

運転環境における高精度関節位置推定と姿勢評価システム

山田 高寛¹ 土門 憲司² 入江 英嗣¹ 坂井 修一¹

概要: 本論文では運転環境下における高精度な人体関節位置推定手法と、高精度な非通常姿勢検知アルゴリズムを提案する。関節位置推定手法については、人体のみならず運転環境を含めたデータおよび知識を学習器に与えることで精度を高めた。また、運転姿勢検知アルゴリズムは複数フレームを神経発火モデルに基づき累積させていくことでノイズに対する頑強性を高めた。更にこれらの提案手法を統合させたシステムを実車環境上に実装して走行試験を行い、非通常姿勢の検知に関する評価を行った。その結果検知率 90 % を達成することができ、運転姿勢検知システムの実車に対する有用性が確かめられた。

High Accuracy Joint Position Estimation in Driving and Posture Detection System

TAKAHIRO YAMADA¹ KENJI DOMON² HIDETSUGU IRIE¹ SHUICHI SAKAI¹

1. はじめに

近年、計算機資源やセンサ類の小型化・高性能化が進み、メディア処理などの高度な技術が様々な機器へ組み込まれるようになり、運転者支援でも活用している例が増えてきている。例えば、[1] では近赤外線カメラを利用し、ドライバのまぶたの開閉具合を検出している。この情報を元にしてドライバの眠気を推定している。また、加速度センサでドライバの腕の動きを利用している例もある [2]。

このように運転者支援において、ドライバ情報は重要な要素である。ドライバの状態が判別できれば、例えばドライバの意識消失時の交通事故予防などの幅広い運転者支援を行うことができる。また今後発展するとされている自動運転の分野では、機械から人間へと操作を移さなければならないタイミングがある [3]。この切り替えの際、人間側が操作を可能かどうかの判断にドライバ情報は不可欠である。

ドライバの情報で未だ用いられていない情報に、ドライバの姿勢がある。今まで用いられてきた情報に加えてドライバの姿勢が利用可能となれば、より外乱にロバストで強

力な運転者支援を行うことができる。このような目的から我々は運転者支援のための辞書照合システムを提案している [4]。このシステムは Kinect などの深度センサを用いて人体の関節位置を取得し、この座標群から姿勢記述モデルを用いて姿勢を判断する。通常運転時に観測された姿勢を辞書に蓄え、辞書にない姿勢が表れたら検知する。既に Microsoft Kinect の API を用いたシステムを実装しており、室内シミュレータ上で予備評価を行っている。

本論文ではドライバの姿勢検知精度を高めるため、運転環境に特化した関節位置推定手法と、時間要素を考慮した運転姿勢検知アルゴリズムを提案する。

非通常姿勢の検知率を高めるために、独自開発した関節位置推定手法を用いる。詳細は後述するが、人体のみならず運転環境ごと学習を行い、また運転環境ゆえの限定条件を利用して運転環境下の関節位置推定精度を高めた。

またノイズ耐性をつけるために、非通常姿勢検知アルゴリズムを改良し時間的連続性を考慮できるようにする。具体的には単純なフレームごとの辞書照合ではなく、複数フレームを神経発火モデルに基づき累積させていくことで、ロバストな姿勢検知を行う。

更にこれらの提案手法を統合させたシステムを実装し、実車環境下において想定した運転姿勢の検知具合に関する評価を行う。検知漏れ、または誤検知が発生しないかを調査することで、統合システムの実車環境に対する有用性を

¹ 東京大学 大学院情報理工学系研究科
Graduate School of Information Science and Technology,
The University of Tokyo

² 東京電気通信大学 大学院情報システム学研究科
Graduate School of Information Systems, The University of
Electro-Communications

確かめる。

本論文の構成は以下の様になっている。2章ではドライバ情報や人体姿勢推定手法などの関連研究について述べる。3章では提案手法の概要を述べる。4章では運転環境下における高精度人体関節位置推定手法について、5章では姿勢検知システムについて提案手法の詳細について述べる。6章では実車試験環境についての詳細を述べ、7章で提案システムを用いた運転姿勢検知に関する評価を行う。そして、8章で結論を述べる。

2. 関連研究

ドライバの情報を利用している先行研究は、生体センサなどの人体に付着させるセンサ方式を利用しているものと、RGBカメラなどの非侵襲なセンサを用いているものにわけられる。

[2]では加速度センサをドライバの腕に装着させることで、ドライバの運転動作の解析を行っている。[5]では生体センサを用いて脳波 (EEG) を計測している。これを用いて脳の覚醒度合いを見積もり、ドライバの疲労具合を検出している。

[1]では、近赤外線カメラを用いて、ドライバの顔画像を取得して瞼の開閉度合いを検出している。[6]ではRGB-Dカメラを用いて、ドライバの顔の向きと目の状態を推定し、ドライバの疲労度合いを推定している。

人体の姿勢推定手法については [7] が大きな成功を収めている。depthセンサを用いて、各ピクセルに深度距離が保存されている深度画像を取得する。各ピクセルに対して、周囲のピクセルの深度距離を用いて特徴量を算出する。その特徴量を Random Forest [8] という複数の決定木から構成される集合学習器にかけて、身体部位を算出する。それぞれの身体部位のピクセルに対して、Mean Shift [9] というクラスタリング手法を用いて代表点を抽出して、人体の関節位置座標を推定する。

このように得られた人体の関節位置座標を姿勢表現に変換するのには [10] がある。得られた関節位置座標を、胸部を原点とする球座標系における角度表現に変換し、ジェスチャ認識を行っている。

3. 提案手法

今回我々が提案する運転姿勢検知システムは図1のようになっている。提案システムでは、姿勢表現の取得段階と、運転姿勢検知段階の大きく2つに分けることができる。以下ではこのシステムの流れを説明する。

まず、ドライバの前方上部に設置されたRGB-Dセンサを用いて運転風景を撮影することで、運転環境の深度情報を取得する。今回はRGB-DセンサとしてMicrosoft Kinect v2を使用した。取得した深度情報を、今回新たに提案する運転環境に特化した人体関節位置推定手法を用いてドライ

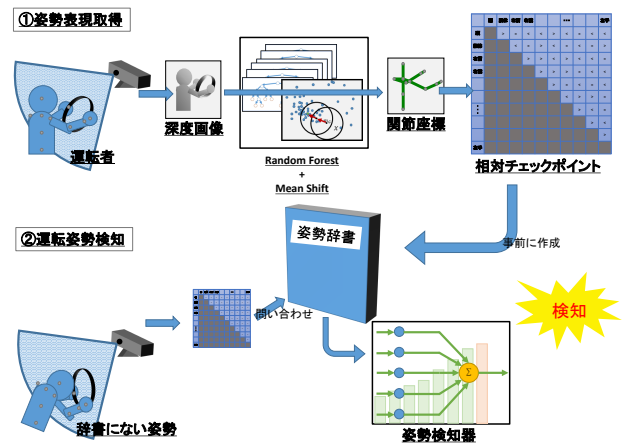


図1 提案システム概要

バの関節位置座標を推定する。推定した関節位置座標を相対チェックポイント [11] を用いて数値表現から姿勢表現に変換する。そしてこの姿勢表現を用いて、検知しなくて良い通常時の姿勢を辞書として事前に保持しておく。

運転姿勢を検知する際には、先と同様の流れで姿勢表現を取得する。その姿勢表現を事前に作成しておいた辞書に問い合わせを行い、通常時の姿勢とどのくらい差異があるのかを算出する。この差異を使って、非通常度というものを算出する。算出した非通常度を時間にしたがって積算していき、ある閾値を超えたところで通常ではない姿勢として検出する。

4. 運転環境における高精度関節位置推定手法

以下では既存手法の運転環境下での問題点を述べ、提案システムのうち、運転環境の深度距離を取得してからドライバの関節位置座標を推定するまでの詳細を述べる。

4.1 運転環境下の関節位置精度の低下

KinectのAPIを用いて、運転環境下で人体関節位置推定を行うとその推定精度は低下する。例えば、図2では頭部と胴体部の位置推定がうまく行えていない。これはシートやハンドルなどの邪魔な障害物が影響しているからであり、運転環境においては避けることができない。

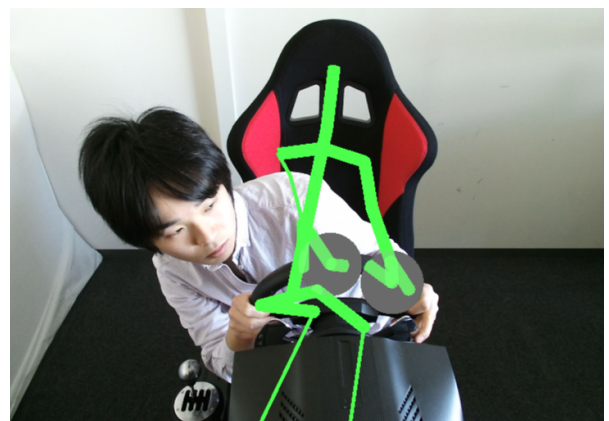


図2 頭部、胴体部の位置推定がうまく行えていない例

このように関節位置推定が上手く行えないと、提案システム後段の姿勢検知段階において悪影響が出てしまう。もし、実際のドライバの姿勢変化が僅かなものでも関節座標位置の推定結果が一意に決定しないと辞書の肥大化、ひいては姿勢検知精度の低下を招く。そこで、既存の関節位置推定手法 [7] を活用し、運転環境下での工夫を施すことで運転環境におけるドライバの関節位置推定精度を高める手法を提案する。

4.2 提案手法の紹介

既存の関節位置推定手法では、Random Forest の学習に人体のみのデータを用いている [7]。そこで提案手法では、人体のみならず、シートまたはハンドルなどの運転環境に特有の障害物を含めた学習を行う。これは運転環境下ではシートとハンドルは固定されているものであり、学習の前後で変化しないことを利用した。その他に運転環境に合わせた、以下の様な工夫を行った。

得られた深度画像に対して、ある閾値を超えた背景に当たる部分については推定を行わないことにした。運転中にドライバの深度方向の大幅な移動は考えられないので、人体関節位置推定には影響が出ないはずである。背景という複雑な環境下の情報を削除することで、推定精度が高められると考えた。

また、天井の影響をなくすため、特徴量の計算を行う際に用いるオフセットベクトルを垂直上方向のものは用いないことにした。

さらに今回のアプローチでは、学習を個々の環境下で行う必要が有るため、それに伴う学習データセット作成支援ソフトを開発した。

4.3 学習データセット作成

Random Forest の学習には大量の学習データセットが必要であり、今回新たに学習データセットが必要である。実際には深度画像に対して、身体部位別に色分けをする必要がある。この学習データセットの作成作業を簡便にするた



図 3 身体部位ごとに色分けされた T シャツ

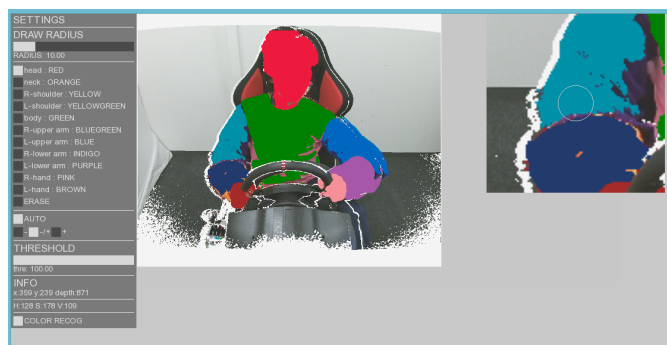


図 4 ラベルデータ作成支援ツールのキャプチャ画面

めに支援ツールを開発した。このツールでは身体部位別にラベルを作成するのに色認識技術による支援を用いた。ラベルデータを作成する手順は以下の通りとなる。

まずデータを収集する際に、図 3 のような身体部位ごとに色分けされた T シャツをドライバに着てもらおう。次に RGB-D センサを使用して、運転している様子の深度画像と RGB 画像を取得する。こうして得られた 2 つの画像を開発したツールに読み込ませる。開発したツールでは読み込み時に RGB 画像を用いて、色認識技術により身体部位ごとに色分けを行う。色認識には HSV 表色系における各値の比較を行うことによって実現した。また色認識が上手く行えないところには、深度画像による情報を利用して 3 次元距離が近いものには同じラベルを付与することを行った。図 4 が実際のラベルデータ作成画面である。色認識と深度情報を利用した身体部位の色分けでも不十分な場合が生じる。その部分は人手による補正を行う。細かな部分の補正を行いやすくするために、注目したい部分領域を拡大して表示できるようにした。

データセットの偏りをなくすために、重複するデータを作成しないようにした。具体的な基準としては [7] に従い、各データ間で同じ関節部位の座標間距離が 5cm 以内に収まっている場合は同じ姿勢とすることにした。この基準をラベルデータ作成の際に適用させるために、作成者に関節位置を指定してもらうことにした。この情報を用いて、今まで作成したデータの中に似通った姿勢が含まれる場合はそのデータに関してはスキップし、ラベルデータの作成を行わないことにした。これらにより、平均して 1 時間あたり 60 枚の学習データセットを作成することができた。

4.4 実装方法

実装は C++ 言語を用いて行った。画像処理に関するところは OpenCV を用いた。Random Forest の各種パラメータには表 1 のものを使用した。

表 1 Random Forest のパラメータ

決定木の数	3
決定木の最大深さ	18
特徴量の候補数	1000
サンプルピクセル数	1000
オフセットの最大値	100 ピクセル
クラスラベル	頭, 右左肩, 胴体, 右左上腕, 右左下腕, 右左手



図 5 室内運転環境シミュレータ

4.5 予備評価:関節位置推定結果

予備評価として実際に図 5 のような室内運転環境シミュレータ上で学習を行い, ドライバの関節位置を推定してみた。

図 6 に推定結果を載せる。これを見ると, 少ないデータ数にもかかわらず既存手法に対して遜色ないレベルで行えた。また, 既存手法では精度が低下していた姿勢に対して, 提案手法では高精度に関節位置の推定を行うことができた。



図 6 室内運転環境シミュレータ上での関節位置推定結果. 左側は提案システム, 右側は既存手法

5. 運転姿勢検知システム

5.1 姿勢表現方法

	頭	胴体	右肩	右腕	...	左手				
頭		>	=	<	<	>	<	=	<	<
胴体			<	=	>	<	=	<	=	>
右肩				>	>	<	>	<	<	=
右腕					>	>	>	=	<	<
...						>	=	<	=	<
...							<	>	>	=
...								>	<	=
...									>	<
左手										>

図 7 相対チェックポイント

提案システムでは, こうして推定した関節座標を数値表現から姿勢表現へと変換する。姿勢表現方式には相対チェックポイント [11] を用いる。相対チェックポイントとは, 関節間の相対的な座標関係を姿勢記述として用いる方法である。絶対的な数値を用いるのではなく, 相対的な表現を用いることにより人体の体格にロバストな姿勢記述を行うことができる。具体的には, 図 7 のように関節間のすべての組み合わせにおいて座標の大小比較を行い, それぞれの組み合わせにおいて, 大きいか, 小さいか, 同じかを姿勢記述として用いる。

また, 相対チェックポイントは 1 つの姿勢を記述するのにデータ量が少量で済む。今回は上半身の姿勢のみを用いるので, 使う関節は 10 個である。すべての 10 の関節の組み合わせは, ${}_{10}C_2 = 45$ 通りで, 各組み合わせに対して x, y, z の 3 次元と, 大, 小, 同程度の 2bit で, 合計 $45 \times 3 \times 2 \times 2 = 270\text{bit}$ のデータ量となる。対して座標数値表現の場合は, 座標数値を浮動小数点型として少なくとも 32bit で表すと, $10 \times 3 \times 32 = 960\text{bit}$ とデータ量が大きくなってしまふ。

5.2 神経発火モデルに基づく運転姿勢検知モデル

この相対チェックポイント記述を用いて, 検知しなくて良い通常時の姿勢だけの姿勢辞書をあらかじめ作成する。先程述べたように相対チェックポイントは 1 つの姿勢を記述するのに, 軽量なので辞書量の爆発は起こらない。

運転姿勢を検知するときには, まず姿勢辞書に含まれる通常運転動作の全姿勢との比較を行う。そして, 辞書内の全姿勢との類似度を算出する。類似度の算出方法は, 姿勢辞書に含まれる各姿勢の相対チェックポイントの一致数とする。

次に, 現在の姿勢がどの程度通常の姿勢と違っているか

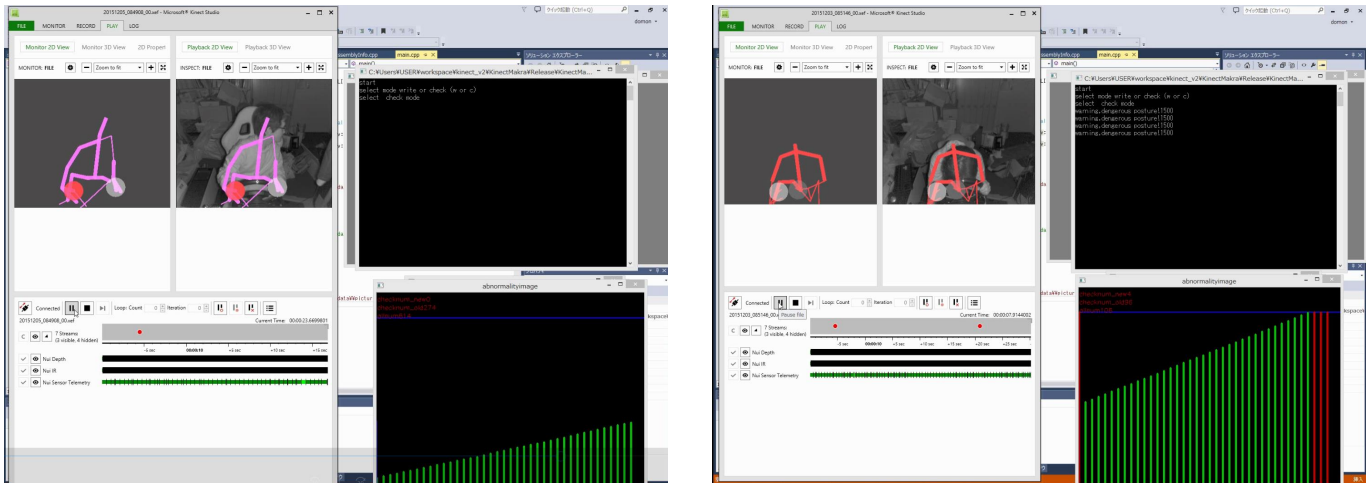


図 8 室内運転環境シミュレータ上での運転姿勢検知の様子

を表す距離を算出する。先程算出した類似度から最大のものを見つける。その最大類似度を厳密度と呼ぶ閾値と比較する。厳密度を超えたものは辞書内の姿勢と類似しているものなので、通常姿勢とする。もし、厳密度を下回っていたら辞書内の姿勢とは違った姿勢とする。この最大類似度と厳密度との差を現在の姿勢と通常時の姿勢との距離とする。

単純に閾値を下回ったものに対して姿勢検知としてもよいが、これではノイズに対して誤検出をしてしまうので神経発火モデルに基づいたモデルを用いる。このモデルでは、時間ごとに非通常度と呼ぶ姿勢を判定するための閾値を累積していくことで、通常姿勢の辞書にない姿勢を検出していく。先程得られた通常時の姿勢との距離が正であるならば、距離の数値を元に累積されている非通常度を減算する。具体的な減算値の算出方法は式 (1) のようになる。

$$(\text{減算値}) = (\text{現在の非通常度}) / (FPS \times T_{\text{observ}}) \quad (1)$$

FPS は現在のフレームレートである。この T_{observ} は累積非通常度をどのくらいの時間で 0 にするかによって決定される。

5.3 シミュレータ上での予備評価

図 8 は室内運転環境シミュレータ上で、実際に運転姿勢検知を行った様子である。図中の左側のように緑色の縦棒が累積非通常度であり、この色が緑色であるかぎり通常姿勢として検知する。累積非通常度がある閾値を超えたとき、図中右側のようにこの縦棒は赤くなり、非通常時姿勢として検知する。

6. 実車試験環境

今回この姿勢検知システムを評価するのに、実車環境において試験を行った。今回の試験では、通常の運転姿勢では想定されない姿勢を検知することを目的とする。また、誤検知の様子を調べるために紛らわしい通常運転姿勢を含め

た。今回の試験の詳細は以下のとおりである。

被験者は男女 1 名ずつの合計 2 名で行った。試験の内容は、高速時試験、低速時試験、非通常運転時の 3 通りである。

高速試験では、通常の運転姿勢に加えて一時的に手をハンドルから離してしまう付属機器類の調整、ギア操作、車線変更時の巻き込み確認を含めた。具体的な頻度は、付属機器類の調整を 5 回、ギア操作を 10 回、車線変更時の巻き込みを 5 回行った。

低速試験でも紛らわしい操作として、ハンドルを大きく回転させる操作、ギア操作、右左折時の巻き込み確認を含めた。こちらも具体的な頻度は、ハンドル操作、ギア操作、巻き込み確認それぞれ 10 回行った。

非通常運転では、右倒れこみ、左倒れこみ、前夫倒れこみ、後方倒れこみ、ずり下がりの 5 つの姿勢を検知することを目的とした。

被験者 1 名に対して、高速試験と低速試験を 1 回ずつ、非通常運転時試験を 2 回分評価した。

7. 評価結果

厳密度を 40, 37, 35 の 3 通りに変化させて評価を行った。

7.1 高速時試験結果

高速時の試験結果は図 9 のようになった。ギア操作、左右確認の姿勢は誤検知されなかったが、厳密度が 40 の時の、付属機器調整の誤検知が目立つ形となった。実際に誤検知した時の様子を調べてみると、左手を使って付属機器の調整をしている時であった。右手で付属機器を操作している時は見つからなかった。これは姿勢辞書に含まれている右手で行うギア操作の姿勢と、右手で付属機器を調整した時の姿勢が似通ったものであり、検知されなかったためである。

7.2 低速時試験結果

低速時の試験において、誤検知は発生しなかった。ハンドル操作の際に、累積非通常度が反応している様子が確認されたが、判別閾値まで達することはなかった。これは単純に姿勢辞書に含まれていない姿勢を検知するのではなく、累積非通常度によって判別を行うこととしたためである。

7.3 非通常運転時試験結果

非通常運転時の試験結果は図 10 のようになった。右方倒れ時、左方倒れ時、前方倒れ時、ずり下がり時の項目に関しては漏れがなく検知することができた。しかし、後方倒れ時の姿勢は検知漏れが発生した。実際の後方倒れ時の姿勢を確認して見ると、検知できたものは頭部位置の変化が大きなもの、検知漏れしたものは頭部位置の変化が乏しい物であった。このことが理由で検知漏れが発生した。実際、判定が厳しい厳密度の数値が大きい方がうまく検知することができている。

8. まとめ

本論文では、運転環境に特化したドライバ関節位置推定

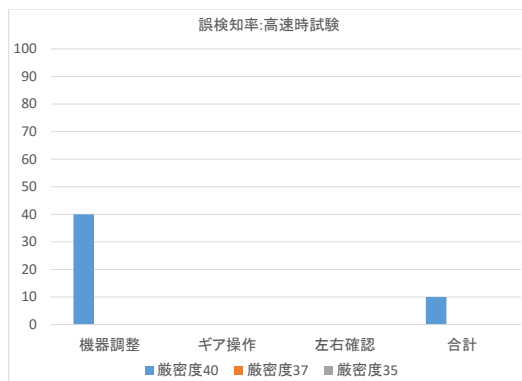


図 9 高速時試験結果

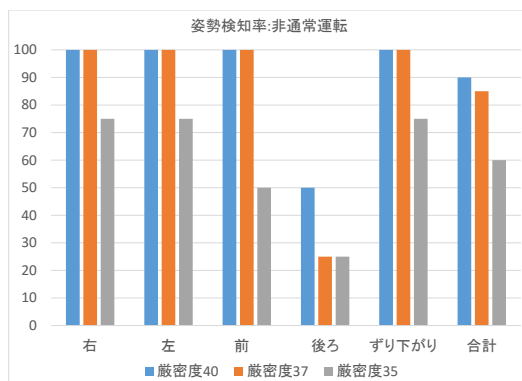


図 10 非通常運転時試験の様子

手法と動きに対応した運転姿勢検知システムを提案し、その実車評価を行った。評価の結果、パラメータの調整にもよるが、誤検知率は最大で 10 %、想定姿勢の検知率は最大で 90 % 達成する事ができた。

また、それぞれの評価結果を詳細に見ていくと、姿勢辞書に保管されていない姿勢を単純に検知するのではなく、前後のフレームの情報を姿勢検知に使うことで誤検知を防ぐことができた。

ただ、想定した姿勢の検知漏れが発生した。これは今回非通常度の算出の際に、各関節に同じ重みづけを行っていたからであるからである。関節ごとに、例えば動きが多い部分については重みを小さくするなどの適切な重みづけをすれば防ぐことができたと考えられる。また、今回相対チェックポイントを姿勢表現に用いたが、相対チェックポイントは姿勢間での関節位置の小さな変化を表現するには難しいため、検出漏れが起きたと考えられる。

今後の課題としては、提案システムの実車試験での有用性が確かめられたため、更なる実車試験を行い普遍的に運転姿勢の検知ができるかを確認する必要がある。また、今後大量のデータが必要になった際には人手では対応できないため、学習データセットの作成が自動的にできるようにすることがあげられる。更には、運転姿勢、ひいては一般的に人体姿勢を記述するのにどういった記述表現が適しているのかを評価する必要がある。

参考文献

- [1] 大見拓寛. 運転者の居眠り状態評価の画像センサ. 人工臓器, 42(1):99-103, 2013.
- [2] 多田昌裕, 納谷太, 岡田昌也, 野間春生, 鳥山朋二, and 小暮潔. 無線加速度センサを用いた模範運転動作からの逸脱検出. 人工知能学会論文誌, 23(3):105-116, 2008.
- [3] 須田義大, 大口敬, 中野公彦, 大石岳史, 小野晋太郎, 吉田秀範, and 杉町敏之. 自動運転システムの社会実装に関する課題と展望. 生産研究, 68(2):95-98, 2016.
- [4] 土門憲司, 野村隼人, 吉見真聡, 入江英嗣, 吉永努, 坂井修一, et al. Rgb-d センサと学習による運転姿勢検知. 研究報告計算機アーキテクチャ (ARC), 2015(20):1-3, 2015.
- [5] Hamzah S AlZu'bi, Waleed Al-Nuaimy, and Nayel S Al-Zubi. Eeg-based driver fatigue detection. In *Developments in eSystems Engineering (DeSE), 2013 Sixth International Conference on*, pages 111-114. IEEE, 2013.
- [6] Liyan Zhang, Fan Liu, and Jinhui Tang. Real-time system for driver fatigue detection by rgb-d camera. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(2):22, 2015.
- [7] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1):116-124, 2013.
- [8] Leo Breiman. Random forests. *Machine learning*, 45(1):5-32, 2001.
- [9] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*,

24(5):603–619, 2002.

- [10] Michalis Raptis, Darko Kirovski, and Hugues Hoppe. Real-time classification of dance gestures from skeleton animation. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics symposium on computer animation*, pages 147–156. ACM, 2011.
- [11] 黒田修平, 放地宏佳, 吉見真聡, 吉永努, 入江英嗣, et al. 相対座標を用いた運動指導システム. *マルチメディア, 分散協調とモバイルシンポジウム 2013 論文集*, 2013:1475–1482, 2013.