撮影方向が異なる多視点映像の閲覧方式

志村 直樹 [1,a)]　亀田 能成 [2,b)]　北原 格 [2,c)]

1. はじめに

複数台のカメラで撮影した多視点映像は，様々な方向からの見え方を記録しているため，３次元物体の詳細な観察に適している[1]． Bullet-Time[1][2]は，被写体の周囲を取り囲むように設置したカメラで撮影した多視点映像を，カメラ配置に従って切り替えながら提示する映像提示手法である．視点移動感の再現が可能であることから，映画やプロモーションビデオなどでの利用が進んでいる[3][4]．上述したように多面的な観察に優れているが，被写体の周囲に設置したカメラで撮影（Outside-In 撮影）した映像の閲覧では，撮影空間内部に没入したような見え方の再現は困難である．一方，撮影空間内部に設置した全方位カメラで撮影（Inside-Out 撮影）すれば，没入視点からの見え方を観察することが可能となる．また，被写体を至近距離から撮影するため，被写体の細部まで観察することができる．しかし，単一視点で撮影するため，視点移動を伴う多面的な観測が困難である．

本稿では，図１に示すように撮影空間の周囲と内部に配置したカメラで撮影した映像を適切に切り替えながら提示することにより，Outside-In／Inside-Out 映像の長所を兼ね備えた映像提示方式を提案する．

2. 関連研究

暦本ら[5]は，テレプレゼンスシステムにおいて，特定の一人称視点映像のみを利用するだけではなく，別のカメラで撮影した三人称視点からの映像も使用する JackIn Space を提案している．従来の一人称視点のみを利用するテレプレゼンスシステムでは，遠隔地の環境把握が困難であるといった問題が存在するが，三人称視点映像も利用することで，遠隔地の現場周囲を見渡すことができる．また別の視点からの映像に切り替える際，空間的な位置把握を支援するため，視点遷移映像を生成している．この視点遷移映像は，作業者に取り付けた広角カメラと，遠隔環境に取り付けた複数のデプスセンサを用いることで生成している．

本研究でも，被写体の周囲に設置したカメラで Outside-In 撮影した映像で撮影空間の全体像を把握するのに加え，撮影空間内部に設置した全方位カメラで Inside-Out 撮影した映像を用いることで，詳細な被写体の観察を可能とする．一方で，Outside-In 撮影と Inside-Out 撮影では，各カメラの撮影方向が大きく異なるため，それらの映像を単純に切り替える場合，図２に示すように，撮影空間の相対的な位置関係が把握しにくくなることが懸念される．撮影対象空間の３次元形状が既知であれば，自由視点映像技術[6][7]を用いることで Outside-In 撮影と Inside-Out 撮影の中間の見え方を再現することが可能であるが，動的なシーンの３次元情報を実時間で復元するためには膨大な計算コストが必要となる．Tsuda らは，あらかじめ生成しておいた撮影空間の静的な物体の３次元モデルを用いて，異なる地点で撮影した多視点映像の間を滑らかに遷移する映像生成提示手法を提案している[8]. 本方式でも壁や床面などの静止領域の３次元形状を用いて視点遷移映像を生成することにより，Outside-In 撮影映像と Inside-Out 撮影映像の見え方を滑らかに補間し，撮影空間の位置関係の把握を支援する．

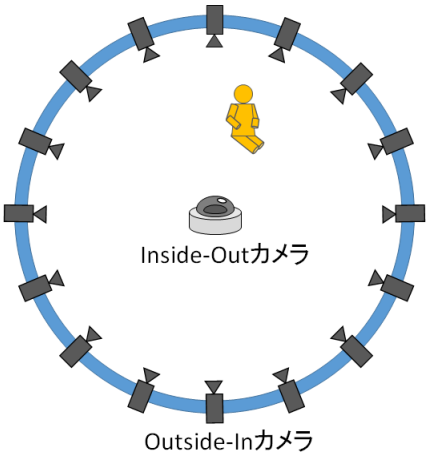

図 1. カメラ配置の一例

図２．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

[1] 筑波大学　大学院システム情報工学研究科
[2] 筑波大学　計算科学研究センター
[a)] s1620791@u.tsukuba.ac.jp
[b)] kameda@iit.tsukuba.ac.jp
[c)] kitahara@iit.tsukuba.ac.jp

図 3. Outside-In カメラの注視点設定処理（上：元画像，下：処理後画像）

3. Outside-In 撮影映像の処理

Outside-In 映像は Bullet-Time 処理を用いて閲覧する．被写体を取り囲むように設置した Outside-In カメラで撮影した映像を，カメラ配置に応じて切り替えることで，被写体の周囲を移動しているような視覚効果を再現する．

各カメラの光軸が一点で交わっていない多視点映像を切り替えながら提示すると，画面上での注目物体の観測位置（注目点）が変動するため，滑らかな視点移動知覚を与えることが難しくなる．本研究では，注目点の３次元位置に基づいて提示映像を撮影する仮想カメラの光軸方向を再設定し（撮影映像に射影変換を施し），注目点が画面上の同一位置で観測されるように Outside-In 映像の見え方を変換する．

変換処理に必要となる多視点カメラの内部・外部パラメータを推定するために，多視点映像に Structure from Motion（SfM）を適用する[9][10]．推定した Outside-In カメラの位置ベクトル$\boldsymbol{M_{out}}$と注目点の位置ベクトル$\boldsymbol{X}$から仮想カメラの光軸の方向ベクトル$\boldsymbol{D}$は式(1)のように表せる．

$$\boldsymbol{D'} = \boldsymbol{X} - \boldsymbol{M_{out}} \tag{1}$$

$$\boldsymbol{D} = \frac{\boldsymbol{D'}}{|\boldsymbol{D'}|} \tag{2}$$

別の Outside-In カメラ映像に切り替える際に，式(1)(2)の再計算を行うことで，図のように常に仮想カメラの光軸と注目点が一致させる．また，切り替え前後で映像の連続性を保つために，注目物体の見かけの大きさを一定にする必要がある．そこで本システムでは，切り替え前後で仮想カメラの画角の再設定を行う．切り替え前後の仮想カメラの画角をα_n，切り替え前の仮想カメラと注目点の距離をd_n，切り替え後の仮想カメラと注目点の注目点の距離をd_{n+1}とすると，切り替え後の仮想カメラの画角をα_{n+1}は，式(3)のように表せる．

$$\alpha_{n+1} = \alpha_n \frac{d_n}{d_{n+1}} \tag{3}$$

これらの処理を施した結果画像を図 3 に示す．

4. Inside-Out 撮影映像の処理

Inside-Out 映像はパノラマ映像として閲覧する． 360 度カメラを用いれば 1 台で全方位を撮影することが可能であるが，空間分解能は低下してしまう．本システムでは，被写体を詳細に観察するために，図 4 に示すリグに固定した８台のカメラを用いて Inside-Out 映像を撮影する．Outside-In カメラの取付け位置とほぼ同じ高さになるように，撮影空間内部にリグを固定し空間を内側から見た 映像を撮影する．８枚の Inside-Out カメラ映像にスティッチング処理を施し１枚のパノラマ映像を生成する．本方式では Kolor 社の Autopano Giga を用いる[11]．図 5 に合成したパノラマ映像の一例を示す．

5. 視点移動処理

Outside-In カメラ映像から Inside-Out カメラ映像に映像を切り替える際，それらの間を疑似的に視点が移動しているような視点遷移映像の生成処理について述べる．

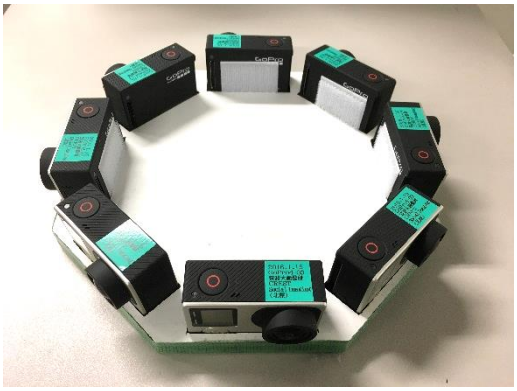

図 4. Inside-Out 撮影で用いたカメラリグ

図 5．パノラマ映像の一例

5.1　３次元モデル生成

視点遷移映像生成には静止領域の３次元モデルが必要である．図 6 に示す撮影空間を例に３次元モデルの生成処理について説明する．この場合，Outside-In カメラは，パネルの所定位置に固定されている SfM で推定したカメラパラメータと撮影空間の概形（疎らな 3D Point Cloud）から，壁・床・パネルの３次元モデルを生成する．床は平面形状とし，Outside-In カメラの高さから規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とする．パネルは，Outside-In カメラの位置姿勢に基づいて規定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 [12].

5.2　視点遷移映像生成

Outside-In カメラと Inside-Out カメラの間で仮想カメラを移動させることにより視点遷移映像を生成する．仮想カメラの位置ベクトル$\boldsymbol{M}$は，Inside-Out カメラの位置ベクトル$\boldsymbol{M_{in}}$と，Outside-In カメラの位置ベクトル$\boldsymbol{M_{out}}$を用いて式(1)で表される．

$$\boldsymbol{M} = k\boldsymbol{M_{in}} + (1-k)\boldsymbol{M_{out}} \qquad 0 \le k \le 1 \tag{4}$$

仮想カメラの光軸の方向ベクトル$\boldsymbol{D}$は，注目物体の位置ベクトル$\boldsymbol{X}$と仮想カメラの位置ベクトル$\boldsymbol{M}$を用いて，式(3)で表される．

$$\boldsymbol{D'} = \boldsymbol{X} - \boldsymbol{M} \tag{5}$$

$$\boldsymbol{D} = \frac{\boldsymbol{D'}}{|\boldsymbol{D'}|} \tag{6}$$

式(3)(4)によって仮想カメラの光軸の方向ベクトル$\boldsymbol{D}$を常に注目物体に向けることで，図 7 に示すように注目物体

図 6．撮影空間

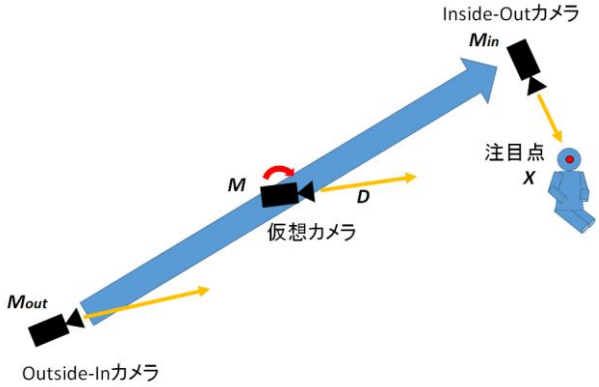

図 7．仮想カメラの動き

図 8．Outside-In カメラから Inside-Out カメラへの視点遷移映像

を捉えつつ Outside-In カメラ映像から Inside-Out カメラ映像に視点を切り替える．処理結果の一例を図 8 に示す．

5.3 適切なカメラの選択処理

Outside-In カメラ映像から Inside-Out カメラ映像に切り替える際には，常に注目物体を捉えつつ視点を切り替えるため，注目物体が映像中に写っている Inside-Out カメラ映像に切り替える．まず，Inide-Out カメラの重心位置から注目物体への方向ベクトル$\boldsymbol{D_X}$を求める．ベクトル$\boldsymbol{D_X}$は Inside-Out カメラ群の重心の位置ベクトル$\boldsymbol{M_C}$を用いて，式(5)のように表される．

$$\boldsymbol{D_X} = \boldsymbol{X} - \boldsymbol{M_C} \tag{7}$$

求めた$\boldsymbol{D_X}$と全ての Inside-Out カメラの光軸の方向ベクトルとの内積を求め，最となる Inside-Out カメラを選択する．

6. 撮影・映像閲覧システム

撮影・映像閲覧システムの実行環境について述べる．撮影空間は 710[cm]×760[cm]の部屋で行った．使用 PC 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 を用いた．Outside-In カメラ・Inside-Out カメラには，無線同期撮影が可能な GoPro 社の GoPro HERO4 ブラック edition をそれぞれ 20 台(Outside-In)・8 台(Inside-Out)を使用し，解像度は 4K の 3840[pixel]×2160[pixel]，フレームレートは 30[fps]，水平画角は 122.6°，垂直画角は 94.4°で撮影を行った．Outside-In カメラは撮影シーンを取り囲むよう円形に配置し，幼児用ゲージのパネルの所定位置に固定する．Inside-Out カメラは撮影空間内部に Outside-In カメラと同じ高さとなるようリグに固定した．SfM カメラキャリブレーションには VisualSFM[13]を用いた．処理に要した時間は，カメラキャリブレーションに 82 秒，掲示処理の時間は初期設定時に 8 秒ほどかかるが，それ以降の閲覧開始後の処理は，ビデオレートで実行可能である．

7. まとめ

本研究では，Outside-In カメラ映像に加え，Inside-Out カメラ映像を用いることにより，双方の長所を兼ね備えた映像閲覧法について述べた．従来の Outside-In カメラ映像のみでは観察できないような，撮影空間内部からの映像を撮影することで，より詳細な観察を可能とした．今後は，本手法の有効性を確認するため，実証実験等を行いたいと考えている．

本研究は，JSTCREST および、科研費（17H01772）の助成を受けたものである。

References

[1] 富山　仁博，岩舘　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, No. 4, P 622-628, 2010

[2] N. Akechi, I. Kitahara, R. Sakamoto, Y. Ohta, “Multi-resolution bullet-time effect,” ACMSIGGRAPH-Asia, December.2014.

[3] http://timeslicefilms.com/

[4] http://tvtropes.org/pmwiki/pmwiki.php/Main/BulletTime

[5] R. Komiyama, T. Miyaki, J. Rekimoto, “JackIn space: designing a seamless transition between first and third person view for effective telepresence collaborations” Proceedings of the 8th Augmented Human International Conference. Article No. 14, 2017.

[6] T. Kanade, P. Rander, P. J. Narayanan, “Virtualized Reality: Constructing Virtual Worlds from Real Scene,” IEEE MultiMedia, Vol.4, No.1, p34-47, 1997.

[7] L.McMillan and G.Bishop, “Plenoptic Modeling: An Image-Based Rendering System ,” Proc. SIGGRAPH 1995, pp.39-46, 1995

[8] 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

[9] C. Wu, “Towards Linear-time Incremental Structure from Motion,” 3DV, 2013

[10] C. Wu, S Agarwal, B Curless, S Seitz, “Multicore Bundle Adjustment,” CVPR, 2011

[11] AutoPano, Alexandre Jenny: http://www.kolor.com/autopano/

[12] M. Segal, C. Korobkin, R. van. Widenfelt, J. Foran, P. Haeberli, “Fast Shadow and Lighting Effects Using Texture Mapping,” Proceedings of SIGGRAPH’92, pp.249-252, 1992

[13] VisualSFM , C. Wu: http://ccwu.me/vsfm/