8方向のラベルを加え、計16種類のラベルによって、文字パターンをラベル列に置換する。

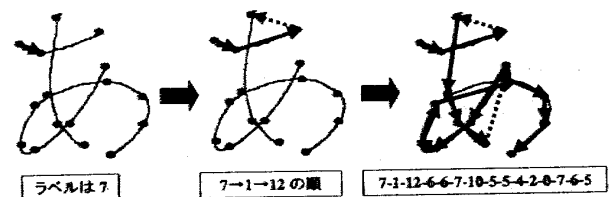


図3 方向特徴の量子化の例

2.2.2.位置特徴の量子化

128×128に正規化された領域を4×4=16個の小領域に分割し、それぞれにラベルを付ける。入力パターンの特徴点がどの小領域に属するか順番にラベル付けをして、量子化を行う。

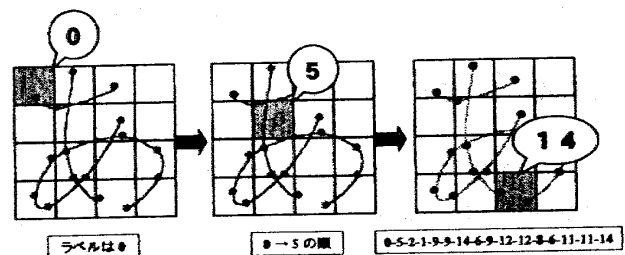


図4 位置特徴の量子化の例

2.3.モデルの学習

モデルの状態数は、少なすぎると文字の構造を十分に表現できない。多すぎると学習時間やメモリ使用量が大きくかかるといったデメリットがある。簡単な予備実験を行った結果から、状態数は平仮名46字種に対して最も適している12に固定する。状態遷移には条件をつけ、自己ループと次の状態への遷移、これとnull遷移の3つだけが有効とする。決定したモデルに対して、Baum-Welchのパラメータ推定アルゴ

リズムにより統計的学習を行う。今回は1カテゴリに対して1モデルとする。揺らぎに強いパラメータとなるように、大量のデータを収束するまで学習させる。

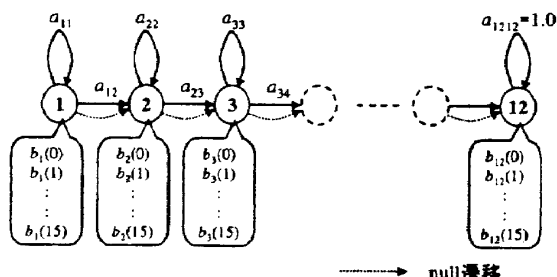


図5 12 状態 16 出力の left-to-right モデル

2.4. 確率計算による認識

未知入力文字パターンは、学習パターンと同じ前処理・ベクトル量子化を行い、ラベル列としてモデルに入力する。この入力に対して、HMM 辞書内の 46 種類の各モデルごとに尤度計算を行い、尤度の大きいものから候補とする。尤度計算は、Viterbi アルゴリズムによって行う。

3. 実験

3.1. 実験で使ったデータ

今回の実験では、学習には nakayoshi_t-97-02 の nky0001~nky0070 の 70 セットの平仮名 46 字種を使用した。以前のものは、各字種に対して文章中に出てくる最初の 1 文字だけを学習対象としていたが、今回は学習オーダーを増やすために、同じ筆記者の 1 字種に対して複数のパターンについて学習させた。評価は kuchibue_d-96-02 の mdb0001~mdb0020 の 20 セットで、1 セットに含まれる平仮名の総数は 3,817 文字より、合計 76,340 文字について認識率を導き出した。

3.2. 異なる特徴による実験

方向特徴だけで計算した認識結果を(a)、位置特徴のみで計算した認識結果を(b)、また方向特徴の計算結果と位置特徴の計算結果に適切に設定した重み付けを行い、それらを加えた評価値を用いて実験した結果を(c)に示す。方向特徴のみでの認識率は最高で 94.7% (mdb0005)、最低で 79.3% (mdb0004) と、データ間(異なる筆記者)でかなりひらきがあった。位置特徴の認識率は方向特徴に比べるとやや劣っている。

位置と方向を組み合わせた HMM での認識率は最高で 96.9% (mdb0009)、最低で 85.3% (mdb0011) と、データ間でのひらきはかなり減少している。また、方向、位置単体による認識より、かなり良い結果が出ている。このことから、2 つの特徴による計算結果を組み合わせることは有効であることがわかる。

表 異なる特徴による実験結果

	認識率	2 位までの 累積認識率	3 位までの 累積認識率
(a)	87.0%	95.9%	97.4%
(b)	81.8%	91.6%	94.7%
(c)	91.6%	97.5%	98.7%

(a) 方向特徴のみ (b) 位置特徴のみ (c) 方向+位置

4. 考察と今後の課題

当研究室のデータベースは、筆記者に特別に制限を付けないため、かなり雑に書かれた文字が多い。しかし、図6のような文字が多いにもかかわらず、3.2 節の実験ではかなり良い成果が得られた。

今後の課題として、

- ・ 認識対象を JIS 第一水準の漢字やアルファベット、カタカナ、記号などに広げる
 - ・ 最適なモデルの構築
 - ・ 学習アルゴリズム高速化への検討
 - ・ 辞書(メモリ使用量)の削減
- などが挙げられる。

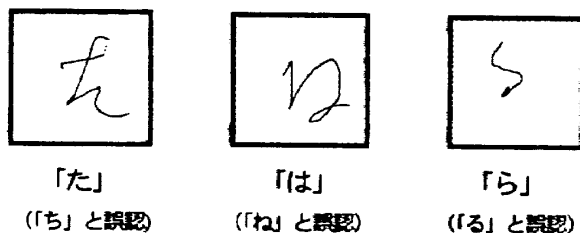


図6 3 位候補までにあがらなかった文字の例

謝辞

本研究は一部、情報処理振興事業協会による創造的ソフトウェア育成事業の補助による。ここに深く感謝する。

参考文献

- [1] G.Rigoll, A.Kosmala, J.Rottland, Ch.Neukirchen: "A Comparison Between Continuous and Discrete Density Hidden Markov Models for Cursive Handwriting Recognition", IEEE Proceedings of ICPR '96, pp.205-209(1996)
- [2] 高橋堅一郎, 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, pp.143-150(1997)
- [3] 中川正樹, 東山孝生, 山中由紀子, 澤田伸一, レー・バン・トゥー, 秋山勝彦: "文章形式字体制限なしオンライン手書き文字パターンの収集と利用", 信学技報, PRU 95-110, pp.43-48(1995)
- [4] 中川聖一: "確率モデルによる音声認識", 電子情報通信学会(1988)