社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

固有変形を用いたオンライン文字認識の検討

三苫　寛人† 　内田　誠一†† 　迫江　博昭††

† 九州大学工学部電気情報工学科
†† 九州大学大学院システム情報科学研究院
〒 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　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．例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ぐ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．
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

Online Character Recognition Using Eigen-Deformations

Hiroto MITOMA†, Seiichi UCHIDA††, and Hiroaki SAKOE††

† School of Engineering, Kyushu University
†† 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, many of misrecognitions are often due to overfitting, which is the phenomenon that a wrong reference pattern is closely fitted an input pattern. In this report, a technique to suppress those misrecognitions is proposed, where frequent deformations of each category, called eigen-deformations, are utilized. In case of overfitting, the matching between the two patterns will diverge from the eigen-deformations of the category of the reference pattern. Thus, the overfitting can be detected by evaluating the divergence. The result of a 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言える．本報告では，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 [1], [2] 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．例えば，「1」への合わせ過ぎを避けるためにカテゴリ「7」の書出し点と最終点の相対位置関係を制約するといった，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 HMM [3] ~ [5] においても同様である．

本報告では，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．一方，書き出し点と最後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，もし固有変形が事前に把握できていれば，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．

固有変形は，主成分分析を用いて統計的に推定される．具体的には，まず同じカテゴリの文字対間で実際に DP マッチングを行い，その結果からそのカテゴリに起こりうる変形を求める．次に，この処理で採集された多数の変形に対し，主成分分析を施すことで，起こりやすい変形，すなわち固有変形が主成分として求まる．後に実験的に示すように，この枠組みにより求められた固有変形は，各カテゴリに実際に生じる変形を適切に表現したものとなる．

こうして求めた固有変形と DP マッチングを如何に組み合わせるかについては，オフライン文字認識の場合 [6], [7] 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 [8], [9]．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，その後に統計量の一種である平均文字パターンを求めるという方法もすでに検討されている [10]．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，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

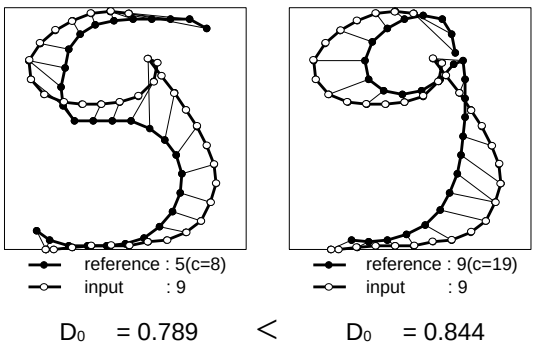

図 1 合わせ過ぎによる誤認識された文字の例．
Fig. 1 Example of overfitting.

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 $\boldsymbol{R}_c$ を時系列

$$\boldsymbol{R}_c = r_1, r_2, \ldots, r_i, \ldots, r_I \tag{1}$$

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

$$\boldsymbol{T}_{c,n} = t_1, t_2, \ldots, t_j, \ldots, t_J \tag{2}$$

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，

$$r_i = (x_i^r, y_i^r, d_i^r)^T \tag{3}$$

$$t_j = (x_j^t, y_j^t, d_j^t)^T \tag{4}$$

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 はそれぞれ $r_i^c, t_j^{c,n}$ のように記述するべきだが，誤解のない限り r_i, t_j と表現する．

2 パターン $\boldsymbol{R}_c$ と $\boldsymbol{T}_{c,n}$ 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]

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

[制御変数 (自然数)]

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

[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 $D_0(\boldsymbol{R}_c, \boldsymbol{T}_{c,n})$ とする．すなわち，

$$D_0(\boldsymbol{R}_c, \boldsymbol{T}_{c,n}) = \min_{j(1),\ldots,j(i),\ldots,j(I)} \sum_{i=1}^{I} \left\| r_i - t_{j(i)} \right\| \tag{5}$$

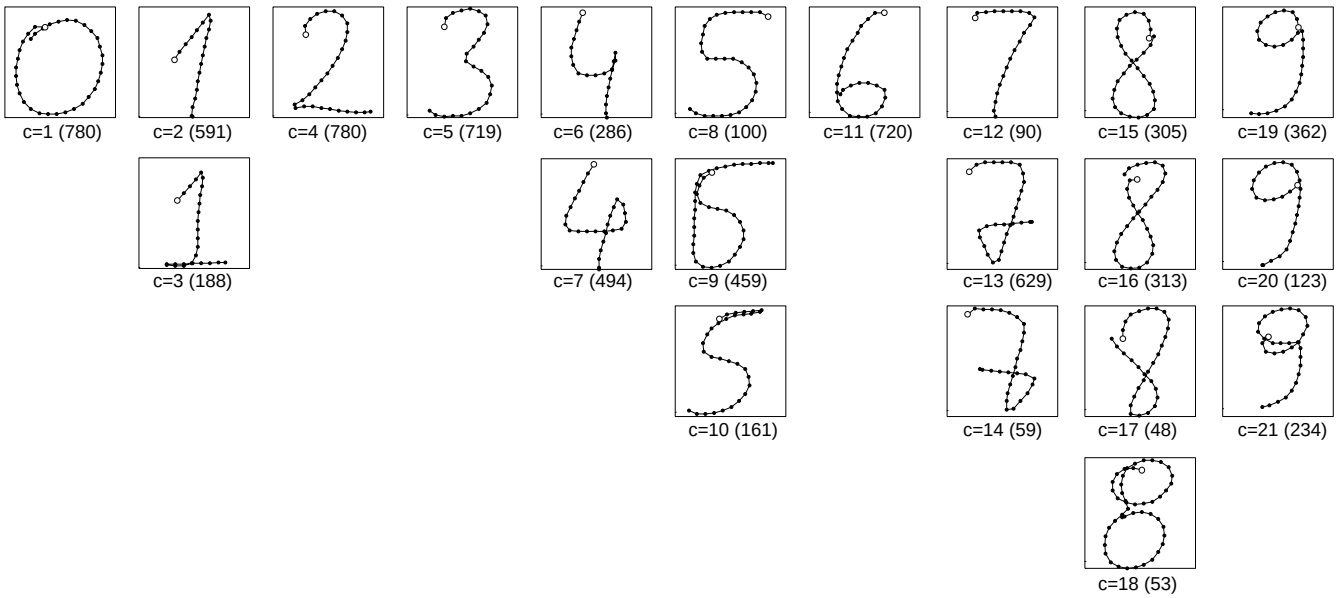

図 2　実験で用いた標準パターン．白点は開始点，括弧内の数字は学習サンプル数 N_c ．
Fig. 2　Reference patterns used in the experiment.

この値は，変形吸収後の $\boldsymbol{R}_c$ と $\boldsymbol{T}_{c,n}$ 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 $j(1),\ldots,j(I)$ 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 $\boldsymbol{T}_{c,n}$ の $2I$ 次元 変位ベクトル $\boldsymbol{v}_{c,n}$ は次のように表現される．

$$\boldsymbol{v}_{c,n} = \big((x^r_1 - x^t_{j(1)}, y^r_1 - y^t_{j(1)}),\ldots, (x^r_i - x^t_{j(i)}, y^r_i - y^t_{j(i)}),\ldots,(x^r_I - x^t_{j(I)}, y^r_I - y^t_{j(I)})\big)^T \quad (6)$$

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 $\boldsymbol{v}_{c,n}$ 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

$$\overline{\boldsymbol{v}}_c = \frac{1}{N_c}\sum_{n=1}^{N_c}\boldsymbol{v}_{c,n} \quad (7)$$

および共分散行列

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

を求める．次にこの Σ_c を固有値展開すれば，各固有ベクトル $\{\boldsymbol{u}_{c,1},\ldots,\boldsymbol{u}_{c,m},\ldots,\boldsymbol{u}_{c,M}\}$ として固有変形が得られる．次数 m は固有変形の寄与 (固有値 $\{\lambda_{c,m}\}$) の大きさ順に対応する．従って第 1 固有変形 $\boldsymbol{u}_{c,1}$ 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. 節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．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固有の変形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 $\boldsymbol{R}_c$ とカテゴリ未知の入力パターン $\boldsymbol{E}$ 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．これを両者の距離が大きくなったと考えると，結局，DP マッチング距離 $D_0(\boldsymbol{R}_c,\boldsymbol{E})$ に，この事後評価で得られる距離

$$P(\boldsymbol{R}_c,\boldsymbol{E}) = (\boldsymbol{v}-\overline{\boldsymbol{v}}_c)^T\Sigma_c^{-1}(\boldsymbol{v}-\overline{\boldsymbol{v}}_c) = \sum_{m=1}^{M}\frac{1}{\lambda_{c,m}}((\boldsymbol{v}-\overline{\boldsymbol{v}}_c)^T\boldsymbol{u}_{c,m})^2 \quad (9)$$

を加えたもの，すなわち

$$D_{\mathrm{eigen}}(\boldsymbol{R}_c,\boldsymbol{E}) = (1-\alpha)D_0(\boldsymbol{R}_c,\boldsymbol{E}) + \alpha P(\boldsymbol{R}_c,\boldsymbol{E}) \quad (10)$$

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．

一般に高次の固有値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 (9) 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，本報告では次の修正マハラノビス距離 [11]

$$P(\boldsymbol{R}_c,\boldsymbol{E}) \sim \sum_{m=1}^{M'}\frac{1}{\lambda_{c,m}}((\boldsymbol{v}-\overline{\boldsymbol{v}}_c)^T\boldsymbol{u}_{c,m})^2 + \sum_{m=M'+1}^{M}\frac{1}{\lambda_{c,M'+1}}((\boldsymbol{v}-\overline{\boldsymbol{v}}_c)^T\boldsymbol{u}_{c,m})^2$$

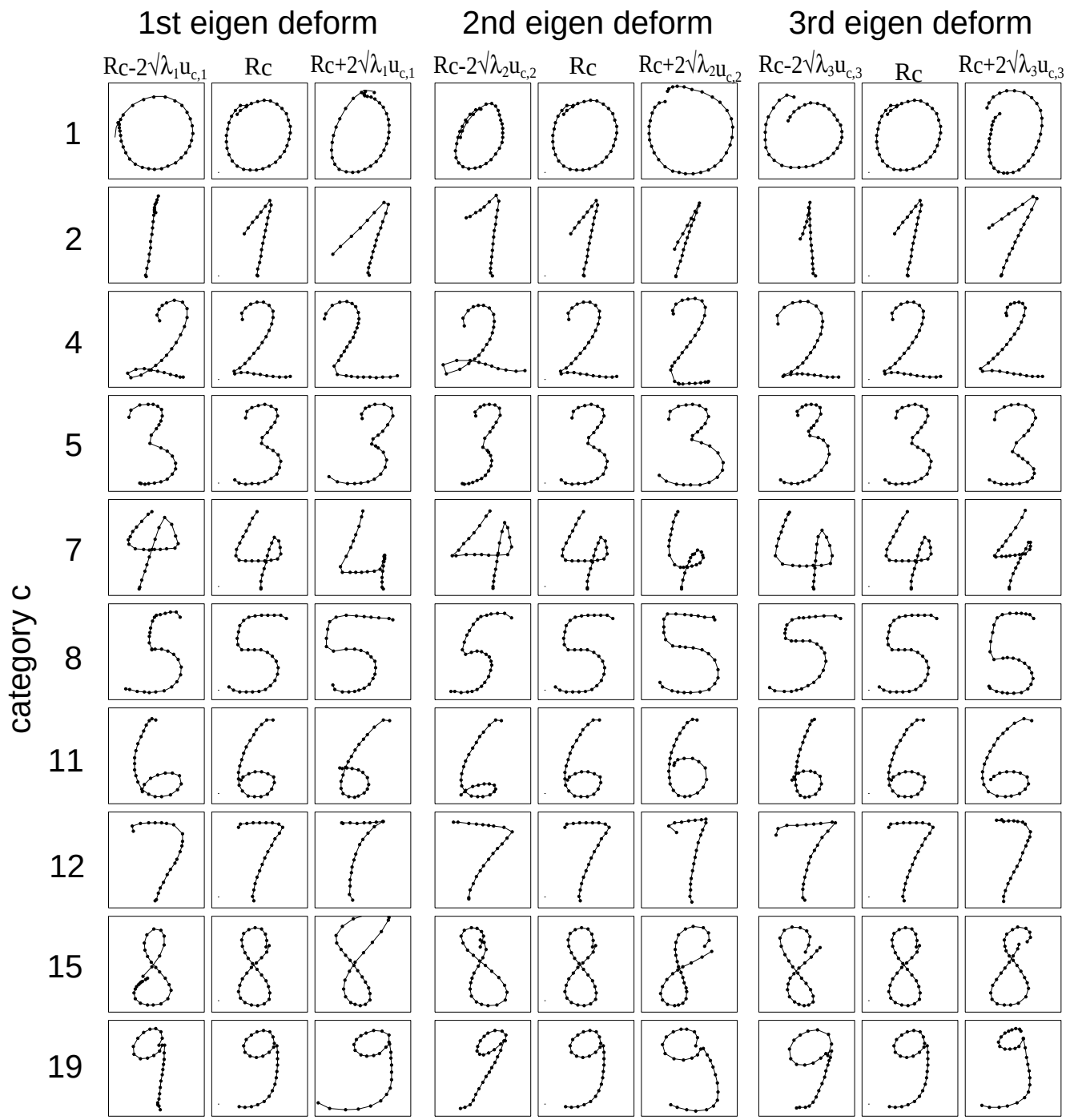

図 3　上位 3 つの固有変形によって変形された標準パターン．

Fig. 3　Reference pattern deformed by the primary eigen-deformations of its class.

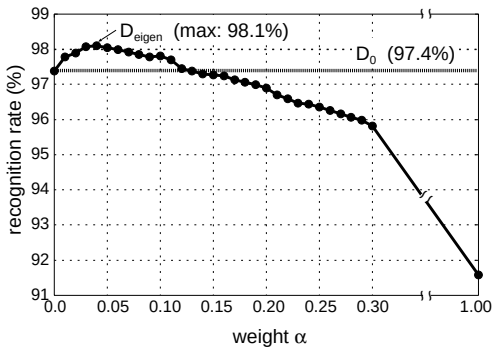

図 4　手書き数字データ (全 3498 サンプル) の認識率．
Fig. 4　Recognition rate as the function of weight α.

$$= \frac{1}{\lambda_{c,M'+1}} \|\boldsymbol{v} - \overline{\boldsymbol{v}}_c\| + \sum_{m=1}^{M'} \left(\frac{1}{\lambda_{c,m}} - \frac{1}{\lambda_{c,M'}}\right) \left((\boldsymbol{v} - \overline{\boldsymbol{v}}_c)^T \boldsymbol{u}_{c,m}\right)^2 \quad (11)$$

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，上述の問題を回避できる．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 θ 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 θ 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 [12] 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 3498 個の学習用サンプルと 7494 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 × 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．

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 $\boldsymbol{R}_c$ 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．

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．

表 1　本手法により低減した誤認識．
Table 1　Decreased misrecognitions.

誤認識	従来法	本手法
「9」→「5」	10	1
「0」→「8」	16	8
「1」→「7」	6	0
「2」→「7」	6	1
「3」→「7」	9	6
「0」→「6」	1	0
「4」→「1」	1	0
「4」→「5」	1	0

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

誤認識	従来法	本手法
「5」→「9」	5	9
「7」→「9」	8	12
「7」→「2」	0	1

4.3 固有変形の推定結果

2. 節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 $\boldsymbol{R}_c$ とその文字の全学習パターン $\boldsymbol{T}_{c,n}$ の間で DP マッチングを行い，変位ベクトル $\{\boldsymbol{v}_{c,1}, \ldots, \boldsymbol{v}_{c,n}, \ldots, \boldsymbol{v}_{c,N_c}\}$ を採集し，主成分分析により最終的に固有変形 $\boldsymbol{u}_{c,m}$ 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 $\boldsymbol{u}_{c,1}$, $\boldsymbol{u}_{c,2}$,$\boldsymbol{u}_{c,3}$ により標準パターン $\boldsymbol{R}_c$ 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

図 4 に，重み α を変えながら測定した，本手法すなわち D_{eigen} に基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．

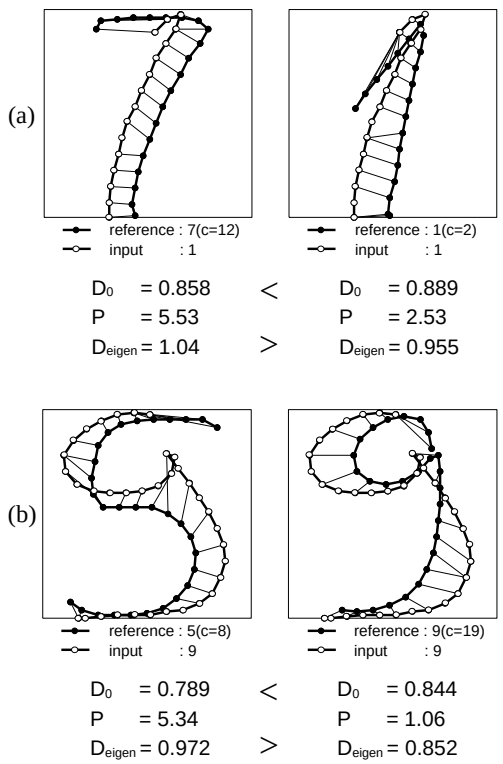

図 5　固有変形の利用による改善例

Fig. 5　Characters correctly recognized by the proposed technique.

本図より本手法により最高で認識率 98.1% 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 0.7%(25 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 9 サンプルの改悪であった．

4.5　合わせ過ぎに起因する誤認識の抑制効果の確認

本手法により低減された誤認識を表 1 に示す．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．図 5(a) に「１」→「７」の誤認識が改善された例を，同図 (b) に「９」→「５」の誤認識が改善された例を示す．同図 (a) では，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_{eigen} による識別では，正解に転じている．また，同図 (b) でも，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ため，「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ので，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_{eigen} による識別では，正しく認識される．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．

また，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，求まっ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

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) のオンライン文字認識における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] J. Hu, M.-K. Brown, and W. Turin, “HMM based online handwriting recognition,” IEEE Trans. PAMI, vol. 18, no. 10, pp. 1039-1045, 1996.

[4] 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,” 信学技法, PRMU97-85, 1997.

[5] 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,” 信学技法, PRMU2000-35, 2000.

[6] S. Uchida and H. Sakoe, “Eigen-deformations for elastic matching based handwritten character recognition,” Pattern Recognit., vol. 36, no. 9, pp.2031-2040, 2003.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] F. Kimura, et al., Modified quadratic discriminant functions and the application to Chinese character recognition, IEEE Trans. PAMI, vol. 9, no. 1, pp. 149-153, 1987.

[12] ftp://ftp.ics.uci.edu/pub/machine-learning-databases/pendigits/