

少クラス分類問題における特徴選択に関する考察

若林 哲史<sup>†</sup> 鶴岡 信治<sup>††</sup> 木村 文隆<sup>†</sup> 三宅 康二<sup>†</sup>

A Study on Feature Selection for Small Class Classification Problems

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

あらまし 正準判別分析は分散比 (F比) を最大化する最も代表的な特徴選択手法であるが, クラス数以上の特徴が選択できないため, 少クラス分類問題に対する有効性に限界がある. この問題を解決するために, 新しい特徴選択手法 (FKL法) を提案し, 手書き数字認識実験によりその有効性を評価する. FKL法は, F比を最大化する正準判別分析と, 次元減少による平均2乗誤差を最小化する主成分分析 (K-L展開) を特殊な場合として含む, より一般的な特徴選択手法である. 正準判別分析, 主成分分析, 正規直交判別ベクトル法 (ODV法) などとの比較実験の結果, 少クラス分類問題ではFKL法により選択された特徴量の識別力が最も高いことを示す.

キーワード 特徴選択, 特徴抽出, 正準判別分析, 主成分分析, K-L展開

1. ま え が き

正準判別分析は分散比 (F比) を最大化する最も代表的な特徴選択手法であるが, クラス数以上の特徴が選択できないため, 少クラス分類問題に対する有効性に限界がある. 例えば字種数の少ない数字認識では識別力に関して順序づけられる特徴量が9個 (字種数-1) となるため, 識別力のより高い10次元以上の特徴ベクトルを得ることができない. この問題を解決するために, 多くの方法が研究されている [1]. 代表的な方法として, 正規直交条件のもとでF比を最大化する方法 [2]~[6], 共分散行列の差異を表す項を追加して特徴評価関数を拡張する方法 [7]~[10], 共分散行列をノンパラメトリック化する方法 [11] などがある.

本論文では, もう一つの新しい特徴選択手法 (Fisher-Karhunen-Loève Method, 以下FKL法) を提案し, 手書き数字認識実験によりその有効性を評価する. FKL法は, F比を最大化する正準判別分析と, 次元減少による平均2乗誤差を最小化する主成分分析 (K-L展開) を特殊な場合として含む, より一般的な特徴選択手法である. 比較のための特徴選択手法としては, 正準判別分析, 主成分分析, 正規直交判別ベクトル法 (Orthonormal Discriminant Vector Method, 以下ODV

法) [3], [5], [6]を用い, 識別には擬似ベイズ識別関数, 2次識別関数, 部分空間法を用いる.

ODV法に代表される, 正規直交条件下でF比を最大化する手法は, 正準判別分析よりF比の大きい特徴を選択することができ, 多クラス分類問題にも適用できるが, 得られる特徴に相関があるため多変量の特徴ベクトルとして識別力が高いとは限らない.

特徴評価関数を拡張する従来の方法は, 基本的に2クラス分類問題を対象とし, 二つの共分散行列の分散差あるいは分散比の項を追加して特徴評価関数を拡張するため, 多クラス問題への拡張が容易でない [12].

共分散行列のノンパラメトリック化は, クラスの代表ベクトルを複数化して級間分散のランク落ちを防ぐ手法であり, 個々のクラスの分布形が複雑でノンパラメトリックな識別系を用いる場合の特徴選択手法としては合理的であるが, 分布形がパラメトリックに表現できる場合や, パラメトリックな識別系を用いる場合に適した手法ではない. また, 分布形が複雑でノンパラメトリックアプローチが適している場合には, FKL法をノンパラメトリック化することもできる.

FKL法は, 特徴評価関数を拡張して正準判別分析と主成分分析を融合する新しい特徴選択手法であり,

(1) 2クラス問題, 多クラス問題のいずれにも区別なく適用できる,

(2) 原特徴量の次元数を超えない範囲で任意の次元数の特徴ベクトルを得ることができる,

<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

(3) 固有値解析問題として数値計算による効率の良い解法が確立している、などの性質を備えている。

判別分析と主成分分析との融合を試みた手法として文献[4]がある。この手法は、正規直交条件を満たす基底ベクトルの各々においてF比あるいは分散のいずれか一方を選択して最大化する手法で、ODV法を拡張したものと考えることができる。しかし、各基底ベクトルにおけるF比と分散の選択が利用者にまかされているため、その応用は、多変量データの可視化や、低次元の特徴ベクトル抽出系の対話的設計に限られている。

## 2. 特徴選択手法

本論文で比較する各特徴選択手法について以下に述べる。なお原特徴量の中から識別に有効な少数の特徴量を選択することを特徴選択と呼び、原特徴量の線形結合などの中から新たな特徴量を選択することを特徴抽出と呼んで両者を区別する場合がある。しかし、本論文では原特徴量の抽出との混同を避けるため、この両者をともに特徴選択と呼ぶことにする。

### 2.1 正準判別分析

$n$ 次元の原特徴ベクトルから正準判別分析によって $n'$ 次元の特徴ベクトルを選択する手順を以下に示す[13]。

まず次式を満たす固有ベクトル行列 $\Phi$ と固有値行列 $\Lambda$ を求める。

$$S_b \Phi = S_w \Phi \Lambda \quad (1)$$

ここで、 $S_b$ は級間変動、 $S_w$ は級内変動であり、 $\Lambda$ は固有値 $\lambda_i$  ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ )を対角要素とする固有値行列、 $\Phi$ は対応する固有ベクトル $\Phi_i$ を列ベクトルとする固有ベクトル行列である。 $S_b$ のランクを $r$ とすると最大 $r$ 個の0でない固有値が得られる。またクラス数を $L$ とすると $r = \min(L-1, n)$ となる。特徴ベクトル $X$ に対して線形結合、

$$y_i = \Phi_i^T X \quad (i=1, 2, \dots, n') \quad (2)$$

によりF比(固有値 $\lambda_i$ )の大きい方から $n'$ 個の正準判別変量を選択し新たな特徴ベクトルとする。 $n'=n$ のとき新しい特徴ベクトルは、

$$Y = \Phi^T X \quad (3)$$

で与えられ、その級間変動 $S'_b$ 、級内変動 $S'_w$ は、

$$S'_b = \Phi^T S_b \Phi = \Lambda \quad (4)$$

$$S'_w = \Phi^T S_w \Phi = I \quad (5)$$

となる。ここで $I$ は $n \times n$ の単位行列である。式(4)、(5)から $Y$ の各要素は $S'_b$ 、 $S'_w$ および全変動 $S'$  ( $= S'_b + S'_w$ )に関して無相関になることがわかる。ま

た、式(5)を満たすとき、固有ベクトル行列 $\Phi$ は $S_w$ に関する正規直交行列と呼ばれ、通常の直交行列ではない。

### 2.2 主成分分析 (K-L 展開)

全変動 $S_i (= S_w + S_b)$ を用いて次式を満たす固有ベクトル行列 $\Phi$ と固有値行列 $\Lambda$ を求める。

$$S_i \Phi = \Phi \Lambda \quad (6)$$

式(2)により分散(固有値 $\lambda_i$ )の大きい順に $n'$ 個の主成分を選択し新たな特徴ベクトルとする。

特徴ベクトル $X$ を $n'$ 個の固有ベクトルでK-L展開したときの平均2乗誤差は、

$$E = \sum_{i=n'+1}^n \lambda_i = \text{trace} \{ \Lambda \} - \sum_{i=1}^{n'} \lambda_i \quad (7)$$

で与えられる。固有ベクトル $\Phi$ は正規直交行列であり、行列のトレースは正規直交変換のもとで不変であるから、式(7)は

$$E = \text{trace} \{ S_i \} - \sum_{i=1}^{n'} \lambda_i \quad (8)$$

と書くことができる。式(8)から分散 $\lambda_i$ を最大化する(打切り)K-L展開による特徴選択が、次元減少による平均2乗誤差を最小化することがわかる[13]。

$n'=n$ の場合新しい特徴の全変動 $S'_i$ は、

$$S'_i = \Phi^T S_i \Phi = \Lambda \quad (9)$$

となり、新しい特徴は $S'_i$ に関して無相関になる。

### 2.3 FKL 法

原特徴量の次元数に比べてクラス数が小さい場合、クラス数程度以下の低次元の特徴量を選択するときは正準判別分析が適しているが、より高次元の特徴量を選択するときは主成分分析が適している[14]。従って、F比と分散の双方を特徴選択の基準とする選択手法を実現すれば、少クラスの分類問題における正準判別分析の制約をとり除くことができると考えられる。

正準判別分析と主成分分析を特殊な場合として含む、特徴選択手法に適した一般固有値解析問題を式(10)に示す。

$$(S_b + \beta S_w) \Phi = \{(1-\beta)S_w + \beta I\} \Phi \Lambda \quad (10)$$

ここで、 $\beta$ は $[0, 1]$ の定数である。式(10)を満たす $\Phi$ 、 $\Lambda$ を求め、式(2)により固有値の大きい順に $n'$ 個の変量を選択して新たな特徴ベクトルとする。この特徴選択手法は、 $\beta=0$ で正準判別分析、 $\beta=1$ で主成分分析になるため、 $\beta$ の値を最適値に選んだときに選択される特徴量の識別力は、両者によって選択される特徴量のうち高い方を(上回るとは限らないが、少なくとも)下回ることはない。 $n' > L$ の場合に実際に

識別力が向上することは、3.において実験により確かめる。また  $0 < \beta < 1$  では、F比と分散の荷重和が特徴選択基準として最大化される。例えば、 $\beta$  が 0 に近いときは、最初の  $L-1$  個の特徴は主に F比を最大化し、残りの特徴は主に分散を最大化すると考えることができる。

なお、FKL法を用いた場合、識別時に必要な式(2)の計算量・記憶容量は、正準判別分析、主成分分析と同一であり、学習時に必要な式(10)の計算量・記憶容量も、正準判別分析とほぼ同一である。

#### 2.4 ODV法

F比を特徴選択基準とする少クラス分類問題に適した特徴選択手法として、正規直交条件のもとでF比を最大化するODV法が提案されている。特徴選択に用いる基底ベクトルは以下の手順で求められる。

(1) 正準判別分析により原特徴空間でF比が最大となる基底ベクトルを求め、これを第1基底ベクトルとする。

(2) 第1基底ベクトルの直交補空間でF比が最大となる基底ベクトルを求め、これを第2基底ベクトルとする。

(3) 以下第*i*基底ベクトルは、第1～第*i*-1基底ベクトルすべてと直交する補空間でF比が最大となる基底ベクトルとする。

ODV法を用いればクラス数を超えてF比が0でない特徴が選択できる。また正準判別分析よりもF比の大きな特徴を選択できることが、理論的に明らかにされている[5]。

なお、上記(1)～(3)の各ステップにおいて、F比以外に分散が最大となる基底ベクトルも選択できるようにすると文献[4]の手法と等価になる。

#### 2.5 その他の一般固有値解析問題

正準判別分析の拡張として、次式の固有値解析問題がある。

$$(S_b + \beta S_w)\Phi = S_w\Phi\Lambda \quad (11)$$

ここで、 $\beta$  は  $[0, 1]$  の定数である。式(11)は  $\beta = 0$  で正準判別分析となり、 $\beta = 1$  では左辺が  $S_b\Phi$  となる。従って  $\beta \neq 0$  とすれば、固有値は一般にすべて正となり、0でない固有値が  $n$  個得られる。しかし、式(11)を変形すると、

$$S_b\Phi = S_w\Phi(\Lambda - \beta I) = S_w\Phi\Lambda' \quad (12)$$

となり、その固有ベクトルは式(1)の固有ベクトルに等しく、固有値は式(1)の固有値に  $\beta$  を加えたものに他ならないことがわかる。従ってこの手法では、

正準判別分析の改善は期待できない。

正準判別分析と主成分分析を特殊な場合として含む最も自然な一般固有値解析問題として式(13)あるいは式(14)が考えられる。

$$\{(1-\beta)S_w^{-1}S_b + \beta S_i\}\Phi = \Phi\Lambda \quad (13)$$

$$\{(1-\beta)S_b + \beta S_w S_i\}\Phi = S_w\Phi\Lambda \quad (14)$$

ここで、 $\beta$  は  $[0, 1]$  の定数である。上式は  $\beta = 0$  で正準判別分析、 $\beta = 1$  で主成分分析の固有値解析問題になる。しかし、式(14)の左辺の行列が非対称行列となるため、固有ベクトルが複素ベクトルとなり、特徴選択には不向きである。

### 3. 比較実験

以下に述べる実験サンプル、原特徴ベクトル、識別関数を用いて認識実験を行い、特徴選択手法の評価・比較を行う。

#### 3.1 実験サンプル

実験には、郵政省郵政研究所が作成した第1回および第2回文字認識技術コンテストの学習用・評価用サンプルを用いる[14]～[16]。このサンプルは実際の郵便番号から収集された3けたの数字からなり、総サンプル数は14,954(44,862文字)である。

画像サイズは横240ドット×縦120ドットで、あらかじめ適当なしきい値で2値化されており、郵便番号枠は除去されている。筆記具は、ペン、ボールペン、サインペン、毛筆など多岐にわたっている。

以後の実験では、先頭から9,961サンプル(29,883文字)を学習用サンプル、残りの4,993サンプル(14,979文字)を評価用サンプルとして用いる。

#### 3.2 特徴の抽出と識別

画像の水平周辺分布と文字枠の位置情報を用いて切り出した数字サンプルから、濃度値こう配を利用した400次元の加重方向指数ヒストグラム[17]特徴を抽出し、原特徴量として用いる。

2.に示した各特徴選択手法を用いて、400次元の原特徴量から  $n'$  次元の特徴量を選択する ( $n' = 3, 6, 9, 12, 20, 30, 64, 100, 144, 196, 256, 400$ )。特徴選択に必要な固有値解析には、コレスキー分解とハウスホルダー・QR法[18]を用いる。

式(11)の方法とFKL法において、 $\beta$ の値は以下に述べるモード4のローテーション法により実験的に定める。すなわち評価用サンプル(14,979文字)を4等分し、そのうち三つを用いる認識実験により最適な  $\beta$  を求め、その値を用いて残りの一つを評価するという

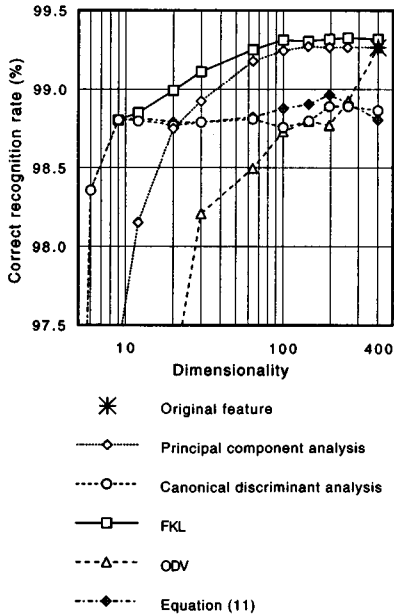


図1 擬似ベイズ識別関数による比較実験  
Fig. 1 Experimental comparison employing a pseudo Bayes discriminant function.

手順を4回繰り返し、平均正読率を求める。従って、 $\beta$ の決定に用いるサンプルと識別率の評価に用いるサンプルに同一のサンプルが含まれることはない。

識別には、文献[14]で総合的に優れた結果を示した擬似ベイズ識別関数以外に、特徴量の次元数と正読率の関係等が理論的によく研究されている2次識別関数および特徴抽出理論の分野でよく研究されている部分空間法を用いる。

識別に用いる平均ベクトル、共分散行列などのパラメータは学習用サンプルから推定し、クラスの事前確率は等確率とした。

擬似ベイズ識別関数で用いる固有ベクトルの数 $k$ は、37 ( $n' \geq 64$ の場合)および $n'$  ( $n' \leq 30$ の場合)とした。また、部分空間法における $k$ の最適値は次元数により異なるため、 $1 \leq k \leq n'$ の範囲で正読率が最高となる値を求めて用いた。

### 3.3 実験結果と考察

図1に擬似ベイズ識別関数による比較実験の結果を示す。横軸は特徴選択後の次元数(対数表示)、縦軸は正読率を表す。正準判別分析による特徴選択では、9次元までは他の手法に比べ高い値を示しているが、それ以上の次元数では正読率が向上しないことがわかる。

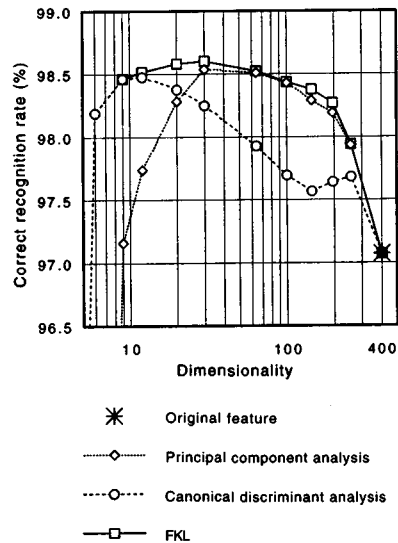


図2 2次識別関数による比較実験  
Fig. 2 Experimental comparison employing the quadratic discriminant function.

また、400次元すべてを選択しても、原特徴量に及ばない。

一方主成分分析により特徴選択した場合の正読率は、20次元までは正準判別分析に及ばないが、30次元以上では正準判別分析を大きく上回っており、400次元で原特徴量と同じになる。また、400次元から100次元まで次元減少しても、正読率の低下はほとんど見られない。

式(11)の方法は正準判別分析とほぼ同じ傾向を示している。

ODV法では、400個の特徴量すべてを選択した場合の正読率は原特徴量に対する正読率と同じになったが、256次元以下では正読率が大きく低下している。

FKL法により特徴選択した場合の正読率は9次元以下では正準判別分析と同等であるが、12次元以上では他のどの特徴選択手法を用いた場合より高い値を示している。特に20次元、30次元では他手法との差が顕著である。また100次元から400次元の特徴ベクトルでは、原特徴量より高い正読率が得られている。400次元で原特徴量より高い正読率が得られることは興味深い結果である。また、64次元程度まで次元減少しても原特徴量と変わらない正読率が得られるため、主成分分析に比べて更に計算時間や記憶容量を節約することができる。

ところで、主成分分析と ODV 法によって 400 個の特徴を選択すると、その正読率は原特徴量と同じになるが、正準判別分析、式 (11) の方法、FKL 法によって 400 個の特徴を選択すると、その正読率は原特徴量と同じにならない。これは、前者が正規直交変換であ

るのに対して後者が斜交変換であることと、擬似ベイズ識別関数が正規直交変換のもとで不変であることによる。後述の部分空間法も同様の性質を示す。

図 2 に 2 次識別関数による比較実験の結果を示す。次元数と正読率の関係を見ると、正準判別分析では 12 次元、主成分分析では 30 次元をピークに正読率が低下している。これは、共分散行列の推定誤差によって尖頭現象 (Peaking phenomenon) が生じているためである。FKL 法では、二つのピークの間を覆うように正読率が向上している。2 次識別関数は一般の線形変換の下で不変であるため、400 次元ではすべての手法で正読率が一致する。

図 3 に部分空間法による比較実験の結果を示す。部分空間法においても FKL 法が最も良い結果を示している。低い次元数において擬似ベイズ識別関数より正読率が低下しているのは、各クラスの部分空間に共通部分が増加するためであると考えられる [14]。

なお、図 1 ~ 図 3 における FKL 法による識別率の向上は、正準判別分析と主成分分析の良い方と比べて 0.05 ~ 0.5 % 程度であり、次元数によっては向上の程度が顕著でない場合もある。しかし、FKL 法に要する計算量・記憶容量や、処理手順の複雑さが正準判別分析と同程度であることを考慮すれば、FKL 法が少クラ

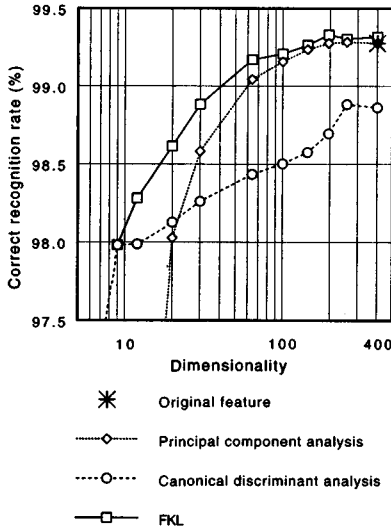
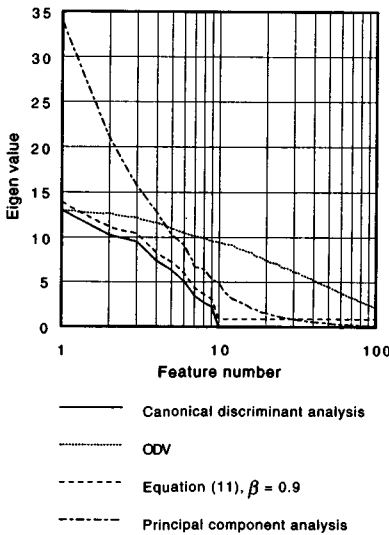
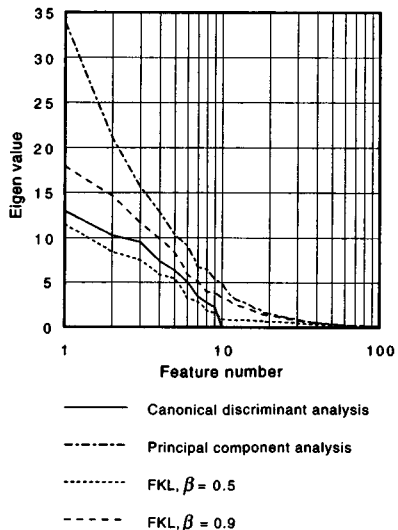


図 3 部分空間法による比較実験

Fig. 3 Experimental comparison employing the subspace method.



(a)



(b)

図 4 各特徴選択手法の固有値

Fig. 4 Eigenvalues of each feature selection method.

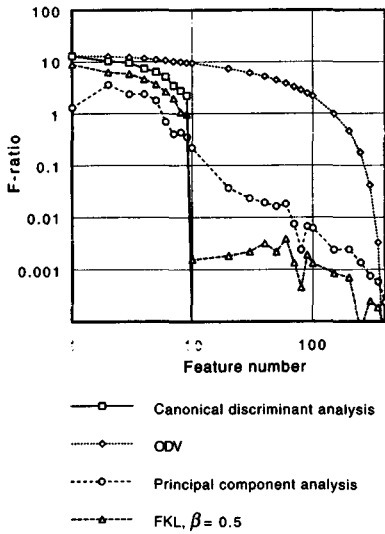


図5 各特徴選択手法のF比  
Fig. 5 F-ratio of each feature selection method.

ス分類問題における特徴選択手法として有効であり、応用上の価値を有することがわかる。

図4は各特徴選択手法で求められる固有値を表したものである。ODV法、式(11)の方法( $\beta=0.9$ )を同図(a)に、FKL法( $\beta=0.5, \beta=0.9$ )を同図(b)に、また正準判別分析、主成分分析は(a),(b)両方に示す。この固有値は正準判別分析、ODV法においては各基底ベクトル上のF比であり、主成分分析においては、各基底ベクトル上の分散である。

図4(a)により、式(11)の方法の固有値は正準判別分析の各固有値に $\beta$ を加えたものになっていることがわかる。また、ODV法の固有値は第2固有値以降、正準判別分析の固有値より大きく、第10固有値以降も顕著に大きな値となっている。また図4(b)に示すFKL法の固有値は、 $\beta=0.5$ では第9固有値までは正準判別分析の固有値より小さいが、第10固有値以降も0ではなくになっている。 $\beta=0.9$ ではより主成分分析の固有値に近づく。

図4では固有値の値を示したが、図5には、正準判別分析、主成分分析、ODV法、FKL法( $\beta=0.5$ )で用いる各基底ベクトル上のF比の値を計算したものを示す。正準判別分析のF比が第9基底ベクトルまで高くそれ以降0であるのに対して、主成分分析のF比は第9基底ベクトルまでは全体的に小さいが、第10基底ベクトル以降も0にならない。FKL法では第9基底ベ

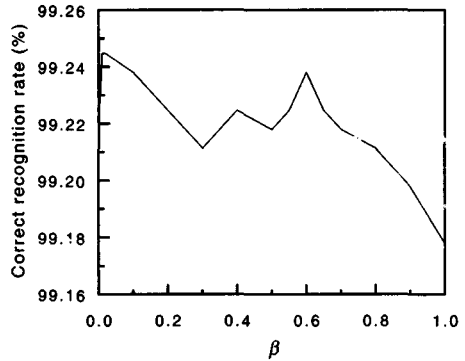


図6  $\beta$ と正読率の関係(64次元)  
Fig. 6 Correct recognition rate v.s.  $\beta$  (64 dimensional case).

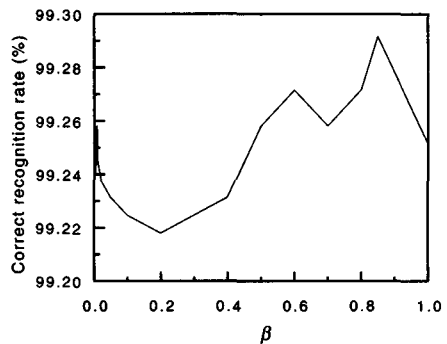


図7  $\beta$ と正読率の関係(144次元)  
Fig. 7 Correct recognition rate v.s.  $\beta$  (144 dimensional case).

クトルまで正準判別分析に近い値でなおかつ第10基底ベクトル以降0とならない。ODV法はF比が顕著に大きい。それにもかかわらずODV法により選択した特徴量に対する正読率が低い理由として、ODV法では級内変動、級間変動が対角化されないために相関の高い冗長な特徴が選択されることが挙げられる。これらの結果から、個々の特徴のF比が大きくても多変量としての特徴ベクトルの識別力が高いとは限らないことがわかる。ODV法の改良手法として、一度抽出された特徴の削除を許すODV法[19]が提案されているが、次元数が大きい場合に計算量が膨大となるため、本論文では触れなかった。

### 3.4 FKL法における次元数と $\beta$ の関係

前節の実験では、 $\beta$ の値はローテーション法により実験的に決定した。ここでは次元数と $\beta$ の関係について考察する。図6、図7にそれぞれ64次元、144次元における $\beta$ と正読率の関係を示す。識別関数には擬似

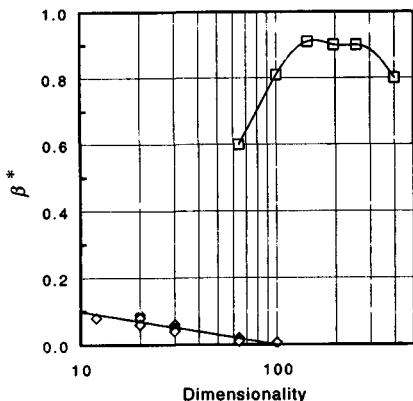


図8 次元数と $\beta^*$ の関係  
Fig. 8  $\beta^*$  vs. dimensionality.

ベイズ識別関数を用い、正読率は評価用サンプルすべてを用いて求めている。

図6、図7を見ると正読率のピークは0に近いところと1に近いところにあり、低い次元数では前者が高く、高い次元数では後者が高いことがわかる。

図8は、次元数と、上記のピークを与える $\beta$ の値( $\beta^*$ )の関係を示したものである。これらの結果から、FKL法における $\beta$ の値は、ある次元数以下では正準判別分析に近く、またその次元数以上では主成分分析に近くなるように設定すればよいことがわかる。

#### 4. む す び

少クラス分類問題における特徴選択手法として、正準判別分析、主成分分析、正規直交判別ベクトル法に加え、新しい特徴選択手法FKL法を提案し、大量の手書き数字サンプルを用いる実験で、特徴量の次元数と正読率の関係を調べた。

一連の実験においてFKL法が最も優れた結果を示しており、その有効性が確認できた。また、選択後の特徴量を用いた場合の正読率が400次元の原特徴量の正読率を上回っており、64次元程度まで次元減少しても原特徴量と変わらない正読率が得られることがわかった。更に特徴選択の基準としてF比の最大化以外に、

(1) 級内変動と級間変動、あるいは全変動に関する無相関性、

(2) 分散の最大化、  
が重要であることがわかった。

FKL法は、

(1) 2クラス問題、多クラス問題のいずれにも区

別なく適用できる、

(2) 原特徴量の次元数を超えない範囲で任意の次元数の特徴ベクトルを得ることができる、

(3) 固有値解析問題として数値計算による効率の良い解法が確立している、

などの性質を備えており、幅広い応用が期待できる。

本論文に述べた方法以外にも筆者らはいくつかの特徴選択手法に関する数字認識の予備実験を行った。それらは、(1) 正準判別分析によって選択された9個の特徴と主成分分析によって選択された特徴を組み合わせる方法、(2) 45組の異なる数字の対に対して正準判別分析を適用して45個の特徴量を求める方法、(3) 主成分分析によって得られる主成分の中から、F比の大きい主成分を選択する方法、などであるが、いずれも特に有効性が認められなかった。その理由として、手法(1)、(2)で選択される特徴間には高い相関が生じる可能性があるためと考えられる。手法(3)に関しては、文献[14]に考察がある。

今後の課題として、少数の類似文字や混同しやすい文字対の識別、低品質データにおける有効性の評価、多変量データの分布の可視化への応用、FKL法の有効性や各次元数における $\beta$ の最適設定に関する理論的考察、FKL法のノンパラメトリック化、および式(14)による特徴選択の検討など興味ある問題が残されている。

謝辞 本研究に関し、手書き郵便番号データを作成・公開された郵政省郵政研究所の皆様、討論をして頂いた研究室の皆様に感謝致します。

#### 文 献

- [1] 浜本義彦, “パターン認識理論の最近の動向,” 信学誌, vol. 77, no. 8, pp. 853-864, Aug. 1994.
- [2] D. H. Foley and J. W. Sammon, Jr., “An optimal set of discriminant vectors,” IEEE Trans. Comput., vol. C-24, no. 3, pp. 281-289, 1975.
- [3] T. Okada and S. Tomita, “An optimal orthonormal system for discriminant analysis,” Pattern Recognition, vol. 18, no. 2, pp. 139-144, 1985.
- [4] D. Duchene and S. Leclercq, “An optimal transformation for discriminant and principal component analysis,” IEEE Trans., vol. PAMI-10, no. 6, pp. 978-983, 1988.
- [5] Y. Hamamoto, T. Kanaoka, and S. Tomita, “On a theoretical comparison between the orthonormal discriminant vector method and discriminant analysis,” Pattern Recognition, vol. 26, no. 12, pp. 1863-1867, 1993.
- [6] 浜本義彦, 金岡泰保, 富田真吾, “パターン認識のための正規直交判別ベクトル法に関する理論的考察,” 信学論(A), vol. J77-A, no. 1, pp. 76-83, Jan. 1994.
- [7] J. Fehlaue and B.A. Eisenstein, “A declustering criterion for feature extraction in pattern recognition,” IEEE Trans.

- Comput., vol. C-27, no. 3, pp. 261-266, March 1978.
- [8] W. Malina, "On an extended fisher criterion for feature selection," IEEE Trans., vol. PAMI-3, no. 5, pp. 611-614, Sept. 1981.
- [9] M. Aladjem, "Parametric and nonparametric linear mappings of multidimensional data," Pattern Recognition, vol. 24, no. 6, pp. 543-553, 1991.
- [10] 浜本義彦, 富田真吾, "パターン認識のための一般化正規直交判別ベクトル法," 信学論(A), vol. J77-A, no. 4, pp. 693-700, April 1994.
- [11] K. Fukunaga and J. M. Mantock, "Nonparametric discriminant analysis," IEEE Trans., vol. PAMI-5, no. 6, pp. 671-678, Nov. 1983.
- [12] M. Aladjem and I. Dinstein, "A multiclass extension of discriminant mappings," Proc. 11th Int. Conf. Pattern Recognition, The Hague, pp. 101-104, 1992.
- [13] K. Fukunaga, "Introduction to Statistical Pattern Recognition, Second Edition," Academic Press, 1990.
- [14] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二, "手書き数字認識における特徴選択に関する考察," 信学論 (D-II), vol. J78-D-II, no. 11, pp. 1627-1638, Nov. 1995.
- [15] 松井俊弘, 山下郁生, 若原 徹, 吉室 誠, "文字認識アルゴリズム複合化の検討, 第一回文字認識技術コンテストの結果より," 信学技報, PRU92-33, Sept. 1992.
- [16] 能見 正, 松井俊弘, 山下郁生, 若原 徹, 吉室 誠, "手書き数字認識における誤読・リジェクトパターンの分析," 信学技報, PRU93-46, Sept. 1993.
- [17] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二, "特徴量の次元数増加による手書き数字認識の高精度化," 信学論 (D-II), vol. J77-D-II, no. 10, pp. 2046-2053, Oct. 1994.
- [18] 名取 亮, "数値解析とその応用," コロナ社, 1990.
- [19] Y. Hamamoto, Y. Matsuura, T. Kanaoka, and S. Tomita, "A note on the orthonormal discriminant vector method for feature extraction," Pattern Recognition, vol. 24, no. 7, pp. 681-684, 1991.

(平成7年11月30日受付, 8年4月25日再受付)



若林 哲史 (正員)

昭60三重大・工・電子卒。昭62同大大学院修士課程了。平2三重大・工・助手, 現在に至る。手書き文字認識, 文書理解, 画像処理, コンピュータグラフィックスの研究に従事。



鶴岡 信治 (正員)

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 (正員)

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1~ミシガン大客員助教授を務める。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。情報処理学会, 日本ME学会, 人工知能学会各会員。工博。



三宅 康二 (正員)

昭35名大・工・電気卒。同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。