

## RDSの考えを用いた奥行き分布検出の試み

三口貴史\*, 三池秀敏\*, 田中恭治\*\*

\*山口大学, \*\*山口県工業技術センター

視覚処理系の再現という立場から、筆者らはランダムドットステレオグラム(RDS)に対する新しいアプローチを提案する。それは、両眼のランダムドットパターンを重ね合わせることにより、左右の画像間での対応付けを決定する手法である。本稿では、同一視差、同一テクスチャを持つ分割面や領域を抽出するために、重ね合わせ演算において排他的論理和(XOR)を用いる手法を紹介する。また、このテクスチャ分割を行なった画像から奥行きを得るために、テクスチャ密度、領域連結、及び、二次元離散 Wavelet 階層化を用いた3つの手法を試みる。

## Detecting depth map from Random Dot Stereogram(RDS) with XOR logic

Takashi Mikuchi\*, Hidetoshi Miike\*, Kyoji Tanaka\*\*

\*Yamaguchi Univ.

\*\*Industrial Tech. Inst. Yamaguchi Pref. Govern.

On the analogy to human visual system, we propose a new approach to random dot stereogram(RDS). By overlaying binocular pictures of random dot, the correspondence between right and left image is determined. In this paper, we introduce an exclusive-or(XOR) logic to extract a segmented plane or region having the same depth and same texture in the overlaying method. Three methods of texture density, region connection and 2D discrete Wavelet transformation are tested to obtain the depth map from the segmented texture image.

## 1 はじめに

視覚計算理論を提唱した D.Marr によると、複雑な情報処理系を理解するには計算理論レベル、表現とアルゴリズムのレベル、ハードウェアによる実現レベルの3つのレベルのとらえ方がある。これらはそれぞれ視覚系研究の3つのアプローチである視覚計算理論、視覚心理学、及び、視覚生理学に相当する<sup>1),2)</sup>。比較的最近になって生理学的知見と解剖学的知見とが統合され、視覚生理学の分野において初期視覚系における構造と機能が次第に明らかにされてきている<sup>2),3),4)</sup>。筆者らは視覚処理系の再現という立場から、視覚情報の流れを処理する各部をモジュール化する試みとして、まず奥行き検出を対象とした。現在、奥行きの知覚には視覚前野の V2 と呼ばれる部位にある両眼視差選択細胞が関与しているとされている。今回提案する手法は、この部位での情報処理の模倣を目的としたものである。すなわち、画像全体をずらしながら重なった部分に対して演算を行なうことにより左右の画像間の対応を求め、そこから視差を算出して奥行き情報を得ようとするものであり、両眼立体視の分野に属する。そこで、この分野の研究者に深い興味を持たれている RDS を例にとり、この手法の有効性を調べる。

## 2 ランダムドットステレオグラム

我々が外界の3次元的情報を検出できる理由の一つとして、目が2つあることが挙げられる。奥行きの知覚は、同一の対象をわずかに異なる視点から観察することにより、両眼に写る網膜像にずれが生じ、そのずれが脳内で検出されることにより生じるとされている。これが両眼立体視の原理である。この両眼立体視の研究を科学的に進めるきっかけとなったものに、B.Julesz が考案し

た RDS(Random Dot Stereogram) がある<sup>5),6)</sup>。RDS とは、一見ランダムな点の集合で構成されていながら、見方によっては物体が立体的に浮かび上がって見える画像のことである (Fig.1)。片目でこの図形を見ても全く形を知覚することはできない。しかし、左のランダムパターンを左目で、右のランダムパターンを右目でみることにより、面が浮かび上がって見える。Fig.1 の場合、手前から長方形、円が知覚されるはずである。この結果から、RDS は奥行き知覚が特徴点(領域)なしでの対応付けが可能であることを示している。

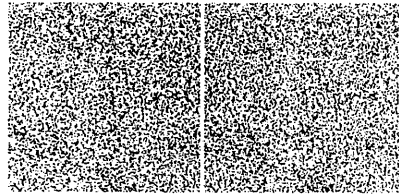


Fig.1 ランダムドットステレオグラム

RDS がランダムな点の集合でありながら立体的に見えるのは、浮かび上がって見える領域と背景となって見える領域とで対応する画像上の領域が異なるため視差を生じるからである。では、この対応付けはどのように行われているのであろうか? 乾らによると<sup>2)</sup>、この対応付けには3つの問題があるという。第一はいかにして両眼の対応付けをとるかということ、第二にその対応からいかにして正確な奥行きを得るかということである。第三の問題として挙げられているのは、なぜ RDS が点ではなく面として知覚されるのか(面の補間)、なぜ面の不連続を知覚できるのかという2つの問題としてとらえることもできる。筆者らは、視差が両眼視差選択細胞という細胞レベルで検出されるという生理学的知見から、視差検出にはそれほど複雑な処理を必要としないと考え、単純に左右の画像を重ね合わせることを考えた。

### 3 両眼立体視

まず、奥行き計算の基礎となる両眼立体視について説明する<sup>6),7)</sup>。基本原理図を Fig.2 に示す。2台のカメラが目に相当する。カメラ間距離を  $x$ 、カメラの焦点距離を  $h$ 、カメラの光軸方向に  $D$  だけ離れた面と左右のカメラの光軸との交点を  $O_1$ 、 $O_2$  とする。図中の点  $P$  は、左の画像では中央から  $u_1$ 、右の画像では中央から  $u_2$ 、右にずれて見える。このとき、三角形の相似関係を考えると、

$$h : D = u_1 : \overline{O_1P} = u_2 : \overline{O_2P} \quad (1)$$

が成り立つ。ここで、

$$\overline{O_1P} - \overline{O_2P} = x \quad (2)$$

を考慮すると、(1) 式は次式のように変形できる。

$$\begin{aligned} D &= \frac{h(\overline{O_1P} - \overline{O_2P})}{u_1 - u_2} \\ &= \frac{hx}{\Delta u} \end{aligned} \quad (3)$$

よって、カメラの焦点距離  $h$  が既知であれば、画像中から視差  $\Delta u = u_1 - u_2$  を求めることにより実際の距離  $D$  が計算できる。

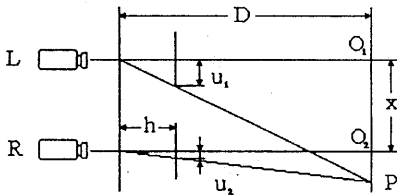


Fig.2 両眼立体視の基本原理解

### 4 奥行き知覚

視差を求めるために筆者らが提案する手法は、左右の画像をずらしながら、その重なり合った領

域に対して演算を行ない (Fig.4)、その結果から視差分布を得ようとするものである。ある奥行きにある対象は、左右の画像においてほぼ同様の輝度分布を持ちながら、奥行きが違う対象とは異なる視差をもって画像中に存在する。よって、左右の画像をその対象の持つ視差  $d$  だけずらして重ね合わせ、各点での輝度値にある種の変換を施せば、その対象の部分が抽出されるはずである。演算を施す重なり合った領域  $\Omega$  上  $(x, y)$  での値  $\omega(x, y, d)$  は、左画像の輝度分布を  $l(x, y)$ 、右画像の輝度分布を  $r(x, y)$  とすると、

$$\omega(x, y, d) = f(l(x + d, y), r(x, y)) \quad (4)$$

で表される。但し、位置  $(x, y)$  は、それぞれの領域の左上を原点とする。

今回の対象となる RDS は二値画像と見なせるので、各ドットに論理演算を施すことにより同一視差にある対象が抽出できる。Fig.1 に対して XOR 演算 (排他的論理和) を施した例を Fig.3 示す。

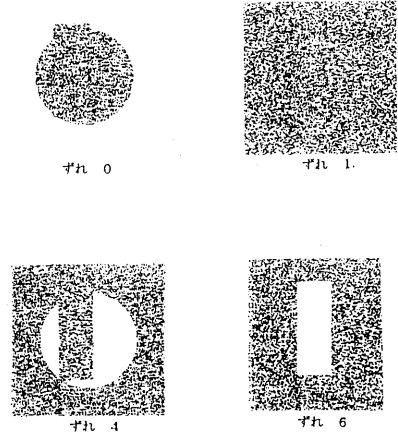


Fig.3 重ね合わせ演算

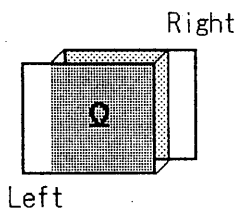


Fig.4 画像の重ね合わせ

#### 4.1 画素密度変化の利用

左右画像をずらしながら重ね合わせ演算を行なうことによって、視差分布検出のための領域分割を行なうことができる。実際に視差分布を求めるには、重ね合わせ演算後の領域 $\Omega$ から情報を抽出しなければならない。第一の手法として、領域 $\Omega$ 上に局所領域を考え、画像のずれによってその領域内での画素密度が変化することを利用する方法を考える。ある代表点におけるずれ $d$ に対する局所領域での画素密度 $J$ の変化の様子を Fig.5 に示す。

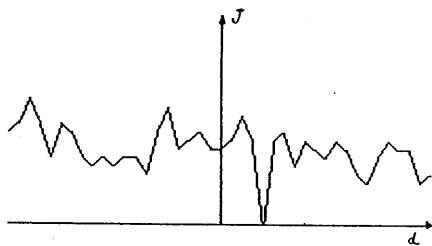
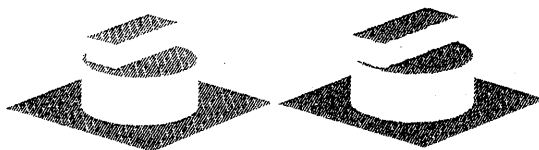


Fig.5 ずれによる局所領域での密度変化

画素密度が最小となるずれをその局所領域での視差とする。局所領域のサイズを  $(2w_x + 1) \times (2w_y + 1)$  としたときの位置  $(x, y)$  での局所密度は次式で表される。

$$J(x, y, d) = \sum_{\Delta x = -w_x}^{w_x} \sum_{\Delta y = -w_y}^{w_y} \omega(x + \Delta x, y + \Delta y, d) \quad (5)$$

Fig.3 に対しては $J$ を最小とする $d$ を位置 $(x, y)$ での視差とすればよい。この手法による結果を Fig.6 に示す。これは、対象とした画像が二値画像であることから、発想の基本は異なるが原理的にはマッチング法と同等の結果となり、局所領域のサイズの設定によって結果が変わり、エッジがぼけるなどの欠点がある。



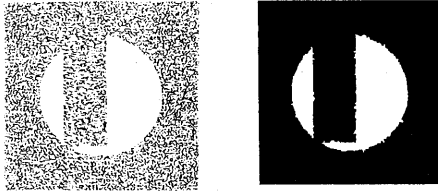
(a) 理論値 (b) 抽出結果

Fig.6 局所領域での密度変化を利用した奥行き分布検出

#### 4.2 領域連続性を利用

別の視差抽出法としては、領域抽出法がある。RDSがドットではなく面が浮かんでいるように知覚されるという事実からも、視差分布を領域で抽出することには意味がある。対象がRDSの場合には、Fig.3を見ても明らかなように対応する視差において連続的な領域が知覚できる。これは、XOR演算を通して対象のテクスチャの違いを明確にし、両眼視差の重要な側面である面の補間を自動的に行なったことに対応している。また、ある大きさ以上の領域しか領域として認識していないこともわかるので、あるずれを与えた領域 $\Omega$ から連続する領域を抽出して、適当なサイズ以上のものをそのずれに対応する視差を持つ領域として選択する。この手法による解析結果を Fig.7 に示

す。特徴として、エッジ部分が他の手法と比べて鮮明になるが、奥行き分布が連続的に変化する場合には連続性の判断が難しいといった問題がある。



(a) 重ね合わせ画像

(b) 抽出領域

Fig.7 領域の連続性の利用

### 4.3 離散 Wavelet 階層化の利用

Fig.3 からの領域抽出は、テクスチャ分離を用いることにより行なえる。ずれに対応した視差の領域は連続的で高周波成分を含んでいないのに対し、それ以外の領域にはランダムドットパターンが現れるため高周波成分を多く含むことになる。そこで、二次元離散 Wavelet 変換による階層化の上位層で、あるしきい値以下の高周波成分を取り除き、逆 Wavelet 変換によって画像を再構成してやると、対象領域と背景部分とのテクスチャを分離することができる。

#### 4.3.1 離散 Wavelet 階層化とは

離散 Wavelet 変換による階層化<sup>8),9)</sup>は次のように定式化される。信号  $f(x)$  の属する空間においてスケール関数  $\phi(x)$  が存在し、

$$\phi_{j,n}(x) = 2^{\frac{j}{2}} \phi(2^j x - n) \quad j, n \in Z \quad (6)$$

はこの空間の直交基底となる。内積  $\langle \cdot, \cdot \rangle$  によって

$$h(n) = \langle \phi_{-1,n}(u), \phi(u-n) \rangle \quad (7)$$

を考え、 $\tilde{h}(n) = h(-n)$  とすれば、信号  $f$  と関数  $\phi_{j,n}$  の内積は、

$$\langle f(u), \phi_{j,n}(u) \rangle = \sum_{k=-\infty}^{\infty} \tilde{h}(2n-k) \langle f(u), \phi_{j+1,k}(u) \rangle \quad (8)$$

で表される。また、 $f(x)$  の属する空間と直交する空間においては、

$$\psi_{j,n}(x) = 2^{\frac{j}{2}} \psi(2^j x - n) \quad j, n \in Z \quad (9)$$

で表される直交基底が存在する。

$$g(n) = \langle \psi_{-1,n}(u), \psi(u-n) \rangle \quad (10)$$

によって関数  $g$  を定義すると、信号  $f$  と関数  $\psi_{j,n}$  の内積は、

$$\langle f(u), \psi_{j,n}(u) \rangle = \sum_{k=-\infty}^{\infty} \tilde{g}(2n-k) \langle f(u), \psi_{j+1,k}(u) \rangle \quad (11)$$

で表される。この  $\psi_{j,n}(x)$  が直交 Wavelet と呼ばれ、式 (8), (11) で表される変換が離散 Wavelet 変換である。 $\tilde{h}$  のフーリエ変換  $H$  は低域通過特性、 $\tilde{g}$  のフーリエ変換  $G$  は高域通過特性を持つ。

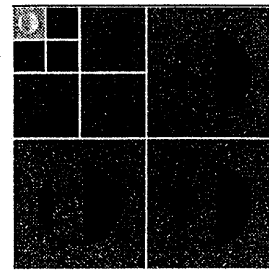


Fig.8 離散 Wavelet による階層化

#### 4.3.2 テクスチャ分離

Fig.9 にずれ4で重ね合わせ演算をした領域に対して、高周波成分除去の方法を変えた結果を示す。Fig.9(a) では、あるしきい値以下の画素において1階層上の画素がしきい値以下のときに高周

波成分を除去し、Fig.9(b)では、しきい値以下の画素の3×3近傍領域において、1階層上の画素がしきい値以下のときに高周波成分を除去している。直線部分の検出は(a)の方がきれいに出ているが、曲線部分も含めて考えると(b)の方がよい結果が出ているといえる。この手法はランダムドットパターンが高周波成分を持つことを利用しているので、実画像において用いるのは困難かもしれない。

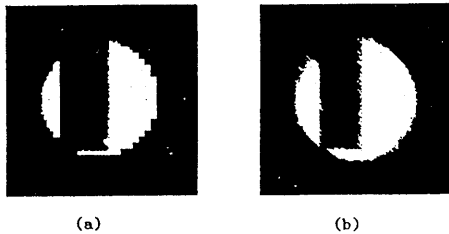


Fig.9 離散 Wavelet 変換の利用

## 5 まとめ

筆者らは、視覚情報処理系の再現という立場から、両眼立体視の一手法として単純な重ね合わせ演算による奥行き検出手法を提案した。RDSという興味深い対象に対する適用結果は満足のいくものであった。しかし今回示した手法は、RDSという二値画像であり、しかも高調波成分を多く含む対象に限定した特殊な処理であることは否定できない。現在、この提案を一般的な実画像へ適用するための手法を模索中である。

## 参考文献

- [1] David Marr(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(1988)
- [2] 川人光男, 行場次郎, 藤田一郎, 乾敏郎, 力丸裕: "認知科学3 視覚と聴覚", 岩波書店, pp.89-127(1994)

- [3] 乾敏郎: "視覚情報処理の基礎", サイエンス社, pp.2-102(1990)
- [4] 平井有三: "視覚と記憶の情報処理", 培風館, pp.29-72(1995)
- [5] Bela Julesz: "Foundations of Cyclopean Perception", The University of Chicago Press(1971).
- [6] 乾敏郎: "Q&A でわかる脳と視覚-人間からロボットまで-", サイエンス社, pp.150-163(1993)
- [7] Berthold K. P. Horn(NTT ヒューマンインターフェイス研究所プロジェクト RVT 訳): "ロボットビジョン-機械は世界をどう見るか", 朝倉書店, pp.329-368(1993)
- [8] 田中恭治, 中村誠, 白上貞三, 堺敏明, 古賀和利, 三池秀敏: "Wavelet 階層化による対象表面テクスチャの抽出"
- [9] Charles K. Chui(桜井明, 新井勉 訳): "ウェーブレット入門", 東京電機大学出版局, pp.1-98(1993)