The University of Osaka
Institutional Knowledge Archive

Title	視覚システムのモデルを用いたパターン認識の研究
Author(s)	庄野, 逸
Citation	大阪大学, 1999, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3155652
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

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

The University of Osaka

視覚システムのモデルを用いた
パターン認識の研究

大阪大学 大学院基礎工学研究科
システム人間系
庄野 逸

論文要旨

生物の視覚システムは非常に優れたパターン認識装置と捉えることが可能である. 近年の計算機科学の発達によりパターン認識という研究分野は大きな進歩を遂げたが, それにも関わらず計算機によってもたらされるパターン認識装置の能力は未だ生物のそれに及ばない. 本論文は, 生物の視覚システムを原理とした工学的なパターン認識装置を構築することを目的としている. 生物の視覚システムを工学的な応用に用いるためには, まず脳の視覚システムを理解しなければならない.

脳における神経系は自己組織化能力を持ち, 学習と呼ばれる環境適応能力によって, そのふるまいを変化させる. 脳の視覚系における学習では, 特に第 1 次視覚領野 (V1 野) の細胞が良く調べられており, 局所的なエッジや直線に特異的に反応する細胞が存在していることが知られている. さらに, これらの細胞を特異性別に分類した分布 (マップ) も調べられており, 最近では運動方向の特異性を示すマップも発見されている. この様なマップの形成は視覚刺激に依存して構築されると考えられているが, この仮説を実証づけるために “競合学習” と呼ばれる学習モデルを用いて細胞の運動方向選択性についてのマップを形成し, これが実際の脳で観察されるマップと非常に良く似た特性を持つことを示した.

V1 野以降の形状認識の情報処理の経路は腹側経路 (V1 野 → V2 野 → V4 野 → IT 野) という経路によって階層的に情報が処理されていくものと考えられている. V1 野においては単純型細胞と複雑型細胞といった 2 種類の細胞が存在し, 方位選択性を持った細胞が存在することが知られている. またサルの IT 野においてはヒトの顔や, 複雑な 3 次元物体を特定の角度からみた画像に対して強い特異性を持つことが明らかにされている. 最近の生理学的データによると IT 野の細胞は, 3 次元物体としての変形 (回転, 拡大・縮小, 平行移動) を, ある程度許容することが定量的なデータとして得られている. 本論文では, V1 野で発見されている単純型細胞および複雑型細胞の関係と, 工学的に提案された神経回路モデルの一種であるネオコグニトロンの基本構造とが一致していることを示し, 皮質の一様性を仮定すると IT 野までの経路の構造がネオコグニトロン状の神経回路で説明できることを示した. その上で, IT 野の細胞にみられるような画像特異性と変形許容能力とが, 3 次元物体の画像を学習させたネオコグニトロンの出力層のユ

ニットでみられる出力特性と非常に類似していることを示すことによって, IT 野のモデルとしての妥当性を示した.

このネオコグニトロン型の神経回路が本当に視覚システムのモデルとして妥当であるならばパターン認識装置としても有用であるはずである. 本論文ではネオコグニトロンの実用性を示すために, 不特定多数によって書かれた大規模な文字画像データベースを学習データに用いてネオコグニトロンを構築した. その結果, ネオコグニトロンの特徴抽出細胞のもつ閾値を適正に設定することによって高い汎化能力を持つパターン認識装置を構築することに成功した.

さらに発展的なパターン認識を行うためにネオコグニトロンにフィードバック経路を導入したモデル (選択的注意機構のモデル) を用いて "英字筆記体文字列" の認識実験を行った. 脳の視覚領野においてもフィードバック型の回路が解剖学的に確認されており, ヒトの "注意" に関連する機構であると言われている. 従来のネオコグニトロンをフィードフォワードの系とすると, これはパターンからシンボルへの変換を行っていることに相当する. フィードバックの系は, シンボルからパターンを生成することが可能なので, システムが現在どのパターンを認識 (シンボル化) しているのかが明確にわかるようになる. これによって複合パターン中の特定パターンのみを切り出すことが出来るようになる. 英字筆記体文字列を個々の文字成分に分離しながら認識していくことはフィードフォワード系のみのシステムでは非常に困難である. しかし選択的注意機構の応用したシステムを用いることによって文字単位で逐次的に認識・切り出しを行うことが原理的に可能であることを示した. このような特定のパターンのみを画像中から切り出すことは, 自然画像を認識させる機械を構築するためには必要な技術であると考えられる. 本論文では, その一手法を提案し, 選択的注意機構のモデルが原理的に優れていることを示した.

目次

1 序論 **1**

1.1 本研究の目的 1

1.2 生理学的にみた視覚システム 5

1.2.1 信号処理の経路 6

1.2.2 V1 野 6

1.2.3 V2 野, V4 野, IT 野 9

1.3 神経細胞の学習, 自己組織化 10

1.3.1 シナプス結合の強化仮説 11

1.3.2 初期視覚領と自己組織化モデル 12

1.4 ニューラルネットによる視覚システムの研究の背景 13

1.4.1 パーセプトロン 14

1.4.2 コグニトロン 16

1.4.3 ネオコグニトロン 17

1.4.4 選択的注意機構のモデル 19

1.5 本論文の構成 20

2 1 次視覚野のディレクションマップの形成 **25**

2.1 研究背景 25

2.2 定式化 29

2.2.1 信号空間 29

2.2.2 射影学習 33

2.3 シミュレーション結果 36
2.4 検討 41

3 視覚経路とネオコグニトロン 43
3.1 研究背景 43
3.2 ネオコグニトロン 48
3.3 Logothetis らの実験 52
3.4 従来の IT 野のモデルとネオコグニトロン 55
3.5 ネオコグニトロンを用いた認識実験 58
3.5.1 回転, 拡大縮小, 位置変化に対する U_{C5} 細胞の反応 60
3.5.2 受容野の結合空間分布の推定 63
3.6 検討 67

4 ネオコグニトロンによる文字パターン認識 69
4.1 研究背景 69
4.2 ネオコグニトロン 71
4.2.1 S 細胞の反応 73
4.2.2 C 細胞層の反応 73
4.2.3 低位段の特徴抽出機構 74
4.3 S 細胞の結合の強化 75
4.4 閾値の設定 78
4.4.1 S 細胞の数学的表現 78
4.4.2 S 細胞の閾値と選択性の関係 81
4.4.3 抑制性結合の強化方式の変更点 83
4.4.4 低い閾値に設定されたネットワークの場合 83
4.4.5 高い閾値に設定されたネットワーク場合 85
4.4.6 閾値の設定方法のまとめ 88
4.4.7 閾値と生成される S 細胞面の個数 89
4.4.8 構造上の制限から来る閾値の上限 91

4.5 認識層について 94
4.6 認識結果の比較 95
4.6.1 文字パターンと学習方法 95
4.6.2 ネオコグニトロンのみによる認識率 99
4.6.3 カテゴリ分類回路 99
4.6.4 カテゴリ分類に LVQ2 を用いた場合 101
4.6.5 分類に誤り訂正学習を用いた場合 102
4.6.6 コスト関数を減少させる方法を用いた場合 104
4.7 検討 105

5 選択的注意機構のモデル化とその応用 107

5.1 研究背景 107
5.2 回路構成の概要 109
5.3 求心性経路の細胞 113
5.4 遠心性経路における細胞 116
5.5 折れ点処理回路 117
5.5.1 折れ点検出回路 117
5.5.2 折れ点復元回路 120
5.5.3 折れ点処理機構を導入する利点 122
5.6 利得の制御と閾値の制御 124
5.6.1 遠心性経路による利得の制御 125
5.6.2 サーチコントローラ による利得制御 126
5.6.3 細胞の閾値制御 128
5.7 計算機シミュレーション 129
5.7.1 回路構成と入力パターン 129
5.7.2 回路の学習 131
5.7.3 認識結果 131
5.8 検討 135

6 むすび **137**
6.1 本論文のまとめ 137
6.2 今後の課題 139

第 1 章

序論

1.1 本研究の目的

多くの生物は視覚という優れたパターン認識装置を持っている．このシステムを解明し工学的な応用に用いるという発想は, ごく自然なものである．しかしながら, その仕組みについては詳しくは分かっていないというのが現状である．ヒトやサルなどの哺乳類の視覚システムは大脳皮質上ではかなりの面積を占めるシステムである．この視覚システムは脳研究のなかでも最もよく研究されている領域であるが, 網膜 → 初期視覚領 → 高次視覚領という解剖学的に示唆された経路での, 信号の伝達のメカニズムや情報表現などが分かってきたのは, ここ数年のことである．電気生理学や培養学の発達により, 脳を構成する一つ一つの素子である神経細胞の特性はかなり分かってきているが, 脳においてはこれらが複雑に結合しあいネットワークを形成することによって, 様々な機能を実現している．このネットワークとしての働きを理解しなければ, 脳を理解したことにはならない．無論, 構成する神経細胞一つ一つの特性を理解することも重要な仕事ではあるが, 本論文ではネットワークとしての機能に焦点を絞って議論を行っていく．

本論文では, 最終的には視覚システムの理解と工学的応用に目標をおいている．物をみて識別するという行為は人間には何の苦もなくできる作業であるが, 工学的な見地に立てば, 画像という情報源から, 自分を取り巻く外界をきちんと識別できる機械というのは未だに実現されてはいない．これは, いかに脳の情報処理能力が優れているかということを示していることに他ならない．

機械による画像理解の研究は計算機の発達とともに行われ, コンピュータービジョンもしくはロボットビジョンと呼ばれる研究分野を築いた. コンピュータービジョンにおける手法としては, 人間が考え出した何らかの特徴量を抽出して行うというものが主だったものである [39]. これらの手法は特定の画像については有効に働くものもあるが, 手法がアドホック的なものなので, より一般の画像を扱うにはうまく働かないことが多い. また, 手法がアドホックなものから派生しているので, その後の研究の発展の見通しが悪いという事実も否めない.

本論文でのアプローチの仕方はこれとは異なり, まず脳に学ぶところからはじめる. すなわち, 既に脳というハードウェアが実現している部分をモデル化してシステムを構築するという手法をとるのである. 無論, 脳に未知の領域が多いということは先にのべた通りなので, 未知の部分は研究者の予想という形で補わなければならない. しかし, その予想によって生じる工学的なモデルの示す性質は, 実際の脳がその予想を反映する性質を持つかどうかということで予想の是非を問うことができる. ひとたび神経回路のモデルが出来上がってしまえば, そのモデルの構造やメカニズムは計算機上でシミュレーション可能である. これは, ただ単に脳との関連性を検証できるというだけでなく, 神経系の長所を取り入れた新しい情報処理システムの設計原理の開発にもつながるはずである. すなわち, 神経回路モデルの研究を行うことで生理学や心理学の成果を工学的な応用に結び付けることが出来るようになるのである.

ただしモデルを組み立てる際に注意しなければならないのは対象となるシステムの構成するユニットの性質を注意深く調べモデル化していかないといけないという点である. これはいたずらにモデルを複雑にするという意味ではなく本質となる部分をなるべくシンプルな形で構成した方が良いという意味である. 一般にシステムに新たなパラメータを導入した場合, 表現できる自由度は上がる. しかし, パラメータの個数が増えるとシステムに適正な動作をさせるようなパラメータを決定するために必要なデータが指数的に増えていくと言うことが “次元の呪い” として知られている [4]. 関数近似の例でも知られているようにパラメータによる自由度が上がっても有限個のデータから元の関数を推定する場合には近似精度が上がるわけではないのである. また工学的な応用を考えた場合, 決定すべきパラメータの数が少ない方が有効であるのは明らかである. したがって “何が本質であるか” を考えるかが非常に重要な問題となる.

本論文では, ヒトやサルなどの視覚システムの生理学的なデータに対する考察をもとにニューラルネットワークモデルの構築を行い, 視覚システムを工学的に再現した柔軟性を持ったパターン認識システムを構築することが目的である. この目的を達成するためには次の三つ条件が必要であると考えられる. それは

- 構築したシステムが, 実際の生理学的なデータをきちんと説明できるかどうかを検証する条件 (モデルの妥当性)
- モデルを計算機上に実装し, シミュレーションによって原理的に問題がないかどうかを確認する条件 (モデルの可能性)
- 構築したシステムがアプリケーションとしてどのくらいの能力を持つか検証する条件 (モデルの有用性)

というものである.

第 1 の条件をクリアするためには, システムを構成するユニットとシステム全体の処理の流れ, という二つの面を考えなければならない.

問題となる脳は神経細胞というユニットから構成されている. すなわち, 脳の一部分である視覚システムという神経回路モデルを構築するためには, まず神経細胞のモデルを構築しなければならない. 神経細胞の最大の特性は学習とよばれる環境に対する適応によって, そのふるまいを変化させることである. 神経細胞の学習についての詳しい研究の背景については, 本章の 1.3 節で述べているが, 現在は神経細胞の活動度によって変化するという節が有力である. 本論文では, まず構築する神経細胞のモデルで用いる学習則である “競合学習仮説” についての検証を行う. これは初期視覚野の特徴抽出細胞と, 競合学習で構築した細胞群との特性を比較することで行う.

各ユニットの持つ性質を満足させた後は, システムの全体構成を考えなければならない. そこで, 一つ一つの神経細胞の性質から離れて今度は脳の視覚システムに目を向けてみる. 脳の視覚系のシステムは解剖学的な性質により様々な領野に分類される. 各領野の持つ性質の詳細や研究背景については, 1.2 節に記述しているが, 大まかな流れとして, 物体の形状認識をつかさどるシステムは V1 野 → V2 野 → V4 野 → IT 野といった階層構造をなすシステムによって形成されている. 各領野では, 抽出していると思われ

る画像の特徴も異なっている. 以上の議論から脳の視覚系のネットワークモデルが満たすべき性質としては,

- 構築されるシステムのパラメータは学習によって決定される.
- ネットワークとして階層性を持つ.
- 各階層の構成要素と生理学的な対応がとれるようにする.

という点があげられる. これは, 生理学的にも確かめられた視点にたってモデルを構築することを意味している. 以上のドグマにしたがってシステムを構築すると Fukushima の提案したモデル “ネオコグニトロン” と非常に良く似た構造を持つことが分かる [23]. ネオコグニトロンの詳細については本文中で述べる. また提案されるに至った工学的なニューラルネットワーク研究の背景については 1.4 章で触れる. 本論文では, 第 2 章で学習アルゴリズムである “競合学習” の妥当性を検討する. 第 3 章では V1 野の細胞間の結合構造とネオコグニトロンの基本構造とが類似している点とサルの IT 野の細胞が示す性質と, ネオコグニトロンの出力ユニットが示す持つ性質の類似性をもって, ネオコグニトロンの IT 野のモデルとしての妥当性を示す.

第 2 の条件をクリアするためには, モデルを計算機上に実装しトイプロブレムのような簡単で少数の例題を用いて, 問題点となることを解決しておく必要がある. なぜなら, このレベルでの問題解決を行わずに, 第 3 の有用性を示すレベルに進んでしまうと, システムに内在する問題点が, 解決のためにモデルの変更を行うべきなのか, パラメータ等の調整で回避できる問題なのかをはっきりさせることが出来ないからである. ネオコグニトロンに関しては, 過去の研究によって様々な問題点が提起され, 解決されてきているので, このレベルに関しては本論文では言及しない [28], [32] [31] . しかし, ネオコグニトロンをさらに発展させた “選択的注意機構のモデル” では, まだこの条件での検証が必要なので, 第 5 章でこのレベルでの検証を行う. 従来のコンピュータービジョンでは用いられてきた手法では, 画像の正規化がなされているという条件が暗黙の内に加えられているものが多かった. 自然画像のような複合パターンにおいては, 認識対象が画像中のどこにあるのかを決定するのは困難な問題である. 認識対象のみを入力画像から切り出す問題は, コンピュータービジョンにおいてはセグメンテーション問題と呼ばれ,

この問題も完全には解決されていない．正規化が必要な手法で, このような複合パターンを処理するには, セグメンテーション問題を解決した上で, 何らかの前処理をしてやらなければならない．ヒトは, 認識対象に “注意” 向けることによって, このような複合パターンのセグメンテーション問題を解決し, 認識能力を向上させていると考えられている．そこで, 第 5 章ではネオコグニトロンにフィードバック型の回路を組み合わせた選択的注意機構のモデルを導入し, このシステムの可能性を探求する．選択的注意機構のモデルは, 複合パターン中の特定パターンに対して注意を向けることが出来るので [24], 英字筆記体文字列のような複合パターンに対しても有効であるかどうかを検討する．

第 3 のレベルは, 第 2 のレベルの延長上にあり, 構築したシステムがどの程度の有用性を持つかを示すというものである．文字パターン認識という観点から言えば, 不特定多数に書いてもらった実画像パターンをシステムに識別させ能力を測定することが, これに当たる．たとえ生理学的な整合性が合致したとしても, それが視覚システムの優れた認識能力を持ち得なければ, 工学的にはあまり意味を持たないのである．本論文ではコンピュータービジョンなどでは工学的な観点から構築されたニューラルネットワークモデルやヒューリスティックな方法を用いて, パターン認識実験が行われている．最近では正規化されたパターンに対しては高い認識能力を持つことが知られている [55] [105]．もともとネオコグニトロンは Fukushima らがパターン認識機械として提案したものだが, 現在までの間に実画像データを用いて, その能力を評価している研究は存在しない．そこで第 4 章で, ネオコグニトロンが工学的なアプリケーションとして役立つかどうかを調べるために, 文字画像データベースを用いてパターンの認識実験を行い有用性を示す．

以下の 1.2 節から 1.4 節までは本研究において記述されたモデルを理解するために必要な研究の背景についての説明を行う．

1.2 生理学的にみた視覚システム

本研究において記述されたモデルを理解するためには, 少なくとも脳の生理学的な性質を知らなければならない．この節ではモデル研究のバックグラウンドとなる生理学的, 解剖学的な性質について, その研究の歴史について必要最低限度に述べる．

大脳は左右の大脳半球に別れており大脳半球を切断してみると表層部に厚さ 2.5mm

程度の灰白質がみられる．この部分を大脳皮質と呼び, 皮質内には多数の神経細胞が密集して神経回路を構成している．皮質内の神経細胞の総数は約 140 億個と言われている．皮質の内側は白質と呼ばれ皮質間の信号を伝達する神経細胞の軸索があつまっている．

大脳皮質を解剖学的に分類する試みは数多くなされてきたが, その一つに Broadmann の脳地図がある [13]. 図 1.1 は Broadman が分類した脳地図で, 52 の領野に分類され各領野には番号がつけられている．大脳皮質の各部分はそれぞれ異なった情報を処理するような機能を分担すると考えられている．

1.2.1 信号処理の経路

ヒトやサルの視覚システムは左右大脳半球の後頭部から始まる．眼球の網膜で捕らえれた光学情報は外側膝状体 (LGN) を介して後頭部の 17 野 (サルなどで言うところの V1 野) に伝達される．軸索の投射の解剖学的な研究から V1 野からの信号は 図 1.2 のような領野に伝達される．視覚情報は大別して二つの経路に分別される．一つは物体や図形の “形” の認識に関わる腹側経路と呼ばれる経路で, V1 野 → V2 野 → V4 野 → IT 野 (PIT, AIT) という経路である．この経路は後頭葉から下部側頭葉 (inferior temporal Cortex) の領野で構成される．もう一方の経路は背側経路と呼ばれる経路で, これは物体や図形の “動き” の認識に関わる経路と言われている．この経路は V2 野から MT 野を経て MST 野へと至る経路である．

1.2.2 V1 野

Hubel と Wiesel はネコの 18 野やサルの V1 野の細胞に微少電極をさして細胞内電位を図ることに成功している [40] [44]. ある一つの神経細胞の反応に影響を与える網膜上の領域 (それに対応する視野上の領域) を神経細胞の受容野と呼ぶが, V1 野の細胞の受容野の大きさは約 1.3 ∼ 5 度と非常に小さいことが知られている．Hubel と Wiesel は V1 野の細胞の受容野に様々な刺激を与え, 特定の方向のエッジや線分に反応する細胞を発見している．細胞が最も強く反応する直線やエッジの方位を, その細胞の最適方位とよぶ. 通常の環境で育ったサルやネコの V1 野の細胞の最適方位は特定の方位に偏ることなく, あらゆる方位のものが一様に分布している．また V1 野の細胞の皮質上で

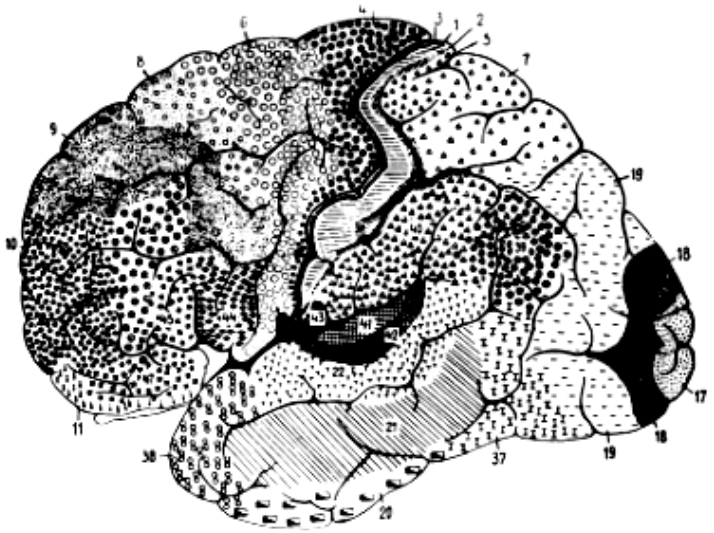

図 1.1: Broadmann の脳地図. ヒトの大脳の左半球を外側からみた図 ([13] より)

の位置と, その細胞の網膜上の受容野の位置との間にはトポロジカルな意味での連続的な対応関係が存在する. すなわち, V1 野皮質上での距離が近い細胞どうしの受容野の中心は網膜上でも近くにあるのである. この幾何学的な対応関係をレチノトピーと呼ぶ. Schwartz は順行性色素を網膜上で格子状に投与し V1 野でどのように投射されるかを調べており, V1 野でのレチノトピーは網膜の歪んだコピーのようになることを示している [89].

V1 野で観察される細胞

ここで, V1 野の細胞についてさらに詳しく述べる. Hubel と Wiesel は, V1 野における細胞を受容野の性質に基づいて単純型細胞と複雑型細胞などに分類した. 彼らは, これらの細胞相互間には 網膜 → 外側漆状体 → 単純型細胞 → 複雑型細胞 という, 階層構造が存在するという階層仮説を提唱している [40] [41] [43] [44].

Hubel と Wiesel によれば単純型細胞は丁度, 直線やエッジに反応するような受容野を持つ. 図 1.3 は, 彼らの実験によって得られた単純型細胞の受容野の例を示したものである. '+' に沿った部分に光刺激が入力されると強い反応を示し, '-' の部分に光刺激が

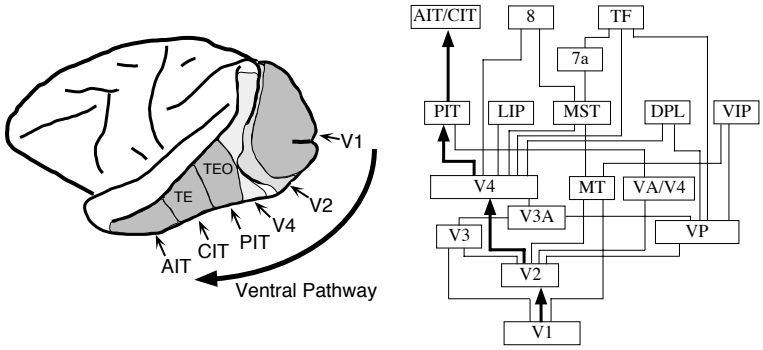

図 1.2: 後頭葉から下部側頭葉にかけての経路 (腹側経路) と, 領野間の階層的な結合関係における腹側経路 ([19] を一部変更)

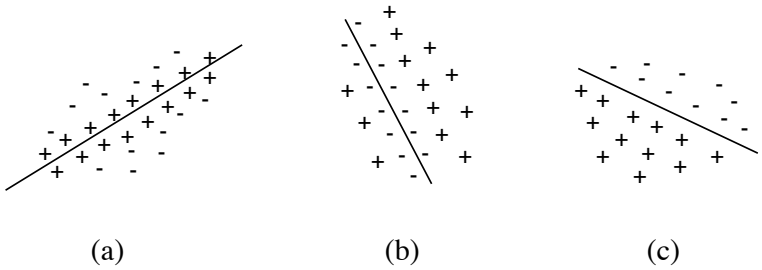

図 1.3: 単純型細胞の受容野の例. ([43] を一部変更)

入力されると細胞の反応が抑制される. これらの細胞は '+' と '-' の配置から考えて, ある特定の傾きを持った直線やエッジを検出するのに役立っていると考えられる. ある細胞が最も強く反応する直線やエッジの方位を, その細胞の最適方位と呼ぶ. 最適方位は細胞によって異なり, ネコやサルを通常の環境で育てた場合にはあらゆる方位のものが一様に存在していると言われている.

単純型細胞では, 刺激として与える直線やエッジの位置が少しでもずれると出力が抑制されてしまう. これに対し複雑型細胞では, 最適方位の直線やエッジが受容野の内部にありさえすれば, その位置に関わらず反応するという特性を持つ. つまり単純型細胞は位相に対して敏感であったのが, 複雑型細胞では受容野内では位相不変な特性を示す. Hubel と Wiesel は, 同一の最適方位を持つ複数の単純型細胞を興奮性結合によって結

び付けられているモデルを複雑型細胞のモデルとして提案している. このモデルにおいて, 複雑型細胞は結合している単純型細胞のいずれか一つでも反応すれば複雑型細胞も発火するものと考えた.

Hubel と Wiesel の階層仮説には様々な反論があるが, このような階層構造をもつ神経回路を構成することによって V1 野の細胞の性質を再現できるのは非常に興味深い考察である.

1.2.3 V2 野, V4 野, IT 野

V1 野以降の腹側経路の情報処理は V2 → V4 野 → IT 野と階層的に続くと考えられている. これらの領野ついては V1 野ほど詳しくは分かっていない. V4 野以降の PIT 野や AIT 野と呼ばれる領域は明確に境界が定まっているわけではない. これらの領野は細胞構築学的に TE 野や TEO 野と呼ばることもある. レチノトピーに関しては, V4 野までは存在することが確認されているがこれ以降の領野に関しては存在しないといわれている. 各領野の細胞の持つ受容野の大きさは V2 野が V1 野の 2 倍強の大きさの 3.2 度, V4 野は約 8.0 度, IT 野ではこれが 20〜50 度であるという報告や [52] [83], IT 野の後部 (PIT 野, もしくは TEO 野と呼ばれる) では 1.5 度から 2.5 度で, IT 野の前部 (AIT 野) に至って 30 〜 50 度の受容野をもつようになるという報告もなされており [10] [107], さまざまな大きさの受容野が存在するとも考えられる. ただ, いずれの報告も処理の階層が上がるにつれ, 受容野が大きくなり, 特異的に反応する特徴も複雑なものになると言われている [36] [52] . Desimone らは, AIT 野の受容野特性について調べており, 線分の長さ, 入り組んだ輪郭, 色の組合わせ, 特定のテクスチャ, などに特異的に反応する細胞を発見している [16]. IT 野の細胞は視覚パターンの形状認識に特に重要と領野と考えられている. サルの AIT 野の破壊実験を行った場合, 特に図形の弁別問題に重度の障害が現れるが, 視覚以外の聴覚, 触覚, 嗅覚による弁別学習には何の異常も見られない [48]. また, IT 野や, その近傍の STP 野ではヒトの顔やサルの顔に特異的に反応する細胞が発見されている [14]. このような細胞は顔から目を取り去ったり, 顔の画像を 16 分割してランダムに並べ替えたものには反応しない. これらの細胞はパターンの大きさが変化したりしても反応に変化はないが, 顔の向きの影響は大きく, 正面からみた顔と横

からみた顔とで大きく反応が変化する [80][61].

最近では Tanaka らが, 画像から IT 野の細胞が強く反応する特徴を探しだし, その特徴をより簡単な特徴に置き換えて最適刺激特徴を探索するという手法を用いて AIT 野の細胞の受容野特性について調べている. それによると 図 1.4 のようなパターンに特異的に反応する細胞が発見されいてる [107] [108].

1.3 神経細胞の学習, 自己組織化

それでは, 受容野に現れるような神経細胞の性質はどのようにして決定されるのだろうか？本節では, 学習と自己組織化についての研究背景をのべる.

結合の大まかな構造は遺伝的な性質によって決定されていると考えられているが, ネコを縦じましか見せないような環境で育成すると V1 野の細胞は, 特に縦方向の直線に反応する細胞が多くなることが知られている. つまり後天的な環境の変化に応じて神経細胞は, その特性を変化させるのである. 成長段階に煩雑に与えられた刺激パターンの特性に対応するような神経回路が形成され, おかれた環境でめったに現れないような刺激パターンに対応するような神経回路が消滅していくということは神経回路の能力を無駄なく効率的に発揮させるのに役立っていると考えられる. カエルや魚類などの網膜から視蓋への神経細胞の結合は最初広く分布しているが, 成長段階においてシナプスが除去されていくことが知られている. この神経細胞の電位依存性チャネルを阻害するとシナプスの除去が行われないことから神経細胞の活動度と密接な関係があると考えられる [59] [90]. この現象は活動依存性シナプス除去と呼ばれ, 発生途上の脊椎動物神経系のほぼ全般にわたってみられる.

このような外部から与えられる刺激に応じて神経回路が変化し自分のおかれた環境に適応するようになっていく過程を, “神経回路の自己組織化 (self-organization)” とよぶ. 自己組織化は, 脳の視覚システムだけでなく他の領野においても生じていると考えられる.

神経細胞は, その入出力に応じて発火特性を変化させることが知られている. この特性変化の原因はおもに神経細胞どうしの結合部であるシナプスで起こると言われており, 神経回路の自己組織化もこのシナプスの可塑性によって引き起こされるものと考えられ

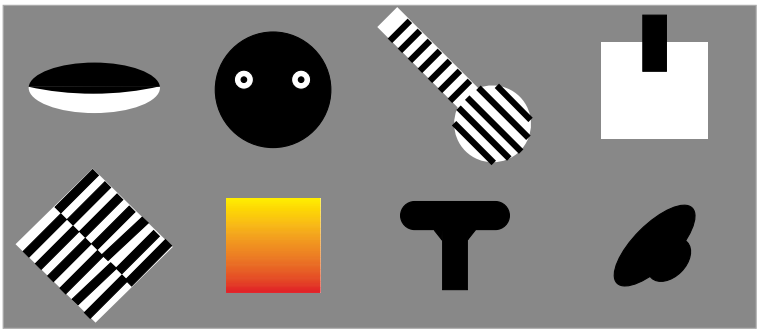

図 1.4: Tanaka らの発見した AIT 野の細胞が特異的に反応する図形パターンの例 ([108] を一部変更)

る. 特にシナプスを介して高頻度の刺激を行うことによってシナプス伝達の効率が長期的に増強する現象 (long term potentiation: LTP) や 長期的に減弱する現象 (long term depression: LTD) が観察されている. このようなシナプスの可塑的な変化の過程を観測する努力は古くから生理学者によって精力的に続けられており, 最近の研究では LTP が実際の反射行動行動を制御している報告がなされている [76].

1.3.1 シナプス結合の強化仮説

シナプス強化による自己組織化の理論的研究は神経回路の研究と生理学的なデータの集積ともに発展してきた. Hebb は "結合している細胞が同時発火したときに結合が強められる" という Hebb 則を提案した [37]. この仮説は Pavlov の古典的条件反射を説明するための仮説として提唱された [78]. この仮説は, "連合作用仮説" と呼ばれる仮説の一つである. 連合作用仮説とはシナプス前細胞 X とシナプス後細胞 Y の状態が適当な組合わせになったときに限って X から Y へのシナプス結合に変化が生じるという仮説である. 図 1.5 (a) のように細胞 Y_1 にシナプスを形成しているのは, 細胞 X_1 だけでなく X_0 や X_2 などの他の細胞もシナプスを形成している場合を考える. この場合に Y_1 と X_1 とが同時発火したときに, この二つの細胞を結び付けるシナプスの結合強度が強められるというのが Hebb の仮説である. 前述の活動依存性シナプス除去も, この連合作用仮説の一種であると考えることができる. なぜならシナプス前細胞 X が発火しても

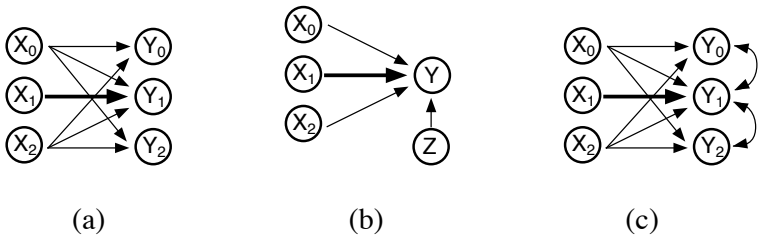

図 1.5: シナプス強化を行うモデル. ([26] より)

シナプス後細胞 Y が発火しなければシナプスの結合が弱まっていくという解釈が可能だからである.

Marr は, 小脳の神経回路モデルとして “強化指令仮説” を提唱した [62]. これは 図 1.5 (b) に示すように, 細胞 Z という, 細胞が存在して, この細胞が丁度教師信号の役割を果たすという仮説である. 細胞 Z が発火すると細胞 Y が必ず発火するような状況を考える. 細胞 Y と 細胞群 X_0, X_1, X_2 とのシナプスが, Hebb 則と同様に, 同時発火によって強化されると考えた場合, 細胞 Z が Y の発火を制御しているので, 初期状態においてシナプス結合が弱くとも自己組織化は進行していく.

“競合学習” もまた連合作用仮説や, 強化指令仮説と同様に理論的考察から導きだされたものである. 競合学習はシナプス前競合とシナプス後競合の二つに分類される. シナプス前競合は, ある細胞に結合できるシナプスの数は限られていて, 結合強度が弱いものは消滅していくというものである. シナプス後競合は 図 1.5 (c) に示すように, 学習すべき細胞群 Y_0, Y_1, Y_2 の間で競合過程が存在して, 競合に勝った細胞のみが, シナプス結合を変化させられる (winner-take-all) という学習規則である [87]. Fukushima が提案している “最大値検出仮説” や Kohonen らが提案している学習規則も競合学習の一種である [22] [53].

1.3.2 初期視覚領と自己組織化モデル

それでは, 上述したような学習モデルによって脳の細胞の性質やふるまいが説明できるのであろうか? この疑問に答えるには, 計算機科学の発達が必要不可欠であった. 1980 年代における計算機科学の著しい発達により, モデルを計算機上で再現することが可能

になり, このような学習モデルを用いた研究が現在でも盛んに行なわれている.

サルの V1 野やネコの 18 野においては, 局所的な線分やエッジに反応する細胞が多数存在する. von der Malsburg は, この生理学的観察結果を Hebb 則を導入した自己組織化モデルによって説明した [114]. von der Malsburg は計算機シミュレーションを行いエッジ特異性を持つ細胞が発現する様子を示している. また, このような方位特異性を持つ細胞が, 目の開いてない時期に既に形成されるという報告に基づき [42] [8], Linsker は Hebb 則を拡張したモデルを用いて多層神経回路モデル上で自己組織化を行っている [58] [57] [56]. Linsker モデルは入力に白色雑音を用いた時に, 第 1 層には LGN の細胞が持つようなメキシカンハット状の受容野, 第 2 層には特定の傾きを持った エッジに対応するような受容野が形成されることを示している. V1 野の細胞の振舞いは結合が遺伝的に決定されていることに起因するという仮説も存在するが, 網膜上の細胞の自発発火を薬品で抑制すると結合強化がみられないという事実から, 上述の学習によるシナプス強化が神経細胞で行われている可能性は高い.

最近では サルの V1 野やネコの 18 野の細胞の発火分布を調べるために, 特殊な色素を大脳皮質表面に投与し視覚刺激を与えたときの皮質表面を CCD カメラで取り込み, 画像処理を行うオプチカルイメージングという解析方法が盛んに行われている. Obermayer らは Kohonen の競合学習アルゴリズム (SOM) を用いて皮質の発火を再現するようなモデルを構築し, V1 野の細胞のオプチカルイメージング像との比較検討を行っている [72]. Obermayer らのモデルは Linsker のモデルと異なり, 入力には特定の方位を持つようなエッジや線分を入力する形になっており, 実際に眼が開いた後の学習について考えたモデルになっている.

現在においては, このようにモデルによる研究と生理学的なデータとを比較し検証していく状態である.

1.4 ニューラルネットによる視覚システムの研究の背景

前節までは実際のヒトやサルの視覚システムと, 神経回路網を構成する結合の学習仮説についてみてきた. この節では, ニューラルネットワークを用いたパターン認識の工学的な応用の研究背景, および本論文で用いているネットワークの構造の概略について

述べる.

1.4.1 パーセプトロン

階層型ニューラルネットを用いてパターン認識を行うという試みは, 心理学者である Rosenblatt のパーセプトロンの研究によってはじめられた [84][85].

Rosenblatt のパーセプトロンは 図 1.6 のような 3 層のネットワークからなる. 各層は S (Sensory) 層, A (Association) 層, R (Response) 層と呼ばれる. 各層は適当な数の細胞 (ユニット) から構成され, 同じ層内の結合はなく, 層間の結合は S 層 → A 層 → R 層という一方向のフィードフォワード型の結合からなっている. このうち S 層から A 層への結合は固定結合で A 層から R 層の結合のみが "学習" により決定される. S 層のユニットは網膜の視細胞のようなものを想定しており, 外界からの刺激パターンは, この層に入力される. A 層 と R 層のユニットは前の層のユニットからの重みつき入力の総和を計算し, これに適当な出力関数をかけたものを出力する. S 層の i 番目のユニットの出力を s_i, A 層の j 番目のユニットの出力を a_j とし, S 層の i 番目のユニットから A 層の j 番目のユニットへの結合 (固定結合) を c_{ji} とすると, a_j は

$$a_j = f(\sum_i c_{ji} s_i - \theta_j^A) \tag{1.1}$$

とかける. ただし, θ_j^A は A 層の j 番目のユニットの閾値であり, f は出力関数である. Rosenblatt は f にユニット関数

$$f(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

を用いている. 同様に R 層のユニットの出力を r, A 層の j 番目のユニットから R 層のユニットへの結合を w_j とすると r は

$$r = g(\sum_j w_j a_j - \theta^R) \tag{1.2}$$

となる. θ^R は R 層のユニットの閾値であり, g は出力関数で, オリジナルのパーセプト

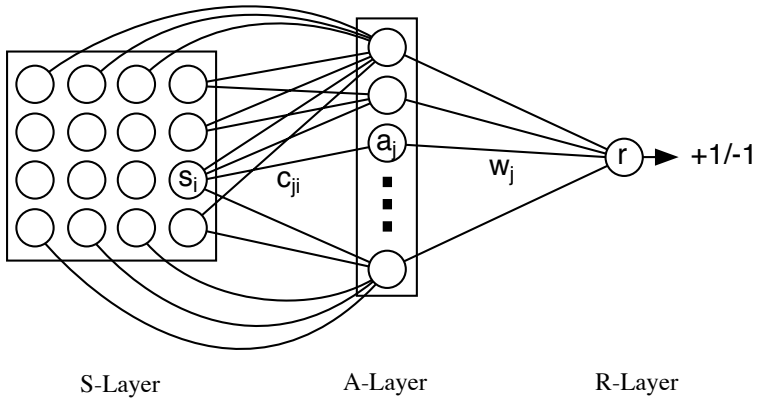

図 1.6: 典型的な 3 層パーセプトロンの例

ロンでは符合関数を用いている:

$$g(x) = \begin{cases} 1 & (x \geq 0) \\ -1 & (x < 0) \end{cases} .$$

結果 S 層に入力されたパターンは $\{-1, +1\}$ という 2 種類のシンボルに変換されることになる. Rosenblatt のパーセプトロンの学習則は, “誤り訂正学習” とよばれ Marr の強化指令仮説と非常に良く似ている [62]. この学習則は, あるパターンが入力され, なおかつ R 層の出力が望ましくない場合に, 重み w_j と閾値 θ^R に対して,

$$w_j = w_j + \eta r a_j \tag{1.3}$$

$$\theta^R = \theta^R - \eta r \tag{1.4}$$

という変更を加えるものである.

Minsky と Papert は特に学習に関連する A 層と R 層のみに特化した “単純パーセプトロン” モデルを用いて, 誤り訂正学習の収束定理を証明し, また, 単純パーセプトロンの能力の限界について言及している [66]. 彼らは, 単純パーセプトロンをある種の論理関数を計算する機械としてとらえ, どのような学習をしても XOR 関数を計算できないことを指摘している. これは逆に考えれば, 多層神経回路モデルの必要性を述べているといえる.

多層パーセプトロンの学習則は Amari によって言及され [1], 1986 年に Rumelhart らによって再発見された [86]. これが現在良く使われているニューラルネットワークの学習規則である一般化誤差逆伝搬アルゴリズム (error back propagation) と呼ばれているものである. これはユニットの出力関数を微分可能な関数に置き換え, ある結合係数を少しだけ変化させたときに出力に対してどれくらい影響があるかを計算する. この操作を全ての結合係数に対して行い, 望ましい結果に向かうように結合を微小変化させる手法である. Rosenblatt のパーセプトロンでは A 層の出力関数が微分不可能な関数だったために S 層から A 層の結合 c_{ji} を変化させることが不可能であるが, これを微分可能な出力関数に置き換えることによって微分のチェーンルールが適用できるようになり, 結合の微小変化に対する応答の変化が計算できるようになる. 一般化誤差逆伝搬アルゴリズムを用いたパーセプトロンは, ソナーによる音響特性の区別や, 文字パターンの認識で評価され, “人は明らかに区別することができるが, 何を手がかりにして区別しているかを明確に定義できないような問題” に対して効果的であることが示されている.

1.4.2 コグニトロン

一般にパーセプトロンのユニットは前層の全てのユニットからの重みつき入力を計算する. 3 層パーセプトロンを考えた場合, S 層が n_S 個, R 層が n_R 個あるような状況では中間層である A 層にユニットを一個追加するだけで $n_S \cdot n_R$ 個の結合が必要になる. 情報処理機械を構築するコストという観点から考えると, 結合の数は非常に重要なファクターであり何らかの制限を加えるのは妥当であると考えられる. Fukushima の提案した “コグニトロン” は多層の神経回路モデルであり, 各ユニットが持つ結合に空間的な局所性を持たせたモデルである (図 1.7) [22]. コグニトロンの各ユニットは 1 層前のユニットの小領域に存在するユニット群から重みつき入力を受け取る. この結合可能な小領域は, 前層の全体を覆うほど大きくはない, しかし層を重ねるごとに入力層 U_0 のより広い範囲からの情報を受けることになり, 最終的な層のユニットは入力層 U_0 全体からの情報を受けるようになっている (図 1.7)(a).

図 1.7 (b) は隣接する層の構造を表したものである. 図には U_{l-1} 層と U_l 層のユニット間の結合のうち, U_l 層の一個のユニット $u_l(\boldsymbol{n})$ に至る結合のみを抜き出して描いてある.

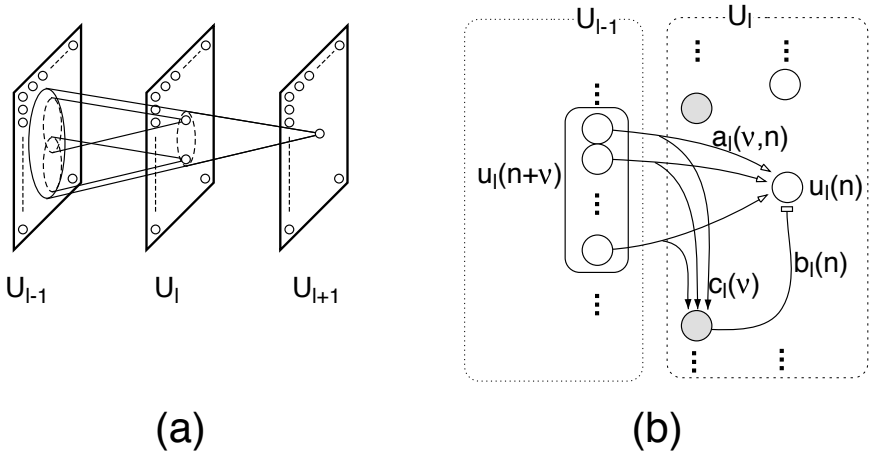

図 1.7: コグニトロンの結合の結合の広がり (a), と隣接する 2 層間の結合 (b)

ユニット $u_l(\boldsymbol{n})$ は結合している小領域 A_l に含まれるユニット群 $u_l(\boldsymbol{n}+\boldsymbol{\nu})$ から $a_l(\boldsymbol{\nu},\boldsymbol{n})$ という重みをかけた入力を受ける. (ただし $\boldsymbol{\nu} \in A_l$ とする) ユニット $v_l(\boldsymbol{n})$ は抑制性のユニットで, やはり領域 A_l から, $c_l(\boldsymbol{\nu})$ という重みをつけた入力を受ける. $v_l(\boldsymbol{n})$ の出力は, 抑制性の信号として $b_l(\boldsymbol{n})$ という重みをつけて $u_l(\boldsymbol{n})$ に伝達される. この $a_l(\boldsymbol{\nu},\boldsymbol{n})$ と $b_l(\boldsymbol{n})$ が学習によって決定される結合である. コグニトロンのユニットはパーセプトロンのユニットとは異なりシャント型の特性を持つ.

コグニトロンのもう一つの特徴は, そのユニットの学習方式であり, これには “最大値検出型仮説” に基づいた競合学習が用いられている. これは, あるユニットの出力がその近傍のユニットの出力に比べて大きいときにのみ結合の更新が起こるというものである. このような winner-take-all 方式の学習はユニットの機能の分業化が自己組織的に起こる. また, ある機能を持つユニットが故障しても, しばらく学習することによって, その機能を復元できるという利点を持つ.

1.4.3 ネオコグニトロン

前述のパーセプトロンやコグニトロンは, 学習させたパターンを別の場所に提示したり, パターンの拡大・縮小をすると全く別のパターンと判断されてしまう. Fukushima は, このような欠点をとり除いた神経回路モデルである “ネオコグニトロン” を提唱した

[23] [28]. ネオコグニトロンは S 細胞と名付けられたユニットの層 U_S と C 細胞と名付けられたユニットの層 U_C とが交互に並べたような階層型の神経回路モデルである. S 細胞は Hubel と Wiesel が発見した単純型細胞をモデル化したユニットであり C 細胞は複雑型細胞をモデル化したユニットである. ネオコグニトロンの各細胞間の結合はコグニトロンと同様に空間的な局所性を持っている (図 1.8). S 細胞は特徴抽出機能を持つユニットで, その入力結合は可塑性を持ち, コグニトロンとのユニット同様に競合学習によって自己組織化することが可能である. S 細胞は同じ特徴を抽出するユニットで "細胞面" を形成している. 一つの細胞面内のユニットは全て同一の空間分布の入力結合をもっている. 同一細胞面内の各ユニットの相違は結合する相手のユニットが面内の位置ずれに相当する距離だけ平行移動しているという点にある. C 細胞は, ある前層の S 細胞層の同一細胞面内のユニットからの結合を受け, 自分の結合している領域の S 細胞がどれか一つでも出力を出した場合に, 出力を出す. これによって C 細胞は複雑型細胞と同様の位相不変性を持つ. 結果的に C 細胞の出力は S 細胞面の発火パターンをぼかしたようなものとなる.

入力パターンの情報は S 細胞層と C 細胞層による操作を多段にわたって受ける. この間に局所的な特徴相互間の位置ずれや変形が徐々に許容されていき, 最終的には相当変形したパターンに対しても正しい認識を行うことができるようになる. 重要なのは位置ずれや変形などの影響を一回で吸収してしまうのではなく, 何段階にもわけて少しずつ吸収していくことである.

Fukushima と Wake はネオコグニトロンを手書き文字に対して適用し, その有効性を示している [32]. 第 4 章で述べるパターン認識を行うシステムのベースとなるモデルもネオコグニトロンである. 第 4 章では, ネオコグニトロンを特徴によるパターン分類機械とみなし, これに前述のパーセプトロンを組み合わせることによってパターン認識機械を構成している.

また LeCun らが提案している数字認識システムに用いられているシステムの一部もネオコグニトロンと非常に良く似たシステムとなっている [55]. LeCun らは, このネットワークを BP (error back propation) 学習を用いて学習させている.

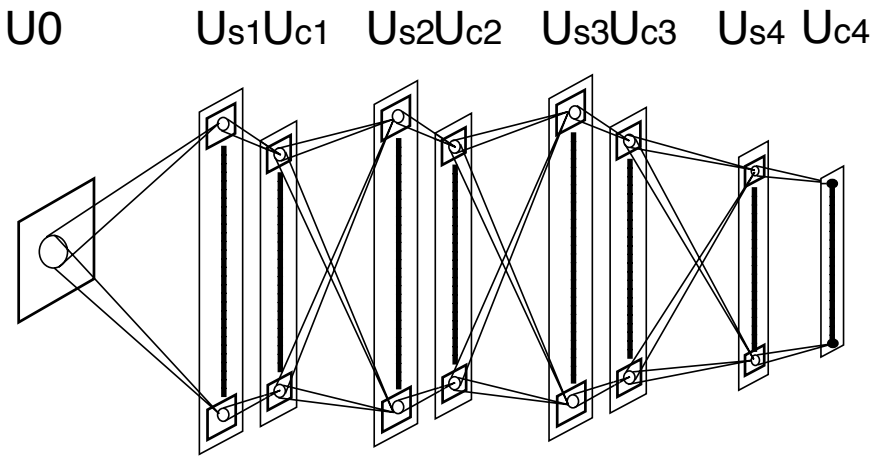

図 1.8: ネオコグニトロンの各階層の細胞間の結合をあらわす概念図 ([23])

1.4.4 選択的注意機構のモデル

階層性を持つ神経回路モデルとしてパーセプトロン, コグニトロン, ネオコグニトロンを例としてあげてきたが, これらの結合は一方向に信号を伝える “フィードフォワード” 型のネットワークであった. フィードフォワード型のネットワークはボトムアップ的なネットワークであるが, これとは逆にフィードバック結合を持つようなトップダウン的なネットワークも考えることが出来る. フィードバック型の歴史は Rosenblatt の逆結合型パーセプトロンに始まる [85]. このモデルは 図 1.6 に示すようなパーセプトロンの A-R 層間の結合がフィードフォワードの結合だけでなく, R 層から A 層に信号を伝えるようなフィードバック結合を持っているようなモデルである. このようなフィードバック結合があるとパーセプトロンはパターン中に含まれる特定の特徴に注意を向けるような性質を持つようになる. Fukushima はこの事実に基づき “選択的注意機構” のモデルを提案している [24]. 選択的注意機構のモデルは 図 1.9 のような 2 方向の信号伝達経路をもつ. フィードフォワードの信号伝達経路は, パターン分類を行う経路として考えられる. すなわちパターンから (文字) シンボルへの変換を行っているものといえる. それでは, あらたに導入されたフィードバックの経路は, どのような機能を実現しているのだろうか？この経路はフィードフォワードの系とは逆に, シンボルからパターン

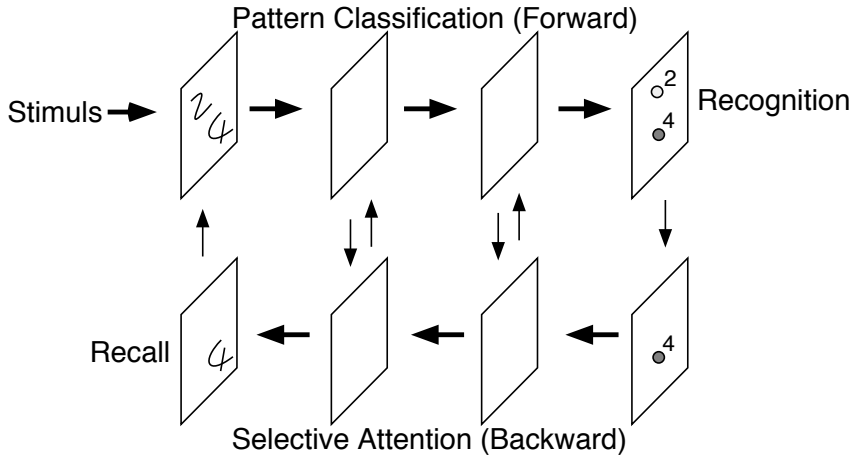

図 1.9: 選択的注意機構の概念図 ([24] を変更)

への逆変換を行ってるものと捉えることができる. すなわち認識している特定のパターンのみを抽出していることになる. 工学的な応用面で考えれば複合パターンを構成する特定パターンの切り出しが可能になることを意味する. 芦田と福島は選択的注意機構のモデルを漢字を部首ごとにわけて認識するシステムを構築している [2].

本論文の第 5 章において議論している選択的注意機構のシステムも Fukushima の選択的注意機構のモデルに基づいて構築されている.

1.5 本論文の構成

以上, 神経回路モデルの研究における生理学的な側面と工学的な側面について述べてきた. 本論文の第 2 章から 第 5 章までの構成は以下の通りである.

第 2 章では, サルやネコなどの生理学実験で良く知られている一次視覚野の学習についての議論を行う. 一次視覚野は, 視覚研究の中でも最も良く研究されている領野であり, 局所的な線分やエッジのような画像パターンに反応する細胞により構成されているといわれている. 近年では画像解析システムなどの計測機器の発達により, ある特定の傾きのエッジや線分に反応する細胞がどのように配置されているかが計測できるようになってきている. このように特定の機能をもつ細胞の配置を皮質の機能地図 (マップ) と

呼ぶが, これは与えられた環境によって様々に変化する. 例えばネコを縦じまだけしか見えないような特殊な環境で育てると縦の線分に応答する細胞が多くなる [7]. つまり, このような細胞の特性は主に学習によって決定される. 最近の研究では, 静的な線分パターンに反応する細胞のマップ (方位マップ) だけでなく, 線分がどの方向に動くかというマップ (方向マップ) も観測されている. 第 2 章では, 方向マップもまた競合学習によって構築できることを計算機実験で示す. 計算機実験で構築した方向マップと実際にサルの V1 野やネコの 18 野で観測される方向マップとは非常に多くの点で類似している.

第 3 章では, サルの視覚経路とネオコグニトロンとの類似性についての議論を行い, 実際の生理学実験によって得られるデータとの比較を行う. モデルとなる視覚経路は物体の形状の認識を行う "腹側経路 (Ventral Pathway)" と呼ばれる経路である. 腹側経路は, V1 野 → V2 野 → V4 野 → IT 野 という領野からなり階層構造を持っている. 視覚情報の信号はこれらの階層をフィードフォワード的に流れるものと仮定している (本論文では, この経路を求心性経路と呼んでいる). 視覚の信号処理は第 5 章でも述べるように, フィードバック経路 (遠心性経路) によっても影響を受けるが, 実際の IT 野の細胞は刺激提示後, 約 200 msec の遅延時間で反応しはじめる. これはニューロンのシナプス遅延を約 1 msec を考えると, およそ 200 個の介在ニューロンが存在すると考えられる. これらのニューロンの軸索の平均長 (数 μm 〜 数百 μm) を 500 μm と仮定しても, 200 msec 後までに信号は 10cm 先にまでしか伝わらない. 実際の皮質上での距離を考えるとフィードバック計算による影響があるのはニューロンの発火がピークになる 400 msec あたりと考えられる. したがって, ここでは近似モデルとして, 求心性経路のみを考えている.

V1 野の多くの機能地図は 第 2 章でも示すように, ある種の競合学習によってモデル化が可能であるが, ネオコグニトロンの各階層に含まれる特徴抽出細胞もまた競合学習によって決定される. 第 3 章では, ネオコグニトロンの各階層を構成する構造が, 生理学的にも妥当であることを示す. Logothetis らは, サルにワイヤクリップ状の物体を回転させたり拡大縮小させた画像を見せて IT 野の細胞の反応を記録している. この Logothetis らの実験とネオコグニトロンの最終層の反応との比較を行い, 非常に良く似た反応が得られることを示す.

Bricolo と Poggio や Risenhuber と Poggio たちも, この IT 野のモデルを提案して

いる [12] [82]. 彼らのモデルも求心性経路のモデルであり関数近似で用いられる RBF ネットで近似されるとしている. しかし彼らのモデルでは結合の局所性と皮質を構成する細胞の一様性については説明できない. ネオコグニトロンは細胞の持つ結合は局所的で, 構造に一様性を持つので彼らのモデルより現実的な近似モデルであると考えられる.

第 4 章では, ネオコグニトロンを工学的に応用したパターン認識器としての側面について述べる. ネオコグニトロンはパターン認識器として知られているが, 従来までは小規模な文字データベースによって, その可能性が示されていたにすぎなかった. この章ではロボットビジョンや RWC(Real World Computing) で用いられているような大規模なデータベースによって認識能力の評価を行っており非常に高い性能が得られることを述べる. 第 4 章ではネオコグニトロンに単純パーセプトロンの様な単層ネットワークを付け加えたシステムによって構成されている. 文字パターンの認識はパターンをシンボル (数字の場合は '0' ∼ '9' までの数) に変換する操作である. ネオコグニトロンはパターンに含まれる形状 (特徴) によって分類するシステムであり, 文字認識を行うためには最終的には分類したカテゴリに対し何らかのシンボルをつけなければならない. 従来この操作は最終層の細胞に学習パターンが属するカテゴリのシンボルをつけることで行っていた. 学習パターンがカテゴリを代表するような特徴を保持している場合や, 学習パターンが少ない場合はこの方式でも問題ないと思われるが, 学習パターンが増えるにつれ他のカテゴリの文字パターンが持つような特徴を持ったパターンが増えてくる. これらのパターンを正しく分類するために, パターンの形状により分類されたネオコグニトロンの出力をシンボルに結び付ける回路が必要になってくる. この最終的なパターンとシンボルとの結びつけを単純パーセプトロンの様なネットワークに行わせている. このようなシステムを構築し, 通産省電子総合研究所が配布している文字パターンデータベース ETL-1 を用いた認識実験を行う.

第 5 章では, ネオコグニトロンに対しフィードバックの経路 (遠心性経路) を導入した "選択的注意のモデル" を用いた画像処理システムを構築し, 従来の計算機モデルでは困難だった複合パターンの認識・切り出しについて実験を行う. 本来の脳の視覚経路にも, 第 3 章で述べた求心性経路とペアになるような遠心性経路が多く観察される. 遠心性経路の機能については様々な仮説が提示されているが, 特に "注意" を向けるという能力に関係するであろうと言われている. 第 5 章で用いたパターンは英字筆記体文字列であ

る．ロボットビジョンや従来のパターンマッチング的手法では, 入力画像パターンがある特定のカテゴリに属する等の暗黙の了解があったが, 実際のカメラから取り込んだ画像パターンは, ほとんどが, このような複合パターンである．パターン認識を行うためには, パターンを切り出す必要があるが, 画像中のどこにパターンが含まれるかを判定するためにはパターン認識を行わなければならないという, “卵と鶏とどちらが先か” と類似した問題にぶつかってしまう．ヒトは無意識のうちにこの種の問題を解いてパターン認識を行っているのである．我々の構築した選択的注意機構を用いたシステムでは求心性経路のサブシステムでパターン認識を行い, 遠心性経路のサブシステムで認識したパターンにのみ注意を向けた想起をさせることによって文字の切り出しを行う．このような認識と切り出しを同時に行うようなシステムを構築することによって, 複合パターンを認識させることを試みている．

第 2 章

1次視覚野のディレクションマップの形成

2.1 研究背景

Hubel と Wiesel が方位選択性を持つ細胞をネコの 18 野や, サルの V1 野で発見して以来 [41], [45], 方位選択性を持つ細胞がこれらの領野でどのように配置されているかについては様々な議論があった. Hubel と Wiesel の微小電極を用いた生理学実験によれば, 電極を皮質に対して垂直に刺して, 細胞の活動電位を観測した場合, 観察される細胞群は, ほぼ同じ位置と傾きを持ったエッジや線分に対して強く反応し, 皮質に対して水平方向に刺した場合には, 細胞のもっとも強く反応する線分の傾き (最適方位) が連続的に変化すると報告されている. これらの観察結果から Huble と Wiesel は皮質には "カラム" と呼ばれる柱状構造が存在し, カラムが機能単位を構成しているという仮説を提案した. 上述の方位や位置のほかにも, サルの V1 野などの初期視覚野のカラムの検出する機能としては, 左右どちらの目からの入力の影響が強いか (眼優位性) や, どのような色に反応するかなどが考えられてきた. カラムを皮質における機能単位と考えたときに, 皮質上でのカラムの配置は機能地図 (マップ) と呼ばれる. 方位に関するマップは方位マップ, 眼優位性に関する機能地図は眼優位マップ, 提示位置に関する機能地図はレチノトピーなどと呼ばれる. 方位マップと眼優位マップに関して Hubel と Wiesel は, "アイスキューブモデル (icecube model)" と呼ばれる構造を提案した [45]. アイスキューブモデルはカラムの方位選択特性が連続的に変化するように直方体のカラムを配置していったようなモデルである (図 2.1 (a)) これに対し, Braitenberg らや, Götz や, Baxter と

Dow は, それぞれある点 (特異点) を中心に一周すると最適方位が連続的に 180 °変化するようなカラム配置を提案した. このモデルは “ピンホィールモデル (pinwheel model)” と呼ばれる (図 2.1 (b)) [11] [35] [3]. 近年, 脳の表面を直接 CCD カメラなどで捕らえて, 画像処理を施すという技法 (オプチカルレコーディング) 等の計測技術の進歩により, 哺乳類の視覚 1 次野のカラムの機能地図の特徴などが解明されてきている. これによると 数 mm × 数 mm という局所的な皮質上ではピンホィールモデルのような特異点を持つ構造であることが知られている [9].

一方, このような機能地図がどのようなメカニズムで自己組織化されるかということも多くの議論をよんできた. Linsker や Miller や Tanaka はカラム間の相関学習によって形成するモデルを提案している [56] [65] [63] [64] [109] [68]. また Obermayer らや, Durbin と Mitchison は, SOM や エラスティックネットモデルといった競合ヘブ学習を用いたモデルによって形成するモデルを提案している [74] [73] [17]. このような数理的な自己組織化モデルは新しい機能地図の特性の予測や, モデルの単純さゆえに形成メカニズムの理解に役立つ. Obermayer と Blasdel はオプチカルレコーディングの技法を用いて方位マップと眼優位マップの関連性に観察し, 方位の特異点は単眼性の領域に存在する傾向があるということを報告している [72]. この性質は SOM の信号空間上での位相幾何学的な解釈を行うことにより理解が可能である [50]

最近の報告によれば, フェレットの 18 野や, サルの V1 野において方位マップだけでなく, 方向マップの存在が確認されている [115] [91]. 方向マップとは, 線分等の動きに対する機能地図のことである. 本章では方向マップと方位マップのジョイントマップを数理的な自己組織化モデルによって形成し, 実際に観察される皮質地図との関連性を調べている. 自己組織化においては, 基本的に Kohonen の SOM を用いているが, 方向マップと方位マップのジョイントマップを考える際には, 従来の SOM では対応できない点があるため, 学習方法を改良し, 射影学習を用いた SOM を提案している. この自己組織化モデルを用いて形成したジョイントマップは, 実際に皮質で観察される方向マップと非常に多くの点で類似性を示している. 特に SOM の信号空間上で位相幾何学的な解釈を行うことによって, 方向に関する特異点 (線) が, 方位に関する特異点の間を結ぶような傾向にあるという現象を説明することに成功した [101].

まず, “方位 (オリエンテーション)” と “方向 (ディレクション)” の関係について考え

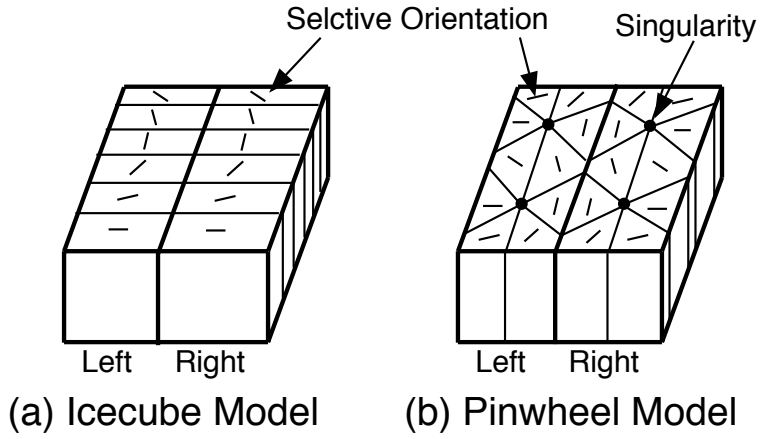

図 2.1: 方位と眼優位性に関する機能地図モデル. (a): アイスキューブモデル. (b): ピンホィールモデル. [45] [35] を変更.

る. ここで述べる方位とは, 提示された線分の傾きを表し, $(0 \sim \pi)$ までの角度で表されるものとする. 方向とは, その線分の移動方向まで区別した言いかたである. したがって, 方向は 0 から 2π までの値をとる. V1 野における細胞は小さな受容野 (約 1.3 度) しか持たないので [83] [52], 小窓問題 (aperture problem) がおきると考えられる. これは線分やエッジの動きが小さな窓を通して観測した場合, 線分の傾きに垂直な成分の速度しか測定できないという問題である. したがって, 本章では方向は線分の方位に対し垂直な成分しか検出できないということを仮定している. したがって一つの方位には正反対の二つの方向が対応している. 皮質上に形成される方位マップと方向マップとは互いに矛盾のないようにできているはずであるが, これらのマップの関係には二通りのありようが可能である.

一つは方位マップ上で, ある方位に対応する一個の領域が一つの方向だけに反応し, それと正反対の方向に反応する細胞は, 別の場所に存在しているという可能性である (図 2.2 の ケース 1). もう一つの可能性は, 一つの方位選択領域が正反対の向きを持つ 2 つの方向領域に分割されているという可能性である. すなわち, 正反対の方向に反応する方向領域が隣接して一つの方位領域をなしている場合である (図 2.2 の ケース 2).

フェレットの 17 野で観測されるオプチカルレコーディングの結果を図 2.3 に示す

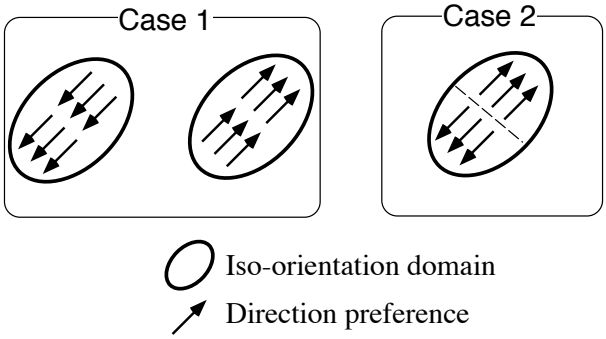

図 2.2: 方位マップのある単一カラム上で考えられる方向マップの在り方. ケース 1: 一つの方位カラムが, 一つの方向カラムによって形成される場合. ケース 2: 一つの方位カラムが, 反対方向の運動に反応するような複数のカラムによって形成される場合.

(図は文献 [115] より抜粋). 図 2.3 のカラースケールは方位マップを表しており, 矢印が方向マップを表している. 図 2.3 下図は, 拡大図で特異点近傍のマップを示している. 同様の方向・方位マップは, ネコの 18 野においても観察されている [91]. ネコ 18 野とフェレット 17 野における方向マップと方位マップとの関係は図 2.3 上図の領域 1, 2, 3 に示されるように, 図 2.2 の Case 2 に対応している.

また Tanaka と Shinbata によって行なわれた自己組織化のシミュレーションでは, マカクザルの MT 野の方向選択性と方位選択性の関係が, フェレットの 17 野で示されたと同様に, 類似の方位に反応する領域が正反対の方向に反応する領域に分割されうることを示している [111][112].

このような皮質マップを形成する自己組織化のモデルとして Kohonen の SOM があげられる [54]. このモデルによる方位マップの形成や, 眼優位性マップ等の形成の計算機シミュレーションがなされており, 計算機実験によれば実際の皮質マップと似たような特徴を持つ機能マップが形成されることが確認されている [75][18][50] [100][99]. しかしながら, 本研究で用いる信号空間を Obermayer らの流儀にしたがって定式化すると信号空間は凸空間にはならず SOM のリファレンスベクトルが信号空間からはみだす恐れがある. そこで本章では, まず, この SOM の学習法を射影学習を用いるように拡張したものを提案する. 次に提案したアルゴリズムを用いて方向マップ及び方位マップを計算

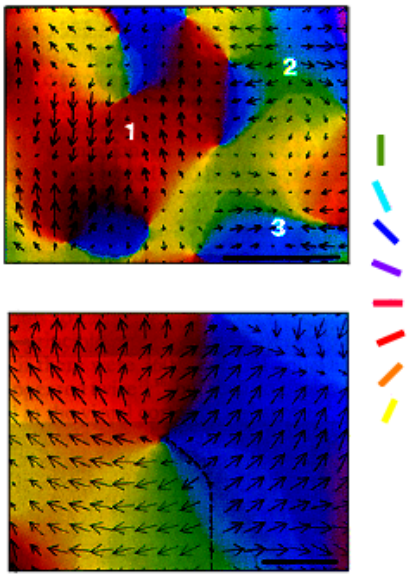

図 2.3: フェレット 17 野における方位マップと方向マップの関係. 図中の色が方位をあらわし, 矢印が方向をあらわしている. 図中のスケールバーの大きさは 1mm ([115] より).

機上で形成したものを示し, フェレットの 17 野や, ネコの 18 野に見られるように一つの方向ドメインが正反対の方向を持つ二つの方向マップの領域に分けられるような機能マップが形成されることを確認する. 最後に, 信号空間の位相幾何学的な解釈を行うことにより方向に関する特異点 (線) と方位に関する特異点の関連を述べる.

2.2 定式化

2.2.1 信号空間

まず SOM における信号空間を考える. 状況としては, ある方位を持った線分を, 適当な速度 v で動かした状況を想定する. 研究の第一段階としては, このような入力全体のなす空間の位相幾何学的構造を再現する, すなわち, できるだけ単純な信号空間を考えることが必要である.

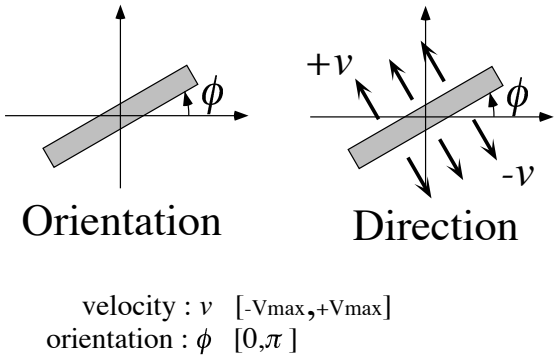

図 2.4: 方向と方位に関するパラメータの定義. 方位に関するパラメータ ϕ は $[0, \pi]$ の範囲で与えられる量であり, 周期性を持つ. 方向に関しては ϕ とは別に, 新たなパラメータ v を導入する. v は線分の ϕ と垂直方向の速度をあらわすパラメータである.

まず始めに, 方向選択性と方位選択性の関係について考えてみる. 方位選択性は方位に対する反応の強さ (選択強度) を除いて考えれば一つの線分の方位 ϕ のみによってあらわされる関数のはずである. また, この関数は線分の方位の性質上, 周期 π の関数でなければならない. これは線分を π だけ回転させてみると, 元の線分とまったく重なってしまうためである. 次に方向選択性について考える. これは線分の方位 ϕ と線分がどちらに動くかということによってあらわされる関数となる. Feret 17 野などの初期視覚野の受容野の大きさを考えると, 小窓問題により, 線分の方位に対して垂直方向の運動のみしか検出されないと考えてもよいはずなので, この関数は線分の方位 ϕ と線分の方位に垂直な運動の速さ v の関数になるはずである. この v の正負で動きの方向を定めるものとする (図 2.4 参照).

図 2.5 は, この線分の運動と信号空間の関係について示したものである. この信号空間は ϕ と v という二つのパラメタで記述されるので帯状の領域となり, パラメタの組 (ϕ, v) で一つの運動を表すことになる.

さて, ここで方位 $\phi = \phi_\epsilon$ の線分が $v > 0$ で動く場合と, 方位 $\phi = \pi - \phi_\epsilon$ の線分が $-v$ で動く場合を考える (図 2.6 参照). この場合 $\phi_\epsilon \to 0$ なる極両者の動きは一致して見えるはずなので, 信号空間上でもこの 2 点を一致させなければならない. すなわち, (ϕ, v)

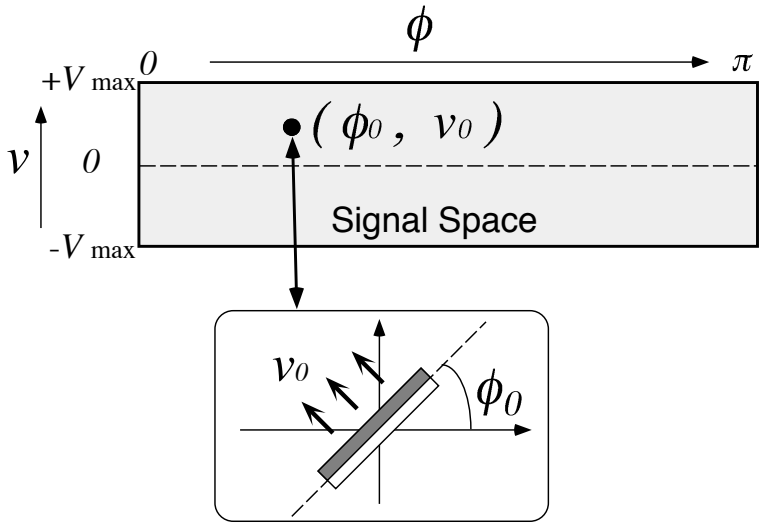

図 2.5: 線分の動きは信号空間上でのある一点に対応する．信号空間は ϕ と v であらわされる 2 次元の空間であり, 傾き ϕ_0, 速度 v_0 の運動は (ϕ_0, v_0) という一意な表現であらわされる．

であらわされる信号空間においては, $(0, +v)$ と $(\pi, -v)$ とは同一の点のはずである．また方位に関する条件より $(0, 0)$ と $(\pi, 0)$ もまた, 同一の入力信号をあらわす点のはずである．

以上の条件を考慮すると方向と方位に関する信号空間は Möbius band によって表現される．方向に関する信号空間をこのように考えればいいことは Tanaka によっても指摘されている．[110].

Möbius Band は 3 次元空間内に埋め込むことができる．一般に SOM の学習では，情報空間の位相構造 (トポロジー) だけが重視され，計量構造は無視されることも多く，最近はこの傾向に関する反省もある [81]．ここでは, 方位 ϕ のシフト (回転) に関する信号空間の対称性をモデルに取り込んでおくことが, シミュレーション上重要であると考えられる．方位 ϕ のシフトに関して不変であるようにするには 4 次元空間が必要となる．つまり

- 方位選択性に関して 2 次元 $(\cos 2\phi, \sin 2\phi)$,

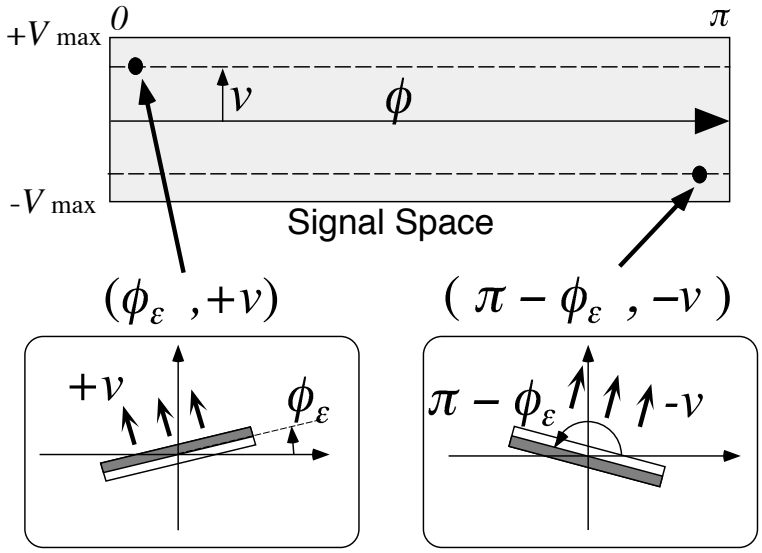

図 2.6: 線分の動き (ϕ_ϵ, v) と $(\pi - \phi_\epsilon, v_0)$ とは $\phi_\epsilon \to 0$ の極限では一致した運動に見える.

- 方向選択性に関して 2 次元 $(v\cos\phi, v\sin\phi)$

の直積をとって $(\cos 2\phi, \sin 2\phi, v\cos\phi, v\sin\phi)$ とする. これに方向・方位選択性に関する感度を表すパラメタ q を乗し, レチノトピー (網膜上の位置) を表す (x, y) を加えると, 神経場の細胞の特性 (リファレンスベクトル), と入力を表す 6 次元ベクトル $\boldsymbol{\xi}$ は以下のように書ける.

$$\boldsymbol{\xi} = \boldsymbol{f}(\boldsymbol{\eta}) = (x, y, q\cos 2\phi, q\sin 2\phi, qv\cos\phi, qv\sin\phi)^{\mathrm{T}}. \tag{2.1}$$

ただし, これらのパラメタ $\boldsymbol{\eta}$ がとりうる値の範囲は

$$Y = \{\boldsymbol{\eta} | x, y \in [0, d], \phi \in [0, \pi], q \in [Q_\epsilon, Q_{\max}], v \in [-V_{\max}, V_{\max}]\}, \tag{2.2}$$

とする.

パラメタ q は値が大きいほど方向・方位選択性の感度が強く, その方向・方位の線分に, より特異的に反応することを示す. ϕ は方位選択性を表すパラメタである. v は方向選択性の強度を表すパラメタで, 正の向きに動いた場合と負の向きに動いた場合が考えられる.

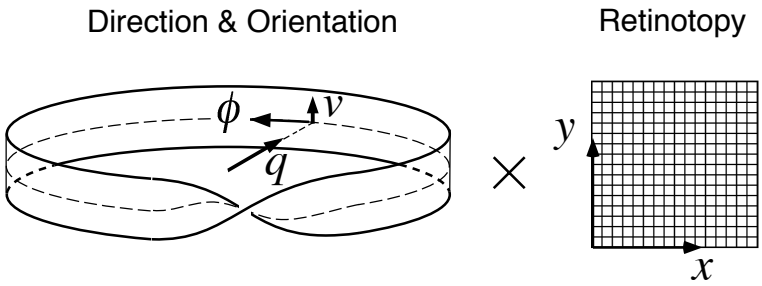

図 2.7: 全ての入力情報を与えたときの信号空間. この空間は Möbius band 上の空間 (方向と方位に関する部分空間) とレチノトピーをあらわすような 2 次元平面の直積空間であらわせる. Möbious band 状の空間は 図 2.5 であらわしたような 2 つのパラメータで描ける.

入力次元に関してもう少し詳しく考えてみると, 方向選択性に 2 次元を割り当てているように見えるが, 方向選択性として新たに導入されたパラメタは v だけである. 結果として信号空間は, 4 次元空間内に埋め込まれた Möbious Band (2 次元) と, その中央に位置する原点を結んでできる 3 次元多様体を 2 次元のレティノトピーと直積したものとなる (図 2.7 参照). この信号空間を M とする:

$$M = \boldsymbol{f}(Y) \tag{2.3}$$

したがって M は $\boldsymbol{R}^6$ に埋め込まれた 5 次元多様体である.

2.2.2 射影学習

パラメタ空間 Y の中にリファレンスベクトルと入力信号を与えて SOM で学習させる方法も考えられるが，それでは Y の持つべき位相幾何学的構造を学習に反映させることが難しい．そこでここでは，6 次元で表現された信号空間 M を使うことにする．信号空間 M を 2 次元の神経場にマップすることになるが, 神経場の位置を $\boldsymbol{r}$ であらわすとすると, 位置 $\boldsymbol{r}$ の細胞が持つリファレンスベクトルは,

$$\boldsymbol{w_r} = (x_{\boldsymbol{r}}, y_{\boldsymbol{r}}, q_{\boldsymbol{r}} \cos 2\phi_{\boldsymbol{r}}, q_{\boldsymbol{r}} \sin 2\phi_{\boldsymbol{r}}, q_{\boldsymbol{r}} v_{\boldsymbol{r}} \cos \phi_{\boldsymbol{r}}, q_{\boldsymbol{r}} v_{\boldsymbol{r}} \sin \phi_{\boldsymbol{r}})^{\mathrm{T}} \tag{2.4}$$

となる. リファレンスベクトル $\boldsymbol{w_r}$ も M 上にある. この状況下で Kohonen の SOM に従って学習させることを考えてみる.

まず, 信号空間 M 内の 1 点をランダムに選択し, これを入力 $\boldsymbol{\xi}$ とする. 次に入力 $\boldsymbol{\xi}$ にもっとも近いリファレンスベクトルをもつ神経場上の細胞を $\boldsymbol{s}$ とする.

$$\boldsymbol{s} = \underset{\boldsymbol{r}}{\operatorname{argmin}} |\boldsymbol{w_r} - \boldsymbol{\xi}| \tag{2.5}$$

ここでの絶対値は $\boldsymbol{R}^6$ 内でのユークリッド距離である. 勝者 $\boldsymbol{s}$ の近傍に含まれる細胞 $\boldsymbol{r}$ で学習をさせることになるが, 通常の SOM の場合の学習は,

$$\boldsymbol{w_r}(t+1) = \boldsymbol{w_r}(t) + \epsilon h(\boldsymbol{r}, \boldsymbol{s})(\boldsymbol{\xi} - \boldsymbol{w_r}(t)) \tag{2.6}$$

で表される. ただし $h()$ は近傍関数で, 通常はガウス関数を用いる.

$$h(\boldsymbol{r}, \boldsymbol{s}) = \exp(-\frac{|\boldsymbol{r} - \boldsymbol{s}|^2}{\sigma^2}) \tag{2.7}$$

ところが, われわれの信号空間 M は $\boldsymbol{R}^6$ の中に浮かんだ 5 次元の曲面であるため, M 上の 2 点を結ぶベクトルとして与えられるリファレンスベクトルの更新要素

$$\Delta \boldsymbol{w_r} = \epsilon h(\boldsymbol{r}, \boldsymbol{s})(\boldsymbol{\xi} - \boldsymbol{w_r}(t)) \tag{2.8}$$

では, 更新後のリファレンスベクトル $\boldsymbol{w_r}(t+1)$ が信号空間 M から飛び出してしまうことがある. そこで, われわれは次に述べる射影学習を提案する. 射影学習は, もともとユークリッド空間で定義された SOM の学習法をユークリッド空間内の曲面上で行なえるよう拡張したものである. これは, 式 (2.6) における微小な更新要素 $\Delta \boldsymbol{w_r} \in \boldsymbol{R}^6$ に代えて, これを M 上に射影したものを新たな更新要素として採用することを方法である (図 2.2.2 参照). 位置 $\boldsymbol{r}$ の細胞に関する更新要素 $\Delta \boldsymbol{w_r}$ を計算しておいてから, これを M 上に射影し, $\boldsymbol{\theta}$ 座標系で表現したベクトルを $\Delta \boldsymbol{\theta} \in \boldsymbol{R}^5$ とする. リファレンスベクトルは $\boldsymbol{\theta} = (x, y, q, v, \phi)^{\mathrm{T}}$ の形で保存し, その更新にも $\Delta \boldsymbol{\theta}$ を用いる. 勝者細胞を決定する際の距離の計算と, 以下に述べる $\Delta \boldsymbol{\theta}$ の計算に 6 次元の表現を用いる.

更新要素の信号空間への射影を求めるために, 信号空間上の $\boldsymbol{w_r}$ 近傍で接平面近似を

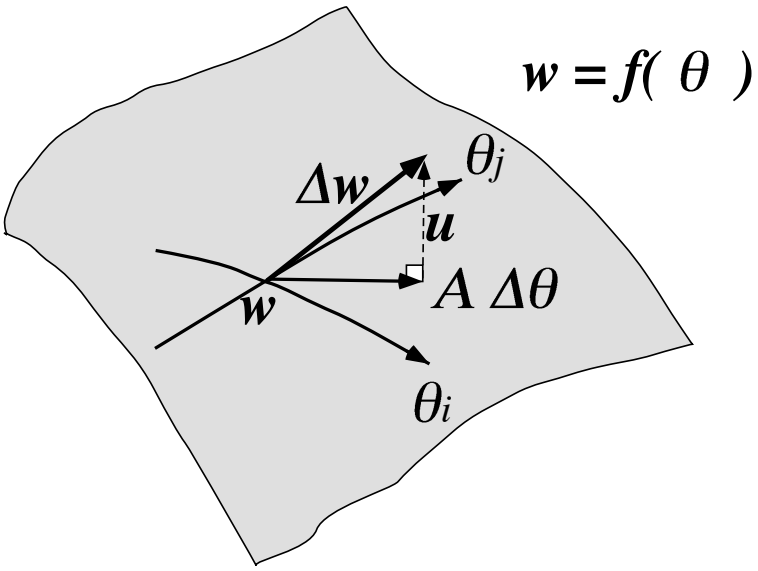

図 2.8: 多様体 M 上に射影した更新ベクトル $\Delta\boldsymbol{w}$ は従来の SOM の更新ベクトルをあらわす. $\boldsymbol{u}$ は $\Delta\boldsymbol{w}$ の M に対する垂直な成分. $A\Delta\boldsymbol{\theta}$ は M に平行な成分をあらわす. 我々は $A\Delta\boldsymbol{\theta}$ を新たな更新ベクトルとして扱う.

おこなう. まず $\boldsymbol{f}(\boldsymbol{\theta})$ の偏微分行列 $A(\boldsymbol{\theta})$ を導入する.

$$
\begin{aligned}
A(\boldsymbol{\theta}) &= \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}} = \left(\frac{\partial f_i}{\partial \theta_j}\right) \\
&= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & \cos 2\phi & 0 & -2q\sin 2\phi \\
0 & 0 & \sin 2\phi & 0 & -2q\cos 2\phi \\
0 & 0 & v\cos\phi & q\cos\phi & -qv\sin\phi \\
0 & 0 & v\sin\phi & q\sin\phi & qv\cos\phi
\end{pmatrix}
\end{aligned}
$$

6 次元空間内の更新要素 $\Delta\boldsymbol{w_r}$ は, $\boldsymbol{\theta_r}$ における M の接平面内のベクトル $A(\boldsymbol{\theta_r})\Delta\boldsymbol{\theta_r}$ と接平面に垂直な成分の $\boldsymbol{u_r}$ とに分解することができる (図 2.2.2 では, 神経場上での位置を示すサフィックス $\boldsymbol{r}$ を省略している). これを式で表すと

$$
\Delta\boldsymbol{w_r} = A(\boldsymbol{\theta_r})\Delta\boldsymbol{\theta_r} + \boldsymbol{u_r} \tag{2.9}
$$

となる. この接平面と, ベクトル $\boldsymbol{u_r}$ とは直交しているので,

$$\boldsymbol{u_r}^{\mathrm{T}} \cdot A(\boldsymbol{\theta_r}) = 0 \tag{2.10}$$

である. 以上の式 (2.9) と式 (2.10) とから

$$\Delta\boldsymbol{\theta_r} = (A_{\boldsymbol{r}}^{\mathrm{T}} A_{\boldsymbol{r}})^{-1} A_{\boldsymbol{r}}^{\mathrm{T}} \Delta\boldsymbol{w_r} \tag{2.11}$$

が導かれる. ただし, ここで $A(\boldsymbol{\theta_r})$ を $A_{\boldsymbol{r}}$ と略記した.

式 (2.6) と式 (2.11) によって得られる $\Delta\boldsymbol{\theta_r}$ によってリファレンスベクトル $\boldsymbol{\theta_r}$ を更新する. ここで, 更新式である 式 (2.6) を以下のように $\boldsymbol{\theta_r}$ に関する更新式へと変更する.

$$\boldsymbol{\theta_r}(t+1) = \boldsymbol{\theta_r}(t) + \Delta\boldsymbol{\theta_r} \tag{2.12}$$

このように SOM の更新式の変更を行なうことによって各細胞が持つリファレンスベクトルが M 上からはみでることはなくなる.

M の境界を適当に貼り合わせ, ϕ のシフトに対して不変なリーマン計量を M 上に定義して, 学習を M の中だけで行うことも不可能ではないが, そうすると M 上の二点間の距離の計算の度に二点間の最短経路を求めることになるので, 膨大な計算時間が必要になり, およそ現実的な方法とはいえない. 逆に, 自明でない計量が与えられた空間に SOM を適用する必要があるとき, もしその空間を高次元のユークリッド空間にうまく埋め込めるなら, ここに提案した射影学習は極めて有効な方法になるだろう.

また $\boldsymbol{f}()$ が恒等写像の場合, $A(\boldsymbol{\theta})$ は単位行列となるので, 従来の SOM の学習法に帰着する.

2.3 シミュレーション結果

計算機シミュレーションでは神経場の大きさを 128×128 細胞とし, 信号空間のサイズを定めるパラメータを $Q_{\max} = 17.0$, $Q_\epsilon = 0.01 \times Q_{\max}$, $V_{\max} = 2.0$ として実験を行なった. 学習は M 上の一様分布に従う $1,000,000$ 個の入力によって行なった. Kohonen マップは一般に収束性が保証されていない. そのため学習強度 ϵ の値を学習回数によって減少させて収束させるなどの手法が用いられるが本研究では学習強度 を一定の値にし

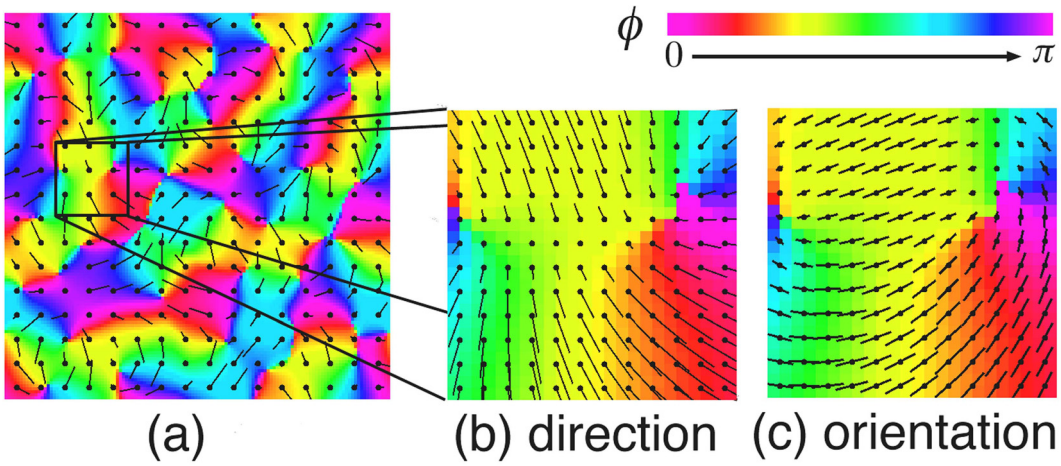

図 2.9: 計算機シミュレーションによって形成したジョイントマップ．図中の色は方位をあらわす．(a): マップ全体をあらわしている, 各グリッド上の棒が選択的な方向を示す．棒の長さは速度選択性をあらわし, 長いほど大きな速度に対応していることを意味する．(b): (a) の一部を拡大表示させたもの．図は方向に関する特異点 (v = 0) となるような領域を含む．この場合特異点は線状になる．(c): (b) と同じ領域を方位マップとして表したもの．図中の棒の傾きが方位をあらわす．棒の長さは方位選択性をあらわし長いほど選択性が高いことを意味する．方向マップで特異点となるような領域でも方位マップでは特異点とはならない．このことから単一の方位選択カラムが二つの正反対の方向選択性を持つカラムによって形成されていることがわかる．

たままでシミュレーションを行なっている. このためマップとしては収束しないが, 入力が 700,000 個を越えた付近でマップの変化は非常にゆっくりとなり, 定性的な性質が変化しないと判断し, 1,000,000 個の入力をしたところでシミュレーションを打ち切った.

シミュレーション結果を図 2.9 に示す. 図 2.9 (a) は全体の方向マップと方位マップを表している. グレースケールは方位マップを表しており, 各グリッドから伸びている線分が方向選択性を示す. 線分の長さはその方位にどれだけ強く反応するか (選択強度) をあらわしている. 図 2.9 (b) は 図 2.9 (a) の一部領域の拡大図を示している. 図 2.9 (c) は 図 2.9 (b) と同じ領域での方位マップを線分で表しており, 線分の方位は選択方位を示し, 線分の長さは選択強度を表している.

図 2.9 (b) の中央に観察される左斜め方向に走る方向特異性が非常に小さい領域は正反対の方向に反応する領域の境界である. この領域を方位選択性マップで見ると 図 2.9 (c) のようになっており, 等方位に反応する領域であることがわかる. すなわち, 図 2.9 (b) では方位選択性マップで, 同じ方位に反応するパッチ状の領域が, 二つの反対向きの方向に反応している領域によって分割されていることが示されている. これは 図 2.3 が示す特性と等しい.

いま方位マップの特異点の周りを一周するような軌道を考えるとする. 特異点の回りでは方位が π だけ変化しているものとする. この軌道は, すなわち図 2.10 下図のように点 A から出発して A に戻ってくるような軌道となる. この, 信号空間での軌道は 図 2.10 上図のように $v = 0$ の点 (点 P) と交わらないと ϕ が π だけずれた時に出発した点 (点 A) に戻ってくることができない. したがって方位の特異点の周囲には少なくとも 1 つ以上の方向の特異点のラインが存在している. Tanaka は位相幾何学的な考察から “方向の特異点の線が方位の特異点間を結ぶように形成される” ことを既に指摘しており [110], 本考察と合致している.

図 2.11 は, 方位に関する特異点 ($q = 0$) を黒であらわし, 方向に関する特異線 ($v = 0$) を表したものである. 図 2.10 のように方位に関する特異点からは, 方向に関する特異線が出ていることが分かる.

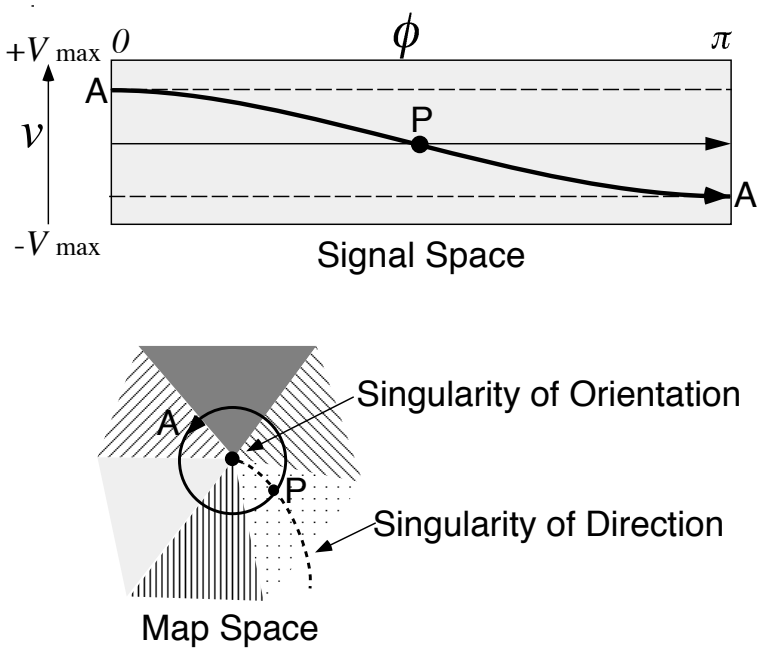

図 2.10: 方位に関する特異点の回りで一周するような経路が与えられたとき, この経路を信号空間上で考える. Möbious band 状の信号空間をなしているので, この経路は必ず一回は $v = 0$ の線を横切る形になる (図中では点 P で交わっている). すなわち方位の特異点からは, 最低でも一本の方向に関する特異線が出ていることになる.

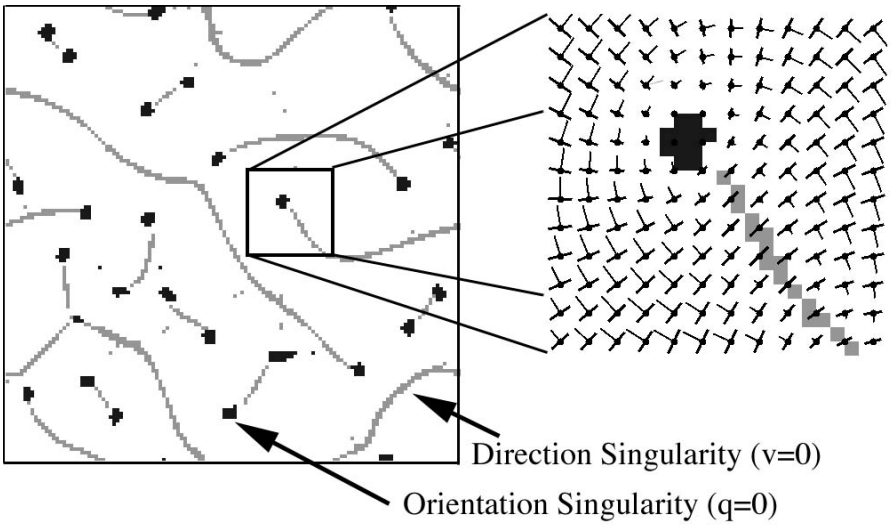

図 2.11: 計算機シミュレーションを用いて方位に関する特異点と方向に関する特異点を表したもの. 暗いグレーであらわした点状の領域方位に関する特異点をあらわしている ($q = 0$). 明るいグレーであらわされた線状の領域が方向に関する特異線をあらわしている ($v = 0$). 方向特異線は, 方向特異点間を結ぶような形になる. 右図は, ある方位特異点の回りを拡大したものを表す. 各グリッド上の棒の組合わせは方向と方位をあらわしており, 線分が長いほど特異性が大きいことをあらわす.

2.4 検討

本章では, SOM を拡張する射影学習ルールを新たに提案し, 方向マップと方位マップに関するジョイントマップを射影学習を用いた SOM によって形成した. その結果, フェレットの 17 野やネコの 18 野で観測されるような方向選択性の特性を持つ機能マップが形成されることを示した. また, 方向マップの特異点が等方位に反応する領域を横切る理由を信号空間上での解釈で説明した.

この研究に関連して Tanaka と Shinbata らは, MT 野のモデルにおいて相関学習を用いるタイプのモデルによる方向選択マップの形成シミュレーションをおこなっており, 細胞が持つ方向選択強度と方位選択強度の比率と, 方向マップの特異点の発現の仕方について述べている [111] [112]. この比率は我々の計算機実験によっても確認されており, 速度の絶対値 $|V_{\max}|$ がある一定値を越えると急激にディレクションマップが形成されるようになる.

また Miyashita と Tanaka のモデルによるシミュレーションにおいてもこのような方位マップの不連続線が発生することが確認されている [69][67]. ただし Miyashita と Tanaka のモデルは, 入力に白色雑音を仮定しカラム間の相関学習によって形成している. 本章で考えた SOM の様な競合ヘブ型のモデルでは入力に実際の線分のようなものを仮定しているので, 実際の V1 野の形成メカニズムを, よりよく再現しているものと考えられる.

第 3 章

視覚経路とネオコグニトロン

3.1 研究背景

人やサルの視覚処理課程においては, 大まかにいって二つの経路が存在することが生理学的, 解剖学的に知られている. これらの経路は “背側経路 (dorsal pathway)” と “腹側経路 (ventral pathway)” と呼ばれている [19]. 物体の動きなどの認知は, 主に背側経路で処理され, 物体の形状などの認知は腹側経路で処理されているといわれている. この章では, 我々が取り扱うモデルであるネオコグニトロンと腹側経路との関係について論じる.

サルの腹側経路は階層構造をもち, 視覚情報は, 網膜 → LGN → V1 野 → V2 野 → V4 野 → IT 野 (inferior temporal cortex) というように, 処理されていいくことが知られている [52].

このうち, 特に V1 野の研究が進んでおり, さまざまな性質が知られている. 中でも Hubel と Wiesel による研究が最も有名である. 当初 Hubel と Wiesel はネコの 18 野 (V1 野に相当する) に微小電極を挿入し, 視覚刺激と神経細胞の発火状態について調べている. 彼らの報告によれば, 18 野には主に 2 種類の細胞が存在し, それらは “単純型細胞” と “複雑型細胞” と呼ばれている. 18 野 や V1 野における単純型細胞の多くは, 受容野内のある特定の位置に提示された特定の傾きをもつような線分や, エッジに対して反応することが知られている. ただし単純型細胞では刺激の位置がずれると反応が小さくなる. これに対して複雑型細胞は, ある特定の傾きを持つ線分やエッジなどが受容野

内のどこに提示されても反応するという性質を持つ.

近年では, 線形応答関数の手法を用いて受容野の時空間的な形状が確認されている [49]. V1 野 における単純型細胞の受容野の時空間的な形状は, Gabor 関数により表現されるものと考えられる. つまり単純型細胞の役割は, 画像に含まれるエッジや線分のような微小要素を Gabor フィルターなどによって抽出するものと考えられる. 同様に複雑型細胞の受容野の空間的な形状も知られている.

ネオコグニトロンは, Fukushima によって提唱された階層型の神経回路モデルの一種である [23], [28]. ネオコグニトロンは主に 2 種類の細胞ユニットから構成される. 一つは 'S 細胞' と呼ばれる細胞で, もう一つは 'C 細胞' と呼ばれる細胞である.

S 細胞 は, 単純型細胞のモデルであり, C 細胞 は複雑型細胞のモデルである. S 細胞の機能は特徴抽出にある. S 細胞 は特定の位置に特定の特徴が提示された時に大きな反応を出し受容野内の特徴の位置に非常に敏感である. S 細胞 の具体的な出力を計算するプロセスは二つのプロセスからなる. まず, S 細胞 が最も特異的に反応する特徴 (最適特徴 と呼ぶ) と受容野に与えられた入力との類似度を計算する. 次のプロセスは, この類似度に対し非線形変換を施し, 最終的な出力を計算する.

一方の C 細胞 は複雑型細胞のモデルである. 複雑型細胞は受容野内の特徴の位置に依存せず反応を出す細胞である. C 細胞 は, 異なる位置に受容野の中心を持ち, 同じ最適特徴を持つ S 細胞 の出力を集めることで実装している (図 3.1 (c)).

近年, Sakai と Tanaka は複雑型細胞のモデルについて考察を行なっている. 彼らは古くからある複雑型細胞の二つのモデル, すなわち "カスケード モデル (cascade model)" と 呼ばれるモデルと "ノンカスケード モデル (non-cascade model)" と呼ばれるモデルの対比を行なっていおり, カスケード モデルの方が複雑型細胞の生理学的データを良く説明することを示している [88]. カスケード モデルは 3 つの連続した計算プロセスからなる. 第一のプロセスは Gabor 関数を用いた線形フィルタのプロセス, 二つめは伝達関数による変調, 最後のプロセスは, 同じ最適方位を持つフィルタを持つ経路からの入力をまとめることによって, 実現されている (図 3.1 (a)). もう一方の ノンカスケード モデル は LGN のような方位選択性のない ON-center や OFF-center タイプのフィルタの出力を最適方位方向に積をとってから和をとるモデルである (図 3.1 (b)). 双方とも, 最適方位を持つような刺激に対しては受容野内の位置に関わらず高い反応を示すが, 複雑

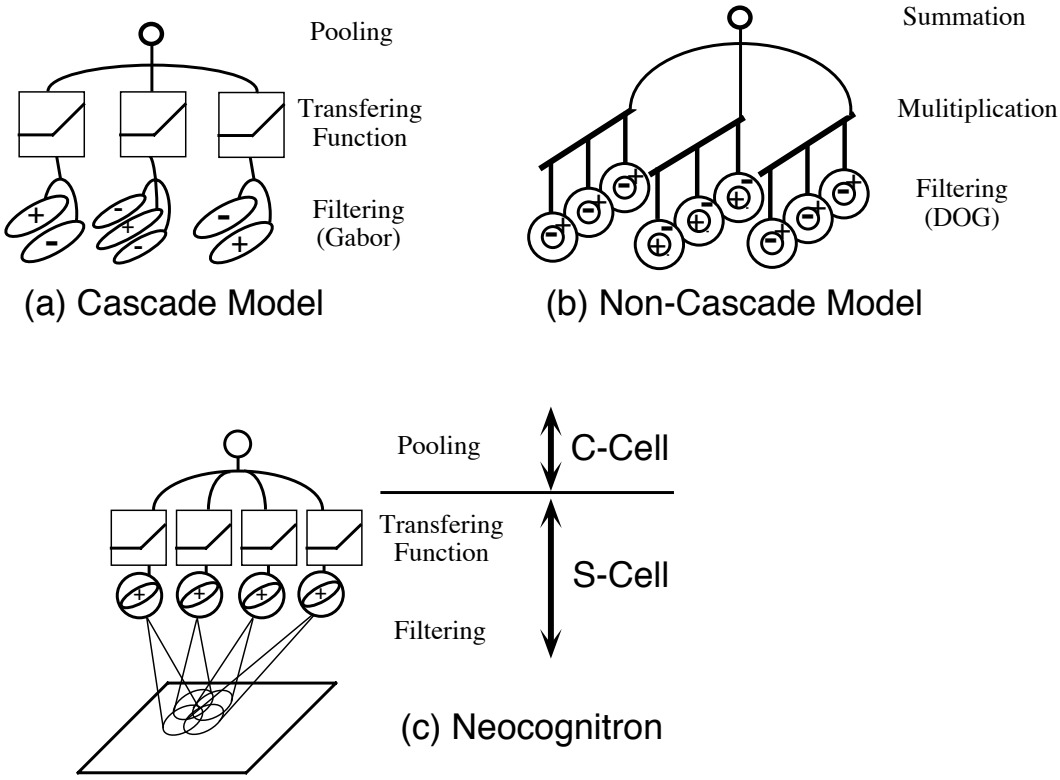

図 3.1: “カスケードモデル” と “ノンカスケードモデル” のアーキテクチャ ([88] を変更) (a): カスケードモデルは, 3つの計算プロセスからなる. すなわちフィルタリングプロセス (Sakai らは Gabor フィルタを用いている), 伝達関数による変換, そして反応の統合というプロセスである. (b): ノンカスケードモデルはフィルタリングプロセス (Sakai らは DOG フィルタを用いている) の出力に対して局所的な積を取り, 全体的な和を取るというプロセスからなる. (c): ネオコグニトロンの段構造をカスケードモデルとした場合の解釈.

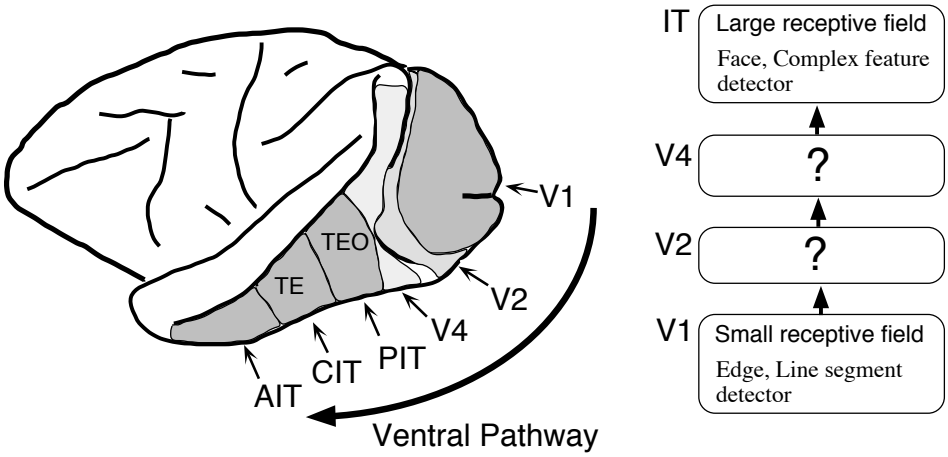

図 3.2: マカクザルの腹側経路を側面から観察した図を摸式化したもの. 各領野の呼称は Felleman と van Essen によるもの [19]

型細胞の受容野内に複数の刺激を提示した場合の応答関数の特性を良く表すのは カスケード モデルである.

ネオコグニトロンにおいて, S 細胞 はフィルタリングのプロセスと変調のプロセスを受け持ち, C 細胞 は複数の S 細胞 の出力をまとめるプロセスを受け持つと考えれば, S 細胞 から C 細胞 への一連のプロセスは, ある種のカスケード モデルと考えることが出来る. したがって, ネオコグニトロンにおける S 細胞 から C 細胞 への計算プロセスは V1 野の複雑型細胞の特性を良く表すものといえる.

残念なことに V1 野以降の腹側経路, すなわち V2 野 → V4 野 → IT 野という経路上の領野の詳細な性質については, あまり良く知られてはいない. そこで我々は V1 野以降の経路における各領野も V1 野 と同じような構造を持つものと仮定した. この仮定を行なうと腹側経路全体はネオコグニトロンと非常に良く似た構造をもつものと考えることが出来る (図 3.2 参照). Rolls らは, 腹側経路上の各領野の細胞の反応時間および, 受容野の大きさの関係に着目して, S 細胞のみを階層化したようなネットワークモデルを提案している [83].

人やサルの IT 野の細胞 (特に AIT 野といわれる領域の細胞) は非常に複雑な図形に特異的に反応することが知られている. Perret らは腹側経路の高次領域である AIT 野

の細胞が顔に反応することを報告している [79]. Logothetis らはこれらの細胞の特性を調べるためにマカクザルにワイヤクリップ状の３次元物体の画像を提示し, そのときの IT 野の細胞の反応を微小電極を用いて測定している [61]. さらに彼らは物体を回転させた画像, 拡大縮小させた画像, 平行移動したような画像を見せて, これらの細胞のコードしている情報を調べている. それによると, 変形に対して不変な細胞というのは少なく, ほとんどの細胞が, ある特定の画像に対して反応して, 物体の回転などによって起こる変形の度合いによって反応が徐々に小さくなっていくという報告がなされている. 本章ではネオコグニトロンを構成する細胞の示す特性と, Logothetis らが示した IT 野における細胞の特性とが非常によく似ていることを示す.

我々の研究に先立って Bricolo らは　IT 野のモデルとして RBF (radial basis function: 放射基底関数) ネットワークを用いたモデルを提案し, Logothetis らの実験を説明しようと試みている [12]. Bricolo らの提案するネットワークは二つの計算プロセスを経て細胞の出力が計算される. 第１のプロセスは, フィルタリングのプロセスであり, Bricolo らは多数のフィルタを予め用意して, 入力された画像をさまざまな成分に分解している. 第２のプロセスは, RBF ネットワークによるパターンの分類である. フィルタリングの出力を RBF ネットワークへの入力とし, RBF ネットワークによって計算される値を最終的な出力としている. 彼らは, この最終出力を出すユニットの特性が Logothetis らの示した IT 野の細胞とよく似ていると主張している. 彼らの提唱しているモデルでは各階層間は非常に広範囲な結合によって実現されているが, 現実の腹側経路における各領野間の結合は局所的なもので構成されており, 彼らのモデルにおける中間層と実際の腹側経路を構成している領野との対応関係がとれない.　最終的には IT 野の細胞は全視野をカバーするくらいの大きさの受容野を持つようになるが, V1 野や V2 野などの受容野は局所的なものでレチノトピーが存在している [52]. このような脳の各領野の階層性を考えれば, 形状を認識する腹側経路のモデルは, 階層性をもち, 各階層で局所的なイメージの情報を徐々に統合していくようなモデルであるほうが望ましい. Riesenhuber らは Bricolo らのモデルを改良して, より腹側経路のモデルとして近いものを提唱している [82]. Riesenhuber らは, Bricolo らの第 1 のプロセスであるフィルタリングのプロセスをレチノトピーを持つ局所的なフィルタリングと抽出した部分特徴を統合するプロセスとで置き換え, それを統合した特徴を RBF ネットワークへの入力として与えてい

る．Risenhuber らはこの局所的なフィルタリングのプロセスを V1 野のモデルとして位置付け, 統合するプロセスを V4 野のモデルとして考えている．しかし, 実際の腹側経路においては各領野は解剖学的には分類されているが, 構成している皮質という意味では均質であると考えられるのでモデルとしては各階層をモデル化したときに, その階層ごとに計算手法が著しく異なるようなモデルは望ましくないと思われる． 残念ながら Riesenhuber らのモデルでは, 各層における計算手法が著しく異なっている．

これに対して我々が提案しているネオコグニトロンによる IT 野のモデルでは, 上述の階層性, 局所性, 均質性の条件を満たす．我々はネオコグニトロンが Logothetis らの実験を説明できることを示し, IT 野までの腹側経路のモデルとして妥当であることを論じる．

3.2 ネオコグニトロン

ネオコグニトロンは階層型のニューラルネットワークモデルの一種である [23], [28]. 本実験で用いたネオコグニトロンを 図 3.3 に示す．ネオコグニトロンは 2 種類の細胞から構成される．この細胞はそれぞれ S 細胞 と C 細胞 と呼ばれている [23].

S 細胞 は視覚領において観察される単純型細胞 [40][41] のモデルであり, 特徴抽出機構として働く．S 細胞 への入力結合荷重は, 可変結合でこれは学習によって決定される．学習終了後においては, S 細胞 はある特定の特徴に対して強く反応するようになる．

C 細胞 は, 同じく視覚領で観察される複雑型細胞のモデルである．C 細胞 に対する入力結合は固定結合である．複雑型細胞は, 受容野内であれば提示される位置に関係なく特定の特徴に対して反応するので, 異なった位置にある, 同じ特徴を抽出する S 細胞の出力を集めることによって実装されている．

ネオコグニトロンは, いくつかの “細胞段” と呼ぶ構造から構成されていて, 一つの細胞段は S 細胞の層 U_{Sl} と, C 細胞の層 U_{Cl} の 2 層構造からなっている．添字の l は, 階層の番号をあらわしている．(図 3.3 参照) 各階層 U_{Sl} および U_{Cl} は “細胞面” と呼ばれる構造から構成されている．細胞面は同じ特徴の抽出を行う細胞を 2 次元平面状に並べたものである．したがって同一細胞面内にある細胞の入力結合の空間分布は同一なものであるが, 特徴を抽出する場所が互いに少しずつ異なっている．よって, 細胞面内で発

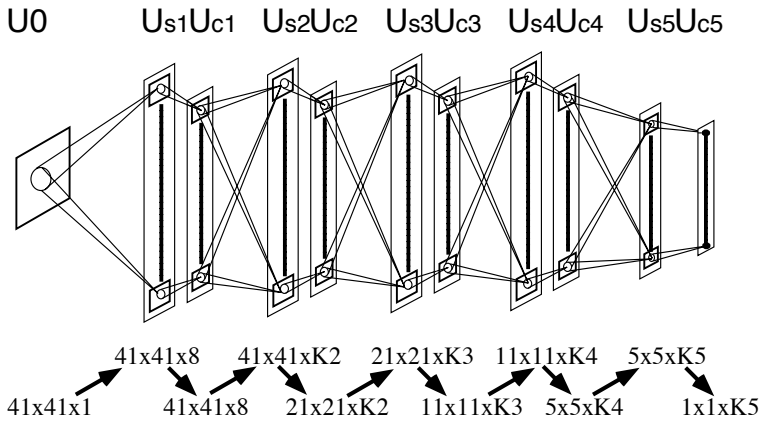

図 3.3: 本実験で用いられたネオコグニトロン. このネオコグニトロンは 5 段構造のネットワークで, 各段は U_{Sl}, と U_{Cl} という 2 層構造からなる. ネットワークの第 1 段目は直線を検出するように結合をつくりつけている.

火した S 細胞 は, そのコード している特徴だけでなく刺激がどこに提示されたかも意味する. また, 同一細胞層に含まれる細胞面内の同じ位置にある細胞は, どの細胞面に属している細胞でも, 入力層上の同じ場所に受容野を持つものとしている. 無論, 解剖学的な立場をとれば皮質内に細胞面のような構造は陽に存在するわけではない. V1 野において細胞はエッジや線分などの微小な特徴に反応することが知られているが, これらの細胞はランダムに配置されているわけではなく, 似た特徴に対して反応する細胞が柱状に配置されている. この構造はカラム構造と呼ばれる. サルの V1 野の場合, カラム内の細胞の抽出するエッジ等の成分の最適方位は, 皮質上で 50 μm 移動すると 10 °の割合で規則的に変化し 1mm 弱の皮質片で 180 °の方位全体をカバーする [40]. 近年の大脳皮質の研究によれば, 視覚皮質の大部分はカラム構造からなり, この構造は V1 野や V2 野だけでなく IT 野などでも観察されることが報告されている [108]. V1 野において, カラムの位置は提示された刺激の位置をコード しており網膜上での像の位置関係 (レチノトピー) が保存されるように配置されている. 細胞面は V1 野のようにレチノトピーが保存されている領野において, 細胞面は異なる位置にある同一な最適方位を持ったカラムを抽出し, 一枚のシート状にレチノトピーを保存するように並べたもの (図 3.4 参照) であると考えることができる [27].

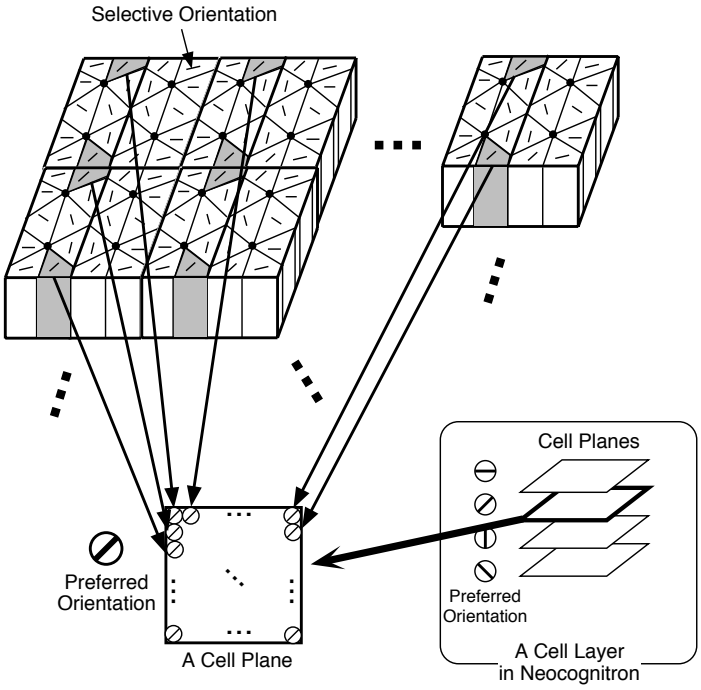

図 3.4: ネオコグニトロンの細胞面は, 実際の同じ特徴 (図では右斜め 45° の線分) を抽出するカラムを一枚のシート状に並べたようなものと考えることができる.

また, ネオコグニトロンの細胞間の結合は空間的に局所的な結合で構成されている. V1 野に対応すると考えている U_{S1} 層では非常に局所的な特徴しか抽出しないように設計しており, 高次 (l が大きい) になるにつれ, この受容野の大きさは徐々に大きくなるようにし大域的な特徴を抽出できるようにしている. 最終層における細胞の受容野は入力面全体をカバーできるような大きさになる.

図 3.3 を見るとわかるように S 細胞層から C 細胞層への計算を行う際に細胞面の大きさを縮小し, 細胞の密度を低くするように設定している. C 細胞は, S 細胞面の局所的な平均を計算する細胞であるので隣接する細胞は似たような出力を出す傾向にある. したがって, 細胞の密度を徐々に低くしていっても, パターンのもつ特徴の位置情報失われにくい.

ここで S 細胞 についてもう少し詳しく述べる. S 細胞 は, ある種の特徴抽出機であり, 第 l 段目の k 番目の細胞で, 細胞面内での位置が $\boldsymbol{n}$ であるようなものの数式表現は

以下のような式で与えられる:

$$u_{Sl}(\boldsymbol{n},k)=\frac{\theta_l}{1-\theta_l}\cdot\varphi\left[\frac{1+\sum_{\kappa}\sum_{\boldsymbol{\nu}}a_l(\boldsymbol{\nu},\kappa,k)\cdot\ u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu},\kappa)}{1+\theta_l\cdot b_l(k)\cdot u_{Vl}(\boldsymbol{n})}-1\right], \tag{3.1}$$

ただし $a(\boldsymbol{\nu},\kappa,k)$ は U_{Cl-1} 層の κ 番目の細胞面からの結合を表している. 関数 φ 内の分子部分は興奮性の入力を表しており, 分母部分が抑制性の入力をあらわしている. 関数　φ 自体は以下のような半波整流形の非線型関数であらわされる:

$$\varphi(x)=\max[x,0], \tag{3.2}$$

結合係数 a_l, と b_l とは学習によって決定される結合係数で, この結合が細胞のふるまいを決定する. 抑制性入力である u_{Vl} は以下のようにかける.

$$u_{Vl}(\boldsymbol{n})=\sqrt{\sum_{\kappa}\sum_{\boldsymbol{\nu}}c_l(\boldsymbol{\nu})\cdot\{u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu})\}^2}, \tag{3.3}$$

ただし, $c_l(\boldsymbol{\nu})$ は固定の結合定数あり, $||\boldsymbol{\nu}||$　に関して単調減少な結合である.

一方, 第 l 段目, k 番目の C 細胞面の位置 $\boldsymbol{n}$ における C 細胞 の発火は以下のように表される. :

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

ただし d_l は同じ l 段目の S 細胞面 からの結合で, $||\boldsymbol{\nu}||$ に関して単調で非負な結合荷重を持つ. 関数 ψ は以下であらわされる:

$$\psi\left[x\right]=\frac{\varphi\left[x\right]}{1+\varphi\left[x\right]}, \tag{3.5}$$

ただし φ は式 (3.2) で定義された半波整流型の関数である. 固定結合である c_l と d_l の具体的な形については 3.5 節で述べる.

式 (3.4) をみてわかる通り, ひとつの C 細胞面の中の C 細胞は同一の S 細胞面からの入力結合のみをもつ. 計算の際には結合 $d(\boldsymbol{\nu})$ で畳み込み和をとることによって受容野内の異なる位置に提示された特徴に対しても反応するような仕組みを持つものとして

いる. 言い換えれば, C 細胞は S 細胞面内の局所的な平均を取ることによって出力を計算している. このような実装によって C 細胞 は S 細胞のみでは許容できない位置ずれなどに対して反応できるようになる.

ネオコグニトロンにおいては S 細胞層と C 細胞層とが交互に並べられており, 特徴抽出と位置ずれの許容プロセスが交互に行われている. ネオコグニトロンは, このプロセスを繰り返し行っていくうちに局所的な特徴が徐々に大域的な特徴として統合され, 最終的にはパターン全体の特徴が抽出できるように設計される.

Sakai と Tanaka は生理実験の結果を支持するモデルとして "カスケードモデル" を支持している. カスケードモデルは, フィルタリング, 非線型関数による変調, そして同じ最適方位を持つさまざまなフィルタ出力の和をとる, という 3 つの計算プロセスからなるモデルである [88]. ネオコグニトロンにおける S 細胞の役割は フィルタリングと非線型関数によるフィルタリングであり, C 細胞の役割はさまざまな位置のフィルタ出力の和をとることであるので, ネオコグニトロンの各段, すなわち U_{Sl} 層と U_{Cl} 層の組み合わせは カスケードモデルの一種であると考えることができる (図 3.1 (c) 参照).

それゆえ, ネオコグニトロンの低次段 ($l = 1$ など) においては, C 細胞は V1 野の性質をよくあらわしてるものと考えられる. 残念ながら V1 以降の構造は V1 野ほど明らかにされていないが, 皮質の均質性などを考えれば, それほど急激に性質が変化しているとも考えにくいので, カスケードモデルのような処理をするプロセスががシーケンシャルに並んだものを腹側経路の主たる計算プロセスと考えてもそれほど不自然ではない. この仮定が成立するならばネオコグニトロンは腹側経路のモデルとして成立しうる.

3.3 Logothetis らの実験

IT 野の細胞に関しては, 顔などの複雑な特徴に反応する報告や [79], 図形を簡単にしていくことによって特定カラム内の細胞の最適刺激を推定する報告などがある [108]. モデルを用いた計算機実験で, これらの報告の現象を再現する場合, 画像中のオブジェクトがいずれも複雑な要素が絡み合ったものなので, オクルージョンや光源の影響によるシェーディングの問題などを解決しておかなければならない. また Perret や Tanaka らの実験は IT 野の細胞の特性を定性的に述べたものであり, IT 野の細胞のオブジェクト

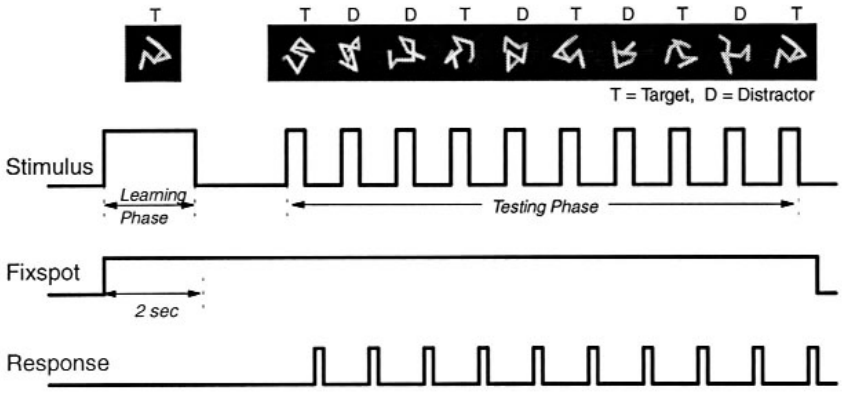

図 3.5: Logothetis らの実験. 弁別訓練したマカクザルに様々な角度からみたターゲットオブジェクト (T) と ディストラクタオブジェクト (D) をみせ IT 野での反応を測定する. ([60] より)

の回転や拡大・縮小などの変形に対する定量的な細胞の反応を調べたものではなかった.

これに対し, Logothetis ら は簡単な物体の画像を生成し, 視点の位置を変えた画像に対して細胞の定量的な反応を測定している. Logothetis らはマカクザルを用いて “クリップ状の物体” の画像イメージをコンピューターグラフィクスを用いて生成し, この画像に対する IT 野の細胞の反応を計測している [60], [61]. 彼らは, まずマカクザルをいくつかのクリップ状の物体 (ターゲットオブジェクトと呼ぶ) の画像イメージで条件付けを行い, マカクザルが教えた物体を識別できるようになるまで訓練を続けた. 訓練後, マカクザルは新たに生成された新奇なクリップ状の物体と, 訓練で用いた物体を区別できるようになった. この新たに生成した新奇な物体をディストラクタオブジェクトとよぶことにする. 実験は弁別すべきターゲットオブジェクトの図形を 2 秒間見せ, 以後, 数百 msec の間見せられた図形を次々と判別していくというものである (図 3.5 参照). このように非常に短い時間の間に提示された図形を判別させていくということは, モデルを考える際に視覚経路のフィードバック結合の影響をあまり考えなくてもいいということを意味している. この実験に際して Logothetis らは条件付けを行ったマカクザルの IT 野に微小電極を刺し込み, ターゲットオブジェクトの, ある特定の位置から見た画像イ

メージに特異的に反応する細胞を発見している． さらに, 彼らは, この細胞がディストラクタオブジェクトに対して, どのようなふるまいを示すのかということと, ターゲットオブジェクトを回転した時の画像イメージや, 拡大縮小させた物体のイメージに対してどのような反応を示すのかということを測定し報告している. 物体を回転させたり拡大縮小させたりすることは, 物体を見る位置を変えることにあたり, 画像としては著しく異なる場合がある. この実験結果によるとディストラクタオブジェクトに対しては IT 野の細胞は弱い反応もしくは, ほとんど反応しないという結果が得られている. また, ターゲットオブジェクトを回転させた画像を提示した場合は, その細胞がもっとも特異的に反応する視点の位置からの回転角が大きくなるほど反応が弱くなっていくことが観測されている. 拡大縮小に関しても同様の結果が示され, 最もよく反応するイメージからの拡大率や縮小率が大きくなるほど, 細胞の反応は徐々に小さくなっていくことが報告されている. また物体が提示される位置の変化に関しては, 細胞の反応は変化せず常に大きい反応を示しており, Ito らの結果と一致している [47].

Logothetis らは, 物体の回転, 拡大縮小に対する反応から, これらの変化に対して, どの程度の不変性があるかをチューニングカーブを描くことで示している. 彼らは, チューニングカーブをガウス関数でフィッティングしてやり, その際に出てくる分散で, 細胞の回転に対する不変性を定量化している. この分散の値が大きいということは回転などの変化に対して許容度が上がることを意味し, 小さいことは, その位置で見た画像に特異的であることをあらわしている. 彼らの測定結果では, 物体の回転に対する分散は $28.87 \pm 12.7°$ であることが判明している. このチューニングカーブとディストラクタオブジェクトの反応の大きさとを比較することによって, 測定している細胞のターゲットオブジェクトとディストラクタオブジェクトとを区別できる限界を知ることができる. これによるとおおよそ $40 \sim 50°$ 程度の回転であれば細胞はターゲットオブジェクトとして識別される.

また, 物体の拡大縮小についても同じようなカーブが描かれ, 拡大率 1 倍がもっとも大きな反応を示し拡大率がこれより大きくなっても小さくなっても反応が低下していくことが報告されている. これもディストラクタオブジェクトに対する反応と比較することによって区別できる限界を知ることができる． これによれば, 拡大率が $0.4 \sim 3.5$ 倍の間であればターゲットオブジェクトとして識別される.

物体の提示位置に関しては $\pm 7.5^\circ$ 位置をずらして提示しても反応は変化しないことが報告されている.

以上の結果から IT 野の細胞はある特定の画像イメージに対して, 特異的に反応し物体の回転, 拡大縮小などの 3 次元的な変化に対しても許容度を持つことがわかる.

3.4 従来の IT 野のモデルとネオコグニトロン

Bricolo らは, Logothetis の実験を説明するためのモデルとして RBF ネットワークを用いたモデルを提案した [12]. 図 3.6 は Bricolo らの提案したモデルの模式図を示す. このネットワークは 3 層構造からなり中間層はフィルタバンクから構成され様々な画像のフィルタをテンプレートとして貯えている. このモデルの問題点として

- 結合が局所的ではない
- 解剖学的な領野との対応がとれない
- 各層間でのプロセスの一様性がない

という点が挙げられる. 第 1 の問題点の根拠は, V1 野から V4 野にかけて徐々に弱くなっていくもののレチノトピーが観察されるところから, 結合は空間的に局所的でなければならないというところにある. Bricolo らのモデルでは 中間層のフィルタバンクにおいて既に入力の面全体からの入力を受け取っており, この層で既にレチノトピーが存在していない. Bricolo らは, 彼らのモデルの最終層の出力が IT 野の細胞と非常に良く似ていると主張しているが, その結果に至る各階層の解剖学的な階層構造とネットワークの構造については何の言及もしていない. したがって解剖学的に観察される V1 野 → V2 野 → V4 野 → IT 野という領野との対応関係がとれない. また Bricolo らのモデルにおける計算プロセスは, フィルタリングと RBF ネットワークの計算という 2 プロセスから成立している. これらの計算プロセスは著しく異なった計算を行なっていることになる. もし仮にこの計算プロセスが腹側経路上で行われているとすれば, おそらく, どこかの領野で構造が著しく変化していると考えた方が自然である. しかしながら Felleman と Essen が示した皮質の生理学的な観察によるとレチノトピーや受容野の大きさなどの

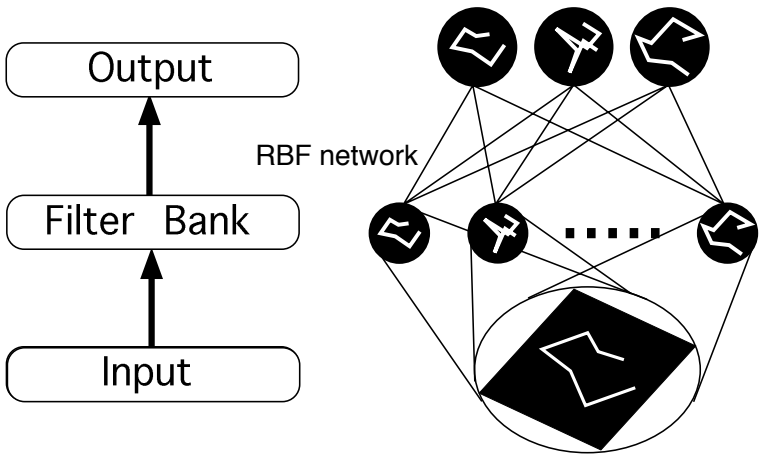

図 3.6: Bricolo らのモデル．フィルタを中間層とする RBF ネットによって形成されている．([12] より)

変化は徐々にしか起きていないことがわかる [19]．したがって Bricolo らのモデルは構造的に腹側経路のモデルとしては, 妥当でないということが言える．

Riesenhuber らは, Bricolo らのモデルを, 生理学的な構造と合致させるために図 3.7 のようなモデルに拡張した [82]. Riesenhuber らのモデルの拡張した点は, 階層を増やし入力層の次の “V1” という層では局所的なフィルタによる特徴抽出を行い, “V4” と呼んでいる層で ‘V1’ 層の最大値を出力するようにしていることである．この拡張により Riesenhuber らは, 腹側経路のモデル化を試みている．Riesenhuber らのモデルは生理学的なモデルとして受け入れやすいモデルとなっており, 結合の局所性と階層の対応付けはなされている．しかし 各階層の計算プロセスは, フィルタ, 最大値検出, RBF とそれぞれの計算プロセスが著しく異なっており, 皮質上での実現可能性を考えた場合, やはり問題が残ると考えられる．

これに対してネオコグニトロンは, 図 3.3 に示すように階層構造を持ち, 各階層の基本的な構造は一様である．各階層を構成する計算プロセスは, 特徴抽出 (S 細胞) と, 空間的なぼかし (C 細胞) という均質なプロセスである．この構成はカスケード モデルであり, Sakai と Tanaka によって V1 野の複雑型細胞のモデルとして, その妥当性が示され

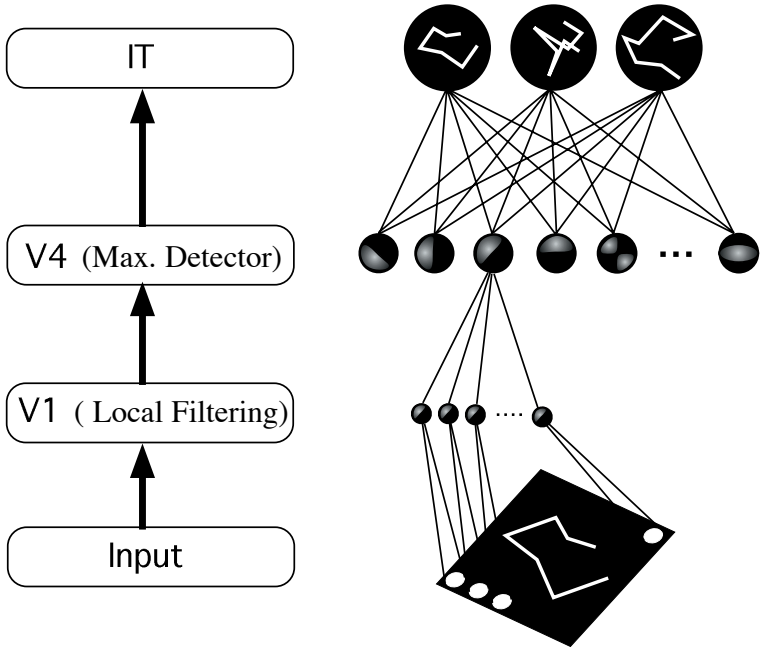

図 3.7: Riesenhuber らのモデル．フィルタを中間層とする RBF ネットによって形成されている．([82] より)

ている [88]. これにより Bricolo や Riesenhuber らのモデルよりも, 生理学的な構造という意味において妥当であると考えられる. さらに機能的な妥当性を検証するためにはネオコグニトロンによって Logothetis らの実験結果を説明することが出来ればよい.

3.5 ネオコグニトロンを用いた認識実験

この実験では, ネオコグニトロンを用いて Logothetis らの実験と同様にクリップ状の物体をネオコグニトロンに提示し学習させ, 学習した物体を回転させたもの, 拡大縮小したもの, さらに学習させてない物体を提示し, ネオコグニトロンに含まれる細胞の反応を調べた. 通常のパターン認識で用いられるネオコグニトロンはさまざまなパラメータ, (たとえば閾値 θ_l) などによって制御されている. Logothetis らの実験結果を, よりよく説明するように細かな調整を行うことも可能であるが, 本実験ではなるべく単純なパラメータを用いてシステムを構築している. 表 3.1 に本実験で用いたパラメータを示す. 各段においても, 同種のパラメータは, なるべく同じ値を持つように設定している.

U_{S1} 層は直線を検出する層で入力された画像に含まれる 8 方位の直線成分を検出するように作り付けている. この層は, V1 野の単純型細胞に対応する. V1 野の単純型細胞のもつ空間フィルタは Gabor 関数で近似でき, 直線成分やエッジを抽出することから, U_{Sl} 層を直線成分としている. したがって U_{S1} 層への入力結合である $a_1(\boldsymbol{\nu}, \kappa, k)$ は固定結合とし, ガウス関数で近似したフィルタを用いている.

これ以外の興奮性の可変結合である $a_l(\boldsymbol{\nu}, \kappa, k) \quad l > 1$ は, ある種の競合学習によって決定される [33]. 同一層内の細胞面で同じ受容野 (中心を $\boldsymbol{n}^*$ とする) をもつ細胞間で競合させ最も反応値が高い細胞を $u_{Sl}(\boldsymbol{n}^*, k^*)$ とする. この場合, 反応の最大値を持つものが勝者となるので以下の更新式によって結合が強化される.

$$a_l(\boldsymbol{\nu}, \kappa, k^*) = a_l(\boldsymbol{\nu}, \kappa, k^*) + q \cdot u_{Cl-1}(\boldsymbol{n}^* + \boldsymbol{\nu}, \kappa), \tag{3.6}$$

特徴抽出細胞の更新に関する更なる詳細は 4 章で述べる.

一方, 抑制性の可変結合である $b_l(k)$ は以下の式によって決定している.

$$b_l(k) = \sqrt{\sum_{\kappa} \sum_{\boldsymbol{\nu}} \frac{\{a_l(\boldsymbol{\nu}, \kappa, k)\}^2}{c_l(\boldsymbol{\nu})}}.$$

表 3.1: 本実験で用いたパラメータ

	0	1	2	3	4	5
S-cell-plane size	N.A.	41×41	41×41	21×21	11×11	5×5
C-cell-plane size	41×41	41×41	21×21	11×11	5×5	1×1
size of $a(\boldsymbol{\nu}), c_l(\boldsymbol{\nu})$	N.A.	7×7				
size of $d_l(\boldsymbol{\nu})$	N.A.	5×5				
$c_l(\boldsymbol{\nu})$	N.A.	$c_l(\boldsymbol{\nu}) = 1$ for all $\boldsymbol{\nu}, l$				
$d_l(\boldsymbol{\nu})$	N.A.	$d_l(\boldsymbol{\nu}) = 1$ for all $\boldsymbol{\nu}, l$				
θ_l	N.A.	$\theta_l = 0.53$ for all l				

これは, 福島が提唱したネオコグニトロンの学習則 [26] とは異なるが, このように結合を決定することにより S 細胞は, 自分がもっとも特異的に反応する特徴 $a_l(\boldsymbol{\nu}, \kappa, k)$ と入力特徴 $u_{Cl}(\kappa, \boldsymbol{n} + \boldsymbol{\nu})$ との距離を方向余弦で測るようになる [25].

図 3.3 からわかるように入力細胞面は 41×41 個の細胞からなり, 段の数 l があがるにつれ, 細胞面の大きさは前段の約 1/2 程度の大きさに減少していくように設計している. 細胞の受容野の大きさは, 段があがるにつれ大きくなるようにしている. 実際のマカクザルの V2 野と V4 野と IT 野における受容野の大きさは Tanaka と Kobatake や Rolls によって計測されており, 領野が進むにつれ大きくなることが報告されており, IT 野では, ほぼ視野全体をカバーするような受容野を持つことが知られている [52] [83]. 各領野の受容野の大きさを図で表すと 図 3.8 のようになる.

図 3.9 は, この実験に用いた入力画像のサンプルを示している. この画像を生成するために, まず 60 個のクリップ状の物体を計算機上で生成し, 第 0 番目の物体 (図 3.9 (a) を参照) を学習させる物体と定めた. 残りの 59 個の物体はディストラクタオブジェクトとして用いた. 図 3.9 (b) はディストラクタオブジェクトの一部を示している. 次に, これらの物体の画像をコンピュータグラフィクスを用いて生成した. 生成する際に学習させる物体に関しては物体を回転させた画像や, 拡大縮小させた画像, 物体の配置位置を変化させた画像を生成させた. このうち $-90°$ から $+90°$ までの間の $30°$ おきの回転

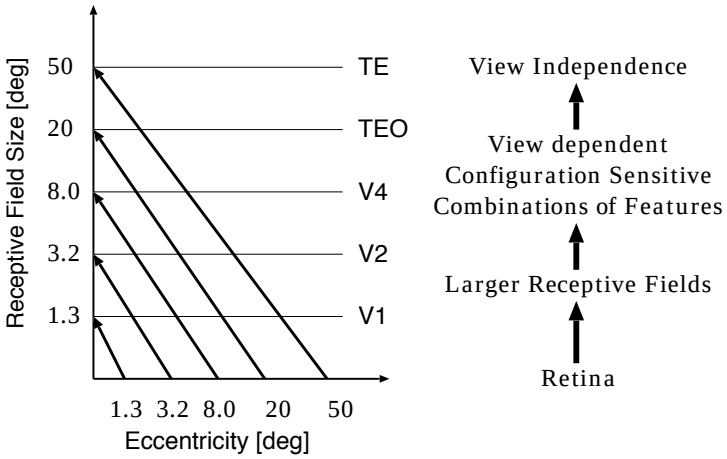

図 3.8: 腹側経路の受容野の大きさ．横軸は視野角度を表し, 縦軸が各領野を表す. ([83] を変更)

させた物体の画像と, 5 種類の ディストラクタオブジェクトを混在させたものを学習パターンとして定めた．この学習パターンを用いて, ネオコグニトロンの可変結合である a_l と b_l を決定した．a_l の学習は和家らのアルゴリズムに基づいた競合学習により決定している.

3.5.1 回転, 拡大縮小, 位置変化に対する U_{C5} 細胞の反応

図 3.10 は, 典型的な U_{C5} 細胞の回転した物体に関するチューニングカーブとディストラクタオブジェクトに対する反応例である．横軸は物体の回転を表し, 縦軸は特異的な画像に対する反応を 1 として正規化した反応の値である．この細胞は $-60.0°$ の画像に対し特異的に反応し, その前後 $40.0°$ の区間で反応を示す．ディストラクタオブジェクトに対する反応は棒グラフで表示されるようにほとんどのディストラクタオブジェクトには反応しない．また反応があっても, その値は非常に小さいものとなっていることがわかる．図 3.11 は, いくつかの細胞の反応を表示したものである．いずれの細胞のチューニングカーブも特異的な画像に対し最も大きい出力を出し, 回転角がその特異的な画像のもつ回転角から離れるほど反応は小さくなっている．また, 直線で結ばれていない反応はディストラクタオブジェクトに対する反応であり, いずれも反応の値は非常

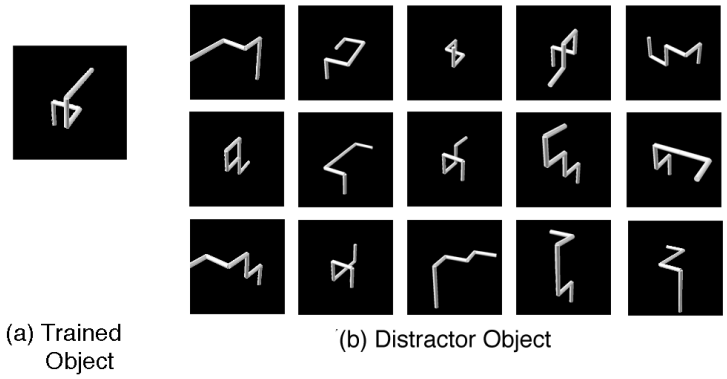

図 3.9: 本実験に用いた刺激画像の例. (a): ターゲットオブジェクトをある角度からみたときの画像. (b): ランダムに生成させたディストラクタオブジェクトの画像例

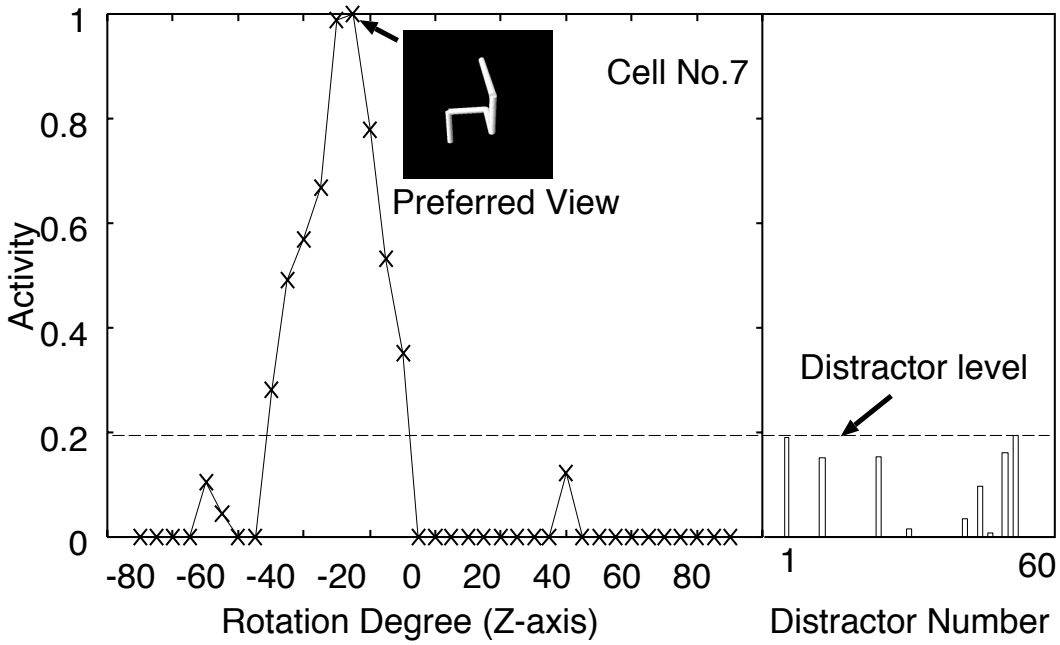

図 3.10: 最終層である U_{C5} 層の細胞の, ターゲットオブジェクトの回転画像に対する, ある細胞の反応. 横軸は, z 軸に対する回転角と, ディストラクタの番号を表している. 縦軸は, ネオコグニトロン最終層でみられる細胞のアクテビティーを表す. 各線は別々の細胞のターゲットオブジェクトを回転させたときに得られるチューニングカーブを表す. このセルは $-30.0°$ 付近の画像に対しもっとも選択性が高い.

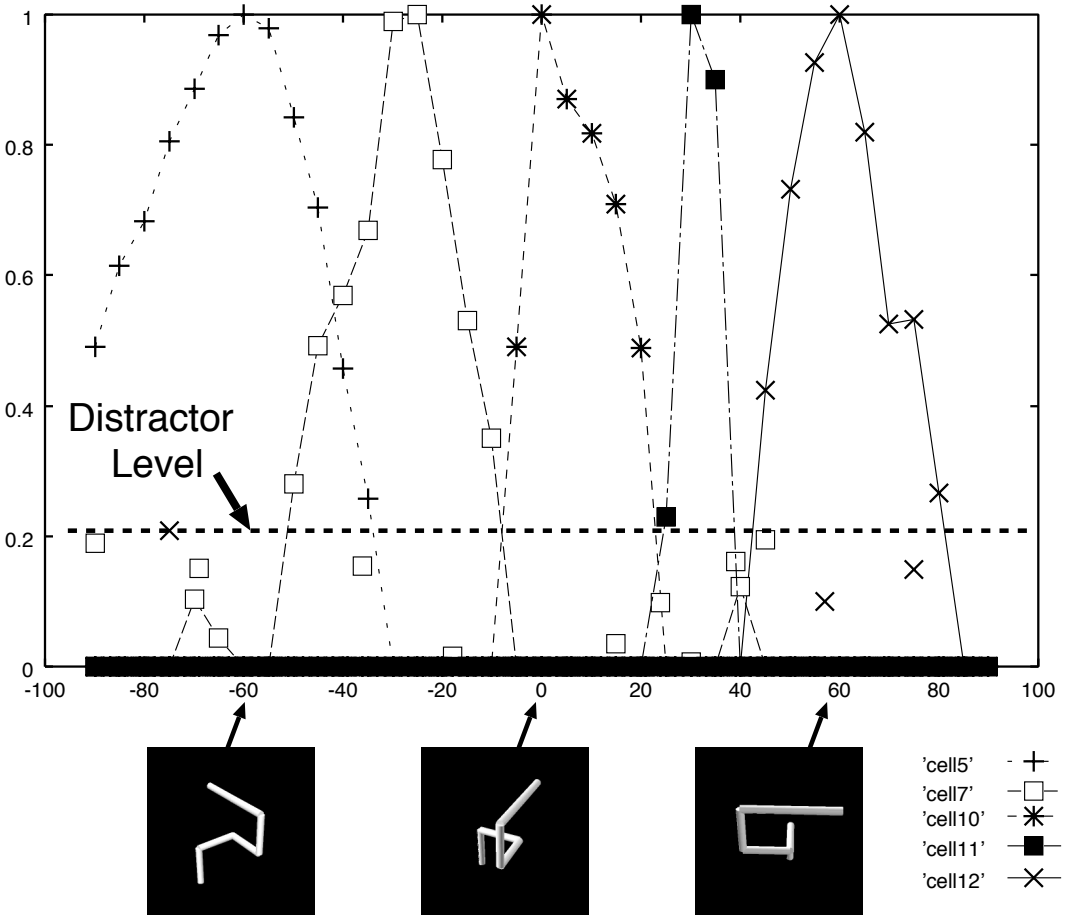

図 3.11: U_{C5} 層の幾つかの細胞の回転に対する反応例．横軸は, z 軸に対する回転角と, ディストラクタの番号を表している．各線は別々の細胞のターゲットオブジェクトを回転させたときに得られるチューニングカーブである．棒グラフはディストラクタオブジェクトに対するアクティビティをあらわしている．ほとんど全ての細胞はターゲットオブジェクトを特定の角度からみた画像に選択的な応答を示し, ディストラクタオブジェクトにはほとんど反応を示さない.

に小さいものとなっている. 回転によって, 画像のもつトポロジーが変化しない場合, 細胞は回転画像に対してチューニングカーブの幅は広いがが, 物体が自己遮蔽を起こすような 30.0° 付近の例では, チューニングカーブの幅は非常に狭くなることもある.

図 3.12 は, 物体の大きさを変化させた場合の細胞が示す反応である. 縦軸は細胞の反応を示すもので, 回転画像に対するグラフと同様に, 特異的な画像に対する反応を 1.0 とした正規化を行っている. 横軸は拡大率を表している. ディストラクタオブジェクトに対する反応のレベルと比較して, この細胞はおおよそ 0.6 倍から 1.4 倍の間, 有効な反応を示すものと思われる. 拡大率が 1.5 倍を超えてしまうと物体自体が入力画像に収まらず, はみ出してしまうので反応は非常に小さくなってしまうが, これは抽出すべき特徴が存在しないので, 画像内に含まれないために起こってしまうものではないかと思われる.

図 3.13 は, 物体の提示位置変化による細胞の出力をあらわしている. ネオコグニトロンは, 結合が並進対称性をもつので, C 細胞においては物体の提示位置が変化しても, 細胞の出力が, それほど変化しないような設計になっているが, 段数が上がるにつれ, 細胞面のサイズが縮小され, 特徴の情報が失われる. 一部の提示パターンについては出力が若干落ちているが, これは, そのことが原因であると考えられる. しかしながらディストラクタオブジェクトに比べれば十分区別できるレベルであるので問題ないと考えられる.

この実験においてネオコグニトロンは, Logothetis らの実験をよく近似するものと考えられる. また構造上の観点から言っても腹側経路のモデルとしての可能性を十分示唆しているといえる.

3.5.2 受容野の結合空間分布の推定

マカクザルなどの V2 野以降の受容野の空間分布は形状が複雑で, 現在でも調査するのが困難である. ここでネオコグニトロンが腹側経路のモデルであると仮定すると, V2 野以降の結合の形状が, ある程度推定できるようになった. これは U_{S2} 層から U_{S5} 層までの細胞のもつ受容野の空間分布というものを推定することによって行った. これは, Fukushima らが提案した “選択的注意機構のモデル” を応用して [24], [46], [95], [97], ある段の S 細胞から信号を逆行方向に流してやり, U_{C0} 層に再生されたパターンを見るこ

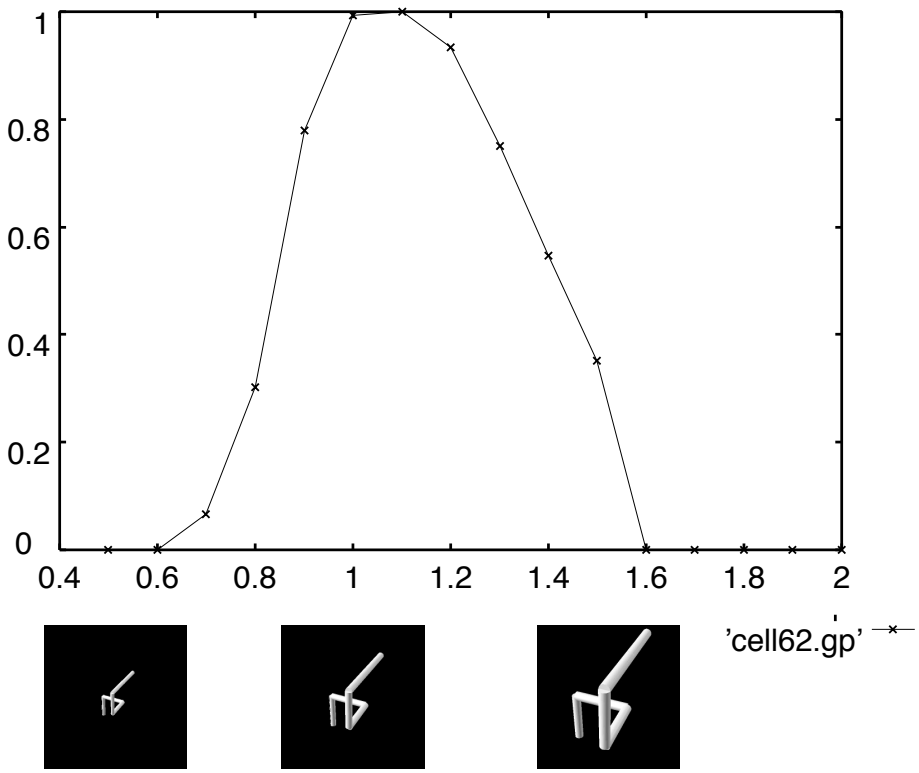

図 3.12: 拡大・縮小に関する反応例. 横軸は画像の拡大率を表す. この細胞は, およそ 0.6 倍から 1.6 倍までの拡大・縮小画像に対して反応することが分かる.

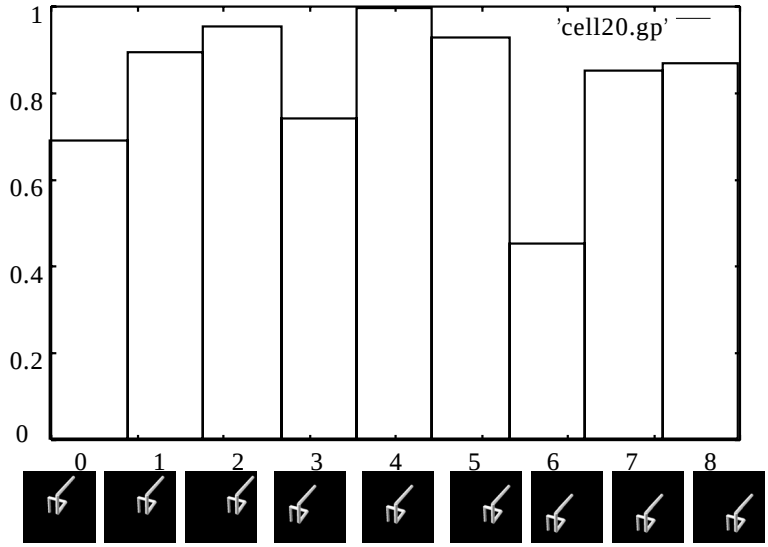

図 3.13: 物体の平行移動に対する細胞の反応. ネオコグニトロンは結合が並進対称性を持つので, 画像の平行移動に関しては, ロバストな識別が可能である. 一部に反応が低下しているのは, 画像を量子化する際の誤差などによるものと考えられる.

とよってに行っている. ただし U_{Sl} 層から U_{Cl} 細胞への結合 D_l は信号を拡散させる働きを持つので U_{Cl} 層から U_{Sl} 層へ逆伝播させる際にデコンボリューションしてやらないと信号が空間的にぼけてしまう. 本推定方法は U_{Sl} 層においてラプラシアン型の核関数と畳み込むことによって信号のシャープニングを行っている.

図 3.14 は, この逆再生の結果を示している. 第 1 行目の列は U_{S1} 層に含まれる S 細胞の空間分布を示している. この層は前述の通り結合が前もって決定されている層であり, 結果として得られた結合分布は, 直線を抽出するような結合が再生されているので, この方法による推定が妥当であることを示している.

第 2 行目の列は U_{S2} 層に含まれる S 細胞の空間分布を表す. これらの細胞は 1 行目の U_{S1} 層の細胞よりも複雑な特徴をコードしていることがわかる. 第 3 行目の列は U_{S3} 層の S 細胞の空間分布を表しており, 物体の角の部分などを中心に見たときの画像らしいものがコードされているのがわかる.

もし, 腹側経路がネオコグニトロンでよく近似できるのであれば, V2 野以降の細胞がもつ受容野の空間分布は, このような形になることが期待される.

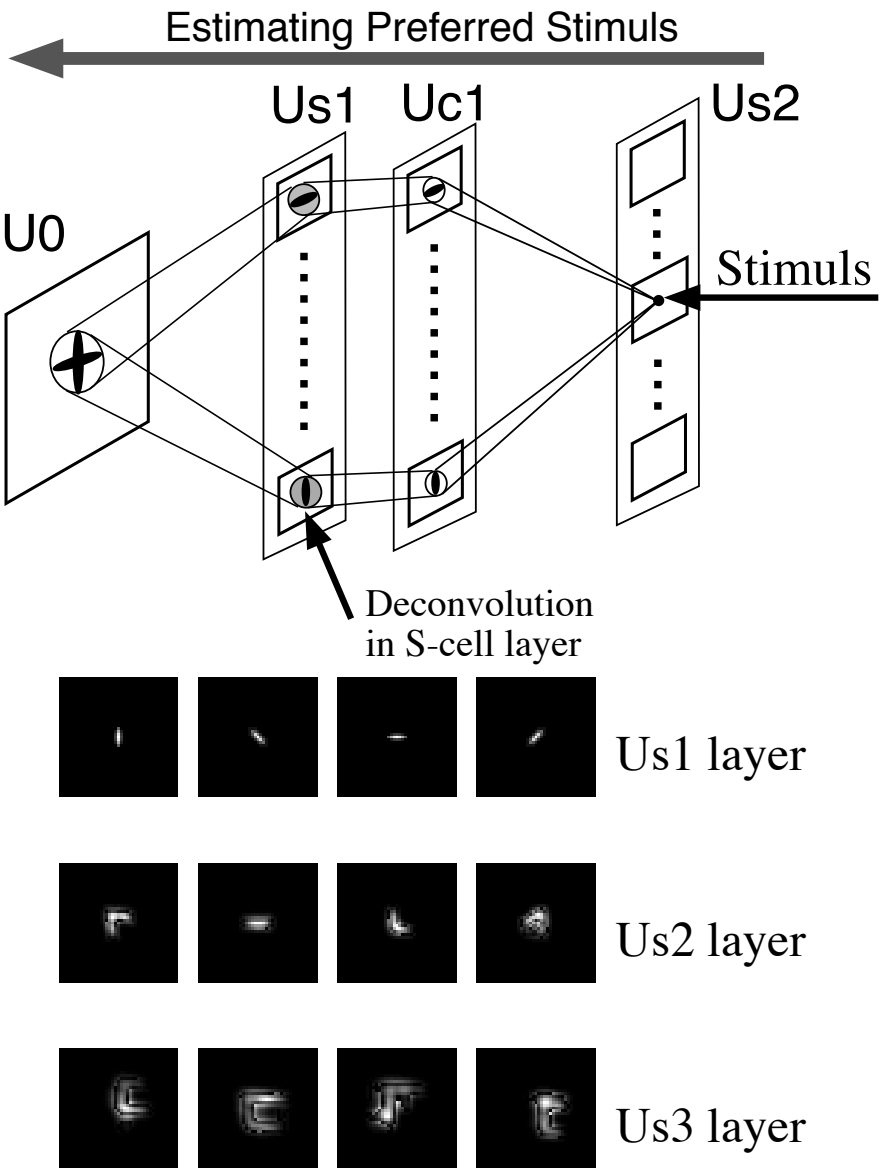

図 3.14: 空間フィルタの推定．S 細胞に刺激を与え信号を通常とは逆方向に伝播させることによって U_{C0} 面にフィルタを推定する．推定されたネオコグニトロンの空間フィルタの例．式 (3.1) の a_l がフィルタにあたり, U_{S1} 層から U_{S3} 層までのフィルタを表している．上段は, U_{S1} 層の空間フィルタを表しており, このアルゴリズムが正しく動作していることを示す．U_{S2} 層と U_{S3} 層のフィルタは, 物体の角の部分などに対応するものが多い．

3.6 検討

我々はネオコグニトロンを用いて腹側経路における IT 野の細胞のモデルを提案した. このモデルは, V1 野の単純型細胞と複雑型細胞のモデル (S 細胞と C 細胞) を拡張したモデルであり, S-Cell と C-Cell の組み合わせは, Sakai と Tanaka が示した生理学的知見を支持している. ネオコグニトロンを用いてクリップ状の 3 次元物体の画像の認識シミュレーションを行った結果, Logothetis らの示す IT 野の細胞の特性をよく表していることが判明した. これらの結果は, ネオコグニトロンの腹側経路のモデルとしての可能性を示唆しているものと考えられる.

Poggio らは RBF ネットワークを用いて IT 野の細胞のふるまいを再現しようとしている [12][82]. 特に Riesenhuber が提案しているモデルは, 第 1 層は直線やエッジなどの低次特徴の抽出を行う層, 第 2 段は抽出した特徴がパターンに含まれるかどうかを判断する層としており, それぞれ V1 野と V4 野のモデルとしている. これは, ネオコグニトロンの S 細胞層と C 細胞層との関係に類似している. 相違点は, ネオコグニトロンが徐々にパターン変換を行っていくのに対し, Riesenhuber のモデルは第 2 層で, 特徴情報が一気に 1 個の素子に統合される点である. 第 2 層の素子は対応する特徴がパターン中に含まれるかどうかを判断している. これは第 2 層で特徴の位置情報が失われるため, パターンを適当な大きさに切ってシャフリングした場合にネットワークが反応する可能性がある.

マカクザルの腹側経路における各領野の面積は V1 野, V2 野あたりが最大の面積を持ち, V4 野以降は面積が徐々に小さくなっていくことが知られている [19]. ネオコグニトロンの各段における細胞の個数も, これと同様に段が上がるにつれて少なくなるようにしている. 各段における細胞の個数は, 細胞面の個数にも依存するが, 一般的にいえば細胞の密度を低くすることによって全体的な個数を減らすことができると考えられる. これは計算機のメモリサイズや計算時間の短縮などの要請によるものだが, 画像データからシンボル化を行うというある種の情報の圧縮を行っている観点からみても興味深い.

今回, 用いた S 細胞の学習は, 細胞面の一個の細胞が学習すると同じ特徴に対して選択性を持つ細胞が強化されるという方式であるが, これは Földiak の考えるような Trace Hebb 則のようなものを考えれば, 実現可能であると考えられる [20] [119]. また S 細胞

から C 細胞への結合は学習によらない固定結合であるが, Fukushima らが提案するような, C 細胞の反応に時間遅れを持たせたような系を考えることによって, この結合も学習によっても実現できると考えられる. これらの方法で結合加重を決定することはたいへん興味深いことであり, これは将来の課題である.

最近の 3 次元物体の認識モデルとしては, 前述の Riesenhuber や Brilcoro 以外にも, 藤田と安藤らによる双方向モジュールを用いたシステムで物体の再構成を行う計算論的なモデルなどが挙げられる [21]. ネオコグニトロンは, ボトムアップ (フィードフォワード) の信号のみによって物体の認識を実現するので, 繰り返し演算を要さない. したがって非常に短時間のうちに解が得られるが, これらのモデルは多量の繰り返し演算の結果として物体の形状を再構成する場合が多く, サルや人間などが非常に短い時間の間に物体の形状を認識するという状況に対応するモデルとしては, 計算時間の観点から見てやや不適合ではなかろうか？　これらのモデルは, むしろトップダウン (フィードバック) の信号を生かして, 自分が認識している対象の確認や, 注意の機構におけるスポットライト効果を実現するようなモデルとして捕らえたほうがよいと思われる.

第 4 章

ネオコグニトロンによる文字パターン認識

4.1 研究背景

Rosenblatt によって パーセプトロンが提案されて以来, ニューラルネットワークを用いた様々なパターン認識システムが開発されてきた [85]. また, 認識システムの発達にともなって, データセットに対する様々な学習方法が提案され, その長所と短所が議論されてきた. なかでも Amari や Rumelhart, McClelland らによって導かれた学習則である誤差逆伝搬法 (error back propargation (BP)) は, 最適なシステムパラメータセットに到達できない可能性がある, という欠点を内包しながらも今日のニューラルネットの発展に大きな役割を果たしている. ニューラルネットに BP 法を組み合わせて文字パターン認識を行なう研究は, 1990 年代に入ってから盛んに行なわれてきているが, なかでも LeCun らのグループによる研究が有名である [55], [105]. LeCun らは, パーセプトロンと ネオコグニトロンの中間のようなネットワーク (Le-net1 と呼ばれる) に, 学習則として BP を用いたシステムを用いている LeCun らは U.S.Postal が, 作成した郵便番号の文字パターンデータベースを用いてパターン認識実験を行なっており, 新奇パターンに対して 95 % 以上の認識率が得られることを報告している.

一方, ネオコグニトロンは, 生物の視覚系のモデルとして考案されたモデルである. 生物の視覚系は, 単純な特徴から次第に複雑なものへと階層性を持って情報処理がなされていくものと考えられている. ネオコグニトロンは, このような階層性を持ち, 入力パターンに含まれる特徴情報を階層が上がるとともに統合していく視覚情報処理のモデル

である [23], [28]. ネオコグニトロンを用いたパターン認識の計算機実験は, 従来より行なわれていたが [33], [31], 計算機の処理能力等の問題でカメラなどから取り込んだ実画像や, 膨大な記憶容量を必要とするデータベースを用いて実験を行なうことが困難であった. しかし, 近年の計算機技術の発達に伴い, これらの制約条件をクリアすること自体は, それほど困難なことではなくなりつつある. そこで我々は, ネオコグニトロンに CCD カメラで取り込んだ実世界の数字パターンのデータベースによる認識を行なわせ, その能力の評価を行なった. データベースには, 通産省電子技術総合研究所の文字データベースである ETL-1 を用いている. 学習には教師なし競合学習を用い ETL-1 に含まれるパターンのうち, 各カテゴリについて 200 パターン (200×10) を学習パターンとして用いて, 学習パターンに含まれない未知パターン (各カテゴリについて 300 パターン) についての認識率の測定を行なったところ, 97.5% の認識率を得ることができた [30][102]. 一般に, 文字認識はパターンの形状のみによる分類ができない. 形が全く違っても, それらのパターンを同じカテゴリに属するとみなす必要がある場合があるからである. ネオコグニトロンは本質的には, パターンを形状によって分類する神経回路モデルであり, 従来のシステムではカテゴリの分類を行なう際に, 教師あり学習 [26] による適当なラベルづけを行なうことによってパターンを分類していた. この方式は一度ラベルづけを行なうと, そのラベルを忘却することが出来ない. すなわち, 学習パターンと教師信号の組合せが適切に選ばれていれば問題なく機能するが, 最初にラベルづけを行なう教師信号にノイズや誤りが含まれている場合には, 誤った認識結果を導く可能性がある. そこで, 我々は, 更に認識率を上げるため, このラベルづけを行なう部分の改良を行なった. ネオコグニトロンを入力パターンの形のみによって分類するネットワークとして位置付けて, 形によって分類される出力情報と, 文字カテゴリとを結び付ける単純な一層の回路を新たに付け加えた. この回路を文字カテゴリ分類回路と呼ぶことにする [103]. このネオコグニトロンと文字カテゴリ分類回路とを組み合わせたネットワークシステムは, Hecht-Nielsen らが提案している forward-only counter propargation network (fCPN) と同様の構造を持つことになる [38]. fCPN は 3 層の構造を持つネットワークで, 中間層で winner take all (WTA) 型の学習を行ない, 出力層で, 誤り訂正学習を行なう神経回路モデルである. Hecht-Nielsen は中間層には Kohonen の SOM [53] を用いて, 出力層には, Hebb 則による誤り訂正学習を行なっている.

我々は, 中間層の役割に当たる回路としてネオコグニトロンを用い, 出力層として, 単純 パーセプトロン のような全結合型の回路を導入している. 出力層の学習則には,

- 誤り訂正学習
- LVQ2
- コスト関数を導入した学習法

のそれぞれを用いて評価を行ない, この新たなカテゴリ分類回路を導入することによってシステムの認識率は, 未学習パターンに対して 98.1% という値が得られた [98].

4.2 ネオコグニトロン

実験に用いたネオコグニトロンは 図 4.1 のような階層構造を持つ多層の神経回路モデルである. 入力層から順に S 細胞と呼ばれる細胞からなる細胞の層 U_S 層と C 細胞と呼ばれる細胞からなる U_C 細胞層が交互に並んでいる. このような 2 種の細胞層 U_S 層と U_C 層との組を "細胞段" と呼ぶ. ネオコグニトロンは細胞段を入力層の後ろに何段もつないだ構造になっている. 本章では, 第 l 段めの S 細胞の層 を U_{Sl}, C 細胞の層を U_{Cl} と表記している.

ネオコグニトロンにおける S 細胞は入力パターンのさまざまな特徴を抽出する. C 細胞は, 前段の U_S 層の反応パターンに空間的な 'ぼかし' を施す. このように特徴抽出と, ぼかし操作を交互に行ない, 特徴相互間の空間的な位置ずれや変形を徐々に許容していくことによって, ネオコグニトロンはパターンの変形などに強く, 汎化能力に富んだパターン認識を実現している. S 細胞は, Hubel と Wiesel が提唱した 単純型細胞のモデルであり, C 細胞は複雑型細胞のモデルである [40][41].

U_{Sl} 層は, 前層の U_{Cl-1} 層から入力を受けており, その入力結合は可塑的であり, 後に述べられる競合学習の一種によって決定される. したがってこの結合が, S 細胞の振舞いを決定する. 学習後, U_S 層内の一個の S 細胞は, 局所的な特徴の一つに選択的に反応を示すようになり, パターンに含まれる特徴を抽出するように形成される. ネオコグニトロンでは, 同じ特徴を抽出する S 細胞を 2 次元平面状に並ぶように形成しており, 特

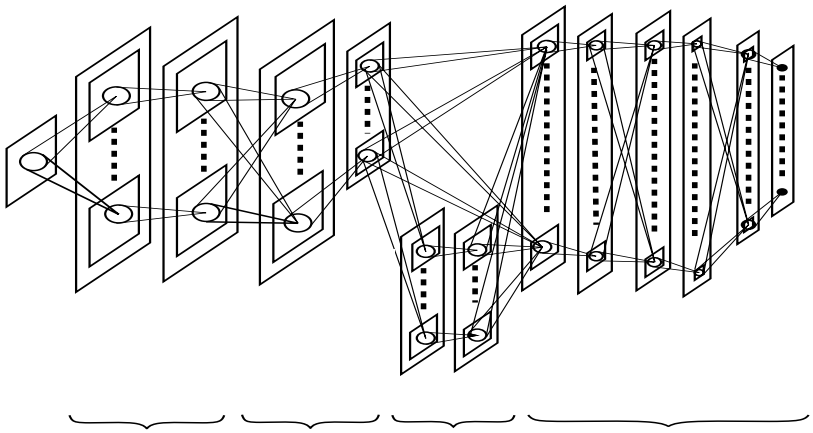

図 4.1: 本実験で用いたネオコグニトロンの構造

徴がどの位置に入力されても, その特徴を抽出することができるようにしている. この平面状に並んだ細胞の集合を “細胞面” と呼ぶ.

U_{Cl} 層は, 前層の U_{Sl} 層から入力を受けており, その結合は固定結合としている. U_C 層内の 1 個の C 細胞は, 同一細胞面内に含まれる複数の S 細胞からの入力を受け, 空間的なぼかし操作を施す. このように S 細胞の出力を集めることによって, C 細胞は, 自分が特異的に反応する特徴が受容野内のどこに提示されても反応するように実装されている.

回路内の細胞間の結合は 図 4.1 のように階層的に行なわれているので上位の層になるほど受容野は大きくなる. ネオコグニトロンの最終出力である C 細胞の受容野は入力層全体を覆う大きさになっている. 各細胞は受容野が大きくなるにつれてパターンの位置ずれの影響を受けにくくなる. そこで, 各層の細胞密度は上位にいくにしたがって減少させ, 最終出力の C 細胞面内の細胞は 1 個になっている.

4.2.1 S 細胞の反応

第 l 段目にある k 番目の細胞面に含まれる S 細胞の反応を考えてみる. 細胞面内の位置 $\boldsymbol{n}$ の位置にある S 細胞は

$$u_{Sl}(\boldsymbol{n},k)=\frac{\theta_l}{1-\theta_l}\cdot\varphi\left[\frac{1+\sum_{\kappa}\sum_{\boldsymbol{\nu}}a_l(\boldsymbol{\nu},\kappa,k)\cdot\ u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu},\kappa)}{1+\theta_l\cdot b_l(k)\cdot u_{Vl}(\boldsymbol{n})}-1\right], \tag{4.1}$$

という出力を出す. ただし $a(\boldsymbol{\nu},\kappa,k)$ は前段の C 細胞層 U_{Cl-1} 層の κ 番目の細胞面からの結合を表す. $theta_l$ は S 細胞の閾値に関連するパラメータである. 関数 φ の分子は興奮性の入力を表し, 分母は抑制性の入力を表す. 関数 φ は

$$\varphi(x)=\max[x,0], \tag{4.2}$$

で表される関数であり, これは半波整流を行なう関数として知られている. したがって, S 細胞の出力は, 非負のアナログ値をとることになる. 結合 a_l と b_l とは, 可塑性な結合で, これらの結合が, S 細胞と, この S 細胞を含む細胞面の振舞いを決定する.

抑制性入力の中に含まれる u_{Vl} は

$$u_{Vl}(\boldsymbol{n})=\sqrt{\sum_{\kappa}\sum_{\boldsymbol{\nu}}c_l(\boldsymbol{\nu})\cdot\{u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu})\}^2}, \tag{4.3}$$

で与えられる. ただし, $c_l(\boldsymbol{\nu})$ の結合係数は固定したものを用いている. この結合は $||\boldsymbol{\nu}||$ にしたがって単調に減少するものを用いる.

4.2.2 C 細胞層の反応

第 l 段目, k 番目の細胞面に含まれる C 細胞 (位置 $\boldsymbol{n}$) の出力は, 以下のようにして計算される:

$$u_{Cl}(\boldsymbol{n},k)=\psi\left[\sum_{\boldsymbol{\nu}}d_l(\boldsymbol{\nu})\cdot u_{Sl}(\boldsymbol{n}+\boldsymbol{\nu},k)\right], \tag{4.4}$$

ただし, d_l は, 同一段の, ある S 細胞面からの結合である. この結合係数の成分は非負で, $||\boldsymbol{\nu}||$ にしたがって単調減少するものとしている. 関数 ψ は飽和特性をもつ関数で:

$$\psi[x] = \frac{\varphi[x]}{1+\varphi[x]}, \tag{4.5}$$

としている. φ は, 式 (4.2) で定められた半波整流型の関数である. S 細胞と異なり C 細胞は同一段のただ 1 種類の細胞面からしか入力を受けとらない. しかも結合を受けとる範囲も結合係数 d_l によって定められたある, 決まった範囲からである. 言い替えれば C 細胞はパターンに含まれる特徴の移動平均をとることによって特徴の存在する位置をぼかし, パターンの位置ずれを許容するようにしている.

ネオコグニトロンの高位の段に含まれる C 細胞層, すなわち U_{C3}, U_{C4}, U_{C5} では, 細胞面に含まれる細胞の密度を減らしている. このように細胞の密度を下げることによって, 抽出した特徴情報を圧縮している. これにより使用する計算機資源を押えることが出来る.

4.2.3 低位段の特徴抽出機構

この節では, 本実験で用いたネオコグニトロンの低位段である入力 U_0 から, 折れ点検出回路 $U_{S2'}$ 層, および $U_{C2'}$ 層までの回路を説明する. この低位段に含まれる S 細胞層, すなわち U_{S1} 層, U_{S2} 層, および $U_{S2'}$ 層に含まれる S 細胞の持つの結合は, 作りつけで決定している.

まず入力層である U_0 層であるが, これは, 他の細胞面と同じく入力細胞を 2 次元平面上に並べたものである. ETL-1 データベースに含まれる入力は 8-bit のグレーレベルを持つパターンであり, バックグラウンド レベルは 0 に近い値を持つ. この値を一旦適当な閾値で処理した後, $[0,1]$ に正規化した値を用いる.

次の層である U_{S1} 層はエッジ抽出を行なう層であり, 入力パターンを 16 方位のエッジ成分に分解する.

U_{S2} 層は直線成分を検出する層であり, これは抽出されたエッジから検出している. 一つの直線成分に対しては, 二つのエッジが対応し, それぞれのエッジの方位は, ちょうど 180° だけ異なっている [33].

$U_{S2'}$ 層は, 折れ点を検出する層であり, U_{S2} 層と U_{S3} 層の間の層としてとらえることが出来る. ここでいう折れ点は, パターンに含まれる曲率が著しく変化する部分のことであり, 端点などもこれに含まれる. 折れ点はパターンに含まれる直線成分の端点として考えられるので, U_{S2} 層の出力のうち, その特異的な方位に沿った成分の端点を抽出する回路を取り付けている. U_{S3} 層は, 直線成分をぼかした U_{C2} 層と, 折れ点成分であるバイパスされた $U_{C2'}$ 層からの入力を受ける.

U_{S3} 層以降は, より複雑な特徴を抽出する回路で, これらの層の結合係数は 4.3 節に述べる学習によって決定される.

4.3 S 細胞の結合の強化

U_{S3} から U_{S5} までの細胞の入力結合は, ある種の競合学習によって強化される. これは, Fukushima と Wake らによって提案された手法で, winner-take-all 学習の一種である [33]. 結合強化のプロセス終了後, S 細胞はある特定の特徴に対してのみ反応を示すようになる. 結合強化のプロセスは, 低位の段から, 高位の段へ向かって行われる. ここでは, ある段の結合強化は, その前の段の S 細胞の結合強化が終了してから行われるようにしている. この結合強化の方法は, 従来のネオコグニトロンの教師なし学習と呼ばれるものとほぼ一致する [33]. 以下この結合強化法の詳細について述べる.

結合の強化アルゴリズム

Require: $U_{S1} \sim U_{Sl-1}$ の S 細胞の結合強化が終了していること.

1: 入力パターンを提示する
2: U_{Cl-1} 層までの計算を行う
3: $K_{Cl-1} \leftarrow U_{Cl-1}$ 層の細胞の数
4: **if** S 細胞面が U_{Sl} に存在する場合 **then**
5: $K_{Sl} \leftarrow U_{Sl}$ 層の細胞の数
6: U_{Sl} 層に含まれる細胞の反応を現状の a_l と b_l とで計算する.

7: **while** $\{\boldsymbol{n}, k \mid u_{Sl}(\boldsymbol{n}, k) > 0\}$ が存在する場合 **do**

8: 以下の条件を満たすような $\boldsymbol{n}^*$, と k^* とを探す.

$$u_{Sl}(\boldsymbol{n}^*, k^*) \geq u_{Sl}(\boldsymbol{n}, k) \quad \text{for}\, \forall(\boldsymbol{n}, k).$$

9: 次式にしたがって結合 a_l と b_l とを強化する

$$a_l(\boldsymbol{\nu}, \kappa, k^*) = a_l(\boldsymbol{\nu}, \kappa, k^*) + q \cdot u_{Cl-1}(\boldsymbol{n}^* + \boldsymbol{\nu}, \kappa), \tag{4.6}$$

ただし κ は $1 \leq \kappa \leq K_{Cl-1}$ を満たす任意の κ とする. また b_l は

$$b_l(k^*) = b_l(k^*) + q \cdot \sqrt{\sum_{\kappa} \sum_{\boldsymbol{\nu}} c_l(\boldsymbol{\nu}) \cdot \{u_{Cl-1}(\boldsymbol{n}^* + \boldsymbol{\nu}, \kappa)\}^2}. \tag{4.7}$$

にしたがって強化される.

10: 強化した位置は勝者となるので他の細胞面の反応を抑制する.

$$u_{Sl}(\boldsymbol{n}, k) = u_{Sl}(\boldsymbol{n}, k) - u_{Sl}(\boldsymbol{n}, k^*)$$

ただし k は, $1 \leq k \leq K_{Sl}$ および $k \neq k^*$ を満たすものとする. .

11: Winner となった細胞を含む k^* 番目の細胞面の反応を抑制する.

12: **end while**

13: U_{Sl} 層内の細胞の反応を計算する.

14: **end if**

15: “シードセル生成面 (seed-selecting plane)” u_{seed} を作成する (図 4.2 参照). これは, U_{Cl-1} 層の反応をぼかして加算を行うことによって作成する.

$$u_{\text{seed}}(\boldsymbol{n}) = \sum_{\kappa} \sum_{\boldsymbol{\nu}} c_{\text{seed}}(\boldsymbol{\nu}) u_{Cl-1}(\boldsymbol{n} + \boldsymbol{\nu}, \kappa),$$

ただし $c_{\text{seed}}(\boldsymbol{\nu})$ はガウス関数のような関数とする.

16: 既に U_{Sl} 層に発火している細胞 ($u_{Sl}(\boldsymbol{n}, k) > 0$) がある場合 $u_{\text{seed}}(\boldsymbol{n})$ の周りに抑制をかける. すなわち

$$u_{\text{seed}}(\boldsymbol{n}) = u_{\text{seed}}(\boldsymbol{n}) - \sum_{\boldsymbol{\nu}} e_{\text{seed}}(\boldsymbol{\nu}) u_{Sl}(\boldsymbol{n} + \boldsymbol{\nu}, k)$$

という操作を行う. ただし係数 $e_{\text{seed}}(\boldsymbol{\nu})$ はガウス関数のような結合係数である

17: **while** $\{\boldsymbol{n} \mid u_{\mathrm{seed}}(\boldsymbol{n}) > 0\}$ が存在する場合 **do**
18: シードセル生成面中で最大の反応を示す位置 $\boldsymbol{n}^{**}$ を探す:

$$u_{\mathrm{seed}}(\boldsymbol{n}^{**}) \geq u_{\mathrm{seed}}(\boldsymbol{n}) \text{ for } \forall \boldsymbol{n}$$

.
19: $K_{Sl} \leftarrow K_{Sl} + 1$.
20: 新しい細胞面を U_{Sl} 細胞層内に生成する. さらに新たな結合 a_l^{new} および b_l^{new} を用意する
21: 結合 a_l^{new} と b_l^{new}, を式(4.6) と 式(4.7) にしたがって強化を行う. ただし, $\boldsymbol{n}^* = \boldsymbol{n}^{**}$ とする.
22: シードセル生成面内の $\boldsymbol{n}^{**}$ 付近の反応を抑制する.
23: **end while**

結合強化のプロセスにおいては, 学習パターンに含まれる, どの部分にある特徴を学習するかを決定する必要がある. 抽出すべき特徴の位置を決定するために Fukushima と Wake らは U_{Cl-1} 層細胞面の反応を加算することによって "シードセル生成面 (seed-cell selecting plane)" を作成し, シードセル生成面内の最大値を持つ位置を特定する. シードセル生成面内で最大値を持つ細胞をシードセル (seed cell) と呼ぶ. 我々も, この方法に従って抽出すべき特徴の位置を決定している. 図 4.2 はシードセル生成面を用いた, 抽出すべき特徴の位置の決定と結合強化のプロセスを模式的に示したものである.

各結合の強化は, 前層 U_{Cl-1} 層の細胞の反応に比例して強化される. なお結合強化のプロセスにおいては, 同じ細胞面内に含まれる細胞の結合係数は, 強化された細胞と同じ分だけ強化されることを仮定している. 計算機上の実装では, ひとつの細胞面内の細胞の持つ結合は, ひとつの結合を共有することによって結合に用いる計算機資源を減らすことが可能になる. また, Földiak らが提案している特徴を動かし Trace Hebb 則のような学習方式を考えることによって, このような結合は実現できる [20] [119].

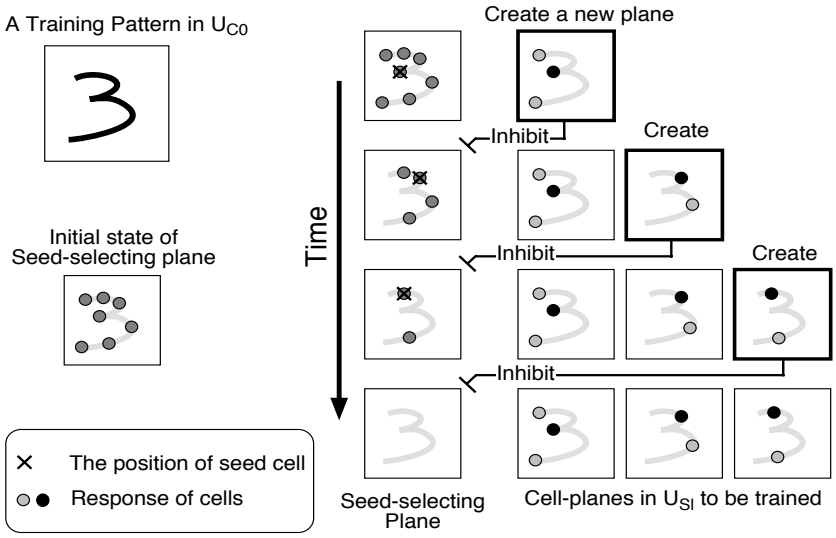

図 4.2: シードセル生成面による代表特徴の抽出方法: 前層の U_{Cl-1} 層の反応をぼかして足し合わせたものがシードセル生成面の初期状態である. シードセル生成面上で反応が大きい細胞を選びだし, その位置に抽出すべき特徴があるものとしている. その特徴を抽出する細胞面を生成しては, シードセル生成面に対して抑制をかけるというプロセスをシードセル生成面上に反応がなくなるまで繰り返す.

4.4 閾値の設定

S 細胞は, 多少の変形した特徴でも受けつけ発火するが, この変形の度合いをどこまで許すかを決定しているのが閾値である. この節では, 閾値と S 細胞との関係を述べる.

4.4.1 S 細胞の数学的表現

ネオコグニトロンの S 細胞の出力は式 (4.1) によって決定される. 簡単のために, ただ 1 個の S 細胞を考える. 図 4.3 は 1 個の S 細胞が入出力関係を模式的に示したものである. 前段の C 細胞からの入力は, $\boldsymbol{x}$ で表すことが出来る. 図に示したような関係の

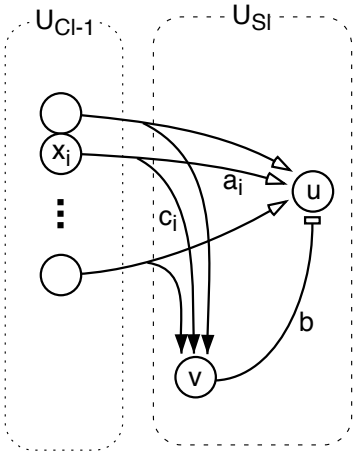

図 4.3: 1 個の S 細胞に対する入力. 前段の C 細胞からの興奮性の結合は a_i としている. 抑制性の入力は V 細胞から結合強度 b によって伝達される.

S 細胞の出力は式 (4.1) を書き換えて

$$u = \frac{\theta}{1-\theta} \cdot \varphi \left[\frac{1 + \sum_i a_i x_i}{1 + \theta b v} - 1 \right], \tag{4.8}$$

とすることが出来る. ただし,

$$v = \sqrt{\sum_i c_i \cdot \{x_i\}^2}. \tag{4.9}$$

としている. 式中の添字 i は式 (4.1) の $\boldsymbol{\nu}$ に対応する. 式 (4.8) は, さらに簡単に

$$u = \gamma \varphi \left[\frac{\sum_i a_i x_i}{bv} - \theta \right] \quad (\text{ if } \ bv > 0 \ ), \tag{4.10}$$

と書くことが出来る [25] [26]. ただし, γ は

$$\gamma = \frac{\theta}{1-\theta} \cdot \frac{bv}{1+\theta b v}. \tag{4.11}$$

としている.

結合 a_i と b とはそれぞれ結合 $a(\boldsymbol{\nu}, \kappa, k)$ と $b(k)$ とに対応する. 4.3 節で述べた結合強化方法は, 以下のように書き直せる: あるパターン $\boldsymbol{x}^*$ に対して勝者になった S 細胞の結合は,

$$\begin{aligned} \Delta a_i &= q \cdot x_i \\ \Delta b &= q \cdot v \ . \end{aligned}$$

と強化され, 敗者になった場合には何の変更も加えられない.

さて, ここである S 細胞が K 個のパターン $\boldsymbol{x}^1, \boldsymbol{x}^2, \cdots, \boldsymbol{x}^K$ に対して勝者になったとする. 結合 a_i と b は上述の式に従って更新されるものとすれば, 結合 a_i と b とは

$$\begin{aligned} a_i &= q \sum_k^K c_i \cdot x_i^k \\ b &= q \sum_k^K \sqrt{\sum_i c_i \cdot \{x_i^k\}^2}. \end{aligned} \tag{4.12}$$

と書ける. 得られた式 (4.12) を 式 (4.8) に代入することによって,

$$u = \gamma\varphi \left[\frac{\sum_k \sum_i c_i x_i^k \cdot x_i}{\sum_k \sqrt{\sum_i c_i \{x_i^k\}^2} \sqrt{\sum_i c_i \{x_i\}^2}} - \theta \right]. \tag{4.13}$$

を得ることが出来る.

ここで, $\boldsymbol{x}$ と $\boldsymbol{y}$ との内積を

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_i c_i x_i y_i.$$

で定義するようなベクトル空間を特徴空間と考えると, 式 (4.13) は,

$$u = \gamma\varphi \left[\frac{||\boldsymbol{X}||}{\sum_k ||\boldsymbol{x}^k||} \frac{(\boldsymbol{X}, \boldsymbol{x})}{||\boldsymbol{X}|| \cdot ||\boldsymbol{x}||} - \theta \right], \tag{4.14}$$

と書き直せる. ただし

$$\boldsymbol{X} = \sum_k \boldsymbol{x}^k.$$

と表すものとする. ここで,

$$s = \frac{(\boldsymbol{X}, \boldsymbol{x})}{||\boldsymbol{X}|| \cdot ||\boldsymbol{x}||}, \tag{4.15}$$

$$\lambda = \frac{\boldsymbol{X}}{\sum_k ||\boldsymbol{x}^k||}, \tag{4.16}$$

という量を導入することによって, 式 (4.8) は最終的に:

$$u = \gamma\lambda \cdot \varphi\left[s - \frac{\theta}{\lambda}\right]. \tag{4.17}$$

と書き直せる. 十分な訓練後, 結合 b が十分大きな値になった場合には, $bv \gg 1$ という関係が得られるので

$$\gamma \sim \frac{1}{1-\theta},$$

という定数と考えることが出来る.

4.4.2 S 細胞の閾値と選択性の関係

前節の数式表現では最終的に式 (4.17) が得られた. この式では, 入力特徴をベクトル $\boldsymbol{x}$ とし強化に用いられたパターンを重ね合わせたベクトルをリファレンスベクトル $\boldsymbol{X}$ としてあらわしてる. 強化によって $a_i = q\sum_k c_i x_i^k$ が得られる結合係数なので, リファレンスベクトル $\boldsymbol{X}$ の第 i 成分と結合係数 a_i との関係は

$$X_i = \sum_k x_i^k = \frac{a_i}{q \cdot c_i}$$

となる. c_i が i に関して変化しない場合には $\boldsymbol{X}$ と $\boldsymbol{a}$ とは平行なベクトルということになる. S 細胞は, これら 2 つのベクトル $\boldsymbol{X}$ と $\boldsymbol{x}$ の間の方向余弦 s を計算し (式 (4.15)

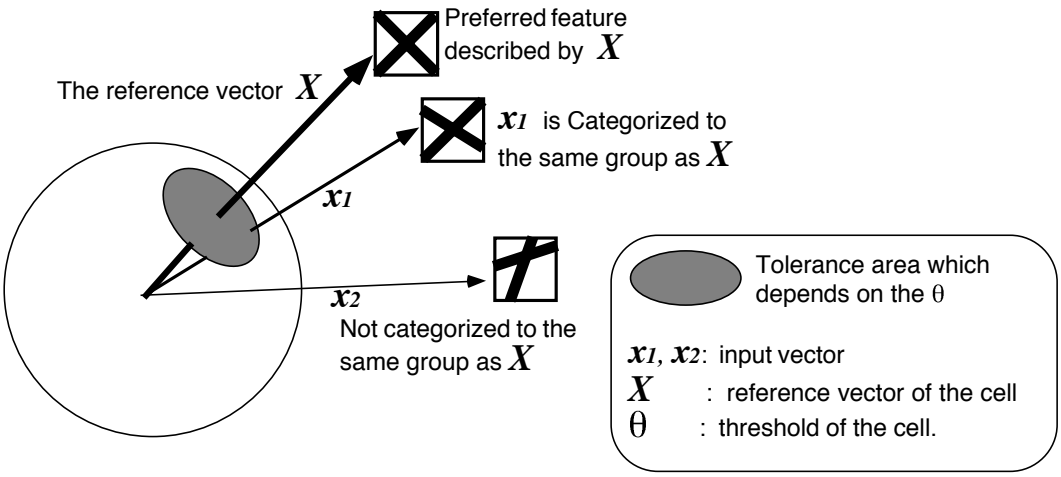

図 4.4: S 細胞の閾値と反応する領域 (許容領域) の幾何学的解釈. 図中の矢印が入力される特徴のベクトルとしている. S 細胞は最も特異的に反応する特徴を持つがそれをリファレンスベクトルとして表している. 単位球面上でのリファレンスベクトルを中心とした斜線で表した領域が許容領域で, 入力がこの領域内にあれば, この S 細胞は発火する.

), この値が θ/λ より大きい場合に出力を出すことがわかる. 図 4.4 は以上の解釈を幾何学的に表したものである. ベクトル $\boldsymbol{X}$ が, 結合強化によって得られたリファレンスベクトルを表し, $\boldsymbol{x}$ が入力ベクトルを表している. 原点を中心とした球を考え, この表面上で, S 細胞が反応を示す領域を考えると図中に示した通りリファレンスベクトルを中心にして, そのまわりの円状の領域の内側となる. この領域を許容領域と呼ぶことにする. この許容領域の中に落ちてくる入力ベクトルであれば, S 細胞は反応するが, 許容領域以外のベクトルであれば, 全く反応を示さないことになる.

この許容領域の半径は, θ/λ で決定される (式 (4.17) 参照). 閾値を低くした場合, この許容領域は広がり, S 細胞はより変形したパターンを, 自分が特異的に反応するパターンと同じものと見なす. 逆に閾値を高くした場合には, 許容領域は小さくなるので S 細胞は, あまりパターンの変形を許さなくなる.

4.4.3 抑制性結合の強化方式の変更点

式 (4.15) にも示す通り, 前述の従来の結合強化方式では許容領域の半径は閾値である θ だけでなく λ にも依存している [33] [26] . λ は式 (4.16) によって得られる値であり, $\lambda \leq 1$ である.

許容領域上で考えた場合, λ は結合強化に用いられたベクトルの集合 $\left\{\boldsymbol{x}^k\right\}$ が, どれだけ広がっているかに依存する量である. 完全に平行なベクトルしか強化に用いなかった場合 $\lambda = 1$ が保持されるが, 強化に用いたベクトルの球面上での分布の半径が広がるほど λ の値は小さくなっていく. もし λ が小さい値をとれば, 式 (4.15) における実効的な閾値は大きくなってしまう. つまり, λ という量は, 学習パターンに含まれる特徴に依存しているので制御不可能なパラメータとなってしまうので望ましくない. そこで, 本実験で用いた結合強化は, 抑制性の結合 b の決定方式を以下のように定めて $\lambda = 1$ を常に保持するように設定している [77]:

$$b = \sqrt{\sum_i \frac{\{a_i\}^2}{c_i}}. \tag{4.18}$$

このように抑制性結合 b を決定することによって, 実効的な閾値を制御できるようになる.

4.4.4 低い閾値に設定されたネットワークの場合

それでは閾値が非常に低く設定された場合の S 細胞とネットワークについて考えてみる. まず S 細胞においては 図 4.5 の左側に示すような許容領域を持つような状態を基準として考える. 基準状態においては, $\boldsymbol{X}$ をリファレンスベクトルに持つような S 細胞は $\boldsymbol{x}_1$ に示す特徴程度の変形には反応し, $\boldsymbol{x}_2$ のように変形が大きい場合には反応しないものとする. この基準状態から閾値を下げると S 細胞においては 図 4.5 に示すように許容領域が広がることになる. 許容領域が広がれば広がるほど, 入力ベクトルがその領域に含まれる可能性が高くなる. 結合強化時において, ある S 細胞の持つ許容領域内に入力ベクトルが存在する場合, その S 細胞 が勝者になり, この勝者の結合のみが強化される. また入力ベクトルが, どの S 細胞の許容領域にも入らなかった場合には, この入力を新たな種類の特徴とみなして, 新たな S 細胞が生成される.

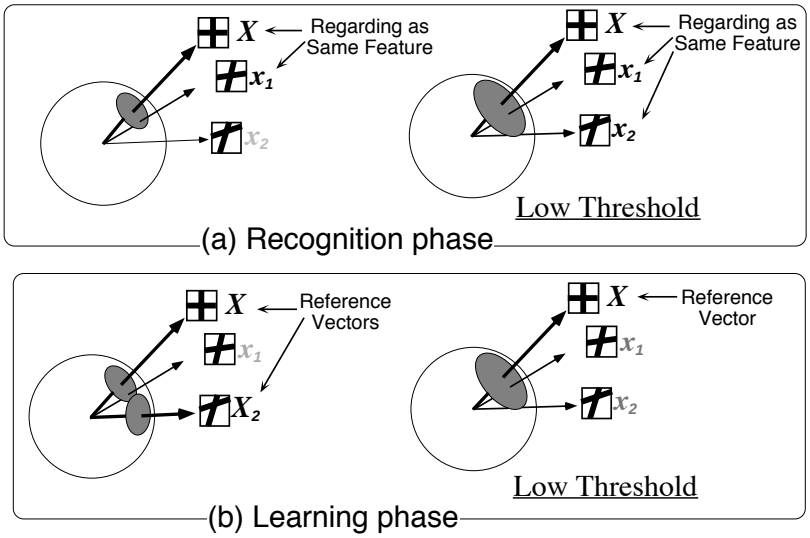

図 4.5: 閾値が低くなると許容領域は大きくなる. (a): 認識時において低い閾値に設定していると, $\boldsymbol{x_2}$ のような, かなり変形した特徴に対しても $\boldsymbol{X}$ の変形したパターンと見なして反応する. (b): 学習時において低い閾値に設定していると, $\boldsymbol{x_2}$ のような特徴が入力されても $\boldsymbol{X}$ が勝者となるため新たな $\boldsymbol{X_2}$ をリファレンスベクトルに持つような S 細胞は生成されない.

結合強化時において閾値が低いと S 細胞は多くの入力ベクトルに対して勝者となる可能性が高くなる. 図 4.5 (b) に示すように, 低い閾値では, $\boldsymbol{X}$ をリファレンスベクトルに持つ S 細胞は, $\boldsymbol{x}_1$ だけでなく $\boldsymbol{x}_2$ も許容領域に入ってくるので, $\boldsymbol{x}_2$ が, 新たな S 細胞のリファレンスベクトルになることはない. したがって, 結合強化時には, 少数の S 細胞しか作成されないことになる. Fukushima と Tanigawa は, 閾値が低い場合には各階層で, 特徴の統合を行なっていく際に特徴情報が失われる可能性について指摘している [31].

図 4.6 は S 細胞の閾値を低く設定してネットワークを構成した場合のシミュレーション例を示している. 図中の小さい四角が細胞面を表し, 細胞面内の黒い点が反応している細胞を表しており, 点の色が黒いほど強い反応である. U_{C5} 層が認識細胞層をあらわしており, 数字カテゴリに対応する認識細胞の発火が認識したパターンをあらわしている. 図 4.6 (a) では, 結合強化に用いられたパターン “7” が入力された場合のネットワークの反応例で, 正しく認識されている. 一方 図 4.6 (b) では, “9” が入力されているが,

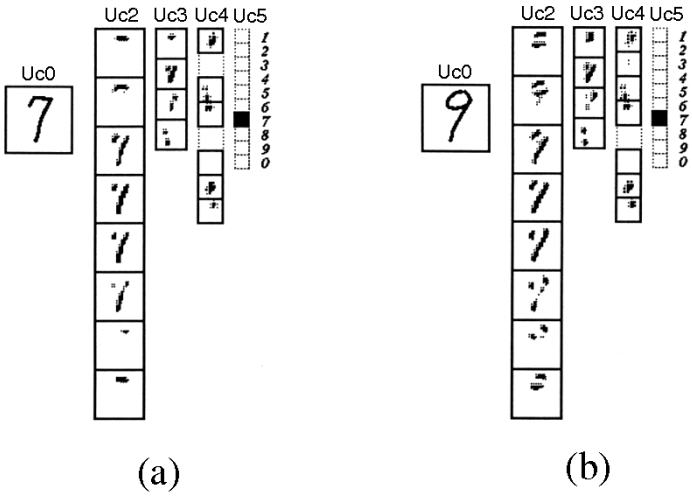

図 4.6: 非常に低い閾値に設定したネオコグニトロンのシミュレーション例. 図中の四角は C 細胞面を表し, 黒い部分が発火している細胞を表す. S 細胞面は表されてないが, C 細胞が S 細胞の反応を空間的にぼかしているだけなので S 細胞のおおよその発火状態は, ほぼ同じものと考えられる. この例では 十分な数の細胞面が作られていないのが作成されていないのがわかる. (a): ある学習パターンを提示したときの反応例. 正しく "7" と分類している. (b): 未学習パターン "9" に対する反応例. "7" と誤認識されているのがわかる. ([31] より)

"7" と誤認識されている. 図 4.6 の (a) と (b) とを比較してみるとわかるように U_{C3} 層で, これらのパターンによる反応が非常に似たものになっていることがわかる. その結果として, それ以降の層ではパターンの違いを区別できず, 誤認識される結果となっているものと考えられる.

4.4.5 高い閾値に設定されたネットワーク場合

それでは, 今度は S 細胞の閾値が高く設定された場合を考えてみよう. 図 4.7 の左側は 図 4.5 と同様に基準状態を示すものとする. 閾値が基準状態より大きくなった場合, 細胞の持つ許容領域の大きさは小さくなる (図 4.7 を参照). 許容領域が小さいということは, 入力ベクトルがその領域外に存在する可能性が高いということを意味する.

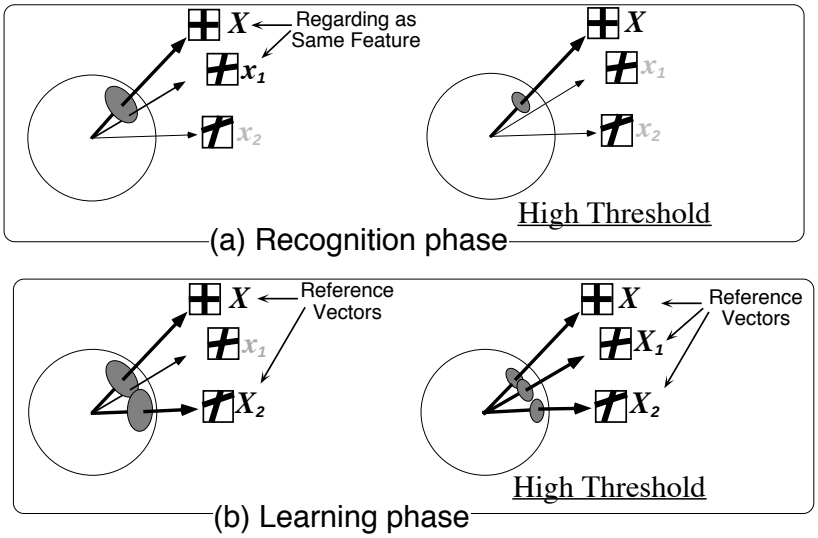

図 4.7: 閾値 θ を高く設定した場合, 許容領域は小さくなる (a): 認識時においては, 特異的に反応する特徴に対する選択性を高める. 閾値を高く設定することによって $\boldsymbol{x_1}$ は $\boldsymbol{X}$ と, おなじ特徴とは見なされなくなる. (b): 学習時においては, $\boldsymbol{X_2}$ で表される特徴だけでなく $\boldsymbol{X_1}$ で表される特徴も $\boldsymbol{X}$ と同じ特徴とは見なされないので, これらをリファレンスベクトルとした S 細胞が生成される.

結合強化時において, ある S 細胞の持つ許容領域内に入力ベクトルが入ってきた場合, その S 細胞 が勝者となり結合が更新される. どの S 細胞の許容領域にも入力ベクトルが入らなかった場合には, 新たな S 細胞が生成される. したがって, 小さい許容領域しか持たない S 細胞では, 結合強化時に入力された全ての特徴をカバーするために多くの種類の S 細胞が必要であることがわかる (図 4.7 (b) 参照). 多数の S 細胞が生成されると, 結合強化時に用いられたパターンセットに含まれる細かな特徴の違いも抽出できるようになる.

一方, 認識時において閾値を高く設定している場合, S 細胞の変形に対する許容度が小さくなるので, 汎化能力が低下することが考えられる (図 4.7 (a) 参照). すなわち, S 細胞は, 学習した特徴のみに対して反応するようになるのである. したがって, 未知の特徴を多く含むパターンに対しては, S 細胞の反応がでなくなり, 誤認識を犯しやすくなる. これは学習セットに対するオーバーフィッティングと呼ばれる現象である.

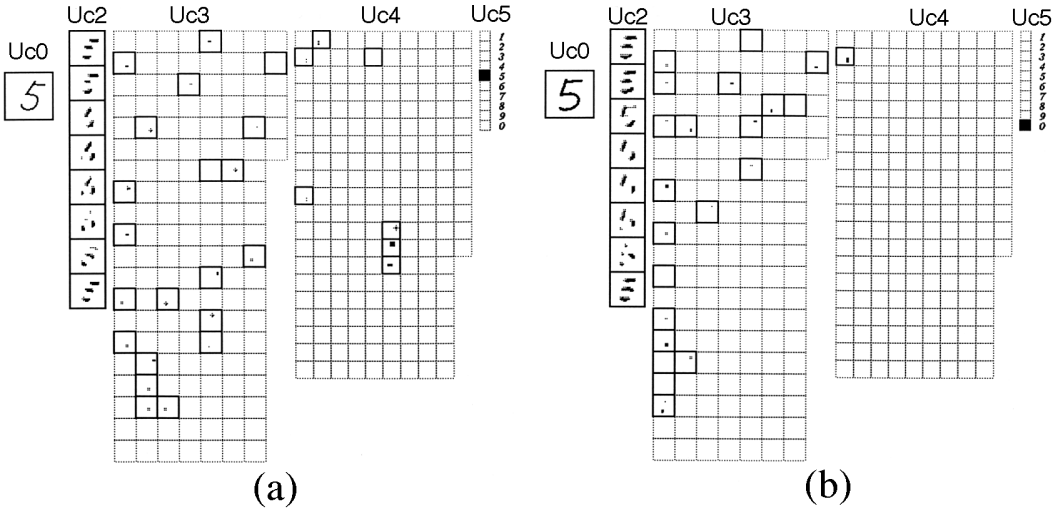

図 4.8: 非常に高い閾値に設定したネオコグニトロンのシミュレーション例. 図中の四角が C 細胞面を表し黒い部分が発火している細胞を表す. 低い閾値の時とは異なり非常に多くの細胞面が生成されていることがわかる. (a): ある学習パターン “5” を入力したときの反応例. 正しく “5” と分類されているのがわかる. (b): おなじ “5” でも変形している場合は正しく認識されない. ([31] より)

図 4.8 は閾値を非常に高く設定した時のネオコグニトロンの C 細胞の反応を示している. 図中の小さい四角が細胞面を表し, 細胞面内の黒い点が反応している細胞を表している. 図 4.8 (a) は結合強化に用いたパターン “5” を入力に与えた時の反応例である. ネオコグニトロンは正しく, このパターンを識別している. 図 4.8 (b) は, 少し変形させた “5” を入力に与えた場合の反応例である. 認識結果は誤認識であった. 図中 (b) の U_{C4} 層では, ほとんど細胞が反応していないことがわかる. したがって, 図に示したシステムにとって (b) で与えたパターンには未知の特徴が含まれていることになる. 認識結果を改善するためには, より多数のパターンを使った結合強化が必要なのがわかる. 図中のシステムは, もっと多くのパターンを用いて結合強化を行なえば, より多数のパターンに対しても正しい認識結果を出すようになるが, それは単に変形パターンを結合強化によって覚えさせているにすぎず, 汎化能力という点では, 余り高い能力は期待できない. また認識時の計算時間も無視できないので, 認識時において, 高い閾値に設定するの

は余り現実的な方法とはいえない.

4.4.6 閾値の設定方法のまとめ

ここで, 閾値の設定と S 細胞の変化についてまとめる. 簡単のために直線の検出を行うような S 細胞を 4.3 節で示した結合強化アルゴリズムで生成させることを考える. 結合強化時には, 閾値が低いと (図 4.9 の一番下の行) では, ある細胞がほとんどの方位に反応しするので, 直交するような方位を検出するような細胞しか作られない. すなわち, 図中灰色であらわした方位に特異的な細胞は作られず, この方位は矢印の先であらわした勝者の S 細胞によって表現されることになる. この状態では, 直交する方位の違い位しか区別できないことになる. 閾値をあげていくと, 許容領域が小さくなっていき, 様々な方位に特異的な S 細胞が生成される. 閾値を非常に高くした場合, 全ての方位に対する直線検出細胞が生成される.

次に, この直線検出細胞の認識時における動作を考える. 方位を横軸に取り, 細胞の反応を縦軸に取ると細胞のもつチューニングカーブを描くことができる. 閾値が高い場合 (図 4.10 の上段) では, 各 S 細胞の持つチューニングカーブは重なっておらず, 矢印で示した位置の方位を持つ直線が入力されたとしても, どの細胞も反応しない. 閾値が低い場合 (図 4.10 の下段) では各 S 細胞のチューニングカーブが重なっており, 矢印の位置の方位を持つ直線が入力された場合, 真ん中と右端の S 細胞が反応することが分かる. すなわち, 入力特徴がその細胞のもっとも特異的に反応する特徴でなくても, 複数の細胞の反応 (ポピュレーションコーディング) によって表現することが出来る. これによりロバストな直線検出が行えると考えられる. 認識時においては, 閾値が低いことは特徴の変形に対して許容力が増えることを意味する. これは汎化能力の向上に繋がるものと考えられる. 一般的に, パターン認識では汎化能力は必要不可欠な能力である. 閾値を低く設定することの問題点は, むしろ結合強化時に十分な量の S 細胞が生成されないことにあるのである.

学習時と認識時において, 閾値を変更することによって得られる効果をまとめると 表 4.1 のようになる.

閾値設定の問題点は, 認識時に閾値が低いのはいいが, 結合強化時にも閾値が低いと十

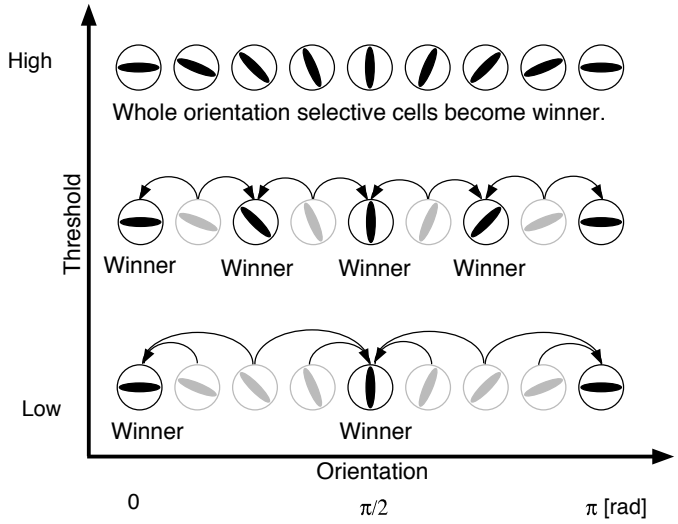

図 4.9: 結合強化時に, 閾値が低い (一番下の行) となかなか勝者になれず直交する方位を抽出する S 細胞くらいしか作られない. 図中の灰色で示した方位抽出細胞は作られないことをあらわす. この方位の直線成分は矢印であらわされる勝者の S 細胞によって表現される. 閾値が上がるにつれ (2 行目), 徐々に勝者になることができるようになり, 閾値が非常に高い場合 (一番上の行) では, 全ての方位に対する S 細胞が生成される.

分な数の S 細胞を生成できない, ということであった. Fukushima と Tanigawa は, ネオコグニトロンに競合学習を用いて学習させる時に, 閾値を学習時において高く設定し, 認識時においては低く設定することを提唱した [31]. この方式をとると, 一つの特徴やパターンに対して複数個の細胞 (面) が発火するようになる. Fukushima と Tanigawa らは, 二つの閾値を用いる方式を, 小さいデータベースによって検証しているが, 我々は実画像を用いている ETL-1 データベースを使ってこの方式を検証した [102][30][29] [70][71].

4.4.7 閾値と生成される S 細胞面の個数

ETL-1 データベースを用いて, 結合強化時の閾値と生成される S 細胞 (面) の関係を U_{S3} 層で調べたのが 図 4.11 である. 図 4.11 の横軸は学習パターンの数を表し, 縦軸が生成された S 細胞面の個数である. 図の曲線を学習曲線と呼ぶことにする. 閾値の大き

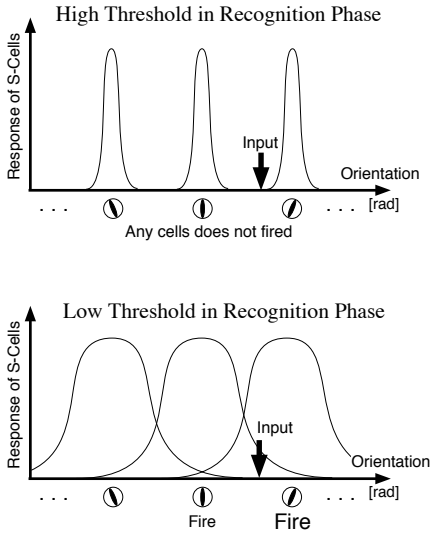

図 4.10: 細胞のチューニングカーブ. 上段の図は閾値が高い場合を表す. 各細胞の持つチューニングカーブは重なっておらず, 矢印の位置の方位の直線が入力されても, どの細胞も反応を出さない. 下段の図は閾値が低い場合であり, 各細胞のチューニングカーブは重なっている. 矢印の位置の方位の直線が入力されると 2 つの細胞が反応することが分かる.

さが $\theta = 0.75$ の場合, 学習曲線は, およそ 1,000 個のパターンを与えられたところで飽和しているのがわかる. これより大きな閾値 ($\theta = 0.8$) の場合 1,000 パターンを与えただけでは学習曲線は飽和状態にはならない. それとは逆に, これより低い閾値 ($\theta = 0.5$) の場合には, もっと少ないパターン数で飽和状態になる.

さて, それでは結合強化時のパターンに, 1,000 個のパターンが与えられた時にはどのような閾値に設定するのが良いのであろうか？もし $\theta > 0.75$ であれば, 学習曲線は飽和状態にならないであろう. 飽和状態にならないということは, まだ抽出できていない特徴もパターンセットの中に存在するということを意味する. したがって, この場合の閾値は 0.75 以下の値が望ましいということになる. しかしながら, あまり低い値に設定し過ぎると学習曲線は, もっと早い段階で飽和してしまうのでパターンセットを十分に活用しているとはいい難い. したがって, この場合は 0.75 程度にするのが良いであろう.

表 4.1: S 細胞の結合強化時と認識時における閾値の設定のまとめ

	低い閾値	高い閾値
結合強化時 (競合学習)	十分な数の特徴抽出細胞が形成されない	多くの特徴抽出細胞が形成される
認識時	汎化能力が高い	学習セットに対しオーバーフィッティングしてしまう.

4.4.8 構造上の制限から来る閾値の上限

前節では, 結合強化時のパターンセットの大きさによって望ましい閾値には上限があることを述べた. この節ではネオコグニトロンが持つ構造上の要請からくる上限について述べる. ネオコグニトロンについて説明した際に述べたように, U_{C3}, U_{C4}, U_{C5} の各層では細胞の密度が前層の S 細胞層よりも下がっている. 図 4.12 は, この密度が減らしている様子を 1 次元のネオコグニトロンで表したものである. ここで, 議論を簡単にするために結合も細胞の出力もバイナリなものを考える. つまり結合も出力も $\{0, 1\}$ の 2 状態しかとらないものとする. さらに 図 4.12 中の U_{Cl} の C 細胞は, 結合している S 細胞が発火していれば発火するものとし, U_{Sl+1} の持つ結合は, $\{1, 1, 1\}$ であると仮定している.

図 4.12 (a) のように U_{Sl} 層の, ただ 1 個の細胞 ((a) では 5 番めの S 細胞) が発火したとすると, このパターンをぼかしたようなパターンが U_{Cl} 層に現れる. U_{Cl} 層では, 2, 3, 4 番め C 細胞が発火することになる. この場合には, U_{Sl+1} 層の 3 番めの細胞が強く反応することになる. その他の S 細胞 (1, 2, 4, 5 番めの細胞) が発火するかどうかは, 閾値に依存している.

図 4.12 (a') のように U_{Sl} 層での発火が 2 個分下にずれたようには, U_{Cl} 層では, 3, 4, 5 番めの C 細胞が発火し, U_{Sl+1} 層では, 4 番めの S 細胞が発火する.

それでは, 図 4.12 (b) のように U_{Sl} 層での発火が 1 個分下にずれた時について考えて

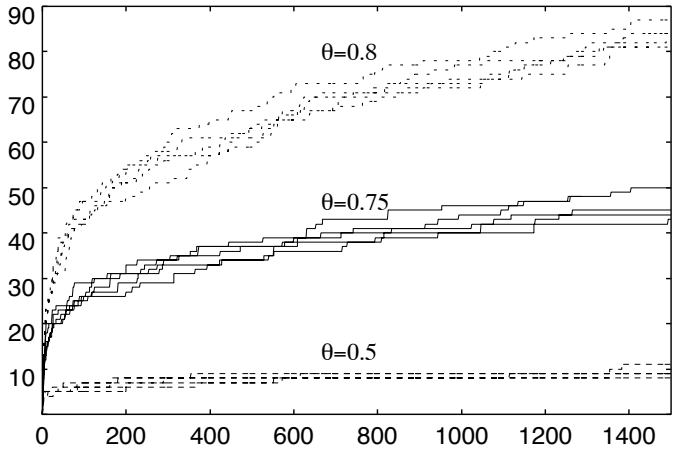

図 4.11: 結合強化回数に対する S 細胞面の生成. 横軸に結合強化回数をとり, 縦軸に S 細胞面の個数をとっている. 曲線は閾値によって変化するがある程度強化を繰り返していくと飽和していく.

みよう. この場合は, 図 4.12 (a) と図 4.12 (a') とのちょうど中間の状態を考えていることになる. U_{Cl} 層では, 3, 4 番めの C 細胞が発火する. 結合の形状を考えると U_{Sl+1} 層で, S 細胞が発火するかどうかは, 閾値次第ということがわかる. つまり閾値を高くとりすぎると図 4.12 (b) の状態では, 発火しなくなってしまうのである. ネオコグニトロンの設計思想からいえば, パターンの平行移動に対して連続的な応答をさせるように設定した方が望ましいのは明らかである. では, その上限とは, どの程度なのであろうか. もう少し定量的な議論をしてみよう. U_{Cl} 層の表現をベクトルで表した場合, 図 4.12 (a) と図 4.12 (b) は, それぞれ:

$$\begin{aligned} \boldsymbol{x}_{(\mathrm{a})} &= (0,1,1,1,0,0)^{\mathrm{T}} \\ \boldsymbol{x}_{(\mathrm{b})} &= (0,0,1,1,0,0)^{\mathrm{T}}. \end{aligned}$$

とかける. $\boldsymbol{x}_{(\mathrm{a})}$ と $\boldsymbol{x}_{(\mathrm{b})}$ との方向余弦は

$$\frac{\boldsymbol{x}_{(\mathrm{a})} \cdot \boldsymbol{x}_{(\mathrm{b})}}{||\boldsymbol{x}_{(\mathrm{a})}|| \cdot ||\boldsymbol{x}_{(\mathrm{b})}||} = \frac{2}{\sqrt{6}} \sim 0.817. \tag{4.19}$$

となる. パターンの平行移動に対して, U_{Sl+1} 層でも連続的な応答が望ましい場合には,

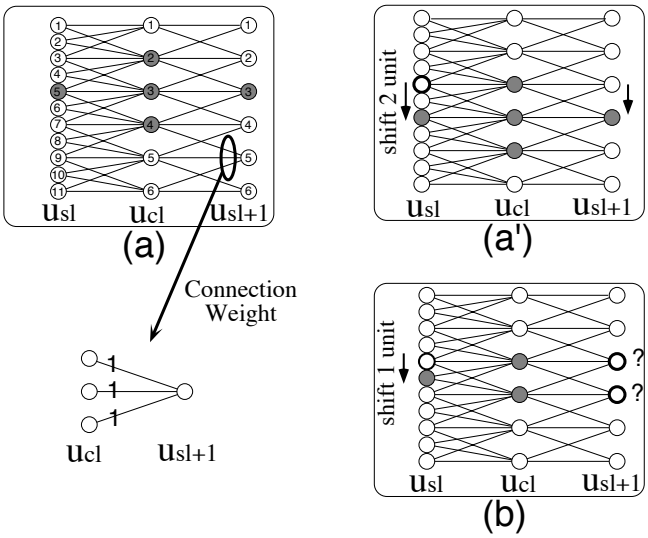

図 4.12: (a): U_{Sl} 層の 5 番目の S 細胞 (グレーの丸で表されている) が発火した場合, U_{Cl} 層の発火は, このパターンをぼかしたようなものになる. すなわち 2, 3, 4 番目の細胞が発火するもし, U_{Cl} から U_{Sl+1} への結合が $\{1, 1, 1\}$ であれば, U_{Sl+1} 層においては 3 番目の細胞が反応する. (a)': U_{Sl} 層で与えられるパターンが (a) であたえられるものより下に 2 個シフトしたようなものの場合. U_{Sl} 層では, 7 番目の S 細胞が発火し U_{Cl}, U_{Sl+1} 層では, (a) で現れるパターンを 1 個下にシフトしたようなものとなる. (b): U_{Sl} 層で与えられるパターンが (a) で与えられるものより下に 1 個シフトしたようなものの場合. U_{Cl} 層で現れるパターンは (a) のものとも (a)' のものとも異なる. この場合 U_{Sl+1} で 3 番目もしくは 4 番目の S 細胞が発火するかは閾値に依存している.

閾値 θ_{l+1} を,

$$\theta_{l+1} < 0.817. \tag{4.20}$$

に設定しなければならない. さもなければ平行移動した特徴を S 細胞が同じ特徴ではないと見なすようになる. この閾値の上限は, U_{Sl} 層から U_{Cl} 層の密度比と, C 細胞への入力結合の広がりに大きく依存している. U_{Cl} 層での細胞密度が疎になるほど, この上限は低くなり, 結合の広がりが大きくなるほど高くなる. また, この上限値は 1 次元のネオコグニトロンで考えたものあり, パターン認識実験で用いるような 2 次元のものを考えると更に低い値となる.

この実験で用いたネオコグニトロンは, U_{S2} 層から U_{C2} 層の密度比が 3:1 となり, 通常の入力結合の広がりのままでは, かなり低くなる. 我々は U_{C2} 層への入力結合の広がりを大きくして, 結合強化時の高い閾値でも, この問題が起こり難くなるようにしている.

4.5 認識層について

U_{S5} の S 細胞の結合も, 他の層の S 細胞と同じ結合強化法によって更新される. ETL-1 データベースに含まれるパターンは様々なパターンを含んでいるので, これを用いて U_{S5} 層の S 細胞の結合を強化した場合には, 一つの文字カテゴリに対して複数種類の細胞面が生成される. U_{S5} 層の反応は, U_{C5} 層へ伝達される. U_{C5} の細胞面の大きさは 1×1 である. この層の細胞は認識細胞として働く. パターン認識のためには, この U_{C5} 層の反応を '0' から '9' までの文字シンボルに結び付ける操作が必要である. 従来のネオコグニトロンでは, 結合強化時に U_{C5} 層の細胞一つ一つに与えられたパターンが持っているシンボル情報のラベルをつけ, 認識時に U_{C5} 層の中で最大値を出す細胞の対応しているラベルによって文字パターンを分類していた [31], [70], [71].

我々は, このパターンとラベルとの間の対応づけ作業に対し, 新たにパターン分類回路 U_6 層を導入した [103] [98]. パターン分類回路は, 単純パーセプトロン [85] のような全結合型の回路である. したがって, システム全体の構成は, 図 4.13 のようになる. このシステムの中で, ネオコグニトロンはパターンを特徴によって分類するサブシステムとして働く. パターン分類回路は, 前述の通りパターンとシンボル情報を結び付けるサ

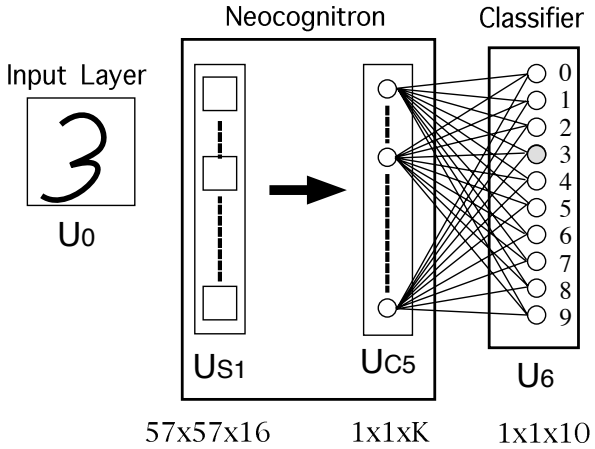

図 4.13: パターン分類回路の導入. ネオコグニトロンをパターンの形状による分類機械と見なし, 文字シンボルとの結びつけを行う.

ブシステムとして働く. パターン分類回路の学習には,

- 誤り訂正学習
- LVQ2
- コスト関数を導入した学習法

を用いて, 各結果の認識率を比較してみた.

4.6 認識結果の比較

4.6.1 文字パターンと学習方法

図 4.14 に ETL-1 に含まれる文字パターンの一例を図を示す.

我々は, “学習パターン”, “バリデーションパターン”, “テストパターン” という 3 種類のパターンセットのクラスを以下のように定義している [102][30][29] [70][71].

学習パターンは, S 細胞への入力結合荷重を決定するために用いられるパターンセットである. 学習パターンは各 U_{S3} 〜 U_{S5} 層に 5 回ずつ提示をして, 結合の競合学習に

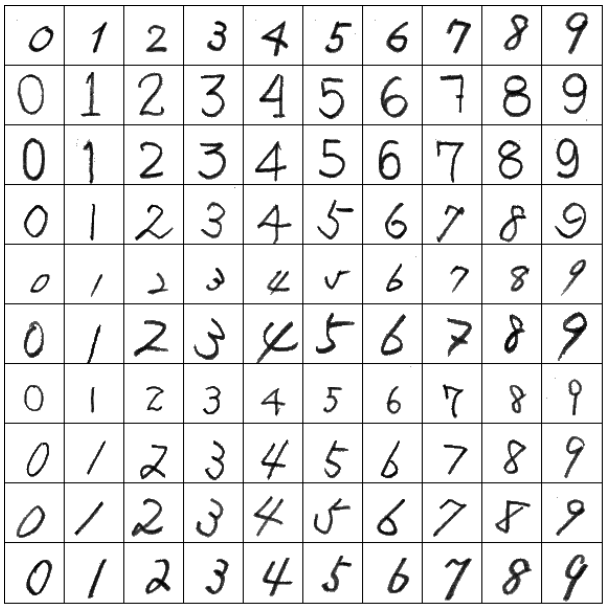

図 4.14: ETL-1 データベースに含まれるパターンの例

よる自己組織化を行なった. パターンは, 各カテゴリについて各々 100 パターン (1,000 パターン) をランダムに選択した.

バリデーションパターンは学習時の閾値と認識時の閾値を最適化するために用いている. このパターンは S 細胞の入力結合係数の学習には用いてはいない. パターンは各カテゴリについて各々 100 パターン (1,000 パターン) を用意しているが, これらのパターンは学習パターンに含まれるパターンと重複がないようにしている.

最後のテストパターンはネオコグニトロンの認識率を測定するためのパターンセットである. このパターンは学習パターンやバリデーションパターンと重複がないように, 各カテゴリについて各々 300 パターン (3,000 パターン) を用意している. したがってテストパターンは, 提案する認識システムにとっては, 新奇なパターンセットとなる.

学習パターンとテストパターン以外にバリデーションパターンを用意したのは次のような理由からである. もし, 学習パターンのみで閾値を最適化する場合には, 学習時の閾値を大きくとるだけで十分である. ところが, これでは学習パターンにオーバーフィッティングしてしまう可能性があり, 学習パターンに対する認識率が向上してもテストパ

ターンのような未知パターンに対して認識率が向上するとは限らない．そこで, 学習パターンとは別のバリデーションパターンを用意してシステムが学習パターンにオーバーフィットしないように閾値を設定している．このアイディアは統計学でいうところの“クロスバリデーション”と呼ばれる手法に基づいている [6] [106].

全体の学習過程は以下に示すようなものとなる.

認識率の評価アルゴリズム

記号の意味:

θ_l^L: l 段目の結合強化時の閾値

θ_R^L: l 段目の認識時の閾値

$\boldsymbol{\theta}$: 閾値の組み合わせ, すなわち $(\theta_3^L, \theta_3^R, \theta_4^L, \theta_4^R, \theta_5^L, \theta_5^R)$.

$\boldsymbol{\Theta}$: 検証すべき $\boldsymbol{\theta}$ の集合.

認識率 R を $R = 0$ に初期化する.
最大認識率 R_{opt} を $R_{\mathrm{opt}} = 0$ に初期化する
repeat
  $\boldsymbol{\Theta}$ の中からある $\boldsymbol{\theta}$ を選択する.
  for $l = 3$ to 5 **do**
    U_{Sl} 層の結合強化を閾値 θ_l^L の下で学習パターンを用いて行う.
    U_{Sl} 層の結合強化終了後, この層の細胞の閾値を θ_l^R にする.
  end for
  パターンと文字シンボルとの結び付け操作を行う. (従来のラベリング, パターン分類回路の学習など)
  バリデーションパターンを用いて認識率を計測し R に記録する.
  if $R > R_{\mathrm{opt}}$ **then**
    $\boldsymbol{\theta}_{\mathbf{opt}} \leftarrow \boldsymbol{\theta}$
    $R_{\mathrm{opt}} \leftarrow R$
  end if
until $\boldsymbol{\Theta}$ に含まれるすべての状態を測定し終えるまで
測定された $\boldsymbol{\theta}_{\mathrm{opt}}$ の条件下で, テストパターンに対する認識率を測定する.

4.6.2 ネオコグニトロンのみによる認識率

まず, 予備実験としてネオコグニトロンのみの場合についての認識率を計測した. ただし, 最上位段 U_{S5} 層と認識細胞である U_{C5} 層の間の細胞のラベルづけは以下のようにおこなった. 入力層 U_0 にパターンが提示され, U_{S5} 層の細胞面が新たに生成されると, その細胞面からは, 入力パターンのカテゴリを表す U_{C5} 層の 10 個の C 細胞に対して結合を作り, U_{S5} 層の細胞面がどのカテゴリに属するかを決定する. U_{S5} 層の S 細胞間での競合は同一カテゴリを示す S 細胞との間だけで行ない, 競合に勝った S 細胞が存在する細胞面の細胞だけが重ね合わせ学習を行なう. もし提示したパターンのカテゴリを示す U_{S5} 層の細胞が発火していない場合には新しい細胞面を形成し学習させる. この場合, 他のカテゴリに属する S 細胞が大きな出力を出していても, その反応は無視する.

この測定の結果, 最も良いと思われるネオコグニトロンの閾値を表 4.2 に示す.

認識時には, まず U_{S5} 層の S 細胞の出力を計算し, U_{S5} 層に含まれる全ての S 細胞と競合させ勝ち残った 1 個の S 細胞と結合している C 細 胞を発火させるようにしている. もし 1 個の S 細胞も発火しない場合にはパターンを分類不能のグループ (リジェクト) として扱うようにしている.

4.6.3 カテゴリ分類回路

我々は更に認識率を上げ, 誤答率を下げるためにカテゴリを分類する一層の回路 U_6 層を導入した.

従来の認識層 U_{C5} 層は, 出力を出している U_{S5} 細胞面に対してカテゴリのラベルづけ

表 4.2: ネオコグニトロンの場合の閾値の設定値

	$l=3$	$l=4$	$l=5$
θ_l^L	0.75	0.65	0.76
θ_l^R	0.66	0.53	0.45

表 4.3: 従来手法のラベル付けで学習させたときのネオコグニトロンの認識率

	認識率	誤答率	リジェクト
学習パターン	99.80	0.20	0.00
バリデーションパターン	98.10	1.90	0.00
テストパターン	97.83	2.13	0.03

をするものとして扱ってきた. 本研究では, U_{C5} 層は他の C 細胞の層と同じくパターンの形状を分類する役割をさせるにとどめ, カテゴリの分類は新たな回路である U_6 層を付け加えることによって行なっている. したがって U_{C5} 層の細胞の反応は, 必ずしもカテゴリに対応しているのではなくパターンの形状が類似したものに対応し, たとえ同一カテゴリに含まれるパターンでも字体が異なっていれば反応は異なったものになる. つまり, ネオコグニトロンを入力パターンの形状によってカテゴライズするシステムと考え, これの出力ベクトルを分類する回路を付け加えるというネットワークシステムを構成した (図 4.13 参照).

ネオコグニトロンの出力層 U_{C5} 層は競合学習によって K 個の細胞が生成されるものとする. これらの反応を $x_j\,(j=1,\cdots,K)$ とする. 新たに付け加える認識層 U_6 層は, 0 〜 9 までのカテゴリに対応した 10 個のユニットからなる. これらのユニットを $y_i\,(i=0,1,\cdots,9)$ とする. これらのユニットを結合する重みを w_{ij} とし, 全結合の回路を考え, 以降の学習則にしたがって, w_{ij} の更新を行なう.

この結合荷重の更新方法は LVQ2, 謝り訂正学習, コスト関数の減少, という 3 種類の方法で学習させ, どの方法がよいかを調べた. この新たに付け加えたカテゴリ分類回路を学習させるために用いたパターンセットは, 学習パターンのみである.

4.6.4 カテゴリ分類に LVQ2 を用いた場合

Kohonen らによれば, LVQ2 を使った場合, カテゴリの境界がベイズ決定理論による境界に近くなることが報告されている [54] . LVQ2 による結合更新ルールは以下の通りである. U_{C5} 層の K_{C5} 個の細胞の反応 $\boldsymbol{u}_{C5} = (u_{C5}(1), \cdots, u_{C5}(K_{C5}))^T$ と, U_{C5} と認識層 U_6 層の i 番めの各カテゴリ細胞を繋いでいる結合 $\boldsymbol{w}_i = (w_{i1}, \cdots, w_{iK_{C5}})^T$ をベクトルと考える. この場合, U_6 層の細胞の出力

$$y_i = \boldsymbol{w_i} \cdot \boldsymbol{x} = \sum_{\kappa} w_{i\kappa} u_{C5}(\kappa) \tag{4.21}$$

のうちで最大値をとるものを, 入力パターンのカテゴリとしている. μ 番めのカテゴリに属するパターンが入力された場合の学習方法は以下の通りで, まず, 第 1 勝者ユニット i^* と, 第 2 勝者ユニット i^{**} を決定する.

$$i^* = \underset{i}{\operatorname{argmin}}(|\boldsymbol{w}_i - \boldsymbol{u}_{C5}|) \tag{4.22}$$

$$i^{**} = \underset{i, i \neq i^*}{\operatorname{argmin}}(|\boldsymbol{w}_i - \boldsymbol{u}_{C5}|) \tag{4.23}$$

次に第 1 勝者ユニットのリファレンスベクトルは

$$\Delta w_{i^*\kappa} = \begin{cases} +q(u_{C5}(\kappa) - w_{i^*\kappa}) & \text{if}(i^* = \mu), \\ -q(u_{C5}(\kappa) - w_{i^*\kappa}) & \text{else} \end{cases} \tag{4.24}$$

とする. さらに第 1 勝者ユニットの分類が誤っていて $(i^* \neq \mu)$, なおかつ第 2 勝者ユニットが正しく分類している $(i^{**} = \mu)$ 場合に限り

$$\Delta w_{i^{**}\kappa} = \begin{cases} +q(u_{C5}(\kappa) - w_{i^{**}j}) & \text{if}(i^{**} = \mu), \\ 0. & \text{else} \end{cases} \tag{4.25}$$

とする.

認識結果を評価する際には,

- 最大出力を出している細胞のカテゴリが正しい場合を正答とする
- 最大出力を出している細胞のカテゴリが誤っている場合を誤答とする

- 出力がない場合をリジェクトとする

というルールで評価した. この場合の結果は 表 4.4 のとおりである. 残念ながら, この方法ではそれほどよい結果は得られなかった.

表 4.4: LVQ2 を用いたときの認識率

	認識率	誤答率	リジェクト
学習パターン	100.00	0.00	0.00
バリデーションパターン	97.90	2.00	0.10
テストパターン	97.57	2.27	0.17

4.6.5 分類に誤り訂正学習を用いた場合

続いて, 結合荷重 $w_{i\kappa}$ の更新をパーセプトロンの学習方法の一つである誤り訂正学習 [85] によって学習 させた. 誤り訂正学習の学習ルールは以下の通りである. 誤り訂正学習では U_6 層の出力ユニットは

$$y_i = f\left[\sum_{\kappa=1}^{K_{C5}} w_{i\kappa} u_{C5}(\kappa) - \theta_i^6\right] \tag{4.26}$$

という出力を出すものとしている. さらに, 式を簡単にするために常に抑制性の出力 (-1) を出す細胞 $u_{C5}(0)$ を考え, 式 (4.26) を

$$y_i = f\left[\sum_{\kappa=0}^{K_{C5}} w_{i\kappa} u_{C5}(\kappa)\right] \tag{4.27}$$

と表すものとする. すなわち U_6 層の各ユニットの閾値は $\theta_i^6 = w_{i0}$ で表される. 出力関数 $f[\eta]$ は

$$f[\eta] = \begin{cases} 1 & (\eta > 0) \\ 0 & (\eta \leq 0) \end{cases} \tag{4.28}$$

という 2 値関数としている. あるカテゴリ μ に属するパターンが提示された時に, U_6 層の細胞には, $\zeta_i = \delta_{i\mu}$ という反応が出て欲しいものとする. ただし $\delta_{i\mu}$ はクロネッカーのデルタとする. この時, 結合荷重は次のように更新する.

$$w_{i\kappa}^{\mathrm{new}} = w_{i\kappa}^{\mathrm{old}} + \Delta w_{i\kappa} \tag{4.29}$$

ただし,
$i = \mu$ のとき

$$\Delta w_{i\kappa} = \begin{cases} +q \cdot u_{C5}(\kappa) & (y_i = 0) \\ 0 & (y_i = 1) \end{cases}$$

$i \neq \mu$ のとき

$$\Delta w_{i\kappa} = \begin{cases} 0 & (y_i = 0) \\ -q \cdot u_{C5}(\kappa) & (y_i = 1) \end{cases}$$

としている.

認識結果を評価する際には

- 出力を出している細胞が, ただ 1 個で正しい場合を正答とする
- 出力を出している細胞が, ただ 1 個で誤っている場合を誤答とする
- 複数の細胞が反応を出したり, 出力がない場合をリジェクトとする

というルールで評価した. この場合の結果は 表 4.5 の通りである. 認識率は下がったが, 誤答率もまた減らすことに成功している.

表 4.5: 誤り訂正学習を用いたときの認識率

	認識率	誤答率	リジェクト
学習パターン	100.00	0.00	0.00
バリデーションパターン	94.90	0.40	4.70
テストパターン	95.03	0.93	4.03

4.6.6 コスト関数を減少させる方法を用いた場合

最後に, 結合加重の更新をコスト関数を定義して, それを減少させるように結合荷重を更新するような学習則を用いて方法を用いた [117][116][86]. 認識細胞の形式は式 (4.27) と同じものを用いるが出力関数を $f(\eta) = \tanh(\eta)$ とする. これはコスト関数を関数を微分可能にするためである. まず, U_6 層の j 番目の細胞の出力 y_j を以下のように定める.

$$y_j = f\left[\sum_{\kappa=0}^{K_{C5}} w_{j\kappa} u_{C5}(\kappa)\right] \tag{4.30}$$

ここでも表記を簡単にするため, 式 (4.27) のように常に -1 を出力するような U_{C5} 層の第 0 番目の細胞 $u_{C5}(0)$ を導入している. 学習のコスト関数は最小 2 乗誤差

$$E[\boldsymbol{w}] = \frac{1}{2}\sum_{i,\mu}(z_i^\mu - y_i)^2 \tag{4.31}$$

で定義する. ただし μ はパターン番号とし, y_j はカテゴリ分類回路の j 番目の細胞の出力, z_j^μ は教師信号で パターンと j が一致している場合は, 1 を出しその他の場合は 0 であるような値を考える. $\delta_{i\mu}$ はクロネッカーのデルタである. このコスト関数が減少していくように重み荷重を更新していく. すなわちコストが極小になるようにするように偏微分値に比例した量を引いていくことによって行う. これは次式

$$\Delta w_{i\kappa} = -q\frac{\partial E}{\partial w} = -q(z_i^\mu - y_i)(1 - y_i^2)u_{C5}(\kappa) \tag{4.32}$$

で与えている.

認識結果を評価する際には,

- 最大出力を出している細胞のカテゴリが正しい場合を正答とする
- 最大出力を出している細胞のカテゴリが誤っている場合を誤答とする
- 出力がない場合をリジェクトとする

というルールで評価した. この場合の結果は 表 4.6 のとおりであり, 未学習のテストパターンに対して 98.13% の認識率を得ることができた.

表 4.6: コスト関数の極小化を行ったときの認識率

	認識率	誤答率	リジェクト
学習パターン	100.00	0.00	0.00
バリデーションパターン	98.40	1.50	0.10
テストパターン	98.13	1.83	0.03

4.7 検討

この章では, 実画像を用いた文字パターンデータベースを用いてネオコグニトロンにカテゴリ分類回路を導入したシステムの認識性能について論じた. カテゴリ分類回路を導入することにより, 単純なラベルづけを行う従来のネオコグニトロンよりも認識率をあげることが可能になる.

ETL-1 データベースを用いたほかの認識実験としては, 内村らが Gabor フィルタのパラメータを実験計画法に基づいて最適化したモデルがあげられる [113]. 内村らのモデ

ルは, 抽出したガボール特徴のテンプレートマッチングを行うようなモデルであり, ネオコグニトロンに比べて構造的には, かなりシンプルであるが, もっとも最適化したパラメータでも 3% 程度の誤認識率がある. したがって, 認識率という立場で論じればネオコグニトロンの方が優れていると言える.

異なった手書きデータベースである U.S.Postal のデータベースや, NIST のデータベースを用いた認識実験としては, LeCun らの Le-net1 があることは, この章のはじめにも述べたが新規パターンに対しては, やはり 5% 程度の誤り率がある. 最近では Denker らが, Le-net1 を改良し, ネットワークの自由度に制限を加えた Tangent Prop と呼ばれる手法を用いて誤り率を 2% まで減らした報告や [105] [104], Vapnik らが提案した Support Vector Machine を用いて誤り率を 1 % 程度のオーダーまで減らせるという報告などがある [92]. これらのデータベースは文字の大きさ等に対して人間が目視で正規化を行っている. したがって, データベースで学習させたあと, カメラなどからパターンを取り込んで, 上述の認識機械を使うためには, もう一段の前処理が必要である. これに対し, ネオコグニトロンは入力パターンの位置の正規化を行わずに認識が行える. これはデータベースで学習させたあと, カメラなどから取り込んだパターンを何の前処理もせずに処理できることを意味する. また後述する選択的注意機構のモデルを用いれば, 入力面に複数の文字パターンが提示されたときでも文字パターンの認識・切り出しが正確に行われることが期待できる.

第 5 章

選択的注意機構のモデル化とその応用

5.1 研究背景

前章までで述べた通り, ネオコグニトロンはフィードフォワード型の階層構造をもつ神経回路モデルである. この章では, ネオコグニトロンにフィードバック型の処理機構とを組み合わせた神経回路モデルである選択的注意機構のモデルを説明し, その応用について述べる.

フィードバックによって信号が, トップダウンで伝わる機構を有するネットワークでは, ネットワークが今, 現在処理をしている認識対象に対して注意を向けさせ, より精度の高い認識処理をさせることが可能である. トップダウン経路を持つニューラルネットの重要性は, Hinton らの Helmholtz machine [15], Xu らの Ying-Yang machine [118] や, Kawato らのモデル [51] によっても指摘されている. トップダウンの経路は, パターンを生成することが可能であり, 遮蔽されていて, 実際の入力が入ってこない場所などで, その場所に見えるべき, もしくはあって欲しいという画像パターンなどを推定することが可能なので, 能動的なパターン認識には必要不可欠な要素なのである. 実際の脳の視覚経路などにおいてもフィードバック結合が, 多々確認され [120] , 認識対象に注意を向ける機構のメカニズムを担っているのではないかという議論がなされている.

ネオコグニトロンとフィードバック型の神経回路を組み合わせたモデルは Fukushima によって提案され [24], “選択的注意機構のモデル (Selective Attention Model)” と呼ばれている. このモデルはフィードフォワードの系 (求心性経路) とフィードバックの系 (遠

心性経路) とが, お互いの経路の信号の伝達を制御するモデルである. このモデルを用いると, 複数の入力刺激が求心性経路の系に入力された時でもどちらか一方の入力刺激に注意を向けさせることが可能になり, 認識対象の切り出しなどが出来るようになる.

この章での認識対象は, 英文筆記体連結文字列である. 英文筆記体のように数個の文字パターンが連結したような複合パターンを認識する場合, パターンを構成する一つ一つの部分パターンを認識せずにパターンの切出しを行うことは, 非常に困難である. 筆記体文字列パターンは, どこまでが文字の始まりで文字の終わりかが指定しづらいので従来のパターンの正規化を行なうといった前処理を行うことが困難である. さらに文字間のつながり部分の形状などは, 前後の文字の影響で変化したりするため, 通常のパターンマッチング的手法では文字列全体の認識や文字列パターンを構成する各文字パターンの切り出しは困難である. 選択的注意機構のモデルを用いたシステムでは, 入力面に複数のパターンを提示した場合に, ある部分パターンのみに注意を向けさせることができる. したがって, このような英字筆記体のようなパターンの認識に対して, 選択的注意機構のモデルは有効であると考えられる. 先行研究として, 今川らは選択的注意機構のモデルを用いた計算機シミュレーションによって英字筆記体文字列の認識を行わせており, その能力を示している. しかしながら今川らのモデルは文字パターンが変形したような場合には, 誤認識することが多く, 認識能力自体はそれほど高いものではなかった.

そこで我々は, まず選択的注意機構のモデルの求心性経路がネオコグニトロン状のサブシステムであることに着目した. 第 4 章でも示した通り, 文字認識を行うネオコグニトロンに “折れ点検出回路” を追加したネットワークは, 通常のネオコグニトロンよりも文字認識率が向上することが Wake らによって示されている [33]. したがって, 折れ点を処理する機構を選択的注意機構のモデルに組み込むことによって認識能力が向上することが期待される. 選択的注意機構のモデルは, 求心性経路と遠心性経路との間で制御を行うような鏡像関係のようなモデルであるので, 我々は新たに遠心性経路側のサブシステムに “折れ点復元回路” を導入した. この, 折れ点検出回路と折れ点復元回路とを合わせたものを本章では “折れ点処理回路” と呼ぶことにする. 我々は折れ点処理回路を今川らのシステムに導入することによって, 今川らのシステムにおいて問題となっていた文字の認識能力を向上させることに成功した. また, 折れ点復元回路の効果によって文字の切り出し能力をも向上させている.

さらに我々は, 今川らの提案した “大域的注目機構 (サーチコントローラ)” についてもを改良を加えている [46]. 大域的注目機構は, 認識対象の観測すべき空間の範囲をある程度限定し, 認識・切り出しの効率を高めるためのサブシステムである. しかし今川らのシステムでは, 大域的注意機構は求心性経路のみを用いて出力を計算していた. この場合, 選択的注意機構が示す認識対象の位置と, 大域的注目機構が示す認識対象の位置とに矛盾がある場合, 正しく認識・切り出しが出来ない. そこで我々は, 求心性経路だけでなく遠心性経路の系も用いて, この大域的注意機構の出力を計算させている. このようにすることによって選択的注意機構の示す認識対象の位置と大域的注意機構の示す位置とに矛盾がないようにしている.

5.2 回路構成の概要

ここで提唱する文字列認識システムは階層型の多層の神経回路で構成されている. 各層を構成する細胞の出力値は非負のアナログ値をとるものとしている. 隣り合う層の細胞間の結合はすべて局所的結合で, その結合強度は細胞の相対的な位置関係に依存するような並進対称性をもつ. 図 5.1 は, システムの全体の概観を示した図である. 図 5.2 も 図 5.1 と同じくシステム全体を表す図であるが, これはシステムを構成する異なる細胞層間の結合を模式的に表わしたものである. 図 5.2 には同じ種類の細胞を 1 個ずつしか描いていないが, 実際には複数個の細胞を配置している.

このシステムは, 大きく二つの部分に分けて考えることができる. 一つは, 図 5.1 , 図 5.2 の上部の右向きの “求心性経路” で, 入力パターンの認識を行なう部分である. もう一つは, 図 5.1 , 図 5.2 下部の左向きの “遠心性経路” で, 認識した結果からパターンの想起を行なう部分である. 求心性経路においては, 信号が下位から上位の細胞へと流れる. 逆に遠心性経路では, 信号が上位から下位の細胞へと流れる.

求心性経路はネオコグニトロン [23][28] とほぼ同じ構造を持ち, パターンの認識を行なう経路としている. この求心性経路は, いくつかの “細胞段” と呼ぶ構造から構成されていて, 一つの細胞段は u_S 細胞の層 U_S と, u_C 細胞の層 U_C の 2 層構造になっている. u_S 細胞はネオコグニトロンの S 細胞に対応し, u_C 細胞は C 細胞に対応する. さらに各細胞層は, それぞれいくつかの “細胞面” と呼ぶ構造から構成される. 細胞面とは, 同じ局

所的特徴に選択的に応答する細胞が 2 次元平面状に配置された細胞群である. 同一細胞層内における各細胞面は信号の流れから見れば, 互いに並列な関係にある. すなわち同一細胞層に含まれる細胞面内の同じ位置にある細胞は, どの細胞面に属している細胞でも, 入力層上の同じ場所に受容野を持つ.

U_S 細胞層は特徴抽出細胞層で, 下位の層では入力層の局所的な特徴を抽出し, 上位にいくほど大局的な特徴を抽出する. U_C 細胞層は前段の U_S 細胞層で抽出した特徴の相対的な位置ずれを許容するために,U_S 細胞層の出力に対してぼかしをかける役割を持つ.

求心性経路の各段を下位から上位へ信号が流れるにつれ, 特徴の統合, 及び位置ずれに対する許容が繰り返し行なわれる. 求心性経路の特徴抽出細胞の入力結合の強度は, 一部の例外を除き winner-take-all 型の競合学習方式を用いた教師なし学習 [33] によって強化する. 学習終了後, 最上位段の u_{C4} 細胞は認識細胞として働く.

本システムでは求心性経路上に, 折れ点検出回路を新たに導入した. システムに与えられる入力パターンは様々な方位の直線成分から構成されるが, これらの特徴は U_{S1} 層で 8 方位の線分成分に分解される. 折れ点検出回路は, これらの直線成分の端点部分に反応する. したがって, 折れ点検出回路は, 入力パターンでの曲率が著しく大きい位置や直線成分の終端部分で大きな出力を出す. 折れ点検出回路の構成は求心性経路の他の段と同じく, u_{SBend} 細胞からなる層 U_{SBend} と u_{CBend} 細胞からなる層 U_{CBend} という 2 層構造から成っている. U_{SBend} 層は, 直線検出の結果である U_{C1} 層の出力を受けとり, 折れ点を抽出する役割を持つ. U_{CBend} 層は 他の U_C 細胞層と同様に U_{SBend} 層の各細胞の出力を空間的にぼかす役割を持つ. 折れ点検出回路の詳細は 5.5.1 節で述べる.

遠心性経路は, u_{C4} 細胞が認識しているパターンの想起を行う. 遠心性経路には 図 5.2 に示すように, 求心性経路と同様の細胞の配置をしている. すなわち u_S 細胞に対応する細胞には w_S 細胞, u_C 細胞に対応する細胞には w_C 細胞, というように両経路の細胞は, 鏡像関係を保つように配置している. 細胞間の結合係数の分布や強度に関しても鏡像関係が成り立つように設計している. ただし, 各結合を経由して伝えられる信号の流れの方向は, 両経路で互いに逆方向になっている.

求心性経路の認識細胞 u_{C4} が反応すると, その信号は対応する遠心性経路の w_{C4} 細胞に送られる. パターンの想起を行うためには, この w_{C4} 細胞から送られる遠心性信号が, 求心性信号と同じ経路を逆にたどって流れるようにする必要がある. そのために w_S 細

胞は, 対応する位置の u_S 細胞からのゲート信号による制御を受けるようにした [46].

逆に遠心性経路から求心性経路に対しては, u_C 細胞に利得の制御信号を送り “注意の集中” を行なうとともに, 認識したパターンに含まれるべき特徴を特徴抽出細胞 u_S が抽出しやすくするために閾値制御信号を送っている [24].

求心性経路への折れ点検出回路の導入にともなって, 対応する位置に遠心性経路上には新たに折れ点復元回路を導入した. 折れ点復元回路も, 他の細胞段と同じく 2 層構造からなっている. すなわち u_{SBend} 細胞に対応する w_{SBend} 細胞からなる層 W_{SBend} と u_{CBend} 細胞に対応する w_{CBend} 細胞からなる層 W_{CBend} から構成される. 折れ点復元回路は W_{S2} 層からの遠心性信号を受けとり, W_{C1} 層の折れ点のあるべき位置に信号を送る. この詳細については 5.5.2 節で論じる.

これらの折れ点検出回路と折れ点復元回路を合わせて折れ点処理回路と呼ぶことにする.

今川 らのシステムでは, 求心性経路で入力パターンを観測すべき空間をある程度限定するために サーチコントローラ を導入していた [46]. サーチコントローラ は サーチエリア と呼ばれる領域を決定し, サーチエリア 付近の利得を周囲の部分と比べて相対的に大きくしてやることによって信号が周囲より流れやすい状態にする. こうすることによって入力パターンの情報を限定し, 同時に処理すべき情報量を減らしている.

本システムでも 今川 らのシステムと同様に サーチエリア を決定するために図 5.1 , 図 5.2 に示す サーチコントローラ を組み込んでいる. しかしながら, 今川らのシステムとは少し異なり, サーチコントローラ は直線成分だけでなく折れ点の成分をも用いて制御を行っている. 英字筆記体の文字パターンでは, 一般に折れ点成分の密度は, 文字パターン間の連結部分よりも文字部分の方が高い. そこで, このシステムでは注目点が文字部分に来やすくするために, サーチコントローラ で, 直線成分と折れ点成分の空間的密度の高い部位を探し, この部分に サーチエリア の中心 (注目点) がくるように制御している. またサーチコントローラ は, 求心性経路の細胞からだけではなく, 遠心性経路の細胞からも制御信号を受け取る. 遠心性信号からの制御は, ただ一個の文字が想起されているときに有効になる. これらの詳細は 5.6 節で説明する.

さて, 学習後に, 学習パターンに類似したパターンを複数個, 同時に提示した場合を考えよう. 最上位層の認識細胞 u_{C4} が出力を出し, あるパターンが認識されると認識細胞

から遠心性経路に信号が流れる．遠心性経路では, 認識に関与した求心性経路の細胞から対応する位置にある細胞に対してゲート信号が送られるので, 求心性経路と同じ経路を通って逆向きの信号が流れる．その結果, 認識されたパターンが遠心性経路の最下位層 (W_{C0}) に想起される．想起されたパターンは再び入力層に正帰還され, 求心性経路に信号が流される．以下, 後述する “注意の転換” が行なわれるまでは, 信号は求心性経路と遠心性経路とで形成される正帰還ループを流れ続ける．さらに遠心性経路から求心性経路に対しては, 利得制御信号が送られる．利得制御信号によって, 信号の流れている遠心性経路と同じ位置の求心性経路の細胞の利得が高められる (促通作用)．この結果, 求心性経路では認識されたパターンに関与する経路は, それ以外の経路に比べて信号が流れやすい状態になる．こうすることによって注意の集中を行う．

このときパターンの部分的欠損や変形によって一部の特徴が検出されなくても, 従来の選択的注意のモデル [24] の場合と同様, 信号が正帰還ループを流れる間に, 閾値制御機構が働いて想起されるパターンの欠損部分は次第に補間されていく．

最上位層のどの認識細胞にも出力がなく, 何も認識されない場合には, 図 5.1 , 図 5.2 に示した無出力検出機構 (no-response detector) が作動して, 求心性経路内のすべての特徴抽出細胞の閾値を下げるような信号を送りだした上で, 再び認識を試みる．最上位層に出力が出るまで, この操作は繰り返される．

注意の転換は遠心性経路の信号を一瞬間だけ遮断することによって行なう．求心性経路の最上位層 u_{C4} で 1 個の認識細胞だけが出力を出し, ある一つのパターンが認識されたとしよう．回路の反応が定常状態になると自動的に注意を転換して, 入力刺激に含まれる他のパターンを認識するようになっている．すなわち図 5.1 , 図 5.2 に示す Attention Switcher が回路の反応をモニターしており, 定常状態に達したことを検出すると, 遠心性経路へ一瞬間だけ抑制性の信号を送って, 遠心性信号の流れを一時的に遮断する．遠心性経路の信号が遮断されると, 今まで流れていた利得制御信号が消滅し, 促通作用が途絶える．求心性経路の細胞は “疲労” の特性を持っているので, 促通作用が途絶えると, それまで大きな利得制御信号を受けとっていた細胞の利得が下がる．

サーチコントローラ は, この状態で, 次の サーチエリア を探す．利得が下がった部位では細胞の応答が小さくなっているので, そのような場所が サーチエリア に選ばれる確率は低い．サーチエリアが決定されると, 再び認識動作を行ない, 新たなパターンが認識

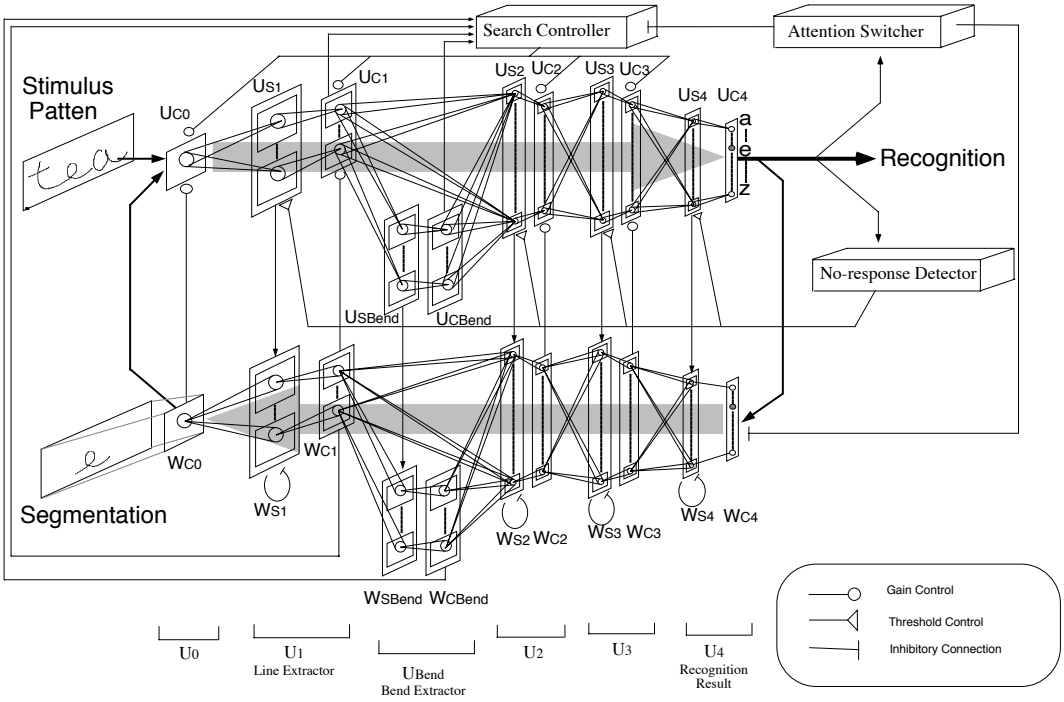

図 5.1: 英字筆記体連結文字列を認識・切り出しを行うシステムの全体図.

されると, そのパターンの想起を行ない, 回路の反応が安定したら注意の転換を行なうという操作を繰り返すことで, 個々のパターンを順次認識し, 切り出していくことができる.

認識の終了条件は, 入力層 U_{C0} の発火率が一定レベル以下になった時である.

以下では, 各層の細胞の反応のうち求心性経路と遠心性経路について, もう少し詳細に説明する. 5.3 節では求心性経路の細胞について, 5.5.1 節では折れ点検出回路について述べている. 遠心性経路の細胞に関しては 5.4 節で述べられており, 折れ点復元回路, については, 5.5.2 節で説明している.

5.3 求心性経路の細胞

ネオコグニトロンの S 細胞と同様に u_S 細胞は局所的な特徴抽出を行なう. 第 l 段目の u_S 細胞 u_{Sl} は, まず側抑制の効果を無視して u_S 細胞の一時的な出力 $\tilde{u}^t_{Sl}$ を求めておき, この出力に対して側抑制を 1 回かけ, その結果を再び細胞の入力と考えるという

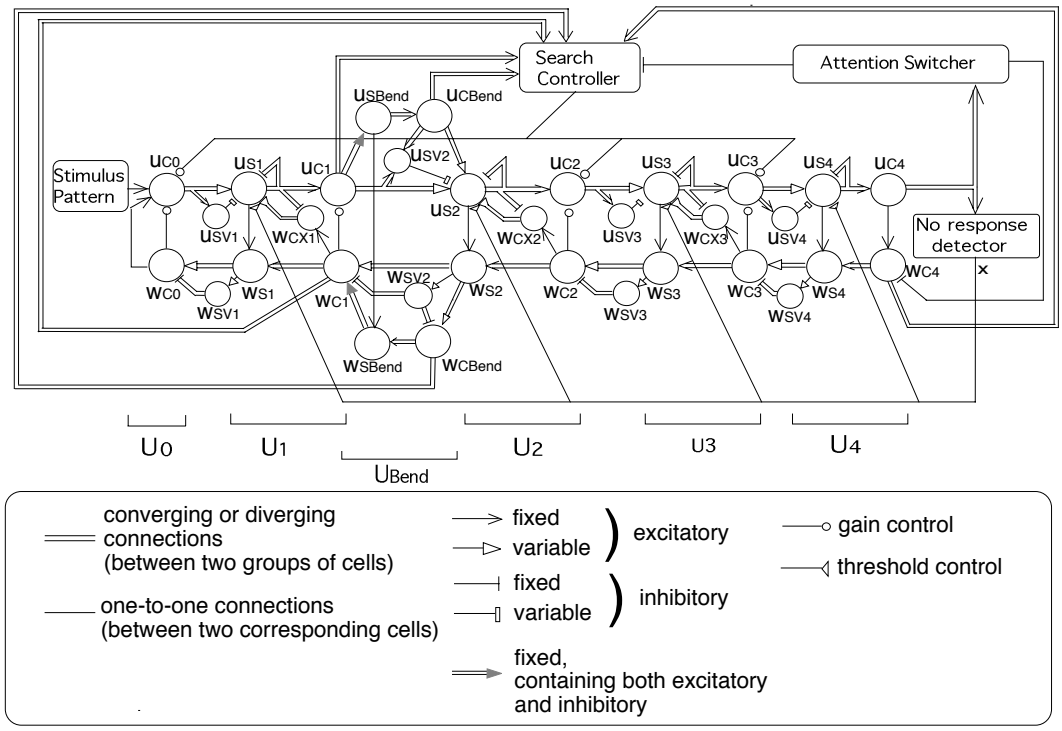

図 5.2: 図はシステムに含まれる異なる細胞間の結合を示す. システムには同種の細胞が複数含まれるが簡単のため一つのシンボルで代表させている. 図中の矢印が結合を表す.

操作を有限回繰り返す. 側抑制をかける理由は, 同一受容野をもつ細胞間で競合をさせることによって, ある位置に存在する特徴が単一であるようにするためのメカニズムである.

$$\tilde{u}^t_{Sl}(\boldsymbol{n},k) = \frac{\theta^t_l(\boldsymbol{n},k)}{1-\theta^t_l(\boldsymbol{n},k)} \cdot \varphi \left[\frac{1 + \sum_{\kappa=1}^{K_{l-1}} \sum_{\boldsymbol{\nu}\in A_l} a_l(\boldsymbol{\nu},\kappa,k) \cdot u^t_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu},\kappa)}{1+\theta^t_l(\boldsymbol{n},k)\cdot b_l(k)\cdot u^t_{SVl}(\boldsymbol{n})} - 1 \right] \tag{5.1}$$

ただし $l = 2$ は折れ点検出回路からの出力を考慮する必要があるので $l \neq 2$ とする. $a_l(\boldsymbol{\nu},\kappa,k)$ は, 第 $l-1$ 段目の U_{Cl-1} 細胞層の κ 番目の細胞面の細胞 $u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu},\kappa)$ か

ら, 第 l 段目の U_{Sl} 細胞層 k 番目の細胞面の細胞 $u_{Sl}(\boldsymbol{n},k)$ への興奮性可変結合の強度を表す. $b_l(k)$ は抑制性細胞 $u_{SVl}(\boldsymbol{n})$ からの抑制性可変結合の強度を表す. $\varphi[\cdot]$ は,

$$\varphi[x] = \max[x, 0] \tag{5.2}$$

で定義される関数で半波整流を行なう. $\theta_l^t(\boldsymbol{n},k)$ は, 時刻 t における閾値の制御変数である. u_{Sl} 細胞の最終的な出力値 $u_{Sl}^t(\boldsymbol{n},k)$ は, 前述の $\tilde{u}_{Sl}^t(\boldsymbol{n},k)$ に対して側抑制をかけたものである. $u_{Sl}^t := \tilde{u}_{Sl}^t$ として次式による再帰計算を行なう.

$$u_{Sl}^t(\boldsymbol{n},k) := \varphi\left[u_{Sl}^t(\boldsymbol{n},k) - \sum_{\boldsymbol{\nu}\in E_l} e_l(\boldsymbol{\nu})\cdot u_{Sl}^t(\boldsymbol{n}+\boldsymbol{\nu},k) - \sum_{\substack{\kappa=1\\ \kappa\neq k}}^{K_l}\sum_{\boldsymbol{\nu}\in E_l} e_l^{'}(\boldsymbol{\nu})\cdot u_{Sl}^t(\boldsymbol{n}+\boldsymbol{\nu},\kappa)\right] \tag{5.3}$$

ただし記号 “:=” は右辺の値を左辺の変数へ代入する演算記号とする. $e_l(\boldsymbol{\nu})$ と $e_l^{'}(\boldsymbol{\nu})$ は $|\boldsymbol{\nu}|$ にのみ依存する関数であり側抑制の空間結合係数の形状である.

側抑制をかけるのは, 同一受容野をもつ特徴抽出細胞間で競合をかけることによって, ある単一の特徴が抽出されるようにするためのメカニズムである. 側抑制を用いることによって求心性経路にはリカレントな回路が形成されるが, シミュレーションでは計算時間を節約するため, まず u_S 細胞が前段の細胞から受け取る信号 (すなわち側抑制の効果を無視した u_S 細胞の出力) $\tilde{u}_{Sl}^t$ を求め, この $\tilde{u}_{Sl}^t$ から u_{Sl} 細胞の最終出力を式 (5.3) によって計算する.

第 l 段目の細胞 u_{SVl} 細胞の出力は, a_l の結合範囲と同じ範囲にある $l-1$ 段目の u_{Cl-1} 細胞の出力の重み付き 2 乗平均に相当し,

$$u_{SVl}^t(\boldsymbol{n}) = \sqrt{\sum_{\kappa=1}^{K_{l-1}}\sum_{\boldsymbol{\nu}\in A_l} c_l(\boldsymbol{\nu})\cdot\{u_{Cl-1}^t(\boldsymbol{n}+\boldsymbol{\nu})\}^2} \qquad (l\neq 2) \tag{5.4}$$

という出力を出す. この u_{SV} 細胞は抑制性細胞で, u_S 細胞に対し抑制出力を行う. $c_l(\boldsymbol{\nu})$ は固定結合の強度を表し, $|\boldsymbol{\nu}|$ に対して単調減少の結合である.

相対的位置ずれを吸収する役目を担う u_C 細胞の出力は, u_S 細胞の出力に対して, 前層の u_S 細胞の出力を空間的にぼかし, その結果に利得変数 $g_l(\boldsymbol{n},k)$ を掛けたものになっ

ている. すなわち

$$u_{Cl}^t(\boldsymbol{n},k) = g_l^t(\boldsymbol{n},k)\cdot\psi\left[\sum_{\boldsymbol{\nu}\in D_l} d_l(\boldsymbol{\nu})\cdot u_{Sl}^t(\boldsymbol{n}+\boldsymbol{\nu},k)\right] \qquad (l\neq 0) \tag{5.5}$$

という出力を出す. $\psi[\cdot]$ は出力の飽和特性を定める関数で,

$$\psi[x] = \frac{\varphi\,[x]}{1+\varphi\,[x]} \tag{5.6}$$

と定義しており, 関数 $\varphi\,[\cdot]$ は, 式 (5.2) で定義された関数である.

U_C 細胞層のうち 入力層 U_{C0} 層は外部からの入力のみならず, 遠心性経路の最下位層 W_{C0} からのフィードバック信号によっても制御される. u_{C0} 細胞は

$$u_{C0}^t(\boldsymbol{n}) = g_0^t(\boldsymbol{n})\cdot\max[p(\boldsymbol{n}), w_{C0}^{t-1}(\boldsymbol{n})] \tag{5.7}$$

という出力を出す. ここに, $p(\boldsymbol{n})$ は外部から入力層 U_{C0} 層へ提示された入力パターンである.

5.4 遠心性経路における細胞

遠心性経路の細胞 w_{Sl} 細胞, w_{Cl} 細胞, および w_{SVl} 細胞の特性は, 今川 が述べたものと同様のものにしている [46]. ある w_S 細胞から, その下段に位置する w_C 細胞への結合強度は, 対応する位置の u_S 細胞に向かってくる求心性の結合強度に比例した値になるように学習の際に連動して調節されるようにした. 従って w_S 細胞から下向する遠心性信号は, 下段の u_C 細胞から u_S 細胞への信号の流れと同じ経路を逆方向に流れることになる. 同様に抑制性細胞の w_{SV} 細胞から下向する信号も, 対応する位置に存在する u_{SV} 細胞への求心性経路を逆にたどることになる. w_C 細胞の出力は,

$$\begin{aligned} w_{Cl}^t(\boldsymbol{n},k) \;=\; \psi\Bigg[\alpha_l\cdot\Bigg\{&\sum_{\kappa=1}^{K_{l+1}}\sum_{\boldsymbol{\nu}\in A_{l+1}} a_{l+1}(\boldsymbol{\nu},k,\kappa)\cdot w_{Sl+1}^t(\boldsymbol{n}-\boldsymbol{\nu},\kappa) \\ &-\sum_{\boldsymbol{\nu}\in A_{l+1}} c_l(\boldsymbol{\nu})\cdot w_{SVl+1}^t(\boldsymbol{n}-\boldsymbol{\nu})\Bigg\}\Bigg] \end{aligned} \tag{5.8}$$

としている. ただし $l=1$ では 折れ点復元を考慮する必要があるので $(l \neq 1)$ とする. α_l は飽和のレベルを定める定数である.

抑制性細胞 w_{SV} は求心性経路の u_{SV} 細胞とは異なり, 線形の入出力特性を持っていて,

$$w^t_{SVl+1}(\boldsymbol{n}) = \theta^0_{l+1} \cdot \sum_{\kappa=1}^{K_{l+1}} b_{l+1}(\kappa) \cdot w^t_{Sl+1}(\boldsymbol{n}, \kappa) \quad (l \neq 1) \tag{5.9}$$

という出力を出す. ただし θ^0_l は閾値の制御変数 $\theta^t_l(\boldsymbol{n}, k)$ の初期値である.

w_S 細胞は

$$w^t_{Sl}(\boldsymbol{n}, k) = \min\left[u^t_{Sl}(\boldsymbol{n}, k), \alpha'_l \cdot \sum_{\boldsymbol{\nu} \in D_l} d_l(\boldsymbol{\nu}) \cdot w^t_{Cl}(\boldsymbol{n} - \boldsymbol{\nu}, k)\right] \tag{5.10}$$

という出力を出す. ここに α'_l は正の定数である. したがって, w^t_{Sl} 細胞の出力は, ちょうど求心性の信号と遠心性の信号の論理積のようなものになっている. 求心性信号がないところは, 発火しないようなゲートの作用を持っている.

5.5 折れ点処理回路

本節では, 新たに導入した折れ点処理回路について詳しく述べる. 折れ点処理回路は, 求心性経路上での折れ点の検出を行う回路と, 遠心性経路上での折れ点の復元を行う回路という二つの回路から成り立つ [93], [94], [96], [95], [97] .

5.5.1 折れ点検出回路

局所的な曲率の急激な変化という意味で折れ点および端点は, パターン認識において重要な特徴であると考えられる. これらの適切な利用は, ネオコグニトロン型ネットワークの能力の向上につながることが示されている [33]. このような経緯から, 求心性経路にネオコグニトロン型のネットワークを持つ本システムに折れ点検出回路を導入した.

本システムに与えられるパターンは, 様々な曲線要素からなるが, これらのパターンは, 第 1 段目で様々な方位の直線成分に分離される. 本研究では, これらの分離された直線成分の両端部分を折れ点と考えている. このような各直線成分の端点を抽出することによって, 入力パターン上で曲率の大きな部分を抽出できる.

直線検出層 U_{C1} の k 番目の細胞面は, 最適方位

$$\alpha = \frac{(k-1)\pi}{K_{C1}} \tag{5.11}$$

の直線成分を抽出する. ここに K_{C1} は U_{C1} 層の細胞面の個数である.

折れ点検出回路は, U_{SBend} 層, 及び U_{CBend} 層から成り, 求心性経路の第 1 段目 (直線検出段) と第 2 段目との間に取り付けられている. 折れ点検出層 U_{SBend} の細胞は, 結合 $a_{Bend}(\boldsymbol{\nu}, k)$ を介して, U_{C1} 層の, ある一つの細胞面上の異なった位置にある 2 群の細胞の出力を受けとり, それらの差をとるような出力を出す (図 5.3).

結合 $a_{Bend}(\boldsymbol{\nu}, k)$ は U_{C1} 層の k 番目の細胞の最適方位の方向に小さな興奮性領域と大きな抑制性領域とを並べた様な形状であり, $\boldsymbol{\nu}$ は結合の中心からの位置である. このような結合を用いると, 興奮性領域に大きな入力があった時 U_{SBend} 層で出力が出る (図 5.3 -(a)(b)). 逆に抑制性領域に入力があると U_{SBend} 層の出力が抑制される (図 5.3 -(c)).

結合 $a_{Bend}(\boldsymbol{\nu}, k)$ は, 最適方位 α を持つ直線成分の端点の一方を検出する [93], [94], [96], [95], [97] . この直線成分のもう一方の端点を検出するためには, 図 5.3 に示す結合の興奮性領域と抑制性領域とを入れ替えたような結合 $a_{Bend}(\boldsymbol{\nu}, k + K_{C1})$ を用いればよい. このように, 折れ点検出段は一つの直線成分の両端を検出するので, U_{SBend} 層と U_{CBend} 層の細胞面の数は, U_{C1} 層の細胞面の個数の 2 倍, すなわち $2K_{C1}$ 個としている. 数式で表現すれば u_{SBend} 細胞は,

$$u^t_{SBend}(\boldsymbol{n}, k) = \varphi\left[\sum_{\boldsymbol{\nu} \in D_{SBend}} a_{Bend}(\boldsymbol{\nu}, k) \cdot u^t_{C1}(\boldsymbol{n} + \boldsymbol{\nu}, \kappa)\right], \qquad \kappa = k \bmod K_{C1}$$

という出力を出す. ただし t は, 他の u_S 細胞の t と同じく認識開始時を $t = 0$ とした離散計算時刻を表わす. $\psi[\cdot]$ は式 (5.6) で定義されたものを用いる.

U_{CBend} 層は, 他の段の U_C 層 (5.3 節参照) と同様に U_{SBend} 層で抽出した特徴の位置に対して, ぼかしをかける役割を持つ. ぼかす小さな領域を D_{Bend} とし, $d_{Bend}(\boldsymbol{\nu})$ を $|\boldsymbol{\nu}|$ に関する単調減少の関数とすると, u_{CBend} 細胞は,

$$u^t_{CBend}(\boldsymbol{n}, k) = \psi\left[\sum_{\boldsymbol{\nu} \in D_{Bend}} d_{Bend}(\boldsymbol{\nu}) \cdot u^t_{SBend}(\boldsymbol{n} + \boldsymbol{\nu}, k)\right] \tag{5.12}$$

という出力を出す.

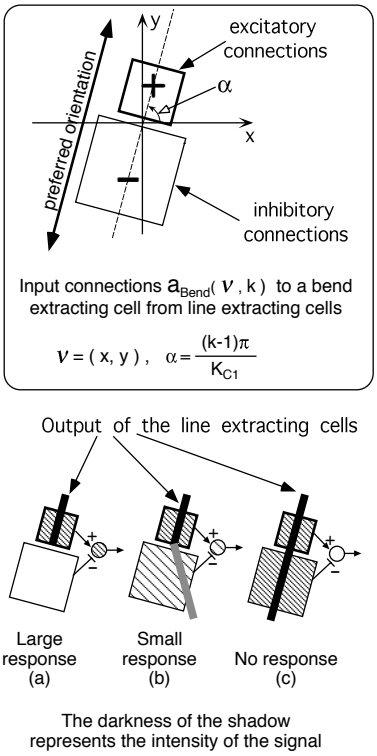

図 5.3: 折れ点検出を行う細胞への入力結合

折れ点検出段の次の段の特徴抽出細胞 u_{S2} は, 直線検出段の出力細胞 u_{C1} と折れ点検出段の出力細胞 $u_{C\,Bend}$ からの信号を受ける. この u_{S2} 細胞が前段の細胞から受け取る信号は次式によって計算される.

$$\tilde{u}^t_{S2}(\boldsymbol{n},k) = \frac{\theta^t_2(\boldsymbol{n},k)}{1-\theta^t_2(\boldsymbol{n},k)} \cdot \varphi\left[\frac{1+I^t_{S_1}(\boldsymbol{n},k)+I^t_{S\,Bend}(\boldsymbol{n},k)}{1+\theta^t_2(\boldsymbol{n},k)\cdot b_2(k)\cdot u^t_{SV2}(\boldsymbol{n})}-1\right] \tag{5.13}$$

ただし $I^t_{S_1}(\boldsymbol{n},k)$ と $I^t_{S\,Bend}(\boldsymbol{n},k)$ は, それぞれ,

$$I^t_{S_1}(\boldsymbol{n},k) = \sum_{\kappa=1}^{K_1}\sum_{\boldsymbol{\nu}\in A_2} a_2(\boldsymbol{\nu},\kappa,k)\cdot u^t_{C1}(\boldsymbol{n}+\boldsymbol{\nu},\kappa) \tag{5.14}$$

および

$$I^t_{S\,Bend}(\boldsymbol{n},k) = \sum_{\kappa=1}^{K_{Bend}}\sum_{\boldsymbol{\nu}\in A_2} a_{2'}(\boldsymbol{\nu},\kappa,k)\cdot u^t_{C\,Bend}(\boldsymbol{n}+\boldsymbol{\nu},\kappa) \tag{5.15}$$

を表す. $a_2(\boldsymbol{\nu},\kappa,k)$ は, 直線検出層 U_{C1} の κ 番目の細胞面から U_{S2} 層の k 番目の細胞面への結合を表す. 同様に $a_{2'}(\boldsymbol{\nu},\kappa,k)$ は, 折れ点検出層 U_{CBend} の κ 番目の細胞面から U_{S2} 層の k 番目の細胞面への結合を表す. これらの結合 $a_2(\boldsymbol{\nu},\kappa,k)$ および $a_{2'}(\boldsymbol{\nu},\kappa,k)$ は, 他の段と同様に学習 [24][33] によって決定される. 式 (5.13) の $u^t_{SV2}(\boldsymbol{n})$ は, u_{SV2} 細胞の出力で,

$$u^t_{SV2}(\boldsymbol{n}) = \sqrt{I^t_{SV1}(\boldsymbol{n}) + I^t_{SVBend}(\boldsymbol{n})} \tag{5.16}$$

という出力を出す. ただし $I^t_{SV_1}(\boldsymbol{n})$ と $I^t_{SVBend}(\boldsymbol{n})$ はそれぞれ,

$$I^t_{SV_1}(\boldsymbol{n}) = \sum_{\kappa=1}^{K_{C1}} \sum_{\boldsymbol{\nu}} c_2(\boldsymbol{\nu}) \cdot u^t_{C1}(\boldsymbol{n}+\boldsymbol{\nu})^2 \tag{5.17}$$

$$I^t_{SVBend}(\boldsymbol{n}) = \sum_{\kappa=1}^{2K_{C1}} \sum_{\boldsymbol{\nu}} c_2^{'}(\boldsymbol{\nu}) \cdot u^t_{CBend}(\boldsymbol{n}+\boldsymbol{\nu})^2 \tag{5.18}$$

を表す. なお結合係数 $c_2(\boldsymbol{\nu})$, $c_{2'}(\boldsymbol{\nu})$ は, 共に $|\boldsymbol{\nu}|$ に関する単調減少の結合係数である. こうして一時的に求められた $\tilde{u}^t_{S2}(\boldsymbol{n},\kappa,k)$ に対して式 (5.3) による再帰計算を行い, 細胞の最終的な出力を求める.

5.5.2 折れ点復元回路

本システムでは, 求心性経路の折れ点検出回路と鏡像関係にある折れ点復元回路を遠心性経路の 1 段目と 2 段目の間に新たに導入した. 折れ点復元回路は W_{S2} 層からの遠心性信号を受けて折れ点成分のあるべき場所を推定する. 折れ点復元回路は, 他の細胞段と同様に 2 層構造からなる. すなわち U_{SBend} 層, と U_{CBend} 層にそれぞれ対応する W_{SBend} 層と W_{CBend} 層である. W_{CBend} 層の細胞 w_{CBend} は,

$$\begin{aligned} w^t_{CBend}(\boldsymbol{n},k) = \psi\Bigg[\alpha_{Bend} \cdot \Bigg\{ & \sum_{\kappa=1}^{K_2} \sum_{\boldsymbol{\nu}\in A_2} a_{2'}(\boldsymbol{\nu},\kappa,k) \cdot w^t_{S2}(\boldsymbol{n}-\boldsymbol{\nu},\kappa) \\ & - \sum_{\boldsymbol{\nu}\in A_2} c_{2'}(\boldsymbol{\nu}) \cdot w^t_{SV2}(\boldsymbol{n}-\boldsymbol{\nu}) \Bigg\}\Bigg] \end{aligned} \tag{5.19}$$

という出力を出す. ここに α_{Bend} は飽和のレベルを定める定数である. なお $w_{SV2}(\boldsymbol{n})$ 細胞の出力 $w^t_{SV2}(\boldsymbol{n})$ は, W_{S2} 層における各細胞面の $\boldsymbol{n}$ に位置する細胞の出力を重み付き

で足し合わせたようなもので, 式 (5.9) で与えられる. $w_{SV2}(\boldsymbol{n})$ 細胞の出力は, 重み付けした線形和に比例する. W_{SBend} 層の細胞は,

$$w^t_{SBend}(\boldsymbol{n},k) = \min\left[u^t_{SBend}(\boldsymbol{n},k),\ \alpha'_{Bend}\cdot\sum_{\boldsymbol{\nu}\in D_{Bend}} d_{Bend}(\boldsymbol{\nu})\cdot w^t_{CBend}(\boldsymbol{n}-\boldsymbol{\nu},k)\right] \quad (5.20)$$

という出力を出す.

ただし α^t_{Bend} は正の定数, $d_{Bend}(\boldsymbol{\nu})$ は式 (5.12) で用いたぼかしをかける結合である. w_{SBend} 細胞は他の段の w_S 細胞と同様に w_C 細胞の出力に対して $|\boldsymbol{\nu}|$ に関する単調減少な結合を介した入力と, 対応する位置の u_S 細胞からゲート信号を受け取る. w_C 細胞の出力をぼかしたような入力のみでは下位の層に遠心性信号が伝わるにつれ, 信号が次第にぼけて伝わることになる. これを避けるために, w_S 細胞は対応する位置の u_S 細胞からゲート信号を受け取るようにしている. w_S 細胞は u_S 細胞と w_C 細胞の両方から入力を受け取ったときのみ出力を出すようになっている.

W_{C1} 層は, 2 種類の細胞 w_{S2} と w_{SBend} からの入力を受け取っている. W_{SBend} 層の各細胞面からは, 求心性経路の折れ点検出に用いた結合 a_{Bend} を経由するフィードバック信号を受け取る (図 5.4). これは, W_{S2} 層から W_{C1} 層に対して結合 a_2 を経由して復元された直線成分の端点の位置を確定する役割を果たしている. 数式で表現すると

$$\begin{aligned} w^t_{C1}(\boldsymbol{n},k) \;=\; \psi\Bigg[\alpha_1\cdot\Bigg\{&\sum_{\kappa=1}^{K_2}\sum_{\boldsymbol{\nu}\in A_2} a_2(\boldsymbol{\nu},k,\kappa)\cdot w^t_{S2}(\boldsymbol{n}-\boldsymbol{\nu},\kappa) \\ &-\sum_{\boldsymbol{\nu}\in A_2} c_2(\boldsymbol{\nu})\cdot w^t_{SV2}(\boldsymbol{n}-\boldsymbol{\nu}) \\ &+\sum_{\boldsymbol{\nu}\in A_{Bend}} a_{Bend}(\boldsymbol{n}-\boldsymbol{\nu},k)\cdot w^t_{SBend}(\boldsymbol{n}-\boldsymbol{\nu},k) \\ &+\sum_{\boldsymbol{\nu}\in A_{Bend}} a_{Bend}(\boldsymbol{n}-\boldsymbol{\nu},k+K_{C1})\cdot w^t_{SBend}(\boldsymbol{n}-\boldsymbol{\nu},k+K_{C1})\Bigg\}\Bigg] \end{aligned} \quad (5.21)$$

としている. 式 (5.21) の 第 1 項と第 2 項とは, W_{C1} 層の k 番目の細胞面に最適方位を持った直線成分の復元に関与する項である. 第 3 項と第 4 項とは折れ点部分があるべき場所に対して, 図 5.4 で示す様に直線成分の端点部分に興奮性の信号を送り, その外側には抑制をかける項である. 興奮性の信号と抑制性の信号とを加えてやることによって直

線成分の端点部分のコントラストがはっきりするようになる． これは, ある直線成分の復元を行なう際に, どこまでも際限なく線が伸びていかないような結合 MRF 的な作用を実現する仕組みとなる [34].

5.5.3 折れ点処理機構を導入する利点

折れ点処理回路を用いた利点を, 図 5.5 を例に取って示す. まずパターン ‘c’ を学習する場合を考える (図 5.5 (a)(b)). 今川らのシステムでは第 2 層は直線成分の学習を行う (図 5.5 (b)). 折れ点検出ありの場合 (図 5.5 (a)) には, 直線成分以外にも折れ点の部位を学習する.

次に折れ点処理機構のない今川らのシステムにおいて, 入力パターン “ct” の ‘c’ の部分を W_{C0} 層に切り出す場合を考える. U_{C0} 層に与えられたパターンは ‘c’ と ‘t’ 連結部分の曲線が 図 5.5 (d) のように直線に近い. このような場合は, 直線成分の復元を行う W_{C1} 層において, この直線成分の端点の位置を決定することができない. その結果, W_{C0} 層に想起されるパターンにおいて ‘c’ と ‘t’ との連結部分の境界が分からず, 求心性信号と遠心性信号とが正帰還ループを回る間に徐々に文字 ‘t’ の一部分を ‘c’ の部分パターンとして切り取っていってしまう場合があった. そのため注意の転換後, ‘t’ の部分のゲインが低下してしまい, 認識に失敗する場合があった [93], [94], [96], [95], [97] .

次に折れ点処理機構のある本システムを用いて, 同様の切り出しを行う場合を考えてみる (図 5.5 (c)). 学習段階では ‘c’ というパターンを構成する各方位の直線成分だけでなく, 折れ点の存在する位置をも学習する (図 5.5 (a)). (図 5.5 (a) では, 特に重要な折れ点を太線の円で示している). 折れ点成分がパターンのどこにあるかをシステムは学習しているので, 文字パターン間の適当な連結部分で折れ点復元回路が働こうとする. 図 5.5 (c) の場合, 文字パターン ‘c’ と ‘t’ との連結部分にわずかでも曲がりがあれば, 折れ点の復元回路が働き, 連結部分のパターンの切り出しはそこで終わる． また, 連結部分に曲がりがなくても, 隣接文字パターンとの合流部分 (図 5.5 (c) の場合は, ‘t’ の付け根の部分) などでは, 方位の異なる複数の直線成分が受容野の中に与えられることになるので, 直線抽出細胞である u_{S1} の検出力が, その周辺部よりも落ちることになる. したがって折れ点検出回路が, この接続部分で出力を出すことになる. つまり文字と文字との連結

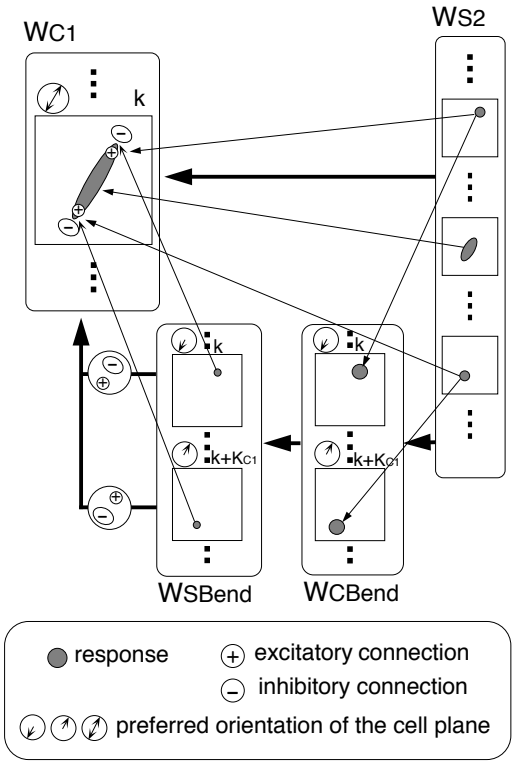

図 5.4: 遠心性経路中の折れ点復元回路による端点や折れ点の推定. 求心性経路と対になっていて, w_{SBend} 層の細胞によって直線成分の端点部分に抑制がかけられる.

部分が確定できるので折れ点復元回路が働き, 最悪の場合でもこの合流部分で切り出しが止まることになる. 従って隣接パターン 't' の一部を 'c' とみなして切り取ってしまうことが少ない.

これは, 直線成分の伸長を適当なところで中断するので Geman らの提案したラインプロセスの一種であると考えることができる. これらの折れ点復元回路の結合は, Geman らが MRF を用いて画像復元を行った際の "turn" や "ending" の線素にあたるものと対応する. [34].

このように折れ点復元回路を用いてシステムを構成すると, 入力パターンに含まれる, ある文字パターンが隣接する文字パターンの一部分を切り出すことが少なくなる. したがって認識・切り出しの効率を上げることができる.

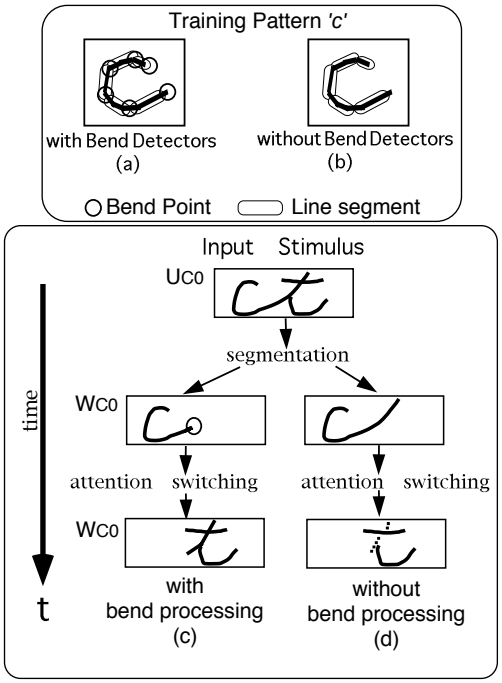

図 5.5: 折れ点復元回路の効果. 折れ点復元回路によって他のパターンを侵食することなく切り出しが行なわれる. 折れ点処理回路がある場合, 学習時には (a) のように 'c' の端点部分も学習される. これにより端点部分に抑制がかかるので, (c) に示すように, 他のパターン ('t') の一部を切り出すことなく 'c' を切り出すことが可能である.

5.6 利得の制御と閾値の制御

求心性経路の u_C 細胞の応答特性を示す式 (5.5) には, 利得変数 $g_l^t(\boldsymbol{n}, k)$ が設けられている. これは, 細胞の応答の強度を制御するための変数で, 遠心性経路 によって制御される成分 $g_{Bl}^t(\boldsymbol{n}, k)$ と, サーチコントローラ からの信号によって制御される成分 $g_{Sl}^t(\boldsymbol{n})$ の 2 つに分けられる. 両者は u_C 細胞に対して次のように積の形で利得を制御する.

$$g_l^t(\boldsymbol{n}, k) = g_{Bl}^t(\boldsymbol{n}, k) \cdot g_{Sl}^t(\boldsymbol{n}) \tag{5.22}$$

遠心性経路による利得は, 現在認識している, パターンの持っているべき特徴の位置に高い利得がかけられるので, 結果としてパターンに含まれる雑音成分の除去を行なうことになる. サーチコントローラ による利得はシステムが現在認識している, おおよその領域を決定し, その周辺の利得を高くしている. サーチコントローラ に関しては, 5.6.2 節で述べる. 一方, 認識が成功した場合でも, 学習パターンが持っていた全ての特徴が検出されているわけではない. これはパターンの変形や雑音の影響などが原因として考えられる. この場合, 検出されなかった特徴を検出してやるためには, その特徴があるべき位置の u_S 細胞の閾値を下げてやる必要がある. これは 5.6.3 節で述べる.

5.6.1 遠心性経路による利得の制御

u_C 細胞に対する利得のうち遠心性経路によって制御される利得

$$g^t_{Bl}(\boldsymbol{n}, k) \ (> 0) \tag{5.23}$$

は, 促通作用をもたらす成分

$$g^t_{B1l}(\boldsymbol{n}, k) \ (\geq 0) \tag{5.24}$$

と, 疲労の作用に対応する成分

$$g^t_{B2l}(\boldsymbol{n}, k) \ (\geq 0) \tag{5.25}$$

から成る. すなわち,

$$g^t_{Bl}(\boldsymbol{n}, k) = 1.0 + \alpha_{B1l} \cdot g^t_{B1l}(\boldsymbol{n}, k) - g^t_{B2l}(\boldsymbol{n}, k) \tag{5.26}$$

と表される.α_{B1l} は $g^t_{B1l}(\boldsymbol{n}, k)$ の強度を定める 1 より大きい定数である. 更に g_{B1l} と g_{B2l} は

(i) $w^t_{Cl}(\boldsymbol{n}, k) > 0$ のとき

$$g^t_{B1l}(\boldsymbol{n}, k) = \gamma_l \cdot g^{t-1}_{B1l}(\boldsymbol{n}, k) + (1.0 - \gamma_l) \cdot w^{t-1}_{Cl}(\boldsymbol{n}, k) \tag{5.27}$$

$$g^t_{B2l}(\boldsymbol{n}, k) = \gamma_l \cdot g^{t-1}_{B2l}(\boldsymbol{n}, k) + (1.0 - \gamma_l) \cdot w^{t-1}_{Cl}(\boldsymbol{n}, k) \tag{5.28}$$

(ii) $w^t_{Cl}(\boldsymbol{n},k)=0$ のとき

$$g^t_{B1l}(\boldsymbol{n},k)=\gamma_{1l}\cdot g^{t-1}_{B1l}(\boldsymbol{n},k) \tag{5.29}$$

$$g^t_{B2l}(\boldsymbol{n},k)=\gamma_{2l}\cdot g^{t-1}_{B2l}(\boldsymbol{n},k) \tag{5.30}$$

と変化する. ここに γ_l, γ_{1l}, γ_{2l} は 0 以上 1 以下の変化定数である. g_{B1l} と g_{B2l} の初期値は 0.0 と仮定して, 遠心性信号が少しでも存在すれば, 同一の変化定数 $(1.0-\gamma_l)$ (シミュレーションでは $\gamma_l=0.1$) で遠心性信号の値に近づいていく. 遠心性信号が切れると, g_{B1l} と g_{B2l} は, それぞれ異なる変化定数で減少する. γ_{1l} は 0.0 に,γ_{2l} は 1.0 に近い値に定めている. 従って, 注意の転換によって遠心性経路の信号が遮断された時には, 促通作用をもたらす g^t_{B1l} は急激に減少し, 疲労の作用をもたらす g^t_{B2l} の効果だけが残る. また, 同じカテゴリーに属するパターンが続けて認識されにくくなることを避けるために, 上位の層に進むに従って γ_{1l} の値は小さく, γ_{2l} 値は大きくしている. このようにして, 認識されたパターンは g^t_{B1l} の効果によって, 順次抑制されていくので, 呈示された各パターンに対して順番に注意を向けていくことが可能となる. また疲労で細胞の出力が下がっていても式 (5.26), (5.27),(5.28) の様にゲインを得る式が和の形で示されているので,u_C 細胞に疲労効果でゲインが下がっていてもこのゲインコントロール制御信号がくれば u_C 細胞を再び発火させ易くすることが出来る.

5.6.2 サーチコントローラ による利得制御

本節では, 遠心性経路の信号も用いるようにしたサーチコントローラ について述べる. 今川らのシステムにおいては, サーチコントローラ はU_{C1} 層をモニターしており, 出力の発火密度が最も高い位置を, システムが注目すべき領域 (サーチエリア) の中心 (注目点 μ と呼ぶ) として定めていた. つまり, 今川らのサーチコントローラは求心性経路の情報のみを用いており, 遠心性経路の情報は用いていないことになる. この場合, 認識結果の情報は用いていないのでサーチコントローラの示すパターンの位置と遠心性経路の出力である W_{C0} 層にあらわれるパターンの位置との間に整合性がとれない場合が存在した. そこで本システムではサーチコントローラは, 主にバックワードの信号を用いるようにした. その結果として, 注目点 μ の値は単語中のある 1 文字の認識・切り出しの操

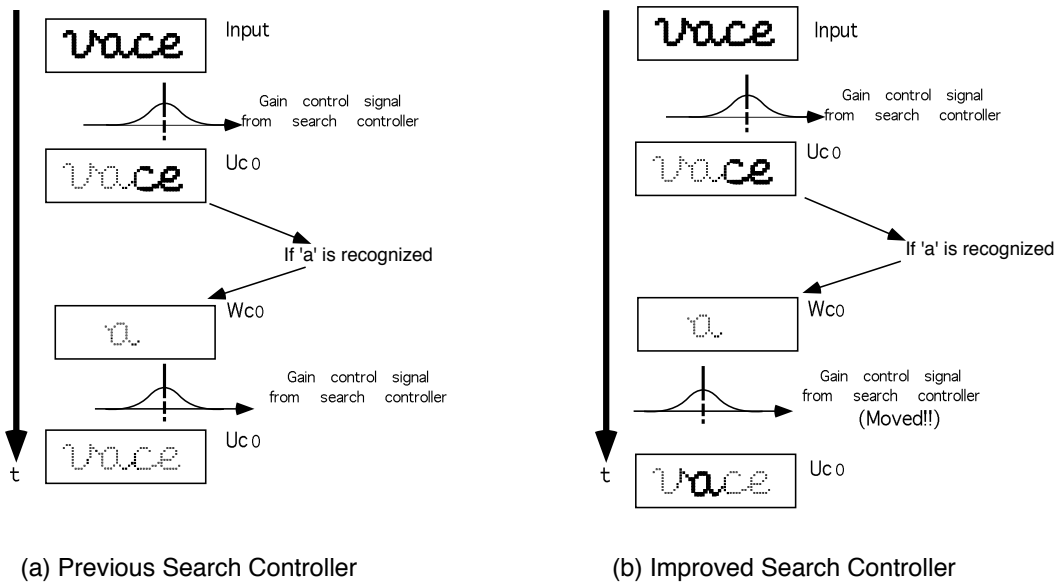

図 5.6: 遠心性信号によるサーチコントローラの制御.

作中ずっと固定されず, 認識結果にに依存して動くようになる. 求心性信号のみを用いて決定した サーチエリア と, 想起された文字とが位置的にずれている場合, この文字パターンに対する利得が不均一になる. この時, 想起した文字が U_{C0} 層に正帰還され, 認識動作が行なわれると, 文字の一部分の利得が小さくなるために認識に失敗することが多い. しかしながら本システムでは, サーチエリア の決定に, 遠心性信号を用いているので, 想起された文字と サーチエリア との位置的なずれが少ない. したがって正帰還後の認識動作でも認識に失敗することが少ない (図 5.6 参照).

W_{C0} 層に文字が想起されない間, または複数の文字が想起されている間, サーチコントローラ は, 求心性経路の U_{C1} 層と U_{CBend} 層とをモニターしている. しかし W_{C0} 層に一つの文字パターンだけが想起された時点で, サーチコントローラ は, モニターする細胞層を遠心性経路の W_{C1} と W_{CBend} とに切替える. すなわち W_{C4} 細胞層 (認識細胞層に対応する遠心性経路の細胞層) の中でただ一つの細胞が発火している場合には, 注目点 μ は遠心性経路の W_{C1} 層と W_{CBend} 層との出力を用いて決定される. この結果 W_{C1} 層 と W_{CBend} 層との出力で発火密度が最も高いところに サーチエリア がくる.

上述の通りサーチコントローラへの入力は, U_{C1} と U_{CBend} との組合わせか, W_{C1} と W_{CBend} との組合わせになっている. 英文筆記体連結文字列を構成する個々の文字パターンでは, 通常の場合, 折れ点の成分が文字間の連結部分より文字部分に多く存在する. したがって文字間の連結部分より文字部分の方が注目点として選ばれやすくなる. その結果, 連結部分の空間パターンを文字として誤認識することが少なくなる. このように, 直線成分だけでなく折れ点成分の発火密度の最も高い場所が注目点 μ として選択されやすくなる.

5.6.3 細胞の閾値制御

認識が成功した場合でも，必ずしもパターンに含まれる部分特徴がすべて検出されているとは限らない．検出されない原因としては，部分特徴の変形や欠損が考えられる．このような場合，W_{C0} 層に再構成されるパターンは，その部分特徴を欠いたものとなる．そこで，存在すべき特徴が検出されていない場合には，検出が行なわれなかった付近に存在する特徴検出細胞の閾値を下げ，特徴を検出しやすくする．

w_{CX} 細胞は存在すべき特徴がを抽出できなかったことを検出する細胞である．ところで，遠心性経路に信号が流れているにもかかわらず，対応する位置の求心性経路に信号が流れていないときには，存在すべき特徴が検出されていないと考えられる．そこで，w_{CX} 細胞は遠心性経路の信号と求心性経路の信号を比較し，その差に応じて出力を出す．すなわち，w_{CX} 細胞の出力は

$$w^t_{CXl}(\boldsymbol{n},k) = \varphi\left[w^t_{Cl}(\boldsymbol{n},k) - \sum_{\boldsymbol{\nu}\in D'_l} d'_l(\boldsymbol{\nu})\cdot u^t_{Sl}(\boldsymbol{n}+\boldsymbol{\nu},k)\right] \tag{5.31}$$

となり，w_C 細胞の応答から同じ段の u_S 細胞の応答の和を差し引いたものになっている．w_C 細胞と u_C 細胞の差を用いていないのは u_C 細胞の出力が利得制御信号 の影響で，変化するためである．このようにして w_{CX} 細胞が応答すると，閾値制御信号 $x^t_{Sl}(\boldsymbol{n},k)$ が変化するようになっており，その変化は,

$$x^t_{Sl}(\boldsymbol{n},k) = \beta_l\cdot x^{t-1}_{Sl}(\boldsymbol{n},k) + \beta'_l\cdot\sum_{\boldsymbol{\nu}\in D_l} d_l(\boldsymbol{\nu})\cdot w^{t-1}_{CXl}(\boldsymbol{n}-\boldsymbol{\nu},k) \tag{5.32}$$

で表される．この信号によって，特徴抽出細胞の閾値 $\theta_l^t(\boldsymbol{n},k)$ は次のように制御される．

$$\theta_l^t(\boldsymbol{n},k) = \frac{\theta_l^0(\boldsymbol{n},k)}{(1-\theta_l^0(\boldsymbol{n},k))(1-x_{Sl}^t(\boldsymbol{n},k)+x_{Xl}^t)+\theta_l^0(\boldsymbol{n},k)} \tag{5.33}$$

ここで，x_{Xl}^t は後述するもう一つの閾値制御信号で，最上位層で出力が出なかった場合に出力される．結果的に，この w_{CX} 細胞が応答した場合には，その付近に位置する u_S 細胞の閾値 $\theta_l^t(\boldsymbol{n},k)$ が下がるようになっている．

入力層にパターンを呈示した場合, 最上位層において, どの u_C 細胞も出力を出さない場合がある. このように, 最上位層で出力が出ないことが検出されると, 式 (5.34) にしたがって, 閾値制御信号 x_{Xl}^t が全段に対して出力される. その結果, 式 (5.33) に従い, すべての段の特徴抽出細胞 u_S の閾値を下げ, 応答が出やすい状態にする.

$$x_{Xl}^t = \begin{cases} x_{Xl}^{t-1} + \beta_{Xl} & \text{if} \quad u_{CL}^{t-1}(\kappa) = 0 \quad \text{for all } \kappa \\ \beta'_{Xl} \cdot x_{Xl}^{t-1} & \text{else} \end{cases} \tag{5.34}$$

ここで, $u_{CL}^{t-1}(\kappa)$ は最上位段の u_C 細胞の出力を表わす.

5.7 計算機シミュレーション

以下では, 計算機実験によって, 提案したシステムが英字筆記体連結文字列認識に有効であることを示している.

5.7.1 回路構成と入力パターン

入力層 U_{C0} は 113×37 個の細胞で構成した. 入力層の解像度が低い場合には粗いサンプリングによる誤差の影響を強く受けて, 直線の方位が正しく検出できなくなることがある. その結果, 認識結果が悪くなるということが従来から指摘されていた [46]. 本システムでは, このような解像度の粗さからくる誤差を小さくするために, 回路のスケールを今川のシステムのほぼ 2 倍としている.

テストに用いた入力パターンは, 学習パターンとテストパターンの両方とも, ワークステーション上で描画した文字パターンを 16 階調に量子化したものを用いている (図 5.7). テストパターンは入力層の大きさに合わせて 113×37 画素からなり, $3 \sim 5$ 文字の連

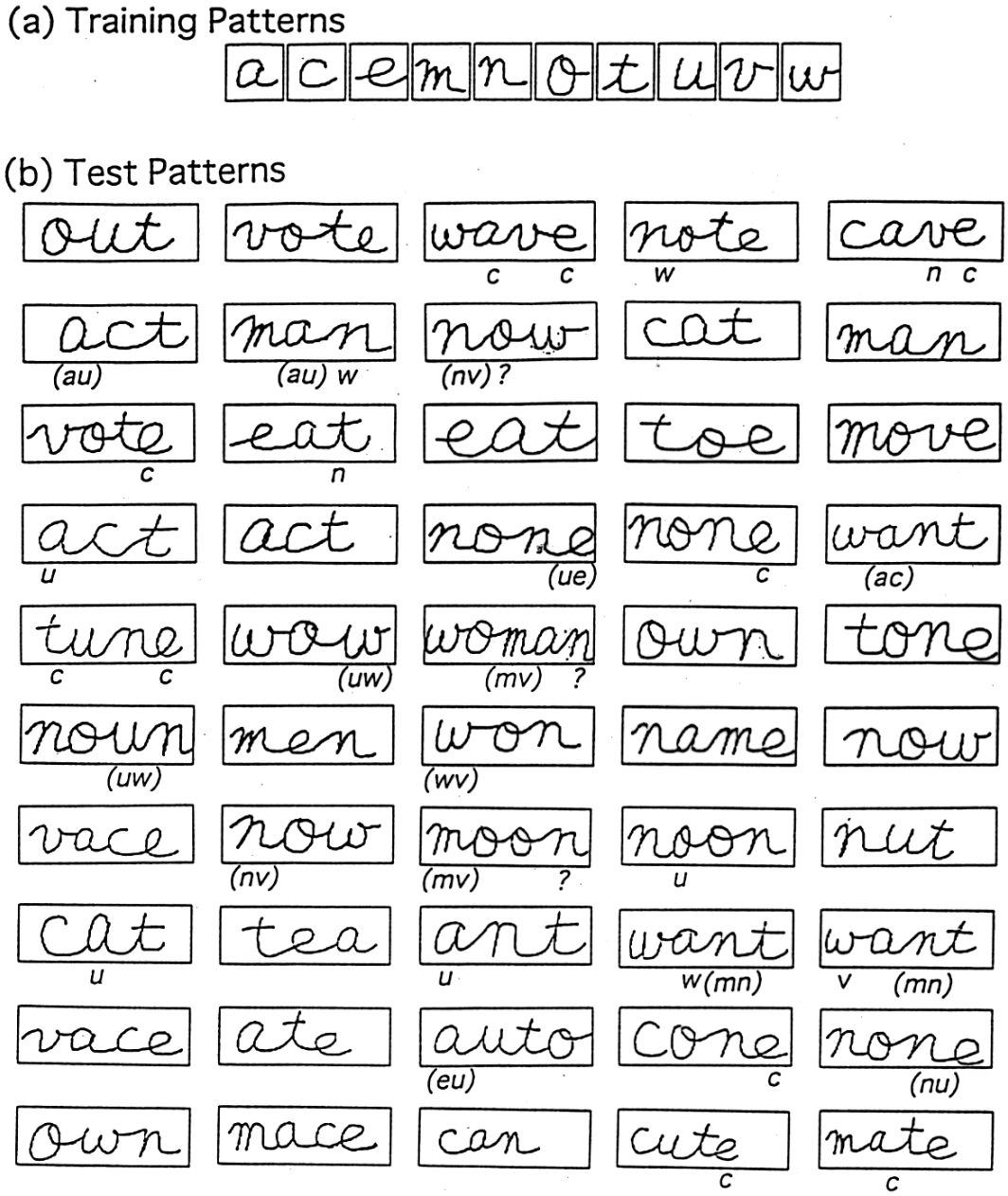

図 5.7: (a) トレーニングパターン (b) テストパターン

結文字列パターンを描いたもの (図 5.7 (b)) であるが, 学習パターンはシステムに学習させる際に 1 文字ずつシステムに与えるので, 1 個の文字パターンが入る大きさの 37×37 画素の大きさにした (図 5.7 (a)).

5.7.2 回路の学習

回路の 2 段目以降の自己組織化は, ネオコグニトロンと同様の教師なし学習 (ネオコグニトロンにおける winner-take-all 型の自己組織化 [33]) によって進めた. 本実験では, 図 5.7 (a) のような少数の学習パターン (10 パターン) を 1 セット用いて自己組織化させた. 学習に用いた文字パターンは, 'a', 'c', 'e', 'm', 'n', 'o', 't', 'u', 'v', 'w' の 10 文字である. 今回用いたのは 10 文字ではあるが, 連結文字列として人間が見た場合に比較的識別しにくいと思われるものを選んである. 学習は, 図 5.7 (a) の各パターンを入力面 U_{C0} 層の中央部付近に提示して行った. 学習は求心性経路についてのみ行い, 遠心性経路の結合は求心性経路の結合に連動して変化する.

今川のシステムでは, 学習パターンとして隣接文字との連結部分を含んだ文字パターンを用い, この連結部分の信号レベルをゆるやかに減らしていくことによって, 連結部分の変化による文字認識への影響を軽減していた [46].

本シミュレーションでは学習中にも サーチコントローラ を作動させ, 学習を行なった. 学習時には遠心性経路を用いていないので, サーチコントローラ は求心性信号のみで制御される. 学習パターンにおいては, サーチエリア の中心である注目点 μ は, ほぼ文字の中央部分になる. 文字と文字との連結部分は μ から離れた部分にあるので, 連結部分に対する利得は小さくなる. すると連結部分の変形の影響が軽減されるので, 連結部分の信号レベルを減らさなくても前述の今川のシステムで行ったような効果が出る.

5.7.3 認識結果

システムへの入力として, 入力層 (U_{C0} 層) に 3 ～ 5 文字の英文字列からなる図 5.7 (b) のようなテストパターンを入力として提示した. 本システムと今川のシステムとでは, 入力層の解像度と認識カテゴリの文字数 (今川のシステムでは 5 文字, 本システムでは 10 文字) とが異なるので同一条件での比較ができない. しかしながら, 予備実験とし

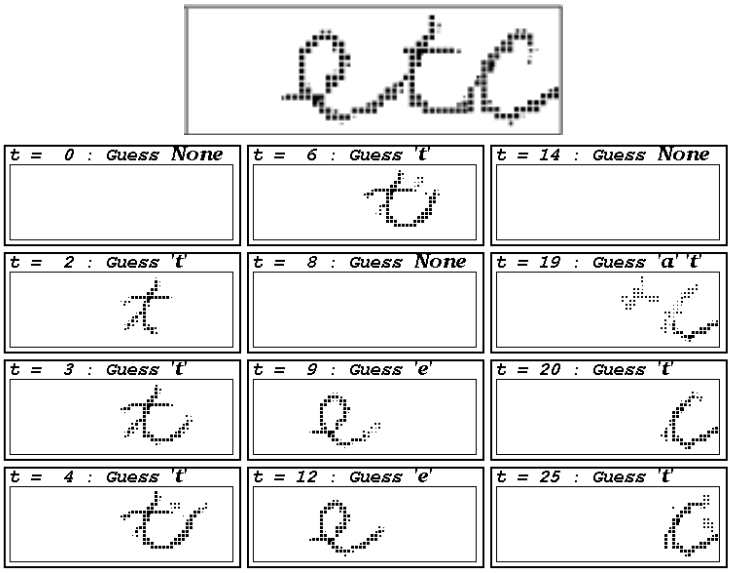

図 5.8: 折れ点処理回路なしの選択的注意機構の認識実験例: 上図は入力パターンを表す．下図は W_{C0} 層に切り出されたパターンを示している．t は経過時間を表す．$t = 4$ で既に “etc” という文字列のうちの ‘c’ の一部を削りとっている．この削りとった部分のゲインが下がるため $t = 19$ において ‘c’ が正しく認識されない

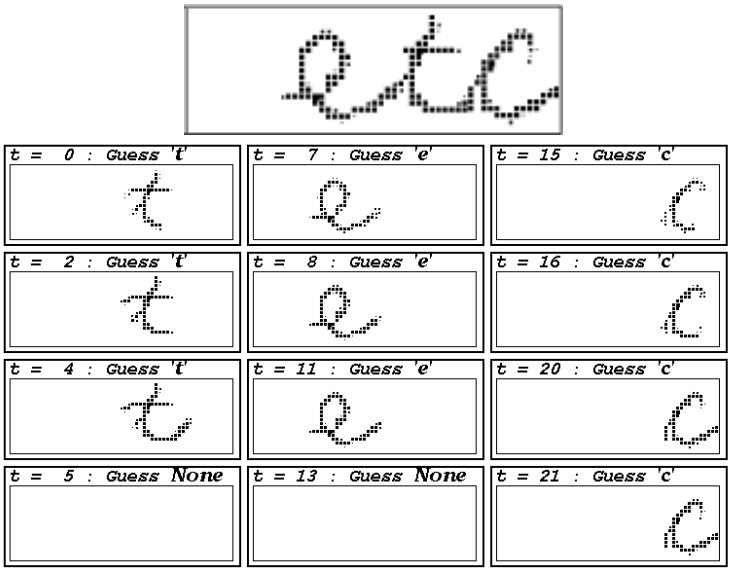

図 5.9: 折れ点処理回路ありの選択的注意機構の認識実験例: 上図は入力パターンを表す. 図は W_{C0} 層に切り出されたパターンを示している. t は経過時間を表す. 図 5.8 と異なり, $t = 4$ においても "etc" という文字列のうちの 'c' の一部を削りとることはない. 結果として $t = 15$ 以降においてもきちんと 'c' の認識・切り出しがなされている

て比較を同じ条件で行っている．これに関しては参考文献 [93] で述べており, 50 パターンの例で試行したところ 今川 のシステムでは全体の文字数の内 53.1% 程度しか認識・切り出しに成功しなかったのに対し, 折れ点処理回路ありのシステムでは,84.8% の文字の認識・切り出しに成功したという結果が出ている．図 5.8 と 図 5.9 は, シミュレーション時における W_{C0} 層の反応を示している．図 5.8 は折れ点処理回路なしのシステムの反応例を示しており, 図 5.9 は折れ点処理回路ありのシステムの反応例を示している．折れ点処理回路がない場合には, 最初の ‘t’ の認識・切り出しによって ‘c’ の一部分を削りとってしまい, 最後で ‘c’ の認識に失敗している．これに対し折れ点処理回路ありのシステムでは, このようなことはなくきちんと認識することが確かめられる．

図に示した 50 種類のテストパターンは, 筆者が無作為に作成したものである．テストパターンに含まれる文字をすべて正しく認識したものもあれば, 一部の文字の認識・切り出しに失敗したものもある．

図 5.7 (b) に示す各テストパターンの下部に何も印がないものは, パターンに含まれるすべての文字に対して認識・切り出しが成功したものである．パターンの下部に文字が書いてあるのは, その箇所が何の文字と間違えて認識されたかを示している．また, 一つの文字に対して複数の認識結果が出た場合には, その結果をすべて括弧で括って示している．‘?’ は, その部分が認識できなかったことを示している．

このシミュレーションでは, 50 個のテストパターンのうち 22 個のパターンに対しては, 含まれるすべての文字の認識・切り出しに成功した．また残りの 28 個のパターンにおいても, 含まれるほとんどの文字は正しく認識されているが, 一部の文字については間違いが生じた．誤りが生じたのは, 特に変形の大きい文字パターンが与えられた場合や, 文字と文字のつなぎ目が別の文字パターンの部分的特徴に似た形状となった場合であった．例えば,‘a’ の下半分が ‘u’ と類似しているので間違える傾向がある．同様な例として ‘e’ と ‘c’, ‘m’ と ‘n’ も挙げられる．このようにある文字パターンが別の文字パターンの一部となっているような場合の認識に関しては, 更なる工夫が必要であると考えている．また文字間のパターンと文字の一部を合わせてを別のパターンとみなすことがあった．例えば 図 5.7 (b) の 第 1 段 4 列目のパターンで ‘n’ の右半分と ‘o’ の下半分を ‘w’ とみな認識している．

5.8 検討

本章では, 折れ点処理回路を用いた選択的注意機構による文字列認識システムを提唱し, 従来から認識が非常に困難であるといわれていた手書きの続け文字の認識に本システムが有効であることをシミュレーションで示した.

本システムでは, 今川 らの作成した選択的注意機構のシステム [46] に対し, 折れ点処理回路を導入した. すなわち, 認識を行う求心性経路に折れ点検出回路を, また想起を行う遠心性経路に折れ点復元回路を加えた. これら折れ点処理回路を導入することによって, パターンの変形に対して強く, またパターンの切り出し能力の向上した文字列認識システムを構築することができた. さらに サーチコントローラ を改良して, 遠心性経路の信号を利用するようにしたところ, 切り出し能力を向上させることができた.

現在は計算機メモリと計算時間の制約のために, アルファベット 26 種の文字でなく, 10 種類の文字カテゴリでシミュレーションを行っている. しかしながら連結文字として見た場合に, 比較的認識が難しいと思われる文字に対してシミュレーションを行っているので, アルファベット 26 種の文字に対してもおそらく満足な結果が得られるであろうと考えている.

第 6 章

むすび

本論文では, 学習, モデルと生理学的な観察結果との比較, そして応用と多岐にわたり視覚システムに関して議論した. 脳研究自体は, まだまだ始ったばかりの学問であり多くのことが未知の処理システムである. 脳の視覚システム一つとっても我々の科学は未だに満足なものを再現できないでいる. それだけに脳というシステムを解明することは重要な仕事であり, また工学的には膨大な応用分野が広がっている. しかし複雑なシステムであるが故に解明していくには道標が必要であり, 本研究で述べたような生理学的知見に基づいたシステムの構成論的手法も必要な研究分野なのである.

6.1 本論文のまとめ

第 1 章では, 実際の脳の生理学的諸性質と提案されてきた学習仮説についての概説を行い, これを単純化したニューラルネットワークシステムについての歴史的背景について述べた.

第 2 章では, 特にサルやネコの一次視覚野の自己組織化モデルを扱った. 学習モデルとしては Kohonen の競合学習アルゴリズム (SOM) を拡張した射影学習の手法を提案した. この射影学習を用いた SOM を用いて脳内に発見されているディレクションマップと似たような性質を持つマップが形成されることを示した. また信号空間上での考察により方位マップの特異点と方向マップの特異点の関連性が理解可能であることを示した. 今後もさまざまな機能マップが発見されると思われるが, SOM を定式化する際の信

号空間上での考察によって, さまざまなマップ間の関連性が理解可能になると期待される. 射影学習を用いた SOM は, 信号空間が凸でないような複雑な形の信号空間でもリファレンスベクトルが信号空間からはみ出ることなくマップを形成できるので, さらに複雑な信号空間を考える際でも適用可能である.

第 3 章では階層型ネットワーク “ネオコグニトロン” を実際に構築し, これと視覚経路との対応付けを議論した. 形状認識を行うと言われている IT 野の細胞と計算機シミュレーションで細胞の性質の比較を行い, 非常に類似した反応特性を持つことを示した. IT 野の細胞は Logothetis らが示したような画像特異性を持つが, これは V1 野の線分特異性の拡張であると考えられる. V1 野のモデルをカスケード接続することによって, この性質が出てくるところを考えれば, このモデルが実際の脳内で実現されている可能性は高い. ただ生理学モデルとして問題となるのは結合の並進対称性である. この様な結合が実現されるには, Földiak が示した様にパターンが並進移動をしていくような状況での学習則 (trace-Hebb 則) を考えてやる必要がある. この点は今後の課題でもある.

第 4 章では, ネオコグニトロンの工学的な応用について示した. ロボットビジョンの分野では他にもさまざまなパターン認識の手法が考えられてきているが, 本研究で用いたシステムの示す認識率は, これらの工学的に考えられたシステムの示す認識率に劣るものではない. 提案しているネオコグニトロンの学習方法では, 閾値パラメータを決定するまでに時間がかかるが, 一旦閾値を決定してしまえば同程度の複雑さを持つパターンセットに適用しても学習にかかる時間は非常に短い. また工学的に提案されている様々なシステムの大部分は, 何らかの前処理や正規化が必要であり, このコストは軽視できるものではない.

第 5 章ではフィードバック型の経路を導入した選択的注意機構のモデルについての議論を行った. 大脳皮質の解剖学的な研究からフィードバック結合が存在していることは確認されている. フィードバック結合の働きについては様々な推測がなされており, 特に “注意” に関する機構との関連性が言われている. 本論文ではフィードバック信号による注意を切り替えながらパターン認識を行うシステムを構築し, 英字筆記体文字列の認識を行った. 英字筆記体文字列は複数の文字パターンが連結した複合パターンである. このような複合パターンの認識を行うためにはパターンの切り出しという問題が付きまとうが, 切り出しを行うためには認識を行わなければならないという循環的な問題に陥っ

てしまう．この問題を選択的注意機構のモデルを用いることによって解決している．選択的注意機構は複合パターンを構成する単独文字パターンに “注意” を当てて認識を行い, 認識を行った文字のパターンを切り出すことが出来ることを示した．本研究で行った計算機実験は, まだトイプロブレム的な性格を持っているが, パターンの図と地の分離問題や自然画像中の特定物体に対し注意を向けるシステムなど様々な応用が考えられる．

6.2 今後の課題

本研究で述べた研究における今後の課題をとしては以下のようなものがあげられる．第 2 章では学習の問題に関して, 特に V1 野の学習モデルについて取り上げたが, これ以降の領野の性質が, 果たして同じ学習ルールで再現できるかどうかは検討する必要がある．第 3 章以降で取り上げているネオコグニトロンベースの神経回路は結合に並進対称性があることを仮定しているが, この仮定を前提としない学習則の構築を行わなければならないであろう．また C 細胞の結合は学習によって決定しているわけではなく, 位相不変な応答を出すために作りつけられた固定結合である．この固定結合は “細胞面” を仮定しているが故にモデルでは簡単に実現できるが, 実際の皮質では陽に細胞面の構造を持つわけではない．したがって C 細胞の様な位相不変な結合を学習によって構成することを考えなければならない．

第 3 章では視覚経路とネオコグニトロンの対応について議論したが, 今後は

- 形状だけでなく, 陰影やオクルージョンの影響を補正すると思われるフィードバック経路の影響
- Biederman らの主張する 3 次元モデルが構成されるのかどうか

についても議論の余地が残されている部分である．第 3 章で述べた視覚経路モデルはフィードフォワードの経路である．しかし実際の脳の形状認識を行う視覚経路にはフィードバックの経路がフィードフォワードの系と, ほぼ対になって存在することが確認されている．当然, このフィードバックの系が画像の補正などに影響することが考えられる．この フィードバックループの問題は, 第 3 章だけの問題ではなく第 5 章にも関わってくる問題である．

Biederman らは Logothetis らの対立意見として脳内に “Geon” と呼ばれるような 3 次元構造を記述するモデルが存在していると主張している [5]. 現在までのところ, この 3 次元構造を記述するモデルを裏付ける生理学的な実験データは存在してない. しかし心理学的な実験データに基づけば, このような 3 次元構造を記述するモデルがあっても不思議ではない. したがって IT 野までの視覚経路で, このモデルが説明できるのか, それとも IT 野まではインテリジェントなパターン変換器として働いていてこれ以上の高次領野 (連合野など) で処理されているのかは非常に興味深いテーマである.

第 4 章においては, 閾値の設定が山登り方に準ずるものになっている. これは非常に時間がかかる上に見つけたパラメータがローカルマキシマムになっている可能性がある. 多次元空間でのパラメータサーチを効率的にするためには, まず次元の縮小を行うのが常道である. したがって細胞面の大きさや細胞段の数などを小さくする必要が出てくるが, 小さくすることによって量子化誤差等の影響が大きくなるので, これも効率との兼ね合いで定めなければならない. また多次元空間で効率的にパラメータを探索する手法として最近注目されている遺伝的アルゴリズム (GA: genetic algorithm) を適用してみてもいいであろう. パターンの認識能力の向上は第 5 章での選択的注意機構の切り出し機構の能力向上にもつながる.

繰り返しになるが, 脳研究は未だに未知の領域が多い. しかし, この複雑なシステムを理解してこそ, 我々は自分自身を知ることができるのではなかろうか？

謝辞

この研究を進めるに当たって, 多くのご指導と助言とを頂いた大阪大学 大学院基礎工学研究科システム人間系 福島 邦彦教授ならびに倉田耕治講師に深謝致します. 福島 邦彦教授には神経回路モデル研究の重要性を学部生ならびに大学院博士前期課程在学中に説いて頂き, さまざまな知識と有益な議論を賜りました. また本論文の主査をして頂き, 本論文をまとめる際にも数々の貴重なご意見を頂きました. 倉田耕治講師には, 神経回路モデルの理論的側面の面白さを教えて頂くだけでなく, さまざまな精神的支援を頂きました. 本論文の副査をして頂きました, 大阪大学大学院基礎工学研究科システム人間系 佐藤 俊輔教授, ならびに中野 馨教授に感謝致します. 本研究を進めるに当たって, 多忙にも関わらず密度の濃い議論をして頂いた科学技術事業団 川人動態脳プロジェクト 岡田真人博士に感謝致します.

また, 本研究の共同研究者である NTT 関西支社 永原 健一氏 と大阪大学基礎工学部生物工学科 古賀 和久氏に感謝致します. 彼らの地道な, データの集積がなければ本研究は成立しえなかったでしょう. 本研究を進めるにあたり, 議論や精神的支援をして頂いた NTT 基礎研究所橋爪 千枝氏, 理化学研究所 脳科学総合研究センター 青西 亨博士, ならびに大阪大学大学院基礎工学研究科システム人間系 菊地 眞之助手に感謝致します. 畑中 信技官には, 計算機の管理などの環境面でいろいろとお世話になりました. 最後に研究を全般にわたってバックアップして頂いた福島研究室のみなさんに感謝致します.

参考文献

[1] S. Amari. A Theory of Adaptive Pattern Classifiers. *IEE Trans. EC-16*, pp. 279–307, 1967.

[2] 芦田英一, 福島邦彦. 選択的注意機構のモデルによる漢字認識. , 電子情報通信学会ニューロコンピューティング研究会, 1993.

[3] W. T. Baxter and B. M. Dow. Horizontal Organization of Orientation Sensitive Cells in Primate Visual Cortex. *Biological Cybernetics*, Vol. 61, pp. 171–182, 1989.

[4] R. Bellman. *Adaptive Control Processes: A Guided Tour.* Princeton University Press, New Jersey, 1961.

[5] I. Biederman and P. C. Gerhardstein. Recognizing Depth-Rotated Objects: Evidence and Conditions for Three-Dimensional Viewpoint Invariance. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 19, pp. 1162–1182, 1993.

[6] C M. Bishop. *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

[7] C. Blakemore and G. F. Cooper. Development of the Brain Depends on the Visual Environment. *Nature*, Vol. 228, pp. 477–478, 1970.

[8] C. Blakemore and R. C. Van Sluyters. Innate and Environmental Factors in the Development of the Kitten's Visual Cortex. *J.Physiol.*, Vol. 248, pp. 663–716, 1975.

[9] G. G. Blasdel and G. Salama. Voltage Sensitive Dyes Reveal a Modular Organization in Monkey Striate Cortex. *Nature*, Vol. 321, pp. 579–585, 1986.

[10] D. Boussaoud and R. Desimone an L. G. Ungerleider. Visual Topography of Area TEO in the macaque. *J.Comp.Neurol*, Vol. 306, pp. 554–575, 1991.

[11] V. Braintenberg and C. Braitenberg. Geometry of Orientation Columns in the Visual Cortex. *Biological Cybernetics*, Vol. 33, pp. 179–186, 1979.

[12] E. Bricolo, T. Poggio, and N. K. Logothetis. 3D Object Recognition: A Model of View-Tuned Neurons. *Advances in Neural Information Processing System*, Vol. 9, pp. 41–47, 1996.

[13] K. Broadman. *Vergleichende Lokalizationslehreder Grosshirnrinde in ihren Prinzipien dargestelltauf Grunde des Zellenbaues.* J.A.Barth, Leipzig, 1909.

[14] C. Bruce, R. Desimone, and C. G. Gross. Visual Properties of Neurons in a Polysensory Area in Superior Temporal Sulcus of the Macaque. *J.Neurophysiol.*, Vol. 46, No. 2, pp. 369–384, 1981.

[15] P. Dayan, G. E. Hinton, and R. M. Neal. The Heomholtz Machine. *Neural Computation*, Vol. 7, pp. 889–904, 1995.

[16] R. Desimone, T. D. Albright, C. G. Gross, and C. Bruce. Stimulus-selective Properties of Inferior Temporal Neurons in the Macaque. *J.Neurosci*, Vol. 4, pp. 2051–2062, 1984.

[17] R. Durbin and G. Mitchison. A Dimension Reduction Framework for Understanding Cortical Maps. *Nature*, Vol. 343, pp. 341–344, 1990.

[18] E. Erwin, K. Obermayer, and K. Schulten. Models of Orientation and Ocular Dominance Columns in the Visual Cortex: A Critical Comparison. *Neural Computation*, Vol. 7, pp. 425–468, 1995.

[19] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in primate cerebral cortex. *Cereb.Cortex*, Vol. 1, pp. 1–47, 1991.

[20] P. Földiak. Learning Invariance from Transformational Sequences. *Neural Computation*, Vol. 3, pp. 194–200, 1991.

[21] 藤田俊史, 安藤広志. 3 次元物体の学習と動的認識モデル. 映像メディア学会誌, Vol. 51, No. 8, pp. 1249–1261, 1997.

[22] K. Fukushima. Cognitron: A Self-Organizing Multilayered Neural Network Model. *Biological Cybernetics*, Vol. 20, No. 3, pp. 121–136, 1975.

[23] K. Fukushima. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by shift in position. *Biological Cybernetics*, Vol. 36, No. 4, pp. 193–202, 1980.

[24] K. Fukushima. Neural Network Model for Selective Attention in Visual Pattern Recogition and Associative Recall. *Applied Optics*, Vol. 26, No. 23, pp. 4985–4992, 1987.

[25] K. Fukushima. Analysis of the Process of Visual Pattern Recognition by the Neocognitron. *Neural Networks*, Vol. 2, pp. 413–420, 1989.

[26] 福島邦彦. 神経回路と情報処理. 朝倉書房, 1989.

[27] 福島邦彦. 神経回路理論. ブレインサイエンス, Vol. 5, No. 3, pp. 73–80, 1994.

[28] K. Fukushima and S. Miyake. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recogn.*, Vol. 15, No. 6, pp. 455–469, 1982.

[29] K. Fukushima, K. Nagahara, and H. Shouno. Training neocognitron to recognize handwritten digits in the real world. In *Proceedings The Second Aizu International Symposium on Parallel Algorithms/Architectures Synthesis*. pAs'97, Mar. 1997.

[30] K. Fukushima, K. Nagahara, H. Shouno, and M. Okada. Training neocognitron to recognize handwritten digits in the real world. In *World Congress on Neural Networks*, pp. 21–24, 1996.

[31] K. Fukushima and M. Tanigawa. Use of Different Threshold in Learning and Recognition. *Neurocomputing*, Vol. 11, No. 1, pp. 1–17, 1996.

[32] K. Fukushima and N. Wake. Handwritten Alphanumeric Character Recognition by the Neocognitron. *IEEE Trans. Neural Networks*, Vol. 2, No. 3, pp. 355–365, 1991.

[33] K. Fukushima and N. Wake. An improved learning algorithm for the neocognitron. In I. Aleksander J. Taylor, editor, *Artificial Neural Networks*, Vol. 1 of *2*, pp. 497–505. Amsterdam: North-Holland, 1992.

[34] S. Geman and D. Geman. Stochastic relaxation, Gibbs Distributions, and the Bayesian restoration of images. *IEEE Trans.PAMI*, Vol. 6, pp. 721–741, 1984.

[35] K. G. Götz. Do "d-blob" and "l-blob" Hypercolumns Tessellate the Monkey Visual Cortex? *Biological Cybernetics*, Vol. 56, pp. 213–223, 1987.

[36] C. G. Gross, C. E. Rocha-Miranda, and D. B. Bender. Visual Properties of Neurons in Inferotemporal Cortex of the Macaque. *Jounal of Neurophysiology*, pp. 96–111, 1972.

[37] D. O. Hebb. *Organization of Behavior.* John Wiley & Sons, NewYork London Sydney, 1949.

[38] R. Hecht-Nielsen. *Neurocomputing.* Addison-Wesley, 1990.

[39] B. K. P. Horn. *Robot Vision.* MIT Press, 1986.

[40] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *J.Physiol.(Lond.)*, Vol. 148, pp. 574–591, 1959.

[41] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J.Physiol.*, Vol. 106, No. 1, pp. 106–154, Jan. 1962.

[42] D. H. Hubel and T. N. Wiesel. Receptive Fields of Cells in Striate Cortex of Very Young, Visually Inexperienced Kittens. *J.Neurophysiol.*, Vol. 26, pp. 994–1002, 1963.

[43] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J.Neurophysiol.*, Vol. 28, No. 2, pp. 229–289, 1965.

[44] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *J.Physiol.*, Vol. 195, No. 1, pp. 215–243, 1968.

[45] D. H. Hubel and T. N. Wiesel. Sequence Regularity and Geometry of Orientation Columns in Monkey Striate C ortex. *J.Comp.Neurol.*, Vol. 158, pp. 267–293, 1974.

[46] 今川太郎, 福島邦彦. 選択的注意機構による英字筆記体連結文字列認識". 信学会論文誌, Vol. J-74 D-II, No. 12, pp. 1768–1775, 1991.

[47] M. Ito, H. Tamura, I. Fujita, and K. Tanaka. Size and Position Invariance of Neuronal Responses in Monkey Inferotemporal Cortex of Monkeys. *J.Neurophysiol.*, Vol. 73, pp. 218–226, 1995.

[48] E. Iwai and M. Mishkin. Further Evidence on the Locus of the Visual Area in the Temporal Lobe of the Monkey. *Exp. Neurol.*, Vol. 25, No. 4, pp. 585–594, Dec. 1969.

[49] J. P. Jones and L. A. Palmer. An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive fields in Cat Striate Cortex. *Journal of Neurophysiology*, Vol. 58, No. 6, pp. 1223–1258, Dec. 1987.

[50] 柏木規正, 倉田耕治. 第 1 次視覚野機能地図形成の Kohonen モデルによるシミュレーション. NC95-181, 電子情報通信学会ニューロコンピューティング研究会, Mar. 1995.

[51] M. Kawato, T. Inui, S. Hongo, and H.Hayakawa. Computational theory and neural network models of interaction between visual cortical areas. TR-A-0105, ATR Technical Report, ATR Kyoto, 1991.

[52] E. Kobatake and K. Tanaka. Neuronal Electivities to Complex Object Features in the Ventral Visual Pathway of the Macaque Cerebral Cortex. *Journal of Neurophysiology*, Vol. 71, No. 3, pp. 856–867, Mar. 1994.

[53] T. Kohonen. Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, Vol. 43, pp. 59–69, 1982.

[54] T. Kohonen. *Self-Organizing Maps*. Springer, 1995.

[55] 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*, Vol. 1, No. 4, pp. 541–551, 1989.

[56] R. Linsker. From Basic Network Principles to Neural Architecture: Emergence of Orientation Columns. *Proc.Natl.Acad.Sci.U.S.A.*, Vol. 83, pp. 8779–8783, 1986.

[57] R. Linsker. From Basic Network Principles to Neural Architecture: Emergence of Orientation Selective Cells. *Proc.Natl.Acad.Sci.U.S.A.*, Vol. 83, pp. 8390–8394, 1986.

[58] R. Linsker. From Basic Network Principles to Neural Architecture: Emergence of Spatial Opponents Cells. *Proc.Natl.Acad.Sci.U.S.A.*, Vol. 83, pp. 7508–7512, 1986.

[59] Y. J. Lo and M .M. Poo. Activity-dependent synaptic competition *in vitro*: heterosynaptic suppression of developping synapses. *Science*, Vol. 254, pp. 1019–1022, 1991.

[60] N. K. Logothetis, J. Pauls, and T. Poggio. Spatial Reference Frames for Object Recognition: Tuning for Rotation in Depth. , M.I.T A.I Memo, 1995.

[61] N. K. Logothetis and D. L. Sheinberg. Visual Object Recognition. *Annual Review of Neuroscience*, Vol. 19, pp. 577–621, 1996.

[62] D. Marr. A Theory of Cerebellar Cortex. *J.Physiol.*, Vol. 202, No. 2, pp. 437–470, 1969.

[63] K. D. Miller. Development of Orientation Columns via Competition between on- and off-center inputs. *Neuro Report*, Vol. 3, pp. 73–76, 1992.

[64] K. D. Miller. A Model for the Development of Simple Cell Receptve Fields and the Ordered Arrangement of Orientation Columns through Activity Dependent Competition between on- and off-center Inputs. *J.Nerosci.*, Vol. 14, pp. 409–441, 1994.

[65] K. D. Miller, J. B. Keller, and M. P. Stryker. Ocular Dominace Column Development: Analysis and Simulation. *Science*, Vol. 245, pp. 605–615, 1989.

[66] M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge, MA, 1969.

[67] M. Miyashita, D-S. Kim, and S. Tanaka. Cortical Direction Selectivity Without Directional Experience. *NeuroReport*, Vol. 8, No. 5, pp. 1187–1192, 1997.

[68] M. Miyashita and S. Tanaka. A Mathematical Model for the Self-Organization of Orientation Columns in Visual Cortex. *NeuroReport*, Vol. 3, No. 1, pp. 69–72, 1992.

[69] M. Miyashita and S. Tanaka. A Model for Self-Organization of Spatio-Temporal Receptive Fields of Visual Cortex Cells. 日本神経回路学会第 7 回全国大会講演論文集, pp. 223–224, Nov. 1996.

[70] 永原健一, 庄野逸, 福島邦彦. ETL データベースに対するネオコグニトロンの認識率. NC95-47, 電子情報通信学会ニューロコンピューティング研究会, Jul. 1995.

[71] 永原健一, 庄野逸, 福島邦彦, 岡田真人. ネオコグニトロンの実用化 —ETL 文字データベースによ る評価—. NC95-149, 電子情報通信学会ニューロコンピューティング研究会, Jul. 1996.

[72] K. Obermayer and G. G. Blasdel. Geometry of Orientation and Ocular Dominance Columns in Monkey Striate Cortex. *The Journal of Neuroscience*, Vol. 13, pp. 4114–4129, 1993.

[73] K. Obermayer, G. G. Blasdel, and K. Schulten. Statistical Mechanical Analysis of Self-Organization and Pattern Formation during the Dvelopment of Visual Maps. *Phys.Rev.A*, Vol. 45, No. 10, pp. 7568–7589, 1992.

[74] K. Obermayer, H. Ritter, and K. Schulten. A Principlefor the Formation of the Spatial Structure of Cortical Feature Maps. *Proc.Natl.Acad.Sci.U.S.A.*, Vol. 87, , 1990.

[75] K. Obermayer, K. Schulten, and G. G. Blasdel. A Comparison between a Neural Network Model for the Formation of Brain Maps and Experimental Data. In D. S. Tourtezky and R. Lippman, editors, *Advances in Neural Information Processing Systems 4*, pp. 83–90, 1992.

[76] Y. Oda, K. Kawasaki, M. Morita, H. Korn, and H. Matsui. Inhibitory long-term potentiation underlies auditory conditioning of goldfish escape behvior. *Nature*, Vol. 394, No. 6689, pp. 182–185, Jul. 1998.

[77] M. Ohno, M. Okada, and K. Fukushima. Neocognitron Learned by Backpropagation. *Systems and Computers in Japan*, Vol. 26, No. 5, pp. 19–28, 1995.

[78] I. P. Pavlov. *Conditioned reflexes: An Investigation of Physiological Activity of the Cortex.* Oxfort University Press, London, 1927.

[79] D. I. Perret, A. J. Mistlin, and A. J. Chitty. Visual Neurons Responsive to Faces. *Trends in Neuroscience*, Vol. 10, pp. 358–364, 1987.

[80] D. I. Perret, P. A. J. Smith, D. D. Potter, A. J. Mistlin, A. S. Head, A. D. Milner, and M. A. Jeeves. Visual Cells in the Temporal Cortex Sensitive to Face View and Gaze Direction. *Proc.Roy.Soc.Lond.B*, Vol. 223, No. 1232, pp. 293–317, Jan. 1985.

[81] M. Riesenhuber, H. U. Bauer, and T. Geisel. Analyzing phase transitions in high dimensional self-organizing maps. *Biological Cybernetics*, Vol. 75, pp. 397–407, 1996.

[82] M. Riesenhuber and T. Poggio. Just One View: Invariances in Inferotemporal Cell Tuning. In *Advances in Neural Information Processing Systems 10*, (in print).

[83] E. T. Rolls. Brain mechanisms for invariant visual recognition and learning. *Behavioural Processes*, Vol. 33, pp. 113–138, 1994.

[84] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol.Rev.*, Vol. 65, No. 6, pp. 386–408, 1958.

[85] F. Rosenblatt. *Principles of Neurodynamics.* Spartan Books, Washington D. C., 1962.

[86] D. E. Rumelhart, J. L. McClelland, and PDP Research Group. *Parallel Distributed Processing: Explorations in Microstructure of Cognition.* MIT Press, Cambridge, MA London, 1986.

[87] D. E. Rumelhart and D. Zipser. Feature Discovery by Competitive Learning. *Cognitive Science*, Vol. 9, pp. 75–112, 1985.

[88] K. Sakai and S. Tanaka. Computational analysis of the nonlinear spatial structure in V1 complex cells. 日本神経回路学会第8回全国大会講演論文集, pp. 79–80, 1997.

[89] E. L. Schwartz. Spatial Mapping in the Primate Sensory Projection: Analytic Structure and Relevance to Perception. *Biological Cybernetics*, Vol. 25, No. 4, pp. 645–669, 1977.

[90] C. J. Shatz. The developping brain. *Scientific American*, Vol. 267, No. 3, pp. 60–67, 1992.

[91] A. Shmuel and A Grinvald. Functional Organization for Direction of Motion and Its Relationship to Orientation Maps in Cat Area 18. *The Jouranl of Neuroscience*, Vol. 16, No. 21, pp. 6945–6964, Nov.1 1996.

[92] B. Shölkopf, K. K. Sung, C. J. C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik. Comparing Suport Vector Machines with Gaussial Kernels to Radial Basis Function Classifiers. *IEEE Trans. on Signal Processing*, Vol. 45, No. 11, pp. 2758–2765, Nov. 1997.

[93] 庄野逸, 福島邦彦. 折れ点検出回路を用いた選択的注意機構による英字筆記体 連結文字列認識. NC92-27, 電子情報通信学会ニューロコンピューティング研究会, Jul. 1992.

[94] 庄野逸, 福島邦彦. 折れ点検出回路を用いた選択的注意機構による英字筆記体 連結文字列認識. NC92-105, 電子情報通信学会ニューロコンピューティング研究会, Mar. 1993.

[95] 庄野逸, 福島邦彦. 折れ点処理回路を用いた選択的注意機構による英字筆記体 連結文字列認識. 電子情報通信学会論文誌, Vol. D-II, No. J77-D-II, pp. 940–950, 5 1994.

[96] 庄野逸, 福島邦彦. 選択的注意機構のモデルを用いた英字筆記体連結文字列認識. NC93-119, 電子情報通信学会ニューロコンピューティング研究会, Mar. 1994.

[97] H. Shouno and K. Fukushima. Connected character recognition in cursive handwriting using selective attention model with bend processing. *Systems and Computers in Japan*, Vol. 26, No. 10, pp. 35–46, Sep. 1995.

[98] H. Shouno, K .Fukushima, and M. Okada. *Intelligent Techniques in Character Recognition: Practical Applications*, chapter Recognition of Handwritten Digits in the Real World by Neocognitron. CRC Press, 1998.

[99] 庄野逸, 倉田耕治. Kohonen model による Direction Map の形成. NC96-45, 電子情報通信学会ニューロコンピューティング研究会, Oct. 1996.

[100] 庄野逸, 倉田耕治. SOM による Direction Map の形成. 日本神経回路学会第 7 回全国大会講演論文集, Sep. 1996.

[101] 庄野逸, 倉田耕治. 射影学習を用いた Kohonen model による Direction Map の形成. 神経回路学会誌, Vol. 4, No. 3, pp. 109–114, 1997.

[102] H. Shouno, K. Nagahara, K. Fukushima, and M. Okada. Handwritten Digit Recognition with a Neocognitron using Different Thresholds in Learning and Recognition. In *Proceedings of International Conference on Neural Information Processing*, Vol. Vol.1, pp. 405–409. ICONIP'96, Sep. 1996.

[103] 庄野逸, 永原健一, 福島邦彦, 岡田真人. ネオコグニトロンの実用化 —大規模データベースによる評価—. NC97-19, 電子情報通信学会ニューロコンピューティング研究会, Jun. 1997.

[104] P. Y. Simard, Y. LeCun, and J. Denker. Efficient Pattern Recognition Using a New Transfomation Distance. In S. J. Hanson J. D. Cowan and C. L. Giles, editors, *Advances in Neural Information Processings 5*, pp. 50–58, 1993.

[105] P. Y. Simard, B. Victorri, Y. LeCun, and J. Denker. Tangent Prop - A Formalism for Specifying Selected Invariances in an Adaptive Network. In D. S. Tourtezky and R. Lippman, editors, *Advances in Neural Information Processing Sytems 4*, pp. 895–903, 1992.

[106] M. Stone. Cross-validation: A review. *Math.Operations.Stat.Ser.Stat*, Vol. 9, No. 1, pp. 127–139, 1978.

[107] K. Tanaka. Neuronal Mechanisms of Object Recognition. *Science*, Vol. 262, pp. 685–688, 1993.

[108] K. Tanaka. Inferotemporal Cortex and Object Vision. *Annual Review of Neruoscience*, Vol. 19, pp. 109–139, 1996.

[109] S. Tanaka. Theory of Ocular Dominance Column Formation. *Biological Cybernetics*, Vol. 64, pp. 263–272, 1991.

[110] S. Tanaka. Topology of Visual Cortical Maps. *Forma*, Vol. 12, No. 2, pp. 101–106, 1997.

[111] S. Tanaka and H. Shinbata. A mathematical model for neural response properties of modulara organization in the motion-processing area of the primate cerebral cortex. *NEC R&D*, Vol. 34, No. 1, pp. 1–11, 1993.

[112] S. Tanaka and H. Shinbata. Mathematical model for self-organization of direction columns in the primate middle temporal area. *Biological Cybernetics*, Vol. 70, pp. 227–234, 1994.

[113] 内村俊二, 水野圭, 浜本義彦, 富田眞吾. 実験計画法による Gabor 特徴の分析. HIP97-3, 電子情報通信学会 HIP 研究会, Jul. 1997.

[114] C. von der Malsburg. Self-Organization of Orientation Sensitive Cells in the Striate Cortex. *Kyberntetik*, Vol. 14, pp. 85–100, 1973.

[115] M. Weliky, W. H. Bosking, and D. Fitzpatrick. A Systematic Map of Direction Preference in Primary Visual Cortex. *Nature*, Vol. 379, No. 22, pp. 725–728, 1996.

[116] B. Widrow. Pattern Recognition and Adaptive Control. *IEEE Trans. Applications and Industry*, Vol. 83, No. 74, pp. 269–277, Sep. 1964.

[117] B. Widrow and M. E. Hoff. Adaptive Switching Circuits. In *In 1960 IRE WESCON Convention Record*, part 4, pp. 96–104, 1960.

[118] L. Xu. A unified learning scheme: Bayesian-Kullback Ying-Yang machine. In D. S. Touretzky M. Mozer M. Hasselmo, editor, *Advances in Neural Information Prcessing Systems 8*. MIT Press Cambridge MA, 1996.

[119] 吉本和哉, 福島邦彦, 庄野逸. ネオコグニトロンの C 細胞の自己組織化. NC97-123, 電子情報通信学会ニューロコンピューティング研究会, Mar. 1998.

[120] S. Zeki and S. Shipp. The Functional Logic of Cortical Connections. *Nature*, Vol. 335, pp. 311–317, 1988.

関連発表論文

学術論文:

1. 庄野 逸　福島 邦彦. 折れ点処理回路を用いた選択的注意機構による英字筆記体連結文字列認識. **電子情報通信学会論文誌**, J77-**D-II**(5):940–950, 1994.

2. 庄野 逸　倉田 耕治. 射影学習を用いた Kohonen Model による Direction Map の形成. **神経回路学会誌**, **4**(3):109–114, 1997.

著書:

1. H. Shouno K. Fukushima and M. Okada. Recognition of Handwritten Digits in the Real World by Neocognitron, CRC Press, In press.

国際会議論文:

1. K. Fukushima and H. Shouno. Cursive Word Recognition using Selective Attention with Bend-processing. In *Proceedings of the International Conference on Artificial Neural Networks*, Volume 2. ICANN'94, May 1994.

2. K. Fukushima and H. Shouno. Visual Pattern Recognition With Selective Attention. In *World Congress on Neural Networks*, Volume I, pages 575–580. WCNN'94, June 1994.

3. K. Fukushima K. Nagahara H. Shouno and M. Okada. Training Neocognitron to Recognize Handwritten Digits in the Real World. In *World Congress on Neural Networks*, pages 21–24, 1996.

4. H. Shouno K. Nagahara K. Fukushima and M. Okada. Handwritten Digit Recognition with a Neocognitron using Different Thresholds in Learning and Recognition. In *Proceedings in Neural Information Processing — Proceedings of International Conference on Neural Information Processing*, Vol.1, pages 405–409. ICONIP'96, Sept. 1996.

5. K. Fukushima K. Nagahara and H. Shouno. Training Neocognitron to Recognize Handwritten Digits in the Real World. In *Proceedings The Second Aizu International Symposium on Parallel Algorithms/Architectures Synthesis*. pAs'97, Mar. 1997.

6. K. Fukushima E. Kimura and H. Shouno. Handwritten Digit Recognition by a Neocognitron with Improved Bend-detectors. In *International ICSC/IFAC Symposium on Neural Computation*. NC'98, in press 1998.

7. K. Fukushima E. Kimura and H. Shouno. Neocognitron with Improved Bend-extractors. In *Proceedings of the 1998 International Joint Conference on Neural Networks in WCCI'98 1998 IEEE World Congress on Computational Intelligence*, pages 1172–1175. WCCI'98, May 1998.

国内学会:

1. 庄野 逸 福島 邦彦. 選択的注意機構を用いた英文筆記体文字列認識. **日本神経回路学会第 *6* 回全国大会講演論文集**, 1995.10.

2. 庄野 逸 倉田 耕治. SOM による Direction Map の形成. **日本神経回路学会第 *7* 回全国大会講演論文集**, 1996.9.

3. 庄野 逸 岡田 真人 福島 邦彦. ネオコグニトロンを用いた 3D 物体認識. **日本神経回路学会第 *8* 回全国大会講演論文集**, 1997.11.

研究会:

1. 庄野 逸 福島 邦彦. 折れ点検出回路を用いた選択的注意機構による英字筆記体連結文字列認識. NC92-27, 電子情報通信学会ニューロコンピューティング研究会, 1992.7.

2. 庄野 逸 福島 邦彦. 折れ点検出回路を用いた選択的注意機構による英字筆記体連結文字列認識. NC92-105, 電子情報通信学会ニューロコンピューティング研究会,

1993.3.

3. 庄野 逸 福島 邦彦. 選択的注意機構のモデルを用いた英字筆記体連結文字列認識. NC93-119, 電子情報通信学会ニューロコンピューティング研究会, 1994.3.

4. 永原 健一 庄野 逸 福島 邦彦. ETL データベースに対するネオコグニトロンの認識率. NC95-47, 電子情報通信学会ニューロコンピューティング研究会, 1995.7.

5. 永原 健一 庄野 逸 福島 邦彦 岡田 真人. ネオコグニトロンの実用化 —ETL 文字データベースによる評価—. NC95-149, 電子情報通信学会ニューロコンピューティング研究会, 1996.7.

6. 岡崎 哲朗 庄野 逸 福島 邦彦. ネオコグニトロン型神経回路モデルを対象としたライブラリの設計と実装. NC96-117, 電子情報通信学会ニューロコンピューティング研究会, 1997.3.

7. 庄野 逸 永原 健一 福島 邦彦 岡田 真人. ネオコグニトロンの実用化 —大規模データベースによる評価—. NC97-19, 電子情報通信学会ニューロコンピューティング研究会, 1997.6.

8. 木村 栄司 福島 邦彦 庄野 逸. 脱抑制を用いた特徴検出回路. NC97-31, 電子情報通信学会ニューロコンピューティング研究会, 1997.7.

9. 吉本 和哉 福島 邦彦 庄野 逸. ネオコグニトロンの C 細胞の自己組織化. NC97-123, 電子情報通信学会ニューロコンピューティング研究会, 1998.3.

10. 木村 栄司 福島 邦彦 庄野 逸. 新しい折れ点抽出機構を持つネオコグニトロン. NC97-168, 電子情報通信学会ニューロコンピューティング研究会, 1998.3.

参考:

1. H. Shouno and K. Fukushima. Connected Character Recognition in Cursive Handwriting using Selective Attention Model with Bend Processing. *Systems and Computers in Japan*, **26**(10):35–46, Sep. 1995.

2. H. Shouno K. Kurata. Formation of a Direction Map using Kohonen's SOM by Projection Learning., *Biological Cybernetics*, 投稿中