

## 中心視—周辺視モデルによるステレオ画像の対応探索

池田昇 中村康弘

防衛大学校情報工学科

本報告では、ステレオ画像における対応探索の精度を改善する一手法を提案する。一般に使用されているブロックマッチング法はステレオ画像間の視点の違いによる注視点付近の歪みに対して誤差が生ずる。人間の網膜では網膜の中心部（中心視）よりも周辺部（周辺視）の方が受容野が広く、解像度が低いという特性があるが、本研究ではこの特性を利用したモデルによってステレオ画像の対応探索を行った。本手法により通常のブロックマッチング法に比較して対応付けの精度が向上した。

## Stereoscopic Image Matching Using the Central-Peripheral Vision Model

Noboru Ikeda Yasuhiro Nakamura

Dept. of Computer Science, National Defense Academy

This paper investigate the method which reform matching of stereoscopic image. Block matching, generally using, make error, which occur by distortion for two different view-point. On human retina, the receptive field which exist on the periphery of retina, peripheral vision, is larger than the receptive field which exist on the center of retina, central vision. We apply the model which has this feature to matching of stereoscopic image. The results indicate that this model make more exact matching than block matching.

# 1 はじめに

近年、非接触で物体の3次元形状を復元する研究が盛んに行われているが、その手法は能動的手法と受動的手法の2つに分けることができる。能動的手法は、精密かつ信頼性が高く主に工業分野において研究開発が進められてきたが、計測装置がレーザー光等を用いる大規模かつ高価なものであり、また計測環境に多くの制約を受ける。それに対して受動的手法は、CCDカメラ等の受動的センサのみで計測が可能であり、安価かつ広範囲な撮影対象を扱うことができる。しかしながら画像を処理する段階で多くの要因による誤差が生ずるため計測精度の向上が難しいという問題がある。

本報告は、対象物体を2方向より撮影したステレオ画像を用いて高精度な受動的計測を行うことを目的としているが、この手法における最も大きな問題は、ステレオ画像の対応探索をいかにして行うかということにある。ステレオ画像の対応探索は左右の画像から特徴的な部分を抽出して、その部分同士で対応付けを行う特徴ベース法[1]と画像間の局所領域を比較する領域ベース法に大別できる。特徴ベース法は特定の人工物のような各部分の特徴がきわだった対象物体の場合には、左右画像の特徴点の位置関係が明確であるため正しい対応付けを行いやすい。しかしながら、自然物のように対象物体中に類似した特徴点が多数現れるもの場合は対応付けが困難である。本報告では広範囲な撮影対象を扱うために領域ベース法による対応探索を研究対象とする。

領域ベース法によるステレオ画像の対応探索は、一般に画像の一部の領域(ウィンドウ)内の局所的パターンを比較することにより行う。これは一般にブロックマッチングと呼ばれている。例えば、左画像の1点 $(x_l, y_l)$ につき、 $(2n+1) \times (2m+1)$ 画素からなるウィンドウを設定し、右画像のエピポーラ線上の各点 $(x_r, y_r)$ を中心とするウィンドウとの類似度を求める。その中で類似度が最も高く、かつしきい値以上のものを対応点とする。

類似度の評価はさまざまあるが、次式で表される相関値 $s$ を用いるものが一般的である[2][3]。

$$s(x_l, y_l, x_r, y_r) = \sum_{i=-n}^n \sum_{j=-m}^m (I_l(x_l+i, y_l+j) - \bar{I}_l) \times (I_r(x_r+i, y_r+j) - \bar{I}_r) / ((2n+1)(2m+1)\sqrt{\sigma^2(I_l) \times \sigma^2(I_r)})$$

$$\times (I_r(x_r+i, y_r+j) - \bar{I}_r) / ((2n+1)(2m+1)\sqrt{\sigma^2(I_l) \times \sigma^2(I_r)})$$

$I_l, I_r$ : 画像の濃度値

$\bar{I}_l, \bar{I}_r$ : ウィンドウ内の平均値

$\sigma^2(I_l), \sigma^2(I_r)$ : ウィンドウ内の標準偏差

ここで、ある左画像の点 $L(x_l, y_l)$ が右画像の点 $R(x_r, y_r)$ と対応関係にある場合、その近傍の点の間にはどのような関係があるであろうか。ステレオ画像は対象物体(3次元)を左右の撮影面(2次元)に投影したものであるが、このとき対応がとれている注視点及びその近傍の点に注目すると、左右の画像間(2次元)では注視点を中心とする回転及び拡大・縮小変換と見なすことができる。ただし、各近傍の点ごとの変換は一律ではない。点 $L, R$ 近傍の対応のとれている各々の点を $Li(x_l + \Delta x_{li}, y_l + \Delta y_{li}), Ri(x_r + \Delta x_{ri}, y_r + \Delta y_{ri})$ とすると、その視差 $\Delta LRi(\Delta x_{li} - \Delta x_{ri}, \Delta y_{li} - \Delta y_{ri})$ の大きさは注視点からの距離が大きくなるにつれて増大する。すなわち、

$$\max_{\theta(r=a)} |\Delta LRi| \propto r$$

$$(r = \sqrt{(\Delta x_{li} - \Delta x_{ri})^2 + (\Delta y_{li} - \Delta y_{ri})^2}, \theta = \tan^{-1} \frac{\Delta y_{li} - \Delta y_{ri}}{\Delta x_{li} - \Delta x_{ri}}, a = \text{常数})$$

の関係がある(図1)。従って、 $(n \times m)$ の長方形の局所領域を比較するブロックマッチングでは、注視点からの距離に比例する視差の増大に対応しきれない。

ところで、現実の人間の網膜では受光細胞によって光を検出し神経節細胞を経て外側膝状体に出力される。このとき1つの神経節細胞の出力に影響を及ぼす範囲は網膜上の円形の領域で、これを受容野と言う。神経節細胞とその受容野の網膜上での分布は一律ではなく、偏心度(網膜の中心からの距離)が増すにつれて分布の密度は減少し同時に受容野の大きさが増すことが知られている[4][5]。したがって網膜の中心部(中心視)では検出される情報の密度は高く、周辺部(周辺視)では低くなる。

この受容野の特性は、先に示した対応探索の注視点近傍の誤差に対してきわめて有効なモデルである。すなわち、対応探索の注視点を網膜の中心と考える

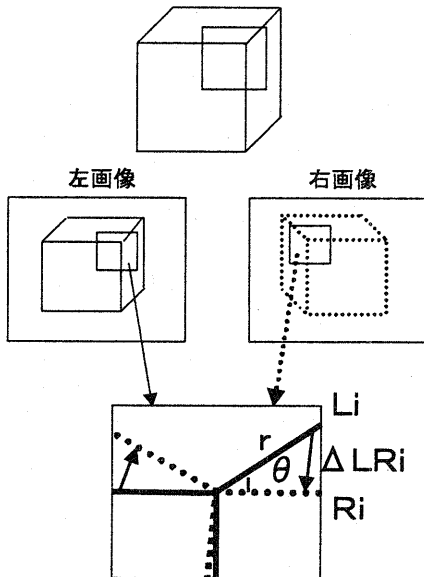


図 1: 対応近傍点の幾何学的性質

と、注視点からの距離が増大するにつれて受容野の出力に影響を与える網膜上の領域が大きくなるので誤差の増大に対応することができる。

そこで本報告では、以下の章でこの特質を利用したモデルを提案する。さらに、このモデルを実際の画像に適用した結果を示す。

## 2 中心視一周辺視モデル

### 2.1 処理の概要

以下の手順で左右の画像の対応探索を行う。(図2)

1. 一方の画像から対応探索を行う注視点を選択
2. もう一方の画像の最大許容視差内から対応点候補となる注視点を選択
3. 左右の注視点付近の受容野の出力を計算
4. 左右の受容野の出力から対応候補の信頼度を算出
5. 2から4をすべての対応点候補について適用し最も信頼度の高いものを対応点として決定

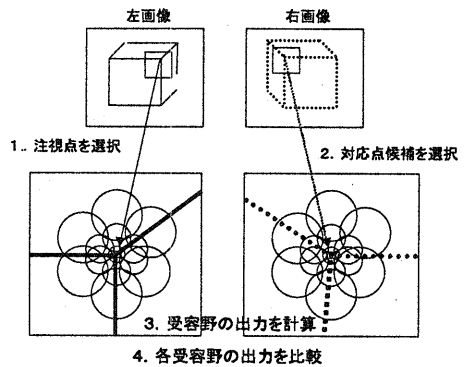


図 2: 処理の概要

### 2.2 各受容野の出力

まず注視点の位置を基準に極座標  $(r, \theta)$  の位置に受容野の中心を設定する。これらの受容野は、それぞれ中心からの距離が  $r(n)$  の位置に、角度  $\Delta\theta(n)$  の間隔で配置される。(n:注視点からの受容野の数)(図3)

次に各受容野の出力を算出する。各受容野の中心を原点とする極座標を  $(r', \theta')$ 、画素値を  $I(r', \theta')$ 、受容野の範囲を  $D$  (面積  $S(D)$ ) とすると、各受容野の出力  $v(r, \theta)$  は以下のように表される。

$$v(r, \theta) = \frac{1}{S(D)} \iint_D I(r', \theta') c(r', n) dr' d\theta'$$

$c(r', n)$  は対象の座標と各受容野の中心からの距離と、その受容野の注視点からの距離に依存する係数である。

### 2.3 信頼度の計算

前項で算出した左右の各受容野の出力を比較し、各対応点候補の信頼度を計算する。信頼度を  $s$  とすると

$$s = 1 - \frac{1}{I_{max}} \left( \frac{1}{N} \sum_r \left( \sum_{\theta} (v_l(r, \theta) - v_r(r, \theta)) \right) \right)$$

$N$ : 受容野の個数

$I_{max}$ : 画素値の最大値

$v_l$ : 左画像の各受容野の出力

$v_r$ : 右画像の各受容野の出力

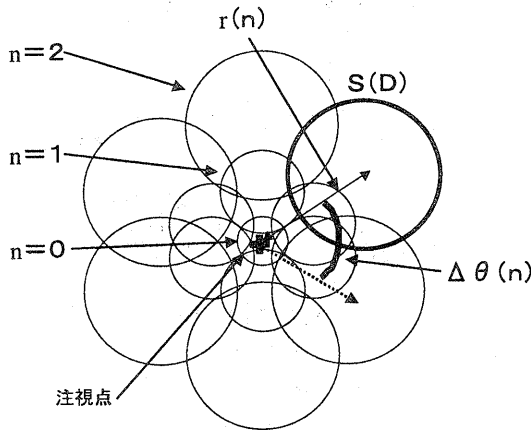


図 3: 各受容野の出力

信頼度  $s$  は左右の受容野の出力がすべて一致するとき最大値 1、画素値の最大値分の差があるとき最小値 0 をとる。

このように各対応点候補について信頼度を計算し、その値が最も大きいものを対応点として採用する。

### 3 実装モデル

前章の中心視一周辺視モデルでは極座標の連続値で処理を記述したが実際の画像は離散的な 2 次元の配列で表現されている。そこで実際の画像において信頼度の計算を以下の要領で実施する。

前処理として 1 つの注視点近傍の全ての受容野を包含するウィンドウの領域を決める。以後の処理はこの領域を対象として行われる。

次にこれらの領域に含まれる画素の座標を読み出し、この画素がどの受容野に包含されるかを計算する。画素が各受容野に与える出力はその画素と各受容野との距離によって決定される。注視点  $(x, y)$  からの位置ベクトルが  $(\Delta x, \Delta y)$  の画素は、注視点からの位置ベクトルが極座標  $(r, \theta)$  の位置にある受容野に  $c(r', n)I(x + \Delta x, y + \Delta y)$  の出力を与える。

$(I(x + \Delta x, y + \Delta y))$ : 画素値,

$$r' = \sqrt{(r \cos \theta - \Delta x)^2 + (r \sin \theta - \Delta y)^2}$$

そこで各画素ごとに各受容野の出力に対する係数  $c(r', n)$  を計算しこれを画素ごとの変換テーブル  $t(\Delta x, \Delta y, m)$  ( $m$ : 受容野の識別番号) に格納する。

対応点候補の信頼度を計算するにあたっては、ウィンドウ内の画素値  $I(x + \Delta x, y + \Delta y)$  を呼び出すごとに変換テーブル  $t(\Delta x, \Delta y, m)$  の値と掛け合わせて、その値を各受容野の記憶領域  $M(m)$  に加算する。このとき各受容野のカウンター  $c(m)$  の値を 1 つ加算する。この計算をウィンドウを構成する全ての画素について行う。最終的な受容野  $m$  の出力は  $M(m)/c(m)$  である。こうして得られた左右の各受容野の値を前章の要領で比較し信頼度を算出する。

ここで、このモデルの注視点近傍の各画素と受容野との関係を検証する。例として、

(1) 最も外側の受容野 ( $n = 4$ )

$$r(n) = 2n, \Delta\theta(n) = \frac{\pi}{4}, c(r', n) = 1.0(|r'| \leq (1.4)^n)$$

(2) 最も外側の受容野 ( $n = 2$ )

$$r(n) = 2n, \Delta\theta(n) = \frac{\pi}{4}, c(r', n) = 1.0(|r'| \leq (1.4)^n)$$

(3) 最も外側の受容野 ( $n = 4$ )

$$r(n) = 2n, \Delta\theta(n) = \frac{\pi}{3}, c(r', n) = 1.0(|r'| \leq (1.4)^n)$$

(4) 最も外側の受容野 ( $n = 4$ )

$$r(n) = 2n, \Delta\theta(n) = \frac{\pi}{4}, c(r', n) = 1.0(|r'| \leq (1.2)^n)$$

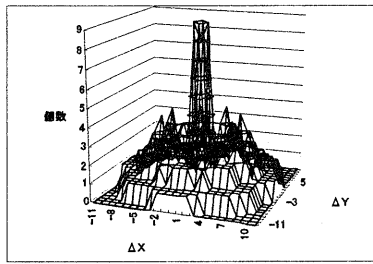
の 4 つを検証する。

上記の 4 つの例について、ウィンドウ内の各近傍点は何個の受容野に含まれるかを図 4 に、近傍点が含まれる最大の受容野の半径を図 5 に示す。各近傍点は、包含される受容野の数が多いほど信頼度に対する重みが大きくなり、包含される受容野の半径が大きいほどより大きな注視点付近の誤差に追従できると考えられる。

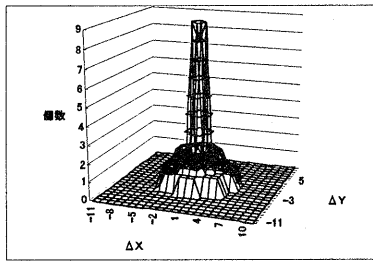
(1) と (2) から (4) を比較すると、受容野の分布範囲が大きいほどより広範囲の近傍点の値が信頼度に反映され、受容野の配置が密で各受容野の半径が大きくなるほど注視点から見て等方位的に近傍点が信頼度の計算に反映されることがわかる。

### 4 実験

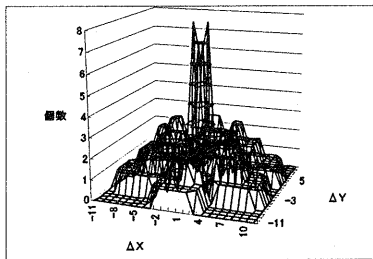
ここで、実際の画像を用いて 3 章で示した中心視一周辺視モデルと一般的なブロックマッチングとの対応探索における精度を比較する。



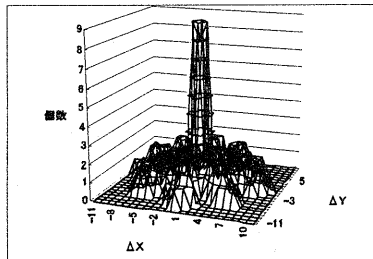
(1)



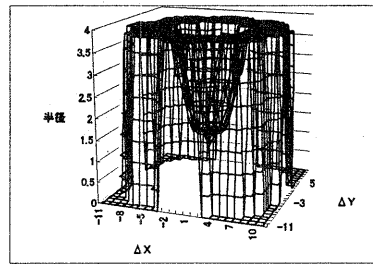
(2)



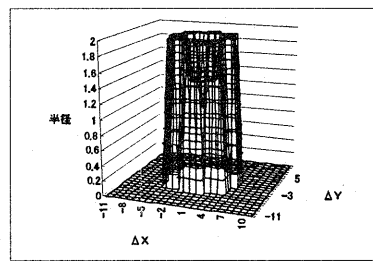
(3)



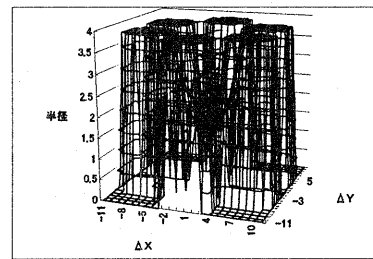
(4)



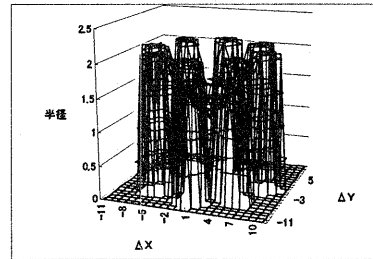
(1)



(2)



(3)



(4)

図 4: 各近傍点が包含される受容野の数

図 5: 各近傍点が包含される最大受容野の半径

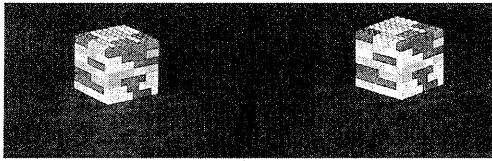


図 6: ステレオ画像

使用するステレオ画像は2台のCCDカメラを用いて模型を2方向から撮影したものである(図6)。ステレオ画像の撮影面上の座標(2次元)と復元する撮影対象(3次元)との変換行列は既知であるものとする。このステレオ画像から人間の目によって対応点を100組抽出する。次にブロックマッチングと中心視一周辺視モデルによって、それぞれ先に抽出された対応点の左画像上の点100個と対応する右画像上の点100個を算出する。各々の手法によって算出された右画像上の点と人間の目によって抽出された対応点との平均対応誤差(単位:画素)をもって対応探索の精度とする。

#### 4.1 ブロックマッチングの対応探索の精度

ブロックマッチングによる対応探索の精度を図7に示す。ウィンドウの大きさを  $n = 8, 9, 10, 11, 12, 13$  の6通りについて対応探索の精度を算出した。

ブロックマッチングでは対応探索の精度はウィンドウのサイズに比例しない。この実験例では  $n = 10$  で誤差が最小となりそれ以上では逆に誤差が増大する。これは1章で示したように注視点からの距離が大きくなると長方形のウィンドウでは視点の違いによる左右の画像の歪みに対応できないからである。

#### 4.2 中心視一周辺視モデルの対応探索の精度

中心視一周辺視モデルによる対応探索の精度を示す。

変数の値は、3章の(1)を基準とし、最も外側の受容野の輪の数を変えた場合 ( $n = 2, 3, 4$ ) (図8)、受容野の配置の間隔を変えた場合 ( $\theta(n) = \frac{\pi}{3}, \frac{\pi}{4}, \frac{\pi}{6}$ ) (図9)、各受容野の半径を変えた場合 ( $c(r, r') = 1.0(|r'| \leq (1.1)^n, (1.2)^n$ ,

$(1.3)^n, (1.4)^n, (1.5)^n$ ) (図10)の各々について実験を行った。

図8、図9より受容野の数及び密度が増大するほど対応探索の精度が向上することが分かる。また、図10より、受容野の半径がある値のときに視点の違いによる誤差に最適な値をとるので対応探索の精度は最良となることが分かる。

これらの結果から、このモデルはブロックマッチングよりも高い精度の対応探索を行うことができると言える。

#### 4.3 復元された3次元形状の比較

ウィンドウのサイズが  $n = 10$  のブロックマッチングによって3次元形状を復元した像(図11)と前項のモデルの基準の値で復元した像(図12)を示す。これらの像は左右の画像のエピポーラライン上の点を各々の手法で対応探索を行い3次元座標を計算したものである。点間の補完はエピポーラライン上で隣接し、かつ距離が一定のものを接合することにより実施した。像以外の背景は濃度値の大きさの低い部分(黒)を対象外とした。

得られた像を比較すると、立方体の左右両側面は視差が小さいため顕著な違いは見られない。立方体の上面に注目すると、視差が大きく類似したパターンがあるため双方とも誤差が大きいが、中心視一周辺視モデルによる対応探索の方がより良好な像が得られたことが分かる。

## 5 まとめ

本報告では、視覚の中心視一周辺視の特性をモデルにした手法によってステレオ画像の対応探索の精度を改善する方法を提案した。

注視点からの網膜上の距離に比例して受容野が大きくなるモデルによって、ブロックマッチングでは追従できないウィンドウ間の誤差に対応する手法を示した。さらに実験として、CCDカメラによって作成したステレオ画像の対応探索を行い、本手法の有効性を示した。

しかしながら、本手法は受容野の配置及び受容野内の係数の最適性を保証してはいない。したがって、他の対応探索の手法に対して数学的に最適であるとは言えない。そこで、今後は視覚情報処理あるいは

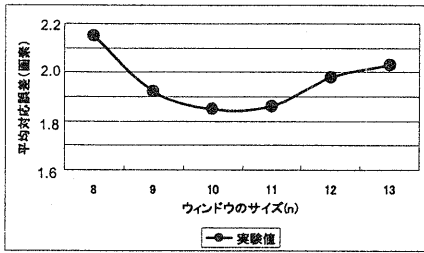


図 7: ブロックマッチングの対応探索精度

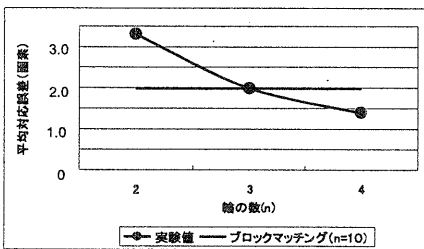


図 8: 最も外側の受容野の輪の数を変えた場合

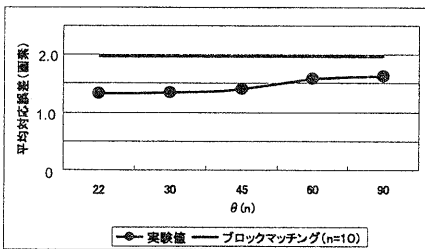


図 9: 受容野の配置の間隔を変えた場合

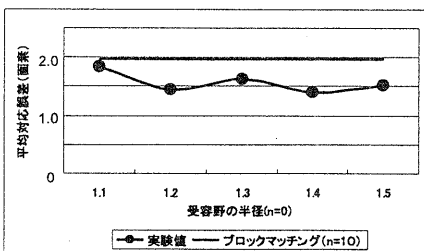


図 10: 各受容野の半径を変えた場合



図 11: ブロックマッチングによる復元

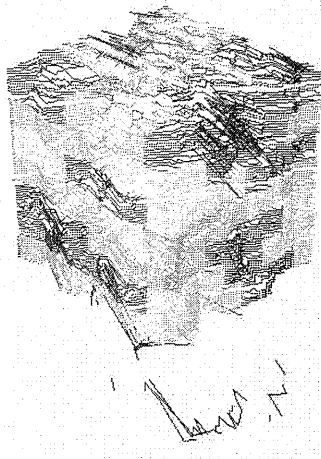


図 12: 中心視一周辺視モデルによる復元

数学的な観点において、より洗練されたモデルを追求する必要がある。

## 参考文献

- [1] 鈴木 英行, 中村 納, "等濃度線分布と新しい  
区間対応法に基づくステレオ対応探索アルゴリズム", 信学論 (D-2), vol.J80-D-2, no.9, pp.2379-  
2392, 1997
- [2] 徐 剛, 辻 三郎, "3次元ビジョン", 共立出版
- [3] 山口 証, 高地 和隆, 井口 征士, "石像の3次元計測のためのステレオ対応付け", 信学論 (D-2), vol.J81-D-2, no.4, pp.716-725, 1998
- [4] 淀川 英司, 東倉 洋一, 中根 一成, "視聴覚の認知科学", 電子情報通信学会
- [5] 近藤 真悟, 佐藤 幸男, "人間の視覚解像度を模した視覚システムの構築", 信学技報, PRUM98-17, no.70, 1998