

ヘッドセットの動的特徴量の DTW 距離に基づく VR ユーザの個人識別

三浦 晃暉†

明治大学総合数理学部 先端メディアサイエンス学科 菊池研究室†

秒のデータで 73.20% の精度で識別した。

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にされる可能性がある。Nair ら [2] や Lieber ら [3] は、VR リズムゲーム “Beat Saber” [4] 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。

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” [4] 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 [2] 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

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 [3] を実施した。Meta Quest 2 ヘッドセットを使用し、VR リズムゲーム “Beat Saber” をプレイする 15 名の被験者からセッション毎 2 回、指定された楽曲をプレイしたモーションデータを 4 セッション収集した。各セッションのデータの位置と回転の統計量を計算した 18 次元の特徴量を基に構築された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、識別性能を評価した。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、高精度な個人識別を実現する手法を提案した [5]。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 距離を用いて個人識別を行い、等価エラー率 (EER) 0.036 の精度を達成した。

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

†Kikuchi Laboratory, Department of Frontier Media Science, School of Interdisciplinary Mathematical Science, Meiji University.

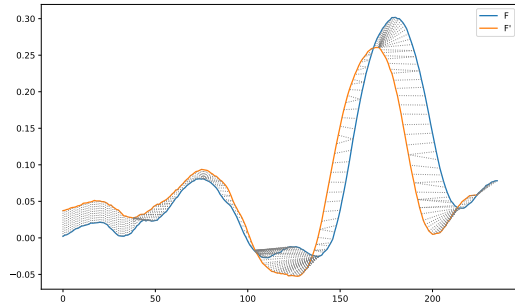


図 1 被験者 F の 1, 2 回目に測定した右手の x 座標のデータの DTW の適用例

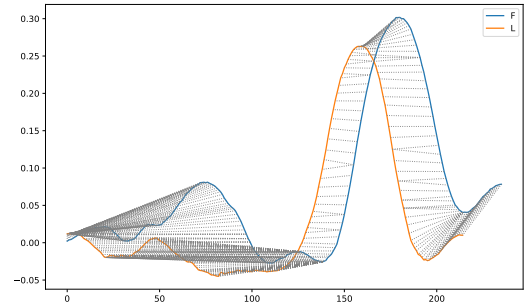


図 3 被験者 F と L の右手の x 座標のデータの DTW の適用例

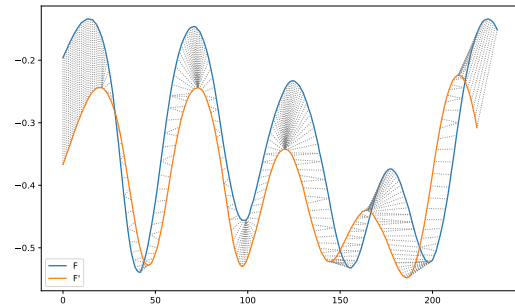


図 2 被験者 F の 1, 2 回目に測定した右手の y 座標のデータの DTW の適用例

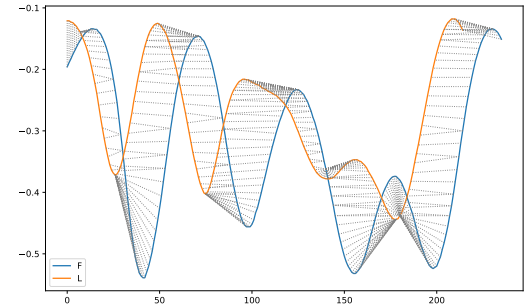


図 4 被験者 F と L の右手の y 座標のデータの DTW の適用例

系列データの DTW 距離を算出して、個人識別を行う。

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 [5] 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 個人識別

与えられた特定の動作のモーションデータに対して、被験者全員のテンプレートデータとの間で DTW 距離

を計算する。最も DTW 距離が小さかったテンプレートデータの該当者を識別結果とする。

4 実験

4.1 実験目的

VR デバイスから DTW 距離に基づく個人識別の精度を明らかにする。

4.2 実験方法

実験環境を表 1 に示す。Meta Quest 3 を用いて HMD・両手のコントローラの 3 次元座標の時系列を得るために、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 人に対し、四拍子を行わせるモーションデータを計測した。四拍子とは図 5 に示す動作である。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を 6 回測定した。

表 1 実験環境

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

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

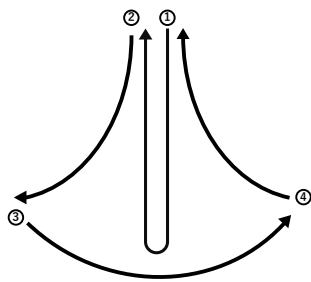


図 5 四拍子の例

また、各モーションデータに対して個人識別を行う。1 回目から 6 回目までのデータを順にテンプレートデータとして使用して、それぞれのテンプレートデータと他のデータとの DTW 距離を計算する。例を挙げると、1 回目のデータをテンプレートデータとした場合、テンプレートデータと 2-6 回目のデータとの DTW 距離を計算する。2 回目のデータをテンプレートデータとした場合、テンプレートデータと 1, 3-6 回目のデータとの DTW 距離を計算する。

4.3 実験結果

1, 2 回目に測定した被験者 F の動作の時系列データに DTW を適用した例を x, y 座標の各動作について図 1, 2 に示す。被験者 F と L の動作の時系列データに DTW を適用した例を x, y 座標の各動作について図 3, 4 に示す。縦軸が座標、横軸がフレーム数を表している。2 曲線間の点線は、DTW 距離を求める際に距離が最小となる 2 点の組み合わせをつないでいる。図 2 の 100 フレーム目で見られるように、1 点が複数の点と対応する場合もある。また、被験者 D, F, L の右手の四拍子の様子を図 6, 7, 8 に示す。これらの図の青点は四拍子の

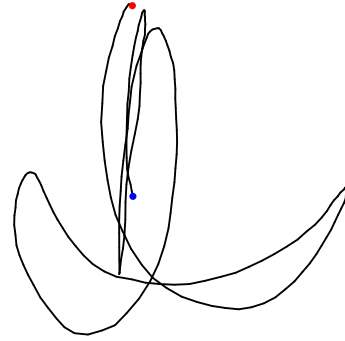


図 6 被験者 D の右手の四拍子の時系列データ (X-Y)

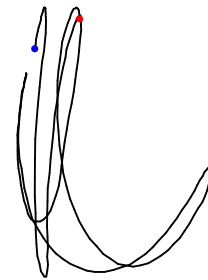


図 7 被験者 F の右手の四拍子の時系列データ (X-Y)

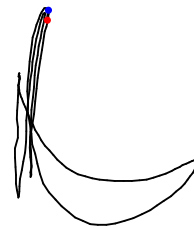


図 8 被験者 L の右手の四拍子の時系列データ (X-Y)

開始点、赤点は四拍子の終了点を示している。被験者 D の四拍子は左右に手を振っているが、被験者 F と L は右側だけに手を振る似た四拍子をしている。四拍子 1 回にかかるフレーム数は約 200 フレームであった。

識別結果を表 3 に示す。計算にかかった時間は 6 時間 12 分 14 秒であった。全体の正解率は 95.56% となった。表 3 を見ると、被験者 C が被験者 I, 被験者 F が被験者 I や L と推定されている。この誤識別の原因は四拍子の手の振り方が似ているからだと考えられる。

表 3 識別結果の混合行列

真 \ 推定	A	B	C	D	E	F	G	H	I	J	K	L	FRR[%]
A	30	0	0	0	0	0	0	0	0	0	0	0	0.00
B	0	30	0	0	0	0	0	0	0	0	0	0	0.00
C	0	0	26	0	0	0	0	0	4	0	0	0	13.33
D	0	0	0	30	0	0	0	0	0	0	0	0	0.00
E	0	0	0	1	29	0	0	0	0	0	0	0	3.33
F	0	0	0	0	0	21	0	0	5	0	0	4	30.00
G	0	0	0	0	0	0	30	0	0	0	0	0	0.00
H	0	0	0	0	0	0	0	30	0	0	0	0	0.00
I	0	0	0	0	0	0	0	0	30	0	0	0	0.00
J	0	0	0	0	0	0	0	0	0	30	0	0	0.00
K	0	0	0	0	0	0	0	0	0	0	30	0	0.00
L	0	0	0	0	0	0	2	0	0	0	0	28	6.67
FAR[%]	0.00	0.00	0.00	0.30	0.00	0.00	0.61	0.00	2.73	0.00	0.00	1.21	

5 おわりに

本実験により、12 名分のデータを 95% を超える精度で識別できることを示した。しかし、計算時間が 6 時間を超えてしまったため、Nair ら [2] の 50,000 を超えるデータに対して本手法で識別を行うのは現実的ではない。DTW と機械学習を組み合わせたアプローチを今後の課題とする。

参考文献

- [1] MM 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.m2ri.jp/release/detail.html?id=600>, 2024 年 6 月参照).
- [2] 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.
- [3] Jonathan Liebers and Christian Burschik, Uwe Gruenefeld, Stefan Schneegass, “Exploring the Stability of Behavioral Biometrics in Virtual Reality in a Remote Field Study”, Virtual Reality Software and Technology 2023, pp.1-12, VRST, 2023.
- [4] Meta, “Beat Saber - VR リズムゲーム”, (https://www.meta.com/jp/quest/beat-saber/?srsltid=AfmB0orFWBiX8b_m4_ZKaXHp9W1F027nLJFgfcEEcbH3HxRlZxnfEX0, 2024 年 6 月参照).
- [5] 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”, DI-COMO2018 シンポジウム, pp.672-680, IPSJ, 2018.
- [6] 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.