

移動ロボットによる 3 次元人物追跡のための 2 台のカメラを用いたパーティクルフィルタによる人物検出

○橋本 慧志 (中央大学), 本多 明彦 (中央大学),
池 勇勳 (中央大学), 梅田 和昇 (中央大学)

Human Detection by Particle Filter Using Two Cameras for Three-dimensional Human Tracking with Mobile Robot

○Satoshi HASHIMOTO (Chuo University), Akihiko HONDA (Chuo University),
Yonghoon JI (Chuo University), and Kazunori UMEDA (Chuo University)

Abstract: This paper proposes human detection and tracking method that uses two cameras mounted on a mobile robot. Particle filter and template matching-based approach are used to detect human region in images and three-dimensional (3D) information of the detected human is obtained by stereo matching.

1. 序論

特定の人物の追跡は多岐に渡る自律移動ロボットに求められる機能の 1 つであり、この機能により、ロボットが荷物や資材を運搬しながら特定の人物を追跡することが可能となる。また、これらの汎用的な機能は、工事や災害現場、人間の日常生活など適用範囲は広く、様々な環境での需要が見込まれる。この機能を実現する為に用いられる外界センサの内、カメラを使用する場合は、ステレオカメラが多く用いられる[1]。一般的なステレオカメラを用いる場合、シーン全体の距離画像を極力正確に取得し、その後に得られた距離画像に対してタスクに応じた処理を行う。しかし、これを実現するには正確なカメラキャリブレーションが必要であり、処理速度にも課題が残っている。これに対して、本多らは、個々のカメラで異なる処理を行い、その処理結果を融合することで、特定の人物など必要な対象のみについての 3 次元情報を取得する手法を提案した[2]。具体的なタスクとしては、物体の 3 次元トラッキングを取り上げ、個々のカメラで物体検出を行い、その結果を統合することで 3 次元情報を取得する手法を構築している。この研究では、個々のカメラ画像における物体検出及びトラッキングのために Cam-shift[3]を用いる。しかし、Cam-shift は、基本的に対象の色情報を元にトラッキングを行う手法である上に、照明条件など、環境に合わせてパラメータを自由に設定することが不可能であるため、移動ロボットによる人物追跡においては、変動する環境下での追跡のロバスト性に課題が残っている。

以上の問題を踏まえ、本研究では個々のカメラ画像

における処理に時系列フィルタリングの 1 つであるパーティクルフィルタ[4,5]とテンプレートマッチングを用い、人物の三次元情報を取得してトラッキングを行う手法を提案する。パーティクルフィルタは、環境に合わせてパラメータ設定を変更可能であるため、従来の手法より追跡精度を向上することが可能である。

2. 提案手法の概要

本手法では、同一のカメラを簡易的に平行に並べることで、平行ステレオを実現しており、高精度なステレオマッチングを必要としないため、個々のカメラとして、一般に普及している安価の web カメラを用いて人物の検出を行う。Fig. 1 に本手法の流れを示す。また、Fig. 2 に用いる機器の外観を示す。まず、ターゲットとなる

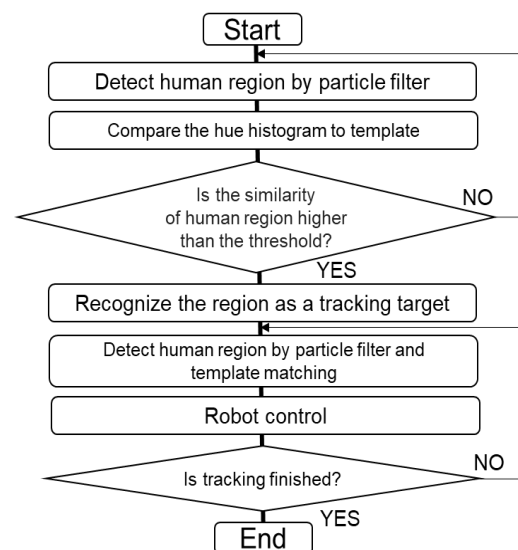


Fig. 1 Flow chart of proposed method.

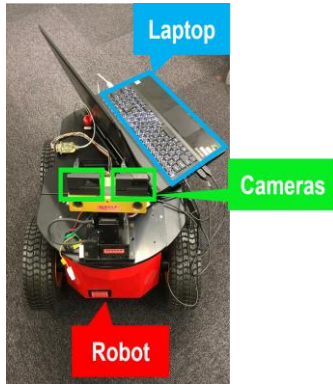


Fig. 2 Cameras arranged in parallel on mobile robot.

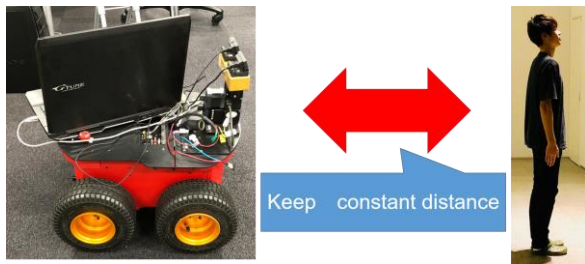


Fig. 3 Example of robot control.

人物の後ろ姿の画像を取得し、画像の色相情報からヒストグラムを作成しテンプレートとする。次に、左のカメラ画像の中で人物領域をパーティクルフィルタにより検出し、得られた人物領域とテンプレート間の色相における類似度を比較する。計算した類似度が一定の値を超えている場合に追跡対象と指定する。その後、左のカメラではパーティクルフィルタ、右のカメラではテンプレートマッチングを用いて追跡対象の検出を行う。同時に、各カメラから検出した対象までの距離計測を行い、算出された距離値を用いて、Fig. 3 のように一定の距離を保つようにロボットの制御を行う。ここで、各カメラにおいてそれぞれ異なった処理を行うことにより、正確な対象人物の 3 次元情報を取得するためのマッチング精度を向上することが可能である。上記の処理を毎フレーム繰り返す。

3. テンプレートの作成

色情報のテンプレートを作成するため、本研究では RGB 画像を HSV 空間に変換して用いる。HSV はそれぞれ色相、彩度、明度を表す。次に、明るさや照明変動の影響を受けにくいとされる色相情報を用いてテンプレートのヒストグラムを作成する。Fig. 4 にテンプレート画像の例並びにその色相のヒストグラムを示す。

4. 人物領域の検出

本章では、追跡対象の人物領域を検出する手法を説明する。まず、左カメラ画像中の、人物領域を、パーティクルフィルタを用いて検出する。すなわち、最初用

意した Fig. 4(a) のようなテンプレート画像を用いて、左カメラ画像にパーティクルフィルタを適用する。その後、推定された座標値周辺の領域を人物領域とする (Fig. 5)。この領域をテンプレートと同様に HSV 空間へ変換し、ヒストグラムを作成する。領域のヒストグラムとテンプレートのヒストグラムの類似度の計算には、バタチャリヤ距離を用いる。2 章で言及した通り、ここで計算した人物領域の類似度が一定の閾値を超えた場合、追跡対象と指定する。パーティクルフィルタにおける詳細の処理は次章に示す。

5. 人物追跡の手法

5.1 左カメラの処理

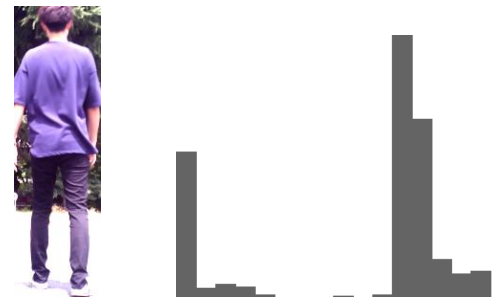
本手法では、人物の検出手法としてパーティクルフィルタを用いる。パーティクルフィルタとは、時系列フィルタリングの一種であり、過去の状態から未来の状態を逐次的に推定できる手法である。本手法でのパーティクルフィルタの処理の流れを以下に示す。

ランダムノイズによる予測

本手法では、次式のようにパーティクルの状態遷移モデルに正規分布に従うランダムノイズ $\mathbf{r}$ を用いる。

$$\mathbf{x}_{t|t-1}^{(n)} = \mathbf{x}_{t-1}^{(n)} + \mathbf{r}_{t-1}^{(n)} \quad (1)$$

ここで、 $\mathbf{x}_{t-1}$ と $\mathbf{x}_{t|t-1}$ はそれぞれ遷移前と遷移後のパーティクルの位置を示す。 $n = 0, 1, \dots, K$ はパーティクルのインデックスであり、 K はパーティクルの数である。



(a) Template image (b) Hue histogram

Fig. 4 Histogram generation for template image.

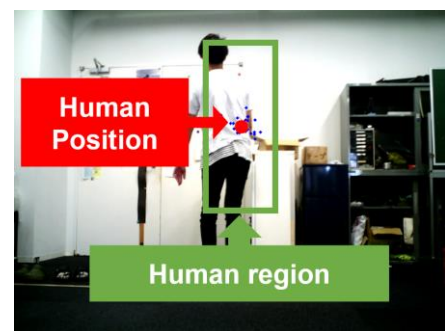


Fig. 5 Detection of human region by particle filter.



Fig. 6 Estimation of likelihood of particles.



(a) Left camera image

(b) Right camera image

Fig. 7 Template matching between two cameras.

観測によるパーティクルへの重みづけ

各パーティクルの重みは、各パーティクルの周辺領域とテンプレート画像との色相の相違度 d をバタチャリヤ距離から K 個のパーティクルに対してそれぞれ計算して評価する。Fig. 5 に処理の例を示す。類似度が得られたら、 K 個のパーティクルに対して、次式を用いて重みを計算する。

$$w_t^{(n)} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\{d(H_1, H_2)_t^{(n)}\}^2}{2\sigma^2}\right] \quad (2)$$

ここで、 H_1 と H_2 は各ヒストグラムを意味し、 σ^2 は正規分布における分散である。

リサンプリング

各パーティクルの重みに基づいて、重みが大きいパーティクルには多くのコピーを生成し、重みが小さいパーティクルは消滅させるか、少ない量のコピーを生成する。

人物の位置推定

パーティクルの重み付き平均を以下の式を用いて計算し、その結果を人物の位置 X_t として出力する。

$$X_t = \frac{\sum_{n=0}^K w_t^{(n)} x_{t|t-1}^{(n)}}{\sum_{n=0}^K w_t^{(n)}} \quad (3)$$

ここで、 $w_t^{(n)}$ は各パーティクルの重み、 $x_{t|t-1}^{(n)}$ は各パーティクルの位置である。

5.2 右カメラの処理

4章で検出した人物領域 (Fig. 5) と、テンプレート画像との色相の類似度を計算し、一定値以上の場合はその領域を追跡対象と指定して、右カメラへの処理を開始する。ここでは、左のカメラで検出した追跡領域を、

右カメラのテンプレートマッチングにおける第 2 テンプレートとして用いる。この際、右カメラのテンプレートマッチングの範囲は、処理の高速化及びマッチング精度向上のため、Fig. 7(b)の赤色矩形で示すように左カメラの追跡領域を少し拡大した範囲で行う。テンプレートマッチングの類似度計算には相関を用いる。

5.3 距離の計測

左のカメラで検出された追跡領域の中心点と、Fig. 7(b)の赤点に示す右のカメラでテンプレートマッチングを行った結果検出された範囲の中心点に着目し、その視差によりターゲットまでの距離 Z を次式により求める。

$$Z = \frac{bf}{\delta(u_l - u_r)} \quad (4)$$

ここで、 $u_l - u_r$ は左右カメラの視差である。 b は左右カメラの基線長、 f はカメラの焦点距離、 δ は画素間隔である。

以上の追跡領域の検出結果に基づき、オンライン処理でロボットを制御し人物を追跡する。ロボットの制御には PID 制御を用い、制御目標値を人物とロボットとの距離として移動ロボットの速度制御を行う。

6. 実験

6.1 実験概要

提案手法の有効性を検証するため、比較実験を行った。実験は屋内の様々な環境下にて行い、個々のカメラ画像における物体検出に Cam-shift を用いた従来手法と、提案手法で 2 台の web カメラを搭載した移動ロボットを用いた人物追跡を行った。Fig. 8 に示す環境の条件は以下の通りである。

- 条件 1：十分に明るく、色相判別に問題ない環境
- 条件 2：十分に明るい、対象が目立たない環境
- 条件 3：暗い領域が存在する環境
- 条件 4：対象に似た色相が存在する環境

Web カメラには ELECOM 製の UCAM-DLI500TN を 2 つ並べて使用した。画像の解像度は 640×480 pixel、画角は 54° であり、基線長は 94 mm、焦点距離は 654 pixel であった。移動ロボットには AdeptMobileRobots 社製の Pioneer3-AT を用いた。パーティクルフィルタの粒子数は 100 個とし、検出された追跡対象が 2.5 m 以上離れると追跡を開始した。

6.2 実験結果

全ての条件において 74 % 以上の精度での追跡を達成することができた。Fig. 9 に実験の様子を示す。本実験では、正確に特定の人物を追跡できた割合を定量的に評価するため、式 (5)、(6) を用いた [1]。

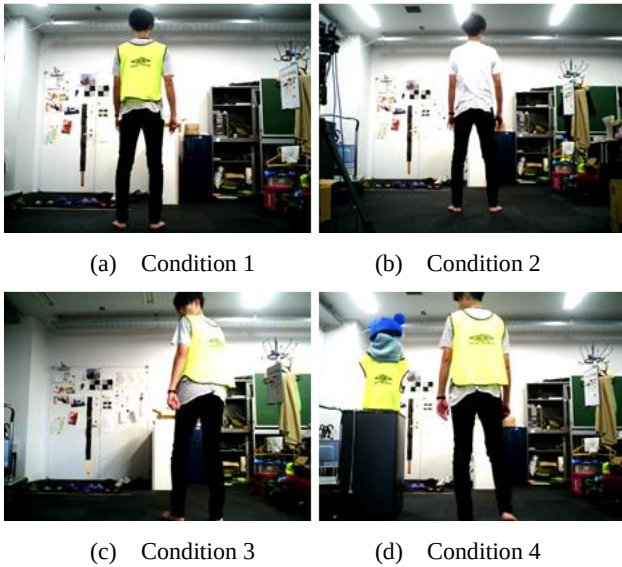


Fig. 8 Examples of environmental conditions.

Table 1 Evaluation of results in indoor environments

Condition	Proposed method		Previous method	
	Precision [%]	Recall [%]	Precision [%]	Recall [%]
1	98	90	100	96
2	100	76	49	45
3	100	74	48	59
4	82	83	27	74

$$P = \frac{A}{A + B} \quad (5)$$

$$R = \frac{A}{A + C} \quad (6)$$

ここで、 A は追跡対象人物が正しく検出されたフレーム数、 B は追跡対象人物以外が検出されたフレーム数であり、 C は追跡対象人物が存在していたが何も検出されなかったフレーム数である。 P は適合率、 R は再現率を意味する。実験の結果得られた各評価値を Table 1 に示す。

6.3 実験考察

全ての条件下で 74 %以上の適合率、再現率が得られた。実験中に一度途中で人物が止まって歩き出した場合も問題なく追跡することができた。条件 2, 3 における再現率の低下は、未検出フレームの多さが原因である。これは、条件 2 では、対象者の服装に色の特徴がないため、人物領域の類似度が上昇せず未検出になったと考えられる。条件 3 では、暗い領域に入った際の、色相情報の若干の変動が類似度の低下を招き、未検出が増加したと考えられる。また、従来手法と比べて、環境に合わせたパラメータ設定を適用が可能となったために、追跡精度の向上につながったと考えられる。

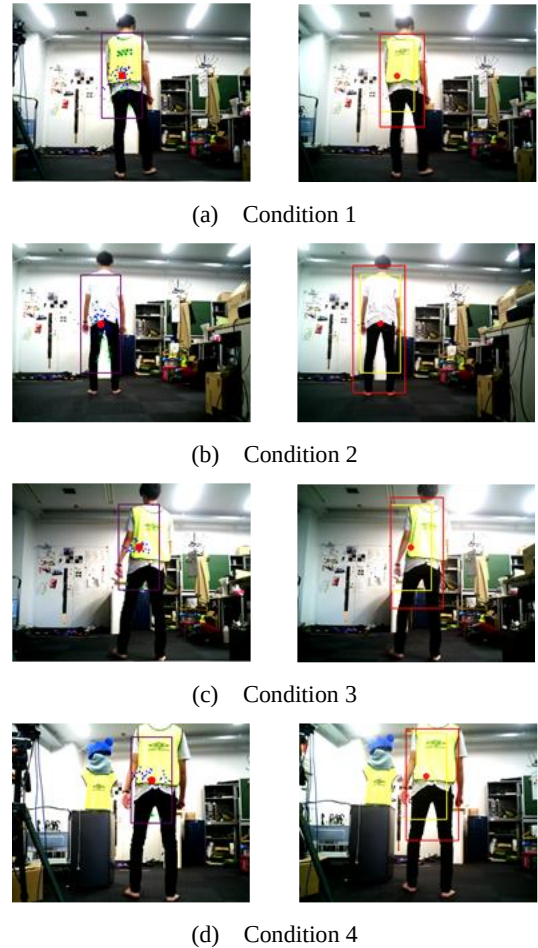


Fig. 9 Experimental results.

今回の実験では、比較的安定した屋内の照明環境下で行ったことや、人物のオクルージョンを考慮していないことから、高い精度での追跡を実現できた。また、屋内の環境であれば多少の変動にもロバストに追跡可能であることが確認でき、条件 2-4 においても従来手法より精度を大きく向上させることが可能であった。

7. 結論及び今後の展望

本報では、個々のカメラで異なる処理を行い、その処理結果を融合することで、必要な対象のみに関しての 3 次元情報を取得し、移動ロボットによる追跡を行う手法を提案した。具体的なタスクとして物体の 3 次元トラッキングを取り上げ、個々のカメラで物体検出を行い、その結果を統合することで 3 次元情報を取得する手法を構築した。さらに、パーティクルフィルタを用いた人物検出により、必要なタスクに適した 3 次元情報を得られ、従来手法より高い精度での特定の人物トラッキングが可能であることを実験により確認した。

今後はパーティクルフィルタの状態遷移モデルの改善や、尤度推定手法の改善を行い、照明変動が顕著に表れる屋外環境下、複数人のオクルージョンが発生する環境下でのロバストな追跡の実現を目指す。

参考文献

- [1] 磯邊 柚香, 増山 岳人, 梅田 和昇, “ステレオカメラを用いた追跡対象領域の抽出に基づく移動ロボットの人物追跡,” 日本機械学会論文集, Vol. 82, No. 834, pp. 1-18, 2016.
- [2] 本多 明彦, 梅田 和昇, “二台のカメラのトラッキングの融合による移動ロボットの3次元人物追跡,” 日本機械学会ロボティクス・メカトロニクス講演会 2018講演概要集, 1A1-M09, p. 30, 2018.
- [3] Gary R. Bradski, “Computer Vision Face Tracking For Use in a Perceptual User Interface,” *Intel Technology Journal*, Q2, pp. 1-15, 1998.
- [4] 樋口知之, “粒子フィルタ,” 電子情報通信学会誌, Vol. 88, No. 12, pp. 989-994, 2005.
- [5] 倉爪 亮, 山田 弘幸, 曾我部 光司, 村上 剛司, 岩下 友美, 長谷 川, “SIR/MCMCパーティクルフィルタを用いた分散カメラとレーザによる複数移動体の同時追跡,” 日本ロボット学会誌, Vol. 28, No. 1, pp. 65-76, 2010.