

人物の移動予測システムに関する実用可能性調査*

長谷川 辰雄**

車載カメラを用いて歩行者を検知し、ドライバーに注意喚起する事故防止システムが実用化されているが、歩行者の進行方向を予測して衝突を防止するシステムは開発の途上にある。停止している歩行者がどの方向に進むかの予測は、未来の位置の観測値が得られないため経路予測を難しくさせている。本研究では、カメラ画像から人物の関節を抽出し、その関節座標の時系列変化から進行方向を予測する実用可能性について調査した。

キーワード：人物移動予測、関節抽出

Feasibility study of human movement prediction system

HASEGAWA Tatsuo

Key words : Human movement prediction, Human joint extraction

1 緒言

カメラ動画から歩行者が未来の時刻にどの場所に移動するかは、その現時点での未来の位置の観測値が得られないため、予測を難しくさせている。また、監視カメラに使われる固定カメラや車載カメラのように移動するカメラなどは、カメラの用途ごとに画像処理の方法が異なり複雑化している。そのため、歩行者のみの情報だけでなく、建物や道路などの歩行者の周りに存在する環境画像と歩行者の姿勢推定を組み合わせる方法が提案されている¹⁻³⁾。ここでは歩行者のみの情報に着目し、歩行者のひざ関節の位置情報のみを使用して、左右方向への移動についての予測の可能性を調査した。

過去の人物姿勢推定は、人物を頭部、胴体、腕、足などのパーツに分け、画像内で各パーツのテンプレートをマッチングする方法で行われていたが、精度向上に課題があった。2017年に公開されたオープンソースソフトウェアのOpenPose⁴⁾は、ニューラルネットワークを用いた関節推定で精度を飛躍的に向上させた。

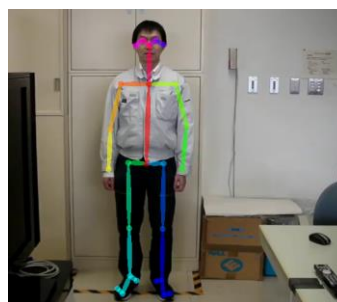
本研究では、関節の動きを用いた人物の移動予測の可能性を検証するための基礎データの取得を目的に、固定カメラとOpenPoseを用いて、人物関節座標の変化から左右移動方向の予測の可能性を調査した。

2 実験方法

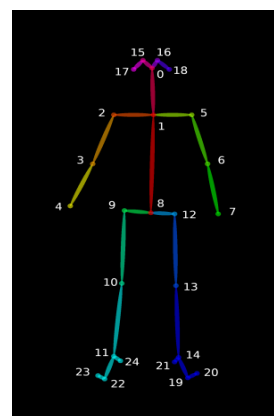
2-1 人物の姿勢推定方法

画像から人物の移動方向を予測するには、各関節の位置を推定し、その移動方向を把握する方法がある。本研究では、歩行の際に大きく変化すると予想されるひざの座標に着目し、OpenPoseを使用して歩行者の動画画像からひざの関節座標の変化を観察した。カメラ画像から人物の関節を推定する方法は幾つか提案されている。

OpenPoseはBottom-up手法と呼ばれ、画面全体の中から最初に関節箇所を全て推定した後に、その関節がどの関節なのかの関係性を決定する方法である。人が密集したシーンでも推定が可能であり、人物の動作の推定精度が良いとの報告⁵⁾があるため、今回採用した。



(a) OpenPoseの関節推定結果



(b) OpenPoseの関節番号

図1 OpenPoseによる関節位置推定

2-2 OpenPose

OpenPose法は人物関節座標を画像から推定する方法である。図1(b)に示す通り肩や肘など18~25カ所の関節座標を推定できる特長がある。OpenPoseのニューラル

* 平成31年度 技術シーズ創生研究事業 育成ステージ

** 電子情報システム部

ネットワークの構造は、VGG-19⁹⁾の畳み込み層16と全結合層3の全19層で構成されている。また、人物の関節がどのあたりに分布しているかを示す信頼度マップを、画像の特徴量を用いて関節の位置が特定できるような出力として得られることが特徴である。

図1(a)はOpenPoseによる25カ所の関節推定の結果であり、図1(b)は関節番号と各関節の接続を色別で表している。実験では、歩行する際のひざの位置(図1(b)の10番と13番)に着目し、これら左右のひざの座標位置を計測した。

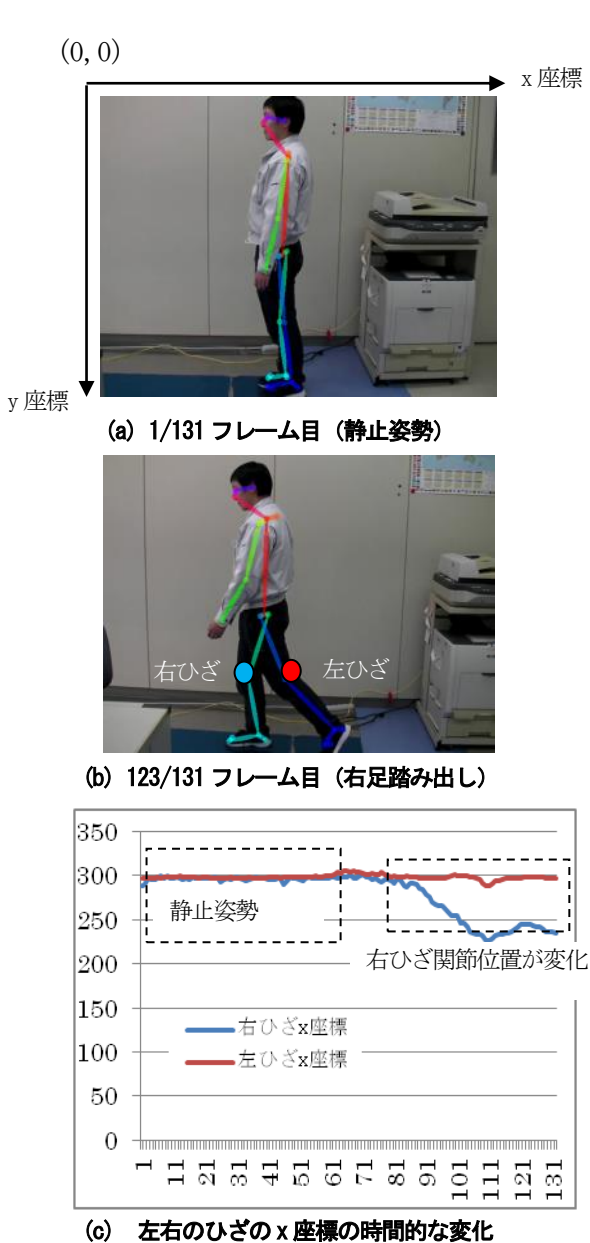


図2 右向き姿勢で右足を踏み出した関節位置の変化

3 実験結果

歩行の際に大きな変化が予想されるひざの位置について、静止状態から一步踏み出した際のひざ関節座標を取得し、歩行者の進行方向が推定可能かを調査した。

人物の左右のそれぞれの静止姿勢から、足を一步踏み出した動画を撮影し、左右のひざ関節座標のグラフ化を行った。図2(a)は右向きの静止姿勢を示しており、図2(b)は右足を一步踏み出した両ひざの座標位置を示している。また、図2(c)は左右のひざの x 座標値の時間的な変化を表しており、右ひざの x 座標が負の方向に増加していることが分かった。同様に図3(a)~(c)は左向きの静止姿勢から左足を一步踏み出した時の左右のひざの座標位置を示しており、左ひざの x 座標が正の方向に増加していることが分かった。

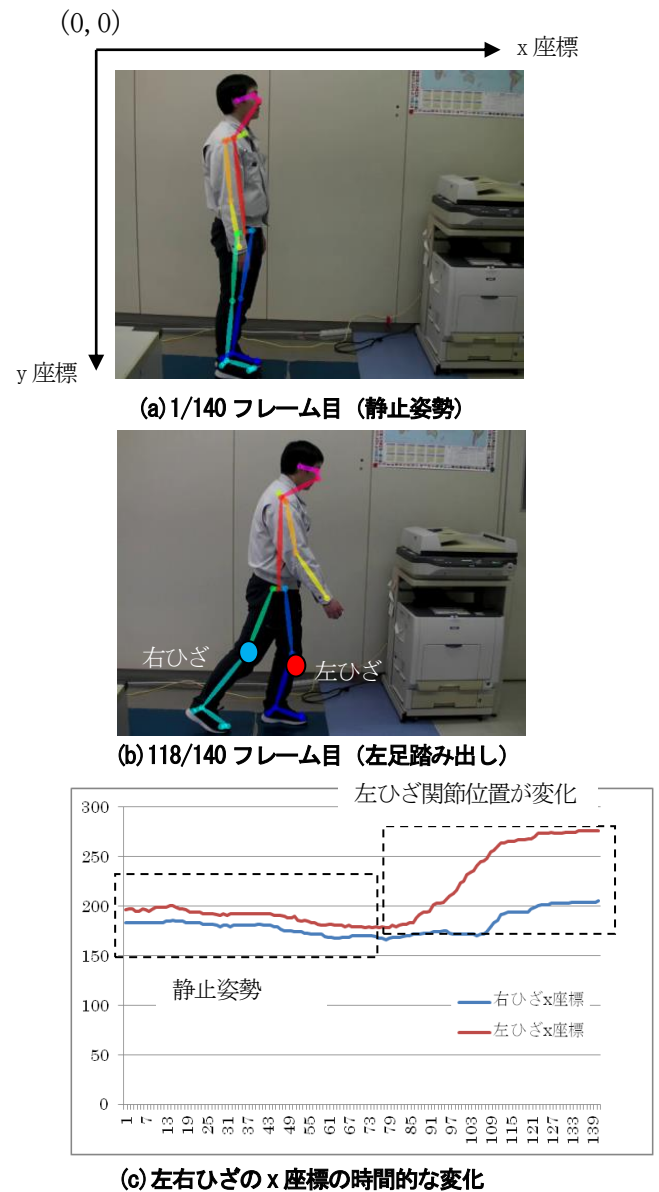


図3 左向き姿勢で左足を踏出した関節位置の変化

4 考察

固定カメラと OpenPose 法を使用して、人物の左右ひざの関節位置を取得し、両ひざ関節の x 座標の正負方向の増減を測ることで、左右の進行方向の推定が可能であることが分かった。しかし、実際には斜め方向や後ろ方向に進む場合など、さらに細かく進行方向を分類して

関節データを取得し解析する必要がある。

また、少数であるが数フレームで関節データが欠損する問題が発生した。データ欠損は画像ノイズが主な原因であり、OpenPose の AI 学習用のデータを網羅的に準備して精度を向上させる方法もあるが、多様な人物の関節を全て網羅することは不可能である。現実的には数フレームの欠損を補正するために、カルマンフィルタのような過去のデータで欠損値を補完する方法が必要である。今回の実験では基礎的な姿勢のため、データ欠損が少なくデータ補完は実施しなかった。

OpenPose 法を商用利用する場合、年間約 280 万円(2018 年時点)のライセンス使用料が発生するため、中小企業の導入が容易ではない問題がある。今回は性能の良さから OpenPose 法を選択したが、無償で商用利用が可能な PoseNet や deep-high-resolution-net を使用することで開発コストを減らすことが可能と考えられる。しかし、これらのソフトウェアは OpenPose 法に比べて、使い方が難しく、開発時間がかかるデメリットがある。

以上の調査の結果、開発コストを抑えて実用化をする場合は、PoseNet や deep-high-resolution-net の使い方習得及びその性能評価のための時間が必要と考える。

5 結 言

人物の行動予測の可能性調査として、歩行者の進行方向を OpenPose 法の関節座標取得で検証した結果、人物の右向き、左向きの基本姿勢に限って推定可能であることが分かった。今回は人物関節を検出し易い左右の姿勢で検証したが、カメラが捉える実際の姿勢は斜め方向や後ろ向きなど様々ある。また、カメラが移動する場合の調査も不足しており、現実的な人物の行動予測は本検証だけでは不十分である。過去の提案では、背景シーンと人物進行方向の抽出を組み合わせる手法が一般的となっており、リアルタイムの行動予測には人物以外の情報が必要となることが分かった。

現状の人物の行動予測は、現時刻から未来の姿勢を予測するものではなく、移動している人物の時系列画像が継続すると仮定して、その数秒先の位置を予測するものである。これまで、このような人物の経路予測には再帰型ニューラルネットワーク RNN (Recurrent neural network) を用いた方法⁷⁾や、長期短期記憶素子 LSTM

(Long short-term memory) を使った方法⁸⁾が提案されている。しかし、これらは難易度が高いため挑戦的な課題となっている。これらの方法では 4~5 秒先の人物の行動位置を予測しており、先端的な予測手法を使っても数秒先の予測が現状となっている。

文 献

- 1) K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," European Conference on Computer Vision, pp. 201-214, 2012.
- 2) S. Huang, X. Li, Z. Zhang, Z. He, F. Wu, W. Liu, J. Tang, and Y. Zhuang, "Deep learning driven visual path prediction from a single image," IEEE Trans. Image Process., vol. 25, no. 12, pp. 5892-5904, 2016.
- 3) A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," Comput. Vis. Pattern Recognit., pp. 961-971, 2016.
- 4) Z. Cao, T. Simon, S-E Wei, Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", <https://arxiv.org/abs/1611.08050>, 2016.
- 5) 高崎智香子, 竹房あつ子, 中田 秀基, 小口 正人, "姿勢推定ライブラリ OpenPose を用いた 機械学習による動作識別手法の検討", DEIM Forum 2019
- 6) K. Simonyan and A. Zisserman, "Very deep convolutional net-works for large-scale image recognition," in *ICLR*, 2015.
- 7) S. Yi, H. Li and X. Wang, "Pedestrian Behavior Understanding and Prediction with Deep Neural Networks", Proc. of the 14th European Conference on Computer Vision (ECCV2016), pp. 263-279, 2016
- 8) A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei and S. Savarese, "Social LSTM: Human Trajectory Prediction in Crowded Spaces", Proc. of the 29th IEEE conference on Computer Vision and Pattern Recognition (CVPR2016), pp. 961-971, 2016