

両眼映像の干渉処理によるRDSからの曲面の再構築

○小川新矢, 宇野公裕, 三池秀敏
山口大学工学部

ロボットビジョンの分野の一つとして両眼立体視による2次元画像からの3次元情報の再構築というテーマがある。物体の形や陰影などの様々な情報を含む現実画像でこの問題を解決するのは非常に困難である。ところがランダムドットステレオグラム(RDS)は中でも視差の情報しか含まず、両眼立体視のメカニズムの解明の手がかりとして好都合である。本研究ではヒトがRDSを両眼立体視で見ようとするとき両眼の画像を重ねようとして見ていることから、2値画像であるRDS画像を僅かにずらし干渉させることにより視差領域を抽出することを試みた。この手法を拡張し多値画像の干渉画像の密度を見ることで視差領域の抽出を試みた。

Curved-Surface Reconstruction from Interference Pattern of Stereo Images in Random Dot Stereogram

○Shinya Ogawa, Kimihiro Uno, Hidetoshi Miike
Faculty of Engineering, Yamaguchi University

One of the most important aim of robot-vision is the reconstruction of 3-D shape from 2-D image. In real image, there are many helpful information to reconstruct 3-D shape, however, it is generally difficult to solve a stereo problem only from a single key. Random dot stereogram (RDS) has only one information of disparity. In this paper, we develop a method to reconstruct 3-D shape from interference pattern of stereo images in RDS. By applying this method to real stereo images with multiple gray levels, we try to reconstruct 3-D shape by use of the density of the interference pattern.

1. はじめに

人間の視覚による画像情報処理能力は素晴らしいものがある。この能力の秘密を探るため今までに様々な視点からの研究がなされてきた[1]。近年、この分野において大きなインパクトを与えたのがDavid Marrである。David Marrの提唱した3つのレベルの計算論的アプローチには、計算理論、表現とアルゴリズム、ハードウェアによる表現があり、視覚系の研究に大きな役割を果たしている[2]。2次元画像から3次元情報を再構築する3次元画像計測の分野での研究は2つに大別される。一つはビジョンと呼ばれ、人間の視覚機能のサイエンスとしての理解を目指す分野である。もう一つは、工学的アプローチであり、人工的に視覚機能を実現し工業に生かそうとする分野である[3]。

3次元情報を得る手がかりとして両眼視差、運動視差、陰影、テクスチャー、輪郭、接点などがあげられる[4]。3次元情報を捉える機能としていくつかのモジュールが脳に存在することが知られているが[5]、我々が日頃見ている世界は実に様々な情報を含んでおり、単一モジュールによるモデル化では限界がある。しかし1960年Bela Juleszの作ったランダムドットステレオグラム(RDS)には視差情報しか含まれず、見た目はランダムな点の集まりであるのに、両眼によって融合することで3次元構造が知覚される[6]。RDSは左右の画像においてドットが対応している。対応しているドットがお互いの画像でずれている大きさが視差となり我々に奥行き感を認知させるのである。RDSを解くことにより、両眼立体視の機能を理解する手助けになることが考えられる。しかし、通常のテンプレートマッチングなどの方法では対応点探査は非常に困難である。最近、現実の濃淡画像に関してはテンプレートマッチングを応用した方法[7][8]など数多くの手法が提案されている。一方、ヒトの視覚機構を模倣した手法としてはGrisword[9]らの研究がある。最近、曲面を表すRDSを解く手法として点对点の対応をニューラルネットワークにより解くことで3次元構造復元の手法が報告されている[10]。

しかし、人間の視覚機能を考察した時、立体視に関しては、学習機能を持つ数値化機構のような複雑な処理を行っているようには思えない。より単純な対応付けのアルゴリズムが期待される。そこで、我々がRDSをみるときに左右の画像をずらし重ねることで対応を探すと仮定し、両眼からの映像を重ね合わせその結果得られた干渉画像から3次元構造を再構築する事を提案してきた[11]。前回の報告ではこの単純なアルゴリズムにより離散段階の3次元情報取得が可能であることを確認している。本研究では更に曲面の再構成を目指し、なめらかに変化する多段階の視差情報を持つRDSに対しこのアルゴリズムを適応した。さらにこの手法を拡張し、対象画像をRDSから現実の多値(濃淡)画像にかえて、3次元構造の再構築を試みた。

2. 原理

2.1 両眼立体視

両眼立体視とは、呼んで字のごとく左右2つの目で物体を見ることにより、3次元情報を再構成することである。2つの目を2つのカメラに見立ててその光学系を示すと図1のようになる。厳密にはヒトの視覚系の場合、図1.aのように両眼の光軸が1点に向かっておりこの系をモデルにした研究も試みているが[12]、ここでは簡単のため図1.bのように左右のカメラの光軸が平行である場合をまず考える。ここで、カメラ間距離を x 、カメラの焦点距離を f 、カメラから D 離れた面左右のカメラの光軸との交点を O_1 、 O_2 、とする。図1中の点 P は、カメラに捉えられた画像 S_L 、 S_R ではそれぞれ光軸の交差する点を中心としてそれぞれ u_1 、 u_2 だけずれて見える。ここで、3角形の相似の関係より

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

が成り立つ。次に式1を変形して、

$$D = \frac{f(\overline{O_1P} - \overline{O_2P})}{u_1 - u_2} = \frac{fx}{d} \quad (2)$$

ただし、

$$x = \overline{O_1P} - \overline{O_2P} \quad (3)$$

よって、画像上での視差 $d = u_1 - u_2$ が解れば実際の距離 D を求めることができる。

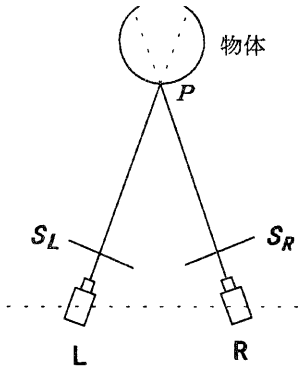


図 1.a ヒトの視点

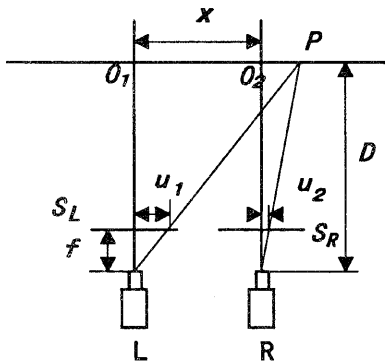


図 1.b 平行な光軸を持つカメラによる両眼立体視

2.2 RDSの解法

前節で述べたように、画像の左右の対応点が解れば実際の距離を求めることができる。このことは左右の画像におけるマッチングに他ならない。しかしテンプレートマッチングに代表されるマッチングの手法ではRDSにおいて対応点を探すのは通常困難であり、図2に示すRDSを単純なマッチングの手法で解くと、図3のようにほとんど mismatches を起こしてしまう。このRDSの左右の画像をOHPシートなどに写し、重ね合わせずらしてゆくと図4の様に隠されているパターンを発見することができる。今回提案し

た方法はこの原理を利用し、2枚の画像全体をある視差ずらして重ね合わせた時に現れる干渉パターンを利用する。

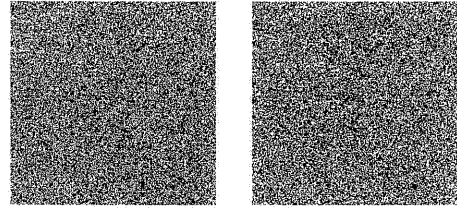


図 2 ランダムドットステレオグラム

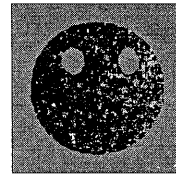
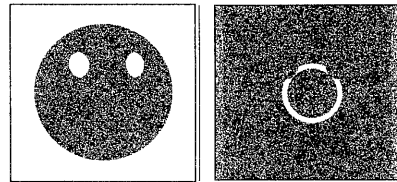


図 3 単純なマッチングの結果



$d = 0$

$d = 37$

図 4 ずらして重ね、XORによる干渉画像

2.3 3次元情報再構築アルゴリズム

今回提案したアルゴリズムの説明を図5に示す。図に示すようにRDSの左右の画像をある大きさ d [pixels] だけずらして重ねる。この d は視差を示しており、平行法で裸眼立体視をしたときRDSの描かれる紙面を基準として正の値の時は飛び出し、負の値の時奥に見えるものとする。そして重なる部分のピクセル値同士でXORの演算を行う。RDS画像は図2の白い点が0、黒い点が1の値を持つ2値画像であるためお互いのピクセル値が同じ値を持つときに0となる。このとき視差 d [pixels] を持つ領域があれば両画像が対応していることから図5に示すように(この

図の場合は丸でかかれた部分) その部分だけ抜けてしまう。この領域をあるサイズの連結領域として抽出することにより、視差 d [pixels] の領域を確定する。次に領域抽出のアルゴリズムの説明する。まず、両画像の XOR 演算を行った画像の左上から走査して行き、注目点とその近傍 8 [pixels] での演算値が 0 である点を探す。次にその点の近傍 8 [pixels] が連続か否か調べるのだが、その条件は図 6 に示す通りである。点 B とその近傍 8 [pixels] の演算値が注目点 A と同じ値を持つとき点 A は点 B と連続であると定義した。連続と見なされる領域がなくなるまで走査を繰り返すことによって 1 つの連続領域を確定した。このときランダムドットの間隙を縫って本来の視差に対応する領域以外の小領域が出現するため、第 1 段階では連続領域が 30 [pixels] に満たない領域は本来の視差領域と見なさないようにした。さらに連続領域の定義では抽出領域がひとまわり小さくなるので、第 2 段階では抽出した領域の周り 1 [pixels] も抽出領域と見なすこととした。

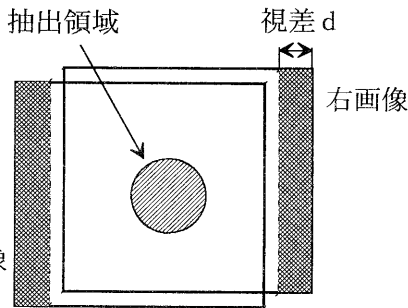


図 5 画像干渉による視差領域の抽出

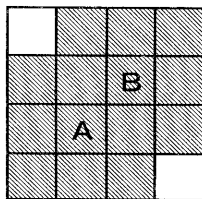


図 6 連続領域抽出アルゴリズム

2. 4 RDS の 3 次元情報の再構築

今回解析に用いた RDS 画像を図 2 に示す。画像サイズは 512×512 [pixels] の 2 値画像である。また、立体視する事で見える画像は視差 0 ~ 30 [pixels] までの範囲に分布する半球状のもので、視差 0 をもつ 2 つの穴がいている。ランダムドットの密度は 30% とした。実際の 3 次元情報を濃淡画像として図 7 に示す。この画像に対し $d = -30 \sim 30$ [pixels] 間での範囲で視差領域の抽出、再構築を行った。この結果を図 8 に示す。図 7、8 とも明るいほど d が大きいことを示している。全体的に領域の一部が欠けているが、これは、図 7 の様に分布する視差情報を含む RDS に直すときに視差領域をわずかにずらしながら重ねてゆくために両画像間の対応がつかなかったためと思われる。また、半球の縁に当たる狭い領域は正しい視差領域が得られていないが、その他の広い領域は正しい視差の値が得られており、ほぼ正しい 3 次元情報が得られた。

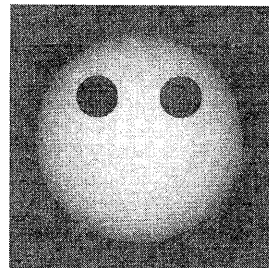


図 7 実際に与えた 3 次元情報

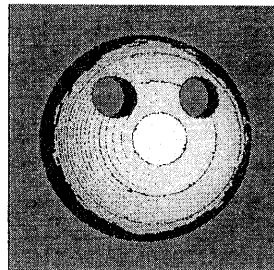


図 8 図 2 に対する第 1 段階の結果

3. 現実画像（多値画像）への拡張

前章にも書いていたように、RDSの左右の画像をずらして重ねてゆくことにより一瞬対応視差領域が現れることが解る。これはある視差 d [pixels]だけずらして重ねてみたときに次々と変化するドットの密度変化を見ていることに変わりない。これを考慮して、多値画像（256段階）に対応したアルゴリズムに拡張した。簡単に書くと、まずはじめに左右の画像全体を視差 d [pixels]ずらし、重なった部分のピクセル値同士で演算を行う。演算方法としては2章で述べたように互いのピクセル値の各ビットのXORを取る方法、または差分を取る、和を取るなどが考えられるが、今回は差分値の絶対値を取ることにした。次に重ねて演算を施した画像に対し、ある大きさを持ったテンプレートを用意する。ここでテンプレート内のテクスチャーの密度を見ることで視差を求める。今回、テクスチャーの密度を評価するのに分散の値を用いた。

3.1 アルゴリズム

まず、図9に示すようにカメラで捉えた左右の画像 $L(x,y)$ と $R(x,y)$ 同士を視差 d [pixels]ずらし、重なる部分の輝度値の差の絶対値を計算する。この画像を $I(x,y,d)$ とする。すなわち、

$$I(x,y,d) = |L(x,y) - R(x+d,y)| \quad (4)$$

である。そして各点を中心として縦横 $(2M+1)$ [pixels]の大きさを持ったテンプレートを考える。次に次式、

$$m(x,y,d) = \frac{1}{(2M+1)^2} \sum_{i=-M}^M \sum_{j=-M}^M I(x+i,y+j,d) \quad (5)$$

よりテンプレートが掛かっている領域の輝度値の平均 $m(x,y,d)$ を計算する。さらに次式より、

$$V(x,y,d) = \sum_{i=-M}^M \sum_{j=-M}^M |m(x,y,d) - I(x+i,y+j,d)| \quad (6)$$

点 $I(x,y,d)$ を中心としたテンプレート領域内の分散 $V(x,y,d)$ を求める。この $V(x,y,d)$ が最小になる視差 d [Pixel]が点 $I(x,y,d)$ での視差とすることで3次元情報を求める。

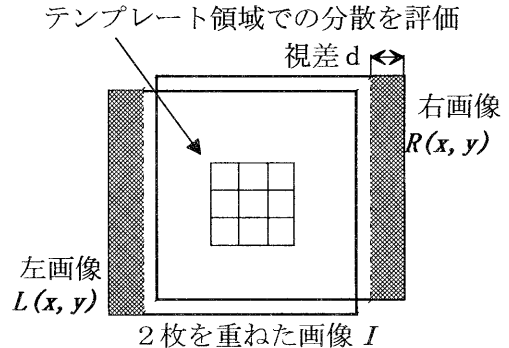


図9 拡張アルゴリズム

3.2 多値画像の3次元情報の再構築

今回実験に使用した画像は図10、11に示す、球のシミュレーション画像と3つの物体が存在している実画像である。1台のカメラで捉えた映像でも、カメラの移動中に対象物体が静止していれば、2台のカメラで捉えたときと同じと見なせる。実画像は図12のように、静止した物体に対し、1台のカメラをある一定の速度 Δt で動かして撮影した動画像の中から、2枚の画像をとりだして両眼の画像としている。どちらも画像サイズは 128×128 [pixels]で256段階の濃淡画像である。テンプレートの大きさは 13×13 [pixels] 探索する視差の範囲は $d = -20 \sim 20$ [pixels]とした。本手法との比較のために、同一条件下で、テンプレートマッチングにより計算を行った。テンプレートマッチングとは、図13に示すようにある大きさ（例えば 3×3 [pixels]など）のテンプレートを用意し、一方の画像におけるテンプレートを固定して他方の画像のテンプレートを移動させながら両方のテンプレート領域の差分を取り、それらを足しあわせた値がもっとも小さいときの視差を結果として得る。ただし、テンプレートマッチングでは対応点がなくなるのをできるだけさけるために、図14の様に対象となる

$t=0$ の画像の領域を分割し3枚の画像を使い、マッチングを行った。

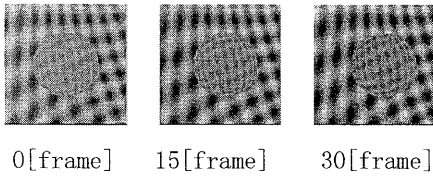


図10 実験に用いたシミュレーション画像



図11 実験に用いた実画像

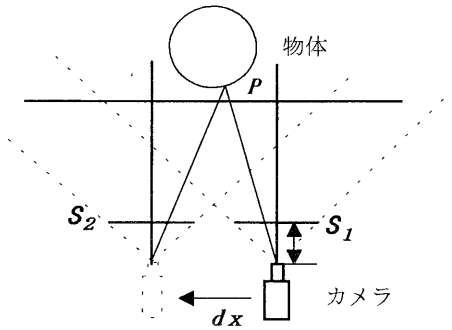


図12 実画像の計測

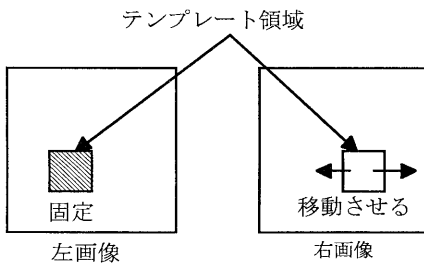


図13 テンプレートマッチング

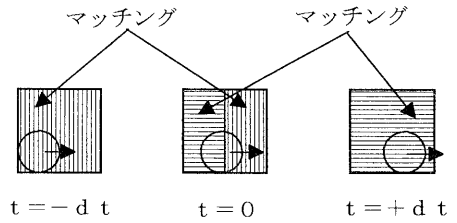
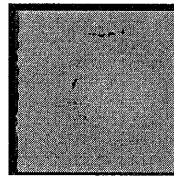


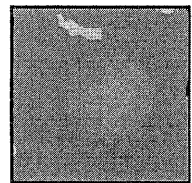
図14 3枚の画像からのマッチング

3.3 実験結果

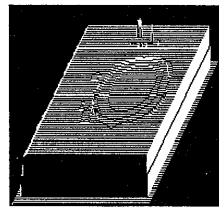
今回提案した手法には図10における15, 30[frame]、図11における15, 25[frame]を両眼画像として用いた。また、マッチング法には図10, 11で3枚すべての画像を用いた。これらの画像を用いて3次元情報を再構築した画像を次の図15, 16に示す。両図とも左の画像が今回提案した手法による結果であり、右の画像がマッチングによって得られた結果を示す。



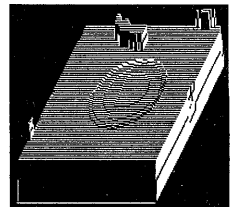
濃淡表示



濃淡表示

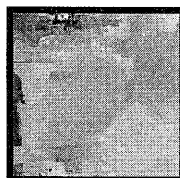


3次元表示
(今回の手法)

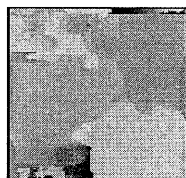


3次元表示
(マッチング法)

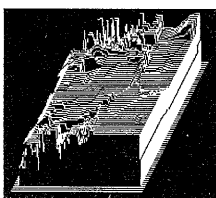
図15 図10の画像に対する結果



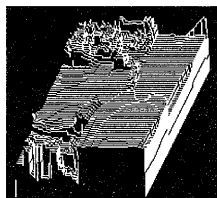
濃淡表示



濃淡画像



3次元表示
(今回の手法)



3次元表示
(マッチング法)

図16 図11の画像に対する結果

シミュレーション画像における3次元情報の再構築については、特に画像上中央部分に誤対応が大きく現れているマッチング法に比べ、提案手法でよい結果が得られた。球の形も球の中央部が欠けているマッチング法に比べより正しく捉えられている。ただ、マッチング法に比べ本手法では領域のエッジ部分がはっきりしていない。次に実画像から得た結果を見比べてみると、やはりマッチング法による結果と比べエッジがかなりぼやけており、改良を要する。

4. まとめ

今回提案した手法は、ヒトが両眼で3次元空間を見るとき、注視点付近では物体がクリアに見え、注視点からはずれたところはぼやけて見える。つまり両眼画像を僅かにずらしつつテクスチャーなどの密度の変化を見て物体の奥行きの情報を得ているのではないかとこの仮定が元になっている。例えばRDSを考えたとき、両眼画像を重ね合わせ少しづつずらして行くことによって、ある視差領域どうしが重なったとき瞬時にランダムドットの密度が下がった領域を認知することができ

る。2値のRDSの場合、重ねてXOR演算をすることで視差領域が得られた。この手法はランダムドットの各点ごとの対応を出す方法に比べて、モデルがシンプルであり簡単な演算で解を得ることができるという点では非常に興味深いことである。しかし、半球状の端の細い領域がうまく検出できなかった。この解決方法として考えられることは、まず領域の広いところから解いて行くことや、連続領域と見なす評価をゆるめることなど改良の余地がある。

つぎに多値画像に対する拡張アルゴリズムについて考えてみる。今回用いた画像は両眼のカメラの光軸が平行な状態で明確な“視点”というものが存在しなかった。しかしヒトはものを見るとき、ある点において両眼の光軸が交わる。この交わった点付近は、両眼画像のマッチングがうまくいっているところに他ならない。この注視点付近が全体の奥行きを認識するための1つの手がかりとなると考えられる。今後カメラの光軸を平行にしないで、光軸がある位置で交差しているような画像に対して提案した手法を適応していく予定である。

参考文献

- [1] J.P. フリスビー著、村山久美子訳：“シーイング 錯視—脳と心のメカニズム”，誠信書房，1982.
- [2] 川人光男，行場次郎，藤田一郎，乾俊郎，力丸裕：“認知科学3 視覚と聴覚”，岩波書店，1994.
- [3] 井口征士，佐藤宏介：“三次元画像計測”，昭晃堂，pp. 29-72，1995.
- [4] 乾俊郎：“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，1993.
- [5] 平井有三：“視覚と記憶の情報処理”，培風館，1995.
- [6] David Marr (乾俊郎，安藤広志訳)：“ビジョン—視覚の計算処理論と脳内表現—”，産業図書，1988.
- [7] 奥富正敏，金出武雄：“複数の基線長を利用したステレオマッチング”，電子情報通信学会論文誌 D-II，J75-D-II，pp. 1317-1327，1992.

- [8]Ting-Chuen PONG, Robert M. HARALICK,
Linda G. SHAPIRO :
" Matching topographic structures in
stereovision" ,Pattern Recognition
Letters **9**(1989),pp.127-136.
- [9]N. C. Griswold, C. P. YEH : " A New
Stereo Model Based upon the
Binocular Fusion Concept" ,COMPUTER
VISION, GRAPHICS, AND IMAGE
PROCESSING **41**, pp. 153-171, 1988.
- [10]塚原朋哉, 平井有三 : " ランダムドット
からの曲面再構築" , 電子情報
通信学会論文誌D-II, **J76-D-II**,
pp. 1676-1683, 1993.
- [11]三口貴史, 三池秀敏, 田中恭治 : " R
DSの考えを用いた奥行き分布検出の
試み" , コンピュータービジョン研究
会, **98-1**, 1996.
- [12]宇野公裕, 三池秀敏 : " A Stereo
Vision through Creating a Virtual
Image using Affine Transformation"
MVA' 96 IAPR Workshop on Machine
Vision Applications, pp. 526-529,
1996.