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.

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

松本　直樹† 　内田　誠一†† 　迫江　博昭††

† 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
〒 812–8581 福岡市東区箱崎 6–10–1
E-mail: †{matumoto,uchida,sakoe}@human.is.kyushu-u.ac.jp

あらまし　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提案する．本手法はクラスタリング法の一種であり，パターン空間における学習パターンの分布の局所的な重心を標準パターンとして用いる．従来手法では重心を求める際の距離尺度にはユークリッド距離を用いているのに対し，本手法では新たに距離尺度として認識時と同じ弾性マッチング距離を用いる．従来手法では標準パターン設定時と認識時の間で距離尺度が一致しないのに対し，本手法では距離尺度が一致することになる．実験により，本手法により設定された標準パターンを用いたほうが高い認識率が得られることが確認された．
キーワード　弾性マッチング，クラスタリング，画像パターン認識

A Technique for Setting Standard Patterns for Image Pattern Recognition Based on Elastic Matching

Naoki MATSUMOTO†, Seiichi UCHIDA††, and Hiroaki SAKOE††

† Graduate School of Information Science and Electrical Engineering, Kyushu University
†† Faculty of Information Science and Electrical Engineering, Kyushu University
6–10–1 Hakozaki, Higashi–ku, Fukuoka–shi, 812-8581 Japan
E-mail: †{matumoto,uchida,sakoe}@human.is.kyushu-u.ac.jp

Abstract　A technique for setting standard patterns for image pattern recognition based on elastic matching is investigated. The proposed technique is a kind of clustering techniques, which generally provide standard patterns as the centroids of the distribution of training patterns in pattern space. In conventional clustering techniques, the centroid is defined as the local center of gravity under the metric of the Euclidean distance. Contrary to this, in the proposed technique an elastic matching distance is newly employed as the metric. Thus, the same elastic matching-based metric is consistently used at the standard pattern setting stage and the recognition stage with the proposed technique, whereas different metrics are inconsistently used in those stages with the conventional technique. From experimental results, it was shown that high recognition rates can be attained with the standard patterns provided by the proposed technique because of the consistency of the metric.
Key words　Elastic matching , Clustering , Image pattern recognition

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．

他の認識手法と同じく，弾性マッチングに基づくパターン認識手法においても標準パターン画像の設定は重要な課題である．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

画像が必要となるので，それらの自動的な設定法の開発が実用上の課題となる．また，画像そのものが標準パターンとして用いることによる利点も考慮する必要がある．

各カテゴリ複数個の標準パターンを自動設定する手法としては，従来よりクラスタリングが用いられている．本論文で言うクラスタリングとは，学習パターンのパターン空間における分布の局所的な重心をカテゴリ毎に求める手法である．この重心を標準パターンとすれば，パターン分布の粗密に応じて標準パターンが効率的に配置されることになり，Nearest Neighbor 法などで識別する際に利用できる．

従来より用いられているクラスタリング法における距離尺度はユークリッド距離であり，従って求まる標準パターンはユークリッドの意味での重心である．しかし，認識の際には弾性マッチング距離を用いるため，クラスタリング時 (標準パターン設定時) と認識時で距離尺度の不一致が起こるという問題点がある．例えば，一つのパターンがあるクラスタにユークリッド距離で属するとされても，弾性マッチング距離だと別のクラスタに属するという事態が起こり，この結果認識性能が低下することが考えられる．

本論文では，距離尺度に弾性マッチング距離を用いたクラスタリング法を提案する．この本手法より得られる効果には以下の 3 点が考えられる．

(1) クラスタリング時と認識時で距離尺度が一致することにより認識性能が向上する．

(2) ぼけのない標準パターンの画像が求まる．

(3) 標準パターンとして学習パターンの生成源となるものが得られる．

効果 1 を得ることが，本手法の主な目的である．4.2 節において，標準数字画像パターンデータベース MNIST を用いて，クラスタリング及び認識実験を行い，画像パターン認識における本手法の有効性を確認する．

効果 2 で言うぼけとは，画像パターンをユークリッド距離によって単純平均した際にパターンの変形に起因して生じるものである．認識を行うだけならこのぼけは特に問題にはならないと考えられるが，標準パターンそのものが画像として意味を持つような場合にはこうしたぼけが問題となる．例えば，標準パターンがそのカテゴリの代表であるとして観察の対象となるような場合，本来学習パターンには存在しないぼけたパターンが標準パターンとして求まることは望ましいことではない．一方，本手法では，位置や変形を正規化した後に平均化するため，こうしたぼけが解消できる．4.3 節において，人工的に生成させた線分パターンを用いたクラスタリング実験を行い，この効果を確認する．

効果 3 は，学習パターン集合がある基本となるパターンが変形されることで生成されているとすれば，本手法によりその基本的なパターンそのものが得られる可能性が高いことを意味している．4.4 節では，ある文字パターンを弾性マッチングの変形モデルに従ってランダムに変形させたパターンを作り，それらを学習パターンとしてクラスタリング実験を行い，この効果を確認する．

本研究に類似した検討としては，オンライン文字認識 [1] や音声認識 [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．さらに，オンライン文字認識においてはクラスタリング時に，パターンの位置を平均化するので，画像を平均化することより扱いやすい．また，2 次元パターンを対象としたものとしては，Tangent 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．

本論文では，クラスタリング手法として分布の局所重心を求める手法を採用しているが，クラスタリング法には他にも識別誤り最小化を基準とした手法がある．本手法の基本となる考え方はこのどちらの手法にも適応可能である．高い識別性能を追求する場合には，むしろ後者の手法が主流であるが，本研究では扱いやすさの点から，前者，後者の順に検討を進めることにした．

2. 弾性マッチングによる認識

本節では本手法の説明に先立ち，弾性マッチングを用いた画像パターン認識の概要を示す．

前述のように，弾性マッチングとは標準パターン $\boldsymbol{R} = \{r(u,v) \mid u,v = 1,2,\ldots,N\}$ を入力パターン $\boldsymbol{X} = \{x(i,j) \mid i,j = 1,2,\ldots,N\}$ に近づけるように非線形に変形させる方法である．この弾性マッチングを用いた認識では，$\boldsymbol{X}$，$\boldsymbol{R}$ 間の距離はユークリッド距離

$$D_r(\boldsymbol{X},\boldsymbol{R}) = \sum_{i,j} \| x(i,j) - r(i,j) \| \qquad (1)$$

の代わりに，次の $u(i,j)$，$v(i,j)$ に関する最適化問題を解くことで与えられる弾性マッチング距離 $D_e(\boldsymbol{X},\boldsymbol{R})$ を用いる (図 1)．

$$D_e(\boldsymbol{X},\boldsymbol{R}) = \min_{\{u(i,j),v(i,j)\}} \sum_{i,j} \| x(i,j) - r(u(i,j),v(i,j)) \| \qquad (2)$$

ここで $u(i,j)$，$v(i,j)$ は，入力パターン $\boldsymbol{X}$ 上の座標 (i,j) に対応する標準パターン $\boldsymbol{R}$ 上の座標である．その最適値を $\tilde{u}(i,j)$，$\tilde{v}(i,j)$ とすれば，$\boldsymbol{R'} = \{\, r(\,\tilde{u}(i,j)\,,\,\tilde{v}(i,j)\,)\,\}$ であり，$D_e(\boldsymbol{X},\boldsymbol{R}) = D_r(\boldsymbol{X},\boldsymbol{R'})$ となる．認識においては，各入力パターン $\boldsymbol{X}$ について，この距離 $D_e(\boldsymbol{X},\boldsymbol{R})$ の意味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．

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，パターン空間上で $\boldsymbol{R}$ 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．

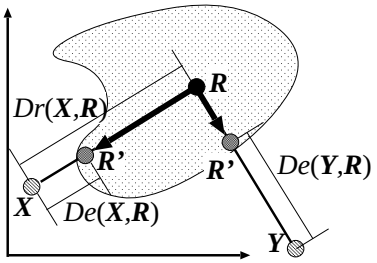


図 1　ユークリッド距離 $D_r(\boldsymbol{X},\boldsymbol{R})$ と弾性マッチング距離 $D_e(\boldsymbol{X},\boldsymbol{R})$

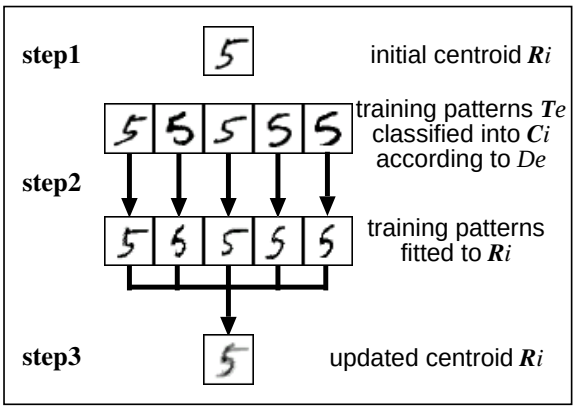


図 2　本手法の標準パターン設定アルゴリズム

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，クラスタリング法が利用できる．従来法では，ユークリッド距離を基準としてパターン分布の重心を求め，それを標準パターンとしていた．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．

そこで本論文では，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，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．ベースとなるクラスタリング法には，任意の手法が使用できるが，本論文ではパターン分布の局所的重心を求める手法としては最も単純な k-means 法を採用した．

本手法の具体的なアルゴリズムは以下の通りである (図 2)．

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 $\{\boldsymbol{T}_l \mid l=1,\ldots,L\}$ 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: 各学習パターン $\boldsymbol{T}_l$ を最も近い標準パターンのクラスタ $\boldsymbol{C}_k(k\in\{1,\ldots,K\})$ に属させる．ただし，このときの距離は弾性マッチング距離 D_e とする．すなわち，

$$\boldsymbol{C}_k=\{\boldsymbol{T}_l \mid \operatorname*{argmin}_m D_e(\boldsymbol{T}_l,\boldsymbol{R}_m)=k\} \tag{3}$$

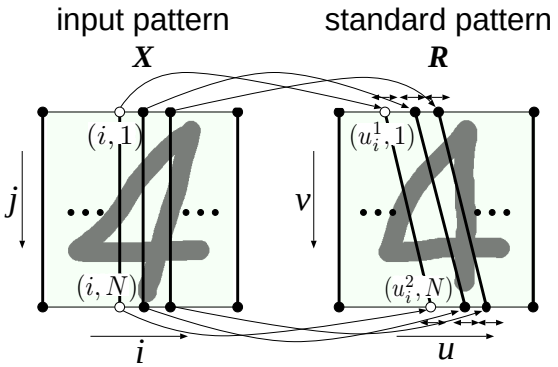


図 3　実験で使用した弾性マッチング法

このとき同時に，各 $\boldsymbol{T}_l$ を最近傍標準パターン R_k に近づくように変形したパターンが求まるが，以下ではそれを $\boldsymbol{T'}_l$ とする．

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

$$\boldsymbol{R}_k=\frac{1}{\mid\boldsymbol{C}_k\mid}\sum_{l\in\boldsymbol{C}_k}\boldsymbol{T'}_l \tag{4}$$

ここで $\mid\boldsymbol{C}_k\mid$ は $\boldsymbol{C}_k$ に属しているとされた学習パターンの個数とする．

step4: **step3** が終った時点で，標準パターンが更新された場合 **step2** へ戻る．標準パターンが更新されなかったら終了．

3.3 求まる標準パターン画像の性質

従来の k-means 法では，標準パターン更新時に標準パターンの近傍にあるパターンをそのまま平均化して新たな標準パターンとするため，学習パターンの変形に起因してぼけが生じたり，存在しないようなパターンが生じるという問題点がある．一方，本手法においては，図 2 からもわかるように平均化する前に弾性マッチングによって各学習パターンを標準パターンに近づけるよう変形しており，従って位置や変形の正規化を行っている．このため，平均化された画像にはぼけが少なくなる．この点については 4.3 節で実験的に確認する．

また学習パターン集合が，ある基本となるパターン (以下，基本パターンと呼ぶ) が変形されることで生成されているとすれば，本手法では標準パターンとしてこの基本パターンが得られる可能性が高い．今，あるクラスタに属する学習パターンを考えると，それらは弾性マッチング距離の意味でそのクラスタの標準パターンに最も近いものである．そのため，得られる標準パターンは，弾性マッチングによってそのクラスタ内の全ての学習パターンに十分近づくよう変形できるものであり，従って基本パターンであると考えられる．

4. 実　　験

4.1 使用した弾性マッチング法

以下全ての実験では，図 3 に示した弾性マッチング法 [4] を共通して用いた．この方法では，全ての $u(i,j),v(i,j)$ を制御する代わりに，入力画像パターン $\boldsymbol{X}$ 各列の両端点のみを制御点とし，その最適な対応先を標準画像パターン $\boldsymbol{R}$ 上に求めることで 2 画像間の最大一致を図る．以下に弾性マッチングの定

式化を示す．

サイズ $N \times N$ の入力画像パターン $\boldsymbol{X} = \{x(i,j) \mid i,j = 1,2,\ldots,N\}$ の第 i 列は，参照画像パターン $\boldsymbol{R} = \{r(u,v) \mid u,v = 1,2,\ldots,N\}$ に $(u_i^1, 1)$ と (v_i^2, N) を両端点とする線分に写像される．この際，同列上の画素 $(i,j)(j = 1,\ldots,N)$ の対応先は線形補間により定められるものとする．すなわち，

$$(i,j) \mapsto \big(u(i,j), v(i,j)\big) \tag{5}$$

$$\text{ただし} \begin{cases} u(i,j) = (u_i^2 - u_i^1)j/N + u_i^1 \\ v(i,j) = j \end{cases}$$

この弾性マッチングの最適化問題は次式のように定義される．

[目的関数]

$$\sum_{i=1}^{N}\sum_{j=1}^{N} || \, x(i,j) - r((u_i^2 - u_i^1)j/N + u_i^1 \ , \ j)|| \rightarrow \min \tag{6}$$

[制御変数]

$$\{u_i^1, u_i^2 \mid i = 1,\ldots,N\} \tag{7}$$

[制約条件]

$$\begin{cases} 0 \leqq u_i^p - u_{i-1}^p \leqq 2 & \ldots \text{単調連続性} \\ i - w \leqq u_i^p \leqq i + w & \ldots \text{範囲制限} \\ u_1^p = 1, \quad u_N^p = N & \ldots \text{境界条件} \end{cases} \tag{8}$$

ここで，p は 1 または 2 を表し，w は正定数である．これら (8) の制約条件は，補償対象とする変形を妥当な範囲に制限するためのものである．

4.2 文字パターンのクラスタリング及び認識実験

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 4 に正規化後のサンプルの例を示す．

これらのうち学習パターンを用いて従来手法 (ユークリッド距離によるクラスタリング法) もしくは本手法 (弾性マッチング距離によるクラスタリング法) によりクラスタリングを行い，標準パターンを求めた．図 5 に，$K = 5$ として求まった数字 0 から 9 の標準パターンの例を示す．ただし，従来手法と本手法で同じ初期パターンを用いた．

こうして求まった 2 通りの標準パターンそれぞれに対して，前述の弾性マッチングによる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．ここで，図中に破線で示した極限値とは全 N 個の学習パターンをそのまま N 個の標準パターンとした場合の認識率である．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．結果より，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 4　手書数字画像データベース MNIST のサンプル例

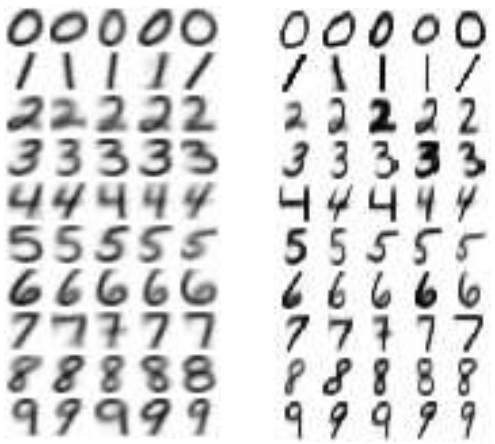


図 5　従来手法と本手法により設定された標準パターン $(K = 5)$

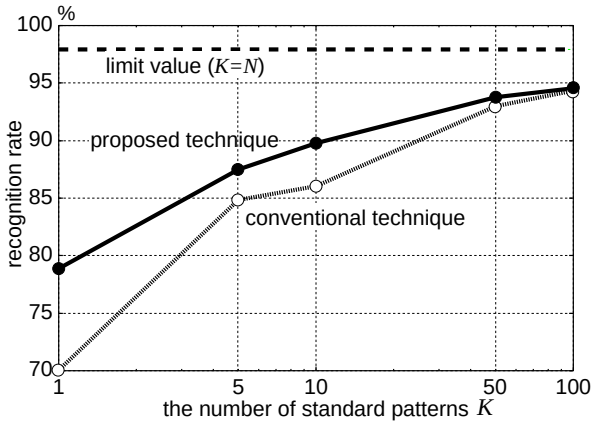


図 6　平均認識率

標準パターン数 K の増加に従って認識性能の差は小さくなっている．これは次のように説明できる．標準パターン数 K が大きくなるにつれて，1 つのクラスタに属する学習パターン数 $|\boldsymbol{C}_k|$ は小さくなる．これに従い従来手法と本手法で求まる標準パターンの差は小さくなり，特に極限すなわち $K = N$ で一致する．その結果，認識率の差も小さくなるものと考えられる．

次に弾性マッチング距離によって計算したクラスタ内の平均の歪み $\left((1/L)\sum_k \sum_{T \in C_k} D_e(\boldsymbol{R}_k, \boldsymbol{T})\right)$ を測定した．学習パターン T，および T をテストパターン X に変えて得られた測定結果をそれぞれ図 7，図 8 に示す．本手法に比べ従来手法に

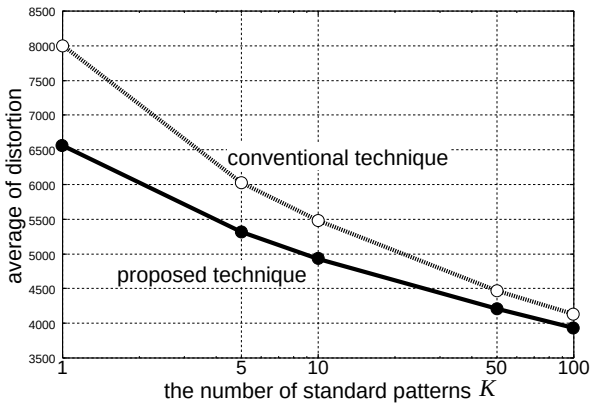


図 7　学習パターン T に対するクラスタ内平均歪み

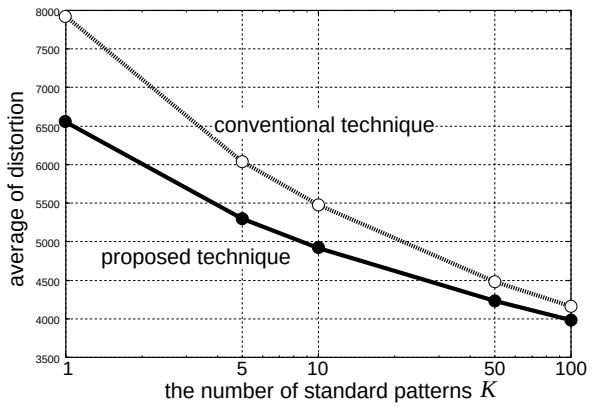


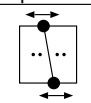
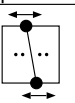
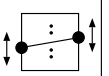
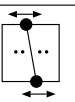
図 8　テストパターン X に対するクラスタ内平均歪み

よるクラスタリング結果にはクラスタ内の歪みが大きく，弾性マッチング距離の意味で標準パターンから遠い学習パターンもしくはテストパターンがクラスタ内に含まれていることがわかる．

以上の実験に加えて，単にクラスタリング時に弾性マッチング距離を用いたために認識性能が向上したのではなく，距離尺度が一致したためであることを示すための実験を行った．具体的には，認識時，クラスタリング時において共に同じ弾性マッチング (4.1 節) を用いた場合 (以下 (a)) と，認識時には変形モデルを 90° 回転した弾性マッチングを用いた場合 (以下 (b)) について比較を行った．このうち (b) については，認識時，クラスタリング時共に弾性マッチングを用いているものの，その伸縮の方向性の違いから，距離尺度としては異なるものを用いたことになる．(a) 及び (b) で $K = 10$ としてクラスタリング及び認識実験を行い，結果を表 1 に示す．場合 (b) による認識率は (a) だけでなく従来手法による認識率 (図 6) よりも低くなっている．これにより，本手法を用いることにより認識性能を向上することができる理由は，クラスタリング時と認識時で距離尺度が一致するためであるということがわかる．

なお，1 節でも述べたように認識率を追及するならば，GLVQ [5] など誤識別率最小化基準に基づくクラスタリング法を用いた方が有利と考えられるが，本論文では簡単のために k-means 法を適用した．本手法の，クラスタリング時と認識時で距離尺度を一致させるという考え方自体は，GLVQ にも適用可能であると考えられている．

表 1　クラスタリング時と認識時で異なる弾性マッチングを用いた場合の認識率 ($K = 10$)

	clustering phase	recognition phase	recognition rate
consistent case (a)			89.40%
inconsistent case (b)			82.23%

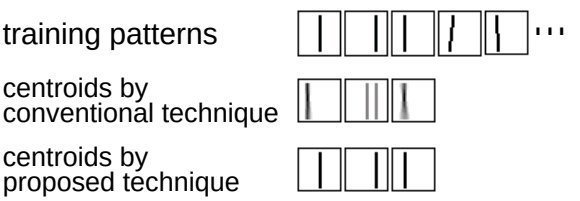


図 9　線分パターンのクラスタリング結果 ($K = 3$)

4.3　線分パターンのクラスタリング実験

図 5 の比較から，本手法 (D_e) により求めた標準パターンは従来手法 (D_r) によるものに見られるぼけが解消されていることがわかる．このことより 1 節で述べた効果 2 が確認される．この効果をさらに明確に示すために，非常に単純なパターンを用いた実験を行った．具体的には，端点の水平位置をランダムに変えて作成した線分パターンを学習パターンとして，従来手法と本手法のそれぞれで $K = 3$ としてクラスタリングを行った．求まった標準パターンを図 9 に示す．

本手法による結果にはほとんどぼけがなく，学習パターンを代表するようなパターンが得られていることがわかる．

一方，従来手法による標準パターンには画像がぼけていたり，学習パターンにはない形状が現れるなどしている．特に，本来 1 本の線分画像を対象としているにもかかわらず，2 本の線分を含む画像が標準パターンとして求まるケースが見られた．この結果は，位置や傾きが異なっていた学習パターン群を単純に平均化してしまう従来手法の限界を端的に表している．以上により，1 節で述べた効果 2 を確認することができた．

4.4　人工的に生成した文字パターンのクラスタリング実験

1 節では効果 3 として，標準パターンは学習パターン集合の基本パターンとなる可能性が高い．このことを確認するため以下の実験を行った．

まず基本パターンとして 1 つの文字パターン (図 10 A) を選び，それを 4.1 節の弾性マッチングの変形モデルにより変形できる範囲内でランダムに変形させてパターンを 1000 個作成する．これら作成されたパターンを学習パターンとしてクラスタリング ($K = 1$) し，従来手法および本手法により標準パターンを求めた．加えて，基本パターンを 2 個 (図 10 A,B) としてそれぞれから学習パターンを 500 個ずつ作成し，これらを混合した合計 1000 個のパターンに対して，$K = 2$ として同様の実験

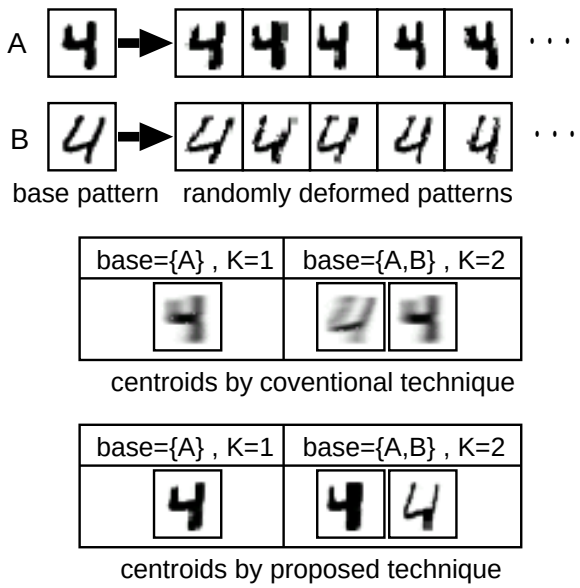


図 10　基本パターンをランダムに変形させて作成したパターンのクラスタリング結果

を行った．得られた結果を図 10 に示す．従来手法による結果では，標準パターンはぼけてしまい，学習パターンの元となった基本パターンとは異なるパターンが得られているのに対し，本手法による結果では，標準パターンには基本パターンに近いパターンが得られていることがわかる．以上により，1 節で示した効果 3 を確認することができた．

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．また本手法の副次的な効果として，ぼけのない標準パターンを得ることができ，さらに標準パターンとして学習パターンの変形の元となるパターンを得ることができることを確認した．

今後の課題としては，(i) より識別を意識したクラスタリング法 (例えば GLVQ) 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が挙げられる．特に (iii) の一例として，弾性マッチングによって得られた変位ベクトルを変位空間内においてクラスタリングする手法の検討を考えている．

謝　　辞

本研究の一部は文部科学省科学研究費補助金 (No.14780293) によった．

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.