

二台のカメラを用いた OpenPose の骨格抽出による 移動ロボットの 3 次元人物追跡

Three-dimensional human tracking of a mobile robot using two cameras
by skeleton extraction using OpenPose

○学 松原 慎也 (中央大) 正 池 勇勲 (中央大)
正 梅田 和昇 (中央大)

Shinya MATSUBARA, Chuo University, matsubara@sensor.mech.chuo-u.ac.jp
Yonghoon JI, Chuo University
Kazunori UMEDA, Chuo University

In this paper, a new stereo vision framework that combines the processing results of two cameras and obtains only the three-dimensional (3D) information necessary for each task is used. Under this framework, in a complex environment where there are multiple people, we aim to realize 3D tracking of a specific person by performing person detection using OpenPose. The human region image is divided into upper body and lower body from skeleton information extracted by OpenPose. A specific person is tracked by identifying the person based on the color of the clothes. In this paper, we propose a tracking method and the effectiveness of the tracking method is verified quantitatively by experiments.

Key Words: Human tracking, Stereo camera, Autonomous mobile robot, Three-dimensional measurement

1. 序論

自律移動ロボットは、人間の生活を支援するロボットとして工場や災害現場、日常生活等、様々な場面での活用が期待されている。これらのロボットに求められる機能の一つに、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[1]。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[2]。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。

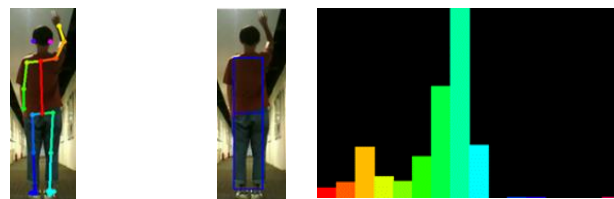
これに対して我々は、二台のカメラがそれぞれでタスクに応じた処理を行い、その後処理結果を融合することで、必要な対象のみに関しての 3 次元情報を取得するという枠組みを提案している[3]。画像全体で距離を求める必要がないことから、誤マッチングを抑制し、必要なタスクにとって最適な 3 次元情報を得ることが可能である。具体的なタスクとして特定の人物の三次元追跡を取り上げる。先行研究[4]では、この新たな枠組みの下で、人物検出に OpenPose[5]を用いた人物の 3 次元追跡を提案した。しかし、この研究で対象としていたのは、画像内に追跡対象者以外の人物が映り込んでおらず、オクルージョンが発生していない、人物を追跡しやすい環境であった。

そこで本論文では、複数人物が存在する環境下での OpenPose を用いた特定人物の 3 次元追跡を実現する。

2. 追跡手法の概要

2.1 初期人物テンプレートの作成

追跡対象者は、意図を伝えるために挙手をするものとする。カメラ起動時に手を挙げている人物の領域を最初のテンプレート画像とする。人物の挙手の検出には OpenPose を用いる。OpenPose とは、CNN (Convolution Neural Network) を用いて単一画像から複数の人間の体や顔などの複数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ライブラリである。左右のカメラそれぞれでテンプレート画像を作成する。テンプレート画像は RGB 画像を HSV 画像に変換して用いる。次に、変換したテンプレート画像の HSV の値をヒストグラム化する。本研究では、明るさや照明変動の影響を受けにくいとされる色相 H のみを用いる。Fig. 1 に初期テンプレート画像並びにその色相のヒストグラムを示す。

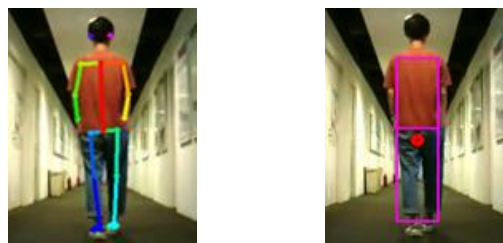


(a) 挙手検出 (b) テンプレート画像 (c) 色ヒストグラム

Fig. 1 テンプレート画像の色ヒストグラム生成

2.2 OpenPose を用いた人物候補領域の検出

人物候補領域の検出にも OpenPose を用いる。OpenPose により検出されたキーポイントのうち、左右の肩と足首の座標を用いてカメラ画像中の人物候補領域を切りとる。Fig. 2 に OpenPose による骨格抽出とその人物の候補領域を示す。



(a) OpenPose による骨格検出 (b) 人物候補領域

Fig. 2 人物候補領域の検出

2.3 人物候補領域の分割

Fig. 3 に示すように、OpenPose によって検出した人物候補領域を分割する。取得したキーポイントのうち腰の座標を用いて上半身と下半身を分割することで、上着と下着それぞれで色の比較が可能となる。

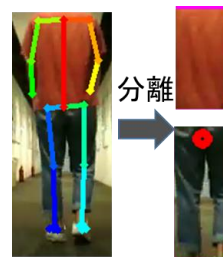


Fig. 3 人物候補領域の分割

2.4 類似度比較

本論文では、服の色の違いから人を識別する．そこで、服の色情報を用いて、検出した人物候補領域画像の色ヒストグラムと作成したテンプレート画像の色ヒストグラムで類似度比較を行う．類似度計算には正規化相互相関を使用する．人物領域画像もテンプレート画像と同様に HSV 画像に変換し、色相ヒストグラムを作成する．作成した現フレームの色相ヒストグラムとテンプレート画像の色相ヒストグラムの類似度 d を以下の式で算出する．

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

H_1 はテンプレート画像のヒストグラム、 H_2 は人物候補領域のヒストグラムである．2つの色ヒストグラムを比較し、類似度が一定の閾値を超えている場合はその人物候補領域を追跡対象とする．

2.5 テンプレート画像作成と更新

現フレームで決定した人物候補領域画像を次フレームでの類似度比較に利用するテンプレート画像とする．つまり、毎フレームでテンプレート画像を更新する．また、テンプレート画像が 1 枚だと急な照明変動に対応できない場合がある．そこで、Fig. 4 に示すように、直近のフレームでのテンプレート画像を複数枚用意する．現フレームの人物領域画像と複数枚のテンプレート画像との類似度を求め、最大の類似度を用いる．以上の処理により、照明環境の変化にロバストな追跡が可能となる．

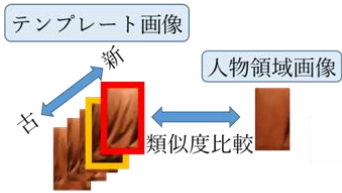


Fig. 4 テンプレート画像と人物領域画像の類似度比較

2.6 人物の検出結果に基づく追跡

以上の手法で左右のそれぞれのカメラで特定の人物領域を検出する．2台のカメラで検出されたそれぞれの人物領域の中心点 (Fig.2(b))における赤点)に着目し、その視差により三角測量の原理を用いて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、 $u_l - u_r$ は左右カメラでの視差である． 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．以上で求めた距離情報を用い、オンライン処理で移動ロボットを制御し、特定の人物を追跡する．人物と移動ロボットとの距離と方位を入力して移動ロボットの速度制御を行う．

3. 実験

3.1 実験条件

カメラには ELECOM 製 UCAM-DLI500TN を使用し、左右カメラの基線長を 95mm とした．移動ロボットには Adept Mobile Robots 製の Pioneer3-AT を用いた．比較的明るくかつ照明変動がある屋内環境において、オクルージョンが 5 回発生する条件の元で、人物追跡を行った．比較時の類似度の閾値は 0.85、テンプレート画像は 5 枚とした．

3.2 実験結果

追跡実験の結果を Table 1, Fig.5~10 に示す．Fig.5, Fig.6 のように急な照明変動が発生した場合も追跡を継続することができた．また、Fig.7~9 のようにオクルージョンが発生した場合にも追跡を復帰することが可能である．Fig.10 は人物候補領域を複数箇所検出したシーンを表している．ここでは、類似度が閾値を超えた人物領域を赤紫枠、閾値を下回った領域を緑枠で描画している．このように、人物が複数人存在する環境下でも人物を識別し、特定の人物を追跡することができている．また、Table 1 からわかるように、高い追跡成功率を実現することができた．しかし、両方のカメラで追跡が成功しないと移動ロボットでの追跡が行えないことから、追跡成功率が低くなった．

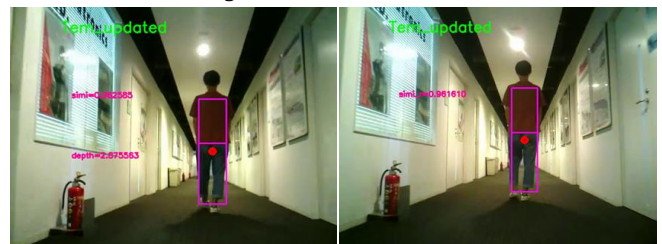
Table 1 対象者追跡成功率

	左カメラ	右カメラ	追跡システム
対象者が存在したフレーム数	167	166	164
対象者追跡失敗数	2	10	11
追跡成功率[%]	98.8	94	93.4



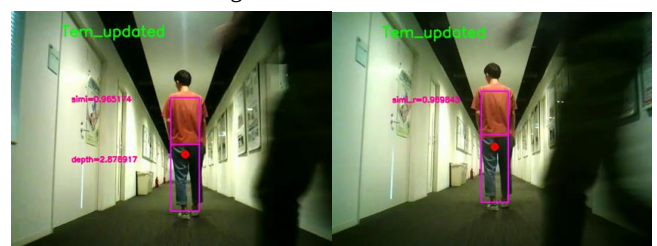
(a) 左カメラ (b) 右カメラ

Fig. 5 比較的明るい環境



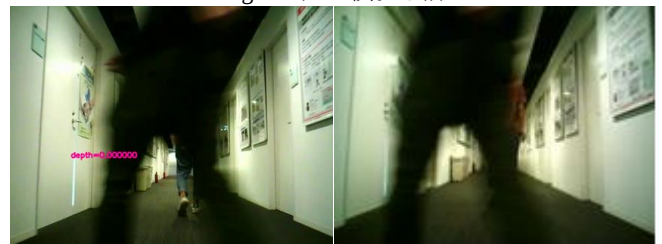
(a) 左カメラ (b) 右カメラ

Fig. 6 比較的暗い環境



(a) 左カメラ (b) 右カメラ

Fig. 7 人が横切る前



(a) 左カメラ (b) 右カメラ

Fig. 8 人が横切っている最中



(a) 左カメラ

(b) 右カメラ

Fig. 9 人が横切った後



(a) 左カメラ

(b) 右カメラ

Fig. 10 人物を複数人検出した場面

4. 結論と今後の展望

本論文では、複数人物が存在する環境下において、人物検出に OpenPose を用いて 2 台の単眼カメラの処理結果を融合した人物の 3 次元追跡を実現し、定量的に評価した。また、人物領域の分割やテンプレート画像を複数枚用意するといった手法を用いることで、画像内に追跡対象者以外の人物が映り込んでいる場面やオクルージョンが発生する場面でも、特定の人物の追跡に成功した。現状では、両眼での追跡が成功しなければロボットによる追跡が行えないという課題がある。今後は、この課題を解決するため、単眼のみでも追跡が行える手法を構築したい。また、色情報以外の特徴量の利用により、服が類似する人物が複数存在する場合など様々な環境での追跡の実現を目指す。

参考文献

- [1] C. Hu, X. Ma, and X. Dai, "A Robust Person Tracking and Following Approach for Mobile Robot", Proc. of the 2007 IEEE Int. Conf. on Mechatronics and Automation, (2007), pp. 3571-3576.
- [2] 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.
- [3] 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.
- [4] 本多 明彦, 池 勇勲, 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", 第 19 回計測自動制御学会システムインテグレーション部門講演会講演論文集 (SI2018), 3D2-15, pp.3017-3020, 2018.
- [5] 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.