カテゴリ毎の変形特性を組み込んだ弾性マッチングと その最適化アルゴリズムに関する検討
Category-Dependent Elastic Matching and Its Optimization Algorithms

内田 誠一 †　　**迫江 博昭** †
Seiichi Uchida　　Hiroaki Sakoe

1. まえがき

弾性マッチングは画像パターン認識の要素技術の１つであり，標準パターンを入力パターンに近づくよう２次元的に最適変形させる方法である．その結果，入力パターンの変形は補償され，２パターン間の変形不変距離が求まる．従って，この距離に基づいて最短距離識別を行えば，変形に頑強な認識系を実現できる．

筆者らはこれまで，各カテゴリ固有の変形方向 (以下，固有変形) を組み込んだ弾性マッチング法 (以下，本手法) について検討してきた [1]．これは弾性マッチングにより補償可能な変形の範囲をカテゴリ毎に規定しようとする試みであり，従来法において問題となっていた変形不足や過変形 (合わせ過ぎ) を低減できると考えられる．

本稿では，一種の非線形最適化問題となる本手法に対し，対照的な２通りの解法を検討する．具体的には，線形近似に基づく解法，および共役勾配法に基づく反復解法を挙げ，手書き文字画像の認識実験を通して，これらの２つの解法の性能比較を行う．

2. 固有変形を組み込んだ弾性マッチング

カテゴリ c の標準パターンを画像 $\boldsymbol{P}_c = \{P_c(x,y)\}$ と表す．このとき，カテゴリ依存変形モデル $s(\alpha, \boldsymbol{P}_c)$ を次のように定義する．

$$s(\alpha, P_c(x,y)) = P_c\Big(x + \sum_{m=1}^{M} \alpha_m X_{c,m}(x,y), \; y + \sum_{m=1}^{M} \alpha_m Y_{c,m}(x,y)\Big) \quad (1)$$

ここで $(X_{c,m}(x,y), Y_{c,m}(x,y))$ は，カテゴリ c の第 m 固有変形による，画素 (x,y) の変位である．ここで固有変形とは，統計的に推定された，カテゴリ固有の変形方向を表現するものである．例えば文字画像「A」ならば，全体的な傾き変形や，水平ストロークの上下動などがそれぞれ固有変形となりうる．このように本手法では，各カテゴリに生じる任意の変形を M 個の固有変形の重み ($\alpha_m \in \alpha$) 付き線形和で表現する．なお固有変形は，何らかの弾性マッチング法により収集した２次元変形に対して主成分分析を施すことで自動推定できる [2,3]．

変形モデル (1) を用いた場合，入力パターン $\boldsymbol{E}$ と標準パターン $\boldsymbol{P}_c$ 間の弾性マッチングは次の目的関数 $J(\alpha)$ の α に関する最小化問題として定式化できる (図 1)．

$$J(\alpha) = \|s(\alpha, \boldsymbol{P}_c) - \boldsymbol{E}\| \quad (2)$$

目的関数の最小値 $\min_\alpha J(\alpha)$ は，変形モデル (1) に従って $\boldsymbol{P}_c$ を最も $\boldsymbol{E}$ に近づけた後の距離なので，一種の変形不変距離である．従ってこの距離に基づいた最短距離識別は，通常の単純重ね合わせ距離に比べ変形に頑強であると考えられる．また，本手法では補償対象となる変形の範囲をカテゴリ毎の固有変形によって規定しているため，全カテゴリで一定の変形特性を仮定した従来の手法に比べ，異なるカテゴリ間での過変形が起きにくいと期待される．

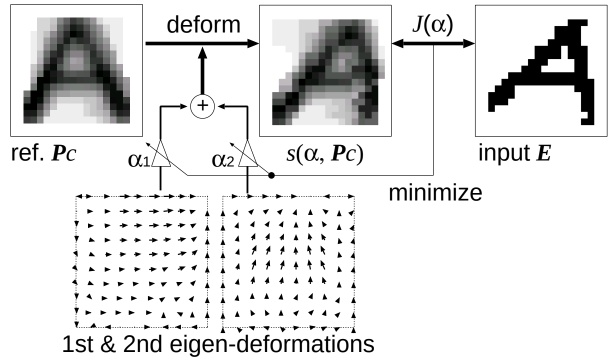

図 1: 固有変形を組み込んだ弾性マッチング ($M = 2$)

3. 最適化アルゴリズム

目的関数 (2) の最小化問題は，制御変数 α が非線形関数 $\boldsymbol{P}_c$ に含まれた，一種の非線形最適化問題となっている．以下ではその２種類の解法について述べる．

3.1 線形近似に基づく解法 [1]

目的関数 (2) の最小化問題は，変形モデル $s(\alpha, \boldsymbol{P}_c)$ の代りに，それを $\alpha = 0$ 近傍でテーラー展開して得られる

$$t(\alpha, \boldsymbol{P}_c) = \boldsymbol{P}_c + \sum_{m=1}^{M} \alpha_m \phi_{c,m} \quad (3)$$

を用いることで線形化できる．ここで，

$$\phi_{c,m}(x,y) = \frac{\partial P_c(x,y)}{\partial x} X_{c,m}(x,y) + \frac{\partial P_c(x,y)}{\partial y} Y_{c,m}(x,y) \quad (4)$$

である．幾何学的には，パターン空間内において M 次元多様体 $s(\alpha, \boldsymbol{P}_c)$ を $\alpha = 0$ における接平面 $t(\alpha, \boldsymbol{P}_c)$ で近似したことになる．この近似 (3) の結果，制御変数 α は $\boldsymbol{P}_c$ の外に出ることになり，式 (2) は最小２乗問題として閉じた形で解くことができる．なお同様の線形近似の考え方は，アフィン変形モデルに基づく弾性マッチング法の一種である tangent distance 法 [4] にも見られる．

本解法の第一の利点は，上述のように閉じた形で解が求まるので，収束性を考慮しなくてよい点にある．第二は，この解を求めるための計算量が $O(M^2)$ 程度と，単純

†九州大学, Kyushu University.

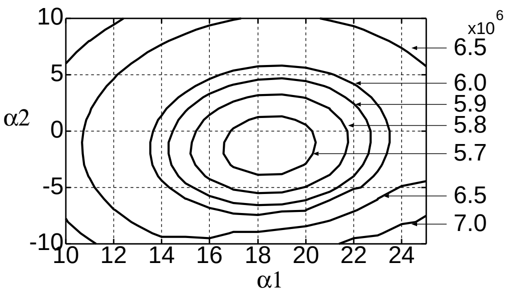

図 2: 目的関数値 $J(\alpha)$ の等高線表示 $(M = 2)$

重ね合わせの高々定数倍程度で済む点である．一方，欠点としては，式 (3) の近似が有効なのがあくまで $\alpha = 0$ の近傍，すなわち変形の小さな領域であり，大きな変形には対応できないことが挙げられる．

3.2　共役勾配法による解法

近似なしに非線形最適化問題 (2) を直接解く方法として，共役勾配法 [5] の利用を検討する．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．反復の停止条件，すなわち極小値の判定条件として，本稿では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する．

共役勾配法の利点としては，変形モデル (1) の持つ補償能力を十分に (すなわち $\alpha = 0$ 近傍という制約無しに) 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，欠点としては，反復毎に勾配を求めたり，その勾配方向に 1 次元探索を行う必要があるため，線形近似の場合に比べ計算量が多くなる点である．前処理により反復回数を減らす方法 [5] もあるが，平均反復回数自体はそう多くはないので (次節の実験では $M = 2$ 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．

共役勾配法などの反復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いと考えられる．図 2 は，2 つの手書き文字画像 (図 1 の $\boldsymbol{P}_c$ と $\boldsymbol{E}$) のマッチングにおける目的関数 $(M = 2)$ を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 $(\alpha_1, \alpha_2) = (18.5, -1.27)$ が得られた．

4.　実験

本手法について，以上述べた 2 つの解法の性能を比較するために，ETL6 の手書き英語大文字 26 文字種 (各カテゴリ 500 サンプル) の認識実験を行った．図 3 は使用する固有変形数 M を変えながら測定した認識率と計算時間の変化である．図より，認識時間の点では，線形近似解法の方が共役勾配法に比べ，3 桁程度高速であることがわかる．実装の更なる工夫により，この差は小さくできると思われるが，優劣に変化はないと考えられる．

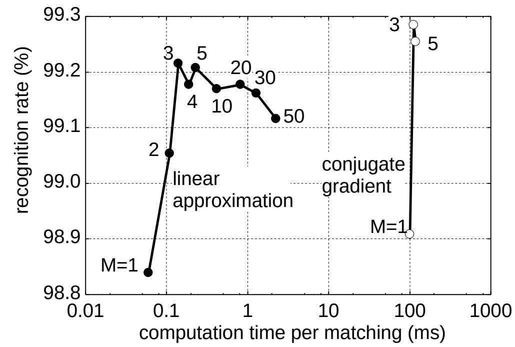

図 3: 各解法における認識率と計算時間の関係

一方，認識率の点では，共役勾配法を用いた方が若干性能が高い．これは変形モデル (1) 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．しかし，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相当数あり，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．これらテスト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” →“Y”) が生じている場合が散見された．本モデル (1) 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おり，今後の検討を要する．

なお，同一のサンプル集合を単純重ね合わせ法で認識した場合，その認識率は 98.1% であった．また，式 (1) の代りに，従来のカテゴリ非依存の変形モデル (affine 変形モデル) を用い，これを 3.1 の線形近似解法で最適化した場合の認識率は，99.0%であった [1]．

5.　むすび

固有変形を組み込んだ弾性マッチング法に対し，線形近似に基づく解法，および反復解法に基づく解法を提案した．また，手書き文字画像を対象とした認識実験により，それらの比較を行った．

謝辞 共役勾配法についてご助言を頂いた京都大学大学院情報学研究科 天野晃先生に感謝致します．本研究の一部は文部科学省科学研究費補助金 (No.14780293) によった．

参考文献 [1] 内田, 迫江, “カテゴリ毎の変形特性を組み込んだ弾性マッチングによる手書き文字認識,” 信学技報, PRMU2002-231, 2003. [2] 内田, 迫江, “固有変形の利用による手書き文字認識の高精度化,” MIRU2002, vol. 1, pp. 391–396, 2002. [3] S. Uchida and H. Sakoe, “Eigen-deformations for elastic matching based handwritten character recognition,” Pattern Recognition, vol. 36, no. 9, pp. 2031–2040, 2003. [4] P. Simard, et al., “An efficient algorithm for learning invariances in adaptive classifiers,” Proc. 11th ICPR, vol. 2, pp. 651–655, 1992. [5] 茨木, 福島, 最適化の手法, 共立出版, 1993.