

パターン空間の構造解析とそれに基づく認識技術に関する研究  
A Study on pattern recognition techniques based on pattern space  
analysis methodology

February 2008

Graduate School of Science and Engineering  
Waseda University

坂野 鋭  
Hitoshi Sakano



# Contents

List of Figures	iii
List of Tables	v
<b>1 序論</b>	<b>1</b>
1.1 研究の背景と目的	1
1.2 本論文の構成	3
<b>2 現代のバタン認識装置の課題</b>	<b>5</b>
2.1 現代のパターン認識装置の基本構成	5
2.2 特徴抽出器の概論	5
2.3 識別器と代数構造	6
<b>3 手書き文字認識における特徴抽出系開発の歴史</b>	<b>11</b>
3.1 概要	11
3.2 様々なヒューリスティックな特徴抽出	11
<b>4 遺伝的アルゴリズムによる文字認識系の解析</b>	<b>15</b>
4.1 はじめに	15
4.2 特徴ベクトルからの文字画像再構成問題	16
4.3 文字画像再構成アルゴリズム	17
4.4 文字画像再構成実験	19
4.5 従来手法による識別系の解析	20
4.6 参照ベクトルと識別境界の可視化	21
4.6.1 最尤推定の場合	22
4.6.2 LVQ の場合	23
4.6.3 議論	25
4.7 まとめと今後の課題	26
4.8 その後の発展	27
<b>5 質問学習による文字認識辞書の学習</b>	<b>29</b>
5.1 はじめに	29
5.2 文字認識系のための質問学習	30
5.2.1 質問学習に関する従来の研究	30
5.2.2 文字認識系への適用	30
5.3 手書数字認識実験	33
5.4 考察	34
5.5 まとめと今後の課題	35

<b>6</b>	<b>物体認識と非線形性</b>	<b>37</b>
6.1	物体認識問題と非線形性	37
6.2	これまでの研究とその問題点	39
<b>7</b>	<b>核非線形相互部分空間法による物体認識</b>	<b>41</b>
7.1	はじめに	41
7.2	相互部分空間法, 非線形主成分分析	42
7.2.1	相互部分空間法	42
7.2.2	非線形主成分分析法	43
7.3	核非線形相互部分空間法	44
7.4	認識実験	47
7.5	考察	47
7.6	まとめと今後の課題	50
7.7	その後の発展	50
<b>8</b>	<b>結論：パターン認識におけるアルゴリズム設計の課題</b>	<b>51</b>
8.1	パターン集合の線形性とアルゴリズムの有効性の関係	51
8.1.1	文字認識アルゴリズムについて	51
8.1.2	物体認識アルゴリズムについて	52
8.2	残された問題	52
8.2.1	より優れた認識系の設計論の実現に向けて	52
8.2.2	人間の様な認識系の実現に向けて	52
<b>A</b>	<b>Coulomb 相互作用モデルによる二次識別関数の学習的構成法</b>	<b>55</b>
A.1	はじめに	55
A.2	Coulomb 相互作用モデル	56
A.2.1	評価関数の定式化	56
A.2.2	Bayes 識別面との関係	58
A.3	学習アルゴリズム	59
A.3.1	重み付きユークリッド距離によるモデル	59
A.3.2	重み変更規則の導出	60
A.3.3	学習の収束規準	61
A.4	識別面のダイナミックス	62
A.4.1	理論の確認	62
A.5	文字認識実験	65
A.6	まとめと今後の課題	67
<b>B</b>	<b>画像再構成のための Affine Mutation</b>	<b>73</b>
B.1	はじめに	73
B.2	GA による画像再構成アルゴリズムと従来の突然変異手法	73
B.2.1	GA による画像再構成	73
B.2.2	画像再構成問題における突然変異手法	74
B.3	Affine Mutation	74
B.4	文字画像再構成実験	75
B.5	考察	75
B.6	まとめと今後の課題	76
	<b>References</b>	<b>81</b>

# List of Figures

2.1	文字画像，顔画像の平均．左辺の文字画像，顔画像はそれぞれ明らかに文字であり顔であるが，右辺の画像は文字画像でも顔画像でもない．このことは濃淡特徴の空間における「文字画像」「顔画像」の集合の線形性の欠如を示唆している．	10
4.1	ePDC 特徴の抽出過程	18
4.2	GA を用いた手法のフローチャート	20
4.3	2次元交叉の概念図	21
4.4	原画像と再構成画像 (a):原画像 (b) 再構成画像	21
4.5	個体の平均二乗誤差の世代発展．誤差棒は個体の二乗誤差の標準偏差	22
4.6	パターンの進化の例．同じ行の 10 パターンは当該世代よりランダムに抽出した	23
4.7	解析に用いる文字認識系の流れ図	24
4.8	学習パターンの中のカテゴリ「0」「2」の全誤読パターン．パターンの下の数字は認識結果を示す	24
4.9	最尤推定で作られた参照ベクトルとその中点の再構成画像．マトリクス対角要素の画像は参照ベクトルに，他の画像はその中点に対応している．	25
4.10	LVQ で作られた参照ベクトルとその中点の再構成画像	27
5.1	提案する質問学習機械の構成	31
5.2	質問学習に伴う認識率の推移	34
5.3	初期の質問に因って作られる不自然な境界の例	35
6.1	画像の回転に伴う非線形性の原因．右の様に 1-2 度程度の変動であれば，特徴空間中での画像の変化は小さく，ほぼ同じ形態になる為，ほぼ線形の分布になると考えられる．一方，左の様に数十度の回転がある場合，回転前後の中点は回転途中の画像と一致しない．すなわち，回転が大きいつきは線形な変動をしないことが分かる．なお，この画像データは UMIST(University Manchester Institute of Science and Technology) データベースを用いた．	38
7.1	認識率の次元数依存性	48
7.2	線形主成分分析による顔画像の散布図	48
7.3	kPCA による顔画像の散布図	49
A.1	誤識別領域の近似的な「体積」	59
A.2	テストに用いたサンプルと従来の手法で得られる識別面	63

A.3	提案手法による誤り訂正の様子 . . . . .	64
A.4	緩和的な方法を用いない場合のポテンシャルの変化 . . . . .	65
A.5	緩和的な方法を用いない場合の誤識別数の変化 . . . . .	66
A.6	緩和的な方法を用いた場合のポテンシャルの変化 . . . . .	67
A.7	緩和的な方法を用いた場合の誤識別数の変化 . . . . .	68
A.8	提案手法によって得られる識別面 . . . . .	69
A.9	学習データが十分に多い場合の識別面 . . . . .	70
A.10	文字認識辞書に適用した場合のポテンシャルの変化 . . . . .	71
A.11	文字認識辞書に適用した場合の誤識別数の変化 . . . . .	72
B.1	提案手法による到達二乗誤差 . . . . .	75
B.2	従来手法による到達二乗誤差 . . . . .	76
B.3	従来手法による到達二乗誤差 (拡大) . . . . .	77

# List of Tables

4.1	最尤推定で作成した参照ベクトルを用いた場合の認識率 . . . . .	22
4.2	学習パターンに対する confusion matrix . . . . .	22
4.3	未知パターンに対する confusion matrix . . . . .	24
4.4	LVQ による認識率 . . . . .	26
5.1	対象とした文字認識系の認識率 . . . . .	33
7.1	Recognition rate of each procedure . . . . .	47
A.1	実験に用いた Gauss 分布のパラメーター . . . . .	62
A.2	正規乱数学習終了時の識別率の比較 . . . . .	64
A.3	手書き漢字学習終了時の識別率の比較 . . . . .	66





# Chapter 1

## 序論

### 1.1 研究の背景と目的

本研究の主な目的は、現存するパターン認識装置の動作原理を解明するとともに、それを指針にしたパターン認識機械の設計方法を確立することにある。

20世紀初頭の計算機が現れる前から、パターン認識の研究は開始され、1950年ごろまでは主としてハードウェアの研究として進められていた[24]。計算機が現れ、一般化するとともに当初はハードウェアのシミュレーションの研究として、そして計算機の進化、つまり小型化、高性能化、低価格化とともに入力装置以外はアルゴリズムの研究として進むことになる。

しかしながら、人間、あるいはほかの動物のパターン認識機能の原理がわかっていない現状では必ずしも科学的な研究がなされてきたとは言えない。よく、「飛行機は羽ばたいて飛ぶわけではないので人工知能も人間と同じメカニズムである必要はない」という反論もなされるが、鳥も飛行機も同じ流体力学の原理に従って飛ぶのであり、メカニズムが異なるからといって原理まで異なっているわけではない[14]。

それに対して、人類がこれまで実現してきたパターン認識機械はまさに人間によって創造されたものでありながらその原理が十分に理解されていないものもある。

例えば、初期のシングルフォントを対象にした文字認識装置は、単純な重ね合わせのアルゴリズムが用いられ、誤認識の原因も物理的な雑音に限られたため、その原理が明らかで技術的な問題への対処方法も明快であった。

これがマルチフォント漢字認識、手書文字認識などの様により複雑な対象を認識する必要が生じてきたときに、徐々に問題が顕在化してくる。

当然のことながらマルチフォントや手書文字では同じ文字種であっても形状が異なるため単純な重ね合わせでは高い認識率を得ることはできない。

また、一方で文字認識以外にも認識対象は広がり、指紋、顔などの現在バイオメトリクスと呼ばれる生体を用いた個人認証技術[33]、交通標識、車線などを対象とした自動車運転支援技術なども研究開発が進み、現時点では既に部分的に実用化している。

これらの認識系について共通する問題は

- 人間はこれらの対象全てを認識することが出来るにもかかわらず、実装されている認識アルゴリズムは著しく異なることが多い。つまり、現代のパターン認識アルゴリズムには汎用性が無い。

- パターン認識の原理がわかっていないため、一部の物理的な問題が明確なもの以外、経験的に認識アルゴリズムが設計されることが多く、その原理がわからないことが珍しくない。

ということである。

こうした現象は、認識対象が困難になるにつれ、認識系、特に特徴抽出系および前処理系が複雑化し、基本原理がわからないながらも認識精度が高くなればよしとしてきたという問題が根源にある。すなわち、人間のパターン認識の原理がわからないながらも結果的に同様な機能を実現しようとする努力のために自ら動作原理の理解できない装置を創造してしまうわけである。

本論文では、これらのパターン認識装置に共通する数学的な動作条件として、

1. パターン集合の線形性
2. 特徴空間のユークリッド性

が実現されている場合であることを実験的に立証することを目標として研究を行った。

これらの条件は、パターン認識の言葉に訳すと、

1. あるカテゴリに属するデータを同士を加算したもの、あるいはスカラー倍したものは、やはり同じカテゴリに属すること。例えば平均操作で作られたパターンはやはり同じカテゴリに属するパターンでなくてはならない。
2. パターンが存在している特徴空間では、パターン同士の類似度が評価出来なくてはならない（その意味では必ずしも「ユークリッド的」である必要はないが、象徴的な意味で「ユークリッド的」という言葉を用いる）また、そのときの類似度が人間の感じる類似度と近ければ、よい認識系を与える。

と言い換えることが出来る。

さらに概念を簡略化し、パターン集合が線形空間である状態をパターン空間と名づけると、パターン認識装置の動作条件として、「特徴空間がユークリッド的であり、その上でパターン集合がパターン空間をなす場合」と定義することが出来る。

このことは一見、自明のことに思われるが、現実のパターン認識装置、時として既に実用に供されているパターン認識装置でも、これらの条件が実験的に示されている例は殆ど無い。

本論文では、現在実用に供されている光学的文字認識装置 (Optical character reader; 以下 OCR) と、現在も活発に様々な研究が行われている物体認識装置を対象に研究を行った。

ただし、これらの装置に関する著者の立ち位置は微妙に異なっている。アルゴリズムの研究に限っても、OCRの研究は既に40年以上の歴史を持ち、既に多くのアプリケーションで実用に供されている。にもかかわらず、OCRが上記のパターン認識装置の動作条件を満足していることは検証されていない。現代の文字認識装置は高度に複雑化し、アルゴリズムの動作を観察すること自体が困難になっているからである。すなわち、上述の「人間が創造したにもかかわらず、人間が動作を理解できない装置」の典型的な例となっている。従って、OCRについては現存するパターン認識装置の動作原理を実験的に探ることが研究の主目的になる。

一方、顔認識系に代表される物体認識の技術の研究が活発化したのは、91年の Turk[97], 94年の村瀬 [54] らの研究からであり、現時点では（その兆候は見え

るものの) OCR ほど不可解なアルゴリズムは用いられていない<sup>1</sup>。つまり, 物体認識系に関しては, 新しいアルゴリズムを上記の動作条件に従って設計することが主眼となる。

最初に現在主として用いられているパターン認識技術のうち学習, 識別装置に関して考察し, これが動作する条件を示す。その上で既存に文字認識アルゴリズム, 特に特徴抽出系についてその性質を解析し, それに基づく学習法を検討する。さらに, 物体認識問題で現れる非線形性を考察し, 撮像条件をも考慮に入れた上での新しい認識アルゴリズムを提案する。すなわち, 本論文では, 既存の文字認識系の性質を解析し, その際的设计論を物体認識問題に適用することで新しい認識アルゴリズム的设计論を確立することを試みたものである。

## 1.2 本論文の構成

本論文は以下のように構成される。

第2章では現代的なパターン認識機械の概論を述べるとともに, その学習部分が動作する条件が特徴空間のユークリッド性, パターン集合の線形性であることを示す。

その上で第3章では現代までの手書文字認識装置における特徴抽出系の歴史的発展をたどり, これらの特徴抽出系が上記の条件を満たしているかが不明確であることを示す。

この問題を明らかにするために第4章では遺伝的アルゴリズムを用いた特徴ベクトルからの文字画像再構成手法を提案し, 実験的に有効性を示すとともにこれを用いて特徴空間やパターン集合の性質を解析した結果を示す。その結果, 現在良好に動作している文字認識装置の特徴抽出系が作る特徴空間がユークリッド的であり, 特徴空間に存在するパターン集合が線形空間をなすことを実験的に示す。また, この方法により誤認識の原因や学習機械の学習プロセスを直感的に理解できることも示す。

この研究の成果を受けて第5章では質問学習のアルゴリズムを文字認識装置に実装する方法を提案する。そして, これにより, 文字認識装置の性能が向上できることを実験的に示す。

次に, 第6章では顔画像認識などの物体認識で現れる姿勢変動に伴う非線形性に関して議論し, その原因とこれまでの研究をまとめる。

第7章ではこれを解決する新しい認識アルゴリズム核非線形相互部分空間法を提案し, 実験的に有効性を示すとともに, このアルゴリズムも2章で示された条件が満足された場合, すなわち特徴空間上のパターン集合が線形である場合に良好に動作することを示す。

最後に第8章で本研究をまとめるとともに今後に残された課題を示す。

<sup>1</sup>現状, 顔認識系より顔検出系でその傾向は顕著である。Gabor フィルタの適用などは方向性特徴の導入と位置づけることが出来るが, Joint Haar feature などはかなりヒューリスティックな要素を含んでいると思われる。



## Chapter 2

# 現代のパターン認識装置の課題

### 2.1 現代のパターン認識装置の基本構成

パターン認識技術はおよそ 100 年前の文字認識の研究に始まり，音声認識，物体認識，バイオメトリクスと呼ばれる体の一部を用いた個人識別方法など多数の分野に広がるとともに，技術も深化し，これらの一部は実用に供されている [22]．

しかしながら，現在研究されているパターン認識技術は，必ずしも人間の行っているそれとは関係が無い．むしろ，例えば飛行機が羽ばたいて飛ぶのではないように，同じ機能（とはいっても人間の認識能力には遠く及ばないのであるが）を実現するために試行錯誤的に研究されてきた技術が結実したものというべきであろう．

現代的なスタイルのパターン認識研究ではパターン認識装置を特徴抽出器と識別器に分けて論じる<sup>1</sup>．

特徴抽出器とはスキャナやカメラから取り込まれた文字画像や物体画像などの認識対象となるデータから特徴ベクトルと呼ばれる高次元データを抽出する装置である．その性質は対象に依存することが多く，特徴抽出系の設計のためには対象となるデータに対する知見が重要な役割を果たす．

このとき抽出される特徴ベクトルは数百次元から数千次元の高次元に上ることも珍しくはない．そのため，特に手書文字認識の様に長く研究されてきた分野における特徴抽出系は，非常に複雑化し，その有効性の根拠が理解できない様な状況に陥っている．しかも，多くの場合，一見異なる特徴抽出系が同様に動作するように見える場合もあり，さらに混乱の原因となっている．

本章では一般論としての特徴抽出器の議論を行い，その上で識別に関する尺度の議論，そしていくつかの学習機械の例を紹介し，それらの機械が動作する理由を考察する．

### 2.2 特徴抽出器の概論

この節ではパターン認識分野で言われる「特徴」とはどういうものかについて，概念的な事項を説明する．

通常，パターン認識技術は構造解析型とパターンマッチング型に分類されるが，構造解析型の方法は文献に譲り [55]，本研究で主として対象としているパターンマッチング型で用いられる特徴について説明する．

<sup>1</sup>あまり本質的ではないが，人間がこのようにしているという考えられているわけではない

画像パターン認識で対象となるのは何らかの画像であるが、通常、同じカテゴリに分類される画像は大きく異なっている場合が多い。パターンマッチング型のパターン認識機会における特徴抽出系の目標はカテゴリ内の相違を除去し、カテゴリ間の違いを明確化することといえる。

また、特徴抽出の方式は画像の性質によっても左右される。例えば文字認識の問題の多くは白黒2値画像の問題として扱われるが、顔画像認識などの物体認識問題の多くはグレースケールの多値画像の問題として扱われる。

2値画像の問題の場合には、図と地の境界が明らかであるために図、あるいは地の形状を何らかの方法で記述する方式がとられる。

多値画像の場合にはグレースケールの勾配などの特徴を圧縮するなどの方法が取られることが多い。

また、いずれの場合も Fourier 変換やアダマール変換、Wavelet 変換などの数理的特徴抽出処理が採用される場合もある。

いずれにせよ、特徴は画像から数値列として取り出される。このため、特徴は一般に「特徴ベクトル」と呼称される。

ただし、こうした特徴ベクトルが本来の目的であるカテゴリ内変動を吸収しているかどうかは殆ど確認された例が無い。多くの文献では認識実験の結果からそのようなことが行われていると解釈されているに過ぎない。

次節では、識別器の動作を通じて「特徴ベクトル」が満たすべき性質を議論する。

## 2.3 識別器と代数構造

識別器は入力されたパターンがどのカテゴリであるかを決定する機械である。一般的には、識別器は何らかの形でそのカテゴリに関する典型的な情報である参照情報<sup>2</sup>を持ち、入力されたパターンと何らかの類似度を計測して、最も類似していると判断されたカテゴリ名を認識結果として出力する。

類似尺度としては、単純なユークリッド距離やマンハッタン距離、あるいはベクトル間の内積などの比較的単純なものや統計的な分散を考慮したマハラノビス距離などが用いられることが多い。最近では、これらの類似度の変わりに核が用いられることも多い。

一方で動的計画法に基づく DP 距離や、順序の交換を考慮した編集距離の様な類似尺度もある。

もうひとつの問題は「あるカテゴリに関する典型的な情報」をどのように保持するかである。一つの考え方はあるカテゴリが確率分布を満たすと考え、その確率モデルを持つことである。今ひとつの考え方は、複数のカテゴリの違いに着目してカテゴリ間の識別面を推定することである。通常、前者を密度推定型、後者を識別型と呼んでいる。

さらに確率密度関数を仮定し、そのパラメータを推定するものをパラメトリックな方法、密度関数を仮定しない方法をノンパラメトリックな方法と呼んでいる。

識別器が学習・識別を行うプロセスをいくつかの簡単なモデルで考察してみよう。

<sup>2</sup>文字認識、文書認識などの分野では「辞書」と呼ばれることが多い。顔画像認識、指紋認識などの個人識別を目的とした応用では「テンプレート」と呼ばれる。これらと同様な意味で「参照ベクトル」という言葉が使われることもある。本論文では、これらには本質的に違いが無いとして議論を進める。なお、音声認識の分野では「モデル」と呼ばれる。

## ユークリッド距離識別器

パラメトリックな識別器で最も簡単なものは等分散の Gauss 分布，

$$p_i(\vec{x}) = \frac{1}{\sqrt{2\pi}|\Sigma|} \exp(-(\vec{x} - \vec{\mu}_i)\Sigma_i^{-1}(\vec{x} - \vec{\mu}_i)^T) \quad (2.1)$$

を仮定した識別器である．ここで  $\vec{x}$  は特徴ベクトル， $\vec{\mu}$  は平均値， $\Sigma$  は分散共分散行列， $i$  はカテゴリを表すインデックスである．式 (2.1) は  $\vec{x}$  がカテゴリ  $i$  に属する確率を表している．

等分散が仮定されているので， $\Sigma$  は各カテゴリで同じになる為識別の際には無視することが出来る．故に，このような識別器の学習は参照情報としての平均値  $\vec{\mu}_i$  を推定する問題となる．すなわち，カテゴリ  $i$  に属する  $N$  個の特徴ベクトル  $\vec{x}_{1i}, \vec{x}_{2i}, \dots, \vec{x}_{Ni}$  が与えられた時に

$$\vec{\mu}_i = \frac{1}{N} \sum_{j=1}^N \vec{x}_{ji} \quad (2.2)$$

を計算することに他ならない．

識別時には入力された特徴ベクトル  $\vec{x}$  と各カテゴリを表現する平均ベクトル  $\vec{\mu}_i$  のユークリッド距離を計算し，距離が最小となったカテゴリ名  $i$  を認識結果として出力する．

 $k$ -最近傍識別器とポテンシャル関数法

よく誤解されるが， $k$ -最近傍識別器は上記のユークリッド距離識別器と同様に，入力されたサンプルがどのカテゴリに属するかの確率を計算することを基本原理としている [64]．

$k$ -最近傍識別器では全ての学習パターンから抽出された特徴ベクトル  $\vec{x}_{1i}, \vec{x}_{2i}, \dots, \vec{x}_{Ni}$  が与えられた時にこれを全て記憶することで，カテゴリに関する参照情報とする．

識別時には入力された特徴ベクトル  $\vec{x}$  から全ての学習パターンとの距離を計算し，最も近いものから上位  $k$  位の中で最も数の多い学習パターンが属するカテゴリを認識結果として出力する．

このプロセスは，ある超球の中にあるカテゴリの確率すなわち  $k_i/k$  を求めていることに他ならない．ここで  $k_i$  は上位  $k$  位の中のカテゴリ  $i$  に属する学習データの数である．これがノンパラメトリックな密度関数推定型と言われる所以である．

関連する，やはり単純なノンパラメトリックな密度関数推定型識別器がポテンシャル関数法，あるいはカーネル法<sup>3</sup>と呼ばれる識別器である [1]．

ポテンシャル関数法では学習時にやはり全ての学習データを学習し，識別時にはポテンシャル関数（もしくは核） $k(\vec{x}, \vec{x}_{ji})$  を計算し，カテゴリごとにその和

$$s_i = \sum_{j=1}^N k(\vec{x}, \vec{x}_{ji}) \quad (2.3)$$

を計算し，最も大きなカテゴリ名を認識結果として出力する．

ポテンシャル関数としては Gaussian などを用いることが多い．つまり，あるパターンが存在したときには，その近傍に類似あるいは同カテゴリのパターンがある程度の確率で存在するという仮定を仮定して，確率密度関数を推定するわけである．

<sup>3</sup>この分野では kernel をカタカナで「カーネル」と呼ぶのが多数派であるが旧来から「核」という和訳があるので以降は「核」もしくは「核関数」という言葉を用いる

## 線形識別器

簡単のために 2 カテゴリーの識別問題を考える．このとき線形識別器は

$$g(\vec{x}) = \vec{a} \cdot \vec{x} + b \quad (2.4)$$

の正負で  $\vec{x}$  のカテゴリを判定するものということが出来る．ここで  $\vec{a}$  は識別器の重み， $b$  はマージンと呼ばれる．

識別器の重みは何らかの評価関数を最適化することで計算される．様々な線形識別器が提案されているが，これらの違いは本質的には評価関数の違いではない場合が多い．例えば，線形判別分析では級間分散と級内分散の比を最大化であり，単純パーセプトロンでは学習パターンの誤識別最小化，サポートベクターマシン (Support vector machine; 以下 SVM) ではマージンの最大化である．

単純パーセプトロンの学習は逐次的に，判別分析や SVM の学習はフォーマルに行われるがこれらは必ずしも本質的な違いではない．

また，多くの非線形識別器はこれらの線形識別器を非線形変換して用いるか，複数用いて非線形の識別境界の区分線形近似を行うものである．前者の代表が非線形 SVM であり，後者の代表が多層パーセプトロン (Multi layered perceptron; 以下 MLP) である．

## 二次識別器

二次識別器は非線形ではあるがフォーマルに構成できる点でほかの非線形識別器とは区別して取り扱う必要がある．最も代表的な二次識別器は Bayes 識別器と呼ばれるもので，式 (2.1) で分散が仮定されない場合を考えればよい．

式 (2.1) の対数を取り  $-1$  を乗ると

$$g_i(\vec{x}) = (\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) - \log \frac{1}{\sqrt{2\pi}} + \log |\Sigma| \quad (2.5)$$

という 2 次識別関数と呼ばれる識別器が得られる．この識別器の学習は平均値，分散共分散行列の推定，分散共分散行列の逆行列の計算で行われる．この識別器の右辺第 2，3 項を無視したものがいわゆる Maharanobis 距離である．また，分散共分散行列の非対角成分を無視した重みつきユークリッド距離という 2 次識別器も実用上は多用される．

二次識別器の逐次学習的な構成法はあまり知られていない，パーセプトロンに 2 次以上の多項式で変換を加えた  $\psi$  マシンや著者らの提案した Coulomb 相互作用モデルがわずかに知られている (付録 A 参照)．

そのほかに実用的にも理論的にも重要な識別器として部分空間法があるが後述する．

## 代数構造

これで基礎的な識別器の学習・識別過程を概観したわけだが，全ての識別器について動作の前提になっているのは，特徴空間で広い意味での内積が計算出来ることであることがわかる．ユークリッド距離識別器や  $k$ -最近傍識別器ではユークリッド距離が，線形識別器であれば参照ベクトルと入力ベクトルの内積が計算できなくてはならない．なお，核も広い意味での内積である．

また，式 (2.2) を見ればわかるとおり「パターンの集合」にも代数構造が必要である．つまり同じカテゴリに属するパターンは和を取ったりスカラー倍しても同じカテゴリのパターンにならなくてはならない．これはパターン集合というものが局所的に出会っても線形空間でなくてはならないことを示している．



すなわち、特徴抽出系がこの条件を満足するように設計されていなくては、明らかに識別器は動作しない。このことは全ての識別器が線形代数を基礎に設計されていることから自明のことに思われる。しかし、一方でこのことを実験的に示した研究は殆ど無いと言ってよい。

実際、図 (2.1) に示すとおり、濃淡特徴を取った場合には「顔画像の集合」「文字画像の集合」はそれぞれ明らかに線形性が無い。

それでは何故現実のパターン認識装置は動作しているのか？その問題を次章以降に、文字画像、物体画像について明らかにしていく。

ただし、ここで問題になるのはその手法である。線形空間であることは元を和、スカラ倍したのも線形空間に属する、と定義されるが、パターン集合があるカテゴリに属するかどうかについては定義がなく、人間が主観的に判断するしかない<sup>4</sup>。

従ってパターン集合が当該の特徴空間の中で線形構造を持っているかを判定するためには、

1. 濃淡特徴などの様に人間が認識可能な特徴の場合には、和やスカラ倍したパターンが当該カテゴリに属しているかを判定する。
2. 人間が認識できない特徴の場合には、何らかの方法で人間が理解できる形に変換してから和やスカラ倍したパターンが当該カテゴリに属しているかを判定する。
3. それも不可能な場合には、低次元の散布図に写像するなどの方法でパターン同士の関係を調べる。

の様な多分に主観に頼った研究方法に頼らざるを得ない。

第1の場合は、単一フォントや比較的フォント数の少ない活字認識の際のメッシュ特徴などが該当する。この場合には問題は比較的単純であり、技術的な障壁も低い。第2の場合でも、扱っている特徴が Fourier 変換や Wavelet 変換の様な可換な変換の場合には、あまり大きな問題は無い。実際、Trier らは Zernike モーメントや Fourier descriptor の場合に、そのようなことが可能であることを示している [96]<sup>5</sup>。

問題なのは現在の文字認識装置で用いられているような複雑で動作原理も不明確な様な特徴であり、この様な場合にはパターン集合の線形性などは殆ど調べられていない。第3の場合については、様々な研究が存在するが、低次元に写像する際に大きな情報欠落が起こるために、重要な情報を捕らえているというための保障が必要であるという困難がある。また、重要な情報が視覚化できる程度の低次元空間に集中していない場合も多い。

次章以降では、これらの方法論をベースに、パターン集合の線形性、特徴空間における内積の妥当性などについて、現時点で可能な方法を模索していった軌跡を提示する。この軌跡は、現時点では不完全ではあるが、将来のあるべきパターン認識装置の設計論に繋がっている。

<sup>4</sup>例えば、文字について定義が欠如していることなどから、この問題はかなり一般的なものであることが理解できるであろう。実際、渡辺憲は「カテゴリ」の実在そのものについて疑問を投げかけている [58]。この意味で現時点までのパターン認識の技術は「科学」と「非科学」の境界にあると言えるかもしれない。

<sup>5</sup>とはいえ、彼らの文献では可視化が可能であることを示しているだけで、そこから特徴空間の構造に関する何らかの知見を得ようとしているわけではない。この点で本研究で目標としている物とは本質的に異なる。

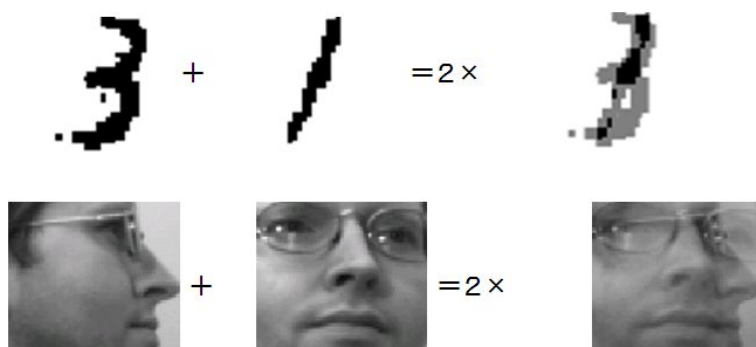


Figure 2.1. 文字画像，顔画像の平均．左辺の文字画像，顔画像はそれぞれ明らかに文字であり顔であるが，右辺の画像は文字画像でも顔画像でもない．このことは濃淡特徴の空間における「文字画像」「顔画像」の集合の線形性の欠如を示唆している．

## Chapter 3

# 手書き文字認識における特徴抽出系開発の歴史

### 3.1 概要

手書き文字には筆記者の個性による様々な変形があるほか、多様な異字体も存在し、これを単一のものとして認識することは困難な問題である。

また、極めて本質的な問題として「文字とは何か」という定義自体が存在しないという事実がある。この問題のため、手書き文字の特徴抽出問題をフォーマルに定義することは難しく、研究者の直感と経験によりノウハウが蓄積されるという形で進められてきた。

歴史的には当初は比較的きれいな手書き文字パターンを対象に Fourier 記述子や Legendre 多項式, Zernike モーメントなどの直行関数展開に代表される数理的な方法が研究されていたが、特に手書き漢字認識の研究が始まると、多様な変形と多数の類似文字の組に対応することが出来ないことが明らかになり次第にヒューリスティックな方法で研究されるようになってきた。

本章ではその歴史を振り返り、現代の OCR で用いられている特徴がいかに複雑化してきたかを眺めることにする。そして、これはおそらく将来顔認識系がたどる道の予測でもある。

なお、本章で用いる認識率の数値は引用する論文で主張された値であり、実験条件が異なるなどの理由から単純に比較できるものではないことを注意されたい。また、本章の内容は [80, 24, 46, 11, 12, 96] の複数の解説から著者が本論文のために必要な歴史をまとめたものであり、一般的な解説として十分なものではないことにも注意されたい。

### 3.2 様々なヒューリスティックな特徴抽出

比較的クラシックな教科書では文字認識の問題を、形状記述の問題として捉え、ストロークなどのプリミティブを抽出し、文法的な記述で文字を認識する方法に多くのページが割かれている。

しかし、このような人工知能的アプローチは、文字とはいかにあるべきかという定義があって初めて成立するものであり、手書き文字認識の研究の初期に捨てられた。

現実の手書き文字には安定したストロークや線分の接続関係などが存在しない

場合が少なくないため、プリミティブを抽出すること自体が困難である上、抽出できたプリミティブも同一カテゴリ内で一定しないためである。

次に現れたのは、印刷漢字認識で有効とされた手法の適用である。印刷漢字認識でもマルチフォントの場合にはある程度の変形があるため、単純な画素のマッチングでは高い認識性能は得られない。また、これらの研究が活発に行われた1970年代後半の計算機では処理量とメモリ量が（現在の感覚からすると）極めて小さかったため、これらの技術はフォントごとの変形を吸収するとともに出来るだけ小さな情報量に圧縮することを要求された。このため研究の目的が交錯しており、現代から当時の問題意識を再現することは難しい。

まず最初に用いられたのは、文字領域を矩形領域に分割し領域内の黒点数を数える「メッシュ特徴」といわれる特徴ベクトルである。これはビットマップのパターンマッチングと同等に思えるが、類似尺度として単純類似度（正規化した内積）を用いると、ゴシックと明朝程度の違いであれば、その違いを吸収することが出来る。

しかし、これで吸収できたのは筆記用具の違いによる変動程度で、筆記者の違いまでは吸収できなかった。

次に用いられたのは Glucksman コードに代表される、文字の複雑さを記述する方法である。Glucksman コードでは、文字中の任意の白点から周囲4方向をスキャンして、どの方向に黒点があるかで、その白点の性質をコード化する。また、上下左右からスキャンして、ストロークの本数を数える方式も用いられた。

これらの方法には明らかに一長一短があるが、どのように形状を記述しているかは割りとはっきりしている。

混乱が発生し始める原因は現在ではよくわからないが、筆者は方向性特徴の有効性の発見と、これらの方法の知見が融合して行く過程ではないかと考えている。

これらの経験的な手法の発展の中で最もドラスティックな変化は方向性特徴の有効性の発見であった。当初、方向性特徴の有効性は文字の周辺分布の利用の形で現れた。統計学における密度関数で行われるのと同様に、文字画像に対して上下左右、斜め方向に積分を行い、その射影を特徴ベクトルとしてもちいる。直感的にわかるように、縦線があれば上方からの周辺分布は縦線のところで大きくなる。左右、斜めについても同様である。

こうした知見から複数の研究者がストロークを方向性マスクのみで分離し、直接的なパターンマッチングを試みたところ、認識性能が劇的に向上するという結果が得られた。

この結果から文字認識研究は、どのように方向性特徴を抽出するかという方向に Drive されていくことになる。

当初のストロークの分離から始まり、単純なエッジの抽出などの研究を経て、方向性特徴は次々に高度化されていった。このときに気をつけて見なくてはならないのは、方向性特徴に関する研究の際にも当初用いられていたメッシュやストロークの複雑さの有効性に関する知見が生かされたことである。つまり、方向性特徴についても、厳密に記述を正確にしていくばかりではなく、ある程度ぼかした方が変動の吸収に有効であること、ストロークの本数などの文字の複雑さを表す情報はやはり有効であることである。

これらの知見が生かされ、80年代に特徴抽出系は結局、萩田らの外郭方向寄与度特徴と鶴岡らの加重方向性ヒストグラム特徴に収束する。これらの研究については以降、小改良が行われ多少の認識率の向上が見られたが、有効性の理由が必ずしも明確ではないため、本質的な進歩はここで終わり、以降、文字認識の研究は特徴空間での統計的学習理論にシフトして行くことになる。

しかし、統計的学習理論にシフトしたことで逆に、その特徴が何かを把握することが極めて重要になった。つまり、どのような学習が行われたかを認識率以

別の方法で確認することが出来なかったからである。次章ではこの問題に対して、特徴ベクトルから文字画像を再構成することで

- これらの複雑な特徴が何を表しているか
- 学習機械が何を学習しているか

を明らかにすることを試みる。



## Chapter 4

# 遺伝的アルゴリズムによる文字認識系の解析

### 4.1 はじめに

文字認識研究の目的は高い認識率を達成することであると集約される．そのため多くの研究者が様々な技術を開発し，部分的には実用に供されている．しかし，これまで実用的な認識率を達成することを目標としてきた文字認識研究では高い認識率が達成された理由や，誤りを引き起こす理由について考察を欠いている面があったことは否定できない．今後，さらなる認識率の向上を目指すためには，文字認識研究の過程における誤読／正読の原因について解析的に調べることが必須と言える．しかしながら，パターンマッチング方式の文字認識アルゴリズムにおける識別系の動作は人間には見えないため，その認識アルゴリズムが何故高性能なのか？なぜ誤読が生ずるのかを議論することはこれまで殆ど行われず，得られた認識率のみを指針に研究が進められてきた．[11][12]

本章では，文字認識研究において，特徴ベクトルを文字画像の形で可視化することで文字識別系の問題点を解析する新しいアプローチを提案する．

まず，従来のパターンマッチング方式における識別系についての問題点を明らかにするために，最近接識別器について考えてみよう．文字認識の特徴抽出系においては，入力された文字パターンから特徴ベクトル  $\vec{v}$  が抽出され，続いて識別系において，何らかの距離尺度  $D(\vec{v}, \vec{\mu}_i)$  が最小となるような参照ベクトル  $\vec{\mu}_i$  を選択し，カテゴリ名  $i$  を認識結果として出力する．このときに識別の基準となる参照ベクトルは，人間には見ることの出来ない特徴空間上に作られている．そのため我々は参照ベクトルに対応する画像を具体的に知ることができない．つまり，識別系が入力パターンをあるカテゴリの文字だと決定した際に，識別系と人間の感覚に乖離があったとしても，それを直接的に検出することは難しい．

このような識別器の性質を正確に把握するためには，何故認識系が正読あるいは誤読したかを人間に理解出来るような形で提示できることが要請される．そのための方法としては，参照ベクトルや識別器が作った識別境界の近傍についての文字データを収集することが考えられる．しかし，識別器を設計する際に，どのような文字パターンが辞書，あるいは境界に位置するかを事前に知ることができない．従って，識別器を設計する度に条件にあったパターンを収集することになり，これはサンプル収集のコストの観点から極めて非効率的である．

本章では，参照ベクトル，識別境界を文字画像の形で可視化することにより文字識別系の問題点を解析するアプローチを提案する．遺伝的アルゴリズム（以下

GA) [69] を用いて特徴空間の任意の点を文字画像の形で可視化することができることを示す。また、このことを通して、拡張外郭方向寄与度特徴における文字パターンの集合がパターン空間をなすことを示し、パターン空間の存在が光学的文字読取装置の動作原理であることを明らかにする。以下、4.2. では、特徴空間からの文字画像再構成問題を定式化し、課題を明確化する。4.3. では、特徴空間から文字画像を再構成するための GA について概説する。4.4. では、GA を用いた文字画像再構成手法の性質を予備実験により示し、4.5. では、可視化によらない従来の解析手法の限界を示す。4.6. では、最尤推定および人工神経回路網の各々の手法で識別系を設計した場合の参照ベクトル、識別境界の可視化を行い、それぞれの識別系の性質について新たな知見が得られることを示す。4.7. では、まとめと今後の課題について述べる。

## 4.2 特徴ベクトルからの文字画像再構成問題

まず、一般の画像再構成問題を考える。雑音がない場合の画像の空間  $x, y$  における原画像  $f(x, y)$  に対する観測過程を

$$v(z) = \int \int A(z; x, y) f(x, y) dx dy \quad (4.1)$$

とする。ここで、 $A(z; x, y)$  は観測過程を表わすオペレーターである。このときの画像再構成問題は、任意のデータ空間  $z$  上の観測データ  $v(z)$  から得られる推定画像  $f_0(x, y)$  と原画像  $f(x, y)$  の誤差、

$$J = \int \int |f(x, y) - f_0(x, y)|^2 dx dy \quad (4.2)$$

を最小化することが目的となる。観測オペレーター  $A(z; x, y)$  についての解析的な表現が得られる場合には、射影フィルタなどの解析的手法が用いられ、Computer Tomography などに代表される様に、多くの成果をあげている [13]。

次に特徴空間からの文字画像再構成問題について考える。通常の画像再構成問題は観測データから原画像を再構成することが目標となるが、識別系の設計者にとって有用な情報は種々の変形を吸収して得られた特徴ベクトルの意味で近い画像であって、必ずしも原画像ではない。ゆえに、この問題において、最小化すべき評価関数は、 $A(z; x, y)$  と  $v(z)$  をそれぞれ特徴抽出オペレーター、特徴ベクトルと看做すと、

$$\begin{aligned} J' &= \int |v(z) - v_0(z)|^2 dz \\ v_0(z) &= \int \int A(z; x, y) f_0(x, y) dx dy, \end{aligned} \quad (4.3)$$

と書ける。ここで、 $v_0(z)$  は解候補である画像  $f_0(x, y)$  から計算された特徴ベクトルである。

文字画像再構成問題をこのような評価関数の最小化問題として考えた場合、

- 一般に特徴抽出系の数学的モデルを得ることは困難であり、場合によっては不可能である。
- 原画像と特徴空間の性質が著しく異なるために式 (4.2) の様な評価関数を用いる意味がない。



- 求められる原画像が一つとは限らない

などの独特の問題がある．

これらの問題を明確にするために例として ePDC 特徴 [45][30] について考えてみよう．ePDC 法における特徴ベクトルは図 4.1 に示す様に，

1. 1 方向の文字枠の各々の画素 (例えば点 P) から内側に向かって走査し，黒点 (外郭点) を探す
2. 黒点 Q に当たったら 8 方向に黒点連結長  $l_k$  を求める．ただし， $k$  は方向のインデックスである (図では上からの走査の例を示す)
3. 黒点連結長から

$$De_k = \frac{l_k}{\sqrt{\sum_{k=1}^8 l_k^2}} \quad (4.4)$$

として，拡張方向寄与度  $De_k$  を計算し，走査を開始した画素に対応するメモリにストアする．

4. 走査方向の文字枠を 8 区間に等分割してその区間内にストアされている方向寄与度の和をとる

という過程で抽出される．

このプロセスは文字枠の周囲 8 方向について，文字枠からの走査で外郭点を 3 回 (外郭深度) まで抽出する．結局，抽出される特徴ベクトルは 8 区間  $\times$  周囲 8 方向  $\times$  8 方向寄与度  $\times$  3 深度 = 1536 次元となる．

こうした特徴抽出器を例えば行列の様な数学的なモデルに置き換えることは困難である．これは，特徴ベクトルは原画像における位置情報を不変にしているため，位置情報が欠落している，特徴ベクトルの各成分は，外郭深度と対応しており，ベクトルは画像の他の成分の存在によって非線形的に変動する，などの理由による．この様に解析的に表現することが難しい問題には人工神経回路網に代表される最急降下法などの方法も適用が困難である．

### 4.3 文字画像再構成アルゴリズム

問題が解析的に表現できないことから，正統的な画像再構成アルゴリズムを導出することはできない．そこで，我々は特徴空間からの文字画像再構成問題をある特徴ベクトルが与えられたとき，同じ特徴ベクトルを発生する画像を探索する問題とみなした．

このような問題に適用可能な探索手法は，Simulated Annealing や GA などいくつか考えられるが，ここでは，実現が容易で，解の探索が比較的高速であることが知られている GA を採用することとした．

図 4.2 に用いる手法の流れ図を示し，以下にそれぞれのプロセスについて解説する．

- (1) 遺伝子コーディング 文字のモデル化という観点からは様々な遺伝子コーディングが考えられるが，今回は，コーディング手法に関する個別の依存性を避けるためにパターンを構成する画素自体を遺伝子とするコーディングを採用した．つまり， $n$  画素で構成される画像を対象にした場合には，各個体は  $n$  ビットで構成されることになる．

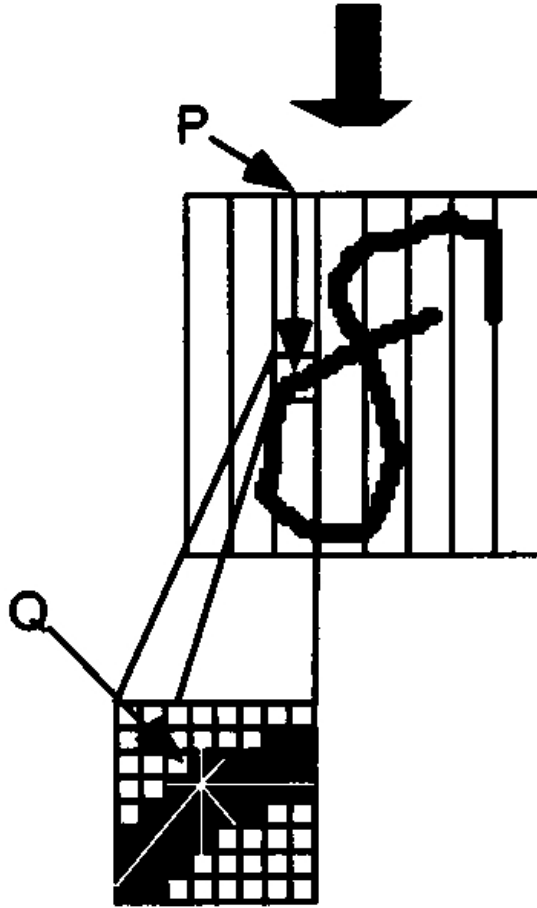


Figure 4.1. ePDC 特徴の抽出過程

- (2) 初期集団生成 通常 GA の適用に当たっては，乱数による初期集団生成が行なわれる．しかしながら，今回の実験に用いた文字パターンは  $64 \times 64$  のマトリックスに格納されており，この時の探索空間は  $64 \times 64$  次元，組み合わせ数としては  $2^{64 \times 64} \approx 10^{1233}$  という莫大なものになる．これだけの探索空間をそのまま扱うことは，計算コストの観点から現実的ではなく，有効な初期値設定法が必要となる．そこで，探索領域を文字領域に限定するために初期集団として文字パターンデータを用いた．さらに，境界の可視化を重点的に行うことを目的として，全対象カテゴリの文字画像を初期個体として用いた．
- (3) 評価 式 (4.4) を最小化することが探索の目的であるから，原画像から得られた特徴ベクトル  $\vec{v}_0$  と個体  $g$  から得られた特徴ベクトル  $\vec{v}_g$ , ( $g = 1, 2, \dots, N$ ) の二乗誤差

$$s_g^2 = |\vec{v}_0 - \vec{v}_g|^2 \quad (4.5)$$

の小さな個体ほど高く評価することが必要である。ただし、後述する適応度比例戦略に基づく個体選択では、適応度の大きな個体ほど選択確率が大きくなる必要があるため、適応度として、シグマ切断

$$F_g = -1.0 \times s_g + (\langle s_g \rangle + 2.0 \times \sigma) \quad (4.6)$$

を用いた。ただし、 $F_g$  の値が 0 以下になる場合には  $F_g = 0$  とした。ここで  $\langle s \rangle$  は当該世代における各個体の二乗誤差の平均値、 $\sigma$  は二乗誤差の標準偏差である。この変換により、二乗誤差の小さな個体ほど大きな適応度を持つことになる。また、標準偏差の 2 倍以上の二乗誤差を持った劣悪な個体は子孫を残せなくなる。このため、最適化が促進される。

- (4) 選択 選択過程では、優秀な個体を確率的に選択する。ここでは、最も広く使われている適応度比例戦略（ルーレット戦略）を用いた。すなわち個体  $g$  の選択確率  $P_g$  として、

$$P_g = \frac{F_g}{\sum_g F_g} \quad (4.7)$$

を与えた。さらに、適応度の比較的低い集団が最優秀個体を破壊してしまうことを避けるために、最優秀個体は無条件で次世代に生存させるエリート保存と呼ばれる手法を併用している。

- (5) 交叉 遺伝子を交叉させる手法は多々提案されている。ここでは高津 [39] らによって提案された 2 次元交叉を採用した。2 次元交叉は画像の局所性を考慮した交叉法で、図 4.3 に示したように、2 個体を選択し、乱数で縦または横に切ることを決め、さらに乱数で切断箇所を決定し、画像を切断して互いに切断箇所を交換するという処理である。
- (6) 突然変異 通常局所最小解からの脱出を企図して用いられる。しかしながら、どのような突然変異手法がこの問題に対して有効か不明なため、今回の実験においては用いていない。<sup>1</sup>

## 4.4 文字画像再構成実験

本節では、前節の文字画像再構成アルゴリズムが、ePDC 特徴から文字画像を再構成できることを示す。また、この実験により、文字画像再構成に必要な世代交代数などのパラメータの検証、初期集団設定手法の妥当性の検証を行う。

実験用のサンプルとしては郵政研の手書き数字データベース IPTP CDROM1 から任意に抽出した 1000 パターンづつの 2 組のデータセットを用いた。それぞれのセットを  $set\#1$ ,  $set\#2$  と名付けた。各データセットは数字 10 カテゴリから構成され、1 カテゴリあたり 100 パターンを持つ。各々のパターンは  $64 \times 64$  の配列に格納し、特徴抽出の前処理として 1 次モーメントによる位置の正規化と 2 次モーメントによる大きさの正規化を施した。

再構成のテストのためのパターンは  $set\#2$  から 0 ~ 9 のそれぞれのカテゴリについて 1 パターンづつ計 10 パターンをランダムに選択し、さらに、初期集団としては  $set\#1$  に含まれる全 1000 パターンを用いることにした。

図 4.4 に原画像と再構成画像を、図 4.5 にカテゴリ「4」の場合について、特徴空間における二乗誤差の平均及び標準偏差の世代発展の例を示す。これらの結果から、世代交代によって次第に二乗誤差を減少すること、100 世代程度の世代

<sup>1</sup>実際、ビット反転などの突然変異手法を適用するとパフォーマンスの低下が見られた

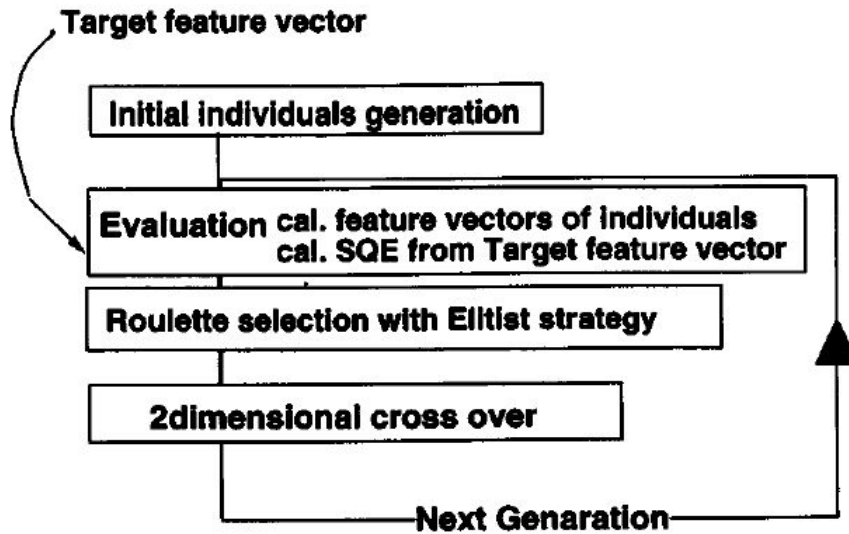


Figure 4.2. GA を用いた手法のフローチャート

交代で文字が再構成されることがわかる。また、この実験の範囲では初期集団の設定についても適切であった。

カテゴリ「4」の場合について画像空間での収束の様子を図 4.6 に示す。図 4.5, 図 4.6 から、大体 40 世代程度でどの個体もほぼ同一画像になり、100 世代では全ての個体が同一になってしまっている。このことから、本手法では 1 回の試行につき一つの画像しか得られないことがわかる [27]。このことについては 7. で議論する。

## 4.5 従来手法による識別系の解析

本節では、従来の主要な解析法である、認識率、confusion matrix、誤読パターン解析の実例を小規模な文字認識実験により示し、解析アプローチの限界を明らかにする。これは 6. 以下で述べる可視化による解析アプローチと比較することが目的である。

実験で用いる認識系のブロック図を図 4.7 に示す。対象は数字 10 カテゴリを考える。特徴抽出系は 1536 次元の ePDC 特徴を用い、識別系はユークリッド距離による最近接規則を用いる。このときの参照ベクトルは最尤推定で作成する。実験用においては set#1 を辞書作成のための学習パターン、set#2 を評価のための未知パターンとする。

表 1. にこの認識系の認識率を、4.2 と 4.3 に学習データ、未知データのそれぞれの confusion matrix を示す。さらに、学習パターンのうち、カテゴリ「0」「2」についての全誤読パターンを図 4.8 に示す。これらのデータから以下の性質を理解することが出来る。

- 認識率は 94%前後である
- かすれ文字や変形の大きな文字は読めない

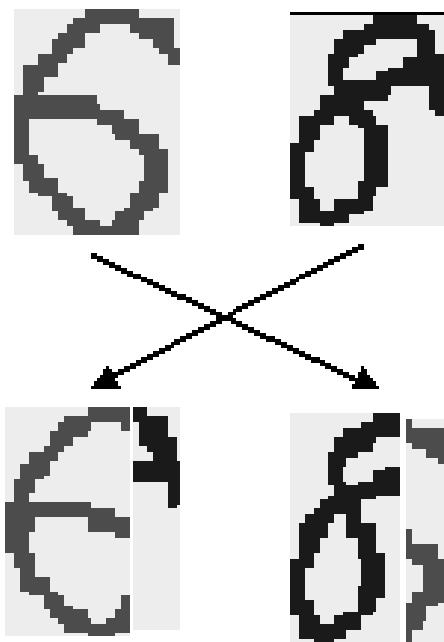


Figure 4.3. 2次元交叉の概念図

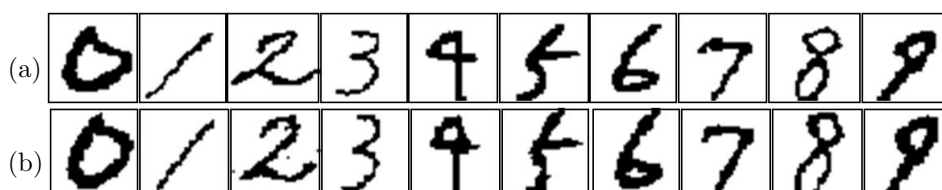


Figure 4.4. 原画像と再構成画像 (a):原画像 (b) 再構成画像

- 「1」、「2」や「9」に誤読する傾向が強いが、学習、未知のそれぞれのパターンについて誤読先カテゴリは必ずしも一致していない

などである。しかしながら、これらの結果からは運用現場におけるOCRの振る舞いを予測するような情報は得られない。さらには誤読パターンのなかにもかなりきれいな文字があり、これらの文字パターンが誤読となった原因を理解することは困難である。

従って、従来のアプローチから、実際にどのように識別系を改善するかの指針を得ることは難しい。

## 4.6 参照ベクトルと識別境界の可視化

前節の文字認識系について、各カテゴリの参照ベクトルと識別境界を文字画像の形で可視化する実験を行った。二つの識別器の設計手法を比較するために統計的手法として最尤推定を、人工神経回路網的な手法として学習ベクトル量子化(以下LVQ)をとりあげる。ただし、大規模実験を避けるために、特に重要と考えら

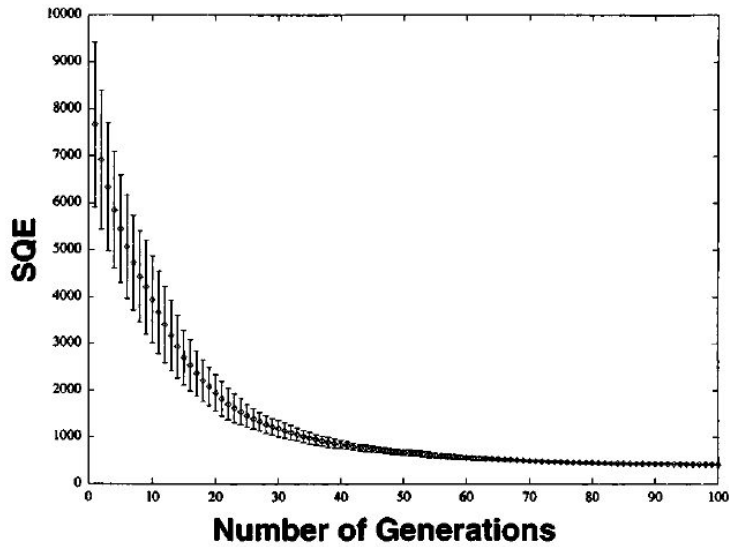


Figure 4.5. 個体の平均二乗誤差の世代発展. 誤差棒は個体の二乗誤差の標準偏差

Table 4.1. 最尤推定で作成した参照ベクトルを用いた場合の認識率

データ	認識率
set #1 (学習)	94.4%
set #2 (未知)	93.9%

れる参照ベクトルとその中点の可視化に焦点をあてた<sup>2</sup>.

#### 4.6.1 最尤推定の場合

図 4.9 は最尤推定で作られた参照ベクトルとそれらの中点の再構成画像である. 世代交代数は予備実験の結果に基づき 100 世代とした. 参照ベクトルの中点  $\vec{l}_{ij}$  は

Table 4.2. 学習パターンに対する confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	91	1	2	1	.	1	3	.	.	1
1	.	99	1	.	.	.	.	.	.	.
2	1	1	94	1	.	.	1	.	.	2
3	.	1	1	95	.	.	.	1	.	2
4	.	.	1	.	95	.	1	1	.	2
5	.	1	.	1	.	97	.	.	1	.
6	1	3	1	.	.	.	95	.	.	.
7	1	.	1	.	.	.	.	94	.	4
8	.	3	.	.	.	1	1	.	93	2
9	.	4	1	.	.	.	.	2	2	91
	0	1	2	3	4	5	6	7	8	9

<sup>2</sup>ユークリッド距離による最近接識別規則では中点は識別境界上にあるがマハラノビス距離の様な距離関数を用いた場合には当然別の点に着目しなければならない.

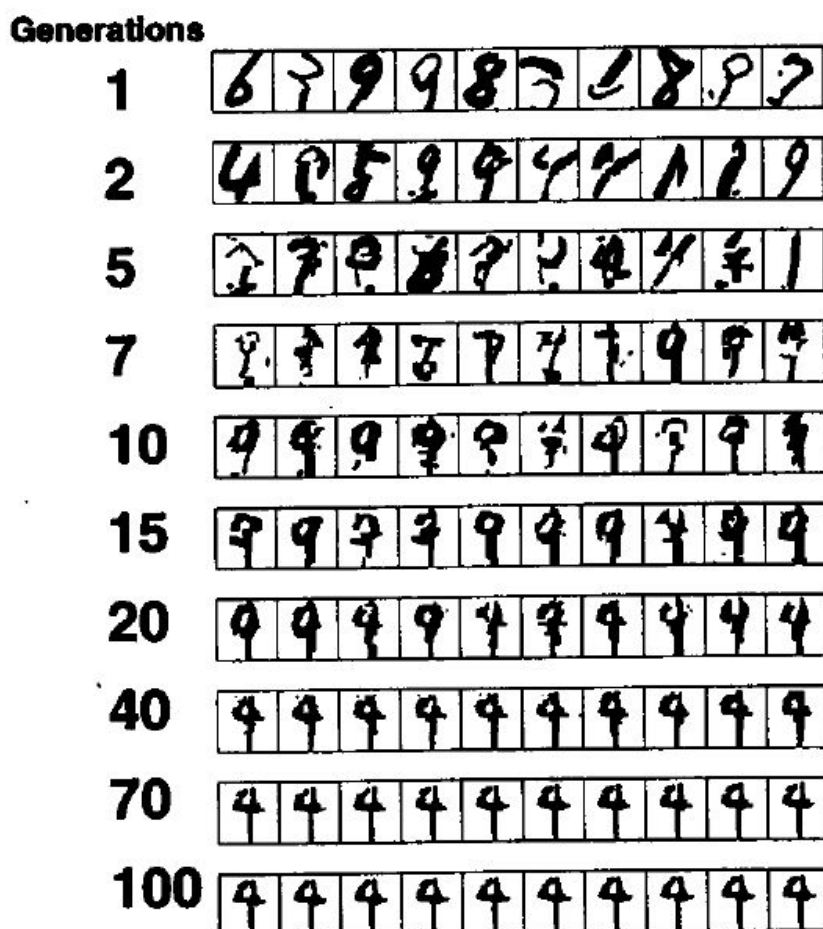


Figure 4.6. パターンの進化の例．同じ行の 10 パターンは当該世代よりランダムに抽出した

$$\vec{I}_{ij} = \frac{\vec{\mu}_i + \vec{\mu}_j}{2}, \quad (4.8)$$

のように定義される．ここで， $\vec{\mu}_i$  and  $\vec{\mu}_j$  はカテゴリ  $i, j$  の参照ベクトル， $\vec{I}_{ij}$  は  $\vec{\mu}_i$  と  $\vec{\mu}_j$  の中点である．図 4.9 中，参照ベクトルに対応する部分ははっきりとそのカテゴリであることを表わす文字になっており，辞書として妥当であると考えられる．一方，中点に対応する画像はいずれか一方のカテゴリに読めてしまうものが多く，識別系が高い認識率を達成するためには不十分であることを示している．

#### 4.6.2 LVQ の場合

次に LVQ を適用した場合の参照ベクトルと識別境界の可視化結果を示す．ここでは LVQ2.1[77] をとりあげる．LVQ の学習ルールは

$$\begin{aligned} \vec{\mu}(t+1) &= \vec{\mu}(t) - \alpha(t)(\vec{x}'(t) - \vec{\mu}(t)) & \text{if } x' \in C \\ \vec{\mu}(t+1) &= \vec{\mu}(t) + \alpha(t)(\vec{x}'(t) - \vec{\mu}(t)) & \text{if } x' \in F, \end{aligned} \quad (4.9)$$

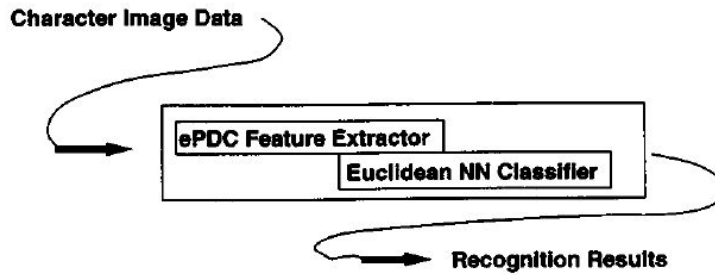


Figure 4.7. 解析に用いる文字認識系の流れ図

Table 4.3. 未知パターンに対する confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	92	1	1	.	.	1	4	.	.	1
1	.	97	1	2	.	.	.	.	.	.
2	.	.	96	3	.	.	1	.	.	.
3	.	.	2	92	.	.	.	2	.	4
4	.	.	2	.	95	.	.	.	.	3
5	.	4	.	2	.	93	1	.	.	.
6	.	4	1	.	.	1	94	.	.	.
7	.	.	2	.	.	.	.	93	.	5
8	.	1	1	1	.	.	1	.	92	4
9	1	.	1	.	.	.	.	2	1	95
	0	1	2	3	4	5	6	7	8	9

の様に与えられる．ここで， $\vec{\mu}$  は参照ベクトル， $C$  は正解カテゴリ， $F$  は誤読カテゴリ， $t$  は学習回数， $\alpha(t)$  は収束を制御する減少関数で  $0 < \alpha(t) < 1$  である．今回の実験では

$$\alpha(t) = \frac{\alpha_0(n-t)}{n},$$

を用いた．ここで  $n$  は全学習回数， $\alpha_0$  は  $0 < \alpha_0 < 1$  なる定数で，我々は事前の検討から  $\alpha_0 = 0.4$  とした．また，カテゴリあたりの参照ベクトル数は 1 個とし，学習の初期値は最尤推定で与えた．表 4.4 は学習後の認識率である．この結果は最尤推定に対する LVQ の優位性を示している．図 4.10 は LVQ の場合の参照ベクトルと中点の再構成画像である．

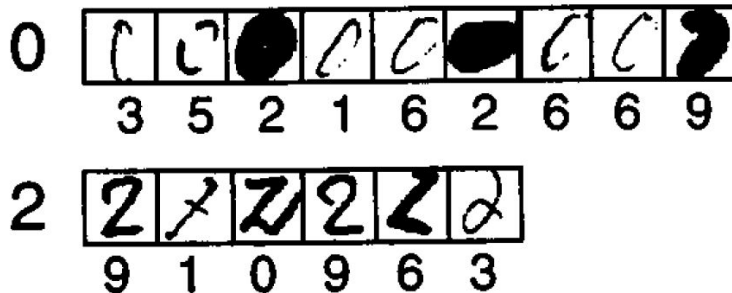


Figure 4.8. 学習パターンの中のカテゴリ「0」「2」の全誤読パターン．パターンの中の数字は認識結果を示す



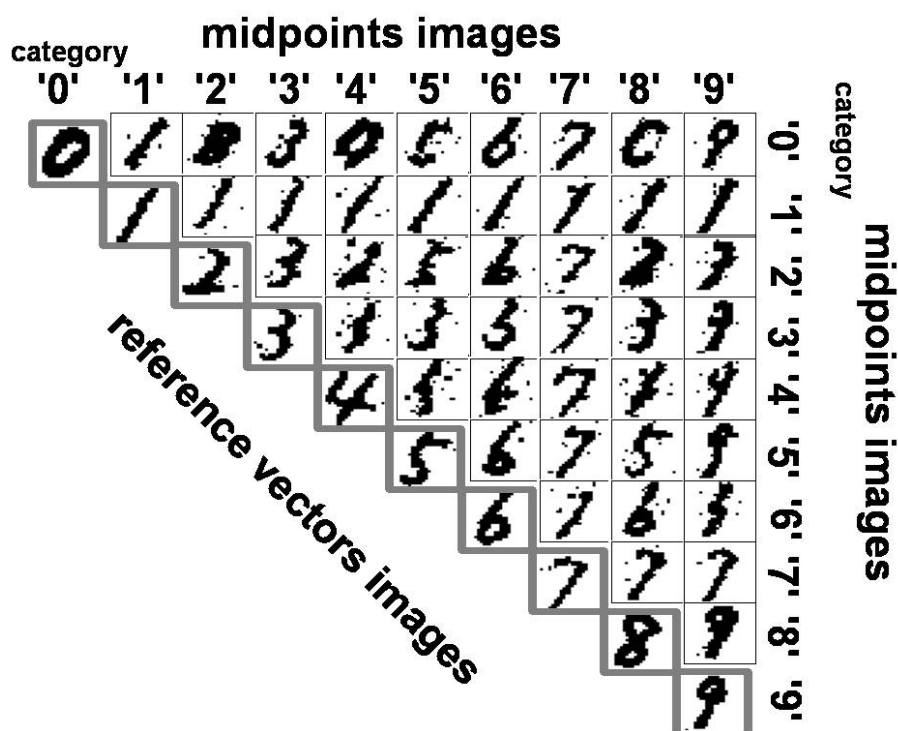


Figure 4.9. 最尤推定で作られた参照ベクトルとその中点の再構成画像。マトリクスの対角要素の画像は参照ベクトルに、他の画像はその中点に対応している。

### 4.6.3 議論

本節では以上の可視化結果から明らかになる、2つの識別系の性質について議論する。

図 4.9 と図 4.10 を比較すると、LVQ の結果の方が最尤推定の場合よりノイズが少ないということがわかる。

これは LVQ の学習が画像空間のごましおノイズの多い領域に対応する特徴空間の領域を避けて進行したことを示唆している。また、図 4.9 の参照ベクトルの画像を、図 4.8 と比較すると、LVQ は最尤推定の誤読パターンの属性をカテゴリーに含む様に学習していることがわかる。例えば、図 4.10 の「0」の参照ベクトルに対応する画像は、図 4.9 の「0」に比して右側が欠けている。これを図 4.8 における「0」の誤読パターンと比較すると LVQ ががすれパターン属性を学習したことがわかる。さらに、これほど顕著ではないが、「2」の参照ベクトルの画像についても、LVQ の学習過程で上部のストロークを学習したらしいことが推測される。

また、中点に対応する画像は、「なぜ、LVQ の方が最尤推定より認識率が高いか」という疑問に対して、一つの可能な解釈を与える。例えば図 4.9 と 4.10 において「1」と「3」の中点の画像を比較してみる。最尤推定の場合の中点はほぼ「1」と読めてしまうのに対して、LVQ の場合の中点は「1」と「3」の中間的

Table 4.4. LVQ による認識率

データ	認識率
set#1 (学習)	100.0%
set#2 (未知)	97.8%

な図形になっている。つまり、LVQ で作られた識別面が、中間的な図形に対応するから認識率が高いという解釈である。もっとも、このことは全てのカテゴリについて言える訳ではなく、例えば図 4.10 の「0」と「4」の midpoint の画像は、文字とはかけはなれた図形になっているが、これがシステムの認識率とどのような関係になるかは明確ではない。

いずれにせよ、これらの結果は

1. 同じカテゴリの文字パターンの平均は同じカテゴリの文字パターンになっている。
2. LVQ の誤り訂正過程では同じカテゴリのパターンの定数倍を加算し、異なるカテゴリのパターンの定数倍を減算することになるが、認識率の向上を示す範囲では参照ベクトルが当該カテゴリの文字であることを逸脱しない

とまとめられる。すなわち、ePDC 特徴の特徴空間上で文字パターンの集合がパターン空間をなすことが確認されたといえる。

## 4.7 まとめと今後の課題

本章では、文字認識アルゴリズムの識別系を解析する新しい方法論として、参照ベクトルと識別境界を可視化するアプローチを提案し、具体的に GA を用いて可視化が実現できることを示した。そして、最尤推定と LVQ により設計した二つの識別系の例について可視化結果を示し、参照ベクトルに対応する文字画像は誤読の可能性のない参照パターンにふさわしいきれいな文字になること、識別面付近には人間にも判読の困難な中間的な文字画像が存在していることを明らかにした。

これらのことは、ePDC 特徴の空間では文字パターンの集合がパターン空間を構成していることを示している。すなわち、ePDC 特長に関する限り、パターン認識装置の動作条件が満たされているために良好に動作することが示された。

また、参照ベクトルの可視化画像と識別結果の誤読パターンの比較によって、誤読原因の把握を容易にすることを示した。カテゴリ間境界の可視化により、識別系が作った識別面と真の識別面のずれの検出が可能になることを示した。

これらの解析結果からさらに新しい識別手法の開発も可能であろう。例えば、識別境界近傍にはどのカテゴリともつかない中間的な図形が存在するという事実は、従来の識別手法がいずれかのカテゴリに強制的に識別していたこと自体の問題を示しているように考えられる。これを考慮した識別手法、リジェクト手法の開発が今後の文字認識アルゴリズムの性能向上のために必要であろう。

しかしながら、残された課題も多い、一つの特徴ベクトルに対応する図形が複数ある場合が考えられるが、4.4. で示した結果は、そのうちの一つが解として選られ、複数の解を同時に得ることができないことを示している。すなわち、GA を用いた今回の手法は、文字認識アルゴリズムの性質の把握のために有効であることを示したが、これで十分ではない。今後どのような文字が特徴空間の同一の点に射像されるのか？そのなかには他のカテゴリの文字が入っていないのか？と

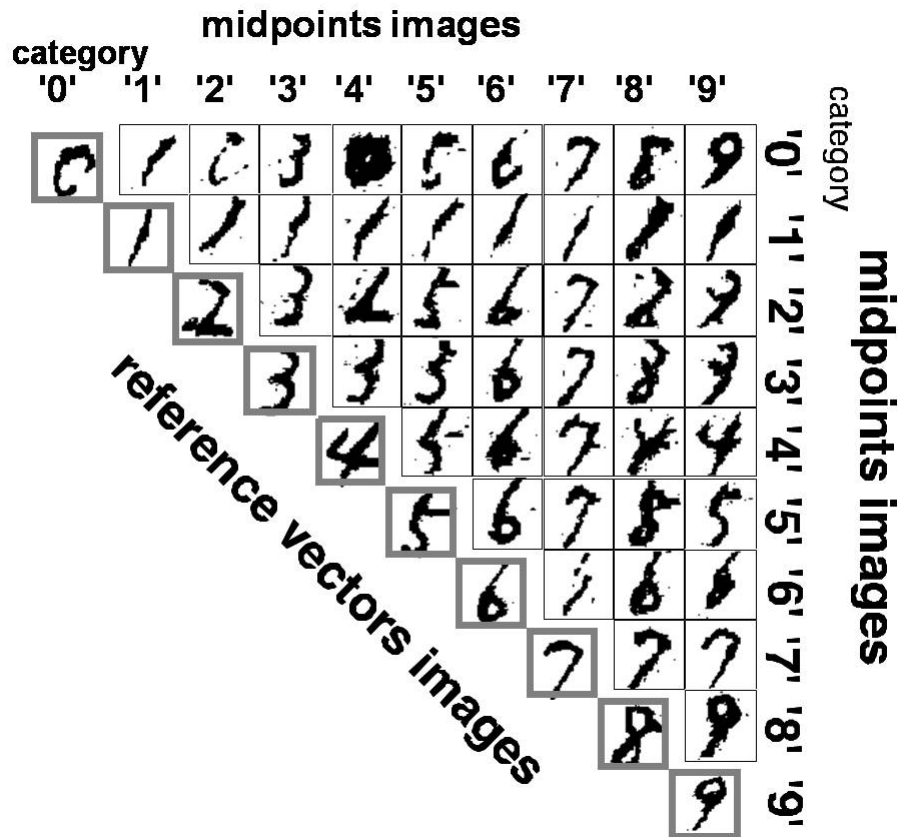


Figure 4.10. LVQ で作られた参照ベクトルとその中点の再構成画像

いう問題を明らかにするためには、多数の解を同時に探索するアルゴリズムの開発が必要である。

#### 4.8 その後の発展

本研究は当初の発表後、殆ど省みられなかったが、近年になりその重要性が再評価されたためか改良研究 [75] や、別の方向の可視化研究 [18] が現れ始めた。

いずれにせよ筆者の提案したアルゴリズムは、一定の特徴抽出系にしか有効なものではなく、認識系の高度化のためにさらに今後の研究が必要であると考えられる。



## Chapter 5

# 質問学習による文字認識辞書の学習

### 5.1 はじめに

文字認識アルゴリズムの研究は初期の特徴抽出系の研究に始まり，識別関数の高度化のフェーズを経て，識別の観点からの学習理論に関する研究にその焦点を移しつつある．文字認識技術において，学習の理論はいかにして与えられた学習パターンから汎化した学習結果，すなわち未知パターンに対して高い認識率を得るかが焦点となっている．しかしながら，有限の学習パターンを用いた学習ではその有限性に起因するバイアスを避けることは原理的に不可能であり，大きな課題になっていた．本章ではこの問題に対し，質問学習を文字認識技術に適用することを提案する．

一般的に「例」のみから成る学習に比して教師に対する「質問」や現実の事象に対する「試行」を行う積極的な学習が有効であることはよく知られている．学習機械に積極的な「質問」や「試行」をさせることで，効率的な学習を目指す枠組みは「能動学習」と呼ばれ計算論的学習理論や人工神経回路網の分野で活発に研究されている [37]．

ロボットの行動学習の様に「試行」に対する結果の評価が直接的に得られる場合には Q-Learning に代表される強化学習が有効であることが知られている．一方で，パターン認識の様に解かれるべき問題に明確な定義が与えられず結果の評価が人間に委ねられる問題に対しては，学習機械の発する質問に教師が逐次的に回答する「質問学習」の枠組みが特に有効であると考えられる．

しかしながら，実用的な文字認識機械の学習にこの学習方法が適用された例はない．その理由は主として実用的な文字認識機械で用いられる特徴抽出器が高度に抽象的で，学習機械が質問を発したとしても，その質問を教師（すなわち人間）に理解できるように翻訳する手法が存在しなかったためである．文字認識の場合にはこの翻訳は特徴空間から文字画像に変換する操作となる．

質問学習を文字認識機械の学習に適用した例としては Kudoh[78] が挙げられるが，彼らの研究では質問の条件にあった文字パターンを乱数的に生成する技術が基礎となっているために識別に関する最適性の観点から問題を残していた．

前章では遺伝的アルゴリズム（以下 GA） [69] を用いて，特徴ベクトルで表現される文字を可読な文字パターンに変換する研究を行ってきた [29]．この手法は，質問学習における翻訳機として使用することが出来る．そこで，我々は，本章において実用的な文字認識系に対する質問学習アルゴリズムを提案する．

質問学習には学習の高速化、学習の高精度化、サンプル収集コストの低減、などの利点が指摘されている。これらの目的のためには人工生成文字を利用して文字認識辞書を設計するという方法も考えられ、すでに、[19], [3] などの優れた研究があるが、本研究では識別器が学習の必要のある部分を指摘するという意味で本質的に異なっており、少数の学習パターンの追加で有為な認識率の向上を可能とする。

本章においては、GA を用いた質問学習を実用的な文字認識系に適用するための枠組みを示し、実験によりその有効性を確認する。まず、5.2. では質問学習を実用的な文字認識系に適用するための手法を具体的に述べる。5.3. では手書き数字認識実験により手法の有効性を検証し、5.4. では実験結果に対する考察を行い、5.5. ではまとめと今後の課題について述べる。

## 5.2 文字認識系のための質問学習

### 5.2.1 質問学習に関する従来の研究

質問学習については、最近 10 年間に膨大な研究が行われている。質問学習の基本的な理念は、学習機械が現在までに与えられた情報から結果の不鮮明な部分を抽出し教師に質問として投げかけ、その回答を用いて再学習を行うことである。

モデル近似問題の場合は実験計画法をベースにした技術が使われることが多く、現在の学習器械の状態から情報量最大化などの観点で次に採取すべきデータを指定し、そのデータと回答のラベルを得て再学習を行う。識別問題の場合には識別困難領域を質問として採用する方式が用いられることが多い [74][60]。ユニークなものとしては複数の学習器械の結果が矛盾する領域を質問とする研究もある [93]。

いずれの場合にも 1) 学習の高速化、2) 学習の高精度化、3) サンプル収集コストの低減、のメリットが報告されている。特に [60] は限定された状況ながら多項式時間で学習を終了できることが理論的に示しており興味深い。[74] においても、ランダムな例からの学習の場合に比べ、1/5 程度の回数の質問学習で良好な結果が得られることを実験的に示している。

### 5.2.2 文字認識系への適用

前節で述べた様に質問学習には多くのメリットがあるが、パターン認識問題に適用された例は多くない、Sung[94] らは、画像からの顔の切り出しに実験計画法型の質問学習を適用し良好な結果を得ている。これは顔切り出し問題が画像そのもののマッチングで解かれる問題であることに因るところが大きい。それに対して、現在実用的に用いられている文字認識アルゴリズムのほとんどはヒューリスティックに設計された特徴抽出器と最近接識別器によって構成されている。例えば著者らは拡張外郭方向寄与度特徴抽出器ユークリッド距離による最近接識別器を用いた文字認識アルゴリズムを用いている。

最近接識別器において識別困難な領域は、認識対象となるパターンを  $\vec{x}$ 、近接する参照ベクトルを  $\vec{\mu}_i, \vec{\mu}_j$  としたとき、 $|\vec{x} - \vec{\mu}_i| - |\vec{x} - \vec{\mu}_j| < \epsilon$  となる部分である。ただし、 $\epsilon$  は微量である。本報告ではこの様な領域を「危険領域」と呼ぶことにする。最近接識別器の場合の質問学習は危険領域を抽出し、質問とすることで達成できる。しかしながら、このような枠組みを文字認識機械に適用しようとした場合には、学習機械が発する質問が特徴ベクトルとなるために質問が人間に理解できないという問題がある。

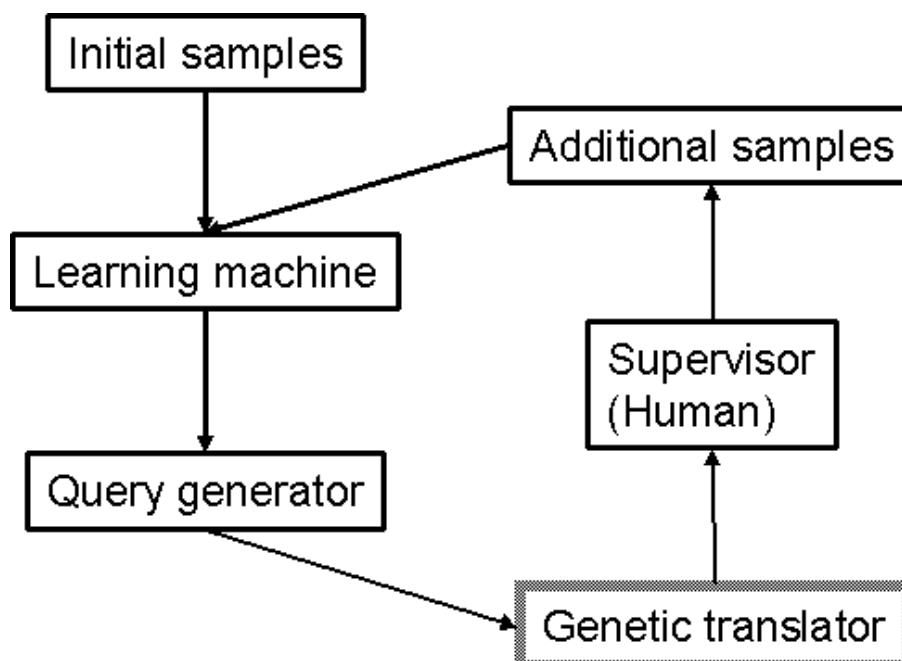


Figure 5.1. 提案する質問学習機械の構成

我々はこれまでに GA を用いて特徴ベクトルから文字パターンを再構成することが可能であること、特に参照ベクトルや識別境界を可視化出来ることを示した。この手法を用いることで、任意の特徴空間において学習機械が発した質問を人間が可読な形に「翻訳」することが出来る。これにより、あらゆる質問学習アルゴリズムは文字認識系に適用可能になる。

従って、我々が提案する質問学習機械は図 5.1 に示す様な、学習機械、質問発生機械、質問翻訳機械、教師へのインターフェースによって構成される。以下、学習、質問翻訳についてアルゴリズムを概説する。

#### 学習アルゴリズム

文字認識装置のための学習機械は統計的学習法を始めとして多岐に渡るが、識別問題に関しては誤り訂正アルゴリズムが有効である。ここでは、我々が用いている文字認識アルゴリズムがユークリッド距離による最近接識別器であることに鑑み、学習ベクトル量子化 (LVQ : Learning Vector Quantization) [77] を採用する。LVQ はユークリッド距離を用いた最近接識別のための学習規則であり、文字認識に関する有効性が示されている。本研究ではさらに改良が施された佐藤による定式化を採用する [35]。

佐藤の定式化に基づく LVQ は以下の手順により実行される。 $\vec{x}$  をサンプルパターンの座標、 $\vec{\mu}_i$  を参照ベクトルの座標、ただし、 $i$  はカテゴリのインデックス、 $\mathcal{C}$  を正解サンプル集合、 $\mathcal{F}$  を誤りサンプル集合、としたときに、任意に選択したサンプルに対して、ユークリッド距離に基づく認識処理を行い、

$$\vec{\mu}_i(t+1) = \vec{\mu}_i(t) - \alpha(t)(\vec{x}'(t) - \vec{\mu}_i(t)) \text{ if } \vec{x}' \in \mathcal{C}$$

$$\vec{\mu}_i(t+1) = \vec{\mu}_i(t) + \alpha(t)(\vec{x}'(t) - \vec{\mu}_i(t)) \text{ if } \vec{x}' \in \mathcal{F}$$

という規則で参照ベクトルを移動する． $\alpha$  は学習係数で  $0 < \alpha(t) < 1$  である．一般的には

$$\alpha(t) = \frac{\alpha_0(n-t)}{n},$$

の様な形で減衰させることで収束を制御する．

#### 文字画像再構成アルゴリズム

本節では我々が用いている GA を用いた文字画像再構成アルゴリズムについて概略を述べる．その基本は，遺伝子として与えられた画像が，与えられた特徴ベクトルを発生させるような画像であるかを評価しながら，交叉，淘汰を続けることにより与えられた特徴ベクトルに対応する画像を探索するものである．

遺伝子コーディング 文字パターンを構成する画素自体を遺伝子と見做す．このため，各個体は  $64 \times 64 = 4096$  ビットで表現されることになる．

#### 初期集団生成 同種カテゴリの文字パターン

評価 原画像から得られた特徴ベクトル  $\vec{v}_o$  と個体  $g$  から得られた特徴ベクトル  $\vec{v}_g$  ( $g = 1, 2, \dots, N$ ) の二乗誤差

$$s_g^2 = |\vec{v}_o - \vec{v}_g|^2 \quad (5.1)$$

の小さな個体ほど高く評価することが必要である．ただし，後述する適応度比例戦略に基づく個体選択では，適応度の大きな個体ほど選択確率が大きくなる必要があるため，適応度として，シグマ切断

$$F_g = -1.0 \times s_g + (\langle s_g \rangle + 2.0 \times \sigma) \quad (5.2)$$

を用いた．ただし， $F_g$  の値が 0 以下になる場合には  $F_g = 0$  とした．ここで  $\langle s \rangle$  はここで  $\langle s \rangle$  は当該世代における各個体の自乗誤差の平均値  $\sigma$  は自乗誤差の標準偏差である．

#### 選択 適応度比例戦略 + エリート保存

#### 交叉 2次元交叉

突然変異 画像の局所性を重視した突然変異手法である Affine mutation を用いた<sup>1</sup>．この手法は大規模な実験によって有効性を確かめられている．

最終的に，我々が提案するアルゴリズムは以下の様に構成される．

- (0) 初期設定 初期参照ベクトルを用意する．
- (1) 質問の発生 危険領域の特徴ベクトルを選択し質問を発生する．
- (2) 質問の翻訳 選択された特徴ベクトルを GA で文字パターンに再構成する．
- (3) 回答 人間などの教師が読んだ結果を再構成したパターンにラベリングする
- (4) 再学習 解答を追加して LVQ による学習を行う

これら (1)-(4) までのステップの繰り返しを行う．

<sup>1</sup>付録参照



sample	Recognition Rate
train.	100.0%
test	97.2%

Table 5.1. 対象とした文字認識系の認識率

### 5.3 手書数字認識実験

提案手法の有効性を示すために認識実験を行った。ただし、提案する枠組みについて、現象を詳細に観察可能にするために可能な限り単純なアルゴリズムとして実装した。

実験は数字 10 カテゴリを対象とし、データベース IPTP CDROM1 より、学習セットとして 2000 パターン / カテゴリ、テストセットとして約 15000 パターンを選択した。初期参照ベクトルは 2000 パターン / カテゴリの学習パターンについて LVQ2.1 で学習した結果得られた参照ベクトルを用いた。このとき、学習係数  $\alpha$  の初期値は予備実験の結果から  $\alpha = 0.5$  とし、収束条件は誤りがなくなるか、全学習パターンについての学習の繰り返し数が 30 回となることとした。参照ベクトルは 1 カテゴリあたり 1 個とし、LVQ の初期値は学習セットのカテゴリごとの平均値を用いた。

認識実験の結果を表 5.1 に示す。この結果は用いている文字認識アルゴリズムが実用的な性能であることを示している。

この系に提案アルゴリズムを適用した。以下、実験のための設定について解説する。

**初期設定** 上記の学習結果で得られた参照ベクトルを初期参照ベクトルとして用いる。

**質問の発生** 質問は危険領域全体をカバーするように発するのが自然であるが、拡張外郭方向寄与度特徴の空間が 1536 次元に及び、さらに GA による画像再構成の処理量が大きいため質問は識別上最も問題となる

$$D = \min |\vec{x} - \vec{\mu}_i| - |\vec{x} - \vec{\mu}_j|$$

となる点、つまり、中点近傍に限定した。従って、10 カテゴリの場合には一回の学習ステップで中点の数の 45 質問が発せられることになる。なお、Baum, Hwang のアルゴリズムでは識別面の両面に質問を発生させることを要求するが、これは翻訳の際に GA を用いる中でバリエーションを持ったパターンを発生させることで実現する。

**質問の翻訳** 質問となる 2 つの参照ベクトルの中点を再構成のための標的特徴点として定め、GA による再構成を行う。ただし、再構成には多大な計算時間を要するため以下の様な簡便法で一回の再構成で複数種類のパターンを採取する。再構成の際、GA は全体として単一のパターンに収束するようにふるまうため、最適化がある程度すすみ、かつ収束が終了しない程度の世代交代を経たあとの個体集団から 2 つの参照ベクトルのそれぞれに近いパターンを 5 つずつ選択する。この方法により、一回の翻訳作業で識別面の両面からサンプルを採取することが可能になると期待される。従って、10 カテゴリの場合には一回の学習ステップで (中点の数  $\times$  採用したパターン数) が質問となり、リジェクトパターンを指定しなかった場合には 1 回につき 450 パターンが再学習に用いられることになる。なお、世代交代数は予備実験により 30 世代に決定した。

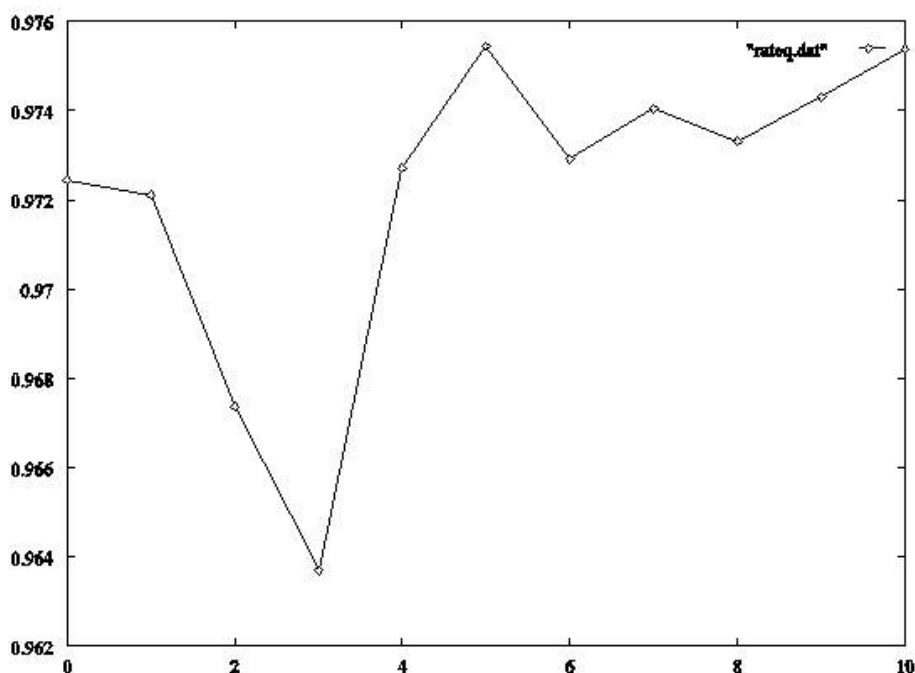


Figure 5.2. 質問学習に伴う認識率の推移

質問の回答 翻訳された質問に対しての解答は一人の人間があたえた。ただし、判読不能パターンは学習パターンから排除することにした。

再学習 回答をラベルされたサンプルを初期学習セットに追加し、再びLVQによる学習を行う。再学習の際には、以前までのステップで得られたパターン及び初期学習パターンのすべてを学習の対象とし、LVQの初期参照ベクトルも初期参照ベクトルの作成時と同様に初期学習データセットのカテゴリごとの平均値を用いた。

実験では以上のステップの繰り返しは10回行った。実験結果を図5.2に示す。グラフの横軸は上記の繰り返しステップ数、縦軸は認識率を示す。

## 5.4 考察

提案手法による認識率の向上はわずかであるが、学習の後半では確実な認識率の向上を示し、有為であると考えられる。実際10回の繰り返しまでに実際に学習に寄与しているパターンは2503パターンすなわち約250パターン/カテゴリであり、ランダムに与えた場合に比して効果的であると考えられる。

提案手法は学習の初期には著しい認識率の低下を招くがこれはさほど意外ではない。図5.3に示すように、中点の様に発生確率の低い点はカテゴリ中で孤立した存在となり、そのために識別面に不自然な傾斜を引き起こすと考えられるためである。

この不自然な傾斜は学習が進み境界領域に多数のサンプルが与えられると徐々に解消し認識率は向上していく事になると考えられる。

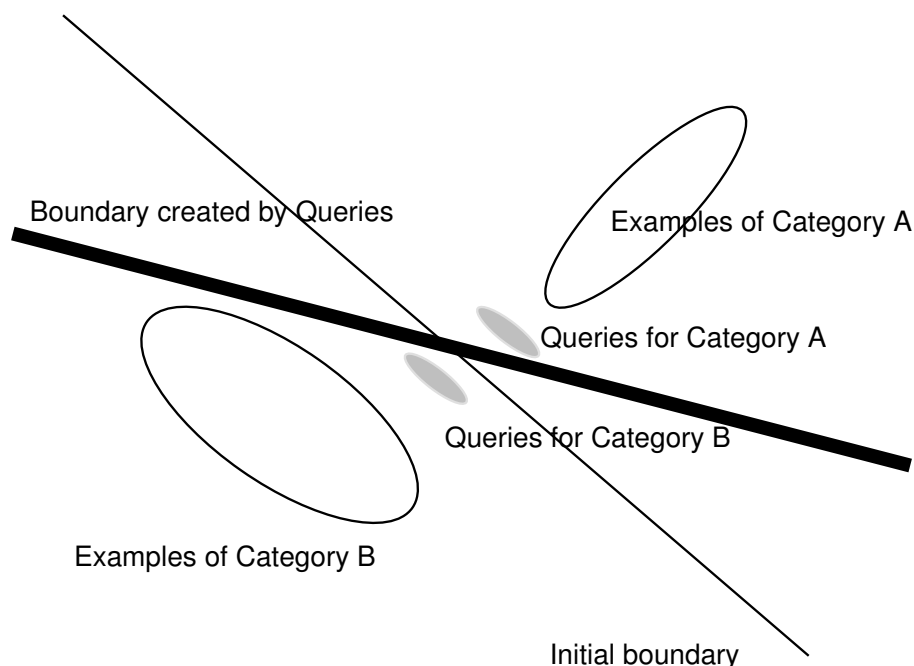


Figure 5.3. 初期の質問に因って作られる不自然な境界の例

しかし一方で、質問として出力されるパターンは繰り返し毎に判読不能パターンを多く含むようになる傾向が見られる。判読不能パターンには二つもしくは3つの文字種の特徴が混ざってしまいどちらとも判別が不可能なものと、全く数字として認識が不可能なものの2種類が現れる。識別の観点からはどちらのパターンが現れることが望ましいかは判然としないため考察し難いが、繰り返し回数が増えると、文字として認識できないランダムノイズに近いパターンが現れる傾向があり、結果としてそれ以上の学習が不可能になってしまう。

この現象の原因としては、LVQ2.1が識別境界付近を重視して学習するアルゴリズムであるために、質問として与えられたパターンに過度に反応し、必ずしも文字領域ではない部分に中点を設定してしまうためと考えられる。この問題の解決のためにはLVQ2.1よりも収束性にすぐれた学習アルゴリズムの採用が必要であると考えられる。

また、中点近傍のみを採用した弊害も散見される。例えば「0」と「3」の境界に「8」が現れるという様な現象が数多く観察される。つまり、質問のフィードバックが必ずしも学習機械の性能向上に繋がらない場合が有り得る。すなわち、単に距離値が近いという以上の質問選択アルゴリズムの必要性を示唆していると考えられる。

## 5.5 まとめと今後の課題

本章においては実用的なOCRのための質問学習の枠組みを提案し、数字認識実験でその可能性を示した。今回は処理量の削減のため中点近傍のみを用いたが、それ以外の点を用いることで学習効率上がることは十分考えられる。また、LVQ以外の学習方法の場合についても実験的な検討が必要だと考えられる。



## Chapter 6

# 物体認識と非線形性

本章では, 物体認識問題において発生する画像分布の非線形性について考察し, その解決手段をレビューする. 本章の考察は, 次章で著者が提案する新しいパターン認識手法, 核非線形相互部分空間法の基礎になる.

### 6.1 物体認識問題と非線形性

本節で対象とする物体は, 剛体であり, どの方向から見てもほぼ凸であり, 滑らかな色調の変化をするようなテクスチャで覆われていることを仮定する<sup>1</sup>. また, 物体画像中において位置と大きさがなんらかの方法で正規化されていることを仮定する.

特徴抽出系としては濃淡特徴を考える. このことは物体認識問題を特殊化してしまうことには当たらない. と言うのは 3 次元再構成を主として考えるいわゆる Computational Vision の技術以外の画像からの特徴抽出系は Fourier 変換や K-L 変換, Gabor, Haar などの wavelet 変換が主であり, この様な線形変換では画像分布の非線形性を本質的に除去することは不可能な為である.

物体認識問題において発生する画像の非線形分布の原因は主として, 物体の回転である.

物体の回転が例えば 1-2 度程度の微小な場合, 物体画像は本質的類似した物になり, その分布は線形になる.

しかし, 物体が数十度のオーダーで回転した場合, 物体画像群は多くの場合非線形な分布をとる.

これらの理由を定性的に説明するのが図 6.1 である. 図 2.1 を見れば明らかな通り, 回転が小さな時は画像変動は小さく, ほぼ線形な分布をするものと考えられる. 一方で回転が大きな時は回転前後の画像を平均したものと回転途中の画像が大きく異なっている. つまり, 特徴空間中の画像群の軌跡は曲線となる<sup>2</sup>.

この現象は本質的には濃淡特徴をそのまま用いていることに起因する. しかし, 一方でこの問題の本質的な解決の為には 3 次元再構成, つまり Computational Vision の技術が必要になるが, 現時点ではその技術自体が認識の前処理として用いるために十分な精度を持っているとは言いがたい. さらに, 文字認識で用いら

<sup>1</sup>最後の仮定は恣意的であるように思えるかもしれないがそうではない, 工学的に有用である人間の顔, 人物自体, 車両などは殆どこの条件を満たしているし, 後述するように非線形変動の主因が物体の回転である為に, 単色の球体の様な場合には非線形分布が発生せず, 考察の対象になりえないからである.

<sup>2</sup>容易に想像できる通り, この様な非線形性は, 画像の併進, 画像内回転, カメラの特性などによっても発生するがここでは議論の単純化のために回転のみを扱うことにする

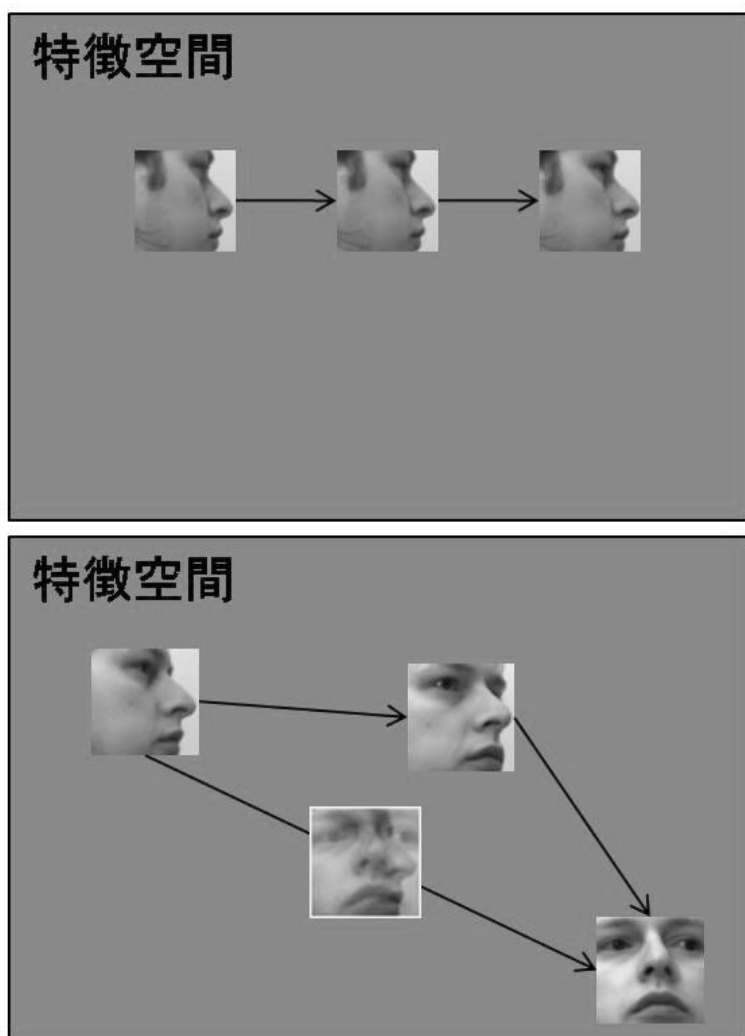


Figure 6.1. 画像の回転に伴う非線形性の原因. 右の様に 1-2 度程度の変動であれば, 特徴空間中の画像の変化は小さく, ほぼ同じ形態になる為, ほぼ線形の分布になると考えられる. 一方, 左の様に数十度の回転がある場合, 回転前後の中間は回転途中の画像と一致しない. すなわち, 回転が大きいときは線形な変動をしないことが分かる. なお, この画像データは UMIST (University Manchester Institute of Science and Technology) データベースを用いた.

れた特徴抽出系が画像の記述方法を変えることにより非線形性を消去することに成功しているのと同様な技術でこの問題を解決することが出来るとは思えない。

次節では、この非線形問題に対するこれまでの研究を概観し、その問題点を明確化する。

## 6.2 これまでの研究とその問題点

画像の姿勢変動に伴う非線形性とそれを意識した最初の認識アルゴリズムは村瀬らによるパラメトリック固有空間法である [54]。村瀬らは認識対象となる画像群を主成分分析で圧縮した上で、画像群の軌跡を Cubic spline で近似することで、物体認識と姿勢推定の双方を同時に可能にするアルゴリズムを提案した。

パラメトリック固有空間法は発表後、爆発的なブームを引き起こし、大量の改良研究を生み出した。

しかし、どの研究もパラメトリック固有空間法の本質的な問題点を解決するものではなかった。パラメトリック固有空間法の問題は変動画像が作る非線形多様体を Spline 関数、すなわち回帰的に記述することにあった。

つまり、パラメトリック固有空間法での学習の為には物体であることのラベルと姿勢のラベルの双方が無ければ学習が不可能であるということにある。

こうした問題に対して積極的な解答を出したのが、石山、坂本による研究である。石山らは顔の 3 次元モデルを用い、学習時に顔のテクスチャを 3 次元モデルに貼り付けることで姿勢、照明変動に対してロバストなアルゴリズムを開発した。しかし一方で、この方法は人間の顔が基本的に類似した形状をしていることを利用しており、顔の識別の様な特別な問題にしか適用できないと言う問題がある [10]。

一方の解決法は前田、村瀬によって提案された核非線形部分空間法である。この方法では核による非線形変換で非線形分布を線形に変換することを可能にしているが、残念なことに物体認識系への応用可能性を示してはいない [51]。

もう一つの解決法は、非線形分布を包摂する線形部分空間で近似するアプローチである。山口らは複数の入力画像を仮定し、これを主成分分析して特徴を抽出したあと辞書である部分空間との正準角を類似度として用いることで認識を可能にする相互部分空間法による物体認識アルゴリズムを提案した。[56]。この方法ではカテゴリの増加とともに認識精度の低下が懸念されるが、入力画像群を主成分分析することでこの問題をある程度解決している。しかし、やはりカテゴリが増加した場合には認識率が低下するものと考えられる。

これらの研究が示唆しているのは、

- 非線形な軌跡を近似するための方法が重要であること
- 特殊な場合には比較的簡単な解法があるが、一般的な場合には Computational vision の技術が必要になるなど、簡単な解法は存在しないこと

である。

次章では、このことを考慮した上での新しい物体認識アルゴリズム、核非線形相互部分空間法を提案し、実験的に有効性を示すとともに、その動作が物体画像の線形な空間への写像になっていることを示す。





## Chapter 7

# 核非線形相互部分空間法による 物体認識

### 7.1 はじめに

本章において、我々は新しい物体認識アルゴリズム、核非線形相互部分空間法を提案する。

画像からの物体認識は、通常、1枚の静止画からの画像認識問題として議論されている。しかしながら、物体認識系の現実的な実験、実装環境では画像入力にはビデオカメラが用いられ、連続的な画像の採取が簡便に行われるばかりではなく、計算機の高速化により、簡単な画像処理であればフレームレート以下の時間での処理も可能となっている。

直感的には複数の入力画像を認識に用いることで対象に関する情報を増やすことが出来れば、認識精度の向上を図ることが出来るはずである。しかしながら、これまで多くの研究では動画像としての性質の利用は、背景差分による物体検出の高速化などの補助的な利用が殆どであり、ジェスチャ認識など動作を認識する場合を除き、動画像としての性質を積極的に利用することは殆ど行われていなかった。

動画像としての性質を認識に利用するには主として二つのアプローチが考えられる。一つは剛体運動のモデル、光学モデルなど対象の事前情報を組織的に利用して、入力画像から対象物体のモデルを作成するものである。今一つのアプローチは、入力される画像系列を単なる複数の入力信号と考え、それらに対する統計モデルに対して、従来の統計的パターン認識の手法を適用することである。

前者のアプローチは、言わば Physics-based vision のモデルをパターン認識の前処理に用いることに相当し、理論、応用の両面から興味深い。現時点での Physics-based vision の研究の多くは必ずしも認識のために研究されているわけではないことから、適用条件の観点から実用的な認識アルゴリズムを構成するのは困難である。例えば、因子分解法 [95] を用いて 3次元再構成を行い、方向不変な物体認識系を構成することなどが考えられるが、顔などの対応点問題を解く事が困難な対象には適用できない。

後者のアプローチは、単純に中心極限定理から雑音除去の効果が期待できるばかりではなく、特徴空間で同一の統計モデル同士を比較するアプローチと考えられ、統計的パターン認識の新しい研究対象として興味深い。こうしたアプローチの先駆的な研究例としては、前田 [52] による相互部分空間法の提案がある。相互部分空間法は入力信号に対しても主成分分析を行い、辞書として登録された部分空間と、入力信号から得られた部分空間の角度を類似度として用いることにより

認識処理を行う。山口は [56] は相互部分空間法を顔の認識に応用し、多数決、単純平均と比較しても高性能な認識系を構成し得ることを実験的に示した。

しかしながら、基本的に相互部分空間法の部分空間構成法は Watanabe の部分空間法 [99] で用いられた主成分分析の方法であり、部分空間法の問題点をそのまま継承している。すなわち、認識対象の分布が S 字状等の非線形構造を持っている場合にはそれに起因する誤認識を避けることが出来ない。

物体認識問題において、物体の姿勢変動に起因する強い非線形構造が存在することは多くの研究者に示唆されてきた。例えば村瀬ら [54] は、画像からの物体の姿勢推定問題を物体の角度をパラメータとした非線形回帰の問題として扱っている。

こうした非線形性を考慮した場合、相互部分空間法の高精度化のために統計学や神経回路網の分野で研究されている非線形主成分分析法を適用することは自然な発想である。しかし、相互部分空間法では、部分空間の基底同士の内積を計算することが必要となるのに対し、多くの非線形主成分分析法では多くの場合、部分空間の基底が明確に定義されておらず、基底同士の内積を計算することが困難もしくは不可能である。

基底が陰的にはあるが明確に定義される非線形主成分分析法として、Scheolkopf らによって導出されたカーネル主成分分析 (kernel principal component analysis; 以下 kPCA) がある [91]。津田 [40]、前田ら [51] は kPCA を部分空間類別法に適用するカーネル非線形部分空間法 (kernel nonlinear subspace method; 以下 KNS) を独立に提案し、実験的にその高い識別性能を示した。

本章では相互部分空間法に対して、核非線形型部分空間法<sup>1</sup>を適用し、新たな物体認識手法、核非線形相互部分空間法を導出する。

以下、7.2. では相互部分空間法、非線形主成分分析についての従来の研究を概観し、7.3. では Scheolkopf の議論に基づき、kPCA を導出する。その上で kPCA で算出される基底の類似度計算公式を理論的に導出し、核非線形相互部分空間法の具体的な表式を与える。7.4. では認識実験の結果を通し、提案手法が従来法より安定した動作を示す。7.5. では実験結果に関する考察を述べ、実験結果が物体の運動の自由度と認識率の関係を無矛盾に説明していることを示す。7.6. はまとめと今後の課題である。

## 7.2 相互部分空間法、非線形主成分分析

### 7.2.1 相互部分空間法

相互部分空間法は前田によりパターン認識の分野に導入された [52]。Watanabe の部分空間法では、学習のために与えられた  $n$  次元空間上の当該カテゴリに属する  $m$  個のベクトルから計算される自己相関行列、もしくは共分散行列を対角化して得られた固有ベクトル  $\vec{v}_i (i = 1, \dots, m)$  より累積寄与率などの基準により選択された  $D$  個のベクトルを基底とする部分空間を認識辞書とする。認識の際には、入力されたパターン  $\vec{x}$  と部分空間との何らかの類似度を計算することにより認識処理を実行する。類似度として部分空間への写像の大きさが用いられる場合には CLAFFIC、二乗誤差が用いられる場合には投影距離法、角度が用いられる場合には複合類似度法と呼ばれる<sup>2</sup>。

部分空間法は、学習データの最も広がり大きな部分をフィッティングする手法となっており、データを線形多様体で記述したことに対応する。そのため、部分

<sup>1</sup>この命名は前田による「カーネル非線形部分空間法」の「カーネル」を日本語の「核」に置き換えたものである。津田は「ヒルベルト空間における部分空間法」と命名しているが、核関数が多項式の場合には、写像される関数空間が有限次元となるため、前田の命名の方が一般性が高いと考えられる。

<sup>2</sup>ここで行った分類は著者の立場によるものであり、分類には他の観点もある。

空間法で用いられる類似度は認識対象の統計的な変動の少ない部分で定義される事になる.

相互部分空間法は, 認識対象画像の入力として複数のデータが利用出来る場合に適用され, 入力された画像を主成分分析し, 辞書として登録された固有ベクトルと入力データから得られた固有ベクトルの類似度を計算することにより認識処理を実行する. つまり, 学習データと認識対象データの変動の少ない量同士を比較するわけで, データの分布が線形近似可能な場合には強力な物体認識手法となりうる.

前田の定式化では部分空間同士の類似度は角度で定義され,  $m'$  個の入力画像列から得られた固有ベクトルを  $\vec{v}_j$ , ( $j = 1, \dots, m'$ ) とすると部分空間同士の角度は, 行列

$$X_{ij} = \sum_{l=1}^D (\vec{v}'_i \cdot \vec{v}_l)(\vec{v}_l \cdot \vec{v}'_j) \quad (7.1)$$

の固有値問題

$$\lambda \vec{x} = X_{ij} \vec{x} \quad (7.2)$$

を解くことにより最大固有値として与えられる [52]. ここで  $D$  は認識に用いる部分空間の次元数である.

### 7.2.2 非線形主成分分析法

物体認識問題において, 図 7.2<sup>3</sup>に示される様な姿勢変動に起因する強度の非線形性が存在することは, 多くの研究者によって示唆されていたが, 非線形性のクラス, その生成メカニズムなどが明らかでないことも有り, 正面から扱われることは少なかった.

村瀬らはこの様な非線形性を物体の角度をパラメータとした非線形回帰の問題として取り扱い, 3 次スプライン補間を用いることで物体の姿勢推定問題を解いた. しかし, 顕なパラメータが与えられない物体認識問題を考えると, 非線形多様体をコンパクトに記述する, 非線形主成分分析の適用が必須となる.

7.1. でも触れた通り, 統計学, 人工神経回路網の分野ではデータの非線形な分布を近似する非線形主成分分析の方法が数多く提案されてきた. 本節ではこれらの方法について概説するとともに, 相互部分空間法に適用する際の問題を指摘する.

統計学の分野では, データに多項式型の変換を施すことにより高次相関を取り込んだ拡張共分散行列を用いるアプローチが検討されている. このアプローチの代表的な例は Gnanadesikan[68] の方法であるが, この方法では高次相関を正面から扱っているため,  $n$  次元の属性を扱うために  $2n + \frac{n(n-1)}{2}$  次の行列を処理する必要があり, 計算量の観点からパターン認識への適用そのものが現実的ではない. これに対して斎藤はガウス確率密度関数の級数展開がエルミット多項式となることに着目し, 高次相関を考慮した非線形主成分分析法 NPCA1 を提案した [90]. これは後に坂野により手書き文字認識に適用され, パターン認識への適用可能性が論じられた [26]. しかし, この方法はそれぞれの次数の相関行列を独立したものと扱い, 主成分をそれらの固有ベクトルの和として表現するために, 部分空間同士の類似度を評価することが困難である.

一方, 人工神経回路網の分野では二つの多層パーセプトロンを結合し, 入出力が一致した場合の結合荷重を主成分として用いる, いわゆる自己想起パーセプトロンの方法が研究されてきた [87],[66][23]. さらに, Hastie は自己組織化ネットワークの形で非線形主成分を行うアルゴリズム Principal Curve を提案している [71]. これらの方法では, 部分空間の基底が顕わに与えられていないために, 類似度の計

<sup>3</sup>詳細は後述

算方法が明らかでないばかりでなく、部分空間の計算にバックプロパゲーションを用いることが多く、リアルタイムで動作することを要求される相互部分空間法のためには処理量が大きすぎるという問題があった。

### 7.3 核非線形相互部分空間法

前節で議論したとおり、過去に提案されていた非線形主成分分析法を相互部分空間法に適用することは、理論的にも実装上也困難である。

本節では Scheolkopf が提案した kPCA について、相互部分空間法への適用が可能である事を理論的に示す。まず Scheolkopf[91] に従い kPCA を導出し、その経過を用いて相互部分空間法へ適用するために基底間の類似度を評価する表式を導出する。

kPCA を一言で説明すると非線形写像でデータを無限次元もしくは極めて高次元の関数空間に写像した後、特異値分解で関数空間での主成分を求める方法と言うことが出来る。

通常、主成分分析は  $n$  次元の特徴空間  $\mathcal{R}^n$  上の  $m$  個の特徴ベクトル  $\vec{x}_i, (i = 1, m)$  から計算される共分散行列

$$C = \frac{1}{m} \sum_{j=1}^m \vec{x}_j \vec{x}_j^T \quad (7.3)$$

についての固有方程式,

$$\lambda \vec{v} = C \vec{v} \quad (7.4)$$

を解く事により、分布の主成分  $\vec{v}_i, (i = 1, m)$  を得る。但し、ここでは全てのデータはデータの重心から計算されたものとする。この主成分は線形近似の下でデータの変動の大きな方向を記述している。上の固有方程式は

$$\lambda \vec{x} = \left[ \frac{1}{m} \sum_{j=1}^m (\vec{x}_j \cdot \vec{x}_j^T) \right] \vec{v} \quad (7.5)$$

$$= \frac{1}{m} \sum_{j=1}^m (\vec{x}_j \vec{x}_j^T) \vec{v} = \frac{1}{m} \sum_{j=1}^m (\vec{x}_j \cdot \vec{v}) \vec{x}_j \quad (7.6)$$

の様に変形できる。このとき  $\vec{v}$  は  $\{\vec{x}_1, \dots, \vec{x}_m\}$  の間にあるので

$$\lambda (\vec{x}_k \cdot \vec{v}) = \vec{x}_k \cdot C \vec{v} \quad (7.7)$$

となる。以上の準備の下で kPCA を導出する。

最初に関数空間への非線形写像

$$\Psi : \mathcal{R}^N \rightarrow \mathcal{F}, \vec{x} \rightarrow \vec{X} \quad (7.8)$$

を考える。すると、空間  $\mathcal{F}$  での共分散行列は

$$\bar{C} = \frac{1}{m} \sum_{j=1}^m (\Psi(\vec{x}_j) \Psi(\vec{x}_j)^T) \quad (7.9)$$

と書ける。特に  $\mathcal{F}$  が無限次元の場合には  $\Psi(\vec{x}_j) \Psi(\vec{x}_j)^T$  は  $\mathcal{F}$  上の線形演算子

$$\vec{X} \rightarrow \Psi(\vec{x}_j) (\Psi(\vec{x}_j) \cdot \vec{X}) \quad (7.10)$$

となる. 但し, ここで  $\mathcal{F}$  は  $L_2$  の関数空間

$$\infty > (\Psi(\vec{x}) \cdot \Psi(\vec{x})^*) \quad (7.11)$$

であることを仮定している.

$\mathcal{F}$  上での主成分  $V \in \mathcal{F}/\{0\}$  を考えると, これは行列  $\bar{C}$  の固有値問題

$$\lambda \vec{V} = \bar{C} \vec{V} \quad (7.12)$$

の解となる.  $\mathcal{R}^N$  の時と同様に, これは

$$\lambda(\Psi(\vec{x}_k) \cdot V) = (\Psi(\vec{x}_k) \cdot \bar{C} \vec{V}) \quad (7.13)$$

と書ける. また, 主成分はデータの線形結合であるから, ある係数  $\alpha$  により

$$V = \sum_{i=1}^m \alpha_i \Psi(\vec{x}_i) \quad (7.14)$$

となることは明らかである.

式 (7.13)(7.14) を連立すると,

$$\begin{aligned} \lambda \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_k) \cdot \Psi(\vec{x}_i)) = \\ \frac{1}{m} \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_k) \cdot \sum_{j=1}^m \Psi(\vec{x}_j) (\Psi(\vec{x}_j) \cdot \Psi(\vec{x}_i))) \end{aligned} \quad (7.15)$$

を得る.

ここで新しい  $m \times m$  行列

$$K_{ij} = (\Psi(\vec{x}_i) \cdot \Psi(\vec{x}_j)) \quad (7.16)$$

を定義すると.

式 (7.15) は

$$\begin{aligned} \lambda \sum_{i=1}^m \alpha_i K_{ji} = \\ \frac{1}{m} \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_k) \cdot \sum_{i=1}^m \Psi(\vec{x}_j) K_{ji}) = \\ \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^m \alpha_i K_{kj} K_{ji} \end{aligned} \quad (7.17)$$

ここで  $\alpha_i$  は結合係数を並べたものなのでベクトルとみなすことが出来て,

$$m\lambda\alpha K = \alpha K^2 \quad (7.18)$$

と書ける. 両辺に  $K^{-1}$  を乗ずると

$$m\lambda\alpha = \alpha K \quad (7.19)$$

なる固有値問題を解く事によって  $\alpha$  の具体的な表式が得られることがわかる。これは  $\vec{V}^i \cdot \vec{V}^j = \delta_{ij}$  の様な制約条件を課すことによって特異値分解などの既知の方法で計算することが出来る。

固有ベクトル  $\alpha$  が求まると、任意の特徴ベクトル  $\vec{x}$  から  $V$  への写像は

$$V \cdot \Psi(\vec{x}) = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i (\Psi(\vec{x}_i) \cdot \Psi(\vec{x})) \quad (7.20)$$

によって計算できることがわかる。

但し、 $\Psi(\cdot)$  は一般に極めて高次元、時には無限次元のベクトルであるため、 $\Psi(\vec{x}) \Psi(\vec{y})$  を計算するためには極めて長時間もしくは無限の時間を必要とする。このため、 $\Psi(\cdot)$  の選択のためには Mercer の条件

$$k(\vec{x}, \vec{y}) = (\Psi(\vec{x}) \cdot \Psi(\vec{y})) \quad (7.21)$$

を満たすような写像を選ぶ必要がある。このような写像が選択できた場合には、 $\Psi(\vec{x}) \cdot \Psi(\vec{y})$  は単に関数  $k(\vec{x}, \vec{y})$  を計算することに帰着される。従って、

$$V \cdot \Psi(\vec{x}) = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i k(\vec{x}_i, \vec{x}) \quad (7.22)$$

の様になり、高次元の固有ベクトル  $V$  を顕わに求めなくても、その写像を計算することが出来るようになる。これを認識辞書に用いたパターン認識法が津田、前田により提案された KNS であり、実験的に高い識別性能が示されている [40][51]。

次に、入力画像系列からも同様の方法で固有ベクトルへの写像を計算した場合に、どの様にその類似性を評価するかを示す。

これまで用いてきた関数空間での固有ベクトル  $V$ 、固有値  $\lambda$ 、学習データ  $\vec{x}_i$  を辞書側と呼び、認識対象として入力された画像系列を  $\vec{x}'$ 、関数空間での固有ベクトルを  $W$ 、固有値  $\nu$ 、と書き、入力側と呼ぶことにする。つまり、 $V$  を「辞書側固有ベクトル」、 $W$  を「入力側固有ベクトル」、などと呼ぶことにする。

$V, W$  はそれぞれ辞書側、入力側データから計算された基底ベクトルであり、ノルムが正規化されていると仮定すると、辞書登録された画像群と入力された画像系列の類似度を評価するためには  $(V \cdot W)$  の値の大きさを評価すればいい事になる。 $(V \cdot W)$  は  $\mathcal{F}$  上の積分であるから、この値を顕わに有限の時間で計算することは出来ない。しかし、式 (7.14) を用いて  $V \cdot W$  の表式を書くと

$$W = \sum_{j=1}^{m'} \alpha'_j \Psi(\vec{x}'_j) \quad (7.23)$$

より

$$V \cdot W = \sum_{i=1}^m \alpha_i \Psi(\vec{x}_i) \sum_{j=1}^{m'} \alpha'_j \Psi(\vec{x}'_j) \quad (7.24)$$

$$= \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j (\Psi(\vec{x}_i) \cdot \Psi(\vec{x}'_j)) \quad (7.25)$$

である。ここで式 (7.21) を代入すると

$$V \cdot W = \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j k(\vec{x}_i, \vec{x}'_j) \quad (7.26)$$

認識手法	CLAFFIC	KNS	MSM	KMS
認識率 (%)	95.7	96.5	99.6	100.0
$\sigma$	-	450	-	410
辞書次元数	5	7	6	4
入力次元数	-	-	2	2

Table 7.1. Recognition rate of each procedure

となり、有限の計算時間で類似性が評価できることがわかる。

上に示した計算手順により kPCA を相互部分空間法に適用することが可能であることが示された。具体的に kPCA で求められた基底同士の角度を計算するためには、前節で引用した式 (7.1)(7.2) に (7.26) を代入すればよい。以上の手順で物体認識を行うアルゴリズムを核非線形相互部分空間法 (kernel mutual subspace method; 以下 KMS) と呼ぶことにする。

## 7.4 認識実験

本節では KMS を物体認識問題に適用し実験的に有効性を示す。実験は顔画像の認識問題を対象とした。顔画像認識が用いられるシチュエーションは、本人認証、ユーザインタフェースなどであるが双方とも連続的に顔画像を採取することは簡便であり、提案手法の応用が有効である。

実験データは、マンチェスター大学理工学研究所で収集された 15 人分のデータで [70] 被験者は無表情、光源一定で正面から横顔までの変化が連続的に撮影されている。これらから顔部分を手動で切り出した上で、 $15 \times 15$  画素の画像に正規化し実験に用いた。学習データは被験者一人あたり 10 枚の画像を様々な角度で等分になるよう選択した、評価データは被験者一人あたりの枚数が異なるが、20 枚から 50 枚を用いた。但し、MSM、KMS では重複する 5 枚の画像を一つの入力として扱い、結局 599 の画像列をテストデータとして用いた。核関数としてはガウス型 RBF

$$k(\vec{x}, \vec{y}) = \exp\left(-\frac{\|\vec{x} - \vec{y}\|^2}{2\sigma^2}\right) \quad (7.27)$$

を用い、予備実験により適切と思われる  $\sigma$  を設定した。

また、比較のために静止画を対象とした部分空間法として、CLAFFIC、KNS について実験を行った。各手法における最高の認識率を表 1. に示す。この実験だけでは有意とは言えないものの KMS が MSM を凌ぐ認識性能を見せている。また、図 (7.1) に部分空間次元数と認識率の関係を示す。次元数を適当に選んだ場合 KMS、MSM の性能はさほど違いがないが MSM では次元数が高すぎても低すぎても十分な性能が出ないのに対して、KMS では次元数への依存性が少なく、安定した性能を見せている。

## 7.5 考察

本節では MSM、KMS の性質の違いを明らかにするために、部分空間次元数の低い場合の解析を行い、認識性能の違いが生ずる原因について考察を試みる。

図 (7.2) は UMIST データベースの人物 1 の画像データを線形主成分分析により得られた第 1 主成分、第 2 主成分の張る平面中の人物 1 の画像の散布図である。テストデータを白点で、学習データは黒点で示し、さらに対応する画像データを提示してある。また、図中、中央にあるのは学習データを平均して作った画像である。

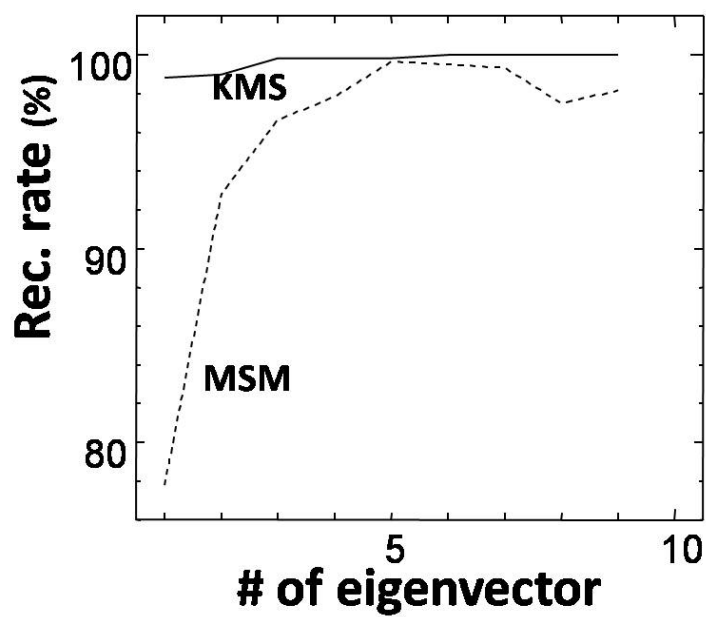


Figure 7.1. 認識率の次元数依存性

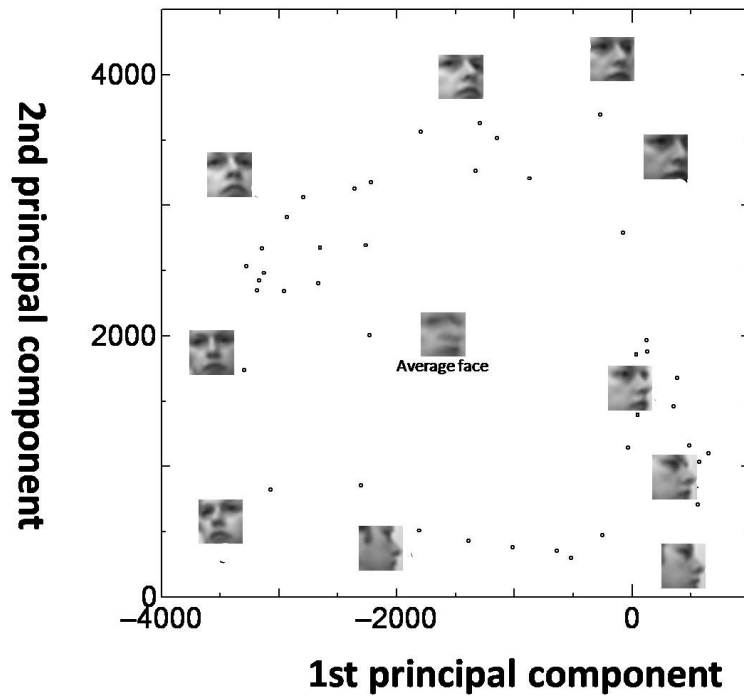


Figure 7.2. 線形主成分分析による顔画像の散布図



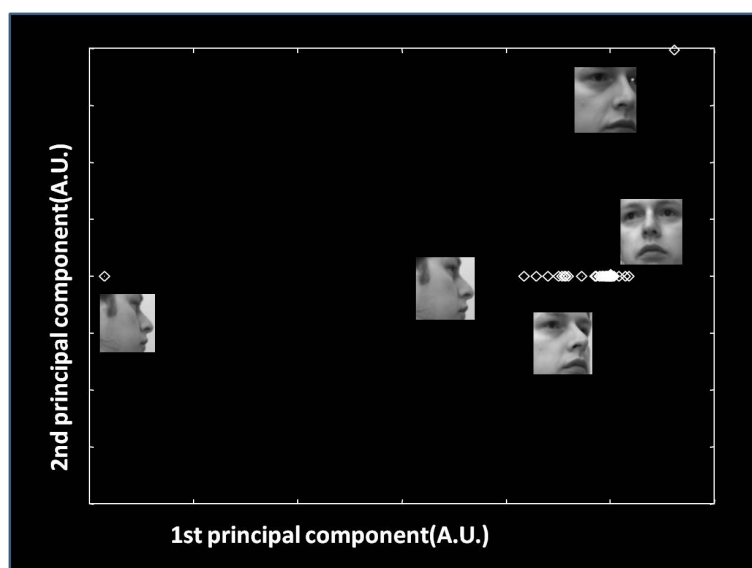


Figure 7.3. kPCA による顔画像の散布図

この散布図において、データの分布は、一つの軸が極端に長い Gauss 分布が湾曲しているものと解釈することが可能である。湾曲した軸上の顔が、回転のパラメータの変化と対応している事が分かる。さらに、平均画像が人物 1 の顔とは見做せない事から、この分布は非線形構造を持つと考えるのが自然である。一方、図 (7.5) は、同じデータを kPCA で写像した結果である。一部の外れ値を除いて分布がほぼ第 1 主成分で近似される様子が見える。

これらの結果は、KMS が 1 次元の部分空間を用いただけで高い認識率を得ているのに対して、MSM が 5-6 次元の部分空間を必要とすることを、矛盾なく説明していると考えることが出来る。即ち、顔の 1 自由度の運動に応じて特徴空間中に生成される 1 次元の軌跡を kPCA が高精度に近似し、さらに KMS の入力画像に対する雑音除去効果の為、1 次元の部分空間のみで高い認識率が得られていると見做せる。一方、MSM は 1 次元の軌跡を線形部分空間を用いて表現するために自由度とは関係ない高い次元数の部分空間が必要となると考えられる。

今回の実験は物体の自由度と認識率の高い部分空間次元数に関する無矛盾な仮説を示したものである。

しかしながら、この実験結果は特徴空間に出来る軌跡の非線形性の原因などについて新たな知見をもたらすものではない。この問題の追及のためにはさらに多様な条件での実験的研究の積み上げが必要である。

また、MSM では次元数が増大した場合には認識率の減少が見られる。これは部分空間の重なりが大きくなって行くためと説明されるが、KMS ではこのような現象は見られない。これは、KMS の認識処理が無限次元空間で行われるため部分空間の重なりの影響が少なくなっているためと考えられ、大カテゴリ問題への有効性を示唆している。

提案手法は処理量の観点からも改善の余地がある。kPCA で求められた基底同士の類似度は式 (7.26) で計算されるが、これは全ての学習データ、認識対象データの間でカーネル関数を計算することを意味しており MSM 等と比べると極めて処理量が多い。このためには少なくとも再計算可能な学習データについては記憶すべきデータを削減する必要がある。具体的な方向は必ずしも明確ではないが

同じく kernel trick を用いた認識手法である Support Vector Machine で用いられているデータ削減手法が有効である可能性もあり, そうした方向からの検討も期待される [62].

## 7.6 まとめと今後の課題

新しい物体認識手法である核非線形相互部分空間法を理論的に導出し, 顔画像認識に適用した. 認識実験の結果は従来の方法と比較して大きな認識率の向上を示してはいないが, 次元数に依存しない安定性, 物体の自由度の様な最低限の記述でカテゴリを表現出来ることなど様々な優れた特性を持つことを示した. また, 従来から示唆されていた画像の非線形構造と認識率の関係について無矛盾な実験結果を得た. 今後は多数の種類の変動を含んだデータについて実験を行い, 有効性を確認するとともに処理量の減少, 最適なパラメーターの決定について検討を進める予定である.

また, 動画画像を統計処理するアプローチを用いるためには連続するフレームから物体をセグメンテーションすることが前提となる. 顔の認識の場合には「顔」の領域をセグメンテーションした後で「誰か」を認識する問題となっているため, このアプローチが適用できるが一般の画像認識問題の場合には, セグメンテーション自体が難しい問題になる. 今後の研究の発展のためには, この課題を解決もしくは回避する方法を見出す必要がある.

## 7.7 その後の発展

筆者が提案した KMS の特徴は 1) 学習データが非線形性を持ち, 2) 認識対象データが連続的に入力される場合に高い認識性能を示すことである.

筆者の実験では認識性能の高さは有為には示されていないが, その後の福井 [47] らの物体認識に関する研究で KMS の非線形問題に関する優位性は明確に確認された.

同様な性質を持った問題として, 唇動作に依る個人認証の問題, 音声による個人認証の問題があるが, これらのについても市野らにより実験的に有効性が確認された [7, 8]. この研究に際し, 市野らは k-平均法を用いて学習サンプル数を減少させることにより, KMS の処理速度を劇的に向上させることに成功している [9].

また, 少々意外な応用として宇宙機のテレメトリデータからの異常検出への応用がある. 藤巻らは独自に KMS と同様なアルゴリズムを導出し人工衛星の異常検出への応用を行った [48].

これらの研究からは KMS にはさらに広い応用可能性があると考えられる. 例えば, 筆者識別, ジェスチャ認識, 手話認識などである. 今後の研究により KMS の有用性はさらに証明されていくものと考えられる.

## Chapter 8

# 結論：パターン認識におけるアルゴリズム設計の課題

本章では、これまでの研究成果を受け、得られた結論とそこから導かれる今後の研究課題を抽出する。

### 8.1 パターン集合の線形性とアルゴリズムの有効性の関係

本論文では、特徴ベクトルから画像を再構成すること、特徴ベクトルを低次元に写像することの二つの手段で、特徴空間とパターン集合の代数構造を調べた。

調べた対象は文字認識、物体認識についてそれぞれ異なっているため、分割して整理する。

#### 8.1.1 文字認識アルゴリズムについて

著者は高精度な文字認識系を構成するための特徴抽出アルゴリズムである拡張外郭方向寄与度特徴について、遺伝的アルゴリズムを用いた文字画像再構成手法を提案し、特徴空間の中で非常に重要な部分である、文字認識辞書とその中点を文字画像として可視化することに成功した。

これを学習手法を変えて比較することで、

- 特徴ベクトルの平均を用いた認識辞書はきれいな当該カテゴリの文字になっていること
- 中点がどちらかのカテゴリに偏っている場合には認識精度が低くなること
- LVQの学習過程では誤認識した文字の傾向にしたがって学習が進行すること

などの現象を観測することが出来た。これらの事実は、拡張外郭方向寄与度特徴の空間では特徴ベクトルは加法的であり、限定的かもしれないがスカラ倍も成立することがわかった。

ゆえに拡張外郭方向寄与度特徴の空間はユークリッド的であり、その中では文字画像集合は線形性を持つことがわかった。

### 8.1.2 物体認識アルゴリズムについて

濃淡特徴を用いた物体認識系に現れる非線形構造について、主成分分析を用いた低次元写像で現象を調べるとともに、原因を考察し、新しい物体認識アルゴリズム、核非線形相互部分空間法を提案し、理論的に動作を保障するとともに実験的に有効性を示した。

また、提案手法が動作する条件が、非線形写像によりパターン集合が線形に分布する場合であることを実験的に示し、物体認識系の場合にもパターン空間の線形性が本質的であることを確認した。

この結論は文字認識における結果と無矛盾であり、パターン空間の線形性の重要性を実験的に明確化した初めての例となった。

## 8.2 残された問題

### 8.2.1 より優れた認識系の設計論の実現に向けて

当初に指摘した最大の問題は、人間が単一の認識系で様々な対象を認識可能であるのに対して、工学的には認識対象によって様々な技術を開発しなくてはならないという事実であった。この問題自体の解決は近未来に望むことは不可能であろう。

しかし、本論文の結論からパターン認識技術が進むべき方向はある程度明確化された。

現代のパターン認識機械はヒューリスティックに開発された特徴抽出機と統計的学習理論に基づく識別器で構成される。本論文の結論から、特徴空間がユークリッド的であり、その中でのパターン集合が線形空間であれば、パターン認識装置は良好に動作することが明らかになった。

しかし、これをどのように確認するか一般論が得られたわけではない。文字認識においてこの結論を出すために用いた遺伝的アルゴリズムを用いた方法は、「似た特徴を抽出できる画像」を探索する仕組みであり、画像再構成のための必要十分条件を満たすものではない。従って、複数の解がある場合にこれを見逃す場合もありうるし、特徴がそもそも画像再構成に十分な情報を持っていない場合にはどのような解が得られるかについても何の保証もない。

また、物体認識系については主として線形、非線形主成分分析の結果から、特徴空間でのパターン分布を観測した。これが可能であるのは顔画像の画素間の相関が非常に高いために、主成分分析によりかなり低次元で分布の特徴が捉えられるからに他ならない。

つまり、著者が採用している方法は文字認識における拡張外郭方向寄与度特徴と顔認識における濃淡特徴の性質に依存した方法になっているという意味で、従来のパターン認識研究の頸城を脱却しているわけではない。

将来は、こうした検討を一般的に行うための枠組みの研究が必要である。そのためには、画像特徴から画像への再構成の一般論、低次元空間への写像の効率的な方法の検討が重要である。

### 8.2.2 人間の様な認識系の実現に向けて

これまでの研究ではパターン集合の代数構造を調べるために用いてきた手段は、パターンを人間が「見る」ことであった。これはあるパターンが文字であるか、顔であるか、あるいはどの文字であるか、誰の顔であるかという判定を人間が確認する以外の手段で特定することが不可能であることによる。

このことは単純であるが重大な問題を提起する．一般的にパターン，あるいは人間がパターンと呼ぶものには厳密な定義がないということである<sup>1</sup>．

故に，あるパターンが何らかのカテゴリに属するかどうかを判定することは人間の主観に頼らざるを得ない．このことは人間が行っているパターン認識プロセスが狭義の現代科学の条件，すなわち論理性，客観性，再現性のうち客観性を持たないという意味で科学的ではないことを示唆する．

しかし，このことはパターン認識の仕組みを解明する上で科学的なアプローチを否定するものではない．

近年，科学の多数の方面で狭義の科学の3条件を満たさない分野が増えている．例えば経済学は従来の需要と供給のバランス，つまり市場原理が価格を決定するという，科学的な過程をおいたマクロ経済学の立場から，消費者の価値観，つまり主観が価格を決定するという超ミクロ経済学の方向にシフトしている [42]．

また，従来からいわれているとおり進化論や宇宙論は再現性を欠くという意味でやはり従来の科学の枠組みには収まらない．

これらの科学に共通する課題は「歴史性」であり，パターン認識の視点に立つと，人間が進化の過程で身に着けてきたパターンの概念，学習や約束で身に着けてきたパターンの概念を再現することは，進化による人間の発生，宇宙物理学的发展による現在の宇宙の姿を再現することと同程度に困難であると考えられる．

こうしたこれまでの科学の枠に収まらない分野を「科学」として扱うための新しい論理（それがどんなものかはわからないが）こそが，パターン認識分野を真の科学にしているために必要に思われる．

例えばペンローズの様な異様なアプローチにもなんらかの真理があるかもしれない [49]．

---

<sup>1</sup>一見，顔には「目鼻などの感覚器がついた場所」の様な定義が可能に見える．しかし，物理的実態ではない顔，例えば絵画，漫画などの顔でも人間は顔だと認識する．その意味ではやはり顔にも厳密な定義はないと考えるべきであろう．



## Appendix A

# Coulomb 相互作用モデルによる二次識別関数の学習的構成法

### A.1 はじめに

パターン認識などに用いられる識別関数の学習的構成法は Rosenblatt の perceptron をはじめとした多数の手法が提案されており，文字認識をはじめとする種々のパターン識別問題に適用されてきた [59][85]．識別関数の学習問題については

1. 識別面の形状
2. 学習ルール
3. 評価関数及び収束規準

などが主な主題となるが，従来の手法に関する議論は，主に識別面の形状と学習ルールに関するものが中心であった．本章では，識別面の構成に関してどのような評価関数が最適か？という問題を出発点に，新たな学習モデルの提案を行う．

従来の手法，例えば多層 perceptron における誤差逆伝播法 [44]，あるいは Kohonen による LVQ (Learning Vector Quantization)[79][6] などの手法は複雑な識別面を構成できるという意味で極めて強力な学習手法であるが，これらは過学習という共通の問題を抱えており，実用上の大きな問題になっている [89]．一方で過学習の可能性の少ない perceptron などの単純な識別系においてはサンプルの線形分離可能性が学習の成否を決定する重要な要素となるが，現実的な問題では事前に線形分離の保証が得られるのはまれで，線形分離可能性を学習の成否で判断せざるを得ないのが現状である．

筆者らは，これらに共通する問題が評価関数および収束規準にあると考えた [25]．これまでに提案されている学習的手法において最も多く用いられている評価関数は perceptron 型評価関数である．これを距離の観点からみると [15]，識別面からの距離にほぼ比例して，誤って識別されたサンプルは強いポテンシャルを持つことになり，孤立したサンプルが存在した場合には，その重要性の高低にかかわらず，学習機械は救済のための努力を行なうことになる．故に，識別面が分離不可能である場合には収束しないばかりでなく，良好なパラメーターに近づくとという保証も存在しない．さらに，多くの場合には，識別誤りを 0 にするという収束基準が用いられているために孤立サンプルを救うためにロバスト性の低い識別面を作ってしまうことがある．

本章では、こうした問題を解決するためにサンプルの重要度を考慮した評価関数と、それに基づく学習モデル、Coulomb 相互作用モデルを提案する。提案手法は、自動的に重要なサンプルとそうでないサンプルを区別し、重要でないサンプルについては救済を諦め、学習を収束させることにより、誤った教師や極端に変形の激しいパターンによる識別関数のオーバーチューニングを緩和することを狙っている。本論文では A1.2. で学習に用いる新しい評価関数の定式化を行い、Coulomb 相互作用モデルを導出する。そして、相互作用の平衡点と Bayes 識別面との関係を議論し、同モデルに基づく学習アルゴリズムが Bayes 識別の観点で準最適解に収束しうること示す。A1.3. では最も単純な 2 次識別関数である重み付きユークリッド距離について学習手法アルゴリズムを提示する。A1.4. では正解の解った問題に対して提案手法の振る舞いを実験的に分析し、そして、A1.5. では現実的なパターン認識問題である手書き文字の識別に適用した場合の実験結果を示す。A1.6. では議論のまとめと今後の課題について述べる。

## A.2 Coulomb 相互作用モデル

### A.2.1 評価関数の定式化

サンプルの重要度を反映した評価関数を導出するため、誤り訂正学習における学習ルールを、パターン間の相互作用による変動ととらえ理論を展開する。まず、個々のサンプルにとって、自分に似ている、という意味で重要なサンプルは距離の近いサンプルであり、その数が多ければ重要性は非常に大きいと考えられる。しかし、距離的に遠いサンプルであっても特定の方向に自分と同じカテゴリのサンプルが多く存在すれば、その方向への変動がそのサンプルの本質的な情報を含んでいるという意味で重要度が高いといつてよいであろう。ゆえに、パターン間の相互作用は以下の性質を持つことが望ましい、

- 距離が遠いときには小さくなる
- 重ね合わせが可能で、小さな値が重なって強くなることが可能である。

このような性質をもつ相互作用は数多く考えられ、例えば、クラスタリング手法に対しては Watanabe[100], 若原 [57] の手法があるが、ここでは、 $n$  次元の特徴空間において各々のパターンがお互いの視野を占める大きさに着目して議論する。つまり、直観的には、近くにあるパターンは大きく見え、また、密なパターン集合は遠くからでも良く見えるという考え方から出発する。

このことの数学的な表現を考えるため、仮想的にパターンに大きさを考える。このとき、方向に対する不変性と、解析の容易さからパターンが単位長さを直径とする  $n$  次元の超球であると考えよう。すると、立体角の考え方から、距離が  $r$  だけ離れたパターンは、視野の中で

$$\frac{\sqrt{(\pi)^{n-1}}}{\Gamma(\frac{1}{2}(n+1))} \frac{1}{r^{n-1}} \quad (\text{A.1})$$

だけの面積を占めることになる [20]。ここで、 $\Gamma$  はガンマ関数、 $n$  は特徴空間の次元数である。このときパターンの視直径は距離が遠いときには小さくなり、遠くても大量にあれば集団としての面積の和が大きくなることから、パターンの視直径をそのまま他のパターンに対する相互作用の強さと考えると、あるパターンから距離  $r$  のところにある他のパターンに働く相互作用は Gauss の法則を満たすと考えられる。すなわち

$$\int \vec{E} \cdot d\vec{s}(r) = \rho \int_0^r f(\vec{x}) dv \quad (\text{A.2})$$



である．ここで， $|\vec{E}|$  は相互作用の強さ， $\vec{s}$  は面積要素， $f(\vec{x})$  は分布の確率密度関数， $\vec{x}$  は相互作用を受けるパターンの座標， $\rho$  は相互作用の強度を決めるパラメータ， $v$  は体積要素である．

式 (A.2) について，カテゴリを囲む  $n$  次元の球を考えると，

$$\vec{E}(\vec{x}) \cdot \vec{s}(\vec{x}) = \rho \int_0^r f(\vec{x}) dv \quad (\text{A.3})$$

とくに一点からの相互作用については

$$E = \frac{\rho}{S(r)} \quad (\text{A.4})$$

と書ける． $S(r)$  を点を囲む超球の表面積と考えると

$$S(r) = n \frac{\sqrt{\pi}^n}{\Gamma(\frac{n}{2} + 1)} r^{n-1} \quad (\text{A.5})$$

であるから，相互作用の表式は結局

$$E(r) = \rho \frac{\Gamma(\frac{n}{2} + 1)}{n \sqrt{\pi}^n r^{n-1}} \quad (\text{A.6})$$

と書ける．例えば， $n = 3$  では

$$E = \frac{\rho}{4\pi r^2} \quad (\text{A.7})$$

となり普通の電場と一致する．

誤り訂正型の学習アルゴリズムにおいて，評価関数はパラメータを変化させる相互作用のポテンシャルとして与えられるので，

$$\vec{E} = -\nabla\phi \quad (\text{A.8})$$

の様に評価関数  $\phi$  を表記することが出来る．

このような評価関数を用いた教師付き学習プロセスはパターン間に同じカテゴリ同士なら引き合い，異なるカテゴリ同士なら反発し合うような相互作用を導入し，その力に従ってパラメータを変更していく過程と考えられる．この相互作用は

- 強さが距離に反比例する
- 次元数が大きくなるほど相互作用が弱くなる

などの意味で我々の識別に関する直感と大枠で一致する．

距離関数を学習するイメージでは，学習の対象となるパラメータ，例えばテンプレート，重みなどのパラメータを学習ごとに変化させる際に，遠くにある誤識別パターンに対しては，あまり大きな変化を起こさないが，それが大量にある場合にはその方向への変動は累積的に大きくなり，最終的にそれらのパターンは救済されることになる．このような評価関数をもつ学習アルゴリズムを静電場との類推から Coulomb 相互作用モデルと名付けることにする．

### A.2.2 Bayes 識別面との関係

ここでは提案したモデルと Bayes 識別との関係について議論する．簡単のために 1 次元の 2 カテゴリ問題について考える．一般的な式には添字を付けず，二つの分布を区別する必要があるときにだけ密度関数，平均値，スケールファクターなどに (1), (2), の上付き添字を付けるものとする．図 A.1 に示すような 2 つの分布がある場合に，誤識別となる部分を 1 次近似，すなわち確率密度関数の Bayes 識別面（この場合は点）からの接線で表現する．なお，ここでは分布は単峰性であると仮定する．すると，Bayes 識別面を  $x_0$  として確率密度関数は近似的に

$$f_{app} = \frac{df(x_0)}{dx}(x - x_0) + f(x_0) \quad (A.9)$$

と書ける．ここで 1 次元での相互作用の表式は式 (A.6) より，

$$E = \rho \frac{\Gamma(\frac{3}{2})}{\sqrt{\pi}} \quad (A.10)$$

の様に定数となるので，相互作用の強さの比較は，図のハッチングされた三角形の面積を比較することで行われる．ここで，1 次近似された確率密度関数と  $x$  軸の交点を  $x_c$ ， $x$  での微分を ' と書くと， $x_0$  と  $x_c$  の間の距離は  $x_0 - x_c = -\frac{f(x_0)}{f'(x_0)}$  となるので，

$$E = \rho \int f_{app} dx = \rho \left| \frac{f(x_0)^2}{f'} \right| \quad (A.11)$$

であるから，結局，Bayes 識別面での相互作用の強さの違いは確率密度関数の微係数の違いで表される．ゆえに，Coulomb 相互作用の平衡点と Bayes 識別面の一致する条件は，

$$\left| \frac{df^{(1)}(x_0)}{dx} \right| = \left| \frac{df^{(2)}(x_0)}{dx} \right| \quad (A.12)$$

で与えられることになる．

この条件は， $f^{(1)}$  と  $f^{(2)}$  の関数形が同じ時には明らかに満たされ，他の場合にはかなり特殊な場合にしか満たされない．しかし，ここで

$$f = f\left(\frac{x - \mu}{\sigma}\right) \quad (A.13)$$

と書くと， $\xi = \frac{x - \mu}{\sigma}$  として，

$$\frac{df}{dx} = \frac{1}{\sigma} \frac{df}{d\xi} \quad (A.14)$$

であるから，微係数は分布のスケールファクターを変更することで変化させられるので，識別関数のスケールファクター，つまり Mahalanobis 距離の共分散行列，重み付きユークリッド距離の重みなどを変更することで Coulomb 型相互作用の平衡点と Bayes 識別面を近似的に一致，もしくは識別関数の拘束条件の下での準最適解に到達させられることが分かる．しかし，最初に仮定した誤識別領域の体積は密度関数の変曲点より内側ではあまり良い近似にならない．実用的には分布が近似的に正規分布と見做せるような場合に標準偏差の内側に識別面があるときには，この近似は正当ではない．つまり，学習前の識別率が 60% を下回るような場合には Coulomb 型の評価関数を用いて良好な識別面が得られるとは言えない．

また，ここでの考察は 1 次元に限定しているが直感的には高次元になれば識別超平面から離れた部分の寄与は減少して行くため Bayes 識別面付近での相互作用の強さはより拮抗してくることが予想される．特に十分に次元数が大きければ真の Bayes 識別面においては相互作用は平衡となる．(付録参照)

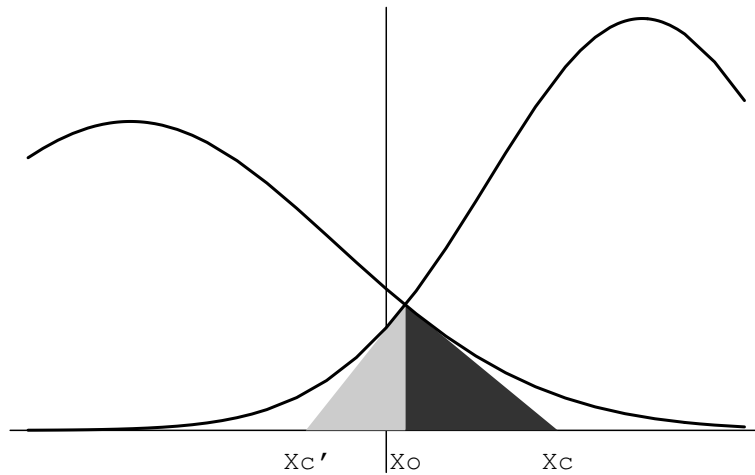


Figure A.1. 誤識別領域の近似的な「体積」

## A.3 学習アルゴリズム

### A.3.1 重み付きユークリッド距離によるモデル

ここでは、2.1で提案した評価関数を用いて識別関数を学習するアルゴリズムを具体的に定式化する。教師つき学習のプロセスは、普通、誤りなどの条件に応じ、標準パターンの属性として辞書に格納されている参照ベクトルなどを逐次変更していく過程として実現される。提案した評価関数を最小化する意味合いではスケールファクターを持った物であればどのような識別関数を対象としてもよいが、実用上の重要性から重み付きユークリッド距離

$$r_w^2 = \sum_{i=1}^N w_i (x_i - \mu_i)^2 \quad (\text{A.15})$$

を考える。

二つのカテゴリ1, 2について、重み付きユークリッド距離で与えられる識別境界は方程式、

$$\sum_{i=1}^N w_i^{(1)} (x_i - \mu_i^{(1)})^2 = \sum_{i=1}^N w_i^{(2)} (x_i - \mu_i^{(2)})^2 \quad (\text{A.16})$$

の解で与えられる。統計的識別の立場では、標準ベクトル  $\mu_i$  はカテゴリを構成するパターンの平均値、重み  $w_i$  はカテゴリの標準偏差の逆数として与えられる。本章では、このうちの重みに着目してアルゴリズムを構成する。

重みの学習は2.1で提案したように誤識別されたパターンについての相互作用を計算し、それを小さくする方向に重みを変化させれば良いが

- 相互作用をどのように計算するか
- 重みの変化の方向をどのように設定するか

等の点を考えなくてはならない。2.で展開した理論通りにアルゴリズムを構成すると、誤りを起こしたパターンは系を構成する全てのパターンと相互作用し、その

相互作用に基づいて識別関数のスケールファクターを変更することになる。しかし、全てのパターンとの相互作用を計算する事は計算時間の点で現実的ではない。

ところで、誤り訂正型の学習アルゴリズムという点に着目すると相互作用を考えるべきパターンは誤りを起こしたパターンとカテゴリを構成するパターンとなる。単峰性の分布を仮定した場合には、誤りを起こしたパターンが十分にカテゴリの中心から離れていれば、カテゴリを構成するパターンの大半はカテゴリの中心付近に集まっていると見てよい。そこで考察すべき相互作用は、カテゴリの中心と、誤識別されたパターンとの間に働くものだけに限定してもさほど悪い近似ではないと考える。この近似の下では誤識別されたパターンがカテゴリの中心と相互作用し、正解パターンに近づき、誤識別カテゴリからは遠ざかるように重みを修正することになる。

この近似を入れることで、分布の形状についての情報が落ちてしまうが、このような近似が成立しないような系では 2. で与えた近似的な estimate も正当ではない場合が多く、いずれにせよこのモデルを使うべきではないと考える。

### A.3.2 重み変更規則の導出

3.2 では重みつきユークリッド距離の重みを変更することにより相互作用を最小化する学習規則について述べる。誤識別されたパターンに対してカテゴリの中心から働く相互作用は

$$E_i = \frac{E(r)x_i}{r} \quad (\text{A.17})$$

と書ける。ここで  $r$  はパターン集合の中心からの距離であるが、学習する距離関数にあわせて、 $r$  として重み付きユークリッド距離を用いる。また、 $x_i$  は座標の各成分である。式 (A.17) に従ってサンプルパターンは正解となるカテゴリの参照パターンに近づき、誤識別したカテゴリから遠ざかるように重みを変更する。ここで簡単のために誤識別されたパターンが動くものとする。このとき、カテゴリの標準パターンからの移動ベクトル  $\Delta\vec{x}$  を

$$\begin{aligned} &\text{正解側} \\ -\Delta\vec{x} &= -\frac{E_c(\vec{x} - \vec{\mu}_c)}{r_c} \\ &\text{不正解側} \\ -\Delta\vec{x} &= \frac{E_f(\vec{x} - \vec{\mu}_f)}{r_f} \end{aligned} \quad (\text{A.18})$$

とする。ここで  $r, E$  などにつく添字  $c, f$  は正解 (correct), 不正解 (false) を表す。すると新しい位置  $\vec{x}^{new}$  は

$$\begin{aligned} &\text{正解側} \\ (\vec{x} - \vec{\mu}_c) - \Delta\vec{x}_c &= (\vec{x}^{new} - \vec{\mu}_c) \\ &\text{不正解側} \\ (\vec{x} - \vec{\mu}_f) + \Delta\vec{x}_f &= (\vec{x}^{new} - \vec{\mu}_f) \end{aligned} \quad (\text{A.19})$$

となる。

重みによる識別面の最適化のためにはサンプルパターンが、式 (A.19) の位置で感じる相互作用と同じ強さの相互作用を感じるように重みを動かせばよい。相互作用が働いたことによる各カテゴリの標準パターンと誤識別パターンの相対位置

の変化は、式 (A.19) で与えられるので、式 (A.16) を考慮して、新しい重み  $\vec{w}^{new}$  は成分で書くと、

$$\begin{aligned} & \text{正解側} \\ & \sqrt{w_{c_i}^{new}}(x_i - \mu_{c_i}) = \sqrt{w_{c_i}}(x_i - \mu_{c_i} - \Delta x_{c_i}) \\ & \text{不正解側} \\ & \sqrt{w_{f_i}^{new}}(x_i - \mu_{f_i}) = \sqrt{w_{f_i}}(x_i - \mu_{f_i} + \Delta x_{f_i}) \end{aligned} \quad (\text{A.20})$$

を満たすように決定すればよい。よって、新しい重み  $\vec{w}^{new}$  は

$$\begin{aligned} & \text{正解側} \\ & \vec{w}_c^{new} = \vec{w}_c \left( \vec{1} - \frac{\Delta \vec{x}_c}{|\vec{x} - \mu_c|} \right)^2 \\ & \text{不正解側} \\ & \vec{w}_f^{new} = \vec{w}_f \left( \vec{1} + \frac{\Delta \vec{x}_f}{|\vec{x} - \mu_f|} \right)^2 \end{aligned} \quad (\text{A.21})$$

のように決定されることになる。ここで  $\vec{1}$  は単位ベクトルである。

結局、式 (A.21) により、重みつきユークリッド距離のための重みの逐次変更ルールが与えられる。

### A.3.3 学習の収束規準

カテゴリ同士が線形分離不可能な場合、perceptron においては、真の識別境界近傍を振動することが知られているが、この「近傍」については明確な定義が与えられず、従って、perceptron 型の学習則ではなんらかの方法で強制収束させたとしても良好な識別面が得られる保証はない。一方、Coulomb 相互作用モデルにおいては、静電場について知られている Earnshaw の定理 [67] により、識別面が安定にならないことがほぼ確実に言える。従って、なんらかの方法で学習を強制的に収束させなくてはならなくなるが、相互作用のポテンシャルの系全体についての和

$$\psi = - \sum_{x_i \in \mathcal{X}} \int_{\infty}^0 \vec{E}_i \cdot d\vec{r}_i \quad (\text{A.22})$$

をそのまま収束規準として用いることができる。ただし、 $\mathcal{X}$  は誤識別されたパターンの集合である。

学習の最終段階では、このポテンシャルの値はある一定値の周辺で振動することになるが、なんらかの方法で振動を減衰させてやればポテンシャルは一定値に落ち着くと考える。通常、このような現象に対しては緩和的な方法による収束が採用されている [63]。

ここで、相互作用の表式 (A.6) と式 (A.18) を連立すると、

$$\Delta \vec{x} = \pm \frac{\rho \Gamma(\frac{n}{2} + 1)}{n \sqrt{\pi^n} r^n} (\vec{x} - \vec{\mu}) \quad (\text{A.23})$$

の形式で書けるので、緩和的な方法で、運動を徐々に小さくしていくためには学習が行われるごとに  $\rho$  の値を小さくしてやればよい。ただし、式 (A.18) から分かるように式 (A.22) で表されるポテンシャルの値は必ずしも正の値には限られない。

結果的にはこのような条件のもとで、ポテンシャルの値が一定値に落ち着くことが収束条件として与えられる。このような収束法は、perceptron をそのまま緩和的に用いたのと同様に見えるが、分離不可能な場合に識別面がどのような場所にいるかを議論できる意味で異なっている。5. で実験的に示すが、緩和型 perceptron では分離不能の場合には良好な解を得ることは出来ない。

手法に含まれるパラメータ  $\rho$  の初期値と  $\rho$  の値を小さくしていくためのスケジューリングについては理論的に予言することは出来ない。さらに、初期の  $\rho$  やスケジューリングのやり方により、重みの値が変動してしまうために、結果的に同じ様な識別面を与えることになったとしても、収束する先のポテンシャルの値は異なってしまうことが考えられる。これらについては実験的にしか決定することができない。

## A.4 識別面のダイナミクス

### A.4.1 理論の確認

これまでの議論で Gauss の法則を満たすようなタイプの相互作用によるパターンの仮想的な運動を考え、それに添った形で識別関数の重みを誤り訂正過程の形で更新していき、相互作用の平衡点を求めることにより良好な識別超平面を得ることが出来る可能性があることを示した。

ここでは、この可能性を検証するとともに、現象を直感的に把握するために、Bayes 識別の意味での最適解がわかっている小規模な系でシミュレーションを行なう。

シミュレーションに用いたのは、図 A.2 に示した 2次元の 2カテゴリ問題である。それぞれのカテゴリは表 A.1 に示すパラメーターで表される相関のある 2次元の Gauss 分布を基に発生させた乱数にラベル付けを行なうことにより得られたデータで、1カテゴリあたり 2200 個のデータを発生させ、この中から 100 個をランダムに抽出して学習データとした。図 A.2 においては  $x, o$  などの文字がサンプルを、破線が Bayes 識別の意味での最適な識別面を、実線が重み付きユークリッド距離による識別面を表している。

Table A.1. 実験に用いた Gauss 分布のパラメーター

	平均 (x)	平均 (y)	標準偏差 (x)	標準偏差 (y)	相関係数
カテゴリ 1 (o)	30	30	14	11	0.8
カテゴリ 2 (x)	70	70	20	10	-0.3

このデータは

- 完全に重なっているサンプルがあり、perceptron などの手法では収束しない<sup>1</sup>
- 相関があるため、重み付きユークリッド距離の場合には改善の余地がある

<sup>1</sup>無論、現実のパターン認識問題ではこのようなパターンは存在しないが、ごく近くに近接したパターンが存在し、学習時の困難となって現われることがある。このような場合の一つの極限として完全に重なったパターンを残した。

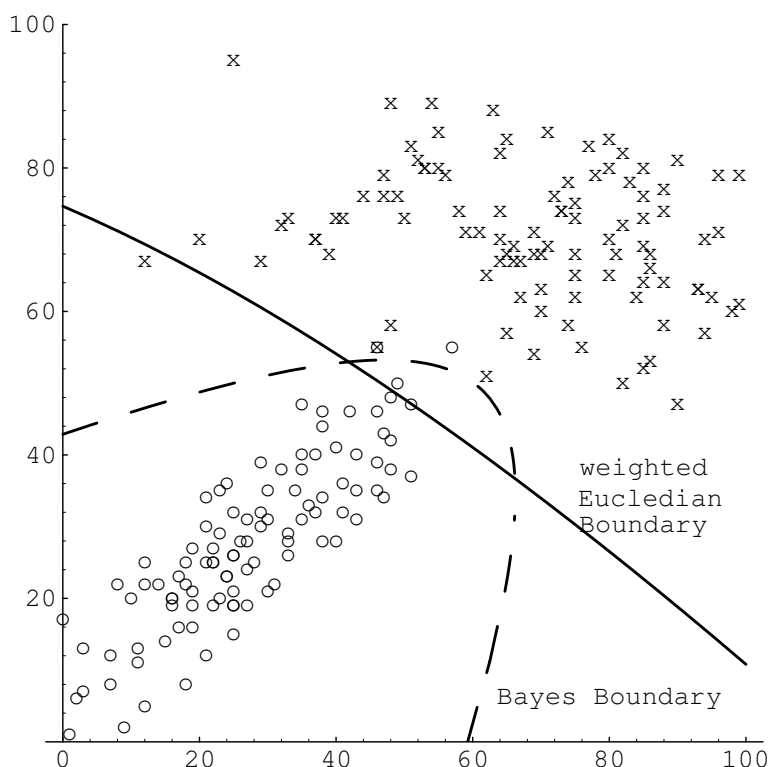


Figure A.2. テストに用いたサンプルと従来の手法で得られる識別面

等の意味でテストのために好適であると考えた。

学習に伴う識別面の移動の様子を図 A.3 に示す。図中、矢印で示したのがそのときに誤識別されたサンプル、破線は誤識別となったときの識別面、実線は誤り訂正のために重みを変更した際の識別面である。図には学習が始まってからの 4 例のみを示したが、これらの結果から、提案手法が基本的には、誤りを訂正するように振る舞っていることがわかる。特に特徴的なのは 4 番目の例で、この場合には分布の中心からずれたサンプルを救うために、識別面が曲率を変えて、分布のより重なった部分での誤識別を増加させないように振る舞っている。

次により大局的な振る舞いを考察する。3. でも述べたように、提案手法においては系全体のポテンシャルの値を調べることで、学習の進行状況ある程度把握することができる。

まず最初に緩和的な方法を用いない場合の提案手法の振る舞いを図 A.4, A.5 に示す。それぞれ、学習の進行に伴うポテンシャルと誤識別数の変化を示している。誤識別の数に大きな変動はなく、正解から大きく外れた部分を探索するようなことにはなっていないと見られるものの、ポテンシャルの変化が、特定の値を中心として振動的に振る舞っている。

そこで、強制的に収束させるために、 $\rho$  の値を減衰させたときの振る舞いを調べた。 $\rho$  の値の減衰には  $\rho_0$  を初期値として、

$$\rho = \rho_0 c_{rate}^{\tau} Prate \quad (\text{A.24})$$

という形のスケジュールを用いた。ここで  $\tau$  は学習回数、 $c_{rate}$  は  $0 < c_{rate} < 1$  なる定数で、このパラメータによって緩和スケジュールが決定されることにな

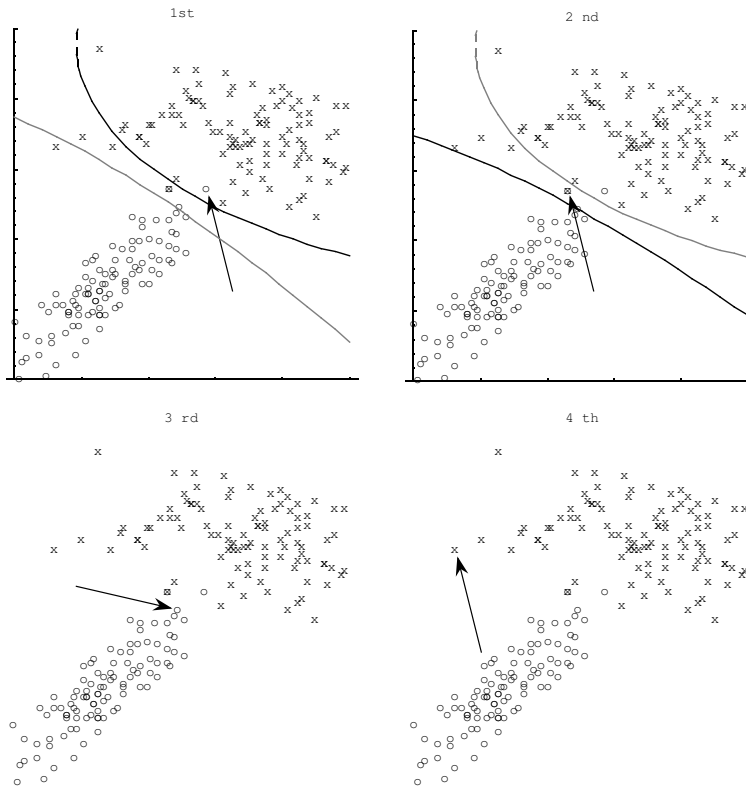


Figure A.3. 提案手法による誤り訂正の様子

る．また， $p_{rate}$  は一つ前の学習サイクルの誤識別率で識別率が上がるごとにパラメーターの修正を小さくする意味合いで導入した．

図 A.6, A.7 に示したのがそれぞれそのような場合のポテンシャルと識別率の変化で，緩和なしの場合に比べて急速に特定の値に収束し，学習パターンの識別率も高く保たれることが分かる．このとき，パラメーターは  $\rho_0 = 1.0, c_{rate} = 0.999$  であった．このような方法で最終的に得られた識別率を表 A.2 に示す．表からわかるように本学習手法においては未知サンプルにおいて，Bayes 識別の場合の識別率 98.77% と学習前の識別率 97.18% の差を 60% ほど吸収した 98.27% という値になっている．識別面も図 A.8 に示すように分布の間が最も重なっていると思われる部分では Bayes 識別面とほぼ一致しており，提案手法の有効性を裏付けている．

Table A.2. 正規乱数学習終了時の識別率の比較

	Bayes 識別の場合	学習前 (重み付きユークリッド距離)	提案手法
学習	99.5%	98.0%	99.0%
未知	98.77%	97.18%	98.27%

さらに，サンプル数が大きく，カテゴリ同士の重なりが大きくなるような場



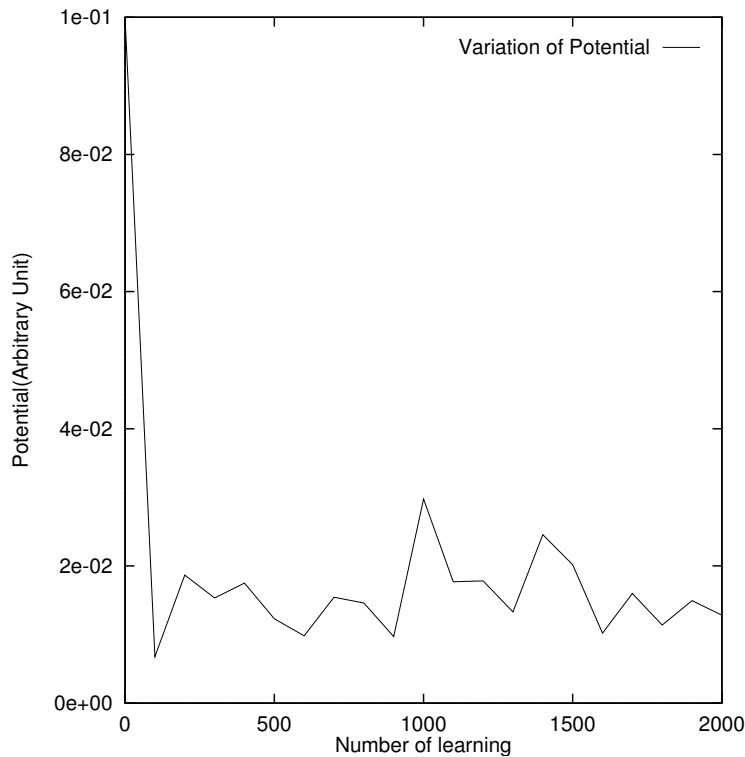


Figure A.4. 緩和的な方法を用いない場合のポテンシャルの変化

合に出来る識別面を図 A.9 に示す．この場合には，分布の密度の濃い部分を避けるように識別面が形成される様子が，さらに顕著に現われている．

## A.5 文字認識実験

我々が提案した手法を現実的なパターン識別問題である手書き文字認識に適用する．

選択したカテゴリセットは「問，問，聞，聞，聞，聞，閉，閉，問，問」の 10 カテゴリで，これらを独自に収集した文字パターンデータベースから 150 パターン / カテゴリを学習パターンとして選択し，1 次，および 2 次モーメントによる位置と大きさの正規化を行った後，外郭方向寄与度法 [45] で特徴抽出し，さらに，JIS 第一水準漢字 2965 カテゴリ全てを用いた判別分析により作られた次元圧縮マトリックスにより 64 次元に圧縮した特徴を用いた．評価データとしては，使用した文字パターンのうち，最も品質の悪いデータセット 20 パターン / カテゴリを用いた．このような問題設定では，

- 類似したカテゴリであり，識別が困難である
- 学習パターンが次元数の 2 倍以上あるため，学習パターンだけでも線形分離可能ではないと予測される．[65]

などの点から，従来の学習手法の適用が困難である．さらに，純然とした文字認識の問題として見た場合にも判別分析による次元圧縮では，特徴空間が直交しな

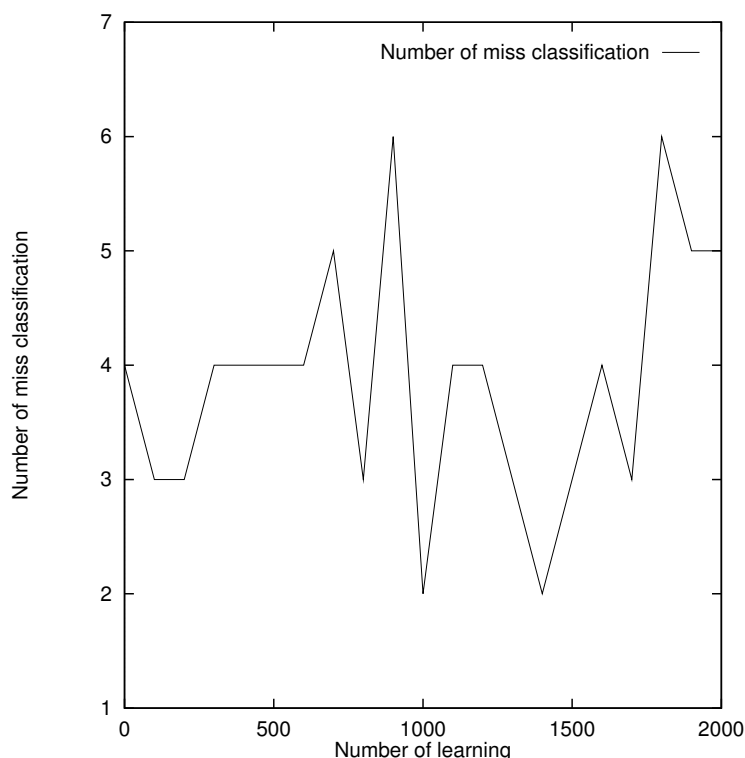


Figure A.5. 緩和的な方法を用いない場合の誤識別数の変化

いため、投影距離、マハラノビス距離などによる高度化は難しい。

提案手法の適用にあたっては、予備実験の結果から  $\rho_0 = 0.01, c_{rate} = 0.999999$  とした。また、比較のために perceptron 的な評価関数を用いた手法で、学習がそれ以上進まなくなるという意味で収束させることが可能なアルゴリズムのうち、Chang[63] の緩和型 perceptron、及び LVQ2.1 について実験を行った。LVQ2.1 は、初期値を  $k$ -平均法で与え、同様なカテゴリでの他の実験結果を踏まえ [5]、テンプレート数を 4 個とした。par 実験結果を表 A.3 に示す。

Table A.3. 手書き漢字学習終了時の識別率の比較

	学習前 (重み付きユークリッド)	緩和型 perceptron	LVQ2.1	提案手法
学習	91.06%	82.20%	92.90 %	91.13%
未知	78.50%	68.50%	66.50 %	81.50%

緩和型 perceptron では、正解の近傍からも離れてしまっているように振る舞っているが提案手法では認識率をわずかながら上昇させている。問題の難しさを考慮すると、未知パターンについてもよりよい解を探索できているという意味で、提案手法の有効性を裏付けている。さらに、学習パターンの認識率では LVQ2.1 が最も高い値を示しているが、未知パターンについてはかなり低い値になってお

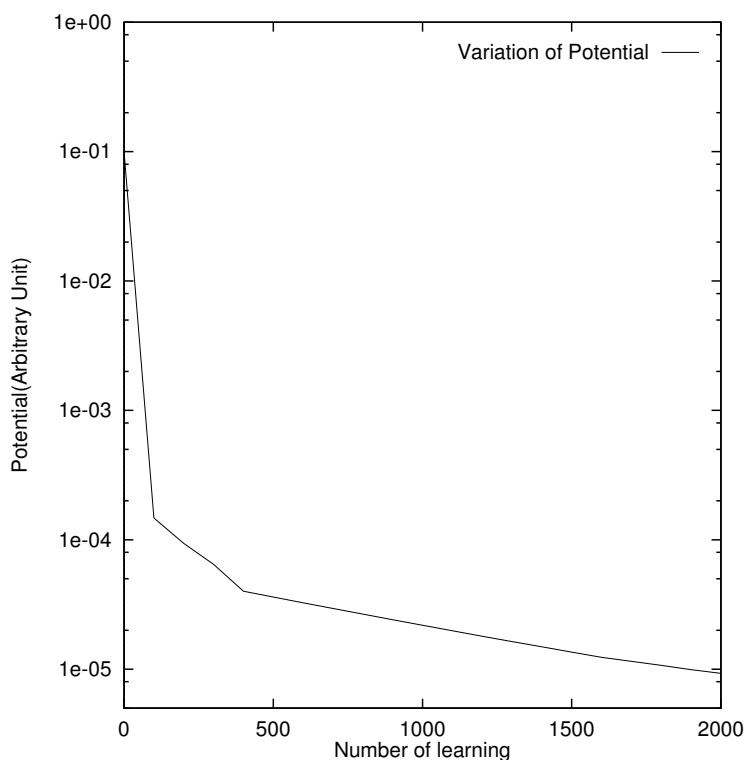


Figure A.6. 緩和的な方法を用いた場合のポテンシャルの変化

り、完全に過学習を起こしているものと見られる。これらの実験から提案手法が、

1. 分離不能なパターン集合に対しても良好な辞書に収束する
2. 困難な問題にも過学習の危険が少ない

ことが示されたと言えよう。

## A.6 まとめと今後の課題

サンプルの重要度を考慮した評価関数の導入により、分離が不可能な場合にも、学習を Bayes 識別面近傍に収束させることが出来る識別関数の学習的構成法を提案し、重み付きユークリッド距離に適用した場合のシミュレーションおよび手書き文字の認識実験によってその効果を確認した。提案手法を用いることで従来の学習的手法では高度化の困難な場合にも高い識別率を達成する識別関数を作ることが出来る。

しかしながら、提案手法は実際にパターン認識辞書の作成に使用するにはまだまだ実用的とは言い難い。例えば (A.6) 式の通りに相互作用を計算することは、外郭方向寄与度法で与えられる 1536 次元という高次元の空間に対しては数値的発散のためにほとんど不可能と見てよい。この困難を回避するためにはパターン間に働く相互作用として Gauss の法則を満たすという条件以外のどのような相互作用の形式が有効かを検討することが必要であると考えられる。

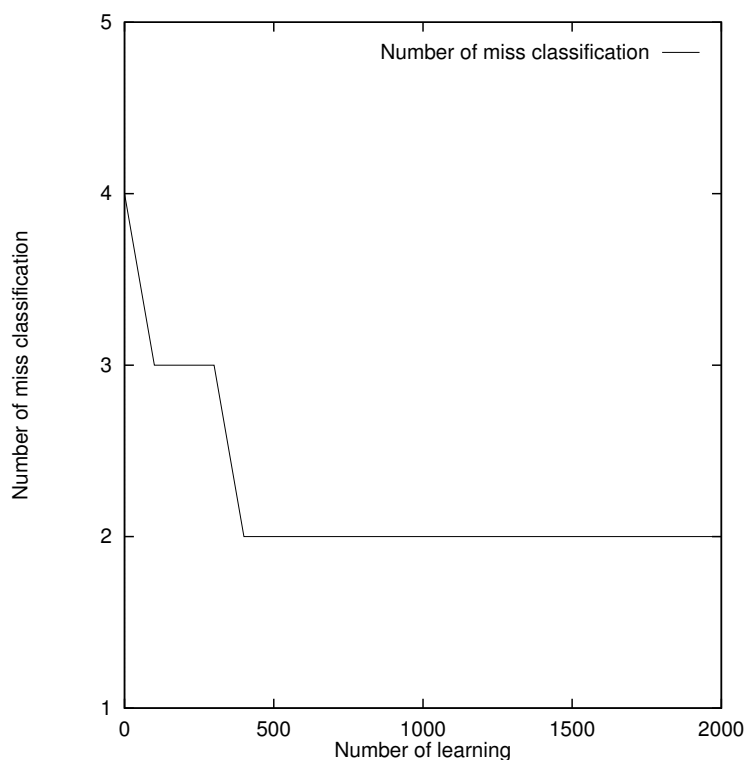


Figure A.7. 緩和的な方法を用いた場合の誤識別数の変化

移動のパラメーター  $\rho$  と緩和スケジュールについては、実験的にせよ理論的にせよ組織立った計算法が必要であり、さらに検討を要する。

また、本章においては緩和型の学習則に対して Coulomb 型評価関数を適用したがこの方法では local minima を避けることはできず学習ダイナミクスとしてノイズを入れた、つまり、シミュレーテッドアニーリングの様なアルゴリズムを採用することも考慮すべきであろう [76][53]。

理論的には、提案手法はポテンシャル関数法を用いて、距離関数を学習する方法と位置づけることも出来、この方向からの改良も考えられる。実験的にも識別超平面に関するダイナミクスについて、さらに詳細に検討する余地がある。特に、分布が異なる場合についてどの程度の分布のズレが理論の適用限界を破るかについてはより多くのケースについて実験的な検証が必要であると考えられる。

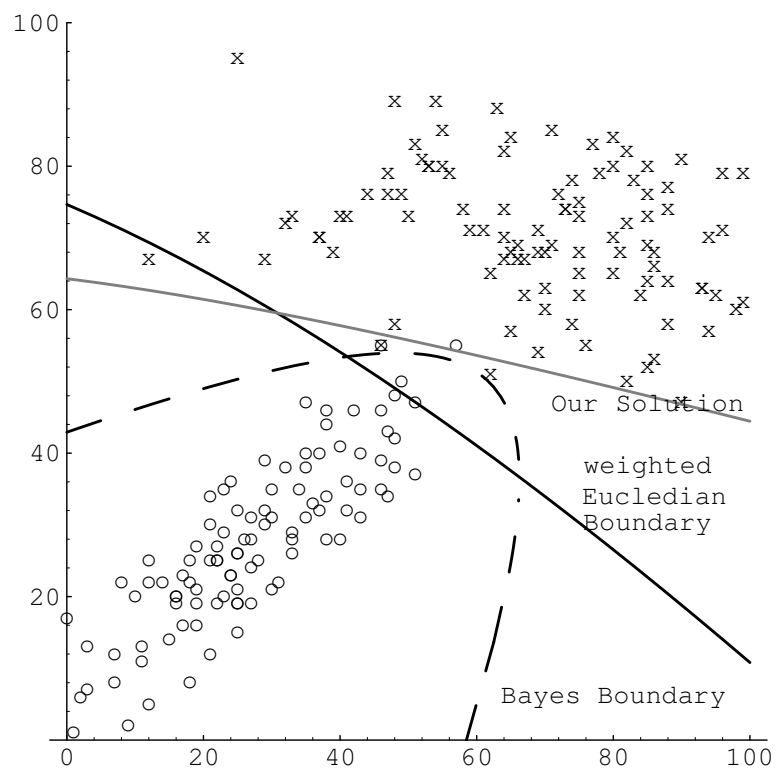


Figure A.8. 提案手法によって得られる識別面

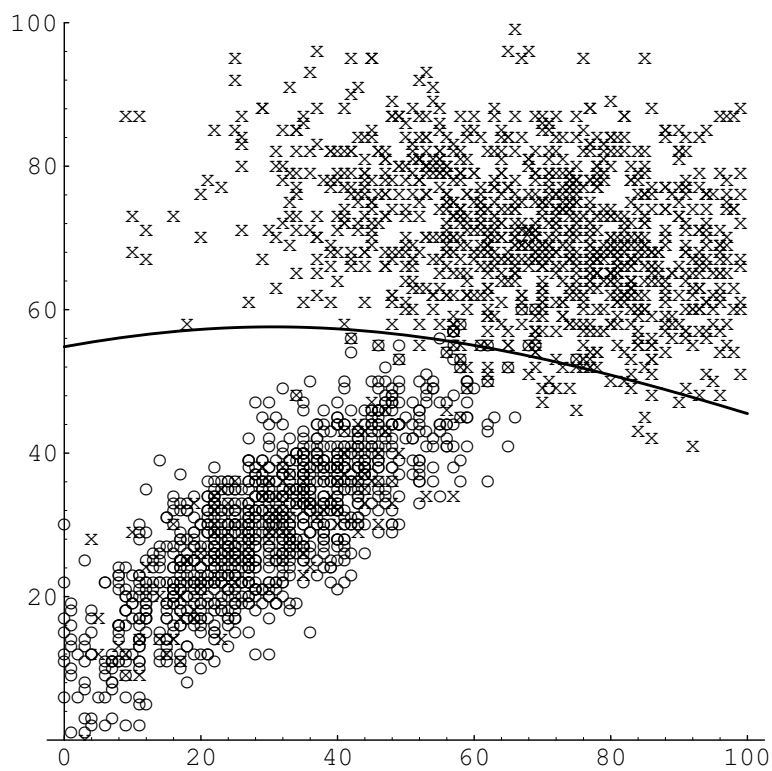


Figure A.9. 学習データが十分に多い場合の識別面

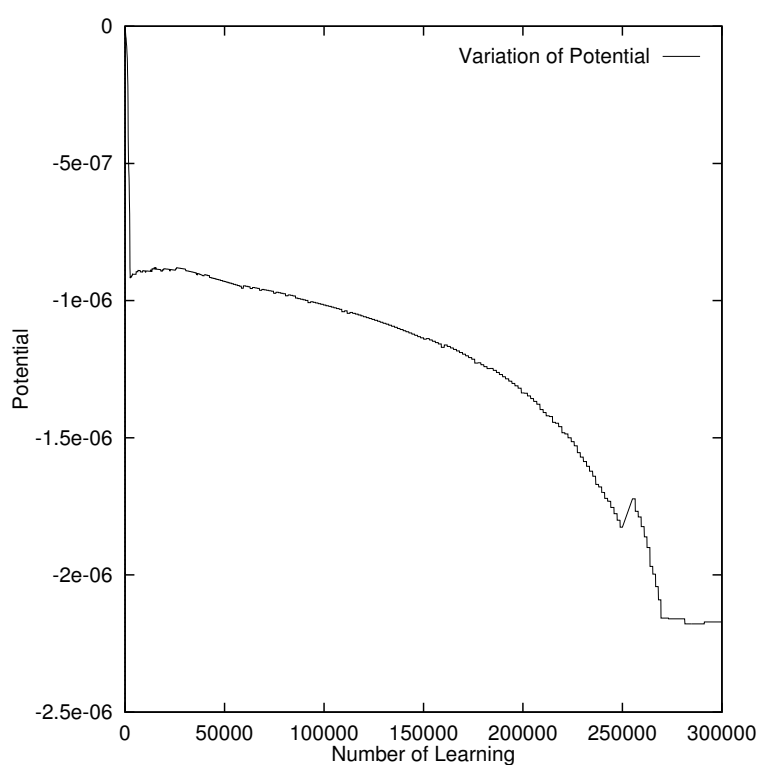


Figure A.10. 文字認識辞書に適用した場合のポテンシャルの変化

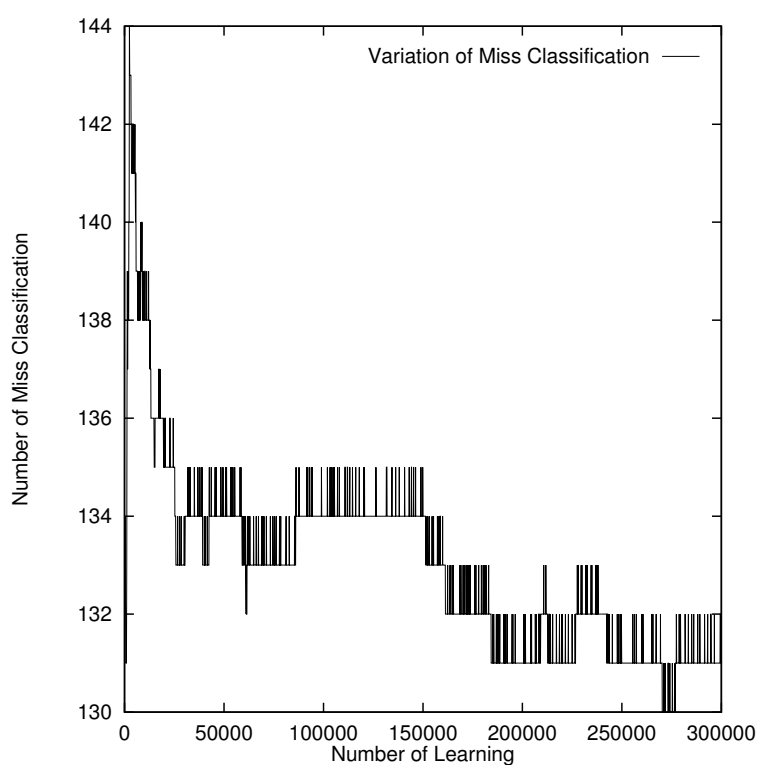


Figure A.11. 文字認識辞書に適用した場合の誤識別数の変化



## Appendix B

# 画像再構成のための Affine Mutation

### B.1 はじめに

遺伝的アルゴリズム（以下 GA）は解の評価が可能であり，解候補がビット列で表現することが可能であれば適用可能であるという広範囲の応用可能性により，各分野で注目を集めている解探索手法である．

画像再構成もしくは画像復元問題に対しては，画像そのものを遺伝子とみなし，画像から得られたデータとの二乗誤差を評価関数とするというスキームで劣化画像からの復元問題，陰影からの形状復元問題，X 線 CT からの画像再構成問題，文字認識用特徴からの文字画像再構成問題など多くの対象に応用されてきた．

しかしながら，筆者らが指摘したように [32]，GA を画像処理に用いた場合に，どのような突然変異手法が有効かという問題は一切検討されていなかった．本章においては，画像再構成問題に対して一般的に有効と考えられる突然変異手法，アフィン突然変異を提案する．

以下，B.2.2 では GA を画像再構成問題に適用する際の一般的なスキームについて述べ，従来の応用について概説する．B.2.3 では提案手法についてその妥当性を検討する．B.2.4 では特徴ベクトルからの文字画像再構成実験の結果を示し，提案手法の有効性と性質を調べる．

### B.2 GA による画像再構成アルゴリズムと従来の突然変異手法

#### B.2.1 GA による画像再構成

まず，一般の画像再構成問題を考える．雑音がない場合の画像の空間  $x, y$  における原画像  $f(x, y)$  に対する観測過程を

$$v(z) = \int \int A(z; x, y) f(x, y) dx dy \quad (\text{B.1})$$

とする．ここで， $A(z; x, y)$  は観測過程を表わすオペレーターである．このときの画像再構成問題は，任意のデータ空間  $z$  上の観測データ  $v(z)$  から得られる推定画

像  $f_0(x, y)$  と原画像  $f(x, y)$  の誤差,

$$J = \int \int |f(x, y) - f_0(x, y)|^2 dx dy \quad (\text{B.2})$$

を最小化することが目的となる。観測オペレーター  $A(z; x, y)$  についての解析的な表現が得られる場合には、射影フィルタなどの解析的手法が用いられ、多くの成果をあげている [13]。しかしながら、多くの問題においてはオペレーター  $A(z; x, y)$  の解析的な表現を得ることが困難もしくは不可能であり、 $A(z; x, y)$  の近似表現を求めるか、推定画像を探索的に求める方法が用いられている。

GA を用いた画像再構成手法は探索的手法の一つと位置づけられる。GA を画像再構成に応用する場合には、画像のビットマップそのものを遺伝子とみなし、 $N$  個の個体  $g_n$ , ( $n = 1, \dots, N$ ) から得られた観測データ  $z_n$  と、再構成対象の観測データの二乗誤差

$$J_a = |z - z_n|^2 \quad (\text{B.3})$$

を最小化する手法として定式化される。

### B.2.2 画像再構成問題における突然変異手法

ビットマップを遺伝子型として用いた GA では、交叉は各個体の有効な部分を交換する操作、突然変異は初期個体に含まれなかった部品を調達するための操作と位置づけることが出来る。

画像再構成問題に GA を適用する際には、画像の基本的な性質に立脚した遺伝的オペレーターを用いることが重要となる。交叉については、高津ら [39] が 2 次元交叉と呼ばれる手法を提案し、画像復元実験を通して有効性を確認している。この方法は画像のマルコフ性に着目した手法であり、2 次元領域を交換することでそれを実現している。

一方で突然変異手法としては、多くの場合、画素に対する点突然変異が用いられている。通常の場合であれば、点突然変異が妥当である場合もあるが、この場合には遺伝子そのものが画像としての意味を持つために、必ずしも妥当とは考えられない。簡単な例として、2 値の文字画像を扱う場合を考えると、点突然変異は多くの場合無意味に黒点もしくは白点を発生させることになり、2 値画像の再構成に大きな影響を及ぼすとは考えづらい。次節ではこの問題を考慮した突然変異手法を提案する。

## B.3 Affine Mutation

突然変異の目的は初期個体群の組み合わせで用意できなかった部品を調達することと位置づけることが出来る。一方で自然画像の最も基本的な特徴はマルコフ性、すなわち、一つの画素の近傍画素は類似した値を取るという性質である。この性質は物体の凝集性に起因し、エッジなど特殊な領域を除く画像の大部分で成立する。

故に、突然変異によって新たに個体群に組み込まれる部品もマルコフ性を持っていなくてはならない、そして、個体群と類似し、かつマルコフ性を維持した部品を導入する最も簡単な手段は任意の個体から連続的な変換によって新たな個体を生成することである。こうした考察から我々は遺伝子 (画像) にアフィン変換を施すことで新たな個体を生成する手法、アフィン突然変異法を提案する。次節では OCR (光学的文字認識装置) で用いられる特徴ベクトルから文字画像を再構成する問題を例にして提案手法の有効性を実験的に示す。

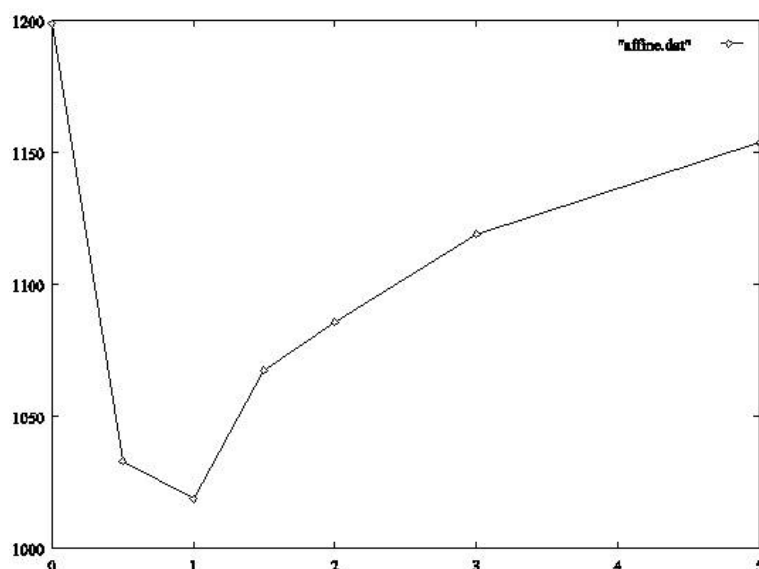


Figure B.1. 提案手法による到達二乗誤差

## B.4 文字画像再構成実験

OCR で用いられる特徴抽出系は強い非線型性を持つために特徴ベクトルからの文字画像再構成問題はその重要性を認識されつつも長く未解決の問題であったが、GA を用いた探索的解法によって近年部分的に解決された本節ではこの問題に提案手法を適用し、その有効性を確認する。実験では特徴抽出系として拡張外郭方向寄与度特徴を用い、個体数を 1000 個体の文字データを用いて実行した。なお詳細なアルゴリズムは第 4 章を参照されたい。提案手法を適用するに当たり、今回はアフィン変換として並進のみを用い並進の大きさを決定するために、ガウス分布  $N(0, \sigma^2)$  を置き、乱数  $x$  が  $x < N(0, \sigma^2)$  を満たすときに幅  $x$  だけの並進を行うこととした。結果的に  $\sigma$  の値が突然変異率に対応する量になる。実験結果を図 B.1 に比較のために行った点突然変異の実験結果を図 B.2, B.3 に示す。グラフはそれぞれ横軸に突然変異率もしくはそれに対応する  $\sigma$ 、縦軸に最終的な個体集団の自乗誤差の標的的特徴ベクトルに対する最小値の 1000 回の試行の平均値を示している。これらの結果は提案手法が最適値で点突然変異を上回るだけでなく、比較的広い範囲の  $\sigma$  で安定に改善効果を出すことを示している。

## B.5 考察

提案手法の法がよいのは画像の局所性に着目した手法だからである。突然変異が起った場合、点突然変異の場合には確実にスキーマは破壊されるがアフィン突然変異の場合は文字画像を構成する位置ずれした部品が供給されることになる。これが提案手法が効果的で、安定な理由である。

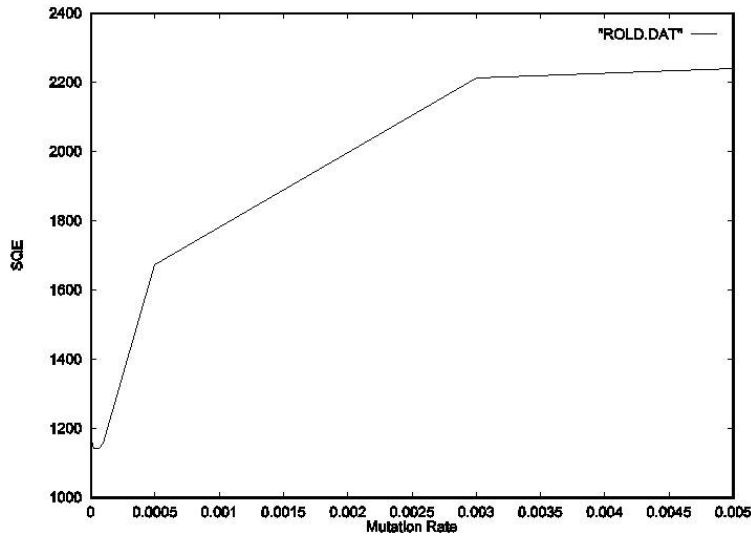


Figure B.2. 従来手法による到達二乗誤差

## B.6 まとめと今後の課題

画像再構成問題に GA を適用する際に一般的に有効と考えられる新しい突然変異手法, Affine Mutation を提案し, 特徴ベクトルからの文字画像再構成実験を通して有効性を確認した. 今回は変異を平行移動だけに限定したが, スケール変換や回転を考慮した場合についても検討が必要である.

また, 陰影からの形状復元 [90] などの他の再構成問題に対しての有効性も実験的に調べる必要があるだろう.

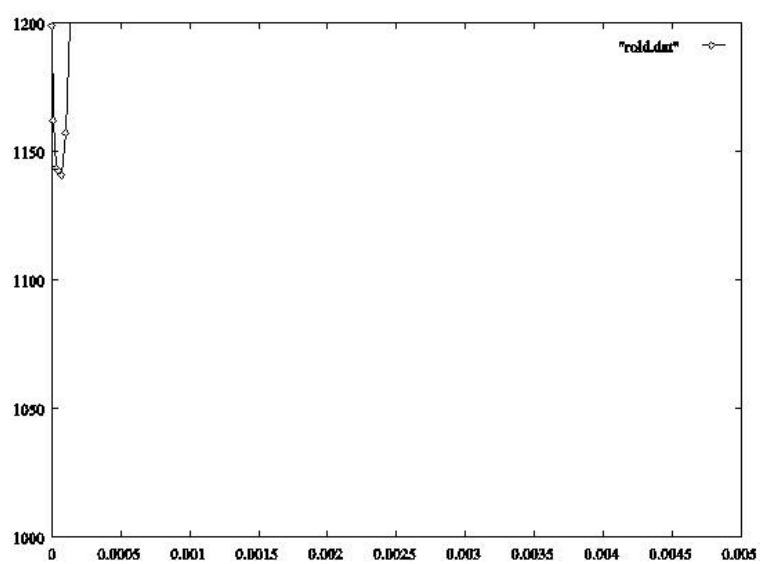


Figure B.3. 従来手法による到達二乗誤差 (拡大)



# 謝辞

本学位論文執筆にあたり、多数の方々の支援を受けました。まず、これまで私を育ててくれた両親に感謝いたします。

主査を引き受けていただき、御指導を頂いた早稲田大学橋本周司教授に感謝いたします。橋本教授の御指導でこれまでに無い視野を広げることが出来ました。さらに副査を引き受けていただいた早稲田大学 森島繁生教授、中島啓幾教授、小松尚久教授に感謝いたします。

また、三枝亮博士、市野正嗣君を始めとした、早稲田大学応用物理学科橋本研究室、早稲田大学 CS 学科小松研究室の学生諸君に感謝いたします。

さらに、私の上司として御指導を頂いた(株)NTTデータの荒川 弘熙博士(現ニューソン株式会社)、阿部一朗氏(現 NTT データ先端技術(株))、木田博巳氏(現 NTT データ PMO)、武川直樹博士(現東京電機大)、岩城修博士、曾根岡昭直博士(現 NTT 第3部門)、島崎勝美氏、大塚作一博士(現鹿児島大学)、関根純博士、三部靖夫博士、に感謝いたします。

また、同僚として日ごろから有益な議論を頂いた(株)NTTデータの同僚諸氏、特に松永務博士、内山俊郎博士(現 NTT サイバーソリューション研究所)に感謝します。

さらに、学会などで本研究に関わる貴重な議論の機会を頂いた東京大学 池上高志准教授、統計数理研究所 伊庭 幸人准教授、九州大学 内田誠一教授、東北大学 大町真一郎准教授、東北文化学園大学 大町方子講師、大阪大学 菊池誠教授、大阪府立大学 黄瀬浩一教授、高知工科大学 木村義正教授、慶応大学 斉藤英雄教授、国立情報学研究所 佐藤真一教授、東京大学 佐藤洋一准教授、京都大学 鷲見和彦教授、大阪大学 時田恵一郎准教授、Bonn 大学 Manfred Fiebig 教授、筑波大学 福井和広准教授、山口大学 浜本義彦教授、法政大学 若原徹教授、に感謝します。

そして、修士課程で御指導頂いた元埼玉大学大学院理学研究科物理学専攻 川崎温教授、宮本芳子助手、学士論文の指導教官としてご指導いただいた中央大学理工学部物理学科 杉本秀彦教授に感謝いたします。研究に関する基本的な態度はこの方々より教わりました。

最後に私の健康上、精神面の支えになってくれた妻、坂野貴子博士に感謝いたします。





# References

- [1] アイゼルマン,et.al, パターン認識と学習制御, 共立出版,(1978)
- [2] 甘利俊一編, 生命・生物科学の数理, 岩波講座 応用数学 [対象 8], 岩波書店, (1993)
- [3] 石井健一郎, 変形文字を用いた識別辞書の設計, 信学論, Vol.J72-D-II, pp. 669-677, (1989)
- [4] 石井健一郎, 変形文字パターン発生法とその応用, 信学論, Vol.J66-D, pp. 1270-1277, (1983)
- [5] 出本, 阿部, 手書き漢字認識におけるサブカテゴリ辞書作成方式の検討, 情報処理学会第 41 回全国大会, 4L-4, (1990)
- [6] 出本浩, 横塚志行, 木田博巳, LVQ を用いた手書き漢字認識辞書作成方式の検討, 信学技報 PRU91-50, (1991)
- [7] 市野将嗣, 坂野鋭, 小松尚久, 核非線形相互部分空間法による話者認識, 電子情報通信学会論文誌 D-II, J88-D-II, (2005)
- [8] 市野将嗣, 坂野鋭, 小松尚久, 話者認識における核非線形相互部分空間法の適用に関する検討, MIRU2006 サテライトワークショップ部分空間法研究会 2006, pp.1-6, (2006.7)
- [9] 市野将嗣, 坂野鋭, 小松尚久, クラスタリングを用いた核非線形相互部分空間法の処理量削減手法, 電子情報通信学会論文誌 D, Vol.90, D No.8 pp.2168-2181
- [10] 石山壘, 濱中雅彦, 坂本静生, 顔の三次元見えモデルを用いた任意姿勢・照明変動下の顔画像認識, 信学論, Vol.J88-D2 No.10 pp.2069-2080, (2005)
- [11] 梅田三千雄, 手書き漢字認識の個性的なサーベイ, 郵政研シンポジウム「文字認識技術の現在・過去・未来」資料, (1993)
- [12] 梅田三千雄, 手書き漢字認識研究の歩み, 信学技報 PRU95-120, (1995)
- [13] 小川英光, 信号と画像の復元, 信学誌, Vol.71, 5-8, p.491, p.593, p.739, p.828, (1988)
- [14] 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)
- [15] 片桐, 学習ベクトル量子化と多層パーセプトロンの統一的处理, 信学技報 MBE88-72, (1988)

- [16] 川谷隆彦, 距離関数の学習による手書き数字認識, 信学論 (D-II), J76-D, No.9, pp. 1851-1859, (1993)
- [17] 北野宏明, 遺伝的アルゴリズム, 産業図書, (1993)
- [18] 木村義正, 文字認識における特徴の可視化システム, 画像電子学会 第34回年次大会, (2006)
- [19] 倉掛正治, 変形文字パターンを用いた手書き文字認識系の評価, 信学論, Vol.J74-D-II, pp. 1208-1216, (1991)
- [20] 例えば, T. コホネン, システム論的連想記憶, p41, サイエンス社. (1980) など, ただし, 久保, 大学演習 熱学・統計力学, 第5章, 裳華房, (1961) にはより詳細な記述がある.
- [21] T. コホネン, 自己組織化と連想記憶, シュプリンガー・フェアラク東京, (1993)
- [22] 例えば【特集: パターン認識の最近の動向】, O plus E, 2003年10月号
- [23] Ryo SAEGUSA, A Study of Nonlinear Principal Component Analysis Using Neural Networks, 早稲田大学博士論文, (2005)
- [24] 坂井利之編, 文字・図形の認識機械, 共立出版, 1967
- [25] 坂野鋭, 横塚志行, 木田博巳, 動力学モデルによる識別超平面の最適化, 信学技報 IE91-129, (1992)
- [26] 坂野 鋭, 横塚志行, 木田博巳, 非線型主成分分析による投影距離法 - 手書き文字認識への応用, 画像の認識と理解シンポジウム MIRU'92.(1992)
- [27] 坂野鋭, 木田博巳, 遺伝的アルゴリズムによる特徴空間からの文字画像再構成, 信学技報 PRU94-99,(1995)
- [28] 坂野 鋭, 出本 浩, 木田 博巳, 手書き文字認識における重みつき距離関数の学習法 -重みつき学習ベクトル量子化-, 画像の認識と理解シンポジウム MIRU'94, pp. I-131-136
- [29] 坂野鋭, 木田博巳, 武川直樹, 遺伝的アルゴリズムの文字画像生成能力に関する検討, 信学技報 PRU95-118,(1995)
- [30] 坂野, 宮本, 拡張外郭方向寄与度法による手書き漢字認識, 信学総大,(1995), D-552
- [31] 坂野鋭, 木田博巳, Coulomb 相互作用モデルによる識別関数の学習的構成法, 信学論 D II ,J77-D-II,pp. 1806-1816 , (1994)
- [32] 坂野 鋭, 斎藤英雄, パターン認識における遺伝的アルゴリズム, その誘惑と悪夢, 信学会誌, Vol.79, No. 10, pp.961-966, (1996)
- [33] 坂野 鋭, バイオメトリック個人認証技術の動向と課題, 信学技報,PRMU-99-29, (1999.6)
- [34] 佐瀬幹哉, 山懸 陽, 小杉幸夫, 遺伝的アルゴリズムによるニューラルネットの問い合わせ学習, 信学論, Vol.J79-D-II, pp.960-968,(1996)

- [35] 佐藤敦, 津雲淳, 増殖型ニューラルネットを用いたひらがな認識, 信学技報 NC93-26, pp.17-24, (1993)
- [36] 佐藤敦, 山田敬二, 一般化学習ベクトル量子化の提案, 信学技報 NC95-26, pp.17-24, (1995)
- [37] 特集 能動学習, 情報処理学会誌, Vol. 38, No. 7, (1997)
- [38] 孫, 安倍, 根本, 改良型方向線素特徴量および部分空間法を用いた高精度な手書き文字認識, 信学技報, PRU94-31, p.33, (1994)
- [39] 高津, 澤井他, 画像のベイズ復元への遺伝的アルゴリズムの適用, 信学論 (D-II), J-77-D-II, p.1768, (1994)
- [40] 津田宏治, ヒルベルト空間における部分空間法, 信学論 (D-II), Vol. J82-D-II, No.4, pp. 592-599, (1999)
- [41] 鶴岡, ほか, 加重方向指数ヒストグラム法による手書き漢字ひらがな認識, 信学論 (D), J70D,7
- [42] 友野典男, 『行動経済学 経済は「感情」で動いている』, 光文社, (2006)
- [43] 中田, 尾川, GA による投影からの画像再構成, 第 25 回画像工学コンファレンス論文集, p.137, (1994)
- [44] 例えば 中野, ニューロコンピュータの基礎, (コロナ社), (1990)
- [45] 萩田, 内藤, 増田, 外郭方向寄与度法による手書き漢字の識別, 信学論 (D), J66-D, No.10 (1983)
- [46] 橋本新一郎編著, 文字認識概論, オーム社, (1977)
- [47] 福井和広, 山口修, カーネル非線形制約相互部分空間法による物体認識, 電子情報通信学会論文誌 (D-II), vol. J88-D-II, no.8, pp.1349-1356, (2005)
- [48] 藤巻遼平, 矢入健久, 町田和雄, カーネルヒルベルト空間における正準角を利用した宇宙機異常検知法, 第 20 回人工知能学会全国大会予稿集, (2006)
- [49] R. ベンローズ, 皇帝の新しい心, みすず書店, (1994)
- [50] D.M.Hawkins, 多変量解析の理論と実際, MPC 出版, (1988)
- [51] 前田英作, 村瀬 洋, カーネル非線型部分空間法によるパターン認識, 信学論 (D-II), Vol. J82-D-II, No.4, pp. 600-612, (1999)
- [52] 前田賢一, 渡辺貞一, 局所構造を導入したパターン・マッチング法, 信学論 (D), Vol. J68-D, No. 3, pp. 345 - 352, (1985)
- [53] 松永, 阿部, 木田, シミュレーテッドアニーリング法を用いた文字認識辞書の最適化, 信学技法, PRU90-39 (1990) .
- [54] 村瀬洋, S. Nayer, 2次元照合による3次元物体の認識とその学習 - パラメトリック固有空間表現 -, 信学技報, PRU93-120, (1994)
- [55] 森俊二, 画像認識入門, オーム社, (1988)
- [56] 山口修, 福井和宏, 前田賢一, 動画像を用いた顔認識システム, 信学技報, PRMU97-50, (1998-06)

- [57] 若原 徹 , "クラスタリングにおける動的凝集モデル" , 信学技法 PRU91-26, (1991)
- [58] 渡辺慧 , 認識とパタン , 岩波新書 , (1978)
- [59] B.G.Batchelor, "Practical Approach to pattern Classification," Plenum Press,(1974)
- [60] E. B. Baum, "Neural Net Algorithms That Learn Polynomial Time from Example and Queries," IEEE Trans. Neural Networks, Vol. 2, No. 1, pp. 5-14, (1991)
- [61] P. B. Belhumeur, D. Kriegman, "What is the set of Images of an Object Under All Possible Illumination Conditions?," International Journal of Computer Vision. Vol. no. 28, Issue No. 3, pp245-260, (1998)
- [62] C. J. C. Burges, B. Schoelkopf, "Improving the Accuracy and Speed of Support Vector Machine," Advances in Neural Information Processing, Vol. 9, pp. 375-381, (1997)
- [63] C.L.Chang, "The accelerated relaxation method for linear inequalities", IEEE Trans.Comput., C-20,2, (1971)
- [64] R.O.Duda and P.E.Hart, "Pattern Classification and Scene Analysis", Wiley-Interscience, (1973)
- [65] T.M.Cover,"Geometrical and statistical properties of system of linear inequalities with applications in pattern recognition," IEEE Trans. Electron. Comput.,EC-14,3,p326,(1965)
- [66] K. I. Diamantaras, et. al. "Principal component neural network," Willey New York, (1996)
- [67] 例えば , R. P. Feynman, et.al. , Feynman Lectures on Physics, (Addison Wesley,(1964)
- [68] R. Gnanadesikan, "Method for Statistical Data Analysis of Multivariate Observation," Jon Willey and Sons, (1977)
- [69] D. E. Goldberg, "Genetic Algorithms - in search, Optimization & Machine Learning," Addison-Wesley, (1989)
- [70] D. B. Graham and N. S. Allinson, "Characterizing Virtual Eigensignatures for General Purpose Face Recognition," in H. Wechsler, et al. ed. "Face Recognition From Theory to Applications", Springer Verlag (1998).
- [71] T. Hastie, el. al. "Principal curves," JASA, Vol. 84, pp. 502-516, (1989)
- [72] J. H. Holland, "Adaptation in Natural and Artificial Systems," Univ. of Michigan Press, (1975).
- [73] Y. S. Huang, K. Liu, C. Y. Suen, A. J. Shie, I. I. Shyu, M. C. Liang, R. Y. Tsay and P. K. Huang, "A Simulated Annealing Approach to Construct Optimized Prototypes for Nearest-Neighbor Classification," in Proc. of 13th ICPR Vol. IV, pp. 483-487

- [74] J-N. Hwang, J. J. Choi, S. Ho, R. J. Marks Jr., "Query Based Learning Applied to Partially Trained Multilayer Perceptrons," *IEEE Trans. Neural Networks*, Vol. 2, No. 1, pp. 131-136, (1991)
- [75] Chihiro Iga, Toru Wakahara, "Character Image Reconstruction from a Feature Space Using Shape Morphing and Genetic Algorithms," In. Proc. IWFHR'04, pp. 341 - 346, (2004)
- [76] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, " Optimization by simulated annealing, " *Science* , 220 , pp.671-680 (1983) .
- [77] T. Kohonen, G. Barna, and R. Chrisley, "Statistical Pattern Recognition with Neural networks : Benchmarking Studies," *IEEE Proc. ICNN*, Vol. 1, pp. 61-68, (1988)
- [78] M. kudo, K. Mizukami, Y. Nakamura and M. Shinbo, "Realization of membership queries in character recognition," *Pattern Recognition Letters*, Vol. 17, pp. 77-82, (1996)
- [79] T. Kohonen, "Improved Versions of Learning Vector Quantization," *Proceedings of International Joint Conference on Neural Networks*, pp I-545-550,(1990)
- [80] J. J. Leimer, "Design Factors in the Development of an Optical Character Recognition Machine," *IRE Trans. on Information Theory*, pp. 167-171, (1962)
- [81] Seong-Whan Lee and Hee-Heon Song, "Optimal Design of Reference Models Using Simulated Annealing Combined with an Improved LVQ3," *Proc.2nd ICDAR*, pp244-249,(1993)
- [82] Seong-Whan Lee and Hee-Heon Song, "Optimal Design of Reference Models for Large-Set Handwritten Character Recognition," *Pattern Recognition*, Vol. 27, No. 9, pp. 1267-1274, (1994)
- [83] Seong-Whan Lee and Jong-Soo Kim, "Multi-lingal, Multi-font and Multi-size Large-set Character Recognition using Self-Organizing Neural Network," *Proc.3rd ICDAR*, pp28-33,(1995)
- [84] G. J. McLachlan, "Iterative reclassification procedure for constructing an asymptotically optimal rule of allocation in discriminant analysis," *J. Amer. Stat. Assoc.*, 70, 365-369, (1975)
- [85] N. J. Nilson, "Learning Machine," Macgrowhil, (1965)
- [86] S. Nozawa, T. Takeshita, "Writer Identification by Human eyes and by Computer," *Proc. 4th International Workshop on Frontiers in Handwriting Recognition*, p.354, (1994)
- [87] E. Oja, "A simplified neuron model as a principal component analyser," *J. Math. Biology*, Vol. 15, pp.267-273, (1982)
- [88] B. Moghaddam, "Principal Manifolds and Bayesian Subspaces for Visual Recognition," In *Proc. Intl. Conf of Computer Vision*, pp. 1131-1136, (1999)

- [89] D.E.Rumelhart,  
"Learning and Generalization: The role of Minimal Networks," ATR  
Workshop Neural Net and PDP model, (1988)
- [90] S. Saito, et. al. "A Generalization of principal component analysis," Jour-  
nal of Japan Statistical Society, Vol. 18, No. 2, (1988)
- [91] B. Scheolkopf, et. al., "Nonlinear component analysis as a Kernel eigen-  
value problem," Neural Computation, Vol. 10, pp. 1299-1319,(1998)
- [92] Amnon Shashua, "Illumination and View Position in 3D Visual Recog-  
nition," Advances in Neural Information Processing, Vol. 4, pp. 404-411
- [93] P. Sollich, "Learning from minimum entropy query in large committee  
machine," Physical Review E , Vol. 53 ,p. 2061, (1996)
- [94] Kah Kay Sung and Partha Niyogi, "A Formulation for Active Learning  
with Application to Object Detection," MIT AI-Memo No. 1438, (1995)
- [95] C. Tomasi, T. Kanade, "Shape and motion from image stream under  
orthography: a factrization method," International Journal of Computer  
Vision, Vol. 9, No. 2, pp. 137-154, (1992)
- [96] O. D. Trier, A. K. Jain and T. Taxt, "Feature extraction methods for  
character recognition – A survey," Pattern Recognition, Vol. 29, No. 4,  
pp. 641-662, (1996)
- [97] M. Turk and A. Pentland, "Eigenface for recognition," J. Cogn. Neurosci.  
Vol. 3, No. 1, pp. 71-86, (1991)
- [98] V. Vapnik , "Statistical Learning Theory," Willy,(1998)
- [99] S. Watanabe and N. Pakvasa, "Subspace method of pattern recognition,"  
Proc. 1st IJ CPR, pp. 25-32, (1973)
- [100] S. Watanabe and E. Harada, "A dynamic model of clustering," in Proc.  
2nd IJ CPR, pp320-322, (1978)

# 業績リスト

本研究にかかわる業績

[学術論文]

1. 坂野 鋭, 木田博巳, Coulomb 相互作用モデルによる識別関数の学習的構成法, 電子情報通信学会論文誌 D-II, Vol. J77-D-II, pp.1806-1816, (1994)
2. 坂野 鋭, 木田博巳, 武川直樹, 遺伝的アルゴリズムによる文字識別系の解析, 電子情報通信学会論文誌 D-II J80-D-II, pp. 1687-1694, (1997)
3. 坂野 鋭, 武川直樹, 中村太一, 核非線形相互部分空間法による物体認識, 電子情報通信学会論文誌 D-II J84-D-II, pp. 1549-1556, (2001)
4. 坂野 鋭, パターン認識における主成分分析-顔画像認識を例として-, 統計数理, 第 49 巻, 第 1 号, pp.20-31, (2001)
5. Hitoshi Sakano, "Query Learning Method for Character Recognition Methods Using Genetic Algorithm," IEICE Trans. Vol.E88-D, No.10, pp.2313-2316, (2005)
6. Hitoshi Sakano, Naoki Mukawa and Taichi Nakamura, "Kernel Mutual Subspace Method and Its Application for Object Recognition," Electronics and Communications in Japan. Vol.E88, No.6, pp.45-53, (2005)

[国際会議論文 (査読あり)]

1. Hitoshi Sakano, Hiromi Kida and Naoki Mukawa, "Seeing the Character Images that an OCR system Sees - Analysis by Genetic Algorithm -," Proc. in 13th ICPR, Vol. 4, pp. 411-416, (1996)
2. Hitoshi Sakano and Naoki Mukawa, "Kernel Mutual Subspace Method for Robust Facial Image Recognition," Proc. in 4th Intl. Conf. Knowledge based Engineering System, Vol. 1, pp. 245-248 (2000)
3. Hitoshi Sakano, and Takashi Suenaga, "Classifiers under continuous observation," In Ed. T. Caeli, et al., Structural, Syntactic, and Statistical Pattern Recognition 2002, Lecture Note on Computer Science 2396, pp.798-805, 2002, Springer-Verlag Berlin Heidelberg 2002 Proc. in IAPR Intl. Workshop. SPR 2002, (2002)
4. Hitoshi Sakano, Genetic Translator: How to apply query learning to practical OCR systems, Proc. in IAPR Intl. Conf. ICPR 2002, (2002)

## [総説]

1. 坂野 鋭, 斎藤英雄, パターン認識における遺伝的アルゴリズム, その誘惑と悪夢, 電子情報通信学会誌, Vol. 79, No. 10, pp.961-966, (1996)
2. 坂野 鋭, 山田敬嗣, 怪奇!!次元の呪い - 識別問題, パターン認識, データマイニングの初心者のために -, 情報処理, Vol. 43, No. 5, pp.562-567, No.6 pp.658-663, (2002)
3. 坂野 鋭, 核非線形相互部分空間法による物体認識, 画像ラボ, pp. 22-25, 2002年6月号
4. 坂野 鋭, 末永高志, 高次元データの可視化技術, 画像電子学会誌, p251, 第32巻第3号(2003年5月)
5. 坂野 鋭, 学習機械は働いているか?, 人工知能学会誌, Vol.19, p. 664, (2004)
6. 坂野 鋭, 高次元データに語らせる技術, CADM News Letter, No.45, May. (2006)

## [招待公演]

1. 坂野 鋭, 「画像認識の問題としてのバイオメトリック個人認証技術」 AVIRG2001年第5回例会, (2001.3.26)
2. Hitoshi Sakano, Facial image recognition technology in Japan. Why Japanese researcher love subspace method, Proc. in Deuche Japan information technology Forum, (2002.4.29)
3. 坂野 鋭, 顔画像認識技術の現状と動向, 電子情報通信学会東京支部セミナー, (2002.5.10)
4. 坂野 鋭, 統計的パターン認識技術と外観検査への応用, 外観検査の自動化ワークショップ View2002, (2002.12.5)
5. 坂野 鋭, パターン認識における統計的方法, 計算機統計学会全国大会特別講演, (2003.5.23)
6. 坂野 鋭, 識別問題のモデル化と数理, 数理モデル化と問題解決シンポジウム, (2004.10)
7. Hitoshi Sakano, Biometrics under continuous observation, In Proc. ICARCV. (2004)
8. 坂野 鋭, 統計的識別の理論と実際 - テキスト分類への応用を目標に - 言語処理学会第11回年次大会チュートリアル, (2005.3.14)

## [国際会議, 査読なし]

1. Hitoshi Sakano, Optimization Method for Discriminant Hyper Surface by Dynamical Model, '93 Oji International Seminar on Complex Systems. Mar. (1993)
2. Hitoshi Sakano, Kernel Mutual Subspace Method for object Recognition, Proc. in New Trend in Optimization and Computation, (2001)



## [研究会, シンポジウム]

1. 坂野 鋭, 横塚 志行, 木田 博巳, 動力学モデルによる識別超平面の最適化, 電子情報通信学会技術報告, IE91-129, (1992)
2. 坂野 鋭, 横塚 志行, 木田 博巳, 非線形主成分分析による投影距離法, 画像の認識と理解シンポジウム MIRU'92, (1992), 電子情報通信学会, 情報処理学会
3. 坂野 鋭, 出本 浩, 木田 博巳, 手書き文字認識における重みつき距離関数の学習法 -重みつき学習ベクトル量子化-, 画像の認識と理解シンポジウム MIRU'94, (1994)
4. 坂野 鋭, 木田 博巳, 遺伝的アルゴリズムによる特徴空間からの文字画像再構成, 電子情報通信学会技術報告, PRU-94-99, (1995)
5. 坂野 鋭, 木田 博巳, 武川直樹, 遺伝的アルゴリズムの文字画像生成能力に関する検討, 電子情報通信学会技術報告, PRU-95-118, (1995)
6. 坂野 鋭, 木田博巳, 武川直樹, OCR の見ている文字を見る -遺伝的アルゴリズムによる文字識別系の解析-, 画像の認識と理解シンポジウム MIRU'96, 電子情報通信学会, 情報処理学会, (1996)
7. 坂野 鋭, OCR の見ている文字を見る - 遺伝的アルゴリズムによる文字認識装置の解析 -, 情報文化学会第 1 1 回マルチメディア研究部会研究会, (1998-5.30)
8. 坂野 鋭, 武川直樹, 質問学習による手書き数字認識, 電子情報通信学会技術報告, PRMU-98-45, (1998)
9. 坂野 鋭, 武川直樹, 核非線型相互部分空間法と顔認識への応用, 電子情報通信学会技術報告, PRMU-99-116, (1999)
10. 坂野 鋭, 武川直樹, 物体認識のための核非線型相互部分空間法, 画像の認識と理解シンポジウム MIRU2000, p.II-169, (2000)

## [全国大会]

1. 坂野 鋭, 宮本 信夫, 拡張外郭方向寄与度法による手書き漢字認識, 電子情報通信学会総合大会発表論文集, (1995.3)
2. 坂野 鋭, 画像再構成のための新しい突然変異手法 -Affine Mutation-, 情報処理学会第 59 回全国大会講演論文集 第 2 分冊 p.265, (1999.9)
3. 坂野 鋭, 大塚 作一, 三部 靖夫, 武川 直樹, 物体認識における時系列観測データの分布間類似度の性質に関する考察, 2006 年度 画像電子学会 第 34 回年次大会, (2006.6.23)

## その他の業績

## [著書]

1. 共著, e-コマースシステム技術体系, フジテクノシステム (2001)
2. Arata Sato, Takashi Suenaga, Hitoshi Sakano, Medical data mining using cluster discriminant analysis, In Ed. H. Motoda, Active Mining, IOS press, (2002)
3. 共著, 情報セキュリティ技術体系, フジテクノシステム (2004)

## [学術誌論文]

1. 末永高志, 佐藤新, 坂野 鋭, クラスタ構造に着目した特徴空間の可視化法-クラスタ判別法-, 電子情報通信学会論文誌 D-II, J85-D-II, No. 5, pp.785-795.(2002)
2. 鈴木賢一郎, 中嶋正臣, 坂野 鋭, 三部靖生, 大塚作一, 動き方向ヒストグラム特徴を用いた映像データからのカット点検出法, 電子情報通信学会論文誌 D-II, J86-D-II, No.4, pp.468-478, (2003)
3. 三枝亮, 坂野 鋭, 橋本周司, 主成分順位を保存する非線形主成分分析, 電子情報通信学会論文誌 D-II, J86-D-II, No.7 pp.943-950, (2003)
4. Ryo Saegusa, Hitoshi Sakano and Shuji Hashimoto, "Nonlinear Principal Component Analysis to Preserve the Order of Principal Components", Neurocomputing, no.61, pp.57-70, 2004.
5. 市野将嗣, 坂野 鋭, 小松尚久, 核非線形相互部分空間法による話者認識, 電子情報通信学会論文誌 D-II, J88-D-II, No. 8, pp.1331-1338, (2005)
6. Ryo Saegusa, Hitoshi Sakano and Shuji Hashimoto, "A Nonlinear Principal Component Analysis on Image Data", IEICE Trans. Vol.E88-D, No.10, pp.2242-2248, 2005
7. 市野将嗣, 坂野 鋭, 小松尚久, クラスタリングを用いた核非線形相互部分空間法の処理量削減手法, 電子情報通信学会論文誌 D, Vol.90,D No.8 pp.2168-2181, 2007

## [国際会議, 査読あり]

1. Yoko Higashi, Arata Sato, Hirotaka Higuchi, Hitoshi Sakano, Toshihiko Morimoto, Tsutomu Matsunaga, Keisuke Ishii and Masaaki Muramatsu, Semi-Automated Calling System for SNPs (single nucleotide polymorphisms), In Proc. Intl. Conf. of Intelligent Systems in Molecular Biology, A249, Aug. (2002)
2. Takashi Suenaga, Arata Sato and Hitoshi Sakano, Cluster Discriminant analysis for feature space visualization, Proc. in Intl. Conf. IEE KES 2002, (2002)
3. Arata Sato, Takashi Suenaga and Hitoshi Sakano, Outlier Detection using Cluster Discriminant Analysis -Data mining from Medical Data-, Proc. in Intl. Conf. IEE KES2002, (2002)

4. Ken-ichiro Suzuki, Masaomi Nakajima, Hitoshi Sakano, Yasuo Sambe and Sakuichi Otsuka, Abrupt Shot Boundary Detection from Video Sequence Using Motion Direction Histogram Feature In Proc. Machine Vision and Applications. (2002)
5. Hiroyuki Tanaka, Hitoshi Sakano and Sakuichi Otsuka, Multi-category Image Retrieval using Parametric Mixture Model, In Proc. ICPR, (2004)
6. Ryo Saegusa, Hitoshi Sakano and Shuji Hashimoto, "A nonlinear principal component analysis on image data", 2004 IEEE International Workshop on Machine Learning for Signal Processing, Sao Luis, Brazil, Sep 29 - Oct 1, 2004
7. Masatsugu Ichino, Hitoshi Sakano, Naohisa Komatsu, Speaker recognition using Kernel Mutual Subspace Method, In Proc. ICARCV. (2004)
8. Masatsugu Ichino, Hitoshi Sakano and Naohisa Komatsu, Multimodal biometrics of lip movements and voice using kernel fisher discriminant analysis, In Proc. ICARCV, P.1552, December 2006

## [学会誌，専門誌解説]

1. 坂野 鋭, バイオメトリック個人認証技術の展開, 日本データ通信, pp. 22-25, 2000年6月号
2. 坂野 鋭, 中村逸一, 認証技術の現在と未来, 情報処理, Vol. 41, No. 7, pp.816-822, (2000)
3. 坂野 鋭, 情報システムにおける個人認証の現状と課題, 映像情報メディア学会誌, Vol. 55, No. 2, pp.185-189, (2001)
4. 坂野 鋭, バイオメトリクス・個人認証のためのパターン認識 OplusE, 2003年10月号
5. 小松 尚久, 内田 薫, 坂野 鋭, 和田 誓一, 池野 修一, バイオメトリクスの精度評価, 情報技術標準 Newsletter, p. 2, No.60, (2003)
6. 長尾, 平池, 岡本, 藤野, 坂野, 道坂, 春山, メディア応用システム, 映像情報メディア学会誌第57巻, 8号, pp.1136-1139, (2004)
7. 小松 尚久, 内田 薫, 坂野 鋭, 和田 誓一, 池野 修一, バイオメトリクスの精度評価, 計測と制御 Vol.43, No.7, (2004)
8. 坂野 鋭, 多重バイオメトリクス, 計測と制御 Vol.43, No.7, (2004)
9. 坂野 鋭, 生体認証技術の最近の動向, 日本法科学技術学会誌, Vol.12, No.1, pp.1-12, Feb, (2007)

## [招待講演，セミナー講師]

1. 坂野 鋭, バイオメトリック個人認証技術の動向と課題, 電子情報通信学会技術報告, PRMU-99-29, (1999.6)

2. 木村義政, 内田薫, 小松尚久, 坂野 鋭, 瀬戸洋一, 中野裕二, パネル討論  
バイオメトリクス認証はインターネット社会の役に立つか?, 電子情報通信  
学会技術報告,PRMU-2000-75, (2000.9)
3. 坂野 鋭, モバイル環境でのセキュリティ技術, 画像電子学会秋季 세미나,  
(2000)
4. 坂野 鋭, バイオメトリック個人認証の新展開, 映像情報メディア学会研究  
報告,ITE-VIS (2002.3.1),
5. 坂野 鋭, 多重バイオメトリクスの研究動向, 電子情報通信学会総合大会発  
表論文集,PD-12-2,(2002.3.28)
6. 坂野 鋭, 多重生体認証技術の研究動向, 日本能率協会セキュリティ技術セ  
ミナ,(2002.11.28)
7. 坂野 鋭, バイオメトリック個人認証技術の世界, 名古屋市産業振興公社福  
祉工学研究会,(2003,3,10)
8. 坂野 鋭, バイオメトリック個人認証技術の最新動向-標準化, 実証実験-, 映  
像情報メディア学会メディア工学シンポジウム,(2003,3,12)
9. 坂野 鋭, 広がるバイオメトリック個人認証技術, ビジネス  
ショー,(2003,5,20)
10. 坂野 鋭, バイオメトリック個人認証技術の標準化 - 世界の動向 -, 情報  
処理学会コンピュータービジョンとイメージメディア研究会, (2003,7,5)
11. 坂野 鋭, バイオメトリック個人認証技術の最新動向, 中国技術振興セン  
ター,(2003,9)
12. 坂野 鋭, バイオメトリック個人認証技術の新展開, 画像センシングシンポ  
ジウム,I-1,(2004,6)
13. 坂野 鋭, SC37/WG5(性能試験)の最新動向, バイオメトリックセキュリ  
ティコンソーシアム 2004 年度第 2 回講演会 ,(2004,7.20)
14. 坂野 鋭, 顔画像認識技術の標準化 -何を標準化するのか -, 電子情報通信  
学会 コピキタネットワーク社会のためのバイオメトリクスセキュリティ  
研究会,(2004,9)
15. 坂野 鋭, バイオメトリック個人認証技術の動向, 日本鑑識科学技術学会特  
別講演, (2004.11.11)
16. 坂野 鋭, バイオメトリック個人認証技術の動向, 月刊カードウェブ主催  
バイオメトリクスセミナー,(2004,12)
17. 坂野 鋭, バイオメトリック個人認証技術の展開, 画像センシングシンポジ  
ウム 2005,(2005,6)
18. 道坂修,坂野 鋭, バイオメトリクス認証による国境管理システムの現状,  
画像センシングシンポジウム 2006,A-2, (2006,6)

[国際会議、査読なし]

1. Hitoshi Sakano, Japanese Standardization of evaluation method of facial authentication system, In Proc. 1st . Asian Biometric Workshop. (2003)
2. Satoshi Haruyama, Yasuaki Nikura, Osamu Dosaka and Hitoshi Sakano, Multimodal Biometrics using Kernel Fisher Discriminant Analysis, In Proc. Biometrics Authentication Symposium in Kyoto, pp. 145-147, 2006.1.16, Kyoto University

[研究会, シンポジウム, 査読なし]

1. 村瀬洋 (NTT), A: 増田 健 (電子技術総合研究所), 谷口倫一郎 (九州大学), B: 井宮淳 (千葉大), 岡田稔 (名古屋大学), C: 今井正和 (奈良先端科学技術大学院大学), 中島昇 (NEC), 黄瀬浩一 (大阪府立大学), D: 斎藤英雄 (慶応義塾大学), 坂野 鋭 (NTT データ通信), 13th ICPR (第 13 回パターン認識国際会議) 報告, 電子情報通信学会技術報告, PRMU96-111, (1996)
2. 坂野 鋭, 武川直樹, AIC による部分空間次元数の決定法, 電子情報通信学会技術報告, PRMU97-173, (1997-11)
3. 中山直樹, 春山智, 坂野 鋭, 摂動法による顔画像の認識, 電子情報通信学会技術報告, PRMU-98-90, (1998)
4. 春山智, 坂野 鋭, 武川直樹, 陰影・幾何拘束を用いた光源・姿勢変動にロバストな顔認識アルゴリズム, 電子情報通信学会技術報告, PRMU-98-, (1998)
5. 劉偉傑, 坂野 鋭, 結合規則と閾値の最適化に基づく多重バイオメトリクス, 電子情報通信学会技術報告, PRMU-99-23, (1999)
6. 新村貴彦, 出本浩, 坂野 鋭, 木田博巳, LVQ における誤認識パターンを重視した学習パターン選択法, 電子情報通信学会技術報告, PRMU-99-146, (1999)
7. 春山智, 坂野 鋭, 武川直樹, 顔認識システムにおける照明変動の影響とその改善手法, 情報処理学会研究報告, CSEC 7-5, (1999)
8. 飯野徹, 岩瀬史幸, 坂野 鋭, 中嶋秀樹, 指紋機能搭載型 IC カードによる本人認証方式, 電子情報通信学会技術報告, ISEC-2000-44, (2000)
9. 坂野 鋭, 有田大作, 大町真一郎, 岡田稔, 栗田多喜夫, 斎藤英雄, 島田伸敬, 藤木淳, 森稔, 国際会議報告: ICPR2000 とその周辺, 電子情報通信学会技術報告, PRMU-2000-146, (2000)
10. 末永高志, 佐藤新, 坂野 鋭, クラスタ構造に着目した特徴空間の可視化 - クラスタ判別法 -, 電子情報通信学会技術報告, PRMU-2001-44, (2001)
11. 鈴木賢一郎, 中嶋正臣, 坂野 鋭, 識別的な手法に基づく映像データからのカット検出法, 電子情報通信学会技術報告, PRMU-2001-47, (2001)
12. 佐藤新, 末永高志, 坂野 鋭, クラスタ判別法の医療データ解析への応用, 人工知能学会研究報告, (2001,11,14)
13. 坂野 鋭, 赤松茂, 鷲見和彦, 坂本静男, 長尾健司, 佐藤勝彦, 矢野浩, 磯部俊弘, 顔画像認証アルゴリズムの精度評価方法 - 標準化に向けて -, 信学技報 PRMU2001 - 164, (2001 . 12)

14. 東陽子, 佐藤新, 樋口裕高, 坂野 鋭, 森本俊彦, 松永務, 石井敬介, 村松正明, クラスタリングによる遺伝子解析 - 遺伝子型の自動決定 -, 信学技報 PRMU2002 - 71, (2002.9)
15. 鈴木賢一郎, 中嶋正臣, 坂野 鋭, 局所動き方向ヒストグラム特徴による映像データからの動物体の検出, 信学技報 PRMU2002 - 110, (2002.10)
16. 末永高志, 坂野 鋭, 松永務, 顔認識系におけるデータ採取地の影響について, 信学技報 PRMU2002 - 147, (2002.12)
17. 春山智, 道坂修, 坂野 鋭, 中村太一, 核非線形判別分析による多重バイオメトリクス, 信学技報 PRMU2002 - 204, (2003.2)
18. 吉澤大樹, 坂野鋭, 橋本周司, 最適化のための粗視化ニュートン法, In Proc. IBIS2002, pp. 184-189, 2002
19. 吉澤大樹, 坂野 鋭, 橋本周司, 最適化のための粗視化ニュートン法, 情処研報 MPS43-6, (2003)
20. 佐藤新, 坂野 鋭, 松永務, 非線形構造に着目した識別ルール抽出法, 信学技報 PRMU2003-34, (2003.6.20)
21. 鈴木賢一郎, 坂野 鋭, 大塚作一, エッジ点の差分特徴を用いた判別分析による映像データからの漸次カット点検出法, 信学技報 PRMU2003-80, (2003.9.8)
22. 田中宏征, 坂野 鋭, 大塚作一, 多重カテゴリに基づく画像検索法の提案, 信学技報 PRMU2003-96, (2003.9.8)
23. 坂野 鋭, データマイニングの鉱脈を探る -パターン認識の視点から-, 信学技報 PRMU2003-93, (2003.9.8)
24. 田中宏征, 坂野 鋭, 大塚作一, パラメトリック混合モデルによる画像検索, 画像の認識と理解シンポジウム MIRU2004, (2004)
25. 三枝亮, 坂野 鋭, 橋本周司, 非線形主成分分析による画像解析, 画像の認識と理解シンポジウム MIRU2004, (2004)
26. 市野正嗣, 坂野 鋭, 小松尚久, 核非線形相互部分空間法による話者認識, 画像の認識と理解シンポジウム MIRU2004, (2004)
27. 近藤功, 坂野 鋭, Steerable Filter によるテクスチャ記述, 信学技報 PRMU-2004-165, (2005)
28. 鈴木賢一郎, 坂野 鋭, 大塚作一, 縁取りに着目したテロップ認識, 信学技報 PRMU-2004-232, (2005)
29. 市野将嗣, 坂野 鋭, 小松尚久, 核非線形相互部分空間法による話者認識, , 2005年 暗号と情報セキュリティシンポジウム予稿集, Vol.3, pp.973-978, Jan.2005
30. 市野将嗣, 坂野 鋭, 小松尚久, 核非線形相互部分空間法の処理量削減に関する検討, 画像の認識と理解シンポジウム MIRU2005, (2005)

31. 市野将嗣, 坂野 鋭, 小松尚久, 唇動作の非線形性に着目した個人認証アルゴリズムの高速化に関する一考察, ユビキタスネットワーク社会におけるバイオメトリクスセキュリティ研究発表会第5回研究会予稿集, pp.1-8, (2005)
32. 市野将嗣, 坂野 鋭, 小松尚久, テキスト提示型話者認識のための音声個人認証方式, 暗号と情報セキュリティシンポジウム (SCIS2006), 2D1-3, 2006.1.18
33. 市野将嗣, 坂野 鋭, 小松尚久, 話者認識における核非線形相互部分空間法の適用に関する検討, MIRU2006 サテライトワークショップ部分空間法研究会 2006, pp.1-6, (2006.7)
34. 市野将嗣, 坂野 鋭, 小松尚久, 唇動作と音声の非線形性を考慮した多重バイオメトリクス認証方式に関する検討, 暗号と情報セキュリティシンポジウム (SCIS2007), 2007年1月
35. 市野将嗣, 坂野 鋭, 小松尚久, 正準角の統合による相互部分空間法の改良に関する一検討, MIRU2007, IS-5-01, (2007.8)
36. Hitoshi Sakano, Linear and Nonlinear Hybrid Kernel Mutual Subspace Method, 信学技報 PRMU2007-256, (2007.12)

## [全国大会]

1. 坂本光徳, 金沢辰樹, 鎌田啓一, 坂野 鋭, 川崎温, 安藤利得, 増崎, 克 5p-K-12 矩形断面導波管を用いた相対論的電子サイクロトロン・メーザーの実験 I 秋の分科会講演予稿集 1989(4), 243, 19890912 (社団法人日本物理学会)
2. 金澤辰樹, 坂本光徳, 鎌田啓一, 坂野 鋭, 川崎温, 安藤利得, 増崎, 克 5p-K-13 矩形断面導波管を用いた相対論的電子サイクロトロン・メーザーの実験 II 秋の分科会講演予稿集 1989(4), 243, 19890912 (社団法人日本物理学会)
3. 坂野 鋭, 渡辺寛, 2次元ペンローズ格子上的DLA, 日本物理学会講演概要集 第3分冊, p.483, (1990)
4. 小田久子, 岸本泰明, 坂本慶司, 前原直, 志甫諒, 坂野 鋭, 川崎温 27p-D-6 CARM (Cyclotron Auto Resonance Maser) による電磁波増幅の数値解析 春の分科会講演予稿集 1991(4), 189, 19910311 (社団法人日本物理学会)
5. 坂野 鋭, 2次元ペンローズ格子上的DLA II, 日本物理学会講演概要集 第3分冊, p.440, (1993)
6. 坂野 鋭, 阿部 一朗, 分布の非対称性を考慮した手書き文字認識, 電子情報通信学会全国大会発表論文集, (1991.3)
7. 磯部俊洋, 坂野 鋭, 劉偉傑, 顔画像と音声を用いた個人認証方式について, 音響学会春季講演論文集, 2-6-9, pp.71-72, (1998.3)
8. 坂野 鋭, 磯部俊洋, 劉偉傑, 春山智, 武川直樹, 顔と音声のスコア統合による個人認証, 電子情報通信学会総合大会発表論文集, (1998.3)
9. 春山智, 坂野 鋭, 武川直樹, 陰影・幾何学的拘束を用いた光源・姿勢にロバストな顔認識アルゴリズム, 電子情報通信学会総合大会発表論文集, D-12-104, (1999.3)

10. 佐藤新, 末永高志, 坂野 鋭, クラスタ判別法の医療データへの応用, 情報処理学会第 63 回全国大会講演論文集 第 3 分冊 p.193,(2001.9)
11. 末永高志, 佐藤新, 坂野 鋭, クラスタ判別法による顔画像データ解析, 電子情報通信学会ソサイエティ大会発表論文集,D-12-2,(2001.9)
12. 末永高志, 佐藤新, 坂野 鋭, クラスタ判別法による文字データ解析, 電子情報通信学会総合大会発表論文集,D-12-2,(2002.3)
13. 坂野 鋭, 末永高志, 佐藤新, 新しい高次元データ可視化法 -クラスタ判別法-, システム制御情報学会,PD-12-2,(2002.12)
14. 吉澤大樹, 坂野 鋭, 橋本周司, ”粗視化ニュートン法における探索点の統計的動作”, 情報処理学会第 65 回全国大会講演論文集 (分冊 1) , pp.169-170 (2003).
15. 鈴木賢一郎・中嶋正臣・坂野 鋭, 動き方向ヒストグラム特徴を用いた瞬時カット点検出法 -映像処理評価用映像データベースによる評価-, 電子情報通信学会総合大会発表論文集,D-12-65,(2003.3.21)
16. 末永高志・坂野 鋭・松永 務, 顔画像検出におけるデータ採取地の影響について, 電子情報通信学会総合大会発表論文集,D-12-55,(2003.3.21)
17. 佐藤 新・部 亜紀子・坂野 鋭・松永 務, 非線形構造に着目した識別ルール抽出法, 電子情報通信学会総合大会発表論文集,D-12-45,(2003.3.20)
18. 部 亜紀子・佐藤 新・坂野 鋭・松永 務, 区分線型構造に着目した識別ルール抽出法, 電子情報通信学会総合大会発表論文集,D-12-44,(2003.3.20)
19. 市野将嗣, 高倉大樹, 坂野 鋭, 小松尚久, 母音音素分布の非線形性について, 電子情報通信学会総合大会,D-14-14, 2004 年 3 月