

5G 網を介した動画像と加速度センサによる人物認証

伊藤鋼一郎[†] 西村 優典[†] 横山 昌平^{†,‡,‡‡}

[†] 東京都立大学 〒191-0065 東京都日野市旭が丘 6-6

[‡] 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

^{‡‡} 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: [†]{ito-koichiro,nishimura-yusuke}@ed.tmu.ac.jp, [‡]shohei@tmu.ac.jp

あらまし 5G 通信が普及しており、低遅延かつ超高速な通信が可能となっている。本手法ではこの 5G 網を利用し、動画像の座標の変化と歩行者の持つスマートフォンの加速度センサデータの変化を統合することによって加速度センサ所持者を特定する手法を提案する。本手法ではサーバに加速度データを送る際に 5G 通信を利用する。また、5G 通信で超高速でデータを送ることができるため、スマートフォンから加速度のデータをサーバに送るだけで人物認証を行うことが可能であり、ユーザは顔や歩容のデータなどの身体情報を事前に登録する必要がない。

キーワード 画像認識, 映像解析, 信号処理

1 はじめに

我々の生活の安全・安心を実現するために人物認証が様々な方法によって行われている。一番身近なところでいうと、スマートフォンのロック画面解除である。パスワードだけでなく、指紋や顔など様々な手法でロック画面を解除することができる。自身のデバイスといった私的な用途だけでなく、空港でのパスポートとの顔照合、オンライン決済時の本人確認など公共の場においても人物認証は様々な方法かつ様々な目的で行われている。

人物認証は我々の生活で頻繁に行われているが、人物認証における認証手法は大別してアクティブ認証とパッシブ認証の 2 つがある。アクティブ認証とはパスワードを入力する等、ユーザの能動的な動作に基づいて認証を行う手法である。それに対しパッシブ認証は IP アドレス・MAC アドレス認証であったり、筆跡鑑定等を用いた認証のような、ユーザ自身に特別な動作を課さない認証の事である。さらに顔や指紋、歩容情報など身体的・行動的特徴を用いた認証手法は生体認証と位置付けられる。

パッシブ認証かつ生体認証の人物認証である歩容認証 [1] は、基本的に個人を識別することを目的としており、データベース内に格納された歩容情報と、観測情報を統合することによって認証を行う。データベースの情報と観測情報を統合するという点は指紋認証や顔認証も同様である。ここで我々は事前に歩容情報や指紋情報などの生体情報を保持しておくことにセキュリティ上の危険が存在していると考え、データベースからの生体情報の抜き取り、偽造された生体情報を用いて認証を突破 [2]、といったように多方面からのセキュリティ上の危険が存在する。また、ユーザがデータベースに情報を保持することに対する心理障壁も考えられる。

前述した人物認証の例では、顔認証や指紋認証のように特定の個人を識別することを例として挙げた。人物認証はその人物の 2 つ以上のデータを照らし合わせることでその人物を特定

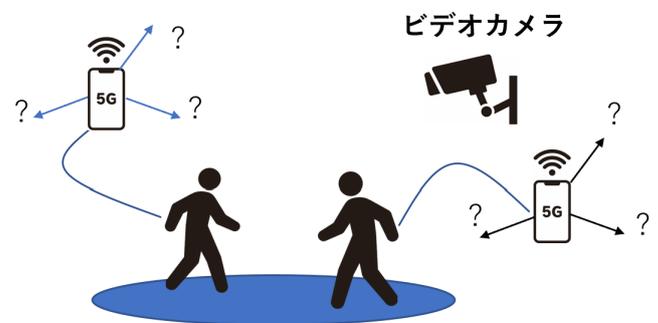


図 1 歩行者をカメラで捉え、加速度を取得するイメージ

する。つまり、その人物に関する情報が 2 つ以上あればそれらを照合できれば人物認証ができる。そして本論文が想定している環境は現実世界の物理的な動きを仮想世界に反映するアプリケーションにおいて個人の認証を行うことを想定している。ここでそのような環境下で認証を行うにあたって、個人の生体情報を用いなければならないのだろうか。個人の生体情報ではない情報を取得する方法として、スマートフォン内蔵の加速度センサとカメラ視野内の人物に関する骨格検出を利用する。図 1 は、カメラで捉えている歩行者の動きを仮想世界に反映するアプリケーションで加速度と骨格検出を用い、人物認証を行うイメージである。

スマートフォンは今や世界全体において広く普及したデバイスとなっており、街中では歩いていると多くの人がスマートフォンを操作している様子が見受けられる。スマートフォンには加速度センサが内蔵されており、加速度センサによって物体の傾きや振動を計測することができる。そして街中では歩きながらスマートフォンを操作したり、スマートフォンを手を持ったまま歩くと見受けられる。このことから歩行者がスマートフォ

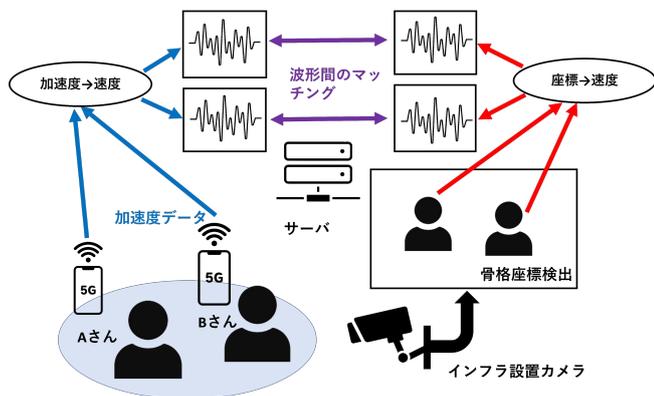


図 2 加速度データと骨格検出データの統合

ンを持つ持ち方は多種多様であり、その腕の振り方についても多種多様である。そうすると、加速度センサの示す値についても歩行者によって異なることが分かる。

本論文では生体情報を事前に保持する危険性を背景に着目し、ユーザがスマートフォンを持つだけで、そのユーザと持っているスマートフォンのペアリングを行う手法を提案する。具体的には手を振るなどの、何か動作をしているユーザが手に持っているスマートフォンに内蔵されている加速度センサのデータを、5G 通信網を用いた高速通信によってサーバに送信する。そしてそのユーザの手の動きを骨格検出によりトラッキングし、動画像上での手の変位情報を取得する。この骨格検出データと加速度データを統合させることによってペアリングを完結させる。提案手法では人物認証という目的を達成するために、あくまでユーザはカメラ視野内でスマートフォンを持てばよいだけであり、生体情報を用いずに人物認証を行うことが可能である。次の図 2 に加速度データと骨格検出データに統合を行うイメージを示す。

また 5G 通信が普及が始まっており、低遅延、多数同時接続での通信環境がスマートフォンでも利用できるようになりつつある。本手法では 5G 通信を用いることで加速度センサの値を低遅延で送ることが可能となっている。本論文では図 1 のようにカメラの視野内で複数の人が歩いているような環境下で、その歩行者 1 人 1 人がスマートフォンを持ち、それらの加速度の値とビデオカメラからの歩行者の運動情報を紐づけ、オンデマンドでどの歩行者の持っているスマートフォンかを判別する。また本論文では実現していないが、5G 通信を用いてビデオカメラのリアルタイム映像をサーバに送ることで、リアルタイムで加速度センサと情報を統合しリアルタイムで人物認証を行う研究も進めている。

2 関連研究

加速度センサを用いた人物認証についての研究については既にいくつか提案手法が存在する。

繁田ら [3] [4] は実際に加速度センサとカメラを用いて人物認証を行っている。繁田らは加速度センサが実際にスマートフォ

ンに付いた想定で実験を行っている。つまり、加速度センサのみの機能を持つ機器を使用しており得られた加速度の値はサーバとの無線通信によって送信される。また映像はサーバに USB でビデオカメラを接続し取得している。カメラと加速度センサの情報の対応付けをでは加速度のノルムの正規化相互相関を計算している。また、[3] は [4] とカメラと加速度センサの対応付けは同じ条件でありながらも、加速度センサの場所を手だけに限定せず、ズボンのポケットの中や、バッグの中に入れ加速度センサを持つ人物が特定する手法を提案している。

これら研究と異なる点は、本手法では加速度センサを手を持って人物認証を行う点である。これは [4] と条件は同じであるが、近年では繁田らの研究していた時代と異なり、スマートフォンが普及している。そして、スマートフォンを手を持ちながら歩く人が多いという近年の歩行者の特徴を活かしている。また、本手法では加速度センサではなく、スマートフォンを用いて実験を行っており、5G 通信を使用しサーバと通信を行う点も異なる点である。[3] [4] では加速度センサのデータを無線で送っていたが、本手法ではスマホから 5G 通信を介してサーバに加速度のデータを送り、サーバ上で加速度のデータを処理する。

また本論文ではビデオカメラからの情報と加速度のデータの処理、統合をオンデマンドで行っている。我々の今後想定している完全リアルタイムでの人物認証が実現した際には、ビデオカメラの映像、加速度のデータの送信を全て 5G 通信網上で行うことができる。そのため、関連研究のような USB 接続によるビデオカメラからの映像の取得、無線による加速度の値の送信から発展し完全オンライン上での処理が可能となる。

長友ら [5] はカメラと加速度センサを用いたデバイスペアリングの手法を提案している。こちらの手法では加速度のデータとカメラからのデータの統合において、加速度、座標ともに速度データに変換させ、その速度データの類似度を計算し、デバイスペアリングを行っている。加速度のデータを用いることで受信信号強度や赤外線といった無線通信を用いたデバイスペアリングと比べ、頑強かつ高精度なペアリングを提案している。

本手法と異なっている点は、まず、人物認証ではなくデバイスペアリングを行っていることが挙げられる。そのため、加速度のデータと統合されるデータはトラッキングした人の動きではなく、実験者が開いているスマートフォンの画面に表示されているマーカーの 4 つ角の座標変化となっている。次に加速度の測定開始と測定終了を手動で行っている点である。本手法では加速度はスマートフォンをサーバに立ち上げたサイトに接続すると、サイトを閉じない限り永遠に送り続ける仕組みとなっており、スマートフォンを持った歩行者がカメラの視野内に入るとトラッキングと人物認証を行う。

3 加速度センサと骨格検出を用いたペアリング方法の提案

3.1 加速度送受信のアプリケーション

まず加速度送信のアプリケーションについて説明する。スマートフォンには x, y, z 方向の 3 方向加速度及び方位角, x, y

表 1 配列に格納された関節の種類

0	鼻	13	左膝
1	首	14	左足首
2	右肩	15	右目
3	右肘	16	左目
4	右手首	17	右耳
5	左肩	18	左耳
6	左肘	19	左足親指
7	左手首	20	左小指
8	左手首	21	左かかと
9	右腰	22	右親指
10	右膝	23	右小指
11	右足首	24	右かかと
12	左腰		

,z 軸の 3 軸回転角を取得するセンサが搭載されている。本手法では動画像とのデータの対応付けを行えばよいため、3 軸回転角に関するデータは取得していない。またカメラからスマートフォンまでの奥行きが測定できないため、加速度の z 方向に関して比較する手段がない。そのため、z 方向に関する加速度も取得していない。

実装には JavaScript と Python を使用した。JavaScript の Devicemotion イベントでスマートフォンの加速度センサの値が変化した際に 3 方向加速度の値を取得することができる。この処理が内部で行われる Web ページにスマートフォンから接続することによってサーバとの加速度の送受信を行う。加速度を測定する web ページを開くと加速度の取得が開始され、同時にサーバとの WebSocket の接続が開始される。加速度の値が変化すると同時に WebSocket を用いてクライアントであるスマートフォンから Python で立てたサーバに送信する仕組みになっている。

加速度センサからのデータはスマートフォンのユーザーごとに別々の csv ファイルに保存される。

3.2 スマートフォンを持つ手のトラッキング

次にスマートフォンを持っている手のトラッキングについて説明する。手のトラッキングをするためにオープンソースのプログラムである、OpenPose¹を用いることを想定する。

OpenPose は人の関節を検出することが可能である。OpenPose によって検出することができる関節は 25 種類であり、検出された関節座標は配列として格納される。表 2 に格納された関節を示す。

スマートフォンは手で持っているが、検出点では手首が一番近い。そのため検出された 25 種類の関節点のうち、スマートフォンを持っている腕の手首の関節についての座標を手の座標とする。そしてその手の座標を 1 フレーム度に取得し手の変位とする。

3.3 加速度データとトラッキングデータからの端末判別

上述したようにして得られた加速度センサのデータと人物の

手首の座標データを統合することによって端末を判別する。その手順を以下に示す。

[step1] スマートフォンを振る動作を撮影し加速度データを取得
加速度データを受信するサーバを立ち上げる。そしてカメラで録画を開始し、スマートフォンからサーバに接続し加速度の送信を開始する。その後スマホを振り 10 秒ほどたってからサーバとの接続を解除し録画を止め、加速度データと動画を取得する。

[step2] 骨格検知を行い、動画内の手の座標を取得

撮影した動画を骨格検知にかけ、手首の座標を取得する。手首の座標に関しても加速度と同じように csv ファイルで取得する。また、動画のフレームレートは 30fps に設定しており、骨格の座標は 1 フレームごとに検出される。つまり、1 フレーム \equiv 0.033 秒ごとの骨格の座標を検出することができる。

[step3] 加速度データのフィルタリング

加速度のデータにのみローパスフィルタをかけノイズを除去する。

[step4] 加速度データのサンプリング

加速度送信の際は加速度センサの値が変化する度にサーバに加速度を送信するようになっている。そのため動画から得られる人物の座標のデータと時間間隔が異なる。時間間隔を揃えるために加速度のデータを動画のフレームレートと同じ周波数の 30Hz でサンプリングする。しかし、動画を撮影している時間と加速度を計測している時間が異なるためこの段階ではデータ数は揃っていない。

[step5] 手を振っている区間の抽出

手を振っている区間を機械的に抽出する方法を実装していないため、動画から手を振り始めと振り終わりのフレームを目視で確認し、その手を振っているフレーム間のデータを加速度、動画データともに抽出する。この段階でデータ数が揃う。

[step6] 動画データの原点の座標を変換

骨格検出から得られる座標で初期設定されている原点は映像の左上となっている。そしてアスペクト比によって手首の座標が変化してしまう。そこで動画データに関する手を振っている区間の x, y それぞれの座標の平均値を求め、その点を原点とする。

[step7] 速度データへの変換

[step5] で得た手の振っている区間のデータに関して加速度のデータを積分し、動画データに関して微分することでそれぞれ速度データに変換する。本研究では加速度と座標の次元を揃えない場合と速度に揃えた場合の、ペアリングに関する比較も行うため、速度データも用いる。

[step8] 全てのデータの正規化

加速度データの座標の単位は m/s^2 であり、骨格検出から得られる座標の単位は $pixel/s$ である。そのため単位を変換せずそのまま比較しても意味がない。そこで 2 つのデータを正規化することによって単位を揃える。ここで [step6] から次元の揃っていない加速度のデータと動画データ、[step7] から次元が速度に揃っている加速度データ、動画データの合計 4 種類のデータが存在する。これらそれぞれのデータの各軸を、時刻の個数の次

1: <https://github.com/CMU-Perceptual-Computing-Lab/openpose>

元のベクトルとみなし、そのベクトルの大きさを単位ベクトルとすることで正規化する。

[step8]DTW

DTW(Dynamic Time Warping) は時系列データ同士の距離・類似度を測ることができる。2つの時系列の各点の距離を総当たりで求め、全て求めた上で2つの時系列が最短となるパスを見つける。そしてDTWの値が小さいほど類似度が高いといえる。

本論文ではまず、[step8]から得た正規化した次元の揃っていない加速度のデータと動画データの二つのDTWを計算し、加速度データと動画データの類似度を測る。次元の揃っていないデータ同士のDTWからデバイスのペアを求め人物認証とする。次に[step8]から得た正規化した次元が速度に揃っている加速度のデータと動画データについてのDTWを計算し、類似度を測る。こちらも速度に次元が揃っているデータ同士のDTWからデバイスのペアを求め人物認証とする。そして次元を揃える前後でのペアリングの精度を比較する。

4 実験

4.1 実験手法

実験では、台の上に置いてあるスマートフォンを持ち上げ、振り、スマートフォンを台に置くという動作を行い、この動作をビデオカメラで撮影した。動画撮影が開始されてから5秒ほど時間を置いて加速度をサーバに送信し始めてからスマートフォンを持ち上げる。またスマートフォンを台に置き、加速度の送受信が終了してから5秒ほど時間を置いてから動画撮影を終了するようにした。このように実験を行い、オンデマンドでスマートフォンを持つ手をトラッキングし加速度センサのデータと統合できるようにした。図3に実験の際に録画したスマートフォンを振る様子を示す。

この実験を複数人で行い、スマートフォンの振り方を変化させ動画の情報からどちらのスマートフォンであるかの判定を行うようにした。その実験概要を表2に示す。

実験から、加速度データが2つ、骨格検出から得られた人物の手首の座標データ(人物データ)が2つの計4つを取得できる。ここで加速度のデータをそれぞれAとB、人物データをそれぞれCとDとする。そうするとDTWを用いて波形の距離を計算する組合せは表3のようになる。

ここで例えば A_x, C_x はそれぞれスマートフォンAのx軸の加速度のデータ、人物Cの手首のx座標のデータを表しており、 $DTW(A_x, C_x)$ はAのx軸の加速度と人物Cのx座標に関するデータのDTWを計算する組合せとなる。 $DTW(A_x, C_x), DTW(A_y, C_y)$ はスマートフォンAと人物Cのx, yそれぞれに関するDTWであり、比較する際には、それぞれを合計しなければならない。そこで $DTW(A, C) = DTW(A_x, C_x) + DTW(A_y, C_y)$ のように加速度データと人物データのx, y全体に関するDTWを求め、その値を用いて加速度データと人物データのペアリングを行う。

例えば $DTW(A, C) > DTW(A, D)$ となるとき、AはDと



図3 実験風景及び、スマートフォンを振る様子

表2 実験の概要

人数	振り方
2人	1人縦1人横
2人	2人とも縦

表3 DTWを計算する組合せ

	C	D
A	$DTW(A_x, C_x)$	$DTW(A_x, D_x)$
	$DTW(A_y, C_y)$	$DTW(A_y, D_y)$
	$DTW(A, C)$	$DTW(A, D)$
B	$DTW(B_x, C_x)$	$DTW(B_x, D_x)$
	$DTW(B_y, C_y)$	$DTW(B_y, D_y)$
	$DTW(B, C)$	$DTW(B, D)$

表4 速度に関するDTWを計算する組合せ

	vC	vD
vA	$DTW(vA_x, vC_x)$	$DTW(vA_x, vD_x)$
	$DTW(vA_y, vC_y)$	$DTW(vA_y, vD_y)$
	$DTW(vA, vC)$	$DTW(vA, vD)$
vB	$DTW(vB_x, vC_x)$	$DTW(vB_x, vD_x)$
	$DTW(vB_y, vC_y)$	$DTW(vB_y, vD_y)$
	$DTW(vB, vC)$	$DTW(vB, vD)$

類似度が高く、スマートフォンAと人物Dがペアリングとして決定する。また、今回の実験でのスマートフォンと人物の手首の座標の正しい組合せは全て(A, C), (B, D)である。

次に実験から加速度データ、人物データともに速度に変換したデータも取得できる。速度データについてのDTWの組合せについて表4に示す。

速度のDTWも先ほどと同様にAのスマートフォンの加速度を速度に変換したvAとCの人物データを速度に変換したvCそれぞれのx, yに関するDTWを合計した $DTW(vA, vC)$ を用いる。例えば $DTW(vA, vC) < DTW(vA, vD)$ となるときvAはvDよりvCの方が類似度が高く、スマートフォンAと人物Cがペアとなる。また、速度のデータの正しい組合せについても(vA, vC), (vB, vD)である。

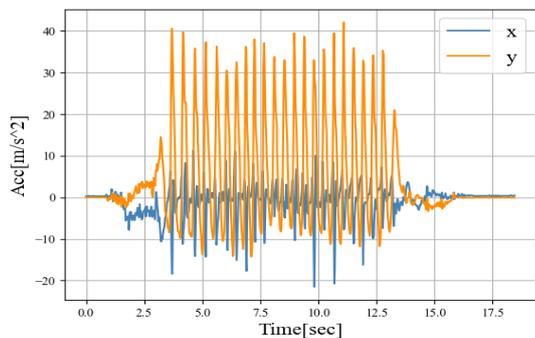


図 4 縦にスマートフォンを振っているときの加速度の変化

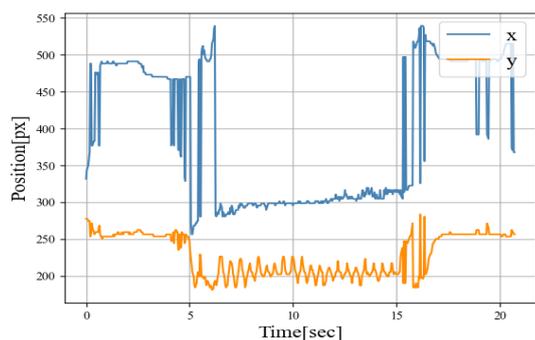


図 5 縦にスマートフォンを振っているときの手首の座標の変化

4.2 実験結果 1 : 1 人横, 1 人縦に振っているとき

まず図 4, 図 5 にそれぞれ前処理を施していない加速度のデータと骨格検出より得られた人物の右手首の座標のデータを示す。また、これらの図は振り方を 1 人横 1 人縦としたときのスマートフォンを縦に振っている人物に関するデータである。

図 3 から分かるように x の変位と y の変位を見比べると y の方が変位が大きく、縦方向に振っていることが分かる。また、図 5 では約 5.5 秒 ~ 15 秒の間を見比べるとこちらも y の方が変位が大きく、縦方向に振っていることが分かる。また、この区間は手を上下に振っている区間であり、0 秒 ~ 5 秒や 15 秒 ~ 20 秒と比べると外れ値は見受けられない。この理由は手の振っている区間以外はカメラから見ると手が体に隠れてしまうタイミングがあり、手首の位置を推定せざるを得ないため外れ値が見受けられる。

手を振っている区間はカメラから体に隠れず手首の位置を検出できる。そのため、手を振っている区間にのみ着目してデバイスとスマートフォンのペアリングの方法について検討する。

次に表 3 の組み合わせで加速度データと人物の座標データについて DTW を行った結果を表 5 に示す。表 5 から加速度センサ A と人物 C, D どちらの DTW が近いかを見ると $DTW(A, C) < DTW(A, D)$ となる。そのため、加速度センサ A と類似度が高いのは人物 C であることが分かる。同様に加速度センサ B と人物 C, D どちらの DTW が近いかを見ると $DTW(B, C) > DTW(B, D)$ となるため、加速度センサ B と人物 D が類似度が高く、今回のスマートフォンと人物の手首の

表 5 縦と横に振ったときの加速度と座標の DTW

	C	D
A	$DTW(A_x, C_x) = 3.0393$	$DTW(A_x, D_x) = 4.4172$
	$DTW(A_y, C_y) = 2.4046$	$DTW(A_y, D_y) = 3.407$
	$DTW(A, C) = 5.4439$	$DTW(A, D) = 7.8242$
B	$DTW(B_x, C_x) = 5.5365$	$DTW(B_x, D_x) = 4.1482$
	$DTW(B_y, C_y) = 4.1941$	$DTW(B_y, D_y) = 2.5404$
	$DTW(B, C) = 9.7306$	$DTW(B, D) = 6.6886$

表 6 縦と横に振ったときの速度の DTW

	vC	vD
vA	$DTW(vA_x, vC_x) = 1.8767$	$DTW(vA_x, vD_x) = 2.0096$
	$DTW(vA_y, vC_y) = 2.4325$	$DTW(vA_y, vD_y) = 2.3967$
	$DTW(vA, vC) = 4.3092$	$DTW(vA, vD) = 4.4063$
vB	$DTW(vB_x, vC_x) = 5.8838$	$DTW(vB_x, vD_x) = 4.2411$
	$DTW(vB_y, vC_y) = 2.5699$	$DTW(vB_y, vD_y) = 2.1640$
	$DTW(vB, vC) = 8.4537$	$DTW(vB, vD) = 6.4051$

座標のペアリングは (A, C) , (B, D) となり、これは正しい組合せとなる。したがって正しくペアリングすることができた。

次に加速度のデータと人物の手首の座標に関するデータを速度に変換したときの DTW を表 6 に示す。表 6 から $DTW(vA, vC) < DTW(vA, vD)$ となるため、加速度センサ A と類似度が高いのは人物 C であることが分かる。同様に $DTW(vB, vC) > DTW(vB, vD)$ となるため、加速度センサ B と人物 D が類似度が高い。よって今回のスマートフォンと人物の手首の座標のペアリングは (vA, vC) , (vB, vD) となり、正しいペアリングとなった。

4.3 実験結果 2 : 2 人とも縦に振っているとき

図 6, 図 7 にそれぞれ前処理を施していない加速度に関するデータと骨格検出より得られた人物の右手首の座標データを示す。また、これらの図は 2 人ともスマートフォンを縦に振っているときの時間変化である。図 6 は 2 つの加速度データの y 軸方向に関する時間変化であり、 y_1 と y_2 を比較すると y_2 の方が変位の変わる間隔が短いことが分かる。そのため y_1 よりも y_2 の方が速くスマートフォンを振っていることが分かる。

図 7 は骨格検出より得られた人物 2 人の右手首の y 座標に関する時間変化である。約 5.5 秒 ~ 15 秒の間を見ると、0 秒 ~ 5 秒や 15 秒 ~ 20 秒の区間と比べると外れ値は見受けられない。これは 1 人横, 1 人縦に振っているときと同様に手を振っている区間は体の正面でスマートフォンを振っているため、正確に骨格検出ができる。そのため、約 5.5 秒 ~ 15 秒の区間の手を振っている区間についてのみ考えていく。図 7 においても手を振っている区間の y_2 の方が y_1 より変位の変わる区間が短く、スマートフォンを速く振っていることが分かる。

表 7 に加速度データと人物の座標データについて DTW を行った結果を示す。表 7 から加速度センサ A と人物 C, D どちらの DTW が近いかを見ると $DTW(A, C) < DTW(A, D)$ となる。そのため、加速度センサ A と類似度が高いのは人物 C であることが分かる。同様に加速度センサ B と人物 C, D どちらの DTW が近いかを見ると $DTW(B, C) < DTW(B, D)$ となる

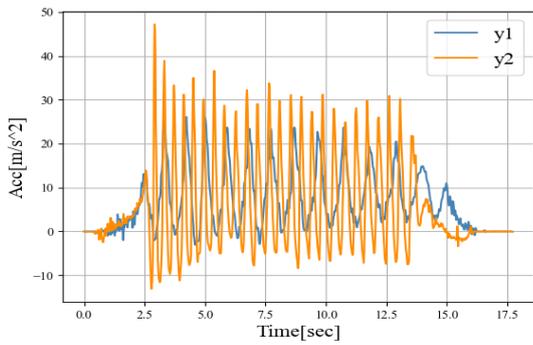


図 6 縦にスマートフォンを振っているときの 2 つの加速度データの変位の比較

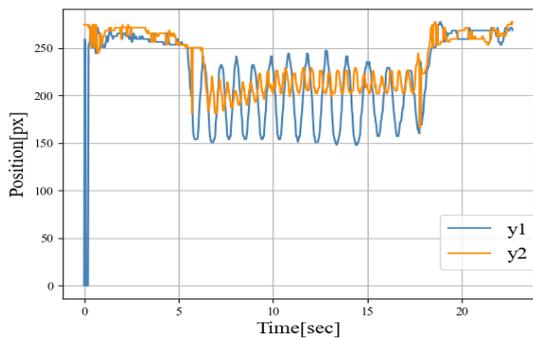


図 7 縦にスマートフォンを振っているときの 2 つの手首座標データの変位の比較

表 7 2 人縦に振ったときの加速度と座標の DTW

	C	D
A	$DTW(A_x, C_x) = 4.1994$	$DTW(A_x, D_x) = 7.5047$
	$DTW(A_y, C_y) = 2.4492$	$DTW(A_y, D_y) = 2.5976$
	$DTW(A, C) = 6.6486$	$DTW(A, D) = 10.1023$
B	$DTW(B_x, C_x) = 3.8608$	$DTW(B_x, D_x) = 7.4939$
	$DTW(B_y, C_y) = 3.6077$	$DTW(B_y, D_y) = 4.4749$
	$DTW(B, C) = 7.4685$	$DTW(B, D) = 11.9688$

ため、加速度センサ B と人物 C が類似度が高く、スマートフォン A, B どちらも人物 C と類似度が高いと判定される。

この場合、ペアリングした際の 2 つのスマートフォンと 2 人の人物の DTW の合計が小さくなる方にすればよい。すなわち $DTW(A, C) + DTW(B, D)$ と $DTW(A, D) + DTW(B, C)$ の大小関係を比較すればよく、 $DTW(A, C) + DTW(B, D) = 18.6174 > DTW(A, D) + DTW(B, C) = 17.5708$ となるため今回のスマートフォンと人物の手首の座標のペアリングは (A, D), (B, C) となり、間違ったペアリングとなった。

次に加速度のデータと人物の手首の座標に関するデータを速度に変換したときの DTW を表 8 に示す。表 8 から $DTW(vA, vC) < DTW(vA, vD)$ となるため、加速度センサ A と類似度が高いのは人物 C であることが分かる。同様に $DTW(vB, vC) > DTW(vB, vD)$ となるため、加速度センサ B と人物 D が類似度が近い。よって今回のスマートフォンと人物

表 8 2 人縦に振ったときの速度の DTW

	vC	vD
vA	$DTW(vA_x, vC_x) = 2.3232$	$DTW(vA_x, vD_x) = 2.4895$
	$DTW(vA_y, vC_y) = 2.0107$	$DTW(vA_y, vD_y) = 2.6482$
	$DTW(vA, vC) = 4.3339$	$DTW(vA, vD) = 5.1377$
vB	$DTW(vB_x, vC_x) = 7.0558$	$DTW(vB_x, vD_x) = 6.5239$
	$DTW(vB_y, vC_y) = 3.3170$	$DTW(vB_y, vD_y) = 2.4583$
	$DTW(vB, vC) = 10.3728$	$DTW(vB, vD) = 8.9822$

の手首の座標のペアリングは (vA, vC), (vB, vD) となり、正しいペアリングとなった。

次元を揃える前では間違ったペアリングとなっている。これは図 7 の y2 の座標変位を見ると、y1 より周波数は大きいに変位が明らかに小さい。そして図 6 の y2 の加速度の変位を見ると y1 より周波数が大きく変位が大きくなっている。そのため DTW を行った際に変位に関して誤差が生じてしまい誤ってペアリングしてしまったと考えられる。

次元を速度に揃えると縦に振ったときであるので、特にスマートフォン B と人物 D の y に関する DTW について、 $DTW(B_y, D_y) = 4.4749$ が $DTW(vB_y, vD_y) = 2.4583$ となり、次元を速度に揃える前後で DTW の値が小さくなった。次元を揃えると加速度の変位を積分、座標の変位を微分することになるため、次元を揃える前の変位の誤差の影響を少なくすることができることが分かった。

5 おわりに

本論文では加速度の値と骨格検出の座標変化から匿名の人物認証をする手法を提案した。提案手法では 5G 通信網を利用して加速度センサデータを送信し、骨格検出の座標変化と情報の紐づけを行った。まず、5G 通信網を通して WebSocket によりオンラインで加速度を低遅延で送受信することが可能であることが分かった。提案手法では、加速度データと座標データを前処理したのち、DTW を用いることによって二つの波形の類似度を測った。これにより、次元の違う加速度データと座標データの類似度の判定を行えるようになった。また、DTW を行う際に、次元を速度に揃えて類似度を測る手法も同時に検討した。

実験より、取得した加速度のデータと座標のデータを次元を揃えずにペアリングする方法と、次元を揃えてペアリングする手法の両方を行い、両者の検討、比較を行った。次元を揃えない方法では 2 人が横と縦に振っている場合は正しくペアリングすることができた。2 人が縦に振っている場合は誤ったペアリングとなった。そして次元を揃える方法では 2 人が横と縦に振っている場合と 2 人が縦に振っている場合、どちらも正しくペアリングできた。

次元を揃える前では加速度の変位、座標の変位の影響を受けてしまい間違ったペアリングとなってしまうことが分かった。そのため速度に次元を揃えることが必要であることが分かった。

今後の展望としてまず今回の実験では 2 人で行ったため、さらに人数を増やして 3 人、4 人の場合に正しくペアリングを行いたいと考えている。また、今回は手を振っている区間を動画

から盗まれる個人情報” <https://www.sankei.com/article/20200305-24TRH7EBQNJ3HHPFVIGF03T5JY/>

- [3] 繁田脩, 鏡慎吾, 橋本浩一. “動画像からの加速度センサ所有者の検出とその性能評価,” 計測自動制御学会東北支部, 249-12, 2009.
- [4] 繁田脩, 鏡慎吾, 橋本浩一. “加速度センサを有する運動物体の固定カメラ視野内における実時間識別,” 第 51 回自動制御連合講演会, セッション ID 730, 2008.
- [5] 長友誠, 油健太郎, 岡崎直宣, 朴美娘. “カメラと加速度センサを用いたデバイスペアリング方式の提案とその評価,” 分散協調とモバイルシンポジウム, pp.172-178, 2018.

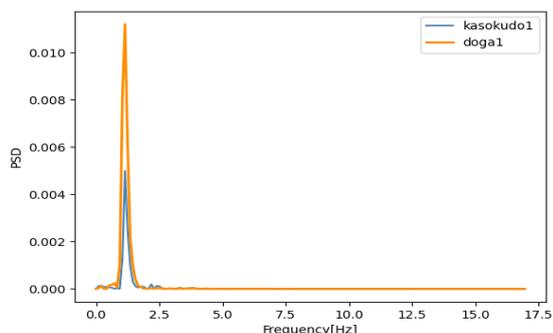


図 8 2 人縦に振っているときの y 方向パワースペクトル密度のペア 1

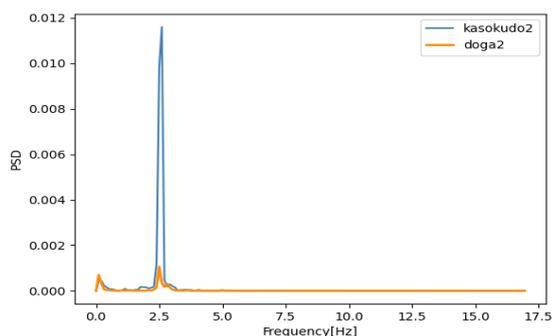


図 9 2 人縦に振っているときの y 方向パワースペクトル密度のペア 2

のフレームを参考に目視で抽出したため、区間の抽出を機械的に行いたいと考えている。さらに並行して研究している 5G 通信網を通してリアルタイムで動画をサーバに送信すると本研究を合わせて、オンデマンドでなくリアルタイムで人物認証を行いたいと考えている。

また図 8, 図 9 に 2 人縦に振っているときの加速度データと座標データのパワースペクトル密度を示す。また, 図 8, 図 9 にはそれぞれ加速度データと座標データのそれぞれの y 方向のパワースペクトル密度であり, 加速度データと座標データは正しいペアとなっている。

図 8, 図 9 より, 加速度データと座標データどちらも正しい組み合わせではほぼ同じ周波数でパワースペクトル密度の値をとっている。またこのデータは加速度データと座標データの次元を揃えていない。このように周波数特性を用いてペアリングを行うことも今後の展望として考えている。

6 謝 辞

本研究の一部は国立情報学研究所 CRIS 委託研究及び東京都立大学「ローカル 5G 環境を活用した最先端研究」の助成を受けたものである。

文 献

- [1] 守脇幸佑, 松村大吾, 八木康史. “人物属性を考慮した歩容認証,” 情報処理学会研究報告, Vol.2018-CVIM-212, No. 22.
- [2] 産経新聞. “ピースサインから指紋、瞳に最寄り駅 S N S