

修士学位論文

仮定と検証に注目した  
視覚情報処理系のモデル化に関する研究

東北大学大学院工学研究科  
電気・通信工学専攻 阿曾研究室

下村 正夫

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 本研究の背景	1
1.2 本研究の目的	3
1.3 本研究の方針	3
1.4 本論文の構成	4
<b>第2章 視覚情報処理系の生理学と心理学</b>	<b>5</b>
2.1 視覚情報の伝達経路	5
2.2 Hubel-Wieselの階層型仮説	6
2.3 大脳内における視覚解析の2つの経路	9
2.4 注視と眼球運動	11
2.5 認識活動における仮定と検証	15
2.6 多解像度を用いた視覚情報処理	15
2.7 まとめ	16
<b>第3章 仮定と検証に基づく位置・変形に頑健なパターン認識モデル</b>	<b>17</b>
3.1 ネオコグニトロン	17
3.1.1 ネオコグニトロンの構造	17
3.1.2 ネオコグニトロンの定式化	18
3.1.3 ネオコグニトロンの学習則	21
3.2 仮定と検証に基づく位置・変形に頑健なパターン認識モデル	22
3.2.1 モデルの概要	22
3.2.2 モデルの構成	24
3.2.3 学習方法	30
3.2.4 認識方法	32
3.3 計算機シミュレーション	36
3.3.1 直交性の高いパターンでの評価実験	36
3.3.2 類似した特徴を持つパターンでの評価実験	39

---

3.4	まとめ	40
<b>第4章</b>	<b>多解像度を利用したパターン抽出モデル</b>	<b>43</b>
4.1	モデルの概要	43
4.2	モデルの構成	44
4.3	学習方法	48
4.4	認識方法	48
4.5	計算機シミュレーション	50
4.5.1	未知パターンを含まない場合の評価実験	51
4.5.2	未知パターンを含む場合の評価実験	55
4.6	まとめ	57
<b>第5章</b>	<b>結論</b>	<b>58</b>
5.1	本研究の成果	58
5.2	今後の課題	59
	<b>参考文献</b>	<b>60</b>

## 目 次

1.1	一般的なパターン認識系の構成；(文献 [3] より引用) . . . . .	2
2.1	網膜から一次視覚野までの視覚情報伝達経路 (文献 [18] より抜粋) . . . . .	6
2.2	Brodmann の脳地図 (文献 [18] より抜粋) . . . . .	7
2.3	単純型細胞の特性 . . . . .	8
2.4	複雑型細胞の特性 . . . . .	8
2.5	単純型細胞と複雑型細胞の階層仮説構造 . . . . .	9
2.6	ヒトの脳内における視覚情報処理の 2 経路 (文献 [22] より抜粋) . . . . .	10
2.7	アカゲザルの視覚情報処理の部位と流れ (文献 [22] より抜粋) . . . . .	11
2.8	図形を単純化し刺激となっている特徴を抽出する過程 (文献 [22] より抜粋) . . . . .	12
2.9	「対象に基づく選択」の実証実験で用いられた刺激例 (文献 [25] より抜粋) . . . . .	13
2.10	特徴統合理論の概念図 (文献 [25] より抜粋) . . . . .	14
3.1	ネオコグニトンの概念図 . . . . .	18
3.2	C 細胞層-S 細胞層間の結合関係 . . . . .	19
3.3	S 細胞層-C 細胞層間の結合関係 . . . . .	20
3.4	位置・変形に頑健なパターン認識モデルの概要図 . . . . .	22
3.5	位置・変形に頑健なパターン認識モデルの構造 . . . . .	24
3.6	網膜モデルにおける 2 次元座標系 . . . . .	25
3.7	特徴抽出部の反応例 . . . . .	26
3.8	特徴種類記憶部の構造 . . . . .	27
3.9	認識部の構造 . . . . .	28
3.10	特徴位置記憶部の構造 . . . . .	30
3.11	学習パターンセットの例 . . . . .	33
3.12	提示パターンと初期切り出し位置 . . . . .	34
3.13	切り出し位置補正量の算出 . . . . .	35
3.14	切り出し位置の修正 . . . . .	36
3.15	直交性の高い学習用パターン . . . . .	38

---

3.16	直交性の高いパターンにおける認識の過程	38
3.17	認識の過程における $U_{C3}$ 層の反応	39
3.18	認識可能な初期切り出し範囲	39
3.19	パターン混在時における認識の過程	40
3.20	類似した特徴を持つ学習用パターン	40
3.21	類似した特徴を持つパターンにおける認識の過程	41
3.22	類似した特徴を持つパターンの認識可能な初期切り出し範囲	42
4.1	パターン抽出における2つのフェーズ	44
4.2	パターン抽出モデルの概要	45
4.3	網膜モデルにおける抑制マスクの働き	46
4.4	発火重心位置の計算における問題点	47
4.5	細胞面内における競合後の発火重心位置	48
4.6	抑制定数と細胞の反応	49
4.7	学習用パターン	50
4.8	学習パターンのみから構成されるテストパターンの一例	51
4.9	パターン抽出成功例 (学習パターンのみの場合)	53
4.10	パターン抽出失敗例 (学習パターンのみの場合)	54
4.11	混入した未知パターン	55
4.12	未知パターンを含むテストパターンの一例	55
4.13	パターン抽出成功例 (未知パターンを含む場合)	56
4.14	パターン抽出失敗例 (未知パターンを含む場合)	57

---

## 表 目 次

3.1	ネオコグニトロンのパラメータ . . . . .	26
3.2	網膜モデルのパラメータ . . . . .	37
3.3	特徴抽出部のパラメータ . . . . .	37
3.4	認識部に関連するパラメータ . . . . .	37
4.1	網膜モデルのパラメータ . . . . .	51
4.2	特徴抽出部のパラメータ . . . . .	52
4.3	認識部に関連するパラメータ . . . . .	52
4.4	未知パターンを含まない場合の抽出率と平均抽出ステップ数 . . . . .	55
4.5	未知パターンを含む場合の抽出率と平均抽出ステップ数 . . . . .	57

# 第1章

## 序論

### 1.1 本研究の背景

ヒトを初めとする諸動物が行う知的情報処理として重要なものに、外界から得た情報を基にして外界の様相を知覚・認識するというパターン認識という行為が挙げられる [1]。このパターン認識という行為を通じて、我々は脳内に作り上げた「外界」と接しながら生きている [2]。すなわち、ヒトの機能を代替可能な人工物を作り上げる際に、パターン認識は欠かせないものである。そのため、従来の機械式計算機に比べ飛躍的に複雑な計算を可能とした電子計算機が 1940 年代に登場して以来、パターン認識の能力を電子計算機に持たせようと幾多の研究が行われてきた。だが、研究の発展に伴い、諸動物が当たり前の事として行っているこの行為が見かけほど簡単には実装できないということも明らかになってきた。しかしながらパターン認識が持つ高い実用価値とヒトの知的情報処理を機械で実現するという知的好奇心に支えられて、パターン認識に関する研究は今なお活発に行われている。

文字認識に代表される工学分野におけるパターン認識の研究は、1950 年代から 60 年代にかけて隆盛を極め、基本的な枠組みが築かれた。その後、1980 年代以降にかけて、処理を行う計算機の飛躍的な速度・記録容量の向上によって実用段階を迎えるに至った [3]。

しかし、図 1.1 で示されるような現在実用化されている純粋に工学的なアプローチによるパターン認識の機構は、研究者の認識対象に対する知識を想定した認識環境に合わせてこれをアルゴリズムに置き換え認識を実現しているため、限られた環境下でのみ有効に機能する。すなわち、環境や認識対象の変化に対して柔軟に対応可能な物とはなっていない。

これに対して、ヒトを初めとする諸動物が行っているパターン認識の能力は非常に高度であり、かつ柔軟性に富む。例えば、認識対象を 2 次元パターンに限定しても、パターンの変化は変形・位置ずれ・拡大縮小・回転・ノイズ混入・テクスチャの違いなど様々なものが考えられ

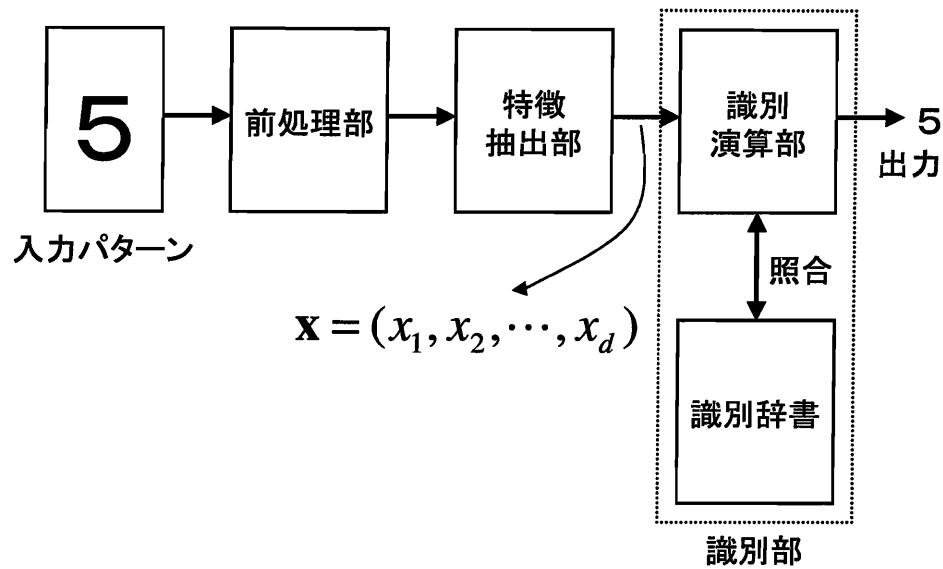


図 1.1: 一般的なパターン認識系の構成；(文献 [3] より引用)

パターンが入力されるとまず前処理部でノイズ除去、正規化などを行う。続いて特徴抽出部で膨大な特徴を持つ原パターンから認識に必要な本質的な特徴のみを抽出する。この特徴を元に識別部では識別処理を行う。識別処理は入力パターンに対して複数のクラスのうち一つを対応させる事によって行う。そのため、前もって識別辞書を用意し、抽出された特徴をこの辞書と照合することにより入力パターンの所属するクラスを出力する。

るが、ヒトはこれらに柔軟に対応している。さらに3次元パターンでは上記の変化に加えて見る距離や方向によって“同じ”パターンが全く違って見える場合があるなどの問題も生じるが、それらに影響されることなく瞬時に認識を行っている。

このような高度な知的情報処理を実現するためには、「その基本原理を脳に求めることで処理機構を明らかにする必要がある」というアプローチが1940年代に登場した。現在では、神経生理学・情報数理学・心理学などの幅広い分野の研究者によって支持されており、学際的な研究が活発に行われている。この中で、計算論的神経科学や神経計算学と呼ばれる分野が最近注目されるようになった。これはMcCullochとPitts[4]による形式ニューロンモデルに端を発するもので、情報数理学の立場からヒトの知的情報処理の解明を目指すものである。具体的には、神経回路網の数理モデルを構築し、モデルの動作を調べることで脳の機構を予見すると共に、工学的にも有用な新しい知的情報処理方式の開発を目的としている。

このアプローチをパターン認識の研究に適用することで、“汎用認識器”、すなわち、環境の変化にも柔軟に対応する視覚情報処理機構の構築が期待される。



## 1.2 本研究の目的

パターン認識を目的とした視覚情報処理系の数理モデルとしては、過去にもいくつかの研究がなされてきた。代表的なモデルとしては、Fukushimaによって提案された階層型神経回路モデルで、実際に文字認識に対して有効性が示されている Neocognitron[5]や、Basakによって提案された心理学的な考察を元にパターン同士の重なりを解決し、複数物体の同時認識を可能とした Psycop[6]などが挙げられる。

しかし、いずれのモデルにおいても、認識対象はすでに切り出された個別のパターンであり、かつパターンの全体を入力したときにのみ有効に動作するモデルであった。ところが、ヒトの視覚情報処理系においては、視界内に入る範囲は外界のごく一部であり、実際にどこにパターンが存在しているのかは認識を通じて把握している。そして視界内に混在したパターン中で重要だと思われる場所へ視点を次々移す行動をとることで外界を把握している。この機構は注視と呼ばれており、効率的な視覚情報処理を実現するために重要な役割を担っている事が知られている [7][8][9]。

また、注視を考慮した視覚神経系のモデルに関する研究は無数に行われている [10][11][12][13]が、視界外に及ぶ注視点の移動とパターン認識を一貫して実行可能なモデルは未だ提案されていない。

そこで本研究では、注視の機構をモデルに組み込む事により、以下のようなパターン認識を可能とする視覚情報処理系のモデルを構築することを目的とする。

- 種々のパターンが混在した広い範囲を視点を移しながら認識。
- 視界からはみ出しパターンの一部しか見えていない場合でも視点を移すことでパターンの位置ずれに対応。
- パターンの変形・ノイズの混入に対する頑健性も保持。

視界範囲に限定されず、認識対象の変化に柔軟に対応するこのような認識機構は、近年メカトロニクス分野で発展を遂げている動物型ロボットのビジョンシステムを構築する際にも重要な役割を果たすものである。

## 1.3 本研究の方針

本研究では上記のような認識機構を構築するにあたって、心理学的に、ヒトの認知活動で利用していることが指摘されている「仮定と検証」という動作を基礎に置いて、研究を進めてゆく。この動作は、まず限定された感覚入力から環境や知識などに基づき認識対象についての仮定が生成され、その仮定から導かれる結論を検証してゆくことで情報処理を進めてゆくものである [14][15]。

この「仮定と検証」という動作を、後述する視覚情報の特徴統合時における注意 [17] と結びつけることで、たとえ認識開始時には入力情報が不完全な場合であっても、仮定を生成し注視する場所を変化させながら検証するという動作によって入力情報を統合し、最適な認識結果を得ることが期待できる。実際に過去の研究では、回転したパターンの認識モデルにおいて「仮定と検証」の有効性が実証されている [16]。

また、神経生理学や情報数理学、心理学などで得られている様々な知見を考慮し、数理モデルを構築する。得られたモデルは計算機上に実装し、種々のパターンにおいて適応するかを実験し、有効性を確認する。

なお、本研究は脳で行っている情報処理と同等の機能を工学的に実現することが目標であり、脳内の各部位における細胞数や部位間の結合など脳の構造をモデルで忠実に再現することが目標ではない。

## 1.4 本論文の構成

本論文の構成は以下の通りである。

**第1章** 序論であり、研究の背景及び目的を述べた。

**第2章** 視覚情報処理系の数理モデルを構築する際に基盤となるヒトの視覚情報処理における生理学的・心理学的事実について述べる。

**第3章** パターンの一部のみ網膜モデルに入力されている場合でもパターンを正しく切り出し・認識を行うことのできる視覚情報処理系の数理モデルを提案する。また、計算機シミュレーションにより評価・考察を行う。

**第4章** 第3章のモデルを元に、複数パターン提示時の注視切り替えが可能となるモデルを提案し、評価・考察を行う。

**第5章** 結論を記す。

## 第2章

# 視覚情報処理系の生理学と心理学

この章では、視覚情報処理系の数理モデルを構築する際に基盤となるヒトの視覚情報処理における生理学的・心理学的な解明事項について述べる。

### 2.1 視覚情報の伝達経路

「見る」という情報処理は、簡単に説明すると図 2.1 のような流れでおこなわれる。まず、目で見ている情報は、光の刺激として目の網膜 (retina) に入る。ヒトの網膜には、1 億数千万もの視細胞 (visual cell; photoreceptor cell) が並んでおり、光はそこに当たって電気信号へと変換される。その情報は視神経を通過して、視神経交叉 (optic chiasma) と呼ばれる部分で右目からの情報と左目からの情報が半分ずつ分かれる。そして、右目で見えた右半分の視野の情報と左目で見えた右半分の視野の情報が左脳へ、同様に、左右両眼の左視野の情報が右脳へと流れてゆく。両眼の各視野の統合された情報は外側膝状体 (lateral geniculate body=LGB;lateral geniculate nucleus=LGN) で中継され、最後に、大脳新皮質の後頭葉、視覚野と呼ばれるところに到達する。この時、外側膝状体の中継細胞の軸索は束にはならず、放射状に発散して視覚野に向かうので、これを視放線 (optic radiation; visual radiation) とよぶ。その後、大脳視覚野で様々な情報処理が行われてゆく。

大脳視覚野は左右の大脳半球の後部 (後頭葉) に存在し、Brodmann の脳地図 (図 2.2) の 17 野が第 1 次視覚野 (visual area I = V1) に対応している。第 1 次視覚野で処理された情報は、その前方に位置する 18 野及び 19 野の視覚前野に送られ、物体の認識処理が行われる。視覚前野は、機能的に異なった多数の領野に分けられ、それぞれ V2(visual area II; 第 2 次視覚野) V3,V3A,V4,MT(middle temporal area),MST(medial superior temporal area) などと名付けられている。

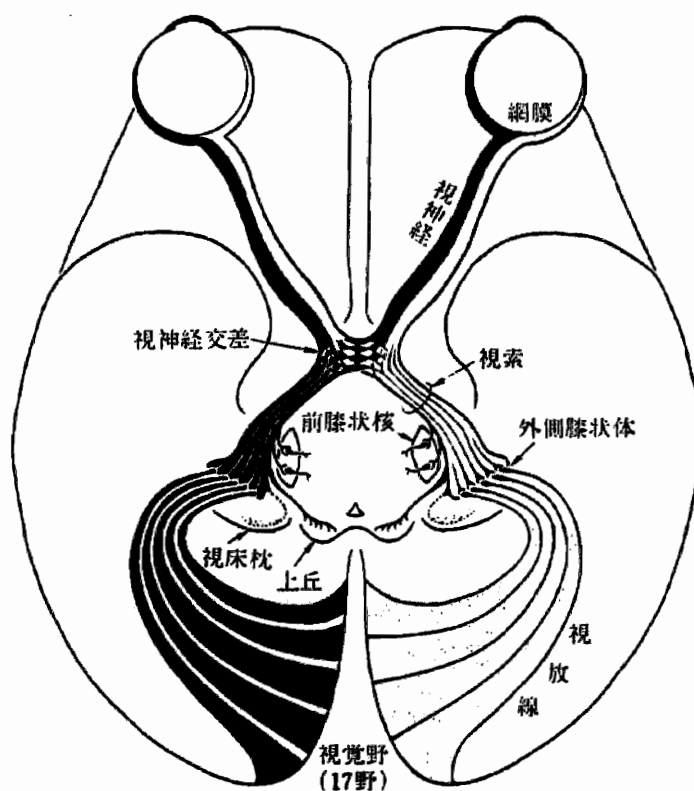


図 2.1: 網膜から一次視覚野までの視覚情報伝達経路 (文献 [18] より抜粋)

なお、視覚に関連した領域は、後頭葉だけでなく下部側頭葉 (inferotemporal cortex; 側頭連合野の下部に位置する) の後半部の PIT 野 (posterior inferotemporal area) や、前半部の AIT 野 (anterior inferotemporal area), 頭頂葉野 (頭頂連合野) の 7a 野, さらには前頭葉などにも及ぶことが近年解明されている。PIT 野と AIT 野とをまとめて IT 野 (inferior temporal area) と呼ぶこともある。

## 2.2 Hubel-Wiesel の階層型仮説

Hubel と Wiesel は、1960 年代にネコやサル の視覚野 (V1, V2) の神経細胞の受容野の性質を調べ [19][20]、これらの細胞を受容野の性質に基づき単純型細胞 (simple cell)、複雑型細胞 (complex cell) などに分類し、これらの細胞間には、単純型細胞 → 複雑型細胞 → 超複雑型細胞、という階層的構造が存在するという仮説を提唱した。

ここでは、特に興味深い性質を示す単純型細胞と複雑型細胞の階層的構造について説明する。

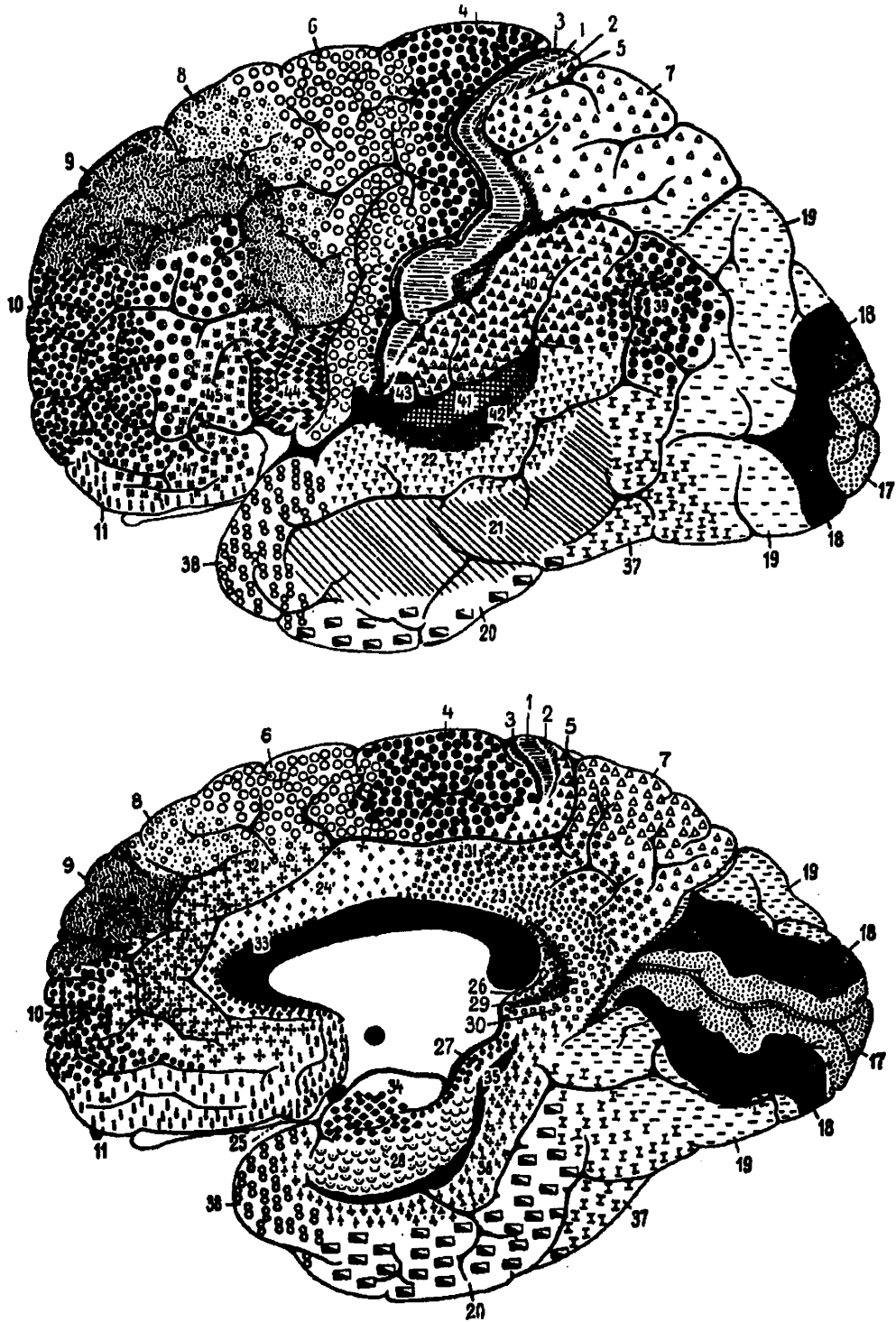


図 2.2: Brodmann の脳地図 (文献 [18] より抜粋)

上図はヒトの脳の左半球を外側から見た図, 下図は右半球を内側から見た図. 図の左側が前, 右側が後.

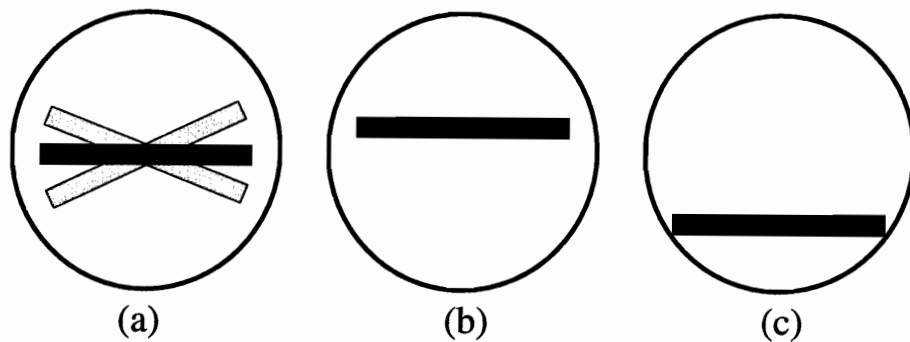


図 2.3: 単純型細胞の特性  
細胞の受容野は円で示される領域である。

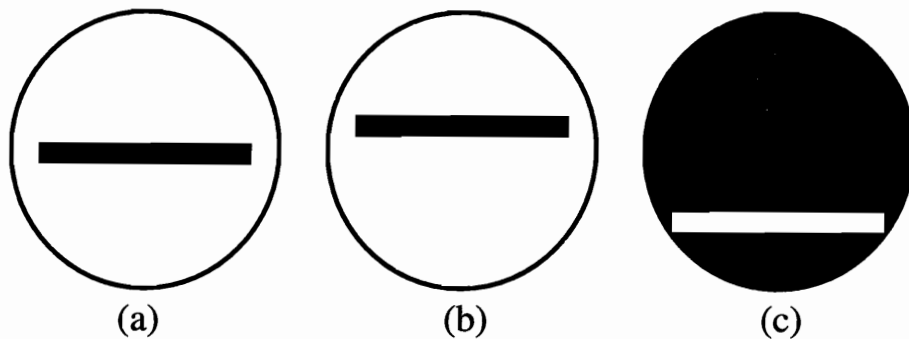


図 2.4: 複雑型細胞の特性  
細胞の受容野は円で示される領域である。

**単純型細胞** 単純型細胞は方位選択性を持つ神経細胞で、ある特定の方位を持つ線分のみに対応する性質を持ち、その方位は各々の細胞で異なる。また、単純型細胞などの視覚に関連する神経細胞は、網膜上のある一部の領域に存在する神経細胞の出力に依存して反応しており、この網膜上の領域を受容野という。例えば、図 2.3(a) の黒い線分で示される方位に対して選択的に反応する単純型細胞は、灰色の線分で示される方位に対しては反応しない。また、図 2.3(b) で示される線分に反応する単純型細胞は、少し位置がずれた図 2.3(c) で示される線分には反応しない。また、方位と位置が同じでも、位相がずれて入力されているものに対しては反応しない。

**複雑型細胞** 単純型細胞では、刺激として与える線分の位置や位相が少しでもずれると出力は抑制されてしまう。これに対して複雑型細胞では、受容野の内部に特定の方位を持つ線分が入ってくるとその位置や位相にはあまり影響されず反応する性質を持つ。例えば、図 2.4(a) で示される線分に反応する複雑型細胞は、図 2.4(b), 図 2.4(c) で示される線分に対しても反応する。

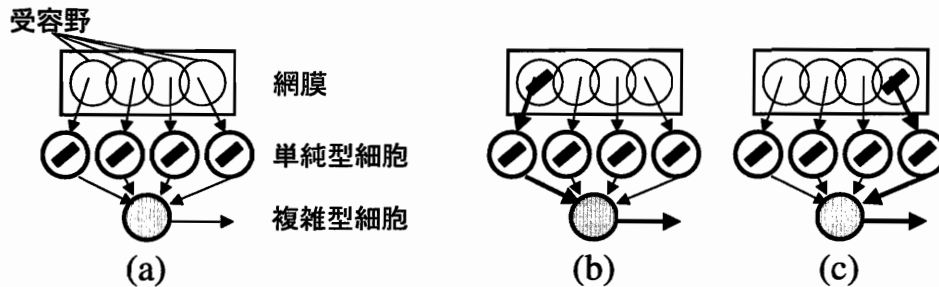


図 2.5: 単純型細胞と複雑型細胞の階層仮説構造

Hubel と Wiesel は、単純型細胞と複雑型細胞が図 2.5(a) で示すような簡単な階層構造を形成していると仮定した。複雑型細胞は、方位選択性は同じであるが受容野が互いに異なる複数の単純型細胞から入力を受けているために上記のような特性を持っている。例えば、図 2.5(a) では複雑型細胞はある方位選択性を持つ 4 つの単純型細胞から入力を受けている。ここで、図 2.5(b) のように網膜の左側に同じ方位の線分が提示されても、図 2.5(c) のように網膜の右側に同じ方位の線分が提示されても、この複雑型細胞は反応する。また、異なる方位を持つ線分が提示された場合には単純型細胞は発火しないため、この複雑型細胞も発火しない。これは、複雑型細胞の受容野の特性とよく適合するものであり、図 2.5 中の複雑型細胞は横長の受容野を持つということになる。位相に関する特性についても、同様の階層型仮説から説明が可能である。

## 2.3 大脳内における視覚解析の 2 つの経路

第 1 次視覚野において解析された情報は、視覚連合野に送られて物体の知覚や視野全体の認知が行われる。ヒトやアカゲザルなどの解剖結果から、Ungerleider らは第 1 次視覚野から視覚連合野に続く情報処理の経路が二つあることを発見した [21]。図 2.6, 図 2.7 で示すように、一つの経路は視覚前野 (extrastriate cortex) から始まり後頭葉 (inferior temporal lobe) で終端する腹側経路 (ventral stream) であり、主に物体形状の知覚に関わっている。もう一つの経路は視覚前野から始まり下側頭葉 (posterior parietal lobe) で終端する背側経路 (dorsal stream) であり、主に物体位置の知覚に関わっている [22]。

視覚前野は、第 1 次視覚野の神経細胞が次に情報を送っている部位であり、複数の領域から構成され、その各領域は方位・動き・空間周波数・視差など特定の特徴に反応する神経細胞から構成されている。また、第 1 次視覚野を先頭とする階層構造をなしていることがわかっているが、詳しい機能については未だ不明なところが多い。

物体形状の知覚を行う腹側経路では、側頭葉の腹側にある下側皮質 (inferior temporal cortex) で形状の解析と色情報の統合が行われ、物体が知覚される。下側皮質は TE 野と TEO 野の 2

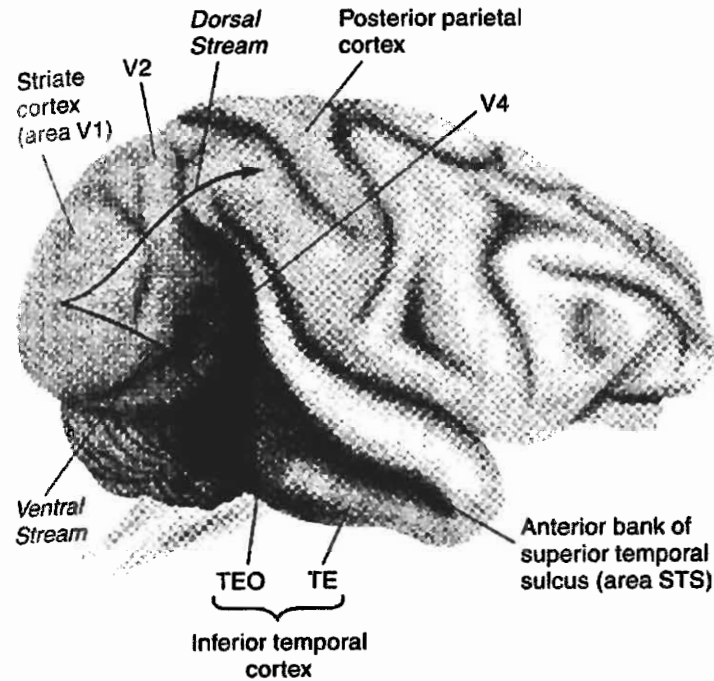


図 2.6: ヒトの脳内における視覚情報処理の2経路 (文献 [22] より抜粋)

つの領域で形成され、これらの領域が破壊されると物体の判別ができなくなる。このうち、TE野は視野の半分にも及ぶ最も大きな受容野を持ち、その神経細胞は3次元物体の提示に対して最もよく反応する。TanakaらによるTE野中の神経細胞の反応特性解析 [23] によると、TE野には図2.8で示すような図形特徴に対応した神経細胞群があることが解明されている。

物体位置の知覚を行う背側経路では、視覚前野の全ての部位からの出力を頭頂葉を經由して下側頭葉で受け取り、空間位置の知覚を行っている。このことは、頭頂後頭領に損傷を受けると発症するBalint症候群などから解明されており、その中でも特にSimultanagnosiaと呼ばれる物体が1つの時は認知できるが複数の物体になると知覚できない現象からは、別々の物体の認識は、同じ視野内にあった場合でも独立に実行されている事が示唆されている。また、背側経路は、眼球運動の制御・物体に手を届かせる動き・指や手の平で物体を捕むといった働きを持つ前頭葉と広範囲に渡って結合していることから、Goodaleらによって背側経路の主要な機能は位置の知覚ではなく行動を導く事ではないかということが提唱されている [24]。

このように、物体形状の知覚と物体位置の知覚は脳内の異なる経路で処理されている。



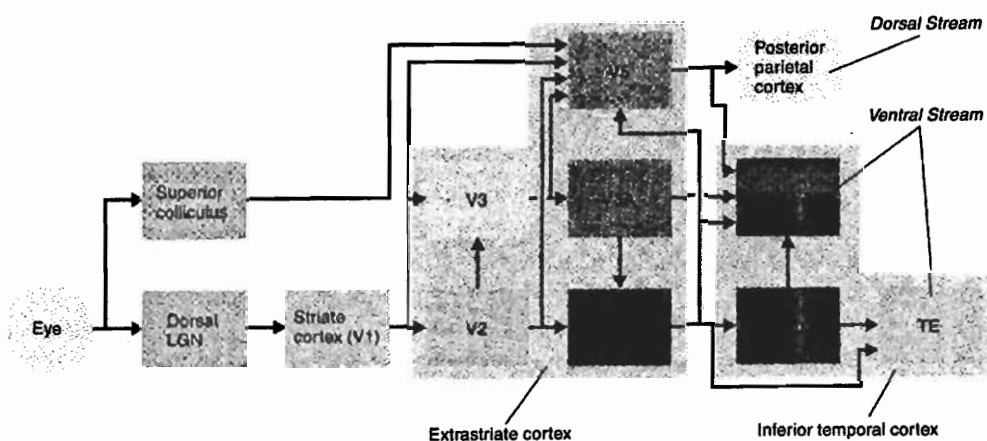


図 2.7: アカゲザルの視覚情報処理の部位と流れ (文献 [22] より抜粋)  
視覚情報処理系のうち特に重要な経路と部位のみを图示している。

## 2.4 注視と眼球運動

ヒトの視覚を考えた場合、網膜に映っている外界は一部分に過ぎず、広い範囲を知覚するためには頭や眼球を動かして見える範囲を移さなくてはならない。また、網膜上に入力された情報においても、低次の部分では並列処理を行っているが高次の処理では一度に可能な処理量は限られている。すなわち、視覚系には空間的な並列処理の限界量があり、その限界範囲内において必要な情報を効率的に選択し、処理する機能が必要となる。この情報選択の機構は視覚的注意や注視 (Visual Attention) と呼ばれる。

注視における視覚情報選択の方法としては、大きく分けて「空間に基づく選択」と「対象に基づく選択」が報告されている [25]。「空間に基づく選択」とは、目標刺激が提示される前にあらかじめ目標提示位置に注意を向けることによって刺激への反応時間が短くなる、という実験から明らかになった選択方法で、空間上において注意をあたかもスポットライトのようにずらしながら認識を進めてゆく方法である。一方で、「対象に基づく選択」とは、空間位置に対して注意を向けるのではなく、認識対象に対して注意を向ける方法である。例えば、図 2.9 で示されるような空間的に重なった位置に 1 つの四角形と 1 本の線分を提示し、そのうちの 2 つの属性を報告するという課題を課した結果、1 つの対象 (例えば線分) に関する 2 つの属性 (線分の傾きと種類) を報告する条件の方が 2 つの対象 (四角形と線分) に関する 2 つの属性 (長方形の縦の長さや線分の傾き) を報告する条件よりも正答率が高い、ということが実験で報告されている。

ところで、ヒトの視覚情報処理系においては網膜から取り入れた情報を形・色・運動などの属性ごとに一度分解し、それらを再統合することによって物体像を認識している。この再統合の

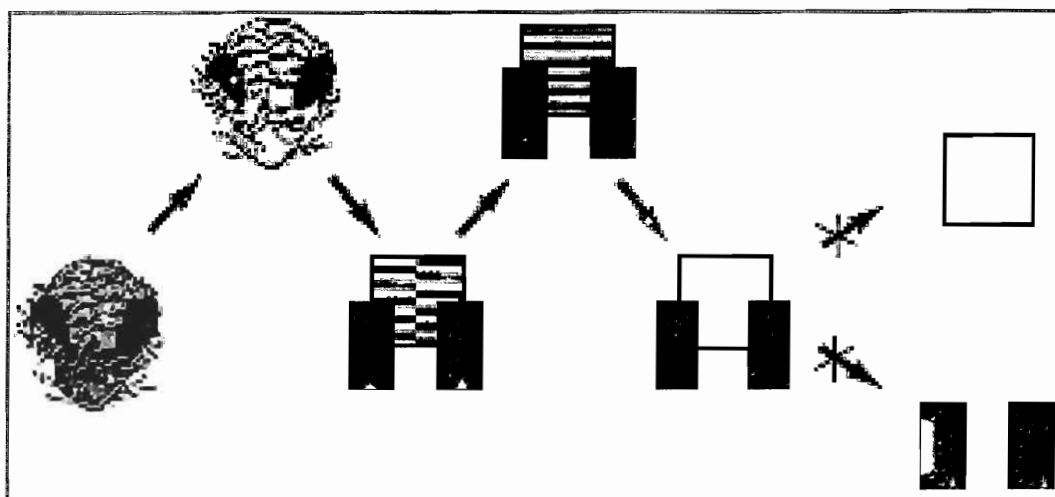


図 2.8: 図形を単純化し刺激となっている特徴を抽出する過程 (文献 [22] より抜粋)

トラの頭に反応した細胞は、白地の正方形の両側に黒い長方形を重ねた図形特徴に反応していることが解明された。さらに分解し、白い正方形や黒い長方形2つになると反応しなくなる。

過程において、注視が重要な役割を果たすという考えが Treisman らによって提唱された [17]。Treisman らは、様々な妨害刺激を持つ場面から特定の目標を検出させる視覚探索という課題実験の結果から、色・形・明るさなどの単純特徴は並列的に処理が可能であり、それらの単純特徴を組み合わせた結合特性の処理は、単純特徴同士を結合する際にその対象物体に注意を集中する必要があるため、逐次的にしか処理できないという結論を出した。この考えを元に提唱された理論が**特徴統合理論**[17]である。特徴統合理論では、図 2.10 で示すように、色や傾きなどの基本的な特徴をそれぞれ空間的に並列に抽出し、それらの特徴ごとの値を記録した特徴マップを形成する。これらの特徴マップは位置の情報のみを持つ位置マップと空間的な対応関係を保って繋がっており、特定の位置に注意を向けることで位置マップに連結した各特徴マップ上に対応する特徴値を呼び出して結合する。この場合、同時に複数の位置に注意を向けることは不可能なので、結合過程は逐次的に行われることになる。

注視を実行する際に欠かせないものに、眼球運動がある。眼球運動は注意を向けた対象物に視線を向け、高解像度で視覚情報処理が可能な網膜中心窩に対象物をとらえるための視線の移動に関わる運動 (gaze-shifting) と、視線を向けた対象物の網膜像が動かないように視線を保持するための運動 (gaze-holding) に大別される [25]。

視線の移動に関わる運動 (gaze-shifting) で生じる速度の速い眼球運動をサッケード (saccadic eye movement; saccade; 跳躍眼球運動) という。目標となる対象物の明るさや網膜の中心である中心窩からの距離、提示方法などによってサッケードが実際に発生するまでの時間である潜時が変化し、その速度は最高で  $1000 \text{ deg/Sec}$  にも達する。また、サッケードの対象物が視野周辺

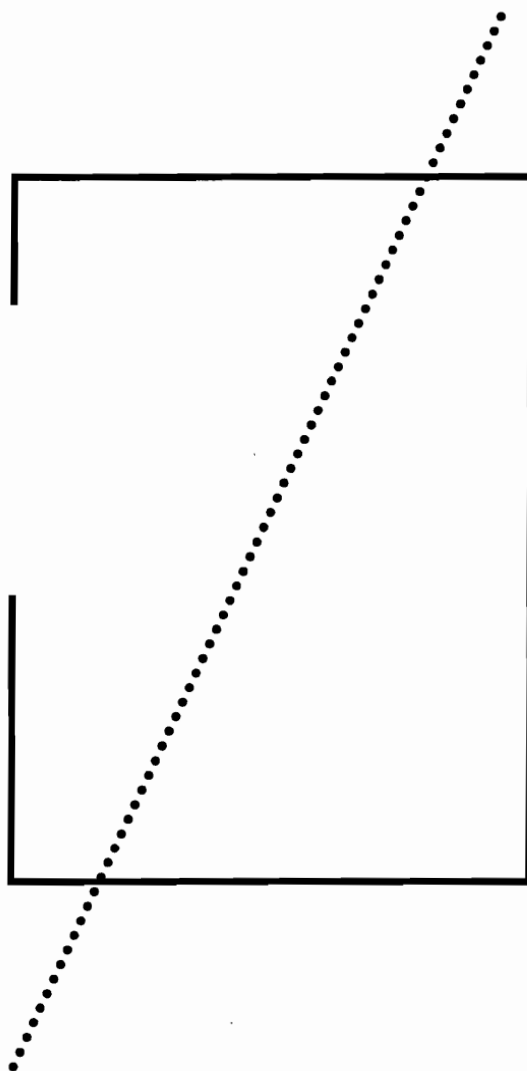


図 2.9: 「対象に基づく選択」の実証実験で用いられた刺激例 (文献 [25] より抜粋)

1つの対象 (例えば線分) に関する2つの属性 (線分の傾きと種類) を報告する条件の方が2つの対象 (四角形と線分) に関する2つの属性 (長方形の縦の長さ and 線分の傾き) を報告する条件よりも高い正答率を得られた。

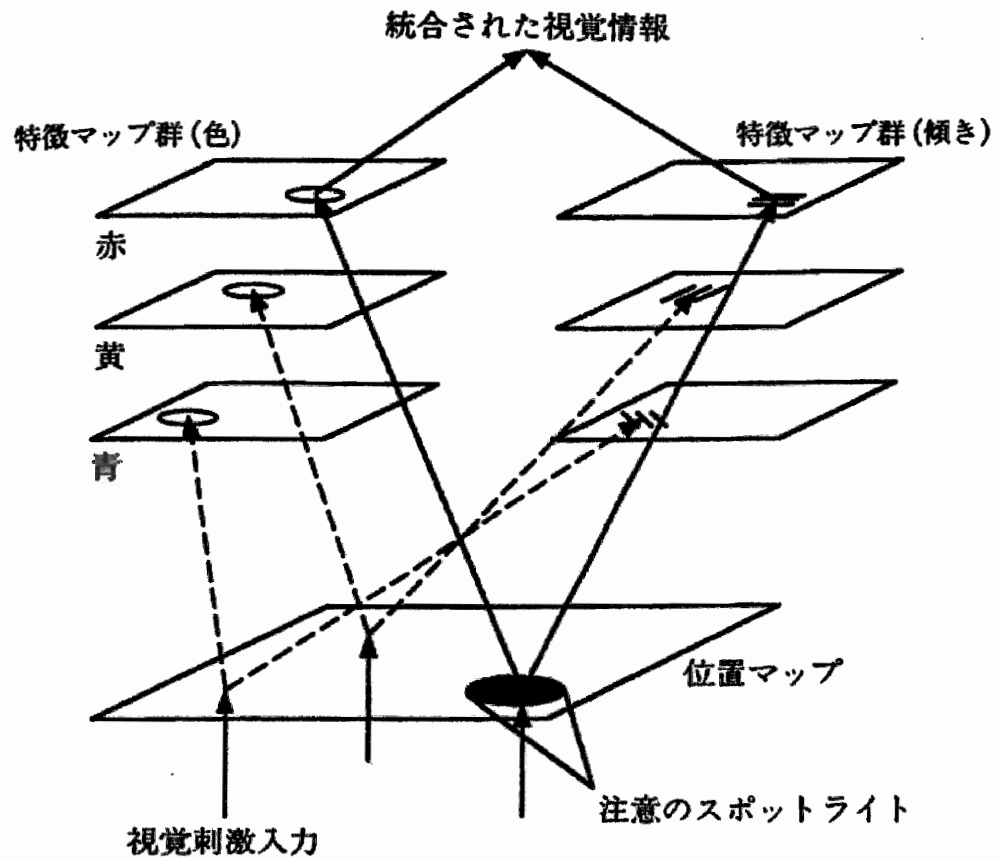


図 2.10: 特徴統合理論の概念図 (文献 [25] より抜粋)  
特徴マップ群の例として、色マップ群と傾きマップ群をとりあげて示す。

部にある場合、最初のサッケードだけでは対象物を補足しきれず、さらに小さなサッケードが生じて対象物をとらえようとする。この小さなサッケードは修正サッケード (correction saccade) と呼ばれ、最初のサッケードの終了から修正サッケード開始までの時間がふつうのサッケードよりもかなり短いことから修正サッケードは最初のサッケードから引き続いていているものとする説 [26] もある。

## 2.5 認識活動における仮定と検証

ヒトの認識活動における「仮定と検証」の利用は、認知心理学者である Neisser によって示唆されたものである [14]。Neisser によると、認知の処理には、低次のレベルから高次のレベルへと順次特徴分析を積み重ねて解釈認識に至るボトムアップ処理 (bottom-up processing) と、過去に学習した知識や経験を利用して仮説を立ててから特徴分析に向かうトップダウン処理 (top-down processing) の二つの処理が存在しているという。認知の過程では、このボトムアップ処理とトップダウン処理が循環的に行われることで、認識対象の階層性を持った概念図式であるスキーマ (schema) が活性化され認知に至る、と説明した。その後も、多数の心理モデルで認知を行う際の「仮定と検証」の有効性が示されている [27][28]。

生理学的に見ても、視覚情報の処理経路はボトムアップの経路だけではなく、視床の傍膝状体核で制御される第1次視覚野から外側膝状体へのフィードバック経路が存在することなどから、高次から低次への伝達経路が存在することが解明されている。この事からも、ボトムアップ処理とトップダウン処理を共に利用して認識を行っていることが予想される。

## 2.6 多解像度を用いた視覚情報処理

ヒトの視覚情報処理系においては多解像度による視覚情報処理が行われていることが解明されている。

生理学的には、網膜から視覚野に至る経路において、均一な情報伝達がなされていないことが挙げられる。例えば、網膜の各部位における細胞の分布は一樣ではなく、中心部と周辺部ではかなり違っている。網膜の出力細胞である神経節細胞に対して、その受容野の大きさと、網膜上の位置との関係を調べると、網膜中心部の中心窩と呼ばれる場所では受容野は小さく、周辺に行くに連れて受容野は大きくなっていることがわかっている。つまり、網膜中心部では解像度が高く、周辺部に行くに連れ解像度は急激に低下している。網膜と同様に視覚野に存在する細胞の受容野の大きさも網膜中心部から離れるにつれ増加していることがわかっている。

また心理学的にも、ヒトに様々な情報を探索させその時空間特性を分析するという視覚探索 (Visual Search) の実験から多解像度を用いていることが示唆されている。これは、被験者に同じ刺激を提示して、刺激に含まれる探索目標の有無・色・形状を別々の課題として実験する方

法である。この結果から形状の認知には高い解像度の注意が必要であるが、それ以外の課題には低解像度の分散的な注意しか必要ないという結果が得られている [29]。

## 2.7 まとめ

ヒトの視覚情報処理系において生理学的・心理学的に解明されている機構を述べた。視覚情報処理系においては、ボトムアップ処理だけではなく、トップダウン処理も使用されており、両者の循環によって認知が行われている。物体の位置と形状を処理する経路は別々に存在し、処理の過程で多解像度を利用している。ばらばらの情報を統合する時の鍵となるのが視覚的注意である。

次章では、これらの知見を元に、パターン認識可能な数理モデルを作成する。

## 第3章

# 仮定と検証に基づく位置・変形に頑健なパターン認識モデル

この章では、第2章で述べた知見に基づき実際にパターン認識の可能な視覚情報処理系の数理モデルを提案する。まず、生理学的に妥当性があり工学的に有効な視覚情報処理系の数理モデルであるネオコグニトロンを紹介する。そして、ネオコグニトロンを特徴抽出器として利用し、パターンの一部のみ網膜モデルに入力されている場合でもパターンを正しく切り出し・認識を行うことのできる数理モデルを構築する。提案したモデルの計算機シミュレーションを行い有効性を評価するとともに考察を行う。

### 3.1 ネオコグニトロン

Fukushimaによって提案されたネオコグニトロン [30] は、パターン認識を目的とする数理モデルであり、2.2節で説明した Hubel-Wiesel の階層型仮説を基にして構成された。

#### 3.1.1 ネオコグニトロンの構造

ネオコグニトロンは、図 3.1 で示すように多層の階層構造を持つ神経回路網である。パターンを提示する入力層である  $U_0$  層を先頭として、単純型細胞に対応する S 細胞で構成される  $U_S$  層と、複雑型細胞に対応する C 細胞で構成される  $U_C$  層とが交互に並んでいる。

各細胞層内では多数の細胞が 2 次元平面上に集まっており、これを細胞面と呼ぶ。1 つの細胞層内にはこのような細胞面が多数存在しており、同一細胞面内の細胞はいずれも同一形状の特徴に対して反応するという性質を持つ。

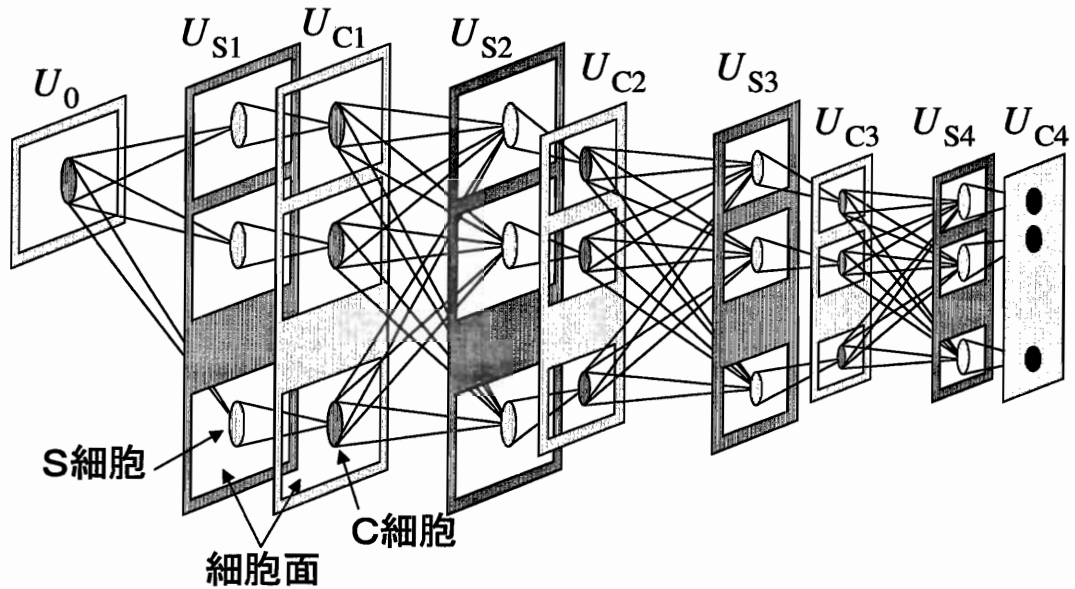


図 3.1: ネオコグニトロン概念図

S細胞は特徴抽出機能を持つ細胞で、入力結合として前段のC細胞との間で可塑性のある興奮性結合を持つ。C細胞は位置ずれの許容を行う性質を持つ細胞で、前段のS細胞との間で固定結合を持つ。また、図中では省略してあるが、S細胞・C細胞以外にV細胞と呼ばれるものも存在し、S細胞との間で抑制性・可塑性の結合を持つ。これら可塑性を持つ結合の強度は、学習時に自己組織的に決定される。

### 3.1.2 ネオコグニトロンの定式化

$l$ 段目のS細胞層を  $U_{Sl}$ 、C細胞層を  $U_{Cl}$  と記す。入力層は  $U_0$  で表される。各細胞層は細胞面から構成されており、この細胞面の数をそれぞれ  $K_{Sl}$ 、 $K_{Cl}$  とする。各細胞面にはS細胞あるいはC細胞が平面上に並んでおり、 $U_{Sl}$ 層の  $k$ 番目 ( $k < K_{Sl}$ )の細胞面に属する1個のS細胞の出力を  $u_{Sl}(\mathbf{n}, k)$ 、 $U_{Cl}$ 層の  $\kappa$ 番目 ( $\kappa < K_{Cl}$ )の細胞面に属する1個のC細胞の出力を  $u_{Cl}(\mathbf{n}, \kappa)$  と記す。同様に、V細胞の出力を  $u_{Vl}(\mathbf{n})$  とする。 $\mathbf{n} = (n_x, n_y)$  は、細胞面内での2次元位置座標を表す。また、細胞の出力を“ $u$ ”で表すのに対して、細胞自体を指し示すのに“ $U$ ”と表記する。

図3.2で示すように、 $l$ 段目のS細胞  $U_{Sl}(\mathbf{n}, k)$  は  $l-1$ 段目のC細胞及び  $l$ 段目のV細胞から結合を受けている。S細胞  $U_{Sl}(\mathbf{n}, k)$  とC細胞  $U_{Cl-1}(\mathbf{n} + \boldsymbol{\nu}, \kappa)$  間の興奮性可変結合は  $a_l(\boldsymbol{\nu}, \mathbf{n}, \kappa, k)$  で示される。V細胞  $U_{Vl}(\mathbf{n})$  は  $U_{l-1}(\mathbf{n} + \boldsymbol{\nu}, \kappa)$  から興奮性固定結合を受け取り、その結合強度



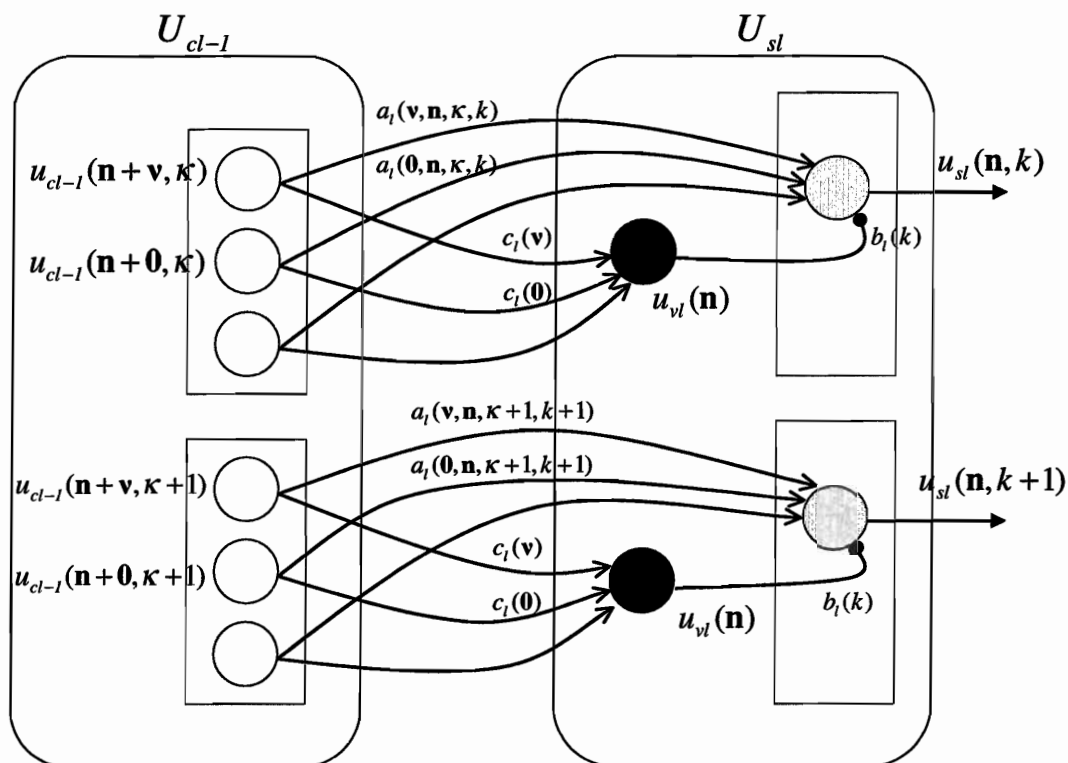


図 3.2: C 細胞層-S 細胞層間の結合関係

実際の各細胞は 2 次的に配置されているが、ここでは簡略化のため 1 次的に配置した場合を考えている。また、図中では  $U_{cl-1}$  層の  $\kappa$  番目の細胞面から、 $U_{sl}$  層の  $k$  番目の細胞面への結合のみを記述しているが、実際には他の全ての細胞面間にも結合が存在する。

を  $c_l(\nu)$  で表す。 $U_{vl}(\mathbf{n})$  は  $U_{sl}(\mathbf{n}, k)$  に抑制性信号を送っており、その可変結合強度は  $b_l(k)$  である。このとき、V 細胞  $U_{vl}(\mathbf{n})$  の出力は

$$u_{vl}(\mathbf{n}) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} c_l(\nu) \cdot \{u_{cl-1}(\mathbf{n} + \nu, \kappa)\}^2}, \quad (3.1)$$

で与えられ、それを用いて S 細胞  $U_{sl}(\mathbf{n}, k)$  の出力は

$$u_{sl}(\mathbf{n}, k) = r_l \cdot \varphi \left[ \frac{1 + \sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} a_l(\nu, \mathbf{n}, \kappa, k) \cdot u_{cl-1}(\mathbf{n} + \nu, \kappa)}{1 + \frac{r_l}{1 + r_l} \cdot b_l(k) \cdot u_{vl}(\mathbf{n})} - 1 \right], \quad (3.2)$$

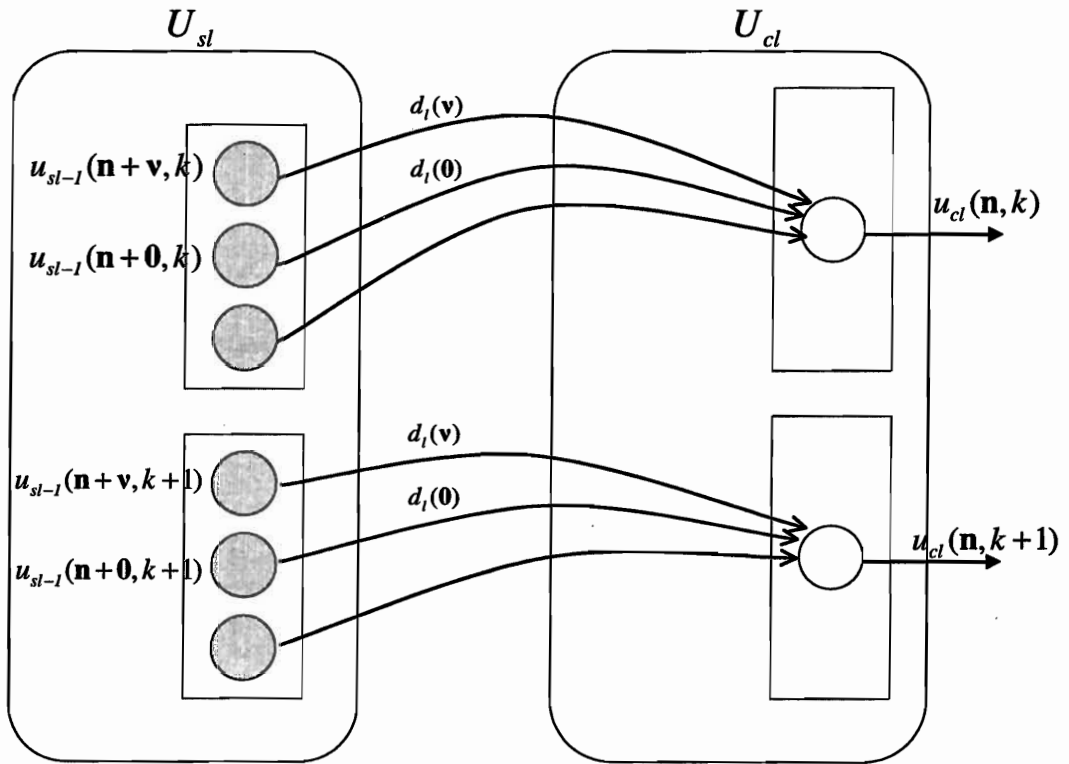


図 3.3: S 細胞層-C 細胞層間の結合関係

S 細胞が前層の全ての細胞面内に存在する C 細胞から入力を受けていたのに対して、C 細胞は同層の同じ細胞面番号  $k$  に属する S 細胞からのみ入力を受ける。

と与えられる。ただし、関数  $\varphi[ ]$  は

$$\varphi[x] = \begin{cases} x & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (3.3)$$

で与えられる折線型 (半波整流形) の関数である。また、 $A_l$  は  $U_{Sl}$  の S 細胞が  $U_{Cl-1}$  の C 細胞と結合する範囲の広さを表し、 $r_l$  は抑制入力を制御する正の定数であり、一般的な神経細胞における閾値に対応する。

$l$  段目の C 細胞  $U_{Cl}(n, k)$  は、図 3.3 で示すように  $l$  段目の S 細胞から結合を受けている。C 細胞  $U_{Cl}(n, k)$  と S 細胞  $U_{Sl}(n + \nu, k)$  間の興奮性固定結合は  $d_l(\nu)$  で示される。このとき、C 細胞  $U_{Cl}(n, k)$  の出力は

$$u_{Cl}(n, k) = \psi \left[ \sum_{\nu \in D_l} d_l(\nu) \cdot u_{Sl}(n + \nu, k) \right], \quad (3.4)$$

で与えられる。ただし  $\psi[\ ]$  は C 細胞の入出力間の飽和特性を定める関数で

$$\psi[x] = \frac{\varphi[x]}{1 + \varphi[x]}, \quad (3.5)$$

で表される。また、 $D_l$  は  $U_{Cl}$  の C 細胞が  $U_{Sl}$  の S 細胞と結合する範囲の広さを表す。

### 3.1.3 ネオコグニトロンの学習則

ネオコグニトロンの学習では、 $U_0$  層に学習パターンを入力し、S 細胞が持つ可変結合の重みを更新する動作を行う。通常、 $U_{S1}$  層及び  $U_{C1}$  層は線分の方位を抽出するように結合重みをあらかじめ調節しておき、学習は  $U_{S2}$  層以降に対して適用される。学習は低次の層より順に進められ、ある段の結合が全て決定した後に次の段の学習へと移る。学習前の段階では  $U_{S2}$  層以降に細胞面は存在せず、学習を進めて行く過程で細胞面が自動的に生成されてゆき、学習パターン中の重要な特徴が獲得されてゆく。

パターン中での重要な局所特徴の選出方法として、一般的に「シードセル生成面を用いた学習」[31]を用いることが多い。この方法は未獲得な局所の特徴を表現するために、winner-take-all 型学習則における winner-cell に相当するシードセルを選出し、可変結合の強化をこのシードセルのみに適用することで重要な特徴を獲得してゆく方法である。いま、シードセルとして S 細胞  $u_{Sl}(\hat{n}, \hat{k})$  が選ばれたとすると、この細胞に繋がる可変結合の強度  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  および  $b_l(\hat{k})$  は次式に従って更新される。

$$\Delta a_l(\nu, \hat{n}, \kappa, \hat{k}) = q_l \cdot c_l(\nu) \cdot u_{Cl-1}(\hat{n} + \nu, \kappa), \quad (3.6)$$

$$\Delta b_l(\hat{k}) = q_l \cdot u_{Vl}(\hat{n}). \quad (3.7)$$

ここで、 $q_l$  は学習定数であり、正の定数である。シードセルにつながる結合強度を更新した後、同じ細胞面内の全ての入力結合をシードセルの入力結合と同じ重みに設定し、細胞面内で同一の空間分布を保つようにする。このシードセルを選択するための指標として、シードセル生成面と呼ばれる細胞面を導入する。

シードセル生成面は、入力パターン中の重要な局所の特徴のある位置に強い反応を示す細胞面で、前段の C 細胞から固定結合で入力を受け取る。シードセルはシードセル生成面中で最大発火した細胞の位置に対応して選出され、式 (3.6), (3.7) に従って新規細胞面の学習を行う。次に、学習した細胞面の反応位置周辺にあるシードセル生成面に対して抑制をかけ、再度シードセル生成面内の最大発火位置からシードセルを選出し、学習を行う。この操作をシードセル生成面内の発火が無くなるまで繰り返す。

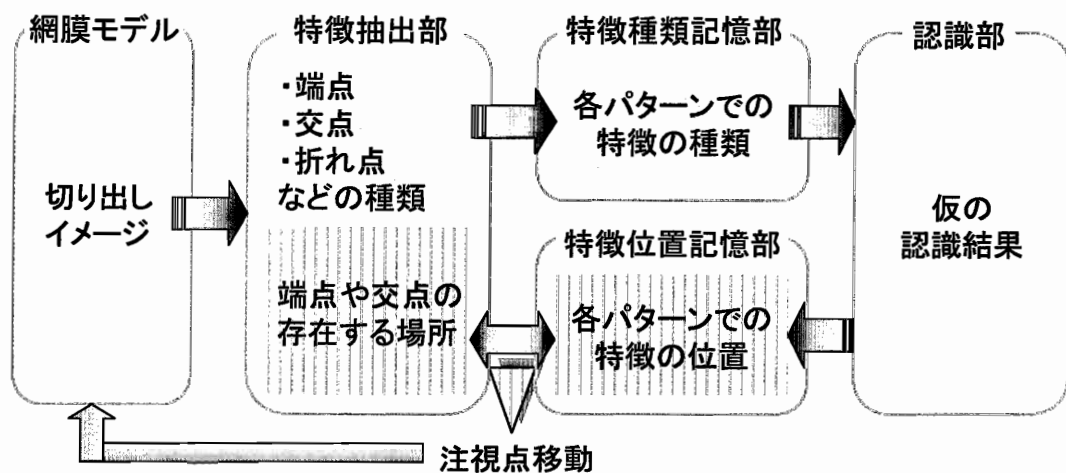


図 3.4: 位置・変形に頑健なパターン認識モデルの概要図

### 3.2 仮定と検証に基づく位置・変形に頑健なパターン認識モデル

本節で提案するモデルでは網膜の受容範囲を逸脱するような大幅な位置ずれを許容するために、入力イメージが持つ**特徴の種類**と**特徴の位置**を分離して処理する方法を採用した。ここで、**特徴の種類**とはパターン中に含まれる折れ点や交点、端点など認識対象となるパターンの特徴的な構成要素がどれだけ含まれているかという情報であり、**特徴の位置**とは、パターン中のどの位置に折れ点や端点などが存在しているかという情報である。

このような方法をとる背景には、ヒトの視覚情報処理系において物体形状の知覚と物体位置の知覚を処理する経路が分かれていること（2.3節参照）を考慮したことが挙げられる。また、Treismanらの特徴統合理論（2.4節参照）においても色・形・明るさなどの単純特徴は位置を鍵として再統合が行われており、**対象物の位置が不確定な場合には位置情報を分離して使用する手段をとることは有効であることが予想される。**

また、**特徴の種類**と**特徴の位置**を統合し、最終的な認識結果を導出するための方法として2.5節で述べた「仮定と検証」を用いる。

まず、提案モデルの概要について述べる。

#### 3.2.1 モデルの概要

##### モデルの構成

本モデルは、図 3.4 で示されるように網膜モデル、特徴抽出部、特徴種類記憶部、特徴位置記憶部、認識部の5つのモジュールで構成される。

これらのモジュールの働きを簡潔にまとめると以下の項目となる。

- **網膜モデル**  
全景中から一部の領域を切り出し特徴抽出部に転送。
- **特徴抽出部**  
切り出されたイメージから下記の2種類の特徴を抽出。
  - － **特徴の種類**  
端点・交点・折れ点などがそれぞれどの程度存在しているのかという情報。
  - － **特徴の位置**  
端点・交点・折れ点などがイメージ中のどこに存在しているのかという情報。
- **特徴種類記憶部**  
学習パターンにおける特徴の種類を記憶。
- **特徴種位置記憶部**  
学習パターンにおける特徴の位置を記憶。
- **認識部**  
現在の切り出し位置でのパターンに対する仮の認識結果を算出。

### モデルの動作

網膜モデルで切り出されたイメージは特徴抽出部に送られ、特徴の種類と特徴の位置が抽出される。その後、特徴種類記憶部の情報を利用して認識部において現在の切り出し位置での仮の認識結果が出力される。続いて、仮の認識結果に基づき、特徴位置記憶部に記憶されている特徴の位置と現在抽出されている特徴の位置とを比較し、切り出し位置の補正を行う。その後、網膜モデルにおいて切り出し位置の補正を行った後に再び特徴を抽出し、認識部においてパターンを認識する。

認識結果が先ほどの仮認識結果と一致しない場合には仮認識結果は棄却され、次候補を新たな仮認識結果として切り出し位置の補正をやり直す。

仮認識結果と補正後の認識結果が同一の場合には仮認識結果が採択され、再度特徴の位置を比較して、切り出し位置の補正を続ける。この手続きを繰り返してゆき、切り出し位置の補正が終わるまで続ける。補正が終わった段階での仮認識結果が最終的な認識結果となり、その位置がパターンの適切な切り出し位置となる。

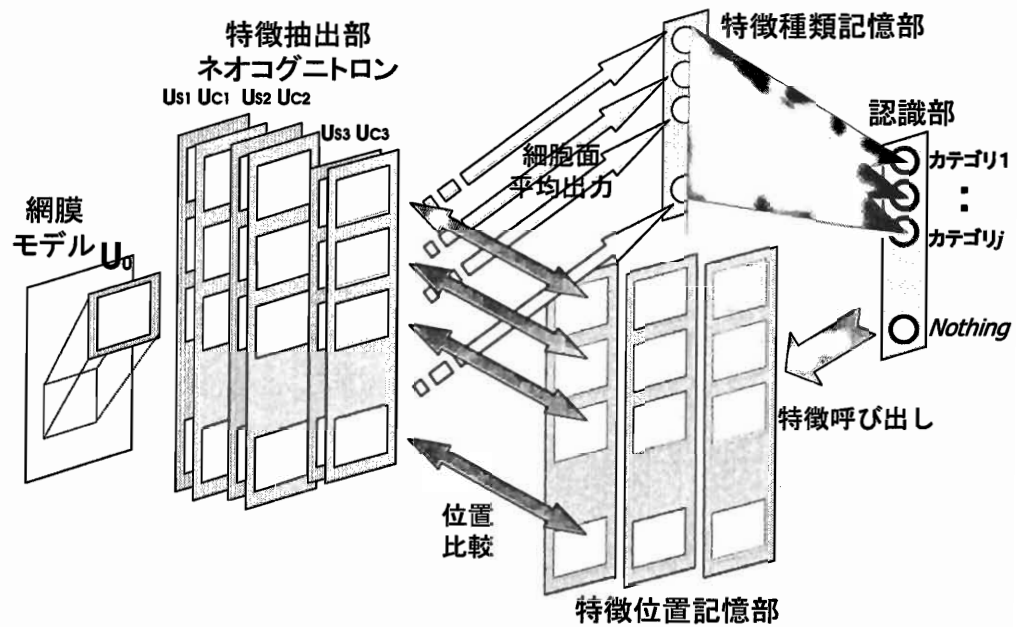


図 3.5: 位置・変形に頑健なパターン認識モデルの構造

### 3.2.2 モデルの構成

ここでは、モデルの具体的な構成について説明する。  
まず、図 3.5 にモデル全体の構造を示す。

- **網膜モデル**  
全景中からある一部を切り出し特徴抽出部の入力層である  $U_0$  層に転送。
- **特徴抽出部**  
3.1 節で紹介したネオコグニトロンで構成。
- **特徴種類記憶部**  
特徴抽出部から  $U_{C3}$  層における各細胞面の発火強度平均を入力。
- **特徴位置記憶部**  
各パターンにおける  $U_{C3}$  層の発火パターンの重心を記憶。
- **認識部**  
特徴種類記憶部と可変結合を生成。

次に、各モジュールの構造について定式化しながら説明する。

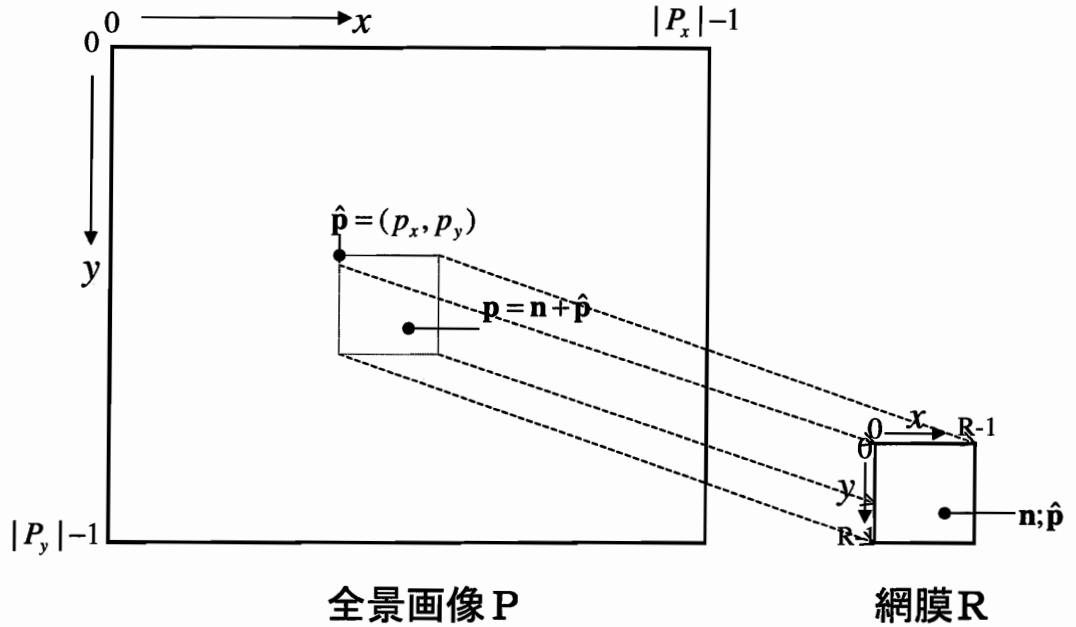


図 3.6: 網膜モデルにおける 2 次元座標系

### 網膜モデル

網膜モデルは、全景中から特徴抽出部の入力層に一定範囲内の視覚情報を投射する役割を担っている。いま、図 3.6 のように座標系を取り、全景画像 (Panoramic view) の空間  $P$  の大きさを  $|P_x| \times |P_y|$  [pixel]、座標  $\mathbf{p} = (x, y)$  における画素値を  $u_p(\mathbf{p})$  とする。網膜 (Retina) の空間  $R$  の大きさを  $R \times R$  [pixel]、網膜上の座標を  $\mathbf{n}$  で表す。このとき、切り出し位置  $\hat{\mathbf{p}} = (p_x, p_y)$  における網膜出力  $u_{gc}(\mathbf{n}; \hat{\mathbf{p}})$  は次式で与えられる。

$$u_{gc}(\mathbf{n}; \hat{\mathbf{p}}) = \begin{cases} \frac{u_p(\mathbf{n} + \hat{\mathbf{p}})}{P_{\max}} & ; \text{if } \mathbf{n} + \hat{\mathbf{p}} \in P \\ 0 & ; \text{otherwise} \end{cases} \quad (3.8)$$

ただし、 $P_{\max}$  は画素値が取りうる最大値である。

$$P_{\max} = \max_{\mathbf{p} \in P} u_p(\mathbf{p}). \quad (3.9)$$

$P_{\max}$  で画素値を正規化することで、網膜出力の取りうる値は  $0 \leq u_{gc}(\mathbf{n}; \hat{\mathbf{p}}) \leq 1$  となる。

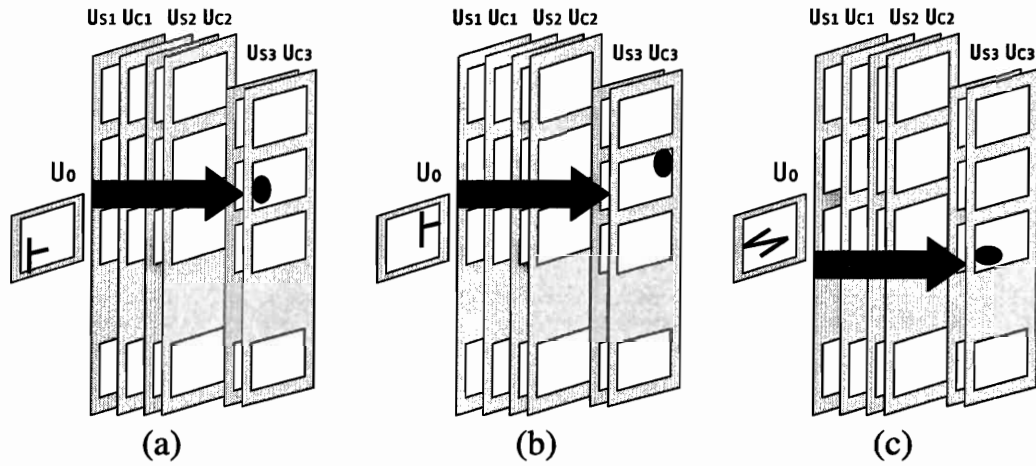


図 3.7: 特徴抽出部の反応例

(a),(b):同一の大局的特徴に対しては同じ細胞面が発火し、発火している細胞の位置が提示された特徴の位置に対応する。(a),(c):異なる大局的特徴に対しては異なる細胞面が発火する。

表 3.1: ネオコグニトロンのパラメータ

パラメータ種類	第0段目	第1段目	第2段目	第3段目
細胞面の大きさ $N_l$	$R \times R$	$n_1 \times n_1$	$n_2 \times n_2$	$R \times R$
結合範囲 $A_l$	—	$a_1 \times a_1$	$a_2 \times a_2$	$a_3 \times a_3$
結合範囲 $D_l$	—	$d_1 \times d_1$	$d_2 \times d_2$	$d_3 \times d_3$
抑制定数 $r_l$	—	$r_1$	$r_2$	$r_3$
学習定数 $q_l$	—	$q_1$	$q_2$	$q_3$

### 特徴抽出部

特徴抽出部は、 $U_0$  層から  $U_{c3}$  層までのネオコグニトロンで形成されている。

学習終了後のネオコグニトロン  $U_{c3}$  層では、学習パターンの中の端点や交点などの局所的特徴が統合された大局的特徴に対して各細胞面が発火するように結合が調整される。異なる大局的特徴に対しては異なる細胞面が発火し、細胞面中での発火する場所は特徴が  $U_0$  層のどの位置に見られるかに対応する (図 3.7)。

表 3.1 にネオコグニトロンのパラメータ表記を示す。入力である第0段及び出力である第3段の各細胞面の大きさは  $R \times R$  とし、その他の層の細胞面の大きさは任意とする。



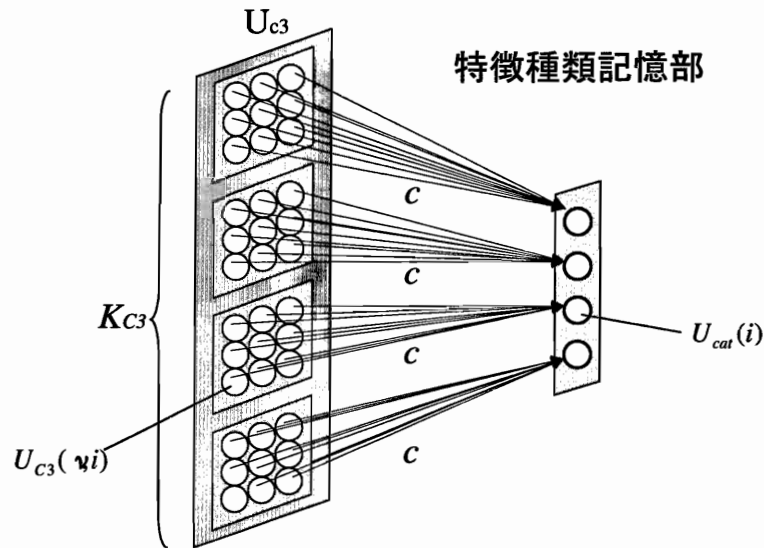


図 3.8: 特徴種類記憶部の構造

特徴抽出部の出力層である  $U_{C3}$  層の対応する各細胞面から強度  $c$  の興奮性固定結合を通じて特徴種類記憶部の各細胞に情報が伝搬される。特徴種類記憶部の細胞出力値は、位置情報を欠落させた  $U_{C3}$  層の各細胞面における発火値に対応する。

$U_0$  層は網膜モデルから次式に従って入力を受ける。

$$u_0(\mathbf{n}) = u_{gc}(\mathbf{n}; \hat{\mathbf{p}}); \mathbf{n} \in R \quad (3.10)$$

また、学習終了後の  $U_{C3}$  層細胞面数を  $K_{C3}$  とする。

このとき、3.1 節の定式化に従って特徴抽出結果である  $U_{C3}$  層の各細胞出力が計算されてゆく。

### 特徴種類記憶部

特徴種類記憶部では、特徴抽出部で抽出されたパターンの大局的特徴に対応して、その提示位置に関係なく発火する細胞が存在する。細胞の個数は学習後の  $U_{C3}$  層における細胞面の数  $K_{C3}$  に等しい。図 3.8 に特徴種類記憶部の構造と、特徴抽出部からの入力結合を図示する。特徴種類記憶部の各細胞には  $U_{C3}$  層の対応する各細胞面から興奮性固定結合  $c$  を通じて情報が伝搬される。

特徴種類記憶部の  $i$  番目の細胞を  $U_{cat}(i)$  とし、その出力を  $u_{cat}(i)$  とすると、出力は次式で与えられる。

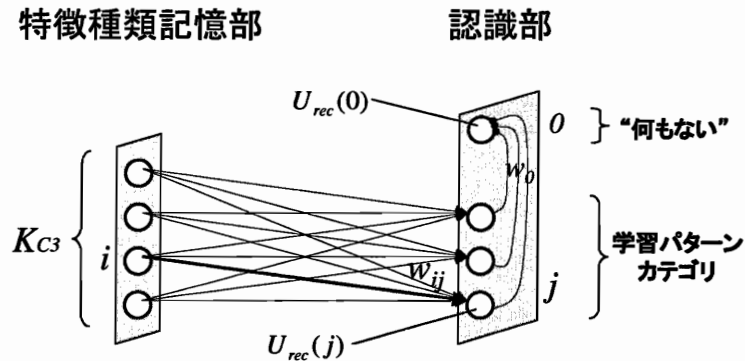


図 3.9: 認識部の構造

特徴種類記憶部の各細胞と、認識部のパターンカテゴリに対応する細胞との間には可変の興奮性全結合  $w_{ij}$  が存在しており、パターンカテゴリ  $j$  に対応する細胞  $U_{\text{rec}}(j)$  と “何もない” 事を示す細胞  $U_{\text{rec}}(0)$  の間には固定の抑制性全結合  $w_0$  が存在している。

$$u_{\text{cat}}(i) = \sum_{\nu \in \mathcal{Z}} c \cdot u_{C3}(\nu, i) \quad (3.11)$$

ただし、

$$c = \frac{1}{R^2} \quad (3.12)$$

であり、 $\mathcal{Z}$  は細胞面内の空間を表す。

すなわち、特徴種類記憶部の各細胞の出力は、 $U_{C3}$  層の各細胞面における平均発火率に対応する。

### 認識部

学習パターンのカテゴリ数を  $K_{\text{learn}}$  としたとき、認識部にはそれぞれのパターンカテゴリに対応する細胞  $K_{\text{learn}}$  個と、現在の切り出し範囲に “何もない” 事を示す 1 個の細胞の計  $K_{\text{learn}} + 1$  個の細胞が存在する。図 3.9 に、認識部の構造と特徴種類記憶部から認識部への入力結合を図示する。特徴種類記憶部の各細胞と、認識部のパターンカテゴリに対応する細胞との間には可変の興奮性全結合  $w_{ij}$  が存在しており、パターンカテゴリに対応する細胞と “何もない” 事を示す細胞の間には固定の抑制性全結合  $w_0$  が存在している。

認識部のパターンカテゴリ  $j$  ( $1 \leq j \leq K_{\text{learn}}$ ) に対応する細胞を  $U_{\text{rec}}(j)$  とし、その出力を  $u_{\text{rec}}(j)$  で表す。同様に “何もない” 事を示す細胞を  $U_{\text{rec}}(0)$ 、出力を  $u_{\text{rec}}(0)$  とする。 $U_{\text{cat}}(i)$  と

$U_{\text{rec}}(j)$  の興奮性可変結合を  $w_{ij}$ 、 $U_{\text{rec}}(j)$  と  $U_{\text{rec}}(0)$  の抑制性固定結合を  $w_0$  としたとき、各細胞の出力は式 (3.13), (3.14) で与えられる。

$$u_{\text{rec}}(j) = \rho(j) \cdot \iota(j) \cdot \sum_{i=1}^{K_{C3}} w_{ij} \cdot u_{\text{cat}}(i) \quad (3.13)$$

$$u_{\text{rec}}(0) = 1 - \sum_{j=1}^{K_{\text{learn}}} w_0 \cdot u_{\text{rec}}(j) \quad (3.14)$$

ただし、 $\rho(j)$  はパターンカテゴリ  $j$  という仮認識結果が棄却 (reject) されていれば 0、そうでなければ 1 を取る関数である。

$\iota(j)$  は思い込み (Illusion) を表す関数で、現在検証中の仮認識結果がパターンカテゴリ  $j$  である場合には定数  $I_l (> 1)$  を、そうでなければ 1 を取る関数である。この関数は学習したパターンが類似した特徴を持っている場合に、検証の初期段階において仮認識結果が適切なものであった時、微小の差で誤って棄却してしまうことを防ぐために導入するもので、定数  $I_l$  が大きいほど提示されているパターンが現在の仮認識結果であると思いつく程度が激しい事に対応する。

また、通常  $w_0 = 1$  とする。

全ての細胞の出力値が求まった後、winner-take-all-rule に従い、現在の仮認識結果に対応する細胞のみが発火を続ける。

$$u_{\text{rec}}^{\text{cont}}(j) = \begin{cases} 1; & \text{if } j = \arg \max_{0 \leq l \leq K_{\text{learn}}} u_{\text{rec}}(l) \\ 0; & \text{otherwise} \end{cases} \quad (3.15)$$

### 特徴位置記憶部

特徴位置記憶部は、パターンカテゴリ数  $K_{\text{learn}}$  と同数の細胞層から構成される。各細胞層は、学習終了時の  $U_{C3}$  層と同じ構造をしており、パターンカテゴリ  $j$  に対応する細胞層を  $U_{\text{loc}}(j)$  とする。

$U_{\text{loc}}(j)$  には、カテゴリ  $j$  の学習パターンを提示したときの  $U_{C3}$  層の各細胞面における発火パターンの重心位置が記憶されている。認識部と特徴位置記憶部は興奮性固定結合  $w$  で結ばれており、 $U_{\text{rec}}(j)$  が発火している場合にのみ  $U_{\text{loc}}(j)$  は記憶した発火パターンの重心位置を想起する。通常  $w = 1$  とする。

すなわち、 $U_{\text{loc}}(j)$  の第  $k$  番目の細胞面における位置  $\mathbf{n}$  の細胞の出力  $u_{\text{loc}}(j, \mathbf{n}, k)$  は

$$u_{\text{loc}}(j, \mathbf{n}, k) = w \cdot u_{\text{rec}}^{\text{cont}}(j) \cdot u_{\text{mem}}(j, \mathbf{n}, k) \quad (3.16)$$

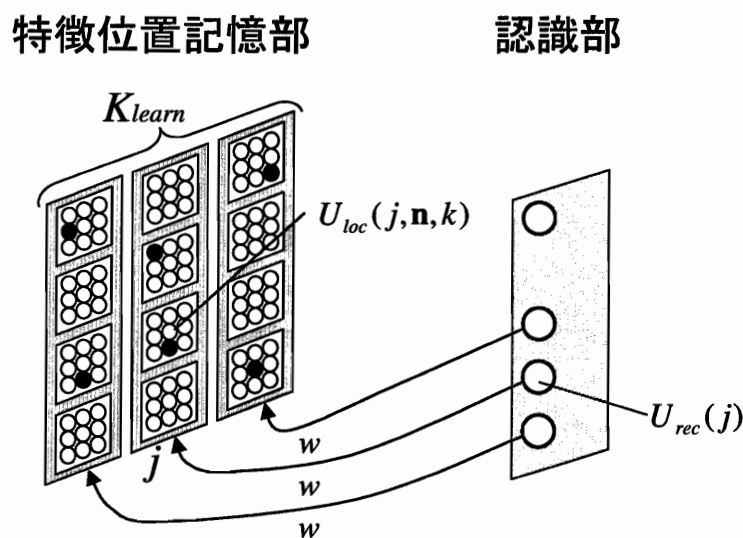


図 3.10: 特徴位置記憶部の構造

特徴位置記憶部は、 $K_{learn}$  個の  $U_{C3}$  層と同じ構造を持つ細胞層から構成される。認識部と特徴位置記憶部は興奮性固定結合  $w$  で結ばれており、カテゴリ  $j$  に対応する認識部の細胞  $U_{rec}(j)$  が発火している場合にのみ  $U_{loc}(j)$  は記憶した発火パターンの重心位置を想起する。●は  $u_{mem}(j, n, k) = 1$  の細胞である。

ただし  $u_{mem}(j, n, k)$  は学習時に獲得される発火パターン重心位置を表し、0 または 1 をとる。 $u_{mem}(j, n, k)$  についての詳細は後述する。

### 3.2.3 学習方法

使用する各学習パターンの大きさを  $R \times R[\text{pixel}]$  とし、学習パターンのカテゴリ数を  $K_{learn}$  とする。カテゴリ  $j$  に属する学習パターンを  $P(j)$ 、 $P(j)$  における座標  $\mathbf{p} = (x, y)$  の画素値を  $p(j, \mathbf{p})$  とする。

学習時には、網膜モデルに与える切り出し位置を

$$\hat{\mathbf{p}} = (p_x, p_y) = (0, 0) = \mathbf{0} \quad (3.17)$$

と固定しておく。つまり、学習時の網膜モデルの出力は常に次式で求められる。

$$u_{gc}(\mathbf{n}; \mathbf{0}) = p(j, \mathbf{n}) \text{ when category } j \text{ is trained.} \quad (3.18)$$

モデルの学習は、以下の3段階を経て行われる。

### 特徴抽出部の学習

モデルの学習は、まず特徴抽出部のネオコグニトロンに対して学習を適応することから始まる。ネオコグニトロンの学習には3.1.3節で説明したシードセル生成面を用いた学習法を利用し、すべてのカテゴリ  $j$  に対して各段順番に学習を行う。ネオコグニトロンの学習がすべてのカテゴリ  $j$  に対して終了した後の  $U_{C3}$  層における細胞面数を  $K_{C3}$  とする。

### 特徴種類記憶部–認識部間の学習

特徴抽出部の学習が終わった後、特徴種類記憶部の細胞を  $K_{C3}$  個だけ用意し、式(3.12)に従って重み  $c$  の固定結合を生成する。同時に、認識部の細胞との可変結合を生成する。認識部の細胞との可変全結合の重み  $w_{ij}$  は、

$$w_{ij} = 1 \text{ for all } i, j \quad (3.19)$$

で初期化する。

次に、認識部の各細胞が対応する学習パターンの特徴の種類を獲得するように可変全結合の重み  $w_{ij}$  を調整する。

まず、カテゴリ  $j$  の学習パターンを網膜モデルに提示し、特徴抽出部において特徴を抽出する。その後、特徴種類記憶部の各細胞での反応を式(3.11)に従って求める。次に、カテゴリ  $j$  に対応する認識部の細胞  $U_{rec}(j)$  の出力を  $u_{rec}(j) = 1$  に、それ以外の認識部の細胞の出力を0に設定し、結合重み  $w_{ij}$  をHebb則に従って強化した後、結合重みの総和が一定値を保つように正規化する。すなわち、Hebb則学習定数を  $\delta Rec$  とすれば、 $U_{rec}(j)$  と  $U_{cat}(i)$  との結合の重みは

$$w'_{ij} = \delta Rec \cdot u_{cat}(i) \cdot u_{rec}(j) + w_{ij}^{old}, \quad (3.20)$$

$$w_{ij}^{new} = \frac{\sum_{i=1}^{K_{C3}} w_{ij}^{old}}{\sum_{i=1}^{K_{C3}} w'_{ij}} \cdot w'_{ij}, \quad (3.21)$$

に従って更新される。同様に他のカテゴリについても  $w_{ij}$  を更新する。この重みの更新を各カテゴリで  $N_{hebb}$  回ずつ繰り返して、カテゴリ  $j$  が提示された際には  $U_{rec}(j)$  が発火するように学習を行う。

### 特徴位置記憶部の学習

最後に、特徴位置記憶部に各カテゴリの学習パターンにおける  $U_{C3}$  層での発火パターンの重心位置を学習させる。

まず、カテゴリ  $j$  の学習パターンを網膜モデルに提示し、特徴抽出部において特徴を抽出する。その後、 $U_{C3}$  層の各細胞面  $k$  について、細胞が発火している場所の重心  $\mathbf{g}_k = (g_x, g_y)$  を次式に従って算出する。

$$\mathbf{g}_k = \frac{1}{\sum_{\nu} \phi[u_{C3}(\nu, k)]} \cdot \sum_{\nu \in Z} \nu \cdot \phi[u_{C3}(\nu, k)] \quad (3.22)$$

ただし、 $\phi[\ ]$  は次式で与えられる 2 値関数である。

$$\phi[x] = \begin{cases} 1 & (x > 0) \\ 0 & (x \leq 0) \end{cases} \quad (3.23)$$

その後、特徴位置記憶部に算出した重心を記録する。

$$u_{\text{mem}}(j, \nu, k) = \begin{cases} 1 & (\nu = \mathbf{g}_k) \\ 0 & (\nu \neq \mathbf{g}_k) \end{cases} \quad (3.24)$$

この動作を繰り返し、全てのカテゴリに対して重心位置の記録を行う。

以上の 3 段階を経て、モデル全体の学習が終了する。

### 3.2.4 認識方法

認識は、大きく分けて下記の項目で示される 2 つのステップで行われる。

- 「仮定の生成」
  - その時点の切り出し位置から得られた**特徴の種類**を基に現在提示されているパターンカテゴリを推定し、仮認識結果を算出。
- 「仮定の検証」
  - 認識結果に基づき**特徴の位置**を利用して適切な位置へと切り出し位置を補正。

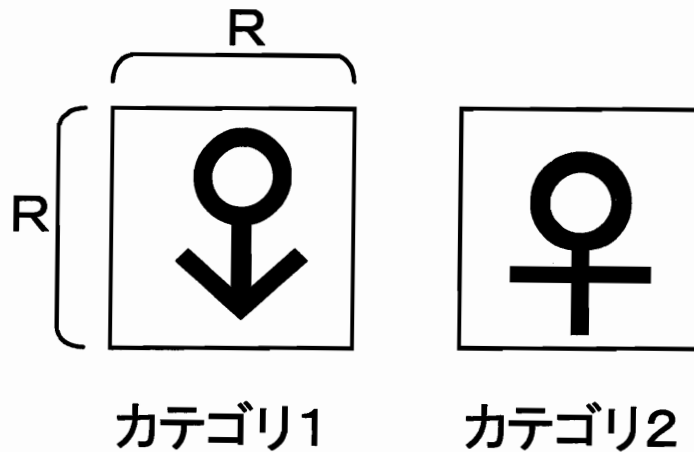


図 3.11: 学習パターンセットの例  
パターンの外枠は網膜の大きさを示している。

- 特徴の種類を利用して仮認識結果が適切な事を確認。

ここでは、具体例を交えながらこれらのステップについて解説する。

いま、図 3.11 で示される 2 カテゴリの学習パターンを用いてモデルの学習が終了しているとする。

ここで、図 3.12 のようにパターンが提示され、その初期切り出し位置が  $\hat{p}_0$  であるとする。

#### 「仮定の生成」

まず、切り出し位置  $\hat{p}_0$  における仮認識結果を求める。式 (3.8) に従い網膜上に映ったパターンは特徴抽出部で特徴抽出が行われ、特徴種類記憶部を経て認識部へと信号が伝搬される。そこで式 (3.15) によって最大発火した細胞が検出され、その結果  $u_{\text{rec}}^{\text{cont}}(j) = 1$  であるカテゴリ  $j$  が仮認識結果として出力される。また、 $u_{\text{rec}}^{\text{cont}}(0) = 1$  であった場合には現在見ている位置には“何もない”と判断され、認識は終了する。

いま、切り出し位置  $\hat{p}_0$  における仮認識結果としてカテゴリ 1 が選出されたとして話を進めてゆく。

#### 「仮定の検証」

「仮定の検証」では、まず「仮定の生成」によって求められた仮認識結果を基に、切り出し位置  $\hat{p}_0$  におけるパターンと、学習パターンとの間の特徴の位置を比較する。

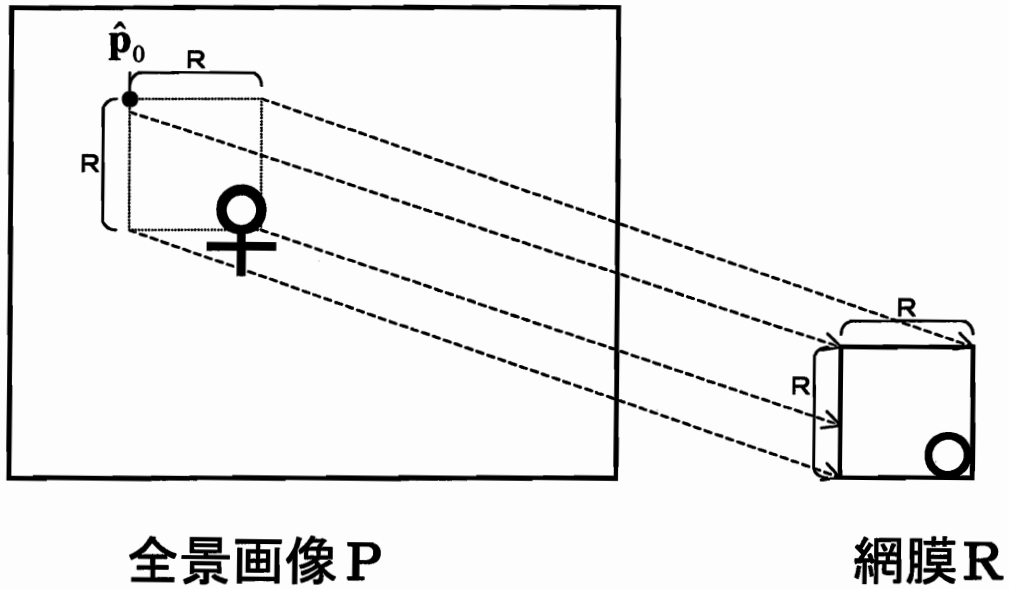


図 3.12: 提示パターンと初期切り出し位置

いま、式 (3.22) に従い、切り出し位置  $\hat{p}_0$  における  $U_{C3}$  層での各細胞面での発火重心  $g_k$  を求める。そして、図 3.13 で示すように、第  $k$  番目細胞面での発火の重心位置を仮認識結果であるカテゴリ  $j$  の学習パターンにおける発火重心位置  $g_k^j$  と比較し、その差から次式に基づき切り出し位置の修正量  $\Delta\hat{p}$  を求める。

$$\Delta\hat{p} = \frac{1}{K_{\text{fired}}} \cdot \sum_{k=1}^{K_{C3}} \Xi(k) \cdot (g_k^j - g_k) \quad (3.25)$$

ここで、 $\Xi(k)$  は第  $k$  番目細胞面で位置の比較を行うべきかを判断する関数で、学習パターンと提示パターンの両方において第  $k$  番目細胞面が発火している時にのみ重心のずれを加算するための関数である。

$$\Xi(k) = \phi \left[ \sum_{\nu \in Z} u_{C3}(\nu, k) \cdot \sum_{\nu \in Z} u_{\text{loc}}(j, \nu, k) \right] \quad (3.26)$$

$K_{\text{fired}}$  は次式で定義される値で、比較すべき細胞面数の総数である。

$$K_{\text{fired}} = \sum_{k=1}^{K_{C3}} \Xi(k) \quad (3.27)$$



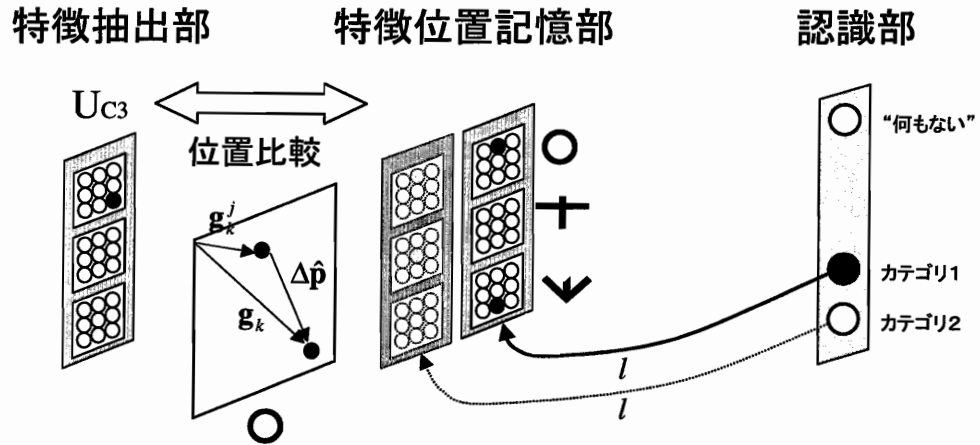


図 3.13: 切り出し位置補正量の算出

図中では、細胞層中で学習パターンと提示パターンの両方において発火している一番上の細胞面のみで重心の比較を行う場合を示す。現在の提示パターンによる発火重心位置  $g_k$  から、学習パターンにおける発火重心位置  $g_k^j$  を引いたものが位置修正量  $\Delta\hat{p}$  となる。

また、 $g_k^j$  は、次式で求められる。

$$g_k^j = \sum_{\nu \in Z} u_{loc}(j, \nu, k) \cdot \nu \quad (3.28)$$

続いて、補正後の切り出し位置  $\hat{p}_1 = \hat{p}_0 + \Delta\hat{p}$  において、再び特徴抽出を行い、特徴種類記憶部を経て認識部において  $\hat{p}_1$  における認識結果  $j'$  を求める。

$j = j'$  の場合には仮認識結果  $j$  は棄却されずにふたたび切り出し位置の補正を進めてゆく。もし、 $n$  回の検証を経て位置補正量  $\Delta\hat{p} = 0$  となったならば、その時点での切り出し位置  $\hat{p}_n$  にカテゴリ  $j$  のパターンが存在すると結論づける。

$j \neq j'$  の場合には仮認識結果は棄却される。すなわち、

$$\rho(j) = 0 \quad (3.29)$$

に設定し、切り出し位置を  $\hat{p}_0$  に戻し仮定の生成をやりなおす。

もし全てのカテゴリにおいて仮定が棄却され、

$$\sum_{j=1}^{K_{learn}} \rho(j) = 0 \quad (3.30)$$

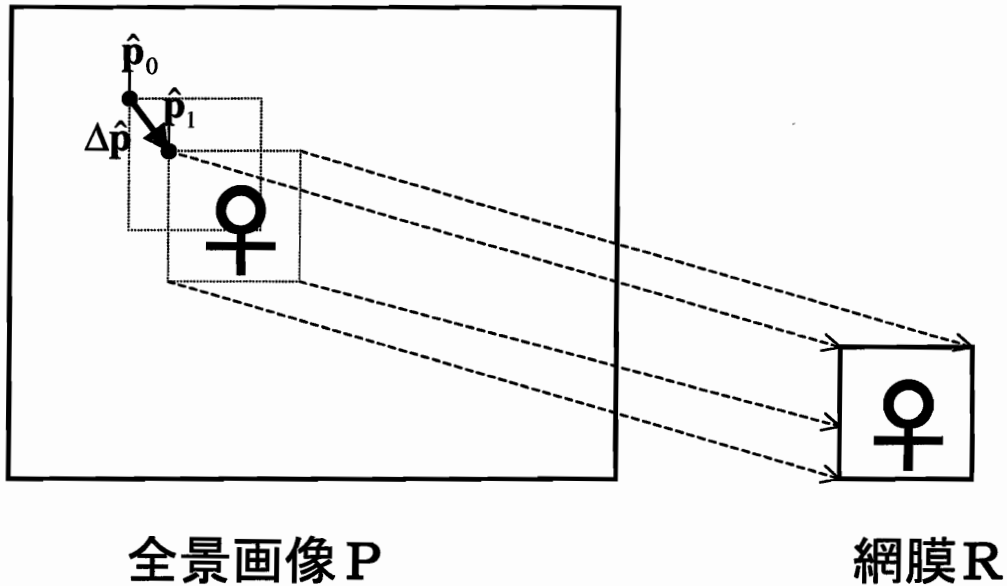


図 3.14: 切り出し位置の修正  
 補正後の切り出し位置  $\hat{p}_1 = \hat{p}_0 + \Delta\hat{p}$  に従い、切りだし位置の修正を行う。

となった場合には、 $\hat{p}_0$  には学習パターンは存在しない、すなわち“未知パターン”が存在すると結論づける。

### 3.3 計算機シミュレーション

仮定と検証に基づく位置・変形に頑健なパターン認識モデルの性能を評価するために、計算機による数値シミュレーションを行う。まず、直交性の高いパターンでの評価実験を行う。次に、類似した特徴を持つパターンでの評価実験を行う。

なお、特徴抽出層であるネオコグニロン部分の実装には宮野らによるネオコグニロンライブラリ [32] を使用した。

#### 3.3.1 直交性の高いパターンでの評価実験

学習パターンとして使用したのは図 3.15 で示す  $16 \times 16$  [pixel] の 3 通りのパターンである。網膜の大きさ  $R \times R$  は学習パターンサイズに合わせる。

実験時の網膜モデルパラメータ値を表 3.2 に、特徴抽出部のパラメータを表 3.3 に、特徴種類記憶部・認識部・特徴位置記憶部に関するパラメータを表 3.4 に示す。

表 3.2: 網膜モデルのパラメータ

パラメータ種類	値
全景画像の大きさ $ P_x  \times  P_y $	$48 \times 48[\text{pixel}]$
網膜の大きさ $R \times R$	$16 \times 16[\text{pixel}]$
最大画素値 $P_{\max}$	255

表 3.3: 特徴抽出部のパラメータ

パラメータ種類	第0段目	第1段目	第2段目	第3段目
細胞面の大きさ $N_l$	$16 \times 16$	$18 \times 18$	$10 \times 10$	$16 \times 16$
結合範囲 $A_l$	—	$3 \times 3$	$5 \times 5$	$5 \times 5$
結合範囲 $D_l$	—	$3 \times 3$	$7 \times 7$	$5 \times 5$
抑制定数 $r_l$	—	3.85	2.35	1.65
学習定数 $q_l$	—	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$
学習後の C 細胞面数 $K_{Cl}$	1	8	13	15

表 3.4: 認識部に関連するパラメータ

パラメータ種類	値
Hebb 則学習定数 $\delta Rec$	3.0
Hebb 則学習回数 $N_{\text{hebb}}$	5
思い込み定数 $I_l$	1.15

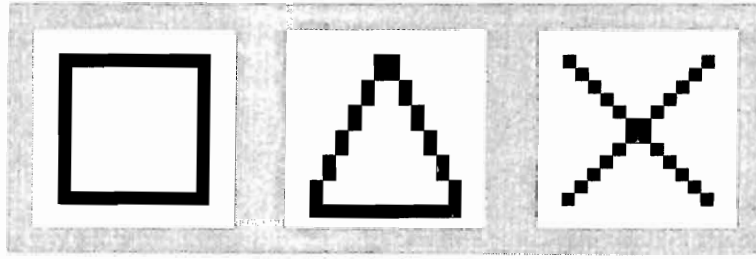


図 3.15: 直交性の高い学習用パターン

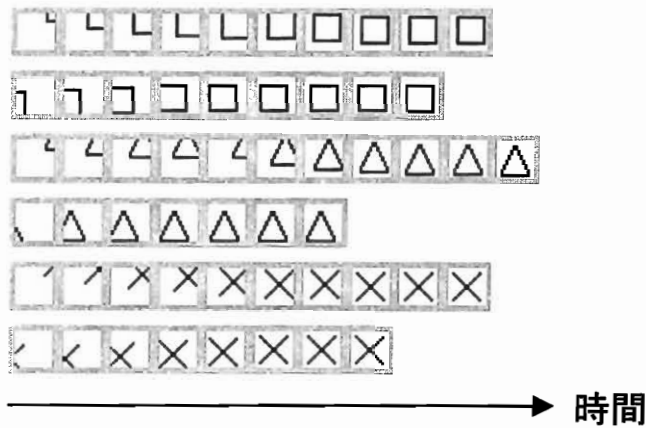


図 3.16: 直交性の高いパターンにおける認識の過程

各段の左端が認識開始時に網膜に入力されたイメージであり、仮定の検証が進むに連れ切りだし位置の修正が行なわれる。

モデル全体の学習後、学習したパターンを中央に配置した  $48 \times 48$  [pixel] の画像を用いて様々な初期切り出し位置  $\hat{p}_0$  から認識を開始する。図 3.16 は、認識の過程の一例である。

また、図 3.17 は、特徴抽出部  $U_{C3}$  層の反応の時系列で、特徴の位置が認識の過程で修正されてゆく様子を示したものである。

図 3.18 で、各パターンにおける正しく切り出し・認識可能な初期切り出し位置  $\hat{p}_0$  を示す。図中で、赤く示した部分が各パターンにおいて正しく切り出し・認識が実行できた初期切り出し位置の中心である。これより、パターンの一部でも入力されていればほとんどの位置で認識可能であることがわかる。

認識で失敗している位置では、初期切り出し位置  $\hat{p}_0$  の画像では特徴抽出部のネオコグニトロンが反応せず未知と判断されてしまっていた。

同一の初期切り出し位置においてパターンが混在している場合の認識過程を図 3.19 で示す。初めに生成された仮定に基づき、余分な特徴情報を使用せずに認識を進めてゆくので、パター

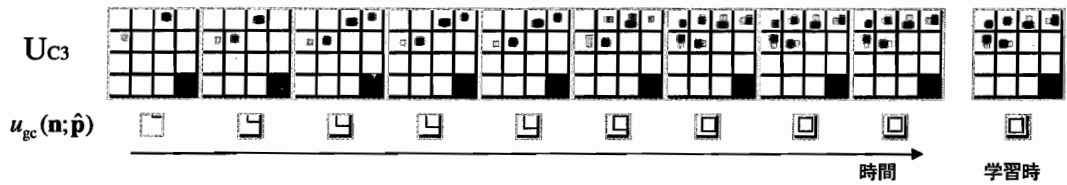


図 3.17: 認識の過程における  $U_{C3}$  層の反応  
 下段が網膜に入力されたイメージ、上段が対応する特徴抽出部  $U_{C3}$  層の反応をまとめて表示したものである。

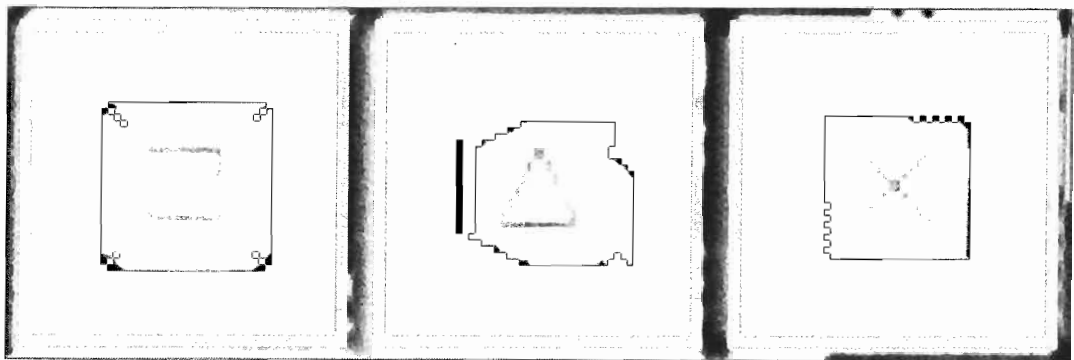


図 3.18: 認識可能な初期切り出し範囲  
 赤で示されている部分は認識に成功した初期切り出し位置の中心点を表し、白の部分は失敗した初期切り出し位置の中心点を表す。

ンが混在していても正しく認識・切り出しが可能である。

### 3.3.2 類似した特徴を持つパターンでの評価実験

学習パターンとして使用したのは図 3.20 で示す  $16 \times 16$  [pixel] の 3 通りのパターンである。実験時の各パラメータの値は表 3.2, 表 3.3, 表 3.4 に示したものと同一値である。

モデル全体の学習後、直交したパターンと同様に学習したパターンを中央に配置した  $48 \times 48$  [pixel] の画像を用いて様々な初期切り出し位置  $\hat{p}_0$  から認識を開始する。図 3.21 は、認識の過程の一例であり、全く同じパターンから開始された場合の認識の過程を示している。

図 3.21 で示されるように、全く同じパターンから開始しても、仮定の生成と検証を活用することで正しく認識可能である。

しかし、本モデルでは初めに仮定候補が決定論的にすべて生成されるている。例えば、この図であれば認識開始時のパターンを元に仮定は右に +, 上に +, 斜めに × のパターンという順で生成されるため、結果として候補順位が低いものに対して認識に時間がかかってしまう。誤っ

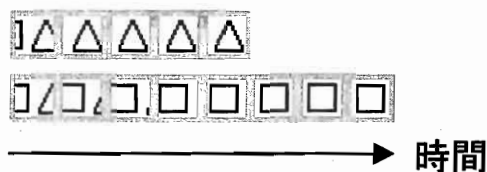


図 3.19: パターン混在時における認識の過程  
各段の左端で示す認識開始時に網膜に入力されたイメージには複数のパターンが混在しているが、仮定を検証し、切りだし位置を補正することで認識可能である。



図 3.20: 類似した特徴を持つ学習用パターン

た仮定の検証途中では開始時よりも多くのパターン情報を得ているので、検証途中での仮定生成や直しなどを加味する必要があるといえる。

また、図 3.22 で、各パターンにおける正しく切り出し・認識可能な初期切り出し位置  $\hat{p}_0$  を示す。図中で、赤く示した部分が各パターンにおいて正しく切り出し・認識が実行できた初期切り出し位置の中心である。この結果から、類似した特徴を持つパターンにおいても、位置ずれに対して非常に頑健であると言える。

認識で失敗している初期切り出し位置では、初期切り出し位置  $\hat{p}_0$  の画像では特徴抽出部のネオコグニトロンが反応せず未知と判断されてしまっていた場合や、仮定の検証途中で仮認識結果が微妙な差で変わってしまい、すべての仮定が誤って棄却され、未知と判断された場合などがあつた。

### 3.4 まとめ

ヒトの視覚情報処理系を参考とし、パターンの全景が網膜範囲内に収まらない場合でも認識可能なパターン認識モデルを提案した。また、計算機シミュレーションにより、実際にパターンの一部分を提示すれば適切に切り出し・認識可能なことが確認された。

しかし、本モデルではある1つのパターンの認識が終了した後に他の部分に注視を向けるこ

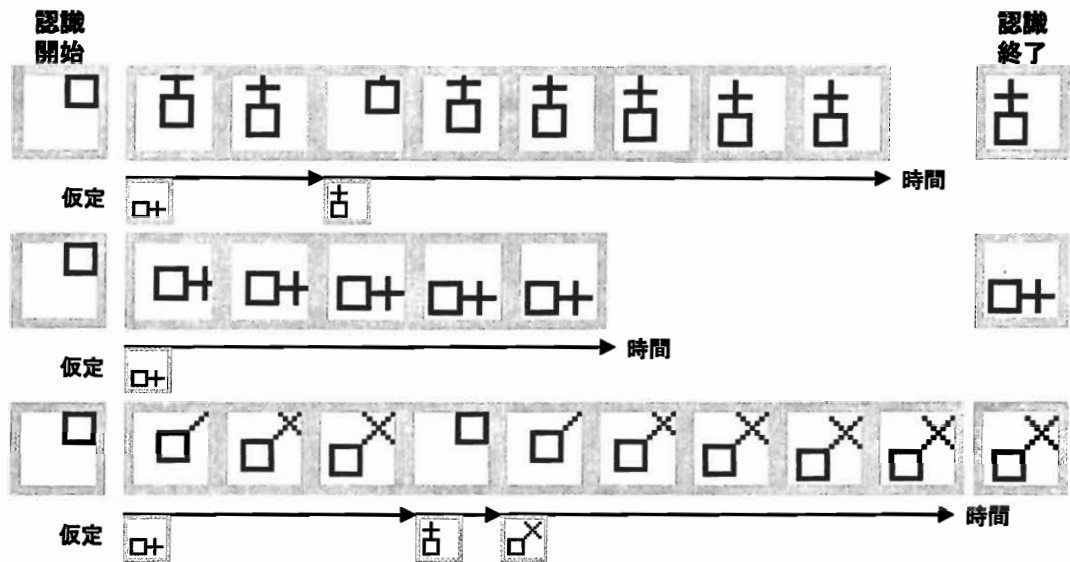


図 3.21: 類似した特徴を持つパターンにおける認識の過程  
各段の左端で示す認識開始時に網膜に入力されたイメージには同一のものであるが、仮定を検証し、切りだし位置を補正することで認識可能である。

とができず、動作が止まってしまう。そこで、第4章では、網膜周辺視野を加味することによって連続的にパターンの認識・切り出しが可能なモデルを提案し、検証する。

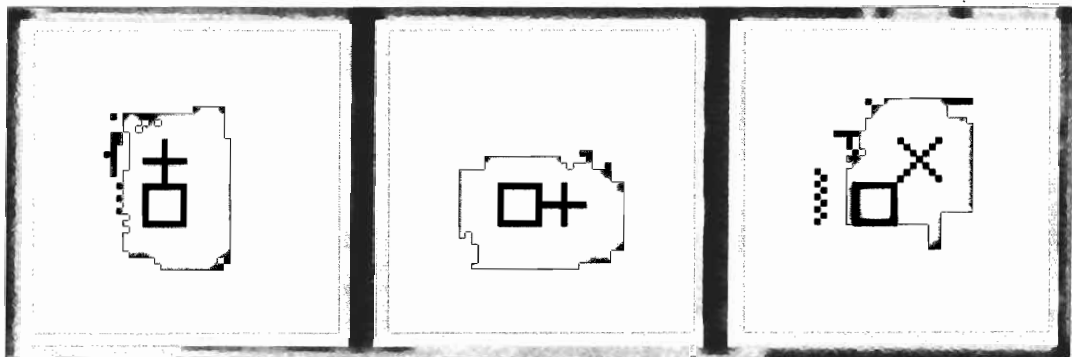


図 3.22: 類似した特徴を持つパターンの認識可能な初期切り出し範囲  
赤で示されている部分は認識に成功した初期切り出し位置の中心点を表し、白の部分は失敗した初期切り出し位置の中心点を表す。



## 第4章

# 多解像度を利用したパターン抽出モデル

この章では、3章で構築したモデルを元にして、パターンが混在した環境下において連続的に効率よくパターンを切り出し・認識、すなわち“パターン抽出”が可能な数理モデルを構築する。また、提案したモデルの計算機シミュレーションを行い有効性を評価し考察を行う。

### 4.1 モデルの概要

第3章で構築したモデルでは、1つのパターンの認識が終了してしまうと他の部分にパターンが存在していても動作が終了してしまっていた。そこで、本章のモデルでは、連続的なパターン抽出を行うために、視覚情報処理系における多解像度性を加味し、1つのパターンを認識後、周辺視野から次のパターンを探し出す動作を組み入れる。

生体上における網膜は解像度が連続的に変化しているが、本モデルでは中心窩に対応する高解像度・局所的な情報を処理する機構と、網膜周辺視野に対応する低解像度・広範囲な情報を処理する機構の2つに単純化してモデル化を行う。そして、高解像度で情報処理を行なう機構を機構を注視機構、低解像度で行なう機構をサッケード機構と呼ぶ。

パターン抽出は、図4.1で示すようにサッケード機構により広範囲からパターンを探し出すサッケードフェーズと、注視機構によりパターンの位置・カテゴリを正確に割り出す注視フェーズの2つのフェーズを逐次的に実行しながら行われる。

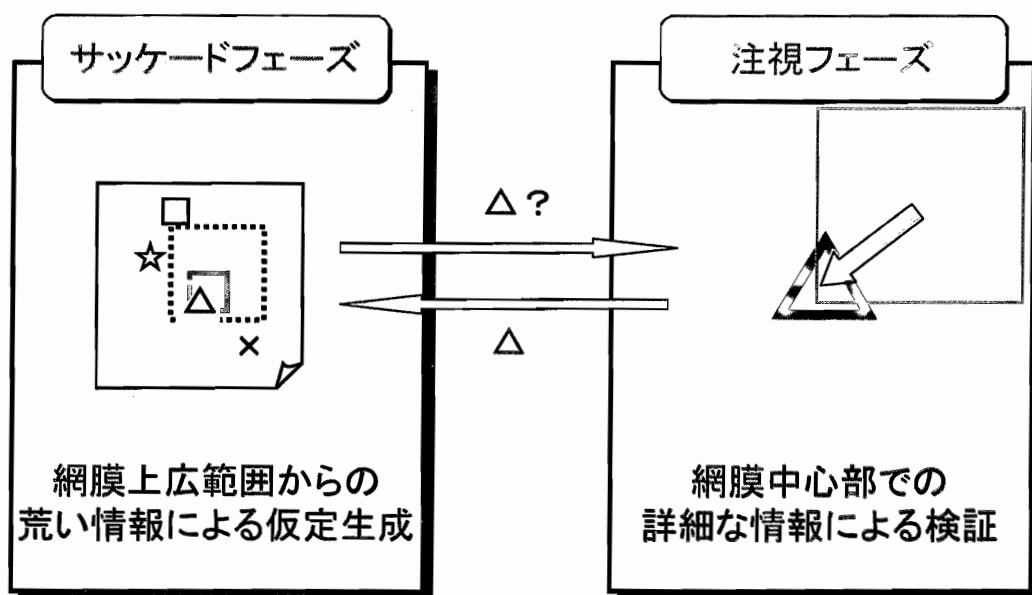


図 4.1: パターン抽出における2つのフェーズ

パターン抽出では、サッケード機構により広範囲からパターンを探し出すサッケードフェーズと、注視機構によりパターンの位置・カテゴリを正確に割り出す注視フェーズの2つのフェーズを逐次的に実行する。

## 4.2 モデルの構成

本モデルは、図 4.2 で示すように、サッケード機構と注視機構の2つの機構で構成される。各機構の構造は、基本的には第3章で構築したモデルと同じである。

ここでは、第3章のモデルとの相違点を説明する。定式化にあたり本章で使用する記号は、3章で用いた記号の右上にサッケード機構 ( $L$ ) または注視機構 ( $H$ ) を表す ( $L$ ) または ( $H$ ) を付けたものとする。

例えば、サッケード機構の特徴抽出部における  $U_{C3}$  層細胞面は  $U_{C3}^{(L)}$ 、注視機構における認識部のパターンカテゴリ  $j$  に対応する細胞は  $U_{rec}^{(H)}(j)$ 、その出力を  $u_{rec}^{(H)}(j)$  などと表す。

### 網膜モデル

網膜モデルにおける相違点は、抽出された領域に対して抑制をかけることで同じパターンが抽出され続けることを防ぐ点である。これを実現するために、図 4.3 のように全景画像と同じ大きさの抑制マスクを新たに導入する。抑制マスクの座標  $p = (x, y)$  における値を  $u_{mask}(p)$  とする。 $u_{mask}(p)$  は

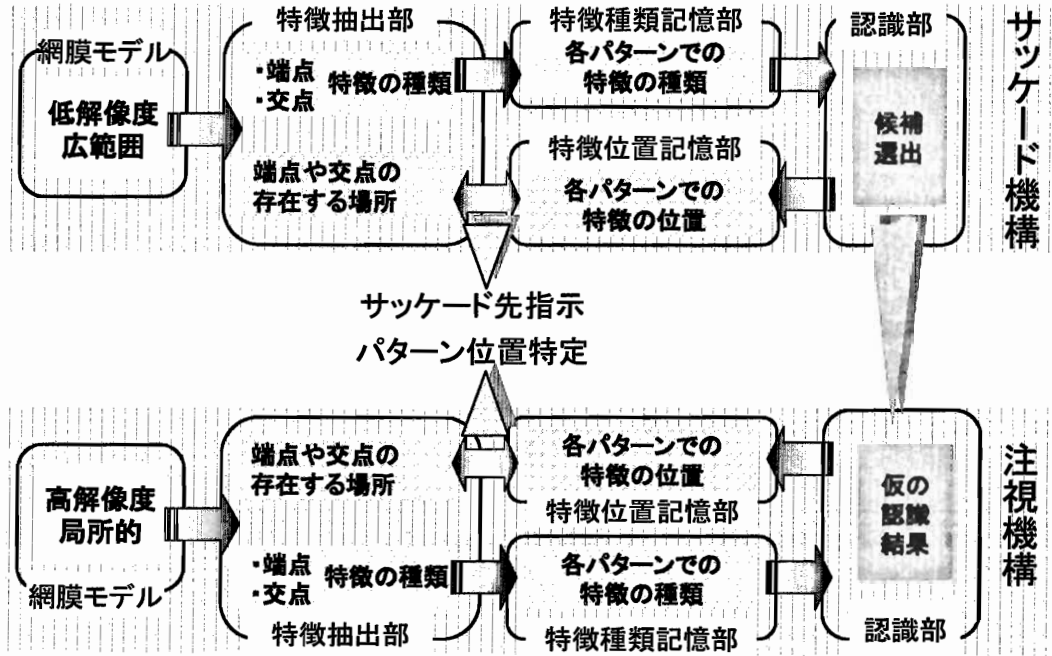


図 4.2: パターン抽出モデルの概要

$$u_{\text{mask}}(\mathbf{p}) = \begin{cases} 0; & \text{If there is a pattern extracted on position } \mathbf{p}. \\ 1; & \text{otherwise.} \end{cases} \quad (4.1)$$

と設定され、これを用いることで、注視機構の網膜出力  $u_{\text{gc}}^{(H)}(\mathbf{n}; \hat{\mathbf{p}})$  は次式で与えられる。

$$u_{\text{gc}}^{(H)}(\mathbf{n}; \hat{\mathbf{p}}) = \begin{cases} u_{\text{mask}}(\mathbf{n} + \hat{\mathbf{p}}) \cdot \frac{u_{\text{p}}(\mathbf{n} + \hat{\mathbf{p}})}{P_{\text{max}}} & ; \text{if } \mathbf{n} + \hat{\mathbf{p}} \in P \\ 0 & ; \text{otherwise} \end{cases} \quad (4.2)$$

また、サッケード機構において全景画像から低解像度画像への変換を行う役割を追加する。いま、簡単のため網膜の大きさは、両解像度において等しい、すなわち  $R^{(L)} = R^{(H)}$  とする。サッケード機構で解像度を全景画像から  $\frac{1}{C_d}$  に落として使用する場合、サッケード機構の網膜出力  $u_{\text{gc}}^{(L)}(\mathbf{n}; \hat{\mathbf{p}})$  は次式で与えられる。

$$u_{\text{gc}}^{(L)}(\mathbf{n}; \hat{\mathbf{p}}) = \begin{cases} u_{\text{mask}}(C_d \cdot \mathbf{n} + \hat{\mathbf{p}}) \cdot \frac{u_{\text{p}}(C_d \cdot \mathbf{n} + \hat{\mathbf{p}})}{P_{\text{max}}} & ; \text{if } C_d \cdot \mathbf{n} + \hat{\mathbf{p}} \in P \\ 0 & ; \text{otherwise} \end{cases} \quad (4.3)$$

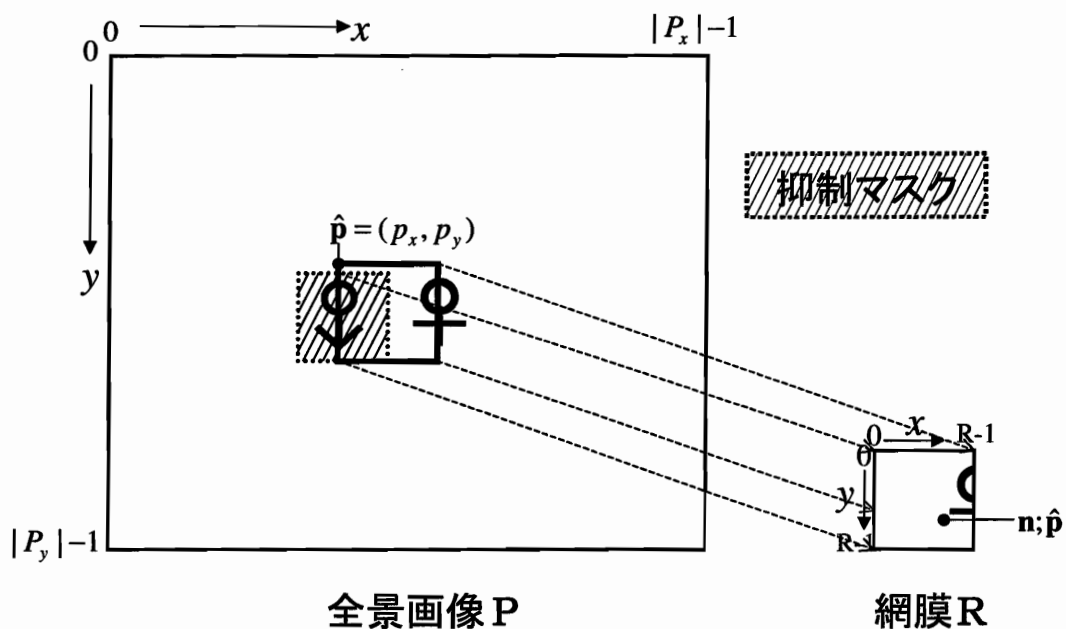


図 4.3: 網膜モデルにおける抑制マスクの働き

図中で、斜線が引かれている領域が抑制マスク  $u_{\text{mask}}(\mathbf{p}) = 0$  の部分である。抑制がかかっている斜線の領域では、網膜にパターンが射影されない。

#### 特徴抽出部

相違点はない。

#### 特徴種類記憶部

相違点はない。

#### 認識部

注視機構の認識部に相違点はない。

サッケード機構においては仮認識結果が棄却される事がないので常に

$$\rho^{(L)}(j) = 1 \tag{4.4}$$

となる。また、思い込みも生じないため、

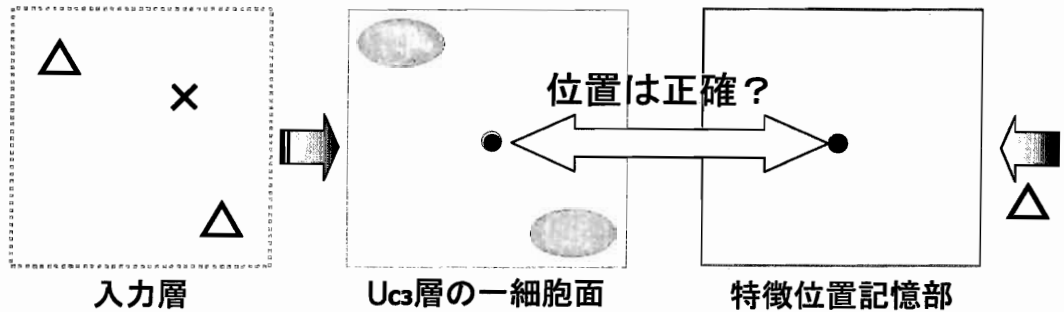


図 4.4: 発火重心位置の計算における問題点

多数のパターンが切り出し範囲内に存在すると、パターンが範囲内の両端に存在している場合などに重心位置が中央に計算されるなどの不都合が生じる。

$$I_i^{(L)} = 1 \tag{4.5}$$

である。

### 特徴位置記憶部

注視機構の認識部に相違点はない。

サッケード機構においては、切り出し範囲内にパターンが数多く混入するが発生する。このとき、式(3.22)に従って発火重心を計算すると、図4.4のように不正確な重心が算出される場合がある。

そこで、サッケード機構では、発火重心位置の計算を行う前に特徴抽出部の  $U_{C3}^{(L)}$  層において細胞面内で競合を発生させる。これにより、競合前には複数の位置で発火が起きていても競合後には図4.5のように最大発火している場所のみが発火を続けることになり、正確な位置の補正を実現する。

モデル中では競合を最大値検出という形で代数的に実行する。すなわち、 $U_{C3}^{(L)}$  層の競合後の細胞出力  $u_{C3}^{(L)'}(\mathbf{n}, k)$  は次式で与えられる。

$$u_{C3}^{(L)'}(\mathbf{n}, k) = \begin{cases} u_{C3}^{(L)}(\mathbf{n}, k) & ; \text{if } \mathbf{n} = \arg \max_{\nu \in Z} u_{C3}^{(L)}(\nu, k) \\ 0 & ; \text{otherwise} \end{cases} \tag{4.6}$$

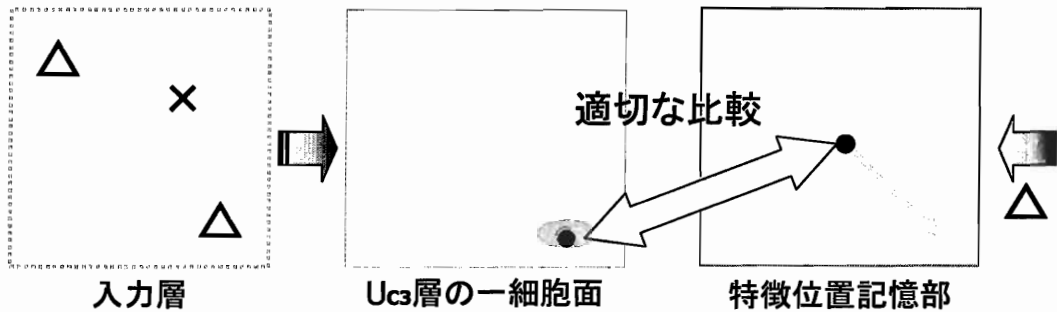


図 4.5: 細胞面内における競合後の発火重心位置

競合後の細胞面では、最大発火している部分だけが発火を続けるため、重心位置を適切に求めることができる。

### 4.3 学習方法

モデルの学習は2つの機構で独立に行われる。各機構の具体的な学習方法は第3章で述べたものと変わらない。

ただし、特徴抽出部の抑制定数は2つの機構において適切に調節する必要がある。図4.6で示すように、学習時に抑制定数を小さく設定すると、パターンの大局的な特徴が抽出される傾向が強まり、細胞面数は減少する。逆に、抑制定数を大きく設定すると、パターン中の局所的な特徴が抽出される傾向が強まり、細胞面数は増大する。

これをモデル全体の振る舞いに置き換えれば、抑制定数が小さいと切り出しの精度は落ちるが許容誤差は増えることになる。逆に、抑制定数が大きいと切り出し精度は上がるが許容誤差は小さくなる。このことから、サッケード機構においては抑制定数を小さく、注視機構においては抑制定数を大きく設定する必要がある。

実際には、サッケード機構においては  $U_{C3}^{(L)}$  層の学習終了後の細胞面数  $K_{C3}^{(L)}$  をパターンカテゴリ数  $K_{\text{learn}}$  と同数程度になるように設定し、注視機構においては  $U_{C3}^{(H)}$  層の学習終了後の細胞面数  $K_{C3}^{(H)}$  をパターンカテゴリ数  $K_{\text{learn}}$  の3~4倍程度になるように設定する。

### 4.4 認識方法

認識では、まずサッケードフェーズにおけるサッケード先の指示が生成され、続いて注視フェーズにおいてパターンの抽出が行なわれる。

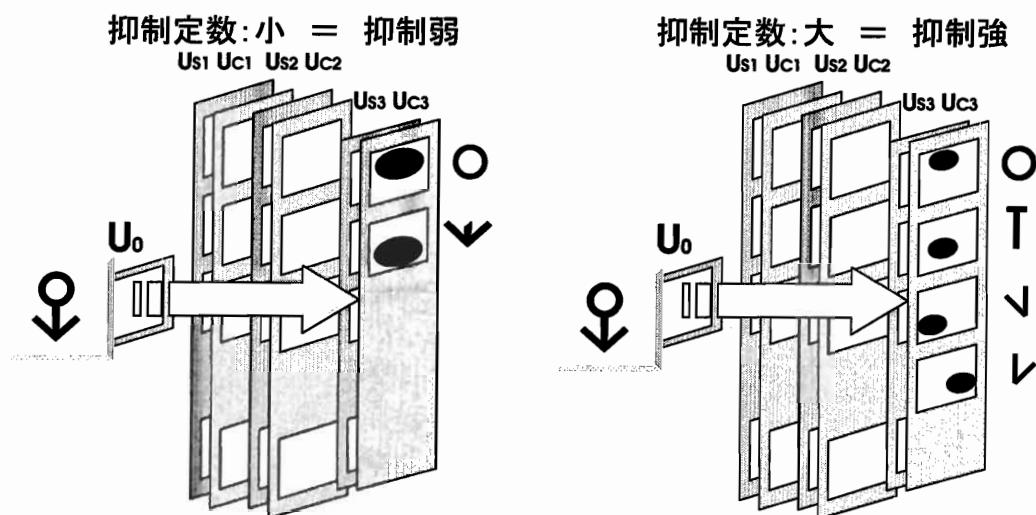


図 4.6: 抑制定数と細胞の反応

図左側のように学習時に抑制定数を小さく設定すると、パターンに含まれる大局的な特徴が抽出される傾向が強まり、生成される細胞面数は減少する。逆に、図右側のように抑制定数を大きく設定すると、パターン中の端点や交点など局所的な特徴が抽出される傾向が強まり、生成される細胞面数は増大する。

### サッケードフェーズ

$n$  回目のフェーズにおいて、サッケード機構の網膜モデルでの切り出し位置  $\hat{p}_n^{(L)}$  で式 (4.3) に従い広範囲の低解像度画像が生成され、特徴抽出部を経て認識部において候補選出が行われたとする。いま、選出された候補をカテゴリ  $j$  とする。この候補  $j$  は注視機構の認識部に送られ、仮認識結果として使用される。同時に、サッケード機構の特徴位置記憶部におけるカテゴリ  $j$  の発火重心位置と、現在の発火重心位置から切り出し位置の修正量  $\Delta\hat{p}^{(L)}$  を算出する。

このとき、サッケード先  $\hat{p}_0^{(H)}$  は次式によって決定される。

$$\hat{p}_0^{(H)} = \hat{p}_n^{(L)} + C_d \cdot \Delta\hat{p}^{(L)} \quad (4.7)$$

また、サッケード機構の認識部において“何もない”細胞が発火した場合には、もう抽出すべきものは無いと判断し、パターンの抽出を終了する。

### 注視フェーズ

注視フェーズでは、切り出し位置  $\hat{p}_0^{(H)}$  に存在する仮認識結果  $j$  のパターンを検証する。

検証の過程で仮定が棄却された場合には、切り出し位置  $\hat{p}_0^{(H)}$  において注視機構で仮定生成が行われ、検証へと移る。

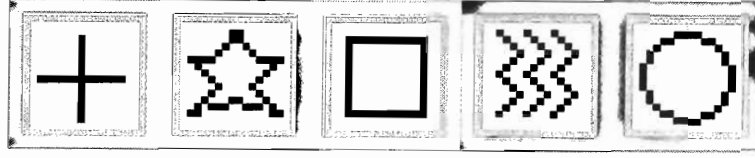


図 4.7: 学習用パターン

各パターンの枠の色は結果表示時におけるパターンの抽出範囲に対応する

注視機構において  $m$  回後の位置補正の後に補正が終了し、位置  $\hat{p}_m^{(H)}$  にカテゴリ  $j$  のパターンが存在すると認識された場合には、次式に従い抑制マスクを変更する。

$$u_{\text{mask}}(\mathbf{p}) = 0; \text{ if } \mathbf{p} \in \hat{\mathbf{p}}_0^{(H)} + \mathbf{R}^{(H)} \quad (4.8)$$

もしも  $m$  回後の位置補正の後に全てのカテゴリにおいて仮定が棄却された場合には、位置  $\hat{p}_m^{(H)}$  には“未知パターン”が存在する可能性があるとして判断し、一時的に式 (4.8) に従い抑制マスクを変更する。ただし、この変更は次の  $n+1$  回目のサッケードフェーズに限って有効であり、 $n+2$  回目には  $u_{\text{mask}}(\mathbf{p}) = 1$  に戻して、誤って未抽出の学習パターンが抑制マスクで抑制されることを防ぐ。

抑制マスクの変更が終了したら、サッケード開始点  $\hat{p}_{n+1}^{(L)} = \hat{p}_m^{(H)}$  に設定し、再びサッケードフェーズへと戻る。

また、3回以上同じ位置  $\hat{p}_k^{(H)}$  で全ての仮定が棄却され、“未知パターン”が存在する可能性があるとして判断された場合には、その位置に“未知パターン”が存在すると結論付け恒久的に式 (4.8) に従い抑制マスクを変更する。

## 4.5 計算機シミュレーション

多解像度を利用したパターン抽出モデルの性能を評価するために、計算機による数値シミュレーションを行う。まず、抽出するパターンが学習パターンのみの場合についての評価実験を行い、次に未知パターンを含む場合の評価実験を行う。

学習パターンとして使用したのは図 4.7 で示す  $16 \times 16[\text{pixel}]$  の 5 通りのパターンである。左から順に、“十字”、“星”、“四角”、“波”、“円” と呼ぶことにする。



表 4.1: 網膜モデルのパラメータ

パラメータ種類	値
全景画像の大きさ $ P_x  \times  P_y $	$80 \times 80[\text{pixel}]$
最大画素値 $P_{\max}$	255
サッケード機構 網膜の大きさ $R^{(L)} \times R^{(L)}$	$16 \times 16[\text{pixel}]$
サッケード機構 画像縮小率 $C_d$	5.0
注視機構 網膜の大きさ $R^{(H)} \times R^{(H)}$	$16 \times 16[\text{pixel}]$

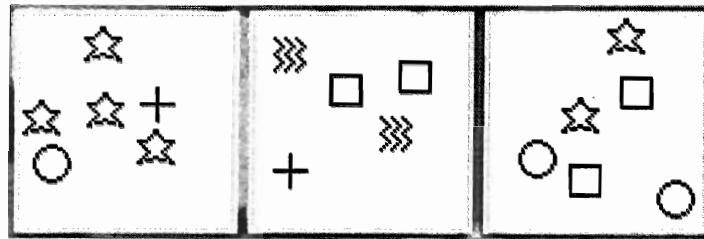


図 4.8: 学習パターンのみから構成されるテストパターンの一例

#### 4.5.1 未知パターンを含まない場合の評価実験

##### 実験条件とパターン抽出例

実験時の網膜モデルパラメータ値を表 4.1 に、特徴抽出部のパラメータを表 4.2 に、特徴種類記憶部・認識部・特徴位置記憶部に関するパラメータを表 4.3 に示す。

モデルの学習後、学習パターンをランダムに配置した  $80 \times 80[\text{pixel}]$  の画像を用いてパターンの抽出を行う。このとき、サッケード機構の初期切り出し位置  $\hat{p}_0^{(L)} = (0, 0)$  とする。図 4.8 は、使用したテストパターンの一例である。

実験では、サンプル 10000 枚に対してパターン抽出を行った。図 4.9 で、パターン抽出に成功したサンプルの例を示す。図中でそれぞれ色が塗られている範囲がパターン抽出された場所を示し、各色は図 4.7 の各パターンの枠の色に対応することでカテゴリ分けの結果を表す。また、灰色の部分はその間に未知パターンが存在すると抽出された区域を示す。適切な場所で適切なカテゴリにパターン抽出を行えることが実証された。

図 4.10 では、パターン抽出に失敗したサンプルの例を示す。図 4.10 の色分けも図 4.9 と同様の意味を表す。

抽出に失敗した例では、一つのパターンがずれて抽出されてしまったためにその周囲のパターンに不必要な抑制がかかり未知パターンと抽出されてしまったものや、2つの並んだパターンの間で未知パターンと抽出されたために失敗したものなどが存在した。また、パターン“四

表 4.2: 特徴抽出部のパラメータ

	パラメータ種類	第0段目	第1段目	第2段目	第3段目
サッケード機構	細胞面の大きさ $N_l^{(L)}$	16 × 16	18 × 18	16 × 16	16 × 16
	結合範囲 $A_l^{(L)}$	—	3 × 3	5 × 5	5 × 5
	結合範囲 $D_l^{(L)}$	—	3 × 3	7 × 7	5 × 5
	抑制定数 $r_l^{(L)}$	—	2.40	1.20	0.80
	学習定数 $q_l^{(L)}$	—	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$
	学習後の C 細胞面数 $K_{Cl}^{(L)}$	1	8	9	6
注視機構	細胞面の大きさ $N_l^{(H)}$	16 × 16	18 × 18	10 × 10	16 × 16
	結合範囲 $A_l^{(H)}$	—	3 × 3	5 × 5	5 × 5
	結合範囲 $D_l^{(H)}$	—	3 × 3	7 × 7	5 × 5
	抑制定数 $r_l^{(H)}$	—	3.20	2.00	1.00
	学習定数 $q_l^{(H)}$	—	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$
	学習後の C 細胞面数 $K_{Cl}^{(H)}$	1	8	28	20

表 4.3: 認識部に関連するパラメータ

	パラメータ種類	値
サッケード機構	Hebb 則学習定数 $\delta Rec^{(L)}$	3.0
	Hebb 則学習回数 $N_{hebb}^{(L)}$	5
	思い込み定数 $I_l^{(L)}$	1.15
注視機構	Hebb 則学習定数 $\delta Rec^{(H)}$	3.0
	Hebb 則学習回数 $N_{hebb}^{(H)}$	5
	思い込み定数 $I_l^{(H)}$	1.15

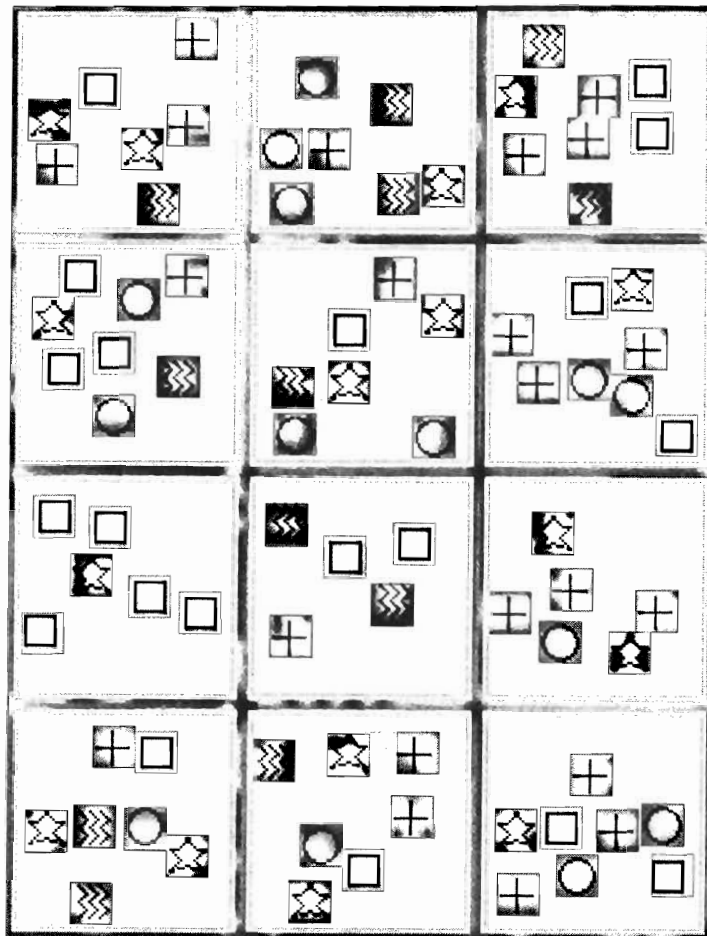


図 4.9: パターン抽出成功例 (学習パターンをみの場合)

角”がずれて入力されたときに、パターン“十字”と誤って抽出してしまうサンプルが多く見られた。

#### 抽出率と抽出ステップ数による評価

モデルの定量的評価を行うために、抽出成功率と抽出ステップ数という二つの尺度を定義する。

抽出率は、サンプルに含まれるパターンをどれだけ正確に抽出できたかを測る尺度である。いま、下記の項目のように抽出精度を3段階で定義する。

- *Correct*

適切なカテゴリに学習パターンとの位置ずれ精度  $\pm 2[\text{pixel}]$  以内で抽出できた場合

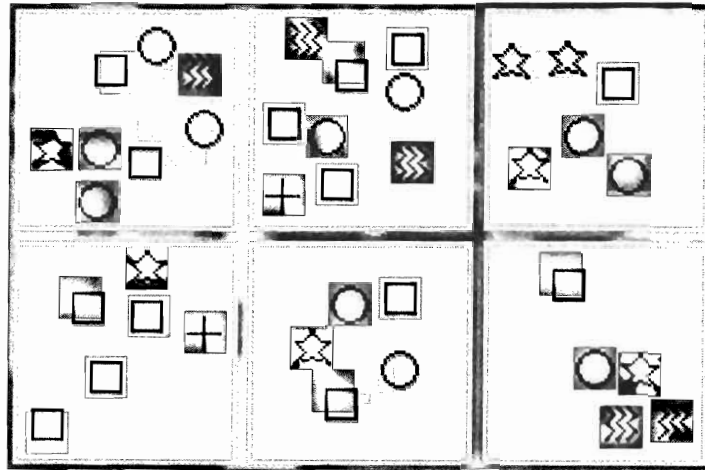


図 4.10: パターン抽出失敗例 (学習パターンのみの場合)

- *Poor*  
適切なカテゴリに学習パターンとの位置ずれ精度  $\pm 6[\text{pixel}]$  以内で抽出できた場合
- *Bad*  
カテゴリを誤って抽出するか、位置ずれが  $\pm 6[\text{pixel}]$  以上で抽出された場合

このとき、それぞれの精度における抽出率は次式で定義される。

$$Ratio_{\text{Correct}} = \frac{\text{Correct 総数}}{\text{パターン総数}} \times 100(\%) \quad (4.9)$$

$$Ratio_{\text{Poor}} = \frac{\text{Poor 総数}}{\text{パターン総数}} \times 100(\%) \quad (4.10)$$

$$Ratio_{\text{Bad}} = \frac{\text{Bad 総数}}{\text{パターン総数}} \times 100(\%) \quad (4.11)$$

また、 $Ratio_{\text{Correct}}$  を抽出成功率と呼ぶ。

抽出ステップ数は、1つのパターンを抽出するときに何回切り出し位置の補正を行ったかを表し、ステップ数が小さいほど効率よく抽出が行えたことを表す。

表 4.4 に、10000 サンプルの画像を処理した結果を示す。平均抽出成功率は 90.3%、平均抽出ステップ数は 6.65 ステップであった。

この結果を見ると、“四角”において抽出成功率が低い。これは、上記の失敗例でも述べたように、位置のずれた“四角”を“十字”と分類してしまう傾向が見られたためである。また、“十字”の平均抽出ステップ数が他と比べて多いのは、抽出時に“四角”という仮定を検証し、棄却することが多いためと考えられる。

表 4.4: 未知パターンを含まない場合の抽出率と平均抽出ステップ数

項目	“十字”	“星”	“四角”	“波”	“円”
$Ratio_{Correct}$	94.3	86.8	74.7	99.2	96.5
$Ratio_{Poor}$	0.2	0.0	7.8	0.3	0.0
$Ratio_{Bad}$	5.5	13.2	17.5	0.5	3.5
平均抽出ステップ数	9.76	5.21	6.88	6.08	5.32

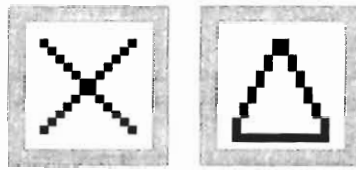


図 4.11: 混入した未知パターン

#### 4.5.2 未知パターンを含む場合の評価実験

##### 実験条件とパターン抽出例

実験時の各パラメータの値は学習パターンのみの場合と同様である。モデルの学習後、学習パターンの他に図 4.11 で示す 2 種のパターンをランダムに配置した  $80 \times 80$  [pixel] の画像を用いてパターンの抽出を行う。このときもサッケード機構の初期切り出し位置  $\hat{p}_0^{(L)} = (0, 0)$  とする。図 4.12 は、使用したテストパターンの一例である。

実験では、サンプル 10000 枚に対してパターン抽出を行った。図 4.13 で、パターン抽出に成功したサンプルの例を示す。未知パターンが混入していても、学習パターンを正しく抽出可能なことがわかる。また、未知パターン領域も部分的に確定できている。

図 4.14 では、パターン抽出に失敗したサンプルの例を示す。抽出に失敗した例では、学習パ

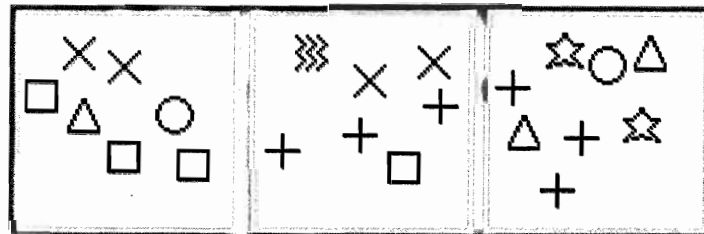


図 4.12: 未知パターンを含むテストパターンの一例

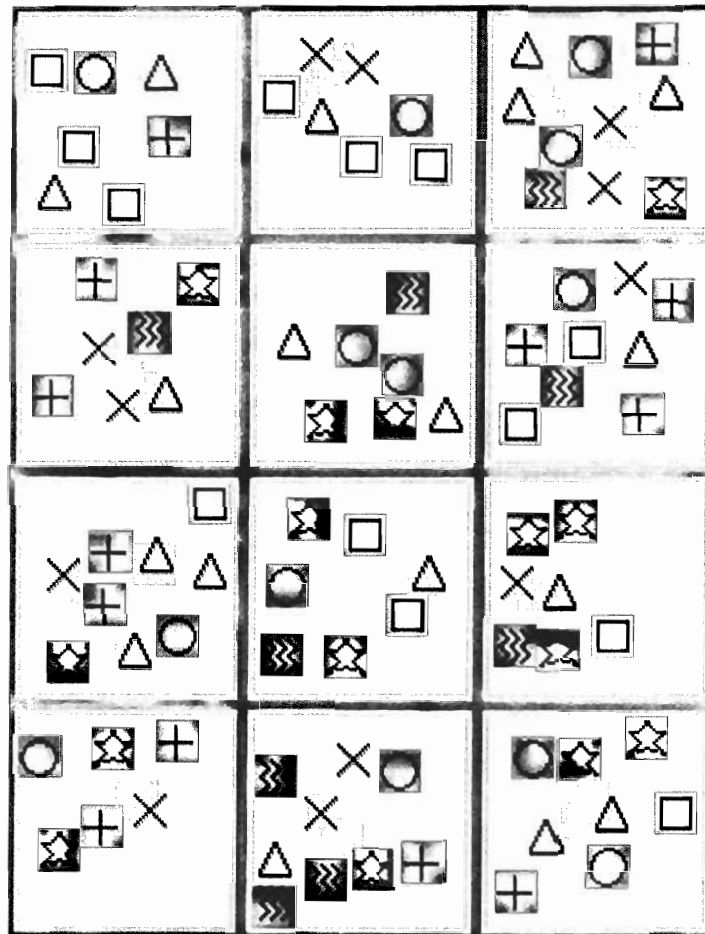


図 4.13: パターン抽出成功例 (未知パターンを含む場合)

ターンに未知パターンが接近したためその中間を誤って未知パターンとして抽出してしまったものが多かった。また、“四角”がずれて“十字”と誤る例もやはり見られた。

#### 抽出率と抽出ステップ数による評価

未知パターンを含む場合においても、抽出率と抽出ステップ数により評価を行った。表 4.5 に、10000 サンプルの画像を処理した結果を示す。

平均抽出成功率は 88.8%、平均抽出ステップ数は 6.76 ステップであった。

結果の傾向は、学習パターンのみの場合と全く変わらず、“四角”において抽出成功率が低く“十字”の平均抽出ステップ数が多いというものであった。

注目すべき点は、未知パターンを含む場合と学習パターンのみの場合において平均抽出成功

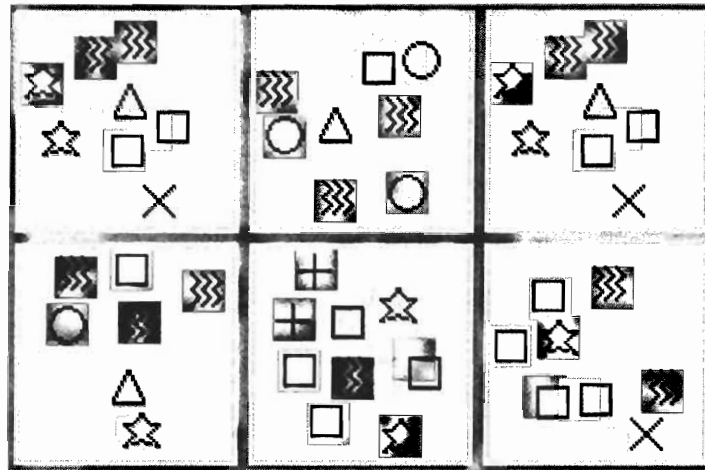


図 4.14: パターン抽出失敗例 (未知パターンを含む場合)

表 4.5: 未知パターンを含む場合の抽出率と平均抽出ステップ数

項目	“十字”	“星”	“四角”	“波”	“円”
<i>RatioCorrect</i>	93.9	84.2	72.3	98.6	95.0
<i>RatioPoor</i>	0.2	0	8.5	0.7	0
<i>RatioBad</i>	5.9	15.8	19.2	0.7	5.0
平均抽出ステップ数	9.76	5.32	7.03	6.27	5.42

率、平均抽出ステップ数ともにほとんど変化が見られないことである。すなわち、本モデルでは未知パターンの混入にも頑健であると言える。

## 4.6 まとめ

低解像度の周辺視野を組み入れることで、連続的にパターンの抽出が可能なモデルを提案した。また、計算機シミュレーションにより、学習パターンのみが存在するサンプルおよび未知パターンを含むサンプルの場合に分けて実験を行い、パターンの抽出が可能なことを確認した。シミュレーション結果からは、未知パターンの混入に影響されにくいモデルであることが示唆された。

# 第5章

## 結論

### 5.1 本研究の成果

本研究は、様々な特長を持つヒトの視覚情報処理系を参考とすることで、次のようなパターン認識を可能とする数理モデルを構築することを目的とした。

- 種々のパターンが混在した広い範囲を視点を移しながら認識
- 視界をはみ出しパターンの一部しか見えていない場合でも視点を移すことでパターンの位置ずれに対応
- パターンの変形・ノイズの混入に対する頑健性も保持

以下、各章で得られた結果を示す。

**第3章** ヒトの視覚情報処理系において生理学的・心理学的に解明されている事を基に、パターンの一部のみ網膜モデルに入力されている場合でもパターンを正しく切り出し・認識を行うことのできる視覚情報処理系の数理モデルを提案した。このモデルは、パターンに含まれる特徴の種類と位置を分離して使用し、「仮定と検証」のプロセスを経て最適な認識結果を得ることで、位置ずれに対して頑健な認識を実現する。計算機シミュレーションの結果、実際にパターンの位置ずれに対する頑健性が確認された。

**第4章** 第3章のモデルを元に、視覚情報処理系の多解像度性を加味することで連続的なパターンの抽出を実現するモデルを提案した。また、計算機シミュレーションにより、適切にパターンの抽出が行えることが実証された。特に、未知パターンの混入に影響されにくいという性質が確認された。



## 5.2 今後の課題

本研究では、主に位置ずれに対する頑健性を考慮に入れモデルの作成を行った。しかし、他にも2次元パターンの多様性としては拡大縮小や回転なども考えられる。そこで、本モデルの考え方を受け継ぎ、まず多様性に対する不変的特徴量を用いて仮定を生成し、仮定を検証する段階で徐々に変動する特徴量を取り入れるという方法で、拡大・縮小・回転等への対応を検討したい。

その際に考えなくてはならないのが、モデルパラメータの自己組織的な生成である。例えば、本モデルでは解像度別に独立して学習を行っているが、拡大縮小に対応するには任意の解像度においてモデルを生成する必要がある。そこで、高解像度の学習パターンによる任意の解像度のモデルパラメータを自動生成する仕組みが必要になると思われる。

また、現在は特徴抽出部にネオコグニトロンを用いているため、認識対象の特徴が抽出できるのは細線化されたパターンである場合に限られている。そこで、網膜モデルをより生体の網膜に近づけ特徴抽出部にも工夫を行うことで、自然画像中からの特徴抽出が可能なモデルの構築を行いたい。そして、認識対象を実際の物体とすることで柔軟なコンピュータ・ビジョンへの応用へとつなげてゆきたい。

現在のパターン認識では、結果の正否は出力されたカテゴリと学習パターンにおけるカテゴリを比較して出すというのが一般的である。

しかし、本来“認識”という活動がなぜ必要なのかを考えると、外界から得た情報を基にして外界の様相を知覚することで生き抜くために最前の行動を生成するという目的が大きい。すなわち、“認識”の結果を評価する時には、単にパターンのカテゴリ分けの成功率などで評価するのではなくて、認識を発端として生成された行動の有効性で評価するべきだと思われる。つまり、**身体性のある認識**を考えてゆくことが、真に“認識”を理解していく上で必要不可欠となってくるものと思われる。

そこで、工学的に身体性を実現する手段であるロボットに着目し、ロボットビジョンにヒトの視覚情報処理系を参考とした機構を組み入れ、行動生成を結果出力とした認識系の構築に関する研究も進めてゆきたいと考えている。

## 参考文献

- [1] 飯島:“パターン認識理論”, 森北出版株式会社,1989.
- [2] M Minsky.The Society of Mind. Simon&Schuster,1986. 心の社会, 安西祐一郎訳, 産業図書.
- [3] 石井健一郎, 上田修功, 前田英作, 村瀬洋:“わかりやすいパターン認識”, オーム社出版局,1998.
- [4] W.C.McCulloch and W.H.Pitts: “A logical calculus of the ideas immanent in neural nets.”,Bull.Math.Biophys.,5,pp.115-133,1943.
- [5] K.Fukushima: “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”,Biological Cybernetics,36,pp.193-202,1980.
- [6] J.basak and S.K.P:“Psyncop—a psychologically motivated connectionist system for object perception”,IEEE Trans.Neural Networks,6,6,pp.1337-1354,1995.
- [7] A.L.Yarbus:“Eye Movement and Vision”,Plenum Press,1967.
- [8] Noton D.and Stark L.:“Scanpaths in saccadeic eye movements while viewing and recognizing patterns”,Vision Res.,11,pp.929-942,1971.
- [9] Schwartz E.L.:“Spatial mapping in the primate sensory projection:Analytic structure and relevance to perception”,Biological Cybernetics,25,pp.181-194,1977.
- [10] 青西亮, 福島邦彦:“網膜と皮質の不均一性を考慮した注視点移動モデル”, 電子情報学会論文誌,J78-D-II,9,pp.1363-1371,1995.
- [11] 山本裕之,Y.Yeshurun,Martin D.Levine:“中心窩センサを用いた視覚システム”, 電子情報学会論文誌,J77-D-II,1,pp.119-130,1994.

- [12] Luiz Pessoa, Sergio Exel, Alexandre Roque, and Ana Leitão: "Attentive Visual Recognition for Scene Exploration", Proc.Int.Conf. on Neural Information Processing, pp.1291-1294, 1998.
- [13] 岩崎正宏, 橋山知訓, 大熊繁: "注意機能に基づいた特徴抽出法の一提案と連想想起への応用", 電気学会論文誌C分冊, Vol.119-C, pp1461-1467, 1999.
- [14] Neisser, U.: "Cognition and reality: principles and implication of cognitive psychology", 1986.
- [15] C.Stanfill and D.Waltz: "Toward memory-based reaspmomg", Communication ACM ,29 ,pp.1213-1228 ,1986.
- [16] Shunji Satoh, Shogo Miyake, Hiroto Aso: "Evaluation of Two Neocognitron-type Models for Recognition of Rotated Patterns", Proc.Int.Conf. on Neural Information Processing, Vol.1, pp.295-299, 2000.
- [17] A.M.Treisman and G.Gelade: "A fracture-integration theory of attention", Cognitive Psychology, 12, pp.97-136, 1980.
- [18] 福島邦彦: "神経回路と情報処理", 朝倉書店, 1989.
- [19] D.H.Hubel and T.N.Wiesel: "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", Journal of Physiology, 160, pp.106-154, 1962.
- [20] D.H.Hubel and T.N.Wiesel: "Receptive fielsa and functional architecture of monkey striate cortex", Journal of Physiology, 195, pp.215-243, 1968.
- [21] Ungerleider, L.G., and Mishkin: "Two cortical visual systems", MIT Press, 1982.
- [22] N.R.Carlson: "Physiology of behavior", Allyn and Barcon, 1998.
- [23] K.Tanaka, H.Saito, Y.Fukada and M.Moriya: "Coding visual images of objects in the inferotemporal cortex of the macaque monkey" , Journal of Neurophysiology , 66, pp.170-189 , 1991.
- [24] Goodale, M.A., and Milner, A.D.: "Separate visual pathways for perception and action" , Trends in Neuroscience, 15, pp.20-25, 1992.
- [25] 日本視覚学会: "視覚情報処理ハンドブック", 朝倉書店, 2000.

- [26] W.Becker and A.F.Fuchs:“Further properties of the human saccadic system;Eye movements and correction saccades with and without fixation point”,*Vision Res.*,9,pp.1247-1258,1969.
- [27] Gustavo Deco,Bernd Schümann,“A hierarchical neural system with attentional top-down enhancement of the spatial resolution for object recognition”,*Vision Res.*,40,pp.2845-2859,2000.
- [28] 横澤一彦:“多解像度モデルによる視覚的注意と視覚探索の分析”,*認知科学*,1 卷 2 号,pp.64-82,1994.
- [29] Brabo,M.J.&Nakayama,k.:”The Role of Attention in Differnt Visual-Search Tasks” , *Perception & Psychophysics*,51(5),pp.465-472,1992.
- [30] K.Fukushima,S.Miyake:“Neocognitron : A Hierarchical Neural Network Capable of Visual Pattern Recognition”,*Neural Networks*, 1(2), pp.119-130,1988.
- [31] 和家伸明, 福島邦彦:“ネオコグニトロンの新しい学習法”, *電子情報学会論文誌*,J75-D-II,11,pp.1892-1899,1992.
- [32] 宮野靖弘, 佐藤俊治, 阿曾弘具, 三宅章吾:“ネオコグニロン型神経回路モデルの開発支援を目的としたライブラリの設計と実装”, *信学技報*,NC99-135,2000.

## 研究業績

下村正夫, 佐藤俊治, 三宅章吾, 阿曾弘具,  
“仮定と検証に基づく位置・変形に頑健なパターン認識モデル,”  
電子情報通信学会 2000 年 情報・システムソサエティ大会,  
2000 年 10 月.

下村正夫, 佐藤俊治, 三宅章吾, 阿曾弘具,  
“仮定と検証に基づくパターン抽出モデル,”  
電子情報通信学会技術研究報告,  
2000 年 3 月.

## 謝辞

本研究を進めるにあたり、研究の機会を与えてくださり、折りに触れて問題点を指摘して頂くなど全般的な御指導を賜りました東北大学大学院工学研究科 電気・通信工学専攻 阿曾弘具教授に心より感謝致します。

本論文をまとめるにあたり貴重な御意見を頂いた東北大学大学院工学研究科 星宮望教授、東北大学電気通信研究所 矢野雅文教授に深く感謝致します。

本研究全般に渡り貴重な専門的御意見・御指導を賜りました東北大学大学院工学研究科 応用物理学科 三宅章吾助教授に心より感謝致します。

東北大学大学院工学研究科 電気・通信工学専攻 大町真一郎助教授には、貴重な時間をいただいて研究方針などについて議論していただいたことを深く感謝いたします。

常日頃から親身になって貴重な御意見、御助言を戴き、この研究分野に携わるきっかけを与えていただいた日本学術振興会特別研究員 佐藤俊治博士に心より感謝致します。

日々の研究におきまして多くの助言をいただき、素晴らしい研究環境を整えていただきました東北大学情報処理計算機センター 後藤英昭助手、宇都宮大学工学部 森大毅助手、東北大学大学院工学研究科 菅谷至寛助手、同 須田貴夫助手に心より感謝いたします。

東北大学電気通信研究所 伊藤真博士、同 石田文彦氏には数多くの御意見、御討論をいただきましたことを深く感謝いたします。

宮野靖弘氏(現在、シャープ株式会社勤務)には研究を進める上で重要なネオコグニトロンクラスライブラリを提供していただきましたことを心より感謝いたします。

また、高田直幸氏(現在、セコム株式会社勤務)、加藤毅氏(現在、東北大学大学院工学研究科博士後期課程在学中)、岩村雅一氏(同)、須藤貴志氏(現在、東北大学大学院工学研究科博士前期課程在学中)、村田正徳氏(現在、総合警備保障株式会社勤務) 女川俊一氏(現在、東北大学大学院工学研究科博士前期課程在学中)、門谷信愛希氏(同)、森島公一氏(同)、加藤学氏(現在、東北大学工学部通信工学科在学中)、佐々木裕児氏(同)、佐藤昌宏氏(同)をはじめとする東北大学大学院工学研究科 電気・通信工学専攻 阿曾研究室の皆様には日頃から研究に限らず多岐に渡ってお世話になりましたことを心より感謝致します。

最後に、長きにわたり温かなご支援を続けていただいた両親、家族、友人に感謝致します。