

H-036

姿勢推定による歩きスマホ認識のための歩行検出 Walking Detection for Texting While Walking Recognition by Pose Estimation

加藤 君丸[†] 渡辺 裕[†]
Kimimaru Kato Hiroshi Watanabe

1. はじめに

近年、スマートフォンの普及により、歩きながらスマートフォンを使用する「歩きスマホ」が増加している。歩きスマホは歩行者の周囲への注意度を低下させるため危険である。特に駅ホームでの歩きスマホは、重大な人身事故を引き起こす可能性があり、問題となっている。従って、鉄道の安全運行を支援するために、歩きスマホの認識が必要である。

2. 関連研究

歩きスマホ認識の技術は過去にいくつか提案されている。皆本らは、盲導犬ロボットへの搭載を目的とし、パーティクルフィルタや HOG 特徴量を用いた歩きスマホ認識手法を提案している[1]。また、新村らは、車載カメラの映像を対象に HOG 特徴量と SVM による歩きスマホ認識手法を提案している[2]。

これらの手法は、対象の人物が歩行者であることを前提とし、人物がスマートフォンを使っているかを認識する手法である。このため、静止した人物がスマートフォンを使用している場合、それを歩きスマホと誤認識してしまうといった問題がある。これを解決するため、歩行検出により対象の人物が歩行者であるか確かめる必要がある。

3. 提案手法

3.1 概要

人物の動画から歩行検出を行う手法を提案する。提案手法は、1. 姿勢推定、2. 関節角度の計算、3. 複数フレームの情報の結合、4. SVM による分類の 4 ステップからなる。図 1 に提案手法の概要図を示す。また、図 2 に姿勢情報と関節角度の例を示す。以下それぞれのステップについて説明する。

3.2 姿勢推定

入力画像から、撮影された人物の姿勢推定を行う。姿勢推定の手法として、Realtime Multi-Person Pose Estimation[3]を用いる。Realtime Multi-Person Pose Estimation は、階層的な Convolutional Neural Network (CNN)を用いたアルゴリズムであり、入力画像から複数人物の姿勢を高速に推定可能である。Realtime Multi-Person Pose Estimation による出力は、図 2(中央)に示す人物の 14 か所の部位の入力画像における座標である。

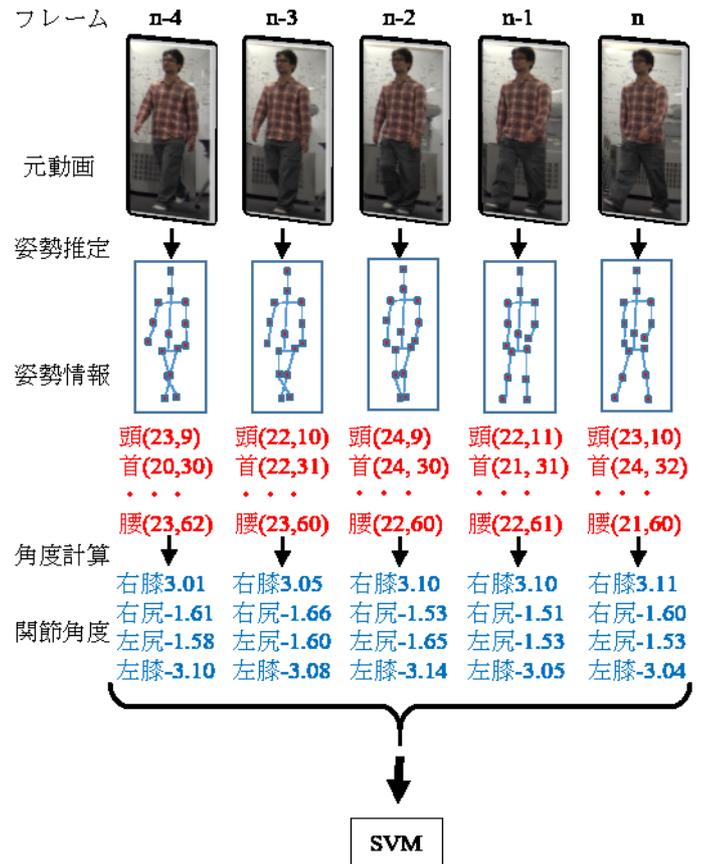


図 1 提案手法の概要図

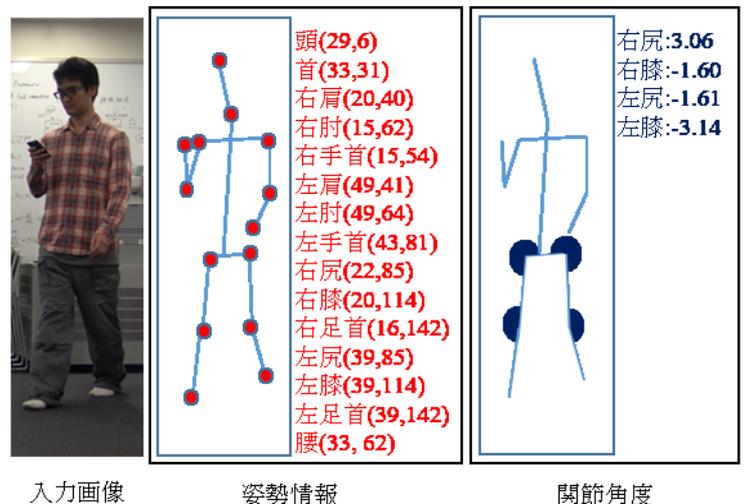


図 2 姿勢情報と関節角度の例

[†] 早稲田大学, Waseda University

3.3 関節角度の計算

推定された姿勢から、人物の膝と尻の関節角度を計算する。右足首-右膝-右尻のなす角、右膝-右尻-左尻のなす角、右尻-左尻-左膝のなす角、左尻-左膝-左足首のなす角の角度を計算し、それぞれ右膝、右尻、左尻、左膝の関節角度とする。図 2(右)に計算された関節角度の例を示す。

3.4 複数フレームの情報の結合

現在処理中のフレームで計算された関節角度に加えて、直近の過去フレーム $N-1$ 枚において計算された関節角度を一つのベクトルに結合する。図 1 の例では $N=5$ である。

3.5 SVM による分類

3.2~3.4 の処理によって得られたベクトルを学習済みの SVM によって「Walking」と「Not Walking」の 2 クラスに分類する。

4. 実験

実験データとして、歩行している人物の動画（歩行動画）と、直立している人物の動画（直立動画）をそれぞれ複数用意した。SVM の学習用データとして、歩行動画と直立動画のそれぞれから 160 フレーム、合計 320 フレームをランダムに選択した。同様に、評価用のフレームを 320 フレーム選択した。評価用のデータは、学習用のデータ選択に使われた動画とは別の動画から選択された。

$N=1$ から $N=9$ の範囲で N を変化させ、提案手法の評価を行った。評価指標として、Precision, Recall, F1-score を採用した。結果を表 1 に示す。また、 $N=5$ における混同行列を表 2 に示す。加えて、 $N=5$ において誤分類された人物画像の例を図 3 に示す。

表 1 において、 $N=1$ から $N=5$ の範囲ではいずれの評価指標も単調増加している。このことから、精度の高い歩行検出には一定数以上のフレームの情報が必要であることが確認できた。

表 2 において、直立している人物を歩行者と誤分類するケースが多い。誤分類された画像の内訳では、図 3(右)のように下半身が前方に傾いた体勢の人物が多かった。

5. おわりに

本稿では、姿勢推定を用いた歩きスマホ認識のための歩行検出手法を提案した。実験より、提案手法の有効性を確認した。

今後の課題として、座っている状態、壁にもたれかかった状態のような、歩行および直立以外の状態を含んだデータを用いた検証があげられる。また、動画における姿勢情報を用いた人物の追跡手法の検討が必要であると考えられる。

謝辞

本研究成果は、国立研究開発法人情報通信研究機構(NICT)の委託研究「ソーシャル・ビッグデータ利活用・基盤技術の研究開発」により得られたものである。

表 1 フレーム数 N と歩行検出の性能

N	Precision	Recall	F1-score
1	0.695	0.684	0.690
2	0.781	0.766	0.773
3	0.884	0.884	0.884
4	0.900	0.900	0.900
5	0.918	0.913	0.915
6	0.915	0.903	0.909
7	0.910	0.897	0.904
8	0.915	0.897	0.906
9	0.904	0.881	0.892

表 2 $N=5$ における混同行列

		SVM による分類		合計
		Walking	Not Walking	
正解	Walking	155	5	160
	Not Walking	23	137	160
合計		178	142	320



図 3 誤分類された人物画像の例
左: Not Walking と誤分類された例
右: Walking と誤分類された例

参考文献

- [1] 皆本光, 佐野睦夫, 歩行者の不注意行動認識 -歩きスマホ検出-, 第 6 回視覚・聴覚支援システム (VHIS) 研究会予稿集 (2015)
- [2] 新村文郷, 川西康友, 出口大輔, 井手一郎, 村瀬洋, 藤吉弘亘, “車載カメラ画像からの「スマホ歩き」認識に基づく歩行者の不注意度推定,” 信学技報, vol. 115, no. 98, PRMU2015-46, pp. 83-88, (2015)
- [3] 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)