

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

-OpenPoseによる人物検出の性能向上-

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

3D Human Tracking Using Mobile Robot by Fusion of Tracking Results from Two Cameras

-Performance Improvement of Human Detection by OpenPose-

○Akihiko HONDA (Chuo Univ.), Yonghoon JI (Chuo Univ.), and Kazunori UMEDA (Chuo Univ.)

Abstract: This paper proposes a novel process that use two cameras to obtain three-dimensional (3D) information of a target object for tracking. Object detection and tracking results from two cameras are integrated to get the 3D information. In this paper, the performance of human detection is improved by OpenPose.

1. 序論

自律移動ロボットは、工場や災害現場、人間の日常生活等、様々な場面での活用が期待されている。このロボットに求められる機能の1つに、特定人物の追跡が挙げられ、この機能の実現を目的とした研究が盛んに行われている。これらの研究にはステレオカメラがしばしば用いられる[1]。

一般的なステレオカメラでは、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っている。しかしながら距離画像を正確に取得するためには正確な事前キャリブレーションが必要であり、誤マッチングが生じることや計算コストがかかるといった問題が存在する。

これに対し、我々は個々のカメラでタスクに応じた処理を行い、その後に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次元情報を取得するという手法を提案している[2]。本論文では個々のカメラでの物体検出・トラッキングにOpenPose[3]を使用して人物検出の性能を向上させることにより、従来の手法では対応できなかった環境下での人物の追跡を実現する。

2. 3次元人物追跡手法の概要

本手法の流れをFig.1に示す。本手法では、移動ロボット上に同一種類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する。同一種類のカメラには現在普及している安価なカメラを用いて人物の検出を行う。処理の流れとしては、まず、ターゲットとなる人物の追跡対象となる後ろ姿の画像を取得し、画像の色

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、OpenPoseライブラリを用いてカメラ画像中の人物領域を検出する。検出した範囲内の色情報から作成したヒストグラムとテンプレートのヒストグラム間の類似度を計算する。その類似度が一定の値を超えている場合、追跡の対象と指定する。その後OpenPoseライブラリを用いて指定した対象の追跡を行う。同時に指定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ロボットの制御を行う。これを毎フレーム繰り返す。

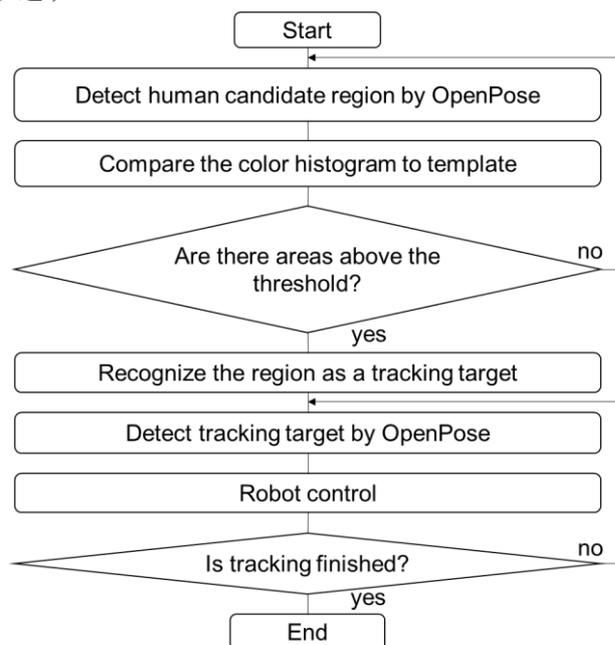


Fig.1 Flow chart

3. 人物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、取得した 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、照明変動の影響を受けにくい色相情報のみを用いてテンプレートのヒストグラムを作成する。本手法では人物の肩より下の部分をテンプレートとして用いる。Fig. 2 にテンプレートの例並びにその色相のヒストグラムを示す。

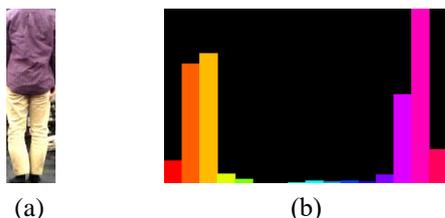


Fig. 2 Histogram generation for template image: (a) Template image and (b) Color Histogram

4. 人物領域の検出

人物領域の検出には、CNN (Convolution Neural Network) を用いて単一画像から複数の人間の体や顔のキーポイントをリアルタイムに検出することができる OpenPose ライブラリを使用する。OpenPose により検出されたキーポイント (Fig. 3(a)) のうち、左右の肩と足首の座標を用いてカメラ画像を切り抜き、切り抜いた領域 (Fig. 3(b)) 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。作成した色相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 d を以下の式で算出する。

$$d(H_1, H_2) = \frac{\sum_l (H_1 - \bar{H}_1) (H_2 - \bar{H}_2)}{\sqrt{\sum_l (H_1 - \bar{H}_1)^2 \sum_l (H_2 - \bar{H}_2)^2}} \quad (1)$$

ここで、 H_1 はテンプレートのヒストグラム、 \bar{H}_1 はテンプレートのヒストグラムの平均値、 H_2 は初期人物候補領域のヒストグラム、 \bar{H}_2 は初期人物候補領域のヒストグラムの平均値である。類似度が一定の閾値を超えている場合は追跡対象とし、テンプレートを更新する。これを毎フレーム繰り返す。

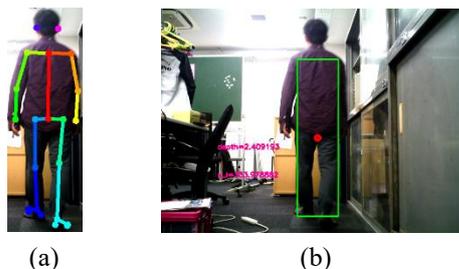


Fig.3 Detection of human area: (a) Detection of human body keypoints by OpenPose and (b) Examples of an area used to calculate histogram similarity

5. 人物の検出結果に基づく追跡

以上の手法で左右のカメラにおいて追跡領域を検出する。片方のカメラで検出された範囲の中心点と、もう片方のカメラで検出された範囲の中心点 (Fig.4 における赤点) に着目し、その視差によりターゲットまでの距離を以下の式より求める。

$$Z = \frac{b \cdot f}{\delta(u_l - u_r)} \quad (2)$$

ここで、 u_l, u_r はそれぞれ左右カメラの中心点の横座標であり、 $u_l - u_r$ 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。

以上の追跡対象人物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方位として移動ロボットの速度制御を行う。



Fig.4 Examples of points used for distance calculation: (a) Left camera image and (b) Right camera image

6. 実験

6.1 実験の概要

提案手法の有効性を検証するため、実験を行った。また、従来の手法[2]と比較して優位性を示すため、従来の手法では追跡が不可能な環境において実験を行った。従来の手法では、照明が変化してしまうと追跡対象とする人物を正しく認識することが出来なくなってしまう、追跡することが不可能であった。そこで今回は照明が変化する環境において実験を行った。

実験は屋内の二箇所にて行い、一箇所は比較的明るくかつ照明が変化する環境、一箇所は比較的暗くかつ照明が変化する環境において Web カメラと移動ロボットを用いた人物追跡を行った。

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 94mm とした。移動ロボットには AdeptMobileRobots 社製の Pioneer3-AT を用いた (Fig.5)。人物とロボットとの間の距離が一定の値を超

えた場合にロボットは人物に向かって移動を開始する。
Fig.6 に実験の概念図を示す。



Fig. 5 Pictures of two identical cameras and robots

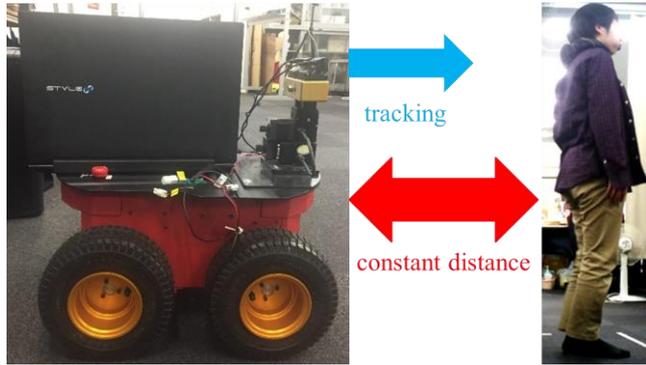


Fig.6 Conceptual diagram of target tracking with a mobile robot

6.2 実験結果

カメラを固定し、実際にロボットに人物を追跡させながら処理を行った。ロボットと対象の人物が2m離れた場合にロボットが進むようにしたところ、どちらの実験においても問題なく追跡することができた。

Fig.7に比較的明るくかつ照明変化する環境で行った実験の様子を、Fig.8に比較的暗くかつ照明変化する環境で行った実験の様子を示す。



(a)



(b)



(c)

Fig.7 Experimental results of two cameras in a relatively bright indoor environment: (a) just before the robot moves, (b) while moving, (c) immediately after the robot stops.



(a)



(b)



(c)

Fig.8 Experimental results of two cameras in a relatively dark indoor environment: (a) just before the robot moves, (b) while moving, (c) immediately after the robot stops.

6.3 考察

一回目の実験 (Fig. 7) では明るい環境であったものの途中で照明の影響を受けて追跡対象の人物が明るく見える場合や暗く見える場合がある。Fig. 7 (a)は追跡開始時の様子であるが、このときはロボットから見て追跡対象の人物が暗く見えている。一方、Fig. 7 (b)はロボットが追跡を開始してから一定の時間が経ち、人物が進んだときの様子であるが、このときはロボットから

見て追跡対象の人物が明るく見えている。今回の実験はこうした照明による変動が何度も起きる場所での実験である。このような照明による変動が起きた場合においても追跡対象の人物が前進して一定距離離れるとロボットが対象の人物を追跡することが出来た。

二回目の実験 (Fig. 8) は一回目の実験と同様に途中で照明の影響を受けて追跡対象の人物が明るく見える場合や暗く見える場合があり、さらに一回目の実験より暗い環境での実験である。Fig. 8(a)は追跡開始時の様子であるが、このときはロボットから見て追跡対象の人物が一回目の実験よりも暗く見えている。一方、Fig. 8(b)はロボットが追跡を開始してから一定の時間が経ち、人物が進んだときの様子であるが、このとき周囲は暗いものの付近にある照明の影響によりロボットから見て追跡対象の人物が明るく見えている。今回の実験はこうした照明による変動が繰り返し起きる場所での実験である。このような照明による変動が起きても人物を追跡することが出来た。またどちらの実験においても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出来た。また、Fig. 7(c)とFig. 8(c)に示すように、どちらの実験でもロボットが停止した地点では人物全体が写っていないが、これは人物が停止した際に人物とロボットが近づき、ロボットが停止すると判断してから実際にロボットが停止するまでの間にロボットが進んでしまうためである。ロボットが停止すると判断した時点では必要な部分が映っていることから、追跡に関して問題はないと考えられる。

7. 結論と今後の展望

本論文では、自律移動ロボットに求められる機能の1つである特定人物の追跡に関して性能の向上を図った。既に提案している手法では、カメラを2台使用し、それぞれのカメラでタスクに応じた処理を行い、その後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次元情報を取得している。今回は、個々のカメラでの物体検出・トラッキングにOpenPoseを使用することで、人物検出の性能を向上させた。その結果、従来の手法では追跡することが不可能であった環境下において人物が追跡できることを示した。

今回の実験は周囲に他の人物が映り込んでおらず、人物を追跡しやすい環境であったと考えられることから、多くの人物が存在する環境下でも追跡可能であるかどうかの検証が必要であると考えられる。

今後は複数人が映っている状況やオクルージョンが発生する状況などの多様な状況での追跡の実現を目指す。

参考文献

[1] Yuzuka Isobe, Gakuto Masuyama, Kazunori Umeda: "Occlusion handling for a target-tracking robot with a

stereo camera", Robomech Journal, Vol.5, No.4, pp.1-13, 2018.

- [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 講演論文集, 1A1-M09, 2018.
- [3] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," Proc.of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.