

## 推薦論文

## 弾性マッチングと固有変形を用いたオンライン文字認識

三 苦 寛 人<sup>†</sup> 内 田 誠 一<sup>†</sup> 迫 江 博 昭<sup>†</sup>

DP マッチングなど、いわゆる弾性マッチングに基づくオンライン文字認識においては、合わせ過ぎによる誤認識が発生する。たとえば、入力パターンが“1”であっても、マッチングによって水平部が非線形伸縮しながら対応付けられた結果、“7”に誤認識される場合がある。本論文では、こうした誤認識の低減手法を提案する。合わせ過ぎの発生原因としては、弾性マッチングが本来そのカテゴリでは起こりえないような変形も補償対象としていることがあげられる。そこで本手法では、あらかじめ各カテゴリに生じやすい変形（固有変形）を統計的手法により求めておき、認識の際のマッチングの結果が固有変形からどれくらい逸脱しているかを評価する。その逸脱量が大きければ、そのマッチングにより合わせ過ぎが起きていると判断できる。オンライン数字データを用いた認識実験により、本手法の有効性を確認した。

## Online Character Recognition Using Elastic Matching and Eigen-deformations

HIROTO MITOMA,<sup>†</sup> SEIICHI UCHIDA<sup>†</sup> and HIROAKI SAKOE<sup>†</sup>

In online character recognition based on elastic matching, such as DP matching, many of misrecognitions are often due to overfitting, which is the phenomenon that a wrong reference pattern is closely fitted an input pattern by the matching. In this report, a technique to reduce those misrecognitions is proposed, where frequent deformations of each category, called eigen-deformations, are employed. In case of overfitting, the matching between the two patterns will not be expressed by the eigen-deformations of the category of the reference pattern. Thus, the overfitting can be detected by evaluating the divergence of the matching result from the eigen-deformations. The results of recognition experiment showed the usefulness of the proposed technique.

## 1. はじめに

オンライン手書き文字認識における誤認識の要因には、筆順変動、画数変動、ストロークの形状変動の3つがある。このうち形状変動（以下、単に変形と呼ぶ）は1ストローク文字にも起こる変動であり、最も基本的なものといえる。本論文では、この変形への対処法を提案する。

変形の吸収法として、DP マッチング<sup>1),2)</sup>がある。DP マッチングは弾性マッチング（非線形伸縮）アルゴリズムの一種であり、音声の時間方向の非線形伸縮を吸収する目的として音声認識の分野で古くから用いられている。オンライン文字データも音声と同じ時系列データなので、DP マッチングによりその変形を吸

収できる。

従来の DP マッチングの問題点として、合わせ過ぎによる誤認識がある。たとえば、入力パターンが“1”であっても、DP マッチングにより“7”に誤認識される場合がある。これは、マッチングによる時間方向の非線形伸縮の結果、“1”の最初の短い水平部が“7”の水平部に対応づけられてしまったためである。こうした合わせ過ぎへの単純な対処法として、マッチングの範囲を各カテゴリごとに制約することが考えられる。しかしながら、その範囲をどう設定すればよいかは必ずしも明らかではない。さらに、DP を用いる限り、マルコフ的な変形すなわち隣接筆点間の局所変形しか制御できない点も問題として残る。つまり、“1”への

本論文の内容は 2004 年 3 月の情報処理学会九州支部主催平成 15 年度火の国情報シンポジウムにて報告され、火の国情報シンポジウム編集委員により情報処理学会論文誌への掲載が推薦された論文である。

<sup>†</sup> 九州大学  
Kyushu University

合わせ過ぎを避けるためにカテゴリ“7”の書き出し点と右上の角の点の相対位置関係を制約するといったような大局的変形の制御はできない。この限界は、DPを確率的な枠組みに発展させた Statistical DTW<sup>3)</sup> や HMM<sup>4)-6)</sup> においても同様である。

本論文では、各カテゴリ固有の変形特性(以下、固有変形)の利用に基づくオンライン文字認識の高精度化法を提案する。本手法では、各カテゴリに生じる変形には一定の傾向があるという点に着目する。たとえば、文字“7”においては、その垂直部の傾き角の変化による変形は頻繁に生じる。すなわち、この変形は“7”の固有変形の1つである。一方、書き出し点と右上の角の点が非常に近くなるような大局的変形、すなわち“1”に近づいてしまうような変形は実際にはあまり起こらない。これは、書き手が他の文字との区別をつけるために意識的・無意識的に変形の範囲を規定しているものと考えることができる。したがって、事前に固有変形が把握できていれば、それからの逸脱として合わせ過ぎを排除できることになる。なお、後に示すように、各カテゴリの固有変形はDPと主成分分析の組合せにより自動的かつ統計的に推定される。

認識処理において固有変形をどのように用いるかについては、オフライン文字認識の場合<sup>7),8)</sup>と同様、2つの方式が考えられる。すなわち、(i)事後評価による方式と、(ii)直接組み込みによる方式、が考えられる。本論文では(i)の方式を採用する。この方式では、まず従来と同じく入力と標準パターンのDPマッチングを行う。次にそのマッチングの結果について固有変形からの逸脱量を求める。これを事後評価量として識別時にペナルティとして用いる。合わせ過ぎが起きれば、この事後評価量は大きくなると考えられるので、結果的に合わせ過ぎにペナルティを与えることになる。この方式は、方式(ii)に比べて考え方が容易であるため、検討の第1段階として採用した。

線パターン間のマッチングを統計量を使って評価する試みは、木村らによって1970年代後半にすでに検討されている<sup>9),10)</sup>。この手法はオフライン文字認識に関するものであり、線分表現された標準パターンと、細線化された入力パターン間のマッチングを、マハラノビス距離を用いて評価している。本手法はこれらの手法を、オンライン文字認識処理の枠組みにおいて発展させたものであるといえる。

## 2. オンライン手書き文字の固有変形の推定

本章では、オンライン手書き文字の固有変形の推定法について述べる。前述のように、固有変形とは、

各文字種固有の変形方向性であり、本論文では標準パターンからの変位として表現される。固有変形は、(i)DPマッチングによる変位の採集と、(ii)その主成分分析、という2つのステップにより自動的に求められる。以下では、それぞれのステップについて詳述する。

### 2.1 DPマッチングによる変位の採集

第 $c$ カテゴリの標準パターン $R_c$ を時系列

$$R_c = r_1, r_2, \dots, r_i, \dots, r_I \quad (1)$$

と表し、同様に第 $c$ カテゴリの第 $n$ 番目の学習パターン $T_{c,n}$ を時系列

$$T_{c,n} = t_1, t_2, \dots, t_j, \dots, t_J \quad (2)$$

と表す。ここで $r_i$ および $t_j$ は筆点の特徴ベクトルであり、

$$r_i = (x_i^r, y_i^r, d_i^r)^T \quad (3)$$

$$t_j = (x_j^t, y_j^t, d_j^t)^T \quad (4)$$

のように $x$ 座標、 $y$ 座標、および方向を要素とする3次元ベクトルとする。なお、正確には $r_i, t_j, I, J$ をそれぞれ $r_{c,i}, t_{c,n,j}, I_c, J_{c,n}$ のように記述すべきだが、以下では誤解のない限りこれらの簡略表現を用いる。

2パターン $R_c$ と $T_{c,n}$ 間の弾性マッチング問題は、次の最適化問題として定義される。

[目的関数]

$$\frac{1}{I} \sum_{i=1}^I \|r_i - t_{j(i)}\| \rightarrow \text{minimize}$$

[制御変数(自然数)]

$$j(1), \dots, j(i), \dots, j(I)$$

[制約条件]

$$\begin{cases} j(i) - j(i-1) \in \{0, 1, 2\} \\ j(1) = 1 \\ j(I) = J \end{cases}$$

ここで記号 $\|x\|$ はベクトル $x$ のユークリッドノルム $\sqrt{x^T x}$ である。よく知られているように、この問題はDPを使うことで効率的に解ける。また、目的関数の最小値を $D_0(R_c, T_{c,n})$ とする。すなわち、

$$D_0(R_c, T_{c,n}) = \min_{j(1), \dots, j(i), \dots, j(I)} \frac{1}{I} \sum_{i=1}^I \|r_i - t_{j(i)}\| \quad (5)$$

とする。この値は、変形吸収後の $R_c$ と $T_{c,n}$ のユークリッド距離であり、従来から識別関数値として用いられている。

制御変数 $j(1), \dots, j(I)$ は2パターンの対応関係を

表す変数であるが、これを基にすれば標準パターンを基準とした場合の学習パターンの変形を変位ベクトルとして求めることができる<sup>12)</sup>。具体的には、 $T_{c,n}$  の  $2I$  次元変位ベクトル  $v_{c,n}$  は次のように表現される。

$$v_{c,n} = \left( (x_1^r - x_{j(1)}^t, y_1^r - y_{j(1)}^t), \dots, (x_i^r - x_{j(i)}^t, y_i^r - y_{j(i)}^t), \dots, (x_I^r - x_{j(I)}^t, y_I^r - y_{j(I)}^t) \right)^T \quad (6)$$

よって、各カテゴリ  $c$  について標準パターン  $R_c$  と  $N_c$  個の学習パターン  $T_{c,n}$  間で DP マッチングを行えば、 $N_c$  個の変位ベクトル  $v_{c,n}$  が求まる。

### 2.2 主成分分析による固有変形の推定

固有変形は頻出する変形であり、変形のなす分布の分散の大きな方向に対応する。したがって固有変形は、前ステップで求めた  $N_c$  個の変位ベクトル  $v_{c,n}$  を主成分分析することで求まる。

具体的には、まず各カテゴリについて、平均

$$\bar{v}_c = \frac{1}{N_c} \sum_{n=1}^{N_c} v_{c,n} \quad (7)$$

および共分散行列

$$\Sigma_c = \frac{1}{N_c} \sum_{n=1}^{N_c} (v_{c,n} - \bar{v}_c)(v_{c,n} - \bar{v}_c)^T \quad (8)$$

を求める。変位ベクトル  $v_{c,n}$  の次元は式 (6) のように  $2I$  で一定しているため、以上の計算が可能となっている。なお、式 (7) と同様の平均変位ベクトルの求め方は、文献 11) にも見られる。

次にこの  $\Sigma_c$  を固有値展開すれば、固有ベクトル  $\{u_{c,1}, \dots, u_{c,m}, \dots, u_{c,M}\}$  として固有変形が得られ、固有値  $\{\lambda_{c,1}, \dots, \lambda_{c,m}, \dots, \lambda_{c,M}\}$  として各固有変形の寄与の大きさが求まる。ここで、次数  $m$  は固有値の大きさ順に対応する。したがって、第 1 固有変形  $u_{c,1}$  が最頻変形であり、また大きな  $m$  に対応する固有変形ほど寄与は小さく、ノイズ的になる。

### 3. 固有変形を用いた認識

従来の DP マッチングの枠組みによるオンライン文字認識、すなわち入力パターンと標準パターン間の距離を  $D_0$  を用いて評価することに基づく文字認識においては、1 章で述べたように合わせ過ぎによる誤認識が問題となる。図 1 に合わせ過ぎによる誤認識例を示す。この誤認識された入力パターン“9”は、標準パターン“5”と不自然に対応付けられている。より詳細には、“5”の円弧部が交差しながら“9”の屈曲部に対応付けられている。その結果、各筆点間の距離の総

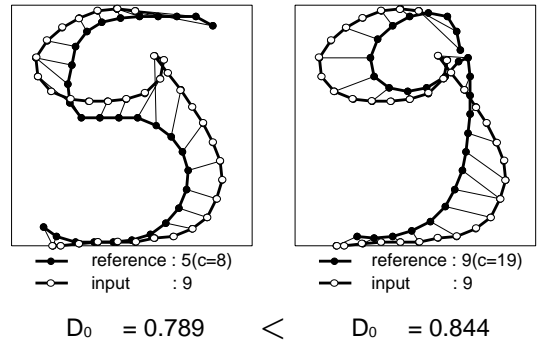


図 1 合わせ過ぎにより誤認識された文字の例 (入力文字“9”→認識結果“5”)。括弧内の  $c$  の値は図 2 の標準パターン番号  $c$  に対応

Fig. 1 Example of overfitting.

和、すなわち DP マッチング距離  $D_0$  が小さくなり、最終的には“5”に誤認識されている。こうした合わせ過ぎが発生する原因は、DP マッチング時に各カテゴリの変形方向特性が考慮されていないためと考えられる。

そこで本章では、DP ベースのオンライン文字認識法において固有変形を利用することで、合わせ過ぎによる誤認識を低減する手法について述べる。具体的には、従来どおり、まず標準パターン  $R_c$  とカテゴリ未知の入力パターン  $E$  間で DP マッチングを行い、距離  $D_0(R_c, E)$  ならびに  $E$  の  $2I$  次元変位ベクトル  $v$  を求める。その後その変位ベクトル  $v$  と固有変形  $\{u_{c,m} \mid m = 1, \dots, M\}$  の距離を評価する。合わせ過ぎを起こした際の変位ベクトル  $v$  は、多くの場合、上位すなわち高頻度の固有変形では表現できないため、両者の差異は大きくなると考えられる。すなわち、 $v$  の距離は大きくなる。本手法では、この事後評価による距離を  $P(R_c, E)$  とすると、それを DP マッチング距離  $D_0(R_c, E)$  に加えたもの、すなわち

$$D_{\text{eigen}}(R_c, E) = (1 - \alpha)D_0(R_c, E) + \alpha P(R_c, E) \quad (9)$$

を用いて最短距離法による識別を行う。合わせ過ぎが起きた場合、距離  $P$  の増加により、距離  $D_{\text{eigen}}$  は距離  $D_0$  を単独よりも大きくなる。したがってそのサン入力パターン  $E$  はカテゴリ  $c$  には認識されにくくなる。この仕組みにより、合わせ過ぎによる誤認識を排除できると考えられる。ここで、 $\alpha$  は重み係数 ( $0 \leq \alpha \leq 1$ ) である。

本論文では、 $v$  が平均  $\bar{v}_c$ 、分散  $\Sigma_c$  の多次元正規分布をなすという仮定の下、事後評価距離  $P$  として次式

$$P(\mathbf{R}_c, \mathbf{E}) = \frac{1}{I} \sqrt{p(\mathbf{R}_c, \mathbf{E})} \quad (10)$$

を用いる．ここで  $p(\mathbf{R}_c, \mathbf{E})$  は次の (2 乗) マハラノビス距離である．

$$p(\mathbf{R}_c, \mathbf{E}) = (\mathbf{v} - \bar{\mathbf{v}}_c)^T \Sigma_c^{-1} (\mathbf{v} - \bar{\mathbf{v}}_c) \\ = \sum_{m=1}^M \frac{1}{\lambda_{c,m}} ((\mathbf{v} - \bar{\mathbf{v}}_c)^T \mathbf{u}_{c,m})^2 \quad (11)$$

式 (10) において，このマハラノビス距離の平方根をとっているのは，式 (9) において線形加算されるユークリッド距離 (5) とオーダを同じにするためである．また  $I$  で割っているのは，式 (5) と同じく，式 (11) における標準パターン長の影響を正規化するためである．

上位すなわち大きな  $m$  に対応する固有ベクトル  $\mathbf{u}_{c,m}$  は推定誤差を含みやすいことが知られている．上位の固有値  $\{\lambda_{c,m}\}$  は非常に小さな値となるので，その上位固有ベクトルの推定誤差はマハラノビス距離 (11) の値に大きく影響する．さらにゼロに非常に近い固有値については計算誤差の影響を受ける場合もある．このため，本論文では式 (11) の代わりに，次の疑似マハラノビス距離<sup>13)</sup>

$$p(\mathbf{R}_c, \mathbf{E}) \\ \sim \sum_{m=1}^{M'} \frac{1}{\lambda_{c,m}} ((\mathbf{v} - \bar{\mathbf{v}}_c)^T \mathbf{u}_{c,m})^2 \\ + \sum_{m=M'+1}^M \frac{1}{\lambda_{c,M'+1}} ((\mathbf{v} - \bar{\mathbf{v}}_c)^T \mathbf{u}_{c,m})^2 \\ = \frac{1}{\lambda_{c,M'+1}} \|\mathbf{v} - \bar{\mathbf{v}}_c\|^2 \\ + \sum_{m=1}^{M'} \left( \frac{1}{\lambda_{c,m}} - \frac{1}{\lambda_{c,M'}} \right) ((\mathbf{v} - \bar{\mathbf{v}}_c)^T \mathbf{u}_{c,m})^2 \quad (12)$$

を用いる．これにより，前述の誤差の影響を抑えることができる．パラメータ  $M'$  の決め方には様々な方法が考えられる．本論文では累積寄与率  $\sum_{m=1}^{M'} \lambda_{c,m} / \sum_{m=1}^M \lambda_{c,m}$  が，あるしきい値  $\theta$  をはじめて超えたときの  $M'$  を用いることとした．後述する認識実験においては，この  $\theta$  は全カテゴリ共通で 0.9 としている．

## 4. 実験

### 4.1 実験試料

本論文では，Unipen フォーマット<sup>14)</sup> に従った公開データベースの 1 つである Ethem Alpaydin Digit<sup>15)</sup> の孤立数字データを実験試料として用いた．本データベースは，あらかじめ 7,494 個の学習パターンと 3,498

個のテストパターンに分けられて提供されている．したがって，各文字種平均で学習パターン 749 サンプル，テストパターン 350 サンプル，総数 10,992 サンプルから構成されている．以下の実験では，学習パターンセットを標準パターン設定と固有変形推定に，テストパターンセットを認識タスクとして用いた．

これらすべてのサンプルについて，画数変動の影響をなくすために，ペンアップ部分も接続して 1 文字 1 画とした．その後，文字全体の大きさを縦横比は保持したまま  $128 \times 128$  に線形正規化し，さらに隣接筆点間の距離が一定になるようにリサンプリングを行った．

本データベースの学習パターンセットとテストパターンセットには同じ筆記者が含まれていなかった．具体的には，学習パターンセットは 30 人の筆記者によって作成されているが，テストパターンセットは完全に別の 14 人の筆記者によって作成されている．このため両データセットの筆記特性には差異が見られた．この差異が認識実験の結果に及ぼす影響については後述する．

### 4.2 標準パターンの設定

すべての学習パターンを観察し，手で標準パターン  $\mathbf{R}_c$  を作成した．その際，筆順変動がある文字種に対しては，複数の標準パターンを用意した．実験試料同様，すべての標準パターンにも，1 画化，正規化，リサンプリングを施した．

図 2 は設定した全 21 カテゴリの標準パターンである．数字 “8” の 4 番目の標準パターン ( $c = 18$ ) に反映されているように，本データベースの学習パターンには，かなり特殊な筆順も含まれていた．

4.1 節でも述べたように，学習パターンとテストパターンでは筆記者が異なっている．その結果，学習パターンを基準として作成した以上の標準パターンとテスト用サンプルとの間には傾向の違いが見られた．たとえば，数字 “8” の 3 番目の標準パターン ( $c = 17$ ) に類似したサンプルは，学習パターンセット中には 48 個存在したのに対し，テストパターンセット中には皆無であった．同様に，数字 “8” の 4 番目の標準パターン ( $c = 18$ ) に類似したサンプルは，テストパターンセット中でわずかに 1 個であった．

### 4.3 固有変形の推定結果

2 章で述べた固有変形の推定は以下の手順で行った．まず，1 つの標準パターンしか用意しなかった文字 (“0”，“2”，“3”，“6”) については，標準パターン  $\mathbf{R}_c$  とその文字の全学習パターン  $T_{c,n}$  の間で DP マッチングを行い， $N_c$  個の変位ベクトル  $\{\mathbf{v}_{c,1}, \dots, \mathbf{v}_{c,n}, \dots, \mathbf{v}_{c,N_c}\}$  を採集し，主成分分析に

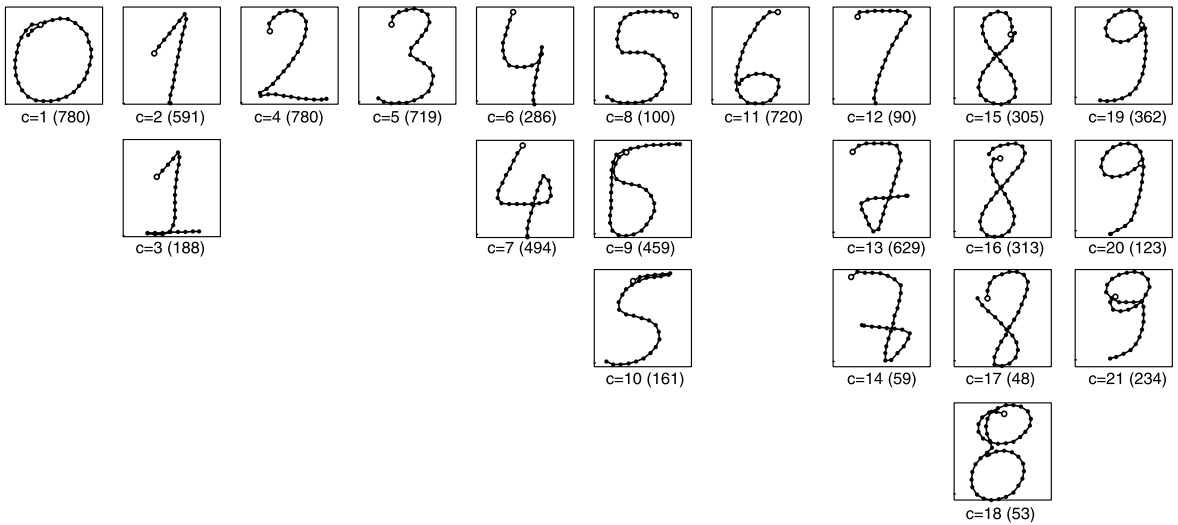


図 2 実験で用いた標準パターン  $\{R_c \mid c = 1, \dots, 21\}$ . 白丸は開始点,  $c$  は標準パターン番号, 括弧内の数字は学習パターン数  $N_c$ .

Fig. 2 Reference pattern used in the experiment.

より最終的に固有変形  $\{u_{c,m}\}$  と固有値  $\{\lambda_{c,m}\}$  を求めた. 一方, 複数の標準パターンを持つ文字 (すなわち “1”, “4”, “5”, “7”, “8”, “9”) については, 同文字の学習パターンを次の手順で分割した. すなわち, 各学習パターンについてそれら複数の標準パターンと DP マッチングを行い, 距離が最も小さくなった標準パターン用の学習パターンとして分割した. その後, 各標準パターンごとに 1 文字 1 標準パターンの場合と同じ手続きで固有変形と固有値を求めた. 図 2 中で各標準パターンの下に付与した括弧内の数字は, その標準パターンのカテゴリ  $c$  に属した学習パターンの数である. このように, 1 カテゴリあたりの学習パターン数は最少でも 48 ( $c = 17$ ) であった.

図 3 は, 寄与の大きな上位 3 つの固有変形  $u_{c,1}$ ,  $u_{c,2}$ ,  $u_{c,3}$  により標準パターン  $R_c$  を変形させたものである. これらの固有変形は, 普段よく見られる変形になっていることが分かる. たとえば, 数字 “0” ( $c = 1$ ) の第 1 固有変形は, 書き出し点の変化であると分かる. また, 数字 “6” ( $c = 11$ ) では, 第 1 固有変形が文字全体の傾きの変化を, 第 2 固有変形がループ部の縦方向の変化を, 第 3 固有変形がループ部の横方向の変化を, それぞれ表していることが分かる.

4.4 認識結果および考察

4.4.1 認識率

図 4 に, 重み  $\alpha$  を変えながら測定した, 本手法すなわち  $D_{\text{eigen}}$  に基づく最短距離識別法による認識率を示す. なお,  $\alpha = 0$  の場合が, 従来法すなわち  $D_0$  に基づく最短距離識別法に対応する.

同図より,  $\alpha = 0.45$  のときに最高認識率 98.2% が得られたことが分かる. 一方, 従来法による認識率は 97.4% であり, これに比べると, 本手法の結果は 0.8% (27 サンプル分) の向上となっている. この向上の内訳を見ると, 35 サンプルの改善と 8 サンプルの改悪であった. その詳細については 4.4.2 項で検討する. また, 本手法において誤認識となった全 64 サンプルについては, 4.4.3 項で考察する. さらに, 学習パターンとテストパターンで筆記者が異なること (4.1 節参照) がこの認識率に及ぼした影響については, 4.4.4 項で考察する.

文献 16) では, 本論文と同じデータベースを用いた認識実験がなされ, 97.09% の最高認識率が得られたと報告されている. 同文献で提案されているのはニューラルネットに基づく認識機構を複数組み合わせた手法であり, 本手法とはかなり異なる認識モデルである. したがって両者の優劣を認識率だけで決めることはできない. しかしながら, 両者の認識率の差異は, 本手法の認識性能の高さを裏付けているといえる.

なお, ユークリッド距離単独での認識率 ( $\alpha = 0$ ) とマハラノビス距離単独での認識率 ( $\alpha = 1$ ) には大きな差が見られる. この一因として, マハラノビス距離の計算で用いる変位ベクトルには方向情報が入っていない点あげられる. これは変位ベクトルの次元数をおさえるための配慮である. 十分な学習サンプルが得られれば, 変位ベクトルに方向特徴の差異も含めることで, マハラノビス距離による識別性能を向上させられると考えられる. この場合  $D_{\text{eigen}}$  は, パターン

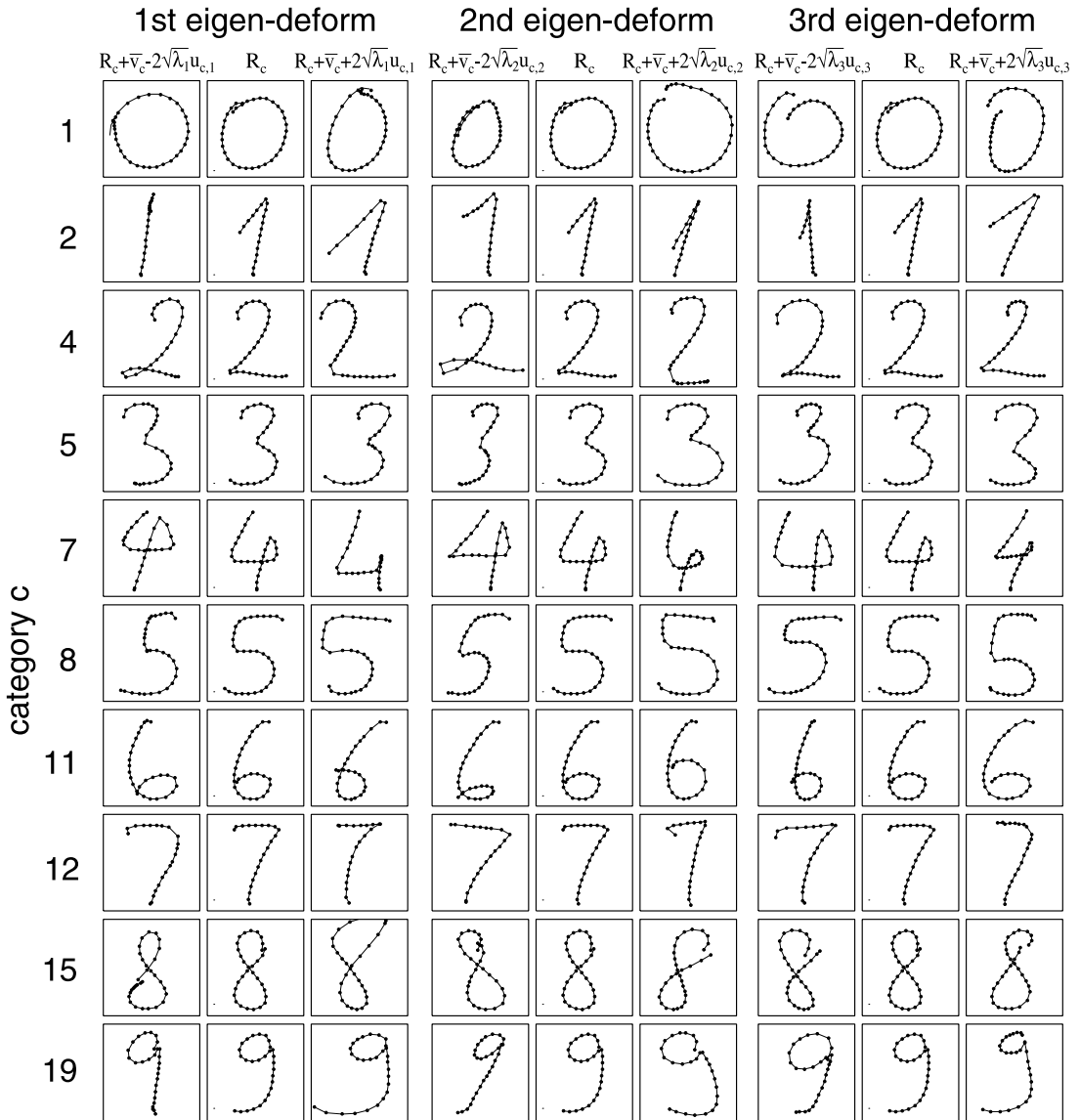


図3 上位3つの固有変形  $u_{c,1}, u_{c,2}, u_{c,3}$  によって変形された標準パターン  
 Fig. 3 Reference pattern deformed by the primary eigen-deformations of its category.

間の座標差と方向差という同一の情報をユークリッド距離とマハラノビス距離という異なる2つの距離尺度で評価し、それらを合算したものに相当する。

4.4.2 改善例および改悪例

従来法では誤認識であったが本手法 ( $\alpha = 0.45$ ) では正しく認識された全35サンプル(改善サンプル)を表1にまとめる。表中で使用されている“9”→“5”という表記は、入力パターン“9”が“5”に誤認識されたことを意味する。それら改善された全35サンプルを観察したところ、明らかに合わせ過ぎに起因する

誤認識が抑制されていることが分かった。その具体例を図5に示す。同図(a)は誤認識“1”→“7”が改善された例、同図(b)は誤認識“9”→“5”が改善された例である。なお、同図(b)の入力パターン“9”は、図1のものと同じである。

図5(a)では、入力パターン“1”の冒頭部と標準パターン“7”の水平部が不自然に対応付けられている。この合わせ過ぎの結果として従来 $D_0$ による識別では入力パターン“1”は“7”に誤認識されている。一方、事後評価距離  $P$  を見ると、正解標準パターン“1”

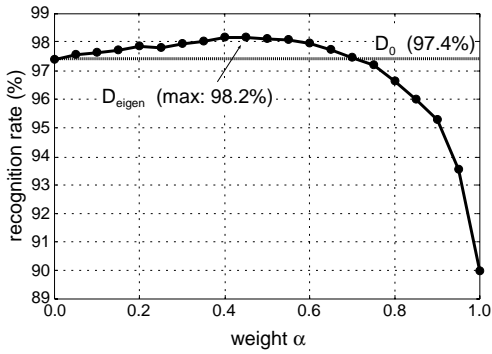


図 4 手書き数字データ (全 3498 サンプル) の認識率  
Fig. 4 Recognition rate as the function of weight  $\alpha$ .

表 1 本手法により減少した誤認識  
Table 1 Decreased misrecognitions.

input → result	#misrecog.		#diff
	$D_0$	$D_{eigen}$	
“9” → “5”	10	1	9
“0” → “8”	16	9	7
“1” → “7”	6	0	6
“2” → “7”	6	1	5
“3” → “7”	9	6	3
“0” → “6”	1	0	1
“4” → “1”	1	0	1
“4” → “5”	1	0	1
“9” → “1”	1	0	1
“9” → “7”	1	0	1
total			35

とのマッチングよりも、この不正解標準パターン “7” とのマッチングの方で大きくなっている。結果的に、 $D_0$  と  $P$  の線形和である  $D_{eigen}$  による識別では、正解に転じている。

図 5 (b) も、事後評価距離  $P$  の効果により正しい認識結果が得られている。これは “9” が “5” に不自然に対応付けられたために距離  $D_0$  が過小評価され、誤認識されていた例である。しかし、図 3 を見ると、標準パターン “5” の円弧部に生じる変形は “9” のように屈曲したものではなく本来平滑であることが分かる。したがって、その円弧部が交差しながら入力パターンに対応付けられるような状況はほとんど生じない。これが入力パターン “9” とのマッチングにおける事後評価距離を大きくしたものと考えられる。図 6 は図 5 (b) の 2 つのマッチングにおける各固有変形成分を示したものである。合わせ過ぎが起きている不正解標準パターン “5” とのマッチングでは、下位すなわち低頻度の固有変形のところで比較的大きな値が出ており、これがペナルティとなって  $P$  を大きくしていることが分かる。

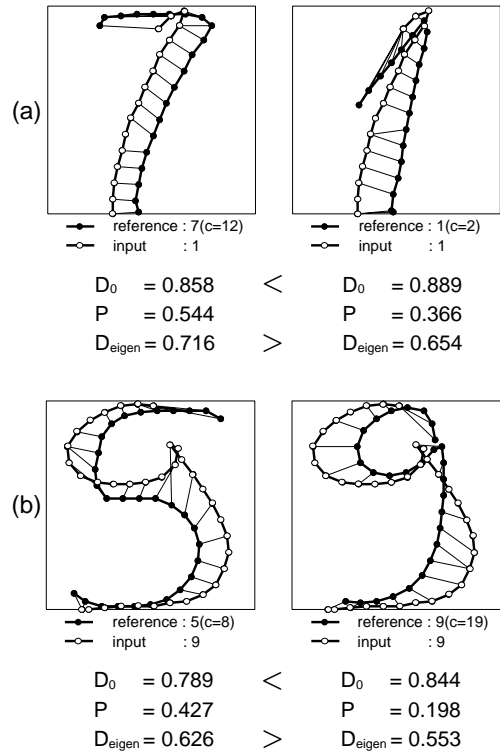


図 5 固有変形の利用による改善例  
Fig. 5 Characters correctly recognized by the proposed technique.

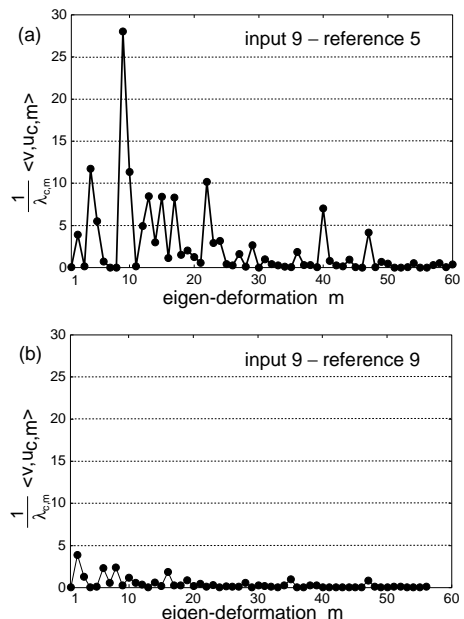


図 6 図 5 (b) のマッチングの際の各固有変形成分の大きさ。(a) 入力パターン “9” と標準パターン “5” のマッチング。(b) 入力パターン “9” と標準パターン “9” のマッチング  
Fig. 6 Detail of posterior evaluation.

表 3 誤認識となった全 64 サンプルの内訳  
Table 3 Detail of all misrecognitions.

reason of misrecog.	#samples	input sample (upper) and recog. result (lower, input→result)
abnormal start/end point	3	 “3”→“9”   “3”→“1”   “9”→“8”
extreme deformation	12	 “1”→“2”   “4”→“2”   “5”→“3”   “9”→“7”
case 1 ( $P_{\text{correct}} < P_{\text{wrong}}$ )	14	 “0”→“8”   “2”→“7”   “3”→“5”   “9”→“5”
case 2 ( $P_{\text{correct}} > P_{\text{wrong}}$ )	35	 “3”→“7”   “4”→“9”   “5”→“9”   “7”→“9”

表 2 本手法により増加した誤認識  
Table 2 Increased misrecognitions.

input→ result	#misrecog.		#diff
	$D_0$	$D_{\text{eigen}}$	
“7”→“9”	8	12	4
“5”→“9”	5	8	3
“6”→“8”	0	1	1
	total		8

従来法では正しく認識されたが本手法では誤認識された全 8 サンプル (改悪サンプル) を表 2 にまとめる。この新たに誤認識された “5” と “7” の入力パターンは学習パターンには見られないものであった。そのため、求めた固有変形ではそれら入力パターンの変形を表現できず、事後評価距離が大きくなり誤認識になったものと考えられる。

#### 4.4.3 誤認識原因の考察

本手法 ( $\alpha = 0.45$ ) で誤認識となった全 64 サンプル (前出の改悪サンプル 8 を含む) の内訳を、考えられる原因とともに表 3 に示す。最初の 2 種の誤認識 (始端・終端異常ならびに難読文字) 計 15 サンプルは、本手法に限らず一般に認識が非常に困難なものである。よって以下では、残る 49 サンプルの誤認識を中心として考察する。

これら 49 サンプルは以下のように case 1 と case 2 に分類される。まず case 1 とは条件  $P_{\text{correct}} < P_{\text{wrong}}$  を満たす誤認識パターンであり、14 サンプルが含ま

れた。ここで、 $P_{\text{correct}}$  は正解標準パターンとの事後評価距離、 $P_{\text{wrong}}$  は不正解標準パターンとの事後評価距離である。この case 1 の誤認識の原因の 1 つとしては、

原因 A: 事後評価距離  $P_{\text{wrong}}$  は有効に機能したが距離  $D_0$  の差が元々大き過ぎた、

ことが考えられる。したがって、この原因 A による誤認識の低減には、正解カテゴリとの  $D_0$  を不正解カテゴリとのそれより小さくする、すなわち従来法の誤認識を低減させることが必要となる。具体的には、標準パターンを新たに増やす、もしくは合理的な設定手法 (たとえば学習による自動的な標準パターン設定法) を採用することにより改善できると考えられる。

次の case 2 とは条件  $P_{\text{correct}} > P_{\text{wrong}}$  を満たす誤認識である。この case 2 の誤認識の原因は次の 2 つが考えられる。

原因 B: 正解標準パターンに対する事後評価距離  $P_{\text{correct}}$  が過剰に大きくなった。

原因 C: 不正解標準パターンに対する事後評価距離  $P_{\text{wrong}}$  が十分大きくならなかった。

原因 B と C は排他的ではなく、両方による誤認識も多く存在する。実際、case 2 の 35 サンプルのほとんどはそうした誤認識であった。具体的には、原因 B を  $P_{\text{correct}} > \bar{P}_{\text{correct}}$  を満たす場合 ( $\bar{P}_{\text{correct}}$  は  $P_{\text{correct}}$  の平均値)、原因 C を  $P_{\text{wrong}} < \bar{P}_{\text{wrong}}$  を満たす場合 ( $\bar{P}_{\text{wrong}}$  は不正解標準パターンに対する  $P$  のうちで最



小のものの平均値)と定義すると, 35 サンプル中 32 サンプルは両原因によるものであった.

原因 B は, 実際にカテゴリに起こっている変形であるにもかかわらず, それがあまり起こらない変形と判断してまったことを意味する. より具体的には, その誤認識されたテストパターンの変形がその正解カテゴリの上位の固有変形では表現できなかったことを意味する. したがって原因 B による誤認識を低減させるには, 学習パターンを拡充することが効果的と考えられる. 前述のように, 公開データベース Ethem Alpaydin Digit のテストパターンセットと学習パターンセットでは筆記者が異なっており, 本実験では特にこの不整合がこの原因 B を顕著なものとしていてと考えられる.

原因 C は, 異なるカテゴリにおいても類似した変形が存在することを意味する. したがって, 変形がカテゴリごとで異なっていることを前提とした本手法では, この原因による誤認識を直接低減することは難しい. しかしながら, 前述のように case 2 に属する誤認識のほとんどすべては, 原因 B と原因 C の両方によって発生したものである. さらに,  $P_{\text{correct}}$  を小さくできれば, 相対的に  $P_{\text{wrong}}$  の効果を大きくできる可能性がある. 以上から, 原因 B に対処することで結果的に原因 C による誤認識も減らせると考えられる.

#### 4.4.4 筆記者の影響

以上の考察においては, テストパターンセットと学習パターンセットで筆記者が異なっており, さらにこの不整合が誤認識を誘発していることを指摘してきた. この不整合は重み  $\alpha$  の設定にも影響する. 実際, 学習パターンで閉じた認識実験での最適値  $\alpha = 0.70$  を, テストパターンの認識実験に用いると, 図 4 から, テスト用サンプルの認識率は 97.5% と, 従来法程度まで低下することが分かる.

そこで, この筆記者の不整合をなくした場合, すなわち筆記者集合を学習時と認識時で一致させた場合の実験を行った. 具体的には, データベースの全パターンをいったん混合し, 再び学習パターン 7494 サンプルと 3498 個のテストパターンに振り分けた. この振り分けにより学習パターンもテストパターンも等しく全 44 人分のサンプルを含むことになる. この新たな学習パターンセットをすべて使って固有変形を求めた後, 新たなテストパターンの認識実験を行った. なお, 実験結果の比較を容易にするため, 標準パターンについては前節までの実験と同じ図 2 の 21 個を用いた.

実験の結果, 本手法によるテストパターンの認識率

は 98.3% となった. これは前述の筆記者が異なっていた場合の結果 (98.2%) よりも 0.1% (5 サンプル分) 程度向上している. 前節の実験と同じ標準パターンを使ったためにあまり大幅な認識率の向上にはなかったものの, 前節までの実験では筆記者の不整合が影響していた可能性があったことが分かる. また, 重み  $\alpha$  については, 学習パターンで閉じた認識実験においてもテストパターンの認識実験においても, 認識率のピークを与えたのは同じ  $\alpha = 0.55$  であった. したがって, 筆記者が一致すると  $\alpha$  の不整合も解消するといえる.

誤認識となった全 59 サンプルのうち, 19 サンプルが case 1 であり, 26 サンプルが case 2 であった. 前節の実験では case 2 が case 1 の 2 倍以上あったのに対し, 本実験ではその差は明らかに縮まっている. これは前節でも考察したように, 筆記者一致の結果, 前節の原因 B による誤認識を低減できたためと考えられる.

## 5. おわりに

オンライン文字認識におけるストロークの形状変動への一対処法として, カテゴリ固有の変形特性 (固有変形) の利用について検討した. 実験の結果, 主成分分析により各カテゴリに実際に生じる変形を適切に表現した固有変形が得られることが分かった. また, 認識実験を通して, 従来法では合わせ過ぎにより誤認識となっていたサンプルも本手法では正しく認識することができ, 合わせ過ぎによる誤認識の抑制効果を確認することができた.

将来的な課題として, 固有変形をマッチング過程に直接組み込む方式の検討があげられる. この方式を採用することで, 認識精度だけでなく計算量的にも有効ではないかと考えている. また, 本論文では手動で行っていた標準パターンの設定についても自動化<sup>(11), (17)</sup>が必要と考えている. さらに多画文字への適用についても検討する予定である.

謝辞 本研究の一部は文部科学省科学研究費補助金 (若手 B No.14780293), および総務省戦略的情報通信研究開発推進制度の支援を受けた.

## 参考文献

- 1) 藤原塩和, 池田克夫, 富永善治, 清野 武: 接線ベクトル列を用いたオンライン手書き文字の認識, 情報処理, Vol.17, No.3, pp.191-199 (1976).
- 2) Yoshida, K. and Sakoe, H.: Online handwritten character recognition for a personal computer system, *IEEE Trans. Consumer Electronics*, Vol.CE-28, No.3, pp.202-209 (1982).

- 3) Bahlmann, C. and Burkhardt, H.: The writer independent online handwriting recognition system *frog on hand* and cluster generative statistical dynamic time warping, *IEEE Trans. PAMI*, Vol.26, No.3, pp.299–310 (2004).
- 4) Hu, J., Brown, M.-K. and Turin, W.: HMM based on-line handwriting recognition, *IEEE Trans. PAMI*, Vol.18, No.10, pp.1039–1045 (1996).
- 5) 伊藤 等, 中川正樹: Hidden Markov Model に基づくオンライン手書き文字認識, 信学技報, PRMU97-85 (1997).
- 6) 嵯峨山茂樹, 中井 満, 下平 博: ストローク HMM に基づくオンライン手書き文字認識方式, 信学技報, PRMU2000-35 (2000).
- 7) Uchida, S. and Sakoe, H.: Eigen-deformations for elastic matching based handwritten character recognition, *Pattern Recognition*, Vol.36, No.9, pp.2031–2040 (2003).
- 8) 内田誠一, 迫江博昭: カテゴリ固有変形の線形結合モデルに基づく弾性マッチング法, 信学論 (D-II), Vol.J87-DII, No.2, pp.639–648 (2004).
- 9) 木村文隆, 吉村ミツ, 三宅康二, 市川真人: ストローク構造解析法による自由手書き片仮名文字認識, 信学論 (D), Vol.J62-D, No.1, pp.16–23 (1979).
- 10) 木村文隆, 堀尾佳弘, 塚本隆啓, 吉村ミツ, 三宅康二: 文字の構造情報抽出に用いる対応付けシステムと平仮名文字認識への応用, 信学論 (D), Vol.J63-D, No.8, pp.650–657 (1980).
- 11) 趙 鵬, 佐藤幸男, 吉村ミツ: オンライン走り書き文字認識における汎用辞書の作成, 情報処理学会論文誌, Vol.34, No.3, pp.418–425 (1993).
- 12) 若原 徹: 局所的 Affine 変換を用いたオンライン手書き文字認識, 信学論, Vol.J71-D, No.2, pp. 379–386, 1988.
- 13) 栗田昌徳, 鶴岡信治, 横井茂樹, 三宅康二: 加重方向ヒストグラムと疑似マハラノビス距離を用いた手書き漢字・ひらがな認識, 信学技報, PRL82-79 (1982).
- 14) Guyon, I., Schomaker, L., Plamondon, R., Liberman, M. and Janet, S.: UNIPEN project of on-line data exchange and recognizer benchmarks, *Proc. Int. Conf. Pattern Recognition*, pp.29–33 (1994).
- 15) ftp://ftp.ics.uci.edu/pub/machine-learning-databases/pendigits/
- 16) Alimoglu, F. and Alpaydin, E.: Combining multiple representations and classifiers for pen-based handwritten digit recognition, *Proc. Int. Conf. Document Analysis and Recognition*, Vol.2 of 2, pp.637–640 (1997).
- 17) Connell, S.D. and Jain, A.K.: Template-based online character recognition, *Pattern Recogni-*

*tion*, Vol.34, No.1, pp.1–14 (2001).

(平成 16 年 7 月 13 日受付)

(平成 16 年 10 月 4 日採録)

## 推薦文

入力文字図形の文字カテゴリを認識する場合, 基本的には, 入力図形と標準文字図形の類似度を計算し, 最も類似している文字カテゴリを入力図形の文字カテゴリとする. ところが, オンライン手書き文字認識では, 個人の癖などによる形状変形のために, 入力図形と正解の標準文字図形との類似度がそれほど高くない場合があり, 誤認識の要因の 1 つとなっている. これを解決する手法として, 非線形伸縮により入力図形を変形し標準図形に合わせる (近づける) DP マッチングがある. 従来の DP マッチングは, 合わせ過ぎによる誤認識の問題があったが, 本論文では, 入力図形を標準図形に合わせて変形する際, 各文字カテゴリごとに固有の変形特性 (固有変形) があることを利用し, 合わせ過ぎの問題を防ぐ手法を提案している. この固有変形も DP と主成分分析の組合せにより自動的に推定される. 手法の正当性については問題がなく, 実験によりその有効性も示されているため, 本論文を推薦するものである.

(火の国情報シンポジウム編集委員 岡崎直宣)



三苦 寛人

平成 16 年九州大学工学部電気情報工学科卒業. 同年同大学院システム情報科学府修士課程入学. 現在に至る. オンライン文字認識に関する研究に従事. 電子情報通信学会会員.



内田 誠一 (正会員)

平成 2 年九州大学工学部電子工学科卒業. 平成 4 年同大学院修士課程 (情報) 修了. セコム (株) 勤務を経て, 現在, 同大学院システム情報科学研究所知能システム学部門助教. 博士 (工学). 画像パターンの解析・認識, および音声の認識・合成に関する研究に従事. IEEE, 電子情報通信学会, 日本音響学会, 映像情報メディア学会各会員.



迫江 博昭（正会員）

昭和 41 年九州工業大学制御工学科卒業．昭和 43 年九州大学大学院修士課程（電子）修了．同年日本電気株式会社入社．平成元年九州大学に移り，現在システム情報科学研究院

知能システム学部門教授．工学博士．音声認識および画像パターン解析の研究に従事．IEEE ASSP Senior Award，電子情報通信学会業績賞，電子情報通信学会米沢ファウンダーズメダル受賞記念特別賞および論文賞（社）日本発明協会・内閣総理大臣発明賞，九工大嘉村記念賞各受賞．IEEE，電子情報通信学会（フェロー），日本音響学会，映像情報メディア学会各会員．

---