

手書き文字認識における特徴量の次元数と変数変換に関する考察

正 員 若林 哲史<sup>†</sup>      正 員 鶴岡 信治<sup>††</sup>

正 員 木村 文隆<sup>†</sup>      正 員 三宅 康二<sup>†</sup>

On the Size and Variable Transformation of Feature Vector  
for Handwritten Character Recognition

Tetsushi WAKABAYASHI<sup>†</sup>, Shinji TSURUOKA<sup>††</sup>, Fumitaka KIMURA<sup>†</sup>  
and Yasuji MIYAKE<sup>†</sup>, Members

あらまし 統計的手法による文字認識の正読率を向上させるためには、各字種の分布の分離性が良い特徴量を用い、最適(ベイズ)識別関数になるべく近い識別関数を用いればよい。本論文では、(1)特徴量の分離性向上のために特徴量の次元数を増加し、(2)識別関数として、正規分布に対する擬似ベイズ識別関数を用い、(3)用いる特徴量を正規分布に近づけるための変数変換について検討し、これらの有効性を JIS 第 1 水準手書き漢字データベース ETL9B に対する認識実験により評価する。その結果、3036 字種を対象とした認識実験で、196 次元の加重方向指数ヒストグラムを用いた場合に、平均 97.74% (未学習文字) の高い正読率が得られた。

キーワード 文字認識, ベイズの決定則, 特徴抽出, 変数変換, 正規性

1. ま え が き

統計的手法による文字認識の正読率を向上させるためには、各字種の分布の分離性が良い特徴量を用い、最適(ベイズ)識別関数になるべく近い識別関数を用いればよい。特徴量の分離性を向上する直接的な方法として、特徴抽出の空間解像度を上げるなどして、特徴量の次元数を増加させる方法が考えられる。しかし、正規分布に対する最適識別関数としてよく知られている 2 次識別関数は、特徴量の次元数を増加させると、共分散行列の推定誤差のために、ある次元数を境に逆に正読率が低下する尖頭現象 (Peaking phenomenon) を起こすために、特徴量の次元数増加による分離性の向上が、必ずしも正読率の向上に寄与しない。この現象は、母集団の共分散を標本共分散で代用した 2 次識別関数が最適識別関数ではない(ベイズ決定でない)ために生じる現象であり、ベイズ決定においては、このような現象が生じないことが示されている<sup>(1)</sup>。

本論文では、(1)特徴量の分離性向上のために、特徴

量の次元数を増加し、(2)尖頭現象を避けるために、識別関数として、共分散行列が未知の正規分布に対するベイズ決定則から導かれる擬似ベイズ識別関数を用い、(3)用いる特徴量を正規分布に近づけるための変数変換について検討し、ETL9B(電総研作成の JIS 第 1 水準手書き漢字データベース: 3036 字種)を対象とする認識実験により、有効性を評価する。

ETL9B を対象とした認識実験では、斎藤らが多値ばかしパターン整合法により 82.93% の正読率を報告している<sup>(2)</sup>。その後津雲らが方向パターンマッチング法を用いて 95.42% を達成し<sup>(3)</sup>、更に文献(4)において 95.59% の正読率を報告している。また、郭らは整形変換により 95.66% とほぼ同等の結果を報告している<sup>(5)</sup>。

筆者らは前報<sup>(6)</sup>において加重方向指数ヒストグラム法と擬似ベイズ識別関数を用いる統計的認識手法を提案し、手書き教育漢字データベース ETL8 に対して高い正読率が得られることを報告した。本論文では特徴量の次元数増加と変数変換により JIS 第 1 水準手書き漢字データベース ETL9B の 3036 字種を対象とした認識実験で、196 次元の加重方向指数ヒストグラムを用いた場合に、平均 97.74% (未学習文字) の高い正読率が得られることを示す。

<sup>†</sup> 三重大学工学部情報工学科, 津市

<sup>††</sup> 三重大学工学部電気電子工学科, 津市

Faculty of Engineering, Mie University, Tsu-shi, 514 Japan

## 2. 認識方式の概略

認識処理の流れと構成を図1に示す。本研究では、文字ごとに切り出された2次元の2値パターン(横64×縦63ドットの文字データベース)を用いたため、文字切出しは行っていない。前処理部では、簡単なフィルタ処理によるノイズの除去と、文字の外接長方形を一定の大きさにする大きさの正規化(外接枠・重心合せ法<sup>6)</sup>)を行い、その後、大きさの正規化によって輪郭線に生じるのこぎり歯状ノイズを、3×3の平均値フィルタによって平滑化する。特徴抽出部では、文字輪郭線の抽出と方向指数化(チェーンコーディング)および方向指数ヒストグラムの計算等を行う。識別部では、字種ごとに擬似ベイズ識別関数の計算を行う。

大きさの正規化後の前処理および特徴抽出は、すべて完全並列処理が可能である。識別関数の計算も、内積計算を主体とするベクトル演算であるため、ベクトル化、パイプライン処理などハードウェア化による高速化が容易である。

## 3. ベイズの決定則と修正2次識別関数

本研究ではベイズ識別関数に近い識別関数(擬似ベイズ識別関数)として次に示す修正2次識別関数(MQDF: Modified Quadratic Discriminant Function)を用いて識別を行った。

$$g(X) = \frac{1}{k^2} \left[ \|X - M\|^2 - \sum_{i=1}^k \frac{\lambda_i}{\lambda_i + k^2} \{\Phi_i^T(X - M)\}^2 \right] + \ln \left\{ \prod_{i=1}^n \lambda_i + k^2 \right\} - 2 \ln P(\omega) \quad (1)$$

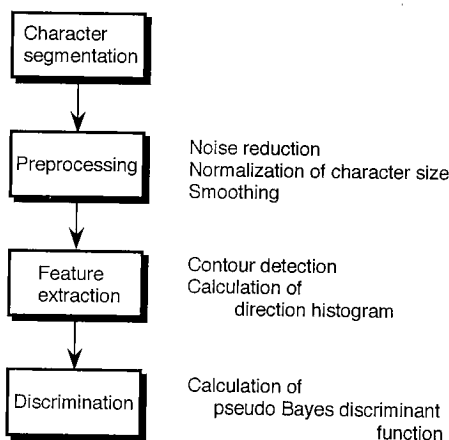


図1 認識処理のフロー

Fig. 1 Flowchart of character recognition process.

ここで、 $M$ は標本平均ベクトル、 $\lambda_i$ 、 $\Phi_i$ は標本共分散行列 $\Sigma$ の第 $i$ 固有値、固有ベクトルであり、 $k^2$ は後述する定数である。また、 $k$ は主要固有ベクトル(dominant eigen vector)の数で、 $k^2$ の値と共に実験的に決める。

以下にはベイズの決定則からMQDFを導出する概略を示す<sup>7)</sup>。

### 3.1 既知の正規分布に対するベイズ識別関数

次式で与えられる2次識別関数(QDF)は、パラメータが既知の正規分布に対するベイズ識別関数としてよく知られている。

$$g_l(X) = (X - M_l)^T \Sigma_l^{-1} (X - M_l) + \ln |\Sigma_l| - 2 \ln P(\omega_l) \quad (2)$$

ここで、 $M_l$ 、 $\Sigma_l$ は第 $l$ クラスの平均ベクトルと共分散行列、 $P(\omega_l)$ はクラス $\omega_l$ の事前確率である。以後クラスに関する添字 $l$ は省略する。現実のパターン認識の問題では、平均、共分散行列が未知のため、代わりに標本平均、標本共分散行列を用いるのが一般的である。しかし、この2次識別関数が実際に文字認識に用いられることはほとんどない。文字認識などの工学的パターン認識では数十から百以上の特徴量を用いることが多く、パラメータ、特に共分散の推定誤差による認識性能の低下が著しいからである。また、QDFが必要とする計算量が多いため、十分な文字読取り速度を達成するのが困難である。

### 3.2 共分散行列が未知の正規分布に対するベイズ識別関数

共分散行列の推定誤差に起因する問題は、共分散行列が未知の正規分布に対するベイズ識別関数を用いることで解決できるはずである。このベイズ識別関数は次式で与えられる<sup>7)</sup>。

$$g(X) = (N_N + n - 1) \ln \left[ 1 + \frac{(X - M)^T \Sigma_N^{-1} (X - M)}{(N_N - 1)} \right] + \ln |\Sigma_N| - 2 \ln P(\omega) \quad (3)$$

但し、

$$N_N = N + N_0$$

$$\Sigma_N = (1 - \alpha) \Sigma + \alpha \Sigma_0$$

$$\alpha = \frac{N_0}{N_N}$$

ここで、 $n$ は特徴量の次元数、 $\Sigma$ は標本共分散、 $\Sigma_0$ は真の共分散行列の初期推定量で、 $N_0$ は $\Sigma_0$ の信頼度定数(confidence constant)である。

### 3.3 修正2次識別関数(MQDF)

$N_N$ が次元数 $n$ より十分に大きい場合、式(3)は以下

のように簡略化できる。

$$g(X) = (X - M)^T \Sigma_N^{-1} (X - M) + \ln |\Sigma_N| - 2 \ln P(\omega) \quad (4)$$

これは共分散行列を除いては、2次識別関数と同じ形になる。実際に式(4)を識別関数として用いるためには、 $N_0$ と $\Sigma_0$ の値を決めて $\Sigma_N$ の値を求めなければならない。 $N_0=0$ とすると $\Sigma_0$ の値は不要となり、一見好都合であるが、こうすると、この識別関数はQDFとほとんど同じものになり、QDFの欠点が是正されない。我々は、 $\Sigma_0 \neq 0$ であることを0でない信頼度で信じていることができるため、 $N_0=0$ という仮定は妥当ではないと考えられる。この点を考慮して、 $\Sigma_0 = \sigma^2 I$ 、 $N_0 < N$ を仮定すると、擬似ベイズ推定

$$\Sigma_N = \Sigma + \alpha \sigma^2 I = \Sigma + h^2 I \quad (5)$$

が得られる。

ここで、

$$h^2 = \alpha \sigma^2 = \frac{N_0}{N + N_0} \sigma^2$$

式(4)と式(5)より、

$$g(X) = (X - M)^T [\Sigma + h^2 I]^{-1} (X - M) + \ln |\Sigma_N| - 2 \ln P(\omega) \\ = \sum_{i=1}^n \frac{[\Phi_i^T (X - M)]^2}{\lambda_i + h^2} + \ln |\Sigma_N| - 2 \ln P(\omega) \quad (6)$$

$i > k$ の場合に  $h^2 \gg \lambda_i$  を仮定すると、式(6)は次式で近似される。

$$g(X) = \sum_{i=1}^k \frac{[\Phi_i^T (X - M)]^2}{\lambda_i + h^2} + \sum_{i=k+1}^n \frac{[\Phi_i^T (X - M)]^2}{h^2} \\ + \ln \left[ \prod_{i=1}^n (\lambda_i + h^2) \right] - 2 \ln P(\omega) \quad (7)$$

以下の関係、

$$\sum_{i=k+1}^n [\Phi_i^T (X - M)]^2 = \|X - M\|^2 - \sum_{i=1}^k [\Phi_i^T (X - M)]^2 \quad (8)$$

を用いれば式(1)で与えられるMQDFが導かれる。MQDFはベイズ識別関数に近いという意味で擬似ベイズ識別関数の一つである。また、MQDFは、共分散行列の推定誤差に起因する性能低下が少なく、計算量・記憶容量もQDFの1~3割程度で、文字認識等への応用に適している<sup>(8)-(11)</sup>。

## 4. 特徴抽出

### 4.1 抽出アルゴリズム

以下に64次元の加重方向指数ヒストグラム特徴を抽出する手順を示す。

(1) 前処理後の2値文字パターンに対して文字部分の輪郭線を8連結で追跡(輪郭線を左回り、ホール部を右回り)し、各輪郭線の方向を45°おきの4方向に量子化する(図2)。

(2) 文字図形を縦7×横7の49個の小領域に分割し、それぞれの小領域内で各方向ごとにヒストグラムを求め、196次元(縦7×横7×4方向)の方向指数ヒストグラムを得る。

(3) この方向指数ヒストグラムに重なりのある2次元ガウスフィルタ(5×5)(図3)を掛けて、64次元(縦4×横4×4方向)に次元圧縮して得られたものを加重方向指数ヒストグラムと呼ぶ。

### 4.2 変数変換

MQDFは、観測値の分布が多変量正規分布であるという仮定のもとでベイズの決定則から導かれる。従って観測値が多変量正規分布でない場合には、その性能

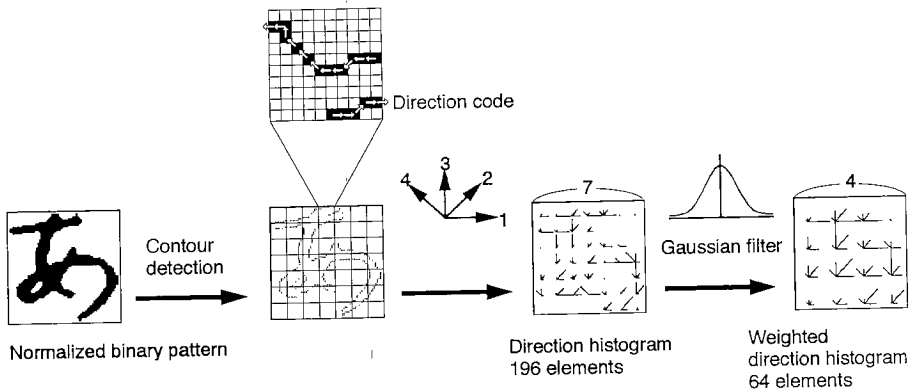


図2 特徴抽出  
Fig. 2 Feature extraction.

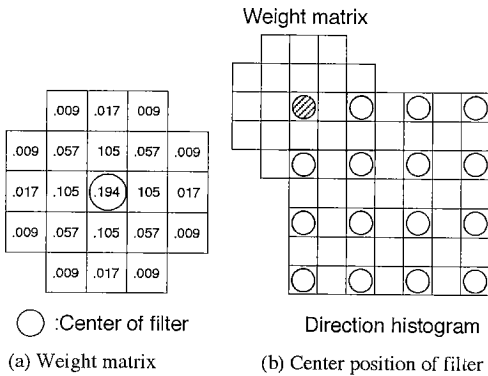


図3 ガウスフィルタ  
Fig. 3 Gaussian filter.

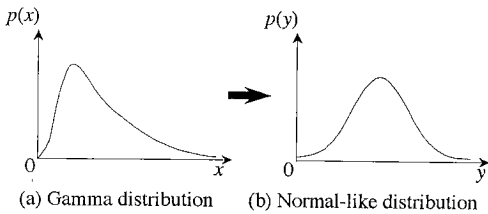


図4 変数変換  
Fig. 4 Variable transformation.

は低下してしまう。このような場合には、データを変換し、“正規分布らしく見える”ようにする変数変換の処理が有効である。

一般的に、カウントすることにより観測された正の値のみからなるヒストグラムは、図4(a)のようにガンマ分布となることが多い。このとき、次式のようなべき変換によって観測値を変換すると、平均の周りの対称性と正規分布への近似がしばしば顕著に改善される(図4(b))。

$$y = x^u \quad (u > 0) \tag{9}$$

ここで、 $x$  は変換される値で、 $u$  はべき乗を示す。文献(12)には  $u=0.4$  のときにガンマ分布が最も正規分布に近くなることが示されている。

一方、対数変換  $y = \ln(x)$  によっても正規分布に近似されることが知られているが、加重方向指数ヒストグラム特徴には  $[0, 1]$  の値がしばしば見られ、変換によって大きな負のすそ野 (long negative tail) が発生するため望ましくない。この欠点を回避するため、次式を用いる。

$$y = \ln(1+x) \tag{10}$$

今簡単のために、変数ごとに式(9)、式(10)の変数変換を行えば、それぞれの周辺分布をおおよそ正規分布

にすることができる。

$$y_i = x_i^u \tag{11}$$

$$y_i = \ln(1+x_i) \tag{12}$$

$$(i=1, 2, \dots, n)$$

ここで、 $n$  は特徴ベクトルの次元数を表す。

正規周辺分布は同時分布が多変量正規分布であることを保証するには十分ではないが、実用的な応用では十分と考えられる。

### 4.3 ひずみ度、尖度による正規性の検定

分布が正規分布からどの程度逸脱しているかを見る統計量にひずみ度と尖度がある。この統計量は直感的にわかりやすいため、加重方向指数ヒストグラム特徴がどの程度多変量正規分布から逸脱しているか、また変数変換により正規性が改善されたかどうかを評価する基準量となる。

今、 $N$  個のデータの平均値を  $\bar{x}$ 、標準偏差を  $S_n$  とすると、分布の非対称度を示すひずみ度  $b_1$  は、

$$b_1 = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{S_n} \right)^3 \tag{13}$$

分布のすそ野の長さを示す尖度  $b_2$  は、

$$b_2 = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{S_n} \right)^4 \tag{14}$$

である。ここでひずみ度  $b_1$  は  $b_1=0$  のとき正規分布、 $b_1 > 0$  のとき右すそ広がり分布、 $b_1 < 0$  のとき左すそ広がり分布を示している。尖度は  $b_2=3$  のとき正規分布、 $b_2 > 3$  のときは  $t$  分布などすそ広がり強い分布、 $b_2 < 3$  のときはすそが途切れた分布を表している。このひずみ度と尖度を加重方向指数ヒストグラム特徴の各変数ごとに計算し、周辺分布の正規性を検定すれば、実用上同時分布の正規性が検定されたとみなしてもよいと思われる。

## 5. 認識実験

### 5.1 使用データ

本研究で使用したデータは、工業技術院電子技術総合研究所において、「JIS 第1水準手書き漢字データベース ETL9」<sup>(13)</sup> を横64×縦63の点で2値化した「ETL9B」である。

ETL9Bには、JIS 第1水準漢字2965字種に平仮名71字種を加えた3036字種が各字種200文字ずつ収められている。そこで、3036字種を各字種1文字ずつをまとめて、1セットとし(計200セット)、10セットをまとめて1巻とする(計20巻)。以後、個々のデータを巻番号とセット番号で区別する。また、全20巻をA、B、

C, D, Eの五つのブロックに分ける(表1).

### 5.2 特徴の正規性に関する考察

#### 5.2.1 変数変換と正規性の検定

ETL9Bのすべてのセットの各字種に対し、4.1を用いて64次元の加重方向指数ヒストグラム特徴を求める。得られた特徴ベクトルの各変数に対し、式(11)の変数変換を行う。 $u_i$ はすべての変数に共通で $u_i=0.3, 0.5, 0.7$ について計算した。 $u_i$ は変数ごとに変える方法もあるが<sup>(16)</sup>、予備実験の結果では一定の方が良い結果が得られた。変換前と変換後の各変数のひずみ度・尖度を計算した(図5)。図5(a)~(d)は平仮名「あ」の、(e)~(h)は漢字「綾」のひずみ度・尖度のべき乗値 $u_i$ に対する変化を表している。横軸は特徴量の各変数である。0付近と3付近の破線は、それぞれひずみ度( $b_1=0$ で正規分布)と尖度( $b_2=3$ で正規分布)を正規分布と仮定したときの有意水準5%を意味している。つま

り、この範囲外であれば正規分布とは言えないが、範囲内では正規分布の可能性を否定できないと言える。(a),(e)に注目すると変数変換なしの場合には「あ」も「綾」もひずみ度が $b_1>0$ の側に寄っており、右すその長いガンマ分布に近い分布状態である。また、尖度は平仮名である「あ」の方が漢字「綾」よりも正規分布から大きくずれており、すその広い分布となっている。これらの傾向は他の平仮名・漢字にも同様に見られ、平仮名が漢字よりもすが広く正規分布からずれる傾向があることを示している。また、変数変換においては「あ」、「綾」ともに $u_i=0.5$ のときに最も正規性が高くなっている。一方、図6は漢字「綾」の196次元特徴量のひずみ度・尖度で、同図(a)は変数変換なし、同図(b)はべき変換( $u_i=0.5$ )の場合をそれぞれ表している。これにより196次元においてもべき変換が正規性の向上に有効に働いていることがわかる。また、図5(e),(g)と比較すると196次元の方が正規性が悪くなっている。196次元の特徴量を求める際には64次元よりも小領域の大きさが小さく、小領域に含まれる輪郭線が少なくなる。そのため、各要素の値が64次元に比べて小さい。また、平仮名は漢字よりも輪郭線が少ない。このことから、漢字よりも平仮名のような簡単な文字ほど、また次元数を上げるほど正規性が悪くなるのは、小領域あたりの輪郭画素の減少が影響してい

表1 ETL9Bのブロック化

ブロック	巻	セット
A	1, 2, 19, 20	1~20, 181~200
B	3, 4, 5, 6	21~60
C	7, 8, 9, 10	61~100
D	11, 12, 13, 14	101~140
E	15, 16, 17, 18	141~180

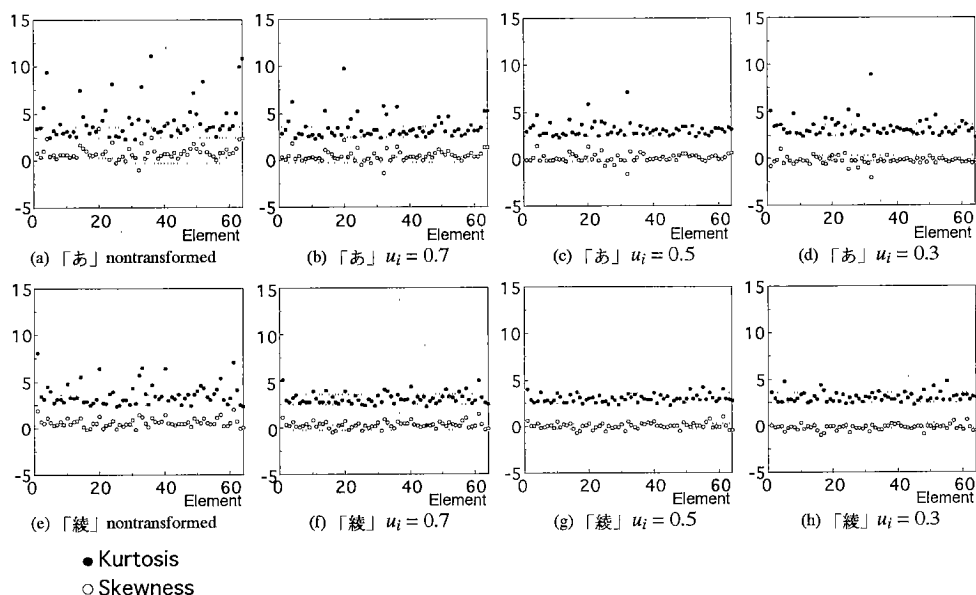
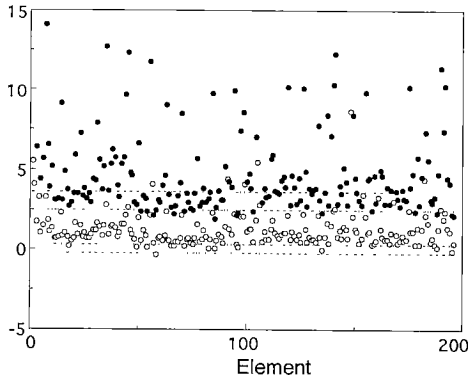
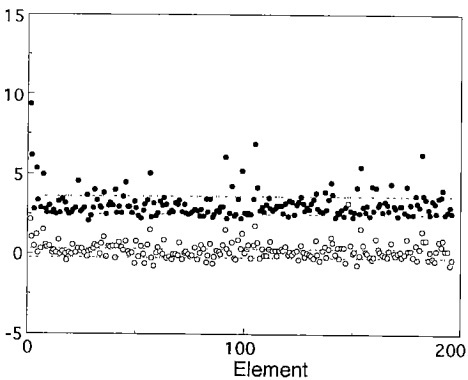


図5 変数変換とひずみ度・尖度

Fig. 5 Skewness and Kurtosis of transformed/nontransformed variables.



(a) 「綾」 nontransformed



(b) 「綾」  $u_i = 0.5$

- Kurtosis
- Skewness

図6 196次元特徴量のひずみ度・尖度

Fig. 6 Skewness and Kurtosis of 196-dimensional feature vector.

るものと思われる。更に、平仮名の場合には輪郭線の位置ずれのために特徴量の変動が大きいことも関与している可能性がある。

### 5.2.2 変数変換による正読率の向上

次にこれらの特徴量を用いて認識実験を行う。表1のETL9BブロックB, C, D, Eの計16巻, 160セット(160文字/字種)を用いて特徴ベクトルの平均値および共分散行列の固有値・固有ベクトルを求め、これを辞書とする。そして、ブロックAの計4巻(40文字/字種)をテストサンプルとして認識実験を行う。MQDF(式(1))の定数は、 $h^2 = \{N_0 / (N + N_0)\} \sigma^2 = \alpha \sigma^2$ とし、 $\sigma^2$ には全字種全固有値の平均を用いる。但し $\alpha$ は経験的に最もよい結果をもたらす値を選んだ。識別に用いる固有ベクトルは20軸( $k=20$ )とし、式(12)による対数変換についても実験した。

表2 変数変換による正読率の向上

	従来方法	$u_i = 0.7$	$u_i = 0.5$	$u_i = 0.3$	$\ln(1+x)$
$\alpha$ (最適値)	0.45	0.55	0.55	0.55	0.5
正読率 (%)	92.38	94.05	94.48	94.03	94.31

表3 ETL9B先頭300字種での実験

	従来	$\chi^{0.5}$	両方で 誤読 (内数)
正読率 (%)	97.48	98.20	
ひらがな→漢字 (個)	14	3	1
漢字→ひらがな (個)	2	4	2

入力文字の特徴ベクトルに対して、MQDFを用いて相違度を求め、最小の相違度の字種が第1位の候補字種となるが、これをすべての字種に対して用いると膨大な計算量となる。そこでシティーブロック距離により大分類を行い、候補字種(100字種)を限定し、この候補字種に対して、MQDFを用いて詳細識別を行った(表2)。

ひずみ度・尖度による検定で最も正規性が高かった $u_i = 0.5$ のときに正読率が最も高くなり、変数変換による分布の正規性の向上が正読率を高めることがわかった。また対数変換でもべき変換とほぼ同等の効果があらることがわかった。

### 5.2.3 変数変換による決定境界の変化

5.2.2で誤読文字に関して検討してみると、変数変換の有無にかかわらず、類似文字による誤読が最も多い。しかし、変数変換なしの場合に、平仮名を複雑な漢字に誤読する現象がしばしば見られたのが、変数変換により減少している。このことを詳しく調べるためにETL9Bの先頭300字種(うち平仮名71字種)を用いて認識実験を行った(表3)。ETL9BブロックB, C, D, E(160文字/字種)を学習に、ブロックA(40文字/字種)をテストに用いている。特徴ベクトルは64次元、固有ベクトルは20軸用いた。平仮名を漢字に誤読していたが、変数変換により正読となった例を図7に示す。

べき変換( $u_i = 0.5$ )を行うことにより平仮名を漢字に誤読する例は大幅に減少する。逆に漢字を平仮名に誤読する例はあまり増加しない。しかも、漢字を平仮名に誤読したのは「乙」のようなもともと画数の少ない漢字ばかりである。図8は横軸に特徴ベクトルの第10

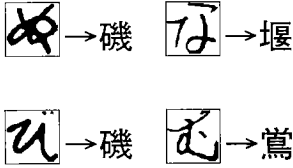


図7 平仮名を漢字に誤読した例

Fig. 7 Examples of confusion between "HIRAGANA" and "KANJI".

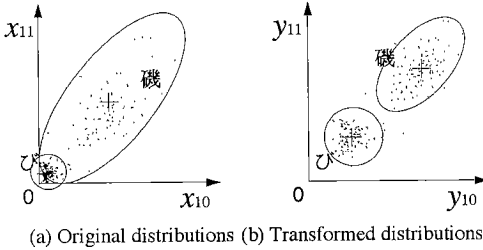


図8 変数変換による決定境界の変化

Fig. 8 Improvement of separability by variable transformation.

要素, 縦軸に第11要素をとった場合のべき変換による分布の変化を表している。同図(a)の「磯」のように非対称に分布している場合, 分布のピークが平均よりも原点寄りにある。そのまま正規分布で近似すると, 分布が原点側に伸びたように解釈され, ベイズの決定境界が「ひ」の方にずれ込んでしまう。このため, 「ひ」を「磯」に間違えやすくなる。しかし, べき変換によりこの現象は解消されている。変数変換はこのように分布の平均の周りの対称性と正規分布への近似を改善するため, MQDFの分離性能を高める作用をもっている。

5.3 次元数に関する考察

5.3.1 次元数と正読率の関係

64次元の特徴量を用いた認識実験で一番の誤読の原因である類似字種に対する対策として, 特徴量の次元数増加が考えられる。対象とする字種数が3036と多いのに対して, 用いた特徴量が64次元では正しく識別するのに不十分であることが予想される。従って次元数を増加させ, より細かな特徴をとらえることにより, 識別関数の特徴量に対する分離性能が向上すると予想される。そこでETL9Bの3036字種に対して4.1で分割数を7×7, 8×8, 9×9, 10×10としてそれぞれ64次元, 100次元, 144次元, 196次元の加重方向指数ヒストグラム特徴を作成し, 認識実験を行った。また識別に用いる固有ベクトルは40軸(k=40)とし, パラメー

表4 ETL9B 3036字種正読率(変数変換なし)

次元数	64	100	144	196
$\alpha$ (最適値)	0.15	0.3	0.5	0.7
ローテーション法 全巻平均(%)	92.46	95.25	96.35	96.83

表5 ETL9B 3036字種正読率(べき変換  $u_i=0.5$ )

次元数	64	100	144	196
$\alpha$ (最適値)	0.25	0.4	0.7	0.9
ローテーション法 全巻平均(%)	94.36	96.54	97.36	97.74

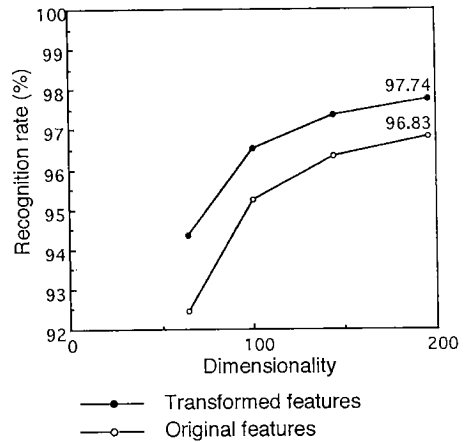


図9 全セットの平均正読率

Fig. 9 Average recognition rate over all sets.

タ  $\alpha$  は経験的に最適となるものを選んだ。変数変換なしの場合の結果を表4に,  $u_i=0.5$ のべき変換の場合の結果を表5に示す。なお, 表にはローテーション法によりA, B, C, D, E全ブロックをテストサンプルとして実験し平均をとった結果が示されている。ローテーション法とは, ブロックAをテストするときにはそれ以外のB, C, D, Eで学習(各字種160文字)を行い, ブロックB, C, D, Eについても同様にそれ以外のブロックを学習サンプルとする方法である。

図9は全ブロック全セットに対するテスト結果の平均をグラフにしたものである。この図から, 次元数が増加するほど正読率がよくなっていることがわかる。また, どの次元でもべき変換( $u_i=0.5$ )により正読率が向上しているが, これはエラー率にするとほぼ30%の減少である。

処理時間を表6に示す。使用した計算機は米国サンマイクロシステムズ社製のワークステーション

表6 処理時間 (ms/文字)

次元数	64	100	144	196
前処理 特徴抽出	81	71	77	77
識別	252	500	688	946

表7 64次元の認識で誤読の多い字種

誤読回数	テスト字種	第1候補字種
18	間	間 闊 閱 闊 潤 間
15	閣	閣 閣 閑 閱 朗
14	業	業 菓 粟 業 栗 某
	東	東 案 栗 車
13	線	線 練 絞 績 腺 綿
	撤	撤 欲 赦 轍 搬
12	粟	粟 統 総 既 絞 縮 粧 疏
	鳥	鳥 異 奥 果 菜 爽
	栗	栗 案 采 票 菓 東 葉
	鳥	鳥 息 鳥
	間	間 陶 聞

表8 196次元の認識で誤読の多い字種

誤読回数	テスト字種	第1候補字種
15	鳥	鳥
13	鳥	鳥 息
10	間	間
9	采	采 菜
	突	突 安 穴 麦
	采	采 禾 米 来

Sparcstation 2 (28.5 MIPS) である。特徴抽出に比べ識別に時間がかかるが、並列計算が可能であり、容易に高速化できる。

### 5.3.2 誤読文字例

表7は64次元・べき変換 ( $u_i=0.5$ ) でブロック A の40セット分の認識実験を行った結果、40個/字種中どのような字種が多く誤読したかを示している。第1候補となる個数が多いものほど左側になるように並べてある。表8は196次元の場合に誤読の多い字種である。表7からわかるように、誤読を起こしやすい字種は形が似ているものや同一の部首をもつものが多い。次元数を上げると類似字種の誤読は減るが、依然として門がまえの字種は誤読回数が多い。ただ、誤読する候補字種は特定のものに限定されている。

64次元で誤読し、196次元では正読した文字「統」の特徴ベクトルと、64次元において第1候補字種となっ

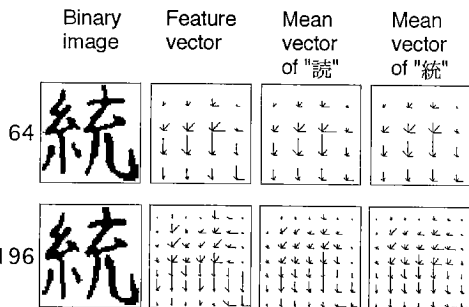


図10 64次元の特徴ベクトルと196次元の特徴ベクトル  
Fig. 10 64- and 196-dimensional feature vectors.

た「統」の平均特徴ベクトルを図10に示す。図に示すように196次元では細かな特徴までとらえている。

## 6. むすび

本研究では、特徴量の分離性向上のために特徴量の次元数を増加し、用いる特徴量を正規分布に近づけるための変数変換について検討した。そして、ETL9Bを用いた計算機シミュレーションにより、以下の結果と知見が得られた。

- (1) 変数変換 (0.5 乗) により特徴量の正規性が向上することを、ひずみ度・尖度を用いた検定で確認した。
- (2) 変数変換によってエラー率が約 30% 減少し正読率が向上した。
- (3) 正規分布で近似した場合の分布の分離性の向上により、平仮名を漢字に誤読する現象が発生しにくくなった。

(4) 特徴量の次元数増加により正読率が向上し、196次元の加重方向指数ヒストグラムを用いた場合に、平均 97.74% の高い正読率が得られた。

(5) 漢字より平仮名のような画数の少ない文字ほど、また特徴量の次元数を上げるほど、特徴量の正規性が悪化することがわかった。

更に正読率を向上させるために、(1)次元数増加と正規性の悪化の関係の解析と解決、(2)学習サンプル数の増加、(3)加重方向指数ヒストグラムと他の特徴量の併用等が考えられるが、これらについては今後検討する予定である。

謝辞 本研究に関し、手書き漢字データベース ETL9B を作成・提供して下さった電総研の皆様、討論をして頂いた研究室の皆様へ深謝致します。

## 文 献

- (1) Van Campenhout J. M. : "On the peaking of Hughes



- mean recognition accuracy : The resolution of an apparent paradox", IEEE Trans. Syst., Man. & Cybern., **SMC-8**, 5, pp. 390-395 (May 1978).
- (2) 斎藤泰一, 山田博三, 山本和彦 : "JIS 題 1 水準手書き漢字データベース ETL9 とその解析", 信学論(D), **J68-D**, 4, pp. 757-764 (1985-04).
  - (3) 津雲 淳 : "方向パターンマッチング法の改良と手書き漢字認識への応用", 信学技報, **PRU90-20** (1990-06).
  - (4) 津雲 淳 : "パターンマッチングに基づいた手書き漢字認識", 信学 '92 春大, D549.
  - (5) 郭 軍, 孫 寧, 根元義章, 佐藤利三郎 : "整形変換を用いた手書き文字データベース ETL9B の高精度認識", 信学技報, **PRU92-37** (1992-10).
  - (6) 鶴岡信治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二 : "加重方向指数ヒストグラム法による手書き漢字・ひらがな認識", 信学論(D), **J70-D**, 7, pp. 1390-1397 (1987-07).
  - (7) Kimura F. and Shridhar M. : "Handwritten Numeral Recognition based on Multiple Algorithms", Pattern Recognition, **24**, 10, pp. 969-983 (1991).
  - (8) Kimura F., Takashina K., Tsuruoka S. and Miyake Y. : "Modified Quadratic Discriminant Function and the Application to Chinese Character Recognition", IEEE Trans. of Pattern Anal. & Mach. Intell., **PAMI-9**, 1 (1987-01).
  - (9) 酒巻 久, 安達宗郎, 木村善将, 久貝正己, 上野修五郎, 矢ヶ崎敏明 : "認識速度 70 字/秒の日本語の OCR, 専用 LSI と DSP のパイプライン処理で高速化", 日経 Electronics 1990.7.23 (1990-07).
  - (10) 中川聖一, 榎本康広 : "時系列パターンの統計的認識法を用いた不特定話者の音韻および単語音声の認識", 信学論(D), **J71-D**, 10, pp. 1977-1983 (1988-10).
  - (11) 浜本義彦, 内村俊二, 金岡泰保, 富田真吾 : "標本数が少ない状況下における識別器の評価", 信学技報, **PRU92-91** (1992-12).
  - (12) Fukunaga K. : "Introduction to Statistical Pattern Recognition, Second Edition", ACADEMIC PRESS (1990).
  - (13) 斎藤泰一, 山田博三, 山本和彦 : "手書き漢字データベースの解析 (VIII)", 電総研彙報 **49**, 7, pp. 487-525 (1985-07).
  - (14) 谷口 修, 鶴岡信治, 木村文隆, 三宅康二 : "加重方向指数ヒストグラム法による JIS 第 1 水準手書き漢字の認識", 信学技報, **PRU88-152** (1989-03).
  - (15) 矢田勝啓, 鶴岡信治, 木村文隆, 三宅康二 : "加重方向指数ヒストグラム法のつぶれ文字への対応", 信学技報, **PRU90-128** (1991-02).
  - (16) Johnson R. A., Wichern D. W., 西田俊夫訳 : "多変量解析の徹底研究", 現代数学社 (1988).
  - (17) 市原清志 : "バイオサイエンスの統計学", 南江堂 (1989). (平成 5 年 4 月 30 日受付)



若林 哲史

昭 60 三重大・工・電子卒, 昭 62 同大大学院修士課程了。同年ダイキン工業株式会社入社。平 2 三重大・工・助手, 現在に至る。手書き文字認識, 画像処理, コンピュータグラフィックスに関する研究に従事。



鶴岡 信治

昭 52 岐阜大・工・電子卒, 昭 54 名大大学院博士(前期)課程了。同年三重大・工・電子助手, 平 1 同大・工・助教授, 現在に至る。この間, 平 3~4 米国ミシガン大デアボーン校客員助教授。工博。手書き文字認識, 文書理解, コンピュータグラフィックス, 医用画像処理に関する研究に従事。情報処理学会, 日本 ME 学会, 人工知能学会各会員。



木村 文隆

昭 48 名大・工・電気卒, 昭 53 同大大学院博士課程了。同年同大・工・助手, 昭 58 三重大・工・助教授, 現在に至る。その間平成 1 年からミシガン大学客員助教授を勤める。文字・パターン認識, 画像処理, コンピュータグラフィックスの研究に従事。情報処理, ME, 人工知能学会各会員。工博。



三宅 康二

昭 35 名大・工・電気卒, 昭 40 同大大学院博士課程了。同年同大・工・助手, 昭 43 同講師, 昭 44 同助教授, 昭 53 三重大・工・教授, 現在に至る。現在, 手書き文字の機械認識, コンピュータグラフィックスおよび医用画像処理の研究に従事。情報処理学会, 日本 ME 学会, 日本人工知能学会, 日本ロボット学会, 電気学会, 日本シミュレーション学会等各会員。工博。