

ビデオテクスチャと投影マッピングによる自由視点映像生成

木村 文哉[†] 宍戸 英彦[‡] 北原 格[‡]

筑波大学理工学群工学システム学類[†] 筑波大学計算科学研究センター[‡]

1.はじめに

撮影機器やネットワークの発展に伴い、多視点映像を計算機内部で統合し、任意視点からの見え方を再現する自由視点映像に注目が集まっている。自由視点映像では被写体は静止した状態で観察されるため、生命感を表現することが難しいという課題が存在する。本研究では、生命感を感ぜさせる要因の一つである“動き”に着目し、その再現可能な自由視点映像生成方式の実現について述べる。被写体の動きは、骨格変動による体全体の動きと、表情変化など体表面の微細な動きに分類される。体全体の動きについては、モーションキャプチャで計測した被写体の動作に合わせて3次元モデルの姿勢を変化させることで再現するアプローチが開発されている。本研究では、骨格変動のような大きな動きではなく、テクスチャ変動により表現される体表面の微細な動きによって被写体に生命感を付与することを目的とする。

複数台のカメラで被写体（人物）の全身を撮影した多視点映像に対して Structure-from-Motion (SfM) [1]を適用し3次元形状を復元する。多視点カメラの中から被写体の生命感表現に影響を与えやすいメインカメラを選択し、その映像を静的な3次元モデルに投影マッピング[2]することで、体表面の微細な動きを表現する。投影するテクスチャは Video Textures[3]を用いて生成する。メインカメラの映像フレーム間の類似度に基づいてフレーム遷移を制御することにより、短い映像から長時間の動きの表現を実現する。

2.関連研究

Video Textures[3]は、短時間の映像を映像品質を維持しつつランダムフレームにジャンプさせることで、単調な繰り返しではない長時間映像を合成する手法である。Video Textures を3次元モデルに投影することにより、体表面の微細な動きにランダム性を盛り込み、生命感の表現力向上を図る。

“Free-viewpoint video generation using video textures and projective mapping”

Fumiya Kimura[†], Hidehiko Shishido[‡], Itaru Kitahara[‡]

Engineering Systems, University of Tsukuba[†],

Center for Computational Sciences, University of Tsukuba[‡]

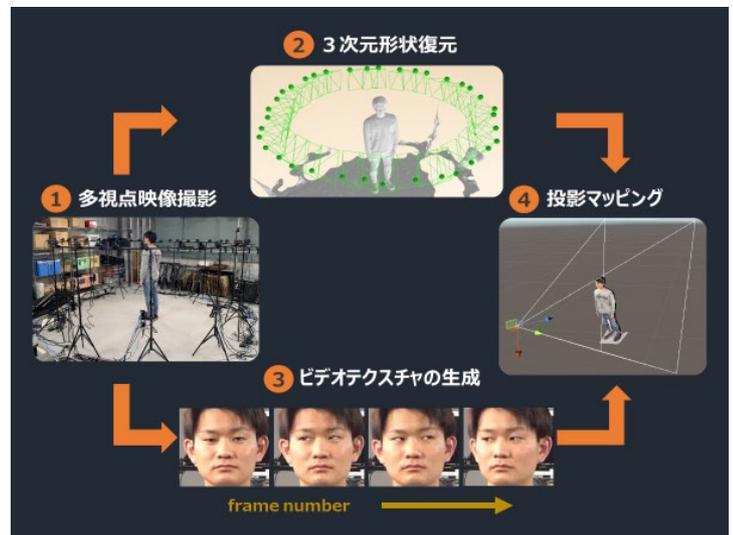


図1：提案する自由視点映像生成の流れ

3.ビデオテクスチャと投影マッピングによる自由視点映像生成法

本節では、多視点撮影から撮影対象の3次元形状復元、ビデオテクスチャの生成、投影マッピングを用いた自由視点映像生成法について述べる。

3.1 多視点映像撮影

図1①に多視点映像撮影の様子を示す。3次元形状復元およびビデオテクスチャ生成に用いる映像を36台のカメラ（ソニーRX0）で同期撮影する。カメラの高さは被写体の視線と同程度とし、撮影対象を囲むように円状に配置する。本研究における被写体の動きは、表情や顔の向き変化など小さなものに限定し、骨格変動を伴う大きな動きは対象外とする。

3.2 3次元形状復元

図1②に3次元形状復元の様子を示す。多視点映像から同時刻で撮影したフレームを切り出し、SfMによってフレーム間の特徴点の対応関係から被写体の3次元形状復元を行うと同時に、多視点カメラパラメータを推定する。多視点カメラの外部パラメータ（位置・姿勢）は、投影マッピングに使用する仮想プロジェクタの位置・姿勢の情報として利用する。

3.3 ビデオテクスチャの生成

多視点カメラの中から被写体の生命感表現に影響を与えやすいメインカメラを選択する。一般的には、被写体を正面から撮影したカメラが選ばれやすい。図2に示すように、カスケード型分類器を用いてメインカメラ映像から検出した顔領域について、フレーム間の類似度を画素値の差分二乗和によって算出し、遷移フレームペアを決定する。遷移フレームペア間で時系列を遡るフレームにジャンプすることをバックワードジャンプと言う。一定確率（本研究では30%）でバックワードジャンプを実行させることにより、見え方変化にランダム性を与えつつ長時間の映像生成を実現する。



図2 遷移フレームの決定

3.4 投影マッピング

VR環境構築プラットフォーム（Unity）を用いた実装例について紹介する。図3に示すように、構築した仮想空間に、被写体の3次元モデルと仮想プロジェクタ8台を配置する。仮想プロジェクタの位置姿勢は SfM で推定した外部パラメータを参照する。提示映像撮影用カメラからの距離が最も小さいプロジェクタを用いて、対応する多視点画像をモデル上に投影する。図3は、プロジェクタ1を用いた投影の結果である。

同様の処理を撮影映像のフレーム毎に繰り返すことで“動き”を提示する。その際、フレーム切替え速度を撮影映像のフレームレートと一致させる。



図3 カメラと投影プロジェクタ

4. 自由視点映像生成の結果

提案手法によってレンダリングした自由視点映像の一例を図4に示す。図4上段は図3に示すプロジェクタ1を用いた投影結果、下段は図3に示すプロジェクタ8を用いた投影結果である。ビデオテクスチャを投影マッピングすることにより表情変化を再現することができており、生命感の表現が可能となっている。

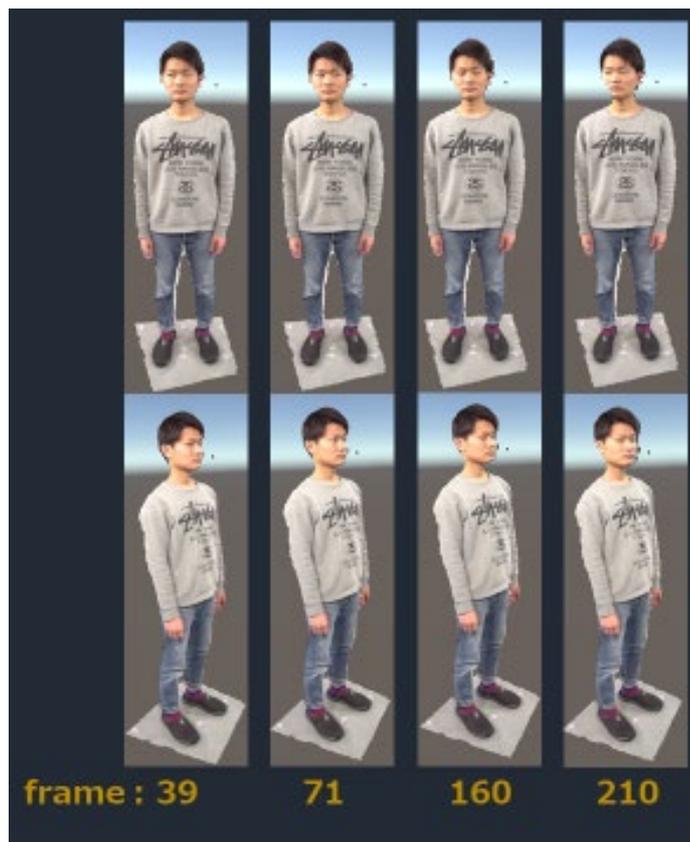


図4 本手法で提示する自由視点映像の一例

5. おわりに

静的な人物3次元モデルに対してビデオテクスチャを投影マッピングすることにより、微細な動きの表現が可能な自由視点映像生成法を提案した。本研究は科研費（17H01772）の助成によるものである。

参考文献

- [1] Johannes L. Schonberger, and Jan-Michael Frahm: "Structure-from-motion revisited", CVPR (2016)
- [2] C. Everitt: "Projective Texture Mapping, White paper", NVidia Corporation (2001)
- [3] Arno Schodl, Richard Szeliski, David H. Salesin, and Irfan Essa: "Video Textures", SIGGRAPH (2000)