

# 逆ボリュームレンダリングを用いた 多視点画像からの 3 次元モデル復元

山崎俊太郎<sup>†</sup> 持丸 正明<sup>†</sup> 金出 武雄<sup>†</sup>

<sup>†</sup> 独立行政法人 産業技術総合研究所  
デジタルヒューマン研究センター

〒 135-0064 東京都江東区青海 2-41-6 産業技術総合研究所臨海副都心センター 3F

E-mail: †{shun-yamazaki,m-mochimaru,t.kanade}@aist.go.jp

あらまし 本論文は、微細な複雑形状を持つ対象物の 3 次元空間情報を、多視点画像からボクセルモデルを用いて表現する方法を提案する。樹木や毛髪などの微細な幾何構造を持つ物体は、対象物を正確に境界面表現することが難しく、通常の画像取得手法では空間解像度が不足するため、計測画像から正確なモデルを作ることが困難である。提案法は、ボクセルモデルを用いて対象物の色と空間占有率をモデル化することによりこの問題を解消する。対象物表面の光の伝播をボリュームレンダリング方程式で記述し、空間占有率をボクセルの不透明度 ( $\alpha$  値) として推定する。実験では、複数の複雑形状物体に対して本手法を適用し実験を行い、生成されたボクセルモデルから元の計測画像を再現できることを示す。

キーワード 形状復元, 多視点画像,  $\alpha$  推定, ボクセルモデル, ボリュームレンダリング

## Inverse Volume Rendering Approach to 3D Reconstruction from Multiple Images

Shuntaro YAMAZAKI<sup>†</sup>, Masaaki MOCHIMARU<sup>†</sup>, and Takeo KANADE<sup>†</sup>

<sup>†</sup> Digital Human Research Center National

Institute of Advanced Industrial Science and Technology (AIST)

Water Front 3F, 2-41-6, Aomi, Koto-ku, Tokyo 135-0064, JAPAN

E-mail: †{shun-yamazaki,m-mochimaru,t.kanade}@aist.go.jp

**Abstract** We propose a novel method of volumetric scene reconstruction from multiple camera-images. It is difficult to create an accurate model of intricately-shaped objects, such as, fur and hair which have complicated geometric structure on their surface. We generalize conventional voxel-based method, attempting to estimate spatial occupancy, as well as color, of an object which we want to model. We have modeled an imaging process by the combination of the matting equation and the volume rendering equation. The volumetric model is generated through EM method. Experimental results show that intricately-shaped objects can be modeled by our proposed method, and the original and novel view-images of the objects can be synthesized by volume rendering.

**Key words** 3D reconstruction, multiple camera-images, opacity estimation, voxel modeling, volume rendering

### 1. はじめに

実世界にある対象物の 3 次元空間情報を、カメラで計測した画像から復元する処理は、コンピュータビジョンにおける主要な研究分野である。近年では特に、画像計測や画像処理の技術が進歩したことにより、多視点で同時に観測した画像

を元に、対象風景を全方向的に 3 次元復元する研究が精力的に行われている。デジタルカラー写真を入力として空間情報をモデル化し、対象物や風景をコンピュータグラフィクスを用いて写實的に再現する技術は、電子商取引、仮想現実感、文化財保存など、幅広い分野における応用が期待されている。現在までに、コンピュータビジョンおよびコンピュータグ

ラフィクスの分野を中心として、多くの3次元復元手法が提案されてきた。ところが、既存手法の多くは、対象物が平面や滑らかな曲面で構成されていることを仮定しており、その適用範囲は限られている。特に、樹木や毛髪などの微細な幾何構造を持つ物体の復元は、依然として困難である。これには主に次の二つの理由が考えられる。

第一に、微細な幾何構造を持つ対象物は、形状を境界面表現することが難しく、正確な幾何モデルを構築できない。既存の多くの形状復元手法は、対象物体の表面位置を画像から推定することを目的としている。対象物の表面が平面や滑らかな曲面で構成される場合、境界面表現は効率のよい表現形式である。ところが対象物の幾何形状が複雑な場合、境界面表現を用いて形状表現することは、不可能でないとしても極端に非効率である。

第二に、デジタルカメラなど通常の光学センサーでは、空間解像度が不足するため対象物を正確に計測することができない。計測画像中で表現できる最小単位は画素であるが、例えば一本一本の髪の毛の太さは画像面上で画素サイズ以下である。したがって、計測画像から対象物の正確な形状を推定することができない。

対象物が複雑な幾何形状を持つ場合、その形状を正確に復元することは困難である。ところが、計測画像中では、空間解像度が不足しているにも関わらず、対象物とその詳細部も含めて写実性を失っていないことがある。すなわち、画像合成による再提示を目的とした場合、対象物の空間情報の復元は、必ずしも正確である必要はないことが分かる。画像合成を行うのに十分な解像度で、3次元空間情報をモデル化すればよい。

本論文では、ボクセルモデルを用いて、対象物の幾何情報を空間占有率を用いてモデル化する方法を提案する。具体的には、対象物体の空間占有率はボクセルの不透明度 ( $\alpha$  値) として推定し、ボリュームレンダリングによって自由視点画像を合成する。実験では、複数の複雑形状物体に対して本手法を適用し実験を行い、生成されたボクセルモデルから元の計測画像や新しい視点画像を合成できることを示す。

## 2. 関連研究

### 2.1 多視点画像に基づく形状復元

複数の画像から対象物の3次元空間情報を復元する方法は、コンピュータビジョンにおける主要な研究テーマであり、様々な組み合わせの計測装置、カメラ配置、対象物に対し、多くの復元アルゴリズムが提案されてきた。

2つ以上の画像から対応点を元に深度情報を復元するステレオ法 [9] は、最も初期に提案された形状復元手法であり、多くの研究者によって視点画像に拡張されている。Okutomiら [11] は、対応点の探索に用いる画像輝度の相関演算を、複数のカメラ間で同時に行うマルチベースランステレオ法を提案し、ノイズなどの影響による誤対応の発生を減少させることに成功した。対象物を全周から計測してステレオ法を適用する際には遮蔽の問題が発生するが、Kang [5] らは遮蔽部分

をあらかじめ検出することによって形状復元精度を向上する方法を提案している。

画素単位での明示的な対応点探索を必要としない形状復元法として、視体積交差法が提案されている [?]。この方法は、対象物の輪郭線の凸包を用いて形状復元する。そのため、遮蔽の問題が発生しないという利点がある一方で、計測画像が少ない場合には復元精度が悪く、凹面は原理的に復元できないという欠点がある。

Seitz らは、対象物を含む3次元空間をボクセル表現し、各画像で色の整合性が保たれるように、各ボクセルに色づけを行うボクセルカラーリング法 [14] を提案した。計測画像の背景をあらかじめ検出しておくことにより、この方法は、視体積交差法の一般化とみなすことができる。色の整合性の判定を行うことにより、ステレオ法と同様に画像間の対応点を決定する必要があるが、ボクセルを一定の順序で処理することにより、遮蔽の影響が暗黙のうちに考慮することができるという利点がある。Seitz らの方法はボクセルの処理の順序やカメラ視点に関して大きな制限があるが、この点はKutulakosら [6] によって提案されたスペースカービング法で解消されている。ただし、これらの方法は対象物を密な微小立方体 (あるいは直方体) の集合で表現するため、微細な幾何構造を扱うことが難しい。ボクセル解像度を上げることによって、これらの方法で微細構造を扱うことは原理的には可能だが、計算機の記憶容量の制限から現実的な方法ではない。

本論文は、スペースカービング法を出発点とし、微細な幾何構造を扱うこと可能な方法を提案する。具体的には、各ボクセルを密な幾何要素として扱うのではなく、ボクセル内に対象物が存在する空間占有率を推定することによって、ボクセルの解像度を上げることなく微細構造を表現可能にしている。

### 2.2 不透明度推定 ( $\alpha$ 推定)

デジタル画像では、表現できる計測対象の詳細度が特定の解像度で制限される。カメラによる計測画像においては、対象物表面上の複数の異なる点から発せられた光線が同一の画像素子に入射し、入射した全光線の平均強度に比例した輝度が記録されるという現象が起こる。

画素中で平均化された複数の光線を分離する方法は、コンピュータグラフィックスのための画像処理方法として研究が始まった。Smith らは、写真中の風景を前景と背景に分類し、背景色を変更することにより計測画像中の各画素の前景色成分  $F_p$  と背景色成分  $B_p$  を分離する方法を提案した [16]。画素色  $C_p$  に対する  $F_p$  と  $B_p$  の関係は、画素中の前景物体色の割合を表す  $\alpha_p \in [0, 1]$  を用いて、次の式で記述される。

$$C_p = F_p + (1 - \alpha_p)B_p \quad (1)$$

式 (1) を *Matting* 方程式と呼ぶ。 $\alpha_p$  は前景の不透明度を表し、 $\alpha_p = 0, \alpha_p = 1$  の場合がそれぞれ完全に透明、完全に不透明な場合に相当する。

画像中で確実に前景、確実に背景の領域をユーザーが明示的に与え、前景色、背景色の局所類似性を仮定して単一画像から  $\alpha_p$  を推定する方法が、Mishima ら [10] によって提案さ

れている。また, Ruzon ら [13], Chuang ら [3] によって, より精巧な確率モデルを使って  $\alpha$  推定を行う方法が提案されている。

画像中の前景物体の  $\alpha$  推定を, 3次元空間で行う方法も提案されている。Szeliski らはステレオ法により深度復元を復元する際に, 各深度に仮想的な平面を考え, その平面上にある前景物体の不透明度を推定する方法を提案した [17]。Bonet らはボクセルを用いて半透明物体の形状を復元する手法として, 各画像に含まれる視線にあるボクセルが, 画素に対して与える色の影響を推定し, これを元にボクセルの不透明度を推定する, Roxels 法を提案した [1]。Broadhurst らはスペースカービング法で色の整合性を評価する際に, ボクセル内の色の分布を確率モデルを用いて表現することによって, 形状復元精度を向上させた [2]。また粗いボクセルモデルを用いて樹木の大きな空間占有率をモデル化する方法が, Reche-Martinez らによって提案されている [12]。

本論文では, 入力画像に対する画像上であらかじめ  $\alpha_p$  を推定しておき, これを用いて対象物の色と空間占有率のボクセルモデルを復元する方法を提案する。

### 3. 逆ポリュームレンダリング

#### 3.1 仮定

本論文では, 時間変化しない対象風景を扱う。また, 対象物における光線の反射は, 均一ではないが等方的であると仮定する (Lambertian 条件)。この仮定の下で, 対象風景内の各点における放射輝度は単一の色として観測される。また, 対象風景は, 3次元空間情報をモデル化したい前景物体  $O$  と, それ以外の背景  $B$  から構成され, 背景成分は第 2.2 節で紹介した  $\alpha$  推定法によってあらかじめ除去できるものとする。

アルゴリズムの入力は,  $O$  の全周の異なる  $N_{view}$  視点で計測したカラー画像である。カメラ位置に制限はないが, 計測されない部分は正しくモデル化できない。カメラの内部, 外部, 色感度パラメータは既知とする。アルゴリズムの出力は, 一辺の長さが  $w$  の立方体ボクセル  $v_i$  から構成されるボクセル集合 (ポリューム) である。各ボクセル  $v_i$  は, RGB 色  $c_i$  と不透明度  $\alpha_i$  を持つ。ポリュームの座標系は,  $O$  がポリューム内部に含まれるように, 適宜設定されているものとする。また, 全てのボクセルは, 計測画像上で画素サイズより十分大きいものとする。

#### 3.2 Volume Rendering 方程式

対象物  $O$  が CCD 素子などを用いてデジタル画像として計測される際, 単一画素で観測される色  $C_p$  は, カメラ中心と画素によって定義される四角錐の内部から入射する放射輝度の和に比例する。四角錐が対象物表面の一部  $S_f$  によって部分的に遮蔽されている場合, 観測される放射輝度は,  $S_f$  とその背後にある対象物表面  $S_b$  の放射輝度の線形和となる。したがって, 四角錐から入射して画素中で観測される色は次の離散積分で記述できる。

$$C_p = \sum_{i=s} \alpha(i)c(i) \prod_{j=0}^i (1 - \alpha(j)), \quad (2)$$

$s$  は計測視点から測った視線上の距離を表すパラメータである。 $c(s)$  は外部光源によってもたらされる距離  $s$  の点における放射輝度 (色),  $\alpha(s) \in [0, 1]$  は距離  $s$  の位置にある対象物が背後の物体を遮蔽する割合であり, 不透明度と同等である。この式は, 2 次反射の影響を無視した場合の光の振る舞いを記述する式とみなす事ができ, Volume Rendering 方程式 [7] と呼ばれる。

いま,  $S_f$  が  $S_b$  を遮蔽する割合が 3次元空間中の座標で表せると仮定すると, Lambertian 条件より, 式 (2) 中の  $c(s)$  と  $\alpha(s)$  は 3次元座標の関数となる。そこで, 空間をボクセル分割して離散化し,  $c(s)$  と  $\alpha(s)$  を各ボクセルに割り当てる。ボクセル  $v_i$  の色を  $c_i$ , 遮蔽度 (不透明度) を  $\alpha_i$  とすると, 次の式を得る。

$$C_p = \sum_{i=s} c_i \alpha_i \prod_{j=0}^i (1 - \alpha_j), \quad (3)$$

#### 3.3 Matting Equation

画像から対象物表面に位置するボクセルの値を推定するために, ボクセル値  $c_i$ ,  $\alpha_i$  を計測画像と関連付ける必要がある。式 (3) はボクセル値と画素値の関係を記述しているが, 一般に  $C_p$  は, 対象物表面以外の放射輝度も含むため, 不要な背景成分を除去する。

ボクセル格子が 3次元空間中に無限に広がっているとすると, 計測画像視点から伸びる視線上にボクセル座標を定義することができる。各視線のうち  $O$  に向かって伸びるもの考えると, 視線上に  $O$  と  $B$  を分離する点が存在する。式 (3) を, この点の手前側 (front) と, 奥側 (back) で分離すると, 次の式を得る。

$$C_p = F_p + (1 - A_p)B_p \quad (4)$$

ただし

$$A_p = 1 - \prod_{k \in \text{front}} (1 - \alpha_k) \quad (5)$$

$$F_p = \sum_{i \in \text{front}} c_i \alpha_i \prod_{j=0}^i (1 - \alpha_j) \quad (6)$$

$$B_p = \sum_{i \in \text{back}} c_i \alpha_i \prod_{j \in \text{back}} (1 - \alpha_j) \quad (7)$$

である。

式 (1) より, 式 (4) は Matting 方程式である。したがって, 第 2.2 節で述べた方法を用いて計測画像からあらかじめ  $A_p$ ,  $F_p$  を計算しておくことにより, 式 (5), 式 (6) を用いてボクセルの色  $c_i$  と不透明度  $\alpha_i$  を推定することができる。本論文では, この推定を逆ポリュームレンダリングと呼ぶ。

#### 3.4 拘束式の導出

通常, 逆ポリュームレンダリングは拘束が不足する逆問題である。その原因としては, 第一に, 計測画像中で対象物が計測されている画素の数は, ボクセル数よりも少ないことであり, 第二に, 完全に不透明なボクセル ( $\alpha_i = 1$ ) の背後にあるボクセルの特性の推定は原理的に不可能であることがあ

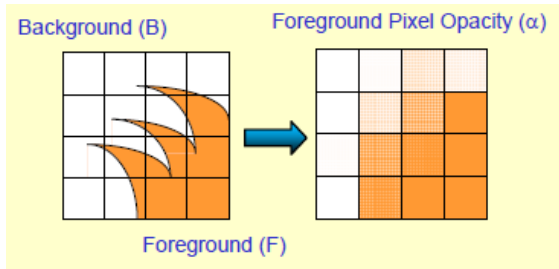


図 1 画素上での光線の混合

Fig. 1 Blending of rays

げられる．この問題を避けるために、われわれは 2 段階の推定方法を提案する．まず第一段階で、式 (5) を用いて  $\alpha_i$  のみを推定し、第二段階では  $\alpha_i$  を第一段階で得られた値に固定し、式 (6) を用いて  $c_i$  を推定する．

第一段階の推定では、計測画像から得られた  $\alpha$  画像中で  $A_p \neq 1$  を満たす各画素に対して、透明度  $T_p = 1 - A_p$ 、ボクセルの透明度  $t_i = 1 - \alpha_i$  を用い、式 (5) の両辺の  $\log$  を取ることで、次の式を得る．

$$\log(T_p) = \sum_i \log(t_i) \quad (8)$$

前処理として計測画像の  $\alpha$  推定を済ませておくことにより、式 (8) は  $\log(T_p)$  が定数、 $\log(t_i)$  が未知数の線形方程式となる．

第二段階の推定では、計算によって得られた  $t_i$  (したがって  $\alpha_i$ ) を定数とみなし、式 (5) を整理することにより次の線形方程式を得る．

$$F_p = \sum_{i \in \text{front}} \left[ \alpha_i \prod_{j=0}^i (1 - \alpha_j) \right] c_i \quad (9)$$

式 (9) 中では、 $[\dots]$  が係数、 $F_p$  が定数、 $c_i$  が未知数である．

## 4. 実装

### 4.1 EM 法による逆投影

通常、式 (8)、式 (9) は未知数が式の数よりも多く、何らかの推定を使うことによって、最適解を選択する必要がある．方程式を解く上で最大の問題は、係数行列のサイズが巨大である点にある．式 (8) や式 (9) の係数は最大で、全ボクセル数と全画像中の画素数の積に等しい．このようなサイズの係数行列を通常の計算機で一度に扱うことはできない．

そこで、本論文では Maximum Likelihood-Expectation Maximization (ML-EM) 法 [8] を用いて式 (8)、式 (9) を解く．この方法は CT や MRI などの断層撮影における立体復元の際に、少ない計測画像から歪みの少ない空間情報を復元するための反復推定法である．

ML-EM 法の統計学的な理論背景に関しては、Lange ら [8] によって詳しく解析されているためここでは省略する．この推定法には以下に述べるような利点がある．

(1) 推定値が負にならない．

(2) 画像が少数であっても、単純な back-projection を行った場合に現れる縞状のノイズが現れにくい．

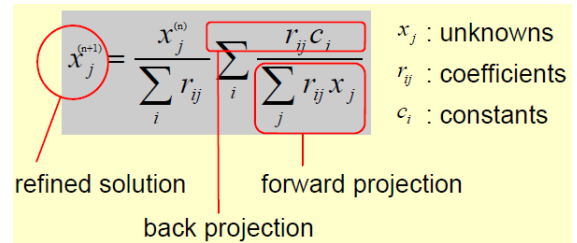


図 2 ML-EM の更新式の解釈

Fig. 2 Interpretation of the update law in ML-EM

ML-EM 法を用いて解くべき連立方程式の未知数の、 $n$  回目の反復における推定結果の集合を  $\{x_j^{(n)}\}$ 、係数行列の要素を  $\{r_{i,j}\}$ 、定数の集合を  $\{c_i\}$  とすると、 $n+1$  回目の反復による推定結果は次の式で表される．

$$x_j^{(n+1)} = \frac{x_j^{(n)}}{\sum_i r_{i,j}} \sum_i \frac{r_{i,j} c_i}{\sum_j r_{i,j} x_j^{(n)}} \quad (10)$$

図 2 に示すように、ML-EM 法による推定は、投影 (forward projection)、逆投影 (back projection)、正規化、推定値の更新、の 4 つの計算要素に分解できる．反復計算の終了判定に関しては明確なルールはないため、再投影によって合成された画像と  $\alpha$  推定によって得られた画像の誤差が、局所最小値に至るまで反復を繰り返す．

ML-EM 法は、収束が遅く、問題のサイズが大きい場合には実用的でない．そこで、計測画像集合をいくつかの組 (subset) に分割しておき、この subset に属するデータだけで投影、逆投影を繰り返すことにより、処理を高速化することができる (Ordered Subset-EM, OSEM 法) [4]．subset に分割することにより、1 回反復の際に推定値を更新する回数が増え、結果として早く収束することが期待できる．実験では、なるべく離れた視点から計測した 4 つ以上の画像の組から成る subset を構成した．

### 4.2 ボクセルの表現方法

ボクセル空間において、対象物が存在しない領域の不透明度  $\alpha_i$  は 0 であり、このようなボクセルの存在は、計測画像の視線にある他のボクセルの  $c_i$  や  $\alpha_i$  の推定に影響を与えない．同様に、完全に不透明なボクセル ( $\alpha = 1$ ) がある場合、視線上でボクセルの背後にあるボクセルも他のボクセルの推定に影響を与えない．そこでこれらのボクセルに関する情報を保持しないことによって、記憶容量の効率を向上させる．

全ての計測画像で  $\alpha$  推定を行った後、 $A_p$  の値から、ボクセル全体を次の 3 種類に分類しておく．

(1) 背景ボクセル: 常に  $\alpha_i = 0$

(2) 内部ボクセル: 常に  $\alpha_i = 1$

(3) 境界ボクセル:  $c_i$  や  $\alpha_i$  を推定するボクセル

ボクセルの分類は次の手順で行う．まず  $A_p = 1$  となる画素の視体積の交差を全画像で考え、凸包内にあるボクセルを内部ボクセルとする．次に、 $A_p > 0$  である画素の視体積の交差を考え、凸包内にあるボクセルのうち内部ボクセルでないものをシェルボクセルとし、残ったボクセルを背景ボクセルとする (図 3)．図 3(a) で赤が内部ボクセル、緑が境界ボク

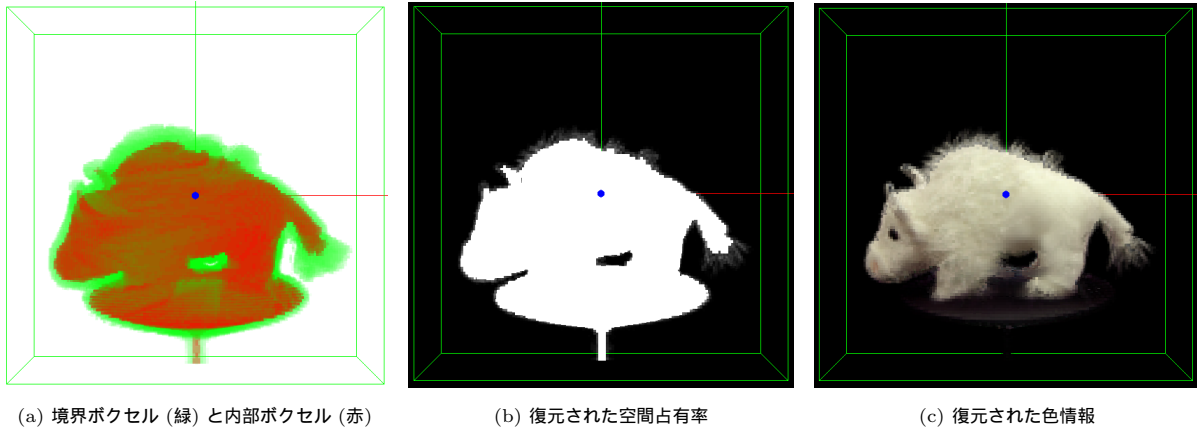


図 3 ボクセルによるモデル化

セルであり、背景ボクセルは表示していない。

### 4.3 LDI による実装

式 (8), 式 (9) の左辺の係数の総数は、(計測画素数)×(ボクセル数)であるため、これらをあらかじめ計算しておくことは、記憶容量の問題から現実的ではない。そこで、ML-EM 法を使って一つの推定値を更新するたびに、必要な係数を  $\alpha_i$  から計算する必要がある。係数の生成には、各画像視点からの光線追跡が必要であり、これを毎回行うことは非効率である。

そこで、全ての画素に対してあらかじめ光線追跡を行っておき、光線と交差するボクセルのリストを LDI [15] として保持しておくことにより、計算を効率化できる。このとき、 $A_p = 0$  の画素は背景ボクセルとのみ交差するため、LDI の構築を省略できる。 $A_p = 1$  の画素を通過する光線は、途中で内部ボクセルと交差するため、ML-EM 推定で必要な逆投影ができないため同様に省略する。

## 5. 実験結果

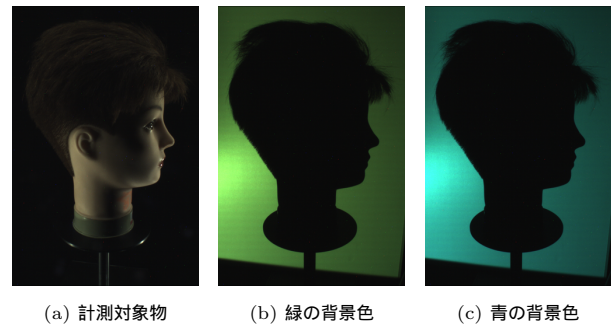
本論文で提案する手法を C++ を用いて実装し、Pentium4 3.4GHz, 主記憶 2GByte の PC 上で実験を行った。計測対象は、毛で覆われた牛の人形、およびマネキンの頭部である。計測時には、対象物を回転台上に設置し、回転台を 10 度間隔で回しながら、ほぼ水平方向に固定したデジタルカメラで 36 回計測した。計測に用いたデジタルカメラは、Tsai [?] の方法を用いてキャリブレーションした。画像解像度は  $640 \times 480$ 、形状復元のボクセルの解像度は  $128^3$  または  $64^3$  を用いた。

### 5.1 $\alpha$ 推定

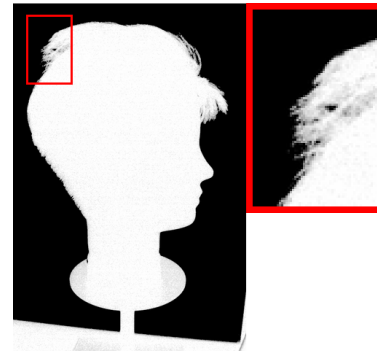
画像から  $\alpha$  推定を行う方法として、2 種類の背景色を利用する Smith らの方法 [16] を用いた。背景色の制御には、液晶プロジェクターを用いた。計測物体の背後に均一な背景色を投影し、各視点において、2 種類の異なる背景色  $C_{k1}, C_{k2}$  (ともに RGB) で対象物を計測する。計測画像中の同一画素における計測色をそれぞれ  $C_{m1}, C_{m2}$  とすると、その画素の不透過度  $A_p$  が次の式で計算できる。

$$A_p = 1 - \frac{(C_{m1} - C_{m2}) \bullet (C_{k1} - C_{k2})}{(C_{k1} - C_{k2}) \bullet (C_{k1} - C_{k2})} \quad (11)$$

ただし演算子 ( $\bullet$ ) は RGB ベクトルの内積を表す。画像の (不



(a) 計測対象物 (b) 緑の背景色 (c) 青の背景色



(d)  $\alpha$  推定で得られた 画像

図 4 複数背景画像を用いた  $\alpha$  推定

Fig. 4 Multi-background matting

透明度込み) 前景色  $F_p$  は次の式で得られる。

$$F_p = (C_{m1} + C_{m2} - (1 - A_p)(C_{k1} + C_{k2}))/2 \quad (12)$$

図 4 に、 $\alpha$  推定の入力画像、および推定された  $\alpha$  画像を示す。

### 5.2 ポリウムレンダリング結果

提案法でボクセルモデルを構築し、ポリウムレンダリングによって仮想視点画像を合成した結果を図 5 に示す。図 5(a) は、計測対象の実画像であり、ボクセルモデル生成には用いていない。図 5(b) は構築したモデルを同様の視点においてポリウムレンダリングした結果である。図 5(c) は別の視点におけるレンダリング結果である。

### 5.3 収 束

第 4.1 節で述べたとおり、OSEM 法の反復計算の終了判定に関しては明確なルールはない。そこで、再投影によって合成された画像と  $\alpha$  推定によって得られた画像の誤差が、局

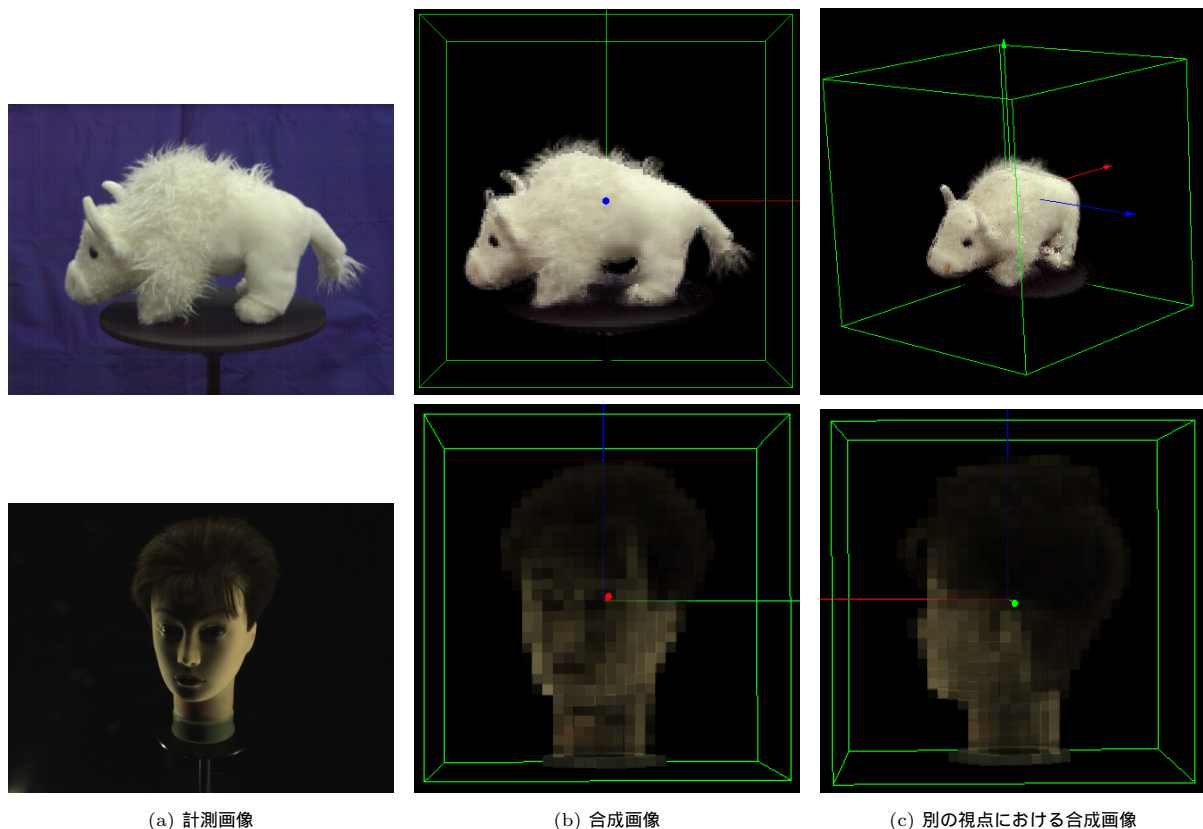


図 5 ボリュームレンダリング結果  
Fig. 5 Results of volume rendering

所最小値に至るまで反復を繰り返した。実験では、実験データに応じて 20 回から 50 回程度で収束するが、実際には最初の数回の反復で、再投影誤差の最小値に十分近い所まで収束しており、それ以上反復を繰り返しても見た目が大きく変わることはなかった。

図 6 上段は、提案法の第一段階である  $\alpha_i$  の反復推定の過程、下段は第二段階である色情報の推定の過程である。再構成ボクセルの解像度は  $128^3$  である。視覚的には、10 回以上の反復を繰り返しても、再投影誤差はほとんど変化しない。

図 7 は、同様の計測データに対する再投影誤差の数値の変動である。ボクセルサイズが  $128^3$  と  $64^3$  のそれぞれの場合の実験結果を示す。最初の数回 (5 回程度) の反復により再投影誤差が極小値に十分近づき、その後、緩やかに収束することが分かる。

#### 5.4 記憶容量と計算時間

計算に要した記憶容量と計算時間は表 1 に示すとおりであった。 $64^3$  と  $128^3$  の 2 種類のボクセル解像度でモデル生成を行った。計算機 (主に記憶容量) の制限により、これ以上のボクセルサイズは扱うことができなかった。また、マネキン頭部モデルに関しては、 $128^3$  ボクセルでも記憶容量が不足して扱うことができなかった。

## 6. おわりに

本論文では、微細な複雑形状を持つ対象物を、ボクセルモデルを用いて表現する方法を提案した。ボクセルを用いた形状モデル化手法を一般化し、対象物が空間中に存在する空間

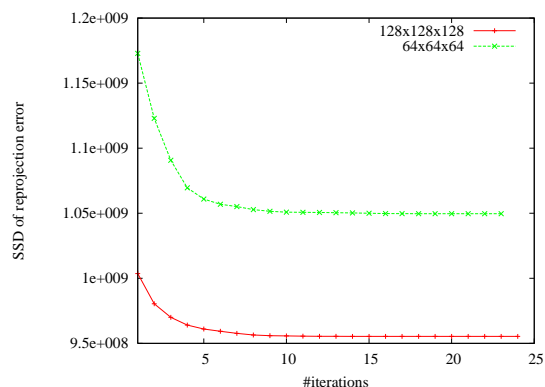


図 7 再投影誤差の収束  
Fig. 7 Convergence

占有率を推定するための、画像とボクセル間に成立する拘束式を導出し、画像からボクセルの色と不透明度 ( $\alpha$  値) を推定する方法を提案した。

図 5 では、生成されたボクセルの  $\alpha$  値が、期待される数

表 1 記憶容量と計算時間. “—” で示す組み合わせは計算できなかった。

Table 1 Performance

計測対象 ボクセル数	境界ボクセル数		記憶容量 (byte)		計算時間 (min)	
	$64^3$	$128^3$	$64^3$	$128^3$	$64^3$	$128^3$
牛の人形	23379	185013	~256M	~2.1G	~30	~186
マネキン頭部	32531	254123	~510M	—	~45	—

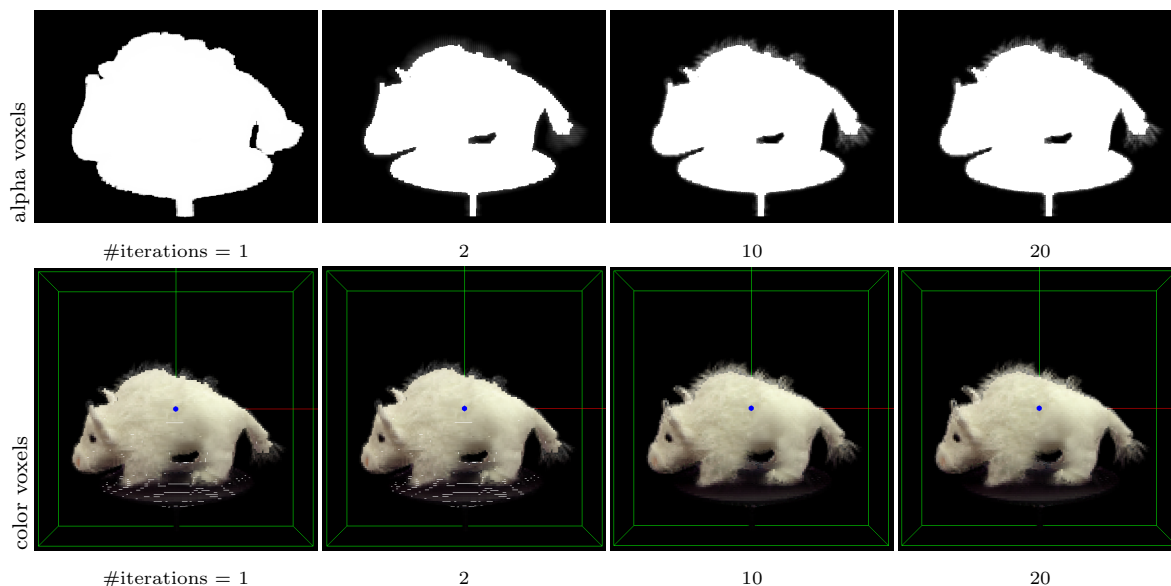


図 6 反復最適化によるボクセルモデルの生成

Fig. 6 Iterative optimization

値よりも高く推定されていることが分かる．ボクセルの  $\alpha$  値の推定の際に対数  $\log(1 - \alpha)$  を用いているため， $\alpha$  が 1 に近い部分で数値が不安定になる．そのため，画像の  $\alpha$  推定の推定誤差が，ボクセルの  $\alpha$  値の推定に大きく影響してしまっていることが原因だと考えられる．特に  $A_p = 1$  付近での誤差に対する影響が大きく，一部の  $A_p$  が誤差により若干低めに推定されることによって，ボクセルの広い範囲で  $\alpha_i$  がかなり低く見積もられるという現状が起こっている．この問題に関して，提案手法は改善の余地が大きい．

提案法は非常に多くの記憶容量を必要とするため，ボクセル空間の解像度を上げることができず，そのために結果画像の画質が損なわれてしまっている．オクツリーを用いた多重解像度表現を使うなどの方法も含め，計算効率の向上，および並列計算機を用いた処理の高速化が今後の課題である．

#### 文 献

- [1] Jeremy S. De Bonet and Paul A. Viola. Roxels: Responsibility weighted 3d volume reconstruction. In *Proc. International Conference on Computer Vision '99*, pages 418–425, 1999.
- [2] Adrian Broadhurst, Tom Drummond, and Roberto Cipolla. A probabilistic framework for space carving. In *Proc. International Conference on Computer Vision 2001*, pages 388–393, 2001.
- [3] Yung-Yu Chuang, Brian Curless, David H. Salesin, and Richard Szeliski. A bayesian approach to digital matting. In *Proc. Computer Vision and Pattern Recognition 2001*, volume 2, pages 264–271, 2001.
- [4] H. Malcolm Hudson and Richard S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(12):601–609, 1994.
- [5] Sing Bing Kang, Richard Szeliski, and Jinxiang Chai. Handling occlusions in dense multi-view stereo. In *Proc. Computer Vision and Pattern Recognition 2001*, pages I:103–110, 2001.
- [6] Kiriakos N. Kutulakos and Steven M. Seitz. A theory of shape by space carving. In *Proc. International Conference on Computer Vision '99*, pages 307–314, September 1999.
- [7] Philippe Lacroute. *Fast Volume Rendering Using a Shear-Warp Factorization of the Viewing Transformation*. PhD thesis, Stanford University, 1995.
- [8] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal of Computer Assisted Tomography*, (8):306–316, 1984.
- [9] D. C. Marr and T. Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London*, B 204:301–328, 1979.
- [10] Y. Mishima. Soft edge chroma-key generation based upon hexoctahedral color space. In *U.S. Patent 5,355,174*, 1993.
- [11] Masatoshi Okutomi and Takeo Kanade. A multiple-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4):353–363, 1993.
- [12] Alex Reche-Martinez, Ignacio Martin, and George Drettakis. Volumetric reconstruction and interactive rendering of trees from photographs. *ACM Transactions on Graphics*, 23(3):720–727, 2004.
- [13] Mark A. Ruzon and Carlo Tomasi. Alpha estimation in natural images. In *Proc. Computer Vision and Pattern Recognition 2000*, pages 24–31, 2000.
- [14] S. M. Seitz and C. M. Dyer. Photorealistic scene reconstruction by voxel coloring. In *Proc. Computer Vision and Pattern Recognition '97*, pages 1067–1073, June 1997.
- [15] Jonathan Shade, Steven Gortler, Li wei He, and Richard Szeliski. Layered depth images. In *Proc. SIGGRAPH '98*, pages 231–242, 1998.
- [16] Alvy Ray Smith and James F. Blinn. Blue screen matting. In *Proc. SIGGRAPH '96*, pages 259–268, 1996.
- [17] R. Szeliski and P. Golland. Stereo matching with transparency and matting. In *Proc. International Conference on Computer Vision '98*, pages 517–523, 1998.