

複数の特徴量を用いた部分空間法による 手書きひらがなの識別

1L-4

島田恭宏 大倉 充 塩野 充 橋本禮治
岡山理科大学 工学部

1. はじめに

筆者らは、複数の特徴量を用いた識別手法の提案を行ってきた[1]。この手法は、Watanabeによって提案された部分空間法の1つであるCLAFIC法[2]に基礎を置くものである。まず、文字パターンから抽出された複数の特徴量を用いて特徴量毎に部分空間を作成し、入力パターンの各部分空間への射影量を求め、各カテゴリーに対して射影量の数の次元を有するベクトル(射影量ベクトル)を定める。識別は、得られた射影量ベクトルのノルムを計算し、最大の値を示すカテゴリーを求めることによって行われる。使用する特徴量は、大分類用に考案された5種類[3]を使用した。これまでに、類似漢字を対象とした小規模な実験で本手法の有効性を示してきた。そこで本稿では、この手法による電総研作成の手書き漢字データベースETL-9[4]に含まれるひらがな全71カテゴリーを用いて行った識別実験について報告する。

2. 実験に使用したデータ

実験には、電総研により作成されたJIS第一水準手書き漢字データベースETL-9(B)[4]よりひらがな71カテゴリーを用いた。識別実験では、1カテゴリー200サンプル中、前半100サンプルを学習サンプル、後半100サンプルを未知サンプルとして用いた。各サンプルの画面次数は63×64であるが、雑音除去の後、1次変換の拡大・縮小の式を用い、32×32としている。

3. 特徴量

3.1 粗いメッシュパターン

粗いメッシュパターン(M特徴)とは、漢字パターンを構成しているストロークの概形情報を表すものである。図1(a)の2値化された漢字パターンをより低次元のベクトル空間で表現するために、 $n \times n$ (本研究では $n=8$)の粗いメッシュ領域に分割する(図1(b))。そして、各メッシュ領域に含まれる文字部(画素値が1の部分)を計数することによって $n \times n (=64)$ 次元の特徴ベクトルを得る。

3.2 粗いペリフェラルパターン

粗いペリフェラルパターン(P特徴)とは、漢字パターンの持つ輪郭部分の特徴に着目し、構成ストロークの大まかな輪郭形状や、位置などの情報を面積に置き換えて抽出するものである。漢字パターンを図2に示すように n 分割($n=8$)し、一方の外接

枠から他方の外接枠に向かって最初の文字部に出会うまでの文字部でない領域(画素値が0の部分)の面積を計数する。この操作を4つの外接枠について行い、 $n \times 4$ 次元の粗い1次P特徴とする。同様に、最初の文字部との出合いまではそのまま計数を継続し、2度目の文字部に出会うまでの面積を計数し、これを $n \times 4$ 次元の2次のP特徴とする。本研究では、1次と2次の2つのP特徴を使用したため、次元数は $2 \times n \times 4 (=64)$ となる。

3.3 ストローク密度関数

ストローク密度(Stroke Density)とは、文字枠内の位置をパラメータとしてストロークの存在の度合を示す特徴である。ストローク密度関数(SDF特徴)は、ストロークの本数に着目することによって、漢字パターンの複雑さを反映する特徴量である。図3に示すように、漢字パターンを水平、垂直方向に走査し、ストロークを横切る回数を計数する。本研究では、用いた漢字パターンの画面次数が 32×32 であることより、得られる特徴量の次元数は、 $2 \times 32 (=64)$ となる。

3.4 局所ストローク密度

局所ストローク密度(LSD特徴)とは、M特徴とSDF特徴を併合し、発展させたものである。図4に示すように、 $n \times n (=4 \times 4)$ に分割された粗いメッシュ領域でXおよびY方向に投影される線密度の平均値を特徴量とするものである。本研究でのこの特徴量の次元数は $2 \times 4 \times 4 (=32)$ となる。

3.5 局所方向寄与度特徴

局所方向寄与度特徴(LDC特徴)とは、方向寄与度特徴を発展させたものである。方向寄与度特徴は、図5のように、文字内のある黒点Bにおいて、8方向に触手を延ばし、求められる黒点の連結長 l を用いて式(1)で定義される。

$$d_m = \frac{l_m + l_{m+4}}{\sqrt{\sum_{j=1}^4 (l_j + l_{j+4})^2}} \quad (1)$$

LDC特徴は、式(1)によって定義される方向寄与度特徴を文字パターンを構成するすべての黒点について求め、 $n \times n (=4 \times 4)$ に分割された粗いメッシュ領域に含まれる全黒点について平均を算出することにより抽出される。本研究でのこの特徴量の次元数は、 $4 \times 4 \times 4 (=64)$ となる。

4. 識別手法

あるカテゴリ $C^{(i)}$ に属する $n (=100)$ 個の学習パターンがあるとし、その第 j 番目のサンプルから抽出された k 番目の特徴量を $f_{jk}^{(i)}$ ($k=1, \dots, 5$) ($\|f_{jk}^{(i)}\|=1$) で表わす。この特徴量の次元数を $D(k)$ で表わす。 $C^{(i)}$ の k 番目の特徴量から作成する相関行列 $Kk^{(i)}$ を式(2)で定義する。

$$Kk^{(i)} = \frac{1}{n} \sum_{j=1}^n f_{jk}^{(i)} f_{jk}^{(i)T} \quad (2)$$

$Kk^{(i)}$ より固有値、固有ベクトルを計算し、固有値の大きなものから順に $d(k) (\leq D(k), n)$ 個までの固有ベクトルを採用し、部分空間を作成する。このため、各カテゴリに対して、特徴量の種類の数だけ部分空間が作成されることになる。カテゴリ決定は、入力パターンから抽出された k 番目の特徴量の k 番目の特徴量に対応した各部分空間への射影量 $P_{kd(k)}^{(i)}$ とし、式(3)で定義する射影量ベクトル $V^{(i)}$ のノルムが最大となるカテゴリを定めることで行われる。

$$V^{(i)} = (P_{1d}^{(i)}, P_{2d}^{(i)}, P_{3d}^{(i)}, P_{4d}^{(i)}, P_{5d}^{(i)})^T \quad (3)$$

5. 識別実験

実験では、棄却処理は行わず、全ての入力パターンに対して正読、誤読の判定を行った。実験はワークステーション SPARCserver670MP を用いて行った。

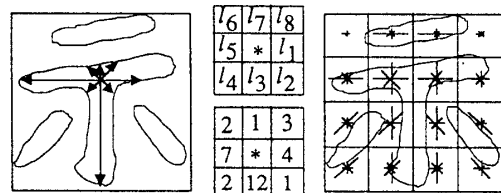
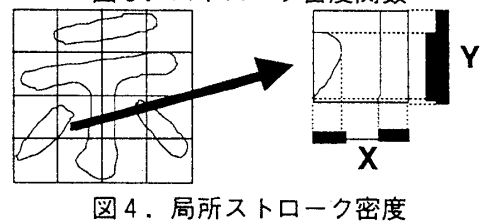
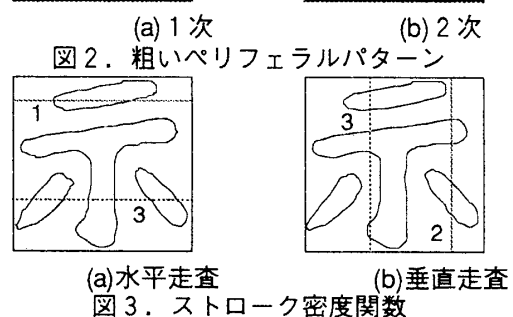
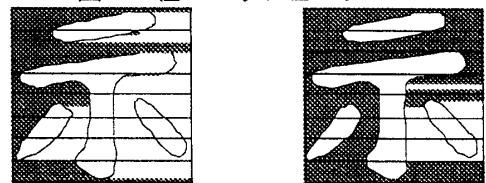
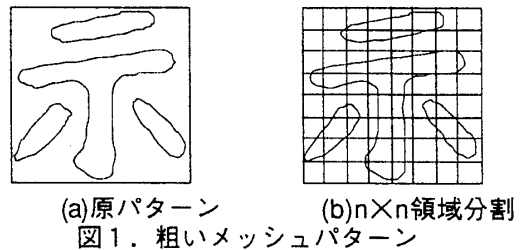
表1に本手法、および、単独の特徴量を用いた部分空間法による識別実験結果を示す。単独の特徴量を用いた部分空間法による実験では、採用する固有ベクトル(基底)数をパラメータとして、これを32まで採用した場合に得られた最高の平均識別率と採用基底数を示した。本手法による実験では、使用する特徴量の数と種類をパラメータとし、5種類の特徴量すべての組み合わせについて識別実験を行った。このときの基底数は、単独の特徴量を用いた部分空間法の実験結果より、最高の平均識別率が得られた基底数を採用した。その結果、最高の平均識別率は、SDF特徴、M特徴、P特徴、LDC特徴を用いた場合であり、93.69%を得た。特徴空間上では、1つのカテゴリに属するサンプル群は、ある代表的なサンプルの近傍に分布しているが、中心から遠くに位置するサンプル群もある。しかしそれらのサンプル群でも特徴空間を変えることによって、分布の中心近傍に位置するサンプルとすることができる。本手法で得られた識別率の上昇は、上記の事柄が識別に有効に作用したためと考えられる。しかし、本実験では5種類の特徴量を使用したにもかかわらず、最高の平均識別率は前出の4種類を使用した場合であり、特徴量の選択方法を考慮する必要がある。

6. おわりに

本稿では、複数の特徴量を用いた部分空間法による手書きひらがなの識別実験を行い、本手法のひらがなの識別性能について報告を行った。今後、採用する特徴量の選択方法について検討する予定である。

表1. 識別実験結果 (未知サンプル)

	本手法	単独の特徴量を用いた部分空間法				
		SDF特徴	M特徴	P特徴	LSD特徴	LDC特徴
平均識別率[%]	93.69	61.18	88.44	89.52	82.03	83.96
基底数	-	3	10	21	5	17



参考文献

[1] 島田恭宏, 大倉 充, 塩野 充, 橋本禮治: "大分類用の特徴量から生成された複数の部分空間による手書き類似漢字識別", 信学技報, PRU92-119(1993-01). [2] 橋本新一郎編著: "文字認識概論", オーム社 (昭57). [3] エルツキ・オヤ著 (小川英光, 佐藤 誠訳): "パターン認識と部分空間法", 産業図書 (昭61). [4] 斎藤泰一, 山田博三, 山本和彦: "手書き文字データベースの解析 (VII)", 電総研彙報, 49, 7, pp.487-525(1985).