弾性マッチングを用いた画像パターン認識のための クラスタリング法に関する基礎的検討

A Study of Clustering Method for Elastic Matching Based Image Pattern Recognition

松本 直樹 † 内田 誠一 ‡ 迫江 博昭 ‡
Naoki Matsumoto Seiichi Uchida Hiroaki Sakoe

1. まえがき

画像パターン認識に用いられている弾性マッチングとは，標準パターン画像を非線形に変形させ，入力パターン画像に近づける手法である．弾性マッチングによって得られる 2 画像間の距離は，変形補償されたものであり，よって変形不変量となる．従ってその距離を用いることで，変形に対して頑強な認識を行なうことができる．

標準パターンの設定は，弾性マッチングに基づく認識手法にとって重要な課題である．特に，弾性マッチングが必要とされるような変形の大きなパターンの認識問題では，各カテゴリに複数個の標準パターンが必要となるので，それらの自動的な設定法の開発が実用上の問題となる．本稿では，複数個の標準パターンの設定法という課題に対し，弾性マッチング距離に基づくクラスタリング法の利用を検討する．また，文字画像を対象とした実験により，本手法の有効性を検証する．

本研究に類似した検討としては，オンライン文字認識 [1] および音声認識 [2] における標準パターンの設定法がある．いずれも 1 次元パターンを弾性マッチングによりクラスタリングする手法であり，特に弾性マッチングの持つパターン次元数 (長さ) の正規化能力に主眼を置いている．また，2 次元パターンを対象としたものとしては，Tanget Distance 法を前堤とした標準パターン設定法 [3] がある．しかし，手法の提示のみに留まり，実験的な検討はなされていない．

2. 弾性マッチングによる認識

前述のように，弾性マッチングとは標準パターン $\boldsymbol{R}$ を入力パターン $\boldsymbol{X}$ に近づけるように非線形に変形させる方法である．変形後の標準パターンを $\boldsymbol{R}'$ とすると，$\boldsymbol{X}$ と $\boldsymbol{R}$ の弾性マッチング距離 D_e は，$\boldsymbol{R}'$ と $\boldsymbol{X}$ のユークリッド距離で与えられる (図 1)．認識においては，各入力パターンについて，この距離 D_e で最も近い標準パターンのクラスをその認識結果とする．距離 D_e は変形不変量であることから，変形に頑強な認識が可能となる．

一般に，弾性マッチングは何らかのモデルに従って定式化されており，このモデルによって補償可能な変形の範囲が規定される．この範囲は，$\boldsymbol{R}$ を中心として等方的 (円形) であるとは限らず，図 1 で模式的に示したように，一般には非等方形となっている．よって，$\boldsymbol{R}$ から同じユークリッド距離だけ離れた 2 パターン (図 1 中の $\boldsymbol{X}$ と $\boldsymbol{Y}$) についても，$\boldsymbol{R}$ との弾性マッチング距離は大きく異なる場合がある．

†九州大学大学院システム情報科学府
‡九州大学大学院システム情報科学研究院

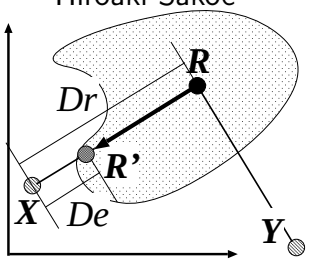

図 1: ユークリッド距離 D_r と弾性マッチング距離 D_e

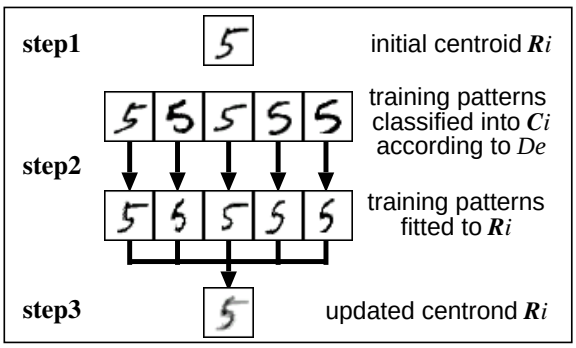

図 2: アルゴリズム

3. 標準パターンの設定アルゴリズム

3.1 クラスタリングによる標準パターンの設定

標準パターンの自動設定法としては，従来よりクラスタリング法が利用されている．このクラスタリング法では，なんらかの距離尺度を最小化することにより標準パターンが選定される．このクラスタリングの距離尺度としては，ユークリッド距離が一般的である．しかしながら，上述のように，認識時に用いられる弾性マッチング距離は非等方的であり，単純にユークリッド距離を用いてクラスタリングを行ったとしても，認識時にこれらの距離の性質の不整合による性能劣化が起こると考えられる．そこで，本稿では，クラスタリング時にも認識時と同じく弾性マッチング距離を用いることを検討する．

3.2 アルゴリズム

前節での考察をもとにして，各カテゴリについて標準パターン $\boldsymbol{R}$ を複数個設定するという課題に対し，本稿では k-means 法に弾性マッチング距離を組み込んだクラスタリング手法を用いることを考える．

本手法の具体的なアルゴリズムは以下の通りである．
step1: 同一クラスの画像パターンによりなる学習パターンセットの中から K 個の初期パターンを適当に選び，それらを初期標準パターン $\boldsymbol{R_1}, \ldots, \boldsymbol{R_K}$ とする．
step2: 全ての学習パターン $\boldsymbol{T_l}(l = 1, \ldots, L)$ をそれぞれに最も近い標準パターンのクラスタ $\boldsymbol{C_i}(i \in \{1, \ldots, K\})$ に属させる．ただし，このときの距離は弾性マッチング

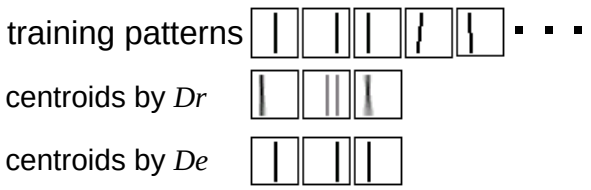

図 3: 線分パターンのクラスタリング結果 ($K = 3$)

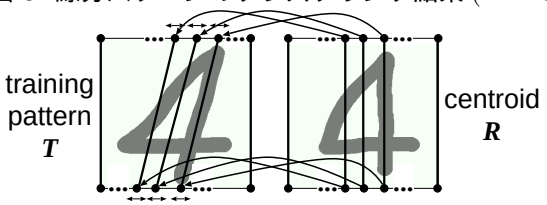

図 4: 実験で用いた弾性マッチング法

距離 D_e とする．またこのとき，$\boldsymbol{T}_l(l = 1, \ldots, L)$ についてそれぞれが属するクラスタの標準パターンに近づくように弾性マッチングを用いて変形したパターンを $\boldsymbol{T}'_l(l = 1, \ldots, L)$ とする．

step3: それぞれのクラスタ $\boldsymbol{C}_i$ について，$\boldsymbol{T}'_l(\in \boldsymbol{C}_i)$ の平均を新たな標準パターン $\boldsymbol{R}$ とする．

step4: **step3** が終った時点で，標準パターンが動いた場合 **step2** へ戻る．標準パターンが動かなかったら終了．

なお本手法では，標準パターンは補正した学習パターンの平均となるので，副次的な効果として平均画像のぼけが少なくなるというメリットがある．実際に，線分パターンを学習パターンとして，従来手法と本手法でそれぞれクラスタリングすることによって求まった標準パターン (図 3) を見ると，本手法にはほとんどぼけがないことがわかる．

4. 実験

4.1 使用した弾性マッチング法

実験では，図 4 に示した弾性マッチング法 [4] を用いた．この方法では，入力画像 $\boldsymbol{T}$ 各列の両端点の座標の最適な対応先を $\boldsymbol{R}$ 上に求めることで 2 画像間の最大一致を図る．

4.2 文字パターンのクラスタリング及び認識結果

標準数字画像データベース MNIST(学習パターン 60000 個，テストパターン 10000 個) を用いて実験を行った．全てのサンプルについて，文字の大きさを 14×14 に線形正規化し，周囲に 1 画素分のマージンを付け，結果的に 16×16 とした．図 5 に $K = 5$ として，従来法 (D_r) および本手法 (D_e) により求まった数字 0 から 4 の標準パターンの例を示す．

このようにして求まった標準パターンを用いて，前述の弾性マッチングを用いた認識実験を行なった．ただし，k-means 法には初期値依存性があるため，初期パターンを 10 通りに変えながらそれぞれクラスタリングおよび認識処理を行なった．その平均認識率を，クラスタ数 K に関するグラフとして図 6 に示す．

図 6 から，同程度の認識率を得ようとする場合，本手法が従来手法に比べより少ない標準パターンで済むということがわかる．これより，弾性マッチングを用いた認識においては，弾性マッチング距離を用いたクラスタリング法が有効であることが確かめられた．

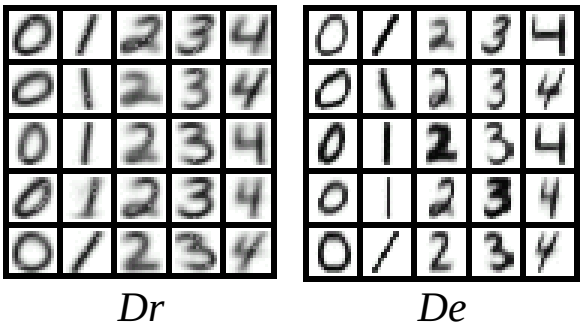

図 5: 数字パターンのクラスタリング結果 ($K = 5$)

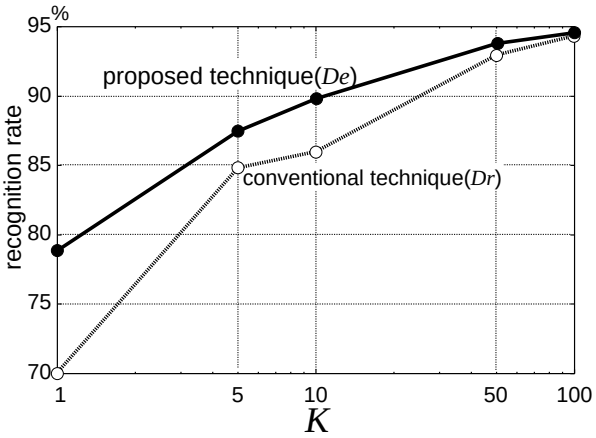

図 6: クラスタリング時の距離尺度による認識率の違い

なお，理論上は K が大きくなるにつれて，従来手法と本手法で求まる代表ベクトルは同じものになっていき，$K = L$ で実際に一致する．K の増加に従って認識性能の差が小さくなっていると考えられる．

5. むすび

弾性マッチングを用いた画像パターン認識のための，標準パターン設定法を検討した．具体的には，弾性マッチングによる距離を用いたクラスタリング手法を提案した．手書き数字画像を用いた認識実験の結果，従来のユークリッド距離に基づくクラスタリング手法に比べ，本手法で得られた標準パターンのほうが少数で高い認識率を与えることが判った．

今後の課題としては，(i) 他のクラスタリング法 (例えば GLVQ [5]) の利用，(ii) 性質の異なる弾性マッチングを使用することの結果への影響の観察，(iii) 弾性マッチングとクラスタリングの他の組み合わせ形態の検討，が挙げられる．

参考文献 [1] 秋山, 石垣, “学習標本による高品質なオンライン文字認識辞書生成の一手法,” 信学技報, PRMU99-235, 2000. [2] 小森, 片桐, “DTW に基づく音声認識のための識別学習,” 信学技報, SP91-10, 1991. [3] T. Hastie, et al., “Learning prototype models for tangent distance,” Advances in Neural Information Processing Systems, vol. 7, pp.999-1006, 1995. [4] 内田他, “区分線形 2 次元ワープによる手書き文字の変形吸収の試み,” 信学技報, PRMU99-228, 2000. [5] 佐藤, 山田, “一般学習ベクトル量子化の提案,” 信学技報, NC95-60, 1995.