

2台のカメラを用いた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 by a Mobile Robot via Skeleton Extraction from Two Cameras

精密工学専攻 57号 松原慎也
Shinya Matsubara

1. 序論

近年、自律移動ロボットは工場現場での運搬作業や高齢者向けの介護支援といった、様々な場面での活用が期待されている。これらの自律移動ロボットにおける重要なタスクの一つに、特定人物の追跡が挙げられ、該当人物を安定して追跡する性能の向上に向けた研究が数多く行われている[1]。この追跡機能を実現するための研究には、ステレオカメラやセンサを用いることが多く、取得した三次元情報を用いることで移動ロボットによる追跡を実現している[2]。

我々は新たなステレオ視の枠組みとして、シーン全体の距離画像は取得せず、単眼カメラ2台がそれぞれでタスクに応じた処理を行った後にその処理結果を融合することで必要な対象のみについての三次元情報を取得するという枠組みを提案した[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三次元情報を得ることが可能である。この具体的なタスクとして特定の人物の3次元追跡を取り上げる。

本論文では、安価なウェブカメラのみを利用したこの枠組みの下で、人物検出に OpenPose[4]を用いた移動ロボットによる3次元人物追跡を提案する。そのとき様々な状況が想定される実利用時を考慮すると、ロバストな人物追跡の実現には様々な課題がある。例えば、急な照明変動などが起こることで2台のカメラのうちどちらか一方のカメラで追跡対象者を見失う、そして曲がり角のように追跡対象者がカメラからフレームアウトすることで追跡の継続が不可能となるという二つの課題がある。そこで2台のカメラのうちどちらか一方が追跡に成功していれば移動ロボットによる3次元人物追跡を行う追跡復帰機能、ならびに追跡対象者の姿勢の変化を骨格抽出により検知して移動ロボットを制御する曲がり角追跡機能を実装し、それらの機能を備えた新たな骨格抽出による人物追跡システムを構築する。

2. 提案手法

2.1 人物追跡手法概要

人物追跡の全体の流れを図1に示す。人物検出、人物識別、距離計測、ロボット制御を毎フレーム繰り返すことで、人物追跡を実現する。図2に示すように、本追跡手法では、人物検出に単一画像から複数の人物の骨格情報をリアルタイムに検出することができる OpenPose[4]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骨格点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。この章では、OpenPoseを用いた人物追跡手法について説明する。

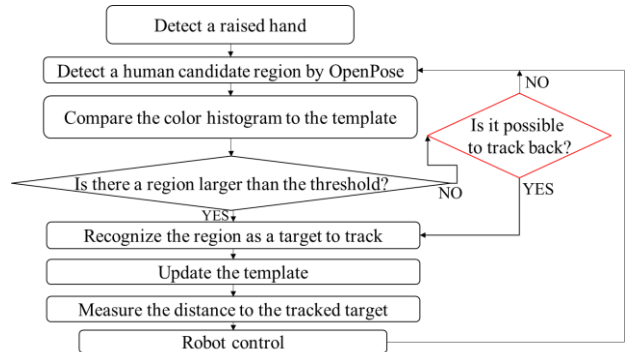


Fig.1 Flow chart of human tracking



Fig.2 Detection of human using OpenPose

2.2 初期テンプレートの作成

まず初めに、追跡対象者には挙手をしてもらう。OpenPoseで手を挙げている人物を検出し(図3(a))、その人物の領域画像を初期テンプレート画像とする(図3(b))。このテンプレート画像をHSV画像に変換し、HSVのうち明るさや照明変動の影響を受けにくいとされる色相(H)のみで表された色ヒストグラム(図3(c))を用いる。このテンプレート画像は後の人物識別を行う際に用いる。

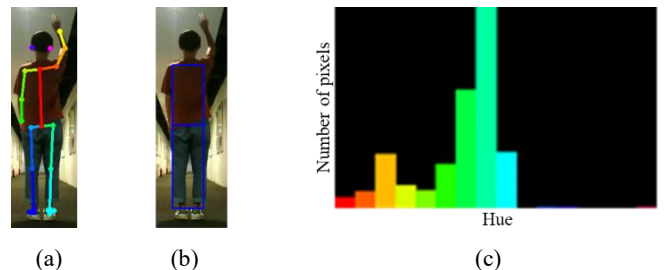


Fig.3 Histogram generation for template image: (a) detection of raised hands (b) template image (c) color histogram

2.3 骨格抽出による人物候補領域の検出と分割

人物候補領域の検出に OpenPose を用いる。図4に示すように、OpenPoseにより得られた骨格情報のうち、左右の肩と

足首の座標を用いて人物領域画像を切り取る。人物候補領域をピンクの矩形で表す。さらに人物候補領域を分割する。取得した骨格座標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り、人物識別精度が向上する。

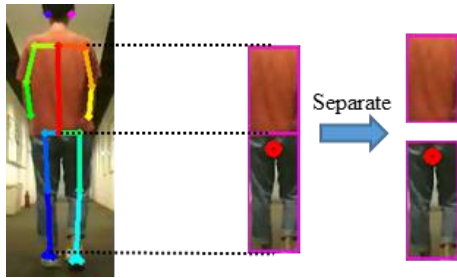


Fig.4 Detection and separation of human region

2.4 服の色比較による人物特定

服の色の違いから人を識別し、追跡対象者を一人に特定する。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を比較し、類似度を計算する。類似度計算には正規化相互相関を使用する。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。

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

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。類似度が一定の値を超えている場合、その人物候補領域に対応する人物を追跡対象者と決定する。

2.5 距離情報を用いた移動ロボットによる人物追跡

ここまでの処理を2台のカメラで行い、それぞれで追跡対象者を特定する。その後、図5に示すように追跡対象人物の領域の中心点(赤点)に着目し、この点までの距離 Z を求める。2台のカメラの視差(横方向のずれ)により三角測量の原理を用いて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。以上で求めた人物と移動ロボット間の距離と位置の情報を用いてロボットを制御する。これを毎フレーム繰り返すことで追跡を可能としている。追跡の概念図を図6に示す。



Fig.5 Examples of points used for distance calculation: (a) left camera image (b) right camera image

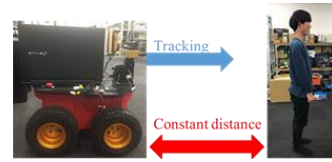


Fig.6 Conceptual diagram of target tracking with a mobile robot

3. 追跡復帰機能

図7に示すように、色ヒストグラム同士の比較には、前フレームの追跡対象者の人物領域画像をテンプレート画像として用いる。テンプレート画像は毎フレーム更新する。このとき、現フレームで決定した人物領域画像を次フレームでのテンプレート画像とするため、フレーム間で急な照明変動が起こると、色ヒストグラムが変化し、追跡対象者を見失う場合がある。この課題を解決するために、提案手法では追跡復帰機能を搭載する。

エピソード拘束により探索範囲を限定したテンプレートマッチングによって追跡復帰を実現する。2台のカメラのうち片方のみが追跡に失敗した場合、三角測量で距離を測ることができず、移動ロボットでの追跡は失敗してしまうという課題がある。この場合にのみ、テンプレートマッチングを用いて、追跡に成功しているカメラの人物領域画像からもう一方の追跡に失敗しているカメラの人物領域画像を探索することで追跡の復帰を試みる。このとき、誤対応を減らし探索時間を短縮させるために、テンプレートマッチングを行う際にエピソード拘束を用いる。図8にこの手法の流れを示す。

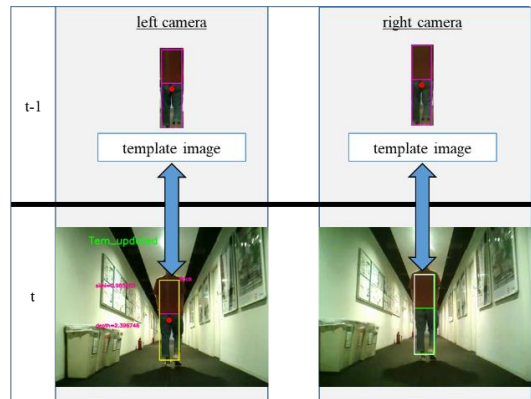


Fig.7 Similarity calculation between adjacent frames

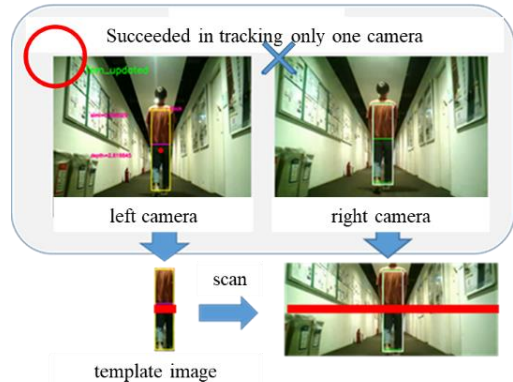


Fig.8 Tracking return function using template matching

4. 曲がり角追跡機能

曲がり角のようなシーンにおいて追跡対象者がカメラからフレームアウトした場合、追跡を継続することが困難となる。この課題を解決するために、骨格情報から人の体の向きを推定し、その推定結果から移動ロボットを制御することで曲がり角のある環境での追跡継続を可能にする。

人が横を向いた時、カメラから見た人の特徴として肩幅が小さくなり鼻が観測できるという点が挙げられる。そこで、OpenPose で取得した骨格情報から人が背面を向いているか、横を向いているかを推定する。また、横向きの場合にはつま先の座標を利用して左右どちらを向いているかを認識する。これらの推定には、図 9 に示すように、抽出した骨格点の中で左右の肩と足首の骨格座標を用いて取得した肩幅と身長のと鼻の座標の有無を利用する。鼻の座標が存在し、その比が一定の値より小さければ横向きと判定する。

三角測量によって求めた移動ロボットから追跡対象者までの距離情報と骨格抽出によって推定した追跡対象者の体の向き情報から移動ロボットを制御する。移動ロボットは追跡対象者が曲がった地点まで直進し、向いた方向に 90° 回転するように自律走行する。

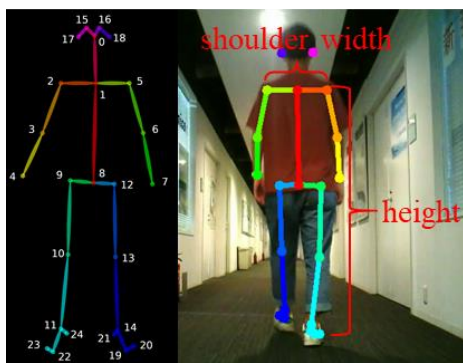


Fig.9 Skeleton points used to estimate body orientation

5. 人物追跡システムの性能検証実験

提案手法の有効性を検証するために二種類の実験を行った。両実験とも屋内環境で行い、被追跡者は一定の速度で歩行動作を行った。実験で使用した装置を図 10 に示す。カメラには ELECOM 製 UCAM-DLI500TN、移動ロボットには AdeptMobileRobots 社製の Pioneer3-AT を使用した。カメラの解像度は 640×480 pixel である。移動ロボット上に 2 つのカメラを平行に並べることで平行ステレオを実現した。基線長は 95mm とした。



Fig.10 Experimental equipment

5.1 追跡復帰機能の検証

追跡復帰機能の有効性を検証するため、二つの実験を行った。一つ目の実験は、オクルージョンが 5 回発生する条件のもと、事前に撮影した動画を用いて、追跡復帰機能を搭載する手法と追跡復帰機能を搭載しない手法それぞれをオフライン環境で行った。二つ目の実験は、オンライン環境かつより長い距離で、移動ロボットによる 3 次元人物追跡を行った。一つ目の追跡実験の結果を表 1 に、二つ目の追跡実験の結果を表 2 に示す。

表 1 からわかるように、追跡復帰機能未搭載の手法では 11 フレームで移動ロボットによる追跡に失敗していた。これは全て 2 台のカメラのうちの片方が追跡に失敗してしまい距離計測が行えないことが原因であった。図 11 に追跡失敗の例を示す。左カメラでは問題なく追跡に成功しているが、右カメラでは照明の変化によって画像中の服の色見え方が変わってしまい、追跡に失敗していることがわかる。ここでは追跡対象者の人物領域をピンク色、追跡対象者以外の人物領域を黄緑色で表している。つまり、右カメラでは追跡対象者以外の人物と判定してしまっている。図 12 には、提案手法を用いて追跡の復帰に成功したフレームの例を示している。ここでは右カメラで一度追跡に失敗したが、左カメラで追跡に成功した人物領域画像をテンプレート画像としたことで、追跡復帰が成功した。これにより従来手法で追跡に失敗していた 11 フレーム全てで追跡復帰に成功した。また、移動ロボットの動作を含めた評価を図 13 に示す。追跡対象者と移動ロボット間の距離は 2.0~3.5m の範囲内で保つことができ、リアルタイムでの追従が可能である。赤く矩形で囲った部分は非追跡対象者が横切ったフレームを示しており、未検出の際はロボットが停止してしまうため、距離が離れてしまっていることがわかる。

表 2 からわかるように、追跡距離を長めにとった 2 つ目の実験では、左右のカメラ合計 32 フレームで一度は追跡に失敗していた。追跡失敗のうち、29 フレームはいずれか一方のカメラで追跡が成功していたため、追跡復帰処理を行うことでそれら全てのフレームで追跡復帰に成功した。追跡復帰に成功したフレームの様子を図 14 に示す。これにより、3 フレームを除いた 378 フレームで移動ロボットによる追跡が成功し、追跡成功率 99.2% という結果を得た。なお、ここでの追跡成功率とは、左右のカメラのうち追跡成功数の少ない方を全フレーム数で割ったものを表す。



Fig.11 A frame that failed to track a person using the method without a return function



Fig.12 A frame that succeeded in returning to track a person with the proposed method

Table 1 Success rate of target tracking in experiment 1

	Without return function	With return function
Total number of frames	164	164
Number of tracking failure frames	11	11
Number of tracking return frames	-	11
Number of successful tracking frames	153	164
Tracking success rate %	93.3	100

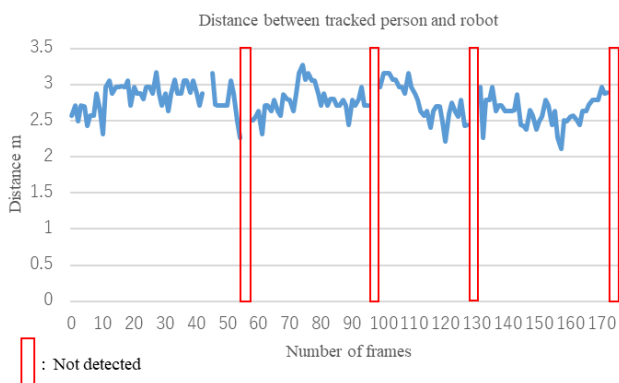


Fig.13 Distance between tracked person and mobile robot



Fig.14 A frame that succeeded in returning to track a person with the proposed method

Table 2 Success rate of target tracking in experiment 2

	Proposed method
Total number of frames	381
Number of tracking failure frames	32
Number of tracking return frames	29
Number of successful tracking frames	378
Tracking success rate %	99.2

5.2 曲がり角追跡機能の検証

曲がり角追跡機能の有効性を検証するため、従来の手法では追跡の継続が不可能であった屋内環境の曲がり角でオンライン実験を行った。追跡対象者は一定の速度で前進し曲がり角を曲がると仮定する。

Fig.15, 16に追跡対象者が曲がり角を曲がる前後における2台のカメラ画像を示す。Fig.15に示すように、OpenPoseで

取得した骨格情報から追跡対象者が右を向いたことを正しく認識することが確認できた。また、Fig.16に示すように、移動ロボットが曲がり角を曲がった後に再度追跡対象者を追従することが確認できた。Fig.17に示すように、追跡対象者が曲がった地点まで直進し、90°向きを変え、追跡対象者を再度追従するという移動ロボットの自律移動を実現した。

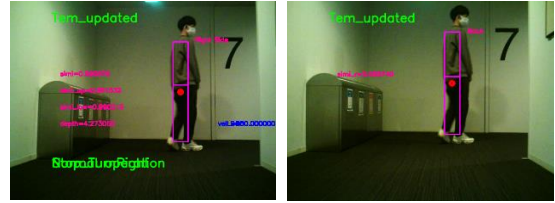


Fig.15 Camera images before turning the corner

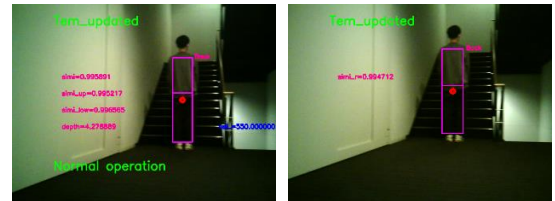


Fig.16 Camera images after turning the corner

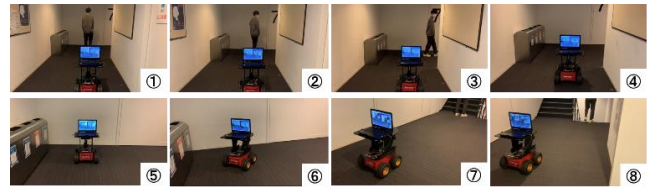


Fig.17 Autonomous driving of mobile robots at a corner

6. 結論

本論文では、2台のカメラを用いた骨格抽出による3次元人物追跡システムを提案し、追跡復帰機能と曲がり角追跡機能を搭載することで頑強な人物追跡を実現した。今後は、服の色での識別は人の姿勢の影響を受けやすいという課題に対し、曲がり角を曲がる前に骨格情報を用いた人物識別手法を取り入れることで追跡精度の向上を目指す。

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.
- [2] Y. Isobe, G. Masuyama, K. Umeda, "Occlusion handling for a target-tracking robot with a stereo camera", Robomech Journal, pp.1-13, 5.1, 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", ロボティクス・メカトロニクス講演会, 1A1-M09, 2018.
- [4] Z. Cao, T. Simon, S. Wei, Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," Proc. of 2017 IEEE conference on computer vision and pattern recognition, pp.7291-7299, 2017.