分散表現による両眼視差抽出の神経回路モデル

非会員 藤井 真人、非会員 伊藤 崇之、非会員 三宅 誠 (NHK)

Neural Network Model with Population Coding for Binocular Disparity Extraction
Mahito Fujii, Takayuki Ito, Sei Miyake, Non-Member (NHK)

A neural network model with population coding for extracting binocular disparity is proposed. The population coding is expected to represent binocular disparity with hyperacuity. In order to solve the false target problem in stereo vision, some constraints, such as compatibility, uniqueness and smoothness, are effective. In general, the uniqueness constraint is realized by selecting an appropriate binocular disparity among some candidates. Accordingly, it is difficult that the uniqueness constraint is realized in a system with population coding.

The presented neural network model realizes the uniqueness constraint and population coding simultaneously. The model also interpolates binocular disparities at pixels where features to extract binocular disparity do not exist. The model is based on physiological findings.

It is shown by computer simulations that the model is able to extract almost correct binocular disparities in terms of center of gravity of distributed binocular disparities from random-dot stereogram forming a sphere and a plane, as well as two planes. In other words, the model is effective to extract binocular disparity from curved objects by means of population coding.

キーワード: 両眼視差、神経回路モデル、分散表現

1. まえがき

ランダムドットステレオグラムを用いることにより、人間が両眼視差のみでも奥行きを感じられることを容易に知ることができる⁽¹⁾。この事実を説明するために、従来から、両眼視差を抽出する神経回路モデルの構築に関心が持たれてきた。特に、計算論的立場から、多くの両眼視差抽出の神経回路モデルが提案されてきている⁽²⁾⁽³⁾⁽⁴⁾⁽⁵⁾⁽⁶⁾⁽⁷⁾⁽⁸⁾。

これらのモデルが解決すべき中心的課題は、不良設定問題の一つである対応点問題である⁽⁸⁾。Marr や Poggio らは、付加情報として自然界における物理的な拘束条件を利用することで、不良設定問題を良設定問題に変更可能なことを示した⁽⁸⁾⁽⁹⁾。特に、Marr が提案した、適合性、一意性、連続性の三つの拘束条件は、現在でも多くのモデルの中で使われ、その有効性が実証されている。

一方、生理学などの知見により、脳の中の情報表現の一形態として、分散表現 (Population Coding) というものが知られている。Lehky らは、両眼視差に対する人間の分解能の心理物理学的データが、分散表現を持つ神経回路モデルにより説明できることを示した⁽¹⁰⁾。両眼視差の分散表現は、画素単位での視差量よりも細かい両眼視差が表現できるものと期待される。

ところで、一意性の拘束条件と分散表現は、厳密な意味においては同時に成り立たない。事実、提案されているほ

とんどの神経回路モデルにおいては、一意性の拘束条件を、各網膜位置に対し一つの細胞モデル (ある特定の両眼視差を抽出している細胞モデル) のみが出力する、という単一表現によって実現している。

そこで、本論文では、緩和された一意性の拘束条件と、分散表現が同時に成り立つような視差抽出の神経回路モデルを提案する。本モデルでは、一意性の拘束条件により誤対応が取り除かれ、かつ分散表現により画素以下の両眼視差の表現が可能となっている。また、両眼視差が抽出できなかった箇所に両眼視差を補間する機構もモデル化する。本モデルの構築に当たっては、できるだけ多くの生理学的知見を取り入れている。

誤対応の多く生じるランダムドットステレオグラムを用いたシミュレーション実験により、本モデルが、球面のように両眼視差が連続的に変化する物体に対しても有効であることを示す。

以下、2章において参考にした生理学的知見について述べ、3章で提案した神経回路モデルについて説明する。4章では、シミュレーション実験の結果を示す。

2. 生理学的知見

この章では、モデルを構築するために参考とした生理学的知見について述べる。

<2.1> 両眼視差抽出のための図形的特徴

ネコやサルは視覚野、特に第1次視覚野では、特定の方向に傾いたエッジや線分に特異的に反応する細胞が多く見

つかる⁽¹¹⁾。また、同じ方向に傾いたエッジや線分を両眼に同時に与えた時、両眼での刺激の位置のずれ(両眼視差)が特定の量になると強く反応する細胞(以下、両眼視差抽出細胞と呼ぶ)も見つかっている⁽⁴⁾⁽¹²⁾。この細胞は、両眼立体視に寄与している細胞であると考えられる。

<2.2> 両眼細胞の方位選択性

両眼視差に特異的に反応する細胞の最適方位について Maske らは、エッジや線分によく反応する細胞の場合には、垂直方向に近い最適方位を持つ細胞のみに両眼視差抽出能力があることを報告している⁽¹³⁾。彼らの報告によれば、これらの細胞が持つ最適方位の分布範囲は、垂直から±約30°以内である。一方、短い線分によく反応する細胞では、最適方位が水平のものでも両眼視差抽出能力を持つという事実も発見した。但し、“線分”と“短い線分”という言葉の本論文での使い分けは、線分の長さが細胞の受容野(細胞が影響を受ける網膜上の範囲)の幅に比べて長いか短いかによっている。線分の長さが受容野の幅より短い場合に“短い線分”と呼ぶ。この短い線分に特異的に反応する細胞の受容野は、受容野中心部の興奮領域を、全周囲に渡って取り囲むような強い抑制領域がある構造であることも調べられている。

さて、エッジ、あるいは線分によく反応する細胞の最適方位分布が垂直から±約30°以内に限定されている理由としては、以下のことが考えられる。エッジや線分は細胞の受容野の幅を越える長さを持つため、細胞側から見ると、水平に横たわるエッジや線分が横方向にずれてもその違いは分からない。従って、左右網膜上の水平位置のずれである水平視差は、これら水平方位のエッジや線分を用いて測ることができない。この状況は、水平に近い方位を持つエッジや線分においてもほぼ同様である。一方、線分の端が受容野内に入る短い線分では、最適方位が水平であっても端点を手がかりにして水平視差を測ることができる。これらの事実に応じ、意味のある両眼視差を抽出することのできる細胞のみが脳の中に存在しているのだと考えられる。

<2.3> ランダムドットステレオグラムに応答する細胞

ランダムドットステレオグラムの両眼視差に特異的に反応する細胞がサルの大脳皮質で発見されている⁽¹⁴⁾。同時に、これらの細胞のほとんどが複雑型細胞に分類されることも報告されている。従って、第一次視覚野の複雑型細胞のレベルにおいて、すでに対応点問題が解かれていることになる。

3. 両眼視差抽出の神経回路モデル

<3.1> モデルの概要

2章で述べた生理学的知見を参考にし、以下の性質を持つ両眼視差抽出の神経回路モデルを構築した。

- 両眼視差を抽出する単純型細胞モデルと、対応点問題を解く役割の複雑型細胞モデルが階層的に結合している。
- 各単純型細胞モデルは、両眼視差抽出の手がかりと

して、最適な特徴(エッジ、線分、短い線分のいずれか)と最適な方位を持つ。

- 複雑型細胞モデルの相互結合により、繰り返し処理の後、誤対応が取り除かれ、かつ、両眼視差の抽出できなかった所に両眼視差が補間される。
- 一つの複雑型細胞モデルにより両眼視差を表すのではなく、複数の複雑型細胞モデルの出力で分散して両眼視差を表す(図1)。

上記(c)、(d)の機構を同時に成立させることが、本神経回路モデル構築の主要テーマであるが、そのメカニズムについては、<3.3> (iii) で述べる。

図1 両眼視差に対する単一表現と分散表現
Fig.1. Representations of one disparity by single cell(interval coding) and multi-cells(population coding).

図2 神経回路モデルの基本構造

Fig.2. Basic hierarchy of the presented neural network model.

<3.2> モデルの基本構造

図2にモデルの基本構造を示す。図中、太線の四角は、一つの複雑型細胞モデル(C-CELL)を中心とした両眼視差

抽出単位を示し、中太線で書かれた四角は、同一処理を行う細胞モデル群を示す。細線の四角は、一つの細胞モデルを示している。ここで図中のカッコ内の L, R は、それぞれ、右眼、左眼からの情報のみを扱っていることを示す。

本神経回路モデルでは、網膜視細胞モデル (PR) に入った明るさ情報が、まず網膜神経節細胞モデル (RG) の持つ DOG フィルターによりコントラスト強調される。その処理結果は、外側膝状体細胞モデル (LGN) と、第一次視覚野四層にあり同心円状の受容野を持つ細胞のモデル (本論文では中継細胞モデル (RELAY) と呼ぶ) を経由して、単純型細胞モデル (S-CELL) へと送られる。

単純型細胞モデルは、方位情報を持つエッジ、線分、あるいは短い線分という特徴のどれかを手がかりとし、左右の画像中の特定の位置にそのような特徴が存在する時、大きな出力を出す。実際には、この細胞モデルは、左右で特定の両眼視差分だけずれた位置に受容野中心を持ち、事前に、左右の結合係数分布中に特徴の形状パターンを記憶している。入力層に画像が入力されると、単純型細胞モデルは、記憶している形状パターンと左右の受容野内の入力形状パターンを比較し、その一致度を出力する。各単純型細胞モデルは、一つの方位と一つの特徴を手がかりとしており、神経回路モデル全体としては、各網膜上の一点において、方位数×特徴の種類数だけの手がかりを利用して両眼視差を抽出することになる。ここで、各特徴ごとの最適方位範囲は、先に述べた Maske らの生理学的知見に従い、エッジ、線分に対しては垂直を基準にして $\pm 30^\circ$ 、短い線分に対しては全方位とした。本細胞モデルの段階では、まだ誤対応が完全には除去できていない。

複雑型細胞モデル (C-CELL) は、単純型細胞モデルから入力を受け、複雑型細胞モデル間の相互作用による繰り返し処理 (一意性の拘束条件の実現) により、誤対応を取り除いた望ましい両眼視差情報を抽出する。また、単純型細胞モデルで両眼視差が抽出できなかった所に両眼視差を補間する (連続性の拘束条件の実現)。この時、複雑型細胞モデルは、分散表現も実現する。

<3.3> 各細胞モデルの機能

以下、本神経回路モデルで用いた各細胞モデルについて述べる。但し、この神経回路モデルでは、外側膝状体細胞モデルと中継細胞モデルを、単に情報の中継をするだけの機能を持つものとして扱い説明を省いた。

(i) 網膜神経節細胞モデル (RG)

左右入力画像をそれぞれ、 $u_{1L}(x, y)$ 、 $u_{1R}(x, y)$ と表したとき、網膜神経節細胞モデルの左右の出力 $u_{2L}(x, y)$ 、 $u_{2R}(x, y)$ はそれぞれ以下の式で表現できる。

$$u_{2L(j)}(x, y) = \varphi \left[(-1)^{j-1} \iint_{S_2} g_2(\xi, \eta) u_{1L}(x + \xi, y + \eta) d\xi d\eta \right] \quad (1)$$

$$u_{2R(j)}(x, y) = \varphi \left[(-1)^{j-1} \iint_{S_2} g_2(\xi, \eta) u_{1R}(x + \xi, y + \eta) d\xi d\eta \right]$$

(2)

ここに、 $j=1$ のときはオン中心型細胞モデル、 $j=2$ のときはオフ中心型細胞モデルの出力を表す。S2 は網膜神経節細胞モデルの受容野領域を表している。

$g_2(\xi, \eta)$ は DOG 関数であり、

$$g_2(\xi, \eta) = \frac{1}{2\pi\alpha^2} \exp \left[-\frac{\xi^2 + \eta^2}{2\alpha^2} \right] - \frac{1}{2\pi\beta^2} \exp \left[-\frac{\xi^2 + \eta^2}{2\beta^2} \right] \quad (3)$$

と表される ($\beta > \alpha > 0$: 定数)。さらに φ は、

$$\varphi[X] = \text{MAX}(X, 0) \quad (4)$$

の形の閾値関数である。

本細胞モデルは、明るさ情報の直流分をカットし、画像のコントラスト強調を行う。この網膜神経節細胞モデルの出力は、外側膝状体細胞モデルと中継細胞モデルによって中継され、単純型細胞モデルへと入力される。

(ii) 単純型細胞モデル (S-CELL)

本細胞モデルは、基本的にはネオコグニトロン⁽¹⁵⁾で用いられている細胞モデルを参考にして考案した両眼視差抽出の細胞モデルと同じである⁽¹⁶⁾。但し、今回のモデルでは、Ohzawa らの知見に基づき⁽¹⁷⁾、受容野内の正負の領域が反転した関係にある細胞間で相互抑制をかけ、さらに、その出力には 2 乗特性を持たせた。本細胞モデルの出力を $u_{3(k,h)}(x, y, d)$ とすると、

$$u_{3(k,h)}(x, y, d) = (\tilde{u}_{3(k,h)}(x, y, d))^2 \quad (5)$$

$$\tilde{u}_{3(k,h)}(x, y, d) = u'_{3(k,h)}(x, y, d) - u'_{3(\hat{k}, \hat{h})}(x, y, d) \quad (6)$$

$$u'_{3(k,h)}(x, y, d) = r\varphi \left[\frac{u_{3e(k,h)}(x, y, d) + \sigma_3}{b_{(k,h)}(d)u_{3i}(x, y, d) + \sigma_3} - 1 \right] \quad (7)$$

と表される。

ここに、パラメータ h は、両眼視差抽出の手がかりとして用いた特徴の種類 (エッジ、線分、短い線分の 3 種類。但し、線分、短い線分に関しては興奮性と抑制性の領域を入れ替えた 2 種類のパターン群が存在する) を表し、 k は、各特徴に対する細胞モデルの最適方位を表す。 d は、各細胞モデルが抽出しようとしている最適両眼視差を表す。また、(6) 式の $\hat{k}$ 、 $\hat{h}$ は、 h 、 k の組合せでできる受容野の興奮性と抑制性の領域が入れ替わった受容野を構成するものであることを示す。 σ_3 は定数である。

$u_{3e(k,h)}(x, y, d)$ は興奮性の入力成分であり、以下の式で計算される。

$$u_{3e(k,h)}(x,y,d) = EL + ER \quad (8)$$

$$EL = \sum_{j=1}^2 \iint_{S_3} a_{3L(j,k,h)}(\xi,\eta) u_{2L(j)}(x+\xi,y+\eta) d\xi d\eta \quad (9)$$

$$ER = \sum_{j=1}^2 \iint_{S_3} a_{3R(j,k,h)}(\xi,\eta,d) u_{2R(j)}(x+d+\xi,y+\eta) d\xi d\eta \quad (10)$$

さらに、 $u_{3i}(x,y,d)$ は抑制性入力を表しており、

$$u_{3i}(x,y,d) = \sqrt{IL + IR} \quad (11)$$

$$IL = \sum_{j=1}^2 \iint_{S_3} g_3(\xi,\eta) (u_{2L(j)}(x+\xi,y+\eta))^2 d\xi d\eta \quad (12)$$

$$IR = \sum_{j=1}^2 \iint_{S_3} g_3(\xi,\eta) (u_{2R(j)}(x+d+\xi,y+\eta))^2 d\xi d\eta \quad (13)$$

$$g_3(\xi,\eta) = \frac{1}{2\pi\gamma_3^2} \exp\left[-\frac{\xi^2 + \eta^2}{2\gamma_3^2}\right] \quad (14)$$

の各式で計算される。ここに、 γ_4 は定数である。

係数 $a_{3L(j,k,h)}, a_{3R(j,k,h)}, b_{(k,h)}(d)$ は、各特徴を検出するのに適した結合分布であり、ネオコグニトロンで用いられている学習手法によって得たものである。但し、ここでは各細胞モデルが記憶すべき形状パターン(学習パターン)を事前にそれぞれ一つ決めている。従って、各細胞モデルは、学習パターンと同じ入力パターンが与えられた時、もっとも強く反応する。図3に今回用いた学習パターンの例を示した。(但し先にも述べたように、線分、短い線分に関しては、図形の白黒を反転した学習パターンも用意した)。

図3 単純型細胞モデルが利用する図形的特徴の例(学習パターンの例)

Fig.3. Features to be used in S-CELL to extract binocular disparities(training patterns).

(iii) 複雑型細胞モデル(C-CELL)

単純型細胞モデルで各特徴と各方位ごとに抽出された両眼視差情報は、この複雑型細胞モデルで統合される。さらに複雑型細胞モデルにおいては、連続性と一意性の拘束条件が適用され、誤対応の除去と両眼視差の補間が行われる。時刻 t の複雑型細胞モデルの出力を $u_4^{(t)}(x,y,d)$ とすると、

$$u_4^{(t)}(x,y,d) = \lambda_1 u_4^{(t-1)}(x,y,d) + \lambda_2 E1 + \lambda_3 E2 - \lambda_4 I \quad (15)$$

によりその出力が得られる。ここで $E1$ と $E2$ は興奮性の入力であり、 I は抑制性の入力である。

興奮性入力 $E1$ と $E2$ は、それぞれ、単純型細胞モデルからのボトムアップの入力と、近傍領域にある他の複雑型細胞モデルからの入力である。

ボトムアップの入力 $E1$ は単純型細胞モデルからの興奮性入力で、以下の式で表される。

$$E1 = \sum_{h \in H} \max_{k \in K} (u_{3(k,h)}^{(t-1)}(x,y,d)) \quad (16)$$

ここに、 H, K はそれぞれ全特徴種類の集合と、全方位の集合を表す。

一方、 $E2$ は、最適な両眼視差が同一で、かつ近傍にある他の複雑型細胞モデルからの出力の合計である。ここでの計算は、ある種の連続性の拘束条件を実現しようとするものである。

$$E2 = \iint_{S_4} g_4(\xi,\eta) u_4^{(t-1)}(x+\xi,y+\eta,d) d\xi d\eta \quad (17)$$

$$g_4(\xi,\eta) = \frac{\exp\left[-\frac{\xi^2 + \eta^2}{2\gamma_4^2}\right]}{\left(\iint_{S_4} \exp\left[-\frac{\xi^2 + \eta^2}{2\gamma_4^2}\right] d\xi d\eta\right) - 1} \quad (18)$$

ここに、 S_4 は本細胞モデルの受容野範囲を示す(γ_4 : 定数)。但し、 S_4 には $\xi = \eta = 0$ の点を含まない。

次に抑制性結合について説明する。一意性の拘束条件を満足するよう、左右の各同一視線上の両眼視差を抽出している複雑型細胞モデル同士に相互抑制がかけられている。但し、分散表現を実現するために、両眼視差 d を選択的に抽出する複雑型細胞モデルの場合、(22) 式で示されるように、 d に近い両眼視差が最適である複雑型細胞モデルとの間の相互抑制が弱くなるようにしてある(図4)。

図4 複雑型細胞モデルの各視線上での抑制分布の例

Fig.4. A distribution of inhibitory coefficients along a line of sight in a C-CELL.

抑制性入力 I は以下の式で表される。

$$I = ISL + ISR \quad (19)$$

$$ISL = \sum_{\zeta = dmin}^{dmax} \epsilon(d - \zeta) u_4^{(t-1)}(x,y,\zeta) \quad (20)$$

$$ISR = \sum_{\zeta=dmin}^{dmax} \varepsilon(d-\zeta) u_4^{(i-1)}(x+d-\zeta, y, \zeta) \quad (21)$$

$$\varepsilon(X) = 1 / \left(1 + \exp \left(\frac{\theta - |X|}{T} \right) \right) \quad (22)$$

ここに、 ISL と ISR は、それぞれ左視線上と右視線上での相互抑制の成分を表す。また、 $dmin \sim dmax$ は両眼視差の探索範囲を表す。以上の複雑型細胞モデル間の興奮性と抑制性の結合を、一つの複雑型細胞モデルを中心にまとめたものが図5である。図中の左右視線の各交点に対し、その位置に対応した複雑型細胞モデルが存在する。図はこれらの細胞同士の結合関係を示したものである。但し、実際の興奮性結合の範囲は2次元的な広がりを持つが、この図では1次元断面のみを示した。

図5 複雑型細胞モデル間の興奮性と抑制性の結合領域

Fig.5. Areas of excitatory and inhibitory interactions in a C-CELL.

4. シミュレーション実験

本モデルの両眼視差抽出能力を検証するために、計算機上で作成された2種類のランダムドットステレオグラムを用い、シミュレーション実験を行った。シミュレーション実験を行うに当たっては、3章に現れるすべての式中の $\int$ を $\sum$ に置き換えて計算を行った。また、両眼視差の探索範囲は $-2 \sim 15$ とした。

<4.1> モデルにおける各パラメータの値

まず、シミュレーション実験で用いた各パラメータの値について述べる。

(i) 網膜神経節細胞モデル

DOG関数の形状を決定する二つのパラメータの値は、 $\alpha = 1.5$ と $\beta = 2.4$ とした。これらは $\beta/\alpha = 1.6$ の関係を満たしている。Marrらの研究により、上記条件を満たす時、DOG関数は、ほぼ ∇G フィルタの分布に等しいことが知られている⁽⁸⁾。また、 $S2$ は半径6画素の近似的円内の範囲とした。

(ii) 単純型細胞モデル

$\gamma_3 = 5.0$ 、 $\sigma_3 = 0.05$ とした。また、 $S3$ も半径6画素の近似的円内の範囲とした。

最適方位角度は、エッジに関しては、垂直を0度とした時に、 $-30 \sim 30$ と、 $150 \sim 210$ の範囲で15度間隔ごとの10方位、線分に関しては $-30 \sim 30$ の範囲で15度間隔ごとの5方位(但し、同じ方位で受容野の正負の構成が逆転しているものも同数)用意し、短い線分に関しては、0度から始めて15度間隔で165度までの12方位を用意した(但し、この場合も線分同様受容野構成の逆転しているものを同数用意した)。

(iii) 複雑型細胞モデル

複雑型細胞モデルにおける $\lambda_1 \sim \lambda_4$ の値は以下のような

考えに従って決めた。

まず、複雑型細胞モデルの出力が満たすべき以下の二つの条件を設定した。

- ・ 複雑型細胞モデルの出力が、単純型細胞モデルの出力に比較的近くなるとともに(データフィッティングの制約)、受容野の位置が近く、かつ同じ最適な両眼視差を持つ複雑型細胞モデル間で、複雑型細胞モデルの出力が比較的滑らかにつながる(滑らかさの制約)。
- ・ 単純型細胞モデルの出力がその場所での視差抽出の確からしさを表していると考えられるため、出力の大きい場所ではデータフィッティングの制約を大きく評価し、出力の小さい場所では滑らかさの制約を大きく評価する。

この二つの条件を満たすために、まばらな観測データから元の全データを復元するための一手法である標準正規化理論から類推し、以下のような逐次式を構成した(最急降下法的解法)⁽⁹⁾⁽¹⁸⁾。

$$u_4^{(i)}(x, y, d) = u_4^{(i-1)}(x, y, d) - \delta \times (\Delta e_1 + \Delta e_2) \quad (23)$$

$$\Delta e_1 = p \{ u_4^{(i-1)}(x, y, d) - E_1(x, y, d) \} \quad (24)$$

$$\Delta e_2 = q \{ u_4^{(i-1)}(x, y, d) - E_2(x, y, d) \} \quad (25)$$

Δe_1 はデータフィッティングの誤差を減少させる項であ

り、 Δe_2 は滑らかさを促進させる項である。

また、 p 、 q は、各 (x, y, d) でデータフィッティングと滑らかさの評価の割合を制御する項である。それらの値は、以下の式で表されるように、繰り返し計算の中では定数である。

$$p = \frac{E_1(x, y, d)}{E_{MAX}} \quad (26)$$

$$q = 1 - \frac{E_1(x, y, d)}{E_{MAX}} \quad (27)$$

ここに、 E_{MAX} はすべての E_1 の中の最大値を表す。

さて、実際には 3 章で述べたように抑制性の結合もあるので、この逐次計算の中に抑制項も含める必要がある。これらをまとめて書くと、

$$u_4^{(i)}(x, y, d) = u_4^{(i-1)}(x, y, d) - \delta \times (\Delta e_1 + \Delta e_2) + \lambda_4 I \quad (28)$$

となる。この式と (15) 式を比較すると、 $\lambda_1 = (1 - \delta)$ 、 $\lambda_2 = q\delta$ 、 $\lambda_3 = p\delta$ となり、 δ 、 p 、 q より値が計算できる。また、 $\lambda_4 = 0.25\delta$ とした。興奮性の結合の広がりを決めるパラメータに関しては、 $\gamma_4 = 12.0$ および $S4$ の半径を 6 画素の近似的円内とした。

δ は、最急降下法中のステップ量に相当する。本シミュレーションでは、収束を早めるために δ の値を次式に従って変化させた。

$$\delta^{(t)} = 2 - \frac{(\text{active-cell\#})^{(t-1)}}{C} \quad (29)$$

ここに、 $(\text{active-cell\#})^{(t-1)}$ は時刻 $t-1$ の時に正の反応出力を出している複雑型細胞モデルの全個数である。 C は定数であり、 $C = 220000$ とした。

最後に、抑制の強さの分布に関するパラメータに関しては、 $\theta = 4.0$ および $T = 0.5$ とした。

<4.2> シミュレーション実験結果

(i) 両眼視差の異なる 2 平面を持つ入力画像の場合

まずはじめに、両眼視差の異なる 2 平面(両眼視差 4 と両眼視差 8)を構成するようなランダムドットステレオグラムを入力画像として用いた(図 6)。画像サイズは 193 画素 $\times$ 129 画素である。図 7 は与えた両眼視差分布である。但し、オクルージョン部分は、都合上、中間の両眼視差(6)で表してある。

図 6 入力画像(平面)

Fig.6. Input images(planes)

図 8 (a)、(b) に、初期の複雑型細胞モデルの出力と繰り返し回数 10 回の時の出力をそれぞれ示す。図中の各長方形は、左上隅に示した最適な両眼視差を持つ複雑型細胞モデルの反応値を、対応している左入力画像の座標位置に示

したもので、濃淡値がその出力値の大小を表している。

図 7 与えた両眼視差分布
Fig.7. Given binocular disparities in input images.

(a) と (b) を比べると、初期の段階で両眼視差が抽出されなかった位置に、10 回の繰り返し処理後には、両眼視差が補間されていることがわかる。また、(b) において同じ座標にあり最適な両眼視差が異なる細胞モデル間の出力を見比べると、複数の複雑型細胞モデルが反応していることがわかる。従って、両眼視差の分散表現が実現されていることがわかる。しかし、このままではどのような両眼視差が取り出されているのかわかりにくい。そこで、複雑型細胞モデルによって分散表現されている両眼視差の重心を、左入力画像上の座標毎に計算した((30)式)。

$$d_{cog}^{(i)}(x, y) = \frac{\sum_{d=dmin}^{dmax} d \times u_4^{(i)}(x, y, d)}{\sum_{d=dmin}^{dmax} u_4^{(i)}(x, y, d)} \quad (30)$$

図 9 抽出された両眼視差の重心計算結果
Fig.9. Calculated center of gravity of distributed binocular disparities in C-CELLs.

計算した重心の値を、ワイヤフレームによって表したものが図 9 である。但し、画像端の影響により画像周辺では

(a) 単純型細胞モデルの出力統合直後 (繰り返し処理前)

(b) 繰り返し処理 10 回後

図 8 複雑型細胞モデルの反応

Fig.8. Examples of responses of C-CELLs((a):initial state, (b): after 10 iterations).

図 12 複雑型細胞モデルの反応例 (繰り返し処理 10 回後)

Fig.12. Examples of responses of C-CELLs(after 10 iterations).

両眼視差がうまく抽出できない所が生じるため、本論文では x 座標 21 ~ 171、 y 座標 7 ~ 121 の範囲のみ表示した。この結果をみてわかるように、わずか 10 回程度の繰り返し処理により、全体としては与えた両眼視差とほぼ同じものが抽出されていることがわかる。但し、与えた両眼視差が急激に変化している所では、取り出された両眼視差がスパイク状に大きく外れる箇所がある。この原因としては、これらの場所で、滑らかな効果と一意性のための競合過程の関係が不安定になっているためだと考えられる。

(ii) 球面を持つ入力画像の場合

(i) の例では、知覚される形状が、奥行き感覚において離散的な二つの面を形作づくるものであった。そこで、次に、知覚される形状が滑らかな面を形成するような入力に対してモデルの検証を行った。用いた入力画像は、平面を背景にして球面が浮き出すようなランダムドットステレオグラムである(図 10)。図 11 は、想定される知覚形状を持つ両眼視差分布である。但し、実際に与えた両眼視差は離散的な値(この例では整数値)である。

Left image Right image

図 10 入力画像(球面と平面)

Fig.10. Input images(a sphere & a plane).

図 11 両眼視差分布

Fig.11. Given binocular disparities.

図 12 に繰り返し回数 10 回の時の複雑型細胞モデルの出力例を示す。また、図 13 に複雑型細胞モデルによって分散表現されている両眼視差の重心の計算結果を示す。

図 13 を見ると、球面と平面の境界部分を除き、かなり球面に近い両眼視差分布が得られていることがわかる。従って、両眼視差を分散表現することにより、連続的な表現が可能となることがわかった。

図 13 抽出された両眼視差の重心計算結果
Fig.13. Calculated center of gravity of distributed binocular disparities in C-CELLs.

5. むすび

本論文では、視覚野で見つっている両眼視差に感度のある細胞などの生理学的知見を参考にした両眼視差抽出の神経回路モデルを提案した。本モデルの特徴は、対応点問題を解く上で有効な一意性の拘束条件と、両眼視差を画素以下の解像度まで表現可能な分散表現とを同時に実現した点にある。また、対応点が検出できなかった所に両眼視差を補間する機構も実現した。シミュレーション実験により、本モデルによって分散表現された両眼視差を、それらの重心で代表させることで、かなり望ましい両眼視差が抽出できていることを示した。特に、球面に対する実験により、画素単位での視差量よりも細かい両眼視差に対してもかなりの表現能力を持っていることを示した。このことは、分散表現により、比較的少数の細胞で細かな解像度の両眼視差が表現可能であることを意味する。

ところで、分散表現されている両眼視差を次の段階の表現に変換する時、重心計算は必ずしも好ましい方法とは限らない。本来、分散表現は、複数の細胞によって両眼視差を表現することで、入力の一安定性を吸収できる頑強な表現方法だと考えられる。また、透明視にも有効なものと考えられる。分散表現から次の段階への表現(たとえば、面の表現)への変換に関しては、人間の面知覚にみられるような安定した表現への変換方法を検討する必要がある。本モデルの透明視や実画像への適用可能性については今後の検討課題としたい。

最後に本研究の機会を与えて下さった本研究所西澤台次所長、泉武博前所長に感謝致します。また、日頃から有意義なご議論をして下さった中川俊夫職員に感謝致します。

(平成 7 年 10 月 27 日受付、同 8 年 2 月 7 日再受付)

文 献

- (1) Julesz : "Foundations of Cyclopean Perception", Univ. of Chicago Press, Chicago(1971)
- (2) 平井有三, 福島邦彦: 「両眼像の対応を見いだす神経回路モデル」, 信学論, **59-D**, 133 ~ 140(1976)
- (3) N.Sugie and M.Suwa: "A scheme for binocular depth per-

- ception suggested by neurophysiological evidence", *Biol. Cybern.*, **26**, 1 ~ 15(1977)
- (4) G.F.Poggio and T.Poggio: "The analysis of stereopsis", *Ann. Res. Neurosci.*, **7**, 379 ~ 412(1984)
- (5) S.B.Pollard, J.E.Mayhew, and J.P.Frisby: "PMF: stereo correspondence algorithm using a disparity gradient limit", *Perception*, **14**, 449 ~ 470(1985)
- (6) K.Prazdny: "Detection of binocular disparities", *Biol. Cybernetics*, **52**, 93 ~ 99(1985)
- (7) R.Blake and H.R.Wilson: "Neural models of stereoscopic vision", *TINS*, **14**, 10, 445 ~ 452(1991)
- (8) デビット・マー著, 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)
- (9) T.Poggio, V.Torre, and C.Koch: "Computational vision and regularization theory", *nature*, **317**, 314 ~ 319(1985)
- (10) S.R.Lehky and T.J.Seinowski: "Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity", *J.Neuroscience*, **10**, 7, 2281 ~ 2299 (1990)
- (11) D.H.Hubel: "Eye, Brain, and Vision", Scientific American Library, New York(1988)
- (12) D.H.Hubel and T.N.Wiesel: "Cells sensitive to binocular depth in area 18 of the macaque monkey cortex", *Nature*, **22**, 5, 41 ~ 42(1970)
- (13) R.Maske, S.Yamane, and P.O.Bishop: "End-stopped cells and binocular depth discrimination in the striate cortex of cats", *Proc. R. Soc. Lond.*, **B 229**, 257 ~ 276(1986)
- (14) G.F.Poggio, B.C.Motter, S.Squatrio, and Y.Trotter: "Responses of neurons in visual cortex (v1 and v2) of the alert macaque to dynamic random-dot stereograms", *Vision Res.*, **25**, 3, 397 ~ 406(1985)
- (15) 福島邦彦: 「神経回路と情報処理」, 朝倉書店 (1989)
- (16) 藤井真人: 「両眼視差抽出機構の神経回路モデル - 等視差面における輪郭線の特定 -」, *信学論*, **J73-D-II**, 8, 1288 ~ 1296(1990)
- (17) I.Ohzawa, G.C.DeAngelis, and R.D.Freeman: "Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors", *Science*, **249**, 1037 ~ 1041(1990)
- (18) K.Ikeuchi and B.K.Horn: "Numerical shape from shading and occluding boundaries", *Artif. Intell.*, **17**, 141 ~ 184 (1981)

藤井 真人

(非会員) 1958年8月6日生まれ。81年3月名古屋大学工学部電気工学科卒業。83年3月同大学院修士課程了。同年NHK入局。87年より同放送技術研究所に勤務。主に視覚のバイオサイバネティクス、立体画像情報処理の研究に従事。電子情報通信学会、テレビジョン学会、日本神経回路学会各会員。

伊藤 崇之

(非会員) 1954年8月5日生まれ。77年3月京都大学工学部電気工学科卒業。79年3月同大学院修士課程了。同年NHK入局。81年より同放送科学基礎研究所、および放送技術研究所に勤務。主に視覚のバイオサイバネティクス、画像認識、立体画像情報処理の研究に従事。電子情報通信学会、テレビジョン学会、日本神経回路学会各会員。現在、同

ヒューマンサイエンス研究部副部長。工博。

三宅 誠

(非会員) 1947年7月21日生まれ。70年3月早稲田大学理工学部電気通信工学卒業。同年NHK入局。74年より同放送科学基礎研究所、および放送技術研究所に勤務。主に視覚のバイオサイバネティクス、パターン認識、並列処理の研究に従事。電子情報通信学会、テレビジョン学会、日本神経回路学会各会員。現在、同ヒューマンサイエンス研究部部長。工博。