

## 複数任意視点からの Octree 表現による 3次元形状復元

澤 弘義<sup>†</sup> 坂本 尚久<sup>†</sup> 岩成 英一<sup>†</sup> 岡田 至弘<sup>†</sup> 金出 武雄<sup>††</sup>

<sup>†</sup> 龍谷大学 理工学部 電子情報学科

<sup>††</sup> カーネギーメロン大学 ロボティクス研究所

あらまし

本研究では、任意視点から撮影した複数の画像を基に3次元形状復元が可能なシステムを提案する。まずここでは、3次元空間上と2次元画像上の対応する点を一つの行列で結びつけた中心射影行列によるカメラキャリブレーションを行う。これによって、撮影者は可搬性に富んだ任意の視点、及び焦点距離での撮影ができ、よりオクルージョンが少ない画像を得ることが可能となる。次に、色情報を用いた閾値処理により対象物体のシルエットを抽出する。そして、複数視点から撮影した入力画像のシルエット情報とキャリブレーションデータを基に、octree表現によって骨格オブジェクトの3次元形状復元を行う。

キーワード 3次元復元, 骨格オブジェクト, 任意視点, shape from silhouette, octree表現

## Reconstruction of 3D Objects by Octree Representation from Multiple Any Views

Hiroyoshi Sawa<sup>†</sup> Naohisa Sakamoto<sup>†</sup> Eiichi Iwanari<sup>†</sup>

Yoshihiro Okada<sup>†</sup> Takeo Kanade<sup>††</sup>

<sup>†</sup>Department of Electronics and Informatics, Ryukoku University

<sup>††</sup>The Robotics Institute, Carnegie Mellon University

**Abstract**

In this research, we propose a system to reconstruct the 3D objects based on the images from any views. First, we calibrate the camera by using a perspective projection matrix which combines the corresponding points in the 3D space and in the 2D image. And then, we are able to take pictures from any views in any focal length, and obtain images with high mobility and fewer occlusion. Next, we extract object silhouette by the threshold method using color information. Finally, we reconstruct the 3D skeleton object by octree representation based on silhouette information and calibration data of the input images from multiple any views.

**key words** 3D reconstruction, skeleton object, any views, shape from silhouette, octree representation

## 1 はじめに

3次元形状復元の研究は、CADなどの工業分野だけでなく、医用、アパレル、土木・建築、さらに考古学や映画のCG制作など幅広い分野で必要とされている。従来はハードウェア的なアプローチによる3次元形状復元が一般的であり、ロボットアームによる接触型や、レーザー光やスリット光などを用いたレンジファインダと呼ばれる非接触型の3次元形状計測装置によって3次元データを獲得した。これらの中には数mm以下のオーダーで計測できるものもあり、データの信頼性もよい。ただし、これらの計測装置は高価であり、機動性に欠け、また装置に応じて計測できる物体の大きさに制限がある。

そこで、これらの計測工学よりのアプローチではなく、CV(Computer Vision)の分野から、カメラを人間の視覚情報の入力装置と考え、得られた複数の画像からソフトウェア的に物体の3次元形状を復元する研究が行われている。従来のCVにおける3次元形状復元はShape from Xとしてまとめることができる。Xには3次元形状復元を行う際に注目する画像特徴が記され、{shading, contour, highlight, motion, silhouette}などがある。また、それぞれの特徴量について分析し、{motion, silhouette, shading}を統合した手法[1]や、{motion, contour, highlight}を統合した手法[2]などが提案されている。しかし様々な物体、及び環境下において、個々の特徴量が安定ではないため根本的な解決には至っていない。

そこで、近年は最も安定な特徴量であるシルエット情報を用いて、対象物体を小立方体の集合で表現する研究が盛んに行われている。これには、回転テーブルを用いて、octree表現によって複雑な形状を持った物体に対しても復元を行った研究[3][4][5]や、ボクセル表現によって復元を行った研究[6][7]などがある。また、光線追跡法によりそれぞれの立方体に色属性をあたえるボクセルカラリング[8]やテクスチャマッピングによりリアリティの向上が目指されている。

また、近年は仮想空間の発展に伴い、手軽に実世界の物体を3次元データとして利用したい必要性から、従来の回転テーブルから開放された任意視点からの撮影によって復元を行う手法などが提案されている[9][10]。

本研究では回転テーブルを用いずに、任意の視

点・焦点距離で物体を撮影し、より正確な復元を行うために、オクリュージョンが発生する箇所において別視点からの撮影を行う。そして、物体のシルエット情報からoctree表現によって、複雑な骨格オブジェクトに対しても復元可能なシステムを提案する。

## 2 本手法の特徴

ここでは、本手法における3つの特徴を示し、処理の流れを図1に示す。

### ● 任意視点におけるカメラキャリブレーション

我々は、3次元空間上と2次元画像上の対応する点を1つの中心射影行列によって直接関係づけることで、任意視点からの撮影を可能とする。従来、撮影に用いるCCDカメラは業務用のものが多く、理想的なピンホールカメラであるか、それに相当する焦点距離を精度よく求めなければならなかった。ここでは、カメラの位置・姿勢の自由度に合わせて、ズームによる撮影も可能となり、撮影者の意図に沿ったモビリティな画像の入力が可能となる。

### ● カラー情報を用いた物体領域の抽出

複数の任意視点から撮影した全ての画像において、対象物体のシルエットを抽出し、物体の形状情報を獲得する。これは、物体領域と背景領域とを分離するという比較的簡単な処理のみであるため、ピクセルレベルの精度でロバストにデータを得ることが可能であり、複雑な形状の物体に対しても有効である。ここでは、従来の濃淡値からの閾値処理に加えて、カラー情報を用いた色空間閾値処理を行う。これによって、照明光に影響されずに、どのような撮影環境下においてもロバストな抽出が可能となる。

### ● octree表現を用いた骨格モデルの復元

我々は、カメラで撮影が可能であり、画角に収まるものであれば、自然界に存在するあらゆる物体に対して3次元形状の復元が可能となる手法の確立を目指す。例えば、対象物体にソリッドな物体はもちろん、植物や昆虫などの複雑な形状や表皮を持ったものも対象とする。また、針金によって組み立てられた骨組みや生物の骨格などの骨格オブジェクトに対しても、正確な復元が可能であることを示す。

そこで、任意視点でのそれぞれのキャリブレーションデータと物体のシルエット情報から、

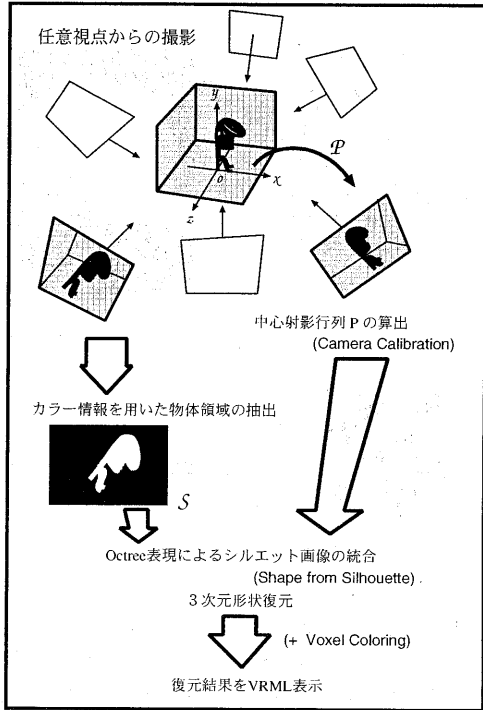


図1: 本手法の処理の流れ

octree 表現によってすべてのシルエットを計算機内で統合する。octree 表現を用いることで、複雑な形状の物体に対しても正確な復元が可能になる。

### 3 原理

#### 3.1 中心射影行列によるカメラキャリブレーション

##### 3.1.1 カメラキャリブレーション法

任意視点で撮影を行うために、それぞれの視点でカメラキャリブレーションを行う。

ここで、ピンホールカメラモデルにおけるカメラキャリブレーション法には、一般に大きく分けて2つの方法がある。一つは、Tsai に代表される外部変数と内部変数を独立に求める方法 [11] で、安定にカメラパラメータを求めることができる。このときの内部変数は、CCD の内部データを用いた物理的な方法によって算出する。

もう一つは、世界座標系から画像座標系までの光学的・幾何学的な関係を1つの行列で表す方法である。これは、射影幾何学に由来し、 $3 \times 4$  の中心射影行列によって表すことができる。カメラの外部

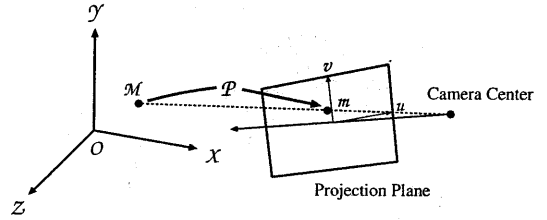


図2: 中心射影行列による座標変換

変数、及び内部変数を含んだ行列として、対象から画像までの対応を記述することができる。これは、AR(拡張現実感: Augmented Reality)[12]における仮想物体と実空間とを対応づける際のキャリブレーションとしてよく用いられる手法である [13]。

##### 3.1.2 中心射影行列の算出

中心射影行列は焦点距離、及び画素サイズを一つの行列に内在することができ、任意の視点・焦点距離からの撮影が可能である。中心射影行列は、6つの3次元空間上の点と2次元画像上の点との対応づけから求めることが可能である。ここで、中心射影行列を  $P$  とし、3次元上の点を  $M = \{X, Y, Z, 1\}^T$ 、画像上の点を  $m = \{u, v, 1\}^T$  とすると、以下の関係式が成り立つ。

$$sm = PM \quad (1)$$

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \quad s: \text{スカラー}$$

実際には、6点以上の点から式(1)から得られた解を初期値として、式(2)の  $C$  を最小にする評価関数を非線形最小2乗法によって解く [14]。

$$C = \sum_{i=1}^N \left( \frac{p_{11}^T M_i + p_{14}}{p_{31}^T M_i + p_{34}} - u_i \right)^2 + \left( \frac{p_{21}^T M_i + p_{24}}{p_{31}^T M_i + p_{34}} - v_i \right)^2 \quad (2)$$

##### 3.1.3 キャリブレーションパターン

キャリブレーションポイントである点  $M$ 、 $m$  を得るために、図3のような底1枚、横2枚からなるキャリブレーションボードを用いる。底ボードの中心に対象物体を置き、そこを世界座標系の原点とする。ボードにはチェックパターンが描かれており、エッジ検出により特徴点  $m$  を抽出し、3次元上の点  $M$  との対応をとる。また、撮影者が360度あらゆる角度から撮影できるように、横ボードと底ボードは切り離すことができ、撮影に応じて横ボードを90度ずつ移動させる。撮影方向がわかるように、図

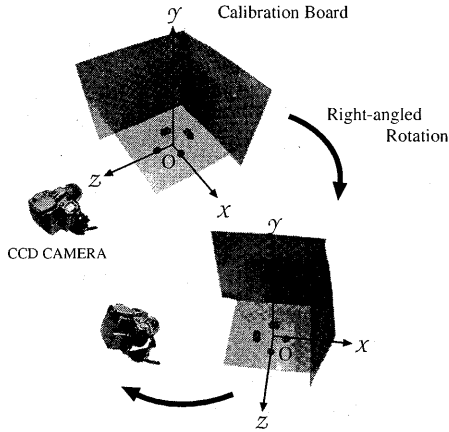


図3: キャリブレーションボード

view angle	0	$\pi/2$	$\pi$	$3\pi/2$
mark set in pattern				

図4: 撮影方向推定のためのランドマーク

4のようなランドマークを底ボードに貼っておく。さらに、対象物体が小さい場合には、横ボードを原点に近付けてキャリブレーションを行い、中心射影行列 $P$ を求める。

### 3.2 色空間閾値処理による物体領域の抽出

複数の画像から物体領域を抽出し、そのシルエット情報から3次元形状を復元する手法を、一般にShape from silhouette法と呼ぶ。ただし、凹面やオクルージョンの発生する箇所では、図5のようなモデリングエラーが発生し、正確な復元結果を得ることができない。我々は任意視点から撮影することによって、このモデリングエラーを減らす。

まず、正確に物体領域と背景領域とを分離するために、RGB色空間からHSV色空間へ変換し、濃淡値の強度、色相値の両面から閾値処理を行う。HSV色空間に色空間閾値処理をほどこすことで、照明光や様々な物体色に左右されずに良好な物体のシルエットを抽出できる。

物体の背景は、2色からなるチェックパターンが描かれたキャリブレーションボードである。図6に示すようなチェックパターンの2色に、濃淡値がわずかに異なった濃い緑色を用いる。これは人間の知覚に基づいて、多くのCCDカメラにおいてRGB

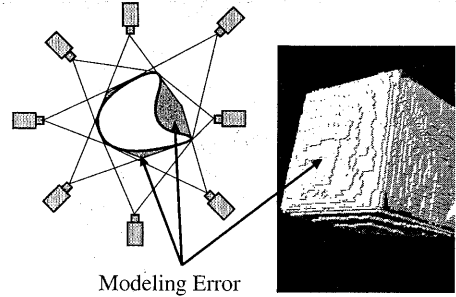


図5: モデリングエラー

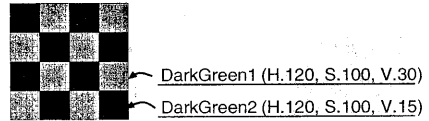


図6: チェックパターン

固体撮像素子に緑色系を多く割り当てており、緑色系が最も忠実に画像で再現できるからである。さらに、濃淡値の低い緑色を用いることによって、濃淡値の高いものに比べて、照明光に対する色相値の変化を軽減することができる。

### 3.3 Octree 表現による3次元形状復元

#### 3.3.1 Octree 表現によるシルエット画像の統合

octree 表現とは、3次元空間上のデータを大きさの異なる立方体の組み合わせで表現する手法である。これらの立方体は、図7のように八分木による再帰的な構造になっている。octreeのそれぞれの立方体には、分木が存在するかどうかの属性を持っている。これには3つの属性があり、それぞれ、データが完全に含まれる立方体を“BLACK”、全くデータが含まれない立方体を“WHITE”、一部にデータが含まれる立方体を“GRAY”と定義する。さらに属性が“GRAY”である立方体は8つの小

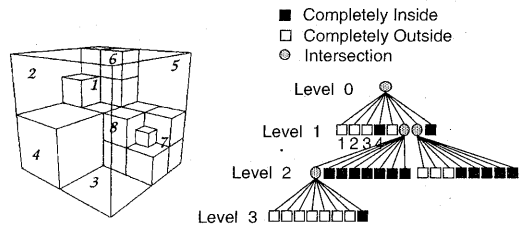


図7: octreeの構造

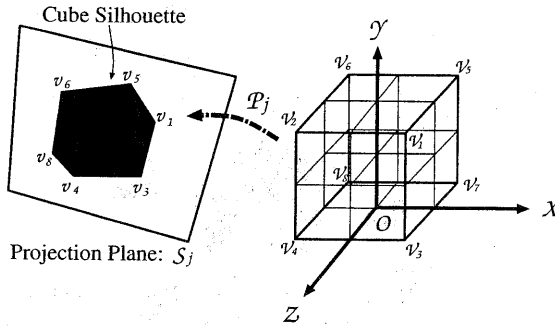


図8: 投影面に octree を投影

な小立方体に分割され、それぞれ属性を持つようになる。最終的には、データがすべて含まれる立方体“BLACK”と、全くデータが含まれない立方体“WHITE”とに分割される。

以下、octree 表現によるシルエット画像の統合処理について、フローの流れに沿って説明する。まず、それぞれの任意視点  $j = 1, 2, \dots, n$  で撮影した際の射影行列を  $P_j$ 、シルエット画像を  $S_j$  と定義する。

次に、対象物体をすべて含むような最も大きな立方体“ROOT-CUBE”を1つ定義し、LEVEL 0 とする。再帰の回数に応じて LEVEL 0 から LEVEL  $l$  まで階層レベルを与える。LEVEL が1つ増えるごとに、立方体の体積は  $1/8$ 、個数は8倍になる。

次に、ROOT-CUBE を8つの SUB-CUBE に等分割し、LEVEL を1つ増やす。そして、この8つの SUB-CUBE それぞれで対象物体のシルエットに対して交差判定を行い、それぞれの立方体に“1”、“0”の属性値、または再帰の命令を与える。

交差判定は、立方体をそれぞれの視点  $j$  に投影したときのシルエットと、対象物体のシルエットとの交差状態を判定する。立方体のシルエットが物体のシルエットに完全に含まれているならば“BLACK”、また交差しているならば“GRAY”、また全く交差していないならば“WHITE”となり、属性が“WHITE”になれば、その立方体領域内にはデータが存在しないことになり処理を終了する。これらの交差判定をすべての視点で行い、最終的にその立方体の属性を決定する。ただし、1枚でも交差しなければ、属性は“WHITE”となる。

この交差判定では、図8のように立方体のそれぞれの頂点  $V$  を、投影面  $S_j$  に投影する。このときの投影された8つの頂点の画像座標は  $v = P_j V$  となり、これらを頂点とする立方体の多角形シルエットを生成する。このように、指定した階層レベルまで再帰的に処理を繰り返すことによって、おおまかな形状から詳細な形状へと復元される。

【Octree 表現によるシルエット画像統合のプロセス】

ROOT-CUBE を定義

LEVEL = 0

Octree

LEVEL + 1

for SUB-CUBE = 1 to 8 do

for 視点:  $j = 1$  to  $n$  do

交差判定

if 完全な内部 then 属性値 1 に決定

if 完全な外部 then 属性値 0 に決定

if 交差 then 再帰的に Octree を繰り返す

交差判定

for 立方体の頂点:  $V = 1$  to 8 do

$v = P_j V$

8つの  $v$  から立方体の多角形シルエットを生成

if 物体シルエット内に多角形シルエット

then BLACK

if 物体シルエット外に多角形シルエット

then WHITE

if 物体シルエットと多角形シルエットが交差

then GRAY

### 3.3.2 復元の精度と信頼性

octree 表現による復元精度は画像の解像度、及び octree の階層レベルに依存する。octree の精度に最も大きく関係し、計算機の処理能力に応じて設定可能なのは、階層レベルである。一般には、1つ階層レベルを上げることで、3次元的に精度は8倍になる。この精度は、入力画像から得られたシルエット画像が持つ形状情報をいかに正確に表現できるかのレベルである。よって、入力画像の解像度以上の精度は得られない。例えば、解像度  $800 \times 600$  の入力画像の場合、LEVEL 10 以上再帰を繰り返しても精度は変わらない。これは ROOT-CUBE を投影した画像の縦横の解像度が octree の分解度より

も小さくなるためである。従って、LEVEL 10 では画像解像度が 1024x1024、LEVEL 11 ではさらに 2048x2048 もの画像解像度が必要となる。一般には、以下の条件を満たす必要がある。

$$\diamond \text{ROOT-CUBEを投影した縦横の解像度} > 2^{\text{LEVEL}}$$

我々は復元精度の向上のために、復元処理のパラメータとして曖昧度を設定する。一般に、物体のシルエットと octree のシルエットとの交差判定は、撮影したすべての画像に対して行う必要がある。しかし、キャリブレーションによる計測誤り、陰影による不完全なシルエット抽出を考慮すべきである。そこで、交差判定をすべての画像に対して行わず、数枚の余裕を持たせる。つまり、たとえ交差判定が“WHITE”になろうとも、曖昧度として定義した撮影枚数までは処理を終了させずに続行する。この余裕は、処理速度が若干遅くなるが、データの欠落した箇所や誤差に対して非常に有効に働き、全体としての信頼性は下がるが、精度は向上する。

## 4 実験

### 4.1 カメラキャリブレーション精度

中心射影行列によるカメラキャリブレーションの精度評価を実際の撮影環境下で行う。キャリブレーションボードは 34cmx34cm の 3 枚からなり、2cmx2cm のチェックパターンが描かれている。評価方法は、3 次元上の点  $M$  と画像上の点  $m$  から、式 (1) より中心射影行列  $P$  を求める。そして、得られた  $P$  行列から再び 3 次元点  $M$  により、その点の画像座標  $m'$  を算出し、すべての特徴点での平均誤差  $E(\text{pixel})$  を求める。

$$sm = PM \rightarrow P \quad (3)$$

$$m' = PM \quad (4)$$

$$\text{平均誤差 } E = \frac{1}{N} \sum_{i=1}^N |m'_i - m_i| \quad (5)$$

この結果、中心射影行列による平均誤差は、 $E = 0.117(\text{pixel})$  となった。この誤差は手動で特徴点の画像座標  $m$  を求めた際のずれであると考えられる。また、別視点から同様に誤差を求めた結果、 $E = 0.141(\text{pixel})$ ,  $E = 0.178(\text{pixel})$  となり、誤差は  $0.1 \sim 0.2(\text{pixel})$  の範囲内に収まる良好な結果を得ることができた。

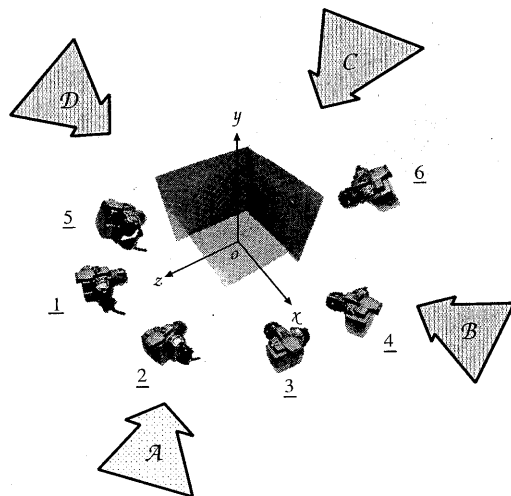


図9: 撮影視点

### 4.2 任意視点からの3次元形状復元

ここでは、ソリッドな物体と骨格オブジェクトを用いて復元を行う。撮影には CCD カメラ、MINOLTA RD-175(175 万画素) を用い、画像解像度を 800x600(pixel) とする。

まず、物体をキャリブレーションボードの中心に置き、A から水平方向に 4 視点、上方から 2 視点、撮影を行った。次に、横 2 枚のキャリブレーションボードを 90 度移動させながら、それぞれ B、C、D から 6 視点ずつ、計 24 の任意視点から撮影を行った(図9)。

図 10 に 2 種類の物体の復元結果を示す。復元の際の曖昧度は  $0.125(= 3/24)$  とした。それぞれ、撮影した入力画像中の 4 枚、そのシルエット画像、本手法により復元した 3 次元データを VRML で記述、表示したものである。

Teapot ではデータ量が増えるために、階層レベル 8 までの復元を試みた。骨格オブジェクトの Bird では、頭部や胸部など、従来の回転テーブルでは復元できなかった箇所についても正確な復元結果を得た。

## 5 おわりに

本研究では、octree 表現を用いることで、ソリッドな物体だけではなく、骨格オブジェクトなどの複雑な物体に対しても正確な復元結果を得ることができた。また、任意の視点・焦点距離で撮影するこ

とによって、従来の回転テーブルでは復元が困難であったオクリュージョン部分に対しても本手法の有効性を実証できた。

ここで、本手法における復元結果の精度ついて様々な視点から考察を重ねる。復元の精度に関しては幾つかの要因があるが、まず第一に、安定した中心射影行列の算出が挙げられる。本手法では、任意視点からの画像を生成するために、中心射影行列によるカメラキャリブレーション法を行い、その有効性を示した。しかし、任意視点からの撮影には、カメラから物体までの距離、また焦点距離がすべてにおいて異なるためにカメラ歪みを考慮していない。そこで、広角レンズで撮影した場合には、キャリブレーション、シルエット抽出の誤差が無視できないものとなる可能性がある。今後、任意視点の撮影に適したカメラ歪みの算出法が必要となる。

また、今回手動で行ったカメラキャリブレーションでの3次元空間上の点と、2次元画像上の点との対応づけを自動化する必要がある。さらに、中心射影行列 $P$ の算出では、初期値の解が不安定であり、6つの点の取り方により初期値の値が大きく変わり精度に影響を及ぼした。今後は3次元点の特徴点の算出についてさらなる考察が必要である。

また第二の要因として、物体のシルエット抽出の精度が挙げられる。本手法により、キャリブレーションとシルエット抽出を同時に行うことが可能になった。今後は閾値の自動化と、抽出誤りの判定法について考えていく必要がある。

最後の要因は、入力画像枚数を増やすことによって、モデリングエラーを減らすことである。今回は従来の回転テーブルに比べ、撮影枚数が少なかったために多くの箇所でもデリングエラーが発生した。今後は、任意視点からの撮影枚数と可搬性をいかに両立させるかが課題である。

なお今後は、精度・応用の両面からこれらの問題を解決するために、全方位からのビデオカメラによる撮影を行う。そして、動画像による動き拘束からキャリブレーションパターンの特徴点の追跡を行い、安定したカメラパラメータの算出、ロバストなシルエット抽出を行う予定である。また、データ量の削減やリアルティの向上のために、物体の形状に応じたテクスチャマッピング、ボクセルカラリングについてさらなる研究を行う。

## 参考文献

- [1] 日浦, 佐藤, 井口: “対象物体の回転による形状と反射率の同時計測,” 情処論, Vol.36, No.10, pp.2295-2302 (1995)
- [2] 大塚, 鄭, 田中, 安部: “多種類の視覚特徴の分離と融合による3次元モデルの復元,” 信学論 (D-II), Vol.J80-D-II, No.9, pp.2324-2332 (1997)
- [3] Potmesil, M.: “Generating Octree Models of 3D Objects from Their Silhouettes in a Sequence of Images,” *CVGIP*, pp.1-29 (1987)
- [4] Richard Szeliski: “Real-Time Octree Generation from Rotating Objects,” *Digital Equipment Corporation Cambridge Research Lab, Technical Report* (1990)
- [5] A.W.Fitzgibbon, G.Cross, A.Zisserman: “Automatic 3D Model Construction for Turn-Table Sequences,” *Proc.SMILE Workshop*, pp.154-170 (1998)
- [6] W.Niem: “Robust and Fast Modelling of 3D Natural Objects from Multiple Views,” *SPIE Proceedings "Image and Video Processing II"*, Vol.2182, pp.388-397 (1994)
- [7] Y.Matsumoto, H.Terasaki, K.Sugimoto, T.Arakawa: “A Portable Three-dimensional Digitizer,” *Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pp.197-204 (1997)
- [8] S.M.Seitz, C.R.Dyer: “Photorealistic scene reconstruction by voxel coloring,” *Proc. Computer Vision and Pattern Recognition Conference*, pp.1067-1073 (1997)
- [9] S.J.Gortler, R.Grzeszczuk, R.Szeliski, M.F.Cohen: “The lumigraph,” *Proc. SIGGRAPH 96*, pp.43-54 (1996)
- [10] W.Niem, J.Wingberm: “Automatic Reconstruction of 3D Objects Using a Mobile Monoscopic Camera,” *Proceedings of the International Conference on Recent Advances in 3-D Imaging and Modeling* (1997)
- [11] R.Y.Tsai: “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision,” *IEEE Conf. CVPR*, pp.364-374 (1986)
- [12] S.Feiner, B.MacIntyre and D.Seligmann: “Knowledge-Based Augmented Reality,” *Comm. ACM*, Vol.30, No.7, pp.53-62 (1993)
- [13] 佐藤, 山本, 田村: “現実空間と仮想空間の位置合わせ手法 - ステレオカメラと3次元センサの組合せ -,” 画像の認識・理解シンポジウム (MIRU'98) (1998)
- [14] 徐剛, 辻 三郎: “3次元ビジョン,” 共立出版 (1998)

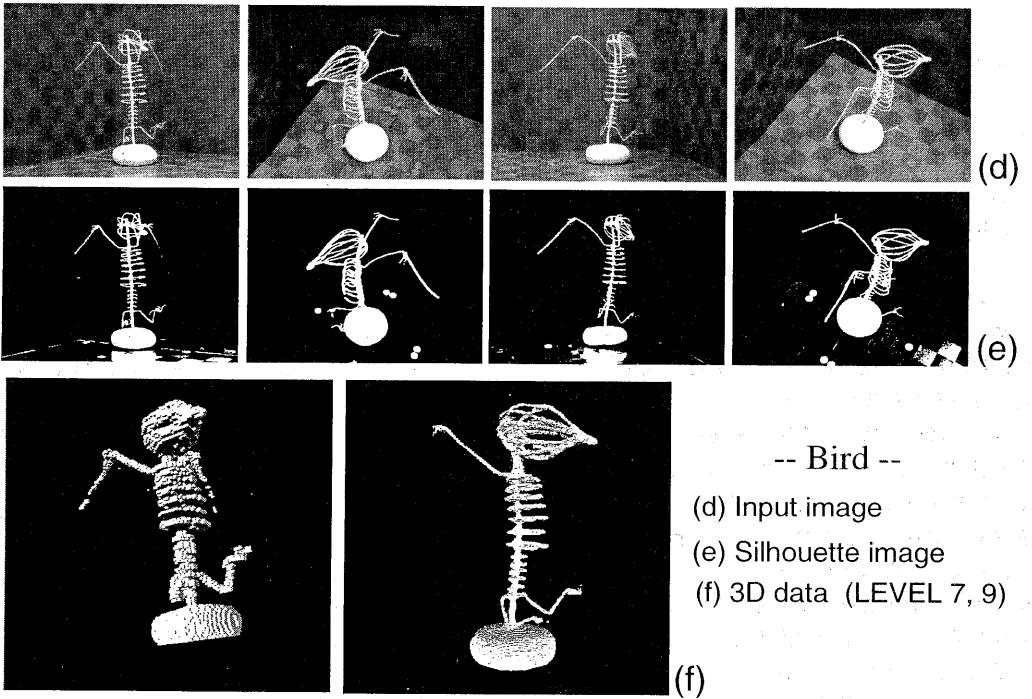
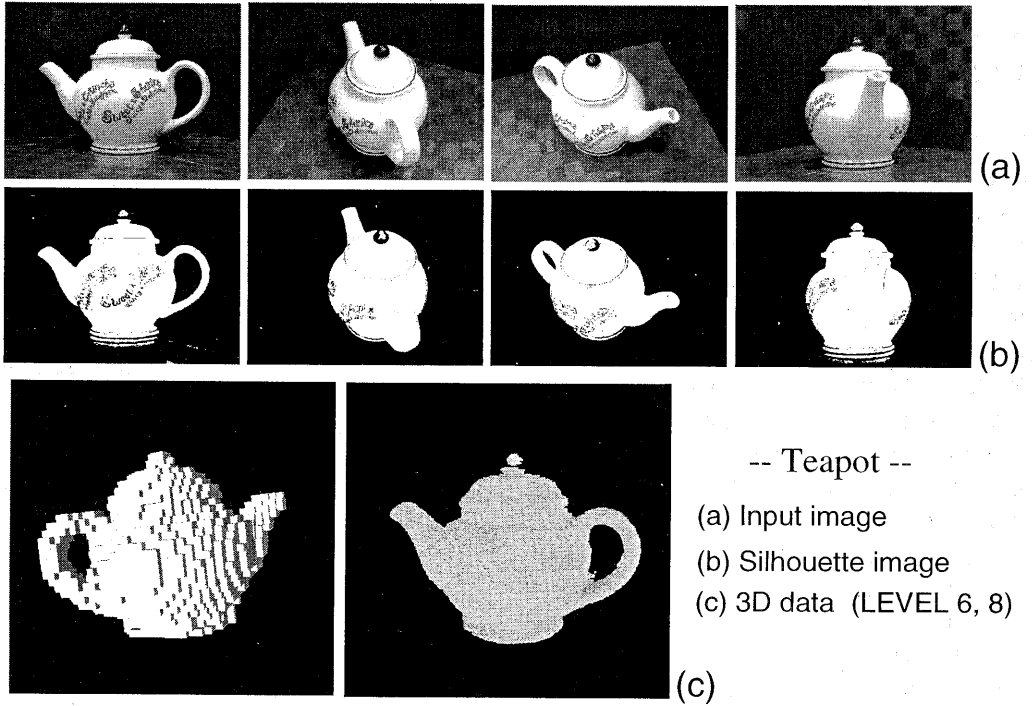


图 10: 復元結果