

多眼ステレオカメラを用いた 直接法による全周サーフェス生成

塩田 陽介^{†1} 杉本 茂樹^{†1} 奥富 正敏^{†1}

本論文では、多眼ステレオカメラで撮影した画像を利用して、撮影対象のサーフェスモデルを直接法により生成する手法を提案する。提案手法では、微小三角メッシュを対象のサーフェスモデルとし、空間中のある一点からメッシュ各頂点までの距離をモデルパラメータとする。そして、画像間の輝度差と形状の滑らかさに基づいたコストを定義し、各カメラからの各平面の可視性を考慮しつつ、高速平面パラメータ推定法を拡張した最適化手法によりコストを効率的に最小化する。さらに、最小化の際に階層メッシュを利用し、階層間での平面姿勢の類似度をコストに加えることにより、安定なサーフェス生成を実現する。様々な形状と拘束項に関する比較実験を通じ、提案手法の有効性を示す。

A Direct Method for 3D Reconstruction from Multi-View Stereo

YOSUKE SHIOTA,^{†1} SHIGEKI SUGIMOTO^{†1}
and MASATOSI OKUTOMI^{†1}

In this paper, we propose a direct method for 3D reconstruction from multi-view stereo images. In the proposed method, we reconstruct a 3D surface by estimating the parameters of a piecewise planar surface model, which are the distances from the world origin to all vertices of the triangular patches. We minimize the cost function composed by the pixel value differences between images and smoothness constraints of the surface shape by an efficient optimization method extended from the fast plane-parameter estimation proposed recently. Furthermore, we realize robust estimation by using hierarchical mesh and the additional cost of similarity measure of plane orientation between hierarchical levels. The validity of the proposed method is demonstrated through results of experiments using synthetic images.

1. はじめに

伝統的なステレオ計測では、実物体の3次元形状を、空間中の点群として取得する。一方、3次元形状をサーフェスモデルによって表現することは、CGによる可視化だけでなく、3次元データベースにおけるメモリ量の低減や、形状認識などのアプリケーションにおいても有用である。

サーフェスを取得するための代表的なアプローチは、多眼ステレオ画像を用いた計測などによって得られた3次元空間中の点群データを、ポリゴンメッシュやB-スプラインなどのサーフェスモデルにあてはめる方法である^{2),4),5),10)}。しかし、ステレオ計測によって得られた点群データには一般に誤差が含まれており、このアプローチによって適切なサーフェスを生成するには複雑な処理が必要となる⁵⁾。ボクセルベースの3次元形状復元を利用すると、Marching Cubes⁹⁾などにより、ポリゴンメッシュを容易に生成することができる。Level Set法^{1),11),12)}やGraph Cuts法^{6),7),15),16)}を利用して、多眼ステレオから高精度なボクセルデータを取得することが可能であるが、精度の高いサーフェスを得るためには細かいボクセル空間が必要なため、大きなメモリ量を必要とし計算コストが大きい。

一方で、ステレオ画像から直接的にサーフェスモデルを生成する手法が提案されている^{3),13),14),17)}。これらの手法では、ステレオ計測のように画素単位の奥行きを推定するのではなく、物体形状としてポリゴンメッシュを導入し、そのモデルパラメータを画像から直接的に推定する。例えば、文献^{3),17)}では、空間をサンプリングすることにより多眼ステレオ画像間の輝度差を評価し、そのコストを最小化するパッチ頂点の3次元位置を推定している。ただし、この方法では問題の自由度が大きくなり、最適化に要する計算コストも大きい。一方、文献¹⁴⁾では、2眼ステレオ画像を用いて、各パッチに関する平面射影変換に基づいて画像間の輝度差を評価し、そのコストを最小化するようなパッチ頂点の奥行きを推定している。この方法では、Gauss-Newton法を適用した際の高速な計算手法を提供しているものの、基準カメラ座標系で定義されたサーフェスパラメータを生成しているため、多方向から物体を観測する多眼ステレオのケースに容易に拡張できないなどの問題がある。

本論文では、多眼ステレオ画像を利用して、全周のサーフェスを安定かつ効率的に推定する手法を提案する。提案手法では、微小三角メッシュを対象のサーフェスモデルとし、ワー

^{†1} 東京工業大学大学院理工学研究科機械制御システム専攻

Department of Mechanical and Control Engineering, Graduate School of Science and Engineering,
Tokyo Institute of Technology

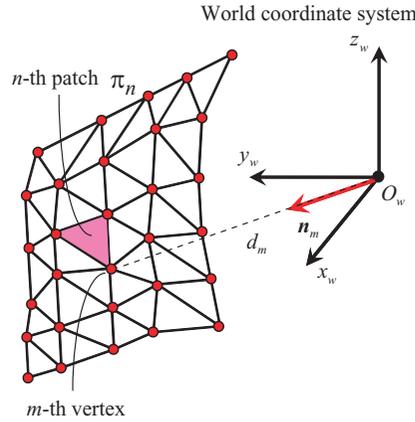


図1 ワールド座標系を基準とする微小三角メッシュモデル

Fig.1 Piecewise triangular mesh model based on world coordinates.

ワールド座標原点からメッシュ各頂点までの距離をモデルパラメータとする．そして，画像間の輝度差と形状の滑らかさに基づいたコストを定義し，そのコストを効率的に最小化する．さらに，最適化の際に階層メッシュを利用し，階層間での平面姿勢の類似度をコストに加え，パッチに関する可視性を考慮することにより，安定なサーフェス生成を実現する．本論文では，合成画像を用いて様々な形状と拘束項に関する比較実験を行い，提案手法の有効性を示す．

2. 形状モデル

提案手法では，物体のサーフェスを微小三角メッシュによって表現する．すなわち，サーフェスは微小な三角パッチが構成するメッシュによって表現され，三角パッチの各頂点は隣接パッチと共有される．

メッシュを構成する三角パッチのインデックスを n , ($n = 1, 2, \dots, N$)，メッシュ頂点（三角パッチの頂点）のインデックスを m , ($m = 1, 2, \dots, M$) とする．すなわち， n 番目の三角パッチ π_n は，頂点インデックス m の集合 $\{i, j, k\}$ を持つ．この集合を $S(n) = \{i, j, k\}$ と書く．

各頂点の3次元座標はワールド座標系において定義され，その座標は $x_m = (x_m, y_m, z_m)^T = d_m n_m$ と表わされる．ただし， n_m は，ワールド座標原点 O_w からの

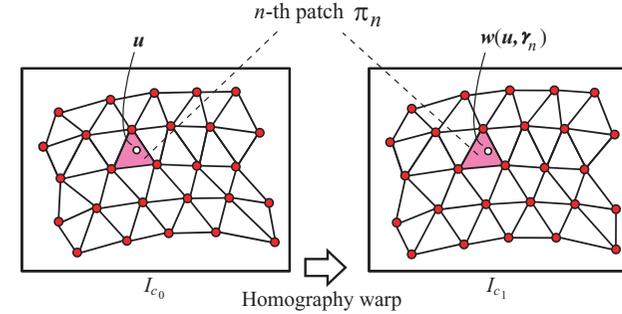


図2 ステレオ画像間のメッシュ形状変化

Fig.2 Mesh on stereo images.

方向を表す単位法線を示し， d_m は原点からの距離である．単位法線 n_m は，空間中のメッシュを単位球に投影した際の各頂点の座標を表わすものであり，ワールド座標原点の位置 O_w とともに，ユーザが定義するものとする．このとき，サーフェスを一意に定めるパラメータベクトルを，各頂点の原点からの距離 d_m の逆数を用いて，次のように表わす．

$$\Gamma = (1/d_1, 1/d_2, \dots, 1/d_M)^T \quad (1)$$

本論文では， Γ を求めるべきサーフェスパラメータとする．

3. コスト関数

提案手法では，以下のコスト関数 E を最小化することにより，サーフェス生成を行う．

$$E = E_d + E_s + E_h \quad (2)$$

ただし， E_d, E_s, E_h は，それぞれ，画像間の輝度差に基づくデータ項，空間的な平滑化項，階層的メッシュにおける平滑化項を表す．以下で，各項の詳細について述べる．

3.1 データ項

パッチ π_n の3頂点の距離 $d_m, m \in S(n)$ が定まると，画像間の平面射影変換が算出できる． E_d は，パッチ π_n について，基準画像と他の画像を平面射影変換によって変形した結果から得られる輝度のSSD (Sum of Squared Differences) を，全画像ペアについて加算し，更にそれを全パッチについて加算したものである．

いま，多眼ステレオ画像中の c 番目の画像を $I_c, (c = 1, 2, \dots, C)$ とする．任意の2枚の画像を I_{c_0}, I_{c_1} とし， c_0 は基準画像のインデックス， c_1 は参照画像のインデックスを表すものとする．このとき，パッチ π_n に関する2枚の画像間のSSDは，次式で表わされる．

$$E_{d,n}^{(c_0,c_1)} = \sum_{\mathbf{u} \in \mathcal{R}_n} (I_{c_0}[\mathbf{u}] - I_{c_1}[\mathbf{w}(\mathbf{u}, \gamma_n)])^2 \quad (3)$$

ただし、 \mathbf{u} は基準画像上の正規化画像座標を表し、 $\mathbf{w}(\mathbf{u}, \gamma_n)$ は、 \mathbf{u} を平面射影変換した座標を表す（図 2 を参照）。この平面射影変換は、3 頂点のパラメータベクトル $\gamma_n = (1/d_i, 1/d_j, 1/d_k)^T, \{i, j, k\} \in S(n)$ と、カメラ間の外部パラメータから算出される。また、式 (3) は、空間中の三角パッチ π_n を基準画像上に投影した三角形領域 \mathcal{R}_n についてのみ計算するものとする。なお、全カメラペア間の外部パラメータは、各カメラの内部パラメータとともに、キャリブレーションされているものとする。

いま、画像ペアの全集合を \mathcal{P} とする。パッチ π_n の可視性を考慮して、すべての画像ペア $\{c_0, c_1\} \in \mathcal{P}$ について式 (3) を足し合わせると、パッチ π_n に関するコスト $E_{d,n}$ は、次式のように表わされる。

$$E_{d,n} = \sum_{\{c_0, c_1\} \in \mathcal{P}} \delta_n^{(c_0)} \delta_n^{(c_1)} E_{d,n}^{(c_0, c_1)} \quad (4)$$

ただし、 $\delta_n^{(c)}$ は、 c 番カメラからのパッチ π_n の可視性を表す二値関数であり、 c 番カメラからパッチが見える場合は 1、それ以外は 0 とする。この可視性は、カメラの光軸ベクトルとパッチの単位法線ベクトルとの内積により判断する。

式 (4) を、全パッチについて重み付加算することにより、次式で表されるコスト E_d を得る。

$$E_d = \sum_n \omega_{d,n} E_{d,n} \quad (5)$$

ただし、 $\omega_{d,n}$ は、パッチごとの重みを表す。この重みについては、他のコストにおけるパッチごとの重みとともに、後述する。

3.2 空間的平滑化項

上述した画像間の輝度差に基づくコストのみによってサーフェス生成が可能であり、文献¹⁴⁾の方法では、輝度差によるコストのみを 2 眼ステレオ画像を用いて最小化している。しかし、精度を高くするためにメッシュを細かくすると、画像に含まれるノイズなどの影響により、推定が不安定になる傾向がある。そこで、提案手法では、形状の滑らかさに基づくコストを導入し、推定の安定化を図る。

パッチ π_n の平面パラメータを q_n とする。ただし、 q_n は、ワールド座標系 x において、平面の姿勢を一意に定めるパラメータであり、このとき平面の方程式は $q_n^T x = 1$ と表わされる。また、 π_n に隣接する 3 個のパッチ $\pi_{n_\mu}, (\mu = 1, 2, 3)$ を考え、それらの平面パラメータ

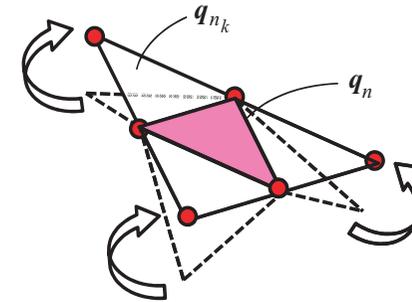


図 3 近接パッチ間の拘束を利用した空間的平滑化拘束
Fig.3 Smoothness constraint on adjacent patches.

を q_{n_μ} と表す（図 3 を参照）。そして、パッチ π_n について、形状の滑らかさを表すコスト $E_{s,n}$ を次式で定める。

$$E_{s,n} = \sum_{\mu=1}^3 \|q_n - q_{n_\mu}\|^2 \quad (6)$$

このコストは、近接する 3 つのパッチ π_{n_μ} が、パッチ π_n がなす平面と同じ平面であるときに 0 となる。よって、このコストは、サーフェスを滑らかにする役割を担い、推定を安定化できる。

式 (6) を、全パッチについて重み付加算することにより、次式で表されるコスト E_s を得る。

$$E_s = \sum_n \omega_{s,n} E_{s,n} \quad (7)$$

ただし、 $\omega_{s,n}$ は、パッチごとの重みを表す。この重みについては後述する。

3.3 階層的メッシュにおける平滑化項

提案手法では、文献¹⁴⁾の方法と同様に、最初は粗いメッシュを利用し、そのメッシュによる推定が終了すると、次に細かいメッシュを利用して推定を行うという、階層的メッシュを利用してサーフェス生成を行う。

この方法は、初期値に対する推定の安定性を高める効果があり、かつ徐々にメッシュを細かくすることで高精度な推定を実現できる。しかし、メッシュが細かくなると、解の次元数が増える一方で、各パッチ内に含まれる画素数が減少するため、推定に失敗しやすくなる。

上述の空間的平滑化項の利用により、推定の安定化を実現することができるが、その拘束は真の形状を適切に表現したのではなく、拘束をあまり強くすると推定結果を過度に平滑化してしまう可能性がある。一方、粗いメッシュによる推定結果は、実形状の大まかな形状を表現した結果であり、メッシュを細かくした推定結果は、粗いメッシュの結果から大きく異なる可能性は少ないと考えられる。そこで、前の階層における粗いメッシュの推定結果を、次の階層における細かいメッシュの推定結果に対する拘束とすることにより、物体の実形状を考慮したロバストな推定が実現できると考えられる。

パッチ π_n に関する階層間の類似度に基づくコスト $E_{h,n}$ を、次式で定める。

$$E_{h,n} = \left\| \mathbf{q}_n^{(l)} - \hat{\mathbf{q}}_n^{(l-1)} \right\|^2 \quad (8)$$

ただし、 $\hat{\mathbf{q}}_n^{(l-1)}$ は、1つ前の階層メッシュレベルにおいて推定された平面パラメータであり、 $\mathbf{q}_n^{(l)}$ は、現在のメッシュレベルの平面パラメータを表す。 $\hat{\mathbf{q}}_n^{(l-1)}$ は、基本的に期待される平面姿勢パラメータであればよく、たとえば Visual hull⁸⁾ による推定結果を利用することが考えられるが、ここでは前のメッシュの階層における推定結果を用いている。

式 (8) を、全パッチについて重み付加算することにより、次式で表されるコスト E_h を得る。

$$E_h = \sum_n \omega_{h,n} E_{h,n} \quad (9)$$

ただし、 $\omega_{h,n}$ は、パッチごとの重みを表す。この重みについては次節で述べる。

3.4 パッチごとの重み

式 (5), (7), (9) それぞれに示したパッチごとの重み $\omega_{d,n}, \omega_{s,n}, \omega_{h,n}$ を適応的に変化させることで、推定のロバスト性が向上できると考えられる。

まず、式 (5) の重み $\omega_{d,n}$ について述べる。提案手法では、この重みを次式によって適応的に定める。

$$\omega_{d,n} = \frac{\alpha + \epsilon_n^{(\text{tex})} / \text{ave}(\epsilon_n^{(\text{tex})})}{\alpha + \epsilon_n^{(\text{mse})} / \text{ave}(\epsilon_n^{(\text{mse})})} \quad (10)$$

ただし、 α はユーザ定義の定数であり、 $\epsilon_n^{(\text{tex})}$ と $\epsilon_n^{(\text{mse})}$ は、三角パッチ π_n を画像に投影した領域内のテクスチャの豊富さ、および画像レジストレーションの良好さを示し、それぞれ次式で表わされる。

$$\epsilon_n^{(\text{tex})} = \frac{1}{\nu} \sum_{\mathbf{u} \in \mathcal{R}_n} \|\nabla I_{c_0}\|^2 \quad (11)$$

$$\epsilon_n^{(\text{mse})} = \frac{1}{\nu} \sum_{\mathbf{u} \in \mathcal{R}_n} (I_{c_0}[\mathbf{u}] - I_{c_1}[\mathbf{w}(\mathbf{u}, \gamma_n)])^2 \quad (12)$$

ただし、 ν はパッチ π_n を画像に投影した際の三角領域内の画素数を表わす。また、 $\text{ave}(\cdot)$ は、全パッチに関する平均操作を表わす。

式 (10) の重みは、最適化における繰返し計算ごとに評価され、テクスチャが豊富なパッチや、良好なレジストレーションが行われているパッチについては、空間的平滑化項や階層的平滑化項よりも、データ項の寄与が大きくなる。

一方、式 (7) の空間的平滑化項と、式 (9) の階層的平滑化項における重みは、提案手法では一定とし、それぞれ次式で定義する。

$$\omega_{s,n} = \nu \omega_s, \quad \omega_{h,n} = \nu \omega_h \quad (13)$$

ただし、 ω_s および ω_h は、それぞれの重みに対してユーザが定義する定数であり、画素数 ν を積算しているのは、データ項との間の正規化をするためである。

4. モデルパラメータの推定

提案手法では、式 (2) を、Gauss-Newton 法を用いて最小化する。すなわち、パラメータベクトル Γ を、現在の推定値 $\bar{\Gamma}$ と微小変化量 $\Delta\Gamma$ を用いて $\Gamma = \bar{\Gamma} + \Delta\Gamma$ とおき、式 (2) を構成する二乗和内の関数を一次近似した上で、式 (2) を $\Delta\Gamma$ で微分することにより、 $\Delta\Gamma$ を導出する。そして、その導出結果を利用して、繰返し演算ごとに $\Delta\Gamma$ の計算と $\bar{\Gamma} \leftarrow \bar{\Gamma} + \Delta\Gamma$ による更新を、所定の収束条件を満たすまで繰り返す。

注目すべきは、式 (7) で示した空間的平滑化項と、式 (9) の階層的平滑化項は、それぞれの二乗和内を一次 Taylor 近似した上で微分すると、そこから導かれるヘッセ行列は現在の推定値 $\bar{\Gamma}$ に依存せず、どちらも定数行列となることである。すなわち、これらの拘束項を付与したことによって生じるヘッセ行列の変化は一定であり、その変化を繰返しごとに再計算する必要はない。

一方、式 (5) のデータ項は、それを一般的な Gauss-Newton 法によって最小化すると、非常に大きな計算コストを必要とする。ただし、文献¹⁴⁾ では、2眼ステレオに用いた微小三角メッシュサーフェス生成に対して、Gauss-Newton 法を適用した場合の高速な計算方法を提供しており、提案手法では、この計算方法を拡張することにより、効率的なサーフェス生成を行う。以下では、詳細は省きつつ、この方法の概要と効果を述べる。

式 (3) は、次式のように変形できる．

$$E_{d,n}^{(c_0,c_1)} = \sum_{\mathbf{u}} (I_{c_0}[\Delta \mathbf{w}(\mathbf{u}, \Delta \gamma_n)] - I_{c_1}[\mathbf{w}(\mathbf{u}, \bar{\gamma}_n)])^2 \quad (14)$$

ただし、 $\Delta \mathbf{w}(\mathbf{u}, \Delta \gamma_n)$ は、 $\mathbf{w}(\mathbf{u}, \gamma_n) = \mathbf{w}(\mathbf{u}, \bar{\gamma}_n) \circ \Delta \mathbf{w}(\mathbf{u}, \Delta \gamma_n)^{-1}$ の関係にある平面射影変換行列である．式 (14) を用いて、式 (5) のデータ項を表現すると、式 (2) を最小化する $\Delta \Gamma$ は、最終的に次式のように書ける．

$$\Delta \Gamma = - \left(\sum_n \mathbf{H}_n \right)^{-1} \left(\sum_n \mathbf{b}_n \right) \quad (15)$$

where

$$\mathbf{H}_n = \omega_{d,n} \mathbf{H}_{d,n} + \omega_{s,n} \mathbf{H}_{s,n} + \omega_{h,n} \mathbf{H}_{h,n} \quad (16)$$

$$\mathbf{b}_n = \omega_{d,n} \mathbf{b}_{d,n} + \omega_{s,n} \mathbf{b}_{s,n} + \omega_{h,n} \mathbf{b}_{h,n} \quad (17)$$

ただし、 $\mathbf{H}_{d,n}$ 、 $\mathbf{H}_{s,n}$ 、 $\mathbf{H}_{h,n}$ および、 $\mathbf{b}_{d,n}$ 、 $\mathbf{b}_{s,n}$ 、 $\mathbf{b}_{h,n}$ は、式 (4)、(6)、(8) から得られるヘッセ行列と最急降下勾配ベクトルをそれぞれ表す．上述のように、 $\mathbf{H}_{s,n}$ 、 $\mathbf{H}_{h,n}$ は繰返しごとに変化せず一定であるが、データ項から得られる $\mathbf{H}_{d,n}$ は、 $\bar{\Gamma}$ に依存するため、繰返しごとに再計算が必要である．ただし、 $\mathbf{H}_{d,n}$ は、次式のように表わされる．

$$\mathbf{H}_{d,n} = \sum_{\{c_0,c_1\} \in \mathcal{P}} \kappa_{n,c_0,c_1} \delta_n^{(c_0)} \delta_n^{(c_1)} \mathbf{H}_{d,n,c_0,c_1} \quad (18)$$

ここで、 κ_{n,c_0,c_1} は、カメラ間の外部パラメータと π_n の平面パラメータに依存するスカラー値である．式 (18) で、 $\bar{\Gamma}$ に依存するのは、スカラー値 κ_{n,c_0,c_1} のみであり、 \mathbf{H}_{d,n,c_0,c_1} は繰返しごとに変化しない．すなわち、パッチ π_n に関するヘッセ行列は、可視性やパッチごとの重みの変化があるものの、各画像ペアから得られる固定ヘッセ行列の重み付加算のみによって得られる．これは、提案手法では、画素ごとの演算は必要としないことを意味しており、一般的な Gauss-Newton 法による最適化よりも、大幅に効率的な最適化を実現できる．

一方で、式 (15) により $\Delta \Gamma$ を求めるためには、ヘッセ行列 $\sum_n \mathbf{H}_n$ の逆行列を計算する必要がある．この行列は、頂点数 $M \times$ 頂点数 M の正方行列であり、その逆行列の計算コストは比較的大きい．ただし、この対称行列は、ほとんどの要素がゼロとなる疎行列であることから、前処理付き共役勾配法を用いて効率的に逆行列を計算できる．

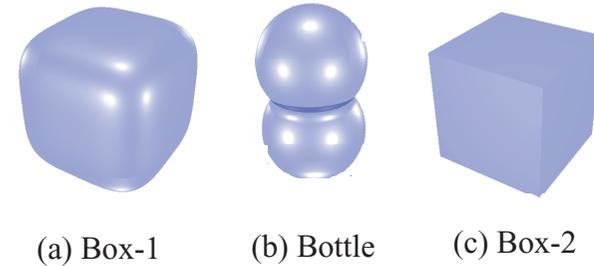


図 4 真の形状
Fig. 4 Actual shape.

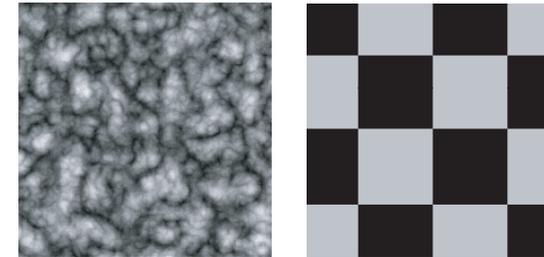


図 5 テクスチャ
Fig. 5 Texture.

5. 実験結果

本節では提案手法の効果を検証するため、さまざまな形状の物体について合成画像実験を行った結果を示す．

合成画像はサイズ 240×240 の多眼ステレオ画像 102 枚を利用した．図 4 に対象の形状を示す．同図 (a) は、表面が滑らかで形状が単純な形状 (Box-1) を示し、(b) は表面が滑らかで凹凸を含む形状 (Bottle)、(c) は表面に角をもつ形状 (Box-2) を示している．それぞれの物体に図 5(a)(b) で示すテクスチャを張り付け、ほぼ全周を網羅するようなカメラ配置を想定して合成画像を作成し、実験を行った．

表 1 各階層レベルにおけるメッシュの構成
Table 1 Mesh at each level.

Mesh level	Number of patches	Number of vertices
Level 0	80	42
Level 1	320	162
Level 2	1280	642
Level 3	5120	2562

同図 (a) は、テクスチャが豊富で比較的推定がしやすいもの (Granite), (b) は格子の境界部分のみしか輝度変化をもたないテクスチャ (Checker) を示している。

サーフェス生成における設定パラメータは, $\alpha = 10.0, \omega_s = 200.0, \omega_h = 1600.0$ とし, 最適化の際の収束条件は, $\|\Delta I\| < 0.015$ とした. また, 階層的メッシュにおける各階層のメッシュの細かさは, 表 1 に示す通りである.

図 6 は, 滑らかな表面形状を持つ Box-1 に豊富なテクスチャ (Granite) を貼り付けた対象に対し, メッシュの各階層における推定結果を示している. 階層的メッシュを利用することにより, 高精細な形状が段階的に復元されていることがわかる.

また, 図 7 は, Box-1 と Bottle に Granite を貼り付けた対象に対し, 拘束項の違いによる推定結果を比較したものである. 同図 (a) と (b) は, Box-1 に対する提案手法の復元結果を, 視点を変えて表示したものであり, 同図 (c) と (d) は, 空間的平滑化のコスト E_s を利用しなかったときの結果である. 空間的平滑化項を利用した提案手法により, 形状のノイズが取り除かれた高精度な復元結果が得られていることが分かる. また, 同図 (e) と (f) は, Bottle に対する提案手法の結果を示し, 同図 (g) と (h) は, 階層的平滑化のコスト E_h を利用しなかったときの結果を表している. 階層的平滑化項により, ロバストな推定結果が得られていることがわかる.

図 8 は, 角をもつ Box-2 に対し, 豊富なテクスチャ (Granite) と格子模様のテクスチャ (Checker) を張り付け, それらに対して復元した結果を示している. 同図 (a) と (b) は, テクスチャ Granite に対する提案手法の結果を異なる視点から表示したものであり, (c) と (d) は, それらの結果に対し, データ項におけるパッチごとの重み $\omega_{d,n}$ の大きさを色温度によって表わしたもので, 暖色であるほど大きく, 寒色であるほど小さい値を示している. 一方, 同図 (e) と (f) は, テクスチャ Checker に対する結果を示しており, 同図 (g) と (h) は, データ項におけるパッチの重みを表わしている. 立方体の稜やテクスチャの少ない領域では重みが小さくなっており, データ項のコストに比べて, 平滑化拘束によるコストをより強く利用

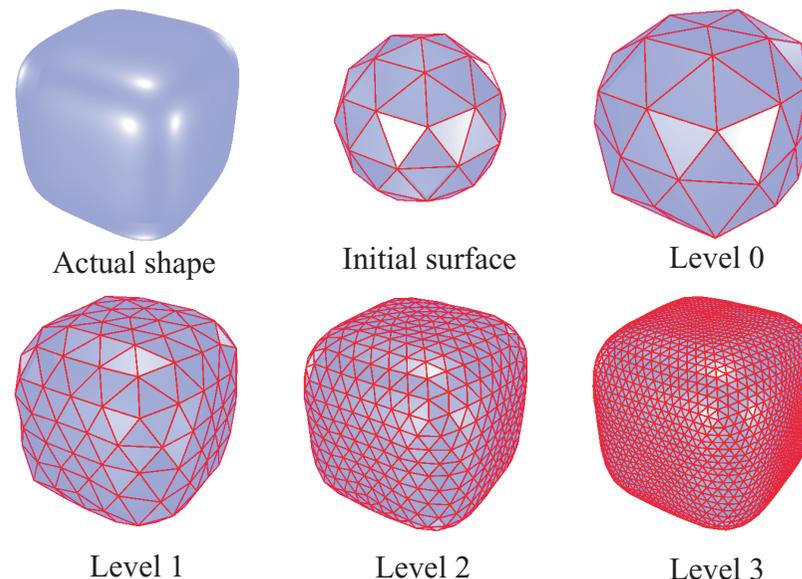


図 6 Box-1 に対する各メッシュ階層における推定結果
Fig. 6 Surface of Box-1 estimated at each mesh level.

し, 良好な復元が実現されている.

プログラムは, C 言語を用いて実装し, Pentium-IV 2.8GHz の Linux PC 上で動作させた. 提案手法では, 最適化において 1 回の繰返しに要した時間は 1 秒程度であった. 推定までの計算時間は, 実形状や初期形状などによって変化するが, 全てのケースで 1 分以内で形状復元ができた. サーフェス生成の既存手法では, 数十枚の画像を利用して, 推定に数時間を要するものが多いことと比べると, 提案手法が非常に高速であることがわかる.

6. おわりに

本論文では, 多眼ステレオカメラで撮影した画像を利用して, 撮影対象のサーフェスモデルを直接法により生成する手法を提案した. 提案手法では, 微小三角メッシュモデルに対し, 空間中のある一点からメッシュ各頂点までの距離をパラメータとして, 画像間の輝度差と形状の滑らかさに基づいたコストを定義した. そして, 各カメラからの各平面の可視性を

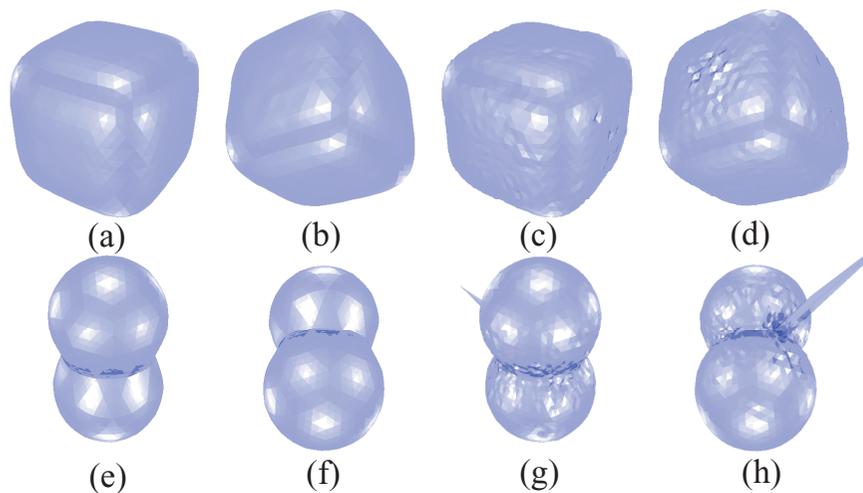


図 7 Box-1 と Bottle に対する拘束項の効果
Fig. 7 Effect of constraint term.

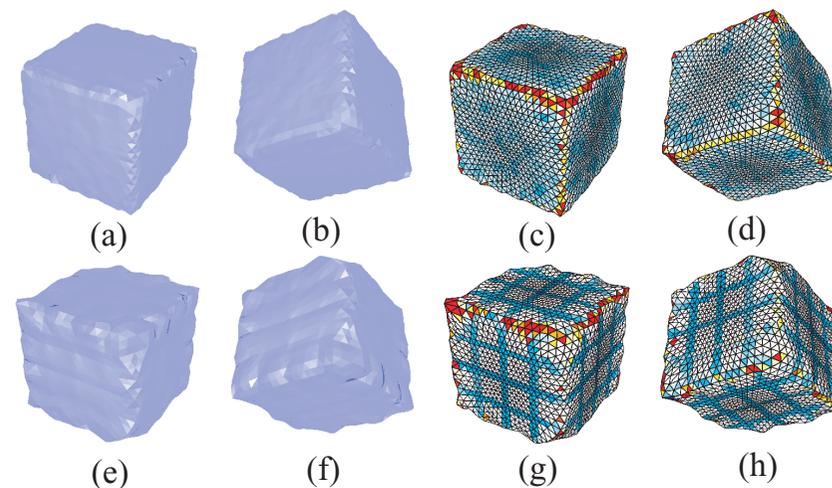


図 8 Box-2 に対する拘束項の効果と各パッチの重み
Fig. 8 Effect of constraint term and patch weight.

考慮しつつ、高速平面パラメータ推定法を拡張した最適化手法によりコストを効率的に最小化した。最小化の際に階層メッシュを利用し、階層間での平面姿勢の類似度をコストに加えることにより、合成画像に対して安定なサーフェス生成を実現した。

今後は、実画像に対する提案手法の有効性を確認することが課題である。また、パッチごとの並列化や、GPUを用いた演算により、より高速にサーフェスを復元する方法等を検討する予定である。

参 考 文 献

- 1) O.D. Faugeras and R.Keriven. Complete dense stereovision using level set methods. In *European Conference on Computer Vision*, Vol.I, pp. 379–393, 1998.
- 2) P.Fua. From multiple stereo views to multiple 3D surfaces. *International Journal of Computer Vision*, Vol.24, No.1, pp. 19–35, 1997.
- 3) P.Fua and Y.G. Leclerc. Object-centered surface reconstruction: Combining multi-image stereo and shading. *International Journal of Computer Vision*, Vol.16, No.1, pp. 35–56, 1995.
- 4) B.F. Gregorski, B.Hamann, and K.I. Joy. Reconstruction of B-spline surfaces from

- scattered data points. In *Computer Graphics International*, pp. 163–170, 2000.
- 5) H.Hoppe, T.DeRose, T.Duchamp, J.McDonald, and W.Stuetzle. Mesh optimization. In *Computer Graphics (SIGGRAPH)*, pp. 19–26, 1993.
- 6) A.Hornung, B.Zeng, and L.Kobbelt. Image selection for improved multi-view stereo. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.
- 7) V.Kolmogorov and R.Zabih. *Visual Correspondence with Occlusions Using Graph Cuts*. PhD thesis, Stanford Univ, 2002.
- 8) A.Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.16, No.2, pp. 150–162, 1994.
- 9) W.E. Lorensen and H.E. Cline. Marching cubes : A high resolution 3d surface construction algorithm. *ACM SIGGRAPH Computer Graphics*, Vol.21, No.4, pp. 163–169, 1987.
- 10) A.Pentland and S.Sclaroff. Closed-form solutions for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.13, pp. 715–729, 1991.
- 11) J.P. Pons, R.Keriven, and O.Faugeras. Modelling dynamic scenes by registering

- multi-view image sequences. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 822–827, 2005.
- 12) J.P. Pons, R. Keriven, and O. Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, Vol.72, No.2, pp. 179–193, 2007.
 - 13) A. Rockwood and J. Winget. Three-dimensional object reconstruction from two-dimensional images. *Computer-Aided Design*, Vol.29, No.4, pp. 279–285, 1997.
 - 14) 杉本茂樹, 奥富正敏. ステレオ画像からの高速な微小平面 3D サーフェス直接生成法. 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol. 48, No. SIG16(CVIM19), pp. 38–50, 2007.
 - 15) G. Vogiatzis, C. Hernandez, P.H.S. Torr, and R. Cipolla. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, No.12, pp. 2241–2246, 2007.
 - 16) G. Vogiatzis, P.H.S. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 391–398, 2005.
 - 17) L. Zhang and S. Seitz. Image-based multiresolution shape recovery by surface deformation. In *SPIE: Videometrics and Optical Methods for 3D Shape Measurement*, pp. 51–61, 2001.