

弾性マッチングに基づく画像パターン認識のための
標準パターン設定法に関する検討

松本 直樹[†](学生員) 内田 誠^{††a)}(正員)

迫江 博昭^{††}(正員:フェロー)

A Clustering Algorithm for Elastic Matching-Based Image
Pattern Recognition

Naoki MATSUMOTO[†], *Student Member*,
Seiichi UCHIDA^{††a)}, *Member*, and Hiroaki SAKOE^{††}, *Fellow*

[†]九州大学大学院システム情報科学府, 福岡市

Graduate School of Information Science and Electrical Engineering,
Kyushu University, 6-10-1 Hakozaki, Higashi-ku,
Fukuoka-shi, 812-8581 Japan

^{††}九州大学大学院システム情報科学研究院, 福岡市

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

a) E-mail: uchida@is.kyushu-u.ac.jp

あらまし 弾性マッチングに基づく画像パターン認識のための標準パターン設定法について述べる。本手法はクラスタリング法の一つであるが、従来法がユークリッド距離を基準としているのに対し、本手法では識別時と同じ弾性マッチングによる距離を基準とする。

キーワード クラスタリング, 弾性マッチング, 距離尺度, 画像パターン認識

1. まえがき

画像パターン認識において、標準パターンを入力パターンに非線形に整合させるための手法として、弾性マッチング法の利用が検討されている。理論的には、非線形整合後の2パターン間の距離(以下、弾性マッチング距離と呼ぶ)は、弾性マッチングで補償され得る範囲の幾何学的ひずみに対して不変となる。したがって、弾性マッチング距離を識別に用いることで、変形に対して頑健な画像パターン認識が実現できる。

画像パターン認識一般において重要な問題の一つに標準パターンの設定がある。この設定に際しては、計算量の増加を避けると同時に十分な認識精度を保つために、なるべく少数で入力パターン全体をカバーすることが要求される。このことは弾性マッチングに基づく認識手法にも当てはまる。一つの標準パターンを中心とした変形補償範囲は有限であるため、したがって1カテゴリ当り複数の標準パターンの設定が必要となる。

こうした要求にこたえながら適切な標準パターンを自動的に設定する方法として、クラスタリング法が従来より利用されている。ここでいうクラスタリング法とは、あるカテゴリの学習パターンの分布を何らかの

基準で最適に代表するパターンを選定する方法であり、パターン認識ではこれら代表パターンを標準パターンとして用いることになる。

従来のクラスタリング法は、弾性マッチングに基づく認識手法のための標準パターン設定法としては適切ではない。これは、標準パターン設定時と識別時の距離尺度に不一致が生じるためである。具体的には、従来のクラスタリング法で求めた標準パターンは、ユークリッド距離尺度のもとで最適化されたものであり、したがってユークリッド距離尺度に基づく識別関数を用いることを想定しているためである。

本論文では、ユークリッド距離の代わりに弾性マッチング距離を用いたクラスタリング法により、標準パターン設定時と識別時で距離尺度を一致させることを提案する。また、それが有効であることを、手書き文字画像を対象とした認識実験を通して示す。本論文の主たる目的はこの距離尺度の一致の重要性を確認することにあるので、その点の見通しよくするためにクラスタリング法も弾性マッチング法も非常に単純なものを採用することとする。なお、より識別問題に有効とされる GLVQ [1] などのクラスタリング法に対しても、本論文の考え方は有効であろうと考えている。

弾性マッチングに基づく従来のパターン認識手法のほとんどは、標準パターン設定問題を重要視していなかった。例えば、Jain らの試み [2] においては、計算量を度外視し、すべての学習パターンをそのまま標準パターンとして用いている。少数ながら標準パターン設定に注目した試みも存在するが [3]、そうした試みも距離の不一致を問題視し、それを回避することを主たる目的としておらず、したがってその問題について十分な理論的・実験的検討はなされていない。

2. 弾性マッチングによる画像パターン認識

提案する標準パターン設定法の紹介に先立ち、本章では弾性マッチングによる画像パターン認識の概要について述べる。画像 $X = \{x(i, j) \mid i, j = 1, 2, \dots, N\}$ を未知の入力パターン、 $R_k = \{r_k(u, v) \mid u, v = 1, 2, \dots, N\}$ をある一つのクラスでの k 番目の標準パターンとする。ここで、画像サイズは $N \times N$ とする。このとき弾性マッチング距離 $D_{EM}(X, R_k)$ は、次の制約付き画素対応最適化問題を解くことで得られる。

$$D_{EM}(X, R_k) = \min_{\{(u_{i,j}, v_{i,j})\} \in \mathcal{M}} \sum_{i,j} \|x(i, j) - r_k(u_{i,j}, v_{i,j})\| \quad (1)$$



図 1 入力パターン X , 標準パターン R_k , 及び弾性マッチングにより X に整合された R_k
 Fig. 1 Input X , prototype R_k , and the prototype fitted to the input by EM.

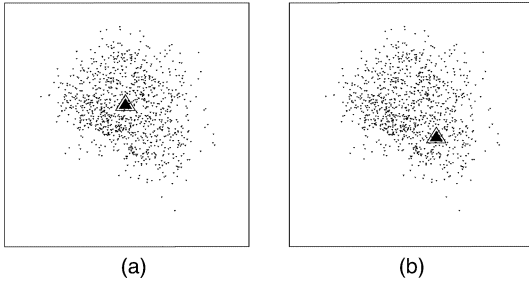


図 2 ユークリッド距離 (a) 及び弾性マッチング距離 (b) のもとでの手書き文字パターン集合の代表パターン (いずれも 2 次元部分空間に射影して表示)
 Fig. 2 The centroids of handwritten character patterns under (a) Euclidean distance and (b) EM distance. Note that all patterns are displayed in two-dimensional subspace.

ここで \mathcal{M} は変形モデルであり, $(u_{i,j}, v_{i,j})$ はこのモデルの制約のもとで X 上の座標 (i, j) と対応する R_k 上の画素を表す. 理論的には, 弾性マッチング距離 $D_{EM}(X, R_k)$ は \mathcal{M} によって許される範囲の幾何学変形に対して不変である. そのため, 画像パターン認識においては弾性マッチング距離を距離尺度として用いることにより, ユークリッド距離

$$D(X, R_k) = \sum_{i,j} \|x(i, j) - r_k(i, j)\| \quad (2)$$

を用いる場合に比べ高い認識性能が期待できる. 図 1 に弾性マッチングの例を示す.

3. 弾性マッチング距離の非等方性

ユークリッド距離 $D(X, R_k)$ は等方的であるため, R_k からの等距離なパターンの集合 (等距離面) は画像空間において超球をなす. 一方, 弾性マッチング距離は非等方的となり, 等距離面はモデル \mathcal{M} に依存した不規則な曲面をなす.

図 2 はこの弾性マッチング距離の非等方性を実験的に示したものである. 図中の点は $256 (= 16 \times 16)$ 次元で表される手書き文字パターンを 2 次元部分空間に射影して表示したものである. ここでの二つの軸はこれら文字パターンの主成分分析により求めた第 1, 2

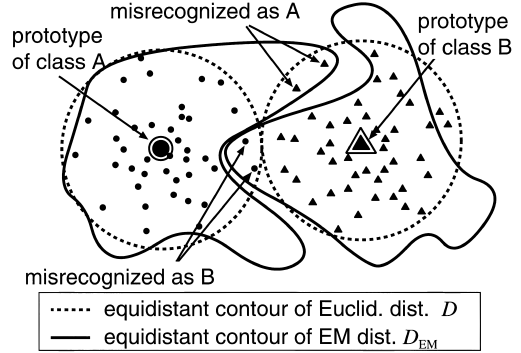


図 3 標準パターン設定時と識別時で距離尺度が異なった場合に発生する誤認識
 Fig. 3 Equidistant contours showing the reason of misrecognition when prototypes are set under Euclidean distance and discrimination is performed under EM distance.

主軸である. 三角形が文字パターンの分布の重心, すなわち文字パターンとの距離の合計が最小となる点である. ユークリッド距離によって求めた重心は文字パターン分布の中心位置に存在する (図 2 (a)). 一方, 弾性マッチング距離 (6.1 で述べるモデルを使用) によって求めた重心は分布の中心からは外れた位置に存在しており (図 2 (b)), 距離尺度の非等方性を示している.

4. 従来のクラスタリング法の問題点

標準パターンを自動的に設定する方法としてはクラスタリング法が従来から用いられている. ここで $T = \{T_l | l = 1, \dots, L\}$ をある一つのクラスの学習パターン集合, C_k をそのクラスの k 番目のクラスタ ($k = 1, \dots, K$) とする. 各学習パターンは必ずいずれか一つのクラスタに属するものとする. クラスタリングの目的は, クラスタ $\{C_k\}$, 及びその代表パターンとしての標準パターン $\{R_k\}$ をある基準に従い最適化することである.

従来のクラスタリング法では距離尺度としてユークリッド距離が用いられている. 例えば, k -means 法は以下に示す目的関数 J を最小化基準として, $\{R_k\}$ と $\{C_k\}$ を求める手法である [4].

$$J = \sum_k \sum_{T_l \in C_k} D(T_l, R_k) \quad (3)$$

この場合, 標準パターン R_k は図 2 (a) のように各クラスタの中心付近に設定される. GLVQ や他の複雑なクラスタリング法においても, 何らかの形でユークリッド距離が用いられていることが多い.

これらの従来法は、弾性マッチングに基づく認識用の標準パターン設定法としては適切ではない。これは前章で述べた、ユークリッド距離と弾性マッチング距離の違いのため、図3に示すように誤認識が生じてしまうためである。このように、ユークリッド距離に基づいて最適化された標準パターンはユークリッド距離に基づく識別には適当であるが、弾性マッチング距離に基づく識別には適さないと考えられる。

5. 弾性マッチング距離に基づくクラスタリング法

本論文は弾性マッチング距離に基づくクラスタリング法による標準パターン設定法を提案する。基本とするクラスタリング法には様々なものが考えられるが、本論文では扱いやすさという点から前述の k -means 法を用いることとする。

弾性マッチング距離を用いた標準パターン設定法は、 $\{R_k\}$ と $\{C_k\}$ を制御変数とした以下の目的関数 J_{EM} の最小化問題である。

$$J_{EM} = \sum_k \sum_{T_l \in C_k} D_{EM}(T_l, R_k) \quad (4)$$

この最小化問題を解くことにより求めた標準パターン $\{R_k\}$ は弾性マッチング距離における重心となり、図3に示した誤認識を極力回避できると考えられる。

目的関数 J_{EM} の最小化は J の場合よりも更に複雑である。これは $D_{EM}(T_l, R_k)$ の計算の際、 T_l と R_k 間の画素対応関係も最適化する必要があるためである。そこで本論文では、 $\{R_k\}$ 、 $\{C_k\}$ 、及び画素対応関係を逐次的に最適化する近似アルゴリズムを用いる。このアルゴリズムを図4に示す。具体的には、Step 2における $\{C_k\}$ と画素対応関係の最適化の際には $\{R_k\}$ を固定して計算する (D_{EM} の計算の際に画素対応関

係の最適化も行っている。) 一方、Step 3における標準パターン R_k の最適化及び更新の際には、 $\{C_k\}$ と画素対応関係を固定する。なお、Step 3における R_k の更新手順の妥当性を付録に示す。

6. 実験

6.1 諸元

標準手書き数字画像データベース MNIST (学習パターン 60,000 個, テストパターン 10,000 個) を用いて認識実験を行った。前処理としてすべてのサンプルについて、画像の大きさを 14×14 に線形正規化し、周囲に 1 画素分のマージンを付け、 $N = 16$ とした。

図5に実験で用いた弾性マッチング \mathcal{M} を示す。このモデル \mathcal{M} は局所的なせん断変換を組み合わせた 2 次元非線形変換を表現している。したがって (1) の最適化問題では、各列の上・下端画素の対応関係のみが制御変数となり、それ以外の画素対応は線形補間により与えられる従属変数となる。最適化には DP アルゴリズム [5] が利用できる。なお、本手法の考え方自体は原理的に任意の弾性マッチングについて適用可能である。

6.2 認識結果

本手法と従来法でそれぞれ標準パターンを作成し、認識実験を行った。ここでの従来法とはユークリッド距離を用いた k -means 法である。認識には弾性マッチング距離に基づく最小距離識別法 (1-NN 法) を用いた。ここで、すべての学習パターンを標準パターンとして用いた場合 ($K = L \sim 6,000$) の認識率は 97.90% であった^(注1)。

実験によって得られた認識率を K に関するグラフとして図6に示す。いずれも標準パターン設定時の初期パターンを 10 通りに変えて得られた平均認識率である。このグラフから、本手法により作られた標準パターンの方が従来法のものより高い認識率となること

[Step1: Initialization]

Choose K initial centroids R_1, \dots, R_K from a training pattern set $\{T_l \mid l = 1, \dots, L\}$.

[Step2: Partitioning]

For each training pattern T_l , find its nearest centroid $R_{k'}$, where $k' = \operatorname{argmin}_k D_{EM}(T_l, R_k)$. Then assign T_l to cluster $C_{k'}$.

[Step3: Updating]

Each R_k is updated by replacing $r_k(u, v)$ by the average value of (u, v) 's corresponding pixels on T_l in C_k .

[Step4: Convergence check]

If R_k is changed by Step3 go to Step2; otherwise, terminate the algorithm.

図4 弾性マッチング距離を用いたクラスタリング手法

Fig. 4 The proposed clustering algorithm.

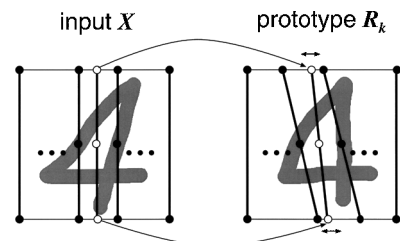


図5 実験で用いた弾性マッチング法
Fig. 5 Elastic matching in the experiment.

(注1): 認識時にユークリッド距離を用いた場合の認識率は 95.70% であった。このことは弾性マッチングの有効性を示している。

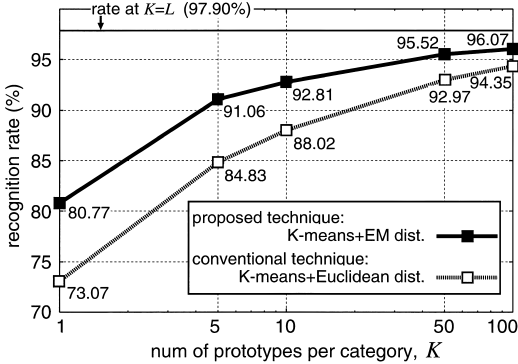


図 6 手書き数字 (MNIST) の認識率
Fig. 6 Recognition results of MNIST.

が分かる(注2)。この結果は (i) 標準パターン設定時と認識時の間での距離尺度は一致している必要があること, (ii) 本手法は弾性マッチングに基づく認識において有効であることを示唆している。

7. むすび

弾性マッチングに基づく画像パターン認識のための標準パターン設定法として、弾性マッチング距離に基づくクラスタリング法を提案した。本手法では標準パターン設定時と認識時の間の距離尺度不一致の問題を回避するために、標準パターン設定時の距離尺度を従来のユークリッド距離の代わりに弾性マッチング距離を用いた。手書き数字の認識実験により、本手法により作成した標準パターンが従来法によるものに比べ高い認識率を示すことを確認した。

今後の課題としては、より認識問題に適したクラスタリング法や、より精度の高い弾性マッチング法を利用した場合にも同じ結論が得られることを確認したいと考えている。

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), 及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] 佐藤 敦, 山田敬嗣, “一般学習ベクトル量子化の提案”; 信学技報, NC95-60, Oct. 1995.

(注2): 図 6 から、K が大きくなるにつれて本手法の優位性が小さくなることが分かる。これは K が大きくなるにつれて、クラスタリング法の意義がうすれ(ほとんどのクラスタがたった一つの学習パターンしか含まなくなる)、得られる認識率は K = L の場合のそれに収束していくためである。

[2] A.K. Jain and D. Zongker, “Representation and recognition of handwritten digits using deformable templates,” IEEE Trans. Pattern Anal. Mach. Intell., vol.19, no.12, pp.1386-1391, Dec. 1997.
[3] T. Hastie, P.Y. Simard, and E. Säckinger, “Learning prototype models for tangent distance,” Advances in Neural Information Processing Systems, vol.7, pp.999-1006, 1995.
[4] R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley Interscience, 1973.
[5] 内田誠一, 迫江博昭, “区分線形 2 次元ワーブ法の検討”; 信学論 (D-II), vol.J83-D-II, no.12, pp.2622-2629, Dec. 2000.

付 録

標準パターンの更新手続き

図 4 中の Step 3 の目的は、クラスタ $\{C_k\}$ と画素対応関係が固定された状態で (4) を最小化する R_k を求めることである。そのような R_k は以下の式を満たす必要がある。

$$\frac{\partial}{\partial r_k} \sum_{T_l \in C_k} \sum_{i,j} \|t_l(i,j) - r_k(u_{i,j}^l, v_{i,j}^l)\| / \partial r_k(u, v) = 0$$

ここで $(u_{i,j}^l, v_{i,j}^l)$ は、 T_l と R_k の間の画素対応関係であり、Step 2 の D_{EM} の導出の過程で既に求まったものである。上式は以下のように変形できる。

$$\sum_{T_l \in C_k} \sum_{i,j} (t_l(i,j) - r_k(u_{i,j}^l, v_{i,j}^l)) \frac{\partial r_k(u_{i,j}^l, v_{i,j}^l)}{\partial r_k(u, v)} = 0$$

更に関係

$$r_k(u_{i,j}^l, v_{i,j}^l) = \sum_{u,v} r_k(u, v) \delta(u_{i,j}^l - u, v_{i,j}^l - v)$$

を用いることで、以下の式を得ることができる。

$$r_k(u, v) = \frac{\sum_{T_l \in C_k} \sum_{i,j} t_l(i,j) \delta(u_{i,j}^l - u, v_{i,j}^l - v)}{\sum_{i,j} \delta(u_{i,j}^l - u, v_{i,j}^l - v)}$$

この式は、更新される標準パターンの画素値 $r_k(u, v)$ はその画素が対応した $T_l \in C_k$ 上の画素値の平均であることを意味する。以上より、図 4 の Step 3 の更新手順が得られた。

(平成 16 年 1 月 14 日受付)