

歩容特徴を用いた人物追跡の研究

三浦 真和[†] 神戸 英利[‡]東京電機大学大学院理工学研究科^{†‡}

1. はじめに

近年、ネットワークカメラ(以後 IP カメラ)を用いたシステムが増加し、需要が高まっている[1]. IP カメラのシステム市場の推移を図 1 に示す.



注1. ベンダー出荷金額ベース
 注2. 2021年は見込値、2022年以降は予測値
 注3. カメラ、サーバ、エンコーダ、NVR/DVR録画装置、VMSソフトウェア、VCAシステム、その他システムを構築するハードウェア、ソフトウェアを対象として算出した。但し、工事費や保守メンテナンス料は含まない。

図 1 監視カメラシステム世界市場規模・予測 ([1]から引用)

IP カメラの具体的な応用方法として、人物追跡がある。通常、IP カメラを用いた人物追跡は複数のカメラ映像を目視で確認するため、時間と労力がかかっていたが、昨今では機械学習技術の発展により、人力での作業が短縮され、的確な分析ができるようになった[2]。しかし人物追跡における課題として、人物の追跡に必要な特徴について容易に活用できない点がある。既存の人物追跡に必要な特徴として顔や服装がある。顔はプライバシーの問題や撮影角度に条件があるため、実用化が難しい。また服装を用いた手法は主に色を用いて人物の追跡を行なっているが、追跡中に着替えなど行い、服装が変わると追跡が困難になる。

2. 研究目的

現状の人物追跡手法は追跡に必要な特徴を取得する上で撮影角度や追跡特徴の変化について対応できない課題がある。そこで本研究では外見

A study of person tracking using gait features
 Masakazu Miura[†] Hidetoshi Kambe[‡]
 Graduate School of Science and Engineering,
 Tokyo Denki University^{†‡}

の変化に左右されない、人物の歩く姿(以後歩容)を人物追跡に使用する特徴とし、異なる角度からの撮影でも個人の識別を可能とすることで追跡を実現することを目的とする。

3. 提案手法

作成予定のシステム概要図を図 2 に示す。



図 2 システム概要図

各 IP カメラを設置し、映像として情報を取得する。取得できた映像はサーバに保存し、システム管理者が確認できるようにする。管理者はサーバ上の映像から追跡対象人物を指定する。指定した人物から歩容特徴を学習し、各 IP カメラ映像から追跡人物の歩容特徴と類似した人物をピックアップする。ピックアップした人物の撮影時間、撮影エリアを管理者に知らせることで、追跡を可能にする。

歩容を識別する方法として、主に 2 種類[3]あり、本研究では複数角度からの撮影を想定しているため、人物の関節を用いたモデルベース手法を使用する。モデルベース手法は人物の関節情報を使うことで、歩容の識別を行う。

本研究では、映像から人物の関節座標を取得する方法として、画像 AI を用いた姿勢推定技術を用いる。この手法で取得できる関節座標は (x, y) の 2 次元座標である。2 次元座標の場合、別角度で撮影された場合に対応できない。

そこで、検出された 2 次元の関節座標について 3 次元化を行う。 (x, y, z) の 3 次元化を行うことで、異なる角度から撮影された場合でも対応するこ

とが可能となる。3次元化手法について、多視点の画像と単眼の画像の2種類があるが、複数のIPカメラで同地点を撮影する状況は限られるため、単眼の画像から関節座標の3次元化を行う。

3次元の関節座標を用いて歩容の識別を行う。使用する特徴量として、下半身の動きが大きく、画像AIから推定を行いやすいため、足幅と角度を取得する。また、歩容全体の動きも特徴とすることでより差別化を図る。

4. 実装

映像を取得するためにマイクロコントローラを用いて映像の制御、及びデータ送信機能を持った監視カメラを作成した。

作成した監視カメラから映像を取得し、映像中から人物を検出する必要がある。そこで人間の物体検出が含まれているHRNetを用いることで、人物の検出及び人物の関節座標の推定を行った。実際に姿勢推定を行なった様子を図3に示す。

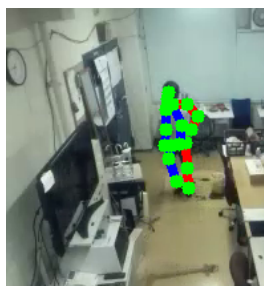


図3 姿勢推定結果の一例

図3から人物の関節座標が推定できているのが確認できる。この取得した関節座標を[4]の手法を用いて3次元化を行なった。図3の姿勢推定を3次元化した結果を図4に示す。

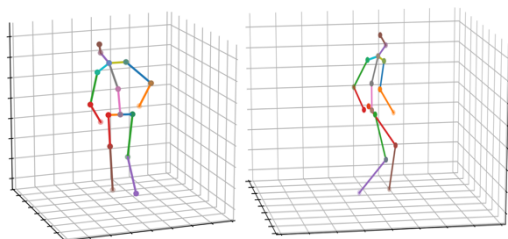


図4 図3の3次元推定結果

図4から図3の関節座標を3次元化できており、複数角度から確認できる。識別器の作成について、今後学習データも含め作成予定である。

5. 評価

歩容の識別精度を評価する上で、識別器の識別精度と歩容情報の取得精度について評価を行う。識別器の出力結果の精度だけでなく、どの精度が

最適なのかについて適合率、再現率を用いて評価を行う。実用を考慮した場合、IPカメラの映像から追跡人物をピックアップできないことを避ける必要がある。そのため、誤検知が増加してでも追跡に必要な人物はピックアップすることを条件に精度評価を行う。また、機械学習技術を用いて関節座標の検知及び3次元化を行なっているため、関節の推測結果についてどれほど妥当であるかについて評価を行う必要がある。

6. 考察

姿勢推定についてHRNetはTop-down手法の姿勢推定手法であるため、検出精度は良いが、物体などに遮られると適切に関節座標を取得できない場合がある。そのため、条件として周囲に物体がない場面での使用に限られてしまう。関節座標の3次元化について単眼画像から行っていることから、距離について適切に推定できない場合がある。原因として検出前後の動作が不安定なことや、距離の推定が適切に行われていない可能性がある。

7. まとめと今後の課題

本研究では、歩容特徴を用いて人物追跡を行うことで、撮影角度や追跡特徴の変化に対応した手法を確立することを目的とした。歩容情報の取得について精度を上昇させる工夫などを行い、評価を行うためシステムの実装を行う。

参考文献

- [1] 監視カメラシステム世界市場に関する調査を実施 (2021年) | ニュース・トピックス | プレスリリース | 市場調査とマーケティングの矢野経済研究所
https://www.yano.co.jp/press-release/show/press_id/2813 (参照 2022-8-4)
- [2] 日立製作所: Hitachi Multifeature Video Search: 日立,
<https://www.hitachi.co.jp/Prod/comp/app/mvs/> (参照 2022-6-1)
- [3] 槇原靖, 村松大吾, 八木康史, “歩容認証とその科学捜査への応用”, IEICE Fundamentals Review. 2021. Vol.14 No.4 pp. 318-328
- [4] Wenhao Li, Hong Liu, Runwei Ding, Mengyuan Liu, Pichao Wang, and Wenming Yang, “Exploiting Temporal Contexts with Strided”, IEEE Transactions on Multimedia 2022