

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

Three-dimensional target tracking with a mobile robot by fusion of tracking of two cameras

○ 本多 明彦 (中央大) 正 梅田 和昇 (中央大)

Akihiko Honda, Chuo University, honda@sensor.mech.chuo-u.ac.jp
Kazunori UMEDA, Chuo University

In this paper, target tracking is one of the important functions for autonomous mobile robots. To realize this function, stereo cameras are often used. In a general stereo camera, a range image of the entire scene is acquired as precisely as possible, and the image is processed. On the other hand, in this paper, processing is performed by individual cameras according to the requirements, and the results are integrated to acquire three-dimensional (3D) information of only necessary objects. 3D tracking of an object is a specific task. Object detection and tracking are performed by individual cameras, and the results are integrated to acquire 3D tracking information. As a result, erroneous matching can be greatly suppressed, and optimum 3D information for necessary work can be obtained. The effectiveness of the proposed method is verified through target-tracking experiments.

Key Words: Target tracking, Stereo camera, Object detection

1 序論

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。

これに対して、本研究で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する。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・トラッキングを行い、その結果を統合することで3次元のトラッキング情報を取得する手法を構築する。

タスクに応じた処理を各画像に行い識別性の高い高次特徴を抽出してマッチングを行うこと、画像全体で距離を求めることを要求しないこと、さらに各画像に対する処理結果と距離計測との相互処理（フィードバック）を行うことで、必要なタスクに適した3次元情報を得ることが可能であることを実験により検証する。

ステレオカメラは特定の人物の追跡の実現を目的とした研究にしばしば用いられる[1]。この機能は多岐に渡る自律移動ロボットに求められる機能の一つであり、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。汎用的な機能であるため、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。

2 提案手法の概要

2.1 人物追跡の流れ

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する。

本手法のフローを Fig. 1 に示す。本手法では、カラー画像が取得可能なカメラを用いて人物の検出を行う。必要とするカラー画像には高いスペックを要求しないため、現在普及しているカメラであればどのようなものでも取得することがで

きる。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、カメラ画像の中で、人物だと思われる候補領域をフレーム間差分を用いて検出し、テンプレートと比較して一定の値を超えている場合は追跡の対象と指定する。その後指定した対象を Cam-shift を用いて追跡を行う。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。

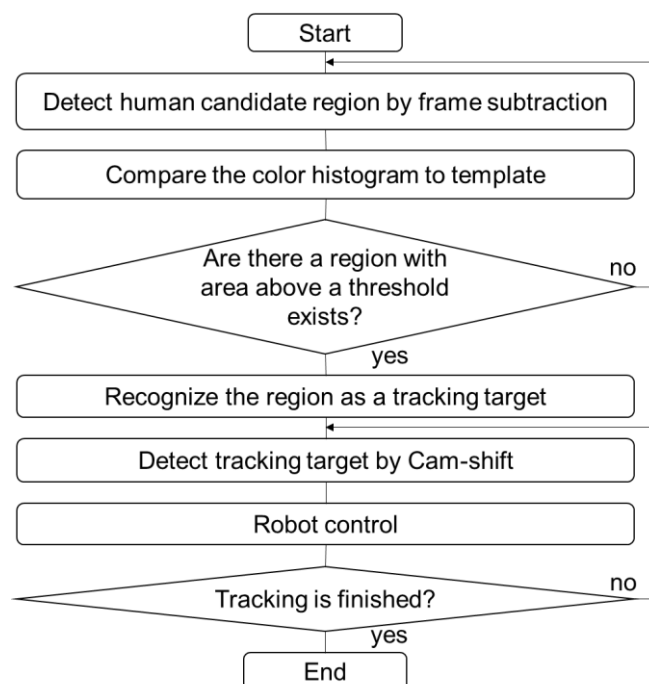
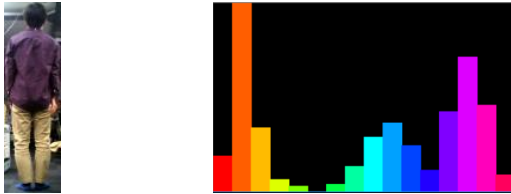


Fig.1 Flow chart

2.2 人物テンプレートの作成

まず色情報のテンプレートを作成する。テンプレートには RGB 画像を HSV 変換して用いる。HSV はそれぞれ色相、彩度、明度を表しており、RGB 値から算出される。次に、変換したテンプレートの HSV の値をヒストグラム化する。本研究では、明るさ、照明変動の影響を受けにくいとされる色相を主に用いて検出をする。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。



(a) Template image (b) Color Histogram
Fig. 2 Template of target

2.3 初期人物領域の検出

次に、カメラ画像の中で、人物だと思われる候補領域をフレーム間の差分画像を用いて検出する。最初に差分画像の人物のテンプレートを用意し、カメラ画像から得た差分画像とのテンプレートマッチングを行う。この結果検出された領域を初期人物候補領域とする (Fig. 3). (a)は差分画像の人物のテンプレート、(b)は差分画像にて検出された領域、(c)は初期人物候補領域の例である。この領域をテンプレートと同様に HSV 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を以下の式で算出する。

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

H_1 はテンプレートのヒストグラム、 H_2 は初期人物候補領域のヒストグラムである。類似度が一定の閾値を超えている場合、追跡対象の初期人物領域とする。



(a) (b) (c)
Fig. 3 Detection of human candidate region

(a) Person template of subtraction image, (b) The area detected in the subtraction image, (c) Example of human candidate region

2.4 追跡領域の検出

2.3 節で決定した初期人物領域を Mean-shift[2,3]と呼ばれるロバストな統計的データ解析手法を拡張した Cam-shift[4]を用いて追跡する。追跡の流れは以下になっている。

- (1) 探索窓を作成し、初期位置を設定する。本研究では初期人物領域をもとに作成する。
- (2) 探索窓内の重みの分布を計算し、探索窓内の重心を計算す

る。

- (3) (2)において計算された重心を用いて、探索窓を設定する
- (4) (2),(3)を収束するまで、もしくは決められた一定回数繰り返す、求められた重心をもとに、探索窓の姿勢及び大きさを変化させ、次のフレームの探索窓を作成する。

- (5) (2)から(4)を毎フレーム繰り返す。

2.5 追跡対象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。その後片方のカメラで検出した追跡領域に外接する矩形をテンプレートとし、もう片方のカメラで検出した追跡領域より少し大きい範囲の画像との間でテンプレートマッチングを行う (Fig. 4). 検出された範囲に外接する矩形 (Fig. 4(a)における黄色矩形) をテンプレートとしてテンプレートマッチングを行う。

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テンプレートマッチングを行った結果検出された範囲 (Fig. 4(b)における黄色矩形) 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。

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

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。



(a) Left camera (b) Right camera
Fig. 4 Template matching between two cameras

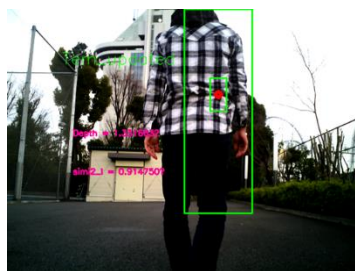
3 実験

提案手法の有効性を検証するため、実験を行った。実験は屋外と屋内の二箇所にて行い、どちらも比較的明るくかつ照明が大きく変化しない環境において 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 Web カメラ 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]とした。移動ロボットには AdeptMobileRobots 社製の Pioneer3-AT を用いた。

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 1.5m 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. 5 に屋外環境で行った実験の様子を、Fig. 6 に屋内環境で行った実験の様子を載せる。



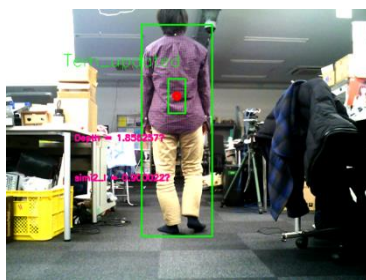
(a) Human region recognized as a tracking target



(b) The robot is moving as seen from the two cameras



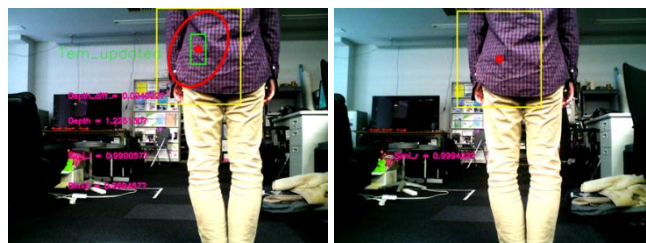
(c) Immediately after the robot stops as seen from the two cameras
Fig.5 Experimental results in outdoor environment



(a) Human region recognized as a tracking target



(b) The robot is moving as seen from the two cameras



(c) Immediately after the robot stops as seen from the two cameras
Fig.6 Experimental results in indoor environment

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それを追跡することが出来た。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出来た。

今回の実験はどちらの実験でも直射日光や建物の影といった照明の変動が少ない環境で行ったため、比較的検出しやすい環境であったと考えられる。さらに今回は初期人物を特定する際に、周囲に他の人物が映り込んでいないことから、初期人物を特定しやすい環境であったと考えられる。

今回は追跡対象以外の部分なるべく映らないように対象から 1.5m 離れると追跡を開始するように設定したが、一度人物が停止して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と追跡対象に近づきすぎてしまい、対象全体が映らなくなっているため、1.5m 以上の距離に設定して実験を行い、追跡対象以外の部分が多く映っても問題がないかを検証する必要があると考えられる。

上記のように、検出、追跡が安定した理由はいくつか考えられるものの、実際の屋外環境や屋内環境で大きな問題なく追跡が出来たことはこの手法が有効であると考えられる点である。

4 結論と今後の展望

本論文では、カメラを二台使用し、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し、必要な対象のみに関しての 3 次元情報を取得した。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・トラッキングを行い、その結果を統合することで 3 次元のトラッキング情報を取得する手法を実装した。実験の結果、特定環境下において人物が追跡できることを示した。今後は照明環境が変動する環境などの多様な環境での追跡、複数人が映っている状況をはじめとする多様な状況での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, 2016.
- [2] Yizong, C., “Mean Shift, Mode Seeking, and Clustering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.17, No.8, pp.790-799, 1995.
- [3] 都築勇司 他, “SIFT 特徴量に基づく Mean-Shift 探索による特徴点追跡”, 情報処理学会論文誌: コンピュータビジョンとイメージセンサ, Vol.49, No.SIG6 (CVIM20), pp.35-45, 2008.
- [4] Gary, R. et al., “Computer Vision Face Tracking For Use in a Perceptual User Interface”, Intel Technology Journal, Q2, pp.1-15, 1998.