

## 双方向レイトレースモデリングによる形状推定と

### Screened Voxel による形状表現

財満 久義  
(株) 日立製作所 公共情報事業部  
h-zaima@jkk.hitachi.co.jp

山本 強  
北海道大学 工学部  
yamamoto@eng.hokudai.ac.jp

マルチメディア、テレコミュニケーション分野で待望視されている三次元オブジェクトの生成方法と表現方法について提案を行う。三次元オブジェクトは、そのニーズが高い一方、データの作成方法やデータ容量など解決すべき問題が多く残されており、広く普及しているとは言い難い現状にある。三次元オブジェクトをマルチメディアのコンテンツとして位置付けた場合、その生成方法はユーザフレンドリーであることが望まれ、そのデータ構造はインフラフレンドリーであることが望まれる。本研究は、既存物体をいかに簡易にモデリングし、いかに低コストで配信、利用するかに焦点をあてる。提案手法では三次元オブジェクトを双方向レイトレースモデリングという生成方式を用いてモデル化する。この生成方法によりフリーハンドで撮影した複数枚の画像より三次元の幾何情報の推定を行う。また、*Screened Voxel* という表現形式により簡易にデータを記述する。この表現方法では、コアとなるデータ容量が小さいため、現行のネットワーク環境下で低コストの配信が可能となる。

## A Reconstruction using Bi-directional Ray Traversal Modeling

### and a Geometry Expression using Screened Voxels

Hisayoshi Zaima  
Information Systems Group, Hitachi Ltd.

Tsuyoshi Yamamoto  
Faculty of Engineering, Hokkaido University

We propose an approach for the generation and distribution of three-dimensional scenes that is practical on existing networks. In spite of needs and expectations from many fields, three-dimensional visualization is not yet commonly used. Our approach is based on the idea that multimedia data should be both infrastructure-friendly and user-friendly. We represent three-dimensional objects using screened voxel expressions. This representation keeps the data size small enough so that existing network bandwidth is sufficient, and can be rendered efficiently enough (using the 3DDDA algorithm) so that low cost client computers can be used. In this paper we explain how to represent scenes using screened voxel expressions and give some examples.

## 1. はじめに

三次元オブジェクトを計算機内に復元 (Reconstruction) するにはレンジファインダーなどの計測機器を用いることが精度の面で優位にある。しかし、機器そのものが高価であり、広く普及しているとは言いがたい上に、屋外での使用や対象物の大きさなどの制限がある。よって、対象物を撮影した画像情報を基に推定により復元を試みる研究が行われている。あるいはモデリング技術に頼ることになる。しかし、3次元オブジェクトの生成や配信が行えるのはモデリング技術に精通したパワーユーザに限定される上、それらを利用する場合にもデータ量が重いなど、既存のネットワーク上で扱うには不利な点が多いとされてきた。

本報告では、三次元オブジェクトの生成技法として双方向レイトレースモデリング (BR<sub>TM</sub>: Bidirectional Ray Traversal Modeling) と表現手法として *Screened Voxel Model* について報告する。

## 2. Related Works

古典的な幾何情報推定手法として、ステレオ視やSFM (Structure from Motion) などの研究が挙げられる。これらの手法では、幾何情報をいかに正確に復元できるかに焦点があてられてきた。今日では、必ずしも幾何情報を必要とせず、ビュー操作により新規画像を生成する仕組みを提案し、その写実性を評価する研究が盛んになっている。

複数の画像情報から被写体の幾何情報を推定する手法の多くは、被写体表面上の特徴点の抽出と追跡を高精度で推定することにより実現する。Voting手法は、特徴点の追跡を行わないという点で従来手法と異なり、高い評価を得ている<sup>1)2)</sup>。

また、写実性重視という側面からは、イメージベースのアプローチやソリッドテクスチャなどの研究が盛んである。原島らは、イメージベースレンダリングの技術を応用し、ポリゴンモデルの空間に光線空間を生成する手法を提案している<sup>3)</sup>。これにより高い成果を出してはいるものの、データ量が大きい点などが課題としてあげられている。ソリッドテクスチャの研究も写実性向上に寄与している<sup>4)</sup>。仮想空間内にボクセルベースのデータを記述しレンダリングを行う点では本報告の提案

手法と共通する点がある。その他、Voxel Coloring手法などもボクセル空間を利用したオブジェクトの表現手法として注目を集めている<sup>5)</sup>。

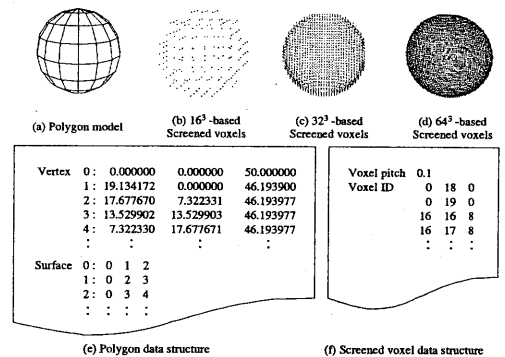


図1 ボクセルモデルとSVM

## 3. 提案手法の概要

### 3.1 Screened Voxel Models

3次元CGにおけるモデリングは、ジオメトリベースのポリゴンモデルとボクセル空間ベースのボクセルモデルとに大別される。提案手法では後者のボクセルモデルを応用する。ここでいうボクセルモデルは、しきい値による等値面表現を可視化する手法とは異なる。提案手法では、初期状態として3次元正方格子のボクセルを等間隔に定義し、各ボクセルのプロパティを特定し、最終的には対象物表面近傍に位置するボクセルのみを *Screened Voxel* として抽出し、そのモデルを *SVM* (*Screened Voxel Model*) という。(図1参照)

初期状態として、対象物を内包するように三次元の均一密度のボクセル空間を定義する。この状態下より、個々のボクセルを対象物表面上 (あるいは表面近傍) に位置しているか否かを特定し、クラス分けを行う。すなわち、ボクセルプロパティ (状態変数) をもって幾何情報を表現する。ボクセルプロパティは次の2種類を用意する。

- *OnVoxel*: 対象物表面上を表す
  - *OffVoxel*: 対象物内部または外部を表す
- 各ボクセルは *OnVoxel* か *OffVoxel* のいずれか一方の状態が与えられる。しかし、幾何情報として意

味を持つのは *OnVoxel* のみである。(図 2 参照)  
*OnVoxel* は、コアデータとして抽出され、配信やレンダリングの対象となる。

本表現方法の特徴を 3 点述べる。

第一に、表現精度の自由度が高い点である。幾何情報の記述精度を追求するならば、形状の局視性に従ってデータの空間的な分布にも粗密が発生することになる。*SVM* では、ボクセル数およびボクセルピッチという 2 つのパラメータを与えることにより表現精度が決定される。すなわち、対象物形状の複雑さの局所性に関わらず、2 つのパラメータにのみに依存した表現精度を持つことになる。形状の局所性に関わらず、一様なデータ分布をもつことになる。画面表示や印刷を対象とする場合は、一様分布が有効となる場合も多い。

第二は、コアデータ容量が小さいことである。ポリゴンモデルでは浮動小数点の頂点データを必要とするため、データ量が膨大になる傾向がある。しかし、*SVM* の場合は、頂点データ、浮動小数点データを必要としないため、コアとなるデータ容量が小さくなる。

第三に、レンダリングの負荷が小さく、アルゴリズムの記述も容易に行える点である。基本的には、デジタル直線の発生とコアデータの参照により、新規ビューのピクセル値を決定し、レンダリング画像を得ることとなる。ローエンドマシンでのレンダリングも可能となる。

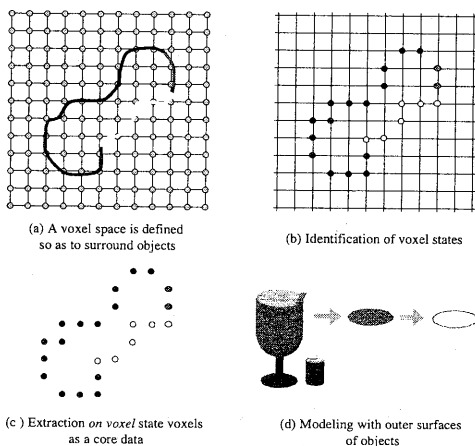


図 2 Screened Voxel によるモデル表現

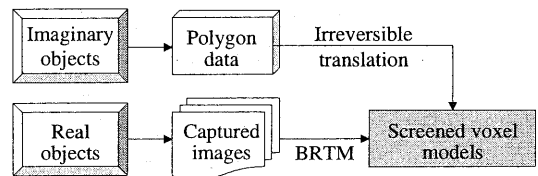


図 3 SVMs の生成

### 3.2. 双方向レイトレースモデリング

モデルの生成は、アプリケーションを用いてポリゴンモデルを記述するものと、画像情報からアプリケーションを用いて推定を行う方法がある。本報告では、後者のアプローチによる双方向レイトレースモデリング (*BRTM*) について説明する<sup>6)</sup>。

推定に必要な情報はフリーハンドビデオ撮影によって対象物画像を取得するものとする。撮影は屋内外を問わず、自然環境下で行うものとする。空間座標の定義、カメラキャリブレーションのためのグラフシートを置くこと以外は背景には手を加えない。よって、被写体背後にはランダムな背景が存在するものとする。ここで、グラフシートに従って、ボクセル空間を定義する。この状態下で、個々のボクセル属性 (*OnVoxel* / *OffVoxel* および *OnVoxel* の場合は色情報も) を特定することによりモデル化を行う。

その基本概念を図 4 に示す。図 4 (a) に示すとおり、初期状態として対象物を内包するようにボクセル空間を定義する。ここで、

- 最終的には *OnVoxel* となるべきボクセル (対象物表面近傍に存在するボクセル)
- 最終的には *OffVoxel* となるべきボクセル (対象物表面外に位置するボクセル)

では、異なる視点からの見えかた (基本的にはピクセル値) の一貫性に違いが生ずる。図 4 (b) に示すとおり、対象物表面近傍に存在するボクセルを異なる視点から参照しようとする、その見え方は基本的には安定している。一方、図 4 (c) に示すとおり、対象物表面外に位置するボクセルでは、異なる視点からの見えかたに一貫性が保証されない。右側の視点からは背景色を得ることとなり、左側の視点からは対象物表面色を得ることになる。

*BRTM* を形状復元技法 (reconstruction) として

評価した場合、特徴点を扱う必要が無いことが第一の特徴としてあげられる。従来の多くの推定手法では、物体形状の境界やテクスチャの境界等を特徴点として抽出する。異なる画像間で特徴点の追跡を行うことにより、その点の奥行情報を得ようとする。よって、十分な数の特徴点をいかに抽出し、誤対応なくいかに追跡するかが問題となる。また、推定結果が特徴点の数や密度に影響される。*BRTM*では、特徴点の抽出と追跡を行わないため、対象物や撮影環境において自由度が高い。

### 3.3 BRTM によるモデル推定

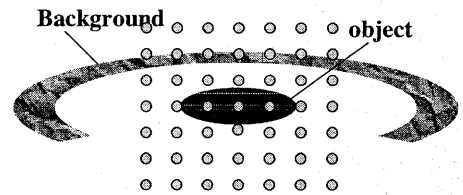
個々のボクセルのクラス分けは *BRTM* の内部処理にて行う。このプロセスは大きく次の3ステップに分類され、ステップは逐次的に処理される。

- バックワードプロセス
- 第一フォワードプロセス
- 第二フォワードプロセス

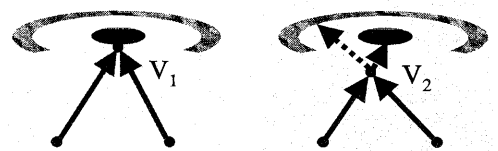
バックワードプロセスでは、図5(a)に示すとおり、バックワードレイを用いた処理を行う。バックワードレイは、各ボクセルの座標とカメラ位置座標により定義される直線である。この直線が投影面と交われば、その交錯点でのピクセル値が獲得できる。例えば、あるボクセルから20個の視点に対してバックワードレイを放った場合、最大で20個のピクセル値を獲得することになる。対象物表面近傍に位置するボクセルであれば、獲得したピクセル値の中に、対象物表面色が含まれている可能性がある。このピクセル色獲得処理を全ボクセルについて行う。これにより、ボクセル毎の候補色リストを獲得することになる。

第一フォワードプロセスでは、カメラ位置座標とピクセル座標（空間座標）により定義される半直線による処理を行う。このフォワードレイは3DDDAによるデジタル直線であり、探索ボクセルリスト（配列）として表現される。

本プロセスの目的は、前プロセスで生成した候補色リスト内の矛盾を発見し、その色候補を失効させることにある。すなわち、異なる視点からの見えかたに関する矛盾を発見し、各ボクセルが図4に示す(b)または(c)のいずれの状態にあるかを検証するものである。レイが持つピクセル色と、



(a) Target object and a voxel space



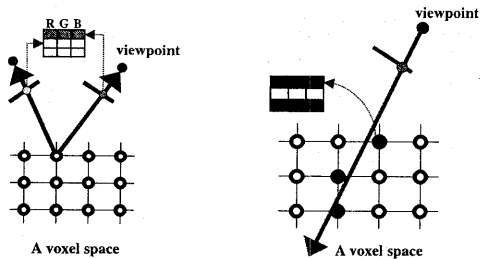
(b)  $V_1$  is on the surface of the object      (c)  $V_2$  is off the surface of the object

図4 BRTMの基本的な考え方

レイが通過するボクセルの候補色の比較照合を行い、見え方に関する一貫性を検証する。比較照合によって、ピクセル色と異なると判断されたボクセル候補色はその時点で「対象物表面色である可能性」を失うこととなり、候補色リストから外される。候補色を全て失ったボクセルは対象物表面外に位置するものと判断され、*OffVoxel*の状態となる。一方、候補色を一色以上保持したボクセルは対象物表面内に位置するものと推定され、*OnVoxel*の状態を持つ。

第二フォワードプロセスは、前述の第一フォワードプロセスと同じレイを用いる。本処理の目的は、*OnVoxel*のもつボクセル候補色を最終的に一色に決定することと、対象物内部に存在するボクセルを一括して失効させ、対象物の外面のみを取り出すことにある。

これら *BRTM* の3段階の処理により対象物表面を抽出し、*Screened Voxel*による三次元オブジェクトが生成できる。



(a) The backward process (b) The forward process

図 5 バックワード処理とフォワード処理

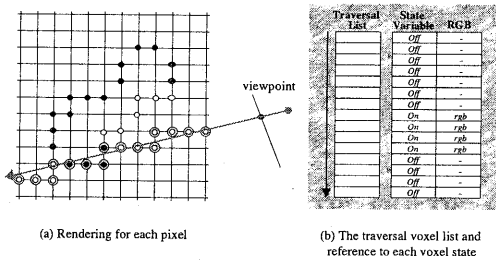


図 6 Screened Voxel Model のレンダリング

#### 4. SVM のレンダリング

SVM のレンダリングは、レイトレーシングと同様、新規視点から各ピクセルを経由してレイを放出し、通過ピクセルの色情報を獲得することにより実現できる。レイトレーシングと異なる点は、3DDDA を用いてレイを表現すること、二次レイ以降は考慮しない点が挙げられる。ボリュームレンダリングとも異なり、レイキャストも行わない。図 6 に示すとおり、視点からレイを発し、各ピクセル経由してボクセル空間内のレイ探索を行う。各レイは 3DDDA によってデジタル直線として放出され、通過ボクセルの ID がボクセル探索リストに格納される。視点から近い順に通過ボクセルの ID がリスト上に格納される。

ここで、コアデータを参照し、通過ボクセルの ID が *OnVoxel* か否かを順次検証する。ボクセル探索リスト上で初めに現れる *OnVoxel* の色情報をも

って新規ビューのピクセル値とする。

#### 5. 評価

モデリングおよびレンダリングの実験例を示す。図 8 (a) に示すとおり、カメラキャリブレーションを行うためにグラフ用紙に対象物を配置している。キャリブレーションは Tsai のアルゴリズムを用いた。

- モデリング  
画像枚数 : 20  
解像度 : 640×480  
24 bit colors  
空間解像度 : 128<sup>3</sup>
- レンダリング  
解像度 : 640×480

図 7 は、対象物の表面抽出が行われる過程を前述の 3 プロセス毎に示している。各段階で *OnVoxel* と状態設定されているボクセルを可視化したものである。

図 8 は、屋外で撮影した 20 フレームの画像より被写体をモデル化し、レンダリングを行った結果である。このモデルの場合、最終段階で、*OnVoxel* と認識されているボクセル数は約 59K 個である。すなわち、空間として定義した 128<sup>3</sup> ボクセルに対して、3.8% が *OnVoxel* として残ったことになる。

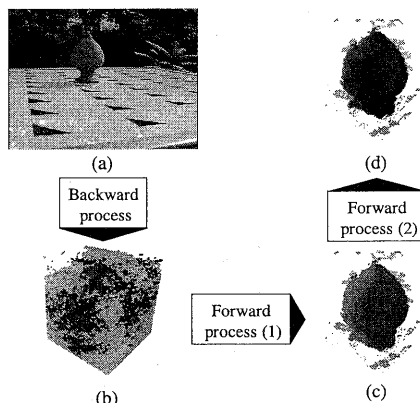
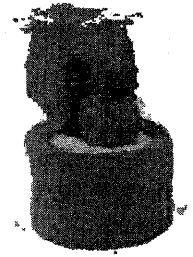
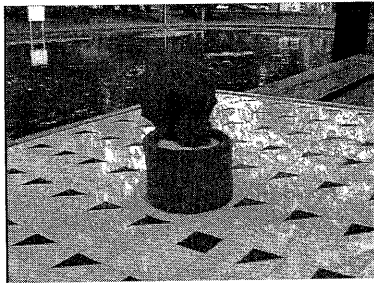


図 7 各プロセス直後の状態



(a) Examples of the input frames

(b) A result of modeling and rendering

図 8 BRTMによる推定事例

## 6. 課題

今後の課題を述べる。

第一に、カメラキャリブレーション環境が挙げられる。キャリブレーションのためにグラフシートを対象物下に敷くため、汎用性が著しく欠けることになる。また、グラフシートは汎用性の欠如とともに、アーチファクトの生成を助長する。このグラフシートは全フレームに共通の背景として撮影されることになり、フォワードプロセスで排除しきれない場合にアーチファクトとして発現することになる。

第二に、各種光学現象への対応である。鏡面反射、ハイライト、半透明処理などがこれにあたる。半透明処理などは SVM のデータ構造を含めて検討を行う必要がある。

## 7. まとめ

三次元オブジェクトは生成の容易性がその優劣を左右する一因と考える。生成した 3D コンテンツの可用性は、そのデータ容量が優劣を左右する一因になると考える。これら双方の視点から検討を行った *Screened Voxel* による表現手法は、検討事項をいくつか残すものの、今後のマルチメディア、テレコミュニケーションの分野で 3D コンテンツを支える一要素としての位置づけが行えたと思う。

## 参考文献

- 1) 浜野輝夫, 安野貴之, 石井健一郎: "空間への voting による 3 次元環境情報抽出手法", 信学論 (D-II), J75-D-II, pp. 342-350 (1992)
- 2) 川戸慎二郎: "空間への 2 段階投票による 3 次元情報の抽出", 信学論 (D-II), J77-D-II, pp. 2334-2341 (1994)
- 3) 内山, 片山, 田村, 苗村, 金子, 原島: "光線空間理論に基づく実写データと CG モデルを融合した仮想環境の実現", 3 次元画像コンファレンス'96, 1-3, pp. 13-18 (1996. 7).
- 4) N. Chiba and K. Muraoka: "Rendering of Forest Scenery Using 3D Textures", The Journal of Visualization and Computer Animation, Vol. 8, pp. 191-199, (1997)
- 5) S. M. Seitz and C. R. Dyer: "Photorealistic Scenes Reconstruction by Voxel Coloring", Proc. of Computer Vision and Pattern Recognition Conference, pp. 1067 - 1073, (1997)
- 6) H. Zaima and T. Yamamoto: "Representation of Three-Dimensional Scenes using Screened Voxel Expressions", Multimedia Information Systems in Practice, pp. 202-212, 1998