

# 弾性マッチングに基づくオンライン文字認識における固有変形の利用

三苫寛人<sup>†</sup> 内田誠一<sup>††</sup> 迫江博昭<sup>††</sup>

<sup>†</sup> 九州大学大学院システム情報科学府

<sup>††</sup> 九州大学大学院システム情報科学府

〒 812-8581 福岡市東区箱崎 6-10-1

E-mail: {mitoma, uchida, sakoe}@human.is.kyushu-u.ac.jp

あらまし 弾性マッチング (非線形伸縮マッチング) 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。例えば入力パターンが「1」であっても、弾性マッチングによって水平部が非線形伸縮された結果、「7」に誤認識される場合がある。本報告では、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には、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吸収の対象としていることが挙げられる。そこで本手法では、あらかじめ各カテゴリに生じ易い変形 (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、その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過変形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。

キーワード オンライン文字認識, 固有変形, DP マッチング, 主成分分析

## Elastic Matching-Based Online Character Recognition with Eigen-Deformations

Hiroto MITOMA<sup>†</sup>, Seiichi UCHIDA<sup>††</sup>, and Hiroaki SAKOE<sup>††</sup>

<sup>†</sup> Graduate School of Information Science and Electrical Engineering, Kyushu University

<sup>††</sup> Faculty of Information Science and Electrical Engineering, Kyushu University

Hakozaki 6-10-1, Higashi-ku, Fukuoka-shi, 812-8581 Japan

E-mail: {mitoma, uchida, sakoe}@human.is.kyushu-u.ac.jp

**Abstract** In online character recognition based on elastic matching, such as DP matching, 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.

**Key words** Online character recognition, eigen-deformations, DP matching, principal component analysis

### 1. ま え が き

オンライン手書き文字認識における文字線 (ストローク) の非線形変形の吸収法として、DP マッチング [1] ~ [3] ならびにその確率的拡張である HMM [4] ~ [6] が利用されている。DP マッチングは弾性マッチング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より用いられている。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。例えば、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の各筆点毎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

間の局所的変形しか制御できない点も問題として残る。つまり、「1」への合わせ過ぎを避けるためにカテゴリ「7」の書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[3]やHMM[4]~[6]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。例えば、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。従って、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。後に示すように各カテゴリの固有変形は、DPと主成分分析を組み合わせて用いて統計的に推定できる。

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[7],[8]と同様、2つの方針が考えられる。すなわち、(i)事後評価による方針と、(ii)直接組み込みによる方針が考えられる。本報告では(i)の方針を採用する。この方針では、まず通常通り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。この方針は、方針(ii)に比べて考え方が容易であるため、検討の第一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[9],[10]。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。関連して、DPマッチングで2つのオンライン文字パターンを整合し、その後統計量の一つである平均文字パターンを求めるといった方法もすでに検討されている[11]。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言える。

## 2. オンライン手書き文字の固有変形の推定

本節では、DPマッチングと主成分分析の組合せによるオンライン手書き文字の固有変形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の方向性であり、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

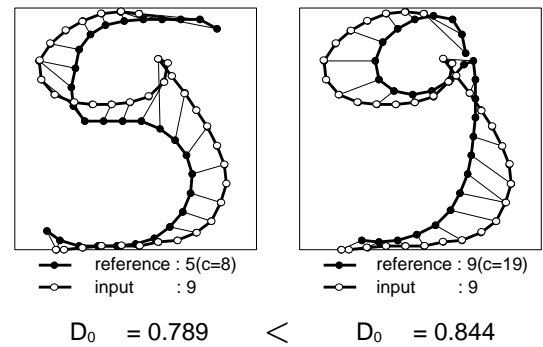


図1 合わせ過ぎにより誤認識された文字の例。括弧内の $c$ の値は図2の標準パターンの番号 $c$ に対応。

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

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、方向を要素とする3次元ベクトルとする。なお、 $r_i, t_j, I$ はそれぞれ $r_{c,i}, t_{c,n,j}, I_c$ のように記述するべきだが、誤解のない限り簡略化して $r_i, t_j, I$ と表現する。

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(整数)]

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

[制約条件]

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

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、

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

この値は、変形吸収後の $R_c$ と $T_{c,n}$ の距離であり、これを識別基準とすることで変形に頑強な認識が可能となる。

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は変位ベクトル[12]として表すことができ

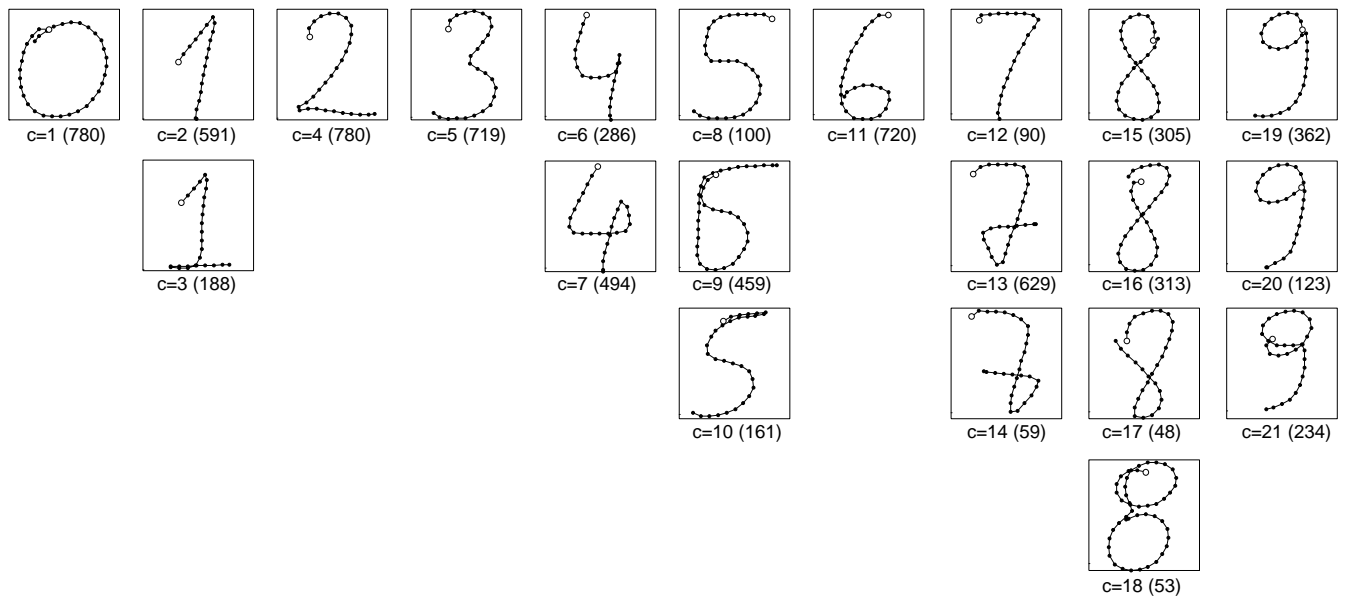


図 2 実験で用いた標準パターン．白点は開始点，括弧内の数字は学習サンプル数  $N_c$ ．

る．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，従って変形の成す分布の分散の大きな方向，いわゆる主軸に対応する．従って固有変形は，前ステップで求まった  $N_c$  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．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  $\{\lambda_{c,1}, \dots, \lambda_{c,m}, \dots, \lambda_{c,M}\}$  として各個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．従って，第 1 固有変形  $u_{c,1}$  が最頻変形であり，また高次の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においては，1. 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 に合わせ過ぎにより誤認識された文字の例を示す．図のように入力パターン「9」は，標準パターン「5」と不自然に対応づけられているが，各筆点間の距離の総和，すなわち DP マッチング距離  $D_0$  が最も小さくなった「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の方向特性が考慮されていないためと考えられる．

そこで本節では，DP ベースのオンライン文字認識法における固有変形の一利用法について述べる．具体的には，従来通り，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，その後そのマッチングの結果と固有変形の整合度を事後評価するという方針を採用する．合わせ過ぎを起こした際のマッチング結果は固有変形では表現できないはずであるから，両者の整合度は低くなると考えられる．すなわち，両者の距離は大きくなる．この事後評価で得られる距離を  $P(R_c, E)$  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

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

を新たな距離として用いれば，合わせ過ぎが起きた際は距離値が大きくなる．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 ( $0 < \alpha < 1$ ) である．

事後評価の距離として，マハラノビス距離

$$P(R_c, E) = (v - \bar{v}_c)^T \Sigma_c^{-1} (v - \bar{v}_c) = \sum_{m=1}^M \frac{1}{\lambda_{c,m}} ((v - \bar{v}_c)^T u_{c,m})^2 \quad (10)$$

が考えられる．ところで，一般に高次の固有値  $\{\lambda_{c,m}\}$  は非常に小さな値となり，数値誤差も含みやすい．よって，式 (10) のようにその逆数を使用した場合，数値誤差が拡大し，その影響が無視できなくなる．このため，本報告では次の疑似マハラノ

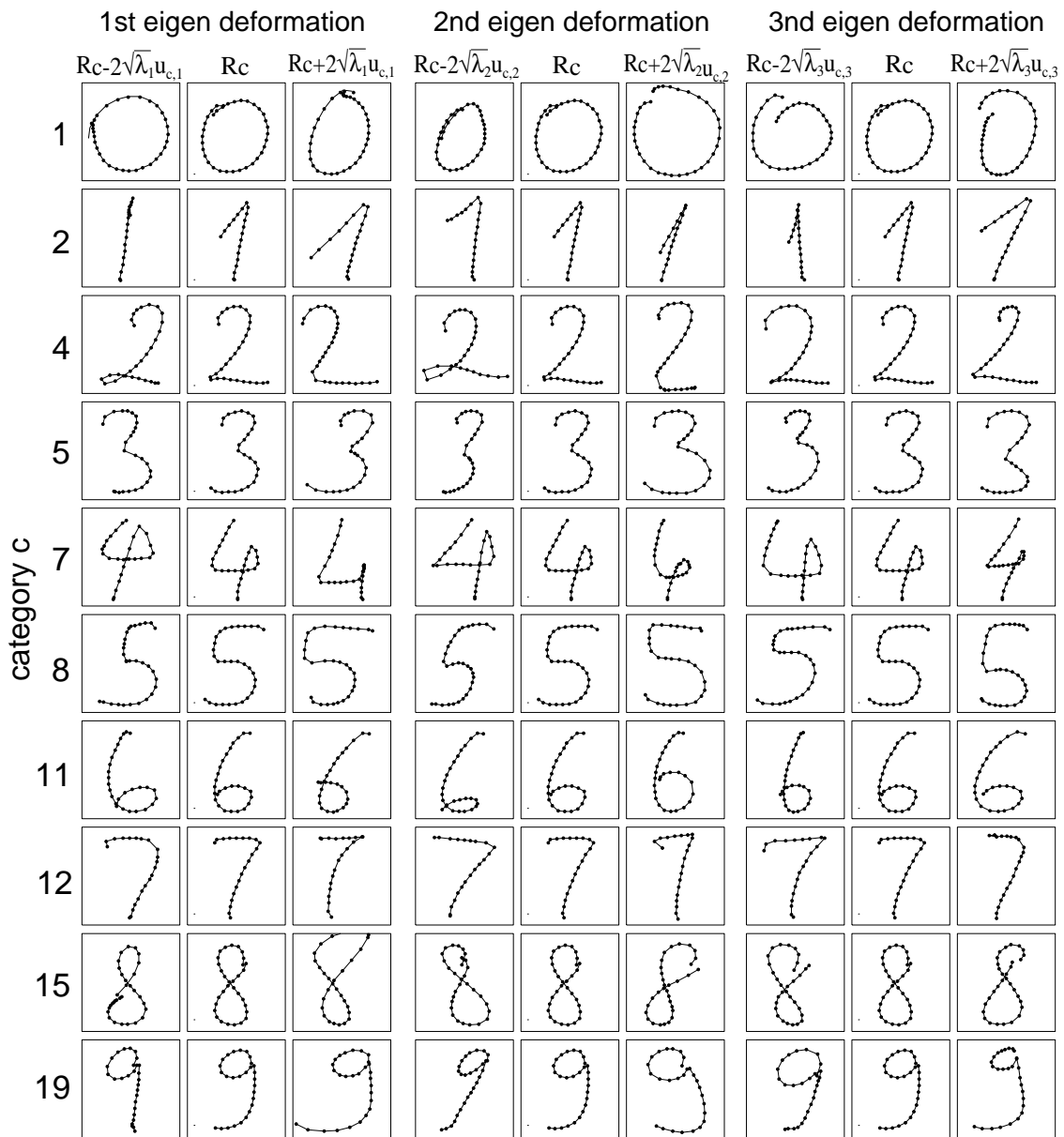


図 3 上位 3 つの固有変形  $u_{c,1}, u_{c,2}, u_{c,3}$  によって変形された標準パターン

ビス距離 [13]

$$\begin{aligned}
 P(R_c, E) & \\
 & \sim \sum_{m=1}^{M'} \frac{1}{\lambda_{c,m}} ((v - \bar{v}_c)^T u_{c,m})^2 \\
 & + \sum_{m=M'+1}^M \frac{1}{\lambda_{c,M'+1}} ((v - \bar{v}_c)^T u_{c,m})^2 \\
 & = \frac{1}{\lambda_{c,M'+1}} \|v - \bar{v}_c\| \\
 & + \sum_{m=1}^{M'} \left( \frac{1}{\lambda_{c,m}} - \frac{1}{\lambda_{c,M'}} \right) ((v - \bar{v}_c)^T u_{c,m})^2 \quad (11)
 \end{aligned}$$

を用いる。これにより、高次の固有値に含まれる数値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が、本報告では累積寄与率  $\sum_{m=1}^{M'} \lambda_{c,m} / \sum_{m=1}^M \lambda_{c,m}$  が、あるしきい値  $\theta$  をはじめて越

えた時の  $M'$  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 とした。

## 4. 実験

### 4.1 実験試料

本報告では、実験試料として Ethem Alpaydin Digit [14] の孤立数字データベースを用いた。本データベースは、予め 7494 個の学習用サンプルと 3498 個のテスト用サンプルに分けられている。従って、各文字種平均で学習用 749 サンプル、テスト用 350 サンプル、総数 10992 サンプルから構成されている。以下の実験では、学習用のサンプルを固有変形推定に、テスト用サンプルを認識タスクとして用いた。

これら全てのサンプルについて、画数変動の影響を無くす為に、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

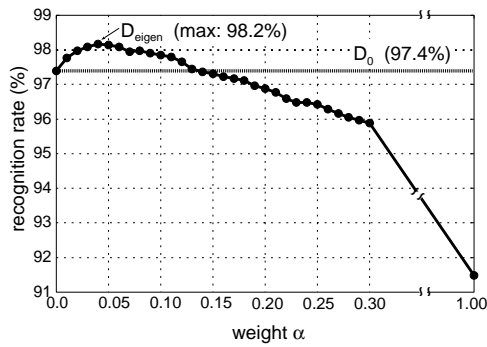


図 4 手書き数字データ (全 3498 サンプル) の認識率

表 1 本手法により低減した主な誤認識

input→result	#misrecognitions	
	by $D_0$	by $D_{eigen}$
「9」→「5」	10	1
「0」→「8」	16	8
「1」→「7」	6	0
「2」→「7」	6	1
「3」→「7」	9	6

表 2 本手法により増加した誤認識

input→result	#misrecognitions	
	by $D_0$	by $D_{eigen}$
「5」→「9」	5	8
「7」→「9」	8	11
「7」→「2」	0	1

ングを行った。

#### 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、手動で標準パターン  $R_c$  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。

なお、学習サンプルを観察したところ、テストサンプルとは傾向が異なる場合があることがわかった。例えば  $c = 17$  のような、逆向きに書かれた「8」については、学習パターンには 48 サンプルあったにもかかわらず、テストサンプルには皆無であった。また、 $c = 18$  についてもテストサンプルにはわずかに 1 サンプルのみであった。以下の実験結果には、この不整合による悪影響が多少あると思われる。

#### 4.3 固有変形の推定結果

2. 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 (カテゴリ) しか用意しなかった文字 (すなわち「0」「2」「3」「6」) については、標準パターン  $R_c$  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  $\{v_{c,1}, \dots, v_{c,n}, \dots, v_{c,N_c}\}$  を採集し、主成分分析に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

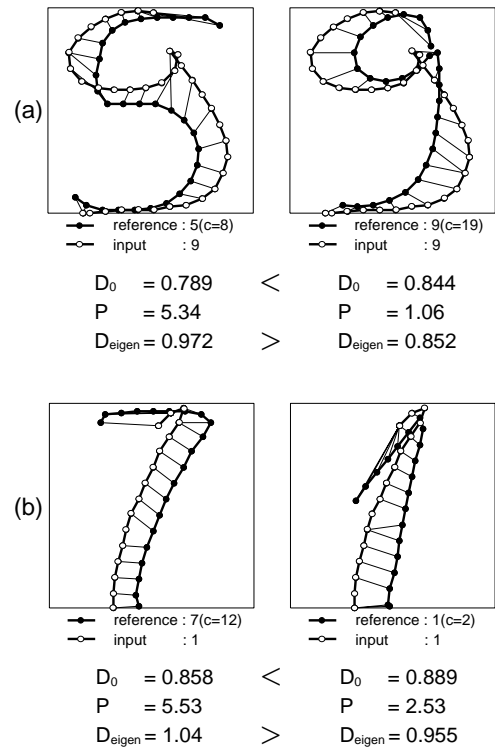


図 5 固有変形の利用による改善例

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のカテゴリの学習パターンとして分割した。その後、各標準パターン毎に、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。なお、図 2 に示したように、この分割の結果、1カテゴリあたりの学習パターン数は最少でも 48 であった。

図 3 は、10 カテゴリについて、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わかる。例えば、 $c = 1$  (「0」) の第 1 固有変形は、書き出し点の変化であるとわかる。また、 $c = 11$  (「6」) 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わかる。

#### 4.4 認識結果および考察

図 4 に、重み  $\alpha$  を変えながら測定した、本手法すなわち  $D_{eigen}$  に基づく最短距離法を用いた場合の認識率を表す。なお、 $\alpha = 0$  の場合が、固有変形を用いない従来法に対応する。

本図より本手法により最高で認識率 98.2% が得られたことがわかる。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、34 サンプルの改善と 7 サンプルの改悪であった。

本手法により低減された誤認識を表 1 に示す。ただし、表中で使用されている「9」「5」という表記は、入力が「9」で認識結果が「5」のサンプルを表す。それら改善されたサンプルをすべて観察してみると、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わかった。その具体例を図 5 に

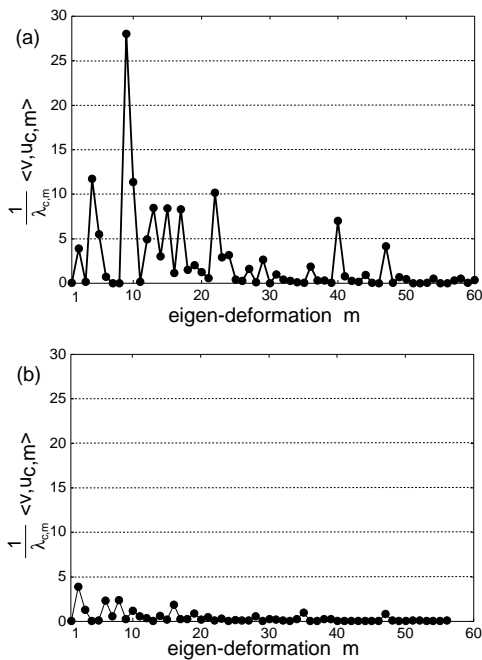


図 6 (a) 標準「5」と入力「9」、および (b) 標準「9」と入力「9」のマッチングにおける各固有変形の成分

示す．同図 (a) は「9」「5」の誤認識が改善された例，同図 (b) は「1」「7」の誤認識が改善された例である．同図 (a) では， $D_0$  に基づく識別では，標準パターン「9」より標準パターン「5」の方が全体的に入力パターン「9」との各筆点間の距離が近いので「5」に誤認識されている．ところが「5」には上の水平部と下のカーブを描く部分が近付くような変形は起こらない．その結果，図 6(a) に示すように「5」と「9」のマッチングにはほとんど起こらないような高次の固有変形成分が発生している．一方，図 6(b) に示すように「9」と「9」のマッチングには高次の固有変形成分はほとんど発生しない．これにより，事後評価量  $P$  は「5」と「9」のマッチングの方が大きくなっている．よって， $D_{\text{eigen}}$  による識別では，正しく認識されている．また，図 5(b) でも，標準パターン「7」の水平部が，非常に短い入力パターン「1」の水平部に不自然に対応づけられているが， $D_0$  における影響は少なく，結果として従来法，すなわち  $D_0$  に基づく識別では，誤認識されている．一方，事後評価量  $P$  を見ると，こうした不自然な対応付けを含む「1」と「7」のマッチングの方が大きくなっている．結果的に， $D_{\text{eigen}}$  による識別では，正解に転じている．このように，本手法を用いることにより，合わせ過ぎに起因する誤認識を抑制できていることが確認できた．

本手法により増加した誤認識を表 2 に示す．その結果「7」「9」のサンプルや「5」「9」のサンプルは学習パターンには見られないものであった．そのため，求めた固有変形にはそのような変形は含まれておらず，事後評価での距離が大きくなり誤認識になったものと考えられる．これは学習サンプル数を増やすことにより改善されると思われる．

## 5. まとめ

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 1 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わかった．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へ直接組み込む方針の検討が挙げられる．この方針を採用することで，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．

## 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．

## 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] K. Yoshida and H. Sakoe “Online handwritten character recognition for a personal computer system”, IEEE Trans. Consumer Electronics, vol. CE-28, no. 3, pp. 202-209, 1982.
- [3] C. Bahlmann and H. Burkhardt, “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] J. Hu, M.-K. Brown, and W. Turin, “HMM based online 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] S. Uchida and H. Sakoe, “Eigen-deformations for elastic matching based handwritten character recognition,” Pattern Recognit., 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,” 情処学論, 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] <ftp://ftp.ics.uci.edu/pub/machine-learning-databases/pendigits/>