

物体表面の勾配を考慮したステレオマッチング

服 部 寛†

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が画像間で変形しないと仮定し、同一の正方形のウィンドウで異なる画像上の小領域間の類似度を計算する。しかし、実際には物体表面の局所勾配によって対応領域の輝度パターンは変形し、この変形が対応精度を悪化させる原因となる。本稿では、物体表面の局所勾配を考慮したステレオマッチング法を提案する。提案法では、各点の局所勾配に応じてウィンドウをアフィン変形させながら対応づけを行う。この際、計算コストの増大を回避するため、対応点の周囲の輝度勾配から局所勾配を直接計算する。実画像を用いた実験により、提案法が従来方式に比べて高精度な形状復元を実現し、また、様々な画像に対して適用可能であることを示す。

Stereo Matching with Direct Surface Orientation Recovery

HIROSHI HATTORI†

We propose a new stereo matching algorithm which computes the depth and surface orientation simultaneously. While area-based stereo matching is an essential technique in the recovery of the dense 3D shape, conventionally it uses square windows based on the implicit assumption that intensity patterns surrounding corresponding points have no deformations between images from different views. In practice, however, the local surface orientations deform the intensity patterns and such deformations often give rise to poor estimation of the 3D shape. To solve this problem, we formulate a new algorithm that allows a matching window to locally deform according to the surface orientation, which we propose to compute directly from intensity gradients within the window. Through experiments we demonstrate that our algorithm indeed realizes more precise recovery of the 3D shape than do conventional ones while being applicable to various images.

1. はじめに

物体の3次元形状復元はロボット視覚、VR、CG等、様々な分野への適用が可能な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、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、求めるべき形状がユークリッド形状でなければならぬか、非ユークリッド形状⁴⁾で十分であるかはタスクによって異なるが、3次元空間中で同一である点を複数画像上で対応づけることは、あらゆるタスクに共通する課題である。この対応づけは、複数画像間に存在する幾何学的拘束、すなわち、エピポーラ拘束を利用すると奥行きを未知とする1次元の探索問題に帰着する。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、これまでに様々な方式が提案されているが^{3),21)}、これらは特徴ベース法と領域ベース法の2つに分類することができる。特徴ベース法は、エッジ、線分、角点等の特徴を各画像から抽出し、これらの特徴どうしを対応づける。この方法は計算コストが比較的低く、対象によってはロバストな形状復元が可能であるが、得られるのは基本的に疎な奥行き情報である。一方、領域ベース法は、画像上の小領域間の輝度の相関等から対応領域を求める。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うことを主眼に、領域ベースのステレオマッチングについて述べる。

領域ベース法では通常、同じサイズの正方形のウィンドウ（以下では正方ウィンドウと呼ぶ）を用いて異なる画像上の小領域間の類似度を評価して対応点を探索する。各画像において同一の正方ウィンドウを用い

† 株式会社東芝研究開発センター

TOSHIBA Corporate Research & Development Center

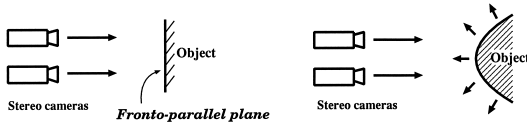


図1 従来法は物体表面と画像面が局所的に平行と暗黙のうちに仮定している(左)。しかし、実際には物体表面は任意の傾きを持つ(右)。

Fig.1 The conventional method is based on the implicit assumption that the 3D surface is locally fronto-parallel (left). In practice, however, any 3D surface has arbitrary surface orientations (right).

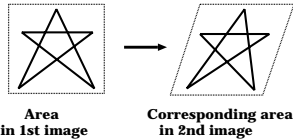


図2 対応領域のステレオ画像間の変形

Fig.2 Deformation between matching areas.

ということとは、対応点周囲の輝度パターンはステレオ画像間で変形しないと仮定していることを意味する。これは、いい換えれば、図1左に示すように“物体表面は局所的に画像面と平行な平面である”と仮定していることになる。しかし、実際には図1右のように物体表面は任意の勾配(傾き)を持つため、対応領域の輝度パターンは図2に示すように表面の局所勾配によって変形する。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、ある程度大きなサイズの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対応領域の変形が精度低下の原因となる。

本稿では、通常用いられる、物体表面が局所的に画像面と平行であるという仮定を一般化し、“物体表面は局所的に任意勾配の平面である”という仮定の下で対応づけを行う方式を提案する。従来の正方ウィンドウに基づく方式が奥行きのみを未知とするのに対し、提案法では奥行きと物体表面の勾配を未知とし、ウィンドウの形状を各点の表面勾配に応じて変形させながら対応づけを行う。この際、ウィンドウの変形をアフィン変形で近似し、輝度勾配から表面勾配を直接的に求めることにより計算コストの増大を回避する。このアフィン変形を許容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^{7),22)}

と呼ぶ。局所アフィン変形ウィンドウにより、表面の傾きによらず、従来よりも精度の高い対応づけが可能となる。以下に2章で対応領域の変形について言及した従来研究と本研究の相違について述べた後、3章でステレオ画像間の対応領域の変形を解析し、4章において物体表面の局所勾配を輝度勾配から直接計算する方式について述べる。これらに基づき、5章において新しいステレオマッチング法を提案し、6章で実験結果を示した後、7章において、まとめと今後の課題について述べる。

2. 従来研究との相違

物体表面の局所勾配による対応領域の変形がもたらす対応づけの精度低下を防ぐためには、以下の2つのアプローチが考えられる。1つは、ウィンドウの大きさを適応的に変化させ、間接的に変形に対処するアプローチである。Okutomiら¹²⁾は局所的な輝度変動と視差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法を提案した。しかし、この方式では対応領域の変形を探索ウィンドウを小さくすることによって扱わざるをえない。一般にウィンドウのサイズが小さくなると対応の曖昧性が増加するため、正しい対応点を特定するのが困難になるという問題がある。

もう1つは、奥行きのみならず表面勾配も求めるアプローチである。本稿で提案する方法もこちらのアプローチに属する。ステレオ視による物体表面の局所勾配の計算に関する研究には以下のものがある。Gårdingら⁶⁾は2次モーメント行列(windowed second moment matrix)により対応領域の局所変形を評価した。この方法によると左右画像の輝度勾配から直接的に視差勾配を求めることが可能である。Jonesら⁹⁾は複数の異なる空間フィルタを用いて視差勾配を求める方法を示した。Robertら¹⁶⁾は明示的な3次元再構成を行うことなく、視差勾配から平面領域を求める方法を提案している。しかし、これらの研究は対応づけの後、いかにして画像から直接的に視差勾配を求めるかという問題に言及しており、対応探索問題を扱っていない。Devernayら²⁾は視差勾配に応じてウィンドウ形状を変化させながら対応づけを行う enhanced correlation method を提案している。この方法は我々のアプローチと枠組みを共有するが、視差勾配の計算方法が明らかでない。Mainomeら¹⁰⁾は位相を用いたステレオ対応づけ手法において局所変形の問題に言及しているが、視差勾配を探索により求めているため計算コストが高

3章で詳述するが、画像間の幾何学的関係(カメラ姿勢の相違)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。

い。また、これら 5 つの従来研究^{2),6),9),10),16)}は対応領域の変形を 3 次元空間中の表面勾配（奥行き勾配）ではなく画像上の視差勾配をベースに定式化している。2 眼ステレオでは両者に相違はないが、3 枚以上の画像を入力とする場合、視差や視差勾配はステレオ画像対ごとに異なるため、これらの方式に対応の曖昧性を減少させるのに有効な多眼視^{13),20)}へ拡張することが困難であるという問題がある。

一方、動画上のフロー計算や特徴点追跡においても輝度パターンの局所変形の問題に取り組んだ研究がある。Rehg ら¹⁵⁾は目標物追跡を行うための多項式変形モデルを考案し、Fuh ら⁵⁾はフロー計算のためのアフィン変形モデルを提案した。しかしながら、これらの方式は物体形状だけでなくカメラの相対的な動きも未知とするため計算コストが高いたけでなく、両者を同時に計算することは本質的に困難であることが指摘されている^{1),17)}。このため、以下のような現実的なアプローチが採用されている。Shi ら¹⁷⁾はアフィン変形モデルを用いてフローの信頼度を評価し、Bergen ら¹⁾はカメラの動きを既知としてアフィン変形モデルにより平面フローを計算している。Manmatha¹¹⁾は、空間フィルタを変形することによって、相似変換（スケール変化と光軸まわりの回転）を評価する方式を示した。

本稿ではステレオマッチングにおける対応領域の輝度パターンの変形問題について言及する。エビポラ拘束を利用して対応領域の変形を物体表面の勾配 (p, q) を未知とする 2 自由度のアフィン変形で表現し、 (p, q) を輝度勾配から直接計算する。これにより、計算コストを増大させることなく変形を考慮した対応づけが可能となる。さらに、対応領域の変形を視差勾配ではなく、3 次元空間中の表面勾配（奥行き勾配）に基づいて定式化することにより、繰返しパターンやオクルージョンに対して有効な複数のステレオ画像対の利用が可能となる。

3. 画像の局所変形

本章では、ステレオ画像間の対応点周囲の輝度パターンの変形について解析する。提案法では物体上の局所領域を任意勾配の平面で近似する。平面の画像間の変換は一般には 2 次元の射影変換で表現されるが^{16),18)}、局所的にはアフィン変換で近似できる。この線形化により、輝度勾配を用いた表面勾配の直接計算が可能となる。

図 3 に示すようなステレオカメラを考える。第 1 のカメラの座標系を基準とし、 $X = (X, Y, Z)^T$ と

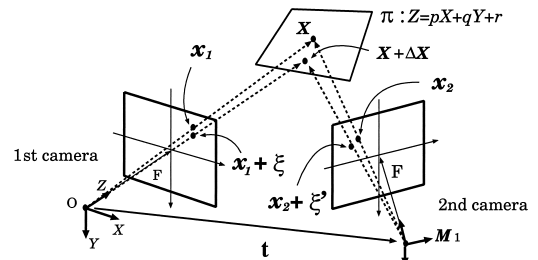


図 3 座標系の定義

Fig. 3 The stereo geometry.

$X + \Delta X = (X + \Delta X, Y + \Delta Y, Z + \Delta Z)^T$ を空間中の平面 π 上の点とする。これら 2 点の第 1 の画像への投影点を各々 $x_1 = (x, y)^T$, $x_1 + \xi = (x_1 + \xi, y_1 + \eta)^T$ として、焦点距離を F とすると、

$$x_1 = \frac{F}{Z} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (1)$$

$$x_1 + \xi = \frac{F}{Z + \Delta Z} \begin{bmatrix} X + \Delta X \\ Y + \Delta Y \end{bmatrix} \quad (2)$$

という関係式が成り立つ。同様に点 X と $X + \Delta X$ の第 2 の画像への投影を各々 $x_2 = (x_2, y_2)^T$, $x_2 + \xi' = (x_2 + \xi', y_2 + \eta')^T$ とし、第 2 のカメラの姿勢を表す $M_1 = (m_{11}, m_{12}, m_{13})^T$, $M_2 = (m_{21}, m_{22}, m_{23})^T$ を図 3 のように定義して $M = (M_1^T, M_2^T)^T$ とすると、

$$x_2 = \frac{F}{Z'} M(X - t) \quad (3)$$

$$x_2 + \xi' = \frac{F}{Z' + \Delta Z'} M(X + \Delta X - t) \quad (4)$$

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。平面 π の方程式を $Z = pX + qY + r$ とおくと、 X と $X + \Delta X$ は平面 π 上の点だから $\Delta Z = p\Delta X + q\Delta Y$ となる。これと式 (1) ~ (4) から、

$$\xi' = A\xi = (A_c + A_v)\xi \quad (5)$$

を得る。ここで、 $A = A_c + A_v$,

$$A_c = \frac{Z}{Z'} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \quad (6)$$

$$A_v = \frac{Z}{Z'} \begin{bmatrix} pm_{13} & qm_{13} \\ pm_{23} & qm_{23} \end{bmatrix} \quad (7)$$

であり、導出過程で $\Delta Z \ll Z$, $\Delta Z' \ll Z$ とした。式 (5) は図 4 に示すような対応点周囲の輝度パターンのス

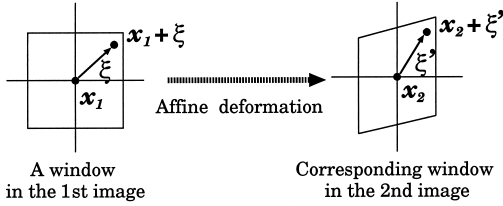


図4 ウィンドウの変形
Fig. 4 Window deformation.

テレオ画像間の変形を表す．提案法では，ウィンドウを式(5)により変形させながら対応づけを行う．式(5)の行列 A_c は第2のカメラの姿勢 $M = (M_1^T, M_2^T)^T$ とスケールファクタ $s = Z/Z'$ にも依存する． M は既知であり， s も奥行き探索とともに得られるパラメータであるので， A_c は容易に決定できる．一方， 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 (p, q) を含むため，式(5)を用いてウィンドウを変形させるためには (p, q) を評価する必要がある．行列 A_v は通常の正方ウィンドウを用いた対応づけでは，物体表面は局所的に画像面と平行であるという仮定 ($p = q = 0$ ，つまり $A_v = O$) に基づいて無視されるが，提案法ではマッチング精度向上のために A_v を考慮してウィンドウを変形させる．

4. 局所勾配の直接計算

奥行き Z と同様に物体表面の勾配 (p, q) も探索により求めると，3次元の探索が必要となるため計算コストが多大なものとなる．本章では，対応点周囲の輝度勾配から直接的に表面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⁸⁾を変形した方法である．本方法では，以下に示す対応領域の輝度の線形変化を仮定する⁵⁾．

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

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，図4に示すように 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吸収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 ， b は画像の場所ごとに異なる．勾配法により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)において $a = 1$ ， $b = 0$ 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．この拡張は時系列画像と比べると，対応点の位置ズレが大きいステレオ画

像にとって重要である．

$$\begin{aligned} \text{式(8)の右辺に式(5)を代入して Taylor 展開すると,} \\ g = g(x_2 + A_c \xi + A_v \xi') \\ = g(x_2 + A_c \xi) + \nabla g(x_2 + A_c \xi) A_v \xi' \quad (9) \end{aligned}$$

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．

$$\Phi \alpha = G(\xi) \quad (10)$$

ただし，

$$\Phi = \begin{bmatrix} \Gamma(\xi)\xi & \Gamma(\xi)\eta & F(\xi) & -1 \end{bmatrix} \quad (11)$$

ここで $\Gamma(\xi) = s\{m_{13}g_x(x_2 + A_c \xi) + m_{23}g_y(x_2 + A_c \xi)\}$ ， $F(\xi) = -f(x_1 + \xi)$ ， $G(\xi) = -g(x_2 + A_c \xi)$ ， $\alpha = (p, q, a, b)^T$ 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，各点に関する方程式を連立させれば未知パラメータ α を得る．つまり，局所勾配 (p, q) は最低2枚の画像から求めることができる．3枚以上の画像を用いる場合も同様に局所勾配 (p, q) を求めることが可能である．

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，

$$a'f(x_1 + \xi) + b' = h(x_3 + \xi'') \quad (12)$$

同様の式変形を行うと，

$$\Phi' \beta = H(\xi) \quad (13)$$

ただし，

$$\Phi' = \begin{bmatrix} \Gamma'(\xi)\xi & \Gamma'(\xi)\eta & F(\xi) & -1 \end{bmatrix} \quad (14)$$

を得る．ここで $\Gamma'(\xi) = s'\{m'_{13}h_x(x_3 + A'_c \xi) + m'_{23}h_y(x_3 + A'_c \xi)\}$ ， $H(\xi) = -h(x_3 + A'_c \xi)$ ，さらに $\beta = (p, q, a', b')^T$ である．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(13)が成り立つので，これら $2N$ 本の線形連立方程式を $\gamma = (p, q, a, b, a', b')^T$ について解くことにより局所勾配 (p, q) を得る．この方法によれば，対応点の周囲の輝度勾配から物体表面の勾配 (p, q) が直接得られるため，比較的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が可能となる．また，反復計算で勾配の精度を向上させることも可能である． (p_k, q_k) を k 回の反復後の計算値とすると，この際には前述のウィンドウ変形を表す行列を

$$A' = A'_c + A'_v \quad (15)$$

と分解する．ただし，

以下で a' 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．

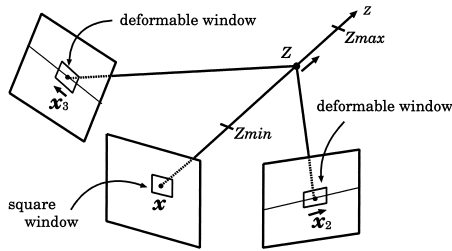


図5 3枚の画像を用いたステレオマッチング

Fig. 5 The stereo matching using three images.

$$A'_c = A_c + s \begin{bmatrix} p_k m_{13} & q_k m_{13} \\ p_k m_{23} & q_k m_{23} \end{bmatrix} \quad (16)$$

$$A'_v = s \begin{bmatrix} \Delta p m_{13} & \Delta q m_{13} \\ \Delta p m_{23} & \Delta q m_{23} \end{bmatrix} \quad (17)$$

である。この式から同様に修正値 $(\Delta p, \Delta q)$ を求め、局所勾配を $(p_{k+1}, q_{k+1}) = (p_k, q_k) + (\Delta p, \Delta q)$ と更新する。

5. アルゴリズム

本章では、これまでの解析に基づいて、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法を示す。提案法は一般の $N(\geq 2)$ 眼ステレオ画像を入力として用いることが可能であるが、ここでは図5に示すように3枚の画像を用いた場合について示す。

- (1) 基準画像上の点 x の奥行き Z の初期推定値 Z_0 をウィンドウ内の奥行きが一樣であるという仮定 ($p = q = 0$) により計算する。
- (2) x の奥行きを Z_0 として勾配 (p, q) を求める。
- (3) x の奥行きを Z_0 の周囲の Z と仮定し、 Z と (p, q) を式(5)に代入して、各参照画像上のウィンドウのアフィン形状を表す行列 $A(j)$ ($j = 2, 3$) を求め、奥行き Z の妥当性 $C(Z)$ を計算する。 $C(Z)$ は基準画像と各参照画像間の正規化相互相関 (normalized cross correlation) の和とする。
- (4) 探索範囲 ($Z_0 - \Delta Z < Z < Z_0 + \Delta Z$) 内の各奥行き Z に対してステップ(3)を実行し、 $C(Z)$ が最大となる Z^* を x の真の奥行きとする。

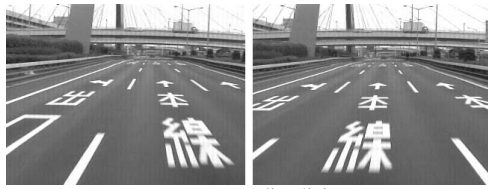
ここでは計算コストを考慮して奥行き初期値を用いる方法を示しているが、探索範囲の各奥行き Z に対して局所勾配を計算するという方法も考えられ、こちらの方式もほぼ同等の性能があることを確認している。

6. 実験

本章では提案法を実画像に適用した実験について述べる。まず、物体表面の局所勾配による対応領域の変形を考慮して対応点を探索する提案法と、勾配による変形を考慮しないで対応点を求める従来法の視差に対する相関値 C の変化を比較した。図6(a)に示す道路シーンのステレオ画像 (320×240) の基準画像 (右) からピックアップした3点 A, B, C (図6(b)) の各対応点を探索したときの視差に対する相関値 C の変化を図6(c), (d), (e)に示す。探索ウィンドウのサイズは 15×15 画素とした。実線と破線が各々、従来法と提案法による相関値の変化を示しており、横軸 d は真の視差で0となるように設定している。つまり、相関値 C のピーク位置が $d = 0$ に近いほど良い。点 A, B, C のいずれの場合も、提案法 (破線) が従来法 (実線) に比べて $d = 0$ に近い位置で最大となっている。たとえば、点 A の例では提案法は $d = 0$ の位置で正しく最大となっているが、従来法では $d = -21$ で最大値をとる。また、点 B の例では従来法では高さほぼ同じであるピークが複数存在し、しかも $d = -27$ で最大となるが、提案法では真の視差付近 ($d = 1$) で明確なピークを持つ。これらの結果から、物体表面の局所勾配によるウィンドウの変形を考慮すると、相関値が正しい対応位置のより近くで、より明確なピークを持つことが分かる。

同様の探索を道路領域全体に対して行って求めた視差の平均誤差を図7(a)に示す。ここでは、3車線分 (走行車線と左右両隣の車線) を対象領域とし、視差の真値は対象領域が平面であるとして求めた。縦軸が視差の平均誤差であり、横軸が探索ウィンドウのサイズである。実線 ($-o-$) が表面勾配による対応領域の変形を考慮しない従来法 (カメラの幾何学的関係による変形のみ考慮) による実験結果であり、2本の破線がアフィン変形ウィンドウを用いた場合の実験結果である。2本のうち、“-+”は局所勾配を反復計算を使わずに求めた場合の結果であり、“- -”は反復計算を用いた結果を示している。反復回数は最大で5回とした。 7×7 より小さいウィンドウでは、十分な輝度の

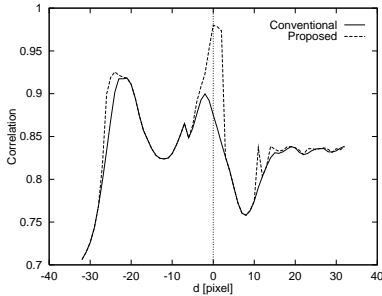
これらの結果ならびに本章で示す提案法による対応づけ精度の向上は、単に輝度の線形変化 (式(8)) を仮定したことに起因することはあり得ない。提案法では5章で述べたようにマッチングの評価関数として正規化相互相関 C を用いており、 C は輝度の線形変化に対して不変な関数であるため、対応領域の形状変化を仮定しなければ輝度変動は C によって吸収される。道路上の特徴点を用いて道路面に対するステレオ画像間の変換パラメータを求め、このパラメータから対象領域の視差を求めた。



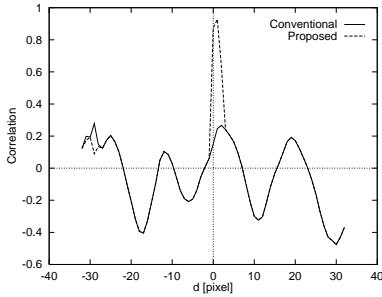
(a) ステレオ画像 “道路 1”



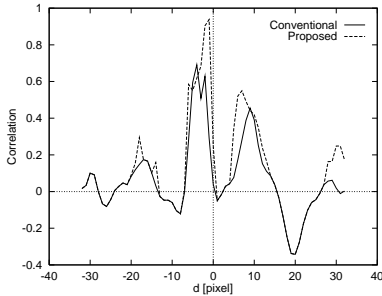
(b) サンプル点 A, B, C



(c) 点 A における相関値の変化



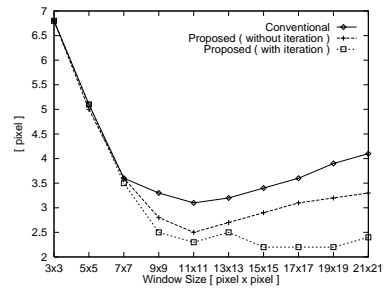
(d) 点 B における相関値の変化



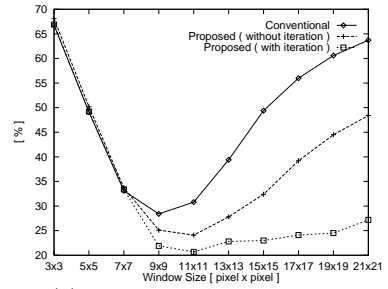
(e) 点 C における相関値の変化

図 6 入力画像 (a) のサンプル点 (b) における視差に対する相関値の変化 (c), (d), (e)

Fig. 6 (a) “Road1” stereo images. (b) Points A, B and C. (c), (d), (e) Correlation values versus disparity.



(a) Average disparity error.



(b) Percentage of gross error.

図 7 精度の比較

Fig. 7 Comparison of the precision.

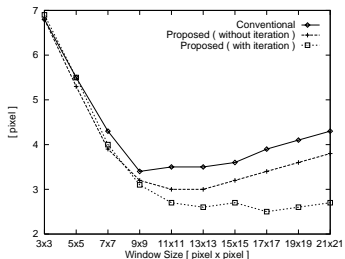
変動がないので、いずれの方法でも誤差が大きい。したがって、視差を精度良く求めるためには、より大きなウィンドウを用いることが不可欠である。9×9 以上のいずれのサイズのウィンドウでも、提案法の結果が従来法に比べて精度が高いことが分かる。また、表面勾配を反復計算により求めた場合、精度が向上していることが分かる。図 7 (b) には誤対応 (比較的大きな対応誤り) の割合を示す。ここでは、真値からの 3 画素以上の視差の相違を誤対応と定義した。平均誤差の場合と類似した傾向を持ち、9×9 以上のいずれのサイズのウィンドウでも提案法が従来法よりも誤対応が少なく、従来法では最も誤対応が少ない場合 (9×9) でも 28.4% の誤対応を含むのに対し、局所勾配の反復計算をとらなう提案法では、最小値が 20.7% (11×11 ウィンドウ) となっている。

図 8 (a) のステレオ画像 “道路 2” に対する同様の定量評価結果を図 8 (b), (c) に示す (走行車線のみを対象領域)。“道路 1” と同様に提案法が従来法に比べて、全体的に精度が高く、かつ、誤対応が少ないことが分かる。

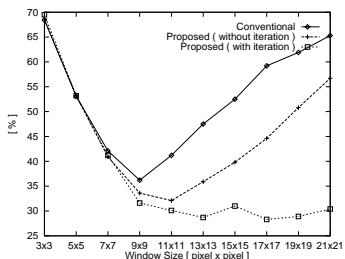
次に 3 枚の画像を入力として行った実験について述べる。ここでも詳細な精度評価を行うため、まず、図 9 に示す形状既知の物体 (野球の軟式ボール、半径 $R = 3.5 \text{ cm}$) を用いた。この物体から約 50 cm 離れた位置から焦点距離 25 mm のレンズを装着した CCD カメラを用いて、480×480 画素の画像を 3 枚撮影し



(a) ステレオ画像“道路2”.



(b) Average disparity error.



(c) Percentage of gross error.

図 8 精度の比較

Fig. 8 Comparison of the precision.

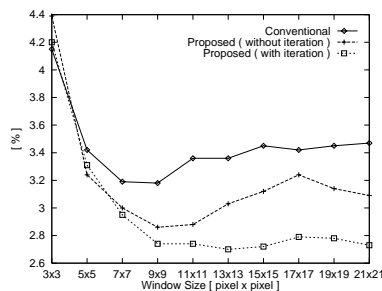


図 9 入力画像“ボール”(480 × 480).

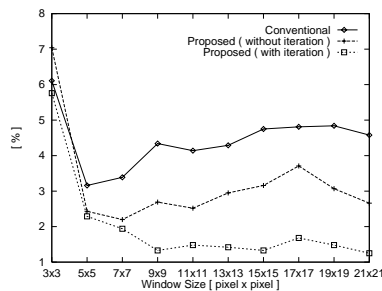
Fig. 9 “Ball” images (480 × 480).

た．中央の画像が基準画像であり，この画像上の各点の奥行き（提案法では局所勾配も）を計算した．また，真値は球体であるとして求めた．

図 10 に“道路 1”，“道路 2”と同様の定量評価結果を示す．すなわち，実線（-○-）が表面勾配を考慮しない従来法（カメラの幾何学的関係による変形のみ考慮）による実験結果であり，2本の破線が提案法による実験結果である（“-+”... 反復計算なし，“- -”... 反復計算使用）．ただし，図 10(a)の縦軸は物体のサイズ，つまり，半径 R で正規化した奥行き（平均誤差）を示す．ここでも小さいウィンドウ（3 × 3 や 5 × 5）では，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．3 × 3 の場合，提案法は従来法より



(a) Average relative error.



(b) Percentage of gross error.

図 10 精度の比較

Fig. 10 Comparison of the precision.

もわずかに大きな誤差を含む．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，従来法による誤差も大きく，精度良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．5 × 5 以上のいずれのサイズのウィンドウでも，提案法の結果が従来法に比べて精度が高いことが分かる．また，表面勾配を反復計算により求めた場合，誤差が小さくなっていることが分かる．図 10(b)には誤対応（大きな対応誤り）の割合を示す．ここでは，真値からの 20 画素（約 4 mm に相当）以上の奥行き（相違）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，5 × 5 以上のいずれのサイズのウィンドウでも提案法の方が従来法よりも誤対応が少なく，従来法が 3 ~ 5% の誤対応を含むのに対し，勾配の反復計算をともなう提案法では誤対応が 1 ~ 2% となっている．図 11 に 11 × 11 画素のウィンドウを用いた提案法による復元形状を示す．

次に顔，手といった画像に対し提案法を適用した結果を図 12, 13, 17, 18 に示す．各対象物とカメラとの距離は約 50 cm である．これらの対象物は比較的テクスチャが少なく，形状復元が困難であるが，精度の高い形状が復元されており，これらの結果は提案法が様々な画像に対して適用可能であることを示す．図 14 には“顔”の奥行き画像を示す．同図 (a), (b)

奥行きを輝度で表現した画像．明るいほど近いことを示す．

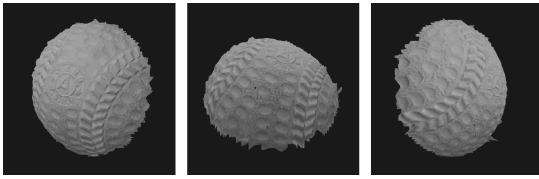


図 11 “ボール”の復元形状

Fig. 11 Recovered shape of “Ball” by the proposed algorithm.

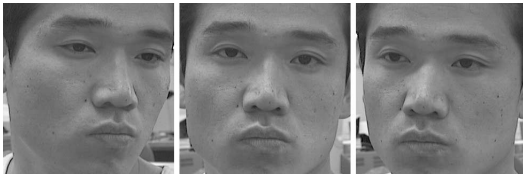


図 12 入力画像“顔”

Fig. 12 “Face” images.

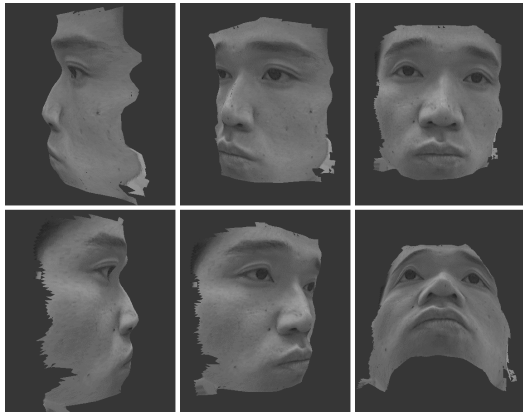
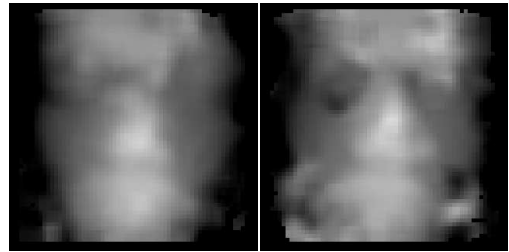


図 13 “顔”の復元形状

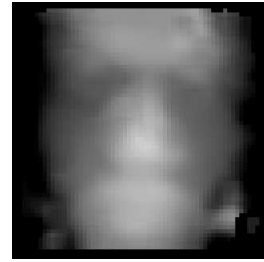
Fig. 13 Recovered shape of “Face”.

は各々、左画像と基準（中央）画像、右画像と基準画像のみを使って得られた結果であり、(c)は3枚の画像を入力とした場合の結果である。(a),(b)に比べて3枚の画像を入力とした結果(c)は全体的に妥当な形状が復元されている。たとえば、図15(a)に示す両目頭の間の奥行き断面を同図(b)に示す。この図の横軸、縦軸は各々、(a)に示すライン上の水平位置と奥行きを示す。横軸は(a)に示すように最も盛り上がっている部分が原点となるように設定した。実線(—◇—)が3枚の画像を用いて得られた奥行き断面であり、2本の破線が2枚の画像を入力として得られた結果である(“-+”...左画像と基準画像、“- -”...基準画像と右画像)。他の2つの結果に比べて画像を3枚用いた場合には、原点を中心とした左右ほぼ対称で滑らかな起伏が再現されており、これらの結果は2眼ステレ



(a) 左画像+基準画像

(b) 基準画像+右画像



(c) 基準画像+左右画像

図 14 奥行き画像の比較

Fig. 14 Comparison of the depth map. The lighter, the closer.

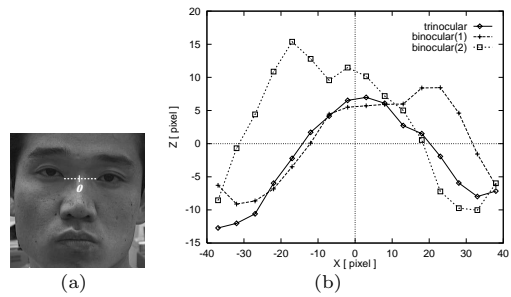
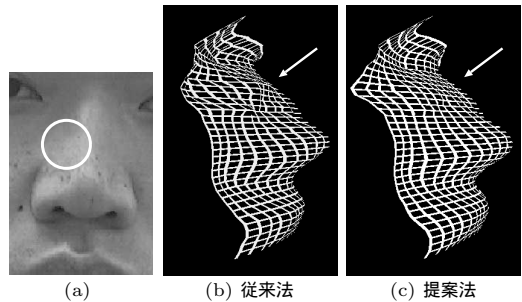


図 15 “顔”の両目頭間 (a)の奥行き断面 (b)

Fig. 15 (a)Target part in “Face”. (b) Recovered profiles by the proposed algorithm and conventional ones.



(a)

(b) 従来法

(c) 提案法

図 16 “顔”の鼻周辺部 (a)の復元形状 (b),(c)

Fig. 16 (a) Nose area in “Face”. (b) Recovered shape by the conventional algorithm. (c) Recovered shape by the proposed algorithm.

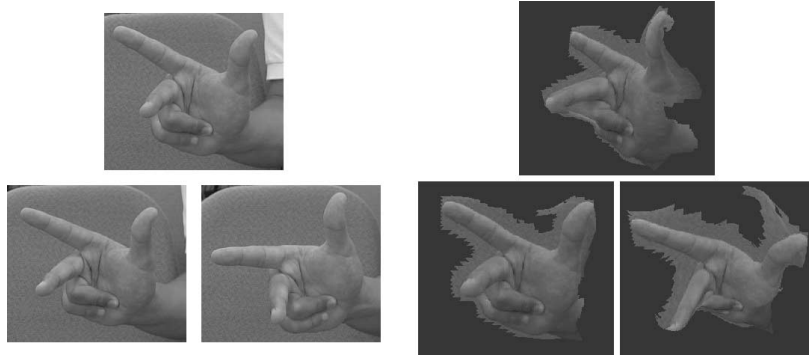


図 17 入力画像“手 1” (左) と復元形状 (右)
 Fig. 17 “Hand 1” images (left) and recovered shape (right).

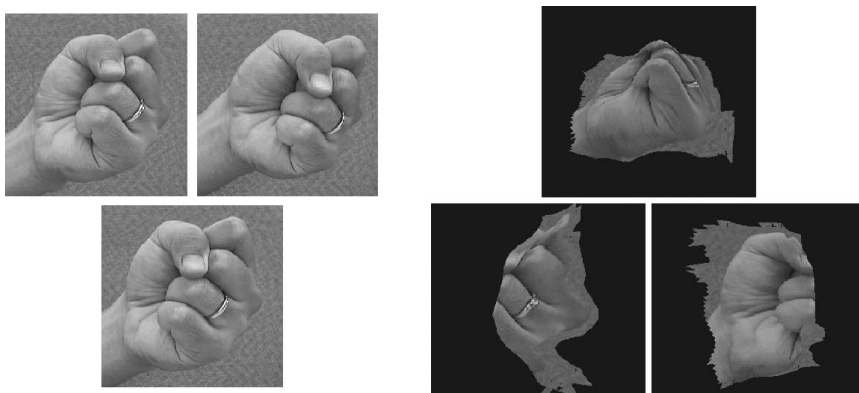
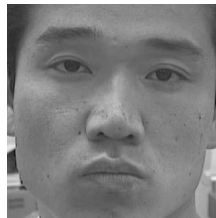
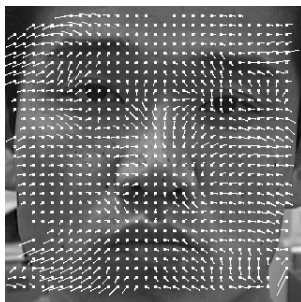


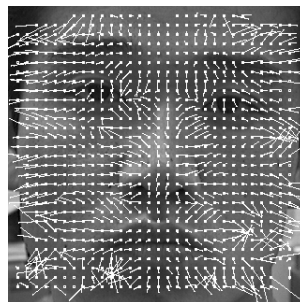
図 18 入力画像“手 2” (左) と復元形状 (右)
 Fig. 18 “Hand 2” images (left) and recovered shape (right).



(a) Original.



(b) Proposed.



(c) From depth map.

図 19 “顔” の表面勾配の復元結果の比較
 Fig. 19 Recovered surface orientations of “Face”.

オに対する多眼ステレオの有効性を示している。

図 16 には “顔” の鼻周辺部 (a) の復元形状を示す。(b) が局所勾配を考慮し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。(a) に示す白い円で囲んだ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に対し, 提案法ではより正確な形状が復元されている。

提案法は各点の奥行きだけでなく, 表面勾配も出力する。図 19 (a) に提案法によって得た, 入力画像 “顔” の各点の局所勾配を示す。また, 比較のため, 得られた奥行きマップに局所的に平面をあてはめることによって計算した表面勾配を図 19 (b) に示す。奥行きマップから求めた場合, 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, 提案法においては, 画像から直接的に表面勾配を得るため, より正確に計算することができる。この結果は, 提案法が物体表面の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。

7. ま と め

本稿では, 各点の局所勾配を考慮したステレオマッチング法を提案した。従来方式が, 物体表面を局所的に画像面と平行な面で近似するのにに対し, 提案法においては, 物体表面を局所的に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, 求めた局所勾配を用いてウィンドウを適応的に変形させながら対応づけを行う。実画像を用いた実験を通して, 提案法が従来方式に比べて高精度な形状復元を実現すること, ならびに, 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, 一般的な投影モデル下での structure-from-motion の手法^{14), 19)} 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。

参 考 文 献

- 1) Bergen, J.R., Anandan, P., Hanna, K.J. and Hingorani, R.: Hierarchical model-based motion estimation. *Proc. 2nd ECCV*, pp.237–252 (1992).
- 2) Devernay, F. and Faugeras, O.: Computing differential properties of 3D shapes from stereoscopic images without 3D models. Technical report 2304, INRIA, France (1994).
- 3) Dhond, U.R. and Aggarwal, J.K.: Structure from stereo — a review, *Trans. Systems, and*

- Man Cybernetics*, 19:1489–1510 (1989).
- 4) Faugeras, O.D.: What can be seen in three dimensions with an uncalibrated stereo rig?, *Proc. 2nd ECCV*, pp.563–578 (1992).
- 5) Fuh, C. and Maragos, P.: Motion displacement estimation using an affine model for matching, *Optical Engineering*, Vol.30, No.7, pp.881–887 (1991).
- 6) Gårding, J. and Lindeberg, T.: Direct estimation of local surface shape in a fixating binocular vision system, *Proc. 3rd ECCV*, pp.365–376 (1994).
- 7) Hattori, H. and Maki, A.: Stereo matching with direct surface orientation recovery, *Proc. 9th British Machine Vision Conference*, pp.356–366 (1998).
- 8) Horn, B.K.P. and Schunk, B.G.: Determining optical flow, *Artificial Intelligence*, Vol.17, pp.185–203 (1981).
- 9) Jones, D.G. and Malik, J.: Determining three-dimensional shape from orientation and spatial frequency disparities, *Proc. 2nd ECCV*, pp.661–669 (1992).
- 10) Maimone, M.W. and Shafer, S.A.: Modeling foreshortening in stereo vision using local spatial frequency, *Proc. International Conference on Intelligent Robots and Systems*, pp.519–524 (1995).
- 11) Manmatha, R.: Measuring the affine transform using gaussian filters, *Proc. 3rd ECCV*, pp.159–164 (1994).
- 12) Okutomi, M. and Kanade, T.: A locally adaptive window for signal matching, *IJCV*, Vol.7, No.2, pp.143–162 (1992).
- 13) Okutomi, M. and Kanade, T.: A multi-baseline stereo, *PAMI*, Vol.15, No.4, pp.353–363 (1993).
- 14) Pollefeys, M., Koch, R. and Gool, L.V.: Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, *Proc. 6th ICCV*, pp.90–95 (1998).
- 15) Rehg, J.M. and Witkin, A.P.: Visual tracking with deformation models, *Proc. International Conference on Robotics and Automation*, pp.844–850 (1991).
- 16) Robert, J. and Hebert, M.: Deriving orientation cues from stereo images, *Proc. 3rd ECCV*, pp.377–388 (1994).
- 17) Shi, J. and Tomasi, C.: Good features to track, *Proc. CVPR*, pp.593–600 (1994).
- 18) Torr, P., Fitzgibbon, A.W. and Zisserman, A.: Maintaining multiple motion model hypotheses over many views to recover matching and structure, *Proc. 6th ICCV*, pp.485–491 (1998).

- 19) Triggs, B.: Factorization methods for projective structure and motion, *Proc. CVPR*, pp.845-851 (1996).
- 20) Satoh, K., Kanamura, Y., Matsuura, T. and Ohta, Y.: Occlusion detectable stereo — occlusion patterns in camera matrix, *Proc. CVPR*, pp.371-378 (1996).
- 21) 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).
- 22) 服部：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, 情報処理学会研究報告, 98-CVIM-109, pp.33-40

(1998).

(平成 12 年 9 月 1 日受付)

(平成 13 年 1 月 19 日採録)

(担当編集委員 長尾 健司)



服部 寛 (正会員)

昭和 41 年生。平成 3 年早稲田大学理工学部応用物理学卒業。同年 (株)東芝入社。コンピュータビジョンの研究開発に従事。日本ロボット学会会員。