

射影方向線素による多種字体印刷漢字認識

正 員 北島 宗雄[†]

Recognition of Multi-Font Kanji Characters Using Projected Direction Segments Feature

Muneo KITAJIMA[†], *Regular Member*

あらまし 本論文では、視覚障害者用読書器における文字認識手法の開発を目的として、明朝体、ゴシック体、および教科書体を対象とする多種字体印刷漢字認識のための詳細分類法と分類実験について述べる。詳細分類は、文字パターンから安定な方向を持つ線素を抽出し、さらに各方向に垂直な軸上に射影して作成される射影線素分布の動的な重ね合わせによって行われる。射影線素は、ほぼ忠実に原文字パターンの方向を持ったストロークの位置と長さを表わして、字体の違いによる字形の影響を受けにくい特徴となっており、教科書体を含めることによって生じる著しい字形変形を吸収することが可能である。また、分類実験は、文字の輪郭の大まかな形状を表わす粗外形パターンコードと、文字の縦長、横長を表わす縦横比コードを用いた大分類に、本手法の詳細分類を組み合わせた2段階分類方式を用いて、教育漢字881を対象に行われ、その結果、認識率99.16%、第2位分類率99.89%が得られ、本手法の多種字体印刷漢字認識における有効性が確認された。

1. ま え が き

昭和55年度厚生省統計によると、わが国における視覚障害者は約34万人である。このうち約50%の人は視覚による読書は不可能で、点字や録音テープによるか、または晴眼者による朗読で読書を代行している。しかしながら、点訳図書による読書には訓練が必要であり、また入手できる点訳本や録音テープの種類が少ないなど問題があり、視覚障害者用の安価で個人使用のできる読書器の開発が望まれている。

読書器が実際に使われる場として小学校程度の教育の場を想定した場合、明朝体、ゴシック体、教科書体を対象とした多種字体印刷文字の認識を行う必要がある。また、読書器のコスト低減を図るためには、辞書容量の小さい認識手法をとらなければならない。

さて、多種字体印刷文字認識の試みの一つとして、目黒らはベリフェラル特徴とメッシュ特徴とを併用した分類に、差分ストロークによる識別を加えたマルチフォント漢字認識シミュレーションシステムを構成し⁽¹⁾、明朝体、ゴシック体のタイプ印刷漢字を対象にその有効性を確認し、実験装置作成にまで至っている⁽²⁾。こ

の手法が教科書体に適用可能なことは十分予想されるものの、辞書容量は300KB(1929文字)と大きく、上記読書器実現のためには、より小さな辞書でしかも字形変形に強い手法が必要となる。

筆者は、先の報告で、視覚障害者用読書器実験システム⁽³⁾における多種字体印刷文字認識手法として、文字パターンを上下左右の4方向から眺めたときに文字線によって全ての方向から遮られる部分を黒く塗りつぶして作成される外形パターンを利用した大分類と、平均線数特徴を用いた詳細分類から成る2段階分類方法を提案したが⁽⁴⁾、本論文では、詳細分類法としてあらたに射影線素を用いた方法を提案し、その能力を明らかにする。平均線数特徴が、文字パターンを4×4メッシュの領域に分割したときに各メッシュ領域に含まれる縦線および横線の本数を特徴としているのに対し、射影線素照合法では、漢字パターンが近似的に、方向を持った線素の集合によって構成されているという考えの基に、文字パターンから水平、垂直、対角、逆対角の各方向を持つ線素を抽出し、これらをそれぞれの方向に垂直な軸上に射影して作成される射影線素分布の重ね合わせによって照合を行う。このような文字線素の方向に着目した手法は、手書き漢字認識にしばしば利用され、有効性が確認されており^{(5)~(7)}、多種字体印刷文字認識においても有効であると期待される。

[†]製品科学研究所人間工学部、茨城県
Human Factors Engineering Division, Industrial Products
Research Institute, Ibaraki-ken, 305 Japan

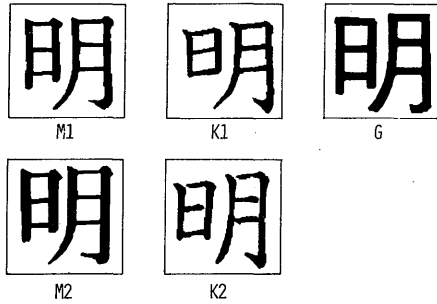


図1 印刷文字データの例

Fig.1- Examples of Kanji characters printed in various type fonts.

また、照合に際しては、射影線素分布間の対応が動的になされるので、平均線数特徴に見られるようなメッシュ化による弊害を取り去ることができると考えられる。

射影線素抽出法と辞書の作成について2.に、また、明朝体、ゴシック体、および教科書体を対象とする分類実験と結果を3.に述べる。

2 射影線素分布の抽出と辞書の作成

文字印刷には種々の書体が用いられているが、一般の書物に最も広く使われているのは明朝体とゴシック体である。一方、小学校用教科書印刷には主に教科書体とゴシック体が使われており、明朝体はほとんど使用されていない。さらに、書体が同じであってもそれに含まれる多くの字体が存在する。これらの字体の中から代表的なものとして、中明朝体(M1)、太明朝体(M2)、教科書体2字体(K1, K2)、および中ゴシック体(G)を選び認識実験に使用しているが、その一例を図1に示す。図から、字体が異なることによって構成ストロークの形状や配置が大きく変化していること、特に手書き文字に近い教科書体と他の字体との差が顕著なことがわかる。

本章では、図1に見られるような字体間の差を吸収しつつ、かつその文字をよく表わす特徴として一定の方向を持つ線素を取りあげ、その抽出法とそれを垂直な軸上に射影して得られる射影線素間の距離の定義を行う。また、各文字カテゴリーを代表する射影線素分布辞書の作成について述べる。

なお、ここで用いられる文字パターンは、写植印刷文字(50級)をテレビカメラにより128×128メッシュの分解能で撮影し2値化したものである。

2.1 方向コード化

抽出する方向の角度は、水平方向を 0° として 0° 、 45° 、 90° 、 135° (水平、対角、垂直、逆対角)の4方向である。これらをそれぞれ θ_k ($k=0,1,2,3$)とし、またそれらの反対方向を θ_{k+4} とする。いま、入力文字パターン上の一つの黒点の座標を (x_B, y_B) とし、この点を始点として θ_k 方向に走査して行ったときにはじめて現われる白点までの黒画素数を $d_{\theta_k}(x_B, y_B)$ とする。また、この点に対し θ_k 方向のランレングスは、

$$l_{\theta_k}(x_B, y_B) = d_{\theta_k}(x_B, y_B) + d_{\theta_{k+4}}(x_B, y_B) - 1 \quad (k=0,1,2,3)$$

と表わされる。入力文字パターン上の各黒点について4方向のランレングスを求め、それらの中で最大となるランの方向 m に対し、条件

$$\begin{cases} l_{\theta_m}(x_B, y_B) > 2 \cdot l_{\theta_{m_\perp}}(x_B, y_B) \\ d_{\theta_m}(x_B, y_B) - d_{\theta_{m_\perp+4}}(x_B, y_B) = 0 \text{ or } 1 \end{cases}$$

ただし、 $m = \text{mod}(m+2, 4)$

を同時に満たす点を特徴点として抽出し、 m を方向コードとして与える。ここで、 m_\perp は m に垂直な方向である。これらの条件を満たす点は、その点において m 方向と m_\perp 方向のランレングスの比が2より大きく、かつ m_\perp 方向の中央となる点であり、方向が安定した黒領域を代表する点である。図2に抽出された方向コードパターンの一例を示す。

2.2 射影線素分布

方向コードパターンは単に安定な方向を持つ点の集合である。これらの点から文字の特徴、例えばストロークの存在する位置と長さやストローク間の構造的関係を抽出し認識に利用することが考えられるが、これらによる認識は一般に複雑となる。ここでは、できるだけ単純で見通しの良い方法で認識を行うために、2次元パターンである方向コードパターンをそのまま使わずに、各方向に垂直な軸上に射影して得られる1次元パターン(射影線素分布)を作成し、ストロークの射影軸上における位置と長さを求め認識に利用する。

射影された方向コードパターンの分布は、ストロークの凹凸や微かな傾きの影響を受けて多少の広がりをもっている。射影線素分布の中から一つのストロークに対応する部分を見つけるために、射影軸上で隣り合う射影線素間の距離が w 以下のものは、一つのストロークに属するとして統合処理を行う。すなわち、方向 k の射影線素分布を $S_k(x)$ ($k=0,2$ のとき $x=1, \dots, 128$ 、 $k=1,3$ のとき $x=1, \dots, 255$)とすると、

$$\max(0, x_0 - w) < x < x_0 \text{ で } S_k(x) = 0 \quad \text{かつ } S_k(x_0) \neq 0$$

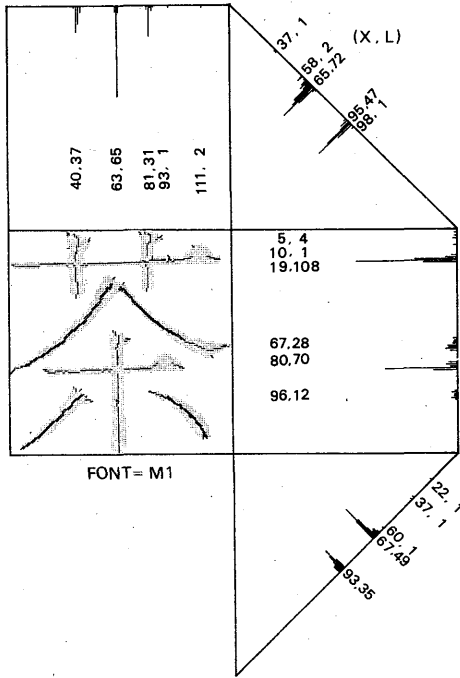


図2 方向コードパターンと射影線素
Fig.2- Direction pattern generation and extraction of projected direction segments.

を満たす点 x_0 から, $x_1 \geq x_0$ で

$$S_k(x_1) \neq 0$$

かつ $x_1 < x < \min(x_k^e + 1, x_1 + w)$ で $S_k(x) = 0$, ($x_0^e = x_2^e = 1.28$; $x_1^e = x_3^e = 2.55$) となる最初の点 x_1 までの領域を一つのストロークに対応する部分とする。また, この領域に対応するストロークの長さは, 区間 $[x_0, x_1]$ における射影線素分布の総和

$$l = n_k \sum_{x=x_0}^{x_1} S_k(x) \quad (1)$$

また, 位置はこの区間における射影線素分布の重心

$$\bar{x} = \frac{n_k \sum_{x=x_0}^{x_1} x S_k(x)}{l / n_k} \quad (2)$$

で与えられる。ここで, n_k は対角方向の射影線素を正規化するための定数で, $n_0 = n_2 = 1, n_1 = n_3 = 1/\sqrt{2}$ である。図2に射影線素分布 (projected segments) と, $w=3$ として統合処理をして得られた位置と長さを示す。本手法により, 各方向ともほぼ忠実にストロークの位置と長さを抽出することができる。今後, 統合処理をした結果を射影線素分布と呼ぶことにする。

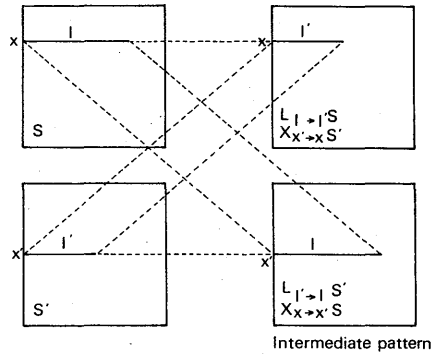


図3 中間パターン
Fig.3- Illustration of intermediate pattern.

また, 統合の幅 w は3とする。

2.3 距離

2本の射影線素 S, S' の位置と長さをそれぞれ (x, l), (x', l') とし, これら2本の射影線素間の距離が, 位置ずれ $\Delta x = |x - x'|$ と長さずれ $\Delta l = |l - l'|$ に適当な重み付けをした和で表わされるとすると, 距離 $d(S, S')$ は,

$$d(S, S') = \alpha \Delta l + (1 - \alpha) \Delta x \quad (3)$$

$(0 \leq \alpha \leq 1)$

となる。ここで α は, 位置ずれと長さずれをどのような割合で距離に反映させるかを決定するパラメーターである。 α は以下に示す手順で, 文字データを用いた統計的処理により決定される。

まず, 射影線素間の基本的な演算子として位置を x から x' に移動する演算子 $X_{x \rightarrow x'}$, 長さ l から l' に変える演算子 $L_{l \rightarrow l'}$ を導入する。対応付けられた射影線素対の i 番目のものを $S_i = (x_i, l_i), S'_i = (x'_i, l'_i)$ とすると, それらは演算子を用いて

$$S'_i = X_{x_i \rightarrow x'_i} L_{l_i \rightarrow l'_i} S_i = L_{l'_i \rightarrow l_i} X_{x_i \rightarrow x'_i} S_i$$

と表わされる。これを書きかえると

$X_{x'_i \rightarrow x_i} S'_i = L_{l_i \rightarrow l'_i} S_i$; $L_{l'_i \rightarrow l_i} S'_i = X_{x_i \rightarrow x'_i} S_i$ となり, これらの式の各項は S_i から S'_i に変形する際に現われる中間的なパターン S_i^{int} を表わしている。

S_i, S'_i の中間パターン (intermediate pattern) は図3に示すように, 一方の位置と他方の長さをもつ射影線素である。この中間パターンと, S_i, S'_i との距離が等しくなるように重み付けパラメーター α_i を定めることにすると, S_i と S_i^{int} との距離が $\alpha_i \Delta l_i, S'_i$ と S_i^{int} との距離が $(1 - \alpha_i) \Delta x_i$ となるので, $\Delta x_i + \Delta l_i \neq 0$ のとき,

$$\alpha_i = \Delta x_i / (\Delta x_i + \Delta l_i)$$

となる。対応付けられた射影線素対に対して得られる α_i の平均値

$$\bar{\alpha} = \frac{1}{N} \sum_{i=1}^N \alpha_i$$

は、重み付けパラメータ α を用いて対応付けを行ったときに、 S_i および S'_i とそれらの中間パターンとの距離をできるだけ等しくする値を与える。ここで、 N は射影線素対の総数である。距離の表式に表われる α とここで得られた $\bar{\alpha}$ は一致しなければならないが、一致しない場合には、新たに $\bar{\alpha}$ を(3)式の α と置きかえ $\alpha = \bar{\alpha}$ となるまでこの手続きを繰り返す。なお、射影線素対の対応付けについては次節で述べる。

中明朝体、教科書体、および中ゴシック体の3字体について α を求めた結果、 $\alpha_{M1, K1} = 0.309$, $\alpha_{K1, G} = 0.326$, $\alpha_{G, M1} = 0.329$ を得た。これらの値は概ね一致しており、長さずれと位置ずれを1:2の比で重み付けをすればよいことを示している。今後 $\alpha = 1/3$ として距離計算を行う。

2.4 射影線素分布の対応付け

正規化された2つの文字データから得られる方向 k の射影線素分布を

$$S_k = \{(x_i, l_i) | 1 \leq i \leq n, x_1 < x_2 < \dots < x_n\}$$

$$S'_k = \{(x'_j, l'_j) | 1 \leq j \leq n', x'_1 < x'_2 < \dots < x'_{n'}\}$$

とする。ただし、 n, n' は抽出された射影線素数である。射影線素分布はその作成の仕方から明らかなように、必ずしも実在するストロークと1対1に対応せず、複数のストロークが射影軸上で一本に縮退したり、逆に視覚的にはストロークと認められない部分も短い射影線素として抽出されてしまう。したがって、2つの射影線素分布の対応付けを行う際には、射影線素対の1対1対応だけでなく、多対1対応も考慮する必要がある。ここでは簡単のために、一方の射影線素が、他方の射影線素分布の1本、あるいは2本に対応する場合と、対応付けがされない場合、すなわち、1対1、1対2、1対0対応のみを考えることにする。これらの対応の仕方に対して、距離はそれぞれ

$$\left. \begin{aligned} \text{1対0} : d((x_i, l_i); (x'_j = x_i, l'_j = 0)) &= \alpha l_i \\ &\quad (1 \leq i \leq n; n \geq 1) \\ \text{1対1} : d((x_i, l_i); (x'_j, l'_j)) &= \alpha |l_i - l'_j| + (1 - \alpha) \\ &\quad \cdot |x_i - x'_j| \\ &\quad (1 \leq i \leq n, 1 \leq j \leq n'; n \geq 1, n' \geq 1) \\ \text{1対2} : d((x_i, l_i); (x'_j, l'_j), (x'_{j+1}, l'_{j+1})) \\ &= \alpha |l_i - (l'_j + l'_{j+1})| \\ &\quad + (1 - \alpha) |x_i - (x'_j l'_j + x'_{j+1} l'_{j+1}) / (l'_j + l'_{j+1})| \end{aligned} \right\}$$

$$\left. \begin{aligned} (1 \leq i \leq n, 1 \leq j \leq n' - 1; n \geq 1, n' \geq 2) \end{aligned} \right\} \quad (4)$$

と定義される。ここで、 $\alpha = 1/3$ である。1対0対応の場合は、 (x_i, l_i) に対応する射影線素が、位置 x_i にある長さ0の射影線素と考えた場合の表式であり、また1対2対応の場合は、 (x_i, l_i) に対応する射影線素として、隣り合う2本の射影線素 (x'_j, l'_j) , (x'_{j+1}, l'_{j+1}) の重心に位置し、長さがそれらの和となるものを考えて得られる。これは、2.2で述べた統合の幅 w を大きくした場合に相当する。1対0、1対2対応は、1対1対応の自然な拡張となっていることがわかる。なお、方向 k の射影線素数が共に0 ($n = n' = 0$) のときの射影線素分布間の距離は、0と定めておく。以降、1対0、1対1、1対2の各対応関係をそれぞれ、 $i \leftrightarrow \phi$, $i \leftrightarrow j$, $i \leftrightarrow j \& j+1$ のように表わす。

2つの射影線素分布 S_k と S'_k の対応は、可能な対応の仕方の中で距離が最小となるものである。全ての対応の仕方について距離を求めるのは、射影線素数が多い場合には膨大な計算量になるので、以下に述べる方法により最適対応を探索する。まず、その準備として、全射影線素間の距離を計算しておく。これを距離マトリクス (distance matrix) と呼ぶ。図4は、中明朝体と教科書体の文字“明”について水平方向の射影線素間の距離を(4)式によって求めた結果を示している。

次に、最適対応を得る前に、最適ではないかもしれないが、最適に近い準最適対応を求めておく。これは、距離マトリクス上で距離の小さい順に対応付けを矛盾しないように行って得られ、準最適対応の距離 \tilde{d} は、この対応付けの平均距離である。図4では、実線で示すように、対応2 & 3 \leftrightarrow 3, 6 \leftrightarrow 5, 4 & 5 \leftrightarrow 4, 1 \leftrightarrow 1 & 2 が順に選ばれ、 $\tilde{d} = 4.3$ である。

さて、最適対応は、距離マトリクス上で距離の小さい順に対応付けを行っていき、もし、既に使われている射影線素が再び選ばれたならばその時点で新しい探索木を生成し、全ての探索を行って距離が最小となるものを探して得られる。新しい探索木は、その木が生成される時の平均距離が \tilde{d} より小さい場合に限り生成される。これにより、大幅に探索木の数を減らすことができる。図4では、準最適対応のほかにもう一つの対応2 & 3 \leftrightarrow 3, $\phi \leftrightarrow$ 5, 4 & 5 \leftrightarrow 4, 1 \leftrightarrow 1 & 2, 6 \leftrightarrow ϕ , $d = 4.8$ が見出されている。これは、同図に点線で示されている。この例では、準最適対応が最適な対応となっている。図5に文字パターンと対応付け

FONT = K1

STROKE X, L	0	1 7,41	1&2 11,76	2 16,35	2&3 21,101	3 40,66	3&4 56,136	4 71,70	4&5 78,24	5 114,14
0	-	13.7	-	11.6	-	22.0	-	23.3	-	4.6
1 9,70	23.3	11.0	5.0	16.4	23.0	22.0	53.0	41.3	60.3	88.7
1&2 19,105	-	33.0	-	33.7	-	23.3	-	42.6	-	90.0
2 40,35	11.6	23.9	32.6	16.0	30.0	10.4	44.0	32.4	51.3	56.3
2&3 43,62	-	31.4	-	27.4	-	3.7	-	21.0	-	63.0
3 47,27	9.0	31.4	39.9	23.3	37.4	17.7	46.7	30.3	49.3	49.0
3&4 61,63	-	42.6	-	38.7	-	14.3	-	11.0	-	52.4
4 73,36	12.0	45.7	54.3	38.3	51.7	32.0	45.1	12.7	30.3	34.7
4&5 76,66	-	54.3	-	50.3	-	23.9	-	4.6	-	42.6
5 79,30	10.0	51.7	60.3	43.6	57.7	38.0	50.9	18.7	32.4	28.7
5&6 94,48	-	63.4	-	59.4	-	45.1	-	25.7	-	24.3
6 118,18	6.0	81.7	90.4	73.7	87.7	68.0	81.0	48.6	39.0	4.0

DISTANCE MATRIX

図4 距離マトリクスの例

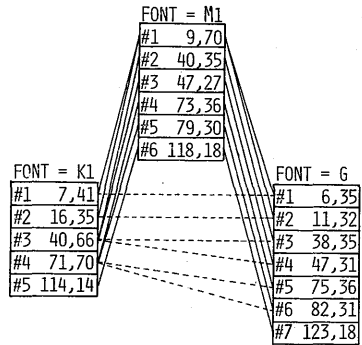
Fig.4-An example of distance matrix.

の結果を示した。

なお、2.3で述べた重み付けパラメータの決定は、中明朝体、教科書体、中ゴシック体の各文字の水平、および垂直方向の射影線素分布について、1対1対応の準最適対応をとって行った。

2.5 辞書の作成

射影線素を用いた認識用の辞書を、中明朝体、教科書体、中ゴシック体の3字体を用いて教育漢字881字について作成した。辞書には、3つの字体で矛盾なく対応がとれる射影線素の位置と長さの平均値を登録する。すなわち、中明朝体と教科書体、教科書体と中



DICTIONARY

No.	X, L (M1 K1 G)
#1	9,71 (1 1&2 1&2)
#2	41,64 (2&3 3 3&4)
#3	75,67 (4&5 4 5&6)

図6 射影線素辞書の作成

Fig.6- Generation of projected segments dictionary.

ゴシック体、中ゴシック体と中明朝体の3通りの場合について最適対応をとり、どの字体にも含まれかつ対応がうまくとれる射影線素のみを、その文字を特徴づける射影線素として登録する。したがって、字体固有のストロークや、対応がうまくとれないものは辞書に載らないことになる。図6に文字“明”の水平方向の射影線素について得られた結果を示す。この例では、3組の対応付けが見つかっており、その平均値が辞書に登録されている。また、中明朝体の6番目の射影線素は、教科書体の5番目と中ゴシック体の7番目の射影線素と対応がとれているが、教科書体と中ゴシック体

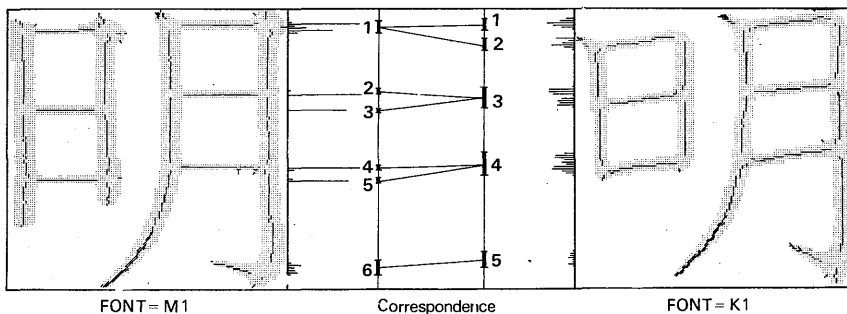


図5 射影線素の対応

Fig.5-Correspondence of projected segments.

の間で対応がとれていないので、辞書には登録されていない。

3. 分類実験と結果

分類は、粗外形パターンコードと縦横比コードを用いた大分類、および射影線素分布を用いた詳細分類の2段階で行う。この大分類法については、文献(4)に詳しく述べられているので、ここでは簡単に触れるに留める。

3.1 縦横比コード

縦横比は、文字パターンに外接する方形の高さ h と幅 w の比

$$r = w/h$$

である(図7)。これを用いて縦長あるいは、横長の文字を分離し、詳細分類にかけられる候補文字数を減らすことができる。このために、縦横比の大小に応じて次のように3ビットの符号を各文字に与える。

$$r_c = \begin{cases} '001' & : r > 1.3 \\ '010' & : 0.7 \leq r \leq 1.3 \\ '100' & : r < 0.7 \end{cases}$$

文字カテゴリ f の標準縦横比コード \bar{r}_c^f は、字体 α の文字 f に対して得られる縦横比コード $r_c^{f\alpha}$ の全字体にわたる論理和として定義され、

$$\bar{r}_c^f = r_c^{f1} \vee r_c^{f2} \vee \dots \vee r_c^{f\nu}$$

となる。ここで ν は標準コード作成に用いられる字体数である。また、文字 g の縦横比コード r_c^g と標準縦横比コード \bar{r}_c^f の距離 $D_r(f, g)$ は、

$$D_r(f, g) = \begin{cases} 0 & : r_c^g = \bar{r}_c^f \wedge r_c^g \\ 1 & : r_c^g \neq \bar{r}_c^f \wedge r_c^g \end{cases}$$

である。

3.2 粗外形パターンコード

粗外形パターンコードは、正規化された文字パターンを、上下左右の4方向から眺めたときに、文字を構成するストロークによって全ての方向から遮られる部分を黒点にした外形文字パターンをコード化したものである。外形文字パターンは文字の輪郭形状を表わしているが、漢字のつづし文字の心理実験⁽⁸⁾や、文字パターンのエントロピー評価実験⁽⁹⁾によって、文字の大きな特徴がその輪郭に表われていることが確かめられており、外形文字パターンから文字の有効な情報が得られると考えられる。粗外形パターンコード(outline pattern code)は、外形文字パターンを 4×4 の粗いメッシュ領域に分割したときの各メッシュ領域の黒点数を、メッシュ毎に定められた閾値と比較して、その

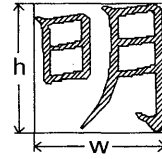


図7 縦横比

Fig.7- Ratio of Kanji character's width to its length.

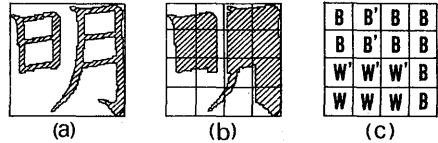


図8 粗外形パターンコードの抽出

Fig.8-Extraction of outline pattern code.

大小に応じて2ビットの符号 '10', '01' を各メッシュに与えたものである。いま、字体 α の文字 f の粗いメッシュ領域 (i, j) における黒点数を $M_{ij}^{f\alpha}$ ($i, j = 1, \dots, 4$; $f = 1, \dots, N$; $\alpha = 1, \dots, \nu$) とすると粗外形パターンコード $C_{ij}^{f\alpha}$ は、

$$C_{ij}^{f\alpha} = \begin{cases} '10' & : M_{ij}^{f\alpha} \geq \bar{M}_{ij} \\ '01' & : M_{ij}^{f\alpha} < \bar{M}_{ij} \end{cases}$$

となる。ここで \bar{M}_{ij} は、領域 (i, j) における黒点数の平均、

$$\bar{M}_{ij} = \frac{1}{\nu N} \sum_{\alpha=1}^{\nu} \sum_{f=1}^N M_{ij}^{f\alpha}$$

また、 N, ν はそれぞれコード作成に使われる文字カテゴリ数、および字体数である。

文字カテゴリ f を代表する標準粗外形パターンコード \bar{C}^f は、 $C^{f\alpha}$ の全字体にわたる論理積をとったものとして定義され、

$$\bar{C}^f = C_{ij}^{f1} \wedge C_{ij}^{f2} \wedge \dots \wedge C_{ij}^{f\nu}$$

となる。したがって、標準コードには、安定な黒領域、安定な白領域、または不安定な領域に対応して、それぞれ、コード '10', '01', '00' が与えられることになる。文字 g から得られる粗外形パターンコード C^g と標準粗外形パターンコード \bar{C}^f との距離 $D_c(f, g)$ は、各メッシュ領域の距離、

$$d_{ij}^{fg} = \begin{cases} 0 & : \bar{C}_{ij}^f = \bar{C}_{ij}^f \wedge C_{ij}^g \\ 1 & : \bar{C}_{ij}^f \neq \bar{C}_{ij}^f \wedge C_{ij}^g \end{cases}$$

の和として与えられ、

$$D_c(f, g) = \sum_{i,j=1}^4 d_{ij}^{fg}$$

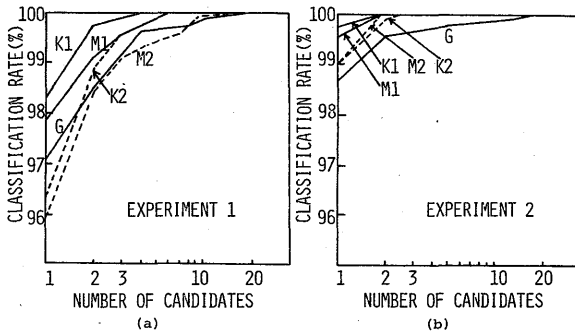


図9 分類実験結果
Fig.9-Classification results.

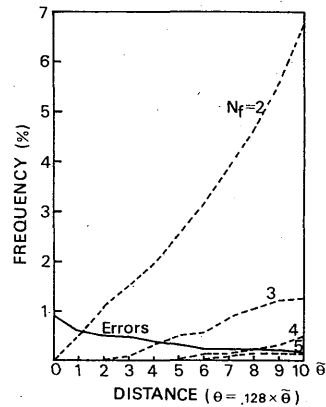


図10 識別の余裕
Fig.10-Recognition margin.

となる。なお、図8(a)に正規化された文字パターン、(b)に外形文字パターン、(c)に粗外形パターンコードを示した。図8(c)で、*W*および*B*は、それぞれコード'01', '10'に対応し、また'*l*'は、標準粗外形パターンコードで'00'が与えられる領域である。

3.3 大分類

標準縦横比コードと標準粗外形パターンコードは、中明朝体、太明朝体、教科書体2字体、中ゴシック体の教育漢字881字を用いて作成された。文字カテゴリ-*f*の標準コードと入力文字*g*の距離は、

$$D(f, g) = D_r(f, g) + D_o(f, g)$$

となり、文字*g*に対する候補文字カテゴリ-*f*は、 $D(f, g) = 0$ となるものである。なお、この大分類により、候補文字数が平均20文字程度になることが示されている⁽⁴⁾。

3.4 詳細分類

詳細分類は、大分類で選ばれた候補文字について、4方向の射影線素分布の準最適対応をとることによって行われる。ここで距離は、各方向の距離の平均とする。また、最適対応をとらないのは、計算時間の短縮のためと、認識では距離だけが必要で詳細な対応関係を知る必要がないためである。

3.5 分類実験と結果

中明朝体、太明朝体、教科書体2字体、中ゴシック体の教育漢字881字について分類実験を行った結果を図9(a)に示す(実験I)。この図で実線は射影線素辞書を作成するのに用いた3字体、破線は他の2字体に対応している。これより、第1位分類率、すなわち識別率は学習サンプルで平均97.72%、未知サンプルで平均96.13%となっている。

ここで得られた識別率は決して高い値ではない。そ

の原因として、射影線素分布辞書を作成する際、実際に存在する線素が字体M1、K1、G間で矛盾なく対応するという条件を満たせずに辞書に載っていないことが挙げられる。例えば、図6において目視では明らかに対応付けられる字体M1の6番目、字体K1の5番目、字体Gの7番目の射影線素は、K1、G間の対応がとれないためにこれらに対応する射影線素が辞書に登録されていない。そこで、この例のように目視により対応付けが行え、しかも既に作成した辞書との違いがある場合に限り辞書に手を加え識別率の向上を図った。ここで辞書修正の対象となる文字カテゴリ-*f*は、各字体で誤認識をしたものである。この修正辞書を用いた分類実験の結果を図9(b)に示す(実験II)。これより平均識別率99.16%、第2位分類率99.89%、また順位が3位以下になる文字はわずか5文字となっており、高い分類率が得られている。実験Iと比べると著しく分類率が向上し、辞書修正の効果が大きいことがわかる。

実験IIでは、中ゴシック体の分類率が他の4字体に比べて悪いのが目立っている。これは射影線素を抽出する際に、複数のストロークが合流する点に対して特別な処理を行っていないことにより、中ゴシック体の射影線素の長さが対応するストロークの長さよりもかなり短くなっていることが原因と考えられる。例えば、文字“書”の中央の縦ストロークの実際の長さに対する抽出された長さの割合は、中ゴシック体では0.43、また他の4字体ではほぼ等しく0.7程度となっており、中ゴシック体での実際の長さとの差が目立っている。本手法によると、抽出されるストロークの長さは実際

の長さより短くなるが、交点数が多く、またゴシック体のように線幅が大きい場合に特に顕著になる。中ゴシック体の識別率を上げるには、セグメント結合を行って、この影響を取り除くか、または誤認識した文字について辞書を追加する必要があると考えられる。

本手法の識別能力を、累積分類率と異なった視点、つまり各入力文字と第1位文字カテゴリーの距離と、下位文字カテゴリーとの距離の差によって示される識別の余裕 (recognition margin) によって見ることもできる。いま、入力文字を f 、大分類で得られた候補文字カテゴリーを g_i ($i=1, \dots, n$)、 f と g_i の射影線素分布間の距離を $d(f, g_i)$ とすると、距離が最小となる文字カテゴリー r に対し、

$$d(f, g_i) - d(f, g_r) \leq \theta$$

を満たす文字カテゴリー数 N_f は、識別の余裕が θ のときに混入してくる文字数である。 $N_f=1$ のときは、認識結果は一意に定まるが、 $N_f \geq 2$ の場合、複数の文字カテゴリーが結果として出力される。図10は、実験Ⅱの場合について、混入してくる文字数が N_f となる文字数の割合を示したものである。ここで、 θ は0.128刻みの値をとる。図から、余裕を増やすと共に混入してくる文字数がほぼ直線的に増大しているのがわかる。また、誤認識率は $\bar{\theta}$ が6までの間で急速に減少し、その後は緩やかに減少している。これは、射影線素分布が類似しているために起こる誤認識が $\bar{\theta}=6$ 程度のところまでに多く存在し、 $\bar{\theta} > 6$ における誤認識は、字体固有の変形が大きいため本手法では対応しきれないために生じていることを示している。したがって、これらの文字に対して個別の辞書を用意すれば、 $\bar{\theta}$ を6程度に設定することにより、正解文字カテゴリーを常に出力結果に含ませられると考えられる。 $\bar{\theta}$ を6とした場合に、認識結果が一意に定まらない文字数は、881文字中35文字とかなり少なくなっており、本手法の有効性を示している。また、辞書容量は大分類用5Kバイト、詳細分類用25Kバイトと非常に小さくなっている。

4. むすび

本論文では、中明朝体、太明朝体、中ゴシック体、および教科書体2字体の多種字体印刷文字を対象に、射影線素分布を用いた詳細分類法を示し、縦横比コードと粗外形パターンコードによる大分類を組み合わせた2段階分類により、教育漢字881字について分類実験を行った結果、識別率9.916%、第2位分類率

9.989%を得て、本手法の有効性が確認された。

本手法の特長は、

(1) 安定な方向を持つ線素を抽出し垂直な軸上に射影して得られる分布を特徴量とするので、字体の違いによる字形変形の影響を受けにくい。

(2) 射影線素分布間の対応付けが、多対1対応まで考慮して動的になされるため、柔軟な対応付けができる。

などが挙げられる。また、問題点として、

(1) 文字線が交差する部分でセグメント結合をしていないため、交点数が多い文字や線幅の太い字体で抽出される射影線素の長さが、対応するストロークの長さよりかなり短くなり、分類率の低下をもたらす。

(2) 微かな形状変化に鈍感なため、類似文字の識別が困難である。

などである。今後の課題として、

(1) 2段階分類により一意に結果が定まらない文字に対する識別アルゴリズムの開発

(2) 文字線交差部の影響を除くこと

などがあり、これらの点については現在検討中である。

謝辞 日頃、有益な御討論頂く篠原正美氏に感謝致します。

文 献

- (1) 目黒, 梅田: “マルチフォント印刷漢字の認識”, 信学論(D), J65-D, 8, pp. 1026-1033(昭57-08).
- (2) 目黒, 梅田: “マルチフォント印刷漢字認識実験装置”, 信学技報, PRL82-46(1982-11).
- (3) 北島, 吉岡, 佐川, 増田: “文字・音声変換実験システムの概要”, 昭58信学総全大, 1454.
- (4) 北島宗雄: “教科書体を含むマルチフォント印刷文字認識”, 信学論(D), J66-D, 4, pp. 400-406(昭58-04).
- (5) 馬場口, 北村, 塩野, 真田, 手塚: “細線化を用いない文字パターン方向セグメント抽出の一手法”, 信学論(D), J65-D, 7, pp. 874-881(昭57-07).
- (6) 馬場口, 相原, 真田, 手塚: “文字パターンにおける方向セグメントの接続形状と幾何学的特徴部の抽出”, 信学論(D), J66-D, 5, pp. 495-502(昭58-05).
- (7) 斉藤, 山田, 山本: “手書漢字の方向パターンマッチング法による解析”, 信学論(D), J65-D, 5, pp. 550-557(昭57-05).
- (8) 森, 青沼: “文字の構造(1)”, 信学会パターン認識と学習研資, PRL73-85(1974-01).
- (9) 坂井, 森: “漢字パターンの大分類”, 信学会パターン認識と学習研資, PRL73-14(1973-05).

(昭和58年6月13日受付, 9月8日再受付)