

教科書体を含むマルチフォント印刷文字認識

正 員 北島 宗雄[†]A Recognition Method for Multifont Chinese Characters
Including Kyokasho-Type for School TextbookMuneo KITAJIMA[†], *Regular Member*

あらまし 教科書を含む一般の書物の文字認識を目的として、明朝体、ゴシック体、および、教科書体を対称とするマルチフォント印刷文字認識のための分類方法について述べる。分類は二段階で行なわれる。まず、大分類では、二つの特徴パターンコード、すなわち、印刷文字を構成しているストロークの大きな存在形状や位置を表わす粗メッシュパターンコード、および、文字の輪郭の大きな状態を表わす粗外形パターンコードの各コードによる分類の可能性について検討した。その結果、粗外形パターンコードの方が大分類に適しており、これと縦横比コードを併用することにより、平均候補文字数を全文字カテゴリー数の約1/50に絞ることができた。次に、細分類は、縦・横ストロークの平均的な本数を表わす32次元の平均線数ベクトルを用いて、類似度を尺度として行なわれ、認識率98.27%、第4位分類率99.96%が得られた。

1. ま え が き

計算機による文字認識の研究は盛んに進められ、字体を限定した印刷文字に関しては、実用化の域にまで達している。現在では、文字認識機械の性能をさらに高めるために、認識対象とする字体を限定しないマルチフォント印刷文字認識の研究が行なわれている。

印刷文字は同一字種であっても字体が違うことによって、その文字を構成しているストロークの存在する位置や形状に何らかの差異があるために、多種字体の印刷文字を対象とする場合には、単一字体の印刷文字を対象とした場合よりも字形の変形に対して強い認識手法が要求される。マルチフォント印刷文字認識への一つのアプローチとして、従来単一字体の印刷文字を対象に用いられていた方法をそのまま適用することが考えられる。その例として、三宅らが行なった複合類似度法を用いたマルチフォントの印刷数字の認識があり、かなり高い正読率を得ている⁽¹⁾。もう一つのアプローチは、最初からマルチフォント印刷文字を対象として、字体の違いによる変形に影響されない特徴を抽出し、認識を行なおうとするものである。このような

考えのもとに、梅田は64次元の粗いメッシュパターンと、粗いベリフェラルパターンを作成し、まず16安定点による大分類を行ない、次に、そこで得られた候補文字について、各特徴パターンの全要素を用いた細分類を行なっている^{(2),(3)}。この方法は、明朝体とゴシック体の混在する印刷文字を対象にその有効性が確認され、現在ハード化が検討されている⁽⁴⁾。

認識対象として、主に明朝体とゴシック体で印刷されている一般の書物のほかに、教科書を加える場合には、さらに教科書体を含んだマルチフォント印刷文字認識を行なう必要がある。教科書体は、明朝体およびゴシック体が主に直線から構成され、高度にデザイン化されているのに比べ、曲がったストロークが多く、手書き文字に近い形状をしている。したがって、認識対象字体に教科書体を含めることにより、字体の違いによる変形は、それを含まない場合に比べて、著しく増大することが予想され、その結果、明朝体、ゴシック体、および教科書体を対象とするマルチフォント印刷文字認識方法としては、従来のものよりも、変形に対し、さらに安定なものが必要になる。

本論文では、明朝体、ゴシック体、および教科書体の混在するマルチフォント印刷文字を効率よく認識するための、大分類、細分類から成る二段階分類方法について述べる。すなわち、大分類法として、文字パターンから抽出される二つの特徴パターンコード、つまり、

[†]工業技術院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
論文番号：昭58-論192[D-48]

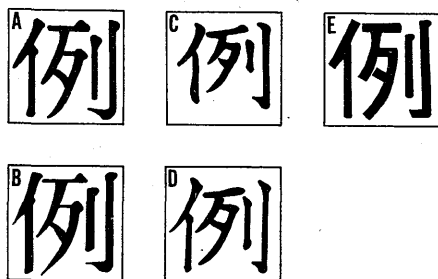


図1 印刷文字データの例

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

- (1) 文字を構成するストロークの大まかな存在位置を表わす粗メッシュパターンコード
- (2) 文字パターンから作成される輪郭パターンの大まかな形状を表わす粗外形パターンコード

の、各特徴パターンコードの変形に対する安定性、および分類効率を検討し、さらに、マルチフォント印刷文字、手書き文字の認識において有効性が確かめられている線密度情報^{(6),(6)}を利用した細分類を行なう。

2. 印刷文字データ

文字印刷には、種々の書体が用いられるが、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、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。この点を勘案し、印刷文字サンプルとして、一般の書物の本文に使われる中明朝体、見出しに使われる太明朝体、および中ゴシック体、教科書に使われる教科書体2種類の、合計5字体を選んだ。字種は、教育漢字881文字、平仮名、片仮名142文字、数字10文字からなる1033文字である。

印刷文字データは、文字の切れ、かけ等が少なく質の良い写植印刷文字(50級)をテレビカメラを用いて、128×128メッシュの分解能で撮影し、照明むらの補正⁽⁷⁾をした後、各文字毎にしきい値を定め、2値化して作成された。一例を図1に示す。字体はそれぞれ、A：中明朝体、B：太明朝体、C：教科書体1、D：教科書体2、E：中ゴシック体である。なお、ここで用いた写植印刷文字は、全てM社のものであり、各書体で最もよく使われる字体を選んだ。

図1に示したように、書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が、特に、手書き文字に近く曲がったストロークの多い教

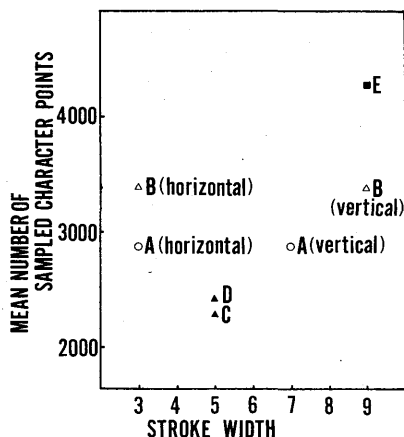


図2 各フォント文字パターンの平均黒点数と線幅

Fig. 2 - Mean number of sampled character points and stroke width of Kanji characters in each type font.

科書体と、主に文字のバランスを考えて設計され、直線の多い明朝体、ゴシック体との違いが顕著である。また、書体が同じでも、字体が異なることによって上記の変形が認められる。

字体の違いによる変形の一例として、各字体の全カテゴリーパターンについて求めた平均の黒点数、および文字パターンを縦方向と横方向に走査して得られる黒ラン数の頻度分布から求めた線幅を図2に示す。明朝体では、縦ストロークと横ストロークの線幅が全く違うので、その両方を示した。平均黒点数は、縦・横ストローク共に細い教科書体が最も小さく、縦・横ストローク共に太い中ゴシック体が最大で、その差は、約12%(2000メッシュに相当)にも及んでいる。また、線幅は、縦ストロークについては、教科書体が5メッシュと最も細く、これは、中ゴシック体と太明朝体の線幅の約半分である。一方、横ストロークについては、その差はさらに大きくなり、中明朝体、太明朝体の線幅が3メッシュであるのに対し、中ゴシック体の線幅は、その3倍にもなっている。このように、平均黒点数、線幅だけを見ても、字体の違いによる変形がいかに大きいかかわかる。

3. 特徴パターン

3.1 粗メッシュパターンコード

文字を構成しているストローク、即ち黒点の存在位置によって文字を特徴づけようとする場合、字体の違いによる変形の影響を受けにくくするためには、128

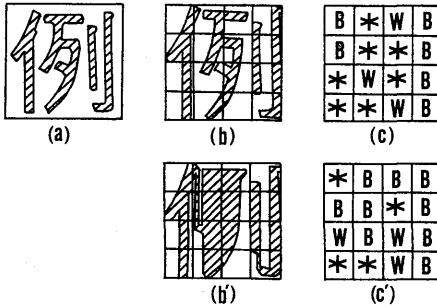


図3 粗メッシュパターンコードと粗外形パターンコードの作成

Fig. 3-Extraction of feature pattern codes.

×128メッシュの原文字パターンを適当に粗くする必要がある。

最も単純な方法は、原文字パターンそのものを粗くして、粗いメッシュパターンを作成するものである。粗いメッシュパターンは次のようにして作成される。まず、文字の大きさによる影響を取り除くために、全ての文字パターンを128×128メッシュの大きさにそろえる。次に、それを、 $n \times n$ の粗いメッシュ領域に分割して、各メッシュ領域に含まれる黒点数を計数し、それをその領域の値 M_{ij}^{α} ($i=1, \dots, n; j=1, \dots, n$)とする。したがって、粗いメッシュパターンは、 n^2 個の成分を持つベクトルとして表わされる。このようにして作られた粗いメッシュパターンは、印刷文字を構成しているストロークの大きかな存在位置を表わしてはいるが、線幅によってその値は大きく変化すると考えられる。先に述べたように、字体の違いによる線幅の変動はかなり大きいので、線幅に強く依存する粗いメッシュパターンが、十分に変形を吸収しているとは言い難い。そこで、粗いメッシュパターンを変形に対してさらに強くするために、次のように粗メッシュパターンコード(mesh pattern code)を定めた。

まず、字体 α の文字 f の粗いメッシュ領域(i, j)における黒点の計数値 M_{ij}^{α} ($i, j=1, \dots, n; f=1, \dots, N; \alpha=1, \dots, \nu$)のその領域における平均値

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

を求める。ここで、 N, ν はそれぞれコード作成に使われる文字カテゴリー数、および、字体数である。次に、 M_{ij}^{α} と \bar{M}_{ij} を比較し M_{ij}^{α} が \bar{M}_{ij} より大きい場合、小さい場合に応じて、それぞれコードとして、2 bitの値、'10'および'01'を割り当てる。即ち、字体 α の文字 f に対する粗メッシュパターンコード $C_{ij}^{Mf\alpha}$

の(i, j)成分 $C_{ij}^{Mf\alpha}$ は、

$$C_{ij}^{Mf\alpha} = \begin{cases} '10' & : M_{ij}^{\alpha} \geq \bar{M}_{ij} \\ '01' & : M_{ij}^{\alpha} < \bar{M}_{ij} \end{cases} \quad (1)$$

となる。各文字カテゴリーを代表する標準粗メッシュパターンコードは、 $C^{Mf\alpha}$ の全字体にわたる論理積をとったものとして定義できる。つまり、標準粗メッシュパターンコードを \bar{C}^{Mf} とすると、その成分 \bar{C}_{ij}^{Mf} は、

$$\bar{C}_{ij}^{Mf} = C_{ij}^{Mf1} \wedge C_{ij}^{Mf2} \wedge \dots \wedge C_{ij}^{Mf\nu} \quad (2)$$

となる。 \bar{C}_{ij}^{Mf} は、その構成のしかたから明らかなように、全ての字体について M_{ij}^{α} がその平均値 \bar{M}_{ij} より大きいときには、'10'、小さいときには、'01'、その他の場合には、'00'という値をとり、各粗メッシュ領域に、字体の違いによる文字の変形によらず安定にストロークが存在するかどうかを表わしている。

分割数 n を4、文字カテゴリー数 N を1033、字体数 ν を5とした場合に標準粗メッシュパターンコードが作られる様子を図3に示す。まず、原文字パターン(図3(a))を128×128メッシュの大きさに正規化する。次に、それを4×4の粗いメッシュ領域に分割し(図3(b))、各粗メッシュ領域に含まれる黒点数を計数する。これを同一文字カテゴリーの5種類の字体について行ない、(1),(2)式を用いて標準粗メッシュパターンコードを得る(図3(c))。図3(c)では、標準粗メッシュパターンコード'10','01','00'をそれぞれ文字'B','W','*'で表わした。なお、各粗メッシュ領域の平均値 \bar{M}_{ij} は表1に示されている。

3.2 粗外形パターンコード

粗いメッシュパターンは、印刷文字を構成しているストロークの位置情報を全て利用しているために、ストロークの位置変動に対し、非常に弱いという欠点もっている。そこで、この欠点を取り除くために、印刷文字を外側から眺めたときに見える外形文字パターンを作成する。外形文字パターンは、文字の輪郭形状を表わしているが、漢字のつぶし文字の心理実験⁽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か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期待される。

外形文字パターンは、原文字パターンを、上下左右の4方向から眺めたときに、文字を構成してい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パターンであり、森、青沼のつぶし文字パターン作成法⁽⁸⁾を参照した。一例を図3(b')に示す。図からわかるように、外形文字パターンは文字の輪郭を構成

表1 各メッシュ領域の平均値と相関比
(上段:平均値, 下段:相関比)

		j			
		1	2	3	4
粗 メ ッ シュ パ タ ー ン	1	1 8 9.6	2 8 1.1	3 0 9.9	2 1 8.6
		0.6 6 4	0.6 5 4	0.6 3 6	0.7 6 4
	2	2 6 4.6	3 2 2.7	3 4 2.1	2 6 0.3
		0.6 6 3	0.6 6 0	0.6 3 9	0.6 8 1
	3	2 5 0.3	2 9 6.0	3 2 5.3	2 2 9.5
		0.6 7 8	0.6 9 8	0.6 7 0	0.7 2 5
	4	2 2 3.2	2 4 6.7	2 6 2.9	2 6 2.9
		0.6 7 2	0.6 0 8	0.6 6 5	0.7 8 3
外 形 パ タ ー ン	1	2 3 3.7	4 4 0.3	4 9 3.8	2 7 8.0
		0.7 7 6	0.8 3 4	0.8 6 5	0.8 4 4
	2	3 8 6.4	7 1 4.8	7 4 4.5	3 7 8.9
		0.8 2 2	0.8 6 2	0.8 7 7	0.8 1 9
	3	3 4 8.1	6 5 9.4	7 0 2.4	3 4 8.6
		0.8 0 0	0.8 8 0	0.8 8 4	0.8 4 7
	4	2 6 2.5	4 1 5.8	4 4 8.2	3 3 6.8
		0.7 7 5	0.8 2 3	0.8 4 9	0.8 5 1

するストロークの存在状態を表わして、内部のストロークの影響を受けないので、ストロークの位置変動に対し原文字パターンよりも安定になっていると考えられる。

原文字パターンから字体の違いによる変形の影響を受けにくくするために粗メッシュパターンコードを作成したように、外形パターンから同様の手続きにより、粗外形パターンコードを定めることができる。一例として、文字'例'に対して得られた粗外形パターンコード(outline pattern code)を図3(c')に示す。

4. 大分類

4.1 粗い特徴パターンコードによる分類

このようにして作成された標準粗メッシュパターンコード \bar{C}^{Mf} と文字 g から得られる粗メッシュパターンコード C^{Mg} の距離 $D_a(f, g)$ は粗メッシュ領域 (i, j) における距離 d_{ij}^{fg}

$$d_{ij}^{fg} = \begin{cases} 0 & : \quad \bar{C}_{ij}^{Mf} \equiv \bar{C}_{ij}^{Mf} \wedge C_{ij}^{Mg} \\ 1 & : \quad \bar{C}_{ij}^{Mf} \neq \bar{C}_{ij}^{Mf} \wedge C_{ij}^{Mg} \end{cases}$$

を用いて

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

と定義される。この定義から明らかなように、粗メッシュ領域 (i, j) における距離 d_{ij}^{fg} は、文字 g の黒地

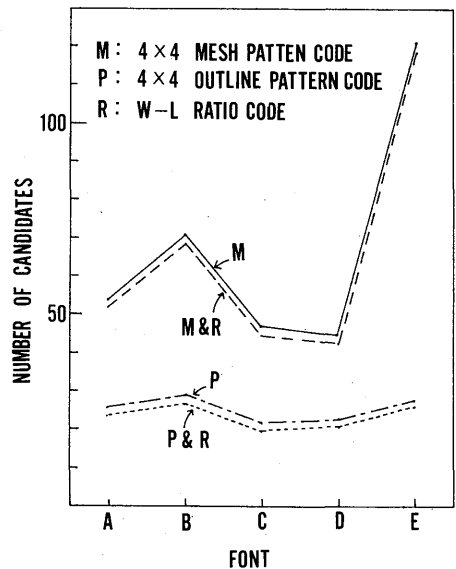


図4 二つの特徴パターンコードと縦横比コードを用いた分類結果

Fig. 4 - Classification results by using two 4x4 feature pattern codes and W-L ratio code.

(白地)の部分为标准パターン上で、字体の違いによって変動しない安定な黒地(白地)になっているとき、あるいは、不安定で白地とも黒地とも断定できないメッシュ領域となっている場合に0となり、逆に、文字 g の黒地(白地)が標準パターン上の安定な白地(黒地)になっているとき1となる。したがって、各粗メッシュ領域における距離の和として定義される距離 $D_a(f, g)$ は、文字 g の黒地・白地領域が、標準パターンの安定な黒地・白地領域と合致していない領域数を表わしている。なお、外形パターンの場合も、距離は、全く同様に定義される。

さて、文字 g に対する候補文字 f は、標準粗パターンコードと、文字 g から得られる粗パターンコードの距離 $D_a(f, g)$ が0となる文字カテゴリーである。全文字に対し、二つの特徴パターンを用いて候補文字数を求め、字体毎に平均して得られる平均候補文字数を図4に示す。この図から、粗メッシュパターンを用いて分類すると、どの字体についてもその平均候補文字数が、粗外形パターンを用いた場合よりも著しく大きくなっていることがわかる。また、前者によった場合、平均候補文字数は、字体の違いによってかなり変動し、線幅が大きい中ゴシック体、および天明朝体では、特に多くなっている。一方、後者では、そのような傾向は見られず、平均候補文字数は字体によらずほとんど

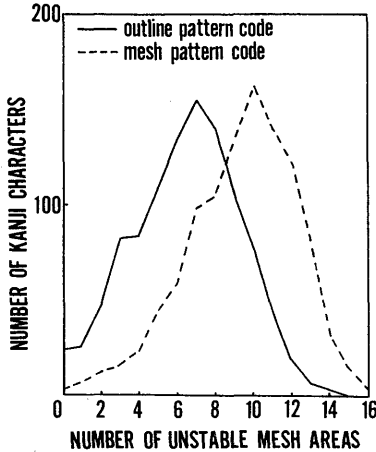


図5 不安定なメッシュ領域数の頻度分布
Fig. 5-Distribution of the number of unstable mesh areas.

一定である。このことから、印刷文字を構成しているストロークに注目して作られた粗メッシュパターンコードよりも、文字の輪郭情報を利用し、内部のストロークの状態の影響を受けにくい粗外形パターンコードの方が、より高い分類能力をもち、また、字体の違いによる文字の変形をよく吸収していることがわかる。

ここで作成した二種類の特徴パターンが、どの程度字体の違いによる文字変形に対して強いを示す指標の一つに、各メッシュ領域の相関比がある。相関比 η^2 は、全分散 σ_f^2 に占める文字カテゴリー間分散 σ_b^2 の割合と定義されるので、 η^2 が1に近い程、同一文字カテゴリーの字体間分散が小さく、したがって、字体の違いによる文字変形が小さいことになる。表1に、5字体各1033文字について得られた、粗メッシュパターン、および粗外形パターンの各メッシュ領域における相関比を示す。相関比は、全ての領域で一様に、粗メッシュパターンよりも粗外形パターンの方が大きくなっており、その差は、最大0.24にもなっている。このことは、粗メッシュパターンよりも粗外形パターンの方が、字体の違いによる文字変形の影響を受けにくい特徴パターンであることを示している。

また、字体の違いによる文字変形に対する粗メッシュ領域の不安定性は、標準パターンコード上において、黒地とも白地とも断定できない不安定なメッシュ領域の個数に反映されている。二種類の標準パターンコードについて、不安定メッシュ領域数の頻度分布を求めた結果を図5に示す。平均不安定メッシュ領域数は、粗メッシュパターンコードでは、9.3メッシュとなり、



図6 縦横比
Fig. 6-Ratio of Kanji character's width to its length ($W-L$ ratio).

候補文字を選ぶのに使われない不安定なメッシュの方が、有効な情報を持っているものよりも多くなっている。一方、粗外形パターンコードでは、平均不安定メッシュ領域数は、6.5メッシュとなり、粗メッシュパターンコードに比べかなり小さくなっている。候補文字選択に有効な情報を持たないメッシュ領域が多い程候補文字になる場合が増えるが、粗メッシュパターンコードを用いた時の平均候補文字数が、粗外形パターンコードの場合よりも多くなるのはこのためである。

このようなことから、粗外形パターンコードは、字体の違いによる文字変形を比較的良好に吸収しており、特に、粗メッシュパターンコードでは吸収しきれなかったストローク幅の変動の影響をあまり受けないといえることができる。

4.2 縦横比

これまでに述べた特徴パターンは、大きさを正規化した文字パターンに対し得られるものであるが、正規化する前の原文字パターンの高さと幅の比(縦横比, $W-L$ ratio)を利用することにより、さらに候補文字数を減らすことができる。縦横比 r は、文字を囲む外接方形の高さ h と幅 w の比

$$r = w/h$$

と定義される(図6)。中明朝体の1033文字について r を求め、その値の大きい順に並べた結果を図7に示す。図から、ほとんどの文字は縦横比が1に近く正方形に近い形をしており、約5%程度の文字が縦長、あるいは横長の形状をしていることがわかる。これらの文字を安定に分離するために、縦横比の大小に応じて、次のように3 bitの符号 r_c を各文字に与える。

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

文字カテゴリー 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 $r_c^{\alpha f}$ の全字体にわたる論理和として定義され、

$$\bar{r}_c^f = r_c^{\alpha 1} \vee r_c^{\alpha 2} \vee \dots \vee r_c^{\alpha \nu}$$

表2 各文字に割り当てられた縦横比コード

コード	文字カテゴリー
001	一つへべ*ハへユ*バ*
011	五口三四二ハエ心入以必いづべ*マル*ロバ
010	(他の文字のカテゴリー)
110	月日目多角自さしそとまもよをぐイ4
100	うくらりトミ*リド123567890

(* ; 片仮名)

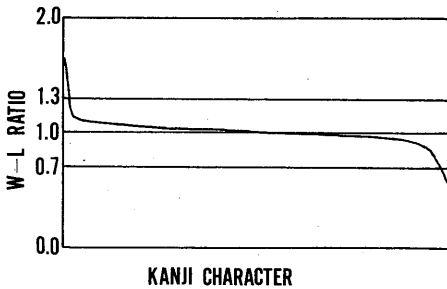


図7 中明朝体の各文字の縦横比

Fig. 7-W-L ratios of the Kanji characters in Chu-mincho type font.

となる。表2に1033の各文字カテゴリーに与えられた標準縦横比コードを示す。表に見られるように、縦横比によって、縦長あるいは横長の文字として分離される文字カテゴリーは、平仮名、片仮名、数字に集中しており、漢字は、ほとんど正方形に近い形状をしていることがわかる。

文字 g の縦横比コード r_g^f と標準縦横比コード \bar{r}_g^f の距離 $D_r(f, g)$ は、

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

と定義される。標準パターンコードと標準縦横比コードを併用したとき、標準パターン f と入力文字パターン g の距離 $D(f, g)$ を、両コード上の距離の和、

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

と定義すると、文字 g に対する候補文字カテゴリーは $D(f, g) = 0$ となるものである。全文字に対し候補文字数を求め、字体毎に平均して得られる平均候補文字数を図4に示す。図から、標準粗パターンコードだけを用いた場合に比べ、どの字体についても、2文字程度、平均候補文字数が減っていることがわかる。

以上のことから、標準コードとして粗外形パターンコードと縦横比コードを併用して大分類を行なうことにより、比較的少ない候補文字カテゴリーを、極めて

簡単な論理演算で高速に選び出せることがわかる。ここで得られた平均候補文字数は約20文字であるが、これは全文字カテゴリー数の約1/50に当たり、この方法が優れた大分類法であることを示している。

5. 細分類

平均線数を用いた分類

前章で述べたように、入力文字パターン g に対して標準コードとの距離 $D(f_i, g)$ が0の文字カテゴリー $f_i (i = 1, \dots, n)$ が候補文字として選ばれてくるが、文字 g を認識するためには、さらに分類を進め、候補文字を絞る必要がある。ここで用いる方法としては、第一段階の候補文字選択に使った特徴(粗外形パターンと縦横比)と相補的な特徴、つまり、文字輪郭線の内部の特徴を利用する方法で、また、字体の違いによる文字変形を十分に吸収できるものが望ましい。このような点から、文字を構成しているストロークの本数によって定められる平均線数を第二段階の分類に用いる。

平均線数は次のように定義される。まず、大きさの正規化を行ない、文字パターンを 128×128 メッシュの大きさにそろえる。正規化された文字パターンを、

$$f(x, y) = \begin{cases} 1 & (\text{黒点}) \\ 0 & (\text{白点}) \end{cases} \quad (x, y = 1, \dots, 128)$$

とすると、横線パターン $S_h(x, y)$ は、原パターンとそれを縦方向に1メッシュ平行移動したパターンとの差の絶対値、

$$S_h(x, y) = |f(x+1, y) - f(x, y)| \quad (\text{ただし, } f(129, y) = 0; x, y = 1, \dots, 128)$$

同様に、縦線パターン S_v は、

$$S_v(x, y) = |f(x, y+1) - f(x, y)| \quad (\text{ただし, } f(x, 129) = 0; x, y = 1, \dots, 128)$$

と定義される。平均線数は、このようにして得られた横および縦線パターンの 4×4 粗メッシュ領域における計数値を成分とする32次元のベクトル $S = (S_1, \dots, S_{32})$ である。文字カテゴリー f の標準平均線数ベクトル \bar{S}^f は、収集した5字体の平均線数を平均して得られ、また、入力文字パターン g の平均線数ベクトル S^g と標準平均線数ベクトル \bar{S}^f との類似度 $\rho(f, g)$ は、

$$\rho(f, g) = \frac{(\bar{S}^f, S^g)}{\|\bar{S}^f\| \|S^g\|}$$

で与えられる。なお、ここで導入した平均線数ベクトルは、内藤ら⁹⁾が用いているストローク密度に、より

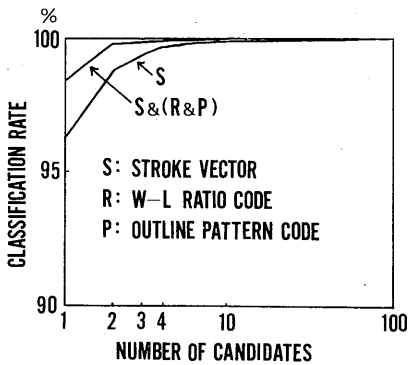


図8 分類結果

Fig. 8 - Classification results.

細かい位置情報を加えたものである。このことによって、ストロークの位置変動に対する安定性は減少するものの、その反面、情報量は多くなり、細分類を行なうのに必要なかなり詳細な情報を持つことができると考えられる。

収集した全文字パターンから標準パターンを作成し、それを用いて分類実験を行なった結果を図8に示す。この図は、まず粗外形パターンと縦横比を用いて入力文字パターン g に対する候補文字 $f_i (i=1, \dots, n_g)$ を求め、それらについて類似度 $\rho(f_i, g)$ を計算し、分類率を得たものである。これより、第1位分類率、つまり、識別率は平均98.27%となり、かつ候補文字として第4位までとれば、候補文字から対象文字カテゴリがもれるのは、わずか2文字となりほとんど完全に分類を行なえることがわかる。同じ図に、粗外形パターンと縦横比を用いた大分類をせずに、平均線数だけで分類を行なった結果を示す。図から、識別率平均96.13%となり、また、大分類を行なったときの第4位分類率を達成するためには、約50個の候補文字をとらなければならないことがわかる。

このように、外形パターン、縦横比と、平均線数が互いに相補的な特徴をとらえているために、前者を用いて大分類することにより、平均線数では近い類似度を持つ文字カテゴリをふるい落とすことができ、その結果、平均線数による分類効率が著しく向上させることができる。

6. むすび

中明朝体、太明朝体、中ゴシック体、および教科書体2字体の計5字体のマルチフォント印刷文字を対象に、二段階分類方式について検討した。まず、大分類では、 4×4 の粗いメッシュパターンと粗い外形パターンを作成

し、字体の違いによる変形の影響を受けにくくするためにコード化を行なった。分類実験を行なった結果、粗いメッシュパターンコードよりも粗外形パターンコードの方が、変形に対し安定で、分類効率が良いことがわかった。また、コード化することにより辞書の容量を1文字カテゴリ当り32bitと小さくでき、また、辞書との照合を簡単な論理演算だけで行なえるので高速な分類処理ができた。また、縦横比を併用することにより、平均候補文字数が約2文字減少することがわかった。

さらに、細分類を32次元の平均線数ベクトルを用いて、類似度を尺度として行なった。分類実験を行なった結果、識別率98.27%、第4位分類率99.96%を得た。また、粗外形パターンコードと平均線数が互いに相補的な特徴であり、両者を順次分類に使うことによって識別率が大きく向上することが確認された。

現在、各種教科書を対象とした分類実験を行なっており、本方式の有効性を検討している。

今後は、さらに対象とする字体、字種を増やし、教科書体を含むマルチフォント印刷文字認識システムの構成を目ざしていく予定である。

謝辞 日ごろ、有益な御討論頂く篠原正美氏に感謝致します。

文 献

- (1) 吉沢, 三宅, 田辺, 小宮: “マルチフォント認識における複合類似度法の効果について”, 昭57 信学総全大, 1329.
 - (2) 梅田三千雄: “マルチフォント印刷漢字認識のための粗分類”, 信学論(D), J62-D, 11, pp. 758-765 (昭54-11).
 - (3) 梅田三千雄: “マルチフォント印刷漢字の多段階分類”, 信学技報, PRL78-53(1978-11).
 - (4) 目黒, 梅田: “マルチフォント印刷漢字認識装置構成の検討”, 昭57 信学総全大, 1282.
 - (5) 萩田, 梅田, 増田: “三つの概形特徴を用いた手書き漢字の分類”, 信学論(D), J63-D, 12, pp. 1096-1102(昭55-12).
 - (6) 大田, 森, 坂井: “漢字パターン列の特徴パラメータによる検索”, 信学論(D), J64-D, 11, pp. 997-1004 (昭56-11).
 - (7) Kooi, R. and Lin, W. C.: “An On-Line Microcomputer-Based system for Reading Printed Text Aloud”, IEEE Trans. Syst. Man & Cybern., SMC-8, 1, pp. 57-62(Jan. 1978).
 - (8) 森, 青沼: “文字の構造(1)”, 信学技報, PRL73-85(1974-01).
 - (9) 坂井, 森: “漢字パターンの大分類”, 信学技報, PRL73-14(1973-05).
 - (10) 内藤, 淀川: “手書き漢字のストローク密度関数による分類”, 信学技報, PRL79-3(1979-04).
- (昭和57年5月10日受付, 11月5日再受付)