



撮影距離の異なる 2 台のカメラ間の視点移動表現法

志村直樹¹⁾, 亀田能成²⁾, 北原格²⁾

1) 筑波大学大学院 システム情報工学研究科

(〒305-8573 茨城県つくば市天王台 1-1-1, s1620791@u.tsukuba.ac.jp)

2) 筑波大学 計算科学研究センター

(〒305-8577 茨城県つくば市天王台 1-1-1, {kameda, kitahara}@iit.tsukuba.ac.jp)

概要: 撮影空間を取り囲むように配置した多視点カメラと撮影空間内部に設置した全方位カメラを併用することにより, 外側からでは観察できない内側からの見え方の観察を可能とする映像の撮影提示方式に関する研究に取り組んでいる. 内外カメラで撮影された見え方の大きく異なる 2 種類の映像を単純に切り替えた場合, 撮影距離の違いによって被写体の見え方が急激に変化するため, 視点移動を知覚しにくくなるのが懸念される. 本発表では, 撮影距離の異なる 2 台のカメラ間の視点遷移映像を生成することにより, 視点移動感を表現する手法を提案する.

キーワード: 多視点映像, 全方位カメラ, 視点遷移映像, Inside-Out/Outside-In 撮影

1. はじめに

複数のカメラで撮影した多視点映像は, 単一視点映像よりも多くの視覚情報を有する (より多面的な観測が可能である) ため, スポーツシーンの閲覧などの被写体の詳細な観察に用いられている [1]. 多視点映像を提示する手法の一つである **Bullet-Time** [2] は, 被写体の周囲を取り囲むように設置した複数のカメラを用いて撮影した多視点映像を, カメラを配置した順に切り替えることで, 視点の移動感を伴う被写体の観察を可能とする手法である. しかし, 被写体の周囲に設置したカメラで撮影 (**Outside-In** 撮影) した映像を切り替えるだけでは, 撮影空間内部に入り込んだ見え方を生成することは困難である.

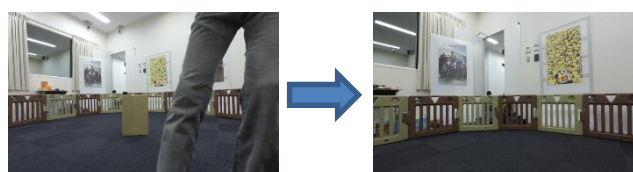
一方, 撮影空間内部に全方位カメラを設置することにより, 撮影空間を内部から観察可能なパノラマ映像を撮影 (**Inside-Out** 撮影) することができるが, 視点を移動させることは難しい.

我々は, **Outside-In** 撮影を行う多視点カメラと **Inside-Out** 撮影を行う全方位カメラを相補的に組み合わせることで双方の長所を兼ね備えた撮影方式の研究に取り組んでいる. 多視点カメラに加えて, 撮影空間内部に全方位カメラを設置することで, 被写体と近い距離からの映像も撮影することができ, 没入感のある映像生成を行うことができる. しかし, 図 1 に示すように, それらの映像を単に切り替えるだけでは, 多視点カメラと全方位カメラの位置関係の把握が困難であり, 映像切り替えの際, 視点移動を知覚しながらの被写体の連続的な観察が難しくなることが懸念される.

本研究では, 被写体までの距離が異なる, 2 種類のカメラ映像を切り替える際, それらの間を視点が疑似的に移動したような視点遷移映像を生成・内挿することによって, カメラの位置関係の把握を支援する手法を提案する. 閲覧者は切り替え前後のカメラの位置関係を把握しやすくなり, 切り替え前後の映像において被写体を連続的に観察することが可能となる.

2. 関連研究

あるカメラ映像から他のカメラ映像へ切り替える際に, 切り替え前後のカメラの設置位置や撮影方向が大きく異なると, 切り替えによって被写体や背景領域の見え方が急激に変化するため, カメラの位置関係の把握が困難となる問題が生じる. 津田ら[3]は, 視点が 2 台のカメラ間を仮想的に移動する CG 映像 (視点遷移映像) を合成することで, 閲覧者がカメラ間の位置関係を認識しつつ, 閲覧映像を切り替える手法を提案している. 視点遷移映像は, あらかじめシーンに存在する主な建物と地面のおおまかな 3 次元モデルを作成し, そこに切り替え前後のカ



Outside-In カメラ映像

Inside-Out カメラ映像

図 1 撮影距離の異なる 2 台のカメラで撮影した映像の切り替え

メラ映像を投影テクスチャマッピングすることで生成する。

本手法でも同様に、撮影空間の壁、床、物体のおおまかな3次元モデルを作成し、カメラの映像を投影テクスチャマッピングにより、モデルに投影することによって視点移動映像を生成する。

3. Inside-Out カメラと Outside-In カメラ間の視点遷移映像

3.1 3次元モデルの生成

本研究では、視点移動表現を行うために、撮影空間のおおまかな3次元モデルを用意する。3次元モデルを用いて視点移動表現を行うことで、運動視差の再現可能な映像の生成提示が可能となり、閲覧者がカメラの位置関係を理解しやすくなる。

図2に Outside-In カメラと Inside-Out カメラの配置を示す。Outside-In カメラは、撮影区間の中央付近の一点を向くようにほぼ同じ高さのパネルの所定位置に固定されている。それらのカメラを用いて撮影した多視点映像に Structure from Motion (SfM) を適用し、カメラの内・外部パラメータを推定する。これらのカメラパラメータを用いて、壁・床・物体のおおまかな3次元モデルを生成する。

床は平面形状とし、Outside-In カメラの高さから一定値だけ下方に設置する。壁の3次元形状は、Outside-In カメラの光軸が交わる点を中心とし床面に直立する円筒で表現する。パネルは、Outside-In カメラの位置姿勢に基づいて、あらかじめ採寸したサイズの直方体を配置する。

推定した3次元形状に投影テクスチャマッピングを用いて多視点映像を貼り付ける。投影テクスチャマッピングとは、図3のようにライトを照射するようにCG物体表面にテクスチャ情報を付与する手法である[4]。

3.2 仮想カメラの移動

Outside-In カメラと Inside-Out カメラの間を移動する仮想カメラの設定について述べる。SfM ベースのキャリブレーションから求めた撮影空間の3次元モデル上におけるカメラの位置から、移動する仮想カメラの位置姿勢を求める。仮想カメラの位置ベクトル \mathbf{P} は、Inside-Out カメラの位置ベクトル \mathbf{P}_{in} と、Outside-In カメラの位置ベクトル

\mathbf{P}_{out} を用いて式(1)のように表される。

$$\mathbf{P} = k\mathbf{P}_{in} + (1-k)\mathbf{P}_{out} \quad 0 \leq k \leq 1 \quad (1)$$

また、カメラの光軸の向きも同様に、仮想カメラの光軸の方向ベクトル \mathbf{R} は、Inside-Out カメラの光軸の方向ベクトル \mathbf{R}_{in} と、Outside-In カメラの光軸の方向ベクトル \mathbf{R}_{out} を用いて式(2)のように表される。

$$\mathbf{R} = k\mathbf{R}_{in} + (1-k)\mathbf{R}_{out} \quad 0 \leq k \leq 1 \quad (2)$$

式(2)のように仮想カメラの光軸の方向ベクトルを変化させることによって、連続的な視点移動映像を生成することができる。式(1), (2)において変数 k を変化させることによって仮想カメラは図4に示すような動きをする。

3.3 テクスチャのブレンディング

視点移動映像の見え方は、切り替え前後の Inside-Out カメラと Outside-In カメラから得られる2台のカメラ映像を仮想カメラの位置に応じた重みに応じて3次元モデルに投影することで生成する。Outside-In カメラ映像に対応する画素値を src_{out} 、Inside-Out カメラ映像に対応する画素値を src_{in} とすると、投影するテクスチャの画素値 dst は、式(3)のように表される。

$$\text{dst} = k\text{src}_{out} + (1-k)\text{src}_{in} \quad (3)$$

式(2)の範囲で変数 k を変化させることで、仮想カメラの

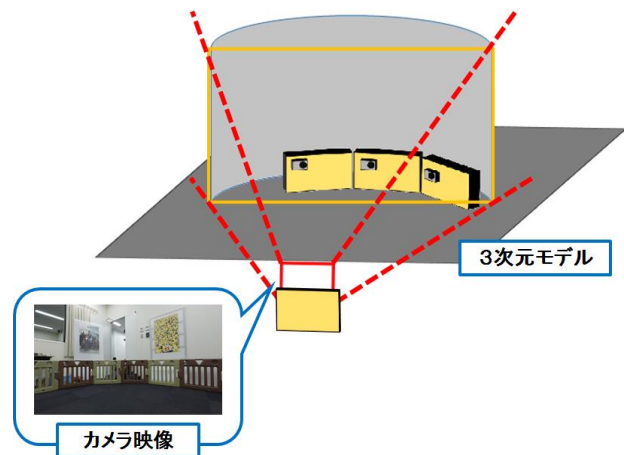


図3 投影テクスチャマッピング

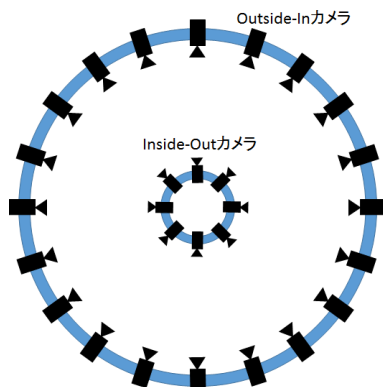


図2 カメラ配置の一例

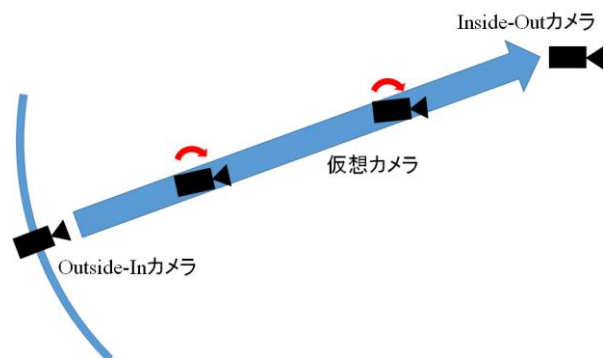


図4 仮想カメラの動き

移動と同時に、投影映像のテクスチャのブレンディングを行う。

カメラのキャリブレーション誤差や被写体形状の近似誤差により、Inside-OutカメラとOutside-Inカメラから投影されるテクスチャにズレが生じるが、二つのテクスチャのブレンディングを行うことにより、この違和感を軽減しつつ、自然な見え方の切り替えが実現される。

4. 実証実験

4.1 実験環境

撮影空間は図5の710cm×760cmの部屋で行った。本実験では、CPU：Intel Corei7 3.40GHz、GPU：NVIDIA GeForce GTX 550 Ti、メモリ：8.00GB RAMを搭載したデスクトップPCを用いて、多視点映像のキャリブレーション、3次元モデル生成、視点遷移映像生成の処理を行った。Inside-OutカメラとOutside-InカメラにはGoPro社のHERO4ブラックエディションを計19台使用した。今回の撮影では4Kでの撮影を行い、4Kの撮影時の水平画角は122.6度で垂直画角は94.4度、焦点距離は14mm、画像の解像度は4Kの3,840画素×2,160画素、撮影時のフレームレートは30fpsである。SfMカメラキャリブレーションにはVisualSfM[5]を用いた。本実験では、図6のようにカメラを設置した。カメラを取り付けるパネルは、幅71cm×高さ56cm×厚み5cmのものを12台用いた。そこに円形

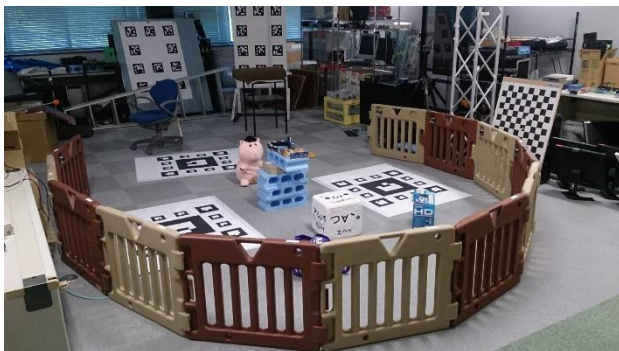


図5 撮影の様子

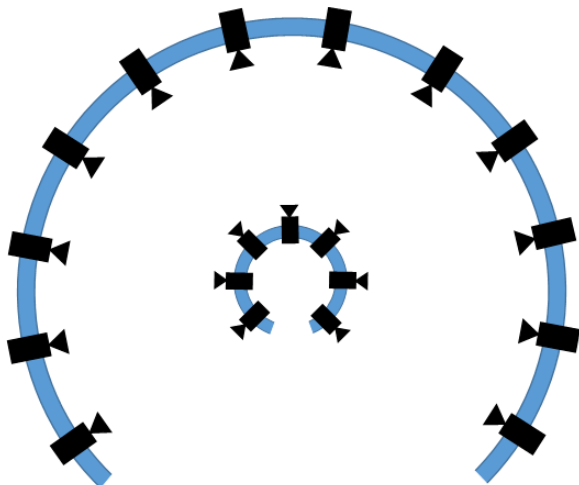


図6 カメラの配置図

にパネルを12台設置し、12台のOutside-Inカメラを取り付けた。また、Inside-Outカメラは7台用い、円の中心にOutside-Inカメラの取り付けられた高さと同様な高さになるよう、台の上に円形に載せ、撮影を行った。

4.2 多視点映像のキャリブレーション

本実験では、SfMベースのキャリブレーションを行うため、特徴点を多くとることで、精度の高いキャリブレーションを行うことができる。そこで、キャリブレーション用のシーンを撮影するため、壁や床にマーカや物体を設置し、この撮影映像を用いてキャリブレーションを行った。

図7は、Outside-InカメラとInside-Outカメラで撮影された計19枚の画像をVisualSfMによってキャリブレーションした結果画像である。推定されたカメラの位置姿勢はおよそ正しく推定されているとわかる。

4.3 3次元モデル生成

図8は、3章の手順で生成した3次元モデルである。撮影映像のテクスチャは、ほぼ正しい位置に投影されていることがわかる。3次元モデルに投影されたテクスチャが少しずれているように見えるのは、キャリブレーションの際

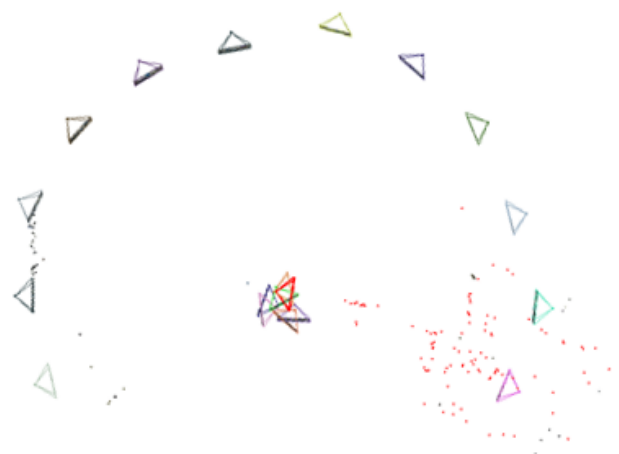


図7 キャリブレーション結果

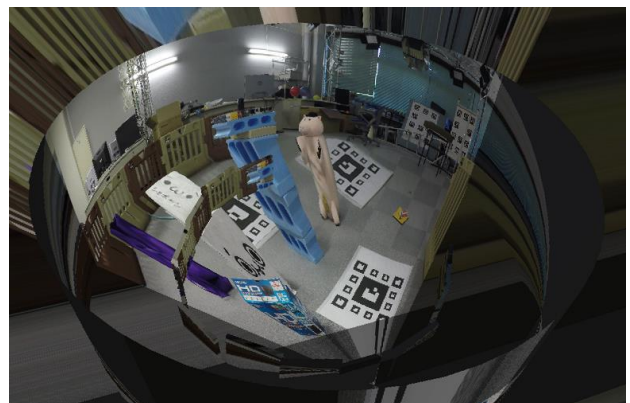


図8 3次元モデルの俯瞰図

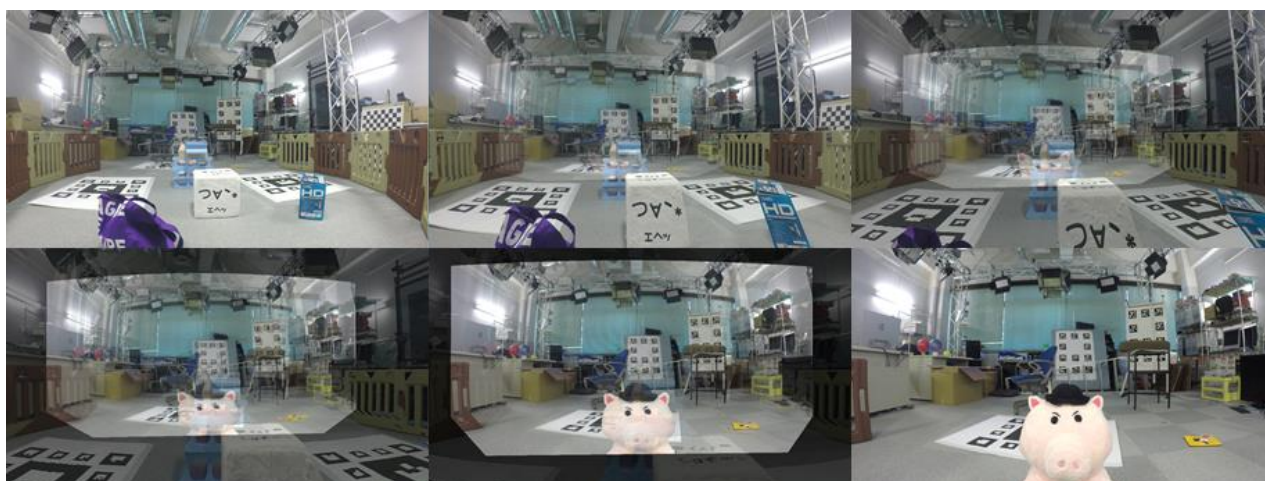


図 9 Outside-In カメラからの視点切り替え

の推定誤差が原因だと考えられる。

4.4 視点遷移映像生成

図 9 は提案手法による実行結果である。移動前後の撮影映像のテクスチャは、ほぼ正しい位置にマッピングされていることがわかる。しかし、撮影空間内に 3 次元モデル生成を行っていない物体が存在する場合、視点移動が行われた際に、テクスチャが大きく歪んでしまうことがある。これは、3 次元モデルの生成が行われていない物体のテクスチャが、正しい位置に投影されていないためである。しかし、結果映像から、壁やパネルの位置の移動で、本研究の目的である映像切り替え前後のカメラ間の位置関係の把握には十分であると考えられる。また、視点移動中に異なるカメラ映像のテクスチャが同時に投影されている際、同一の物体のテクスチャがずれていることがある。この原因として、二つの問題が挙げられる。一つ目は、壁の 3 次元モデルを円筒で表現しているため、3 次元モデル上の正しい位置にテクスチャが貼りついていないためである。二つ目にキャリブレーションの際の推定誤差のために、カメラの位置や姿勢が実際のものとは異なることがあるためである。これらの問題は、より正確な 3 次元モデル生成、及び SfM キャリブレーションの精度向上によって解決可能であると考えている。

5. おわりに

本研究では、Outside-In カメラと Inside-Out カメラ間での映像切り替わり時に、カメラ間を疑似的に視点移動する映像を生成・挿入することによって、視点移動の把握を支援する映像提示手法を提案した。カメラキャリブレーション処理によって得られた 3 次元点群情報から撮影シーンのおおまかな 3 次元形状を復元し、それに投影テ

クスチャマッピングを適用することにより撮影空間の 3 次元モデルを生成した。この 3 次元空間モデルを用いて視点移動映像を生成した。今後、Inside-Out カメラに全方位カメラを用いた撮影実験に取り組む予定である。Inside-Out カメラを全方位カメラに変更し、撮影を行ったシーケンスでも同様の映像切り替えを行えるシステムの実装を検討する。

本研究は、科学技術振興機構(JST)の戦略的創造研究推進事業(CREST)における研究領域「人間と調和した創造的協働を実現する知的情報処理システムの構築」の研究課題「ソーシャル・イメージング：創造的活動促進と社会性形成支援」により行ったものである。

参考文献

- [1] Itaru Kitahara, Yuichi Ohta, “Scalable 3D Representation for 3D Video Display in a Large-scale Space,” Proc. of IEEE Virtual Reality 2003 Conference (VR2003), pp.45-52, (2003)
- [2] 富山仁博, 岩館祐一, “多視点ハイビジョン映像生成システムの開発,” 映像情報メディア学会誌, Vol.64, No4, pp.622-628, (2010)
- [3] T. Tsuda, I. Kitahara, Y. Kameda, Y. Ohta, “Smooth Video Hopping for Surveillance Cameras,” The 33rd International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH2006) Sketches Article No. 129, (2006)
- [4] C.Everitt, “Projective Texture Mapping,” White paper, Nvidia Corporation, (2011)
- [5] C. Wu, “VisualSFM : A Visual Structure from Motion System,” <http://ccwu.me/vsfm>, access:2016/7/11