

## 視覚障がい者のための歩きスマホ検出システム

常山 菜美 花泉 弘  
法政大学情報科学部

## 1. はじめに

近年デジタル化の進展に伴い、歩きスマホが社会的問題になっている。既に解決策として歩きスマホをしている人のスマートフォンのカメラから周囲の状況を確認し警告を画面上に出すシステム [2] や、センサーや画面の動きから歩きスマホを検知し画面上に警告を出すシステム [3] が提案されている。しかし、それらのシステムは歩きスマホをしている人自身がダウンロードしなくては意味がない。そこで本研究では、視覚障がい者の方が身に付ける歩きスマホ検出システムを提案する。システム対象者を被害を受ける側にするすることで、より効果的に衝突事故を防ぐことにつながるのではないかと考えている。

## 2. 原理

本研究での歩きスマホの特徴は肘の角度、顔の向きである。この2つの特徴を満たしたとき、かつ一定距離近づいてきたときに衝突回避アラートを出す。図1のように、首元にステレオカメラを設置し近づいてくる人を正面から撮影することを想定する。

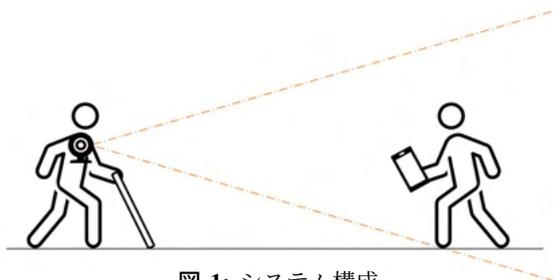


図1: システム構成

OpenPose によって撮影した動画の姿勢推定を行い、得られた肩、肘、手首の3つの関節点座標から肘の角度を算出する。肘  $(x_0, y_0)$ 、肩  $(x_1, y_1)$ 、手首  $(x_2, y_2)$  とすると、 $\vec{a} = (x_1 - x_0, y_1 - y_0)$ 、 $\vec{b} = (x_2 - x_0, y_2 - y_0)$  となる。よって、式1で肘の角度を算出することができる。

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|} \quad (1)$$

2つ目の特徴である顔の向きは、InsightFace によって判定する。InsightFace で推定された回転行列  $R$  からオイラー角度 Pitch を式(2)を用いて求め、顔の向きの特徴として得る。

$$\text{Pitch} = \tan^{-1}(r_{21}/r_{11}) \quad (2)$$

これらの方法で、歩きスマホ、正面を見て歩いているとき、荷物を持って歩いているときの肘の角度と顔の向き

を算出する。算出した2つの特徴の値を式(3)の  $x$ (肘の角度)、 $y$ (顔の向き)に代入し、重回帰分析によって定数項  $a_0$ 、肘の角度の係数  $a_1$  と顔の向きの係数  $a_2$  を求め、一つの平面の式を算出する。 $z = 1$  は歩きスマホ、 $z = 0$  は正面を見て歩いていると判別する。

$$z = a_0 + a_1x + a_2y \quad (3)$$

$z = 1$  のとき歩きスマホと認識し、かつ深度推定による距離が3mより近づいたときアラートを発する。深度推定の手順は次のとおりである。キャリブレーションによってステレオカメラの左右の画像をマッチングし平行化する。カメラ座標  $(u, v)$  とワールド座標系  $(X, Y, Z)$  の関係は、画像座標へと変換する内部パラメータ  $A$  を用いて、式(4)で表すことができる。 $f_x, f_y$  は  $x$  軸、 $y$  軸方向の焦点距離、 $(c_x, c_y)$  は画像面と光軸の交点である。

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A(R|t) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

ステレオカメラの左画像を  $(u_l, v_l)$ 、右画像を  $(u_r, v_r)$ 、左右カメラ間の距離  $B$ 、焦点距離  $f$  とすると、式(5)を用いて2次元座標からワールド座標系  $(X, Y, Z)$  に変換することができ、深度  $Z$  を得ることができる。深度推定の対応点は首の関節点座標とする。

$$X = \frac{Bu_2}{u_1 - u_2}, \quad Y = \frac{Bv_2}{u_1 - u_2}, \quad Z = \frac{Bf}{u_l - u_r} \quad (5)$$

## 3. 実験と結果

ステレオカメラで10m離れたところから撮影を行った。被験者は20人であり、スマートフォンを右手、左手、両手に持ち歩きスマホをしているときの動画を2回ずつ(計120動画)、正面を向いて歩いているときの動画を2回(計40動画)、荷物を持って正面を見て歩いているときの動画を2回(計40動画)を、カメラを固定した条件下と撮影者の首元に設置した条件下で全400の動画データを撮影した。撮影者の身長は156cmであり、被験者らとの身長差は-5cm~+30cmである。出力結果の例を図2に示す。



(a) 歩きスマホ

(b) 正面を見て歩行

図2: Open Pose の出力結果の例

A system to detect an approaching smartphone zombie for visually impaired person

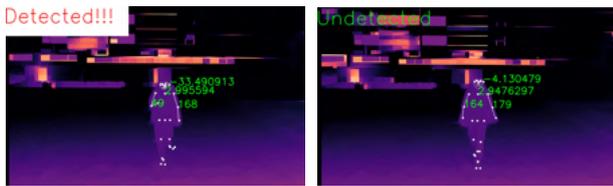
†NamiTsuneyama, †HiroshiHanaizumi

Faculty of Computer and Information Science, Hosei University

10人の歩きスマホと正面を見て歩いているとき、荷物を持って歩いているときのデータ計30個の動画データのうち、ランダムに80%のトレーニングデータ(計24個)を選び重回帰分析を行った結果、式(6)の係数を得た。

$$z = 0.211 - 0.002x - 0.026y \quad (6)$$

実際に歩きスマホ検出した結果は図3aのとおりである。正面を見て歩いている場合、一定距離近づいてきても図3bのように歩きスマホを検出なかったため、正確な判定結果が出力されている。左上のDetectionの表示は歩きスマホを検出したことを示している。体についている白い点は、OpenPoseで得られた関節点をプロットしており、腕の白い線は肩、肘、手首を結んだ線である。首元にある数値は距離、肘の横にある数値は肘の角度、そして顔の横にある数値は顔の向き(Pitch)である。



(a) 歩きスマホ (b) 正面を見て歩行

図3: 歩きスマホ検出結果

#### 4. 考察

トレーニングデータを用いて求めた式6のモデルの性能をROC曲線を用いて評価を行った結果、図4のようになった。ROC曲線は真と偽に分類する際の閾値を変更させ、真陽性率TPRと偽陽性率FPRの値をプロットしたものである。真陽性TP、真陰性TN、偽陽性FP、偽陰性FNを用いて、TPRとFPRはそれぞれ式(7)で算出することができる。それぞれの閾値の点と座標軸が作る短形の面積を比較した結果、図7に示した矢印の閾値の面積が最大であるため、最も性能がいい閾値の値は0.425であるとわかった。

$$FPR = \frac{FP}{FP+TN}, \quad TPR = \frac{TP}{TP+FN} \quad (7)$$

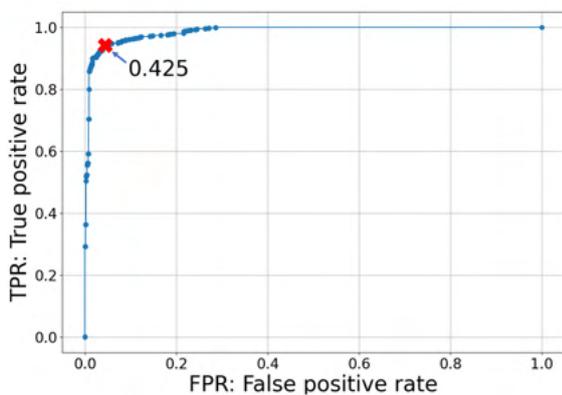
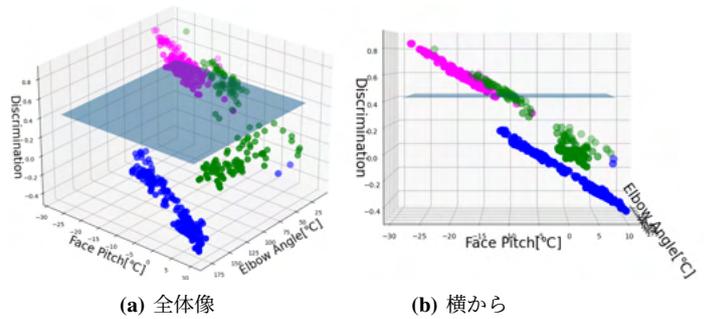


図4: ROC曲線

得られた最適な閾値で平面を引き、テストデータをプロットした結果、図5のようになった。ピンクが歩きスマホ、青が正面を見て歩いている、緑が荷物を持って歩いているデータである。また、テストデータの分類についての混同行列は表1のようになった。



(a) 全体像 (b) 横から

図5: テストデータプロット図

表1: 混同行列(テストデータ)

|     |   | 予測値  |     |
|-----|---|------|-----|
|     |   | 0    | 1   |
| 実測値 | 0 | 1148 | 52  |
|     | 1 | 33   | 533 |

図5から、緑点で示す「荷物を持って歩いている」データの一部が誤検出されてしまっていることがわかる。原因はマスクによって顔の向きが誤判定されてしまっていることが考えられる。荷物のデータがどれだけ分類に影響を与えているかを評価するため、表1の混同行列の値を式(8)に代入したところ正解率0.952、適合率0.911、再現率0.942を得た。このことから、本手法によって高精度で歩きスマホの分類を行うことができることがわかった。

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}, \quad (8)$$

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}$$

#### 5. おわりに

従来の歩きスマホシステムは歩きスマホをしている人自身が危険を回避するためのものであったが、本研究では視覚障がい者のための歩きスマホ検出システムを提案し、効果的な衝突事故削減へとつながる高精度なシステムを開発することができた。今後の課題は、混雑しているところでも歩きスマホ検出ができるようにすることである。人の進路を推定することでどのタイミングで衝突するか推定することができるのではないかと考えている。

#### 参考文献

- [1] KODAMA SHO, ENOKIBORI YU, MASE KENJI. Examination of Safe-Walking Support System for "Texting While Walking" Using a Time-of-Flight Range Image Sensor, IPSJ SIG Technical Report, Vol.2016-UBI-50 No.3, pp487-492
- [2] Tianyu Wang, Giuseppe Cardone, Antonio Corradi, Lorenzo Torrè-sani, Andrew T.Campbell. WalkSafe: A Pedestrian Safety App for Mobile Phone Users Who Walk and Talk While Crossing Roads, University of Bologna, February 2012
- [3] Yu-Chin Tung, Kang G.Shin. Use of Phone Sensors to Enhance Distracted Pedestrians' Safety, IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL.17, No.6, JUNE 2018
- [4] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh. OpenPose: Real-time multi-person 2D Pose Estimation Using Part Affinity Fields, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL.43, NO.1, JANUARY 2021
- [5] Jia Guo, Jiankang Deng, Alexandros Lattas, Stefanos Zafeiriou. Sample and Computation Redistribution for Efficient Face Detection, arXiv:2105.04714v1, May 2021