

# 複数シルエットからの制約を利用した奥行エッジに起因する 距離画像復元誤差の解消法

飯山 将晃<sup>†a)</sup>      角所 考<sup>†</sup>      美濃 導彦<sup>†</sup>

Robust Depth Map Acquisition against Depth Edges with Silhouettes Consistency

Masaaki Iiyama<sup>†a)</sup>, Koh Kakusho<sup>†</sup>, and Michihiko Minoh<sup>†</sup>

あらまし 照度差ステレオより得られる法線マップを利用することで視体積交差法では計測することのできない物体の滑らかな曲面や物体の凹面を計測する手法を提案する。法線マップより距離画像を復元する際に奥行エッジに起因する距離画像の誤りが発生する。この誤差を軽減させるために、他方向からのシルエットを利用する。奥行エッジを考慮せずに復元された距離画像は他方向から得られたシルエットとの整合性を欠く場合が多いという性質を利用し、距離画像とシルエットとの整合性及び距離画像と法線マップとの整合性を表すエネルギーを定義し、そのエネルギーを最小とする距離画像を得ることにより距離画像の誤りを軽減させる。シミュレーションデータ及び実物体に対して提案手法を適用した実験により、提案手法の有効性を評価した。

キーワード 視体積交差法, 照度差ステレオ, 奥行エッジ

## 1. ま え が き

物体形状を手軽に計測することに対する要求が高まっている。例えば、個人の人体形状を計測し、それをオーダーメイド衣服のデザインに用いることや、工業製品の設計・開発の際、モックアップの形状を計測し、それを CAD データへと即座に変換することによって開発期間やコストを削減することなどが挙げられる。

従来、物体形状計測手法として普及している光切断法や空間コード化法 [1] などでは、スリットレーザやパターン光照射装置など高価な機材を用意する必要があった。また、そのような機材を用いて得られる形状はある一方向から見た距離画像であり、物体の全周の形状を計測しようとするか回転台などを用意する必要があった。上で述べたように手軽に物体形状を計測すること想定した場合、このような高価な機材を使用することなく物体形状を計測する手法が望ましい。光切断法では精密な形状を計測するためにレーザ走査時間が長くなるという問題や、レー

ザをうまく反射しない物体に対する形状獲得の困難さがあり、空間コード化法でも精密な形状を計測するためには、照射するパターンを密にすることや照射光の拡散によるボケへの対処が必要となることなどの課題がある。

このような状況の中、近年低価格となったカメラを複数台用いて物体形状を計測する手法が注目されている。複数台のカメラを用いた形状計測手法としてはマルチベースラインステレオ [2] や Space Carving [3]、視体積交差法 [4] などがある。

その中で、マルチベースラインステレオや Space Carving は色情報を用いてカメラ間で画素の対応付けをとる手法であり、計測対象となる物体の表面が texture-less である場合、対応付けの誤りによって形状が正確に得られなくなるという問題がある。先ほど例として挙げた、人体やモックアップなどの texture-less な物体への適用にはこれらの手法は適さない。

一方、視体積交差法は複数のカメラから得られたシルエットから物体形状を獲得する手法であり、texture-less な物体に対しても適用可能であるという利点がある。しかしながら、視体積交差法によって得られる形状は物体に外接する凸包となる [5] ため、物体の滑らかな曲面や物体の凹面を計測することは困難若し

<sup>†</sup> 京都大学学術情報メディアセンター, 京都市  
Academic Center for Computing and Media Studies, Kyoto  
University, Kyoto-shi, 606-8501 Japan

a) E-mail: iiyama@media.kyoto-u.ac.jp

くは不可能であるという欠点がある。従来この問題に対し、時系列で得られた視体積を統合することによって仮想的に視点の数を増やし、滑らかな曲面を計測する手法 [6] ~ [9], 色情報に基づいてカメラ間での対応付けをとることにより物体の凹面を計測する手法 [3], [10], [11] が提案されている。時系列の視体積を統合する手法では、滑らかな表面の計測は可能であるが、物体の凹面を計測できないという視体積交差法の原理的な問題は解決できない。また、色情報を用いた手法は、texture-less な物体への適用が可能であるという視体積交差法の利点を損ねてしまうという問題がある。

そこで本研究では、カメラで観測される物体の陰影情報に着目し、物体のシルエットだけではなく陰影情報も利用することで視体積交差法では計測することのできない物体の滑らかな曲面や物体の凹面を計測する手法を提案する。本手法により、texture-less な物体への適用が可能であるという視体積交差法の利点を残しつつ、視体積交差法では計測することのできない物体の滑らかな曲面や物体の凹面を計測することができる。

物体の陰影を利用した形状計測手法として、照度差ステレオ [12] がある。照度差ステレオは三つ以上の異なる方向から照明を照射して得られた (3 枚以上の) 画像組をもとに物体表面の法線方向を法線マップ (needle diagram) として獲得する手法であり、texture-less な物体に対しても適用可能であり、画像間の対応付けを行うことなく物体表面の法線を密に獲得することができる。この法線マップを利用して、視体積交差法では計測することのできない物体の滑らかな曲面や物体の凹面を計測する。

複数台設置したカメラそれぞれで、照度差ステレオによって得られた法線マップをもとに距離画像を復元し、各カメラで得られた距離画像を三次元空間上で位置合せすることで物体形状を計測することができることが Cho ら [13] や Chen ら [14] によって示されている。法線マップから距離画像の復元は、距離画像から計算される物体表面の法線と法線マップとの整合性が高くなるような距離画像を得ることによって実現され、また、距離画像の位置合せは、シルエットとの整合性が高くなるような距離画像のオフセット値を探索することによって実現される。

法線マップから距離画像を復元する際に問題となるのが奥行エッジの存在である。奥行が不連続に変化する

箇所 (奥行エッジ) では距離画像から法線が計算できないため、本来は法線マップとの整合性をとることができない。このとき、奥行エッジを考慮せず、奥行エッジが生じている箇所でも奥行が連続的に変化しているとみなして物体表面の法線を計算し法線マップとの整合性をとってしまうと、奥行エッジ付近で距離画像に誤りが生じてしまう。

そこで、奥行エッジに起因する距離画像の誤りを軽減させるために、他方向からのシルエットを利用することを考える。奥行エッジを考慮せずに復元された距離画像は、他方向から得られたシルエットとの整合性を欠く場合が多い。本手法ではこの性質を利用し、距離画像と法線マップとの整合性及び距離画像とシルエットとの整合性を表すエネルギーを定義し、そのエネルギーを最小とする形状を得る。これにより、各カメラから得られるシルエットとの整合性を保ちつつ、法線マップとの整合性が高くなるような距離画像を獲得し、奥行エッジに起因する距離画像の誤りを軽減させる。

本研究と関連した研究として、Cho らや Chen らが提案した手法では、法線マップをもとに距離画像を獲得し、複数台のカメラからそれぞれ得られた距離画像を三次元空間上で位置合せすることで形状を獲得している。それら関連研究では、距離画像をいかに正確に位置合せするかについて検討されているが、距離画像を獲得する際に発生する奥行エッジの影響については考慮されていない。一方、我々の手法は奥行エッジの影響に対処するという点に特色がある。我々の手法ではシルエットとの整合性を距離画像の位置合せに用いるだけではなく距離画像復元にも用いることで奥行エッジの影響を軽減させており、奥行エッジが存在するような物体形状に対する獲得形状の精度向上が期待される。

本論文の構成は以下のとおりである。まず、2. で提案手法のベースとなる視体積交差法と照度差ステレオについて述べる。3. では距離画像と法線マップとの整合性と距離画像とシルエットとの整合性を表すエネルギーを定義し、そのエネルギーを最小化する手法について述べる。4. では、8 台のカメラを用いて実物体に対して提案手法を適用した実験により、提案手法の有効性を評価する。5. はまとめである。

## 2. 視体積交差法と照度差ステレオ

本章では、提案手法のベースとなる、視体積交差法

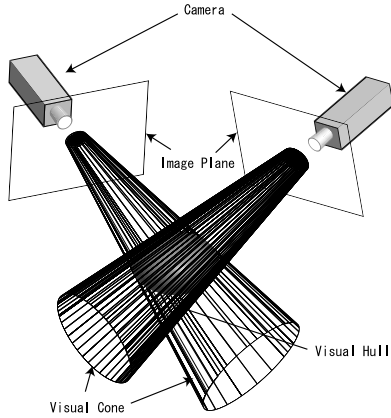


図1 視体積交差法  
Fig.1 Visual hull.

と照度差ステレオについて述べる。

### 2.1 視体積交差法

視体積交差法(図1)について説明する。視体積交差法は、複数台のカメラから得られるシルエットより物体の三次元形状を獲得する手法である。

まず、各カメラで物体を撮影した画像と、あらかじめ物体を配置しない状態で撮影して得られた背景画像との差分に基づいて、画像中で物体が投影されている領域(シルエット)を抽出する。以降、カメラ台数を  $C$  台とし、カメラ  $c(c=1, \dots, C)$  で得られるシルエットを  $S_c$  と表記する。

各カメラがピンホールカメラモデルで表され、その射影行列が既知であるとする。カメラのレンズ中心を端点としシルエットを含む錐体状の開空間(視錐体)がカメラごとに得られる。各カメラから得られた視錐体の積領域を視体積と呼び、これが視体積交差法によって得られる物体の復元形状となる。

視体積交差法には texture-less な物体に対してもシルエットさえ得られれば形状が獲得可能であるという利点がある。しかしながら、視体積交差法によって得られる視体積は、物体の真の形状を内包する凸包であり、物体の凹面を計測することは原理的に不可能である。また、凹面でなくとも物体の滑らかな面を計測しようとする、多くの台数のカメラが必要となる。

### 2.2 照度差ステレオ

視体積交差法では獲得が困難な物体の凹面や滑らかな面を計測する手法として、照度差ステレオを考える。

照度差ステレオは、光源を切り換えながら異なる方向から光源を照射して得られた画像組をもとに、物体

表面の法線を法線マップとして獲得する手法である。原理的には異なる3方向から照射した画像が得られれば、法線方向の推定が可能である。

物体表面の反射特性、光源に対して以下の条件を設定する。

- 光源は平行光源(若しくは平行光源とみなせるほど物体から十分離れた点光源)である。
- 光源の照射方向が既知であり、少なくとも三つの照射方向ベクトルが一次独立である。
- キャストシャドウが生じない。
- 物体表面は拡散反射面である。

カメラ  $c$  について、物体に対してカメラ  $c$  と同じ側から照射している光源の数  $N_c$ 、各光源の照射方向(物体表面から光源へ向かうベクトル)  $\mathbf{l}_j^c (j=1, \dots, N_c)$  が既知であるものとする。

照射方向  $\mathbf{l}_j^c$  から入射した光は物体表面上で拡散反射しカメラで観測される。そのとき、 $S_c$  に含まれる  $M_c$  個のピクセル中のあるピクセル  $m_i^c (i=1, \dots, M_c)$  で観測される反射光の強さ  $I_{ij}^c$  は、光源から物体表面に入射する入射光の強さを  $L_j^c$ 、ピクセル  $m_i^c$  で観測される物体表面における法線を  $\mathbf{n}_i^c$ 、拡散反射係数を  $\rho_i^c$  とすると、以下の式で表される。

$$I_{ij}^c = L_j^c \rho_i^c \frac{\mathbf{l}_j^c \cdot \mathbf{n}_i^c}{|\mathbf{l}_j^c| |\mathbf{n}_i^c|} \quad (1)$$

ここで、 $I_{ij}^c$  を並べて作成した  $M_c$  行  $N_c$  列の行列  $\mathbb{I}$  を考えたとき、 $\mathbb{I}$  は式(2)で表されるように、物体表面の法線を含む  $M_c$  行3列の行列  $\mathbb{N}$  と光源の照射方向を含む3行  $N_c$  列の行列  $\mathbb{L}^T$  で構成され、 $\mathbb{I}$  を  $\mathbb{N}$  と  $\mathbb{L}^T$  に分解することによりピクセル  $m_i^c$  が観測する物体表面の法線  $\mathbf{n}_i^c$  を得ることができる。

$$\mathbb{I} = \begin{pmatrix} I_{11}^c & \dots & I_{1N_c}^c \\ \vdots & \ddots & \vdots \\ I_{M_c 1}^c & \dots & I_{M_c N_c}^c \end{pmatrix} = \mathbb{N} \mathbb{L}^T = \begin{pmatrix} \rho_1^c \frac{\mathbf{n}_1^{cT}}{|\mathbf{n}_1^c|} \\ \vdots \\ \rho_{M_c}^c \frac{\mathbf{n}_{M_c}^{cT}}{|\mathbf{n}_{M_c}^c|} \end{pmatrix} \begin{pmatrix} L_1^c \frac{\mathbf{l}_1^c}{|\mathbf{l}_1^c|} & \dots & L_{N_c}^c \frac{\mathbf{l}_{N_c}^c}{|\mathbf{l}_{N_c}^c|} \end{pmatrix} \quad (2)$$

### 2.3 法線マップからの距離画像復元

照度差ステレオには、先に述べた物体表面の反射特性・光源に対する条件さえ満たせば texture-less な物

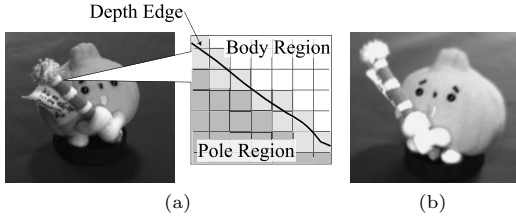


図 2 奥行エッジ  
Fig. 2 Depth edge.

体に対しても法線マップを得ることができるという利点がある．しかしながら，照度差ステレオによって得られるのは物体表面の法線を表す法線マップであり，物体形状そのものではない．そのため，後処理として法線マップから距離画像を復元し，物体形状を獲得する処理が必要となる．

なお，法線マップから距離画像の復元は，距離画像から計算される物体表面の法線と法線マップとの整合性が高くなるような距離画像を決定することにより実現される．その際，奥行エッジが生じる個所については形状が不連続なため，本来法線が計算できず，法線マップとの整合性を取ることができないことに注意する必要がある．

例えば，図 2(a) において，マスコットの胴体とマスコットがもつ棒の部分とは不連続であるため，胴体と棒が隣り合う個所で法線を計算することはできない．法線マップ上でも，胴体と棒が隣り合う個所では二つの異なる面が同一の画素に投影されるため，そこで得られる法線は誤りを含むものとなる．このとき，奥行エッジを考慮せずに胴体と棒が連続しているとみなし，誤りを含んだ法線をもとに距離画像を復元すると，胴体と棒が滑らかにつながるような形状が得られてしまう．

### 3. シルエット・法線マップとの整合性の利用

前章で述べた奥行エッジの影響を軽減させるため，他方向からのシルエットを利用することを考える．

例として，図 2(a) では胴体と棒が隣り合う個所が，別方向から観測した画像（図 2(b)）では隣り合っていない．図 2(a) を用い，本来連続していない面を連続しているとみなして得られた形状を別方向から観測すると，図 2(b) との整合性がとれなくなることが予想される．

そこで，距離画像と法線マップとの整合性，距離画

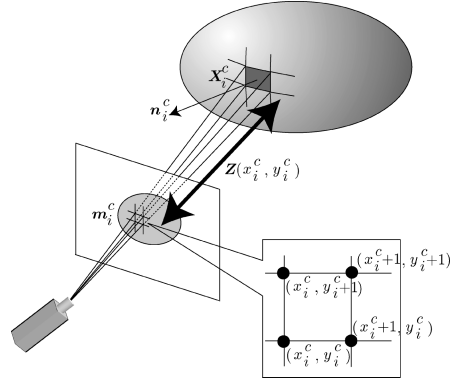


図 3 法線マップとの整合性  
Fig. 3 Needle map vs surface normal.

像とシルエットとの整合性をそれぞれエネルギー関数として定義し，これら二つのエネルギー関数の和を最小化することによって，法線マップ・シルエットそれぞれと整合性のとれた距離画像を生成し，奥行エッジの影響を軽減させる．

以降，シルエット  $S_c$  に含まれるピクセル  $m_i^c$  が占めている画像上の正方領域を  $([x_i^c, x_i^c + 1], [y_i^c, y_i^c + 1])$  とし，点  $(x_i^c, y_i^c)$  を  $m_i^c$  の代表点と呼ぶ．距離画像は，ピクセル  $m_i^c$  の代表点に投影される物体表面上の点  $X_i^c$  からカメラの焦点までの奥行  $Z(x_i^c, y_i^c)$  を推定することにより得られる．

#### 3.1 法線マップとの整合性

まず，距離画像と法線マップとの整合性について考える．

距離画像のあるピクセル  $m_i^c$  における奥行  $Z(x_i^c, y_i^c)$  と，そのピクセルが観測する物体表面の法線  $\mathbf{n}_i^c = (p_i^c, q_i^c, -1)^T$  が与えられたとしよう（図 3）．

$m_i^c$  に投影される物体表面が法線  $\mathbf{n}_i^c$  の平面であるとすると，距離画像上での三つの点  $(x_i^c + 1, y_i^c), (x_i^c, y_i^c + 1), (x_i^c + 1, y_i^c + 1)$  における奥行  $Z_{\pm 0}, Z_{\pm 0}, Z_{-1}$  は，ピクセルの代表点  $(x_i^c, y_i^c)$  での奥行  $Z(x_i^c, y_i^c)$ ，法線  $\mathbf{n}_i^c$ ，そしてカメラ  $c$  の焦点距離  $f_c$  を用いてそれぞれ以下のように表される．

$$Z_{\pm 0}(x_i^c + 1, y_i^c) = Z(x_i^c, y_i^c) \left( 1 + \frac{p_i^c}{f_c} \right) \quad (3)$$

$$Z_{-1}(x_i^c, y_i^c + 1) = Z(x_i^c, y_i^c) \left( 1 + \frac{q_i^c}{f_c} \right) \quad (4)$$

$$Z_{-1}(x_i^c + 1, y_i^c + 1) = Z(x_i^c, y_i^c) \left( 1 + \frac{p_i^c}{f_c} + \frac{q_i^c}{f_c} \right) \quad (5)$$

このことより、距離画像上の点  $(x_i^c, y_i^c)$  に対応する奥行は、 $Z(x_i^c, y_i^c)$ 、式 (3) を用いて  $(x_i^c - 1, y_i^c)$  から計算される奥行、式 (4) を用いて  $(x_i^c, y_i^c - 1)$  から計算される奥行、そして式 (5) を用いて  $(x_i^c - 1, y_i^c - 1)$  から計算される奥行の 4 通りの計算方法によって得ることができる。そこで、これら 4 通りの方法で得られる奥行が一致すればするほど、距離画像と法線マップとの整合性が高いことに基づき、距離画像と法線マップとの整合性を表すエネルギーを次式で定義する。なお、整合性が高いほどエネルギーは小さくなる。

$$E_N = \sum_{m_i^c \in S_c} \left( \left| Z(x_i^c, y_i^c) - Z_{\pm 0}(x_i^c, y_i^c) \right|^2 + \left| Z(x_i^c, y_i^c) - Z_{-1}(x_i^c, y_i^c) \right|^2 + \left| Z(x_i^c, y_i^c) - Z_{-1}(x_i^c, y_i^c) \right|^2 \right) \quad (6)$$

### 3.2 シルエットとの整合性

次に、距離画像とシルエットとの整合性について考える。シルエットと整合性のとれた形状を獲得する手法として、2.1 で述べた視体積交差法がある。視体積は物体に外接する形で存在し、視体積をカメラに投影して得られる投影像とシルエットとが一致するという性質が Laurentini [5] によって示されており、この性質を利用してカメラ間の対応付けをとる手法 [7] や、視体積構築に必要なシルエットの抽出誤差に対処する手法 [15] が提案されている。我々は奥行エッジに起因する距離画像の復元誤差を解消するために、視体積のこの性質を利用し、距離画像とシルエットとの整合性を評価する。

#### 3.2.1 視線分制約

シルエットに含まれるピクセルのうち、シルエットと背景との境界に位置するピクセル（境界ピクセル）に対し、カメラの焦点を端点とし境界ピクセルの代表点を通る視線を考える（図 4）。なお、ここでは境界ピクセルの数が多いほど制約として用いる視線の数が増え、より正確に形状が獲得できるとの考えのもと、8 近傍にシルエットに含まれないピクセルが存在するようなシルエット中のピクセルを境界ピクセルとする。

この視線を他のカメラへ投影すると、その投影像

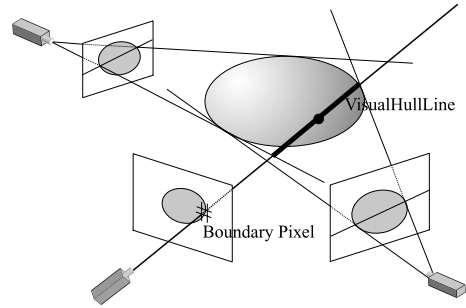


図 4 視線分  
Fig. 4 Visual hull line.

（投影像も直線となる）はそのカメラでのシルエットと交差する。ここで、ある境界ピクセルの代表点を通る視線のうちその投影像が他のすべてのカメラでのシルエットに含まれるような区間を考えこれを視線分と呼ぶ。視線分は（最大で 1 画素の大きさ分の誤差はあるが）視体積の表面の一部となる。

この視線分と物体形状との関係を考えてとき、視体積の投影像とシルエットとが一致するという性質から、以下の二つの制約（視線分制約）が得られる。

- 物体はどの視線分とも交差しない。
- 物体は各々の視線分に対し、少なくとも 1 点で接する。

一つ目の制約が満たされない場合、視線分は視体積の表面であることから、視体積が物体に外接しているという制約が満たされなくなる。また、二つ目の制約が満たされない場合、この視線線上には物体は存在しない、つまり、この視線を構成する境界ピクセルがシルエットに含まれないこととなり、視体積とシルエットとの整合性がとれなくなる。

距離画像とシルエットとの整合性は、この視線分制約をどの程度満たしているかによって評価することができる。

#### 3.2.2 視線分制約の利用

ここで、視線分のうち、どの視線分が距離画像に対して制約を与えるかを考える。視線分の中には、あるカメラから観測できるものもあれば、他の視線分によって隠されそのカメラからは観測されないものもあり、視線分がそのカメラからどの程度観測できるのかによって、そのカメラから得られる距離画像に対して制約を与えるかどうか異なる。

ある視線分について、その全体がカメラ A から観測可能である場合（図 5(a)）、その視線分はカメラ A

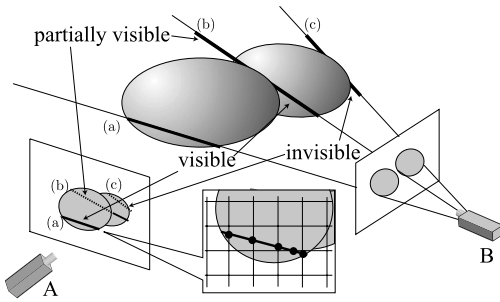


図 5 視線分の観測可能性  
Fig. 5 Visible/Invisible visual hull line.

での距離画像に対して二つの視線分制約の両方を与える．一方、視線分がカメラから観測不可能である場合（図 5 (c)）、距離画像に対して何の制約も与えない．また、視線分の一部だけが観測可能である場合（図 5 (b)）、視線分制約の一つ目の制約（物体は視線分と交差しないという制約）については距離画像に対して与えられるが、二つ目の制約（物体は視線分の少なくとも 1 点で接するという制約）については、視線分のうちカメラから観測されない部分で物体と接している可能性があるため、距離画像に対する制約として利用することができない．なお、ある視線分が観測可能であるかどうかは、いったんすべての視線分を画像に投影することによって Z バッファを構築し、Z バッファとの比較を行うことによって判定することができる．

以上のことを踏まえ、視線分を距離画像上に投影し、距離画像の画素格子との交点（図 5）を抽出し、視線分制約をどの程度満たしているかどうかを、カメラから視線分までの奥行と距離画像から得られるカメラから物体表面までの奥行を各交点において比較することによって評価する．

視線分のうち観測可能な部分の投影像と画素格子との交点を  $(v_{x1}, v_{y1}), \dots, (v_{xK}, v_{yK})$  とする．ある交点  $(v_{xk}, v_{yk})$  におけるカメラから物体表面までの奥行  $Z(v_{xk}, v_{yk})$  は、代表点を  $([v_{xk}], [v_{yk}])$  とする画素がもつ奥行  $Z([v_{xk}], [v_{yk}])$  と、その画素に対応する法線  $\mathbf{n}_k = (p_k, q_k, -1)^T$  を用いて以下の式で計算することができる．

$$Z(v_{xk}, v_{yk}) = Z([v_{xk}], [v_{yk}]) \cdot \left( 1 + \frac{v_{xk} - [v_{xk}]}{f_c} p_k + \frac{v_{yk} - [v_{yk}]}{f_c} q_k \right) \quad (7)$$

視線分制約の一つ目の制約より、視線分のうち観測

可能な部分については視線分より手前には物体表面は存在しないため、すべての交点において、視線分の奥行の方が物体表面までの奥行よりも小さくなる．

また、視線分制約の二つ目の制約よりある視線分全体が観測可能であれば、交点のうち少なくとも 1 点で、視線分の奥行と物体表面までの奥行が一致する．

式 (7) によって得られるカメラから物体表面までの奥行と、交点  $(v_{xk}, v_{yk})$  におけるカメラから視線分までの奥行との差を  $\delta Z_k, \delta Z_k$  の中で最小となるものを  $\delta Z_{min}$  とすると、視線分制約の一つ目の制約より、すべての  $k$  に対して  $\delta Z_k \geq 0$  となり、また、視線分制約の二つ目の制約より、視線分全体が観測可能であれば  $\delta Z_{min} = 0$  となる．以上のことを踏まえ、この視線分が距離画像に与える制約に基づくエネルギーを以下の式で定義する．

$$E_{VL} = \begin{cases} 0 & \text{if } \Delta Z_{min} \geq 0 \text{ かつ} \\ & \text{視線分の一部が観測不可} \\ \Delta Z_{min}^2 & \text{otherwise} \end{cases} \quad (8)$$

これをすべての視線分に対して合計したものが距離画像とシルエットとの整合性に基づくエネルギーとなる．

### 3.3 エネルギー最小化

3.1 で定義した距離画像と法線マップとの整合性に基づくエネルギーと 3.2 で定義した距離画像とシルエットとの整合性に基づくエネルギーとの和  $E_{all}$  をアニーリング法 [16] によって最小化し、距離画像を得る．

$$E_{all} = E_N + \lambda E_{VL} \quad (9)$$

ここで  $\lambda$  は視線分制約の重みである． $\lambda$  に十分大きな値を与えておくことによって、視線分制約を満たす範囲内で法線マップとできるだけ整合するような距離画像が得られる．これにより、奥行エッジが生じている部分ではシルエットとの整合性を重視して奥行値が推定され、奥行エッジが生じていない部分では法線マップとの整合性に基づいて奥行値が推定される．

アニーリング法では、 $S_c$  に含まれるピクセルをランダムに選択し、選択されたピクセルの奥行値を変動させたときのエネルギー  $E_{all}$  の変化を評価する．エネルギーが減少するか、若しくはエネルギーの増加がアニーリングスケジュールによって定められた値（温度）以下である場合には、選択されたピクセルの奥行値を更新する．これを温度を少しずつ低下させながら

繰り返すことにより、 $E_{all}$  を最小化する。

### 3.4 距離画像の統合

各カメラについて得られた距離画像を統合することにより三次元形状を得る。ピクセルの代表点  $(x_i^c, y_i^c)$  に対応する物体表面の三次元位置  $\mathbf{X}_i^c$  は、奥行  $Z(x_i^c, y_i^c)$  とカメラ  $c$  の焦点距離  $f_c$ 、及びカメラのレンズ中心  $(e_x^c, e_y^c)$  を用いて次式で表される。

$$\mathbf{X}_i^c = \begin{pmatrix} \frac{Z(x_i^c, y_i^c)}{f_c} (x_i - e_x^c) \\ \frac{Z(x_i^c, y_i^c)}{f_c} (y_i - e_y^c) \\ Z(x_i^c, y_i^c) \end{pmatrix} \quad (10)$$

式 (10) を用いて距離画像から計算される物体表面の三次元位置はカメラの焦点を原点としたカメラ座標系であるため、カメラ  $c$  におけるカメラ座標系から世界座標系への変換を表す回転行列  $R_c$  と平行移動ベクトル  $\mathbf{T}_c$  を用いて  $\widetilde{\mathbf{X}}_i^c = R_c \mathbf{X}_i^c + \mathbf{T}_c$  とすることによって世界座標系での物体表面の三次元位置  $\widetilde{\mathbf{X}}_i^c$  を得る。

## 4. 実 験

提案手法の有効性を検証するため、形状が既知のシミュレーションデータを用いた実験と、実物体を用いた実験を行った。

図 6(a), (b), (c) で示される恐竜のシミュレーションデータに対して 20 方向から照明を照射した様子で、この対象を取り囲むように配置した 9 台の仮想カメラ (640 × 480 画素) で撮影した画像を入力とした。

図 6 の (e) ~ (j) はそれぞれシルエットとの整合性のエネルギー  $E_{VL}$  のみ (図 6(e), (h)), 法線マップとの整合性のエネルギー  $E_N$  のみ (図 6(f), (i)), シルエットとの整合性と法線マップとの整合性のエネルギーの和  $E_{all}$  (図 6(g), (j)), をあるカメラ (図 6(d)) について最小化して得られた距離画像より復元される物体形状を表している。また、図 6(e) ~ (g) は図 6(a), 図 6(h) ~ (j) は図 6(c) と同じアングルで生成した画像である。

図 6(e), (h) と図 6(g), (j) より、シルエットとの整合性のみを考慮した場合と比較して、法線マップの整合性を用いた方が滑らかな面を復元できていることが確認できる。また、画像 (図 6(d)) で左右の脚の境界部分で奥行エッジが生じている。その結果、シルエットの整合性を考慮しない場合、左右の脚が隣り合う箇所が滑らかにつながるように形状が構築され、本来の形状と異なってしまう。図 6(f), (i) はシルエットの整合性を用いなかったときの復元結果であるが、本来

接触していない右脚と左脚が図 6(f), (i) では接触してしまっていることが図 6(a), (c) との比較から確認できる。それに対し、シルエットの整合性を考慮した場合、奥行エッジの大部分が他のカメラでのシルエットとの整合性をとることにより検出される。図 6(f), (i) では接触してしまっている右脚と左脚が、シルエットとの整合性を用いて復元された図 6(g), (j) では離れていることが確認され、奥行エッジを有する物体に対する本手法の有効性が確認できる。

獲得された距離画像中の各画素  $m_i^c$  に対し、奥行値の推定誤差を計算した。それを対数ヒストグラムで表したものを図 6(k) に、推定誤差の平均値、及び、誤差が 0.01, 0.02, 0.05 以下となる画素の総数をシルエット全体の画素総数で割ったものを表 1 に示す。なお、奥行値の推定誤差の大きさを、物体の全長を 1 としたときの長さで表現している。例えば誤差 0.01 とは、物体の全長の 1% に相当する長さ分だけ奥行誤差が生じていることを示している。

シルエットとの整合性のみを利用した場合、画素の半数以下しか推定誤差 0.01 の範囲に収まらなかったのに対し、法線マップとの整合性を利用することにより、85% の画素において推定誤差が 0.01 の範囲に収まっていることが確認でき、法線マップの利用による精度向上が確認できる。また、法線マップとの整合性のみを用いた結果と提案手法を用いた結果とを比較すると、提案手法の方が全般的に良い精度で距離画像を獲得できていることが表 1 より明らかになっている。また、0.05 以上の誤差をもつ画素数が法線マップとの整合性のみを用いた結果と比較して大きく減少していることが図 6(k) から読み取れる。これは、左右の脚が隣り合う箇所が滑らかにつながるように構築されていた形状が提案手法によって改善されたことに起因するものであると考えられる。

次に、実物体を用いて実験を行った。90 cm 立方のフレームの中央に、計測対象となる幅約 4 cm の物体を設置し、640 × 480 画素のカラー CCD カメラ (PointGray 社 Dragonfly) を物体の前方側のフレームに 4 台、後方側のフレームに 4 台の計 8 台設置し、物体の前方側のフレームに 12 個、後方側のフレームに 12 個の計 24 個の白熱電球 (40 W) を設置した。白熱電球は点光源であるが照射方向のぶれは約 3° ほどであり、ほぼ平行光源とみなすことができる。

24 個の照明を順に一つずつ点灯させ、その様子を 8 台のカメラで観測することにより照度差ステレオに必

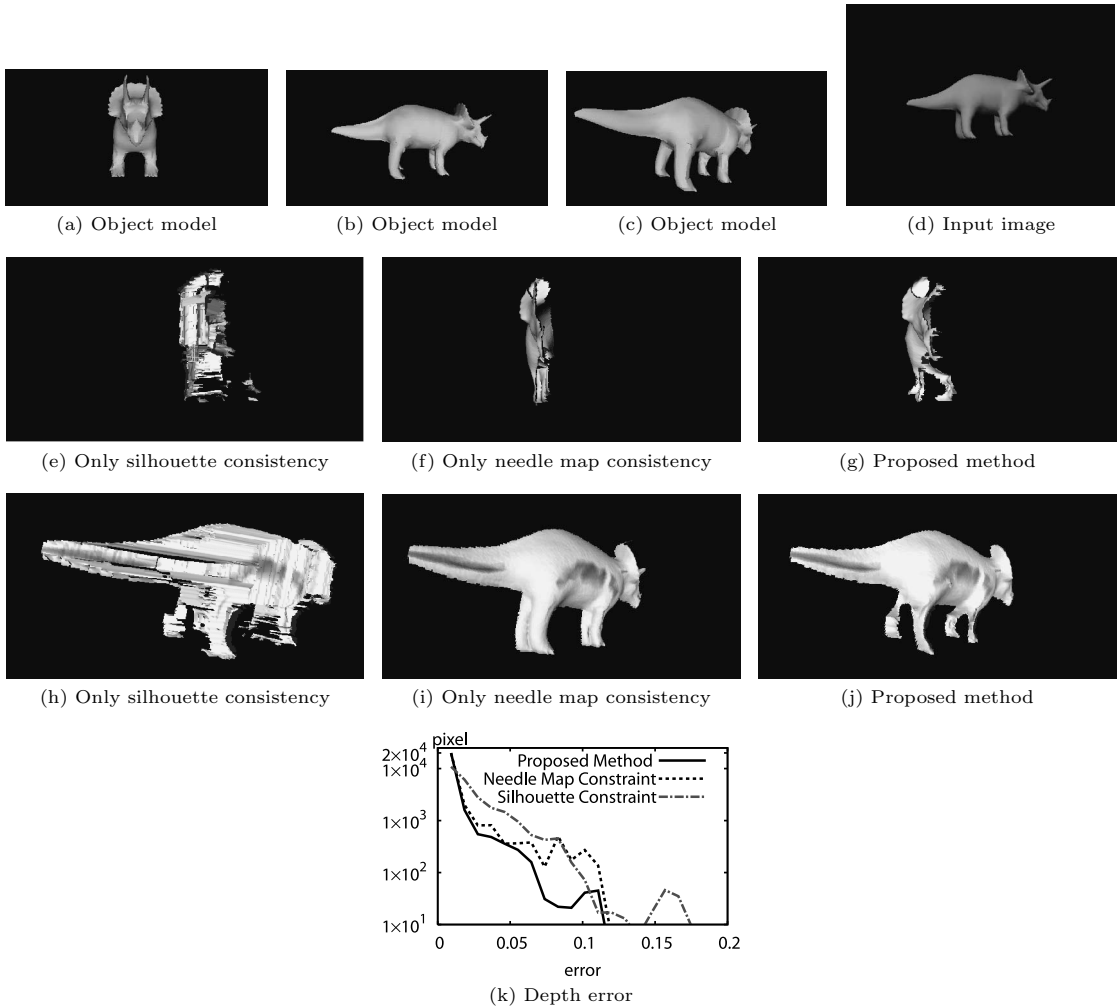


図 6 距離画像から復元された部分形状 (シミュレーションデータ)  
Fig. 6 Partial shape from a depth map. (synthetic data)

表 1 奥行値の推定誤差  
Table 1 Depth error.

	シルエットとの整合性のみ	法線マップとの整合性のみ	提案手法
平均誤差	0.019	0.011	0.0063
誤差 0.01 以下	44.6%	78.5%	85.0%
誤差 0.02 以下	68.1%	85.2%	91.6%
誤差 0.05 以下	91.1%	93.0%	97.6%

要な画像列を得た。また、視体積交差法に必要なシルエットは背景差分法を用いて抽出した。

マスコットを計測対象とし実験を行った。あるカメラ (図 7(a)) について、 $E_{VL}$  のみ (図 7(d), (g)),  $E_N$  のみ (図 7(e), (h)),  $E_{all}$  (図 7(f), (i)) のそれ

ぞれを最小化することによって得られた距離画像をもとに物体形状を復元した結果を図 7 に示す。図 7(b), (c) はマスコットを図 7(d) ~ (f) 及び (g) ~ (i) とほぼ同じアングルから撮影したものである。なお、復元された距離画像に対して、図 7(a) のカメラで観測されたテクスチャを図 7(d) ~ (i) に貼り付けている。

図 7(d), (g) と図 7(f), (i) より、シミュレーションデータを用いた実験結果と同様、シルエットとの整合性のみを考慮した場合と比較して、法線マップの整合性を用いた方が滑らかな面を復元できていることが確認できる。また、画像上でマスコットの胴体とマスコットがもつ棒が隣り合う個所で奥行エッジが生じている。この奥行エッジによる影響をシルエットとの整





図 7 距離画像から復元された部分形状 (実物体)  
 Fig. 7 Partial shape from a depth map. (real data)

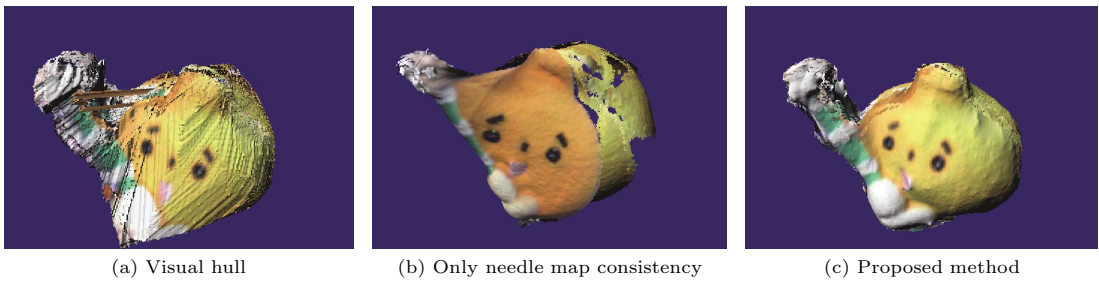


図 8 復元形状  
 Fig. 8 Recovered shape.

合性を利用することによって対処できていることが、  
 図 7(e), (h) と図 7(f), (i) 及び図 7(b), (c) との比較により確認できる。

また、各カメラから得られた距離画像を統合した結果を図 8(c) に示す。シルエットとの整合性しか利用しない視体積交差法 (図 8(a)) では獲得することの困難な、物体の滑らかな表面や、凹んだ表面が提案手

法によって計測できていることが確認できる。

また、シルエットとの整合性を考慮せずに獲得された物体形状 (図 8(b)) と、提案手法による獲得結果 (図 8(c)) との比較により、奥行エッジが発生している個所付近で獲得形状が異なっていることが確認でき、奥行エッジが存在する物体に対する提案手法の有効性が示された。

## 5. む す び

本研究では、物体のシルエットに加え、カメラで観測される物体の陰影情報に着目し、陰影情報を利用することで視体積交差法では計測することのできない物体の滑らかな曲面や物体の凹面を計測する手法を提案した。照度差ステレオによって得られた法線マップより距離画像を生成する際、奥行エッジを考慮しないと、奥行エッジが発生する個所付近で距離画像に誤差が生じる。奥行エッジを考慮せずに復元された距離画像は、他方向から得られたシルエットとの整合性を欠く場合が多いという性質を利用し、距離画像とシルエットとの整合性及び距離画像と法線マップとの整合性を表すエネルギーを定義し、そのエネルギーを最小とする形状を得ることにより距離画像の誤りを軽減させた。

今後の課題として、異なるカメラから得られた距離画像同士の重ね合わせの問題が挙げられる。距離画像同士の重なりをずらす原因として、法線の推定誤差やカメラ位置の推定誤差、シルエットの抽出誤差などが考えられる。これらの誤差をいかに軽減させるかについて今後研究する必要がある。また、異なる方向から計測された距離画像の重ね合わせを行う研究が既にいくつか行われており、それらの手法がどの程度本手法に適用可能か今後実験していく必要がある。

## 文 献

- [1] 佐藤宏介, 井口征士, “空間コード化による距離画像入力,” 信学論 (D), vol. J68-D, no. 3, pp. 369–375, March 1985.
- [2] M. Okutomi and T. Kanade, “A multiple-baseline stereo,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 15, no. 4, pp. 353–363, 1993.
- [3] K.N. Kutulakos and S.M. Seitz, “A theory of shape by space carving,” Int. J. Comput. Vis., vol. 38, no. 3, pp. 199–218, 2000.
- [4] W.N. Martin and J.K. Aggarwal, “Volumetric descriptions of objects from multiple views,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 5, no. 2, pp. 150–158, 1983.
- [5] A. Laurentini, “How far 3D shapes can be understood from 2D silhouettes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 2, pp. 188–195, 1995.
- [6] S. Vedula, S. Baker, S. Seitz, and T. Kanade, “Shape and motion carving in 6D,” Proc. Computer Vision and Pattern Recognition, vol. 2, pp. 592–598, 2000.
- [7] K.M. Cheung, S. Baker, and T. Kanade, “Visual hull alignment and refinement across time: A 3D reconstruction algorithm combining shape-from-silhouette with stereo,” Proc. Computer Vision and Pattern Recognition, pp. 375–382, 2003.

## 付 録

- [8] 豊浦正広, 飯山将晃, 角所 考, 美濃導彦, “視体積交差法における時系列画像の統合による三次元復元形状の再現性の向上,” 信学論 (D-II), vol. J88-D-II, no. 8, pp. 1530–1538, Aug. 2005.
- [9] 築澤宗太郎, 杉本晃宏, 松山隆司, “装着型視覚センサを用いた対象の 3 次元形状復元,” 情処学研報, CVIM, vol. 2002, no. 34, pp. 71–78, 2002.
- [10] 富山仁博, 折原 豊, 岩館祐一, “多視点画像を用いた高精度 3 次元モデルの高速生成手法,” 画像の認識・理解シンポジウム 2004, vol. II, pp. 85–90, 2004.
- [11] 延原章平, 和田俊和, 松山隆司, “弾性メッシュモデルを用いた多視点画像からの高精度 3 次元形状復元,” 情処学 CVIM 研報, vol. 43, SIG11, pp. 53–63, 2002.
- [12] R.J. Woodham, “Photometric method for determining surface orientation from multiple images,” Opt. Eng., vol. 19, no. 1, pp. 139–144, 1980.
- [13] C. Cho and H. Minamitani, “3-D reconstruction using photometric stereo and silhouette informations,” IECON’94, vol. 2, pp. 806–809, 1994.
- [14] C.Y. Chen, R. Klette, and C.F. Chen, “3D reconstruction using shape from photometric stereo and contours,” Proc. IVCNZ, pp. 251–255, 2003.
- [15] K. Forbes, A. Voigt, and N. Bodika, “Using silhouette consistency constraints to build 3D models,” Proc. 14th Annual Symposium of the Pattern Recognition Association of South Africa (PRASA 2003), pp. 33–38, 2003.
- [16] W. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes in C, Cambridge University Press, 1992.

## 式 (3), (4), (5), (7) の補足

カメラの画角が狭く、また、カメラと対象とが十分離れていると仮定し、カメラの投影モデルを弱中心射影で近似する。カメラの焦点から  $Z(x_i^c, y_i^c)$  離れた平面上では、画像上での 1 ピクセル幅に相当する長さが  $\frac{Z(x_i^c, y_i^c)}{f_c}$  に相当する<sup>(注1)</sup>。  $(x_i^c, y_i^c)$  に投影され、かつ、カメラの焦点から  $Z(x_i^c, y_i^c)$  離れた点  $\mathbf{X}_i^c$  と、  $(x_i^c + \alpha, y_i^c + \beta)$  に投影される点  $\bar{\mathbf{X}}_i^c$  が法線ベクトル  $\mathbf{n}_i^c = (p_i^c, q_i^c, -1)^T$  で表される同一平面上に存在する場合、  $(\mathbf{X}_i^c - \bar{\mathbf{X}}_i^c)\mathbf{n} = 0$  が成り立つ。

画像上で 1 ピクセル幅に相当する長さが  $\frac{Z(x_i^c, y_i^c)}{f_c}$  に相当することをを用いると、点  $\bar{\mathbf{X}}_i^c$  の  $X$  座標・ $Y$  座標は  $\mathbf{X}_i^c$  からそれぞれ  $\frac{Z(x_i^c, y_i^c)}{f_c}\alpha$ ,  $\frac{Z(x_i^c, y_i^c)}{f_c}\beta$  離れた座標となる。この結果と  $(\mathbf{X}_i^c - \bar{\mathbf{X}}_i^c)\mathbf{n} = 0$  より、  $(x_i^c + \alpha, y_i^c + \beta)$  に投影される点の奥行は  $Z(x_i^c, y_i^c)(1 + \frac{p_i^c}{f_c}\alpha + \frac{q_i^c}{f_c}\beta)$  で

(注1): 画角が大きい場合、画像の端の方ではこの近似は成り立たないが、本手法で想定しているような、物体から離れたところから狭い画角で撮影しているような環境ではこの近似が成り立つ。

与えられる。

(平成 17 年 11 月 14 日受付, 18 年 3 月 3 日再受付)



飯山 将晃 (正員)

平 15 京大大学院情報学研究科博士課程指導認定退学。同年京都大学学術情報メディアセンター助手。コンピュータビジョン・3次元データ処理の研究に従事。IEEE, 情報処理学会各会員。



角所 考 (正員)

平 5 阪大大学院工学研究科博士課程了。平 5~6 スタンフォード大客員研究員。平 6 大阪大学助手。平 9 京都大学総合情報メディアセンター助教授。平 4 同学術情報メディアセンター助教授。工博。視覚メディア処理, コミュニケーション関係の研究に従事。



美濃 導彦 (正員)

昭 53 京大・工・情報工学卒。昭 58 同大大学院博士課程了。同年工学部助手, 昭 62~63 マサチューセッツ州立大学客員研究員, 平元 京都大学工学部付属高度情報開発実験施設助教授, 平 9 京都大学総合情報メディアセンター教授, 平 14 京都大学学術情報メディアセンター教授。画像処理, 人工知能, 知的コミュニケーション関係の研究に従事。工博。IEEE, ACM, 情報処理学会, 画像電子学会, 日本ロボット学会各会員。