

簡便な対話的操作による背景モデルの生成

An interactive room modeling system for 3D video

一色 彩 坂本 竜基 北原 格 キム ハンソン 鳥山 朋二 小暮 潔*

Summary. 本稿では、3次元形状を復元するアプローチを用いた自由視点映像撮影環境において、背景領域の形状モデルを対話的操作により生成する手法を提案する。一般に前景領域の形状モデルは、視体積交差法などの手法を用いて自動的に生成されるが、背景領域の形状モデルは手動で生成することが多い。提案する手法は、前景領域の3次元形状推定の際に用いられる多視点カメラのキャリブレーション情報と、カメラ画像内で指定された平面から背景領域の形状モデルを生成する。本手法によると、簡便な対話的操作によって、乱雑な背景になりがちな実環境においても柔軟かつ頑健に背景領域の形状モデルの生成が可能となる。

1 はじめに

コンピュータビジョン分野における自由視点映像生成技術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, 4]。複数のカメラ映像を前景領域とそれ以外の背景領域に分割し、前景領域のシルエット情報を計算機内部で統合することにより3次元形状を再構成する視体積交差法[2]は、ステレオ処理による形状処理が困難な、カメラ配置が疎な場合に有利であり、前景領域の自由視点映像を生成する際にしばしば用いられる。一方で、背景領域の形状モデルは、手作業によって生成されることが多い。本稿では、このような自由視点映像中の背景領域となる部屋等の形状モデルを、簡便な対話的操作により生成する手法を提案する。

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。これらのカメラ画像において自動的に対応点探索を行い、3次元形状を推定する手法[6]が提案されているが、多視点カメラが分散的に配置された環境では、照明条件などによる対応点の見え方や、カメラの配置による死角の発生状況がカメラ間で大きく異なるため、頑健に対応点を設定することは困難である。これに対して本手法では、ユーザが各多視点画像間で対応関係にある平面を指定することにより、3次元形状の生成を行う。本手法は、画像中に対応点の設定が困難な場合においても、人間の能力を活用して平面を仮定することができるため、柔軟なモデル生成が可能である。また、自由視点映像を生成する際の映像を背景領域の形状モデルのテクスチャとして用いるた

め、前景と背景でテクスチャの明るさや色合いなどの見た目が馴染みやすいことも利点である。

2 背景モデリング

文献[3]は、スタジアムなどの大規模な空間における背景モデルの生成手法を提案している。この手法では、遠方に存在する背景領域の形状を、ついたて状の直立する複数枚の平面で表現可能であると仮定し、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領域のモデル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、我々が撮影対象とする室内環境の背景領域は、形状がより複雑なため、ついたて状の平面群だけでは表現することが困難である。本手法では、複数の画像上に共通して写りこんだ平面群を指定することにより、背景領域の形状をより柔軟に表現することが可能である。

ユーザは、多視点画像から同一の壁や床領域が写り込んでいる画像2枚を選び、画像上で観測されるその領域中の対応点を 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[5]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。

この補正されたカメラ画像上の対応点の座標から、3次元空間中の座標を算出する。ここで、この補正された各カメラ画像上の点を $l(u_1, v_1, 1)^T$, $l'(u_2, v_2, 1)^T$ とし、上記2点で観測される3次元空間中の点を $w(x, y, z, 1)^T$ 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。

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

Copyright is held by the author(s).

* Aya Isshiki, ATR 知識科学研究所, 奈良女子大学 理学部 情報科学科 / Ryuki Sakamoto, Hangsung Kim, Tomoji Toriyama and Kiyoshi Kogure, ATR 知識科学研究所 / Itaru Kitahara, ATR 知識科学研究所, 筑波大学 大学院 システム情報工学研究科

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

ここで λ 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, 解を一意的に算出することができない. 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. この平面にカメラ画像をテクスチャとしてマッピングすることにより, 複数の平面で構成された背景領域の形状モデルが生成される. 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, 前景領域を取り囲む空間の 3 次元モデルを構築する.

3 実験

2 章の手法を実装したアプリケーションを用いて部屋モデリングを行った結果について述べる. 図 1 に示すように, 実験環境は, $5m(D) \times 5m(W) \times 2.8m(H)$ の部屋であり 7 台のカメラが環状に配置されている. 本環境は自由視点映像生成用の環境であり [4], 各カメラはキャリブレーション済みである. ユーザは, 各画像から平面の頂点として任意の 4 点をマウスで指定し, その平面が多視点画像間に対応していることを指定する. ステレオ視により各平面の頂点の 3 次元位置が算出され, ポリゴン群が描画される. その上に画像上で指定された平面領域をテクスチャとしてマッピングすることにより 3 次元モデルを生成し, 自由視点映像をレンダリングする. 床面も含めて 19 枚分の平面を指定して, 背景領域をモデリングした結果を図 2 に示す.

謝辞

本研究の一部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.

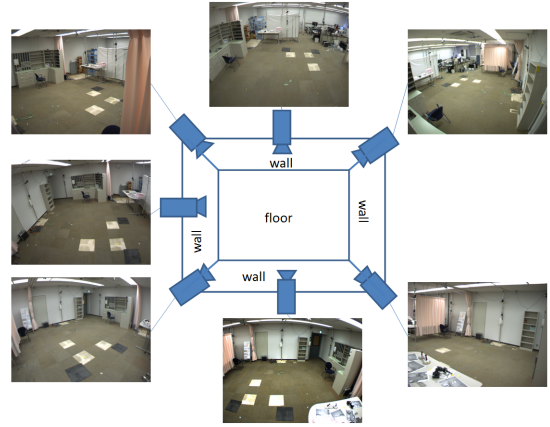


図 1. 撮影環境



図 2. モデリング結果

- [2] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan. Image-based visual hulls. In *SIGGRAPH00*, pp. 369–374. ACM Press, 2000.
- [3] 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.
- [4] 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.
- [5] 徐剛, 辻三郎. 3 次元ビジョン. 共立出版, 1998.
- [6] 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.