

単一の RGB-D カメラで撮影した人物と死角領域の点群データ補間 Interpolationing RGB Pointcloud Outside Sensor Area

鈴木 幹大¹⁾ 江藤 謙¹⁾ 小篠 裕子¹⁾
Mikihiro Suzuki Ken Eto Yuko Ozasa

1 はじめに

RGBD カメラは、単一の方向から撮影すると、図 1 に示すように撮影対象の背面や、背後に位置する壁や床のデータが欠損する。

欠損したデータを補間する研究は数多く提案されているが、RGB 情報もしくは点群情報のみを補完する研究がほとんどであり、RGB 情報付きの点群 (以降 RGB 点群) を補完する研究は未だ少ない。

本研究では、人物を単一方向から撮影した RGBD 画像から生成した RGB 点群において、欠損した人物背面や背後の RGB 点群を補完する。具体的には、まず、RGBD 画像から RGB 点群を生成する。次に、RGB 画像のみから人物の三次元モデルを推定し、RGB 点群化する。RGB 点群から人物部分を取り除き、得られた背面の点群の欠損を補間する。欠損を補間した背面の点群と、RGB 画像のみから生成した人物の RGB 点群を合成する。

検証実験より、提案方法によって人物の姿勢や背面が複雑でない場合は、ある程度自然に欠損を補間した有用性の高い RGB 点群が入手できることがわかった。

2 提案方法

図 2 に提案方法の概要を示す。提案手法では、人物を撮影した RGBD 画像から RGB 点群を生成し、人物部分の点群を取り除いて、背景の点群のみを残す。背景の点群の欠損部分を補完する。RGB 画像から人物の三次元形状復元をし、得られた三次元人物モデルと、欠損を補完した背景を合成する。

人物の三次元形状の推定では、単一の RGB 画像から人物の 3D モデルを推定する手法の 1 つである PIFu[2] を用いる。単一の RGB 画像から人物の 3D モデルを推定する手法は多くあるが [4, 3, 6]、そのほとんどが深度情報のみの推定であり、RGB 情報を推定していない。よって、本研究では、人物の RGB 情報も推定している PIFu を採用する。

本研究で PIFu を用いる際にはいくつかの前処理が必要である。まず、PIFu による推定には別途復元対象のマスク画像が必要なため、U2-Net[1] によりマスク画像を生成する必要がある。さらに、PIFu で推定した RGB 点群は、RGBD カメラにより取得した RGB 点群とスケールが異なるため、RGBD カメラにより取得した RGB 点群を元に、PIFu で推定した RGB 点群をスケール調整する必要がある。

背景の補間には、機械学習を用いたインペインティング手法の 1 つである DeepFillv2[5] を用いる。DeepFillv2 では、まず RGBD カメラより取得した RGB 画像と深度画像をそれぞれネットワークに入力することで、対象物体の死角となり欠損した背景の RGB 点群を補間する。次に、出力した RGB 画像と深度画像から RGB 点群を生



図 1 欠損した点群データの例

成する。DeepFillv2 は、PIFu と同様に、対象物体のマスク画像が必要である。RGBD カメラの計測時における、RGB 画像と深度画像の空間方向の誤差を考慮し、PIFu の場合よりもマスク領域に余裕を持ってマスク画像を出力する。

生成した人物の 3D モデルと、欠損を補間した背景の RGB 点群を合成する際は、既にスケール済みである人物の RGB 点群の重心を元に、位置合わせする。ここで、重心による位置合わせは、PIFu により推定した RGB 点群と DeepFillv2 により補間した RGB 点群の座標系が同じ場合のみ有効であることに注意する。本稿では、PIFu により推定した人物の RGB 点群の重心座標を、RGBD カメラにより取得した元データである RGB 点群の重心座標に合わせることで位置合わせする。RGBD カメラにより取得した RGB 点群と、DeepFillv2 により補間した背景の点群の座標系は同一であるため、点群をそのまま足し合わせることで点群は合成できる。

3 検証実験

RGBD カメラで撮影したデータを用いて提案手法の有効性を検証する。本稿では、RGBD センサである RealSense D455 を用いて、協力者を 1 名毎に単一の方向から撮影した。協力者には立位の姿勢をとってもらい、センサは協力者正面に設置した。

図 2 に提案方法を用いて欠損を補完した結果例を示す。図 2 中の (a) は、入力となる RGBD 画像から生成した RGB 点群、(b) は、欠損を補完した背面の RGB 点群、(c) は RGB 画像から推定した人物の三次元モデルから生成した RGB 点群、(b) + (c) は、(b) と (c) を合成した結果となっている。提案手法によって、ある程度自然に欠損を補完した RGB 点群を入手できたことがわかる。

図 3 に人物の三次元形状モデルがうまく推定できなかった例を示す。PIFu は、複雑な姿勢には対応できない。複雑な姿勢にも対応できるような他手法を参考に新

1) 東京電機大学

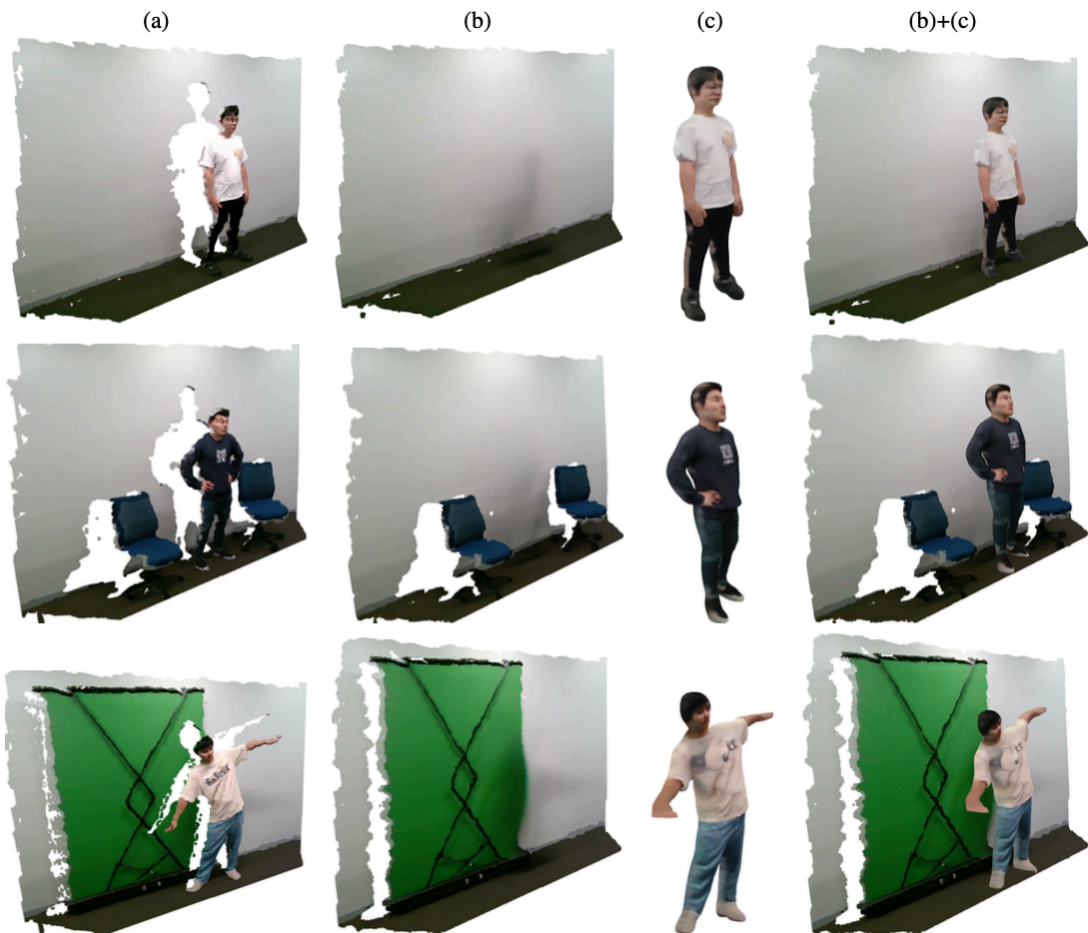


図 2 (a) 撮影した RGBD 画像から生成した RGB 点群 (b) 補間した背景の RGB 点群 (c) RGB 画像から推定した人物 3D モデルの RGB 点群 (b)+(c) の合成結果

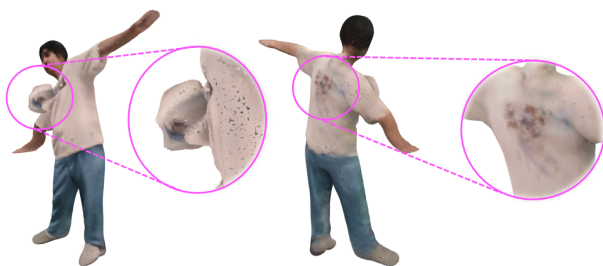


図 3 PIFu により生じた問題

たな手法の開発が望まれる。

4 おわりに

本稿では、単一の方向から RGB-D カメラを用いて対象物体を撮影し、その際に欠損した点群を補間した。具体的には、人物を RGB 画像から推定した三次元モデルから生成した RGB 点群と、人物背面に位置する壁と床の点群を補間した点群とを合成することで欠損のない RGB 点群を得た。提案方法により、人物の姿勢や、背面が複雑でない場合は、有用性の高い RGB 点群を得ることができた。今後は、リアルタイムに、より複雑な姿勢の人物や背面に対応する手法を提案する予定である。

参考文献

- [1] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R Zaiane, and Martin Jagersand. U2-net: Going deeper with nested u-structure for salient object detection. Vol. 106, p. 107404. Elsevier, 2020.
- [2] Shunsuke Saito, Zeng Huang, Ryota Natsume, Shigeo Morishima, Angjoo Kanazawa, and Hao Li. Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. 2019.
- [3] Shunsuke Saito, Tomas Simon, Jason Saragih, and Hanbyul Joo. Pifuhd: Multi-level pixel-aligned implicit function for high-resolution 3d human digitization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 2020.
- [4] Yuliang Xiu, Jinlong Yang, Dimitrios Tzionas, and Michael J. Black. Icon: Implicit clothed humans obtained from normals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13296–13306, June 2022.
- [5] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas Huang. Free-form image inpainting with gated convolution. 2019.
- [6] Zerong Zheng, Tao Yu, Yebin Liu, and Qionghai Dai. Pamir: Parametric model-conditioned implicit representation for image-based human reconstruction. 2020.