

## 自由視点映像における背景モデル作成のための一手法

一色 彩<sup>†\*</sup> 坂本 竜基<sup>‡</sup> 北原 格<sup>†\*</sup> 小暮 潔<sup>†</sup> 城 和貴<sup>\*</sup>

*issyoku@ics.nara-wu.ac.jp*

<sup>\*</sup> 奈良女子大学 大学院 人間文化研究科

<sup>†</sup> ATR 知識科学研究所

<sup>‡</sup> 和歌山大学 システム工学部 情報通信システム学科

<sup>\*</sup> 筑波大学 大学院 システム情報工学研究科

### 概要

自由視点映像における背景領域の形状モデルを簡便に生成する手法を提案する。前景領域の自由視点映像は、視体積交差法などの手法により3次元形状を復元することで実現されている。一方で、背景領域の形状モデルはCADなどの3Dモデリングソフトを用いて作成されることが多い。この場合、詳細なモデルの作成には手間がかかり、また、簡易的に作成した単純な形状のモデルを使用すると映像が不自然なものとなっていた。本提案手法では、背景領域の形状を複数の平面で近似することとし、多視点画像に写り込んだ平面の頂点を対応点として指定するという簡単な処理により背景モデルを生成する。各対応点の3次元座標は、前景領域の3次元推定に用いられる多視点カメラのキャリブレーション情報を基にステレオ処理により算出される。実際の形状に比較的近いモデルを簡易に生成することが可能である。

## A Room Modeling Method for 3D Video

Aya Isshiki<sup>†\*</sup> Ryuuki Sakamoto<sup>‡</sup> Itaru Kitahara<sup>†\*</sup> Kiyoshi Kogure<sup>†</sup> Kazuki Joe<sup>\*</sup>

<sup>\*</sup> Graduate School of Human Culture, Nara Women's University

<sup>†</sup> ATR Knowledge Science Laboratories

<sup>‡</sup> Department of Computer and Communication Sciences, Wakayama University

<sup>\*</sup> Graduate School of Systems and Information Engineering, Tsukuba University

### Abstract

This paper proposes a method to reconstruct the 3D shape model of background regions simply for generating a 3D video. As the progress of computer vision technologies continues, many 3D video systems have been developed that tend to focus on automatic 3D reconstruction to capture target (foreground) objects. On the other hand, the 3D shapes of background regions such as walls and floors are generally made manually using 3D modeling software, i.e., CAD, or as a textured cube that just assumes the room. If we make background models using 3D modeling software, the quality is very high but it costs a lot of time and money. On the other hand, just making a cube is too simple to represent the outer space with sufficient appearance. To solve this problem, we propose a method to reconstruct the background model with simple operations but whose quality remains higher than cubes. In our method, a user needs only to segment multiple background images into several boxes by indicating their vertices as the corresponding points. From these data, the 3D coordinates of the vertices are calculated by using a stereo vision algorithm.

## 1 はじめに

近年、コンピュータビジョン分野における自由視点映像は応用段階に移りつつあり、いくつかのアプリケーションが提案されている [1, 2]。自由視点映像とは、視点の異なる複数のカメラからの映像（以下、多視点映像と呼ぶ）を用いて、3次元空間における任意の視点から見た2次元映像を自動生成する技術である。自由視点映像生成のアプローチに、2次元画像から3次元

形状モデルを再構築した後、レンダリングするものがある。このアプローチの代表例である視体積交差法 [3] では、多視点映像を被写体などの前景領域とそれ以外の背景領域とに分割し、前景領域のシルエット情報から前景の3次元形状を復元する。このような形状復元を行うアプローチでは、前景領域の3次元形状モデルは自動的に生成される。一方で、背景領域の形状モデルはCADなどのモデリングソフトを用いた手作業に

よって生成される場合が多い。しかし、モデリングソフトを用いる場合、詳細なモデルの作成には手間がかかり、また、単純な形状のモデルを使用すると、実際の撮影環境の背景領域の形状と大きく異なるため、映像が不自然になると考えられる。

本稿では、視体積交差法を用いた自由視点映像の背景領域の形状モデルを直方体より実際の形状に近く、かつ簡易に生成する手法を提案する。

## 2 関連研究

本提案手法では、多視点映像に写る点の3次元位置の推定結果から、背景の3次元形状を推定する。文献[4]では、多視点画像において自動的に対応点探索処理を行い、画像に写りこんだ領域の形状を推定する手法が提案されている。しかしこの対応点探索は、照明状況などにより対応点の見え方がカメラ間で大きく変化する場合や、カメラに死角が発生した場合は、対応点を少数しか検出することが出来ない。検出される対応点が少ない、または一部に偏って存在する場合は正しい形状として復元できない。これに対して本提案手法では、ユーザが各多視点画像中に写りこんだ背景領域の対応情報を入力する。これにより上述した原因による対応点の検出が困難な場合においても、人間の能力を活用して対応点を仮定することが出来るため、柔軟かつ頑健にモデルの生成が可能である。

画像中の背景領域に存在するオブジェクトは、主に壁や床などで構成されているため、複数の平面のモデルで表現できると仮定し、背景領域の指定は、複数の対応点を頂点とする平面群で指定する。文献[5]は、スタジオなどの大規模な空間における背景モデルを生成対象としている。この研究は大規模空間の背景領域は、視聴者の注目している前景物体より遠方に存在し、モデルの形状は詳細である必要がないため、背景モデルは複数枚の衝立状の直立する平面で表現可能と仮定し、単一画像上の地面と背景領域の境界線上の2点の座標値を用いて背景モデルの形状を算出している。この手法は、大規模空間を対象としているため上記の仮定が適用可能であるが、一般的な室内環境は、前景と背景とが比較的近接して存在し、また背景の形状が複雑になりがちなため、直立する平面のみで表現すると不自然な映像になる。本提案手法では、背景領域を複数の対応点を頂点とする平面群で指定するため、より柔軟にモデルの生成を行うことが可能である。

## 3 背景モデリング

### 3.1 背景モデリング手法

一般的に視体積交差法を用いる場合は、各カメラはキャリブレーション済みである。本提案手法では、ユーザが各多視点画像中に写りこんだ背景領域の対応情報

を入力し、前景領域の形状モデル生成の際に用いたキャリブレーション情報を基に、ステレオ視[6]を適用することで背景領域の3次元座標を算出する。

ユーザは、多視点画像から同一の壁や床などの背景領域中の平面が写り込んでいる画像2枚を選び、画面上で観測されるその平面の頂点を対応点として $n$ 個指定する。この対応点情報に以下に述べるステレオ処理を適用することにより、3次元座標を導出する。なお、屋内環境で撮影を行う場合、カメラから被写体までの距離が十分に確保できないため、画角が広く歪の大きなレンズを利用せざるを得ない、この影響を除去するために、各カメラ画像には歪み補正処理を適用する。補正後のカメラ画像上の対応点2点を $l(u_1, v_1, 1)$ ,  $l'(u_2, v_2, 1)$ とし、観測される3次元空間中の点を $w(x, y, z, 1)$ とする。このとき、各画像への射影変換は以下のように表される。

$$\lambda l = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} w$$

$$\lambda l' = \begin{pmatrix} p'_{11} & p'_{12} & p'_{13} & p'_{14} \\ p'_{21} & p'_{22} & p'_{23} & p'_{24} \\ p'_{31} & p'_{32} & p'_{33} & p'_{34} \end{pmatrix} w$$

ここで $\lambda$ はスケールファクター、 $p_{3 \times 4}$ ,  $p'_{3 \times 4}$ は各カメラの射影変換行列の各要素を表す。

上記2式では、3個の未知数に対し4本の式が存在するため、解を一意に算出する事ができない。そこで最小二乗法を用いて未知数である3次元座標 $w(x, y, z)$ を求める。具体的には、以下に示すように、行列 $B$ の擬似逆行列 $B^+$ を用いて上記2式を解く。

$$B = \begin{pmatrix} p_{31}u_1 - p_{11} & p_{32}u_1 - p_{12} & p_{33}u_1 - p_{13} \\ p_{31}v_1 - p_{21} & p_{32}v_1 - p_{22} & p_{33}v_1 - p_{23} \\ p'_{31}u_2 - p'_{11} & p'_{32}u_2 - p'_{12} & p'_{33}u_2 - p'_{13} \\ p'_{31}v_2 - p'_{21} & p'_{32}v_2 - p'_{22} & p'_{33}v_2 - p'_{23} \end{pmatrix}$$

$$w = B^+ \begin{pmatrix} p_{14} - p_{34}u_1 \\ p_{24} - p_{34}v_1 \\ p'_{14} - p'_{34}u_2 \\ p'_{24} - p'_{34}v_2 \end{pmatrix}$$

このようにして求めた $n$ 個の3次元点を頂点とする平面を3次元空間中に設置する。各平面には撮影した画像をテクスチャとしてマッピングする。多視点画像に写り込んでいる背景領域全てに対して、同様の処理を繰り返すことにより、複数の平面で構成された背景モデルが生成される。

### 3.2 平面間の隙間の除去

3.1節の手法に加え、キャリブレーションの誤差やユーザによる対応点の指定の誤差による平面モデル間

の隙間の除去を行う。視体積交差法では、全カメラから見えている領域をモデルの生成対象とするため、キャリブレーションは主に部屋の中心部で行う場合が多い。しかし本研究のモデルの生成対象である背景領域は、部屋の中心部から遠方に存在するため、キャリブレーションの誤差、特に外部パラメータの誤差が大きくなる。この誤差により、求めた3次元座標と実際の3次元座標とに差が生じ、本来ならば連続しているはずの平面間に隙間が生じる。また、ユーザによる対応点の指定の際の画素単位の誤差により隙間が生じる場合や、カメラの死角により一部の背景領域の指定が困難な場合も考えられる。隙間が存在すると背景モデルの見え方が不自然になるため、以下の隙間を埋める手法を2種類提案する(図1)。

**手法1：不整合をもつ頂点同士の平均をとる方法**

各平面の求めた3次元頂点座標が最も近い平面同士を、隣り合う平面と仮定し、その中点を頂点とする。その後、新しい頂点についてテクスチャ座標を再計算する。

**手法2：平面を隣の平面方向に延長する方法**

各平面の左右の頂点を通る直線の延長線上に平面を拡大し、平面同士を交差させる。この手法も同様にテクスチャ座標を再計算する。

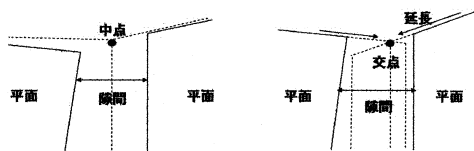


図1: 左: 隣り合う平面の頂点を中点に統合(手法1), 右: 平面を延長(手法2)

## 4 実験結果

本手法を実装したアプリケーションを用いて自由視点映像を撮影している部屋を対象に背景形状のモデリングを行った。撮影環境は自由視点映像生成用の環境であり[2]、図2に示すように $5m(D) \times 5m(W) \times 2.8m(H)$ の部屋にキャリブレーション済みのXGAの解像度をもつIEEE1394カメラ7台が撮影空間の外周を取り囲むように配置されている。ユーザは、各カメラ画像上で平面の頂点として指定したい任意の4点をマウスで指定し、その平面が多視点画像間で対応していることを指定するために、対応する平面に同じIDを付ける。図3は対応する平面を指定するインターフェースである。この入力インターフェースを用いて、計17枚分の平面を指定して形状のモデリングをおこなった結果を図4、図5左に示す。

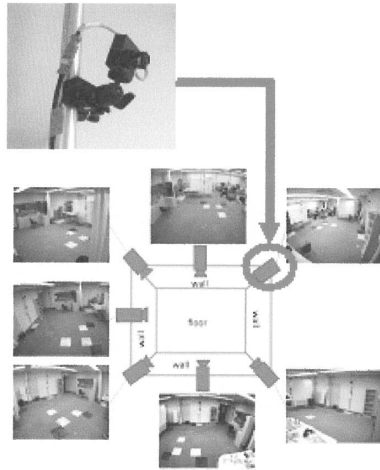


図2: 撮影環境：カメラ7台と多視点画像

図5左のモデリング結果を見ると、3.2節で述べた通り、求めた3次元座標には誤差が生じており、本来ならば連続しているはずの平面間に隙間が存在している。実際の3次元空間中の平面の位置と、ステレオ視で求めた位置の距離が最大で80cmのずれが見られた。この隙間を除去するため3.2節で提案した2種類の手法をそれぞれ適用した(図5中央, 右)。両手法とも平面間の隙間が埋まり、適用前と比べ背景モデルの見え方が比較的自然になっていることが分かる。しかし、手法1では3.1節の手法で求めた3次元座標が実際の3次元座標に近い場合、実際の3次元座標からのずれが逆に大きくなる。また、手法2では3.1節の手法を用いて求めた3次元座標と実際の3次元座標の誤差が大きい場合、ずれも拡大され、全体の形状がユーザの意図していた形状と著しく異なるものとなっている。

また、前景領域のモデルとの合成結果を図6に示す。

## 5 まとめ

本研究では、自由視点映像における背景領域の3次元形状モデルを、多視点画像上で背景領域を対応点を頂点とする平面群で指定することにより生成する手法を提案した。従来はCADなどのモデリングソフトを用いて手間をかけて行っていた背景モデルの生成を簡便に行うことが可能である。また、指定する平面を増やすことで、複雑な形状である実環境にも柔軟に対応し、頑健なモデルの生成が可能である。

また、外部パラメータの誤差による平面間の隙間の除去のため、不整合をもつ頂点同士の平均をとる手法と、平面を隣の平面方向に延長する手法の2つの手法

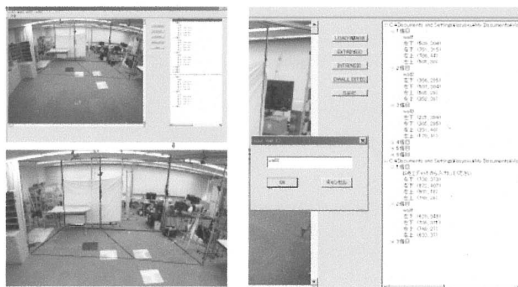


図 3: 左上: 平面指定画面, 左下: カメラ画像を表示し, 平面を指定する画面, 右: 指定された平面の各頂点座標を表示するリストボックス

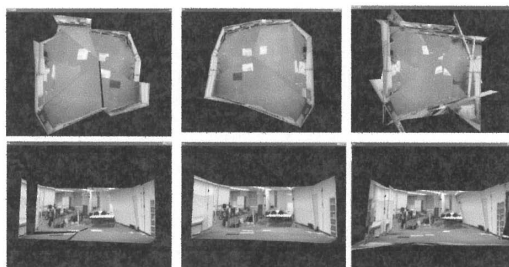


図 5: 上段: モデルを上から見た結果, 下段: 側面から見た結果, 左: 3.1 節のレンダリング結果, 中央: 3.2 節の手法 1, 右: 3.2 節の手法 2 を用いて平面間の隙間の除去を行った結果

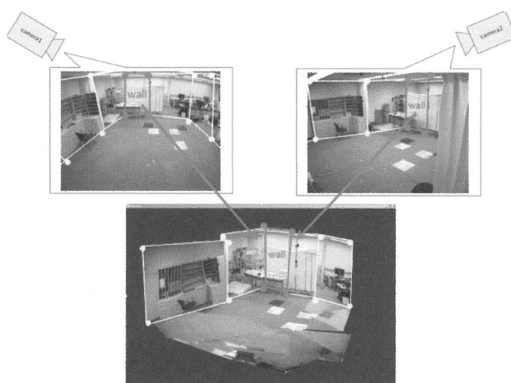


図 4: 上: 多視点画像上の対応する平面, 下: モデリング結果

を提案した。これらの手法を適用することで、実際の 3 次元座標とステレオ視により求めた 3 次元座標に誤差が生じる場合でも、背景モデルの見え方を比較的自然にすることが可能である。

## 謝辞

本研究は情報通信研究機構の委託研究により実施したものである。

## 参考文献

[1] K. Kimura and H. Saito.: *Video synthesis at tennis player viewpoint from multiple view videos*, In IEEE VR2005, pp. 281-282 (2005).



図 6: 前景モデルとの合成結果

[2] R. Sakamoto, I. Kitahara, M. Satomi, K. Tanaka and K. Kogure.: *Cinematized reality: cinematographic camera control in 3D videos*, In SIGGRAPH2005 Sketches, p. 103, ACM Press (2005).

[3] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan.: *Image-based visual hulls*, In SIGGRAPH00, pp. 369-374, ACM Press (2000).

[4] 菅谷保之, 金澤靖, 金谷健一: エピ極線幾何学による 2 画像間の密な点対応の生成, 情報処理学会研究報告, 2005-CVIM-148-19, pp. 145-152 (2005).

[5] N. Nomura, I. Kitahara, Y. Kameda, and Y. Ohta.: *A Background Modeling Method with Simple Operations for 3D Video*, In Proc. Of 3DTV Conference 2007 (2007).

[6] 徐剛, 辻三郎: 3 次元ビジョン, 共立出版 (1998).