

頭部と両手腕のモーションデータのDTW距離に基づくVRユーザ識別

三浦 晃暉^{1,a)} 菊池 浩明^{1,b)}

概要: 近年、バーチャルリアリティ (VR) 技術の普及率が上昇しており、メタバースが次世代のインターネットの利用形態として注目されている。教育やゲーム、観光といった様々な用途やコンテンツがある VR では、利用者は匿名の存在であり、架空の人物として振る舞うことができる。しかし、詳細に記録された VR ユーザのモーションデータはそのユーザに関する多くの特徴を含み、現実世界で該当する個人が識別される可能性がある。例えば、Nair らや Liebers らは、VR リズムゲーム “Beat Saber” の公開リプレイデータから、9 割を超える精度で個人を識別できることを示した。しかし、個人識別の精度に最も影響を与えたのは身長や手の長さなどといった静的特徴量であった。腕の動かし方のくせなどの動的な特徴量を用いれば、より高い精度で識別できる可能性がある。そこで、本研究では、VR 体験中のモーションデータから DTW (Dynamic Time Warping) 距離を用いて、動的特徴量を要因とする VR 空間における個人識別のリスクを定量化する。

VR user identification based on DTW distance of head and arms motion data

Abstract: In recent years, the adoption rate of virtual reality (VR) technology has been on the rise, and the metaverse is attracting attention as a next-generation form of internet usage. VR offers a variety of applications and content, such as education, gaming, and tourism, where users can remain anonymous and behave as fictional characters. However, there is a possibility that individuals in the real world can be identified from motion data that records VR users' head-and-hands movements in detail, represented in time-series data. Nair and Lieber demonstrated that publicly available replay data from the VR rhythm game “Beat Saber” can identify individuals with over 90% accuracy. Nevertheless, the features that had the greatest impact on identification accuracy were static attributes such as height and arm length. Therefore, in this study, we attempt to identify individuals in VR domain based on dynamic features such as users' distinctive ways of moving their arms by employing the DTW (Dynamic Time Warping) distance derived from motion data recorded during VR experiences.

1. はじめに

近年、バーチャルリアリティ (VR) 技術の普及率が上昇しており、メタバースが次世代のインターネットの利用形態として注目されている [1]。教育やゲーム、観光といった様々な用途やコンテンツがある VR では、利用者は匿名の存在であり、架空の人物として振る舞うことができる。

しかし、VR ユーザの動作を詳細に記録したリプレイデータから、現実世界で個人が識別される可能性がある。Nair ら [3] や Liebers ら [4] は、VR リズムゲーム “Beat Saber” [5]

の公開リプレイデータから、9 割を超える精度で個人を識別できることを示した。しかし、個人識別の精度に最も影響を与えたのは身長や手の長さなどといった静的特徴量であった。

動作などの動的特徴量を用いると識別精度を向上できる可能性がある。そこで、本研究では、時系列データの DTW (Dynamic Time Warping) 距離を用いて、腕の動かし方のくせなどの動的な特徴量を要因とする個人識別を試みる。VR 体験中のモーションデータから個人を識別することを目的とする。動的な特徴量の粒度を細かくすればするほど、他人と誤識別する確率は減るが、同一人物による複数の試行を他人と誤って拒否する確率が増える。従って、本人らしさを残しつつ、複数回のパターンが固有になるようなモーションを用いることが望まれる。Nair ら [3] は、Beat Saber

¹ 明治大学総合数理学部

School of Interdisciplinary Mathematical Sciences, Meiji University

^{a)} ev210597@meiji.ac.jp

^{b)} kkn@meiji.ac.jp

におけるオブジェクトの打ち方の特徴に着目した。鈴木ら [6] は、コントローラを持ち上げる自然な動作の個人差を用いた個人認証を行った。我々は本研究において、四拍子の指揮を用いた特徴量を提案する。指揮では何度も同じパターンを繰り返すので、安定した特徴が得られることが期待できる。

2. 関連研究

2.1 階層的分類モデル

2023 年, Nair らは VR リズムゲーム “Beat Saber” [5] をプレイ中の頭部と手の動きのデータから, 50,000 を超える VR ユーザを高精度で識別する手法 [3] を提案した。ナイーブに機械学習を用いた手法で識別をすると, 大規模 VR ユーザを識別するには莫大な計算コストと記憶コストがかかる。そこで, Nair らは, 位置・回転の統計量といったモーションデータとゲーム内オブジェクトとの相互作用を表すコンテキストデータを統合した 232 次元のハイブリッド特徴量を用いて, LightGBM を用いた階層的分類モデルを構築した。各ユーザ 5 分間のモーションデータを学習し, 100 秒のテストデータを 94.33%, 10 秒のデータで 73.20% の精度で識別した。

2.2 マルチクラス分類器

2023 年, Liebers らは VR 環境における行動的バイオメトリクスの安定性と時間経過による識別性能の変化を評価するため, リモートフィールドスタディ [4] を実施した。15 名の被験者を対象に, VR リズムゲーム “Beat Saber” の同じ楽曲を 8 週間に渡りプレイさせ, その動作データをランダムフォレストによるマルチクラス分類器を用いて評価した。ランダムフォレストでは, HMD と左右のコントローラそれぞれの回転量 (クォータニオン) および, HMD に対する各コントローラの相対座標の統計量 (最小値, 最大値, 平均, 標準偏差) の計 72 特徴量を用い, ハイパーパラメータはデフォルトのまま使用されている。初回セッションのデータセットの 80% で訓練したモデルを, 残りの 20% のデータで評価した場合, 識別精度 (F1 スコア) は 86% であった。また, 初回セッションのデータセットで訓練したモデルを, 後続セッションのデータセットで評価した場合, 時間が経過したセッションで評価するにつれて識別精度は減少した。

2.3 DTW 距離による歩容識別

2018 年, 森らは Kinect を用いて収集した歩容データの時系列情報に DTW を適用し, 高精度な個人認証を実現する手法を提案した [7]。提案手法は, モーションキャプチャデバイスである Kinect を活用し, 被験者に装置を装着させることなく関節の動きから得られる身体の 25 箇所の関節の 3 次元スケルトンデータを解析して, 特定の関節の動

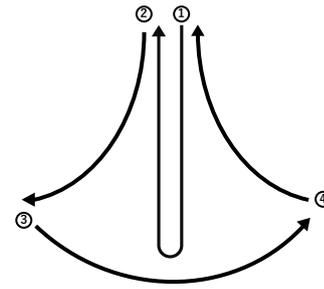


図 1 四拍子の指揮の例

きを統合し, 最適な特徴量の組み合わせを検討した。このデータに DTW を適用することで, 時間軸を調整した類似性評価を可能とした。実験では, 31 人の被験者を対象に, 関節間の統合 DTW 距離を用いて個人認証を行い, 等価エラー率 (EER) 0.036 の精度を達成した。

2.4 コントローラのリフト動作を用いたユーザ認証

2022 年, 鈴木らは VR ヘッドマウントディスプレイ (HMD) のユーザ認証方法として, ユーザの自然な動作を活用したジェスチャ認証を提案した [6]。本研究では, VR 空間におけるモーションデータが個人識別につながる可能性を四拍子の動作から検証し, そのプライバシーリスクを明らかにするを目的としている。一方, 鈴木らの研究では, HMD のログイン認証におけるユーザビリティとセキュリティを両立する新たなジェスチャ認証方式の提案と評価を行うことを目的としている。10 名の被験者を対象に, HMD 利用時によく行われるコントローラのリフト動作とピンチ動作を行わせ, 個人を認証する手法の有効性を検証した。DTW, SVM (Support Vector Machine), ランダムフォレストの 3 種類の認証モデルを比較した結果, SVM モデルが最も高い精度を示し, リフト動作では平均 0.991, ピンチ動作では平均 0.989 の認証精度を達成した。しかし, DTW は 3 種類の認証モデルの中で, 最も低い精度となった。

3. 提案個人識別

3.1 個人識別手法

本研究では, VR デバイスのセンサから得られた HMD と両手のコントローラの 3 次元座標を測定し, 動作の時系列データの DTW 距離を算出して, 個人認証を行う。また, 動作として, 四拍子の指揮を採用する。四拍子の指揮とは図 1 に示す動作である。この動作は, 単純な動作であるが個人差が存在し, 何度も同じパターンを繰り返すため, 安定した特徴が得られると考えられる。

3.2 DTW 距離の計算

DTW は, 音声などの波形識別に用いられる手法で, 長

さの異なる時系列データの類似度を測る。森ら [7] と同様に、長さ n_p の時系列データ $P = (p_1, p_2, \dots, p_{n_p})$ と長さ n_q の時系列データ $Q = (q_1, q_2, \dots, q_{n_q})$ の DTW 距離 $d(P, Q)$ を $d(P, Q) = f(n_p, n_q)$ と定義する。ここで $f(i, j)$ は、

$$f(i, j) = \|p_i - q_j\| + \min \begin{cases} f(i, j-1), \\ f(i-1, j), \\ f(i-1, j-1), \end{cases}$$

$$f(0, 0) = 0, f(i, 0) = f(0, j) = \infty,$$

と定める。また、 $\|p - q\|$ は、3次元ベクトルのユークリッド距離である。

3.3 DTW の適用

図 2 と図 3 に、4.2.1 節で取得した被験者 D の異なる試行における右手コントローラの x 座標と y 座標の時系列データに DTW を適用した結果を示す。また、図 4 と図 5 に、4.2.1 節で取得した異なる被験者 C と D の右手コントローラの x 座標と y 座標の時系列データに DTW を適用した結果を示す。図中の青線とオレンジ線は、それぞれ異なる時系列データを表す。点線は、DTW によって対応付けられた点同士を結んでいる。

図 2 と図 3 について、本人同士では時系列にずれはあるものの、似た手の振り方の四拍子の指揮をしたことがうかがえる。図 4 と図 5 について、同一被験者の場合と比較して、2つの波形の形状が大きく異なる。特に、 y 座標では、波形のピークの数や位置が異なっており、DTW による対応付けも複雑になっている。また、 x 座標では被験者 D は C より左右に手を振る動作をしていることが分かる。ここで、図 6 に、被験者 C と D の四拍子の指揮の右手コントローラの軌跡 (x - y 平面) を示す。青点は四拍子の指揮の開始点、赤点は終了点を表す。被験者 C と D の四拍子の指揮の軌跡は異なり、被験者 C と D の間に前述のような特徴が見られる。

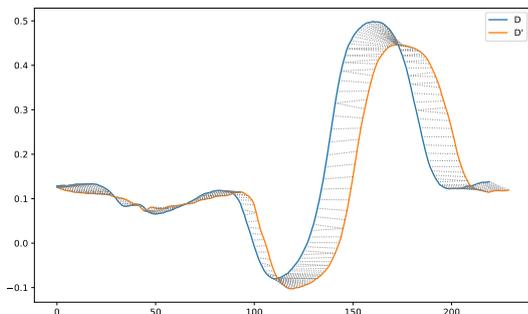


図 2 被験者 D 同士の右手コントローラ x 座標への DTW 適用例

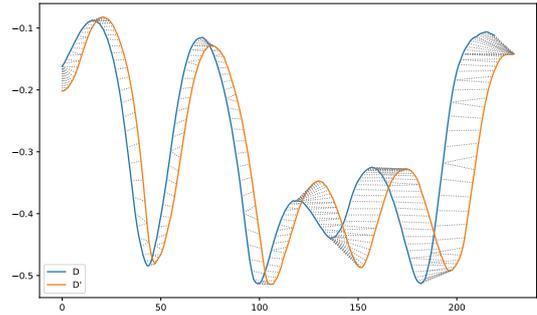


図 3 被験者 D 同士の右手コントローラ y 座標への DTW 適用例

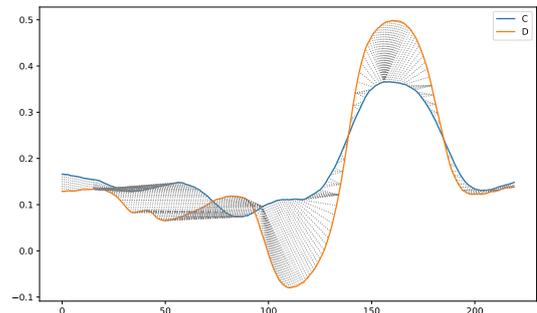


図 4 被験者 C と D の右手コントローラ x 座標への DTW 適用例

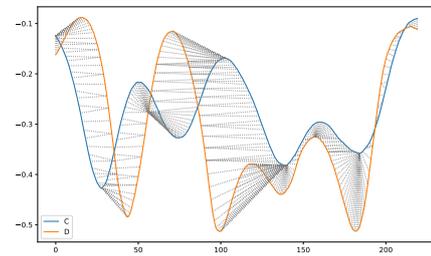


図 5 被験者 C と D の右手コントローラ y 座標への DTW 適用例

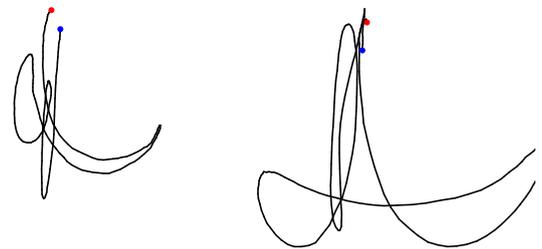


図 6 被験者 C (左) と D (右) の四拍子の指揮の軌跡

3.4 提案方式

次の手順で個人を識別する。

- (1) 全てのユーザ $i = 1, \dots, n$ についてモーションデータ y_i を測定し、 i のテンプレートとする。

表1 測定環境

項目	環境
VR デバイス	Meta Quest 3
開発環境	Unity 2022.3.28f1
サンプリングレート	72 fps

表2 被験者の情報

項目	情報
参加者年齢	20-50 代
参加者性別	男性 11 名 女性 1 名
参加人数	12 名

(2) 未知のユーザ j のモーションデータ x_j を測定し, n 個のテンプレートとの DTW 距離を求め, 次の j^* を推定ユーザとする.

$$j^* = \operatorname{argmin}_{i \in \{1, \dots, n\}} d(y_i, x_j)$$

4. 実験

4.1 実験目的

本実験の目的は, VR デバイスから得られる HMD と両手コントローラの 3 次元座標の時系列データに対し, DTW 距離を用いた個人識別手法の有効性を検証することである. 提案手法が先行研究の手法と比べ, どの程度の識別精度を達成できるかを定量的に評価する. さらに, 本研究では以下のリサーチクエスチョンを設定する.

RQ1. 四拍子のリズムが正確に繰り返されているかどうか, リズムに対する個人差はあるか?

RQ2. DTW 距離を用いた提案方式と静的特徴量による従来方式 [4] とで, 個人識別リスクが高いのはどちらか?

4.2 実験 1

4.2.1 実験方法

四拍子の指揮のデータで, 3.4 節の提案方式と 2.2 節の Liebers らの方式を評価する.

動作の測定環境を表 1 に示す. Meta Quest 3 を用いて HMD・両手のコントローラの 3 次元座標と回転量の時系列データを得るために, Unity のアプリケーションを通して, HMD・両手のコントローラの 3 次元座標と回転量を取得するシステムを C# を用いて実装した.

被験者の情報を表 2 に示す. 被験者 12 名に対し, 四拍子の指揮の動作を行わせ, モーションデータを計測した. 2024 年 11 月 1 日から 2024 年 11 月 15 日にかけて一人あたり両手の四拍子の指揮を 6 回測定した. 測定したデータについて, 被験者により開始位置と正面の向きが異なっていたため, 開始位置と正面の向きを揃える操作を行った.

測定後, 各参加者の 6 回の試行データを 3 つのグループに分割する. 各グループについて, そのグループをテストデータ, 残りの 2 つのグループをテンプレートデータとして各手法を適用する. 3 回の評価結果の平均正解率, 平均適合率, 平均再現率, 平均 F1 スコアで手法同士の比較を行う.

4.2.2 取得データの統計量分析

被験者全体の各特徴量の統計量 (平均値, 標準偏差) を表 3 に示す. hmd はヘッドマウントディスプレイ, rcon は

右手コントローラ, lcon は左手コントローラ, pos は位置座標, quat はクォータニオンを表している. コントローラの位置座標は, xy 方向の標準偏差が大きい. これは, 四拍子の指揮が主に手の上下左右の運動を含むためと考えられる. 右手コントローラと左手コントローラを比較すると, 上下方向 (y 軸) の平均値は絶対値が近いことから, 両手の動きが等しい高さを維持していると考えられる.

被験者ごとの動作データのフレーム数の統計量 (最小値, 最大値, 平均値, 標準偏差) を表 4 に示す. 表 4 から, フレーム数には個人差があることが分かる. また, 各被験者内での標準偏差は比較的小さい.

従って, **RQ1** について, 各被験者は比較的一定のリズムで四拍子の指揮が行っており, 個人によって四拍子の指揮のリズムが異なると言える.

表3 被験者全体の各特徴量の統計量

特徴量	平均値	標準偏差
hmd-pos-x	0.00	0.01
hmd-pos-y	0.00	0.00
hmd-pos-z	0.01	0.01
rcon-pos-x	0.11	0.14
rcon-pos-y	-0.33	0.15
rcon-pos-z	0.27	0.07
lcon-pos-x	-0.18	0.13
lcon-pos-y	-0.33	0.14
lcon-pos-z	0.24	0.08
hmd-quat-x	0.02	0.07
hmd-quat-y	0.05	0.09
hmd-quat-z	-0.01	0.02
hmd-quat-w	-0.99	0.01
rcon-quat-x	0.02	0.21
rcon-quat-y	0.00	0.20
rcon-quat-z	-0.05	0.14
rcon-quat-w	-0.04	0.94
lcon-quat-x	-0.06	0.21
lcon-quat-y	-0.17	0.20
lcon-quat-z	-0.05	0.08
lcon-quat-w	0.93	0.06

4.2.3 DTW 距離分析

図 7 に, 本人同士と他人同士の DTW 距離の確率分布を示す. 本人同士の DTW 距離は, 他人同士の DTW 距離に比べて, 明らかに小さい値に集中している. 本人同士の DTW 距離は, おおよそ 50 未満に分布している. 一方, 他人同士の DTW 距離は, より広い範囲に分布しており, 50 を下回る値も多く見られる.

表4 各被験者のフレーム数統計量

被験者	最小値	最大値	平均値	標準偏差
A	216	241	232.5	8.3
B	165	200	180.3	13.3
C	210	251	224.3	13.4
D	211	246	225.7	11.6
E	200	221	208.5	7.2
F	180	236	202.7	20.6
G	206	271	227.7	21.1
H	201	251	222.2	15.9
I	165	211	186.5	15.3
J	231	260	246.8	9.1
K	160	191	167.7	11.8
L	210	240	221.2	10.6

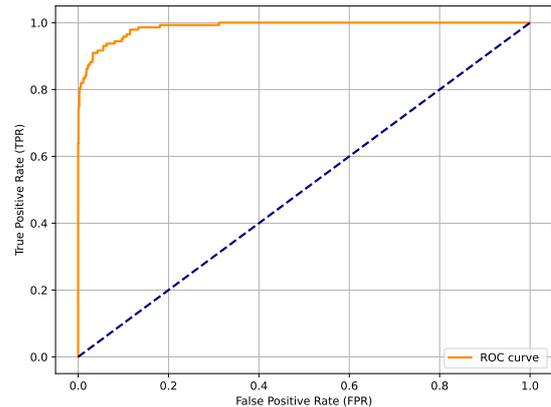


図9 ROC 曲線

図8に、DTW距離の閾値に対するFAR（他人受入率）とFRR（本人拒否率）の変化を示す。また、図9にROC曲線を示す。EERは0.0644であり、その際のDTW距離閾値は52.49であった。DTW距離の閾値を小さくすると、FARは低下するが、FRRは上昇する。逆に、閾値を大きくすると、FRRは低下するが、FARは上昇する。

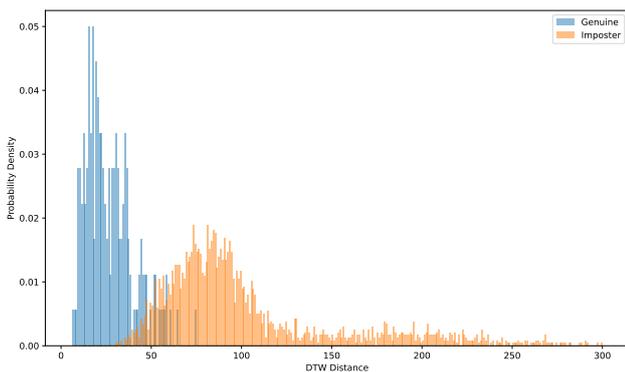


図7 本人同士と他人同士のDTW距離の確率分布

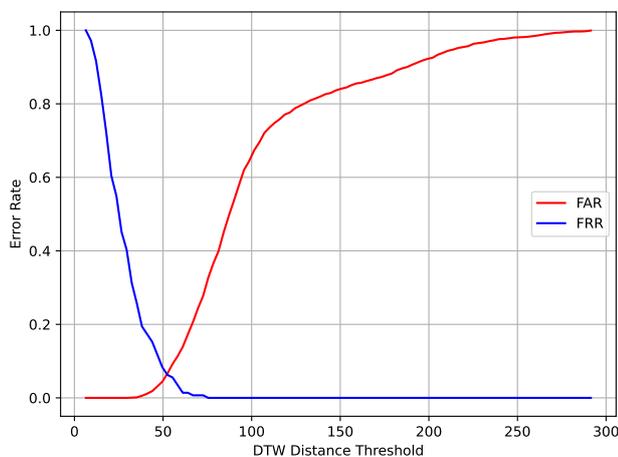


図8 FARとFRRのトレードオフ

4.2.4 実験結果

各評価指標（正解率、適合率、再現率、F1スコア）の平均値を表5に示す。また、DTWとランダムフォレストそれぞれによる識別結果の混同行列を表6、7に示す。

表5をみると、DTWはすべての評価指標において、ランダムフォレストを上回っている。また、表6、7をみると、DTWはランダムフォレストよりも1名分多く識別している。

これらの結果から、DTWを用いた提案手法は、四拍子という単純な動作の個人識別において、先行研究で用いられたランダムフォレストよりもわずかに高い性能を示すことが確認された。

表5 実験1におけるDTWとランダムフォレストの性能比較

手法	正解率	適合率	再現率	F1スコア
DTW	0.98	0.99	0.98	0.98
ランダムフォレスト	0.97	0.97	0.98	0.97

表6 DTWによる識別結果の混同行列

真 \ 推定												
	A	B	C	D	E	F	G	H	I	J	K	L
A	6	0	0	0	0	0	0	0	0	0	0	0
B	0	6	0	0	0	0	0	0	0	0	0	0
C	0	0	6	0	0	0	0	0	0	0	0	0
D	0	0	0	6	0	0	0	0	0	0	0	0
E	0	0	0	0	6	0	0	0	0	0	0	0
F	0	0	0	0	0	5	0	0	0	0	0	1
G	0	0	0	0	0	0	6	0	0	0	0	0
H	0	0	0	0	0	0	0	6	0	0	0	0
I	0	0	0	0	0	0	0	0	6	0	0	0
J	0	0	0	0	0	0	0	0	0	6	0	0
K	0	0	0	0	0	0	0	0	0	0	6	0
L	0	0	0	0	0	0	0	0	0	0	0	6

表7 ランダムフォレストによる識別結果の混同行列

推定 \ 真	A	B	C	D	E	F	G	H	I	J	K	L
A	6	0	0	0	0	0	0	0	0	0	0	0
B	0	6	0	0	0	0	0	0	0	0	0	0
C	0	0	6	0	0	0	0	0	0	0	0	0
D	0	0	0	6	0	0	0	0	0	0	0	0
E	0	0	0	1	5	0	0	0	0	0	0	0
F	0	0	0	0	0	6	0	0	0	0	0	0
G	0	0	0	0	0	0	5	0	0	0	0	1
H	0	0	0	0	0	0	0	6	0	0	0	0
I	0	0	0	0	0	0	0	0	6	0	0	0
J	0	0	0	0	0	0	0	0	0	6	0	0
K	0	0	0	0	0	0	0	0	0	0	6	0
L	0	0	0	0	0	0	0	0	0	0	0	6

4.3 実験2

4.3.1 実験方法

Liebers ら [4] の Beat Saber のデータで、3.4 節の提案方式と 2.2 節の Liebers らの方式を評価する。

Liebers らは、収集したデータを用いて以下の2つの実験を行っている。

- (1) **単一セッション学習:** 最初のセッションのデータのみを用いてモデルを学習し、後続のセッションのデータでテストする。
- (2) **複数セッション学習:** あるセッションのデータでテストする際に、それ以前のすべてのセッションのデータを用いてモデルを学習する。

これらの実験は、データを8セッション以上提出した被験者8名 ($N = 8$) とデータを4セッション以上提出した被験者15名 ($N = 15$) の2つのグループに対して実施された。被験者により、セッション毎に提出したデータの数が異なるため、評価指標として重み付き F1 スコアを用いている。本実験では、Liebers らの公開データを用い、楽曲をプレイしてから初めてスコアが加算された後の 1.5 秒間の動作に DTW を適用し、同様の手順で得た結果を Liebers らの結果と比較する。

4.3.2 実験結果

表 8, 9 に $N = 8$ の場合、表 10, 11 に $N = 15$ の場合のそれぞれの識別結果を示す。表の数字は重み付き F1 スコアとその標準偏差を示している。DTW とランダムフォレストともに、学習(テンプレート)データが増えるにつれて、全体的にテストデータの F1 スコアが向上するのが見られる。ランダムフォレストは、テストデータのセッションが進むにつれて、全体的に F1 スコアが低下している。対して DTW は、テストデータのどのセッションにおいても F1 スコアを維持していることから、DTW は時間経過による動作の変化の影響を受けにくいことが分かる。また、F1 スコアの標準偏差においても、DTW はランダムフォレストより全体的に小さいことが分かる。

従って、RQ2 について、4.2.4 節と 4.3.2 節により、DTW

距離を用いた提案方式は、静的特徴量に基づく従来方式と比較して、個人識別リスクが高いと言える。

4.4 考察

4.4.1 実験1の誤識別

4.2.4 節において、表 6 を見ると、被験者 F が L に誤識別されている。

表 12 に、被験者 F が L と誤識別されたテストデータの DTW 距離を示す。被験者 F の 1 回目のデータと被験者 L の 3 回目のデータの DTW 距離が最も短いことが分かる。ここで、図 10 に被験者 F の 1 回目と 4 回目、L の 3 回目の四拍子の指揮の右手コントローラの軌跡 (x-y 平面) を示す。被験者 F の 4 回目の四拍子の指揮は、手の位置が下がっており、左右のブレが 1 回目より激しいことが分かる。対して被験者 L の 3 回目の四拍子の指揮は、被験者 L の 1 回目と四拍子の指揮の形状が似ている。これらの理由から、被験者 F の 1 回目と L の 3 回目のデータ間の DTW 距離が小さくなった原因と考えられる。

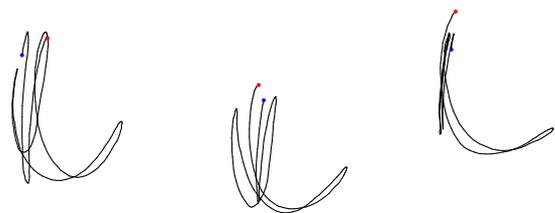


図10 被験者 F_1 (左) と F_4 (中央) と L_3 (右) の四拍子の指揮の軌跡

4.4.2 時間経過による動作の変化の影響

4.3.2 節において、DTW はランダムフォレストと比べ、時間経過による動作の変化の影響を受けにくいことが分かった。Beat Saber では、同じ曲、同じ難易度であれば、ノーツが流れてくるタイミングが基本的に同じであり、プレイヤーはノーツを切ってスコアを上げる。一般的に、試行回数を増やすにつれ、プレイヤーのプレイは上達していくと予想できる。DTW は、プレイの基本的な時間構造が同じであっても、プレイの上達による細かいタイミングのずれや、動作速度の変化に強いと言える。

4.4.3 動作の個人差

鈴木らの研究 [6] では、リフト動作とピンチ動作を用いた認証において、SVM が最も高い識別精度を示し、次いでランダムフォレスト、DTW の順であった。一方、本研究の四拍子の指揮を用いた実験では、DTW が最も高い識別精度を示した。リフト動作やピンチ動作は、個人差が表れにくい動作であると考えられるが、四拍子の指揮は、動作に個人差が表れやすい。そのため、本研究では鈴木らの研究に比べ、DTW が有利になったと考えられる。

表 8 N = 8 の場合の識別結果 (DTW)

トレーニングデータ	テストデータ						
	セッション 2	セッション 3	セッション 4	セッション 5	セッション 6	セッション 7	セッション 8
セッション 1	0.69 (0.16)	0.74 (0.34)	0.70 (0.33)	0.62 (0.22)	0.70 (0.19)	0.72 (0.22)	0.66 (0.37)
セッション 1-2		0.86 (0.15)	0.85 (0.15)	0.58 (0.31)	0.74 (0.18)	0.78 (0.18)	0.74 (0.30)
セッション 1-3			0.91 (0.09)	0.76 (0.17)	0.77 (0.21)	0.83 (0.19)	0.80 (0.16)
セッション 1-4				0.80 (0.16)	0.85 (0.13)	0.83 (0.19)	0.80 (0.16)
セッション 1-5					0.89 (0.11)	0.86 (0.19)	0.85 (0.15)
セッション 1-6						0.92 (0.12)	0.85 (0.15)
セッション 1-7							0.85 (0.15)

表 9 N = 8 の場合の識別結果 (ランダムフォレスト)[4]

学習データ	テストデータ						
	セッション 2	セッション 3	セッション 4	セッション 5	セッション 6	セッション 7	セッション 8
セッション 1	0.71 (0.33)	0.65 (0.41)	0.55 (0.41)	0.42 (0.37)	0.48 (0.41)	0.34 (0.44)	0.27 (0.39)
セッション 1-2		0.96 (0.08)	0.70 (0.34)	0.57 (0.39)	0.73 (0.34)	0.57 (0.40)	0.34 (0.40)
セッション 1-3			0.78 (0.33)	0.62 (0.39)	0.82 (0.34)	0.57 (0.38)	0.49 (0.44)
セッション 1-4				0.74 (0.33)	0.91 (0.19)	0.59 (0.42)	0.57 (0.48)
セッション 1-5					0.83 (0.35)	0.55 (0.42)	0.61 (0.51)
セッション 1-6						1.00 (0.00)	0.75 (0.46)
セッション 1-7							0.71 (0.44)

表 10 N = 15 の場合の識別結果 (DTW)

テンプレートデータ	テストデータ		
	セッション 2	セッション 3	セッション 4
セッション 1	0.59 (0.26)	0.58 (0.33)	0.62 (0.34)
セッション 1-2		0.77 (0.21)	0.74 (0.32)
セッション 1-3			0.80 (0.27)

表 11 N = 15 の場合の識別結果 (ランダムフォレスト)[4]

学習データ	テストデータ		
	セッション 2	セッション 3	セッション 4
セッション 1	0.71 (0.42)	0.63 (0.42)	0.48 (0.34)
セッション 1-2		0.81 (0.26)	0.48 (0.36)
セッション 1-3			0.62 (0.36)

表 12 F と L 間の DTW 距離

テンプレート	テスト	
	F ₁	L ₁
F ₃	64.69	
F ₄	57.34	
F ₅	58.46	
F ₆	58.62	
L ₃	39.10	
L ₄	44.23	
L ₅	44.18	
L ₆	40.62	

5. おわりに

本研究では、VR 環境におけるユーザの HMD と両手コントローラの 3 次元座標の時系列データから、DTW 距離を用いて個人を識別する手法を提案した。実験では、12 名の四拍子の指揮の時系列データと 15 名の Beat Saber のプレイの時系列データに DTW を用いた提案手法を適用し、

従来方式と比較して、多くの評価指標において上回る性能を示した。

今後の課題として、被験者数の拡大や 1 動作あたりのフレーム数増減による影響の検証、さらなる識別性能の向上のための DTW と機械学習の組み合わせの検証が必要である。

また、VR 空間でのモーションデータによる識別を防止するには、インターネット上にアップロードするモーションデータの数値をそのまま公開するのではなく、詳細な動きの情報を除去し、元の動作に沿った大まかな動きの情報にするのが有効だと考える。一般化やフレーム率を下げるといった、より粗い粒度のデータにすることで、動作の個人差を低減させることが期待できる。また、モーションデータにノイズを付加することで、統計量を特徴量とした機械学習による識別に対する対策も考えられる。

参考文献

- [1] MRI 三菱総合研究所, “三菱総合研究所, 国内のメタバースの認知・利用に関する研究成果を発表”, 三菱総合研究所, 2023-3-30, (<https://www.mri.co.jp/knowledge/column/20230330.html>, 2024 年 6 月参照).
- [2] Report Ocean, “ヘッドマウントディスプレイ市場は 2030 年に 2,212 億米ドルに達すると予測”, Report Ocean, 2021-11-19, (<https://newscast.jp/news/8368543>, 2024 年 7 月参照).
- [3] Vivek Nair and Wenbo Guo, Justus Mattern, Rui Wang, James F. O’ Brien, Louis Rosenberg, Dawn Song, “Unique Identification of 50,000+ Virtual Reality Users from Head & Hand Motion Data”, the 32nd USENIX Security Symposium, pp.895-910, USENIX, 2023.
- [4] Jonathan Liebers and Christian Burschik, Uwe Gruenefeld, Stefan Schneegass, “Exploring the Stability of Behavioral Biometrics in Virtual Reality in a Remote Field Study”, Vir-

- tual Reality Software and Technology 2023, pp.1-12, VRST, 2023.
- [5] Meta, “Beat Saber - VR リズムゲーム”, (https://www.meta.com/jp/quest/beat-saber/?srsltid=AfmB0orFWBiX8b_m4_ZKaXHp9W1F027nLJFgfcEEcbH3HxRLZxnfVEX0, 2024 年 6 月参照).
- [6] 鈴木芽依, 飯島涼, 野本一輝, 森達哉, “ユーザの自然な動作を利用したヘッドマウントディスプレイユーザ認証方法の提案と評価”, CSS2022, pp.667-674, IPSJ, 2022.
- [7] 森駿文, 菊池浩明, “歩容データの DTW 距離に基づく個人識別手法の提案と外乱に対する評価”, DICOMO2018 シンポジウム, pp.672-680, IPSJ, 2018.
- [8] 森駿文, 菊池浩明, “歩容データの DTW 距離に基づく個人識別における複数部位のフュージョン手法”, CSS2018, pp.1-8, IPSJ, 2018.
- [9] 森駿文, 菊池浩明, “複数の歩容特徴量の DTW 距離に基づくロバストな個人識別手法の提案, 情報処理学会論文誌ジャーナル, 60 号, pp.1538-1549, IPSJ, 2018.
- [10] 當麻僚太郎, 谷口輝海, 菊池浩明, “歩容に基づく個人識別における Kinect と OpenPose の多人数追跡評価”, 情報処理学会研究報告 (2023-DPS-194), Vol.38, pp.1-7, IPSJ, 2023.