

こう配特徴ベクトルと変動吸収共分散行列による手書き漢字認識の高精度化

澤 和宏^{†*} 若林 哲史[†] 鶴岡 信治[†] 木村 文隆[†]
三宅 康二^{††}

Accuracy Improvement by Gradient Feature and Variance Absorbing Covariance Matrix in Handwritten Chinese Character Recognition

Kazuhiro SAWA^{†*}, Tetsushi WAKABAYASHI[†], Shinji TSURUOKA[†],
Fumitaka KIMURA[†], and Yasuji MIYAKE^{††}

あらまし 筆者らは、前報で濃淡画像のこう配から抽出した特徴ベクトルを用いることにより、手書き数字認識の高精度化が可能になることを示した。本論文では、手書き漢字を対象としたこう配特徴ベクトルの抽出方法と変動吸収共分散行列による強化学習法を提案し、その有効性を評価する。こう配特徴ベクトルの抽出においては、加重フィルタの精細化、画像の拡大が認識精度向上に有効であることも示す。変動吸収共分散行列とは、ずれや傾きなどすべての字種に共通な傾向をもって生じる誤差の変動を先験的知識として標本共分散行列に加えたものであり、学習標本の範囲外にある文字パターンの認識をねらいとした強化学習に利用できる。これらの提案手法を用いて ETL9B 全セットを対象とした認識実験を行ったところ、未学習文字に関して 99.41% の認識率が得られた。

キーワード 手書き漢字認識、特徴抽出、こう配、ETL9B、変動吸収特性核

1. ま え が き

本論文では、こう配特徴ベクトルと変動吸収共分散行列による手書き漢字認識の高精度化について考察する。

文字認識に用いる特徴ベクトルの抽出では、標準化、量子化、ブロック化(領域分割)などによって離散化とデータの圧縮・削減が行われることが多いが、これらの処理によって、文字の詳細な形状特徴が失われて類似文字等の分離性が悪化し、認識率の低下を招いている可能性がある。このような問題に対処するために、手書き漢字を対象としたこう配特徴ベクトルの抽出方法と、ブロック化のアンチエイリアスに用いる加重フィルタの精細化について検討を行う。筆者らは、

文献 [1] で濃淡画像のこう配から抽出した特徴ベクトルを用いることにより、手書き数字認識の高精度化が可能になることを示した。このときこう配を用いた方向量子化数の増加による特徴ベクトルの次元数増加が有効であった。しかし、ストローク数が少なく曲線部分が多い手書き数字認識と、ストローク数が多く曲線部分の少ない手書き漢字認識とは、各々適した文字画像の正規化手法や空間分割数及び方向量子化数が異なる。

一方、手書き漢字はクラス数が多く、すべてのクラスに十分な学習サンプルを収集することが困難である。手書き漢字データベース ETL9B においても 1 クラス当りのサンプル数は 200 程度であり、特徴ベクトルの高次元化による認識精度の向上を阻んでいた。そのため文献 [2] では正準判別分析による特徴ベクトルの次元圧縮と摂動非線形正規化による人工的な学習サンプルの生成方法及びその有効性について考察した。

また、文献 [2] では文字画像の非線形正規化が統計的手法による手書き漢字認識にも有効であることを示したが、今までの非線形正規化手法では文字の傾きが

[†] 三重大学工学部，津市
Faculty of Engineering, Mie University, Tsu-shi, 514-8507
Japan

^{††} 中部大学工学部情報工学科，春日井市
Department of Computer Science, College of Engineering,
Chubu University, Kasugai-shi, 487-8501 Japan

* 現在，三重県立津高等技術学校電子制御情報科

十分に補正されない．この問題に対して，文献 [3] では，部分傾き検出による画像変換を用いて文字周辺部の傾きによる変動を抑え，ETL9B に対して 99.08% の認識率が得られている．更に文献 [4] では，類似文字間の差違を考慮した混合マハラノビス関数を提案し 99.31% の認識率が得られている．

個人の書き癖による文字のずれや傾きなどの変動はすべての字種に共通な傾向をもって現れる．これらの変動に対処するには，上記の非線形正規化や傾き補正のように画像処理的に評価パターンの変動を少なくする手法のほかに，認識辞書に変動を学習させるアプローチがある．単純な方法としては学習文字画像に傾きやずれなどの変形を加えた人工的なサンプルを発生させ，それを学習に用いる方法が考えられるが，文献 [5] では自己相関行列を変動に関する知識を利用して変形した変動吸収特性核を提案し，平仮名を対象とした認識実験でその効果を確認している．文献 [5] では，濃淡画像の変形としてブロック単位の変形のみを仮定して 2 値 (0 または 1) の変形行列を用いているが，本論文では，ブロック化されたこう配画像の 1 ブロック以下の変動にも対処するために，変形行列を実数行列に拡張することを試みる．

本論文の 2. では手書き漢字を対象としたこう配特徴ベクトルの抽出方法と加重フィルタの精細化及び画像の拡大，3. では変動吸収特性核を発展させた変動吸収共分散行列による強化学習法を提案し，それらの有効性を ETL9B [6] に対する認識実験により評価する．

なお本論文では以降，正準判別分析による次元圧縮前の特徴ベクトルを特に原特徴ベクトルと呼び，次元圧縮後の特徴ベクトルと区別する．また，原特徴ベクトルを抽出するプロセスを原特徴抽出と呼ぶことにする．

2. 原特徴抽出の改良

2.1 こう配特徴ベクトルの抽出

以下に濃淡画像のこう配から原特徴ベクトルを抽出する従来手法の手順を示す．画像の正規化サイズ，原特徴ベクトルの次元数等は，文献 [7] に示される文字輪郭点のチェーンコードを用いた原特徴抽出と同じである．

(1) 原画像 (2 値) を 2×2 の平均値フィルタで平滑化し濃淡画像を作成する．原画像のサイズは後述の ETL9B の場合，横 64 \times 縦 63 である．

(2) 山田の改良型非線形正規化 [8] により原画像

から 2 次元的線密度を求め，線密度の縦横方向への投影をそれぞれ均等化する変換関数を作成し，濃淡画像を非線形に正規化する．正規化画像のサイズは 78×78 とする．78 は後述の (7) のブロック数 13 の倍数となるよう定めたものである．

(3) 3×3 の平均値フィルタを施して正規化画像を平滑化し，濃淡画像を得る (図 1 (a))．

(4) 平均値が 0，最大値が 1 となるように濃淡画像の濃度値を正規化する．

(5) Roberts フィルタにより，濃度値曲面のこう配の強さと向きを求める (図 1 (b), (c))．

(6) こう配の方向を $\pi/16$ 刻みで向きを考慮した 32 方向に量子化する．

(7) 量子化した方向別に 32 枚の画像を作成し，各画像を 13×13 のブロックに分割すると，1 ブロック当り 6×6 画素となる．各ブロック内でこう配の強さを加算して， $32 \times 13 \times 13$ 次元の局所方向ヒストグラムを作成する．

(8) 加重フィルタ $[1 \ 4 \ 6 \ 4 \ 1]$ を 1 方向おきに掛けることにより，ヒストグラムを 32 方向から 16 方向に再標本化し，更に加重フィルタ $[1 \ 2 \ 1]$ を 1 方向おきに掛けて 16 方向から 8 方向にする．

(9) 5×5 の 2 次元ガウス型分布に基づく空間的加重フィルタ (図 2 (a)，以後 2 次元ガウスフィルタあるいは単にガウスフィルタと呼ぶ) を縦横 1 ブロックおきに施して (図 2 (b))，ブロック数を 13×13 から 7×7 に削減し，392 次元 ($8 \times 7 \times 7$) の原特徴ベクトルを求める．

(10) 変数変換 (べき乗) により，特徴量の分布を正規分布に近づける [9]．

2.2 空間分割数の増加とガウスフィルタの精細化
本節では，前節の従来手法に対して，(7) の空間分

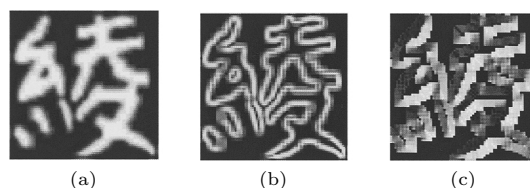


図 1 濃淡画像とこう配 ((a) 非線形正規化とフィルタリング処理後の濃淡画像，(b) こう配の強さ，(c) こう配の向きのシェーディング表示)

Fig. 1 Gray scale image and gradient ((a) gray scale image after non-linear normalization and filtering, (b) strength of gradient, (c) direction of gradient (shading representation)).

割数 (ブロック数) を増加させ, (9) で用いるガウスフィルタを精細化する方法について検討する.

サンプリング定理によれば, ナイキスト周波数 (連続時間信号に含まれる上限の周波数の 2 倍の周波数) より低い周波数でサンプリングを行うと, エリアシングと呼ばれる重なりによる余分な周波数成分を含んでしまう現象を生じる. 画像の縮小 (標本点数の削減) を行う場合, 単純サンプリング (ある領域内の一点を代表させる方法) では高域成分をもつ画像に対して美しい画像が得られず, ジャギーやモアレ等の雑音を生じるのはこのためである. これに対しては, 高域成分を除去する平滑化フィルタを施してから単純サンプリングを行えば良質な縮小画像が得られることが知られている.

2.1 では, 方向ごとの画像に対して, (7) で空間分割により標本点数を縦横それぞれ $1/6$ ($78 \times 78 \rightarrow 13 \times 13$) に削減し, 更に (9) で互いに重なりのある 2 次元ガウスフィルタによって縦横それぞれ約 $1/2$ ($13 \times 13 \rightarrow 7 \times 7$) に削減する処理を行っている. 空間分割による標本点数の削減は, 一様重みフィルタによるフィルタリング処理後の単純サンプリングと考えることができる. 同様にガウスフィルタによる標本点数の削減も, ガウスフィルタによるフィルタリング処理後の単純サンプリングである. ただし空間分割においては, 低いサンプリング周波数で標本点数を大きく削減するにもかかわらず, サンプリングする標本点に対応する一様重みフィルタの領域が 6×6 画素と小さく互いに重なりがないため, 高域成分除去が不十分となる. 標本点数の削減時に用いる低域フィルタとしてはガウスフィルタが優れていることが知られており, 良質な原特徴ベクトルを得るためには, 空間分割を行

ずにガウスフィルタのみによる標本点数の削減を行うのが理想的である. しかし, この場合にはフィルタリング処理における計算量の大幅な増加を招くため, ここでは空間分割数の増加とガウスフィルタの精細化によってより良い原特徴ベクトルを抽出する方法を提案する.

具体的には図 3 (b) のように空間分割数を増加させ, 同図 (a) のようなより精細化された 2 次元ガウスフィルタを適用する. 図 3 の例では, 空間分割数は 49×49 であり, 31×31 のフィルタを縦横 7 ブロックおきに施すことにより標本点数は 7×7 に削減される. 78×78 画素の画像を 49×49 ブロックに分割する場合, 画素 (i, j) の属するブロック (k, l) は次式により求めるものとする. ただし, k, l は四捨五入により整数化する.

$$k = \frac{48}{77}(i - 1) + 1$$

$$l = \frac{48}{77}(j - 1) + 1 \tag{1}$$

したがって, 画像のサイズの縦横が空間分割数の整数倍でないときは, 等分割にはならない.

2.3 2次元ガウスフィルタの設計

図 2, 図 3 のように 2 次元ガウスフィルタは互いに重なりをもって配置される. フィルタの設計にあたり, 2 次元ガウス型分布は等方的であるとし, その分散 σ^2 は重なったフィルタの和がほぼ一定になるように, つまり各ブロック成分の原特徴ベクトルに対する寄与度がほぼ均一になるように定める. これは, 場所ごとの最適な寄与度に関する事前知識がない状況下では, もとの情報をできるだけ忠実に反映する再標本化が望ましいとの考えに基づいている.

図 4 に示すように四つの 2 次元ガウス型分布が, A, B, C, D を中心にして互いに重なっている場合を考えると, 点 (x, y) における四つの分布の和は次式で表

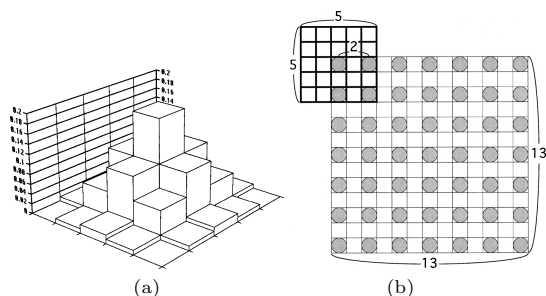


図 2 5 × 5 の 2 次元ガウスフィルタと標本点

Fig. 2 2-dimensional Gaussian-like filter (5 × 5) and sampling points.

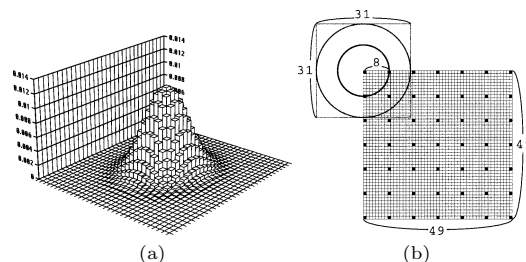


図 3 31 × 31 の 2 次元ガウスフィルタと標本点

Fig. 3 2-dimensional Gaussian-like filter (31 × 31) and sampling points.

される .

$$f(x, y) = g_0(x - a, y - a) + g_0(x + a, y - a) + g_0(x + a, y + a) + g_0(x - a, y + a)$$

$$g_0(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad (2)$$

$-a \leq x \leq a, -a \leq y \leq a$ の領域内で $f(x, y)$ が完全に一定となるのは $\sigma^2 = \infty$ の場合以外には存在しないので、ここでは分布の中心位置 A, B, C, D と点 (0, 0) における四つの分布の和が等しくなるような分散 σ^2 を以下のようにして求める .

$$f(0, 0) = f(\pm a, \pm a) \quad (3)$$

より、

$$\frac{1}{2\pi\sigma^2} \left\{ 4 \exp\left(-\frac{a^2}{\sigma^2}\right) \right\}$$

$$= \frac{1}{2\pi\sigma^2} \left\{ 1 + 2 \exp\left(-\frac{2a^2}{\sigma^2}\right) + \exp\left(-\frac{4a^2}{\sigma^2}\right) \right\} \quad (4)$$

$$t = \exp(-a^2/\sigma^2) \text{ とおくと、}$$

$$(t - 1)(t^3 + t^2 + 3t - 1) = 0 \quad (5)$$

であるから、 $t = 1$ または $t \approx 0.295598$ が得られる . $t = 1$ の場合には $\sigma^2 = \pm\infty$ となるため、後者の数値解を採用すると求める分散は次式で与えられる .

$$\sigma^2 = -\frac{a^2}{\ln t} \quad (t \approx 0.295598) \quad (6)$$

このとき $f(x, y)$ の値は $-a \leq x \leq a, -a \leq y \leq a$ の領域内でほぼ一定となり、 $f(a/2, a/2)$ の値は $f(a, a), f(0, 0)$ と比べ 4.6% だけ大きくなる . 更に分布の中心間の距離 (サンプリング周期) を d とすると $d = 2a$ より、

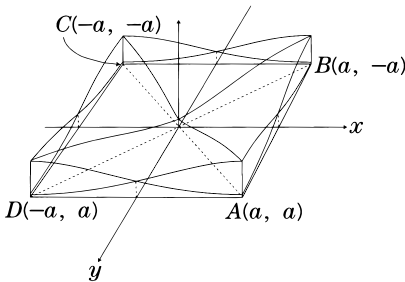


図4 四つのガウス分布の重なり

Fig. 4 Overlapping of four Gaussian filters.

$$\sigma^2 = -\frac{d^2}{4 \ln t} \quad (7)$$

となり分散の値は d^2 に比例して大きくなる .

図2の 5×5 のガウスフィルタは、 $d = 2$ として式(7)より求めた σ^2 と $-2 \leq x \leq 2, -2 \leq y \leq 2$ の座標値を式(2)の $g_0(x, y)$ に代入して作成した . 同様に図3の 31×31 のガウスフィルタは $d = 8$ と $-15 \leq x \leq 15, -15 \leq y \leq 15$ の座標値を用いて作成したものである . ガウスフィルタの縦横のサイズは、その影響範囲が一つ隣の標本点には及ぶが二つ隣の標本点には及ばない範囲で適当に選んだ . これは、式(7)の導出において二つ隣の標本点上のガウスフィルタの影響を考慮していないためであるが、フィルタリングの際の計算量を削減するねらいもある .

2.4 画像正規化サイズの拡大

ここでは文字画像を拡大・平滑化して、擬似的に観測系の精度を向上させた原特徴抽出手法について検討する .

具体的には 2.1(2)において非線形正規化画像のサイズを 147×147 とし、(3)において 3×3 の平均値フィルタを施す回数を3回とする . 拡大画像より求めたこう配の強さと方向を図5に示す . その後は図3のように空間分割数を 49×49 とし、精細化したガウスフィルタによって原特徴ベクトルを求める .

2.5 原特徴ベクトルの比較実験

チェーンコード、こう配のそれぞれに対して、従来のガウスフィルタを用いた場合、精細化したガウスフィルタを用いた場合、画像を拡大したのち精細化したガウスフィルタを適用した場合の3通りの原特徴ベクトルを作成し、手書き漢字データベース ETL9B を対象とした比較実験を行う .

ETL9Bには3036字種を1セットとする文字サン

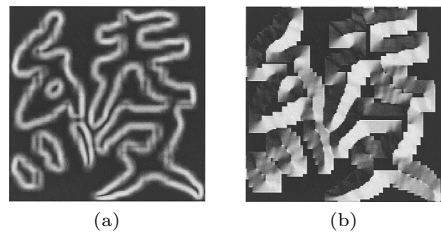


図5 拡大画像 (147×147) より求めた (a) こう配の強さと (b) こう配の向き

Fig. 5 Gradient for enlarged image (147×147) ((a) strength of gradient, (b) direction of gradient (shading representation)).

ブルが 200 セット含まれている．この 200 セットを十等分し，それぞれ 20 セットからなる a, b, c, d, e, f, g, h, i, j を得る．ここでは b~i の 160 セットを学習用に，比較的品質の良い a と品質の悪い j を合わせた 40 セットを評価用に用いることで平均的な認識率を求める．

学習用サンプルより抽出した原特徴ベクトルを用いて正準判別分析を行い，得られた基底ベクトルのノルムを 1 に正規化する．この基底ベクトルの中から判別比の大きい 256 個の基底ベクトルを選択する．392 次元の学習・評価用原特徴ベクトルを選択された基底ベクトルに射影して，256 次元の特徴ベクトルに次元圧縮する．学習用の特徴ベクトルで大分類・詳細識別用辞書を作成し，評価用特徴ベクトルに対する大分類，詳細識別を行う．大分類は線形識別関数で行い，F 比の大きい方から 128 次元の特徴量を用いる．詳細識別は式 (8) の擬似ベイズ識別関数を用いて行う．この識別関数は，分布パラメータのうち母集団の共分散行列が未知の正規分布に対する最適識別関数 [10] から導出された近似式で，識別精度を損なうことなく計算量を大幅に削減できるのが特長である [11] ．

$$g(X) = (N + N_0 + 1) \ln \left[1 + \frac{1}{N_0 \sigma^2} \left[\|X - M\|^2 - \sum_{i=1}^k \frac{\lambda_i}{\lambda_i + \frac{N_0}{N} \sigma^2} \Phi_i^T (X - M)^2 \right] \right] + \sum_{i=1}^k \ln \left(\lambda_i + \frac{N_0}{N} \sigma^2 \right) - 2 \ln P(\omega) \quad (8)$$

ここで X は入力文字の n 次元特徴ベクトル， M は母集団の平均ベクトル， λ_i ， Φ_i はそれぞれ標本共分散行列の第 i 固有値，第 i 固有ベクトル， k は識別に用いる固有ベクトルの数， N は学習サンプル数， $P(\omega)$ はクラス ω の事前確率である．また文献 [10] の最適識別関数はベイズ学習に基づく識別関数であるため事前分布が導入されているが，式 (8) 中の σ^2 は特徴ベクトル X の事前分布を球状と仮定した場合の分散であり， N_0 は σ^2 の信頼度を表す定数で信頼度定数と呼ぶべきものである．ただし，正準判別分析による基底ベクトルは通常（正規でない）斜交系となるため，特徴抽出によって事前分布の球状性が失われてしまう．そこで，基底ベクトルの 2 本ずつの組すべてに対して

内積を求める予備実験を行ったところ，本研究で用いるデータでは基底ベクトルはほぼ直交しており，そのスケールのみが大きく異なる状態であることがわかった．ここでノルムを 1 に正規化すれば，基底ベクトルが正規直交系に近づき，特徴抽出後の特徴ベクトルの事前分布の球状性を極力保存する（ユークリッド距離を極力保存する）ことが可能になると考えられる [7] ．

実験では， M は標本の平均ベクトルで代用し， σ^2 の値としては，全字種，全固有値の平均を用いる．また $P(\omega)$ を含む項はクラス間で共通として省略し，固有ベクトル数は $k = 100$ とする．ただしここでは N_0 の値は，

$$N_0 = \frac{\alpha}{1 - \alpha} N \quad (9)$$

とし， α を $0 < \alpha < 1$ の範囲で変化させて認識率が最も高いものを選ぶことにする．

表 1 に示した実験結果より，チェーンコードよりもこう配から求める原特徴ベクトルを用いた方が認識率が高いことがわかる．また，こう配においてガウスフィルタの精細化と画像の拡大をそれぞれ単独で用いた場合，双方とも 99.31% と同程度の効果が認められる．ところがチェーンコードにおいては，精細化ガウスフィルタ単体では有効性が認められるものの，画像拡大単体では認識率低下を招いている．これはチェーンコードが，拡大によって増幅される輪郭のジャギに対してより敏感であるためと推測される．しかしながら，チェーンコード，こう配ともに，ガウスフィルタの精細化と画像の拡大を併用することで，最も高い認識率が得られている．

図 6 に，チェーンコードより求めた原特徴ベクトル，こう配より求めた原特徴ベクトル，拡大画像のこう配より精細化したガウスフィルタを用いて求めた原特徴ベクトル，の正準判別変量の F 比（判別比）を示す．このグラフから，こう配の利用やガウスフィルタの精細化，画像の拡大により，クラス間の分離性が高い良質な原特徴ベクトルが抽出され，認識率の向上に

表 1 各原特徴ベクトルの認識率 (%)
Table 1 Recognition rates for original feature vectors.

	チェーンコード	こう配
従来のガウスフィルタ (5 × 5)	99.14	99.29
精細化ガウスフィルタ (31 × 31)	99.23	99.31
従来のガウスフィルタ + 画像拡大	99.03	99.31
精細化ガウスフィルタ + 画像拡大	99.26	99.34

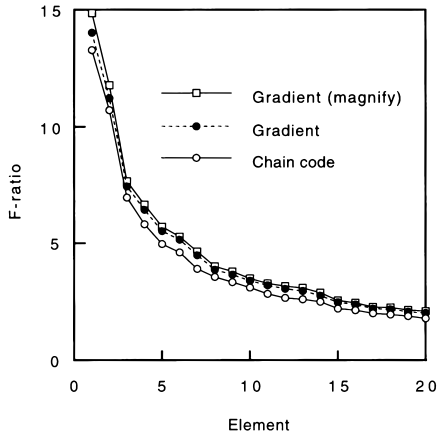


図6 正準判別変量のF比の比較
Fig. 6 Comparison of F-ratios.

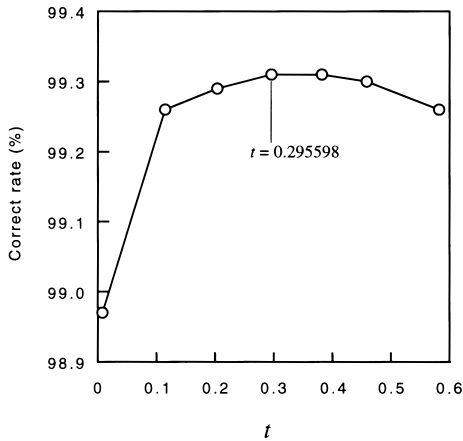


図7 パラメータ t と認識率の関係
Fig. 7 Correct rate vs. parameter t .

寄与したことがわかる。

また図7に、拡大画像のこう配と精細化したガウスフィルタによる特徴抽出において、式(7)の t を変化させた場合の認識率を示す。図より、 $t = 0.295598$ (再標本化の際のガウスフィルタの寄与度がほぼ一定)の場合に、認識率がピークとなっていることが確認できる。

3. 変動吸収共分散行列による強化学習

変動吸収共分散行列とは、ずれや傾きなどすべての字種に共通な傾向をもって生じる誤差の変動を先験的知識として標本共分散行列に加えたものであり、学習標本の範囲外にある文字パターンの認識をねらいとし

た強化学習に利用できる。文献[5]では、濃淡パターンベクトルより求めた自己相関行列に変形行列を作用させ、手書き平仮名認識によってその効果を確認したが、本手法では方向特徴ベクトルより求めた共分散行列に変形行列を作用させ、手書き漢字認識によってその効果を確認する。

3.1 変動吸収共分散行列の定義

あるクラスに属する m 個の学習サンプル X_j ($j = 1, 2, \dots, m$) に対する標本共分散行列 Σ は次式で定義される。

$$\begin{aligned} \Sigma &= \frac{1}{m} \sum_{j=1}^m (X_j - M)(X_j - M)^T \\ &= S - MM^T \end{aligned} \quad (10)$$

ただし、

$$M = \frac{1}{m} \sum_{j=1}^m X_j, \quad S = \frac{1}{m} \sum_{j=1}^m X_j X_j^T$$

で、 S は自己相関行列である。

本節では、特徴ベクトル X_j をそのまま用いるのではなく、 l 種類の変形行列 D_i ($i = 1, 2, \dots, l$) を掛けた $D_i X_j$ に対して次式の変動吸収共分散行列 Σ^* を定義して用いる。

$$\Sigma^* = \frac{1}{ml} \sum_{j=1}^m \sum_{i=1}^l (D_i X_j - M^*)(D_i X_j - M^*)^T \quad (11)$$

ただし、

$$M^* = \frac{1}{l} \sum_{i=1}^l D_i M$$

更に式(11)は、

$$\Sigma^* = \frac{1}{l} \sum_{i=1}^l D_i S D_i^T - M^* M^{*T} \quad (12)$$

と変形できるため、 Σ^* を求めるのに $D_i X_j$ を逐一計算する必要はなく、自己相関行列 S と平均ベクトル M 及び M^* を計算しておき、用意した l 種類の変形行列を作用させて平均をとったものから算出すればよい。

3.2 変形行列の作成

2. で抽出した 392 次元の原特徴ベクトルの要素は、横 7 × 縦 7 ブロックの特徴量が 8 方向分順に並んだものとなっている。本論文では簡単のためにまず 7 × 7 ブロックの濃淡パターンの位置的変動に対する変形行列 d_i を作成し、それを次式のように対角上に 8 個配置したものを D_i とする。

$$D_i = \begin{pmatrix} d_i & & & & & & 0 \\ & d_i & & & & & \\ & & \ddots & & & & \\ 0 & & & & & & d_i \end{pmatrix} \quad (13)$$

d_i は 49×49 , D_i は 392×392 の行列となる。より一般的には、方向別に d_i を変えたり、非対角要素に 0 以外の値を入れて輪郭線やこう配の方向的変動を表現することも可能である。

文献 [5] では、 d_i の要素を 1 か 0 の整数値としていたが、本論文では実数値を用いることで 1 ブロック以下の変動にも対応させる。

まず筆者らは、手書き文字でよく見られる変形として図 8 に示すようないくつかの概念的変形を考えその変形行列を作成した。また単位行列（変形なし）も変形行列として用意した。予備的な認識実験において、単位行列と数種類の移動量や角度で発生させた同図 (a) ~ (g) の変形行列を単独で用いた場合、あるいは複数組み合わせ合わせた場合など様々な試行錯誤を行った。ここでは予備実験での結果が良好であった以下の 2 種類の变形行列を紹介する。

(1) 上下左右の平行移動

手書き文字には特有の変動である筆先の流れや止ま

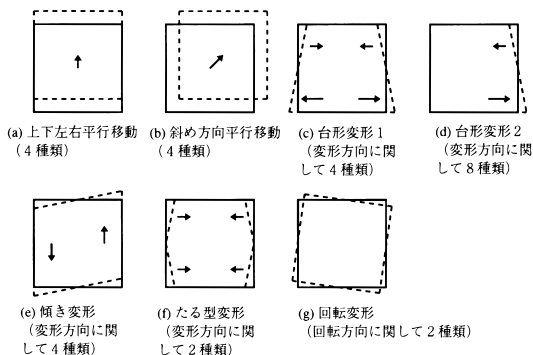
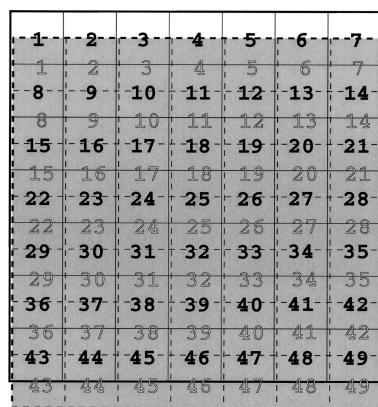


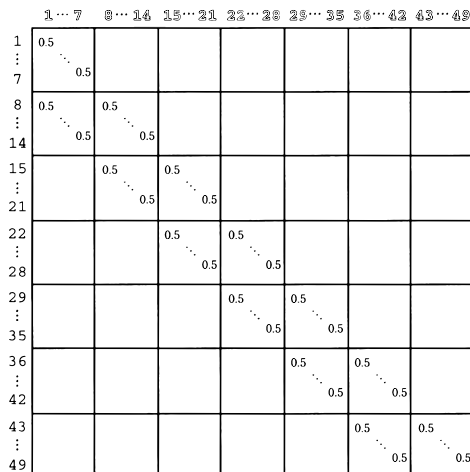
図 8 手書き文字でよく見られる変形

Fig. 8 Representative deformation in handwriting.

り等があり、文字記号として意味のある主要部分が中心より偏った場所に位置するという現象が起きる。この位置ずれは文字画像の非線形正規化によってある程度抑えられるが、十分ではない。そこでこのような位置変動を共分散行列において吸収させる目的で X_i を上下左右の 4 方向にそれぞれ 0.5 ブロック平行移動させる変形行列を作成した。図 9 に下方向へ 0.5 ブロック移動させる変形行列を示す。同図 (a) の点線と網掛けの部分は、下方向への特徴ブロック群の移動を意味する。同図 (b) は (a) に対応する変形行列で、横方向の白抜き文字は移動した特徴ブロック群のブロック番号、縦方向はもとのブロック番号である。行列の要素



(a)



(a)

(b)

図 9 下方向 0.5 ブロックの平行移動変形行列

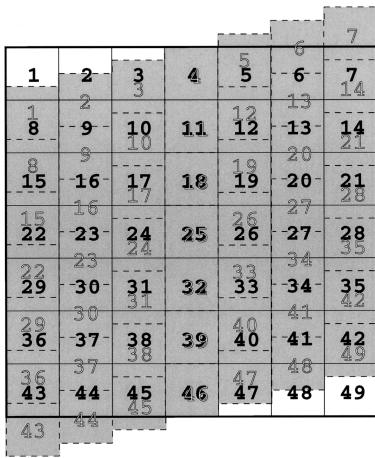
Fig. 9 Deformation matrix for 0.5 block downward translation.

は移動ブロックからもとのブロックへの寄与度を表し、空白部分は0である。

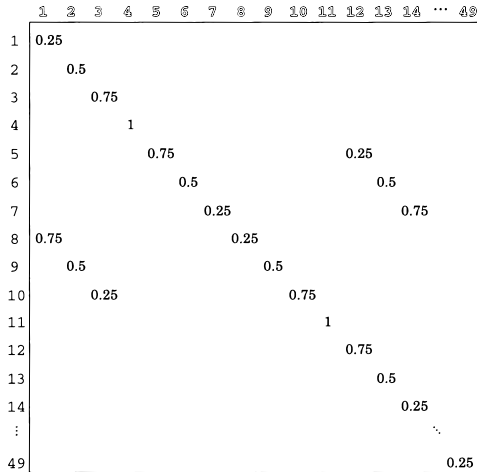
(2) 傾き変形

右上がり、右下がり、左傾き、右傾きなどの個人の癖は、すべての字種に共通の変動を生じさせる。本論文で用いた山田の改良型非線形正規化法ではこれらの傾き変形を十分に補正できない。そこでこの変動を吸収するために上記四つの傾き変形に対して 14.03° の角度をもつ変形行列をそれぞれ作成した。図 10 に右上がり 14.03° の変形行列を示す。

なお、変形によって文字枠外にはみ出す部分は無視



(a)



(a)

(b)

図 10 右上がり 14.03° の傾き変形行列

Fig. 10 Deformation matrix for 14.03 degree right up shearing.

することとした。

3.3 変動吸収共分散行列を用いた認識実験

2.1 の手法で求めたこう配特徴ベクトルを用いて認識実験を行う。認識辞書の作成に用いる標本共分散行列を変動吸収共分散行列に置き換える以外は 2.5 の手順と同じである。変形行列は原特徴ベクトルに作用させなければならないため、正準判別分析により抽出した特徴ベクトルに対応する変動吸収共分散行列 Σ^{**} を求めるには次式による次元削減を行う。

$$\Sigma^{**} = \Psi^T \Sigma^* \Psi \tag{14}$$

ここで $\Psi = [\Psi_1, \Psi_2, \dots, \Psi_{n'}]$ は正準判別分析により求めた基底ベクトル行列で、各基底ベクトルのノルムは 1 に正規化してある。実験では $n' = 256$ とする。

変動吸収共分散行列を、3.2(1)の上下左右の平行移動 4 個と単位行列 2 個の計 6 個の変形行列で作成した場合、同じく 3.2(2)の傾き変形 4 個と単位行列 2 個の計 6 個の変形行列で作成した場合、更に 0.5 ブロックの平行移動 4 個と 14.03° の傾き変形 4 個、そして単位行列 2 個の計 10 個で変動吸収共分散行列を作成した場合について認識実験を行った結果を表 2 に示す。99.40%という認識率は表 1 の標本共分散行列を用いた場合の 99.29%に比べ 0.11%高く、誤読は約 15%減少している。0.5 ブロックの平行移動単独での認識率向上が大きい、この理由として、文字の特徴を決定づけるある画素群が標本点付近にある場合と二つの標本点の中間付近にある場合とで生じる原特徴ベクトルの差異を吸収する効果も推測される。上下左右の平行移動、傾き変形以外の変形行列を単独で単位行列 2 個に加えて用いた場合にも、認識率は 99.31~99.37%の範囲で向上するが、これを上記組合せに追加しても認識率の向上はほとんど見られなかった。他の誤読要因である、つぶれ、かすれ、ノイズなどに対しては、先験的な知識を用いて変動吸収共分散行列を作成するのは困難であるため、これらの変動を含む実際のサンプルまたは画像処理により人工的に発生させたサンプルを学習に加える方が容易である [12]。また、

表 2 変動吸収共分散行列の比較実験結果

Table 2 Comparative experimental results on variance absorbing covariance matrices.

変動吸収共分散行列	認識率 (%)
平行移動 0.5 ブロック	99.39
傾き変形 14.03°	99.34
平行移動 (0.5) + 傾き変形 (14.03°)	99.40

くずし字など筆記者に依存する変形には、特定筆記者の平均ベクトルを加える方法 [13] の利用が考えられる。類似文字間の誤読に対しては、強化学習での対処は難しく、特徴抽出や識別系そのものの改良が必要である。

4. ETL9B 全体での認識実験

ここでは 2.5 で示した a~j のグループを用いるモード 10 のローテーション法で、ETL9B 全体を対象とした認識実験を行う。すなわちある一つのグループ (20 セット) を評価する際には、残りの九つのグループ (180 セット) を学習用サンプルとして認識率を求めるという手順を 10 回繰り返し、平均認識率を求めるといものである。原特徴ベクトルには、(1) チェインコードより求めた原特徴ベクトル、(2) こう配より求めた原特徴ベクトル、(3) 画像拡大後こう配より精細化ガウスフィルタで求めた原特徴ベクトル、の三つを選び、標本共分散行列を用いる場合と変動吸収共分散行列を用いる場合の比較を行う。式 (10) の α の値は、b~i の 160 セットを学習用に、a, j の 40 セットを評価用に用いた予備実験で認識率が最も高くなるように定め、それを全体の評価に用いた。その他の条件は 2.5, 3.3 と同じである。表 3 に ETL9B 全体での認識率比較結果を、表 4 に 1 文字当りの認識時間と認識速度を示す。なお実験はヒューレットパッカー社のワークステーション C180 を用いて行い、処理時間は CPU 時間を計測した。

表 3 ETL9B 全セットに対する認識率 (%)
Table 3 Recognition rates for all set of ETL9B.

	標本共分散行列	変動吸収共分散行列
チェインコード	99.07	99.22
こう配	99.26	99.38
こう配 (画像拡大+精細化 ガウスフィルタ)	99.31	99.41

表 4 1 文字当りの処理時間と認識速度
Table 4 Processing time and recognition speed per character.

	チェインコード	こう配	こう配(画像拡大+精細化ガウスフィルタ)
原特徴抽出	5.9 ms	5.1 ms	28.4 ms
特徴抽出	1.8 ms		
大分類	19.2 ms		
詳細識別	21.2 ms		
計	48.1 ms	47.3 ms	70.6 ms
認識速度 (文字/s)	20.8	21.1	14.2

表 3 より、ETL9B 全セットを対象とした認識実験においても、チェインコードよりこう配から求めた原特徴ベクトルを用いた場合の認識率が高いこと、変動吸収共分散行列を用いる強化学習がどの原特徴ベクトルに対しても有効であることがわかる。その結果、こう配と変動吸収共分散行列を用いた場合の認識率は、従来手法であるチェインコードを用いた場合の 99.07% に比べて、99.38% と大きく向上している。このとき、チェインコードに対してこう配の原特徴抽出時間は約 14% 減少し、変動吸収共分散行列による強化学習は認識時間には影響を及ぼさない。更に画像拡大と精細化ガウスフィルタにより、認識率は 99.41% に向上する。この場合には原特徴抽出に要する計算時間が大幅に増加するが、もともと認識時間全体に占める割合が小さいため、全体では 1.5 倍程度の増加にとどまる。なお、認識率が 99.07% から 99.41% に向上することで、誤読は 37% (2064 文字) 減少したことになる。

図 11 は、こう配特徴ベクトルで標本共分散行列と変動吸収共分散行列を用いた場合の誤読文字数の内訳を示したものである。図中の () 内の数字は、そのうちの平仮名の誤読文字数を示している。全テストサンプル 607,200 文字に対して、変動吸収共分散行列を用いることで新たに認識されたのは 1192 文字であったが、副作用により誤認識されたのは半分以下の 488 文字であった。ただし、平仮名に関しては、新たに認識されたものとはほぼ同数の文字が副作用で誤認識となっている。同様に図 12 は、画像拡大と精細化ガウスフィルタを用いたこう配特徴ベクトルの誤読の内訳である。変動吸収共分散行列により新たに認識されたのは 1082 文字、副作用により誤認識となったのは 431 文字であった。しかも、平仮名で新たに認識されたのが 49 文字であるのに対して、副作用により誤認識となったのは 29 文字にとどまっており、画像拡大と精細化

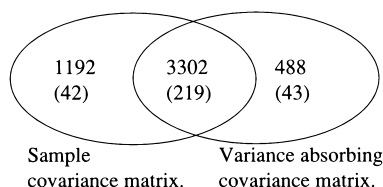


図 11 全テストサンプル 607,200 文字に対する誤読文字数の内訳(こう配)。図中 () 内は平仮名の誤読数
Fig. 11 Results of failure analysis with gradient feature. Numbers in the parenthesis show the errors of hiragana.

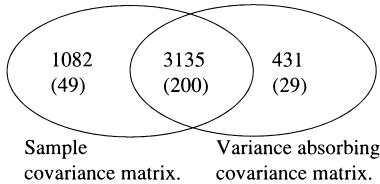


図 12 全テストサンプル 607,200 文字に対する誤読文字数の内訳 (こう配+画像拡大+精細化ガウスフィルタ). 図中 () 内は平仮名の誤読数
 Fig. 12 Results of failure analysis with gradient feature, image magnification and refinement of the weighted filter. Numbers in the parenthesis show the errors of hiragana.



図 13 変動吸収共分散行列により新たに認識された文字 (こう配). 上段は原画像. 下段は非線形正規化画像
 Fig. 13 Samples newly recognized by variance absorbing covariance matrices.

ガウスフィルタを用いれば, 平仮名に対しても変動吸収共分散行列が有効であることがわかった. 以上より本論文における提案手法が手書き漢字認識の高精度化に有効であることが確認できる.

図 13 に, こう配特徴ベクトルで変動吸収共分散行列により新たに認識された文字例を示す. 上段は原画像で下段は非線形正規化画像である. ノイズ除去が不完全で, 非線形正規化後に位置が画像の上下左右に偏ってしまった文字, 傾きの強い文字などが新たに認識されており, 3.2 で示した変形行列の効果が現れている. 図 14 には, 変動吸収共分散行列の副作用で誤認識された文字例を示す. 濁点, 半濁点が不明瞭な平仮名, つぶれた文字などが誤認識となっている. また, 図 15 には, 変動吸収共分散行列を用いても依然として誤認識となる文字例を示す. 極端なかすれや「鳥」と「鳥」や「平」と「乎」などの類似文字, くずし字などに対する誤認識がみられる. このように人間にとっても正しく読むことが困難な文字がかなりある一方で, 人間なら間違いなく読める文字 (図 14 の「ば」, 図 15 の「一」, 「化」など) もかなり残っている. このような機械認識固有の誤読の原因としては, ノイズ等の影響による文字切出しの誤りや非線形正規化の副作用のほか, 原因の特定できない誤読もある. これらの原因を



図 14 変動吸収共分散行列の副作用で誤認識された文字 (こう配). 上段は原画像. 下段は非線形正規化画像
 Fig. 14 Samples newly misrecognized by variance absorbing covariance matrices.

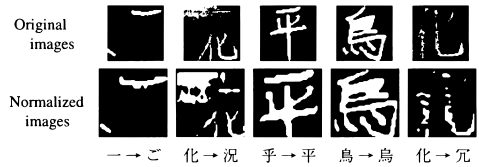


図 15 変動吸収共分散行列を用いても依然として誤認識となる文字 (こう配). 上段は原画像. 下段は非線形正規化画像
 Fig. 15 Error samples by both sample covariance matrices and variable absorbable covariance matrices.

解明し, 更に認識率を向上させて人間の認識率に近づけることは今後に残された興味ある課題の一つである.

5. むすび

本研究では, 手書き漢字を対象としたこう配特徴ベクトルの抽出手法と変動吸収共分散行列による強化学習法を提案し, その有効性を検討した. その結果, 手書き漢字認識においても, チェインコードから求めた原特徴ベクトルより, こう配から求めた原特徴ベクトルを用いた方がクラスの分離性が向上し, 高い認識率が得られ, 処理速度も向上することがわかった. 更に, 加重フィルタの精細化と画像の拡大もクラスの分離性と認識精度の向上に有効であることを示した. また, 標本共分散行列に 0.5 ブロックの平行移動と傾き変形の変形行列を作用させて変動吸収共分散行列を作成することにより, 認識率の向上が見られることを示した.

ETL9B 全セットを対象とした認識実験では, こう配特徴ベクトルと変動吸収共分散行列を用いた場合に 99.38%の認識率と 21 文字/秒の認識速度が, 更に加重フィルタの精細化と画像の拡大により 99.41%の認識率と 14 文字/秒の認識速度が得られ, 提案手法の有効性が確認できた.

謝辞 手書き漢字データベース“ETL9B”を作成, 配布されました, 電子技術総合研究所の皆様にご感謝致

します。

文 献

- [1] 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.
- [2] 若林哲史, Yan Deng, 鶴岡信治, 木村文隆, 三宅康二, “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,” 信学論 (D-II), vol.J79-D-II, no.5, pp.765–774, May 1996.
- [3] M. Suzuki, N. Kato, H. Aso, and Y. Nemoto, “A handprinted character recognition system using image transformation based on partial inclination detection,” IEICE Trans. Inf. & Syst., vol.E79-D, no.5, pp.504–509, May 1996.
- [4] 鈴木雅人, 大町真一郎, 加藤 寧, 阿曾弘具, 根元義章, “混合マハラノビス関数による高精度な類似文字識別手法,” 信学論 (D-II), vol.J80-D-II, no.10, pp.2752–2760, Oct. 1996.
- [5] 村瀬 洋, 木村文隆, 吉村ミツ, 三宅康二, “パターン整合法における特性核の改良とその手書き文字認識への応用,” 信学論 (D), vol.J64-D, no.3, pp.276–283, March 1981.
- [6] 斉藤泰一, 山田博三, 山本和彦, “JIS 第 1 水準手書き漢字データベース ETL9 とその解析,” 信学論 (D), vol.J68-D, no.4, pp.757–764, April 1985.
- [7] F. Kimura, T. Wakabayashi, S. Tsuruoka, and Y. Miyake, “Improvement of handwritten Japanese character recognition using weighted direction code histogram,” Pattern Recognit. J., vol.30, no.8, pp.1329–1337, Aug. 1997.
- [8] 山田博三, 斉藤泰一, 山本和彦, “非線形正規化法の改良,” 昭 63 信学春季全大, D-439.
- [9] 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.
- [10] D.G. Keehn, “A note on learning for Gaussian properties,” IEEE Trans. Inf. Theory, vol.IT-11, no.1, pp.126–132, Jan. 1965.
- [11] 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.
- [12] 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をつぶれ文字への対応,” 信学技報, PRU90-128, Feb. 1991.
- [13] 鶴岡信治, 森田裕之, 木村文隆, 三宅康二, “筆記者に対して適応機能をもった自由手書き文字認識,” 信学論 (D), vol.J70-D, no.10, pp.1953–1960, Oct. 1987.

(平成 13 年 3 月 21 日受付)



澤 和弘 (正員)

昭 57 三重大・工・電子卒。昭 59 同大大学院修士課程了。昭 59 伊勢電子工業 (株) 入社, 薄膜 EL の研究に従事。平 2 三重県立津高等技術学校電子制御情報科勤務。平 13 三重大学院博士 (後期) 課程了。文字認識の研究に従事。



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 3 三重大・工・助手。平 10 同助教授, 現在に至る。その間平 10~11 米国レンセリア工科大学 (RPI) 客員研究員を勤める。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。情報処理学会会員。工博。



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。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。工博。



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。平 10 同教授, 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。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 43 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及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。