

デジタルサイネージ間の連携を想定した人物同定手法の提案

Person identification method for coordination among digital signages

池田 優哉^{*1}, 安留 誠吾^{*2}, 井上 裕美子^{*2}

Yuya IKEDA^{*1}, Seigo YASUTOME^{*2}, Yumiko INOUE^{*2}

^{*1} 大阪工業大学大学院情報科学研究科, ^{*2} 大阪工業大学情報科学部

Email: m1m21a03@st.oit.ac.jp

あらまし：複数のデジタルサイネージの訪問履歴に応じたコンテンツを提供するにあたって、従来手法では複数のデジタルサイネージ間の連携のための人物同定は利用者のデバイスに依存する手法ないしは利用者の個人的特徴を用いる手法などがあり、利用におけるハードルが高くなっていった。

そこで、利用者のデバイスに依存することなく、利用者の人物同定をするために Web カメラで得られる動画のみを利用する。またプライバシーを考慮し、前処理として顔・容姿に結びつかない加工を施した後に機械学習済みのデータを用いる人物同定の手法を提案する。

キーワード：デジタルサイネージ、人物同定、機械学習

1. はじめに

デジタルサイネージはコンテンツを不特定多数の利用者に見せることを目的としており、ディスプレイ技術の発達やデジタルネットワークや無線 LAN の普及に伴って様々な進化を遂げている。また、特定の層に対してコンテンツを選ぶことができる手法としてインタラクティブデジタルサイネージが存在する。対話的なアプローチによって、特定の層のみではあるが、興味を引きやすいという特徴がある。

本研究では、このようなインタラクティブサイネージに着目し、効果をさらに引き出すために、特定個人情報を用いずに人物の同定をする手法を提案する。

2. 関連研究

2.1 利用者に応じたコンテンツの提供をする手法

中島ら[1]の研究では、デジタルサイネージと利用者のスマートフォンの BLE ビーコン機能を用いた連携によって利用者の人物同定を行っている。だが、人物の同定をする場合は、そのためだけに利用者のデバイスにアプリケーションをインストールすることは、利用者に負担を強いることになる問題点がある。

2.2 歩容データによる人物同定手法

森ら[2]の提案手法ではモーションキャプチャーセンサーの Kinect を用いて動作の観測を行い、骨格（スケルトンデータ）の時系列データに（Dynamic Time Warping）DTW を適応し、パターン認識をすることで人物の同定を行っている。だが Kinect は現在販売がされておらず、後継の Azure Kinect は高価であり、人物の同定を行うには複数個必要なため本研究では不適切であると考えた。

3. 提案手法

本研究では歩行を行う人物のスケルトンデータを多クラス分類することで個人を識別する手法を提案

する。

3.1 実験環境

本研究で使用した環境は以下の通りである。

- ・ OS : Ubuntu 20.04 LTS
- ・ 言語 : Python
- ・ GPU ライブラリ : CUDA 11.6
- ・ 機械学習ライブラリ : Pytorch 1.13.1

3.2 データ収集方法

本研究では、図 1 のように、高さ 90cm の位置にカメラを固定し、カメラから 510cm 離れた位置から歩行を開始し、150cm の地点で歩行を終了する映像データを用いる。



図 1 カメラの位置

3.3 3次元骨格推定モデルの選定

本研究の実験環境で用いる 3 次元骨格推定モデルを mmPose[3] の 3 次元姿勢推定モデル(以下 mmPose) と Pre-Trained Spatial Temporal Many-to-One Model for 3D Human Pose Estimation(以下 P-STMO)[4] の 2 つを比較した。

図 2 と図 3 はそれぞれのモデルが同じ映像データを推定した結果のうち、右足の Key Point のカメラに向かってくる方向の軸の遷移である。

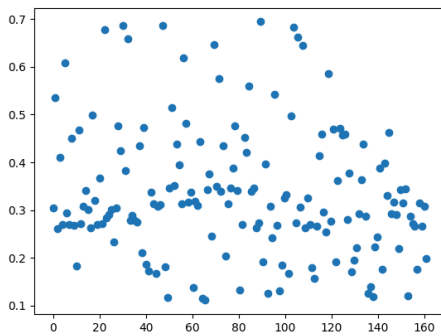


図 2 mmpose の 3 次元姿勢推定結果の遷移

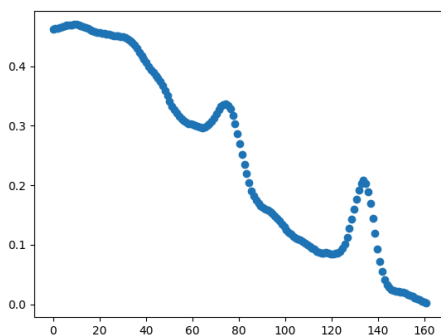


図 3 P-STMO の 3 次元姿勢推定結果の遷移

P-STMO の推定結果は連続して遷移しているが、mmpose の推定結果は非連続な遷移である。

よって本研究で用いる 3 次元骨格推定モデルは P-STMO とした。

3.4 データ前処理

3.2 節で取得した映像データを 3.3 節で選定した 3 次元骨格推定モデルを用いて、3 次元座標におけるスケルトンデータに変換する。

3.5 機械学習モデル

本研究で用いる人物同定を行う機械学習モデルは、51 次元の LSTM 層と全結合層 1 層を繋げたものである。このモデルで多クラス分類として、教師あり学習をする。

4. 実験

4.1 実験内容

実験内容は 3 章の 3.2 節で示した方法で 7 人の映像データを用いて、3.5 節で示した機械学習モデルの精度評価を行う。このときの機械学習のハイパーパラメータとしてエポック数は 10,000 回、100,000 回と 1,000,000 回とした。バッチサイズは機械学習ライブラリの Pytorch により自動設定を用いた。

4.2 評価方法

本研究の実験の評価方法は評価用映像データで人

物同定させて、その結果を Macro-F1 スコアで評価する。

5. 結果と考察

5.1 実験結果

エポック数 100,000 回の結果を表 1 に示す。

表 1 エポック数 100,000 回の際の評価

Recall	1	0	0	0.50	0	0.25	0	
Presition	1	0	0	1	0	1	0	
F1	1	0	0	0.67	0	0.4	0	0.30

5.2 考察

結果より、Macro-F1 スコアは 0.30 となった、このスコアは多クラス分類における精度を評価するものである。0.30 は良いスコアではないが、学習が行えたことを示すことができた。

6. 結論

本研究では、このようなインタラクティブサイネージの効果さをさらに引き出すために、特定個人情報を用いずに人物の同定をする手法を提案した。

今回行った実験では、歩行を行う人物のスケルトンデータを多クラス分類することで個人を識別する手法の有効性を検証するため、7 人の映像データを用いて、3.5 節で示した機械学習モデルの精度評価を行った。結果は Macro-F1 スコアが 0.30 のため、提案手法の有効性は確認されない結果となった。理由としては、サンプル数の少なさと映像データの撮影日が冬で他の季節と比較し厚着のために服で隠れる部分が多くなったため、骨格推定の後の特徴が減ったことだと考えられる。そのため、今後の展望としてはサンプル数の増加と夏に映像データの撮影を行うことで、有効性が確認できるのではないかと考える。

謝辞

本研究の一部は JSPS 科研費 19K03039 の助成を受けたものである。

参考文献

- (1) 中島博敬 ほか :”意図しない BLE 接続を防ぐ デジタルサイネージ・スマートフォン連携システムの構築”, マルチメディア,分散, 協調とモバイルシンポジウム 2016 論文集, vol. 2016, pp.1195–1199(2016)
- (2) 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”, マルチメディア,分散, 協調とモバイルシンポジウム 2018 論文集, vol, 2018,, pp. 672–680(2018)
- (3) mmpose <<https://github.com/open-mmlab/mmpose>> 参照 2023/1/30
- (4) Pre-Trained Spatial Temporal Many-to-One Model for 3D Human Pose Estimation <<https://github.com/patrick-swk/p-stmo>> 参照 2023/1/30