

## 車載カメラ映像を用いた自転車の姿勢推定と後方確認動作検出

山崎丈<sup>†</sup> 鎬木崇史<sup>†</sup>  
国際基督教大学<sup>†</sup>

## 1. はじめに

近年、世界の多くの企業が自動運転に関連する技術の開発に取り組んでいる。日本は諸外国と比較して自転車の交通手段分担率が高いことから、自動運転技術の実装において自転車の行動予測は重要な課題である。また、日本の自転車走行帯は駐車車両とスペースを共有していることが多く、自動車との接触事故のケースとして駐車車両の追い越しシーンが挙げられる。自転車が進路上の障害物を回避する際にどのような軌道を取るかは後方確認の有無が大きく影響することが実証されており[1]、後方確認動作を検出することで自転車のより正確な行動予測とより安全な運転支援システムの実現が可能であると考えられる。

一般道や交差点を横断する歩行者や自転車を動画から検出する研究や、その将来的な動きを予測する研究は長年行われてきており、特に近年は Deep Neural Network (DNN) を用いたアルゴリズムが高い精度を実現している。Fang らは畳み込みニューラルネットワーク (CNN) を用いて自転車運転手の姿勢を推定し、手信号を検出することでその後の行動を高精度に予測できることを実証した[2]。しかし、日本において自転車が手信号を用いて意思表示することは一般的でない。

以上の状況を踏まえ、本研究は自転車の追い越し挙動の予測精度向上を目的とし、車載カメラの映像から姿勢推定を用いて後方確認動作の有無を検出する手法を提案する。

図1に本研究の概略図を示す。

## 2. 顔向き角度と姿勢推定

## 2.1 提案手法

はじめに、顔が向いている角度と姿勢推定により得られる情報の対応関係を把握するための実験を行う。自転車運転手を想定した被験者が椅子に座り、真左 (0°) から真右 (180°) まで首を動かさず様子を三脚に固定したカメラで撮影する。固定カメラは椅子の真後ろに設置し、顔全体と肩を写せるよう距離を 220cm 取る。三脚の高さは一般的な乗用車のドライブレコーダーを想定し、120cm とする。

撮影した映像から人物の姿勢推定を行うため、先行研究で Fang らが手信号の検出に用いた OpenPose [3] を利用する。OpenPose は同時検出に優れたボトムアップ型手法であり、Part Affinity Fields (PAFs) と呼ばれるベクトル空間を用いることで高精度かつ高速な検出を実現している。OpenPose によって得られる体のパーツ全 25 点の x 座標, y 座標, 確率をそれぞれ時系列データとして分析し、顔の角度と姿勢情報の対応関係を整理する。なお、座標は入力したフレームの左上を原点とし、x 座標は画面右、y 座標は画

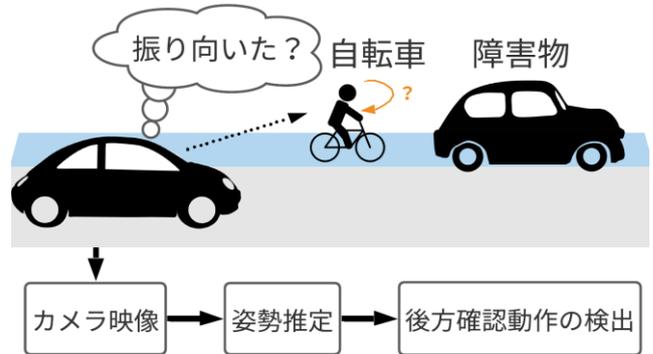


図1. 研究の概略図

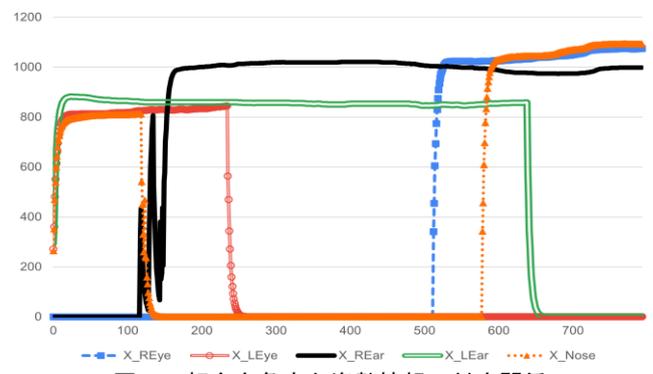


図2. 顔向き角度と姿勢情報の対応関係

面下へいくほど大きくなるほか、検出できなかった点の座標は (0, 0) が出力される。また、出力値の瞬間的な誤検出の軽減のため 1 フレーム前, 2 フレーム前との 3 点移動平均を取り影響を和らげる。

## 2.2 実験条件

屋内において Panasonic DMC-FZ200 を用いて 30fps で撮影を行い、動画を OpenPose に通して姿勢に関する情報を得た。姿勢推定は Google Colaboratory 上で実施し、GPU は Tesla K80 を使用した。なお、後述の 3 章においても同様の条件で実験を行った。

## 2.3 実験結果

得られた情報を分析したところ、全 25 点のパーツのうち顔の 5 点 (右目, 左目, 右耳, 左耳, 鼻) の検出の有無が顔の向きに大きく影響していることが確認できた。検出の有無は x 座標, y 座標, 確率の全てから判断することができるが、顔の細かい動作を最も捉えていたのは x 座標であった。以降、5 点の x 座標をそれぞれ  $REye(t)$ ,  $LEye(t)$ ,  $REar(t)$ ,  $LEar(t)$ ,  $Nose(t)$  と変数で表す。また、3 点移動平均を用いて誤検出の影響を減らすことができた。

図2に顔の角度と出力値 5 点の分析結果を示す。横軸が

フレーム(t), 縦軸が x 座標(px)となっている。左を向いた状態では左耳, 左目, 鼻が検出でき, 正面を向くにつれて右耳が見え始め, 鼻, 左目の順で x 座標が減少し見えなくなった。正面を向くと左目も検出できなくなり, 両耳のみが検出された状態となり, 右を向くにつれて右耳の座標が急激に上昇, 続いて鼻の座標も上昇し, 逆に左耳は減少し見えなくなっていった。

### 3. 後方確認動作検出

#### 3.1 提案手法

顔向き角度の実験を踏まえ, 実際の交通シーンを想定した上で後方確認動作の検出手法について考察する。なお, 本研究では左側通行を想定しているため, 被験者が顔を右側へ向けて後方確認する動作を後方確認動作と定義する。また, 動画撮影, 姿勢推定, 後方確認動作検出を全て別々に実施するが, リアルタイムでの検出も可能なようにアルゴリズムを設計する。

車載カメラを想定した固定カメラを用いて, 被験者が椅子に座り後方確認を行う様子を後方から撮影する。誤検出への頑健性を計るため, 障害物の追い越しを想定して部屋の中を動く様子や, 後方確認せずに首を左右に動かす様子も撮影する。カメラの位置は実際の自転車と乗用車の位置関係を想定し, 被験者の斜め右後方 17°, 230cm の位置に設置する。高さは顔向き角度の実験と同様, 120cm とする。

撮影した映像に対し OpenPose によって姿勢推定を行い, 顔5点の x 座標を用いた場合やそれぞれを組み合わせで計算した場合に, どの程度の精度で後方確認の検出が可能か比較検証する。精度は後方確認動作中のフレームにおける出力結果の最大値から, その他のフレームの最大値を引いた値を正規化して評価する。この値が大きいほど検出の精度が良く, 値が小さい場合は誤検出が発生していると判断できる。なお後方確認が動画内で複数回行われている場合は, 各動作における出力の最大値から最小のものを代表値として用いる。動作の開始フレームと終了フレームは目視で確認して入力するほか, 前後 150 フレームは動作移行フレームとして精度の計算に利用しない。

#### 3.2 実験結果

はじめに, 図3に後方確認1回分の出力結果を示す。被験者が右側へ顔を回すと共に  $LEar(t)$  の値が下がっていき, 逆に  $LEye(t)$  と  $Nose(t)$  の値が上昇した。後方確認時は  $LEye(t)$  が最も高く,  $Nose(t)$ ,  $REye(t)$  が続く形となった。一方で  $REar(t)$  は全体を通して大きな変化は見られなかった。被験者が後方確認を行わず首を動かす場面では  $LEye(t)$  は動作前後で変化しなかったものの,  $REye(t)$ ,  $Nose(t)$ ,  $LEar(t)$  は確認動作を行う際と同様の反応を示した。

次に, 顔5点の x 座標を単体で用いて検出を実施したところ, 体全体が動く際に誤検出が発生した。これは絶対的な座標を用いたことにより被験者の位置による影響が大きくなったためである。

一方で複数の情報を利用して相対的な位置関係に着目したところ, 体全体が動く際の誤検出を減らすことに成功した。表1に顔5点のうち2点を利用した10通りの検出結果を示す。精度は  $LEye(t)$  と  $LEar(t)$  の座標を利用したものが最

表1. 2点の位置関係に着目した際の精度

	REye_LEar	LEye_LEar	REar_LEar	Nose_LEar	REye_REar
Score	0.93	1.87	0.88	0.99	-0.81
	LEye_REar	Nose_REar	REye_Nose	LEye_Nose	REye_LEye
Score	-0.82	-0.79	-0.33	-0.92	-0.99

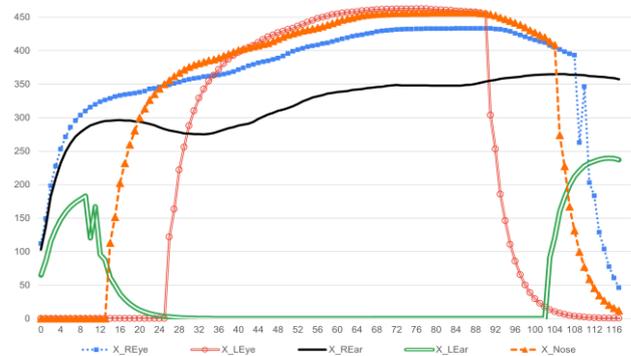


図3. 後方確認1回分の出力

も良く,  $Nose(t)$  と  $LEar(t)$  の座標を利用したものが続いた。これら以外の8通りの組み合わせについては, 首を左右に動かした際の誤検出が大きく, 適切な検出ができない結果となった。 $LEye(t)$  と  $LEar(t)$  の座標を利用したものが最も精度が良かったのは, 顔の角度と各点の検出の有無に着目した際, 左目の検出される角度が最も大きいためである。

以上の結果より, 姿勢推定によって得られた顔5点の x 座標のうち, 2点の位置関係に着目することで後方確認動作の検出が可能であることが確認できた。また, 右目や鼻よりも左目と左耳の位置関係を利用すると誤検出の可能性が低いことが検証できた。さらに, 右側通行を仮定して被験者が顔を左側へ向けて後方確認した際の出力に着目すると, 動作の左右が逆転するため  $REye(t)$  と  $REar(t)$  の座標を利用したものが最も精度が良かった。このため, 右側通行の国においても本提案手法は有効であると考えられる。

#### 4. まとめ

本稿では, 姿勢推定を用いて自転車の後方確認動作を検出する手法を提案し, 精度の比較検証により検出の有効性を示すとともに, 姿勢情報の組み合わせにより精度が向上することを実証した。本研究は自転車の行動予測モデル精度向上への足がかりとなるほか, より安全な運転支援システムの開発への応用が可能であると考えられる。今後は, 実際の交通シーンを想定して屋外の映像を用いて後方確認動作の検出を行う。

#### 参考文献

- [1] 鈴木美緒; 細谷圭介; 屋井鉄雄. “自転車の走行空間整備に向けた駐車車両追い越し挙動に関する基礎的研究.” 交通工学論文集 2.1 (2016): p. 21-30.
- [2] Fang, Z.; Lopez, A.M. “Intention recognition of pedestrians and cyclists by 2d pose estimation.” IEEE Transactions on Intelligent Transportation Systems. 2019, 21.11: p. 4773-4783.
- [3] Cao, Z. et al. “OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields.” IEEE transactions on pattern analysis and machine intelligence. 2019, 43.1, 14p.