

## スマートフォンを使った対人距離測定システムの実装と評価 Implementation and Evaluation of an Interpersonal Distance Measurement System Using Smartphones

吉江 治紀<sup>†</sup>  
Harunori Yoshie

田村 仁<sup>‡</sup>  
Hitoshi Tamura

### 1. はじめに

#### 1.1 研究背景

2020 年春に確認されて以来、現在に至るまで新型コロナウイルス(COVID-19)による被害は拡大を続けており世界各地で甚大な被害を生んでいる。

そういった数々の対策のなかでも三密の回避、特にソーシャルディスタンスの維持が公共機関から民間、個人に至るまで意識されている対策である。具体的な対策として飲食店や劇場などでは座席の制限や指定などで利用客間の距離の維持を行えているが、博物館や庭園、イベント会場といった利用客が一つの場所に留まらないような施設では一度の入場者数を制限し施設内では利用者の自己判断での距離の維持に任せるに留まっている。そこで個人が各自で身を守るための個人での利用が可能な周囲の群集の密集度を判定し過密状態を警告するシステムが有用であると考えた。また個人が使用可能なデバイスの候補は複数挙げられるがコロナ対策であることから非接触式であることが望ましい。対象から距離を取って測定を行う手法として距離センサを使用するといった手法やステレオカメラの使用が考えられる。そあそ Rui Pascoal らの[1]の調査によると調査対象者の 92%は新規の AR を使用したシステムに求める要素としてスマートフォンで動作することを望んでいるという結果からわかる通り新規のデバイスを使用したシステムは一般には普及させることは困難な事からすでに一般的に普及しているスマートフォンで動作することが望ましいため今回はスマートフォンのみで動作する手法を考案する。

#### 1.2 関連研究

カメラ映像から群集の密集度を計測する技術の前例として米司健一らの[2]がある。これは施設モニタカメラのカメラ映像を解析し施設内の混雑度や人流を計測するシステムである。このシステムは映像を用いた施設内の人流計測技術つまり映像中の動き情報を解析することで、映像内の人数や移動速度、移動方向などを抽出する技術を元としている。この研究に使用される人流計測技術は映像中から Tracklet と呼ばれる細かい動き情報(図 1)を抽出し、抽出した動き情報を解析することで、流動を計測する手法が使用されている。

これらの情報を組み合わせ混雑情報を可視化したアプリケーションが図 2 である。(a) 通過人数、(b) 滞留人数、(c) 混雑度割合グラフ、(d) 人物アイコン化画像、(e) 混雑度ヒートマップの 5 つの手法により混雑度を可視化さ



図 1 Tracklet の例 ( 図中の線が Tracklet を示す )  
(引用:[1])

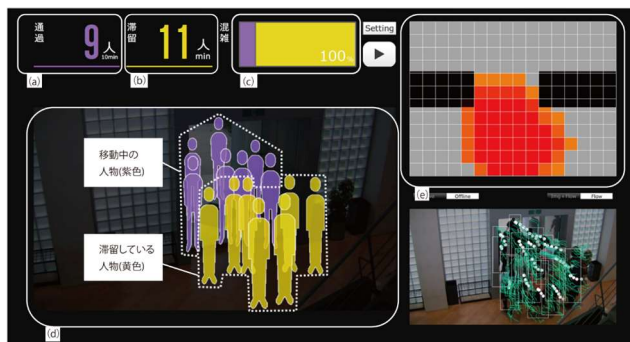


図 2 混雑度可視化アプリケーション  
(引用:[1])

#### 1.3 研究目的

本研究は、屋内屋外問わず、イベント会場等で直立した状態の集団が滞留する状態での運用を想定し、対象となる人物から人物の距離を計測することで集団の密集度を判断し十分な距離の確保を行うよう注意喚起する距離計測システムの作成を目的とする。また不特定多数の人々が手軽に利用可能な媒体としてスマートフォンのみで一連の計測が可

<sup>†</sup> 日本工業大学 機械システム工学専攻,  
Nippon Institute of Technology Department of Mechanical Systems Engineering

<sup>‡</sup> 日本工業大学 先進工学部 ロボティクス学科,  
Nippon Institute of Technology Department of Advanced Engineering Faculty of Robotic

能なシステムとして開発する。このシステムの作成にあたり必要となる要素として

- (1) 対象人物の立ち位置の検出
- (2) 対象人物までの距離の測定の 2 点が必要となる。

## 2. 提案手法

### 2.1 対象人物の検出

本システムを実装するにあたり必要となる要素として対象人物の立ち位置の検出がある。この機能を実装する方法として第一フェイストラッキングを使用することを考えた。

しかし今回の想定される運用として美術館やイベント会場等で展示物等を鑑賞している観客を対象としているため対象人物の正面へと移動する必要がある顔認識では不適切である。

そこで対象者の向きや姿勢に影響を受けにくい人物の検出方法として姿勢推定を行うことを考えた。現在姿勢推定のシステムとして主流なシステムとして OpenPose[3]と PoseNet[4]が存在する。OpenPose はカーネギーメロン大学 (CMU) の Zhe Cao らが発表した深層学習を用いて人物のポーズを可視化してくれる手法であり単一の画像上で人体、手、顔面のキーポイント (全部で 130 個のキーポイント) を共同で検出することが可能である。この OpenPose は静止画を入力するだけで人間の関節点を検出することが可能であり動画内に複数人の人物がいる場合でもリアルタイムに検出することも可能(図 3)である。



図 3 OpenPose での姿勢検出  
(引用:[3])

このように複数人の対象の姿勢を検出することが可能でありカメラ画像のみで推定を行うことが可能である。しかし問題点としてこのシステムを動作させる推奨されているスペックがスマートフォンで動作させることを想定すると少々高い。

もう一つの候補である PoseNet は Google から発表されている機械学習向けに開発されたエンドツーエンドのオープンソースプラットフォームである TensorFlow にて発表されているブラウザ・node.js で機械学習のトレーニングおよび実行が出来るライブラリである TensorFlow.js を使用した WEB カメラなどの入力からブラウザ上でリアルタイムにユーザーの姿勢検出ができるシステムである。OpenPose と同様、人の関節や目、鼻、耳を推論するという仕組みであるがこちらは 17 か所のキーポイントを推論しており図 4 のように姿勢の推定が可能である。上記の通りブラウザ上で動作するという点と OpenPose と比較して動作が軽量なことから今回の重要な運用想定としてのスマートフォン

で容易に運用が可能という点を満たしている。このことから人物の検出には PoseNet を使用する。

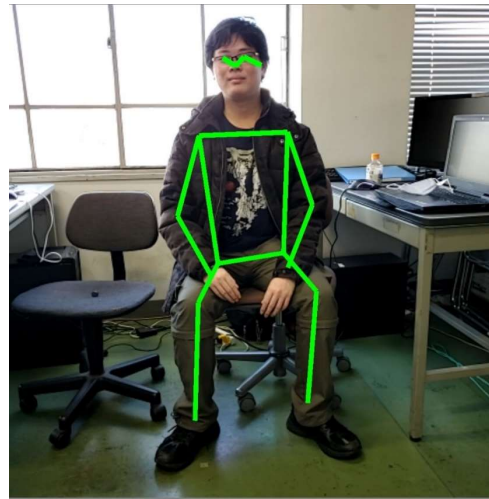


図 4 PoseNet での姿勢推定  
(引用:[5])

### 2.2 対象人物までの距離の測定

本システムを実装するにあたり必要となる対象人物までの距離の測定を行う手法として周辺環境の 3D スキャンを考えた。

スマートフォンを使用した周辺環境の 3D スキャンの方法として Google から発表されていた AR プラットフォームである Project Tango を使用した[6]が存在する。こちらでは図 5 のように部屋全体のスキャンを行い 3D データとして生成が可能である。しかし Tango を利用可能なスマートフォンは深度センサを搭載した機種のみで極めて限られる。また Tango の正式なサポートもすでに終了していることから簡易的に利用可能という要件を満たさない。

そこで今回は同じ Google から発表されている ARCore[7]を使用する。まずスマートフォンに内蔵された各種センサとカメラのみで周辺環境の 3D スキャンを行うには使用機材の自己位置推定と周辺の物体の特徴点の抽出の 2 点が必要となる。そして今回使用する ARCore はカメラの画像を分析して、物体の表面とする場所にマーキングし、座標を使って三次元上での物体の位置を把握するモーショントラッキングとカメラ映像と IMU センサーを使用して水平面を検出する機能を有しており必要となる要件を同時に満たしている。また ARCore はスマートフォンの内蔵カメラやモーションセンサーのみの使用で上記の機能を使用することが可能で対応機種が極めて多いことも今回重要視する容易に利用可能であるという点を満たす。



図 5 生成された 3D データ  
(引用:[6])

### 2.3 実装方法

本研究では、屋内屋外問わず、イベント会場等で直立した状態の集団が滞留する状態での運用を想定し、対象となる人物から人物の距離を計測することで集団の密集度を判断し十分な距離の確保を行うよう注意喚起することを目的とする距離計測システムを提案する。また不特定多数の人々が手軽に利用可能な媒体としてスマートフォンのみで一連の計測が可能なシステムとして開発する。

ARCore で現実紐付いた三次元情報を持つ特徴点の集まりである点群(point cloud)を取得し得られたデータから周囲の平面といった三次元情報の認識を行い対象人物を PoseNet で検出し ARCore で認識した三次元情報を元に対象人物の位置を判定する。ここで対象人物との距離を測定するための基準点としては PoseNet は上半身の姿勢推定が下半身よりも精度が優れているため頭部や肩部を基準とするのが望ましいが ARCore での距離の測定は検出した平面を基準として距離を測定するため基本的な基準となる地面から測定対象が離れるほど精度が低下するため頭部から脚部までの中でもっとも精度が高くなる部位、または測定された距離を平均した数値を対人距離として使用することが考えられる。

### 3. 実験

制作したアプリケーションを使用しての距離の測定が実用的か確認するために以下の実験を行った。

1. PoseNet で推定した立ち位置の精度の確認
2. ARCore を使用した距離測定の精度の確認
3. 方向が ARCore での距離測定の精度に与える影響の確認
4. 背景が ARCore での距離測定の精度に与える影響の確認

#### 3.1 実験 1

PoseNet で推定した立ち位置の精度の確認のために対象人物の頭部直下の地面をその人物の立ち位置として PoseNet で推定された両足首を基準とする立ち位置との誤差を計測を行った。また同時に対象人物の服装が精度にどのような影響を与えるか確認するため異なる服装の人物(図 6)で同様の検証を行った。

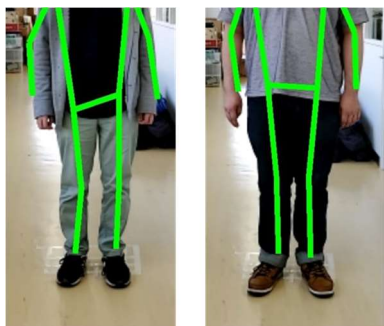


図 6 測定対象

#### 3.2 実験 2

ARCore を使用した距離測定が実用的な精度を持つか否かを確認するために ARCore を元で作成した周囲の平面を認識し指定した点から点までの距離を計測するシステム

(図 7)を制作し測定を行った。制作したシステムを使用して直径 300[mm]の対象物を地面に設置しその対象物からスマートフォンを 1[m]刻みで 1[m]から 5[m]まで距離を離しながら 1 の計測を 10 回行った。



図 7 計測画面  
(引用:[5])

#### 3.3 実験 3

方向が ARCore での距離測定の精度に与える影響の確認のために ARCore を用いて指定した基準点から

1. 左側 1000(mm)の地点までの距離
2. 右側 500(mm)手前方向 1000(mm)移動した地点までの距離

を計 10 回測定し実際の距離との誤差を計測する。

#### 3.4 実験 4

実験 2 とは異なる床面(図 8)で同様の実験を行い精度に差が生じるか検証した。実験 2 での床面よりも特徴となる模様が少ない地点で実験を行った。



図 8 測定に使用した床面

## 4. 結果

### 4.1 実験 1

脚部の輪郭線が比較的鮮明な対象人物の計測結果の誤差平均が 187.8 (mm).

脚部の輪郭線が比較的不鮮明な対象人物の計測結果の誤差平均が 289.4(mm)となった。

## 4.2 実験2

計測した結果から対象物との距離と誤差に正の相関関係があることがわかった。この結果を元にした近似直線を図8に示す。対象物から1[m]離れるごとに約9[mm]の誤差が生じることがわかった。

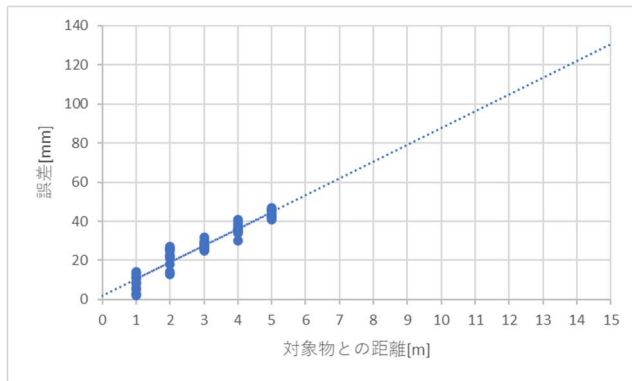


図 9 対象物との距離と誤差の近似直線

## 4.3 実験3

測定を行った所結果は

1. 左に 1000(mm)の地点までの距離の計測で発生した誤差の平均が 76.2(mm)・計 7%の誤差
2. 左に 1000(mm)の地点までの距離の計測で発生した誤差の平均が 76.2(mm)・7%の誤差となった。また全ての計測で本来より数値が大きくなった

## 4.4 実験4

双方の床面の結果を比較すると後者の特徴の少ない床面の誤差の平均が前者と比較して 19.3%だけ増加した

## 5. 考察

実験 1 にて脚部の輪郭線が比較的鮮明な対象人物でも約 200(mm)もの誤差が発生した主要原因として PoseNet で推定された立ち位置は両足首の間が基準であるので踵の高さの分だけ本来の立ち位置からズレが生じるためであり姿勢推定によって推定される対象者の身長に合わせて足首下に線を延長することで解決が可能だと考えられる。

実験 2 にて一般的にソーシャルディスタンスとして必要な距離は 3[m]とされている。このことから測定対象の人物とは 3[m]以上の距離をとった状態での運用が想定される。そのため最低でも 30[mm]の誤差が生じるが対象が周囲と 3[m]以上の距離を保っているかを判別することを目的としているため誤差として許容可能な範囲だと考える。

また使用する媒体がスマートフォンであるため対象との距離がある場合でも適切な距離への移動が容易な点からも今回の実験で判明した誤差が許容範囲だとする理由である。

実験 3,4 にて発生した誤差は運用想定では使用機材がスマートフォンであり対象人物と測定者が共に頻繁に移動することから許容範囲だがコロナ対策としては距離が実際の距離よりも離れて表示されることは問題である。

## 6. おわりに

本稿では、対象となる人物から人物の距離を計測することで集団の密集度を判断し十分な距離の確保を行うよう注意喚起する距離計測システムの作成を目的とし、実験から実用可能な精度であることは確認できた。

また今回は両足首の間を立ち位置としたが感染防止の観点からは頭部の位置を基準とするほうが望ましいため人物の位置の判定に使用する部位の選定を今一度行う必要がある。今後の予定として、複数人の人物を対象とした精度と処理速度の検証を行う。また群集の密集度を判断するための基準を設定し当システムの完成を目指す。

### 参考文献

- [1] Rui Pascoal; Bráulio Alturas; Ana de Almeida, Rute Sofia "A survey of augmented reality: Making technology acceptable in outdoor environments", 2018 13th Iberian Conference on Information Systems and Technologies (2018).
- [2] 米司 健一,松原 大輔,秋良 直人,飯田 都,村上 智一,額賀 信尾"駅構内モニタカメラを用いた混雑度可視化技術"デジタルプラクティス 8(2), 152-159(2017)
- [3] 「CMU-Perceptual-Computing-Lab/openpose」  
<https://github.com/CMU-Perceptual-Computing-Lab/openpose> (参照 2021-1-21)
- [4] 「Pose estimation | TensorFlow Lite」  
[https://www.tensorflow.org/lite/models/pose\\_estimation/overview](https://www.tensorflow.org/lite/models/pose_estimation/overview) (参照 2020-10-16)
- [5] 吉江 治紀,田村 仁"スマートフォンを用いた対人距離の測定"情報処理学会 第 83 回全国大会講演論文集 Vol.2 No.89-90(2021)
- [6] Abdoulaye A. Diakite, Sisi Zlatanova "FIRST EXPERIMENTS WITH THE TANGO TABLET FOR INDOOR SCANNING", ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-4, (2016)
- [7] 「ARCore - Google Developers」  
<<https://developers.google.com/ar>> (参照 2020-7-10)