

差分ステレオとパーティクルフィルタによる人物追跡

Tracking of People Using Particle Filter and Subtraction Stereo

○学 川下 雄大 (中央大) 学 柴田 雅聡 (中央大)
 生形 徹 (中央大) 有江 誠 (中央大)
 正 寺林 賢司 (静岡大) 正 梅田 和昇 (中央大)

Takehiro KAWASHITA, Chuo University, kawashita@sensor.mech.chuo-u.ac.jp
 Masatoshi SHIBATA, Chuo University, shibata@sensor.mech.chuo-u.ac.jp
 Toru UBUKATA, Chuo University, ubukata@sensor.mech.chuo-u.ac.jp
 Makoto ARIE, Chuo University, arie@sensor.mech.chuo-u.ac.jp
 Kenji TERABAYASHI, Shizuoka University, tera@eng.shizuoka.ac.jp
 Kazunori UMEDA, Chuo University, umeda@mech.chuo-u.ac.jp

In this paper, a tracking system based on Particle Filter is presented. Color and distance information are used. Distance information are obtained by "Subtraction Stereo", which restricts stereo matching to foreground regions extracted by subtraction. The effectiveness of the proposed system is verified by experiments.

Key Words: Stereo Camera, Subtraction Stereo, Tracking, Particle Filter

1. 緒言

近年, 防犯意識の高まりやカメラの低価格化に伴い, 街中に多くの監視カメラが設置されている. 監視カメラ設置の目的の多くは異常行動の検出や不審者の追跡, 人流計測による安全管理である. しかし, 現在使用されている監視カメラの用途の多くが撮影したデータを保存し, 事案の発生後にデータを検証するタイプのものである. リアルタイムで不審者の検出・追跡等を行うためには, 常にカメラの映像を監視する人物が必要となる. さらに, カメラ設置台数に伴い, 多くの人材・人件費が必要となる. また, 監視中のヒューマンエラーも考えられるため, 人物によるリアルタイムの監視は, 難しいと考えられる. そのため, PC の処理能力の向上等の技術の進歩に伴い, カメラ映像からリアルタイムで人物を検出・追跡を行うといった研究が多数行われている[1,2,3].

監視カメラの用途を想定すると, 映像内の各人物の行動の時系列データの正確な取得が重要となる. 追跡にあたり, 時系列フィルタリングが有効であることが知られている. 時系列フィルタとは, 過去の観測データから現状態や未来の観測値を推定する手法である. その中でも, カルマンフィルタ[4]やパーティクルフィルタ[5, 6]がよく知られている. 特に, 最近ではパーティクルフィルタを用いた追跡手法の有効性が多数報告されている. 片岡らは, 複数台のカメラを用いることで, オクルージョンを減らし, 高精度な人物追跡を可能にしている[2]. しかし, 同じ箇所を複数台のカメラで計測する環境は, コスト等を考慮すると現実的ではない. そこで本手法では, 3次元情報を取得できるステレオカメラを用いて計測を行う. 先行研究において我々は, 運動領域に注目することでステレオカメラのロバスト化を実現する差分ステレオ(Subtraction Stereo)を提案している[7].

本論文では, 差分ステレオにより得られた 3 次元距離情報と人物の色情報を利用したパーティクルフィルタを実装する. また, 実環境下において本手法の人物追跡に対する有用性を実験により実証する.

本論文の構成は以下の通りである. まず, 第 2 章で, 本研究での人物検出の基盤となっている差分ステレオと影検出について説明する. 第 3 章では, 3次元距離情報と色情報を利用したパーティクルフィルタによる人物追跡手法について説明する. 第 4 章では, 本手法を用いて行った人物追跡実験につ

いて説明する. また実験より, 本手法の有効性を示す. 最後に第 5 章では, 結論と今後の展望について述べる.

2. 人物検出

2.1 差分ステレオ

本論文では, カメラ画像中から人物を検出する手法として, 差分ステレオという手法を用いる. 差分ステレオの基本アルゴリズムを図 1 に示す. また, 差分ステレオにより得られた視差画像を図 2 に示す. 通常のステレオカメラでは, 左右のカメラの画像をマッチングすることで視差画像を得る. 対して, 差分ステレオでは, まず左右のカメラの画像をそれぞれ背景差分を行い, 前景領域を抽出し, 抽出された画像内でマッチングを行うことで距離情報を得る. 背景差分を行うことで, 用途が定点カメラに限られるが, マッチングする領域を背景差分で得られた前景領域に限定することで, マッチングの誤対応と処理時間を削減することができる.

2.2 影検出

差分ステレオによって検出される前景領域には, 人物の影となる領域も含まれてしまう. 人物を正確に検出するためには, 影となる領域を前景領域から除去する必要がある. 以下に手法を示す[8].

画像座標 (x, y) における輝度値を $I(x, y)$ とし, 背景画像における同位置の輝度値を $I'(x, y)$ とする. 影を判定する評価関数として, 次式を用いる.

$$\theta_{(t+1, x, y)} = \begin{cases} \alpha \Psi_{(x, y)} + \beta \Lambda_{(x, y)} + (1 - \alpha - \beta) \theta_{(t, x, y)}, & \text{if } \frac{I_{(x, y)}}{\eta} < I'_{(x, y)} \\ \infty, & \text{otherwise} \end{cases} \quad (1)$$

ここで, θ は影と判断するためのスコアを表し, θ が閾値以下となる画素を影と判断して前景領域から除去する. Ψ は前景の近傍画値と背景の近傍画素値の相違度, Λ は前景の色相と背景の色相の相違度を表す. また, α, β, η はそれぞれの項に対して重みを与える定数である.

影検出を行った結果を図 3 (c) に示す. 図中において, 影と認識された領域を緑色で表わし, 前景領域から除去している.

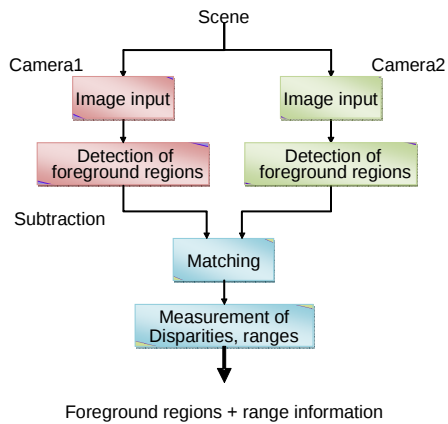


Fig.1 Flow of the subtraction stereo

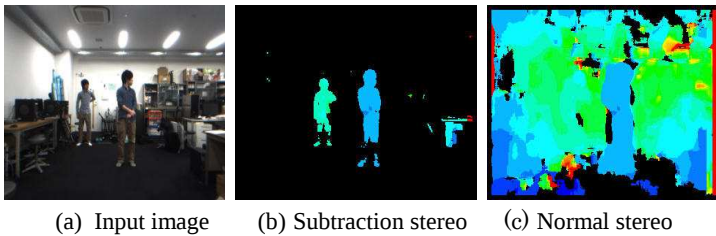


Fig.2 Disparity image

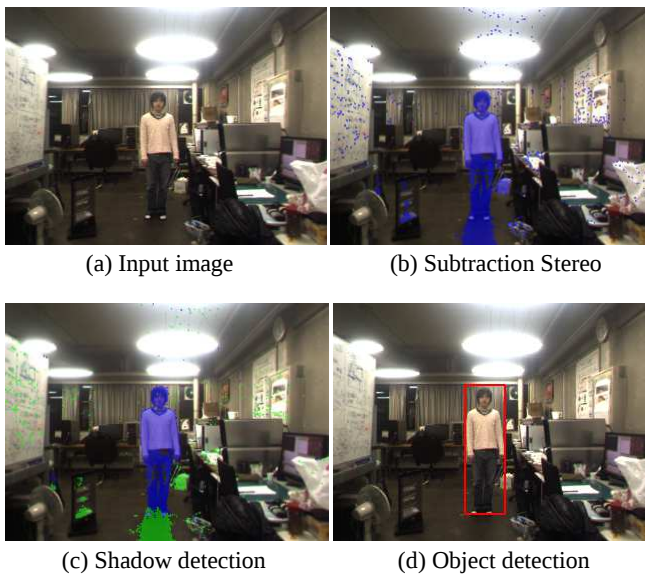


Fig.3 Flow of the object detection

3. 人物追跡

本節では、距離情報と色情報を組み合わせたパーティクルフィルタのアルゴリズムについて説明する。パーティクルフィルタとは、前状態から起こりうる多数の現状態を、多数の粒子に見立て全パーティクルの尤度に基づいた重みつき平均を現状態として求める。さらに、対象の位置を逐次予測しながら追跡を行なっていく手法である[5]。主な処理は、予測、観測、対応付け、選択の4つであり、これらの処理を繰り返すことで、人物の追跡を行う。本研究では、差分ステレオに

よって検出された人物に対して、パーティクルフィルタを適用し追跡を行う。処理の流れを図4に示す。図4(a)の点は、散布されたパーティクルを表している。図4(c)のパーティクルを中心とした円は各パーティクルの尤度を視覚的に表したもので、半径が大きいほど尤度が大きいということを表している。また、図の赤点は、人物の重心位置を表している。

(1) 初期化

人物検出時にパーティクルを人物周辺にランダムに散布する。

(2) 予測

人物の前フレームの移動量から等速直線運動を仮定して、現在のフレームの人物の位置を予測し、それに乱数を加え、パーティクルを遷移させる。

(3) 観測

人物の重心位置と各パーティクルのユークリッド距離より各パーティクルの尤度を計算する。各パーティクルの尤度 L は次式より定義する

$$L = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (2)$$

d は、人物の重心位置と各パーティクルのユークリッド距離、 σ は、標準偏差を表している。

(4) 対応付け

各フレームでパーティクルが追跡中の人物の判別を行い、対応付けを行う。人物の判別は、パーティクル群と人物の重心位置間のユークリッド距離と色の類似度から行う。色情報には、HSV モデルのうち、H(色相)を用い、ヒストグラムを用いて表現する。色情報の取得は、人物の重心位置から上半身・下半身を矩形で指定し、矩形内の前景領域のみを走査して色相の値をカウントし、色ヒストグラムを作成することで行う。この処理を人物が検出されてから5フレーム分行い、得られた5フレーム分の上半身・下半身の色ヒストグラムの平均をモデルヒストグラムとする。また、モデルヒストグラム完成時に、モデルヒストグラムを作成した際に追跡を行っていたパーティクルに対応させる。そして、モデルヒストグラムと各フレームで得た人物の上半身・下半身のヒストグラムを比較し、類似度を求める。類似度は、まず各モデルヒストグラムと各フレームで計算される人物の上半身・下半身の色ヒストグラム間の Bhattacharyya 係数を求める。Bhattacharyya 係数 ρ は次式により定義される。

$$\rho = \sum_{u=1}^m \sqrt{p_u q_u} \quad (3)$$

ここで、 p , q は比較対象となる正規化された色相分布、 u は色相番号、 m は色相の成分数を表す。Bhattacharyya 係数 ρ より色距離 D_c を次式で求める。

$$D_c = \sqrt{1 - \rho} \quad (4)$$

式(3)から上半身・下半身の色距離 D_{c-upp} , D_{c-low} を求め、3次元空間のユークリッド距離 D_E を足しあわせ、最終的な距離 D_s とする。

$$D_s = D_E + D_{c-upp} + D_{c-low} \quad (5)$$

式(4)の D_s が最小の値を持つ人物を各フレームにおいてパーティクルが追跡中の人物とする。また、対応付け終了後に対応づけられた人物とパーティクルの尤度の再計算を行う。

(5) 選択

各パーティクルの尤度に従い選択を行う。尤度の大きいパーティクルが存在する位置にはパーティクルを尤度に従い再配置し、尤度が小さいパーティクルは、その位置に人物重心位置がある確率が低いとみなし消滅させる。この処理により、人物の重心位置周辺にパーティクルが残る。これにより、追跡が続くにつれ、パーティクルは人物の重心位置に収束し、安定した人物の追跡が可能となる。

以上の処理を繰り返すことで、人物の追跡を行う。また、人物がオクルージョン等によって検出されない場合、検出された最終フレームの人物の速度・ベクトルを用いて予測を行う。これにより再検出された場合でも、追跡中の人物の大半の位置が把握できるため、再追跡が可能となる。

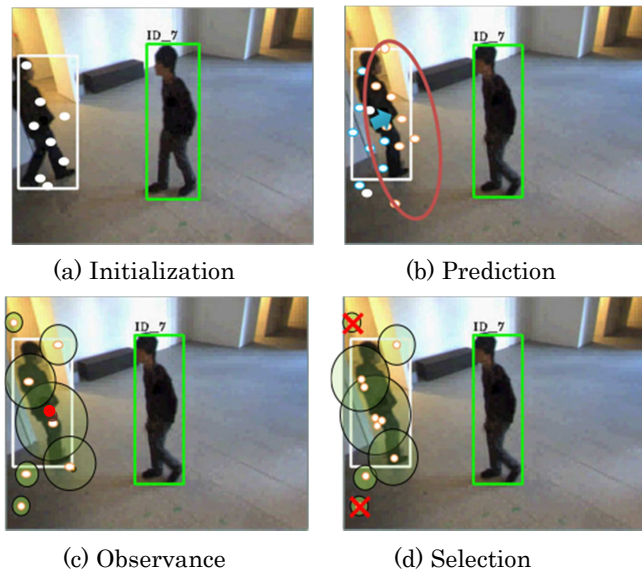


Fig.4 Flow of the Particle Filter

4. 実験結果

4.1 屋内での人物追跡実験

本章では、前章で説明した手法の有効性を検証するため、図 5 に示す異なる 2 シーンにおいて人物追跡実験を行った。2 シーンは、それぞれ屋内・屋外に監視カメラを設置したことを想定した。実験装置は、Bumblebee2(Point Grey Research, カラー, $f=3.8\text{mm}$, $b=120\text{mm}$, VGA, 30fps)で保存した動画に対して処理を行った。処理を行う際に使用した PC は、Intel Core2 Extreme (2.93 GHz), メモリが 3 GB RAM を搭載したものを使用した。カメラの設置環境は、地面からの高さ: 2.3m, 仰角: 30° とした。屋内での実験環境を図 5(a)に示す。各図の出力画像にある矩形は、人物が検出されていることを表しており、矩形上部には人物毎の識別番号である ID が記されている。矩形毎の色は、視認性向上のため、ID 毎に色分けを行ったものである。また、俯瞰図にある色のついた点群は散布中のパーティクルである。色は、ID 毎に色分けを行った結果である。また、図中にある黒点は差分ステレオから得られた各人物の観測値を表している。

今回実装した距離情報と色情報を用いた人物追跡手法の有用性を検証するため、色情報を加えた場合と距離情報のみを用いた場合での比較実験を行った。各追跡対象に 700 個パーティクルを散布し、追跡を行う。また、人物の追跡成功の判定基準は、人物が検出されて識別番号である ID が与えられてから、画面外に出るまで ID が変化しないことである。逆に、追跡中にオクルージョン等が発生し、ID の入れ替えや変化が起きると、その人物に対しての追跡は失敗と見なす。実験結果を表 1, 実験中の様子を図 6 に示す。

実験結果より、色情報を組み合わせることで、人物追跡精度が向上することがわかった。これは、服の色が異なる人物同士であると、すれ違い等が生じ人物同士の距離が近い場合などでも色距離を用いることで、判別ができるためである。追跡精度も 92.5%と高いことから、本手法が複数人物の追跡に有用であると考えられる。

4.2 屋外での人物追跡実験

本手法の実用性を検証するため屋外においての実験も行った。本節での実験も保存した動画に対して処理を行った。カメラの設置環境は、地面からの高さ: 8.3m, 仰角: 50° とした。屋外での実験環境を図 5(b)に示す。

パーティクルの散布数は、屋内での実験と同じ 700 個である。表 1 に示す結果より、色情報を組み合わせることで追跡精度が向上することがわかった。これは、4.1 節で説明した理由と同様であり、色情報を用いることで、人物の判別がより正確に行えたことが原因だと考えられる。また、追跡精度もほぼ 90%を達成出来た。しかし、屋内での実験に比べ追跡精度は下がってしまった。これは、人物検出が屋内に比べうまくいかず、誤検出・未検出が多かったことが原因として考えられる。図 7(b)では、二人の人物が一人の人物として検出されてしまっている。これは、観測対象の距離が屋内に比べ遠く、距離の情報が正確に取得できず、距離情報によるセグメンテーションがうまくいったことが原因である。本システムでは、追跡中の人物が未検出の際には、検出できた最終フレームの速度・ベクトルを用いて、再び人物が検出されるまで位置予測を行う。未検出が短いフレームの間だと、予測と観測値のずれが小さく正確に追跡が行えるが、未検出が長い期間で起こると、予測値と観測値のずれが大きくなってしまい、再検出された際に対応付けがうまく行かず、追跡の失敗が起きてしまうと考えられる。

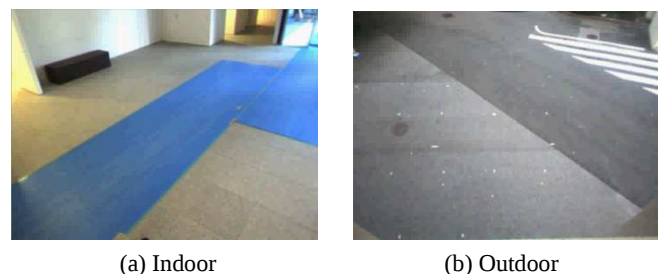


Fig.5 Experimental Scenes

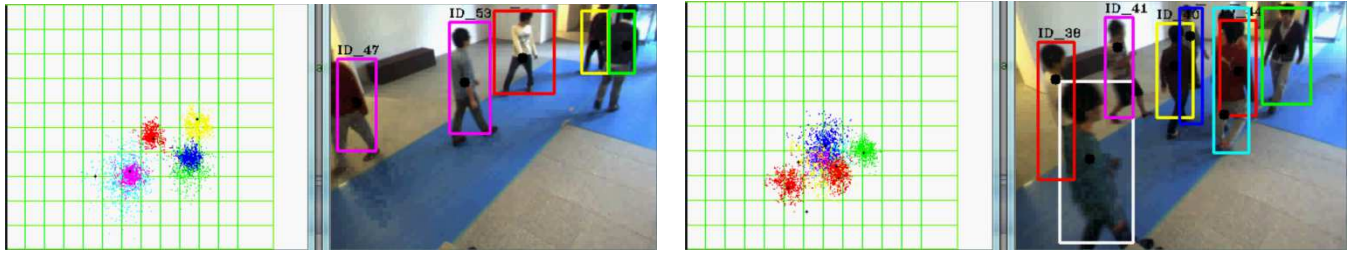


Fig.6 Experimental result (Indoor)
(Left: overhead view, Right: output Image)

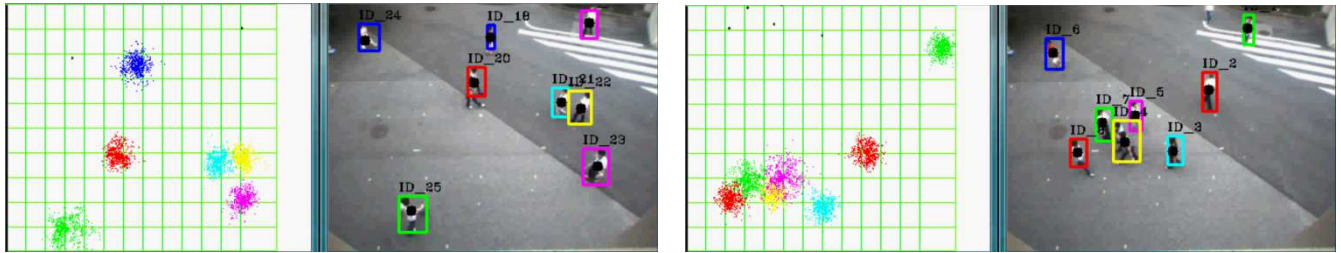


Fig.7 Experimental result (Outdoor)
(Left: overhead view, Right: output Image)

Table 1 Experimental results

実験環境	検出人数[人]	特徴量	追跡精度[%]
屋内	40/43	距離+色	92.5
		距離のみ	87.5
屋外	33/38	距離+色	88.0
		距離のみ	81.5

5. 結論と今後の展望

本論文では、差分ステレオにより得られた 3 次元距離情報と人物の色情報を利用したパーティクルフィルタを用いた人物追跡手法を実装した。また、本手法の人物追跡に対する有用性を検証するため、屋内と屋外の異なる 2 シーンで実験を行った。結果より、両シーン共に、ほぼ 90%と精度の高い結果が得られ、本手法が人物追跡に有用であることが示された。

今後は、実験の結果から人物の検出がうまくいかないことで追跡が失敗してしまう問題に対して、対策を考案する。人物検出精度向上のために、まず背景差分の精度を向上させることと近年人物検出手法として多数有効性が報告されている Joint HOG[9]を組み合わせることが挙げられる。人物検出精度が向上することで、人物の時系列データ(3次元位置等)を詳細に取得できるため、追跡精度も向上することが考えられる。

あるいは、追跡中の人物が検出されていない時の人物の位置予測の精度を上げることで追跡精度が上がると思われる。他に、式(5)で最終的な距離を求める際に、状況により色と距離の重み付けを行うことでより正確な追跡が可能になると考えられる。

文 献

- [1] 大西正輝, 依田育士, “ 大型複合施設における長期間にわたる人流比較の可視化手法 ”, 電子情報通信学会論文誌, 93-D(4), pp.486-493, 2010.
- [2] 片岡裕雄 他, “ 歩行者予防安全のための CoHOG 特徴量改善による検出と Particle Filter による追跡 ”, 電気学会論文誌, Vol.132, No.7, 2012.
- [3] 西濃拓郎 他, “ 複数尤度を用いた 3 次元パーティクルフィルタによる選手の追跡 ”, 画像の認識・理解シンポジウム, MIRU2010, IS1-39, pp.307-312, 2010.
- [4] S.Thum, W.Burgard, and D.Fox, “ Probabilistic Robotics ”, MIT press, 2006.
- [5] 加藤丈和, “ パーティクルフィルタとその実装法 ”, 情報処理学会研究報告, CVIM, Vol. 2007(1), 2007, pp.161-168.
- [6] 倉爪他, “ 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 28(1), pp.65-76, 2010.
- [7] Umeda K., et al., “ Subtraction Stereo -A Stereo Camera System That Focuses On Moving Regions -” ,Proc. of SPIE-IS&T Electronic Imaging, Vol.7239 Three-Dimensional Imaging Metrology, 723908, 2009.
- [8] A. Moro, et al., “ Auto-adaptive threshold and shadow detection approaches for pedestrians detection ”, In Proc AWSVCI, pp. 9-12, 2009.
- [9] 三井相和, 山内悠嗣, 藤吉 弘亘, “ Joint 特徴量を用いた 2 段階 Boosting による物体検出 ”, 電子情報通信学会論文誌, Vol.92-D, No.9, pp. 1591-1601, 2009.