

非線形正規化と特徴量の圧縮による手書き漢字認識の高精度化

若林 哲史[†] 鄧 巖^{††} 鶴岡 信治^{††} 木村 文隆[†]
三宅 康二[†]

Accuracy Improvement by Nonlinear Normalization and Feature Compression
in Handwritten Chinese Character Recognition

Tetsushi WAKABAYASHI[†], Yang DENG^{††}, Shinji TSURUOKA^{††}, Fumitaka KIMURA[†],
and Yasuji MIYAKE[†]

あらまし 本論文では、統計的漢字認識の前処理における非線形正規化、原特徴量の高次元化、および特徴選択によって手書き漢字認識の認識精度を向上させることを検討する。また、処理の高速化のために、線形識別関数による高速大分類の手法と有効性について検討する。その結果、ETL9Bの全セットを対象とする認識実験で、99.05%の高い正読率と約9文字/秒の読取り速度が得られることを報告する。また、非線形正規化によって統計的漢字認識手法の正読率が向上する理由を、特徴量の分離性と正規性の観点から考察する。

キーワード 文字認識, 手書き漢字認識, 非線形正規化, 正準判別分析, 特徴抽出, 特徴選択, ETL9

1. ま え が き

ハードウェア技術の進歩により文字認識技術の応用範囲は、専用のOCR装置から、パーソナルコンピュータで利用できるソフトウェアOCR、携帯端末のペン入力などへと広がっている。使用字種数が比較的少ない印刷英文OCRでは、既にいくつかの商用ソフトウェアが利用されており、それらの性能評価に関する研究も行われている[1], [2]。印刷日本語OCRも商用ソフトウェア化が進んでいるが、使用字種数が多いため、十分なコストパフォーマンスを実現し手書き文字認識機能を導入・拡張するためには、文字認識技術のいっそうの発達が望まれる[3]。

統計的漢字認識手法の認識精度を向上させるためには、原特徴量の高次元化によって特徴空間における各クラスの分離性を向上させることが有効である。しかし、原特徴量の高次元化には以下のような問題がある。

- (1) 分布形や分布パラメータの推定誤差の増加による識別性能の劣化
- (2) クラス内の分散(変動)が大きく分離性の悪い特徴量の増加
- (3) 正規性(正規分布らしさ)の悪い特徴量の増加
- (4) 必要な計算量、記憶容量の増加

これらの問題には一般的に特徴選択による次元削減が有効であり、また上記(2)および(3)に対しては、文字画像の非線形正規化が有効であると考えられる。

本論文は、統計的漢字認識の前処理における非線形正規化、原特徴量の高次元化、および特徴選択によって認識精度を向上させることを検討したものである。

手書き文字の形状の変動を吸収・正規化するために、文字画像の非線形正規化法が提案され大きな効果があることが報告されている[4]~[6]。筆者らは文献[7]で非線形正規化と分散共分散を考慮した統計的識別手法の相乗効果によって、手書き漢字データベースETL9Bに含まれる3036字種の未学習サンプルを対象とした小規模な認識実験で98.5%の高い正読率が得られることを示した。また、文献[8]では非線形正規化と部分空間法の組合せによって97.76%、文献[9]では非線

[†] 三重大学工学部情報工学科, 津市
^{††} 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
^{†††} (株)ジャストシステム, 徳島市
JustSystem, Tokushima-shi, 770 Japan

形正規化と改良型マハラノビス距離によって98.24%の平均正読率がETL9Bの全サンプルに対して得られている。しかし、非線形正規化によって正読率が向上する理由は、直感的には理解されているが、これまでのところ十分に詳しく調べられていない。

文字認識の高精度化の他の手段として筆者らは、原特微量の高次元化と特徴選択の有効性について考察してきた[10]~[12]。文献[11]、[12]では、曲線部分が多く字種数の少ない数字の認識においては、文字輪郭線の方向量子化数の増加と、主成分分析による特徴選択が有効で、400次元の原特微量から得られる140次元程度の特微量によって、郵政省郵政研究所作成の手書き数字データベースに対して99.2%の高い正読率が得られることを示した。本論文では、直線部分が多く、字種数の多い漢字の認識に適した原特微量の高次元化と特徴選択の手法について考察する。

その他、処理の高速化のために、線形識別関数による高速大分類の手法と有効性について検討し、学習サンプル数の不足を補うために、非線形正規化を利用して学習サンプル数を増加させる方法などについて考察する。

2. に本研究において最も高い正読率が得られた認識手法を示し、その認識手法を決定するために行った予備実験については3. に述べる。4. では、ETL9Bの全サンプルに対する実験結果について述べる。

2. 認識手法

提案する認識手法は非線形正規化、特徴抽出、特徴選択、大分類、詳細識別の各過程から構成されている。各過程について2.1から2.5に述べる。

2.1 非線形正規化

以下に非線形正規化の手順を簡単に示す。

(1) 原画像(2値)を 2×2 の平均値フィルタで平滑化し、2値化する。

(2) 各画素での線密度を求める。線密度の定義は文献[6]による。

(3) (2)で求めた線密度を均等化する変換関数を作成し、平滑化画像を非線形に正規化する。

(4) 3×3 の平均値フィルタで平滑化し2値化する。

2値画像に対する非線形正規化によって得られる画像は、文字輪郭部のジャギーが誇張されるため、輪郭線の方向を特徴量とする認識手法に対して悪影響を及ぼす。この悪影響を低減するために、非線形正規化前

後の画像にそれぞれ 2×2 および 3×3 の平滑化フィルタを適用する。また、処理手順を工夫することにより、外接枠・重心合せ[13]による位置・大きさの正規化と同程度の処理速度が実現できる。

2.2 特徴抽出

手書き数字認識に対する文献[11]の考察では、領域分割数および方向量子化数が縦 $5 \times$ 横 5×16 方向のときに最も高い正読率が得られ、領域分割数を更に増やしても正読率が向上しないことがわかった。一方、ETL9Bに対する手書き漢字認識では、縦 $7 \times$ 横 7×4 方向のときに最も高い正読率が得られているが、同じ条件下で方向量子化数を更に増加させた場合の考察はまだされていない[10]。本論文では、方向量子化数が8の場合に得られる392次元(縦 $7 \times$ 横 7×8 方向)の特徴ベクトルを原特微量として用いる。

以下に特徴抽出の手順を簡単に示す。

(1) 非線形正規化画像に対して文字部分の輪郭線を8近傍で追跡し、輪郭点のチェーンコードを求める。隣接する二つのベクトル和から向きの違いを含めて16方向に量子化する。

(2) 正規化画像を169個(縦 $13 \times$ 横 13)の小領域に分割し、領域別、方向別に輪郭画素数を計数し、2704次元(縦 $13 \times$ 横 13×16 方向)のヒストグラムを得る。

(3) 2次元ガウスフィルタを縦横1領域おきに施して、領域数を 13×13 から 7×7 に削減する。同様に $[1 \ 2 \ 1]$ の加重フィルタを1方向おきに施して、方向量子化数を16方向から8方向に削減し、392次元(縦 $7 \times$ 横 7×8 方向)の特徴ベクトル(加重方向指数ヒストグラム)を得る。

(4) 変数変換($y = x^{0.5}$)により、特徴ベクトルの分布を正規分布に近づける[14]、[10]。

2.3 特徴選択

特徴ベクトルの次元数を増加させると、特徴空間におけるカテゴリー(字種)の分離性は良くなるが、一方で処理に必要な計算量・記憶容量が増大する。また次元数に対して学習サンプル数が相対的に不足し、共分散行列の推定誤差等の影響が大きくなり認識精度が低下する場合がある。

これらの問題を解決するために、統計的手法により識別に適した少数の特徴量を選択し、認識精度を低下させることなく次元数を減少させることが有効である。ここでは、正準判別分析により求めた変換行列を用いて、高次の原特徴ベクトルから低次の特徴ベクトルを求める。

なお、特徴抽出理論の分野では、原特徴量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、原特徴量のすべてを用いて識別に有効な少数の特徴量を新たに生成することを特徴抽出と呼んでいる。しかし本論文においては、前節の文字画像からの特徴抽出との混同をさけるため、この両者をともに特徴選択と呼ぶことにする。

以下に、 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示す。

まず全クラスの学習サンプルを用いて次式を満たす固有ベクトル行列 Φ と固有値行列 Λ を求める。

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

ここで、 S_b は級間分散、 S_w は級内分散である。また Λ の対角要素を λ_i ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$)、対応する固有ベクトルを Φ_i ($\Phi = [\Phi_1 \Phi_2 \dots \Phi_n]$) とする。特徴ベクトル X に対して式 (2) の線形結合で n' 個の正準判別変量を求め、新たな特徴ベクトルとする。

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

特徴選択は n 次元の内積計算 n' 回で完了するため、その計算量は漢字・平仮名認識システム全体の計算量に比べて十分に小さい。また、次元数を削減することで以降に述べる大分類や詳細識別に必要な計算量や記憶容量を大きく減らすことができる。

2.4 大分類

手書き漢字・平仮名はカテゴリ数が多く、計算量の多い識別関数を全カテゴリに対して適用することは認識速度の制約上困難である。そのため、まず計算量の少ない識別関数で大分類を行い、候補字種を絞った後に詳細識別を行うことによって認識速度を向上させる必要がある。計算量の少ない識別関数としては、シティーブロック距離、ユークリッド距離、線形識別関数等が考えられる。ここでは、線形識別関数として次式で定義されるものを考える。

$$g_l(X) = W_l^T X + W_{l0} \quad (3)$$

$$W_l = S_w^{-1} M_l$$

$$W_{l0} = -\frac{1}{2} M_l^T S_w^{-1} M_l$$

ここで、 l はクラスを表す添字、 M_l は第 l クラスの平均ベクトル、 S_w は全クラスの級内分散である。

線形識別関数の 1 カテゴリー当りの計算量は、 n 次元の内積計算 1 回 (積和演算が n 回) と W_{l0} の加算 1

回である。(2乗) ユークリッド距離の計算量は減算・積和演算が n 回であり、シティーブロック距離の計算量は減算・絶対値演算・加算が n 回である (実際に要する計算時間は、乗算、加減算、絶対値演算の実行時間に依存するが、積和演算命令をもつ浮動小数点演算装置を備えたワークステーション等においては、線形識別関数の計算時間が最も少ない。実際にサンマイクログシステムズ社のワークステーションを用いた簡単な実験で計算時間を計測したところ、線形識別関数に比べて、(2乗) ユークリッド距離が約 2 倍、シティーブロック距離が約 3 倍の計算時間を要した)。

一方、これらの識別関数の識別性能 (大分類の効率) は、一般に線形識別関数が最も高く、ユークリッド距離とシティーブロック距離はほぼ同程度となる。

ところで、正準判別分析により選択した特徴ベクトルの級内分散は単位行列となるため、式 (3) は次式のように書き直すことができる。

$$g'_l(X) = M_l^T X - \frac{1}{2} M_l^T M_l \quad (4)$$

これは、(2乗) ユークリッド距離を展開した式からクラス間で共通な項を除いたものと同じである。よって正準判別分析により選択した特徴空間においては、ユークリッド距離と線形識別関数は決定則としては等価となり、識別性能も同じになる。更に、ユークリッド距離とシティーブロック距離の識別性能に大差がないことから、シティーブロック距離でも線形識別関数に近い識別性能を得ることができる。以上の考察の結果から、ワークステーションを用いる以下の実験では式 (4) を大分類に用いる。式 (4) の計算に用いる平均ベクトル M_l は、次節の詳細識別で用いる擬似ベイズ識別関数と共用が可能で記憶容量を節約できる。

また、式 (4) を用いることにより、シティーブロック距離に比べて約 3 倍の速度での大分類が可能となる。しかも原特徴空間においては、シティーブロック距離で十分な大分類率を確保するためには候補を上位 100 字種程度選ぶ必要があるが、正準判別分析後の特徴空間では候補字種を 10~20 字種に絞ることができるため、詳細識別に要する時間も短縮することができる。

2.5 詳細識別

詳細識別には次式で定義される擬似ベイズ識別関数 [11], [12] を用いた。

$$g(X) = (N + N_0 + n - 1) \ln \left[1 + \frac{1}{N_0 \sigma^2} \left\| X - M \right\|^2 - \sum_{i=1}^k \frac{\lambda_i}{\lambda_i + \frac{N_0}{N} \sigma^2} \{ \Phi_i^T (X - M) \}^2 \right] + \sum_{i=1}^k \ln \left(\lambda_i + \frac{N_0}{N} \sigma^2 \right) \quad (5)$$

ここで、 X は入力文字の特徴ベクトル、 n は特徴ベクトルの次元数、 N は各クラスの学習サンプル数、 M は平均ベクトル、 Φ_i 、 Λ_i はそれぞれ共分散行列の第 i 固有ベクトルと第 i 固有値、 N_0 は σ^2 の信頼度定数、 k は識別に用いる固有ベクトルの数である。また、 σ^2 は事前分布を等方的正規分布と仮定した場合の分散で、その値として以下の実験では、全字種、全特徴量の分散の平均を用いる。加重方向指数ヒストグラムは物理的に同一種類の特徴によって構成されているため、事前分布を等方的正規分布と仮定する。必要な計算量・記憶容量は、線形識別関数の約 k 倍、2次識別関数の約 k/n 倍であり、部分空間法と同程度である。また、共分散行列の推定誤差に起因する性能低下が少ないため、文字認識等への応用に適している。更に、部分空間法に比べクラス間の共有部分空間における識別率の低下が少ない[12]ため、字種の多い漢字認識に適している。

3. 予備実験

手書き漢字・平仮名の認識に適した特徴選択法、非線形正規化法、原特徴量の高次元化、および大分類の手法を決定するための予備実験についてそれぞれ3.2～3.5に述べる。また、3.3では非線形正規化による特徴量の分離性と分布形の変化を調べ、非線形正規化の効果について考察する。

3.1 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」[15]を横64×縦63画素に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。各字種から1文字ずつ取り出したものをまとめて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)。また全20巻をA(1, 2, 19, 20巻)、B(3～6巻)、C(7～10巻)、D(11～14巻)、E(15～18巻)の五つのブロックに分ける。

以下3.2～3.5の予備実験では、ブロックB、C、D、Eの計160文字/字種を学習用サンプル、ブロックAの40文字/字種を評価用サンプルとして用いる。

3.2 特徴選択手法の比較

筆者らは文献[12]で、特徴選択の手法として主成分分析と正準判別分析を比較した場合、特徴ベクトルの次元数に比べてカテゴリ数の少ない数字認識では前者が適していることを示した。しかし、カテゴリ数が多い漢字・平仮名に対する認識には、正準判別分析の方が適していると考えられる。本節ではETL9B 3036字種を対象とした認識実験で両者を比較する。

文字画像に対し位置・大きさの正規化を行った後、196次元の加重方向指数ヒストグラム[10]を抽出し、変数変換により分布を正規分布に近づける。正準判別分析、主成分分析を用いて原特徴ベクトルから196次元、144次元、100次元、64次元の特徴ベクトルを作成し、擬似バイズ識別関数により識別を行った結果を図1(a),(b)に示す。同図(c)には特徴選択を行わず特徴抽出時の領域分割数を変化させることで求めた196次元、144次元、100次元、64次元の原特徴量に対する正読率も併せて示してある。擬似バイズ識別関数の計算に用いる固有ベクトル数は40とし、非線形正規化は行っていない。

特徴選択なしの場合は次元数の減少と共に正読率が大きく低下するが、主成分分析を用いて特徴選択した場合には正読率の低下が緩やかである。144次元の正読率は196次元と大差がない。更に正準判別分析を用いた場合には、より低次元の特徴ベクトルにおいて196次元の原特徴量よりも高い正読率が得られている。この理由については、5.で考察する。

以上の結果から次のことがわかる。

(1) 同じ次元数ならば、その次元数の原特徴量よりも、より高次元の原特徴量から特徴選択した圧縮特徴量を使用する方が正読率が高い。

(2) 字種の多い手書き漢字・平仮名に対する特徴選択には、主成分分析より正準判別分析の方が有効である。

3.3 非線形正規化手法に関する予備実験

3.3.1 非線形正規化手法の比較

手書き文字パターンの変形吸収のために、線密度の

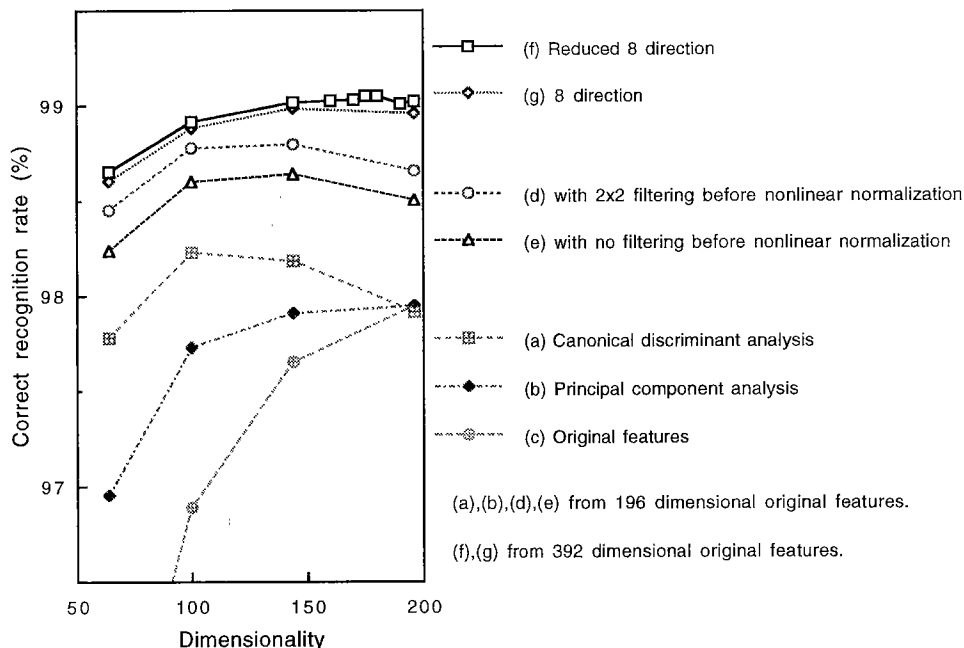


図1 予備実験の結果
Fig.1 Results of preliminary experiments.

特性量を利用する非線形正規化法がいくつか提案されている。本研究では代表的な手法として、山田の線密度イコライゼーション[4] (手法1)、津雲のストローク間隔による正規化法[5] (手法2)、山田の改良型非線形正規化法[6] (手法3)を選び、比較実験を行った。

手法1では、線密度として縦方向および横方向の走査線と黒成分との交差数を定義し、線密度を均一化するように文字画像を正規化する。一方、手法2では線間隔の逆数を局所的な線密度として定義し、同様に正規化を行う。また手法3でも線間隔の逆数としての局所的な線密度を用いるが、手法1・手法2が縦方向、横方向でそれぞれ独立に非線形変換関数を求めるのに対し、両方向の線密度を統合した2次元局所的線密度を用いて変換関数を作成する点が異なっている。

実験データには平仮名を含むETL9Bの先頭300字種を用いる。文字画像に対し各手法で非線形正規化を行った後、64次元の加重方向指数ヒストグラム特徴[10]を抽出し、擬似バイズ識別関数を用いて識別を行う。非線形正規化なしの場合を含めて行った実験の結果を表1に示す。但し、識別に用いる固有ベクトル数は20で、非線形正規化前の2x2の平均値フィルタリングと特徴ベクトルの変数変換は行っていない。

表1 非線形正規化手法の比較実験 (300字種)
Table 1 Comparison on nonlinear normalization (300 categories).

	非線形正規化手法			
	なし	手法1	手法2	手法3
正読率 (%)	97.55	97.81	98.07	98.17

表1の結果から以下のことがわかる。

(3) 輪郭線の方向を特徴量とする統計的的文字認識手法においても、文字画像の非線形正規化が有効である。

(4) 三つの非線形正規化法の中では、手法3を用いた場合に最も高い正読率が得られた。

よって本研究では、非線形正規化手法として手法3を用いることにした。

3.3.2 非線形正規化前の平滑化の効果

非線形正規化前の平滑化の効果を調べるために、ETL9B 3036字種を対象とした認識実験を行った。非線形正規化前の平滑化を行う場合と行わない場合について3.2と同様に196次元の特徴ベクトルを求め、正準判別分析による特徴選択で196次元、144次元、100次元、64次元の特徴ベクトルを作成し、認識実験を行った。結果を図1(d),(e)に示す。識別に用いる固有ベクト

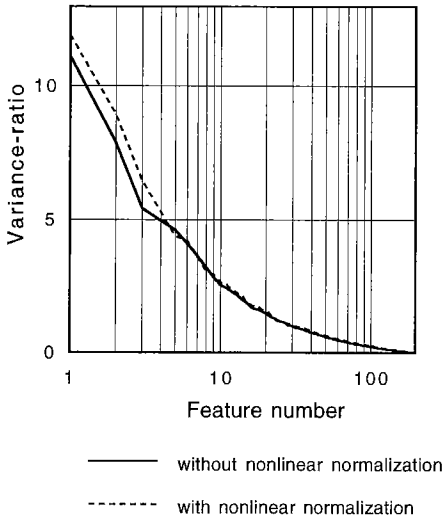


図2 正準判別変量の分散比
Fig.2 Variance-ratio of canonical discriminant variables.

ル数はここでは80(64次元の場合には64)とした。

図1(d),(e)の結果を3.2の結果(同図(a),(b),(c))と比較すると、使用する固有ベクトル数に違いはあるが、非線形正規化の効果が大きいことがわかる。また、(5)非線形正規化前にも平滑化を行った方が文字の輪郭が滑らかに保たれ、正読率が高いことがわかる。

3.3.3 特徴量の分離性の改善効果

非線形正規化による特徴量の分離性の改善効果を調べるために、196次元の特徴ベクトルから得られる正準判別変量の分散比(式(1)の λ_i)を非線形正規化手法を用いる場合と用いない場合について比較した(図2)。この結果から、(6)非線形正規化により分散比が増大し、分離性がよくなっていることがわかる。

3.3.4 特徴量の分布形の改善効果

非線形正規化による特徴量の分布形の改善効果を調べるために、196次元の特徴ベクトルの各特徴量に対して分布のひずみ度・尖度を要素ごとに計算し、横軸に要素番号、縦軸にひずみ度・尖度を図示した(図3、図4)。ひずみ度は分布の対称性を示す値で、正規分布では0になる。尖度は分布のすその広がりを示す値で、正規分布では3になる。0付近と3付近の破線は、それぞれひずみ度と尖度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。

図3は160個のサンプルを用いて計算した漢字“霞”に対するひずみ度・尖度で、(a)が変数変換・非線形正規化なし、(b)が変数変換のみあり、(c)が変数変換・非

線形正規化ありの場合を示している。図4は同様に漢字“林”に対するひずみ度・尖度を示している。

“霞”は縦横のストロークが多く、非線形正規化が比較的容易な字種と考えられる。“林”はストローク数が少なく斜めのストロークの割合が大きい、同様に非線形正規化の効果が表れている。

図3、図4を見ると、次のことがわかる。

(7)変数変換により分布は正規分布に近づくが、非線形正規化により更に分布形が改善される。

擬似ベイズ識別関数は、正規分布に対して準最適であることから、正規性の改善が正読率の向上に寄与していると考えられる。

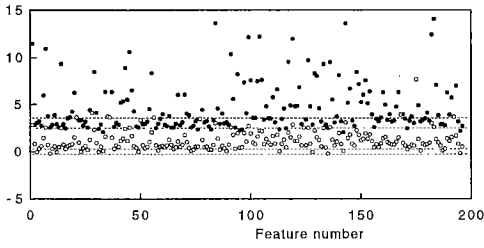
3.4 原特徴量の高次元化(方向量子化数の増加)

3.2(1)によれば、原特徴量の次元数をより大きくすれば更に正読率が向上する可能性がある。一方、予備実験の結果では、196次元(縦7×横7×4方向)の特徴ベクトルに対して、これ以上領域分割数を増やしても効果がないことがわかっている。そこで次元数増加の手段として方向量子化数の増加[11]を試みた。非線形正規化の後、8方向と縮小8方向の、392次元(縦7×横7×8方向)の特徴ベクトルを作成する。ここで8方向とは、輪郭点のチェーンコードを向きを考慮した8方向に量子化したもの、縮小8方向とは2.2に示したように、向きの違いを含めて16方向に量子化して求めたヒストグラムに対し、[1 2 1]の加重フィルタを用いて8方向に縮小したものである。

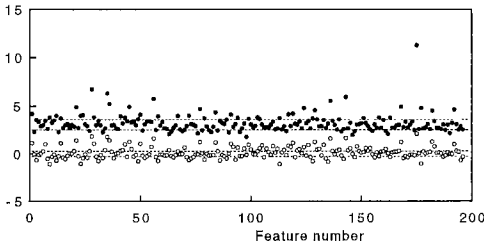
特徴ベクトルに変数変換を施した後、正準判別分析により求めた変換行列を用いて、196次元、144次元、100次元、64次元の特徴ベクトルを作成し、識別を行う。認識実験の結果を図1(f),(g)に示す。識別に用いる固有ベクトル数は80(64次元の場合には64)である。

3.3.2の結果(図1(d),(e))と比較すると、方向量子化数を4方向から8方向に増やして原特徴量の次元数を2倍にしたことにより、同じ次元数でも正読率が大きく向上していることがわかる。また、8方向より縮小8方向の正読率が高いことも示されている。縮小8方向に関しては、より適した次元数を求めるために160次元、170次元、175次元、180次元、190次元の特徴ベクトルに対しても認識実験を行った。その結果175次元に正読率のピークが見られた。この結果から、

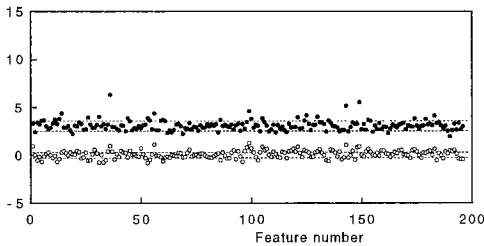
(8)手書き漢字・平仮名を対象とした認識実験においても、方向量子化数の増加による原特徴量の次元数増加が正読率向上に有効であることがわかった。



(a) Original features.



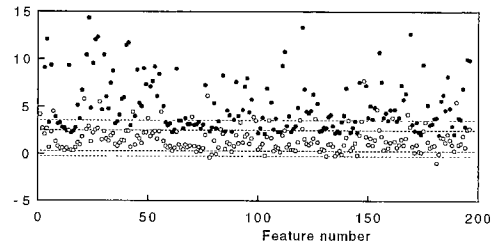
(b) Features with variable transformation.



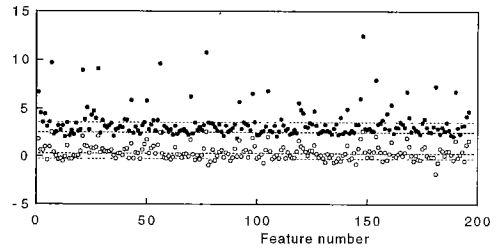
(c) Features with nonlinear normalization and variable transformation.

○ Skewness ● Kurtosis

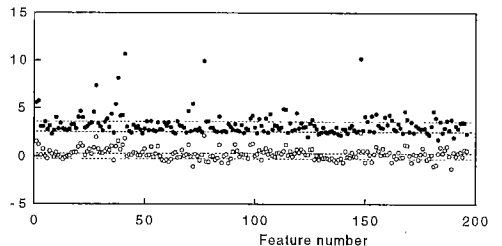
図3 特徴量のひずみ度と尖度 (“霞”)
Fig.3 Skewness and kurtosis of features.



(a) Original features.



(b) Features with variable transformation.



(c) Features with nonlinear normalization and variable transformation.

○ Skewness ● Kurtosis

図4 特徴量のひずみ度と尖度 (“林”)
Fig.4 Skewness and kurtosis of features.

3.5 大分類の予備実験

正準判別分析による特徴選択の前後における、線形識別関数、ユークリッド距離、シティーブロック距離による識別率を表2に示す。実験データにはETL9Bの先頭600字種を用いた。また原特徴量は、非線形正規化前に平滑化を行った場合の196次元特徴ベクトルである。

この結果から、(9)原特徴量に対しては線形識別関数の識別率が最も高いが、特徴選択後はユークリッド距離の識別率が線形識別関数と同じになり、シティーブロック距離の識別率も向上していることがわかる。

表2 大分類に用いる関数のETL9B(先頭から600字種)に対する識別率(%)

Table 2 Performance of pre-classifiers (%) (600 categories).

	特徴選択後の次元数				原特徴
	64	100	144	196	196
線形識別関数	98.20	98.43	98.50	98.45	98.45
ユークリッド距離	98.20	98.43	98.50	98.45	97.78
シティーブロック距離	97.74	98.05	98.09	98.03	97.50

4. ETL9B全セットに対する認識実験

予備実験の結果により、2.で述べた非線形正規化と特徴量の圧縮に基づく認識手法の有効性が確認された。ここではETL9B全セットに対する認識実験を行

い、学習文字数と正読率の関係を調べ、本手法により高い正読率が得られることを示す。

4.1 学習文字数と正読率の関係 (実験1)

学習文字数と正読率の関係を調べるために、1字種当りの学習文字数として(a) 100文字、(b) 160文字、(c) 190文字の三つの場合の認識実験を行った。

(a)では ETL9B 全 20 巻を奇数巻と偶数巻の 2 ブロックに分け、ローテーション法を用いる (あるブロックの評価時にはそれ以外の全ブロックを学習用として認識実験を行い、全ブロックの平均正読率を求める)。

(b)では A, B, C, D, E の五つのブロックに対してローテーション法を用いる。

(c)では全 20 巻を 20 ブロックとしてローテーション法を用いる。特徴ベクトルは 392 次元から 175 次元に圧縮したものを用い、識別に使用する固有ベクトル数は 100 とした。

なお、4.2 で述べる α については、事前の調査で最も効果の高かった 1.25 を用いる。

4.2 非線形正規化による学習サンプルの増加 (実験 2)

山田の改良型非線形正規化法では、標本点 (x_p, y_p) の 2 次元局所的線密度 $\rho(i, j)$ を求めた後、次式で x 方向、 y 方向への線密度の投影 h_x, h_y を求め、それを平均化するように画像の変換を行っている。但し、 $i=1, 2, \dots, I, j=1, 2, \dots, J$ である。

$$h_x(x_i) = \sum_{j=1}^J \rho(i, j) + \alpha$$

$$h_y(y_j) = \sum_{i=1}^I \rho(i, j) + \alpha \quad (6)$$

ここで α は非線形正規化の度合を調節するパラメータで、値が大きいほど非線形性が弱まり、線形変換に近づく。図 5 は α の値として 1.25, 6.25, 1000 を選んだ場合の正規化画像を示している。 $\alpha = 1000$ ではほぼ線形変換になっている。

このように変換の非線形性を変化させることで、一つの画像から正規化の程度が異なる複数の画像を発生させることができる。これらの正規化画像から特徴抽出して学習用の特徴ベクトルを求めることにより、変形が大きく、非線形正規化だけでは十分正規化できない入力文字の識別誤りを減少させる効果が期待できる。そこで学習用データに対しては、 α を 1.25, 6.25, 1000 とした 3 通りの前処理をした後、各々の正規化画像から特徴抽出して参照ベクトル (平均ベクトル, 固有ベクトル) を求め、評価用データに対しては、 $\alpha =$

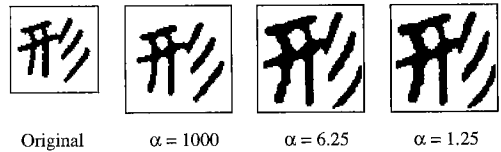


図 5 パラメータ α に対する非線形正規化画像の変化
Fig.5 Nonlinear normalization for different values of α .

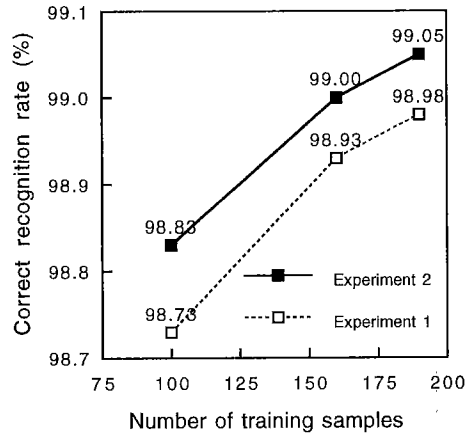


図 6 全セットの平均正読率
Fig.6 Average recognition rate over all sets.



図 7 誤読文字例
Fig.7 Examples of misrecognized samples.

1.25 とした前処理した後、特徴抽出して認識実験を行った。但し、特徴選択用の変換行列を計算する際には、 $\alpha = 1.25$ の正規化画像のみを用いた。また、4.1 と同様に、学習文字数が(a) 100文字、(b) 160文字、(c) 190文字の三つの場合について評価した。

4.3 実験結果と考察

実験 1, 実験 2 の結果を図 6 に示す。非線形正規化のパラメータを変化させて人工的に作成した文字画像を学習に加えることで、(a), (b), (c) のいずれにおいても正読率が 0.06~0.1% 程度向上していることがわかる。学習文字数が 190 の場合には、実験 1 (c) で 98.98%

表3 処理時間の内訳
Table 3 Breakdown of processing time.

	特徴抽出	特徴選択	識別	計
処理時間 (ms/文字)	12.7	5.1	95.3	113.1

であった正読率が、実験2(c)で99.05%となった。

また、実験2(c)の誤読文字5781サンプルを目視により確認したところ、筆記者の記入欄間違い、誤字（類似文字、存在しない文字等）によって誤読となったものが約700サンプル存在した。これらの誤記を除外して正読率を計算すると、99.19%となる。

図7に誤記以外の誤読文字例を示す。つぶれやかすれ、著しい変形のために、これらのサンプルの中には人間でも判別が困難なものが少なくない。

処理時間の内訳を表3に示す。認識速度は約9文字/秒である。計測には、サンマイクロシステムズ社のSparc Station 20 model 61を用いた。

5. む す び

統計的漢字認識の前処理における非線形正規化、原特徴量の高次元化、および特徴選択による認識精度の向上について検討した。主要な結果を以下に要約する。

(1) 輪郭線の方向を特徴量とする統計的漢字認識手法においても、文字画像の非線形正規化が有効であり、その理由として、クラス内分散の低減による分離性の向上と共に、特徴量の分布形が正規分布に近づくことを示した。

(2) 同じ次元数ならば、その次元数の原特徴量よりも、より高次元の原特徴量から特徴選択した圧縮特徴量を使用の方が正読率が高いことを示した。そのための原特徴量の高次元化の手段としては、方向量子化数の増加が有効であることと、字種の多い手書き漢字・平仮名に対する特徴選択には、主成分分析より正準判別分析の方が有効であることを示した。その他、

(3) 大分類に用いる識別関数として、原特徴量に対しては線形識別関数の識別率がユークリッド距離やシティーブロック距離より高いが、正準判別分析による特徴選択後はユークリッド距離の識別率が線形識別関数と同じになり、シティーブロック距離の識別率も向上すること、などを示した。

以上の考察の結果に基づいて構成された識別系を用いて、ETL9Bの全セットを対象とする認識実験を行っ

た結果、これまでに得られている中で最高の正読率（平均99.05%、未学習文字）と9文字/秒の読取り速度が得られた。

ところで、図1(a)の正準判別分析により得られた196次元の特徴空間における正読率が、より低次元の特徴空間より低下している。このような尖頭現象（Peaking phenomenon）が生じるのは一般に、1)共分散行列等のパラメータの推定誤差[10]、2)特徴量の正規分布からのずれ[11]、3)事前分布の非等方性、が次元数と共に増加するためと考えられる。特に、正準判別分析は正規直交変換ではないため、原特徴量の事前分布が等方的正規分布でも、変換後は非等方的となる。実験結果では、正準判別分析による特徴選択の場合だけ尖頭現象が起きていることから、事前分布の非等方性が主な原因として考えられる。この問題を解決すればより低次元の特徴空間における正読率も更に向上する可能性がある。この問題については、他の特徴選択手法[16]との比較・検討と共に今後考察する予定である。

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、熱心に討論をして頂いた研究室の皆様に深く感謝致します。

文 献

- [1] J. Kanai, T.A. Nartker, and S.V. Rice, "Performance metrics for document understanding systems," Proc. 2nd International Conference on Document Analysis and Recognition, pp.424-427, 1993.
- [2] S.V. Rice, "An evaluation of OCR accuracy," 1993 Annual Research Report, ISIR, UNLV.
- [3] (編)小川英光, "パターン認識・理解の新たな展開; 挑戦すべき課題," 電子情報通信学会, pp.94, 1994.
- [4] 山田博三, 齊藤泰一, 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.
- [5] 津雲 淳, "手書き漢字認識における非線形正規化処理," 昭62信学情報・システム大全, 68.
- [6] 山田博三, 齊藤泰一, 山本和彦, "非線形正規化法の改良," 昭63信学春季大会, D-439.
- [7] Y. Deng, 若林哲史, 鶴岡信治, 木村文隆, 三宅康二, "手書き文字認識における非線形正規化法の改良," 信学技報, PRU93-126, Jan. 1994.
- [8] 孫 寧, 安部正人, 根本義章, "改良型方向線素特徴量および部分空間法を用いた高精度な手書き文字認識システム," 信学技報, PRU94-31, 1994.
- [9] 孫 寧, 安部正人, 根本義章, "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," 信学技報, PRU94-94, Dec. 1994.
- [10] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二, "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," 信学論(D-II), vol.J76-D-II, no.12, pp.2495-2503, Dec. 1993.

- [11] 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.
- [12] 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.
- [13] 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), vol.J70-D, no.7, pp.1390-1397, July 1987.
- [14] K. Fukunaga, “Introduction to Statistical Pattern Recognition, Second Edition,” Academic Press, 1990.
- [15] 斉藤泰一, 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, no.7, pp.487-525, July 1985.
- [16] T. Okada and S. Tomita, “An optimal orthonormal system for discriminant analysis,” Pattern Recognition, vol.18, no.2, pp.139-144, 1985.

(平成7年5月19日受付, 10月23日再受付)



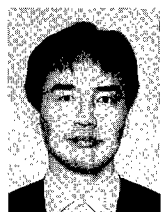
木村 文隆 (正員)

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1~3ミシガン大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.



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.



鄧 巖

1984中国北京大学・理・計算数学卒. 同年, 中国天津市United Software Company入社. 計算方法に関する研究に従事. 1985天津市医学計算センター助手. 統計分析に関する研究に従事. 平3来日. 平6三重大大学院修士課程了(電子工学). 手書き文字認識に関する研究に従事. 同年, (株)システムサーブ・ジャパン入社. 平8(株)ジャストシステム入社, 現在に至る. システム開発に従事.



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.