

東北大学大学院 工学研究科 博士学位論文  
階層的パターン統合処理に基づく視覚情報  
処理モデルに関する研究

佐藤俊治

東北大学大学院 工学研究科 電気・通信工学専攻 阿曾研究室  
学籍番号 97TD3005

2000年2月

# 目次

<b>第1章 序論</b>	<b>6</b>
1.1 本研究の背景	6
1.2 本研究の目的	9
1.3 本稿の構成	11
<b>第2章 ネオコグニトロン</b>	<b>12</b>
2.1 単純型細胞と複雑型細胞の階層型仮説	12
2.2 ネオコグニトロンの性質および構造	15
2.2.1 ネオコグニトロンの構造	15
2.2.2 細胞間の結合法則	17
2.2.3 各細胞の数式表現	19
2.2.4 細胞面	20
2.2.5 特徴検出と位置ずれ許容	23
2.3 学習方法	23
2.4 回転パターンに対する評価実験	25
2.5 まとめと考察	25
<b>第3章 回転対応型ネオコグニトロン</b>	<b>28</b>
3.1 回転パターンの認識に関する生理学的・心理学的考察	28
3.2 従来型ネオコグニトロンの再構成	32
3.2.1 細胞面群	32
3.2.2 回転対応型ネオコグニトロンの構造および性質	34
3.2.3 各細胞の数式表現	34
3.3 学習方法	39
3.4 回転したパターンの認識機構	41
3.4.1 S-細胞による回転した局所パターンの検出	44
3.4.2 C-細胞による回転によるずれの吸収	46
3.5 評価実験	46
3.5.1 回転パターンに対する評価実験	48
3.5.2 手書き数字パターンに対する認識性能	50
3.5.3 回転した手書き数字パターンに対する認識性能	53
3.6 まとめと考察	53

<b>第4章</b>	<b>高速生成アルゴリズム</b>	<b>55</b>
4.1	パラメータ設定に関する問題点	55
4.2	S-細胞のベクトル空間上での解析	57
4.2.1	ベクトル空間上での解析	57
4.2.2	しきい値とS-細胞の発火条件の関係	59
4.3	学習過程の解析	60
4.4	細胞面群融合法による高速生成アルゴリズム	65
4.4.1	処理の流れ	65
4.4.2	細胞面群融合法	67
4.5	評価実験	70
4.5.1	最適なしきい値集合の決定	70
4.5.2	実行時間	70
4.5.3	従来法との比較	71
4.6	まとめ	73
<b>第5章</b>	<b>連想想起を伴う回転パターンの認識</b>	<b>78</b>
5.1	「回転に対して単純なパターン」の同定	78
5.2	回転したパターンを認識するための回路構成	81
5.3	認識モデルと連想想起モデル	85
5.3.1	提案モデルの処理過程	85
5.3.2	認識モデル	85
5.3.3	連想想起モデル	87
5.3.4	提案モデルの数式表現	87
5.4	評価実験	90
5.4.1	回転したパターンに対する認識性能	90
5.4.2	認識モデルのみによる回転したパターンの認識性能	91
5.4.3	鏡像反転したパターンに対するモデルの動作	91
5.5	認識モデルと連想想起モデルの融合	91
5.5.1	認識モデルと連想想起モデルの融合	92
5.5.2	回転対応型ネオコグニトロンとの比較	93
5.6	まとめ	95
<b>第6章</b>	<b>結論</b>	<b>97</b>
6.1	本研究の成果	97
6.2	今後の課題	98

# 目次

1.1	代表的な工学的手法による認識方法の流れ図	7
1.2	アカゲザルの視覚情報処理の部位および流れ	8
2.1	方位選択性を持つ単純型細胞と複雑型細胞の特性	13
2.2	Hubel と Wiesel による階層型仮説	14
2.3	ネオコグニトロンの構造	16
2.4	S-, C-, V-細胞の結合関係	18
2.5	S-細胞によるパターンの検出	22
2.6	ネオコグニトロンの評価実験で用いる学習用パターン	26
2.7	ネオコグニトロンの回転したパターンに対する認識結果	27
3.1	心理学実験による回転物体の認識の研究	29
3.2	サルの下側頭皮質における回転パターンに対する反応	31
3.3	細胞面群の導入	33
3.4	回転対応型ネオコグニトロンの構造	35
3.5	辞書式順序集合 $A_l$ の要素に関する説明	37
3.6	シードセル以外の興奮性可変結合の強化方法	42
3.7	回転対応型ネオコグニトロンの $U_{S_2}$ 層における回転した局所 パターンの検出	45
3.8	C-細胞の回転ずれの許容範囲	47
3.9	回転対応型ネオコグニトロンの回転したパターンに対する認 識結果	49
3.10	手書き数字認識に対する評価実験に用いる学習用パターン	51
3.11	正しく認識できた手書き数字の一例	52
3.12	誤認識した手書き数字の一例	53
4.1	C-細胞と S-細胞の結合関係	57
4.2	S-細胞の発火条件 (1)	60
4.3	S-細胞の発火条件 (2)	61
4.4	ベクトル空間内における学習過程の考察	64
4.5	高速生成アルゴリズム	66
4.6	細胞面群融合法	68

4.7	高速生成アルゴリズムで生成した回転対応型ネオコグニトロンの認識率	71
4.8	通常のアゴリズムで生成した回転対応型ネオコグニトロンの認識率	72
4.9	高速生成アルゴリズムで生成された回転対応型ネオコグニロンと通常のアゴリズムで生成された回転対応型ネオコグニトロンの認識率の差	72
4.10	$U_{S2}$ 層における興奮性可変結合の類似度	74
4.11	$U_{S3}$ 層における興奮性可変結合の類似度	75
4.12	$U_{S4}$ 層における興奮性可変結合の類似度	76
5.1	心理実験による回転した文字の認識の研究	79
5.2	回転に対して単純なパターンを同定するための心理実験の様子	82
5.3	認識モデルと連想想起モデル	84
5.4	認識モデルと連想想起モデルの処理の流れ	86
5.5	正しく認識されたパターンの例	90
5.6	物体の認識および連想時における大脳皮質の反応の様子	92
5.7	認識モデルと連想モデルを融合したモデル	93
5.8	回転対応型ネオコグニロンと融合モデルの細胞数の比較	94

# 表 目 次

2.1	ネオコグニトロンの各細胞と結合の表記法 . . . . .	17
2.2	ネオコグニトロンで使用される記号と説明 . . . . .	21
2.3	細胞面数と細胞面内の細胞数に関するパラメータの値 . . . . .	26
2.4	結合野の大きさ . . . . .	26
2.5	S-細胞のしきい値および学習定数 . . . . .	26
3.1	回転対応型ネオコグニトロンの各細胞と結合の表記法 . . . . .	35
3.2	回転対応型ネオコグニトロンで使用される記号と説明 . . . . .	36
3.3	細胞面数と細胞面内の細胞数に関するパラメータの値 . . . . .	48
3.4	結合野の大きさ . . . . .	48
3.5	S-細胞のしきい値および学習定数 . . . . .	48
3.6	細胞面数と細胞面内の細胞数に関するパラメータの値 . . . . .	50
3.7	S-細胞のしきい値および学習定数 . . . . .	50
5.1	パターンの複雑性と心的回転依存性, およびネオコグニトロ ン型モデルとの対応. . . . .	80
5.2	細胞面群数, 細胞面数および細胞数 . . . . .	88
5.3	結合野 $A_l$ と $D_l$ の範囲 . . . . .	88

# 第1章 序論

## 1.1 本研究の背景

過去、ヒトからヒトへの情報は記録することで後世にその情報を残してきた。そこで使用される媒介物は主に紙と筆記用具であり、伝達すべき情報は主に手書の文字および活字で表現されてきた。音声による情報表現も記録や配布のために使用されているが、時間的コスト及び価格コストが文字によるものと比較して高いためにその量は文字によるもの程多くはない。ヒトは得られた情報を共有するために紙面に描かれた文字を認識し、理解する手法を選択してきたが、文字に限らず、情報の共有・伝達には常に情報の「記録・配布・認識」が伴うため、それら個々のコストを下げる研究が多くなされてきた。例えば本による情報の共有を行なうためには印刷による記録、完成した本の搬送による配布、実際に書かれた情報を得る文字認識が必要となる。情報の記録に関しては、活版印刷技術の普及により同一の情報を非常に短時間で文字として記録することができ、配布に関しては交通網の発展や近年の計算機ネットワークの普及により短時間・低価格で実現できるようになった。しかし近代においても複製された文字の認識を行なうのはあくまでもヒトであったため、認識に要するコストは記録・配布と比較すると非常に大きい。よって電子計算機が1940年代に登場して以来、逐次処理を基本とするノイマン型計算機に人間の知的活動の一つであるパターン認識の能力を計算機に持たせようとする試みが、多くの科学者にとって最も興味のある課題の一つとなったことは不思議なことではない。更に、近年の情報化社会の発展に伴ない前述の「記録・配布・認識」に要する時間的・価格コストを一層低下させるための計算機による文字認識技術の社会的重要性は一層高まってきている。

文字認識を含むパターン認識活動は、ヒトが通常最も頻繁に行なっている知的情報処理であり、ヒトを始めとする諸動物が外界から得た様々な情報を基にして、外界の様相を知覚・認識する機能をいう [1]。その機能は非常に高度であり、柔軟性に富み、様々な具体的場面において認識結果をほとんど無意識のうちに瞬時に引き出すことができる。パターン認識活動を文字の認識という場面に限定しても文字の色、テクスチャ、変形、大きさ、位置ずれ、拡大縮小、回転、ノイズ、飾り及び複雑な背景に影響されずに瞬時に認識結果が得られる。

文字認識の研究はこれら文字の多様性に影響されない認識機能を計算機に、

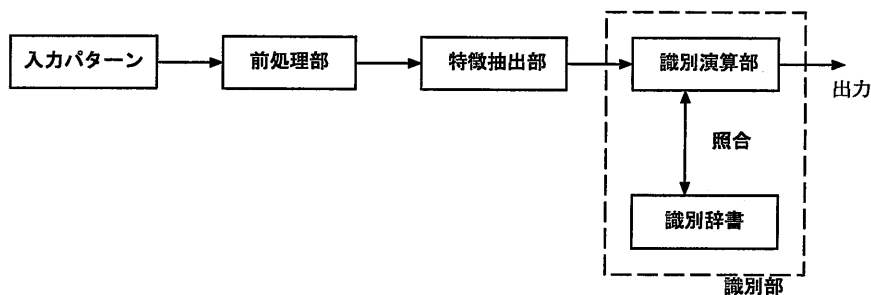


図 1.1: 代表的な工学的認識手法の流れ図. 前処理部ではあらかじめ作成しておいた識別辞書との整合性を向上させるために, 大きさの正規化やノイズ除去等を行なう. 前処理された入力パターンは特徴抽出部において高い識別能力を得るために設定された特徴の抽出を行なう. 一般に抽出された特徴はベクトル化され, 同様にベクトル化された辞書との類似度を計算することにより認識 (判別) を行なう. (文献 [2] より抜粋)

「工学的」に代用させる試みから出発した. すなわち実際のヒトの認識機能や機構を考慮せずに, 入力された文字を含む画像をあるクラス (概念) に対応付けするものである. 代表的な処理の流れを図 1.1 に示す. 実際「工学的」な文字認識の研究成果である OCR (Optical Character Reader: 光学的文字読み取り装置) は市場に出回るまでに発達しており, 発売当初と比較して認識率が格段に向上しただけではなく, 変形した文字や大きさの違う文字の混在というような, 様々な具体的場面に適用が可能となっている. しかし文字認識を含むパターン認識の問題が完全に克服されてはおらず, むしろこれまでの成果は当初目的とされてきたヒトの知的情報処理の実現とは全く異質なものに思われる. 事実, 現在実用化されている OCR は単に限定された環境下での文字当ての技術であり, 想定した環境下以外では全く機能せず, ヒトの知的情報処理との大きな違いとなっている. この問題を克服するために, 言語知識等を適用することで改善を試みる研究も音声認識を中心に数多くなされているが, これもやはりヒトが持つ知識獲得のための機能を無視する立場を取っており, 未だヒトの知的活動を完全に代用するには至っていない [3, 4]. これらいわゆる工学的なアプローチに共通する問題の原因はヒトの認識活動をほとんど無視したことにある. このアプローチにおいては研究者の対象 (文字や音声) に対する知識を計算機のプログラムに置き換えることで実現しているために, その知識を逸脱した環境下では全く機能しない.

一方, 電子計算機の登場と同時期の 1940 年代は, ヒトの知的活動, すなわち脳による知的情報処理を実現するためにはその基本原理を脳に求めることで, 知的処理機構を明らかにする必要があるという研究アプローチが登場した時期でもある. このアプローチは, 解剖学, 神経生理学, 情報数理学及



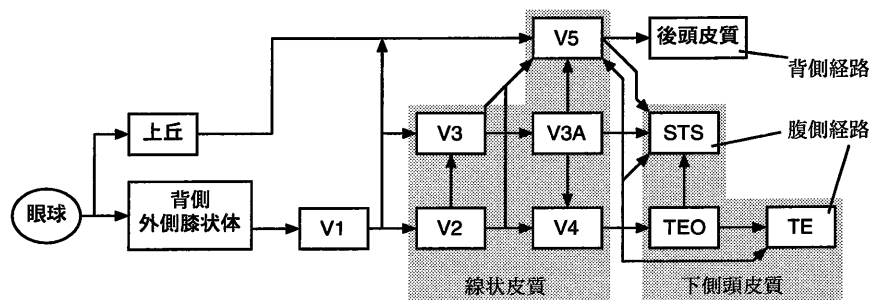


図 1.2: アカゲザルの視覚情報処理の部位および流れ。重要な経路と部位のみを示しており、省略している部分もある。矢印は視覚情報の流れの方向を示している。背側経路では物体の位置を（‘where’ part）、腹側経路では物体の種類を（‘what’ part）認識しているものと考えられている。文献 [5] の図 6.35 を改変。

び心理学等の、様々な分野の研究者に支持されており、現在も学際的な研究がなされている。

例えば神経生理学においては、脳による情報処理の基本構成要素となっている神経細胞（neuron）について様々な生理学的事実が明らかにされており、特に視覚に対する研究は盛んである。脳内における視覚情報の処理の経路はかなり明らかにされている。図 1.2. にその主な経路を図示する。個々の部位における神経細胞の特性についても数多くの研究がなされている。膨大な研究成果があるためここでは例を挙げないが、文献 [5] にまとまった記述があるので参考にさせていただきたい。

次に心理学における研究例を挙げる。心理学者の手法は様々な視覚パターンを被験者に見せ、それらがどのように見えるかを被験者に質問し、回答を得ることによって、観察者（被験者）の知覚的な経験を決める知覚的体制化の一般法則を見出すことにある [6]。その一例としてゲシュタルト派として知られている心理学者たちにより 1920 年代に行なわれたものがあるが、彼等の関心は知覚の際に個々の刺激がどのようにまとまって全体、つまりゲシュタルト（Gestalts）になるのかという点にあった。これらの研究から得られた法則には、近接の法則、同類の法則、運命共同の法則、よい連続の法則、閉合の法則などがある [7]。このような「見え」に関する研究の他に、「認識」に関する研究も盛んに行なわれている。Shepard と Metzler らは被験者に 3 次元物体を描いた線画対を提示した [8]。この対の一方は回転しており、被験者はこの 2 つの刺激を回転させてそれが同じものかどうかを決めなければならなかった。Shepard らは決定までに要する時間、すなわち認識時間を測定した。この実験より、ヒトは自分の脳内のイメージを実際に回転させていることを見だし、これを心的回転と呼んだ。また Cooper と Shepard は先の実

験を発展させ、結果として、ある種の認知的な課題を行なう際、被験者がイメージ的な記憶表象を作り変換しているという証拠を与えた [9, 10]。また回転した物体の認識に対する心理学的実験は [11] に詳しい。

情報数理学においては、ニューロンを数学的にモデル化して知的情報処理を実現し、工学的に応用しようとする研究が行なわれている。このような研究はバイオニクス（生物工学）という一つの研究分野となるまでに発展し、1960年には第1回のバイオニクス会議が開かれるまでになった。情報数理学の立場によるヒトの知的情報処理の解明を目指す研究は McCulloch と Pitts [12] による形式ニューロンモデルに始まる。彼らはニューロンを簡略化し、0-1の離散入出力関係から構成される論理演算子に対応付けした。Hebb は学習を神経細胞同士の結合の度合を表わすシナプスの可塑性が学習の効果をもたらすと仮定し、学習に関する基本原理を提案した [13]。この原理は非常に単純なものであり、ある入力を与えられたときにそれと同時に興奮している神経細胞間の結合を強化するというものである。Hebb の学習則は単純ではあるが現在の研究にも大きく影響しており、実際に神経生理学の研究によって彼の仮定が成立している神経細胞が発見されている。また、図 1.2 に示す各部位のモデル化およびシミュレーション実験も盛んに行なわれている。

以上のようにヒトの視覚情報処理を解明するために様々な実験が行なわれているが、実際にヒトの知的情報処理の一つである文字認識を計算機に代用させるためには学際的なアプローチが必要となってくる。すなわち神経生理学で得られた処理の基本単位となる個々のニューロンの特性および構成、心理学で得られた包括的な脳の機能、情報数理学による数値計算機への実装方法および工学的意味、「工学的」アプローチによるパターン認識理論、これらを組み合わせてはじめて工学的・社会的意味のある知的な計算機の構成が可能となる。

## 1.2 本研究の目的

文字認識を行なう脳の数理モデルに関する研究は多くないが、例を以下に挙げる。Fukushima によって提案されたネオコグニトロンは、階層型神経回路網であり、提示された文字を学習することで文字認識を行なう神経回路モデルである [14, 15, 16]。このモデルは Hubel と Wiesel の一連の研究 [17, 18, 19] による第1次視覚野の単純型細胞と複雑型細胞の階層型仮説を基にして構成された。ネオコグニトロンは  $U_1$  細胞段と呼ばれる低次の層に第1次視覚野で行なわれている線分の抽出機構である Hubel らの階層型仮説を取り入れたが、更に複雑な図形を抽出している  $V_2$ ,  $V_4$ ,  $TEO$ ,  $TE$  野の経路に至っても同様の階層型仮説が成立していると考え、 $U_2$  細胞段以降の高次の層でも  $U_1$  細胞段と同様の階層型の構造を持つモデルを構成した。これは非常に大胆な仮説であり、実際、神経生理学的にはこの階層型構造の仮説は部分的に否定

されている [20]. しかし神経細胞網の情報処理の機能としてこのモデルが否定されたわけではない [21]. Fukushima は実際に手書き数字および手書アルファベットのデータベースを用いて、ネオコグニトロンが大きさ、位置ずれ、拡大縮小に影響されない認識を行なえることを確認した。しかし、文字の回転に対する耐性は考察されていない。

その他の神経回路モデルとしては Reid らによって提案されたモデル [22], Widrow らによって提案されたモデル [23] がある。これらのモデルは神経回路網を文字パターン識別器として工学的に適応したものである。これらの神経回路モデルには回転した文字パターンを認識させる際に神経細胞モデルの組み合わせの数が爆発的に増大するといった問題や、 $90^\circ$  毎に回転したパターンしか認識できないという問題点がある。また Fukumi らによって提案されたモデルは特徴抽出を行なう機構の構成に心理学的アプローチを取っている非常に興味深いモデルであり、回転したパターンの認識を効率的に行なうだけでなく、角度の推定も可能とするモデルである [24]。しかし Fukumi らのモデルは変形したパターンの認識は不可能であるという欠点がある。同様に心理学的考察を基に構成されたモデルとして重なり合った複数のパターンを同時に認識することができるモデルが Basak らによって提案されている [25]。しかしこのモデルはパターンの変形に対する頑健性はない。

本研究では上記の数ある神経回路モデルのうち、生理学的に妥当であり工学的に有効なネオコグニトロンに注目し、個別の文字認識を行なう数理モデルを構築する際の基礎とする。また文字の多様性に起因する主な変化には、変形、大きさ、位置ずれ、拡大縮小、回転、ノイズが挙げられるが、これまでの文字認識に関する研究ではこれらの一部にしか頑健性がない。例えばネオコグニトロンは回転に対して頑健性がなく、Fukumi らのモデルは変形に対して頑健性がない。そこで本研究では上記の変形、大きさ、位置ずれ、拡大縮小、回転、ノイズの全てに頑健な数理モデルを構築することを目的とする。また、モデルの学習に要する時間の大幅な削減、細胞数の減少といった工学的な要求にも答えることも目的としている。また、近年、視覚情報処理を行なう脳内の細胞レベルおよび処理の流れに関する研究は特に盛んであり様々な成果が得られているが、文字認識を行なう数理モデルを構成するための十分な知見は未だ得られていない。そこで本研究では、ヒトの認識機構で明かにされていない事項に関しては心理学的実験を行なうことで新しいモデルの構築の際の足掛かりにする。本研究の目的と方針をまとめると以下の項目となる。

- ネオコグニトロンを基礎とし、神経生理学、心理学および情報数理学で得られた知見を考慮しつつ、ヒトが行なう個別文字認識を実現する数理モデルを構成する。数理モデルには以下のものを満すようにする。
  - 提案する数理モデルは文字の変形、位置ずれ、拡大縮小、回転、ノイズに影響されないモデルにする。

– 工学的要求を満たすモデルおよび学習方法を構築する。

- 得られた数理モデルを数値計算機に実装しその有効性を確認する。
- これまでの研究で明らかにされていない認知現象がある場合には新たに心理実験を行なう。
- 得られた心理実験の結果から新たな数理モデルを構成し、数値計算機に実装しその有効性を確認する。

本研究では個別文字認識に議論を限定しているが同様の研究アプローチを取ることで、文字の切り出し、注視、連想といった、ヒトが持つより高度な知的視覚情報処理が実現できるものと考えられる。すなわち本研究は知的視覚情報処理機構を明らかにするための基礎的研究と位置付けている。

### 1.3 本稿の構成

第1章 序論であり、研究の背景および目的を述べた。

第2章 ネオコグニトロンの説明をし、回転パターンに対する評価実験を行なう。

第3章 ヒトの回転したパターンの認識に関する考察を行ない、ネオコグニトロンの変形に対する頑健性を回転に対して適応した回転対応型ネオコグニトロンを提案する。また数値シミュレーションにより評価、考察を行なう。

第4章 学習時間を大幅に削減する新しい高速生成アルゴリズムを提案し、評価、考察を行なう。

第5章 新しく行なった心理実験の結果を基に、トップダウン回路を持つ数理モデルを提案する。

第6章 結論であり、本研究の成果と今後の課題について述べる。

## 第2章 ネオコグニトロン

本研究では前章の考察より、視覚情報を統合し認識する数理モデルの基礎としてネオコグニトロンに着目する。本章ではネオコグニトロンの性質、構成、各種神経細胞の数式表現、学習方法について述べ、文字の多様性、特に回転に対する頑健性を調査することを目的とする。

### 2.1 単純型細胞と複雑型細胞の階層型仮説

ネオコグニトロンは Hubel と Wiesel による階層型仮説に従いモデルを構成しているため、本節ではまずこの仮説について説明する。Hubel と Wiesel らの視覚皮質に関する革命的な研究 [17, 18] によると第 1 次視覚野の皮質内には左右の眼に由来する眼優位性コラムが交互に存在し、この眼優位性のコラム内には特定の方位に対して選択的に反応するニューロンによって形成された方位選択性コラムが存在する。また異なる特性を持つ単純型細胞 (simple cell) および複雑型細胞 (complex cell) が存在し、それぞれ非常に興味深い性質を持つことが明らかになった。図 2.1 を用いて方位選択性を持つ神経細胞の特性について説明する。方位選択性とはある特定の方位を持つ線分のみ反応 (発火) する性質という意味である。その方位は個々の細胞で異なり。例えば図 2.1(a) の黒い線分で示される方位に対して選択的に反応する単純型細胞は、灰色の線分で示される方位に対しては反応しない。例えば図 2.1(b-1) に示すような網膜からの入力に対して反応する単純型細胞は、(b-2) に示す入力には反応しない。また方位と位置が同じでも位相がずれている入力に対しても反応しない。一方、方位選択性を持つ複雑型細胞は入力的位置と位相には影響されずに反応する。例えば、図 2.1(c-1) の入力に対して反応する複雑型細胞は、(c-2) および (c-3) の入力に対しても反応する。また視覚に関係する神経細胞は網膜上の全ての神経細胞の出力に影響されているのではなく、網膜上のある一部の領域に存在する神経細胞の出力にのみ依存して反応する。この領域を受容野と言い、図 2.1 では円で表している。

また、Hubel と Wiesel は単純型細胞と複雑型細胞が図 2.2 に示すような簡単な階層構造を形成していると仮定した。これを単純型細胞と複雑型細胞の階層型仮説と言う。Hubel と Wiesel は複雑型細胞が、方位選択性は同じであるが受容野が互いに異なる複数の単純型細胞から入力を受けているために上記に示した特性が複雑型細胞に現れると仮定した。図 2.2(a) ではある複

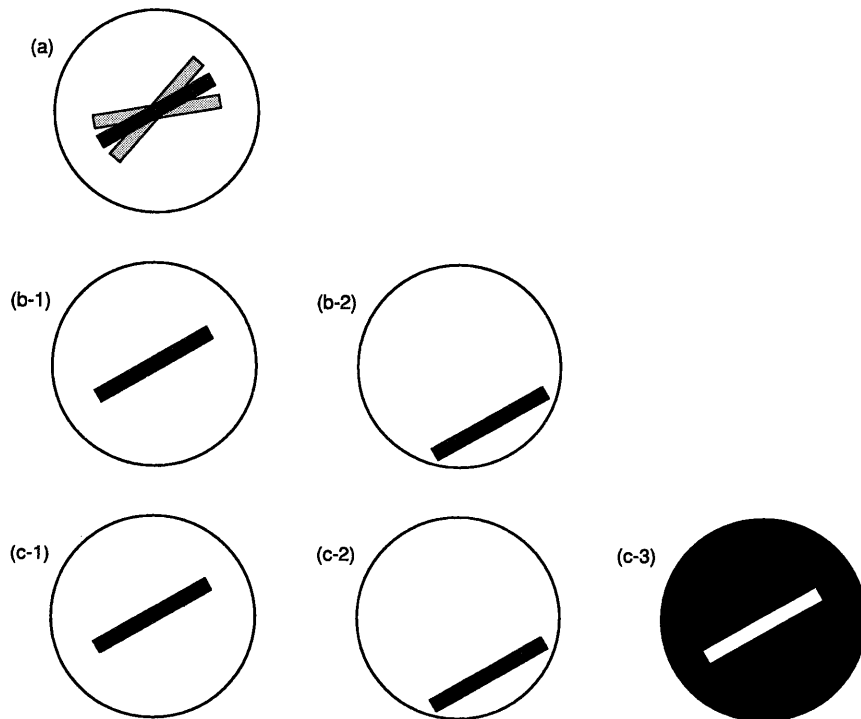


図 2.1: 方位選択性を持つ単純型細胞と複雑型細胞の特性. (a): 方位選択性を持つ第一次視覚野の神経細胞はある特定の方位を持つ線分のみ反応（発火）する. 例えば黒い線で示す方位の線分には反応するが, 灰色の線で示す方位の線分には反応しない. 単純型細胞は網膜上の線分の方位と位置に敏感に反応する. 例えば (b-1) に示す線分に反応する単純型細胞は, (b-2) に示す線分には反応しない. 一方, 複雑型細胞は線分の方位には敏感であるが, 位置および位相によらず反応する. 例えば複雑型細胞は (c-1), (c-2), (c-3) で示す線分には全て反応する.

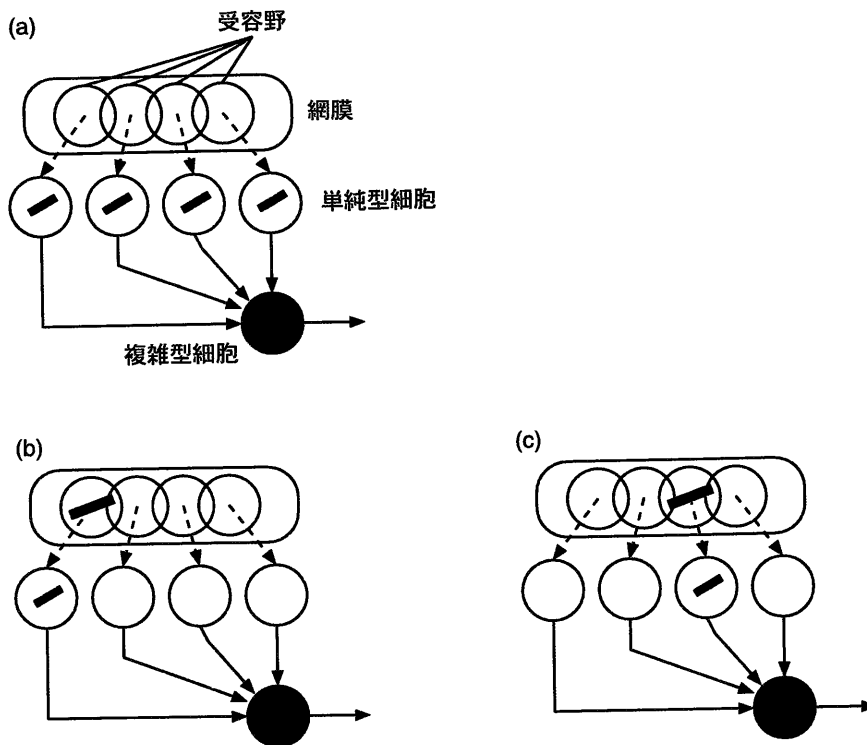


図 2.2: (a): Hubel と Wiesel による単純型細胞と複雑型細胞の階層型仮説. 単純型細胞は網膜上の受容野内の光受容器から入力を受ける. 複雑型細胞は方位選択性は同じであるが受容野が互いに異なる複数の単純型細胞から入力を受ける. 階層型仮説によると, 複雑型細胞は複数の単純型細胞からの入力を受けているので, その中の一つの単純型細胞が発火するだけで複雑型細胞が発火する. 複雑型細胞は, (b): 網膜上の左側に線分が提示された場合, (c): 右側に線分が提示された場合のいずれの場合においても発火することが可能になる. この仮説により複雑型細胞が入力の位置によらずに発火するという特性が説明できる. 位相に関する特性も同様の仮説で説明できる.

雑型細胞が  $20^\circ$  の方位選択性を持つ4つの単純型細胞から入力を受けている。この時、図 2.2(b) のように網膜上の左側に  $20^\circ$  の線分が提示されていても、また図 2.2(c) のように網膜上の右側に  $20^\circ$  の線分が提示されていてもこの複雑型細胞は発火する。これは複雑型細胞の反応が受容野内であれば入力的位置に依存しないことを説明するものである。当然、図中の4つの単純型細胞以外で同じ方位選択性を持つ細胞が発火しても、この複雑型細胞は発火しない。これは複雑型細胞の受容野の特性を説明するものであり、図 2.2 に示す複雑型細胞は横長の受容野を持つことになる。位相に関する特性についても同様の階層型仮説で説明できる。

## 2.2 ネオコグニトロンの性質および構造

### 2.2.1 ネオコグニトロンの構造

ネオコグニトロンは多層の神経回路網であり、教師なし学習および教師あり学習によって構造を自動的に獲得する数理モデルである。このモデルは学習時に使用した文字（学習パターン）だけではなく、大きさが異なる文字、変形、位置ずれ、ノイズが存在する未学習文字に対しても頑健に認識することができるモデルとして知られている。本研究ではパターンを文字に限定しているため、文字を単にパターンと呼ぶことにする。図 2.3 にネオコグニトロンの構造を図示する。学習パターン、および評価時に使用されるテストパターンは、網膜のモデルである  $U_0$  層に提示される。 $U_0$  層以降には複数の細胞層が配置されており、 $U_S$  層と  $U_C$  層が交互に並んでいる。 $U_S$  層と  $U_C$  層をまとめて細胞段と呼び、 $U_{S_l}$  層と  $U_{C_l}$  層をまとめて第  $l$  細胞段と呼ぶ ( $U_l$  と記す)。図 2.3 では4段、8層の細胞層で構成されるネオコグニトロンを示している。段数は認識するパターンの複雑さおよび解像度によって異なる。例えば英数字の認識を目的とする場合は一般に4段構造をなし、漢字の認識を目的とする場合はより多くの段数を必要とする。

各層には細胞面を複数含んでおり、 $U_S$  層の細胞面を S-細胞面、 $U_C$  層の細胞面を C-細胞面と呼ぶ。また S-細胞面には特徴抽出細胞である S-細胞、C-細胞面には位置ずれ許容細胞である C-細胞が2次元平面内に配置されている。その他に V-細胞が存在するが図 2.3 では省略してある。

$U_{S_l}$  層の S-細胞は  $U_{C_{l-1}}$  層の C-細胞と興奮性可変結合を形成している。また、この S-細胞は V-細胞と抑制性可変結合を形成している。これら可変結合の強度は学習時に自動的に決定される。

図 2.3 において最終層である  $U_{C_4}$  層の C-細胞の発火頻度は  $U_0$  層に提示されたパターンの認識結果を示す。例えば0から9の数字を認識する場合には  $U_{C_4}$  層には10個の C-細胞が配置されており、発火している C-細胞に対応するクラスが認識結果を示す。



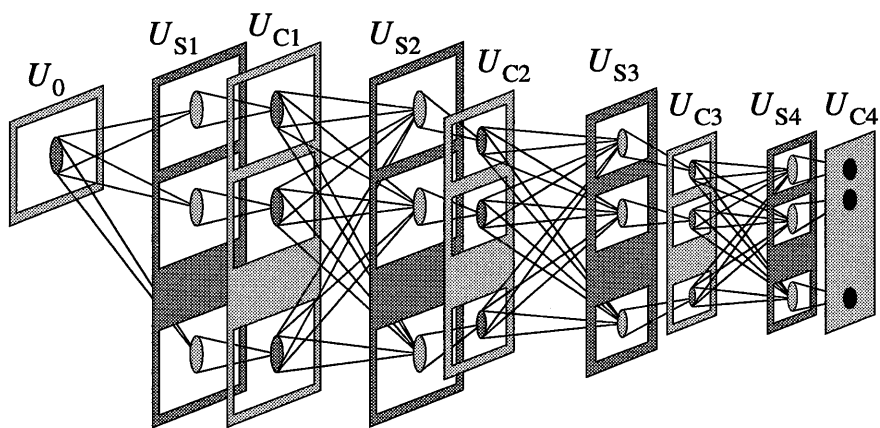


図 2.3: 一般的なネオコグニトロンの構造. 多層の神経回路網となっている.  $U_0$  層は網膜のモデルであり,  $U_0$  層に続いて数段の細胞層から構成される. 図では 8 つの細胞層から構成されており,  $U_S$  層と  $U_C$  層が交互に並んでいる.  $U_S$  層と  $U_C$  層をまとめて細胞段と呼び,  $U_{S1}$  層と  $U_{C1}$  層をまとめて第 1 細胞段と呼ぶ. 各層には透明の四角形で描かれている細胞面を含んでいる.  $U_S$  層の細胞面を S-細胞面,  $U_C$  層の細胞面を C-細胞面と呼ぶ. また S-細胞面には特徴抽出細胞である S-細胞, C-細胞面には位置ずれ許容細胞である C-細胞が 2 次元平面内に配置されている. その他に V-細胞が存在するが省略してある.

表 2.1: ネオコグニトロン各細胞と結合の表記法

表記法	意味
$U_{Cl-1}(n, \kappa)$	第 $l-1$ 細胞段の C 層で, 細胞面の番号が $\kappa$ であり, 細胞面内の位置 $n$ に配置してある C-細胞. $n = (n_x, n_y) \in \mathbf{Z}^2$
$U_{Vl}(n)$	第 $l$ 細胞段内で, 位置 $n$ に配置してある V-細胞.
$U_{Sl}(n, k)$	第 $l$ 細胞段の S 層で, 細胞面の番号が $k$ であり, 細胞面内の位置 $n$ に配置してある S-細胞.
$a_l(\nu, n, \kappa, k)$	$U_{Cl-1}(n + \nu, \kappa) - U_{Sl}(n, k)$ 間の興奮性可変結合
$b_l(k)$	$U_{Vl}(n) - U_{Sl}(n, k)$ 間の抑制性可変結合
$c_l(\nu)$	$U_{Cl-1}(n + \nu, \kappa) - U_{Vl}(n)$ 間の興奮性固定結合
$d_l(\nu)$	$U_{Sl}(n + \nu, k) - U_{Cl}(n, k)$ 間の興奮性固定結合

次節より各細胞間の結合関係を詳細に説明する.

### 2.2.2 細胞間の結合法則

S-, C-, V-細胞はそれぞれ細胞面内に 2 次的に配置されているが, ここでは説明のために各細胞が 1 次的に配置してあるものとして図 2.4 を用いる (一般性は失なわれない). 各細胞は次の 4 種類の情報で特定される.

- 細胞の種類 (S, C, または V),
- 細胞が属する層番号  $l$ ,
- 細胞が配置してある 第  $k$  細胞面,
- 細胞面内の位置  $n$ .

これらの情報を用いた各細胞の表記法を図 2.1 に示す. 各細胞の出力 (発火頻度) は “ $u$ ” で表現する. 例えば 細胞  $U_{Cl-1}(n, \kappa)$  の出力は  $u_{Cl-1}(n, \kappa)$  となる.

図 2.4(a) は C-S 細胞間および V-S 細胞間の結合を図示している<sup>1</sup>. S-細胞  $U_{Sl}(n, k)$  は前層の C-細胞と同層の V-細胞と結合している. S-細胞  $U_{Sl}(n, k)$  と C-細胞  $U_{Cl-1}(n + \nu, \kappa)$  間の興奮性可変結合は  $a_l(\nu, n, \kappa, k)$  で示される. V-細胞  $U_{Vl}(n)$  は  $U_{Cl-1}(n + \nu, \kappa)$  と興奮性固定結合を構成しており結合強度は  $c_l(\nu)$  で示される.  $U_{Vl}(n)$  は  $U_{Sl}(n, k)$  に抑制性の信号を送っており,

<sup>1</sup>「C-S 細胞間」とは情報の流れが C-細胞から S-細胞へと流れるとき, その C-細胞と S-細胞の間という意味である.

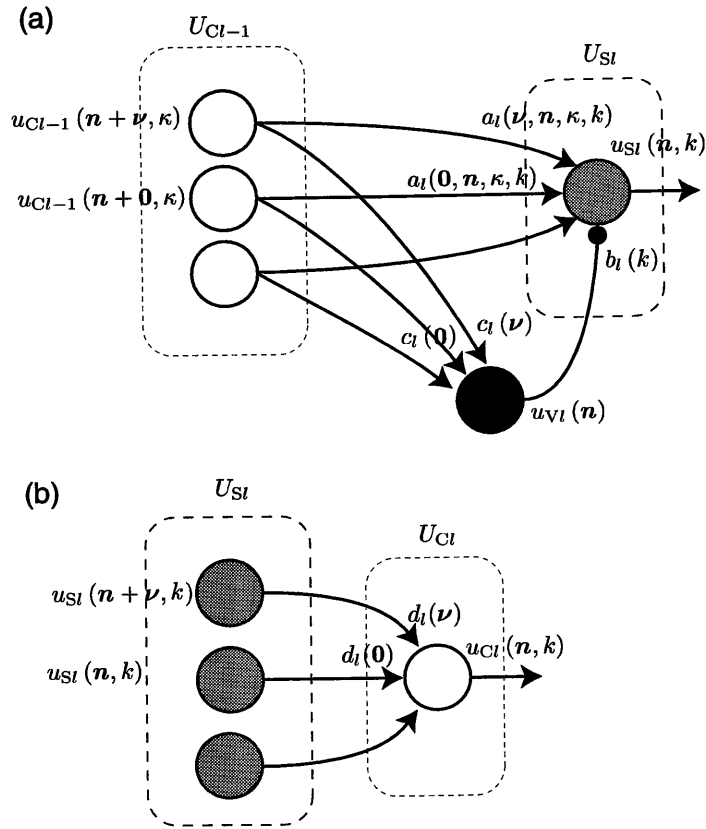


図 2.4: S-, C-, V-細胞の結合関係. 実際には各細胞は 2 次的に配置してあるが, 説明のために 1 次的に表示している (一般性は失われない). “U” は細胞の「名前」を意味し, “u” はその細胞の「出力 (発火頻度)」を意味する. (a): C-S 細胞間および V-S 細胞間の結合. S-細胞  $U_{Sl}(n, k)$  は前層の C-細胞と同層の V-細胞と結合している. S-細胞  $U_{Sl}(n, k)$  と C-細胞  $U_{Cl-1}(n+\nu, \kappa)$  間の興奮性可変結合は  $a_l(\nu, n, \kappa, k)$  で示される. V-細胞  $U_{Vl}(n)$  は  $U_{Cl-1}(n+\nu, \kappa)$  と興奮性固定結合を構成しており結合強度は  $c_l(\nu)$  で示される.  $U_{Vl}(n)$  は  $U_{Sl}(n, k)$  に抑制性の信号を送っており, 結合強度は  $b_l(k)$  である. (b): S-C 細胞間の結合. C-細胞  $U_{Cl}(n, k)$  は S-細胞  $U_{Sl}(n+\nu, k)$  と  $d_l(\nu)$  で示される興奮性固定結合を構成している.

結合強度は  $b_l(k)$  である。図 2.4(b) は S-C 細胞間の結合を図示している。 $U_{Cl}(n, k)$  は  $U_{Sl}(n + \nu, k)$  と  $d_l(\nu)$  で示される興奮性固定結合を構成している。各結合の表記法についても表 2.1 にまとめてあるので参照していただきたい。

### 2.2.3 各細胞の数式表現

前節までに、各細胞の表記と結合の関係を述べた。ここでは各細胞の入出力関係を記述する。

#### S 細胞

$U_{Sl}(n, k)$  細胞の出力  $u_{Sl}(n, k)$  は次式で与えられる。

$$u_{Sl}(n, k) = r_l \cdot \phi \left[ \frac{1 + e}{1 + \frac{r_l}{1 + r_l} \cdot i} - 1 \right], \quad (2.1)$$

$$\phi[x] = \max(x, 0), \quad (2.2)$$

ここで関数  $\phi$  は S-細胞の出力関数であり電子回路の分野における半波整流関数の意味を持つ。 $r_l$  はしきい値の効果を与える正の値である。 $e$  は C-細胞 ( $u_{Cl-1}(n + \nu, \kappa)$ ) から S-細胞への興奮性入力との総和、 $i$  は V-細胞 ( $u_{Vl}(n)$ ) からの抑制性入力をそれぞれ表わし、次式で与えられる。

$$e = \sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} a_l(\nu, n, \kappa, k) \cdot u_{Cl-1}(n + \nu, \kappa), \quad (2.3)$$

$$i = b_l(k) \cdot u_{Vl}(n). \quad (2.4)$$

ここで  $K_{Cl-1}$  は  $U_{Cl-1}$  層における細胞面の数である。 $A_l$  は  $U_{Sl}$  の S-細胞が結合する  $U_{Cl-1}$  層の C-細胞面内の範囲を示すの集合である。これを「結合野」と呼ぶことにする<sup>2</sup>。 $A_l \subset Z^2$  である。ただし  $Z$  は整数の集合である。例えば  $|A_l| = 9$  の場合は ( $|\cdot|$  は要素数を表わす) 前層の  $9 \times K_{Cl-1}$  個の C-細胞と結合することになる。式 (2.1), (2.3), (2.4) および図 2.4 からわかるように S-細胞は V-細胞および前層の C-細胞から入力を受けている<sup>3</sup>

S-細胞はパターン検出細胞であり、Hubel と Wiesel によって発見された単純型細胞 (simple cell) の数理モデルとなっている。従って、検出すべきパターンの位置が少しでもずれると発火しなくなる (代わりにずれた方向に位置する S-細胞が発火する)。

<sup>2</sup>結合野は受容野とは異なる。受容野は網膜上の結合範囲である。

<sup>3</sup>ネオコグニトロンでは、同一の細胞面にある S-細胞は全て同じ空間分布をもつ。よって S-細胞の興奮性可変結合は位置  $n$  に依存しないため  $a_l(\nu, \kappa, k)$  と記すべきである。実際、文献 [14, 15, 16, 26] 等では  $n$  に依存しない表記方法を使用しているが、本稿では後に提案する新しい数理モデルとの整合性を取るためにあえて  $n$  に依存する表記を用いた。

## V-細胞

$U_{Sl}$  層の S-細胞に抑制性の信号を送っている V-細胞  $U_{Vl}(\mathbf{n})$  の出力  $u_{Vl}(\mathbf{n})$  は次式で与えられる.

$$u_{Vl}(\mathbf{n}) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} c_l(\nu) \cdot \{u_{Cl-1}(\mathbf{n} + \nu, \kappa)\}^2}, \quad (2.5)$$

ここで  $c_l(\nu)$  は興奮性固定結合の強度を表わす.  $U_{Vl}$  細胞は  $U_{Cl-1}$  層の C-細胞から入力を受けている.

## C-細胞

C-細胞  $U_{Cl}(\mathbf{n}, \kappa)$  は  $U_{Sl}$  層内の S-細胞から入力を受けている. ただし S-細胞が前層の全ての細胞面内に位置する C-細胞から入力を受けていたのに対し, C-細胞は同層の同じ細胞面番号  $\kappa$  を持つ S-細胞からのみ入力を受ける.  $u_{Cl}(\mathbf{n}, \kappa)$  は次式で与えられる.

$$u_{Cl}(\mathbf{n}, \kappa) = \psi \left[ \sum_{\nu \in D_l} d_l(\nu) \cdot u_{Sl}(\mathbf{n} + \nu, \kappa) \right], \quad (2.6)$$

$$\psi[x] = \frac{\phi[x]}{1 + \phi[x]}, \quad (2.7)$$

ここで  $d_l(\nu)$  は興奮性固定結合の強度を表わし,  $D_l$  は  $U_{Cl}$  層の C-細胞が  $U_{Sl}$  層の S-細胞と結合する範囲を示すの集合である.  $D_l \subset \mathbf{Z}^2$  である.

C-細胞は Hubel と Wiesel による階層型仮説に従って構成されている. 図 2.2 の階層型仮説と, 図 2.4(b) のネオコグニトロン of S-細胞と C-細胞の結合関係が対応していることに注意していただきたい. C-細胞は複雑型細胞 (complex cell) と似た性質を持ち, 検出すべきパターンの位置が  $D_l$  で定められる結合野内で移動しても発火しつづける.

表 2.2 にネオコグニトロンで使用される記号についてまとめてあるので参照していただきたい.

### 2.2.4 細胞面

ネオコグニトロンの重要な構成要素の一つに細胞面 (cell-plane) がある. 前節までに細胞面が 1 つの層に複数存在することを説明したが, ここでは細胞面の性質についてより詳しく説明する. 以下に細胞面の性質を列挙する.

- 細胞面内には複数の細胞が 2 次元平面上に配置されている. 例外的に最終層内の細胞面内には C-細胞が 1 つだけ存在する.

表 2.2: ネオコグニトロンで使用される記号と説明

表記法	意味
$K_{S_l}$	$U_{S_l}$ 層内の細胞面の数. 細胞面の番号は $k$ (または $\kappa$ ) で参照する.
$K_{C_l}$	$U_{C_l}$ 層内の細胞面の数. 一般に $K_{C_l} = K_{S_l}$ である. 細胞面の番号は $\kappa$ (または $k$ ) で参照する.
$A_l$	$U_{S_l}$ 層内の S-細胞が $U_{C_{l-1}}$ 層内の C-細胞と結合する範囲を与える集合. 集合の要素 $\nu$ は 2次元相対位置座標 ( $\nu \in \mathbb{Z}^2$ ) である.
$D_l$	$U_{C_l}$ 層内の C-細胞が $U_{S_l}$ 層内の S-細胞と結合する範囲を与える集合. 集合の要素 $\nu$ は 2次元相対位置座標 ( $\nu \in \mathbb{Z}^2$ ) である.

- ある細胞面内に存在する細胞が持つ前層との結合強度は全て空間的に同一である. これは次の 3 点を意味する.
  - ある細胞面内に存在する細胞は全て同じパターンを検出する.
  - 異なる細胞面内にある細胞は異なるパターンを検出する.
  - 検出すべき局所パターンが移動すると, 発火している細胞もそれに対応して移動する.

図 2.5 に図示するように各細胞は細胞面内の 2次元平面上に配置されている. 実際には各層に複数の細胞面があるが図では 1つの細胞面だけを図示している (実際には  $U_{S_l}$  層の S-細胞は  $U_{C_{l-1}}$  層内の  $K_{C_{l-1}}$  個の細胞面内の C-細胞と結合している).  $U_{S_l}$  層の S-細胞は,  $A_l$  で定義される局所的範囲内に位置する  $U_{C_{l-1}}$  層の C-細胞と結合している. ある細胞面内の S-細胞の結合強度は全て空間的に同一である. 例えば図 2.5 に示す  $U_{S_l}$  層の S-細胞は T-型の結合分布をした興奮性可変結合,  $a_i(\nu, n, \kappa, k)$ , で前層の C-細胞と結合している. すなわち図 2.5 の S-細胞は T-型の局所的パターンを検出する細胞の集団であり, 発火位置が検出された T-型パターンの位置に対応する. 異なる細胞面内の S-細胞は, T-型とは異なる結合分布をしているため, 違うパターンを検出する (例えば L-型). 各細胞面の結合分布は一般に教師なし学習によって自動的に決定される.

$U_{S_l}(n, k)$  と同一の細胞面の番号  $k$  をもつ C-細胞,  $U_{C_l}(n, k)$  もまた, 同じ局所パターンを検出する. 例えば図 2.5 の C-細胞  $U_{C_l}(n, k)$  も T-型の局所パターンに対して発火する. ただし検出すべき局所パターンの位置が  $D_l$  で定められる範囲内ですべて発火しつづける. これは C-細胞が複雑型細胞の性質を持つことを意味する. C-細胞が複雑型細胞の性質を持つことは重要であり, ネオコグニトロンが位置ずれや変形に頑健である理由の一つである.

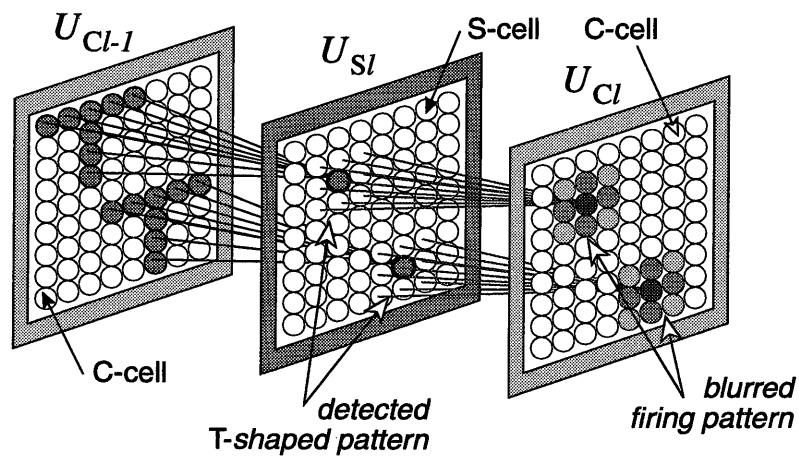


図 2.5: S-細胞によるパターンの検出. 簡単のために V-細胞は省略してある.  $U_{Sl}$  層の S-細胞は,  $A_i$  で定義される局所的な範囲内に位置する  $U_{Cl-1}$  層の C-細胞と結合している. ある細胞面内の S-細胞の結合強度は全て空間的に同一である. 例えば図に示す  $U_{Sl}$  層の S-細胞は T-型の結合分布を持つ. すなわち図中の S-細胞は T-型の局所的なパターンを検出する細胞の集団であり, 発火する位置が T-型パターンの位置を表現する. 同一の細胞面の番号をもつ C-細胞もまた, 同じ局所パターンを検出する. ただし検出すべき局所パターン, ここでは T-型パターンの位置が  $D_i$  で定められる範囲内でずれても発火しつづける.

## 2.2.5 特徴検出と位置ずれ許容

ネオコグニトロンでは  $U_S$  層と  $U_C$  層が階層構造をなしているため、

- S-細胞による位置に敏感な局所的なパターンの検出、
- C-細胞による位置ずれを許容する局所的なパターンの検出、

の2種類の操作が繰り返し行なわれる。よって各細胞は後層になるに従い、

- S-細胞が検出するパターンはより複雑なものになり、
- C-細胞による位置ずれの範囲はより広くなる

という性質を持つ。故に、最終層のC-細胞の受容野は  $U_0$  全体となり、検出するパターンは英数字のような複雑なものになる。

## 2.3 学習方法

ネオコグニトロンの学習とは、学習用パターンを  $U_0$  層に提示することで可変結合を強化し、モデルの構造を変化させる過程を指す。 $U_{S1}$  層および  $U_{C1}$  層は線分の方位を検出するようにあらかじめ各可変結合の強度を決定しておく、実際の学習は  $U_2$  細胞段以降で行なわれる。学習は前層より順に進められていく。例えば  $U_{S3}$  層の学習は、 $U_{S2}$  層の学習が終了したあとに行なわれる。学習前には  $U_2$  細胞段以降に細胞面は1つも存在せず、学習過程において細胞面が自動的に生成される。すなわち、検出すべき局所パターンは学習によって自動的に決定される。ネオコグニトロンにおける学習アルゴリズムをアルゴリズム1に示す。

検出すべき局所パターンの選択はネオコグニトロンの教師なし学習の1つである「シードセル生成面を用いた学習方法」を用いるのが一般的である [27]。シードセルとは winner-take-all 型の学習則における winner-cell に相当し、実際の学習、すなわち可変結合の強化はただ1つの細胞（シードセル）のみに適応される。シードセルが  $U_{Si}(\hat{n}, \hat{k})$  であるとする可変結合の1回の学習における可変結合の変化量は以下の式で示される。

$$\Delta a_l(\nu, \hat{n}, \kappa, \hat{k}) = q_l \cdot c_l(\nu) \cdot u_{Cl-1}(\hat{n} + \nu, \kappa), \quad (2.8)$$

$$\Delta b_l(\hat{k}) = q_l \cdot u_{Vl}(\hat{n}), \quad (2.9)$$

ただし  $q_l$  は学習定数で正の値をとる。実際に学習を行なうのはシードセルのみであるが、シードセルが属している細胞面内にある他のS-細胞の興奮性可変結合は、シードセルの興奮性可変結合の変化量、 $\Delta a_l(\nu, \hat{n}, \kappa, \hat{k})$ 、と同じ量だけ強化される。言い換えると、シードセルで得られた興奮性可変結合は



---

**アルゴリズム 1** ネオコグニトロンの学習方法

---

**Require:** All the S-cell layers up to  $U_{Sl-1}$  are already reinforced.

- 1: **while** New training patterns still exist **do**
  - 2:   Input a new training pattern to  $U_0$ .
  - 3:   Calculate responses of the cells in the preceding layer  $U_{Cl-1}$ .
  - 4:   Make seed-cells using seed-selecting plane.
  - 5:   Insert positions of the seed-cells in a queue,  $Q$ .
  - 6:   **while**  $Q$  is not empty **do**
  - 7:      $\hat{n} \leftarrow$  dequeue  $Q$ .
  - 8:     Search  $\hat{k}$  such that  $u_{Sl}(\hat{n}, \hat{k}) \geq u_{Sl}(\hat{n}, k)$  for any  $k$
  - 9:     **if**  $U_{Sl}(\hat{n}, \hat{k})$  is firing, that is,  $u_{Sl}(\hat{n}, \hat{k}) > 0$  **then**
  - 10:       Reinforce  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  and  $b_l(\hat{k})$  subject to (2.8) and (2.9).
  - 11:       Reinforce  $a_l(\nu, n, \kappa, \hat{k})$  subject to (2.10) for all  $n \neq \hat{n}$
  - 12:     **else**
  - 13:       Make a new cell-plane stack
  - 14:        $\hat{k} \leftarrow$  the serial number of the new cell-plane.
  - 15:       Reinforce  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  and  $b_l(\hat{k})$  subject to (2.8) and (2.9).
  - 16:       Reinforce  $a_l(\nu, n, \kappa, \hat{k})$  subject to (2.10) for all  $n \neq \hat{n}$
  - 17:     **end if**
  - 18:   **end while**
  - 19: **end while**
-

同一細胞面内にある他の S-細胞にコピーされることになり、以下の式で示される。

$$\Delta a_l(\nu, n, \kappa, \hat{k}) = \Delta a_l(\nu, \hat{n}, \kappa, \hat{k}) \quad n \neq \hat{n} \quad (2.10)$$

式 (2.10) は前述の細胞面の性質である、「細胞面内における S-細胞の興奮性可変結合の空間同一性」を保つことを意味する。

## 2.4 回転パターンに対する評価実験

ネオコグニトロンが変形、位置ずれ、拡大縮小、ノイズに対して頑健であることは既に多数の実験により知られている [14, 15, 16, 27, 28, 29, 30]。本節ではネオコグニトロンが正立した（回転していない）学習用パターンを学習しただけで、どの程度回転に対する頑健性があるかを数値シミュレーションにより調査する。使用する学習パターンは図 2.6 に示す 6 通りのパターンである。細胞段は 4 段（8 層）であり、細胞数に関連するパラメータ値は表 2.3、結合に関するパラメータ値は表 2.4、しきい値および学習定数は表 2.5 にそれぞれ示す。

学習後、図 2.6 の学習用パターンを Alan らによって提案されたアルゴリズム [32] を用いて計算機上で回転させた評価用パターンを提示し、認識結果を調査した。結果を図 2.7 に示す。図 2.7 の灰色で囲われた範囲が、正しく認識できた回転パターンの範囲である。これよりネオコグニトロンは回転に対してはほとんど頑健性がないことがわかる。「2」および「6」に対する認識範囲が広く、「4」「5」「7」に対する認識範囲は特に狭い。これは曲線を多く含む局所パターンは回転してもその特徴が変化しないためである。極端な例を挙げると、数字の「0」は曲線で構成される局所パターンのみで構成されているため、任意の角度回転していても認識できる。

## 2.5 まとめと考察

ネオコグニトロンは Hubel と Wiesel によって提案された第 1 次視覚野における階層型仮説を取り入れた階層型視覚情報統合モデルであるが、回転したパターンに対する頑健性はほとんどないことが数値シミュレーションにより明らかになった。我々ヒトは提示されたパターンが重要なもの（例えば「顔」）であれば回転していても瞬時にそれを正しく認識することができる。次章ではネオコグニトロンを基本としたパターンの回転に頑健な視覚数理モデルについて考察する。

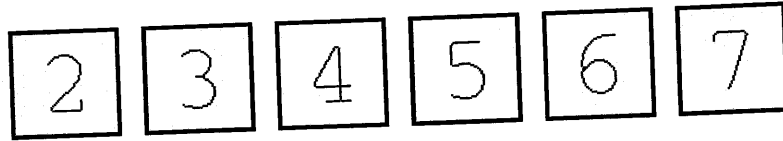


図 2.6: ネオコグニトロンの評価実験で用いる学習用パターン。このパターンは XLFD 記法 [31] で, `--courier-medium-r--*-40-*****` で示されるフォントに対して骨格化処理を施し線幅を 1 にしたものである。

表 2.3: 細胞面数と細胞面内の細胞数に関するパラメータの値。\* で示される値は定義されていないことを示す。また, 第 2 細胞段以降の細胞面は学習により生成されるので学習前に細胞面数を決定することができない。括弧内の値は学習終了後の値である。

	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$K_{Sl}$	*	1	(5)	(7)	(6)
$N_{Sl}$	*	59	17	13	3
$K_{Cl}$	1	1	(5)	(7)	(6)
$N_{Cl}$	61	19	17	10	1

表 2.4: 結合野の大きさ。

	$l=1$	$l=2$	$l=3$	$l=4$
$A_l$	$3 \times 3$	$3 \times 3$	$5 \times 5$	$3 \times 3$
$D_l$	$5 \times 5$	$3 \times 3$	$5 \times 5$	$3 \times 3$

表 2.5: S-細胞のしきい値および学習定数。

	$l=1$	$l=2$	$l=3$	$l=4$
$r_l$	3.8	2.3	2.0	1.9
$q_l$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$

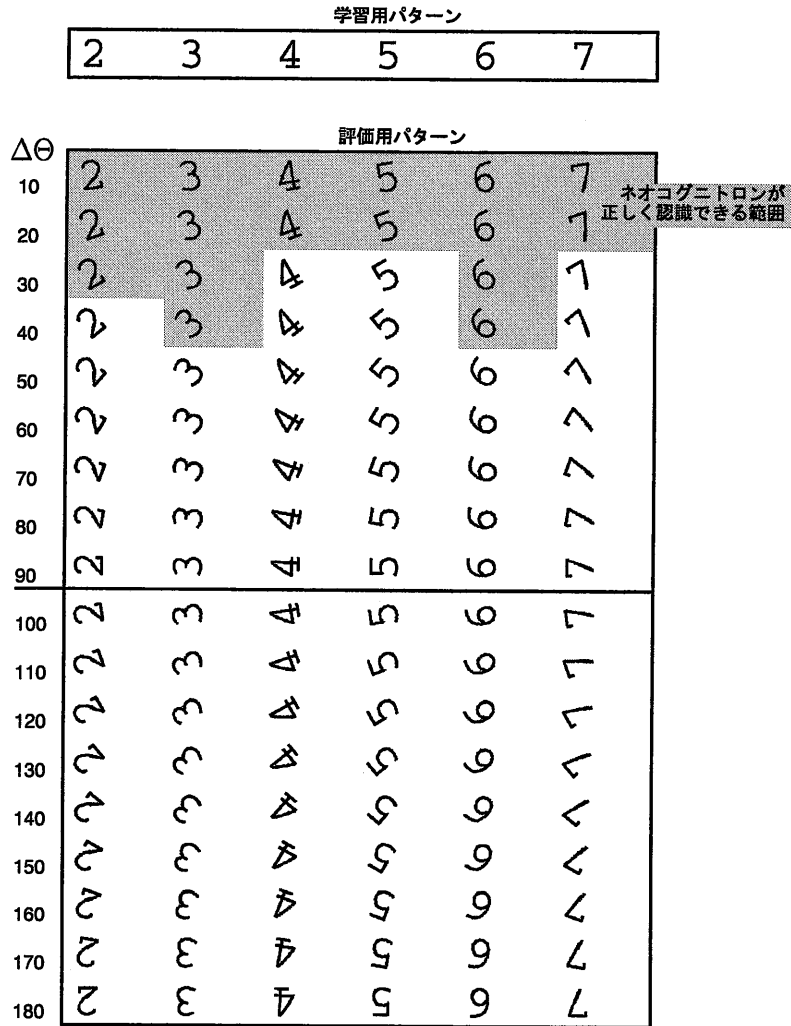


図 2.7: ネオコグニトロンの回転したパターンに対する認識結果。灰色の部分が正しく認識できた回転パターンの範囲を示す。パターンの回転に対してはほとんど頑健性がないことがわかる。

## 第3章 回転対応型ネオコグニトロン

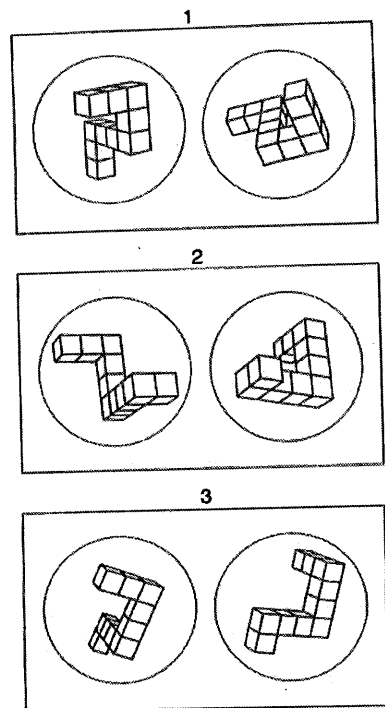
前章でネオコグニトロンがパターンの回転にはほとんど頑健性がないことが明らかになった。信頼性の高いパターン認識を行なう数理モデルには回転の頑健性も要求される。本章ではまず、ヒトが回転したパターンをどのように認識するかについての生理学および心理学的考察を行なう。本章の後半では得られた結論から新しい数理モデルを構築する。

### 3.1 回転パターンの認識に関する生理学的・心理学的考察

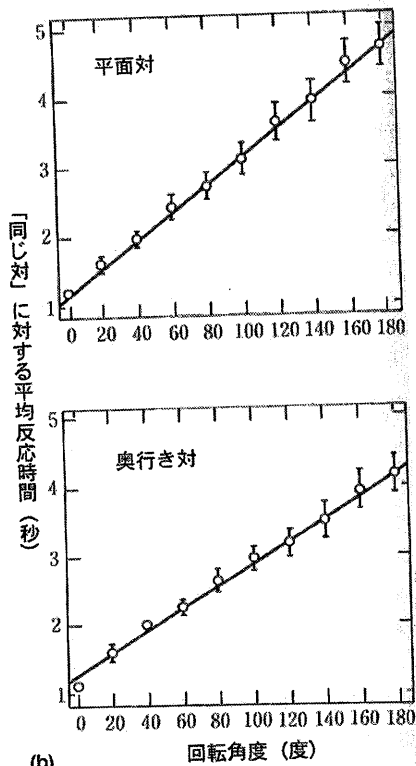
心理学的アプローチによる回転したパターンの認識に関する研究の一つに、Shepard と Metzler による実験 [8] がある。文献 [8] において Shepard らは各試行で、被験者に立方体からなる3次元物体の線画を提示した。図 3.1 の (a)-1 の線画対は絵の平面で回転させると一致し、(a)-2 の線画対は絵の奥行き方向で回転させると一致する。一方、図 3.1 の (a)-3 の線画対は鏡像関係にありどのように回転させても一致しない。Shepard らは被験者に図のような2つの線画対を提示し、それらが同じものかどうかを判断させ、その判断に要する時間を測定した。結果として、(a)-1 や (a)-2 のような回転させると一致する線画対に対しては反応時間が回転角の1次関数となった。Shepard らはこれらの結果から次のことを示唆している。

- 脳内のイメージ上で一方の物体を他方の物体と同じ方向に回転させる速度は一定。
- 刺激となる物体の構造を分析して記号化し、記号化された記述を基に比較し、識別する可能性はない。

換言するとヒトは外界に実在する物体の画像情報をイメージとして脳内に再現し、そのイメージを操作（ここでは回転）することができる。しかもその操作は物理的法則に従っており、イメージの回転は離散的なものではなく連続的な回転をしていることがわかる。Shepard らはこれを「心的回転」と呼んだ。



(a)



(b)

図 3.1: Shepard と Metzlar の回転物体の認識に関する研究. 被験者には 2 つの刺激対を与えられイメージ内で回転させてそれらが同じかどうかを判断しなければならない. Shepard らは判断にかかる時間を測定した. (a): 被験者に与えられた刺激対. “1” の立体は絵の平面で回転させると一致する. “2” の立体は奥行き方向で回転させると一致する. “3” の立体はどのように回転させても一致しない. (b): “1” や “2” のような対の場合には反応時間は回転角度の 1 次関数となる. 文献 [8, 6] を和文化的に.

Shepard らの主張は以上であるが、さらに考察を加えてみる。彼等の与えた線画対は全く未知の物体である。これは物体に対する事前知識を与えていないことに等しく、それにもかかわらず回転角と判断時間は非常に「きれいな」1次関数の形をとっている。これは被験者が線画対に対するなんらかの手掛りを用いて、2つの線画対の回転角および回転方向を実際にイメージで回転する前に推定しているものと考えられる。なぜならば、もし回転角の推定が行なわれなければ時計回りに 45° 回転した線画対に対して、イメージを反時計回りに 315° 回転させて一致するかどうかを判断する可能性があり、この場合には判断時間が1次関数の形にはならないはずだからである。

では、この手掛りとはどのような視覚情報であろうか。回転角の大きさにかわらず図 3.1(b) の結果が得られるということは、手掛りは回転角に依存せず瞬時に一致するかしないかを判断できる視覚情報でなければならない。すなわち回転していても心的回転のようなトップダウン的知識を用いずに認識できる物体や図形が存在するはずである。実際に神経生理学で、回転している個々の物体自体を直接検出する細胞の集団が発見されている。図 3.2 は Wang らによるサルの下側頭皮質における機能構造の研究結果の一つである。下側頭皮質は図 1.1 にあるように、物体の形を判断・認識する部位である。彼等は図に示すような回転した人形の頭をサルに提示し、サルの下側頭皮質の細胞の活動レベルを光学的記録手法を用いて調べた。図 3.2 の中段の矢印が活動レベルが高い細胞集団の位置を示している。このような黒い部分は他の細胞と比較してより多くの酸素を消費している細胞を示している。図から人形の頭の回転角に応じて異なる細胞の活動レベルが上昇し、更に互いの領域が重なっていることがわかる。この結果からサルの下側頭皮質には回転した個々の人形の頭を事前知識を用いることなく直接検出できる細胞が存在することがわかる。サルや人間にとって顔は生存していく上で非常に重要な視覚パターンであり、実際に顔だけを抽出する細胞が TE 野と STS 野で発見されている。

この結果から、(顔のように) 生物にとって重要な視覚パターンに対してはそれが回転したパターンそのものを検出する細胞が存在すると推測できる。本研究ではそのような重要な視覚パターンは単に「単純な視覚パターン」でも起り得ると仮定する。すなわち

- 3次元物体の線画の構成要素である複数の線分からなる視覚パターン。  
例えば立方体の角、
- 文字の構成要素である角や曲線、例えば L-型や T-型といった複数の線分から構成される単純な視覚パターン、

のような視覚パターンで回転しているパターンはそれ自体を検出する細胞が存在すると仮定する。

この仮定を用いると Shepard らの実験結果が説明できる。例えば、被験者が3次元物体の線画対の回転角を推定する際には、立方体の角のような単純

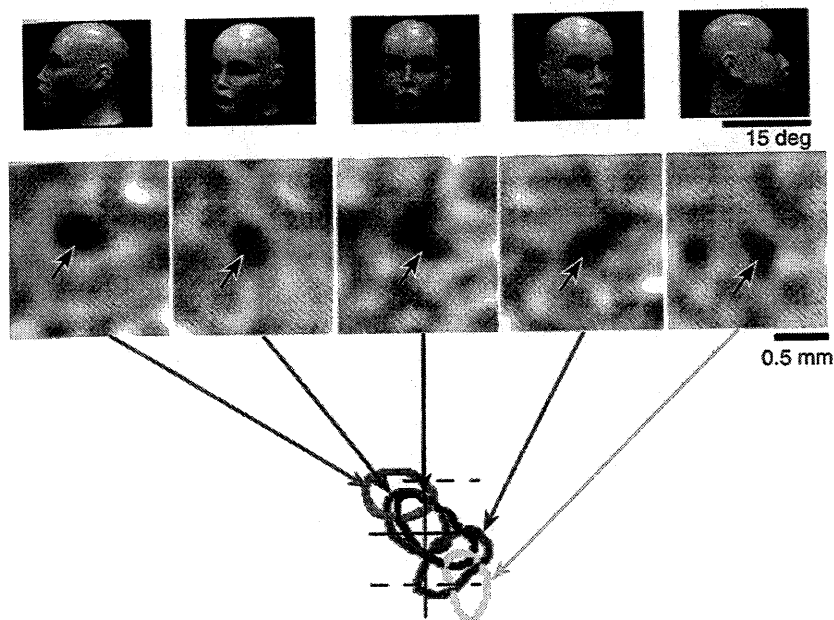


図 3.2: 光学記録手法を用いた Wang らによるサルの下側頭皮質における機能構造の研究 (文献 [33] より抜粋). 回転したいくつかの人形の頭をサルの眼下に提示し, 下側頭皮質の細胞の活動レベルを測定した. 図の矢印で示されている黒い部分が活動している神経細胞の集団であり, 回転に伴ない活動している細胞群の位置がずれていることがわかる. このサルは回転した個々の人形の頭を事前知識を用いることなく検出できていることを示す.



な視覚パターンの回転角はトップダウン的な情報を用いなくともそれを検出する細胞が存在するので瞬時に回転角を推定できると説明ができる。

## 3.2 従来型ネオコグニトロンの再構成

前節で「単純な視覚パターン」はそれを回転させた視覚パターン自体を検出する細胞が存在することを示唆した。本節ではこの事実を基にしてネオコグニトロンを再構成した新しい視覚数理モデルを提案する。提案するモデルを回転対応型ネオコグニトロンと呼ぶことにし、Fukushima によって提案されたモデルを従来型ネオコグニトロンと呼ぶことにする。回転対応型ネオコグニトロンは従来型ネオコグニトロンと同様にボトムアップ型の階層型神経回路網モデルであり、正立した（回転していない）学習用パターンを学習しただけでそれを回転したパターンや、回転・変形を同時に施したパターンを直接認識することが可能となるモデルである。回転に対して単純ではないパターンを認識するための数理モデルについての考察は第 5 章で行なう。

従来型ネオコグニトロンとの主な相異点は細胞面を 3 次元に拡張した細胞面群を導入したことであり、新しく付け加えられた次元は局所パターンの回転角度を表現している。次に細胞面群の説明をする。

### 3.2.1 細胞面群

図 3.3(a) で示すように従来型ネオコグニトロンの  $U_{S1}$  層では位置は異なっても同じ方位をもつ線分は同じ特徴量とみなし、同じ細胞面内で処理されていた。つまり、

- 一つの細胞面内の各細胞はすべて同じ方位の線分に対して反応（検出）し、検出された線分の位置情報は反応する細胞の位置  $\mathbf{n} = (n_x, n_y) \in \mathbf{Z}^2$  で表わされ、
- 異なる方位の線分は異なる特徴であるとみなし、それぞれ異なる細胞面で検出される。特徴の番号は細胞面の通し番号  $k$  で表現される。例えば  $0^\circ$  の線分を検出する細胞面の番号は  $k = 1$ 、 $22.5^\circ$  の線分を検出する細胞面の番号は  $k = 2$  としている。

しかし方位の異なる線分も単に方位が異なるだけであり、短い線分という同一の特徴をもっているとみなすこともできる。ここで図 3.3(b) のように細胞面に新しい次元  $n_\theta$  を付け加え、各細胞が 3 次元的に配置されるように複数の細胞面を並び換える。この並び換えた複数の細胞面をまとめて細胞面群と呼ぶことにする。すると従来型ネオコグニトロンの複数の細胞面は 1 つの細胞面群で表現され、

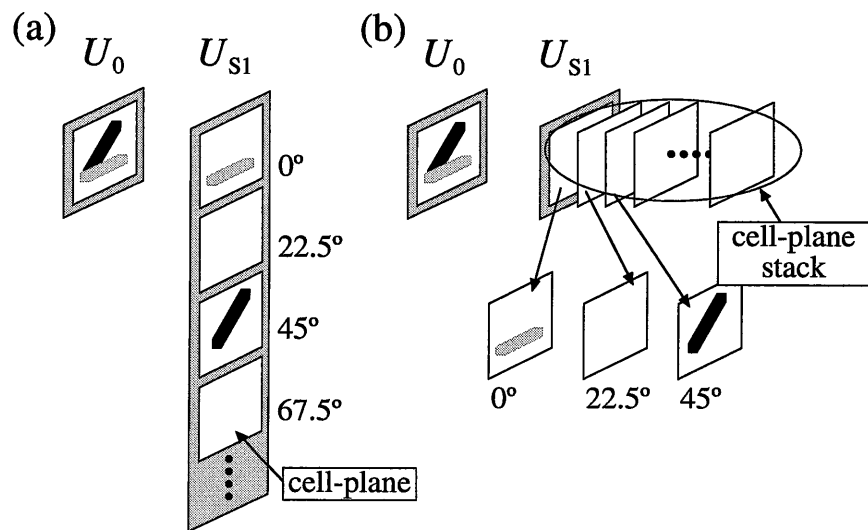


図 3.3: 細胞面から細胞面群への拡張. (a): 従来型ネオコグニトロン $U_{S1}$ 層による異なる方位を持つ線分の検出. (b): 細胞面群による線分の検出. 各細胞が 3 次元的に配置されるように複数の細胞面を並び換える. 新しく付け加えられた次元は  $n_\theta$  と表記され, 線分の回転角を表現する次元となる. この並び換えた複数の細胞面をまとめて細胞面群と呼ぶことにする. 細胞面群の導入により, 異なる方位を持つ線分も同様の性質を持つ局所パターンとして扱われる.

- 細胞面群内に存在するある 1 つの細胞面はすべて同じ方位の線分に対して反応し、線分の位置情報は反応する細胞の位置  $\mathbf{n} = (n_x, n_y) \in \mathbf{Z}^2$  で表わされ、
- 異なる方位の線分も同じ特徴とみなし、1 つの細胞面群内の細胞で検出される。特徴の番号（線分の方向）は新しく付け加えられた次元  $n_\theta$  で表現される、

という性質を持つことになる。  $U_{S1}$  層に細胞面群を導入することは、線分を単純な視覚パターンと仮定していることに等しく、図 3.2 の Wang らの実験結果と定性的に対応している。

### 3.2.2 回転対応型ネオコグニトロンの構造および性質

$U_{S2}$  層以降の細胞層に対しても細胞面群を導入する。これにより再構築した従来型ネオコグニトロンが回転対応型ネオコグニトロンであり、従来型ネオコグニトロンの細胞面を全て細胞面群で置き換えた数値モデルとなっている。図 3.4 に回転対応型ネオコグニトロンの構造を図示する。従来型ネオコグニトロンと同様に複数の細胞層からなる階層型神経回路網である。従来型ネオコグニトロンと同様に  $U_0$  層は網膜のモデルであり、 $U_0$  層に続いて数段の細胞層から構成される。図では 8 つの細胞層から構成されており、 $U_S$  層と  $U_C$  層が交互に並んでいる。各細胞層は 1 つ、もしくは複数の細胞面群を含み、 $U_S$  層の細胞面群を S-細胞面群、 $U_C$  層の細胞面を C-細胞面群と呼ぶ。また S-細胞面群には特徴抽出細胞である S-細胞が、C-細胞面には位置ずれ・回転ずれ許容細胞である C-細胞が 3 次元的に配置されている。その他に V-細胞が存在するが図では省略してある。

回転対応型ネオコグニトロンは正立した（回転していない）学習用パターンを学習するだけでそれを回転した未知のパターンを正しく認識することが期待される。また従来型ネオコグニトロンの性質を完全に継承した構成であるために、従来型ネオコグニトロンの性能も完全に継承している。更に従来型ネオコグニトロンで得られた様々な知見やアルゴリズムをそのまま適応できることが期待される。

### 3.2.3 各細胞の数式表現

各細胞の数式表現を与える。各細胞および各結合強度の表記方法は表 3.1 に、その他の記号の意味は表 3.2 にまとめてあるので参照していただきたい。従来型ネオコグニトロンとの主な相異点は、細胞の位置を表す  $\mathbf{n}$ 、および相対位置を表す  $\nu$  が  $\mathbf{Z}^3$  の要素となったことである。また本稿では説明を容易にするために、回転方向の座標値  $n_\theta, \nu_\theta$  に対して  $n_\theta = 0, 1, \dots$  と表記する以外

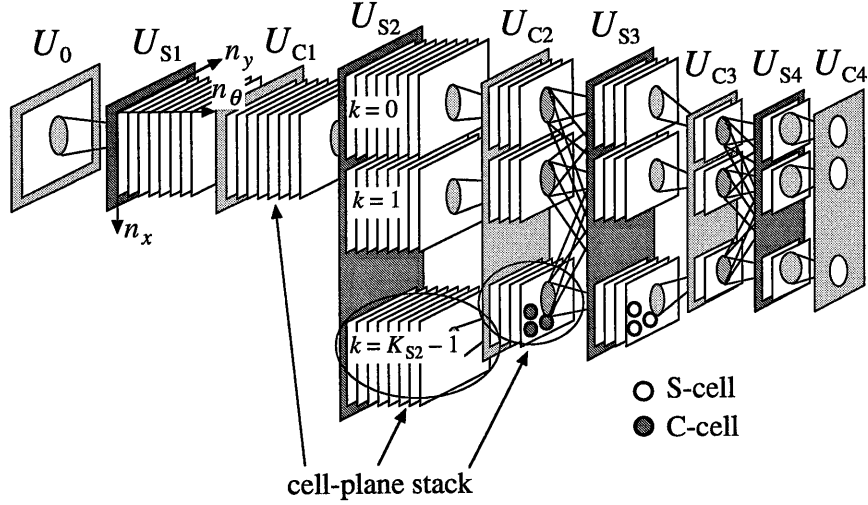


図 3.4: 回転対応型ネオコグニトロン構造. 各細胞層は細胞面群を含む. 従来型ネオコグニトロンと同様に  $U_0$  層は網膜のモデルであり,  $U_0$  層に続いて数段の細胞層から構成される. 図では 8 つの細胞層から構成されている.  $U_S$  層の細胞面群を S-細胞面群,  $U_C$  層の細胞面を C-細胞面群と呼ぶ. また S-細胞面群には特徴抽出細胞である S-細胞が, C-細胞面には位置ずれ許容細胞である C-細胞が 3 次元的にそれぞれに配置されている. その他に V-細胞も存在するが省略してある.

表 3.1: 回転対応型ネオコグニトロンの各細胞と結合の表記法

表記法	意味
$U_{Cl-1}(n, \kappa)$	第 $l-1$ 細胞段の C 層で, 細胞面群の番号が $\kappa$ であり, 細胞面群内の位置 $n$ に配置してある C-細胞. $n = (n_x, n_y, n_\theta) \in \mathbf{Z}^3$
$U_{Vl}(n)$	第 $l$ 細胞段群内で, 位置 $n$ に配置してある V-細胞.
$U_{Sl}(n, k)$	第 $l$ 細胞段の S 層で, 細胞面群の番号が $k$ であり, 細胞面群内の位置 $n$ に配置してある S-細胞.
$a_l(\nu, n, \kappa, k)$	$U_{Cl-1}(n \oplus \nu, \kappa) - U_{Sl}(n, k)$ 間の興奮性可変結合
$b_l(k)$	$U_{Vl}(n) - U_{Sl}(n, k)$ 間の抑制性可変結合
$c_l(\nu)$	$U_{Cl-1}(n \oplus \nu, \kappa) - U_{Vl}(n)$ 間の興奮性固定結合
$d_l(\nu)$	$U_{Sl}(n \oplus \nu, \kappa) - U_{Cl}(n, k)$ 間の興奮性固定結合

表 3.2: 回転対応型ネオコグニロンで使用される記号と説明

表記	意味
$K_{Sl}$	$U_{Sl}$ 層内の細胞面群の数. 細胞面群の番号は $k$ (または $\kappa$ ) で参照する.
$K_{Cl}$	$U_{Cl}$ 層内の細胞面群の数. 細胞面群の番号は $\kappa$ (または $k$ ) で参照する.
$T_{Sl}$	$U_{Sl}$ 層内の 1 つ細胞面群内に存在する細胞面数.
$T_{Cl}$	$U_{Cl}$ 層内の 1 つ細胞面群内に存在する細胞面数.
$A_l$	$U_{Sl}$ 層内の S-細胞が $U_{Cl-1}$ 層内の C-細胞と結合する範囲を与える集合. 集合の要素 $\nu$ は 3次元相対位置座標 ( $\nu \in \mathbf{Z}^3$ ) である.
$D_l$	$U_{Cl}$ 層内の C-細胞が $U_{Sl}$ 層内の S-細胞と結合する範囲を与える集合. 集合の要素 $\nu$ は 3次元相対位置座標 ( $\nu \in \mathbf{Z}^3$ ) である.

に,  $n_\theta = 0^\circ, 45^\circ, \dots$  と表記する場合もある.  $U_{Sl}(\mathbf{n}, k)$  細胞の出力  $u_{Sl}(\mathbf{n}, k)$  は次式で与えられる.

$$u_{Sl}(\mathbf{n}, k) = r_l \cdot \phi \left[ \frac{1+e}{1 + \frac{r_l}{1+r_l} \cdot i} - 1 \right], \quad (3.1)$$

$$\phi(x) = \max(x, 0), \quad (3.2)$$

ここで関数  $\phi$  は S-細胞の出力関数である.  $r_l$  はしきい値の効果を与える正の値である.  $e$  は C-細胞から S-細胞への興奮性入力 of 総和を表わし,  $i$  V-細胞からの抑制性 of 入力を表わし, 次式で与えられる.

$$e = \sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} a_l(\nu, \mathbf{n}, \kappa, k) \cdot u_{Cl-1}(\mathbf{n} \oplus_{T_{Cl-1}} \nu, \kappa), \quad (3.3)$$

$$i = b_l(k) \cdot v_{Sl}(\mathbf{n}). \quad (3.4)$$

ここで二項演算子  $\oplus_M$  は次式で定義される.

$$\begin{cases} n_x \oplus_M \nu_x & \stackrel{\text{def}}{=} n_x + \nu_x, \\ n_y \oplus_M \nu_y & \stackrel{\text{def}}{=} n_y + \nu_y, \\ n_\theta \oplus_M \nu_\theta & \stackrel{\text{def}}{=} (n_\theta + \nu_\theta) \bmod M, \end{cases} \quad (3.5)$$

余剰演算  $\bmod$  の値域は  $[0, M)$  とする.  $\oplus_M$  は  $n_\theta$ -方向 of 周期性を保つために導入された演算子である. 例えば  $n_\theta + \nu_\theta = 400^\circ$  である場合にはこれを  $60^\circ$  にする必要があり, これを表現しているのが  $\oplus_M$  である. 単に  $\oplus$  と記す場合もある.  $K_{Cl-1}$  は  $U_{Cl-1}$  層における細胞面群 of 数を表す.

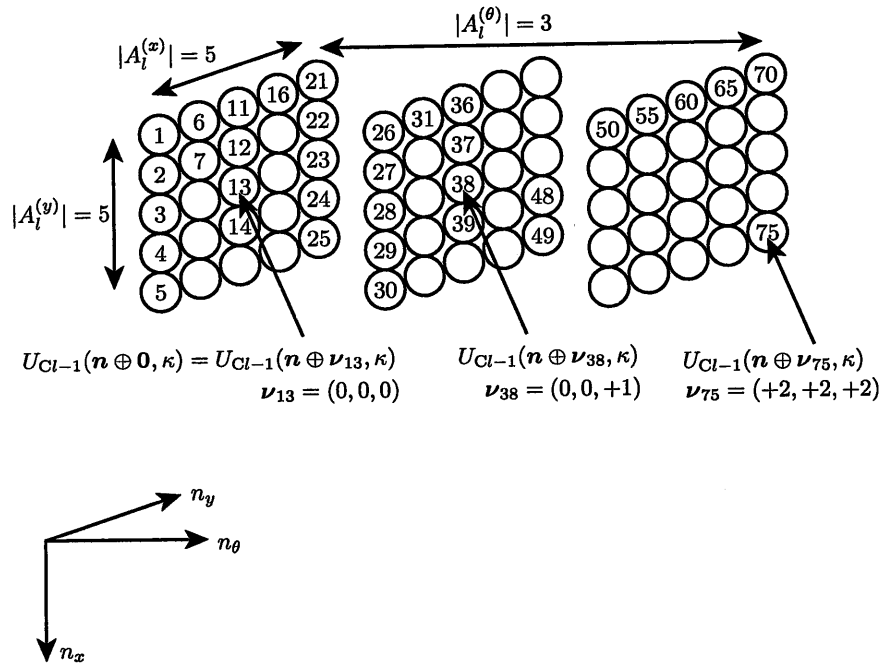


図 3.5: 辞書式順序集合  $A_l$  の要素  $\nu_i$  に関する説明. 白い円は S-細胞  $U_{Sl}(n, k)$  と結合している C-細胞を表わし, この全体が集合  $\{U_{Cl-1}(n \oplus \nu_i, \kappa) \mid \nu_i \in A_l\}$  に対応する. S-細胞は  $A_l$  で決定される領域内に存在する C-細胞のみと結合しており, 図では位置  $n$  (13 番細胞) を中心とした  $5 \times 5 \times 3$  の大きさを持つ領域内に存在する C-細胞を示している. この領域の  $n_x$ -方向の大きさは  $|A_l^{(x)}|$ ,  $n_y$ -方向の大きさは  $|A_l^{(y)}|$ ,  $n_\theta$ -方向の大きさは  $|A_l^{(\theta)}|$  とそれぞれ表現する. 円内の数字は  $A_l$  の要素  $\nu_i$  の番号  $i$  を指している. この番号付けの規則は  $\nu_i$  の辞書式順序に従っている (本文参照).

$A_l$  は  $U_{Sl}$  の S-細胞が  $U_{Cl-1}$  層の C-細胞と結合する範囲を与える集合である. 3次元に拡張した細胞面群を導入しているため  $A_l$  の要素  $\nu$  は  $\nu \in \mathbb{Z}^3$  となり, 次のように表わす.

$$A_l \stackrel{\text{def}}{=} \{\nu_1, \dots, \nu_i, \dots, \nu_j, \dots, \nu_{|A_l|}\}. \quad (3.6)$$

ただし記号  $|\cdot|$  は集合の要素数を意味する. 集合  $A_l$  は辞書式順序集合であるとし, 各要素  $\{\nu_1, \nu_2, \dots, \nu_{|A_l|}\}$  は辞書式に番号付けされているとする. すなわち関係  $<_L$  を  $\mathbb{Z}^3$  に対する辞書式順序 (lexicographic order) とすると, 各要素が  $\nu_i <_L \nu_j$  となるように番号  $i, j$  が付けられている. 図 3.5 を用いて説明する. 白い円は, S-細胞  $U_{Sl}(n, k)$  と結合する  $U_{Cl-1}$  層の C-細胞を表現している. S-細胞は  $A_l$  で決定される領域内に存在する C-細胞のみと結合しており, 図では  $|A_l| = 5 \times 5 \times 3$  とした場合の C-細胞を示している.  $A_l$  の

$n_x$ -方向の座標値の集合を  $A_l^{(x)}$ ,  $n_y$ -方向の座標値の集合  $A_l^{(y)}$ ,  $n_\theta$ -方向の座標値の集合を  $A_l^{(\theta)}$  と記す. 図の場合は  $A_l^{(x)} = A_l^{(y)} = \{-2, -1, 0, +1, +2\}$ ,  $A_l^{(\theta)} = \{0, +1, +2\}$  であり,

$$A_l = A_l^{(x)} \otimes A_l^{(y)} \otimes A_l^{(\theta)},$$

となる. 図は  $|A_l^{(x)}| = 5, |A_l^{(y)}| = 5, |A_l^{(\theta)}| = 3$  の場合である. 一般に

$$|A_l| = |A_l^{(x)}| \times |A_l^{(y)}| \times |A_l^{(\theta)}|,$$

である. 各円内の数字は  $A_l$  の要素  $\nu_i$  の番号  $i$  を指しており, これは各要素を辞書式順序で並べたときの番号に等しい. 図の例では

$$\begin{aligned} \nu_1 &= (-2, -2, 0) \\ <_L \nu_2 &= (-1, -2, 0) \\ <_L \nu_3 &= (0, -2, 0) \\ <_L \nu_4 &= (+1, -2, 0) \\ <_L \nu_5 &= (+2, -2, 0) \\ <_L \nu_6 &= (-2, -1, 0) \\ <_L \nu_{13} &= (0, 0, 0) \\ <_L \nu_{29} &= (-1, -2, +1) \\ <_L \nu_{38} &= (0, 0, +1) \\ <_L \nu_{74} &= (+1, +2, +2) \\ <_L \nu_{75} &= (+2, +2, +2) \end{aligned}$$

となる.

## V-細胞

$U_{Sl}$  層の S-細胞に抑制性の信号を送っている V-細胞  $U_{Vl}(\mathbf{n})$  の出力  $u_{Vl}(\mathbf{n})$  は次式で示される.

$$u_{Vl}(\mathbf{n}) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{\nu \in A_l} c_l(\nu) \cdot \left\{ u_{Cl-1}(\mathbf{n} \oplus_{T_{Cl-1}} \nu, \kappa) \right\}^2}, \quad (3.7)$$

ここで  $c_l(\nu)$  は興奮性固定結合の強度を表わす.  $U_{Vl}$  細胞は  $U_{Cl-1}$  層の C-細胞から入力を受けている. また,  $U_{Vl}$  細胞は  $U_{Sl}$  層の全ての細胞面群  $k$  に共通である.

## C-細胞

C-細胞  $U_{Cl}(\mathbf{n}, \kappa)$  は  $U_{Sl}$  層内の S-細胞から入力を受けている. 従来型ネオコグニトロンと同様, S-細胞が前層の全ての細胞面内に位置する C-細胞から

入力を受けていたのに対し、C-細胞は同層の同じ細胞面番号  $\kappa$  を持つ S-細胞からのみ入力を受ける。  $u_{Cl}(\mathbf{n}, \kappa)$  は次式で与えられる。

$$u_{Cl}(\mathbf{n}, \kappa) = \psi \left[ \sum_{\boldsymbol{\nu} \in D_l} d_l(\boldsymbol{\nu}) \cdot u_{Sl}(\mathbf{n} \oplus_{T_{Sl}} \boldsymbol{\nu}, k) \right], \quad (3.8)$$

ここで  $d_l(\boldsymbol{\nu})$  は興奮性固定結合の強度を表わし、  $D_l$  は  $U_{Cl}$  層の C-細胞が  $U_{Sl}$  層の S-細胞と結合する範囲の集合を与える。  $D_l \subset \mathbf{Z}^3$  でありその要素を次の記す。

$$D_l \stackrel{\text{def}}{=} \{\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \dots, \boldsymbol{\nu}_{|D_l|}\}. \quad (3.9)$$

$D_l$  の  $n_x$ -方向の座標値の集合を  $D_l^{(x)}$ ,  $n_y$ -方向の座標値の集合  $D_l^{(y)}$ ,  $n_\theta$ -方向の座標値の集合を  $D_l^{(\theta)}$  と記す。

$$D_l = D_l^{(x)} \otimes D_l^{(y)} \otimes D_l^{(\theta)}$$

であり、一般的には次式が成立する。

$$|D_l| = |D_l^{(x)}| \times |D_l^{(y)}| \times |D_l^{(\theta)}|.$$

### 3.3 学習方法

学習方法は従来型ネオコグニトロンの学習アルゴリズム 1 とほぼ同様である。相異点は細胞面群の導入に起因する。回転対応型ネオコグニトロンの細胞面群では複数の細胞面から構成されており (i) : 細胞面内の S-細胞 ( $n_\theta$  方向の位置が同じ S-細胞) の興奮性可変結合は空間的に均一であるが、(ii) : 同じ細胞面群に属する他の細胞面内の S-細胞 ( $n_\theta$  方向の位置が異なる S-細胞) の可変結合は回転角  $n_\theta$  に合せて回転させた空間分布を持つ。この関係を数式表現すると次のようになる。

$$a_l(\boldsymbol{\nu}, \mathbf{n}, \kappa, \hat{k}) = a_l(\boldsymbol{\nu}, \hat{\mathbf{n}}, \kappa, \hat{k}) \quad \text{if } n_\theta = \hat{n}_\theta, \quad (3.10)$$

$$a_l(\boldsymbol{\nu}, \mathbf{n}, \kappa, \hat{k}) = R(n_\theta - \hat{n}_\theta) \cdot a_l(\boldsymbol{\nu}, \hat{\mathbf{n}}, \kappa, \hat{k}) \quad \text{if } n_\theta \neq \hat{n}_\theta, \quad (3.11)$$

ただしシードセルを  $U_{Sl}(\hat{\mathbf{n}}, \hat{k})$  とし、  $\hat{\mathbf{n}} = (\hat{n}_x, \hat{n}_y, \hat{n}_\theta)$  としている。演算  $R(n_\theta - \hat{n}_\theta)$  については後述する。すなわち強化すべき S-細胞の位置  $\mathbf{n}$  が条件  $\hat{\mathbf{n}} = \mathbf{n}$  を満している場合には従来型ネオコグニトロンの学習則と同じ様にアルゴリズム 1 を適応する。一方、条件  $n_\theta \neq \hat{n}_\theta$  を満す S-細胞の可変結合は、シードセルで得られた可変結合を「回転させた可変結合」となるように強化する学習則とした。シードセルの可変結合を S-細胞の回転方向の位置に応じて回転させることで、回転した局所的なパターンを効果的に検出することができる (図 3.7 を参照)。

回転対応型ネオコグニトロンの学習方法を数式的に表記するために  $|A_l^{(x)}|$  行  $|A_l^{(y)}|$  列の行列  $A_l(\boldsymbol{\nu}_\theta, \mathbf{n}, \kappa, k)$  を定義する。この行列の要素  $A_l(\boldsymbol{\nu}_\theta, \mathbf{n}, \kappa, k)_{m,n}$



は次式で定義される.

$$\begin{aligned} \mathcal{A}_l(\nu_\theta, \mathbf{n}, \kappa, k)_{(m,n)} &\stackrel{\text{def}}{=} a_l(\nu_{(m,n)}, \mathbf{n}, \kappa, k), \\ \nu_{(m,n)} &= \nu_{m+(n-1) \times |A_l^{(x)}| + \nu_\theta \times |A_l^{(x)}| \times |A_l^{(y)}|}. \end{aligned} \quad (3.12)$$

この様に定義された行列  $\mathcal{A}_l(\nu_\theta, \mathbf{n}, \kappa, k)$  は下記の性質を持つ.

- S-細胞  $U_{Sl}(\mathbf{n}, k)$  が持つ興奮性可変結合の集合  $\{a_l(\nu_i, \mathbf{n}, \kappa, k) \mid \nu_i \in A_l, \kappa = 1, \dots, K_{Cl-1}\}$  (要素数は  $|A_l| \times K_{Cl-1}$ ) の中で,
- 相対位置  $\nu_i = (\nu_{i,x}, \nu_{i,y}, \nu_{i,\theta})$  の回転方向の値  $\nu_{i,\theta}$  が  $\nu_{i,\theta} = \nu_\theta$  であり,
- 細胞面群の番号が  $\kappa$  である興奮性可変結合を行列要素として持つ集合.

図 3.5 を用いて  $\mathcal{A}_l(\nu_\theta, \mathbf{n}, \kappa, k)$  具体的に説明する. 図 3.5 は  $K_{Cl-1} = 1$ ,  $A_l^{(x)} = A_l^{(y)} = 5$ ,  $A_l^{(\theta)} = 3$  の場合を例示している. また図中の白い円は, S-細胞  $U_{Sl}(\mathbf{n}, k)$  と結合している  $U_{Cl-1}$  層内の 75 ( $= 5 \times 5 \times 3$ ) 個の C-細胞を表現している. またこの S-細胞は 75 個の興奮性可変結合  $\{a_l(\nu_i, \mathbf{n}, \kappa, k), i = 1, \dots, 75\}$  を持つ. この 75 個の興奮性可変結合の中で回転方向が 0 ( $\nu_\theta = 0$ ) の位置にある 25 個の C-細胞,

$$\{U_{Cl-1}(\mathbf{n} \oplus \nu_i, \kappa) \mid \nu_{i,\theta} = 0, i = 1, \dots, 25\},$$

と結合を構成している興奮性可変結合は

$$\{a_l(\nu_i, \mathbf{n}, \kappa, k) \mid \nu_{i,\theta} = 0, i = 1, \dots, 25\},$$

の 25 種類である. よって  $\mathcal{A}_l(\nu_\theta = 0, \mathbf{n}, \kappa, k)$  は次式で表わされる.

$$\mathcal{A}_l(\nu_\theta = 0, \mathbf{n}, \kappa, k) = \begin{pmatrix} a_l(\nu_1) & a_l(\nu_6) & a_l(\nu_{11}) & a_l(\nu_{16}) & a_l(\nu_{21}) \\ a_l(\nu_2) & a_l(\nu_7) & a_l(\nu_{12}) & a_l(\nu_{17}) & a_l(\nu_{22}) \\ a_l(\nu_3) & a_l(\nu_8) & a_l(\nu_{13}) & a_l(\nu_{18}) & a_l(\nu_{23}) \\ a_l(\nu_4) & a_l(\nu_9) & a_l(\nu_{14}) & a_l(\nu_{19}) & a_l(\nu_{24}) \\ a_l(\nu_5) & a_l(\nu_{10}) & a_l(\nu_{15}) & a_l(\nu_{20}) & a_l(\nu_{25}) \end{pmatrix},$$

ただし紙面の都合上  $a_l(\nu_i, \mathbf{n}, \kappa, k)$  を  $a_l(\nu_i)$  と略記してある. 同様に,

$$\mathcal{A}_l(\nu_\theta = 1, \mathbf{n}, \kappa, k) = \begin{pmatrix} a_l(\nu_{26}) & a_l(\nu_{31}) & a_l(\nu_{36}) & a_l(\nu_{41}) & a_l(\nu_{46}) \\ a_l(\nu_{27}) & a_l(\nu_{32}) & a_l(\nu_{37}) & a_l(\nu_{42}) & a_l(\nu_{47}) \\ a_l(\nu_{28}) & a_l(\nu_{33}) & a_l(\nu_{38}) & a_l(\nu_{43}) & a_l(\nu_{48}) \\ a_l(\nu_{29}) & a_l(\nu_{34}) & a_l(\nu_{39}) & a_l(\nu_{44}) & a_l(\nu_{49}) \\ a_l(\nu_{30}) & a_l(\nu_{35}) & a_l(\nu_{40}) & a_l(\nu_{45}) & a_l(\nu_{50}) \end{pmatrix},$$

$$\mathcal{A}_l(\nu_\theta = 2, \mathbf{n}, \kappa, k) = \begin{pmatrix} a_l(\nu_{51}) & a_l(\nu_{56}) & a_l(\nu_{61}) & a_l(\nu_{66}) & a_l(\nu_{71}) \\ a_l(\nu_{52}) & a_l(\nu_{57}) & a_l(\nu_{62}) & a_l(\nu_{67}) & a_l(\nu_{72}) \\ a_l(\nu_{53}) & a_l(\nu_{58}) & a_l(\nu_{63}) & a_l(\nu_{68}) & a_l(\nu_{73}) \\ a_l(\nu_{54}) & a_l(\nu_{59}) & a_l(\nu_{64}) & a_l(\nu_{69}) & a_l(\nu_{74}) \\ a_l(\nu_{55}) & a_l(\nu_{60}) & a_l(\nu_{65}) & a_l(\nu_{70}) & a_l(\nu_{75}) \end{pmatrix},$$

である。

行列  $A_l(\nu_\theta, n, \kappa, k)$  を用いると、回転対応型ネオコグニトロンの学習における可変結合の変化量は以下の式で示される。

$$\Delta a_l(\nu_i, \hat{n}, \kappa, \hat{k}) = q_l \cdot c_l(\nu_i) \cdot u_{Cl-1}(\hat{n} \oplus \nu_i, \kappa), \quad (3.13)$$

$$\Delta b_l(\hat{k}) = q_l \cdot u_{vl}(\hat{n}), \quad (3.14)$$

$$A_l(\nu, n, \kappa, \hat{k}) = R(n_\theta - \hat{n}_\theta) \cdot A_l(\nu, \hat{n}, \kappa, \hat{k}) \quad (3.15)$$

ただし  $R(n_\theta - \hat{n}_\theta)$  は行列  $A_l(\nu_\theta, \hat{n}, \kappa, \hat{k})$  の要素を2次元画像の画素値とした場合に、その画像を  $n_\theta - \hat{n}_\theta$  [deg] 回転させる演算である。

この興奮性可変結合を要素とする行列  $A_l(\nu_\theta, \hat{n}, \kappa, \hat{k})$  を回転させる演算  $R(n_\theta - \hat{n}_\theta)$  についての例を図3.6を用いて説明する。図3.6(a)は式(3.13)に従って強化されたシードセル  $U_{Sl}(\hat{n}, \hat{k})$  の興奮性可変結合の集合の一部  $A_l(\nu_\theta = 0^\circ, \hat{n}, \kappa, \hat{k})$  を例示したものである<sup>1</sup>。ただし図ではその要素  $a_l(\nu_i, \hat{n}, \kappa, \hat{k})$  を  $a_l(\nu_i)$  と略してある。シードセルの興奮性可変結合の強化が完了すると、シードセルが属している細胞面群  $\hat{k}$  内の他のS-細胞の興奮性可変結合は式(3.15)に従って強化される。例えば図3.6(b)にあるように、シードセルの位置  $\hat{n}$  と比較して45°ずれた位置  $n$  に配置されているS-細胞  $U_{Sl}(n, \hat{k})$  の興奮性可変結合の集合  $A_l(\nu_\theta = 0, n, \kappa, \hat{k})$  を強化する場合を考える。すなわち  $n_\theta - \hat{n}_\theta = 45^\circ$  とする。ただし図では要素  $a_l(\nu_i, n, \kappa, \hat{k})$  を  $a'_l(\nu_i)$  と略してある。このとき  $A_l(\nu_\theta = 0, n, \kappa, \hat{k})$  は式(3.15)に従い、演算子  $R(n_\theta - \hat{n}_\theta = 45^\circ)$  を適応して  $A_l(\nu_\theta = 0, \hat{n}, \kappa, \hat{k})$  を45°回転したものとなる(図3.6(c))。  $A_l(\nu_\theta = 0, n, \kappa, \hat{k})$  の要素である興奮性可変結合  $a_l(\nu_i, n, \kappa, \hat{k})$  の値を算出する方法は様々存在するが、本論文では4点線形補間法によってこれを算出する。例えば図3.6(c)にあるように、回転後の興奮性可変結合  $a'_l(\nu_{14})$  を算出する際には、隣接する4要素である  $a_l(\nu_8), a_l(\nu_9), a_l(\nu_{13}), a_l(\nu_{14})$  の各値を線形に補間した値とする。これを式として表現すると以下のようなになる。

$$a'_l(\nu_{14}) = \begin{bmatrix} x & 1-x \end{bmatrix} \begin{bmatrix} a_l(\nu_8) & a_l(\nu_{13}) \\ a_l(\nu_9) & a_l(\nu_{14}) \end{bmatrix} \begin{bmatrix} y \\ 1-y \end{bmatrix}.$$

ただし  $x, y$  は図3.6(d)にあるように  $a_l(\nu_8), a_l(\nu_9), a_l(\nu_{13}), a_l(\nu_{14})$  からの距離に基づく内分比を示すものである。

上記の学習則を用いた回転対応型ネオコグニトロンの学習アルゴリズムをアルゴリズム2に示す。

### 3.4 回転したパターンの認識機構

本章では回転対応型ネオコグニトロンの回転したパターンの認識方法を説明する。まず細胞面群内の特徴抽出細胞であるS-細胞が、局所パターンの回

<sup>1</sup>この場合、興奮性可変結合の「全体」は  $\{A_l(\nu_\theta = 0, \hat{n}, \kappa, \hat{k}), A_l(\nu_\theta = 1, \hat{n}, \kappa, \hat{k}), A_l(\nu_\theta = 2, \hat{n}, \kappa, \hat{k})\}$  である。

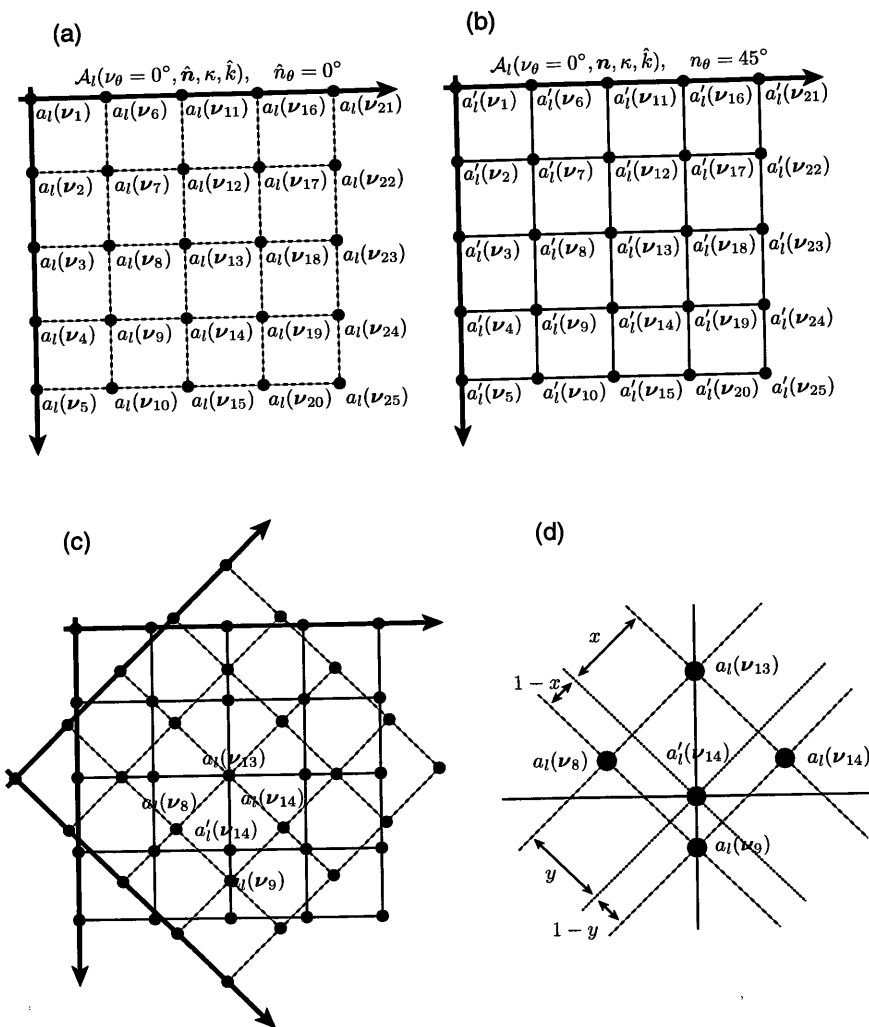


図 3.6: シードセル以外の興奮性可変結合の強化方法. (a): 強化されたシードセル  $U_{Si}(\hat{n}, \hat{k})$  の興奮性可変結合の配置  $A_i(\nu_\theta = 0, \hat{n}, \kappa, \hat{k})$ . ただし行列の要素  $a_i(\nu_i, \hat{n}, \kappa, \hat{k})$  を  $a_i(\nu_i)$  と略してある. (b): シードセルの位置  $\hat{n}$  と比較して  $45^\circ$  ずれた位置  $\mathbf{n}$  に配置されている S-細胞  $U_{Si}(\mathbf{n}, \hat{k})$  の興奮性可変結合を要素として持つ行列  $A_i(\nu_\theta = 0, \mathbf{n}, \kappa, \hat{k})$ . ただし要素  $a_i(\nu_i, \mathbf{n}, \kappa, \hat{k})$  を  $a'_i(\nu_i)$  と略してある. (c):  $U_{Si}(\mathbf{n}, \hat{k})$  の興奮性可変結合はシードセルの興奮性可変結合を画像として回転したものをコピーする. (d): (c) を拡大した図. 4 点線形補間法による興奮性可変結合の値  $a'_i(\nu_{14})$  の決定方法を示す.  $a'_i(\nu_{14})$  の値は 4 つの近隣要素の値から線形に内挿して求める.  $x, y$  は距離に基づく内分比を示すものである.

---

**アルゴリズム 2** 回転対応型ネオログニトロン<sup>1</sup>の学習方法

---

**Require:** All the S-cell layers up to  $U_{Sl-1}$  are already reinforced.

- 1: **while** New training patterns still exist **do**
  - 2:   Input a new training pattern.
  - 3:   Calculate responses of the preceding layer  $U_{Cl-1}$ .
  - 4:   Make seed-cells using seed-selecting plane stack.
  - 5:   Insert positions of the seed-cells in a queue,  $Q$ .
  - 6:   **while**  $Q$  is not empty **do**
  - 7:      $\hat{n} \leftarrow$  dequeue  $Q$ .
  - 8:     Search  $\hat{k}$  such that  $u_{Sl}(\hat{n}, \hat{k}) \geq u_{Sl}(\hat{n}, k)$  for any  $k$
  - 9:     **if**  $U_{Sl}(\hat{n}, \hat{k})$  is firing **then**
  - 10:       Reinforce  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  and  $b_l(\hat{k})$  subject to (3.13) and (3.14).
  - 11:       Rotate  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  and copy the rotated connections subject to (3.15).
  - 12:     **else**
  - 13:       Make a new cell-plane stack
  - 14:        $\hat{k} = k^* \leftarrow$  the serial number of the new cell-plane stack
  - 15:       Reinforce  $a_l(\nu, \hat{n}, \kappa, k^*)$  and  $b_l(k^*)$  subject to (3.13) and (3.14).
  - 16:       Rotate  $a_l(\nu, \hat{n}, \kappa, k^*)$  and copy the rotated connections subject to (3.15).
  - 17:     **end if**
  - 18:   **end while**
  - 19: **end while**
-

転情報を保持したままどのようにして特徴検出を行なうかを説明する。その後、細胞面群内の C-細胞の特性について説明する。

### 3.4.1 S-細胞による回転した局所パターンの検出

3.2.1 節において  $U_1$  細胞段に細胞面群を導入することで方位を持つ線分が単一の特徴として扱うことが可能となることを説明した。今一度  $U_1$  細胞段における細胞面群の性質をまとめると以下ようになる。

- 細胞面群は複数の細胞面からなり、細胞面群内には 3 次元的に S-細胞および C-細胞が配置されている。
- 各細胞の位置を示す  $\mathbf{n}$  は  $Z^3$  の要素であり、 $\mathbf{n} = (n_x, n_y, n_\theta)$  と表わされる。
- $U_1$  細胞段の各細胞面内の細胞は図 3.3(b) に示すような方位選択性を持ち、あらかじめ定められた方位の線分を検出する。すなわち  $U_0$  層に提示されたパターンを構成する線分の方位は、発火している細胞の位置  $n_\theta$  で表現される。
- 同様に、線分の位置は  $n_x$  および  $n_y$  で表現される。

回転対応型ネオコグニトロンではより高次の  $U_2$  細胞段にも細胞面群を導入している。ここでは図 3.7 を用いて、 $U_{S2}$  層の S-細胞がどのようにして回転した局所パターンを検出するかを説明する。図 3.7(a) は  $U_0$  層に  $\angle$ -型の局所パターンが提示され、S-細胞  $U_{S2}(\hat{\mathbf{n}}, \hat{\mathbf{k}})$ ,  $\hat{\mathbf{n}} = (\hat{n}_x, \hat{n}_y, \hat{n}_\theta = 0^\circ)$  がシードセルとして選択されたときの学習の様子を示している。このときのシードセルの興奮性可変結合  $a_l(\nu, \hat{\mathbf{n}}, \kappa, \hat{\mathbf{k}})$  は図 3.7(a) の実線の矢印のような空間分布を持つようになる (アルゴリズム 2-10, 式 (3.13))。また、シードセルと同じ細胞面群に属しているが細胞面が異なる S-細胞の興奮性可変結合、例えば図 3.7(b) に示すように  $\mathbf{n} = (n_x, n_y, n_\theta = 22.5^\circ)$  に位置する S-細胞の興奮性可変結合はシードセルの興奮性可変結合を  $22.5^\circ$  回転したものとなる (アルゴリズム 2-11, 式 (3.15))。この状態で未知のパターンである  $22.5^\circ$  回転した  $\angle$ -型のパターンが  $U_0$  層に提示されると、 $U_{S2}(\mathbf{n}, \hat{\mathbf{k}})$ ,  $\mathbf{n} = (\hat{n}_x, \hat{n}_y, 22.5^\circ)$  に位置する S-細胞が最大に発火することになる。すなわち、

- $U_0$  層に  $\Delta\theta$  回転したパターンが提示されると、発火する S-細胞の位置も  $n_\theta$ -方向に  $\Delta\theta$  の量だけずれる。すなわち  $U_0$  層に提示されたパターンの回転角は発火している細胞の位置  $n_\theta$  で表現される。
- 従来型ネオコグニトロンと同様に、 $U_0$  層内でパターンが平行移動すると発火する S-細胞の位置も  $n_x, n_y$ -平面内で平行移動する。すなわち  $U_0$  層に提示されたパターンの位置は発火している細胞の位置  $(n_x, n_y)$  で表現される。

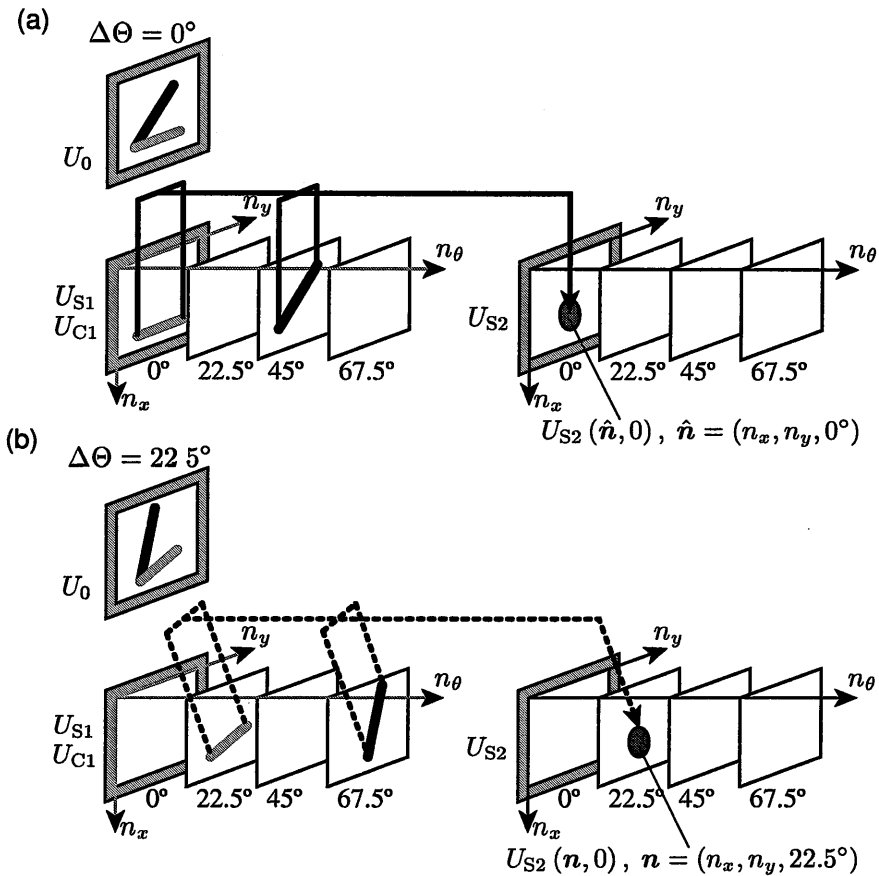


図 3.7: 回転対応型ネオコグニトロン の  $U_{S2}$  層における回転した局所パターンの検出. (a):  $U_0$  層に  $\angle$ -型の局所パターンが提示され, S-細胞  $U_{S2}(\hat{n}, \hat{k})$ ,  $\hat{n} = (\hat{n}_x, \hat{n}_y, 0^\circ)$  がシードセルとして選択されたときの学習の様子を示している. 強化されたシードセルの興奮性可変結合  $a_l(\nu, \hat{n}, \kappa, \hat{k})$  は図中で実線の矢印で示している. (b): 点線の矢印はシードセルが属する細胞面群内の  $n = (n_x, n_y, 22.5^\circ)$  に位置する S-細胞の興奮性可変結合を表わしている. 式 (3.15) に従い, シードセルの興奮性可変結合を  $22.5^\circ$  回転した結合となっている ( $n_\theta = 22.5^\circ, \hat{n}_\theta = 0^\circ \Rightarrow R(n_\theta - \hat{n}_\theta = 22.5^\circ)$ ).

このように  $U_{S2}$  層の S-細胞も  $U_{S1}$  層の S-細胞と同様に、局所パターンの位置と方位を同時に検出する細胞であることがわかる。これら S-細胞の特徴は更に後層の  $U_{S3}$  層においても同じである。

回転対応型ネオコグニトロンでは発火している細胞が  $n_\theta$ -方向にずれることは検出している局所パターンが回転していることを意味する。

### 3.4.2 C-細胞による回転によるずれの吸収

C-細胞も S-細胞と同様に、細胞面群内に 3 次元的に配置されている。回転対応型ネオコグニトロンにおける  $U_{C1}$  層の C-細胞のずれを許容する範囲  $D_l$  も 3 次元に拡張されている。すなわち  $n_\theta$  方向へのずれも許容していることになるが、これは  $U_{S1}$  層で検出された局所的パターンがある程度回転しても C-細胞が発火しつづけることを意味する。回転対応型ネオコグニトロンの C-細胞の特徴をまとめると以下の項目になる。

- $U_{S1}$  層で検出された局所的パターンが  $D_l^{(x)}, D_l^{(y)}$  で定義される範囲内で平行移動しても  $U_{C1}$  層の C-細胞は発火し続ける。
- 同様に局所的パターンが  $D_l^{(\theta)}$  で定義される範囲内で回転しても  $U_{C1}$  層の C-細胞は発火し続ける。

図 3.8 に  $U_{C3}$  層の C-細胞を例にして上記の特徴を説明する。図 3.8(a) は  $U_{C3}$  層のある S-細胞  $U_{S3}(n, k)$  が白い円内に示す 3 つの局所パターンを学習し A-型のパターンを学習した状態を示している。C-細胞  $U_{C3}(n, k)$  の位置ずれの許容範囲は灰色の円で、回転ずれの許容範囲は矢印で示している。従来型ネオコグニトロンでは各局所パターンが灰色の円内に存在すれば  $U_{C3}(n, k)$  は発火し続ける (図 3.8(b))。回転対応型ネオコグニトロンでは更に各局所パターンが矢印で示す範囲で回転していても発火しつづける (図 3.8(c))。

回転対応型ネオコグニトロンではこのような回転ずれの許容機能を導入しているのでパターンの変形に更に頑健であることが期待される。また従来型ネオコグニトロンでは大きく変形したパターンも最終層の 1 つの C-細胞で検出することが可能であった。回転対応型ネオコグニトロンでは最終層の 1 つの C-細胞で任意の角度に回転したパターンを検出可能となることが期待される。

## 3.5 評価実験

回転対応型ネオコグニトロンの性能を評価するために数値シミュレーションを行なう。まず 2.4 節で従来型ネオコグニトロンに対して行なったパターンに対する評価実験と同様の実験を行ない、次に手書き数字データベースを用いた評価実験をする。

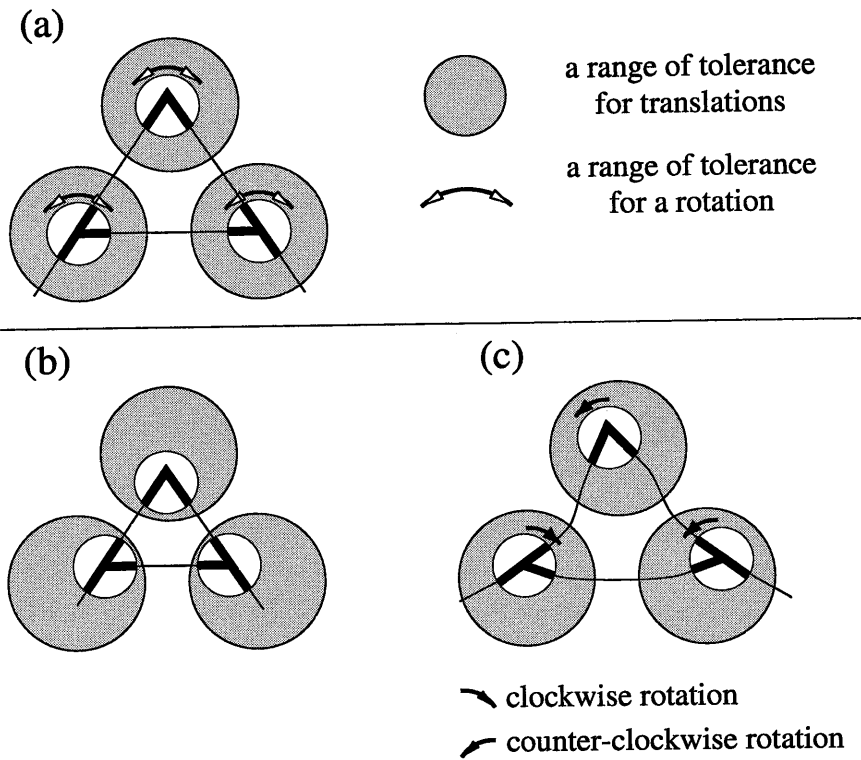


図 3.8: C-細胞の回転ずれの許容範囲. (a):  $U_{S3}$  層のある S-細胞  $U_{S3}(n, k)$  が 3つの局所パターンで構成される A-型のパターンを学習したときの  $U_{C3}$  層の C-細胞  $U_{C3}(n, k)$  の許容範囲を示す. 灰色の円が位置ずれの許容範囲, 両矢印が回転ずれの許容範囲を示す. (b): 従来型ネオコグニトロンでは各局所パターンが灰色の円内に存在すれば  $U_{C3}(n, k)$  発火し続けたが, (c): 回転対応型ネオコグニトロンでは局所パターンが矢印で示す範囲で回転しても発火し続ける.



表 3.3: 3.5.1 節の数値実験で用いた細胞面数と細胞面内の細胞数に関するパラメータの値. \* で示される値は定義されていないことを示す. また, 第2細胞段以降の細胞面は学習により自動的に生成されるので学習前に細胞面数を決定することができない. 括弧内の値は学習終了後の値である.

	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$K_{Sl}$	*	1	(4)	(5)	(6)
$N_{Sl}$	*	59	17	13	3
$T_{Sl}$	*	16	8	4	2
$K_{Cl}$	1	1	(4)	(5)	(6)
$N_{Cl}$	61	19	17	10	1
$T_{Cl}$	1	8	4	2	1

表 3.4: 結合野の大きさ. 各次元における要素の濃度を示す. 各項目において  $|A_l^{(x)}| \times |A_l^{(y)}| \times |A_l^{(\theta)}|$ , または  $|D_l^{(x)}| \times |D_l^{(y)}| \times |D_l^{(\theta)}|$  という表記方法を用いている. 例えば  $A_3 = 5 \times 5 \times 4$  は  $|A_3^{(x)}| = 5, |A_3^{(y)}| = 5, |A_3^{(\theta)}| = 3$  となり,  $|A_3| = 75$  である.

	$l=1$	$l=2$	$l=3$	$l=4$
$A_l$	$3 \times 3 \times 1$	$3 \times 3 \times 8$	$5 \times 5 \times 4$	$3 \times 3 \times 2$
$D_l$	$5 \times 5 \times 3$	$3 \times 3 \times 3$	$5 \times 5 \times 3$	$3 \times 3 \times 2$

### 3.5.1 回転パターンに対する評価実験

2.4 節で行なった評価実験を回転対応型ネオコグニトロンに対して行なう. 各パラメータを表 3.3, 3.4, 3.5 に示す. 学習用パターンは図 2.6 を使用した. 評価用の回転したパターンと回転対応型ネオコグニトロン認識結果を図 3.9 に示す. この結果から回転対応型ネオコグニトロンは正立したパターンだけを学習したにもかかわらず, パターンの回転に対して完全に頑健であることがわかる.

表 3.5: S-細胞のしきい値および学習定数.

	$l=1$	$l=2$	$l=3$	$l=4$
$r_l$	3.8	2.7	3.5	1.9
$q_l$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$



図 3.9: 回転対応型ネオコグニトロンの回転したパターンに対する認識結果。灰色の部分は従来型ネオコグニトロンが認識できる範囲を示す。回転対応型ネオコグニトロンはパターンの回転に対して完全に頑健であることがわかる。

表 3.6: 3.5.2 節で用いた細胞面数と細胞面内の細胞数に関するパラメータの値. \* で示される値は定義されていないことを示す. また, 第 2 細胞段以降の細胞面は学習により生成されるので学習前に細胞面数を決定することができない. 括弧内の値は学習終了後の値である.

	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$K_{Sl}$	*	1	(6)	(7)	(9)
$N_{Sl}$	*	59	17	13	3
$T_{Sl}$	*	16	8	4	2
$K_{Cl}$	1	1	(6)	(7)	(9)
$N_{Cl}$	61	19	17	10	1
$T_{Cl}$	1	8	4	2	1

表 3.7: S-細胞のしきい値および学習定数.

	$l=1$	$l=2$	$l=3$	$l=4$
$r_l$	3.8	1.5	1.86	1.22
$q_l$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$	$1.0 \times 10^5$

### 3.5.2 手書き数字パターンに対する認識性能

前節では人工的なパターンを使って評価を行なった. 提案する数理モデルは手書き数字のような実世界のパターンに対しても頑健であることが期待される. 本節では手書き英数字データベース ETL1 [34] を用いて回転対応型ネオコグニトロン性能評価実験を行なう. 使用した学習用パターンの一例を図 3.10 に示す. また前節のパターンの実験で回転対応型ネオコグニトロンはパターンの回転に対して完全に頑健であることが確認されたので, パターン '6' とパターン '9' を識別できないものと思われる. 本実験では '6' と '9' は同じクラスに属するパターンとして認識の成功/失敗を判断している.

その他のパラメータを表 3.6, 3.7 に示す. 結合に関するパラメータは前節の実験で用いたもの (表 3.4) と同じである.

学習用パターンを用いて回転対応型ネオコグニトロンを構築した後, 1,000 種の評価用パターンを用いて認識実験を行なった. 図 3.11 に正しく認識できたパターンの一例を示す. 変形したパターンも認識可能であることが確認できた. これは回転対応型ネオコグニトロンが従来型ネオコグニトロンの特徴を継承していることを示す. 評価用パターンの認識率は全体で 96.4% であった. 一方, 従来型ネオコグニトロンで同様の学習用パターンおよび評価用パターンを用いて実験を行なった場合の認識率が 88.2% であることを確認して

0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9

図 3.10: 手書き数字認識に対する評価実験に用いる学習用パターンの一例。電子技術総合研究所発行の ETL1 データベースの一部である。実際には各字種に対して 50 種、合計 500 パターンを学習用に使用した。

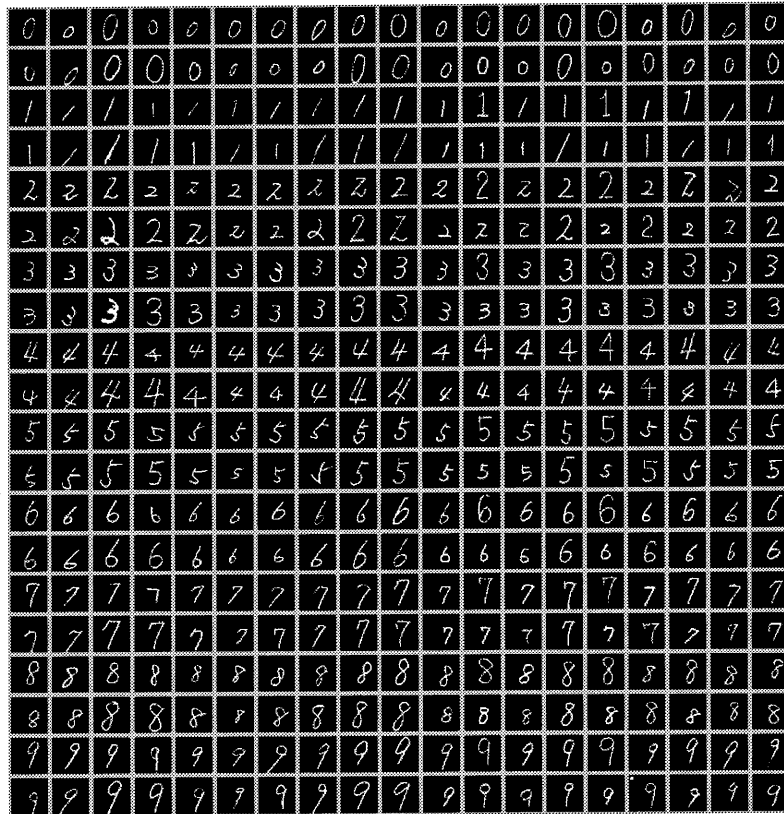


図 3.11: 回転対応型ネオコグニトロンが正しく認識できた手書き数字の一例. ただし回転対応型ネオコグニトロンは回転に対して完全に頑健であるために '6' と '9' を識別できない. ここでは '6' と '9' を同一の字種とみなして認識の成功/失敗を判断している.

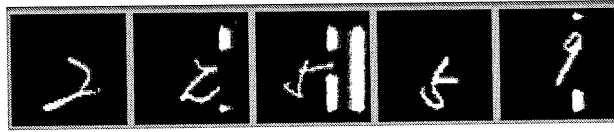


図 3.12: 回転対応型ネオコグニトロンが誤認識した手書き数字の一例。パターン  
の極度の変形が主な原因である。

いる。この結果より回転対応型ネオコグニトロンが変形に対して更に頑健であることがわかった。

誤認識したパターンの例を図 3.12 に示す。誤認識した原因はそのほとんどがパターンの極度の変形であった。稀に、変形は大きくないが小さすぎるパターンが存在し特徴を検出できないことが原因である場合もあった。

### 3.5.3 回転した手書き数字パターンに対する認識性能

次に ETL1 データベース中のパターンを計算機上で回転させたパターンを評価用パターンとして用いて認識性能を評価する。具体的には前節で構築した回転対応型ネオコグニトロンに、図 3.11 に例示したような評価用パターンを文献 [32] で提案されているアルゴリズムを用いて人工的に回転パターンを作成して評価した。作成した回転パターンは各クラス毎に  $30^\circ, 60^\circ, \dots, 330^\circ$  回転させたものであり、全体で 22,000 パターンである。この人工回転パターンに対する認識率は 95.790% であった。この結果は回転対応型ネオコグニトロンが手書き数字の回転に対しても非常に頑健であることを示している。

## 3.6 まとめと考察

ヒトの回転したパターンの認識機構を生理学的および心理学的に考察し、従来型ネオコグニトロンを基本とした回転対応型ネオコグニトロンを提案した。実際に数値シミュレーションにより回転対応型ネオコグニトロンを構築し、人工パターンおよび、実際の手書き数字に対してもパターンの回転・変形・平行移動・拡大縮小に非常に頑健であることが確認された。ネオコグニトロンを用いて回転したパターンを認識するには単純に考えれば各回転角に対応した複数のネオコグニトロンを構築すればよいが、この場合、細胞数の増加や学習に長い時間を要するといった問題点がある。

回転したパターンを直接認識する数理モデルとして回転対応型ネオコグニトロンを提案し、英数字を学習用および評価用パターンとして用いて実験を行なった。しかし実際には Cooper と Shepard らによる心理実験によると英数字は「単純なパターン」ではないことがわかっている。すなわちヒトは英

数字を回転に対して単純とはみなしておらず，回転した英数字を認識するにはトップダウン的な知識を必要とする．以降，第5章では以上の事実を考慮に入れた回転対応型ネオコグニトロンを基にした数理モデルを構築する．

次章，第4章では従来型ネオコグニトロンおよび回転対応型ネオコグニトロンを様々なパターンに適応する際の問題点を調査し，その問題点を解決するための新しいアルゴリズムを提案し検証する．

## 第4章 高速生成アルゴリズム

本章では従来型ネオコグニトロンおよび回転対応型ネオコグニトロンに共通する内容を記述するため、両者をネオコグニトロン型モデルと呼ぶことにする。第1章で述べたようにヒトの認識機能の実現を目的として提案される数理モデルは工学的にも有効なモデルである必要がある。本章ではネオコグニトロン型モデルを様々なパターンに適應する際の問題点を調査し、それを解決する新しいアルゴリズムを提案する。また、その有効性を数値シミュレーションにより評価する。

### 4.1 パラメータ設定に関する問題点

ネオコグニトロン型モデルは手書き数字の認識だけではなくアルファベットや日本語の認識、楽譜の認識 [35] 等に適應可能である。適應する対象が異なる場合にはパラメータの値もそれぞれに対して適切なものに変更する必要がある。ネオコグニトロン型モデルはS-細胞による特徴抽出と、C-細胞によるずれの許容を数回繰り返すことでパターンの多様性に頑健となるモデルであるが、特にS-細胞のしきい値  $r_i$  ( $U_{Si}$  層のS-細胞のしきい値) はネオコグニトロン型モデルの数あるパラメータの中でも非常に重要な役割を果しており、認識性能に最も影響を及ぼすものである。

一般に  $r_i$  を解析的に設定する統一的な方法は現時点 (2000年2月) では存在しないため、

1. 各  $U_{Si}$  層において  $r_i$  の可能な組み合わせ (パラメータセット) を用意し、
2. 全パラメータセットに対応するネオコグニトロン型モデルを大量に構築し、
3. 構築した全てのネオコグニトロン型モデルに対して学習用パターンによる評価実験を行ない、
4. 全ネオコグニトロン型モデルの中で最も高い認識性能を示すパラメータセットを選定する、

以外に方法はない。例えば、構築するネオコグニトロン型モデルが4つの細胞段から構成され、各層において10種類のしきい値を用意すると、パラメータセットの数は1,000個となり、各パラメータセットを持つ1,000種のネオコグニトロン型モデルに対して学習および認識性能の評価を行なわなければならない。すなわちこの方法はパラメータ空間の全探索であるために膨大な



時間を要する。Lovell らはこの全探索を特に SHOP (Selectivity Hunting to Optimize Performance,[36]) と呼んでいるが文献 [37] において次のような指摘をしている (式は以降の議論との対応のために改変している)。

Execution time cannot be ignored in our assessment of SHOP's feasibility. If  $t_T$  is the time take to train the neocognitron,  $t_P$  is the time taken to propagage activity through the network and  $N_P$  is the number of examples in the validation set, then the time taken to execute SHOP is

$$t_{\text{test}}^{(\text{SHOP})} = |r_1| \times |r_2| \times \cdots \times |r_L| \times N_P \times t_T, \quad (4.1)$$

Fukushima cites values of  $t_T = 780[\text{sec}]$ ,  $t_P = 3.3[\text{sec}]$  for a simulation of the neocognitron, written in FORTRAN and running on a SUN Sparcstation. If we specified ten possible selectivity values for each layer of a four-layer network and used a validation set with 400 patterns, SHOP would take about **eight months** to come up with a good set of selectivities on Professor Fukushima's computer. In a commercial/industrial situation, a SHOP trained neocognitron would have to be implemented with dedicated hardware (*e.g.*, [38]).

ここで  $r_l$  は  $U_{Sl}$  層におけるしきい値  $r_l$  の可能な値を要素とする集合であり、

$$r_l = \{r_l^{(1)}, r_l^{(2)}, \dots, r_l^{(|r_l|)}\}, \quad (4.2)$$

である。Lovell らは 1,000 種のネオコグニトロン型モデルを作成し評価するには約 8ヶ月の時間を要するため、専用ハードウェアで実現する必要があると主張しているが、実際にはソフトウェア上で、すなわち数値シミュレーションによって事前に性能評価をする必要がある。また近年の急速な CPU の性能向上に伴ない実行時間の問題は徐々に薄れていくかもしれないが、それでも全探索を行わなければならないので根本的な解決にはならない。

そこで本章ではネオコグニトロン型モデルの学習過程の解析を行ない、その結果を用いて学習時間に関する問題の根本的解決を試みる。具体的には 1つのネオコグニトロン型モデルを生成する時間を劇的に減少させるための新しい生成アルゴリズムを提案する。ネオコグニトロン型モデルの解析は S-細胞の特性に限り Fukushima [26] によって行なわれているため次節でこれを紹介する。その後で学習過程のベクトル空間上における解析を試みる。

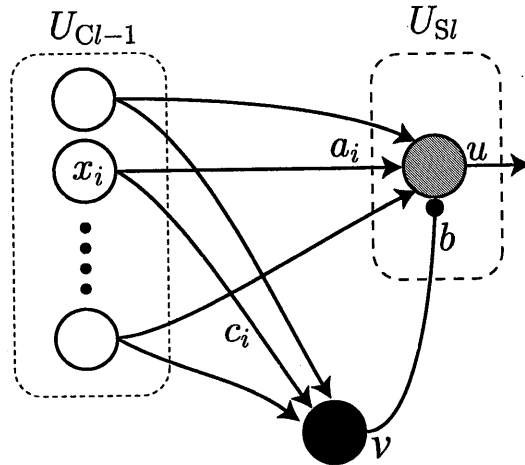


図 4.1: C-細胞と S-細胞の結合関係. 説明を簡単にするために各細胞と結合の記述を簡単な形にしてある.  $x_i$  は  $U_{Cl-1}$  層の C-細胞の発火頻度,  $u$  は  $U_{Sl}$  層の S-細胞の発火頻度,  $v$  は V-細胞の発火頻度をそれぞれ表わす.  $a_i$  は C-S 細胞間の興奮性可変結合,  $b$  は V-S 細胞間の抑制性可変結合,  $c_i$  は C-V 細胞間の興奮性固定結合をそれぞれ表わす.

## 4.2 S-細胞のベクトル空間上での解析

### 4.2.1 ベクトル空間上での解析

図 2.4(a) に  $U_{Cl-1}$ - $U_{Sl}$  間の結合を示したが, 説明を簡単にするために図 4.1 に示す表記を用いることにする (一般性は失われない). 各表記は次の意味を持つ.  $x_i$  は  $U_{Cl-1}$  層の C-細胞の発火頻度,  $u$  は  $U_{Sl}$  層の S-細胞の発火頻度,  $v$  は V-細胞の発火頻度をそれぞれ表わす.  $a_i$  は C-S 細胞間の興奮性可変結合,  $b$  は V-S 細胞間の抑制性可変結合,  $c_i$  は C-V 細胞間の興奮性固定結合をそれぞれ表わす.

以上の表記を用いて式 (2.1) および (2.5) を書き替えると次の式になる.

$$u = r \cdot \phi \left[ \frac{1 + \sum_i a_i x_i}{1 + \frac{1}{r} b v} - 1 \right], \quad (4.3)$$

ただし,

$$v = \sqrt{\sum_i c_i \{x_i\}^2} \quad (4.4)$$

であり, 添字  $i$  は集合  $A_l$  の要素  $v$  の番号である. 関数  $\phi$  に関して  $\phi(r y) =$

$r \cdot \phi(y)$  が成立するので、式 (4.3) を更に書き替えると、

$$u = \gamma \phi \left[ \frac{\sum_i a_i x_i}{bv} - \frac{r}{1+r} \right] \quad \text{if } bv > 0, \quad (4.5)$$

となる。ただし  $\gamma$  は次式で与えられる。

$$\gamma = \frac{rbv}{1 + \frac{r}{r+1}bv} \quad (4.6)$$

可変結合  $a_i$  と  $b$  は学習によってその値が変化する。仮にあるパターン  $\mathbf{x}$  に対してこの S-細胞がシードセルもしくはシードセルが属する細胞面群となった場合（以下「勝者となる」と言う）には、以下の式で値が変化する。

$$\Delta a_i = q \cdot c_i \cdot x_i, \quad (4.7)$$

$$\Delta b = q \sqrt{\sum_i c_i \cdot x_i^2}. \quad (4.8)$$

ここでこの S-細胞が学習用パターンの中で  $K$  個のパターン  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(K)}$  に対して勝者になった場合を考える。すると、学習終了後の  $a_i$  と  $b$  は次式で表わすことができる。

$$a_i = q \sum_k c_i \cdot x_i^{(k)}, \quad (4.9)$$

$$b = q \sum_k \sqrt{\sum_i c_i \cdot \{x_i^{(k)}\}^2}.$$

式 (4.9) を式 (4.5) に代入すると次式が得られる。

$$u = \gamma \phi \left[ \frac{\sum_k \sum_i c_i \cdot x_i^{(k)} \cdot x_i}{\sum_k \sqrt{\sum_i c_i \{x_i^{(k)}\}^2} \sqrt{\sum_i c_i \{x_i\}^2}} - \frac{r}{1+r} \right] \quad (4.10)$$

ここで任意のベクトル  $\mathbf{x}$  とベクトル  $\mathbf{y}$  の内積を次式で定義する。

$$\langle \mathbf{x}, \mathbf{y} \rangle \stackrel{\text{def}}{=} \sum_i c_i \cdot x_i \cdot y_i. \quad (4.11)$$

すると式 (4.10) は次式のように書き替えられる。

$$u = \gamma \phi \left[ \frac{\|\mathbf{X}\| \langle \mathbf{X}, \mathbf{x} \rangle}{\sum_k \|\mathbf{x}^{(k)}\| \|\mathbf{X}\| \cdot \|\mathbf{x}\|} - \frac{r}{1+r} \right]. \quad (4.12)$$

ただし

$$\mathbf{X} \stackrel{\text{def}}{=} \sum_k \mathbf{x}^{(k)}, \quad (4.13)$$

$$\|\mathbf{y}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}, \quad (4.14)$$

である。更に  $s$ ,  $\lambda$ ,  $\theta$  を次式で定義する。

$$s \stackrel{\text{def}}{=} \frac{\langle \mathbf{X}, \mathbf{x} \rangle}{\|\mathbf{X}\| \cdot \|\mathbf{x}\|}, \quad (4.15)$$

$$\lambda \stackrel{\text{def}}{=} \frac{\|\mathbf{X}\|}{\sum_k \|\mathbf{x}^{(k)}\|}, \quad (4.16)$$

$$\theta \stackrel{\text{def}}{=} \frac{r}{1+r}. \quad (4.17)$$

式 (4.3) は次式に書き替えられる。

$$u = \gamma \cdot \lambda \cdot \phi \left[ s - \frac{\theta}{\lambda} \right]. \quad (4.18)$$

また一般に  $\lambda \sim 1$  としてよい [26]. 常に  $\lambda = 1$  となるようにするには抑制性可変結合  $b$  の値を以下の式で与えればよい. [39].

$$b = \sqrt{\sum_i \frac{\{a_i\}^2}{c_i}}. \quad (4.19)$$

以降  $\lambda = 1$  として議論を進めるために、抑制性可変結合の値  $b$  は上記の式 (4.19) で与えられるものとする。  $\lambda = 1$  であるので最終的に式 (4.3) は次式のような非常に単純な書き替えられる。

$$u = \gamma \cdot \phi [s - \theta]. \quad (4.20)$$

#### 4.2.2 しきい値と S-細胞の発火条件の関係

前節の解析結果からある 1 つの S-細胞の発火条件を図示したものが図 4.2 である。式 (4.20) より、S-細胞は  $s$  の値が  $\theta$  を越えたときにのみ発火することがわかる。S-細胞と結合している C-細胞の発火頻度を要素としてもつベクトルをパターンベクトル  $\mathbf{x}$  と呼ぶことにすると、 $s$  の定義式である式 (4.15) は S-細胞に入力されたパターンベクトル  $\mathbf{x}$  と  $\mathbf{X}$  の正規化された類似度である。以降、 $\mathbf{X}$  を S-細胞の参照ベクトルと呼ぶことにする。また  $\theta$  は S-細胞の実質的なしきい値と考えることができる。以降  $\theta$  を実効しきい値、または単にしきい値と呼ぶことにする。

S-細胞が発火を持続するパターンベクトル  $\mathbf{x}$  の範囲を許容範囲と呼ぶことにする。図 4.2 はこの関係を図示したものであり、 $\theta$  の値によって定められ

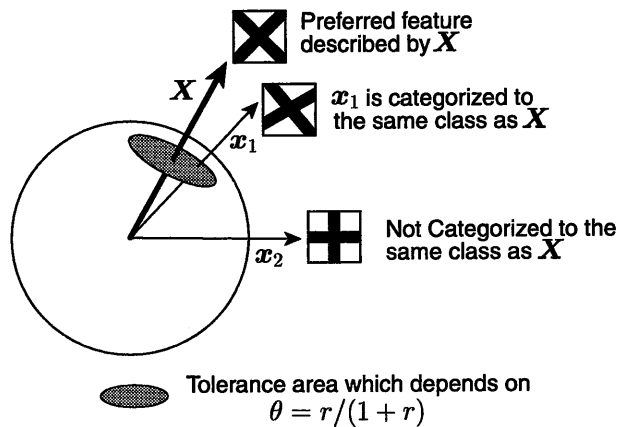


図 4.2: S-細胞の発火条件.  $X$  は S-細胞の参照ベクトル.  $x_1$ ,  $x_2$  は S-細胞に  
 入力されるパターンベクトルである. パターンベクトル  $x_1$  は  $X$  を持つ S-  
 細胞により検出され, この S-細胞は  $x_2$  に対しては反応しない.

る許容範囲 (図 4.2 の灰色の楕円) 内にパターンベクトル  $x$  が入っていると  
 ときにのみ反応する. 例えば図 4.2 に示す  $x_1$  がパターンベクトルである場合  
 にはこの S-細胞は発火する. 一方,  $x_2$  に対しては許容範囲外にあるので発火  
 しない<sup>1</sup>. この許容範囲は超球内における角度  $\tau$  で表現することができ, 次  
 式で与えられる.

$$\tau = \cos^{-1} \theta = \cos^{-1} \frac{r}{1+r}.$$

この許容範囲は実効しきい値  $\theta$  で与えられる. 図 4.3(a) はしきい値が高い  
 場合を図示したものであり, 許容範囲は狭くなっている. この場合パターン  
 ベクトル  $x$  が少しでも参照ベクトル  $X$  からずれるとこの S-細胞は発火しな  
 い. 逆に図 4.3(b) はしきい値が低い場合を図示しており, 許容範囲は広くな  
 る. この場合はパターンベクトルがかなり大きく変形したとしてもこの S-細  
 胞は発火し続ける.

### 4.3 学習過程の解析

次に回転対応型ネオコグニトロン<sup>1</sup>の学習アルゴリズム (アルゴリズム 2) と  
 しきい値の関連を調査する. 回転対応型ネオコグニトロン<sup>1</sup>の学習アルゴリ  
 ズムは従来型ネオコグニトロン<sup>1</sup>の学習アルゴリズムを包含しているため, アル  
 ゴリズム 2 だけを考慮しても一般性は失われない.

<sup>1</sup>図 4.2 は 1 つの S-細胞について図示したものである. 回転対応型ネオコグニトロン<sup>1</sup>におい  
 ては  $w_2$  はこの S-細胞で検出されることはないが, 図にある S-細胞が属する細胞面群内の他の  
 S-細胞が  $w_2$  を検出することになる.

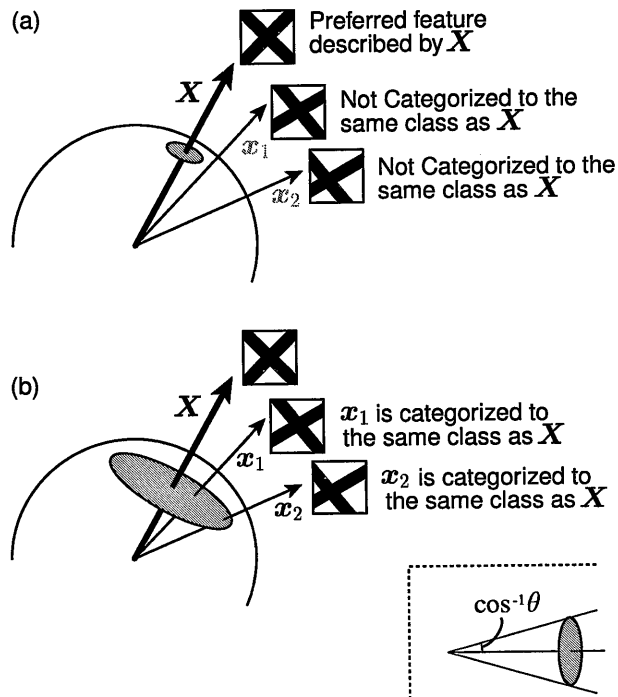


図 4.3: S-細胞の発火条件.  $X$  はある S-細胞の参照ベクトル,  $x_1, x_2$  は S-細胞へのパターンベクトルである. (a): しきい値が高い場合. S-細胞が発火を持続するパターンベクトルの許容範囲は狭くなる. この場合参照ベクトルから少し変形したパターンに対しても発火を停止する. しきい値が高過ぎると変形に対する頑健性が低下する. 高過ぎるしきい値は汎化性能の低下を招く. (b): しきい値が低い場合. 許容範囲は広がり, 大きく変形したパターンに対しても発火を持続する. しきい値が低過ぎると適切に特徴を識別できなくなりモデル全体の性能を低下させる.

前節の Fukushima による解析をより一般的なものとするために以下の式を与える. ある S-細胞  $U_{Sl}(\mathbf{n}, k)$  が持つ参照ベクトルを  $\mathbf{a}_l(\mathbf{n}, k)$  と表すことにすると

$$\mathbf{a}_l(\mathbf{n}, k) \stackrel{\text{def}}{=} \begin{bmatrix} a_l(\boldsymbol{\nu}_1, \mathbf{n}, 1, k) \\ \vdots \\ a_l(\boldsymbol{\nu}_{|A_l|}, \mathbf{n}, 1, k) \\ a_l(\boldsymbol{\nu}_1, \mathbf{n}, 2, k) \\ \vdots \\ a_l(\boldsymbol{\nu}_{|A_l|}, \mathbf{n}, 2, k) \\ \vdots \\ a_l(\boldsymbol{\nu}_1, \mathbf{n}, K_{Cl-1}, k) \\ \vdots \\ a_l(\boldsymbol{\nu}_{|A_l|}, \mathbf{n}, K_{Cl-1}, k) \end{bmatrix}, \quad (4.21)$$

と書くことができる. また  $U_{Sl}(\mathbf{n}, k)$  の入力であるパターンベクトルを  $\mathbf{u}_{Cl-1}(\mathbf{n})$  と表記すると,

$$\mathbf{u}_{Cl-1}(\mathbf{n}) \stackrel{\text{def}}{=} \begin{bmatrix} u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_1, 1) \\ \vdots \\ u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_{|A_l|}, 1) \\ u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_1, 2) \\ \vdots \\ u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_{|A_l|}, 2) \\ \vdots \\ u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_1, K_{Cl-1}) \\ \vdots \\ u_{Cl-1}(\mathbf{n} \oplus \boldsymbol{\nu}_{|A_l|}, K_{Cl-1}) \end{bmatrix}, \quad (4.22)$$

となり、S-細胞に関する興奮性固定結合  $c_l$  は次のように書ける。

$$c_l \stackrel{\text{def}}{=} \begin{bmatrix} c_l(\nu_1) \\ \vdots \\ c_l(\nu_{|A_l|}) \\ c_l(\nu_1) \\ \vdots \\ c_l(\nu_{|A_l|}) \\ \vdots \\ c_l(\nu_1) \\ \vdots \\ c_l(\nu_{|A_l|}) \end{bmatrix}. \quad (4.23)$$

また次のベクトル  $a'_l(\mathbf{n}, k)$  を定義する。

$$a'_l(\mathbf{n}, k) \stackrel{\text{def}}{=} \begin{bmatrix} a_l(\nu_1, \mathbf{n}, 1, k) / c_l(\nu_1) \\ \vdots \\ a_l(\nu_{|A_l|}, \mathbf{n}, 1, k) / c_l(\nu_{|A_l|}) \\ a_l(\nu_1, \mathbf{n}, 2, k) / c_l(\nu_1) \\ \vdots \\ a_l(\nu_{|A_l|}, \mathbf{n}, 2, k) / c_l(\nu_{|A_l|}) \\ \vdots \\ a_l(\nu_1, \mathbf{n}, K_{Cl-1}, k) / c_l(\nu_1) \\ \vdots \\ a_l(\nu_{|A_l|}, \mathbf{n}, K_{Cl-1}, k) / c_l(\nu_{|A_l|}) \end{bmatrix}. \quad (4.24)$$

以上の記法より、式 (4.15) の類似度は以下の式になる。

$$s_l(\mathbf{n}, k) \stackrel{\text{def}}{=} \frac{\langle a'_l(\mathbf{n}, k), u_{Cl-1}(\mathbf{n}) \rangle_l}{\|a'_l(\mathbf{n}, k)\| \cdot \|u_{Cl-1}(\mathbf{n})\|}, \quad (4.25)$$

ただし、

$$\langle \mathbf{x}, \mathbf{y} \rangle_l = \sum_i x_i \cdot c_{l,i} \cdot y_i, \quad (4.26)$$

である。式 (4.25) は  $U_{Sl}(\mathbf{n}, k)$  が持つ参照ベクトルとパターンベクトルとの類似度を意味する。

一般的に高いしきい値で学習すると細胞面群の数が増大し、逆に低いと減少することが知られている。この現象はパターンベクトルが分布している領域を何個の許容範囲で包含できるかという問題から発生する。すなわち高いしきい値を持つS-細胞は許容範囲が狭いためにパターンベクトルが分布している領域を包含するためにはより多くの参照ベクトルと許容範囲が必要とな



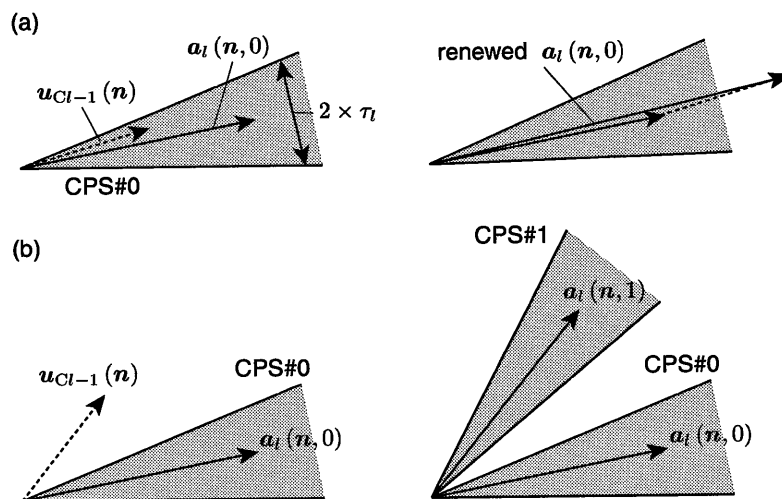


図 4.4: ベクトル空間内における学習過程の考察. (a): 左側の図は学習過程で既に  $U_{Sl}$  層内に 1 つの細胞面群 (CPS#0,  $k = 0$ ) が生成されており, その参照ベクトルが  $a_l(n, 0)$  である S-細胞の許容範囲内に, 学習用パターンのベクトル  $u_{Cl-1}(n)$  が入力された状況を表している. この場合はこの学習用パターンのベクトルは CPS#0 内の S-細胞によって学習が行なわれる. (b): 一方,  $u_{Cl-1}(n)$  が  $k = 0$  の細胞面群 (CPS#0) の許容範囲外にある場合は新たな細胞面群 (CPS#1) が生成される.

り, 逆に低いしきい値の場合には少数の参照ベクトルと許容範囲で済む. この現象はアルゴリズム 2 の 9 行目と 12 行目の条件分岐がそれぞれ

**9 行目** パターンベクトルに対して位置  $n$  の位置にある S-細胞が発火しているならば, その中で最大発火している S-細胞に対して学習則を適応し,

**12 行目** 位置  $n$  の位置にある S-細胞がどれも発火していないのであれば新しく細胞面群を作成し, その細胞面群内の位置  $n$  に配置されている S-細胞に学習則を適応する.

という意味であり, 発火している/していないという条件が学習アルゴリズムの中に含んでいることから理解できる.

図 4.4 にベクトル空間内における学習過程の様子を示す. 図 4.4(a) の左側の図は (i): 学習過程で既に  $U_{Sl}$  層内に 1 つの細胞面群 (CPS#0,  $k = 0$ ) が生成されており, (ii): その参照ベクトルが  $a_l(n, 0)$  であり, (iii): S-細胞の許容範囲内にある学習用パターンのベクトル  $u_{Cl-1}(n)$  が入力された状況を表している<sup>2</sup>. この場合は新たな細胞面群は生成されず,  $k = 0$  の細胞面群

<sup>2</sup>細胞面群を “Cell-Plane Stack” と訳すため, 以降細胞面群を簡単に “CPS” と記すことがある.

(CPS#0) で学習が行なわれ、図 4.4(a) 右の状態になる。一方、図 4.4(b) は  $u_{Cl-1}(n)$  が  $k = 0$  の細胞面群の許容範囲外にある状態を図示したものであり、この場合は新たな細胞面群 (CPS#1,  $k = 1$ ) が生成される。

## 4.4 細胞面群融合法による高速生成アルゴリズム

前節までの解析結果を用いて大量のネオコグニトロン型モデルを高速に生成するアルゴリズムを提案する。まず全体の処理の流れを説明してから詳細を述べる。

### 4.4.1 処理の流れ

以降しきい値  $r_l$  ではなく実効しきい値  $\theta_l$  を用いて議論するために、ある 1 つのネオコグニトロン型モデルが持つしきい値の集合を

$$\theta = \{\theta_1, \theta_2, \dots, \theta_L\} = \left\{ \frac{r_1}{1+r_1}, \frac{r_2}{1+r_2}, \dots, \frac{r_L}{1+r_L} \right\} \quad (4.27)$$

と記す。ただし  $L$  は細胞段の数であり、本論文では  $L = 4$  としている。また  $U_{Si}$  層における可能な実効しきい値の集合を

$$\theta_l = \{\theta_l^{(1)}, \theta_l^{(2)}, \dots, \theta_l^{(|\theta_l|)}\} \quad (4.28)$$

とする。以上より生成すべきネオコグニトロン型モデルの数、すなわち実効しきい値の集合 (パラメータセット) の数は、

$$|\theta| = |\theta_1| \times |\theta_2| \times \dots \times |\theta_L| \quad (4.29)$$

で表わすことができる。また  $|\theta|$  個ある実効しきい値の 1 つを特定して指定するとき、

$$\theta^{(n)} = \{\theta_1^{(n_1)}, \theta_2^{(n_2)}, \dots, \theta_L^{(n_L)}\}, \quad (n_l = 1, 2, \dots, |\theta_l|) \quad (4.30)$$

$$\in \theta_1 \otimes \theta_2 \otimes \dots \otimes \theta_L \quad (4.31)$$

と表記する。

図 4.5 は高速生成アルゴリズムの処理の流れを図示したものである。まず図 4.5(a)-(b) の処理で通常の方法を用いてネオコグニトロン型モデルを学習する。この時、与える各  $U_S$  層のしきい値は十分に高い値にする。このしきい値の集合を、

$$\theta^{(\text{high})} = \{\theta_1^{(\text{high})}, \theta_2^{(\text{high})}, \dots, \theta_L^{(\text{high})}\} \quad (4.32)$$

とする。高速生成アルゴリズムでは実際に学習用パターンを与えて学習則を適応するのはこの  $\theta^{(\text{high})}$  を持つネオコグニトロン型モデルだけであり、そ

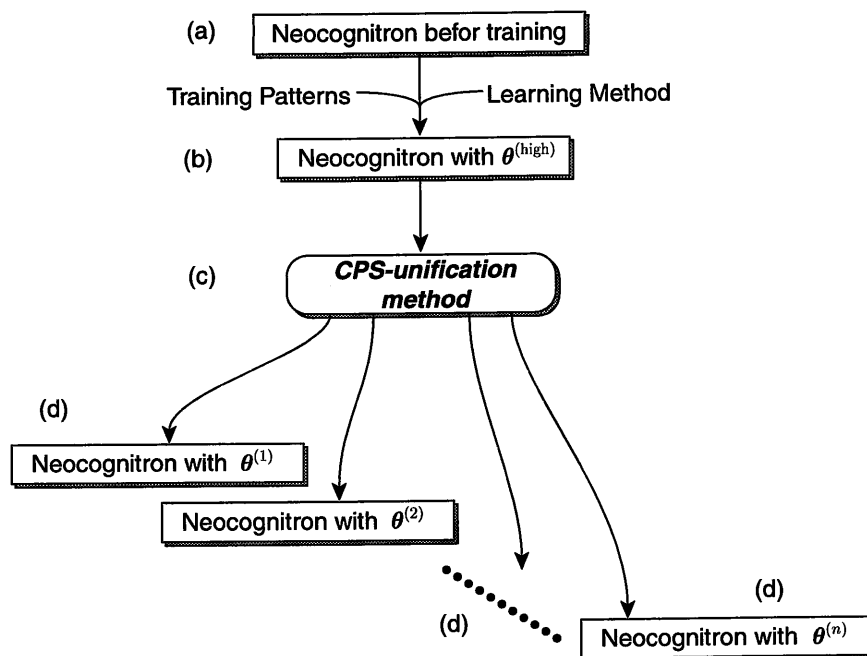


図 4.5: 高速生成アルゴリズムの処理の流れ. (a)-(b): しきい値の集合  $\theta^{(high)}$  を持つネオコグニロン型モデルを通常の方法で学習し構成する. (c)-(d): 他のしきい値集合  $\theta^{(n)}$  を持つネオコグニロン型モデルは細胞面群融合法によって高速に生成する. 図中の “CPS-unification method” は細胞面群融合法の英名である.

の他の大量のネオコグニトロン型モデルを生成する際には学習用パターンおよび学習則は適応しない。次に図 4.5(c)-(d) の段階ではしきい値集合  $\theta^{(\text{high})}$  を用いて生成したネオコグニトロン型モデルを材料として、次節で説明する細胞面群融合法を用いて  $\theta^{(\text{high})}$  より低い値のしきい値集合を持つ多数のネオコグニトロン型モデルを生成する。(c)-(d) の段階で1つのネオコグニトロン型モデルの作成に要する時間 ( $N_P \times t_T$ ) は(a)-(b) の段階でネオコグニトロン型モデルを作成する時間と比較して非常に短い ( $\epsilon$ )。なぜならば、(c)-(d) の段階では具体的な学習用パターンや学習方法を適用する必要がないからである。

この高速生成アルゴリズムで  $|\theta| = |\theta_1| \times \dots \times |\theta_L|$  種の異なるしきい値を持つ  $|\theta|$  種のネオコグニトロン型モデルを生成する時間  $t_{\text{test}}^{(\text{FGA})}$  は、

$$t_{\text{test}}^{(\text{FGA})} = |\theta| \times \epsilon + N_P \times t_T \quad (4.33)$$

$$= |\theta_1| \times \dots \times |\theta_L| \times \epsilon + N_P \times t_T \quad (4.34)$$

$$= |r_1| \times \dots \times |r_L| \times \epsilon + N_P \times t_T \quad (4.35)$$

$$\ll t_{\text{test}}^{(\text{SHOP})} \quad (4.36)$$

となるため Lovell らが指摘した問題を解決できる。

#### 4.4.2 細胞面群融合法

本手法は細胞面群を持つ回転対応型ネオコグニトロンだけではなく、細胞面を持つ従来型ネオコグニトロンにも適応できることを強調しておきたい。本節では高速生成アルゴリズムの本質である細胞面群融合法について図 4.6 を用いて説明する。図 4.6(a) は高い値を持つしきい値集合  $\theta^{(\text{high})}$  でネオコグニトロン型モデルを生成したときの  $U_{SI}$  層の細胞面群の様子を示している。ここでは2つの細胞面群 (CPS#0 と CPS#1) が生成されているものとする。この場合の  $u_{SI}$  細胞の許容範囲を角度  $\tau_i^{(\text{high})}$  で表わすと以下の式になる。

$$\tau_i^{(\text{high})} = \cos^{-1} \theta_i^{(\text{high})} = \cos^{-1} \frac{r_i^{(\text{high})}}{1 + r_i^{(\text{high})}} \quad (4.37)$$

一般のしきい値  $r_i^{(n)}$  に対しては以下の式になる。

$$\tau_i^{(n)} = \cos^{-1} \theta_i^{(n)} = \cos^{-1} \frac{r_i^{(n)}}{1 + r_i^{(n)}} \quad (4.38)$$

次に  $\theta_i^{(\text{high})}$  を持つネオコグニトロン型モデルのしきい値を  $\theta_i^{(n)}$  に低下させた ( $\theta_i^{(n)}$  に置き換えた) 様子が図 4.6(b) である。すると各細胞面群の参照ベクトルが互いの許容範囲に入り込むことがわかる。この状態は通常の学習則を用いた場合には決して起らない状態である (矛盾状態と言う)。なぜならば 4.3

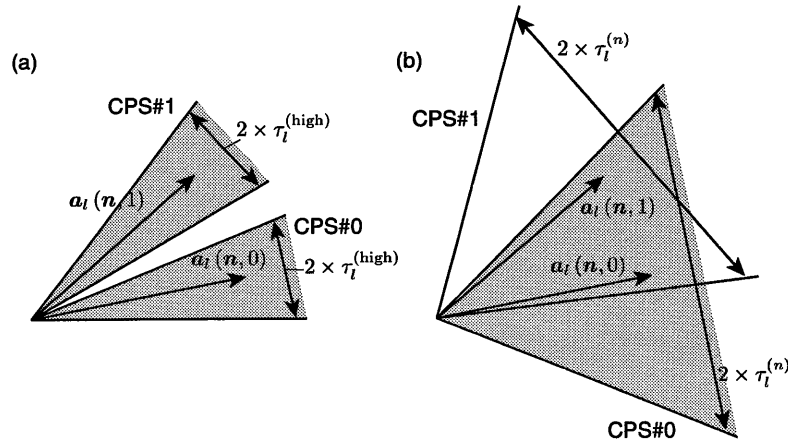


図 4.6: 細胞面群融合法. (a): しきい値の集合  $\theta^{(\text{high})}$  を用いてネオコグニトロン型モデルを学習した際,  $U_{Sl}$  層内において2つの細胞面群が生成された様子を示す. (b):  $\theta^{(\text{high})}$  を持つネオコグニトロン型モデルのしきい値集合を  $\theta^{(n)}$  に低下させると, 参照ベクトルがお互いの許容範囲に入り込む状態が発生することがある. これは通常の学習では発生しない状態 (矛盾状態) であるため, この2つの細胞面群を融合させる.

節で解析したように, パターンベクトルから作成された  $a_i(n, 1)$  が CPS#0 の許容範囲に含まれるときは, そのパターンベクトルは細胞面群 CPS#0 内の S-細胞で学習されるため新しい細胞面群は生成されないからである. すなわちしきい値  $\theta_i^{(n)}$  で学習した場合には細胞面群は1つだけが生成されるはずである.

細胞面群融合法は低いしきい値に置き換えた時に矛盾条件である2つの細胞面群を検出し, それらの細胞面群を1つに融合する方法である. 具体的には2つの細胞面群, CPS# $j$  と CPS# $k$  が

$$\cos^{-1} \frac{\langle a_i(n, j), a_i(n, k) \rangle_l}{\|a_i(n, j)\| \cdot \|a_i(n, k)\|} < \tau_i^{(n)}, \quad (4.39)$$

を満す場合にはこれらの細胞面群を1つの細胞面群, CPS# $i$  に融合する. 新しく作成された CPS# $i$  の可変結合は学習過程を考慮すると単純に以下の式で表わすことができる.

$$a_i(\nu, n, \kappa, i) = a_i(\nu, n, \kappa, j) + a_i(\nu, n, \kappa, k), \quad (4.40)$$

$$b_i(i) = b_i(j) + b_i(k). \quad (4.41)$$

この融合は後ろの層である  $U_{Sl+1}$  層の結合にも影響する.  $U_{Sl+1}$  層内の結合は以下の式で表すことができる.

$$a_{l+1}(\nu, n, i, k') = a_{l+1}(\nu, n, j, k') + a_{l+1}(\nu, n, k, k'). \quad (4.42)$$

---

**アルゴリズム 3** Fast-generative algorithm using CPS-unification method in order to make a number of neocognitrons.

---

**Require:** A trained neocognitron with a threshold set,  $\theta^{(\text{high})}$ , using any learning method, *e.g.* Algorithm 1 or Algorithm 2.

- 1:  $L$  = the number of module.
  - 2:  $|\theta| = |\theta_2| \times \cdots \times |\theta_L|$
  - 3:  $\theta^{(n)} = \{\theta_2^{(n_2)}, \dots, \theta_L^{(n_L)}\}$ , ( $n_2 = 1, \dots, |\theta_2|$   $n_L = 1, \dots, |\theta_L|$ )
  - 4: **for**  $\theta^{(n)}$ ,  $n = 1$  to  $|\theta|$  **do**
  - 5:     **for**  $l = 2$  to  $L$  **do**
  - 6:         Calculate  $\tau_l^{(n)}$  by use of (4.38).
  - 7:         **for all** combination of  $j$  and  $k$  ( $j \neq k$ ) **do**
  - 8:             **if** CPS# $j$  and CPS# $k$  satisfy the condition (4.39) **then**
  - 9:                 Make a new cell-plane stack, CPS# $i$ .
  - 10:                 Make new connections,  $a_l(\nu, n, \kappa, i)$  and  $b_l(i)$ .
  - 11:                 Apply (4.40) and (4.41).
  - 12:                 Remove CPS# $j$  and CPS# $k$ .
  - 13:                 Make a new connection,  $a_{l+1}(\nu, n, i, k')$ .
  - 14:                 Apply (4.42).
  - 15:                 Remove  $a_{l+1}(\nu, n, j, k')$  and  $a_{l+1}(\nu, n, k, k')$ .
  - 16:             **end if**
  - 17:         **end for**{loop for  $j$  and  $k$ }
  - 18:     **end for**{loop for  $l$ }
  - 19:     Output the neocognitron with  $\theta^{(n)}$ .
  - 20: **end for**
-

アルゴリズム 3 に高速生成アルゴリズムを記す。

## 4.5 評価実験

高速生成アルゴリズムの有効性を数値シミュレーションで評価する。用いた学習用パターン，評価用パターン，しきい値以外のパラメータは 3.5 節の評価実験で用いたものと同じである。

### 4.5.1 最適なしきい値集合の決定

高速生成アルゴリズムを用いて大量の回転対応型ネオコグニトロンを高速に作成し，学習用パターンを用いて認識実験を行ない，その認識性能を評価することで最適なしきい値の集合  $\theta^{(opt)}$  を決定する。可能な実効しきい値を以下のように各層において 12 種類用意した。すなわち  $1,728 (= 12^3)$  種の回転対応型ネオコグニトロンを生成する。

$$\theta_2 = \theta_3 = \theta_4 = \{0.40, 0.45, \dots, 0.95\}$$

$U_{S1}$  層の最適なしきい値， $\theta_1^{(opt)}$  は各細胞が適切な方位に反応するように設定すれば良いので簡単に決められる ( $\theta_1^{(opt)} = 0.8$  とした)。学習用パターンに対して最大の認識率を示したしきい値集合は

$$\theta^{(opt)} = \{0.8, 0.60, 0.65, 0.55\}$$

であった。3.5 節の実験結果はこのパラメータを持つ回転対応型ネオコグニトロンを対象として行なった。 $\theta^{(opt)}$  を持つ回転対応型ネオコグニトロンの評価用パターンに対する認識率は 96.4% であった。

### 4.5.2 実行時間

可能なしきい値の組み合わせは 1,728 種となり，高速生成アルゴリズムを用いて異なるしきい値を持つ 1,728 種の回転対応型ネオコグニトロンを作成する。高速生成アルゴリズムによってこの大量の回転対応型ネオコグニトロン作成するのに要した時間は約 2 時間 ( $t_{test}^{(FGA)} = 2[\text{hour}]$ ) であった。従来の方法が 1ヶ月と 26 日要していた ( $t_{test}^{(SHOP)} \sim 2[\text{month}]$ ) のに比較して非常に短時間で生成可能であることがわかる。上記の時間は Pentium 333[MHz] の CPU を用いて実験した結果である。

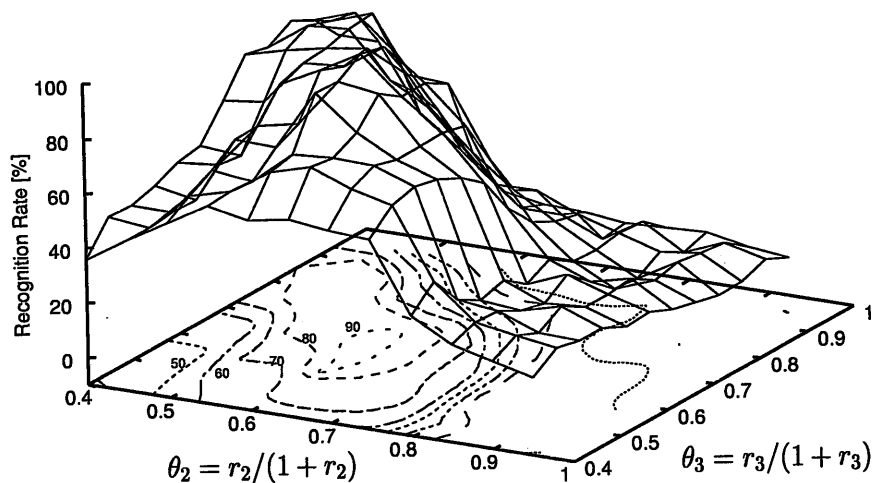


図 4.7: 高速生成アルゴリズムで生成した回転対応型ネオコグニトロン認識率.  $\theta_4^{(opt)} = 0.55$  に固定し,  $\theta_2$  と  $\theta_3$  を変化させたときの高速生成アルゴリズムによって作成された回転対応型ネオコグニトロン 144 種の認識率.  $\{\theta_2, \theta_3\} = \{0.60, 0.65\}$  で最大認識率 96.4% を示している.

### 4.5.3 従来法との比較

#### 認識率の比較

高速生成アルゴリズムで生成された回転対応型ネオコグニトロンと, 実際に学習用パターンと学習アルゴリズムを適応して作成した回転対応型ネオコグニトロン認識性能を比較した.

図 4.7 は  $\theta_4^{(opt)} = 0.55$  に固定し,  $\theta_2$  と  $\theta_3$  を変化させたときの高速生成アルゴリズムによって作成された回転対応型ネオコグニトロン 144 種の認識率である. また図 4.8 は同様に通常の学習アルゴリズムによって作成された回転対応型ネオコグニトロン 144 種の認識率である. どちらの場合も  $\{\theta_2, \theta_3\} = \{0.60, 0.65\}$  で最大の認識率を示している.  $\theta = \{0.60, 0.65, 0.55\}$  として通常の学習アルゴリズムを用いて学習した場合の認識率は 95.8% であった.

図 4.9 は図 4.7 と図 4.8 の差分を取ったものである. 認識率の差は最大で 1.8% であり, 性能にほとんど差がないことが確認できる. これは高速生成アルゴリズムが非常に短時間で, 通常の学習アルゴリズムを用いて作成された回転対応型ネオコグニトロンと同等の性能を持つ回転対応型ネオコグニトロンを生成できることを意味する.



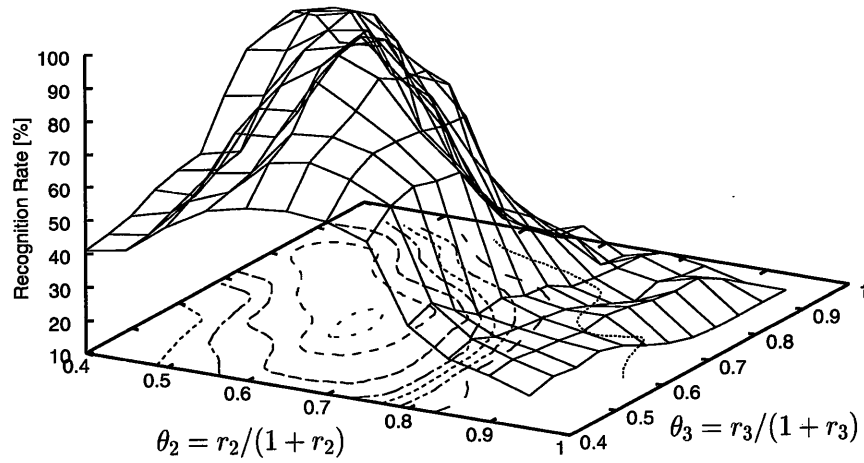


図 4.8: 通常のアルゴリズムで生成した回転対応型ネオコグニトロン認識率.  $\theta_4^{(\text{opt})} = 0.55$  に固定し,  $\theta_2$  と  $\theta_3$  を変化させたときの通常アルゴリズムによって作成された回転対応型ネオコグニトロン 144 種の認識率.  $\{\theta_2, \theta_3\} = \{0.60, 0.65\}$  で最大認識率 95.8% を示している.

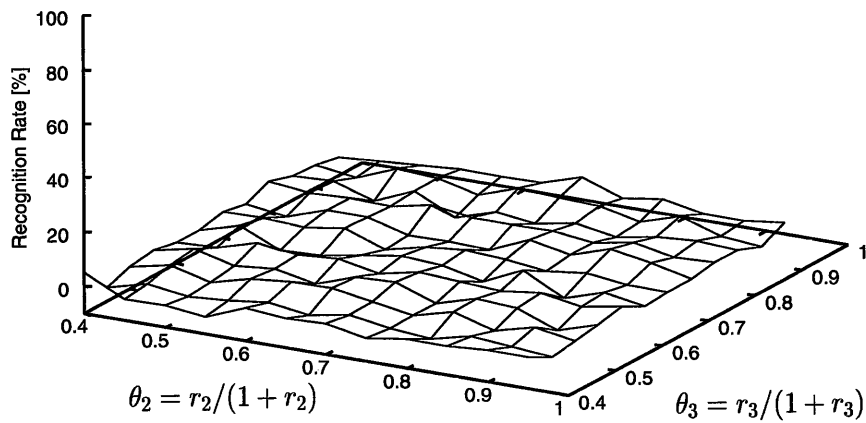


図 4.9: 高速生成アルゴリズムで生成された回転対応型ネオコグニトロンと通常アルゴリズムで生成された回転対応型ネオコグニトロンの認識率の差. 認識率にはほとんど差がないことがわかる.

## 構造の比較

高速生成アルゴリズムで生成された  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンと、実際に学習用パターンと学習アルゴリズムを適応して作成した  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンの構造の比較を行う。次の2点について調査した。

1. 各層における細胞面群数。この数が両モデルにおいて同数である場合、両モデルの細胞数が同数であることを意味する。
2. 細胞面群内の興奮性可変結合  $a_i(\nu, n, \kappa, k)$  の比較。興奮性可変結合が同じであれば、検出する局所パターンも一致する。上記の細胞面群数も同数であれば2つのモデルは完全に構造が一致することを意味する。

結果、細胞面群数においては両者とも同数であり、 $K_{S2} = K_{C2} = 6$ ,  $K_{S3} = K_{C3} = 7$ ,  $K_{S4} = K_{C4} = 9$  であった。

興奮性可変結合の値を比較するために、両モデルの各細胞面群の興奮性可変結合の類似度を評価した。高速生成アルゴリズムで生成した回転対応型ネオコグニトロンの  $U_{S_l}$  層の第  $k$  細胞面群の興奮性可変結合をベクトル  $\mathbf{a}_i^{(F)}(k)$ 、通常の学習則を用いて生成した回転対応型ネオコグニトロンの興奮性可変結合をベクトル  $\mathbf{a}_i^{(N)}(k)$  と表記し、以下の式で類似度  $s_l(k)$  を定義した。

$$s_l(k) = \frac{\langle \mathbf{a}_i^{(F)}(k), \mathbf{a}_i^{(N)}(k) \rangle}{\|\mathbf{a}_i^{(F)}(k)\| \cdot \|\mathbf{a}_i^{(N)}(k)\|},$$

ただし演算  $\langle \cdot, \cdot \rangle$  は通常の内積であり、 $\|\cdot\|$  はベクトルのユークリッドノルムである。各層における類似度を図 4.10, 図 4.11 および図 4.12 に示す<sup>3</sup>。各図の縦軸が類似度  $s_l(k)$  である。各層の各細胞面群において興奮性可変結合の類似度はほぼ1であることがわかる。以上の結果より、高速生成アルゴリズムで生成した回転対応型ネオコグニトロンは、通常の学習則を用いて生成した回転対応型ネオコグニトロンと構造がほぼ同一であることがわかる。また、構造が同一であるために認識性能にも差がないことがわかる。

## 4.6 まとめ

作成した視覚数理モデルは工学的にも有効なものである必要がある。これまで、ネオコグニトロン型モデルの適切なしきい値を決定する際には膨大な時間を必要としていた。本章ではこの問題を解決するために学習過程におけるネオコグニトロン型モデルの挙動を解析し、その結果から高速生成アルゴリズムを提案した。数値シミュレーションにより、高速生成アルゴリズムは高速に、しかも従来のアルゴリズムを用いて作成した回転対応型ネオコグニ

<sup>3</sup> $U_{S1}$  層は適切な方位選択性を持つようあらかじめ設定しておくので類似度  $s_1(k)$  は常に1である。

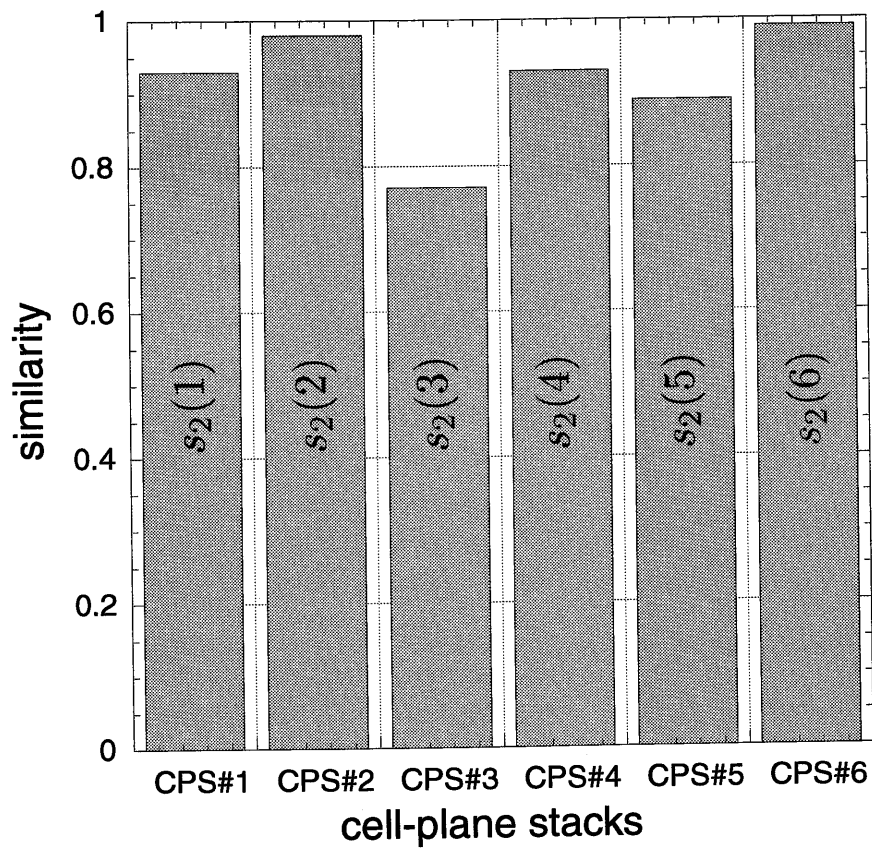


図 4.10: 高速生成アルゴリズムで生成された  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンと、実際に学習用パターンと学習アルゴリズムを適応して作成した  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンの  $U_{S_2}$  層における興奮性可変結合の類似度,  $s_2(k)$ ,  $k = 1, \dots, 6$ . 作成された細胞面群は 6 種類 ( $K_{S_2} = 6$ ) であった. 類似度がほぼ 1 であり, 高速生成アルゴリズムで生成した回転対応型ネオコグニトロンの  $U_{S_2}$  層は通常の学習則を用いて生成したものとほぼ同じ構造であることがわかる.

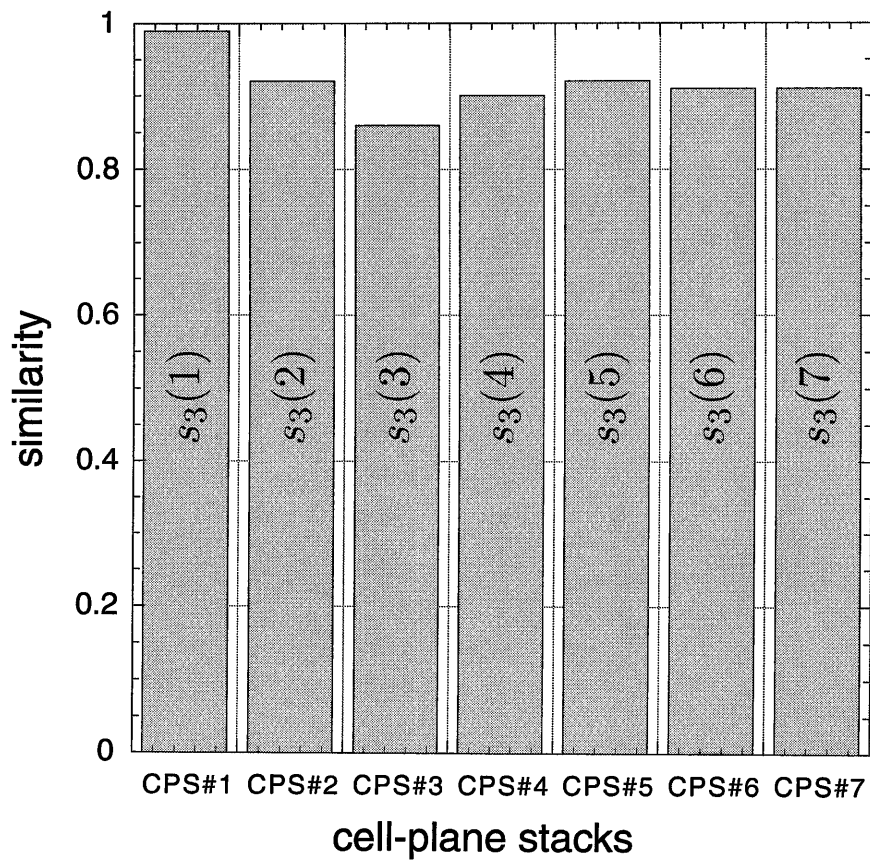


図 4.11: 高速生成アルゴリズムで生成された  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンと、実際に学習用パターンと学習アルゴリズムを適応して作成した  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンの  $U_{S_3}$  層における興奮性可変結合の類似度,  $s_3(k)$ ,  $k = 1, \dots, 7$ . 作成された細胞面群は 7 種類 ( $K_{S_3} = 7$ ) であった. 類似度がほぼ 1 であり, 高速生成アルゴリズムで生成した回転対応型ネオコグニトロンの  $U_{S_3}$  層は通常の学習則を用いて生成したものとほぼ同じ構造であることがわかる.

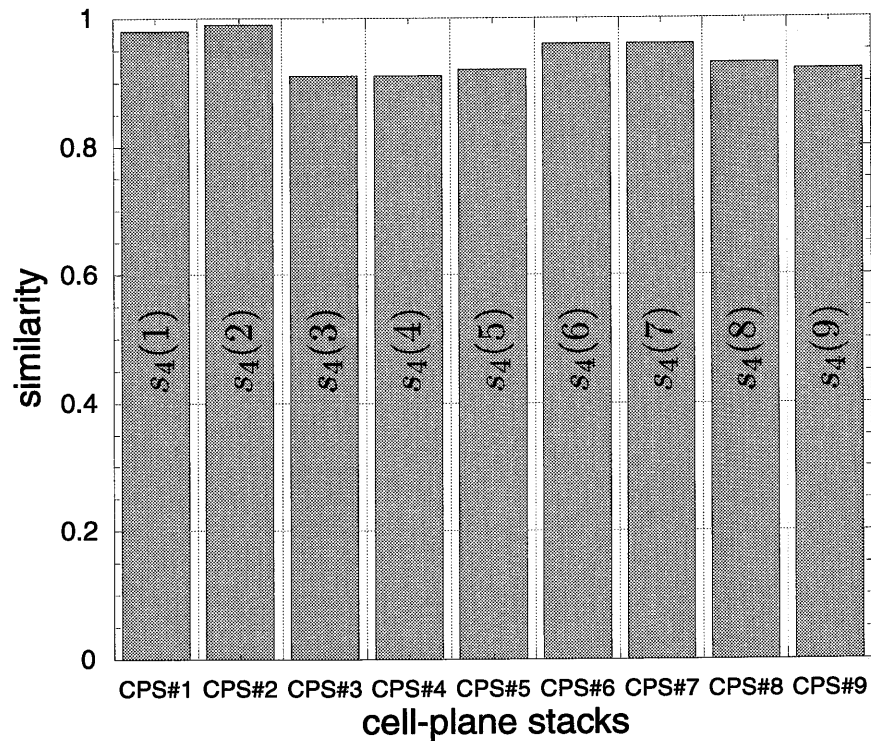


図 4.12: 高速生成アルゴリズムで生成された  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンと、実際に学習用パターンと学習アルゴリズムを適応して作成した  $\theta^{(\text{opt})}$  を持つ回転対応型ネオコグニトロンの  $U_{S_4}$  層における興奮性可変結合の類似度,  $s_4(k)$ ,  $k = 1, \dots, 9$ . 作成された細胞面群は 9 種類 ( $K_{S_4} = 9$ ) であった. 類似度がほぼ 1 であり, 高速生成アルゴリズムで生成した回転対応型ネオコグニトロンの  $U_{S_4}$  層は通常の学習則を用いて生成したものとほぼ同じ構造であることがわかる.

トロンと同等の性能および構造をもつ回転対応型ネオグニトロンを生成することが確認された。

## 第5章 連想想起を伴う回転パターンの認識

第3章では「回転に対して単純」なパターンを認識するための数理モデルとして回転対応型ネオコグニトロンを提案した。本章では「回転に対して単純ではない」パターンを認識するため数理モデルを提案する。

### 5.1 「回転に対して単純なパターン」の同定

3.1節において、回転したパターンの認識機構の解明の一つとして Cooper と Metzler の実験に対する考察を行ない、

- 文字の構成要素である角や曲線，例えば L-型や T-型といった複数の線分から構成される単純な視覚パターンは，それを回転したパターンそれ自体を検出する細胞が存在し，
- このような視覚パターン（回転に対して単純なパターン）は回転していても瞬時に認識することができ，
- 認識の際にトップダウン的な処理機構を必要としない，

という結論を得た。第3章ではこの考察を基に回転対応型ネオコグニトロンを提案した。このモデルは「回転に対して単純なパターン」の認識を行なうモデルであり，回転したパターンを直接（ボトムアップ型処理で）認識することができる。すなわち，回転したパターンを検出する細胞が存在する。実際に数字を用いた数値シミュレーションの結果，回転に対して完全に頑健であることが示された。

では「回転に対して単純ではない」パターンを認識するための数理モデルにはどのような構成が必要となるであろうか。この疑問に関連する実験結果として Cooper と Shepard の回転した英数字に対する心理実験 [9] がある。図 5.1 に Cooper らの実験結果の一部を挙げる。彼等は回転した文字，および回転し鏡像反転させた文字を異なる条件下で被験者に提示しそれが正常か鏡像かを判断させ，判断に要する時間を測定した。使用した文字は 'R'，'J'，'G'，'2'，'5'，'7' である。条件は次の3種類である。

C-条件 提示する文字のクラスの情報と向きの情報を与え，その3秒後に実際の文字を提示する。

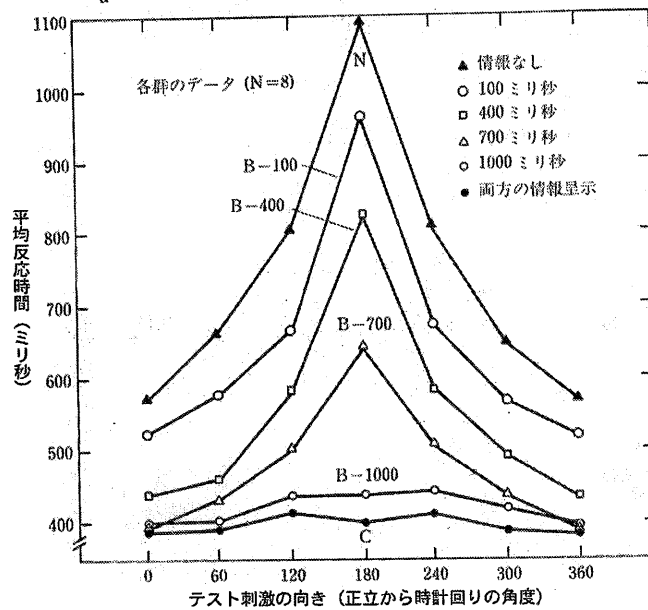
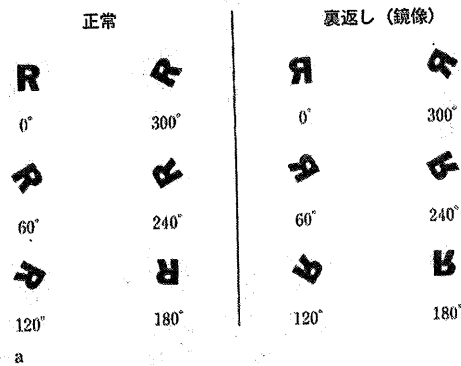






図 5.1: Cooper と Shepard による回転した文字の認識の研究結果の一部。彼等は回転した文字, および鏡像反転し回転させた文字を異なる条件下で被験者に提示しそれが正常か鏡像かを判断させ, 判断に要する時間を測定した。使用した文字は 'R', 'J', 'G', '2', '5', '7' である。条件は次の 3 種類である。C-条件: 提示する文字のクラスの情報と向きの情報を与え, その 3 秒後に実際の文字を提示する。B-条件: C-条件とほぼ同じであるが文字を提示するまでの時間が短い (100, 400, 700, 1000[ms])。N-条件: 情報を与えない。文献 [6] より引用。



表 5.1: パターンの複雑性と心的回転依存性, およびネオコグニトロン型モデルとの対応.

細胞段	$U_1$	$U_2$	$U_3$	$U_4$
検出されるパターン				
回転に対する依存性	単純		非単純	

B-条件 C-条件とほぼ同じであるが文字を提示するまで時間が短い (100, 400, 700, 1000[ms]).

N-条件 文字クラスの情報, 向きの情報のいずれも与えない.

N-条件の結果から文字回転角が大きい程判断に時間を要していることがわかる. これは回転した文字を瞬時に認識できないことを意味している. C-条件では被験者は提示される文字のクラスおよびその回転角がわかっており, 更に実際の文字提示までに3秒間のブランクがあるために(脳内の)イメージ内で回転した非鏡像文字を作り上げることができる. よってC-条件における判断時間は回転角に依存しない. これらの結果より英数字は回転に対して単純なパターンではないことがわかる<sup>1</sup>. また回転した英数字を認識するためにはイメージ内でパターンを作り上げるといったトップダウン的な機能が必要であることもわかる.

ではどのようなパターンがヒトの脳において回転に対して単純なのであろうか. 本研究ではどの程度単純なパターンであれば心的回転のようなトップダウン的な処理を用いずに正しく認識できるかを心理実験により調査した. 特にパターンをネオコグニトロン型モデルの各層で検出するパターンに限定し, どの程度の複雑度であれば回転に対して単純であるかを確認する. また, ネオコグニトロン型モデルの高次層で検出される局所パターンを複雑度が高く, 低次層で検出される局所パターンを複雑度が低いと定義した. 実験条件は以下の通りである.

1. 線分, 線分から構成されるパターンといった各種パターンを, ネオコグニトロン型モデルでの各細胞段  $U_i$  で検出されるパターンに応じて分類する. 例を表 5.1 に示す (各段に4つのパターンを例示してある).
2. 表 5.1 に示すようなパターン, その鏡像パターンおよび鏡像反転し回転したパターンを用意する.

<sup>1</sup>本節の考察から数字は「回転に対して単純ではない」ことが明らかになったが, これは第3章で提案した回転対応型ネオコグニトロンを否定しているのではない. 回転対応型ネオコグニトロンはあくまでも「回転に対して単純」なパターンを認識する数理モデルであり, モデルの評価をするためのパターンとして数字を使用しただけである.

3. 各被験者にはパターン提示後にできるだけすぐに（時間をおかずに）パターンのクラスを報告するように要請する。
4. 被験者に用意したパターンをランダムに提示する。

実験の様子を図 5.2 に示す。被験者の人数は 6 歳から 24 歳の 14 人であったが被験者全員から次の結果が得られた。

- (a)  $U_1$  および  $U_2$  に属する複雑度を持つパターンは全て正しく認識される。
- (b)  $U_3$  および  $U_4$  に属する複雑度を持つパターンで大きく回転しているパターンに対しては、それが鏡像パターンであっても非鏡像パターンのクラスとして誤認識する。例えば図 5.2 のパターンを “2” と報告する。
- (c) 報告までの時間に制限を与えない場合は正しく認識する（鏡像であることを後から気付く）。

上記の結果 (a) より、被験者は  $U_1$  および  $U_2$  に属する複雑度を持つパターンは心的回転の操作をすることなく正しく認識可能であり、回転したパターンを既に学習および記憶している可能性がある。すなわち  $U_1$  および  $U_2$  に属するパターンはヒトの脳では回転に対して単純であることがわかる（表 5.1 参照）。一方結果 (b) より、 $U_3$  および  $U_4$  に属する複雑度を持つパターンは心的回転の操作なしでは正しく認識できないことがわかる。このことは、 $U_3$  および  $U_4$  に属するパターンを回転させて得られたパターンを学習および記憶していない可能性を示唆しており、 $U_3$  および  $U_4$  に属するパターンはヒトの脳においては回転に対して単純ではないと推定される（表 5.1 参照）。

この結果から、ヒトの脳における回転に対して単純ではないパターンを認識するための視覚モデルのボトムアップ型回路を、ネオコグニトロンに基礎を置いて構築するのであれば次のような構成となる。

- $U_1$  と  $U_2$  細胞段を回転したパターンの認識が可能な回転対応型ネオコグニトロンで構成し、
- $U_3$  と  $U_4$  細胞段を従来型ネオコグニトロンで構成する。

この視覚モデルは後半を従来型ネオコグニトロンで構成しているので、回転したパターンを直接認識することはできない。

## 5.2 回転したパターンを認識するための回路構成

では実際に回転に対して単純ではないパターンを認識するにはどのような回路が必要であろうか。3.1 節での考察と図 5.1 に示した Shepard らの実験結果より、正立したパターンと提示された回転パターンの回転方向をなんらかの「手掛り」を用いて推定しているはずである。また回転方向を推定する

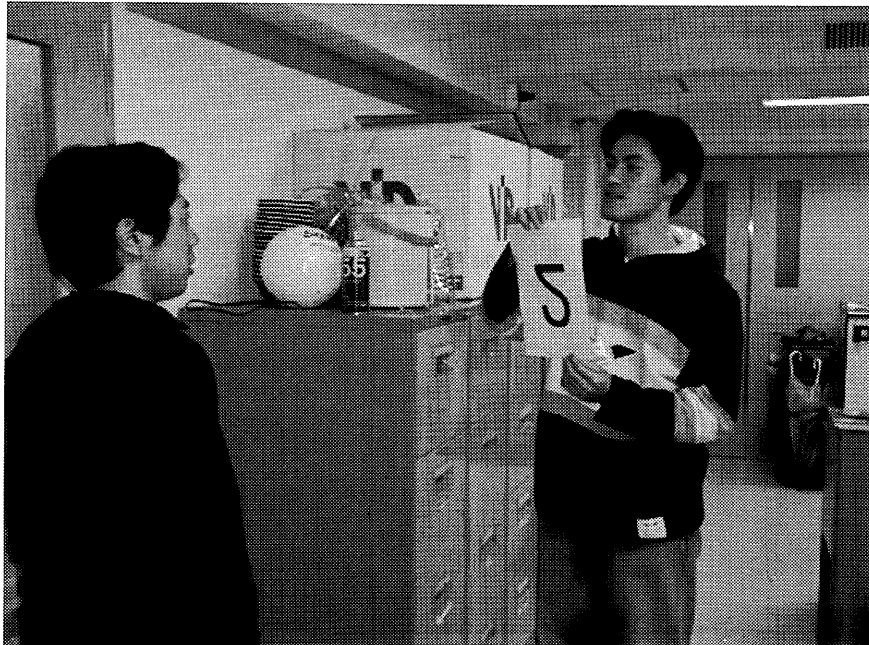


図 5.2: 回転に対して単純なパターンを同定するための心理実験の様子. ここでは被験者に鏡像反転し回転させた数字の '2' を提示している. この被験者 H.U. はこのパターンを '2' と報告した. 回転している '2' を既に学習し記憶しているのであればこのパターンを '2' と答えるはずがない.

には提示された回転パターンの仮の認識結果が必要となる<sup>2</sup>。なぜならば回転角の推定には、仮の認識結果に対する正立したパターン（基準となる回転していないパターン）と提示された回転パターンとの比較が必要不可欠だからである。もし「仮の認識結果」が得られなければ、正立パターンという基準が存在しないために回転角の推定はできず、図 3.1 や図 5.1 に示すような結果にはならないはずである。

仮の認識結果に対応する正立パターンを得るためにはトップダウン的な処理である連想想起が必要である。本研究では正立パターンの連想想起を行なうトップダウン型回路として Fukushima によって提案された神経回路網モデルに着目する [40]。

また、回転角推定のために正立パターンと回転パターンを比較しなければならないが、比較するパターン、すなわち「手掛り」をネオコグニトロン型モデルの  $U_2$  層で検出される局所パターンであるとする。これは「手掛り」が心的回転に依存しないパターンでなければならないという理由からである。

まとめると、(ネオコグニトロンを基礎と置いた場合に) 回転した単純ではないパターンを認識する視覚数理モデルを構築するには以下の条件を満たさなければならない。

- ボトムアップ的な処理だけではなくトップダウン的な処理が必要。
- ボトムアップ処理により、回転したパターンに対する仮の認識結果（仮定）を生成。
- トップダウン処理により、仮の認識結果に対応するクラスの正立パターンを連想。
- ボトムアップ処理による認識回路は  $U_1$  細胞段と  $U_2$  細胞段を回転対応型ネオコグニトロン、 $U_3$  細胞段と  $U_4$  細胞段を従来型ネオコグニトロンで構築する。
- 回転方向の推定は、 $U_2$  細胞段で検出される「回転に対して単純」なパターンを用いて行なう。

以上の考察より、回転したパターンを認識する新しい視覚数理モデルを提案する。図 5.3 にその構造を示す。図 5.3 の上段の神経回路網がボトムアップ型の認識モデルであり、下段が「仮の認識結果」に対応する正立パターンを連想するトップダウン型の連想想起モデルである。連想想起モデルは認識モデルと同様の構造をしているが、信号の流れる方向が遠心性（図では右から左）である。以降、上段の神経回路網を認識モデル、下段の神経回路網を連想想起モデルと呼ぶ。次節より各モデルの動作を詳しく説明する。

<sup>2</sup>ここで言う「仮の認識結果」とはその結果が必ずしも正しい結果でなくともよいという意味である。

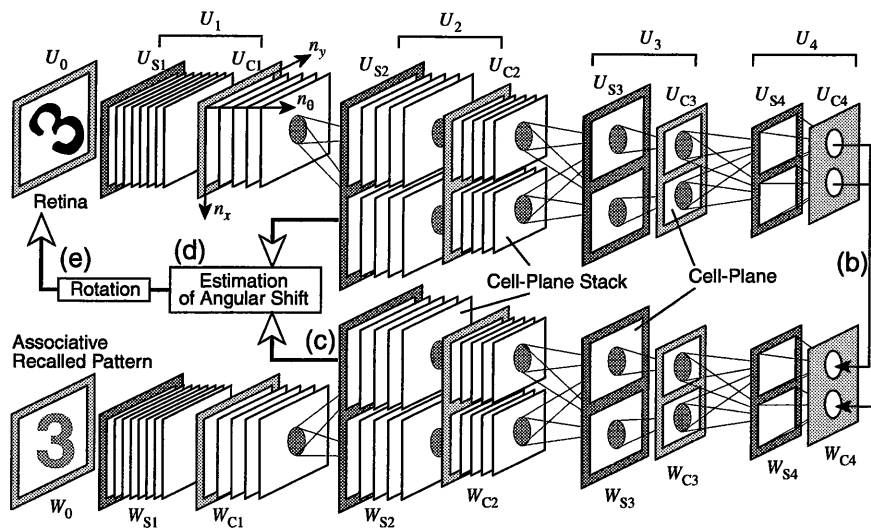


図 5.3: 認識モデルと連想想起モデルの構造. 上段の神経回路網がボトムアップ型の認識モデルであり, 下段が「仮の認識結果」に対応する正立パターンを連想するトップダウン型の連想想起モデルである. 連想想起モデルは認識モデルと同様の構造をしているが, 信号の流れる方向が遠心性 (図では右から左) である. 上段の神経回路網を認識モデル, 下段の神経回路網を連想想起モデルと呼ぶ. (a)-(e) の記号は図 5.4 に示すモデルの動作の説明との対応のために記してある.

## 5.3 認識モデルと連想想起モデル

### 5.3.1 提案モデルの処理過程

まず全体の処理過程を説明する。図 5.3 の数理モデルの動作を説明するために図 5.4 を用いて説明する。図 5.3 と図 5.4 に記述してある (a) から (e) の記号はそれぞれ対応している。

- (a) 回転したパターン（図では回転した '3' を例にしている）が認識モデルに入力されると認識モデルはそのパターンを直接認識できないため、
- (b) 回転したパターンに対して仮の認識結果を生成する。
- (c) 仮の認識結果は連想想起モデルに伝えられ、その結果に対応するクラスの正立したパターンを連想する。
- (d) 連想したパターンと、回転したパターンの回転角を「回転に対して単なるパターン」を比較して推定する。
- (e) 実際に入力された回転パターンを推定した回転角に応じて回転する。

(e) の動作が完了すると、回転を補正されたパターンは再び認識モデルによって認識される。このときの認識結果と「仮の認識結果」が異なる場合は仮定が否定されたことになるため、新しい「仮の認識結果」を生成して (c) 以降の動作が再度行なわれる。一致する場合には仮定が立証されたことになるために処理は停止する。全ての仮の認識結果が否定された段階で処理は停止し、認識結果は未知のパターンであるとして認識を終了する（一般に「リジェクトされる」と言う）。また (a) の処理において認識細胞が通常の認識過程で発火する場合には (b) 以降の処理は行なわれない。これは正立したパターンが入力された場合に起こる。

この「仮の認識結果」とその「検証」の繰り返しによる手法は文献 [41] にある、the technique of iterative hypotheses verification の一種とも捉えられる。

### 5.3.2 認識モデル

回転したパターンの回転角を推定するにはボトムアップ型の認識モデルのみを用いて「仮の認識結果」を得る必要がある。しかし認識モデルの後層は従来型ネオコグニトロンで構成されているために、学習時に使用されたパラメータ値のままでは大きく回転したパターンに対する認識結果を得ることができない。すなわち  $U_{C4}$  層に存在する全ての認識細胞が発火しない。5.1 節で行なった心理実験によると心的回転の実行を抑制した条件下では、あるパターンを大きく回転させたパターンとその鏡像パターンを同一カテゴリーとして誤認識することが明らかになった。これはパターンを構成する局所的パターン、すなわち  $U_2$  や  $U_3$  細胞段で検出されるパターンの絶対位置  $n$  を無

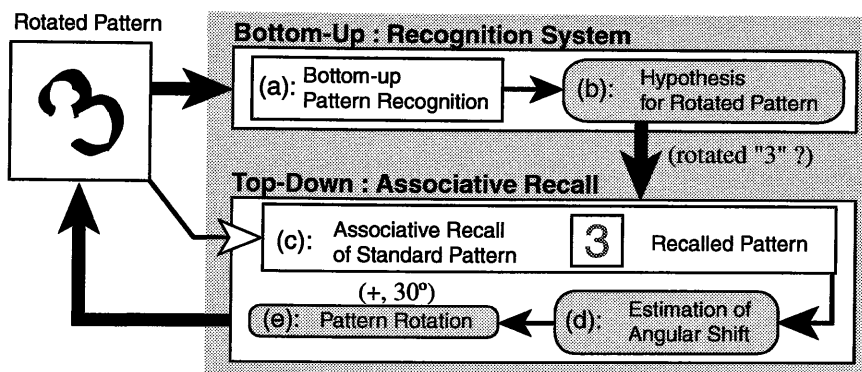


図 5.4: 認識モデルと連想想起モデルの処理の流れ. (a): 回転したパターン (図では回転した '3' を例にしている) が認識モデルに入力されると認識モデルはそのパターンを直接認識できないため, (b): 回転したパターンに対して仮の認識結果を生成する. 仮の認識結果は連想想起モデルに伝えられ, (c): その結果に対応するクラスの正立したパターンを連想する. (d): 連想したパターンと, 回転したパターンの回転角を「回転に対して単純なパターン」を比較して推定する. (e): 実際に入力された回転パターンを推定した回転角に応じて回転する. (e) の動作が完了すると, 回転を補正されたパターンは再び認識モデルによって認識される. このときの認識結果と「仮の認識結果」が異なる場合は仮定が否定されたことになるため, 新しい「仮の認識結果」を生成して (c) 以降の動作が再度行なわれる. 一致する場合には仮定が立証されたことになるために処理は停止する. 全ての仮の認識結果が否定された段階で処理は停止し, 認識結果は未知のパターンであるとして認識を終了する.

視した認識を行なっていると考えられる。

認識モデルの  $U_2$  細胞段は、細胞面群で構成されており、各細胞は3次元的に配置されている。各細胞の位置は  $(n_x, n_y, n_\theta)$  で表現される。局所的パターンの絶対位置はその局所的パターンを検出し発火している細胞の  $xy$  平面における座標  $(n_x, n_y)$  で表現され、局所的パターンの回転角は  $n_\theta$  で表現される。すなわち局所的パターンの絶対位置を無視した認識を行なうには、 $(n_x, n_y)$  で表現されている局所パターンの絶対位置の情報を無視した局所パターンの検出を行なえばよいことになる。これは回転対応型ネオコグニトロンや従来型ネオコグニトロンにおいて、ぼかしの範囲を定めるパラメータ  $D_2$ ,  $D_3$  および  $D_4$  を拡大することに等しい。また絶対位置情報の無視は、特徴選択性を定めるパラメータである  $r_3$ ,  $r_4$  の値を低下させることでも実現できると考えられる。認識モデルでは以上2種類の処理を取り入れている。

### 5.3.3 連想想起モデル

仮の認識結果が得られた後、連想想起モデルによって正立パターンが想起される。本研究では Fukushima によって提案されたネオコグニトロン型の連想想起モデル [40] に基づいたモデルを構築した (図 5.3 下部)。ただし Fukushima の連想想起モデルの機能の一つである選択的注意視は本モデルに組み込まれていない。

連想想起モデルの各層の表記法は図 5.3 に示す通りである。連想想起モデルの細胞の配置および細胞数は認識モデルと同じであるが、信号の流れる方向が遠心性になっている。結果的に連想層の  $W_0$  層では、ぼけた正立パターンが連想されることになる。ただし  $W_0$ ,  $W_{S1}$ ,  $W_{C1}$  層は実質的に必要でないため後で行なう数値シミュレーションでは組み込まれていない。

回転角の推定は  $U_{S2}$  層と  $W_{S2}$  層における S-細胞の発火パターンを  $n_\theta$ -方向で比較することで行なわれる。すなわち正立パターンと回転パターンの角度のずれを、発火パターンの角度のずれに対応させている。

### 5.3.4 提案モデルの数式表現

ここでは各細胞とモデルの動作の数式表現を与える。

#### 認識モデル

認識モデルは回転対応型ネオコグニトロンと従来型ネオコグニトロンを結合させたものであるので第2章および第3章で与えたものと同じである。学習側についても同様である。細胞数に関するパラメータの値を表 5.2 に示す。



表 5.2: 細胞面群数, 細胞面数および細胞数. 括弧内の数字は学習終了後の数  
を表す. \*は定義されていないことを示す.

	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$K_{Sl}$	*	1	(6)	(22)	(10)
$N_{Sl}$	*	59	17	13	3
$T_{Sl}$	*	16	8	4	2
$K_{Cl}$	1	1	(6)	(22)	(10)
$N_{Cl}$	61	17	17	10	1
$T_{Cl}$	1	8	4	2	1

表 5.3: 結合野  $A_l$  と  $D_l$  の範囲.  $D_l$  の上段は通常の結合野の範囲であり, 下  
段は仮の認識結果を得る場合の拡大された結合野の範囲を示す.

	$l=1$	$l=2$	$l=3$	$l=4$
$A_l$	$3 \times 3 \times 1$	$3 \times 3 \times 8$	$5 \times 5 \times 1$	$3 \times 3 \times 1$
$D_l$	$5 \times 5 \times 3$	$3 \times 3 \times 3$	$5 \times 5 \times 1$	$3 \times 3 \times 1$
	$5 \times 5 \times 3$	$5 \times 5 \times 8$	$7 \times 7 \times 1$	$5 \times 5 \times 1$

### 仮の認識結果の生成

認識段階において大きく回転したパターンが  $U_0$  層に提示されると, 全ての  
認識細胞  $u_{C4}(n, k)$  が発火しなくなる. 5.3.2 節で議論したように認識モデル  
は全ての認識細胞が発火しない状態を検出すると, 以下に示す 2 種類のパラ  
メータ調整を行なうことで少なくとも 1 つの認識細胞を強制的に発火させる;

1. 特徴選択性パラメータ  $r_3$  および  $r_4$  を次式に従って低下させる.

$$r_l := \gamma \cdot r_l, \quad (l = 3, 4). \quad (5.1)$$

ただし  $:=$  は PASCAL の代入演算子の意味であり,  $\gamma$  は  $0 < \gamma < 1$  の減  
衰定数である.

2.  $U_{C2}$ ,  $U_{C3}$  および  $U_{C4}$  層のぼかしの範囲  $D_2$ ,  $D_3$  および  $D_4$  を拡大す  
る. 拡大後の範囲を表 5.3 中  $D_l$  の下段に示す.

それでも発火しない (仮の認識結果が得られない) 場合は, 上記 (1) を繰り返  
し適応する.

### 連想想起モデル

連想想起モデルは  $W_0, W_{S_l}, W_{C_l}$  ( $l = 1, 2, 3, 4$ ) の9層から構成されている。認識モデルの  $U_{C_4}$  層が出力する仮の認識結果は連想想起モデルの  $W_{C_4}$  層に戻され、遠心性経路を辿って連想層である  $W_0$  層に到達する。

遠心性経路における  $W_{S_l}$  層内の  $w_{S_l}$  細胞の発火頻度は次式で与えられる。

$$w_{S_l}(\mathbf{n}, k) = \beta_l \cdot \sum_{\nu \in D_l} d_l(\nu) \cdot w_{C_l}(\mathbf{n} \ominus_{T_{C_l}} \nu, k), \quad (5.2)$$

ただし  $\beta_l$  は正の定数である。

$W_{C_l}$  層内の  $w_{C_l}$  細胞の発火頻度と、 $W_{S_l}$  層内の  $w_{V_l}$  細胞の発火頻度は次式で与えられる。

$$w_{C_l}(\mathbf{n}, k) = \psi[\alpha_l \cdot (e' - i')], \quad (5.3)$$

$$e' = \sum_{\kappa=1}^{K_{S_{l+1}}} \sum_{\nu \in A_{l+1}} a_{l+1}(\nu, \mathbf{n}, k, \kappa) \cdot w_{S_{l+1}}(\mathbf{n} \ominus_{T_{S_{l+1}}} \nu, \kappa), \quad (5.4)$$

$$i' = \sum_{\nu \in A_{l+1}} c_{l+1}(\nu) \cdot w_{V_{l+1}}(\mathbf{n} \ominus_{T_{S_{l+1}}} \nu), \quad (5.5)$$

$$w_{V_{l+1}}(\mathbf{n}) = \frac{r_{l+1}}{1 + r_{l+1}} \sum_{\kappa=1}^{K_{S_{l+1}}} b_{l+1}(\kappa) \cdot u_{S_{l+1}}(\mathbf{n}, \kappa), \quad (5.6)$$

ここで  $\alpha_l$  は正の定数である。











連想想起モデルの学習は認識モデルの学習後に続いて行なわれる。すなわち認識モデルで新しい細胞面群や細胞面が生成されると、連想想起モデルにおいても同じ細胞面群や細胞面が生成される。連想想起モデルの遠心性可変結合は次のように自動的に強化される: 求心性可変結合の強化の終了後、 $w_S$  細胞から降下していく遠心性結合は対応する  $u_S$  細胞から上昇していく求心性結合と等しくなるように自動的に強化される。

### 回転角推定

入力された回転パターンの回転角を推定行なうには回転パターンと連想想起された正立パターンの回転角の差を推定する必要がある。局所的パターンの絶対位置は発火細胞の位置  $(n_x, n_y)$  で表現されているため、角度の推定は  $U_{S_2}$  と  $W_{S_2}$  層の細胞の発火パターンをそれぞれ  $n_\theta$  方向に射影した発火パターン同士を比較する。推定角  $\Theta$  は次式で与えられる。

$$\Theta = \underset{\substack{\Delta\theta=0,\pm 1,\dots, \\ \pm T_{S_2}-1}}{\operatorname{argmax}} \left\{ u_{S_2}(n_\theta, k) \cdot w_{S_2}(n_\theta \oplus_{T_{S_2}} \Delta\theta, k) \right\}, \quad (5.7)$$

図 5.5: 正しく認識されたパターンの例.  $\Theta^{(\text{org})}$  は正立パターンからの実際の回転角.  $\Theta^{(\text{est})}$  は角度推定部により推定された回転角.

評価用 パターン					
$\Theta^{(\text{org})}$	+90°	+90°	+70°	-80°	-50°
$\Theta^{(\text{est})}$	+90°	+90°	+90°	-90°	-45°
評価用 パターン					
$\Theta^{(\text{org})}$	+90°	-30°	-45°	-50°	-30°
$\Theta^{(\text{est})}$	+90°	-45°	-45°	-45°	-45°

$$u_{S2}(n_\theta, k) = \sum_{n_x, n_y \in N_{S2}} u_{S2}((n_x, n_y, n_\theta), k), \quad (5.8)$$

$$w_{S2}(n_\theta, k) = \sum_{n_x, n_y \in N_{S2}} w_{S2}((n_x, n_y, n_\theta), k). \quad (5.9)$$

## 5.4 評価実験

提案した数理モデルの認識性能の評価を行なう. 使用した学習用パターンおよび評価用パターンは 3.5 節で使用したものと同一である. また適切なしきい値を決定するために第 4 章で提案した高速生成アルゴリズムを用いた. 得られたしきい値は次の通りである.

$$\theta^{(\text{opt})} = \{0.8, 0.60, 0.55, 0.45\}$$

### 5.4.1 回転したパターンに対する認識性能

学習終了後に正しく認識したパターンの一例を示す. 図 5.5 中の角度  $\Theta^{(\text{org})}$  はデータベース中の画像を回転した際の実際の回転角を示し,  $\Theta^{(\text{est})}$  は提案モデルが実際に入力イメージを回転した角度を示す.  $\Theta^{(\text{org})}$  と  $\Theta^{(\text{est})}$  の値の差は, 提案モデルの角度推定の解像度に起因する. これは角度推定部が  $U_{S2}$  層と  $W_{S2}$  層の発火パターンの角度ずれを検出して回転角を推定しており, 全方位が  $T_{S2} = 8$  方位に量子化されているためである. しかし手書き数字は執筆者によって形状が異なり, 実際には正確な回転角を知ることはほとんど不可能である<sup>3</sup>. すなわちこの角度の誤差は形式上のもので, 認識性

<sup>3</sup>筆跡鑑定をすればある程度知ることができる.

能としての誤りを示すものではない。例えば 5.5 のパターン “9” の回転角が  $-30^\circ$  か  $-45^\circ$  かを判断することは不可能である。

評価用パターン中で計算機上で回転させていない 1,000 パターンに対する認識率は 97.3% で、回転した 15,000 パターンに対する認識率は 96.39% であった。提案モデルは回転パターンを学習していないにもかかわらず回転パターンに対する認識率も高いことは、提案モデルの有効性を示している。

#### 5.4.2 認識モデルのみによる回転したパターンの認識性能

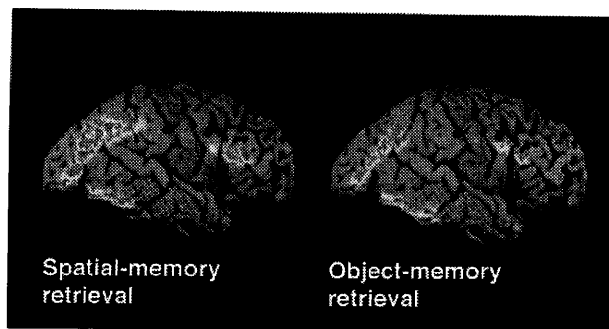
認識モデルのみを使用した場合の回転していないパターンに対する認識率は 97.4% で、回転したパターンに対する認識率は 14.67% であった。この結果から、ボトムアップ型の認識モデルが回転しているパターンを学習および記憶していない場合、連想想起モデルと角度推定による入力パターンの回転操作が有効に作用していることがわかる。

#### 5.4.3 鏡像反転したパターンに対するモデルの動作

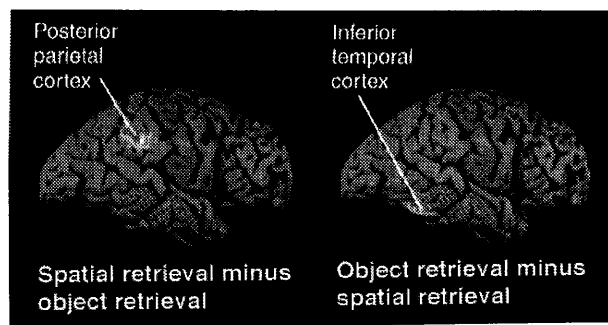
鏡像回転パターンを計算機上で 1,500 パターン作成し、 $U_0$  層に提示した。認識モデルの認識細胞による仮の認識結果は、非鏡像パターンに対するクラスを示したが、入力パターンの回転後に認識細胞の発火頻度が再び抑えられ、最終的に鏡像回転パターンは未知パターンとして排除（リジェクト）された。これは人を対象とした心理実験の結果と酷似しており、非常に興味深い結果である。また、鏡像パターンおよび鏡像回転パターンを直接学習しないにもかかわらず、それに対応するクラスを仮の認識結果として出力したことは本モデルの構成の際に直接意図していなかったことであり、本モデルの妥当性を示す根拠の一つである。

### 5.5 認識モデルと連想想起モデルの融合

前節までに、認識モデルと連想想起モデルによる回転パターン認識のための数理モデルを提案し、その構造を図 5.3 に示した。図からもわかるように認識モデル（図 5.3 上段）と連想想起モデル（図 5.3 下段）は構造的に独立である。本節ではまず認識と連想に関する生理学的知見を調査する。その結果これまで構造的に独立としてきた両モデルが融合可能であることを示唆し、融合した視覚モデルが細胞数の観点から工学的にも有効であることを示す。



(a)



(b)

図 5.6: 物体の認識および連想時における大脳皮質の反応の様子（文献 [42] より抜粋）。赤い部分に存在する細胞が他の部位と比較して高い反応を示している。(a)：右はある物体を連想している場合、左はその物体の位置を連想している場合の反応。下側頭葉（inferior temporal cortex）に存在する細胞は物体を「認識」する細胞として知られており（(b) 右）、後頭葉（posterior parietal cortex）は物体の位置を検出する細胞として知られている（(b) 左）。

### 5.5.1 認識モデルと連想想起モデルの融合

#### 認識と連想に関する生理学的知見

図 5.6 は物体の認識時と、その物体の連想時の大脳皮質の反応を PET を用いて調査したものである。赤い部分に存在する細胞が他の青い部分と比較して高い反応を示している。図 5.6(a) の右の図は、ある物体の形状を連想している場合、左はその物体の位置を連想している場合の反応である。下側頭葉（inferior temporal cortex）に存在する細胞は物体を「認識」する細胞として知られており（(b) 右）、後頭葉（posterior parietal cortex）は物体の位置を検出する細胞として知られている（(b) 左）。この実験結果から認識時および連想時共にほぼ同じ部位の細胞が反応していることが洞察される。すなわ

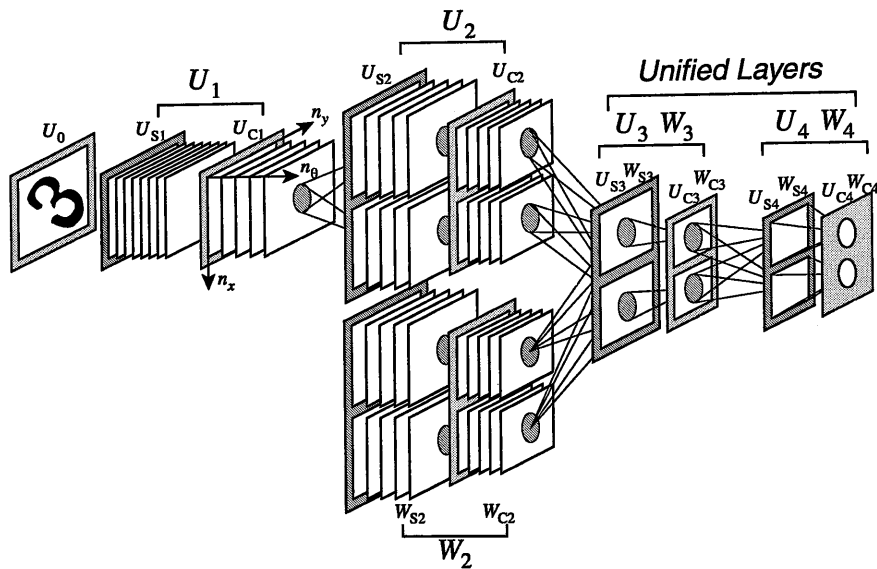


図 5.7: 認識モデルと連想モデルを融合したモデル.  $U_3$  層,  $W_3$  層,  $U_4$  層および  $W_4$  層を融合している. このモデルでは認識時と連想時に同じ細胞が反応することになる. この融合された層は下側頭葉に対応する.

ち認識と連想は脳内において同じタスクであり, どちらの場合も同じ細胞がこれを実現している. 前節で構成したモデルでは認識は認識モデルで行ない, 連想は連想モデルで行なうというように別の細胞で実現されていたが, 実はこれらを融合の方が生理モデルとしてより妥当であることがわかる.

### 認識モデルと連想モデルの融合による新しいモデル

図 5.7 に認識モデルと連想モデルを融合した新しいモデル (融合モデルと呼ぶことにする) を示す. このモデルでは  $U_3$  層と  $W_3$  層, および  $U_4$  層と  $W_4$  層がそれぞれ融合され, 認識時と連想時に同じ細胞が反応する. またこの融合された層は下側頭葉に対応することになる. 融合モデルにおける動作や各細胞の数式表現は前節で説明したものと全く同じである. ただし認識時と連想想起時では使用する発火関数が異なり, 神経生理学で得られた知見と一致しない.

### 5.5.2 回転対応型ネオコグニロンとの比較

回転対応型ネオコグニロンは「回転に対して単純」なパターンを認識するための数理モデルであり, 融合モデルは「回転に対して非単純」なパターン

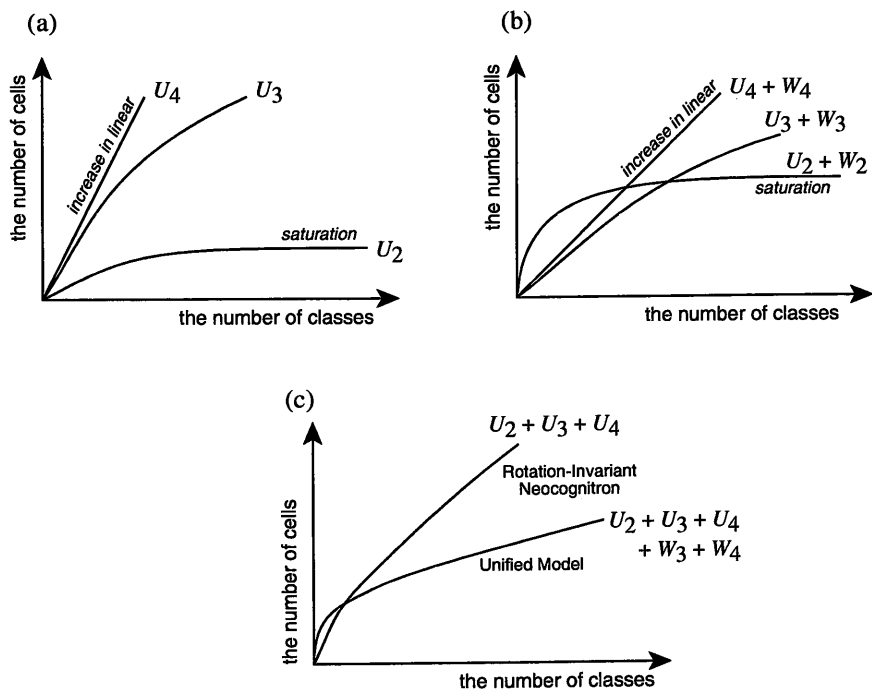


図 5.8: 回転対応型ネオコグニトロンと融合モデルの細胞数の比較. (a): 回転対応型ネオコグニトロンの各層における細胞面群の増加と, (b): 融合モデルの細胞面および細胞面群の各層における増加の様子を図示したものである. 横軸は学習するパターンのクラス数である. (c): 回転対応型ネオコグニトロンと融合モデルの全体の細胞数の増加の様子. クラス数が少ない場合には回転対応型ネオコグニトロンの方が細胞数の観点から有効であるが, 多い場合には融合モデルの方が有効である.

を認識するための数理モデルである。ここでは回転に対して非単純なパターンを認識する際に、どちらのモデルが工学的に有効であるかについて述べる。具体的には細胞数による比較をする。

回転対応型ネオコグニロンと融合モデルは前述の通りシードセル生成面を用いた学習方法により学習が行なわれる。このとき学習すべきパターンのクラスが増加すると、検出すべき局所的パターンも増加するために生成される細胞面および細胞面群が増加する。クラス数増加に対する細胞面および細胞面群の増加の様子を図 5.8 に示す。図 5.8(a) は回転対応型ネオコグニロンの細胞面群の増加を、図 5.8(b) は融合モデルの細胞面および細胞面群の増加を図示したものである。横軸は学習するパターンのクラス数である。 $U_2$  層と  $U_3$  層では学習すべきパターンのクラス増加に伴ない細胞面群数も増加するがあるクラス数に達するとそれ以上増加しない。これはクラス数が増えても検出すべき局所的パターンがすでに検出されていることに相当する。この現象は回転対応型ネオコグニロンおよび融合モデルの両者に共通するものである。また  $U_4$  層の細胞数はクラスの増加に伴ない線形増加する。これは  $U_4$  層の細胞はいわゆる grand mother cell として設計されているからである。

さて、融合モデルの第 2 細胞段は  $U_2$  層と  $W_2$  層を含むために回転対応型ネオコグニロンと比較して 2 倍の細胞を必要とする。そのため図 5.8(c) に示すように、パターンのクラス数が少ない場合には回転対応型ネオコグニロンの方が少ない細胞数で構築できる。しかし、回転対応型ネオコグニロンは第 3 モジュール以降において回転したパターンを直接検出する細胞を保持しているために、クラス数が多い場合には回転対応型ネオコグニロンの方がより多くの細胞数を必要とする。すなわち認識すべきパターンのクラス数が多い場合には融合モデルの方が細胞数の観点から有効であることがわかる。

## 5.6 まとめ

心的回転の考察を基礎として回転パターンを有効かつ高精度に認識可能なモデルを提案した。このモデルは心理実験から得られた結果を参考に構成されており、回転パターンを明示的、非明示的に学習することなく認識することが可能である。実際に手書き数字データベースを用いた数値シミュレーションにより、本モデルの有効性が確認された。

本章では心的回転に関する実験結果を基にしてモデルの構築を行なったが、心的回転の実現が目的ではない。心的回転とはイメージ内でパターンを回転する際にヒトが必要としている連続した操作であり、計算機上でパターンを回転させるためには心的回転のような（時間を要する）操作は不必要である。本研究ではパターンの回転操作を計算機上で行うため、仮の認識結果が正しい認識結果であれば回転したパターンの認識に要する時間は回転角に依存しない。計算機上でパターンを回転させる際に心的回転のような連続した回転



操作を取り入れるのであれば心理学的実験で得られた結果に一致するものと思われるが、工学的な重要性はない。

また第3章で提案した回転対応型ネオコグニトロンは「回転に対して単純」なパターンを認識するための数理モデルであり、本章で提案したモデルは「回転に対して非単純」なパターンを認識するための数理モデルである。両モデル共、その評価実験の際に数字パターンを使用したのが、これは単に評価を行なうために使用しただけである。ここで強調しておきたいことは、回転対応型ネオコグニトロンは回転に対して単純なパターンを認識するためのモデルであり、その第3章における評価実験の際に、(ヒトにとって) 回転に対して非単純な数字パターンを使用したこと自体に意味はない。

## 第6章 結論

### 6.1 本研究の成果

本研究は文字の回転、変形、位置ずれ、拡大縮小、ノイズに対して頑健な認識モデルの構築を目的とした。また認識モデル構築の際にはヒトの認識機構を神経生理学および心理学的に調査しその結果を基礎とすることで、より柔軟で信頼性の高い認識モデルの構築を目指した。以下各章で得られた結果を記す。

**第2章** 神経生理学的に妥当であり、工学的も有効なネオコグニトロンに注目し認識性能を評価した。その結果、文字の回転にはほとんど耐性がないことが明らかになった。

**第3章** 神経生理学および心理学で得られた実験結果を基にして、回転したパターンの認識を行なう回転対応型ネオコグニトロンを提案した。このモデルは局所的パターンの回転角を表現する軸をもつ細胞面群で構成される。細胞面群は従来型ネオコグニトロンの細胞面を拡張したものである。数値シミュレーションの結果、このモデルが文字の回転、変形、位置ずれ、拡大縮小、ノイズに対して非常に頑健であることが明らかになった。これら文字の多様性の全てに対して耐性がある単一の認識システムの構築に成功したのは本研究が初めてである。

**第4章** 構築した認識システムは工学的にも有効なものではなくてはならない。従来型ネオコグニトロンや回転対応型ネオコグニトロンが持つ重要なパラメータとしてしきい値があるが、適切なしきい値を決定するにはこれまで非常に長い時間を要してきた。そこで学習過程の解析を行ない、その結果から新しい生成アルゴリズムを提案した。このアルゴリズムは様々なしきい値を持つ大量のネオコグニトロン型モデルを非常に短時間で生成することができるアルゴリズムである。モデルの生成時間が劇的に減少できる上に従来の方法で生成したモデルと同等の性能を持っていることを、実際に数値シミュレーションにより確認した。

**第5章** 回転したパターンの認識に関する心理学的考察を行なった。また認識時に回転に依存しない局所的パターンを同定するために心理実験を行なった。これらの結果からトップダウン構造を持つ新しいモデルを提案した。このモデルが回転したパターンを明示的・非明示的に学習するこ

となく認識可能であることを数値シミュレーションにより確認した。鏡像パターンに対して提案モデルがヒトと同じような動作をすることは、非常に興味深い結果である。

## 6.2 今後の課題

本研究はネオコグニトロンを基礎にして議論をしてきた。ネオコグニトロンは Hubel と Wiesel の階層型仮説を採用して作成されたモデルであるために回転対応型ネオコグニトロンも階層型仮説を採用していることになる。ところで Hubel らは第 1 次視覚野に方位選択性を持つ神経細胞を発見したが、ある神経細胞の近隣の神経細胞はよく似た方位選択性を持つことも発見した [19]。Hubel らはこの結果から、方位選択性を持つ細胞の配置に関するモデルを立て、アイスキューブモデルと名付けたが、このモデルは  $U_1$  層の細胞面群と非常に良く似た性質をもっている。すなわち、アイスキューブモデルにおいてもパターン（線分）が回転すると発火する細胞がある方向に平行移動する。Hubel らの仮説に従って構成したネオコグニトロンを、回転に対しても頑健であるように拡張すると第 1 次視覚野におけるモデルと類似することは非常に興味深いことではあるが、Götz らによってアイスキューブモデルは否定され、ピンホイールモデルが新たに提案された [43]。Götz らのモデルは神経生理学的に裏付けられている。工学的観点から見ればどちらのモデルを採用しても同じ結果が得られると考えられるが、考察すべき事実ではある。

また第 5 章において、「仮定」とその「検証」によって回転したパターンを認識するモデルを提案したが、この考えをパターンのその他の多様性に適応できる可能性がある。例えば本研究ではすでに切り出された個別のパターンの認識に限定したが、実際にはどこにパターンが存在するかを知り得ない。この場合、パターンへの注意視が問題となる。注意視の問題に関する研究は無数にあるが、「仮定」と「検証」によるパターンの切り出しに関する研究はない。検討する価値は十分にあると思われる。

最後に、本研究の動機は従来の文字認識の研究方法に関する疑問から発生している。すなわち想定していない環境下においては動作しないことに疑問を感じたことから始まっている。これまでの研究はある認識システムや認識モデルがある環境下で動作しない場合にはその想定していない環境下で動作するようにシステムやモデルを改良、もしくは新しい機能を付加していく立場をとっている。しかし、この方法では文字が書かれている状況における全環境の記述が完了しない限り、すなわち物理現象を全て記述しない限り、いわゆる認識率 100% には達しない<sup>1</sup>。本研究で得られた数理モデルも当然想定されていない環境下で動作しないが、仮定と検証に基づく概念がそれを解決するものと思われる。なぜならば「仮定」とはモデルが保持している内部

<sup>1</sup> 「認識率 100%」の定義自体、研究者ごとに異なる。

世界であり、想定されていない環境と内部世界を「検証」することによって、想定されていない環境であるのかそうでないかを、少なくとも検出することはできるからである。モデル自身が「何を知り、何を知らないか」を検出することができれば想定されていない環境をモデルの内部世界に射影する際の手掛かりとなる可能性がある。すなわちこれまでの認識に関する研究で発生した問題の根本的な原因は、ある環境を想定して構築したモデルが、想定されていない環境に置かれた場合に、その環境を「想定した環境」なのか「想定していない環境」なのか判断できない（しない）ことにあると思われる。

今後は環境を自動的に内部世界に取り込む方法、すなわち知識の獲得方法に関する研究も行なう予定である。

## 関連図書

- [1] 飯島：“パターン認識理論”，森北出版株式会社 (1989).
- [2] 石井, 上田, 前田, 村瀬：“パターン認識”，オーム社出版局 (1998).
- [3] C. E. Shannon: “A mathematical theory of communication”, Technical report, Bell System Technical Journal (1948).
- [4] 森：“知識に依存しない言語モデルに関する研究”，博士学位論文, 東北大学大学院工学研究科電気・通信工学専攻 (1997).
- [5] N. R. Carlson: “Physiology of behavior”, Allyn and Bacon (1998).
- [6] K. T. Spoehr and S. W. Lehmkuhle: “Visual information processing.”, W.H.Freeman and Company (1982).
- [7] M. Wertheimer: “Principles of perceptual organization.”, Princeton (1958).
- [8] R. N. Shepard and J. Metzler: “Mental rotation of three-dimensional objects.”, Science, **171**, pp. 701-703 (1971).
- [9] L. A. Cooper and R. N. Shepard: “The time required to prepare for a rotated stimulus.”, Memory & Cognition, **1**, pp. 246-250 (1973).
- [10] L. A. Cooper and R. N. Shepard: “Demonstration of a mental analog of an external rotation.”, Perception & Psychophysics, **19**, pp. 296-302 (1976).
- [11] R. Lawson: “Achieving visual object constancy across plane rotation and depth rotation.”, Acta Psychologica, **102**, pp. 221-245 (1999).
- [12] W. C. McCulloch and W. H. Pitts: “A logical calculus of the ideas immanent in neural nets.”, Bull. Math. Biophys., **5**, pp. 115-133 (1943).
- [13] D. O. Hebb: “The Organization of Behaviour”, New York: Wiley-Interscience (1949).

- [14] K. Fukushima: "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biological Cybernetics*, **36**, pp. 193-202 (1980).
- [15] K. Fukushima, S. Miyake and T. Ito: "Neocognitron: A neural network model for a mechanism of visual pattern recognition", *IEEE Trans. Sys., Man, and Cybern.*, **13**, pp. 826-834 (1983).
- [16] K. Fukushima: "Neocognitron: A hierarchical neural network capable of visual pattern recognition", *Neural Networks*, **1**, pp. 119-130 (1988).
- [17] D. H. Hubel and T. N. Wiesel: "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", *Journal of Physiology*, **160**, pp. 106-154 (1962).
- [18] D. H. Hubel and T. N. Wiesel: "Receptive fields and functional architecture of monkey striate cortex", *Journal of Physiology*, **195**, pp. 215-243 (1968).
- [19] D. H. Hubel and T. N. Wiesel: "Sequence regularity and geometry of orientation columns in the monkey striate cortex", *Journal of Comparative Neurology*, **158**, pp. 267-293 (1974).
- [20] D. Ferster: "Spatially opponent excitation and inhibition in simple cells of the cat visual cortex", *Neuroscience*, **8**, pp. 1172-1180 (1988).
- [21] S. Zeki, J. D. Watson, C. Lueck, K. Friston, C. Kennard and R. S. Frackowiak: "A direct demonstration of functional specialization in human visual cortex", *Neuroscience*, **11**, pp. 641-649 (1991).
- [22] M. B. Reid, L. Spirkovska and E. Ochoa: "Rapid training of higher order neural networks", *Proc. Int. Joint Conf. Neural Networks*, Vol. 1, pp. 689-692 (1989).
- [23] R. G. W. B. Widrow and R. A. Baxter: "Layered neural nets for pattern recognition", *IEEE Trans. Acoust., Speech, Signal Processing*, **36**, (1988).
- [24] M. Fukumi, S. Omatu and Y. Nishikawa: "Rotation-invariant neural pattern recognition system estimating a rotation angle", *IEEE Trans., Neural Network*, **8**, pp. 569-581 (1997).
- [25] J. basak and S. K.P.: "Psychop—a psychologically motivated connectionist system for object perception", *IEEE Trans. Neural Networks*, **6**, 6, pp. 1337-1354 (1995).

- [26] K. Fukushima: "Analysis of the process of visual pattern recognition by the neocognitron", *Neural Networks*, **2**, pp. 413-420 (1989).
- [27] K. Fukushima and N. Wake: "An improved learning algorithm for the neocognitron", *Artificial Neural Networks*, **1**, 2, pp. 497-505 (1992). Amsterdam: North-Holland.
- [28] K. Fukushima and N. Wake: "Improved neocognitron with bend detecting cells", *Joint Conf. on Neural Networks*, Vol. 4, pp. 190-195 (1992).
- [29] K. Fukushima, M. Okada and K. Hiroshige: "Neocognitron with dual c-cell layers", *Neural Networks*, **7**, 1, pp. 41-47 (1994).
- [30] K. Fukushima and M. Tanigawa: "Use of different thresholds in learning and recognition", *Neurocomputing*, **11**, 1, pp. 1-17 (1996).
- [31] A. Nye: "Xlib programming manual", Vol. 1, chapter Appendix A, O'Reilly & Associates, Inc. (1995).
- [32] A. Paeth: "A fast algorithm for general raster rotation", *Graphics Interface*, pp. 77-81 (1986).
- [33] G. Wang, T. Tanaka and M. Tanifuji: "Optical imaging of functional organization in the monkey inferotemporal cortex", *Science*, **272**, pp. 1665-1668 (1996).
- [34] "電子技術総合研究所". <http://www.etl.go.jp/>.
- [35] O. Yadid-Pecht, M. Gerner, L. Dvir, E. Brutman and U. Shimony: "Recognition of handwritten musical notes by a modified neocognitron", *Machine vision and applications*, **9**, 2, pp. 65-72 (1996).
- [36] D. R. Lovell: "The neocognitron as a system for handwritten character recognition: limitations and improvements", PhD thesis, Univ. Queensland, Brisbane, Australia, <ftp://archive.cis.ohio-state.edu/pub/neuroprose/Thesis> (1994).
- [37] D. R. Lovell, T. Downs and A. C. Tsoi: "An evaluation of the neocognitron", *IEEE Trans. Neural Networks*, **8**, pp. 1090-1105 (1997).
- [38] A. M. Chiang and M. L. Chuang: "A ccd programmable image processor and its neural network applications", *IEEE J. Solid-State Circuits*, **26**, pp. 1894-1901 (1991).

- [39] L. C. Jain and B. Lazzerini: "Knowledge-based Intelligent Techniques in Character Recognition", CRC Press (1999). ISBN:0-8493-9807-X.
- [40] K. Fukushima: "Neural network model for selective attention in visual pattern recognition and associative recall", *Applied Optics*, **26**, 23, pp. 4985-2992 (1987).
- [41] C. Stanfill and D. Waltz: "Toward memory-based reasoning", *Communication ACM*, **29**, pp. 1213-1228 (1986).
- [42] M. Moscovitch, S. Kapur, S. Koehler and S. Houle: "Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography study in humans", *Proc. of the National Academy of Sciences*, Vol. 20, pp. 3721-3725 (1995).
- [43] K. G. Götz: "Do "d-blob" and "l-blob" hypercolumns tessellate the monkey visual cortex?", *Biological Cybernetics*, **56**, pp. 107-119 (1987).



## 研究業績

### 学術論文

Shunji Satoh, Jousuke Kuroiwa, Hirotomo Aso and Shogo Miyake,  
“A rotation-invariant neocognitron,” *Systems and computers in Japan*,  
**30**, 4, pp. 31–40, 1999.

佐藤俊治, 黒岩丈介, 阿曾弘具, 三宅章吾,  
“回転対応型ネオコグニトロン,” *電子情報通信学会論文誌*, **J81-D-II**,  
6, pp. 1365–1374, 1998.

### 査読付き国際会議

Shunji Satoh, Hirotomo Aso, Shogo Miyake and Jousuke Kuroiwa,  
“Pattern recognition system with top-down process of mental rota-  
tion,” *Proceeding of the fifth International work-conference on arti-  
ficial and natural neural networks*, **1**, pp. 816–825, 1999 (invited  
session).

Shunji Satoh, Hirotomo Aso, Shogo Miyake and Jousuke Kuroiwa,  
“Recognition of hand-written patterns by rotation-invariant neocog-  
nitron,” *Proceedings of the fifth international conference on neural  
information processing*, **1**, pp. 295–299, 1998.

Shunji Satoh, Jousuke Kuroiwa, Hirotomo Aso and Shogo Miyake,  
“Recognition of rotated patterns using neocognitron,” *Proceedings of  
the fourth international conference on neural information processing*,  
**1**, pp. 112–116, 1997.

### 投稿予定の学術論文

Shunji Satoh, Hirotomo Aso, Shogo Miyake and Jousuke Kuroiwa,  
“Rotation-invariant Neocognitron and a new generative algorithm,”  
*IEEE Trans. on Neural Networks*.

## 学会発表等

佐藤俊治, 阿曾弘具, 三宅章吾, 黒岩丈介,  
“心的回転を基に構成されたパターン認識モデル,” 電子情報通信学会  
ニューロコンピューティング研究会, 1999年3月, 於玉川大学.

佐藤俊治, 黒岩丈介, 阿曾弘具, 三宅章吾,  
“回転角に対応する可変結合を持つ回転対応型ネオコグニトロン,” 電子  
情報通信学会総合大会, 1997年3月, 於関西大学.

佐藤俊治, 黒岩丈介, 阿曾弘具, 三宅章吾,  
“回転対応型ネオコグニトロンの改良,” 電子情報通信学会総合大会, 1996  
年9月, 於金沢大学.

佐藤俊治, 黒岩丈介, 阿曾弘具, 三宅章吾, 猪苗代盛,  
“回転パターンを認識するネオコグニトロン,” 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年3月, 於玉川大学.

佐藤俊治, 黒岩丈介, 阿曾弘具, 三宅章吾, 猪苗代盛,  
“回転パターンを認識する改良型ネオコグニトロン,” 電気関係学会東北  
支部連合大会, 1995年8月, 於岩手大学.

## 書籍

L.C. Jain and B. Lazzerini(Editor), “Intelligent character recognition,”  
chapter 3, CRC Press LLC, 1999. ISBN: 0-8493-9807-X

## その他

佐藤俊治,  
“視覚情報の統合過程とそのモデル化,” 東北大学数物談話会, 於東北大  
学応用物理学科棟, 1997年5月.

## 謝辞

本研究を進めるにあたり、長年にわたり数多くの御指導とともにこの研究の機会を与えて下さりました東北大学大学院工学研究科電気・通信工学専攻 阿曾弘具教授に心から感謝致します。東北大学大学院工学研究科応用物理学科 三宅章吾講師には日々の研究生活において貴重な御意見 御指導を戴きましたことを、また、貴重な時間をいただいて議論していただいたことを深く感謝致します。広島大学総合科学部数理情報科学 黒岩丈介助手にはこの研究分野に興味を与えて下さるとともに適切な助言、御指導を戴きましたことを心より感謝致します。東北大学工学研究科電気・通信工学専攻 大町真一郎助教授には研究方針や研究に対する姿勢等を教えて頂きましたことを深く感謝致します。

本論文をまとめるにあたり貴重な御意見を頂いた東北大学大学院工学研究科電子工学専攻 星宮望教授、東北大学電気通信研究所 矢野雅文教授に心から感謝致します。電気通信大学電気通信学部情報通信工学科 福島邦彦教授、ERATO Project 岡田真人博士には貴重な御指導、御意見を戴きましたことを深く感謝致します。日本大学工学部情報工学科 猪苗代盛教授には貴重なお時間を割いていただき議論して戴いたことを深く感謝致します。

日々の研究におきましては、多くの助言ならびに研究環境を整えて戴きました東北大学情報処理教育センター 後藤英昭助手、宇都宮大学工学部 森大毅助手、東北大学情報科学研究科情報基礎科学専攻 鈴木基之助手、阿曾研究室 菅谷至寛氏に感謝します。東北大学電気通信研究所 伊藤真氏、また、東北大学工学研究科電気・通信工学専攻阿曾研究室 高田直幸氏、同 宮野靖弘氏、須藤貴志氏、下村正夫氏を始めとする阿曾研究室の皆様にも数多くの御意見、御討論を戴きましたことを深く御礼申し上げます。

本研究を最後までやり遂げることができたのも佐藤哲二様、私の両親、弟、故 西島こと様、西島英男様、露久保政子様、露久保和江様、鈴木雅美嬢、及び友人の方々の温かなご支援の賜物であります。ここに深く感謝します。