

TR-H-015

0044

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

Masahiko SHIZAWA
志沢 雅彦

1993. 7. 21

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

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

志沢 雅彦

ATR 人間情報通信研究所

1993年7月21日

あらし

透明視の普遍的計算理論である重ね合わせの原理に基づいて、両眼立体視における透明視（トランスペアレンシー）の計算理論と計算モデルを提案する。トランスペアレンシーとは、一般に、複数の重なり合った半透明な表面の視知覚であって、従来の工学的なステレオビジョンのアルゴリズムは、全くこの様な状況を想定していなかった。本論文では、輝度ベースのステレオ視のエピポーラ線上における整合の拘束条件をトランスペアレンシーに拡張し、画像面の各点に複数の視差が存在する場合に成立する基本拘束方程式を導出する。この方程式は、左右の画像の輝度の差と、それぞれの画像の輝度変化の間の拘束になる。さらに、その一次近似に基づいた2重視差の準線形アルゴリズムと一撃計算モデルを提案する。従来の多重運動の計算理論では、 n 個の運動視差に対して、最低 $n + 1$ 枚の画像が必要であったので、これをステレオ視に適用することができなかった。本論文で提案する計算モデルは、左右2枚からなるステレオ画像対におけるエピポーラ線上の各点において、2個の視差を線形フィルタの出力から直接一撃的に計算できる。したがって、従来の様に、アドホックな特徴抽出過程をステレオ整合に前置する必要がない。提案するアルゴリズムは、工学的には、SSD (Sum of Squared Difference) 法を2重視差に拡張したものと考えることができるが、正則化理論を用いた実現も可能である。

キーワード:

ステレオビジョン、両眼立体視、視差、トランスペアレンシー、透明視、一撃アルゴリズム、重ね合わせの原理

Computational Theory of Binocular Stereo Transparency and a Single-shot Model for Computing Two-fold Disparities

Abstract

A closed-form single-shot stereo disparity estimation algorithm is proposed that can compute multiple disparities due to transparency directly from signal differences and variations on epipolar lines of a binocular image pair. The *transparent stereo constraint equations* have been derived by using a novel mathematical technique, the *principle of superposition*. A computationally tractable single-shot algorithm is derived by using the first-order approximation of the constraint equations with respect to disparities. The algorithm can compute multiple disparities from only *two* images in contrast to the previous algorithms for motion transparency which needed at least $n + 1$ frames for n simultaneous motion estimates. The derived algorithm can be viewed as the SSD (sum of squared differences) for signal matching extended to deal with multiple disparities. However, the constraints are not dedicated solely to the SSD method and several other implementations are possible. These possibilities are also discussed in this paper.

Keywords: Binocular stereo vision, Disparity, Transparency, Single-shot algorithm, Principle of superposition.

もくじ

1	はじめに	4
2	ステレオ・トランスペアレンシーの形式的計算理論	6
2.1	ステレオ整合の基本拘束条件の演算子形式による記述	6
2.2	重ね合わせの原理を用いたステレオ・トランスペアレンシーの基本拘束条件の導出	7
3	2重視差の準線形アルゴリズムと一撃計算モデル	8
3.1	ステレオ・トランスペアレンシーの1次近似拘束方程式と準線形アルゴリズム	8
3.1.1	1次近似拘束方程式	8
3.1.2	準線形アルゴリズム	9
3.2	局所ガウススケール微分を用いた多重スケール局所画像情報表現とステレオ・トランスペアレンシーの計算モデル	10
4	シミュレーションによる検証	11
4.1	実験1：1次元信号を用いたシミュレーション	11
4.2	実験2：2次元信号を用いたシミュレーション	12
5	その他の可能なアルゴリズムについて	12
5.1	周波数領域における拘束方程式	12
5.2	正則化理論の適用	12
6	結論	13
	謝辞	14

1 はじめに

両眼立体視 (ステレオ・ビジョン) のための両眼視差の検出は、視覚の計算理論において相当な注目を集めてきた [3] [7] [12] [16] [18]. ステレオ・ビジョンは、シーンの奥行き、あるいは3次元構造を受動的に獲得する方法として工学的にも盛んに研究されてきている [4] [11] [13]. しかし、これらすべてのステレオ整合アルゴリズムは透明視 (トランスペアレンシー)[8]、すなわち、同一の網膜上の点あるいは、画像投影面上の点に複数の表面を復元することはできない。近年、人間のステレオ視におけるトランスペアレンシーの能力と限界が明らかにされてきた [1] [17] [23]。トランスペアレンシーは、オクルージョンや、物理的に光透過性をもつ物体、カメラの解像度以上の細かいオクルージョンなどが存在する様な、複雑な環境を扱わなければならないコンピュータ・ビジョンの重要な課題でもある。これらの複雑な環境は、フェンス、樹木、藪、水面、そして、窓ガラスなど、自然環境においても、人工環境においても共通に存在する。

近年、人間のステレオ視において、パターンの繰り返し重畳によって作られる解釈が曖昧なステレオグラムに関するトランスペアレンシーを計算論的に説明することが Weinshall [23] と Pollard et al.[17] によって論じられてきた。そこでは、ステレオ整合の一意性拘束 (uniqueness constraint) がトランスペアレンシーの場合に用いられているか否かが議論されている。しかし、いずれにしろ、まだ、十分に計算理論として深められているとはいえ、現象論的な説明にとどまっている。

Prazdny は、物体表面の不連続は、不連続の近傍に2枚の表面があるとみなすことができることを指摘した [18]。彼は、また、Burtらによって発見された人間の視差勾配限界 (disparity gradient limit)[3] の拘束に基づいた特徴点ベースのステレオ整合アルゴリズムを提案した。Prazdny のアルゴリズムは、Marr & Poggio のアルゴリズム [12] の様に近傍の視差に対する抑制的な相互作用を持たないので、複数の表面を扱うことができる。Pollard らも、これに類似したステレオ整合アルゴリズム (PMF アルゴリズム) を提案しており [16]、人間の曖昧な多重表面の知覚現象をも説明できると主張している [17]。

しかし、これらは、すべてランダムドットステレオグラムを想定した特徴点ベースのステレオ整合であり、信号レベルでトランスペアレンシーを扱っていないため、純粋に信号の重畳によって生じるトランスペアレンシーの場合 (例えば、窓ガラスの透過像と反射像) には適さない。従来の特徴検出演算子は暗黙のうちに表面の非透明性を仮定しているからである。すなわち、この場合には、各表面のパターンが画像生成過程で重畳されたあとで特徴抽出が行われる形になる。このため、従来のエッジやかどの特徴抽出演算子では、複数の表面の特徴間で干渉が起こり、誤った特徴が混入することになる。そこで、この場合には、画像生成過程の分析結果に基づいた輝度ベースの整合がより適切である。実際、人間は、ランダムな輝度パターンの重畳によって作成された図1のステレオグラム (以下、ランダム輝度ステレオグラムと呼ぶ) から多重表面を

知覚することができる。

従来、視覚の計算理論で取り上げられてきたステレオ視のトランスペアレンシーは、すべて密度の低いランダムドット（従って、ドット同士の重なりが少なく、空間的に特徴が分離されていて、特徴検出器の空間解像度の方がドット密度よりも十分高い）によって作られたステレオグラムを対象としており、ここで問題になる各表面パターン間の干渉作用は無視できた [18] [17]。一方、輝度ベースの整合 [11] [13] では、局所的に複数の視差を扱うことは、従来考えられてこなかった。

図 2 は、もっとも簡単な平行視ステレオの場合の多重表面によるトランスペアレンシーを説明したものである。手前の表面は、半透明なパターンを持つ面で、奥の表面は、透明でないパターンを持つ面である（一般に、人間は、もっとも奥の表面だけ不透明に感じ、その手前にある表面は透明に感じる）。左右の画像の各点に各表面に対応する 2 つの視差を検出することによって複数の表面の 3 次元奥行き情報を復元することができる。

現状のコンピュータ・ビジョンでは、このようなトランスペアレンシーを含む複雑な環境を正しく扱う決定的な理論や道具はまだ確立されていない。実際、トランスペアレンシーは従来から難問とされているセグメンテーションをその特殊な場合として含み、したがってそれよりもさらに困難な問題である。本論文では、セグメンテーションを経ずに、両眼の画像対から複数の視差を直接一撃的に復元するアルゴリズムを提案する。

そのための計算理論として、画像の輝度が 2 枚の表面のパターンの加法的重ね合わせによって生成され、画像面の各点に 2 個の視差が存在するという仮定のもと¹、ステレオ画像対のエピポーラ線上の輝度変化に関する直接的拘束を表す線形基本拘束方程式を、透明視の普遍的計算理論として著者が提案した重ね合わせの原理 [20] を用いて導出する。

提案するアルゴリズムは、多重オプティカルフロー推定アルゴリズム [20] の特殊な場合ではない。多重オプティカルフローの場合には、 n -重のフロー推定のためには n -階の時空間偏微分が必要であり、 n -階の時空間偏微分には、最小 $n+1$ 枚の画像フレームが必要であったからである。ここで提案するアルゴリズムは、たった 2 枚の画像からエピポーラ線上の各点において複数の視差を計算できる。このアルゴリズムは、2 枚の画像フレームからの多重オプティカルフロー検出に比較的容易に拡張することができる。

以下では、まず、第 2 章において、ステレオ視におけるトランスペアレンシーについて説明した後、重ね合わせの原理にもとづいて、演算子形式によるステレオ・トランスペアレンシーの形式的理論を展開して基本拘束方程式を導く。次に、第 3 章で、基本拘束方程式の 1 次近似を導出する。この拘束方程式は、ステレオ画像対の左右の画像

¹一般化重畳 [14] に拡張することによって、乗法的重ね合わせの場合を含んだかなり広い範囲の重畳に対応できる。また、厳密には重畳の条件を満たさなくても、特定の周波数サブバンドでは、近似的に重畳の条件が成立する場合が多い。

の輝度微分に関して線形であり、その係数は、一般に、複数視差の基本対称式についての線形パラメータの形になる。この拘束方程式の局所最適化問題を変分法で解くことによって、2個の視差を計算するアルゴリズムを導出し、それに基づいた一撃計算モデルを提案する。第4章では、1次元信号およびランダムドット、ランダムテクスチャのステレオ画像対を用いたアルゴリズムの検証を行う。第5章では、拘束方程式の周波数空間における表現と正則化による実現について述べる。第6章で本研究をまとめる。

2 ステレオ・トランスペアレンシーの形式的計算理論

本章では、まず、輝度ベースのステレオ整合拘束条件を線形演算子に基づいた形式を用いて書き直す。そして、重ね合わせの原理と合成演算子の対称化操作を用いて、ステレオ・トランスペアレンシーの基本拘束条件を導出する。

なお、本論文では、2重のトランスペアレンシーの場合のみを述べるが、3重以上の場合に理論を拡張することも可能である。

2.1 ステレオ整合の基本拘束条件の演算子形式による記述

不透明な表面に対する、輝度ベースのステレオ整合の拘束条件は次式で表される。

$$L(x, y) - R(x - D, y) = 0, \quad R(x, y) - L(x + D, y) = 0, \quad (1)$$

ここで、 $L(x, y)$ と $R(x, y)$ は、それぞれ、左右の画像の輝度分布関数を表し、 D は x 方向の視差を表す。 x はエピポーラ線にそった座標で、 y はそれに垂直な方向の座標である。式 (1) は、次の様に、振幅演算子 $a(D)$ とデータ分布関数 $f(x, y)$ を用いて書き直すことができる。

$$a(D)f(x, y) = 0 \quad (2)$$

ここで、

$$a(D) \stackrel{\text{def}}{=} \begin{bmatrix} 1 & -D(D) \\ -D(-D) & 1 \end{bmatrix}, \quad f(x, y) \stackrel{\text{def}}{=} \begin{bmatrix} L(x, y) \\ R(x, y) \end{bmatrix}. \quad (3)$$

$D(D)$ は、 x 軸方向への D だけのずらし演算子であり、例えば $f(x)$ を $f(x - D)$ に変換する。(1) の2個の方程式は互いに同値であるが、演算子の中で、左右を同等に扱うために、両方を用いて定式化した。図3に、ステレオ画像対に対する振幅演算子の働きを図示した。

$f(x - D)$ を D に関して x のまわりでテイラー展開すると次のようになる。

$$f(x - D) = f(x) - D\partial_x f(x) + \frac{D^2}{2!}\partial_x^2 f(x) - \dots = \sum_{i=0}^{\infty} \frac{(-D)^i}{i!} \partial_x^i f(x), \quad (4)$$

ここで、 $\partial_x^i f(x)$ は x に関する $f(x)$ の i 階の偏導関数を表す。これから、ずらし演算子は微分演算子

$$\mathcal{D}(D) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} \frac{(-D)^i}{i!} \partial_x^i = \exp(-D\partial_x) = 1 - D\partial_x + \frac{D^2}{2!} \partial_x^2 - \frac{D^3}{3!} \partial_x^3 \dots \quad (5)$$

として書かれることがわかる。したがって、式 (4) は $f(x-D) = \mathcal{D}(D)f(x)$ と書くことができる。このずらし演算子は線形演算子であることに注意する。すなわち、 $\mathcal{D}(D)(f_1(x) + f_2(x)) = \mathcal{D}(D)f_1(x) + \mathcal{D}(D)f_2(x)$ および $\mathcal{D}(D)0 = 0$ が成立する。さらに、ずらし演算子はその合成積に関して閉じており、

$$\mathcal{D}(D_\alpha)\mathcal{D}(D_\beta) \equiv \mathcal{D}(D_\beta)\mathcal{D}(D_\alpha) \equiv \mathcal{D}(D_\alpha + D_\beta) \quad (6)$$

が成り立つ。

拘束方程式 (2) は、画像の輝度分布関数 $L(x, y)$ と $R(x, y)$ の高階の導関数に関する拘束に拡張することができる。 $L(x, y)$ と $R(x, y)$ の (p, q) 階偏導関数を $L^{(p,q)}(x, y) = \partial_x^p \partial_y^q L(x, y)$ および $R^{(p,q)}(x, y) = \partial_x^p \partial_y^q R(x, y)$ とし、これらを合わせて $f^{(p,q)}(x, y) = [L^{(p,q)}(x, y), R^{(p,q)}(x, y)]^T$ と書くことにする。すると、拘束方程式 (2) は $p, q = 0, 1, 2, \dots$ に対して、

$$a(D)f^{(p,q)}(x, y) = 0, \quad (7)$$

と拡張できる。一般に、ある画像点におけるすべての (p, q) 階偏導関数の集合はその点の周りの画像の局所構造を完全に特徴付ける [9] ので、この拡張によって、特徴ベースの整合と輝度ベースの整合が統一される。

ずらし演算子の線形性を用いると振幅演算子 $a(D)$ もまた線形演算子であることがわかる。

2.2 重ね合わせの原理を用いたステレオ・トランスペアレンシーの基本拘束条件の導出

前節のステレオ整合の定式化に基づいて、加法的 2 重トランスペアレンシーの場合に多重オプティカルフローの計算理論において提案された重ね合わせの原理 [20] を形式的に適用すると、次の拘束方程式が得られる。

$$a(D_2)a(D_1)f^{(p,q)}(x, y) = 0, \quad (8)$$

ここで、 $f^{(p,q)}(x, y) = f_1^{(p,q)}(x, y) + f_2^{(p,q)}(x, y)$ であり、各 $f_i^{(p,q)}(x, y)$ は $a(D_i)f_i^{(p,q)}(x, y) = 0$ ($i = 1, 2$) と拘束されている。

残念ながら、この場合、振幅演算子の合成積が可換でない。実際、

$$a(D_2)a(D_1) - a(D_1)a(D_2) = \begin{bmatrix} -K & 0 \\ 0 & K \end{bmatrix} \quad (9)$$

となる。ここで、

$$K \stackrel{\text{def}}{=} \mathcal{D}(D_1 - D_2) - \mathcal{D}(D_2 - D_1) \\ = 2 \left\{ -(D_1 - D_2)\partial_x - \frac{(D_1 - D_2)^3}{3!}\partial_x^3 - \frac{(D_1 - D_2)^5}{5!}\partial_x^5 - \dots \right\} \quad (10)$$

であるから、演算子 $a(D_1)$ と $a(D_2)$ は、 $D_1 \neq D_2$ である限り、可換でない。したがって、振幅演算子の積の順序を変えると拘束方程式の意味が変化してしまい、かつ、基本拘束方程式 (8) の左辺は一般の信号に対してけっして零にはなりえなくなる。前者の問題は、このトランスペアレンシーの振幅演算子の積に対して、視差 D_1 、と D_2 に関する対称化操作を施すことによって解決される。後者の問題についても、式 (8) の左辺の 2 乗エネルギーをとり、その最小化問題として捉えることによって回避することができる。

対称化された振幅演算子は次式となる。

$$\bar{a}^{(2)}(D_{(1)}, D_2) \stackrel{\text{def}}{=} \frac{1}{2} \{a(D_1)a(D_2) + a(D_2)a(D_1)\} \quad (11)$$

ここで、添字に付けられた括弧は、それに囲まれた添字に関して対称化の操作がなされていることを表す。

3 2 重視差の準線形アルゴリズムと一撃計算モデル

前章で導出したステレオ・トランスペアレンシーの基本拘束方程式は、視差 D_1, D_2 に関してきわめて非線形性が高いので、そのままでは計算機による実現に適さない。本章では、振幅演算子の一次近似をおこなうことによって導かれた両眼ステレオ視の準線形一撃アルゴリズムを述べる。

3.1 ステレオ・トランスペアレンシーの 1 次近似拘束方程式と準線形アルゴリズム

本節では、2 重視差の場合を述べるが、単一視差の場合の一次近似と従来手法との関連については付録 1 を参照せよ。

3.1.1 1 次近似拘束方程式

ずらし演算子 $\mathcal{D}(D)$ に対する一次近似

$$\mathcal{D}(D) \approx 1 - D\partial_x. \quad (12)$$

を用いると 2 重視差の場合の対称化された振幅演算子は次の形に近似される (導出は付録 2 を参照せよ)。

$$\bar{a}^{(2)}(D_{(1)}, D_2) \approx \bar{\alpha}^{(2)}(D_{(1)}, D_2) = \begin{bmatrix} 2 - D_1 D_2 \partial_x^2 & -2 + (D_1 + D_2)\partial_x \\ -2 - (D_1 + D_2)\partial_x & 2 - D_1 D_2 \partial_x^2 \end{bmatrix} \quad (13)$$

したがって基本拘束方程式の残差関数は次式となる。(前章で述べたように、0 にならないので残差関数と呼ぶ。)

$$\begin{aligned} r_L^{(2)}[p, q](x, y) &\stackrel{\text{def}}{=} 2 \{L^{(p, q)} - R^{(p, q)}\} + (D_1 + D_2)R^{(p+1, q)} - D_1 D_2 L^{(p+2, q)} \\ r_R^{(2)}[p, q](x, y) &\stackrel{\text{def}}{=} 2 \{R^{(p, q)} - L^{(p, q)}\} - (D_1 + D_2)L^{(p+1, q)} - D_1 D_2 R^{(p+2, q)} \end{aligned} \quad (14)$$

ただし、 $L^{(p, q)} = L^{(p, q)}(x, y)$, $R^{(p, q)} = R^{(p, q)}(x, y)$ という省略を用いた。これらの式を観察してみると、最初の式は、左右の輝度 (の (p, q) 階微分) の差 $L^{(p, q)}(x, y) - R^{(p, q)}(x, y)$ 、右画像 (の (p, q) 階微分) の x 方向 1 階微分 $R^{(p+1, q)}(x, y)$ および、左画像 (の (p, q) 階微分) の x 方向 2 階微分 $L^{(p+2, q)}(x, y)$ の間の近似的な線形従属関係を表している。また、2 番目の式は最初の式の左右の画像の役割を入れ換えたものになっている。

3.1.2 準線形アルゴリズム

点 (x_0, y_0) における視差 D_1 および D_2 を推定するために、次のエネルギー関数の局所最小化を行う。

$$\begin{aligned} E^{(2)}(D_1(x_0, y_0), D_2(x_0, y_0)) &= \iint \{w_{p, q}(x - x_0, y - y_0)\}^2 \\ &\quad \times \left[\{r_L[p, q](x, y)\}^2 + \{r_R[p, q](x, y)\}^2 \right] dx dy \end{aligned} \quad (15)$$

ここで、 $w_{p, q}(x - x_0, y - y_0)$ は積分領域と画像微分量の統計的信頼度に応じて拘束式の寄与度を定める重み関数である。一般に高次の画像微分量は誤差が大きいため、それに応じて重みは小さくする必要がある。本論文のシミュレーションでは、微分の階数 $(p + q)$ は固定して考えることにするので、以後、重み関数の決め方には言及しない。エネルギー関数 (15) は、2 つの視差に関する基本対称式 $s_1 = \frac{1}{2}(D_1 + D_2)$ および $s_2 = D_1 D_2$ に関して 2 次形式になっている。変数 D_1 および D_2 に関するエネルギー最小化を行う代わりに、 s_1, s_2 に関する最小化を行う。エネルギー関数の最小化条件 $\frac{\partial E^{(2)}}{\partial s_1} = 0$ および $\frac{\partial E^{(2)}}{\partial s_2} = 0$ から、次の連立線形方程式 (正規方程式) を得る。

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (16)$$

ここで、

$$\begin{aligned} a_{11} &= 4 \sum_{p, q} \iint (w_{p, q})^2 \left\{ (R^{(p+1, q)})^2 + (L^{(p+1, q)})^2 \right\} dx dy, \\ a_{12} &= 2 \sum_{p, q} \iint (w_{p, q})^2 (L^{(p+1, q)} R^{(p+2, q)} - L^{(p+2, q)} R^{(p+1, q)}) dx dy, \end{aligned}$$

$$\begin{aligned}
a_{22} &= \sum_{p,q} \iint (w_{p,q})^2 \left\{ (R^{(p+2,q)})^2 + (L^{(p+2,q)})^2 \right\} dx dy, \\
b_1 &= 4 \sum_{p,q} \iint (w_{p,q})^2 (R^{(p+1,q)} + L^{(p+1,q)}) (R^{(p,q)} - L^{(p,q)}) dx dy, \\
b_2 &= 2 \sum_{p,q} \iint (w_{p,q})^2 (L^{(p+2,q)} - R^{(p+2,q)}) (L^{(p,q)} - R^{(p,q)}) dx dy \quad (17)
\end{aligned}$$

で定義される。ただし、 $w_{p,q} = w_{p,q}(x - x_0, y - y_0)$ と省略した。二つの視差 D_1 と D_2 ($D_1 \geq D_2$) は、次の d に関する 2 次方程式の 2 つの実根として得られる。

$$d^2 - 2s_1 d + s_2 = 0. \quad (18)$$

これらの根は、

$$D_1 = s_1 + \sqrt{s_1^2 - s_2}, \quad D_2 = s_1 - \sqrt{s_1^2 - s_2} \quad (19)$$

と書くことができる。ここで、根は、実数でなければならないので、

$$\Delta^{(2)} \stackrel{\text{def}}{=} s_1^2 - s_2 \geq 0, \quad (20)$$

が必要である。 $\Delta^{(2)}$ は、代数方程式 (18) の判別式を表す。単一視差の場合は次の条件が成立する。

$$\Delta^{(2)} \stackrel{\text{def}}{=} s_1^2 - s_2 = 0, \quad (21)$$

単一視差の場合には、視差は、 $D = s_1$ として求められる。視差の数を決定することは多重視差検出アルゴリズムにとって重要な問題である。ステレオ・トランスペアレンシーの場合、条件 (21) を検査することによって単一視差の場合と 2 重視差の場合を判別できる。実際には、絶対値があるしきい値よりも小さいことを検査することによって判定する。この他に、エネルギー関数 (15) の残差や、正規方程式の行列の階数を検査することによって、それ以外の視差計算不能の場合を検出することができる。さらに、誤差モデルを仮定して統計的検定法を適用することも考えられる。本論文では、基本拘束式 (14) を検証することを主眼とするので、これらの問題をこれ以上深く論ずることはしない。

3.2 局所ガウススケール微分を用いた多重スケール局所画像情報表現とステレオ・トランスペアレンシーの計算モデル

前節では、画像の輝度の微分を計算できることを前提としてアルゴリズムを導出したが、微分演算は不良設定問題であるから、なんらかの正則化手法を用いないとロバストなアルゴリズムとはならない。ここでは、ステレオ整合の局所画像表現として、局所ガウススケール微分 [9] を用いる。この方法は、線形基底フィルタ集合の出力をステレオ整合の基本要素とした Jones & Malik[6] のアプローチと近いものである。しかし、画像表現それ自身は、ステレオ・トランスペアレンシーの計算理論とは独立である。本論文で、局所ガウススケール微分を表現として用いるのは、前章で導いたアルゴリズム

に必要な情報である輝度微分を直接得られるからである。もちろん、生物の視覚系における画像の初期表現モデルとしてしばしば用いられるガボール基本関数系のような他の画像表現を用いることも可能である。

(p, q) 階の局所ガウススケール微分は、次の様に、ガウス関数の (p, q) 階微分と画像の畳み込み積を計算することによって得られる。

$$\begin{aligned} L_{\sigma}^{(p,q)}(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{\sigma}(x-u, y-v) L^{(p,q)}(u, v) dudv \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{\sigma}^{(p,q)}(x-u, y-v) L(u, v) dudv, \\ R_{\sigma}^{(p,q)}(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{\sigma}(x-u, y-v) R^{(p,q)}(u, v) dudv \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{\sigma}^{(p,q)}(x-u, y-v) R(u, v) dudv \end{aligned} \quad (22)$$

ここで、2次元ガウス関数 $G_{\sigma}(x, y)$ は、

$$G_{\sigma}(x, y) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \quad (23)$$

と定義される。

図4は導出されたアルゴリズムをネットワーク計算モデルとして図示したものである。ここでは、微分階数 p, q および、スケール σ に関する統合をも取り入れている。 $w_{p,q,\sigma}$ は、この統合の際の重み係数を表す。

4 シミュレーションによる検証

本章では、まず、1次元信号によるシミュレーションの結果を示す。さらに、ランダムドット、および、ランダムテクスチャーによるステレオ画像対から2次元の2重表面を復元した例を示す。

4.1 実験1：1次元信号を用いたシミュレーション

乱数とガウス関数による畳み込みによって生成した信号を用いて、図5に示す2つの視差マップを持つ1次元ステレオ対を作り、それらを左右独立に輝度の和をとって適当にサブサンプルした信号を生成した（図6に生成した信号を示す。）。

実験には、 x 方向への5階までの局所ガウススケール微分を用いた。また、シミュレーションでは、ガウススケールパラメータ σ は4[pixel]に固定した。

エネルギー積分の重み関数 $w_{p,q}(x)$ は、幅 $M = 121$ [pixel] の矩形窓とした。また、計算は、 $p = 0, 1, 2, 3$ のそれぞれの場合に対して独立に行った。図7は、アルゴリズムの最終出力をそれらの真値とともに重畳してプロットしたものである。

ここで、強調しておきたいことは、これらの結果は一次近似のもとで計算されたものであり、かつ、視差のなめらかさ拘束を利用していないということである。（窓の中で視差が一定であるという仮定はしているが。）

4.2 実験2：2次元信号を用いたシミュレーション

2次元信号を用いたシミュレーションは、超並列計算機 CM-2 上でおこなった。図8は、図1と同様のランダム輝度ステレオグラムに対する結果である。ただし、視差のパラメータは、 $D_1 = -0.5[\text{pixel}]$ 、および、 $D_2 = 0.5[\text{pixel}]$ とした。フィルタは、 $p + q = 2$ 、 $\sigma = 1.6[\text{pixel}]$ のものだけを用いた。単一視差の判定は、判別条件 (21) を適当なしきい値(この実験では、0.11 という値を用いた。)でテストすることによって実現した。なお、局所エネルギー積分領域は、 $25 \times 25[\text{pixel}]$ の矩形領域とした。図8a,b は、計算された視差を高さで表した視差マップである。枠の領域では、正しく単一視差の判定が行われている。中心領域では、ノイズ成分は多くなるが、2重視差が計算されている。図8c,d は、視差の平均化操作($20 \times 20[\text{pixel}]$ の窓による移動平均)を施したものである。視差の変化に滑らかさの拘束条件を課したものに相当する。

5 その他の可能なアルゴリズムについて

5.1 周波数領域における拘束方程式

ステレオ・トランスペアレンシーの1次近似残差関数 (14) はフーリエ変換により、周波数領域における表現に変換することができる ($p = q = 0$ の場合を示す)。

$$\begin{aligned}\tilde{r}_L^{(2)}(\omega_x, \omega_y) &= 2 \left(\tilde{L} - \tilde{R} \right) + (D_1 + D_2)(2\pi i)\omega_x \tilde{R} - D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{L} \\ \tilde{r}_R^{(2)}(\omega_x, \omega_y) &= 2 \left(\tilde{R} - \tilde{L} \right) - (D_1 + D_2)(2\pi i)\omega_x \tilde{L} - D_1 D_2 (2\pi i)^2 \omega_x^2 \tilde{R}\end{aligned}\quad (24)$$

ただし、 $\tilde{L} = \tilde{L}(\omega_x, \omega_y)$ 、 $\tilde{R} = \tilde{R}(\omega_x, \omega_y)$ と省略した。これらの残差関数は例えばガボール関数を用いたアルゴリズムを導くのに有用である。

5.2 正則化理論の適用

第3章で述べたアルゴリズムは2乗誤差の和の最小化に基づいた SSD(Sum of Squared Difference) 法によるものであった。したがって、局所窓における視差は定数であると仮定されている。第2章で提案した多重視差を求めるために、この不良設定問題に対して、正則化理論 [15] を適用することができる。この場合、視差パラメータ D_i になめらかさの拘束を課すと、基本拘束方程式に一次近似をおこなってもエネルギー関数が2次でなくなり、最適化が難しくなる。しかし、 D_i の代わりに s_i になめらかさの拘束を課せば、少なくとも、拘束方程式を1次近似した場合には、エネルギー関数が2次となり、線形逆問題に帰着される。

この様なパラメータ付けを行った場合、 n 重視差モデルに対する正則化のエネルギー関数は次式の様になる。

$$\mathcal{E}^{(n)} \left(s^{(n)}(x, y) \right) =$$

$$\iint \left[\left\{ \sum_{p,q,\sigma} w_{p,q,\sigma}^2 \left\| G_\sigma \otimes \bar{a}^{(n)} \left(s^{(n)}(x,y) \right) f^{(p,q)}(x,y) \right\|^2 \right\} + \left\{ S s^{(n)} \right\}^T \left\{ S s^{(n)} \right\} \right] dx dy \quad (25)$$

ここで、 $s^{(n)} = (s_1^{(n)}, s_2^{(n)}, \dots, s_n^{(n)})$ は、視差 $D_i (i = 1, \dots, n)$ の基本対称式からなるベクトルである。 \otimes は畳み込み演算を示す。 $n = 2$ の場合には、 $s_1^{(2)} = \frac{1}{2} (D_1 + D_2)$ 、 $s_2^{(2)} = D_1 D_2$ である。ただし、このような置き換えをおこなった場合には、正則化演算子 S をどの様にして決めたらよいかということが課題として残される。

6 結論

両眼立体視におけるトランスペアレンシーの計算理論およびアルゴリズムと多重視差の一撃計算モデルを提案した。この計算理論は、演算子形式で表現された透明視の普遍的計算理論である重ね合わせの原理に基づいて導出された。導出された基本拘束方程式を各視差に対応する振幅演算子について1次近似し、局所エネルギー最小化問題を解くことにより準線形一撃アルゴリズムを導いた。計算機で生成された1次元信号を用いたシミュレーションによってこの一撃アルゴリズムがよい初期値を計算できることを示した。

ここで示したアルゴリズムは、ステレオアルゴリズムの一要素である。より精密な視差の推定が欲しければ、繰り返し画像位置合わせ法 [11] の様な繰り返し修正アルゴリズムに発展させることもできるであろう。大きな視差に対しては、局所ガウススケール微分のスケールパラメータの制御が重要になる。窓関数に関しては、統計モデルに基づいた適応的な窓の大きさの選択法を応用することもできよう [13]。正則化理論による実現も有力な手法である。奥行きの不連続におけるオクルージョンの解析も本アルゴリズムと統合される必要がある [10]。

また、本稿の計算理論とアルゴリズムは、2枚の画像フレームからの多重オプティカルフロー推定に発展させることもできる。これは、視差を x 方向のみならず y 方向にも定義すればよい。この拡張によって、2重運動に関する3フレームアルゴリズム [2] の欠点である加速度に対する弱さを克服することができよう。

本論文では、2重トランスペアレンシーの基本拘束方程式とその検証のためのアルゴリズムの導出に重きをおいて、推定誤差の理論的・実験的考察や、このアルゴリズムに適した画像表現などは検討しなかった。今後は、近年開発され洗練されてきたコンピュータビジョンにおける計算結果の信頼度を表す共分散の計算法などの誤差や偏差の一般的解析方法や、ステレオ整合に適した画像表現のための基底核関数の設計理論、非線形の重ね合わせや、表面の滑らかさ拘束の導入などについて検討し、実画像ステレオグラムへの適用性を検証することが必要である。

謝辞

ご討論頂いた ATR 人間情報通信研究所 安藤広志博士、早川秀樹氏、NTT 基礎研究所西田真也博士に感謝致します。本研究は、ATR 視聴覚機構研究所および ATR 人間情報通信研究所において行われた。本研究の機会を与えていただいた同認知機構研究室下原勝憲室長（現在、ATR 人間情報通信研究所第 6 研究室室長）、ATR 人間情報通信研究所 東倉洋一社長、同 赤松茂第 2 研究室室長に感謝致します。

参考文献

- [1] Akerstrom, R.A. and Todd, J.T. 1988. The perception of stereoscopic transparency. *Percept. & Psychophys.* 44(5): 421-432.
- [2] Bergen, J.R., Burt, P., Hingorani, R., and Peleg, S. 1992. A three-frame algorithm for estimating two-component image motion. *IEEE Trans. PAMI*, 14,9 pp.886-896.
- [3] Burt, P.J. and Julesz, B. 1980. A disparity gradient limit for binocular fusion. *Science*, 208: 615-617.
- [4] Dhond, U.R. and Aggarwal, J.K. 1989. Structure from stereo — a review. *IEEE Trans. Sys. Man, & Cybern.* 19(6): 1489-1510.
- [5] Hubel, D.H. and Wiesel, T.N. 1962. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol.* 160: 106-154.
- [6] Jones, D.G. and Malik, J. 1992. Computational framework for determining stereo correspondence from a set of linear spatial filters. *Image and Vision Comput.* 10(10): 699-708.
- [7] Julesz, B. 1960. Binocular depth perception of computer generated patterns. *Bell Sys. Tech. J.* 38: 1001-1020.
- [8] Kersten, D. 1991. Transparency and the cooperative computation of scene attributes. In *Computational Models of Visual Processing* (Landy M.S. and Movshon J.A. eds.), MIT Press, Cambridge, MA.
- [9] Koenderink, J.J. and van Doorn, A.J. 1990. Receptive field families. *Biol. Cybern.* 63: 291-298.
- [10] Little, J.J., Gillett, W.E. 1990. Direct evidence for occlusion in stereo and motion. *Image and Vision Comput.* 8(4): 328-340.
- [11] Lucas, B.D. and Kanade, T. 1981. An iterative image registration technique with an application to stereo vision. *Proc. Image Understanding Workshop*, pp.121-130.
- [12] Marr, D. and Poggio, T. 1979. A theory of human stereopsis. *Proc. R. Soc. London Ser.B* 204: 301-328.

- [13] Okutomi, M. and Kanade, T. 1992. A locally adaptive window for signal matching. *Int. J. Comput. Vision* 7(2): 143-162.
- [14] Oppenheim, A.V. 1967. Generalized Superposition. *Inf. & Control*, 11, pp.528-536.
- [15] Poggio, T., Torre, V., and Koch, C. 1985. Computational vision and regularization theory. *Nature*, 317: 314-319.
- [16] Pollard, S.B., Mayhew, J.E.W., and Frisby, J.P. 1985. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Percept.* 14: 449-470.
- [17] Pollard, S.B. and Frisby, J.P. 1990. Transparency and the uniqueness constraint in human and computer stereo vision. *Nature*. 347: 553-556.
- [18] Prazdny, K. 1985. Detection of binocular disparities. *Biol. Cybern.* 52: 93-99.
- [19] Shizawa, M. 1992. On visual ambiguities due to transparency in motion and stereo. In Sandini G.(ed.):*Computer Vision — ECCV92, 2nd European Conference on Computer Vision*, Lecture Notes in Computer Science 588, Springer-Verlag, pp.411-419.
- [20] 志沢雅彦、間瀬健二 1993. 多重オブティカルフロー：基本拘束方程式と運動透明視・運動境界検出の統一計算理論. 信学論 (D-II), J76-D-II, 5, pp.987-1005 (1993-05).
- [21] Shizawa, M. 1993. A Quasi-linear, single-shot computational model for stereoscopic pure transparency. *Investigative Ophthalmology and Visual Science Supplement (ARVO 1993)*, 34(4):1186.
- [22] Shizawa, M. 1993. Direct estimation of multiple disparities for transparent multiple surfaces in binocular stereo. *Proc. IEEE 4th International Conference on Computer Vision (ICCV 1993)*, Berlin, Germany, IEEE Computer Society Press, pp.447-454.
- [23] Weinshall, D. 1991. Seeing “ghost” planes in stereo vision. *Vision Res.* 31(10): 1731-1748.

付録1 単一視差の場合の拘束と従来法との関連

近似 (12) を用いると単一視差 D の振幅演算子 $\mathbf{a}(D)$ の1次近似は次式に示す演算子 $\alpha(D)$ となる。

$$\mathbf{a}(D) \approx \alpha(D) = \begin{bmatrix} 1 & -1 + D\partial_x \\ -1 - D\partial_x & 1 \end{bmatrix}. \quad (26)$$

したがって、一般化基本拘束方程式 $\mathbf{a}(D)f^{(p,q)}(x,y) = 0$ は、方程式 $\alpha(D)f^{(p,q)}(x,y) = 0$ で近似される。成分ごとに書くと次式になる。

$$\begin{aligned} r_L^{(1)}[p,q](x,y) &\stackrel{\text{def}}{=} \{L^{(p,q)}(x,y) - R^{(p,q)}(x,y)\} + DR^{(p+1,q)}(x,y) = 0 \\ r_R^{(1)}[p,q](x,y) &\stackrel{\text{def}}{=} \{R^{(p,q)}(x,y) - L^{(p,q)}(x,y)\} - DL^{(p+1,q)}(x,y) = 0 \end{aligned} \quad (27)$$

$p = q = 0$ の場合は繰り返し修正によるステレオ整合 [11][13] における視差の修正量の計算のために用いられた方程式に等価である。(ただし、これらの論文では、式 (27) のうち、片方のみを用いている。しかし、本論文では、二つの拘束方程式を同時に用いて、左右の対称性を持たせている。)

付録2 二重視差の振幅演算子の1次近似の導出

視差 D_1 および D_2 に対する振幅演算子 $\mathbf{a}(D_1)$ および $\mathbf{a}(D_2)$ の一次近似振幅演算子 $\alpha(D_1)$ および $\alpha(D_2)$ の合成演算子を計算すると、

$$\begin{aligned} \alpha(D_2)\alpha(D_1) &= \begin{bmatrix} 1 & -1 + D_2\partial_x \\ -1 - D_2\partial_x & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 + D_1\partial_x \\ -1 - D_1\partial_x & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 + (D_1 - D_2)\partial_x - D_1D_2\partial_x^2 & -2 + (D_1 + D_2)\partial_x \\ -2 - (D_1 + D_2)\partial_x & 2 - (D_1 - D_2)\partial_x - D_1D_2\partial_x^2 \end{bmatrix}. \end{aligned} \quad (28)$$

ここで、 D_1 と D_2 を交換すれば、

$$\alpha(D_1)\alpha(D_2) = \begin{bmatrix} 2 + (D_2 - D_1)\partial_x - D_1D_2\partial_x^2 & -2 + (D_1 + D_2)\partial_x \\ -2 - (D_1 + D_2)\partial_x & 2 - (D_2 - D_1)\partial_x - D_1D_2\partial_x^2 \end{bmatrix}. \quad (29)$$

従って、これら2個の行列を加算して2で割って、式 (13) を得る。

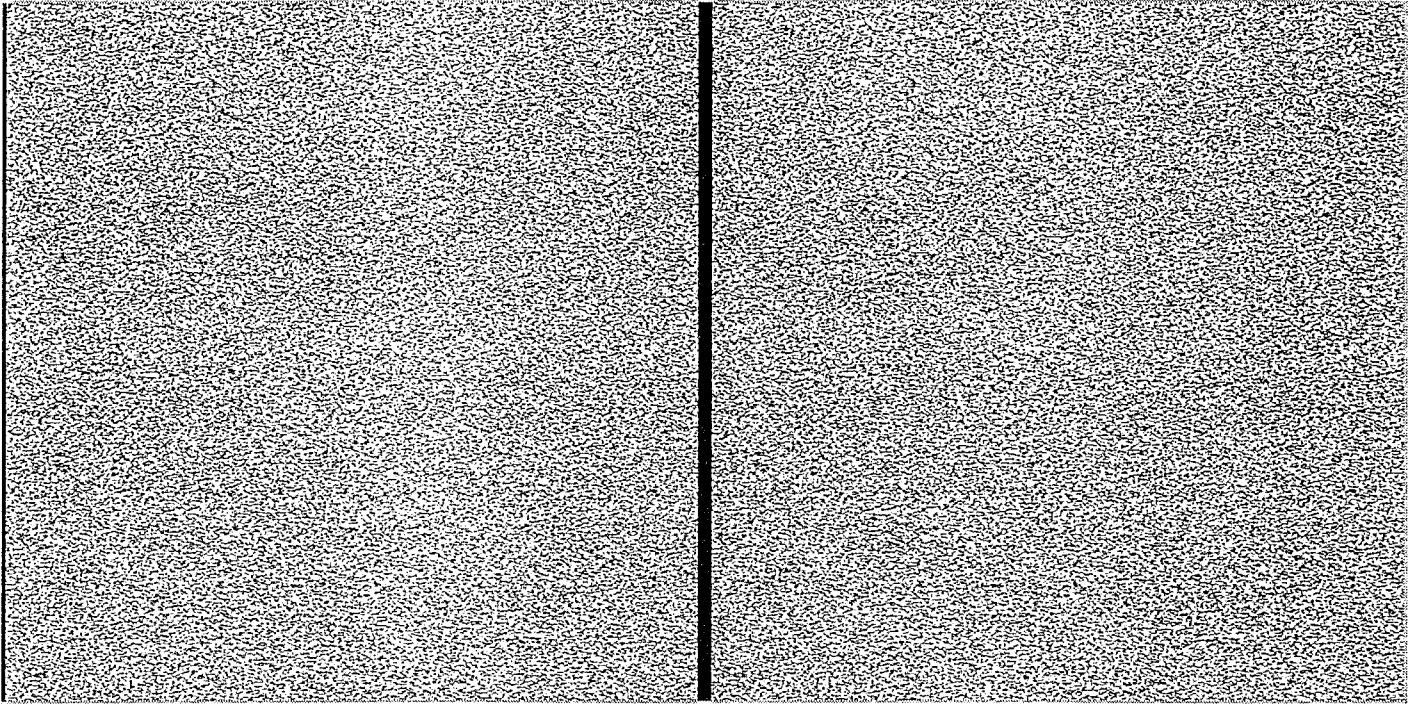


図 1 純粋のトランスペアレンシーを含むランダム輝度ステレオグラム (交差法) [各画像の大きさは、 512×512 [pixel]、各面の視差は、それぞれ、背景： $D_1 = -5.0$ [pixel]、浮き出た中心部の矩形： $D_2 = 5.0$ [pixel]]

Fig. 1 Random intensity stereogram involving pure transparency (for cross-fusion) [Size of the images is 512×512 [pixel]. Disparities of the background and the pop-out square at the center are $D_1 = -5.0$ [pixel] and $D_2 = 5.0$ [pixel], respectively.]

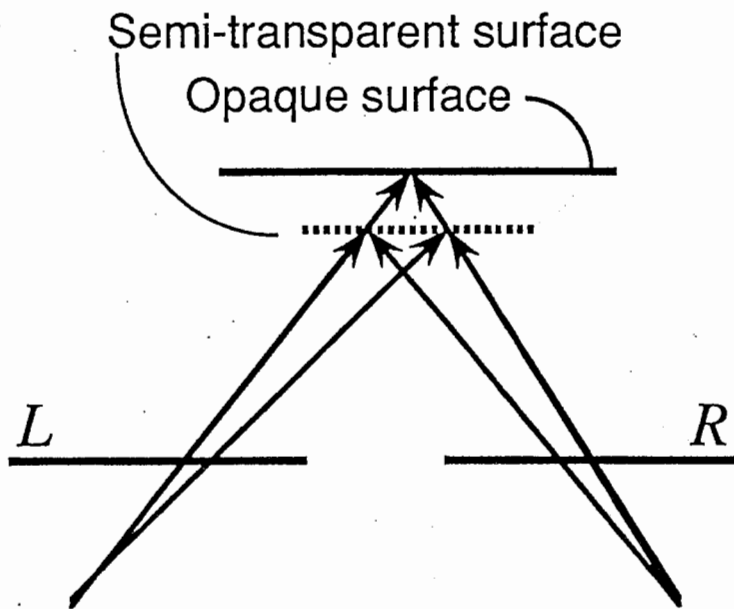


図 2 平行2眼ステレオとトランスペアレンシー
Fig. 2 Parallel binocular stereo and transparency.

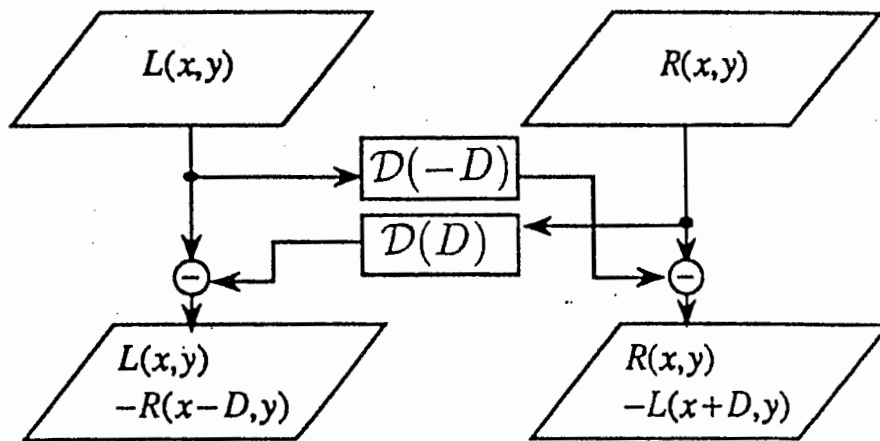


図 3 振幅演算子のはたらき

Fig. 3 Function of the amplitude operator.

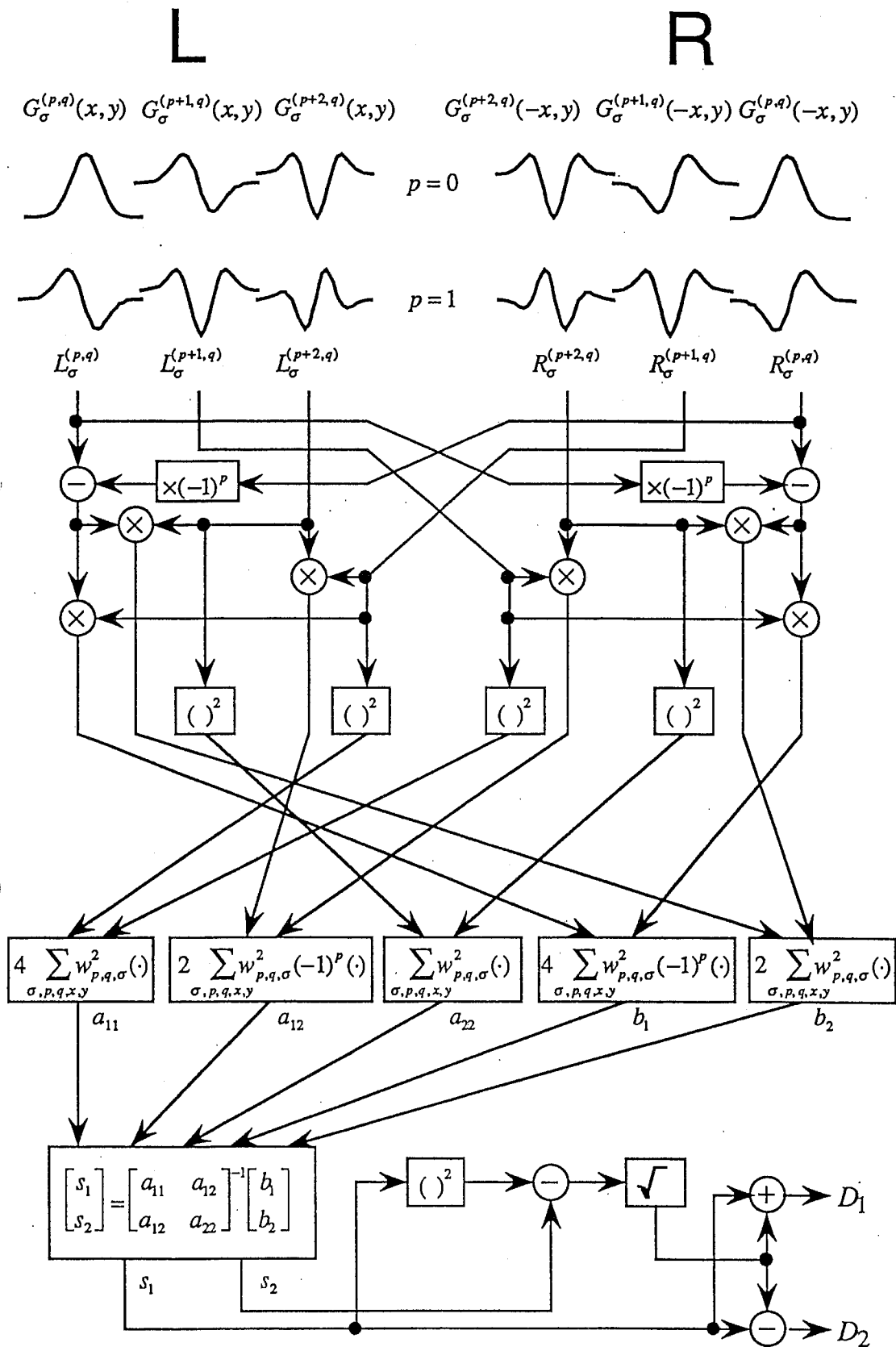


図 4 2 重視差に対する一撃アルゴリズムのネットワーク表現

Fig. 4 Network representation of the derived single-shot algorithm for two-fold disparities.

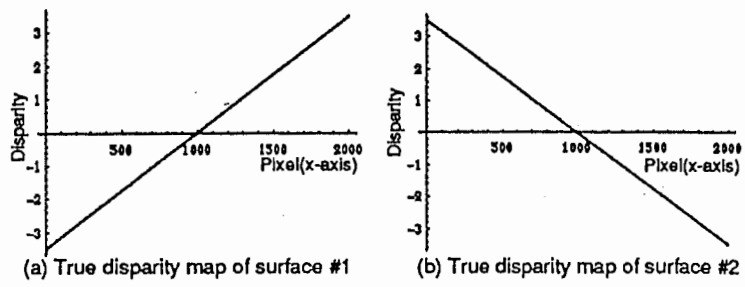


図 5 一次元信号によるシミュレーションで用いた2つの視差マップ
 Fig. 5 Two disparity maps for the computer simulation using 1-D signals.

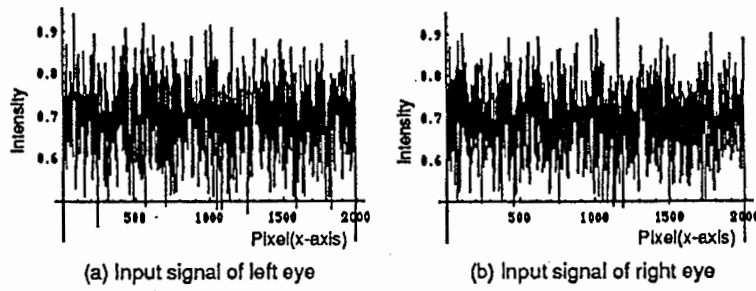


図 6 シミュレーションのために生成された 2 重視差を含む一次元信号

Fig. 6 1-D signal containing two-fold disparities generated for the simulations.

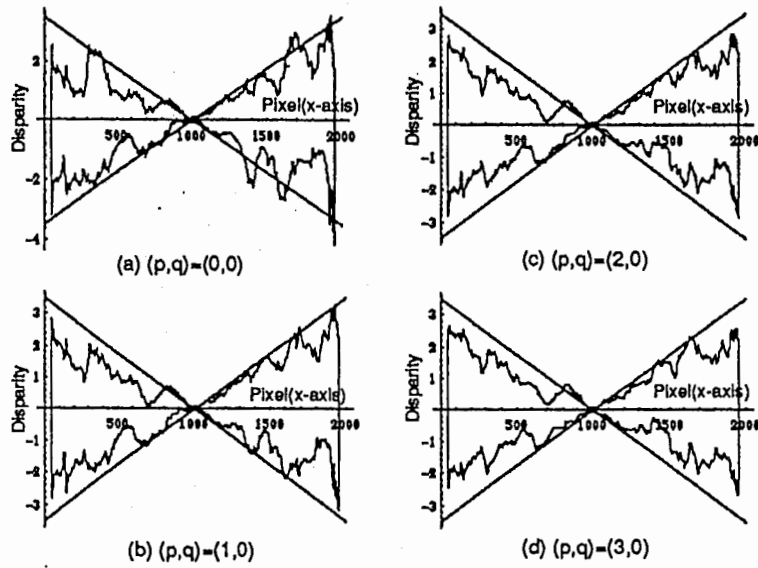
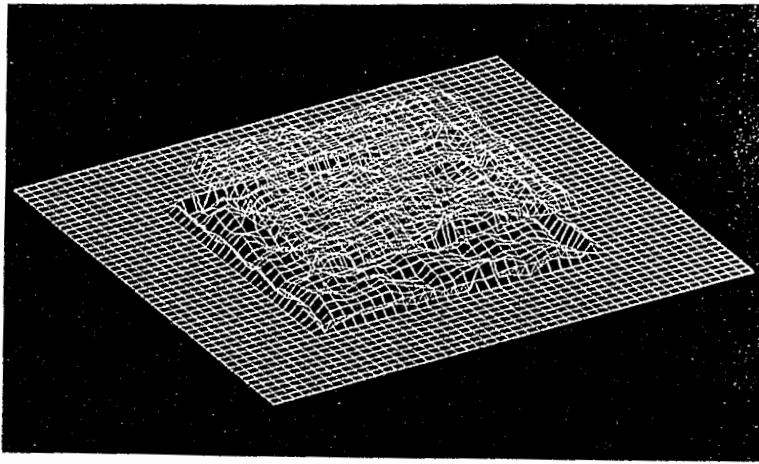


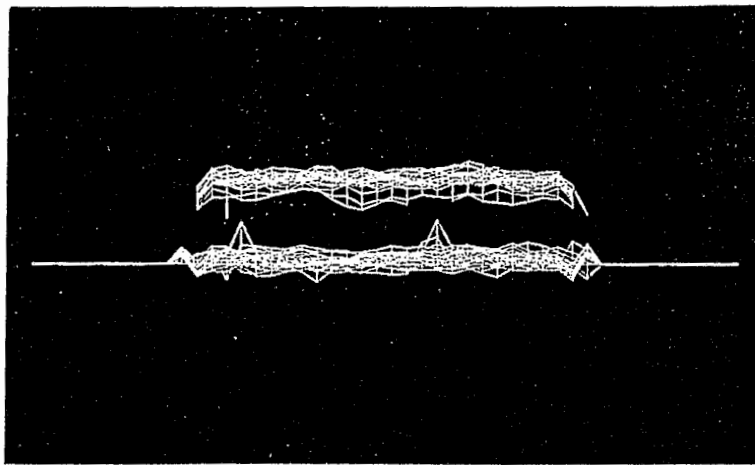
図 7 一撃アルゴリズムによる 2 重視差推定結果 (真の視差マップと
の重畳表示)

Fig. 7 Estimated two-fold disparities by the single-shot algorithm.
The disparity estimates are shown with superimposed true
disparity maps.



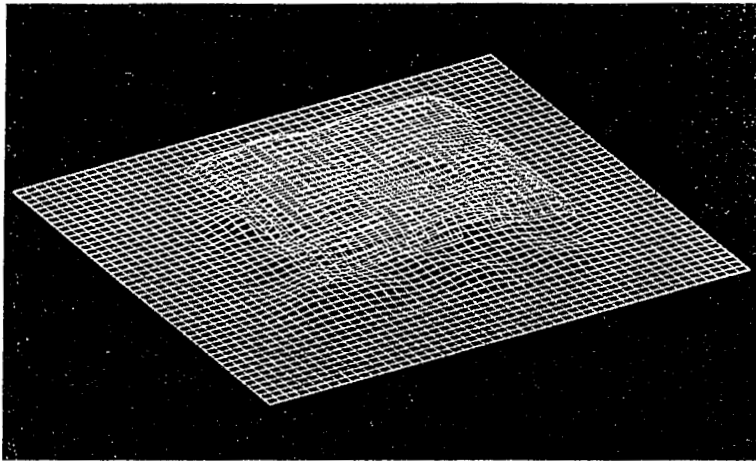
(a) Raw disparity map

図 8 2次元のランダム輝度ステレオグラムに対する推定結果
Fig. 8 Estimated disparity fields for a random intensity stereogram(See text).



(b) Raw disparity map (side view)

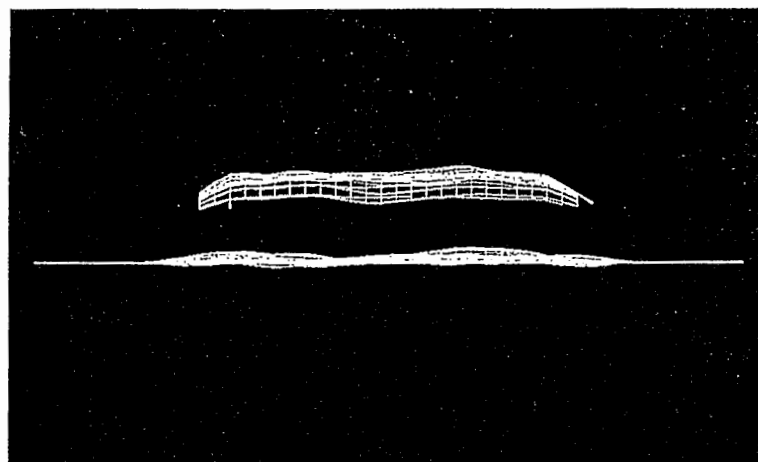
図 8 2次元のランダム輝度ステレオグラムに対する推定結果
Fig. 8 Estimated disparity fields for a random intensity stereogram(See text).



(c) Smoothed disparity map

図 8 2次元のランダム輝度ステレオグラムに対する推定結果

Fig. 8 Estimated disparity fields for a random intensity stereogram(See text).



(d) Smoothed disparity map (side view)

図 8 2次元のランダム輝度ステレオグラムに対する推定結果
Fig. 8 Estimated disparity fields for a random intensity stereogram(See text).