論　文

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

内田　誠一[†]　　迫江　博昭[†]

Category-dependent elastic matching based on a linear combination of eigen-deformations

Seiichi UCHIDA[†] and Hiroaki SAKOE[†]

あらまし　画像パターンの認識において，パターンに生じた変形を補償するための手法として，弾性マッチングの利用が検討されている．従来法がすべてのカテゴリに共通の変形特性を仮定していたのに対し，本論文では各カテゴリに固有の変形特性を組み込んだ手法を提案する．具体的には，各カテゴリの任意の変形をそのカテゴリに固有ないくつかの変形の線形結合で表現する．その結果，各カテゴリ内に生じる変形だけが適切に補償されることになり，過変形の抑制および計算効率の向上といった効果が得られる．本手法は，一種の非線形最適化問題として定式化される．本論文ではその解法についても述べ，実験を通して有効性を検証する．

キーワード　弾性マッチング, 画像パターン認識, 固有変形, 最適化, 手書き文字

1. まえがき

手書き文字，顔，指紋といった変形を伴う画像パターンの認識において，弾性マッチング法の利用が検討されている．弾性マッチング法は 2 つの画像パターンの最適な非線形整合を目的とした最適化問題として定式化される．最適整合の下での 2 パターン間の距離は変形補償後の 2 画像間の距離であり，従って変形不変量となる．よってそれを識別関数として用いることで，変形に対して頑強な識別系を実現できる．

弾性マッチング法の性質は，ベースとなる変形モデルに依存する．変形モデルとは補償しうる変形の範囲を規定するもので，制約条件や変形の関数表現といった形で最適化問題に組み込まれる．従来この変形モデルとしては，カテゴリ非依存，すなわちすべてのカテゴリで同一のものが用いられていた． これはすべてのカテゴリが同一の変形特性を持つという仮定に基づいている． 例えばアフィン変形モデルを用いた場合，アフィン変換の範囲で許される変形が，すべてのカテゴリにおいて同じように発生することを仮定していることになる．ところが，実際の画像パターンに生じた変形を見ると，こうした仮定は必ずしも妥当ではない．例えば手書き文字パターン “M” の 2 本の縦ストロークはしばしば “ハ” の字状に傾くが，こうした変形は “H” には見られない (図 1)．この単純な例からも予想されるように，カテゴリ非依存の変形モデルと用いると，実際に生じる変形と弾性マッチングの変形補償特性の間にずれが生じる．その結果，あるカテゴリにおいては類似カテゴリへの合わせ過ぎ (過変形) が生じ，同時に他のカテゴリでは変形補償能力の不足が生じるといった問題が発生する．

本論文では，カテゴリ依存の変形モデルおよびそれに基づく弾性マッチング法を提案する．本モデルにおいては，各カテゴリに生じる任意の変形を，そのカテゴリに生じる代表的な幾つかの変形 (以下，固有変形と呼ぶ) の線形結合によって表現する．本変形モデルを用いることで，カテゴリ内に生じる変形だけを適切に補償するような弾性マッチング法を実現でき，前述の問題を回避できると考えられる．本手法の効果として，まず認識率の向上が期待できる．さらに，あるカテゴリに発生しない変形は補償範囲すなわち弾性マッチングの解空間から排除されるので，計算効率の向上も見込める．本論文ではこれら 2 つの効果について，本手法の従来法に対する優位性を実験的に検証する．

本手法の実装に関しては，2 つの課題がある．第 1

[†] 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
Faculty of Information Science and Electrical Engineering, Kyushu University
Hakozaki 6–10–1, Higashi-ku, Fukuoka-shi, 812–8581 Japan

図 1　各カテゴリに固有の変形方向が存在することの例．
Fig. 1　Category-dependency of deformation.

は，いかに各カテゴリの固有変形を求めるかという課題である．これに対して本論文では，文献 [1], [2] の手法を利用することとする．第 2 の課題は，本変形モデルに基づく弾性マッチング問題の実際の解法，すなわちアルゴリズムの検討である．本手法は非線形最適化問題として定式化されるので，その厳密解を解析的に求めることは難しい．そこで本論文では，Simard ら [3] ~ [5] による tangent distance 法 (TD) の考えを援用した線形近似解法を提案する．また，他に目的関数の勾配を利用した反復解法についても触れる．

本手法は原理的に任意の画像パターンに適用可能であるが，本論文では手書き文字パターンを例として採用する．従って，この手書き文字パターンの認識問題を通して本手法の性能を評価することにする．

パターンに固有な変形方向を推定し利用するという考え方は，Cootes ら [6] による Point Distribution Models (PDM) に見ることができる．PDM は当初 1 次元パターン (輪郭) を 2 次元画像に最適整合させるためのものとして開発され，その後は顔認識や文字認識 [7] に利用されている．筆者らは，この PDM を 2 次元的に拡張したものを固有変形と呼んで，その自動推定法を提案し，手書き文字認識に利用した [1], [2]．ただしその認識過程において，固有変形は従来の弾性マッチング (すなわちカテゴリ非依存の変形モデルに基づく弾性マッチング) の事後評価において利用されていた．こうした間接的な利用形態に対し，本手法では，固有変形を弾性マッチングに埋め込むことで，より直接的に利用する．なお，5. 1 節では再びこの従来法 [1], [2] と本手法の比較を行う．

2. 定　式　化

2. 1　弾性マッチング

本論文で提案するカテゴリ依存の変形モデルに基づく弾性マッチング法は，他の弾性マッチング法と同じく，一種の最適化問題として定式化される．本節では，本手法の定式化に先立ち，一般の弾性マッチング法が基づくこの最適化問題ついて述べる．弾性マッチング法はパラメトリックな変形モデルに基づくものとノンパラメトリックな変形モデルに基づくものに大別される．前者は画素の変位を幾つかのパラメータからなる関数で表現し，そのパラメータを制御変数として最適整合を図る．後者は各画素の変位を直接制御変数として扱い，最適化を図る．本手法はこのうち前者に属するので，以下ではそれに限定して議論する．

パラメトリックな変形モデルに基づく 2 画像パターン間の弾性マッチングは，一般に次の目的関数 $J(\alpha)$ の最小化問題として定式化される．

$$J(\alpha) = \|s(\alpha, \boldsymbol{P}_c) - \boldsymbol{E}\| \tag{1}$$

ここで $\boldsymbol{P}_c = \{P_c(x,y)\}$ はカテゴリ c の標準パターンであり，$\boldsymbol{E} = \{E(x,y)\}$ は入力パターンである．両パターンは共に $I \times I$ 画像 (すなわち $x, y = 1, \ldots, I$) とする．関数 $s(\alpha, \boldsymbol{P}_c)$ は，M 個のパラメータ $\alpha = \{\alpha_m | m = 1, \ldots, M\}$ によって制御される変形モデルであり，何らかの 2 次元-2 次元写像により $\boldsymbol{P}_c$ の変形画像を与えるためのものである．パラメータ α を制御して $J(\alpha)$ の最小化を図ることは，$\boldsymbol{P}_c$ をモデル $s(\alpha, \boldsymbol{P}_c)$ に従って変形し，$\boldsymbol{E}$ に最も近づけることに相当する．

目的関数 $J(\alpha)$ の最小値を $d(\boldsymbol{P}_c, \boldsymbol{E})$ とする．すなわち，

$$d(\boldsymbol{P}_c, \boldsymbol{E}) = \min_{\alpha} J(\alpha) \tag{2}$$

この値 $d(\boldsymbol{P}_c, \boldsymbol{E})$ は，標準パターン $\boldsymbol{P}_c$ と，$\boldsymbol{P}_c$ を基準として変形補償された入力パターン $\boldsymbol{E}$ との単純重ね合わせ距離である．もし変形モデルが $\boldsymbol{E}$ に生じている変形を補償できるならば，$d(\boldsymbol{P}_c, \boldsymbol{E})$ は $\boldsymbol{E}$ の変形に依らず一定となるので，それを識別関数とすれば変形に致して頑強な認識系を構成できる．

2. 2　カテゴリ依存変形モデル

カテゴリ c において m 番目に代表的な変形方向をカテゴリ c の第 m 固有変形と呼ぶ．この固有変形は，それによる画素 (x,y) の変位 $(X_{c,m}(x,y), Y_{c,m}(x,y))$ を列挙した $2I^2$ 次元ベクトル $\boldsymbol{u}_{c,m}$ により表わされる．すなわち，

$$\begin{aligned} \boldsymbol{u}_{c,m} = & ((X_{c,m}(1,1), Y_{c,m}(1,1)), \\ & \ldots, (X_{c,m}(x,y), Y_{c,m}(x,y)), \\ & \ldots, (X_{c,m}(I,I), Y_{c,m}(I,I))) \end{aligned} \tag{3}$$

本手法では，この固有変形を組み込んだ次式のカテゴリ依存な変形モデルを利用する．

$$s(\alpha, P_c(x,y))$$

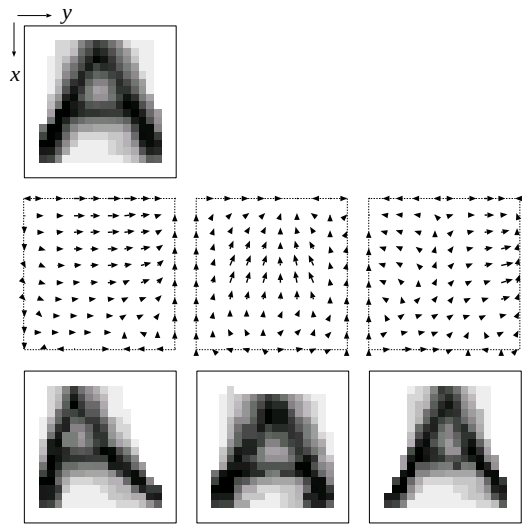

図 2　カテゴリ “A” の標準文字パターン $\boldsymbol{P}_c$ (上段), 同カテゴリの第 1,2,3 固有変形 $\boldsymbol{u}_{c,1}$, $\boldsymbol{u}_{c,2}$, $\boldsymbol{u}_{c,3}$ (中段), および各固有変形により変形された標準文字パターン $\{P_c(x + X_{c,m}(x,y), y + Y_{c,m}(x,y))\}$ (下段).

Fig. 2　Top three eigen-deformations of character “A”.

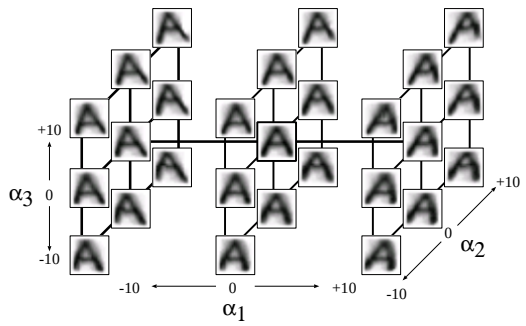

図 3　カテゴリ依存変形モデル (4) による標準パターンの変形 (第 1,2,3 固有変形を使用した場合)

Fig. 3　Character patterns created by the proposed deformation model (4).

$$= P_c\Big(x + \sum_{m=1}^{M} \alpha_m X_{c,m}(x,y),\ y + \sum_{m=1}^{M} \alpha_m Y_{c,m}(x,y)\Big) \qquad (4)$$

このように，本変形モデルでは，カテゴリ c に生じる任意の変形をそのカテゴリの M 個の固有変形 $\boldsymbol{u}_{c,1}, \ldots, \boldsymbol{u}_{c,m}, \ldots, \boldsymbol{u}_{c,M}$ の線形結合で表現している．

従来のカテゴリ非依存変形モデルの代りに，本変形モデル (4) を用いることで，2 つの効果が見込める．第 1 は，認識率の向上効果である．図 1 の例からもわかるように，すべてのカテゴリに一定した変形特性を仮定すると，カテゴリによっては変形補償能力の過不足が生じる．これに対し各カテゴリ毎にその変形特性が調整されている本モデルを用いれば，そのカテゴリ内に生じる変形を過不足なく補償されると期待でき，結果的に認識率が向上すると考えられる．第 2 は計算効率の向上効果である．従来のカテゴリ非依存変形モデルを使ってすべてのカテゴリの変形を補償しようとすれば，モデルの変形補償能力をある程度高くしておく必要があり，結果的に制御パラメータ数も多くなる．これに対し，本モデルを用いる場合は各カテゴリに生じる変形だけを考慮すればよいので，平均的に少いパラメータ数で済み，結果的に計算量を低減できると考えられる．

次節 2.3 で述べるように，固有変形は実際の画像パターンの集合から統計的に推定できる [1], [2]．図 2 は，手書き文字データ (カテゴリ “A”) について実際に第 1∼3 固有変形を求めたものである．同図中段の矢印は，各固有変形による各画素の変位 $(X_{c,m}(x,y), Y_{c,m}(x,y))$ を表している．この図より文字カテゴリ “A” の第 1 固有変形 ($\boldsymbol{u}_{c,1}$) は，全体的な傾斜変形であり，次いで水平ストロークの上下変位，「∧ 」部の膨らみ変化となっていることがわかる．

図 3 は，これら第 1∼3 固有変形をモデル (4) に従って線形結合し，その際の重み α_m を正負に変えて $\boldsymbol{P}_c$ を変形させたものである．これから，本モデルによって実際に存在するような文字パターンが生成されることがわかる．また，ここで生成された文字パターンと類似した入力パターン $\boldsymbol{E}$ については，弾性マッチング距離 $d(\boldsymbol{P}_c, \boldsymbol{E})$ は 0 が近くなるので，正解カテゴリに識別できることになる．

本モデル (4) には，重み α の範囲に関する制約がない．すなわち，変形の方向はカテゴリ毎に規定されているが，その変形の大きさは規定されていない．このため，わずかではあるが過変形の危険性は残る．また，本モデルではすべての固有変形を平等に扱っているが，次節で述べる固有変形の推定手順から明らかになるように，固有変形の寄与すなわち重要度は本来その次数 m によって異なる．後述する実験結果を見る限り，これらの点が深刻な問題にはなっていないが，今後さらにモデルの精度を上げるためには，α_m に対して範囲制約を課するような仕組みを導入する必要があると考

えられる．

2. 3 固有変形の推定

各カテゴリの固有変形 $\{\boldsymbol{u}_{c,m}\}$ の推定については，文献 [1], [2] の手法を用いる．これは，従来の (カテゴリ非依存なモデルに基づく) 弾性マッチング法と主成分分析を組み合わせることにより，固有変形の自動推定を可能とした手法であり，以下に示すように 2 ステップから成る．

[ステップ 1: 変形の採集]　各カテゴリ N 個の画像パターン $\boldsymbol{T}_{c,n}(n = 1, \ldots, N)$ を学習パターンとして準備し，それぞれに含まれる変形を採集する．この変形は，従来の弾性マッチングにより得られる標準パターン $\boldsymbol{P}_c$ と各 $\boldsymbol{T}_{c,n}$ 間の最適画素対応から求まる．例えば，$\boldsymbol{P}_c$ の画素 (x, y) が弾性マッチングの結果 $\boldsymbol{T}_{c,n}$ の画素 (x', y') に対応したならば，$\boldsymbol{T}_{c,n}$ は $\boldsymbol{P}_c$ を画素 (x, y) において $(x' - x, y' - y)$ だけ変形 (変位) させたものと言える．これを画像全体について見ると，学習パターン $\boldsymbol{T}_{c,n}$ の変形は変位 $(x' - x, y' - y)$ を全 I^2 画素について列挙した $2I^2$ 次元ベクトルとして与えられる．

[ステップ 2: 変形の主成分分析]　固有変形 $\{\boldsymbol{u}_{c,m}\}$ 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 個の変位ベクトルを主成分分析することで求まる．具体的には N 個の変位ベクトルの共分散行列を固有値展開すれば，各固有ベクトル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) の大きさ順に対応し，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．

ステップ 1 で用いる従来の弾性マッチング法は，(i) 補償可能な変形の範囲が広い (回転等も補償可能)，(ii) 大局的に最適な画素対応を与えるアルゴリズムが存在，(iii) 変位ベクトルの実効次元数が $2I^2$ よりも小さい，という 3 つの要求を満たすことが望ましい．要求 (i)(ii) は正確な変形すなわち画素対応を採集するためのものである．ところで，これら (i)(ii) を満たすような弾性マッチング法をそのまま認識に用いると，過変形の危険性が高く，また一般に多くの計算量を要する点が問題となる．しかし本手法の枠組みにおいて利用する限り，これら 2 点は問題にならない．これはこの従来の弾性マッチング法が，同一カテゴリの文字パターン間のマッチングのみに利用され，かつ認識処理以前の学習処理において利用されるためである．要求 (iii) は限られた学習サンプル数でなるべく信頼性の高い固有変形を推定するためのものである．

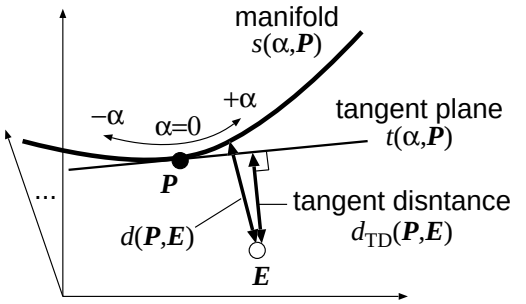

図 4　変形モデル $s(\alpha, \boldsymbol{P})$ の成す多様体，その接平面 $t(\alpha, \boldsymbol{P})$，および $\boldsymbol{E}$ と接平面との距離 $d_{TD}(\boldsymbol{P}, \boldsymbol{E})$．

Fig. 4　The manifold by the deformation model $s(\alpha, \boldsymbol{P})$, its tangent plane $t(\alpha, \boldsymbol{P})$, and the tangent distance $d_{TD}(\boldsymbol{P}, \boldsymbol{E})$.

図 5　濃淡画像表現したカテゴリ "A" の接ベクトル ϕ_1, ϕ_2, ϕ_3．(画素値は 0 から 255 の範囲に正規化してある．)

Fig. 5　Top three tangent vectors of character "A".

3. 線形近似に基づく解法

変形モデル (4) に基づく式 (1) の最小化問題は，最適化すべき制御パラメータ α が非線形関数 $P_c(x, y)$ に内包された，一種の非線形最適化問題となっていて，最適解を解析的に求めるのは困難である．そこで本節ではこの問題に対する線形近似解法について述べる．

非線形最適化問題 (1) の解き方としては，この線形近似解法以外にも種々考えられる．本論文 4. 6 節では，その一つとして，共役勾配法に基づく反復解法について触れる．

3. 1 問題の線形化

この線形近似解法は，Simard らによる tangent distance (TD) [3] ~ [5] の考え方を援用したものである．この近似解法を式 (1) の最適化に用いることで，次の 3 つの利点が得られる．すなわち，(i) 解析的に近似解が求まる．従って解の収束性等に関する議論が不要で

ある．(ii) 任意の変形画像が $M+1$ 枚の画像の線形結合により表現されるという，再び簡潔で見通しのよい形式が得られる．(iii) 単純重ね合わせの数倍程度の計算量で済む．

この解法では，$s(\alpha, \boldsymbol{P})$ を近似することで，非線形関数 (1) を線形化する (本節では，混乱のない限りカテゴリを表す添字 c を省略する)．具体的には $J(\alpha)$ を次のように近似する．

$$J(\alpha) \sim \|t(\alpha, \boldsymbol{P}) - \boldsymbol{E}\| \tag{5}$$

ここで，$t(\alpha, \boldsymbol{P})$ はモデル $s(\alpha, \boldsymbol{P})$ を $\alpha = 0$ の周りでテイラー展開して 2 次の項以降を無視したものであり，次式で与えられる．

$$t(\alpha, \boldsymbol{P}) = \boldsymbol{P} + \sum_{m=1}^{M} \alpha_m \phi_m \tag{6}$$

ただし，

$$\phi_m = \left[\frac{\partial s(\alpha, \boldsymbol{P})}{\partial \alpha_m} \right]_{\alpha=0} \tag{7}$$

である．このように制御変数 α は非線形関数 $\boldsymbol{P}$ の外に出るため，式 (5) はいわゆる最小 2 乗問題として容易に解くことができる．ただし，以上の近似は $\alpha = 0$ すなわち $\boldsymbol{P}$ の近傍でのテイラー 展開に基づくため，大きな α すなわち大きな変形に対しては近似誤差が大きくなるという欠点を持つ．

以上の線形近似解法は以下のように幾何学的に説明できる (図 4)．まず， 変形モデル $s(\alpha, \boldsymbol{P})$ に従って変形された標準パターン $\boldsymbol{P}$ の集合が，I^2 次元のパターン空間内で，M 次元多様体を成すことに着目する．このとき，$t(\alpha, \boldsymbol{P})$ は多様体 $s(\alpha, \boldsymbol{P})$ の $\alpha = 0$ における接平面であり，ϕ_m は接ベクトルとなっている．また，近似された目的関数 (5) の最小値を $d_{\mathrm{TD}}(\boldsymbol{P}, \boldsymbol{E})$ とすれば，この $d_{\mathrm{TD}}(\boldsymbol{P}, \boldsymbol{E})$ は接平面 $t(\alpha, \boldsymbol{P})$ と点 $\boldsymbol{E}$ との距離になる．なお，接平面 $t(\alpha, \boldsymbol{P})$ は斜交基底 $\phi_1, \phi_2, \ldots, \phi_M$ の張る超平面であるので，$d_{\mathrm{TD}}(\boldsymbol{P}, \boldsymbol{E})$ は一種の投影距離 [8] であると言える．

接ベクトル ϕ_m は $\boldsymbol{P}$ や $\boldsymbol{E}$ と同じく I^2 次元であり，画像として表現できる．ここで接平面 $t(\alpha, \boldsymbol{P})$ がパターン $\boldsymbol{P}$ の変形集合 $s(\alpha, \boldsymbol{P})$ の近似表現であることを考えると，画像 $\boldsymbol{P}$ の任意の変形パターンが M 枚の画像 $\phi_1, \phi_2, \ldots, \phi_M$ と $\boldsymbol{P}$ の線形結合で見通しよく近似表現されていることがわかる．

3.2 詳　　細

以上の近似解法の枠組みにおいて，モデル $s(\alpha, \boldsymbol{P})$ として，提案するカテゴリ依存変形モデルを利用することを考える．すなわち，式 (7) に式 (4) を代入する．その結果，接ベクトル ϕ_m の第 (x, y) 要素は

$$\begin{aligned} &\phi_m(x, y) \\ &= P_x(x, y) X_m(x, y) + P_y(x, y) Y_m(x, y) \end{aligned} \tag{8}$$

として与えられる．ここで，$P_x(x, y)$ と $P_y(x, y)$ は $P(x, y)$ を x および y に関して偏微分したものである．図 2 に示したカテゴリ “A” の最初の 3 つの固有変形をそれぞれ式 (8) に代入して得られた接ベクトル ϕ_1, ϕ_2, ϕ_3 を図 5 に示す．

離散的な $P(x, y)$ の偏微分値 $P_x(x, y), P_y(x, y)$ の求め方には幾つかの方法が考えられる．本論文では単純な差分近似ではなく，文献 [3] ~ [5] と同様，$\boldsymbol{P}$ をガウス関数 $G_\sigma(x, y)$ でぼかしたものを偏微分するという方法を採用する．

$$\begin{aligned} P_x(x, y) &\sim \frac{\partial P(x, y) * G_\sigma(x, y)}{\partial x} \\ &= P(x, y) * \frac{\partial G_\sigma(x, y)}{\partial x} \end{aligned}$$

ここで，記号 $*$ は畳み込み演算を表す．この方法には，誤差の大きな差分近似を回避できるという利点がある．以下では，ガウス関数の標準偏差 σ を実験的に 1.25 に設定した．文献 [4] にも同様の指摘があるように，この σ の設定は微妙なものではなく，実際，我々の実験では $\sigma \in [0.75, 1.5]$ の範囲ではほぼ同じ結果が得られた．

以上の線形近似の下，式 (5) を最小化して距離 $d_{\mathrm{TD}}(\boldsymbol{P}, \boldsymbol{E})$ を求めるための計算量は，$O(MI^2 + M^2)$ となる．後に **4.** で示されるように，使用固有変形数 M は実際には非常に小さな数 (~ 3) で済む．従って，計算量はおよそ $O(MI^2)$ すなわち単純重ね合わせ法の M 倍程度となり，従来の弾性マッチング法に比べて同程度もしくは相当少い計算量で済む．

4. 実　　験

4.1 実 験 試 料

変形を伴う画像パターンの例として手書き文字パターンを用い，本手法の評価実験を行った．実験では，標準手書き文字データベース ETL6 に含まれる英語大文字 26 カテゴリ各 1100 サンプルを用いた．まず，すべてのサンプルについて，文字大きさを 16×16 に

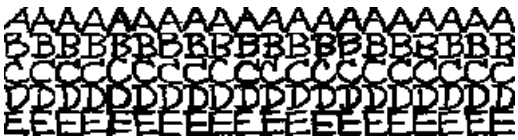

図 6　ETL6 の文字データの例．大きさは 16×16 に線形正規化されている．
Fig. 6　Character samples from ETL6.

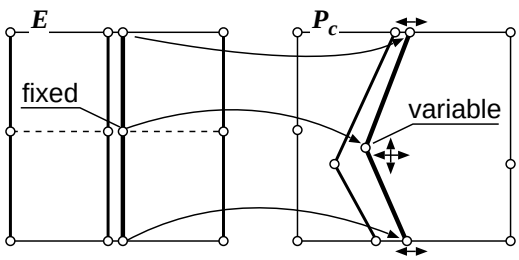

図 7　区分線形 2 次元ワープ [10]．
Fig. 7　Piecewise linear two-dimensional warping.

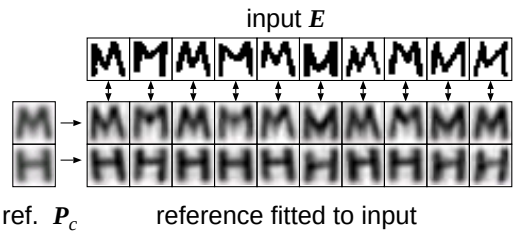

図 8　“M”-“M” および “H”-“M” のマッチング結果．
Fig. 8　Matching results between “M” and “M”, and “H” and “M”.

線形正規化した後，周囲に 2 画素分のマージンを付け，結果的に $I = 20$ とした．図 6 に線形正規化後のサンプルの例を示す．次に，各カテゴリ最初の 100 サンプルを単純平均して標準パターン $\boldsymbol{P}_c$ を作成した．次の 500 サンプルは，固有変形推定用の学習パターン $\boldsymbol{T}_{c,n}$ として用いた．残る 500 サンプル (全カテゴリで 13000 サンプル) は，4. 4–4. 6 節での認識実験におけるテストサンプル $\boldsymbol{E}$ として用いた．

実験では，画素特徴 ($P_c(x, y)$ および $E(x, y)$) として，輝度特徴と 4 方向特徴 [9] からなる 5 次元ベクトルを用いた．なお，前節までの議論をこの高次元特徴ベクトルに対応させるのは容易である．

4. 2　固有変形の推定

2. 3 節の要求 (i)-(iii) を満たす従来の弾性マッチング法として，本実験では区分線形 2 次元ワープ法 [10] を採用した．これはノンパラメトリックな弾性マッチング法の一種であり，画像パターンの各列中央の画素と上下端の画素の変位を制御変数として 2 画像間の最大一致を図る．制御変数以外の画素の変位は線形補間で与えられる (図 7)．この手法を用いた場合の変位ベクトルの実効次元数は，その制御変数の少なさから $I = 20$ の時で $74 (\ll 2I^2 = 800)$ となり，要求 (iii) も満たしている．

実験では 2. 3 節で述べた手順に従い，まず区分線形 2 次元ワープにより前節で述べた $\boldsymbol{T}_{c,n}$ と $\boldsymbol{P}_c$ の弾性マッチングを行い，各 $\boldsymbol{T}_{c,n}$ の変位ベクトルを自動採集した．次にこうして集めた 500 個の変位ベクトルに対して主成分分析を施し，各カテゴリの固有変形を得た．なお，図 2 や図 3 で用いた固有変形 $\boldsymbol{u}_{c,m}$ はこの手順で実際に求まったものである． これらの図からもわかるように，以上の手順により自動的に適切な固有変形を推定できる．なお，文献 [1], [2] には全 26 カテゴリの第 1 ~ 3 固有変形が図示されている．

4. 3　変形補償能力のカテゴリ依存性の確認

本手法の効果を確認するために，図 1 で例に挙げた “H” と “M” に関するマッチングを実際に行った．具体的には，入力パターンとしてカテゴリ “M” の 10 サンプルを用い，これらに対しカテゴリ “M” と “H” の変形モデル ($M = 3$) がどのように追従すなわち変形補償するかを観察した．結果を図 8 に示す．カテゴリ “M” の変形モデルは入力パターンの多様な変形を適切に補償していることがわかる．また，それに比べてカテゴリ “H” の”M” に対する変形補償能力は鈍いことがわかる．特に図 1 で例示した，“M” の垂直ストロークの「ハ」の字状の変形についてはほとんど追従できていないことがわかる．以上より，本手法には，カテゴリ内に起る変形は適切に補償し，一方でカテゴリに生じない変形については無視するような所期の能力があると言える．

4. 4　認 識 結 果

図 9 左に，使用固有変形数 M を変えながら測定した，本手法の認識率と計算時間の関係を示す．なお本実験においては，比較等を簡単にするため，すべてのカテゴリで固有変形数 M を同一とした．最高認識率は 99.21%となった．これは単純重ね合わせ (rigid matching, 本手法において $\alpha = 0$ もしくは $M = 0$ の場合に相当) に比べ 1.12% の改善となっていた．この改善 (144 サンプル分) の内訳を見ると，本手法により初めて正しく認識されたものが 152 サンプル，本手法により誤認識に転じたものが 8 サンプルであった．

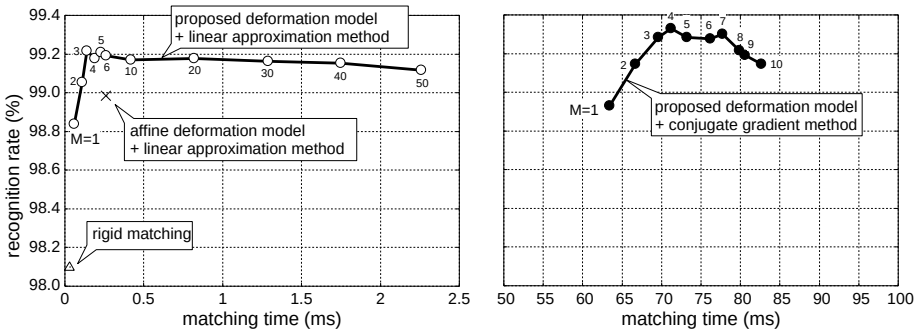

図 9　認識率と計算時間の関係．
Fig. 9　Recognition rates and computation times.

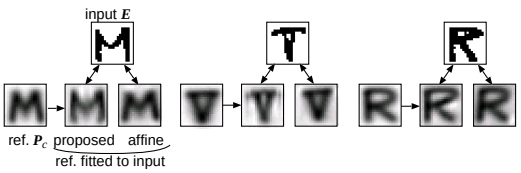

図 10　変形モデルのカテゴリ依存/非依存性によるマッチング結果の差異．
Fig. 10　Comparison between category-dependent and category-independent deformation models.

使用固有変形数 M と認識率の関係を見ると，認識率は M の増加に従って急激に向上し，$M=3$ で最高となっていることがわかる．このことは，式 (4) のカテゴリ依存変形モデルを用いることで，各カテゴリに発生する変形を非常に少ないパラメータ数で効率的に表現できることを示している． さらに，M が非常に小さいことで，計算が効率的に行えることがわかる．一方，M をそれ以上大きくしても，認識率は向上しないことがわかる．これは寄与の小さな高次の固有変形を用いることに意味がないためと考えられる．

4.5　カテゴリ非依存の変形モデルに基づく弾性マッチングとの比較

提案したカテゴリ依存変形モデルのカテゴリ非依存変形モデルに対する優位性を確認するための実験を行った．具体的には，式 (4) の代りに，通常用いられることの多いアフィン変形モデル

$$s(\alpha, P_c(x,y)) = P_c(x + (\alpha_1 x + \alpha_2 y + \alpha_3), \\ y + (\alpha_4 x + \alpha_5 y + \alpha_6)) \quad (9)$$

をすべてのカテゴリにおいて用いた弾性マッチング法による認識実験を行った．このアフィン変形モデルを用いた場合も，本手法と同様に接平面近似 (5) に基づく線形近似解法を導出できる．この場合，6 つのカテゴリに依らない接ベクトル ϕ_m が得られることになる．なお，このアフィン変形モデル (9) とその線形近似に基づく認識手法は，Simard らの tangent distance 法 [3] ~ [5] と同じものになる．

図 9 左に示したように，その認識率は 98.98%であり，本手法の 99.21%よりも低い．さらにこれらのパラメータ数 M は，それぞれ 6 と 3 であるから，本手法のほうが半分のパラメータ数で高い認識率を与えたことがわかる．この結果より，カテゴリ依存変形モデル (4) を用いたほうが，カテゴリ非依存変形モデルを用いるよりも，各カテゴリの生じる変形をより的確にかつ効率的に表現できていると言える．なお，アフィン変形モデルに対する本手法の改善率は 0.23%と一見小さいが，これは誤認識数を約 3/4 に低減できたことに相当する．

図 10 に，本手法で正しく認識され，かつアフィン変形モデルに基づく方法で誤認識されたテストサンプル $\boldsymbol{E}$ およびそのマッチング結果を示す．これから，本手法のほうが $\boldsymbol{P}_c$ を $\boldsymbol{E}$ に近づくよう変形できており，変形補償能力が適切であることがわかる．

4.6　解法の影響

以上では目的関数 (1) の最小化法として，3. の線形近似解法を用いてきたが，他の解法の利用も考えられ，結果も異なってくると思われる．そこで共役勾配法 [11] を用いて同様の認識実験を行った．共役勾配法は，反復に基づく最適化法の一種であり，初期解から出発して目的関数の勾配 $(\partial J/\partial\alpha_1, \ldots, \partial J/\partial\alpha_m, \ldots, \partial J/\partial\alpha_M)^T$ を利用しながら，目的関数の極小値を探索する方法である．以下では，反復の停止条件として目的関数値の更新量がしきい値 (10^5) 以下になるという基準を採用

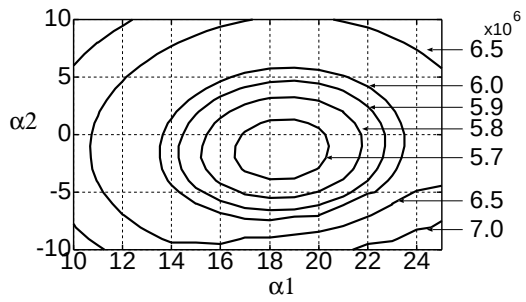

図 11　目的関数値 $J(\alpha)$ の等高線表示 $(M = 2)$
Fig. 11　The value of the objective function $J(\alpha)$.

した．また初期解は $\alpha_1 = \cdots = \alpha_M = 0$ とした．

共役勾配法を用いる利点としては，変形モデル (4) 自体には近似を行わないため，その補償能力を十分に発揮できる点がある．このため，カテゴリ内の変形の補償能力については，線形近似の場合よりも優位と考えられる．一方，線形近似の場合に比べ計算量が多くなるという欠点がある．これは，反復毎に勾配を求めたり，その勾配方向に 1 次元探索を行う際に多数回目的関数値を計算する必要があるためである．前処理により反復回数を減らす方法もあるが [12]，平均反復回数自体はそう多くはないので (以下の実験では $M = 2$ と 15 でそれぞれ平均 3.5 回，4.8 回)，1 桁以上の高速化は困難と思われる．

この共役勾配法を用いた場合の認識実験の結果を図 9 の右側に示す．最高認識率 $(M = 4$ のとき $99.33\%)$ を与えた場合について同図左の線形近似解法と比較すると，共役勾配法を用いた方が若干高い認識率 (16 サンプル分) を与えている．これは変形モデル (4) の持つ変形補償能力を近似なしに活用できたためと考えられる．実際，共役勾配法のみで正しく認識された 37 個のテストサンプルを見ると大きく変形しているものが多く，共役勾配法に比べて線形近似解法によるマッチングでは十分に変形が補償されていない (すなわち $\boldsymbol{E}$ と $t(\alpha, \boldsymbol{P}_c)$ に差異が残っている) 場合が多かった．

ただし，線形近似解法のみで正しく認識されたテストサンプルも 21(=37-16) 個あり，現状では共役勾配法が認識率の点で圧倒的に優れているとは結論できない．これら 21 サンプルの共役勾配法によるマッチング結果には，過変形 (例えば "T" が "Y" に類似するように過変形が 8 例) が生じている場合が散見された．これは変形補償能力が悪用された結果であるが，本モデル (4) を用いてカテゴリ内の変形方向を規定しても，少数ながら依然として過変形が発生しうることを意味している．

図 9 から，共役勾配法と線形近似解法の計算時間には，同じ固有変形数 M で比べると 2 桁以上の差があることがわかる．前述したように， 線形近似解法では $O(MI^2)$ という少ない計算量で解析的に解が求まるが，共役勾配法では反復ごとに目的関数および勾配を多数回求める必要があり，その差異がそのままこの結果に反映されているものと考えられる．

共役勾配法などの反復解法に共通する問題点としては，目的関数が非凸な場合に生じる局所最小解が考えられる．しかし本問題の目的関数 $J(\alpha)$ を観察したところ，ほぼ凸関数となる場合が多く，従って局所解はあまり問題にならなかったと考えられる．図 11 は，カテゴリ "A" の 2 画像のマッチングにおける目的関数値を $M = 2$ の場合について等高線表示したものであるが，ほとんど単峰であることがわかる．実際にこの図の目的関数に対し共役勾配法を適用した結果，ほぼ最適解と思われる $(\alpha_1, \alpha_2) = (18.5, -1.27)$ が得られている．また，誤認識の場合のマッチング結果を観察したところ，局所最適解に陥ってしまったようなものは見当たらなかった．よって，目的関数 (1) の最小化問題における共役勾配法の解探索能力には特に問題がなかったものと考えられる．

このように，共役勾配法により目的関数 (1) の最小解がほぼ常に得られていたとなれば，それによって得られた実験結果は，本変形モデル (4) による認識性能の限界を表していることになる．前述のように，本実験結果から過変形が少数ながら誤認識の要因となっていることが明らかになっており，それがこの限界の現れであるとすれば， 今後は 2.2 節後半での指摘と関連して重み α の範囲制約を検討する必要があると思われる．

なお，非線形最適化問題の反復解法としては，他にも逐次代入法などがある．Wakahara ら [13] は，アフィン変形モデルに基づく弾性マッチングを同様の非線形問題として定式化し，それを逐次代入法によって解く手法を提案している．

5. 他の手法との関連

5.1 固有変形を事後評価において利用する手法 [1], [2] との関連

弾性マッチングを用いた文字認識における，固有変形の他の利用形態として，事後評価における利用が提

案されている [1], [2] . 具体的には，まず従来法の弾性マッチングを行った後，その際の変位を固有変形と比較照合し，逸脱が見られた場合を過変形としてペナルティを与えるというものである . 文献 [1], [2] では，この間接的な利用形態によっても，十分高い認識率が得られると報告されている .

しかしながら，計算効率の面では，本手法のほうが有利と考えられる . 事後評価を行う方式で用いられた従来の弾性マッチングでは，すべてのカテゴリに生じる変形を補償できる程度の能力を確保しておかなければならない . よって，1 に示した “H” の両縦ストロークを傾けるような不要な変形も対象としながら，計算コストをかけて無駄な探索を一旦は行うことになる . これに対し本手法は，変形モデルに直接固有変形を組み込んだ新しい弾性マッチング法であり，結果的に無駄な変形を根本的を排除し，効率的/高速なマッチングを可能としている .

また，事後評価型の利用形態においては，マッチングコストと事後評価値をの比重を定めるパラメータの値が重要であり，それにより性能が大きく左右される . 一方，本手法にはそうしたパラメータは不要である .

5. 2 部分空間法 **[14]** との関連

線形近似された変形モデル (6) と部分空間法 [14] は，任意の変形画像を複数の画像 (基底画像) の線形結合で近似表現するという点で類似している . 一方で，以下のような違いがある . 第一に，本手法は弾性マッチング法であり，求められた α から 2 パターン間の画素対応関係が判明するという意味で構造解析的パターン手法の側面も持っている . 一方，部分空間法にはそういう意味合いはない . 第二に，モデル (6) の基底画像 ϕ_m はパターン空間内において互いに直交していないが，部分空間法のそれは直交している . もちろん $\{\phi_m\}$ の直交化は可能であるが，もともと ϕ_m は変位空間の基底 $\boldsymbol{u}_{c,m}$ をパターン空間内で表現したものであり，本質的に部分空間法で用いる基底とは異なっている .

このように両手法には差異があるが，相対する立場にあるのではなく，相補的な手法となり得るものと考えている . 例えば，本手法の基底画像 ϕ_m と部分空間法の基底画像が，パターンの分布をで表現する上で異なった役割を相補的に担う可能性もあり，両者の組み合わせを含めて，今後比較検討が必要と考えている .

5. 3 カテゴリ毎の変形特性を組み込んだ他の弾性マッチング法

カテゴリ毎の変形方向を学習し，モデルに組み込む試みは，本手法以外にも幾つか検討されている . 文献 [15] ~ [17] で提案されている文字認識手法では，スプライン曲線モデルを用い，その節点の変位の確率分布をカテゴリ毎に学習している . 本手法で用いる固有変形も変位の分布を表現するものであり，従って両手法には共通点があると言える . ただし，弾性マッチング法としての性質には大きな違いがある . 具体的には，本手法では最適整合を得るために固有変形の重みパラメータ α を制御するのに対し，文献 [15] ~ [17] の手法では全節点の変位を直接制御する . すなわち，前者はパラメトリックな弾性マッチング法 (2. 1 節参照) であるのに対し，後者はノンパラメトリックである . また，本手法が 2 次元-2 次元弾性マッチングであり濃淡画像のマッチングに直接適用できるのに対し，この従来法は 1 次元-2 次元弾性マッチング法であるという点も異なる .

文献 [18] では，パターンを構成する各線分の方向や長さなどの統計量を整合度の評価値に組み込んだ 1 次元-2 次元弾性マッチング法が提案されており，文字認識に利用されている . この手法も文献 [15] ~ [17] の手法と同様にノンパラメトリックである .

弾性マッチング法を HMM の枠組みで実現する試みも提案されている [19], [20] . この場合，各カテゴリ毎に用意された HMM の状態遷移確率や出力確率として，変形特性が表現されることになる . ただし，マルコフ性の制約により学習の結果得られるのは隣接画素間の関係だけであり，大局的な変形の方向性を直接組み込むことはできない .

6. む す び

変形に対して頑強な認識系の実現を目的として，カテゴリ依存変形モデルに基づく弾性マッチング法を提案した . この変形モデルでは，あるカテゴリの任意の変形をそのカテゴリに固有な幾つかの変形方向の線形結合で表現する . 本モデルを用いることで，そのカテゴリに生じる変形だけを効率的に補償でき，その結果，認識率の向上や計算量の低減という 2 つの効果が得られると期待される . 非線形最適化問題として定式化される本手法に対し，本論文では線形近似に基づく解法を提案した . 加えて共役勾配法による解法についても検討した . 手書き文字を対象として実際に認識実験を行い，所期の効果が得られることを確認した .

謝辞　実験では，電子技術総合研究所 (現 産業技術総合研究所) の文字データベース ETL6 を利用させて

頂いた．ここに感謝する．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) によった．

文　献

[1] 内田誠一, 迫江博昭, "固有変形の利用による手書き文字認識の高精度化," 画像の認識・理解シンポジウム (MIRU2002), I-391–396, 2002.

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

[3] P. Simard, Y. Le Cun, J. Denker, and B. Victorri, "An efficient algorithm for learning invariances in adaptive classifier," *Proc. 11th ICPR*, vol. 2, pp. 651–655, 1992.

[4] P. Simard, Y. Le Cun, and J. Denker, "Efficient pattern recognition using a new transformation distance," *Advances in Neural Information Processing Systems*, vol. 5, pp. 50–58, 1993.

[5] P. Simard, Y. Le Cun, and J. Denker, "Memory-based character recognition using a transformation invariant metric," *Proc. 12th ICPR*, vol. 2, pp. 262–267, 1994.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Comput. Vis. Image Und.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[7] D. Shi, S. R. Gunn, and R. I. Damper, "Handwritten Chinese radical recognition using nonlinear active shape models," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 25, no. 2, pp. 277–280, Feb. 2003.

[8] 池田正幸, 田中英彦, 元岡 達, "手書き文字認識における投影距離法,"　情処学論，vol. 24, no. 1, pp. 106–112, Jan. 1983.

[9] 斉藤泰一, 山田博三, 山本和彦, "手書き漢字の方向パターン・マッチング法による解析," 信学論 (D), vol. J65-D, no. 5, pp. 550–557, May 1982.

[10] 内田誠一, 迫江博昭, "区分線形 2 次元ワープ法の検討," 信学論 (D-II), vol. J83-D-II, no. 12, pp. 1643–1649, Dec. 2001.

[11] 坂和正敏, 非線形システムの最適化, 森北出版, 1986.

[12] 茨木俊秀, 福島雅夫, 最適化の手法, 共立出版, 1993.

[13] T. Wakahara, Y. Kimura, and A. Tomono, "Affine-invariant recognition of gray-scale characters using global affine transformation correlation," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 23, no. 4, pp. 384–395, April 2001.

[14] 石井健一郎, 上田修功, 前田英作, 村瀬 洋, わかりやすいパターン認識, オーム社, 1998.

[15] M. Revow, C. K. I. Williams, and G. E. Hinton, "Using generative models for handwritten digit recognition," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 18, no. 6, pp. 592–606, June 1996.

[16] K. -W. Cheung, D. -Y. Yeung, and R. T. Chin, "A Bayesian framework for deformable pattern recognition with application to handwritten character recognition," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 20, no. 12, pp. 1382–1388, Dec. 1998.

[17] 加藤 毅, 大町真一郎, 阿曽弘具, "伸縮変形モデルを用いた手書き文字認識," 信学論 (D-II), vol. J83-D-II, no. 12, pp. 2578–2586, Dec. 2000.

[18] 中村香織, 小林 隆, 杉山岳弘, 阿部圭一, "輪郭線の方向を用いたマッチングによる手書き数字認識," 信学技報, PRMU99-220, Feb. 2000.

[19] E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," *Proc. ICASSP*, vol. 3, pp. 149–152, 1992.

[20] S. Kuo and O. Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-d hidden Markov models," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 16, no. 8, pp. 842–848, Aug. 1994.

（平成 x 年 xx 月 xx 日受付）

内田　誠一　（正員）

平 2 九大・工・電子卒．平 4 同大大学院修士課程（情報）了．セコム（株）勤務を経て，現在，同大システム情報科学研究院知能システム学部門助教授．博士（工学）．画像パターンの解析・認識，及び音声の認識・合成に関する研究に従事．IEEE，日本音響学会，情報処理学会，映像情報メディア学会各会員．

迫江　博昭　（正員）

昭 41 九工大・制御卒．昭 43 九大大学院修士課程（電子）了．同年日本電気株式会社入社．平成元年，九大に移り，現在システム情報科学研究院知能システム学部門教授．工博．音声認識及び画像パターン解析の研究に従事．IEEE ASSP Senior Award，本会業績賞，本会米沢ファウンダーズメダル受賞記念特別賞及び論文賞（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授賞．IEEE，日本音響学会，情報処理学会，映像情報メディア学会各会員．本会フェロー．