

5S-01

群衆における特定人物の自動検索手法に関する研究

西田義人[†] 梅原喜政[‡] 坂本一磨^{†‡} 鳴尾丈司^{‡‡} 田中成典^{†‡‡} 飛田和輝^{†‡‡} 松村峻平^{†‡‡}
 金沢工業大学基礎教育部[†] 摂南大学経営学部[‡] 公立小松大学生産システム科学部^{†‡}
 関西大学先端科学技術推進機構^{†‡} 関西大学総合情報学部^{†‡‡} 関西大学大学院総合情報学研究科^{†‡‡}

1. はじめに

我が国では、防犯カメラの設置台数が増加し、複数のカメラ映像から人物を追跡して防犯対策やマーケティング戦略に利用[1]されている。しかし、群衆から特定人物を迅速に見つける作業は、多大な人件費や時間的コストに課題がある。近年では、要素技術である FlipReID[2]により、深層学習を用いて人物の外見特徴から人物を同定する技術が注目されている。しかし、検出した人物画像（以下、クエリ画像）が、データベース（以下、人物ギャラリー）に存在することを前提としており、特定人物が人物ギャラリーに未登録の場合、人物ギャラリーのいずれかの人物として、誤った推定をする課題がある。そのため、群衆から特定人物を検索するためには、人物ギャラリーに登録した画像（以下、ギャラリー画像）とクエリ画像の類似度から特定人物が登録済みかどうかを判定する必要がある。そこで、本研究では、群衆から特定人物のみを検索できる手法を提案する。

2. 研究の概要

本研究は、群衆の動画から特定人物を自動で検索する。入力データは、検索対象者の人物ギャラリー構築用動画と群衆を撮影した動画、出力データは人物検索結果とする。本提案手法（図1）は、1) 人物ギャラリー構築機能と 2) 人物検索機能により構成される。

2.1 人物ギャラリー構築機能

本機能では、様々な角度から撮影した検索対象者の画像を選定することで人物ギャラリーを構

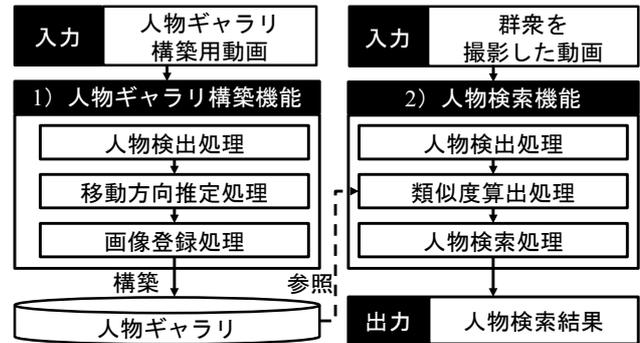


図1 提案手法の流れ

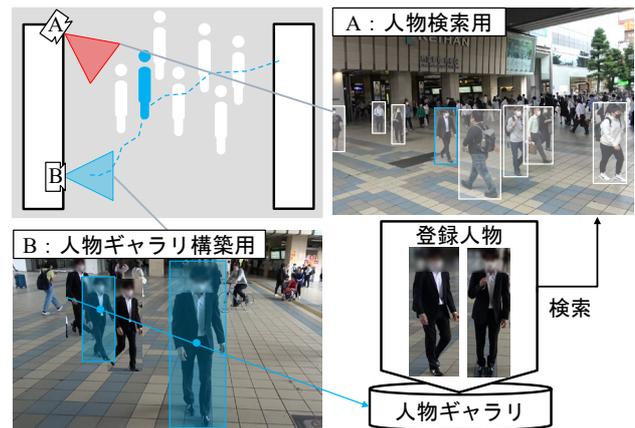


図2 人物検索機能の概要

築する。まず、物体検出手法のYOLOX[3]により人物ギャラリー構築用動画からギャラリー画像を取得する。次に、ViTPose[4]を用いて骨格情報から、多クラス分類モデルより各ギャラリー画像の移動方向を推定する。そして、推定した移動方向毎に骨格情報の各部位の推定信頼度の合計が高い上位2件のギャラリー画像を人物ギャラリー登録する。

2.2 人物検索機能

本機能（図2）では、群衆から取得したクエリ画像と人物ギャラリーを照合して群衆から自動で人物を検索する。まず、YOLOXにより群衆からクエリ画像を取得する。次に、クエリ画像とギャラリー画像をFlipReIDに入力し、外見特徴から類似度を算出する。そして、算出した類似度、クエリ画像の検出位置、画像サイズや色情報を考慮した二値分類モデルにより、人物ギャラリーの人物を群衆から検索する。

Research for Auto-Retrieving Specific Individuals in Crowd
[†] Yoshito Nishita
 Academic Foundations Programs, Kanazawa Institute of Technology
[‡] Yoshimasa Umehara
 Faculty of Business Administration, Setsunan University
^{†‡} Kazuma Sakamoto
 Faculty of Production Systems Engineering and Sciences, Komatsu University
^{‡‡} Takeshi Naruo
 Organization for Research and Development of Innovative Science and Technology, Kansai University
^{†‡‡} Shigenori Tanaka and Ryohei Matsumura
 Faculty of Informatics, Kansai University
^{†‡‡} Kazuki Hida
 Graduate School of Informatics, Kansai University

3. 実証実験

本実験では、JR 京橋駅と京阪京橋駅の乗り換え口（図 3）を対象とし、提案手法の有用性を評価する。

3. 1 実験内容

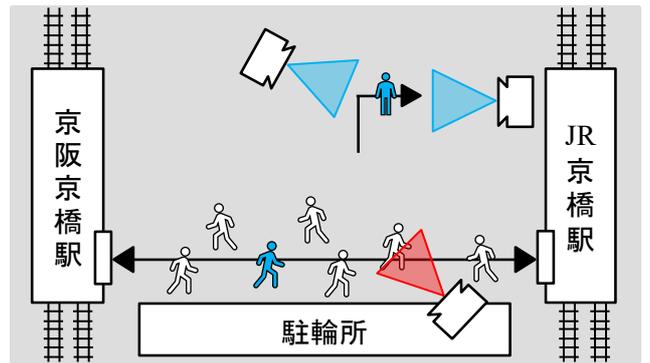
本実験では、通勤時間帯（午前 8 時～午前 9 時）の屋外において、監視カメラの設置位置（高さ 2m～2.5m）を想定してカメラを固定した。まず、検索対象者 3 名（図 4）のギャラリー画像を登録することで、人物ギャラリーを構築する。次に、その 3 名が群衆内で登場する動画からクエリ画像を取得し、人物検索処理を実施することで人物の検索を行う。人物検索処理においては、検索対象の人物を除いた 2 名のデータで二値分類モデルを構築し、人物を検索する。ここで、二値分類モデルの説明変数を人物の類似度、クエリ画像取得位置とカメラ設置位置との距離、画像サイズ、アスペクト比、Lab 色空間の各平均値と定める。そして、群衆から取得したクエリ画像を検索対象として人物を特定する手法を評価し、提案手法の有用性を確認する。

3. 2 結果と考察

人物の検索精度を表 1 に示す。検索精度より、3 名中 2 名において F 値 0.68 以上で検索できたことから、提案手法の一定の有用性を確認した。特に、人物 3 の F 値に着目すると、0.891 と最も高精度で人物の検索ができた。これは、群衆に登場する多くの人物がスーツを着用していたことで、人物 3 の特徴量の差が大きかったため、人物 1 と人物 2 よりも高精度に検索ができたと考えられる。また、人物 1 と人物 2 の再現率に着目すると、0.5 程度と精度が低い。これは、2 人の服装がスーツであり、周りとの特徴量の差が小さかったことが考えられる。さらに、人物 2 は適合率も精度が低い。これは、群衆から取得したクエリ画像において、人物 2 のクエリ画像の多くはオクルージョンの発生により、他の人物も一緒に写りこんだクエリ画像が多かったことに起因している。そのため、人物の検索において、オクルージョン判定から、人物の特徴を正確に抽出できるクエリ画像のみで検索を行い、追跡技術と組み合わせることで、検索精度の向上が期待できる。

4. おわりに

本研究では、カメラごとの二値分類モデルにより、人物ギャラリーに登録されている人物を群衆から検索する手法を提案し、実証実験により一定の有用性を確認した。今後は、様々な現場での追加実験により提案手法の汎用性とオクルージョン時での精度向上を目指す。



凡例
 →: 歩行経路 ▲: 人物ギャラリー構築用
 ◻: カメラ設置箇所 ▲: 人物検索用

図 3 歩行経路



図 4 ギャラリー画像とクエリ画像（一部抜粋）

表 1 人物の検索精度

	適合率	再現率	F 値
人物 1	0.985	0.520	0.681
人物 2	0.425	0.580	0.491
人物 3	1.000	0.804	0.891

参考文献

- [1] 経済産業省：カメラ画像利活用ガイドブック，経済産業省（オンライン），入手先（<https://www.meti.go.jp/press/2021/03/20220330001/20220330001-1.pdf>）（参照 2024-01-11）。
- [2] Ni, X. and Rahtu, E.: FlipReID: Closing the Gap between Training and Inference in Person Re-Identification, 2021 9th European Workshop on Visual Information Processing, IEEE, pp.1-6 (2021).
- [3] Ge, Z., Liu, S., Wang, F., Li, Z. and Sun, J.: YOLOX: Exceeding YOLO Series in 2021, arXiv(online), available from (<https://arxiv.org/pdf/2107.08430.pdf>) (accessed 2024-01-11).
- [4] Xu, Y., Zhang, J., Zhang, Q. and Tao, D.: ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation, Advances in Neural Information Processing Systems (2022).