

パターン認識の高精度化および高速化に関する研究

大町真一郎

東北大学大学院工学研究科

あらまし パターン認識技術は文字認識、音声認識、顔認識、ロボット視覚など様々な応用分野がある。文字認識や音声認識など一部実用化されたと言われている技術も、その精度は人間には遠く及ばないものであり、高精度化が望まれている。また、ロボット視覚など未だ実現の難しい分野では、パターン認識技術のさらなる高精度化および高速化が急務となっている。本研究では、現実世界の問題を対象とし、様々な目的のために利用できる汎用パターン認識手法を開発している。開発された手法は認識対象によらずに適用でき、認識の高速化・高精度化を実現できるという利点がある。

1 はじめに

パターン認識は、対象物を観測し、それが何であるかを機械で自動的に判断する処理のことであり、人間の見る、聞くといった行動をコンピュータによって実現するために必要不可欠な技術である。文字認識や音声認識のように一部実用化されている技術もあるが、最近ではセキュリティのための顔認識技術やロボットに環境を認識・理解させるための画像認識技術など、パターン認識の応用技術が注目を浴びており、そのための高精度かつ高速な認識手法が求められている。特に、机上の理論ではなく、ノイズや変形を含む現実世界の認識問題に耐えうる方法が必要となっている。

パターン認識は、一般に、対象物を観測して数値化する過程（特徴抽出過程）と数値化されたベクトルから対象物を特定する過程（識別過程）に分類される。本研究では、これらのそれぞれについて高精度かつ高速な認識を実現するための、現実的かつ一般的な理論を構築することを目的とする。具体的には、パターン認識における特徴抽出過程と識別過程のそれぞれについて高精度かつ高速な認識を実現するための方策を検討した。

2 特徴抽出法

特徴抽出過程では、ノイズや変形によらず、対象の持つ本質的な特徴をいかに抽出するかが高精度化の鍵となる。本研究では画像を対象とし、画像の本質的な特徴を抽出する手法について検討した。

その結果、複数解像度の画像を用いた新たな特徴抽出法の開発に成功した。これは、画像を対象とし、元画像に複数のパラメータでぼかしを加えた画像群を用いることにより画像の本質的な特徴を抽出する手法である。解像度の違う複数の画像から幾何的な特徴を用いることにより対象物

の大局的な構造を抽出し、次にこれを補間することにより対象物の本質的な構造を抽出する。

以下では、飾り文字を対象とし、具体的な手順を述べる [1]。飾り文字は新聞・雑誌等の見出しに用いられ、読者の目を引き付ける、レイアウトのデザイン性を高める等の特殊な目的を持つため、既成のフォントにとらわれないさまざまな形状を形作っている。

2.1 飾り文字の構造抽出

2.1.1 飾り文字の構成と構造抽出の概要

図1に飾り文字の例を示す。(A)はロゴ、(B)はさまざまなフォントの“A”の画像¹である。これらのうち、(A)および(B)の(a)(b)は一般的なテキストで文字を表したもので、(B)の(c)~(j)は細かい飾り等により文字を修飾したものである。いずれも、大局的な構造が文字の概形を表し、局所的な構造は飾りを表している。従って、大局的な構造を抽出することで飾り文字の構造を抽出できると考えられる。

画像の大局的な構造を得るためには、画像にぼかしを与えることが考えられる。図2に例を示す。(A)は“IBM”という文字列、(B)は“A”の文字の画像を表す。いずれも、(a)が原画像、(b)がぼかしを与えたもの、(c)が(b)の画像をあるしきい値で二値化したものである。(A)はうまく構造が抽出されているが、そのようにするためにはぼかしの程度や二値化のしきい値を人間が見て決める必要がある。(B)の場合はぼかしによって文字の構造が変化してしまったことが分かる。従って、一様なぼかしによって人間が介在せずに飾り文字の構造を抽出することは困難である。

本稿では、さまざまなスケールでぼかしを与えた画像を用いて必要な情報を得ることを検討する。複数のスケール

¹文献 [50] より。

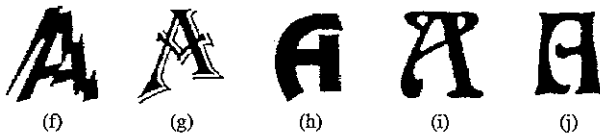


(a) IBM (b) NEW

(A) Logotype images



(a) (b) (c) (d) (e)



(f) (g) (h) (i) (j)

(B) Character images

図 1: 飾り文字の例



(a) Original image (b) Blurred image (c) Binarized image

(A) Logotype "IBM"



(a) Original image (b) Blurred image (c) Binarized image

(B) Character "A"

図 2: ぼかしと二値化による構造抽出

でぼかしを与えた画像群をここではマルチスケール画像と呼ぶ。飾り文字においては大局的な構造が文字の構造を反映し局所構造が飾りを表すことから、大局的な構造を抽出して文字の概形を得た後、補間により連結図形を得ることで、文字の構造を抽出する。

2.1.2 マルチスケール画像

ガウスマルチスケール画像等によって画像にぼかしを与えると、ぼかしの程度(スケール)によって画像の濃淡が変化する。スケール空間は各スケールにおける画像の濃淡特徴を記述する手法である。画像 $f(x, y)$ のスケール空間表現 $L(x, y; t)$ とは、 $f(x, y)$ と分散 t のガウス関数 $g(x, y; t)$ とのたたみ込みであり、以下の式で与えられる [51]。

$$L(x, y; t) = g(x, y; t) * f(x, y) \quad (1)$$

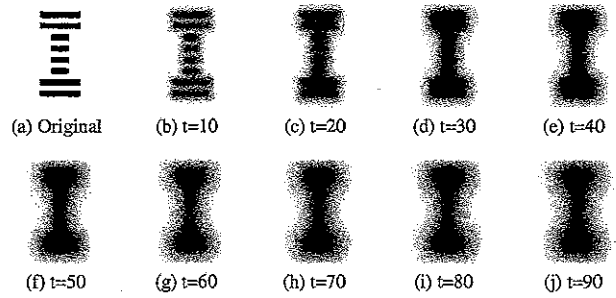


図 3: マルチスケール画像

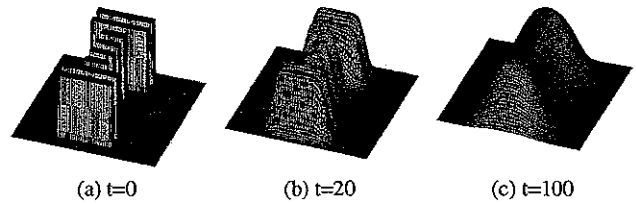


図 4: 濃淡曲面

ただし、実際には画像は離散であるから、以下に示すフィルタを画像全体に $t/\Delta t$ 回繰り返して作用させることにより $L(x, y; t)$ を求める [52]。 Δt は微量なぼかし量である。

$$\begin{pmatrix} \frac{1}{12}\Delta t & \frac{1}{3}\Delta t & \frac{1}{12}\Delta t \\ \frac{1}{3}\Delta t & 1 - \frac{5}{3}\Delta t & \frac{1}{3}\Delta t \\ \frac{1}{12}\Delta t & \frac{1}{3}\Delta t & \frac{1}{12}\Delta t \end{pmatrix} \quad (2)$$

原画像に対してさまざまな t の値でぼかしを与えることにより、マルチスケールの画像を得る。図 3 に、文字 "T" のマルチスケール画像を示す。(a) が原画像で、(b)~(j) がさまざまなスケール t でぼかしを与えた画像である。また、図 4 には、 $t = 0$, $t = 20$, $t = 100$ の場合の濃淡曲面を示す。

2.1.3 大局構造の抽出

図 4 より、飾り文字にぼかしを与えた画像では、濃淡曲面で尾根 (ridge) に相当する部分が大局構造を表すと予想される。従って、適当なスケールにおける尾根の抽出により大局構造の抽出を行う。画像中のある点の 2 次微分の絶対値が最大となる方向を p 、 p と直交する方向を q とする。ただし、簡単のため p, q は 8 方向 ($45^\circ \times n, 0 \leq n \leq 7$) に量子化して求めた。点 (x, y) がスケール t において尾根となっている条件は、

$$\begin{cases} \frac{\partial L(x, y; t)}{\partial p} = 0 \\ \frac{\partial^2 L(x, y; t)}{\partial p^2} < 0 \end{cases} \quad (3)$$

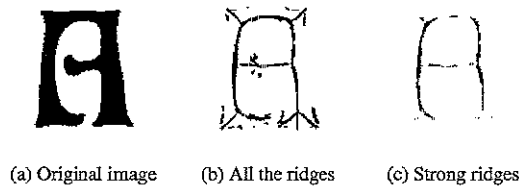


図 5: 尾根の抽出

である。ただし、構造抽出のためには式(3)を満たす点のうち、画像の特徴を表すのに適切な尾根を選択することが必要となる。本稿では尾根の強度を以下の式により定義し、強度が極大となるスケール t を選んで尾根の選択を行うものとする。

$$S(x, y; t) = \left\{ \frac{\partial^2 L(x, y; t)}{\partial p^2} - \frac{\partial^2 L(x, y; t)}{\partial q^2} \right\}^2 \quad (4)$$

点 (x, y) において強度が極大となるのは

$$\begin{cases} \frac{\partial S(x, y; t)}{\partial t} = 0 \\ \frac{\partial^2 S(x, y; t)}{\partial t^2} < 0 \end{cases} \quad (5)$$

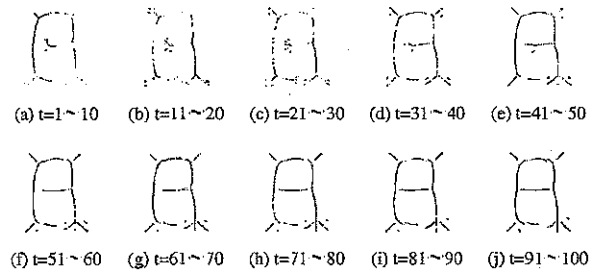
が満たされる場合である。

しかしこれを直接飾り文字に用いれば、飾り部分の局所構造まで抽出されてしまい、飾り文字が表す文字の構造を得ることは難しい。図 5 に例を示す。図 5(a) の画像に対して、式(3)および式(5)を満たす点を抽出したものが図 5(b) である。原画像から、文字“A”の大局構造が抽出されているが、同時に飾り部分の局所構造も抽出されていることが分かる。飾り文字の構造抽出においては、飾り部分の構造は不要であり、大局構造を表す尾根のみを抽出することが重要である。

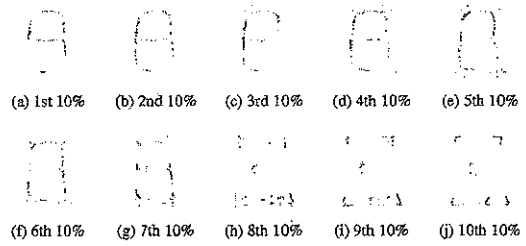
大局構造を表す尾根のみを選ぶには、

- 大きなスケールでぼかしを与えた画像のみを用いて尾根を抽出する。
- 尾根の強度(式(4)の値)が大きい点のみを用いる。

等の方法が考えられる。これらの手法の妥当性を検証するための予備実験を行った。実験は、ぼかしを与えるスケールの大きさを変えたものと、抽出する尾根点の強度を変えたものの2種類について行った。結果を図 6 に示す。図 6(A) は構造を抽出するスケールの大きさを $t = 1 \sim 10$, $t = 11 \sim 20$, ..., $t = 91 \sim 100$ の 10 段階に変え、式(3)を満たす点を抽出したものである。図より、スケールが小さい場合は局所構造、スケールが大きい場合は大局構造が得られるという傾向はあるものの、大きなスケールの場



(A) Various scales.



(B) Various strength.

図 6: 尾根の形状変化

合にも局所構造が抽出されていることが見て取れる。一方、図 6(B) は抽出する尾根点の強度を変えたものである。まず原画像に対して $t = 1 \sim 100$ でぼかしを与え、式(3)を満たす点をすべて抽出する。抽出された点の総数を N とし、尾根強度の大きい順にソートする。これを $N/10$ 個ずつ 10 個の組に分割し、尾根強度の最も大きい組(「1st 10%」と表記)から最も小さい組(「10th 10%」と表記)まで、各組に含まれる点のみをプロットした。図より、尾根強度の大きい点は大局構造を、尾根強度の小さい点は局所構造を表す傾向が顕著に現れていることが分かる。従って、大局構造を得るためには尾根強度の大きい点を抽出することが妥当であると考えられる。

そこで、本稿では以下のアルゴリズムによって大局構造のみを抽出する。まず、対象とする画像にスケールを $t = 1$ から $t = 100$ まで変えてぼかしを与え、マルチスケール画像を得る。そして、式(3)および式(5)を満たす点を抽出する。抽出された点の総数を N とし、これらの点から式(4)により表される強度の大きい順に $N \times \theta$ 個の点を抽出する。ただし、 θ は定数であり、 $0 < \theta \leq 1$ である。

提案手法により図 5(a) の画像から大局構造を抽出したものが図 5(c) である。ただし、 $\theta = 0.4$ とした。原画像の飾り部分の局所構造が消去され、大局構造のみ抽出されていることが分かる。

しかし、大局構造を抽出した後の画像(図 5(c))を見る限り、線分が連続した構造を抽出することができておらず、途切れているのが確認できる。飾り文字の認識を可能にす

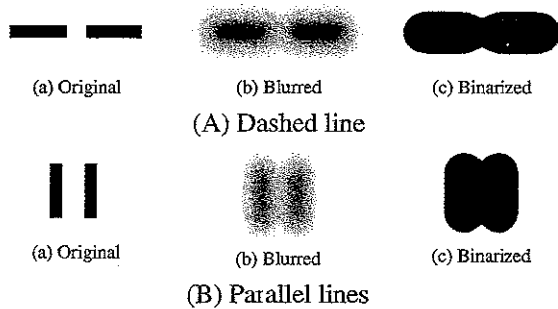


図 7: ぼかしによる構造の補間

るには、文字の特徴を表現する構造を得ることが重要であり、不連続な構造では不十分であると考えられる。次節では、不連続な構造の線分間を補間することによって、連続した構造を得る手法を提案する。

2.1.4 再帰的な谷抽出による構造補間

構造を補間するための手段のひとつに、画像にぼかしを与えることが挙げられる。ぼかしを与えて二値化することで不連続な線分の埋めるべき箇所を補間できると考えられる。しかし、これはいまよくない。以下に例を示す。補間すべき構造の典型的な例を図 7 に示す。(A) は図 5(c) のように線分が途切れたもの、(B) はテキストで構成された飾り文字からの大局構造によく見られる、本来一本の線分であるはずのものが複数の並行する線分となってしまったものである。いずれも、(a) が原画像、(b) がぼかし(blur)を与えたもの、(c) が (b) の画像をあるしきい値で二値化したものである。これらの例では線分間が広いので、補間させるためにはぼかしのスケールを大きくしなければならない。そのため、線分間のみを補間するという本来の目的が達成できず、結果として期待される補間後の画像とは形状が大きく異なってしまふ。単純なぼかしによる補間では、このような形状の大きな変化によって埋めてはいけない文字の特徴まで埋めてしまう可能性がある。

そこで補間が必要な線分間の谷 (ravine) を再帰的に抽出し、線分間を狭めることによりぼかしのスケールを最小限に抑える方法を提案する。谷の抽出手法は尾根の場合とほぼ同様である。点 (x, y) がスケール t において谷となる条件は、

$$\begin{cases} \frac{\partial L(x, y; t)}{\partial p} = 0 \\ \frac{\partial^2 L(x, y; t)}{\partial p^2} > 0 \end{cases} \quad (6)$$

である。ただし、式 (4) に相当する強度は考慮しないもの

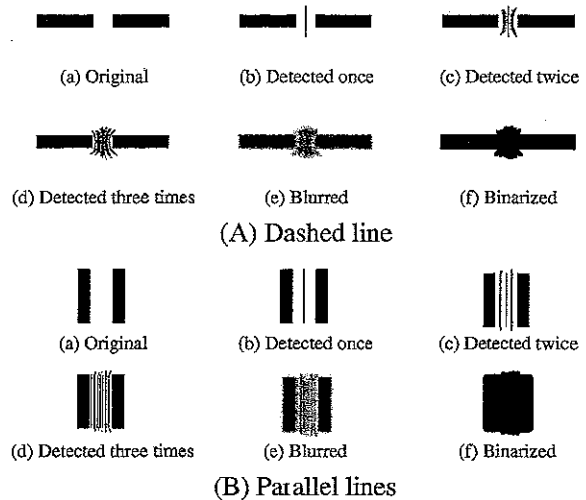


図 8: 再帰的な谷の抽出と構造の補間

とする。スケールを大きくとるほど、より幅が広い線分間の谷を抽出することになる。

谷抽出による補間のアルゴリズムを以下に示す。まずスケールの初期値 t_1 を設定する。このスケールで谷を抽出できる間隔が補間の最大幅となる。谷を抽出して、それを黒画素の線 (谷線) として加えた画像からもう一度谷を抽出する。このときその谷線を得たスケールの $1/2$ の値で谷を抽出する。つまり谷を抽出できる線分間隔が半分になったことになる。このように谷の抽出とスケールの半減を繰り返すことにより、文字に必要な特徴を埋めてしまう等の過度の補間を避けることが可能となる。上記の手順によってある程度谷を抽出し終えた画像に対し、ごく小さなスケールでぼかしを与えて二値化することにより、大幅な形状の変化なしに補間を行うことが可能になる。

図 8 に構造の補間の例を示す。(A), (B) がそれぞれ図 7 の (A)(a), (B)(a) の画像に対する処理結果である。いずれも、(a) が原画像、(b)~(d) が谷を再帰的に抽出したもの、(e) がぼかしを与えたもの、(f) は適当なしきい値で二値化したものである。(b)~(d) では、1 回ごとに谷抽出のスケールを半減させ、谷を抽出できる幅を半分に行っている。図 7 と比較して、原画像を大幅に変更させることなく線分間のみが補間されていることが分かる。

2.1.5 細線化による平滑化

前節の手法を実際の文字画像に適用した結果の画像は、線幅が一定ではなく飾り部分の局所構造が一部残り輪郭に凹凸がある等の特徴が見られた。これは文字の構造を表

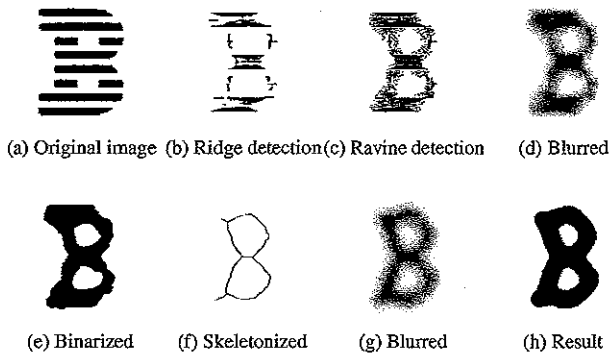


図 9: 飾り文字の構造抽出の流れ

す図形としては不適當であり、平滑化を行う必要がある。そこで、谷抽出による補間を行った画像に対して、細線化を行った後小さいスケールでぼかしを与えることで線幅を一定にし、輪郭を平滑化する。

2.1.6 飾り文字の構造抽出アルゴリズム

2.1.3, 2.1.4, 2.1.5 より、飾り文字の構造抽出アルゴリズムは次のようになる。

1. 飾り文字画像からマルチスケール画像を求め、尾根抽出により大局構造を得る。
2. 得られた大局構造から再帰的に谷を抽出し、画像に加える。
3. 谷抽出後の画像全体をごく小さいスケールでぼかしを与えた後二値化することによって補間する。
4. 細線化を行い、再度小さいスケールでのぼかしと二値化の操作を加え、飾り文字の構造を得る。

図 9 に飾り文字“B”の構造抽出を行った過程を示す。尾根抽出により大局構造が抽出され、谷抽出による構造補間によって連続図形が得られている。さらに、細線化により輪郭の凹凸が平滑化され、“B”の文字情報としての構造が得られた様子が分かる。

2.2 飾り文字の構造抽出の実験

提案手法を用いてさまざまな飾り文字に対して構造を抽出する実験を行なった。対象とした画像は図 1(A) のロゴを各文字ごとに切り出したもの(計 6 種類)と、図 1(B) の各文字であり、合計 16 種類である。

対象とする画像に対して $t = 1$ から $t = 100$ までのスケールでぼかしを与え、大局構造となる尾根を抽出した後、再帰的な谷抽出により構造を補間した。なお、大局構

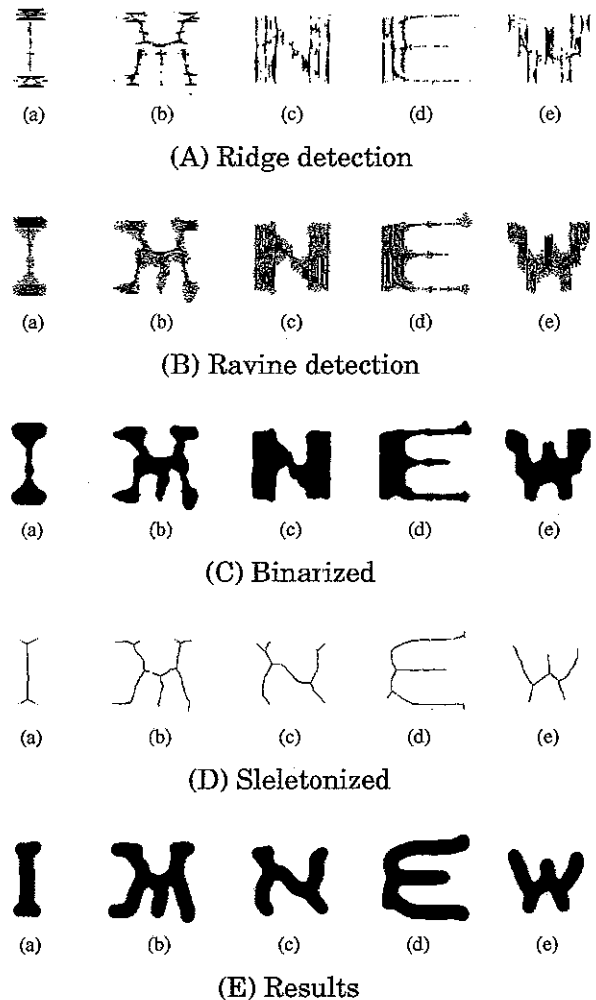


図 10: 飾り文字の構造抽出結果 (1)

造の抽出に用いるしきい値は $\theta = 0.4$ とした。谷抽出におけるスケールの初期値は $t_1 = 30$ とし、再帰的な谷抽出を 5 回繰り返した。結果を図 10, 図 11 に示す。

図 10 は図 1(A) の 6 種類の文字に対して実験を行った結果 (“B”については図 9 に示されているので省略) である。これらは主にスリット状のテクスチャで構成された飾り文字であったが、本来の文字の構造が得られていることが分かり、尾根抽出による大局構造の抽出、再帰的な谷抽出による構造の補間が有効に作用したことが確認された。

図 11 は、図 1(B) に示されるさまざまな飾り文字“A”に対して実験を行なった結果である。(a)(b) はテクスチャで構成された文字、(c) は字形が大きく異なるもの、(d)~(j) はさまざまな飾りで修飾したものと多様な字体を扱ったにもかかわらず、図の (E) を見ると多くの場合に“A”の構造を表す図形が得られており、提案手法が有効に働いていることが分かる。ただし、(c) の左下部分のように大局構造

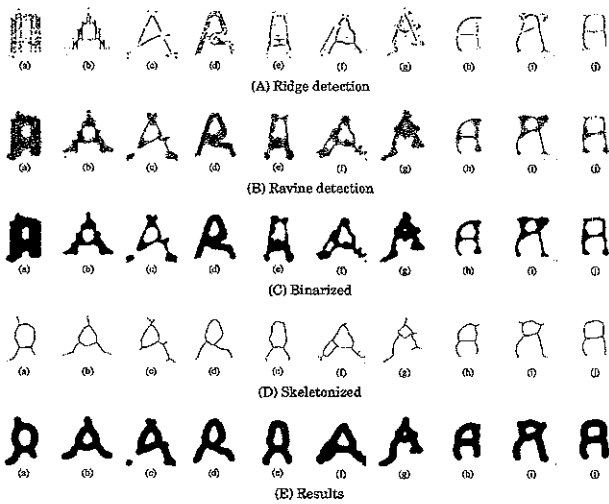


図 11: 飾り文字の構造抽出結果 (2)

を補間しきれなかったものや (i) の右下部分のように飾り部分に新たなノイズが発生したものも見られた。特に (c) は、原画像 (図 1(B)(c)) では連結図形であるにもかかわらず本手法によって分離してしまったことから、大局構造のみ抽出することによる悪影響が生じた例である。これは、抽出対象の飾り文字の形状を考慮にいれず一様なパラメータで構造抽出を行なったことが原因であると考えられる。対象とする飾り文字それぞれに対してパラメータを設定することができれば、より多様な形状の飾り文字にも対応できるようになるはずである。

図 11(E) より、(c), (i) を除いてほぼトポロジが同様の図形が得られていることが分かる。また、(c), (i) も他の図形の形状に孤立点加わっただけの形状であるから、構造解析的手法による認識 [44] が可能になるものと期待される。

3 識別手法

識別手法としては、あらかじめ何であるかが分かっている対象物を複数個用意しておき、特徴抽出過程により得られた特徴の値が分布する確率密度関数を求めておき、未知の対象物が観測された場合、確率的に最も確からしいものを識別結果とするのが一般的である。しかしこの手法では、未知の対象物がノイズ等によりあらかじめ分かっている対象物と大きく異なってしまった場合に識別が正しく行われないう問題がある。この問題を解決する手法について検討した。

その結果、ノイズに頑健な新たな識別手法の開発に成功した [4]。これは、対象物のノイズによって識別手法を補

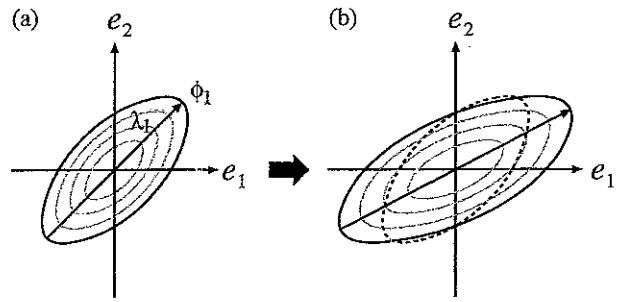


図 12: 分布形状の変化

正することで高精度化を実現する手法である。従来法はノイズに影響されにくい手法というだけでノイズの有無にかかわらず同じ手法を用いていたのに対し、提案手法はノイズによる影響を積極的に補正する手法を提案したものであり、これまでとは全く異なる新しいアプローチである。

3.1 分布形状の変化を反映した識別関数

本章では、雑音による分布形状の変化を反映した識別関数について述べる。今、特徴ベクトルのうち一部の要素にのみ雑音に加わり、他の要素は雑音がないものとする。この場合、雑音を含む要素の標準偏差が雑音の程度によって増加すると考えられる。この観点から、雑音のあるサンプルの分布形状がどのように変化するかを定量化し、雑音のないパターンによって推定された分布を雑音による変化に応じて動的に変更することを検討する。

3.1.1 定義

まず n 次元正規分布における 2 次識別関数 (先験確率はすべてのカテゴリで等しいものとする) とマハラノビス距離を考える。 μ と Σ を平均ベクトルと $n \times n$ の共分散行列とする。2 次識別関数は以下の式で表される。

$$g = (x - \mu)^t \Sigma^{-1} (x - \mu) + \log |\Sigma| \quad (7)$$

また、 μ から x への 2 乗マハラノビス距離は、以下の式で定義される。

$$d^2 = (x - \mu)^t \Sigma^{-1} (x - \mu) \quad (8)$$

$$= \sum_{i=1}^n \frac{1}{\lambda_i} ((x - \mu)^t \phi_i)^2 \quad (9)$$

ここで、 λ_i は Σ の第 i 固有値、 ϕ_i は λ_i に対応する固有ベクトルである。以下、2 乗マハラノビス距離のことを単にマハラノビス距離とよぶ。

基本的考えの説明のため、2次元正規分布を考える。図12(a)のように、 e_1 と e_2 を元の座標系における座標軸とする。 ϕ_1 と λ_1 を第1主成分に対応する固有値、固有ベクトルとする。平均0の正規分布に従う雑音が e_1 要素にのみ加わり、 e_2 要素には加わらないとした場合、 e_1 要素の標準偏差は r_b 倍になり、 e_2 要素は変化しない。ここで、 b は雑音の程度であり、 r_b は雑音の程度が b の要素の、雑音のない要素に対する標準偏差の比である。この場合、図12(b)のように分布形状が変化し、この変化を補正するために、式(7)、(8)の e_1 要素に相当する項は $1/r_b^2$ 倍にする必要がある。

n 次元の場合には、 $b(i)$ を第 i 要素の雑音の程度とし、対角行列 K を次のように定義する。

$$K = \begin{bmatrix} 1/r_{b(1)} & & & 0 \\ & 1/r_{b(2)} & & \\ & & \ddots & \\ 0 & & & 1/r_{b(n)} \end{bmatrix} \quad (10)$$

行列 K を補正行列とよぶ。雑音のない要素については、 $r_{b(i)} = 1$ である。

分布形状の変化を反映する識別関数として、補正行列 K を含む以下の関数を提案する。

$$\hat{g} = (\mathbf{x} - \boldsymbol{\mu})^t K \Sigma^{-1} K (\mathbf{x} - \boldsymbol{\mu}) + \log |\Sigma| \quad (11)$$

$$\hat{d}^2 = (\mathbf{x} - \boldsymbol{\mu})^t K \Sigma^{-1} K (\mathbf{x} - \boldsymbol{\mu}) \quad (12)$$

$$= \sum_{i=1}^n \frac{1}{\lambda_i} ((\mathbf{x} - \boldsymbol{\mu})^t K \boldsymbol{\phi}_i)^2 \quad (13)$$

K は \mathbf{x} を求める過程で分かるものとする。実際、文字認識においては、 \mathbf{x} を与える文字画像から K が決まる。

3.1.2 人工データを用いたシミュレーション

人工データを用いた2カテゴリ識別問題に式(11)と式(12)の識別関数を適用するシミュレーション実験を行い、提案した関数の有効性を確認する。

最初に、適当な n 次元正規分布に従う2つのカテゴリ(カテゴリAとカテゴリBとよぶ)を与える。これらの分布を $N(\boldsymbol{\mu}_A, \Sigma_A)$ と $N(\boldsymbol{\mu}_B, \Sigma_B)$ とする²。 $\boldsymbol{\mu}_A = (\mu_{A1}, \mu_{A2}, \dots, \mu_{An})$ とし、 Σ_A の対角要素を $\sigma_{A1}^2, \sigma_{A2}^2, \dots, \sigma_{An}^2$ と記す。

²実際は、ETL9B[45]の「あ」の200個の画像から求めた方向線素特徴量[55]の集合から $\boldsymbol{\mu}_A$ と Σ_A を、「お」の200個の画像から求めた方向線素特徴量の集合から $\boldsymbol{\mu}_B$ と Σ_B を求めた。

評価用データ

乱数を用い、分布が $N(\boldsymbol{\mu}_A, \Sigma_A)$ であるような N_E 個の n 次元ベクトルを生成する。このベクトルの集合を E_A とする。また、分布が $N(\boldsymbol{\mu}_B, \Sigma_B)$ であるような N_E 個のベクトルから成る別の集合 E_B を作る。ここで $n = 196$ であり、 $N_E = 10000$ とした。

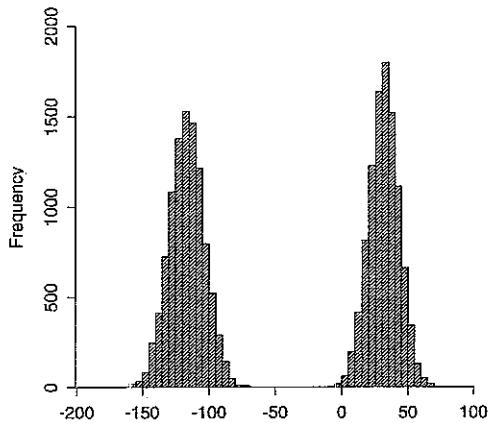
次に、 E_A の各ベクトルに以下のようにして雑音を加えることにより、集合 E'_A を作成する。各 $\mathbf{v} = (v_1, v_2, \dots, v_n) \in E_A$ に対して、 \mathbf{v} の n 個の要素から、ランダムに k 個の要素を選択する。選択された要素を $v_{\alpha_1}, v_{\alpha_2}, \dots, v_{\alpha_k}$ とする。要素 v_{α_i} ($1 \leq i \leq k$)の値を、1次元正規分布 $N(\mu_{A\alpha_i}, (r\sigma_{A\alpha_i})^2)$ に従い、乱数によって発生された値と置き換える。すなわち、 v_{α_i} の分布を決めている平均値 $\mu_{A\alpha_i}$ は同じで、標準偏差を $\sigma_{A\alpha_i}$ から r 倍の $r\sigma_{A\alpha_i}$ に変えている。ここで、 r は3.1.1で述べた r_b に対応する比の値であり、各 v_{α_i} について同一にしている。雑音を加わったベクトル毎に K が決まる。同様にして、 E_B の各ベクトルに雑音を加えることにより集合 E'_B を得る。

E_A と E_B の性質を調査した。Fisherの線形判別法[56]により E_A と E_B を最適に識別する軸を求め、 E_A と E_B に含まれるベクトルをこの軸上に射影した。軸上のヒストグラムを図13(a)に示す。図より、 E_A と E_B は線形識別関数により識別が可能であることが分かる。次に、平均ベクトル $\bar{\boldsymbol{\mu}}_{E_A}$ 、 $\bar{\boldsymbol{\mu}}_{E_B}$ および共分散行列 $\bar{\Sigma}_{E_A}$ 、 $\bar{\Sigma}_{E_B}$ を E_A と E_B より計算する。式(8)を識別関数として用いた場合、誤識別率は0%であった。

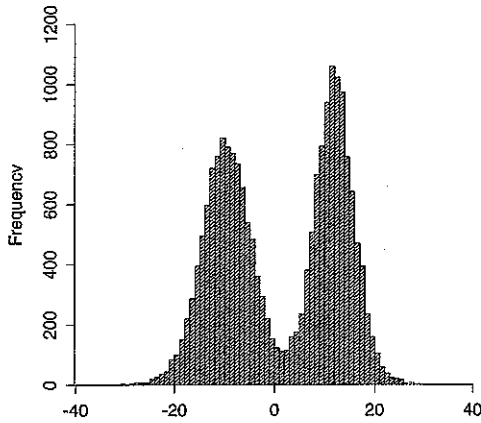
$k = 4$ 、 $r = 10.0$ の条件下で E'_A と E'_B を生成した後、これらを最適に識別できる軸上に E'_A と E'_B に含まれるベクトルを射影した。そのヒストグラムを図13(b)に示す。さらに、図13(c)は $k = 12$ 、 $r = 20.0$ の場合である。これらの場合では、 E'_A と E'_B は線形識別不可能である。式(8)を識別関数として用いた場合、誤り率は(b)で7.51%、(c)で21.4%であった。これは、 E'_A と E'_B がマハラノビス距離を用いても識別不可能であることを意味する。

訓練データ

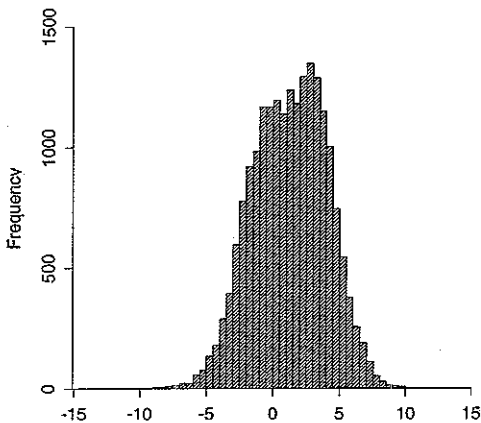
認識では、評価用データとは別に訓練データを用いて識別関数用の辞書を作成する。そこで、そのための訓練データを評価用データと同様の方法で作成して実験を行う。乱数により、 $N(\boldsymbol{\mu}_A, \Sigma_A)$ に従う N_T 個のベクトルと $N(\boldsymbol{\mu}_B, \Sigma_B)$ に従う N_T 個のベクトルを作成し、 T_A 、 T_B とする。ここで、 $N_T = 10000$ とした。 T_A 、 T_B は E_A 、 E_B とは異なる集合である。また、 T_A と T_B に E'_A 、 E'_B と同様の方法で雑音を加えることにより、 T'_A と T'_B を得る。こ



(a) E_A と E_B



(b) E'_A と E'_B ($k=4, r=10.0$)



(c) E'_A と E'_B ($k=12, r=20.0$)

図 13: 評価用データの分布

表 1: 4 種類の実験

手法	識別関数	平均ベクトル, 共分散行列	
		カテゴリ A	カテゴリ B
M1	式 (7)	$\tilde{\mu}_{T_A}, \tilde{\Sigma}_{T_A}$	$\tilde{\mu}_{T_B}, \tilde{\Sigma}_{T_B}$
M2	式 (7)	$\tilde{\mu}_{T'_A}, \tilde{\Sigma}_{T'_A}$	$\tilde{\mu}_{T'_B}, \tilde{\Sigma}_{T'_B}$
M3	式 (11)	$\tilde{\mu}_{T_A}, \tilde{\Sigma}_{T_A}$	$\tilde{\mu}_{T_B}, \tilde{\Sigma}_{T_B}$
M4	式 (11)	$\tilde{\mu}_{T'_A}, \tilde{\Sigma}_{T'_A}$	$\tilde{\mu}_{T'_B}, \tilde{\Sigma}_{T'_B}$

れら 4 個の集合 (T_A, T_B, T'_A, T'_B) を訓練データとし, 平均ベクトル ($\tilde{\mu}_{T_A}, \tilde{\mu}_{T_B}, \tilde{\mu}_{T'_A}, \tilde{\mu}_{T'_B}$) と共分散行列 ($\tilde{\Sigma}_{T_A}, \tilde{\Sigma}_{T_B}, \tilde{\Sigma}_{T'_A}, \tilde{\Sigma}_{T'_B}$) を求め, 辞書とする。

シミュレーション手法

シミュレーションにより, 雑音による補正を 2 次識別関数へ適用する効果 (式 (7), (11)) およびマハラノビス距離へ適用する効果 (式 (8), (12)) を確認する。これらはほぼ同じ操作であるため, ここでは 2 次識別関数へ適用する効果を確認するシミュレーションの手法について述べる。

訓練データから求めた辞書を用いて, 2 次識別関数 (式 (7)) と提案した識別関数 (式 (11)) により E'_A と E'_B を識別する 4 種類の実験 (M1, M2, M3, M4) を行う。 k と r の値としては様々な値を用いた。表 1 は各手法における平均ベクトル, 共分散行列, 識別関数の組み合わせを表す。各手法の定性的な意味は以下ようになる。

- M1** 雑音のないデータで辞書を作成し, 従来の 2 次識別関数を識別関数として用いる。
- M2** 雑音のあるデータで辞書を作成し, 従来の 2 次識別関数を識別関数として用いる。
- M3** 雑音のないデータで辞書を作成し, 提案した識別関数を用いる。
- M4** 雑音のあるデータで辞書を作成し, 提案した識別関数を用いる。

なお, 各評価用ベクトルの補正行列 K は E'_A と E'_B を作る時点で求められている。

結果

図 14 は 2 次識別関数, 図 15 はマハラノビス距離を用いた場合の, 4 つの手法による誤り率を表す。各図において, (a) は雑音を加える要素数を 4 に固定し, r として様々な値を用いたもの, (b) は r を 10 に固定して雑音を加える

要素数を様々に変えたものである。また、それぞれの図において左図がすべての場合の誤り率を表示したものであり、右図は左図の一部を拡大したものである。

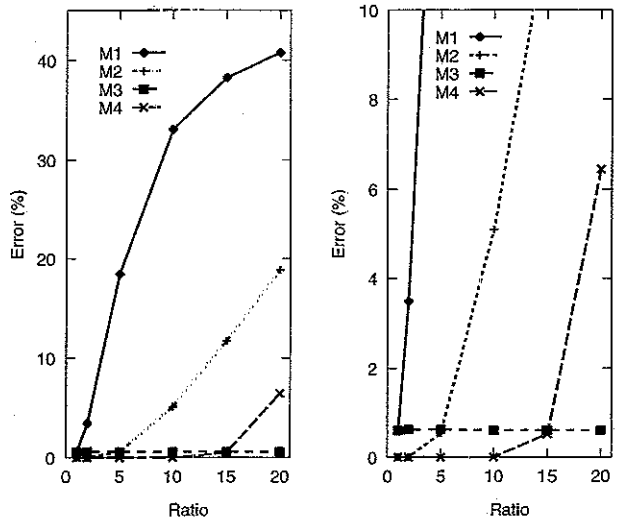
2次識別関数の場合もマハラノビス距離の場合も同じような結果が得られた。図から分かるように、M1とM2の誤り率は k や r が大きくなるに従って増加する。M3の誤り率は k が大きくなるに従って大きくなるが、 r にはほとんど依存しない。定性的には、雑音の加わる要素数が多くなれば誤り率は増加するが、雑音の程度には依存しないことになる。M4の誤り率は、マハラノビス距離の場合は実験したすべてのパラメータで0%であった。2次識別関数の場合は、図14(a)から分かるように r の値が小さい範囲ではM3より小さいが、 r の値が大きくなるとM3より大きくなる。これは、 r が大きいと式(11)の $\log|\Sigma|$ の項が正しく求まらないことが原因であると考えられる。

M1の誤り率は、 k や r が小さくても比較的高い。M1とM2の結果を比較すると、M2の誤り率がかなり低く、雑音のあるデータを認識するためには雑音のあるデータを用いて辞書を作るのが有効であることが示された。また、M2とM3の結果を比較すると、 r が小さいときはM2の方が良いが、M3は r にほとんど依存しないため r が大きい場合はM3の方が良くなる。 $r \geq 5$ の場合はM3の誤り率はM2の誤り率よりも小さい。これは、雑音のひどいデータを認識するためには雑音のあるデータを用いて辞書を作っただけでは不十分であり、データの雑音を検出して補正することが重要であることを意味する。

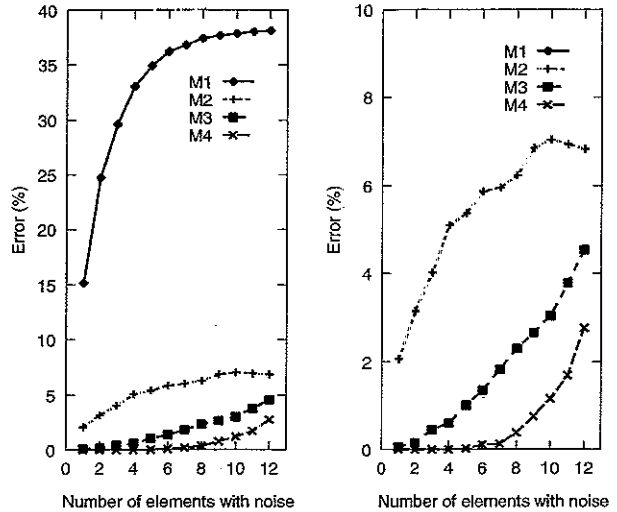
さらに、マハラノビス距離を補正した式(12)を用いた場合、 $k = 12$ 、 $r = 20.0$ の場合でさえもM4の誤り率は0%であった。先に述べたように、 E'_A と E'_B の誤り率は $k = 12$ 、 $r = 20.0$ のとき21.4%である(このときの分布は図13(c)に示されている)。これは、従来の統計的な手法では識別が不可能なカテゴリ対も、雑音を補正することで識別が可能となることを意味する。つまり、提案手法は既存の統計的な手法を補い、高い識別性能を実現するものである。

3.2 低品質文字認識への適用

漢字の画像を用い、提案した手法の文字認識における有効性を確認する実験を行った。先の考察より、雑音がひどい時は2次識別関数よりもマハラノビス距離を用いた方が良い結果が得られるため、認識にはマハラノビス距離を用いる。



(a) 雑音を加える要素数を固定 ($k = 4$)



(b) 比の値を固定 ($r = 10.0$)

図14: シミュレーション結果(2次識別関数)

3.2.1 マハラノビス距離の様々な変形

文字画像を認識する場合、2次識別関数やマハラノビス距離は計算量が膨大でありサンプル数が少ない場合に誤差が生じる等の欠点がある [57]。これらの問題を解決するために、マハラノビス距離の様々な変形が提案されている。典型的なものとして、擬似マハラノビス距離 (Quasi-Mahalanobis distance; QMD)[58]、改良型マハラノビス距離 [59] (Modified Mahalanobis distance; MMD)、簡素化マハラノビス距離 [5] (Simplified Mahalanobis distance; SMD) 等がある。

これらのうち、SMD はマハラノビス距離と同じ特性を持ち、その識別性能は EIL9B [45] を用いた識別実験により確認されている [5]。SMD を説明するために、式 (9) を以下のように書き直す。

$$d^2 = \sum_{i=1}^m \frac{1}{\lambda_i} ((x - \mu)^t \phi_i)^2 + \sum_{i=m+1}^n \frac{1}{\lambda_i} ((x - \mu)^t \phi_i)^2 \quad (14)$$

SMD は式 (14) の第 2 項の λ_i を $i = m + 1, \dots, n$ における λ_i の平均値で置き換えたものであり、以下の式で与えられる。

$$\begin{aligned} d_S^2 &= \sum_{i=1}^m \frac{1}{\lambda_i} ((x - \mu)^t \phi_i)^2 \\ &+ \sum_{i=m+1}^n \frac{1}{\lambda} ((x - \mu)^t \phi_i)^2 \quad (15) \\ &= \sum_{i=1}^m \frac{1}{\lambda_i} ((x - \mu)^t \phi_i)^2 \\ &+ \frac{1}{\lambda} \left\{ \|x - \mu\|^2 - \sum_{i=1}^m ((x - \mu)^t \phi_i)^2 \right\} \quad (16) \end{aligned}$$

ここで、

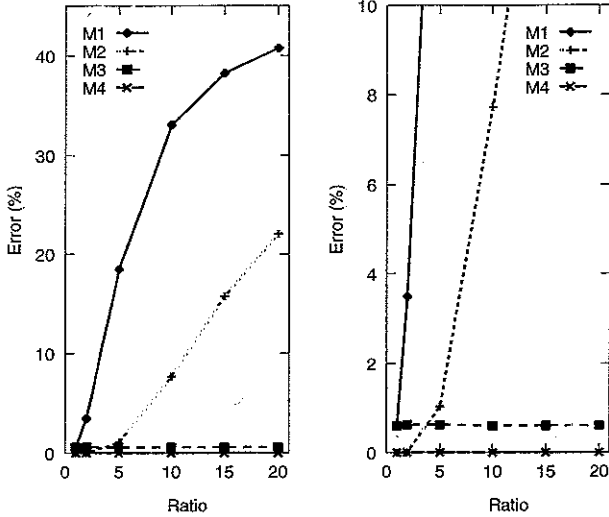
$$\lambda = \frac{1}{n - m} \sum_{i=m+1}^n \lambda_i \quad (17)$$

$$= \frac{1}{n - m} \left\{ S - \sum_{i=1}^m \lambda_i \right\} \quad (18)$$

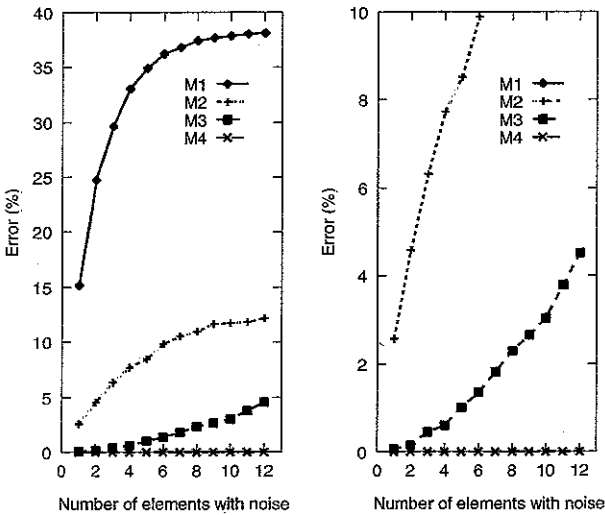
であり、 S は Σ の対角要素の和である。

3.1.1 で定義した補正行列 K を SMD の式に導入したものを \hat{d}_S で表すと、

$$\begin{aligned} \hat{d}_S^2 &= \sum_{i=1}^m \frac{1}{\lambda_i} ((x - \mu)^t K \phi_i)^2 \\ &+ \frac{n - m}{S - \sum_{i=1}^m \lambda_i} \left\{ \|K(x - \mu)\|^2 \right. \\ &\left. - \sum_{i=1}^m ((x - \mu)^t K \phi_i)^2 \right\} \quad (19) \end{aligned}$$



(a) 雑音を加える要素数を固定 ($k = 4$)



(b) 比の値を固定 ($r = 10.0$)

図 15: シミュレーション結果 (マハラノビス距離)

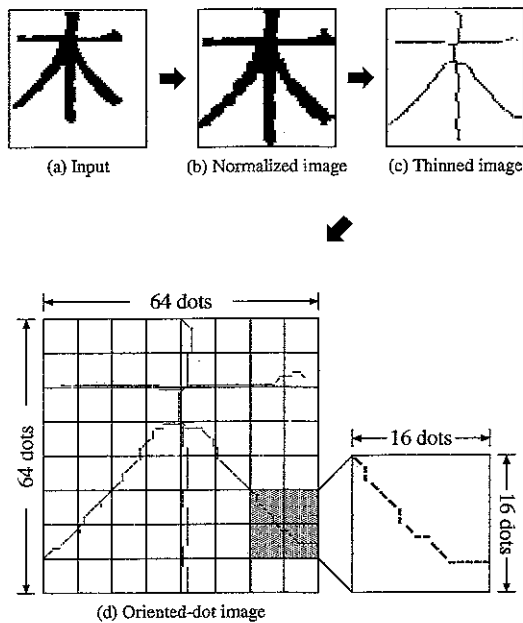


図 16: 方向線素特徴量

となる。式 (19) を適応型 SMD (Adaptive SMD) とよび、低品質文字の認識に適した識別関数として用いることにする。

3.2.2 方向線素特徴量

文字認識の特徴量としては方向線素特徴量 [55] を用いた。方向線素特徴量を求めるには、まず入力画像を 64 ドット × 64 ドットの大きさに正規化し、細線化・線素化処理を施す。線素としては縦 (|), 横 (—), 右上斜め (／), 左上斜め (＼) の 4 種類を割り当てる。入力画像から線素化後のパターンを求めるまでの例を図 16 に示す。次にこれを 8 ドット × 8 ドットの正方形領域に分割し、その隣り合う 4 個ずつを一つの小領域 (例えば、図 16(d) の網かけをした領域) とする。全部で 49 個の小領域 (部分領域) となる。各小領域ごとに線素の数を重み付きで数えることで特徴量とするので、 $49 \times 4 = 196$ 次元のベクトルとなる。

3.2.3 雑音の定量化

雑音の程度を表す尺度としては、つぶれ度 [24] を用いる。これは、文字画像の細線化処理の過程で求める。細線化の方法としては、Hilditch の方法 [60] とほぼ同様の結果が得られる 10 近傍法 [61] を用いた。この方法では線幅が 2 ドット以上のとき、それを輪郭から 1 ドットずつ削ることを繰り返すことで、最終的に線幅が 1 のパターンを得る。この繰り返し回数を制限すると、つぶれのある領域は

表 2: つぶれ度と標準偏差の比の関係

つぶれ度	比
0	1.0
1	5.0
2, 3	8.3
4 ~ 6	10.0
7	12.5
8	20.0

線幅が 1 にならずに残るため、残った黒画素の数を部分領域ごとに数えて量子化し、つぶれ度とする。つぶれ度は 0 から 8 までの値を取り、0 であることはつぶれがほとんどないことを意味し、つぶれ度の値が大きくなるほどつぶれがひどくなることを意味する。

つぶれない領域のベクトルに対する、つぶれのある領域の特徴ベクトルの標準偏差の比がどのように変化するかを調査した [24]。結果を表 2 に示す。表より、つぶれがひどいほど標準偏差の比の値が大きくなる事が分かる。

3.2.4 認識実験

実験には JIS 第一水準の 2,965 種類の漢字を用いた。単一フォントの 10 種類の大きさの印刷文字 (6 ポイントから 22 ポイント) を訓練データとして用いた。すべてのサンプルパターンは 400dpi のイメージスキャナで読み込まれ、特徴ベクトルに変換される。評価データとしては、3 種類の大きさの印刷文字 (6, 7, 8 ポイント) を 2 つのモード (薄いモードと濃いモード) で複写したものをを用いた。文字画像は複写によりつぶれが生じ、特に濃いモードで複写したものはつぶれのひどい画像となっている。9 ポイント以上の文字画像はほとんどつぶれがないため、評価データとしては選ばなかった。

識別関数としては、式 (16) の SMD と式 (19) の適応型 SMD を用いた。ただし、訓練データ数 10 を考慮して $m = 5$ とした。先に述べたシミュレーション実験では、式 (16) を用いたものは M2 に、式 (19) を用いたものは M4 に対応する。各未知入力文字画像から、文字画像の各小領域ごとのつぶれ度により表 2 を用いて補正行列 K を定める。実験結果 (誤り率) を図 17 に示す。適応型 SMD を用いることで、すべての場合において、誤り率が減少する (または変化しない) ことが確認された。この結果により、提案した分布形状の変化を反映する識別関数が文字認識に有効であることが確認された。

表 3 に、従来の SMD では不正解であったが適応型 SMD により正しく認識された文字の例を示す。表には、入力された文字画像、正解、SMD によって得られた候補文字が

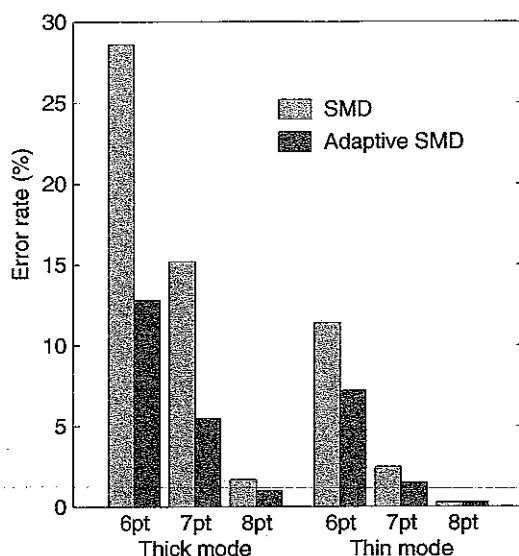


図 17: 実験結果

表 3: 適応型 SMD により正しく認識された文字

	(a)	(b)	(c)	(d)
文字画像				
正解	葦	胎	闇	憶
SMD による候補	繁	胎	簡	攪

示されている。すべての組み合わせにおいて、正解と候補は非常に良く似た文字であり、雑音のある(つぶれた)部分が誤認識の原因になっていると思われる。適応型 SMD を用いることで、相対的に綺麗な部分が識別に重要な役割を果たすことになり、正しく認識された。

例えば (b) の場合、画像の右側(旁)が比較的綺麗であるため、SMD による候補文字として右側に同じ構造を持つ文字が得られている。この文字画像では、左上が綺麗であり、この部分に関して正解文字と候補文字とが異なるため、適応型 SMD では正しく認識されたと考えられる。

4 むすび

本稿では、パターン認識の高精度化および高速化を目的とし、パターン認識における特徴抽出過程と識別過程のそれぞれについて高精度かつ高速な認識を実現するための研究結果について述べた。

まず、特徴抽出においては、さまざまな飾り文字の構造を表す画像を得る手法を提案した。まず、マルチスケール画像を用いて尾根を抽出した後、尾根強度の強い点のみ抽出することで飾り文字の持つ大局構造を得る手法を提案

した。そして、再帰的な谷の抽出によって線分の不連続な構造を補間する手法を提案した。さらに、細線化による平滑化と合わせ、飾り文字の構造を得るアルゴリズムを提案し、飾り文字の構造抽出実験を行った。実験結果から、本手法により文字の構造を表す画像が抽出でき、同じ文字についてはさまざまな飾り文字からほぼ同様のトポロジを持つ図形が得られることが確認された。飾り文字には、もとの字体をテキストで表したもの、もとの字体を変形させたもの、飾りを加えたり字体の一部を削ったもの等、実にさまざまなものがあり、一様なばかし等により構造抽出を行うことは困難である。しかし、本手法によりもとの字体の構造を表す画像を得ることができ、同一字種ではほぼ同じトポロジの図形を得ることができた。これは、本手法が飾り文字の構造抽出に有効であり、さまざまな飾り文字に柔軟に対応できることを示すものである。本稿では限られた英文字を対象とした。今後、漢字等での有効性を確認する必要がある。また、本手法により得られた画像を用いた認識手法を考案することは今後の課題である。さらに、今回はロゴとして用いられている文字列については人間が一文字ずつ切り出しを行ったが、認識の前段階として文字列から個々の文字を切り出す手法の検討も必要である。

また、識別手法としては、ノイズに頑健な新たな識別手法について述べた。従来の統計的パターン認識手法では、特徴ベクトルの分布形状が最も重要であった。しかし、データに雑音を加わった場合、ベクトルの分布形状が変化し、誤認識の大きな原因となっていた。これを解決するため、雑音の特徴を考慮し、雑音による分布形状の変化を動的に補正することのできる識別関数を提案した。分布形状の変化を表すために、雑音による標準偏差の変化を定量化した場合の補正行列を提案し、それをベイズ識別器やマハラノビス距離等の識別関数にいかん適用すべきかを検討した。そして、人工データを用いたシミュレーション実験と実際の文字画像を用いた認識実験により、その効果を確認した。適切な K が推定できれば、提案した識別関数を文字認識以外に適用することが可能であるが、それは今後の課題である。また、文字認識については、自然言語等の知識を用いた後処理と組み合わせる実用的な認識システムを構築することが今後の課題である。

パターン認識技術は、多くの応用分野を持ち、文字認識や音声認識など一部実用化されたと言われている技術もその精度は人間には遠く及ばないものであり、高精度化が望まれている。本研究で得られた成果をより現実世界の問題と密着したものとし、パターン認識に関連した様々な問題を解決していくことが今後の課題である。

謝辞 本研究を進めるにあたり、多大な御指導を賜った東北大学大学院工学研究科阿曾弘具教授に感謝する。

文献

特徴抽出法に関するもの

- [1] S. Omachi, M. Inoue, and H. Aso, "Structure Extraction from Decorated Characters Using Multiscale Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 315-322, March 2001
- [2] S. Omachi, M. Inoue, and H. Aso, "Structure Extraction from Various Kinds of Decorated Characters Using Multi-Scale Images," *Proceedings of the 15th International Conference on Pattern Recognition*, vol. 4, pp. 455-458, September 2000
- [3] 狩野芳正, 大町真一郎, 阿曾弘具, "特徴選択によるテキストチャ画像の教師なし領域分割," *電子情報通信学会論文誌 D-II*, vol. J86-D-II, no. 7, pp. 988-995, July 2003

識別手法に関するもの

- [4] S. Omachi, F. Sun, and H. Aso, "A Noise-Adaptive Discriminant Function and Its Application to Blurred Machine-Printed Kanji Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 314-319, March 2000.
- [5] F. Sun, S. Omachi, and H. Aso, "Precise selection of candidates for handwritten character recognition using feature regions," *IEICE Trans. Inf. & Syst.*, vol. E79-D, no. 5, pp. 510-515, 1996.
- [6] H. Ujiie, S. Omachi, and H. Aso, "A Discriminant Function Considering Normality Improvement of the Distribution," *Proceedings 16th International Conference on Pattern Recognition (ICPR 2002)*, vol. 2, pp. 224-227, August 2002
- [7] M. Iwamura, S. Omachi, and H. Aso, "A Method to Estimate the True Mahalanobis Distance from Eigenvectors of Sample Covariance Matrix," *Lecture Notes in Computer Science*, vol. 2396 (Joint IAPR International Workshops SSPR 2002 and SPR 2002), pp. 498-507, August 2002
- [8] I. Kato, S. Omachi, and H. Aso, "Asymmetric Gaussian and Its Application to Pattern Recognition," *Lecture Notes in Computer Science*, vol. 2396 (Joint IAPR International Workshops SSPR 2002 and SPR 2002), pp. 405-413, August 2002
- [9] I. Kato, S. Omachi, and H. Aso, "Mixture of Asymmetric Gaussians and Its Maximum Likelihood Estimator," *New Trends in Optimization and Computational Algorithms (NTOC2001)*, pp. 105-106, December 2001
- [10] F. Sun, S. Omachi, N. Kato, H. Aso, S. Kono, and T. Takagi, "Two-Stage Computational Cost Reduction Algorithm based on Mahalanobis Distance Approximations," *Proceedings 15th International Conference on Pattern Recognition (ICPR2000)*, vol. 2, pp. 700-703, September 2000
- [11] M. Iwamura, S. Omachi, and H. Aso, "A Modification of Eigenvalues to Compensate Estimation Errors of Eigenvectors," *Proceedings 15th International Conference on Pattern Recognition (ICPR2000)*, vol. 2, pp. 378-381, September 2000
- [12] S. Omachi, F. Sun, and H. Aso, "A New Approximation Method of the Quadratic Discriminant Function," *Lecture Notes in Computer Science*, vol. 1876, pp. 601-610, August/September 2000

- [13] S. Omachi, F. Sun, and H. Aso, "A Discriminant Function for Noisy Pattern Recognition," *Proceedings of the 11th Scandinavian Conference on Image Analysis*, pp. 793-800, June 1999.
- [14] F. Sun, S. Omachi, and H. Aso, "An Algorithm for Estimating Mixture Distribution of High Dimensional Vectors And Its Application to Character Recognition," *Proceedings of The 11th Scandinavian Conference on Image Analysis (SCIA'99)*, pp. 267-274, June 1999
- [15] F. Sun, S. Omachi, and H. Aso, "An Algorithm for Constructing a Multi-template Dictionary for Character Recognition Considering Distribution of Feature Vectors," *Proceedings 14th International Conference on Pattern Recognition (ICPR'98)*, pp. 1114-1116, August 1998
- [16] S. Omachi, F. Sun, and H. Aso, "Precise Recognition of Blurred Chinese Characters by Considering Change in Distribution," *Proceedings of The 10th Scandinavian Conference on Image Analysis*, pp. 501-506, June 1997.
- [17] 大町真一郎, "文字認識のためのパターン認識," *O plus E*, vol. 25, no. 10, pp. 1143-1149, October 2003
- [18] 岩村雅一, 大町真一郎, 阿曾弘具, "パターン認識における予測分布の偏りに関する一考察," *電子情報通信学会論文誌 D-II*, vol. J87-D-II, no. 2, pp. 407-415, February 2004
- [19] 氏家秀紀, 大町真一郎, 阿曾弘具, "クラスの確率分布を仮定しない特徴変換に基づく2次識別関数," *電子情報通信学会論文誌 D-II*, vol. J86-D-II, no. 4, pp. 441-449, April 2003
- [20] 岩村雅一, 大町真一郎, 阿曾弘具, "標本共分散行列の固有ベクトルを用いた真のマハラノビス距離の推定法," *電子情報通信学会論文誌 D-II*, vol. J86-D-II, no. 1, pp. 22-31, January 2003
- [21] 大町真一郎, 阿曾弘具, "低品質文字認識におけるつづれを動的に補正する部分空間法," *電子情報通信学会論文誌 D-II*, vol. J82-D-II, no. 11, pp. 1930-1939, November 1999
- [22] 大町真一郎, 阿曾弘具, "分枝限定法と計算回数削減の推定に基づく近隣識別法の高速化," *電子情報通信学会論文誌 D-II*, vol. J82-D-II, no. 4, pp. 641-649, April 1999
- [23] 孫方, 大町真一郎, 加藤寧, 阿曾弘具, "特徴量の要素の相関を考慮した高速・高精度な識別関数と文字認識への応用," *電子情報通信学会論文誌 D-II*, vol. J81-D-II, no. 9, pp. 2027-2034, September 1998
- [24] 大町真一郎, 阿曾弘具, "つづれを考慮した低品質印刷文字の高精度認識," *信学論 (D-II)*, vol. J79-D-II, no. 9, pp. 1534-1542, 1996.
- [25] 大町真一郎, 孫方, 阿曾弘具, "カテゴリ間分布を考慮した文字認識用マルチテンプレート辞書の構成法," *電子情報通信学会論文誌 D-II*, vol. J79-D-II, no. 9, pp. 1525-1533, September 1996

実際の問題への応用に関するもの

- [26] N. Kato, M. Suzuki, S. Omachi, H. Aso, and Y. Nemoto, "A Handwritten Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 258-262, March 1999
- [27] S. Taguchi, S. Omachi, and H. Aso, "Fast Visual Search Using Simplified Pruning Rules - Streamlined Active Search," *Proceedings 17th International Conference on Pattern Recognition (ICPR 2004)*, vol. 4, pp. 937-940, August 2004

- [28] S. Omachi, F. Sun, and H. Aso, "Precise Estimation of High-Dimensional Distribution and Its Application to Face Recognition," Proceedings 17th International Conference on Pattern Recognition (ICPR 2004), vol. 1, pp. 220-223, August 2004
- [29] T. Kato, S. Omachi, and H. Aso, "Precise Hand-Printed Character Recognition Using Elastic Models via Nonlinear Transformation," Proceedings 15th International Conference on Pattern Recognition (ICPR2000), vol. 2, pp. 364-367, September 2000
- [30] H. Aso, S. Ohmachi, N. Sun, Y. Katsuyama and M. Kimura, "SEIUN: Systolic Elaborated Intelligent Ultra-high-speed Neosystem for Character Recognition," 1991 Workshop on Computer Architecture for Machine Perception (CAMP'91), pp. 471-482, December 1991
- [31] H. Aso, S. Ohmachi, Y. Katsuyama and M. Kimura, "Systolic Architecture for Fast and Intelligent Character Recognition Systems," Proceedings of Pacific Rim International Conference on Artificial Intelligence (PRICAI'90), pp. 805-810, November 1990
- [32] 大町真一郎, 布田寿康, 阿曾弘具, "文書画像中の折れ線グラフの認識," 画像ラボ, vol. 15, no. 4, pp. 38-41, April 2004
- [33] 布田寿康, 大町真一郎, 阿曾弘具, "連結成分追跡による文書画像中の折れ線グラフの認識," 電子情報通信学会論文誌 D-II, vol. J86-D-II, no. 6, pp. 825-835, June 2003
- [34] 加藤 毅, 大町真一郎, 阿曾弘具, "伸縮変形モデルを用いた手書き文字認識," 電子情報通信学会論文誌 D-II, vol. J83-D-II, no. 12, pp. 2578-2586, December 2000
- [35] 大町真一郎, 横山弘子, 阿曾弘具, "遺伝的アルゴリズムを用いた対象のユークリッド空間への配置法," 電子情報通信学会論文誌 D-II, vol. J82-D-II, no. 12, pp. 2195-2202, December 1999
- [36] 大町真一郎, 阿曾弘具, "品質を考慮した印刷文字認識手法," 情報処理学会論文誌, vol. 38, no. 12, pp. 2490-2498, December 1997
- [37] 鈴木雅人, 大町真一郎, 加藤 寧, 阿曾弘具, 根元義章, "混合マハラノビス関数による高精度な類似文字識別手法," 電子情報通信学会論文誌 D-II, vol. J80-D-II, no. 10, pp. 2752-2760, October 1997
- [38] 張 新学, 大町真一郎, 阿曾弘具, "RCE モデルに基づくネットワークを用いたファジィ回帰分析," 日本ファジィ学会誌, vol. 9, no. 4, pp. 533-540, August 1997
- [39] X. ZHANG, S. OMACHI, and H. ASO "Fuzzy Regression Analysis Using RFLN and Its Application," Proceedings Sixth IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'97), pp. 51-56, July 1997
- [40] 阿曾弘具, 大町真一郎, 木村正行, 勝山 裕, "高速高精度知的認識システム SEIUN," 電子情報通信学会論文誌 D-II, vol. J76-D-II, no. 3, pp. 474-484, March 1993
- [41] 大町真一郎, 阿曾弘具, 木村正行, "高速文字認識前処理用ストリックアレー," 電子情報通信学会論文誌 D-II, vol. J73-D-II, no. 2, pp. 167-174, February 1990
- [42] 大町真一郎, 阿曾弘具, 木村正行, 勝山 裕, "超高精度知的文字認識システム SEIUN," 画像ラボ, vol. 5, no. 1, pp. 22-26, January 1994

参考文献

- [43] 山本和彦, "弛緩整合法による手書き教育漢字認識," 電子情報通信学会論文誌 (D), vol. J65-D, no. 9, pp. 1167-1174, 1982.
- [44] J. Rocha and T. Pavlidis, "A shape analysis model with application to a character recognition system," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 4, pp. 393-404, 1994.
- [45] 齊藤泰一, 山田博三, 山本和彦, "JIS 第 1 水準手書漢字データベース ETL9 とその解析," 電子情報通信学会論文誌 (D), vol. J68-D, no. 4, pp. 757-764, 1985.
- [46] 若林哲史, 鄧 巖, 鶴岡信治, 木村文隆, 三宅康二, "非線形正規化と特徴量の圧縮による手書き漢字認識の高精度化," 電子情報通信学会論文誌 (D-II), vol. J79-D-II, no. 5, pp. 765-774, 1996
- [47] 横瀬善正, "形の心理学," 名古屋大学出版会, 1986.
- [48] M. Sawaki and N. Hagita, "Recognition of degraded machine printed characters using a complementary similarity measure and error-correction learning," IEICE Trans. Inf. & Syst., vol. E79-D, no. 5, pp. 491-497, 1996
- [49] R. M. Haralick, L. I. Watson, and I. J. Laffey, "The topographic primal sketch," International Journal of Robotics Research, vol. 2, no. 1, pp. 50-72, 1983
- [50] 小川 英光 編著, "パターン認識。理解の新たな展開—挑戦すべき課題—," 電子情報通信学会, 1994.
- [51] Tony Lindeberg, "Edge detection and ridge detection with automatic scale selection," Computer Society Conference on Computer Vision and Pattern Recognition, pp. 465-470, 1996.
- [52] 本谷秀聖, 出口光一郎, "画像濃淡を表す曲面の主曲率に着目したマルチスケール形状記述手法," 情報処理学会研究報告, 96-CVIM-102, 1996.
- [53] 森 健一, "文字読取装置における画像処理," テレビジョン, vol. 26, no. 11, pp. 949-958, 1972.
- [54] N. Sun, M. Ito, J. Kudoh, Y. Nemoto, H. Aso, and M. Kimura, "A method of binarization based on information of character structure," Advanced Session of IEEEJ on Document Proceeding and Digital Imaging, 1992
- [55] 孫 寧, 田原 透, 阿曾弘具, 木村正行, "方向線素特徴量を用いた高精度文字認識," 電子情報通信学会論文誌 (D-II), vol. J74-D-II, no. 3, pp. 330-339, 1991.
- [56] 大津展之, "パターン認識における特徴抽出に関する数理的な研究," 電総研研究報告, no. 818, pp. 78-91, 1981.
- [57] 竹下鉄夫, 木村文隆, 三宅康三, "マハラノビス距離の推定誤差に関する考察," 電子情報通信学会論文誌 (D), vol. J70-D, no. 3, pp. 567-573, 1987.
- [58] 栗田昌徳, 鶴岡信治, 横井茂樹, 三宅康三, "加重方向ヒストグラムと疑似マハラノビス距離を用いた手書き漢字・ひらがな認識," 信学技報, PRL 82-79, pp. 105-112, 1983
- [59] 加藤 寧, 安倍正人, 根元義章, "改良型マハラノビス距離を用いた高精度な手書き文字認識," 電子情報通信学会論文誌 (D-II), vol. J79-D-II, no. 1, pp. 45-52, 1996
- [60] C. J. Hilditch, "Linear skeleton from square cupboards," In: *Machine Intelligence 6*, B. Meltzer & D. Michie, Eds., Univ. Press, Edinburgh, pp. 403-420, 1969.
- [61] 阿曾弘具, "並列処理向き細線化アルゴリズム," 電子情報通信学会論文誌 (D-II), vol. J76-D-II, no. 9, pp. 2148-2150, 1993.