クラスタ間情報に基づくマハラノビス距離による文字認識

岩村雅一　　大町真一郎　　阿曽弘具

東北大学大学院工学研究科

〒 980-8579 仙台市青葉区荒巻字青葉 05

Tel:022–217–7088　　Fax:022–263–9418

E-mail:masa@aso.ecei.tohoku.ac.jp

あ ら ま し　パターン認識において識別関数としてマハラノビス距離を用いる場合，一般に学習サンプルから得られた標本共分散行列が用いられる．しかし，特徴量の次元数に比べて十分な学習サンプルを用意することが困難な場合，標本共分散行列の固有値・固有ベクトルに推定誤差が生じ，誤識別が生じることが知られている．この問題を解決するため，これまで各クラスタの分布情報をクラスタ毎に補正する手法が提案されている．本論文では，不足している情報をクラスタ間の情報を用いることで補う手法を提案する．具体的には，学習サンプルが少ないために求まる固有値・固有ベクトルの数が不十分な場合において，クラスタ間の情報を用いて新しい軸を作り，擬似的な分散を与える手法を提案する．そして，手書き文字を用いた認識実験により従来法と比べて認識率が向上することを示す．

キーワード　文字認識，マハラノビス距離，クラスタ間情報，推定誤差，ETL9B

Character Recognition with Mahalanobis Distance Based on Between-Cluster Information.

Masakazu IWAMURA, Shin'ichiro OMACHI, and Hirotomo ASO

Department of Electrical and Communication Engineering,
Graduate School of Engineering, Tohoku University
Aoba 05, Aramaki, Aoba-ku, Sendai-shi, 980-8579 Japan

Tel:022–217–7088　　Fax:022–263–9418

E-mail:masa@aso.ecei.tohoku.ac.jp

Abstract　In the case of using the Mahalanobis distance as discriminant function, usually the covariance matrix calculated from training samples is used. However, it is extremely difficult to prepare enough training samples if the dimension of feature vector is large. Therefore, estimated eigenvalues and eigenvectors of covariance matrix will include errors that cause misclassification. In this paper, a new method to construct an effective discriminant function is proposed by considering between-cluster information. In the proposed method, if the number of calculable eigenvalues and eigenvectors is not enough because of less training samples, some new axes are constructed and then the pseudo-variances are computed based on the between-cluster information. The effectiveness of this method is shown by the experiments with handwritten characters.

key words　character recognition, mahalanobis distance, between-cluster information, estimate error, ETL9B

1　はじめに

マハラノビス距離は真の分布が既知ですべてのカテゴリで同じ正規分布であるとき最適な識別関数である [1]．しかしマハラノビス距離をパターン認識に用いる場合，真の分布が既知であることはほとんどなく，パターンの分布を表わす情報は学習サンプルから推定されるのが一般的である．

ところがこうして得られた分布情報を用いて認識を行なっても，期待されたほどの判別性能が得られない．その主な理由として挙げられるのが固有値・固有ベクトルの推定誤差である．特に高次の固有値・固有ベクトルが大きな推定誤差を含むことが知られ [2]，これに対処するために多くの研究が報告されている [3] [4] [5]．これらの研究は大きく 2 つに分けることができる．

一つは推定誤差が大きく信頼性の低い固有値・固有ベクトルの使用を避ける方法である．加藤ら [3] は，固有値にバイアスを加えた改良型マハラノビス距離 (MMD) を提案している．木村ら [4] は，マハラノビス距離とユークリッド距離の荷重和からなる修正 2 次識別関数 (MQDF) を示している．これらの方法ではバイアスやユークリッド距離の重みといったパラメータを与える必要があり，その最適値は理論的には求められず，認識実験によって決められる．

もう一つは，固有値に推定誤差を打ち消すような補正をして，真の分布の正しい推定を目指す方法である．酒井ら [5] は，標本共分散行列から推定される分布形状 (固有値) の偏りを補正する手法を提案した．しかしこの方法では分布形状の補正に少数の学習サンプルのみが用いられるため，得られる情報には限りがあると考えられる．

そこで著者らは，学習サンプルから分布を推定する際に発生する誤差などの悪影響を避け，認識精度を向上させることを目的とし，クラスタ間情報に着目し，一つのクラスタで不足している分布情報を他のクラスタの分布情報を用いて誤認識が起こらないように補う手法を提案する．ここでいうクラスタ間情報とは，各クラスタの分布や相対的な位置などの情報のことである．すなわち，各クラスタの分布情報に，当該学習サンプルだけでなく，全学習サンプルの情報を反映させるのである．

具体的には対象クラスタの軸の分散を他のクラスタの分布情報を用いて修正し，修正した分散を用いて認識を行なう．特に学習サンプル数が次元数より少ない場合を考える．このケースでは全ての軸が求まらないので，各クラスタに新しい軸を作成し，クラスタ間の分布に着目して擬似的な分散 (擬似固有値) を与える．

本手法のメリットとして，分布の推定誤差が原因で引き起こされる誤認識を回避できることが挙げられる．これまでの研究では各クラスタの分布はそれぞれのカテゴリに属する学習サンプルから得られるが，高次の固有値・固有ベクトルに推定誤差が集中していることがわかっている [2] ため，信頼性の低い成分を修正する際に，他のクラスタの信頼性の比較的高い成分が持つ情報を利用する．また，この方法はさらに実験的に最適パラメータを求める必要のない手法としても大いに期待できる．なぜなら [3] や [4] で用いられているバイアスやユークリッド距離の重みなどはクラスタ間の距離や分布といったクラスタ間に存在する情報に依存していると考えられるため，本手法で包含できると考えられるからである．

2　少数学習サンプルの影響

2.1　マハラノビス距離

マハラノビス距離 $d^2(\boldsymbol{x})$ は $\boldsymbol{x}$ を未知入力ベクトル，$\boldsymbol{\mu}$ を標本平均ベクトル，$\boldsymbol{\Sigma}$ を標本共分散行列とすると，

$$d^2(\boldsymbol{x}) \;=\; (\boldsymbol{x}-\boldsymbol{\mu})^t\,\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu}) \qquad (1)$$

と表わされる．ここで，λ_k を標本共分散行列の第 k 固有値 ($\lambda_1 \geq \cdots \geq \lambda_D$)，$\boldsymbol{\varphi}_k$ を λ_k に対応する固有ベクトルとすれば，

$$d^2(\boldsymbol{x}) \;=\; \sum_{k=1}^{D} \frac{1}{\lambda_k}\,(\boldsymbol{\varphi}_k\cdot(\boldsymbol{x}-\boldsymbol{\mu}))^2 \qquad (2)$$

と表わすこともできる．ここで D は次元数であるが，学習サンプルが少数の場合には次元数分の固有値，固有ベクトルが求まらない [6]．そのため実際には，

$$\tilde{d}^2(\boldsymbol{x}) \;=\; \sum_{k=1}^{M} \frac{1}{\lambda_k}\,(\boldsymbol{\varphi}_k\cdot(\boldsymbol{x}-\boldsymbol{\mu}))^2 \qquad (3)$$

を用いて認識を行なうことになる．ここで M は求まる固有値・固有ベクトルの数であり，学習サンプル数を N とすると $M \leq \min\{D, N-1\}$ である．本論文では (3) 式をマハラノビス距離ということにする．

2.2　起こりうる誤認識例

固有ベクトルと固有値は標本が分布する領域を特徴ベクトル空間内の超楕円体ととらえるときの軸べ

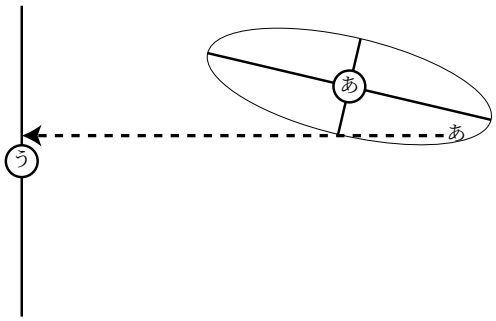


図 1: 誤認識の例

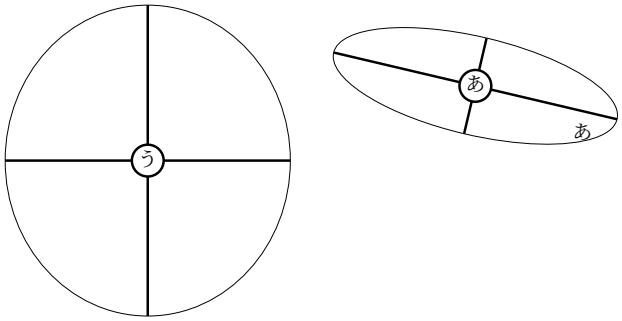


図 2: 改善例

クトルとその軸上の半径 (の 2 乗) を与える．この観点からは，(3) 式は固有ベクトルが存在しない方向の分散 (その方向の半径) を無限大とみなすことを表わしている．そのため図 1 のように，明らかに「あ」に属すと思われる入力文字が「う」と誤認識される例が考えられる．そこで，このような誤認識が起こるのを防ぐ方法として，求まらない固有ベクトルの代わりに適当な軸を作成し，各軸に適切な分散を与えることが考えられる．それを行なったのが図 2 であり，図 1 では誤認識された未知入力文字が正しく「あ」と認識されるようになる．

今後，新しく作成した軸を擬似固有ベクトル，その分散を擬似固有値と呼ぶことにし，3 章でその与え方を述べる．

3 擬似固有ベクトル・擬似固有値

3.1 擬似固有ベクトルの作成

標本共分散行列から求まった固有ベクトルに以下の条件を満たす軸を追加し，合計 D 本の軸を用意する．作成した擬似固有ベクトルは今後，第 $M+1$ 固有ベクトルから第 D 固有ベクトルであるかのように扱う．第 k 固有ベクトルを $\boldsymbol{\varphi}_k = (\varphi_{k1}, \cdots, \varphi_{kD})^t$ と記述する．

以下に擬似固有ベクトルの具体的な作成法を示す．計算量削減のため，擬似固有ベクトルとして標準基底が使用できる場合は積極的に利用することにする．

3.1.1 標準基底の利用

$(1, 0, \cdots, 0)$, $(0, 1, \cdots, 0)$ といった標準基底の中で，全ての固有ベクトルと直交するものは新たな固有ベクトルとして使用できる．特徴量の選び方や字種によっては，固有ベクトルが

$$\boldsymbol{\varphi}_k = \left(\varphi_{k1}, \cdots, \varphi_{k(n-1)}, 0, \varphi_{k(n+1)}, \cdots, \varphi_{kD}\right)^t \quad (k = 1, \cdots, i) \quad (4)$$

のようにある次元が全て 0 となる場合がある．この場合には第 $i+1$ 固有ベクトルとして，

$$\boldsymbol{\varphi}_{i+1} = (0, \cdots, 0, 1, 0, \cdots, 0)^t \quad (5)$$

をとることができる．このようなベクトルは全て擬似固有ベクトルとして採用する．

3.1.2 擬似固有ベクトルの一般的な作成法

標準基底が利用できる条件は限られているので，それ以外の場合には既に求まっている固有ベクトル全てに直交するベクトルを求めることになる．一般に D 次元空間で i 本のベクトルと直交するベクトルは $D-i$ 本あり，これらを 1 本ずつ作成する．

固有ベクトルと擬似固有ベクトルが合計 i 本求まっているとすれば，第 $i+1$ 固有ベクトルに相当する擬似固有ベクトルはノルムが 1 で，第 $i+2$ 要素以降を 0 にしたものを考える．すなわち，

$$\begin{bmatrix} \varphi_{11} & \cdots & \varphi_{1(i+1)} \\ \vdots & \ddots & \vdots \\ \varphi_{i1} & \cdots & \varphi_{i(i+1)} \end{bmatrix} \begin{bmatrix} \varphi_{(i+1)1} \\ \vdots \\ \varphi_{(i+1)(i+1)} \end{bmatrix} = \boldsymbol{O} \quad (6)$$

$$\varphi_{(i+1)1}^2 + \cdots + \varphi_{(i+1)(i+1)}^2 = 1 \quad (7)$$

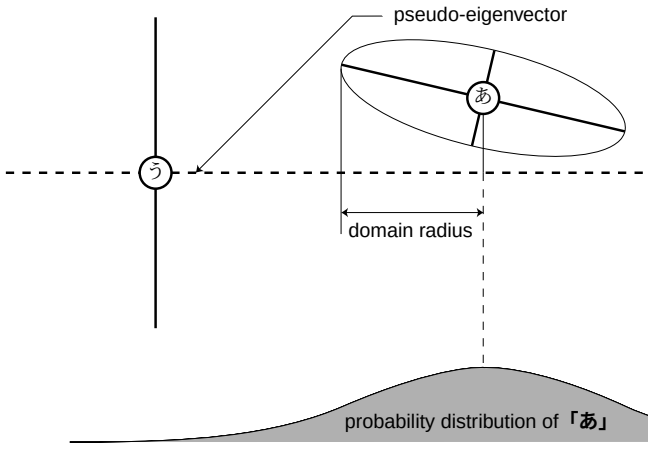
$$\varphi_{(i+1)(i+2)} = \cdots = \varphi_{(i+1)D} = 0 \quad (\text{if exists}) \quad (8)$$

を満たすベクトルである．

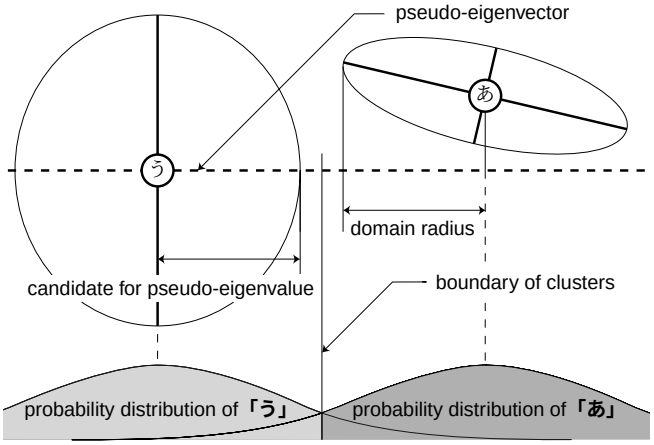
3.2 擬似固有値の作成

作成した擬似固有ベクトルそれぞれに特徴領域を表わす超楕円体の半径となる，擬似固有値を与える．擬似固有値の与え方は種々考えられるが，本論文では分散を与える必要のある軸に対して以下の方法で与えることにする．

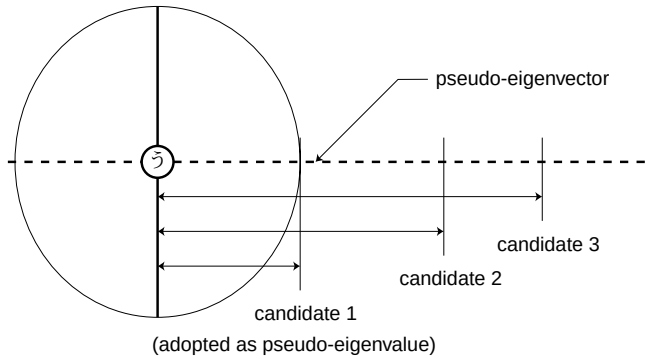
1. 対象とする擬似固有ベクトル上に他のクラスタの分布を射影する (図 3(a)).
2. 射影された他のクラスタの分布の軸上の分散と同じ値を当該クラスタの擬似固有値の候補とする (図 3(b)). これは 2 つのクラスタの軸上での識別境界をクラスタ間の中点に置くことを意味している.
3. 全ての擬似固有値の候補のうち，最小のものを擬似固有値として採用する (図 3(c)). 2 の手順で，(クラスタ数 − 1) 個の擬似固有値の候補が決まるが，この中から最小のものを選ぶことにより，図 1 のような，他のクラスタに属する未知入力文字がこのクラスタに属するというような誤認識を全てのクラスタについて防ぐことができる.



(a) 他クラスタの射影



(b) 候補選出



(c) 擬似固有値の決定

図 3: 擬似固有値作成法

4 認識実験

提案手法の有効性を調べるために認識実験を行なった．実験の比較対象として，(3) 式で表わされるマハラノビス距離及び (10) 式で表わされる識別関数を用いた．MQDF[4] は 2 次識別関数に修正を加えたものであるが，この識別関数はマハラノビス距離に MQDF の考えを導入したものである．今後これを従来手法と呼ぶこととし，h^2 と認識率の関係を調べた．この手法は提案手法の擬似固有値を全て同じ値 h^2 にしたのと同じである.

$$
\begin{aligned}
g(\boldsymbol{x}) &= \sum_{k=1}^{m} \frac{1}{\lambda_k} \left(\boldsymbol{\varphi}_k \cdot (\boldsymbol{x}-\boldsymbol{\mu})\right)^2 \\
&\quad + \sum_{k=m+1}^{D} \frac{1}{h^2} \left(\boldsymbol{\varphi}_k \cdot (\boldsymbol{x}-\boldsymbol{\mu})\right)^2 \qquad (9) \\
&= \frac{1}{h^2}\|\boldsymbol{x}-\boldsymbol{\mu}\|^2 \\
&\quad - \sum_{k=1}^{m} \left(\frac{1}{h^2} - \frac{1}{\lambda_k}\right) \left(\boldsymbol{\varphi}_k \cdot (\boldsymbol{x}-\boldsymbol{\mu})\right)^2 (10)
\end{aligned}
$$

4.1 使用サンプル

本実験では手書き文字データベース ETL9B[8] のうち，ひらがな 71 字種を使用した．全 200 セットのうち，第 1 〜第 20 セットを認識用サンプル，第 21 〜第 200 セットまでを辞書作成用の学習サンプルとした．各サンプルは文字画像を 64×64 の大きさに非線形正規化 [7] をした後，32×32 の大きさの画像に

変換した．変換は，64×64 の画像の重複しない 4 画素 (2×2) の黒画素数を 32×32 の画像の 1 画素の値とすることで行なう．こうして作成した 32×32 の画像の各画素を 1 次元として 1024 次元特徴量とした．なお，ノイズ除去やスムージング，正準化等の前処理は一切行なっていない．

4.2 結果と考察

図 4 にマハラノビス距離の認識率のグラフを示す．これは第 1 固有値・固有ベクトルから第 n 固有値・固有ベクトルを認識に使用した場合のグラフ [9] で，横軸が n，縦軸が認識率を表わしている．この結果，$n = 155, 156$ のときに最も高い 85.70%の認識率を示し，その後認識率は落ちて $n = 179$ の場合では 53.94%となった．マハラノビス距離の認識率は理論的に次元数が増えるにつれて単調増加すると考えられるので，認識率の上げ止まり，低下は，固有値・固有ベクトルに推定誤差が多く含まれているためであると考えられる．以後の認識実験は固有値・固有ベクトルは第 1 固有値・固有ベクトルから第 150 固有値・固有ベクトルを使用するものとする．(10) 式においては $m = 150$ である．

図 5 に従来手法の認識率，表 1 に提案手法の認識率を示す．提案手法の辞書 1 とは，第 1 固有値・固有ベクトルから第 150 固有値・固有ベクトルと，874 個の擬似固有値・擬似固有ベクトルを辞書とした場合である．辞書 2 とは，874 個の擬似固有値・擬似固有ベクトルのみを辞書として与えた場合であり，性能評価用である．図 5 中の点線に提案手法 (辞書 1) の認識率を示している．

従来手法では h^2 が小さくなるにつれ認識率が上昇し，$h^2 = 0.9$ のときに認識率が 95.49%で最高となった．しかし h^2 が 0.9 より小さくなると認識率は下がり，0.1 より小さくなると一定値 95.07%となった．h^2 が小さくなると，固有値を h^2 で置き換えた成分である (9) 式における第 2 項が大きくなる．距離に占めるこの成分が第 1 項を無視できるほど大きくなると，(10) 式は (11) 式で表わされる．

$$\tilde{g}(\boldsymbol{x}) \quad = \quad \sum_{k=m+1}^{D} \frac{1}{h^2} \left(\boldsymbol{\varphi}_k \cdot (\boldsymbol{x} - \boldsymbol{\mu})\right)^2 \qquad (11)$$

辞書 1 を用いた認識実験で本手法はマハラノビス距離，従来手法よりも高い認識精度を示し，本手法が有効であることが示された．辞書 2 を用いた認識実験においても，提案手法が (11) 式よりも高い認識精度が得られることを示している．図 5 で認識率が

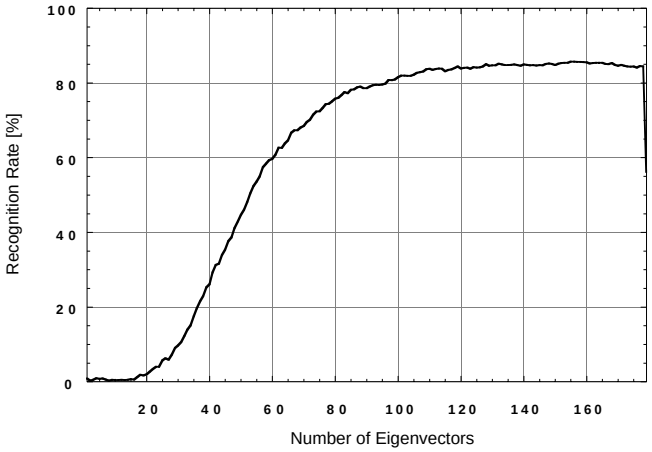


図 4: マハラノビス距離の認識率

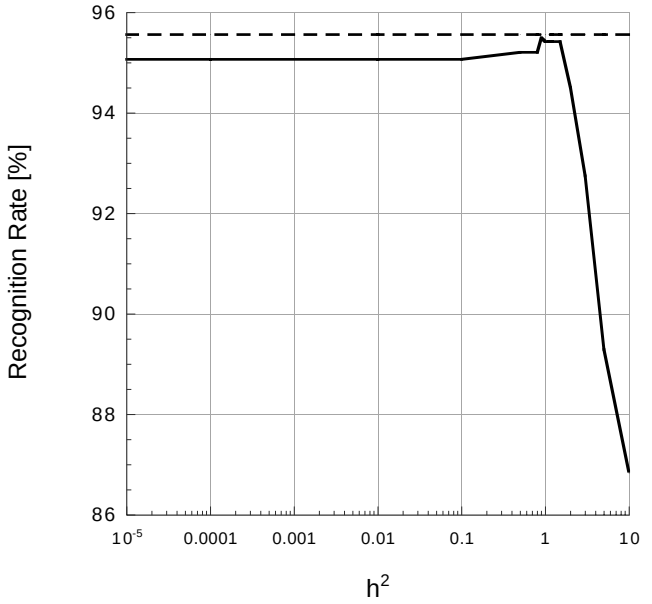


図 5: 従来手法の認識率

高くなっている部分が (9) 式の第 1 項と第 2 項のバランスがとれている部分であり，従来手法が活きる部分である．この認識率のピークが表われる付近では h^2 の小さな変動に対する認識率の変動が大きく，その範囲も狭い．しかし，本手法では実験的なパラメータを用いることなく従来手法での最高の認識性能と同等以上の認識が可能である．これらの実験，考察により，本手法の有効性が確認された．

5 結び

本論文では，クラスタ間情報に着目し，一つのクラスタで不足している分布情報を他のクラスタの分布情報を用いて誤認識が起こらないように補う手法を

手法	認識率
辞書 1	95.56
辞書 2	95.28

表 1: 提案手法の認識率

提案した．特に学習サンプル数が次元数より少ない場合について，擬似固有ベクトルと擬似固有値の与え方を提案した．提案手法の有効性を確認するためにマハラノビス距離及びマハラノビス距離に MQDF の考えを導入した識別関数との比較実験を行ない，その有効性を確認した．

比較実験に用いた，マハラノビス距離に MQDF の考えを導入した手法は，パラメータ h^2 が狭い範囲のときにのみ高い認識性能を示すが，提案手法では予備実験をすることなく，この性能と同等以上の性能を発揮する．

しかし，擬似固有値の作成法にはまだまだ改良の余地がある．本論文では特にクラスタ間の境界に注目し，なるべく誤認識が起こらないように擬似固有値を作成したが，より識別性能の高い擬似固有値を作成することが今後の課題である．

参考文献

[1] Richard O. Duda, Peter E. Hart, “Pattern classification and scene analysis,” A Wiley-Interscience Publication, pp.24–31, 1973.

[2] 竹下鉄夫, 木村文隆, 三宅康二, “マハラノビス距離の推定誤差に関する考察,” 信学論 (D), vol.J70-D, no.3, pp.567–573, 1987.

[3] 加藤　寧, 安倍正人, 根元義章, “改良型マハラノビス距離を用いた高精度な手書き文字認識システム,” 信学論 (D-II), vol.J78-D-II, no.6, pp.922–930, June 1995.

[4] 木村文隆, 高階健治, 鶴岡信治, 三宅康二, “2 次識別関数のピーキング現象とその防止に関する考察,” 信学論 (D), vol.J69-D, no.9, pp.1328–1334, Sep. 1986.

[5] 酒井　充, 米田政明, 長谷博行, “多次元で有効な新しい 2 次識別関数,” 信学技報, vol.PRMU98-43, June 1998.

[6] K. Fukunaga, “Introduction to statistical pattern recognition,” 2nd edition, Academic Press, pp.39–40, 1990.

[7] 山田博三, 斉藤泰一, 山本和彦, “線密度イコライゼーション—相関法のための非線形正規化法,” 信学論 (D), vol.J67-D, no.11, pp.1379–1383, Nov. 1984.

[8] 斉藤泰一, 山田博三, 山本和彦, “JIS 第 1 水準手書漢字データベース ETL9 とその解析,” 信学論 (D), vol.J68-D, no.4, pp.757–764, 1985.

[9] 岩村雅一, 大町真一郎, 阿曽弘具, “認識率に寄与する文字画像の固有ベクトル,” 平成 10 年度電気関係学会東北支部連合大会, 2G24, p.286, 1998.