



Title	両眼視による奥行知覚の神経回路モデルに関する研究
Author(s)	渡部, 修
Citation	大阪大学, 1999, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.11501/3155486">https://doi.org/10.11501/3155486</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

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

The University of Osaka

両眼視による奥行知覚の  
神経回路モデルに関する研究

渡部 修

1999年1月

## 概要

本論文は両眼視による奥行知覚の神経回路モデルに関する研究をまとめたものである。視覚によって得られる奥行手がかりには様々なものが存在するが、両眼視から得られる両眼視差はこの中でも最も重要な手がかりの一つである。両眼視差とは、両眼網膜に投影された像のずれのことで、このずれの量から物体の奥行を知ることができる。しかし、視覚による外界の構造復元にはオクルージョン(遮蔽)という問題が存在する。両眼視では、オクルージョンによって、一方の眼からは見えているが、もう一方の眼からは手前の面に遮蔽されて見えなくなる、両眼間非対応点が生じる。従来のほとんどの両眼視のモデルでは、両眼視差のみが奥行手がかりとして用いられ、偽対応しか生成しない両眼間非対応点はノイズとして無視されてきた。この研究で特に重要な点は、オクルージョンによって生じる両眼間非対応点を、奥行手がかりとして用いる手法を提案している点である。

本論文では、まず、両眼視の情報処理の基礎となる、脳の両眼性細胞のモデルを提案している。このモデルは、アナログ AND 回路に基づいて構成されている。そして、ネオコグニトロン型の神経回路を用いて、網膜から V1 の両眼性細胞までの階層型神経回路モデルを提案している。

本論文では、次に、両眼間対応点および両眼間非対応点双方を奥行手がかりとして用いる両眼視のモデルを提案している。これは、V1 以降の初期視覚領野のモデルに対応する。入力像に両眼間非対応点が存在するとき、推定される面は、入力像に存在する両眼間非対応点を生じさせることができ可能な構造に制限されると考えられる。本論文では、両眼間非対応点がどのように奥行手がかりとして用いられ、何が計算されなくてはならないかについて論じ、拘束条件を提案している。そして、この理論を実現するための両眼視アルゴリズムを提案している。このアルゴリズムは、網膜からの入力データを両眼間対応点と両眼間非対応点の二種類に区別し、両眼間対応点のところでは両眼視差を用いて、両眼間非対応点のところではオクルージョンの幾何学的な制約を用いて、視差の推定を行なう。さらに、計算機シミュレーションを行ない、従来の両眼視のモデルでは扱えなかった、両眼間非対応点が奥行知覚に大きな役割を果たすステレオグラムも説明できることを示している。

# 目 次

<b>第1章 序論</b>	<b>7</b>
1.1 初期視覚研究の背景 . . . . .	7
1.1.1 初期視覚のモジュール性 . . . . .	7
1.1.2 初期視覚の不良設定性と計算理論 . . . . .	10
1.1.3 標準正則化理論 . . . . .	16
1.1.4 MRF とラインプロセス . . . . .	19
1.1.5 視覚大脳皮質の計算理論 . . . . .	22
1.2 オクルージョン研究の背景 . . . . .	26
1.2.1 初期視覚におけるオクルージョンの影響 . . . . .	26
1.2.2 両眼間非対応点の心理学的知見 . . . . .	27
1.2.3 両眼間非対応点を扱うアルゴリズム . . . . .	32
1.3 本研究の目的 . . . . .	35
1.4 本論文の構成と概要 . . . . .	37
<b>第2章 両眼性細胞のモデル</b>	<b>41</b>
2.1 両眼性細胞の種類 . . . . .	41
2.2 エネルギーモデル . . . . .	43
2.3 アナログ AND モデルの基本構成 . . . . .	45
2.3.1 tuned cell の基本回路 . . . . .	45
2.3.2 near cell, far cell の基本回路 . . . . .	47
2.4 視差検出細胞の階層型神経回路モデル . . . . .	49
<b>第3章 両眼視におけるオクルージョンの拘束条件</b>	<b>57</b>
3.1 オクルージョンの拘束条件と奥行拘束領域 . . . . .	57
3.2 非対応点の視差推定 . . . . .	61
3.3 特殊なステレオグラム . . . . .	67
3.3.1 Narrow Occlusion . . . . .	67
3.3.2 Anderson のステレオグラム . . . . .	70
3.4 オクルージョン以外の原因で生じる非対応点 . . . . .	71

<b>第4章 オクルージョン手がかりを用いた両眼視アルゴリズム</b>	<b>75</b>
4.1 オクルージョンの計算理論の枠組	75
4.2 両眼視アルゴリズム	78
4.2.1 入力画像の前処理	80
4.2.2 面の構造復元	83
4.2.3 オクルージョン手がかりを与える領域の検出	88
4.2.4 エネルギー関数による表現	91
4.3 計算機シミュレーション	92
4.3.1 Wheatstone ステレオグラム	92
4.3.2 da Vinci ステレオグラム	94
4.3.3 Panum's limiting case	94
4.3.4 自然画像	99
<b>第5章 結論</b>	<b>101</b>

# 第1章 序論

この章では、まず 1.1 節で、本論文の主題である両眼立体視の研究を中心に、これまで行なわれてきた初期視覚の理論的な研究を概説する。続いて、1.2 節で、オクルージョン(遮蔽)が存在するときに、視覚系でどのような処理が行なわれているかを示唆する心理物理学実験について概説する。第 1.3 節で本論文の研究の目的を述べる。ここで、視覚系におけるオクルージョンの処理の重要性と、これまでの両眼視アルゴリズムが抱えていたオクルージョンが存在する状況での問題点について考察する。最後に、1.4 節で本論文の構成と概要を述べる。

## 1.1 初期視覚研究の背景

### 1.1.1 初期視覚のモジュール性

David Marr は 1982 年、その著書 “Vision” [53] で、脳内での視覚情報処理を初期視覚、中間視覚、そして高次視覚という三つの段階に大別した。ここで初期視覚(early vision)とは、2 次元の網膜像中に存在する様々な奥行き手がかり、例えば両眼視差(binocular disparity)<sup>1</sup>、運動視差(motion parallax)、陰影(shading)、遮蔽輪郭(occluding contour)、テクスチャー(texture)などから、3 次元の外界の構造を推定することを目的とする、視覚情報処理の初期段階を指している。

初期視覚の各奥行き手がかりの中でも、両眼視差による奥行き知覚の研究は非常に昔から活発に行なわれてきた。両眼視差とは外界のある点が左右眼網膜に投射されるときの「ずれ」の量のことである。三角測量と同様の原理で視差から奥行きを計算できる。つまり視差が奥行き手がかりとなり得るということは、理論的には 17 世紀の Kepler 以降知られてきた。図 1.1 は、両眼視(stereopsis)の幾何学を説明するための模式図である。簡単のため、仰角や回旋角は考慮しない。いま、点 F

---

<sup>1</sup> 心理学の分野では、binocular parallax を両眼視差、binocular disparity を両眼非対応と訳し、厳密に区別されている。binocular parallax は式(1.1)で表される量のことであり、binocular disparity は中心窓から等距離、等方向にない対応組、即ち binocular parallax が 0 でない対応組のことである。本論文では、どちらも区別なく両眼視差と訳し、式(1.1)の量を表すとする。また後述のように、本論文では、「非対応」は片眼からは見えるがもう片眼からは見えないという、オクルージョンによって生じる点を表すこととする。

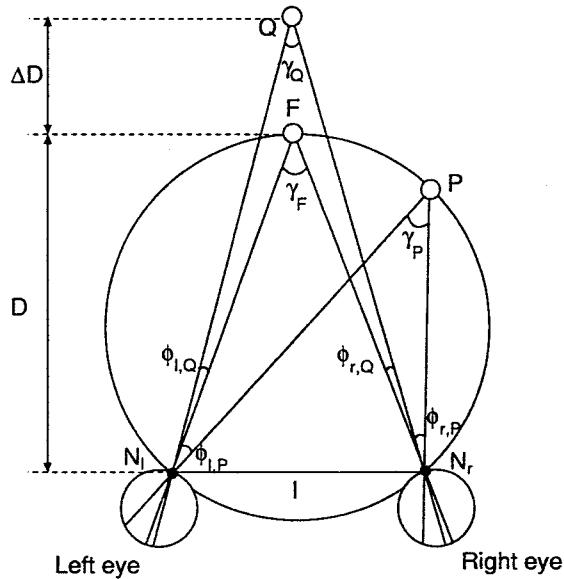


図 1.1: 両眼視の幾何学を説明する模式図. F は注視点,  $N_l$ ,  $N_r$  はそれぞれ左右眼の結節点(レンズ中心)を表す.

を注視しているとする. この点を通る両眼の視線のなす角  $\gamma_F$  を輻輳角 (vergence angle) と呼ぶ. さらに, 注視点 F 以外の点を通る両眼の視線のなす角  $\gamma$  を考えると, この点の視差は,

$$d = \gamma - \gamma_F = \phi_r - \phi_l \quad (1.1)$$

で定義される. 注視点 F と両眼の結節点(レンズ中心)  $N_l$ ,  $N_r$  を通る円周上の点 P は,  $\gamma_F = \gamma_P$  となるため視差 0 となる. この視差 0 を与える円を Vieth-Müller 円と呼ぶ<sup>2</sup>. Vieth-Müller 円の外側の点 Q は,  $\gamma_Q < \gamma_F$  であるため, 負の視差を与える. Vieth-Müller 円の外側(注視点より遠く)の点の持つ負の視差を非交差視差, 内側(注視点より近く)の点の持つ正の視差を交差視差と呼ぶ. 視差と奥行との関係式は, 厳密に記述すると複雑なものになるが ([15, 48, 51, 59, 91, 102, 108] 等参照), 注視点までの距離  $D$  が, 両眼間距離  $I$ , 及び  $D$  と求めたい点までの距離との差  $\Delta D$  に比べて十分長い(即ち  $D \gg I, \Delta D$ ) と仮定すると, 近似的に

$$d = \gamma - \gamma_F \approx \frac{I}{D + \Delta D} - \frac{I}{D} \approx -I \frac{\Delta D}{D^2} \quad (1.2)$$

<sup>2</sup>仰角も考慮すると, 視差 0 の点を与える「面」を考えることができる. この面はホロプター(Horopter)と呼ばれる.

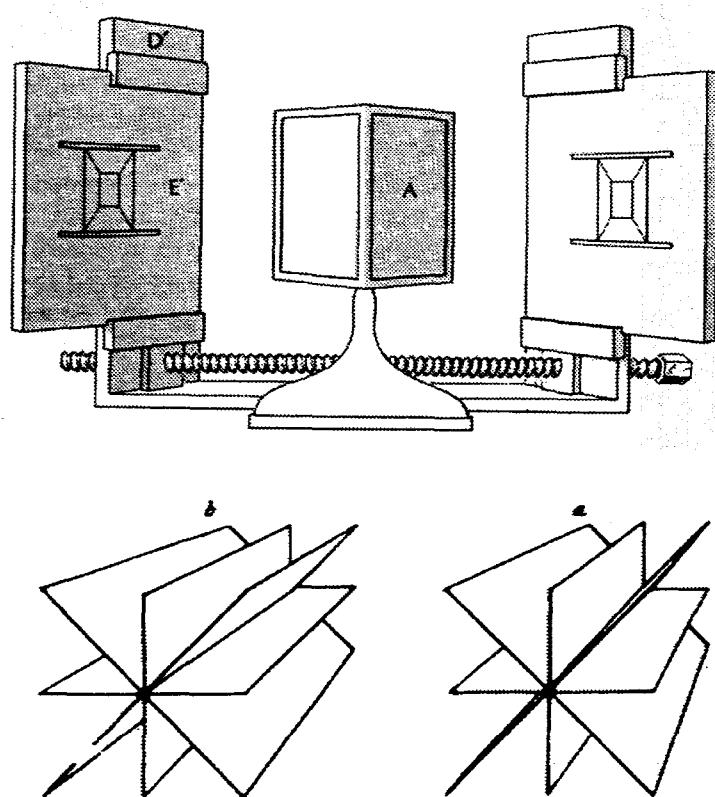


図 1.2: Wheatstone のステレオスコープ (上図) とステレオグラム (下図). 下図の左側の図が左眼画像, 右側の図が右眼画像を表す. (Howard & Rogers, 1995 [31] を改変して転載)

という関係式が得られる ([38, 90] 等参照). ここで,  $I$  は既知とみなせ,  $\Delta D$  は眼球の位置情報から計算可能である<sup>3</sup>. (1.2) 式は, 両眼視差の逆 2 乗法則として知られる式である.

両眼視差が視覚系で実際に奥行手がかりとして用いられているということは, 19世紀に Wheatstone によって実験的に示された [107]. Wheatstone は, 現在よく知られているステレオグラム (stereogram) と, 左右眼に独立な像を同時に提示できるステレオスコープ (stereoscopoe) を考案し, これによって両眼視の体系的な研究が促進された (図 1.2).

1960 年になって, Bela Julesz は図 1.3 のようなランダムドットステレオグラム (random dot stereogram) を考案した [36]. 図 1.2 のようなステレオグラムでは, 両

<sup>3</sup> $D$  は, 眼球の焦点調節からも知ることができる [68].

眼視差以外の様々な情報が混在しているため、両眼立体視がどのような機構で行なわれているのか明確にならない。例えば、左右眼それぞれの画像において、まずパターン認識の処理を行なってから、次にその認識結果を基に両眼の対応組を決定するという機構を考えることもできる [71, 72]。この場合は、単眼で見えない物は両眼でも見えないということになる。図1.3のランダムドットステレオグラムには、両眼視差以外の情報は存在しないため、両眼立体視の機構をより明確に知ることができる。図1.3を単眼で見るとランダムなドットパターンしか見えないが、融合視すると幾何学的な图形(図1.3の下の图形)が浮き上がって知覚できる。従って、まずパターン認識を行なうという前述の機構は却下される。さらに、ランダムドットステレオグラムは両眼視差しか手がかりを持たないので、両眼視差はそれのみで奥行を知覚させることができるということも示している。

ランダムドットステレオグラムによって、両眼視差手がかりのみで奥行知覚が可能であるということが証明された。そして、このことによって、両眼立体視機能のみを狙いとした理論的な研究が可能になった。初期視覚における他の奥行手がかりについても、単独で奥行を知覚させることができると考えられている。このことから、初期視覚は、各手がかりから単独で3次元の外界の構造を推定できるモジュール(module)が並列・独立に存在する、モジュール構造になっていると考えられている。ある奥行手がかりXのモジュールで行なわれる計算を、Xからの形状復元(shape-from-X)と呼ぶ。

### 1.1.2 初期視覚の不良設定性と計算理論

現在の初期視覚の各モジュールの理論的な研究は、Marrらの一連の研究 [52, 53, 55, 56] の影響を強く受けている。もちろん、Marr以前にも初期視覚のモデルの研究は数多くなされている。しかし、Marrらの研究において特筆すべきなのは、(視覚)情報処理課題の理解のための“三つの水準”的思想が根底にあるという点である(図1.4)。この内の第一の水準が、計算理論(computational theory)である。この水準では、情報処理機械の動作が、ある情報から別の情報への写像として特徴づけられる。そして、この写像の抽象的な性質が定義され、問題に対してその理論が適切かつ十分であることが示される。具体的には、(1)何が計算されるか、そして(2)計算の過程においてなぜそのような計算方法を用いるのか、という二つの点が明らかにされる。何が計算されるかという点については、初期視覚の計算理論では、外界の3次元構造を示すパラメータ、例えば相対的な奥行きや面の局所的な表面方向などであるといえる。図1.5は、初期視覚の各モジュールで計算されるパラメータを示したものである。二つめの、なぜそのような計算方法を用いるのかという点については、単にそうすれば正しい解が導き出せるからというだけで

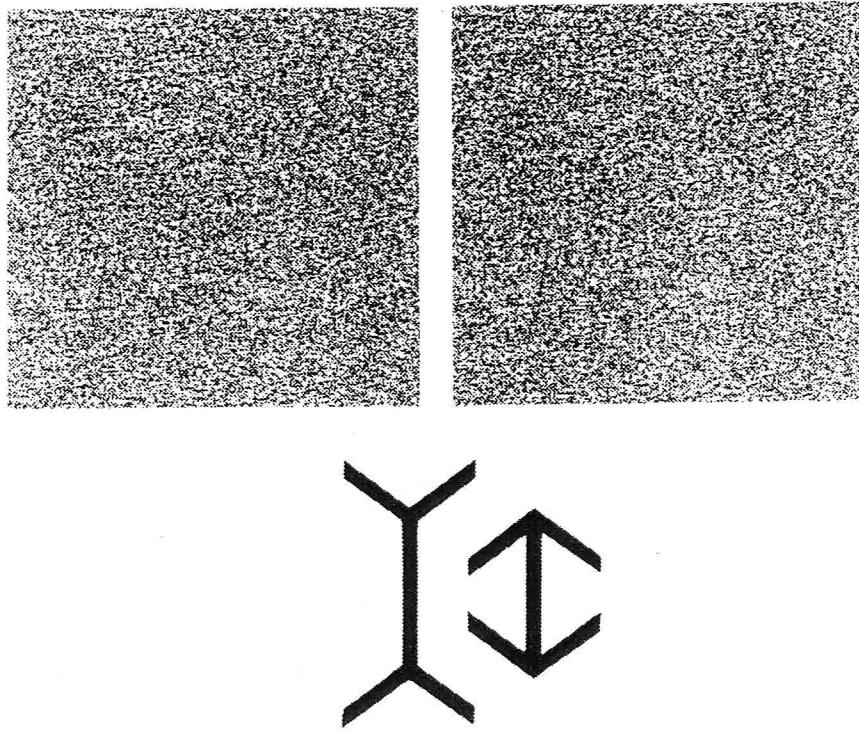


図 1.3: ランダムドットステレオグラム。上のステレオグラムの、左側の図が左眼画像、右側の図が右眼画像である。下の図は、融合視すると浮き上がって見える図形 (Müller-Lyer の錯視図形) の形状を表している。(Julesz, 1971 [37] を改変して転載)

計算理論 (computational theory)	計算の目的は何か、なぜそれが適切なのか、そしてその実行可能な方略の論理は何か。
表現とアルゴリズム (representation and algorithm)	この計算理論はどのようにして実現することができるか。特に入力と出力の表現は何か、そして変換のためのアルゴリズムは何か。
ハード ウェアによる実現 (hardware implementation)	表現とアルゴリズムがどのようにして物理的に実現されるか。

図 1.4: 情報処理課題を実行する機械を理解するのに必要な三つの水準 [53].

処理過程	自然な出力形式
立体視	視差, よって $\delta r, \nabla r, s$
運動方向の選択性	$\nabla r$
運動からの構造復元	$r, \delta r, \nabla r, s$
オプティカルフロー	? , $r, s$
遮蔽輪郭	$\nabla r$
他の遮蔽の手がかり	$\nabla r$
表面方向輪郭	$\nabla s$
表面輪郭	$s$
表面テクスチャー	おそらく $r$
テクスチャー輪郭	$\nabla r, s$
陰影	$\delta s, \nabla s$

図 1.5: 初期視覚の各モジュールの出力形式 [53].  $r$  = (正射影における) 相対的奥行き.  $\delta r = r$  における連続的あるいは小さな局所変化.  $\nabla r = r$  における不連続.  $s$  = 局所的な表面方向.  $\delta s = s$  における連続的あるいは小さな局所変化.  $\nabla s = s$  における不連続.

はなく、なぜそのようにすればうまくいくのかという疑問にも答える必要がある. 導出された解は、外界の様々な物理的な条件を満足させるものでなくてはならない. 従って、外界の物理的な条件を、計算のための拘束条件<sup>4</sup>(constraint)として用いる必要がある.

本論文の主題である、両眼視モジュールについて考えてみる. 両眼視では、まず左右二つの画像の各点の対応を見つけ、その対応する点の「ずれ」の量、即ち両眼視差から、物体の奥行きを知ることができる. よって、両眼視モジュールへの入力は両眼画像、出力は視差でよいということになる. しかし、一般に両眼視では、片眼の画像上のある点がもう片眼の画像上のどの点と対応するのか一意に決定できないという、対応問題(corresponding problem)が存在する. 例えば、左眼画像上のある点に対して、右眼画像のエピポーラ線(epipolar line)上の全ての点が対応候補となるため、多くの偽対応(false match)が生じる(図 1.6 参照). 何か付加的な条件がなければ、偽対応を含むこの複数の対応候補の中から正しい対応を一つ選び出すことはできない. 従って両眼視は、解が一意に決定できない不良設定問題<sup>5</sup>(ill-posed problem)になっているといえる. 初期視覚の各モジュールの

<sup>4</sup> または制約条件と訳す.

<sup>5</sup> (1) ある問題の解が存在し、(2) 解が一意で、(3) 解がデータに連続に依存する、という三つの条件のうち一つでも満たされないような問題を不良設定問題という. この三つの条件が全て満たさ

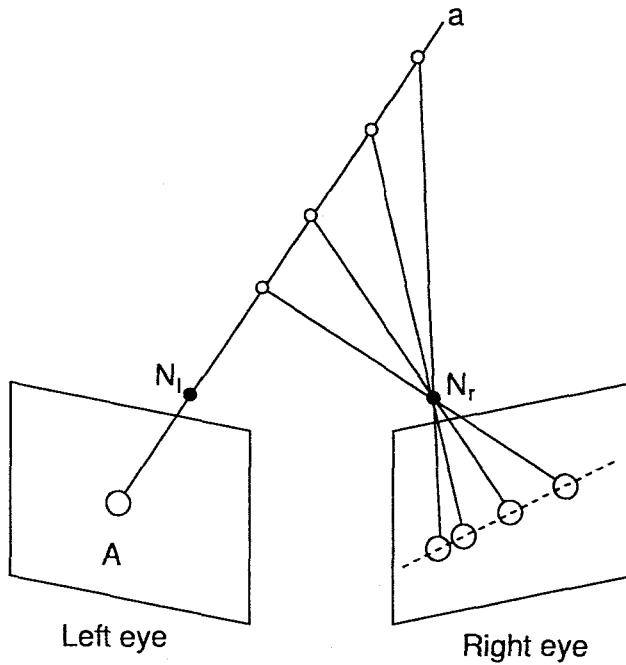


図 1.6: 左眼画像上の点 A に投射される 3 次元空間中の点は、A を通る半直線 a 上に存在する。この半直線 a 上の点を投射してできる右眼画像上の破線をエピポーラ線という [21]。従って、点 A に対応する点はこのエピポーラ線上の点に限定される。しかし、この対応候補の中から正しい対応を決定するには、更に付加的な条件が必要となる。

問題は、全て不良設定問題になっているといえる。従って、初期視覚の計算理論の研究は、問題を良設定にする適切な拘束条件を導き出すことであるといえる。

Marr と Poggio は 1976 年に、両眼視の計算理論、およびアルゴリズムを提案している [55]。彼らは、不良設定問題である両眼視の計算で、一意かつ物理的に妥当な解を導き出すため、以下のような三つの拘束条件を提案した:

- **適合性の拘束条件**(compatibility constraint): 黒点は黒点と、白点は白点とのみ対応し得る。
- **一意性の拘束条件**(uniqueness constraint): 片眼画像上の一つの点は、もう片眼の画像上の唯一の点とのみ対応する。
- **連続性の拘束条件**(continuity constraint)<sup>6</sup>: 点の視差は滑らかに変化する。

れている問題は良設定問題 (well-posed problem) と呼ばれる。

<sup>6</sup> または滑らかさの拘束条件(smoothness constraint)と呼ばれる。

第一の拘束条件は、適合性の拘束条件である。この拘束条件は、二値画像しか考慮していないわけではない。例えば、グレースケール画像のときは、対応要素 (matching primitive) として Laplacian-Gaussian  $\nabla^2 G$  のゼロ交差<sup>7</sup> (zero-crossing) を用いて、同じ符合のゼロ交差とのみ対応しうるとして実現できる。一意性の拘束条件は、画像上の点が、3次元空間中の両眼から見える不透明面から投射されている場合に常に成り立つ、物理的な拘束条件である。連續性の拘束条件は、面の奥行きは滑らかに変化するという仮定から導かれる。これらの拘束条件を用いることによって、不良設定問題である両眼視による構造復元問題から、一意かつ物理的に妥当な解を導くことができる。

同時に、MarrとPoggioは、この計算理論を実現するアルゴリズムとして協調的アルゴリズム (cooperative algorithm) と呼ばれるアルゴリズムを提案している [55]。このアルゴリズムは、特定の位置  $(x, y)$  に特定の視差  $d$  を持った両眼入力があったときに発火する両眼性細胞  $C_{x,y,d}$  が、相互結合した神経回路で構成されている (図 1.7)<sup>8</sup>。このアルゴリズムでは、一意性の拘束条件は左右眼の視線に沿って存在する細胞間の競合で、連續性の拘束条件は近傍の同じ視差に反応する細胞間の興奮性結合で実現されている。このアルゴリズムは次式で表される。

$$C_{x,y,d}^{t+1} = \sigma \left\{ \sum_{x',y',d' \in S(x,y,d)} C_{x',y',d'}^t - \epsilon \sum_{x',y',d' \in O(x,y,d)} C_{x',y',d'}^t + C_{x,y,d}^0 \right\} \quad (1.3)$$

ここで、 $C_{x,y,d}^t$  は、位置  $(x, y)$ 、視差  $d$  に対応する細胞の、時刻  $t$  での状態を表す。 $S(x, y, d)$  は局所的興奮性近傍、 $O(x, y, d)$  は抑制性近傍を表す。 $\epsilon$  は抑制性入力の強さを規定する定数、 $\sigma$  は閾値関数である。 $C^0$  は初期状態であり、定められた視差範囲内にある、偽対応を含む全ての対応候補を表している。

この、MarrとPoggioの協調的アルゴリズム以前にも、同様の両眼視アルゴリズムは提案されていた。しかし、Marrのいう計算理論の解析が伴ったものはほとんどなく、結果として拘束条件が不足していたり誤って実現されていたりした<sup>9</sup>。例

<sup>7</sup>Gauss 関数の Laplacian をとったフィルタを Laplacian-Gaussian フィルタ (Laplacian-Gaussian filter) と呼ぶ。これは次式で表される:

$$\nabla^2 G(r) = -\frac{1}{\pi\sigma^4} \left( 1 - \frac{r^2}{2\sigma^2} \right) e^{-\frac{r^2}{2\sigma^2}}$$

このフィルタで画像をフィルタリングし、値がゼロになる点をゼロ交差点と呼ぶ。これは非常に性質の良いエッジ抽出機構になっており [54, 57]。またこの点を両眼視問題の対応要素とすることで効率的な両眼視アルゴリズムが得られる (詳しくは [56])。

<sup>8</sup>Marr-Poggio の協調的アルゴリズムでは、両眼の視線が並行であると近似して (つまり無限遠点を注視していると仮定して) 垂直視差 (vertical disparity) を考慮していない。これ以降紹介する両眼視アルゴリズム、及び本論文で提案するアルゴリズムも同様である。垂直視差も考慮したアルゴリズムについては、[48, 51, 59, 91, 108] 等を参照

<sup>9</sup>計算論的に妥当なモデルが全く存在しなかった訳ではない。例えば、平井と福島 [26-28] は、

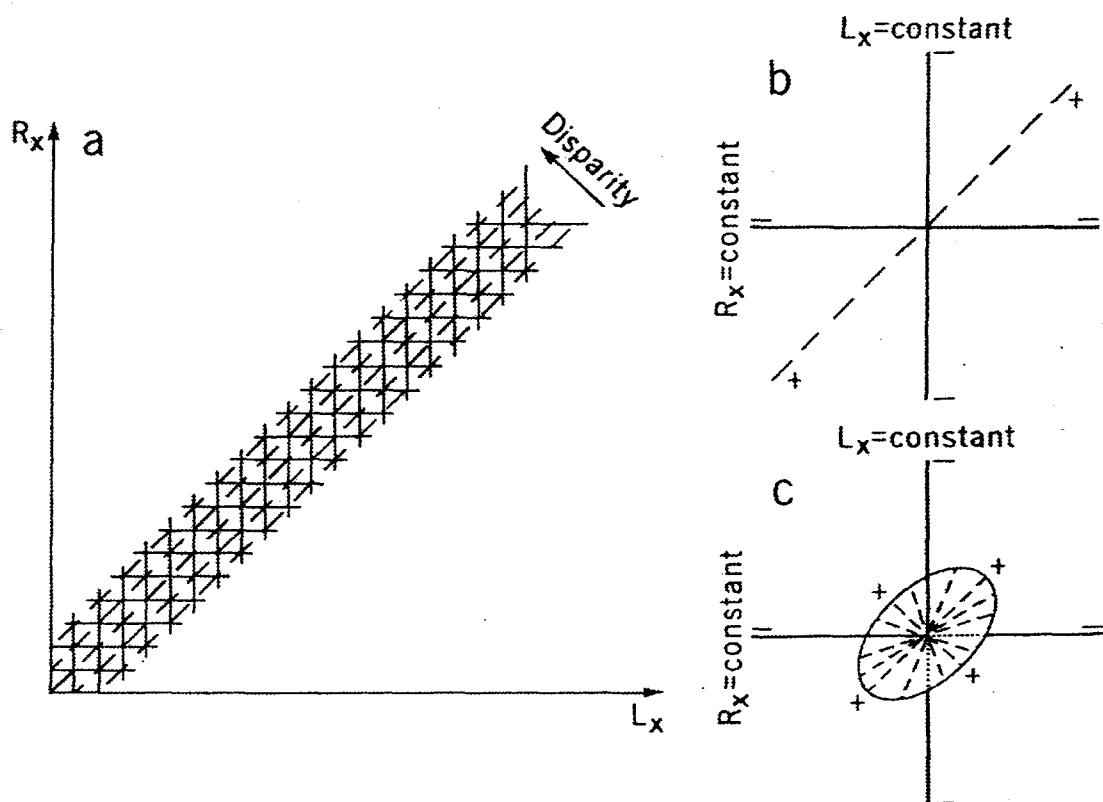


図 1.7: Marr-Poggio の協調的アルゴリズムの模式図. (a)  $L_x$  と  $R_x$  は左右の網膜像の位置を表す. このような図を Keplerian ダイアグラム (Keplerian diagram) と呼ぶ. 垂直及び水平の実線はそれぞれ左右眼からの視線を表す. 斜めの破線は視差一定の線である. Marr と Poggio の協調的アルゴリズムでは, 両眼性細胞が各節点に位置している. 実線が抑制性の相互作用を表し, 点線が興奮性の相互作用を表す. (b) 各節点における局所構造. (c) 2次元に拡張したときの各節点の局所構造. この図における橢円形は, このページの面と直行する2次元の円盤を表す. (Marr & Poggio, 1976 [55] より転載)

えば、1975年に相次いで発表された Dev [10] や Nelson [69] の両眼視アルゴリズムは、物理的に妥当な解を導き出すものではなかった。具体的にいえば、一意性の拘束条件による競合の代わりに、サイクロプスの眼(Cyclopean eye)<sup>10</sup>の視線に沿って対応候補の競合を行なっていた。Marrの功績は、初期視覚モジュールの計算論的な解析の重要性と、外界の物理的な知識を拘束条件として入れることの必要性を示したことにある。

### 1.1.3 標準正則化理論

1.1.2節で触れたように、一般に初期視覚の問題は解が一意に定まらない不良設定問題になっている。Marrの思想は、こうした問題に対して、拘束条件を置くことによって物理的に妥当な唯一の解を導く、というものであった。こうした初期視覚の研究の流れを受けて、Poggio, Torre と Kock は 1985 年、初期視覚を逆問題(inverse problem)として捕え、初期視覚の計算理論の一般的な枠組として標準正則化理論(standard regularization theory)を用いることを提案した [77]( [5, 58] も参照)。

いま、外界の3次元構造を  $z$ 、その網膜に投射された2次元画像を  $y$  とする。 $z$  から  $y$  を求める光学の問題、すなわち順方向の問題は一意に解け、線形オペレータ  $A$  を用いて

$$Az = y \quad (1.4)$$

と書き表すことができる。初期視覚の目的は、与えられた網膜像  $y$  から外界の構造  $z$  を求める逆光学の問題を解くことである。しかし、一般に初期視覚は不良設定問題になっており、一意に解を求ることはできない。 $A$  の逆行列が求められず、 $y$  から  $z$  を

$$y = A^{-1}z \quad (1.5)$$

と一意に求められない場合、標準正則化理論では安定化汎関数(stabilizing functional)  $\|Qz\|^2$  を用いて

$$\|Az - y\|^2 + \lambda \|Qz\|^2 \quad (1.6)$$

一意性及び連続性の拘束条件の代わりに、順序対応の拘束条件(ordering constraint)を用いた両眼視アルゴリズムを提案している。順序対応の拘束条件とは、「対応の結果は入力助順序を変えない」とする拘束条件で、視差の変化に比べて面が十分大きいときに成立つ。この拘束条件は、サイクロピアン座標での視差勾配(disparity gradient)が 2 以上の対応候補と競合することによって実現でき、一意性の拘束条件を含んでいる。このモデルは、一意で物理的に妥当な解を導き出せる。

<sup>10</sup>左右眼間の中心に、仮想的に考えられる眼のこと。左右眼網膜上の対応組の位置をそれぞれ  $x_l, x_r$  とすると、サイクロピアン座標での位置は  $(x_l + x_r)/2$  と表される。

を最小にする  $z$  を求め、これを解とみなす。第1項は解  $z$  の入力データ  $y$  に対するデータフィッティング項(data-fitting term)で、第2項は拘束条件に相当する。ここで、 $\lambda$  は拘束条件の強さを決めるパラメータ、 $Q$  は線形オペレータである。

この標準正則化理論で定式化される初期視覚の諸問題をまとめたものが図1.8である。この図で示される各式は(1.6)式に相当し、エネルギー(energy)と呼ばれる<sup>11</sup>。各式において、第1項は与えられた画像データと推定された解との、データフィッティング項になっている。第2項は拘束条件であり、全て解の滑らかさの拘束条件(smoothness constraint)になっている。標準正則化理論により、初期視覚の問題はエネルギー最小化問題として扱うことができる。

標準正則化理論で、両眼視の問題がどのように定式化されるか見てみよう。図1.8を見ると、両眼視におけるエネルギー関数は次式で与えられている。

$$E = \int \left\{ [\nabla^2 G * (L(x, y) - R(x + d(x, y), y))]^2 + \lambda (\nabla d)^2 \right\} dx dy \quad (1.7)$$

(1.7)式 {} 内の第1項が、データフィッティング項に対応する。この項は、左右眼網膜像に Laplacian-Gaussian フィルタをかけた値の差が、最も小さくなる視差  $d(x, y)$  を求めることを示している。第2項は、滑らかさの拘束条件を実現するスムースネス項(smoothness term)に対応する。この項は視差の1階微分になっており、視差勾配が小さな解を好むことを示している。 $\lambda$  は拘束条件の強さを決めるパラメータである。この式  $E$  を最小にする視差  $d$  が求める解である。(1.1)式と異なり、この式では  $d < 0$  が交差視差、 $d > 0$  が非交差視差を表す。

(1.7)式は、両眼視のアルゴリズムとして提案された Marr-Poggio の第2アルゴリズム [56] を発展させた、Nishihara のモデル [70] の形式的な表現と見なすことができる。一方、このエネルギー関数から、最急降下法を用いて、直ちに問題の解を求めるアルゴリズムを導出することもできる<sup>12</sup>。(1.7)式のエネルギー関数  $E$  を最小にする視差  $d$  が求める解であるので、各視差  $d(x, y)$  を  $E$  を小さくする方向に変化させていくようなアルゴリズムであればよい。従って、各点の視差の時間変化を

$$\frac{\partial d(x, y)}{\partial t} = -\frac{\partial E}{\partial d(x, y)} \quad (1.8)$$

とする繰返しアルゴリズムが考えられる。いま、(1.7)式の第1項を Taylor 展開し

<sup>11</sup> またはコスト関数(cost function)と呼ばれる。

<sup>12</sup> 陰影からの形状復元、及び両眼視のエネルギー関数は2次形式にならないため、ここで説明する方法ではエネルギーの極小値に落ち込み、大域的な最適解が得られないことがある。

問題	正則化原理
エッジ検出	$\int [(S \cdot f - i)^2 + \lambda(f_{xx})^2] dx$
オプティカルフロー (領域)	$\int [(i_x u + i_y v + i_t)^2 + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2)] dx dy$
オプティカルフロー (輪郭)	$\int [(v \cdot N - v^N)^2 + \lambda(\frac{\partial v}{\partial s})^2] ds$
表面再構成	$\int [(S \cdot f - d)^2 + \lambda(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)] dx dy$
時空間近似	$\int [(S \cdot f - i)^2 + \lambda(\nabla f \cdot V + f_t)^2] dx dy dt$
色	$\ I^\nu - Az\ ^2 + \lambda\ Pz\ ^2$
陰影からの構造復元	$\int [(E - R(f, g))^2 + \lambda(f_x^2 + f_y^2 + g_x^2 + g_y^2)] dx dy$
両眼視	$\int \{[\nabla^2 G * (L(x, y) - R(x + d(x, y), y))]^2 + \lambda(\nabla d)^2\} dx dy$

図 1.8: 標準正則化理論によって定式化される初期視覚問題 [77]. 下付の添字は偏微分を表す. 最初のエッジ検出の問題は, 簡単のため 1 次元としている. 離散的な画像強度  $i$  が与えられた時,  $i$  を良く表現し, また 2 項目の拘束条件をよく満たす連続関数  $f$  が推定される.  $S$  はサンプリング作用素である. 2 番目の領域に基づくオプティカルフローの問題では, 画像強度  $i$  が与えられた時, 各点の速度場  $(u, v)$  を推定する. 3 番目の輪郭に基づくオプティカルフローの問題では, 輪郭の法線方向の速度成分  $v^N$  が与えられた時, 真の速度  $v$  を推定する. ここで,  $N$  は輪郭の法線方向の単位ベクトルである. 4 番目の表面再構成の問題では, 離散的な奥行きデータ  $d$  が与えられた時, これを良く表現する連続関数  $f$  が推定される. 5 番目の時空間近似の問題では, 離散的な画像強度  $i$  が与えられた時, これを良く表現する連続関数  $f$  が推定される. 6 番目の色の問題では, 三つの適当な色座標での明るさ  $I^\nu$  ( $\nu = 1, 2, 3$ ) が与えられた時, 照明条件と反射率の要素を含むベクトル  $z$  を推定する.  $A$  は  $z$  から  $I^\nu$  への写像を表す. 第 2 項は, 照明条件が滑らかに変化することと, 反射率が一定値をとるかまたは鋭く変化するかどちらかであるという拘束条件を表す. 以下の二つは, 2 次形式でない正則化汎関数の例である. 7 番目の陰影からの構造復元問題では, 明るさの分布  $E$  が与えられた時, 面の傾き (例えば方位角と傾斜角)  $f, g$  を推定する. ここで,  $R(f, g)$  は面の傾き  $(f, g)$  から明るさを決定する, 非線形な反射率関数である. 8 番目の両眼視の問題では, 左右眼画像  $L(x, y), R(x, y)$  が与えられた時, 各点の視差  $d(x, y)$  を推定する.

て2次以上の項を無視し、第2項を1次差分に置き換えると、

$$\begin{aligned} E = & \int \left[ \nabla^2 G * \left( L(x, y) - R(x, y) - d(x, y) \frac{\partial R(x, y)}{\partial x} \right) \right]^2 dx dy \\ & + \lambda \sum_{x, y} [(d(x, y) - d(x - 1, y))^2 + (d(x, y) - d(x, y - 1))^2] \end{aligned} \quad (1.9)$$

と表せる。 (1.8) 式に (1.9) 式を代入すると、

$$\begin{aligned} \frac{\partial d(x, y)}{\partial t} = & 2 \nabla^2 G * \left( L(x, y) - R(x, y) - d(x, y) \frac{\partial R(x, y)}{\partial x} \right) \cdot \nabla^2 G * \frac{\partial R(x, y)}{\partial x} \\ & + 2\lambda [d(x + 1, y) + d(x - 1, y) + d(x, y + 1) + d(x, y - 1) - 4d(x, y)] \end{aligned} \quad (1.10)$$

というアルゴリズムが得られる。これは、各点  $d(x, y)$  において局所的な計算になっているため、脳のモデルとしても適している。

しかし、この標準正則化理論には限界がある。例えば外界に複数の物体がある場合、その物体間の境界付近では推定する解の真の値、例えば両眼視における視差の値などは大きく変化していると考えられる。標準正則化理論では、たとえ物体の境界に不連続があったとしても、滑らかさの拘束条件によって、不連続が平滑化された解に帰結してしまう。

#### 1.1.4 MRF とラインプロセス

Geman と Geman は 1984 年、MRF(Markov random field; マルコフ確率場)による画像復元アルゴリズムを提案した [18]。ノイズが加えられた画像を復元するという問題も、不良設定問題になっており、何か付加的な条件がない限り解が一意に定まらない。MRF も不良設定問題をエネルギー最小化問題に帰着させるが、エネルギーの形に関しては特別な仮定が置かれていません。Geman らの研究は、初期視覚の問題を確率を取り入れて統一的に扱うことを可能にしたこと、凸関数でないエネルギーの大域的最小値を見つける手法を提案したこと、そして 1.1.3 節の最後でも述べた不連続境界の問題を解決するためのラインプロセスという概念を導入したことで、大きなブレークスルーとなった。

MRF では、Bayes の定理を用いて事後確率を最大にする解を求めるという、MAP 推定(maximum a posteriori estimate; 最大事後確率推定)の方法がとられる。Bayes の定理(Bayes theorem)より、入力データを  $g$ 、推定すべき状態を  $f$  としたときの事後確率(posterior probability)  $\mathcal{P}(f|g)$  は次式で与えられる。

$$\mathcal{P}(f|g) \propto \mathcal{P}(g|f)\mathcal{P}(f) \quad (1.11)$$

ここで、 $\mathcal{P}(g|f)$  は  $f$  という条件のもとで  $g$  が生じる条件付確率(conditional probability)、 $\mathcal{P}(f)$  は  $f$  の事前確率(prior probability)である。事前確率  $\mathcal{P}(f)$  がわからないときはそれらが等確率であると仮定して計算することが多く、この場合は最尤推定と一致することが知られている。

MRFでは、各点の値がその近傍の値にのみ依存して決定されると考える(Markov性の仮定)。これは、初期視覚のモデルとしては妥当な仮定である。このとき、モデルがある状態をとる確率が Gibbs 分布

$$\mathcal{P}(f) = \frac{1}{Z} \exp \left\{ -\frac{1}{T} U(f) \right\} \quad (1.12)$$

で表される。ここで、 $Z$  は全ての状態  $f$  に対する確率の和が 1 になるための規格化定数で

$$Z = \int \exp \left\{ -\frac{1}{T} U(f) \right\} df \quad (1.13)$$

と表される。これは統計力学との対応から分配関数と呼ばれる。 $T$  は定数で、統計力学との対応から温度と呼ばれる。 $U(f)$  はエネルギーに対応する量で、各点の近傍系について定義されるポテンシャルエネルギーの総和になっている。従って、MAP 推定によって事後確率が最大になる状態  $f$  を求めるということは、事後確率のエネルギー  $U(f|g)$  の最小化問題に帰着することができる。そして、この計算は Markov 性から、局所的な並列演算によって実行できる。これは、視覚系のモデルに適した性質である。

前節で述べた標準正則化理論は、事前確率が 2 次形式の Gibbs 分布で与えられ、条件付確率が線形演算に加法的に正規白色雑音を加えたものとして与えられるという、2つの単純化をした MRF の特殊な場合とみなすことができる。つまり、入力データ(網膜像)を  $y$ 、推定すべき状態(外界の構造)を  $z$ 、ノイズの分散を  $\sigma$  とすると、事前確率および条件付確率は以下のようになる。

$$\mathcal{P}(z) = \frac{1}{Z} \exp \{-\|Qz\|\} \quad (1.14)$$

$$\mathcal{P}(y|z) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{\|Az - y\|^2}{2\sigma^2} \right\} \quad (1.15)$$

ここで、 $A, Q$  は線形オペレータで、 $A$  は光学に相当する。すると事後確率は、

$$\mathcal{P}(z|y) \propto \frac{1}{\sqrt{2\pi}\sigma Z} \exp \left\{ -\frac{\|Az - y\|^2 + 2\sigma^2\|Qz\|^2}{2\sigma^2} \right\} \quad (1.16)$$

となり、これを最大化することは指數関数の内部

$$\|Az - y\|^2 + \lambda\|Qz\|^2 \quad (1.17)$$

を最大化することと等価になる。これは、標準正則化理論のエネルギーの式(1.6)そのものである。そして  $\lambda = 2\sigma^2$  より、観測データのノイズが大きければデータフィッティング項より拘束条件の方を信頼するようになることがわかる。MRFは、標準正則化理論と異なり、エネルギー関数の形は制限されていない。ここではその解説は省略するが、非凸関数の最小化を行うために、Gemanらは確率的緩和法(stochastic relaxation)と呼ばれるアルゴリズムも提案している[18]。

Gemanらはまた、面の不連続を扱うため、ラインプロセス(line-process)が導入された2結合型MRF(2-coupled MRF)を提案している。この2結合型MRFは、平均場近似によって Hopfield型の神経回路モデル[29]で計算できることが知られている[16, 47, 89]。

いま、本稿の主題である両眼視を例にラインプロセスの説明を行う。簡単のため、以下のような1次元の場合のエネルギー関数を考える。

$$\begin{aligned} E = & \int [\nabla^2 G * (L(x) - R(x + d(x)))]^2 dx \\ & + \lambda \sum_x (d(x) - d(x - 1))^2 \end{aligned} \quad (1.18)$$

前節で述べた通り、このエネルギー関数を最小化するという方法では、視差に不連続がある場合でも面が滑らかに補間されてしまう。そこで、面の不連続性を表す変数として、ラインプロセス  $l(x)$  を導入する。 $l(x)$  は、隣り合う点  $x, x - 1$  が連続するとき  $l(x) = 0$ 、不連続のとき  $l(x) = 1$  となる変数である。そして、 $l(x) = 1$  のとき、隣り合う点  $d(x), d(x - 1)$  の間を滑らかにつなぐことを禁止する。このラインプロセスもまたMRFになっており、視差を検出する層と相互作用しながら、最も満足のいく解を出す。これは、次式のようなエネルギー関数を最小化することで実現できる。

$$\begin{aligned} E = & \int [\nabla^2 G * (L(x) - R(x + d(x)))]^2 dx \\ & + \lambda \sum_x (1 - l(x))(d(x) - d(x - 1))^2 \\ & + \mu V(l) \end{aligned} \quad (1.19)$$

ここで、 $V(l)$  はラインプロセスについてのエネルギー関数である。1次元の場合、不必要的不連続検出を回避するための付加的コストとして、 $V(l) = \sum_x l(x)$  とすればよい<sup>13</sup>。これは、視差勾配  $|d(x) - d(x - 1)|$  が閾値  $\sqrt{\mu/\lambda}$  より大きいときに

<sup>13</sup>他に、不連続が隣り合って存在するのを抑制する、 $\sum_x l(x)l(x - 1)$  というエネルギー関数も考えられるが、簡単のためここでは用いない。2次元の場合は、不連続境界の局所的なパターンによるポテンシャルも考慮する。

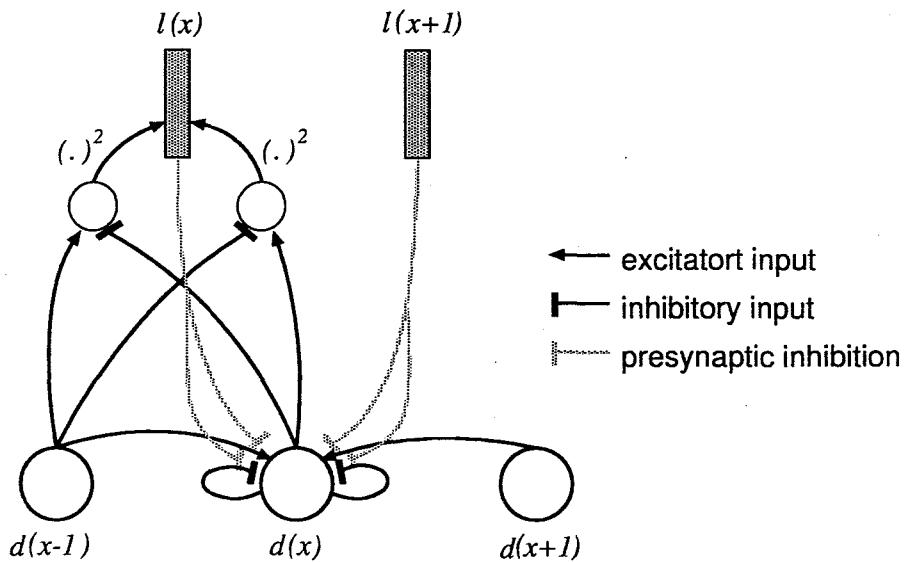


図 1.9: ラインプロセスを導入した両眼視の神経回路モデル.

不連続が検出されることを意味している。これを神経回路モデルで実現すると図 1.9 のようになる。

ラインプロセスと視差の二つの MRF の相互作用は、(1.19) 式のエネルギー関数によって考慮されていることになる。この相互作用は、拘束条件の実行をコントロールしている (constraint adaptation)。不連続境界は一般に画像から直接得られないものであり、2 結合型 MRF はこのような隠れ変数 (hidden variable) を扱うことを可能にする。

(1.19) 式のように、ラインプロセスを導入したエネルギー関数は 2 次形式になっていない。このような 2 次形式になっていないものも含めた計算理論を、一般に、正則化理論 (regularization theory) と呼ぶ。

### 1.1.5 視覚大脳皮質の計算理論

川人と乾は 1990 年、初期視覚、中間視覚、高次視覚の問題を視覚大脳皮質の構造と機能に基づいて統一的に解決する計算理論を提案した [43]([25, 41, 42] も参照)。この理論は視覚大脳皮質の計算理論 (computational theory of the visual cortical areas) と呼ばれている<sup>14</sup>。

<sup>14</sup> または、順逆光学モデル (forward-inverse optics model)、視覚大脳皮質の双方向性理論 (bi-directional theory) と呼ばれる。

彼らの理論の出発点は、Horn [30] の画像放射照度方程式 (image irradiance equation) を一般化したものである。網膜に与えられる 2 次元画像の生成過程をモデル化するときには、様々なレベルでの記述が可能である。低いレベルでは、面の奥行きや向き、各点の反射率、照明光を決めれば Horn の画像照度照射方程式で網膜像が決まる。より高いレベルでも同様に視覚世界を記述できる。3 次元空間の中に、個々の物体がどのように配置され、3 次元像は何で、どのような並進、回転速度を持つかを記述しても網膜像を決定できる。脳内では、これらの各レベルでの記述が全て用いられていると考えられる。これは次式で表される。

$$\begin{aligned} I(\mu, x, y, \lambda, t) &= R(\nabla^2 G * I, dI, d^2 I, v^\perp, sd, r(\lambda), L, md, \nu, C, A, V, N, O) \\ &= R(s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}, s_{13}, s_{14}) \\ &= R(S) \end{aligned} \quad (1.20)$$

この一般化された画像放射照度方程式は、左辺に網膜像  $I$  が、右辺に外界の属性または状態  $S$  が変数として表された方程式である。左辺の網膜像  $I$  は、左右の画像を表す  $\mu$ 、網膜座標  $x, y$ 、波長  $\lambda$ 、時間  $t$  が変数である。一方、右辺では、 $\nabla^2 G$  と網膜像  $I$  との重畠積  $\nabla^2 G * I$ 、網膜像  $I$  のある方向への 1 階及び 2 階微分  $dI$ 、 $d^2 I$ 、局所的な速度成分  $v^\perp$ 、両眼視によって得られる奥行  $sd$ 、波長  $\lambda$  の光に対する反射率  $r(\lambda)$ 、不連続境界  $L$ 、単眼手がかりによって得られる奥行  $md$ 、照明光の波長分布と光源位置  $\nu$ 、 $L$  で区別された個々の 3 次元物体の 3 次元的空間位置  $C$ 、個々の物体の色やテクスチャー  $A$ 、個々の物体の並進や回転の速度ベクトル  $V$ 、観察者の身体や、頭部、眼球の持つ並進や回転の速度ベクトル  $N$ 、3 次元物体の記憶像  $O$  が変数になっている。この両者を関係づける関数  $R$  は、いわば一般化された光学(optics)と呼ぶべきものである。

脳内では、網膜像  $I$  が与えられたとき、そこから右辺の各変数の値を決定しなければならない。従って、脳内では  $R$  の逆関数、即ち逆光学(inverse optics)を計算する必要がある。ここで、脳内ではこの逆光学の計算を極めて速く、おそらく 300ms 以内に計算せねばならない。また、さまざまな条件下で正常に機能するため、その計算はロバストでなければならない。

上述の条件を満足するため、この理論では低次領野から高次領野への順方向(feed-forward)の結合だけでなく、高次領野から低次領野への逆方向(feedback)の結合も用いている。そして、この双方向性(bi-directional)の結合によって、以下のような計算理論を考えている。

網膜像  $I$  から  $S$  の推定に MAP 推定を用い、次のエネルギーを最小化するとい

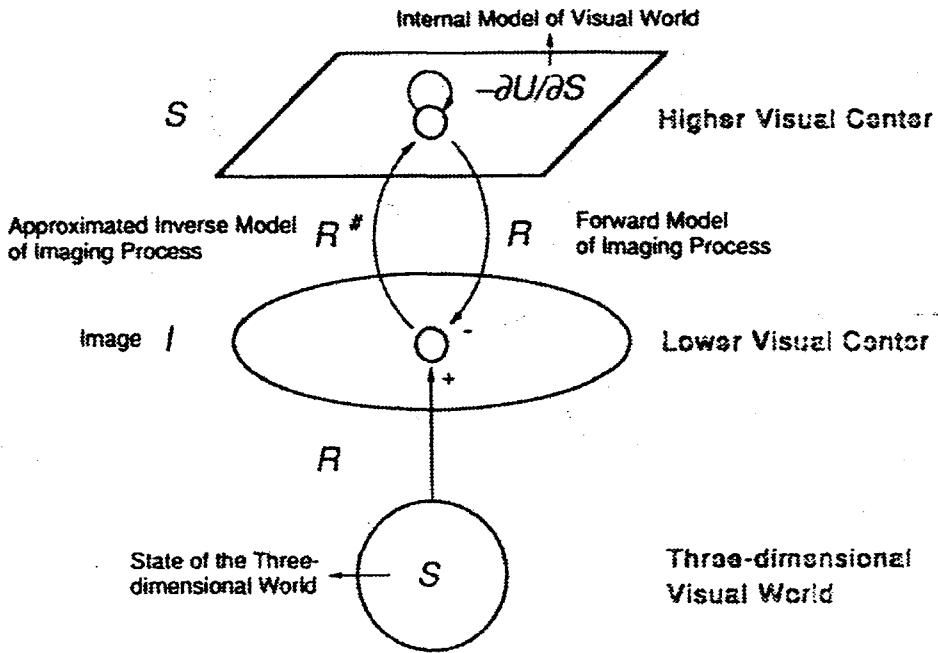


図 1.10: 視覚大脳皮質の計算理論の概念図. (川人, 乾, 1990 [43] より転載)

う方法を考える.

$$\begin{aligned} U(S|I) &= U(I|S) + U(S) \\ &= \frac{1}{2}[R^\# \{I - R(S)\}]^2 + U(S) \end{aligned} \quad (1.21)$$

ここで,  $R^\#$  は光学  $R$  の近似的逆モデルである. 視覚の不良設定性から光学の厳密な逆モデル  $R^{-1}$  は存在しない. しかし, その近似  $R^\#$  は考えられ, 多くの視覚の一撃アルゴリズム (one-shot algorithm) は  $R^\#$  の具体例とみなせる. エネルギー  $U(S|I)$  の最小化には, 前述の Geman らの確率的緩和法を用いることもできるが, これは計算に時間がかかりすぎ脳のモデルとしては適当でない. 従って, 脳の双方向性の結合を用いて, 次式のような神経回路モデルが提案されている.

$$S(0) = R^\#(I) \quad (1.22)$$

$$\frac{dS(t)}{dt} = R^\# \{I - R(S)\} - \frac{\partial U(S)}{\partial S} \quad (1.23)$$

図 1.10 はこの神経回路モデルの概念図である. 低次領野に網膜像  $I$  が与えられるとき, まず低次から高次への順方向の結合で実現される一撃アルゴリズムによって,  $S$  の粗い近似解  $R^\#(I)$  が計算される ((1.22) 式). これは MAP 推定にはなってい

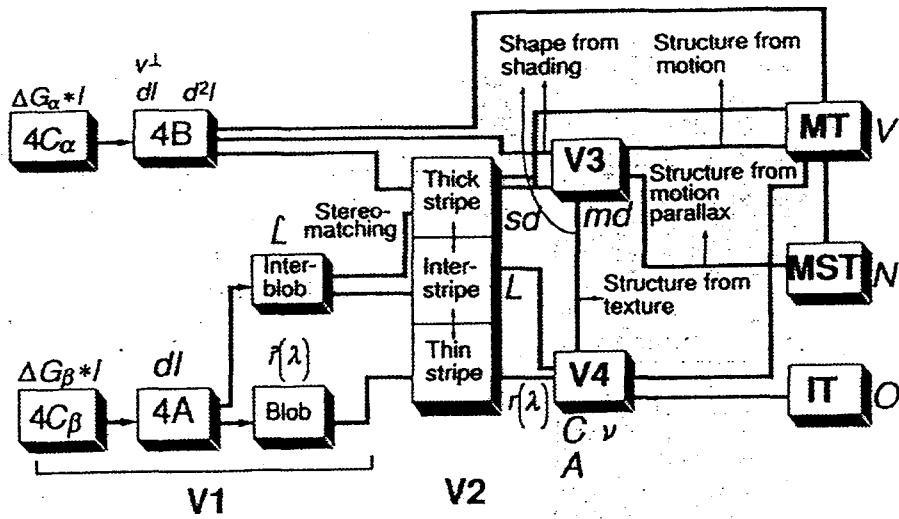


図 1.11: 脳の視覚野が構成する視覚世界の局所並列階層モデル. 矢印のついていない結合は全て双方向性である. (Kawato, Hayakawa, & Inui, 1993 [42] より転載)

ない. 続いて, この粗い近似解  $R^\#(I)$  を初期値として, (1.23) 式で表される順方向, 逆方向のループによる繰返し計算にはいる. 高次から低次への逆方向結合によって現在の推定値  $S$  から網膜像の推定値  $R(S)$  を計算し, 低次領野で実際の網膜像  $I$  と比較して誤差  $I - R(S)$  を求める. この誤差が順方向結合を通して高次領野に戻され,  $R^\#\{I - R(S)\}$  が入力される. 一方, (1.23) 式の第 2 項は高次領野内の固有神経結合を表しており, 外界についての拘束条件が実現されている. もし近似的に  $\partial\{R^\#R(S)\}/\partial S$  が恒等写像になれば, (1.23) 式は (1.21) 式の最急降下法を与えていることがわかる. この繰返し計算は, 粗い近似解を初期値として, これを改善していくように働くので, 多数の繰返し計算は必要でない.

図 1.11 は, 生理学や解剖学の知見に基づいて (1.20) 式の右辺の各変数が表現されている解剖学的位置とその相互結合関係を示したものである [43].  $V$ ,  $N$ ,  $O$  などの推定は高次視覚に対応する.

本稿の主題である, 両眼視による奥行  $sd$  の計算は,  $V1$  の  $4B$  層と  $V2$  の太い縞の間の結合で行なわれていると考えられている. また, この計算は,  $V1$  のプロップ間や  $V2$  の縞間隙で表現されている不連続境界と, 強く相互作用している. 川入らは, この視覚大脳皮質の計算理論に基づいた両眼視モデルでは,  $V1$  の  $4B$  層から  $V2$  の太い縞への順方向結合の計算に, 対応問題を解く一撃アルゴリズムである, Marr と Poggio の第 2 アルゴリズム [56] を用いることを提案している [43]. Marr と Poggio の第 2 アルゴリズムは,  $\nabla^2 G * I$  のゼロ交差を対応要素として用いるアルゴリズムである.

## 1.2 オクルージョン研究の背景

### 1.2.1 初期視覚におけるオクルージョンの影響

第1.1節で、両眼立体視を中心に、初期視覚の理論的な研究の流れを概観した。それを見ると、初期視覚の研究は、各モジュールで奥行手がかりを得る際に生じる不良設定問題を解決するための手法の研究であったことができる。両眼視モジュールにおいては、この不良設定性は対応問題という形で生じている。従って、これまでの理論的な研究では、両眼視モジュールの目的は対応問題の解決であるとみなされていたということができる。

しかし、視覚系が外界の構造復元を行なうときには、オクルージョン(occlusion; 遮蔽)<sup>15</sup>という問題が存在する。オクルージョンとは、手前にある物体が、それより奥にある物体(の一部)を覆い隠す状態である。空間知覚を成立させる感覚様式には視覚、聴覚、運動感覚の三つがあるが、このうち視覚だけが遮蔽の知覚を直接に成立させる。実際、聴覚による通信が視覚による通信に優れているのは、遮蔽物に遮られないという点である。

両眼視モジュールでは、オクルージョンによって、一方の眼からは見えている領域が、もう一方の眼からは手前の面に遮蔽されて見えなくなるという状況が起こり得る。このような領域上の点は、両眼間非対応点(interocularly unpaired point)<sup>16</sup> [64]と呼ばれる。図1.12は、両眼間非対応点が生じる状況の模式図である。このような非対応点は、片眼からしか見ることができないため、正しい対応が存在しない。従って、対応問題を考える際には、これらの点は偽対応しか生成しない。そのため、対応問題の解決を目的とする従来の両眼視のモデルでは、非対応点は計算を困難にするノイズであると見なされてきた。

このような、オクルージョンによって生じ、従来の理論では計算を困難にするノイズとして扱われる特徴は、他の初期視覚モジュールにも存在する。例えば運動視のモジュールにおいては、外界の物体、あるいは観察者自身の運動によって、これまで見えていた領域が手前の面に遮蔽されて見えなくなったり、逆に今まで他の物体の背後にあって見えなかった領域が見えるようになるという状況が起こり得る。これらの領域は、直後、あるいは直前の画像との対応がとれないため、速度や運動方向を検出することができない。さらに、正しい対応が存在しないため、両眼視における非対応点と同様に、計算を困難にするノイズになる。このような領域は生成/消滅領域(accretion/deletion region) [20, 39]と呼ばれる。陰影からの形状復元においては、影(shadow)という問題がある。この影は、両眼視にお

<sup>15</sup>遮蔽は、視界から消える(going out of sight)現象の一つである。視界から消える現象についての生態光学(ecological optics)的な研究は、Gibson [19, 20] を参照。

<sup>16</sup>または、半遮蔽点(half-occluded point)とも呼ばれる [3]。

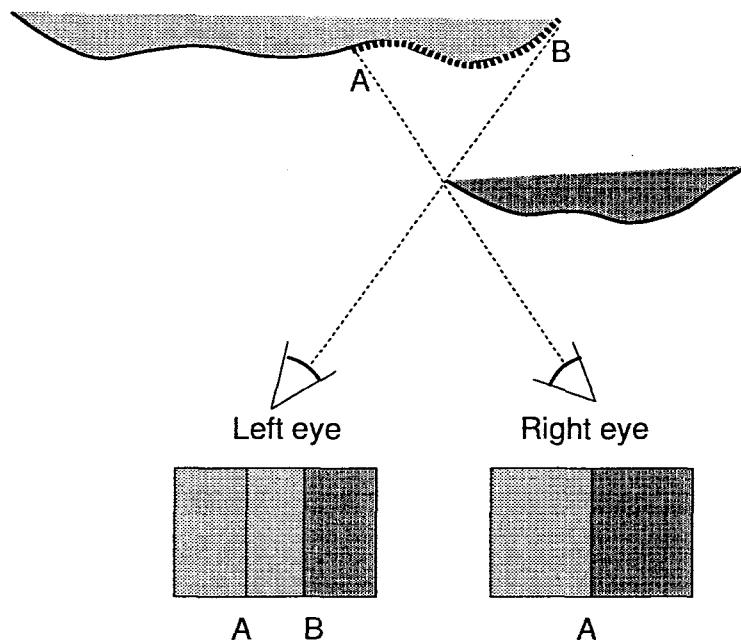


図 1.12: 両眼間非対応点の例. 上図: 外界の構造の鳥瞰図. 下図: 網膜像. 奥の面(明灰色)上の領域 AB は、左眼からは見ることができるが、右眼からは手前の面(暗灰色)に遮蔽されてみることができない. このような領域 AB 上の点を、両眼間非対応点と呼ぶ.

ける非対応点や運動視における生成/消滅領域のように、眼球からの視線が遮蔽されて起こるのではなく、光源からの光線が遮蔽されることによって生じる。陰影からの構造復元問題においても、影の生じている領域では、面の局所的な形状とは無関係に陰影の分布が変化するため、やはり計算が困難になる。

以上で述べたように、オクルージョンによって生じる特徴は、これまでの初期視覚の理論においては計算を困難にするノイズになる。そのため、オクルージョンによって生じる“ノイズ”を除去するための手法もいくつか提案されている。次節では、人間の視覚系がこれらのノイズをどのように扱っているのかを示唆する心理学的な知見を紹介する。

### 1.2.2 両眼間非対応点の心理学的知見

前節で、初期視覚におけるオクルージョンの影響について概説し、これまでの初期視覚の理論的研究では、オクルージョンによって生じる両眼間非対応点などの特徴は計算を困難にするノイズとみなされることを述べた。しかし、近年の心

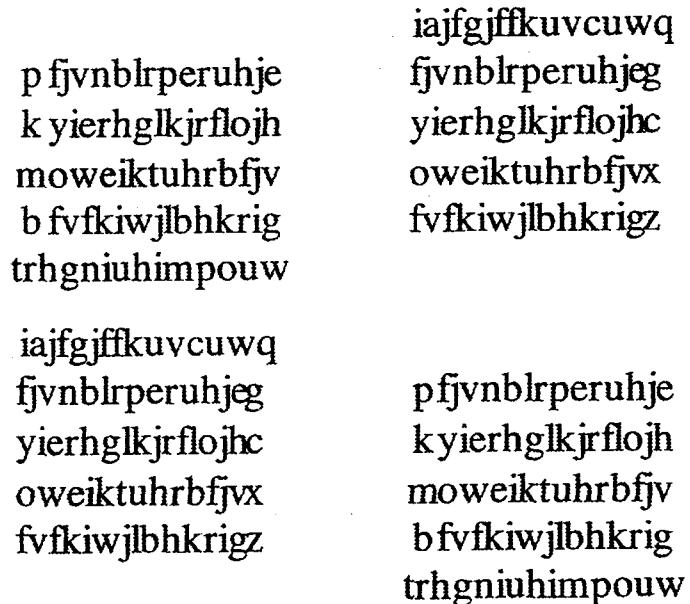


図 1.13: Kaufman のステレオグラム。上下それぞれのステレオグラムについて、左側の図が左眼画像、右側の図が右眼画像である。(Howard & Rogers, 1995 [31] より転載)

理学的な知見では、オクルージョンによって生じるこの“ノイズ”は、実はノイズではなく、脳内では奥行手がかりとして用いられているということが示唆されている。この節では、脳の初期視覚モジュールで、オクルージョンがどのように処理されているかを調べた心理実験を概説する。

両眼間非対応点の存在が両眼視による奥行知覚に影響を与える、正しい対応の存在する両眼間対応点 (interocularly paired point) の視差のみが両眼視モジュールの奥行手がかりとして用いられていると考えると説明できないような知覚が生じることは、現象としては 30 年近く前から知られていた。

Kaufman (1965) [40] は、図 1.13 のようなステレオグラムを考案している。図 1.13 の上のステレオグラムは、中央の文字は左右で対応して視差を得ることができるが、左眼像の一番下の行 (trhgniuhiimpouw) と一番左の列 (pkmb)，および右眼像の一番上の行 (iajfgjffkuvvcuwq) と一番右の列 (gcxz) には対応する文字が存在しない。しかし、このステレオグラムを融合視すると、中央の正しい対応の存在する文字は、上下左右の非対応の文字列より手前にあるように知覚される。両眼視モジュールで用いられる奥行手がかりは、対応点から得られる視差のみであると考える従来の理論では、この知覚を説明することができない。なぜなら、このステレオグラムで視差が得られるのは対応の存在する中央の文字のみで、しかも

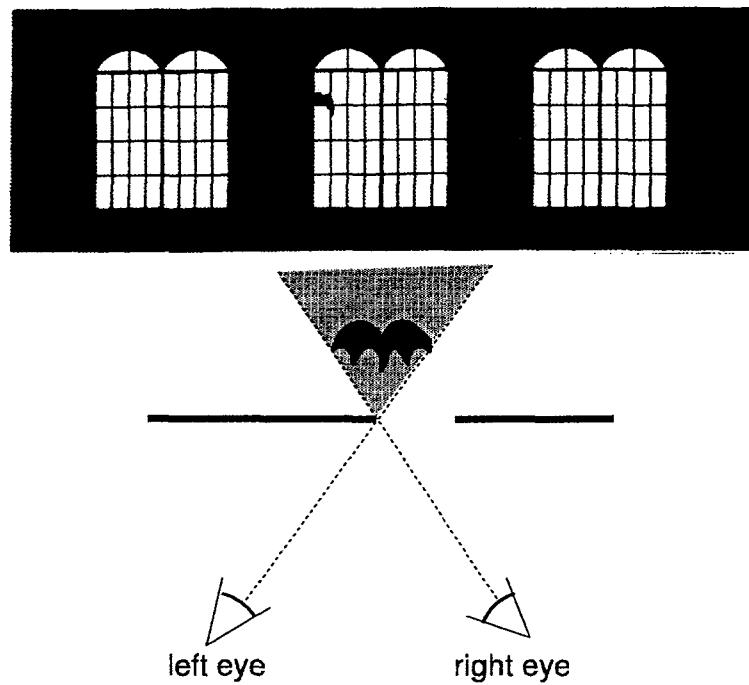


図 1.14: 上図: Kaye のステレオグラム。左右の図が左眼画像、中央の図が右眼画像である。並行法で融合視する場合は左図と中央図を、交差法で融合視する場合は中央図と右図を用いる (Kaye, 1978 [44] より転載)。下図: Kaye のステレオグラムを生成する 3 次元構造の鳥瞰図。

この視差は一定である。そして、従来の理論では、視差情報を得ることができない非対応点の奥行は、近傍の対応点の視差からの補間によって推定されることになる。しかし、このステレオグラムには非対応の文字の持つ奥の視差を与えるような対応点は存在しないからである。

Kaye (1978) [44] も同様のステレオグラムを考案している。図 1.14 のステレオグラムは、窓の外にコウモリが飛んでいるという状況を描いたものである。しかし、左眼からは窓の周囲の壁に遮蔽されるため、右眼からしかコウモリを見ることができない。このステレオグラムでも、対応特徴である窓枠から得られる視差は一定であり、非対応特徴のコウモリから視差を得ることはできないが、融合視するとコウモリが窓の外にいるように知覚される。

ここで、図 1.13 のステレオグラムでは、アルファベットという意味のある図形が用いられている。また、図 1.14 では窓とコウモリが描かれており、片眼像のみを見ても理解できる情景をしている。つまり、これらのステレオグラムには豊富な視覚情報が含まれているということになる。従って、確かに融合視して生じる

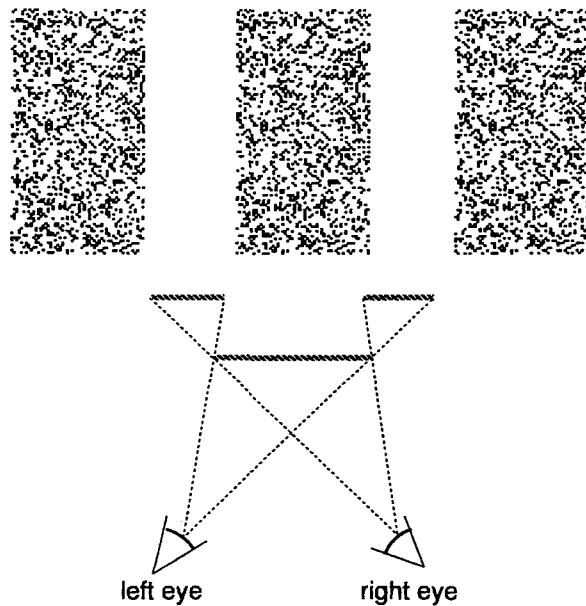


図 1.15: 上図: ランダムドットで作られた da Vinci ステレオグラム. 左右の図が左眼画像, 中央の図が右眼画像である. 中央のドットは視差 0 で対応する. 左眼像の左端, 及び右眼像の右端にあるドットは, 対眼に対応する点を持たない非対応点である. 並行法で融合視する場合は左図と中央図を, 交差法で融合視する場合は中央図と右図を用いる. 下図: 上のステレオグラムを生成する 3 次元構造の鳥瞰図. 図の面上にドットパターンが描かれているが, 奥の面は片眼からしか見ることができない.

知覚は両眼視差だけでは説明できないが, その知覚に影響を与えていたのは非対応点ではないという可能性もある. 例えば, 図 1.14 では, コウモリは窓の外を飛ぶものであるという知識が奥行知覚に影響を及ぼしているのかもしれない. この可能性を排除するためには, ちょうど Bela Julesz が行なったように, 図 1.15 のようなランダムドットステレオグラムを作つてみればよい. このステレオグラムの中央のドットは視差 0 で両眼対応する. そして, 左眼像の左端, 及び右眼像の右端にあるドットは, 対眼に対応する点を持たない非対応点である. これは, 図 1.13 や図 1.14 と同様の状況である. しかし, このステレオグラムには, 両眼対応するかしないかということ以外の情報は存在しない. 融合視すると, 両端のドットは視差 0 の面より奥に知覚される. よって, 非対応点が奥行知覚に影響を与えていたことが分かる.

以上のように, 視差情報を得ることのできない非対応点が奥行知覚に影響を与えることは知られていたが, これはあくまで現象として知られていたに止まっていた.

心理学的に、非対応点が脳内の両眼視モジュールでどのように処理されているかが調べられたのは、1990年代に入ってからである。これには、Nakayama, Shimojoらの一連の研究 [62–65, 87] や、Anderson [3], Liuら [50] の研究がある。NakayamaとShimojo (1990) [64] は、図1.13や図1.14、図1.15のように、非対応点の存在が奥行知覚に強い影響を与え、従来の両眼視の理論ではその知覚を説明できないようなステレオグラムを用いて心理実験を行っている。彼らはこのようなステレオグラムを、非対応点が奥行知覚に影響を与えるという説を最初に記した Leonard da Vinci にちなんで、da Vinci ステレオグラムと呼んでいる<sup>17</sup>。Nakayamaらの行った心理実験の詳細は第3章で述べるが、彼らの研究からは次のことが示唆される：

1. 脳では、両眼間対応点だけでなく、オクルージョンによって生じる両眼間非対応点の検出も行なわれている。そして、脳の高次の領野では片眼優位の細胞はほとんど存在しないという生理学的な知見から、非対応点の検出は初期の領野(おそらくV1)で行なわれていると考えられる。
2. 両眼間非対応点は、物体の遮蔽関係の手がかりとして、外界の構造復元に用いられている。そして、物理的なオクルージョンの幾何学的な条件(オクルージョンの拘束条件)を満たすような面が再構成される。この条件は、例えば画像上のある点に非対応点が与えられたとき、脳内で再構成される面もそこに非対応点を生じさせるような形状に制限するというものである。

図1.14のステレオグラムを例に取ると、物理的に正しい光学過程によってこのような両眼像を生成可能な3次元構造は、コウモリが図1.14の下図の灰色の領域にいる場合に限定される。従って、オクルージョンの拘束条件によって、コウモリがこの灰色の領域内に位置しているような構造が再構成される。これによって、窓の外にコウモリが存在するという知覚を説明することができる<sup>18</sup>。

対応点からは窓枠の視差しか得られないので、非対応点を奥行手がかりとみなさない従来の両眼視のモデルでは、コウモリは窓枠と同じ視差で補間されると考えられる。

心理学的な知見から、上記のようなことは両眼視モジュールだけでなく、他の初期視覚モジュールでも示唆されている。運動視モジュールにおいては、Gibson, Kaplanら [20, 39] や Anderson と Sinha [4] によって、生成/消滅領域の存在が複数の面の前後関係を知覚させる手がかりになることを示している。陰影からの形状

---

<sup>17</sup>これに対し、非対応点の奥行は対応点の視差から補間されているという考えでも説明できる従来のステレオグラムは、1.1.1節で述べた Wheatstone の名をとって、Wheatstone ステレオグラムと呼ばれている [64]。

<sup>18</sup>但し、この説明では、コウモリが灰色の領域中のどの位置に知覚されるかは特定できない。このような場合、脳の両眼視モジュールでどのように奥行が一意に決定されているかについては、3.2節で述べる。

復元モジュールにおいても、影が奥行手がかりになることが心理学的に示されている。Kerstenら [45, 46] は、画像中の影のみを動かすことによって、その影を生成する物体が運動しているように知覚されることを示している。このことは、光源位置が変化しないと仮定して、物理的に正しい光学過程によって背景面にその影を生じさせるような位置に物体が存在するという拘束条件が脳内で用いられていることを示唆している。

従って、初期視覚の各モジュールにおいて、視差を与える両眼間対応点のように外界の面の構造を直接記述するパラメータを与える手がかり（これを直接手がかり (direct cue) と呼ぶことにする）のみではなく、両眼間非対応点のように複数の面の相対的な遮蔽関係を与える手がかり（これをオクルージョン手がかり (occlusion cue) と呼ぶことにする）も用いる理論が必要である。

### 1.2.3 両眼間非対応点を扱うアルゴリズム

これまでに両眼視のモデルは数多く提案されているが、オクルージョンが考慮されたものはほとんど存在しない。この節では、いくつか提案されている両眼間非対応点の存在を考慮した両眼視モデルを概説する。

前節でも述べたように、両眼間非対応点は正しい対応を持たず、偽対応しか生成しないため、従来の両眼視のモデルではノイズとみなされる。両眼間非対応点を考慮したモデルの一つは、非対応点を検出し、これを単に対応候補から除外するものである。非対応点を対応候補から除外することによって、非対応点の生成する偽対応をなくすことができ、視差推定の精度を向上させることができる。

Nasrabadiら (1989) [67] は、運動視モジュールで得られるオプティカルフロー (optical flow) を利用して非対応点を検出して除外する MRF モデルを提案している。彼らのモデルでは、対応する左右眼網膜上の点のオプティカルフローを比較して、その二乗誤差がある閾値以上であれば、この点を非対応点とみなす。そして非対応とみなされた点では、データフィッティング項を用いずに、スムーズネス項のみで視差推定を行うことによって、非対応点が生成する偽対応の影響をなくすことができる。このモデルでは、非対応点の視差は近傍の対応点の視差からの補間によって推定されることになる。しかし、ステレオグラムのように、オプティカルフローのデータが得られない場合にはこのモデルを用いることはできない。

Chang と Chatterjee (1993) [8] や Yang と Yuille (1995) [109] は、両眼視モジュールで扱われるデータのみを用いて非対応点を検出する方法を提案している。彼らのモデルでは、まず非対応点の存在を考慮せずに対応問題を解き、視差の推定を行う。その後に、得られた視差を基にして非対応点の検出を行う。いま、左眼座標上の点  $(x_l, y_l)$  の視差を  $d_l(x_l, y_l)$ 、右眼座標上の点  $(x_r, y_r)$  の視差を  $d_r(x_r, y_r)$  と

すると、この両眼の点が対応点のとき、

$$d_l(x_l, y_l) = d_r(x_r - d_l(x_l, y_l), y_r) \quad (1.24)$$

が成り立つ。彼らのモデルでは、この式を満たさない点を非対応点とみなす。Changらは、(1.24)式の条件を dual matching condition と呼んでいる<sup>19</sup>。脳の両眼視モジュールのモデルとしてみた場合、彼らのモデルにおける問題点は、非対応点を検出する前にまず対応問題を解き、視差を推定しておかなければならぬことである。1.2.2節で紹介した心理実験から、両眼視モジュールでは対応点と同時に非対応点も奥行手がかりとして用いられていると示唆されるため、非対応点を検出する前に対応問題を解いていては意味がない。

McLoughlin と Grossberg [22, 23, 61] は、視差推定と非対応点の検出を同時に行なうモデルを提案している。彼らのモデルは、両眼視における画像のコントラストの影響を調べた心理実験 [24, 60, 88] を説明することを目的としたものである。彼らのモデルは、Marr と Poggio の協調的アルゴリズムのように特定の位置及び視差に選択性を持つ両眼性細胞と、非対応点を検出する単眼性細胞から構成されている。両眼性細胞は左右眼画像上の対応特徴の類似度を、単眼性細胞は各眼の対応特徴そのものを入力として受けとる。そして、左右眼網膜上の同一の点から入力を受ける細胞間で競合を行なうことによって対応問題を解き、同時に左右眼の対応特徴の類似度が低いものは非対応点として検出される。

以上で概説したモデルは、いずれも非対応点を検出するが奥行手がかりとして用いてはおらず、非対応点の視差は近傍の対応点の視差からの補間によって推定される。

両眼間非対応点を考慮したもう一つのモデルは、非対応点を奥行手がかりとみなし、視差推定に用いるものである。

Geiger, Ladendorf と Yuille (1995) [17] は、オクルージョンによって非対応点が生じる時の幾何学的な条件から、

エピポーラ線に沿った片眼の不連続はもう一方の眼の非対応点に対応する。逆もまた成り立つ。

というオクルージョンの拘束条件を提案し、さらにこの拘束条件を用いて非対応点を奥行手がかりとして扱うアルゴリズムを提案した。このオクルージョン拘束は、Wheatstone ステレオグラムの場合にしか成立しない。Geiger らのアルゴリズムは、前述の McLoughlin らのアルゴリズム [22, 23, 61] と同様に、特定の位置及び視差に選択性を持つ両眼性素子と、左右眼網膜上の非対応点を検出する素子からなる。そして、彼らのオクルージョン拘束は、実際の計算においては図 1.16 のよう

---

<sup>19</sup> または、back matching と呼ばれる [74]。

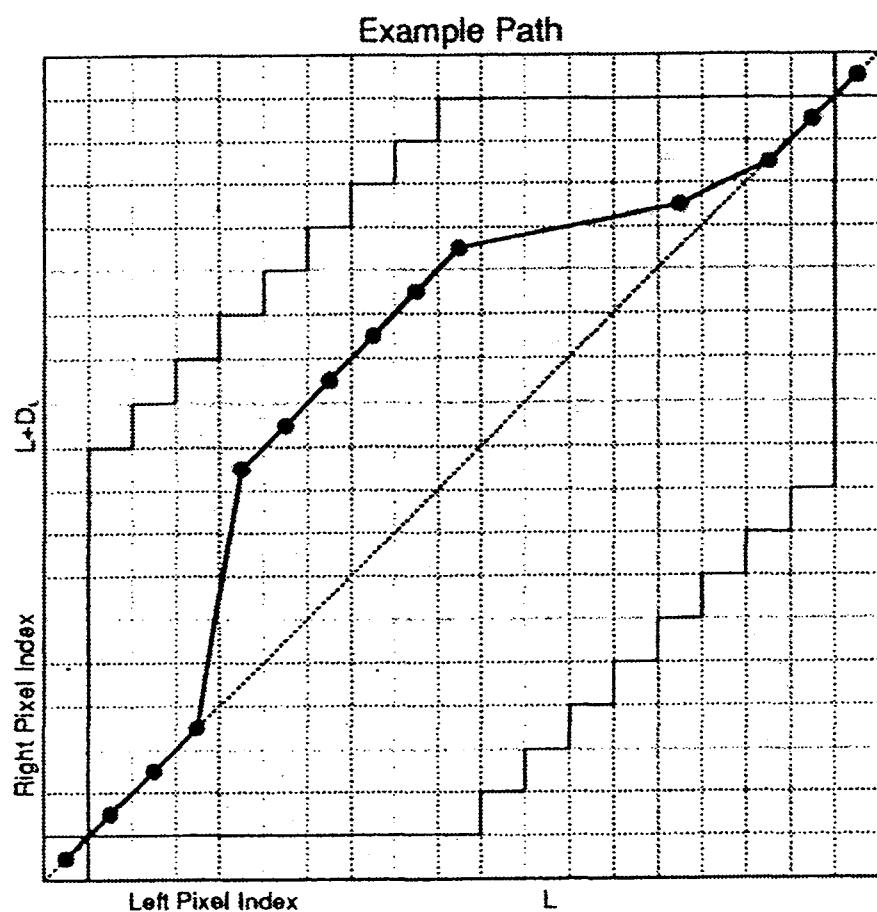


図 1.16: 経路探索問題の解の例. 対応点(黒丸)が検出されている領域では、スムースネス拘束から前額並行面(front parallel plane)が推定される. これに対し, Geiger らのオクルージョン拘束から、左眼の非対応領域では右眼の視線に沿った経路が、右眼の非対応点では左眼の視線に沿った経路が推定される. (Geiger, Ladendorf, & Yuille, 1995 [17] より転載)

な1次元の経路探索問題という形で実現されている。IntilleとBobick [33–35] も、同様の両眼視アルゴリズムを提案している。ただし、彼らのモデルでは図1.13や図1.14、図1.15のようなda Vinciステレオグラムの与える知覚を説明することはできず、非対応点の脳内での役割を説明するモデルとしては不十分である。また、非対応領域の両端の視差をつなぐ視差勾配が2の経路を探索することから、非対応点で知覚される奥行も説明することができない。

### 1.3 本研究の目的

第1.1節で概説したように、従来の初期視覚の理論的な研究では、各モジュールが扱う奥行手がかりは唯一であると考えられてきた。例えば、両眼視モジュールの奥行手がかりは両眼視差、運動視モジュールの奥行手がかりは運動視差と考えられている。なぜなら、そもそも単独で奥行知覚を生じさせることができる奥行手がかりの各々について、モジュールが構成されていると考えられているためである。従って、両眼視モジュールにおける対応問題や、運動視モジュールにおける窓問題のように、各モジュールが唯一用いていると考えられている奥行手がかりを得る際に生じる問題を解決することが、そのモジュールの目的であるとみなされてきた。

しかし、第1.2節で概説したように、両眼視のみで奥行知覚がなされているにも関わらず、奥行手がかりとして両眼視差しか得られないと考えると、その知覚を説明することができないステレオグラム、da Vinciステレオグラム [64] が考案されている。このda Vinciステレオグラムは、両眼視によって得られる奥行手がかりは両眼視差のみではなく、オクルージョンによって生じる両眼間非対応点も奥行手がかりとして用いられていることを示唆している。

非対応点は、再構成される面を、その非対応点自身が生じるためのオクルージョンの幾何学的な条件(オクルージョンの拘束条件)を満たすような構造に制限するという方略で、外界の構造復元に影響を与えていていると考えられる。ただし、図1.14で示したように、オクルージョンの拘束条件によって非対応特徴であるコウモリが窓枠より奥に位置することは推測できるが、これだけではコウモリの奥行を一意に決定することはできない。極端な例として、図1.17の非対応点しか存在しないステレオグラムを融合視してみても、人間は奥行知覚を得ることはできない。このことは、非対応点は奥行手がかりではあるが、単独で奥行知覚を生じさせる手がかりではないことを示している。

両眼間非対応点は、両眼視において対応点から両眼視差を得ることを考える際に相補的に生じる特徴である。従って、非対応点は単独のモジュールで処理されるのではなく、これと相補的な両眼視差を扱うモジュールで同時に処理されてい

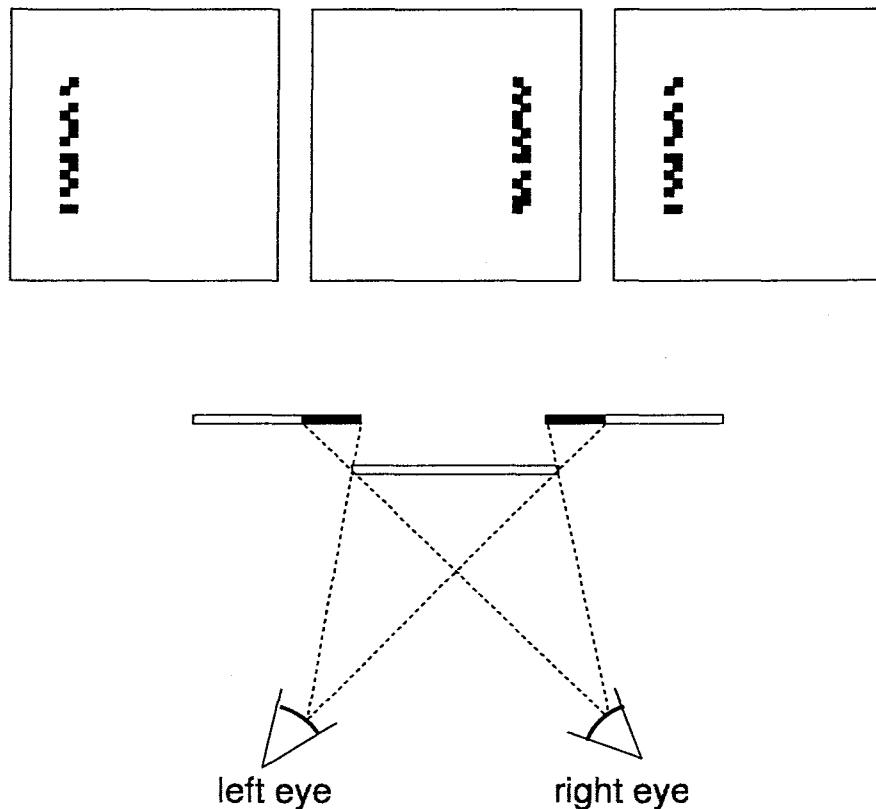


図 1.17: 上図: 非対応点しか存在しないステレオグラム. 左右の図が左眼画像, 中央の図が右眼画像である. 下図: このステレオグラムを生成する3次元構造の鳥瞰図. 両眼から見ることのできる対応領域(白い領域)にはテクスチャがなく, 対応特徴が全く存在しない. 非対応領域(黒い領域)にのみ黒いドットが存在する.

初期視覚モジュール	直接的手段	オクルージョン手段
両眼視モジュール	両眼視差	両眼間非対応領域
運動視モジュール	オプティカルフロー	生成/消滅領域
陰影モジュール	陰影	影
輪郭線モジュール	遮蔽輪郭	T接合, 端点

図 1.18: 初期視覚の各モジュールにおける直接手段とオクルージョン手段.

るのではないかと考えられる。言い換えると、両眼視モジュールで用いられている奥行手がかりは両眼視差のみではなく、非対応点も奥行手がかりとして用いられているのではないかと考えられる。従って、非対応点を奥行手がかりとして用いるための理論を考えるときは、同時に両眼視差からの構造復元の理論も考慮する必要がある。このことは他の初期視覚モジュールについても当てはまると考えられる。図1.18は、各モジュールで扱われる考え方である。

本論文では、初期視覚の中でもオクルージョンの影響が心理学的に最もよく調べられている両眼視モジュールについて、直接手がかりである両眼視差、及びオクルージョン手がかりである両眼間非対応点の双方を奥行手がかりとして用いる計算理論、及びアルゴリズムを提案する。従来提案してきたほとんどの両眼視アルゴリズムは、非対応点を偽対応しか生成しないノイズとして扱っている。非対応点の存在を考慮したモデルでも、非対応点を単に対応候補から除外するだけで、奥行手がかりとしては用いていなかった。Geigerら[17]は、非対応点を奥行手がかりとして扱うモデルを提案している。しかし、彼らのモデルでは、非対応点自体の奥行を推定することはできず、また、非対応点の存在が奥行知覚に大きな影響を与えるda Vinciステレオグラムの知覚を説明することもできないため、両眼視における非対応点の役割を説明するモデルとしては不十分であった。

本論文は以下の心理学的知見に基づいている：

- 視覚系では両眼間対応点のみならず、両眼間非対応点の検出も行なわれている。
- 両眼間非対応点の検出は、初期視覚のレベル(おそらくV1)で行なわれている。
- 視覚系では両眼間非対応点も奥行手がかりとして用いられている。

さらに、本論文で用いられる仮定は以下の通りである：

- 両眼間非対応点は、その非対応点自身が生じるためのオクルージョンの幾何学的な条件(オクルージョンの拘束条件)を満たすような面を再構成するという方略で、外界の構造復元に影響を与えている。
- 両眼視モジュールで、両眼間対応点及び両眼間非対応点双方を奥行手がかりとして用いた構造復元が行なわれている。

## 1.4 本論文の構成と概要

本論文は5章から構成されており、第1章では、研究の背景と目的、そして論文の構成について述べている。第2章では、両眼視の情報処理の基礎となる、脳の

両眼性細胞のモデルを提案している。これは、網膜からV1の両眼性細胞までのモデルに対応する。第3章は、両眼間対応点、及び両眼間非対応点双方を奥行手がかりとして用いるための拘束条件について論じたものである。続く第4章で、この拘束条件を実現するアルゴリズムを提案している。これは、V1以降の初期視覚領野のモデルになる。第5章は、本論文の結論である。第2章から第5章までの概要は、以下に示す通りである。

第2章では、両眼視の情報処理の基礎となる、脳の両眼性細胞のモデルについて述べる。これは、網膜からV1の両眼性細胞までをモデル化したものである。両眼性細胞のモデルとしては、両眼入力のAND演算を行なうANDモデルがある。しかし、実際のデータはアナログ値であると考えられるので、アナログAND演算を行なう必要がある。この章では、アナログAND回路に基づく両眼性細胞のモデルを提案する。このモデルでは、tuned cellおよびnear cell, far cellが、ほぼ同様の回路構成で実現できる。このモデルのnear cell, far cellは、視差選択性の幅が広い両眼性細胞として定義される。この章では、まずアナログAND回路の基本的な回路構成を示した後、Neocognitron型の両眼性細胞を用いた階層型神経回路モデルを提案する。

第3章では、非対応点を奥行手がかりとして用いて構造復元を行なうための、拘束条件を提案する。この章で、どのような論理で、何が計算されなくてはならないかという、計算理論を明確にする。入力両眼像に非対応点が存在するとき、脳の両眼視モジュールでは、オクルージョンの幾何学的な条件に矛盾しない面が再構成されると考えられる。つまり、推定される面は、両眼像に存在する非対応点を生じさせることができない構造に制限されると考えられる。この章では、まず、両眼視におけるオクルージョンの幾何学的な解析に基づいて、オクルージョンの物理的な拘束条件を提案し、非対応点がどのように奥行手がかりとして用いられ、再構成される構造がどれだけ制約されるのかについて述べる。また、両眼視の理論としては、非対応点のみではなく対応点も同時に扱え、一意な解が得られるものでなくてはならない。この章では、さらに、対応点および非対応点双方を奥行手がかりとして構造復元を行なうために必要な、他の拘束条件も提案する。この拘束条件とは、オクルージョンを考慮して拡張した一意性の拘束条件と、凝集性の拘束条件である。そして、この拘束条件は、da Vinciステレオグラムをはじめ、特殊なステレオグラムにおいても矛盾しないことが示される。

第4章では、前章で提案した対応点および非対応点双方を奥行手がかりとして用いる理論を実現するための、アルゴリズムを提案する。そして、このアルゴリズムで何を計算しなくてはならないか、そしてどのような表現を用いれば良いかについて論じる。これは、V1の両眼性細胞以降の初期視覚領野のモデルになる。このアルゴリズムは、網膜からの入力データを対応点と非対応点の二種類のデー

タに区別して扱う。そして、データが対応点のところでは両眼視差を用いて、非対応点のところではオクルージョンの幾何学的な条件を満たすようにするという方法を用いて、外界の構造復元を行なう。従って、このアルゴリズムは(1)外界の構造復元、及び(2)直接手がかりを与える領域とオクルージョン手がかりを与える領域との区別、という二つの問題を解くことになる。このアルゴリズムでは、推定した視差を表す細胞と、各点が対応点か非対応点かを区別する細胞の、二種類の細胞が用いられる。そして、この二種類の細胞が互いに相互作用しながら、繰返し計算によって推定を行なう。さらに、このアルゴリズムの計算機シミュレーションによって、従来の両眼視アルゴリズムでは扱えなかった da Vinci ステレオグラムの知覚も説明できることが示される。

第5章は本論文の結論であり、本論文のまとめを行なうと共に、本研究が関連研究へ与える影響などについて述べる。

## 第2章 両眼性細胞のモデル

この章では、両眼視の情報処理の基礎となる、脳の両眼性細胞のモデルを提案する。まず、2.1節で、両眼性細胞の種類やその性質についての生理学的な知見を概説する。続く2.2節では、複雑型の両眼性細胞のモデルとして知られている、エネルギーモデルについて概説する。視差検出細胞の性質を説明するモデルとしては、両眼入力のAND演算を行なう、ANDモデルが提案されている。2.3節では、ANDモデルの考え方を基に拡張した、アナログANDモデルの基本的な回路構成を提案する。アナログANDモデルでは、tuned cellやnear cell, far cellと呼ばれる視差検出細胞が、ほぼ同様の回路構成で実現できる。そして、2.4節で、このアナログAND回路の考え方に基づいて、Neocognitronを用いた網膜からV1の両眼性細胞までの階層型神経回路モデルを提案する。

### 2.1 両眼性細胞の種類

左右眼網膜に投影された像は、LGN(lateral geniculate nucleus)を経由して、V1の細胞に送られる。V1以降の領域では、両眼から入力を受ける両眼性細胞が存在する。

両眼入力を受ける細胞の種類については、Bishopら[6]がネコのV1で調べたものがよく知られている。彼らは、通常の両眼性細胞の他に、ゲート細胞といわれる特性を持った細胞があることを報告している。通常の両眼性細胞は左右眼いずれからの単眼入力に対しても反応するが、両眼を同時に刺激した場合には、左右眼からの単眼入力を足し合わせたものよりも大きな反応がみられる。これに対し、ゲート細胞は、左右いずれか一方の眼(優位眼)の刺激に対しては片眼刺激でも反応するが、優位でない眼からは抑制性入力を受けており、両眼刺激を与えた場合は優位眼に与えられた刺激による反応も抑制されてしまう。つまり、非優位眼からの情報によって、優位眼からの情報がゲートされているように見える。

さらに、両眼性細胞の中には、単眼入力に対してはほとんど反応せず、最適な視差を持つ両眼入力が与えられたときだけ大きな反応が見られる、視差検出細胞とよばれるものが存在する。視差検出細胞は、PoggioとFischer[75]が反応特性により、tuned cellとnear/far cellの二つのカテゴリーに分類している。図2.1は、こ

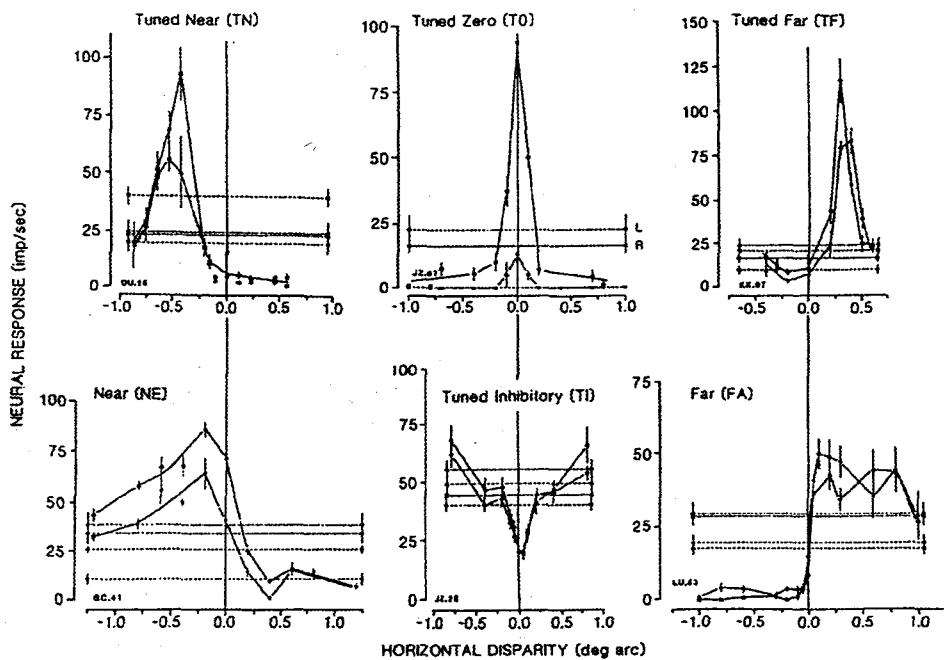


図 2.1: サルの視覚野の両眼性細胞の視差特性. (Poggio, Gonzalez, & Krause (1988) [76] より転載)

これらの視差検出細胞の、視差に対する反応特性をプロットしたものである。各グラフとも、横軸が視差(単位は度)、縦軸が発火頻度(1秒当たりのパルス数)を表す。実線が両眼入力を与えたときの発火頻度である。破線は左眼にのみ、点線は右眼にのみ視覚刺激を与えたときの発火頻度を表す。

tuned cellは、注視点の前後の比較的小さな視差に対して反応し、非常に鋭い視差選択性を持っている。このうち、tuned excitatory cellは、最適視差と一致する入力を与えたときのみ反応する細胞である。図2.1の上の三つのグラフがこの細胞に相当する。tuned inhibitory cellはこれとちょうど反対の特性を持ち、ある特定の視差のときに反応が抑制される。図2.1下段の中央のグラフが tuned inhibitory cellである。これに対し、near cell, far cellは、入力が視差の値によらず注視点より近く、または遠くに与えられたときに反応するという特徴を持っている。図2.1下段の左のグラフが near cell、右のグラフが far cellである。Ferster [12]は、視差検出細胞の特性を詳細に調べ、V1, V2野にみられる near cell, far cellはオリエンテーション選択性を持ち、垂直に近いオリエンテーションを持つものは主に水平視差に、水平に近いオリエンテーションを持つものは主に垂直視差に反応することを報告している。また、片眼から興奮性入力、もう片眼から抑制性入力を受けるmonocular near cell, monocular far cellがあることも報告している。これは、near

cell, far cell のような特徴を持ったゲート 細胞を考えることができる。検出できる視差範囲については、tuned cell では最適視差範囲が最大で  $0.5^\circ$  までと狭い視差範囲しか検出できないのに対し、near cell, far cell では最大で  $4.0^\circ$  にまでわたることが報告されている。

これらの様々な種類の両眼性細胞は、脳内でどのような役割を果たしているのであろうか。両眼視による奥行き知覚を行うには、Marr [53] が述べているように眼球運動が不可欠であると考えられる。これは、粗いが広い視差 (coarse stereopsis) を検出する細胞が輻輳性眼球運動を駆動し、精密だが狭い視差 (fine stereopsis) しか検出できない細胞によって奥行きが計算されるという考え方である。即ち、near cell, far cell が注視点以外の箇所の大域的な奥行きを認識して輻輳性眼球運動を駆動し、tuned cell が注視点付近の奥行きを高解像度で認識していると考えられる。

また、near cell, far cell は奥行き知覚の他にも、自己運動 (self-motion) の知覚に使われているという報告もある。Roy [83] らは、自己運動の知覚を行っていると考えられている MSTd の神経細胞を調べ、その 90% 近くの細胞が near cell または far cell であったことを報告している。また彼らは、その細胞のほとんどが最適速度 (preferred speed)、最適運動方向 (preferred direction) を持つことも報告している。これは、自己運動知覚に適した特性である。

## 2.2 エネルギーモデル

前節で述べた様々な種類の両眼性細胞の各々について、単純型細胞 (simple cell) および複雑型細胞 (complex cell) が存在する。単純型両眼性細胞は、この細胞の最適視差および最適特徴を持つパターンが受容野内に与えられたときでも、入力パターンの提示位置が受容野の興奮性領域から外れていれば反応しない。これに対し、複雑型両眼性細胞は、最適視差を持つ入力パターンが受容野内に与えられたら、その提示位置によらず常に反応できる。複雑型両眼性細胞は、同一の最適視差を持ちながら、受容野の位置がわずかに異なる複数の単純型両眼性細胞から入力を受けるという、階層構造によって実現されていると考えられている (Hubel-Wiesel の階層仮説 [32])。この単純型細胞 → 複雑型細胞という階層構造によって複雑型細胞の特性を説明するモデルには、様々なものが提案されている。両眼性の単純型細胞から両眼性の複雑型細胞を構成するモデルとしては、エネルギーモデルが提案されている。この節ではエネルギーモデルを概説する。

エネルギーモデルは、そもそも Adelson と Bergen (1985) [1] が運動検出のモデルとして提案したものである ([104] も見よ)。しかし、容易に両眼性細胞のモデルにも応用することができる (詳しい解析は [81,82] を参照)。このモデルでは、入力パターンの提示位置や単純型細胞の最適位置を、位相で表現している。そして、複

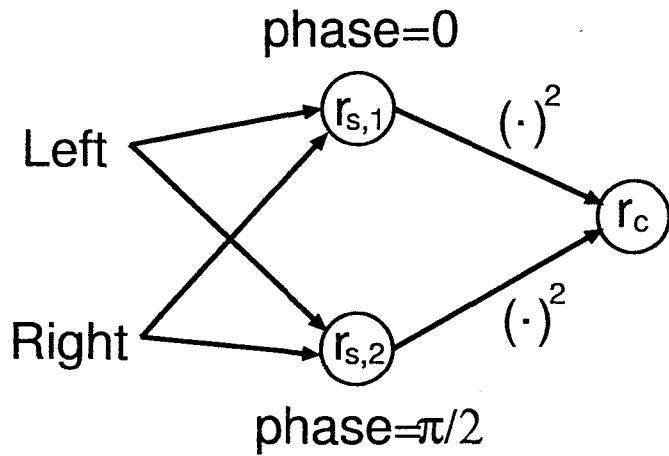


図 2.2: エネルギーモデルの基本回路構成.

雑型両眼性細胞では、入力パターンの位相によらない、パワースペクトルが計算されていると考えている。このエネルギーモデルは、複雑型両眼性細胞を、必要最小限の細胞数で構成することを目指したモデルである。

以下、簡単のため1次元で考えることにする。単純型両眼性細胞の左右眼受容野  $f_l(x)$ ,  $f_r$  を、次式のような1次元 Gabor 関数で定義する。

$$f_{l/r}(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos(\omega x + \phi_{l/r}) \quad (2.1)$$

ここで、 $\omega$  は最適周波数、 $\phi_{l/r}$  は左右眼受容野の位相を表す。この細胞の最適視差は、位相差  $\phi_r - \phi_l$  で定義される。この細胞の出力は、コンボリューション

$$r_s = \int [f_l(\xi)I_l(x - \xi) + f_r(\xi)I_r(x - \xi)] d\xi \quad (2.2)$$

で定義される。ここで、 $I_{l/r}$  は左右眼網膜像である。

複雑型両眼性細胞は、受容野の位相がちょうど  $\pi/2$  異なる二つの単純型両眼性細胞  $r_{s,1}$ ,  $r_{s,2}$  から入力を受けとる。二つの単純型細胞の出力は二乗されて複雑型細胞へ送られる。複雑型細胞の出力

$$r_c = (r_{s,1})^2 + (r_{s,2})^2 \quad (2.3)$$

は、画像のパワースペクトルに相当する。Adelson らは、これをエネルギーと呼んでいる。また、このモデルは直角位相の細胞組を用いるため、quadrature pair モデルとも呼ばれる。エネルギーモデルは、最小限二個<sup>1</sup>の単純型両眼性細胞から複

<sup>1</sup>細胞出力  $r_s$  が負の値を取らないとすると四個必要。

雑型両眼性細胞を構成できることを示してある。図2.2に、エネルギーモデルの回路図を示す。Freemanら[9,73]は、このモデルで受容野形状として用いられるGabor関数で、生理学的な実験で得られた両眼性細胞の受容野のデータを良く近似できることを報告している。

エネルギーモデルでは、最適視差は左右眼受容野の位相差で決定され<sup>2</sup>、tuned inhibitory cellは左右眼受容野の位相がちょうど反転している(位相差 $\pi$ の)細胞として定義される。このモデルでは、near cellやfar cellは、単に大きな視差を検出する(位相差 $\pm\pi/2$ 近くの)細胞とみなされる。受容野の空間周波数が低い場合、生理学的なデータとよく似た視差感受性曲線が得られる。これは、tuned cellとnear/far cellの性質には本質的な違いではなく、単に受容野の空間周波数が異なるだけの同じ種類の細胞であるとみなすことに相当する。

## 2.3 アナログ AND モデルの基本構成

第2.1節で概説した視差検出細胞の一つ、tuned cellのモデルとして、ANDモデルが提案されている。tuned cellは、単眼入力ではほとんど反応せず、最適視差を持つ両眼入力が与えられたときのみ大きな反応を示す。ANDモデルは、2入力のAND回路を用いて、左右眼から同時に入力がある場合にのみ反応するという特性をモデル化している。この考え方に基づいたモデルは、これまでにもHiraiとFukushima[26]やMarr[53]などによって提案されている。また、このようなANDゲート的特性を持つモデルは、両眼性細胞以外にも、エッジ検出細胞など様々な細胞のモデルとして用いられている[54]。

この節では、tuned cellをはじめ、near cell, far cellなどの視差検出細胞モデルの、基本的な考え方を提案する。ここで提案する回路構成は、ANDモデルの考え方を基に拡張した、アナログAND回路を用いている。この回路では、tuned cellとnear/far cellの特性を、ほぼ同様の回路構成で実現できる。そして続く2.4節で、この考え方に基づいて、Neocognitron型の神経回路[13,14]を用いた網膜からV1の視差検出細胞までの階層型神経回路モデルを提案する。

### 2.3.1 tuned cellの基本回路

tuned cellの神経回路をAND回路でモデル化するという考え方は、これまでにもHiraiとFukushima[26]やMarr[53]などによって提案されている。これは、2入力のAND回路を用いて、左右眼から同時に入力がある場合にのみ、細胞を発火

---

<sup>2</sup>左右眼の受容野中心の位置の差で、エネルギーモデルの最適視差を決める場合の解析はQianとZhu(1996)[82]によって行なわれている。

させるという考え方である。ただし、神経細胞への入力値はアナログ値であると考えられるので、これもアナログ AND 回路を用いねばならない。アナログ AND 演算は、2 入力変数の小さいほうの値を出力する演算であると考えることができる。

左右眼からの入力値を  $e_l, e_r$  とする。これは、例えばエネルギーモデルの場合と同様に、網膜入力  $I_{l/r}$  を適当な受容野関数  $f_{l/r}$  でコンボリューションを取った値

$$e_{l/r} = \int f_{l/r}(\xi) I_{l/r}(x - \xi) d\xi \quad (2.4)$$

と考えれば良い。ここで、左右眼受容野  $f_l, f_r$  にはどのような関数を用いててもよい。従って、生理学的に得られたデータに良くフィットする Gabor 関数を用いることも可能である。細胞の入出力特性として、関数

$$\varphi[u] = \begin{cases} u & (u > 0) \\ 0 & (u \leq 0) \end{cases} \quad (2.5)$$

を用いる。単に両眼からの入力が加算されるだけの単純型両眼性細胞の出力を

$$r_s = \varphi[e_l + e_r] \quad (2.6)$$

で定義する。また、優位眼からの片眼刺激に対しては反応するが、両眼刺激を与えた場合は非優位眼からの抑制性入力によって反応が抑制されるという特性を持つゲート細胞を

$$r_{g,l/r} = \varphi[e_{l/r} - w \cdot e_{r/l}] \quad (2.7)$$

で定義する。ここで、抑制性結合の結合荷重は  $w \leq 1$  とする。 $r_{g,l}$  は左眼が優位眼のゲート細胞、 $r_{g,r}$  は右眼が優位眼のゲート細胞である。

すると、上記の三つの両眼性細胞から入力を受ける細胞の出力

$$r_c = \varphi[r_s - (r_{g,l} + r_{g,r})] \quad (2.8)$$

は、 $e_l$  と  $e_r$  のアナログ AND 演算の結果になっている。これだけでは、 $r_c$  は単純型細胞の特性しか持たない。しかし、複数の  $r_s, r_{g,l/r}$  の細胞組から入力を受けることによって、複雑型細胞の特性を得ることができる。この細胞  $r_c$  が、tuned cell のモデルになる。図 2.3(a) は、アナログ AND 回路で実現される tuned cell の模式図である。

ゲート細胞の抑制性結合が  $w = 1$  のときには、 $r_c = 2 \cdot \min[e_l, e_r]$  となる。tuned cell は両眼から大きな入力を受けているとき、つまり入力刺激が最適位置にあるとき、より大きな出力を出すという特性を持っている。 $w < 1$  のときは、AND 回路

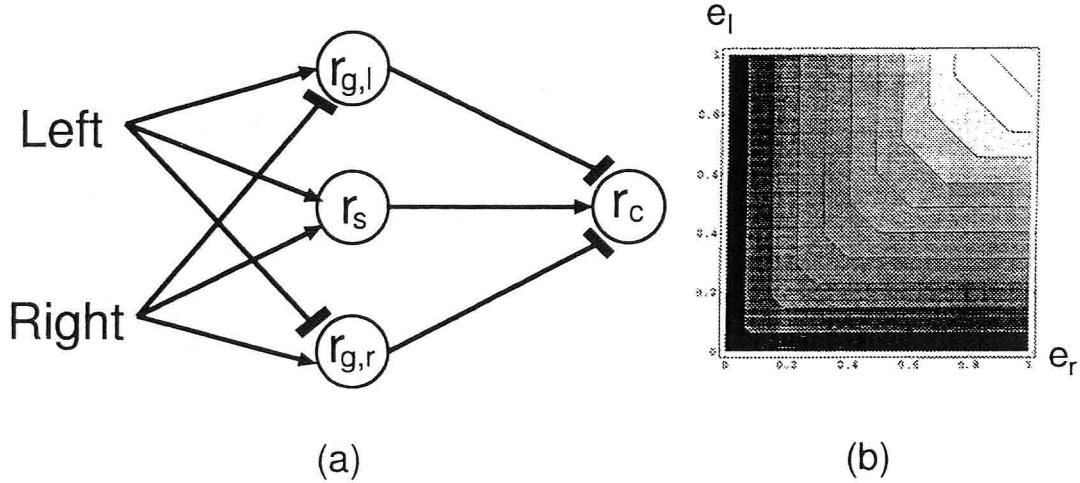


図 2.3: (a) tuned cell の基本回路構成. (b) 入出力特性. 明るい色が大きな出力を表す.

の基本的な特性は保持したまま、両眼入力の最小値  $\min[e_l, e_r]$  が等しくても総和  $e_l + e_r$  が大きいときにより大きな出力をするという特性を得ることができる。これは、左右眼からの入力値にあまり差がない ( $w < e_l/e_r < 1/w$ ) とき、抑制性入力を送るゲート細胞  $u_{g,l/r}$  が両方とも発火するためである。 $w = 0.8$  としたときの入出力特性を図 2.3(b) に示す。これは、左右眼からの入力値  $e_l, e_r$  をそれぞれ縦軸と横軸にとって、そのときの  $r_c$  の出力の大きさをグレースケールで表したものである。

### 2.3.2 near cell, far cell の基本回路

tuned cell は両眼の特定の位置に同時に入力があったときに発火するので、2入力に大きな値が与えられているとき出力するアナログ AND 回路でモデル化できた。これに対し、near cell, far cell は広い範囲の視差に対して発火することが知られている。

これらのモデルとしてまず考えられているものに、様々な視差に選択性を持つ tuned cell からの入力を受けるモデルがある。これによって、確かに広い範囲の視差に対して反応する細胞が得られるが、生理学的な知見では tuned cell が検出できる視差は、near cell, far cell に比べて視差 0 の近傍の狭い範囲に限られている。従って、このモデルは大きな視差にも反応する near cell, far cell のモデルとして適切ではないと考えられる。さらに、特徴選択性の低い tuned cell を near cell, far

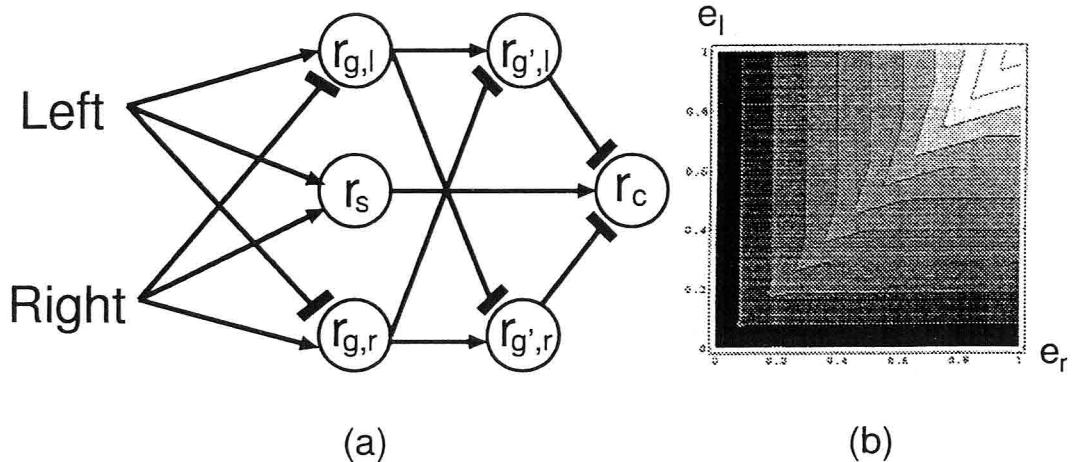


図 2.4: (a)near cell, far cell の基本回路構成. (b) 入出力特性. 明るい色が大きな出力を表す.

cell のモデルとする考え方もある [49]. これによって、対応する特徴の位置が視差を増減させる方向にずれても出力する細胞を構成することができる. しかし、このモデルは、左右眼の入力が対応点かどうかを判断する能力も低くなると考えられる. つまり、特徴選択性が低いために、対応しないパターンを入力した時でも発火することがある. 最後に、エネルギーモデルで考えられているように、最適空間周波数の低い tuned cell を near cell, far cell とみなす考え方がある. このモデルでは、tuned cell と near/far cell には本質的な性質の違いはなく、単に最適空間周波数の違う同じ細胞であるとみなすことになる. しかしこのモデルでは、tuned cell と near/far cell の受容野の幅はほとんど同じであるという知見 [12] を説明できない.

ここで提案するモデルは、near cell, far cell は、tuned cell とは異なる性質を持っている細胞と考える. そして、near cell, far cell は、tuned cell より視差選択性の幅が広い細胞であると考える. しかし、単に特徴選択性を低くしただけでは、視差選択性の幅は広くなるが、対応しない入力が与えられたときにも発火する. near cell, far cell の神経回路に必要な条件は、両眼から入力のあるときにしか反応しないという AND 回路の基本的な機能の他に、両眼からの入力値が等しいときは大きな出力を出すという機能である. 両眼入力  $e_l, e_r$  の値は、左右眼への入力位置および入力パターンによる. すると、入力像の視差が変化してパターンが最適な位置からずれているときは  $e_l, e_r$  は両方とも小さな値になるが、左右眼からの入力が対応する点ならばほぼ  $e_l = e_r$  となる. 左右眼に対応しないパターンが与えられたときは  $e_l \neq e_r$  となる. 従って、near cell, far cell は、両眼からの入力値が等し

ければ大きな出力を出すという特性が必要であることがわかる。

最初の層の細胞として, tuned cell のモデルと同じく, 単に両眼入力が加算される両眼性細胞  $r_s$  と, それぞれ左右眼を優位眼とするゲート 細胞  $r_{g,l}$ ,  $r_{g,r}$  を考える。ゲート 細胞の抑制性結合の結合荷重は  $w < 1$  とする。

図 2.3(a) の tuned cell のアナログ AND 回路に, 一方のゲート 細胞から興奮性入力を, もう一方のゲート 細胞からは抑制性入力を受け取る細胞

$$r_{g',l/r} = \varphi[r_{g,l/r} - r_{g,r/l}] \quad (2.9)$$

を付け加え, 図 2.4(a) で表されるような回路を考える。

この回路の出力は

$$r_c = \varphi[r_s - (r_{g',l} + r_{g',r})] \quad (2.10)$$

で表される。図 2.4(a) の回路で  $w = 0.8$  としたときの入出力特性を図 2.4(b) に示す。この回路は  $e_l \neq e_r$  のときは出力が強く抑制されるという特性を持つ。なぜなら, 左右眼からの入力値にあまり差がない ( $w < e_l/e_r < 1/w$ ) とき, ゲート 細胞  $r_{g,l/r}$  は両方とも発火する。従って,  $e_l = e_r$  のときは, 細胞  $r_{g',l/r}$  への入力が互いに打ち消されて,  $r_c$  へ抑制性入力が全く送られなくなるためである。よって, この回路構造を用いて near cell, far cell をモデル化する。

以上を基本的な考え方として構成した階層ネットワークのアルゴリズムを次節で示す。図 2.3(a) で示した tuned cell の神経回路モデルも, ゲート 細胞から入力を受ける中継細胞を付け加えれば, near cell, far cell と同じ階層構造を持った神経回路モデルで形成することができる。次節では, near cell, far cell 形成回路と tuned cell 形成回路は同じ階層構造にしている。

## 2.4 視差検出細胞の階層型神経回路モデル

前節で, アナログ AND モデルの基本的な考え方を提案した。このモデルでは, tuned cell は, 両眼から興奮性入力を受ける両眼性細胞から興奮性入力を, 片眼から興奮性入力を受けるゲート 細胞から抑制性入力を受けることによって実現できる。これによって, 片眼のみから大きな入力を受ける場合には発火しないという性質が得られる。near cell, far cell は, 左右眼優位のゲート 細胞の出力を互いに打ち消し合うことによって実現できる。これによって, 両眼に同程度の入力が与えられたときには抑制されなくなる。

この節では, 前節で述べたアナログ AND 回路の基本的な考え方に基づいて, 網膜から V1 の両眼性細胞までの階層型神経回路モデルを提案する。このモデルには,

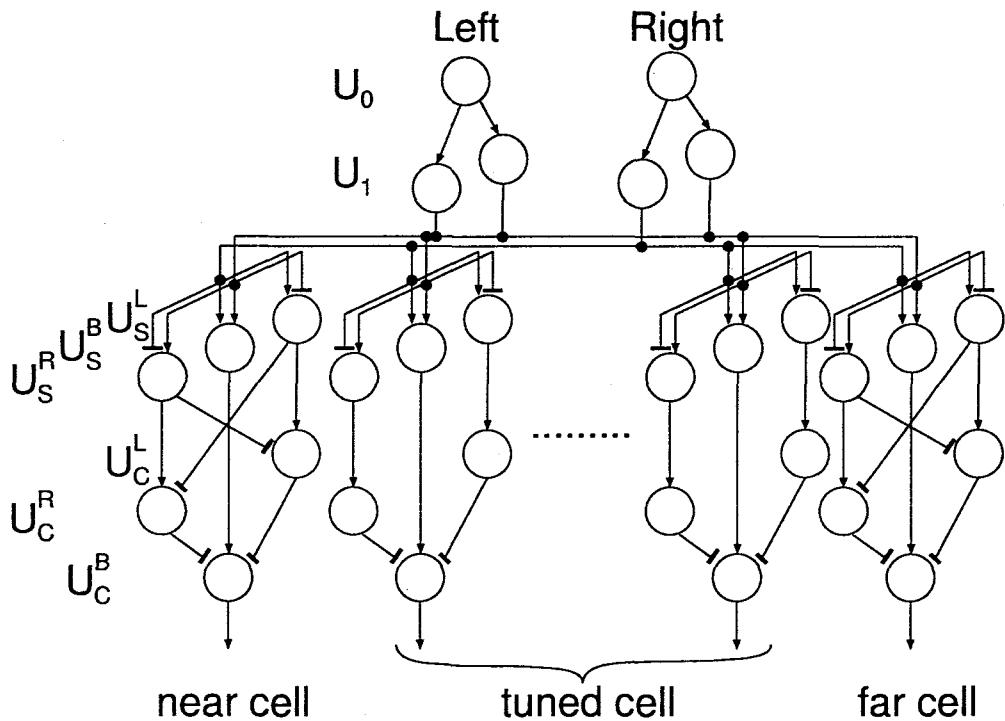


図 2.5: 階層神経回路モデルの模式図.

脳の階層構造のモデルとして提案されている Fukushima [13, 14] の Neocognitron 型の神経回路モデルを用いた、ネットワークの構造を図 2.5 に示す。視差検出細胞には、near cell, far cell と、 $0, \pm 1, \pm 2$  画素の視差を検出する tuned cell を用意した。以下、各層の説明を行なう。

### 光受容細胞層 ( $U_0$ )

$U_0$  は入力層で、左右眼の光受容細胞に相当する。入力としてはグレースケール画像を考え、左眼網膜の位置  $x$  の出力を  $u_0^L(x)$ 、右眼網膜の位置  $x$  の出力を  $u_0^R(x)$  とする。

### 神経節細胞層 ( $U_1$ )

$U_1$  は左右眼の神経節細胞に相当する細胞で、それぞれオン中心型細胞 (on-center cell), オフ中心型細胞 (off-center cell) を持つ。この細胞の出力  $w_1^\mu(x, \kappa)$  は次式で

定義する。

$$u_1^\mu(\mathbf{x}, \kappa) = \varphi \left[ \sum_{\xi} g(\xi, \kappa) \cdot u_0^\mu(\mathbf{x} + \xi) \right] \quad (2.11)$$

ここで、 $\mu$  は左右眼網膜を表し、 $\mu = L, R$  である。 $\kappa$  は神経節細胞の特性を表し、 $\kappa = 1$  のときオン中心型細胞、 $\kappa = -1$  のときオフ中心型細胞を表す。 $g(\xi, \kappa)$  は神経節細胞の受容野形状を表し、DOG(Difference Of Gaussian) 関数

$$g(\mathbf{x}, \kappa) = \kappa \cdot \left\{ \frac{1}{2\pi\sigma_e^2} \exp \left( -\frac{\mathbf{x}^2}{2\sigma_e^2} \right) - \frac{1}{2\pi\sigma_i^2} \exp \left( -\frac{\mathbf{x}^2}{2\sigma_i^2} \right) \right\} \quad (2.12)$$

で近似できる。

### 両眼性細胞層 ( $U_S$ )

$U_S$  は V1 の両眼性単純型細胞に相当する。この層の細胞には Neocognitron 型の S 細胞 [13, 14] を両眼性に拡張したものを用いた。S 細胞は単純型細胞のモデルであり、教師あり、または教師なし学習のいずれによっても、適当な特徴へ選択的に反応する特性を学習できる。またパラメータの設定によって特徴選択性も制御できる。両眼性の S 細胞の出力  $u_S(\mathbf{x}, k)$  は、抑制性入力を送る V 細胞の出力  $u_V(\mathbf{x}, k)$  を介して、次のように表せる。

$$u_S(\mathbf{x}, k) = r \cdot \varphi \left[ \frac{1 + \sum_{\mu=L,R} \sum_{\kappa} \sum_{\xi} a^\mu(\xi, \kappa, k) \cdot u_1^\mu(\mathbf{x} + \nu^\mu + \xi, \kappa)}{1 + \frac{r}{1+r} \cdot b(k) \cdot u_V(\mathbf{x})} - 1 \right] \quad (2.13)$$

$$u_V(\mathbf{x}, k) = \sqrt{\sum_{\mu=L,R} \sum_{\kappa} \sum_{\xi} c(\xi) \cdot \{u_1^\mu(\mathbf{x} + \nu^\mu + \xi, \kappa)\}^2} \quad (2.14)$$

ここで、 $k$  は S 細胞層の細胞面(最適方位)を表す番号である。 $\nu^L, \nu^R$  は、S 細胞の位置と、左右眼それぞれの前層への結合領域の中心とのずれを表す。左右眼の受容野のずれは、これを用いて  $\nu^L - \nu^R$  と表せる。また、 $a^\mu(\xi, \kappa, k)$  は左右眼それぞれの  $\kappa$  の細胞面からの興奮性可変結合の強度、 $b(k)$  は V 細胞からの抑制性可変結合の強度、 $c(\xi)$  は前層から V 細胞への興奮性固定結合の強度を表し、2 次元 Gauss 関数で定めている。 $r$  は、V 細胞からの抑制性入力の効き方を制御する正の定数で、S 細胞の特徴選択性を制御できる。

S細胞の可変結合の強度  $a^\mu(\xi, \kappa, k)$ ,  $b(k)$  は、教師あり学習では、次のように獲得できる。まず、教師あり学習の場合は winner-take-all 型の競合学習を用い、細胞面の中で学習させたい細胞を“核細胞”として指定する [13]。学習させたいパターンを入力したとき、核細胞として  $u_S(\hat{x}, \hat{k})$  を選んだとすると、これと同一の細胞面に含まれる全ての細胞の可変結合の強度を次式に従って強化する。

$$\Delta a^\mu(\xi, \kappa, \hat{k}) = q \cdot c(\xi) \cdot u_1^\mu(\hat{x} + \nu^\mu + \xi, \kappa) \quad (2.15)$$

$$\Delta b(\hat{k}) = q \cdot u_V(\hat{x}) \quad (2.16)$$

$U_S$  には 3 種類の細胞が存在する。 $U_S^B$  は通常の両眼性細胞のモデルである。この細胞は両眼から興奮性入力を受ける、単眼入力でも発火できるが両眼入力ではより大きな反応を示す。 $U_S^L$ ,  $U_S^R$  はそれぞれ、左眼および右眼を優位眼とするゲート細胞のモデルである。この細胞は優位眼から興奮性入力、非優位眼から抑制性入力を受ける、優位眼からの入力を非優位眼がゲートする。

この S 細胞は、学習させるパターンによって通常の両眼性細胞かゲート細胞に分かれる。今、ベクトル表示を用いて、入力パターンを  $p$ 、学習パターンを  $P$  とする。これは、左右眼それぞれの入力パターンを  $p^L$ ,  $p^R$ 、学習パターンを  $P^L$ ,  $P^R$  とするとき、両ベクトルをつないでそれぞれ 1 つのベクトル  $p = [p^L, p^R]$ ,  $P = [P^L, P^R]$  として表したものである。すると、(2.13) 式は次のように近似できる [14]。

$$u_S = (1+r) \cdot \varphi \left[ s - \frac{r}{1+r} \right] \quad (2.17)$$

$$s = \frac{(P, p)}{\|P\| \|p\|} \quad (2.18)$$

ここで、 $s$  は  $P$  と  $p$  の内積を両ベクトルのノルムで正規化したものと解釈できる。従って、多次元ベクトル空間内で考えると、 $s$  は 2 つのベクトル  $P$  と  $p$  の方向余弦を計算していることができる。 $u_S > 0$  となるのは、(2.17) 式の [] 内が正になるとき、即ち

$$s > \frac{r}{r+1} \quad (2.19)$$

のときである。多次元ベクトル空間内で考えると、入力ベクトル  $p$  の方向が、学習ベクトル  $P$  の周りの  $s = \cos \theta > r/(1+r)$  を満たす範囲内にあるときに  $u_S > 0$  になるということができる。特徴選択性を制御する定数  $r$  を大きくすると、(2.19) 式の閾値  $r/(1+r)$  の値は 1 に近づいていく。このとき、パターン  $p$  と  $P$  の類似度  $s$  が非常に高くないと、 $u$  は  $p$  に対して反応しなくなる。つまり、 $r$  を大きくすると、学習パターン  $P$  だけに選択的に反応する傾向が強まっていく。

学習パターンとして両眼入力  $\mathbf{P} = [\mathbf{P}^L, \mathbf{P}^R]$  を用いたとき,  $r$  が十分大きければ, 入力として学習パターンの片眼成分  $\mathbf{p} = [\mathbf{P}^L, \mathbf{0}]$  または  $\mathbf{p} = [\mathbf{0}, \mathbf{P}^R]$  を与えても, 両眼の学習パターン  $\mathbf{P} = [\mathbf{P}^L, \mathbf{P}^R]$  との類似度は低いとみなされるので,  $u_S$  は反応しない. これは tuned cell に似た特性だが,  $r$  が大きすぎると入力パターンの変形に弱くなり, 学習パターンとほとんど同じ入力を与えないと発火しなくなるので, 一般的のステレオ画像の検出に用いることはできない.  $r$  をある程度小さくすると, 単眼入力でも発火するが, 両眼入力を与えたときはより大きな値を出力するという, 通常の両眼性細胞のモデルとして適した性質を得ることができると考えられる. しかし,  $r$  を小さくしすぎると, 単眼入力でも非常に大きな反応を起こす上, 対応しないパターンにも発火することになる.

では,  $r$  はどの程度の値に設定すればよいだろうか. 学習パターンの片眼成分  $[\mathbf{P}^L, \mathbf{0}]$  と  $[\mathbf{0}, \mathbf{P}^R]$  は, 多次元ベクトル空間内では直交しているので, 単眼パターン  $\mathbf{P}^L, \mathbf{P}^R$  のノルムがほぼ等しいとすると, 両眼性 S 細胞が獲得した学習パターン  $\mathbf{P}$  と片眼成分とのなす角は  $\pi/4$  となる. 従って, (2.19) 式より  $r/(r+1) < \cos(\pi/4)$ , 即ち  $r < 1 + \sqrt{2}$  を満たすとき, 両眼性 S 細胞は単眼入力でも発火できるようになる. さらに通常の両眼性細胞は, 両眼刺激による出力の促進作用, 即ち左右眼それぞれへの単眼刺激による出力の和より, 両眼刺激による出力の方が大きくなるという特性を持っている. このような特性を持たせるためには, 単眼入力時, 即ち  $s = \cos(\pi/4)$  のとき, 出力が最大出力の半分以下, 即ち  $u_S < 1/2$  になるように  $r$  を設定しなくてはならない. (2.17) 式より, このとき  $r > \sqrt{2}/2$  となる. 以上より,  $\sqrt{2}/2 < r < 1 + \sqrt{2}$  を満たすとき, 両眼性 S 細胞は通常の両眼性細胞のモデルとして用いることができる.

次に, 学習パターンとして片眼入力  $\mathbf{P} = [\mathbf{P}^L, \mathbf{0}]$  または  $\mathbf{P} = [\mathbf{0}, \mathbf{P}^R]$  を与えたとき, これはゲート細胞を形成する. すなわち, 与えられた学習パターンのうち, 何らかの入力が存在したほうの眼は優位眼としてはたらき, 逆の眼は入力によってゲートを行う非優位眼となる. この細胞に両眼刺激を入力パターンとして与えたとき, 非優位眼への入力は抑制性にしか働くなくなるので, 出力はゲートされることになる. ゲート細胞は, 単眼入力や検出したい視差と異なる両眼入力が与えられたときなどに発火する. 従って, 通常の両眼性細胞の出力をゲート細胞の出力でゲートすることによって, 入力パターンのうち検出したい視差を持った部分だけが得られる.

シミュレーションでは, 3 方位のエッジを学習パターンに用い, 逆符号のエッジも含めて 6 種類のパターンを学習させた. これによって, 対応可能な特徴を減らすことができ, 偽対応が減少する.

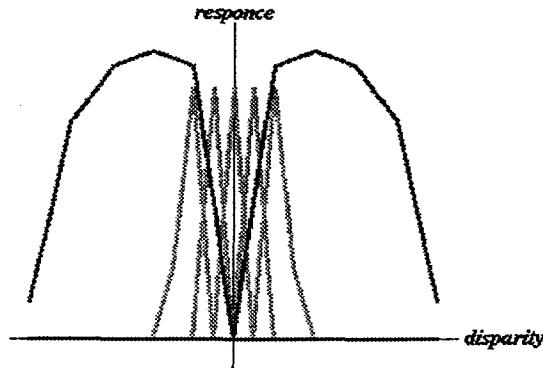


図 2.6: 視差検出細胞の視差感受性曲線.

### 視差検出細胞層 ( $U_C$ )

$U_C$  は視差検出細胞に相当し、複雑型細胞の特性を持っている。この層の細胞にはネオコグニトロン型の C 細胞 [13, 14] を用いた。C 細胞は複雑型細胞のモデルである。 $U_C$  にも 3 種類の細胞が存在する。 $U_C^L, U_C^R$  は前層のゲート細胞から入力を受け、それぞれ優位眼が異なる。図 2.5 のように、tuned cell と near/far cell ではこの細胞への入力が異なっている。これらの細胞は、単眼入力や、検出したくない視差を持つ両眼入力が与えられたとき発火する。従って、tuned cell においてはそれぞれ優位眼の異なるゲート細胞、near/far cell においてはそれぞれ優位眼の異なる monocular far/near cell に相当すると考えられる。ただし、検出したくない視差のとき発火するので、near cell に結合しているのが monocular far cell、far cell に結合しているのが monocular near cell になる。 $U_C^B$  は tuned cell, near/far cell といった視差検出細胞に相当する。この細胞は  $U_S^B$  から興奮性入力を受けていると同時に、 $U_C^L, U_C^R$  から抑制性入力を受けており、単眼入力や望ましくない視差を持つ両眼入力であれば反応しない。

また、これらの細胞は偽対応を除去するための相互結合を持っている。今、左右眼で対応を取ることが可能な特徴が、単位面積当たり密度  $\rho$  で入力されているとする。すると、正しい対応点の密度は  $\rho$  であるのに対し、偽対応を含む全ての対応点の密度はわずか  $\rho^2$  である。従って、near cell, far cell のような広い範囲の視差を検出できる細胞は、その出力を比較することによってどちらが正しい奥行きか決定することができる [53]。つまり、near cell, far cell が偽対応も含め、全ての対応点を検出していながらも、正しい奥行きを有している細胞の方からより大きな出力を得られると期待できる。さらに、特徴選択性が低いながらも、S 細胞によってある程度対応可能な点を減らすことができるので、実際には、偽対応も含

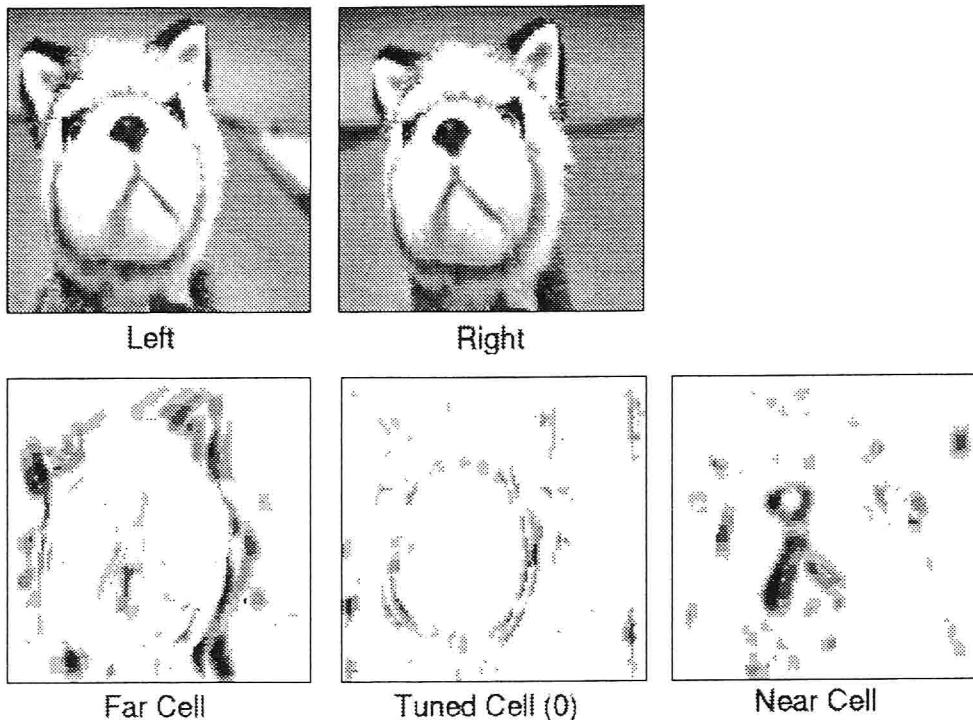


図 2.7: 犬のぬいぐるみのステレオ写真.

め検出される全ての対応点の密度は  $\rho^2$  より低い。従って、near cell, far cell については、同じ位置に受容野中心を持つ near cell と far cell とで競合を行っている。tuned cell どうしについても同様に競合を行なっている。

### シミュレーション

このモデルで構成された near cell, far cell, 及び tuned cell のモデルの視差特性を図 2.6 に示す。これは、学習パターンを入力として与えたときの出力をプロットしたものである。

また、犬のぬいぐるみのステレオ写真を入力として与えたときの出力を図 2.7 に示す。犬の口のまわりにある円形の輪郭のあたりに注視面があり、ここが視差 0 になっている。それぞれの視差に対して、最適方位の異なった複数個の細胞面が用意されている。この図の出力結果は、同じ視差を検出している細胞の出力を最適方位に関係なく全て合計したものである。視差 0 画素を最適視差を持つ tuned cell のモデルの出力は、この円形の輪郭を検出している。near cell のモデルは、注

視面より近くにある鼻と口の部分を, far cellのモデルは, 注視面より遠くにある顔の外側の輪郭と耳の部分をそれぞれ検出している。

## 第3章 両眼視におけるオクルージョンの拘束条件

前章で、両眼視の情報処理の基礎となる両眼性細胞について論じ、網膜からV1までの階層モデルを提案した。V1以降の初期視覚領野では、これらの両眼性細胞から得られる視覚情報を基に、外界の構造復元を行なっている。本論文で特に重要なのは、両眼視差のみではなく、両眼間非対応点も奥行手がかりとして用いる点である。そのためのモデルを考える際には、まず、どのような論理で、何が計算されなくてはならないかという、計算理論を明確にする必要がある。この章では、非対応点も奥行手がかりとして用いるための、両眼視の拘束条件を提案する。3.1節では、両眼視の幾何学的な解析に基づいて、両眼視モジュールにおけるオクルージョンの物理的な拘束条件を提案する。そして、このオクルージョンの拘束条件によつて、非対応点が存在するとき、推定される面の取り得る形状を限定できる、つまり非対応点が構造復元の手がかりとなり得ることを述べる。さらに、非対応点から得られる手がかりの能力や限界を明らかにする。続いて、3.2節で、Nakayamaらの示したda Vinciステレオグラムの知覚についての心理物理学的な知見に基づいて、非対応点の手がかりの限界を克服するために脳が用いていると考えられる仮説を提案する。da Vinciステレオグラムでは、非対応点の手がかりが奥行知覚に大きな影響を与えてるので、脳で用いられている仮説を探るのに非常に適している。第3.3節では、オクルージョンが存在する特殊なステレオグラムにおいても、本章で提案する拘束条件が矛盾しないことを示す。ところで、様々な視覚環境を考慮に入れると、非対応点はオクルージョン以外の原因で生じることもある。その原因の一つに、逆光がある。第3.4節では、脳は逆光を考慮しておらず、逆光が原因で生じる非対応点もオクルージョンによって生じているとみなしていることを示す心理物理学的知見を詳解する。

### 3.1 オクルージョンの拘束条件と奥行拘束領域

第1章で述べたように、両眼視モジュールでは、視差情報を与える両眼間対応点のみではなく、両眼間非対応点も奥行手がかりとして用いられている。この非

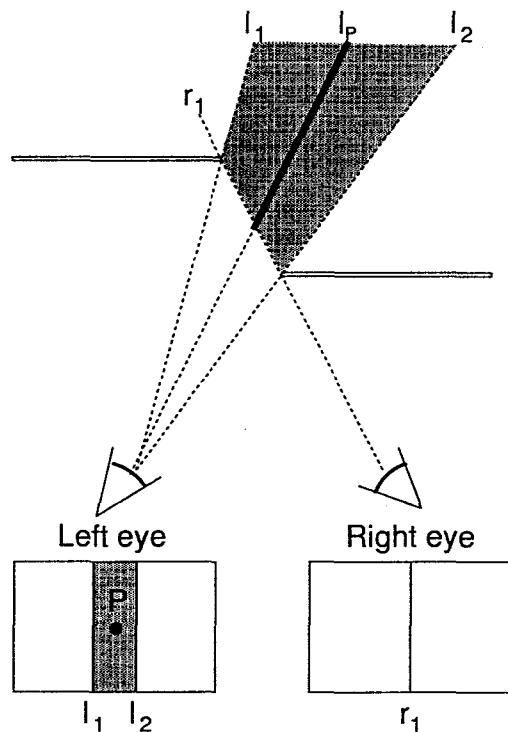


図 3.1: 上図: 奥行の異なる二つの面が存在する状況の鳥瞰図. 下図: 生成される両眼像.  $l_1l_2$  間に存在する点は、左眼のみに見える非対応点になる.

対応点は、面が遮蔽関係にあるということを示すオクルージョン手がかりになる。そして、オクルージョンの幾何学的な条件に矛盾しない面が再構成されると考えられる。推定される外界の構造は、入力画像に存在する非対応点を生じさせることができ可能な構造に限定される。

両眼間非対応点がどのように奥行手がかりとして用いられ、再構成される構造がどれだけ制約されるのかを理解するためには、物理的な光学に従ったオクルージョンの幾何学を解析することが重要である。

まず、図 3.1 に示すような、両眼から見える、奥行の異なる二つの面が存在する状況を考える。図 3.1 の上図はこの状況の鳥瞰図を表し、下図は網膜像を表す。下図に示す右眼網膜像では、 $r_1$  のところで面の不連続が生じる。これに対し、左眼網膜像では、 $l_1l_2$  間 (灰色の領域) に存在する点は左眼からしか見えない左眼非対応点になる。

いま、このような左眼非対応点を生成する外界の点が位置することが可能な範囲を考えてみる。左眼網膜像上に非対応点を生成する点は、図 3.1 の上図の左眼からの視線  $l_1l_2$  間に存在する。しかし、オクルージョンの幾何学から、実際に左眼に

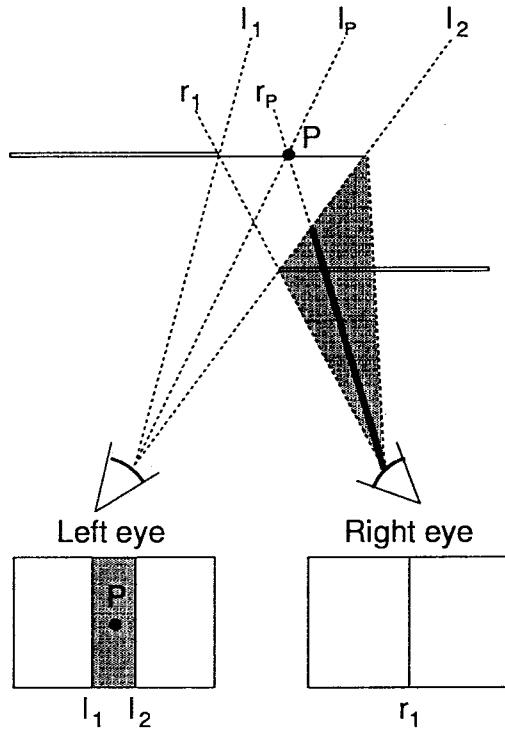


図 3.2: 上図: 奥行の異なる二つの面が存在する状況の鳥瞰図. いま、左眼非対応点  $P$  の奥行が左眼の視線  $l_P$  と右眼の視線  $r_P$  の交点の位置に推定された場合を考える. 下図: 生成される両眼像.

非対応点を生成する点の存在する範囲はより制限される. まず、(1) 左眼非対応点は右眼からは遮蔽されて見えないため、左眼非対応点を生成する点は手前の遮蔽面より奥に存在する. さらに、(2) 奥に存在する面は右眼から遮蔽されることなく見えねばならないため、左眼非対応点を生成する点は右眼の視線  $r_1$  より左に制限される. 図 3.1 の場合、この(1), (2) のどちらか一方の条件を用いただけでも、左眼非対応点を生成する点の存在する範囲は上図の灰色の領域中に限定される. この灰色の領域のような、非対応点を生成可能な点が存在し得る領域を、奥行拘束領域(depth constraint zone)と呼ぶ [64]. よって、図 3.1 の下図に示す左眼非対応点  $P$  を生成する点が位置することが可能な範囲は、網膜像上の点  $P$  から伸びる左眼の視線  $l_P$  上の一部(太線で示した半直線)に制限される.

次に、図 3.2 のように、何らかの条件によって左眼非対応点  $P$  の奥行が上図の黒丸の位置、即ち奥の面と同じ奥行に推定されている場合を考える. 左眼非対応領域  $l_1l_2$  間にある他の点も、点  $P$  と同じ奥行に推定されているものとする.

このとき、点  $P$  を右眼から遮蔽している遮蔽物の存在し得る範囲を考えてみる.

オクルージョンの幾何学から、点  $P$  は右眼からは遮蔽されて見えないことが要請されるので、点  $P$  と右眼との間に必ず遮蔽点が存在する。まず、(1) 左眼非対応点は右眼からは遮蔽されて見えないため、遮蔽点は点  $P$  を通る右眼の視線  $r_P$  上の点  $P$  より近い側に存在する。さらに、(2) 左眼非対応点は左眼から遮蔽されることなく見えねばならないため、遮蔽面は左眼の視線  $l_2$  より右に制限される。よって、(1), (2) より点  $P$  を遮蔽する点は、点  $P$  を通る右眼の視線  $r_P$  上の一部（太線で示した半直線）に存在しなければならない。点  $P$  以外の左眼非対応点も考慮すると、遮蔽面は図 3.2 の上図の灰色の領域中に存在しなければならない。この領域を、遮蔽面についての奥行拘束領域と呼ぶ。

以上、具体例を用いて、非対応点や遮蔽物となる対応点の取り得る奥行が奥行拘束領域内に制限できることを述べてきた。ここで、奥行拘束領域を求めるための一般的な方法を与える。

図 3.1 及び図 3.2 の奥行拘束領域を決定するために用いた条件 (1) は、対応点と非対応点の遮蔽関係から導かれるものである。この遮蔽関係の条件を、オクルージョンの拘束条件(occlusion constraint)と呼ぶ。これは次のように表現できる：

オクルージョンの拘束条件：左(右)眼非対応点は遮蔽されて右(左)眼から見ることができないため、左(右)眼非対応点と右(左)眼とを結ぶ視線上に他の点が少なくとも一つ存在する。

これは、非対応点が見えない方の目からの視線に沿った拘束条件になっている。ここで、上で述べた「視線上の他の点」の種類は特に制限されないことに注意すべきである。つまり、必ずしも対応点である必要はない。このオクルージョンの拘束条件によって、非対応点の持つ遮蔽関係の情報を、奥行拘束領域として外界の構造復元に用いることができる。

このオクルージョンの拘束条件を用いて、非対応点及び遮蔽物の存在し得る奥行の範囲が導かれる。この範囲は、非対応点の存在を考慮した一意性の拘束条件を適用することによってさらに狭められることがある。図 3.1 及び図 3.2 の奥行拘束領域を決定するために用いた条件 (2) は、この非対応点の存在を考慮した一意性の拘束条件から導かれる制限である。1.1.2 節で述べた、Marr と Poggio [55] の提案した一意性の拘束条件は、各視線に沿って唯一の点しか存在しないという条件であり、「あらゆる点が遮蔽されずに両眼から見えなければならない」という仮定に基づいている。つまり、これは非対応点の存在を無視したものになっている。非対応点の存在も考慮した、拡張された一意性の拘束条件は次のように表される：

拡張された一意性の拘束条件：両眼間対応点及び左(右)眼非対応点は遮蔽されることなく左(右)眼から見えねばならないため、両眼間対応点または左(右)眼非対応点と、左(右)眼とを結ぶ視線上には他の点は存在しない。

ここで、この条件は、問題の点を通る視線全体にわたる拘束条件ではなく、問題の点より眼球に近い側だけを扱っていることに注意する必要がある。また、上で述べた「視線上の他の点」の種類が特に制限されることにも注意すべきである。この拡張された一意性の拘束条件は、問題の点が左眼非対応点の場合は左眼からの視線のみに、右眼非対応点の場合には右眼からの視線のみにしか適用されない。そして、入力画像上に対応点しか存在しないときは、Marr と Poggio の一意性の拘束条件と同じ働きをする。

オクルージョンの拘束条件は、視線上に他の点が存在することを要請し、奥行拘束領域を設定する。オクルージョンを考慮した一意性の拘束条件は、視線上に他の点が存在しないことを要請するため、この領域を狭められることがある。これらの条件はいずれも、両眼からの視線に沿って物体の可視性を検証する。

但し、非対応点から与えられる情報には限界が存在する。まず、非対応点及び遮蔽物の奥行拘束領域を決定するためには、それぞれ遮蔽物及び非対応点の視差が与えられている必要がある。なぜなら、非対応点か遮蔽物のどちらか一方の3次元空間上の位置を基準にしなければ、オクルージョンの拘束条件によって奥行拘束領域を定めるための視線を引くことができないためである。さらに、一般に、オクルージョンの幾何学によって、非対応点や遮蔽物の取り得る奥行の範囲を制限することはできるが、奥行を一意に決定することはできない。なぜなら、オクルージョンの拘束条件からは面の遮蔽関係しか得ることができず、視差の絶対値が与えられる訳ではないからである。奥行拘束領域を決定するためには視差が与えられている必要があるが、オクルージョンの幾何学からは視差の絶対値が得られないということから、非対応点単独では解が存在しないということがいえる。この意味で、非対応点のみによる奥行知覚は不良設定問題になっている。これが、1.3節で述べた、非対応点が単独で奥行知覚を生じさせることができない理由である。

## 3.2 非対応点の視差推定

前節で、非対応点は、非対応点及び遮蔽物の取り得る奥行の範囲を限定する手がかりになることを述べた。しかし、この物理的な解析から導かれた条件のみを用いていては、非対応点の手がかりには限界が存在することも述べた。

問題の一つは、非対応点や遮蔽物の奥行拘束領域を決定するためには、それぞれ遮蔽物や非対応点の視差が必要なため、非対応点単独では解が存在しない不良設定問題になるということである。よって、奥行知覚を得るために、非対応点以外に視差の絶対値を与える対応点が存在しなければならない。従って、まず、どれだけの視差が与えられれば良設定問題になるのかということを明らかにする必要がある。オクルージョンの存在する最も単純な状況は、奥行の異なる二つの前

額並行面が存在する場合である。この場合、二つの奥行だけが存在する。脳では、この内の遮蔽面の奥行ただ一方が与えられるだけで、非対応点を奥行手がかりとして用いた奥行知覚が可能になる。これが、da Vinci ステレオグラムである。

心理学的には、遮蔽面の奥行さえ与えられれば問題が良設定になるといえる。しかし、物理的な解析によって得られる拘束条件のみでは、非対応点の奥行を制限することはできても一意に決定することはできないので、その意味ではいくら視差が与えられても良設定問題にはならない。これが、もう一つの問題である。脳内では、遮蔽面の奥行が与えられるだけで、非対応点の奥行を一意に決定して問題を良設定にできるような、何らかの仮定が用いられていると考えられる。従って、次に、非対応点の奥行を一意に決定するために、脳がどのような仮定を用いているか明らかにする必要がある。

まず、遮蔽面、及び被遮蔽面の視差ともに十分に与えられている場合を考えてみる。これには、両眼視以外の視覚手がかりを排除するため、例えば図1.3のようなランダムドットステレオグラムを考えてみれば良い。このようなステレオグラムを融合視したとき、非対応点は明らかに奥の面上に存在するように知覚される。このことから、単純に、非対応点の奥行は滑らかさの拘束条件を用いて、奥の面の視差を補間することによって決定されているという考え方ができる。オクルージョンの拘束条件から、非対応点の取り得る奥行は遮蔽面より奥に制限されるので、奥の面の視差で補間するのが滑らかさの仮定に最も適合する。滑らかさの仮定は一つの考え方であるが、他の仮定でも説明できる可能性がある<sup>1</sup>。脳で用いられている仮定を明らかにするためには、心理物理学的な実験に基づく必要がある。

従って、次に、解を得ることのできる最小限の情報が与えられる状況、即ちda Vinci ステレオグラムの知覚の場合を、心理物理学的な知見を基に考えてみる。この状況では、脳で用いられている仮定がより明確な形で現れる。

まず、Nakayama と Shimojo (1990) [64] の、da Vinci ステレオグラムの知覚についての心理物理学実験について述べる。彼らは、da Vinci ステレオグラムの非対応点の視差を人間がどう知覚しているかを調べている。図3.3は、この心理実験に用いたda Vinci ステレオグラムである。上下二つの長方形は両眼対応可能である。上の長方形の横には片眼からしか見えない非対応の棒線が、下の長方形の横には両眼対応する棒線が提示される。被験者は、下の両眼対応する棒線の視差を自由に変化させることができる。実験では、被験者に、下の棒線の視差を、上の非対応の棒線で知覚されている奥行と一致するように調整させる。二つの棒線の提示位置、及び非対応棒線の提示眼を様々に変化させる。ここで、長方形の右端に左眼非対応点を提示した場合、及び左端に右眼非対応点を提示した場合は、幾

---

<sup>1</sup> 例えば、1.1.2節で述べたように、従来の対応点のみを考慮した両眼視のモデルでは、滑らかさの拘束条件の代わりに順序対応の拘束条件を用いることも可能である。

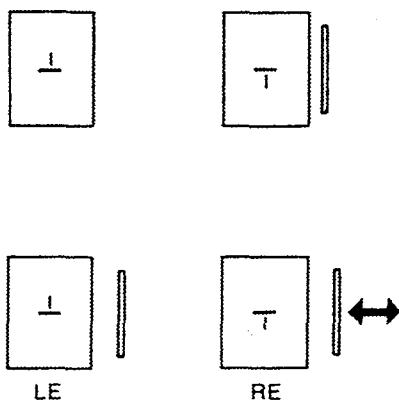


図 3.3: da Vinci ステレオグラムの非対応点で知覚されている奥行を調べるために視覚刺激。下の棒線の視差は被験者が調整可能である。実験では、上の非対応の棒線の奥行に一致するように調整させる。(Nakayama & Shimojo (1990) [64] より転載)

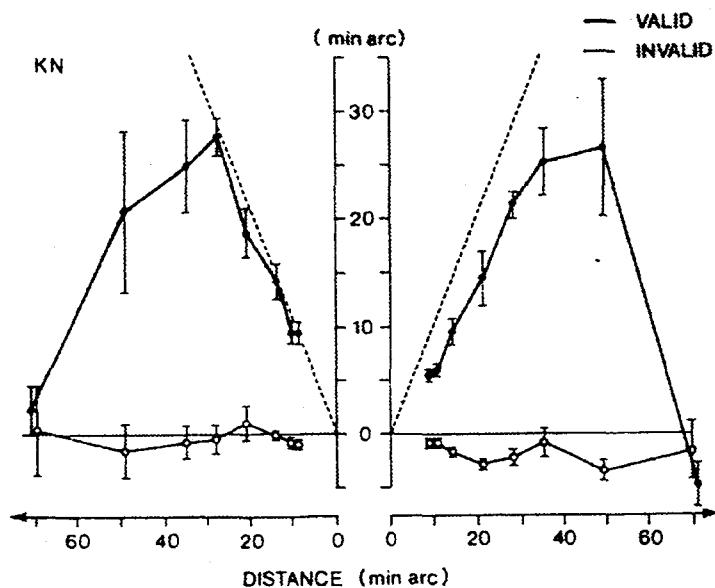


図 3.4: 非対応の棒線で知覚される視差。両眼対応する長方形の左右端からの距離の関数としてプロットされている。点線より上の領域が奥行拘束領域になる。白丸は矛盾した非対応刺激、黒丸は矛盾しない非対応刺激を表す。(Nakayama & Shimojo (1990) [64] より転載)

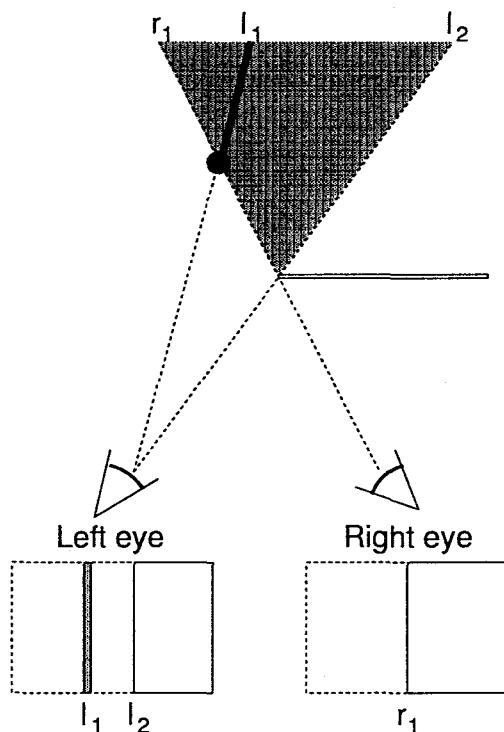


図 3.5: 上図: 両眼から見える面の左側に、左眼からしか見えない非対応の文線が提示されている状況の鳥瞰図。下図: 両眼像。

何学的に矛盾したステレオグラムになる。なぜなら、このステレオグラムには遮蔽物となり得るのは両眼対応する長方形しか存在しないからである。例えば、長方形の右側には遮蔽物となる物体は存在しないため、長方形の右端には本来左眼非対応点は生成し得ない。

図 3.4 はこの心理実験の結果である。グラフの横軸は長方形の端から非対応刺激の位置までの距離(単位は分)、縦軸は非対応刺激で知覚された視差(単位は分)である。右のグラフは長方形の右側に非対応刺激を、左のグラフは長方形の左側に非対応刺激を提示したときの結果である。黒丸でプロットされたデータは幾何学的に矛盾しない非対応刺激が、白丸でプロットされたデータは幾何学的に矛盾した非対応刺激が提示された場合である。グラフに描かれた傾き 1 の点線より上の領域が、オクルージョンの拘束条件から導かれる奥行拘束領域になる。結果は、矛盾した非対応刺激から知覚された視差はほとんど 0、すなわち長方形のと同じ視差であるが、矛盾しない非対応刺激の視差は、距離が 30-40 分までは奥行拘束領域の内で最も近くに見える限界に知覚されていた。

この心理実験から、非対応点の奥行を一意に決定するために、脳でどのような

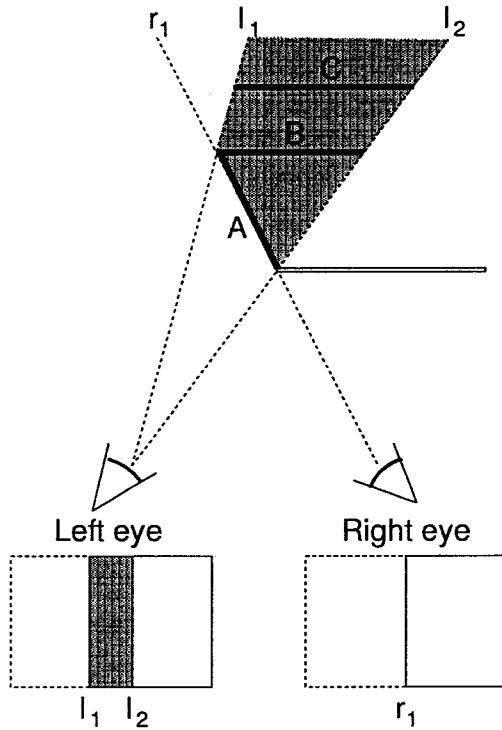


図 3.6: 上図: 両眼から見える面の左側に、左眼からしか見えない幅を持った非対応領域が提示されている状況の鳥瞰図。下図: 両眼像。

仮定が用いられているのかを考察することができる。

まず、図 3.5 に示すような、幾何学的に矛盾しない非対応の棒線  $l_1$  が与えられたときを考える。これは、Nakayama らの実験で用いられた視覚刺激と同じ状況である。オクルージョンの拘束条件から、この非対応の棒線の奥行拘束領域は、図 3.5 上図の太線上に制限される。非対応の棒線の視差手がかりは全く存在しないので、視差が図 3.5 上図の太線上のどこに推定されていても物理的には正しいが、逆にどこに推定されていても根拠が存在するわけではない。このとき、人間は非対応の棒線を図 3.5 上図の黒丸の視差に知覚する。従って、脳で用いられている仮定は、奥行拘束領域内の最も近くに見える視差、即ち遮蔽物に最も近い視差を選ぶようなものであると考えられる。

次に、図 3.6 のように幅を持った非対応領域  $l_1 l_2$  が与えられた場合を考える。これは、図 1.14 のコウモリのステレオグラムと同じ状況である。各々の非対応点について、奥行拘束領域内の最も近くの視差が選ばれるとすると、この非対応領域の奥行として面 A のような傾いた面が選ばれる。しかし、図 1.14 を融合視すれば分かるように、コウモリは傾いて知覚されることではなく、前額並行に知覚される。

従って、様々な前額並行面の候補(面BやC)の中から、可能な限り遮蔽面の近くに位置する前額並行面(面B)が選ばれると考えられる。このことから、脳で用いられている仮定は、前額並行面を好むという滑らかさの仮定の性質を持つと考えられる<sup>2</sup>。そして、複数の前額並行面の候補が存在するときには、奥行拘束領域内の最も近くの面が選ばれる。

最後に、矛盾した非対応点が与えられた場合を考える。このときは、非対応点の遮蔽物となり得る面が存在しない。従って、非対応点の取り得る奥行は、オクルージョンの拘束条件によって特定の奥行拘束領域内に制限されることはない。遮蔽面と被遮蔽面の視差が共に与えられる場合の考察で、脳で用いられている仮定は、被遮蔽面の視差を非対応点に補間するように働くと述べた。今の場合もそれと同様のプロセスで、両眼対応可能な面の視差が非対応点に補間され、対応点と非対応点が同じ視差に知覚されると考えられる。

以上の考察から、脳で用いられている仮定は、奥行拘束領域の中で遮蔽面に近い視差を選び、前額並行面を好むという性質を持っているといえる。これは、MarrとPoggioの両眼視の計算理論[55]や標準正則化理論[5, 58, 77]で用いられている、連続面の局所的な視差の変化で定義される滑らかさの仮定とは異なる。なぜなら、図3.5の遮蔽面と非対応点のように、連続しない離れた点の間でも相互作用する必要があるためである。ここで考えている仮定には、Prazdny[80]の提案した凝集性の原理(coherence principle)が最も適している。凝集性の原理とは、「外界は無秩序に奥行が変化する点から構成されているのではなく、3次元上の特定の位置を占める物体によって構成されている」という原理である。この原理から導かれる凝集性の拘束条件(coherence constraint)は、次のように定義できる：

**凝集性の拘束条件:** 同一の3次元物体に属する点の近傍の視差は類似している。

連続性の拘束条件は、画像上の近傍の点は必ず同一物体に属することを仮定している。これに対し、凝集性の拘束条件は、画像上の近傍点が必ずしも同一物体に属することを要請するわけではない。従って、複数の面が重なり合う場合や、面の不連続の場合なども、例外とすることなく扱える。これは、遮蔽面と非対応点とが重なり合う、オクルージョンの存在する構造を復元するのに適している。凝集性の拘束条件では、近傍の点(隣り合った点だけでなく離れた点も含む)の間で、似た視差を持った点どうしが互いに促進し合い、視差が似ていない場合は単に相互作用しない。これに対し、連続性の拘束条件では、画像上の隣り合った点は視差が大きく異なる場合でも相互作用し、画像上の離れた点の視差は伝播による間

<sup>2</sup> 視差曲率(2次微分)の小さな面を好むという滑らかさの仮定を考えれば、面Aも十分滑らかであるということができる。しかし、両眼視モジュールにおける滑らかさの仮定は視差勾配(1次微分)の小さな面を好むというものであると考えられている。

接的な影響しか与えない。凝集性の拘束条件は、視差に曖昧さのある点では、視差勾配に対して単調減少し、画像上の距離に対しても単調減少するサポート関数(support function)によって最も大きなサポートが与えられる視差を選ぶ、というアルゴリズムで実現できる。この方法は、Prazdny のアルゴリズム以外でも用いられている(例えばPMFアルゴリズム[81]など)。

この凝集性の拘束条件は、前額並行面を好み、かつ、隣り合っていない点にも視差の差が小さくなるような影響を及ぼす拘束条件である。従って、遮蔽面と被遮蔽面の視差が共に与えられれば、離れた面からの影響は相対的に小さくなつて、ほとんど無視できる。しかし、被遮蔽面の視差を決める手がかりが全く存在しない、da Vinci ステレオグラムの場合は、この小さな影響が視差を決める効果を及ぼす。この凝集性の拘束条件を用いると、図3.5や図3.6の場合にも、非対応点の視差として人間の知覚と同じものが選ばれる。ただし、遮蔽面と非対応点は一般に同一の3次元物体に属しているわけではないので、凝集性の拘束条件の本来の意味からは外れているということもできる。しかし、被遮蔽面の視差が全く与えられていないときは、遮蔽面と非対応点が可能な限り凝集して存在するように推定される。

この凝集性の拘束条件は、対応点、非対応点の区別なく適用されることに注意する必要がある。凝集性の拘束条件は、元々対応点のみを考慮した両眼視のモデルで用いられていた。そして、この条件は、非対応点の視差を奥行拘束領域内から一意に決定するためにも用いられるということを述べた。このことは、対応点か非対応点かに関わらず、視差の曖昧さの除去のために、凝集性の拘束条件が用いられることを示している。

対応点の視差は、一意性の拘束条件と凝集性の拘束条件を用いれば推定可能である。従って、ここまで述べた三つの拘束条件で、対応点と非対応点双方を奥行手がかりとして用いることができる。

### 3.3 特殊なステレオグラム

この節では、前節までで提案した拘束条件が、オクルージョンが生じている特殊なステレオグラムに対しても矛盾しないことを示す。

#### 3.3.1 Narrow Occlusion

コンピュータビジョンの分野では、二つの面の視差の差が、手前に存在する遮蔽面の水平方向の幅より大きい状況をnarrow occlusionと呼び、他のオクルージョンの状況と区別することがある。narrow occlusionは、例えばフェンス越しに

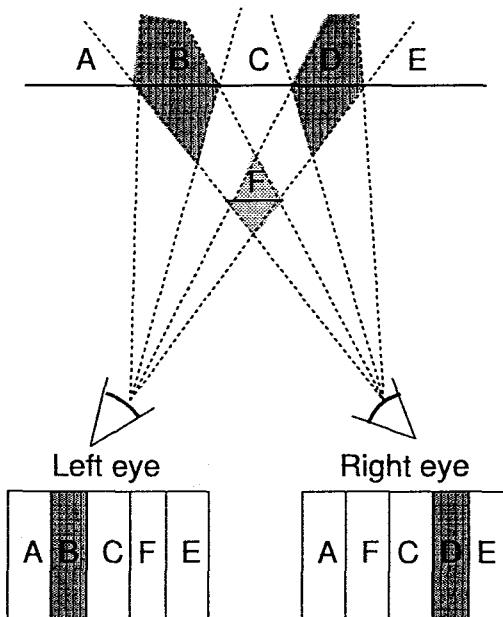


図 3.7: 上図: 細い物体により背景の物体が隠される narrow occlusion の状況の鳥瞰図. 領域 B は左眼のみに見える非対応領域, D は右眼のみに見える非対応領域である. 暗灰色の領域は非対応点の奥行拘束領域, 明灰色の領域は遮蔽面 F の奥行拘束領域を表す. 下図: 両眼像. 灰色の領域は非対応領域を表す. 対応領域 C と F の順序が左右眼で異なっている.

物体を見る場合などに起こり, 擬透明 (pseudo transparency) と呼ばれることがある. この narrow occlusion は, 模式的に図 3.7 のように描ける. この状況の特徴は, 図 3.7 の下図に示されるように, 左右の画像上で対応する領域の順序が入れ替わることである. 従って, 特に順序対応の拘束条件を用いた両眼視アルゴリズムでは問題となり, narrow occlusion は特殊な, 困難な状況として扱われている.

しかし, 3.1 節で提案したオクルージョンの拘束条件は, 面の形状などによらずオクルージョンによって非対応点が生じるときに常に成り立つ物理的な拘束条件であるので, narrow occlusion である場合とそれ以外の場合とを特に区別する必要はない. 前節で述べた方法を用いて, 非対応点の奥行拘束領域は図 3.7 の上図の暗灰色の領域に, 遮蔽面 F の奥行拘束領域は明灰色の領域に決定できる.

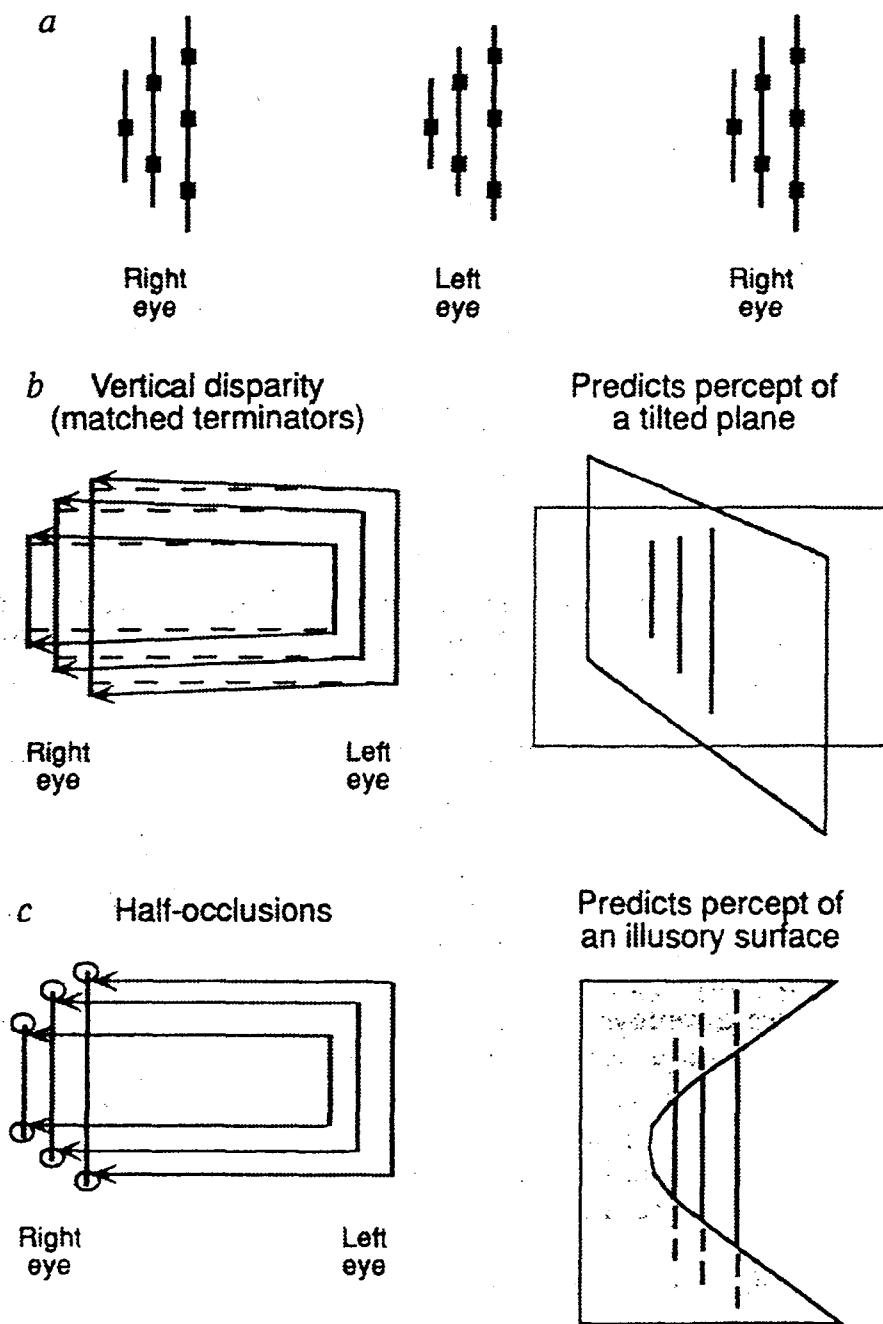


図 3.8: a: 主観的輪郭によって遮蔽面が知覚されるステレオグラム。右眼像の線分は、左眼像の線分よりわずかに長くなっている。b: 端点同士が対応すると考え、垂直視差を考慮した場合に予測される知覚。c: 線分の長い部分が非対応点とみなされ、主観的な遮蔽面が存在すると考えた場合に予測される知覚。(Anderson (1994) [3] より転載)

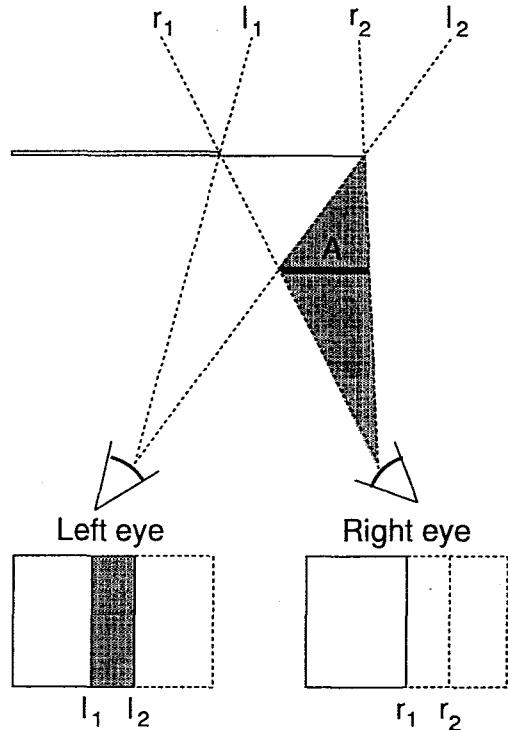


図 3.9: 上図: 遮蔽面の対応点が明示的に与えられない状況の鳥瞰図. 非対応点の視差は, 両眼対応する面と同じ視差に補間される. 下図: 両眼像.

### 3.3.2 Anderson のステレオグラム

第3.2節で紹介した心理実験で示されるように, オクルージョンの幾何学に矛盾する, 遮蔽面が存在しない非対応点は, 近傍の対応点の視差と同じ視差に知覚されてしまう. これは, 既に述べたように本論文で提案する理論で説明可能である. オクルージョンの幾何学に矛盾した非対応点は, オクルージョンの拘束条件によって奥行が拘束されないため, オクルージョンを全く考慮しない他の多くの両眼視モデルと同様, 近傍の対応点の視差と同じ視差で補間される.

Anderson (1994) [3] は, 遮蔽面上の対応点が明示的に与えられない, 図 3.8a のようなステレオグラムを考案している. このステレオグラムには三本の線分が描かれているが, 右眼像の線分が左眼像の線分よりわずかに長くなっている. 両眼像の端点同士が対応すると考え, 垂直視差を考慮に入れると, 図 3.8b のような3次元構造が知覚されるものと考えられる. しかし, 実際に融合視すると, 図 3.8c のように右眼像の線分の長く伸びた部分は非対応領域と知覚され, 主観的な面によって直線が遮蔽されているように知覚される. 図 3.8 の b も c もこのステレオグ

ラムを生成する3次元構造の候補であるが、線分の視差が黒色の小さな四角形で固定されていることから、cの方がより妥当な3次元構造であるとみなされて知覚されたと考えられる。

このステレオグラムは、3.2節の Nakayama と Shimojo (1990) [64] の実験と矛盾しない。なぜなら、確かに遮蔽面になる両眼間対応点は明示的に与えられていないが、線分の端点によって遮蔽面が存在するという手がかりが得られるからである。本論文で提案する理論が、このときどのように働くか考える。図3.9は、この状況を模式的に表したものである。まず、非対応点の視差は対応点と同じ視差に補間されると考えられる。遮蔽面が存在するという情報が得られているので、オクルージョンの拘束条件から、図3.9の灰色の領域のように、遮蔽面の奥行拘束領域が定められる。図3.8Aのステレオグラムでは、端点をつないで構成される主観的輪郭が両眼対応して、遮蔽面の境界の視差が得られる。そして、主観的輪郭の視差と滑らかにつながり、奥行拘束領域内に存在する、図3.9の面Aが遮蔽面として選ばれる。

図3.8のステレオグラムは、端点の手がかりから遮蔽面が存在することが示唆されるため、遮蔽面上の両眼間対応点が明示的に存在しなくても図3.8cのような知覚が得られた。しかし、この知覚を説明するためには、他の初期視覚モジュールも考慮に入れなくてはならない。従って、本論文では以降、このような他の要素がからむステレオグラムは考慮しない。

### 3.4 オクルージョン以外の原因で生じる非対応点

オクルージョンによって生じる非対応点のみを考慮したときには、遮蔽面の右側に左眼非対応点が存在したり、遮蔽面の左側に右眼非対応点が存在したりするステレオグラムは矛盾しているとみなされる。しかし、本当にこのような両眼像は物理的に妥当な光学過程によって生成し得ないのでしょうか。いま、図3.10のような状況を考える。この状況を逆光下で両眼視するという特殊な条件下では、左眼からは奥の面と手前の物体の両方のシルエットを見ることができるが、右眼では手前の物体は面の影になって見ることができない。このときの両眼像は、図3.11A(1)のようになっている。オクルージョンの幾何学からは、このような非対応点を生成する両眼像は生じ得ない。しかし、逆光下で両眼視する場合の「シルエットの幾何学」を考慮すると、この両眼像は物理的に矛盾しない。

両眼間非対応点には、オクルージョンが原因で生じるもの以外にも、逆光によって生じるもののが存在する。本論文で提案する理論では、非対応点を特に区別することなく、全て遮蔽関係の情報をもつたオクルージョン手がかりとみなしている。しかし、純粹に理論的には、オクルージョンの幾何学的制約からオクルージョン

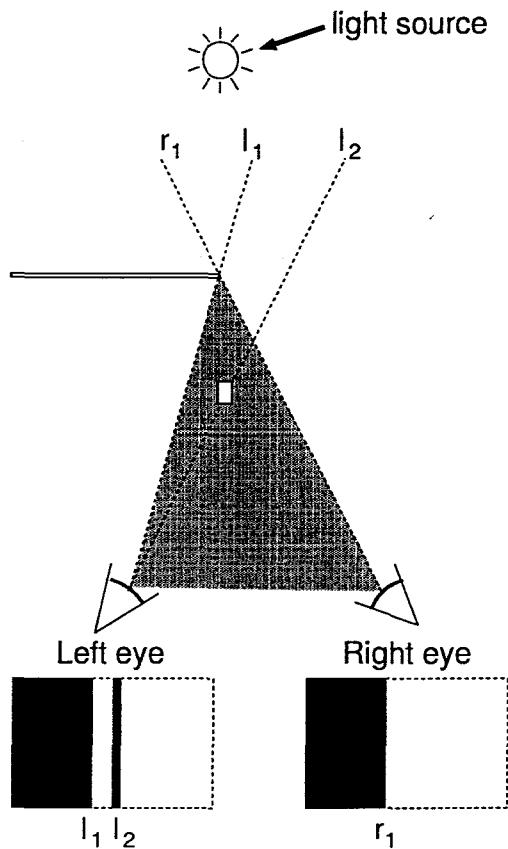


図 3.10: 上図: 逆光の状況の鳥瞰図、下図: 両眼像。

の拘束条件を導いたのと同じ方法で、「シルエットの幾何学」からもシルエットの拘束条件を導くことができる。図 3.10 の灰色の領域は、「シルエットの拘束条件」にから導かれる非対応点の奥行拘束領域である。実際の脳の両眼視モジュールでは、非対応点はオクルージョンによって生じたものと、逆光によって生じたものとに区別されているのだろうか。そして、逆光によって生じた非対応点には、シルエットの拘束条件が用いられているのだろうか。

Nakayama と Shimojo (1990) [64] は、脳は逆光の場合を考慮していないことを心理物理学的に示している。図 3.12 はこのことを示す心理実験の結果である。この実験は、図 3.4 と同様の方法で行なわれている。点線 O より上がオクルージョンの拘束条件から導かれる奥行拘束領域、点線 S より上が「シルエットの拘束条件」から導かれる奥行拘束領域になる。左側のグラフは、長方形が白色で、逆光下とはみなされない図 3.11B のステレオグラムを用いたときの結果で、黒丸がオクルージョンの幾何学に矛盾しない場合 (図 3.11B(2))、白丸が矛盾する場合 (図 3.11B(1)) で

Cases	Left Eye	Right Eye	Occlusion Constraints	Silhouette Constraints
(1)			Invalid	Valid (Front)
(2)			Valid (Back)	Invalid
B				
(1)			Invalid	Invalid
(2)			Valid (Back)	Invalid

図 3.11: 図 3.12 の心理実験で用いたステレオグラム. (Nakayama & Shimojo (1990) [64] より転載)

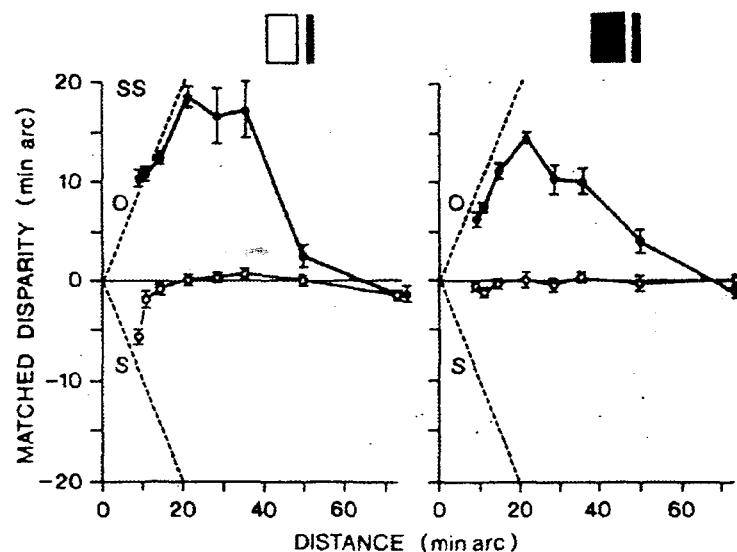


図 3.12: シルエットの拘束条件が存在しないことを示す心理実験. 右のグラフの黒丸がオクルージョンの幾何学に矛盾しない場合, 白丸がシルエットの幾何学に矛盾しない場合. 左のグラフは対照実験. (Nakayama & Shimojo (1990) [64] より転載)

ある。この結果は、図3.4で得られた結果と同様である。右側のグラフは、長方形と棒線が共に黒色で、逆光下にあるとみなすこともできる図3.11Aのステレオグラムを用いた場合の結果である。黒丸のデータがオクルージョンの幾何学には矛盾するがシルエットの幾何学には矛盾しない場合、白丸のデータがオクルージョンの幾何学には矛盾せずシルエットの幾何学には矛盾する場合である。結果は左のグラフと定性的に同じものであり、逆光によって非対応点が生じているとされる白丸の場合の視差はほぼ0になっている。つまり、脳内では「シルエットの幾何学」は考慮されていないということになる。言い換えると、脳内では常に非対応点をオクルージョンの手がかりとして用いているとみなすことができる。

画像処理のアプリケーションを考える際には、この「シルエットの幾何学」も考慮すべきかもしれない。しかし、本論文のように脳の両眼視モジュールの理論を考える際には、考慮する必要はない。

# 第4章 オクルージョン手がかりを用いた両眼視アルゴリズム

この章では、前章で提案した、両眼間非対応点を奥行手がかりとして用いる両眼視の理論を実現するためのアルゴリズムを提案する。これは、V1以降の初期視覚領野のモデルに対応する。まず、4.1節で、直接手がかりとオクルージョン手がかりを同時に用いるためのアルゴリズムの枠組を提案する。そして、このアルゴリズムで何を計算しなくてはならないかについて論じる。求めたい出力は入力画像の各点の視差であるが、これを求めるためには各点が対応点か非対応点かを区別する計算も必要である。次に、4.2節で、対応点および非対応点双方を奥行手がかりとして用いる両眼視アルゴリズムを提案する。そして、このアルゴリズムに適した表現方法について論じる。このアルゴリズムでは、各点の奥行は、特定の視差に選択的に反応する tuned cell のような特性を持つ細胞によって表現される。また、網膜上の各点が対応点か非対応点かを区別する細胞も必要である。そして、この二種類の細胞が相互作用しながら繰り返し計算を行なうことによって、推定が行なわれる。4.3節では、このアルゴリズムの計算機シミュレーションの結果を示す。そして、このアルゴリズムが da Vinci ステレオグラムの知覚も説明できることを示す。

## 4.1 オクルージョンの計算理論の枠組

初期視覚モジュールの目的は、入力像の各点  $x \in I$  の奥行パラメータ（両眼視なら視差） $d(x)$  を推定することである。ここで、 $I$  は網膜像上の全ての点の集合を表す。従来のほとんどの両眼視の理論は、両眼間対応点の与える両眼視差のみを奥行手がかりとして扱ってきた。これらのモデルは、網膜像上の全ての点  $x \in I$  を対応点とみなし、データフィッティング項を適用するということを示している（例えば式(1.7)を参照）。

これに対し、本論文で提案するモデルは、網膜からの入力データを対応点と非対応点の二種類のデータに区別して扱う。いま、 $P \subset I$  は対応点、 $U \subset I$  は非対応点の集合を表し、 $P \cup U = I$ ,  $P \cap U = \emptyset$  とする。そして、データが対応点であ

る点  $x \in P$  では従来の理論で用いられているデータフィッティングを、非対応点である点  $x \in U$  では第3章で詳解したオクルージョンの拘束条件を適用する。

しかし、一般に各点  $x \in I$  が対応点か非対応点かという区別は事前に与えられないため、この区別も行なわなくてはならない。従って、各点  $x \in I$  で、奥行パラメータ  $d(x)$  だけではなく、直接手がかりとオクルージョン手がかりを区別する隠れパラメータ  $\phi(x)$  も推定しなくてはならない。ここで、 $x$  が対応点のとき  $\phi(x) = 0$ 、非対応点のとき  $\phi(x) = 1$  とする。そして、データが対応点である点  $\phi(x) = 0$  でデータフィッティングを、非対応点である点  $\phi(x) = 1$  でオクルージョンの拘束条件を適用する。このアルゴリズムでは、二つのパラメータ  $(d(x), \phi(x))$  を推定する必要がある。つまり、(1) 外界の構造復元、及び (2) 直接手がかりを与える領域とオクルージョン手がかりを与える領域との区別、という二つの問題を解かなければならない。

オクルージョン手がかりを用いた視覚情報処理は、初期視覚の枠組の中で行なわれていると考えられる。なぜなら、前述のように両眼視における非対応点は低次の領野（おそらく V1）で検出されているということが、生理学的な知見から示唆されるからである。さらに、両眼視以外のモジュールにおいても、網膜像からのデータ入力という初期の過程で直接手がかりを与える領域とオクルージョン手がかりを与える領域との区別が必要なため、低次の領野でオクルージョン手がかりを与える領域の検出が行なわれていると考えられる。

しかし一般に、オクルージョン手がかりを与える領域を、構造復元の過程を経ずに網膜像から直に検出することはできない。例えば、ランダムドットステレオグラムから、対応問題を解くことなしに非対応点を指示示すことができるか見て見ればよい。網膜像の各点を対応点と非対応点に区別するためには、網膜像の各点がもう一方の網膜像のどの点と対応しているか、あるいはどの点とも対応していないかを調べなくてはならない。そして、そのためには対応問題を解かなければならぬ。しかし、オクルージョン手がかりは外界の構造復元に用いたいので、なんらかの方法で構造復元を行なった後にオクルージョン手がかりを与える領域を検出していては意味がない。従って、外界の構造復元と領域の区別という二つの問題を同時に、相互作用しながら繰返し計算によって解くことを考える。

図 4.1 にモデルの概要を示す。一番下の層は網膜に対応し、入力画像が与えられる。入力像に適切な前処理を行なったデータが中央の層に送られる。中央の層は低次の視覚野（両眼視モジュールでは V1）に対応する。この層では、入力像の各点を、直接手がかりを与える領域とオクルージョン手がかりを与える領域とに区別する。適切な前処理を行なうことによって、明らかに区別可能な点（例えば両眼視において対応候補が全く存在しないような点）については、繰返し計算の前の時点での区別できる。そして、直接手がかりを与える領域とオクルージョン手がかりを

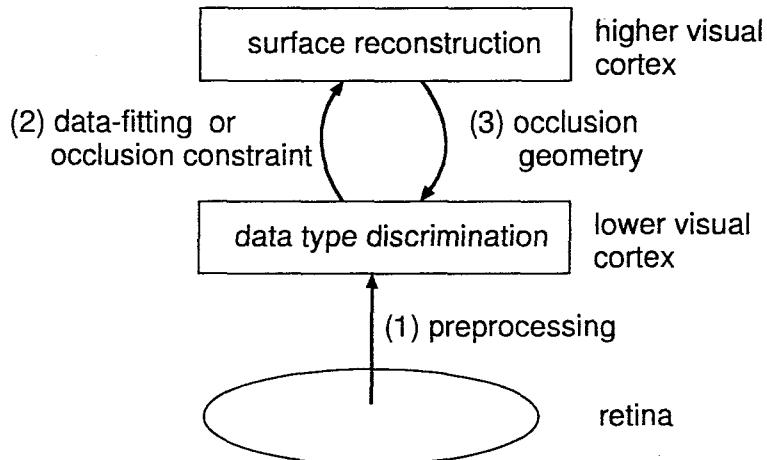


図 4.1: オクルージョン手がかりを用いた初期視覚モデルの概要.

与える領域は、一番上の層に異なるデータを与える。例えば両眼視では、直接手がかりを与える領域は両眼間対応領域に相当し、この領域から高次の層へは視差のデータが送られる。オクルージョン手がかりを与える領域は両眼間非対応領域に相当し、視差のデータは得られないが、オクルージョンの幾何学的な条件から、再構成される面のとり得る形状を制約する。一番上の層は高次の視覚野に対応する。この層では、低次の層から与えられるデータ、および内部の拘束条件を用いて、外界の構造復元が行なわれる。さらに、高次の構造復元を行なう層から低次の層への、逆向きの結合が存在する。この逆向きの計算で、高次の層で再構成されている構造をもとに、直接手がかりを与える領域とオクルージョン手がかりを与える領域との区別が行なわれる。この計算は、光学に基づいたオクルージョンの幾何学の計算をすればよい。この繰返し計算によって、高次の層と低次の層の計算結果が、最終的に、オクルージョンの幾何学にてらして矛盾のない形になる。

この枠組を用いて具体的なモデルを構築する際に考えねばならない処理は、大きく分けて次の三つになる：

1. 入力画像の前処理(図 4.1(1)).
2. オクルージョンの拘束条件を用いて推定される面の形状を制約する処理(図 4.1(2)).
3. 推定された面の形状からオクルージョン手がかりを与える領域を検出する処理(図 4.1(3)).

次節でオクルージョン手がかりを用いた両眼視モデルの具体的な定式化を行なう。

最後に、この理論で鍵となる考え方をまとめておく。初期視覚における外界の構

造復元には、直接手がかりだけではなく、オクルージョン手がかりも用いられている。すると、解くべき問題は構造復元と、オクルージョンによって生じる領域の検出という二つになる。この二つの問題は互いに密接な関係がある。この理論では、ボトムアップの流れで構造復元を、トップダウンの流れでオクルージョンによって生じる領域の検出を行う。これは、面復元モデルにおけるフィリングインプロセスとラインプロセスの関係に類似している。また、この枠組みは川人、乾の理論 [43] と同様の形式になっている。

## 4.2 両眼視アルゴリズム

この節では、4.1 節で提案したオクルージョンの計算理論の枠組に基づいて、両眼間非対応点を奥行手がかりとして用いる両眼視アルゴリズムを提案する。

第1章で述べたように、両眼視アルゴリズムで用いられている視差表現には、大きく分けて二種類のものがある。

一つは、Marr と Poggio の協調的アルゴリズム [55] のように、特定の位置に特定の視差を持つ物体が存在するときのみ発火する、tuned cell に似た特性を持つ細胞で表現する方法である。この表現方法を、high dimensional representation と呼ぶ。この表現形式を用いる利点は、一意性の拘束条件から導かれるような、視線方向に沿った細胞間の結合を構成しやすいという点が挙げられる。さらに、多重表面の表現が容易である点や、生理学的な知見との対応が取りやすい点などがある。

この表現を用いたアルゴリズムの計算は、Hopfield 型の神経回路モデル [29] によって最適化問題を解いているとみなすことができる ([7, 66] など参照)。例えば、(1.3) 式で表される Marr と Poggio の協調的アルゴリズムは、次式で表されるエネルギー

$$\begin{aligned}
 E &= -\frac{1}{2} \sum_{x,y,d} \sum_{\substack{x',y',d' \\ \in S(x,y,d)}} C_{x,y,d}^t C_{x',y',d'}^t + \frac{\epsilon}{2} \sum_{x,y,d} \sum_{\substack{x',y',d' \\ \in O(x,y,d)}} C_{x,y,d}^t C_{x',y',d'}^t \\
 &\quad - \sum_{x,y,d} C_{x,y,d}^0 C_{x,y,d}^t \\
 &= -\frac{1}{2} \sum_{x,y,d} \sum_{x',y',d'} \left( W_{x,y,d}^S - \epsilon W_{x',y',d'}^O \right) C_{x,y,d}^t C_{x',y',d'}^t \\
 &\quad - \sum_{x,y,d} C_{x,y,d}^0 C_{x,y,d}^t
 \end{aligned} \tag{4.1}$$

を最小化する Hopfield モデルとみなすことができる。ここで、

$$W_{x,y,d}^S = \begin{cases} 1 & x', y', d' \in S(x, y, d) \\ 0 & \text{otherwise} \end{cases} \quad (4.2)$$

$$W_{x,y,d}^O = \begin{cases} 1 & x', y', d' \in O(x, y, d) \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

である。

もう一つの表現方法は、ある点の視差を細胞の発火頻度で表現するというものである。この表現方法を、low dimensional representation と呼ぶ。この表現方法を用いる利点は、素子数が少なくてすむ点や、視差を連続値で表現できる点などが挙げられる。この表現を用いたアルゴリズムには、1.1.3 節で示した、標準正則化理論のエネルギー関数から直接導かれるアルゴリズム（式（1.10））がある。

本論文で提案するアルゴリズムの表現には、前者の high dimensional representation を用いる。その理由の一つは、オクルージョンの拘束条件を容易に実現できるためである。オクルージョンの拘束条件は、「非対応点と、この非対応点が見えない眼とを結ぶ視線上に、他の点が少なくとも一つ存在する」という視線に沿った拘束条件であるため、high dimensional representation の方が実現しやすい。もう一つの理由は、全ての非対応点の視差を容易に表現できるためである。一般に、左眼座標や右眼座標、サイクロピアン座標など、どのような座標系で視差を表現しても、非対応領域の一部は遮蔽面と重なってしまう。従って、非対応点の視差を全て表現するためには、多重表面の表現が行なえなくてはならない。多重表面の表現には、high dimensional representation の方が適している<sup>1</sup>。

本論文で提案する両眼視アルゴリズムの構造は、図 4.2 のようになっている。一番下の層は両眼網膜に対応し、入力画像が与えられる。中央の層は V1 に対応する。この層には 2 種類の細胞が存在する。一つは tuned excitatory cell に似た両眼性細胞  $\beta$  で、偽対応も含む全ての対応候補について、その両眼パターンの類似度を出力する。もう一つは、左右眼網膜上の各点が対応点か非対応点かを表す細胞  $\phi$  である。非対応点かどうかを表す細胞は片眼優位の特性を持つと考えられ、右眼非対応点を表す細胞と、左眼非対応点を表す細胞が存在する。前述のように、このような片眼優位性のある細胞は V2 以上の領野ではほとんど見られないため、細胞  $\phi$  は V1 に存在すると考えられる。一番下の層からこの層への入力は、入力画像の

---

<sup>1</sup> 一般に low dimensional representation では、各点の視差を单一の細胞の発火頻度で表現するため、複数の視差を表現することができない。志沢 [84-86] は、各点に複数の細胞を用意することによって、low dimensional representation で多重表面を表現する計算理論を提案している。しかし、この理論を用いたアルゴリズムは非常に複雑になるため、本論文で提案するアルゴリズムでは用いない。

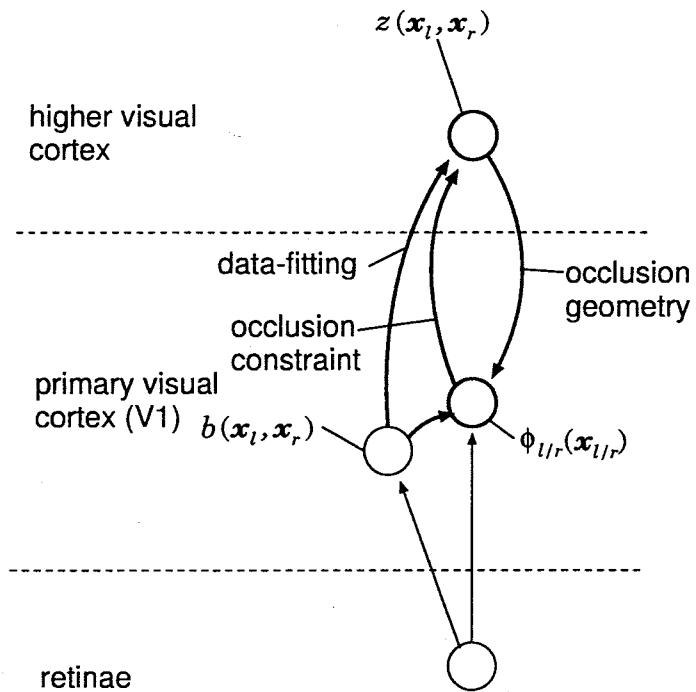


図 4.2: 両眼視アルゴリズムの模式図.

前処理に相当する。この計算は 4.2.1 節で述べる。一番上の層は高次の視覚野に対応する。この層にも tuned excitatory cell に似た両眼性細胞  $z$  が存在する。細胞  $z$  は、低次の層の細胞  $b$  および  $\phi$  から入力を受ける。この層の細胞は、推定される外界の構造を表現する。この計算は 4.2.2 節で述べる。また、 $z$  から  $\phi$  への逆向きの結合によって、 $\phi$  による非対応点の推定が行なわれる。この計算は 4.2.3 節で述べる。この繰返し計算による相互作用によって、 $z$  及び  $\phi$  が更新される。

#### 4.2.1 入力画像の前処理

ここでは入力画像の前処理について述べる。前処理では、偽対応を含めた全ての対応候補を検出する。そして、次のステップで正しい対応を推定する手がかりとして用いるため、対応する左右眼特徴の類似度に相当する値を計算する。また、明らかに対応候補が一つも存在しない点は非対応点として検出する。この処理は、脳内では網膜から V1 の両眼性細胞までの経路で行なわれると考えられる。

2 次元の左右眼網膜像の座標をそれぞれ  $x_l = (x_l, y_l)$ ,  $x_r = (x_r, y_r)$  , この座標の輝度をそれぞれ  $I_l(x_l)$ ,  $I_r(x_r)$  で表す。

全ての対応候補を検出し、その類似度に相当する値を計算するという処理は、第2章で述べた tuned excitatory cell によって行なわれると考えられる。tuned excitatory cell は、最適特徴(オリエンテーション)を持った両眼パターンが最適視差で提示されるとき最も大きな値を出力し、特徴や視差が最適のものと異なると出力は急激に減少するという特性を持つため、この処理には非常に適している。しかし、計算機シミュレーションを行なう際には、様々な最適特徴を持つ tuned excitatory cell を並列・独立に用意しなければならない。また、第2章で述べたモデル自体も複雑なので、簡単のために以下に述べるような単純化したモデルを用いることにする。

いま、 $x_l, x_r$  に受容野中心を持つ両眼性細胞を考える。ここで、簡単のため  $y_l = y_r$  とし、垂直視差は考慮しない。この細胞の示す視差は  $d = x_l - x_r$ 、サイクロプス座標での位置は  $x_c = (x_l + x_r)/2$  で与えられる。この細胞の出力  $b(x_l, x_r)$  は、左右眼網膜上の受容野内のパターンの類似度に相当する値を出力する。類似度として、次式で表される二つのパターンの重みつきの方向余弦を考える。

$$C(x_l, x_r) = \frac{\sum_{\xi} g(\xi) I_l(x_l + \xi) I_r(x_r + \xi)}{\sqrt{\sum_{\xi} g(\xi) \{I_l(x_l + \xi)\}^2} \sqrt{\sum_{\xi} g(\xi) \{I_r(x_r + \xi)\}^2}} \quad (4.4)$$

ここで、重みづけ関数  $g(\xi)$  を Gauss 関数

$$g(\xi) = \exp\left(-\frac{\|\xi\|^2}{2\sigma_g^2}\right) \quad (4.5)$$

で定義する。式(4.4)を用いて、この細胞の出力を

$$b(x_l, x_r) = s_l(x_l) \cdot s_r(x_r) \cdot \varphi\left[\frac{C(x_l, x_r) - \theta}{1 - \theta}\right] \quad (4.6)$$

と定義する。ここで、入出力関数  $\varphi$  は

$$\varphi[u] = \begin{cases} u & (u > 0) \\ 0 & (u \leq 0) \end{cases} \quad (4.7)$$

と表される。 $\theta$  はこの細胞の閾値で、 $|\theta| < 1$  である。この細胞は  $C(x_l, x_r) > \theta$  のとき出力する。また、この前処理の目的から、tuned excitatory cell のように単眼入力のときは発火しないという性質が必要である。この性質は、(4.6)式中の変数  $s_l(x_l), s_r(x_r)$  で実現されている。 $s_l(x_l), s_r(x_r)$  は、それぞれ左右眼に対応特徴となる入力が与えられているとき 1、与えられないとき 0 をとる変数である。図4.3 は、 $s_l(x_l)$  の取る値を説明する図である。いま、図のように二値パターン  $I_l(x)$  が与えられているとする。図の中央より左側には黒の点しかなく、対応特徴となる

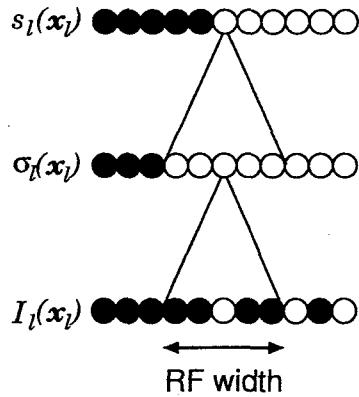


図 4.3:  $s_l(x_l)$  のとる値。簡単のため 1 次元で表す。入力  $I_l(x_l)$  は二値とする。 $\sigma_l(x_l)$ ,  $s_l(x_l)$  は 0 より大きな値をとる点を白丸、0 を黒丸で表す。

パターンが存在しない。特徴の有無を、受容野内のパターンの分散で測ることにする。分散が正のとき、特徴が与えられているとみなす。いま、中心が  $x_l$ 、半径が 2 画素の左眼受容野内のパターンの分散を  $\sigma_l(x_l)$  とする。すると、図のように、特徴が無くなる境界から受容野半径分だけ左側まで分散が正になる。よって、さらに左眼受容野内の各点  $x_l + \xi$  ( $|\xi| \leq 2$ ) を中心とする分散  $\sigma_l(x_l + \xi)$  が全て正のとき  $s_l(x_l) = 1$ 、それ以外では  $s_l(x_l) = 0$  とする。すると、図のように特徴が無くなる境界から  $s_l(x_l) = 0$  となる。右眼特徴の有無  $s_r(x_r)$  も同様に計算する。両眼の受容野のパターンが共に対応特徴を持たないとき、つまり单一の輝度しか持たないときでも  $C(x_l, x_r) = 1$  となる。しかし、このとき  $s_l(x_l) = s_r(x_r) = 0$  となるため、この場合の細胞の出力を 0 にできる。

これで、両眼パターンの特徴によらずに、類似度を計算することができる。(4.6) 式の両眼性細胞  $b(x_l, x_r)$  は、あらゆる最適特徴を持った、最適視差が同一の tuned excitatory cell の集合と見なすことができる。この細胞は、類似度として重みつき方向余弦を用い、変数  $s_l(x_l)$ ,  $s_r(x_r)$  が単眼入力のとき反応を抑制するというアナログ AND モデルのゲート細胞の役割を果たすため、2.4 節で示したモデルを単純化したものとみなすことができる。

次に、左右眼網膜像の点  $x_l$ ,  $x_r$  が非対応点かどうかを表す細胞を  $\phi_l(x_l, t)$ ,  $\phi_r(x_r, t)$  で表す。ここで  $t$  は時刻を表す。それぞれ、非対応点とみなされるときは 1、対応点とみなされるときは 0 を出力する。この細胞を、非対応点検出細胞と呼ぶ。

いま、右眼網膜上の点  $x_r$  が非対応点かどうかを表す、右眼非対応点検出細胞  $\phi_r(x_r, t)$  について考える。右眼網膜上の点  $x_r$  に対応特徴が与えられている(即ち

$s_r(\mathbf{x}_r) = 1$  である)が、 $\mathbf{x}_r$  を受容野中心に持つ両眼性細胞が一つも発火していない場合(即ち  $\forall \mathbf{x}_l$  で  $b(\mathbf{x}_l, \mathbf{x}_r) \approx 0$  の場合)、この点は右眼の非対応点であるとみなすことができる。このとき、右眼非対応点検出細胞  $\phi_l(\mathbf{x}_l, t)$  の  $t = 0$  での初期値が 1 になるように定義する。 $\phi_r(\mathbf{x}_r, 0)$  は、次式のように定義される。

$$\phi_r(\mathbf{x}_r, 0) = s_r(\mathbf{x}_r) \cdot f \left[ \mu - \sum_{\mathbf{x}_l} b(\mathbf{x}_l, \mathbf{x}_r) \right] \quad (4.8)$$

ここで、入出力関数  $f$  として、次のようなステップ関数

$$f[u] = \begin{cases} 1 & u > 0 \\ 0 & \text{otherwise} \end{cases} \quad (4.9)$$

を用いる。 $\mu$  は正の定数である。左眼非対応点検出細胞  $\phi_l(\mathbf{x}_l, t)$  も同様に定義される。

この細胞は、右眼に特徴が与えられたとき発火するが、左眼にこの特徴と対応する特徴が存在すれば出力が抑制される。従ってこの細胞は、様々な最適特徴や最適視差を持つ、右眼優位の tuned inhibitory cell の集合とみなすことができる(このことについては第5章で再び議論する)。 $\phi_l(\mathbf{x}_l, t), \phi_r(\mathbf{x}_r, t)$  は、4.2.3 節で述べる繰返し計算によって最終的な値が推定される。

#### 4.2.2 面の構造復元

ここでは高次の、外界の構造復元を行なう層について述べる。この層では、特定の視差に選択性を持つ tuned cell のような細胞が相互に結合している回路を考える。この細胞を、視差検出細胞と呼ぶことにする。

いま、 $\mathbf{x}_l, \mathbf{x}_r$  に受容野中心を持つ細胞を考える。ここで、垂直視差は考慮しないので  $y_l = y_r$  とする。この細胞の膜電位を  $u(\mathbf{x}_l, \mathbf{x}_r, t)$  とし、出力  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  を

$$z(\mathbf{x}_l, \mathbf{x}_r, t) = (s_l(\mathbf{x}_l) + s_r(\mathbf{x}_r) - s_l(\mathbf{x}_l) \cdot s_r(\mathbf{x}_r)) \cdot f[u(\mathbf{x}_l, \mathbf{x}_r, t)] \quad (4.10)$$

と定義する。この細胞の膜電位の初期値は  $b(\mathbf{x}_l, \mathbf{x}_r)$  とする。また、両眼対応に用いる特徴が存在しない領域では面の再構成をしないようにするために、 $s_l(\mathbf{x}_l) = 0$  かつ  $s_r(\mathbf{x}_r) = 0$  のときは発火しないようにしている。非対応点の視差も表現したいため、片眼にしか特徴が与えられないときは発火するようになっている。 $s_l(\mathbf{x}_l), s_r(\mathbf{x}_r)$  を用いるということは、疎らなドットパターン間を補間するかどうかを、ドット間の距離が受容野の直径以下かどうかで定めることと同値である。

膜電位  $u(\mathbf{x}_l, \mathbf{x}_r, t)$  の時間変化は、次式のように定義される。

$$\frac{\partial}{\partial t} u(\mathbf{x}_l, \mathbf{x}_r, t) = -u(\mathbf{x}_l, \mathbf{x}_r, t) + U_d + \lambda_o U_o + \lambda_u U_u + \lambda_c U_c \quad (4.11)$$

$\lambda_o, \lambda_u, \lambda_c$  は正の定数である。図4.4は、この細胞と、近傍の視差検出細胞や非対応点検出細胞との位置関係を、Keplerianダイアグラムで模式的に示したものである。

(4.11)式の第2項は、この細胞の受容野の位置にある特徴が対応点のときに与えられるデータを表す。これは、対応点におけるデータフィッティング項に対応する。 $\mathbf{x}_l, \mathbf{x}_r$  が共に対応点のときは、低次の両眼性細胞  $b(\mathbf{x}_l, \mathbf{x}_r)$  から、対応する両眼パターンの類似度が入力される。これは、次式で表される。

$$U_d = (1 - \phi_l(\mathbf{x}_l, t)) \cdot (1 - \phi_r(\mathbf{x}_r, t)) \cdot b(\mathbf{x}_l, \mathbf{x}_r) \quad (4.12)$$

$\mathbf{x}_l, \mathbf{x}_r$  のどちらかが非対応点のときは、 $b(\mathbf{x}_l, \mathbf{x}_r)$  からの入力がブロックされ、代わりにオクルージョンの幾何学にそった面復元を行なうための入力が与えられる。

(4.11)式の第3項が、非対応点から与えられるデータを表す。この項は、オクルージョンの拘束条件を実現している。

この細胞の左眼受容野の位置にある特徴が左眼非対応点と推定される場合、つまり左眼非対応点検出細胞が  $\phi_l(\mathbf{x}_l, t) = 1$  のときは、単に  $\phi_l(\mathbf{x}_l, t)$  から抑制性入力が与えられる。しかし、細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  の表す3次元上の点を、右眼から遮蔽し得る点を表す細胞  $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l > x_l$ ) が発火しているときは、この抑制性入力はブロックされる。ここで  $\xi_l = (\xi_l, y_l)$  とする。これによって、非対応点についての奥行拘束領域外の点が発火しにくくなり、オクルージョンの幾何学に矛盾しない遮蔽関係を持った面が再構成され易くなる。

次に、この視差検出細胞の表す3次元上を通る右眼の視線上に存在し、かつ、この点より遠くの点を表す細胞  $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l < x_l$ ) を考える。視差検出細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  が表す3次元上の点は、 $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l < x_l$ ) が表す点を右眼から遮蔽する点となる。 $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l < x_l$ ) が発火し、かつ左眼非対応点とみなされている(つまり  $\phi_l(\xi_l, t) = 1$  である)が、右眼の遮蔽物となり得る点を表す細胞が全く発火していない場合、 $z(\mathbf{x}_l, \mathbf{x}_r, t)$  に興奮性入力が与えられる。これによって、遮蔽面についての奥行拘束領域内の点が発火しやすくなる。

右眼非対応点についても同様のことがいえる。この項は、次式で表される。

$$\begin{aligned} U_o = & -\phi_l(\mathbf{x}_l, t) \cdot (1 - N_r(\mathbf{x}_l, \mathbf{x}_r)) + \sum_{\xi_l < x_l} \phi_l(\xi_l, t) \cdot z(\xi_l, \mathbf{x}_r, t) \cdot (1 - N_r(\xi_l, \mathbf{x}_r)) \\ & - \phi_r(\mathbf{x}_r, t) \cdot (1 - N_l(\mathbf{x}_l, \mathbf{x}_r)) + \sum_{\xi_r > x_r} \phi_r(\xi_r, t) \cdot z(\mathbf{x}_l, \xi_r, t) \cdot (1 - N_l(\mathbf{x}_l, \xi_r)) \end{aligned} \quad (4.13)$$

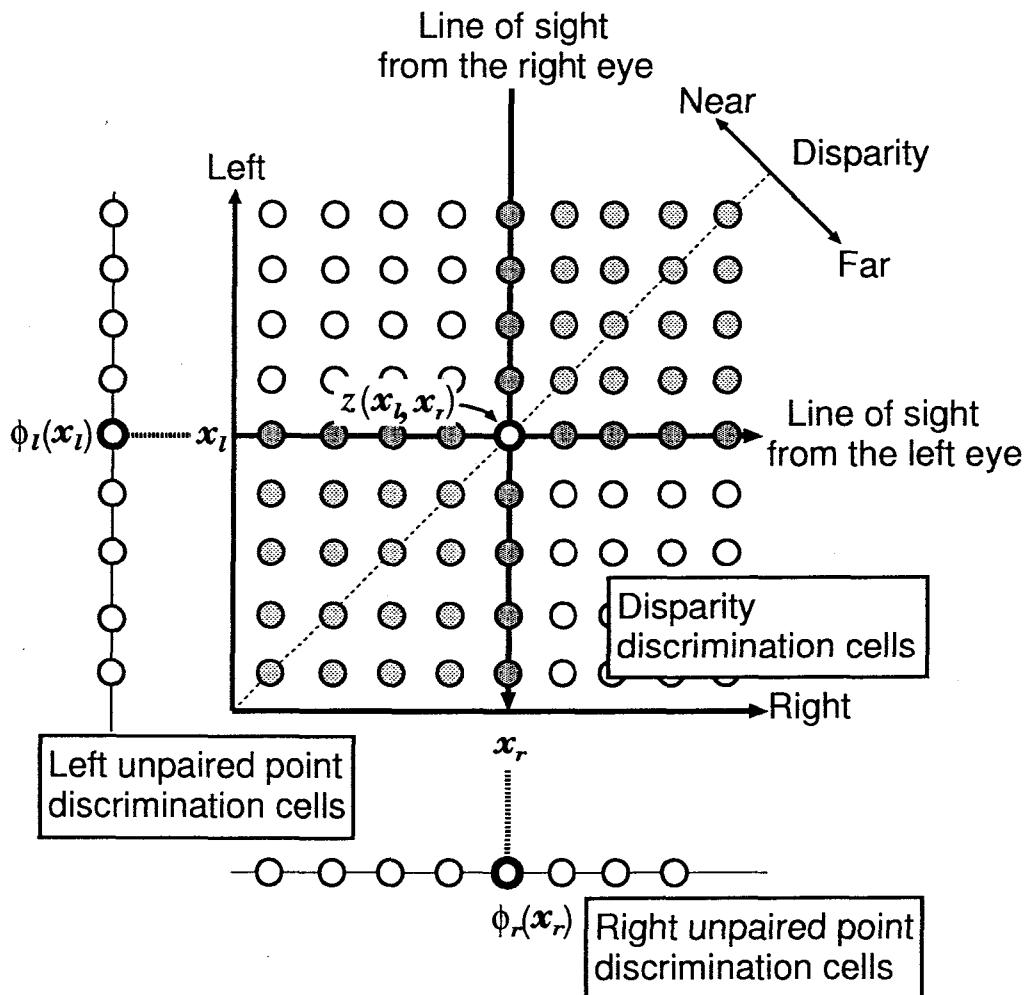


図 4.4: 視差検出細胞および非対応点検出細胞の位置関係の模式図。簡単のため、1 次元の Keplerian ダイアグラムで表示している。縦軸が左眼座標  $x_l$ 、横軸が右眼座標  $x_r$  である。視差検出細胞を通る垂直線が右眼からの視線、並行線が左眼からの視線に対応する。中心に描かれている視差検出細胞  $z(x_l, x_r)$  が非対応点かどうかは、非対応点検出細胞  $\phi_l(x_l), \phi_r(x_r)$  によって表される。一意性の拘束条件より、 $z(x_l, x_r)$  と暗灰色の細胞との間に抑制性の結合がある。この抑制性入力は、非対応点検出細胞によりブロックされることがある。また、凝集性の拘束条件より、視差勾配が 2 以下である暗灰色および明灰色の細胞と興奮性の結合がある。

ここで、 $N_l(\mathbf{x}_l, \mathbf{x}_r)$ ,  $N_r(\mathbf{x}_l, \mathbf{x}_r)$  はそれぞれ、点  $(\mathbf{x}_l, \mathbf{x}_r)$  と左眼及び右眼との間に他の点が存在するとき 1, それ以外のとき 0 を取り,

$$N_l(\mathbf{x}_l, \mathbf{x}_r) = f \left[ \sum_{\xi_l < x_r} z(\mathbf{x}_l, \xi_r, t) \right] \quad (4.14)$$

$$N_r(\mathbf{x}_l, \mathbf{x}_r) = f \left[ \sum_{\xi_l > x_l} z(\xi_l, \mathbf{x}_r, t) \right] \quad (4.15)$$

と表される。 (4.13) 式の第1項と第2項が左眼非対応点が存在するときに与えられる入力, 第3項と第4項が右眼非対応点が存在するときに与えられる入力である。

図 4.5 は、このオクルージョンの拘束条件の働きを示したものである。いま、図 4.5 の右眼非対応点検出細胞  $\phi_r(\mathbf{x}_r, t)$  が発火し、右眼画像中の点  $\mathbf{x}_r$  が右眼非対応点であると推定されているとする。すると、 $\mathbf{x}_r$  を通る右眼からの視線(垂直線)上の視差検出細胞は、この右眼非対応点の視差を表していることになる。図のように遮蔽面が存在するとき、暗灰色で示された点の細胞は左眼から見えないので、オクルージョンの拘束条件を満たしている。そして、これら以外の垂直線上の点は発火しないように、抑制性入力が与えられる。逆に、いま、中心に描かれている視差検出細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  が発火しているとする。オクルージョンの拘束条件からは、この点を通る左眼からの視線(並行線)上の、この点より近くの視差を表す細胞(明灰色の細胞)が発火していることが要求される。従って、明灰色の細胞が一つも発火していない場合は、これらの細胞に興奮性入力が与えられる。

(4.11) 式の第4項が、非対応点を考慮した一意性の拘束条件を実現している。一意性の拘束条件は、同一の視線上にある細胞からの抑制性入力で実現できる。ただし、細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  が左眼非対応点とみなされている(即ち  $\phi_x(\mathbf{x}_l, t) = 1$  である)とき、同一の右眼視線上にあり、この細胞の表す点より近い点を表す細胞  $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l > x_l$ ) からは抑制性入力は受けない。また、同一の右眼視線上にあり、この細胞の表す点より遠い点を表す細胞  $z(\xi_l, \mathbf{x}_r, t)$  ( $\xi_l < x_l$ ) が左眼非対応点とみなされている(即ち  $\phi_x(\xi_l, t) = 1$  である)とき、この細胞からは抑制性入力は受けない。右眼非対応点の場合も同様である。この入力は次式のように表される。

$$\begin{aligned} U_u = & - \sum_{\xi_l > x_l} (1 - \phi_l(\mathbf{x}_l, t) + \phi_r(\mathbf{x}_r, t)) \cdot z(\xi_l, \mathbf{x}_r, t) \\ & - \sum_{\xi_l < x_l} (1 - \phi_l(\xi_l, t) + \phi_r(\mathbf{x}_r, t)) \cdot z(\xi_l, \mathbf{x}_r, t) \\ & - \sum_{\xi_r < x_r} (1 - \phi_r(\mathbf{x}_r, t) + \phi_l(\mathbf{x}_l, t)) \cdot z(\mathbf{x}_l, \xi_r, t) \\ & - \sum_{\xi_r > x_r} (1 - \phi_r(\xi_r, t) + \phi_l(\mathbf{x}_l, t)) \cdot z(\mathbf{x}_l, \xi_r, t) \end{aligned} \quad (4.16)$$

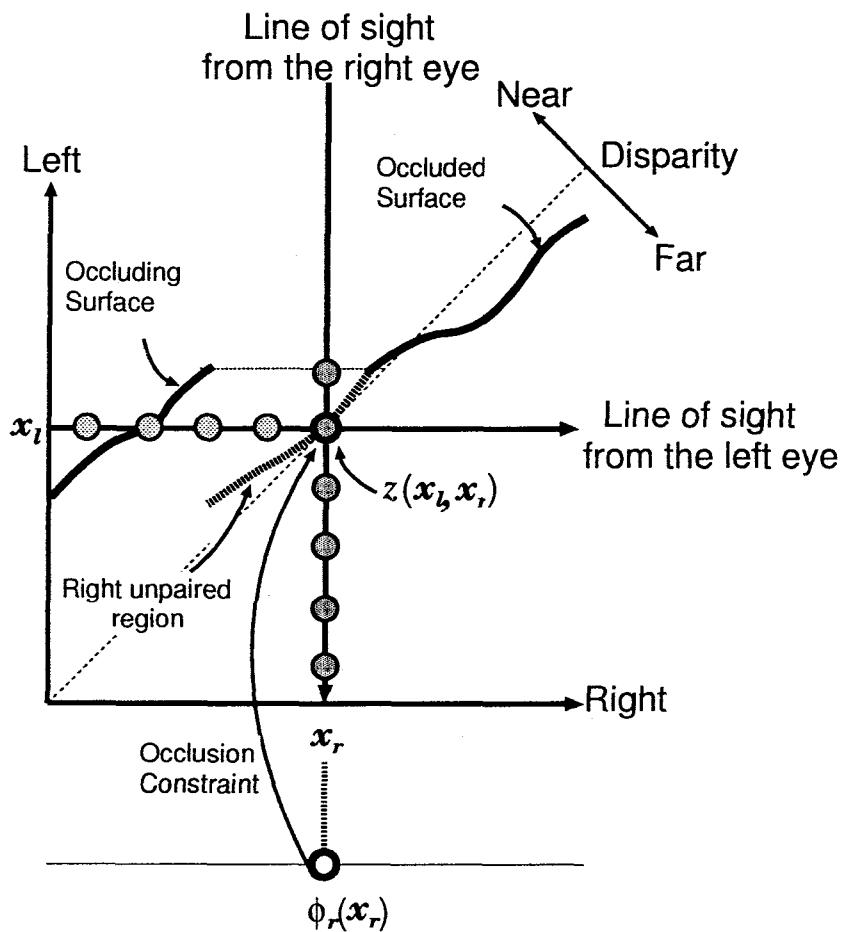


図 4.5: オクルージョンの拘束条件の効果を示す模式図。簡単のため、1次元の Keplerian ダイアグラムで表示している。図中の太線のように遮蔽面及び非遮蔽面が存在し、右眼非対応点検出細胞  $\phi_r(x_r)$  によって右眼座標上の点  $x_r$  が非対応点であると推定されているとする。オクルージョンの拘束条件は、 $x_r$  の視差として、遮蔽物の存在する暗灰色の細胞以外の細胞が発火しないようにする。逆に、 $x_r$  の視差を表す細胞  $z(x_l, x_r)$  が発火しているが、この点の遮蔽物になる明灰色の細胞が一つも発火していない時は、これらの細胞が発火するように働く。

第1項、第2項は、同一の左眼の視線上にある細胞からの抑制性入力を、第3項、第4項は、同一の右眼の視線上にある細胞からの抑制性入力を表す。図4.4は、この項によって  $z(x_l, x_r, t)$  と結合する細胞を模式的に示したものである。左眼非対応点と推定され、かつ右眼非対応点とも推定されている点は、矛盾した点であるので、この場合は非対応点とはみなさずに抑制性入力を受けるようになっている。

(4.11) 式の第5項が、凝集性の拘束条件を実現している。凝集性の拘束条件は、近傍の近い視差を表す細胞からの興奮性入力で実現できる。この入力は次式で表される。

$$U_c = \sum_{\xi_l, \xi_r \neq x_l, x_r} w(\xi_l - x_l, \xi_r - x_r) \cdot z(\xi_l, \xi_r, t) \quad (4.17)$$

この式では  $\xi_{l/r} = (\xi_{l/r}, \eta_{l/r})$  とし、 $y$  軸方向の結合も考える。興奮性結合の重み  $w$  は次式で定義される。

$$w(\xi_l - x_l, \xi_r - x_r) = \begin{cases} \exp\left(-\frac{\Delta d^2}{2\sigma_{\Delta d}^2}\right) \cdot \exp\left(-\frac{|\Delta x_c|^2}{2\sigma_{\Delta x}^2}\right) & \frac{|\Delta d|}{|\Delta x_c|} \leq 2 \\ 0 & \text{otherwise} \end{cases} \quad (4.18)$$

$\Delta x_c$  は近傍の細胞とのサイクロピアン座標での位置の差で、 $\Delta x_c = \{(\xi_l + \xi_r) - (x_l + x_r)\}/2$  と表される。また、 $\Delta d$  は近傍の細胞との視差の差で、 $\Delta d = (\xi_l - \xi_r) - (x_l - x_r)$  と表される。興奮性結合は、視差勾配  $|\Delta d|/|\Delta x_c|$  が 2 以下の細胞間にのみ存在する。つまり、Panum の融合域 (Panum's fusional area) に存在する点の間で結合していることになる。図4.4は、この項によって  $z(x_l, x_r, t)$  と結合する細胞を模式的に示したものである。式(4.18)のGauss関数は、近傍の細胞との視差勾配  $|\Delta d|/|\Delta x_c|$  に対して単調減少の関数になっている。また、距離が離れた細胞との結合は小さくなっている。

#### 4.2.3 オクルージョン手がかりを与える領域の検出

この節では、低次領野の、直接手がかりを与える領域とオクルージョン手がかりを与える領域とを区別するための処理について述べる。この区別は、左眼非対応点検出細胞  $\phi_l(x_l, t)$  と右眼非対応点検出細胞  $\phi_r(x_r, t)$  によって行なわれる。ここでは、右眼網膜像の点  $x_r$  について非対応点なら発火し、対応点なら発火しない細胞  $\phi_r(x_r, t)$  について考える。左眼非対応点検出細胞  $\phi_l(x_l, t)$  も同様の方法で定義できる。

右眼網膜上の点  $x_r$  が非対応点なら発火する細胞の出力  $\phi_r(x_r, t)$  を、膜電位  $v_r(x_r, t)$  を用いて

$$\phi_r(x_r, t) = s_r(x_r) \cdot f[v_r(x_r, t)] \quad (4.19)$$

と定義する。受容野内に特徴が存在しないとき、つまり  $s_r(x_r) = 0$  のときは発火しないようにしている。

この膜電位の時間変化は

$$\frac{\partial}{\partial t} v_r(x_r, t) = -v_r(x_r, t) + V_d + \lambda_o V_o + \lambda_u V_u + \lambda_h V_h \quad (4.20)$$

と表される。

上式の第2項は入力網膜像から与えられるデータを表す。式(4.8)で示したように、前処理の段階では入力網膜像から両眼間の類似度を計算し、対眼のどの点との類似度もほとんど0の点は非対応点と推定され、興奮性入力が与えられた。 $t > 0$ のときは、面の復元を行なう層で、実際に対応が存在すると推定されている点のみを考慮する。即ち、視差検出細胞  $z(x_l, x_r, t), \forall x_l$  が発火し、かつ左眼非対応点ではない(即ち  $\phi_l(x_l, t) = 0$  である)対応の類似度のみを考慮する。これは、次式のように表される。

$$V_d = \mu - \sum_{x_l} (1 - \phi_l(x_l, t)) \cdot z(x_l, x_r, t) \cdot b(x_l, x_r) \quad (4.21)$$

と表される。 $t = 0$  のときは、視差検出細胞の膜電位の初期値が  $b(x_l, x_r)$  と定義され、左眼非対応点検出細胞  $\phi_l(x_l, t)$  の情報を用いることができないことを考えると、上式は初期値を定義した式(4.8)と一致する。

式(4.20)の第3項と第4項は、面の構造復元を行なう層からの入力で、現時点では推定されている外界の構造から、オクルージョンの幾何学に基づいて、各点が非対応点かどうかを推定する。右眼非対応点と推定されるときに興奮性入力が、それ以外のときに抑制性入力が与えられる。

式(4.20)の第3項は、同一の視線上に遮蔽物となる点が存在しているときに与えられる入力を表す。この項では、今考慮している非対応点検出細胞  $\phi_r(x_r, t)$  によって非対応点かどうかを示される視差検出細胞  $z(x_l, x_r, t), \forall x_r$  が発火し、かつ、この点と左眼とを結ぶ視線上に遮蔽物となる点を表す細胞が発火しているとき、この点は右眼の非対応点とみなすことができるので興奮性入力が与えられる。逆に、この点と右眼とを結ぶ視線上に遮蔽物となる点を表す細胞が発火しているときは、この点は左眼非対応点とみなせるので、抑制性入力が与えられる。これは、次式で表される。

$$V_u = \sum_{x_l} z(x_l, x_r, t) \cdot \left[ \sum_{\xi_r < x_r} z(x_l, \xi_r, t) - \sum_{\xi_l > x_l} z(\xi_l, x_r, t) \right] \quad (4.22)$$

式(4.20)の第4項は、同一の視線上に遮蔽物となる点が存在していないときに与えられる入力を表す。今考慮している非対応点検出細胞によって非対応点かど

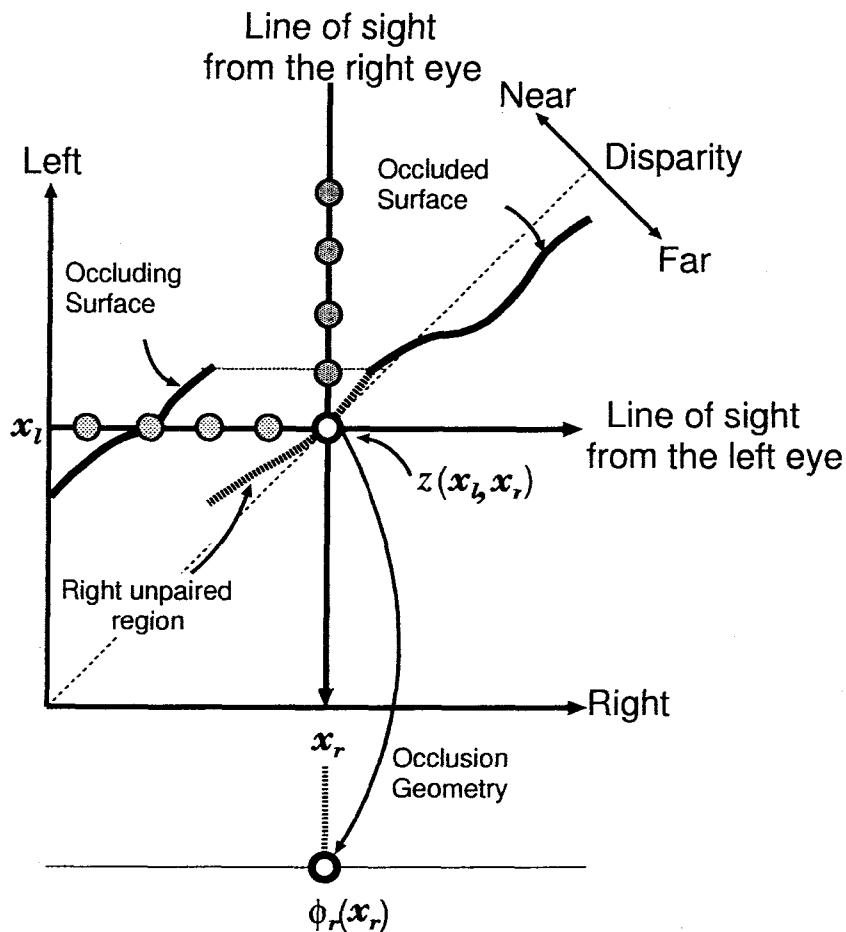


図 4.6: オクルージョンの幾何学による非対応点推定を表す模式図。簡単のため、1次元の Keplerian ダイアグラムで表示している。視差検出細胞  $z(x_l, x_r)$  が発火している場合を考える。この細胞が右眼非対応点かどうかは、右眼非対応点検出細胞  $\phi_r(x_r)$  によって表される。オクルージョンの幾何学から、 $z(x_l, x_r)$  を左眼から遮蔽する細胞(明灰色の細胞)が発火しているとき、この点は右眼非対応点と推定されるので  $\phi_r(x_r)$  に興奮性入力が与えられる。逆に、明灰色の細胞が一つも発火していないとき、および  $z(x_l, x_r)$  を右眼から遮蔽する細胞(暗灰色の細胞)が発火しているときは、右眼非対応ではないと推定されるので、抑制性入力が与えられる。

うかが示される視差検出細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t), \forall \mathbf{x}_l$  と、左眼とを結ぶ視線上に、遮蔽物となる点を表す細胞が一つも発火していないとき、この点は非対応点ではないとみなされるので、抑制性入力が与えられる。この項は次式で表される。

$$V_o = - \sum_{\mathbf{x}_l} z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot (1 - N_l(\mathbf{x}_l, \mathbf{x}_r)) \quad (4.23)$$

図4.6は、式(4.20)の第3項と第4項で表される、オクルージョンの幾何学に基づいた推定を説明する図である。いま、図4.6の視差検出細胞  $z(\mathbf{x}_l, \mathbf{x}_r, t)$  が発火しているとする。この細胞が、右眼非対応点かどうかを表す右眼非対応点検出細胞  $\phi_r(\mathbf{x}_r, t)$  への入力を考える。 $z(\mathbf{x}_l, \mathbf{x}_r, t)$  を左眼から遮蔽する細胞（明灰色の細胞）が発火しているとき、この点は右眼非対応点と推定されるので  $\phi_r(\mathbf{x}_r)$  に興奮性入力が与えられる。明灰色の細胞が発火していないときは、右眼非対応点ではないと推定されるので、抑制性入力が与えられる。さらに、この細胞を右眼から遮蔽する細胞（暗灰色の細胞）が発火しているときも、この点は右眼非対応ではないと推定されるので、抑制性入力が与えられる。

式(4.20)の第5項は、非対応点についての事前知識から導かれる拘束条件を実現している。事前知識としては、非対応点の凝集性の仮定が考えられる。これは、非対応点は凝集して存在し、孤立した非対応点は生成され難いという仮定である。この仮定は、点のような小さな遮蔽面の存在を考慮しないという仮定と等しい。非対応点の凝集性を仮定することによって、入力像に付加されたノイズの影響を除去することができる。この項は、次式のような8近傍の細胞からの興奮性入力で実現する。

$$V_h = \sum_{\xi_r \in S} \phi_r(\xi_r, t) \quad (4.24)$$

ここで、 $S$  は  $\mathbf{x}_r$  についての8近傍を表す。

#### 4.2.4 エネルギー関数による表現

4.2.2節から4.2.3節で述べた繰返し計算アルゴリズムは、次式で表されるエネルギー関数を最小化するアルゴリズムとみなすことができる。

$$E = E_d + \lambda_o E_o + \lambda_u E_u + \lambda_c E_c + \lambda_h E_h \quad (4.25)$$

ここで、 $E_d$  はデータフィッティング項を表し、次式で定義される。

$$\begin{aligned} E_d = & - \sum_{\mathbf{x}_l, \mathbf{x}_r} b(\mathbf{x}_l, \mathbf{x}_r) \cdot (1 - \phi_l(\mathbf{x}_l, t)) \cdot (1 - \phi_r(\mathbf{x}_r, t)) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \\ & - \mu \sum_{\mathbf{x}_l} \phi_l(\mathbf{x}_l, t) - \mu \sum_{\mathbf{x}_r} \phi_r(\mathbf{x}_r, t) \end{aligned} \quad (4.26)$$

$E_o$  はオクルージョンの拘束条件を表す.

$$\begin{aligned} E_o = & \sum_{\mathbf{x}_l, \mathbf{x}_r} \phi_l(\mathbf{x}_l, t) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot (1 - N_r(\mathbf{x}_l, \mathbf{x}_r)) \\ & + \sum_{\mathbf{x}_l, \mathbf{x}_r} \phi_r(\mathbf{x}_r, t) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot (1 - N_l(\mathbf{x}_l, \mathbf{x}_r)) \end{aligned} \quad (4.27)$$

$E_u$  は、非対応点の存在を考慮して拡張した、一意性の拘束条件を表す.

$$\begin{aligned} E_u = & \sum_{\mathbf{x}_l, \mathbf{x}_r} \sum_{\xi_l > \mathbf{x}_l} (1 - \phi_l(\mathbf{x}_l, t) + \phi_r(\mathbf{x}_r, t)) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot z(\xi_l, \mathbf{x}_r, t) \\ & + \sum_{\mathbf{x}_l, \mathbf{x}_r} \sum_{\xi_r < \mathbf{x}_r} (1 - \phi_r(\mathbf{x}_r, t) + \phi_l(\mathbf{x}_l, t)) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot z(\mathbf{x}_l, \xi_r, t) \end{aligned} \quad (4.28)$$

$E_c$  は凝集性の拘束条件を表す.

$$E_c = -\frac{1}{2} \sum_{\mathbf{x}_l, \mathbf{x}_r} \sum_{\xi_l, \xi_r \neq \mathbf{x}_l, \mathbf{x}_r} w(\xi_l - \mathbf{x}_l, \xi_r - \mathbf{x}_r) \cdot z(\mathbf{x}_l, \mathbf{x}_r, t) \cdot z(\xi_l, \xi_r, t) \quad (4.29)$$

$E_h$  は、ノイズの影響を除去するために前節で導入した、非対応点の凝集性の仮定を表し、次式で与えられる.

$$E_h = - \sum_{\mathbf{x}_l} \sum_{\xi_l \in S} \phi_l(\mathbf{x}_l, t) \cdot \phi_l(\xi_l, t) - \sum_{\mathbf{x}_r} \sum_{\xi_r \in S} \phi_r(\mathbf{x}_r, t) \cdot \phi_r(\xi_r, t) \quad (4.30)$$

(4.11) 式および (4.20) 式で示したアルゴリズムは、このエネルギーを減少する方向に変化させる。ただし、明らかにこのアルゴリズムは、このエネルギーを最小にする保証はなく、一般に極小値に収束してしまうと考えられる。しかし、次節で示す計算機シミュレーションの結果から、このアルゴリズムでも十分満足できる結果が得られることが示される。

### 4.3 計算機シミュレーション

この節では、前節で提案した両眼視アルゴリズムの計算機シミュレーションの結果を示す。パラメータは、それぞれ  $\lambda_o = 0.2$ ,  $\lambda_u = 2$ ,  $\lambda_c = 0.7$ ,  $\lambda_h = 0.1$ ,  $\mu = 0.1$  とした。また、両眼パターンの方向余弦を計算する時の閾値は  $\theta = 0.9$  とした。

#### 4.3.1 Wheatstone ステレオグラム

Wheatstone ステレオグラムとは、近傍の対応領域の視差を補間することによって非対応領域の視差が推定されるという、従来の両眼視の理論の考え方でもその

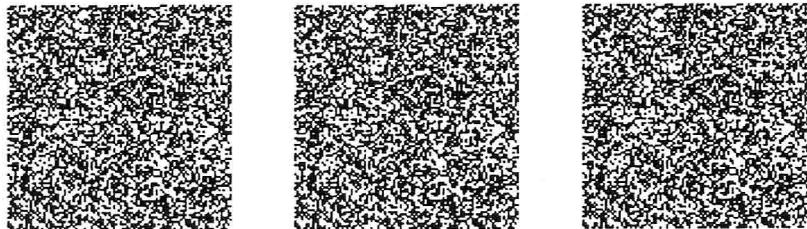


図 4.7: Wheatstone ステレオグラム。左右の図が左眼画像、中央の図が右眼画像である。並行法で融合視する場合は左図と中央図を、交差法で融合視する場合は中央図と右図を用いる。融合視すると、中央部分に四角形が浮き上がって知覚される。

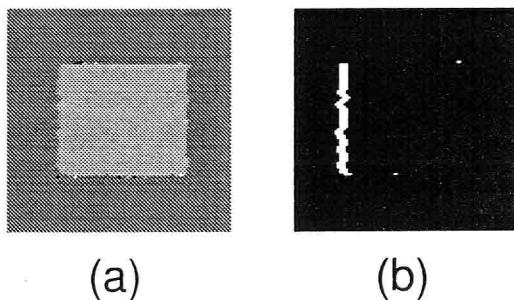


図 4.8: Wheatstone ステレオグラムのシミュレーション結果。左眼座標で表されている。(a) 視差の推定結果。視差はグレースケールで表され、明るい色が近い視差を表す。(b) 左眼非対応点の推定結果。白点が左眼非対応点が存在すると推定された点を表す。

知覚を説明できるステレオグラムである。このステレオグラムでは、遮蔽面と非遮蔽面の視差、及び非遮蔽面上に存在する非対応点が与えられる。

図 4.7がシミュレーションに用いたステレオグラムである。画像のサイズは  $100 \times 100$  画素、ドット密度は 50 % である。このステレオグラムを融合視すると、中央部分に四角形が浮き上がって知覚される。背景部分の視差は 0 画素、中央の四角形の視差は 4 画素になっている。この四角形の左側の境界付近に左眼非対応点が、右側の境界付近に右眼非対応点が存在する。

図 4.8がシミュレーション結果である。図 4.8(a)が、視差検出細胞によって検出された視差を表す。視差はグレースケールで表され、明るい色はより近い視差を、暗い色はより遠い視差を表す。図は左眼座標で表している。図 4.8(b)は、左眼非対応点検出細胞によって検出された左眼非対応点を表す。白色の点が、左眼非対

応点が検出された位置である。視覚系の左側の境界付近で検出できているのが分かる。

### 4.3.2 da Vinci ステレオグラム

da Vinci ステレオグラムとは、非対応点と遮蔽面の視差しか与えられていないようなステレオグラムである(図 1.13, 1.14, 1.15 参照)。これらのステレオグラムを融合視すると、非対応点は遮蔽面より奥に存在するように知覚される。この知覚は、非対応点の奥行が近傍の対応点の視差からの補間によって決まるとする、従来の両眼視アルゴリズムでは説明できない。なぜなら、ステレオグラム中には、遮蔽面より奥の視差を与えるような対応点は存在しないからである。本論文で提案するアルゴリズムは、オクルージョンの拘束条件から、非対応点の取り得る奥行が遮蔽面より奥に制限されるため、この知覚を説明できる。

図 4.9 の上図が、シミュレーションに用いた da Vinci ステレオグラムである。このステレオグラム中のほとんどの点は視差 0 画素で対応するが、左眼画像の左側、及び右眼画像の右側にそれぞれ、両眼対応しない非対応点が描かれている。このようなステレオグラムは、図 4.9 の下図のような 3 次元構造から生成される。下図に示した手前の面には、面全体にドットが描かれている。しかし、奥の面には、片眼からしか見えない小領域にしかドットが描かれていない。ステレオグラム中のドットが描かれている領域のサイズは  $54 \times 100$  画素、ドット密度は 50 % である。このステレオグラムを融合視すると、両端の非対応領域が中央のドットより奥に知覚される。

図 4.10 がシミュレーション結果である。図 4.10(a) が、視差検出細胞によって検出された視差を表す。図は左眼座標で表している。左眼非対応領域の視差は、両眼対応する領域の視差より奥に推定されている。図 4.10(b) は、左眼非対応点検出細胞によって検出された左眼非対応点を表す。

### 4.3.3 Panum's limiting case

Panum's limiting case とは、図 4.11 の上図のステレオグラムのように二つの対応候補が存在し、どちらかを偽対応と決定する手がかりのないような状況のことである。このようなステレオグラムは、図 4.11 の下図のように、同一視線上に複数の物体が存在するときに生成し得る。もちろん物理的には、右眼像の線分は手前の物体の像であるので、左眼像の右側の線分との対応が正しい対応であり、左側の線分との対応は偽対応であるということができる。しかし、このステレオグラムを融合視すると、ちょうど図 4.11 の下図に示したような 3 次元構造が知覚さ

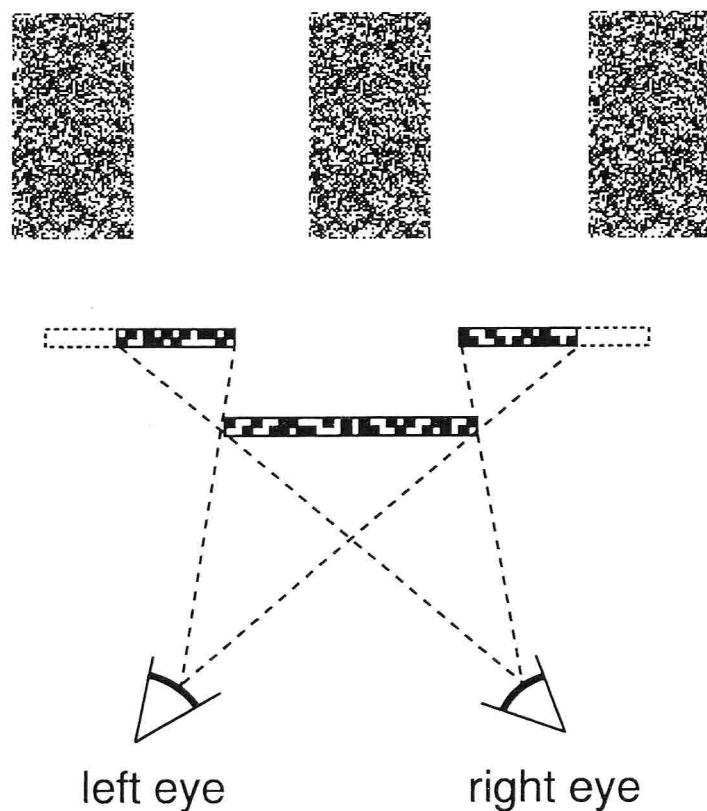


図 4.9: 上図: da Vinci ステレオグラム。左右の図が左眼画像、中央の図が右眼画像である。左眼画像の左端、及び右眼画像の右端にそれぞれ、両眼対応しない非対応点が存在する。下図: 上のステレオグラムを生成する 3 次元構造の鳥瞰図。奥の面には、片眼からしか見えない小領域にしかドットパターンが描かれていない。

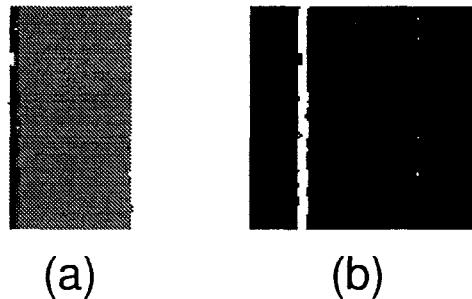


図 4.10: da Vinci ステレオグラムのシミュレーション結果。左眼座標で表されている。(a) 視差の推定結果。視差はグレースケールで表され、明るい色が近い視差を表す。(b) 左眼非対応点の推定結果。白点が、左眼非対応点が存在すると推定された点を表す。

れる。この知覚は、二つの対応候補がどちらも除去されずに知覚されているとみなすことで説明可能である。このことから、両眼視モジュールでは一意性の拘束条件が完全には実現されておらず、多重対応 (multiple match) が許容されているという提案がなされている [2, 78, 105, 106]。

本論文で提案した両眼視アルゴリズムは多重対応を許容しない。しかし、非対応点を考慮していることから、手前の対応を正しい対応として残して、奥の対応を構成していた左眼像の左側の線分を左眼非対応点とみなすことが可能である。この場合、手前の対応が遮蔽物となる。この非対応線分の奥行拘束領域は、左眼像の左側の線分を通る左眼の視線と、右眼像の線分を通る右眼の視線との交点の一点に確定する。従って、多重対応が起こって二つの対応候補がそのまま残る場合の3次元構造と、全く同じ構造を再構成できる。このアルゴリズムは、前述のように手前の対応が正しい対応で、左眼像の左側の線分が対応を持たないという物理的に正しい解釈によって構造復元を行なっている。

図 4.11 のステレオグラムを用いたシミュレーション結果を図 4.12 に示す。図 4.12(a) が推定された視差を左眼座標で表したものである。右側の線分は手前に、左側の線分は奥に推定されている。図 4.12(b) が、検出された左眼非対応点である。左側の線分は左眼非対応とみなされていることが分かる。

次に、図 4.13 のようなステレオグラムを考えてみる。このステレオグラムは図 4.11 と同じ 3 次元構造から生成される。そして、融合視して得られる知覚も図 4.11 の場合と全く同じである。しかし、この図では二つの線分のコントラストは逆になっている。コントラストの極性が逆の場合は、両眼融合しないことが知られている [24]。つまり、このステレオグラムでは、始めから白い線分同士で作られる一組の対応しか得られない。従って、このステレオグラムから得られる知覚は、図

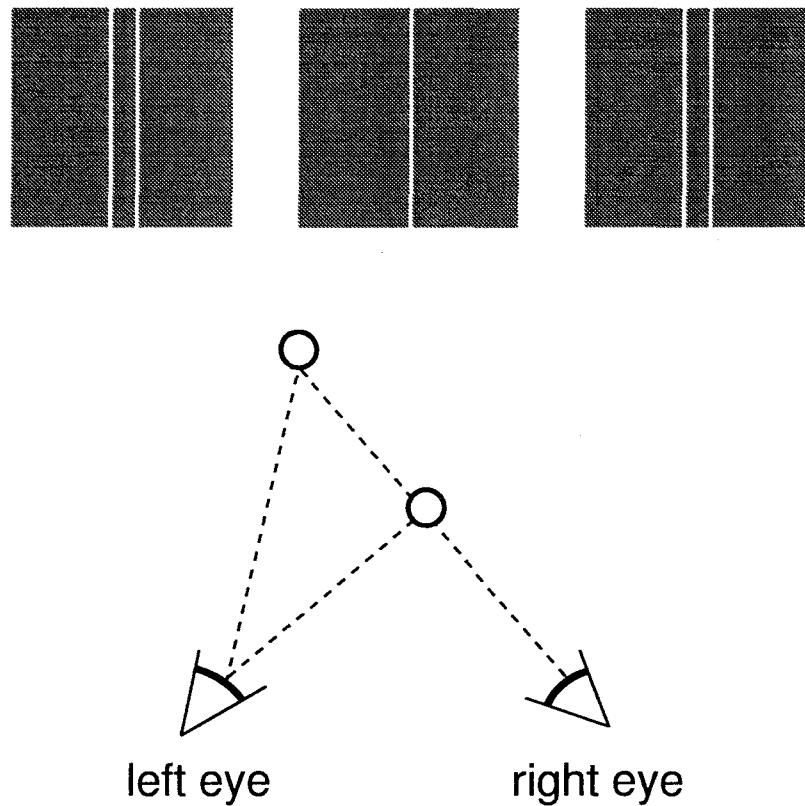


図 4.11: Panum's limiting case. 上図: Panum's limiting case のステレオグラム。左右の図が左眼画像、中央の図が右眼画像である。下図: 上のステレオグラムを生成する 3 次元構造の鳥瞰図。

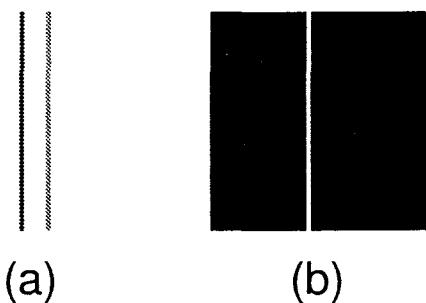


図 4.12: Panum's limiting case のステレオグラムのシミュレーション結果。左眼座標で表されている。(a) 視差の推定結果。視差はグレースケールで表され、明るい色が近い視差を表す。(b) 左眼非対応点の推定結果。白点が、左眼非対応点が存在すると推定された点を表す。

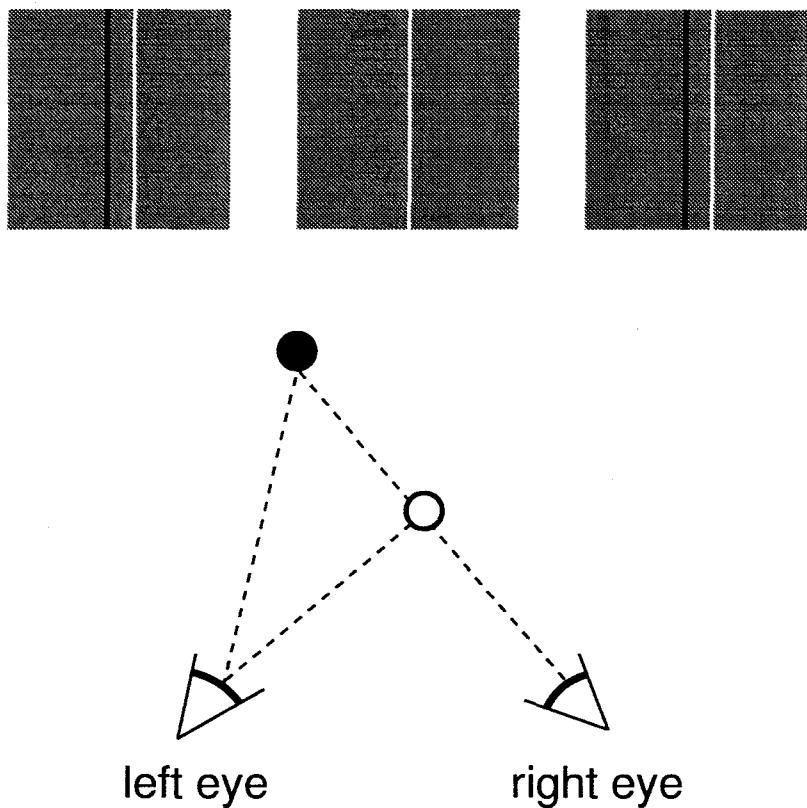


図 4.13: 二つの線分のコントラストが異なる Panum's limiting case. 上図: Panum's limiting case のステレオグラム. 下図: 上のステレオグラムを生成する 3 次元構造の鳥瞰図.

4.11 のように、二つの対応候補がいずれも除去されずに残って知覚されるという、多重対応では説明できない。

本論文で提案するアルゴリズムは、図 4.13 のステレオグラムも、図 4.11 の場合と全く同じ解釈で構造復元を行なう。異なる点は、左眼像の左側の線分が繰返し計算の始めから非対応点とみなされているという点のみである。

図 4.14 は、図 4.13 のステレオグラムを入力として用いたときのシミュレーション結果である。この結果は、図 4.12 の結果と全く同じになっている。図 4.11 のシミュレーションとの違いは、前処理の段階で既に左眼画像の左側の線分が左眼非対応点として検出されるため、繰返し計算の初期値が異なっているという点のみである。

脳の両眼視モジュールでは、Panum's limiting case をどのように処理しているのであろうか。いま、ガラス越しに物体を見るという状況を考えると、両眼視に

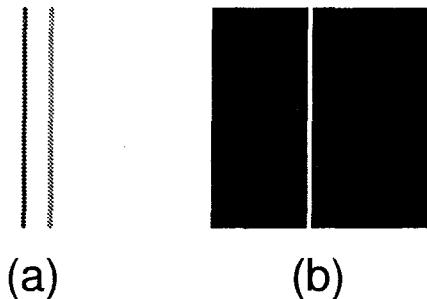


図 4.14: 図 4.13 のステレオグラムのシミュレーション結果。左眼座標で表されている。(a) 視差の推定結果。視差はグレースケールで表され、明るい色が近い視差を表す。(b) 左眼非対応点の推定結果。白点が、左眼非対応点が存在すると推定された点を表す。

よって透明面が重なった多重表面が知覚できるのは明らかである。従って、両眼視モジュールには、一意性の拘束条件を適用せず、多重対応を許容するための何らかの機構は存在するはずである。一方、図 4.13 のようなステレオグラムでは、明らかに左眼像の左の黒い線分は非対応点となり、多重対応による説明は不可能なので、本論文で提案したアルゴリズムと同等の機構によって知覚が成立していると考えられる。図 4.11 に示す Panum's limiting case は、この二つ中間の状況にあり、どちらの解釈を用いても説明可能である。物理的にいえば、本論文で提案したアルゴリズムのように、手前の対応のみを正しい対応とみなし、左眼像の左の線分は非対応点とみなすのが正しい解釈である。しかし、透明視を行なうための、多重対応を許容する機構が働いているという可能性も否定できない。心理物理学的に、図 4.11 のステレオグラムでどちらの機構が働いているのかを明らかにすることは、両眼視モジュールの機構を解明する上で非常に重要であると考えられる。

#### 4.3.4 自然画像

最後に、図 4.15 に示す、自然画像のステレオ写真を用いてシミュレーションを行なう。画像のサイズは  $256 \times 256$  画素で、8ビット(256階調)のグレースケールで表現されている。このステレオグラムを融合視すると、中央の建築物の部分が浮き上がって知覚される。

図 4.16 がシミュレーション結果である。図 4.16(a) が推定された視差を、(b) が推定された左眼非対応点を表す。この結果から、自然画像に対してもうまく働くといふことがいえる。



図 4.15: Pentagon のステレオ写真. 左右の図が左眼画像, 中央の図が右眼画像である.

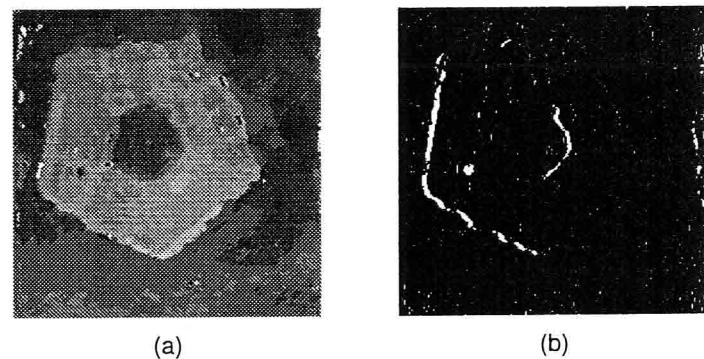


図 4.16: Pentagon のステレオ写真的シミュレーション結果. 左眼座標で表されている. (a) 視差の推定結果. 視差はグレースケールで表され, 明るい色が近い視差を表す. (b) 左眼非対応点の推定結果. 白点が, 左眼非対応点が存在すると推定された点を表す.

## 第5章 結論

本論文では、大脳視覚野で行なわれている両眼視による奥行知覚についての、理論的な研究を行なった。

本論文では、まず、両眼視の情報処理の基礎となる、脳の両眼性細胞のモデルを提案した。このモデルは、アナログ AND 回路を基にして構成したものである。まず、アナログ AND モデルの基本的な回路構成を提案し、このモデルで tuned cell および near cell, far cell が同様の回路構造で実現できることを示した。このモデルでは、near cell, far cell は視差選択性の幅が広い両眼性細胞として定義される。そして、Neocognitron 型の神経回路モデル [13, 14] を用いて、網膜から V1 の両眼性細胞までをモデル化した階層型神経回路モデルを提案した。

次に、対応点および非対応点双方を奥行手がかりとして用いた両眼視のモデルを提案した。これは、V1 以降の初期視覚領野のモデルになる。この研究で特に重要な点は、オクルージョンによって生じる両眼間非対応点を奥行手がかりとして用いている点である。従来のほとんどの両眼視のモデルは、非対応点を偽対応しか生成しないノイズと見なして無視していた。非対応点の存在を考慮したモデルでも、偽対応しか生成しない非対応点を単に対応候補から除外するだけで、奥行手がかりとしては用いていなかった。Geiger ら [17] は、非対応点を奥行手がかりとして扱うモデルを提案している。しかし、彼らのオクルージョンの拘束条件では、非対応点自体の視差を正しく推定できず、また非対応点の存在が奥行知覚に大きな影響を与える da Vinci ステレオグラムの知覚を説明することができないため、両眼視における非対応点の役割を説明するモデルとしては不十分であった。本論文では、まず、非対応点を奥行手がかりとして用いるための計算理論を提案し、どのような拘束条件を用いるべきか論じた。さらに、この計算理論を実現するためのアルゴリズムを提案し、どのような表現を用い、何を計算すべきかについて論じた。非対応点を奥行手がかりとして用いるためには、(1) 外界の構造復元、及び (2) 直接手がかりを与える領域とオクルージョン手がかりを与える領域との区別、という二つの問題を解かなければならない。このアルゴリズムでは、視差検出を行なう細胞と、非対応点検出を行なう細胞が相互作用しながら、繰返し計算を行なう。さらに、計算機シミュレーションを行ない、このアルゴリズムが da Vinci ステレオグラムなどの奥行知覚も説明できることを示した。

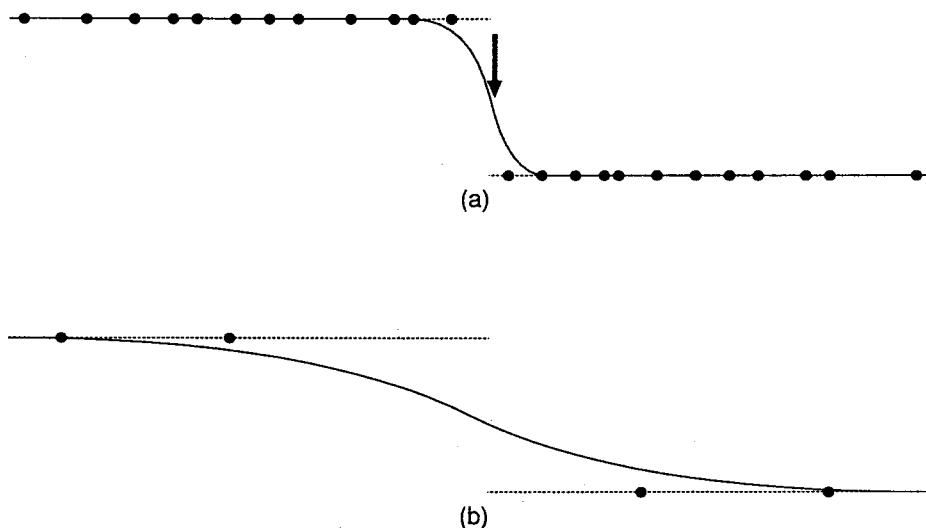


図 5.1: (a) 対応点が十分に与えられている場合. 黒点が対応点の存在する位置を表す. 破線は実際の面, 実線は滑らか拘束による補間を表す. ラインプロセスでは矢印で示した位置で不連続が検出される. (b) 対応点が十分に与えられない場合. 補間された面には視差勾配の大きな点が存在せず, 不連続点が検出できない.

今まで, オクルージョンを考慮した両眼視の理論的な研究はほとんどなされてこなかった. 本論文は, この分野の理論的な研究に対して, 先導的な役割の一端を担ったものと期待できる. さらに, 本研究では両眼視モジュールのみを対象としているが, 本研究で行なったアプローチは, オクルージョンによって生じる特徴が存在する, 運動視など他の初期視覚モジュールの理論的な研究にも応用できると期待できる.

自然環境における両眼像を考えると, 一般に, 非対応領域は画像中の小領域を占めるに過ぎない. 脳でこのような小領域のための処理を行なう意義としては, 非対応点に適用されるオクルージョンの拘束条件が, 不連続点の検出に役立つということが挙げられる. ほとんどの初期視覚のモデルでは, 不良設定問題を解くために滑らかさの拘束条件を用いている. 従って, 本来不連続である点も滑らかに補間されてしまう. この問題を解決するため, 1.1.4節で概説したラインプロセスが用いられている. このラインプロセスは, 滑らかさの拘束条件で補間される面の視差勾配が大きな点で不連続点を検出するという方法をとる. 従って, ラインプロセスは「滑らかさの拘束条件が満たされない点」という不連続点との間接的な手がかりを用いているといえる. よって, 両眼像に視差を与える対応点が十分に存在しない場合, ラインプロセスでは不連続点を検出することができない(図 5.1 を参照). その極端な例が da Vinci ステレオグラムである. da Vinci ステレオグラ

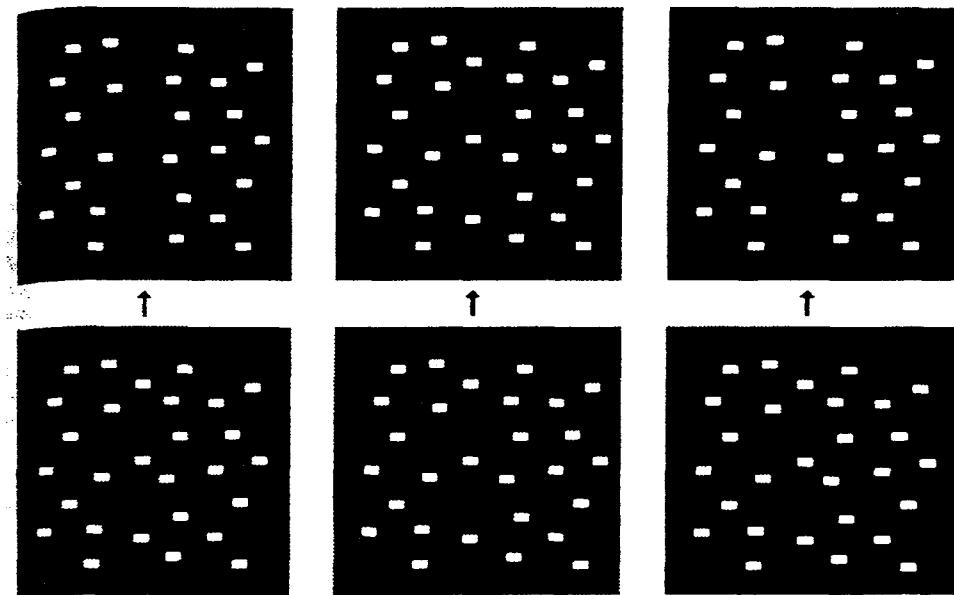


図 6.2: knife-edge ステレオグラム。左右の図が左眼画像、中央の図が右眼画像である。下図: 全ての白点が正しい対応を持つ。対応点の数が少なく、融合視しても二つの面の境界をはっきりと知覚することができない。上図: 下図とほとんど同じだが、中央の三つの白点を非対応点としている。このステレオグラムを融合視すると、面の境界がはっきり知覚できる。このことは、非対応点が不連続境界の検出に寄与していることを示している。(Nakayama & Shimojo, 1990 [64] より転載)

ムでは二つの面のどちらか一方の視差しか与えられないため、ラインプロセスでは面の境界を検出することは不可能である。

これに対し、非対応点は不連続点の直接的な手がかりになり得る。なぜなら、オクルージョンの幾何学的な拘束条件は、非対応点の位置で二つの面がオーバーラップすることを要請するためである。従って、両眼像に対応点が十分に存在しない場合でも、非対応点が検出できれば不連続境界を見つけることができる。よって、ラインプロセスとオクルージョンの拘束条件の両方を用いれば、より正確な不連続検出が可能ではないかと期待できる。実際に非対応点が不連続の手がかりになることが、Nakayama と Shimojo (1990) [64] の考案した knife-edge ステレオグラム(図 6.2)で示されている。

さらに、画像処理のアプリケーションとして用いることを考えた場合にもオクルージョンの研究は意義がある。従来は、オクルージョンによって生じる非対応領域は、画像全体から見ると相対的に小さい領域であるということから、ノイズとみなして無視されてきた。しかし、近年盛んになってきた 3D 映像などの新しい映

像生成のためには、物体の境界部がきれいに表現されることが必要である。しかし、物体の境界部では一般にオクルージョンによる非対応領域が生じるため、このような応用ではオクルージョンを考慮した処理が重要性を増す。本論文で提案する両眼視アルゴリズムは、前述のようにオクルージョンの幾何学的な拘束条件から物体の境界検出が容易になる。

最後に、非対応点を考慮した両眼視モジュールの研究において、今後の課題となる事柄についてまとめる。

### 非対応点検出細胞

まず、非対応点を奥行手がかりとして用いる情報処理の、脳のハードウェアによる実現について考える。

非対応点が奥行手がかりとなり得ることが心理物理学的に示されているため、脳には非対応点を検出するための何らかの機構が存在することは明らかである。本論文で提案したアルゴリズムでは、非対応点が存在するときに発火する非対応点検出細胞を仮定した。それでは、実際に脳ではどこで非対応点が検出されているのであろうか。

非対応点を奥行手がかりとして用いるためには、右眼のみに見える非対応点、および左眼のみに見える非対応点を検出する必要がある。従って、非対応点は、片眼優位の特性を持つ細胞によって検出されていると考えられる。V2より高次の領野ではほとんどの細胞が両眼性であることが知られているが、それらの細胞では既にどちらの眼に視覚情報が与えられたかを区別することはできない。従って、非対応点の検出は、片眼優位性を持つV1の両眼性細胞によって行なわれていると考えられる。Nakayama [62]は、心理物理学実験に基づいた考察として、非対応点を検出する細胞の候補にV1のtuned inhibitory cellを挙げている。本論文で用いた非対応点検出細胞も、様々な最適特徴、最適視差を持つ、片眼優位のtuned inhibitory cellの集合とみなすことができる。この非対応点検出細胞は、明らかに非対応点である入力が与えられた場合は前処理の段階で発火しているが、それ以外の場合には繰返し計算によって非対応点かどうかが推定されることになる。

いま、実際の脳のtuned inhibitory cellについて考えてみる。非対応点が偽対応を作るようなステレオグラムを提示すると、非対応領域に受容野を持つtuned inhibitory cellは、提示直後には偽対応によって反応が抑制されていると考えられる。本モデルと同様の計算が脳で行なわれているとすると、融合視が成立した(即ち繰返し計算が終了した)時点でこのtuned inhibitory cellが発火していれば、この細胞が非対応点を検出する機能を持つということが支持されるのではないかと考えられる。

実際に、どの細胞で非対応点が検出されているかは、いまだ明らかになっていない。しかし、今後得られるであろう非対応点の検出に関する生理学的な知見を基に、より実際の脳に近いモデルを構築することは重要である。

### 透明視

オクルージョンによって生じる非対応点は、遮蔽面が不透明の時のみ生じる。遮蔽面が(半)透明である時は、非対応点は生じず、多重表面が知覚される。

本論文で提案したアルゴリズムは、不透明面のみを考慮しており、多重対応を許容しない。従って、より多様な視覚環境に対応するためには、透明面も扱えるようモデルを拡張する必要がある。しかしこのとき、5.3節でも述べたように、多重表面を表現する機構と、非対応点を奥行手がかりとして用いる機構とを、どのように融合させたら良いのかが問題になる。特に、Panum's limiting caseのように、明らかに遮蔽面は不透明であるが、多重対応による説明も可能な場合は、脳ではどちらの機構が用いられているのかが問題になる。

### 視方向の知覚

視方向(visual direction)の知覚については、一般的に視方向の原理が成り立つといわれている。即ち、両眼中央の仮想視点(サイクロプスの眼)を原点とし、物体は、(1)両眼融合時には各眼の視軸からの偏位角を平均した方向、(2)単眼の場合には单眼の視軸からの偏位角を仮想視点の方向軸からの偏位角として投影した方向、に知覚される。

しかし、オクルージョンが存在する両眼視環境の場合には、サイクロプスの眼の位置はもちろん、3次元空間中のどの位置に仮想視点をおいたとしても、非対応領域の一部は遮蔽物と重なってしまう。従って、もし視方向の原理が厳密に成り立っているならば、非対応領域の一部は知覚されないことになる。

大塚と矢野[71, 72]は、オクルージョンの存在する環境において、立体視によって知覚される視方向を測定する心理物理学実験を行なっている(Erkelenzとvan de Grind[11]も参照)。そして、(1)非対応点は全て知覚される、(2)非対応領域を含む被遮蔽面の視方向は、遮蔽面の視方向と比較して相対的に水平方向に圧縮して知覚される、という結果を報告している。このことは、オクルージョンの存在する環境下では、厳密には視方向の原理が成立しないことを示している。そして、視空間では、被遮蔽面の対応領域を水平方向に圧縮して新たに非対応領域全体を配置し、眼球から見える領域が全て知覚されるようにするという再配置の機構が存在すると示唆される。

このような脳内での空間表現の機構を解明することは、非対応点を考慮した両眼視のモデルを研究する上で重要であると考えられる。

## 謝辞

本研究は、大阪大学大学院基礎工学研究科システム人間系専攻人間情報科学講座福島研究室で行なわれました。

本研究を進めるにあたり、福島邦彦教授に御指導を賜わりました。ここに謹んで深謝の意を表します。また、数々の御討論、御協力を賜わりました倉田耕治講師、庄野逸助手、菊池眞之助手、畠中信技官に深く感謝致します。科学技術振興事業団の岡田真人博士には、本研究を進めるにあたって有益な御討論、御助言を賜わりました。ここに記し深く感謝致します。

# 索引

- アナログ AND analog AND, 45  
 安定化汎関数 stabilizing functional, 16  
 AND, 45
- 一意性の拘束条件 uniqueness constraint, 13  
 拡張された—, 60, 92
- 一撃アルゴリズム one-shot algorithm, 24
- S 細胞 S-cell, 51  
 X からの形状復元 shape-from-X, 10  
 エネルギー energy, 17, 21, 44, 78, 91  
 エネルギーモデル energy model, 43  
 エピポーラ線 epipolar line, 13  
 MRF Markov random field, 19  
 LGN, 41
- 奥行拘束領域 depth constraint zone, 59  
 オクルージョン occlusion, 26  
 オクルージョン手がかり occlusion cue, 32, 36  
 オクルージョンの拘束条件 occlusion constraint, 33, 60, 92  
 オフ中心型細胞 off-center cell, 50  
 オプティカルフロー optical flow, 32  
 オン中心型細胞, 50
- 確率的緩和法 stochastic relaxation, 21  
 隠れ変数 hidden variable, 22, 76
- 影 shadow, 26, 32  
 画像放射照度方程式 image irradiance equation, 23
- 擬透明 pseudo transparency, 68  
 Gibbs 分布 Gibbs distribution, 20  
 逆光学 inverse optics, 23  
 逆方向 feedback, 23, 77, 80  
 逆問題 inverse problem, 16  
 凝集性の原理 coherence principle, 66  
 凝集性の拘束条件 coherence constraint, 66, 92  
 協調的アルゴリズム cooperative algorithm, 14
- quadrature pair モデル quadrature pair model, 44
- 計算理論 computational theory, 10  
 ゲート細胞 gate cell, 41, 46  
 Keplerian ダイアグラム Keplerian diagram, 15
- 光学 optics, 23  
 交差視差 crossed disparity, 8  
 拘束条件 constraint, 12  
 コスト関数 cost function, ⇒ エネルギー
- サイクロプスの眼 Cyclopean eye, 16, 105

- 最大事後確率推定,  $\Rightarrow$  MAP 推定  
サポート関数 support function, 67
- C 細胞 C-cell, 54  
視界から消える going out of sight, 26  
視覚大脳皮質の計算理論 computational theory of the visual cortical areas, 22  
自己運動 self-motion, 43  
事後確率 posterior probability, 19  
視差勾配 disparity gradient, 16, 17, 21, 35, 67, 88  
事前確率 prior probability, 20  
視方向 visual direction, 105  
遮蔽,  $\Rightarrow$  オクルージョン  
順逆光学モデル forward-inverse optics model,  $\Rightarrow$  視覚大脳皮質の計算理論  
順序対応の拘束条件 ordering constraint, 16  
順方向 feedforward, 23  
条件付確率 conditional probability, 20  
初期視覚 early vision, 7  
シルエットの拘束条件 silhouette constraint, 72  
垂直視差 vertical disparity, 14  
ステレオグラム stereogram, 9  
ステレオスコープ stereoscope, 9  
スムースネス項 smoothness term, 17  
生成/消滅領域 accretion/deletion region, 26, 31  
正則化理論 regularization theory, 22  
生態光学 ecological optics, 26  
制約条件,  $\Rightarrow$  拘束条件  
ゼロ交差 zero-crossing, 14  
前額並行面 front parallel plane, 34, 66
- 双方向性 bi-directional, 23  
双方向性理論 bi-directional theory,  $\Rightarrow$  視覚大脳皮質の計算理論
- 対応点,  $\Rightarrow$  両眼間対応点  
対応問題 corresponding problem, 12  
対応要素 matching primitive, 14  
da Vinci ステレオグラム da Vinci stereogram, 31, 35, 62, 94  
単純型細胞 simple cell, 43, 51  
tuned inhibitory cell, 42, 45, 83, 104  
tuned excitatory cell, 42, 45, 82  
直接手がかり direct cue, 32
- データフィッティング項 data-fitting term, 17, 91  
適合性の拘束条件 compatibility constraint, 13  
dual matching condition, 33
- DOG difference of Gaussian, 51  
knife-edge ステレオグラム knife-edge stereogram, 103  
滑らかさの拘束条件 smoothness constraint, 13, 17, 62, 102  
narrow occlusion, 67  
near cell, 42, 47  
2結合型 MRF 2-coupled MRF, 21  
偽対応 false match, 12, 26
- ネオコグニトロン neocognitron, 50  
high dimensional representation, 78  
back matching, 33  
Panum's limiting case, 94  
Panum の融合域 Panum's fusional area, 88

- 半遮蔽点 half-occluded point, ⇒ 両眼間非対応点
- 非交差視差 uncrossed disparity, 8  
非対応点, ⇒ 両眼間非対応点
- 標準正則化理論 standard regularization theory, 16
- far cell, 42, 47
- Vieth-Müller 円 Vieth-Müller circle, 8
- V1, 41
- 複雑型細胞 complex cell, 43, 54
- 輻輳角 vergence angle, 8
- 不良設定問題 ill-posed problem, 12, 16, 26, 102
- Bayes の定理 Bayes theorem, 19
- Wheatstone ステレオグラム Wheatstone stereogram, 31, 92
- ホロプター Horopter, 8
- MAP 推定 maximum a posteriori estimate, 19, 23
- 窓問題 aperture problem, 35
- マルコフ確率場, ⇒ MRF
- モジュール module, 10
- ラインプロセス line-process, 21, 102
- Laplacian-Gaussian フィルタ Laplacian-gaussian filter, 14
- ランダムドットステレオグラム random dot stereogram, 9
- 両眼間対応点 interocularly paired point, 28, 57
- 両眼間非対応点 interocularly unpaired point, 26, 35, 57
- 両眼視 stereopsis, 7
- 両眼視差 binocular disparity, binocular parallax, 7
- 両眼性細胞 binocular cell, 41
- 良設定問題 well-posed problem, 13
- 連続性の拘束条件 continuity constraint, ⇒ 滑らかさの拘束条件
- low dimensiona representation, 79

## 参考文献

- [1] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, Vol. 2, No. 2, pp. 284–299, 1985.
- [2] R. A. Akerstrom and J. T. Todd. The perception of stereoscopic transparency. *Perception & Psychophysics*, Vol. 44, No. 5, pp. 421–432, 1988.
- [3] B. L. Anderson. The role of partial occlusion in stereopsis. *Nature*, Vol. 367, pp. 365–368, 1994.
- [4] B. L. Anderson and P. Sinha. Reciprocal interactions between occlusion and motion computations. *Proceedings of the National Academy of Science*, Vol. 94, pp. 3477–3480, 1997.
- [5] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, Vol. 76, No. 8, pp. 869–889, 1988.
- [6] P. O. Bishop, G. H. Henry, and C. J. Smith. Binocular interaction field of single units in the cat striate cortex. *Journal of Physiology*, Vol. 216, pp. 39–68, 1971.
- [7] A. Branca, G. Convertino, E. Stella, and A. Distante. Geometrically constrained optical flow estimation by an hopfield neural network. In C. von der Malsburg, W. von Seelen, J. C. Vorbrüggen, and B. Sendhoff, editors, *Artificial Neural Networks — ICANN 96*, Vol. 1112 of *Lecture Notes in Computer Science*, pp. 735–740, Bochum, Germany, July 1996. Springer-Verlag.
- [8] C. Chang and S. Chatterjee. Ranging through gabor logons — a consistent, hierarchical approach. *IEEE Transactions on Neural Network*, Vol. 4, No. 5, pp. 827–843, 1993.

- [9] G. C. DeAngelis, I. Ohzawa, and R. D. Freeman. Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature*, Vol. 352, pp. 156–159, 1991.
- [10] P. Dev. Perception of depth surfaces in random-dot stereograms: A neural model. *International Journal of Man-Machine Studies*, Vol. 7, pp. 511–528, 1975.
- [11] C. J. Erkelens and W. A. van de Grind. Binocular visual directions. *Vision Research*, Vol. 34, No. 22, pp. 2963–2969, 1994.
- [12] D. Ferster. A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *Journal of Physiology*, Vol. 311, pp. 623–655, 1981.
- [13] K. Fukushima. Neocognitron: A hierarchical network capable of visual pattern recognition. *Neural Networks*, Vol. 1, pp. 119–130, 1988.
- [14] K. Fukushima. Analysis of the process of visual pattern recognition by the neocognitron. *Neural Networks*, Vol. 2, pp. 413–420, 1989.
- [15] J. Gårding, J. Porrill, J. E. W. Mayhew, and J. B. Frisby. Stereopsis, vertical disparity and relief transformations. *Vision Research*, Vol. 35, No. 5, pp. 703–722, 1995.
- [16] D. Geiger and F. Girosi. Parallel and deterministic algorithms from mrf's: Surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 5, pp. 401–412, 1991.
- [17] D. Geiger, B. Ladendorf, and A. Yuille. Occlusions and binocular stereo. *International Journal of Computer Vision*, Vol. 14, pp. 211–226, 1995.
- [18] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 721–741, 1984.
- [19] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA, 1979. (古崎 敬, 古崎 愛子, 達敬一郎, 村瀬 晃 (訳)『生態学的視覚論—ヒトの知覚世界を探る—』, サイエンス社, 1985).
- [20] J. J. Gibson, G. A. Kaplan, H. N. Reynolds, Jr., and K. Wheeler. The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics*, Vol. 5, No. 2, pp. 113–116, 1969.

- [21] W. E. L. Grimson. *From Images to Surfaces*. MIT Press, Cambridge, MA, 1981.
- [22] S. Grossberg. A solution of the figure-ground problem for biological vision. *Neural Networks*, Vol. 6, pp. 463–483, 1993.
- [23] S. Grossberg and N. P. McLoughlin. Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks*, Vol. 10, No. 9, pp. 1583–1605, 1997.
- [24] J. M. Harris and A. J. Parker. Independent neural mechanisms for bright and dark information in binocular stereopsis. *Nature*, Vol. 374, pp. 808–811, 1995.
- [25] 早川秀樹, 和田安弘, 川人光男. 画像生成過程の順モデル・近似逆モデルに基づく非線形視覚問題の一解法. 電子情報通信学会論文誌 D-II, Vol. J77-A, No. 12, pp. 1683–1693, 1994.
- [26] Y. Hirai and K. Fukushima. A model of neural network extracting binocular parallax. *Biological Cybernetics*, Vol. 18, pp. 19–29, 1975.
- [27] 平井有三, 福島邦彦. 両眼像の対応を見いだす神経回路網のモデル. 電子情報通信学会論文誌, Vol. 59-D, No. 3, pp. 133–140, 1976.
- [28] Y. Hirai and K. Fukushima. An interface upon the neural network finding binocular correspondence. *Biological Cybernetics*, Vol. 31, pp. 209–217, 1978.
- [29] J. J. Hopfield and D. W. Tank. “Neural” computation of decisions in optimization problems. *Biological Cybernetics*, Vol. 52, pp. 141–152, 1985.
- [30] B. K. P. Horn. Obtaining shape from shading information. In P. H. Winston, editor, *The Psychology of Computer Vision*, pp. 115–155. McGraw-Hill, New York, 1975. (白井 良明, 杉原 厚吉 (訳) 『コンピュータービジョンの心理』, 産業図書, 1979).
- [31] I. P. Howard and B. J. Rogers. *Binocular Vision and Stereopsis*. No. 29 in Oxford Psychology Series. Oxford University Press, New York, 1995.
- [32] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular and functional architecture in the cat’s visual cortex. *Journal of Physiology*, Vol. 160, pp. 106–154, 1962.

- [33] S. S. Intille and A. F. Bobick. Disparity-space images and large occlusion stereo. Technical Report 220, The Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, 1993.
- [34] S. S. Intille and A. F. Bobick. Incorporating intensity edges in the recovery of occlusion regions. Technical Report 246, The Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, 1993.
- [35] S. S. Intille and F. Bobick, A. Disparity-space images and large occlusion stereo. In J.-O. Eklundh, editor, *Computer Vision — ECCV '94*, Vol. 801 of *Lecture Notes in Computer Science*, pp. 179–184, Stockholm, Sweden, May 1994. Springer-Verlag.
- [36] B. Julesz. Binocular depth perception of computer generated patterns. *Bell System Technical Journal*, Vol. 39, pp. 1125–1162, 1960.
- [37] B. Julesz. *Foundations of Cyclopean Perception*. University of Chicago Press, Chicago, 1971.
- [38] 柿崎祐一. 心理学的知覚論序説. 倍風館, 1993.
- [39] G. A. Kaplan. Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics*, Vol. 6, No. 4, pp. 193–198, 1969.
- [40] L. Kaufman. Some new stereoscopic phenomena and their implications for theories of stereopsis. *American Journal of Psychology*, Vol. 78, pp. 1–20, 1965.
- [41] 川人光男. 脳の計算理論. 産業図書, 1996.
- [42] M. Kawato, H. Hayakawa, and T. Inui. A forward-inverse optics model of reciprocal connections between visual cortical areas. *Network*, Vol. 4, pp. 415–422, 1993.
- [43] 川人光男, 乾敏郎. 視覚大脳皮質の計算理論. 電子情報通信学会論文誌 D-II, Vol. J73-D-II, No. 8, pp. 1111–1121, 1990.
- [44] M. Kaye. Stereopsis without binocular correlation. *Vision Research*, Vol. 18, pp. 1013–1022, 1978.

- [45] D. Kersten, D. C. Knill, P. Mamassian, and I. Bülthoff. Illusory motion from shadows. *Nature*, Vol. 379, p. 31, 1996.
- [46] D. Kersten, P. Mamassian, and D. C. Knill. Moving cast shadows induce apparent motion in depth. *Perception*, Vol. 26, pp. 171–192, 1997.
- [47] C. Koch, J. Marroquin, and A. Yuille. Analog “neuronal” networks in early vision. *Proceedings of the National Academy of Science*, Vol. 83, pp. 4263–4267, 1986.
- [48] J. J. Koenderink and A. J. van Doorn. Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics*, Vol. 21, pp. 29–35, 1976.
- [49] S. R. Lehky and T. J. Sejnowsky. Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience*, Vol. 10, pp. 2281–2299, 1990.
- [50] L. Liu, S. B. Stevenson, and C. M. Schor. Quantitative stereoscopic depth without binocular correspondence. *Nature*, Vol. 367, pp. 66–69, 1994.
- [51] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, Vol. 293, pp. 133–135, 1981.
- [52] D. Marr. Analysis of occluding contour. *Proceedings of the Royal Society of London B*, Vol. 197, pp. 441–475, 1977.
- [53] D. Marr. *Vision — A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman and Company, New York, 1982. (乾 敏郎, 安藤 広志 (訳) 『ビジョン — 視覚の計算理論と脳内表現—』, 産業図書, 1987).
- [54] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B*, Vol. 207, pp. 187–217, 1980.
- [55] D. Marr and T. Poggio. Cooperative computation of stereo vision. *Science*, Vol. 194, pp. 283–287, 1976.
- [56] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London B*, Vol. 204, pp. 301–328, 1979.

- [57] D. Marr, T. Poggio, and S. Ullman. Bandpass channels, zero-crossings, and early visual information processing. *Journal of the Optical Society of America*, Vol. 69, pp. 914–916, 1979.
- [58] J. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association*, Vol. 82, No. 397, pp. 76–89, 1987.
- [59] J. E. W. Mayhew and H. C. Longuet-Higgins. A computational model of binocular depth perception. *Nature*, Vol. 297, pp. 376–378, 1982.
- [60] S. P. McKee, M. J. Bravo, D. G. Taylor, and G. E. Legge. Stereo matching procedes dichoptic masking. *Vision Research*, Vol. 34, No. 8, pp. 1047–1060, 1994.
- [61] N. P. McLoughlin and S. Grossberg. Cortical computation of stereo disparity. *Vision Research*, Vol. 38, No. 1, pp. 91–99, 1998.
- [62] K. Nakayama. Binocular visual surface perception. *Proceedings of the National Academy of Science*, Vol. 93, pp. 634–639, 1996.
- [63] K. Nakayama, Z. J. He, and S. Shimojo. Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn and D. N. Osherson, editors, *Visual Cognition*, Vol. 2 of *An Invitation to Cognitive Science*, pp. 1–70. The MIT Press, Cambridge, MA, second edition, 1995.
- [64] K. Nakayama and S. Shomojo. da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, Vol. 30, No. 11, pp. 1811–1825, 1990.
- [65] K. Nakayama and S. Shomojo. Toward a neural understanding of visual surface representation. In T. Sejnowski, E. R. Kandel, C. F. Stevens, and J. D. Watson, editors, *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. LV, pp. 911–924. Cold Spring Harbor Laboratory Press, New York, 1990.
- [66] N. M. Nasrabadi and C. Y. Choo. Hopfield network for stereo vision correspondence. *IEEE Transactions on Neural Networks*, Vol. 3, No. 1, pp. 5–13, 1992.

- [67] N. M. Nasrabadi, S. P. Clifford, and Y. Liu. Integration of stereo vision and optical flow by using energy-minimization approach. *Journal of the Optical Society of America A*, Vol. 6, No. 6, pp. 900–907, 1989.
- [68] S. K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 8, pp. 824–831, 1994.
- [69] J. I. Nelson. Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, Vol. 49, pp. 1–88, 1975.
- [70] H. K. Nishihara. Practical real-time imaging stereo matcher. *Optical Engineering*, Vol. 23, No. 5, pp. 536–545, 1984.
- [71] K. N. Ogle. Theory of stereoscopic vision. In S. Kock, editor, *Psychology: A Study of a Science*, Vol. 2, pp. 362–394. McGraw-Hill, New York, 1959.
- [72] K. N. Ogle. The optical space sense. In H. Davson, editor, *The Eye*, Vol. IV, pp. 209–417. Academic Press, New York, 1962.
- [73] 大塚作一. 遮蔽のある立体視条件下での視方向の知覚. 電子情報通信学会技術研究報告, No. PRU95-84, HIP95-12, 1995.
- [74] 大塚作一, 矢野澄男. 立体画像の知覚ひずみと Poggendorff 錯視. テレビジョン学会技術報告, 第 18 卷, pp. 25–30, 1994.
- [75] I. Ohzawa, G. C. DeAngelis, and R. D. Freeman. Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, Vol. 249, pp. 1037–1041, 1990.
- [76] 奥富正敏. ステレオ視. 松山隆司, 久野義徳, 井宮淳 (編), コンピュータビジョン: 技術評論と将来展望, pp. 123–137. 新技術コミュニケーションズ, 1998.
- [77] G. F. Poggio and B. Fischer. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, Vol. 40, No. 6, pp. 1392–1405, 1977.
- [78] G. F. Poggio, F. Gonzalez, and F. Krause. Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *The Journal of Neuroscience*, Vol. 8, No. 12, pp. 4531–4550, 1988.

- [79] T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, Vol. 317, pp. 314–319, 1985.
- [80] S. B. Pollard and J. P. Frisby. Transparency and the uniqueness constraint in human and computer stereo vision. *Nature*, Vol. 347, pp. 553–556, 1990.
- [81] S. B. Pollard, J. E. W. Mayhew, and J. P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, Vol. 14, pp. 449–470, 1985.
- [82] K. Prazdny. Detection of binocular disparities. *Biological Cybernetics*, Vol. 52, pp. 93–99, 1985.
- [83] N. Qian. Computing stereo disparity and motion with known binocular cell properties. *Neural Computation*, Vol. 6, pp. 390–4–4, 1994.
- [84] N. Qian and Y.-D. Zhu. Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Computation*, Vol. 8, pp. 1611–1641, 1996.
- [85] J.-P. Roy, H. Komatsu, and R. H. Wurtz. Disparity sensitivity of neurons in monkey extrastriate area mst. *Journal of Neuroscience*, Vol. 12, No. 7, pp. 2478–2492, 1992.
- [86] M. Shizawa. Direct estimation of multiple disparities for transparent multiple surfaces in binocular stereo. In *Proceedings of IEEE 4th International Conference on Computer Vision*, pp. 447–454, Berlin, Germany, May 1993. IEEE Computer Society Press.
- [87] 志沢雅彦. ステレオ視におけるトランスペアレンシーについて. 電子情報通信学会技術研究報告, No. PRU92-137, 1993.
- [88] 志沢雅彦. 両眼立体視におけるトランスペアレンシーの計算理論と2重視差の一撃計算モデル. 電子情報通信学会論文誌 D-II, Vol. J77-D-II, No. 7, pp. 1245–1254, 1994.
- [89] S. Shimojo and K. Nakayama. Real world occlusion constraints and binocular rivalry. *Vision Research*, Vol. 30, No. 1, pp. 69–80, 1990.
- [90] H. S. Smallman and S. P. McKee. A contrast ratio constraint on stereo matching. *Proceedings of the Royal Society of London B*, Vol. 260, pp. 265–271, 1995.

- [91] D. Terzopoulos. Constraints on deformable models. *Artificial Intelligence*, Vol. 36, pp. 91–123, 1988.
- [92] 魚森謙也, 山田光穂. 立体・奥行知覚. テレビジョン学会誌, Vol. 48, No. 12, pp. 1502–1508, 1994.
- [93] A. Verri and V. Torre. Absolute depth estimate in stereopsis. *Journal of the Optical Society of America A*, Vol. 3, No. 3, pp. 297–299, 1986.
- [94] 渡部修. 視差検出細胞の階層型神経回路モデル. Master's thesis, 大阪大学, 1996.
- [95] O. Watanabe and K. Fukushima. A hierarchical neural network model for disparity discrimination cells. In *Proceedings of the 5th Tohwa University International Symposium “New Concepts to Uncover Higher Brain Functions”*, p. 137, Fukuoka, Japan, 1995.
- [96] 渡部修, 福島邦彦. Near Cell, Far Cell の神経回路モデル. 電子情報通信学会論文誌 D-II, Vol. J78-D-II, No. 9, pp. 1404–1413, 1995.
- [97] 渡部修, 福島邦彦. Near Cell, Far Cell の神経回路モデル. 3次元画像コンファレンス'95 予稿集, pp. 156–161, 1995.
- [98] 渡部修, 福島邦彦. Near Cell, Far Cell の神経回路モデル. 電子情報通信学会技術研究報告, No. NC94-87, 1995.
- [99] O. Watanabe and K. Fukushima. A neural network model of near and far cells. In *Proceedings of Fourth IBRO World Congress of Neuroscience*, p. 505, Kyoto, Japan, 1995.
- [100] 渡部修, 福島邦彦. 視差検出細胞の階層型神経回路モデル. 3D 映像, Vol. 9, No. 4, pp. 20–25, 1995.
- [101] 渡部修, 福島邦彦. オクルージョンを考慮した視差推定アルゴリズム. 電子情報通信学会技術研究報告, No. NC95-146, 1996.
- [102] 渡部修, 福島邦彦. 両眼間非対応領域を用いた両眼視差からの形状復元アルゴリズム. 日本神経回路学会第 7 回全国大会予稿集, pp. 215–216, 1996.
- [103] 渡部修, 福島邦彦. オクルージョン拘束を用いた両眼視アルゴリズム. 日本神経回路学会第 8 回全国大会予稿集, pp. 81–82, 1997.

- [104] 渡部修, 福島邦彦. 点対応を用いない両眼視からの構造復元アルゴリズム. 電子情報通信学会技術研究報告, No. NC96-187, 1997.
- [105] O. Watanabe and K. Fukushima. Stereo algorithm that extracts a depth cue from interocularly unpaired points. *Neural Networks*, to appear.
- [106] A. B. Watson and A. J. Ahumada. Model of human visual-motion sensing. *Journal of the Optical Society of America A*, Vol. 2, No. 2, pp. 322–342, 1985.
- [107] D. Weinshall. Perception of multiple transparent planes in stereo vision. *Nature*, Vol. 341, pp. 737–739, 1989.
- [108] D. Weinshall. Seeing “ghost” planes in stereo vision. *Vision Research*, Vol. 31, No. 10, pp. 1731–1748, 1991.
- [109] C. Wheatstone. Contributions to the physiology of vision – Part the first. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, Vol. 128, pp. 371–394, 1838.
- [110] R. P. Wildes. Direct recovery of three-dimensional scene geometry from binocular stereo disparity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 8, pp. 761–774, 1991.
- [111] Y. Yang and A. L. Yuille. Multilevel enhancement and detection of stereo disparity surfaces. *Artificial Intelligence*, Vol. 78, pp. 121–145, 1995.

## 関連発表論文

### 学術論文

1. 渡部 修, 福島 邦彦. Near Cell, Far Cell の神経回路モデル. 電子情報通信学会論文誌 D-II, J78-D-II, 9, pp. 1404–1413 (1995).
2. O. Watanabe and K. Fukushima. Stereo algorithm that extracts a depth cue from interocularly unpaired points. *Neural Networks*, to appear.

### 国際会議論文

1. O. Watanabe and K. Fukushima. A neural network model of near and far cells. *Proceedings of Fourth IBRO World Congress of Neuroscience*, p. 505 (1995).
2. O. Watanabe and K. Fukushima. A hierarchical neural network model for disparity discrimination cells. *Proceedings of the 5th Tohwa University International Symposium “New Concepts to Uncover Higher Brain Functions,”* p. 137 (1995).

### 学会大会発表

1. 渡部 修, 福島 邦彦. Near Cell, Far Cell の神経回路モデル. 3次元画像コンファレンス'95 予稿集, pp. 156–161 (1995).
2. 渡部 修, 福島 邦彦. 両眼間非対応領域を用いた両眼視差からの形状復元アルゴリズム. 日本神経回路学会第7回全国大会予稿集, pp. 215–216 (1996).

3. 渡部 修, 福島 邦彦. オクルージョン拘束を用いた両眼視アルゴリズム. 日本神経回路学会第8回全国大会予稿集, pp. 81-82 (1997).

## 研究会発表

1. 渡部 修, 福島 邦彦. Near Cell, Far Cell の神経回路モデル. 重点領域研究「脳の高次情報処理」冬のワークショップ (1995).
2. 渡部 修, 福島 邦彦. Near Cell, Far Cell の神経回路モデル. 電子情報通信学会技術研究報告, NC94-87 (1995).
3. 渡部 修, 福島 邦彦. 視差検出細胞の階層型神経回路モデル. 3D 映像, 9, 4, pp. 20-25 (1995).
4. 渡部 修, 福島 邦彦. オクルージョンを考慮した視差推定アルゴリズム. 重点領域研究「脳の高次情報処理」冬のワークショップ (1996).
5. 渡部 修, 福島 邦彦. オクルージョンを考慮した視差推定アルゴリズム. 電子情報通信学会技術研究報告, NC95-146 (1996).
6. 渡部 修, 福島 邦彦. 点対応を用いない両眼視からの構造復元アルゴリズム. 電子情報通信学会技術研究報告, NC96-187 (1997).