

## 複数視点画像からの3次元形状復元

— 局所アフィン変形ウィンドウによるステレオマッチング —

服部 寛

(株) 東芝 関西研究所

〒658 神戸市東灘区本山南町 8-6-26

Tel.078-435-3105 kan@krl.toshiba.co.jp

複数視点画像から物体の3次元形状を復元する問題について述べる。密な3次元形状を復元するためには、領域ベースのステレオマッチングが不可欠である。従来の領域ベースのステレオマッチングは、対応点の周囲の輝度パターンは画像間で変形しないという仮定に基づいている。しかし、物体の局所勾配によって輝度パターンは画像間で変形し、この変形が対応精度を悪化させる原因となる。本稿では、奥行きと局所勾配を同時に求めるステレオマッチングアルゴリズムを提案する。局所勾配に応じてウィンドウをアフィン変形させ、対応づけを行なう。局所勾配は対応点の周囲の輝度勾配から直接計算する。実画像を用いた実験により、提案アルゴリズムが従来のアルゴリズムよりも高精度な形状復元が可能であり、また、様々な画像に対して適用可能であることを示す。

## 3D Shape Recovery from Multiple Images

— Stereo Matching Algorithm with Locally Affine-Deformable Window —

Hiroshi Hattori

Toshiba Kansai Research Laboratories

6-26, Motoyama-Minami-cho, 8-chome, Higashinada-ku, Kobe 658

This paper addresses the problem of recovering a 3D shape from multiple images. Area-based stereo matching is an essential technique in order to recover the dense 3D shape. Conventionally, area-based stereo matching uses the implicit assumption that intensity patterns surrounding corresponding points have no deformations between images from different views. In practice, however, the surface orientations locally deform the intensity patterns and these deformations often cause poor estimation of the 3D shape.

We propose a new stereo matching algorithm which computes the depth and surface orientation simultaneously. The proposed algorithm locally deforms a window according to the surface orientation which is directly computed from intensity gradients within the window. Experiments demonstrate that the proposed algorithm enables more precise recovery of the 3D shape than do conventional algorithms, and is applicable to various images.

## 1 はじめに

対象物の3次元形状復元は、様々な分野への適用が期待できるコンピュータビジョンの中心的課題である。2次元の画像から3次元の形状情報を復元するためには、互いに異なる視点から対象物を撮影した複数枚の画像を用いるのが有効である。この際に問題となるのは、3次元空間中で同一となる、異なる画像上の点を対応づけることである。この対応づけは、複数画像の撮影位置・撮影方向等に関するパラメータが未知である場合には2次元の探索問題となる。しかし、既知である場合には、これらのパラメータを拘束として用いると、奥行きのみ未知とする1次元の探索問題に帰着し、よりロバストな形状復元が可能となる。この拘束はエピポーラ拘束と呼ばれ、エピポーラ拘束を用いる複数画像間の対応づけを、ステレオマッチングと呼ぶ。

ステレオマッチングは、比較的古くから研究されているテーマであり、これまでに多くのアルゴリズムが提案されている[1, 2, 3]。これらは特徴ベースと領域ベースの2つに大きく分類することができる。特徴ベースのステレオマッチングは、エッジ、線分、コーナ等の特徴を各画像から抽出し、これらの特徴を対応づける。この方法は、計算コストが比較的小さく、対象によってはロバストな形状復元が可能であるが、抽出した特徴同士を対応づけるため、得られるのは基本的に疎な奥行き情報である。一方、領域ベースのステレオマッチングは、画像上の小領域間の輝度の相関等をとることによって、対応領域を求める。この方法は特徴抽出を行わないので、密な形状復元が可能である。本稿では、密な形状復元を行なうという目的から、領域ベースのステレオマッチングについて述べる。

領域ベースのステレオマッチングは、通常、同一の正方形のウィンドウを用いて、異なる画像上の小領域間の類似度、あるいは、非類似度を評価して、対応づけを行なう。各画像において同一の正方形のウィンドウを用いるということは、物体表面は局所的に画像面と平行であり、対応領域の輝度パターンは画像間で変形しないと仮定していることを意味する。しかし、一般には、物体表面は任意の局所勾配(傾き)を持つため、その局所勾配によって対応領域の輝度パターンは変形する<sup>1</sup>。十分小さいウィンドウであれば、この変形は無視できるが、ウィンドウが小さいと、対応候補が多くなるため真の対応点を特定するのが困難となる。したがって、対応の曖昧性を少なくするために、ある程度大きなウィンドウを用いる必要があるが、そのようなウィンドウでは、局所勾配による対応領域の輝度パターンの変形が対応精度低下の原因となる。

<sup>1</sup> 第3章で詳述するが、画像間の幾何学的関係によっても輝度パターンは変形する。しかし、ステレオマッチングにおいては、この幾何学的関係は既知であり、その変形は容易に補正可能である。

本稿では、奥行きと局所勾配を同時に求める新しいステレオマッチングアルゴリズムを提案する。本アルゴリズムにおいては、物体表面を局所的に任意勾配の平面で近似し、ウィンドウの形状をその局所勾配によって適応的に変化させ、対応点を求める。この際、局所勾配は輝度勾配を用いて直接的に計算する。これにより、局所勾配に依らない、従来よりも精度の高い対応づけが可能となる。実画像を用いた実験により、本アルゴリズムによって従来よりも精度の高い形状復元が可能となること、ならびに、本アルゴリズムが様々な画像に対して適用可能であることを示す。

## 2 従来研究

一般に、異なる画像上の対応領域の輝度パターンは局所勾配により変形する。領域ベースのステレオマッチングにおいて、この変形による対応精度の低下を防ぐためには、少なくとも以下の2つのアプローチが考えられる。

1つは、奥行き変動に応じてウィンドウのサイズを局所的に変化させるアプローチである。Okutomiらは、局所的な輝度変動と視差(すなわち奥行き)変動を評価することによって、画像の場所ごとに適当なウィンドウサイズを選択する適応型ウィンドウによるステレオマッチングアルゴリズムを提案した[4]。しかし、このアルゴリズムでは、画像上のある点の視差を計算するための適当なウィンドウサイズを決める際に、その点の周囲の点の視差を用いるので、これらの大部分が最初に正しく与えられなければならない。

もう1つは、奥行きのみならず、局所勾配も求めるアプローチである。本稿で提案するアルゴリズムもこちらのアプローチに属する。ステレオ視により、物体表面の局所勾配を計算する研究には以下のようなものがある。Gårdingは、重み付き2次モーメント行列(Windowed second moment matrix)を用いて、輝度パターンの局所変形の評価を行なった[5]。この方法によると、左右画像の輝度勾配の局所的な統計量から直接的に物体表面の局所勾配を求めることができる。Jonesらは、複数の異なる空間フィルタを用いて、局所勾配を求めた[6]。しかし、これらの方法は、いずれも対応づけを行なった後、いかにして画像から局所勾配を求めるかという問題を扱っており、対応づけの問題を取り上げていない。Devernayは、Enhanced correlation methodを提案し、視差と視差勾配、つまり、奥行きと局所勾配を画像データから直接求めた[7, 8]。この方法は、我々のアプローチと枠組を共有するが、局所勾配を計算するために反復計算を必要とし、その計算方法も明らかでない。

本稿で提案するステレオマッチングは、画像上の各点の奥行きと局所勾配を同時に求める。ウィンドウは局所勾配に応じて動的に変形させる。局所勾配は輝度勾配から直接計算するので、比較的小さい計算コストで局所勾

配による対応領域の変形を考慮して、対応点を求めることができる。さらに、上述の従来のアルゴリズムが2枚の画像のみを用いているのに対し、提案アルゴリズムでは、3枚の画像を用いて対応の曖昧性を減らし、よりロバストな形状復元を行なう。

### 3 画像の局所変形

左右カメラが平行でない、一般的なステレオカメラの配置を考える。この場合、対応点周囲の輝度パターンは、左右画像の幾何学的関係と物体表面の局所勾配によって変形する。この変形は $2 \times 2$ のアフィン変形行列で近似することができる。本章では、この変形について解析する。

図1に示すようなステレオカメラを考える。ワールド座標系の原点を第1のカメラの視点にとり、 $X, Y, Z$ 軸を図のように設定する。 $X = (X, Y, Z)^T$ と $X + \Delta X = (X + \Delta X, Y + \Delta Y, Z + \Delta Z)^T$ を3次元平面 $\pi$ 上の点とする。点 $X, X + \Delta X$ の第1の画像への投影点を各々、 $x = (x, y)^T$ ,  $x + \xi = (x + \xi, y + \eta)^T$ とすると、

$$x = \frac{F}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad x + \xi = \frac{F}{Z + \Delta Z} \begin{bmatrix} X + \Delta X \\ Y + \Delta Y \end{bmatrix} \quad (1)$$

という関係式が成り立つ。 $\xi$ について解くと、

$$\xi = \frac{F}{Z} \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix}. \quad (2)$$

ここで $F$ は焦点距離であり、 $|\Delta Z| \ll Z$ と仮定した。

同様に、点 $X$ と $X + \Delta X$ の第2の画像への投影点を各々、 $x' = (x', y')^T$ ,  $x' + \xi' = (x' + \xi', y' + \eta')^T$ とし、ワールド座標系に対する第2のカメラの姿勢 $M_1 = (m_{11}, m_{12}, m_{13})^T$ ,  $M_2 = (m_{21}, m_{22}, m_{23})^T$ を図1のように定義すると、

$$x' = \frac{F}{Z'} \begin{bmatrix} M_1^T (X - t) \\ M_2^T (X - t) \end{bmatrix}, \quad (3)$$

$$x' + \xi' = \frac{F}{Z' + \Delta Z'} \begin{bmatrix} M_1^T (X + \Delta X - t) \\ M_2^T (X + \Delta X - t) \end{bmatrix} \quad (4)$$

という関係式が成り立つ。ここで、 $t = (t_x, t_y, t_z)^T$ は第2のカメラの視点のワールド座標系に対する3次元位置を表し、 $Z'$ と $Z' + \Delta Z'$ は、各々、第2のカメラに対する $X$ と $X + \Delta X$ の奥行きである。式(3)と(4)を連立させて $\xi'$ について解くと、

$$\xi' = \frac{F}{Z'} M \Delta X \quad (5)$$

を得る。ここで、 $M = (M_1^T, M_2^T)^T$ であり、 $|\Delta Z'| \ll Z'$ と仮定した。3次元平面 $\pi$ の方程式を $Z = pX + qY + r$ とおくと、 $X$ と $X + \Delta X$ はこの平面上にあるので、

$$\Delta Z = p\Delta X + q\Delta Y \quad (6)$$

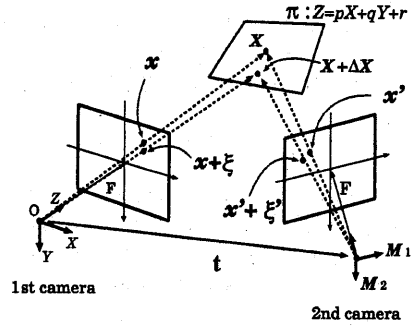


図1: 座標系の定義

となる。式(2)と(6)から、式(5)の $\Delta X$ を消去すると、

$$\begin{aligned} \xi' &= \frac{Z}{Z'} \begin{bmatrix} m_{11} + pm_{13} & m_{12} + qm_{13} \\ m_{21} + pm_{23} & m_{22} + qm_{23} \end{bmatrix} \xi \\ &= (A_c + A_v)\xi = A\xi \end{aligned} \quad (7)$$

となる。ここで、

$$A_c = \frac{Z}{Z'} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \quad A_v = \frac{Z}{Z'} \begin{bmatrix} pm_{13} & qm_{13} \\ pm_{23} & qm_{23} \end{bmatrix},$$

$A = A_c + A_v$ である。式(7)は、図2に示すような対応点周囲の輝度パターンの2枚の画像間の変形を表す。提案アルゴリズムでは、ウィンドウをこの式によって変形させる。行列 $A_c$ は、第2のカメラの姿勢 $M$ と、スケールファクタ $s = Z/Z'$ から計算できる。 $M$ は既知であり、 $s$ も奥行き探索とともに求めることができるので、 $A_c$ は容易に決定できる。一方、行列 $A_v$ は、画像の場所ごとに変化するパラメータ、すなわち、物体表面の局所勾配を含むため、式(7)を用いてウィンドウを変形させるためには、局所勾配を評価する必要がある。行列 $A_v$ は通常、物体表面は局所的には画像面と平行であるという仮定( $p = q = 0$ , すなわち $A_v = O$ )に基づいて無視されるが、提案アルゴリズムでは、 $A_v$ を考慮してウィンドウを変形させる。

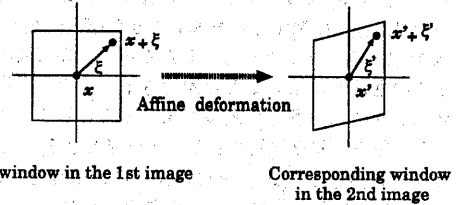


図2: ウィンドウの変形

### 4 局所勾配の直接計算

奥行き探索のように局所勾配も探索により求めると、計算コストは非常に大きくなる。本章では、対応点周囲

の輝度勾配から直接、局所勾配を求める方法について述べる。これは、オプティカルフローの計算法の一つである勾配法 [9, 10] と類似した方法である。本方法では、以下のような対応領域の輝度の線形変化を仮定する。

$$af(x_1 + \xi) + b = g(x_2 + \xi'), \quad (8)$$

ここで、 $f(x)$ 、 $g(x)$  は画像上の点  $x = (x, y)^T$  における輝度である。また、 $x_1$  と  $x_2$  は対応点であり、 $x_1 + \xi$  と  $x_2 + \xi'$  は各々、 $x_1$  と  $x_2$  の周囲のウィンドウ内の対応点である。パラメータ  $a$  は各画像のコントラストの相違を考慮するためのスケールファクタであり、 $b$  は輝度のバイアス変化に対するパラメータである。パラメータ  $a$  と  $b$  は、画像の場所ごとに異なる。勾配法によってフローを計算する際には、通常、輝度不変の仮定 (brightness constancy assumption) を用いる。これは、式 (8) で、 $a = 1$ 、 $b = 0$  としたものである。したがって、式 (8) は輝度不変の仮定より一般的な仮定を表している。この拡張は、連続画像と比べると、対応点の位置のズレが大きいステレオ画像にとっては重要である。

式 (8) の右辺に式 (7) を代入してテイラー展開すると、

$$\begin{aligned} g &= g(x_2 + A_c \xi + A_v \xi') \\ &= g(x_2 + A_c \xi) + \nabla g(x_2 + A_c \xi) A_v \xi', \end{aligned} \quad (9)$$

ここで、 $\nabla g = (g_x, g_y)$  は輝度勾配であり、 $A_v \xi$  の 2 次以上の項は無視した。

この展開を用いると、式 (8) は以下のように整理することができる。

$$[\Gamma(\xi)\xi \quad \Gamma(\xi)\eta \quad F(\xi) \quad -1] \alpha = G(\xi), \quad (10)$$

ここで、 $F(\xi) = -f(x_1 + \xi)$ 、 $G(\xi) = -g(x_2 + A_c \xi)$ 、 $\alpha = (p, q, a, b)^T$  であり、さらに、

$$\Gamma(\xi) = s\{m_{13}g_x(x_2 + A_c \xi) + m_{23}g_y(x_2 + A_c \xi)\}$$

である。式 (10) はウィンドウ内の任意の点について成り立つので、各点に関する式を連立させれば、未知パラメータ  $\alpha$  を求めることができる。したがって、局所勾配は最低 2 枚の画像から求めることができる。しかし、奥行き計算と同様に、多くの画像を用いることによって、ロバストに局所勾配を評価することが可能となる。

3 枚目の画像を  $h(x)$ 、その画像上の対応点を  $x_3$  とする。同様に対応領域の輝度の線形変化を仮定し、

$$a'f(x_1 + \xi) + b' = h(x_3 + \xi''). \quad (11)$$

同様の式変形を行なうと、

$$[\Gamma'(\xi)\xi \quad \Gamma'(\xi)\eta \quad F(\xi) \quad -1] \beta = H(\xi) \quad (12)$$

を得る。ただし、 $H(\xi) = -h(x_3 + A'_c \xi)$ 、 $\beta = (p, q, a', b')^T$ 、

$$\Gamma'(\xi) = s'\{m'_{13}h_x(x_3 + A'_c \xi) + m'_{23}h_y(x_3 + A'_c \xi)\}$$

である<sup>2</sup>。ウィンドウ内の任意の点  $\xi = \xi_i (i = 1, 2, \dots, N)$  について、式 (10) と (12) が成り立つので、

$$\Phi \gamma = \Psi \quad (13)$$

を得る。ここで、 $\gamma = (p, q, a, b, a', b')^T$ 、

$$\Phi = \begin{bmatrix} \Gamma(\xi_1)\xi_1 & \Gamma(\xi_1)\eta_1 & F(\xi_1) & -1 & 0 & 0 \\ \Gamma'(\xi_1)\xi_1 & \Gamma'(\xi_1)\eta_1 & 0 & 0 & F(\xi_1) & -1 \\ \Gamma(\xi_2)\xi_2 & \Gamma(\xi_2)\eta_2 & F(\xi_2) & -1 & 0 & 0 \\ \Gamma'(\xi_2)\xi_2 & \Gamma'(\xi_2)\eta_2 & 0 & 0 & F(\xi_2) & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Gamma(\xi_N)\xi_N & \Gamma(\xi_N)\eta_N & F(\xi_N) & -1 & 0 & 0 \\ \Gamma'(\xi_N)\xi_N & \Gamma'(\xi_N)\eta_N & 0 & 0 & F(\xi_N) & -1 \end{bmatrix},$$

$$\Psi = [G(\xi_1) \ H(\xi_1) \ G(\xi_2) \ H(\xi_2) \ \dots \ G(\xi_N) \ H(\xi_N)]^T.$$

式 (13) を  $\gamma$  について解くと、局所勾配  $(p, q)$  を得る。この方法は、対応点の周囲の輝度情報のみを用いており、基本的に繰り返し計算は不要であるので、少ない計算量で局所勾配に応じてウィンドウ形状を変化させながら、対応点を探索することを可能にする。この局所勾配に応じて変形するウィンドウを、局所アフィン変形ウィンドウ (Locally affine-deformable window) と呼ぶ。

## 5 局所アフィン変形ウィンドウによるステレオマッチング

本章では、これまでの解析に基づいて、本稿で提案するステレオマッチングアルゴリズムを示す。最低 2 枚の画像から適用可能であるが、第 4 章でも述べたように、ロバストな形状復元を行なうために 3 枚の画像を用いる。

異なる視点から得られた複数の画像から形状復元を行なうためには、複数画像間の幾何学的関係、すなわち、複数画像の互いの撮影位置、撮影方向等の関係を求める必要がある。この幾何学的関係は、画像情報のみから求めることができる。これは Structure-from-motion の問題であり、以下のような手順で解くことができる。まず、各画像上でコーナを検出し、コーナの対応づけを行なう。次に、これらの対応情報から幾何学的関係を得る。容易に幾何学的関係を得るために、ここではカメラモデルとして正射影モデルを仮定する。正射影モデルの下では、コーナの対応情報が正しく得られれば、行列の特異値分解 [11] を用いて幾何学的関係を復元することができる。しかし、この段階での対応づけは、2 次元探索であり、誤対応が含まれることが多い。この誤対応が幾何学的関

<sup>2</sup>ここで、 $s'$  等の ' (プライム) は、そのパラメータが第 3 の画像  $h(x)$  に関するものであることを示す。

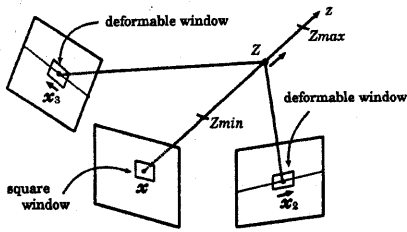


図 3: 3 枚の画像を用いた 3 次元形状復元

係の復元精度低下の原因となるため、RANSAC法を用いて、誤対応を除去し、幾何学的関係を復元する [12].

複数画像間の幾何学的関係が求まると、この関係を拘束として用いる 3 次元形状復元、すなわち、ステレオマッチングが可能となる。以下に、本稿で提案する、局所アフィン変形ウィンドウによるステレオマッチングを示す。

### 3D Shape Recovery Algorithm

1. 基準画像上の点  $x$  の奥行き  $Z_0$  の初期推定値を求める<sup>3</sup>。  $Z_0$  は、ウィンドウ内の奥行きが一樣であるという仮定 ( $p = q = 0$ ) のもとで、従来のステレオマッチングにより計算する。
2.  $x$  の奥行きを  $Z_0$  として、式 (13) より、  $x$  の局所勾配  $(p, q)$  を求める。カメラモデルとして正射影モデルを仮定したので、スケールファクタは無視することできる ( $s = s' = 1$ )。
3.  $x$  の奥行きを  $Z_0$  の周囲の  $Z$  と仮定し、  $Z$  と  $(p, q)$  を式 (7) に代入して、局所アフィン変形行列  $A(j)$  ( $j = 2, 3$ ) を計算し、奥行き  $Z$  の妥当性  $C(Z)$  を計算する。  $C(Z)$  については後述する。
4. 探索範囲 ( $Z_0 - \Delta Z < Z < Z_0 + \Delta Z$ ) 内の各々の奥行き  $Z$  に対し、  $C(Z)$  を計算する。基本的に  $C(Z)$  が最も高くなる  $Z^*$  を  $x$  の真の奥行きとする。

奥行き  $Z$  の妥当性  $C(Z)$  は、ウィンドウの変形を考慮した 2 つの正規化相互相関の和として定義する。すなわち、

$$C_1 = \frac{1}{N} \sum \frac{\{f(x + \xi) - \mu\}\{g(x_2 + \xi') - \mu'\}}{\sigma \times \sigma'},$$

$$C_2 = \frac{1}{N} \sum \frac{\{f(x + \xi) - \mu\}\{h(x_3 + \xi'') - \mu''\}}{\sigma \times \sigma''},$$

$$C(Z) = C_1 + C_2, \quad (14)$$

ここで、  $\mu, \mu', \mu''$  は各々、ウィンドウ内の輝度の平均であり、  $\sigma, \sigma', \sigma''$  は各々、ウィンドウ内の輝度の標準偏差である。  $N$  はウィンドウ内の画素数である。通常、正規

<sup>3</sup>厳密にいうと、この  $Z$  はアフィンカメラモデル (正射影モデルを包含するモデル) の下では必ずしも奥行きではない。しかし、  $Z$  は奥行きと 1 対 1 に対応するので、ここでは「奥行き」と呼ぶ。

化相互相関は、  $\xi = \xi' = \xi''$  として計算する。これに対し、提案アルゴリズムでは、輝度パターンのアフィン変形を考慮するために、画像ごとに異なるベクトル  $\xi, \xi', \xi''$  を用いる。

## 6 実験

本章では、提案アルゴリズムを実画像に適用した実験結果を示す。

### 6.1 従来法との復元精度の比較

まず、提案アルゴリズムと従来のアルゴリズムの性能を比較する実験を行なった。詳細な誤差評価を行なうために、図 4 に示すような既知形状物体を用いた。これは、野球の軟式ボールであり、半径は約 3.5cm である。この物体から約 50cm 離れた位置から 25mm のレンズを装着した CCD カメラを用いて、  $480 \times 480$  画素の画像を 3 枚撮影した。中央の画像が基準画像であり、この画像において、縦、横、5 画素おきに奥行き (提案アルゴリズムでは局所勾配も) を計算した。真値は、この物体が完全な球体であるとして求めた。比較対象の従来法としては、以下の 2 つを用いた。

1. 局所勾配によるウィンドウの変形を考慮しない。これを「従来法 1」と呼ぶ。
2. 奥行き  $Z$  の初期値から局所勾配を求め、それをもとにウィンドウを変形させる。これを「従来法 2」と呼ぶ。

図 5(a) に様々なウィンドウサイズで復元した奥行きの平均誤差を示す。縦軸が物体のサイズ、すなわち、半径  $R$  で正規化した平均誤差であり、横軸が対応づけに用いたウィンドウのサイズである。  $3 \times 3$  や  $5 \times 5$  のような小さいウィンドウでは、ウィンドウ内に対応づけに十分な輝度の変動がないので、いずれの方法でも誤差が大きいことがわかる。このような小さなサイズのウィンドウでは、提案アルゴリズムは従来よりも大きな誤差を含む。これは、小さなサイズのウィンドウでは局所勾配を精度良く計算できないことに起因する。しかしながら、従来のアルゴリズムの誤差も大きく、精度よく形状を復元するためには、より大きなウィンドウを用いることが不可欠である。  $7 \times 7$  以上のウィンドウでは、対応の曖昧性が減少するため、いずれの方法でも誤差が小さくなる。しかし、

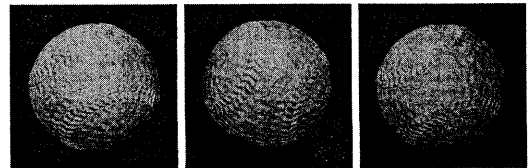
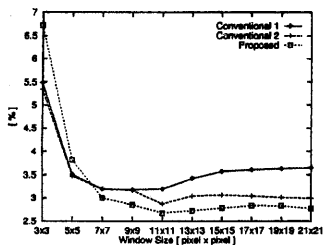
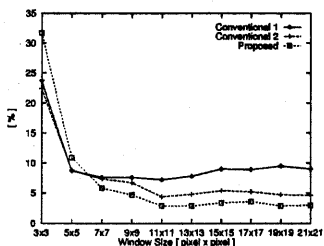


図 4: 入力画像「ボール」 ( $480 \times 480$ )



(a) 平均誤差



(b) 誤対応の割合

図 5: 精度の比較

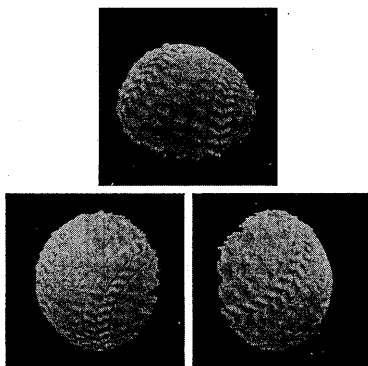


図 6: “ボール”の復元形状

従来法 1 の局所勾配による対応領域の変形を考慮しない場合は、局所勾配によってウィンドウ形状を変化させる他の 2 つの方法と比べると誤差が大きい。また、局所勾配を、奥行き初期値から計算する場合 (従来法 2) と、画像情報から直接求める場合 (提案法) とでは、後者の方が誤差が小さい。これは、奥行き初期値から局所勾配を計算すると、奥行き初期値に含まれる誤差に直接影響を受けるためである。11 × 11 のウィンドウの提案アルゴリズムの平均誤差が最も小さく、最小値は 2.7% であり、これは約 0.95mm に相当する。

図 5(b) に誤対応 (大きな対応誤り) の割合を示す。ここでは、奥行き初期値の 15 画素 (約 3mm) 以上のずれを誤対応と定義した。平均誤差の場合と類似した傾向を持ち、7 × 7 以上のいずれのサイズのウィンドウでも提案アルゴリズム

が最も誤対応が少ない。この場合も 11 × 11 の提案アルゴリズムが最も成績が良く、最小値は 1.9% である。図 6 に 11 × 11 の提案アルゴリズムによる復元形状を示す。

## 6.2 様々な画像への適用

提案アルゴリズムを顔、手、山といった様々な画像に対し適用して得られた結果を以下に示す。図 7~15 に入力画像と復元形状を示す。顔と手 1, 2 のカメラとの距離は約 50cm であり、山は約 1,500m 離れた位置から撮影した。図 9 には、顔の鼻周辺部の復元形状を示す。(b) が、局所勾配によってウィンドウの形状を変化させない従来法による復元形状であり、(c) が提案法による復元形状である。

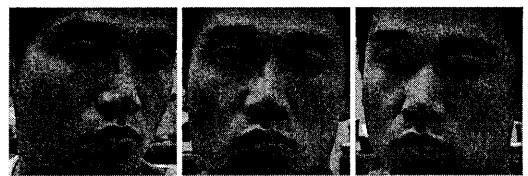


図 7: 入力画像 “顔”

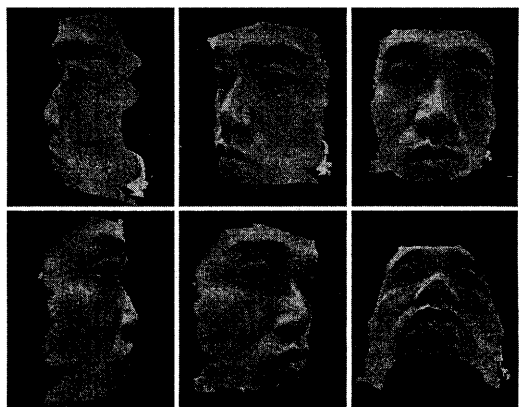
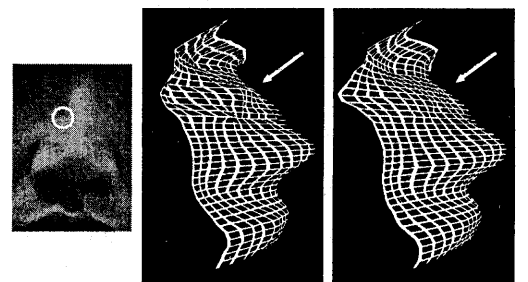


図 8: “顔”の復元形状



(a) (b) 従来法 (c) 提案法

図 9: “顔”の鼻周辺部 (a) の復元形状の比較 (奥行き方向のスケールを 2 倍にして表示)



図 10: 入力画像 “手 1”

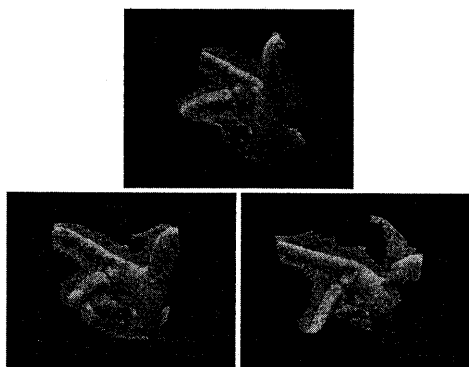


図 11: “手 1” の復元形状



図 12: 入力画像 “手 2”

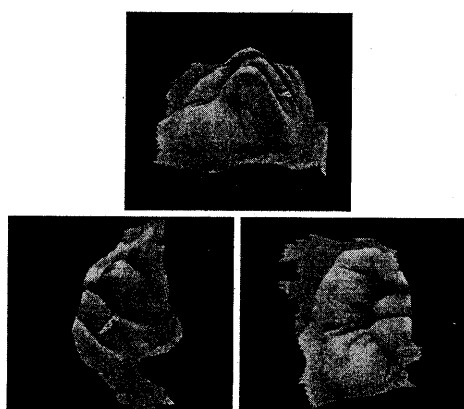


図 13: “手 2” の復元形状

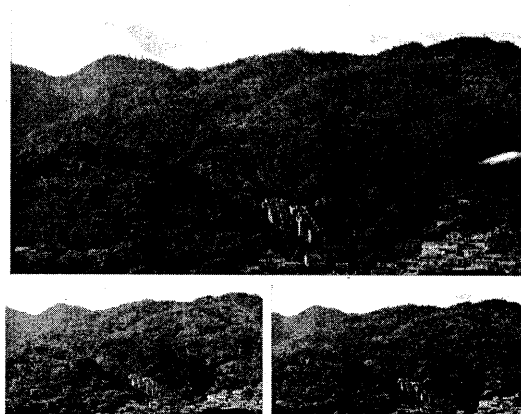


図 14: 入力画像 “山” ( 基準画像のみ大きく表示 )



(a) Front view

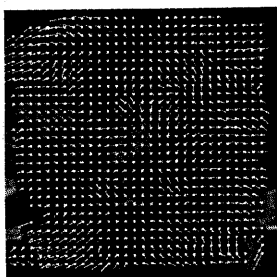


(b) Side view

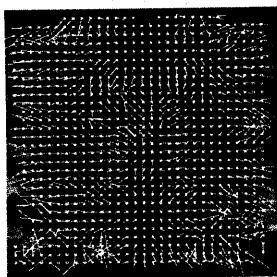
図 15: “山” の復元形状

顔や手の画像はテクスチャが少なく、山の画像は大部分の領域で繰り返しパターンが存在するため、これらの画像からの形状復元は容易ではない。しかしながら、精度の高い形状復元を実現しており、これらの結果は、提案アルゴリズムが多くの画像に対して適用可能であることを示す。また、図9(a)の円付近の鼻の側面の復元形状(矢印)を比較すると、従来法による復元形状には不自然な凹凸が存在するのに対し、提案法ではより正確な形状が復元されていることがわかる。

提案アルゴリズムは、形状のみならず、局所勾配を出力する。図16(a)に提案アルゴリズムによって得た、“顔”の局所勾配を示す。また、比較のため、図16(b)に、奥行きマップに局所的に平面を当てはめることによって計算した局所勾配を示す。これは、第6.1節の従来法2で用いた局所勾配と同一のものである。奥行きマップから得る方法は、奥行き推定結果に含まれる誤差の影響を直接受けるため、得られた結果はノイズが多い。一方、提案アルゴリズムにおいては、画像から直接、局所勾配を得るため、より正確に計算することができる。この結果は、提案アルゴリズムが局所勾配を得るという観点からも適切であることを示している。



(a) 画像から



(b) 奥行きマップから

図 16: “顔”の局所勾配

## 7 まとめ

本稿では、奥行きと局所勾配を同時に求める新しいステレオマッチングアルゴリズムを提案した。従来のアルゴリズムが、物体表面を局所的に画像面と平行な面で近

似するのに対し、提案アルゴリズムにおいては、物体表面を任意勾配の平面で近似する。その近似平面の勾配を輝度勾配から直接的に計算し、求めた局所勾配に応じて、ウィンドウを適応的に変形させて、対応づけを行なう。実画像を用いた実験により、提案アルゴリズムの、奥行きのみを求める従来のステレオマッチングに対する有効性を示し、さらに、様々な画像に対して適用可能であることを示した。今後の課題としては、階層的なアプローチを用いて効率性を上げること考えている。また、より一般的な投影モデル下での Structure-from-motion の手法[13, 14]と今回提案したアルゴリズムを結合することを検討している。この拡張により、シーンや画像に対する拘束の少ない3次元形状復元が可能になると考えている。

## 参考文献

- [1] U.R.Dhond, J.K.Agarwal : “Structure from Stereo - A Review”, IEEE Trans. Systems, Man Cybern., vol.19, No.6, pp.1489-1510, 1989.
- [2] O.D.Faugeras : “Three-Dimensional Computer Vision : A Geometric Viewpoint”, MIT Press, 1993.
- [3] 奥富 : “CVCV-WG 特別報告: コンピュータビジョンにおける技術評論と将来展望 (X) —ステレオ視 (Stereo Vision) —”, 情報処理学会研究報告, 96-CVIM-102, pp.57-64, 1996.
- [4] M.Okutomi, T.Kanade : “A Locally Adaptive Window for Signal Matching”, *International Journal of Computer Vision*, 7:2, pp.143-162, 1992.
- [5] J.Gårding, T.Lindeberg : “Direct Estimation of Local Surface Shape in a Fixating Binocular Vision System”, *Proc. 3rd European Conference on Computer Vision*, pp.365-376, 1994.
- [6] D.G.Jones, J.Malik : “Determining Three-Dimensional Shape from Orientation and Spatial Frequency Disparities”, *Proc. 2nd European Conference on Computer Vision*, pp.661-669, 1992.
- [7] F.Devernay, O.D.Faugeras : “Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models”, Technical Report 2304, INRIA, 1994.
- [8] F.Devernay, O.D.Faugeras : “Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models”, *Proc. Computer Vision and Pattern Recognition*, pp.208-213, 1994.
- [9] B.Horn, B.Schunck : “Determining Optical Flow”, *Artificial Intelligence*, 17, pp.185-203, 1981.
- [10] J.Weber, J.Malik : “Robust Computation of Optical Flow in a Multi-Scale Differential Framework”, *International Journal of Computer Vision*, 14, pp.67-81, 1995.
- [11] C.Tomasi, T.Kanade : “Shape and Motion from Image Streams under Orthography : a Factorization Method”, *International Journal of Computer Vision*, 9:2, pp.137-154, 1992.
- [12] C.S.Wiles, A.Maki, N.Matsuda, M.Watanabe : “Hyper-Patches for 3D Model Acquisition and Tracking”, *Proc. Computer Vision and Pattern Recognition*, pp.1074-1080, 1997.
- [13] C.J.Poelman : “The Paraperspective and Projective Factorization Methods for Recovering Shape and Motion”, Technical Report CMU-CS-95-173, Carnegie Mellon University, 1995.
- [14] B.Triggs : “Factorization Methods for Projective Structure and Motion”, *Proc. Computer Vision and Pattern Recognition*, pp.845-851, 1996.