不確かさの自己学習と
その画像処理回路網への応用

Self-Learning of Uncertainty and its Applications to Neural Networks for Image Processing.

萩原 義裕
Yoshihiro Hagihara

指導教官　小畑 秀文 教授

1996年3月

①

不確かさの自己学習と
その画像処理回路網への応用

Self–Learning of Uncertainty and its Applications to Neural Networks for Image Processing.

萩原 義裕
Yoshihiro Hagihara

指導教官　小畑 秀文 教授

1996 年 3 月

論文概要

神経回路網を用いて複雑な情報処理を行うためには、多数の細胞のパラメータやコネクションを目的に応じて設定する必要がある。本論文の目的は、多数の細胞のパラメータ等を簡単に設定できるような神経回路モデルを提案し、工学的に実現することにある。

まず、それほど複雑でない情報処理を行う単純な神経回路モデルの存在に着目し、単純なモデルを組み合わせてやや複雑な情報処理を行うことを検討する。それを同じ方法でさらに組み合わせることによって、より複雑な情報処理を行うことができるなら、多数の細胞を利用した神経回路網のパラメータ等を容易に設定することが可能になる。そのためには、組み合わせるのが容易な基本回路網を実現することが重要となる。

本論文では神経回路網を構成するために必要な事柄を分析し、各章に分けて検証する。本論文は6章で構成される。第1章は本研究の位置づけ、第2章から第4章は神経回路網モデルの各構成要素に関する提案と検証、第5章はこれらの要素の統合及びシステムの検証、そして第6章は本研究の成果のまとめである。

本論文で提案する大規模神経回路網は、複数のモジュールで構成されている。入力信号は複数に分割され、各モジュールに与えられる。各モジュールの出力は統合されて回路網の出力結果となる。各モジュールの内部もさらに小さなモジュールで構成されている。モジュールの最小単位は小数の細胞から成り立つ基本的な素子である。そのため、回路網を構成するために必要な機能を第2章から第5章で個別に検討した。第2章は各モジュールの出力の統合、第3章は基本素子モデル、第4章は各モジュールの入出力の実現方法、第5章はシステム全体に関する研究となっている。

なお、各章の概要は以下の通りである。

第1章「序論」

神経回路網に関する過去の研究と現状を調査、検討した上、大規模神経回路網モデルの構成を示し、問題点を分析した。

第 2 章 「動的 n–k 最近接近傍法による画像のスムージング」

複数の神経回路網から出力される多次元情報をひとつの情報として統合する際、最も有効と思われる k 最近接近傍フィルタを取り上げ、神経回路網のシミュレーテッドアニーリング的な手法を利用してフィルタを高性能化する手法について述べた。画像ノイズの除去にこの手法を応用したとき、従来のフィルタより雑音除去効果が大きいことを確認した。

第 3 章 「不確定性を考慮した神経回路モデル」

神経回路網の中で基本素子となる単純なモデルを提案した。多数の基本素子を用いて回路網を構成する場合、すべての基本素子を正常に動作させることは困難である。この場合、いくつかの基本素子の出力は有効な情報を持たないことが予想される。従って、基本素子には、出力結果の有効な情報であることを確認できる機能が必要である。本章では、情報の内容と共にその情報の有効性を出力する基本素子モデルを提案した。これにより、神経回路網は、有効な情報のみ取り出すことが可能になる。

第 4 章 「多層パーセプトロンによる確率密度関数の自己学習」

単純な神経回路網をモジュール化する際、各モジュールの出力は簡単に統合化できなければならない。各モジュールの出力情報が有効であることを確認するのは、統合化する際に最も重要と考えられる。ここで、入力信号の確率密度が出力結果の有効性を直接反映していることに着目する。情報の内容とともに確率密度を出力するような学習を行うことにより、統合は容易になる. このような学習を行う場合、従来のモデルでは確率密度をあらかじめ算出し、学習させる必要がある。提案手法は、入力に乱数を与えることによって、事前に算出せずに確率密度を学習するものである。多層パーセプトロンに排他的論理和を学習させることにより、本手法が有効であることを確認した。また、提案手法は、確率密度をあらかじめ算出する方法と比べ、記憶に必要な細胞数が少ないため、メモリ効率が良いことを示した。

第 5 章 「確率密度を自己学習する MLP のアプリケーション」

確率密度関数を自己学習する MLP など、各要素技術を組み合わせた神経回路網の詳細な構成を述べ、画像の雑音除去及び手書き文字認識に適用した計算機シミュレーションについて述べた。複数の物体を撮影した画像や、雑音源の周囲で録音した音声は、複数の情報が重畳した信号である。また、手書き漢字は、分類対象となるパターン数が多いため、学習を失敗する回路網が多いことが予測される。この信号から情報を取り出すためには、信号を分割して有効な要素だけをとりだした後に再び統合化する必要がある。従来手法は、細分化された情報同士に確率的な関係があることを利用している。そのため、雑音が大きい場合など、確率的な関係が切れる場合は著しく認識率が低下するという欠点がある。本章は、確率的な

関係によらない神経回路網を構成することにより、認識率の低下を防ぐことを目的とする。まず、確率密度を出力する神経回路を複数用いて、分割した信号から情報を取り出すとともに、取り出された情報の有効性を判定する。次に、有効な情報のみ取り出して、重ね合わせることにより出力を統合化する。雑音が重畳した印刷パターンの雑音除去と手書き文字認識に本手法を適用し、この手法が有効であることを確認した。

第 6 章 「結論」

本論文全体の考察及び総括をおこない、本研究で提案する神経回路網モデルの有効性を示した。

目次

第 1 章　序論 **1**

1.1　神経回路モデルに関する研究史 1

1.2　近年の研究動向と大規模神経回路網の問題点 2

1.3　局所的最小解による問題 3

1.3.1　大域的最小解が複数存在する問題 5

1.3.2　類似パターンの分離に関する問題 7

1.3.3　情報の複雑化への対応 7

1.4　大規模神経回路網システム 12

1.5　本研究の目的と方針 15

第 2 章　動的 n–k 最近接近傍法による画像のスムージング **20**

2.1　はじめに 21

2.2　ヒストグラムからの検討 21

2.3　動的手法の導入 27

2.3.1　動的 k 最近接近傍フィルタ 27

2.3.2　動的 n–k 最近接近傍フィルタ 27

2.3.3　具体的な処理手順 27

2.4　実験と考察 30

2.4.1　k 最近接近傍フィルタの特性比較 30

2.4.2　1 次元信号での比較 30

2.4.3　2 次元画像における本手法の最適パラメータの測定と性能比較—1 枚の画像での実験 33

2.4.4　複数の画像による性能比較 38

2.5　本章のまとめ 42

第 3 章　不確定性を考慮した神経回路モデル **45**

3.1　はじめに 46

3.2　従来の神経回路モデルの問題点 46

3.3　情報の内容と不確定性 48

3.4 不確定性を考慮した神経回路モデル 48
3.5 連想記憶モデルの学習則 50
3.6 計算機シミュレーション 51
3.7 本章のまとめ 51

第 4 章 多層パーセプトロンによる確率密度関数の自己学習 55
4.1 はじめに 56
4.2 学習サンプルの確率密度関数を推定できる MLP 57
4.2.1 学習サンプルの確率密度関数 57
4.2.2 関数の学習法 58
4.2.3 出力と学習サンプルの確率密度関数との関係に対する検討 59
4.3 コンピュータシミュレーション 62
4.3.1 1 次元信号のカテゴリ分類による性能確認 62
4.3.2 S/N 比による出力変化 67
4.3.3 排他的論理和型のカテゴリ分類実験 67
4.4 本章のまとめ 70

第 5 章 確率密度を自己学習する MLP のアプリケーション 75
5.1 はじめに 76
5.2 画像や音声における諸問題 77
5.3 複数パターンからの信号抽出 77
5.4 画像処理への応用 81
5.4.1 画像修復のためのネットワーク 81
5.4.2 ネットワークの詳細 84
5.4.3 最適パラメータの測定 84
5.4.4 従来の手法との比較 85
5.5 手書き文字認識への応用 94
5.5.1 手書き文字認識システム 94
5.5.2 細線化および特徴ベクトルの抽出 94
5.5.3 パターン分類 97
5.5.4 出力の統合 97
5.6 ETL9B を用いた認識実験 100
5.6.1 認識実験の概要 100
5.6.2 各ネットワークのパターン分類実験 100
5.6.3 出力の統合実験 100
5.7 本章のまとめ 108

第 6 章 結論 111

図目次

1.1 局所的最小解から脱出できないケース 4
1.2 解が複数あるケース 6
1.3 本質的に分離不可能なケース 8
1.4 多層パーセプトロン 9
1.5 統合型ニューラルネットワーク 10
1.6 選択型ニューラルネットワーク (Comb NET) 11
1.7 Nestor Learning System(NLS) 13
1.8 大規模神経回路網のブロック図 14
1.9 各技術の位置づけ 16

2.1 k 最近接近傍フィルタの原理 22
2.2 ガウス雑音が付加された画像のヒストグラム 24
2.3 局所的ピークのある画像のヒストグラム 25
2.4 局所的ピークを乗り越えるパラメータ設定 26
2.5 動的 k 最近接近傍フィルタ 28
2.6 動的 n–k 最近接近傍フィルタ 29
2.7 k 最近接近傍フィルタの k と二乗平均誤差の関係 31
2.8 1 次元信号の雑音除去実験 32
2.9 雑音除去実験に用いた画像 35
2.10 n と k の比と二乗平均誤差の関係 36
2.11 様々な画像パタン 37
2.12 平滑化処理の比較 39

3.1 連想記憶の問題点 47
3.2 連想記憶モデル 49
3.3 学習の過程 52
3.4 連想記憶のシミュレーション 53

4.1 出力と確率密度の関係 61
4.2 1 次元信号の分割実験 63

4.3 確率密度関数の比較 64
4.4 中間細胞数による出力の変化 66
4.5 学習サンプルとランダムデータの比と出力の関係 68
4.6 2次元入力の分割実験 69
4.7 中間細胞数による出力の変化 71

5.1 複数パターンによって構成される画像モデル 78
5.2 1つのパターンと雑音からなる画像 80
5.3 神経回路網システム 82
5.4 神経回路網システムの構成 83
5.5 n に対する k の割合と二乗誤差の関係 86
5.6 シミュレーションに用いた文字パターン 87
5.7 雑音を付加したパターン 88
5.8 確率密度を自己学習する MLP と k 近傍フィルタによる処理結果 89
5.9 確率密度を自己学習する MLP と全平均フィルタによる処理結果 90
5.10 従来型型 MLP と k 近傍フィルタによる処理結果 91
5.11 従来型型 MLP と全平均フィルタによる処理結果 92
5.12 文字認識システムのブロックダイアグラム 95
5.13 細線化および特徴抽出 96
5.14 パターン分類を行うネットワーク 98
5.15 出力の統合 99
5.16 実験に用いた手書き文字パターンの一部 101
5.17 細線化された文字パターン 102
5.18 ネットワークの出力の一例（従来型 MLP） 104
5.19 ネットワークの出力の一例（確率密度自己学習型 MLP） 105
5.20 従来型 MLP による認識結果 106
5.21 確率密度自己学習型 MLP による認識結果 107

表目次

2.1　1次元信号の雑音除去結果 34
2.2　2次元信号の雑音除去結果（その1） 40
2.3　2次元信号の雑音除去結果（その2） 41

5.1　雑音除去結果の比較 93
5.2　ネットワークのパターン分類失敗率 103

第 1 章

序論

1.1 神経回路モデルに関する研究史

近年、神経回路網の研究は急激に発展し、その成果は様々な形で社会に還元されている。生物の神経回路の働きは、その思考や行動の基準となるものであり、様々な視点から見て興味深い研究対象である。脳細胞の化学的な働きを解明し、障害の治療に役立てることを目的とするもの、電気計測によって脳の働きを解明するもの、生物の行動から脳の構造の仮説を立てるもの、神経回路の働きを情報幾何学の立場から検討するもの、工学的なモデルを用いて仮説を証明するもの、神経回路モデルを基本にコンピュータによる高度情報処理を目指すものなど、無数のアプローチがなされている。

工学的な視点に限定すると、神経回路網の研究は 3 つの流れに大別される [1]。階層的ネットワーク、相互結合型ネットワーク、そして Boltzmann Machine である。なかでも階層的ネットワークの歴史は古く、1940 年代まで遡ることができる。McCulloch と Pittz は 1943 年ニューロンモデルを提案し、神経回路網を用いて論理関数の計算が可能であることを示した [2]。Hebb は心理学の立場からシナプス結合の変化規則を提案し、学習の仮説を立てた [3]。そして、Rosenblat が提案した Perceptron によって、神経回路網の工学的研究は第 1 の隆盛期を迎える [4, 5]。Perceptron と名付けられたパターン識別装置の基本原理は、3 層の神経回路網であり、現在実用化されている多くのニューロコンピュータにも利用されている。

しかし、Perceptron を直接応用する研究は 1969 年を境に急激に衰えた。1969 年に Minsky と Papert が Perceptron の学習限界を指摘したためである [6]。これにより Perceptron に対する過大な期待はなくなり、神経回路網の研究は衰退期を迎える。

歴史的には衰退期と呼ばれるものの、地道な研究は着実に行われ、1980 年代の第 2 隆盛期の基礎となる理論が生まれる。時代は遡るが 1967 年、甘利は、より高性能な学習理論を提案している [7, 8]。さらに、甘利は 1972 年、自己組織化ネットワーク、1974 年、神経統計力学など後年の研究に大きな示唆を与える提案を行った。

この時代の研究は多様な方向性を持つ。1969 年、Marr と Albus は小脳のモデルを提案した。Kohonen らは 1972 年から 1974 年にかけて 1 層の連想記憶モデルを、1976 年、相互結

合型の連想記憶モデルを提案した [9, 10, 11, 12, 13]。また、中野も 1972 年、相互結合型の回路網による連想記憶モデルを提案した [14]。一方、福島らは、多層ネットワークで自己組織化を行う Cognitron、Neocognitron を提案した [15, 16, 17]。また、平井は 1 層のネットワークと相互抑制的なネットワークを組み合わせた連想記憶モデルを提案した [18, 19]。

これらの研究は 2 つの流れを形作った。Cognitron に代表される協調系と競合系の結合による神経回路モデルと、相互結合型の回路網による連想記憶モデルである。協調系と競合系の結合による神経回路モデルは、甘利らのカエルの虫検出回路網モデルや両眼立体視モデルなどへと発展する [20]。これらは生理学的知見と矛盾しないモデルを考え、シミュレーションや数学的解析によってモデルを検証するという研究手法を、より明確に打ち出している。一方、1982 年、相互結合型連想記憶モデルとして提案された Hopfield モデルは、NP 完全問題などに適用可能なため注目され、ニューラルコンピュータの第 2 次隆盛期を興した [21, 22]。その後、Touloue et al.、Peretto、Kinzel、Amit et al.、Shinomoto、Meir and Domany など、多くの研究者が物理学的視点からネットワークの研究を行っている。また、Farlman et al. は、確率的に動作する Boltzman Machine を提案した [23]。ここで提案された焼き鈍し法は局所的最小値を回避する手段として有効であることが知られている。ほぼ同時期、Geman らは、マルコフ場で焼き鈍し法を用いる統計的緩和法によって画像フィルタの統一的理論を確立した [24]。統計的緩和法は画像の雑音除去などに利用できる。

ニューラルコンピュータの第 2 次隆盛期を興したもう 1 つの主役は、1986 年に Rumelhart らが提案した誤差逆伝搬法である [25]。これは Perceptron を多層化させた多層 Perceptron(MLP) の学習則であり、実用性が高いため産業界にも大きな影響を与えた。また、1986 年、Kohonen は学習ベクトル量子化 (LVQ) を提案した [26]。これは、パターン空間を有限個に分割し、各パターン空間を参照ベクトルで量子化するベクトル量子化法を学習によって実現するものである。学習ベクトル量子化法は、音声認識や画像認識に応用すると好成績を得ることができる。特に、手書き文字認識など、認識対象となるパターンが多い場合、多層 Perceptron では困難な大分類に有効性があることが確認されている。学習ベクトル量子化法を用いて大分類を行い、その結果を基に小分類を行う Comb-Net は、手書き文字認識などに有効であることが確認されている [27, 28]。

1.2 近年の研究動向と大規模神経回路網の問題点

生物の神経回路は極めて複雑であり、働きを解明することも工学的に利用することも限られた範囲でしか実現していない。工学モデル、生物学モデルの共通の欠点は、生物がいとも簡単に行う複雑な思考を再現できない点にある。従来のモデルを用いて複雑な思考がどのように行われているかを説明することはできない。同様に、複雑な思考を要するシステムを実現することも困難である。

単純な働きをする素子を用いて情報処理を行うためには、処理に応じた数の素子を利用しなければならない。しかし、一般的なモデルでは素子数が多いときには正常に働かないこと

が知られている。素子数が少なければ複雑な情報処理ができないし、素子数を多くすれば制御できないという矛盾が存在するのである。

特に、類似パターンが増加したり、パラメータが不適切だと識別能力は低下する。識別の応力の低下が起きるのは次のような理由による。

- 局所的最小解に陥りやすくなる。
- 大域的最小解が複数存在する確率が増える。
- 類似パターンが増加する。

これらは情報量の多寡に関わらず一般的な神経回路モデルに内在する問題である。情報量の増加に伴って局所的最小解、大域的最小解の自由度、類似パターンが増加するため回路網の識別能力が低下する。このような問題が生じる可能性は与えられる情報の性質によって決定される。

1.3 局所的最小解による問題

局所的最小解から抜け出せなくなるのは、情報量の増加に伴って局所的最小解も増えるためである。神経回路網は誤差が減少するように学習を行う。図 1.1 は、学習によって誤差が減少する過程を模式的に示したものである。図 1.1 (a) のように自由度が小さい場合には局所的最小解から抜け出すことができるが、自由度が大きい図 1.1 (b) はパターンが複雑で超空間上の距離が大きいため抜け出すことができない。情報が複雑になるに従って解の自由度が大きくなり、局所的最小解に陥る可能性も増加する。

局所的最小解から脱出する方法には、次のようなものがある。

- シミュレーテッドアニーリングなどの手法を用いて入出力の応答や学習係数を変化させる。
- 学習関数にモーメント項を組み込む。
- 学習を行った後、ネットワークの重み係数の一部に外乱などを与える。

シミュレーテッドアニーリングは、入出力の応答を変化させる方法である。また、同様な効果を得る手法として学習係数（重み係数の変化量）を変化させる方法も考えられる。学習の初期では大きな値の学習係数を用い、学習が進むに従って次第にその値を小さくする。学習の初期では局所的最小解に陥らないようにするため、重み係数の変化量は大きい方がよい。しかし、重み係数の変化量が大きいと振動によって大域的最小解に収束しないため、学習が進むに従って変化量を小さくする必要がある。この方法は、学習の初期において重み係数の変化量を大きくするため、局所的最小解に陥るのを防ぐと共に学習速度向上にも役立つ。

一方、学習関数にモーメント項を組み込む方法は、前回の重み係数の変化を今回の重み係数に加味するものである。神経回路網の学習では、誤差を減少させる方向に重み係数を変化

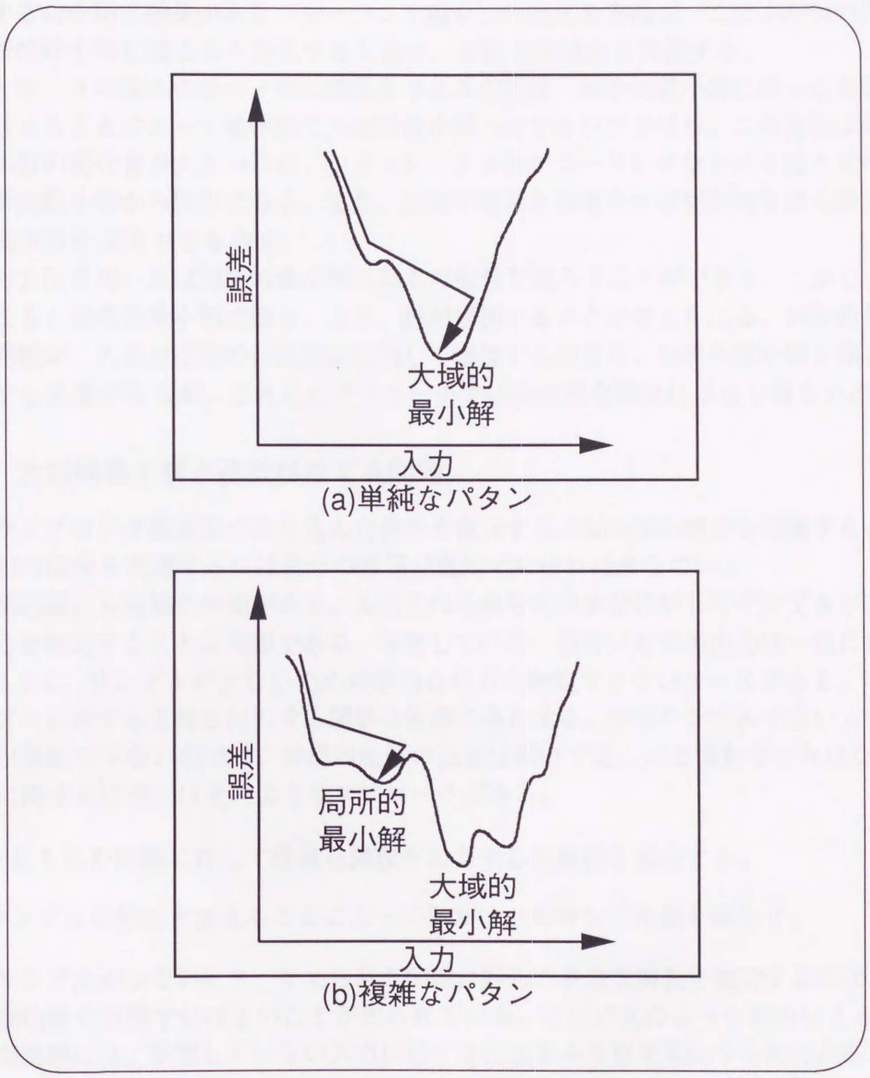


図 1.1: 局所的最小解から脱出できないケース

させる。しかし、真の大域的最小解に収束させるためには、一次的に誤差が増加する方向に重み係数を変化させなければならないことがある。両者のベクトルが同じ方向の時、前回の重み係数を今回の重み係数に加えれば、重み係数の変化量は大きくなり、局所的最小解から脱出することができる。また、両者のベクトルが正反対の時には、重み係数が振動して学習が停止するのを防ぐ効果がある。モーメント項をつけ加える方法は、このふたつの作用により、局所的最小解に陥るのを防止すると共に、学習の高速化を実現する。

ネットワークの重み係数の一部に外乱を与える方法は、局所的最小解に陥った回路網にゆらぎを与えることによって確率的に大域的最小解へとジャンプさせる。この方法は外乱による重み係数の変化量が大きいため、シミュレーテッドアニーリングなどの方法と比べ、より広い局所的最小解から脱出できる。また、忘却の効果を利用すれば局所的な最小解に陥らせずに構造学習を実現できる [29]。

以上の方法を用いれば局所的最小解に陥る可能性を減らすことができる。しかし、情報が複雑になると局所的最小解の深さ、広さ、数が増加することが考えられる。局所的最小解に陥る可能性が、入出力空間の自由度に比例して増加するのなら、局所的最小解に陥る可能性を 0 にする必要があるが、これらのアプローチでは決定的な解決にはなり得ないのである。

1.3.1 大域的最小解が複数存在する問題

粗なサンプリング周波数で取り込んだ信号を復元するには未知の部分を補間する必要がある。未知の部分を補間するには信号の性質が既知でなければならない。

神経回路網にも同様な問題がある。入力される信号の自由度に対してサンプルが少ないとき、出力を特定することは困難である。学習していない信号に対する出力は一意に定まらない。図 1.2は、サンプルが少ないため理想的な出力が特定できないケースである。与えられたサンプルに対する正解を出力する関数は無数に存在する。学習サンプルでない入力に対して出力が保証できないのでは、神経回路網の価値は極めて乏しいと言わなければならない。

これに関する研究には次のようなアプローチがある。

- 与えられた問題に対して最適な関数を出力する回路網を推定する。
- サンプルに外乱を加えることによって見かけ上のサンプル数を増やす。

学習サンプルが少ないとき、与えられた問題に対して最適な関数を決定するには、回路網の中間細胞数を調節すればよいことが知られている。これは次のような理由による。

神経回路網には、学習していない入力に対する出力をある程度推定する能力がある。学習サンプルの近傍の未知入力が与えられたとき、神経回路網は学習サンプルが与えられたときと同様に振舞う。一般的な学習では、この能力を利用して未知入力に対する出力を保証する。ここで、「近傍」の範囲は回路網の記憶容量、すなわち中間細胞数によって決定する。細胞数が不必要に多い場合関数の自由度が大きくなりすぎるため不適切な関数（無意味に複雑すぎる関数）を出力する。十分な学習サンプルが与えられなくても、中間細胞数を適切な数に設定できるとき、神経回路網は理想的な関数を出力する。

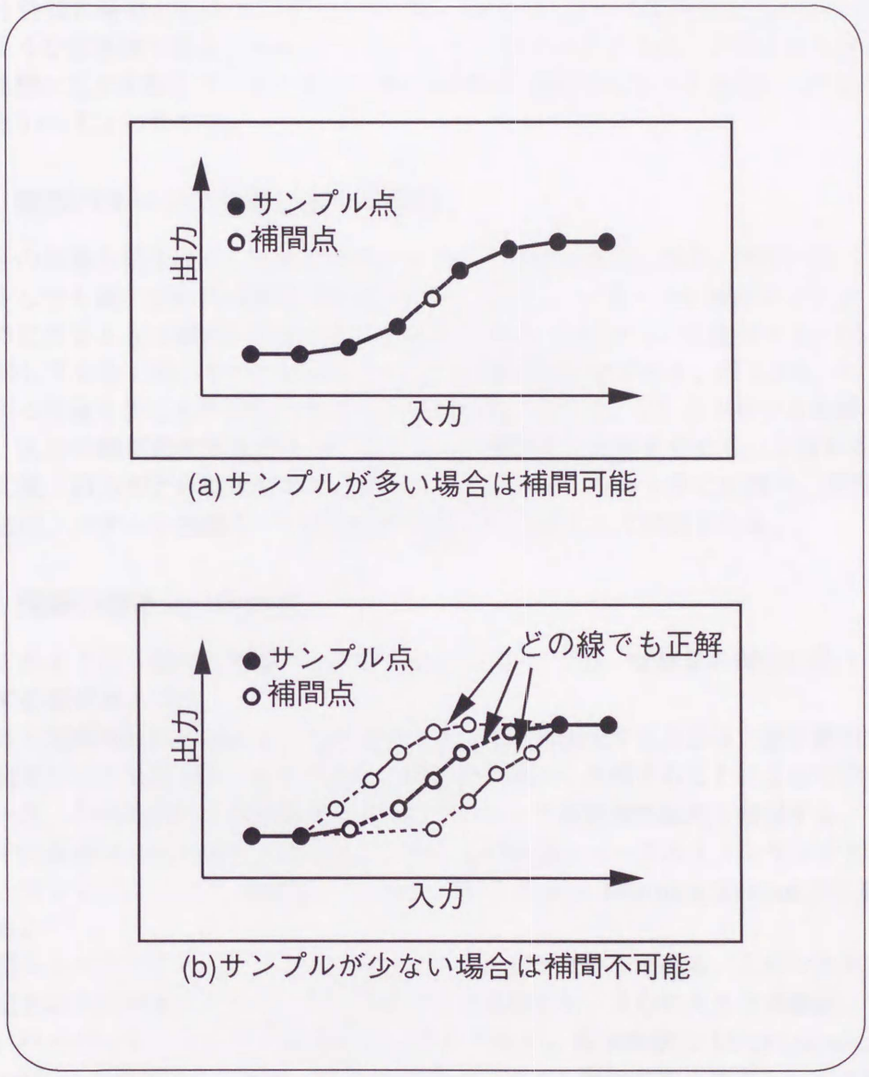


図 1.2: 解が複数あるケース

一方、サンプル数が少ないなら増やせばよいという考え方もある。サンプルに外乱を加え、やや歪んだ信号を作れば、見かけ上のサンプル数が増加する。これによって、サンプル数の不足を補い、未知入力に対する出力を保証する。

これら以外にも、様々なアプローチがある。画像の位置ずれは対象物体の特徴に関与しないという性質に着目したネオコグニトロンは、位置ずれという変化に対して出力の変化を抑えるような回路網である。ネオコグニトロンに代表されるように、入力信号の性質に着目し、回路網に工夫を凝らすことによって適切な関数を構成することを目的としたものは無数にあるといってよいだろう。

1.3.2 類似パターンの分離に関する問題

「双子の兄弟を見分けることができるか」という問題である。本来、類似パターンであっても、少しでも違いがあれば識別は可能である。しかし、パターンの変形などによって全く同じものになるときは識別不可能である。極めて似通ったパターンを識別するには、パターンが変形しても全く同じものにはならないという条件が必要である。図 1.3は、パターンの変形によって全く同じものが発生するケースである。これは、どんな方法でも分離不可能であるが、入力の確率密度を推定することにより、確率的に正解を求めることは可能である。出力の尤度（出力がどの程度信頼できるかという度合い）をあらかじめ調べ、学習させる。この方法は、パターン認識などで誤認識率を減らす方法として利用される。

1.3.3 情報の複雑化への対応

図 1.4 のような一般的な多層パーセプトロン (MLP) では、情報量の増加に伴って素子数も増やす必要がある [30]。

複数の小規模神経回路網によって大規模な神経回路網構成する方法は、素子数の増加に対する解決策として提案されたものである。回路網を細かく分割することによって学習を容易にする一方、各回路網の出力を統合化することによって高度情報処理を実現する。このような手段で大規模神経回路網を実現するモデルには統合型ニューラルネットワーク [31]、選択型ニューラルネットワーク (Comb NET)[27, 28]、 Nestor Learning System (NLS) [32] などがある。

統合型ニューラルネットワークは、図 1.5 のようなシステムである。このシステムは、能力が限定された回路網（サブネット）の出力を複数統合し、さらに大きな回路網（スーパーネット、ハイパーネット）で大統合を行うものである。各回路網は LVQ(Learning Vector Ouantization) で構成されており、その回路網が担当する範囲である確からしさの指標として参照ベクトルとのユークリッド距離を用いている。単純化された課題を学習させるため、学習が容易になる。個々の課題による局所的最小値に陥いる危険が減少するし、正解が一意に決定するようにパラメータを設定することも容易である。従って、課題の複雑化や入力パターンの増加への対応が可能になる。

Comb NET は、図 1.6 のようなシステムである。 このシステムでは、始めに大分類、次

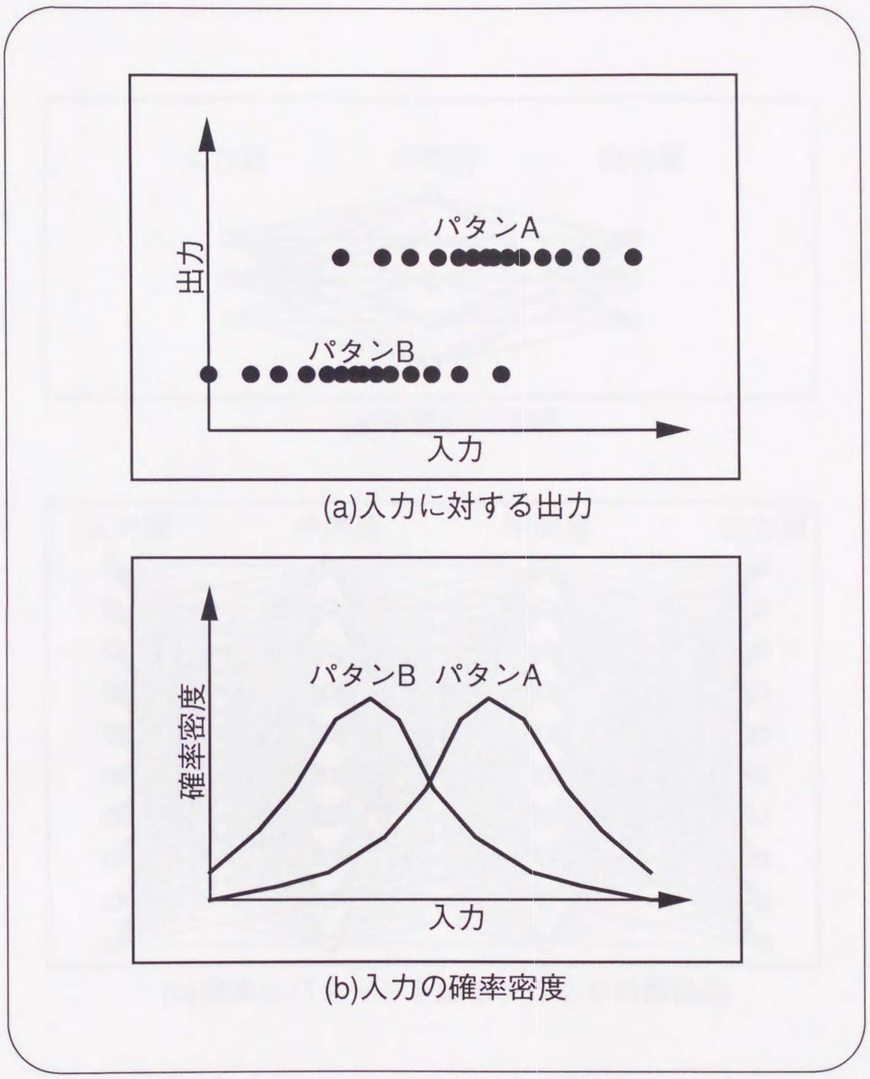


図 1.3: 本質的に分割できないケース

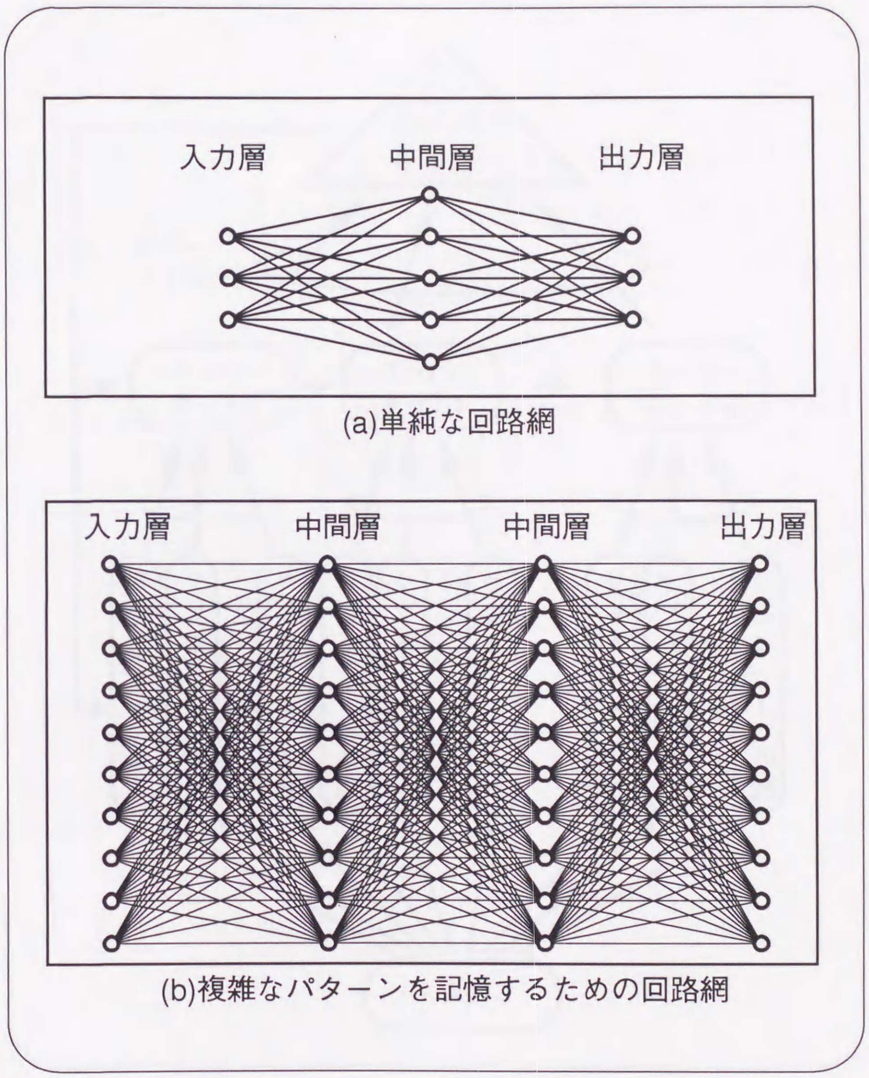


図 1.4: 多層パーセプトロン

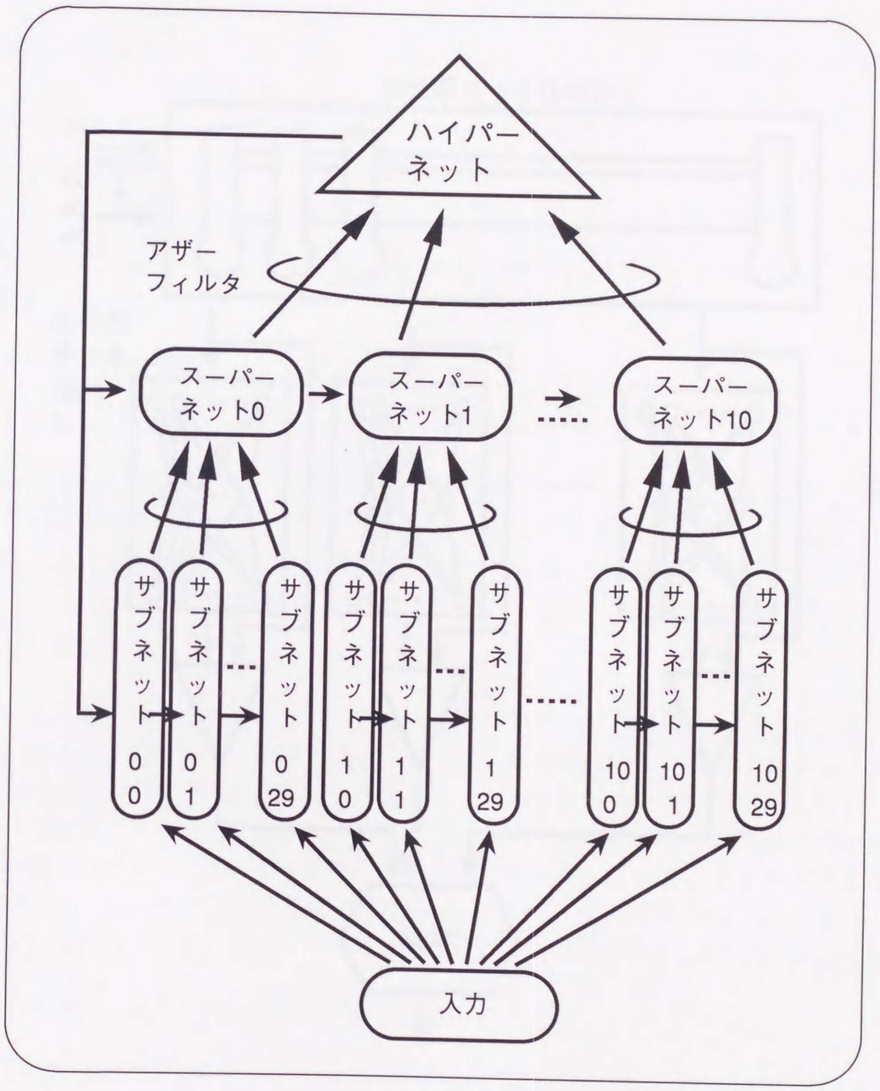


図 1.5: 統合型ニューラルネットワーク [31]

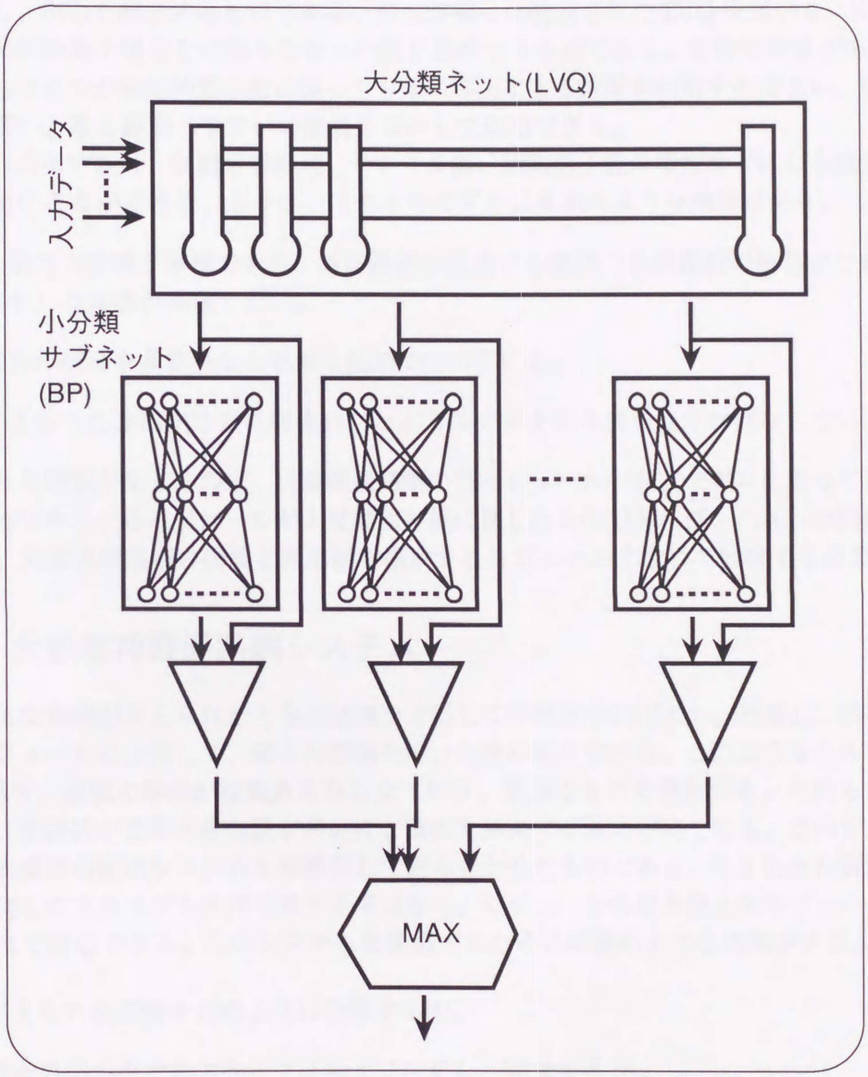


図 1.6: 選択型ニューラルネットワーク (Comb NET)[27]

に小分類を行った後、出力を統合する。大分類にはLVQ、小分類にはMLPを用いる。この方法は、各回路網が限られた範囲の判別を担当するので学習の収束性、認識率の点で優れている。

NLSは、図 1.7 のようなシステムである。このシステムは、複数の特徴を個別の回路網で処理し、出力を統合するものである。各回路網には改良されたLVQを用いる。NLSの長所は、局所的最小解などに陥らなかった解を選択できる点である。複数特徴量がある場合、そのうちひとつが局所的最小値に陥っても他の解法による結果を利用すればよい。複数の特徴量を用い、最も識別しやすい特徴量を選択して利用できる。

これらのモデルは、課題の複雑化、サンプル数、細胞素子数の増加などによる識別能力の低下を防ぐことができる。しかし、これらのモデルにも次のような問題がある。

- 入出力の仕様が単純であり、各回路網が担当する範囲（各回路網の解答がどれだけ確実か）の指標が明確でない。
- 出力の統合を多数決など単純な処理で行っている。
- 与えられた課題がさらに複雑化したときモデルを作り直さなければならない。

これらの問題が生じるのは、大規模回路網の各モジュールを既存モデルによって構成しているためである。各モジュールが大規模回路網に適したものであれば、これらの問題は解決できる。大規模回路網の構成と回路網を構成するモジュールを改めて検討する必要がある。

1.4 大規模神経回路網システム

複合的な課題が与えられたときは課題を分解して単純化すればよい。同様に、回路網も複数のモジュールに分解して、個々の課題を解いた後に統合化する。このようなシステムを図 1.8 に示す。課題の解法が複数あるなら全て解き、最適なものを選択する。これらが可能であれば、回路網が複雑で細胞数が多いほど課題を解決する能力が高くなる。このシステムは従来の大規模回路網システムを単純化して組み合わせたものである。与えられた課題がさらに複雑化したときモデルを作り直す必要はなく、モジュールの置き換えやモジュール数を増やすことで対応できる。このシステムを構築するためには次のような問題がある。

- 与えられた課題をどのように分解するか。
- 個々の出力をどのようにフィルタリングし、統合するか。
- 個々のモジュールにはどのような入力を与え、どのように出力を制御すればよいか。
- 個々のモジュールにはどの程度の能力が必要か。
- モジュールを構成する基本素子はどのようなものか。

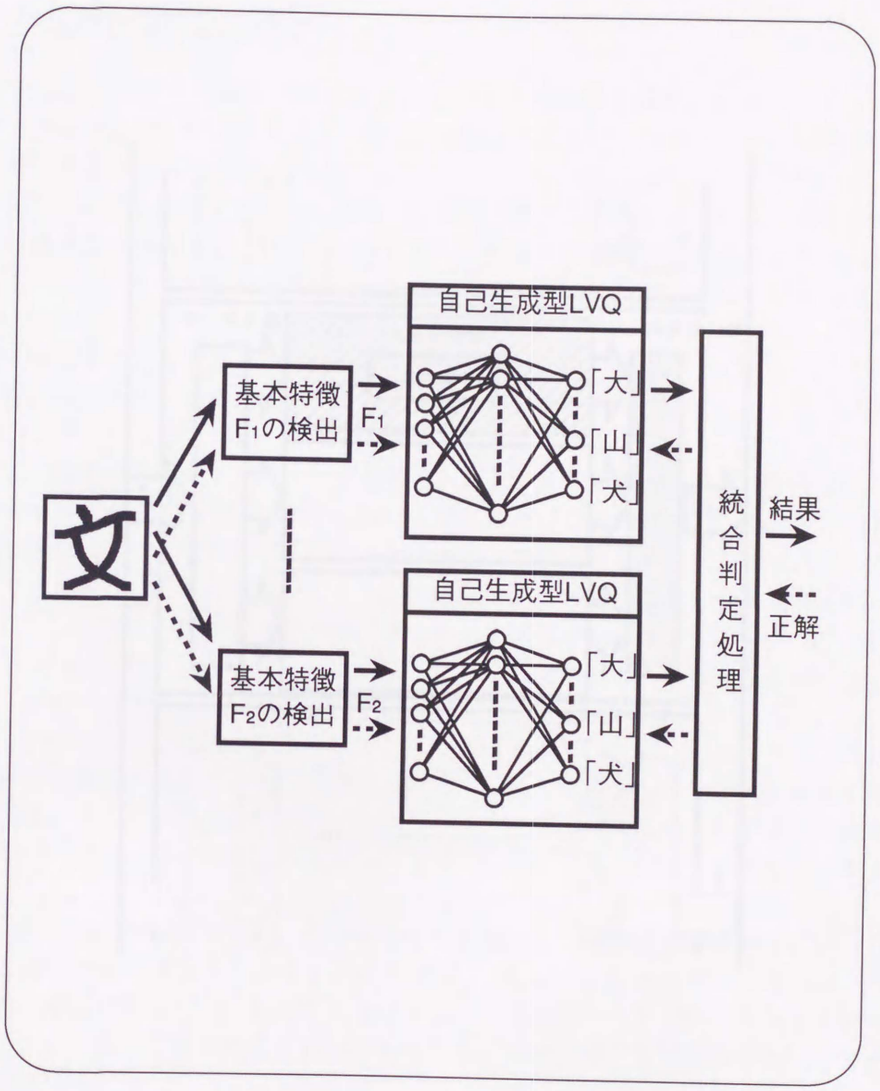


図 1.7: Nestor Learning System(NLS)[32]

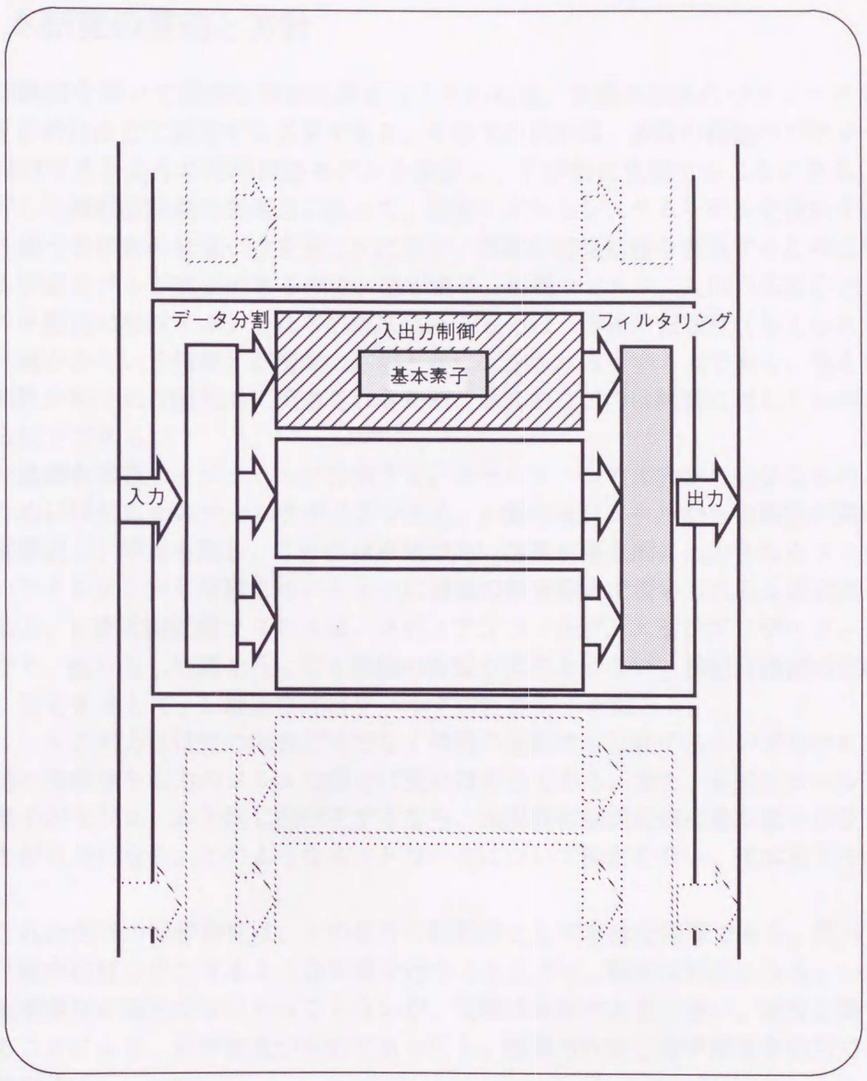


図 1.8: 大規模神経回路網のブロック図

与えられた課題をどのように分解するかは課題によって異なる。個々のモジュールに必要とされる能力も同様である。本論文では、入出力の制御、情報フィルタ、基本素子モデルについて検討する。

1.5 本研究の目的と方針

神経回路網を用いて複雑な情報処理を行うためには、多数の細胞のパラメータやコネクションを目的に応じて設定する必要がある。本論文の目的は、多数の細胞のパラメータ等を簡単に設定できるような神経回路モデルを提案し、工学的に実現することにある。そのため、分析した神経回路網の問題点に従って、要素モデルとシステムモデルを検討する。本論文で取り扱う各技術の位置づけを図 1.9 に示す。複雑な情報処理を実現するためには基本部品となる要素モデルが重要であるので、基本素子、情報フィルタ、入出力応答など、特に重要なものを個別に検討する。それらの根底にあるものは、情報の信頼性（与えられた情報がどれだけ確かかという指標）が情報の内容と共に伝達されるという点である。与えられた情報の信頼性が高ければ優先的に出力されるため、最終的な出力は雑音に対してロバストなものになるはずである。

神経回路網を複数のモジュールに分割する。各モジュールの出力から必要なものだけを統合するためには何らかのフィルタが必要である。n 個のモジュールから信頼性が高いものだけを k 個選択し、平均を取る。これは信頼性が高い情報が優先的に出力されるフィルタである。このフィルタと同じ原理を用いるものに画像の雑音除去で用いられる k 最近接近傍フィルタがある。k 最近接近傍フィルタは、メディアンフィルタ、エッジプリザービングフィルタと異なり、繰り返し処理を行っても画像の特徴が損なわれない。神経回路網の技術である焼き鈍し法を生かして、k 最近接近傍フィルタの性能向上を試みる。

モジュールの出力は情報の内容だけでなく情報の信頼性も反映するものでなければならない。情報の信頼性を出力のノルムで表せば統合は容易である。また、各モジュールを構成する基本素子がモジュールと同じ動作をするなら、大規模神経回路網は基本素子単位まで分割することが可能になる。このようなネットワークについて検討を行い、基本素子モデルを提案する。

入力された信号の確率密度は、その信号の信頼性として有効な情報である。従って、情報の内容と確率密度を出力するような学習を行うことにより、統合は容易になる。一般の学習法では確率密度が既知でなければならないが、実際は未知のときが多い。信号と雑音を交互に与えることにより、確率密度が未知であっても、情報の内容と確率密度を出力できる学習方法を検討する。

さらに、各要素モデルを組み合わせて、神経回路網を実際に構成してみる。検討した全ての要素を盛り込むのは困難なので、必要最小限の要素のみ用いて実用的な回路網を構成する。ターゲットにするアプリケーションは、印刷文字の雑音除去や手書き文字認識である。これらのターゲットに提案手法を適用し、その有効性を確認する。

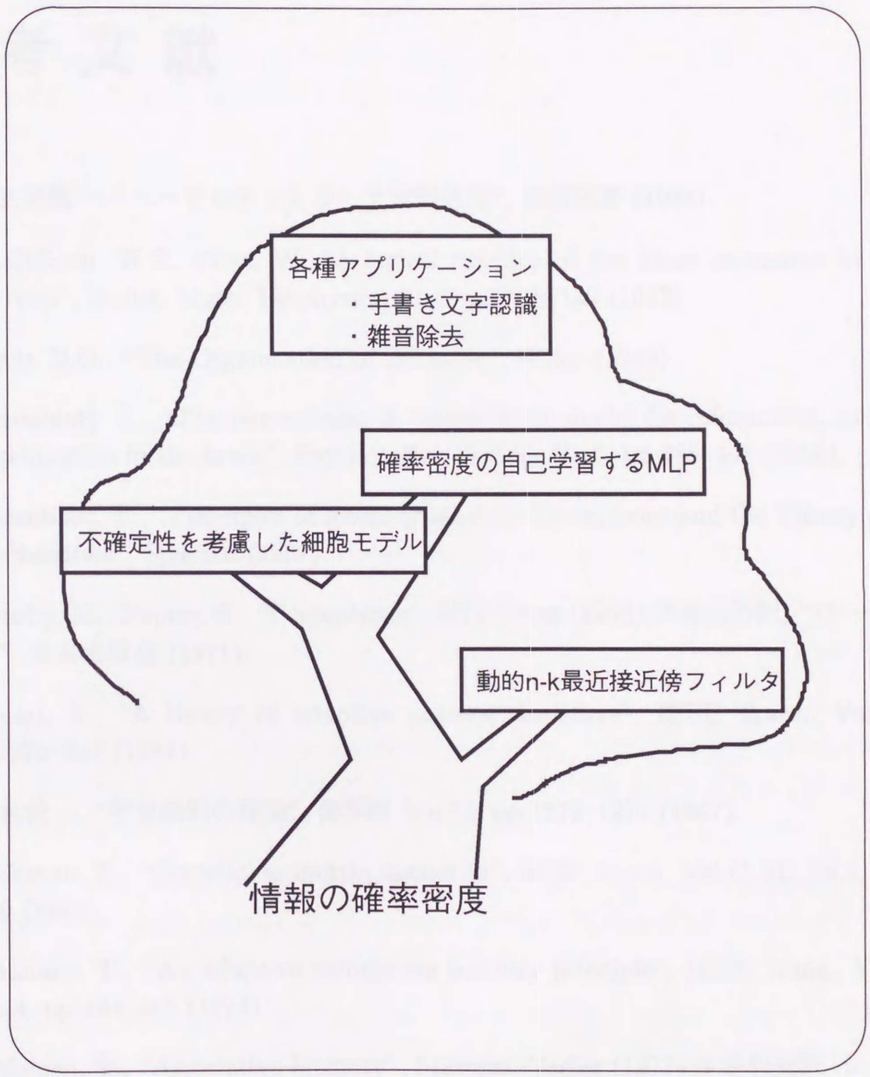


図 1.9: 各技術の位置づけ

参考文献

[1] 麻生英樹, "ニューラルネットワーク情報処理", 産業図書 (1988).

[2] McCulloch, W.S., Pitts, W. "A logical calculus of the ideas immanent in nervous activity", Bullet. Math. Biophysics, No.5, pp.115–133 (1943).

[3] Hebb, D.O., "The Organization of Behavior", Wiley (1949).

[4] Rosenblatt, F., "The perceptron: A probabilistic model for information sorage and organization in the brain", Psychol. Rev. Vol.65, No.6, pp.386–408 (1958).

[5] Rosenblatt, F., "Principles of Neurodynamics—Perceptrons and the Theory of Brain Mechanisms", Spartan (1961).

[6] Minsky, M., Papert, S., "Perceptrons", MIT Press (1969) 斉藤正男訳, "パーセプトロン", 東大出版会 (1971).

[7] Amari, S., "A theory of adaptive pattern classifiers", IEEE Trans., Vol.EC–16, pp.279–307 (1967).

[8] 甘利俊一, "学習識別の理論", 信学誌 Vol.50, pp.1272–1279 (1967).

[9] Kohonen, T., "Correlation matrix memories", IEEE Trans., Vol.C–21, No.4, pp.353–359 (1972).

[10] Kohonen, T., "An adaptive associative memory principle", IEEE Trans., Vol.C–23, No.4, pp.444–445 (1974).

[11] Kohonen, T., "Assosiative Memory", Springer–Verlag (1977) 中谷和夫訳, システム論的連想記憶, サイエンス社 (1980).

[12] Kohonen, T., "Self–organization and Associative Memory", Springer–Verlag (1984).

[13] Kohonen, T., Ruohonen, M., "Representation of associative data by matrix opetators ", IEEE Trans., Vol.C–22, No.7, pp.701–702 (1973).

[14] Nakano, K.,"Associatron—a model of associative memory", IEEE Trans., Vol.SMC–2, pp.380–388 (1972).

[15] Fukushima, K., "Cognitron: a self-organizing multilayerd neural network", Biol. Cybernetics, Vol.20, pp.121–136 (1975).

[16] Fukushima, K., Miyake, S. "A self-organizing neural network with a function of associative memory: Feedback-type cognitron", Biol. Cybernetics, Vol.12, pp.204–215 (1978).

[17] Fukushima, K., Miyake, S., Ito, T, "Neocognitron: a neural network model for a mechanism of visual pattern recognition", IEEE Trans., Vol.SMC–13, pp.826–834 (1983).

[18] Hirai, Y., "A model of human associative processor(HASP)", IEEE Trans., Vol.SMC–13, pp.851–857 (1983).

[19] Hirai, Y., "Mutually linked HASP's: a solution for constraint–satisfaction problems by associative processing", IEEE Trans., Vol.SMC–15, pp.432–442 (1985).

[20] Amari, S., Arbib, M.A., "Competition and cooperation in neural nets", In Systems Neuroscience, J. Metzler, ed., pp.119–165, Academic Press (1977).

[21] Hopfield, J.J., "Neural networks and physical systems with emergent collective Computational abilities", Preceedings of the National Academy of Sciences USA 79, pp.2254–2258 (1982).

[22] Hopfield, J.J., "Neurons with graded responce habe collective computational properties like those of two–state neurons", Proceedings of the National Academy of Sciences USA 81, pp.3088–3092 (1984).

[23] Farlman, S.E., Hinton, G.E., Sejnowski, T.J., "Massively parallelarchitectures for AI: NETL, Thistle, and Boltzman Machines", In Proceedings of the National Conference on Artificial Intelligence AAAI–83, pp.109–113 (1983).

[24] Geman, S., Geman, D., "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", IEEE Trans., Vol.PAMI–6, pp.721–741 (1984).

[25] Rumelhart, D.E., Hinton, G.E., Williams, R.J., "Learning represontations by back–propagation errors", Nature, Vol.323–9, pp.533–536 (1986).

[26] Kohonen, T., "Learning vector quantization for pattern recognition", Helsinki University of Technology, Laboratory of Computer and Information Science, Report TKK–F–A–601 (1986).

[27] 岩田彰, 當麻孝志, 松尾啓志, 鈴木宣夫, "大規模 4 層ニューラルネット Comb NET", 信学論 (D–II), Vol.J73–D–II, No.8, pp.1261–1267 (1989).

[28] 堀田健一, 岩田彰, 松尾啓志, 鈴村宣夫, "大規模ニューラルネット "CombNET–II" ", 信学論 (D–II), Vol.J75–D–II, No.3, pp.545–553 (1992).

[29] 石川真澄, "忘却を用いたコネクショニストモデルの構造学習アルゴリズム", 人工知能学会誌, Vol.5, No.5, pp.595–603 (1990).

[30] 久間和生, 中山高, "ニューロコンピュータ工学", 工業調査会 (1992).

[31] Y.Mori, K.Joe, "A Large–Scale Neulal Network witch Recognizes Handwritten Kanji Characters", Advances in Neural Information Processing Systems, 2, pp.415–421 (1989).

[32] 依田, 根岸, 岡田, 小林, "複合 LVQ ニューラルネットワークによる印刷漢字認識方式の検討", 平成 3 年度後期情報処理学会予稿集, pp.1267 (1991).

第 2 章

動的 n–k 最近接近傍法による画像のスムージング

本章のあらまし

k 最近接近傍フィルタは、画像の雑音除去に有効なフィルタであるが、パラメータの値によって効果が左右される。本研究は、k 最近接近傍フィルタの振舞いを近傍領域のヒストグラムと画像の性質というふたつの立場から検討することによって、その欠点の原因を探り、それを克服した新しい動的 k 最近接近傍法を提案する。すなわち、k 最近接近傍フィルタのスケジューリング方法について新しい提案を行う。それは、ニューラルネットワークにおけるシミュレーテッドアニーリングを応用して、クリークの要素数 k と、近傍領域の画素数 n を徐々に減少させるという方法である。この方法では、アルゴリズム上の負担が少ない上、k 近傍フィルタでは困難なパラメータの設定が容易であり、かつ雑音除去効果においてすぐれたものとなる。これらを実験を通して確認する。

2.1 はじめに

画像処理の分野において、画像のスムージングは重要かつ基本的なものであり、これまでにも多くの研究例がある [1]。近傍領域の平均をとるだけの単純平滑化フィルタは、画像の雑音除去に大きな効果があるが、物体のエッジをぼかしてしまうという欠点がある。エッジ抽出を目的とした場合この欠点は致命的であるため、エッジを保持したまま画像のスムージングを行う方法がさまざまな形で提案されてきた。代表的なものとしてエッジプリサービングフィルタ [2]、メディアンフィルタ [3]、k 最近接近傍フィルタ [4] などがあり、エッジ抽出の前処理としてのスムージングフィルタとして広く知られている [5]。

最近の傾向としては、画像をマルコフ場としてとらえて、統計的緩和法によるスムージングを行う方法についての検討が盛んに行われている。エネルギー最小化問題を画像の初期知覚に対応させる方法は、Geman & Geman により提案されたが、その有効性は広く認められ [6, 7]、この理論によって画像の初期知覚に関するさまざまな理論を統一的に行うことが可能になった。

この方法の特徴は画像に対する事前知識を導入し、それを有効に活用するところにある。しかし、この事前知識は学習などによって与えられれなければならない [8]。事前知識が与えられていない特徴に対しては、この方法の適用は望ましくない。そのため、複雑な図形を認識するときには、事前知識の量を増やし、さまざまなケースに対応させる必要があるが、すべてのケースに対応できるようにすると事前知識自体が存在意義を失う。一方、k 最近接近傍フィルタは、メディアンフィルタやラインプロセスを用いた統計的弛緩法に比べて雑音に弱いことが欠点であるが、事前知識を必要とせず、複雑なパターンにも応用が可能である。しかし、この方法は雑音の大きさと kの値に密接な関係があるため、kの値が不適切であると雑音が除去されなかったり、エッジがなまってしまったりするという欠点がある。

本論文では、k 最近接近傍フィルタの振舞いを近傍領域のヒストグラムと画像の性質というふたつの立場から検討し、kの値及び近傍領域のサイズ n を動的に変化させる新しい動的 n–k 近傍フィルタの提案を行う。また、1 次元信号と 2 次元画像の雑音除去実験によってこの手法の効果を確かめることとする。

2.2 ヒストグラムからの検討

k 最近接近傍フィルタの原理は次のようなものである。図 2.1のような画像の雑音除去を行う場合を考える。画像がいくつかの領域（クリーク）に分けられると仮定する雑音を除去するためには、注目点の真の値を周辺の輝度値から推定する必要がある。注目点と同じ性質（輝度）を持つ画素（クリーク要素）を注目点の近傍領域から取出し、平均を取ることによって真の値を推定する。k 最近接近傍フィルタでは、n 個の近傍画素の内にクリーク要素が k 個以上あると仮定する。クリーク要素の輝度値は注目点の輝度値 d に近い値を持つ。従って、n 個の近傍画素の中から輝度値が d に近いものを k 個取り出せば、これらはクリーク要素であるといえる。注目点の真値は、k 個のクリーク要素の輝度平均である。

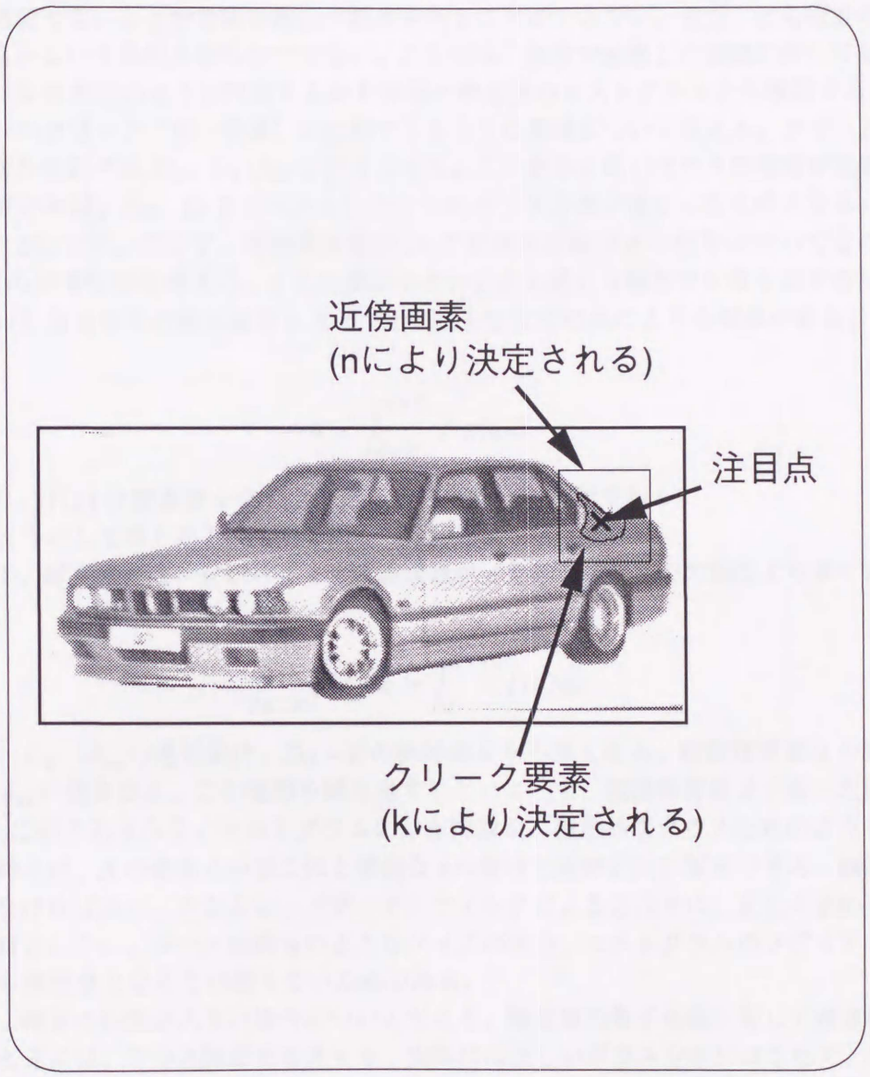


図 2.1: k 最近接近傍フィルタの原理

k 最近接近傍フィルタのアルゴリズムは次のようなる。まず、注目点の輝度に対して n 個の近傍領域内の画素の輝度値との差を求める。次に、それらの絶対値が小さい方から順に k 個の画素を取り出し、それらの輝度の平均値を結果とする。このようにして絶対値が小さいものを用いてスムージングをする事によって、雑音の除去が可能となる。この方法では、k の値が適切でないと十分な雑音除去の効果を得ることができない。また、どの程度の雑音に耐え得るかという条件も明らかではない。ここでは、雑音が重畳した画像に対して k 最近接近傍フィルタがどのように作用するかを画像の輝度値のヒストグラムから検討する。

ふたつのクリーク（同一領域）に分割できる小さな領域について考える。クリークの理想的な輝度をそれぞれ D_0、$D_1(D_0 < D_1)$ とする。この信号に弱いガウス性雑音が重畳した場合の輝度分布は、D_0、D_1をピークとした 2 つのガウス分布が重なったものとなる。その様子を図 2.2に示す。ここで、理想輝度値が D_0で観測輝度値が d の信号についてどのような kを選ぶのが適切かを考える。d との差が小さい方から順に k個までを取り出すのであるから、そのとき d からの最大偏差を Kとすれば、kと Kには次のような関係がある。

$$k = \int_{d-K}^{d+K} f(x)dx \tag{2.1}$$

ただし、$f(x)$ は画素値 x のヒストグラムを表す分布関数である。

このようにして取り出した k個の画素の平均輝度を d_{av}とする。

ここで、輝度差が $d-K$から d までの画素数が d から $d+K$までの個数より多ければ、すなわち、

$$\int_{d-K}^{d} f(x)dx > \int_{d}^{d+K} f(x)dx \tag{2.2}$$

のとき、$D_0 - d_{av}$の絶対値は、$D_0 - d$ の絶対値より小さくなる。観測輝度値 d の地点の輝度値を d_{av}に置き換え、この処理を繰り返すことによって、観測輝度値 d であった画素は次第に D_0に引き込まれる。ヒストグラムの分布関数が、理想的なガウス分布のような単調な関数の時には、Kの値あるいはこれと等価な kの値は比較的自由に設定できる。極端に大きな値でなければ良い。ちなみに、メディアンフィルタによる方法では、正しく求められない点に注目されたい。ガウス性雑音のようなノイズの場合、ヒストグラムのメディアンの値が必ずしも理想値となるとは限らないためである。

次に、雑音の分散が大きい場合について考える。輝度値の量子化幅に対して雑音の分散が大きいときには、ガウス雑音を与えても、実際には正しいガウス分布にはならず、局所的なピークが生じたヒストグラムを示すことが多い。そのような分布を示す例を図 2.3に示す。

この場合は、kの値の決定は重要である。同図に示す位置に注目画素があるとしよう。このとき、小さな kを用いると、$d-K$から d までの個数が d から $d+K$までの個数より多くないため、D_0には引き込まれずに最も近い局所的なピークに引き込まれてしまう。この処理を繰り返してもそれ以上進まず、局所的なピークが多数残ることになる。

ここで、図 2.4のように、ある程度大きな kを用いてみる。その場合小さな局所的ピーク

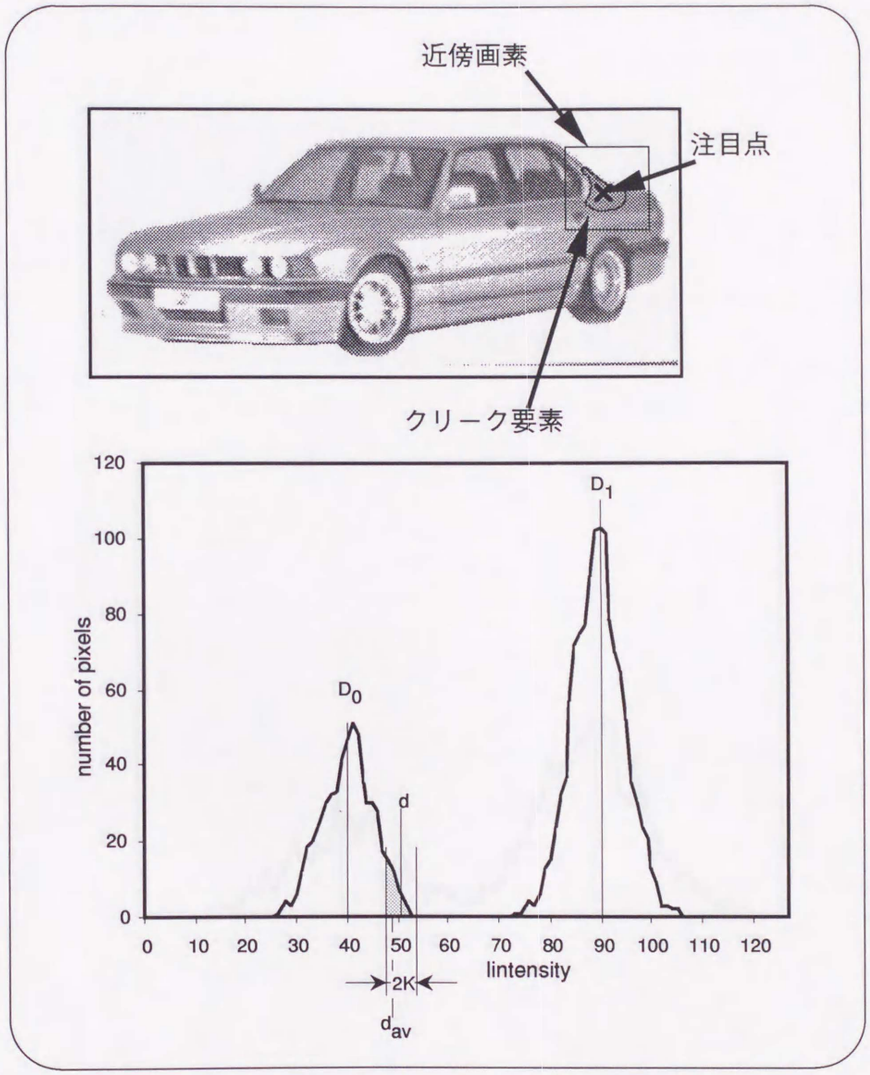


図 2.2: ガウス雑音が付加された画像のヒストグラム

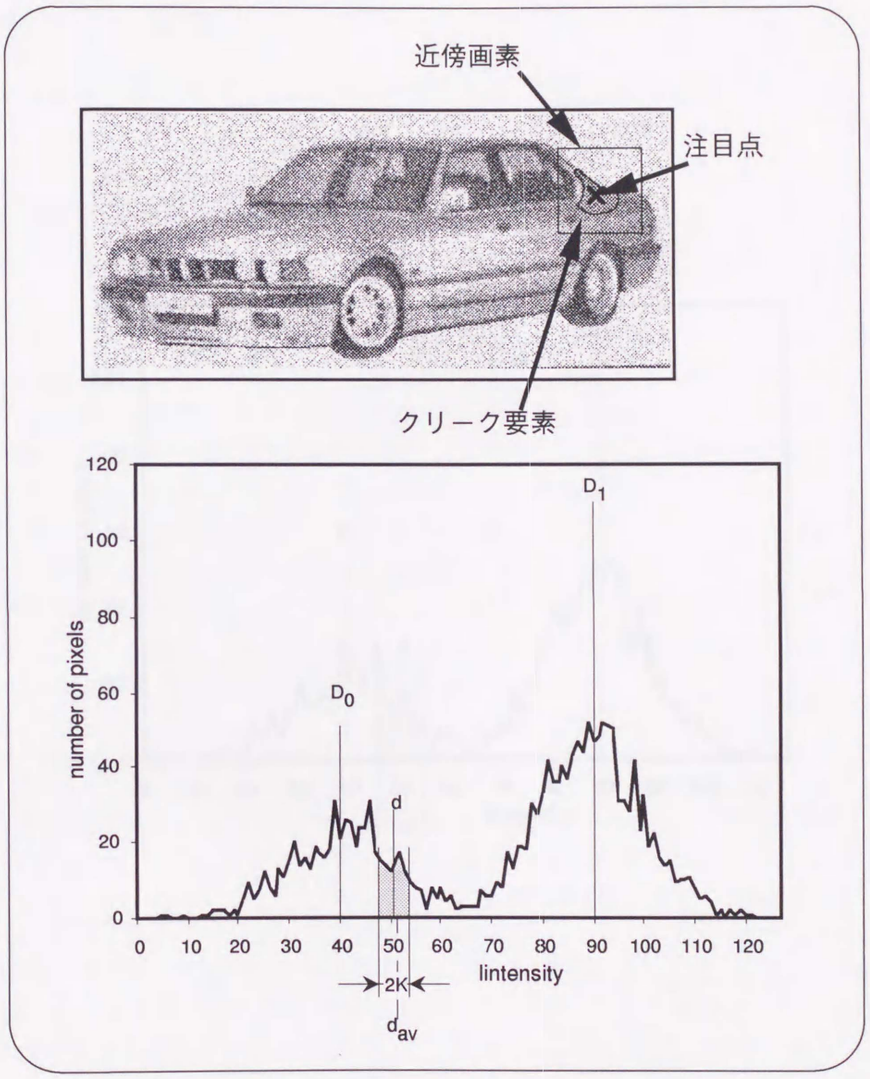


図 2.3: 局所的なピークのある画像のヒストグラム

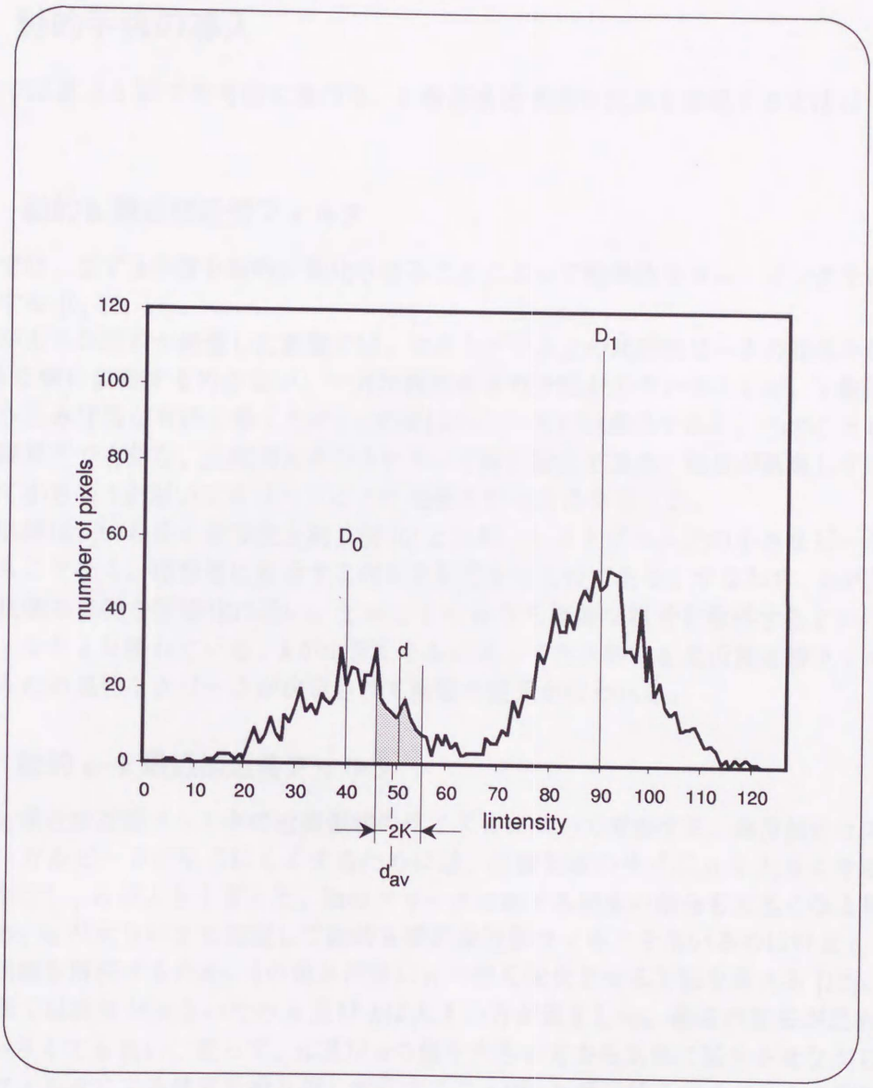


図 2.4: 局所的ピークを乗り越えるパラメータ設定

は無視され、大きなピーク D_0に向かって引き込まれる。ただし、kを極端に大きくとると D_0と D_1もひとつのピークになるように互いに引き込みあうため、この操作を何回も繰り返すのは危険である。

2.3 動的手法の導入

本節では第 2.2 節での考察に基づき、k 最近接近傍法の欠点を克服する方法について考える。

2.3.1 動的 k 最近接近傍フィルタ

ここでは、まず kの値を動的に変化させることによって効果的なスムージングを行う方法を提案する [9, 10, 11]。

分散が大きな雑音が重畳した画像では、ヒストグラム上の局所的ピークの存在を考慮して kは大きな値に設定するのが良い。一方比較的雑音の分散が小さいときには、k 最近接近傍方の引き込み作用が有効に働くため、kの値は小さい方が効果的である。このことから、最初は高雑音下で有効な、比較的大きな kを用いて雑音除去を進め、雑音が低減してゆくにしたがって小さな kを用いてクリークごとの弛緩を行う方法を考える。

この処理は、いわゆる疑似焼き鈍し法 [6] と同様、ヒストグラム上の小さなピークに引き込まれることなく、理想値に接近する効果を期待するものである。すなわち、kが大きいときは、画像の単純な平滑化に近い。しかし $k < n$ なら極端な雑音を除外するという点で平均値フィルタより優れている。kが小さくなるに従って古典的な k 最近接近傍フィルタの特性になるため適切なクリークが決定されて画像の復元が行われる。

2.3.2 動的 n–k 最近接近傍フィルタ

動的 k 最近接近傍フィルタの近傍領域のサイズ n について考察する。輝度値のヒストグラムにローカルピークが生じにくくするためには、近傍領域のサイズ n を大きくする必要がある。ただし、n が大きすぎると、他のクリークに属する画素の割合も大きくなる可能性が増すため、n が大きいまま固定して動的 k 最近接近傍フィルタを用いるのは好ましくない。

この問題を解決するため、kの値と同様に n の値も変化させる方法を考える [12]。処理の初期段階では雑音が大きいため n 及び kは大きい方が望ましい。雑音の除去が進めば n 及び kは小さくても良い。従って、n 及び kの値を大きい方から次第に減少させながら k 最近接近傍フィルタによる処理を繰り返し実行することが、k 最近接近傍フィルタの適切なスケジューリングであるということができる。

2.3.3 具体的な処理手順

動的 n–k 最近接近傍フィルタによるスムージングの具体的な処理手順は次のようになる。

図 2.5: 動的 k 最近接近傍フィルタ

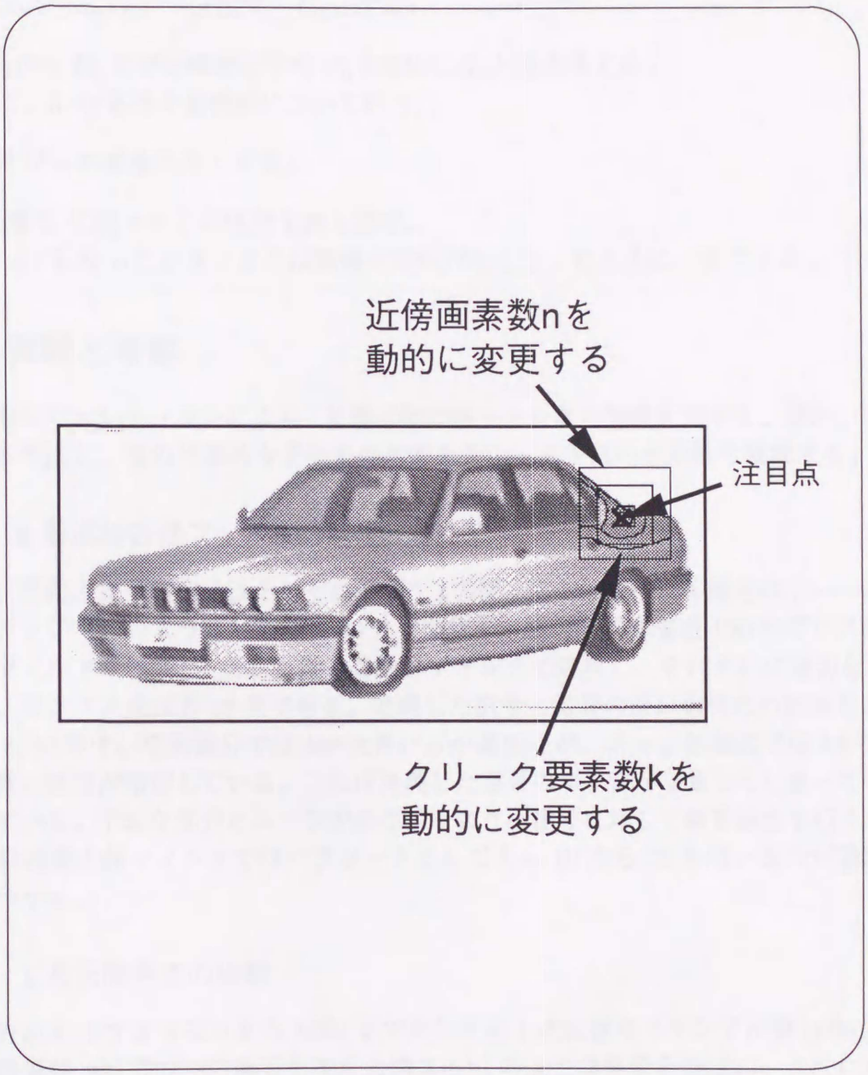


図 2.6: 動的 n–k 最近接近傍フィルタ

1. クリークサイズ k、近傍領域サイズ n を適当に設定する。
2. 注目点の輝度 G_0に対して n 個の接近傍領域内の画素について、輝度の差が小さい順に列べ変える。

   これを $G_0, G_1, \cdots, G_k, \cdots, G_n$とする。
3. G_0から D_kまでの輝度の平均 G_0'を求め、G_0と置き換える。
   (2.、3. の処理を全画素について行う。)
4. k及び n の値を小さくする。
5. 処理 2. に戻ってこの処理を繰り返す。
   $n = 1$ になったとき（または画像の変化がなくなったときに）終了する。

2.4 実験と考察

計算機シミュレーションにより、k 最近接近傍フィルタの特性を調べる。また、本論文で提案する手法と、他の代表的な手法との比較を行い、本手法の有効性を確認する。

2.4.1 k 最近接近傍フィルタの特性比較

エッジ部周辺と平坦部分はそれぞれ適切な kの値を持つ。まったく変化のない一次元信号と、ステップ関数のように 0 から 1000 まで変化する信号に標準偏差 100 のガウス性雑音を加え、フィルタサイズ 21 点の k 最近接近傍フィルタで処理し、それぞれに適切な kの値を調べた。サンプル点は各 20 点である。処理した信号と雑音のない信号との誤差を求めた結果を図 2.7に示す。平坦部分では kが大きい方が適切だが、エッジ部周辺では kが大きすぎると急激に誤差が増加している。これは処理した信号のエッジがなまってしまっていることを示している。平坦な部分とエッジが共存するような信号に対して雑音除去を行う場合、従来の k 最近接近傍フィルタではパラメータとして $k = 10$ から 12 を用いるのが適切であることがわかる。

2.4.2 1 次元信号での比較

図 2.8(a) に示すような 0 から 1000 まで変化する 1 次元信号（サンプル数 1800 点）に対して標準偏差 100 のガウス雑音を加えた図 2.8(b) のような信号を作成し、これに対して雑音除去実験を行った。(動的 n–k 最近接近傍フィルタ以外のフィルタサイズはいずれも 21 点。) 平均値フィルタとメディアンフィルタの処理結果をそれぞれ図 2.8(c)、図 2.8(d) に示す。平均値フィルタでは境界の鋭さは失われている。一方、メディアンフィルタ処理では境界の鋭さが完全に失われてしまうわけではないがややなだらかになっている。

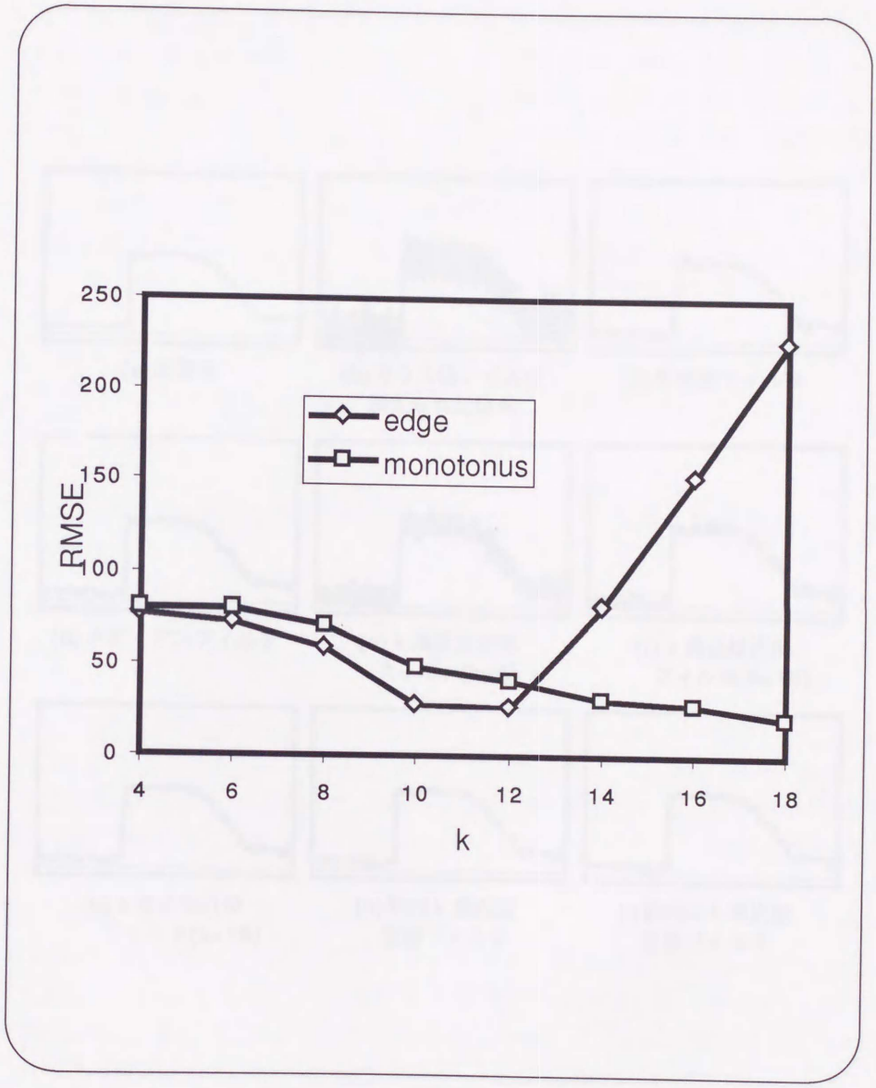


図 2.7: k 最近接近傍フィルタの k と二乗平均誤差 (RMSE) との関係

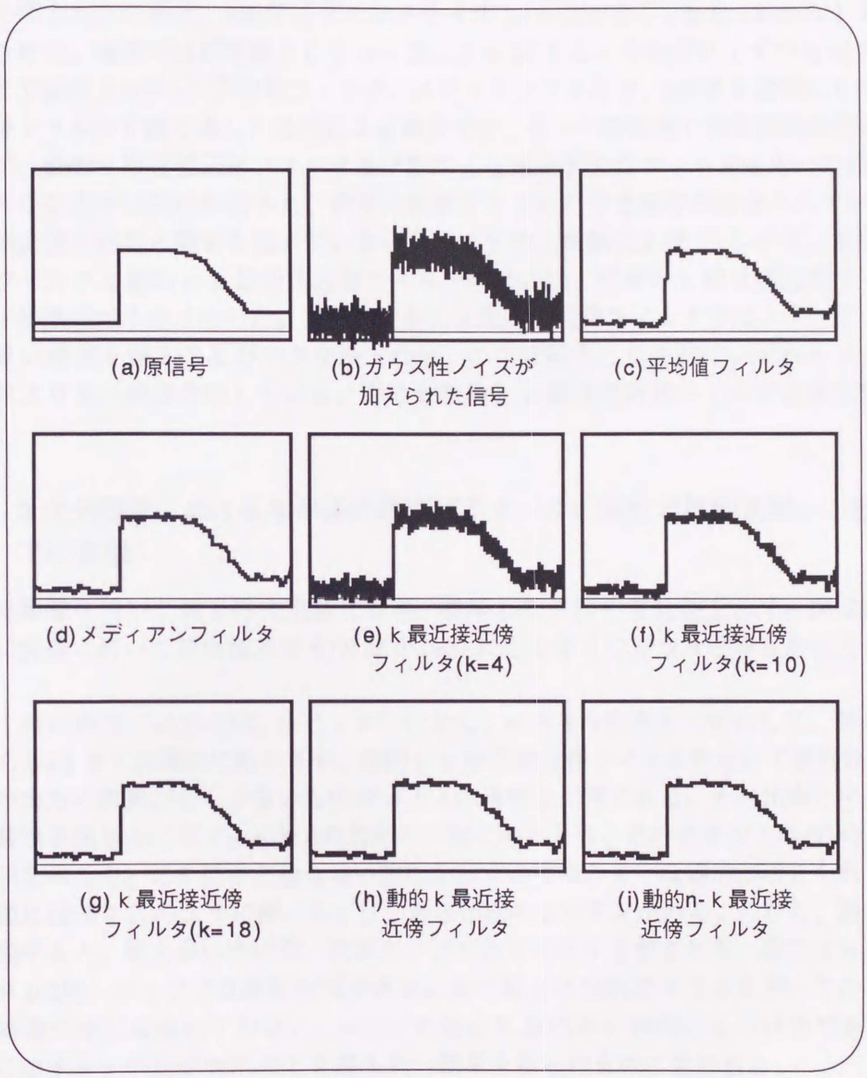


図 2.8: 1 次元信号の雑音除去実験

図 2.8(e)〜(g) に k 最近接近傍フィルタによる処理結果を示す。(繰り返し回数はいずれも 5 回。) 図 2.8(g) のように比較的適切な値 ($k = 10$) を用いた場合にはエッジは残り、雑音もある程度除去されているが、完全には除去されていない。

動的 k 最近接近傍フィルタの処理結果を図 2.8(h) に、動的 n–k 最近接近傍フィルタの処理結果を図 2.8(i) に示す。(前者はフィルタサイズ (n = 21) とし、k を 18 から 1 まで 1 ずつ減少させた。後者では初期値として $n = 21$、$k = 18$ とし、それぞれ 1 ずつ $n = 7$、$k = 4$ になるまで減少させた。) 平均値フィルタ、メディアンフィルタ、kの値を固定したまま k 最近接近傍フィルタを繰り返した処理による結果では、各々の特徴通り境界領域の形がゆがんでいるが、動的 k 最近接近傍フィルタ及び動的 n–k 最近接近傍フィルタを用いた結果では、このような欠点が比較的解消され、信号の特徴をとどめたまま雑音が除去されている。

雑音除去後の信号と雑音を加えていない信号の平均二乗誤差を表 2.1に示す。動的 k 最近接近傍フィルタと動的 n–k 最近接近傍フィルタの出力は、従来の k 最近接近傍フィルタより二乗平均誤差が小さくなった。また、従来の k 最近接近傍フィルタではメディアンフィルタより良い効果を得ることができなかったが、ここで提案した手法はいずれもメディアンフィルタより良い結果を示している。中でも動的 n–k 最近接近傍フィルタは最も効果的であった。

2.4.3 2 次元画像における本手法の最適パラメータの測定と性能比較—1 枚の画像での実験

2 次元画像を用いた雑音除去実験により、提案した手法の有効性を示す。図 2.9(a)、図 2.9(b) に実験に用いた原画像及び S/N 比が 0dB になるようにガウス性雑音を加えた画像を示す。

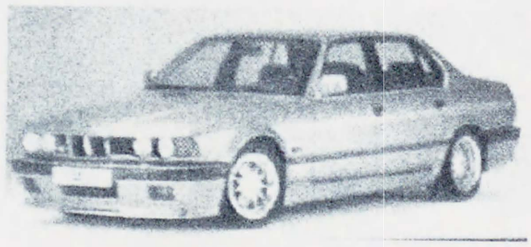
この 1 枚の画像 (512×256、8 ビット) に対し、n と kの比率を一定にして、近傍領域を 9×9 から 3×3 まで段階的に縮小させ、動的 n–k 最近接近傍フィルタを用いて雑音除去を行った。この出力の誤差が最も少ないものが n と kの適切な比率である。その比率と平均二乗誤差との関係を図 2.10に示す。n と kの比率が一定であるとき、その比率が 100:60 のとき最も良い結果を示した。この比率が最も良い性能を示すのは次のような理由が考えられる。一般的な画像には図 2.11のように面、エッジ、角など様々なパタンがある。しかし、画像を小領域に分割すると、最も多いのは面、次がエッジであることが予想される。面では n に対する k の比が 100%、エッジでは最低 50%である。角の場合は 50%以下になるが、このようなパタンは画像の中には極めて少ない。エッジの場合も 50%から 60%のものは希である。従って、n に対する k の比が 60%のとき最も良い結果を示したものと思われる。

次に本手法と従来の手法との比較実験について述べる。従来の手法として、平均値フィルタ、メディアンフィルタ、エッジプリザービングフィルタ及び従来の k 最近接近傍フィルタを用いた。エッジプリザービングフィルタには SPIDER [13] のサブルーチンを用いた。k 最近接近傍フィルタは、kの値が大きいとき ($k = 35$)、小さいとき ($k = 25$)、適切なとき ($k = 30$) の 3 種類を用い、9 回繰り返した。いずれも近傍領域として 7×7 のマトリクスを

表 2.1: 1 次元信号の雑音除去結果

フィルタ	n	k	平均二乗誤差
雑音信号	—	—	102.51
平均値フィルタ	21	—	39.67
k 最近接近傍フィルタ	21	4	77.04
k 最近接近傍フィルタ	21	10	42.11
k 最近接近傍フィルタ	21	18	33.32
メジアンフィルタ	21	—	26.89
動的 k 最近接近傍フィルタ	21	18〜1	26.62
動的 k–n 最近接近傍フィルタ	21〜7	18〜4	25.27

(a) 原画像



(b) 雑音を加えた画像

図 2.9: 雑音除去実験に用いた画像

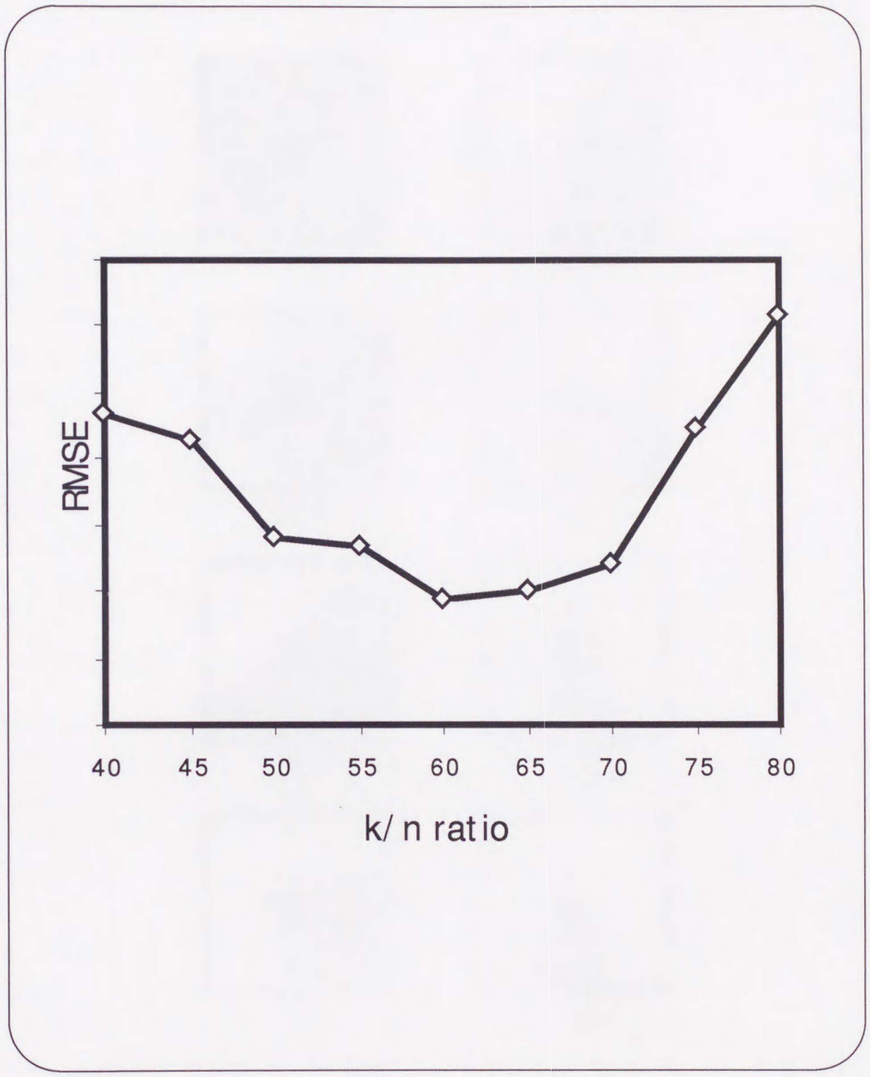


図 2.10: n–k 最近接近傍フィルタにおける n と k の比と二乗平均誤差 (RMSE) の関係

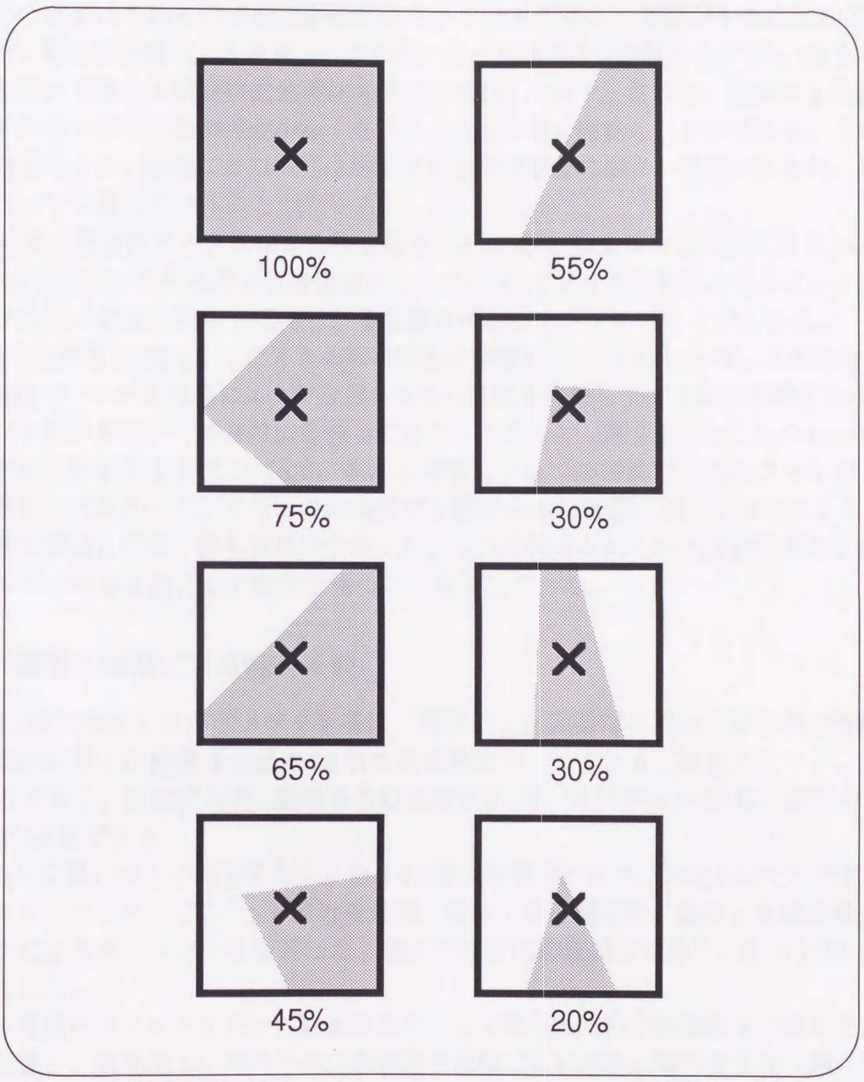


図 2.11: 様々な画像パタン

用いた。

平均値フィルタ、メディアンフィルタ、エッジプリザービングフィルタによる雑音除去結果を図 2.12(a),(b),(c) に示す。平均値フィルタ、メディアンフィルタでは雑音が十分に除去されていない上に、画像全体がぼやけたものとなってしまっている。エッジプリザービングフィルタの結果はそれらの点が比較的改善されているが細かい情報がやや失われている。

従来の k 最近接近傍フィルタで、この画像における適切な kの値として $k = 30$ をパラメータとして用いた場合の雑音除去処理結果を図 2.12(d) に示す。これは、従来の k 最近接近傍フィルタの中で最も二乗誤差が少なくなるように kの値を設定したものである。このフィルタの出力を見ると、領域ごとの境界は明瞭であるが原画像の細かい情報が失われ、平面的な画像となってしまっていることがわかる。

動的 k 最近接近傍フィルタによる雑音除去（$k = 33$ から $k = 5$ に変化）を行った結果を図 2.12(e) に示す。この処理も近傍領域として 7×7 のマトリクスを用いている。この結果ではエッジなどの本来の情報を保持したまま雑音が除去されていることがわかる。

n と kの比率を約 60%としたときの雑音除去の結果を図 2.12(f) に示す。この場合も動的 k 最近接近傍フィルタと同様にエッジなどの本来の情報を保持したまま雑音が除去されている。

各種フィルタを用いて雑音除去を行って得られた出力と、雑音を付加していない原画像との間の平均二乗誤差を表 2.2に示す。新たに提案した 2 つの手法では他のフィルタより雑音除去の効果が大きかった。また、この実験でも動的 n–k 最近接近傍フィルタによる出力は 1 次元信号の実験と同様、最も効果的であった。この結果はこれらの手法が従来の k 最近接近傍フィルタの限界を越える可能性を持つことを示している。

2.4.4 複数の画像による性能比較

第 2.4.3節で測定したパラメータを基に、複数の 2 次元画像を用いて雑音除去性能を測定した。実験に用いた画像は全部で 48 枚の白黒画像で、いずれも 1 画素 8 ビット、サイズは 256×256 である。内訳は人間、動物を含む風景が 15 枚、風景のみが 12 枚、静物が 6 枚、人間の顔が 15 枚である。

これらの画像に対して雑音を加え、第 2.4.3節と同様のパラメータを用いて平均値、メディアン、エッジプリザービング、k 最近接近傍、動的 k 最近接近傍、動的 n–k 最近接近傍の各フィルタによるスムージングを行った。加えた雑音は白色雑音であり、S/N 比は 1dB となるようにした。

上記 6 種類のフィルタを用いて雑音除去を行って得られた出力画像および雑音を加えた処理前の画像と、原画像との間の平均二乗誤差を画像ごとに求め、雑音が少ない順に 1 位から 7 位までの順位付けを行った。

各フィルタごとの平均二乗誤差の合計と順位の合計を表 2.3に示す。新たに提案した 2 つの手法の内、動的 k 最近接近傍フィルタによる出力は順位の合計得点による評価ではあまり高くないが、雑音の絶対値が大きいものでは良い効果が得られたため、平均二乗誤差はそれほど大きくない値となっている。また、動的 n–k 最近接近傍フィルタによる出力は総合的

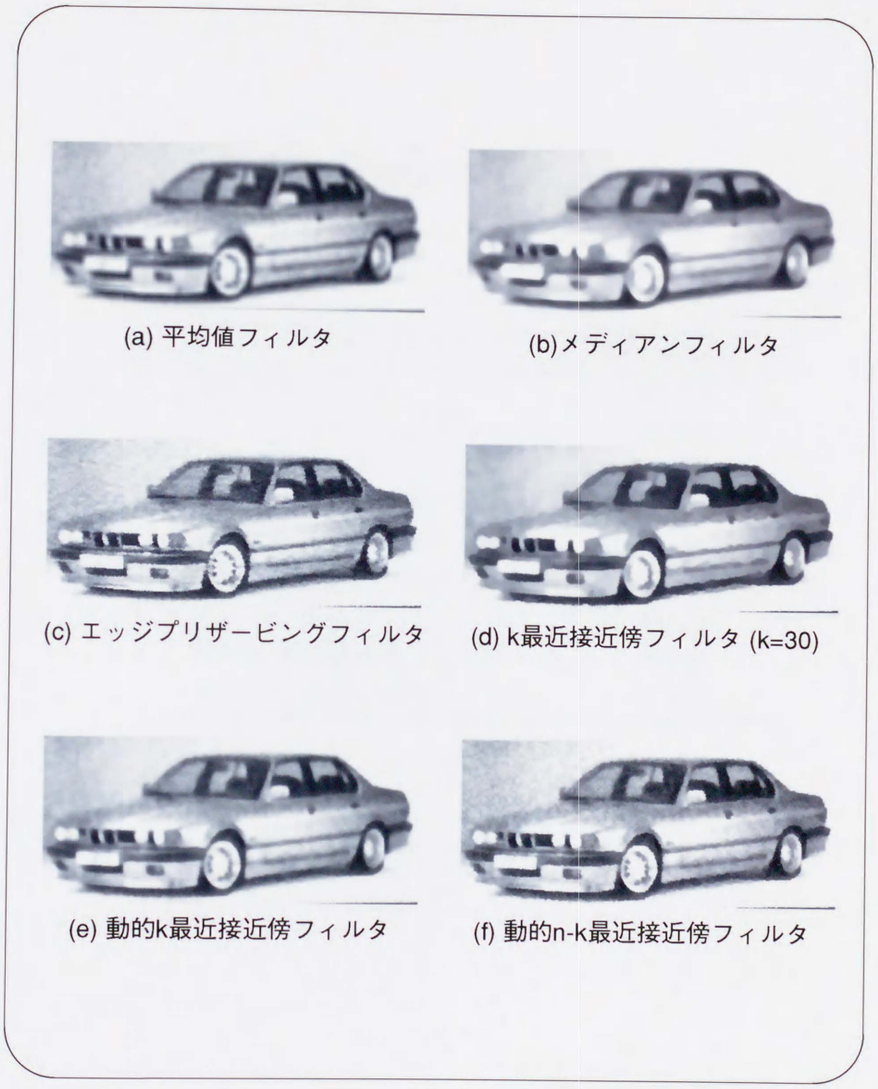


図 2.12: 平滑化処理の比較

表 2.2: 2 次元信号の雑音除去結果（その 1）

フィルタ	平均二乗誤差
雑音画像	207.512
エッジプリザービングフィルタ	158.248
平均値フィルタ	152.086
k 最近接近傍フィルタ（$k = 25$）	115.228
k 最近接近傍フィルタ（$k = 30$）	109.374
k 最近接近傍フィルタ（$k = 35$）	125.496
メジアンフィルタ	113.542
動的 k 最近接近傍フィルタ	88.523
動的 k–n 最近接近傍フィルタ	75.222

表 2.3: 2 次元信号の雑音除去結果（その 2）

フィルタ	平均二乗誤差	順位合計
雑音画像	251.206	335
エッジプリザービングフィルタ	72.982	279
平均値フィルタ	49.804	178
k 最近接近傍フィルタ（$k = 30$）	43.903	178
メジアンフィルタ	48.201	127
動的 k 最近接近傍フィルタ	56.281	216
動的 k–n 最近接近傍フィルタ	40.749	71

に最も効果的であった。この結果は第 2.4.3節で求めた動的 n–k 最近接近傍フィルタのパラメータ調節が他の画像に対しても有効に利用できることを示している。

2.5 本章のまとめ

k 最近接近傍法においてクリークの要素数を変化させる方法を提案し、そのスケジューリング方法について検討した。そして、本手法によるスムージング処理が類似の手法より有効であることを述べ、実際の画像処理に用いたとき、この方法が有効であることを確認した。

本手法によるスムージングは、メディアンフィルタなど他の一般的な手法と比較して、その簡便さが劣ることなく、雑音除去に関する有効性に於いては、高い効果が期待できる。そのため、画像認識等の前処理として利用することによって、高速かつ高性能なスムージングフィルタとして利用することが可能である。

近接近傍法による引き込み作用は、事前知識によらないセグメンテーションを可能にする。これは、事前知識を必要としないクラスタリングのひとつの方法であるといえる。本論文では、輝度値のスムージングに関する効果しか述べていないが、エッジのリラクゼーションのような問題にも応用が容易である。そのような問題への応用によって本手法はさらに利用価値を高めることができるだろう。k 最近接近傍法は、認識などにも用いられる手法であるので、階層構造モデルに応用することによって事前知識に依存する度合いが少ない認識システムが可能になる。

参考文献

[1] J.S.Weszka, "A survey of threshold selection techniques", CGCP, 7–2, pp.259–265 (1978).

[2] M.Nagao and T.matsuyama, "Edge preserving smoothing", CGIP, 9–4, pp.394–407 (1979).

[3] T.S.Hung, G.J.Yang, and G.Y.Tang, "A Fast two-dimensional median filtering algorithm", PRIP'78, pp.128–131 (1978).

[4] 浅野, 横矢, 大久保, 田中, "領域解析に適したフィルターとその比較" 信学技法, PRL 77–13, pp.69–78 (1977).

[5] G.Mastin, "Adaptive Filters for Digital Image Noize Smoothing: An Evalution", Computter Vision, Graphics and imageprocessing, No.31, pp.103–121 (1985)

[6] S.Geman, D.Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", IEEE Trans. Pattern Analysys. Machine Intell., PAM1–6, pp.721–741 (1984).

[7] D.Geman, "Stochastic model for boundary detection", Image and Vision computing, Vol.5, No.2, pp.61–65 (1987).

[8] 本郷, 川人, 乾, 三宅, "エネルギー学習をする局所並列確率アルゴリズムを用いた輪郭線抽出", 信学技法, MBE88–181, pp.151–156 (1989).

[9] 萩原義裕, 小畑秀文, 大森隆司, "動的K近傍法による画像のリラクゼーション", 第22回画像工学コンファレンス論文集, pp.73–75 (1991).

[10] 萩原, 小畑, 大森, "動的 k 近傍法 (IKNOWSK) による画像復元", 神経回路学会第2回全国大会講演論文集, pp.199–200 (1991).

[11] Y.Hagihara, H.Kobatake, T.Omori, "Image restoration with Iterative K-nearest Neighbor Operation With the Scheduled K size(IKNOWSK)", Proceedings of the Third Australian Conference on Neural Networks, pp.178–181 (1992).

[12] 萩原, 小畑, 大森, “動的 n–k 近傍法による画像復元”, 信学論 (D–II), J76–D–II, No.6, pp.1280–1288 (1993).

[13] SPIDER 作業グループ, “SPIDER USER'S MANUAL”, 協同システム開発, (1982).

第 3 章

不確定性を考慮した神経回路モデル

本章のあらまし

神経回路網の中で基本素子となる単純なモデルを提案する。そのモデルは、情報の内容と共にその情報の確かさを盛り込むことができる。

3.1 はじめに

顔全体は見えなくても特徴のある部分を見るだけで、その人の顔を思い出すことは可能である。何も見えなくても声を聞くだけで顔を思い出すことさえできる。その人に対する顔全体の情報量と比較すると、顔の一部や声の情報量は少ないが、想起を行うためにはこの程度の情報量で十分なのである。生物の脳は、極めてわずかな情報からも柔軟に想起を行うことができるからである。すなわち、生物の脳は情報の大規模な欠損に対して安定であるといえる。これは、生物の脳が持つ重要な情報の 1 つである。

しかし、従来の神経回路網モデルは、情報の著しい欠損に対して脆弱である。バックプロパゲーションや、ネオコグニトロンなど、一般的な階層型モデルでは、入力信号は全て得られているという前提で行われているからである。また、連想記憶モデルは、アナログ信号を扱うとき情報に著しい欠損があると正しい結果は想起することができない。一部に欠損があっても誤り訂正能力によって吸収されるため、このような現象は注目されることが少ない。しかし、誤り訂正能力で全てを解決することはできない。赤穂は正則化原理に基づき入力信号のベクトル化を図ることによってこの問題の解決を試みた [1]。本研究は、信号の内容そのものを示すと共に「その信号がどの程度信頼できるか」という指標（確定率）を盛り込むことができるモデルを提案する [2, 3]。このモデルの特徴はベクトル化した出力の全部または一部がフィードバックする再帰型ネットワークである点と入力信号を単純な強調学習によって出力に反映させるという点である [4, 5]。これによって、情報の著しい欠損に脆弱であるという従来モデルの欠点を克服することができる [6, 7]。

3.2 従来の神経回路モデルの問題点

一般的なアナログ連想記憶モデルは、情報が存在するかどうかということと、情報の内容が何を示しているかということを明確に区別していない。アナログ画像を記銘想起する場合について考える。図 3.1に画像の一例を示す。画像は濃淡画像（2 値でない画像）であるとする。「C」のパターンの一部を入力したとき、どのようなパターンが想起されるだろうか。パターンの一部を入力するとき、他の部分（情報が存在しない部分）に代入する値によって想起結果が異なることが予想できる。図 3.1において代入する値が「A」の背景パターンと同じ色のときには結果も「A」となり認識に失敗する可能性がある。代入する値が「C」の背景パターンと近ければ認識は成功することが予想できる。

画像の一部しか与えられていなくても、人間が見れば全体像を思い出すのは容易である。しかし、一般的な神経回路網モデルでは情報が存在しない部分に代入する値によって想起結果が異なる場合がある。従って、条件によって想起結果が異なるという例である。想起を成功させるためには、代入するための最適な値を設定しなければならない。しかし、画像によって最適値は異なる。画像ごとの最適値を設定しなければならないなら、画像の一部か全体を想起できるとはいえない。従来のモデルでは、情報が不足した画像から全体像を想起することはできないのである。

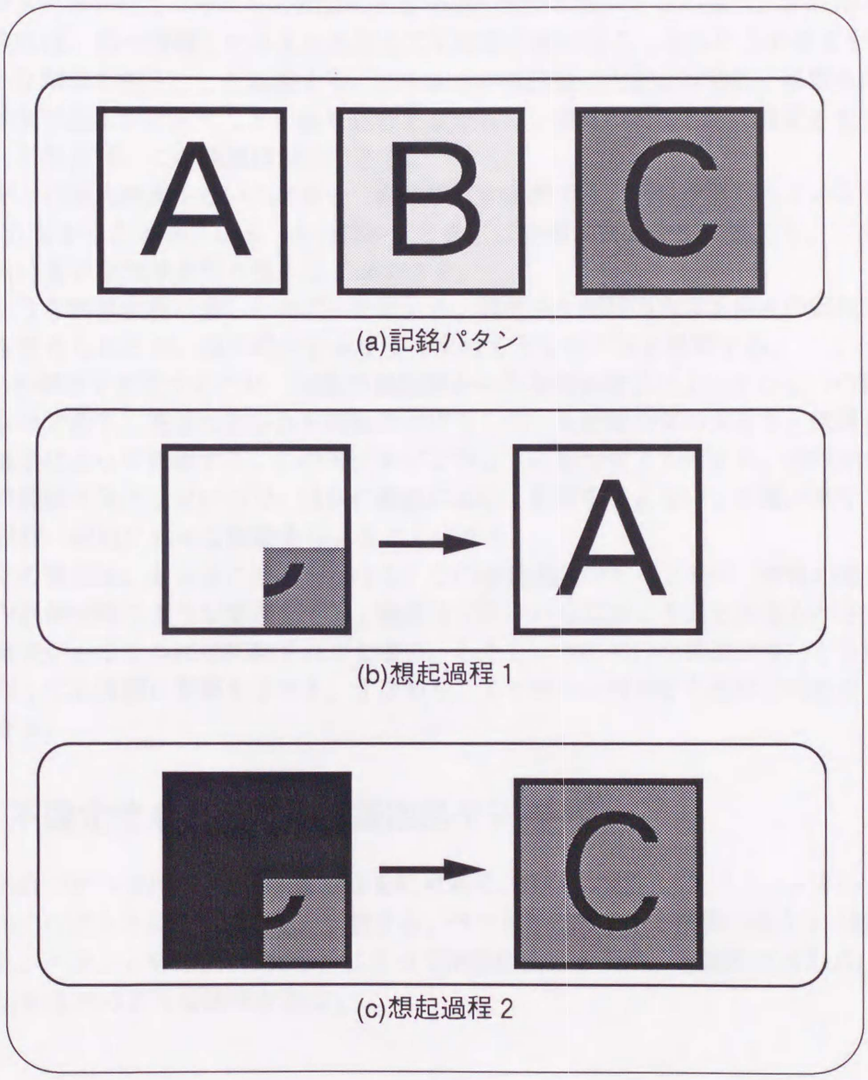


図 3.1: 連想記憶の問題点

3.3 情報の内容と不確定性

このような現象が生じるのは、従来のモデルでは、情報がない部分からの影響を防げないからである。部分的な情報から全体の情報を想起するためには、情報が存在しない部分に何らかのマスクをかけて、それらの部分の影響を他の部分に及ぼさないようにすればよい。これができれば、部分情報しか与えられなくても想起可能になる。さらにこの考えを進めて、あいまいな情報を扱うことを論議する。この場合には情報の内容とは別に、情報の内容がどの程度確実であるかということを盛り込む必要がある。情報の内容とその確実さを 2 次元の情報として扱えば、この問題は解決できる。

情報がどの程度確実かという度合い（確定率）を定義する。情報が得られているときには 1、得られなかったときには 0、あいまいなときには中間的な値を割り当てる。このようにすればあいまいな情報を取り扱うことができる。

そのような情報表現に適したモデルを考える。確実さと情報の内容を別々の細胞で表現する方法も考えられるが、脳の構造をふまえて次のようなモデルを提案する。

2 次元の情報を表現するため、複数の細胞群からなる記憶素子によってひとつの情報を表現するものである。情報の確かさを細胞群が出力している刺激全体の大きさ、情報の内容を出力の偏差によって表現する。このモデルがどのように動作するか考える。情報がないときにはどの細胞も発火しないので、ほかの細胞には全く影響を与えない。情報が来たときにはじめて周囲の細胞に様々な影響を与えることになる。

ラジオの電波は、搬送波に変調をかける。この神経回路モデル上では、情報の確実さが搬送波、内容が変調のような働きをする。搬送波が弱ければ周囲に与える影響が小さく、より強力な電波があるときには抑制されてしまう。しかし、他に強力な電波がないときには、搬送波が弱くても周囲に影響を与える。すなわち、より確かな情報から想起を行わせることが可能になる。

3.4 不確定性を考慮した神経回路モデル

具体的なモデルは図 3.2に示すようなものである。F ユニットと T ユニットという 2 つの細胞の出力ベクトルによって情報を表現する。ベクトルの大きさを情報の確かさ（確定率β_j）と定義し、ベクトルの方向（角度θ）によって情報の内容を表す。各細胞の出力 O_{t_j}、O_{f_j}と確定率β_jには次のような関係がある。

$$\beta_j^2 = O_{t_j}^2 + O_{f_j}^2 \tag{3.1}$$

これによって中間的な値も不確かな信号も扱えることになる。出力情報の正確さに応じて出力信号のパワーを決定するのである。確定率β_jの決定は次のように行う。まず、入力信号 I_t、I_f、重み係数ωより次のような値 M_{t_j}、M_{f_j}を求める。

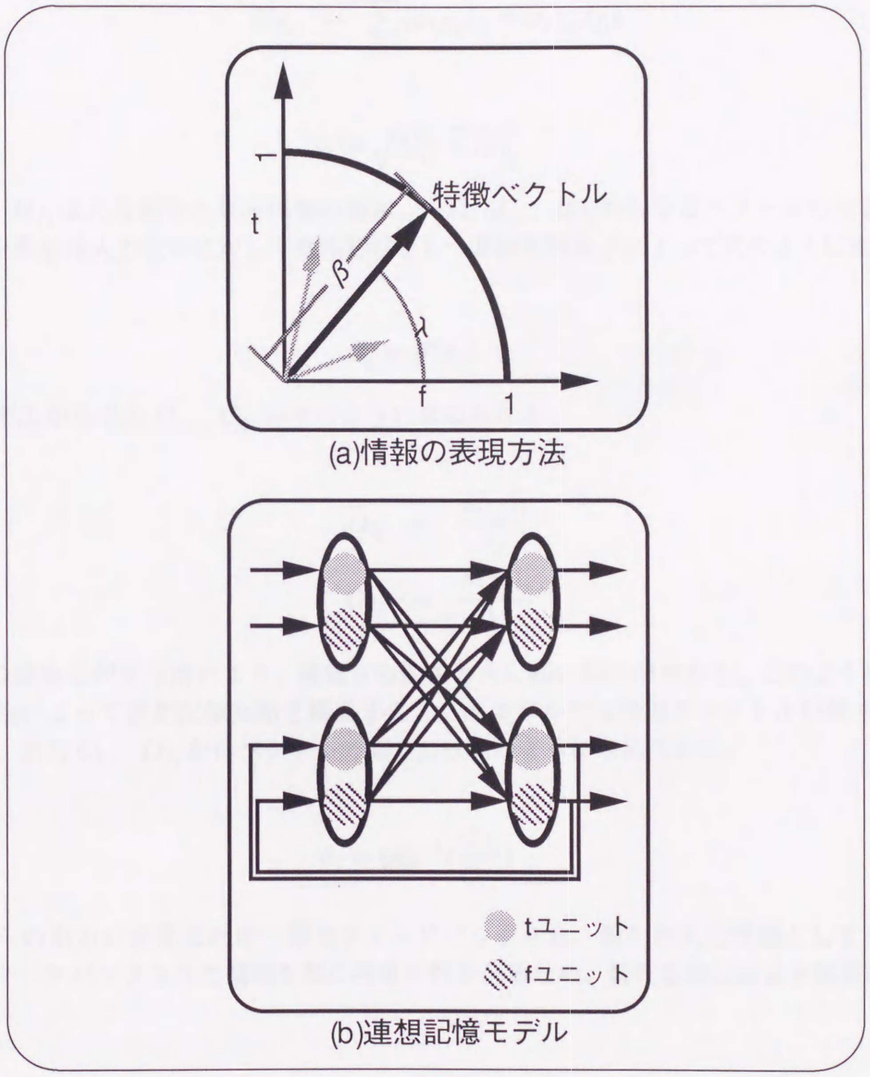


図 3.2: 連想記憶モデル

$$M_{t_j} = \sum_{i=1}^{n}(\omega_{tt_{ij}} I_{t_i} + \omega_{ft_{ij}} I_{f_i}) \tag{3.2}$$

$$M_{f_j} = \sum_{i=1}^{n}(\omega_{tf_{ij}} I_{t_i} + \omega_{ff_{ij}} I_{f_i}) \tag{3.3}$$

$$\lambda_j = \sqrt{M_{t_j}^2 + M_{f_j}^2} \tag{3.4}$$

M_{t_j}、M_{f_j}は入力信号と重み係数の積和、λ_jは M_{t_j}、M_{f_j}からなるベクトルの大きさである。確定率β_jは入力信号に対して飽和特性をもつ非線形関数 Fによって次のように定義する。

$$\beta_j = F(\lambda_j) \tag{3.5}$$

確定率β_jから出力 O_{t_j}、O_{f_j}は次のように求められる。

$$O_{t_j} = \frac{M_{t_j}\beta_j}{\lambda_j} \tag{3.6}$$

$$O_{f_j} = \frac{M_{f_j}\beta_j}{\lambda_j} \tag{3.7}$$

2 つの細胞の相互作用により、確実さの飽和と共に弱い抑制がかかる。このような性質を持つ細胞によって連想記憶回路を構成する。このモデルでは情報をベクトルの傾きθによって表す。出力 O_{t_j}、O_{f_j}からベクトルの傾きθ_jは次のように求められる。

$$\theta_j = \tan^{-1}(\frac{O_{f_j}}{O_{t_j}}) \tag{3.8}$$

これらの出力の全部または一部はフィードバックされ、新たな入力情報として利用される。フィードバックされた情報を基に再考が行われるため、新たな出力はより確実な情報となる。

3.5 連想記憶モデルの学習則

この連想記憶モデルの学習則を考える。ネットワークの出力 O_{t_j}、O_{f_j}から出力ベクトルの角度θ_{O_j}を、教師データ U_{t_j}、U_{f_j}から教師ベクトルの角度θ_{U_j}をそれぞれ求める。重み係数の変化量$\Delta\omega$を次のように決める。

$$d_{t_j} = \theta_{O_j} - \theta_{U_j} \tag{3.9}$$
$$d_{f_j} = \theta_{U_j} - \theta_{O_j} \tag{3.10}$$

$$\Delta\omega_{tt_{ij}} = I_{t_i} d_{t_j} K + C \tag{3.11}$$
$$\Delta\omega_{ft_{ij}} = I_{f_i} d_{t_j} K + C \tag{3.12}$$
$$\Delta\omega_{tf_{ij}} = I_{t_i} d_{f_j} K + C \tag{3.13}$$
$$\Delta\omega_{ff_{ij}} = I_{f_i} d_{f_j} K + C \tag{3.14}$$

このとき、ネットワークの一部に人為的なマスクをかけることによって、学習する部分を選択できる。学習済みのネットワークに後から別の問題を学習させ、順序処理を行わせることも可能である。このモデルのの学習過程を図 3.3に示す。初期状態では小さなベクトルからなるでたらめな出力だが、学習を繰り返すことによって理想状態に近づく。それにともなって出力ベクトルの大きさも大きくなっていく。

3.6 計算機シミュレーション

このモデルに基づき、計算機上でシミュレーションを行った。6×6 ドットの濃淡画像（3 値）を三枚用意して学習させた後、画像の一部から全体像を想起させてみた。非線形関数にはシグモイド関数を用いた。

実験に用いたパターンと想起結果を図 3.4に示す。画像が大きく欠損しており、全体の 4 分の 1 程度しか与えられなかった場合にも、想起結果のベクトルの向きは正しい。これは、情報の欠損に対して安定に想起できたことを示している。この結果によって、提案したモデルが情報の欠損に対して有効であることがわかる。

3.7 本章のまとめ

ここでは、従来の神経回路網モデルが持つ欠点を示し、不確定性を考慮した神経回路モデルを提案した。従来のモデルは情報の著しい欠損に対して脆弱であるという弱点があった。不確定性を考慮した神経回路網モデルは、この弱点を克服することができる。さらに、計算機シミュレーションによってこのモデルの有効性を確認した。

このモデルは、2 つ以上の細胞によって 1 つの概念を表し、その概念がどの程度確実かという情報を盛り込むことができる。従って、情報そのものの内容と、その情報の確実さ（確定率）を伝達することが可能になる。シミュレーションによって、アナログ情報でも情報欠損の多寡に関わらず想起可能であることを確認した。これによって、脳の機能のうち、欠損の激しい情報にも耐え得るという特徴を工学的に実現することができた。

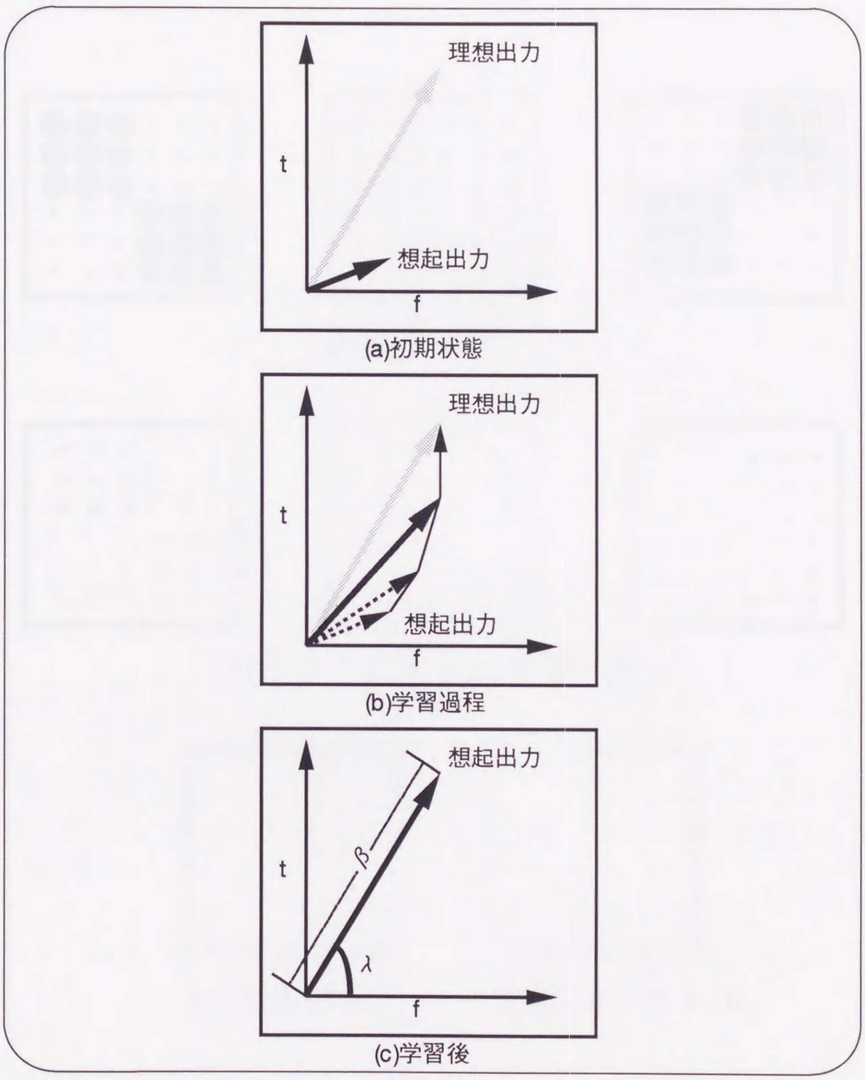


図 3.3: 学習の過程

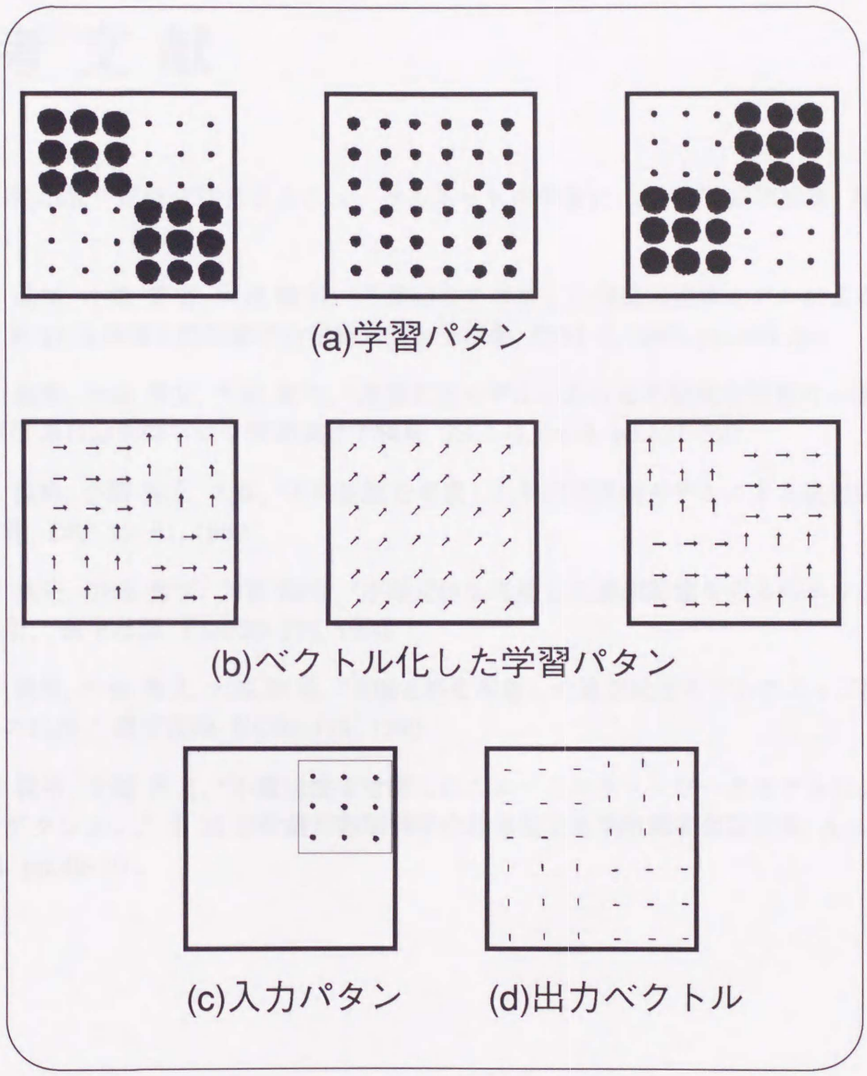


図 3.4: 連想記憶のシミュレーション

参考文献

[1] 赤穂昭太郎, "正則化法によるニューラルネットの学習について", 信学技報, NC91-76 (1991).

[2] 萩原 義裕, 小畑 秀文, 大森 隆司, "不確定性を考慮した神経回路網モデルによる連想記憶," 第 29 回計測自動制御学会学術講演会予稿集, JS24-2, 1990, pp.263-264.

[3] 萩原 義裕, 小畑 秀文, 大森 隆司, "連想記憶モデルにおける不確実な情報の一表現," 第 28 回計測自動制御学会学術講演会予稿集, JS51-5, 1989, pp.521-522.

[4] 萩原 義裕, 小畑 秀文, 大森, "不確定性を考慮した神経回路網モデルによる連想記憶," 信学技報, DSP 90-61, 1990.

[5] 萩原 義裕, 小畑 秀文, 大森 隆司, "不確定性を考慮した連想記憶モデルのエッジ抽出への応用," 信学技報, PRU90-131, 1990.

[6] 萩原 義裕, 小畑 秀文, 大森 隆司, "不確定性を考慮した連想記憶モデルのエッジ抽出・補正への応用," 信学技報, NC90-114, 1990.

[7] 萩原 義裕, 小畑 秀文, "不確定性を考慮したニューラルネットワークモデルによるノイズリダクション," 第 25 回計測自動制御学会北海道支部学術講演会論文集, A-16, No.1, 1993, pp.49-50 .

第 4 章

多層パーセプトロンによる確率密度関数の自己学習

本章のあらまし

ニューラルネットワークにおけるオーバーラーニングは、関数近似を失敗する原因となる。これはニューラルネットワークの汎化能力の限界によるものである。この問題を解決するためには、学習サンプルの確率密度関数をニューラルネットワークに学習させ、学習サンプルが無い領域に属する入力が与えられた場合には「解答不可能」という結果を出力する方法が有効とされている。本研究では、学習サンプルとランダムな値を交互に入力することによって、学習サンプルの確率密度を関数の出力に反映する方法を提案する。さらに、シミュレーションによって、提案した方法が多層パーセプトロン (MLP) で正しく動作したことを報告する。本手法を用いたとき、S/N 比、出力のノルム、学習サンプルの確率密度はそれぞれ密接に関係する。

4.1 はじめに

ニューラルネットワークを用いて学習サンプルから出力関数を推定する場合、学習回数を増やしても適切な応答を実現できないばかりか、かえって悪い結果を出してしまうことがある。この現象はオーバーラーニングと呼ばれ、ニューラルネットワークの問題点のひとつである。オーバーラーニングが生じるのは、入力空間上で学習サンプルが十分に与えられていない部分である。一般の多層パーセプトロン (MLP) では、入力空間上で学習サンプルが十分与えられた部分に対しては適切な応答を実現できる。反面、学習サンプルが与えられていない部分では出力は不定となる。この出力は、中間細胞数、学習回数、重み係数の初期値に依存し、試行ごとに異なる結果となる。中間細胞数が過剰なとき、学習を繰り返すに従って学習サンプルが十分与えられていない空間の出力応答が歪むことが、オーバーラーニングの原因である。したがって、学習サンプルが与えられない部分を近似する方法が課題となる。

汎化能力の向上を目指した学習法は数多く提案されている。小川らは、汎化概念を明確に定義し、過学習が起きる条件やそれを防ぐ方法を提案した [1]。また、川人らは、中間細胞数を検討することにより出力応答の歪みを生じさせないというアプローチで汎化能力を検討した [2]。これらの方法は、学習対象の性質が既知であることが必要となる。

一方、学習サンプルに揺らぎ（雑音など）を加えることにより学習サンプルが無い部分に疑似データを与える方法や、入力の変動に対し出力の変動を最小化する学習法により疑似サンプルを与えずに同様な効果を得る方法も提案されている [3, 4, 5, 6, 7]。これらの方法は、学習サンプル近傍の空間に対して適切な応答を保証するが、入力空間上の全ての部分に適切な応答を与えることは困難である。例えば、入力空間上の全ての部分に適切な応答を与えるような疑似サンプルを与えるためには学習サンプルの揺らぎを非常に大きくする必要があり、かえって能力が低下することが予想できる。

この問題を解決するためには、学習サンプルの確率密度関数をニューラルネットワークに学習させ、学習サンプルが無い領域に属する入力が与えられた場合には「解答不可能」という結果を出力する方法が有効とされている。この方法を用いれば、出力関数はデータの分布に依存するため、学習サンプルが与えられない部分ででたらめな出力を出すことはない。本章では、学習サンプルが存在しない領域に属するような入力が与えられた場合には、「解答不能」を出力する MLP の新しい学習法について述べる [8]。さらに、本手法により学習された MLP の出力は、入力の確率密度を与えるものであることを示す。MLP に確率密度を学習させることができれば、学習サンプルが存在しない領域に属する入力に対して、でたらめな応答が生じるのを防ぐことができる。

入力の確率密度関数を学習するニューラルネットワークはすでにいくつかの方法が提案されている。1 つは別途計算した確率密度をニューラルネットワークに学習させる方法である。中川らは、ベクトル量子化 (VQ) を用いて確率密度関数を計算し、MLP の教師データとして用いる方法を提案した [10]。この方法は、シンプルなネットワークを用いるため、計算量や記憶量が効率的かつ実用的であるが、確率密度を事前に計算し教師データとして与えなければならない。もう 1 つは、確率密度関数を自己学習するするニューラルネットワークモデ

ルを用いるものである [11, 12, 13]。福水らは、4 層のネットワークの第 1 中間層と第 2 中間層の間に抑制細胞を入れることによって確率競合を起こさせ、入出力の信頼度を判定できる Probability Competition Neural Network(PCNN) を提案した。このモデルは、ネットワークが複雑で、学習・応答効率や学習能力が問題となるが、確率密度関数を自己学習するため、確率密度を別途計算して教師データとして与える必要はなく、学習サンプルが与えられない領域に対して適切な応答を得ることができる。

これらの方法に比べ、本章で提案する方法は、極めて簡単な方法で実現できるところにその特徴がある。すなわち、学習サンプルとランダムデータを交互に与えることにより、従来の BP を用いて学習サンプルの確率密度を出力に反映することができる。本章では、確率密度の推定に注目した場合に本手法の評価をヒストグラム法に対して行った。その結果についても述べる。

4.2 学習サンプルの確率密度関数を推定できる MLP

4.2.1 学習サンプルの確率密度関数

学習サンプルの確率密度に出力の正確さが依存することを利用し、学習サンプルが疎である場合には「解答不可能」という出力を得ることを考える。

MLP の入力空間が k 個の微小領域に分割できるとき、各領域に対して全て教師データが与えられれば正常な学習が行われるが、教師データを与えられない領域では学習を行うことができない。各領域は微小であり、入力ベクトル $X_j (j = 1, 2, \cdots, k)$ で表せるとする。このとき、ある入力 X_j に対して教師データ T_j が与えられれば、その入力に対する出力は決定できるが、教師データを与えられない場合には出力は定義できず不定となる。

教師データが与えられていない入力に対して、適当な評価関数を設定できれば、その関数は任意の入力に対して有効である。入力 X_j に対して正しい教師データ T_j が確率的に与えられるとして、任意の入力に対して有効な関数を検討する。入力 X_j に対して正しい教師データ T_j が与えられる確率 $P_{(T_j=\mathrm{True}|X_j)}$ は、教師データが与えられたとき 1、教師データが与えられないとき 0 となる。評価基準として、次のような関数を用いる。

$$G = \sum_{j=1}^{k} \left(\frac{1}{2} \sum_{i=1}^{h} \left(P_{(T_j=\mathrm{True}|X_j)} \cdot t_{ji} - z_{ji} \right)^2 \right) \tag{4.1}$$

教師データ T のノルムは 1 になるように設定する。T の次元は h、T_j の i 番目 $(i = 1, 2, \cdots, h)$ の要素を t_{ij}、それに対する MLP の出力を z_{ij} とする。

すべての入力に対して教師データが与えられているとき、この関数は従来の MLP で用いられる評価関数と同じである。一方、教師データが得られないような入力が与えられたときには、$P_{(T_j=\mathrm{True}|X_j)} \approx 0$ であるから、出力のノルムが 0 となるように学習することになる。すなわち、出力のノルムが 0 のときは「解答不可能」であることが明示される。

この方法は、教師データが与えられた領域とそうでない領域を分割することを可能とするが、これを実現するためには、教師データが与えられているかを全ての入力空間について調べる必要がある。この処理は多次元ヒストグラムを算出し、確率密度を教師データとして与えたときとほぼ等価である。学習サンプルの X_j における確率密度は、学習サンプル群からランダムに 1 つ取りだしたとき、それが X_j である確率である。この値は、教師データが与えられない領域では 0 であり、学習サンプルが密な領域では大きな値を取るため、出力がどの程度信頼できるかという指標となる。

しかし、入力次元が多い場合、多次元ヒストグラムの算出は困難である。そこで、この処理を行わずに出力に学習サンプルの確率密度を反映させる方法を考える。入力空間に対し一様分布のランダムデータを学習サンプルと共に学習させる。ランダムデータ入力に対応する教師データのノルムは 0 とする。ランダムデータの入力空間上の確率密度は定数である。ここで、ランダムデータの X_j における確率密度と、学習サンプルの X_j における確率密度の比を $P_{(T_j=\mathrm{True}|X_j)}$ として再定義する。ランダムデータの確率密度は定数であるため、出力のノルムは学習サンプルの確率密度のみに依存する値となる。したがって、この式は出力に学習サンプルの確率密度を反映させるものとなる。

4.2.2 関数の学習法

この評価関数に基づいて学習を行う実践的な方法として次のような手法を提案する。ランダムデータ N と学習サンプル S を交互に適当な割合で与え、学習サンプルが与えられたときには $P_{(T_j=\mathrm{True}|X_j)}=1$、そうでないとき $P_{(T_j=\mathrm{True}|X_j)}=0$ として学習させる。ランダムデータが入力されたときには評価関数として次のような式を用いる。

$$G=\sum_{j=1}^{k}\left(\frac{1}{2}\sum_{i=1}^{h}(-z_{ji})^2\right) \tag{4.2}$$

これは出力のノルムが 0 になるように学習させることを示している。また、学習サンプルが与えられたときには従来の評価関数で学習させる。

$$G=\sum_{j=1}^{k}\left(\frac{1}{2}\sum_{i=1}^{h}(t_{ji}-z_{ji})^2\right) \tag{4.3}$$

入力値 X_j として、学習サンプルとランダムデータとを混合して与えることを考える。すなわち、学習サンプル群の中から適当に 1 つ取り出して (4.2) 式を用いて学習させる一方、ランダムデータを作り (4.3) 式を用いて学習させる。この 2 種類の学習を適当な割合で行う。入力 X は、学習サンプル群から選択される場合とランダムデータ群から選択される場合がある。ある入力 X_j が学習サンプル群から選択されたとき、それを学習サンプル入力 S_j、ランダムデータから選択されたときランダムデータ入力 N_j と再定義する。(値自体は $X_j=S_j=N_j$

で、全て等しい。) ある入力値が与えられたとき、それが学習サンプル群から選択されたものである確率を $P_{(S_j|X_j)}$とする。また、その入力値がランダムデータ群から選択されたものである確率を $P_{(N_j|X_j)}$とする。

X_jについて学習サンプルとランダムデータとを混合して学習させた場合、出力のノルムは学習時に用いた各教師データのノルムの平均となると考えられる。(4.2) 式では出力のノルムが 1、(4.3) 式では 0 なので、この方法を用いて学習を行えば (4.2) 式が適用される確率 $P_{(S_j|X_j)}$がそのまま出力のノルムとなる。すなわち、出力のノルムは、ランダムデータ N_j と学習サンプル S_jが与えられる回数の比に依存するのである。

なお、ランダムデータ群から 1 つのデータを選択したとき、それが N_jである確率 $P_{(N_j|N)}$ は、N_jに依らない定数である。そのため、$P_{(S_j|X_j)}$は入力の確率密度 $P_{(N_j|N)}$のみによって決定される (証明は後述)。したがって、この方法によって学習させることは、(4.1) 式を用いて学習を行うことと等価である。

4.2.3 出力と学習サンプルの確率密度関数との関係に対する検討

一様乱数 Nと学習サンプル Sを適当な割合で与え、前述のような学習を行った場合、出力のノルムは $P_{(S_j|X_j)}$となる。ここでは、この出力のノルムが学習サンプルの確率密度関数を与える重要な意味を持つことを示そう。MLP にランダムデータを与える確率を $P_{(N)}$、学習サンプルを与える確率を $P_{(S)}$とする。データとして与えられるのはランダムデータか学習サンプルのいずれかであるので両者が与えられる確率の和は 1 となる。

$$P_{(N)} + P_{(S)} = 1 \tag{4.4}$$

入力細胞数が n 個で、入力 Xのノルムは 0 から 1 までの値をとるとする。いま、データ X_jの近傍領域として、$X_j \pm \Delta X$なる微小な超立方体を考え、その中にランダムデータ N_j が入る確率を $P_{(Nj|N)}$とすると、$P_{(Nj|N)}$はこの超立方の体積 $(2\Delta X_j)^n$と入力がとり得る超立方の体積 1^nより求められる。

$$P_{(N_j|N)} = \frac{(2\Delta X)^n}{1^n} \tag{4.5}$$

ただし、ΔXは X_jによらない一定の範囲を取るとする。

N_jはランダムデータ群から選ばれるので $P_{(N_j)}$は $P_{(N)}$に従属である。したがって $P_{(N_j)}$は次のように求められる。

$$\begin{aligned} P_{(N_j)} &= P_{(N_j|N)} \cdot P_{(N)} \\ &= (2\Delta X)^n \cdot P_{(N)} \end{aligned} \tag{4.6}$$

これにより、$P_{(N_j)}$は jに依存しない定数であることがわかる。

一方、S_jは学習サンプル群から選ばれるので$P_{(S_j)}$は$P_{(S)}$に従属である。したがって$P_{(S_j)}$は次のようになる。

$$P_{(S_j)} = P_{(S_j|S)} \cdot P_{(S)} \tag{4.7}$$

入力X_jが学習サンプルである確率$P_{(S_j|X_j)}$は次のように表せる。

$$\begin{aligned} P_{(S_j|X_j)} &= \frac{P_{(S_j)}}{P_{(N_j)} + P_{(S_j)}} \\ &= 1 - \frac{P_{(N_j)}}{P_{(N_j)} + P_{(S_j)}} \end{aligned} \tag{4.8}$$

この式を整理すれば、$P_{(S_j|X_j)}$から$P_{(S_j)}$を求めることができる。

$$P_{(S_j)} = \frac{P_{(N_j)}}{1 - P_{(S_j|X_j)}} - P_{(N_j)} \tag{4.9}$$

$P_{(S_j|X_j)}$は、ネットワークの出力のノルムとして得られるので、この値から入力信号の確率密度$P_{(S_j|S)}$を求めることができる。

$$\begin{aligned} P_{(S_j|S)} &= \frac{P_{(S_j)}}{P_{(S)}} \\ &= \frac{\frac{P_{(N_j)}}{1-P_{(S_j|X_j)}} - P_{(N_j)}}{P_{(S)}} \\ &= \left(\frac{1}{1 - P_{(S_j|X_j)}} - 1\right) \cdot \frac{P_{(N_j)}}{P_{(S)}} \\ &= \left(\frac{1}{1 - P_{(S_j|X_j)}} - 1\right) \cdot \frac{(2\Delta X)^n \cdot P_{(N)}}{P_{(S)}} \end{aligned} \tag{4.10}$$

この式のうち、$P_{(S_j|X_j)}$以外は入出力に関係ない定数である。これにより、出力のノルムから入力信号の確率密度$P_{(S_j|S)}$を求められることが示された。

出力のノルムと学習サンプルの確率密度の関係の一例を図 4.1に示す。これは、(4.10) 式における定数項を 0.1 とした理論値である。例えば、$(2\Delta X)^n$を 0.1 としたとき、サンプルデータとランダムデータを 1:1 で与えればこのようになる。図 4.1からわかるように学習サンプルの確率密度は出力のノルムに対して単調増加となる。したがって、出力ベクトルがどの程度確実であるかという尺度として、出力のノルムを直接利用することができる。

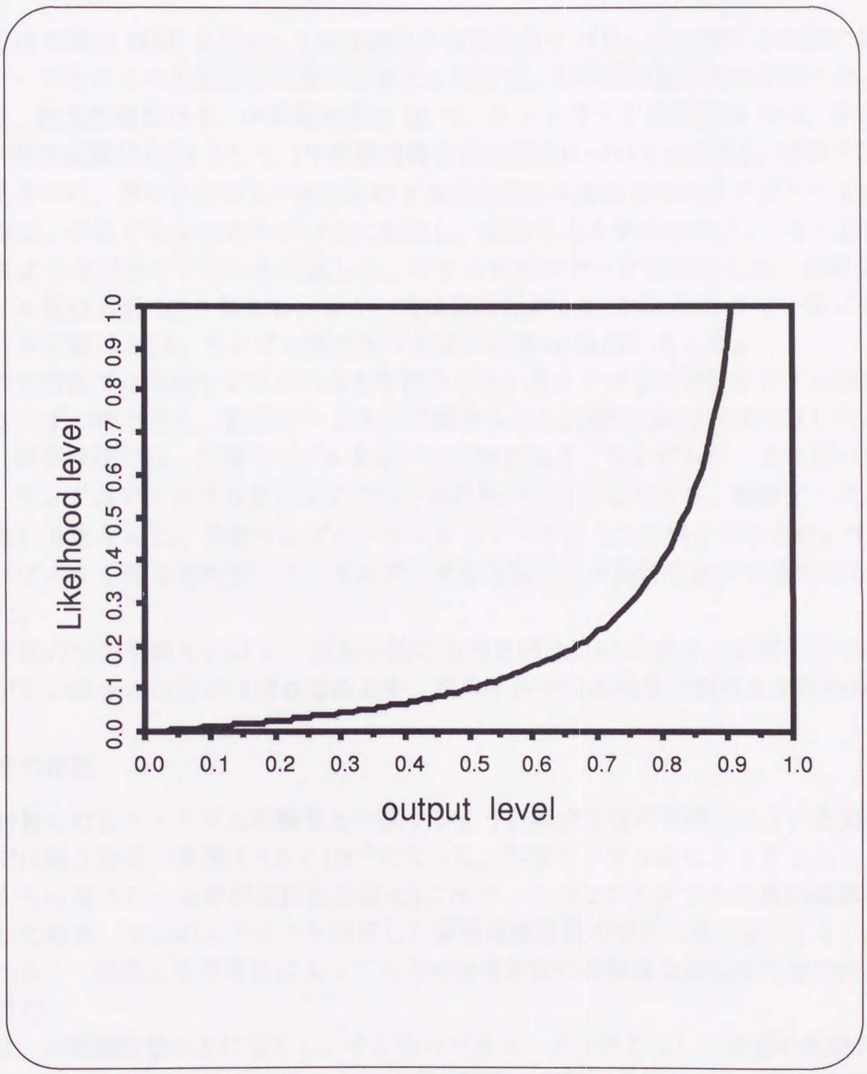


図 4.1: 出力と確率密度の関係

4.3 コンピュータシミュレーション

4.3.1 1 次元信号のカテゴリ分類による性能確認

基本性能の確認

ここでは 3 層の MLP を用い、1 次元信号を複数のカテゴリーに分割する実験に対してランダムデータを与えたときと与えないときとを比較し、本手法の基本性能を調べた。入力細胞数は 1、出力細胞数は 2、中間細胞数は 10 で、ネットワークの温度は 10.0、学習係数は 1.00、学習加速度は 0.50 とした（中間細胞数の設定理由については後述）。学習サンプルとして与えるのは、ガウス分布と一様分布の 1 次元入力からなる 2 つのカテゴリーである。各出力細胞は、それぞれ 1 つのカテゴリに対応し、対応するカテゴリでは 1、それ以外では 0 を与えるような学習サンプルを用意した。ガウス分布のデータは平均 0.25、標準偏差 0.03 でサンプル数は合計 1000 個とし、もう一方は入力値が 0.65 から 0.85 まで一様に分布するデータ（平均値は 0.75、サンプル数はガウス分布同様 1000 個）とした。

従来の学習法では学習サンプルのみを学習させた。各カテゴリの学習サンプル群からランダムに 1 つずつ取り出し、教師データと共に学習させる処理を 50 万回繰り返した。

一方、提案手法では、学習サンプルを用いた学習に加え、ランダムデータを用いた学習を行った。ランダムデータは 0 から 1 までの一様乱数である。このとき、教師データとして全ての細胞に 0 を用いた。学習サンプルとランダムデータを 1:2 の割合で与えた。すなわち、学習サンプルを 1 回与えた後、ランダムデータを 2 回与える処理を従来手法同様 50 万回繰り返した。

提案手法の出力を図 4.2(a) に、従来手法の出力を図 4.2(b) に示す。提案手法では学習サンプルがない部分の出力がほぼ 0 であるが、従来手法では単純な分割のみが行われている。

確率密度の推定

別途計算したヒストグラムの概形と一致するように超立方体の体積 $(2\Delta X)^n$ を調べた。この実験では超立方体の体積は 8.3×10^{-2} となった。学習サンプルのヒストグラムと、出力のノルムから計算された確率密度関数を図 4.3に示す。このヒストグラムを真の確率密度関数と見なした場合、出力のノルムから計算した確率密度関数の平均二乗誤差は、4.1×10^{-3} と非常に小さく、提案した学習法によって入力の確率密度の高精度な推定が可能であることが確認できた。

さらに、中間細胞数を 5 に減らし、その他のパラメータを同じにして同様の実験を行った。このとき、出力のノルムから計算した確率密度関数の平均二乗誤差は、中間細胞数が 10 のときと同様 4.1×10^{-3} となった。入力細胞数 1、中間細胞数 5、出力細胞数 2 のとき、MLP 内の結合数は、24 である。すなわち、平均二乗誤差を 4.1×10^{-3} にするためには、24 個のメモリが必要である。一方、一般的なヒストグラムで、入力を 25 等分したときの平均二乗誤差は、4.6×10^{-3} となった。このとき必要となるメモリの数は 25 個であり、提案手法より

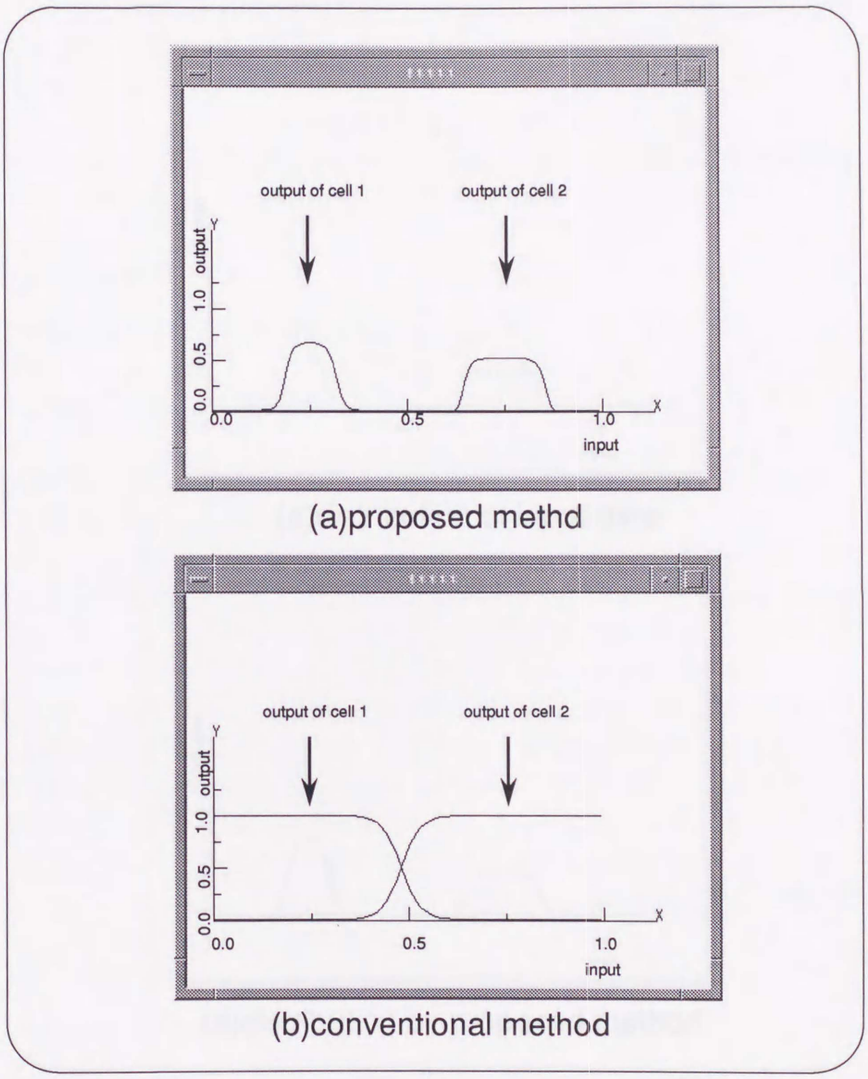
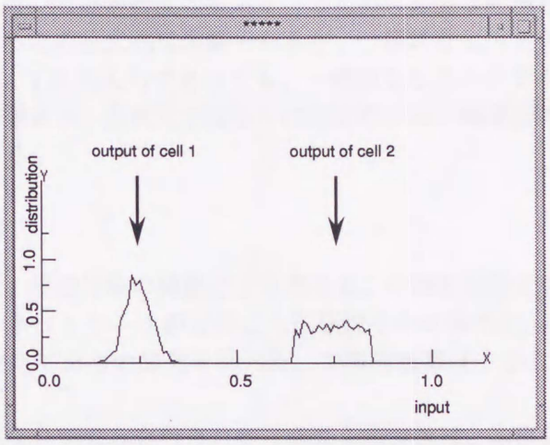
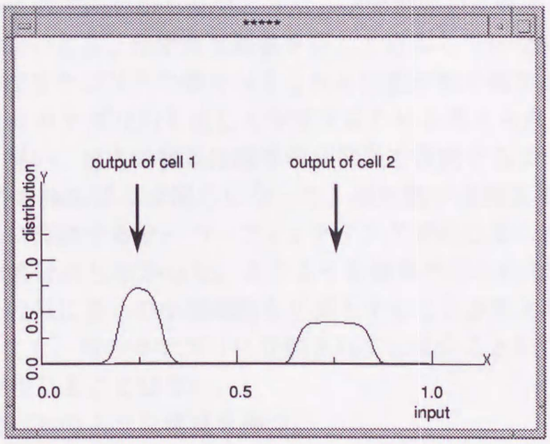


図 4.2: 1 次元信号の分割実験



(a)distribution of input data



(a)distribution by proposed method

図 4.3: 確率密度関数の比較

多い。以上より、提案手法をこの課題に適用したとき、一般的なヒストグラムを用いた場合と同等あるいはやや優れた結果が得られることが明らかになった。

入力が多次元の場合、確率密度を求める為に必要な記憶容量が少なくなるという点で、本手法はヒストグラムを用いた場合より有効性が高い手法となる。一般的なヒストグラムを用いた多次元の確率密度を求めるために必要な記憶容量は、入力の次元数に対し指数的に増加する。一方、本手法を用いた確率密度の推定を行うために必要な記憶容量は、入力の確率密度の性質に強く依存するため定式化は困難であるが、一般的なヒストグラムのように指数的に増加することはない。1 次元入力であっても、一般的なヒストグラムと同等以上の確率密度推定が可能という結果より、多次元ではさらに有効性が高い確率密度関数の推定が可能であるということができる。

中間細胞数による影響

適切な中間細胞数は、学習対象の関数により異なる。中間細胞数が少なすぎるとき、または過剰であるときに、ネットワークがどのように振舞うかを調べた。学習回数を 50 万回とし、中間細胞数のみ変えて出力の変化を調べた。中間細胞数は 1 から 10 まで、1 ステップとした。

従来の学習法で単純な分割のみを行うために必要な細胞数は、本実験で用いたサンプルの場合、1 である。提案手法では複雑な（シグモイド関数で表現しにくい）確率密度関数を近似するために、より多くの中間細胞を必要とする。中間細胞数に対する出力の変化を第 4.4 節に示す。細胞数が少ないとき、確率密度関数を正しく近似していないが、カテゴリ分類には成功している。これはカテゴリ内の値を与えられる回数が他の値を与えられる回数より多いため、カテゴリ外よりカテゴリ内を正しく学習するためと考えられる。

細胞数が多くなるに従い、出力の概形は確率密度関数を反映するようになる。中間細胞数が 4 以上のとき、出力の概形はほぼ同じになった。細胞数が過剰な場合は、学習サンプルの確率密度を必要以上に反映するオーバーフィッティングが生じることが予想されるが、本実験ではそのような障害は生じなかった。このような結果が得られた理由のひとつとして、確率密度関数の正確な近似に多くの中間細胞を必要とすることが考えられる。仮にオーバーフィッティングが生じても、他のカテゴリに分類されてしまうことはないため、従来の学習法のように分類ミスが生じることはない。

以上より、提案手法は次のような性質を持つ。

(1) 中間細胞数が足りないときは従来の学習法と同様に振舞う。

(2) 中間細胞数が適切なら学習サンプルの確率密度関数を反映する。

(3) 確率密度関数の概形を学習させるには多くの細胞を必要としないが、正確に反映するには多くの細胞を必要とする。

(4) 中間細胞数が過剰なとき学習サンプルの分布を正確に反映するため、オーバーフィッティングが起こり得るが、この場合でも他のカテゴリに分類されてしまうことはない。

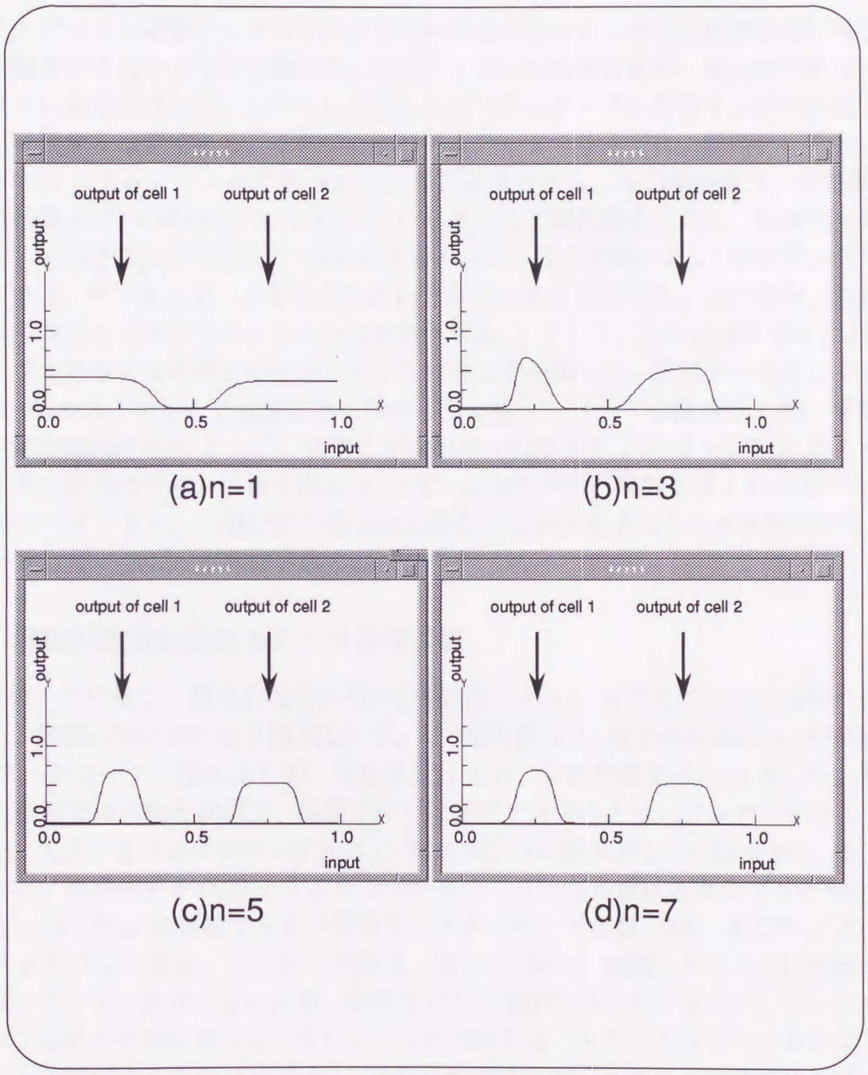


図 4.4: 中間細胞数による出力の変化

したがって、中間細胞数の自由度は、提案手法の方が従来の学習法より大きいということができる。

4.3.2 S/N 比による出力変化

ランダムデータと学習サンプルを与える比率を変化させて、その比率が出力に対して理論通りの影響を与えるかどうかを調べた。学習サンプルの確率密度が一様分布であって出力のカテゴリーによらなければ、入力として与えるランダムデータと学習サンプルの比は出力のノルムと一致するはずである。これは、ランダムデータを与えずに、教師データとして学習サンプルとランダムデータの比を与えたときと等価である。入力細胞数 1、出力細胞数 1、各中間細胞数 10 の 4 層 MLP に 0.00 から 1.00 までの正規乱数を入力し、教師データとして 0 または 1 を特定の比率で与えて 5000 回学習を行い、出力を調べた。(学習サンプルが与えられたとき 1、ランダムデータが与えられたとき 0 であると考える。この場合、教師データとして 1 を与える比率がそのまま出力に反映される。) そして、この比率を様々に変化させ、学習サンプルを与える比率と出力がどのようになるかを調べた。教師データとして 1 を与える比率は 0%から 100%まで 10%ごとに設定した。ネットワークの温度は 1.00、学習係数は 0.04、学習加速度は 0.00 とした。比率と平均出力 (0.00 から 1.00 まで、0.01 ステップで入力したときの出力の平均) の値を図 4.5に示す。出力の平均二乗誤差は 1.1×10^{-2}となった。この結果とグラフより、一様分布の場合は比率と出力が比例することが実験的にも示され、理論と一致していることが確認された。

4.3.3 排他的論理和型のカテゴリ分類実験

2 次元の入力に対し、排他的論理和型の分割実験を行い、本手法の工学的応用への可能性を調べた。実験に用いたのは 3 層 MLP で、入力細胞数は 2、出力細胞数は 2、中間細胞数は 40 で、ネットワークの温度は 1.00、学習係数は 0.05、学習加速度は 0.10 とした。学習方法は 1 次元信号での実験と同様で、提案手法では学習サンプルとランダムデータを 1:2 の割合で与えた。入力に与える学習サンプルは $(0.25, 0.25)$、$(0.25, 0.75)$、$(0.75, 0.25)$、$(0.75, 0.75)$ を平均とし、標準偏差を 0.03 とするガウス分布で、サンプル数は各カテゴリとも合計 1000 個である。図 4.6(a) は実験で与えた学習サンプル (サンプル数 1000) 及びサンプル数 1 千万のヒストグラムである。(これらの図は、図 4.1に従い、強調したもの。) 実験で与えたデータは、サンプル数が少ないため、裾野部分の分布がなめらかでないが、サンプル数が多いものは、なめらかな分布となっている。各出力細胞は、対角にある 2 つのカテゴリにそれぞれ対応し、対応するカテゴリでは 1、それ以外では 0 を与えるような学習サンプルを用意した。

一方、図 4.6(b) は提案手法の出力である。MLP のスムージング効果により、分布の裾野付近でもなめらかな出力となっている。

図 4.6(c) は従来型学習法の出力である。この例では $(0.5, 0.5)$ を入力したとき、$(0.25, 0.25)$ や $(0.75, 0.75)$ を入力したときと同じ値をとなった。学習サンプルが与えられた領域では正

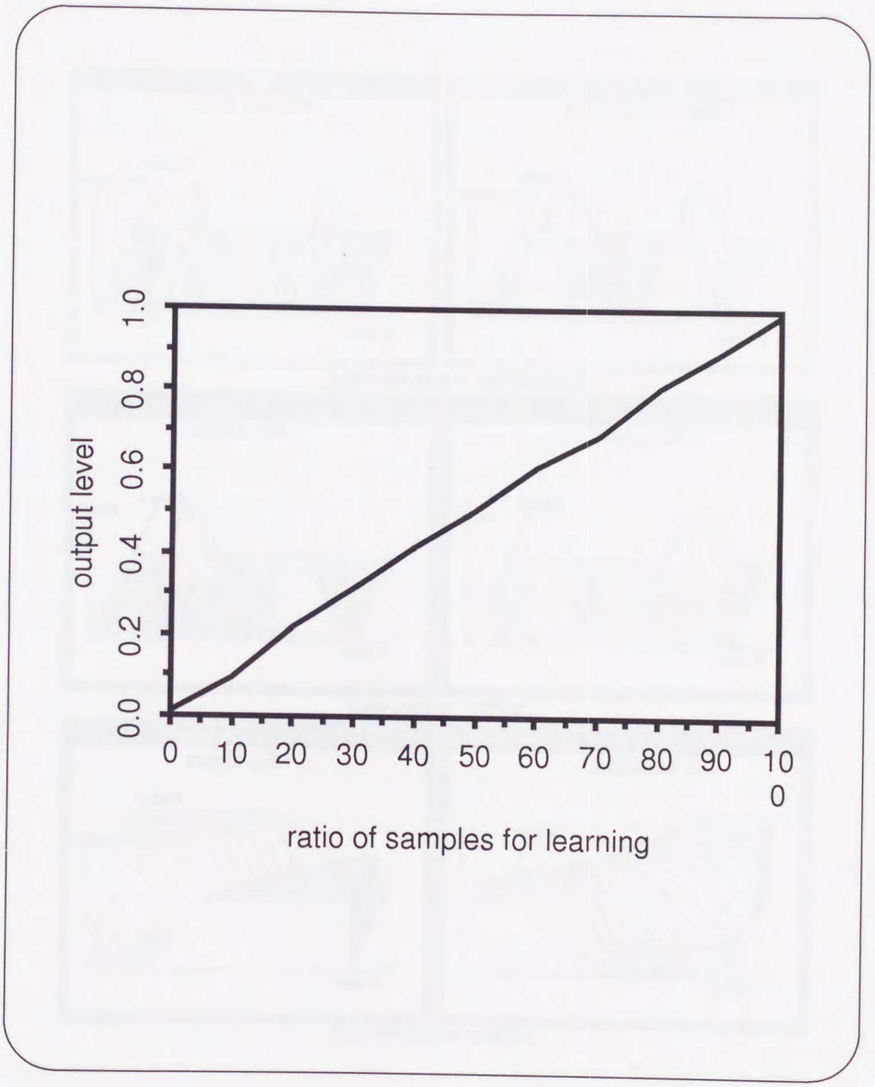


図 4.5: 学習サンプルとランダムデータの比と出力の関係

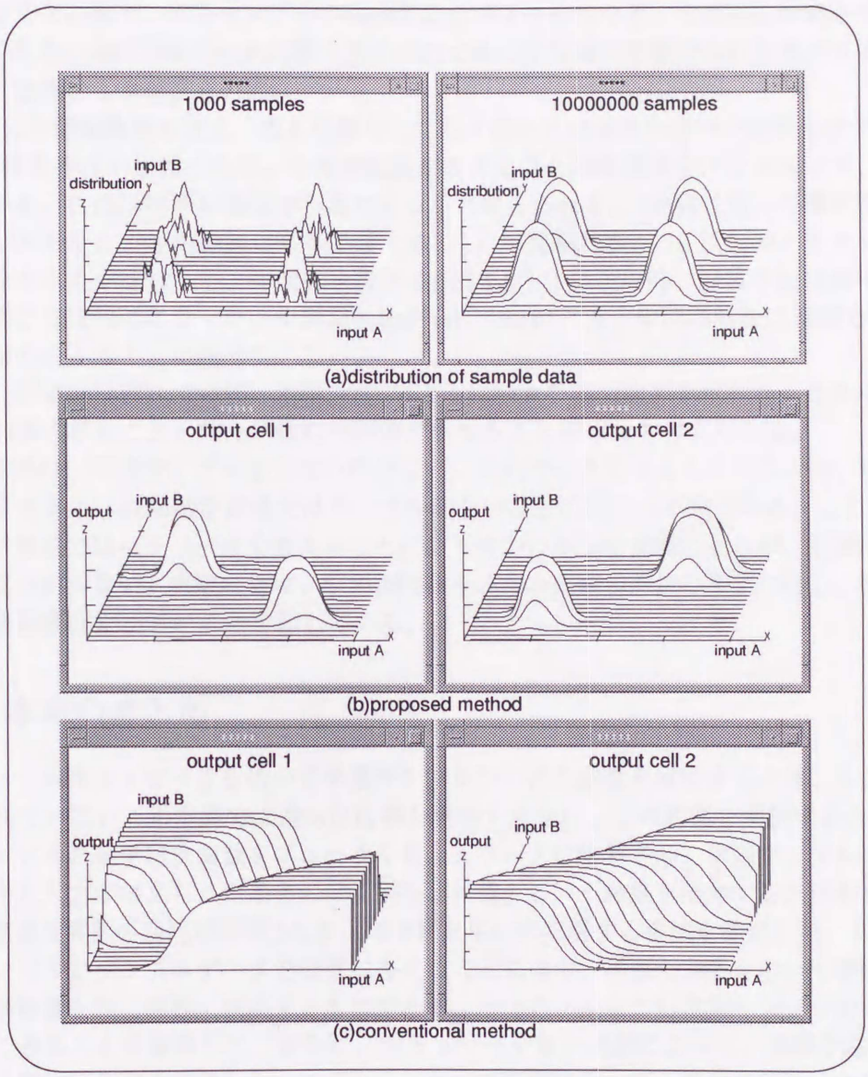


図 4.6: 2 次元入力の分割実験

しく振る舞うが、学習サンプルが与えられていない部分では動作が保証されていないことがわかる。

提案手法の出力は、4 つの領域の部分以外では出力のノルムが 0 となっており、「解答不可能」な領域が学習されていることがわかる。従来の MLP では確率密度関数を教師データとして与えない限り、学習サンプルの確率密度の形（対称性など）を出力に反映させることは困難である。本手法を用いれば多次元入出力であっても確率密度分布の学習が可能であることが、実験により確認された。

図 4.7に中間細胞数を変え、重み係数として各々異なる初期値を用いた結果を示す。(a)～(c) は、提案手法の出力である。中間細胞数が 3 のとき、確率密度を十分反映せず、出力が歪んでいる。これは中間細胞数が不足であったと考えられる。（実験に用いた確率密度分布が円形であるため、中間細胞は多数必要である。）中間細胞数が 10 以上のときサンプルがない部分の出力が 0 になり、中間細胞数が 40(図 4.6(b)) 以上の時、提案手法は確率密度関数の概形とほぼ同じになった。中間細胞数が 60、100 のとき、中間細胞数は過剰であるが、出力曲面の歪みはとくに生じていない。

一方、従来型学習法では図 4.7(d)～(f) のように様々な出力曲線を示した。これは、中間細胞数が多すぎること、重み係数の初期値が異なることが原因と考えられる。

提案手法は、学習サンプルが少ない部分にランダムデータを与えることにより、出力を一意に決定できる。従来型学習法ではサンプルがない部分の出力は不定である。したがって、従来型学習法ではパラメータを変えることにより出力は多様な曲面となるが、提案手法では比較的変化が少ない。実験結果は、中間細胞数及び重み係数初期値の変化に対し、提案手法の出力曲面変化が小さいことを示している。

4.4 本章のまとめ

ニューラルネットワークを用いて学習サンプルから出力関数を推定するとき、サンプルが与えられていない入力空間では適切な応答を実現できない。この問題を解決するためには、学習サンプルの確率密度関数をニューラルネットワークに学習させ、学習サンプルが無い領域に属する入力が与えられた場合には「解答不可能」という結果を出力する方法が有効である。確率密度関数の自己学習を MLP 上で BP を用いて実現する手法を提案した。本手法は、学習サンプルとランダムデータを交互に与えることにより、従来の BP を用いて学習サンプルの確率密度を出力関数に反映するものである。出力のノルムから学習サンプルの確率密度を計算できることを証明した。さらに、シミュレーション実験によって、提案手法が MLP で正しく動作する事を確認した。シミュレーション実験において、提案手法は、中間細胞数が少ないとき従来手法と同様にふるまい、中間細胞が十分なときには確率密度関数を反映することを示した。これにより、提案手法は従来の学習法の上位互換であり、カテゴリ分類等に有効であることが確認できた。

本手法を用いれば、データの与え方に手を加えただけで、入力の確率密度関数を簡単に求めることができる。本手法は MLP を利用しているため、学習速度向上やローカルミニマ

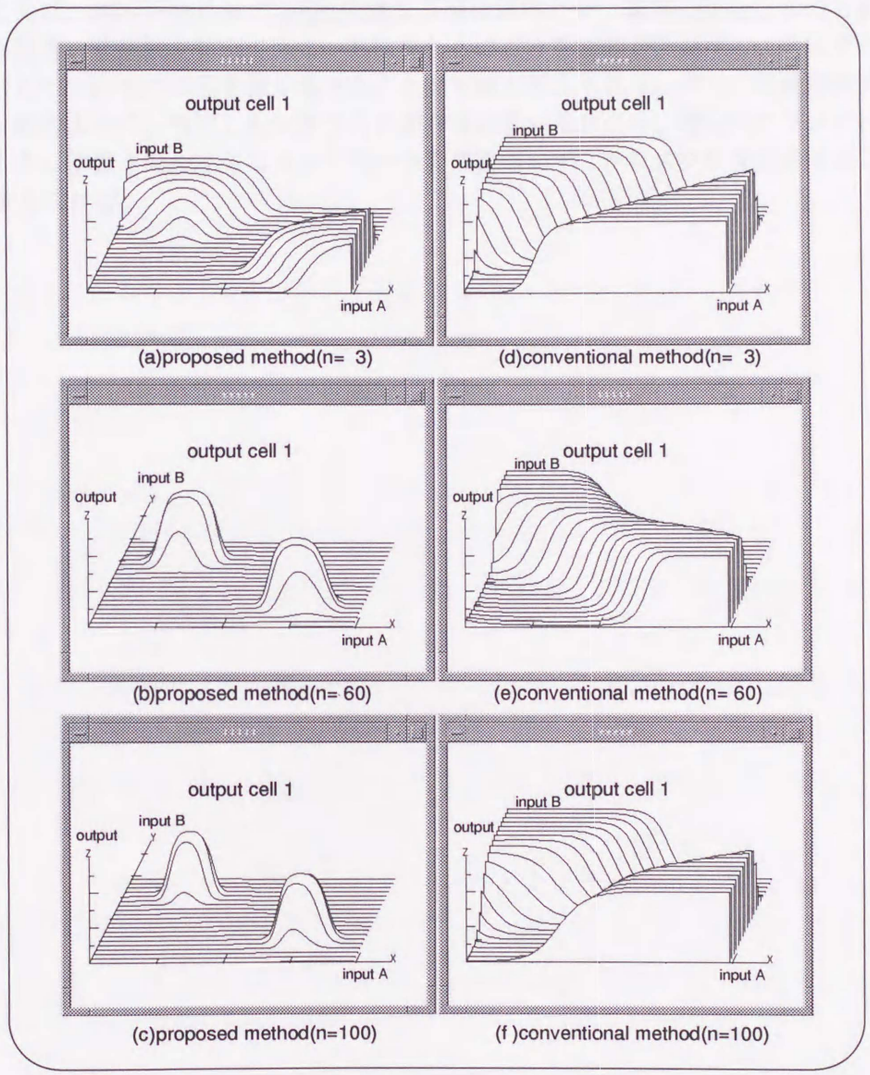


図 4.7: 中間細胞数による出力の変化

ム回避など学習サンプル数に依存しない問題については従来の手法がそのまま適用できる。また、本手法は、あらかじめ確率密度関数を教師データとして用意する必要がない。従来のBP プログラムパッケージがあれば、学習時の入出力にわずかな手を加えるだけで良く、プログラム作成の手間もほとんど必要がない。したがってパターン分類への応用は容易である。たとえば、癌の自動診断で 100%確実な分類は困難だが、確実に安全なものと異常なものを取り除き、計算機で判定できない不確実なものだけを医師が診断することにすれば、安全を維持しつつ医師の負担を減少させることが可能と考えられる。また、衛星画像データのクラス分類のように、判定したいクラスタが非常に多い場合には、特定のクラスタにだけ反応するように学習させた小さなネットワークを多数用いて、クラスタを推定させることもできると考えられる。

参考文献

[1] 小川英光, 山崎一孝, "過学習の理論", 信学論 (D–II), Vol.J76–D–II, No.6, pp.1280–1288 (1993).

[2] 和田安弘, 川人光男, "新しい情報量規準と Cross Validation による汎化能力の推定", 信学論 (D–II), Vol.J74–D–II, No.7, pp.955–965 (1991).

[3] 松岡清利, "誤差逆伝搬法の汎化問題に対する一手法", 信学論 (D–II), Vol.J73–D–II, No.6, pp.897–905 (1990).

[4] 鹿山昌宏, 阿部重夫, "汎化能力向上を目的としたクラスタリング用ニューラルネットの学習方式", 信学論 (D–II), Vol.J76–D–II, No.4, pp.863–872 (1990).

[5] 赤穂昭太郎, "正則化法によるニューラルネットの学習について", 信学技報, Vol.NC91–76 (1991).

[6] 木村義政, "誤差及び出力変動を最小化する学習アルゴリズムとその低品質手書き漢字認識への応用", 信学論 (D–II), Vol.J73–D–II, No.6, pp.840–847 (1990).

[7] 鈴木紳, 河原英紀, "平均局面に基づいた神経回路網の評価基準", 信学論 (D–II), Vol.J75–D–II, No.3, pp.637–645 (1992).

[8] 萩原, 小畑, "多層パーセプトロンによる 確率密度関数の自己学習", 信学論 (D–II), Vol.J78–D–II, No.10, pp.1491–1500, 1995.

[9] 萩原, 小畑, 大森, "不確定性を考慮した連想記憶モデルのエッジ抽出・補正への応用", 信学技報, Vol.NC90–114 (1990).

[10] 中川聖一, 小野義之, "ニューラルネットワークによる確率密度関数・事後確率の推定と母音認識", 信学論 (D–II), Vol.J76–D–II, No.6, pp.1081–1089 (1993).

[11] Specht D., "Probabilistic Neural Networks", Neural Networks,Vol.3, pp.109–118 (1990).

[12] 福水健次, 渡部澄夫, “統計的推論を実現するネットワークとそのパターン認識への応用”, 信学技報, Vol.NC92–36 (1992).

[13] ed. Kosko B., “Neural Networks for Signal Processing”, Prentice–Hall International, pp.199–223 (1992).

第 5 章

確率密度を自己学習する MLP のアプリケーション

本章のあらまし

各要素技術を組み合わせた神経回路網の詳細な構成と画像処理への応用について述べる。確率密度を自己学習する MLP を用い、これと k 近傍フィルタなどの組み合わせによって、複雑な調整無しで実現できる画像処理回路網システムを提案する。このシステムを用いれば記憶容量が小さいネットワークモジュールを用いて、多くのパターンを記憶させることができる。そして、このシステムを手書き文字認識や画像の雑音除去に本手法を用いたとき、従来型 MLP などを用いたときより有効であることを確認する。

5.1 はじめに

多層パーセプトロン (MLP) の学習にはバックプロパゲーションアルゴリズム (BP) が多く用いられており、パターン認識などの分野で有効性が確認されている。しかし、これらの成果の多くは、分類カテゴリーが少ない比較的小規模なネットワークで行われており、画像処理用のシステムなど大規模で実用的なネットワークにそのまま拡張したときの有効性は疑問視されている [1]。手書き文字認識など、分類カテゴリーが多いときは、ネットワークが大きくなるため、学習時にローカルミニマムに陥るなど収束性が悪く、また、膨大な計算量が必要であるという問題がある。ネットワークに有効な能力を持たせるためには学習係数、細胞数、学習回数などのネットワークパラメータを適切に設定しなければならない。これらの設定が不適切なときには、オーバーラーニングやローカルミニマムなどの現象が生じ、認識能力が低下してしまうという問題がある。

この問題を解消するためには、パターン認識システムを学習が行いやすい小規模なネットワークに分割し、それらを統合したネットワークを構築することが有効である。個々の小規模ネットワークは、構造が簡単でローカルミニマムやオーバーラーニングが生じることも少ないため、ネットワーク全体がローカルミニマムに落ち込むのを防ぐことができる。また、小規模なネットワークに分割することにより、ネットワーク全体の結合数を減らすことができるため、学習速度が高速化するという効果もある。このような手段で大規模な神経回路網を実現するモデルには統合型ニューラルネットワーク [2, 3]、選択型ニューラルネットワーク (Comb NET)[4, 5, 6, 7, 8]、Nestor Learning System (NLS) [9]、ファジー推論アーキテクチャに基づく NARA [10] などがある。

統合型ニューラルネットワークは、能力が限定された回路網（サブネット）の出力を複数統合し、さらに大きな回路網（スーパーネット、ハイパーネット）で大統合を行うものである。各回路網には LVQ で構成されており、その回路網が担当する範囲である確からしさの指標として参照ベクトルとのユークリッド距離を用いている。この方法は、分割、単純化された課題を学習させるため学習が容易になるという特徴がある。

Comb NET は、始めに大分類、次に小分類を行った後、出力を統合する。大分類には LVQ、小分類には MLP を用いる。この方法は、各回路網が限られた範囲の判別を担当するので学習の収束性、認識率の点で優れている。

NLS は、複数の特徴を個別の回路網で処理し、出力を統合するものである。各回路網には改良された LVQ を用いる。NLS の長所は、局所的最小解などに陥らなかった解を選択できる点である。複数特徴量がある場合、そのうちひとつが局所的最小値に陥っても他の解法による結果を利用すればよい。複数の特徴量を用い、最も識別しやすい特徴量を選択して利用できる。

これらのネットワークは、大規模化が困難であるという従来の神経回路モデルの弱点を克服したが、ネットワークの構造が複雑である。また、入出力情報の性質が変化したとき、ネットワークの基本的な構成を変えると共に、細分化されたネットワークモジュール同士の結合強度や結合方式を調整し直さなければならないという欠点がある。本章では、確率密度

を自己学習する MLP を用い、これと k 近傍フィルタなどの組み合わせによって、複雑な調整無しで実現できる回路網システムを提案する。このシステムを用いれば記憶容量が小さいネットワークモジュールを用いて、多くのパターンを記憶させることができる。本手法を手書き文字認識と画像の雑音除去シミュレーションに用い、従来型 MLP などを用いたときより有効であることを確認する。

5.2 画像や音声における諸問題

画像情報や音声情報を学習させるには、学習パターン数や入力次元数に応じた回路網を構成しなければならない。これらの情報は、学習パターン数や入力次元数が多い上、学習対象となるパターンを変化させる必要に迫られることも頻繁に生じる。従来型の MLP はローカルミニマムやオーバーラーニングなどの問題が生じやすいため、このような情報を記憶させるには不向きである。統合型ニューラルネット、Comb NET、NLS などの回路網は、学習パターン数や入力次元数の増加に対しては頑強であるが、学習対象の変動に伴って回路網の構成を再検討しなければならない。画像情報や音声情報などマルコフモデルは複数の情報が重なり合っているという性質がある。この性質に基づいて、ネットワークをモジュールに分割し、各モジュールの出力を統合化する回路網を構成する必要がある。

画像情報や音声情報は、複数の情報が重なり合って構成されていると考えることができる。複数の情報が重なり合った信号から特定の情報を取り出して認識するためには、不必要な情報を雑音とみなして取り除く必要がある。マルコフモデルでは、必要な情報を周囲の情報から確率的に取り出すことを可能にしている。これらのモデルでは各特徴量の同時生起確率をあらかじめ決めておく必要がある。しかし、複数の特徴量が重なり合った場合には相互に関係する確率が変動する。

2 次元のマルコフモデルで利用される、平面、エッジ、交点などの特徴量を考える。2 次元画像において複数の平面が存在するとき、平面と平面の境界に境界線が発生する。また、複数の境界線が存在するときには交点が発生する。境界線の形は直線とは限らず、様々な形となる。また、交点に関しても同様である。境界線が交差した場合など、複数の特徴量が重なり合ったときには、同時生起確率は変動するのである。従って、あらかじめ全ての可能性を考えて確率を決定しておくことは困難であるが、複数の近傍領域から注目点の値を推定することによって、この問題を最小限に抑えることにする。

5.3 複数パターンからの信号抽出

観測信号 Xから出力信号 Yを取り出すことについて考える。

観測される信号は、図 5.1のように様々なパターンが重なり合っていると考えることができる。この中から必要なパターンだけを取り出すことを考える。基になる出力パターンは p 個存在し、$Y_1, Y_2, \cdots, Y_p$であるとする。観測される Xは、パターンが重なり合っており、次のような関係があるとする。

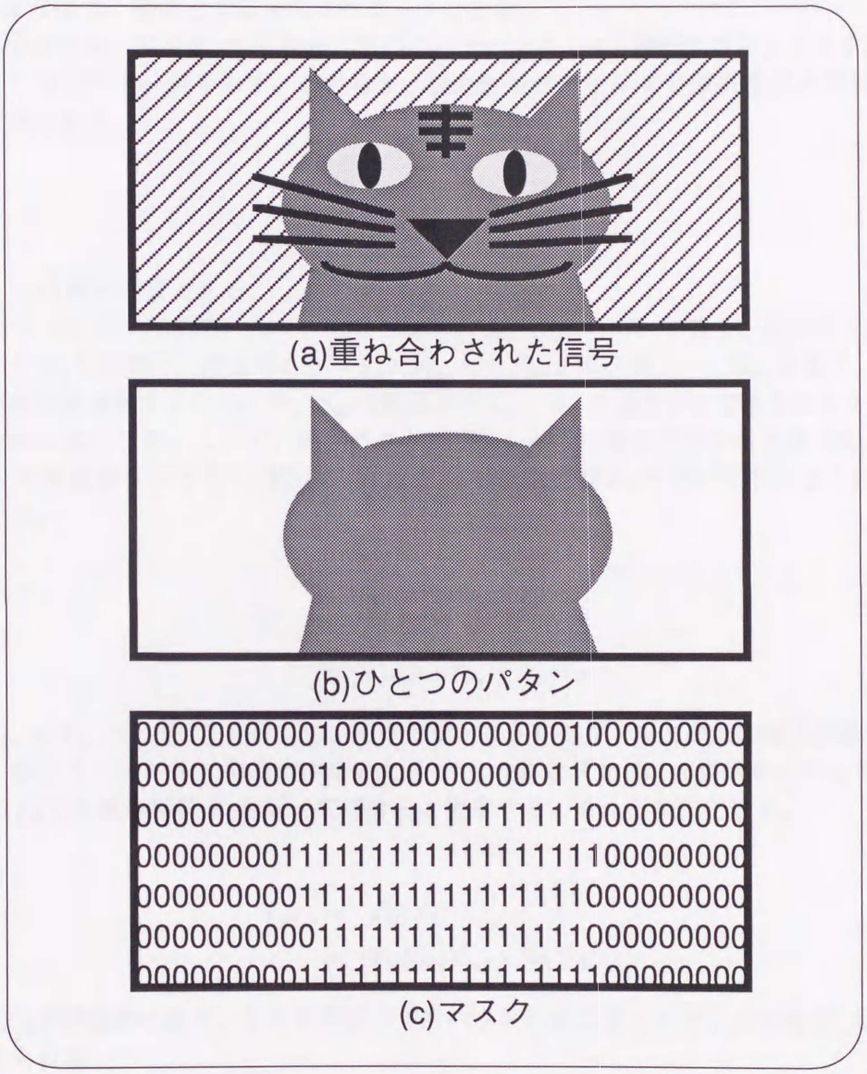


図 5.1: 複数パターンによって構成される画像モデル

$$X = \sum_{i=1}^{n} S_i Y_i \tag{5.1}$$

Sは 0 または 1 を要素とするマスクベクトルである。

簡単にするため、XがY_αとY_βなるふたつのパターンによって構成されると考える。図 5.2 にそのような信号の一例を示す。この場合、観測信号ベクトル Xと原信号 Yの関係は、次のように表される。

$$X = S_\alpha Y_\alpha + S_\beta Y_\beta \tag{5.2}$$

なお、Y_βは雑音要素である。

観測信号 Xを細かく区分けして、部分ベクトル $X_1, X_2, \cdots, X_n$と置く。原信号 Y_α、マスクベクトル S_αも同様に、部分ベクトル $Y_{\alpha_1}, Y_{\alpha_2}, \cdots, Y_{\alpha_n}$、$S_{\alpha_1}, S_{\alpha_2}, \cdots, S_{\alpha_n}$と置く。観測信号から原信号を推定するためには、$S_{\alpha_i}$の要素が全て 1、$S_{\beta_i}$の要素が 0 であるなら $Y_{\alpha_i} = X_i$ という単純な式になる。ここで、部分ベクトル $X_j (j \neq i)$ に雑音が少ないとき（S_{α_j}の要素が 1、S_{β_j}の要素が 0 のとき）、Y_{α_i}は、入力 X_jと非線形関数 F_{ji}を用いて次のように導けると仮定する。

$$\begin{aligned} Y_{\alpha_i} &= F_{ji}(X_j) \\ &= F_{ji}(S_{\alpha_j} Y_{\alpha_j} + S_{\beta_j} Y_{\beta_j}) \end{aligned} \tag{5.3}$$

F_{ji}は $X_{\alpha_h} \approx Y_{\alpha_h}$のとき Y_{α_i}を出力し、そうでないとき 0 ベクトルとなる非線形関数である。

一方、部分ベクトル $\mathrm{X_i}$に雑音が少ないとき（$\mathrm{S_{\alpha_i}}$の要素が 1、$\mathrm{S_{\beta_i}}$の要素が 0 のとき）、$\mathrm{Y_{\alpha_i}}$は、入力 $\mathrm{X_i}$と非線形関数 $\mathrm{F_{ii}}$を用いて導けると仮定する。それを次式で表す。

$$\begin{aligned} \mathrm{Y_{\alpha_i}} &= \mathrm{F_{ii}(X_i)} \\ &= \mathrm{F_{ii}(S_{\alpha_i} Y_{\alpha_i} + S_{\beta_i} Y_{\beta_i})} \end{aligned} \tag{5.4}$$

F_{ii}と F_{ji}が排他的であり、どちらかが必ず 0 ベクトルとなるならば、原信号 Y_{α_i}は次のように求められる。

$$\begin{aligned} Y_{\alpha_i} &= F_{ji}(X_j) + F_{ii}(X_i) \\ &= F_{ji}(S_{\alpha_j} Y_{\alpha_j} + S_{\beta_j} Y_{\beta_j}) + F_{ii}(S_{\alpha_j} Y_{\alpha_j} + S_{\beta_i} Y_{\beta_i}) \end{aligned} \tag{5.5}$$

すなわち、X_jまたは X_iのどちらか一方（両方でも可）が与えられれば原信号 Y_{α_i}を推定することは可能である。これらを排他的に 0 または 1 の値をとるδなる値と、$X_{\alpha_h} \approx Y_{\alpha_h}$の

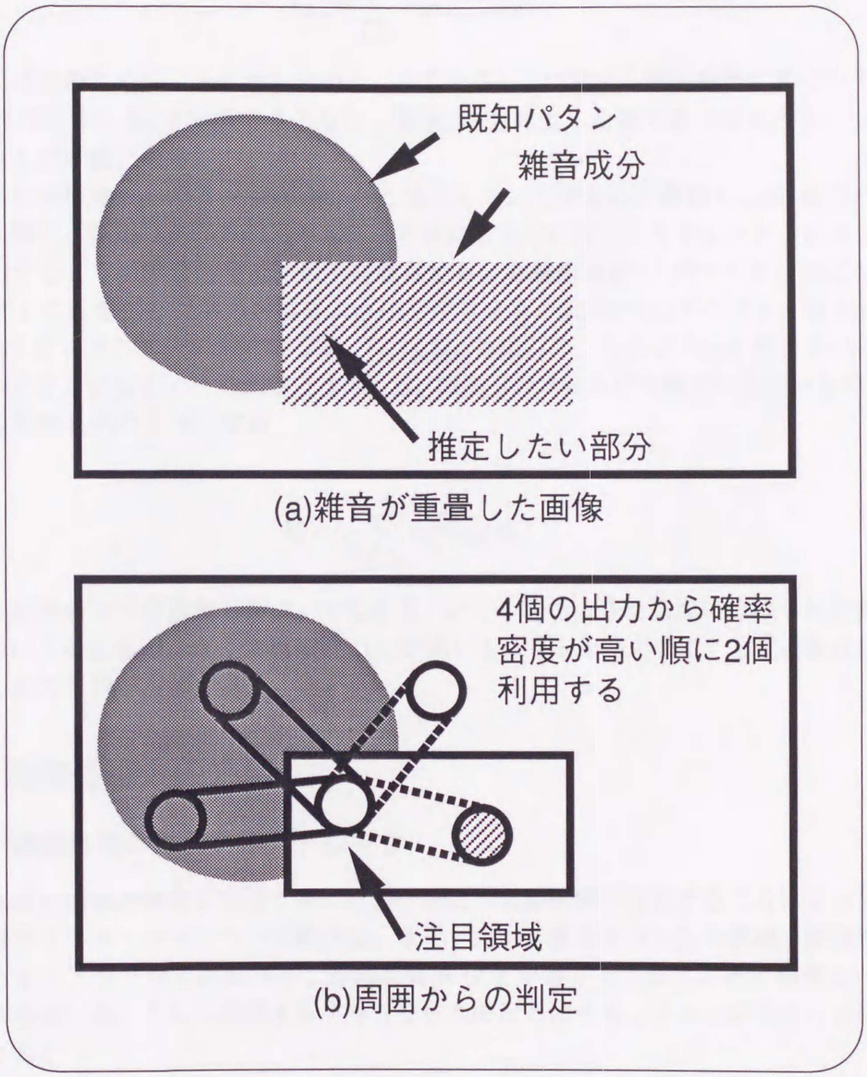


図 5.2: 1 つのパターンと雑音からなる画像

とき Y_{α_i} を出力し、そうでないとき 0 ベクトルとなるような関数 G を使って書き直すと一般式は次のようになる。

$$Y_{\alpha_i} = \sum_{h=1}^{n} \delta_{hi} G_{hi}(x_h) \tag{5.6}$$

この式は次のようなことを示している。すなわち、αに関する事前知識に基づいて非線形関数 G 及びしきい値δを定義できるなら、雑音βがどのような形であってもパターン Y を認識することが可能になる。

実際に雑音除去やパターン認識を行う場合について考える。関数 G は既知のパターンを複数記憶し、既知パターンが入力されたときにはラベルを、そうでないときには 0 ベクトルを出力するように学習させた非線形関数とする。非線形関数 G がラベルを出力しているときには 1 に正規化した出力を得るように学習させる。実際には G のノルムは 0 または 1 の値とならず、それに近いものとなる。δ_{hi}は G_{hi}の出力のノルムが G_{ki} より大きいとき 1 そうでないとき 0 になるような関数とする。G_{ki}は出力のノルムが k番目に大きいものである。入出力の関係を次のようにする。

$$Y_i = \frac{1}{k} \sum_{h=1}^{n} \delta_{hi} G_{hi}(x_h) \tag{5.7}$$

これは次のような意義を持つ。すなわち、いくつかの小さな領域に分けられた信号のうち、雑音による変形が少なくて信頼性 [11] が高いものを k 個取り出し、k 最近接近傍法 [12] によって出力を得るものである。

5.4 画像処理への応用

5.4.1 画像修復のためのネットワーク

この処理を画像の修復に応用する。近傍の画像から原画像を推定することによって画像の修復を行うネットワークについて述べる。まず、互いに重なり合った小領域の画像の雑音除去を行うネットワークを複数作り、非線形関数 G を学習させておく。ある画素について修復を行う場合には、それに関係するネットワークの出力のうち、ノルムが大きいもの k 個の平均をとる。

このような神経回路網システムを図 5.3に示す。小領域の画素から推定値を出力する複数の確率密度を自己学習する MLP と、その出力を統合する k 最近接近傍フィルタによりシステムを構成する。

この神経回路網システムは図 5.4のようなブロックに分けれられる。すなわち、入力画像を小さな領域に分割する分配部、小領域の雑音除去を行う確率密度を学習する MLP、そしてその出力を k 最近接近傍法によって統合する統合部である。

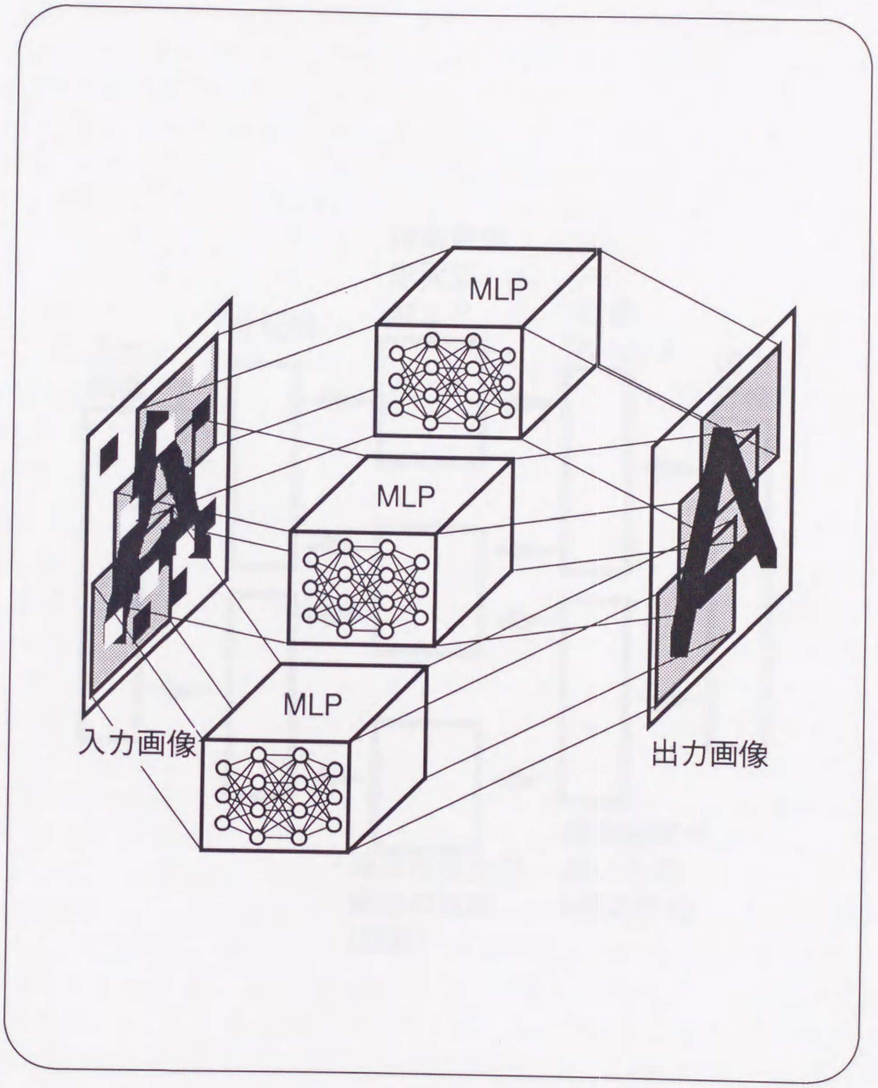


図 5.3: 神経回路網システム

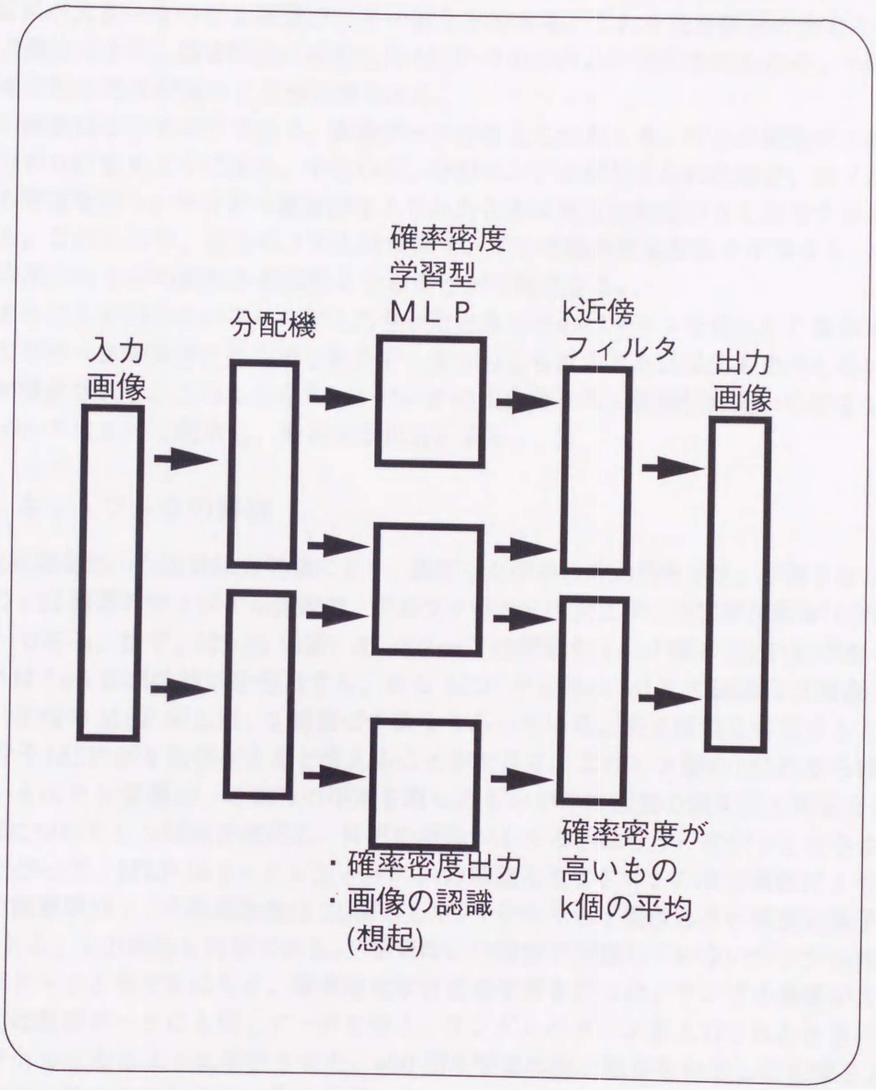


図 5.4: 神経回路網システムの構成

入力された画像を小さな領域に分割して雑音を除去する。その領域は互いに重なり合っており、ひとつの画素は複数の MLP よって雑音が除去される。雑音除去結果は各 MLP によって異なることが予想されるが、より効果的に雑音除去が行われたものを採用する。

各画素の出力は次のように計算される。注目画素に関係する複数の MLP のうち、出力の 2 乗の総和が大きいものを k 個選び、その値を平均する。これを注目画素の出力とする。k 最近接近傍法により、雑音除去に成功した MLP の出力のみが利用されるため、一般的な方法より雑音除去効果が高いことが予想される。

入出力細胞は 2 つでペアであり、画像データが与えられたとき、片方の細胞が 1 のときはもう一方が 0 になるようにする。すなわち、学習サンプルが与えられた場合にはノルムが一定となる学習を行う。ランダム雑音が与えられたときは両方の細胞が 0 を出力するように学習させる。これにより、出力のノルムは学習サンプルの確率密度関数を学習する。従って、各細胞の出力の 2 乗の総和を信頼性とすることが可能になる。

このようにして既知のパターンが入力されたときはそのパターンを復元した結果を、極端にゆがんだデータや雑音に打ち消されたデータが与えられたときには何も出力しないネットワークが構成される。このようなネットワークの出力のうち、信頼性の高いものを k 最近接近傍フィルタによって統合し、最終的な出力とする。

5.4.2 ネットワークの詳細

2 次元画像を用いた雑音除去実験により、提案した手法の有効性を示す。学習させたパターンは、12×24 画素のディジタル画像で、アルファベットの大文字、小文字と数字（合計 62 パターン）である。まず、12×24 画素、62 パターンの画像を 4×10 個の MLP に学習させる。各 MLP は 5×5 画素の領域を担当する。ある MLP と近傍の MLP の距離は 2 画素である。つまり、近傍の MLP 同士は、3 画素づつ重なりあっている。ある画素に着目すると、雑音除去を行う MLP が 9 個存在すると考えることができる。これら 9 個の MLP から確率密度が大きいものを k 個選び、これらの平均を取ったものがその画素の雑音除去結果となる。

1 画素について 2 つ細胞を使用し、片方の細胞が 1 のときはもう一方が 0 となるようにする。したがって、MLP は $5 \times 5 \times 2(+1)$ の入力細胞と $5 \times 5 \times 2$ の出力細胞により構成される。中間層数は 1、中間細胞数は 11 個とした。（そのうち 1 個はしきい値決定素子で常に 1.0 を与える。入力細胞も同様である。）学習時には雑音を付加していないサンプル画像とランダムパターンとを交互に与え、確率密度学習型の学習を行った。サンプル画像が入力されたときには教師データにも同じデータを与え、ランダムパターンを入力されたときには出力が 0 ベクトルになるように学習させた。400 回の学習の後、雑音を付加した画像を入力し、どの程度雑音除去が行われているかを調べた。

5.4.3 最適パラメータの測定

このようにして学習させたネットワークを用いて雑音除去実験を行い、kの値の最適値を求めた。画素の候補数を n とし、n に対する kの比をパラメータとして最適な値を調べた。

この実験を行った結果を図 5.5に示す。n に対する kの比が 30%のとき、最も雑音が小さくなった。この値から、学習させた画像には次のような性質があることが考えられる。

ひとつの MLP が学習したパターン数は、62 パターンである。これは、全パターン数の 1.6%に相当する。MLP が学習した文字パターン（その MLP が担当する文字パターン）と学習しなかった文字パターン（別の MLP が担当する文字パターン）の性質が全く異なるなら、最も雑音が小さくなるのは n に対する kの比が 1.6%を越えないはずである。しかし、ひとつのネットワークが担当する領域は小さいため、学習させなかったパターンに対しても雑音除去効果を持つことが考えられる。例えば、「C」というパターンの左側を受け持つネットワークは、「O」というパターンの雑音除去にも有効であることが予想できる。n に対する kの比が 30%の時最適となるのは、そのネットワークに学習させなかったパターンの内、学習させた入力パターンと局所的に類似するパターンが 28%程度存在することを示している。

この実験に用いたプログラムは k の最小値を 1 としている。本実験における n/kが 0%のときは $k=1$ となる。これは、複数の候補から最も確率密度が高いものを取り出す処理の一種であるということができる。n に対する kの比が小さすぎると雑音除去効果が低いということは、一般的な確率密度推定による雑音除去より、本手法の効果が高いことを示している。

5.4.4 従来の手法との比較

確率密度を自己学習する MLP を用いたときと従来の MLP を用いたときとの比較実験について述べる。実験に用いた画像およびシステムは最適パラメータを測定したときと同じものを用いた。局所的な雑音除去を行うネットワークとして確率密度を自己学習する MLP と従来型 MLP、結果の統合出力として k 平均と全平均を用い、これらの組み合わせによって 4 種類の方法で雑音除去を行った。シミュレーションに用いた文字パターンと実験結果を図 5.6から図 5.11に示す。

この結果では「E」、「F」、「W」のようにいくつかのパターンで画像の修復に失敗している。「E」と「F」は類似のパターンであるため、パターンの学習に失敗したことが考えられる。また、「W」のようなパターンは、パターンの形が複雑であるため学習に失敗したものと思われる。シミュレーションで用いた MLP ではこれらのパターンを学習させるには能力不足であり、学習が適切に行われなかったものと考えることができる。学習回数が少ないことも要因のひとつであろう。これらのパターンを適切に処理するには、学習回数を増やす、細胞数を増やす、学習パラメータをチューニングするなどの対応策を行う必要があると思われる。

各手法の出力の二乗誤差を表 5.1に示す。この結果では、確率密度を自己学習する MLP と k 平均によるフィルタの組み合わせが最もよい値となった。確率密度を自己学習する MLP を用いたときはいずれも従来型 MLP よりよい結果を示している。従来型 MLP は学習させてないパターンが与えれたときは、予期できない（でたらめな）結果を出力するためである。なお、従来型 MLP では出力のノルムは無意味である。従来型 MLP の k 平均は、ネットワーク出力からランダムに k個取り出して平均を求めているに過ぎない。そのため、従来

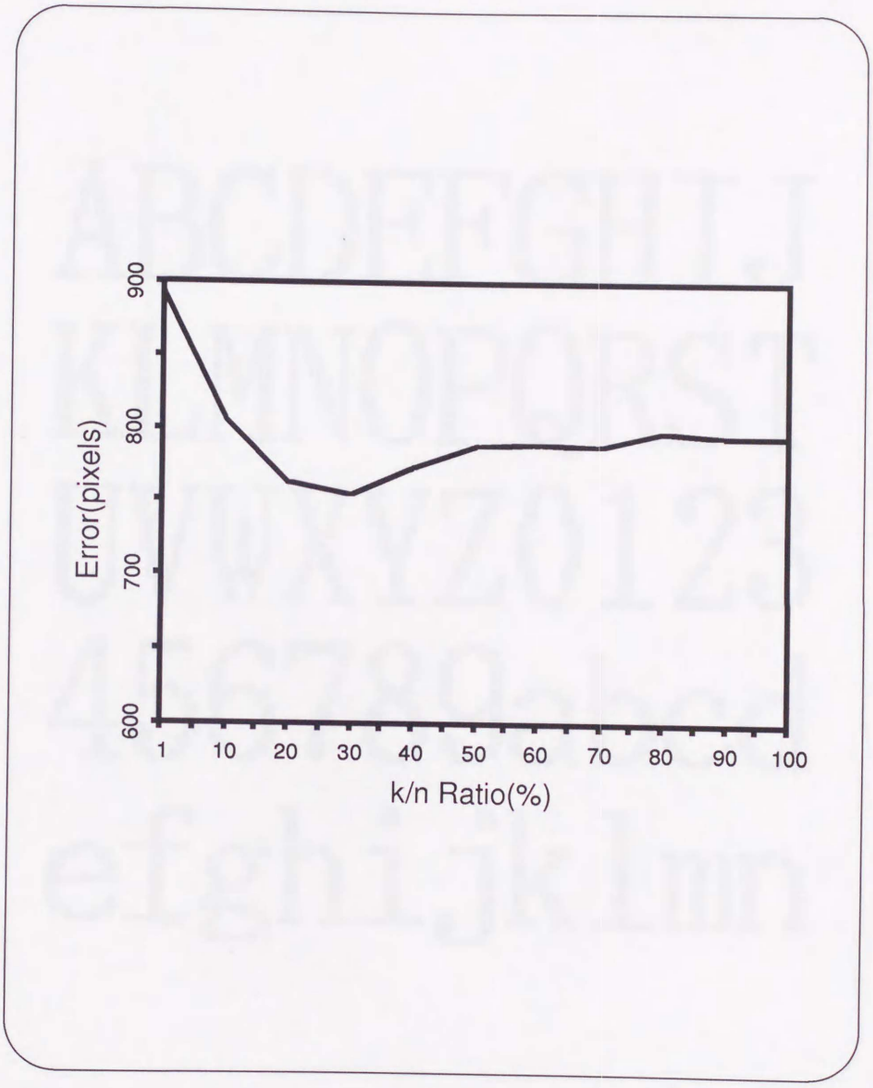


図 5.5: n に対する k の割合と二乗誤差の関係

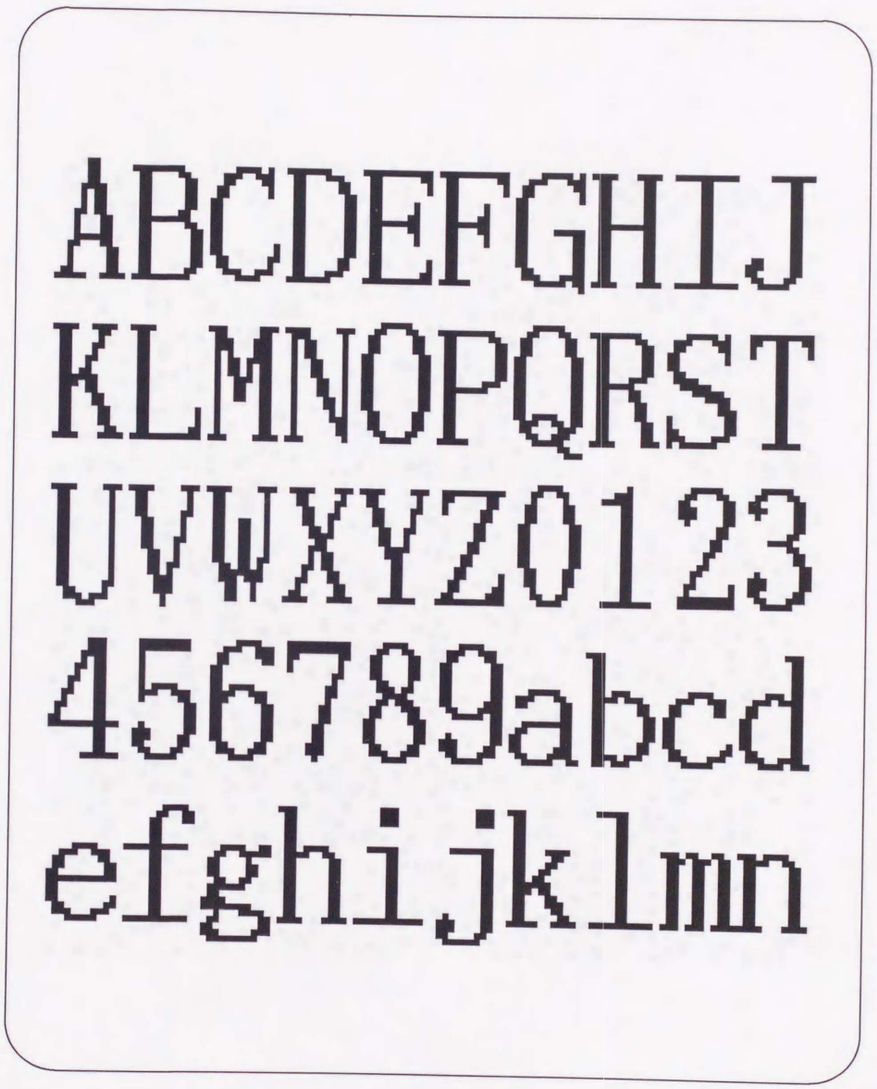


図 5.6: シミュレーションに用いた文字パターン

図 5.7: 雑音を付加したパターン

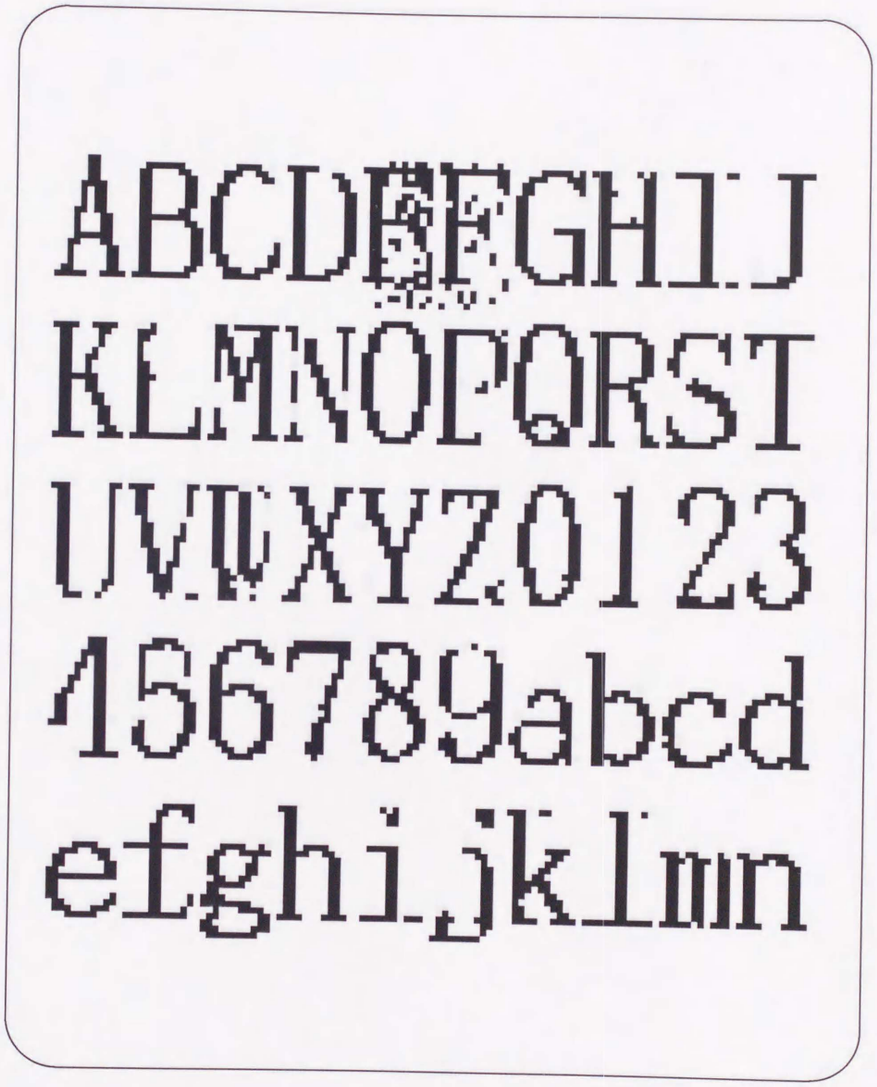


図 5.8: 確率密度を自己学習する MLP と k 近傍フィルタによる処理結果

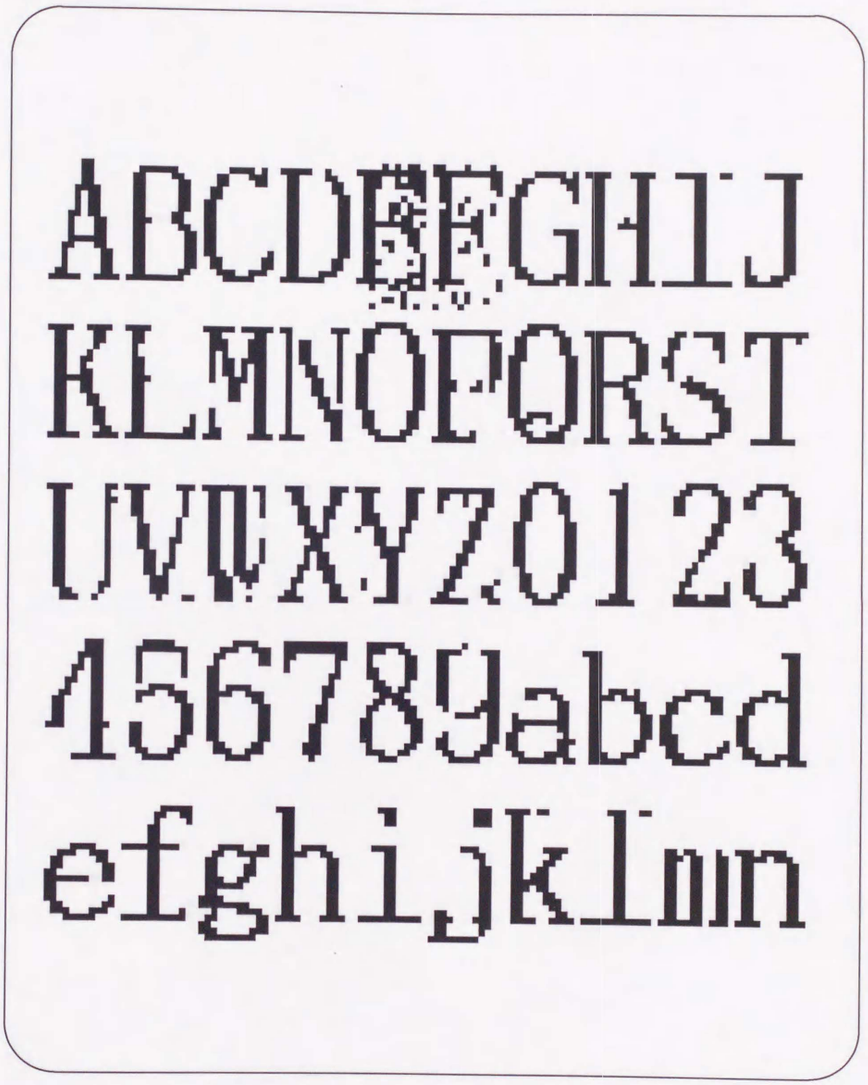


図 5.9: 確率密度を自己学習する MLP と全平均フィルタによる処理結果

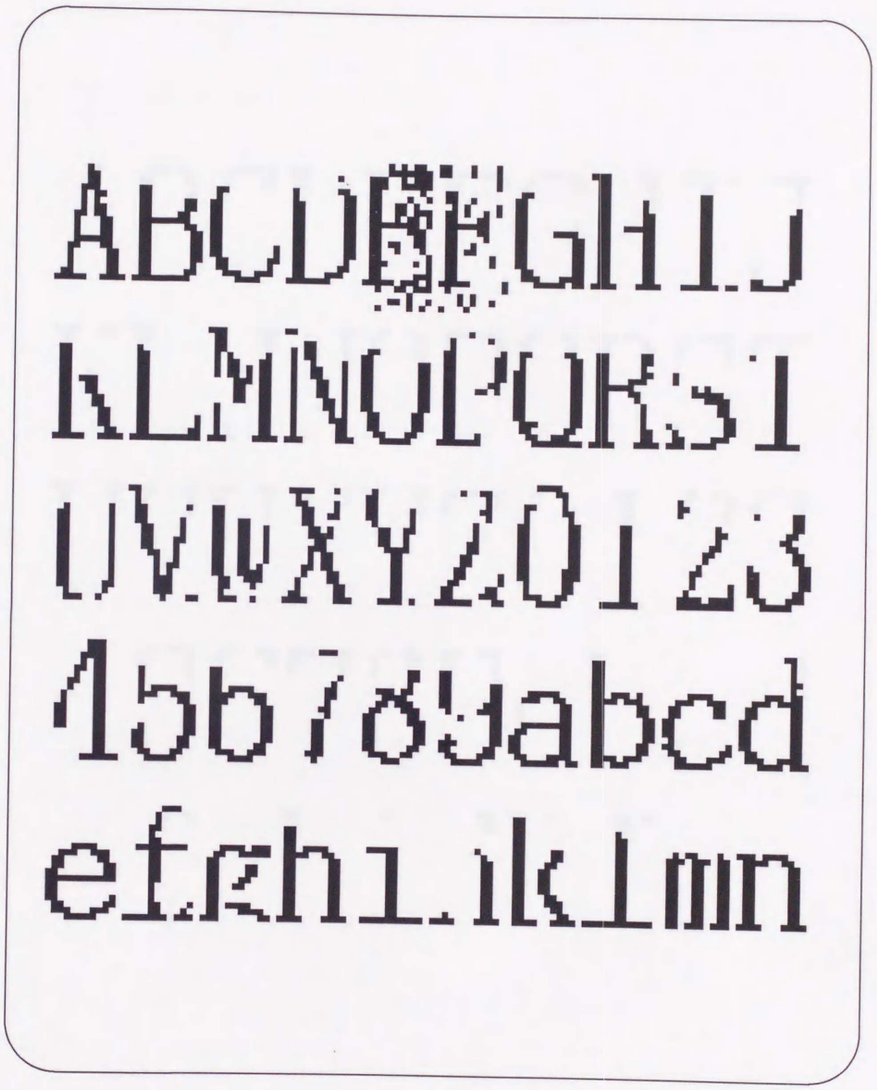


図 5.10: 従来型 MLP と k 近傍フィルタによる処理結果

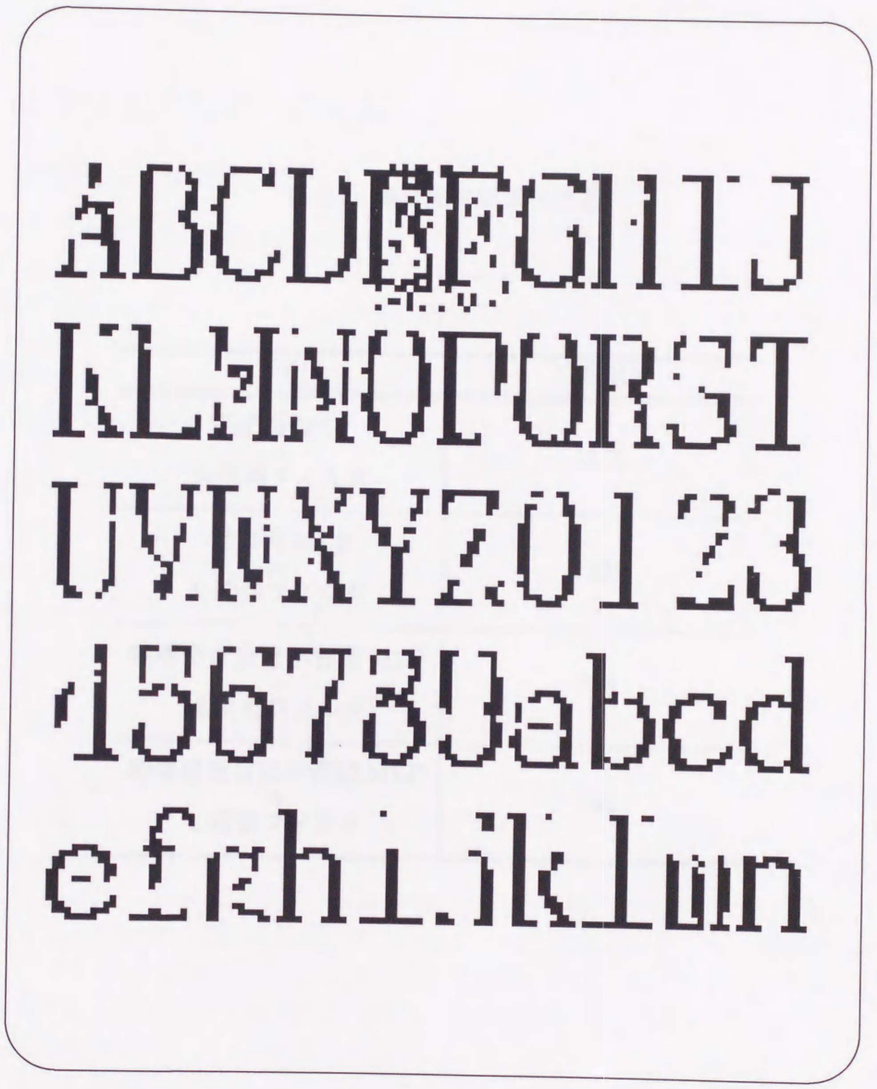


図 5.11: 従来型 MLP と全平均フィルタによる処理結果

表 5.1: 雑音除去結果の比較

手法	二乗誤差
従来型 MLP + 全平均フィルタ	1175
従来型 MLP + k 近傍フィルタ	1272
確率密度自己学習型 MLP + 全平均フィルタ	795
確率密度自己学習型 MLP + k 近傍フィルタ	755

型 MLP を用いたとき、k 平均の結果は全平均の結果より雑音除去効果が低い。従来型 MLP ではパターン学習とは別のプロセスで確率密度推定を行って統合化しなければならない。この実験のような単純な結合では効果を得ることが困難であることが結果より判る。すなわち、実験結果は、確率密度を自己学習する MLP の k 平均を出力として用いたとき、従来手法よりネットワークの組み合わせが容易で、雑音除去効果が高いことを示している。

5.5 手書き文字認識への応用

5.5.1 手書き文字認識システム

本手法を手書き文字認識に応用し、確率密度を自己学習する MLP を用いて高精度な分類が容易に実現できることを示す。手書き文字認識の問題点は、認識対象となるパターン数が多く分類が困難なことと、文字パターンの変形により特徴ベクトルの抽出が困難なことである。本研究では、認識対象となるパターンが多い場合に、確率密度を自己学習する MLP を用いて学習を行うことにより、分類精度を向上させることを目的とする。

本研究で構成した手書き文字認識システムのブロックダイアグラムを図 5.12に示す。本システムは、細線化、特徴ベクトル抽出、パターン分類、分類結果統合の処理により構成される。認識対象とする手書き文字パターンは、電総研の手書き文字データ ETL9B[13] である。はじめに、マスクキングの繰り返しを用いて [14] 細線化を行う。次に、特徴ベクトルを抽出し、196 次元の特徴ベクトルを抽出する。特徴ベクトルの抽出には、手書き文字認識システム DASH[15, 16] で用いられた方法を用いる。パターンの分類には、確率密度を自己学習する MLP を用いる。

5.5.2 細線化および特徴ベクトルの抽出

ETL9B は 63×64 画素の 2 値の手書き文字データである。手書き文字の認識精度は、文字の太さの相違や変形に影響される。この影響を軽減するため、細線化処理及び特徴ベクトルの抽出処理を行う。この処理は、本研究の主目的でないため、手書き文字認識システム DASH をベースに、簡単なアルゴリズムで構成した。処理手順を以下に述べる。

はじめに文字の太さの影響を軽減するために入力文字の細線化を行う。細線化のアルゴリズムは、マスクキングの繰り返し法を用いた。これは、3×3 画素のマスクパターンと注目画素近傍を比較して削除条件に合う画素を削除する処理を繰り返す方法である。マスクパターンには、8 種類の 2 値パターンを用いた [17]。

次に、文字の変形の影響を軽減し特徴量を圧縮するために、方向線素特徴量を求める。方向線素特徴量は、画像内の小領域における線素の方向である。これは、図 5.13に示すように、各小領域について縦横斜め方向の 4 種類のパターンから成り立つ。細線化画像内の小領域の数は 7×7 個である。方向線素特徴量の抽出処理により、各小領域について 4 次元の特徴量が求められる。ひとつの文字に対し、小領域は 49 個あり、各小領域が 4 次元であるた

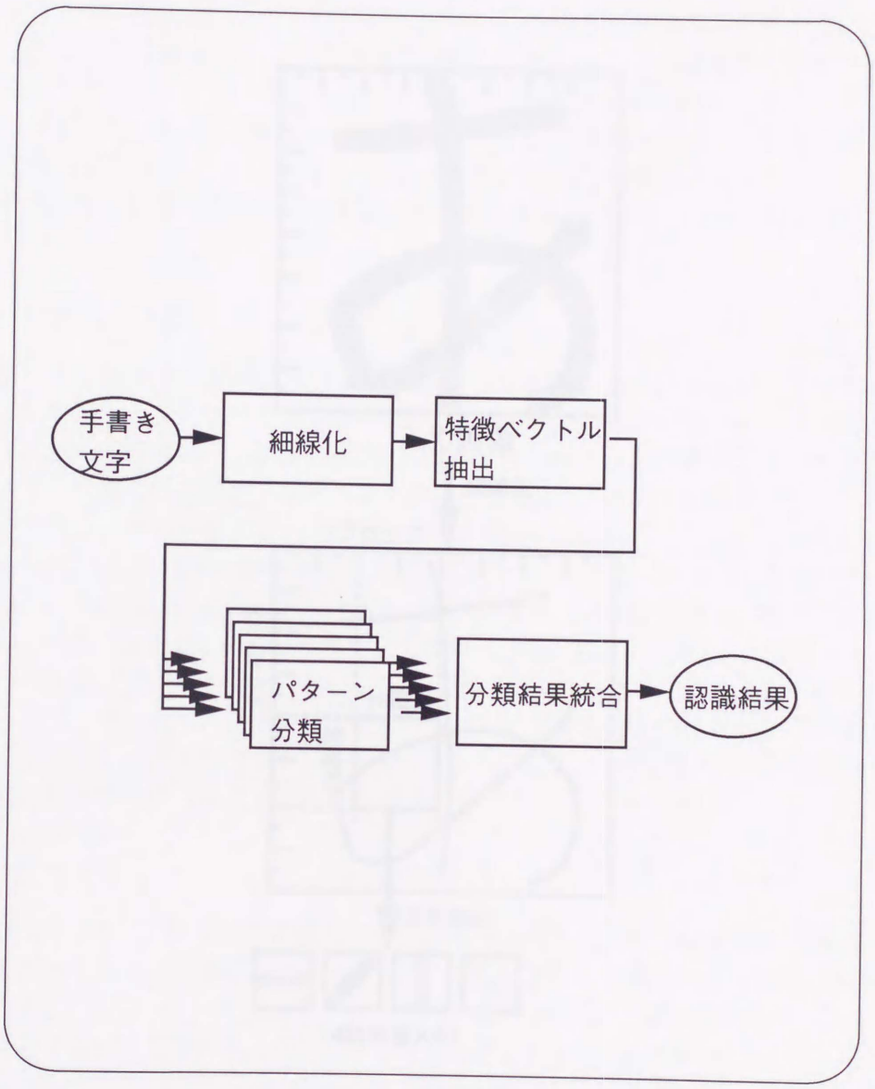


図 5.12: 文字認識システムのブロックダイアグラム

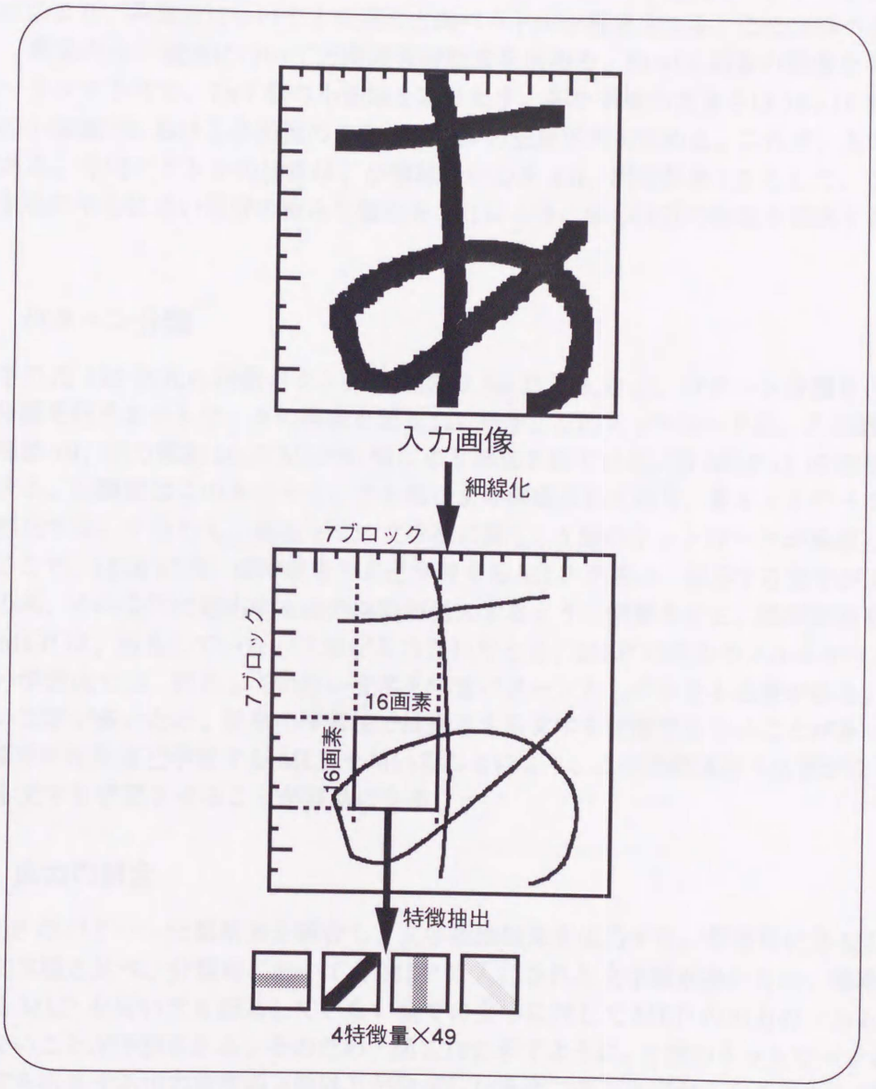


図 5.13: 細線化および特徴抽出

め、その文字全体の特徴量は 196 次元となる。方向線素特徴量を求めるための具体的な手順を以下に述べる。

まず、全画素について縦横斜め方向の 4 種類のパターンを抽出する。注目画素の近傍 8 画素を調べ、近傍画素の値が 1 のときは、注目点から近傍画素への方向のベクトルがあるとする。これにより、各画素について 4 次元の方向ベクトルが構成される。この方向ベクトルを用いて、画像内の小領域について方向線素特徴量を求める。64×64 画素の画像を半分ずつオーバーラップさせて、7×7 個の小領域を取り出す。各小領域の大きさは 16×16 画素である。この小領域内における各画素の方向ベクトルの加重平均を求める。これが、方向線素特徴量である。方向ベクトルの加重は、小領域の中心を 4.0、周辺部を 1.0 とした。この方法は、小領域の中心に近い部分の重みを強めることにより、中心付近の特徴を強調する効果がある。

5.5.3 パターン分類

抽出された 196 次元の特徴ベクトルを複数の MLP に入力し、パターン分類を行う。パターン分類を行うネットワークの構成を図 5.14に示す。このネットワークは、入力細胞 196、中間細胞数 10、出力細胞 10 の MLP95 個により構成されている。各 MLP は 10 種類の文字を担当する。分類部はこのネットワーク 6 個により構成されており、各ネットワークは全文字種を担当する。すなわち、ある 1 つの文字種に関し、5 個のネットワークが重複して担当する。ここで、MLP には、確率密度を自己学習する MLP を用い、担当する文字が入力されたときのみ、その文字に対応する出力細胞が発火するように学習させる。確率密度を自己学習する MLP は、担当していない文字が入力されたとき、MLP の出力のノルムが 0 となる。

従来の学習法では、担当していない文字を学習パターンとして与える必要がある。担当していない文字が多いため、従来の学習法では担当する文字を記憶できないことがあった。しかし、確率密度を自己学習する MLP を用いることにより、この処理は行う必要がなくなり、担当する文字を学習させることが容易になる。

5.5.4 出力の統合

各 MLP のパターン分類結果を統合し、文字認識結果を出力する。学習時に各 MLP が担当する文字種と比べ、分類時において各 MLP に入力される文字種が多いため、確率密度を学習する MLP を用いても担当していない全ての文字に対して MLP の出力のノルムが 0 とはならないことが予想される。そのため、図 5.15に示すように、5 個のネットワークのうち、ある文字を担当する出力細胞の k個以上が発火した場合、その文字が入力されたと判定する方式を用いた。本実験では、$k = 5$、すなわちすべてのネットワークが発火した場合に、その文字と判定することとした。

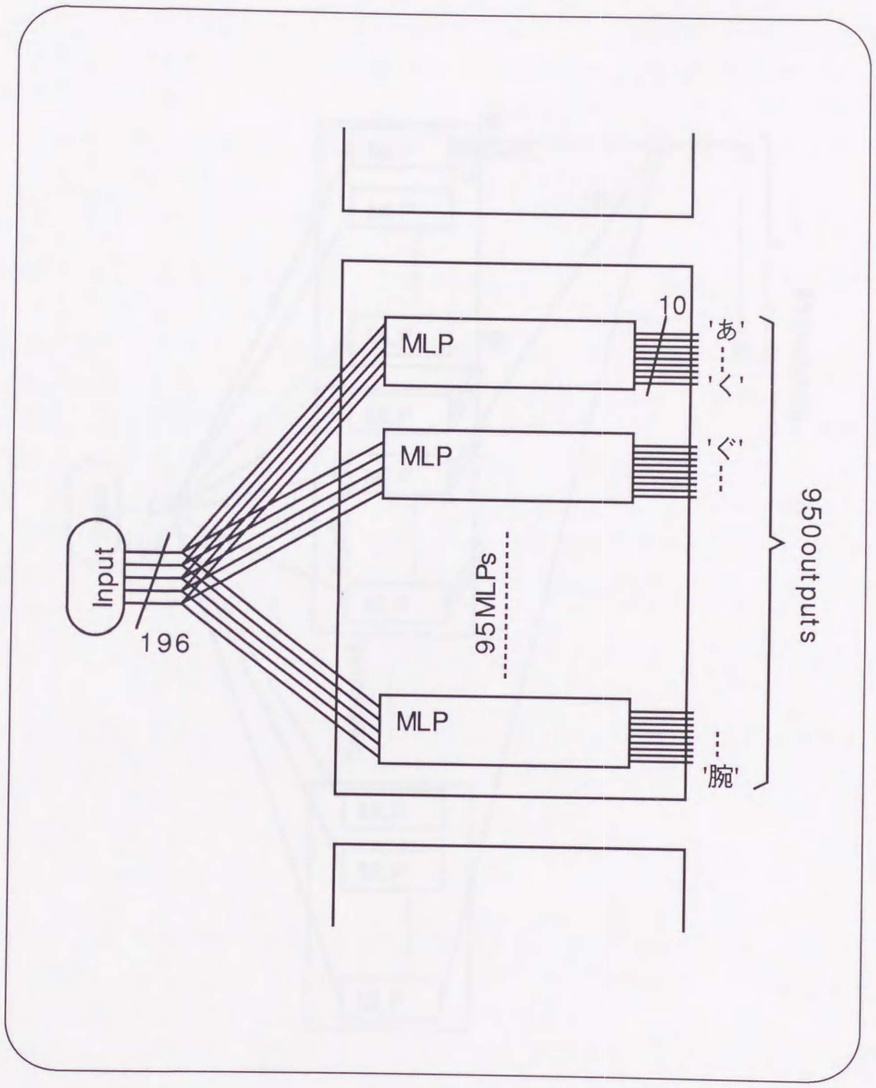


図 5.14: パターン分類を行うネットワーク

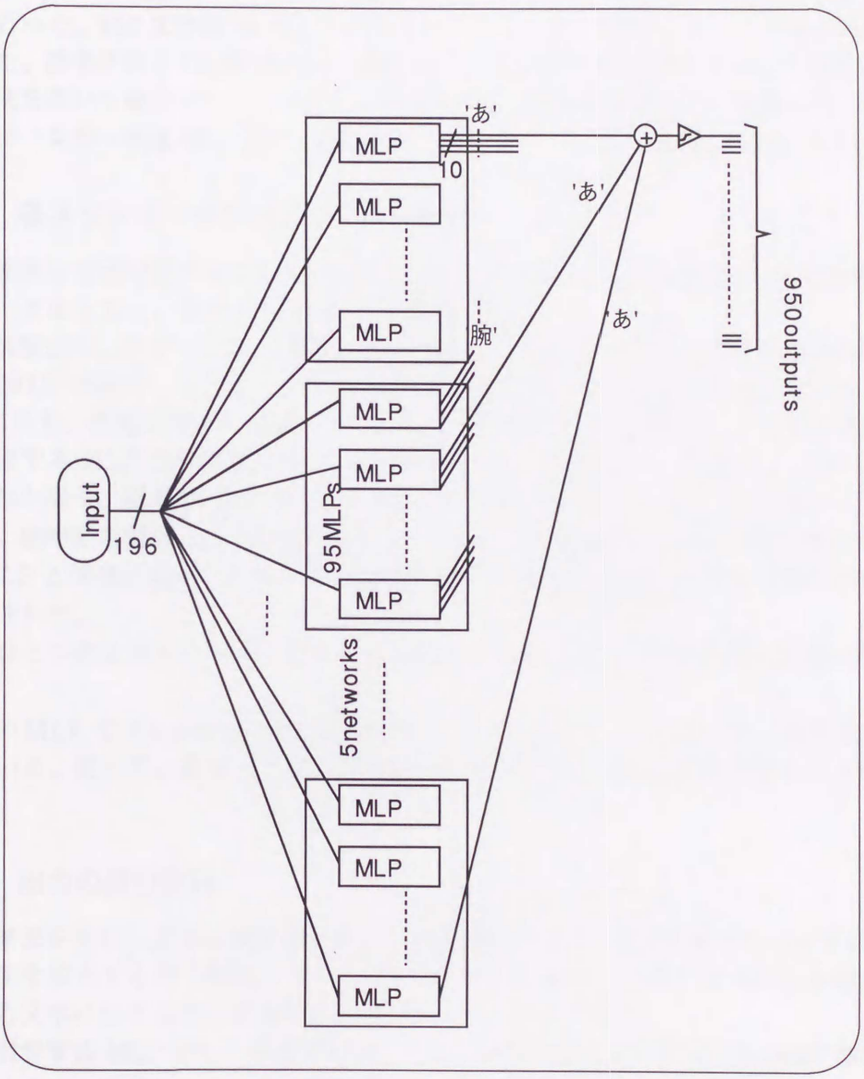

Input
196
MLP
MLP
MLP
MLP
MLP
MLP
MLP
MLP
MLP
95MLPs
5networks
10
'あ'
'腕'
'あ'
'あ'
'あ'
950outputs

図 5.15: 出力の統合

5.6 ETL9B を用いた認識実験

5.6.1 認識実験の概要

手書き文字データ ETL9B のうち、教育漢字 950 文字種を用いて文字認識システムの精度評価を行った。950 文字種 10 セットを学習サンプルとし、他の 1 セットを用いて認識実験を行った。提案手法との比較のために確率密度を自己学習する MLP を用いた場合と、従来型学習法を用いた場合についてパターン分類及び統合結果を評価した。実験に用いた手書き文字パターンの一部を図 5.16に、細線化後の文字パターンの一部を図 5.17に示す。

5.6.2 各ネットワークのパターン分類実験

確率密度を自己学習する MLP と従来型学習法 MLP ともに、各 MLP に 5 文字種 10 セットのデータを入力し、各セットを 300 回学習させた。

分類実験に用いたデータは、学習に用いなかった 1 セットである。担当していない文字に対して MLP が発火してしまったり、担当する文字に対して MLP が出力しなかった確率を表 5.2に示す。従来型学習法 MLP では、その平均確率は 24.6%となった。また確率密度を自己学習する MLP の平均確率は、17.7%となった。この結果は、担当していない文字種が入力された場合、従来型 MLP の方が予期しない結果が出力されるケースが多いことを示している。確率密度関数を自己学習する MLP では、中間細胞数が少ないなどの場合には、従来型 MLP と同様の結果となるが、本実験の場合、中間細胞数が 10 でも有効性があることが確認された。

あるひとつのネットワークにおける、入力文字に対する出力結果を図 5.18及び図 5.19に示す。

複数の MLP でノルムが 0 とならないため、ひとつのネットワークでは、候補文字が複数現れている。従って、各ネットワークの結果をそのまま文字認識に応用することはできなかった。

5.6.3 出力の統合実験

分類結果を元に、文字認識結果を出力する実験を行った。5 個のネットワークにおいて、ある文字を担当する出力細胞がすべて発火した場合、その文字が入力されたと判定するとき、入力文字に対する出力結果を図 5.20および図 5.21に示す。

従来型学習法 MLP では、認識率は 70.21%、確率密度を自己学習する MLP の認識率は 93.26%となり、提案システムが手書き文字認識に有効であることが確認された。

各ネットワークで分類に失敗する確率は、従来型 MLP と確率密度を自己学習する MLP とで 7%程度しか差がなかったが、5 つのネットワークの統合結果との差は、20%以上となった。この理由は、従来型 MLP では学習していない類似パターンの分離が困難であること、出力を統合するために各ネットワークの失敗が指数的に減少するため誤差が増幅されたこと、などが考えられる。

図 5.16: 実験に用いた手書き文字パターンの一部

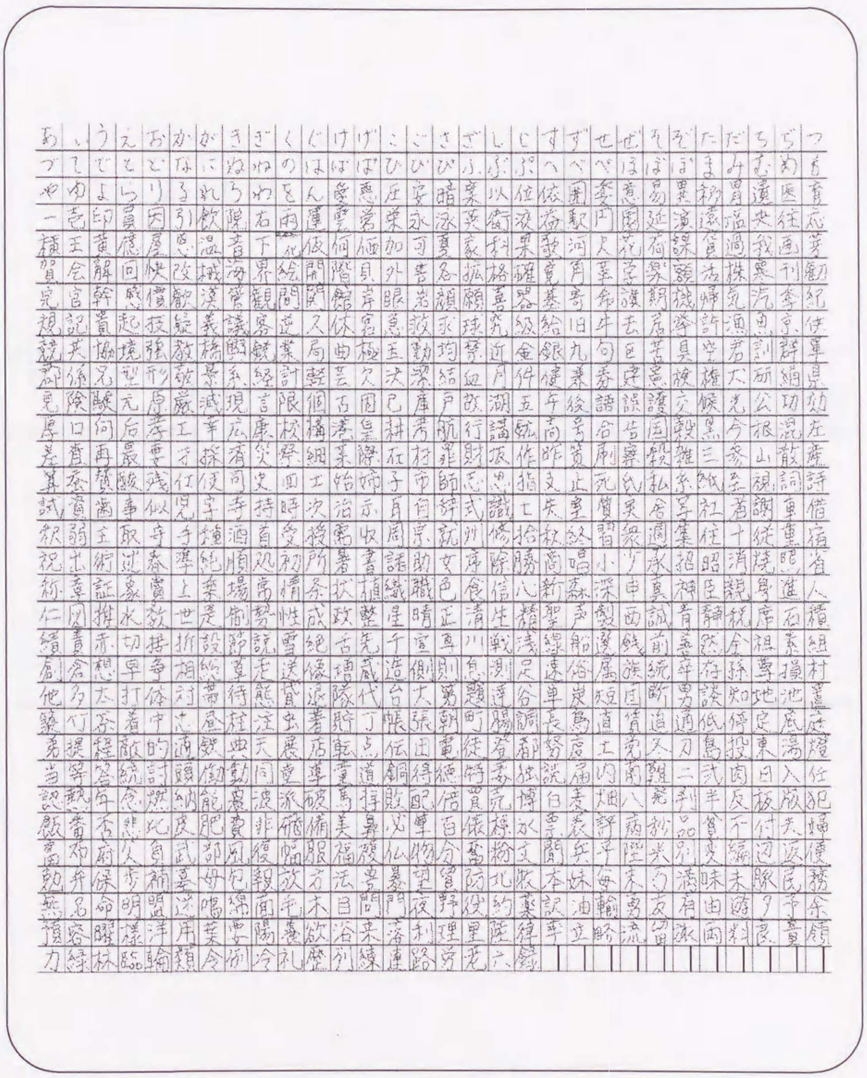


図 5.17: 細線化された文字パターン

表 5.2: ネットワークのパターン分類失敗率 (%)

ネットワーク番号	ネットワークのパターン分類失敗率	
	従来型 MLP	確率密度自己学習型 MLP
0	24.2726	17.7452
1	24.1817	17.1579
2	24.5019	17.4648
3	25.1801	18.0022
4	24.7457	18.1208
5	24.6172	17.6233
平均値	24.5832	17.6857

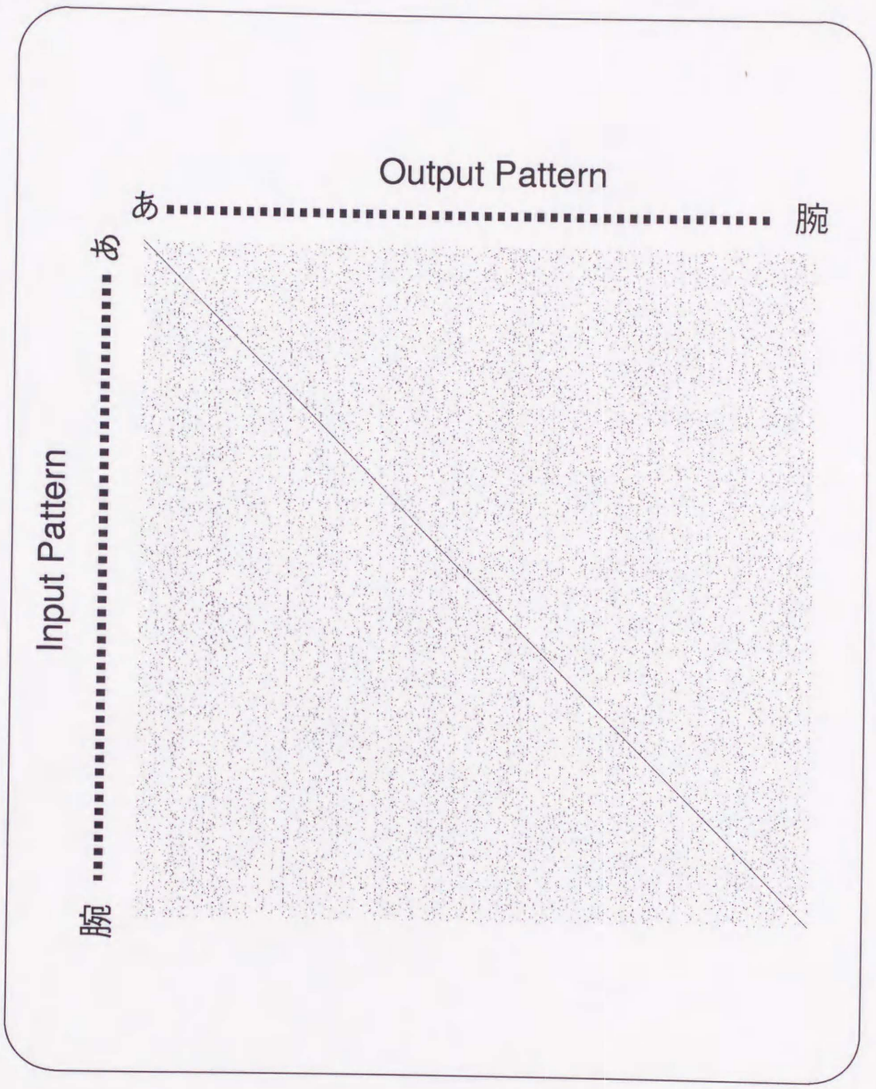


図 5.18: ネットワークの出力の一例（従来型 MLP）

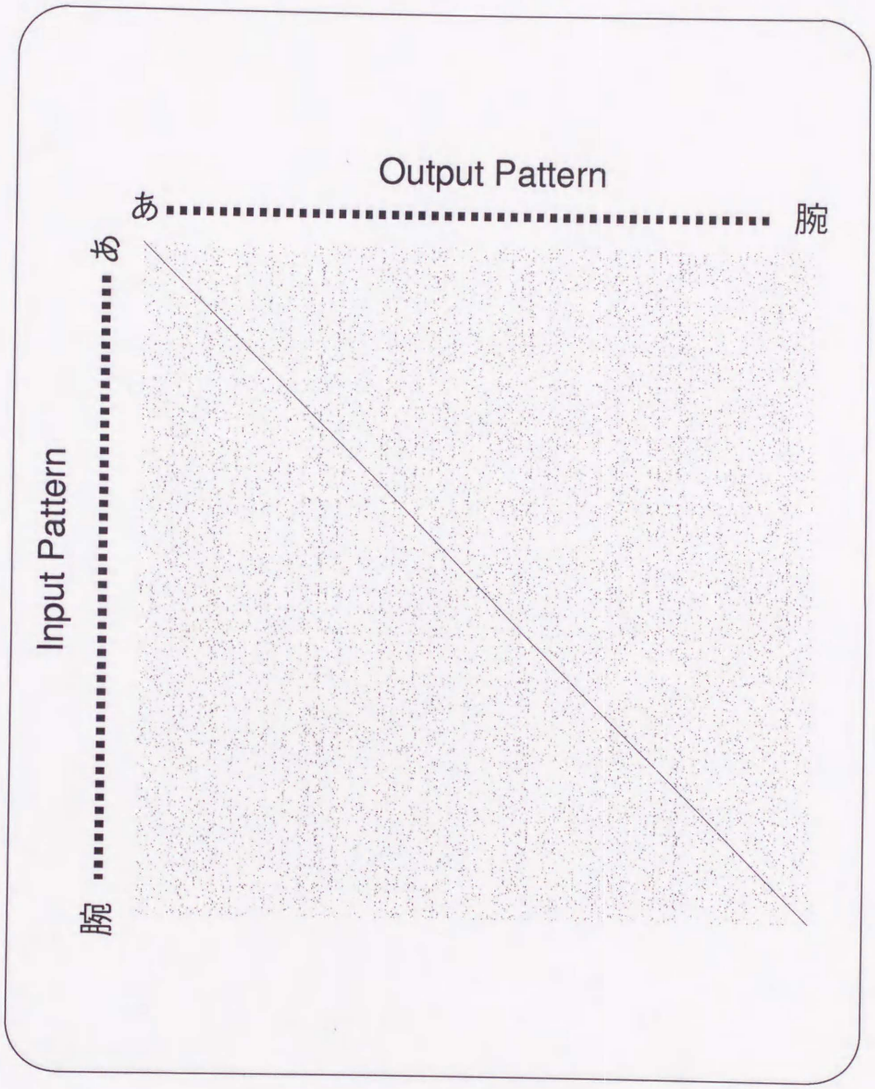


図 5.19: ネットワークの出力の一例（確率密度自己学習型 MLP）

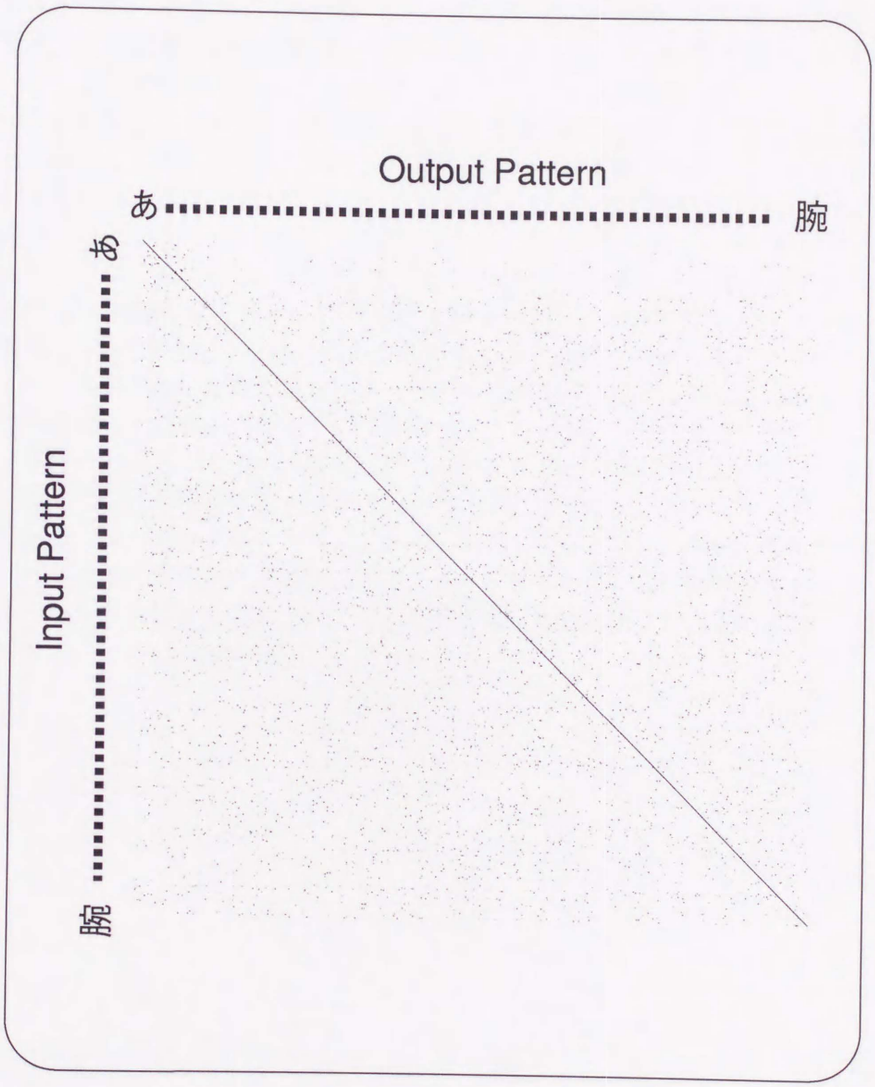


図 5.20: 従来型 MLP による認識結果

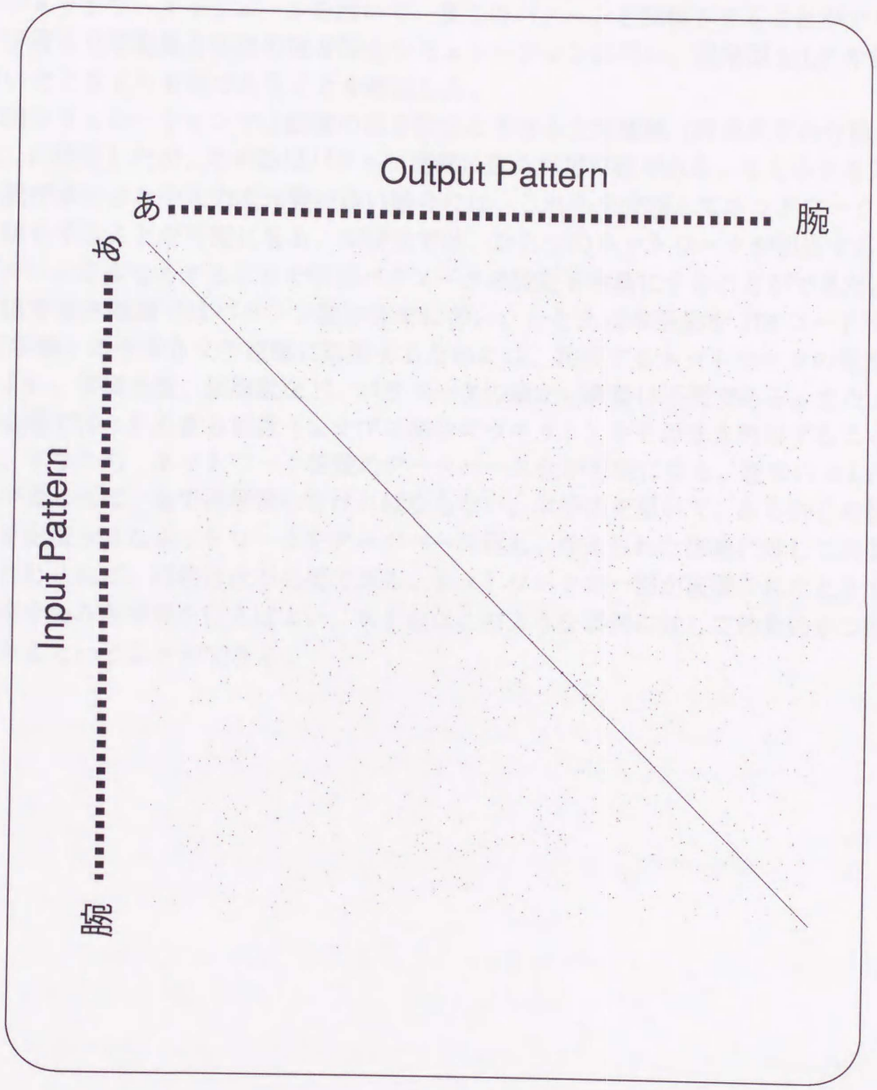


図 5.21: 確率密度自己学習型 MLP による認識結果

5.7 本章のまとめ

確率密度を自己学習する MLP を用い、これと k 近傍フィルタなどの組み合わせによって、複雑な調整無しで実現できる回路網システムを提案した。このシステムを用いれば記憶容量が小さいネットワークモジュールを用いて、多くのパターンを記憶させることができる。本手法を手書き文字認識と画像の雑音除去シミュレーションに用い、従来型 MLP や最尤推定法を用いたときより有効であることを確認した。

計算機シミュレーションでは画像の雑音除去と手書き文字認識（教育漢字の分類、約 950 文字種）に限定したが、本手法はパターン認識に広く応用可能である。与えられる入力のパターン数が多いときや入力次元数が多い場合には、これらを分割してネットワークに与え、出力を統合することが可能になる。本研究では、ひとつのネットワークが担当する画素数、文字パターンを少なくすることで学習パラメータの設定を容易にすることができた。手書き漢字認識や音声認識ではパターン数が非常に多い。たとえば本手法を JIS コード全て（約 3000 文字種）の手書き文字認識に応用するためには、利用するネットワークの数を増やすだけでよい。学習係数、細胞数など、パラメータの細かい調整は不要である。また、本研究の実験結果で得られた重み係数（シナプス結合のウエイト）をそのまま利用することも可能である。すなわち、ネットワーク係数のデータベース化が可能になる。従来の MLP では学習対象が変われば、全て再学習しなければならない。本手法を用いて、あらかじめ様々なパターンを記憶させたネットワークをデータベース化し、与えられた課題に対して必要なものを組み合わせれば、再学習は不必要である。ネットワークの一部が破壊されたときも、破壊された部分のみ再学習を行えばよい。本手法はこのような事例に対して効果的かつ効率的な方法であるということができる。

参考文献

[1] 久間和生, 中山高, "ニューロコンピュータ工学", 工業調査会 (1992).

[2] Y.Mori, K.Joe, "A Large–Scale Neulal Network witch Recognizes Handwritten Kanji Characters", Advances in Neural Information Processing Systems, 2, pp.415–421 (1989).

[3] 森吉弘, "大規模ニューラルネットワークの構想に向けて", 信学技報, Vol.PRU88–59, pp.87–94 (1988).

[4] 岩田彰, 當麻孝志, 松尾啓志, 鈴木宣夫, "大規模 4 層ニューラルネット Comb NET", 信学論 (D–II), Vol.J73–D–II, No.8, pp.1261–1267 (1989).

[5] 堀田健一, 岩田彰, 松尾啓志, 鈴村宣夫, "大規模ニューラルネット "CombNET–II" ", 信学論 (D–II), Vol.J75–D–II, No.3, pp.545–553 (1992).

[6] Iwata, A., Ino, Y., Hotta, K., Suzumura, N., "Hand-written Japanese kanji character recognition by a structured self–growing neural network "CombNET–II" ", Artificial Neural Networks 2 (Proc. Int. Conf. on Artificial Neural Networks), pp.1189–1192 (1992).

[7] 巍, 北村, 岩田, 鈴村, "CombNET–II による 1000 単語音声認識", 信学論, Vol.J75–D–II, No.4, pp.689–696 (1992).

[8] 新井正之, 王晋申, 奥田健三, 宮道壽一, "Honeycomb ネットによる多字種の手書き漢字認識", 信学論 (D–II), Vol.J–76–D–II, No.11, pp.2316–2323(1993).

[9] 依田, 根岸, 岡田, 小林, "複合 LVQ ニューラルネットワークによる印刷漢字認識方式の検討", 平成 3 年度後期情報処理学会予稿集, pp.1267 (1991).

[10] 小島, 香田, 丸野, 坂上, 山本, 〆木, "構造化ニューラルネットワーク (NARA) によるマルチフォント漢字認識", 信学技報, Vol.NC91–31 (1991).

[11] 萩原、小畑、大森, "不確定性を考慮した神経回路網モデルによる連想記憶", 第 29 回計測自動制御学会学術講演会予稿集, Vol.JS24–2, pp.263–264 (1990).

[12] Y.Hagihara, H.Kobatake and T.Omori, "Image restoration with Iterative K–nearest Neighbor Operation With the Scheduled K size(IKNOWSK)", Proceedings of the Third Australian Conference on Neural Networks, pp178–181 (1992).

[13] 斉藤泰一、山田博三、山本和彦、"JIS 第一水準手書き漢字データベース ETL9B とその解析", 信学論 (D), Vol.J–68–D, No.4, pp.757–764 (1985).

[14] 中田和夫、南敏 "音声・画像工学", 昭晃堂, pp.145–147(1987).

[15] 郭軍, 孫寧, 根本義章, 斉藤利三郎, "整形変換を用いた手書き文字データベース ETL9B の高精度認識", 信学論 (D–II), Vol.J–76–D–II, No.5, pp.1015–1022 (1993).

[16] 孫寧, 阿部正人, 根本義章, "改良方向線素特徴量および部分空間法を用いた高精度手書き認識システム", 信学論 (D–II), Vol.J–78–D–II, No.6, pp.922–930 (1995).

[17] 高木幹雄, 下田陽久, "画像解析ハンドブック", 東京大学出版会, pp577–578 (1991).

第 6 章

結論

神経回路網を用いて複雑な情報処理を行うためには、多数の細胞のパラメータやコネクションを目的に応じて設定する必要がある。本論文の目的は、多数の細胞のパラメータ等を簡単に設定できるような神経回路モデルを提案し、工学的に実現することにある。そのため、神経回路網の問題点を分析し、要素モデルとシステムモデルを提案した。多数の細胞のパラメータ等の設定を容易にするためには基本部品となる要素モデルが重要であるので、基本素子、情報フィルタ、入出力応答など、特に重要なものを個別に検討した。これらの根底にあるものは、与えられた情報の信頼性が情報の内容と共に伝達されるという点である。与えられた情報の信頼性が高ければ優先的に出力されるため、最終的な出力は雑音に対してロバストなものになる。

雑音を加えることによってロバストな応答を得るという視点で行われた研究は多い。しかし、それらは「類問を解くことによって応用力を付ける」という考え方で行われている。本論文第 3 章、第 4 章で示したのは「全く関係ないことを学ぶことによって、正しいものの見方を身に付ける」という発想である。このモデルが心理学知見と矛盾しないという点は興味深いといえる。

システムモデルを構成するに当たり、生理学的、心理学的知見と矛盾することがないよう心がけた。一方、本質的でない部分では工学的な要素を大胆に利用し、工学的実用性を持つものにした。従って、本論文で提案したモデルは、多数の細胞を制御する生理モデルとして、また、複雑な情報処理を行うニューロコンピュータとして利用することが可能である。

本研究は、修士課程、博士課程を通して継続的に行ってきた研究の集大成である。第 2 章、動的 n–k 最近接近傍法による画像の雑音除去実験では予想以上の好結果を得ることができた。一般に、画像の雑音除去では対象とする画像の性質によって必ずしも好結果を得ることができない場合が多い。対象とする画像が、模様のない壁のような画像か、布地のように細かい模様がある画像かなどにより、雑音除去のための最適パラメータも変動させなければならない。従って、一枚の画像で最適パラメータを調べても別の画像に対しては無力であると考えなければならない。しかし、本論文で提案した動的 n–k 最近接近傍法は様々な性質の画像に対して一様に好結果をもたらした。

本論文で検証した画像データの収得では、写真をTVカメラとA/Dコンバータで入力するという方法を用いている。また、重畳させた雑音は人工的なガウス雑音であり、熱雑音や電磁誘導などをシミュレートしたものであるが実際の雑音ではない。不特定なハードウェアと実際の雑音を用いて検証しない限り、本手法が直ちに工学的応用に耐え得ると言い切ることはできない。しかし、限定された実験とはいえ、一般的な雑音除去フィルタと比べて普遍的な高性能が得られたということは意外でさえあった。画像の雑音除去というのは歴史的な分野であり、パラメータに少々細工しても限界性能はそれほど高くないであろうと思われたので、期待を大きく越える成果が得られたことに満足できた。

第 3 章の基本素子モデルは修士課程での成果を発展させたものである。当初、神経回路網の中で基本素子となる単純なモデルが最も重要であると考え、これに関する研究を進めようとしたが、生体モデルに近い形のものにこだわったため、高度な信号処理に関して良好な学習結果が得られなかった。現在、他の研究機関では、生体モデルとはやや離れた形で、複素ニューラルネットモデルの研究が行われている。これらは、本論文で提案したモデルとほぼ同じものであるが、学習過程には生体モデルでなく、数学的な誤差最小化法を用いている。これらは、情報のベクトル表現にこだわる一方、生体の学習モデルを捨てたものと位置づけることができる。本研究では基本素子自体を第 4 章の確率密度関数の自己学習という形で発展的に解消させることができた。

本研究を進めた順序は、本論文の構成とは異なり、(1) 基本素子モデル、(2)n–k 最近接近傍法、(3) 大規模回路網モデル、(4) 多層パーセプトロンによる確率密度関数の自己学習という過程を辿った。(1) から (3) までの研究は、開始直後に壁にぶつかり、発展の見込みが全くつかなかった。従って、本研究の過程は発案と撤退の繰り返しであった。基本素子モデルにこだわった期間は非常に長く、この研究が役に立たなければ多くの努力が徒労だったことになる。その意味で、基本素子モデルの研究が確率密度関数の自己学習という形で発展的に解消させることができたことは有意義であったと考える。

第 4 章の確率密度関数自己学習法では、自分にとって 2 つの大きな成果を得ることができた。第一は、極めて単純な手法を用いてユニークな結果を導き出せたことである。そして第二は、確率密度関数の自己学習法を確立することにより、行き詰まりを見せていた各研究に全て日の目を見せることが可能になったことである。

確率密度関数自己学習法は「全く関係ないことを学ぶことによって、正しいものの見方を身に付ける」という発想に基づいている。人間が何かを記憶するとき、記憶対象のデータのみ与えられているとは考えられない。多くの雑音も記憶対象のデータと共に与えられているはずである。例えば、ある人の顔を記憶するときも、顔という情報だけでなく、背景など様々な雑音が与えられている。これらの雑音を抑制するという操作が何らかの形で行われなければならない。

確率密度関数自己学習法はこのような操作を導入したものである。学習の方法は極めて簡単であり、プログラム作成の手間もほとんど必要ない。人間が知識を得るとき、必要な情報だけを与えられているなら、その範囲内だけしか考えが及ばないのは当然である。もし、全く別の情報を与えられても、無理にその範囲内から答を引きだそうとするだろう。様々な雑

音を与えられることによって、無理に答を引き出すのではなく確実なものだけを取り出すことができるようになる。一見、不必要と思われるようなことを学ぶことも、決して無駄というわけでなく、思わぬ効用があるものだというのはいささか言い過ぎであろうか。いずれにせよ、このモデルが心理学知見と矛盾しないという点は興味深い。

確率密度関数自己学習法では、雑音が多すぎると出力させる確率密度が低くなり悪影響を及ぼす。雑音が与えられる回数が非常に多いときには出力のノルムが小さくなり、パターンの分離が困難になるからである。また、単位時間当たりの学習回数が決まっている場合、サンプルデータを与えられる機会が相対的に少なくなるため学習速度が低下する。これも、人間社会では経験的に知られている通りである。遊びもほどほどにといったところだろう。

教師データのノルムを 0 としてランダムな値を入力する方法は、基本素子及び大規模回路網の研究過程で苦し紛れに用いていた方法である。この方法を用いれば、既知のサンプルでベクトルを大きく、未知サンプルではベクトルが小さくなる。しかし、既知サンプルと同じ入力値がランダムデータとして与えられるとき、教師データのノルムが 0 または 1 という 2 種類の値になってしまう。このような形で教師データを与える方法は従来の手法にないため、正しい学習をする可能性は低いであろうと考えていた。

この方法で学習を行ったときの出力から入力の確率密度関数が導き出せることに気づいたことは、本研究の最も大きな成果の 1 つといえる。パターン認識は、入力ベクトルの確率密度関数を求めることによって容易に行えるためである。本研究で提案した手法により、神経回路網によるパターン認識の可能性が大きく広がる。本手法を発展させたいくつかのアイデアを実現する機会が得られることを希望する。

第 5 章のような大規模な神経回路網の研究は工学的な有効性が高いため、文字認識、音声認識の研究の一環として行われる場合が多い。これらは、一般に並列演算が可能な高速計算機を用いて計算を行う。ワークステーションでシミュレーションを行うにはいささか無理があり、本研究では単純な実験を通して、教育漢字 950 文字の分類など必要最小限の有効性を確かめるにとどまった。高速計算機を用いた実験には、JIS 漢字（第一水準）を全て記憶させたものもある。このような研究では認識モジュールのつながりに工夫をするものが多いが、本研究ではモジュール構成にまで立ち入り、高精度化を目指した。従って、本研究の成果をパターン認識等に用いることにより高い効果が期待できる。音声、画像認識など、様々なアプリケーションで本手法の有効性を確認することも興味深い研究対象である。

以下に、本論文で得られた主要な結論を述べる。

k 最近接近傍フィルタを取り上げ、神経回路網のシミュレーテッドアニーリングを利用してフィルタを高性能化する手法について述べた。画像ノイズの除去にこの手法を応用したとき、従来のフィルタより雑音除去効果が大きいことを確認した。

情報の内容と共にその情報の確率密度を盛り込むことができるモジュールが、大規模神経回路網を構成する際の基本単位になることを示した。神経回路網の中で基本素子となる単純なモデルを提案し、このモデルがモジュールと同様の働きをする事を述べた。

入力に乱数を与えることによって、確率密度が未知の場合も確率密度関数の学習が可能であることを示した。多層パーセプトロンに排他的論理和を学習させたとき、本手法が有効で

あることを確認した。

各要素技術を組み合わせた神経回路網を実際の構成し、手書き文字認識および画像の雑音除去の実験を行って有効性を示した。

以上により、多数の細胞を簡単に扱える神経回路モデルのひとつを提案し、またこのモデルの工学的な有効性を確認できた。

謝辞

本研究を行い、博士論文をまとめるに当たり、東京農工大学工学部の小畑秀文教授、大森隆司助教授には、指導教官・副指導教官として終始変わらぬ御指導、御鞭撻を賜りました。また、東京農工大学工学部の伊東正安教授、小林駿介教授、山田晃助教授、篠田裕之講師には、博士論文審査委員として、具体的なご指導をいただきました。諸先生に深く感謝いたします。

小畑教授をはじめ電子情報工学科の先生方には、未熟者である筆者に対し、長い間有形無形の様々な御指導を賜りました。東京農工大学電子情報工学科に入学しなければ、私は博士過程に進学することすら考えなかったのではないかと思います。研究者として、教育者として、また、人生の先輩として、ご相談にお答えいただいた、電子情報工学科の諸先生に重ねて御礼申し上げます。

本研究を行うにあったって共同研究者である森井茂夫氏（現 IHI)、佐次洋作氏（現住友電工)、には大きな御協力を戴きました。

森井茂夫氏の研究は確率密度関数を学習する MLP の研究に、佐次洋作氏の研究は動的 n-k 最近接近傍フィルタの研究にそれぞれ開花させることができました。

また、大森研究室の広池敦氏 (現日立製作所) には博士論文執筆に関する情報を様々な形で戴きました。

現在、各方面で御活躍の諸氏に深く感謝いたします。

そして、電気通信大学の渡邉由香里氏には、資料収集や博士論文編集にご協力をいただきました。忙しい中、乱暴な依頼にも誠実に対処していただき、大きな心の支えを得ることができました。

ほかにも、名前をあげて感謝したい先生、同期、先輩、後輩はたくさんいらっしゃいますが、本論文はここで終わりにいたします。非力な私が本研究をなんとか形にすることができたのは、ひとえにみなさまのご支援、ご協力の賜物と考えます。御協力戴いたみなさまに心より感謝申し上げます。ありがとうございました。