

点群データからの人物の抽出と姿勢推定に関する研究

井上晴可[†] 梅原喜政[‡] 今井龍一^{††} 神谷大介^{‡‡} 田中成典^{‡‡†} 合川舞^{‡‡†}大阪経済大学情報社会学部[†] 関西大学先端科学技術推進機構[‡] 法政大学デザイン工学部^{††}琉球大学工学部^{‡‡} 関西大学総合情報学部^{‡‡†}

1. はじめに

近年、建設現場では事故が多発[1]しており、作業員の安全管理が重要視されている。こうした背景の下、現場の安全性向上に向けた取り組みとして、画像処理などのICTを用いて作業員を自動で認識する研究が多数行われている。例えば、既存研究[2]では、カメラで撮影した動画像から深層学習を用いて人物を認識することで安全管理を支援する手法が提案されている。しかし、動画像は2次元フレーム画像の集合であり、人物の3次元的位置や姿勢を正確に把握することができないため、熱中症や転倒等の異常検知といった高度な安全管理技術には適していない。そこで、昨今、物体の3次元形状を点群データとして計測できるLiDARが注目されている。これらの機器を活用することで、3次元的に人物の位置を把握できる。そこで、本研究では、高度な安全管理の基礎技術として、点群データから人物を抽出し、その姿勢を推定する手法を提案する。

2. 研究の概要

本システム(図1)は、姿勢推定モデル構築機能と姿勢推定機能により構成される。入力データは学習用のフレームと推定用のフレームとする。出力データは人物の姿勢推定結果とする。

2.1 姿勢推定モデル構築機能

本機能では、人物を計測した点群データをConvPoint[3]により学習することで、人物の点群データを姿勢ごとに推定するモデル(以下、姿勢推定モデル)を構築する。ConvPointとは、点

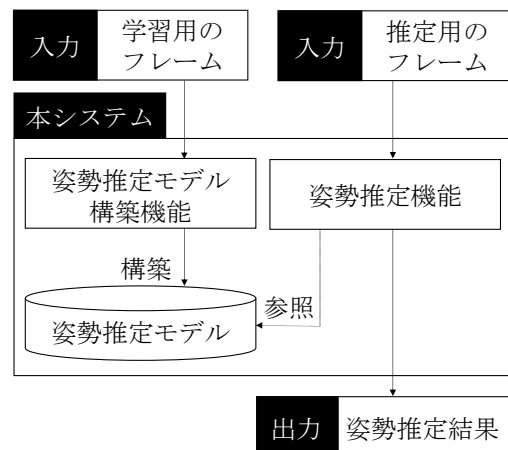


図1 本システムの概要

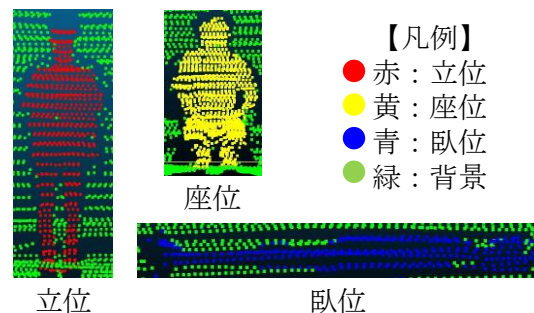


図2 姿勢推定モデル構築機能

群データから3次元的に物体を識別することができる畳み込みニューラルネットワークである。本研究では、人物の基本的な姿勢である立位、座位、臥位と背景の4つのクラスを識別対象とする。点群データに対して立位を赤、座位を黄、臥位を青、背景を緑で着色(図2)して姿勢推定モデルを構築する。なお、本研究で対象とするLiDARは、安価で計測密度が高いLivox Horizon[4]を採用する。

2.2 姿勢推定機能

本機能では、姿勢推定モデルを用いて点群データを解析することで、姿勢ごとに人物を抽出する。まず、計測している人物点群データを姿勢推定モデルに入力することで、人物を抽出する。そして、点群データに付与されたクラスに応じて着色し、姿勢推定結果を出力する。

Research for Extracting Person Posture Using Point Cloud Data

† Haruka Inoue

Faculty of Information Technology and Social Sciences,
Osaka University of Economics

‡ Yoshimasa Umehara

Organization for Research and Development of Innovative
Science and Technology, Kansai University

†† Ryuichi Imai

Faculty of Engineering and Design, Hosei University

‡‡ Daisuke Kamiya

Faculty of Engineering, University of the Ryukyus

‡‡† Shigenori Tanaka and Mai Aikawa

Faculty of Informatics, Kansai University

3. 実証実験

本実験では、提案手法の姿勢推定結果と正解データを比較し、適合率、再現率とF値により本システムの姿勢推定精度を確認する。

3.1 実験内容

本実験では、まず、高さ6mの位置にLivox Horizon[4]を設置(図3)し、計測機器から10m, 15m, 20m, 25mの4地点において、立位, 座位, 臥位の3姿勢を計測する。評価対象の人物は、学習データ, 評価データ共に2名, 内1名は共通の人物とする。計測環境は、学習データ, 評価データ共に同地点としたが、評価データにおける計測では、周囲に障害物を配置した。次に、学習データとして計測した点群データを用いて姿勢推定モデルを生成する。この際、0.2秒間ごとに、計測した点群を1フレームとし、2名および4地点, 3姿勢それぞれにおいて12フレーム, 合計288フレームの人物を学習させる。次に、評価データとして計測した点群データから学習データと同数の人物を抽出する。そして、抽出した人物に人手で姿勢情報を付与することで正解データとする。最後に、評価データから抽出した人物に提案手法を適用し、その結果と正解データの姿勢情報を比較することで姿勢および計測距離ごとの推定精度を確認する。

3.2 結果と考察

計測距離および姿勢ごとの推定精度を表1と表2, 姿勢推定結果と正解データの可視化結果を図4に示す。まず、計測距離ごとのF値に着目(表1)すると、距離の増加に伴い推定精度は多少減少したが、距離による推定精度の違いはほとんど見受けられなかった。次に、姿勢ごとのF値に着目(表2)すると、立位が0.909, 臥位が0.863と高精度に姿勢を推定できたが、座位は0.155と極端に低かった。座位は、再現率が0.969と高精度であることに対して、適合率が0.086と極端に低かった。可視化結果を確認(図4)すると、座位を正しく推定できているが、全フレームにおいて、周囲に配置した障害物を座位と誤推定していることがわかった。これは、学習データに障害物を配置していないフレームのみを採用したためだと考えられる。そのため、様々な環境下で計測したフレームを学習データに使用することで、推定精度を向上できると考えられる。

4. おわりに

本研究では、点群データを用いることで人物の3次元的な位置や姿勢を推定する手法を提案し、実証実験により有用性を確認した。今後は、学習データの増強や、様々な現場での追加実験により提案手法の汎用性の向上を目指す。

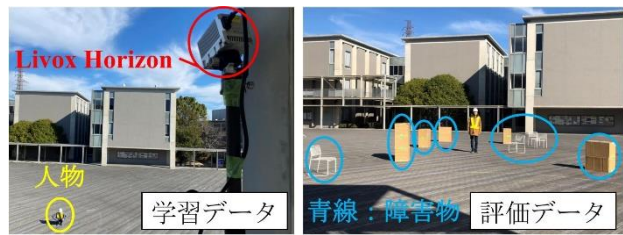


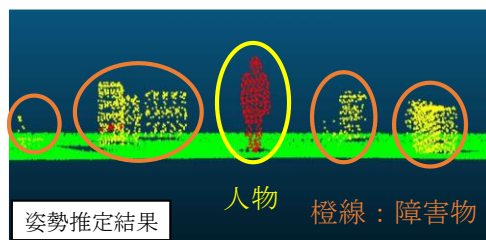
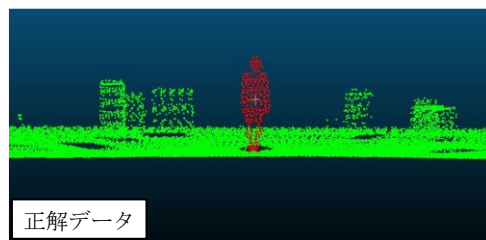
図3 実験環境

表1 計測距離ごとの推定精度

計測距離	適合率	再現率	F値
10m	0.652	0.937	0.700
15m	0.612	0.938	0.640
20m	0.607	0.916	0.621
25m	0.561	0.954	0.602

表2 姿勢ごとの推定精度

クラス	適合率	再現率	F値
立位	0.873	0.973	0.909
座位	0.086	0.969	0.155
臥位	0.871	0.868	0.863
背景	0.999	0.966	0.982



【凡例】赤: 立位, 黄: 座位, 青: 臥位, 緑: 背景

図4 姿勢推定結果と正解データの比較

参考文献

- [1] 厚生労働省: 労働災害発生状況, 厚生労働省 (オンライン), 入手先 (<https://www.mhlw.go.jp/bunya/roudoukijun/anzensei11/rousai-hassei/index.html>) (参照 2022-1-7) .
- [2] 今井龍一, 神谷大介, 井上晴可, 田中成典, 藤井琢哉, 三村健太郎, 伊藤誠: 深層学習を用いた模様付きヘルメットを装着した人物識別に関する研究, 土木学会論文集 F3 (土木情報学), Vol.77, No.2, pp. I_58-I_66 (2021) .
- [3] Boulch, A.: ConvPoint, arXiv(online), available from (<https://arxiv.org/abs/1904.02375>) (accessed 2022-1-7).
- [4] Livox: Horizon, Livox(online), available from (<http://www.livoxtech.com/jp/horison>) (accessed 2022-1-7).