

## 応用論文

# ハンドヘルドビデオカメラを用いた撮影支援インタフェースを有するインタラクティブ三次元モデリングシステム

不殿 健治<sup>\*1</sup> 佐藤 智和<sup>\*1</sup> 横矢 直和<sup>\*1</sup>

Interactive 3-D Modeling System with Capturing Support Interface  
Using a Hand-held Video Camera

Kenji Fudono<sup>\*1</sup>, Tomokazu Sato<sup>\*1</sup> and Naokazu Yokoya<sup>\*1</sup>

**Abstract** – Recently, a number of methods for 3-D modeling from images have been developed. However, the accuracy of a reconstructed model depends on camera positions and postures with which the images are obtained. In most of conventional methods, some skills for adequately controlling the camera movement are needed for users to obtain a good 3-D model. In this study, we propose an interactive 3-D modeling interface in which special skills are not required. This interface consists of “indication of camera movement” and “preview of reconstruction result.” In experiments, we have confirmed this interface makes it possible for users to obtain a good 3-D model easily without a special training.

**Keywords** : interactive 3-D modeling, user interface, hand-held video camera, shape from silhouette

## 1 はじめに

近年、高性能で安価なグラフィクスチップが普及し、一般的に広く利用されている PC や家庭用ゲーム機においても高品位な三次元 CG を表示することが可能となっている。また、エンターテインメントやインターネット広告の分野では、これら三次元 CG のコンテンツとして実物体の三次元モデルが用いられ始めている。これら実物体の三次元モデルを Web ページやインターネットオークションに用いれば、二次元の写真では分からない物体の詳細形状や質感といったものを比較的容易に表現できるため、一般的な PC の利用者においても実物体の三次元モデル利用に対する需要が高まっている。しかし、現在、このような実物体の三次元モデルの作成は、主に三次元モデルの計測に関する専門知識や専門機材を持つ熟練者によって行われており、専門知識や機材を持たない一般の利用者が実物体の三次元モデルを作成することは困難である。そこで本論文では、一般の利用者が用いることができるような実物体の簡便な三次元モデリング手法を開発することを目的とする。

ここではまず、従来の三次元モデリング手法について概説し、問題点を明らかにする。実物体の三次元モデル計測の従来手法は、レーザ距離計測装置 [1, 2] やプロジェクタなどを用いて対象にレーザやスリット光

を照射し、それらを観測することで対象の三次元形状を計測する能動的な計測手法 [3, 4, 5] と、ビデオカメラやデジタルカメラなどで撮影される画像のみを用いて画像解析によって三次元計測を行う受動的な計測手法 [6, 7, 8, 9, 10] に大別できる。

一般に、能動的計測手法では、受動的な手法に比べて良好なモデルを得ることができると、映画やゲーム、工業製品のデザイン等の高精度な三次元計測を必要とする分野で用いられている。また、これら能動的計測手法においては、モデリングシステムが利用者に対象物体の推奨測定位置を提示することで、効率的な三次元モデルの獲得を支援する手法が複数提案されている [3, 4, 5]。しかし、比較的高価な特殊機材を必要とすることや、計測機器の取り扱い、データの統合等に特殊な専門知識を必要とするため、専門知識を持たない一般の利用者がこれらの手法を用いることは困難である。

一方、受動的な手法では、計測機器として一般に広く普及しているデジタルカメラやビデオカメラを用いることができ、特殊な機材を必要としないことから近年盛んに研究されている [6, 7, 8]。また、個人向けのモデリングツールとして販売されているものも存在する [9, 10]。これらの手法では、安価にシステムを構成できるという特長があるが、カメラや対象の動かし方によって三次元モデルの推定精度が変化するため、撮影方法の習熟が必要とされる。しかし、三次元復元処理に長い計算時間を要するために、対象物体の撮影中

<sup>\*1</sup>奈良先端科学技術大学院大学 情報科学研究科

<sup>\*1</sup>Graduate School of Information Science, Nara Institute of Science and Technology

や撮影直後に三次元計測結果を確認することができず、物体の撮影方法を効率的に学習できないという問題がある。加えて、これら受動的な計測手法においては、計測結果のプレビュー表示の有効性や、一般的な利用者のための計測の支援方法については検討されて来なかった。

そこで本論文では、ハンドヘルドビデオカメラにより撮影された動画をを入力とする実物体のリアルタイム三次元モデリングを実現し、加えて、一般の利用者が実物体の三次元モデルを効率よく獲得するための撮影支援機能と推定されたモデルのプレビュー機能を提供するインタラクティブ三次元モデリングシステムを提案する。また、これら計測支援機能の有用性について実験により評価を行う。利用者は、システムの指示に従って対象物体を撮影することで容易に三次元モデルを獲得できるため、撮影に関する特別な練習や三次元計測に関する専門知識を必要としない。また、利用者は撮影と同時に生成されるモデルの形状を確認することができ、短時間で良好なモデルを作成することが可能となる。

以下、2章では提案するインタラクティブモデリングシステムの設計方針とシステムの各処理について説明し、3章では試作システムを用いたモデリング結果を示す。4章では専門知識を持たない一般の利用者による被験者実験を行い、提案手法の有効性を検証する。最後に、5章でまとめと今後の課題について述べる。

## 2 撮影支援による三次元モデリングシステム

本章では、まず、モデリングに関する特殊な機器や知識を持たない一般利用者のためのモデリングシステムの設計方針と、実際に構築したシステムの概要について述べる。次にモデリングシステムの処理手順について順に詳述する。

### 2.1 システムの設計方針

本研究では、実物体の三次元モデリングに関する特殊な知識や機材を持たない利用者が、短時間で良好なモデルを作成できるモデリングシステムを構築することを目的とする。このような目的を達成するためには、以下に示す3つの要件を満たさなくてはならない。

- (a) 市販の安価な機材によるリアルタイム物体形状計測の実現
- (b) 良好なモデル作成のためのリアルタイム計測支援の実現
- (c) 利用者に対する形状計測状況提示の実現

本研究では、これらの要件を満たすシステムを構築するために、図1に示すように、マーカ用紙の上にモデ

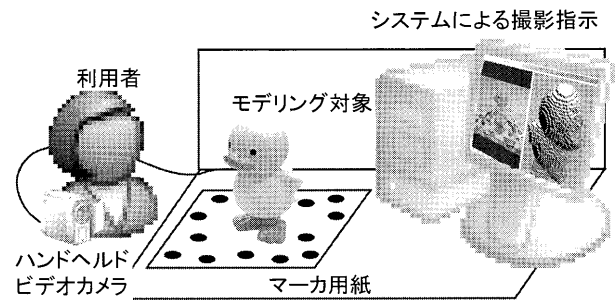


図1 提案システムの利用形態  
Fig.1 Configuration of proposed system.

リング対象を置き、利用者が計算機の画面を確認しながらハンドヘルドビデオカメラを操作するという利用形態を想定する。また、本システムでは、先の要件(a)~(c)に対応した以下の機能を実現する。

- (1) ビデオカメラによるリアルタイム形状推定：一般的なハンドヘルドビデオカメラを用いて物体を撮影し、比較的高速に三次元形状を推定可能な視体積交差法によって、対象の三次元形状をリアルタイムで推定する。これは要件(a)に対応する。
- (2) 撮影方法の指示：システムは、良好なモデルを生成するために必要なカメラの撮影位置を算出し、その撮影位置までのカメラの動かし方を計算機のディスプレイを通して利用者に逐次提示する。利用者は提示された指示に従うことで容易にモデリングを行うことができる。これは要件(b)に対応する。
- (3) モデルのプレビュー表示：対象物体の三次元モデルは、物体の撮影と同時に更新され、利用者は待ち時間なしにモデルの生成状況を確認できる。これは要件(c)に対応する。

なお、本システムでは撮影に用いるカメラの内部パラメータはキャリブレーション済みで既知とする。また、シルエット画像生成処理における精度を上げるため、マーカ用紙は単色の表面を持つ机の上に置かれ、物体の背後に設置される壁面は机面または紙面と同色とする。

### 2.2 提案システムの処理の概要

提案するシステムの処理手順を図2に示す。また、システムによって提示される撮影指示およびプレビュー表示のインタフェース画面を図3に示す。利用者はまず、図3に示すモデリングインタフェース画面内の撮影開始ボタンをマウスでクリックし、対象の撮影を開始する。モデリングが開始されると、システムはフレーム毎にフェーズAの処理を実行し、三次元復元処

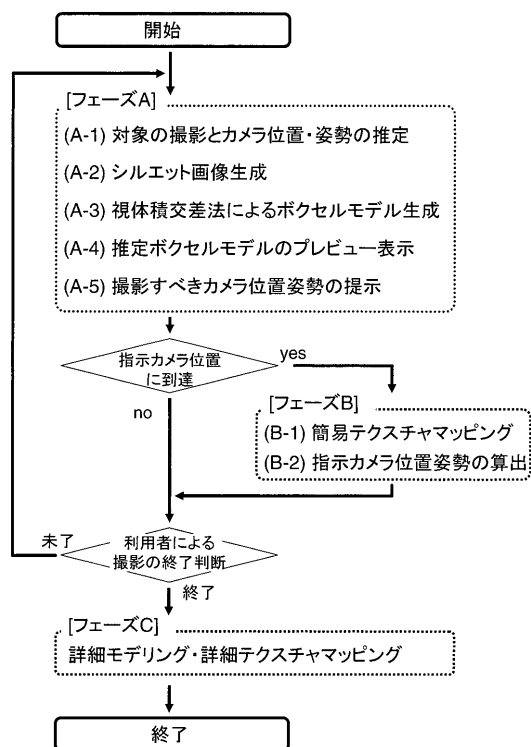


図2 提案システムの処理の流れ

Fig.2 Flow diagram of proposed system.

理, 生成されるモデルのプレビュー表示, カメラの移動指示の表示を行う. 利用者によってカメラが目的の位置まで動かされると, システムはフェーズ B の処理を実行する. フェーズ B では, フレーム毎に実行することが困難なテクスチャマッピング処理と, 次に撮影すべきカメラ位置の算出を行う. 提案システムでは, これらフェーズ A, B の処理を繰り返すことで, 利用者による対象物体のモデリングを行う. 利用者が対象物体の三次元形状が良好に得られたと判断すれば, モデリングインタフェース画面内の撮影終了ボタンをクリックし, 利用者による対象物体の撮影を完了する. 最後に, システムはフェーズ C として, 撮影された全てのフレームの画像を用いて空間解像度の高い詳細モデルの推定をオフライン処理で行う. 以降, 各フェーズの処理について述べる.

### 2.3 フェーズ A: フレーム毎の処理

フェーズ A では, 三次元復元処理, プレビュー表示および撮影すべきカメラ位置の指示をリアルタイムで行う. システムはまず, 利用者が持つカメラによって撮影された画像からマーカを検出し, マーカ座標系におけるカメラの位置・姿勢を推定する (A-1). 次に, 入力画像から対象物体のシルエットを抽出し (A-2), 視体積交差法によって対象の三次元形状を推定する (A-3). 最後に, 現在までに得られているボクセルモデルによる三次元復元形状のプレビュー表示 (A-4) および, 利用者が撮影すべきカメラ位置の指示を行う (A-5). 以

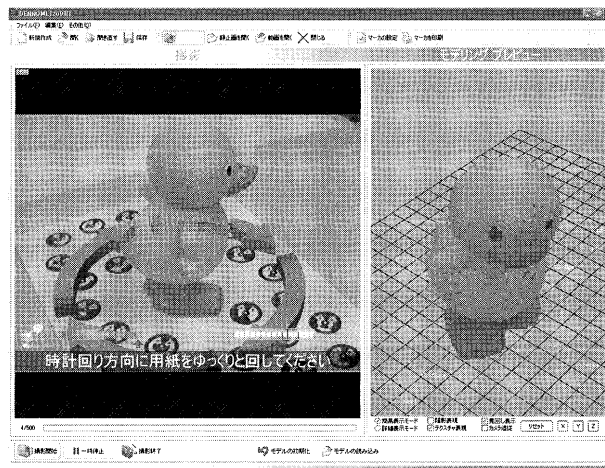


図3 撮影指示およびプレビュー表示のインタフェース

Fig.3 Interface of indication of camera movement and preview.

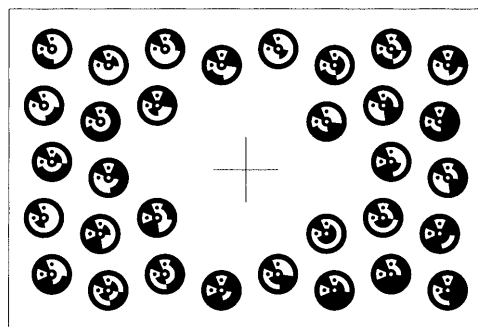


図4 マーカ用紙

Fig.4 Marker paper.

下では各処理について詳述する.

#### (A-1) 対象の撮影とカメラ位置・姿勢の推定

画像内に撮影されたマーカ群を用いて, 利用者の持つカメラの位置・姿勢を推定する. 本システムで用いるマーカ用紙を図4に示す. マーカ用紙には, Naimarkら [11] が提案する円形マーカが複数印刷されている. この円形マーカは, 照明条件の変化に強く, 画像からのマーカ抽出, およびマーカコードの認識を高速に行えるという特徴を持つ. 個々のマーカには 6bit の id が割り当てられており, id から一意にマーカを識別することができる. さらに, 一つのマーカから計 3 点の実空間座標値が識別できる. システムは, まず入力画像中の複数のマーカを画像中から探索し, 個々の id を認識することで画像から一意に識別可能な点群を取得する. 次に, 得られた 4 点以上の画像上のマーカ点群の実空間座標値と, その画像上の座標から, 平面上の PnP 問題 [12] を解くことで, マーカによって張られる実空間座標系に対するカメラの位置・姿勢パラメー

タを推定する。なお、提案システムはマーカ用紙をプリンタに出力する機能を備えており、各マーカの位置関係、用紙の大きさについてはマーカ印刷時の情報により既知とする。また、カメラの内部パラメータは事前にキャリブレーション済みとする。

#### (A-2) シルエット画像の生成

視体積交差法による三次元形状推定の前処理として、画像上での対象物体領域を表すシルエット画像を生成する。ここでは、まず処理 (A-1) で推定されたカメラの位置・姿勢情報に基づき、個々のマーカ周辺に存在する画素の色を平均化することで用紙の色を取得し、また、用紙周辺に存在する画素の色を平均化することで机面領域の色を抽出する。次に、入力画像中から用紙と机面領域の色およびマーカとして検出された領域を排除することで、物体領域のシルエット画像を生成する。

#### (A-3) 視体積交差法によるボクセルモデルの生成

推定したカメラ位置・姿勢とシルエット画像から、視体積交差法を用いて対象物体のボクセル形状を推定する。図5に示すように、視体積交差法では、シルエットを実空間に投影した視体積の中に対象物体が含まれるというシルエット制約に基づき、視体積領域の積である Visual Hull を対象の三次元形状として推定する [13]。本研究では、モデリング対象物体を包含する直方体領域内のボクセルを初期ボクセルとして空間上に配置しておき、フレーム毎に視体積外のボクセルを削除していくことで、逐次的な三次元復元処理を行う [14]。ここでは、ウラ [15] によって提案されている平面間投影に基づく Visual Hull の算出手法を用いることで、視体積外のボクセルを高速に排除し、リアルタイム処理を実現する。本システムでは、マーカ用紙上に置くことが可能な小型のモデリング対象を想定し、あらかじめ初期ボクセルの空間サイズや解像度をシステム内に保持しているが、モデリング対象がシステムの初期値より大きい場合や、計算機的能力不足で処理が重い場合には、モデリングインタフェース内のダイアログを通してボクセルの空間サイズおよび解像度を調整することができる。

#### (A-4) 推定形状のプレビュー表示

視体積交差法によって生成されたボクセルモデルをレンダリングし、モデリング中の物体をディスプレイ上に常時プレビュー表示する。本システムでは、ボクセルの高速なレンダリングを実現するために、各ボクセルを大きさが一定の二次元矩形として近似的に描画する。一般的に行われているボクセルの立方体描画に比べてレンダリングにおける正確さは若干低下するが、描画すべき面の数が6分の1となり、リアルタイム処理における実行レートの低下を防ぐことができる。た

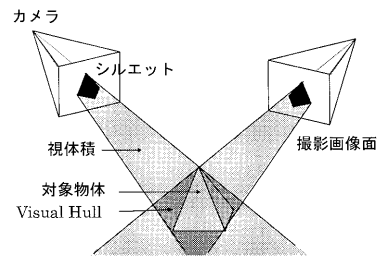


図5 シルエット制約  
Fig.5 Silhouette constraint.

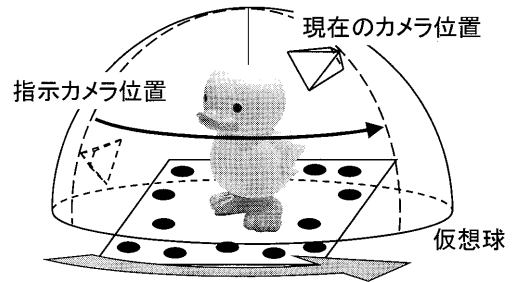


図6 マーカ用紙の回転による撮影位置の移動  
Fig.6 Rotation of marker sheet.

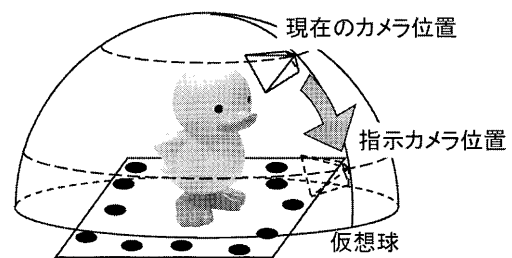


図7 カメラの上下移動による撮影位置の移動  
Fig.7 Up-and-down movement of camera.

だし、利用者はプレビュー表示におけるボクセルの描画を随時立方体による正確なレンダリングに切り替えることができ、またマウスによりモデルを自由に見回すことでモデリング状況を確認することができる。プレビュー表示によって良好なモデルが生成されたと利用者が判断すれば、モデリングインタフェース画面内の撮影終了ボタンをクリックすることで、いつでも撮影作業を完了することができる。

#### (A-5) 撮影すべきカメラ位置の提示

本手法では、一般的に拡張現実感 (AR) の分野で行われているように、ビデオカメラで撮影された実画像上に、動かすべきカメラの方向を CG で重畳合成し、利用者に提示する。これにより、利用者にとって直感的な撮影方法の指示を実現する。また、本システムでは図6, 7に示すように、対象撮影位置の移動指示をマーカ用紙・物体の回転による回転移動とカメラの上下移動に分離し、利用者に指示する。これにより、三

次元モデリングやビデオ撮影の経験が無い一般の利用者においても、比較的簡単に撮影指示に従うことができる。物体の回転およびカメラの上下移動に対応する指示画像の例をそれぞれ図8, 9に示す。図中の矢印は、マーカの回転方向とカメラの移動方向を表しており、図中下部のインジケータおよび矢印の色は、撮影すべきカメラ位置までのカメラの移動量を示している。本システムでは、目標カメラ位置までの距離が小さいほど矢印の色が赤色から緑色に変化する。なお、撮影開始時等、マーカが画面に写っていない場合には、カメラの位置・姿勢が推定できず、撮影すべきカメラの位置・姿勢の提示が困難となるため、図10に示すように、例示画像の提示とともにカメラの初期位置合わせの指示を行い、マーカが画像内に撮像されるようにカメラの移動を促す。

#### 2.4 フェーズ B: 間欠的なモデリング処理

フェーズ A において利用者のカメラの位置がシステムが指示するカメラ位置に到達した場合、システムは間欠的な処理であるフェーズ B を実行する。フェーズ B では、フレーム毎に実行することが困難なテクスチャマッピング処理および次の撮影指示位置の算出を行う。以下では、フェーズ B における各処理について詳述する。

##### (B-1) テクスチャマッピング

フェーズ A で推定されたボクセルモデルの各ボクセルに色を付けることでテクスチャマッピングを行う。ボクセルの色は、それまでに撮影された入力画像上に各ボクセルを投影し、平均色を算出することで決定する。本システムでは、物体表面のボクセルを検出し、表面ボクセルのみに色づけを行うことで計算量を削減する。また、Bresenham の線分描画アルゴリズム [16] を用いて各入力画像におけるボクセルの可視判定を高速に行い、可視となるボクセルのみを用いて色づけを行うことで、遮蔽を考慮した色づけを行う。

##### (B-2) 撮影指示に用いるカメラ位置の算出

テクスチャマッピング処理の結果に基づき、利用者に指示する撮影位置を算出する。本モデリング手法においては、図11に示すように、入力画像上に一度も撮像されないために色が決定できない表面ボクセルが存在する。そこで、本手法では最も多くのボクセルの色を新たに決定できるカメラ位置を算出し、利用者に提示する撮影指示位置として用いる。これにより、利用者が撮影指示に従うことで、色を決定不可能なボクセル数が徐々に減少し、最終的に大多数のボクセルの色が決定された良好な三次元モデルを生成することができる。ただし、一般に、色が決定していないボクセルが最も多く映るカメラ位置・姿勢を算出するためには、カメラ位置・姿勢の網羅的な走査が必要となり、計算



図8 撮影対象の回転指示の例  
Fig. 8 Example of indicating object rotation.

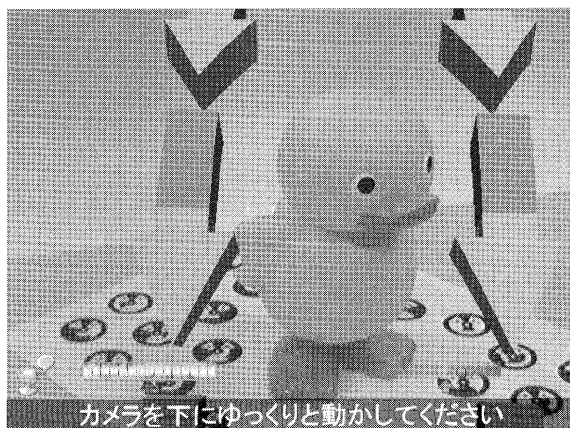


図9 カメラの上下移動指示の例  
Fig. 9 Example of indicating up-down camera movement.

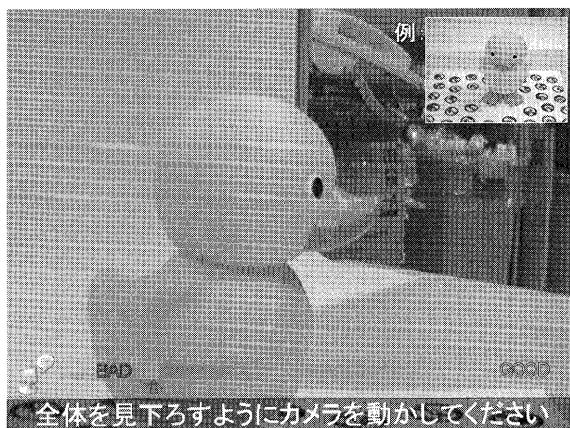


図10 初期位置合わせの指示の例  
Fig. 10 Example of indicating initial camera adjustment.

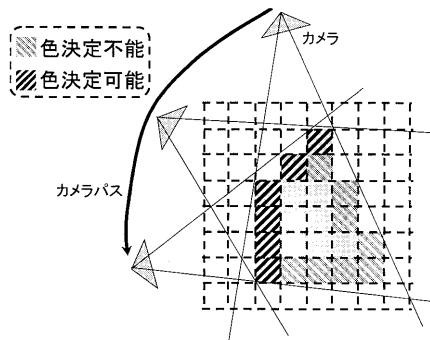


図 11 色を決定不可能なボクセル  
Fig.11 Voxels whose colors cannot be determined.

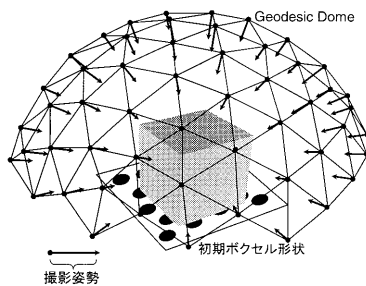


図 12 撮影指示位置の候補 (一部)

コストの問題が発生する。本研究では、この問題を回避するため、撮影指示位置の候補として、図 12 に示す測地ドーム (Geodesic Dome) の頂点を用い、各指示候補位置のカメラ姿勢は初期ボクセルモデルの中心方向に設定する。また、半球体の半径は撮影画角内に初期ボクセルモデルが収まるように設定する。なお、利用者による撮影が既に完了している頂点については、その頂点位置を指示候補から除外する。

### 2.5 フェーズ C: オフライン詳細モデリング

利用者がプレビュー表示から良好なボクセルモデルが推定されたと判断すれば撮影を終了し、オフライン処理であるフェーズ C に移行する。フェーズ C では、リアルタイムモデリング処理では不可能な、より空間解像度の高い詳細モデリング処理を行い、その後、詳細なボクセル形状モデルに対して再度テクスチャマッピング処理を行う。ここではまず、フェーズ A, B で推定した簡易ボクセルモデルにおいて、撮影されたカメラの位置に最も接近するボクセルを探索し、そのボクセルに対してカメラの 1 画素の幅を逆投影する。逆投影した 1 画素の空間中での幅を、新たに単位ボクセルサイズとすることで、視体積交差法によって推定可能な最大空間解像度でモデリング処理を行う。次に、得られた空間解像度の高いボクセルモデルに対して、遮蔽を考慮したテクスチャマッピング処理を行う。これによって得られる詳細なボクセルモデルを最終的な

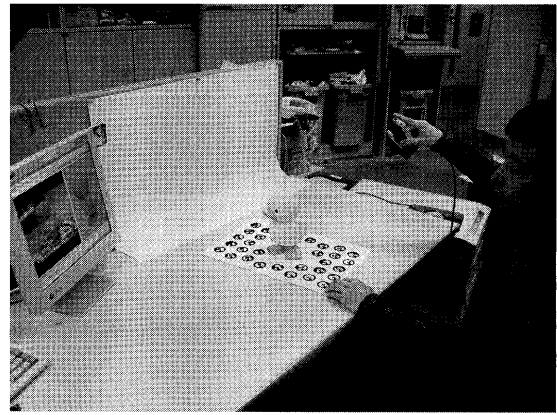
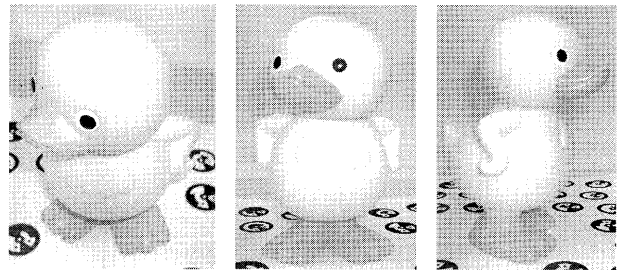


図 13 モデリング環境  
Fig.13 Appearance of interactive modeling system.



(a) 物体 A



(b) 物体 B

図 14 モデリング対象  
Fig.14 Modeling targets.

システムの出力とする。フェーズ C においては、最終的なボクセル出力までに長い待ち時間が発生するが、システムが利用者に何らかの入力操作を要求することはないため、利用者が処理中に拘束されることは無い。

### 3 試作システムによるモデリング実験

本章では、試作システムを用いた実際のモデリング例を示す。本実験では、図 13 に示すように、モデリング対象の下にマーカを印刷した用紙を敷き、本システムによるモデリング手法を熟知した利用者によってモデリングを行った。モデリング対象物体として、図 14(a) に示す凹部分が少なく形状が単純な物体 A と、同図 (b) に示す複雑な形状の物体 B を用いた。

試作システムは表 1 に示す機器から成る。カメラの



不殿・佐藤・横矢：ハンドヘルドビデオカメラを用いた撮影支援インタフェースを有するインタラクティブ三次元モデリングシステム

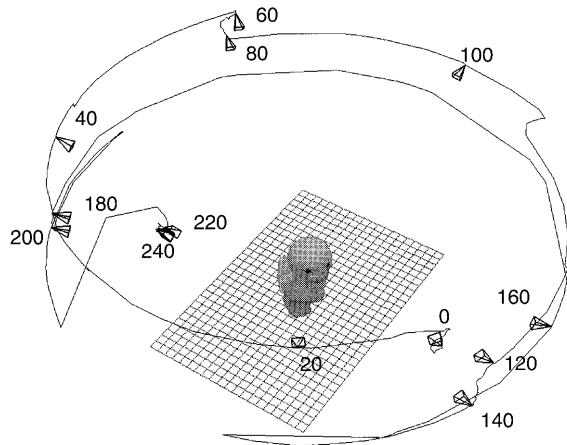


図 15 撮影時のカメラパス (物体 A)  
Fig. 15 Camera path for modeling (object A).

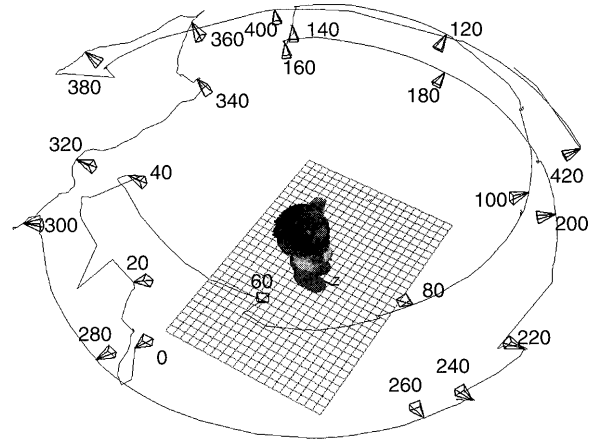


図 17 撮影時のカメラパス (物体 B)  
Fig. 17 Camera path for modeling (object B).

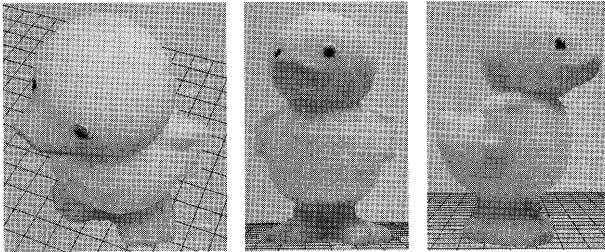


図 16 モデリング結果 (物体 A)  
Fig. 16 Result of modeling (object A).

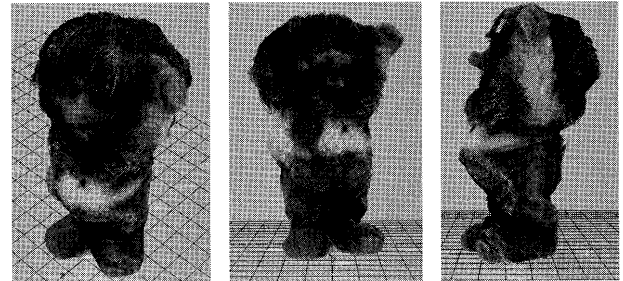


図 18 モデリング結果 (物体 B)  
Fig. 18 Result of modeling (object B).

内部パラメータは事前に Tsai の手法 [17] により算出した。また、表 1 に示した計算機でフェーズ A をリアルタイム処理可能なボクセル解像度として、 $W64 \times H80 \times D64$  個 ( $W200 \times H250 \times D200\text{mm}$ ) のボクセル空間を初期ボクセルとして設定した。

まず物体 A を対象として、システムの指示に従うことでモデリングを行った。システムの指示に従いながら撮影した利用者のカメラパスを図 15 に示す。図中の四角錐は 20 フレーム毎の撮影姿勢を、数値はフレーム番号を表す。同図から、撮影対象を中心として、様々な位置からの撮影が行われていることが分かる。本実験において、フェーズ A でカメラの位置・姿勢が推定され、モデリングに用いられた入力画像枚数は

合計 259 枚であり、各フレームは平均 132 ミリ秒 (7.6 fps) で処理された。またフェーズ B の処理に要した時間は平均 360 ミリ秒であった。利用者が撮影開始ボタンをクリックしてから撮影完了ボタンをクリックするまでの時間は 72 秒 (マウス操作による推定形状の確認時間を含む) であり、この間フェーズ B の処理は 8 回実行された。フェーズ C では、初期ボクセルサイズが自動的に  $W390 \times H576 \times D419$  個 ( $W140 \times H206 \times D150\text{mm}$ ) に設定され、最終的なモデルの出力には 58 分を要した。図 16 にフェーズ C によって生成された詳細モデルを示す。同図から、比較的単純な対象物体 A については良好な結果が得られていることが確認できる。

同様に、物体 B を対象としてモデリングを行った。利用者が操作したカメラのパスを図 17 に示す。本実験において、モデリングに用いられたフレーム数は合計 423 枚であり、撮影完了までに要した時間は 121 秒であった。また、フェーズ B の処理は 10 回実行された。フェーズ C において設定された初期ボクセルサイズは  $W551 \times H796 \times D479$  個 ( $W160 \times H231 \times D139\text{mm}$ ) であり、最終的なモデルの出力には 84 分を要した。図 18 に生成された詳細モデルを示す。対象物体 B においては、形状が凹状になっている人形

表 1 試作システムの機器構成  
Table 1 Specifications of interactive modeling system.

PC	
CPU	Pentium-4 3.2GHz
Memory	2GB
Graphic card	RADEON9800PRO 128MB
IEEE1394 camera: Aflux C104T	
Resolution	640×480 pixels
View angle	31.7° × 24.1°
Maximum frame rate	15 frames/sec

の顔と腕の間部分, および髪の毛の部分でテクスチャが不鮮明となり, 良好な結果が得られなかった. これは, 本システムが用いている視体積交差法が(1)原理的に凹形状を正しく推定できないこと, および(2)背景が透けるような細かな形状には対応できないこと, に起因している. これらの問題に対応するためには, 今後フェーズCにおいて, ステレオ法やスペースカービング法[18]などの凹物体に対応可能な三次元形状推定手法を組み合わせたことや, 物体の透過度を表すアルファ値を考慮した視体積交差法である Opacity Hull による推定手法[19]を用いることが必要であると考えられる.

#### 4 被験者実験による評価

提案システムを用いて複数の被験者によるモデリング実験を行った. 被験者は, 20歳代, 30歳代の男女10名であり, 全員コンピュータを日常的に扱っているが, 3次元モデリングの経験は持たない. 本実験では, システムによる撮影指示の有用性を, モデリング作業の簡便性, 生成されたモデルの満足度を基に複数の被験者により主観的に評価してもらうことを目的とする. 実験に用いた機材や環境は, 前章で示したものと同一であり, また, モデリング対象物体として図14(a)に示した物体Aを用いた.

本実験では, システムを利用してもらう前に, 提案システムの利用方法や撮影に関する注意点について5分ほど口頭により説明した. 実験では, 撮影指示の有用性を確認するため, 各被験者に対して指示機能を有効にした状態でのモデリング(指示あり)と, 指示機能を無効にした状態の自由撮影によるモデリング(指示なし)を行ってもらい, フェーズCでの詳細モデリング終了後にアンケート評価を行った. アンケートでは, 生成されたモデルの精度について, インターネットオークションのための三次元商品カタログや, ゲームやWEBページなどに登場させる三次元オブジェクトへの利用を前提に, [1. 利用に耐えない 2. 大幅な修正作業が必要である 3. 若干の手直しが必要である 4. 直接利用できる]の中から1つを選択させた. また撮影労力については, [1. 大変気になる 2. やや気になる 3. どちらとも思わない 4. 気にならない 5. 全く気にならない]の中から1つを選択させた. なお, 被験者が撮影方法を学習することによって発生する評価の変動を防ぐため, 指示なし, 指示ありの順でモデリング実験を行うグループA(5名)と, 逆に, 指示あり, 指示なしの順で実験を行うグループB(5名)に被験者を分け, 評価を行った. ただし, 計算機のメモリ容量の制限から, 計算機に蓄積される画像枚数の上限を999フレームとし, 上限に達した場合には強制的にフェーズ

Cに移行するものとした. また, 実験時間短縮のため, フェーズCで行う詳細モデリング処理に用いるボクセルの解像度を $W150 \times H150 \times D150$ 個の比較的粗い設定に固定した. この場合のフェーズCに対する待ち時間は, およそ6分程度に短縮される.

モデリング実験における被験者ごとのモデリング時間, 撮影フレーム数と, アンケート評価による主観的なモデルの生成精度, 撮影労力について図19に示す. また, これらの平均値について表2に示す. ここで, モデリング時間は被験者が撮影開始ボタンをクリックしてから撮影終了ボタンをクリックするまでの時間であり, マウス操作による物体のプレビュー確認時等, カメラを対象に向けていない時間も含まれる. 一方, 撮影フレーム数は, 画像内のマーカをシステムが検出することによってカメラ位置・姿勢が推定されたフレーム枚数を表し, カメラがマーカを撮像していないフレームはカウントしない. ここではまず, 生成されたモデルの精度について考察する. 生成されたモデルの精度について, 10名中7名の被験者が撮影指示機能を有効にすることで無効の場合よりも結果が改善するという回答を行った. 特に, 本研究が目的としているWeb等での三次元オブジェクトへの利用について, 指示なしでは「4. 直接利用できる」と回答した被験者が1名であったのに対して, 指示ありでは6名が直接利用できる」と回答している. 逆に, モデルの精度が提案手法によって悪化したと答えた被験者は1名(被験者B-2)であった. この被験者は, 実験アンケートに「カメラの操作に不慣れなために操作が難しかった」とコメントしており, 先に行った「指示あり」の実験では撮影対象を常に画像の中心付近に捉えることが困難な様子であった. これは, 一般的なデジタルカメラ等のプレビュー画像と異なり, 本システムにおいてカメラの撮影画像が離れた場所にあるディスプレイに表示されることに起因していると考えられ, 利用者がこのような撮影操作に不慣れな場合に起こる問題である. 一方, 撮影労力については10名中6名で改善が見られ, 指示ありの撮影方法で労力が「4. 気にならない」と回答した被験者は7名であった. これらの結果から, 多数の被験者においては提案システム(指示あり)を用いることで, 容易に高品位の三次元モデルを生成できていることが分かる.

モデリング時間・撮影フレーム数に関しては, 多くの被験者において指示ありが指示なしを上回っている. これは, 被験者がモデリング時間の大半を撮影指示の確認・実行に費やしたためであり, プレビュー表示の確認による適切な終了判断を行わなかったことが原因であると考えられる. 加えて, 本システムの撮影候補位置までの移動指示手法では移動経路上に存在する撮

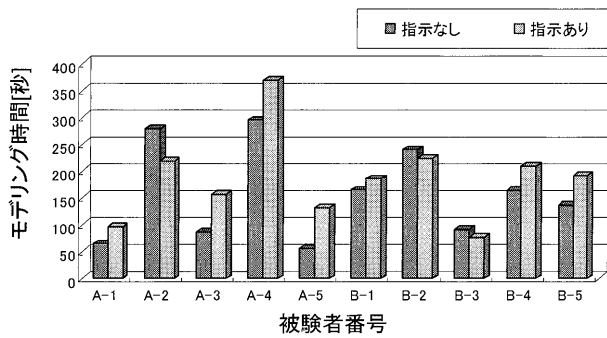


不殿・佐藤・横矢：ハンドヘルドビデオカメラを用いた撮影支援インタフェースを有するインタラクティブ三次元モデリングシステム

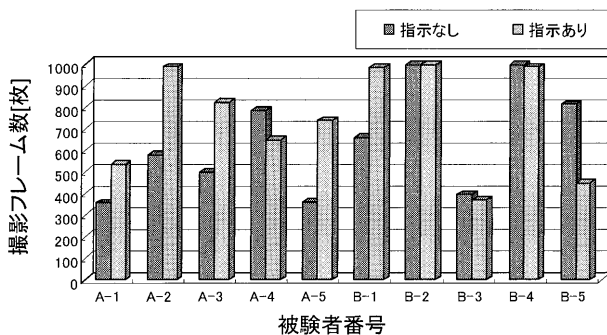
表 2 モデリング時間、撮影フレーム数とアンケート結果の平均

Table 2 Averages of modeling time, captured frame, and questionnaire scores.

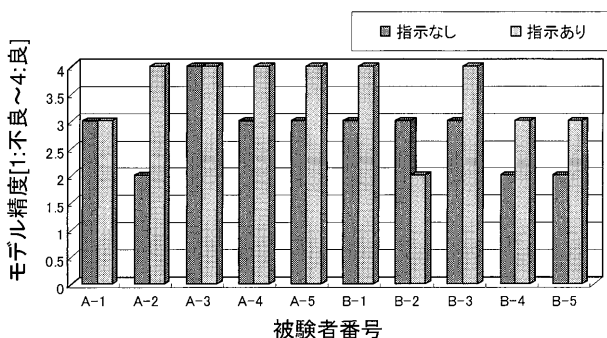
	指示	グループ A	グループ B	全体
モデリング時間 [秒]	なし	156.2	159.4	157.8
	あり	194.6	177.0	185.8
撮影フレーム数 [枚]	なし	515.4	772.4	643.9
	あり	745.8	757.0	751.4
モデリング精度 [1:不良 - 4:良]	なし	3.0	2.6	2.8
	あり	3.8	3.2	3.5
撮影労力 [1:大 - 5:小]	なし	3.0	2.2	2.6
	あり	3.6	3.2	3.4



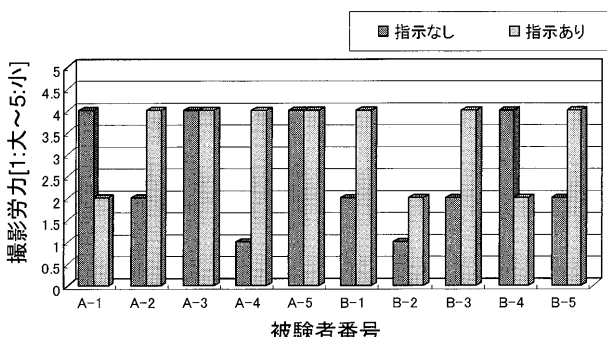
(a) モデリング時間



(b) 撮影フレーム数



(c) 生成モデルの精度



(d) 撮影労力

図 19 個々の被験者に対するモデリング時間、撮影フレーム数とアンケート結果

Fig. 19 Modeling time, captured frame, and questionnaire scores for each subject.

景済みのカメラ位置を考慮していないために、指示ありの場合には、撮影済みの箇所を再度撮影するという冗長な撮影が発生しているものと考えられる。指示なしの場合には、被験者が直感的に撮影済みの箇所を避けて撮影することがあり、結果としてモデリング時間が短くなる場合が見られた。今後、モデリング時間の短縮を実現するためには、最適な撮影指示経路の探索および利用者が撮影終了を明確に判断できる何らかの指標の提示が必要である。

また、被験者自身が最終的な終了判定を行わず、撮影フレーム数が設定した上限の 999 枚に達した被験者が 2 名 (B-2 指示あり・なし, B-4 指示なし) 見られた。これらの被験者は、先に述べたカメラの操作や、複数のタスク (カメラの操作、撮影画像・指示の確認、プレビューの確認等) の同時実行という行為に不慣れた被験者であり、本システムを用いて効率的にモデリングを行うためには多少の習得期間が必要であると考えられる。ただし、本システムはリアルタイムでモデリング結果を確認することができるため、短い習得期間で効率的にシステムの操作を学習することができる。また、これらの考察から、今回の実験上の制約とした撮影フレーム数の上限 999 枚については、上限枚数をさらに増加させてもモデルの生成精度は向上せず、撮影労力のみが増加したものと考えられる。

## 5 まとめ

本論文では、対象物体の撮影方法の指示と推定された三次元形状のプレビュー表示によって支援を行うことで、一般的な利用者が特別な練習なしに実物体の三次元モデルを作成できるインタラクティブモデリングシステムを提案した。実験では、三次元モデリングの経験を持たない一般の利用者によるモデリング実験を行うことで、本システムが専門知識を持たない一般利用者にとって有用であることを確認した。カメラの操作や複数タスクの実行について不慣れた利用者に対しては、システムの操作に関する習熟の必要性が認めら

れたが、このような利用者についても、本システムが提供する支援機能によって効率の良い撮影方法の学習が可能であると考えられる。また今後、複雑な形状に対応するためには、ステレオ法やスペースカービング法の導入が必要であり、加えて、生成した三次元モデルを手軽に利用するためにはボクセルモデルからポリゴンモデルへの変換が必要である。

## 参考文献

- [1] Leica Geosystems HDS LLC: "Hds2500," <http://hds.leica-geosystems.com/>, 2000.
- [2] コニカミノルタ株式会社: "Vivid 910," <http://konicaminolta.jp/>, 2002.
- [3] J. E. Banta, L. M. Wong, C. Dumont and M. A. Abidi: "A next-best-view system for autonomous 3D object reconstruction," *IEEE Trans. Systems, Man and Cybernetics*, Vol. 3, No. 5, pp. 589-598, 2000.
- [4] M. Matsumoto, M. Imura, Y. Yasumuro, Y. Manabe and K. Chihara: "Support system for measurement of relics based on analysis of point clouds," *Proc. Int. Conf. on Virtual Systems and Multimedia*, p. 195, 2004.
- [5] 芳賀浩嗣, 佐藤宏介: "ハンドプロジェクトによる支援光を用いた形状計測", 第9回パターン計測シンポジウム資料, pp. 35-38, 2004.
- [6] L. Zhang, B. Curless, A. Hertzmann and S. M. Seitz: "Photometric method for determining surface orientation from multiple images," *Proc. IEEE Int. Conf. on Computer Vision*, Vol. 1, pp. 618-625, 2003.
- [7] G. G. Slabaugh, W. B. Culbertson, T. Malzbender, M. R. Stevens and R. W. Schafer: "Methods for volumetric reconstruction of visual scenes," *Int. J. on Computer Vision*, Vol. 57, No. 3, pp. 179-199, 2004.
- [8] H. Kim and I. Kweon: "Optimal photo hull recovery for the image-based modeling," *Proc. Asian Conf. on Computer Vision*, Vol. 1, pp. 384-389, 2004.
- [9] 株式会社NTTデータ三洋システム: "Cyber modeler handy light," <http://www.nttd-sanyo.co.jp/>, 2002.
- [10] UZR GmbH & Co KG: "imodeller 3D," <http://www.imodeller.com/en/>, 2001.
- [11] L. Naimark and E. Foxlin: "Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker," *Proc. IEEE/ACM Int. Symp. on Mixed and Augmented Reality*, pp. 27-36, 2002.
- [12] R. Klette, K. Schluns and A. Koschian Eds.: *Computer Vision: Three-dimensional Data from Image*, Springer, 1998.
- [13] H. Baker: "Three-dimensional modeling," *Proc. Int. Joint Conf. on Artificial Intelligence*, Vol. 2, pp. 649-655, 1977.
- [14] Y. Kuzu and V. Rodehorst: "Volumetric modeling using shape from silhouette," *Proc. Turkish-German Joint Geodetic Days*, pp. 469-476, 2001.
- [15] ウ小軍, 和田俊和, 東海彰吾, 松山隆司: "平面間透視投影を用いた並列視体積交差法", *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. 42, No. SIG6(CVIM2), pp. 33-43, 2001.
- [16] J. E. Bresenham: "Algorithm for computer control of a digital plotter," *IBM Systems J.*, Vol. 4, No. 1, pp. 25-30, 1965.
- [17] R. Y. Tsai: "An efficient and accurate camera calibration technique for 3D machine vision," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 364-374, 1986.
- [18] K. N. Kutulakos and S. M. Seitz: "A theory of shape by space carving," *Int. J. of Computer Vision*, Vol. 38, No. 3, pp. 199-218, 2000.
- [19] W. Matusik, H. Pfister, A. Ngan, P. Beardsley and R. Ziegler: "Image-based 3D photography using opacity hulls," *Proc. ACM SIGGRAPH*, pp. 427-437, 2002.

(2005年6月14日受付)

## [著者紹介]

## 不殿 健治



2003年奈良工業高等専門学校専攻科電子情報工学専攻修了。2005年奈良先端科学技術大学院大学情報科学研究科修了。修士(工学)。現在、日本ビクター株式会社勤務。

## 佐藤 智和



1999年大阪府立大学工学部情報工学科卒。2003年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。現在、同大情報科学研究科助手。コンピュータビジョンの研究に従事。2001年電子情報通信学会学術奨励賞受賞。電子情報通信学会、情報処理学会、IEEE各会員。

## 横矢 直和 (正会員)



1974年大阪大学基礎工学部情報工学科卒。1979年同大大学院博士後期課程了。同年電子技術総合研究所入所。以来、画像処理ソフトウェア、画像データベース、コンピュータビジョンの研究に従事。1986~87年マッギル大・知能機械研究センター客員教授。1992年奈良先端科学技術大学院大学・情報科学センター教授。現在、同大情報科学研究科教授。1990年情報処理学会論文賞受賞。2005年情報処理学会フェロー。電子情報通信学会フェロー。工博。電子情報通信学会、情報処理学会、人工知能学会、日本認知科学会、映像情報メディア学会、IEEE各会員。