

## 両眼立体視のモデル

緒形昌美

ATR 視聴覚機構研究所  
〒619-02 京都府相楽郡精華町乾谷・三平谷

### 1. まえがき

外界に存在する物体を見た場合、その像は左右の網膜上において、若干異なる位置に投影される。この位置の差（両眼視差）は、観測者と対象物体との距離によって生ずるもので、我々が奥行きを知覚するための重要な手がかりとなっている。

生体が視差を検出するメカニズムについては、視覚心理学のテーマとして古くから研究されてきた。また、このような両眼による立体視は、ロボット等への応用が期待され、工学の分野でも多くの研究がなされてきた。こうして、今日迄に様々な視差検出アルゴリズムが提案されてきたが、視差検出を特徴点の対応問題（特徴点マッチング）に帰着させているという点では、どれも共通している。すなわち、あらかじめ左右の画像からなんらかの特徴点（局所的な空間パターンの場合もある）を抽出しておき、一方の画像の特徴点が、他方のどの特徴点に対応するかを見つけ出すことで視差を検出している。従って、これらの研究の中心課題は、用いる特徴点の選択と、それらの対応付けの方法を開発することにあった。

このようなマッチング的な手法には、計算量の問題がつきまとう。なぜならば、ある特徴点に対応する点を決定するためには、複数の候補点に対してなんらかの評価関数を計算し、さらにそれらと比較することによって最適値を与えるものを捜さなければならないからである。また、視差が、特徴点の存在する位置でしか得られないことも大きな欠点となっている。

これらの問題に対し、筆者らはこれまでに画像の輝度変化の位相情報を用いた視差検出法を提案した<sup>1</sup>。これは、位相特性の異なる1対のガボア・フィルタによって、左右の各画像における輝度変化の局所的な位相を検出し、それらの差（位相差）から視差を計算するもので、特徴点マッチング法が抱える対応問題を回避でき、画像のいたるところで視差を求めることができる。さらに、平均輝度やコントラスト変化に影響されにくいという利点を持っている。しかし、この方法は生体のモデルとしては妥当性に欠ける。なぜならば、各画像の位相を求めるためには、 $\tan^{-1}$ の計算が必要となるが、このような関数が生体中で計算されているとは考えにくいからである。また、実際に計算機上にインプリメントする場合、関数  $\tan^{-1}$  の計算には四則演算等と比べて多くの時間を必要し、これも従来法の欠点といえる。

ここでは筆者らが提案した従来の視差検出法を改良し、左右各画像の位相を計算することなく、両画像の位相差の符号（-、0、+）のみから視差を計算する方法を提案する。本方法では、各空間周波数チャンネルは、求めるべき視差の取り得る範囲を限定するだけであるが、複数のチャンネルで得られた視差範囲を考慮することによって、最終的に視差を求める。はじめに、従来法について簡単に述べ、3章では、これを改良して、位相差の符号を検出する方法について説明する。また4章では、複数の空間周波数チャンネルで得られた位相差符号を、低周波数チャンネルから高周波数

チャンネルへ伝搬することによって視差を求める方法を示す。5章ではランダム・ドット・ステレオグラムを用いたシミュレーションによって、本方法と我々の従来法との比較を行う。

## 2. 位相情報による視差検出の原理

ここでは、筆者らがこれまでに提案した、位相情報による視差検出法について簡単に述べる。この従来法では、両眼画像における輝度変化の位相差として視差を計算する。このために、まず各画像の位相を求める。ただし、奥行きに関係する視差は、画像中において水平方向のずれとして現れるので、ここでは水平方向の輝度変化の位相を求める。また、視差は位置によって変化するために、位相は各位置において局所的な輝度変化から求めなければならない。このような、局所的な水平方向の位相（以後、単に位相と呼ぶ）を求めるために、入力画像に次のような2種類の空間フィルタを施す。

$$g_c(x, y) = a_c(x)b(y), \quad (1)$$

$$g_s(x, y) = a_s(x)b(y). \quad (2)$$

ここで、

$$a_c(x) = \exp[-(x/\sigma)^2] \cos(2\pi u_s x), \quad (3)$$

$$a_s(x) = \exp[-(x/\sigma)^2] \sin(2\pi u_s x), \quad (4)$$

$$b(y) = \exp[-(y/\sigma)^2]. \quad (5)$$

(1)~(5)式は2次元のガボアフィルタと呼ばれ、生体の視覚皮質における単純細胞の受容野関数の近似としてよく用いられており、図1に示すような形をしている。x方向（水平方向）はガウス関数とコサインまたはサイン関数との積であり、y方向はガウス関数である。これらは、x、y両方向のガウス関数によって、その広がり $\sigma$ に応じた空間的な局所性を与えられている。また、周波数面上でも局所性を持っており、周波数 $(u_s, 0)$ を中心とした帯域通過型フィルタになっている。従って、これらのフィルタは、局所領域において、水平方向に周波数 $u_s$ で変化する成分を抽出する。

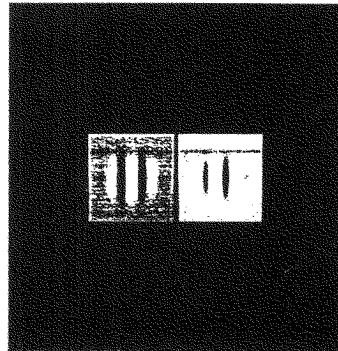


図1 ガボアフィルタ対

ところで、ある位置 $(x, y)$ におけるフィルタ $g_c, g_s$ の出力値 $o_c(x, y), o_s(x, y)$ を考えると、それらは $(x, y)$ を中心とした局所的な輝度変化の、関数 $\cos(2\pi u_s x)$ と $\sin(2\pi u_s x)$ への射影成分と考えることができる。従って、それらの比から、次のように周波数 $u_s$ 成分の位相を計算することができる。

$$p(x, y) = \tan^{-1}(o_s(x, y)/o_c(x, y)) \quad (6)$$

両眼の各画像に対してこのように位相 $p_l(x, y), p_r(x, y)$ を求め、それらの差

$$dp(x, y) = pl(x, y) - pr(x, y) \quad (7)$$

から、次のように視差を計算することができる。

$$d(x, y) = (dp(x, y) / 2\pi) \lambda \quad (8)$$

$\lambda$  はフィルタの中心周波数に対応する波長 ( $\lambda = 1/u_s$ )、 $d(x, y)$  は位置  $(x, y)$  における視差である。

ここで、(8)式が正しく視差を計算するためには、求めるべき視差  $d$  に対して

$$-\lambda/2 < d < \lambda/2 \quad (9)$$

を満たすような  $\lambda$  を持つフィルタ対を用いる必要がある。これは、位相差として求められる範囲が

$$-\pi < dp(x, y) < \pi \quad (10)$$

だからである。

### 3. 従来法の改良

2章で述べた視差検出法は、特徴マッチング法に較べて処理が簡単であり、また平均輝度やコントラスト変化等に強いという利点を有するが、位相を求めるために  $\tan^{-1}$  の計算が必要となり、生体モデルとしての妥当性や計算コストの面で問題を残している。ここでは、位相差から正しい視差を求めるのではなく、位相差の符号をもとに視差に関するおおまかな情報、すなわち、視差の解が取り得る範囲を求めることを考える。

#### 3.1 位相差符号の意味

いま、中心周波数  $u_s (=1/\lambda)$  のガボア・フィルタ対を考えると、このフィルタで検出可能な視差範囲は(9)式で与えられ、この範囲の視差に対しては、左右の画像の位相差の符号は視差の符号（注視点に対して奥か手前かに対応する）と一致する。従って、位相差符号が正であれば、求めるべき視差は

$$0 < d < \lambda/2 \quad (11)$$

の範囲に存在すると考えられる、またそれが負であれば視差が取り得る範囲は

$$-\lambda/2 < d < 0 \quad (12)$$

に限定される。このように、位相差符号を計算することによって、求めるべき視差が取り得る範囲を、検出可能な全範囲の半分に絞ることができる。

#### 3.2 位相差符号の検出

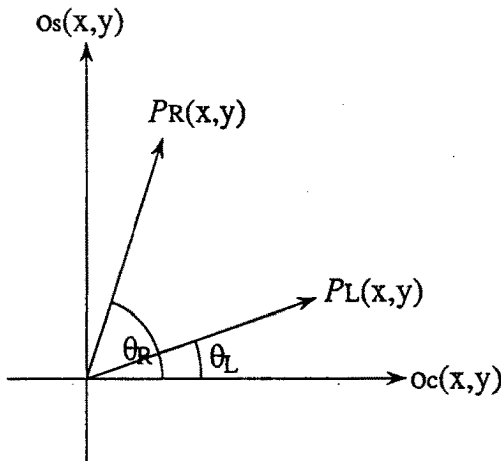


図2 ガボアフィルタ対からの出力ベクトル

図2に示すように、左右の画像に対するガボア・フィルタ  $g_c, g_s$  の出力をそれぞれ

$$P_L(x, y) = (o_{Lc}(x, y), o_{Ls}(x, y)) \quad (13)$$

$$P_R(x, y) = (o_{Rc}(x, y), o_{Rs}(x, y)) \quad (14)$$

なるベクトルとして考える。このとき、各画像の位相は、それぞれのベクトルが水平軸となす角度  $\theta_L, \theta_R$  で表され、両画像の位相差の符号とは、文字どおり  $\theta_L - \theta_R$  の符号である。しかし、この符号は2つのベクトルの相対的な位相関係、すなわち、 $P_L$  に対して  $P_R$  が反時計回りの方向に存在するか、時計回りの方向に存在しているかによって決まるもので、両出力ベクトルの外積

$$\begin{aligned} P_L(x, y) \times P_R(x, y) &= |P_L| |P_R| \sin(\theta_L - \theta_R) \\ &= o_{Lc}(x, y) \cdot o_{Rs}(x, y) - o_{Ls}(x, y) \cdot o_{Rc}(x, y) \end{aligned} \quad (15)$$

の符号として検出できる。よって、位相差符号を求めるだけであれば各位相を計算する必要はなく、 $\tan^{-1}$  の計算を回避できる。

ここで、(15)式の値が0となる場合、これは必ずしも位相差0を意味しない。位相差 $180^\circ$  の場合にも外積は0となる。また、たとえ実際の位相差が0であっても、ノイズなどの影響で外積が正確に0になるとは限らない。そこで、位相差0を検出するために、出力ベクトルの内積を計算する。

$$\begin{aligned} P_L(x, y) \cdot P_R(x, y) &= |P_L| |P_R| \cos(\theta_L - \theta_R) \\ &= o_{Lc}(x, y) \cdot o_{Rc}(x, y) + o_{Ls}(x, y) \cdot o_{Rs}(x, y) \end{aligned} \quad (16)$$

この値が大きいほど位相差0の可能性が高い。我々は、(15)式、(16)式を比較し、次のように位相差符号を決定する。

$$\begin{aligned} \text{内積} > |\text{外積}| &\quad \dots \quad \text{位相差符号 } 0 \\ \text{内積} \leq |\text{外積}| & \\ \text{外積} > 0 &\quad \dots \quad \text{位相差符号 } + \\ \text{外積} < 0 &\quad \dots \quad \text{位相差符号 } - \end{aligned} \quad (17)$$

この関係を図3に示す。検出された位相差符号が+、または-の場合には、求めるべき視差は(1)式または(2)式の範囲に存在するものとし、位相差符号0の場合にはその範囲を、

$$-\lambda/4 < d < \lambda/4 \quad (18)$$

とみなすことにする。位相差符号の検出法の構成を図4に示す。

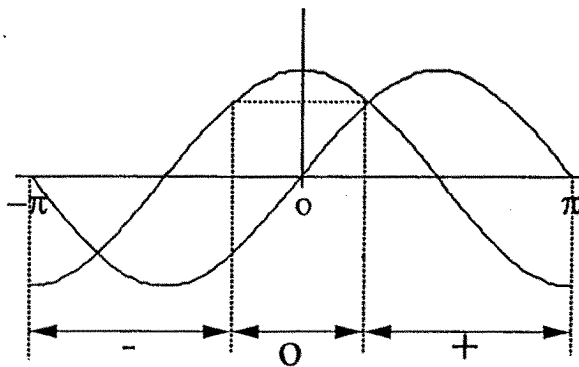


図3 位相差符号

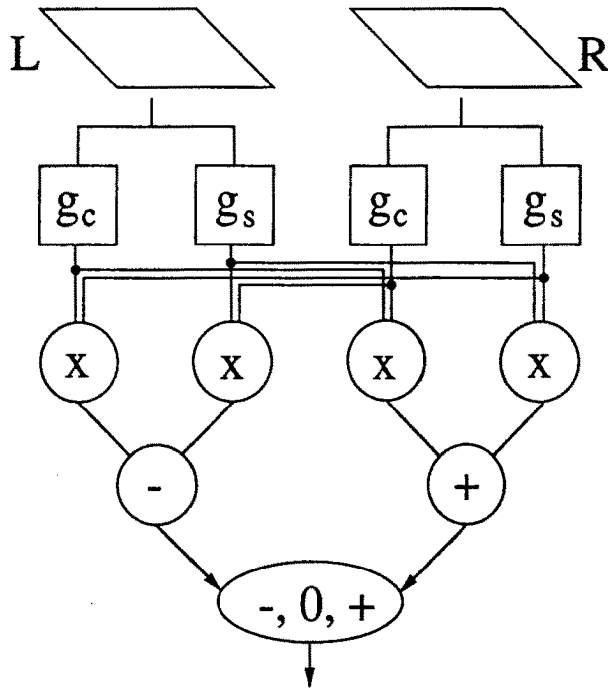


図4 位相差符号検出法

#### 4. 空間周波数チャンネルの統合

3.2で述べた方法によって求めた位相差符号が、正しく視差符号に対応するためには、用いたガボア・フィルタの中心周波数が、求める視差に対して(9)式を満たしていなければならない。すなわち、大きい視差を扱うためには、低い中心周波数を持つフィルタを用いる必要がある。一方、検出された位相差符号が示す視差候補の範囲の大きさは、(11)式または(12)式からも明かなように、フィルタの中心周波数に反比例する。従って、高周波数のフィルタは求めるべき視差を狭い範囲に絞ることができるが、低周波数フィルタでは解の曖昧性が大きくなる。これらのことから、求まる視差の曖昧性と、扱うことのできる視差の上限はトレード・オフの関係にある。ここでは、中心周波数が異なる複数のフィルタ対（空間周波数チャンネル）を用い、各チャンネルから得られる位相差符号を統合することによって、大きな視差に対しても曖昧性の少ない解を求める方法を提案する。

図5に本方法の概略を示す。なお、各チャンネルは、中心周波数が小さいものから順に  $CH_1$ ,  $CH_2, \dots$  と表すことにする。また、 $CH_1$ の中心周波数  $u_{s1}$ （波長  $\lambda_1 = 1/u_{s1}$ ）は十分に小さいものとし、入力画像に含まれるすべての視差は、(9)式の範囲に存在すると仮定する。いま、このチャンネルで検出された位相差符号を+とすると、求めるべき視差は、図中の黒で示された範囲に限定される。そこで、次のチャンネル  $CH_2$ では、この範囲の細分化を行う。これは、位相差符号を検出するときに、左画像の位置  $(x, y)$  に対して右画像の位置  $(x+d_0, y)$  におけるフィルタ出力を用いて外積  $P_L(x, y) \times P_R(x+d_0, y)$ 、内積  $P_L(x, y) \cdot P_R(x+d_0, y)$  を計算することによって実現できる。ここで、 $d_1$ は、チャンネル  $CH_1$ で得られた視差範囲の中心の値であり、図5の場合  $\lambda/4$ である。 $CH_2$ は、 $CH_1$ に較べて小さい視差しか扱うことができないが、その中心周波数  $u_{s2}$  または波長  $\lambda_2$ が

$$u_{s2} < 2u_{s1} \quad \text{または} \quad \lambda_2 > \lambda_1/2 \quad (19)$$

を満たしていれば、CH<sub>1</sub>によって与えられた範囲のすべての視差に対して正しい位相差符号を与えることができる。しかも、このチャンネルCH<sub>2</sub>は、より狭い範囲に解を絞ることができる。

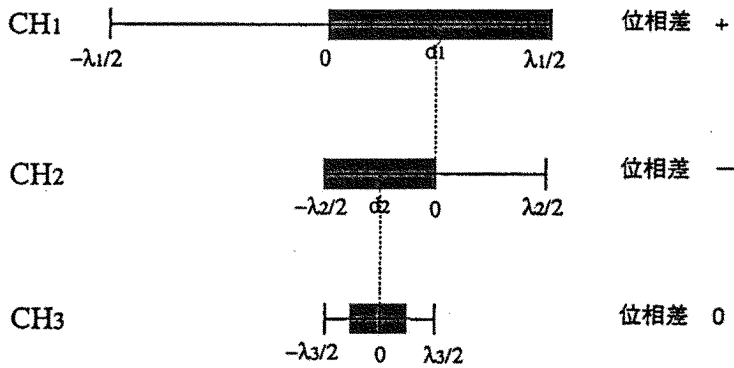


図5 周波数チャンネル間の位相差符号の伝搬

一般にチャンネルCH<sub>n</sub>では、前のチャンネルCH<sub>n-1</sub>で得られた視差範囲の中心値d<sub>n-1</sub>を用いて、外積  $p_L(x, y) \times p_R(x+d_{n-1}, y)$ 、内積  $p_L(x, y) \cdot p_R(x+d_{n-1}, y)$  を計算し、位相差符号を求める。これによって、得られる視差範囲の中心値d<sub>n</sub>を次のチャンネルに伝搬する。このように、低い周波数チャンネルで得られた視差範囲を高い周波数チャンネルで細分化していくことによって、大きな視差に対しても曖昧性を少なくすることができる。十分高い空間周波数チャンネルを用意することにより、精度の高い視差を求めることが可能となる。

ここで述べた複数チャンネルの統合法は、Marr が彼の第2アルゴリズムの中で、眼球運動と対応づけて提案した方法<sup>2</sup>と等価である。

## 5. 計算機シミュレーション

ランダム・ドット・ステレオグラムを用いて、本方法と3章で述べた従来手法の計算時間と計算精度を比較した。なお、以下では本手法をPDS (Phase Difference Sign) 法、従来法をPD (Phase Difference) 法と呼ぶことにする。計算には、UNIX マシンのMASSCOMP5600を用い、計算精度は、計算された視差と、入力画像に与えた視差が一致している画素数を全画素数で割ることによって求めた値 (以後、正答率と呼ぶ) で表した。

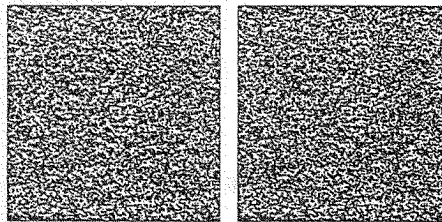


図6 ランダムドットステレオグラム

図6に入力データの例を示す。各画像の大きさは256pixels×256pixelsである。左画像は50%のランダム・ドット・パターンであり、右画像は左画像の中央128pixels×128pixelsの正方領域（ターゲット）を水平方向にずらして作られている。つまり、このずれの量がターゲットの視差となり、その他の部分（ターゲットの周辺）は視差0である。

視差検出には、中心周波数が1オクターブずつ離れた3つの空間周波数チャンネルを用いた。各チャンネルの中心周波数  $u_s$  はそれぞれ、0.0625, 0.125, 0.25cycles/pixel であり、検出可能な視差範囲はそれぞれ、8, 4, 2pixel 未満になる。従って、最も高い周波数チャンネルは、視差の解を、-1, 0または+1に限定でき、位相差符号のみを用いる本方法でも、±7pixels までの視差ならば一意に求めることができる。また各チャンネルの広がり  $\sigma$  は  $0.795/u_s$  とした。これによって、それらの周波数領域上における広がり  $\sigma$  が1オクターブになっている。

図7、図8に、ターゲットの視差が1、及び3の場合の視差検出結果を示す。図中の各画像は、各手法によって推定された視差が  $d$  であった位置を示している。また、表1に、それぞれの計算時間と正当率を示す。どちらの視差に対しても、計算時間に関しては PDS 法が2倍以上速いが、正答率は PD 法の方が2%程度優れている。

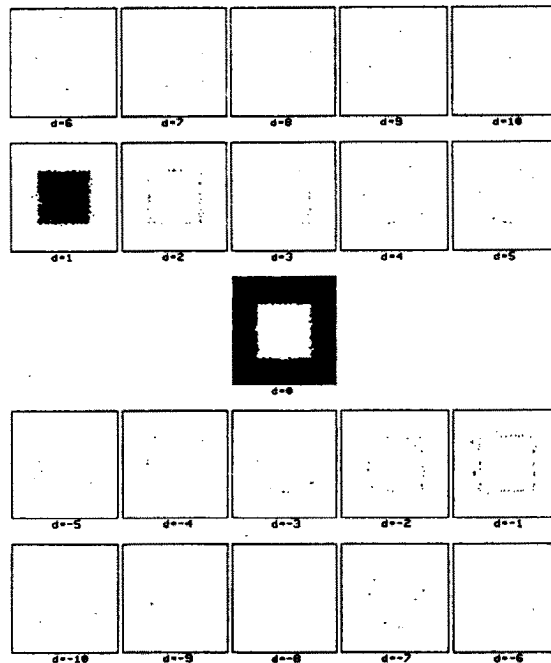


図7(a) PD法による視差検出結果（視差1,3チャンネル）

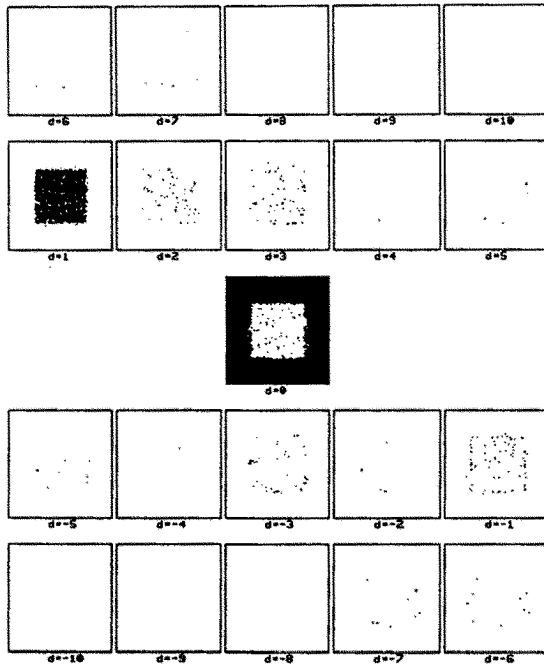


図7(b) PDS法による視差検出結果(視差1,3チャンネル)

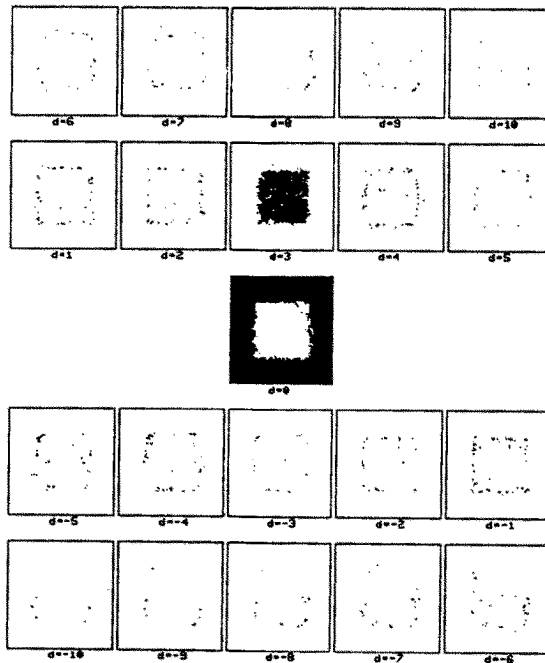


図8(a) PD法による視差検出結果(視差3,3チャンネル)



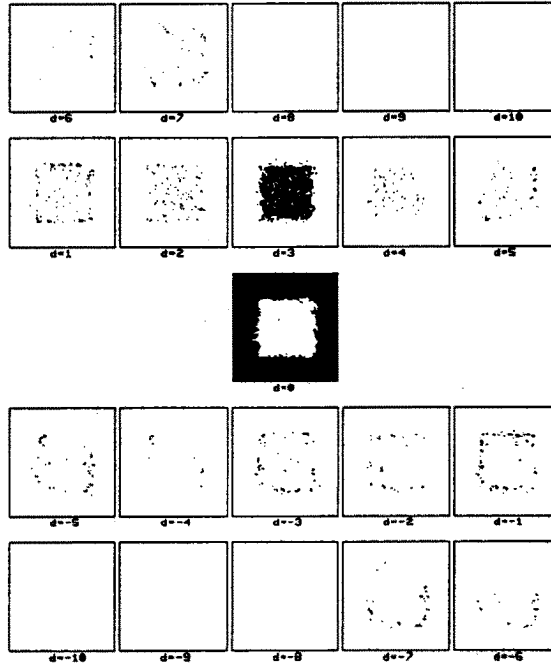


図8 (b) PDS 法による視差検出結果 (視差3, 3チャンネル)

表1 PD 法と PDS 法性能比較 (3チャンネル)

		視差 1	視差 3
PD法	正答率	98.0%	94.5%
	計算時間	40sec	
PDS法	正答率	96.6%	92.6%
	計算時間	17sec	

この正答率の違いは、次のような理由によるものと考えられる。すなわち、PD法は、各チャンネルで視差を計算するので、低い周波数チャンネルでの計算誤差をより高いチャンネルが修復することが可能であるが、PDS法では、1オクターブ毎にチャンネルを用意した場合、低い周波数チャンネルが一旦視差符号を間違えると、それをより高い周波数チャンネルで修復することは不可能となる。このようなPDS法の欠点を克服する1つの方法としては、隣接するチャンネルの中心周波数の間隔を小さくすることが考えられる。そこで、上記の3チャンネルの他に、中心周波数が、0.0883883、0.1767767cycles/pixelである2つのチャンネルを用意し、5つのチャンネルで視差検出を行ってみた。この場合、チャンネルの中心周波数間隔は約0.5オクターブになっている。正答率、及び計算時間

を表2に示す。PD法による正答率はほとんど変化していないが、PDS法では、PD法とほぼ同程度まで正答率の改善がなされている。しかも、5チャンネルによるPDS法は3チャンネルのPD法よりも計算時間が短く、総合的に評価した場合、PDS法はPD法と同程度、あるいはそれ以上の性能を持つと考えられる。5チャンネルを用いたときの両手法による視差検出結果を図9、図10に示す。

表2 PD法とPDS法性能比較（5チャンネル）

		視差 1	視差 3
PD法	正答率	97.9%	95.0%
	計算時間	70sec	
PDS法	正答率	97.8%	94.8%
	計算時間	26sec	

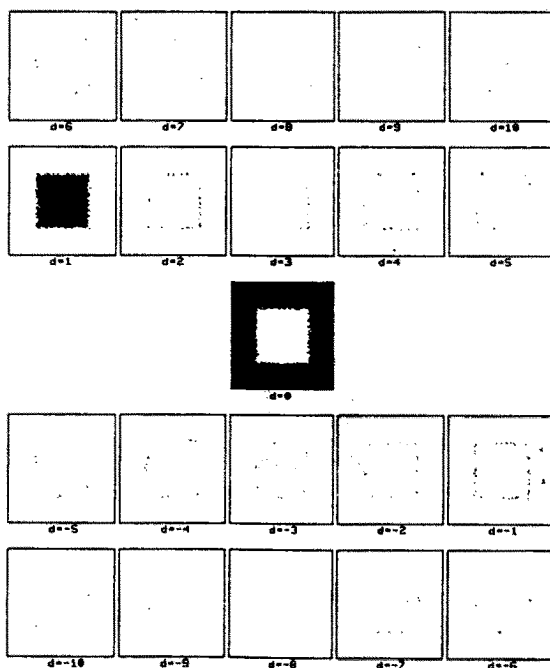


図9(a) PD法による視差検出結果（視差1,5チャンネル）

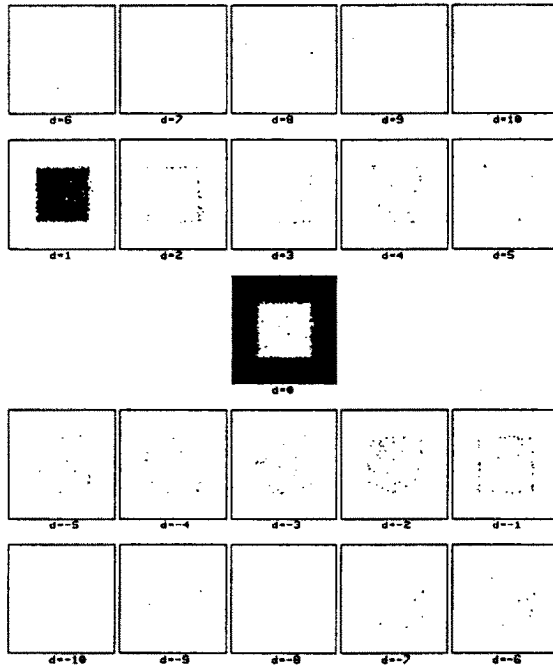


図9(b) PDS法による視差検出結果（視差1,5チャンネル）

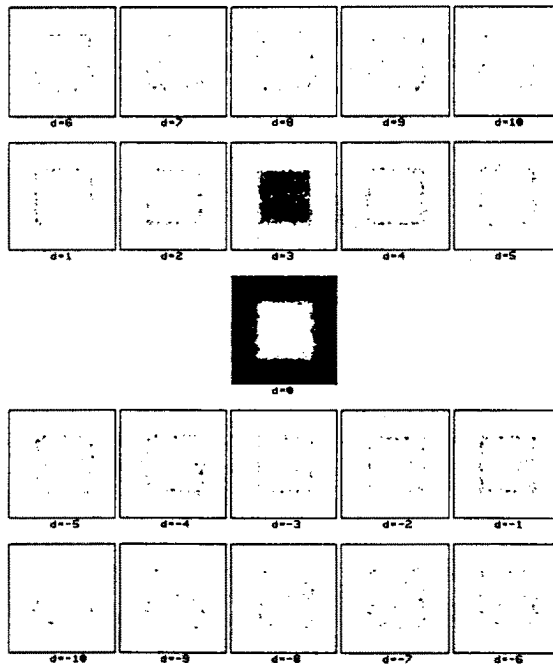


図10(a) PD法による視差検出結果（視差3,5チャンネル）

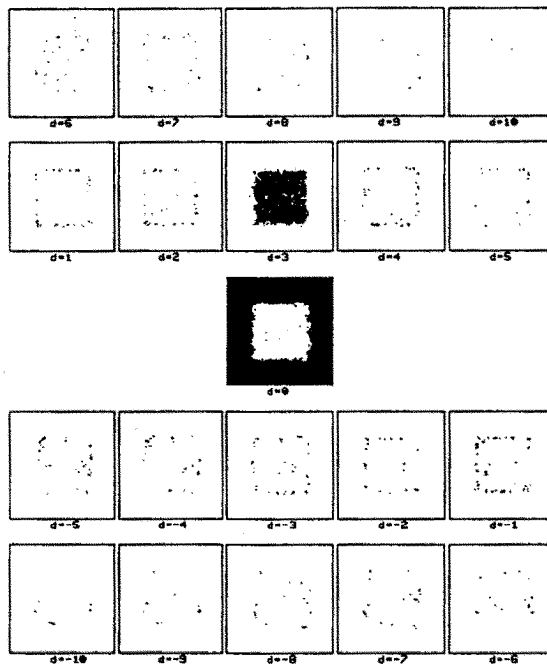


図10(b) PDS 法による視差検出結果 (視差3,5チャンネル)

## 7. むすび

両眼立体視検出法として、左右画像における輝度変化の位相差符号（-、0、+）から視差を求める方法を提案した。本方法では、単一の空間周波数チャンネルは、求めるべき視差が取り得る範囲を限定するだけであるが、これを複数の空間周波数チャンネル間で伝搬することによって最終的に視差を求めることができる。本方法は、我々の従来法と同様、輝度変化の位相情報を用いているために、特徴マッチング法が抱える対応問題を解く必要がない。また、平均輝度やコントラストの変化に影響されにくいという利点をもつ。一方、従来法とは異なり  $\tan^{-1}$  の計算を必要としないため、計算速度が速く、しかも生体モデルとしての可能性がより高いと思われる。また、適当な視差に対しては、位相差の符号は、視差の符号（注視点を中心に奥にあるか手前にあるか）と等しく、それを検出することは、生理学の研究において両眼立体視に関与する細胞として発見されている near cell、far cell の機能を実現しているようにも考えられる。

ランダム・ドット・パターンを用いたシミュレーションでは、位相差から視差を計算する我々の従来法との比較を行い、計算時間と計算精度を考慮した総合的な性能にはあまり差がみられないことが確かめられた。しかし、図7、8、9、及び10の視差検出の結果からわかるように、両手法では誤差の現れ方が異なっている。すなわち、正答率が同じでも、位相差符号のみを用いた本方法の結果には局所的な（孤立点的な）誤差が目だつ。これは、両手法の視差0を検出する能力が、他の視差を検出する能力よりも優れていると仮定することによって次のように説明できる。従来法では、低い周波数チャンネルでも視差を計算できるので、そのチャンネルで正しい視差が求まる可能性がある。しかも、視差一定の領域がある程度広がっている場所では、低周波数チャンネルのほうが積分効果によって精度がよくなる。もし、正しい視差が低周波数チャンネルで求まってしまえば、よ

り高い周波数チャンネルでは視差0を求めればよい。これに対し本方法では、一般的には最も高い周波数チャンネルまでは正しい視差を求めることはできない。従って、このチャンネルは常に0以外の視差を算出することが要求され、低周波数チャンネルの積分効果の恩恵（視差0を求めること）を受けることができない。上記の仮定（0視差検出の優位性）の真偽については今後検討する必要があるが、シミュレーション結果として現れた孤立点ノイズを除去するためには、各周波数チャンネルにおいてなんらかの空間的相互作用を導入する必要がある。

このような視差検出アルゴリズムが、生体によって実際に行われているかどうかはさだかでないが、心理実験などを通してその可能性を調べることは興味深いと思われる。

## 文 献

1. 緒形、佐藤，“位相検出による両眼立体視モデル”，信学技報，MBE89-129，pp. 29-36(1990)。
2. D. Marr，“VISION”，W. H. Freeman and Company(1982)，乾，安藤，（訳）ビジョン：視覚の計算理論と脳内表現，産業図書(1987)。