

## 適応ウィンドウ法を用いた石像計測のためのステレオ対応付け

山口 証, 高地和隆, 井口征士

大阪大学基礎工学部  
システム工学科

### 概要

領域ベースのステレオ法の中心的問題は、最適なウィンドウサイズを選択にある。ウィンドウサイズが大きいと分解能の低下や計算コストの増大を招く。逆にウィンドウサイズが小さいと多対応問題を生じる。我々は適応ウィンドウ法に基づく領域ベースのステレオ対応付け手法を提案する。左ウィンドウに対して右画像の対応可能領域で相関を計算し、相関グラフを作る。我々は相関グラフが最も明確なピークを持つウィンドウサイズを探す。このウィンドウサイズは最も確からしい視差の見積もりを保証する。このアルゴリズムはマルチスケールのテクスチャ表面を持った石像に対して効果的に働く。

## Stereo Matching For Stone Statues Using Adaptive Window Size

*Akashi YAMAGUCHI, Kazutaka KOCHI, Seiji INOKUCHI*

Department Of Systems Engineering  
Osaka University

### Abstract

A central problem in a region-based stereo method lies in selecting an appropriate window size. If the window is large, the resolution of the range becomes low and the computation cost increases. If it is small, a multiple correspondence problem occurs. We propose a region-based stereo algorithm based on an adaptive window size method. A correlation graph is made by computing correlation values of a left window in a correspondence-possible region on a right image. We search the window size whose correlation graph has the clearest peak. This window size provides the most certainly estimation of disparity. This algorithm effectively works for stone statues which have multiple scaled texture surfaces.

## 1 はじめに

世界には様々な石の文化が存在する。これらの石像の形状を計測し、データベースを作成することが文化人類学や考古学、民族学などの分野で重要な課題となっている<sup>[1]</sup>。

両眼立体視は、能動的な距離計測法<sup>[2][3][4]</sup>に比べて特殊な装置を必要としないため、石像のような非工業製品の計測により向いている。両眼立体視では、空間内の同一点に対する左右画面上の点を対応付けなければならない。この対応探索問題を解決するための様々な手法が開発されているが、これらは特徴ベース法と領域ベース法に大別することができる。

特徴ベース法は、画像から特徴を抽出し、特徴同士の対応付けを行う方法である<sup>[5][6][7]</sup>。石像のように表面に細かなテクスチャパターンがある物体に対しては、テクスチャパターンから類似した特徴点が多数現れるため多対応問題を生じ易く、特徴ベース法は適さない。

領域ベース法は、画面上の小領域（ウィンドウ）間の明度の相関をとることにより対応領域を求める問題である<sup>[8]</sup>。領域ベース法で中心的な問題は、最適なウィンドウサイズを選択にある。ウィンドウサイズが小さいと複数の領域と対応する多対応問題を生じる。逆にウィンドウサイズが大きいと分解能の低下や計算コストの増大を招く<sup>[9]</sup>。

本稿では、適応ウィンドウ法に基づく領域ベースのステレオ対応付け手法を提案する。左ウィンドウに対して右画像の対応可能領域で相関を計算し、相関グラフを作る。我々は相関グラフが最も明確なピークを持つウィンドウサイズを探す。このウィンドウサイズは最も確からしい視差の見積もりを保証する。このアルゴリズムはマルチスケールのテクスチャ表面を持った石像に対して効果的に働く。

## 2 相関に基づくステレオ対応付け

今回はステレオカメラモデルとして、左右のカメラの光軸が平行なパラレルステレオモデルを仮定する。対応付けは $m \times m$ 画素（ $m$ は奇数）の矩形のウィンドウごとに実行される。まず、左右両画像上にウィンドウを配置し、左画像上のウィンドウと右画像上のウィンドウの間で

るさの相関を計算する。相関関数は次の通りである。

$$C = \frac{1}{m^2} \frac{\sum_{j=1}^m \sum_{i=1}^m (I_L(i,j) - \mu_L)(I_R(i,j) - \mu_R)}{\sqrt{\sigma_L^2 \sigma_R^2}} \quad (1)$$

$$= \frac{\sigma_{LR}^2}{\sqrt{\sigma_L^2 \sigma_R^2}}$$

$I_L(i,j)$  と  $I_R(i,j)$  はそれぞれ左右画像上のウィンドウ内の明度であり、 $\mu_L$ 、 $\mu_R$ 、 $\sigma_L$ 、 $\sigma_R$  はそれぞれ左右両ウィンドウの明度の平均および分散である。 $\sigma_{LR}$  は両ウィンドウの明度の共分散である。 $C$  は -1 から +1 まで変化する。もし  $C$  が大きければ、両ウィンドウの明度パターンが類似するとして、左右画像間の対応する局所領域と見なされる。

パラレルステレオモデルにおいては、対応点の探索範囲を限定する事ができる。我々は左画像を基準画像として設定し、左画像上の各点に対する対応点を右画像上で探索する。図1に右画像上の探索範囲を示す。

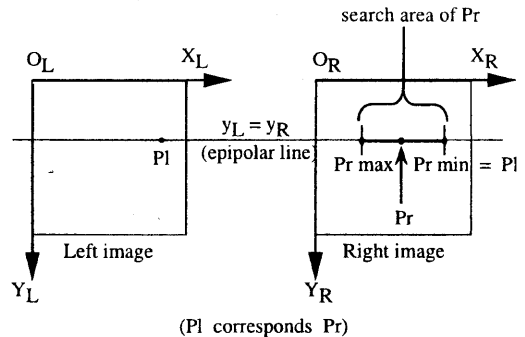


図1 対応点の探索範囲  
Fig.1 Search Area of Corresponding Point

図1において、左画像上の点  $P_L(x_L, y_L)$  に対応する右画像上の点  $P_R(x_R, y_R)$  を探索する。このとき、画像上の全てのエピポーラ線は画面座標系の  $x$  軸と平行となるため、 $P_L$  と  $P_R$  の  $y$  座標値は等しい ( $y_R = y_L$ )。  $P_R$  は右画像上での  $P_L$  と等しい位置よりも左側に存在する。すなわち、 $x_R \leq x_L$  である。よって  $P_L$  と  $P_R$  の間の視差  $D$  は次のように表される。

$$D = x_L - x_R \quad (2)$$

Dの最小値は0である。Dの最大値はカメラシステムのシステムパラメータとカメラから対象物体までの距離に依存する。Prの探索領域は $y_L=y_R$ かつ $x_R=x_L \sim x_L-D_{max}$ の範囲と決定される。この探索範囲内で相関値Cが計算され、水平軸が視差を、垂直軸が相関値をそれぞれ表す相関グラフが作られる。そして、対応位置はこの相関グラフのピーク位置と決定される。

### 3 相関グラフの評価

相関グラフは対応の確からしさを表すと考えられる。相関グラフが明確で大きい唯一のピークを持つ時は、その対応は確からしいと考えられる。しかし、もしピークが小さかったり、複数の似たような高さのピークが現れたり、ピークがはっきりしないような場合は、対応付けの確からしさは低い。

石像の表面の場合、石の組成から生じる材質のむらや表面の細かな凹凸の大きさから、テクスチャパターンの空間分解能が決定される。しかし、これは石像によって様々であり、また一つの石像でも部位により変化する。

石像の画像に対して相関をとる場合、相関グラフの形状は、石像の表面のテクスチャの空間分解能と、相関をとるウィンドウのサイズに依存する。そして、最も対応が確からしいような相関グラフを得られる最適なウィンドウサイズが存在すると考えられる。そのサイズは、異なる石像間や、おなじ石像の画像でも画像上の部位によって変化する。

そこで、次のような評価関数の値を用いて相関グラフから対応の確からしさを評価する事ができる。

$$P = (w_1 P_1 + w_3 P_3) \times w_2 P_2 \times w_4 P_4 \quad (\text{if all } P_i > 0) \\ = -1 \quad (\text{if any } P_i < 0) \quad (3)$$

ただし、

$$P_1 = \begin{cases} C_1 - \gamma_1 & (C_1 > \gamma_1) \\ -1 & (C_1 \leq \gamma_1) \end{cases}, P_2 = \begin{cases} C_2 - \gamma_2 & (C_2 > \gamma_2) \\ -1 & (C_2 \leq \gamma_2) \end{cases} \\ P_3 = \begin{cases} C_3 - \gamma_3 & (C_3 > \gamma_3) \\ -1 & (C_3 \leq \gamma_3) \end{cases}, P_4 = \begin{cases} \gamma_4 - C_4 & (\gamma_4 > C_4) \\ -1 & (\gamma_4 \leq C_4) \end{cases} \quad (4)$$

とする。

ここで、C1は最大ピークの高さ。C2は最大ピークと第2ピークの比。C3は最大ピークの両側の谷との高さの差の小さい方。C4は最大ピークの幅。 $\gamma_1 \sim \gamma_4$ はC1~C4のそれぞれのしきい値。 $w_1 \sim w_4$ は各評価値の重み係数。今回は全部1とする。

この評価値が最も高くなるようなウィンドウサイズを求めて、対応付けに用いる。

相関グラフの評価の概念を図2に示す。

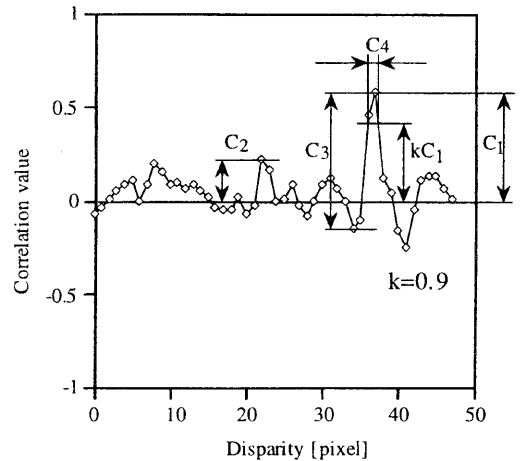


Fig.2 Concept of Correlation Graph Estimation  
図2 相関グラフの評価概念図

### 4 対応付けアルゴリズム

上で述べたような最適なウィンドウサイズを選ぶ方法をステレオ対応付けに適用する。対応付けアルゴリズムの概念を図3に示す。

原画像をウィンドウサイズ毎に別プレーンに分けて、各プレーン毎にウィンドウの評価値と視差を求め、各画素毎に最も評価値の高いウィンドウサイズを選ぶ。

まず用いるウィンドウサイズを決める。次に、各ウィンドウサイズ毎のプレーンを考え、各プレーンの左画像全体にウィンドウを配置する。ただし、ウィンドウ形状は正方形とする。また、奥行きがなめらかに変化する面において隣接するウィンドウの境界で視差がなめらかに変化するよう、ウィンドウは上下、左右とも50%オーバーラップさせる。

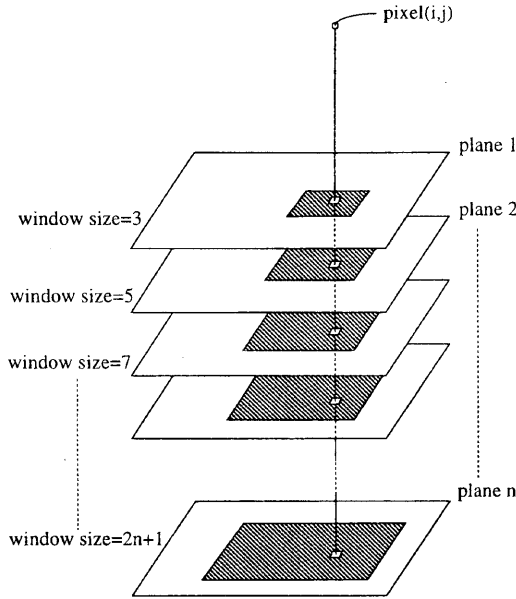


Fig.3 Concept of matching algorithm  
 図3. 対応付けアルゴリズムの概念

つぎに、各プレーン毎にウィンドウの相関を取り、相関グラフの評価値を求める。相関グラフの評価値が0より大きいウィンドウでは、相関グラフの最大ピークを与える視差をそのウィンドウの視差とする。評価値が0より小さいウィンドウでは、ウィンドウの視差を与えない(視差-1とする)。最大ピークの位置は、相関グラフをスプライン補完することにより、サブピクセル単位で求められる。

全てのプレーンでウィンドウの評価値が求められたら、左画像上の各画素について、評価値に基づいて最適なウィンドウを選び、視差を決定する。一つのプレーン上ではウィンドウをオーバーラップして配置してあるため、ある画素を含むウィンドウは最大9個になる。そこでまず、各プレーン毎に、着目する画素を含む全てのウィンドウで、評価値が0より大きいものの多数決をとる。0より大きいウィンドウの方が多ければ、それらの平均値をそのプレーンでの画素の評価値とする。-1のウィンドウの方が多ければ、そのプレーンでの評価値は-1とする。

すべてのプレーンでの評価値が得られたら、その中で評価値が最大となるプレーンを選ぶ。選ばれたプレーンで、着目している画素を含み、かつ評価値が0より大きい全てのウィンドウの

視差の平均をとり、その画素の視差とする。

左画像上のすべての画素で以上の処理を行い、視差画像を生成する。

## 5 実験結果

今回は図4に示すような石の画像に対して対応付けをおこない、視差画像を生成した。画像サイズは320画素×240画素である。対象物体は花崗岩で、おおむね四角柱の形状をしている。寸法はW 90mm×H 130mm×L 330mmである。ステレオカメラは、図5に示すように、2台の三脚で支持された水平なレール上で1台のビデオカメラをスライドさせて実現した。

今回の実験に用いたステレオ画像を撮像したときの位置関係を図6に示す。これらの位置関係から視差と奥行きの変換式を求め、視差情報を距離情報に変換した。

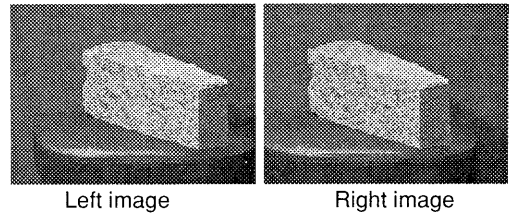


Fig.4 Real stereo image  
 図4 石のステレオ画像

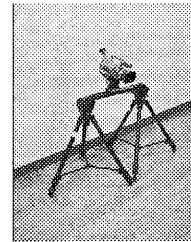


Fig.5 Stereo camera system  
 図5 ステレオカメラシステム

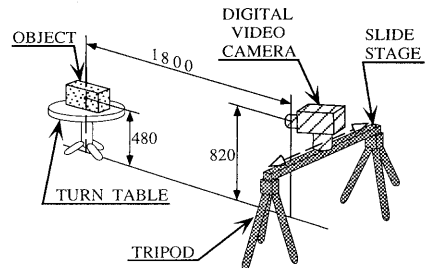


Fig.6 Experimental apparatus  
 図6 実験装置の構成

表1 使用したパラメータ値一覧

Table 1 Parameter Table

$\gamma_1$	0.70
$\gamma_2$	1.30
$\gamma_3$	0.20
$\gamma_4$	7pixel

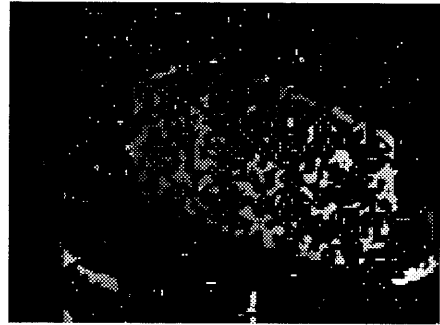
実験に用いた各パラメータを表1に示す。これらの値は一般性を失わない範囲でチューニングされた値である。

最小のウィンドウサイズは3画素とした。これは、3画素がウィンドウを形成する最小の大きさであるからである。また、左画像でウィンドウの自己相関をとり、画像全体で自己相関グラフの評価値を平均すると、ウィンドウサイズがある大きさ以上で平均評価値が飽和する。今回はこの飽和レベルの65%となるウィンドウサイズを最大ウィンドウサイズとした。この値は本画像では17画素であった。また、ウィンドウの画素単位での中心対称性を保つため、異なるウィンドウサイズ間のピッチは2画素とした。

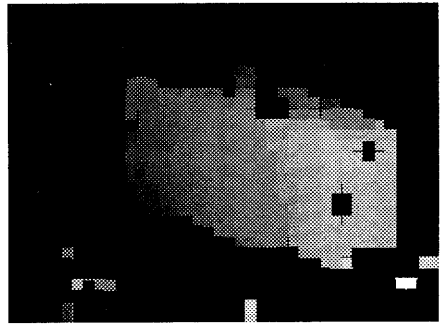
対応付けの結果得られた視差画像を図7に示す。(a) (b) はそれぞれウィンドウサイズを5画素および15画素で固定した結果である。(c) は本手法を適応した結果である。いずれも、明度が高いほど視差が大きい。対象物体の視差は最大で41.75画素、最小で35.75画素である。

対象物体領域内の対応点数を表2に示す。表2より、複数のウィンドウサイズのプレーンから評価値最大のプレーンを選ぶ方式が最も対応点数が多いことが分かる。

図7 (c) の視差画像を距離画像に変換した結果を図8にしめす。また、パターン光投影法によるアクティブレンジファインダ(液晶レンジファインダ)<sup>[10]</sup>を用いて対象物体の三次元形状を計測し、距離画像を得た。得られた距離画像を図9に示す。図8と図9を比べると、本手法では対象物体の概略三次元形状が得られていることが分かる。



(a)Fixed 5 pixels



(b)Fixed 15 pixels



(c)Proposed method

Fig.7 The result disparity images  
図7 結果の視差画像

表2 対応点数と対応率の結果

Table 2 Corresponding Points and Correct Ratio

image	corresponding points	correct ratio
(a)	3175	81.9%
(b)	16204	87.8%
(c)	16882	85.7%
original	20550	100%

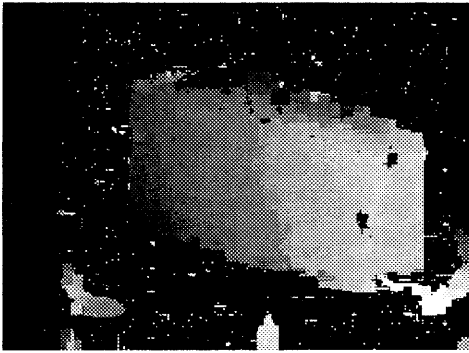


Fig.8 The depth image by stereo matching  
図8 ステレオ対応付けによる距離画像

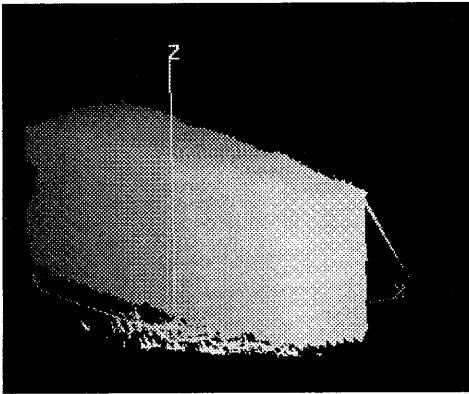


Fig.9 The depth image by active range finder  
図9 液晶レンジファインダによる距離画像

## 6 まとめ

適応ウィンドウ法に基づく石像計測のための領域ベースのステレオ対応付け手法を提案した。石像の表面に最も適したウィンドウ形状を相関グラフの評価値から決定する。さらに多段階のウィンドウサイズで対応付けを行い、対応率を上げる。

実際の石のステレオ画像を取得し、本手法を適用した。その結果、相関グラフの評価値の高いウィンドウサイズを選ぶ手法がウィンドウサイズを固定する手法に比べて有効であることが分かった。

今後は、実際の石像画像から、石像の種類とウィンドウ形状および相関グラフ形状の関係を明確化する。そして、汎用性の高いしきい値を決定し、石像の3次元計測データを作成する。

## 参考文献

- [1] K. Sato, H. Yamamoto and S. Inokuchi, "3-D Shape Measurement of Megalithic Statue Moai", Proc. of 8th ICPR, pp.675-677 (1986)
- [2] Y. Shirai, "Three-Dimensional Computer Vision", Springer-Verlag (1987)
- [3] R.A.Jarvis, "A Laser Time-of-Flight Range Scanner for Robotic Vision", PAMI, Vol.5, No.5, pp.505-512 (1983)
- [4] K. Sato, A. Yokoyama, and S. Inokuchi, "Silicon Range Finder -- A Realtime Range Finding VLSI Sensor --", Proc. IEEE Custom Integrated Circuits Conference, pp.339-342 (1994)
- [5] W. Eric, L. Grimson, "Computational Experiments with a Feature Based Stereo Algorithm", PAMI, Vol.PAMI-7, No.1, pp.17-34 (1985)
- [6] D. Marr, "Vision", W. H. Freedman and Company, New York (1987)
- [7] Y. C. Kim, J. K. Aggarwal, "Positioning Three-Dimensional objects Using Stereo", IEEE Journal of Robotics and Automation, Vol.RA-3, No.4, pp.362-373 (1987)
- [8] M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo", PAMI, Vol.15, No.4, pp.353-363 (1993)
- [9] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window : Theory and Experiment", PAMI, Vol.16, No.9, pp.920-932 (1995)
- [10] K.Sato and S. Inokuchi, "Range-Image System Utilizing Nematic Liquid Crystal Mask", Proc. of 1st International Conference on Computer Vision, 657-661 (1987)