OUKA
The University of Osaka
Institutional Knowledge Archive

Title	視覚情報認識のためのニューラルネットワークモデルに関する研究
Author(s)	津村, 徳道
Citation	大阪大学, 1995, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3081438
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

視覚情報認識のためのニューラルネットワークモデルに関する研究

Neural Network Models for Visual Recognition

1994 年

津村　徳道

Norimichi Tsumura

Dedicated to my parents

目次

第１章　序論 …… 1
　１．１　視覚情報認識 …… 1
　１．２　ニューラルネットワークと視覚情報認識 …… 2
　１．３　光ニューラルネットワークと視覚情報認識 …… 3
　１．４　実世界でニューラルネットワークを用いる場合の問題 …… 4
　１．５　本研究の目的と内容 …… 5

第２章　ニューラルネットワークによる視覚情報認識 …… 7
　２．１　緒言 …… 7
　２．２　基本的なニューラルネットワークと学習法 …… 8
　　２．２．１　基本構造の認識能力 …… 8
　　　A.　単一ユニットの認識能力 …… 9
　　　B.　階層型構造の認識能力 …… 11
　　２．２．２　誤差逆伝搬学習法 …… 13
　　　A.　学習原理と手続き …… 13
　　　B.　特徴と問題点 …… 14
　　２．２．３　RCE学習法 …… 14
　　　A.　RCEネットワーク …… 15
　　　B.　認識手続きと学習手続き …… 16
　　　C.　特徴と問題点 …… 18
　　２．２．４　棄却能力 …… 20
　　　A.　汎化問題 …… 20
　　　B.　認識率と誤認識率 …… 21
　２．３　視覚情報認識のためのニューラルネットワーク構造 …… 23
　　２．３．１　位置不変結合階層型構造 …… 23
　　２．３．２　モジュール型構造 …… 25
　　２．３．３　調和モデル型構造 …… 27
　２．４　大規模光ニューラルネットワーク …… 29
　　２．４．１　ホログラム型光ニューラルネットワーク …… 29
　　２．４．２　空間光変調器型光ニューラルネットワーク …… 31
　２．５　結言 …… 32

第３章　高い棄却能力を持つニューラルネットワークモデル …… 33
　３．１　緒言 …… 33
　３．２　二重超球ネットワーク …… 33
　　３．２．１　構造と伝達関数 …… 33

3.2.2　学習手続きと認識手続き …… 35
3.2.3　学習過程の例 …… 36
3.2.4　RCEネットワークとの棄却能力の比較 …… 38
3.3　数字データベースによる実験 …… 40
3.3.1　実験 …… 40
3.3.2　実験結果 …… 40
3.4　結言 …… 43

第4章　ニューラルネットワークによる視覚情報の能動的認識モデル …… 44
4.1　緒言 …… 44
4.2　関係統合ニューラルネットワーク …… 44
4.2.1　視点の移動とズーム操作に基づく情報表現 …… 44
4.2.2　構造 …… 45
4.2.3　認識過程 …… 48
4.2.4　学習過程 …… 54
A.　学習のための視点の移動 …… 55
B.　パターン照合モジュールの学習方法 …… 58
C.　関係ユニットに関する学習方法 …… 62
4.3　顔の認識についての計算機実験 …… 66
4.3.1　パターン木の収集とパターン照合モジュールの学習 …… 66
4.3.2　実験結果 …… 66
4.4　結言 …… 67

第5章　光学的位置合わせの不要な光ニューラルネットワーク …… 69
5.1　緒言 …… 69
5.2　自己調整型光ニューラルネットワーク …… 69
5.2.2　構造と認識過程 …… 69
5.2.3　学習過程における光学系の自己調整機能 …… 71
5.3　実験 …… 74
5.3.1　スパース符号化ネットワークによる記憶容量の向上 …… 74
5.3.2　光学系 …… 76
5.3.3　実験結果 …… 77
5.4　結言 …… 79

総括 …… 82

謝辞 …… 87
参考文献 …… 88
著者発表論文 …… 95

第１章　序論

１．１　視覚情報認識

1950年代の商用電子計算機の出現[1)]により，人間の代わりに視覚情報の認識を行う自動機械が試作され，1960年代初めには光学的文字読み取り装置が初めて商用化された[2)]．1970年代には，傷検査装置等への産業応用が活発に行われ[3)]，現在では，様々な産業で視覚情報の認識を行う自動機械が用いられている[4)]．このような視覚情報の認識に関する研究は，パターン認識[5,6)]，コンピュータビジョン[7)]，ロボットビジョン[8)]の分野で，画像認識，画像理解の研究として行われた．最近では，視点の移動などにより，画像の枠を超えて視覚情報を獲得する能動的な視覚情報の認識技術[9-11)]が注目されている．

図1-1(a)に，文献3,5,6を参考にして画像を情報とした視覚情報認識の基本となる処理形態をまとめた．画像を入力として，特徴抽出を行う前処理部，画像内の物体を識別するパターン認識部，画像内の物体間の相互関係による意味付けを行う画像理解部から構成されている．通常，前処理部において，画像フィルタによるエッジ抽出，テクスチャ解析による領域抽出，標準要素との重ね合わせによる線の交差点の抽出等が行われる．パターン認識部において，最短距離法，ベイズ推定法等の古典的手法や学習機能をもつニューラルネットワーク等により画像内

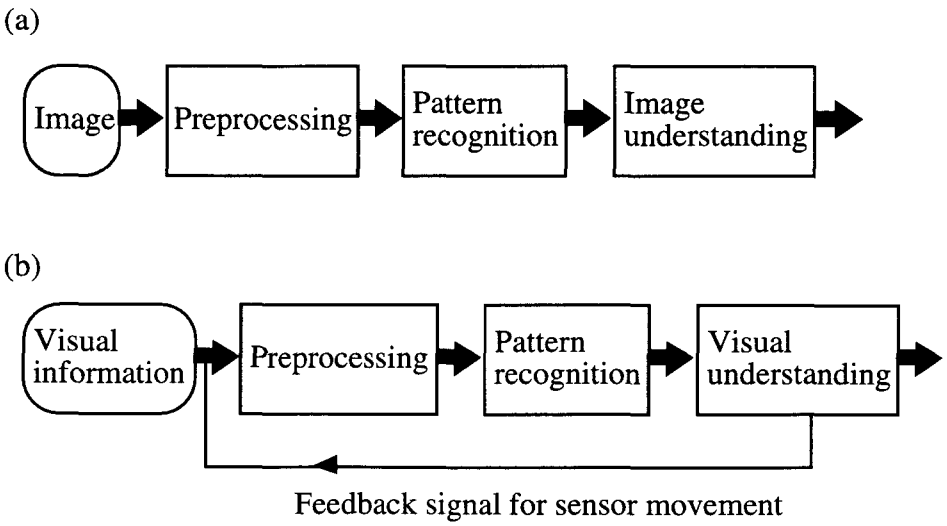

図1-1 視覚情報認識のための基本的な認識形態；(a)画像認識，
(b)能動的視覚情報認識

の物体の識別作業が行われる．画像理解部において，記号に基づく人工知能的手法等による知識情報処理が行われる．図1-1(b)に，視覚情報の能動的な認識のための基本的な認識形態をまとめた．図1-1(a)において，認識する対象は画像であるのに対して，図1-1(b)において，認識する対象は画像の枠を超えた視覚情報である．知識情報処理を行う視覚情報理解部から視覚情報を取り込む部分へのフィードバック信号により視点の移動等が制御される．対象の高速な認識や高い認識能力を得るための効率的な視点移動の制御が，知識情報処理により行われる．

1.2　ニューラルネットワークと視覚情報認識

生物の脳の情報処理形態を模倣したニューラルネットワークは，現在，図1-1のパターン認識部において有効な手法の一つとして用いられている．ニューラルネットワーク[12-16]を画像認識に用いた歴史は古く，1958年にRosenblattによって提案されたパーセプトロンによって，既に簡単な文字の認識が行われた[15,17]．しかし，1969年にMinskyとPapert[18]によりパーセプトロンの限界が明らかにされたために，ニューラルネットワークによる画像認識の研究は1970年代にはあまり行われなかった．一方，コンピュータ技術の進歩を背景に，パターン認識部に重きを置かず，人工知能的な知識情報処理を重視した研究[7]が盛んに行われた．1980年代に入って，Hopfliedの提案したモデル[19]により火がつけられたニューラルネットワークブームにより，再びニューラルネットワークによる画像認識の研究が盛んになった．ブームの直前の1980年にFukushima[20]が，ネオコグニトロンと呼ぶ手書き文字認識ニューラルネットワークを提案していた．Hopfliedモデルが提案されたのと同じ1982年に，Reilyらはパターン認識用の優れた学習アルゴリズムを持つRCE(restricted coulomb energy)[21]ネットワークを提案した．　RumelhartとMcClelandとPDP Research Groupにより1986年に出版された「並列分散処理」[22]の中で述べられた誤差逆伝搬学習法[23]は，非常に強力な学習法であるためにニューラルネットワークの学習に最も良く用いられた．1989年にLeCunらは，この誤差逆伝搬学習法により学習したネオコグニトロンに似た構造を持つニューラルネットワークを用いて，郵便番号の認識を行った[24]．Zhangらも同時期に，誤差逆伝搬学習法によるネオコグニトロン型のニューラルネットワークの学習を実現した[25]．1990年代になって，振動子として働くユニット同士の同期を利用した物体認識モデルが新

しく提案されている[26,27].

現在，２１世紀における高齢化社会を目前にして，従来行われてきた産業への応用だけでなく，実世界で活動するサービスロボット等のための視覚情報の認識手法が要求されている．近年，記号処理に基づく認識手法は，不完全な情報を含む実世界においてうまく機能しないことが指摘される中[28-36]，ニューラルネットワークが，パターン認識の後に行われる知識情報処理においても，不完全な情報を扱うことの出来る新しい認識手法[9,32,34,36-40]の一つとして期待されている[22,36]．ニューラルネットワークは，記憶の分散性や処理の分散性，学習能力を持つことから，不完全な情報を取り扱うことが可能である．先述したネオコグニトロン[20]や振動子を用いた物体認識モデル[26]は，局所的な部分の組み合わせに基づく知識情報処理的な能力を必要とする認識を既に実現している．現在，全体と部分の関係で記述される階層的な知識に基づく認知作業を二層のネットワークで取り扱う調和モデル[41]等，様々な認知モデル[22,36]が提案されている．

１.３　光ニューラルネットワークと視覚情報認識

ニューラルネットワークにより視覚情報を認識する場合，ネットワークは大規模になる．大規模なニューラルネットワークを実現する並列ハードウエアには，非常に多数のユニットと呼ばれる処理単位とそれらを結ぶ膨大な配線を必要とする．しかし，多数のユニットを１本ずつ導線を用いて配線することは不可能に近い．光ニューラルネットワークは，それらのユニット間の膨大な接続を，並列性，無誘導性等の優れた特性を持つ光を用いて実現する信号処理系である[14]．Hopfieldモデルが提案されて３年後の1985年に，Farhatらによってその光学的実現方法の一つが提案された[42]．その後，わずか５年の間に多くの光ニューラルネットワークが提案された[43-49]．それらは，ホログラム型[45-47]と空間光変調器型[42-45,48,49]の２つに大別することができる．ホログラム型は，ホログラムへの多重記録を利用してユニット間の複雑な配線を記録する．空間光変調器型は，レンズアレイ等を用いた入力信号光の複製による空間的な展開により複雑な配線を行う．ホログラム型は，理論的には，非常に大規模な光接続を実現できるが，学習時の多重露光の困難さ等の問題点を持つ[47]．空間光変調器型は，空間光変調器の分解能の限界，レンズ系の収差の影響，光学的位置合わせの困難さ等の問題点を持つ．1990年代に

入って，それぞれの型について，実用化のための研究[50-57]が進められているが，現在のところ，上で述べた問題点が実用化の大きな障壁となっている．

1.4　実世界でニューラルネットワークを用いる場合の問題

ニューラルネットワークによる視覚情報の認識手法を人間の生活する実世界で用いるためには，多数の解決すべき問題が存在する．現在のモデルにおける主要な問題を二つ挙げる．

（1）学習に用いた例以外の入力に対して，ネットワークに如何に望みの出力をさせるか．

（2）従来の記号処理に基づく画像認識[7]のように，複雑な対象の認識が実現できるか．

（1）の問題は汎化問題と呼ばれ，認識率や棄却能力に関する問題ともいえる．棄却能力とは，誤認識の原因となる入力を棄却する能力のことである．認識結果の信頼性を向上するためには，高い棄却能力が必要である．視覚情報の認識装置が，自動車の自動運転のように人間の生活・生命に関わる仕事を行う際に，認識結果の信頼性は非常に重要である．1989年頃から汎化問題は研究され[58-66]，ユニット間の結合の大きさ[61,64,66]やネットワークの規模[62,63]に制約を加えることにより汎化能力が向上することが報告されている．しかし，これらは，学習した対象と同じクラスに属する入力に対する汎化能力のみを対象としている．実世界で通用する視覚情報の認識手法において信頼性の高い認識を行うためには，学習されていないクラスに属する未知入力や，どのクラスにも属さない未知入力に対して，誤った認識を行わずに棄却する必要がある．棄却能力に関する問題はニューラルネットワークを視覚情報の認識に用いる場合だけではなく，従来の文字・文書画像認識技術においても大きな課題となっている[67,68]．（2）の複雑な対象の認識に関する問題は，実世界において複雑な視覚情報が大半を占めることから避けられない問題である．現在までに，複雑な視覚情報に対する認識処理を，モジュール化されたネットワークにより分割して扱う試みが行われている[69-72]．また，各モジュールへ仕事を効率よく振り分ける学習法[73-74]が近年提案されいる．しかし，図1-1(b)に示したような視覚情報の能動的な認識を実現するためには，仕事の分割法のみでなく，時系列で得られた情報の統合や選択的な行動を行う必要がある．

これらに対して，視点の移動により得られた物体に関する時系列情報の統合を行うニューラルネットワークモデル[76]が提案されている．しかし，提案されているモデルでは，選択的に視点を移動せず，単純な図形のみを認識対象としているのが現状である．

1.5　本研究の目的と内容

本研究の目的は，ニューラルネットワークを用いた視覚情報の認識手法に関して，その進歩を妨げている問題を解決することである．本研究ではこれまで述べてきた，棄却能力に関する問題，複雑な対象の認識に関する問題，大規模光ニューラルネットワークにおける光学的位置合わせに関する問題に関し，以下の研究を行った．

(1)ニューラルネットワークの棄却能力を向上させる研究．

(2)ニューラルネットワークにより，視点の移動に基づく能動的な視覚情報の認識を実現する研究．

(3)大規模光ニューラルネットワークにおいて，光学的位置合わせの不要な手法を実現する研究．

本論文において，各研究の成果をそれぞれ第3章，第4章，第5章で述べる．以下に，本論文の構成と内容を示す．

第2章において，本研究の背景となるニューラルネットワークによる視覚情報の認識手法について紹介する．基本的なニューラルネットワーク，視覚情報処理のためのニューラルネットワーク，光ニューラルネットワークの3つに分けて，それぞれ，第3章，第4章，第5章の基礎となる事項について説明する．

第3章において，棄却能力の高いニューラルネットワークモデルを提案する．提案するネットワークモデルは，未知入力を適度に棄却することができる．X線フィルム内に記録された患者ID番号から切り出された数字データベースを用いて，本ネットワークの棄却能力を，従来法の中で棄却能力が高いRCEネットワークの棄却能力と比べて評価する．

第4章において，視点の移動とズーム操作により視覚情報の能動的な認識を行うニューラルネットワークモデルを提案する．基本的な構造は，階層的な知識を取り扱う調和モデル[41]に類似している．ニューラルネットワークを構成するユニッ

ト間の相互作用により，効率よく視点を移動するための選択的な視点の移動が実現できる．顔の認識を例に取り，本ネットワークモデルを構築する．６１名の実際の顔画像を用いて実験を行った結果について述べる．

第５章において，光軸に垂直な方向に光学的な位置合わせの不要な空間光変調器型の光ニューラルネットワークを提案する．空間光変調器上で画像間のAND演算を行うことにより，ユニット間を結ぶ正しい位置に接続の重みを記憶することが出来る．初期実験によりその有効性を確認している．

総括において，本研究の研究成果についてまとめ，今後の研究課題について述べる．

第2章　ニューラルネットワークによる視覚情報認識

2.1　緒言

視覚情報と，は光エネルギーにより網膜の感覚細胞に刺激として与えられた情報である．人間は，物体の形状，色等に関する視覚情報を外界から得る．また，目には見えない温度分布や，生きている人体の断面構造等も画像化して視覚情報として捉えることが出来る．視覚情報は，情報の一部の欠落により不完全なものが多く，現在の計算機による逐次的手法では認識することが困難な場合が多い．しかし，人間の脳はこのような視覚情報を比較的容易に認識することができる．人間の脳の情報処理形態をまねたニューラルネットワークは，これらの不完全な視覚情報を人間のように比較的容易に認識するための手法の一つとして期待されている．

認識(recognition)と似た言葉に検出(detection)や識別(discrimination)，分類(classification)，認知(cognition)，同定(identification)という言葉がある．本論文において，これらの言葉を以下のように区別する．検出とは，対象が存在することを表す信号を得ることである．識別とは，対象を他の対象と区別することである．分類とは，既に得られた基準をもとに対象の属するクラスを得ることである．クラスは，なんらかの類似性を持つ集合を表す．分類には，識別が不可欠であり，対象を分類するというかわりに対象を識別するともいう．認知とは，未知の対象の情報を得る過程である．同定は，対象を既知のものと同一視することである．認識とは，対象を同定する過程である．分類の結果得られたクラスも既知であるため，認識の対象となる．対象を分類しその結果を認識することを単に認識すると呼ぶことが多い．

ニューラルネットワークは，多数のユニット(unit)と呼ばれる非線形処理単位から構成されている．ニューラルネットワークの基本的な構造に，ユニット間が相互に結合された相互結合型構造と，階層的に結合した階層型構造がある．相互結合型構造は，これまで主に連想記憶[19,77]，最適化問題[78,79]等に用いられている．階層型構造は，これまで主に視覚情報認識，制御[80-82]，音声認識[83]等に用いられている．三層以上の階層数と十分な数のユニットを持つ階層型構造は，任意の画像を

認識する能力を持つ[84,85]．しかし，単純な階層型構造においては，学習能力の限界により活字数字程度の複雑さの視覚情報しか認識することができない．学習とは，ニューラルネットワークのユニット間の結合の大きさ等のパラメータを，与えられた例から決定することをいう．視覚情報は，手書き文字や漢字等のように変形，位置ずれ，クラスの多さ等の特有の性質を持つ．そのような視覚情報を認識するために，視覚情報認識のために構築されたニューラルネットワーク構造が用いられる．

本章において，主に文字を認識対象として，ニューラルネットワークによる既存の視覚情報認識手法の説明を行う．本章において，視覚情報は画像として与えられる．すなわち，本章における視覚情報認識は画像認識と等価である．2.2において，基本的なニューラルネットワークである階層型構造を用いた視覚情報認識の原理と，その代表的な学習法である誤差逆伝搬学習法[23]，RCE学習法[21,86-89]について説明する．また，棄却能力についても汎化問題と関係づけて述べる．2.3において，視覚情報認識のための三つのニューラルネットワーク構造についてまとめる．2.4において，視覚情報認識のための大規模光ニューラルネットワークについて述べる．

2.2　基本的なニューラルネットワークと学習法

相互結合型構造と階層型構造のうち視覚情報認識に用いるのは主に階層型構造である．本節において，階層型構造をしたニューラルネットワークに限って，その視覚情報の認識原理，学習法について説明する．階層型構造をしたニューラルネットワークを階層型ニューラルネットワークと呼ぶ．また，ニューラルネットワークの棄却能力についても述べる．

2.2.1　基本構造の認識能力

ニューラルネットワークを構成するユニットは，それ自体で画像を二つのクラスに分割する能力を持つ．ユニットを階層的に並べて結合した場合，三層以上の階層と十分な数のユニットを持つニューラルネットワークは入力画像を任意のクラスに分類する能力を持つ．

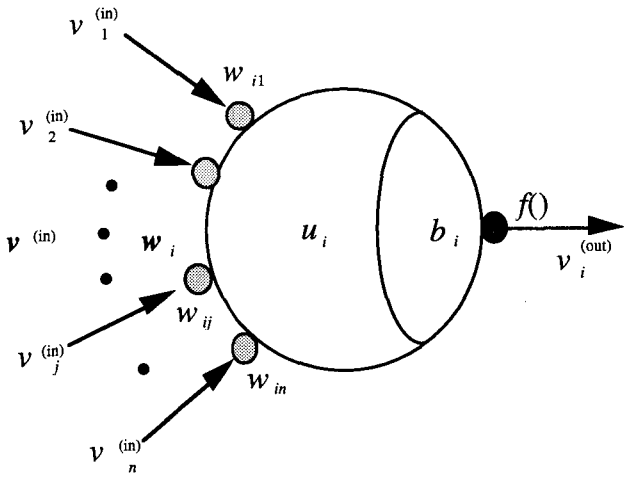

図2-1 ニューラルネットワークを構成するユニット

A. 単一ユニットの認識能力

図2-1に生物の神経細胞（ニューロン）を模倣したユニット i の構造を示す．ユニットは，多入力一出力の非線形素子である．入力信号を$v^{(in)}_1$~$v^{(in)}_n$，出力信号を$v^{(out)}_i$とする．u_i, b_iはそれぞれユニットの内部値，しきい値を表す．また，j番目の入力信号とi番目のユニット間の結合荷重はw_{ij}で表す．ユニットへの入力信号とユニットの結合荷重は，それぞれまとめて入力ベクトル$\boldsymbol{v}^{(in)}$，結合ベクトル$\boldsymbol{w}_i$で表す．ユニットは，伝達関数の異なる相関型と距離型の二種類が用いられる．$f(\cdot)$を非線形関数であるヘービサイド関数（図2-2(a)）とした場合，相関型ユニットの伝達関数は以下のように表せる．

$$u_i = \boldsymbol{w}_i \bullet \boldsymbol{v}^{(in)}, \tag{2-1}$$

$$v^{(out)}_i = f(u_i - b_i). \tag{2-2}$$

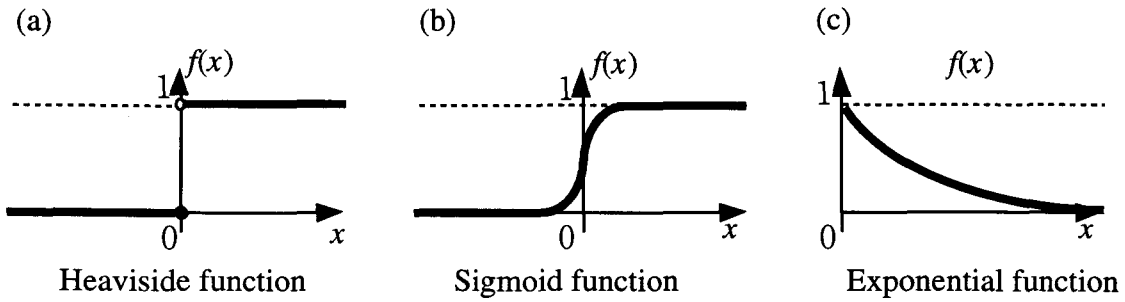

図2-2 ユニットの伝達関数に用いられる非線形関数；(a)ヘービサイド関数，(b)シグモイド関数，(c)指数関数

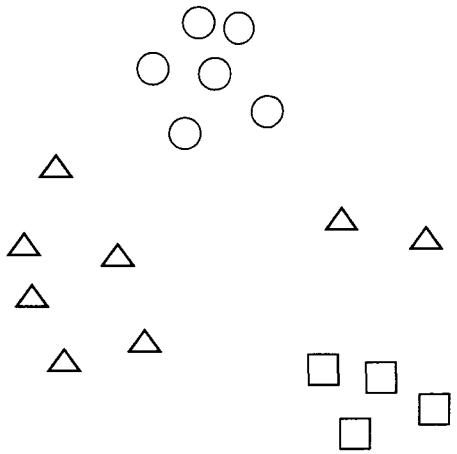

図2-3 画像ベクトル空間内の画像ベクトルの分布例

•はベクトル間の内積を表す．一方，距離型ユニットの伝達関数は以下のように表せる．

$$u_i = \| \boldsymbol{w}_i - \boldsymbol{v}^{(\mathrm{in})} \|, \qquad (2\text{-}3)$$

$$v^{(\mathrm{out})}_i = f(b_i - u_i). \qquad (2\text{-}4)$$

$\| \cdot \|$は，ベクトルのユークリッドノルムを表す．ユニットの出力は，1か0のいずれかをとる．ユニットの出力が1のとき，そのユニットは活性化しているという．

縦n画素，横m画素の画像の各画素を成分とした$n \times m$次元のベクトルを考える．このベクトルを画像ベクトルと呼ぶ．$n \times m$次元の認識対象の画像ベクトルは，$n \times m$次元超空間である画像ベクトル空間内に分布する．図2-3に，画像ベクトルの分布の例を，簡単化のために二次元画像ベクトル空間で示す．画像ベクトルを

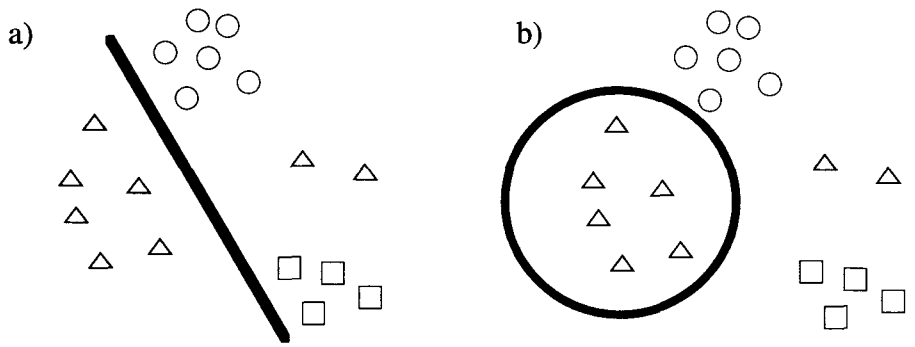

図2-4 単一ユニットによる画像ベクトル空間の分割の例；
(a)相関型ユニットの場合，(b)距離型ユニットの場合

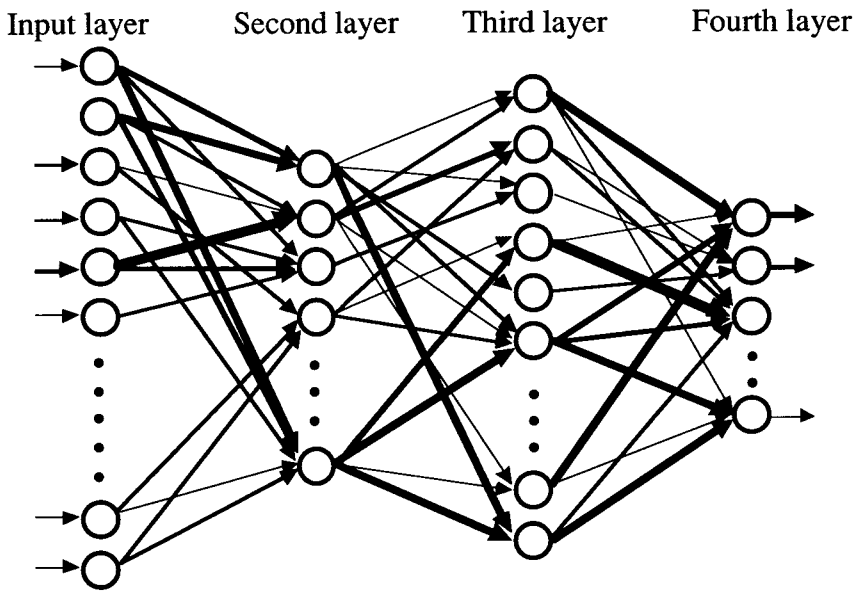

図2-5 四層階層型ニューラルネットワーク

○，△，□で表す．○，△，□は，それぞれ異なるクラスに属する画像ベクトルを表すとする．画像認識は，画像ベクトル空間をクラスに応じて分割する問題と捉えることが出来る．単一ユニットへの入力ベクトルを画像ベクトルとした場合，そのユニットの出力が活性化するか否かで画像ベクトル空間が二つに分割される．これは，単一ユニットが画像を二つのクラスに分割する能力を持つことを意味する．図2-4(a)(b)にそれぞれ相関型と距離型の単一ユニットによる画像ベクトル空間の分割の例を示す．相関型ユニットでは，画像ベクトル空間は超平面により分割され，距離型ユニットでは超球面により分割される．

B.　階層型構造の認識能力

図2-5に階層型ニューラルネットワークの構造の一例を示す．階層型構造においては，層間にはユニット間の結合が存在するが層内には結合が存在しない．第一層は入力層(input layer)と呼ばれ，各ユニットが入力ベクトルの各成分を出力する．最終層のどのユニットが活性化するかにより入力画像の属するクラスを認識する．

相関型ユニットにより構成される階層型ニューラルネットワークについて，各層のユニットが分割する画像ベクトル空間内の領域の形状について述べる[90]．第二層の各ユニットは，それぞれ２．２．１Ａで述べたように超平面による画像ベクト

ル空間の二分割を行う．図2-6(a)に，正方形で囲まれる領域が画像ベクトル空間の一部であるとした場合の二分割の例を改めて示す．第三層の各ユニットは，第二層の複数の二分割の結果としてのユニットの出力を入力とし，それらの重み付き線形総和演算と非線形演算を行う．その結果，図2-6(b)のように，第三層の各ユニットは，画像ベクトル空間を二分割の合成により得られた凸型の領域に分割する．第二層のユニットの数を増やすと凸型の領域の頂点の数を増やすことが可能である．さらに，第四層の各ユニットは，画像ベクトル空間をそれら凸型の領域の合成により得られる凹型の領域に分割することができる．図2-6(c)に，分割される凹型の領域の例を示す．従って，四層の階層型のニューラルネットワークは，中間層と呼ぶ第二層，第三層のユニット数に制限が無いとした場合，画像ベクトル空間を任意の領域に分割することができる．これは，四層の階層型ニューラルネットワークは，任意の画像を分類する能力を持つことを意味する．一般的に，ユニットの非線形関数にシグモイド関数（図2-2(b)）などの飽和型非線形関数を用いた場合は，四層の必要はなく，三層のニューラルネットワークで任意の領域を分割できることが知られている[84,85]．

第二層が距離型ユニット，第三層が相関型ユニットにより構成される三層の階

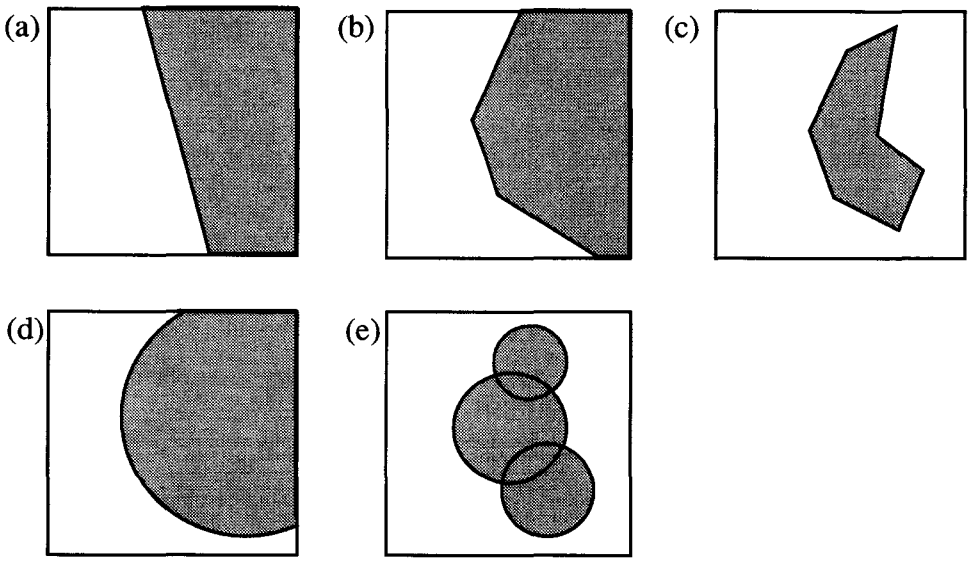

図2-6 ユニットの伝達関数の型，階層性による画像ベクトル空間の分割の違い；相関型ユニットを用いた(a)二層，(b)三層，(c)四層のネットワークの場合と，距離型ユニットを用いた(d)二層，(e)三層のネットワークの場合

層型ニューラルネットワークでは，第二層のユニットは，それぞれ２．２．１Aで述べたように超球による画像ベクトル空間の二分割を行う．図2-6(d)に改めて，超球による分割の例を示す．第三層のユニットは，図2-6(e)に示すように，画像ベクトル空間をそれら超球で合成された領域に分割する．第二層のユニット数に制限を加えない場合，このネットワークは画像ベクトル空間内を任意の領域に分割することができる．距離型ユニットの非線形関数に指数関数（図2-2(c)）を用いた場合には，RBF(radial basis function)ネットワーク[91-93]と呼ばれ，有限の実変数を持つ任意の関数の近似を行うことが出来ることが知られている[94]．

２．２．２　誤差逆伝搬学習法[23]

階層型ニューラルネットワークの学習法として，誤差逆伝搬学習法は，現在もっともよく用いられている学習法である．ただし，実際の利用に際して多くの経験的知識を必要とする．誤差逆伝搬学習法の原理と問題点について述べる．

A.　学習原理と手続き

誤差逆伝搬学習法には，関数の極小値を求める手法の一つである最急降下法が用いられる．最急降下法は，関数の微分値の大きさに従って，関数の値がより小さくなる方向に関数の引数を修正することを繰り返して極小値を得る手法である．学習には，学習用の画像ベクトルとそれが入力された時の理想的な出力ベクトルの対を用いる．誤差逆伝搬学習法において極小値を求める関数Eは，i番目の学習用の画像ベクトルを$\boldsymbol{v}^{(\mathrm{in})}{}_i$，そのときのネットワークの理想的な出力ベクトルを$\boldsymbol{t}_i$，ネットワークの実際の出力ベクトルを$\boldsymbol{v}^{(\mathrm{out})}{}_i$としたとき，以下のように表される．

$$E = \Sigma_i \| \boldsymbol{v}^{(\mathrm{out})}{}_i - \boldsymbol{t}_i \|^2. \qquad (2\text{–}5)$$

関数Eは誤差関数と呼ばれる．誤差関数の値が小さいほど，ネットワークが望みの結果を出力していることを表す．ここで，ネットワークのパラメータを誤差関数の引数として，最急降下法を用いる．ネットワークのパラメータである全ての結合荷重としきい値をまとめてパラメータベクトル$\boldsymbol{p}$で表した場合，i番目の出力ベクトル$\boldsymbol{v}^{(\mathrm{out})}{}_i$は，入力された画像ベクトル$\boldsymbol{v}^{(\mathrm{in})}{}_i$とパラメータベクトル$\boldsymbol{p}$の関数となる．式(2-5)より，誤差関数$E$は学習用の画像ベクトルと理想的な出力ベクト

ルの対の集合 $\boldsymbol{v}^{(in)}{}_1, \boldsymbol{t}_1, \ldots, \boldsymbol{v}^{(in)}{}_N, \boldsymbol{t}_N$ とパラメータベクトル $\boldsymbol{p}$ の関数となる．Nは，学習用の画像ベクトルの総数とする．パラメータベクトル $\boldsymbol{p}$ の修正は，最急降下法により次式に従って行われる．

$$\boldsymbol{p}' = \boldsymbol{p} - c\partial E(\boldsymbol{p}, \boldsymbol{v}^{(in)}{}_1, \boldsymbol{t}_1, \ldots, \boldsymbol{v}^{(in)}{}_N, \boldsymbol{t}_N)/\partial \boldsymbol{p}. \qquad (2\text{-}6)$$

$\boldsymbol{p}'$は修正後のパラメータベクトル，cは学習係数とする．式(2-6)を各パラメータごとに書き直し整理すると，最終層での誤差を信号の伝搬方向とは逆方向に伝搬させることで，各層におけるパラメータの修正値が計算される．誤差信号を信号の伝搬方向とは逆方向に伝搬させることから，この方法は誤差逆伝搬学習法と名付けられた．学習係数を適当に設定することで，誤差関数の極小値を得ることが出来る．ユニットの非線形関数には，微分可能なシグモイド関数(図2-2(a))が用いられる．

B.　特徴と問題点

誤差逆伝搬学習法が現在最も良く使われている理由は，簡単な学習手順で多層構造のニューラルネットワークの学習を行えるからである．出力層での誤差を認識時の信号伝達方向とは逆方向に伝搬させることにより，容易に全ての層の結合荷重等のパラメータの修正値が求められる．また，誤差逆伝搬学習法で学習された多層構造のニューラルネットワークは，ネットワークサイズが的確である場合，ユニットを効率よく使った認識処理が行われることが知られている[95]．誤差逆伝搬学習法が情報圧縮などに用いられるのもこのためである．

一方，誤差逆伝搬学習法には幾つかの問題点がある．誤差逆伝搬学習法の行う基本的な最急降下法では，誤差関数の極小値の一つしか求まらない．そのため，誤差の大きな極小値に収束し，学習できない場合がある．現在，適当な初期値をあらかじめ経験的に与えているのが現状である．また，ネットワーク規模（階層数，ユニット数）の決定は経験的に行うか，複雑な規模自動設定法[62,63]を用いる必要がある．ネットワーク規模の決定は，２.２.４で述べる認識結果の信頼性にかかわる重要な問題である．誤差逆伝搬学習法は，パラメータの漸近的な修正を行うため，学習が完了するまで非常に時間がかかることも問題である．

２.２.３　RCE学習法[21,86,89]

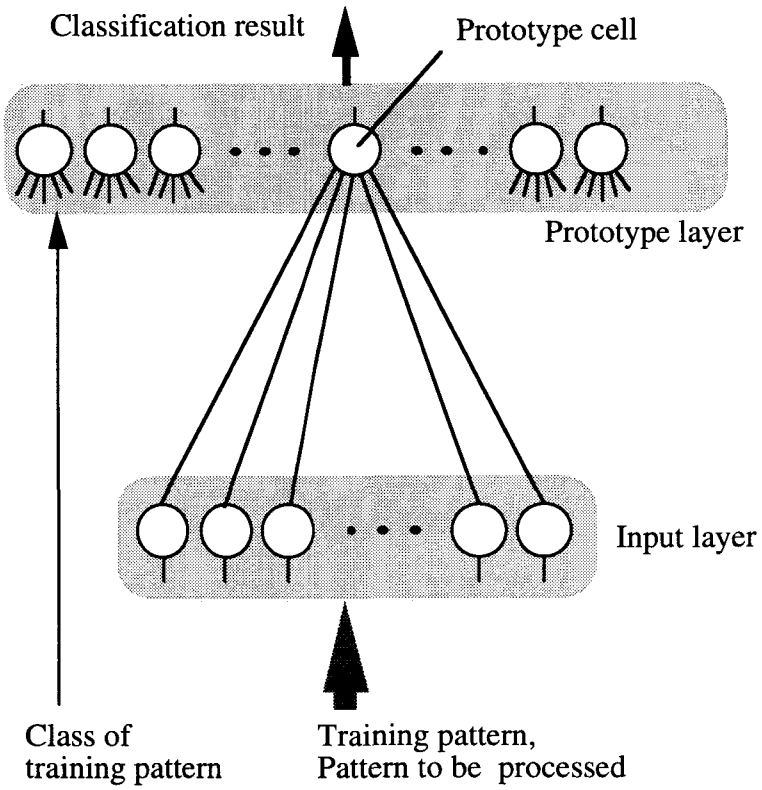

図2-7 二層のRCEネットワーク

階層型ニューラルネットワークの学習法として，与えられる例からユニット数を自動的に決定し，学習時間が短いという特徴を持つRCE学習法について説明する．RCE学習法で学習されるネットワークをRCEネットワークと呼ぶ．

A. RCEネットワーク

RCE学習法は，2.2.1Bで述べた第二層の距離型ユニット，第三層の相関型ユニットより構成される三層階層型ニューラルネットワークの学習を行う．RCE学習の結果，第二層のユニットは画像ベクトル空間内に同じクラスに属する学習用の画像ベクトルのみを囲んだ超球を形成する．第二層から第三層は，その超球の囲む入力ベクトルがどのクラスに属するかについての写像を行う．第二層の各ユニットにクラスに関する記号を与えることにより，二層のニューラルネットワークで三層のネットワークと等価な処理が行える[87]．以後，二層のRCEネットワークを用いてRCE学習法を説明する．

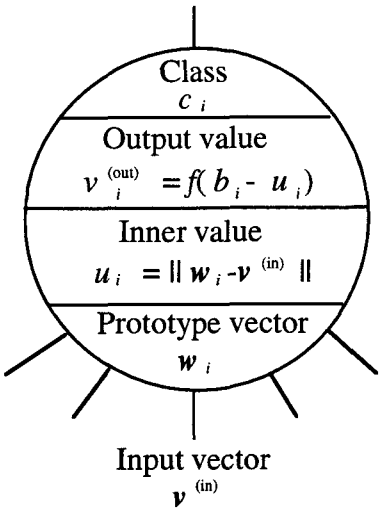

図2-8 RCEネットワークの原型ユニットの構造

図2-7に，説明に用いる二層のRCEネットワークを示す．第一層を入力層，第二層を原型層(prototype layer)と呼ぶ．原型層のユニットとそのユニットへの結合ベクトルは，それぞれ原型ユニット(prototype unit)と原型ベクトル(prototype vector)と呼ぶ．原型ユニットの構造を図2-8に示す．全ての原型ユニットは同じ構造をしている．それぞれの原型ユニットは，パラメータとして，原型ベクトル(prototype vector) w_i，内部値(inner value) u_i，しきい値(threshold) b_i，出力値(output value) $v^{(out)}_i$，クラス(class) c_i を持つ．既に述べたように，距離型の原型ユニットの伝達関数は式(2-3),(2-4)で表される．

B. 認識手続きと学習手続き

認識すべき対象がネットワークに与えられたとき，各原型ユニットの信号伝達処理の結果，原型層で活性化したユニットの中で内部値が最小の値を持つユニットの属するクラスが分類の結果となる．もし，ユニットが一つも活性化しない場合は，その画像は学習されたどのクラスにも属さないものとして棄却される．

学習用の画像ベクトルとその属するクラスの対（学習対）の集合（学習集合）$(v^{(in)}_j, t_j)$ $j = 1 \sim N$ がネットワークの学習に使われるとする．この時，N は学習集合内の学習対の数である．学習対はランダムにネットワークに与えられる．ネットワークの初期状態において，原型ユニットは一つも存在しない．学習により原型ユニットが加えられ，原型ユニットのしきい値が修正される．

今，j番目の学習対がネットワークに入力されたときのi番目の原型ユニットの

パラメータの修正を考える．式(2-3),(2-4)に従い，ユニットは，入力された学習用の画像ベクトルを伝達する．その時，ユニットが活性化し，かつ，そのユニットの属するクラスが入力された学習対のクラスと異なる場合のみ，そのユニットのしきい値 b_i を u_i に修正する．上記の修正後，再び式(2-4)の演算のみを行った結果，全ユニットの中で活性化するユニットが全く存在しない場合，新しい原型ユニットがネットワークに加えられる．新しく加えられたユニットの原型ベクトルは，入力された学習用の画像ベクトルと同じベクトルに設定される．しきい値は，入力された画像ベクトルのクラスとは異なるクラスに属するユニットが持つ内部値の中で最小の値に設定される．クラスは，入力された画像ベクトルのクラスに設定される．原型ユニットが全く存在しない初期状態の場合，しきい値は表現可能な出来るだけ大きな値に設定される．全ての学習対がネットワークに与えられるまでの間を1エポックと呼ぶ．学習は1エポックの間，パラメータが全く修正されなくなるまで続けられる．

RCE学習法の具体的な例を示す．図2-9(a)-(g)に，入力が二画素からなる二次元画像ベクトルである場合のネットワークの学習の進行例を示す．白い四角（□）と白い三角（△）はそれぞれ第1番目と第2番目のクラスに属する学習用の画像ベクトルを表し，黒い四角（■）と黒い三角（▲）は，それぞれのクラスに関してラベル付けされた第1番目と第2番目のクラスに属する原型ベクトルを表すものとする．原型ユニットによって形成される超球は，二次元画像ベクトル空間では円を形成する．認識対象の画像ベクトルが，あるユニットの円の内側にある場合に，そのユニットは活性化する．各状態において，図中の×印は，その時にネットワークに提示されている学習用の画像ベクトルを表す．

図2-9(a)は，原型ユニットが全く存在しないネットワークの初期状態を示す．図2-9(b)において，最初の原型ユニットが，黒い三角で表すように与えられている．円は半径が大きすぎるため図では描かれていない．図2-9(c)は，入力ベクトルと異なるクラスを持つ活性化した黒い三角で表すユニットのしきい値の修正と黒い四角で表す新しい原型ユニットの追加の状態を示す．図2-9(d)のように，次の入力が正しく分類されたとき，パラメータは全く修正されない．図2-9(e)において，活性化したユニットが存在しないため新しい原型ユニットを追加した状態を示す．図2-9(f)において，誤って活性化したユニットのしきい値のみを修正した状態を示

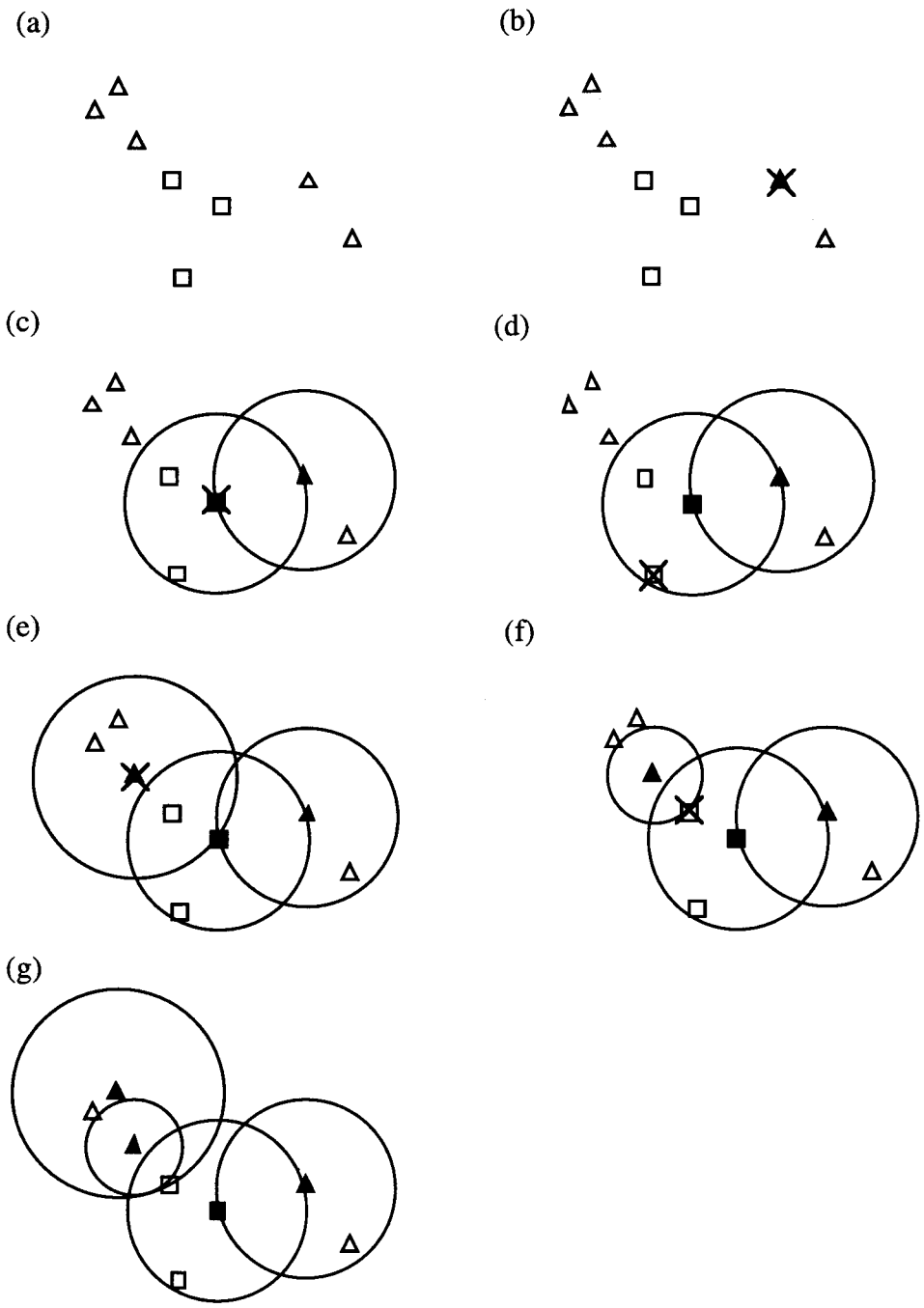

図2-9 RCE学習の進行例

す．図2-9(g)はネットワークの最終状態を示す．

C.　特徴と問題点

RCE学習法の主な二つの特徴と二つの問題点を挙げる．第一番目の特徴は学習

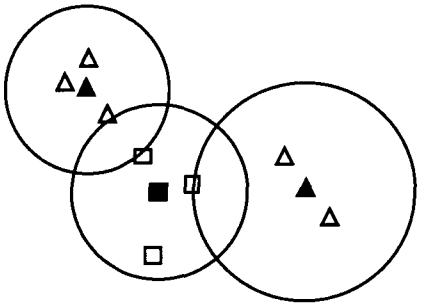

図2-10 原型ベクトルの最適な配置

の高速性である．RCE学習法において，生成されたユニットの結合荷重は学習によって修正されない．また，唯一修正されるしきい値も学習の各時点で最適な値が与えられる．そのため，図2-9に示した例のように，学習が高速で，通常は数エポックで学習が終了する．第二番目の特徴は，ネットワーク規模が問題に応じて自動拡大されることである．RCE学習法において，画像ベクトル空間上に分布した画像ベクトルの分布の複雑さに応じてユニットが生成される．ユニット数に制限を加えない限り，画像ベクトルのどのような分布も学習することができる．ただし，同じ画像ベクトルが2つ以上のクラスに属するような分類は，現モデルにおいて行うことが出来ない．

RCE学習法の第一番目の問題点は，生成されたユニットの原型ベクトルが修正されないため，画像ベクトル空間内で無駄な原型ベクトルの配置を行いがちなことである．図2-10に，図2-9(g)の例の原型ベクトルの最適と考えられる配置を示す．現在，原型ベクトルの最適な配置のために，RCE学習法を基本として誤差逆伝搬学習法により原型ベクトルの修正を行う学習法が提案されている[96]．ユニットの非線形関数に指数関数を用いたRBFネットワーク[91-93]においても，RCE学習法的な学習と誤差逆伝搬学習法を融合した学習法[97,98]が原型ベクトルの最適な配置のために提案されている．しかし，これらは最急降下法を用いた漸近的な原型ベクトルの修正であるため，学習速度が遅くなることが欠点である．RCEネットワークと非常によく似たDystalネットワーク[99]では，漸近的な手法を用いずに原型ベクトルの修正を実現している．この手法は，第3章で提案するネットワークモデルの学習に用いられる．第二番目の問題点は，距離型ユニットを用いるため画像ベクトル空間内の分割すべき領域を超球で覆わざるを得ないことである．そのため，相関型ユニットを用いたニューラルネットワークにより小規模なネットワ

ークで学習できる簡単な問題でも，距離型ユニットでは多くのユニットが生成されることがある．距離型ユニットで多くのユニットが生成される問題に，画像の局所的な領域のみを基に分類する問題がある．この時，その局所的な領域以外はどのような値もとることが可能である．その画像ベクトルは画像ベクトル空間内で画像の局所的な領域以外の成分方向に細長く広がる．細長い領域を，その形状に応じて覆うには多数の超球を必要とする．

2.2.4　棄却能力

ニューラルネットワークを，実世界における認識結果の信頼性の高い視覚情報認識装置として用いるためには，誤認識の原因となる入力画像を棄却する必要がある．ニューラルネットワークの棄却能力は，これまでよく研究されてきた汎化問題と関係している．しかし，従来の汎化問題に対する取り扱いでは棄却能力を十分に議論できない．

A.　汎化問題

ニューラルネットワークは，学習が完全に行われ，学習終了時のネットワークのパラメータが認識時のパラメータと同じならば，学習集合に含まれる入力をすべて正しく分類することが出来る．しかし，学習集合に含まれない入力を正しく処理する保証はない．この学習集合に含まれない入力に対する処理の問題は，汎化問題と呼ばれている．また，学習集合に含まれない入力に対して正しい処理を行う能力は汎化能力と呼ばれている．第1章でも述べたが，学習集合に含まれない入力は以下の二種類に分けることができる．

(A)学習集合を構成するクラスに属する入力．

(B)学習されなかったクラスに属するか，どのクラスにも属さない未知入力．

(A)のタイプの入力に対しては正しく分類し，(B)のタイプの未知入力に対しては棄却することにより，全ての入力に対して汎化能力の高いネットワークが実現できる．現在まで，(A)のタイプの入力に対する汎化能力の向上については，かなり研究されてきた[58-66]．

(A)のタイプの入力に対する汎化能力は，一般にネットワークのもつパラメータ数と学習集合の要素数によって決定される．Amariらによって提案されたネット

ワークの汎化能力とネットワークのパラメータ数，学習集合の要素数の関係を与える理論[58]は，汎化能力の高いニューラルネットワークを構築するための指針を与える．学習集合に含まれない入力に対する誤認識率をε，学習集合の要素数をt，ネットワークのパラメータ数をnとした時，誤認識率εは，

$$\varepsilon \approx n/t, \tag{2-7}$$

で与えられる．この時，学習により，学習集合に含まれる全ての入力が正しく分類できたとする．また，(A)のタイプの学習集合を構成するクラスに属する入力のみを考えるとする．式(2-7)より，(A)のタイプの入力に対する汎化能力は，学習集合の要素数に比例し，ネットワークのパラメータ数に反比例することが分かる．得られる学習集合の要素数には限界が存在するため，ネットワークのパラメータ数を小さくすることが汎化能力の向上には有効である．誤差逆伝搬学習法において既に述べたユニット数を自動設定する手法[62,63]やRCE学習法[21,86-89]等は，この指針に基づく汎化能力の高い学習法である．

B.　認識率と誤認識率

ここで，誤認識率を認識率とともに定義することにより，棄却能力の向上のために棄却するべき誤認識の原因となる入力画像を明確にする．また，棄却能力と汎化能力の関係を述べる．正しく分類されるべき２．２．４Ａで示した(A)のタイプの入力画像は，ニューラルネットワークにより，正しく分類されるか，誤って分類されるか，誤って棄却されるかのいずれかである．棄却されるべき(B)のタイプの入力画像は，ニューラルネットワークにより，誤って分類されるか，正しく棄却されるかのいずれかである．この時，認識率は，認識対象である(A)のタイプの入力画像の中で，正しく分類された入力画像の割合である．従って，認識率に関しては，２．２．４Ａで述べた(A)のタイプの入力画像に対する従来の汎化能力で議論することが出来る．一方，誤認識率は，(A)と(B)の両方のタイプの入力画像の中で，誤って分類された入力画像の割合で定義される．誤認識率は，(A)のタイプの入力画像に対する誤認識と(B)のタイプの入力画像に対する誤認識を考える必要がある．すなわち，今まで扱われていなかった(B)のタイプの入力画像に対する汎化能力の向上が，棄却能力の向上に必要である．また，(A)のタイプの入力画像に対する汎化能力においても，誤認識の原因となる画像の棄却を考える必要があ

る．ここで，(B)のタイプの入力画像を未知画像と呼ぶのに対して，誤認識された(A)のタイプの入力画像を曖昧な画像と呼ぶことにする．

例を用いて，ニューラルネットワークの認識率，誤認識率について具体的に説明する．図2-11(a)-(d)の外枠に囲まれる領域が画像ベクトル空間の入力対象となる全体の領域であるとする．認識処理において，入力対象となる領域内の入力画像の生起確率はすべて一定であるとする．図2-11(a)の灰色の領域は，学習に用いた二つのクラスの領域であるとする．クラスの違いは灰色の濃さの違いで表した．図2-11(b)の灰色の領域は，ネットワークがクラスに属する画像ベクトルとして分類する領域であるとする．この時，ネットワークにより正しく分類される入力画像の占める領域を図2-11(c)に黒色で表し，誤分類される入力画像の占める領域を図2-11(d)に黒色で表す．画像ベクトル空間中の入力対象となる領域全ての体積をV_w，図2-11(a)の灰色の領域である(A)のタイプの入力画像の占める領域の体積をV_c，図2-11(c)の黒色の領域である学習後に正しく分類される入力画像の占める領域の体積をV_r，図2-11(d)の黒色の領域であるネットワークにより誤分類される入力画

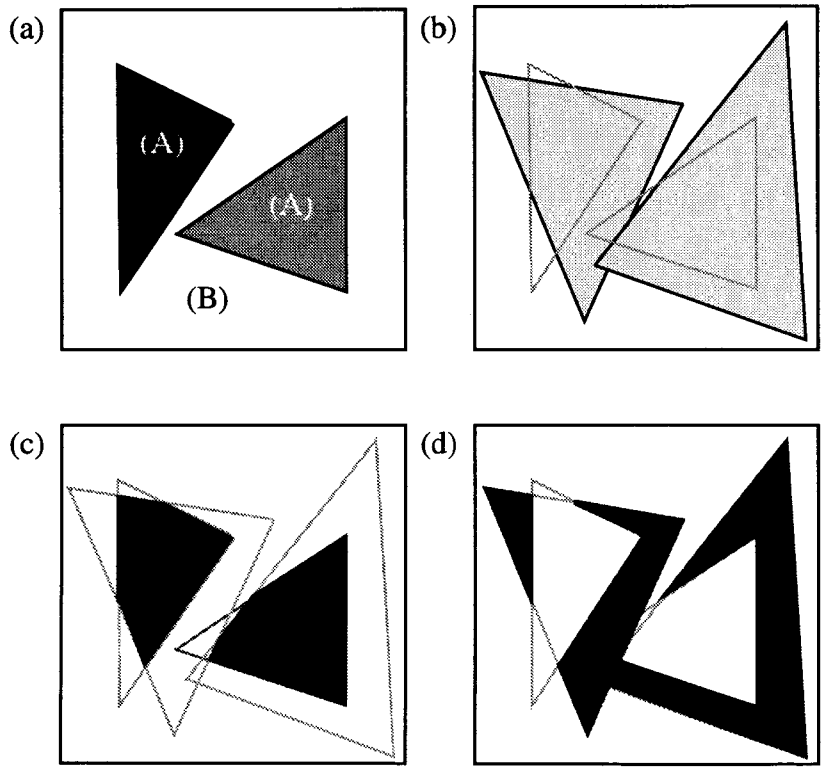

図2-11 二次元の画像ベクトル空間における学習の例

像の占める体積をV_mとした時，認識率（R），誤認識率（E）は以下のように表される．

$$R = V_r / V_c, \tag{2-8}$$

$$E = V_m / V_w. \tag{2-9}$$

式(2-8),(2-9)から，図2-11(c)の黒色の領域が大きいほど認識率が高く，図2-11(d)の黒色の領域が小さいほど誤認識率が低いことが分かる．図2-11(d)の誤認識される領域である黒色の領域の中で，(A)のタイプに含まれる画像が曖昧な画像である．既に述べたように，誤認識率には，(A)のタイプの入力画像に関する誤認識と(B)のタイプの入力に関する誤認識の両方が関係している．そこで，第三章において提案するネットワークの棄却能力を評価する場合，(A)のタイプの画像に対する誤認識率の低さと，(B)のタイプの画像に対する棄却率の高さにより棄却能力を評価する．

RBFネットワークを用いて，入力画像が未知画像や曖昧な画像であるかを識別する手法が提案されている[100]．画像ベクトル空間内で，入力画像ベクトルの存在する位置における学習用の画像ベクトルの確率密度の検出を行う．その確率密度の低い場合は，入力画像は未知画像または曖昧な画像として棄却される．

2.3　視覚情報認識のためのニューラルネットワーク構造

本節において，変形の大きい手書き文字のための位置不変結合型構造，クラスの多い活字漢字や機能的に分割できる認識過程等のためのモジュール型構造，階層的な知識表現のための調和モデル型構造の三つの視覚情報認識に有効なニューラルネットワーク構造について説明する．

2.3.1　位置不変結合階層型構造

手書き文字のように，同一クラスの画像間に大きな変形を含む画像ベクトルは，画像ベクトル空間内で互いのクラスが複雑に入り組んだ分布を呈する．既に述べたように，階層型ニューラルネットワークは，原理的に画像ベクトル空間のどのような分割でも実現することができる．しかし，互いのクラスが複雑に入り組んだ分布を正しく分割するためには，その分布を表せるだけの学習集合が必要になる．そのような学習集合を得ることは困難であり，また，仮に得たとしても非常

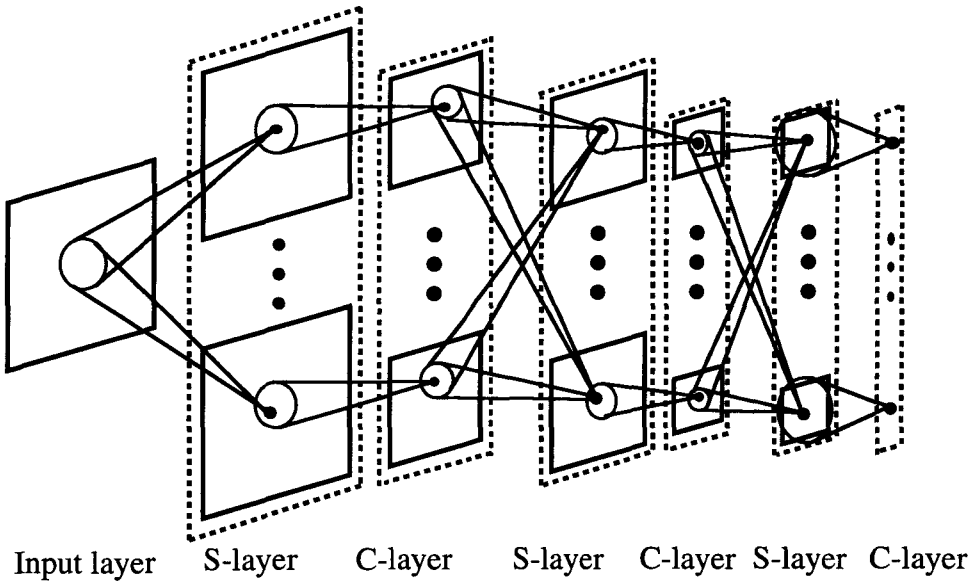

図2-12 ネオコグニトロンの構造

に大きな集合になり，学習時間を考えると実質的に学習が不可能であると考えられる．文字の変形に関して出力結果が左右されない認識処理を行う構造をあらかじめネットワークに与えることにより，少ない学習集合でネットワークの学習が容易に実現できる．

図2-12に，変形の大きい文字等の認識に適したニューラルネットワーク構造の一例を示す．これは，Fukushimaの提案したネオコグニトロン[20]の構造である．各層内のユニットは，幾つかのクラスタを形成している．前後する層に含まれるクラスタ間は，位置不変な結合荷重を持つ接続が行われている．位置不変な結合荷重では，図2-13に示すように，前層のユニットの二次元座標を(i,j)，後層の二次元座標を(k,l)，各位置のユニット間の結合荷重を$w(i,j;k,l)$とした場合，周辺部を除いて以下のような式が成り立つ．

$$w(i,j;k,l) = w(i-k, j-l; 0,0) . \qquad (2\text{-}10)$$

i, j, k, l は$-n$からnまでの間の整数とする．位置不変な結合荷重を持つ階層型構造を，本論文において，位置不変結合階層型構造と呼ぶ．ネオコグニトロンの場合は，S層，C層と呼ばれる働きの異なる層が互い違いに配置された特殊な構造をしている．S層は，前層の局所的な特徴を検出する層で，前層の全てのクラスタと結合している．C層は，前層で検出された特徴の位置をぼかす働きをする層であり，前層の対応するクラスタとのみ結合している．後方の層のユニットほど入力

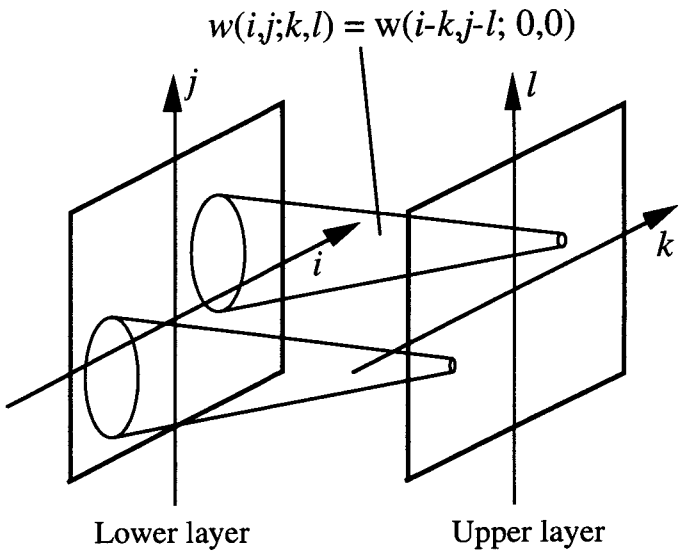

図2-13 位置不変結合

層の大きな領域から信号を得るため，入力画像の大きな特徴に対して活性化する．最終層において，どのユニットが活性化されるかにより，入力画像の属するクラスを認識する．S層の特徴検出能力と，C層の位置ずれの大きさに関する許容能力により，局所的な特徴の位置関係が大きく変化した画像に対しても変化前の画像と同じ分類結果を得る構造をしている．したがって，各クラスに対して単一の画像のみでネットワークを学習させたとしても，各画像の局所的な特徴の位置関係が大きく変化する画像に対して同じ出力を得ることが出来る．Fukushimaは，ネオコグニトロンにより0から4までの変形のない数字を用いて自己組織化学習法による学習を行った．その結果，変形した数字に対しても正しい認識を実現している．

2.3.2　モジュール型構造

モジュール型構造とは，複数のニューラルネットワークを組み合わせた構造である．モジュール型構造のニューラルネットワークをモジュール型ニューラルネットワークと呼ぶ．モジュール型ニューラルネットワークは以下の3つの場合に特に有効である．

(A)膨大な数のクラスの存在する情報を記憶する必要がある場合．

(B)複雑な認識過程を複数の簡単な認識過程に分解できる場合．

(C)複数の情報を利用する場合．

はじめに，(A)の場合について説明する．実世界の視覚情報には，大きな変形が存在するだけでなく，非常に多くのクラスが存在する場合が多い．単一のネットワークで多くのクラスを学習することは，学習時間，ネットワークの規模等の観点から実質的に困難である．認識すべきクラスを分割して複数のニューラルネットワークモジュールで分担する手法がこの場合有効である．Iwataらは，JIS第一水準の印刷漢字2965文字を，第一層をKohonenの自己組織化学習法[101-103]で学習した単一のネットワークと，第二層から第四層までを誤差逆伝搬法で学習した144個のネットワークモジュールとで構成される大規模四層ニューラルネットワークにより認識した[104]．

次に(B)について説明する．複雑な認識処理を単一のネットワークで学習するこ

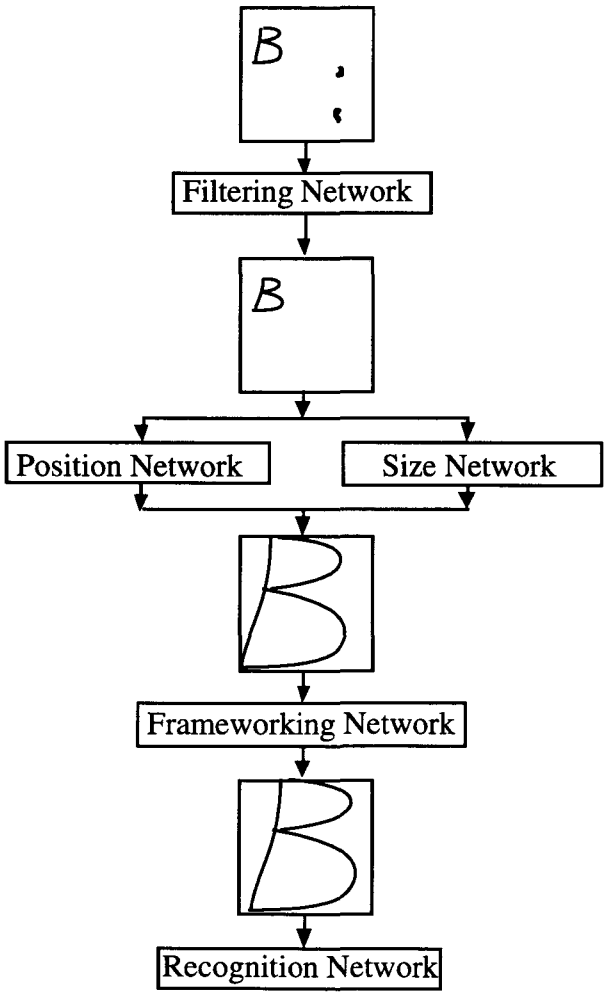

図2-14 モジュール型ニューラルネットワークによる機能分割

とは困難である．しかし，簡単な認識過程を行うネットワークを組み合わせることで複雑な認識処理を実現できる場合がある．例えば図2-14に示すように，任意の位置に書かれた文字の認識を，文字の切り出し機能と切り出された文字の認識機能等に分割し，それぞれの機能をネットワークモジュールで実現する試みが行われている[70]．

最後に(C)の場合について説明する．複数の情報源がある場合，それらをまとめて一つのニューラルネットワークに与えて認識するより，各情報を別々のニューラルネットワークで処理し，最後に統合するほうが効率よく認識処理が出来る場合がある．RCE学習法を用いたNLS(Nestor Learning System)学習法はその一例である[105]．また，文献14の中で，依田らが，LVQ[103]を改良し，それを基にしたモジュール型ネットワークにより印刷漢字の濃淡特徴とCDC特徴（Countour Disribution Code: 文字の輪郭の方向成分の分布）を各モジュールへの入力として認識実験を行った結果，学習・認識に関する効率の良さを確認したことが述べれている．

2.3.3　調和モデル[41]型構造

複雑な構造をした視覚情報は，大きな特徴（全体）とそれに含まれる小さな特徴（部分）の関係に基づく階層構造で記憶することが効率的である．2.3.1で既に述べた位置不変結合階層型構造においても，入力の大きな特徴に活性化する後方の層のユニットは，それに比べて小さな特徴に活性化する前層のユニットの信号を受けることにより，全体と部分の関係に基づく階層構造を記憶している．ここでは，複数の階層性を二層のネットワークに折り畳んだ調和モデルと呼ばれるネットワーク構造を紹介する．調和モデルは，対象の階層性を意識せずに階層的知識を記憶することができる．調和モデルの基本的構造は，第4章で提案する能動的な視覚情報認識モデルにおいて用いられる．

簡単な文字知覚モデルを例とした調和モデルの構造を図2-15に示す．第一層は表現層と呼ばれ，第二層は知識層と呼ばれる．表現層は特徴ノードと呼ばれるユニットから，知識層は知識アトムと呼ばれるユニットから構成されている．調和モデルには，階層的な知識の断片が層間の結合として記憶されている．例えば，図2-15では，ABLEという単語がA, B, L, Eという文字から構成されていることを，知識アトム一つと単語ABLE, 文字A, B, L, Eに関する各特徴ノードとの間の相互結

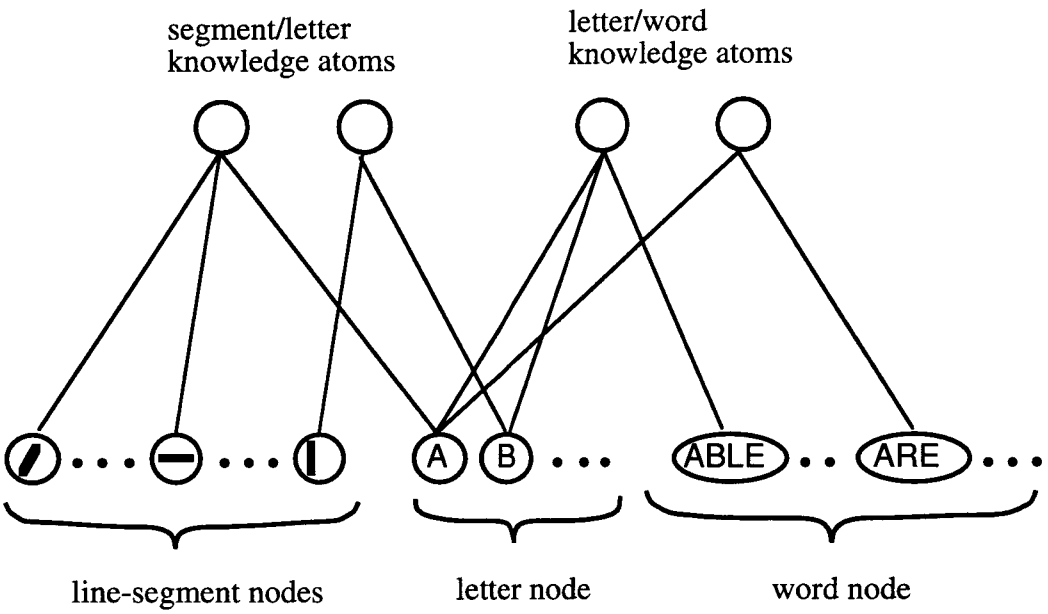

図2-15 調和モデルによる単語知覚モデル

合により記憶している．複数の知識の断片を記憶するだけで，二層のネットワークに階層的な知識が記憶される．各単語ノードは知識アトムを通して文字ノードと結合し，各文字ノードは同様に線分ノードと結合することにより，階層性を構成している．図2-16に，図2-15を階層性が明らかになるように書き直したネット

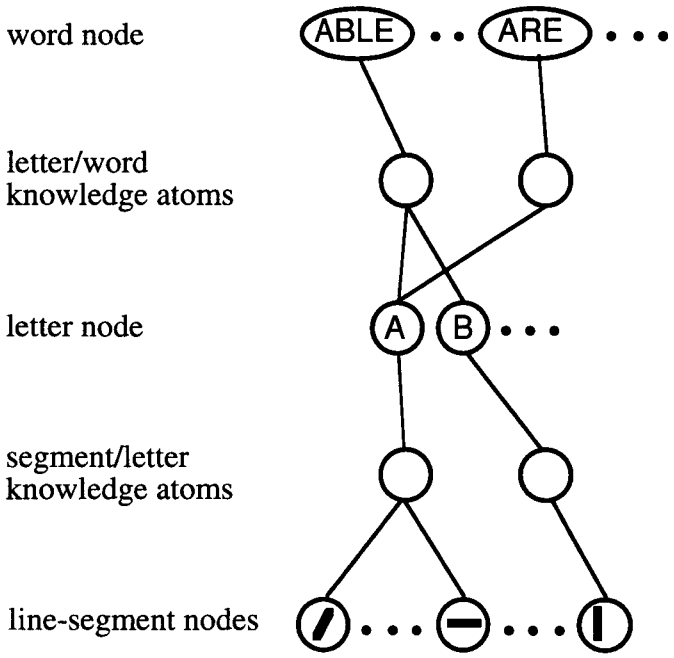

図2-16 調和モデルの階層的表現

ワークを示す．

調和モデルの各ノードは，確率的処理の加わった伝達関数を持つ相関型ユニットによって信号を処理する．対象に関する入力情報は，対応する特徴ノードを活性化させる．その信号が二層間を行きかい，ネットワークに記憶された複数の知識の断片と調和した状態でネットワークは安定する．そのとき活性化されている特徴ノードを調べることにより，入力情報と知識に調和した認識結果が得られる．

2.4　大規模光ニューラルネットワーク

光ニューラルネットワークは，視覚情報の認識に必要とされる大規模なニューラルネットワークを実現する手法として期待されている．ホログラム型と空間光変調器型の光ニューラルネットワークについて，その構造，認識過程，学習過程について説明し，大規模化について考察する．

2.4.1　ホログラム型光ニューラルネットワーク[45-47)]

ホログラム型光ニューラルネットワークの典型的な二層構造を図2-17に示す．認識過程にしたがってこの構造を説明する．入力層にはユニットの出力に応じて強度を変調した光を放射する発光素子(optical transmitter)が配置されている．発光素子からの出力信号光が，ホログラムによって回折され，出力層のユニットに付随する受光素子(optical receiver)に入射する．受光素子への回折効率により，その

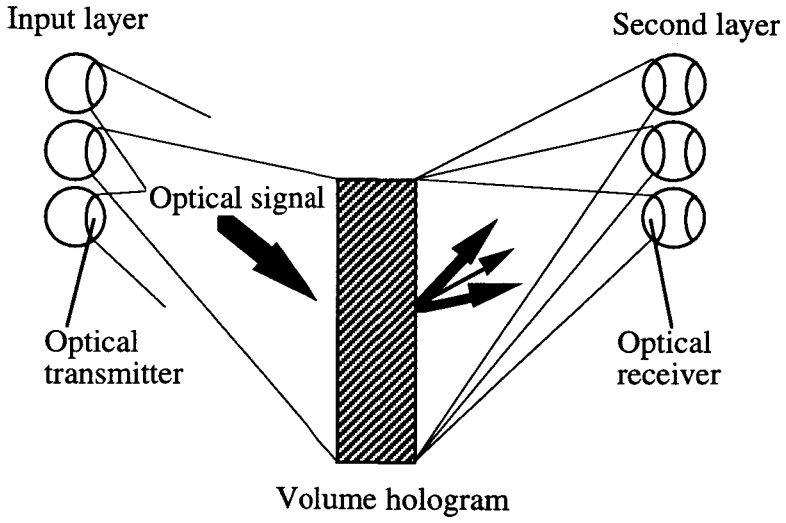

図2-17 ホログラム型光ニューラルネットワーク

受光素子への結合荷重の大きさが決まる．複数の発光素子からの光信号を受けた第二層目のユニットでは，得られた信号に非線形処理を施し，ユニットの出力を得る．三層以上のニューラルネットワークを実現する場合には，第二層のユニットは，得られた出力信号に応じて入力層と同じように信号光を第三層に向けて放射する．

学習は，ホログラムの多重記録により行われる．図2-18(a)(b)に，偏光スイッチ型の体積ホログラムと位相共役鏡を用いた，光学的位置合わせが不要である光ニューラルネットワークの学習法を示す[46]．偏光スイッチ型の体積ホログラムは，入射した光の強度分布に応じて異常光に対する屈折率分布が記録されるフォトリフラクティブ結晶である．常光は，ある向きを基準にして０度の直線偏光，異常光は，90度の直線偏光となるように結晶を配置するものとする．図2-18(a)は，ホログラムの記録方法を示す．入力層からは，学習用の入力画像によりユニットの出力値に応じた信号光（入力信号光）が，第二層からは，理想の出力，または，理想の出力と実際の出力との差に応じた信号光（教師信号光）が放射される．この時の各信号光の偏光状態は，入力信号光が45度の直線偏光，教師信号光が90度の直線偏光であるとする．また，入力信号光と教師信号光は互いにコヒーレントな光とする．入力信号光のうち，0度の偏光成分である常光は，そのまま位相共役

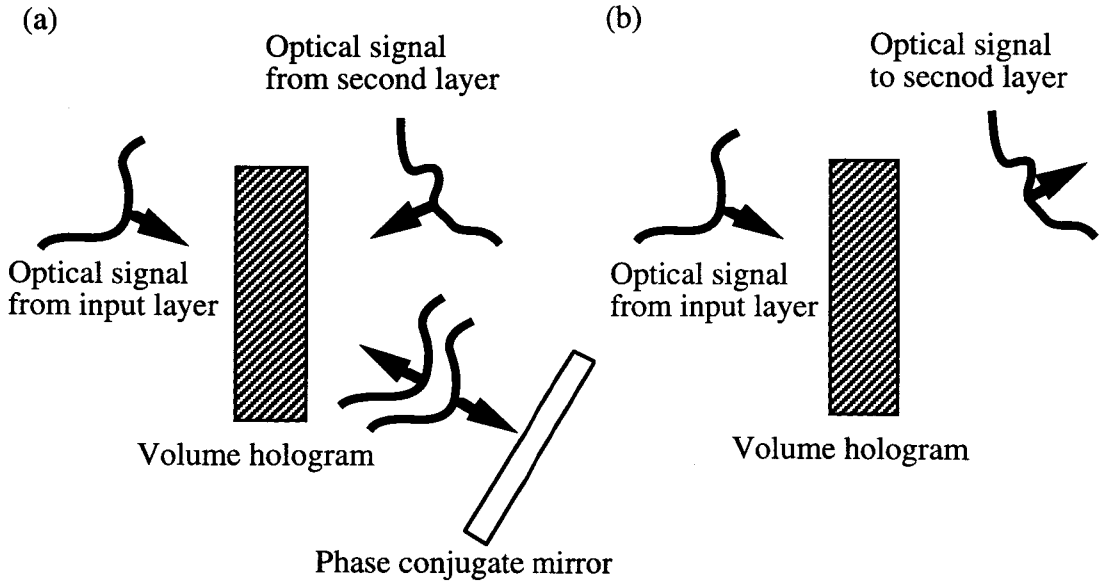

図2-18 ホログラム型光ニューラルネットワーク の
光学的位置合わせが不要となる学習方法

鏡に入射する．位相共役鏡により入力信号光と共役な光が反射される．ホログラムと位相共役鏡の間に45度の偏光回転素子を挿入しておくと，ホログラムに再入射する反射光の偏光方向は90度になる．この時，教師信号光と同じ偏光状態になるため，反射光と教師信号光による干渉縞がホログラムとして記録される．信号再生時（信号処理時）には，図2-18(b)に示すように，体積ホログラムに入射する入力信号光の90度方向の偏光成分である異常光は，生成された回折格子により回折され，教師信号光に共役な回折光が生成される．この教師信号光に共役な回折光の再生は，光を放射した入力素子，出力素子の間を光学的位置合わせを行わずに光学的に接続したことを意味する．このようなホログラムへの記録を多重化することにより結合荷重の学習が行われる．

ホログラム型は，理論的には，1立方センチメートルの体積ホログラムで10^{10}本程度の大規模な光接続を実現できるといわれている[47]．しかし，現状では多重露光の回数に制限があり，多重露光の回数を増加させて多くの情報を記憶することが出来ない．そのため，実質的に学習できる学習集合の大きさには，厳しい制限がある．

2.4.2　空間光変調器型光ニューラルネットワーク[42-45,48,49]

空間光変調器型の光ニューラルネットワークの典型的な二層構造を図2-19に示す．入力画像の空間光変調器上における展開法に，Ishikawaらによって提案された空間符号化法[48]を用いる．空間符号化法により，二次元の画像を一次元のベクトルに変換せず，二次元画像のままベクトル-マトリックス演算を扱うことが出来る．認識過程に従ってこの構造を説明する．入力層，第二層のユニットはホロ

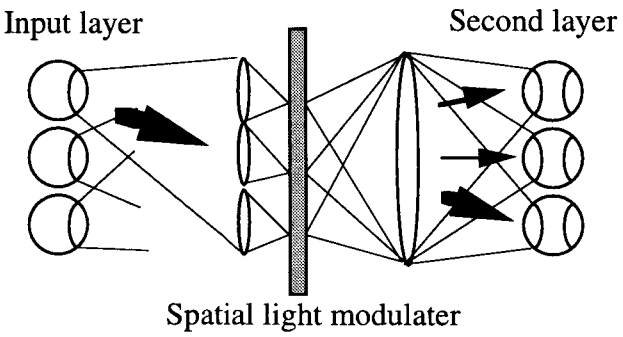

図2-19 空間光変調器型光ニューラルネットワーク

グラム型と同じ働きをする．入力画像は，レンズアレイ等の複製結像系により空間光変調器上に複製される．ここで，透過型の空間光変調器を考える．空間光変調器上の複製像は，空間光変調器により強度に関する変調を受けたあと第二層の受光素子に入射する．この時，空間光変調器上の変調の大きさ（透過率）により結合荷重の大きさが決まる．学習は，空間光変調器上に荷重の修正分を追加記録することにより行われる．空間光変調器型の光ニューラルネットワークは，現状では空間光変調器の分解能の限界やレンズ系の収差等により，大規模化に制限がある．また，ホログラム型のように光学的位置合わせの不要な光学系は提案されていない．しかし，ホログラム型に比べて，空間光変調器への追加書き込みは容易であり，学習集合の大きさの増加が期待できる．

2.5　結言

本章において，本論文における研究の背景となるニューラルネットワークによる視覚情報認識に関する従来法について説明した．2.2において基本的なニューラルネットワークの能力，学習法，棄却能力について述べた．学習法では，第3章で提案するモデルの基本となるRCEネットワークについて，学習の例を挙げて特に詳しく説明した．2.3において，視覚情報認識のためのニューラルネットワーク構造として，位置不変結合階層型構造，モジュール型構造，調和モデル型構造を取り上げ，それらの構造の持つ意味について説明した．調和モデル型構造は，現在までのところ視覚情報の認識には用いられていないが，第4章で提案するモデルの基礎となるので取り上げた．2.4において，大規模光ニューラルネットワークについて説明した．特に，既に提案されているホログラム型光ニューラルネットワークにおける光学系の自己調整機能の実現方法について詳しく述べた．

第３章　高い棄却能力を持つニューラルネットワークモデル

３．１　緒言

本章において，２．２．４Ｂで述べた，学習されなかったクラスに属するかどのクラスにも属さない未知画像や，学習集合を構成するクラスに属するが誤認識の原因となる曖昧な画像を棄却することが出来る新しいニューラルネットワークモデルを提案する[106,107]．本章で提案するモデルは，RBFネットワークを用いたモデル[100]のように，学習用の画像ベクトルの密度を推定することなく，未知画像や曖昧な画像を棄却することができる．この新しいネットワークモデルは，RCEネットワークを基に考案した．新しいネットワークは，RCEネットワークに比べて主に二つの特徴がある．第一の特徴は，原型ユニットが類間しきい値，類内しきい値と呼ばれる二つのしきい値を持つことである．各ユニットの二つのしきい値は，それぞれ画像ベクトル空間に二重の超球を形成する．二重の超球をもとに，分類される領域と棄却される領域の境界となる境界超球が決定される．境界超球の大きさは，望まれる棄却の程度に従い調整可能である．二重の超球を用いることから，新しいネットワークを二重超球(DH:double hypersphere)ネットワークと呼ぶ．第二の特徴は，原型ベクトルが修正されることである． DHネットワークにおける原型ベクトルの修正則は， Dystalネットワーク[99]における修正則と同じである．この修正により，原型ベクトルは未知画像や曖昧な画像を効率よく棄却するために近似的に最適化される．

３．２において，DHネットワークの構造，入力画像の分類法，ネットワークの学習法について述べる．また，二入力ユニットの場合のDHネットワークにおける学習過程の例を示す．この例により，ネットワークの学習方法を具体的に示す．また，この例を用いて，RCEネットワークとDHネットワークの棄却能力の比較を行う．３．３において，実用的なデータを用いたときの，RCEネットワークとDHネットワークにおける分類と棄却に関する評価を行う．

３．２　二重超球ネットワーク

３．２．１　構造と伝達関数

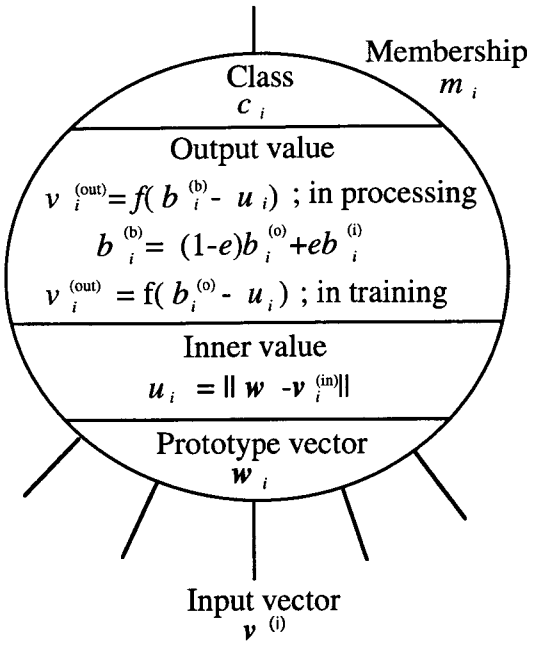

図3-1 二重超球ネットワークの原型ユニットの構造

二重超球(DH)ネットワークの構造はRCEネットワークの構造と同じである．しかし，図3-1に示すように，DHネットワークにおける原型ユニットの構造はRCEネットワークの構造とは異なる．それぞれの原型ユニットiはパラメータとして，原型ベクトル(prototype vector)w_i，内部値(inner value)u_i，類間しきい値(inter-class threshold)$b^{(o)}_i$，類内しきい値(intra-class threshold)$b^{(i)}_i$，境界しきい値(boundary threshold)$b^{(b)}_i$，出力値(output value)$v^{(out)}_i$，要素数(membership)m_iとクラス(class)c_iを持つ．$v^{(in)}$を原型ユニットへの入力ベクトルとした時，その原型ベクトルの伝達関数は，ネットワークの学習時には，

$$u_i = \| w_i - v^{(in)} \|, \tag{3-1}$$

$$v^{(out)}_i = f(b^{(o)}_i - u_i), \tag{3-2}$$

入力画像の分類時には，

$$v^{(out)}_i = f(b^{(b)}_i - u_i), \tag{3-3}$$

である．このとき，$\| \cdot \|$と$f(\cdot)$はそれぞれユークリッドノルム，ヘビサイド関数（図2-2(a)）を表すものとする．境界しきい値$b^{(b)}_i$は，

$$b^{(b)}_i = (1-e)b^{(o)}_i + eb^{(i)}_i, \tag{3-4}$$

で与えられる．このときeは0から1の間の値をとるパラメータである．パラメータeにより棄却の程度を変えることができる．

3.2.2　学習手続きと認識手続き

2.2.3Bと同じように学習対$(\boldsymbol{v}^{(in)}_j, t_j)$ $j=1\sim N$がネットワークの学習に使われ，ランダムな順序でネットワークに与えられる．ネットワークの初期状態において原型ユニットは一つも存在しない．学習により，新しい原型ユニットが加えられたり，現存する原型ユニットの類間しきい値，類内しきい値，要素数，原型ベクトルが修正される．これらの修正は全ての原型ユニットで同じ規則が用いられる．

j番目の学習対がネットワークに与えられたとき，i番目のユニットのパラメータの修正を考える．ユニットは式(3-1),(3-2)に従い学習用の画像ベクトルに応答する．その時，ユニットが活性化されない場合はなにも修正されない．ユニットが活性化され，そのクラスが学習画像のクラスと同じであり，かつ，その内部値が全ての活性化した原型ユニットの中で最大の値である場合，そのユニットの類内しきい値，要素数，原型ベクトルは以下のように修正される．$b^{(i)}_i{}'$, m_i', $\boldsymbol{w}_i'$がそれぞれ$b^{(i)}_i$, m_i, $\boldsymbol{w}_i$の修正後の値とすると，要素数は

$$m_i' = m_i + 1, \tag{3-5}$$

に，原型ベクトルは，

$$\boldsymbol{w}_i' = \boldsymbol{w}_i + (\boldsymbol{v}^{(i)}_j - \boldsymbol{w}_i)/m_i', \tag{3-6}$$

に，類内しきい値は，

$$b^{(i)}_i{}' = \max\{b^{(i)}_i, \| \boldsymbol{v}^{(i)}_j - \boldsymbol{w}_i' \|\}, \tag{3-7}$$

に設定される．これより，原型ベクトルは，画像ベクトル空間内で，そのユニットを活性化させた複数の画像ベクトルの平均ベクトルに修正される．一方，ユニットが活性化され，そのクラスが学習用の入力画像のクラスと異なる場合，類間しきい値$b^{(o)}_i$はu_iに修正される．この修正は，入力画像と異なるクラスに属するユニットの活性化を抑える．この時，更新された類間しきい値が類内しきい値より小さくなった場合，そのユニットはネットワークから取り除かれる．この不要なユニットの除去は，線形非分離問題の学習に不可欠である．

上記の修正の後に再び式(3-2)の演算を行った結果として原型ユニットが全く活性化されない場合，新しい原型ユニットがネットワークに加えられる．新しく加えられたユニットの原型ベクトルは，入力された学習用の画像ベクトルと同じベクトルに設定される．類間しきい値は，入力された画像ベクトルのクラスとは異

なるクラスに属するユニットが持つ内部値の中で最小の値に設定される．クラスは，入力された画像ベクトルのクラスに設定される．類内しきい値と要素数はそれぞれ0, 1の値に設定される．この時，原型ユニットが全く存在しない場合，類間しきい値は表現可能なできるだけ大きな値に設定される．

RCEネットワークでは，学習は，１エポックの間，全てのパラメータが全く修正されなくなるまで続けられるのに対して，DHネットワークでは，学習は，１エポックの間，原型ユニットが全く加えられなくなるまで続けられる．

認識時には，すべての原型ユニットで式(3-1),(3-3)に従い内部値u_iと出力値$v^{(o)}_i$を計算する．原型層で内部値が最小である活性化されたユニットの示すクラスが分類の結果である．もし，ユニットが一つも活性化されない場合は，その画像はどのクラスにも属さない画像として棄却される．

３.２.３　学習過程の例

２.２.３Bの図2-9と同じ例を用いて図3-2(a)-(g)にDHネットワークの学習過程の例を示す．２章と同じく，白い四角（□）と白い三角（△）はそれぞれ第１番目と第２番目のクラスに属する学習用の画像ベクトルを表し，黒い四角（■）と黒い三角（▲）は，それぞれクラスに関してラベル付けされた第１番目と第２番目のクラスに属する原型ベクトルを表すものとする．各状態において，図中の×印はその時ネットワークに提示されている学習用のベクトルを示す．一つの原型ユニットの二つのしきい値によって形成される二重の超球は，二次元画像ベクトル空間において二重の円を形成する．実線と破線の円は，それぞれ類間しきい値と類内しきい値により形成される．学習時に，入力画像ベクトルが，あるユニットの実線の円の内側にある場合に，そのユニットは活性化される．

図3-2(a)は，DHネットワークの原型ユニットが全く存在しない初期状態を示す．最初の原型ユニットが，図3-2(b)において加えられている．実線と破線の円はそれぞれの半径が無限大と０のため，図には描かれていない．次の入力が正しく分類されたとき，原型ベクトルと類内しきい値は図3-2(c)のように修正される．ここでも，類間しきい値は，無限大のために図には描かれていない．図3-2(d)において，修正された類間しきい値が類内しきい値より小さくなったため，その黒い三角で示す原型ユニットはネットワークから取り除かれ，新しい黒い四角で表す

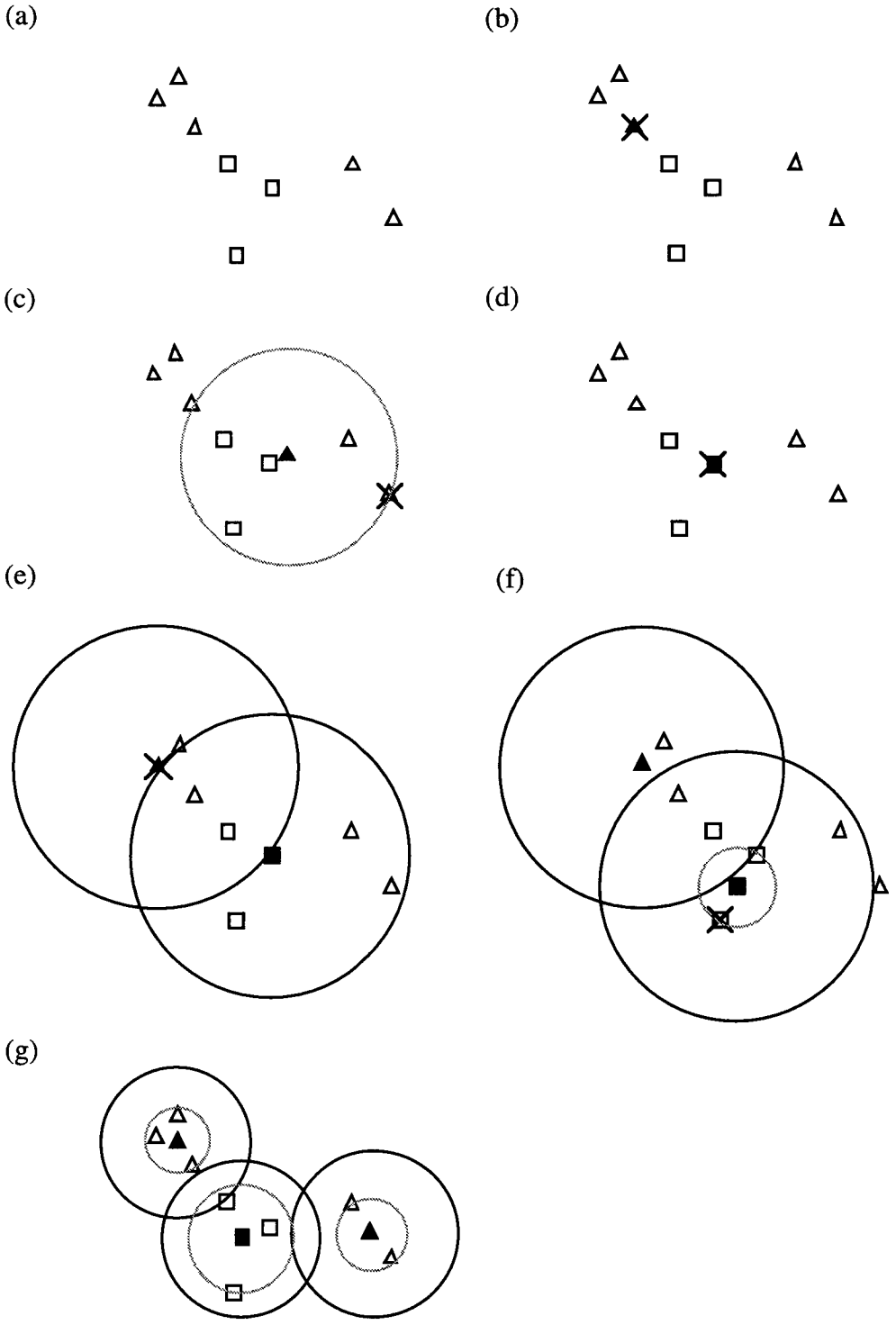

図3-2 二重超球ネットワークの学習の進行例

原型ユニットが加えられている．図3-2(e)は入力ベクトルと異なるクラスに属する活性化した黒い四角で表す原型ユニットの類間しきい値の修正と，黒い三角で

表す新しい原型ユニットのネットワークへの追加を示す．図3-2(f)において，黒い四角で表す原型ベクトルと類内しきい値を修正した状態を示す．図3-2(g)はネットワークの最終状態を示す．この時，それぞれの原型ベクトルは円で囲まれる学習用の画像ベクトルの集合の中心に位置している．また，それぞれの類内しきい値により形成される破線の円の半径は，その円が同じクラスに属する学習用の画像ベクトルをできるだけ多く含むという条件のもとでほぼ最小の大きさとなる．

RCEネットワークの学習過程において，新しく加えられた原型ユニットは，ネットワークに提示されている学習用の画像ベクトルを原型ベクトルとして記憶する．また，学習過程の間，この原型ベクトルは修正されない．この時，原型ベクトルとして利用された学習用の画像ベクトルは，常に正しく分類されることは明らかである．この場合，学習用の画像ベクトルの数には限界があるため，原型ユニットの追加は必ず停止する．以下，DHネットワークの学習過程の収束性について，問題を線形分離問題と線形非分離問題に分けて議論する．線形分離問題では，原型ベクトルは修正されるが，原型ベクトルとして利用された学習用の画像ベクトルは必ず正しく分類される．これは，それぞれの原型ベクトルが画像ベクトル空間内の正しい領域からはみ出すような修正が，行われないためである．しかしながら，線形非分離問題では，図3-2(c)に示すように，修正により原型ベクトルが正しい領域からはみ出すことがある．類内しきい値の修正後，類内しきい値が類間しきい値より大きくなった場合に，それぞれの原型ベクトルは正しい領域外の原型ユニットとして検出される．この場合，図3-2(d)に示すように検出されたユニットは取り除かれ，新しいユニットが加えられる．この新しく加えられたユニットが，原型ベクトルが正しい領域から，上記の例と同じように正しい領域から外へ飛び出すのを防ぐため，この過程は繰り返されない．従って，原型ユニットの追加は必ず停止し，学習は収束する．今までの実験において，収束しなかった学習過程はなかった．

3.2.4　RCEネットワークとの棄却能力の比較

図3-3(a)(b)は，それぞれRCEネットワークとRCE-2ネットワーク[87]で，図3-2と同じ学習対で学習した場合の最終状態を示す．図3-3(a)は図2-9(g)と同じ図である．RCE-2ネットワークは，Hudakによって提案され，学習法を改良したことにより

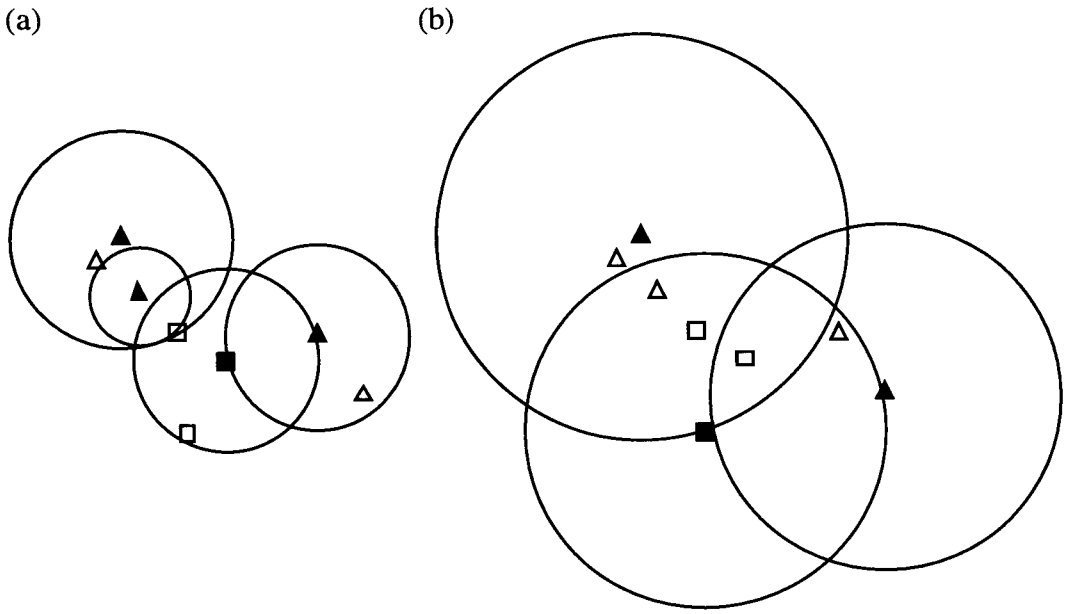

図3-3 学習後の様子；(a)RCEネットワーク，(b)RCE-2ネットワーク

認識率の向上が図られたネットワークである．RCE-2ネットワークは，活性化したユニットのクラスが学習対のクラスと異なる場合でも，ネットワーク全体で認識結果として正しい結果が得られる場合は，しきい値が修正されないように改良されている．図3-3(a)(b)において，円の内側の画像ベクトルは何らかのクラスに分類される．この分類される領域は，学習用の画像ベクトルの存在する領域よりはるかに大きく広がっている．この広がった分類される領域により，学習用の画像ベクトルの存在する範囲より大きくはずれた画像もどれかのクラスに分類され易くなるため，棄却能力は低下すると予想される．

DHネットワークの場合，図3-2(g)を見れば，破線の円によって囲まれる領域は学習用の画像ベクトルの分布する範囲に適して広がっていることがわかる．パラメータ e は，式(3-4)に従い，境界しきい値 $b^{(b)}$ を類内しきい値 $b^{(i)}$ と類間しきい値 $b^{(o)}$ の間で変化させるために使われる．パラメータ e が 0 の場合は，境界しきい値が類間しきい値と同じになるため，棄却の程度は最も低く設定される．パラメータ e が 1 の場合は，境界しきい値が類内しきい値と同じになるため，棄却の程度が最も高く設定される．パラメータ e が0である棄却の程度が最も低い時でさえ，DHネットワークは，図3-2(g)，図3-3(a)(b)の例を見るかぎり，RCEやRCE-2ネットワークに比べて棄却の程度が高い．この時，DHネットワークの分類結果はRCEネッ

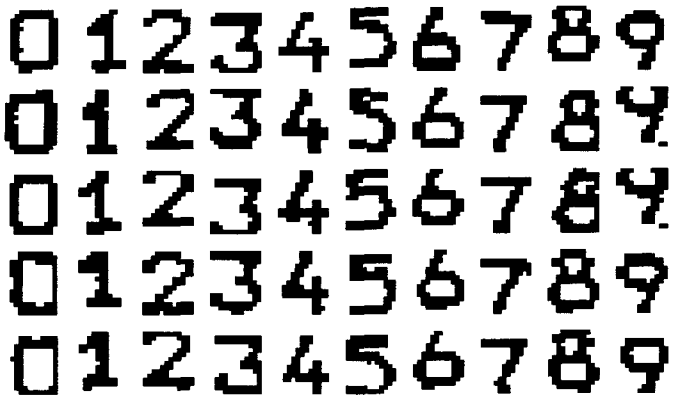

図3-4 X線フィルム内の患者ID番号から切り出された
数字データベースの一部

トワークより信頼できるといえる．

３．３　数字データベースによる実験

３．３．１　実験

X線フィルムの隅に記録されている患者ID番号から切り出された数字データベースを用いる．データベースは，139枚のX線フィルムから切り出された3184個の数字から構成されている．切り出された数字の例を図3-4に示す．各数字の画像は12×14画素の二値画像である．この画像には，記録やディジタル化，切り出しの際に，文字の変形や位置ずれが生じている．X線フィルム内の患者ID番号の自動認識は，大量のX線フィルムをディジタル化する際に不可欠である．認識された患者ID番号は，画像ファイルのヘッダに記憶される．患者ID番号の誤認識は患者の生命にかかわる場合があり，信頼性の高い認識が必要である．

上記のデータベースを用いてRCE, RCE-2, DHネットワークの能力を比較する．全ての偶数文字の内，960個は学習用の画像，952個はテスト用の画像に用いる．棄却の程度を評価するために，1272個のすべての奇数文字を未知画像として用いる．曖昧な画像は，テスト用の偶数文字に含まれているとする．偶数文字の中で誤認識された画像は，ネットワークによって棄却されなかった曖昧な画像と考えられる．

３．３．２　実験結果

表3-1 RCE, RCE-2, DHネットワークの実験結果
（DHネットワークのパラメータ e は 0.0 に設定している．）

	Rate of correct class. against even digits(%)	Rate of incorrect class. against even digits(%)	Rate of rejection against odd digits(%)	No. of prototype cells	No. of training epochs
RCE	99.27±0.28	0.13±0.10	67.43±9.38	39.1±2.5	4.0±0.5
RCE-2	99.51±0.22	0.31±0.19	33.37±13.11	32.6±2.8	3.0±0.6
DH (e=0.0)	99.49±0.17	0.12±0.06	84.57±2.79	33.0±1.7	4.1±1.0

表3-1にRCE, RCE-2, DHネットワークにおける実験結果を示す．DHネットワークのパラメータ e は，最も棄却の程度の低い 0 に設定した．表に，偶数文字の認識率と誤認識率，未知画像としての奇数文字の棄却率，生成された原型ユニットの数，学習にかかったエポック数を示す．平均と標準偏差は初期値の異なる100回の試行により計算されたものである．２.２.４Bより，奇数文字の棄却率が高く，偶数の誤認識率が低いネットワークは，棄却能力の高いネットワークであるということができる．

表3-1より，次のことが分かる．RCE-2ネットワークは，３つのネットワークの中で偶数文字に対する最も高い誤認識率と奇数文字に対する最も低い棄却率を持つため最も棄却能力が低い．RCEネットワークは，RCE-2ネットワークより高い棄却能力を持つ．しかしながら，未知画像である奇数文字について，40%近くを棄却できていない．一方，DHネットワークの奇数文字の棄却率は，RCEネットワークの棄却率に比べて約17%高く，すぐれた棄却能力を持つ．DHネットワークの認識率は，RCE，RCE-2ネットワークの認識率とほとんど同じである．DHネットワークにおいて形成され，最後に残った原型ユニットの数は，RCE, RCE-2ネットワークにおいて形成された原型ユニットの数とほとんど同じである．そのため，DHネットワークの使う記憶容量の大きさは，RCE, RCE-2ネットワークの使う記憶容量の大きさと変わらない．DHネットワークの学習に求められるエポック数は，RCE, RCE-2ネットワークの学習に求められるエポック数と同じくらい少ないため，DHネットワークは，RCE, RCE-2ネットワークと同程度の速さで学習される．実際に，計算機で学習に要した時間は，サンマイクロシステムズ社のSPARCstation 2を用いた場合，RCE, RCE-2, DHはそれぞれ平均3.1，1.9，2.6秒で

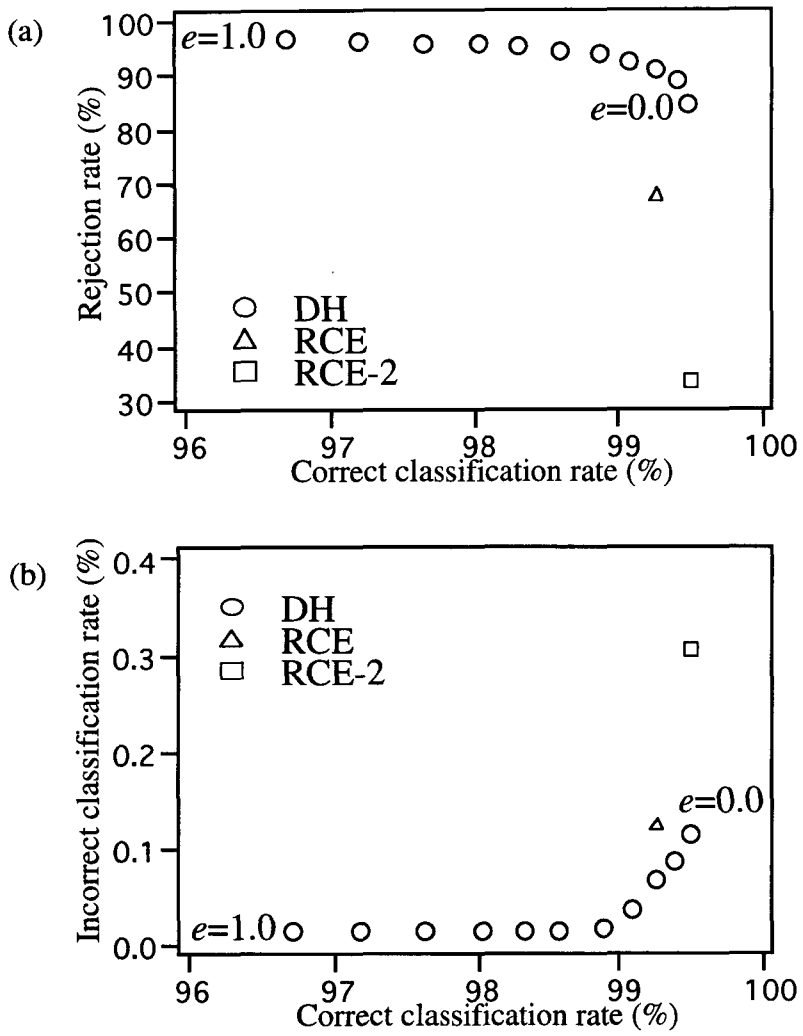

図3-5 パラメータeの変化にともなう棄却能力の変化；(a)偶数文字の認識率と奇数文字の棄却率，(b)偶数文字の認識率と誤認識率

あった．以上より，DHネットワークは，RCE, RCE-2ネットワークの持つ利点を失わずに高い棄却能力を得ることが可能なことが分かる．RCE, RCE-2ネットワークに比べて唯一劣るのは，DHネットワークは，より多くのパラメータを調整する必要があるため，学習手順がより複雑になっていることである．

DHネットワークにおいて，パラメータeによる棄却能力の変化を調べる．図3-5(a)と(b)に，それぞれ偶数文字の認識率と奇数文字の棄却率, 偶数文字の認識率と偶数文字の誤認識率の関係を，パラメータeを変えながらプロットした．それぞれの値は初期値の異なる100回の試行による結果の平均である．パラメータeは0.0から1.0まで0.1刻みで変化させた．表3-1におけるRCE, RCE-2ネットワークの結果もまた図3-5(a)(b)にプロットする．パラメータeを0.0から1.0に変えるに

従って，奇数文字の棄却率が高くなり，偶数文字の誤認識率が下がることが分かる．これにより，ネットワークのより高い棄却能力を求めると，ネットワークの認識率が低下することが分かる．パラメータeが 1.0 の時，棄却率が最も高く96.8%である．

3.4　結言

実用的な画像認識において，分類結果の信頼性は非常に大切である．高い信頼性のためには，未知画像や曖昧な画像は棄却されるべきである．本章において，高い棄却能力を持つ新しいネットワークモデルを提案した．提案したネットワークは第二層のユニットに類間・類内しきい値と呼ばれる二つのしきい値を持つ．棄却能力はネットワークのパラメータを変化させることで容易に調整可能である．簡単な例により，提案したネットワークとRCEネットワークの学習過程を比較し，その棄却能力の高さを示した．実用的なデータによる実験を，X線フィルム内の患者ID番号から切り出された数字データベースを用いて行った．その結果，提案したネットワークは，RCEネットワークの持つ利点を失わずに高い棄却能力を得ることが可能なことが分かった．また，より高い棄却能力をネットワークに与えた場合，認識率が低下することが分かった．

第4章　ニューラルネットワークによる視覚情報の能動的認識モデル

4.1　緒言

ものを見るとき，人間の目は，1秒間に約3回もの頻繁な視点の移動を行う[108]．人間は，得られた空間位置の情報や経験による知識等に基づいた視点の移動により，選択的に外界から視覚情報を獲得していると考えられる．知能ロボットにおいても，目的に応じた視覚情報を視覚センサーを動かして選択的に獲得することが必要である．こうした視点の移動による能動的な視覚システムは，1980年代後半から研究されている[9-11]．

本章において，ニューラルネットワークにより，選択的な視点の移動を行う能動的視覚情報認識モデルである関係統合ニューラルネットワークを提案する[109]．視覚情報として，前章までのように画像ではなくパターン木と呼ばれる情報を用いる．提案するモデルは，視野内のパターンをモデルの記憶するパターンと照合するモジュール，視覚センサの制御を行うモジュール等の各種機能モジュールと，各機能モジュールを接続する関係モジュールで構成されている．関係モジュールは，調和モデル[41]と同じ構造をしているが，調和モデルとは異なり，層間の方向性のある接続により順序の関係する処理を実現している．4.2において，能動的な視覚情報認識を行う関係統合ニューラルネットワークについて述べる．特に，取り扱う情報の表現方法，提案するモデルの構造，認識過程，学習過程について，順を追って説明する．4.3において，計算機実験により，顔画像の認識実験を本モデルを用いて行った結果について述べる．

4.2　関係統合ニューラルネットワーク

顔の認識を例にとり，提案するニューラルネットワークによる視覚情報の能動的認識モデルについて説明する．顔の認識だけを扱う場合，目，鼻等のように要素となるパターンの位置関係がすべての認識対象に対してほぼ同一であると考えることができる．

4.2.1　視点の移動とズーム操作に基づく情報表現

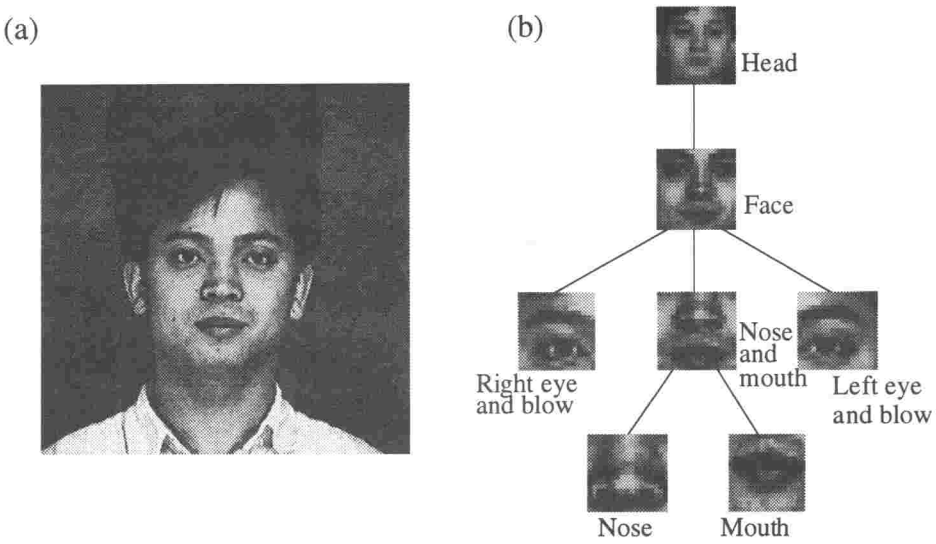

図4–1 顔に関する視覚情報；(a)通常の画像(480×480画素),
(b)パターン木(17×17画素×7)

図4-1(a)に，対象となる人の顔に関する480×480画素の通常の画像を示す．同じ人の顔から，視覚センサの移動とズーム操作を想定して切り出した特徴的な複数の局所パターン(local pattern)を図4-1(b)に示す．図4-1(b)に示す複数の局所パターンは，全体をまとめてパターン木(pattern tree)と呼ぶ[10,110]．パターン木には視点の移動のための位置情報とズーム操作のための倍率の情報が付随しているものとする．パターン木による視覚情報の表現方法は，通常，特徴抽出等により獲得されるが，ここでは認識処理を中心に論議するため，これらが与えられているものとして話を進める．図4-1(b)において，パターン木の上から1段目の局所パターンは，外界の視覚情報から切り出された頭部の画像である．頭部からさらに細かい構造を切り出したのが上から2段目の顔面の画像である．同様に，顔面からは右目・眉，鼻・口，左目・眉の3つの画像が，鼻・口からは鼻，口の2つの画像が切り出される．各局所パターンの画素数は，想定した視覚センサの画素数で決まり，図4-1(b)では17×17画素である．想定した視覚センサの画素数は非常に少ないが，小さな領域において高い解像度が得られる．また，物体をパターン木として記憶すると，システムの記憶容量を大幅に節約することが可能である．

4.2.2　構造

提案するモデルにおいて，パターン木に関する知識をニューラルネットワーク

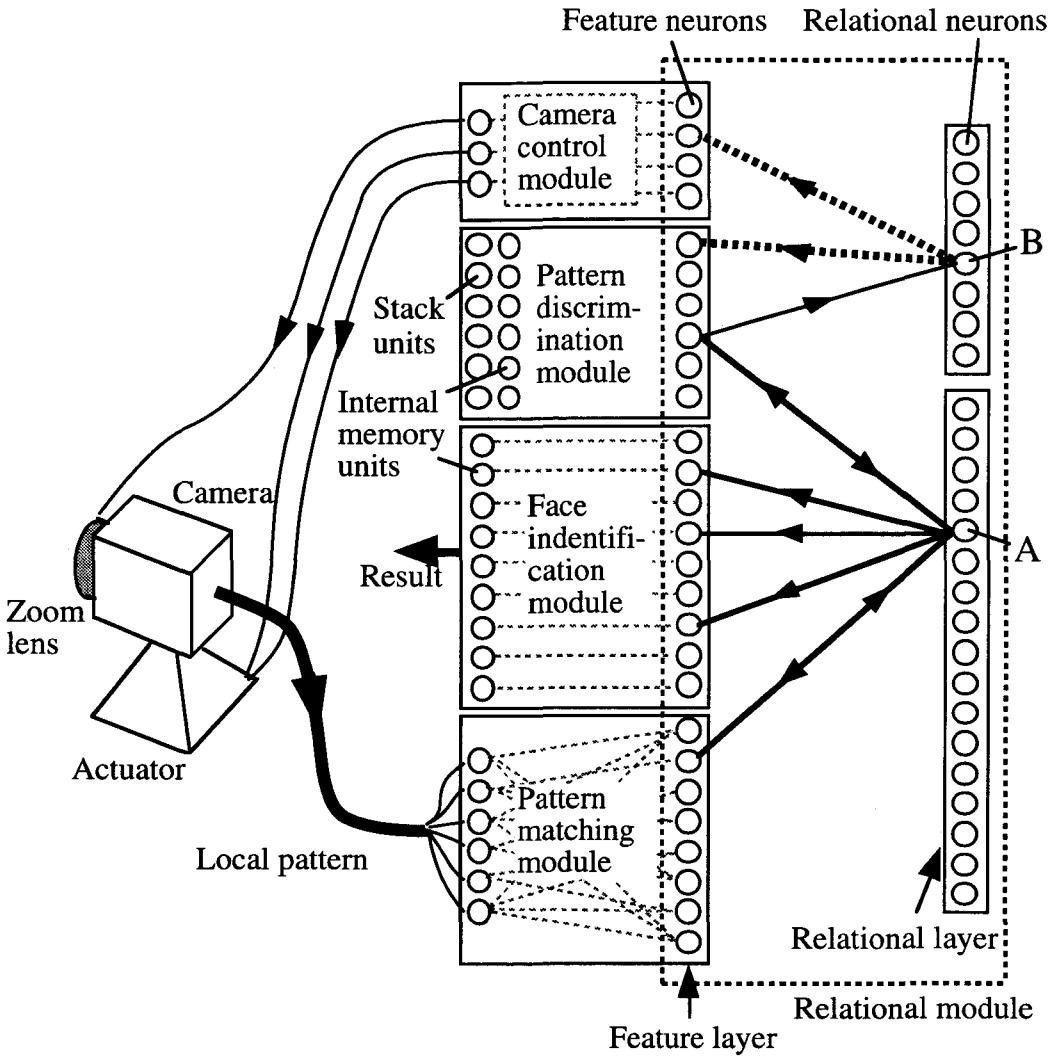

図4-2 ニューラルネットワークによる視覚情報の能動的認識モデル

を構成する処理ユニット間の信号結合強度の分布として記憶し，ユニット間の相互作用により顔の認識を行う．図4-2に，本論文で実現した顔の認識モデルの基本構造を示す．４つの機能モジュール(functional modules)と，それらと相互に作用する関係モジュール(relational module)で構成されている．以下に，各モジュールの構造，役割について述べる．

パターン照合モジュール(pattern-matching module)

入力層，出力層からなる二層のニューラルネットワークである．出力層は関係モジュールの特徴層の一部を成す．本モジュールの入力層，出力層で二層の改良型のRCEネットワークを構成する．本モジュールでは，局所パターンを入力とし，モジュールのもつ複数の参照パターンとの照合結果を出力層を通して関係モジュ

ールに伝える．一つの入力に対して，出力層の一つの特徴ユニットが活性化されるようなネットワークが学習により構成される．

顔同定モジュール(face-identification module)

認識すべき個人の顔一人一人に対応した特徴ユニットと内部メモリユニット(internal memory unit)により構成されている．この特徴ユニットは，関係モジュールから局所パターンの照合結果を受け取る．特徴ユニットの値と内部メモリユニットの値に基づき顔の同定を行う．認識すべき顔に対して，複数の候補があるために顔の同定ができない場合には，候補となる顔の中からどれかの顔を仮説として取り上げ，その結果を関係モジュールに伝える．各内部メモリユニットの値には，処理を始めてから現在までに，現在候補に上がっている顔が仮説として取り上げられた回数が記録されている．詳しい動作については，４．２．３で述べる．

パターン識別モジュール(pattern-discrimination module)

関係モジュールの特徴層の一部を成す一層と学習時に用いられるスタックユニット，内部メモリユニットで構成されている．関係モジュールから局所パターンの照合結果を受け取り，局所パターンが頭部，顔面，右目・眉等のクラスのうち，どのクラスに属するかを識別する．または，関係モジュールから次に視点を移すべき局所パターン（注意視野パターン）に関する情報を受け取り，そのクラスを識別する．各クラスに対し，一つずつの特徴ユニット，スタックユニット(stack unit)，内部メモリユニットが対応するため，７組の特徴ユニット，スタックユニット，内部メモリユニットで構成される．

カメラ制御モジュール(camera-control module)

関係モジュールの特徴層の一部を入力層とし，カメラの位置，ズームレンズの調整のための信号を出力するモジュールである．入力層の特徴ユニットの一つ一つは，特定の視点の移動とズーム操作を指示するユニットである．今回の実験において，実際には視点の移動やズーム操作は行わない．図4-1(b)のパターン木において，局所パターン間を結ぶ直線は6本ある．そのため，パターン木中での視点の移動は全部で12種類必要である．したがって．カメラ制御モジュールは12種の視点移動のための12個の特徴ユニットを備えている．これらのうち一つが活性化した場合，それに固有の位置で得られる局所パターンをパターン照合モジュールに与える．

関係モジュール(relational module)

特徴ユニット(feature unit)の集まりである特徴層(feature layer)と，関係ユニット(relational unit)の集まりである関係層(relational layer)の二層で構成されている．関係モジュールを構成するユニットは，応答値 (response value) r, 推論値 (inferring value) iという2つの独立した内部値を持つ．各値は，顔同定モジュール中の特徴ユニットを除き，すべて0または1の値を持つ．顔同定モジュール中の特徴ユニットの推論値は，負の値を持ち得る．ユニットの応答値または推論値が1の場合，それぞれ，ユニットが応答値，推論値に関して活性化しているという．認識処理において，応答値は実際に存在する入力からの信号を保持し，推論値は推論時に仮説により形成される信号を保持する．応答値は，各層のユニット間を，定められた方向に伝搬する．これに対し推論値はその逆方向に伝搬する．

関係ユニットには2つの種類が存在する．一つは，入力された局所パターンが，どの顔のどのクラスに属するかを記憶する関係ユニット(A)である．図4-2には，この関係ユニットと，顔識別，パターン識別，パターン照合モジュールとが相互結合されている様子が示されている．もう一つは，視線の移動前の局所パターンのクラスから移動後のクラスへのカメラ制御に関する情報を記憶する関係ユニット(B)である．図には視線の移動前のクラスを示すパターン識別モジュール中の特徴ユニットからこの関係ユニットへ，この関係ユニットから移動後のクラスを示す特徴ユニットとカメラ制御モジュールへの結合が形成されている様子が示されている．これらの関係ユニットは，関係層においてそれぞれ独立に信号を受・発信する二つの野と呼ぶユニットの集合を構成している．前者の関係ユニットで構成される野は，注意視野推論野(field for inferring visual field of attention)，後者の関係ユニットで構成される野は視点制御推論野(field for inferring eye movement)と呼ぶことにする．今回，注意視野推論野の関係ユニット(A)は，学習により構成するが，視点制御推論野の関係ユニット(B)は，既に構成されているとする．

4.2.3　認識過程

本節において，まず，すべての学習が完了しているとして，モデル全体の認識動作を説明する．関係モジュールと，内部メモリユニットのみを示した図4-3を用いて説明する．各層の状態は，応答値ベクトル，推論値ベクトルで表わされる．

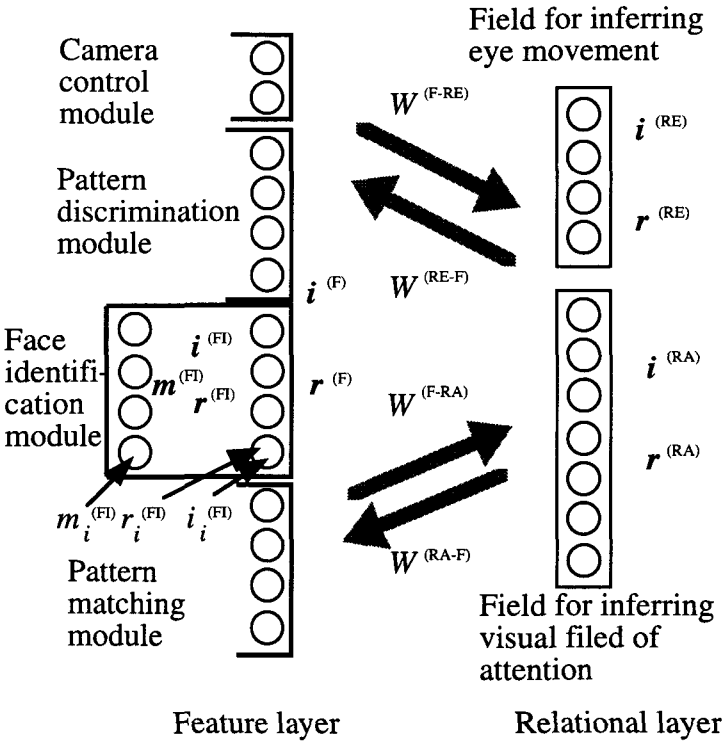

図4-3 関係モジュールと顔識別モジュールの内部メモリユニット

以下に，説明で用いられるパラメータや関数等の定義を示す．

パラメータ，関数等の定義

$\boldsymbol{r}^{(F)},\boldsymbol{i}^{(F)}$　：特徴層の応答値ベクトル，推論値ベクトル．

$\boldsymbol{r}^{(RA)},\boldsymbol{i}^{(RA)}$　：注意視野推論野の応答値ベクトル，推論値ベクトル．

$\boldsymbol{r}^{(RE)},\boldsymbol{i}^{(RE)}$　：視点制御推論野の応答値ベクトル，推論値ベクトル．

$\boldsymbol{r}^{(FI)},\boldsymbol{i}^{(FI)}$　：顔同定モジュールに属する特徴層の応答値ベクトル，推論値ベクトル（$\boldsymbol{r}^{(F)},\boldsymbol{i}^{(F)}$の一部）．

$\boldsymbol{m}^{(FI)}$　：顔同定モジュールの内部メモリユニットの状態ベクトル．

$r^{(FI)}_i,i^{(FI)}_i,m^{(FI)}_i$　：番号 i の顔に対応する，$\boldsymbol{r}^{(FI)},\boldsymbol{i}^{(FI)},\boldsymbol{m}^{(FI)}$の成分．

$W^{(F\text{-}RA)},W^{(F\text{-}RE)}$　：特徴層から注意視野推論野，視点制御推論野への結合行列．

$W^{(RA\text{-}F)},W^{(RE\text{-}F)}$　：注意視野推論野，視点制御推論野から 特徴層への結合行列．

$Step(\boldsymbol{x})$　：$\boldsymbol{x}$の各成分について，0以下の時は0，正のときは1を成分とするベクトルを返すしきい値関数．

$\otimes$　：ベクトルの成分間の積（アダマール積）演算子．

$WTA(\boldsymbol{x})$　：$\boldsymbol{x}$内で活性化した成分がただ一つになるように他の活性化し

た成分を抑制して0としたベクトルを出力する関数.

モデル全体の認識動作は，一連の三つの認識過程により構成されている．以下で，各認識過程について順に説明する．各説明の後に，認識動作の簡単な例を示す．認識の初期状態において，視覚センサは局所パターンのどれかを捉えており，全ての内部メモリユニットは1の値を持つとする．

認識過程 (1)　パターン照合モジュールにおいて，入力された局所パターンに対して，一つの特徴ユニットが応答値に関して活性化される．これを含む応答値ベクトル$\boldsymbol{r}^{(F)}$が，次式に従い関係層へ伝搬し，$\boldsymbol{r}^{(RA)}$が得られる．

$$\boldsymbol{r}^{(RA)} = Step(\, W^{(F\text{-}RA)}\boldsymbol{r}^{(F)}). \qquad (4\text{-}1)$$

この$\boldsymbol{r}^{(RA)}$が，次式に従い再び特徴層の顔同定，パターン識別モジュールに伝搬し，$\boldsymbol{r}^{(F)}$が得られる．

$$\boldsymbol{r}^{(F)} = Step(\, W^{(RA\text{-}F)}\boldsymbol{r}^{(RA)}). \qquad (4\text{-}2)$$

この$\boldsymbol{r}^{(F)}$に含まれる顔同定モジュール内の応答値ベクトル$\boldsymbol{r}^{(FI)}$は，$(\boldsymbol{r}^{(FI)})'$を更新後のベクトルとしたとき，

$$(\boldsymbol{r}^{(FI)})' = Step(\boldsymbol{r}^{(FI)} \otimes \boldsymbol{m}^{(FI)}), \qquad (4\text{-}3)$$

に更新される．同時に，内部メモリベクトル$\boldsymbol{m}^{(FI)}$は，$\boldsymbol{m}^{(FI)'}$を更新後のベクトルとしたとき，

$$\boldsymbol{m}^{(FI)'} = \boldsymbol{r}^{(FI)} \otimes \boldsymbol{m}^{(FI)}, \qquad (4\text{-}4)$$

に更新される．顔同定モジュール内で，修正後の応答値ベクトルのノルム $\|(\boldsymbol{r}^{(FI)})'\|$ の値に従い，以下の顔の同定処理が行われ，この認識過程を終える．

顔の同定処理

(a) $\|(\boldsymbol{r}^{(FI)})'\| = 0$ の場合，対象とする顔は未知の顔であると判断される．その結果，全ての認識過程を終了する．

(b) $\|(\boldsymbol{r}^{(FI)})'\| = 1$ の場合，対象は，唯一つ活性化したユニットの示す顔に識別される．その結果，全ての認識過程は完了する．

(c) $\|(\boldsymbol{r}^{(FI)})'\| > 1$ の場合，現在まで得られた対象に関する情報だけでは，顔を同定できない．さらに次の認識過程(2)に進む．

図4-4に上記の認識過程の簡単な例を示す．図4-4(a)に，A, B 2名の顔のパターン木の例を示す．説明の簡単化のため，図4-1(b)の顔面，右目・眉，鼻・口，口の4つの局所パターンのみを考え，モデルに，A, B 2名のパターン木のみを記憶させるものとする．図4-4(a)の局所パターンの左側に，その局所パターンの入力に対して活性化するパターン照合モジュール内の特徴ユニットの番号(1, 2, 3, 4, 5)を記した．局所パターンの間には，それらの間の視点移動を行うために選ばれるべき特徴ユニットの番号(p, q, r)を記した．顔A, 顔Bでは，口の局所パターンのみが異なる(4, 5)．図4-4(b)に顔A, Bに関する結合行列$W^{(F\text{-}RA)}, W^{(RA\text{-}F)}$を記憶したモデルの一部を示す．視点制御に関する結合行列$W^{(F\text{-}RE)}, W^{(RE\text{-}F)}$もあらかじめ記憶させている．図4-4(b)において，ユニット間の接続を表す線の種類は，接続の方向を区別する．パターン識別モジュールの，顔面，右目・眉，鼻・口，口の4つのクラスに対応する4つの特徴ユニットを，それぞれ f, eb, nm, m とする．まず初めに，顔Aを認識の対象とし，最初に顔面のパターンを視野が捉えている場合を考える．ユニットの上半分の部分に応答値，下半分に推論値を記す．図4-4(c)のように，上で述べた認識過程(1)に従い，顔同定モジュールのA, Bの特徴ユニットが応答値に関して活性化される．顔同定モジュールにおいて複数の特徴ユニットが活性化している($\|(\boldsymbol{r}^{(FI)})'\| > 1$)ため，次の認識過程に進む．

認識過程 (2)　最大の値をもつ内部メモリユニットの中の一つの番号をjとする．この時，対応する番号jの顔同定特徴ユニットの推論値$i^{(FI)}_j$に1を与えることにより，対象とする顔がj番目の特徴ユニットが示す顔であるという仮説を設ける．同時に，内部メモリユニットjの値$m^{(FI)}_j$を1だけ大きくする．仮説からはずれた応答値が正の特徴ユニットの推論値には，負の値$-n$を与える．この時，nは顔同定モジュール内の特徴ユニット数より大きい値とする．応答値が0の特徴ユニットの推論値には，0を与える．これらを含む推論値ベクトル$\boldsymbol{i}^{(F)}$が，次式に従い関係層に伝搬し，$\boldsymbol{i}^{(RA)}$が得られる．

$$\boldsymbol{i}^{(RA)} = Step(W^{(RA\text{-}F)} \boldsymbol{i}^{(F)}). \qquad (4\text{-}5)$$

注意視野推論野の推論値ベクトルのノルム$\|\boldsymbol{i}^{(RA)}\|$の値に従い，以下の視点を移すべき注意視野パターンの検出処理が行われる．

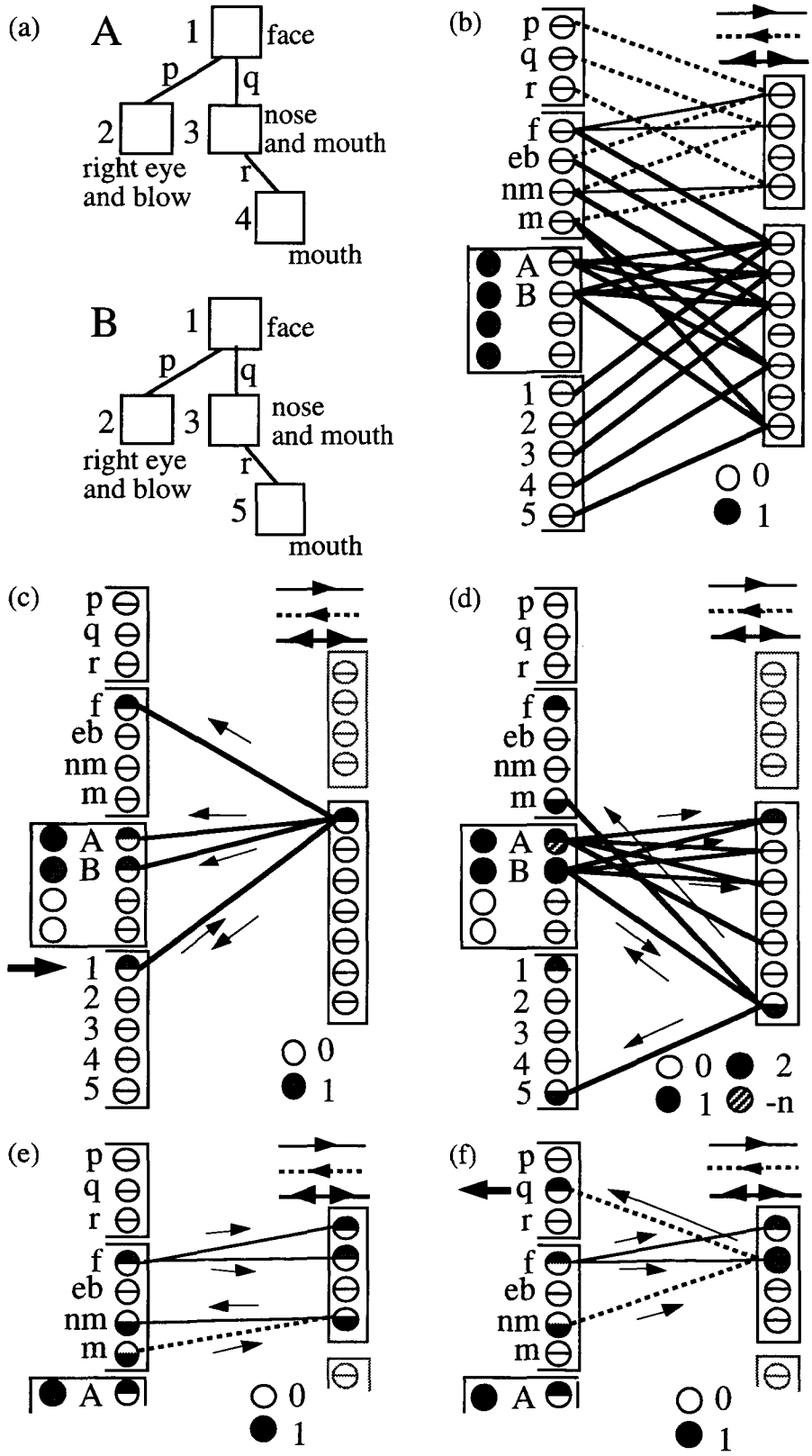

図4-4 顔の認識過程の例；(a)パターン木，(b)記憶された結合，(c)顔の同定処理，(d)注意視野の検出，(e)視点移動の決定過程，(f)視点移動の決定

注意視野パターンの検出処理

(a) $\|\boldsymbol{i}^{(RA)}\| > 0$ の場合，注意パターンが検出され，$\boldsymbol{i}^{(RA)}$が次式に従い，パターン識別，パターン照合モジュールに伝搬し，$\boldsymbol{i}^{(F)}$ が得られる．

$$\boldsymbol{i}^{(F)} = Step(\, W^{(F\text{-}RA)}\boldsymbol{i}^{(RA)}\,). \qquad (4\text{-}6)$$

$\boldsymbol{i}^{(F)}$ が得られた後，この認識過程を終了し，次の認識過程(3)に進む．

(b) $\|\boldsymbol{i}^{(RA)}\| = 0$ の場合，注意視野パターンが検出されるように仮説の拡張が行われる．現在，推論値が $-n$で，内部メモリユニットの値が最も大きい特徴ユニットjの推論値$i^{(FI)}_j$に 1 を与え，再び，式(4-5)に基づく信号の伝搬，注意パターンの検出処理を行う．この時，m_jの値を 1 だけ大きくする．

図4-4(d)において，本認識過程に従い，ユニットAの推論値に$-n$，ユニットBの推論値に 1 を与えている様子が示されている．また，注意視野パターンの検出処理により，次に視野に捉えるべきクラスの特徴ユニット m と，その局所パターンに対応する特徴ユニット 5 の推論値が活性化されている．この例では，対象とする顔の候補が二つしかないので，一度の仮説により，注意パターンが検出されている．しかし，実際には 4.3 の実験で示すように，一つの局所パターンに対して顔の候補は複数あり，仮説の拡張は頻繁に行われる．

認識過程(3)　特徴層の応答値，推論値ベクトル$\boldsymbol{r}^{(F)}$,$\boldsymbol{i}^{(F)}$が，それぞれ次の(4-7),(4-8)式に従い伝搬し，$\boldsymbol{r}^{(RE)}$，$\boldsymbol{i}^{(RE)}$が得られる．

$$\boldsymbol{r}^{(RE)} = Step(\, W^{(F\text{-}RE)}\boldsymbol{r}^{(F)}\,), \qquad (4\text{-}7)$$

$$\boldsymbol{i}^{(RE)} = Step(\, W^{(RE\text{-}F)}\boldsymbol{i}^{(F)}\,). \qquad (4\text{-}8)$$

この応答値ベクトル$\boldsymbol{r}^{(RE)}$は，$(\boldsymbol{r}^{(RE)})'$を修正後のベクトルとしたとき，

$$(\boldsymbol{r}^{(RE)})' = \boldsymbol{r}^{(RE)} \otimes \boldsymbol{i}^{(RE)}, \qquad (4\text{-}9)$$

に修正される．修正後の応答値ベクトルのノルム$\|(\boldsymbol{r}^{(RE)})'\|$の値に従い，以下のように注意視野パターンを捉えるために最初に行う視点の動作に関する決定処理が行われる．

視点動作の決定処理

(a) $\|(\boldsymbol{r}^{(RE)})'\| > 0$ の場合，$WTA(\,(\boldsymbol{r}^{(RE)})'\,)$が，次の(4-10)式に従い特徴層のカメラ制

御モジュールに伝搬し，$r^{(F)}$が得られる．

$$r^{(F)} = Step(W^{(RE-F)} WTA((r^{(RE)})')). \quad (4\text{-}10)$$

$r^{(F)}$に従いカメラの制御が行われ，次の局所パターンを捉える．認識過程(1)からの認識処理を続ける．

(b) $\|(r^{(RE)})'\| = 0$ の場合，視点動作が決定されず，推論値ベクトル$i^{(RE)}$が，次式に従い特徴層に伝搬し，$i^{(F)}$が得られる．

$$i^{(F)} = Step(W^{(F-RE)} i^{(RE)}). \quad (4\text{-}11)$$

これをもとに，再びこの認識過程(3)を繰り返す．

図4-4(e)の状態において，視点動作の決定処理の際，視点制御推論野において応答値，推論値両方に関して活性化している関係ユニットが存在しない．したがって，(4-9)式の左辺の $(r^{(RE)})'$ のノルム $\|(r^{(RE)})'\|$ は0となり，上記(b)のとおりに，(4-11)式による推論値ベクトルの伝搬が行われている．図4-4(f)において，2度目の視点動作の決定処理の結果，応答値，推論値両方に関して活性化している関係ユニットが存在する($\|(r^{(RE)})'\| > 0$)．そのため，(4-10)式による応答値ベクトルの伝搬が行われ，カメラ制御モジュールの特徴ユニットqが発火し，選択された注視視野パターンへ向かう視点の移動が行われる．

この例ではその後，視点の移動により得られた局所パターンをもとに，上記と同じ認識過程(1)(2)(3)を経て，口の部分への視点の移動を行う．その結果，顔同定モジュールにおいてAに関するユニットのみが活性化し，対象はAと認識される．

4.2.4　学習過程

本節において，パターン照合モジュールの学習法と，注意視野推論野の関係ユニットとそれに関する結合荷重の学習法について述べる．視点制御モジュールとパターン識別モジュールを接続する視点制御推論野の関係ユニットは既に形成されているとする．これは，顔に関する局所パターンの位置関係が既知であることを示す．学習時には，与えられた顔画像に関するすべての局所パターンを順に視野に捉えることを目的とした視点の移動を行う．視点の移動により得られた局所パターンを受けて，パターン照合モジュールの学習と注意視野推論野の関係ユニットに関する学習が行われる．パターン照合モジュールでは，従来のRCEネットワ

ークと異なり，第3章で提案した二重超球ネットワークと同じように第二層のユニットが二つのしきい値を持つ．これにより，パターン照合モジュールの学習と注意視野推論野の関係ユニットに関する学習を同時に行うことが出来る．以下に，学習のための視点移動過程，パターン照合モジュールの学習，関係ユニットに関する学習について順に説明する．各説明のあとに簡単な例を示す．

A.　学習のための視点の移動

以下に，4.2.3に示したパラメータ，関数の定義に加えて本節で用いられるパラメータや関数等の定義を示す．

パラメータ，関数等の定義

$\boldsymbol{r}^{(FD)},\boldsymbol{i}^{(FD)}$　：パターン識別モジュールに属する特徴層の応答値ベクトル，推論値ベクトル（$\boldsymbol{r}^{(F)},\boldsymbol{i}^{(F)}$の一部）．

$\boldsymbol{m}^{(FD)},\boldsymbol{s}^{(FD)}$　：パターン識別モジュール内の内部メモリユニットの状態ベクトルとスタックユニットの状態ベクトル．

Θ　：二つのベクトルの対応する成分間の抑止演算を行う演算子．（抑止演算とは，演算子の前の値が1で後ろの値が0の場合のみ演算結果に1を与える演算を意味する．）

学習のための視点の移動過程　教師信号により，パターン識別モジュールの特徴層において現在視野が捉えている局所パターンに対応する特徴ユニットのみが応答値に関して活性化される．同時に，対応する内部メモリユニット，スタックユニットが活性化される．また，その時，内部メモリユニットの値が，対応する特徴ユニットの推論値に与えられる．

$$\boldsymbol{i}^{(FD)}=\boldsymbol{m}^{(FD)}, \tag{4-12}$$

応答値，推論値ベクトルが，それぞれ式(4-7),(4-8)に従い伝搬し，$\boldsymbol{r}^{(RE)}$，$\boldsymbol{i}^{(RE)}$が得られる．この応答値ベクトル$\boldsymbol{r}^{(RE)}$は，$(\boldsymbol{r}^{(RE)})'$を修正後のベクトルとした時，

$$(\boldsymbol{r}^{(RE)})'=\boldsymbol{r}^{(RE)}\Theta\boldsymbol{i}^{(RE)}, \tag{4-13}$$

に修正される．修正後の応答値ベクトルのノルムの値 $\|(\boldsymbol{r}^{(RE)})'\|$ に従い，以下のように視点を移動する．

学習のための視点移動の決定処理

(a) $\|(\boldsymbol{r}^{(RE)})'\| > 0$ の場合，$WTA((\boldsymbol{r}^{(RE)})')$が，(4-10)式に従い特徴層のカメラ制御モジュールに伝搬し，$\boldsymbol{r}^{(F)}$が得られる．$\boldsymbol{r}^{(F)}$に従いカメラの制御が行われ，次の局所パターンを捉える．この時，顔面から目・眉へというように下位層への視点の移動が行われる．

(b) $\|(\boldsymbol{r}^{(RE)})'\| = 0$ の場合，次に示す手順で目・眉から顔面へというように上位層への視点の移動が行われる（視点移動の後戻りと呼ぶ）．

視点移動の後戻り過程　はじめに，スタックユニットの値を対応する特徴ユニットの推論値に与える．

$$\boldsymbol{i}^{(FD)} = \boldsymbol{s}^{(FD)}. \qquad (4\text{-}14)$$

応答値，推論値ベクトルが，それぞれ式(4-7),(4-8)に従い伝搬し，$\boldsymbol{r}^{(RE)}$，$\boldsymbol{i}^{(RE)}$が得られる．この応答値ベクトル$\boldsymbol{r}^{(RE)}$は，$(\boldsymbol{r}^{(RE)})'$を修正後のベクトルとした時，式(4-9)に従い修正される．また，推論値ベクトル$\boldsymbol{i}^{(RE)}$は，$(\boldsymbol{i}^{(RE)})'$を修正後のベクトルとした時，$(\boldsymbol{r}^{(RE)})'$と同じ値に設定される．修正後の応答値ベクトルのノルム$\|(\boldsymbol{r}^{(RE)})'\|$に従い，以下のように視点移動の後戻りが行われる．

視点移動の後戻り決定処理

(a) $\|(\boldsymbol{r}^{(RE)})'\| > 0$ の場合，(4-10)式に従い，$(\boldsymbol{r}^{(RE)})'$が特徴層のカメラ制御モジュールに伝搬し，$\boldsymbol{r}^{(F)}$が得られる．得られた$\boldsymbol{r}^{(F)}$に従いカメラの制御が行われ，視点移動の後戻りが完了する．同時に，スタックベクトル$\boldsymbol{s}^{(FD)}$は，$(\boldsymbol{s}^{(FD)})'$を修正後のスタックベクトルとした時，次式に従い修正される．

$$(\boldsymbol{s}^{(FD)})' = \boldsymbol{s}^{(FD)} - Step(W^{(F\text{-}RE)}(\boldsymbol{i}^{(RE)})'). \qquad (4\text{-}15)$$

(b) $\|(\boldsymbol{r}^{(RE)})'\| = 0$ の場合，全ての局所パターンに視点を移動したことになり視点の移動を終了する．

以上の視点移動過程や視点移動の後戻り過程を繰り返すことにより，対象とする顔内の全ての局所パターンを捉えることが出来る．

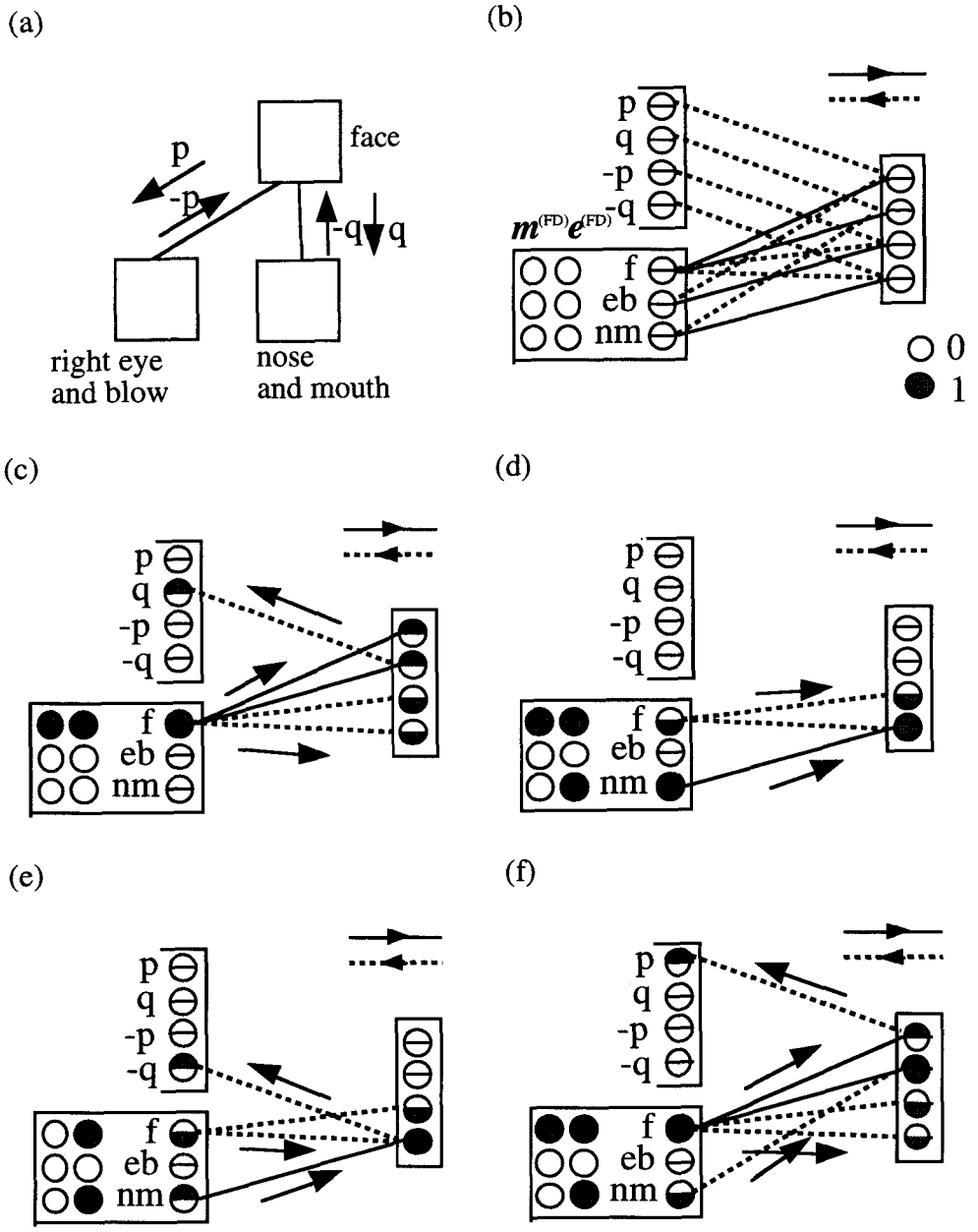

図4-5　学習のための視点移動過程の例；(a)パターン木，(b)記憶された結合，
(c)最初の視点移動の検出，(d)視点移動が検出されなかった場合，
(e)視点の後戻りの検出，(f)視点の後戻り後の視点移動の検出

図4-5に図4-4と同じ要領で学習のための視点移動過程の簡単な例を示す．説明の簡単化のため図4-5(a)に示すように顔のパターン木の一部を用いる．図4-4と異なり，下位の層から上位の層への視点の移動を考慮した．図4-5(b)に図4-5(a)のパターン木の位置関係を記憶したモデルの一部を示す．図4-5(c)のように，最初，顔面の局所パターンが入力されるとする．この時，鼻・口への視点の移動が検出

されている．図4-5(d)において，視点移動の結果により鼻・口の局所パターンが入力されたが，それより下位の層がないため次の下位層への視点移動が検出されない状態を示す．図4-5(e)において，視点の後戻りが検出されている．図4-5(f)において，視点の後戻り後に，再び視点移動の検出が行われている．この時，図4-5(c)において検出された鼻・口への視点移動は抑制される．

B. パターン照合モジュールの学習方法

改良型のRCEネットワークで構成されるパターン照合モジュールの構造を図4-6に示す．改良型のRCEネットワークの第二層目の特徴ユニットは，パラメータとして，結合ベクトル $\boldsymbol{w}_i$，内部値 u_i，類間しきい値 $b^{(o)}{}_i$，類内しきい値 $b^{(i)}{}_i$，ユニット値 s_i，応答値 r_i と局所パターンが何で誰の顔であるかを示す番号 $c^{(what)}{}_i$，$c^{(who)}{}_i$ を持つ．推論値は用いない．本章で用いる改良型のRCEネットワークは，第3章で提案した二重超球ネットワークと同じく，各ユニットは二つのしきい値を持つ．しかし，二重超球ネットワークとは異なり，以下に示すように学習により結合ベクトル（第3章では原型ベクトルと呼んでいる）の修正は行われない．また，本章で用いる改良型のRCEネットワークの各ユニットはクラスとして二つの変数を持つ．

改良型のRCEネットワークの学習法について述べる．順次得られる局所パターンとそれが何で誰の顔であるかを示す教師信号の組の集合を $(\boldsymbol{v}^{(in)}{}_j, t^{(what)}{}_j, t^{(who)}{}_j)$ $j = 1\sim N$ とする．Nは，学習に用いられる局所パターンの数を表す．初期状態にお

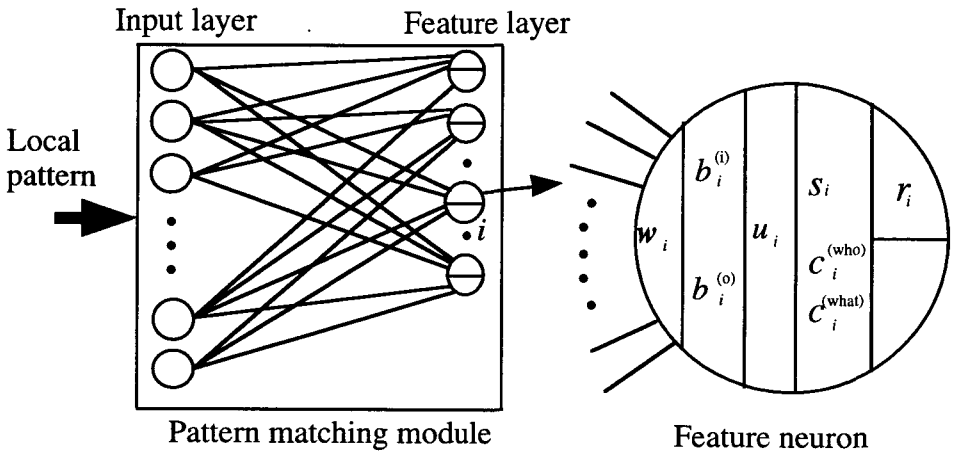

図4-6 パターン照合モジュールの構造

いて，特徴ユニットは一つも存在しない．学習により，特徴ユニットが加えられ，特徴ユニットのしきい値が修正される．

j番目の局所パターンがネットワークに入力されたときのi番目の特徴ユニットのパラメータの修正を考える．特徴ユニットiにおいて以下の演算が行われる．

$$u_i = \| \boldsymbol{w}_i - \boldsymbol{v}^{(\mathrm{in})}{}_j \|, \tag{4-16}$$

$$s_i = f(b^{(\mathrm{o})}{}_i - u_i). \tag{4-17}$$

$f(\cdot)$は，ヘービサイド関数であるとする．入力された局所パターンの教師信号$t^{(\mathrm{what})}{}_j$, $t^{(\mathrm{who})}{}_j$と特徴ユニットiの変数$c^{(\mathrm{what})}{}_i$, $c^{(\mathrm{who})}{}_i$と内部値u_iに応じて，各特徴ユニットの状態を以下の3つの場合に分ける．

（状態1）$(c^{(\mathrm{what})}{}_i \neq t^{(\mathrm{what})}{}_j)$または$(c^{(\mathrm{who})}{}_i \neq t^{(\mathrm{who})}{}_j$ かつ $u_i > b(c^{(\mathrm{what})}{}_i, t^{(\mathrm{what})}{}_j))$.

（状態2）$(c^{(\mathrm{what})}{}_i = t^{(\mathrm{what})}{}_j)$ かつ$(c^{(\mathrm{who})}{}_i = t^{(\mathrm{who})}{}_j$ または $u_i \leq b(c^{(\mathrm{what})}{}_i, t^{(\mathrm{what})}{}_j))$ かつ (u_iがネットワーク内で最小の値).

（状態3）$(c^{(\mathrm{what})}{}_i = t^{(\mathrm{what})}{}_j)$ かつ$(c^{(\mathrm{who})}{}_i = t^{(\mathrm{who})}{}_j$ または $u_i \leq b(c^{(\mathrm{what})}{}_i, t^{(\mathrm{what})}{}_j))$ かつ (u_iがネットワーク内で最小の値ではない).

この時，$b(c^{(\mathrm{what})}{}_i, t^{(\mathrm{what})}{}_j)$は，$c^{(\mathrm{what})}{}_i$と$t^{(\mathrm{what})}{}_j$に応じて決まる類間しきい値の下限値である．この下限値は，従来のRCEネットワークや第3章で提案した二重超球ネットワークでは考えていない．本章で提案する顔の認識モデルにおいて，特徴ユニットを異なる個人の間で共有するために，類間しきい値の下限値は設けられた．

さらに，ユニット値s_iの状態も加えて次の四つの場合に分けて特徴ユニットiのパラメータの修正が行われる．

特徴ユニットiのパラメータ修正

(a)s_i=0 の場合.

修正は全く行われない.

(b)s_i=1 で状態1の場合.

類間しきい値$b^{(\mathrm{o})}{}_i$をu_iに修正する.

(c)s_i=1 で状態2の場合.

類内しきい値$b^{(\mathrm{i})}{}_i$は$b^{(\mathrm{i})}{}_i$, u_iの大きい方の値に修正される．その時，更新さ

れた類内しきい値が類間しきい値より大きくなった場合，類内しきい値を0にリセットする．また，この時，関係ユニットと顔同定モジュール間の結合の修正を行う．その方法は，4.2.4 C(a)において説明する．

(d)s_i=1で状態3の場合．

修正は全く行われない．しかし，この時，関係ユニットと顔同定モジュール間の結合の修正を行う．その方法は，4.2.4 C(b)において説明する．

類間しきい値と類内しきい値の二つのしきい値を用いるのは，(c)において注意視野推論野の関係ユニットに関する結合荷重の修正のための信号を得るためである．

上記の修正後再び式(4-17)の演算のみを行った結果，全ユニットの中でユニット値が1であるユニットが全く存在しない場合，新しい特徴ユニットがネットワークに加えられる．新しく加えるユニットnの結合ベクトルは入力された局所パターンと同じベクトル$\boldsymbol{v}^{(i)}_j$に，類間しきい値$b^{(o)}_n$は，状態1の特徴ユニット中で最小の内部値を持つユニットの持つ類間しきい値に，類内しきい値は0に，各クラスは入力された局所パターンの各クラス$t^{(what)}_j, t^{(who)}_j$に設定される．特徴ユニットが全く存在しない初期状態の場合，類間しきい値は表現可能な出来るだけ大きな値に設定される．

全ての顔の局所パターンがネットワークに与えられるまでの間を1エポックと呼ぶ．学習は，1エポックの間，パラメータが全く修正されなくなるまで行われる．

認識時に，局所パターンがネットワークに与えられたとき，各特徴ユニットは式(4-17),(4-18)にもとづく信号伝達処理の結果，状態2の特徴ユニットが応答値に関して活性化する．その信号は，関係ユニットに伝えられる．

改良型RCEネットワークの具体的な学習例を示す．図4-7(a)-(j)に入力が二次元ベクトルである場合のネットワークの学習の進行例を示す．白い四角（□）と白い三角（△）はそれぞれ鼻と口に属する学習用の入力ベクトルとする．四角と三角の左側に記した番号はそれぞれの属する顔の番号を示すものとする．黒い四角（■）と黒い三角（▲）は，それぞれラベル付けされた第二層の原型ベクトルを示すものとする．第二層の特徴ユニットのにつのしきい値によって，二次元ベクトル空間で二重の円が形成される．類間しきい値で形成された円は実線で，類内

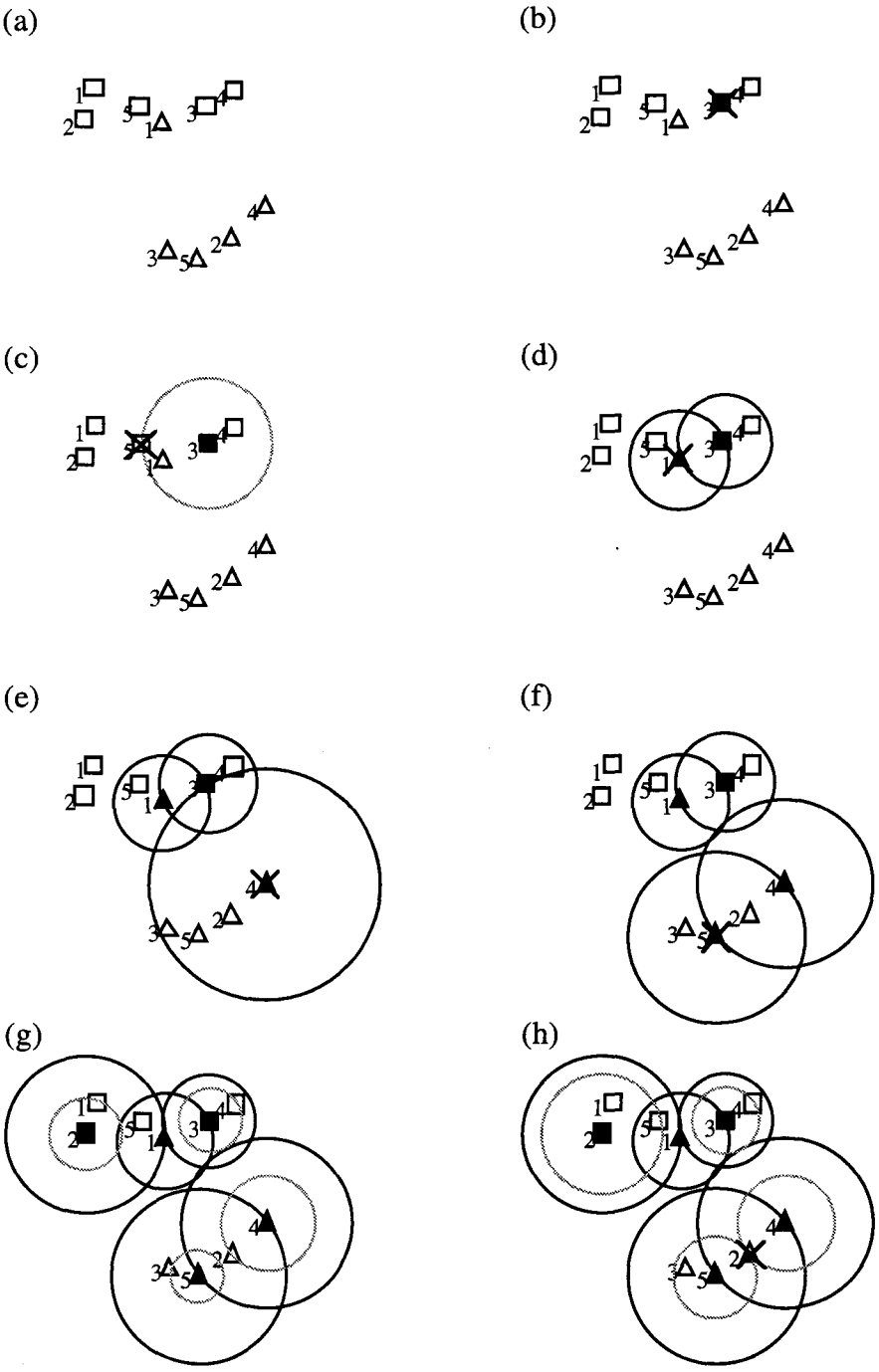

図4-7 入力が2次元ベクトルである場合のパターン照合モジュールにおける学習の進行例

しきい値で形成された円は破線で描く．各状態において，図中の×印はネットワークにその時に提示されている学習用の入力ベクトルを示す．類内しきい値の下限値は適当に設定した．

図4-7(a)は，ネットワークの特徴ユニットが全く存在しない初期状態を示す．最初の特徴ユニット(■3)が，図4-7(b)において加えられている．この時，類間しきい値で形成される円は半径が大きすぎるため，図には描かれていない．また，類内しきい値も0であるため，図には描かれていない．図4-7(c)には，入力ベクトルに対して，内部値が類間しきい値の下限値を下回っている状態2で，ユニット値が1の特徴ユニット(■3)が存在する．この時，その特徴ユニット(■3)の類内しきい値が4.2.4 B(c)に従い修正されている．図4-7(d)において，特徴ユニット(■3)が，その類間しきい値の修正により類間しきい値が類内しきい値より小さくなったため，その類内しきい値がリセットされている．また，新しい特徴ユニット(▲1)が加えられている．図4-7(e)において，新しいユニット(▲4)が追加されている．図4-7(f)において，適当に設定した類内しきい値の下限値より，内部値が大きな値を得ている状態1で，ユニット値が1の特徴ユニット(▲4)が存在する．この時，その特徴ユニット(▲4)の類間しきい値が4.2.4 B(b)に従い修正されるとともに，新しい特徴ユニット(▲5)が追加されている．図4-7(g)は，1回目のエポックが終了した状態を示す．図4-7(h)において，内部値が適当に設定した類内しきい値の下限値より小さな値を得ている状態2で，ユニット値が1の特徴ユニット(▲5)が存在する．この時，この特徴ユニット(▲5)の類内しきい値が4.2.4 B(c)に従い修正されている．図4-7(h)は，また，ネットワークの最終状態と同じである．

C. 関係ユニットに関する学習方法

図4-8に示すように，特徴層のユニット j と注意視野推論野の関係ユニット i 間の相互結合の結合荷重の値を $w^{(F)(RA)}_{ij}$ ($=w^{(F\text{-}RA)}_{ij}=w^{(RA\text{-}F)}_{ij}$)として，注意視野推論野の関係ユニットと特徴層を構成する各モジュールの間の結合荷重の学習方法を以下に説明する．教師信号は，それぞれのモジュールで，現在視野が捉えている局所パターンに対応するユニットを指すものとする．

パターン照合モジュールにおいて特徴ユニット k が生成された時，同時に関係ユニット k が一つ生成される．その際，この関係ユニットは，次の三式に従い生成されたパターン照合モジュールの特徴ユニット k，顔識別モジュールにおいて教師信号が指す特徴ユニット l，パターン識別モジュールにおいて教師信号が示

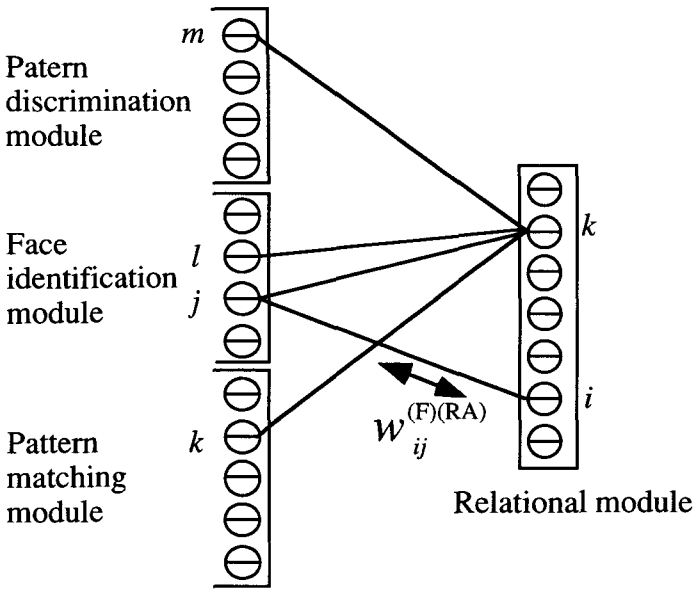

図4-8 特徴層のユニットと関係層のユニットの結合

す特徴ユニット m と相互接続される．

$$w^{(F)(RA)}{}_{kk} = 1, \tag{4-18}$$

$$w^{(F)(RA)}{}_{kl} = 1, \tag{4-19}$$

$$w^{(F)(RA)}{}_{km} = 1. \tag{4-20}$$

上記の結合の生成の他に，関係ユニット k と顔識別モジュールの特徴ユニット間の接続は，パターン照合モジュール内での学習に従って以下の二つの場合に修正される．

関係ユニット k と顔識別モジュール間の接続の修正

(a) 4．2．4 B(c)の条件を満たす特徴ユニット k が存在する場合．

その特徴ユニット k に接続する関係ユニット k と教師信号が示す顔識別モジュールの特徴ユニット l の間に接続が存在しない場合，新たに相互の接続を式(4-19)に従って形成する．しかし，この時，パターン照合モジュールにおいて，類間しきい値が類内しきい値より小さくなり類内しきい値のリセットが行われた場合，その関係ユニット k と顔識別モジュール間のすべてのユニット a (a=1,2,3...N)との間の接続を次式に従って切断する．

$$w^{(F)(RA)}{}_{ka} = 0;\ (a=1,2,3...N). \tag{4-21}$$

Nは，顔識別モジュール内の特徴ユニットの総数とする．

(b) 4.2.4 B(d)の条件を満たす特徴ユニット k が存在する場合.

その特徴ユニット k に対応する関係ユニット k と顔識別モジュール l の教師信号が示す特徴ユニットとの接続がある場合は，次式に従ってその接続を切断する.

$$w^{(F)(RA)}{}_{kl} = 0. \qquad (4\text{-}22)$$

図4-9(a)-(f)に，図4-8と同じ学習用の入力パターンの例を用いて，注意視野推論野の関係ユニットに関する学習過程の例を示す．各図の左側には，上からパターン識別，顔同定，パターン照合モジュールを，右側には注意視野推論野を表す．パターン識別，顔同定モジュールの特徴ユニットの左側には，それぞれのユニットに対応する図4-8の例で用いた番号を記す．パターン照合モジュールの番号は，原型ベクトルに用いられた学習用の入力ベクトルの番号を記す．図4-8(b)に示したパターン照合モジュールの学習と同時に行われる関係ユニットに関する学習を，図4-9(a)に示す．パターン照合モジュールにおいて特徴ユニットが生成されると同時に，関係ユニットが生成され，顔同定モジュール，パターン識別モジュールと生成された関係ユニットとの間の結合が形成される．図4-9(b)には，図4-8(c)に示した学習と同時に行われる4.2.4 C(a)に示す修正に従った，顔同定モジュールの5番目のユニットと以前に生成された関係ユニットとの間の結合の生成の様子が示されている．図4-9(c)には，図4-8(d)に示した類内しきい値のリセットと同時に行われる4.2.4 C(a)に示す修正に従った，顔同定モジュールと関係ユニット間の結合の切断の様子が示されている．また，パターン照合モジュールにおける新しい特徴ユニットの追加とともに，新しい関係ユニットとそれに関する結合が生成されている．図4-9(d)には，図4-8(g)と同じ1エポック終了後のネットワークの接続の様子を示す．図4-9(e)には，図4-8(h)に示したパターン照合モジュールの学習と同時に行われる4.2.4 C(b)に従った，顔同定モジュールと関係ユニット間の一つの結合の切断の様子が示されている．また，同時に4.2.4 C(a)に従って顔同定モジュールと関係ユニット間の結合が一つ形成されている．図4-9(e)において，結合の切断と生成の様子を分かりやすく示すために他の結合の線を細く描いた．図4-9(f)にネットワークの最終状態を示す.

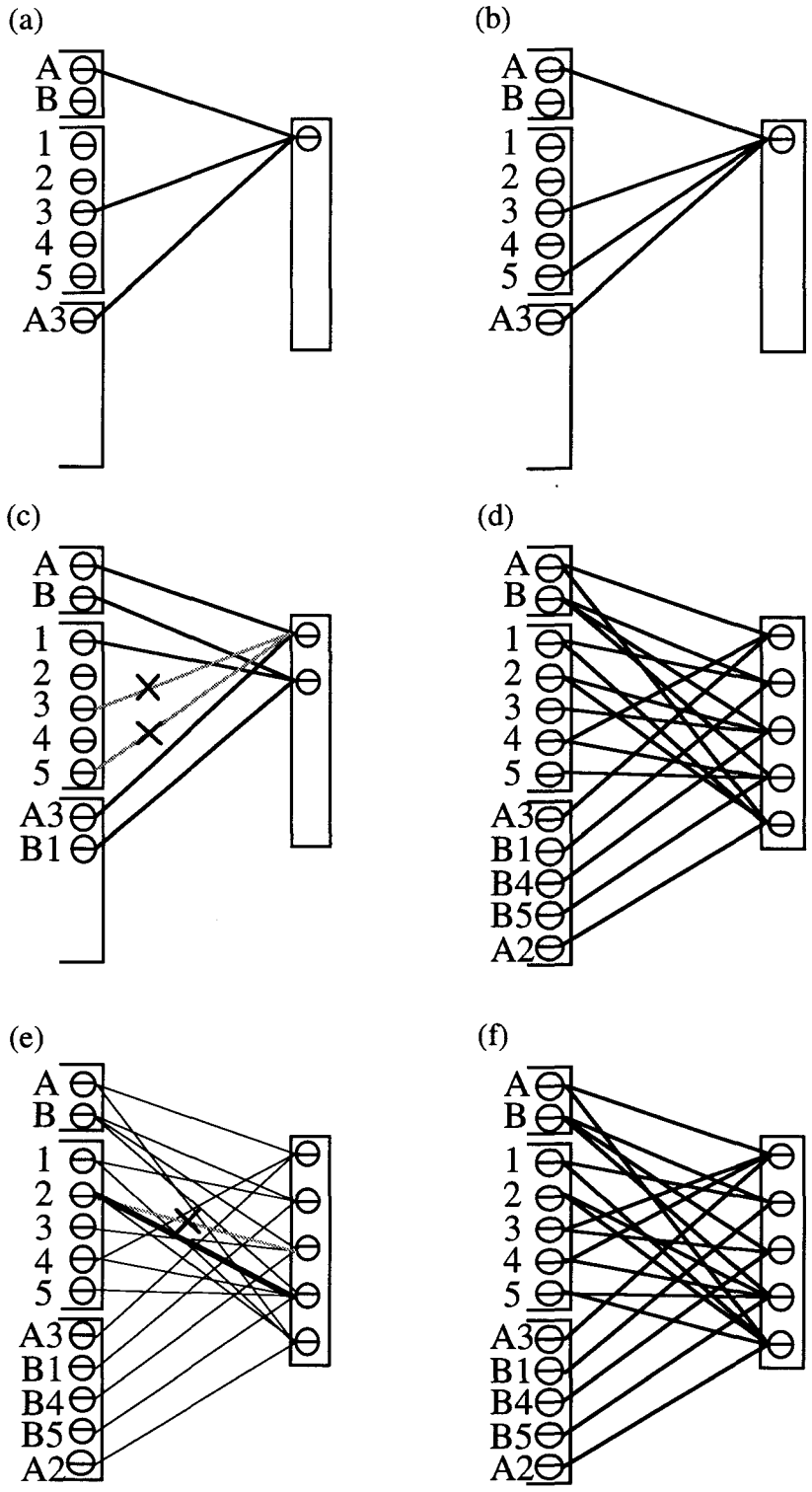

図4-9 注意視野推論野の関係ユニットに関する学習過程の例

4.3　顔の認識についての計算機実験

4.3.1　パターン木の収集とパターン照合モジュールの学習

男子60名，女子1名，計61名の平常顔をビデオカメラにて撮影した．図4-1(a)のような顔画像を61枚用意する．単純なパターンマッチングにより，一人につき7個の白黒の局所パターンからなる図4-1(b)のようなパターン木を得る．この時，局所パターンにおいて，画素値の平均，分散に関する規格化が行われる．

こうして得られたパターン木をもとに，パターン照合モジュールと関係モジュールの注意視野推論野の学習を行った．その結果，形成されたパターン照合モジュールの特徴ユニット数は49個であった．これは，61名のパターン木を，7×61枚の局所パターン中の49枚の局所パターンのみで記憶したことを意味する．改良型RCEネットワークのしきい値の下限値を，異なるクラス間では0に設定した．同じクラス間では，異なる顔の間で7つの局所パターンに対する特徴ユニットをすべて共有しないという条件のもとで，全体として特徴ユニット数が最小になるように人為的に設定した．生成された特徴ユニットの内訳は，頭部の局所パターンに対しては14個，顔面6個，左目・眉8個，右目・眉5個，鼻・口3個，鼻9個，口4個であった．学習にかかったエポック回数は4回であった．

4.3.2　実験結果

本モデルにおいて，視点制御推論野の12個の関係ユニットにより12種の視点移動を，注視パターン推論野の49個の関係ユニットにより61人の顔の特徴を記憶している．本モデルにより記憶した顔画像を選択的な視点の移動により認識した．図4-10(a)-(d)に，4名の顔に関して，通常の画像，パターン木，認識処理が完了するまでの視点移動の軌跡を示す．各局所パターンの左側に，パターン照合モジュールにおいてその入力に対して活性化する特徴ユニットの番号を記した．図4-10(a)の顔は，2度目の視点移動で選択的に鼻・口のパターンに視点を移動している．これは，注意視野パターンの検出により，図4-10(a)と図4-10(b)-(d)とを区別するクラスとして鼻・口が検出されたためである．図4-10(b)と(c)の顔は，鼻と口の局所パターンのみが，互いに異なる特徴ユニットを活性化させる．その結果，最終的に口のパターンまで視点の移動を行った．一度，右目・眉に視点を移動しているのは，図4-10(d)の顔でないことを確認するためである．図4-10(d)は，無駄

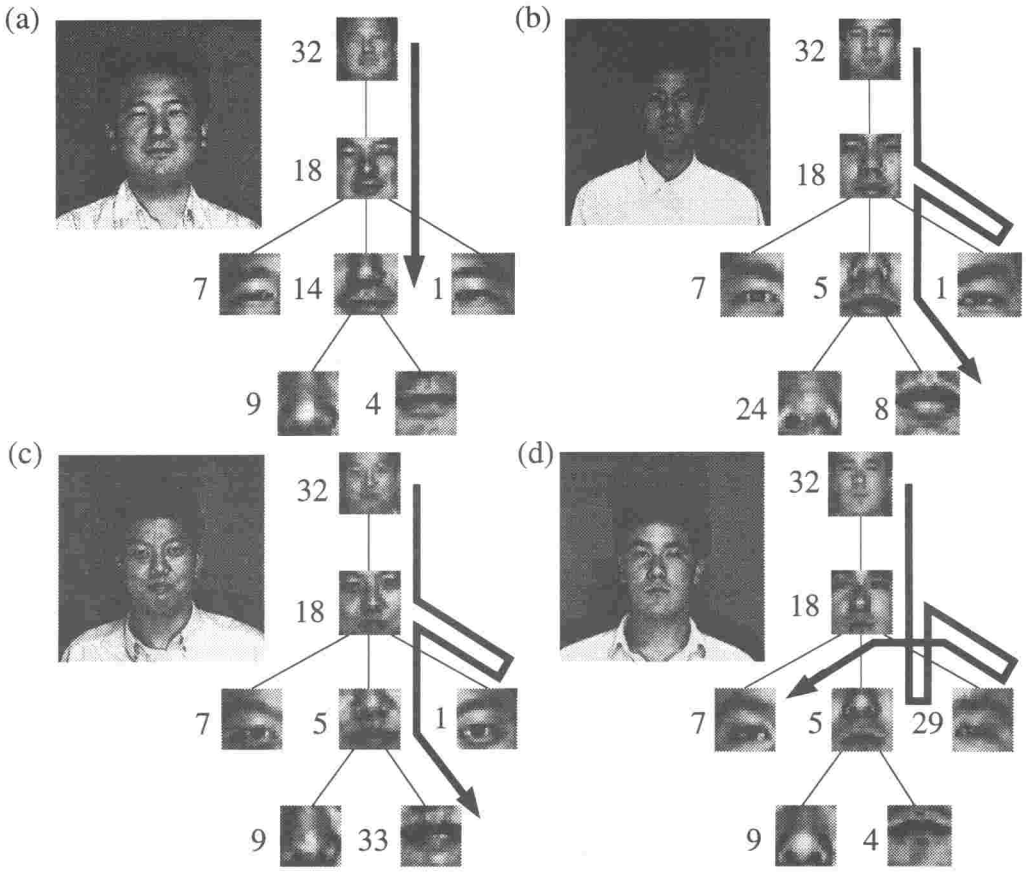

図4-10 顔の認識のための選択的な視点の移動

な視点の移動をしているように見えるが，これは図4-10に示さなかった複数の顔が，図4-10(d)と大部分の局所パターンで似ているために行われたものである．

本モデルにより，61名に関して顔を認識するまでの視点移動の回数は平均2.89回であった．それに対して，視点の移動に基づき，局所パターンの全数検査（縦型探索[111]）により顔の認識処理を行った場合の視点移動の回数は平均4.05回であった．ニューラルネットワークにより，効率の良い視点の移動が行われている．

4.4　結言

本研究において，カメラの移動とズーム操作による視覚情報の能動的認識のためのニューラルネットワークモデルを提案した．提案するモデルを実際の顔の認識に適用し，知識をもとにした選択的な視点の移動が行われることを確認した．

一般に用いられるニューラルネットワークモデルでは，全ての入力情報はあら

かじめ与えられている．本モデルでは，パターン木という階層的な知識源から視点の移動により逐次的に情報を集める．そのため，三つの認識過程に基づく逐次的な認識を行う．階層的な知識を，単純な二層構造で記憶できるため，どのレベルにおいても全く同じ三つの認識過程が用いられる．階層の順序に関する知識は，方向性のある結合で保持されている．全体と部分の関係で記述できる対象に対して，このような階層間での時間分割処理が有効であることは，Hintonによっても主張されている[112]．

第５章　光学的位置合わせの不要な光ニューラルネットワーク

５.１　緒言

視覚情報は膨大であるため，それらを扱うニューラルネットワークは大規模になる．大規模なニューラルネットワークを実現する大規模光ニューラルネットワークにおいては，処理ユニットとして働く素子の空間分布が高密度になり，素子間の位置合わせが困難になる．特に，光学系に歪曲収差等を含む場合は致命的な問題となる．

本章において，光軸に垂直な方向に関する光学的位置合わせの不要な空間光変調器型の光ニューラルネットワークを提案する[113,114]．この光ニューラルネットワークを自己調整型光ニューラルネットワークと呼ぶ．本章ではじめて提案する光学系の自己調整機能に基づいて，学習時に結合荷重を空間光変調器に記録する．本光ニューラルネットワークは，Willshawのネットワーク[115]を光学的に実現したものである．Willshawのネットワークは，入力，出力，結合荷重が全て二値の値をとる二層のニューラルネットワークである．光学実験により，提案する光学系の自己調整機能の有効性を確認した．実験では，スパース符号化法[116]を記憶容量の向上のために用いた．スパース符号化に，ADAM (advanced distributed associative memory)[117]の前部のネットワークを改善したネットワークを用いた．

５.２　自己調整型光ニューラルネットワーク

この節において自己調整型光ニューラルネットワークの原理を述べる．このネットワークの認識過程は従来の空間光変調器型の光ニューラルネットワークと同じであるが，学習過程は異なる．光軸に垂直な面内での厳密な位置合わせがこの学習過程により不要になる．

５.２.２　構造と認識過程

Willshawのネットワーク[115]において光学系の自己調整機能を実現する．Willshawのネットワークの認識過程においては，二値のベクトル-マトリックス演算としきい値処理が必要である．$\boldsymbol{v}^{(in)}$と$\boldsymbol{v}^{(out)}$をそれぞれ入力層，出力層のユニッ

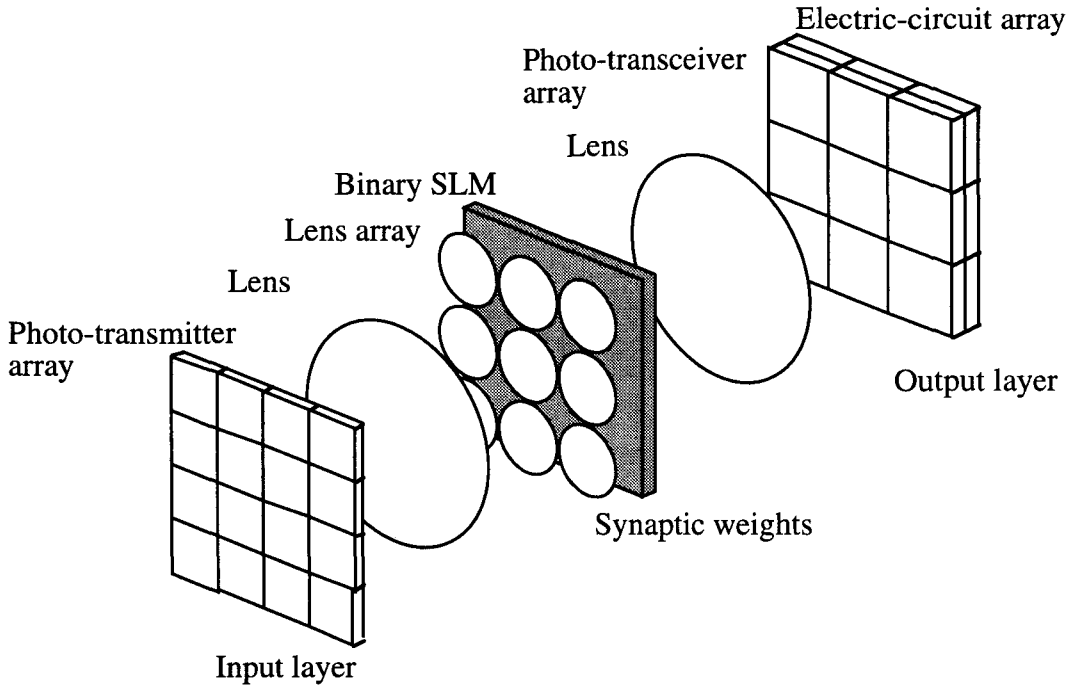

図5-1 自己調整型光ニューラルネットワークの理想的な構造

トの出力ベクトル，Wをユニット間の結合行列とした場合，Willshawのネットワークの伝達関数を以下に示す．

$$\boldsymbol{v}^{(\mathrm{out})} = f(\boldsymbol{W}\boldsymbol{v}^{(\mathrm{in})}). \qquad (5\text{-}1)$$

$f(\cdot)$はベクトルの成分ごとのしきい値演算である．式(5-1)の演算を光学的に行うために，２．４．２で説明したようにIshikawaらによって提案された空間符号化法[48]に基づく空間光変調器型ニューラルネットワークを用いる．

図5-1に自己調整型光ニューラルネットワークの理想的な構造を示す．２．４．２で述べた従来の空間光変調器型光ニューラルネットワークとほぼ同じ構成をしている．ただし，出力層の各ユニットは発光と受光の両方を同じ面で行える素子（受発光素子:optical transceiver）を持つものとする．認識過程も従来の光ニューラルネットワークと同じであるが，二値の値のみを扱うことに注意して，以下に，本光ニューラルネットワークの認識過程を説明する．

入力画像が入力層に与えられたとき，入力層の発光素子アレイから空間光変調器に0, 1の二値の入力信号光が伝達される．単レンズとレンズアレイにより，入力画像の多数の複製が空間光変調器上に結像される．空間光変調器には，学習によって得られた結合荷重が，二次元パターンである透過率分布として記憶されて

いる．空間光変調器を透過した信号光は，受発光素子アレイ上に結像される．それぞれの受発光素子の受光部は，その開口に入射する光の全パワーを検出する．VSTEP(vertical-to-surface transmission electro-photonic device)[118]が理想的な受発光素子の例である．検出された光信号は電子回路アレイ内でしきい値と比較される．この回路アレイは，出力として二値画像を生成する．

５．２．３　学習過程における光学系の自己調整機能

学習過程において，学習対は逐次的にネットワークに与えられる．学習対は，ここでは入力画像 $\boldsymbol{v}^{(in)}$ と理想的な出力画像 $\boldsymbol{v}^{(out)}$ の対を表す．ネットワークを入力画像の分類に用いる場合は，出力画像においてその入力画像のクラスに対応づけた一画素のみを活性化させる．

学習過程において，入力層と出力層の発光素子アレイは，光信号を空間光変調器に同時に送る．図5-2に，光学的位置合わせが正しくない場合の結合荷重の記録例を示す．入力層の4×4のユニットの出力値で構成される入力画像$\boldsymbol{v}^{(in)}$の複製像が空間光変調器上に結像されている．また，3×3の出力層のユニットの出力値で構成される理想的な出力画像$\boldsymbol{v}^{(out)}$が空間光変調器の反対側に結像されている．空間光変調器は，入力画像の複製像と出力画像との間で各点ごとにAND演算を行い，その結果を保持する．すなわち，結合荷重は，空間光変調器上で複製された3×3個の$\boldsymbol{v}^{(in)}$と出力層からの$\boldsymbol{v}^{(out)}$の両方の光信号が重なる場所に記録される．入力層のユニット j と出力層のユニット i との間の結合荷重 w_{ij} は，図5-2の黒色の部分に表すように，信号 $v^{(in)}_{j}$, $v^{(out)}_{i}$ で決められる空間光変調器上の位置に記録される．光学系の位置合わせが正しく行われている場合は，複製像のそれぞれは，理想的な出力画像からの結像画像の各画素に一致する．この場合，記録された結合荷重は，結合荷重の正しい空間符号化[48]が実現される．しかし，図5-2の場合は位置合わせが正しく行われていないため，結合荷重 w_{ij} は四つの部分に分割される．この場合でも，信号は認識過程で正しく処理される．認識過程で入力信号 $v^{(in)}_{j}$ が入力層に与えられた時を考える．この時，入力信号光は，学習時に記録された結合荷重 w_{ij} と同じ位置を正確に透過する．透過した信号光は光の相反性[119]に基づき出力層のユニット i に伝達する．即ち，光学系の位置合わせが正しく行われていない場合でも正しくベクトル-マトリックス演算 $W\boldsymbol{v}^{(in)}$ が実行される．この性質は，学

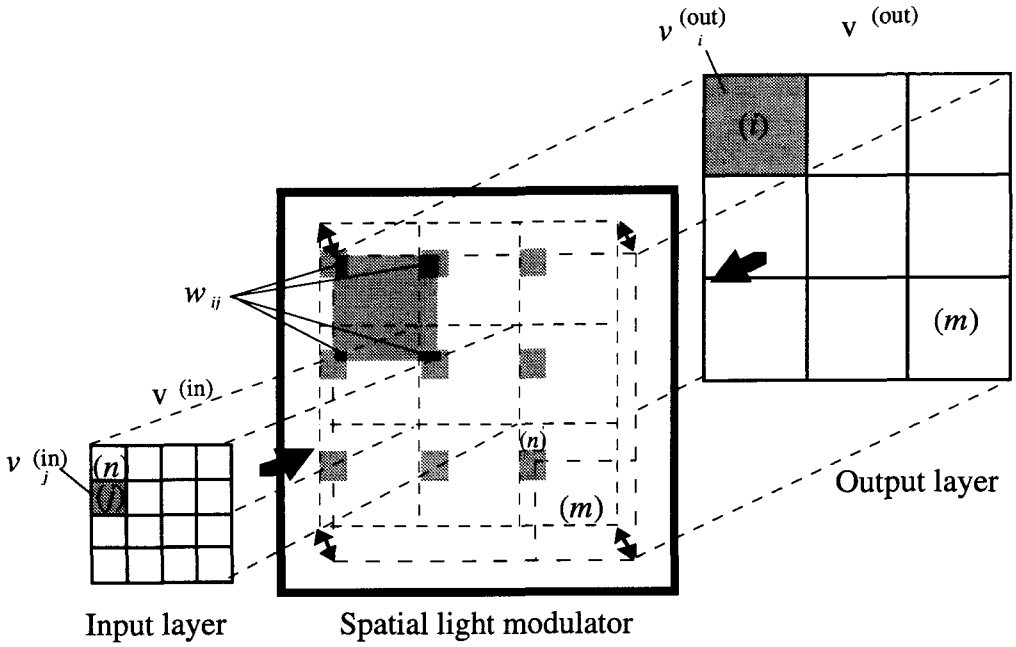

図5-2 結合荷重の記録機構

習過程の後，光学系が変形しない限り保存される．

自己調整型光ニューラルネットワークにおいて，光学的位置合わせがある条件以上に大きく狂った場合，幾つかの入力層のユニットと出力層のユニット間が光学的に接続されない．入力と出力ユニット間が光学的に接続されるためには，空間光変調器上で出力ユニットから結像された領域が，入力ユニットから結像された領域と一部分でも重なるという光接続の条件が必要である．例えば，図5-2でユニット n が入力層の左上の角に，ユニット m が出力層の右下の角に位置すると考える．これらのユニット間の場合，光接続に関する上記の条件を満たしていないため，ユニット n と m は光学的に接続されていない．図5-3に全てのユニット間で光接続の条件が満たされる手法を二つ示す．図5-3は入力層から空間光変調器上に複製結像された像と出力層のある特定のユニットからの像を表す．第一の手法においては，図5-3(a)に示すように複製された入力画像の間にマージンを与える．マージンの長さを空間光変調器上で M とした時，光学系の位置合わせの際，空間光変調器上でのずれを長さ M まで許容する．第二の手法は，レンズアレイに加えられた付加的なレンズにより，図5-3(b)に示すように複製像を増加させて結像された領域を拡張する方法である．拡張された領域の幅を空間光変調器上で D とし

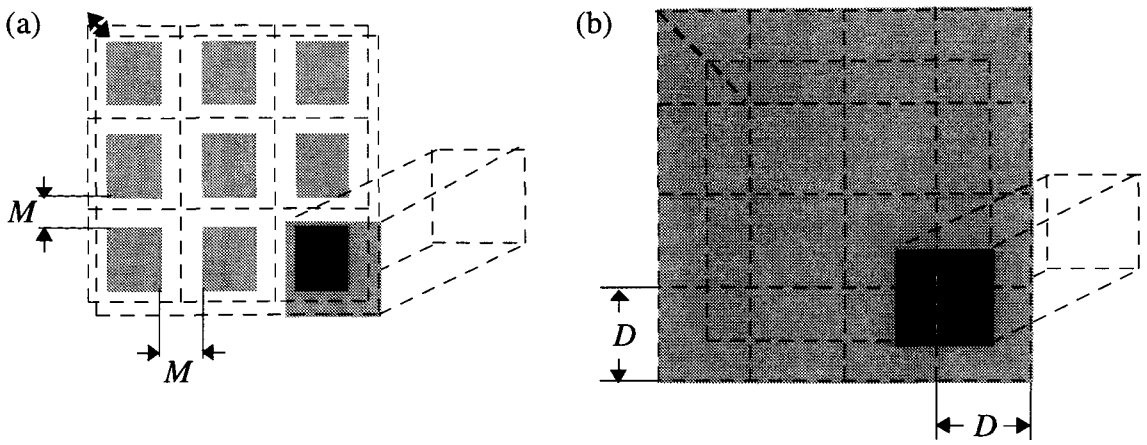

図5-3 光接続の条件を満たす方法

た場合，光学系の位置合わせは，空間光変調器上でのずれを長さDまで許容する．次節に示す実験では，図5-3(a)の手法が用いられる．

逐次的に与えられる学習対に対するAND演算の繰り返しにより，空間光変調器上に次々と演算結果の追加書き込みが行われると仮定する．この仮定は，記録されている空間光変調器上の画像と現在の演算結果の間でOR演算が実行されることを意味する．n番目の学習対がネットワークに与えられた時を考える．$v^{(in)(n)}{}_j$と$v^{(out)(n)}{}_i$をそれぞれ入力層のユニットj，出力層のユニットiからの出力信号とすると入力層と出力層のユニットiとjを接続する新しい結合荷重$w^{(n)}{}_{ij}$は以下のように書ける．

$$w^{(n)}{}_{ij} = w^{(n-1)}{}_{ij} \cup (v^{(in)(n)}{}_j \cap v^{(out)(n)}{}_i) \qquad (5\text{-}2)$$

この時，全ての変数は二値であり，∩と∪はそれぞれANDとOR演算を表す．結合荷重の初期値は0に設定される．この学習方法により，Willshawの結合荷重の設定方法によって決められた結合荷重の値と同じ値の結合荷重を得ることが出来る．

光書き込み型の強誘電性液晶空間光変調器[120]は，上で述べたANDとOR演算を実行することが出来るため，自己調整型光ニューラルネットワークにおける空間光変調器として用いることが出来る．偏光子と検光子で挟まれた空間光変調器に電気的制御パルスが印加された場合，入射光強度があるしきい値（記録しきい値）を越えている領域は，強誘電性液晶分子の回転により透過率が1に変化する．記録しきい値は，電気パルスの電圧の高さや幅を制御することで調節することが出来る．記録しきい値は，入力層からの複製像と出力層からの像の間で正確にAND演算が実行できる値に調整する．この空間光変調器において，液晶分子が回転し

た場合，逆方向に電圧が加えられるまでその状態は保持される．したがって，空間光変調器上で追加的に強度変調の空間分布を記憶することによりOR演算が実行できる．

５．３　実験

本章で提案する光ニューラルネットワークの自己調整機能に関する原理確認実験を行った．このとき，ネットワークの記憶容量を向上させるためにネットワークの直前にスパース符号化部を付加した．このスパース符号化部を先に説明し，次に光学系，実験結果の順で説明する．

５．３．１　スパース符号化ネットワークによる記憶容量の向上

Willshawのネットワークにおいて出力層のユニットは一定のしきい値を持っている．このしきい値は，入力層の活性度に従って設定される．活性度は，層内の活性化したユニットの数，または二値の入力画像ベクトルの中の１の値を持つ画素の数で定義される．入力画像の活性度が画像ごとに変化する場合，出力層に適

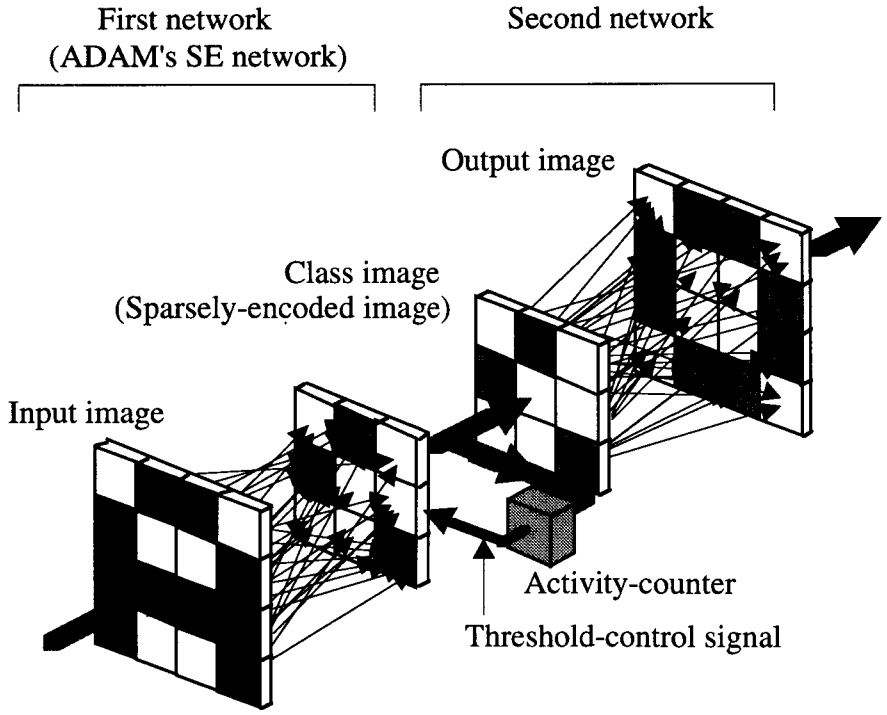

図5-4 ADAMの構造

当なしきい値を設定することが困難になる．ADAMと呼ばれるネットワーク[117]において，この問題点が取り除かれている．ADAMの構造を図5-4に示す．ADAMの前部のネットワークの出力層において，各ユニットのしきい値は，活性度検出器からの信号によって，活性度が一定に保たれるように常に制御される．すなわち，前部のネットワークは，入力画像を一定の活性度を持つ画像（クラス画像）に変換する．ADAMの記憶容量は，スパース符号化法[116]により向上する．スパース符号化法とは，記憶したい画像の活性度が小さいほど，ニューラルネットワークにおいて記憶できる情報の量が増加することを利用した入力画像の符号化法である．スパース符号化のために，クラス画像の活性度を小さく設定する．このようにスパース符号化を行う場合のADAMの前部のネットワークをスパース符号化ネットワークと呼ぶことにする．スパース符号化ネットワークの出力層を，ここではスパース符号化層，クラス画像はスパース符号化画像と呼ぶことにする．

実験では，ADAMのスパース符号化ネットワークに学習機能を与えるために改良したスパース符号化ネットワークを用いる．改良したスパース符号化ネットワークのユニット間の結合荷重は，学習用の入力画像を用いて学習される．初期の結合荷重として，二値でかつランダムな値が与えられる．n番目の学習対がネットワークに提示され，信号がスパース符号化層まで伝達したとする．その時，$v^{(\mathrm{in})(n)}{}_j$と$f^{(\mathrm{out})(n)}{}_i$がそれぞれ入力層のユニット$j$とスパース符号化層でのユニット$i$の出力信号を表すものとする．この時，これらのユニットの結合荷重$m^{(n)}{}_{ij}$は以下のように修正される．

$$m^{(n)}{}_{ij} = \mathrm{m}^{(n-1)}{}_{ij} \cap (v^{(\mathrm{in})(n)}{}_j \cup \bar{f}^{(\mathrm{out})(n)}{}_i). \qquad (5\text{-}3)$$

すべての変数は二値であり，$\bar{f}$ はfの論理的否定を表す．この学習に従って結合荷重が枝刈りされる．このネットワークを認識処理に用いる時には，学習過程で全く活性化しなかったユニットを消去する．

改善されたスパース符号化ネットワークは，本章で提案している光学系の自己調整機能を持つ光ニューラルネットワークにおいて実現することが可能である．強誘電性液晶空間光変調器の消去モードを用いる．消去モードでは，しきい値を越える光信号が入射した領域において，前の記録状態に関わらず光の変調は状態0に変更される．これにより結合荷重の枝刈りが実現される．式(5-3)の否定に対してド・モルガンの定理を用いると，強誘電性液晶の消去モードを用いて，入力

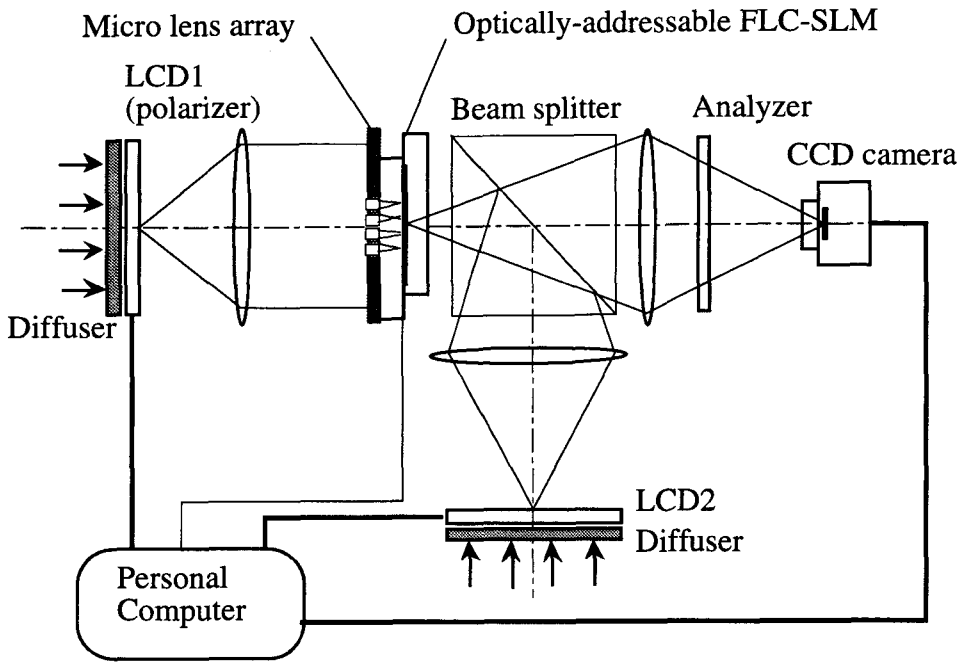

図5-5 実験光学系

画像の反転を作ることにより，式(5-3)で示される演算が実現できることを示すことできる．また，活性度検出部は，電子回路部で電気的に実現することが可能である[121]．

5.3.2　光学系

図5-5に，図5-1の理想的な自己調整型光ニューラルネットワークの有効性を確認するための実験光学系の構成を示す．スパース符号化ネットワークの演算はパーソナルコンピュータ上で行った．光書き込み型の強誘電性液晶として，LAPS-SLM[122]を用いた．レンズアレイとして，4×4の屈折率分布型のマイクロレンズ（セルホックマイクロレンズ[123]）を用いた．マイクロレンズの焦点距離は，レンズ端面がLAPS-SLMに直接接触したときにLAPS-SLMの光感応層に複製像を結像するように設定した．セルホックマイクロレンズの直径は2.0mmである．長さと焦点距離はそれぞれ2.61mmと5.89mmである．発光素子アレイとして，インコヒーレント光源によって照明された液晶ディスプレイ(LCD1: liquid crystal display 1)を用いた．LCD2とCCD(charge-coupled device)カメラは，ビームスプリッタによって空間光変調器から見て光学的に同位置に置くことにより，受発光素子

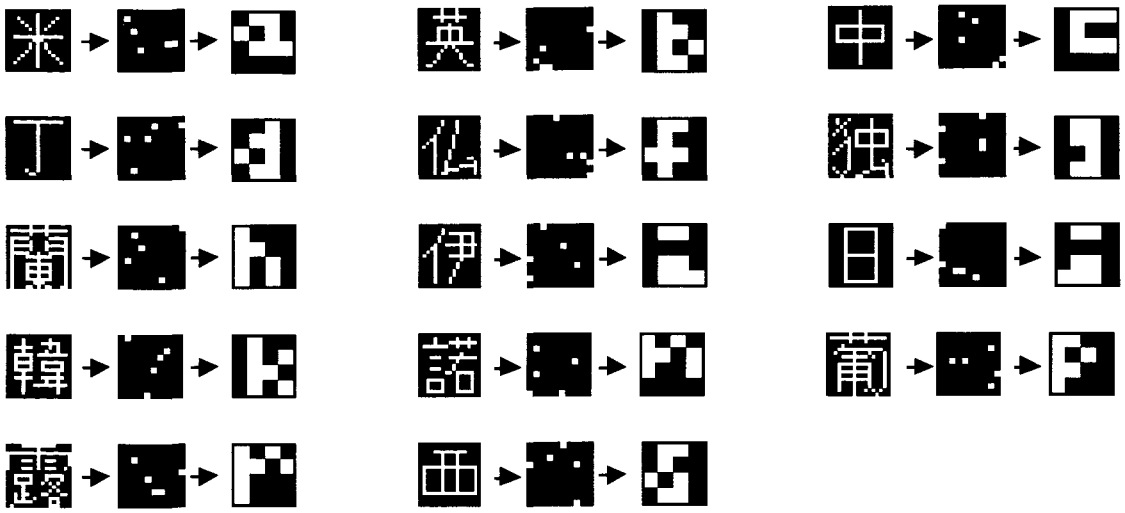

図5-6 実験で用いられる学習対とスパース符号化画像の組

アレイとして用いた．このシステムでは，光接続の条件を図5-3(a)に示す方法で満たした．複製結像された入力画像間のマージンは約1.8mmであるため，空間光変調器上での位置合わせのずれの許容範囲は1.8mmである．

本光学系における学習過程を説明する．スパース符号化ネットワークは，パーソナルコンピュータ内で学習用の画像に対する学習が行われているものとする．スパース符号化ネットワークによって得られたスパース符号化画像がLCD1に表示され，理想的な出力画像がLCD2上に表示されるものとする．パーソナルコンピュータからの制御パルスにより，空間光変調器上に結合荷重を記録する．全ての学習対に対して上記の処理を行い学習が終了する．本光学系における処理過程を説明する．入力画像が，スパース符号化ネットワークにより符号化される．得られたスパース符号化画像がLCD1に表示される．スパース符号化画像は，レンズアレイにより空間光変調器上に複製される．空間光変調器に記憶された結合荷重を表す透過率分布により複製像は変調される．変調された複製像は，CCDカメラで画像としてパーソナルコンピュータに取り込まれる．取り込まれた画像を，パーソナルコンピュータ内で局所的に足し込み，その和を二値化して出力層のユニットの出力を得る．

5.3.3　実験結果

漢字の画像からアルファベットの画像への連想実験を行った．漢字の画像のサ

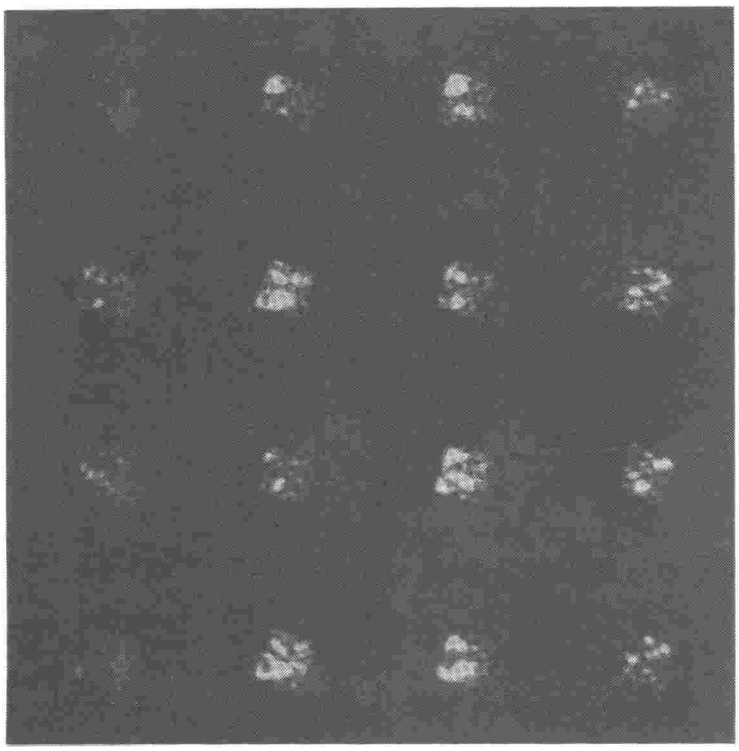

図5-7 学習後のLAPS-SLM上の結合荷重

イズは17×17画素，アルファベットの画像は4×4画素である．スパース符号化層では，15×15のユニットが用いられる．理想的な出力であるアルファベットの画像は，すべてのアルファベット画像が同じ活性度を持つようにした．これは，出力層の活性度も一定に設定することにより，ネットワークの記憶容量が増加することが知られているためである[124]．図5-6に実験で用いられる学習対とそれらのスパース符号化画像を示す．実験において，連想に必要な3600本の結合の実現により，これら14組の学習対を完全に記憶することが出来た．図5-7に，学習後にLAPS-SLMに記録された結合荷重を示す．入力信号光の偏光方向が，LAPS-SLMの液晶分子の配向方向と一致するように，LCD1を約20度傾けているため，LAPS-SLM上でのスパース符号化画像は回転している．図5-7はLAPS-SLM上の約1cm×1cmの領域を示している．記録密度は空間光変調器上で1mm^{2}あたり300画素である．光学系の自己調整機能なしに，実際にそのような厳密な位置合わせを行うのは不可能である．本実験システムの記憶容量は，LAPS-SLM上の結合荷重の記録のにじみによって制限されている．にじみはスパース符号化画像の数回の記録の繰り返し後に起こる．図5-7で，多くのスパース符号化画像が上書きされた位

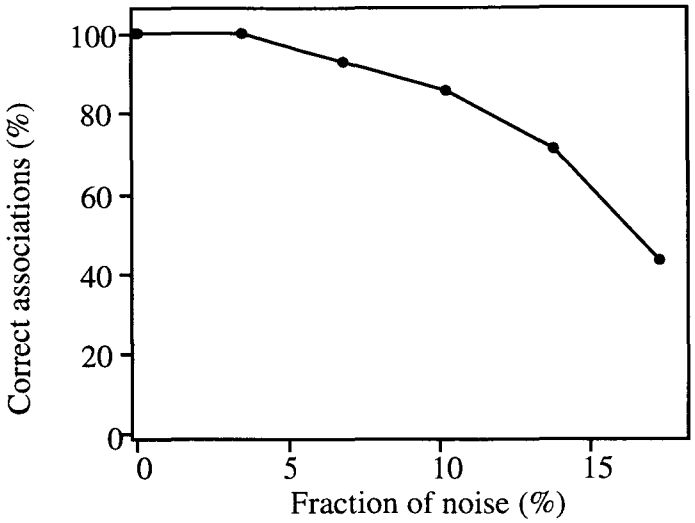

図5-8 ネットワークのロバスト性についての実験結果

置でにじみが起こっているのが分かる．

本光ニューラルネットワークシステムの，ランダムノイズに対するロバスト性について調べた．ノイズのない学習集合で学習されたスパース符号化付き光ニューラルネットワークにランダムノイズを含んだ入力パターンを提示する．図5-8にロバスト性に関する実験結果を示す．正しい連想率をノイズの割合（全画素数に対するノイズの画素数の百分率）に関する関数として表す．本システムは，5%以下の小さい割合のノイズに対して十分ロバストである．入力にノイズの存在する場合の連想結果の例を図5-9に示す．最上行は0% のノイズを含むパターンに対して，中間と最下行はそれぞれ10% と 17% のノイズに対する連想の例である．スパース符号化は 10% のノイズに対してロバストで，スパース符号化画像は全く変化しない．しかしながら，17% のノイズの場合は，スパース符号化画像はわずかに変化する．にもかかわらず，最終的な連想は正確に行われている．ノイズに対するロバスト性は二段階の構造によって増強されることが分かる．スパース符号化ネットワークがない場合，14 組の学習対の中の 3 組しか記憶することが出来なかった．

5.4　結言

空間光変調器型の光ニューラルネットワークにおいて光学系の自己調整機能を提案した．この方法の最も重要な要素は，学習過程において空間光変調器上で

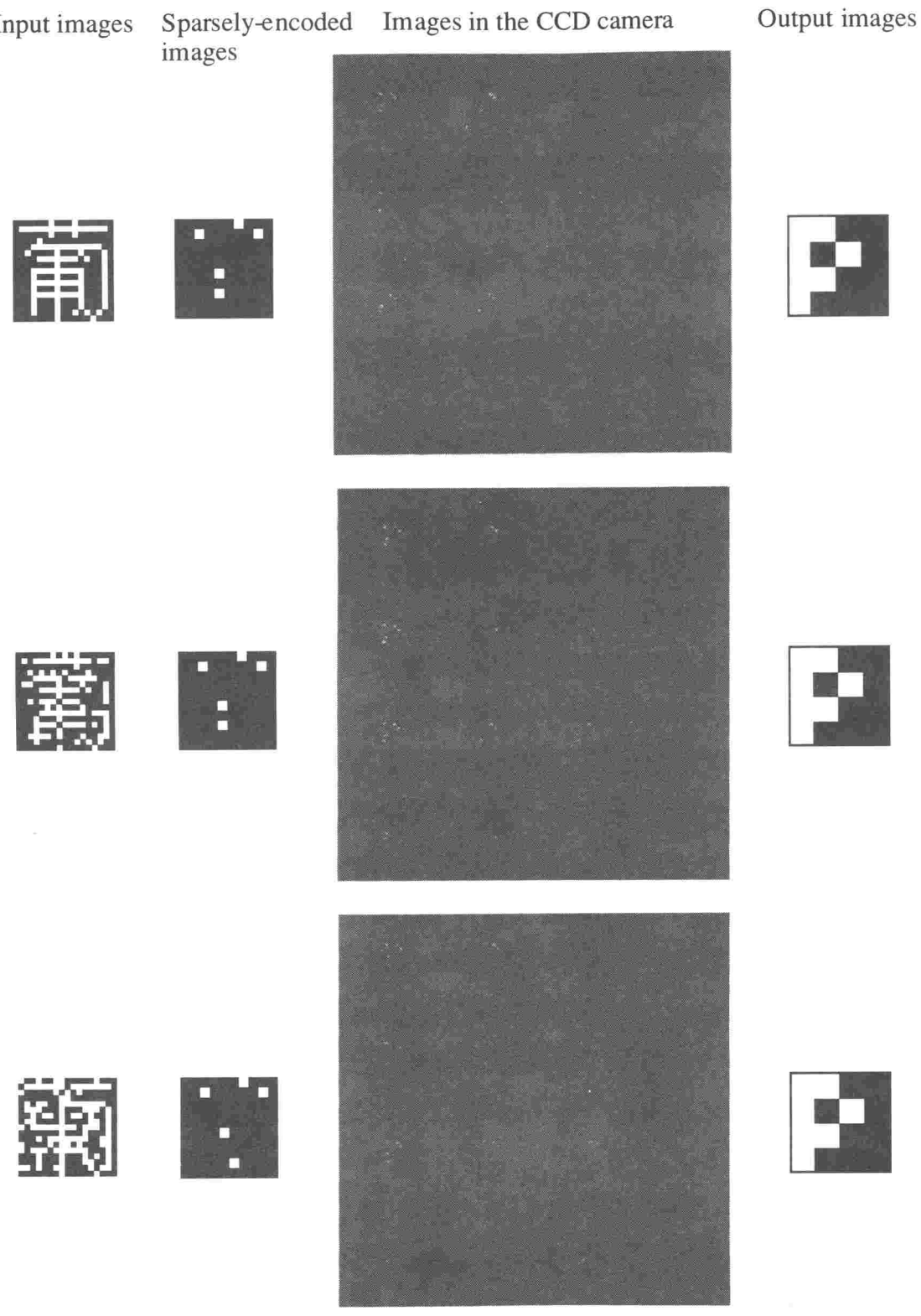

図5-9 ノイズを含んだ画像の連想結果（一列目は０％のノイズ，２列目，３列目はそれぞれ１０％，１７％のノイズが含まれている場合を示す）

AND演算の結果が得られ，その結果が空間光変調器上に同時に書き込まれることである．本章で提案する自己調整機能により，光軸に垂直な面内での光学系の位置合わせが不要になった．また，光学実験により1mm^{2}あたり300画素の高密度な記録が行われることを確認した．

本光ニューラルネットワークにおいて，光学系の光軸に垂直な面内における位置合わせ行う必要はないが，光学系の複製像の焦点だけは厳密に調整する必要がある．複製像のデフォーカスは空間光変調器の持つ高い空間分解能を無駄にしてしまう．その対策としては，光学的パッケージング技術[125,126]を用いれば，焦点面合わせの必要を省くことが出来る可能性がある．

総括

本論文において，ニューラルネットワークによる視覚情報の認識手法を実世界において応用する際に生じる問題に着目し，それらを解決をする新しいニューラルネットワークモデルを提案した．ここでは，本論文で得られた研究成果と新しい知見をまとめ，今後の研究課題について述べる．

第１章において，第一にニューラルネットワークによる視覚情報認識手法の歴史について概説し，現在，モデル化の困難な実世界で用いられる視覚情報認識手法の一つとして，ニューラルネットワークによる手法が期待されていることを述べた．次に，従来法を実世界における視覚情報認識に用いる際に生じる幾つかの問題として，以下の三点を挙げた．

（１）棄却能力に関する問題．

（２）複雑な対象の認識に関する問題．

（３）大規模光ニューラルネットワークにおける光学的位置合わせに関する問題．

第２章において，上の問題に対処する際の基本であり，また本研究の前提となる視覚情報の認識のためのニューラルネットワークに関するこれまでの主な研究をまとめた．基本的なニューラルネットワークとして，相関型または距離型のユニットを用いてニューラルネットワークを構成した場合の認識原理についてまとめた．ニューラルネットワークの学習法としては，特に第３章で提案したモデルの基本となるRCE学習法について詳しく述べた．また，第３章で取り上げた問題である棄却能力の定義を述べた．次に，視覚情報認識のためのニューラルネットワーク構造として，位置不変結合階層型構造，モジュール型構造，調和モデル型構造を概説した．調和モデル型構造は，視覚情報の認識にこれまでは利用されていなかったが，第４章でその基本的構造を用いたので説明した．最後に，大規模光ニューラルネットワークをホログラム型と空間光変調器型に大別して概説した．特に，ホログラム型で行われる光学系の自己調整機能について詳述した．

第３章において，認識結果の信頼性を向上するために，未知画像や曖昧な画像を適度に棄却することが可能であるニューラルネットワークモデルを新しく提案した．提案したモデルは，RCEネットワークを基本とするが，ユニットには，二

つのしきい値を用い，またユニットへの結合ベクトルを学習過程において修正する．ユニットの持つ二つのしきい値は，パターンベクトル空間内で二重の超球を形成し，超球の組み合わせで未知画像や曖昧な画像を適度に棄却することが出来た．X線フィルムに記録された患者ID番号を用いた文字認識の実験において，本論文で提案するニューラルネットワークモデルの棄却能力が高いことを示した．その際，認識率，学習速度，使用メモリ容量の点で，従来のRCEネットワークの利点を損なうことはなかった．ただし，ネットワークのパラメータを変えることにより，さらに棄却能力を向上させた場合，認識率が低下した．

第4章において，複雑な対象を認識するために，大量の情報が含まれる実世界の視覚情報の中から認識に必要な情報を能動的かつ選択的に取得するニューラルネットワークモデルを提案した．モデルは，4つの機能モジュールとそれらを接続する関係モジュールで構成した．視点の移動とズーム操作に基づいて得られる視覚情報を，パターン木と呼ばれる形で表現して取り扱った．本モデルでは，パターン木に含まれるパターンの情報とそのパターン間の位置関係の情報をネットワークを構成するユニット間の接続パターンとして記憶した．関係モジュールと機能モジュール間の相互作用により，効率的に対象を認識するために次に視点を移すべきパターンのクラスと，そのパターンを捉えるために行う最初の視点移動のための制御信号を発生させた．61名の実際の顔画像から得られたパターン木を利用して，本モデルにおいて，能動的かつ選択的な視点の移動により効率よく対象が認識できることを実証した．

第5章において，光学的位置合わせが不要である光ニューラルネットワークを提案した．本ネットワークでは，空間光変調器の異なる面に，それぞれ結像された入力画像の複製像と理想的な出力画像間のAND演算の結果を結合荷重として記録した．この時，光の相反性により記録時の光線と処理時の光線が同じ光路を通るため，結合荷重は正しい処理が行われる位置に自己調整されている．また，本光ニューラルネットワークの記憶容量を増加する手法として，入力画像を活性度の低い画像に符号化するスパース符号化法を用いた．そのスパース符号化は，結合荷重の記録位置の自己調整を行う提案する光学系を適応可能であるニューラルネットワークモデルにより行った．このように，本手法では二段のニューラルネットワークにより認識処理を行うことになる．光学実験において，自己調整機能に

より高密度な結合荷重が正しい位置に記録が出来ることを示した．空間光変調器としては，光書き込み型の強誘電性液晶空間光変調器を用いた．また，実験において，スパース符号化によって，記憶容量が向上できることを示した．

本論文において，ニューラルネットワークによる実世界の視覚情報認識において，課題となる三問題に対する解決例として，新しい三種のニューラルネットワークモデルを提案した．しかし，提案したモデルは，それらの問題の一解決例を示したに過ぎない．以下に本論文で取り上げた各問題について，今後の課題を述べる．

棄却能力

第３章で提案した二重超球ネットワークは，X線フィルムの文字読み取り実験に対して高い棄却能力を得た．しかし，他の問題に対しても同じく棄却能力の高い認識を行うことは保証できない．ここで，二重超球ネットワークを用いて，画像ベクトル空間中の超球状でない領域を占めるクラスを分類する場合を考える．学習原理から考察すると，そのクラスを一つの超球で覆うことを許しているために超球内に未知画像が含まれてしまう．そのため，このような問題に対する棄却能力には限界がある．今後，二重超球ネットワークを様々なデータベースに適用し，その能力が有効に発揮できる応用分野を明確にし，二重超球ネットワークで高い棄却能力の得られない問題に対しては，新たな工夫を行う必要がある．また，提案する二重超球ネットワークは，単純な階層型のニューラルネットワークであるため，第２章で述べたように活字文字程度の認識しか実現できない．今後，２．３で説明した視覚情報認識のためのニューラルネットワーク構造等のように，提案する二重超球ネットワークを構造化していく必要がある．

実世界における視覚情報には，モデル化された世界では起こらない様々な状況が発生する．例えば，学習用の情報が繰り返し与えられない場合，誤った学習用の画像ベクトルが与えられる場合，環境が変化する場合等である．これらの状況においても，結果の安全性のためには信頼性の高い認識を行う必要がある．提案した二重超球ネットワークは，これらの状況に対して信頼性の高い認識を実現することは出来ない．今後，これらの状況を考慮したモデルの構築が望まれる．

人間は視覚情報だけでなく，他の器官から得られた情報をも融合した結果を基に行動している．視覚から得た結果を誤って認識しても，他の器官からの情報等による補正が可能である．このように，視覚情報の認識システムの信頼性も，それが動作する全体のシステムの中で考える必要がある．特に，第3章の実験結果で得られたように，棄却能力をより高めると認識率が低下するような場合は，全体システムにおいて適当な棄却能力と認識率を評価する必要がある．

複雑な対象の認識

第4章で提案した能動的な視覚情報の認識を行うモデルでは，顔に関する局所パターンの位置関係が既に記憶されていると仮定して，その他の結合を学習した．局所パターンの位置関係が記憶されていない物体を記憶する際には，局所パターンの位置関係を学習する必要がある．また，今回のモデルにおいて，視点の移動は，記憶された顔画像に関する情報と現在の入力パターンが何(what)であるかの情報のみを用いて行われた．パターン木の階層性が同じ顔の認識の場合，視点の移動先には対象とする局所パターンが必ず存在した．しかし，パターン木の階層性が多様な物体の認識に本モデルを適用するとき，視点の移動先には対象とする局所パターンが存在しないことが生じる．この問題は，生理学的にも確認されているように[127]，入力パターンから視点を移すべき局所パターンがどこ(where)に存在するかの情報を得る機能モジュール（位置検出モジュール）を本モデルに加えることにより解決できる．しかし，実世界の視覚情報を入力とした位置検出モジュールの実現は現状技術では非常に困難である．現在，文字認識の分野においても，安定な文字の切り出しは大きな課題となっている．今後，視覚情報から物体や局所パターンを安定に切り出す（位置を検出する）研究を進めて行くことが重要である．

本ニューラルネットワークモデルにおいて選択的に視点が移動する例として顔の認識を取り上げた．実際の顔の認識では，顔の向きの変化，表情の変化等に対して不変な認識を行う必要がある．本モデルはそのような認識を行うことが出来ないため，ニューラルネットワークの持つ認識の柔軟性を全く発揮していない．ニューラルネットワークが持つ処理の柔軟性は，記号を，活性化したユニットで構成されるパターンで表現することから発現する．しかし，本論文で提案する能動的認識モデルでは，特徴層において，記号はパターンではなく単一のユニット

で表されているため柔軟な処理が実現できない．唯一，パターン照合モジュールにおける照合のみがニューラルネットワークが持つ処理の柔軟性を発揮している．今後，ニューラルネットワークによる能動的な視覚情報の認識において，現在の機能を失わず，柔軟な認識のために記号情報をパターンで表現したモデルを構築して行く必要がある．

本論文において，複雑な対象の認識法として，視点の移動とズーム操作に基づく能動的認識を行った．能動的認識作業は必然的に動きを伴い，逐次的に処理される．一方，ネオコグニトロン[20]や振動子を用いた物体認識モデル[26]では，局所的な特徴を全て並列に検出したうえで認識している．並列ハードウエアを考えた場合，特徴を全て並列に検出したほうが，高速な認識が実現できる．しかし，膨大な視覚情報を全て取り込み，並列に認識処理を行うことは，装置の規模等の点から限界がある．実世界のための視覚情報認識システムの構築には，本論文で提案したような逐次的な能動的認識手法が不可欠である．今後，実際の問題において，逐次的な能動的認識手法と，並列的な手法の最適な組み合わせを検討する必要がある．

光学的位置合わせ

本論文で提案した光学系の自己調整機能において，空間光変調器上での画像間のAND演算とOR演算が要求される．本光学実験では，画像間のAND演算とOR演算には，光書き込み型の強誘電性液晶を用いた．このとき，光量や書き込みのための制御パルスを適当に調整しないと正確にAND演算が実行できない．また，OR演算を繰り返し行った場合に，にじみが生じるという問題がある．今後，安定したAND演算，OR演算を実行できる空間光変調器の開発が望まれる．

提案した自己調整型光ニューラルネットワークにおいて，結合荷重は0, 1の二値であった．多値の荷重を実現するためには，空間光変調器上で多値画像間の積演算と，その演算結果の重ね書きを行える空間光変調器が必要である．今後，この空間光変調器の開発が課題である．

本論文で提案した光学系の自己調整機能は，収差による像のゆがみや変形に対しても問題なく処理される．しかし，像のぼけ等の画像の分解能を低下させる収差については，提案した光学系の自己調整機能では対処できない．今後，本論文で提案する光ニューラルネットワークに適したレンズ設計を行う必要がある．

謝辞

本研究は，大阪大学大学院工学研究科応用物理学専攻博士後期課程において，一岡芳樹教授の指導の下の行ったものである．終わりに臨み，終始懇切なる御指導と御助言を賜りました一岡教授に深く感謝の意を表すと共に厚くお礼申し上げます．

本学工学部の樹下行三教授，ならびに本学産業科学研究所の豊田順一教授には，本論分作成にあたり御検討頂き，貴重な御意見をいただきました．ここに深くお礼申し上げます．

研究の遂行，論文作成等，常に親身なる御指導，温かい御助言を賜りました本学工学部伊東一良助教授に深く感謝いたします．

一岡研究室の講師谷田純博士には，常に親切な御指導，貴重な御指摘を頂きました．心より厚く感謝いたします．また，本研究室助手の井上卓氏には，いつも温かい御助言を頂きました．心からお礼申し上げます．

本研究室の卒業生の，野村孝徳博士（神戸大学工学部助手），張偉博士（R2 Technology Inc., Manager），浅野晃博士（九州工業大学情報工学部助手），長谷川玲博士（ジョージタウン大学メディカルセンター研究員），谷口正樹博士（大阪工業研究所研究員），岩田昌也博士（電子総合研究所研究員），吉森久博士（古野電気株式会社主任），宮崎大介博士（大阪市立大学工学部助手）には，良き先輩として常に温かい御助言を頂きました．ここに厚くお礼申し上げます．

６年間にわたる研究生活において，良き同輩として励まして頂いたフランシスコ・レネロ氏には心から感謝いたします．本研究室の藤井勇作氏には，第５章について多大なご協力を頂いただけでなく，論文作成の全般にあたり御協力頂きました．深く感謝いたします．第３章の数字データベースに関して研究室卒業生の兼清知之氏（NTT株式会社），柴田健二氏（本学科樹下研究室）には多大な恩恵を頂きました深く感謝いたします．論文作成にあたり，的場修氏，粟辻安浩氏，平井亜紀子氏，請園信博氏，黒川和範氏，椿原一志氏には，多大なる御協力を頂きました．各氏には心から感謝致します．

最後に，６年間の研究生活の中で，様々な形で御協力，御援助していただいた研究室内外の皆様に深く感謝いたします．

参考文献

1) 雨宮真人, 田中譲, コンピュータアーキテクチャ,知識工学講座7 (オーム社, 1988).

2) S. Mori, C.Y. Suen and K. Yamamoto, "Historical Review of OCR Research and Development," Proc. IEEE **80**, 1029-1058 (1992).

3) 尾上守夫, 編, 画像処理ハンドブック(昭晃堂, 東京, 1987).

4) 目視検査の自動化技術調査委員会, 編, 画像処理による目視検査の自動化事例集 (新技術コミュニケーションズ, 東京, 1991).

5) 舟久保登, パターン認識 (共立出版, 1991).

6) 長尾真, コンピュータのパターン認識 (東京大学出版会, 1985).

7) P.R. Cohen and E.A. Feigenbaum, ed., *The handbook of artificial intelligence,* Vol. **III** (William Kaufmann, Inc., Los Altos, 1982).

8) 谷内田正彦, ロボットビジョン,人工知能シリーズ11 (昭晃堂, 1990).

9) R. Bajcsy, "Active Perception," Proc. IEEE **76**, 996-1005 (1988).

10) P.J. Burt, "Smart Sensing within a Pyramid Vision Machine," Proc. IEEE **76**, 1006-1015 (1988).

11) A. Blake and A. Yuille, ed., *Active Vision* (MIT Press, 1992).

12) 合原一幸, ニューラルコンピュータ -脳と神経に学ぶ- (東京電気大学出版局, 1988).

13) 麻生英樹, ニューラルネットワーク情報処理 -コネクショニズム入門, あるいは柔らかな記号にむけて (産業図書, 1988).

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

15) R. Hecht-Nielsen, *Neurocomputing* (Adision-Wesley Publishing Company, Inc., 1990).

16) R. Beale and T. Jackson, *Neural Computing : An Introduction* (IOP Publishing Ltd, 1990).

17) F. Rosenblatt, "The Perceptron: A Probablistic Model for Information Storage and Organization in the Brain," Psychological Review **65**, 386-408 (1958).

18) M.L. Minsky and S.A. Papert, *Percesptrons* (MIT Press, 1969).

19) J.J. Hopfield, "Neural networks and physical syetems with emergent collective computational abilities," Proc. Natl. Acad. Sci. USA **79**, 2554-2558 (1982).

20) K. Fukushima, "Neocognitron: A Self-organizing Neural Network Model for a Machanism of Pattern Recognition Unaffected by Shift in Position," Biol. Cybern. **36**, 193-202 (1980).

21) D.L. Reilly, L.N. Cooper and C. Elbaum, "A Neural Model for Category Learning," Biol. Cybern. **45**, 35-41 (1982).

22) D.E. Rumelhart, J.L. McClelland and PDP Research Group, *Parallel Distributed Processing* (MIT Press, 1986).

23) D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning internal representations by error propagation," *Paralell distributed processing: Explorations in the microstructure of cognition*: Vol **1**. Foundations, ed. D. E. Rumelhart, J. L. McClelland and PDP Research

Group (MIT Press, 1986) 318-362.

24) Y. LeCun, B. Boser, J. S.Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," Neural Computation **1**, 541-551 (1989).

25) W. Zhang, K. Itoh, J. Tanida and Y. Ichioka, "Parallel distributed processing model with local space-invariant interconnections and its optical architecture," Appl. Opt. **29**, 4790-4797 (1990).

26) J.E. Hummel and I. Biederman, "Dynamic Binding in Neural Network for Shape Recognition," Psychological Review **99**, 480-517 (1992).

27) Y. Yamaguchi and H. Shimizu, "Pattern Recognition With Figure-Ground Separation by Generation of Coherent Oscillations," Neural Networks **7**, 49-63 (1994).

28) D.R. Hofstadter, "Walking up from the Boolean dream, or , subcognition as computation," *Metamagical Themas*, (Basic Books, 1985) 631-665.

29) R.A. Brooks, "Intelligence without representation," MIT Technical Report (1988).

30) R.A. Brooks, "Intelligence Without Reason," Proc. IJCAI-91 **1**, 569-595 (1991).

31) 辻三郎, "AIマップ-ロボットから見たAI," 人工知能学会誌 **8**, 404-409 (1993).

32) P.E. Agre and D. Chapman, "Pengi: An Implementation of a Theory of Activity," Proc. AAAI 268-273 (1987).

33) 北野宏明, "AIにおける科学革命," 人工知能学会誌 **8**, 744-751 (1993).

34) P. Smolensky, "Connectionist AI, Symbolic AI, and the Brain," Artificial Intelligence Review **1**, 95-109 (1987).

35) P. Wallich （訳　原田託，溝口文雄), "米国の人工知能最前線," 日経サイエンス 1992年2月号 98-110 (1992).

36) G.E. Hinton, ed., *Connectionist Symbol Processing* (MIT Press, 1991).

37) R.A. Brooks, "A Robust Layered Control System For A Mobile Robot," IEEE Journal of Robotics and Automation **RA-2**, 14-23 (1986).

38) R.A. Brooks, "A Robot that Walks; Emergent Behaviors from a Carefully Evolved Network," Neural Computation **1**, 253-262 (1989).

39) 北野宏明, "実世界における知能," 人工知能学会誌 **9**, 73-81 (1994).

40) 北野宏明, "超並列人工知能," 人工知能学会誌 **7**, 244-261 (1992).

41) P. Smolensky, "Information Processing in Dynamical Systems:Foundations of Harmony Theory," *Paralell distributed processing: Explorations in the microstructure of cognition*: Vol **1**. Foundations, ed. D. E. Rumelhart, J. L. McClelland and PDP Research Group (MIT Press, 1986) 194-281.

42) N.H. Farhat, D. Psaltis, A. Prata and E. Paek, "Optical implementation of Hopfield model," Appl. Opt. **24**, 1469-1475 (1985).

43) D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," Opt. Lett. **10**, 98-100 (1985).

44) A.D. Fisher, R.C. Fukuda and J.N. Lee, "Implementation of Adaptive Associative Optical Computing Elements," Proc. Soc. Photo-Opt. Instrum. Eng. **625**, 196-204 (1986).

45) A.D. Fisher, W.L. Lippincott and J.N. Lee, "Optical implementations of associative networks with versatile adaptive learning capabilities," Appl. Opt. **26**, 5039-5054 (1987).

46) K. Wagner and D. Psaltis, "Multilayer optical learning networks," Appl. Opt. **26**, 5061-5076 (1987).

47) D. Psaltis, D. Brady and K. Wagner, "Adaptive optical networks using photorefractive crystals," Appl. Opt. **27**, 1752-1759 (1988).

48) M. Ishikawa, N. Mukohzaka, H. Toyoda and Y. Suzuki, "Optical associatron: a simple model for associative memory," Appl. Opt. **28**, 291-301 (1989).

49) T. Lu, S. Wu, X. Xu and F.T.S. Yu, "Two-dimentional programmable optical neural network," Appl. Opt. **28**, 4908-4913 (1989).

50) H. Toyoda, N. Mukozawa, Y. Suzuki and M. Ishikawa, "Adaptive optical processing system with optical associative memory," Appl. Opt. **32**, 1354-1358 (1993).

51) C.-H. Wang and B.K. Jenkins, "Subtracting incoherent optical neuron model: analysis, experiment, and applications," Appl. Opt. **29**, 2171-2186 (1990).

52) Y. Qiao and D. Psaltis, "Local learning algorithm for optical neural networks," Appl. Opt. **31**, 3285-3288 (1992).

53) K. Wagner and T.M. Slagle, "Optical competitive learning with VLSI/liquid-crystal winner-take-all modulator," Appl. Opt. **32**, 1408-1435 (1993).

54) M.A. Neifeld and D. Psaltis, "Optical implementations of radial basis classifiers," Appl. Opt. **32**, 1370-1379 (1993).

55) H.-Y.S. Li, Y. Qiao and D. Psaltis, "Optical network for real-time face recognition," Appl. Opt. **32**, 5026-5035 (1993).

56) Y. Owechko, "Cascaded-grating holography for artificial neural networks," Appl. Opt. **32**, 1380-1398 (1993).

57) Y. Kuratomi, A. Takimoto, K. Akiyama and H. Ogawa, "Optical neural network using vector-feature extraction," Appl. Opt. **32**, 5750-5758 (1993).

58) S. Amari, N. Fujita and S. Shinomoto, "Four Types of Learning Curves," Neural Computation **4**, 605-618 (1992).

59) E.B. Baum and D. Haussler, "What Size Net Gives Valid Generalization?," Neural Computation **1**, 151-160 (1989).

60) 篠本滋, "学習の数理," 数理科学 AUGUST, 20-25 (1991).

61) 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 **J73-D-II**, 897-905 (1990).

62) A. Hasegawa, O. Matoba, K. Itoh and Y. Ichioka, "Learning Generalization by Validation Set," Jpn. J. Appl. Phys. Part I **31**, 2459-2462 (1992).

63) Y. Hirose, K. Yamashita and S. Hijiya, "Back-Propagation Algorithm Which Varies the Number of Hidden Units," Neural Networks **4**, 61-66 (1991).

64) C. Ji, R.R. Snapp and D. Psaltis, "Generalizing Smoothness Constraints from Discrete Samples," Neural Computation **2**, 188-197 (1990).

65) E. Levin, N. Tishby and S.A. Solla, "A Stastical Approach to Learning and Generalization in Layered Neural Netwroks," Proc. IEEE **78**, 1568-1574 (1990).

66) W. Zhang, A. Hasegawa, K. Itoh and Y. Ichioka, "Error Back Propagation With Minimum Entropy Weights: A Technique for Better Generalization of 2-D Shift-Invariant NNs," Proc. IJCNN **1**, 645-648 (1991).

67) 小川英光, 編, パターン認識・理解の新たな展開 -挑戦すべき課題-(電子情報通信学会, 1994).

68) 小川英光, "パターン認識・理解の諸問題," 信学誌**76**, 615-631 (1994).

69) 岩田彰, 當麻孝志, 松尾啓志, 鈴村宣夫, "大規模4層ニューラルネットワーク" Comb NET" ," 信学論 **J73-D-II**, 1261-1267 (1990).

70) 今井勝次, 郷原一寿, 内川嘉樹, "3層BPモデルを用いたノイズのある原画像に対する手書き文字認識," 信学論 **J75-D-II**, 314-323 (1992).

71) 中野幸一, 渡辺弥寿夫, 吉村慶之, "2段選択型ニューラルネットワークによる機械部品の認識," 信学論 **J75-D-II**, 1788-1797 (1992).

72) 津村徳道, "多段階認識ニューラルネットワーク -文字の切り出し・認識への応用-," 大阪大学工学研究科修士論文(1992).

73) 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ワーク" CombNET-II" ," 信学論 **J75-D-II**, 545-553 (1992).

74) R.A. Jacobs, M.I. Jordan, S.J. Nowlan and G.E. Hinton, "Adaptive Mixtures of Local Experts," Neural Computation **3**, 79-87 (1991).

75) M.I. Jordan and R.A. Jacobs, "Hierarchical Mixtures of Experts and the EM Algorithm," Neural Computation **6**, 181-214 (1994).

76) K. Nakano, M. Niizuma and T. Omori, "Model of Neural Visual System with Self-Organizing Cells," Biol. Cybern. **60**, 195-202 (1989).

77) J.J. Hopfield, "Neurons with graded responce have collective computational properties like those of two-state neurons," Proc. Natl. Acad. Sci. USA **81**, 3088-3092 (1984).

78) G.A. Tagliarini, J.F. Christ and E.W. Page, "Optimization Using Neural Networks," IEEE Trans. Computer **40**, 1347-1358 (1991).

79) J.J. Hopfield and D.W. Tank, ""Neural" Computation of Decisions in Optimaization Problems," Biol. Cybern. **52**, 141-152 (1985).

80) J.S. Albus, "A New Approach to Manipulator Control: The Cerebellar Model Articulation Controller (CMAC)," Journal of Dynamic Systems, Mesurement, and Control, Trans. of the ASME **97**, 220-227 (1975).

81) T.M. Martinetz, H.J. Litter and K.J. Schulten, "Three-Dimentional Neural Net for Learning Visuomotor Coordination of a Robot Arm," IEEE Trans. Neural Networks **1**, 131-136 (1990).

82) J.A. Walter and K.J. Schulten, "Implementation of Self-Organizing Neural Networks for

Visuo-Motor Control of an Industrial Robot," IEEE Trans. Neural Networks **4**, 86-95 (1993).

83) A. Waibel, T. Hanazawa, G. Hinton, K. Shikano and K. Lang, "Phoneme Recognition: Neural Networks vs. Hidden Markov Models," IEEE International Conference on Acoustics, Speech, and Signal Processing **1**, 107-110 (1988).

84) G. Cybenko, "Approximations by superpositions of a sigmoidal function," Math. Control, Signals, and Systems **2**, 303-314 (1989).

85) B. Irie and S. Miyake, "Capabilities of Three-layered Perceptrons," Proc. IEEE International Conference on Neural Networks **1**, 641-648 (1988).

86) C.L. Scofield, D.L. Reilly, C. Elbaum and L.N. Cooper, "Pattern Class Degeneracy in an Unrestricted Storage Density Memory," *Neural Information Processing Systems*, ed. D. Z. Anderson (American Institute of Physics, 1988) 674-682.

87) M.J. Hudak, "RCE classifiers: theory and practice," Cybern. Syst. **23**, 483-515 (1992).

88) M.J. Hudak, "RCE networks: an experimental investigation," IJCNN-91-Seattle: International Joint Conference on Neural Network **1**, 849-854 (1991).

89) 黒川忠由, "ニューラル・ネットワークとその開発支援ソフト," インターフェース **149**, 239-250 (1989).

90) R.P. Lippmann, "An Introduction to Computing with Neural Nets," IEEE ASSP MAGAZINE April, 4-22 (1987).

91) T. Poggio and F. Girosi, "Networks for Approximation and Learning," Proc. IEEE **78**, 1481-1497 (1990).

92) J. Moody and C. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units," Neural Computation **1**, 281-294 (1989).

93) D. Broomhead and D. Lowe, "Multivariable function interpolation and adaptive networks," Complex Systems **2**, 321-355 (1988).

94) J. Park and I.W. Sandberg, "Universal Approximation Using Radial-Basis-Function Networks," Neural Computation **3**, 246-257 (1991).

95) 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," 信学論 **J73-D-II**, 1173-1178 (1990).

96) A. Hasegawa, K. Shibata, K. Itoh, Y. Ichioka and K. Inamura, "Adapting-Size Neural Network for Character Recognition on X-ray Films", Proceeding of the International Workshop on Applications of Neural Networks to Telecommunications (Lawrence Erlbaum Associates, 1993) 139-146.

97) J.C. Platt, "Learning by Combining Memorization and Gradient Decent," *Advances In Neural Information Processing Systems* **3**, ed. R. P. Lippmann, J. E. Moody and D. S. Touretzky (Morgan Kaufmann, 1991) 714-720.

98) S. Lee and R.M. Kil, "A Gaussian Potential Function Network With Hierarchically Self-Organizing Learning," Neural Networks **4**, 207-224 (1991).

99) K.T. Blackwell, T.P. Vogl, S.D. Hyman, G.S. Barbour and D.L. Alkon, "A New

Approach to Hand-Written Character Recognition," Pattern Recognition **25**, 655-666 (1992).

100) J.A. Leonard, M.A. Kramer and L.H. Ungar, "A neural network architecture that computes its own reliability," Computers chem. Engng. **16**, 819-835 (1992).

101) T. Kohonen, "The "Neural" Phonetic Typewriter," IEEE Computer(magazine) March, 11-22 (1988).

102) T. Kohonen, "Self-Oraganized Formation of Topologically Correct Feature Maps," Biol. Cybern. **43**, 59-69 (1982).

103) T. Kohonen, "The Self-Oraganizing Map," Proc. IEEE **78**, 1464-1479 (1990).

104) A. Iwata, K. Hota, H. Matsuo, H. Sugumura, S. Matsuda and M. Yoshida, "A Large Scale Neural Network "CombNET" on a Neural Network Accelelator(Neuro-Turbo)," Proc. of the Int. Conf. on Fuzzy Logic & Neural Networks, July 329-333 (1990).

105) "複数のニューラル・ネット・モデルを統合した学習システム," ニューロコンピュータ最前線, (日経ＢＰ社, 1988) 102-113.

106) N. Tsumura, K. Itoh and Y. Ichioka, "Artificial neural network that can say "unknown"," Optical Review **1**,41-43 (1994).

107) N. Tsumura, K. Itoh and Y. Ichioka, "Reliable classification by double hyperspheres in pattern vector space," Pattern Recognition (submitted).

108) 池田光男, 目は何を見ているのか-視覚系の情報処理- (平凡社, 1988).

109) 津村徳道, 伊東一良, 一岡芳樹, "並列分散処理による能動的画像認識モデル：顔の認識," 光学**23**,750-757 (1994).

110) P.J. Burt, "Multiresolution techniques for image representation, analysis, and 'smart' transmission," Proc. Soc. Photo-Opt. Instrum. Eng. **1199**, 2-15 (1989).

111) 長尾真, 知識と推論 ,岩波講座ソフトウエア科学14 (岩波書店, 1988).

112) G.E. Hinton, "Mapping Part-Whole Hierarchies into Connectionist Networks," *Connectionist Symbol Processing*, ed. G. E. Hinton (MIT Press, 1991) 47-75.

113) N. Tsumura, Y. Fujii, K. Itoh and Y. Ichioka, "Pattern association by a lateral-alignment-free optical neural network," Optik (to be published).

114) N. Tsumura, Y. Fujii, K. Itoh and Y. Ichioka, "Optical method for generalized Hebbian-rule in optical neural network," Proc. International Joint Conference on Neural Network **1**, 833-836 (1993).

115) D.J. Willshaw, D.P. Buneman and H.C. Longuet-Higgins, "Non-Holographic Associative Memory," Nature **222**, 960-962 (1969).

116) S.-I. Amari, "Characteristic of Sparsely Encoded Associative Memory," Neural Networks **2**, 451-457 (1989).

117) J. Austin, "ADAM: A Distributed associative memory for scene analysis," Proceeding IEEE First International Conference on Neural Networks **4**, (1987).

118) K. Kubota, Y. Tashiro, K. Kasahara and S. Kawai, "Optical Crossbar Interconnection Using Vertical-to-Surface Transmision Electro-Photonic Devices (VSTEP)," Proceedings

of the Society of Photo-Optical Instrumentation Engineers (Optical Computing 88) **963**, 255-259 (1988).

119) M. Born and E. Wolf, Principles of Optics - sixth edition (Pergamon Press, 1980) 381.

120) K.M. Johnson and G. Moddel, "Motivations for using ferroelectric liquid crystal spatial light modulators in neurocomputing," Appl. Opt. **28**, 4888-4899 (1989).

121) T.M. Slagle and K. Wagner, "Winner-take-all spatial light modulator," Opt. Lett. **17**, 1164-1166 (1992).

122) S. Yamamoto, R. Sekura, J. Yamanaka, T. Ebihara, N. Kato and H. Hoshi, "Optical Pattern Recognition with LAPS-SLM (1) / Light Addressed Photoconductor and Smectic C* Liquid Crytal Spatial Light Modulator," Procceding SPIE Computer and Optically Formed Holographic Optics **1211**, 273-283 (1990).

123) K. Matsushita and M. Toyama, "Unevenness of illuminance caused by gradient-index fiber arrays," Appl. Opt. **19**, 1070-1075 (1980).

124) H. Toyoda and M. Ishikawa, "Sparse Encoding Algorithm for Optical Associative Memory Using Bistable Spatial Light Modulator," Proceeding of the 12th International Display Research Conference,October 12-14, 371-374 (1992).

125) D. Miyazaki, J. Tanida and Y. Ichioka, "Reflective block optics for optical computing systems," Optical Computing, 1993, Techinical Digest Series, (Optical Society of America, 1993) 30-33.

126) K. Hamanaka, "Optical bus interconnection system using Selfoc lenses," Opt. Lett. **16**, 1222-1224 (1991).

127) M. Mishkin, L.G. Ungerleider and K.A. Macko, "Object vision and spatial vision: two cortical pathways," Trends In Neurosciences **6**, 414-417 (1983).

著者発表論文

学術論文誌

1) 津村徳道, 山本高志, 大西邦一, 横関俊介, "非球面に対する干渉図形の解析法," 精密工学会誌 **57**, 74-79 (1991).

2) N. Tsumura, K. Itoh and Y. Ichioka, "Artificial neural network that can say "unknown"," Optical Review **1**, 41-43 (1994).

3) 津村徳道, 伊東一良, 一岡芳樹, "並列分散処理による能動的画像認識モデル：顔の認識," 光学 **23**, 750-757(1994).

4) N. Tsumura, Y. Fujii, K. Itoh and Y. Ichioka, "Pattern association by a lateral-alignment-free optical neural network," Optik (in press).

5) N. Tsumura, K. Itoh and Y. Ichioka, "Reliable classification by double hyperspheres in pattern vector space," Pattern Recognition, (submitted).

会議抄録

6) N. Tsumura, T. Kanekiyo, K. Itoh, Y. Ichioka and K. Inamura, "A Neural-Network Approach for Recognition of ID numbers in X-Ray Films," Medical Imaging Technology **10**, 317-318 (1992).[in Japanese]

7) N. Tsumura, Y. Fujii, K. Itoh and Y. Ichioka, "Optical method for generalized Hebbian-rule in optical neural network," Proc. International Joint Conference on Neural Network **1**, 833-836 (Nagoya, 1993).

8) N. Tsumura, K. Itoh and Y. Ichioka,"Artificial Neural Network that can say 'Unknown'," Digest of Topical Meeting of International Commission for Optics, 310 (Kyoto, 1994).

学会発表等

9)津村徳道，横関俊介，一岡芳樹，”非球面に対する干渉図形の解析法，”第37回応用物理学関係連合講演会予稿集, 739(埼玉,1990).

10)津村徳道，新井泰彦，山田朝治，横関俊介，”光学的面に対する干渉図形の解析法に関する基礎研究，”レーザ研シンポジウム1991, (大阪大学レーザ核融合研究センタ,1991).

11)津村徳道，伊東一良，一岡芳樹，”ニューラルネットワーク・フュージョン（１）－文字の切り出し・識別への応用－，”第39回応用物理学関係連合講演会予稿集, 820(千葉,1992).

12)津村徳道，伊東一良，一岡芳樹，”アライメントフリー光ニューラルネットワーク，”光学連合シンポジウム京都'92 講演予稿集, 175-176(京都,1992).

13)津村徳道，伊東一良，一岡芳樹，”ニューラルネットワーク・フュージョン（２）－組み合わせパターンの認識－，”第40回応用物理学関係連合講演会予稿集, 932(東京,1993).

14)津村徳道，藤井勇作，伊東一良，一岡芳樹，”アライメントフリー光ニューラルネットワーク（２）-原理確認実験-，”光学連合シンポジウム旭川'93講演予稿集, 179-180(旭川,1993).

15)津村徳道，伊東一良，一岡芳樹，”より高い拒絶能力をもつ超球分類器”第41回応用物理学関係連合講演会予稿集, 877(東京,1994).

16)津村徳道，藤井勇作，伊東一良，一岡芳樹，”並列分散処理による能動的画像認識モデル：顔の認識”光学連合シンポジウム浜松'94 講演予稿集, 197-198(浜松,1994).