

両眼立体視における奥行き面知覚の分節化問題と 振動的ニューラルネットモデル

村田 勉 (金沢工業大学・電子工学科)

清水 博 (東京大学・薬学部)

§1 序論

知覚分節化

現在のコンピューターと比較したとき、人間の脳はその高速並列情報処理の能力において特徴づけられるだろう。脳が変化に富む複雑な感覚入力パターンから意味のある知覚を得るといふ情報処理を日常的に行なっているのに対し、現在のコンピューター・ビジョンにとって日常風景の分析は大変困難な課題である⁽¹⁾。たとえば、自動車の運転席からの交通信号・標識の認識をコンピューターに実時間で行なわせることは、路上環境の多様性、複雑性ゆえにほとんど不可能であろう。

脳の高速並列処理の一つの鍵となる働きは、画像要素を意味のある知覚物体へと分割し帰属させていく「分節化(画像分割)」と呼ばれる過程である。この分節化によって、脳は入力の複雑さの度合いを軽減させ、知覚的に意味のある領域へのより集中的な分析(「注意(attention)」)を可能にしているのである⁽²⁾。D. Marrが彼の視覚の計算理論において主張したところによると、画像は脳内での表現を得た後に、まず可視表面(その境界も含めて)の情報を抽出するために適切に分節されなければならない。ただし、このとき知覚物体に関する特殊な知識(たとえば、電話、自動車などのメモリー)を用いて、画像をいきなりある物体に対応する領域に分割しようというやりかたは成功しない。それは、画像情報自体のなかには物体に分割するために十分な情報が存在しないので、処理系のほうで物体に関する膨大な知識と仮説検証の過程を必要としてしまうからである。(かつて人工知能的アプローチが極めて限定された状況でしか成功しなかった理由がここにある。)したがって、何の物体に属するのかを決定する前に、(特殊な知識ではなく)外界の物理的性格を反映したより一般的な拘束条件を用いて可視表面の構成をおこなう必要がある。したがって分節化において物体への分割の前に行なわれるべき重要な過程は面の分節化であることになる。(画像から可視表面の表現を得るまでのこれらの過程は初期視覚と呼ばれている。)

両眼立体視の分節化問題

さて、両眼立体視は初期視覚での分節化を含む複雑入力の処理過程の研究対象として大変優れている。両眼立体視とは、左右眼の画像の(水平方向の)わずかなずれに基づいて物体の奥行き情報を抽出する大脳皮質視覚野のはたらきである。図1aのようなランダムドットステレオグラム(RDS)を両眼視すると、入力のドットがまばらであるにもかかわらずそれらを含む連続した奥行き面がその輪郭さえも伴って知覚される。この図のように入力点を欠く狭い領域(ギャップ)があっても、そこを充填(fill-in)して奥行き面が連続して存在するように見える。このひとかたまりの奥行き面のような知覚

をコヒーレントな知覚と呼ぶことにしよう。図1bでは奥行き面が2重にオーバーラップするようにRDSを作っている。奥行きが多重になる場合には、左右の図における重なっている領域を見比べても点の同じ分布パターンをしている所はほとんどない。左右それぞれの図形の中ではすぐ隣同士になっている点であっても異なる奥行き面に帰属されることが起こるからである。言うまでもなく、刺激要素であるドットの一つ一つを考えれば反対眼に呈示されているどのドットとも対応しうるわけであるが、無数の可能な対応づけの中から脳は認識として意味のある対応関係を極めて迅速に検出して奥行き面の知覚をもたらすのである。(これを両眼立体視の対応問題と言うが、可能性をしばらく対応を決定させるために働く拘束条件については§2で述べる。)心理物理学実験によると、両眼にRDSを呈示してから50ミリ秒程度で脳内に奥行き図形知覚の情報が生成されるといわれている⁽³⁾。両眼立体視が人間の脳の高速並列情報処理の優れた例であると思われる理由はこれらのことにある。さらに、図1aと図1bを比較すると、手前に浮かんで見えるドットのセットはどちらも全く同一のものを与えているのだが、奥の面をオーバーラップさせるかさせないかに依存して手前のギャップが充填されたりされなかったりすることがわかる。つまり、同じ奥行き点のセットでも、他の奥行き面との関係によって一つの面に属したり二つの面に分かれたりするのである。したがって、奥行き面の分節化は対応問題が解かれただけではまだambiguous(あいまい)な状態にあり、知覚分節化に関する拘束条件を与えて分節化の起き方を決定しなければならない。このことを我々は両眼立体視の分節化問題と呼ぶことにする。両眼立体視の分節化が初期視覚で行なわれていることを示す心理物理学的研究が近年Nakayama、下條らによって提出されている⁽⁴⁾。また、サルにRDSを見せる電気生理学実験によると、RDSの奥行き情報に対応して発火する神経細胞は大脳皮質への視覚信号の入り口であるV1、V2領域ですでに存在している⁽⁵⁾。これらのことが、両眼立体視が初期視覚における知覚分節化の研究対象として優れていると述べた理由である。

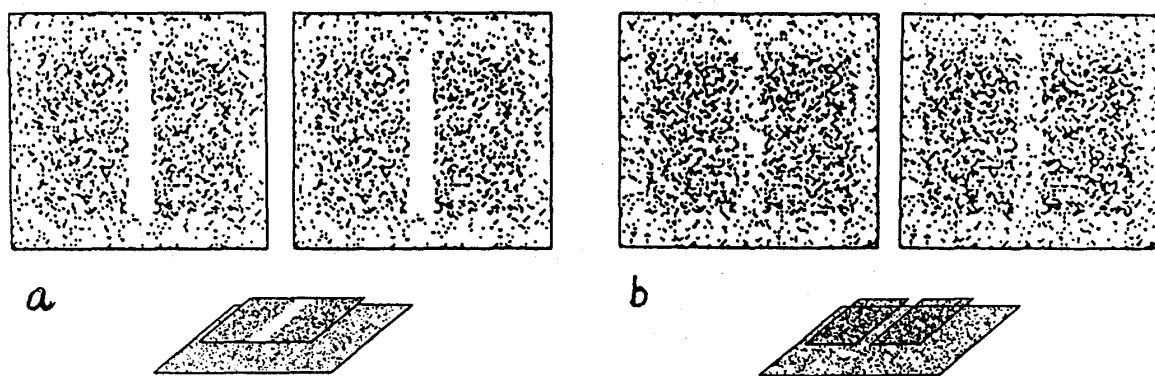


図1：手前の奥行き点のセットは同一であるにもかかわらず、背景の違いによって異なる分節化をするRDS。(非交差で観察してください。)

問題点

さて、脳でおこなわれる分節化の処理を考えると、次のような疑問が生ずる。まず、分節された奥行き面や面間の関係は(脳の)ニューラルネットにおいてはどのように表現されているのか、そして、ランダムドットパターンの入力から分節された奥行き面を

得るまでの過程は、ニューラルネットによってどのように自己組織的に行なわれるのか、ということである。我々の研究グループでは以前より、脳内のコヒーレントな知覚が神経活動の振動の同期（引き込み）によって表現されている可能性を理論モデルによって提唱してきた（ホロビジョン認識モデル⁽⁶⁾）。また、近年の電気生理学実験により、コヒーレントな知覚と神経活動の振動の同期との間の密接な関係が実証された⁽⁷⁾。

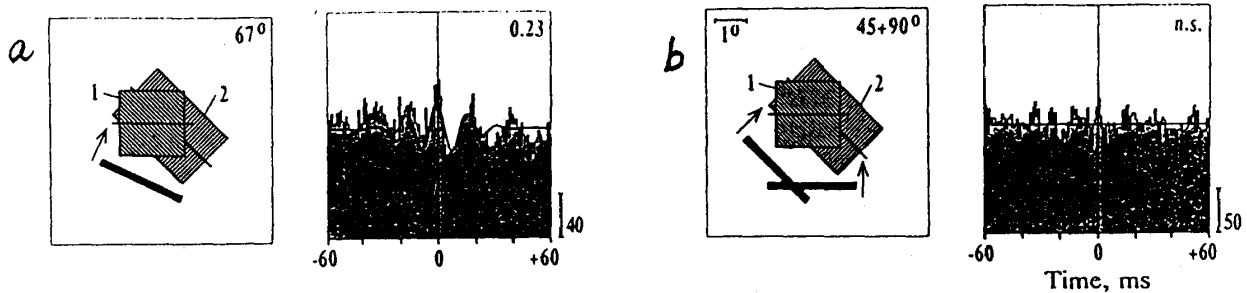


図2：重なった受容野をもつが方位選択性の異なる2つの神経の活動を記録しながら、その受容野に視覚刺激を呈示する。（2つの最適方位の中間の方位を持つ）1本の棒を呈示すると2つの細胞は同期して活動する（a）が、（それぞれの最適方位の）2本の異なる棒が出た時は同期しない（b）。

これらの知見をふまえ、本研究では奥行き面の分節化問題を解決するために両眼立体視の動的ニューラルネットモデルを提出する。このモデルでは、ニューロンは振動的な活動を示し、コヒーレントな奥行き面は奥行きニューロン活動の同期によって表現される。そこで解決すべき問題は、

- (1) 振動的ニューラルネットにおいて両眼立体視の対応問題はどうやって解かれなければならないか、
 - (2) 両眼立体視の分節化問題を解くための適切な拘束条件は何であり、それをニューラルネットの中にどのように実現すべきか、
- ということである。

§ 2 両眼立体視の振動的ニューラルネットモデル

基本構造とニューラルオシレーター

網膜画像は2次元であるが、両眼立体視の立体情報の手掛かりは左右画像の1次元上（多くの場合水平方向）のずれ（視差）である。（垂直視差は奥行き知覚にあまり影響しない。）このため、モデリングは1次元入力としても一般性を失わない。（容易に2次元入力に拡張できる。）

モデルは、大脳皮質視覚野のV1（V2）を想定し、図3aのように左右眼からの入力信号（ドットの有無）をうけとる単眼性細胞と、両眼視差に選択的に反応する両眼性細胞よりなる。各細胞が図3bのように抑制性細胞と相互結合してニューラルオシレーターを構成する。V1、V2において広い範囲の視差に選択的に反応する両眼性細胞が存在すること、および抑制性細胞の偏在性は生理学的に明らかにされている^(8, 9)。

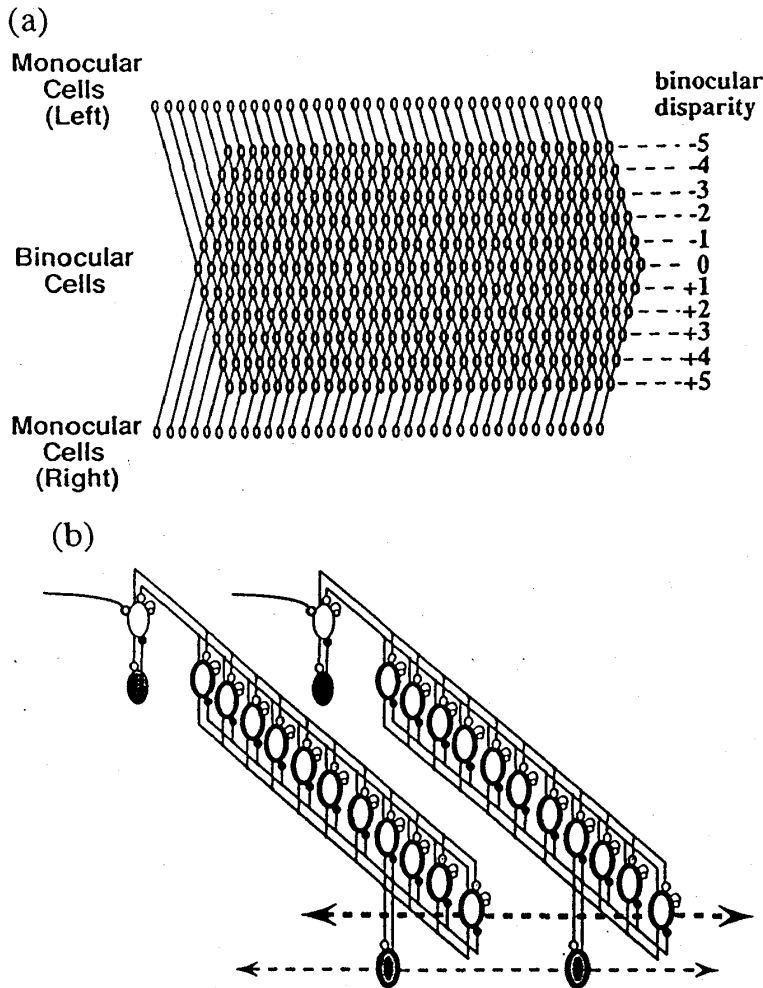


図3：a. 基本構造。

b. 一部分の詳細構造。斜線で塗られた細胞は抑制性細胞である。太い点線は同期性結合、細い点線は脱同期性結合を表す。

対応問題と拘束条件

左右画像の対応付けは、奥行き表面に関する一般的な知識（仮定）を拘束条件として入れることによって解かれる。それらは、1) 適合性（黒いドットは反対眼の黒いドットと対応する）、2) 唯一性（一つのドットは一つのドットと対応する）、3) 連続性（奥行き面は剣山のようにでこぼこではなく、できるだけ滑らかなものを知覚する）、というものである。（これらはやはり D. Marr らによって提案され成果をおさめた⁽¹⁰⁾。）本モデルでは、適合性条件は、左右眼からともに入力を受けた両眼性細胞のみが対応の候補として興奮しうるものとして取り入れられている。連続性条件は、ある近傍での同じ視差の両眼性細胞間の相互興奮結合として取り入れられている。（“同じ視差”というのを”近い視差”とすれば、（視差勾配知覚限界の範囲内で）奥行きの変化する斜面や曲面も扱うことができるが、今回は簡潔化されている。）

シナプス・モジュレーションと振動的競合系

最後に唯一性条件であるが、これは一つの単眼性細胞（左右それぞれ）から入力を受ける一群の両眼性細胞の中から高々一つの細胞が対応付けの答えとして選ばれることを要求しており、基本競合系のしくみによって実現される。甘利らによって提出された基本競合系⁽¹¹⁾は、競合する細胞群の中で最大入力を受ける細胞が抑制性細胞のグローバルなフィードバックを通じて他の全ての細胞を抑制するという winner-take-all機構であ

る。しかし、本モデルのように振動的活動をする細胞群では、そもそも抑制性信号は興奮性細胞を不活化するくらい強いものがあるのでそのまま上記の機構を適用することはできない。そこで、勝者細胞とそれ以外との格差を拡大していくための機構として単眼性細胞から両眼性細胞への入力結合にシナプス・モジュレーションを導入し、シナプス前後の活動相関に依存して結合強度を変化させることにする。これにより、競合細胞数が大きい時も、入力の差が小さいときも唯一の細胞を選び出すことに成功した(図4)。本モデルはこの機構の導入により、入力ドットの密度が50%以下の時、2重面検出において85%以上の正解率をあげ満足できる対応付け能力を示した。短時間スケールで起きるシナプス・モジュレーションは von der Malsburg⁽¹²⁾ によって理論的観点から導入され、他の知覚分節化モデルでも用いられて成果をあげている⁽¹³⁾。

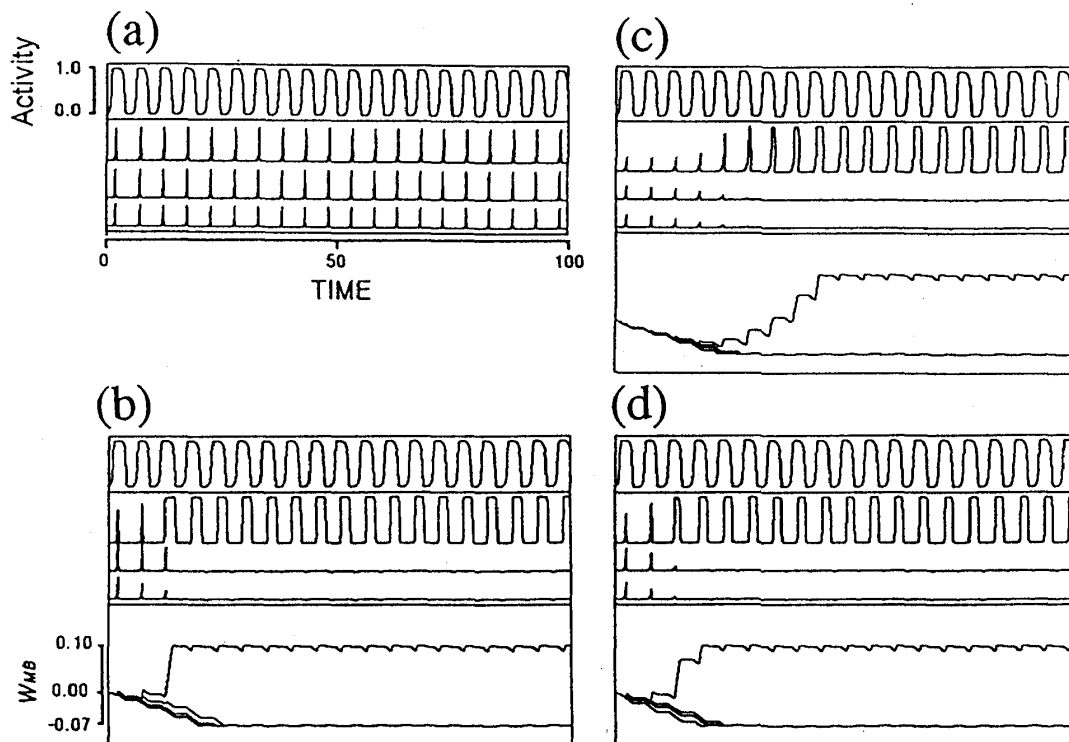


図4：a. シナプス・モジュレーションのないとき。一番上は単眼細胞の活動。その下は競合する両眼細胞5つのうちの3つの活動。b. シナプス・モジュレーションを導入したとき。一番下はシナプス強度の変化。c. 入力強度の差が極めて小さい(1%以下)とき。d. 競合細胞の数が多(30個)とき。

分節化問題と拘束条件

奥行き面の分節化を決めるための拘束条件は、知覚体制化に関する心理学的法則から学び、より一般性のある条件を選ばなければならない。また、選ばれた拘束条件が妥当なものかどうかは、実際に生ずる分節化についての心理物理学的事実の蓄積と照らし合わせて検討されるべきである。本研究では第一歩としてゲシュタルト法則の基本的な2原理⁽¹⁴⁾を拘束条件として採用する：

- 1) 近接性(空間的に近い刺激は同じ面に属する)、
- 2) 最大単純性(より単純な(小さい面積の)面をつくる)。

これらの条件はそれぞれ図3bにおける同期性結合と脱同期性結合として取り入れられている。

ここで、神経活動の脱同期を得るために短なるノイズではなく脱同期性結合を導入した理由は、入力ドットの密度が高いときに同期性結合により全体が同期してしまうことを防ぐためである。同期性結合と脱同期性結合との競合によって適切なサイズの分節が得られるような機構が期待できる。両眼性細胞の活動はパルス状なので、脱同期性結合としての相互抑制結合は anti-phasic な位相関係を安定化させるというよりは、in-phase になることを妨げるように働いていると思われる。

§ 3 結果

奥行き面の分節化

まず、RDS の典型例としてよく用いられる、背景と手前に浮かんだ四角形に対応する入力パターンをモデルに与えてみる (図5a)。モデルはすべての対応を正しく検出した。2つの細胞の活動の cross-correlation をとってみると、手前の面 (上の面) 内、背景の面 (下の面) 内ではそれぞれよく同期していることがわかるが、面間では同期が成り立っていない。全ての細胞の組み合わせについて、cross-correlation から振動の平均周波数における coherency を計算したものが図5cであるが、活動の同期性によって細胞群 1-5 (上の面に属する) と細胞群 6-22 (下の面に属する) とに分けられることがわかる。つまりこのモデルは、入力点に対応づけただけではなく、それらが異なる奥行き面に対応する2つのグループへと分節されたことを明らかに示している。

ギャップに依存した分節化

奥行きが同一であっても入力ギャップによって面が分かれる場合があり、図6はそのような例を扱っている。ギャップを入力点によって埋めてやることによりコヒーレントな面が回復する。

他の奥行き面のオーバーラップによる分節化

図1で扱ったようなケースに対応して、入力ギャップの大きさが同じであっても他の奥行き面をオーバーラップさせることにより面の分節化が引き起こされることが示された (図7)。この場合、脱同期性結合は面間の干渉を生ずるしくみとして不可欠の役割を果たしている。

これらの結果より、同じ神経対であっても (刺激のグローバルな状況に依存して) 同じ面に属する場合には同期した活動をするが、異なる面に属する場合には非同期的になることが示された。このことは、図2の生理学実験で1つの棒刺激か2つの棒刺激かによって同期が生成、消滅することに対応している。

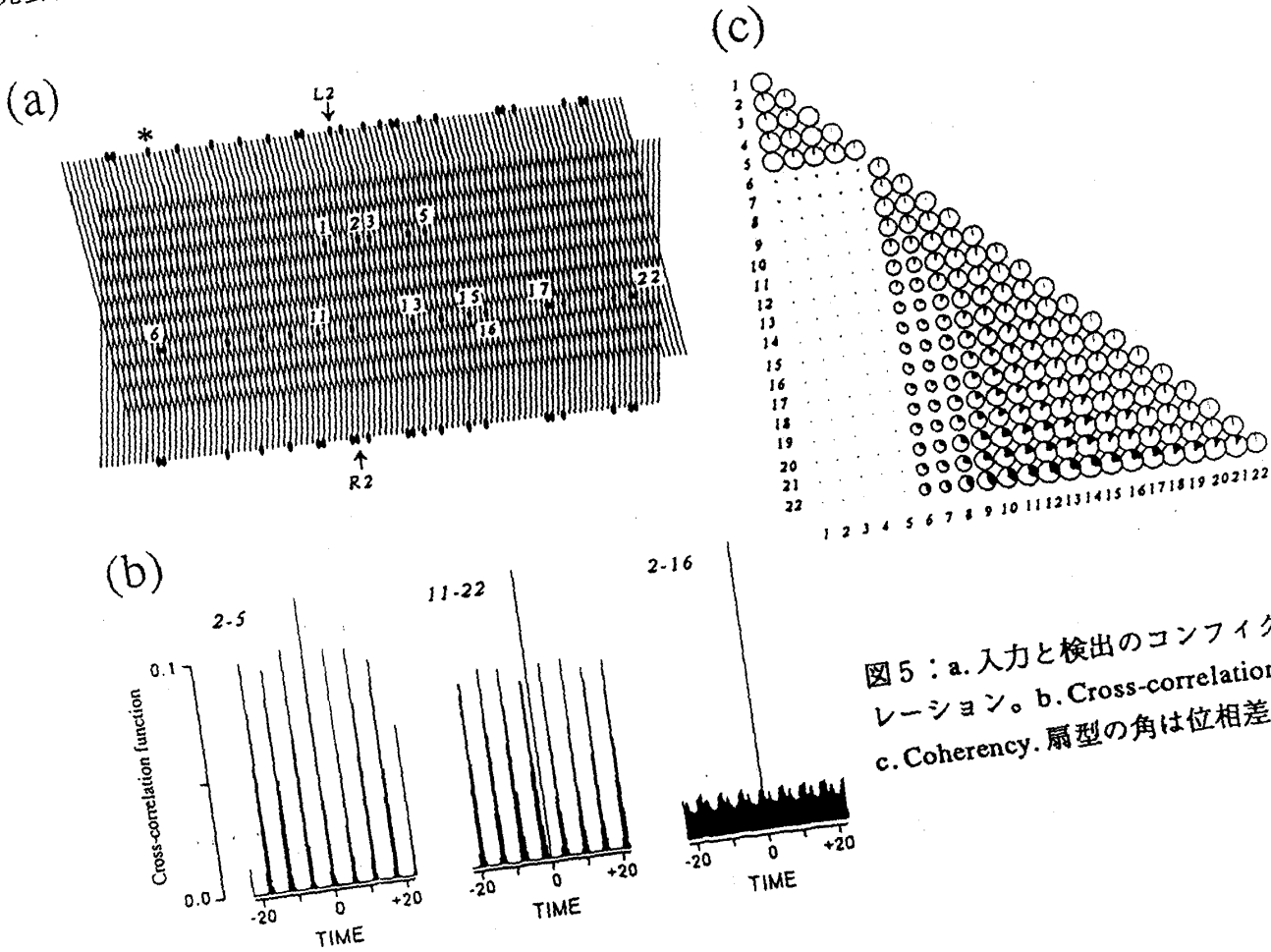


図5：a. 入力と検出のコンフィグレーション。b. Cross-correlation. c. Coherency. 扇型の角は位相差。

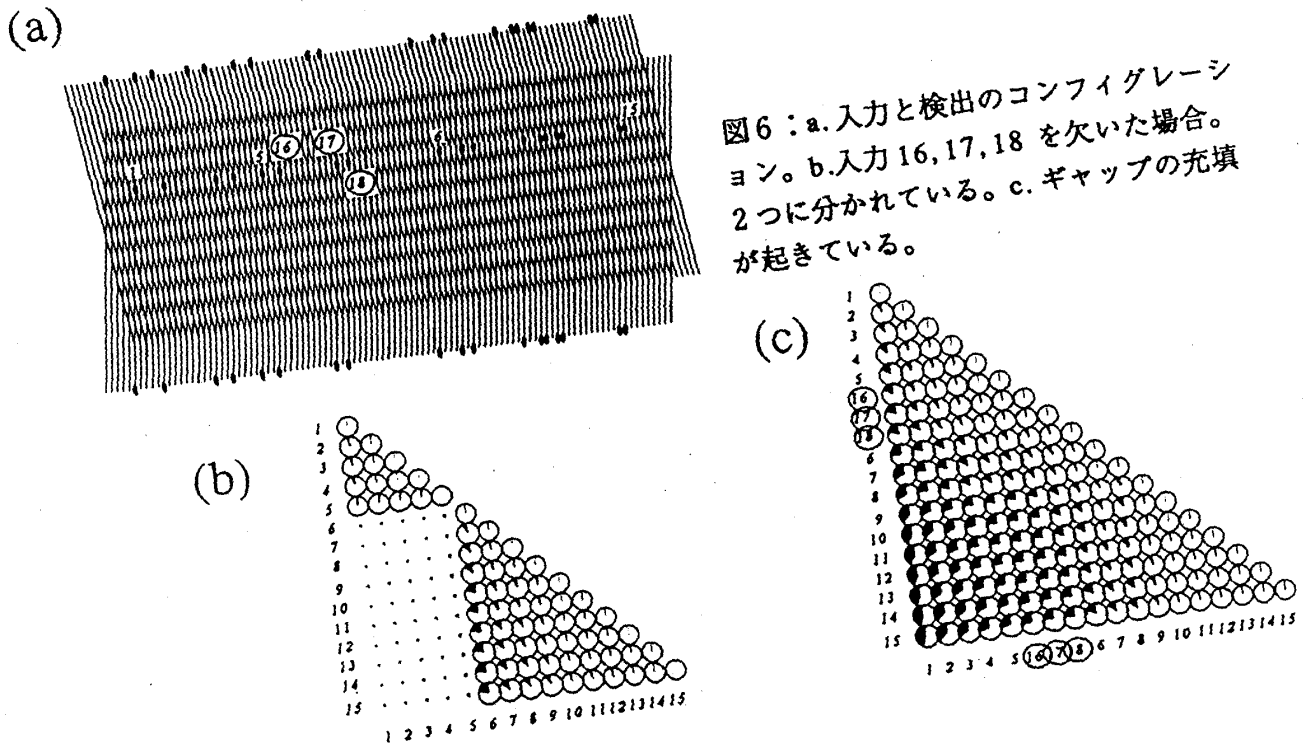


図6：a. 入力と検出のコンフィグレーション。b. 入力16, 17, 18を欠いた場合。2つに分かれている。c. ギャップの充填が起きている。

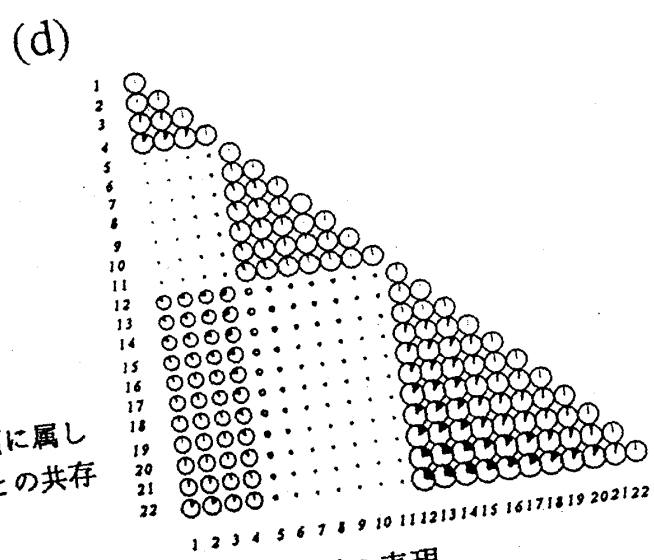
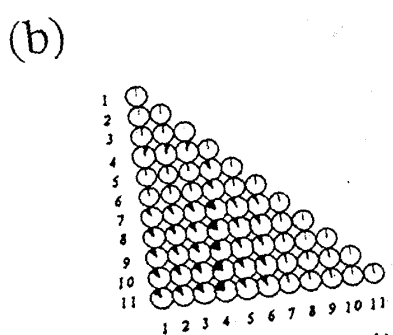
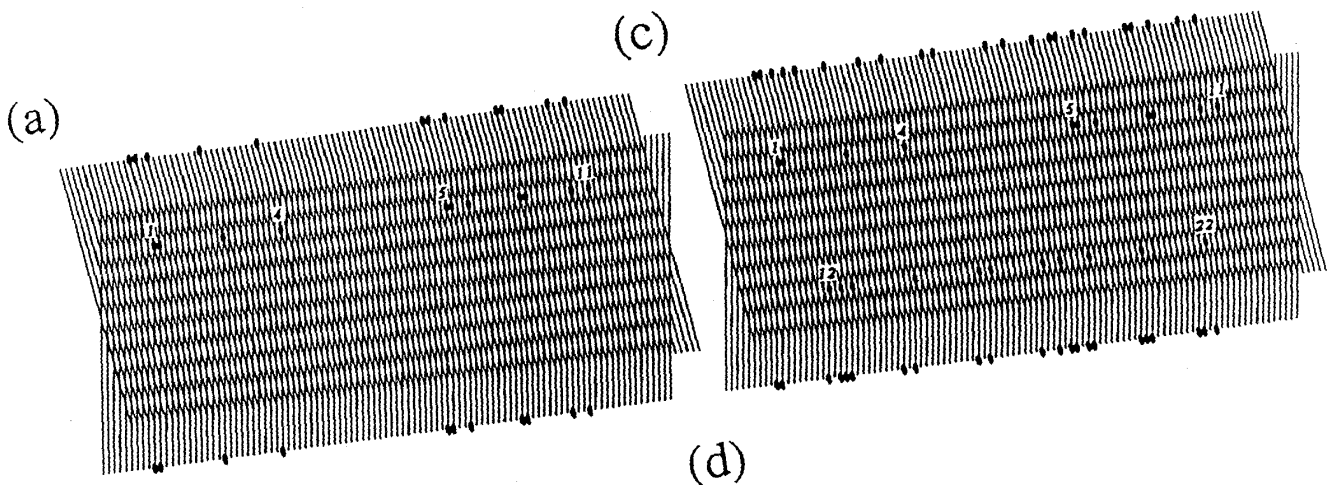


図7：(a, b) 単独奥行きでは一つの面に属していた細胞1-11が、(c, d) 他の面との共存下では1-4と5-11に分かれている。

両眼単眼リカレント結合による単眼性細胞の分節化と帰属の表現

図3bで示したように、このモデルでは両眼性細胞から単眼性細胞へのリカレント結合が入っているため、単眼性細胞同士はなんら相互作用結合を持っていないにもかかわらず、両眼性細胞のレベルでおきた分節化の影響が降りてきてここでも分節化が引き起こされる(図8)。

ここで関心のあることは次の2点である。
 1点めは、両眼対応付けに失敗した非対応の単眼性細胞は分節化のうえでもどのグループにも属さず孤立する。単眼入力非対応性がこのようにしてexplicitな表現を与えられたのは初めてのことである。2点めは他の視覚モジュールの分節化との関係である。このモデルは初期視覚のモデルとして大脳皮質のV1(あるいはV2)を想定していると述べたが、実際、V1の4層の細胞は大脳への視覚信号の入り口にあたり単眼性細胞が多い。視覚信

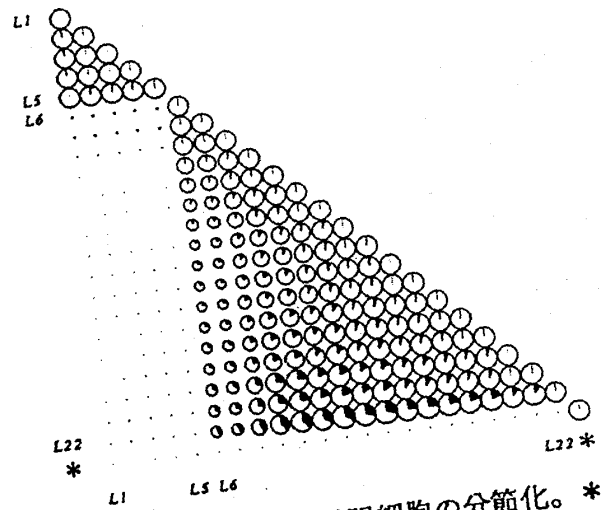


図8：図5aでの単眼細胞の分節化。*は非対応の細胞。

号は4層からそれ以外の層に送られて様々な処理を受けると考えられている。4層以外の層では両眼性細胞の割合が高く、4層へのリカレント結合も見つかっている⁽¹⁵⁾。もし本モデルで示したように、両眼性細胞の分節化が4層の単眼性細胞の分節化をも引き起こすとしたら、ここから信号を受け取っている他の視覚モジュールでもその影響を受けることが予想される。このことは、例えば立体視と情報のパスをかなり共有しているといわれる運動視において、立体視で分節された領域が運動視でも共通運命をもつというような現象に関する心理物理学的な研究と関連させることができれば、大変興味深いものになるだろう。

§ 4 結語・議論

本研究では、両眼立体視において解決すべき問題として対応問題のほかに分節化問題の存在することを明らかにし、これが知覚のコヒーレンスに関するものであり、視覚情報処理過程の中で重要な位置を占めることを示した。そして、神経活動の同期によりコヒーレントな知覚を表現する動的な両眼立体視ニューラルネットモデルを提出し、対応問題を満足できる正解率で解けることを示し、分節化問題のいくつかの例を解決した。さらに、入力要素とコヒーレント知覚との帰属関係の explicit な表現に関する問題を議論した。

面の情報

最後に本モデルの問題点と今後の課題について議論する。本モデルでは生理学実験と対応した同期した神経活動を得ることができたが、その位相関係については実験との不一致が見られる。生理学実験では皮質上のかなり離れた細胞でもコヒーレントな刺激に反応するときにはほぼ完全に in-phase (位相差ゼロ) で同期していることが示されている。このこと自体驚くべきことである。一方、本モデルでは phase-lock しているとはいえ、細胞が離れると位相差を生ずる。振動子集団の理論モデルによると、近傍結合だけでは距離に依存した位相シフトを避けることが難しいが、グローバルなフィードバックを入れることにより in-phase の活動が得られる⁽¹⁶⁾。このことは、本モデルのような知覚系のモデルが近傍同士の同期性結合だけではなく、よりグローバルなフィードバックの仕組みを取り入れなければならないことを要求していると思われる。両眼立体視のモデルにおけるグローバル情報としてまず考えなければならないのが、面そのものの情報である。(本モデルでは面の分節化は扱ったが、面そのものの explicit な表現は与えていない。)しかし、今日の心理学、生理学、脳理論にとって、面の情報がどのように脳内で表現されているのかということは未解決の大問題なのである。本モデルで扱った分節化過程は、面の境界形成にかかわることであり、将来の理論において面の情報のグローバル・フィードバックを取り入れるにあたって、そのフィードバックがどの範囲でおこなわれるべきかを定めるための不可欠な基礎を提供するであろう。

§ 参考文献

- (1) Marr D (1982) *Vision*. Freeman, New York
- (2) Neisser U (1967) *Cognitive Psychology*. Appleton-Century-Crofts, New York
- (3) Julesz B (1971) *Foundations of cyclopean perception*. University of Chicago Press, Chicago
- (4) Nakayama K, Shimojo S, Silverman GH (1989) Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18:55-68
- (5) Poggio GF, Gonzalez F, Krause F (1988) Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J Neurosci* 8:4531-4550
- (6) Shimizu H, Yamaguchi Y (1987) Synergetic computers and holonics- information dynamics of a semantic computer. *Physica Scripta* 36:970-985
- (7) Engel AK, Konig P, Singer W (1991) Direct physiological evidence for scene segmentation by temporal coding. *Proc Natl Acad Sci USA* 88:9136-9140
- (8) LeVay S, Voigt T (1988) Ocular dominance and disparity coding in cat visual cortex. *Visual Neurosci* 1:395-414
- (9) Hirsch JA, Gilbert CD (1991) Synaptic physiology of horizontal connections in the cat's visual cortex. *J Neurosci* 11:1800-1809
- (10) Marr D, Poggio T (1976) Cooperative computation of stereo disparity. *Science* 194:283-287
- (11) Amari S, Arbib MA (1977) Competition and cooperation in neural nets. In: Metzler J (ed) *System neuroscience*. Academic Press, New York, pp 119-165
- (12) Malsburg C von der (1985) Nervous structures with dynamical links. *Ber Bunsenges Phys Chem* 89:703-710
- (13) Sporns O, Tononi G, Edelman GM (1991) Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. *Proc Natl Acad Sci USA* 88:129-133
- (14) Koffka K (1935) *Principles of Gestalt psychology*. Routledge & Kegan Paul, London
- (15) Gilbert CD, Wiesel TN (1979) Morphology and intracortical projections of functionally characterised neurons in the cat visual cortex. *Nature* 280:120-125
- (16) Kammen DM, Holmes PJ, Koch C (1989) Cortical architecture and oscillations in neuronal networks: feedback versus local coupling. In: Cotterill RMJ (ed) *Models of brain function*. Cambridge Univ Press, Cambridge, pp 273-284
- (17) Murata T, Shimizu H (1992) Oscillatory binocular system and temporal segmentation of stereoscopic depth surfaces. Submitted to *Biol Cybern* (revised)