

歩容データを用いた「歩きスマホ」の検出

高松 毅瑠¹ 三好 駿¹ 菊池 浩明¹

概要：歩きながらスマートフォンの操作を行う、いわゆる、「歩きスマホ」が近年問題になっている。歩きスマホを行っている人は注意散漫になり、他の歩行者と衝突してしまう恐れがある。そこで本研究では、モーションキャプチャーによって外部から歩き方を観測し、歩きスマホを自動検出することを試みる。歩様データから複数の関節の組み合わせによる特徴量を用いた機械学習アルゴリズムにより、汎用性のある歩きスマホの検出システムを実装し、121名の歩容データを用いた検出精度を報告する。

Detecting “Smartphone zombie” from gait pattern

Takeru Takamatsu¹ Shun Miyoshi¹ Hiroaki Kikuchi¹

1. はじめに

近年のスマートフォンの普及と共に、歩行中のスマートフォン操作、いわゆる「歩きスマホ」による事故が問題になっている。歩きスマホを行っている人は注意散漫になり、他の歩行者と衝突してしまう恐れがある。平成22年から平成28年までの7年間で歩きスマホ等に係る事故により救急搬送された人数の推移を表した図1からも、事故件数は近年増加傾向にあることが観測できる。加えてスマートフォンの普及率が増加傾向にあることから、今後も増加していくことが予想される。

歩きスマホの検出にはいくつかの研究が先行している。加藤らによる画像処理による姿勢制御[2]や児玉らによる位置センサを用いた障害物検知[4]などが知られている。しかしながら、コストのかかる3次元の画像処理や専用アプリのインストールなどが必要であった。

そこで、本研究では、モーションキャプチャーによって外部から歩き方を観測し、歩きスマホを自動検出することを試みる。スマートフォンが見えない位置から観測した時でも、他の体の部位から推定した歩き方から歩きスマホを検出することができる。

しかしながら、過度に学習データに特化した特徴量を用いると個人差に対応できず汎用性を損ねてしまう。そこで、本稿で提案する手法では、歩容データから、右手と胸の距離等の時系列データを算出し、標準偏差、中央値などの統計量で表す。関節間の距離の統計量を特徴量とする。平均値や中央値では歩く姿勢についての静的な特徴、標準偏差では腕の振幅などの動的な特徴を表すことができる。

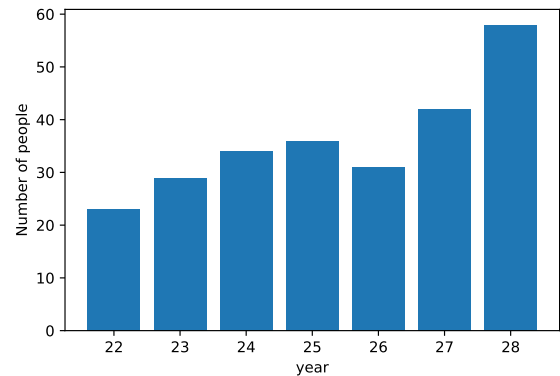


図1 救急搬送人員 [6]

本稿では、機械学習アルゴリズムによって各特徴量の重要度を計算することで歩きスマホの検出のためにどの点に着目する点を明らかにする。歩きスマホ検出システムを実装し、121名の歩容データを用いた検出精度を報告する。

2. 先行研究

歩きスマホ検出の技術は大きく分けて2つある。1つ目は外部から撮影した画像から歩きスマホを検出するシステムである。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[3]。また、加藤らはRealtime Multi-Person Pose Estimationを用いた姿勢推定によって得られた姿勢情報をもとに、スマホ使用姿勢検出、歩行検出、把持物体認識の3ステップによる歩きスマホ検出手法を提案している[2]。8方向を向いた人物の動画による検証を行い、F1-scoreにして0.852という高い精度での検出を報告している。しかし、これらの手法は単眼RGB画像を用いた歩きスマホの判別であり、高度な画像処理と姿勢推定に

¹ 明治大学 総合数理学部 先端メディアサイエンス学科

Department of Frontier Media Science, School of Interdisciplinary Mathematical Science, Meiji University

大きなコストがかかっていた。

2つ目は個人が持っている端末の情報から歩きスマホを検出するシステムである。兄玉らはアンケート調査からどのような時に歩きスマホが危険となるのかを調査した。距離画像センサーを用いて前方の障害物との距離と相対速度を測り、衝突の危険度を計算することで、危険度ごとに個人が持っている端末に警告を行うシステムを提案している[4]。名古屋大学東山キャンパス内で歩きスマホ中の衝突回数を測る実験を行い、衝突率を13.3%から2.2%に減少させた事を報告している。ただし、GPSを使ったマップ機能など移動していることを前提としたアプリケーションを使うことができない問題があった。そこで小野らはスマートフォンの加速度センサを用いた歩きスマホの検出。画面のタッチ頻度からスマートフォンに向ける注意の定量化による歩きスマホ検出、警告のアプリケーションを作成し問題を解決した[5]。しかし、本手法では個人の端末にアプリケーションを入れる事を前提としており、第三者が歩きスマホに対して警告することができない。

3. 提案手法

3.1 概要

本研究では、モーションキャプチャデバイス kinect v2 による関節の3D時系列データからサイクル切り出しを行い、1歩分の歩行データを定める。歩行データに対して歩きスマホに特有の特徴量を抽出、ランダムフォレストによる学習を行い、歩きスマホの検出器を作成する。

3.2 適合率、再現率、F値

本研究では、歩きスマホ検出器の精度を適合率 P と再現率 R 、 P と R の調和平均である F 値により評価する。適合率は歩きスマホと推定したデータのうち、正しいものの割合であり、再現率は歩きスマホの全データのうち正しく検出できたものの割合である。 F 値は適合率と再現率の調和平均をとった値であり、各々次の様に定める。

$$P = \frac{\text{歩きスマホと正しく判定したデータ数}}{\text{歩きスマホと判別したデータ数}}$$

$$R = \frac{\text{歩きスマホと正しく判定したデータ数}}{\text{歩きスマホのデータ数}}$$

$$F = \frac{2}{\frac{1}{R} + \frac{1}{P}}$$

3.3 サイクル切り出し

採取したデータから1周期分を抜き出す。1周期を判別する基準には両足首間の距離を用いる。本研究では、歩行の1周期を片方の足が前に出た時刻から再びその足が前に出るまでの間と定義する。まず右足左足間の距離の時系列

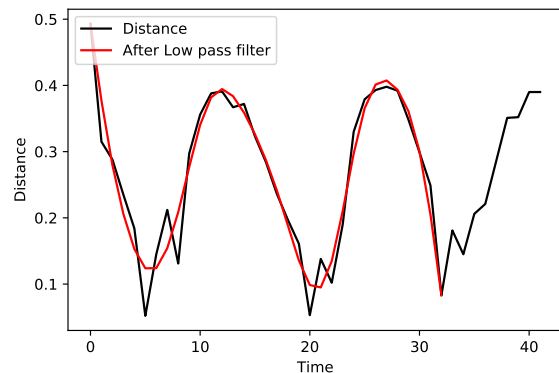


図2 両足間の距離(黒)とローパスフィルター後の距離(赤)の変化

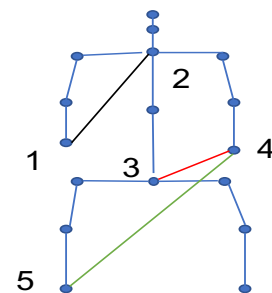


図3 特徴量の例

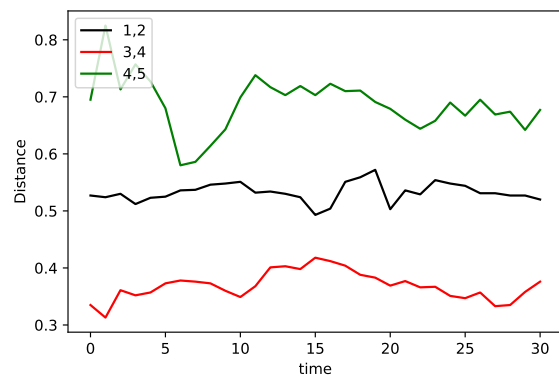


図4 特徴量の一周期間の変化

データに対して、フーリエ変換を適用し、全周波数成分の1/20の低周波数成分を残して残りを0とする。すなわち、ローパスフィルタをかけることでノイズを除去し、そのピーク間を1サイクルとする。以上の処理を行って求めたフィルタ前後の両足間距離の変化を図2に示す。時刻11から28までのピーク間を1サイクルとする。

3.4 特徴量

体の部位の3次元の時系列データを用いて、「歩きスマホ」の状態を推定するには次の問題がある。

(1) 汎用性

歩容には個人差があり、特定の誰かのパターンを細か

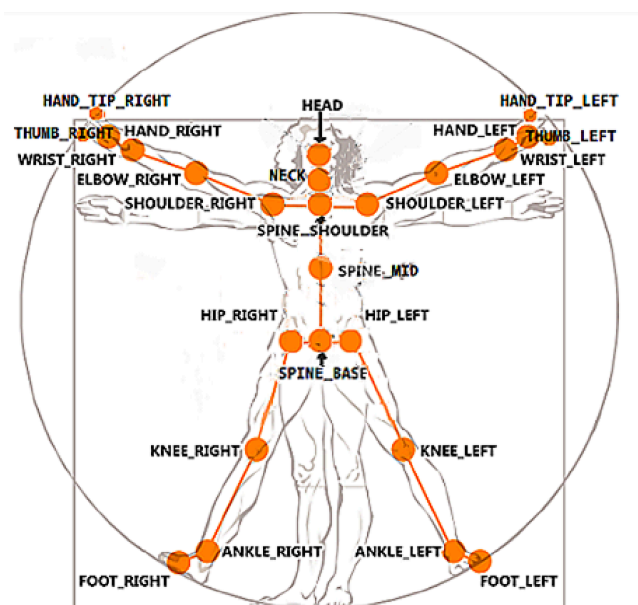


図5 Kinect v2によって認識可能な25の関節 [7]

く学習すれば他の人に適用できなくなる。不特定の被験者に対して、共通に成立する特徴を求める必要がある。

(2) 耐ノイズ性

3次元のの時系列データは動的で常に変動しており、しばしばトラッキング誤りによるノイズが生じる。これに対して頑強な推定方法が求められる。

そこで、本研究では次の方法でこの問題に対処する。

(I) 2つの関節間の相対距離

25個の関節から任意の2つを取る事で300個の特徴量を用意し、その中から最も汎用性の高い特徴量を選定する。

(II) 1周期内で統計量をとる

1周期内の時系列データを平均、標準偏差などの統計量を用いて特徴量とする事でトラッキング誤りによるノイズの影響を受けにくくする。

歩きスマホの判別に特有の最適な特徴量を見つけるため、全身25個の関節から異なる2個(i, j)を組み合わせ、関節間の距離 $\alpha_{i,j}$ を取る。データ列の、平均、中央値、標準偏差、最大値、最小値を用いて、5種類の統計量を求める。25C2 = 300組のデータについて、5種類の統計量をとるので計1500個の特徴量を用意する。

図3は関節(1,2), (3,4), (4,5)の関節を繋いだ距離であり、関節1は図5のWRIST_RIGHT, 関節2はSPINE_SHOULDER, 関節3はSPINE_BASE 関節4はWRIST_LEFT, 関節5はFOOT_RIGHTである。図4は関節間の距離 $\alpha_{1,2}$, $\alpha_{3,4}$, $\alpha_{4,5}$ の時系列データである。

3.5 クロスバリデーション

モデルの汎化性能の評価のため、本研究ではクロスバリ

表1 実験データ

被験者	通常歩行	歩きスマホ
人数 [人]	40	40
データ数	200	200

デーションを用いて適合率、再現率の計算を行う。クロスバリデーションでは200のデータを50ずつ4つのグループに分割し、それぞれのグループを一度ずつ学習データとし、それ以外のデータをテストデータとして4回検証し、適合率と再現率の平均値を算出する。

3.6 ランダムフォレスト

特徴量を基にランダムフォレストを用いて判別を行なう。ランダムフォレストはgini不純度が0.13以下になるか、1つのノードに所属するデータの個数が10個以下になるまで枝を伸ばす。ここでジニ不純度は次のように定める。

決定木のあるノード t に対し、ノード内のサンプルが n 個、ノード内のクラスが m 個のときを考える。ノード t 内で、クラス i に属するサンプルの個数を n_i とすると、クラス i に属するサンプルの割合 $p_{(it)}$ は

$$p_{(it)} = \frac{n_i}{n}$$

と与える。この時ノード t のジニ不純度 $G(t)$ は

$$G(t) = 1 - \sum_{i=1}^m p_{(it)}^2$$

で求められる分数(不純)の度合いを表す。ランダムフォレストでは特徴量毎の重要度を算出する。ある特長量 F がgini不純度をどれだけ減らしたかを表す重要度は次の様に定める。ある決定木 K_j に特徴量 F を用いて分類を行ったノード t がある時、 K_j に対する F の貢献度 $F(K_j)$ を

$$F(K_j) = W(t)G(t) - (W(t_{\text{left}})G(t_{\text{left}}) + W(t_{\text{right}})G(t_{\text{right}}))$$

と表す。

この時、 t_{left} , t_{right} はノード t によって分岐した左側、右側のノードの部分木である。また、 $W(t)$ はノード t 内のサンプル数である。また、決定木 K_j に特徴量 F を用いて分類を行ったノードがない時、 K_j に対する F の貢献度は $F_K(i) = 0$ である。

100個の決定木を作成し、それらの決定木をアンサンブル処理を経て判別を行う時、特徴量 F の重要度 I_F は

$$I_F = \frac{1}{100} \sum_{j=0}^{100} F(K_j)$$

と定める。

4. 実験

4.1 実験目的

1. 歩容を観察し、歩きスマホの検出のために有用な特徴

表2 統計量毎の再現率, 適合率, F 値

	平均値	中央値	標準偏差	最大値	最小値
適合率 P	0.973	0.971	0.864	0.962	0.945
再現率 R	0.913	0.913	0.846	0.905	0.900
F 値	0.942	0.941	0.855	0.932	0.926

表3 2組の統計量を組み合わせ特徴量に追加した時の F 値

	平均	中央値	標準偏差	最大値	最小値
平均		0.947	0.949	0.945	0.939
中央値	0.947		0.951	0.941	0.940
標準偏差	0.949	0.951		0.938	0.927
最大値	0.945	0.941	0.938		0.938
最小値	0.939	0.940	0.927	0.938	

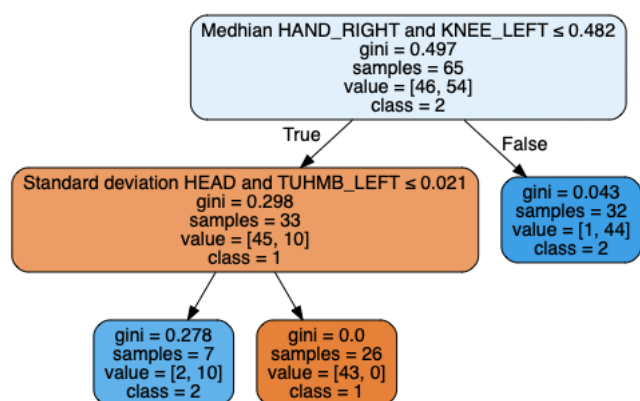


図6 決定木の例

量を明らかにする。

2. 提案手法による歩きスマホの判別精度を明らかにする。

4.2 実験環境

データ取得には Microsoft 社のモーションキャプチャデバイス, Kinect v2 を用いる。Kinect v2 には RGB カメラ, 深度センサ, マイクなどが搭載されており, 体の関節の3次元座標を推定し, その変化に基づいて体の25の関節の3次元座標を測定する。Kinect v2 によって推定できる関節を図5に示す。

4.3 データ取得

本研究では2018年7月, 明治大学中野キャンパス多目的室において, 実験協力者の同意を取って, 121名の歩容データを取得した。床から0.9mの位置に固定した Kinect v2 から5.5mの離れた位置を歩行開始地点とし, 1mの位置を歩行終了地点とする。歩容の測定は4.5m地点から2m地点までの区間で行う。121名に5回歩行してもらい, 測定を行った。121名の内訳を, 表1に示す。80名の400データを実験に用いる。

4.4 実験結果

ランダムフォレストにより作成した決定木の1つを図6

表4 重要度の高い特徴量上位10個

順位	関節の組み合わせと統計量	重要度
1	Median HAND_RIGHT and KNEE_LEFT	0.0534
2	Median FOOT_LEFT and HAND_TIP_RIGHT	0.0359
3	Median WRIST_RIGHT and ANKLE_LEFT	0.0353
4	Median ANKLE_LEFT and THUMB_RIGHT	0.0303
5	Median ANKLE_LEFT and HAND_TIP_RIGHT	0.0268
6	Median WRIST_RIGHT and FOOT_LEFT	0.0259
7	Median WRIST_RIGHT and KNEE_LEFT	0.0255
8	Median FOOT_LEFT and THUMB_RIGHT	0.0252
9	Median SPINE_BASE and HAND_TIP_RIGHT	0.0230
10	Median SPINE_BASE and HAND_RIGHT	0.0224

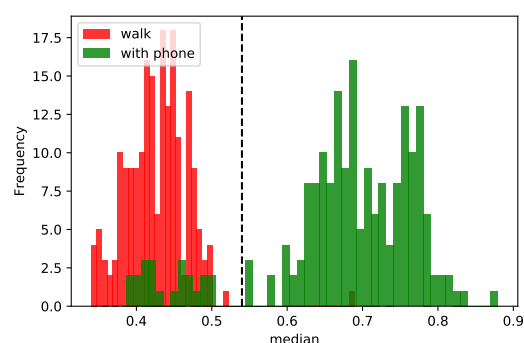


図7 右手と左膝の中央値の分布

に示す。ここで, class 2 が「歩きスマホ」, 1 が通常歩行である。gini はジニ不純度, samples はノード内の全サンプル数, value はサンプルの数である。また, class 1 の割合が増える程オレンジ色に近づく, class 2 の割合が増える程青色に近づく。

統計量毎の適合率, 再現率を表2に示す。この時, 統計量毎に300個の特徴量を判別に用いた。最もF値が高くなった統計量は平均値であり, 適合率は0.973, 再現率は0.913である。

各統計量を組み合わせたランダムフォレストによる判別の結果を表4に示す。この時, 表4の縦軸と横軸2つの統計量を用いて計600個の特徴量を判別に用いた。最もF値が高くなった統計量の組み合わせは, 中央値と標準偏差の組み合わせで適合率0.990, 再現率0.915である。この時最も重要度が高くなった特徴量は右手と左膝の中央値である。

重要度の高い特徴量の上位10個を表4に示す。最も重要度が高い右手と左膝の中央値のヒストグラムを図7に示す。図7より, 通常歩行(赤)と歩きスマホ(緑)の分布が明確に分離しているのが明らかである。

4.5 右手と左膝の中央値の分析

最も重要度が高いと判断された右手と左膝の中央値の分布を図7に表す。図7より閾値を0.54とした時に最もF値が高くなり, 適合率0.994, 再現率0.895であった。閾値

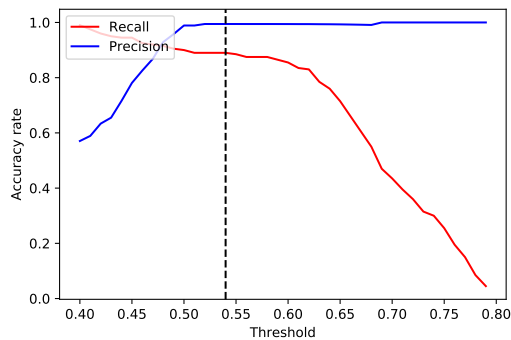


図8 閾値と適合率, 再現率
(右手と左手の中央値)

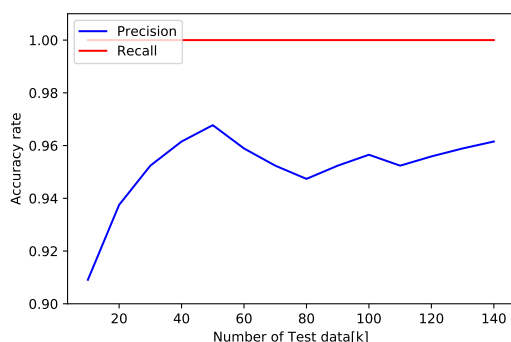


図9 テストデータの数を変動させた場合の適合率と再現率

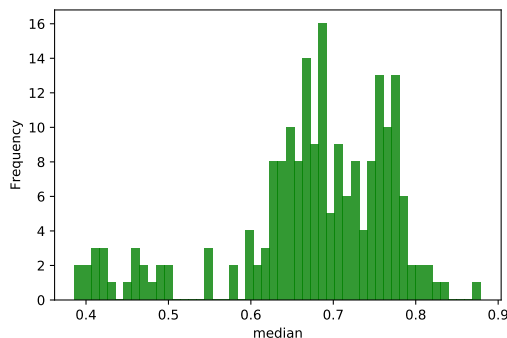


図10 歩きスマホをしている時の右手と左膝の中央値の分布

を変えた時の適合率, 再現率の変化を図8に示す。

4.6 データの増加を想定した解析

データの増加に対する再現率, 適合率の収束を調べる, ランダムに選んだ100個のデータを学習データとし, テストデータの数 k を40から280まで, 20ずつ増やしながら再現率, 適合率を測定した。結果を図9に示す。 k が増えるにつれて P と R , も増加し, $R > 120$ でほぼ収束している。テストデータの数が大きくなったときにも同様であると考えられる。

4.7 考察

表2, 4より, 特徴量単体では適合率, 再現率が $P = 0.993$, $R = 0.913$, $F = 0.942$ であったのに対し, 標準偏差と中央値を組み合わせることでF値で0.951まで上がる。その理由として平均値や中央値などの統計量では関節間の距離の静的なデータしかとることが出来ないのに対し, 標準偏差を組み合わせることで腕の振幅などの動的なデータの特徴を考慮することが出来たことが考えられる。

図7より, 多くの人が歩きスマホの検出には右手と左膝の中央値を使ってほぼ判別することが出来る。図10より, 0.5から0.6の間に掛けてデータが二分されており, 0.5以下のデータは200個中20個であった。誤差の原因は左利きの被験者のデータであると考えられ, 利き手の情報を含む特徴量の追加が必要である。

5. 終わりに

本研究では, 歩きスマホの特徴的な行動を調査し, 歩きスマホ検出器を作成し, 80名の被験者による実験を行った。本研究により次のことが示された。中央値と標準偏差を組み合わせたランダムフォレストによって適合率0.990, 再現率0.915で歩きスマホが検出可能である。歩きスマホの検出には右手と左膝の中央値が有効である。1つの特徴量で適合率0.994 再現率0.895の精度で歩きスマホが検出可能である。

参考文献

- [1] 森 駿文, 菊池 浩明, "複数の歩容特徴量の k 近傍による「歩きスマホ」にロバストな個人識別手法の提案", 暗号と情報セキュリティシンポジウム (SCIS 2019), pp. 1-7, 2019.
- [2] 加藤君丸, 渡辺裕, "姿勢情報を用いたカメラ映像からの歩きスマホ検出", 2017年度 早稲田大学大学院 修士論文.
- [3] 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-" 2016年度, 第6回視覚・聴覚支援システム (VHIS) 研究会予稿集.
- [4] 児玉翔, 榎堀優, "距離画像センサを用いた"安全歩きスマホ"システムの検討", 2016年度 情報処理学会 インタラクシオン 2016.
- [5] 小野有矢, 橋本修平, "危険な歩きスマホ抑制のためのタッチイベントに基づく集中度判定" 2019年度 情報処理学会 第81回全国大会.
- [6] 歩きスマホ等に係る事故に注意! - 東京消防庁「2020年5月4日参照」
<https://www.tfd.metro.tokyo.lg.jp/life/topics/201602/mobile.html>.
- [7] Kinect for Windows - Microsoft 「2020年5月4日参照」
<https://developer.microsoft.com/ja-jp/windows/kinect/>