

福井大学審査
学位論文 [博士(工学)]

ニューラルネットワークによる
パターン認識に関する研究

平成22年3月

竹谷 尚

内容梗概

ニューラルネットワーク (NN) は、学習で獲得した知識を活用する生体の脳をモデルとしたものである。NN は、優れたパターン認識能力を持ち、学習により入出力間の関係をニューロン間の結合重みとして蓄える能力を有する。その情報処理過程は、従来のパターン認識のそれとは異なる特徴を利用しており、パターン認識の分野にとどまらず、組み合わせ最適化問題など多くの分野にわたり研究がなされている。

本論文は、大学在籍中および津山高等専門学校に在職中に行った NN の研究を纏めたものである。すなわち、(1)NN の課題である汎化能力の強化に関するパターン位置の抽出と正規化の研究、(2)NN のパターン認識への応用として、(I)3次元血管内超音波内視鏡を指し多素子超音波トランスデューサによる水中物体の画像化、(II)IC ウェハの製造工程で起こる IC 表面の切り粉汚れの目視検査の自動化、(III) 膝蓋骨亜脱臼の自動識別および (IV) 囲碁印刷総棋譜の自動読み取り、について述べたものであり、全体は、7章からなる。

第1章では、NN によるパターン認識に関して問題の所在を明らかにし、本論文の位置づけと意義について述べる。

第2章では、位置ずれを伴うパターンの正規化手法の提案を行う。文字認識などを行う場合、位置ずれのあるパターンに対し前処理としてこのずれを補正しておくこと NN の認識能力は強化される。本章では、ランダムパターンの学習を行う実験を行い、学習ができない原因を明らかにする。次にその問題点を2つのネットワークを組み合わせることで解決できることを示す。

第3章では、3次元血管内超音波内視鏡の基礎実験として研究グループで試作した、円形単板圧電素子を36分割した多素子超音波トランスデューサによる水中物体の画像化手法を提案する。ネットワークの入力には受信した超音波エコーのエコーピーク時間を用い、3次元形状をもつ未知パターンに対して本手法の有効性を確認した。その際、ネットワークの構成方法に関して情報量基準による評価を行い、広い範囲の認識・再構成するための最小の学習用基本パターンについて検討を行っている。

第4章では、IC ウェハの製造工程で起こる IC 表面の切り粉汚れの目視検査の自動化について扱う。まず、マルチスペクトル光学フィルターと CCD カメラを使用して、異なる照明条件で IC ウェハの画像を撮影し、データの特徴空間を作成する。これらについて BP によるフィードフォワードニューラルネットワーク、最近傍抽出に適した光学波長について調査した。結果として、特徴要素を区別する効果的な画像の取得方法を開発した。

第5章では、ニューラルネットワークによる膝蓋骨亜脱臼の自動識別について述べる。臨床診断においては、医用画像の読影診断のため、医師に対して長期にわたる教育と訓練が必要となる。したがって、画像診断の自動化に大きな期待が懸けられている。ここでは、確定診断つき膝の CT 画像を対象としてニューラルネットワークによる膝蓋骨亜脱臼の判定を行っている。その結果、膝蓋骨亜脱臼のニューラルネットワークによる自動診断の可能性を示せたものと考えている。

第6章では、低品質文字の認識に関するものである。印刷物の囲碁総棋譜は一般的に200手以上あって、小さいページ領域中に小さい文字で記録されている。数字順の棋譜読み取りはかなり手間がかかるだけでなく、ゲーム進行の臨場感もない。囲碁総棋譜の自動読み取りとそれを利用したパソコンによる再現は、囲碁愛好者に強く望まれるところである。また、自動読み取りにより膨大なデータのデータベース化が図れるが、その必要性・価値にも関わらず、自動読み取りに関する研究はなされていない。ここでは、スキャナーで読み込んだ総棋譜の前処理後、FPM(Fuzzy Partition Model) ユニットおよびメッシュ特徴の利用による数字認識率の向上と学習速度の高速化を図っている。

第7章は、2章から6章までの研究で得られた結論を纏め、本研究の位置付けと意義について述べている。

目次

第 1 章	緒論	1
1.1	パターン位置の抽出と正規化	1
1.2	多素子超音波トランスデューサを用いた反射波による物体の 3 次元画像化	2
1.3	パターン認識への応用	4
1.3.1	IC ウェハの切り粉汚れ判定と撮影条件の決定	4
1.3.2	ニューラルネットワークによる膝蓋骨亜脱臼の自動識別	4
1.3.3	低品質文字の認識	5
第 2 章	ニューラルネットワークによるパターン位置の抽出と正規化	9
2.1	緒言	9
2.2	3 層 BP ネットによる位置ずれ学習	10
2.2.1	任意ランダムパターン信号の位置ずれ学習	10
2.2.2	あいまい性	10
2.2.3	位置信号の挿入	12
2.3	位置信号の抽出とパターン正規化	14
2.3.1	位置信号抽出 NET (NET1)	14
2.3.2	二つのネットワークの組合せによる位置の正規化	14
2.4	2 次元への拡張	16
2.4.1	2 次元の位置ずれ学習ネットワーク (位置信号 2 次元化法)	16
2.4.2	位置信号を二つの 1 次元信号に分解した構成 (位置信号 1 次元化法)	16
2.4.3	1 次元正規化ネットの 2 段縦続構成 (1 次元ネット 2 段縦続法)	20
2.4.4	実験結果	20
2.5	結言	22
第 3 章	円形単板圧電素子を分割した多素子超音波トランスデューサによる水中物体の 3 次元画像化	24
3.1	緒言	24
3.2	3 次元物体画像化システム	25
3.2.1	画像化対象	26
3.2.2	受信エコーと前処理	28
3.3	超音波信号の 3 次元画像化	28
3.3.1	観測モデル	28
3.3.2	ニューラルネットワークによる超音波信号の画像化学習	29
3.3.3	未知ターゲットの画像化	32
3.4	特徴の縮約化	34
3.5	考察	35

3.6	結言	36
第4章	ニューラルネットワークによるICウェハーの切り粉汚れ判定と撮影条件の決定	41
4.1	緒言	41
4.2	撮影条件と特徴抽出	42
4.2.1	狭帯域干渉フィルタ	42
4.2.2	照明条件	42
4.2.3	データ抽出と特徴抽出	43
4.3	分類手法	46
4.3.1	ニューラルネットワークによる分類	47
4.3.2	最短距離法	47
4.3.3	最尤法	48
4.4	実験結果	48
4.4.1	分類手法の評価	49
4.4.2	分類に有効な特徴量	50
4.5	結言	51
第5章	ニューラルネットワークによる膝蓋骨亜脱臼の自動識別	57
5.1	緒言	57
5.2	対象と画像処理方法	57
5.3	神経回路モデル(ニューラルネットワーク)	60
5.4	認識実験	61
5.5	結言	62
第6章	PDPモデルによる印刷囲碁総棋譜の自動読み取り	64
6.1	緒言	64
6.2	棋譜数字抽出	65
6.2.1	交点座標および黒石白石の抽出	65
6.2.2	相関法による数字認識	66
6.2.3	数字切り出しとラベリング	67
6.3	FPMを用いた数字認識	68
6.3.1	FPM	69
6.3.2	メッシュ特徴とFPMの構成	70
6.3.3	学習速度比較	71
6.3.4	数字認識	71
6.4	結言	72
第7章	結論	74
付録A	表4.2の画像例(汚れなし)	77
付録B	表4.2の画像例(汚れあり)	81

第1章 緒論

現代社会において多種多様な情報が流通し、その情報を処理する機器およびソフトウェアの発達には目を見張るものがある。現在利用されている情報処理機器のほぼすべてがプログラム内蔵方式であり、目的に応じた手順をプログラムとして記述することにより目的を達成している。したがって処理内容は論理的に明確にされて始めて、プログラム化することが可能となり、パターン認識の分野でも様々な課題に対して、様々なアルゴリズムが検討されてきた。しかし現代のコンピュータ技術の発展にも関わらず、機械によるパターン認識能力は人間のそれと比較すると格段のへだたりがある。

ニューラルネットワーク (Neural Network) は、本来人間をはじめとする生態の神経細胞 (ニューロン) の結合により構成される回路網であり、ニューロン間の結合を教師あり学習により修正する、適応型処理として研究者達に注目された。その後、Minsky ら [1] によりパーセプトロン型のニューラルネットワークの限界が証明され、第1次のニューラルネットワークブームの終焉となった。1980年代に入り、多くのネットワークモデルが提案され、ニューラルネットワークが再び注目されるようになった。その中でも

- (1) 多層パーセプトロン型ネットワークの学習法として最急降下法を多層に拡張したバックプロパゲーションアルゴリズム [2] の発表により、多層ニューラルネットワークが線形分離不可能な XOR 問題などの非線形識別問題を扱えるようになったこと、
- (2) 相互結合型であるホップフィールド型ニューラルネットワーク [3] がエネルギー関数を最小化することにより多くの最適化問題の近似解を実時間で与えること、

が大きな成果として受け入れられ、研究の対象となった。バックプロパゲーションアルゴリズムによる階層型ニューラルネットワークは入力ベクトル空間から出力ベクトル空間への写像を学習により獲得することができる。ニューラルネットワークの学習能力により、論理的な記述が困難なパターン認識などの分野でこれにより論理的な処理が困難なパターン認識などの分野で多くの成果を挙げている。加えてニューラルネットワークは従来のソフトウェア開発コストを減少させることが可能な情報処理のアプローチである。その適用分野は、パターン認識にとどまらず、ホップフィールド型ニューラルネットワークを使った組み合わせ最適化問題など多くの研究がなされており、その果たす役割はますます拡大している。ニューラルネットワークによるパターン認識での課題は、学習による知識の獲得と汎化能力の強化である。計算能力に優れたコンピュータの最も不得意とするのがパターン認識の分野である。本論文はパターン認識に関する研究を纏めたもので、(1) パターン位置の抽出と正規化、(2) 超音波反射波からの3次元再構成、(3) パターン認識への応用から構成される。

1.1 パターン位置の抽出と正規化

同一パターンでも位置ずれがあると、コンピュータでは異なったパターンと認識される。パターンの位置ずれ対策は大きな課題である。パターン認識の際、位置ずれのあるパターンに対し、前処理としてこのずれを補正しておくこととBPNの認識・連想能力は強化される。本章ではニューラルネットワークによるパターン位置の正規化について検討する。

単にパターン位置を正規化するだけならば、パターンの重心を求め、それをもとにシフト回路などの固定的回路で位置シフトを行うことにより位置正規化は行える。しかしながら実際のパターンデータは、雑音やパターンひずみを生じているため種々の対処策が必要になる。ニューラルネットワークによる入力パターンの位置ずれの解決策としては、

- (1) 位置ずれをしたパターンをも含めた形で学習する。
- (2) ノイズを添付したデータ若しくはガウシフィルタなどでぼかしたデータで学習を行う。
- (3) 必ずしも学習によらずに位置ずれを抽出・補正する人為的メカニズムを開発する。

などが考えられる。

(1) の場合、パターンが増えたとき必要な学習パターンが膨大になる問題点があり、従来研究の対象とされなかった。(2) は、簡単には識別空間上で識別面が緩やかに形成されることになり、これによって認識が可能になると解釈されている。これは、位置ずれというよりパターンの変形に有効である。(3) の場合については、Fukushima [4] は特徴の位置ずれを受容野内で許容することにより入力文字の位置ずれ、変形や大きさの違いに対処した。また B. Widrow は、1960 年に提案した Adaline を基本素子とした正規化システム Madaline を提案している [5]。このシステムは、多数の Adaline(Adaptive Linear Neuron) を用いたニューラルネットである。パターンの正規化回路は複数の Adaline を並列に並べ (Slab と呼ばれる)、MAJ と呼ばれる多数決素子と連結することにより構成される。ここでは上下左右や斜方向の位置ずれ、および回転に対して正規化された出力パターンが得られる。また、H. Yang [6] は BP 多層ネットによるパターンの回転について実験を行っている。文献 [7] には位置や形状正規化を行ういくつかの回路網が示されている。しかしこのようなネットワークでは構成が大規模になってしまう問題点がある。また、いずれも人為的に設計された正規化回路である。

そこで、パターン位置の抽出と正規化をニューラルネットワークで実行することにより、以下のようなメリットが期待できる。

- (1) 均一な構成で実現でき、将来汎用的なニューラルネットワークがチップ化されるような場合に効果がある。
- (2) ニューラルネットワークにより位置正規化の普遍的性質が抽出される可能性がある。
- (3) 他の正規化(回転、大きさ、ひずみ)や、より複雑な正規化、またそれらの組合せが学習済み重みを入れ換えるだけで同一回路で処理できる。

第 2 章では 3 層で位置ずれパターンに対して学習できるネットワークの一つの構成法を提案する [8]。まずランダムパターンの位置ずれを補正するネットワークの学習実験を行い学習できない原因を明らかにする。次にその問題点、を二つのネットワークを組み合わせることによって解決し、学習により正規化回路を実現できることを示す。更にその際、重み分布に周波数展開的な性質が出てくることを示す。最後に 1 次元を 2 次元に拡張する三つの手法について示し、それらの比較検討を行った。

1.2 多素子超音波トランスデューサを用いた反射波による物体の 3 次元画像化

トランスデューサによる水中物体の 3 次元画像化について、一般に超音波は、X 線や MRI 画像に比べて分解能が悪いが、非観血的・非侵襲的に体内の組織・器官の状態をリアルタイムに映し出せることから、臨床では胎児診断や心臓内の血流分布などの観察・計測によく用いられている。最近問題になっている成人病の原因の 1

つに動脈硬化がある．この治療法には，大きく分けて内科的および外科的の2つがある．本研究を始めた目的は，このうちの外科的治療に属するバルーン血管形成術において，心臓および血管内に置いたトランスデューサの前方対象物の画像化を行うことである．そのため超小型トランスデューサおよびそのデータ処理法を開発し，最終的には血管内に挿入して3次元な超音波内視鏡としての臨床応用を行うことが最終目標である．固定した超音波アレイセンサーを用いて前方物体の画像化を行う研究には次のようなものがある．

- (1) 1次元フェーズドアレイを2次元に拡張したもの；すなわち格子状トランスデューサアレイの位相制御により細い走査ビームを得るもの [9–11] ．
- (2) 指向性をもったトランスデューサを平面状に並べることにより，画像化を行うのみ；たとえば格子状配置指向性トランスデューサアレイの個々のエコー遅延時間データ (3次元画像) をニューラルネットワーク処理により修復画像化する方法 [12] ．
- (3) 音響ホログラム手法；これには同心円状に配置したトランスデューサを直交関数で駆動・受信し，相関を利用して画像化を行うもの [13, 14] ，同様のトランスデューサを用いて，球面状反射波を考慮した一種の時空間的逆投影法により3次元物体形状の復元を試みたもの [15] ，固定した超音波源から放射された超音波の反射波面をトランスデューサアレイで観測し，逆フーリエ変換手法により粗い物体像を得た後，ニューラルネットワーク処理により，修復するもの [16] ．

など，が提案されている．

しかしながら，これらについては次のような問題がある．(1) は2次元走査に時間を要し，動物体の実時間観測に不利である．(2) は空間的に分離された超音波ビームの独立性を利用して画像化を行うので，処理自体は簡単であるが，小型化が難しい．(3) は複雑な処理を必要とする．またこれらの手法は同一のトランスデューサ素子を並べることを一応の前提としている．そのため，単板分割により得た形状が異なる (したがって指向性などの特性が異なる) 複数のトランスデューサ素子を用いた場合，その補正が必要となってくる．ただし，ニューラルネットワークを利用するものについては，その段階で差異を吸収できる可能性がある．本研究では，そのような柔軟性・適応性に加えて，実時間処理，小型化および処理の簡略化を目指して，ニューラルネットワークを利用することにした．

従来法であるプログラム計算方式は，達成される処理が既知の手順あるいは既知の規則セットの形で記述できる場合のみ使うことが出来る．現行のコンピュータは全て論理的に基づいて動作しており，ソフトウェア開発には，設計，試験，改良の徹底的な繰り返しが要求され，ソフトウェア開発は時間と費用のかかるものとなる．これに対して，ニューラルネットワークはアルゴリズムや規則の開発を必要とせず，ソフトウェアの量を大幅に減らすことが可能な新しい情報処理のアプローチである．

第3章では，3次元血管内超音波内視鏡の基礎実験として円形単板圧電素子を36分割した多素子超音波トランスデューサによる水中物体の画像化手法を提案する [17] ．画像化には階層型ニューラルネットワークを用い，バックプロパゲーションアルゴリズムにより学習を行った．ネットワークの入力は受信した超音波エコーのエコーピーク時間を用いた．3次元形状をもつ未知パターンに対して画像化実験を行い，本手法の有効性を確認した．その際，ネットワークの構成方法に関して情報量基準による評価を行い，与えられた学習パターンセットに対する入力パターンベクトルの圧縮方の検討を行った．最後に本システムを拡張したより実用的な3次元超音波画像化システムについて検討を行った．

1.3 パターン認識への応用

1.3.1 IC ウェハの切り粉汚れ判定と撮影条件の決定

一般に IC ウェハの汚れの有無をコンピュータにより判定することはパターンの多様性からかなり難しい問題である。第 4 章では、IC ウェハの製造工程で洗浄を必要とするかどうかを決定するために切り粉汚れの有無をニューラルネットワークを用いて判定する。工業生産において、品質管理は最終製品の信頼性を保障するために、正確な検査と歩留まりのマネジメントは非常に重要である。多くの場合、製品の目視検査は不可欠であるが、人間による検査はスピードが遅く、高価であり、ミスも多いため、目視検査の自動化 [18, 19] は望まれている。IC の製造工程で起こる汚れには洗浄不足による切り粉汚れと異物の混入によるものがある。不完全なパターンや印刷工程に欠陥を引き起こす IC ウェハの顕微鏡検査に関する研究には文献 [20, 21] がある。しかし、ウェハの切り粉汚れの検査については生産にかなり影響があり、IC 製造 [22] において検出しなければならない。

一方、照明条件について一般に、画像処理および画像認識のためには、できるだけきれいな入力画像を撮影する必要がある。特に、IC ウェハの配線部を位置合わせ用テンプレートとして使用する場合には、アルミ表面のランダムな凹凸(ヒロック)による影響を受け易い。これらは一般的に画像中では雑音となり、適合位置においてもテンプレートと入力画像の相関値(Q 値)が低く、位置合わせに失敗する確率を大きくする。しかしながら、適切な照明を与えると、雑音が消え、大きな適合相関値を得ることが知られている。

ウェハの切り粉汚れは IC ウェハ配線部と関連しており目視検査の自動化のためには、その定量化の評価基準とともに照明条件を決定しなければならない。そこでニューラルネットワークを利用することにした。ここで、本研究での第一の目標は、汚れの度合いを計測することではなく、ウェハの汚れの有無を判断することであるため、切り粉汚れの有無についての分類手法を対象とする。本研究では切り粉汚れの有無のウェハのサンプルが入手可能であったので、本研究では教師あり学習による分類を選択することとした [23]。すなわち、ニューラルネットワークの学習で獲得した知識を活用する優れたパターン認識能力を利用する。

ここでは 3 つの従来の分類手法、バックプロパゲーションアルゴリズムによるフィードフォワードニューラルネットワーク、最近傍法、最尤法について評価し、それぞれのパフォーマンスについて比較を行った [24]。加えて、特徴空間の重要な要素として適切な光学的な周波数を調査した。結果として、特徴要素を区別する効果的な方法を画像の取得方法を開発した。また、バックプロパゲーションニューラルネットワークが IC ウェハの汚れの有無の判断に有効であることを示す。

1.3.2 ニューラルネットワークによる膝蓋骨亜脱臼の自動識別

膝蓋骨亜脱臼の診断においては関節造影 CT, MRI などを用いて、医師の読影による膝蓋骨損傷の診断が試みられている [25, 26]。図 1.1 は、関節造影 CT 画像の例である。左膝屈曲角度 15° の状態で蓋骨にほぼ垂直に膝関節を撮影されたものである。(a) においては、蓋骨が膝の関節のほぼ中央に位置し正常である。(b) においては、蓋骨が膝の関節の中央の位置からずれており異常である。膝蓋骨亜脱臼の診断において医師は、患者の膝の外形から診断できるが、コンピュータによる正常および異常の診断が出来ればその利便性は大きい。図 1.1 において、膝蓋骨亜脱臼の目視による判定は容易に行えるが、コンピュータで行う場合、その評価基準をどのように決定するかが困難になる。例えば、膝蓋骨に個人差があるばかりでなくどこまで正常・異常の判定をするかである。

第 5 章の目的は、膝蓋骨亜脱臼のスクリーニング検査をレーザなどで計測した膝関節外形形状(図 1.1(c),(d))から行うことを考え、80%程度の識別率を目指して行う計画でいたが、外形形状のみからでは識別率が 50%程

に止まった．そこで3層バックプロパゲーション・ニューラルネットワークを用いて，膝のCT画像より骨領域を抽出した2値画像およびエッジ画像からの膝蓋骨亜脱臼画像の自動診断について検討を行った [27]．

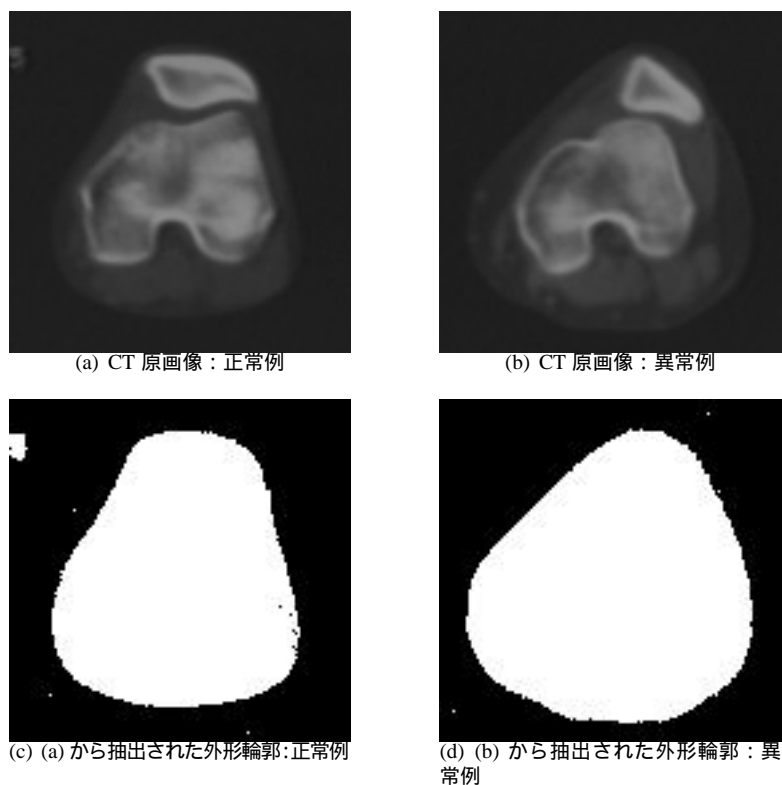


図 1.1: 膝蓋骨亜脱臼診断

1.3.3 低品質文字の認識

印刷物の囲碁総棋譜は一般的に200手以上あって，小さいページ領域中に小さい文字で記録され，読み取り(棋譜理解)はかなり手間がかかるだけでなく，ゲーム進行の臨場感もない．また長年に亘り多く蓄積・記録されている囲碁総棋譜の自動読み取りとそれを利用したパソコンによる一手毎の再現は，囲碁愛好者に強く望まれるところである．しかし，その再現と棋譜データベース化の必要性に関わらずこれまで囲碁総棋譜の自動読み取りに関する研究はなされていない．

棋譜の自動読み取りについては直感的には，すべての手順・白黒石の数字画像をテンプレートとして作成・保存を行い，従来法の相関法で棋譜・交点画像データとの相関値で認識を行う手法が挙げられる．しかし実際の処理を行うと，テンプレート作成画像では，100%の認識率が得られるが，同じ雑誌でも他の棋譜に対して行うと認識率は数%であった．数字のみ取り出して，ボカシ処理も含めて同様の処理を行ったが，数字認識率は50~75%程度であった．

第6章では，囲碁総棋譜の自動読み取りにおいてPDP(Parallel Distributed Processing)モデルによる数字認識を用いることを検討した [28]．文字認識は，(1)文字抽出，(2)認識手法，に分類される．(1)については前処理として，2値化・ラベリング・連結数字の分離・数字の正規化・細線化・メッシュ特徴の利用を行う．(2)につ

いて、FPM(Fuzzy Patition Model) では、評価基準として Kull-back ダイバージェンスを用いユニットの総出力和の正規化を行う。また、教師ベクトルを $(0, \dots, 0, 1, 0, \dots, 0)$ のように 0 と 1 のみで構成することにより、出力層の FPM ユニットに現れる相互抑制を利用する。出力以外は、お互いに抑制し合うことで収束性・認識率の向上を行う。

参考文献

- [1] Minsky M. and Papert S. *Perceptrons*. MIT Press, Cambridge, MA, 1969.
- [2] Rumelhart D.E, McClelland J. L. *Parallel Distributed Processing*. MIT Press, Cambridge, MA, 1987.
- [3] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.*, Vol. 2, pp. 2554–2558, 1982.
- [4] Fukushima K. A neural network for visual pattern recognition. *IEEE Computer*, Vol. 21, No. 3, pp. 65–75, 1988.
- [5] Widrow B. and Winter R. Neural nets for adaptive filtering and adaptive pattern recognition. *IEEE Computer*, Vol. 21, No. 3, pp. 25–39, 1988.
- [6] Yang H. and Guest C. C. Performance of back propagation for rotation invariant pattern recognition. *Proc. IEEE First International Conference on Neural Networks*, pp. IV–365–370, 1987.
- [7] Treisman A. *Visual-cognitive neural networks*. MIT Press, 1987.
- [8] 竹谷, 光本, 田村, 河合, 岡崎, 副井. ユーラルネットワークによるパターン位置の抽出と正規化. 電子情報通信学会論文誌 D-II, Vol. J75-D-II, No. 7, pp. 1260–1270, 1992.
- [9] Pilkington TC and Ideker RE. von Ramm OT. National science foundation/engineering research center of emerging cardiovascular technologies. *Proceedings of the IEEE*, Vol. 81, No. 1, pp. 79–94, 1993.
- [10] Omer Oralkan, A.Sanli Ergun, and etc. Ching-Hsiang Cheng. Volumetric ultrasound imaging using 2-d cmu arrays. *IEEE Tras. ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL*, Vol. 50, No. 11, pp. 1581–1594, 2003.
- [11] Matthew P.Fronheiser and Edward D.Light et al. Real-time, 3-d ultrasound with multiple transducer arrays. *IEEE Tras. ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL*, Vol. 53, No. 1, pp. 100–105, 2006.
- [12] 吉澤信幸. 超音波センサとニューラルネットワークを用いた三次元物体認識. *EMC on.*, Vol. 53, pp. 98–104, 1992.
- [13] 湯浅肇, 石原知明, 高野宰, 他. 波面符合化送信方式による水中撮像装置. 信学技法, Vol. US 94, No. 7, pp. 45–52, 1994.
- [14] 田村安孝. 計算機処理による超音波イメージング. *Med. Imag. Tech*, Vol. 17, No. 1, 1999.
- [15] M.Nambu, M.Do, M.Matani, O.Oshiro, and K.Chihara. A high-speed image acquisition using ultrasonic ring array probe. *Computers in Cardiology*, Vol. 26, pp. 355–358, 1999.

- [16] Watanabe S and Yoneyama M. An ultrasonic visual sensor for three-dimensional object recognition using neural networks. *IEEE Trans Robotics Autom*, Vol. 8, No. 2, pp. 240–249, 1992.
- [17] 竹谷尚, 大田和寛, 田村進一, 岡崎耕三. 円形単板圧電素子を分割した多素子超音波トランスデューサとニューラルネットワークによる水中物体の3次元画像化. *情報処理学会論文誌*, Vol. 48, No. SIG9, pp. 57–66, 2007.
- [18] R.T. Chin and C.A. Harlow. Automated visual inspection: a survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-4, No. 6, pp. 557–573, 1982.
- [19] T.S. Newman and A.K. Jain. A survey of automated visual inspection. *Comput. Vis. Image Understanding*, Vol. 61, pp. 231–262, Nov. 1995.
- [20] X.L. Xie and G. Beni. A validity measure for fuzzy clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 841–847, Nov. 1990.
- [21] M. Nikoonahad, C.E. Wayman, and S.A. Biellak. Defect detection algorithm for wafer inspection based on laser scanning. *IEEE Trans. on Semiconductor Manufacturing*, Vol. 10, No. 4, pp. 459–468, Nov. 1997.
- [22] San Jose. *The National Technology Road-map for Semiconductors*. Semiconductor Industry Association, 1994.
- [23] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, 1996.
- [24] Reza A. Zoroofi, Hisashi Taketani, Shinichi Tamura, Yoshinobu Sato, and Kazuma Sekiya. Automated inspection of ic wafer contamination. *Pattern Recognition*, Vol. 34, pp. 1307–1317, 2001.
- [25] Inoue M., Shino K., and Hirose H. et al. Subluxation of the patella. *Computed Tomography Analysis of Patellofemoral Congruence*, Vol. 70-A, pp. 1331–1337, 1988.
- [26] 中西克之, 井上雅裕, 原田貢士, 他. 膝蓋骨亜脱臼症候群のMRI-膝蓋関節軟骨の病的変化の検討-. *日本医放会誌*, Vol. 51, pp. 387–393, 1991.
- [27] 呉, 井上, 竹谷, 田村, 小野. ニューラルネットワークによる膝蓋骨亜脱臼の自動識別. *Medical Imaging Technology*, Vol. 9, No. 4, pp. 460–463, 1991.
- [28] Hisashi Taketani, Mohd Syafiq Suhaimi, Wang Rong Long, Kozo Okazaki, Kyohei Shimada, and Shinichi Tamura. Auto-reading of IGO ALL-Record by PDP Model. *International symposium on Robotics and Interigent Sensor*, 投稿中, , 2010.

第2章 ニューラルネットワークによるパターン位置の抽出と正規化

2.1 緒言

ニューラルネットワーク (NN: 神経回路網) は、学習で獲得した知識を活用する人間の脳を模擬して作られた。その情報処理過程は、従来のパターン認識のそれとは異なる特徴を利用している。その応用は、パターン認識の分野で文字認識をはじめ幅広く利用され、その効果が報告されている。その課題となっているのは、汎化能力の強化と高速化である。パターン認識の際、位置ずれのあるパターンに対し、前処理としてこのずれを補正しておくことと BPN の認識・連想能力は強化される。本章ではニューラルネットワークによるパターン位置の正規化について検討する。

しかし、この問題を歴史的に見ると、パーセプトロンでは、位置ずれしたパターンの認識は受容野や層数に応じた制約を受けることが Minsky らによって明らかにされた [1]。これらのことが一部では能力の限界を示すものとして否定的に受け取られたが、一般のニューラルネットによる位置ずれ補正を否定したものではない。実際、究極のモデルとしての人間はいろいろな位置に書かれた文字を認識することができ、また運動に伴って視覚系入力が大きく変動するにもかかわらず、静止環境を知覚できる。更に田村らは雑音中の未知波形系列に対し、その統計的性質を利用すると、同期化可能従って位置補正が可能であることを示している [2]。加えて3層ないし4層ネットワークではそれぞれの任意の論理関数および連続関数を実現できることが示されている。これらのことは階層型ニューラルネットによっても任意パターンの位置の正規化を行える可能性が十分あることを示している。

単にパターン位置を正規化するだけならば、パターンの重心を求め、それをもとにシフト回路などの固定的回路で位置シフトを行うことにより位置正規化は行える。しかしながら実際のパターンデータは、雑音やパターンひずみを生じているため種々の対処策が必要になる。ニューラルネットワークによる入力パターンの位置ずれの解決策としては、

- (1) 位置ずれをしたパターンをも含めた形で学習する。
- (2) ノイズを添付したデータ若しくはガウスフィルタなどでぼかしたデータで学習を行う。
- (3) 必ずしも学習によらずに位置ずれを抽出・補正する人為的メカニズムを開発する。

などが考えられる。

(1) の場合、パターンが増えたとき必要な学習パターンが膨大になる問題点があり、従来研究の対象とされなかった。(2) は、簡単には識別空間上で識別面が緩やかに形成されることになり、これによって認識が可能になると解釈されている。これは、位置ずれというよりパターンの変形に有効である。(3) の場合については、Fukushima [3] は特徴の位置ずれを受容野内で許容することにより入力文字の位置ずれ、変形や大きさの違いに対処した。また B. Widrow は、1960 年に提案した Adaline を基本素子とした正規化システム Madaline を提案している [4]。このシステムは、多数の Adaline (Adaptive Linear Neuron) を用いたニューラルネットである。パターンの正規化回路は複数の Adaline を並列に並べ (Slab と呼ばれる)、MAJ と呼ばれる多数決素子と連結する

ことにより構成される．ここでは上下左右や斜方向の位置ずれ，および回転に対して正規化された出力パターンが得られる．また，H, Yang [5] は BP 多層ネットによるパターンの回転について実験を行っている．文献 [6] には位置や形状正規化を行ういくつかの回路網が示されている．しかしこのようなネットワークでは構成が大規模になってしまう問題点がある．また，いずれも人為的に設計された正規化回路である．

そこで，パターン位置の抽出と正規化をニューラルネットワークで実行することにより，以下のようなメリットが期待できる．

- (1) 均一な構成で実現でき，将来汎用的なニューラルネットワークがチップ化されるような場合に効果がある．
- (2) ニューラルネットワークにより位置正規化の普遍的性質が抽出される可能性がある．
- (3) 他の正規化 (回転，大きさ，ひずみ) や，より複雑な正規化，またそれらの組合せが学習済み重みを入れ換えるだけで同一回路で処理できる．

本章では 3 層で位置ずれパターンに対して学習できるネットワークの一つの構成法を提案する [7]．まずランダムパターンの位置ずれを補正するネットワークの学習実験を行い学習できない原因を明らかにする．次にその問題点，を二つのネットワークを組み合わせることによって解決し，学習により正規化回路を実現できることを示す．更にその際，重み分布に周波数展開的な性質が出てくることを示す．最後に 1 次元を 2 次元に拡張する三つの手法について示し，それらの比較検討を行った．

その結果，パターンおよび位置信号をそのまま 2 次元に拡張する位置信号 2 次元化法が学習の収束性の点では一番良かった．また，1 次元ネットワークを単純に束ねて正規化を行った結果をもう一つの方向に同様のネットワークで正規化する 1 次元ネット 2 段縦続法は，正規化の正解率はやや劣るが，2 次元としての学習は不要であり，その正規化処理速度も一番速かった．

2.2 3層 BP ネットによる位置ずれ学習

ネットワークの各ユニット数は，入力層 7，中間層 7，出力層 7 の構成とする．連続する 3 ユニットからなる正值パターンを基本パターンと呼ぶ．この基本パターンを，順次右にサイクリックに平行移動した信号を入力とする．

2.2.1 任意ランダムパターン信号の位置ずれ学習

図 2.1 に示すように，3 ユニット分の幅をもち，その各ユニットにおける値が任意ランダムな入力パターン (基本パターン) がシフトのみを受けて，出力層の中央の 3 ユニットにそのまま出力されるように学習を行う．入力層の左端は右端と接続され，リング状になっている．

実験システムのホストコンピュータは NEC PC-9801RA である．また，ニューロ計算は NEC ImPP ボードと日本電気インフォメーションテクノロジー社ソフト NEURO-07 道真を使用した．

2.2.2 あいまい性

同期問題若しくは位置抽出問題は特に境界が不鮮明な場合，あいまいさが生じ，統計的学習方式においても，学習効率が落ちる [2]．従って，教育時にどこをパターン位置と指示するかが問題になる．本章では長さ 3 の短い，統計的に形が一様に変動するとみなせるパターンを用いたので，単純にその中央をパターン位置とし

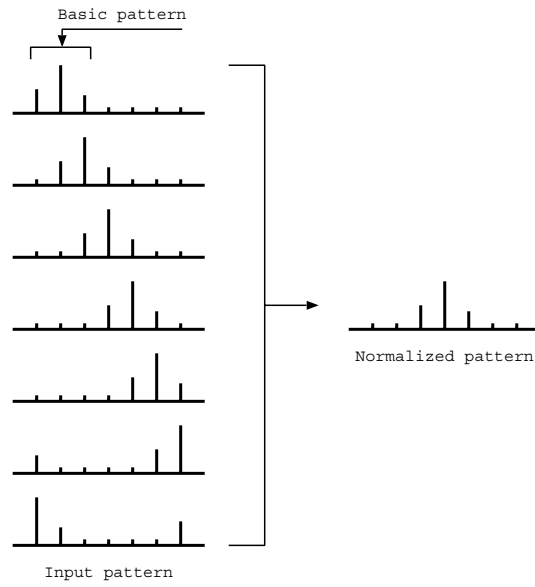
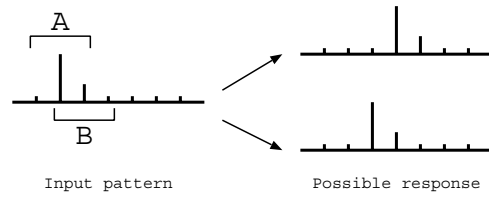
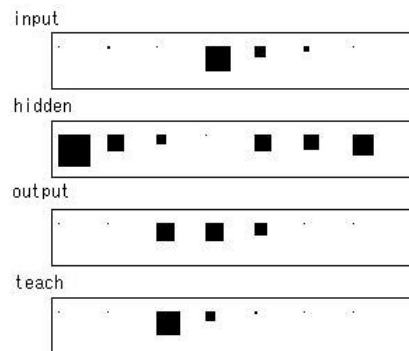


図 2.1: 位置ずれ学習プロフィール



(a) 1 入力に対して複数解が存在する例



(b) 実験例

図 2.2: 位置ずれ学習のあいまい性

表 2.1: 学習状況の比較 (学習終了時の 1 パターンあたりの出力誤差 2 乗和)

位置信号の有無	2 乗誤差和
無	4.78×10^{-2}
有	2.46×10^{-3}

た．しかしながら，本質的に候補が複数個あり，解が一意に定まらない (図 2.4(a)) ため，あいまい性の大きいパターンに対しては学習が進まない可能性がある．実験では 3 点の高さが $\{0.1, 0.2, \dots, 0.9\}$ のどれかをとるランダムなパターン 60 種を 7 箇所の位置にサイクリックにシフトしたものを (60×7) を 3 回繰り返したものを 1 セットの学習用入力パターンデータとし，これを必要回数提示した．端の値が小さいあいまい性の大きいパターンに対しては，出力の誤差 2 乗和は一定値以下にならなかった．

実験例として図 2.3(b) に 220 回の学習後の出力例を示す．入力パターンは $(0.0, 0.0, 0.0, 0.5, 0.2, 0.1, 0.0)$ である．“input” は入力パターン，“hidden” は中間層の活性状態，“output” は出力層の活性状態，“teach” は教師パターンをそれぞれ示す．この場合，上記の理由からパターンの位置のあいまい性で平均されたような出力が得られている．中間層ユニット数を変えて実験したが，パターンの位置ずれ補正はできなかった．

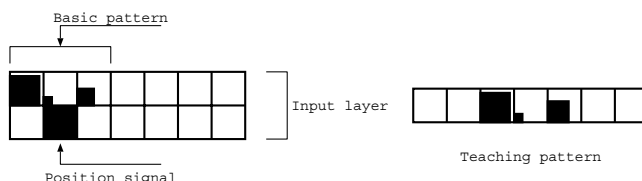


図 2.3: 位置ずれ入力と位置に同期した信号との併用による学習

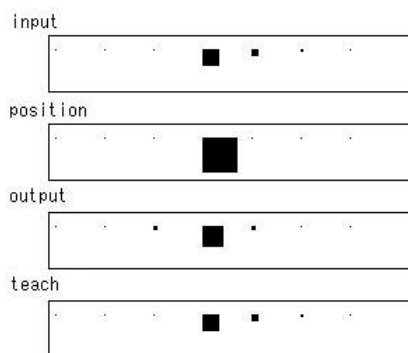
2.2.3 位置信号の挿入

2.2.2 のあいまい性をどの程度除外できるかを検討するために，入力パターンに加えて入力層に更に位置信号層を付加したものを考えた．この位置信号層には，図 2.3 に示すように基本パターンの中央部に対応する位置に “1” が位置信号として入力される．あいまい性のために学習が進まない場合でも，この位置信号の追加により学習が進むことが期待される．表 2.1 に位置信号のなし，ありの場合に対するそれぞれの学習状況の指標として，真値と出力との 2 乗誤差を示す．学習は前節と同じく

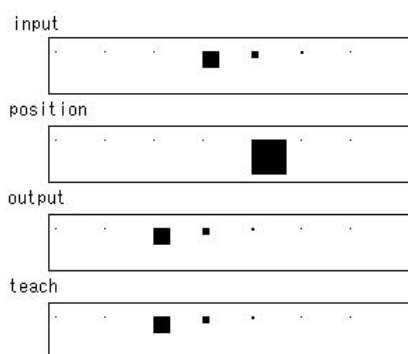
$$\begin{aligned}
 1 \text{ セット} &= 60(\text{基本パターン}) \times 7(\text{シフト回数}) \times 3(\text{パターン提示回数}) \\
 &= 1260 \text{ パターン}
 \end{aligned}$$

を 1 セットとし，2 乗誤差は学習終了時の 1 セット中の 2 乗誤差の平均とした，但し，基本パターンは高さ $\{0.1, 0.5, 0.9\}$ をランダムにとるものを用いた場合の結果である．またこの表の場合，雑音は加えていない．雑音を付加した場合はそれらに応じて誤差は増加する．図 2.4(a),(b) に基本入力パターンをどちらも $(0.5, 0.2, 0.1)$ にした実験例を示す．“position” は位置信号を示す．(a) では位置信号が “0.5” に対応するため，ネットワークは $(0.0, 0.5, 0.2)$ を基本パターンと認識し，それを中央に出力している，(b) は (a) と同じ位置にあるパターンに対して位置信号が “0.2” の位置に対応しているため，“0.5” を中央より左に出力している．以上のように，位置信号

を入力層に挿入することによって2.2で生じるようなあいまい性を取り除くことができた。また学習済みウェイトには空間周波数展開的な傾向が見られた。図 2.5 は入力層から中間層ユニットへの重みを表し、上方向は正、下方向は負、線分の長さは重みの大きさに比例している。また“1”~“8”は中間層ユニット番号，“a”は入力パターン側からの重み(しきい値(左端1個)+各入力パターンユニットからの重み(7個))，“b”は位置信号層側からの重み(7個)を示す。ここでは直流(7-a)，低周波(2-a,4-a,6-a)，中間周波(5-a,8-a)，高周波(1-a,3-a)，それぞれによく反応する重み分布が表現されている[8]。この場合、同じ周波数に属するユニットは位相が反転した重みが対になって出現しているケースが多い。その後、易ら[9]は同様の実験を行い、より明確な空間周波数展開の重み分布を抽出している。易らは入力ユニット数より中間層ユニット数が多い状態で中間層の空間周波数展開の特性を調べ、同一の空間周波数に属する中間層ユニットは2~数個の異なる位相特性をもつものが現れたと報告している。我々は入力層ユニット数と中間層ユニット数がほぼ同じという条件下で実験を行っているため、ユニット数がやや不足して主として位相が反転したものが対になって出現したと思われる。このような周波数分析機能が学習によって自然に獲得できたことはパワースペクトルが位置不変性に関して普遍的である可能性があることを示している。



(a) 基本入力パターン (0.0,0.5,0.2) に対する位置ずれ学習



(b) 基本入力パターン (0.5,0.2,0.1) に対する位置ずれ学習

図 2.4: 位置信号の提示によるあいまい性の除去の実験例

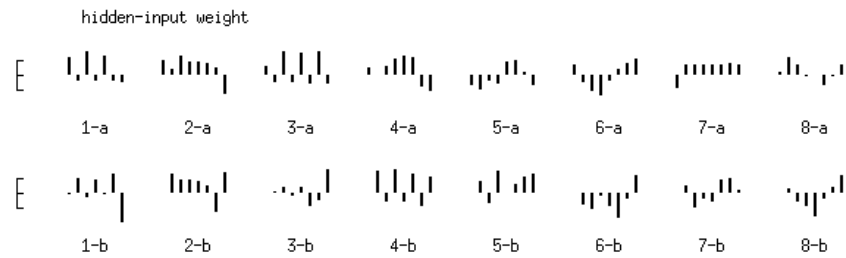


図 2.5: 入力層から中間層ユニットへの重み分布

2.3 位置信号の抽出とパターン正規化

前節では位置信号は教師が与えたが，ここでは更に，位置信号を自動的に抽出するネットワークを考え，その助けにより位置の正規化を行うことを試みる．

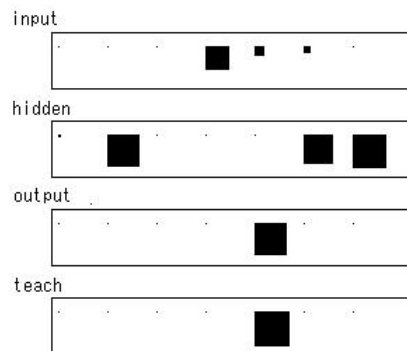


図 2.6: 位置信号抽出ネットワークによるあいまい性のあるパターンの位置信号の検出

2.3.1 位置信号抽出 NET (NET1)

ネットワーク，入力パターンはそれ，それ 2.1 と同じものを使用し，教師パターンには基本パターンの中央位置のみが“1”で，それ以外は，すべて“0”のパターンを用いて学習を行った．学習後の出力例を図 2.6 に示す．2.2 で述べたあいまい性があり，基本パターン部の中央に重心がない入力に対しても位置信号抽出ネットワークにより位置検出がうまく行われた．

2.3.2 二つのネットワークの組合せによる位置の正規化

2.3.1 で用いたネットワークを NET1，2.2.3 で用いたネットワークを NET2 とし，これを組み合わせた図 2.7 のネットワークによる位置の正規化を行った．但しネットワークは，NET1，NET2 を個別に学習を行った結果

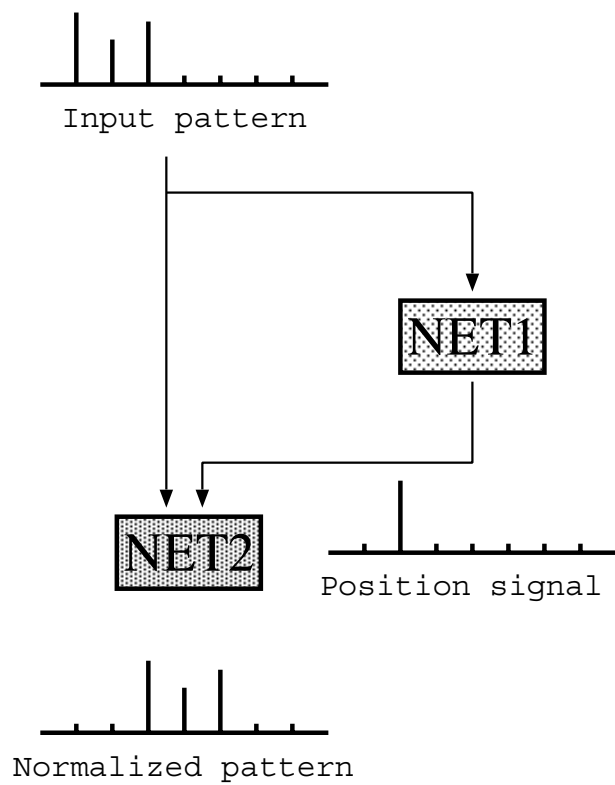


図 2.7: 二つのネットワークの組み合わせ構成による位置の正規化

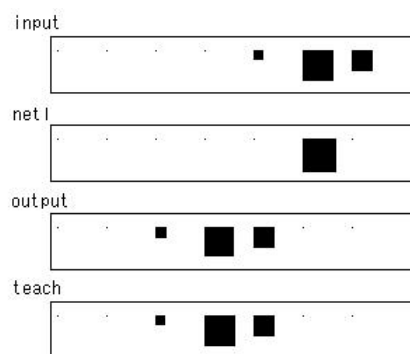


図 2.8: 位置正規化の成功例

のウェイトを用いて構成した．図 2.8～2.10 にさまざまな場合の出力結果を示す．“net1”は NET1 の出力層の活性状態，抽出した位置信号を示す．学習時およびテスト時の雑音において，例えば 10%は [0,0,1] の一様乱数 (正值) を入力パターンに加えたことを示す．

図 2.9 はあいまい性の大きなパターンの正規化の実験例を示す．図 2.9(a) は 2.2.1 と同じパターンを入力とした結果を示している．図 2.8 と比較し，位置信号の活性状態は十分ではないが，パターンの復元はある程度うまく行われている．

図 2.9(b) の例では位置信号が同時に 2ヶ所活性化しているため正規化が行われていない．このような場合，位置信号を一意に決定してやると入力パターンに近いものが対応する位置に出力される．同じ入力パターンに対して位置信号を強制的に決定した例を図 2.9(c),(d) に示す．もう少し大きな雑音を加えた場合の実験結果を図 2.10 に示す．

あいまい性のあるパターンに関しても惑わせるような大きな雑音がない限り NET1 で中央位置をうまく抽出できる場合が多く，組合せネットでは多くの場合正規化がうまく行えた．中央位置の抽出がうまく行えずあいまい性を伴う場合，ぼけたパターンが出力される．また組合せを用いない単一のネットでは，あいまい性がなくても学習が十分収束しないため正規化が十分行えなかった．このように，中間的な結論を得ておくこと学習効率が上がる現象は雑音に埋もれた未知パターンの認識・学習においても見られた．すなわち，雑音が一定以下のときには，パターンの有無を判定 (分類) してからパターン形状の学習を行った方が判定なしに一気に学習を行うよりその効率は高かった [2]．

上に述べた縦続ネットの能力を評価するため，学習，未学習パターンに対する正規化パターン出力の正解率をそれぞれ求めた (図 2.11(a),(b))．図 2.11 において，“square error<0.1”は，この条件を満たすとき正解とみなすことを表している，また，テスト時のノイズ 20%は各入力ユニットでの雑音エネルギー 0.0133 に相当し，それがそのまま出力されるとすると出力 2 乗誤差は 7 ユニット合計で 0.0931 となる．この図より，縦続ネットは単一ネットよりかなりうまくパターンの正規化を行えることがわかる．

2.4 2次元への拡張

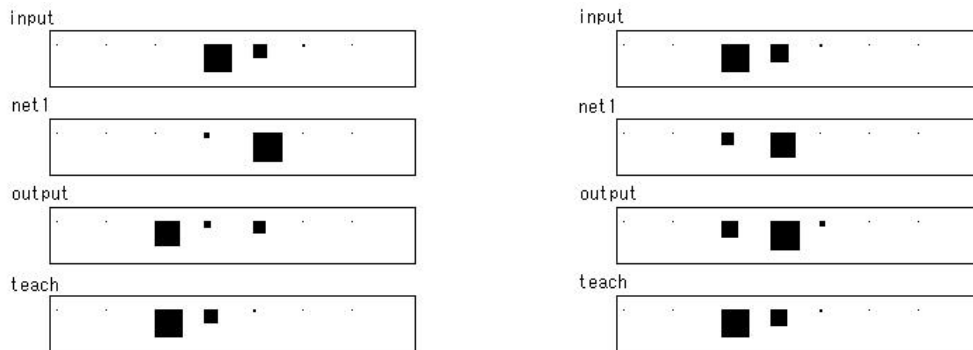
2.4.1 2次元の位置ずれ学習ネットワーク (位置信号 2次元化法)

ネットワークは 1次元の場合のものをそのまま 2次元へ拡張した構成とした．すなわち，入力パターンおよび出力パターンはともに 7×7 である．NET1 のユニット数は入力層 49，中間層 49，出力層 49 である．基本パターンサイズは 3×3 で，パターン $\{0.1, 0.2, \dots, 0.9\}$ のどれかをとりランダムな値を発生させて作成した．その基本パターンを入力層内で移動させ，出力層の中心位置に位置信号を発生させるように学習させた．NET2 のユニット数は入力層 98，中間層可変，出力層 49 である．入力パターン，位置検出信号を移動させ，中心部に基本パターンが出力されるように学習させた．実験はノイズ 0%および 10%について行った．学習時ノイズ 10%，テスト時 0%の場合についての出力例を図 2.12 に示す．これを位置信号 2次元化法とする．

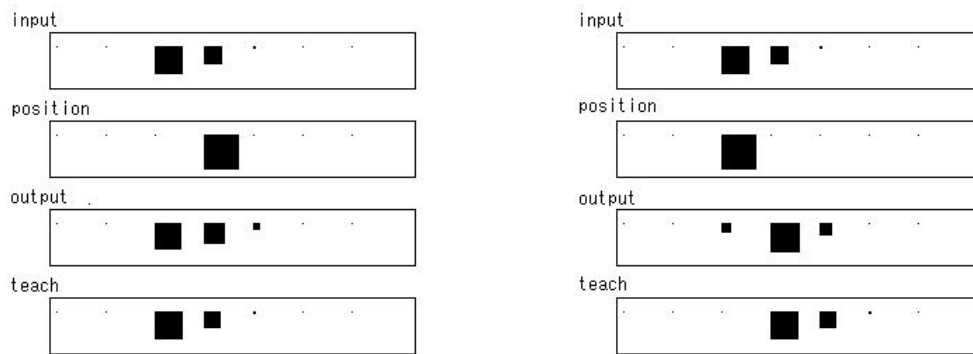
2次元パターンの場合，NET1 は収束性は非常によかったが，NET2 はウェイトの数が多いためか，学習時間がかかると共に収束性も良いとは言えなかった．そこで学習時間を減じるべく，ウェイトの数を減らした 2.4.2 および 2.4.3 の構成を考えた．

2.4.2 位置信号を二つの 1次元信号に分解した構成 (位置信号 1次元化法)

NET1 は入力層 49，出力層 14 である．位置信号は本来の 2次元位置信号 X,Y 軸にそれぞれ投影した信号が得られるよう学習を行った．NET2 のユニット数は入力層 63(入力パターン 49+ 位置検出信号 14)，中間層可



(a) あいまい性を含んだ例 1(ノイズ: 学習時, テスト時ともに 0%) (b) あいまい性を含んだ例 2(ノイズ: 学習時 10%, テスト時 3%)



(c) (b) において NET1 の出力を第 4 ユニットに固定した場合 (d) (b) において NET1 の出力を第 3 ユニットに固定した場合

図 2.9: あいまい性の大きなパターンの正規化実験

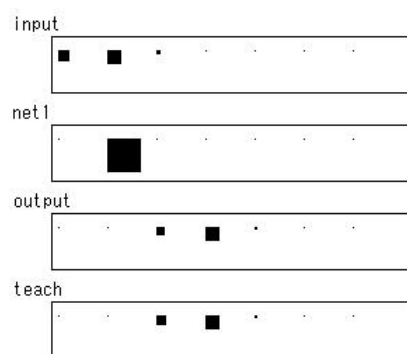
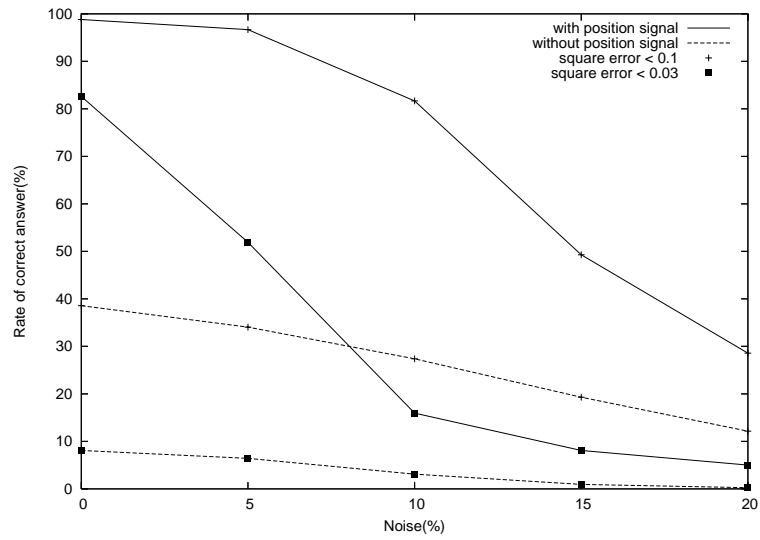
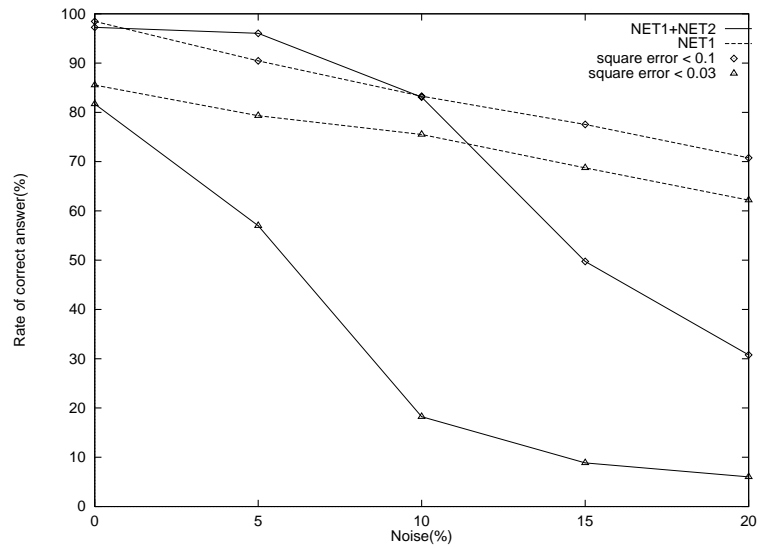


図 2.10: 入力パターンに少し大きなノイズを添付した例(ノイズ: 学習時 10%, テスト時 9%)

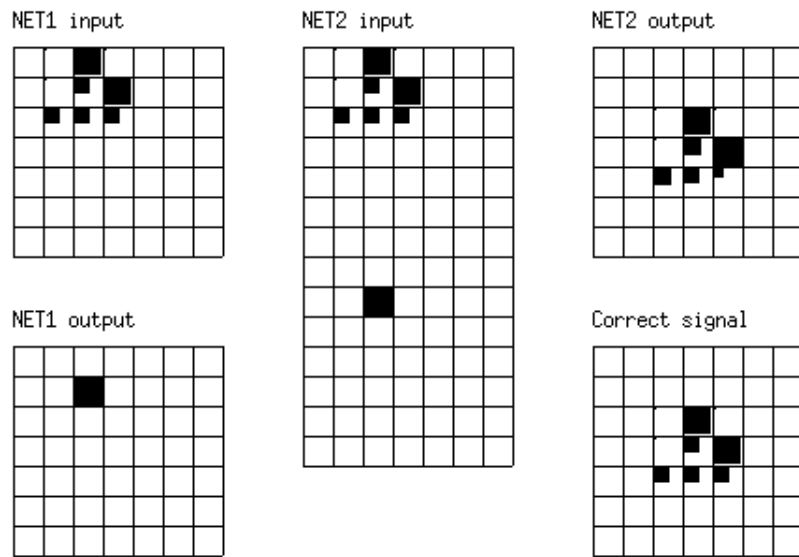


(a) 学習パターンに対する正解率

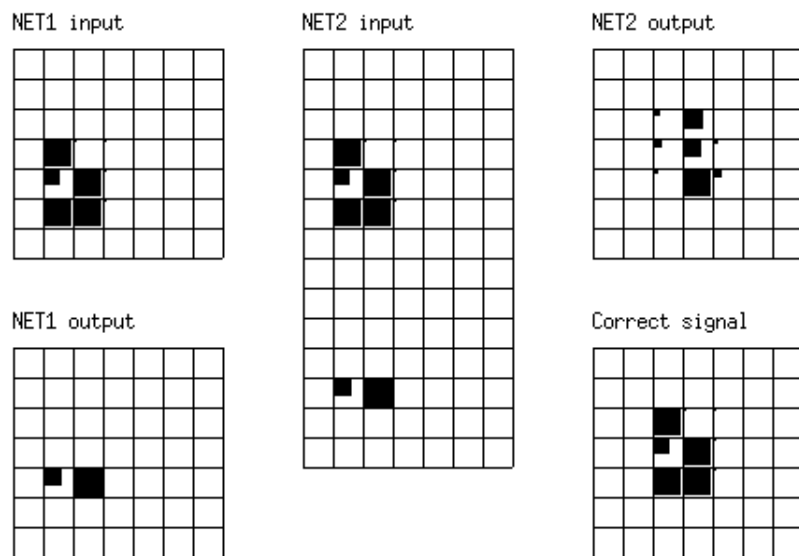


(b) 未学習パターンに対する正解率

図 2.11: 正解率の比較



(a) 成功例



(b) 形状保存が不十分な例

図 2.12: 二つのネットの縦続接続による位置の正規化

変，出力層 49 である。

2.4.3 1次元正規化ネットの2段縦続構成(1次元ネット2段縦続法)

2.3 で使用した 1 次元正規化ネットワークを図 2.13 に示すように 2 段縦続接続し，2 次元パターンの正規化を行おうとするものである。すなわち X 方向に正規化する層と Y 方向に正規化する層で構成する。この際，各層を 2.3 で得られたウェートをそのまま使用してネットワークの構成を行った場合，22.9%の低い正解率しか得られなかった。1 次元で見た場合，パターンの周辺のみを見るネットワークも存在する。そこでこのような零ベクトルパターンに対しては零ベクトルパターンを出力するように学習していただいたところ，正解率は 86.4%に向上した。

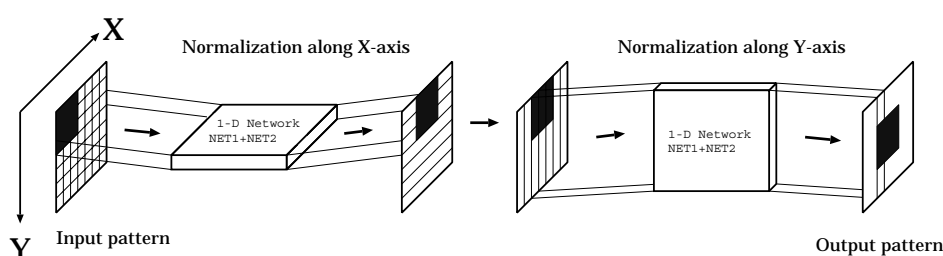


図 2.13: 1次元正規化ネットの2段縦続接続

2.4.4 実験結果

上記の三つの構成法に対してそれぞれのテストを行った結果を図 2.14 に示す。またネットワークの処理時間に関する指標として，各ネットワークの結合数を表 2.2 に示す。学習時間は結合数に学習回数を掛けたものにほぼ比例する。NET2 の学習は位置信号 2 次元化法，位置信号 1 次元化法それぞれ $60(\text{基本パターン}) \times 49(\text{シフト回数}) \times 3(\text{繰り返し回数}) = 8820$ パターンを 1 セットとして 2000 回まで提示，学習を行った。NET1 の学習は二つの構成法とも上記と同様のセットを提示，出力パターンと教師パターンの 2 乗誤差が 1 パターンあたり 0.005 以下になるまで学習を行った。そのため，NET1 により位置信号は位置信号 2 次元化法，位置信号 1 次元化法ともほぼ 100%の確率で正しく抽出された。テストは学習に使われたパターン 60×49 のすべてについて行った。正解率は出力パターンと教師パターンとの 2 乗誤差をしきい値 ($=0.49$) を用いて正解の判定を行った。

位置信号 2 次元化法，位置信号 1 次元化法では中間層ユニット数 49 のとき，それぞれ 97.3%，95.9%の正解率が得られている。これは 2 次元画像でも位置に関する正規化が行えることを示している。

位置信号 1 次元化法では 2 次元信号を二つの軸に分解したものを使用した。1 セットあたりの学習時間はおよそ 3/4 になったが，結合係数の減少に伴い学習の非常に時間がかかり学習の収束性に関するは有益な結果が得られなかった。但し，学習後のテストに関しては位置信号 2 次元化方に対し 3/4 程度の時間で正規化処理が行えた。

1 次元ネット 2 段縦続法では 86.4%の正解率が得られている。この構成法では 1 次元正規化ネットを 2 度使用するが，その際，各行および各列を独立に処理するため，若干正解率が下がった。この中には 1 次元ネットの NET1 があいまい性の影響を受けやすいため位置信号の抽出率が若干低いことによる影響も含まれる。この構成法は若干正解率が低いものの，学習が 1 次元の結果がそのまま使えるのに加え，テスト時間がおよそ 1/4 で済むため，とくに画素数の大きなパターンの正規化処理に向いていると思われる。

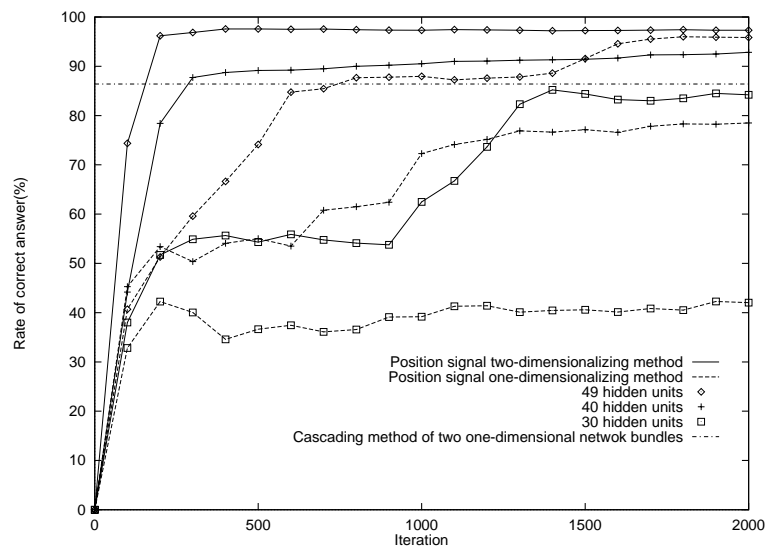


図 2.14: 2次元ネットの正解率の比較

表 2.2: 各ネットワークの結合数

	中間層 ユニット数	位置信号 2次元化法	位置信号 1次元化法	1次元ネット 2段縦続法	
				学習時	認識時
NET1	49	4802	3087	96	1344
NET2	30	4410	3360	147	2058
	40	5880	4480		
	49	7203	5488		

(注)1次元ネット2段縦続法の間層ユニット数はNET1,NET2ともに7である。

2.5 結言

本章では、まず、BPNによる位置の正規化実験について検討を行った。3層の一つのネットワークでは、位置ずれ学習は単純にはできなかった。実際、ニューロ学習問題は逆問題として考えると、ill-posed problem(不良設定問題)になっている [10]。不良設定問題とは、(i) 解が一意に定まらない、(ii) 解が存在しない、(iii) 解がデータに対して連続にならない、のうちいずれかが生じてしまう問題である。従って、解けない場合には何らかの拘束条件を組み入れなければならない。そのとき、その問題についての拘束をネットのニューロンだけでなく、ネットの構造自体にも組み入れることができると考えることができる [11]。

本章で得られた成果は、

- (1) 不良設定問題としての位置ずれ補正に対して、拘束条件をネットを二つ組み合わせた構成の形で入れた方法により解決した。具体的には正規化出力という難しい課題を単一ネットワークに任せてしまうのではなく、ネットワークの一部に位置信号というサブゴールを提示し、かつその出力を適切に使用できるようネットワークの構成を考えることにより、学習を容易にしている。
- (2) 1次元ばかりでなく2次元画像についても、位置の正規化を行うネットの構築を行った。
- (3) 入力層ユニット-中間層ユニット間の重みに空間周波数展開的な傾向が見られた。これは、フーリエ変換と同様の演算をニューラルネットで行っていることが示唆され、学習によりこれらの重みが得られたことは非常に興味深い。

本研究では、NET1, NET2について二つ別々に学習を行い、それを後で組み合わせたが、今後の課題として組合せ構造のまま、BPNで直接学習可能であるかどうか実験したい。また、回転、スケールなどの正規化も行っていきたい。更には、動物の脳におけるように、機能分化という観点からネットワークの構造をとらえたい。これには、進化の過程において、位置信号抽出に相当する入力画像から種々の対象を見つけ出し注目する機能、運動時の三半規管からの位置信号利用による静止外界の知覚機能などをいかに獲得したのかの問題が含まれる。

参考文献

- [1] Minsky M. and Papert S. *Perceptrons*. MIT Press, Cambridge, MA, 1969.
- [2] Tamura S. and Tanaka K. Synchronization unknown signal sequence by learning procedure. *IEEE Trans. Communication*, Vol. COM-20, No. 4, pp. 780–787, 1972.
- [3] Fukushima K. A neural network for visual pattern recognition. *IEEE Computer*, Vol. 21, No. 3, pp. 65–75, 1988.
- [4] Widrow B. and Winter R. Neural nets for adaptive filtering and adaptive pattern recognition. *IEEE Computer*, Vol. 21, No. 3, pp. 25–39, 1988.
- [5] Yang H. and Guest C. C. Performance of back propagation for rotation invariant pattern recognition. *Proc. IEEE First International Conference on Neural Networks*, pp. IV–365–370, 1987.
- [6] Trehub A. *Visual-cognitive neural networks*. MIT Press, 1987.
- [7] 竹谷, 光本, 田村, 河合, 岡崎, 副井. ニューラルネットワークによるパターン位置の抽出と正規化. 電子情報通信学会論文誌 D-II, Vol. J75-D-II, No. 7, pp. 1260–1270, 1992.
- [8] 竹谷尚, 岡崎耕三, 光本浩士, 田村進一, 河合秀夫, 副井裕. ニューラルネットワークによるパターン位置の抽出と正規化. 信学技法 PRU, pp. 89–93, 1990.
- [9] 易建強, 黒木秀一, 松岡清利. 誤差逆伝搬法による位置ずれパターン学習. 信学論 (D-II), Vol. J74-D-II, No. 1, pp. 27–35, 1991.
- [10] 小川英光. 逆問題とニューラルネットワーク. 第 2 回信学会回路とシステム軽井沢ワークショップ論文集, pp. 262–268, 1989.
- [11] 入江文平, 川人光男. 多層パーセプトロンによる内部表現の獲得—テーブルルックアップ法と違うのか?—. 信学技法 NC, pp. 89–15, 1989.

第3章 円形単板圧電素子を分割した多素子超音波トランスデューサによる水中物体の3次元画像化

3.1 緒言

トランスデューサによる水中物体の3次元画像化について、一般に超音波は、X線やMRI画像に比べて分解能が悪いが、非観血的・非侵襲的に体内の組織・器官の状態をリアルタイムに映し出せることから、臨床では胎児診断や心臓内の血流分布などの観察・計測によく用いられている。最近問題になっている成人病の原因の1つに動脈硬化がある。この治療法には、大きく分けて内科的および外科的の2つがある。本研究を始めた目的は、このうちの外科的治療に属するバルーン血管形成術において、心臓および血管内に置いたトランスデューサの前方対象物の画像化を行うことである。そのため超小型トランスデューサおよびそのデータ処理法を開発し、最終的には血管内に挿入して3次元的な超音波内視鏡としての臨床応用を行うことが最終目標である。

固定した超音波アレイセンサーを用いて前方物体の画像化を行う研究には次のようなものがある。

- (1) 1次元フェーズドアレイを2次元に拡張したもの；すなわち格子状トランスデューサアレイの位相制御により細い走査ビームを得るもの [1-3]。
- (2) 指向性をもったトランスデューサを平面状に並べることにより、画像化を行うのみ；たとえば格子状配置指向性トランスデューサアレイの個々のエコー遅延時間データ (3次元画像) をニューラルネットワーク処理により修復画像化する方法 [4]。
- (3) 音響ホログラム手法；これには同心円状に配置したトランスデューサを直交関数で駆動・受信し、相関を利用して画像化を行うもの [5, 6]、同様のトランスデューサを用いて、球面状反射波を考慮した一種の時空間的逆投影法により3次元物体形状の復元を試みたもの [7]、固定した超音波源から放射された超音波の反射波面をトランスデューサアレイで観測し、逆フーリエ変換手法により粗い物体像を得た後、ニューラルネットワーク処理により、修復するもの [8]。

などが提案されている。

しかしながら、これらについては次のような問題がある。(1)は2次元走査に時間を要し、動物体の実時間観測に不利である。(2)は空間的に分離された超音波ビームの独立性を利用して画像化を行うので、処理自体は簡単であるが、小型化が難しい。(3)は複雑な処理を必要とする。またこれらの手法は同一のトランスデューサ素子を並べることを一応の前提としている。そのため、単板分割により得た形状が異なる(したがって指向性などの特性が異なる)複数のトランスデューサ素子を用いた場合、その補正が必要となってくる。ただし、ニューラルネットワークを利用するものについては、その段階で差異を吸収できる可能性がある。本研究では、そのような柔軟性・適応性に加えて、実時間処理、小型化および処理の簡略化を目指して、ニューラルネットワークを利用することにした。

ニューラルネットワークは、本来人間をはじめとする生物のニューロン (neuron: 神経細胞) が織りなす情報伝達 (情報処理) を指すが、このニューロンによる情報処理システムを人工的に作り出す研究が盛んになり、各種のニューロン・モデル (artificial neuron) による (人工的) ニューラルネットワーク (artificial neural network) が工学などの多くの分野で利用されるようになった。

従来法であるプログラム計算方式は、達成される処理が既知の手順あるいは既知の規則セットの形で記述できる場合のみ使うことが出来る。現行のコンピュータは全て論理的に基づいて動作しており、ソフトウェア開発には、設計、試験、改良の徹底的な繰り返しが要求され、ソフトウェア開発は時間と費用のかかるものとなる。これに対して、ニューラルネットワークはアルゴリズムや規則の開発を必要とせず、ソフトウェアの量を大幅に減らすことが可能な新しい情報処理のアプローチである。

文献 [9, 10] には、多素子超音波センサーを試作し、これを用いて得られたエコー信号をニューラルネットワークに入力することにより、センサーの前方においた 3 次元物体形状の学習・認識が試みられている。その際に 2 次元物体での未学習物体の認識については形状のおよび面積的に挟んだ物体で学習を行うと未学習物体の画像化が行えることが報告されている [9]。また 3 次元形状物体の画像化ではニューラルネットワークの内挿能力を利用して、未学習の距離 (奥行き方向) に設置された対象物体を画像化できることが示されている [10]。ニューラルネットワークを利用する場合には、広い範囲の対象に対して適用可能とするため、汎化能力の検討を行う必要がある。

本章では、3 次元血管内超音波内視鏡の基礎実験として円形単板圧電素子を 36 分割した多素子超音波トランスデューサによる水中物体の画像化手法を提案する [11]。画像化には階層型ニューラルネットワークを用い、バックプロパゲーションアルゴリズムにより学習を行った。ネットワークの入力は受信した超音波エコーのエコーピーク時間を用いた。3 次元形状をもつ未知パターンに対して画像化実験を行い、本手法の有効性を確認した。その際、ネットワークの構成方法に関して情報量基準による評価を行い、与えられた学習パターンセットに対する入力パターンベクトルの圧縮方の検討を行った。最後に本システムを拡張したより実用的な 3 次元超音波画像化システムについて検討を行った。

3.2 3次元物体画像化システム

図 3.1 に本システムの概要を示す。超音波の送信-受信は、複数個の素子をもった 1 個の単板小型トランスデューサを用いて行う。トランスデューサの正面像は図 3.2 に示すように、円形単板圧電素子を円周方向に 12 分割、半径方向に 3 分割したもので、全部で 36 素子ある。形状は外径が 20mm、内径は 2mm、各素子の共振周波数は約 9.67MHz (標準偏差:0.096MHz)、表面は共通電極で、各素子の表面積は 8mm² 程度でどの素子も同じ面積になるように分割してある。図に示すように各素子には 1 ~ 36 の番号が振られている。また各素子の形状は異なるため、各素子の放射特性は異なっていることが予想される。本研究ではトランスデューサの小型化よりもデータ処理法の研究に主眼をおいたため、製作と実験の容易さを考えて 36 素子 20mm 径 (中心部 12 素子は約 11mm 径) のトランスデューサを使用した。全体形状は内視鏡に適した円形とした。この際、異形状素子のデータを処理するため、ニューロ処理のような適応的処理が必要となってくる。なお Watanabe ら [8] は、ニューラルネットワークを用いた超音波 3 次元物体認識法を開発しているが、画像復元は主としてフーリエ逆変換法により行っており、ニューラルネットワークは復元画像の修正と認識に用いているのみである。

トランスデューサは直方体のプラスチック容器 (220mm(L)×150mm(W)×90mm(H)) の側面 (W) 中央部に取り付け、ターゲットはトランスデューサの前方 18mm (第 1 画)、20mm (第 2 画)、22mm (第 3 画) の 3 平面内においた。データ入出力はパーソナルコンピュータに接続されたデジタル入出力ポートから行う。

パルス発生部で発生させたパルスを用いて、トランスデューサのどの素子を駆動するか、また各素子に入ってきた受信エコーのうちどの素子のエコーを計測するかは、重複しないように順番にリレーによるデマルチプ

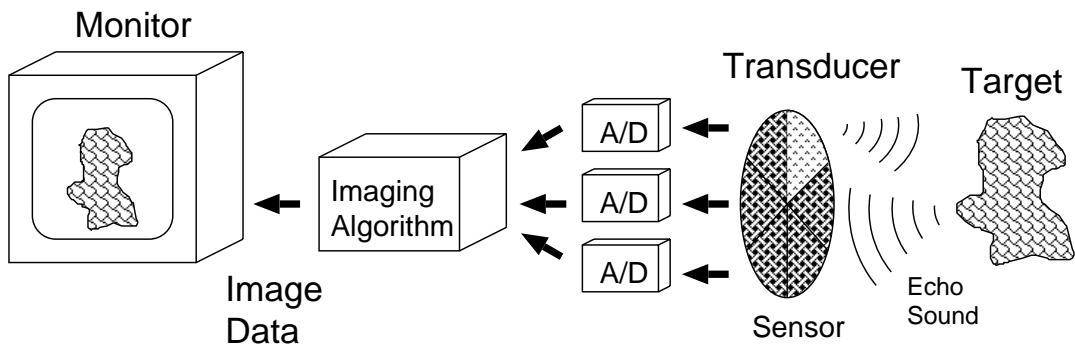


図 3.1: システム概要

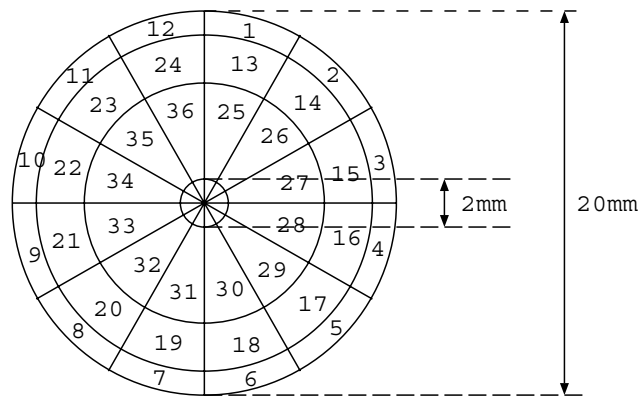


図 3.2: トランスデューサの正面図

レクサおよびマルチプレクサで選択する．駆動には高速パルサーを用い，70V 幅 $0.5\mu\text{s}$ の狭い単一パルスを順番に各素子に送り，他の分割素子 35 個で個別に受信し，1 つのターゲットに対して 1260 組 ($=36 \times 35$) のエコーデータを取得する．

実験に用いたターゲットまでの距離，18,20,22mm は，トランスデューサ前面に施した厚さ約 1.5mm の耐水樹脂面からの距離である．樹脂内部の伝搬速度を水中の 3 倍とすれば，実際のトランスデューサからターゲットまでの超音波の往復時間はそれぞれ最短距離に対して $24.6, 27.3, 29.9\mu\text{sec}$ となる．そこでそれぞれのデータについては，1 個の素子にパルスを印加してから， $23\mu\text{sec}$ 後に選択された受信素子出力のサンプリングを開始し， $43\mu\text{sec}$ 後まで $20\mu\text{sec}$ 間行う．サンプリング速度 100MHz(サンプリング間隔 10nsec) で，サンプリング点数 2048 点で行う．A/D 変換にはデジタルオシロスコープを用い，デジタル化されたエコーデータはオシロスコープの GPIB インターフェースを通してパーソナルコンピュータに転送される．

3.2.1 画像化対象

今回，画像化を行う対象は矩形を組み合わせたものに限定した．実験に用いたターゲットは，厚さ約 3mm のアクリル製の板を加工したもので，その形状を図 3.3(a)~(g) に示す．最小のターゲット A は $4\text{mm} \times 4\text{mm}$ の大きさで，他はそれを矩形状に連結したもので，たとえば F は $12\text{mm} \times 16\text{mm}$ の大きさである．

ターゲット A... 4mm × 4mm
 ターゲット B... 8mm × 8mm
 ターゲット C... 12mm × 12mm
 ターゲット D... 4mm × 16mm
 ターゲット E... 8mm × 16mm
 ターゲット F... 12mm × 16mm
 ターゲット G... 16mm × 16mm

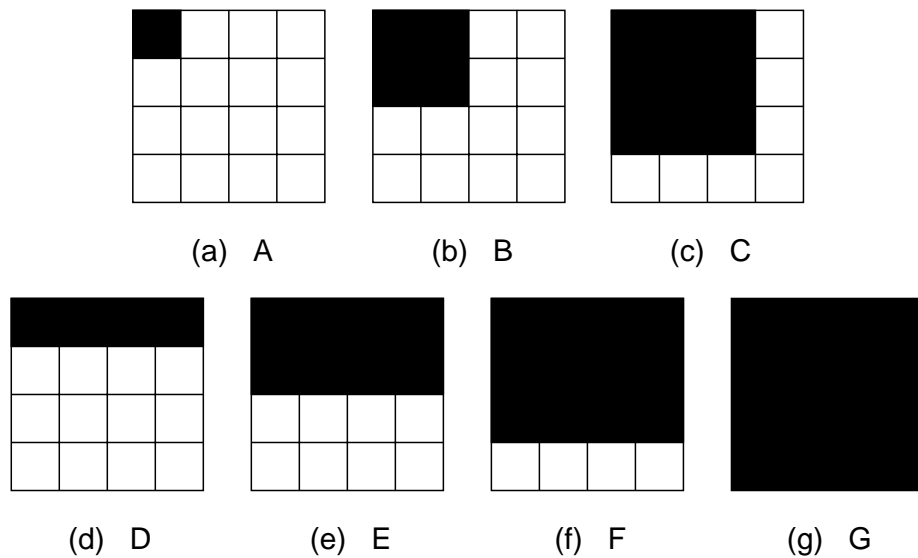


図 3.3: 学習に用いたターゲット

ターゲットの位置は，トランスデューサの前方 18,20,22mm のトランスデューサの中心軸に垂直な 3 つの平面 (第 1 面，第 2 面，第 3 面) 上で，図 3.3 に示すように 16mm×16mm の正方形の領域内に置く．縦横の 1 目盛りは 4mm で 16 個の小正方形からなっている．これはターゲットの基本単位 (画素) を示しており，A~G については基本ターゲットとして画像化することを検討した．大多和ら [10] はニューラルネットワークを用いて未学習の距離に設置された対象物体の画像化を行ったが，本研究ではより実用的な画像を得るため第 2 面にも学習パターンを提示し，学習を行い凹凸のある対象物体の画像化を試みる．

各平面におけるターゲットの配置は，ターゲット A については，図 3.3(a) に示すように，16 個のそれぞれの小正方形の 16カ所に置いた．B,C については，図 3.3(b),(c) の大きさの正方形を，図 3.3 の 16 個の小正方形からなる面上で，縦横に順番に基本単位ずつずらしていき，それぞれ 9,4カ所に置いた．D,E,F,G については，16 個の小正方形からなる正方形内で水平および垂直方向に順番に基本単位ずつずらし，さらに 90 度回転した，8,6,4,1カ所にそれぞれ配置し観測し，合計 18 個 ($= 2 \sum_{i=2}^4 i$) のターゲットを観測した．

3.2.2 受信エコーと前処理

ターゲット C をトランスデューサの正面前方 18mm に置き、ある素子 (No.1) を駆動して、別の素子 (No.2) で受信したエコー信号の波形例を図 3.4 に示す。図 3.4 において、 $1\mu\text{s}$ 付近でパルスを発射している。 $25\mu\text{s}$ 付近の高周波振動を第 1 エコー、 $28\mu\text{s}$ 付近の高周波振動を第 2 エコーとすると、第 1 エコーがターゲット前面で反射したもの、第 2 エコーがターゲットの背面部分で反射したものである。図 3.4 の全体にわたる低周波振動は、本実験で用いたトランスデューサが単板を分割しただけのものであるため、単板全体の振動によるものと考えられる。すなわち高周波振動は素子の厚み (0.4mm) 方向の振動であり、低周波は直径 (20mm) 方向の振動であると仮定すれば直径方向の周波数は約 193kHz となり低周波振動の周波数と一致する。また低周波振動は高周波振動に比べて振幅が小さいため反射波は観測されていないと考えられる。

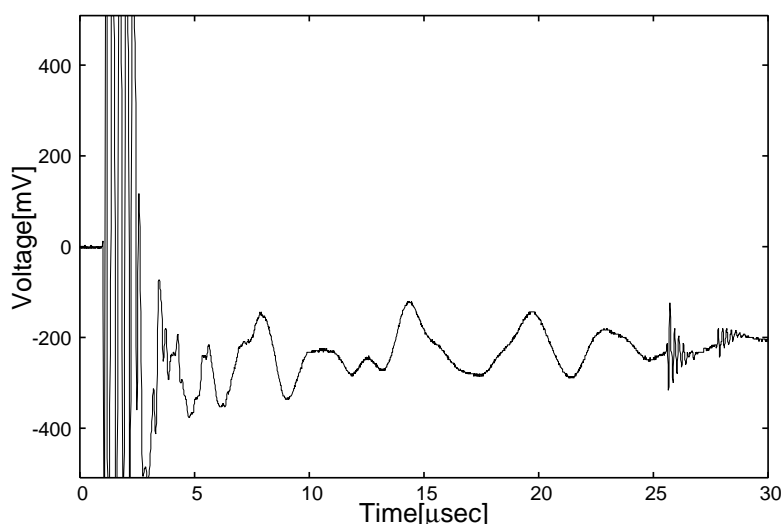


図 3.4: 受信エコー例

これらのデータに対してバンドパスフィルター処理を施すことによりエコーから不要な低周波成分を除去する (図 3.5)。図 3.5 の高周波成分は送受信素子の組み合わせによっては観測されないものもある。図 3.4 の $24\mu\text{s}$ から $20\mu\text{s}$ 間サンプリングを行い、2048 点のデータを得る。これらのデータを用いて対象物体の 3 次元形状の復元を行った。

3.3 超音波信号の 3 次元画像化

3.3.1 観測モデル

図 3.6 において原点上に receiver, Y 軸上の距離 d 離れた位置に transmitter, Z 軸方向に距離 l 離れた位置に target を配置する。ただし送受信を行う素子は指向性 [12] をもち、伝達経路において信号は減衰しないものとする。時刻 $t = 0$ において信号 $f(t)$ が transducer より発射されたとすると、target 上の点 a で観測される信号 $g(t)$ は指向特性 $D_T(\theta_T)$ に従い減衰し、伝播経路の距離に比例した時間遅れを生じ、次式で表される。

$$g(t) = f\left(t - \frac{l \cos \theta_T}{c}\right) D_T(\theta_T)$$

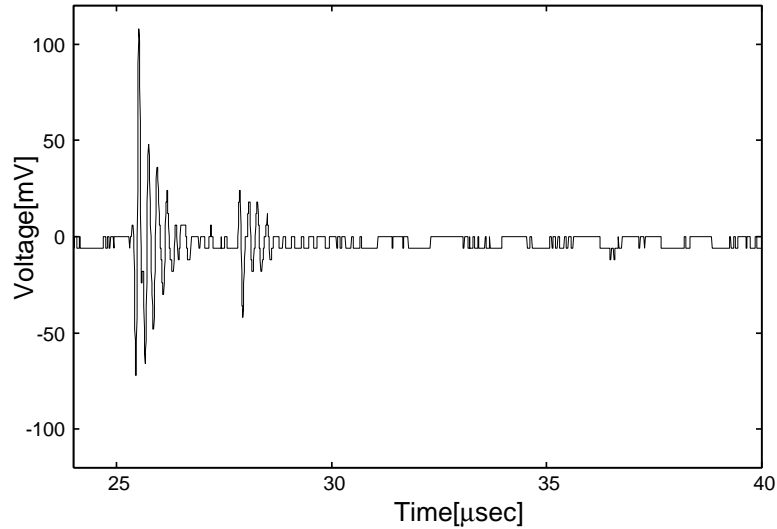


図 3.5: 低周波成分を除去した受信エコー

ただし c は水中での超音波の伝播速度である．また，点 a から発射される信号を $G(t)$ とすると receiver で観測される信号 $h(t)$ は

$$h(t) = G\left(t - \frac{l \cos \theta_S}{c}\right) D_S(\theta_S)$$

となる． $g(t)$ が点 a で乱反射すると仮定すると点 a から発射される信号は $G(t) = g(t)$ となり $h(t)$ は

$$h(t) = f\left(t - \frac{l \cos \theta_T + l \cos \theta_S}{c}\right) D_T(\theta_T) D_S(\theta_S)$$

となる．最終的に sensor において観測される信号 $H(t)$ は次式により示される．

$$H(t) = \iint_{target} h(t) dx dy \quad (3.1)$$

となる．式 (3.1) より target の形状を求めることが可能となれば対象物体の画像化が可能となる．しかし，上式では

(1) 本質的に式 (3.1) は複数の解を持つ (不良設定問題) ．

(2) $f(t)$, $D_S(\theta)$ および $D_T(\theta)$ はトランスデューサの形状等の特性に依存しているため，実測する必要がある．

等の問題点があげられる．上記の問題点に対応するためニューラルネットワークによる適応型処理により画像化を行った．

3.3.2 ニューラルネットワークによる超音波信号の画像化学習

本章では，RF レベルの信号を直接用いて対象物体の形状を推定することはデータ数の増加により困難と考え，最大エコーピーク時間を特徴量として抽出し，画像化を行うことを検討した．

図 3.7 にエコーピーク時間を 2 次元状に配置したもの (エコーピーク時間図と称する) の例を示す．水平方向は超音波パルスを発射した素子の番号，垂直方向は受信した素子の番号に対応し，各点の画素値はエコーピー

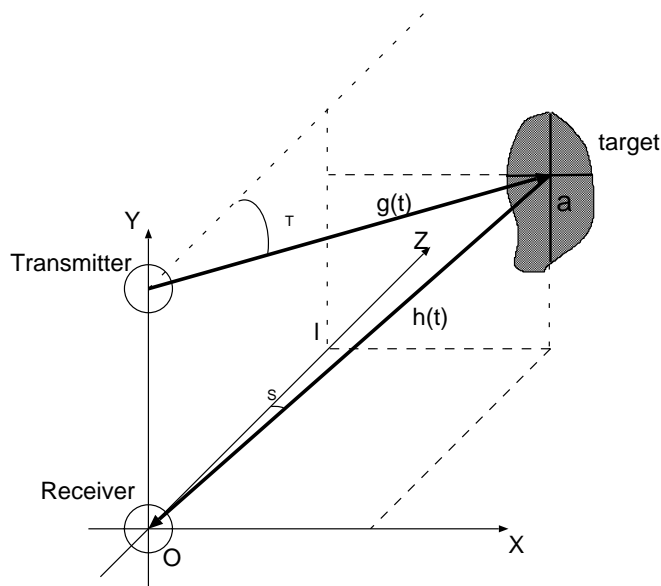


図 3.6: 超音波信号観測モデル

クが受信された時間に比例する（エコーピークが早く観測されるほど値が大きく、図では明るく表現されている）。エコーが観測されない場合は値を 0 とし、図では黒とした。対角成分は送受信の素子の番号が同一であるため、データを得ていない。図 3.7 の水平・垂直方向の線分は分割した圧電素子の形状に関する指向性によるもので、観測するセンサーの正面付近にターゲットがあればエコー信号が観測され、エコーピーク時間図中の濃淡値として表現される。

画像化には 3 層構造のフィードフォワードネットワーク（中間層ユニットおよび出力層ユニットの入出力関数にはシグモイド関数を使用）を使用し、学習にはバックプロパゲーションアルゴリズムを用いた。入力層のユニット数は 1296(36 × 36) とし、エコーピーク時間図の次元数と一致させた。すなわち 1 つの受信サンプリング値系列データに対して、サンプリング開始点からエコー信号の最大振幅位置までの時間を 1 個の入力データとしている。中間層のユニット数は 10~200 ユニットとし、出力層のユニット数は 16 ユニットで 4 × 4 の 2 次元行列表示である。学習には、ターゲット種別（括弧内は移動カ所/1 面）、A(16カ所)、B(9カ所)、C(4カ所)、D(8カ所)、E(6カ所)、F(4カ所)、G(1カ所) の合計 48 パターンとし、これらを

- (1) 第 1 面のみ提示した 48 パターンを学習パターンとしたもの (2 次元情報の学習)
- (2) 第 1,2,3 面の三カ所に提示した 48 × 3 = 144 パターンを学習パターンとしたもの (3 次元情報の学習)

の 2 通りについて学習を行った。なお教師信号の値は第 1 面、第 2 面および第 3 面でそれぞれ 0.95, 0.65, 0.35 とし、その画素に対象物体がない場合は 0.05 とした。学習は出力層ユニットあたりの平均 2 乗誤差が閾値 (=0.0001) 以下になれば学習終了とした。ニューラルネットワークの重みは [-0.1, 0.1] の一様乱数で初期化し、シグモイド関数の傾き係数 0.75、学習パラメータは荷重修正係数 0.001、荷重忘却係数 1.0 とした。これらのパラメータを用いて、1000 万セットの提示を上限として学習を行った。図 3.8 に学習状況を示す。中間層ユニット数を変更した場合も同様の曲線が得られているが、中間層ユニット数によっては閾値に到達しないものもある。

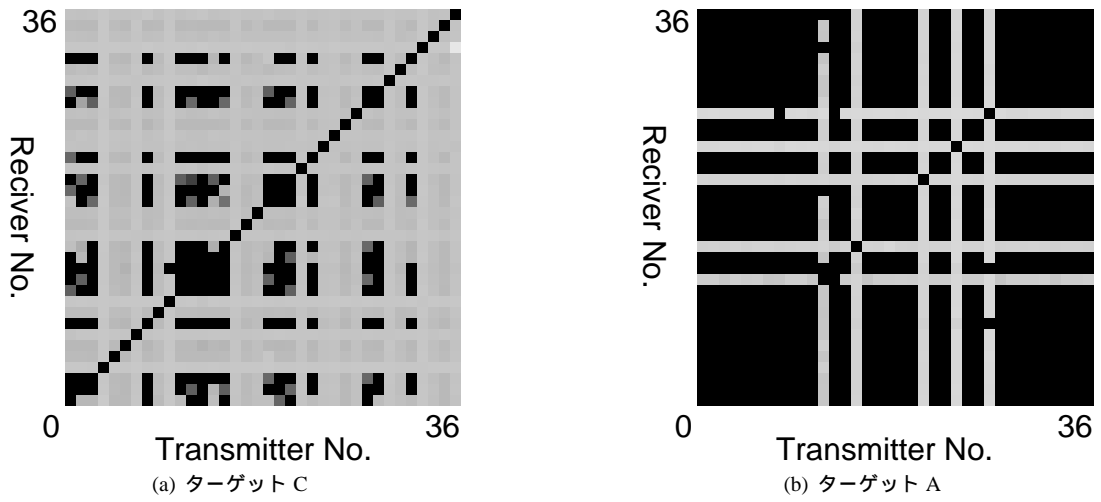


図 3.7: 各ターゲットの 36×36 エコーピーク時間図例

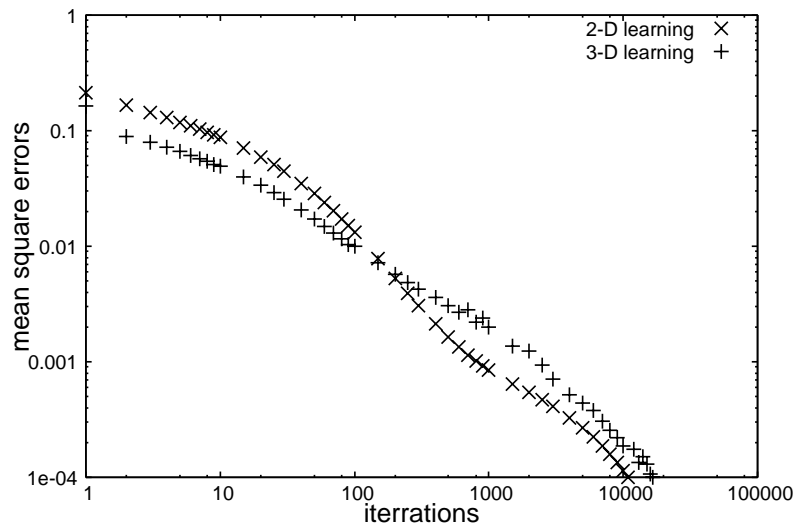


図 3.8: エコーピーク時間図による画像化の学習状況 (中間層ユニット=100 の場合)

3.3.3 未知ターゲットの画像化

図 3.9 に学習済ネットワークに提示した未学習パターンを示す．ターゲットの最小単位は 3.2.1 と同様に $4mm \times 4mm$ とし，穴などが無い単純連結形状のターゲットを対象とした．図 3.9(a)-(h) をターゲットの最前部が第 1 面として観測空間内で平行移動させたものを 3.3.2 の学習済ネットワークに提示し，画像化実験を行った．ただし (e) および (f) についてはターゲットの最前部を第 1 面および第 2 面にセットしたのも未学習ターゲットとして使用した．パターン数はそれぞれ H(24ヶ所)，I(16ヶ所)，J(4ヶ所)，K(6ヶ所)，L(18ヶ所)，M(24ヶ所)，N(4ヶ所)，O(4ヶ所) である．

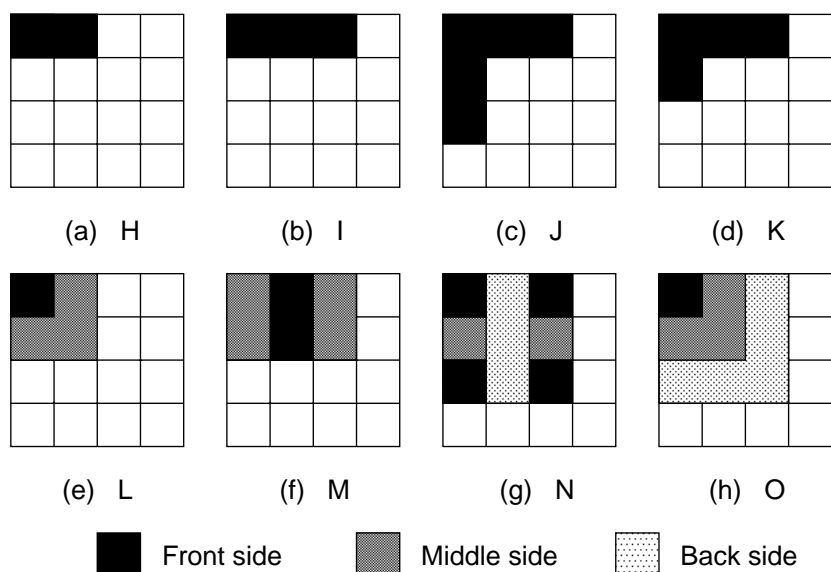


図 3.9: 画像化実験に用いたターゲット

図 3.10(a) に提示したターゲットモデルを示す．図中の左の円形の物体はトランスデューサを示し，左の矩形物体が画像化を行った物体である．図 3.10(b) にニューラルネットワークによる画像化を行った例を示す．図 3.11(a),(b) に別の例を示す．

図 3.10(b), 3.10(b) の例では X-Y 平面内での物体の大きさ，位置および形状ともに良好に画像化されているのが確認できる．しかし Z 方向に関しては十分に復元できているとまでは言えない．この理由として X, Y 方向に関しては独立した個別の入力素子で受信されるのに対して，Z 方向の画像化に使用される情報がエコーピーク時間という同一素子上での連続値として用いられているため十分な分離精度が得られていないことが挙げられる．

また図 3.11(b) より複数の矩形により構成されたターゲットに関しても画像化が行われていることが確認できる．特徴量として採用したエコーピーク時間図はターゲットに対して必ずしも線形性を有しておらず，3 次元画像を再構成する場合に非線型処理が必要となる．本研究では特に困難であるエコーピークの情報把握についても，ある程度はニューラルネットワークの非線型性により対応可能になると考えられる．本研究で採用したターゲットより複雑な形状 (例えば穴をもつ形状や穴を多数もつ網状形状およびそれらに奥行きを加えたもの) をもつターゲットにおいては，複数面で反射したエコーが受信素子上で指向性に従い合成され，エコーピークが観測される．これはかなり非線型な合成となり，学習の困難さが増すが，3 次元情報の獲得に奥行き情報を持つ学習パターンを加えることにより対応できたのと同様に，複雑な形状を一部にもつパターンを加えて学習

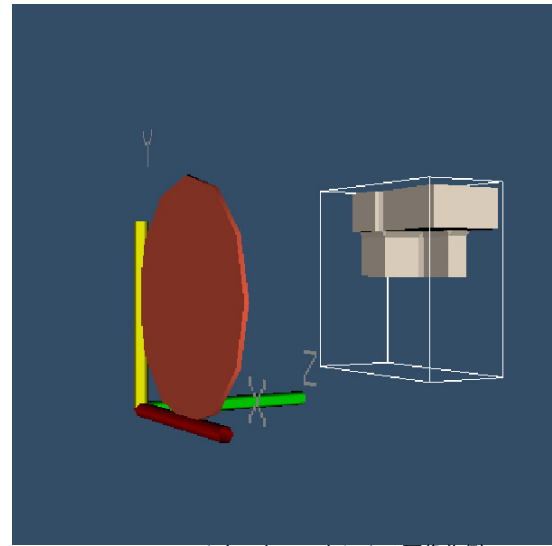
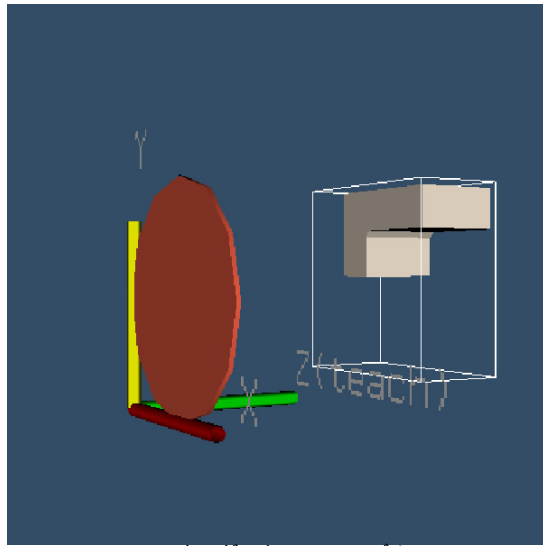


図 3.10: ターゲット“K”の画像化例

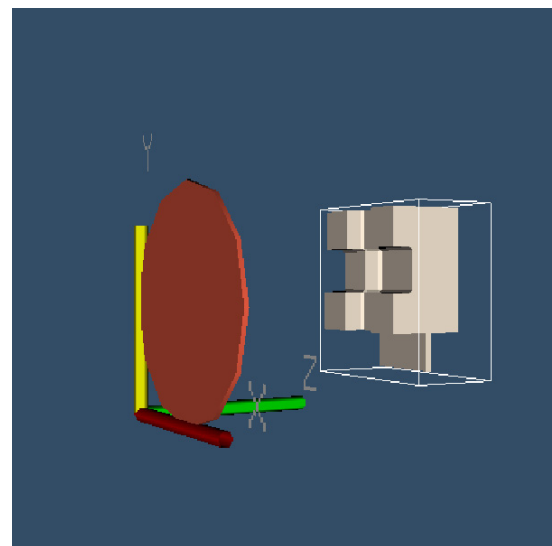
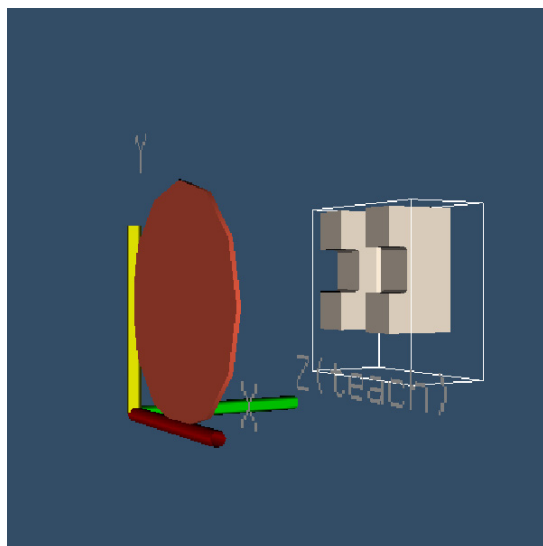


図 3.11: ターゲット“N”の画像化例

を行うことにより、ある程度は本提案手法で対応可能であると考える。

3.4 特徴の縮約化

3.3.3 の実験では入力データは 36 素子を個々に用いているが、ターゲットの複雑さに対して多すぎる可能性がある。そこで受信素子 35 個のピーク時間を平均したデータ (投影データと称する) を作成し、これを用いて学習および未学習パターンによるテストも行った。

図 3.12 に中間層ユニット数を変化させた場合の未学習パターンに対する平均 2 乗誤差を示す。ただし、中間層ユニット数が 30 以下では学習時の平均 2 乗誤差が閾値以下にならなかったが 1000 万セットの学習で打ち切り用いている。平均 2 乗誤差 E は次式を用いた。

$$E = \frac{1}{NK} \sum_{i=1}^N \sum_{k=1}^K (O_{ik} - t_{ik})^2 \quad (3.2)$$

ただし、 N はパターン数、 K は出力層ユニット数、 O_{ik} はパターン i を提示した場合の出力層ユニット k からの出力、 t_{ik} はパターン i の出力層ユニット k の教師信号である。

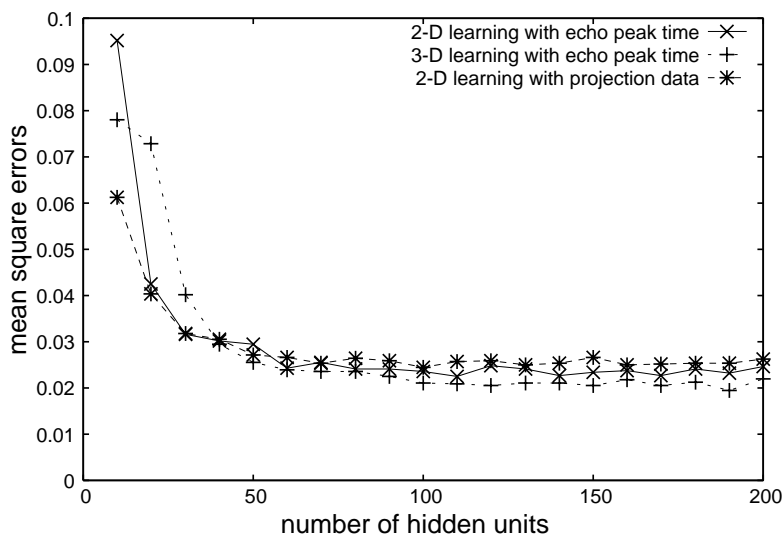


図 3.12: 未学習パターンに対する平均 2 乗誤差 (学習: 2 次元パターン, テスト: 3 次元パターン)

図 3.12 より 2 次元情報のみを学習したネットワークと 3 次元情報を学習したネットワークが同等の内挿能力を有していることが確認できる。

階層型ニューラルネットワークでは中間層ユニット数を多くすると学習パターンについて任意の関数を任意の精度で近似可能である [13] ことが知られているが、本稿で用いた学習データは実測したもので S/N 比が低いため、再構成された画像の精度が向上には限界があることが考えられる。この改善策として

- (1) 同じターゲットに対して多数回観測し、これらを用いて学習を行う。
- (2) 式 (3.1) より理論的にシミュレートした人工的上の観測データを作成し、これを用いて学習を行う。

などが考えられる．

一般に次元数の異なる学習パターンセットに同等の情報量がある場合，次元数が少ない方が入力パターン空間に対する学習パターンの相対密度が上昇するためネットワークの内挿（汎化）能力が向上すると考えられる．そこで本研究では上記のように各受信素子のエコーピーク時間の平均値を特徴量として使用することを考えた．その際に未学習パターンに対する平均 2 乗誤差を図 3.12 に示す．学習パターンは 48 とし，第 1 面にターゲットをおいた場合を学習セットとした．図より投影データを用いて 2 次元情報を学習したネットワークがエコーピーク時間図で学習を行ったものとはほぼ同等の汎化能力をもつことが確認される．これは入力層ユニット数が 1296 から 36 へ減少したにも関わらず，入力パターンに含まれる情報量が減少していないことが示唆される．

3.5 考察

ニューラルネットワークを使用する場合，汎化能力が問題となる．ここでは今回取り扱った超音波の 3 次元画像化の複雑さおよびネットワークの汎化能力を情報量基準に基づく AIC [14] および MDL [15] を適応して評価を行った．

$$AIC = -2\tilde{l}(\{\mathbf{t}_i\}) + 2F \quad (3.3)$$

$$MDL = -\tilde{l}(\{\mathbf{t}_i\}) + \frac{F}{2} \log N \quad (3.4)$$

ただし， F はモデルの自由度とする．最大尤度 $L(\{\mathbf{t}_i\})$ および最大対数尤度 $l(\{\mathbf{t}_i\})$ は次式 [16] で近似している．

$$L(\{\mathbf{t}_i\}) \approx \frac{\exp \left[-\sum_{i=1}^N \sum_{k=1}^K O_{ik}(1-t_{ik}) \right]}{\prod_{i=1}^N \prod_{k=1}^K \{1 + \exp(-O_{ik})\}} \quad (3.5)$$

$$l(\{\mathbf{t}_i\}) \approx -\sum_{i=1}^N \sum_{k=1}^K O_{ik}(1-t_{ik}) - \sum_{i=1}^N \sum_{k=1}^K \log(1 + \exp(-O_{ik})) \quad (3.6)$$

本研究ではこれらの評価基準を入力パターンの次元数の評価に用いた．エコーピーク時間図を使用した場合は入力パターン空間の次元が多すぎるため，学習において適切な入出力関係を獲得するため過大な労力を用いている可能性が強い．これは図 3.12 において平均値による圧縮（投影データ）を用いても出力の平均 2 乗誤差はほとんど変わらないことからいえる．今の場合， 36×36 エコーピーク時間図を用いるニューラルネットワーク構成に対して相対的に学習パターンが少なかったと言える．しかしながら，より多くのパターン種を学習するため，ネットワークを構成する必要最小限の次元数を知る手法を確認しておくことは必要である．

図 3.13 に入力データの次元数を変化させた場合にモデルの自由度 F (重みの数) に対する AIC および MDL を示す．ここでは以下の特徴量を用いた構成方法について比較を行った．

mean36 各受信素子でのエコーピーク時間の平均（36 次元入力）

mean36x3 各受信素子でのエコーピーク時間の円周方向での平均（ 36×3 次元入力）

mean36x12 各受信素子でのエコーピーク時間の半径方向での平均（ 36×12 次元入力）

mean36x15 “mean36x3” と “mean36x12” を足し合わせたもの (36×15 次元入力)

echo peak time エコーピーク時間図 (36×36 次元入力)

3.14 には中間層ユニット数を変化させた場合の未学習パターンに対する AIC およびそのうちの $l(t_i)$ を示す。図より中間層ユニット数が 40 以上では各構成方法とも同程度の値を示しており、本実験で使用したテストパターンに対する同程度の汎化能力を有していることが確認できる。図 3.13 で厳密には “mean36x3” が AIC および MDL とも “mean36x36” と比較してわずかに小さい値を示しているが、これは図 3.14 において中間層ユニット数 100 のときの各方式の変動による差と考えられ、有意な差とは言えない。しかし全体として入力データの次元数は “mean36” および “mean36x3” でも十分であることが確認できる。

AIC および MDL を使用したネットワークの評価では、ネットワークの規模が小規模な場合は問題ないが、本研究で使用する規模のネットワークでは第 2 項の影響が非常に強い。本来評価基準はネットワークの規模に左右されないものが望ましく新しい基準が必要である。

本研究ではデータの前処理、ニューラルネットワークの構築および 3 次元表示にパーソナルコンピュータ (CPU: AthlonXP2800+, OS: Linux) を用いた。実験システムの都合上、現在は 1260 組のデータを取得するのに約 17 分必要であるが、35ヶ所同時にエコーを取得することができれば撮影時間の短縮が可能である。信号がターゲットで反射し、観測されるまでの時間を約 $40\mu\text{sec}$ と仮定すると、 $40\mu\text{sec} \times 36 = 2.52\text{msec}$ となる。各処理時間は前処理: 5.11sec, ニューラルネットワークによる画像化処理 (中間層ユニット数=100): 1.24msec であった。ただし前処理は 1260 組のデータに対して独立した処理であるため、並列化することにより 4.06msec で処理することが可能となる。よって A/D 変換と画像の 3 次元表示を除くシステム全体の処理時間は 7.82msec となり、これはフレームレート 128.9fps に相当し、心臓などの拍動する臓器を対象とした計測に適応可能であると考えられる。

また本システムをより実用的なものへ拡張する場合について検討を行った。設計値として解像度 100×100 , フレームレート 30fps, 計測範囲を前方 $20[\text{mm}]$ 付近と仮定すると超音波エコーの往復時間から送受信素子数が 833 ($= \frac{1}{40[\mu\text{sec}] \times 30[\text{fps}]}$) となる。入力層ユニット数 833, 出力層ユニット 10000 とすると、入出力層のユニット数の比より現在のシステムでは解像度 $21 \times 21 (\approx 36 \times \frac{10000}{833})$ と同等になる。本研究で使用した計測システムは一辺が 4mm の矩形を基本単位とし $16\text{mm} \times 16\text{mm}$ を計測空間としたが、同様の空間で $21 \times 21[\text{pixel}]$ の解像度を必要となり、矩形の基本単位を $0.76[\text{mm}]$ に拡張することに相当する。この時、 $3772 (= 21^2 + 20^2 + \dots + 1^2 + 2(21 + 20 + \dots + 2))$ パターンの学習が必要となるが、実現の可能性についてはさらに議論する必要がある。使用する特徴量としてはさらに検討する必要があるが、ニューラルネットワークの規模が大きくなるにつれて学習に係るコストは膨大になり、サブネットに分割する等の工夫が必要となるであろう。

3.6 結言

第 3 章では、 4×4 のターゲットに対し、9.67MHz36 素子をもつトランスデューサにおける各素子を順次駆動し、それ以外のすべての素子で受信することにより、1260 個の時系列 RF (高周波) データを得た。時系列データ自体は情報量が多いため、エコーの最大ピーク位置を 1 つの特徴としてとして圧縮・表現した。したがって、ニューラルネットワークでは RF レベルでのホログラム的、もしくは開口合成的な処理ではなく、振幅・包絡線レベルでの処理がなされている。秋山ら [17] は同相加算による画像化を行っているが、本研究では各素子の送受信特性が異なるため適応が困難である。今後、開口合成的な処理を可能な限り取り入れることができるならば、走査時間短縮などの性能向上につなげることができるであろう。すなわち、位相制御で 2 次元的走査を行うフェーズドアレイ方式では、1 画面の走査にターゲットの分解能数 × エコー時間かかる。それに対して本方式で RF 信号をそのまま、もしくはそれからの抽出特徴を取り入れ、ホログラム的処理を行うことができるな

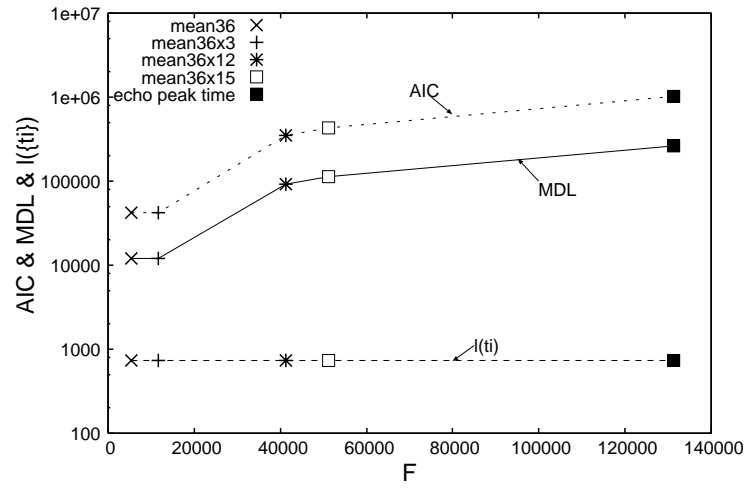


図 3.13: 2次元パターンを学習したネットワークの情報量基準による比較 (中間層ユニット数=100の場合)

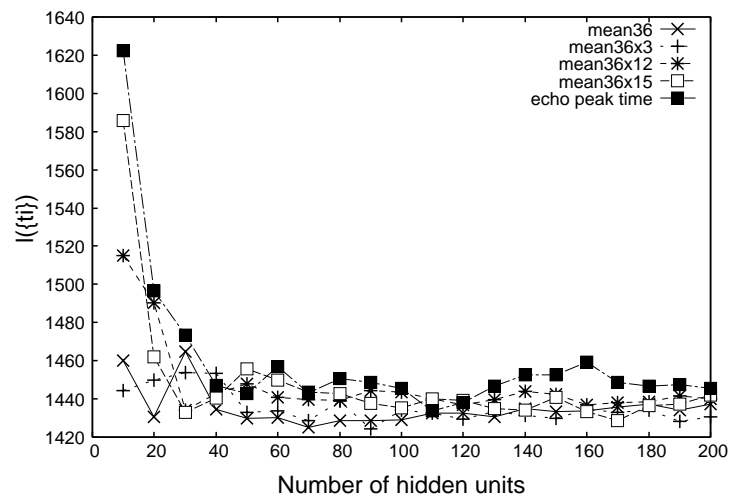


図 3.14: 中間層ユニット数の変化に対する $l(t_i)$ (学習: 2次元パターン, テスト: 3次元パターン)

らば、通常のフェーズドアレイのような送信時のビーム形成ではなく、受信後の後処理で反射波面の形状から対象を再構成することになり、基本的には1素子の駆動だけでよく、1×エコー反射時間で撮像が可能である(現在は素子数×エコーの反射時間で行っている)。眞溪ら [18] は8分割された直径2mmのリングプローブアレイ(振動子中心周波数10MHz)を試作し、プローブ前方の画像化を行っている。本研究で同様のプローブを使用した場合、56組(8×7)のデータを取得することが可能となるが、3.4より十分な次元数であると考えられる。得られた成果を要約すると、

- (1) 従来の内視鏡超音波診断装置が側方横断像を得るのに対して、前方の3次元像を実時間で得る超音波画像診断装置の基礎的な考察を行っている。
- (2) 奥行きをもつターゲットの3次元像について、文献 [10] より実用的な画像化手法について提案し、実験により有効性を確認した。
- (3) その際、実験システムに依存した入力パターンの特徴を利用することにより入力パターン空間の次元数を減少させ、ニューラルネットワークに汎化能力を保ったまま、入力パターン空間の次元数を大幅に減少させたネットワークの構築が可能となった。

となる。

以上、任意の環境と対象に対して使える広い汎化性を得る最終目標に対し、ある程度の汎化能力は得られることが確認できた。しかしながら、もっと別の特徴抽出も考慮して、よりいっそうの性能向上を図りたいと考えている。たとえば正方形だけではなく、生体の目における $\nabla^2 G$ のように、負の反応を持った基本パターンなども考えておく必要がある。これらを含め、任意の3次元形状がより正しく復元できるようできるだけ少数の基本パターンを探す必要がある。

参考文献

- [1] Pilkington TC and Ideker RE. von Ramm OT. National science foundation/engineering research center of emerging cardiovascular technologies. *Proceedings of the IEEE*, Vol. 81, No. 1, pp. 79–94, 1993.
- [2] Omer Oralkan, A.Sanli Ergun, and etc. Ching-Hsiang Cheng. Volumetric ultrasound imaging using 2-d cmut arrays. *IEEE Trans. ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL*, Vol. 50, No. 11, pp. 1581–1594, 2003.
- [3] Matthew P.Fronheiser and Edward D.Light et al. Real-time, 3-d ultrasound with multiple transducer arrays. *IEEE Trans. ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL*, Vol. 53, No. 1, pp. 100–105, 2006.
- [4] 吉澤信幸. 超音波センサとニューラルネットワークを用いた三次元物体認識. *EMC on.*, Vol. 53, pp. 98–104, 1992.
- [5] 湯浅肇, 石原知明, 高野宰, 他. 波面符合化送信方式による水中撮像装置. *信学技法*, Vol. US 94, No. 7, pp. 45–52, 1994.
- [6] 田村安孝. 計算機処理による超音波イメージング. *Med. Imag. Tech*, Vol. 17, No. 1, 1999.
- [7] M.Nambu, M.Doi, M.Matani, O.Oshiro, and K.Chihara. A high-speed image acquisition using ultrasonic ring array probe. *Computers in Cardiology*, Vol. 26, pp. 355–358, 1999.
- [8] Watanabe S and Yoneyama M. An ultrasonic visual sensor for three-dimensional object recognition using neural networks. *IEEE Trans Robotics Autom*, Vol. 8, No. 2, pp. 240–249, 1992.
- [9] 大多和寛, 河合秀夫, 佐藤嘉伸, 他. ニューラルネットワークと多素子超音波トランスデューサによる水中物体の映像化. *Med. Imag. Tech.*, Vol. 13, No. 5, pp. 753–761, 1995.
- [10] 大多和寛, 河合秀夫, 佐藤嘉伸, 他. ニューラルネットワークと多素子超音波トランスデューサによる水中物体の3次元画像化. *Med. Imag. Tech.*, Vol. 14, No. 5, pp. 561–570, 1996.
- [11] 竹谷尚, 大田和寛, 田村進一, 岡崎耕三. 円形単板圧電素子を分割した多素子超音波トランスデューサとニューラルネットワークによる水中物体の3次元画像化. *情報処理学会論文誌*, Vol. 48, No. SIG9, pp. 57–66, 2007.
- [12] 城戸健一, 他. 基礎音響工学, pp. 84–97. コロナ社, 1990.
- [13] 船橋賢一. ニューラルネットワークの capabilities について. *信学技法 MBE*, 1998.
- [14] H.Akaike. A new look at the statistical model identification. *IEEE Trans. Automatic Control*, Vol. AC-19, No. 6, pp. 716–723, 1974.

- [15] J. Rissanen. A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, Vol. 11, No. 2, pp. 416–431, 1983.
- [16] 栗田多喜男. 情報量基準よる 3 層ニューラルネットワークの隠れ層のユニット数の決定法. 電子情報通信学会論文誌 (D-II), Vol. J73-D-II, No. 11, pp. 1872–1878, 1990.
- [17] 秋山いわき, 矢野喜代志. 開口合成と逆投影による半球面の映像化. 日本超音波医学会基礎技術研究部会資料, Vol. BT96-4, pp. 9–12, 1997.
- [18] 眞溪歩, 藤本直登志, 南部雅幸, 他. 超音波リングプローブアレイを用いた前方 3 次元可視化. 電気学会論文集, Vol. 117-E-7, pp. 359–363, 1997.

第4章 ニューラルネットワークによるICウェハ の切り粉汚れ判定と撮影条件の決定

4.1 緒言

本章では、IC ウェハの製造工程で洗浄を必要とするかどうかを決定するために切り粉汚れの有無を判定する。この研究の目的は人間による目視検査を機械による検査に置き換えることである。人間の存在が製造工程で汚れの原因となるので、IC 製造ラインから人間を排除すると、ウェハの洗浄コストを削減することができる。また経験より、ウェハの切り粉汚れは影や色の变化に影響されることがあり、人間が常に色に関して正しい判断をするわけではない [1, 2]。このことは切り粉汚れ検査のエキスパートの養成が非常に困難であることを示している。

工業生産において、品質管理は最終的製品の信頼性を保障するために、正確な検査と歩留まりのマネジメントは非常に重要である。多くの場合、製品の目視検査は不可欠であるが、人間による検査はスピードが遅く、高価であり、ミスも多いため、目視検査の自動化 [1, 3] は望まれている。IC の製造工程で起こる汚れには洗浄不足による切り粉汚れと異物の混入によるものがある。不完全なパターンや印刷工程に欠陥を引き起こす IC ウェハの顕微鏡検査に関する研究には文献 [4, 5] がある。しかし、ウェハの切り粉汚れの検査については生産にかなり影響があり、IC 製造 [6] において検出しなければならない。

一方、照明条件については、一般に、画像処理および画像認識のためには、できるだけきれいな入力画像を撮影する必要がある。特に、IC ウェハの配線部を位置合わせ用テンプレートとして使用する場合には、アルミ表面のランダムな凹凸(ヒロッコ)による影響を受け易い。これらは一般的に画像中では雑音となり、適合位置においてもテンプレートと入力画像の相関値(Q 値)が低く、位置合わせに失敗する確率を大きくする。しかしながら、適切な照明を与えると、雑音が消え、大きな適合相関値を得ることが知られている。

ウェハの切り粉汚れはIC ウェハ配線部と関連しており目視検査の自動化のためには、その定量化の評価基準とともに照明条件を決定しなければならない。そこでニューラルネットワークを利用することにした。ここで、本研究での第一の目標は、汚れの度合いを計測することではなく、ウェハの汚れの有無を判断することであるため、切り粉汚れの有無についての分類手法を対象とする。本研究では切り粉汚れの有無のウェハのサンプルが入手可能であったので、本研究では教師あり学習による分類を選択することとした [7]。すなわち、ニューラルネットワークの学習で獲得した知識を活用する優れたパターン認識能力を利用する。

本研究のアプローチは、顕微鏡、光学フィルター、および CCD カメラ [8] を使用して、異なる照明条件下 [9–14] でいくつかの画像を撮影し、データの特徴空間を作成する。これらの画像の平均値と分散値は分類アルゴリズム [15] の特徴空間の値として使用される。4.2 ではデータ取得と特徴抽出法の詳細について述べる。4.3 では、バックプロパゲーションアルゴリズムによるニューラルネットワーク [16–19] による切り粉汚れの判定を行う。加えて最近傍法 [20] および最尤法 [7, 21] について比較検討を行い、バックニューラルネットワークが IC ウェハの切り粉汚れの有無の判定に有効であることを示す [22]。4.4 ではこれらの分類の実験結果について述べ、最も効果的な光学スペクトルと適切な照明条件を決定する。4.5 では本章のまとめを述べる。

4.2 撮影条件と特徴抽出

ここでは、落射照明と斜光照明を併用した照明条件と狭帯域干渉フィルターを使用して撮影された IC ウェハー画像の評価を行う。そして撮影条件と特徴抽出方法について提案する。

4.2.1 狭帯域干渉フィルタ

狭帯域干渉フィルターとはバンドパスフィルターであり、ある特定の周波数成分 (= 散乱成分) を取り出すことにより、画像中の S/N 比を減少させ認識を容易にすると考えられる。またこれらを用いて認識実験を行うことにより識別に必要な周波数成分が特定できる可能性がある。図 4.1 に実験に使用した狭帯域干渉フィルターの分光特性を示す。

ここではフィルターの効果を調べるために、照明条件を同一にし、フィルターのみを取り替えて撮影した画像に対し、以下の方法で評価を行った。使用したフィルターは 2,3, ..., 15 および R,G,B である。ただし filter1 は撮影に使用していない。また filter16, ..., 22 を使用した画像は非常に暗いため、使用していない。

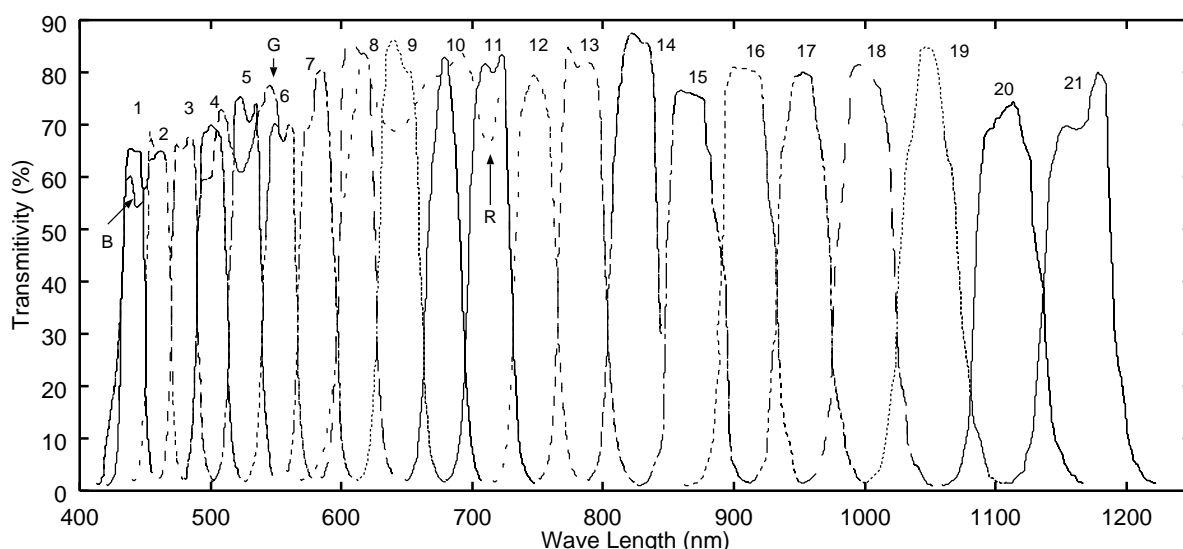


図 4.1: 狭帯域干渉フィルターの分光特性 (グラフ上段の数字はフィルタ番号を示す)

4.2.2 照明条件

- (1) 実験用画像の撮影では照明に落射 (最大光量の 50%) を用い、顕微鏡 (拡大率 10 倍) と CCD-Video カメラを使用して、ウェハー上の汚れ有りと思われる部分を各種のフィルターを用いて撮影を行なった。フィルターを通した画像は 640×480 pixel の大きさの画像として保存した。各画素は 256 階調である。撮影条件の異なる画像例を図 4.2 に示す。この例では同一部位を撮影しているにも関わらず、フィルタ “2” を使用して撮影した画像では洗浄不足部分のテクスチャが確認できる。
- (2) 各々のフィルターを施した画像に対して、画像の背景領域から明るい領域と暗い領域との 2 つの矩形領域 (サイズ: 32×32 pixel) を選択し、これらの領域の画素値の平均を計算する。画像中の明るい領域はウェ

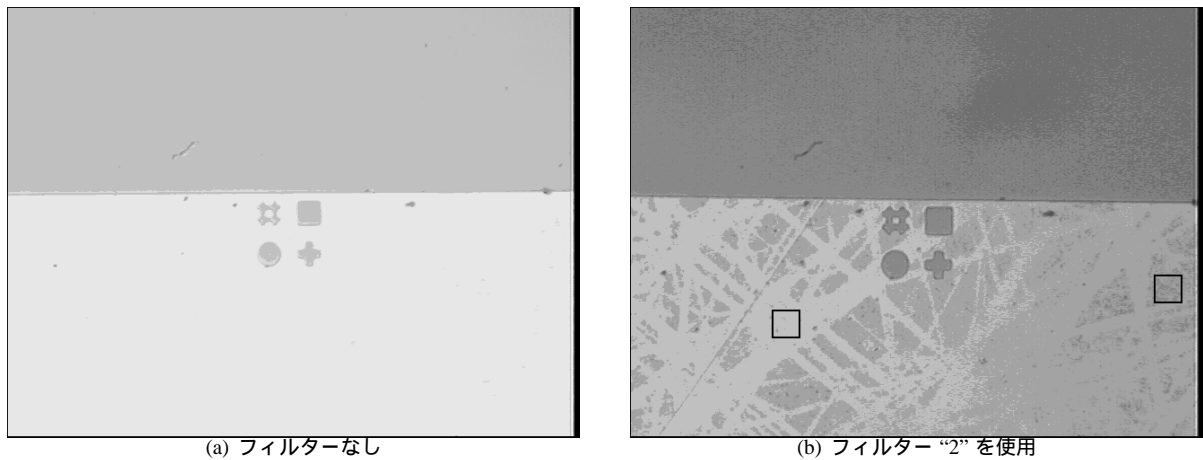


図 4.2: 撮影条件の異なる画像例 (汚れあり)

ウェハ自体の信号 (W) を示し、暗い領域は信号成分が切粉 (ノイズ) により低減された ($W - C$) を表している。ただし C は切り粉汚れのある領域の平均値とした。図 4.2(b) にこれらの領域を示す。右の矩形がウェハ自体の信号領域、左の矩形が切り粉による汚れた領域である。

- (3) 各種のフィルターを用いて撮影された画像について $\frac{W-(W-C)}{W}$ を計算する。得られた値 $\frac{C}{W}$ は最適なフィルターを決定するために使用される。 $\frac{C}{W}$ が大きいことは汚れ (ノイズ) を良く表現しており、より大きな $\frac{C}{W}$ を与えるフィルターが適していると考えられる。
- (4) (2),(3) の処理を矩形領域サイズ 64×64 pixel および 8×8 pixel についても行なった。これらの結果を図 4.3 に示す (図 4.3 では R,G,B filter はそれぞれ “Optical filter No.”, 16,17,18 に対応する)。図 4.3 より領域サイズ 32×32 pixel を使用することとした。
- (5) 図 4.3 より filter2 から最も大きい $\frac{C}{W}$ が得られている。filter3-5 から十分な結果が得られている。filter17(G) および 18(B) と比較して filter16(R) の値が低いことは切り粉汚れ成分が赤色成分を持つことを示している。
- (6) 切り粉汚れありおよび汚れなしのウェハを落射照明のみおよび 45° からの斜光照明を用いて撮影した (図 4.4)。これより斜光照明が効果的であることがわかる。ただし、斜光は切り粉汚れによる散乱成分であるため、画像を形成する光の周波数特性は平坦 (白色) である。したがって光学フィルターの使用は効果的でないと考えられる。

4.2.3 データ抽出と特徴抽出

図 4.3 より切り粉汚れに対して filter2($450 < \lambda < 470(\text{nm}), \lambda$: 波長) が最も効果的であった。これよりフィルターに関しては

- (1) フィルターなし
- (2) $0 < \lambda < 450(\text{nm})$
- (3) $0 < \lambda < 470(\text{nm})$

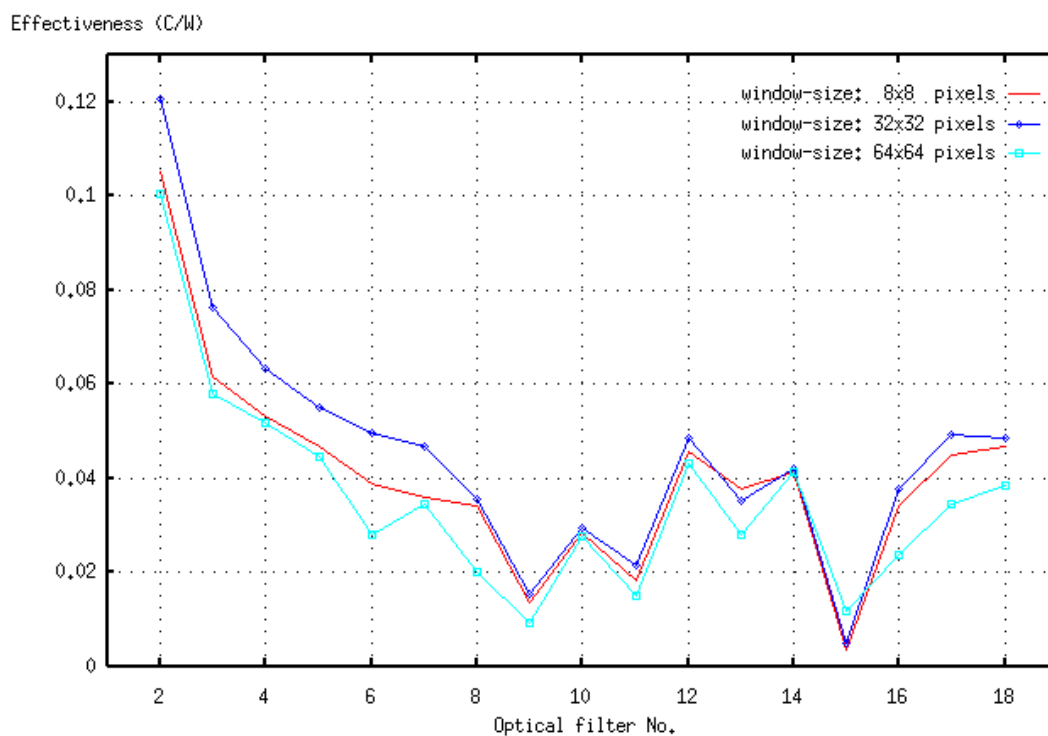


図 4.3: 狭帯域干渉フィルターの効果

(4) $450 < \lambda < 470(\text{nm})$

の 4 種について検討を行う。

落射光量では最大光量を 100%とした場合

- (1) 0%
- (2) 70%
- (3) 100%

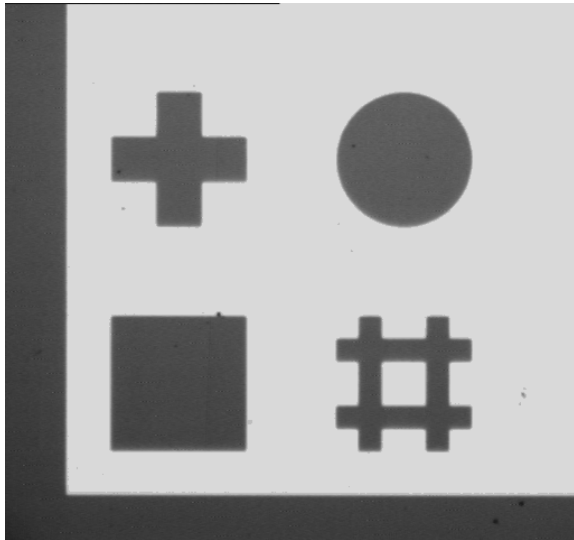
の 3 種について検討を行った。

斜光量では最大光量を 100%とした場合

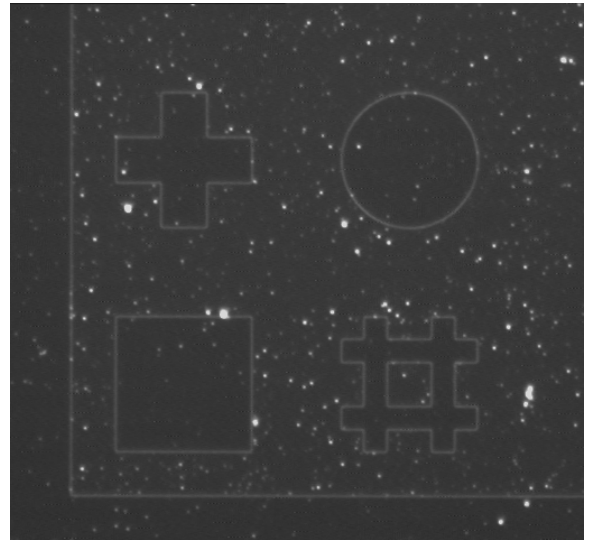
- (1) 0%
- (2) 70%
- (3) 100%

の 3 種について検討を行った。

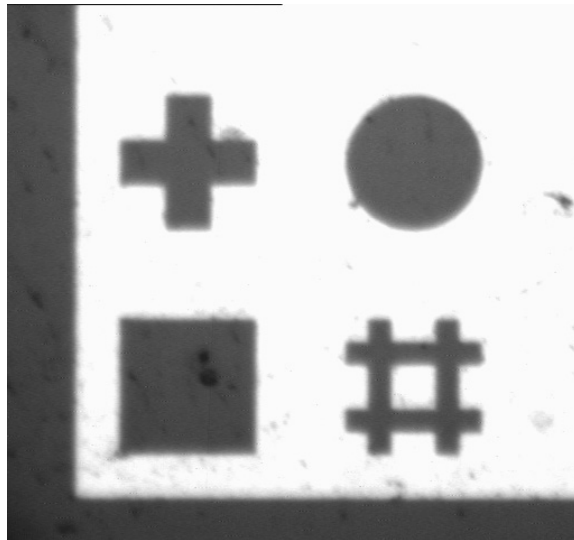
これらの組み合わせにより、1つのサンプルにつき4(フィルター種) \times 3(落射光量) \times 3(射光量)=36種の画像を撮影した。ただし、落射、射光とも0%の場合は照明のない画像となるため、実験で使用された画像は32種である。各濃淡画像の大きさは512 \times 480 pixelで各画素256階調である。汚れなしウェハーの場合と同様に同じ



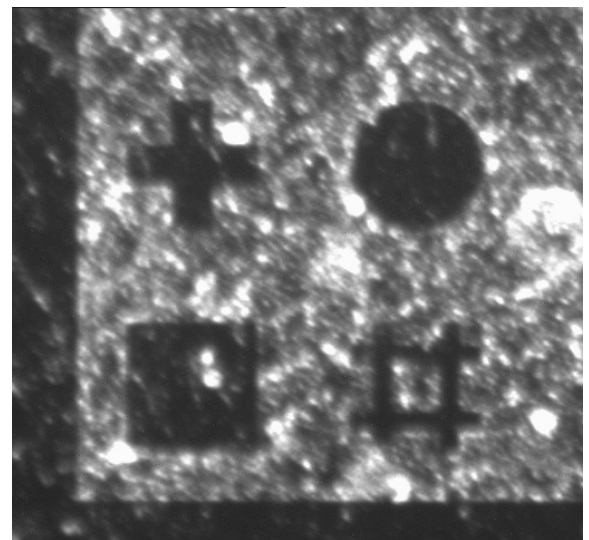
(a) 落射照明のみ (汚れなし)



(b) 斜光照明のみ (汚れなし)



(c) 落射照明のみ (汚れあり)



(d) 斜光照明のみ (汚れあり)

図 4.4: 斜光照明の効果

パターンを持つ領域を汚れありウェハから選択し、撮影を行った。図 4.5 に矩形領域 (32×32pixel) の選択例を示す。

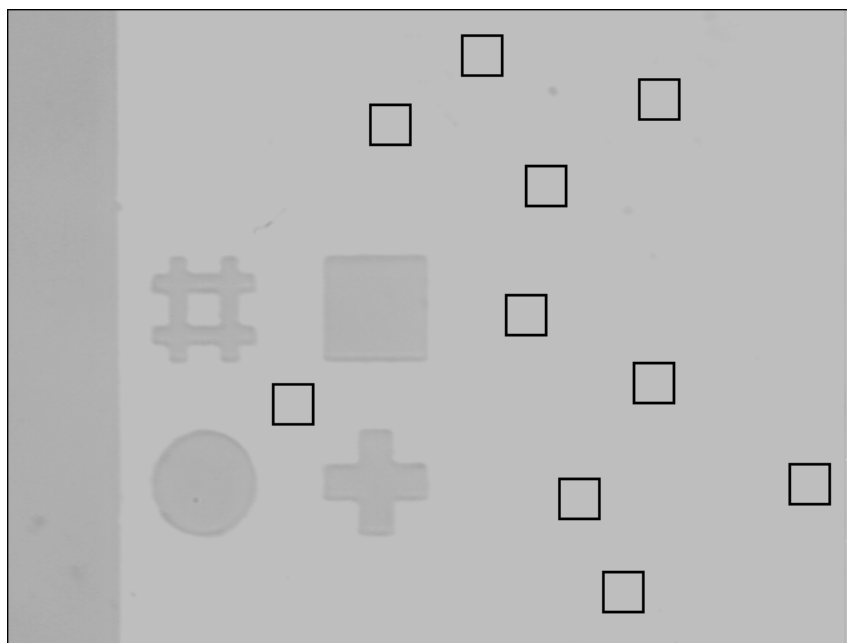


図 4.5: 矩形領域の選択例 (領域数 M=10 の場合)

Mヶ所の矩形領域の画素値の平均と分散を求め [15]、以下の式により正規化を行った。

$$m_n(i) = \frac{m(i) - m_{min}}{m_{max} - m_{min}} \quad (4.1)$$

$$v_n(i) = \frac{\alpha v(i)}{m_{max} - m_{min}} \quad (4.2)$$

ただし、 $m(i)$ は画像 i の正規化されていない平均、 $v(i)$ は同じく分散である。 $m_n(i)$ は画像 i ($i = 1, \dots, 32$) の正規化された平均、 $v_n(i)$ は同じく分散を示す。 m_{min} は 32 枚の画像に対する平均の最小値、 m_{max} は同じく最大値を示し、 α は照明条件に依存した係数である。正規化された 32 種の平均と分散を次節で述べる分類手法の入力データとして使用される。

4.3 分類手法

本研究での第一の目標は、汚れの度合いを計測することではなく、ウェハの汚れの有無を判断することであるため、切り粉汚れの有無についての分類手法を対象とする。本研究では切り粉汚れの有無のウェハのサンプルが入手可能であったので、本研究では教師あり学習による分類を選択することとした [7]。

本研究では従来のパターン認識手法より、統計的なパターン認識手法およびニューラルネットワークによる分類を適用することとした。これらの 2 種のアプローチの特徴は文献 [21] で述べられており、ウェハの切り粉汚れの判定では、利用可能なデータセットに関して具体的な統計的仮定を持つことができなかった。これにより、本研究では ANN 分類手法を適応することとした。加えて、統計的なパターン認識手法として、最短距離法および最尤法を用いた分類を行うこととした。これらの分類手法は 4.4 で実験的に評価を行う。

4.3.1 ニューラルネットワークによる分類

図 4.6 に分類に使用される 3 層構造のフィードフォワードニューラルネットワーク [23] を示す。 L_I, L_H, L_O はそれぞれ入力層, 中間層, および出力層を示し, M, N, P はそれぞれ入力層ユニット数, 中間層ユニット数, 出力層ユニット数を示している。

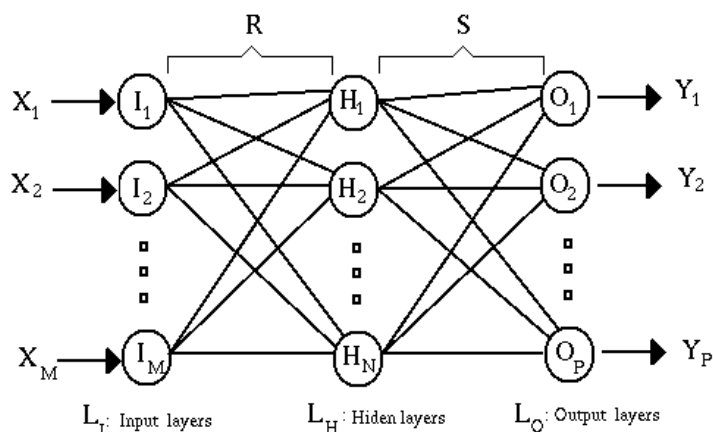


図 4.6: フィードフォワードニューラルネットワーク

$R = |r_{ij}| (i = 1, 2, \dots, M; j = 1, 2, \dots, N)$ は $M \times N$ の行列を示し, r_{ij} は入力層ユニット I_i から中間層ユニット H_j への結合ウェイトを表している。 $S = |s_{ij}| (i = 1, 2, \dots, N; j = 1, 2, \dots, P)$ は $N \times P$ の行列を示しており, s_{ij} は中間層ユニット H_i から出力層ユニット O_j へ結合ウェイトを表している。 $X = (X_1, X_2, \dots, X_M), Y = (Y_1, Y_2, \dots, Y_P)$ はそれぞれ入力ベクトルおよび出力ベクトルを示している。図 4.6 に示される中間層および出力層のすべてのユニットはシグモイド関数を通して出力される [19]。

認識に使用したニューラルネットワークの構成は入力層ユニット 64, 中間層ユニット 3, 出力層ユニット 2 とした。出力層の各ユニットはそれぞれ“汚れなし ($Y_1 = 1, Y_2 = 0$)”, “汚れあり ($Y_1 = 0, Y_2 = 1$)”に対応させ学習を行った。学習には 3 種のサンプル wafer から得られたパターンを用い, 残り 1 種のサンプル wafer から得たパターンを未学習パターンとして評価に用いた。学習にはバックプロパゲーションアルゴリズム [23] を使用し, いずれの場合も誤認識がない状態で終了している。正解の判定は

- $Y_1 - Y_2 > 0.7$ であれば“汚れなし”
- $Y_1 - Y_2 < -0.7$ であれば“汚れあり”
- 上記のどちらにも該当しなければ“不明”

とした。

4.3.2 最短距離法

ここではノンパラメトリックな学習方法である最短距離法について述べる [20, 24]。 M 個の要素を持つ特徴ベクトル $X = (X_1, X_2, \dots, X_M)$ を仮定し, これらを N 個のクラス C_1, C_2, \dots, C_N に分類することを考える。

ただし，クラス C_i の学習パターン数は K_i とする．クラス C_i の平均ベクトルは次式で求めることができる．

$$\overline{X(i)} = \frac{1}{K_i} \sum_{j=1}^{K_i} X(i, j)$$

ただし， $X(i, j)$ はクラス $C_i (i = 1, 2, \dots, N)$ の j 番目のパターンを表す．すべてのクラスの平均ベクトルを計算しておく．

分類では，次式により未知入力特徴ベクトル X とすべてクラスの平均特徴ベクトルのユークリッド距離を計算する．

$$D_i = \|X - \overline{X(i)}\| (i = 1, 2, \dots, N)$$

ここで D_i が最小となるクラス i を分類結果とする．4.2.3 と 4.3.1 で述べた特徴ベクトル ($M = 64$) を最短距離法で分類する．ただし，分類は“汚れなし”，“汚れあり”の2クラスとなる．

4.3.3 最尤法

パラメトリックな学習による分類手法として代表的な最尤法を取り上げる．ベイズの定理として次式が知られている．

$$P(C_i|X) = \frac{p(X|C_i)P(C_i)}{p(X)}, \quad (4.3)$$

ただし， X は特徴ベクトル， C_i は i 番目のクラスを示し， $P(C_i)$ は i 番目のクラスの生起確率を示し， $p(X)$ は観察された特徴ベクトルの確率密度関数を示す．また $P(C_i|X)$ はクラス C_i に分類されるサンプルを与える特徴ベクトル X の事後確率を示し， $p(X|C_i)$ は観察された特徴ベクトルを与える i 番目のクラスに分類されるサンプルのための確率密度関数を示している．式 (4.3) より， X が与えられると $p(X)$ が確定する．事前確率 $p(C_i|X)$ [24] を最大にするために $p(X|C_i)P(C_i)$ を最大にする必要がある． $P(C_i) = 1/N$ (N はカテゴリ数) であれば $p(X|C_i)$ が計算された確率の最大値であれば，特徴ベクトル X がカテゴリ C_i に属していることがわかる．加えて特徴パラメータが連続した結合分布 [25] であれば， $p(X|C_i)$ は対応する確率密度関数に置き換えられる．最終的に $p(X|C_i)$ は次式で表すことができる．最終的に多変量正規密度関数 $p(X|C_i)$ は次式で示される．

$$p(X|C_i) = \frac{1}{(2\pi)^{K/2} |\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(X - M_i)^T \Sigma_i^{-1} (X - M_i)\right], \quad (4.4)$$

ただし， $i = 1, 2, \dots, N$ ， N は分類するカテゴリ数， K は特徴ベクトルの次元数， $M_i = (M_{1i}, M_{2i}, \dots, M_{Ki})$ はカテゴリ i の特徴ベクトルの平均， Σ_i はカテゴリ i に対応する特徴要素の共分散マトリックスである．4.2.3 で示した通り，撮影条件を変えて撮影した 32 枚のウェハー画像から 32 個の平均値 ($M_{1j}, M_{2j}, \dots, M_{32j}$) と同数の分散値 ($\sigma_{1j}^2, \sigma_{2j}^2, \dots, \sigma_{32j}^2$) を算出した．簡単化のために各特徴要素が相関がないと仮定し，これより共分散行列は $\Sigma_i = \sigma_{1j}^2, \sigma_{2j}^2, \dots, \sigma_{32j}^2$ によって表される．現在，トレーニング用として“汚れなし”のサンプル S_1 ，“汚れあり”のサンプル S_2 が入手可能であり，これらのすべてのサンプル ($S_1 + S_2$) が式 (4.4) の分布を持つと考えた．これは S_1 および S_2 のパラメータが正規分布にしたがっていると仮定して算出される．分類フェーズでは疑わしきウェハーの特徴ベクトルを生成した後，すべてのサンプルより式 (4.4) を求め，最大値が“汚れなし”カテゴリーであれば“汚れなし”と判定し，そうでなければ“汚れあり”と判定する．

4.4 実験結果

ここでは，ニューラルネットワーク，最短距離法および最尤法による分類の性能を評価し，IC ウェハーの切り粉汚れの自動判定に最適な光学フィルターと落射照明と斜光照明の条件を選択する．

4.4.1 分類手法の評価

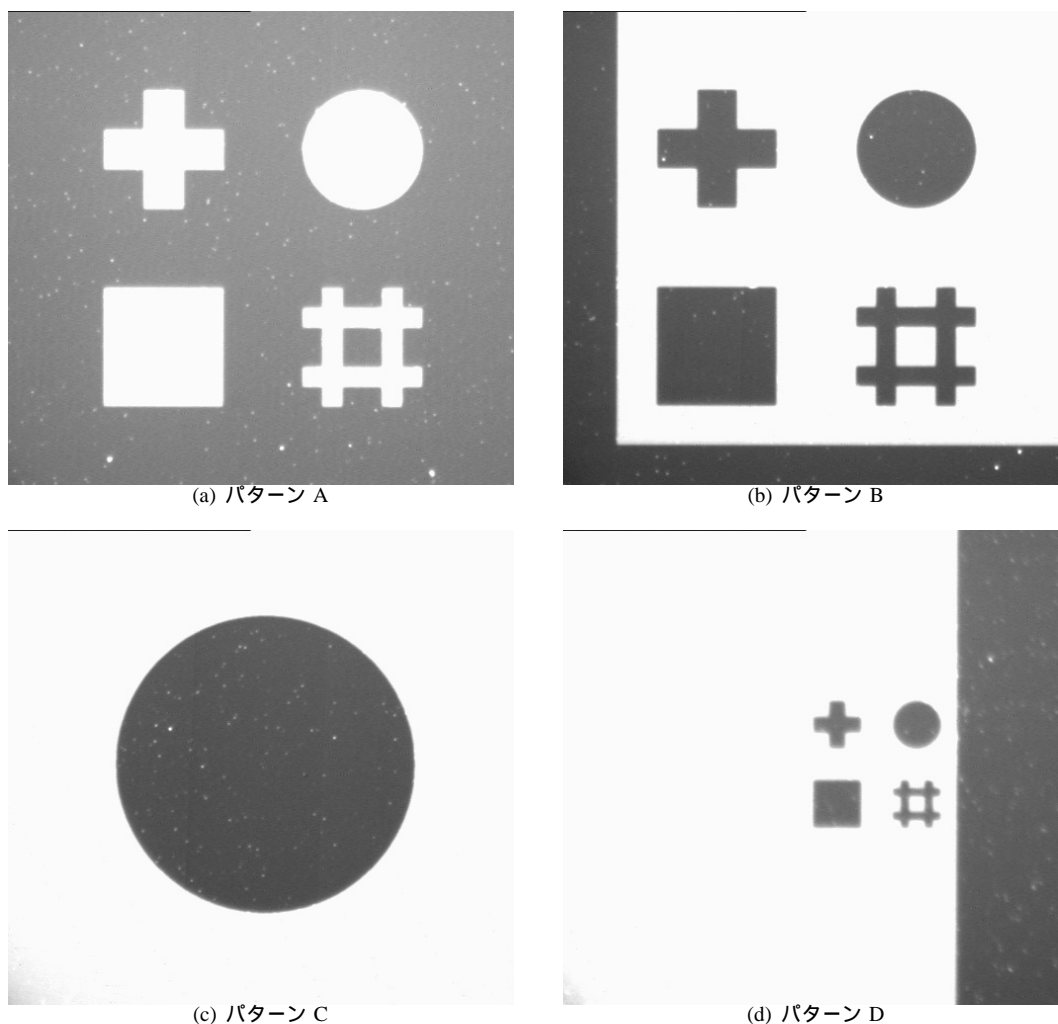


図 4.7: 実験に使用したウェハパターン

図 4.7 は実験に使用される 4 つの IC ウェハのテストパターンである。これらの 4 種のパターンに、“汚れなし” および “汚れあり” のウェハを用意した。これらの 8 枚 (4 × 2) のウェハから、白色部分のみ、黒色部分のみおよびテクスチャを含む部分がサンプルとして選択された。図 4.8 にパターン D(図 4.7(d)) に対して矩形領域を設定した例を示す。ここでは合計 24 種 = 4(ウェハ種) × 2(汚れの有無) × 3(白色, 黒色およびテクスチャ部分) のサンプルを用いて以下の手順により分類手法の評価を行った。

実験 1 24 種のサンプルにより学習を行い、学習に使用されるサンプルのみで分類に精度を評価する。

実験 2 パターン A,B および C の 18 種のサンプルにより学習を行い、パターン D の 6 種のサンプルを用いて評価を行う。

実験 3 パターン A,B および D の 18 種のサンプルにより学習を行い、パターン C の 6 種のサンプルを用いて評価を行う。

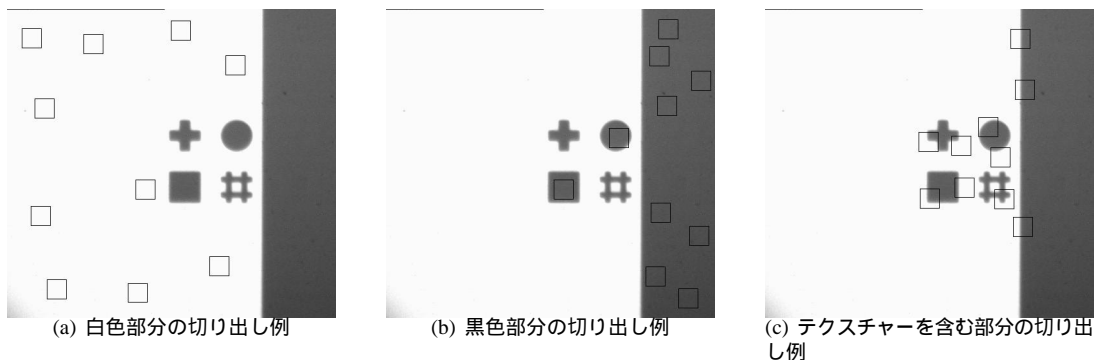


図 4.8: 図 4.7(d) に対する矩形領域の切出し例

表 4.1: 分類手法の比較

実験	正解率 (%)		
	最短距離法	最尤法	BP-ANN
1	87.5	87.5	100.0
2	100.0	66.7	100.0
3	66.7	83.3	83.3
4	100.0	100.0	100.0
5	83.3	100.0	100.0
平均	87.5	87.5	95.8

実験 4 パターン A,C および D の 18 種のサンプルにより学習を行い, パターン B の 6 種のサンプルを用いて評価を行う.

実験 5 パターン B,C および D の 18 種のサンプルにより学習を行い, パターン A の 6 種のサンプルを用いて評価を行う.

使用したニューラルネットワークは, 3 層構造で入力層ユニット数 64, 出力層ユニット数 2, 中間層ユニット数は 5 とした. 実験により中間層ユニット数は増減はネットワークのパフォーマンスに影響しないことがわかった. ニューラルネットワークおよび最短距離法による分類では正規化された矩形領域の画像平均と分散が使用される式 (4.1,4.2). 最尤法式 (4.4) では正規化前の値が使用されている. 上記の 5 種のデータに対して BP-ANN, 最短距離法および最尤法を用いて分類を行った (表 4.1). 表よりニューラルネットワークによる分類が IC ウェハー切り粉汚れの検査において適切であることがわかった.

4.4.2 分類に有効な特徴量

本章では 32 種の条件のうち, 特定のものを削除した状態でネットワークの学習および認識実験を行った. この時, 認識率の低下が確認されれば削除した条件は認識に必要なものである可能性が示唆される. また条件が削除された状態で認識率の低下がなければ, その条件は自動認識にあまり有効でないと考え, 以下の条件で実

験を行った．前述の実験では入力パターンベクトルとして 64 次元のパターンベクトルを用いたが，ここでは 62 次元のパターンベクトルを用いてニューラルネットワークの学習，認識実験を行う．すなわち 32 種の撮影条件のうち，1 種の条件を除いたパターンを学習パターンとした．表 4.2 に実験結果を示す．表 4.3 に使用したフィルターの特性を示す．表 4.2 において，たとえば 29 の画像を削除する場合，光学フィルターは No.1，落射照明は 0%，斜光照明は 70% で撮影した画像を除く 31 枚の画像より得た特徴量を用いて分類を行った．この結果より各 1 種の条件を除いた場合の認識率には優位な差が得られていない．これは 62 次元の入力パターンベクトルが，削除されたベクトル (撮影条件) を補い認識率の低下が確認できなかったと考えられる．

次に光学フィルター，落射照明および斜光照明の最適条件を決定するために実験を行った．表 4.4 に実験に使用した画像の組合わせを示す．ここではフィルターや照明条件の差を明確にするために，撮影条件の同じ画像を集めて認識実験を行った．実験結果を図 4.9 に示す．図 4.9(a) では落射光量が 70% で認識率が最大を示し，図 4.9(b) では斜光量が 100% で認識率が最大となっている．図 4.9(c) では $450 < \lambda < 470(\text{nm})$ で認識率が最大となっており，認識に最適な条件を得ることができた．

4.5 結言

本章では IC ウェハーの切り粉汚れの自動検査を行う場合の分類手法および最適な照明条件に検討を行った．同一条件で 3 種の分類手法を対象に評価を行ったが，ニューラルネットワークによる分類が平均認識率 95.8% となり，切り粉汚れ検出の自動化には有効であることがわかった．

第 4 章の成果を以下に示す．

- (1) 同一条件で 3 種の分類手法を対象に評価を行ったが，ニューラルネットワークによる分類が平均正解率 95.8% となり，切り粉汚れの自動検査には有効であることがわかった．
- (2) ニューラルネットワークによる IC ウェハーの切り粉汚れの自動判定に効果的な照明条件は落射照明 70%，斜光照明 100% であることがわかった．
- (3) 切り粉汚れに有効な周波数成分は $450 < \lambda < 470(\text{nm})$ であるため，これに対応する光学フィルターを使用することにより認識精度の向上が図れる．

今後の課題として本章で開発した手法を拡張して IC ウェハーの切り粉汚れのレベルを計測することにより IC ウェハーを分類することである．

表 4.2: 各画像の撮影条件とそれ以外の画像での認識率

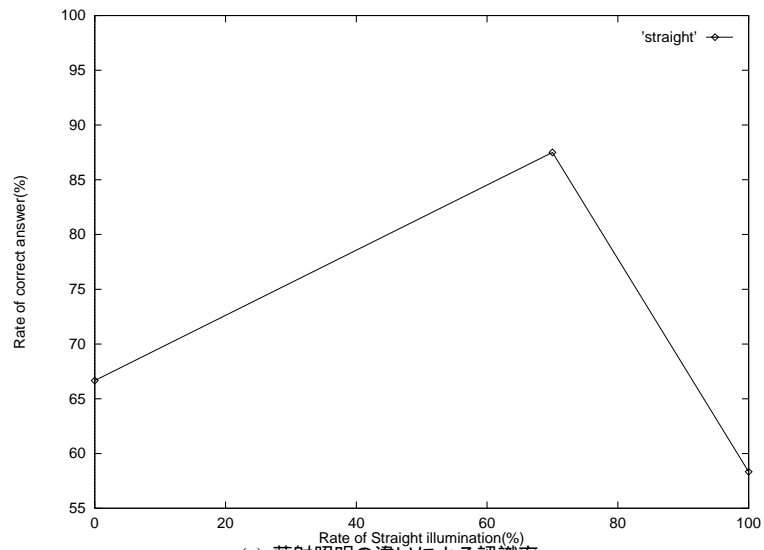
画像番号	撮影条件			BP-ANN 正解率 (%)
	フィルター	落射 (%)	斜光 (%)	
1	1	100	100	95.8
2	2	100	100	91.7
3	3	100	100	87.5
4	4	100	100	91.7
5	1	100	70	95.8
6	2	100	70	95.8
7	3	100	70	95.8
8	4	100	70	83.3
9	1	100	0	91.7
10	2	100	0	87.5
11	3	100	0	91.7
12	4	100	0	83.3
13	1	70	100	91.7
14	2	70	100	95.8
15	3	70	100	91.7
16	4	70	100	87.5
17	1	70	70	91.7
18	2	70	70	83.3
19	3	70	70	91.7
20	4	70	70	91.7
21	1	70	0	87.5
22	2	70	0	87.5
23	3	70	0	95.8
24	4	70	0	95.8
25	1	0	100	83.3
26	2	0	100	87.5
27	3	0	100	87.5
28	4	0	100	83.3
29	1	0	70	79.2
30	2	0	70	83.3
31	3	0	70	87.5
32	4	0	70	83.3

表 4.3: 図 4.9(c) の光学フィルター条件

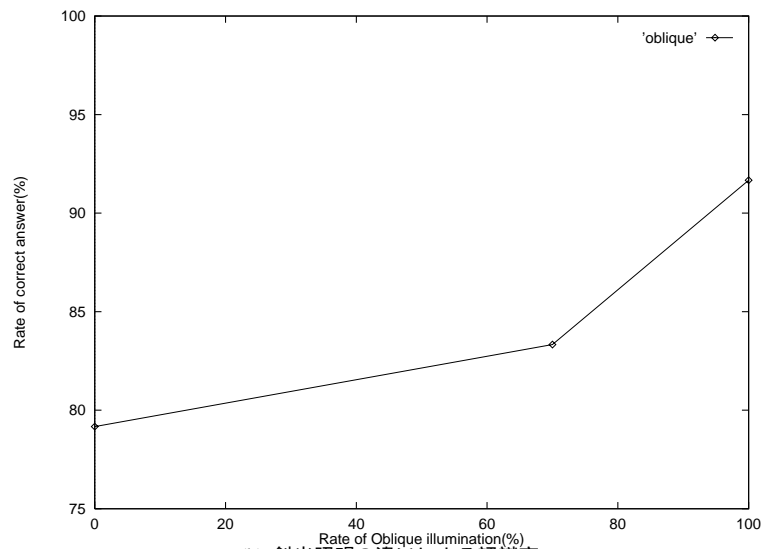
フィルター番号	透過周波数 (λ)
1	no filter used
2	$0 < \lambda < 450[nm]$
3	$0 < \lambda < 470[nm]$
4	$450 < \lambda < 470[nm]$

表 4.4: 図 4.9 の実験に使用された画像

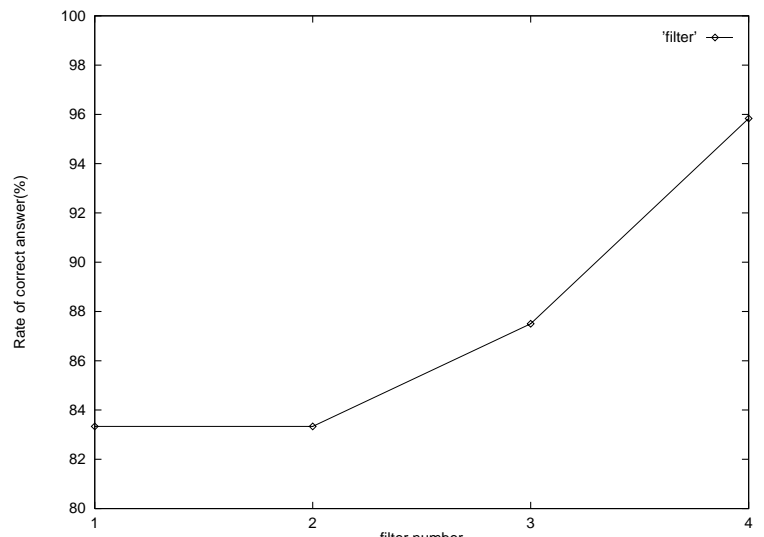
撮影条件		入力画像数	使用された画像番号 (表 4.2)
フィルター	1	8	1,5,9,13,17,21,25,29
	2	8	2,6,10,14,18,22,26,30
	3	8	3,7,11,15,19,23,27,31
	4	8	4,8,12,16,20,24,28,32
落射 (%)	100	12	1,2,3,4,5,6,7,8,9,10,11,12
	70	12	13,14,15,16,17,18,19,20,21,22,23,24
	0	8	24,25,26,27,28,29,30,31,32
斜光 (%)	100	12	1,2,3,4,13,14,15,16,25,26,27,28
	70	12	5,6,7,8,17,18,19,20,29,30,31,32
	0	8	9,10,11,12,21,22,23,24



(a) 落射照明の違いによる認識率



(b) 斜光照明の違いによる認識率



(c) 光学フィルターの違いによる認識率 (横軸は表 4.3 の “Filter no.” に対応する)

図 4.9: ニューラルネットワークによる認識率

参考文献

- [1] T.S. Newman and A.K. Jain. A survey of automated visual inspection. *Comput. Vis. Image Understanding*, Vol. 61, pp. 231–262, Nov. 1995.
- [2] W. Daley and T. Rao. Color vision for industrial inspection. *Proc. of SME Vision Conf. '9*, pp. 12–24, Nov. 1990.
- [3] R.T. Chin and C.A. Harlow. Automated visual inspection: a survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-4, No. 6, pp. 557–573, 1982.
- [4] X.L. Xie and G. Beni. A validity measure for fuzzy clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 841–847, Nov. 1990.
- [5] M. Nikoonahad, C.E. Wayman, and S.A. Biellak. Defect detection algorithm for wafer inspection based on laser scanning. *IEEE Trans. on Semiconductor Manufacturing*, Vol. 10, No. 4, pp. 459–468, Nov. 1997.
- [6] San Jose. *The National Technology Road-map for Semiconductors*. Semiconductor Industry Association, 1994.
- [7] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, 1996.
- [8] G.E. Healey and R. Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 3, pp. 267–276, Mar. 1994.
- [9] D.W. Capson and S. Eng. A tiered-color illumination approach for machine inspection of solder joints. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 3, pp. 387–393, May 1988.
- [10] E.J. Breneman H. Lee and C.P. Schulte. Modeling light reflection for computer color vision,. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 4, pp. 402–409, Apr. 1994.
- [11] B.V. Funt J. Ho and M.S. Drew. Separating a color signal into illumination and surface reflectance components: theory and application. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 10, pp. 966–977, Oct. 1990.
- [12] H. Murase and S.K. Nayar. Illumination planning for object recognition using parametric eigenspace. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 12, pp. 1219–1227, Dec. 1994.
- [13] D. Patel, E.R. Davies, and I. Hannah. Color constancy under varying illumination. *Fifth Int. Conf. on Computer Vision*, pp. 720–725, June 1995.
- [14] Y. Hara, H. Doi, K. Karasaki, and T. Iida. A system for pcb automated inspection using fluorescent light. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 1, pp. 69–78, Jan. 1988.
- [15] V.D. Nguyen, A. Noble, and J. Mundy et al. Exhaustive detection of manufacturing flaws as abnormalities. *Proc. of IEEE-CVPR*, pp. 945–952, 1998.

- [16] M.T. Musavi, K.H. Chan, D.M. Hummels, and K. Kalantri. On the generalization ability of neural network classifiers. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, pp. 659–663, Jun. 1994.
- [17] T. Kohonen. An introduction to neural computing. *Neural Networks*, Vol. 1, pp. 3–16, 1988.
- [18] T. Kohonen. Neural networks for computing ? *Proc. of AIP Conf.: 151*, pp. 1–6, 1986.
- [19] B. Widrow and M. Lehr. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. *Proc. IEEE*, Vol. 78, No. 8, pp. 1415–1441, Sep. 1990.
- [20] S.L. Bartlett, P.J. Besl, and C.L. Cole. Automatic solder joint inspection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 1, pp. 31–43, Jan. 1988.
- [21] I.K. Sethi and A.K. Jain. *Artificial Neural Networks and Statistical Pattern Recognition*, pp. 11–31. Elsevier Science Publishers, New York, 1991.
- [22] Reza A. Zoroofi, Hisashi Taketani, Shinichi Tamura, Yoshinobu Sato, and Kazuma Sekiya. Automated inspection of ic wafer contamination. *Pattern Recognition*, Vol. 34, pp. 1307–1317, 2001.
- [23] J. Rogers. *Object-Oriented Neural Networks in C++*. Academic Press, San Diego, 1997.
- [24] J.T. Tou and R.C. Gonzales. *Pattern Recognition Principles*. Addison-Wesley, 1974.
- [25] R. Duda and O. Hart. *Pattern Classification and Scene Analysis*. Wile, New York, 1973.

第5章 ニューラルネットワークによる膝蓋骨亜脱臼の自動識別

5.1 緒言

臨床診断においては、医用画像の読影診断のため、医師に対して長期にわたる教育と訓練が必要である。したがって、画像診断の自動化に大きな期待がかけられている。一方、目視による診断精度を高めるため、各種の画像処理技術が導入された。これらの処理は、いずれも病変や異常の認識を容易にすることである。

膝蓋骨亜脱臼の診断においては関節造影 CT、MRI などを用いて、医師の読影による膝蓋骨損傷の診断が試みられている [1, 2]。

図 5.1 は、関節造影 CT 画像の例である。左膝屈曲角度 15° の状態で蓋骨にほぼ垂直に膝関節を撮影されたものである。(a) においては、蓋骨が膝の関節のほぼ中央に位置し正常である。(b) においては、蓋骨が膝の関節の中央の位置からずれており異常である。膝蓋骨亜脱臼の診断において医師は、患者の膝の外形輪郭 (図 5.1(c),(d)) から診断できるが、コンピュータによる正常および異常の診断が出来ればその利便性は大きい。図 5.1 において、膝蓋骨亜脱臼の目視による判定は容易に行えるが、コンピュータで行う場合、その評価基準をどのように決定するかが困難になる。例えば、膝蓋骨に個人差があるばかりでなくどこまで正常・異常の判定をするかである。ここでは専門医による確定診断を教師信号としてニューラルネットワークを自動識別の可能性について検討を行った。

本研究では、膝の CT 画像を対象とし、画像認識の第一段階として、最近さまざまな分野で研究が盛んになってきたニューラルネットワークによる膝蓋骨亜脱臼の自動診断の可能性について考察する。ニューラルネットワークは、情報環境に適応して知識を蓄積し、自律的に演算能力を発達させる情報処理システムであり、アルゴリズムや規則の開発を必要としない。医用画像分野におけるニューラルネットワークの応用例はいくつか報告されている。たとえば、われわれの研究グループでは円環状超音波トランスジューサのエコーバックをバックプロパゲーション・ニューラルネットワーク (BPN) で処理して、ターゲット位置の識別を行っている [3]。また、コントラストの悪いヒトの角膜上皮細胞を対象とし、適合度を表す関数の最小化によりその不完全な六角状細胞壁を補完・認識している [4]。

本研究では確定診断つきの膝の CT 画像を対象としてニューラルネットワークによる膝蓋骨亜脱臼の判定を行った [5]。その結果、膝蓋骨亜脱臼のニューラルネットワークによる自動診断の可能性を示せたのではないかと考えている。

5.2 対象と画像処理方法

膝蓋骨亜脱臼診断用の画像はすべて大阪大学医学部附属病院整形外科を受診した患者から撮影した CT 画像である。使用した画像はいずれも左膝屈曲角度 15° 状態である。画像の枚数は正常者のもの 22 枚、異常者のもの 66 枚である。膝画像は、CCD カメラから日本アビオニクス製画像フレームメモリ EXCEL TVIP - 4100 に入力され、 $128 \times 128 \times 8$ ビットデータとなる。入力された CT 原画像の例を図 5.2、図 5.3 に示す。

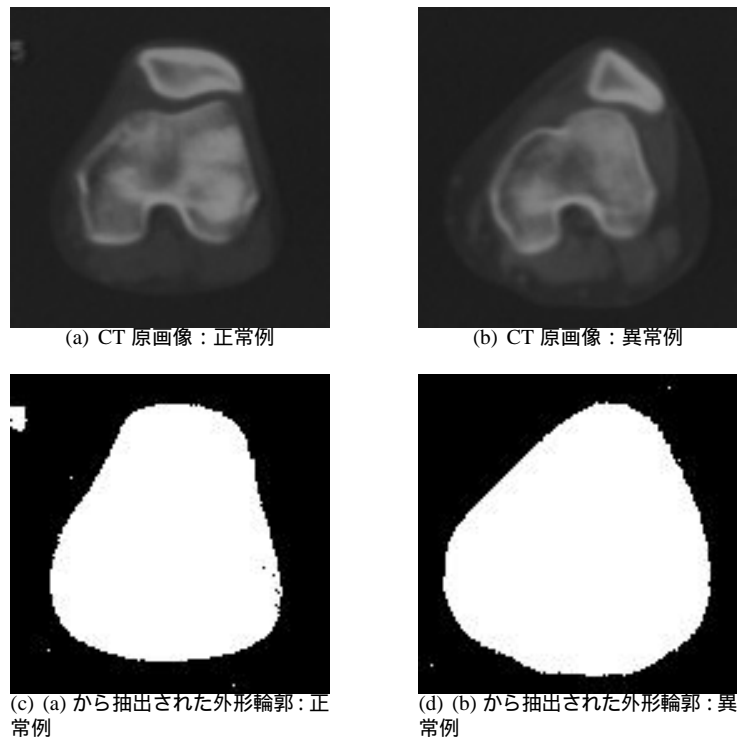


図 5.1: 膝蓋骨亜脱臼診断

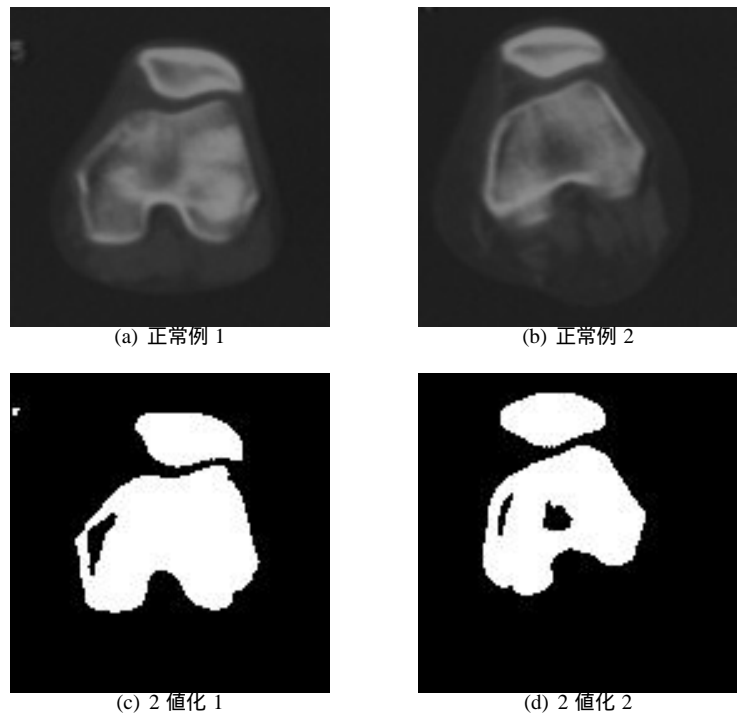


図 5.2: CT 原画像と 2 値化：正常例

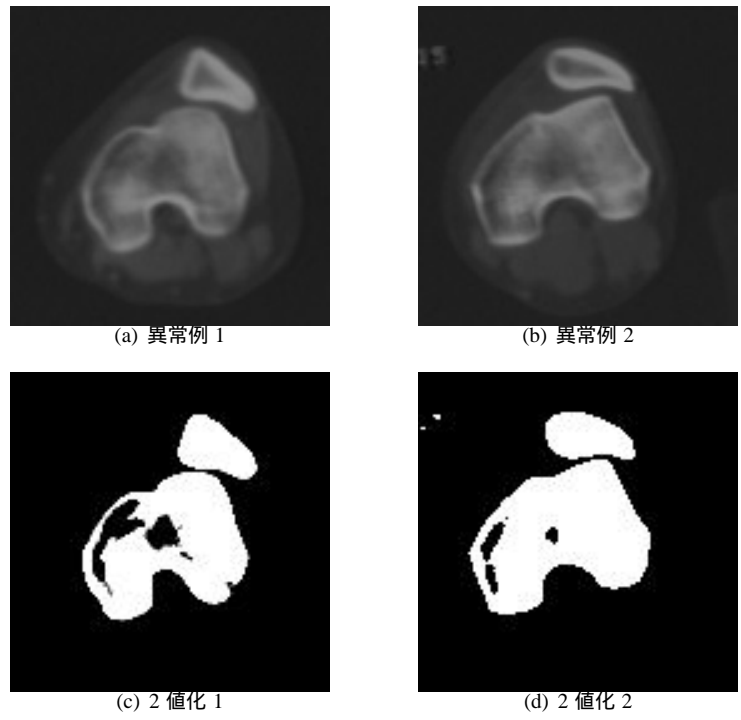


図 5.3: CT 原画像と 2 値化 : 異常例

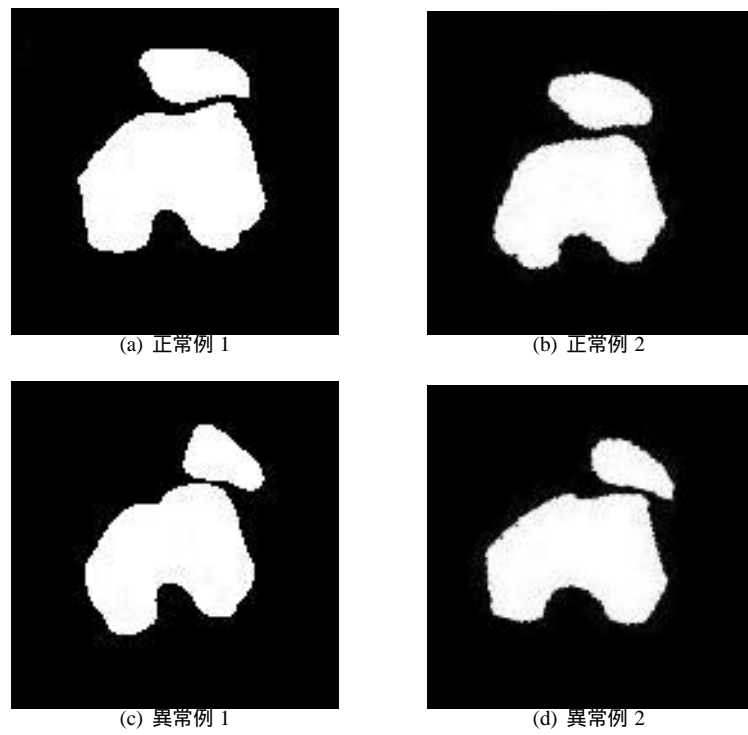


図 5.4: 骨領域の穴埋め・2 値化処理画像例

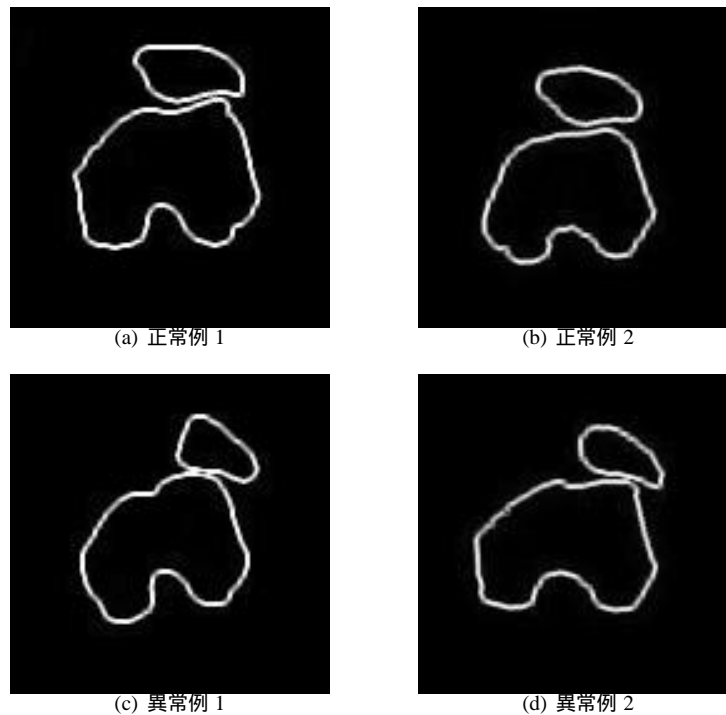


図 5.5: 骨領域のエッジ画像例

処理の流れを図 5.4 に示す．入力原画像 A は，まず，周辺部フィルム傷の不要領域の除去を行った後，局所自動 2 値化処理により骨領域が抽出される (D_0)．続いて，骨領域の穴埋めを行う (D_1)．さらに，2 値化された画像に対して位置，大きさについて正規化を行った．ただし位置，大きさの正規化は重心，水平垂直方向のフェレ径を計測し，垂心を画像中央に，水平方向フェレ径を一定の長さで正規化するように回転，拡大，縮小のパラメータを決定し，これらに基づいて affine 変換を行った (D_2)．図 5.4 にその例を示す．

形状特徴としては今のところ正規化処理後の 2 値化画像および 2 値化画像から抽出した輪郭形状 (エッジ画像 E) を用いた．それらをニューラルネットワークの入力とする．エッジ画像の例を図 5.5 に示す．

5.3 神経回路モデル(ニューラルネットワーク)

本研究では NEC PC-9801 パソコンと“NEURO-07”ニューロ・ボードを用いて 3 層バックプロパゲーションネットワーク (Back Propagation Network; BPN) を構築した．このモデルでは，層内の結合はなく，層間の結合は第 1 層から最終層への一方向の結合のみである．学習は D.E.Rumelhart [6] らの学習方法による．各層のユニットは，前の層のユニットからの重みつき入力を受けて，その総和を計算し，非線形のシグモイド関数を通したものを出力する．はじめは，重みの初期値をランダムに設定する．BPN の学習アルゴリズムではネットワークの出力と教師信号の差を減らす方向に結合の強さを修正する．

形状特徴を持つ画像サイズはシステムの都合上， 128×128 pixel を 32×32 pixel に変換した低解像度画像とした．入力されるユニットの数は 32×32 である．中間層のユニットの数は 50 で，出力層のユニットの数は正常と異常に対応して 2 である．

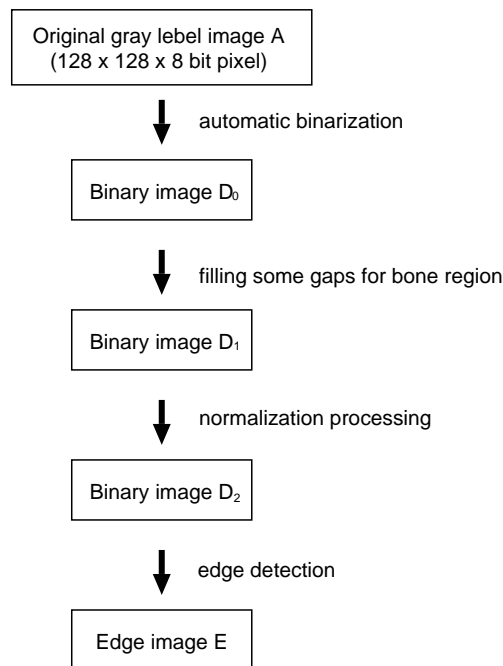


図 5.6: 2 値化処理とエッジ抽出

5.4 認識実験

今回使用したパターンは正常者 22 名と異常者 66 名のものである。この中の正常者 10 名，異常者 15 名の左膝画像を学習パターンとして使った。残りの未学習のものは認識テスト・パターンとした。

以下の 2 種類の画像特徴による認識実験を行った。そのときニューラルネットワークの各パラメータの値および学習結果を表 5.1 にまとめて記す。

- (1) 2 値化画像
- (2) エッジ画像

いずれも正常者 10 名，異常者 15 名，合せて 25 人の左膝画像データを学習パターンとした。各パターンを 2 回提示して， $2 \times 25 = 50$ 回の学習を 1 セッションとし， r セッションの学習が行われた。2 値化画像の場合，200 セッションくり返して学習を行った。ただし，79 セッション目で 100% の認識率となった。エッジ画像の場合，35 セッションくり返して学習を行った。ただし，認識率は 22 セッション目で 100% になった。

つぎに未学習パターンを用いて識別テストを行う。認識率は表 5.2 に示す。2 値化画像の場合，正常者と異常者ともに 75% の認識率を得た。エッジ画像の場合，正常者に対して 83%，異常者に対して 78% の認識率を得た。この結果より，エッジ画像の認識率は 2 値化画像より高いだけでなく，学習回数の点でも優れていることがわかった。この一つの解釈としては次のようなことが考えられる。すなわち，2 値化画像データの場合，領域内の余分の情報が多いため，ニューラルネットワークの学習と認識に対して不利に働いたのではないかと考えられる。それに対して，2 値化画像のエッジ画像にはその図形に関する必要かつ十分な情報が含まれており，したがって，ニューラルネットワークによる図形の特徴の抽出が容易であったと解釈されよう。

表 5.1: ニューラルネットワークの学習およびテストに使用したパラメータ

three layer network	binary image	edge image
input neuron units	1024	1024
hidden neuron units	50	50
output neuron units	2	2
sigmoid activation parameter μ_0	4.5	3.0
learning rate η	0.3	0.3
momentum parameter α	1	1
number of learning pattern P	25	25
learning times of each pattern R	2	2
iterations for learning r	200	35
recognition rate for learning(%)	100	100

表 5.2: 未学習パターンに対する認識率

test pattern	recognition rate(%)	
	binary image	edge image
normal (12)	75.0(9/12)	83.3(10/12)
abnormal (51)	75.4(38/51)	78.4(40/51)
average	74.8	80.9

5.5 結言

本章の目的は、3層バックプロパゲーション・ニューラルネットワークの優れた学習能力に確定診断付き画像を教師信号としてトレーニングを行い、画像のみで異常診断を行うことである。膝蓋骨亜脱臼のスクリーニング検査をレーザなどで計測した膝関節外形形状から行うことを考え、80%程度の識別率を目指して行う計画であったが、外形形状のみからでは識別率が50%程に止まった。そこで膝のCT画像より骨領域を抽出した2値画像およびエッジ画像からの膝蓋骨亜脱臼画像の自動診断について検討を行った。

得られた成果を以下に示す。

- (1) 2値画像を使用した認識では正解率74.8%、エッジ画像を使用した認識では正解率80.9%を得た。

膝蓋骨亜脱臼の自動識別の研究は必要性・有用性が高いが行われておらず、本研究の成果は予備的なものに止まっているが、人工ニューラルネットワークの可能性を示唆するものになると考える。

参考文献

- [1] Inoue M., Shino K., and Hirose H. et al. Subluxation of the patella. *Computed Tomography Analysis of Patellofemoral Congruence*, Vol. 70-A, pp. 1331–1337, 1988.
- [2] 中西克之, 井上雅裕, 原田貢士, 他. 膝蓋骨亜脱臼症候群のMRI-膝蓋関節軟骨の病的変化の検討-. 日本医放会誌, Vol. 51, pp. 387–393, 1991.
- [3] 轟正棟, 田村進一, 河合秀夫, 他. ニュラルネットワークによる超音波を用いた二次元ターゲット位置の識別. *Med. Imag. Tech.*, Vol. 9, pp. 63–67, 1991.
- [4] 光本浩士, 田村進一, 河合秀夫, 他. エネルギー関数によるヒトの角膜上皮細胞の形状抽出・補完. 医用電子と生体工学, 27(特別号), p. 353, 1989.
- [5] 呉, 井上, 竹谷, 田村, 小野. ニュラルネットワークによる膝蓋骨亜脱臼の自動識別. *Medical Imaging Technology*, Vol. 9, No. 4, pp. 460–463, 1991.
- [6] Rumelhart D.E, McClelland J. L. *Parallel Distributed Processing*. MIT Press, Cambridge, MA, 1987.

第6章 PDPモデルによる印刷囲碁総棋譜の自動読み取り

6.1 緒言

世界中には囲碁の愛好者は2000万人～3000万人いると言われる。我々は、画像処理による対局中の囲碁棋譜の自動記録に取り組んでいる [1]。印刷物の囲碁総棋譜は一般的に200手以上あって、小さいページ領域中に小さい文字で記録され、読み取り(棋譜理解)はかなり手間がかかるだけでなく、ゲーム進行の臨場感もない。また長年に亘り多く蓄積・記録されている囲碁総棋譜の自動読み取りとそれを利用したパソコンによる一手毎の再現は、囲碁愛好者に強く望まれるところである。しかし、その再現と棋譜データベース化の重要性にも関わらずこれまで囲碁総棋譜の自動読み取りに関する研究はなされていない。

棋譜の自動読み取りについては直感的には、すべての手順・白黒石の数字画像をテンプレートとして作成・保存を行い、従来法の相関法で棋譜・交点画像データとの相関値で認識を行う手法が挙げられる。しかし実際の処理を行うと、テンプレート作成画像では、100%の認識率が得られるが、同じ雑誌でも他の棋譜に対して行うと認識率は数%であった。棋譜に印刷された数字には多くのフォントが使用されており、スキャナーで読み取る数字の線幅が異なるばかりでなく、白石の円の大きさが異なることが原因であるが、認識処理での位置ずれを考慮して ± 2 画素、 5×5 の範囲で相関の最大値を求める手順を組み込んでもあまり認識率は向上しなかった。一方、数字のみ取り出して、ボカシ処理も含めて同様の処理を行ったが、数字認識率は50%～75%程度であった。

そこで、我々は囲碁総棋譜の自動読み取りにおいてPDP(Parallel Distributed Processing)モデルによる数字認識を用いることを検討した。ニューラルネットワークは、脳を人工的に模擬したモデルであり、学習によって知識は保存される。またパターン認識に優れた能力を持ち、情報処理過程は従来法のそれと異なる [2] ことに注目した。ニューラルネットワークはアルゴリズムや既知の規則セットの開発を必要とせず、ソフトウェアの量を大幅に減らすことが可能な新しい情報処理のアプローチであり、1980年代に入って多くの分野から注目されるようになった。

文字認識は、(1)文字抽出、(2)認識手法、に分類される。(1)については前処理として、2値化・ラベリング・連結数字の分離・数字の正規化・細線化・メッシュ特徴の利用を行う。(2)について、FPM(Fuzzy Partition Model)では、評価基準としてKull-backダイバージェンスを用いユニットの総出力和の正規化を行う。また、教師ベクトルを $(0, \dots, 0, 1, 0, \dots, 0)$ のように0と1のみで構成することにより、出力層のFPMユニットに現れる相互抑制を利用する。出力以外は、お互いに抑制し合うことで収束性・認識率の向上を行う。

まず、雑誌などに記載された囲碁・印刷総棋譜をスキャナーで読み込む。囲碁総棋譜の数字桁判定の前処理を行った後、低品質な総棋譜の数字認識を2つのPDPモデルとしてFPMおよびBP(Back-Propagation)のトレーニングと認識率比較を行う。開発した手法について実験例でその有効性について示している [3]。

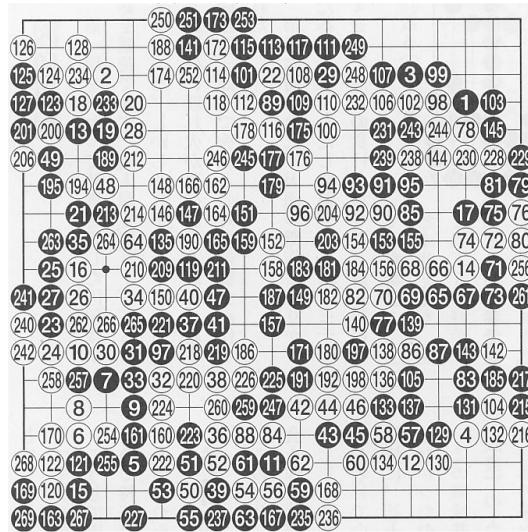


図 6.1: 囲碁印刷総棋譜 (雑誌:月刊 碁ワールド)



図 6.2: 石検出テンプレート (図において灰色の矩形領域は相関値計算は行わない)

6.2 棋譜数字抽出

印刷囲碁総棋譜の特徴は、(1) 図面内文字数が多い、(2) 打石位置は交差点に表示、(3) 白石は白丸中に黒字で、黒石は黒丸中に白抜きで手順が示される、(4) 打石のないところは盤枠線が存在する、(5) 先手黒番、後手白番、(6) 多数の数字フォントが使用されており、3 桁のものでは数字の組み合わせにより異なるフォントが使用される、(7) 印刷物からイメージスキャナーで棋譜を読み取る時、裏面の印刷が雑音として影響する、などの特徴がある。スキャナーで読み取った棋譜画像例を図 6.1 に示す。

利用者の家庭でのパソコンによる利用を主体と考え、イメージスキャナより入力された画像 (解像度 180dpi、約 680×680 pixel) について処理を行う (図 6.1)。

6.2.1 交点座標および黒石白石の抽出

棋譜サイズ検出として、4 隅 (打石されないことが多い) 座標をマニュアルで読み取る。これを用いて交点座標の決定を行う。

次に“白石”、“黒石”および“なし”のテンプレート (図 6.2) を作成し、“白石”、“黒石”および“なし”を検出すると同時にその座標値を求める。

同じテンプレートを用いて、白石、黒石の数字部分の抽出を行う。まず、判別分析 2 値化法 [4] でスキャナー画像の 2 値化を行った後、数字を黒で統一する。実験ではテンプレートによる“白石”、“黒石”および“なし”

く相関値を得たテンプレートを示している。例えば、図 6.4 においてデータ“56”に対し、相関値 0.785 と相関値 0.783 となっている。数字歪が大きく影響していることが窺える。なお、ボカシ処理による認識実験では、32–64%の認識率であった。また、テンプレート作成画像では、上記全ての処理で 100%の認識率であった。単純な相関方法では数字認識は容易でないことが分かった。

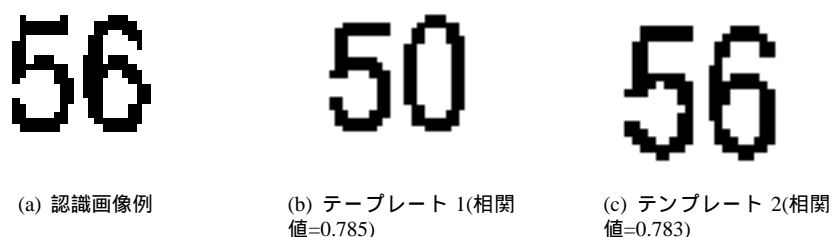


図 6.4: 認識失敗例

6.2.3 数字切り出しとラベリング

手順を示す白石黒石の数字は、黒文字に統一し、各桁に切り出すことにより数字の認識を行う。棋譜から切り出した数字の桁数は 1 桁、2 桁および 3 桁の 3 種類がある。棋譜は 2 値であるが、実際にイメージスキャナー (256 階調: 0~255) で画像として取り込むと多値になっているため、2 値化を行う。この時、閾値の設定により、数字が太くなったり細くなったりする (図 6.3)。ラベリング処理 [5] 後、数字部分の抽出および正規化を行う。

桁判定とラベルの分離 ラベリング処理により数字抽出を行う。ここで、雑音などの影響によりラベリンされた数字の桁判定では、3 桁数字が 2 桁数字に誤判定されることがある。2 桁数字が 1 桁数字になる場合もある。ラベリングの結果、1, 2 桁数字になった場合、連結かどうかは数字幅と画素数を用いて桁数判定を行い、これによって誤ってラベリングされた連結数字は分離を行う必要がある。図 6.6(a) は、2 桁数字が 1 個に連結された原画像例であり、図 6.6(b) は、3 桁数字が 2 個にラベリングされた例である。

ラベリングにおける数字の連結は、次の 3 ケースである。各ケースにつき対処法を示す。

2 桁数字が 1 桁になる場合 (約 3%) 2 桁数字の各桁に対する数字幅の基準 (表 6.2) を基に桁判定・修正を行う。2 桁数字の場合、数字“1”は結合がない。もしラベルに結合があると、その幅は、どの 1 桁数字よりも長くなることを判定基準にする。2 桁数字が 1 桁になったときの特徴を以下に示す。

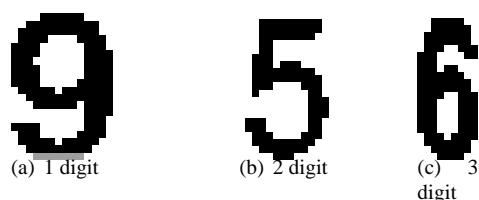


図 6.5: 桁と数字幅例

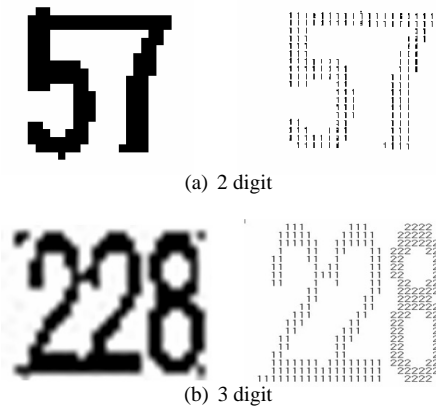


図 6.6: ラベリング

表 6.2: 数字幅

数字	0	1	2	3	4	5	6	7	8	9
1桁 黒	*	12	*	19	*	19	*	18	*	20
白	*	*	20	*	22	*	21	*	22	*
2桁 黒	*	9	16	16	16	15	16	16	16	16
白	18	12	17	18	18	17	17	17	19	18
3桁 黒	11	6	10	10	11	11	11	11	11	11
白	14	8	12	13	13	14	13	12	14	13

- (1) 結合に係る画素数は、1~2である。
- (2) 結合点は、そのラベルのほぼ中央である。

3桁数字が2桁となる場合(約10%) 2つのラベル幅および面積を、それぞれ A_w, A_s (2桁数字) および B_w, B_s (1桁数字) ($A_w > B_w$) とした場合、 $A_w/B_w < 3.0$ であるか、または $B_s/(A_s + B_s) < 0.375$ であれば3桁数字と判定し、分割を行った。これにより実験では、100%正しく分離できた。

3桁数字が1桁となる場合 殆ど生じないので省略した。

規格化 2桁3桁の文字幅にはばらつきがある。文字の特徴を捉えるのが困難であるため、規格化(拡大)処理を行う。2桁3桁の場合は文字が縦長で数字特徴と印刷仕上がりでの美観との関係で、ピッチの微妙な変位と数字の組み合わせで多種のフォントが複雑かつ多様に利用されており、これが数字認識を困難にしている。そこで、画像の重心に一番近い成分に注目し、縦、横へ2倍に引き伸ばした後にそれを細線化することで認識精度の向上を図った。図 6.7 に3桁数字の数字“6”の原画と拡大・細線化処理後の画像例を示す。この処理による入力文字例を図 6.8 に示す。

6.3 FPMを用いた数字認識

加藤, 丹, 江島らは、英数字について Kullback ダイバージェンスで評価した FPM(Fuzzy Partion Model) が従来の BPN(Back Propagation Neural Network) モデルに比べて、はるかに少ない学習回数で、より正確に認識す



図 6.7: 拡大・細線化 (3 桁)

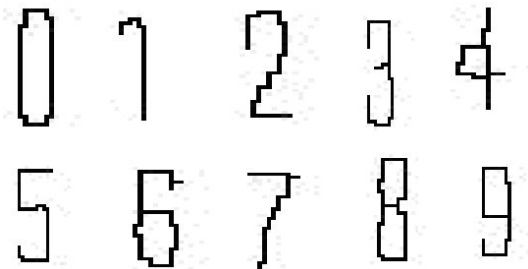


図 6.8: 入力文字例

ることを示している [6] .

FPM の特徴は，ユニットの総出力和の正規化を行い，教師ベクトルを $(0, \dots, 0, 1, 0, \dots, 0)$ のように 0 と 1 のみで構成する．これは文字認識をはじめとするパターン分類によく用いられるケースであり，入力パターンのカテゴリーに対応する出力に対しては 1，その他は 0 に設定する．これにより，出力層の FPM ユニットに現れる相互抑制作用による ??? . 出力以外は，お互いに抑制し合うことで収束性の向上が可能となる [6] .

我々は当初，数字認識について BPN を用いて認識を試みてきたが不十分であった．そこで FPM 素子を用いた学習速度の向上と，認識精度の向上を図った .

6.3.1 FPM

FPM の前身 SVM(Stochastic Vector Machine) は，独自の階層型 PDP(Parallel Distributed Preprocessing) モデルである．SVM モデルをもとに BPN モデルとの対応関係を明確にしたものが FPM である [6] . FPM ユニット (図 6.9) は逆 logit 変換により $(N - 1)$ 個の u に対し N の出力 a を得る .

FPM では出力関数として，逆 logit 関数と呼ばれる ($a = \text{logit}(u^{-1})$) .

$$a_k = \frac{\exp(u_k)}{1 + \sum_{i=1}^{N-1} \exp(u_i)}, k = 1, \dots, N - 1 \quad (6.1)$$

$$a_N = \frac{1}{1 + \sum_{i=1}^{N-1} \exp(u_i)} \quad (6.2)$$

の関係がある．教師ベクトルを $(0, \dots, 0, 1, 0, \dots, 0)$ のように 0 と 1 のみで構成する (教師は，探索空間の端点に配

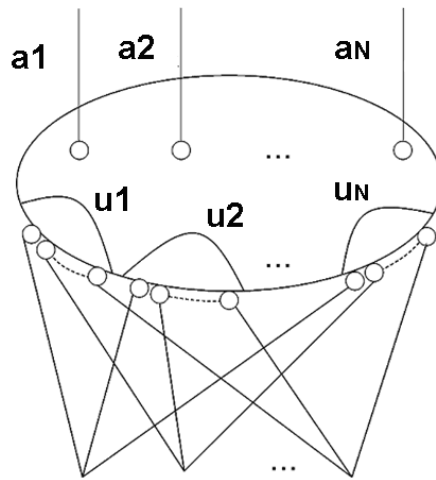


図 6.9: FPM ユニット

置したことになる)。なお，教師と実際の出力との差 $\delta = (t - a)$ の評価について，2 乗誤差和

$$E = \frac{(t - a)^2}{2} = \frac{\delta^2}{2}$$

および Kullback ダイバージェンス

$$D = t \ln \frac{t}{a} + (1 - t) \ln \frac{(1 - t)}{(1 - a)} \approx \frac{\delta^2}{2t(1 - t)}$$

である。D が E より，評価量を多く見積もる。

ここに， u_k は入力， a_k は出力である。ニューラルネットワークの学習は BPN モデルと同様に行い，次式により w を更新させる。

$$\Delta w(t) = \eta \delta a + \varepsilon \Delta w(t - 1) \quad (6.3)$$

右辺第 2 項は慣性項であり， ε は慣性項係数と呼ばれている定数， $\Delta w(t)$ は重みの更新量である。再帰的に用いれば重みは更新され，多層のネットワークの学習ができる。

従来の 2 乗誤差和による最急降下法は，出力が端点に近づくほど学習しにくくなるのが欠点であった。Kullback ダイバージェンスを評価量にすることにより，この点を改善している。

6.3.2 メッシュ特徴と FPM の構成

図 6.10 において，メッシュ特徴を利用した FPM の文字認識への適用を行う。出力層は 0~9 までの 10 種類を出力するため，出力数は 10 とした。出力層は 10 入力 10 出力の FPM 素子を 1 つ用いた。また，中間層は 1 入力 2 出力の FPM 素子を 13 個使用した。

16×32 の 2 値画像をそのまま入力画像として用いた場合，(1) 位置ずれや微妙な変形，(2) 入力ユニット数が多大となる，等の影響があり，認識精度が低下すると考えられる。これに対処するために入力画像を大局的に捕らえる各メッシュでの特徴 [7] を用いた。文字が抜き出せていれば位置や形に多少のばらつきが生じてても，文字要素の分布はおおよそ変わらない。そこで入力画像を 4×8(図 6.11) のメッシュに分け，各メッシュに占める文字要素の割合を新たな文字の特徴に用いることで認識精度の向上の検討を行った。また，メッシュ特徴処理

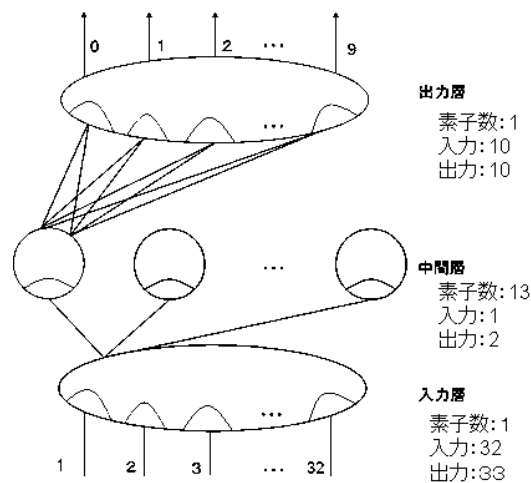


図 6.10: 数字認識に使用した FPM

では入力層は入力画像を 4×8 に 32 等分し、各升目の面積中の文字要素の占める割合を入力値とした。そのため、入力層には 32 入力 33 出力の FPM 素子を 1 つ用いた。



図 6.11: 4×8 メッシュ分割

6.3.3 学習速度比較

学習速度比較のために細線化画像を 10×10 に縮小した入力画像を使用し、FPM と BPN を用いた学習を行う。その際に画像枚数を 10 枚、20 枚、30 枚と変化させたときの学習回数を比較した (表 6.3)。メッシュ特徴 4×8 を用いた時の FPM の学習速度もあわせて測定した。表 6.3 より、FPM は BPN に比べて学習回数をはるかに少ないことが分かる。特にトレーニングデータ件数が 10 件の場合では $1/10$ 以下であり、データ件数が 30 の場合でも BPN と比較して $1/3$ の学習回数である。よって、FPM は、BPN の学習を行うよりもはるかに早く学習を行えることが確認できた。これにより、FPM を用いた場合にはより多くの教師データを学習させることができる。

6.3.4 数字認識

学習には各数字 5 件計 50 件を使用した。FPM および BPN の各手法による認識を 3 桁の数字 (100 ~ 300) までの各 200 件ずつに対して実験を行い、比較検討を。その結果を表 6.4 に示す。

表 6.3: 学習回数の比較結果

教師信号個数	学習回数		
	FPM		BPN
	メッシュ特徴 (4×8)	細線化画像 (10×10)	細線化画像 (10×10)
10	8862	15043	206735
20	25717	73557	228264
30	73672	105619	282765

表 6.4: 認識結果

ネットワーク	認識率 (認識件数)
FPM ニューラルネットワーク	94% (188/200)
BP ニューラルネットワーク	90% (180/200)

表 6.4 より認識結果は FPM の方が約 4% 向上している。これは、切り出しのばらつきや、文字の途切れなどをまだ含む文字について、これまで用いていた BPN での認識と比べて、FPM による認識やメッシュ特徴が有効であったためと考えられる。なお、市販の OCR での数字のみ抽出した画像 (図 6.3) での認識率は 50% 程度であった。

6.4 結言

PDP モデルによる数字認識により囲碁印刷総棋譜の自動読み取りを行った。前処理では細線化処理、メッシュ特徴を用いることで FPM により数字変形に対する学習速度および数字認識率を向上させることができた。

本章の成果を要約を以下に示す。

- (1) FPM による棋譜数字の認識率は 94% が得られ、市販の OCR (正解率 50% 程度) や従来法である相関法 (正解率 75% 程度) に比べ大幅に改善できた。
- (2) 印刷総棋譜数字のラベリング処理において、その桁数字の特徴を利用することにより数字を抽出できた。
- (3) FPM の入力データを、数字の正規化と細線化およびメッシュ特徴を利用することにより、学習速度の高速化と認識率の向上が図れた。

今後の課題は、抽出時に避けられない位置ずれ対策の更なる検討と、多種フォントの学習による認識精度の向上や棋譜のデータベース化などが挙げられる。

参考文献

- [1] 福山忠男, 荻巣高博, 金珍祐, 岡崎耕三. 画像処理による囲碁棋譜の自動記録. 電学論 C, Vol. 126, No. 8, pp. 950–956, 2006.
- [2] I. E. Sethi and A. K. Jain. *Artificial Neural Networks and Statistical Pattern Recognition*, pp. 11–31. Elsevier Science Publishers, New York, 1991.
- [3] Hisashi Taketani, Mohd Syafiq Suhaimi, Wang Rong Long, Kozo Okazaki, Kyohei Shimada, and Shinichi Tamura. Auto-reading of IGO ALL-Record by PDP Model. *International symposium on Robotics and Interigent Sensor*, 投稿中, , 2010.
- [4] 大津展之他. パターン認識. 朝日書店, 1997.
- [5] 谷口 慶治編. 画像処理工学基礎編. 共立出版, 2002.
- [6] 加藤喜永, 丹康雄, 江島俊朗. 英数字認識に対するフィードフォワード形 PDP モデルの比較検討. 電子情報通信学会論文誌 D, Vol. J73-D2, No. 8, pp. 1249–1254, 1990.
- [7] 山本眞司, 中島晃, 中田和男. 階層的パターンマッチング法による漢字認識の実験–印刷漢字認識の研究–. 電子情報通信学会論文誌 D, Vol. 56-D, No. 12, pp. 714–721, 1973.

第7章 結論

本研究ではニューラルネットワークによるパターン認識に関して検討を行った。各章で得られた成果を要約すると次のようになる。

第2章では

ill-posed problem(不良設定問題)である位置ずれ学習問題に対して3層構造のネットワークの組み合わせによる手法を提案した。得られた成果を以下に示す。

- (1) 不良設定問題としての位置ずれ補正に対して、拘束条件をネットを二つ組み合わせた構成の形で入れた方法により解決した。具体的には正規化出力という難しい課題を単一ネットワークに任せてしまうのではなく、ネットワークの一部に位置信号というサブゴールを提示し、かつその出力を適切に使用できるようなネットワークの構成を考えることにより、学習を容易にした。
- (2) 1次元ばかりでなく2次元画像についても、位置の正規化を行うネットの構築を行い、有効性を確認した。
- (3) 入力層ユニット-中間層ユニット間の重みに空間周波数展開的な傾向が見られた。これは、フーリエ変換と同様の演算をニューラルネットで行っていることが示唆され、学習によりこれらの重みが得られたことは非常に興味深い。

第3章では

多素子超音波トランスデューサの前方物体を画像化する手法について提案した。分割された素子より1260個の時系列RF(高周波)データを得たが、時系列データ自体は情報量が多いため、エコーの最大ピーク位置を1つの特徴としてとして圧縮・表現した。得られた成果を要約すると、

- (4) 従来の内視鏡超音波診断装置が側方横断像を得るのに対して、前方の3次元像を実時間で得る超音波画像診断装置の基礎的な考察を行った。
- (5) 奥行きをもつターゲットの3次元像について、実用的な画像化手法について提案し、実験により有効性を確認した。
- (6) その際、実験システムに依存した入力パターンの特徴を利用することにより入力パターン空間の次元数を減少させ、ニューラルネットワークに汎化能力を保ったまま、入力パターン空間の次元数を大幅に減少させたネットワークの構築が可能となった。

となる。

第4章では

IC ウェハの切り粉汚れの自動検査を行う場合の分類手法および最適な照明条件について検討を行った。本章の成果を以下に示す。

- (7) ニューラルネットワークによる分類が平均正解率 95.8%となり、切り粉汚れの自動検査には有効であった。
- (8) IC ウェハの切り粉汚れの自動判定に効果的な照明条件は落射照明 70%、斜光照明 100%である。
- (9) 切り粉汚れに有効な周波数成分は $450 < \lambda < 470$ (nm) であるため、これに対応する光学フィルターを使用することにより認識精度の向上が図れる。

今後の課題として本章で開発した手法を拡張して IC ウェハの切り粉汚れのレベルを計測することにより IC ウェハを分類することである。

第 5 章では

研究の目的は、3 層バックプロパゲーション・ニューラルネットワークの優れた学習能力に確定診断付き画像を教師信号としてトレーニングを行い、画像のみで異常診断を行うことである。膝蓋骨亜脱臼のスクリーニング検査をレーザなどで計測した膝関節外形形状から行うことを考え、80%程度の識別率を目指して行う計画だったが、外形形状のみからでは識別率が 50%程に止まった。膝蓋骨亜脱臼の自動識別の研究は必要性・有用性が高いが、未だ行われていない。本研究の成果は確定診断画像のみを教師データとした予備的なものに止まっているが、人工ニューラルネットワークの可能性を示唆するものになると考える。得られた成果を以下に示す。

- (10) 2 値画像を使用した認識では正解率 74.8%、エッジ画像を使用した認識では正解率 80.9%を得た。

第 6 章では

PDP モデルによる数字認識により囲碁印刷総棋譜の自動読み取りについて検討を行った。本章の成果を要約を以下に示す。

- (11) FPM による棋譜数字の認識率は 94% が得られ、市販の OCR(正解率 50%程度) や従来法である相関法(正解率 75%程度) に比べ大幅に改善できた。
- (12) 印刷総棋譜数字のラベリング処理において、その桁数字の特徴を利用することにより数字を抽出できた。
- (13) FPM の入力データを、数字の正規化と細線化およびメッシュ特徴を利用することにより、学習速度の高速化と認識率の向上が図れた。

謝辞

本研究は、筆者が鳥取大学大学院工学研究科、大阪大学医学部、大阪大学大学院基礎工学研究科を経て津山工業高等専門学校に在職中に完成させたものである。

本研究の遂行および本論文をまとめるにあたり、長期間にわたり直接懇切なる御指導、御鞭撻を賜った福井大学大学院工学研究科電気・電子工学専攻 岡崎耕三教授に心から感謝致します。

本論文をまとめるにあたり、福井大学大学院工学研究科ファイバーアミニティ工学専攻 長谷博行教授、同じく情報・メディア工学専攻 吉田俊之教授、同じく電気・電子工学専攻 王栄龍准教授より適切なお助言を頂きました。謹んで厚く感謝申し上げます。

本研究の遂行にあたり、終始有益な御助言と御鞭撻を賜った大阪大学医学部 田村進一名誉教授(現株式会社エヌビエール)に厚く感謝の意を表します。

また、津山工業高等専門学校電子制御工学科の教職員の諸氏には、日頃より種々の面でお世話になっており、ここに記して感謝致します。

最後に本研究の遂行にあたり、いつも暖かく支えてくれた敬愛する父・猛と母・幸子ならびに親愛なる妻・聖子、娘・歩美、実怜に深く感謝するものです。

付録A 表4.2の画像例(汚れなし)

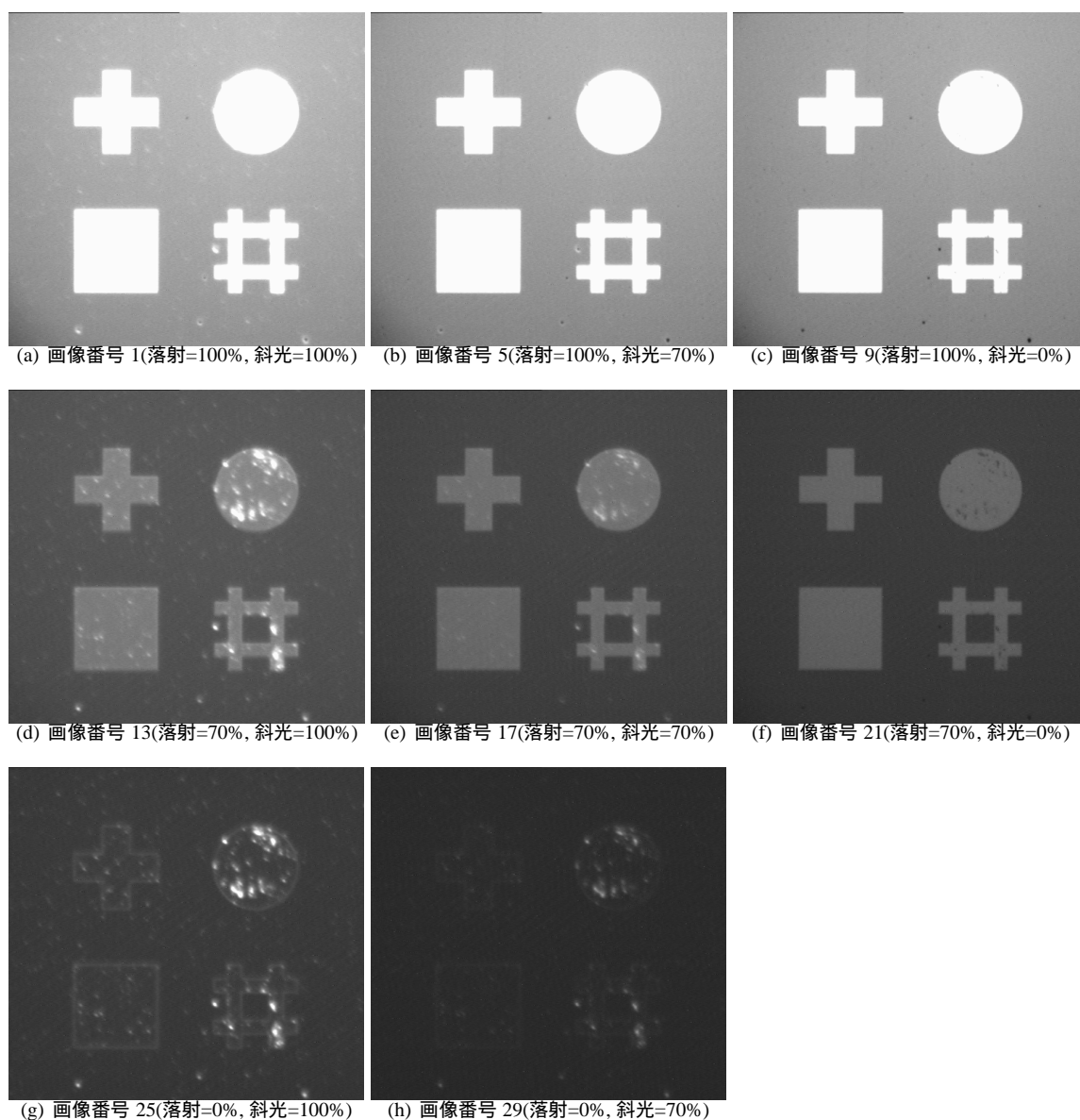


図 A.1: 図 4.7(a) に対する撮影例 (filte No.=1)

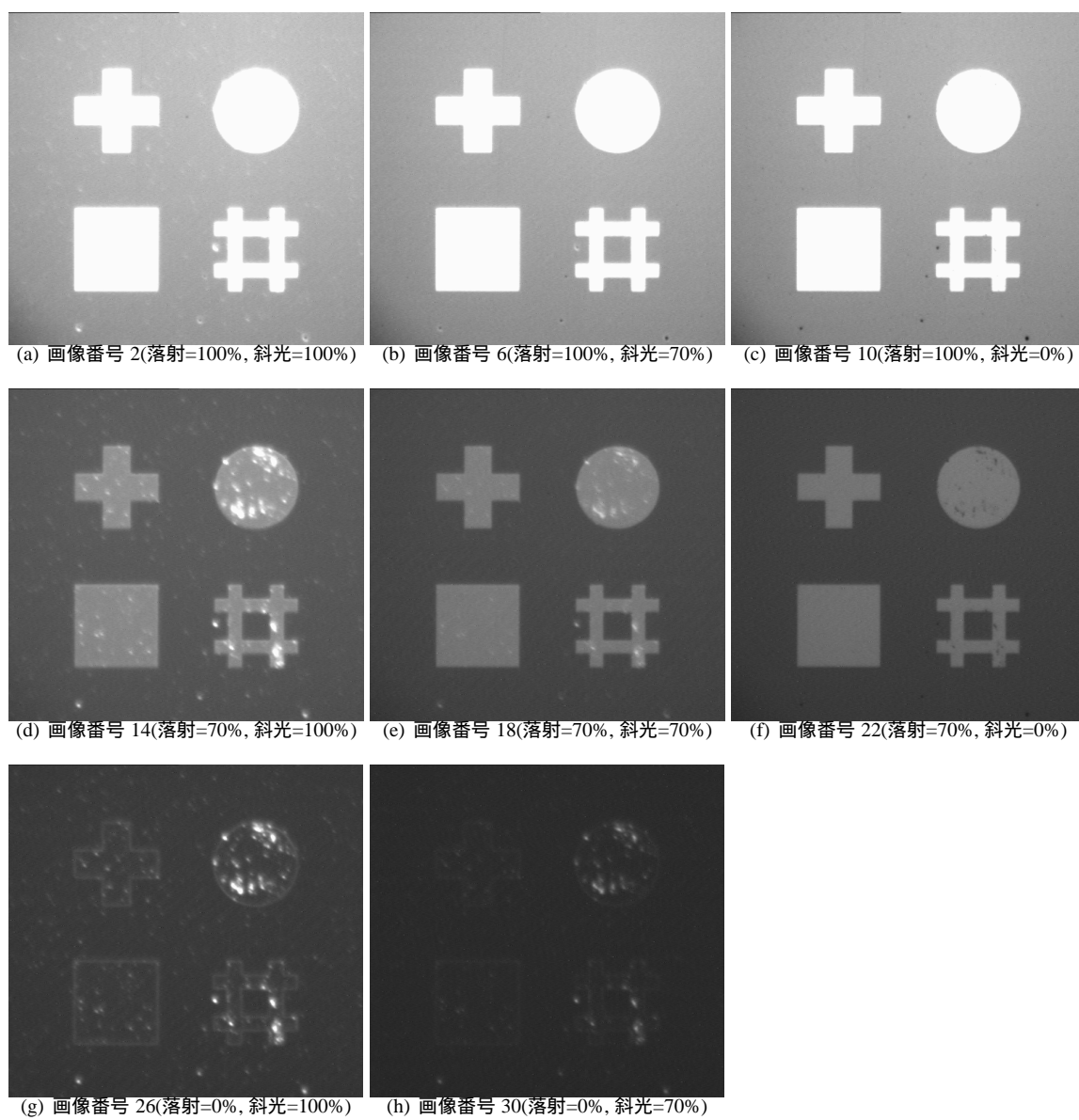


図 A.2: 図 4.7(a) に対する撮影例 (filte No.=2)

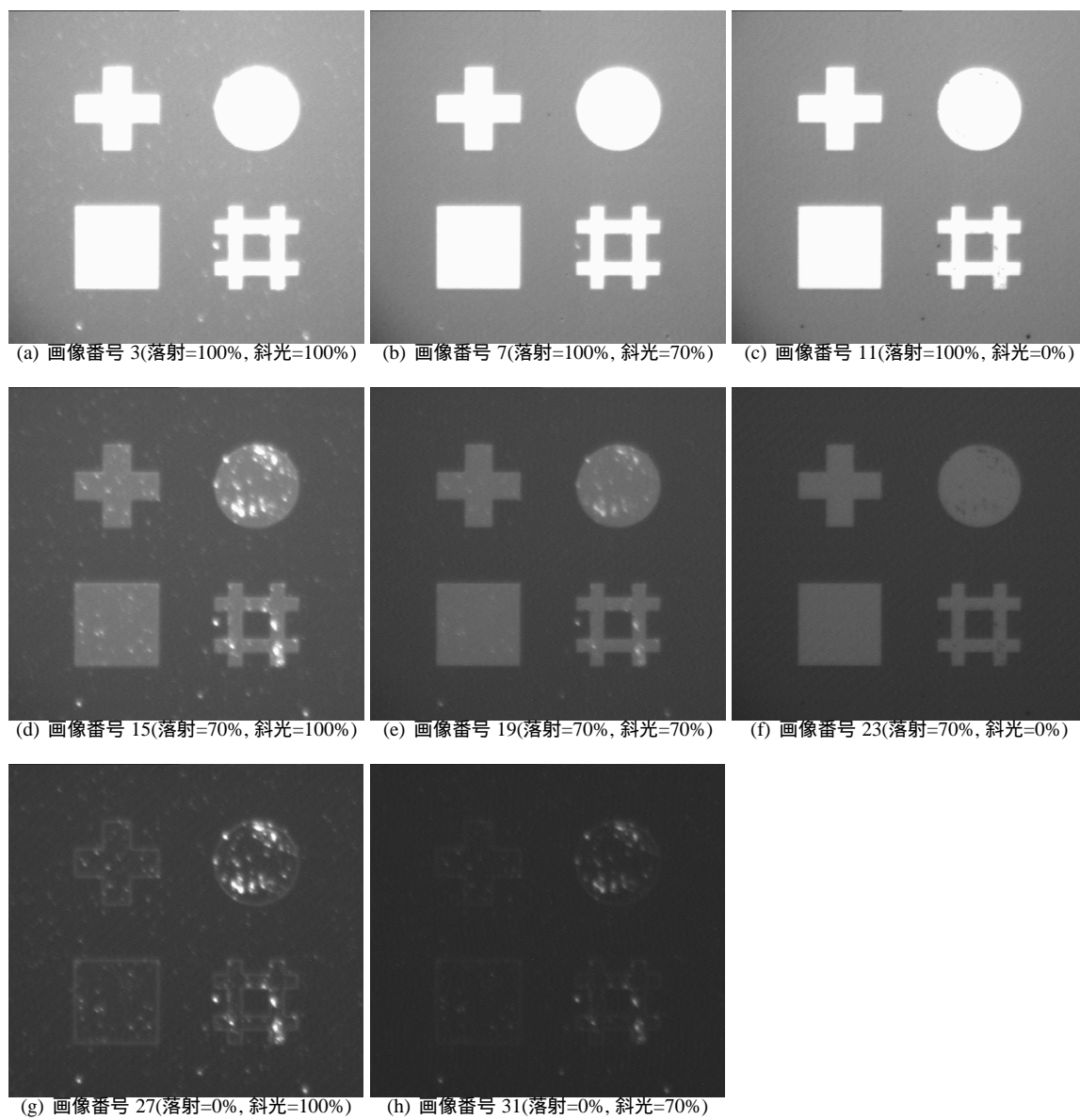


図 A.3: 図 4.7(a) に対する撮影例 (filte No.=3)

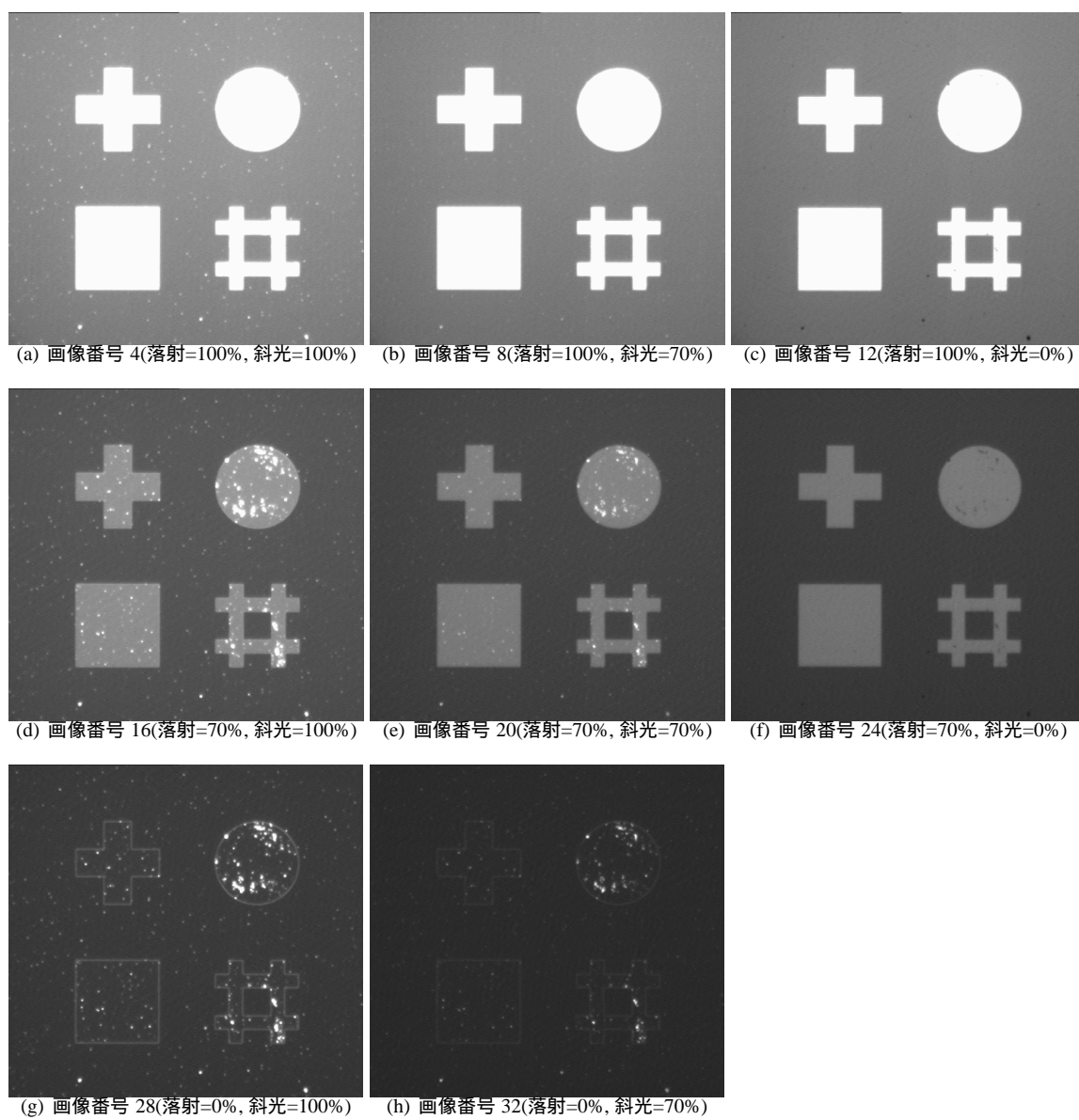


図 A.4: 図 4.7(a) に対する撮影例 (filte No.=4)

付録B 表4.2の画像例(汚れあり)

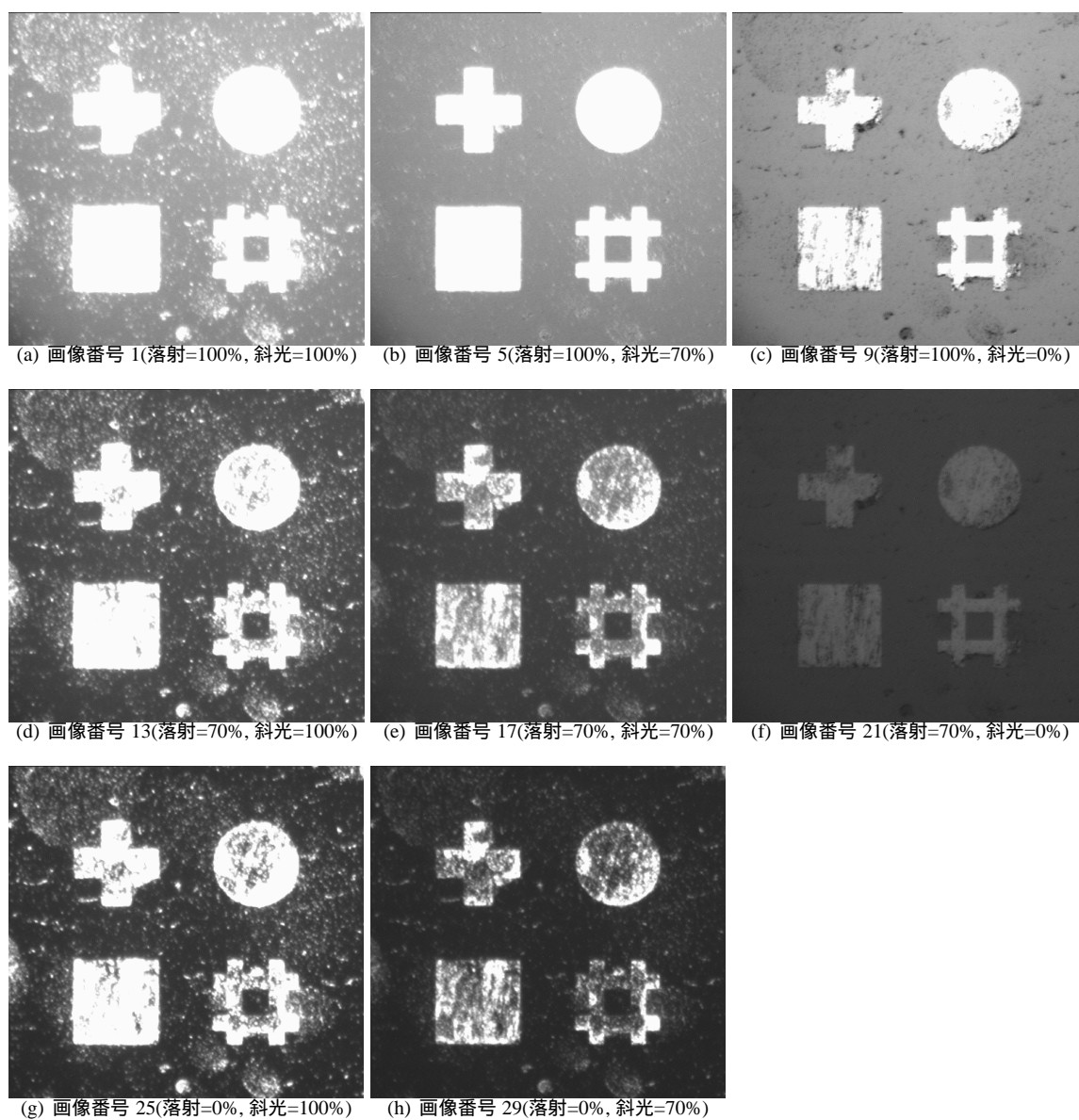


図 B.1: 図 4.7(a) に対する撮影例 (filte No.=1)

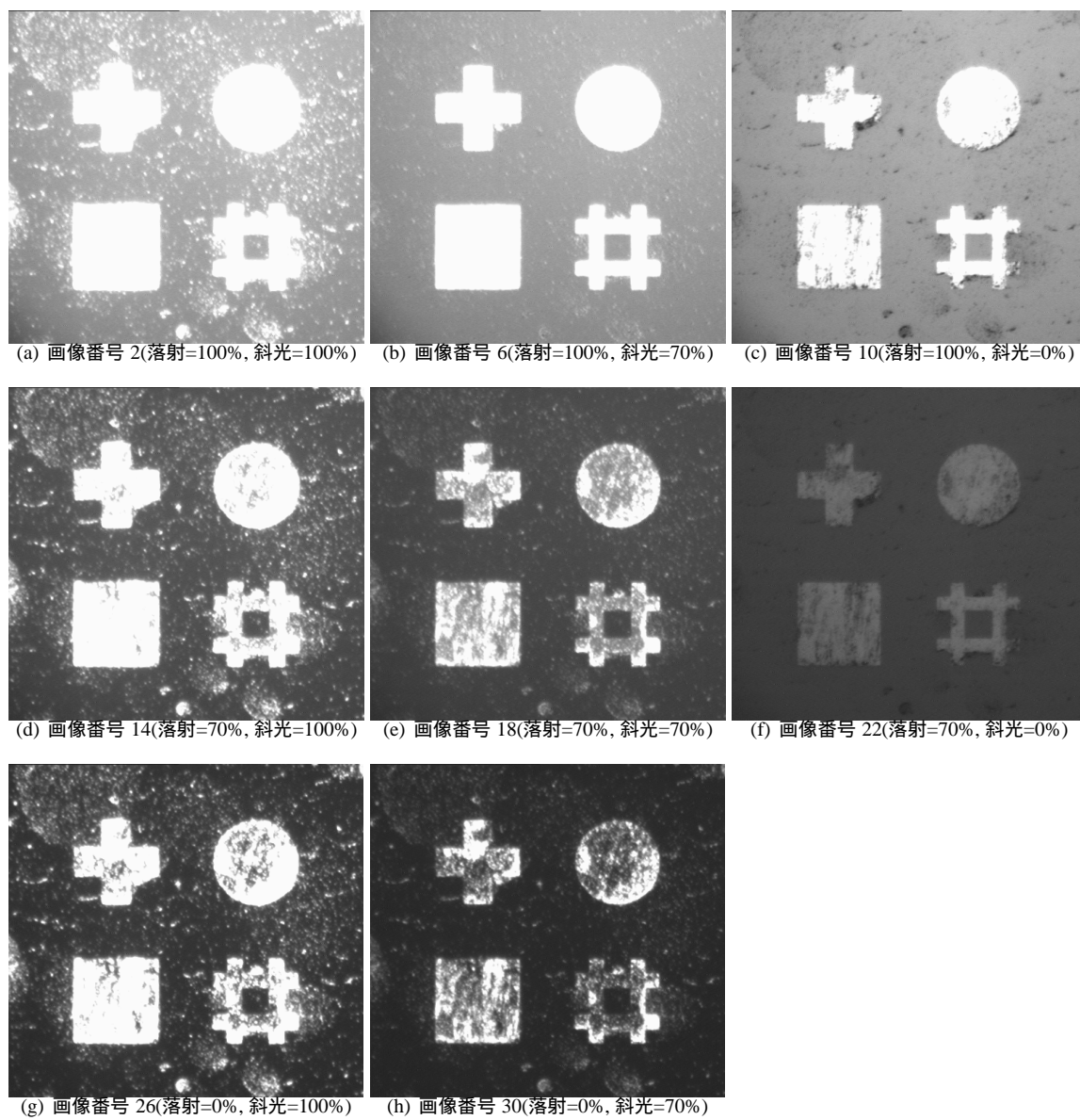


図 B.2: 図 4.7(a) に対する撮影例 (filte No.=2)

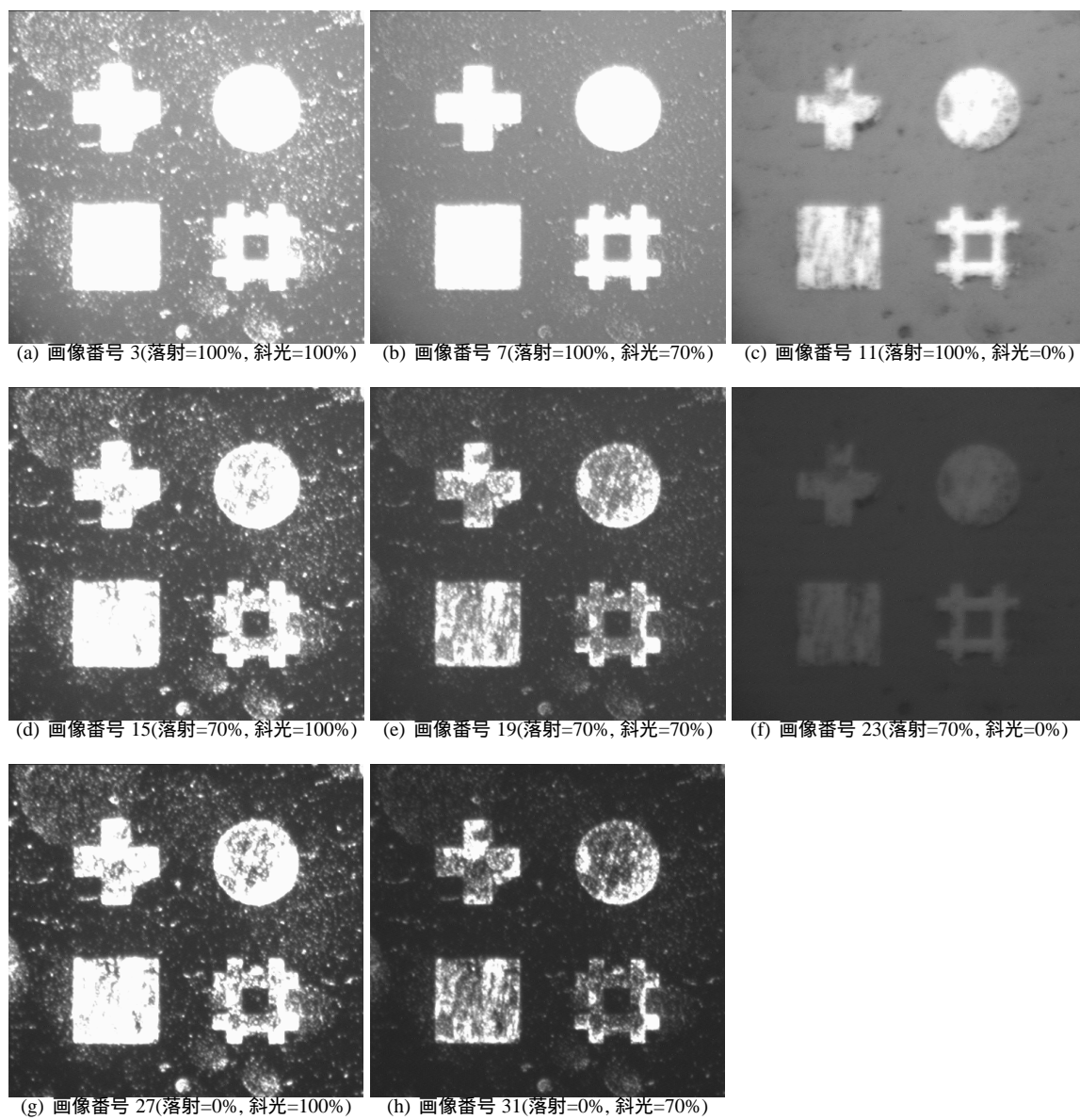


図 B.3: 図 4.7(a) に対する撮影例 (filte No.=3)

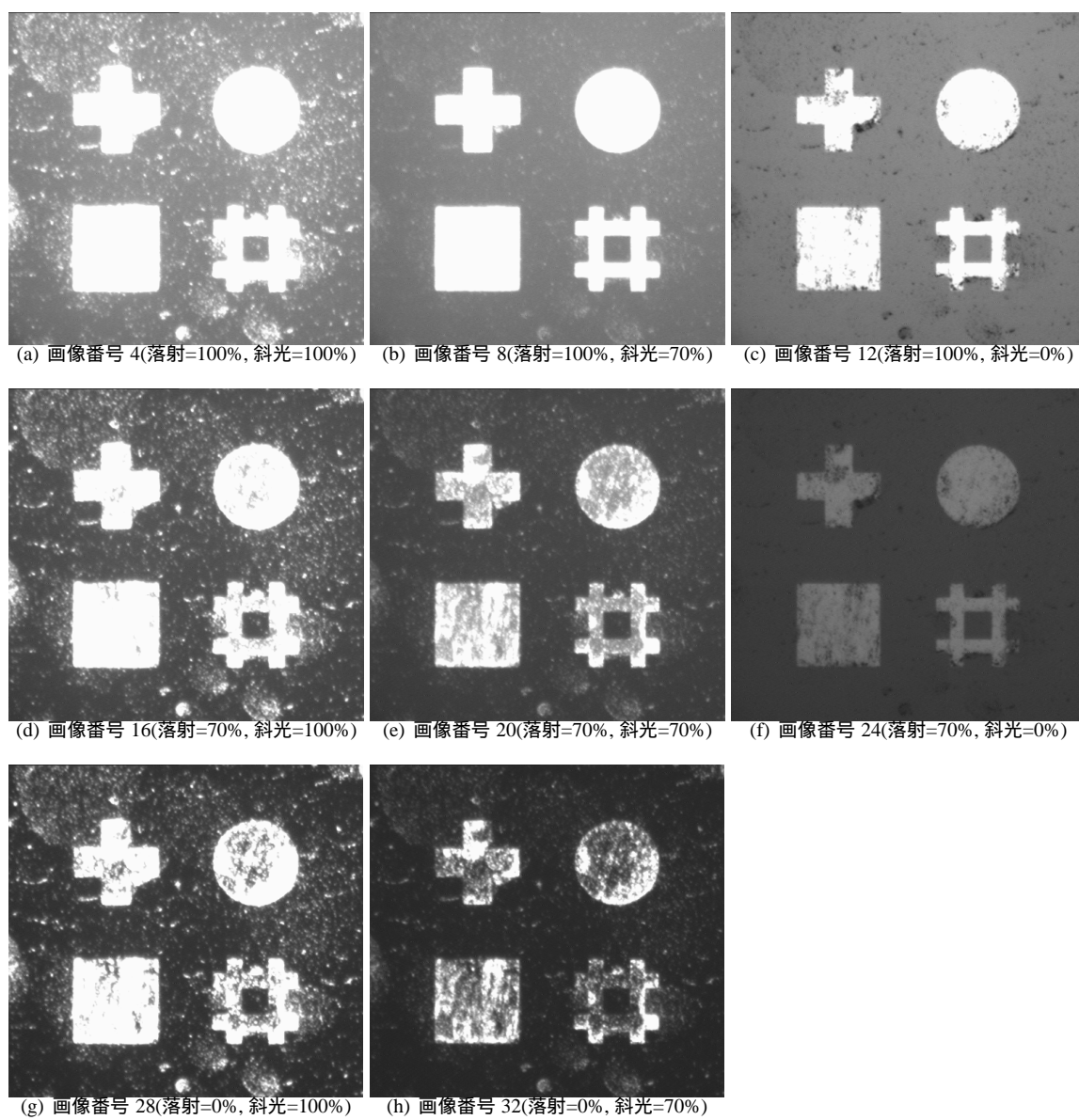


図 B.4: 図 4.7(a) に対する撮影例 (filte No.=4)