

マルコフ決定過程による手書き文字認識法

4 E - 1

松井裕文 塩山忠義

京都工芸繊維大学

1. はじめに

本稿では、マルコフ決定過程¹⁾による手書き文字の認識法を提案する。本認識法では、文字領域を16の部分に分割し、各部分のパターンに基づいて、認識誤りによる期待損失を最小にするように、16段階の決定過程からカテゴリー決定がなされる。この認識過程は、手書き変動の詳細情報である各部分のパターンの分布と部分間のパターン推移の分布を辞書に用いているために、変動の大きな文字に対しても適用できるものと思われる。以下に、その認識アルゴリズムと認識実験について述べる。

2. 認識方法

2.1 学習過程 手書き文字の学習用サンプルは、16×16の領域に分割された一定の枠内に、大きさに関して規格化されている。この領域内で、文字線と背景は濃度値がそれぞれ、1, 0となるように二値化されている。 r 番目のカテゴリーの i 番目の学習用サンプルの領域 (x, y) での濃度値を、 $f_{r,i}(x, y)$; $x, y = 1, \dots, 16, r = 1, \dots, R, i = 1, \dots, L$ で表す。ここで、 R はカテゴリーの数、 L は学習用サンプルの数である。 x, y は、それぞれ列と行を示している。

領域 (x_j, y) ; $j = 1, \dots, m$ が r 番目のカテゴリーの文字線に属す確率 $p^r(x_1, \dots, x_m; y)$; $m = 1, \dots, 16$ は、学習サンプルから次のように求められる。

$$p^r(x_1, \dots, x_m; y) = \frac{1}{L} \sum_{i=1}^L \prod_{j=1}^m f_{r,i}(x_j, y) \prod_{j' \in \{1, \dots, m\}} (1 - f_{r,i}(x_{j'}, y)) / C \quad (1)$$

ここで、 C は規格化定数であり、二値画像の場合には $C = 1$ となる。次に遷移確率を得るために、領域 (x_j, y) ; $j = 1, \dots, m$ かつ領域 $(x_{j'}, y+1)$; $j' = 1, \dots, m'$ が r 番目のカテゴリーの文字線に属す確率として、同時確率 $p^r(x_1', \dots, x_{m'}'; y+1, x_1, \dots, x_m; y)$; $m, m' = 1, \dots, 16$ を定義する。この同時確率は次式のように求められる。

$$p^r(x_1', \dots, x_{m'}'; y+1, x_1, \dots, x_m; y) = \frac{1}{L} \sum_{i=1}^L \prod_{j'=1}^{m'} f_{r,i}(x_{j'}, y+1) \prod_{k' \in \{1, \dots, m'\}} (1 - f_{r,i}(x_{k'}, y+1)) \times \prod_{j=1}^m f_{r,i}(x_j, y) \prod_{k \in \{1, \dots, m\}} (1 - f_{r,i}(x_k, y)) / D \quad (2)$$

ここで、 D は規格化定数であり、今の場合 $D = 1$ である。 y 行目の領域 (x_j, y) ; $j = 1, \dots, m$ が r 番目のカテゴリーの文字線に属しているときに、 $y+1$ 行目の領域 $(x_{j'}, y+1)$; $j' = 1, \dots, m'$ が r 番目のカテゴリーの文字線に属す遷移確率 $p^r(x_1', \dots, x_{m'}'; y+1 | x_1, \dots, x_m; y)$ は、式(1)、式(2)から次のように求められる。

$$p^r(x_1', \dots, x_{m'}'; y+1 | x_1, \dots, x_m; y) = p^r(x_1', \dots, x_{m'}'; y+1, x_1, \dots, x_m; y) / p^r(x_1, \dots, x_m; y) \quad (3)$$

但し、 $p^r(x_1, \dots, x_m; y) \neq 0$ とする。

全てのカテゴリー($r = 1, \dots, R$)に対する式(1)と式(3)の結果より、辞書を構成する。

2.2 認識過程 手書き文字の未知サンプルは、学習サンプルと同様に、大きさに関して規格化されている。領域 (x, y) での濃度値を $f(x, y)$; $x, y = 1, \dots, 16$ とすると、領域 (x_j, y) ; $j = 1, \dots, m$ が未知サンプルの文字線に属している確率 $p(x_1, \dots, x_m; y)$; $m = 1, \dots, 16$ は次のように求められる。

$$p(x_1, \dots, x_m; y) = \prod_{j=1}^m f(x_j, y) \prod_{j' \in \{1, \dots, m\}} (1 - f(x_{j'}, y)) \quad (4)$$

y 行目の状態 $X = (x_1, \dots, x_m)$; $m = 1, \dots, 16$ を、 $j = 1, \dots, m$ に対して $f(x_j, y) = 1$ かつ $j \in \{1, \dots, m\}$ に対して $f(x_j, y) = 0$ の状態として、定義する。このときに決定を k としたときの直接期待損失 $R_k(X; y)$ は、次式で与えられるものとする。

$$R_k(X; y) = |p(X; y) - p^k(X; y)| \quad (5)$$

この損失は、領域 (x_j, y) ; $j = 1, \dots, m$ が未知サンプルの文字線に属しているときに、 k 番

目のカテゴリーであると認識した場合の誤りに起因する。(16-n+1)行目の状態Xから出発して、最後の16行目までの残りn段階での最小期待損失 $V_n(X)$ は、最適性の原理から次式で与えられる。

$$V_n(X) = \min_{k \in A_n(X)} V_n^k(X); n=1, \dots, 16 \tag{6}$$

ここで、

$$V_n^k(X) = R_k(X; 16-n+1) + \sum_{X'} p^k(X'; 16-n+2 | X; 16-n+1) V_{n-1}(X')$$

$$X \equiv (x_1, \dots, x_m); m=1, \dots, 16$$

$$X' \equiv (x_1', \dots, x_{m'}'); m'=1, \dots, 16$$

$$A_n(X) \equiv \{k | k \in \{1, \dots, R\}, p^k(X; 16-n+1) > 0\}$$

である。但し、

$$V_0(X) \equiv 0$$

とする。

学習段階で得られた辞書を用いて、段階n; n=1, ..., 16において $V_n(X)$ が計算され、同時に決定 k_n が得られる。未知サンプルは、最大のヒストグラムをもつ決定に対応するカテゴリー(文字)として認識される。

3. 認識実験

本実験では、認識が困難であると考えられる²⁾グループA(a, e, s, g, 8)と、比較のためのグループB(f, k, o, u, z)の2グループについて認識実験を行った。その結果を図1、図2に示す。認識率は、グループAでは学習サンプル数L=30の場合に、学習サンプルに対し99%、未知サンプルに対し92%、またグループBではL=30の場合、学習サンプルに対し100%、未知サンプルに対し99%であった。図より、学習サンプル数を増加するほど認識率が上昇している傾向が読み取れる。

4. おわりに

本アルゴリズムを用いて手書き文字認識を行った結果、学習サンプル数が非常に少ないと考えられる30の場合で、未知サンプルに対する認識率は、グループAでは92%、グループBでは99%であった。今後、学習サンプル数を増加することにより認識率が向上するものと思われる。

参考文献

- 1) R. A. ハワード, 関根・羽鳥・森共訳: ダイナミックプログラミングとマルコフ過程, 培風館, 1973.
- 2) H. S. Baird, S. Kahan, T. Pavlidis: Components of An Omnifont Page Reader, Proc. 8th Int. Conf. Pattern Recognition, pp. 344-348, 1986.

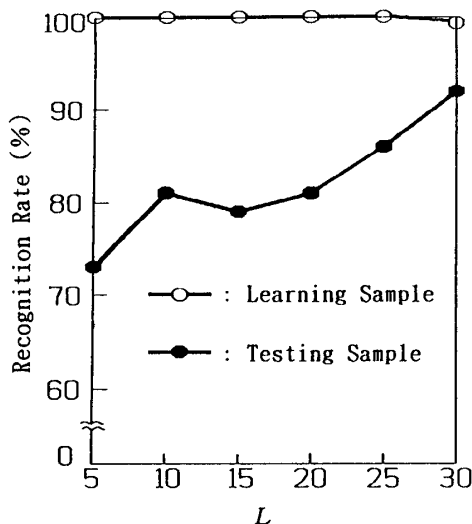


図1. グループAでの認識率

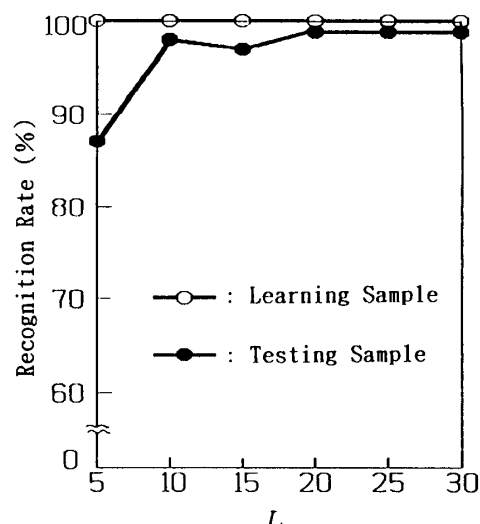


図2. グループBでの認識率