

差分ステレオとパーティクルフィルタによる複数人物追跡*

川下 雄大^{*1}, 柴田 雅聡^{*2}, 増山 岳人^{*3}, 梅田 和昇^{*4}

Multiple human tracking using subtraction stereo and particle filter

Takehiro KAWASHITA^{*1}, Masatoshi SHIBATA, Gakuto MASUYAMA and Kazunori UMEDA

^{*1} Course of Precision Engineering, School of Science and Engineering, Chuo University,
1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

Received 7 November 2013

Abstract

This paper proposes a method for automatic tracking of multiple humans in various scenes using a stereo camera. Surveillance cameras are increasing these days for security. Typically, images of a camera are stored and are surveyed after an incident occurred. It is required to use the cameras online, and real-time automatic human tracking is required. The proposed method detects candidate regions of humans using “Subtraction Stereo”, which restricts stereo matching to foreground regions extracted by subtraction and obtains distance information for the regions. Human tracking is carried out for the extracted regions by using Particle Filter. Particle Filter consists of four steps, i.e., prediction, calculation of likelihood, data association, and resampling. Three features: distance, color, and direction of motion are used in the Particle Filter of the proposed method to achieve robust human tracking. When humans with similar cloth colors pass each other, occlusion occurs and human tracking often fails with distance and color only. The proposed method achieves the robustness to the occlusion by explicitly considering the direction of human motion. The proposed method is evaluated through experiments using a stereo camera that simulates a surveillance camera in real scenes. Tracking accuracy of more than 90% is achieved in three different scenes, which shows the effectiveness of the proposed method for human tracking.

Key Words : Stereo camera, Particle filter, Human tracking, Subtraction stereo, Surveillance, Image processing

1. は じ め に

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は不審者の検出・追跡、異常行動の検出や人流計測などである。現在使用されている監視カメラの多くは撮影したデータを保存し、事案発生後にデータを検証する運用がなされているため、事案発生時に迅速な対応を行うことは難しい。一方、リアルタイムの監視では、事案発生時に迅速な対応が可能となる。しかし、リアルタイムで不審者の検出や追跡を行うためには、常にカメラの映像を監視する人物が必要となるが、カメラ設置台数の増加に伴い、多くの人件費が必要となる。また、監視中のヒューマンエラーは避けられないが、犯罪の監視という問題の性質上、見逃しを可能な限り減らすことが重要となる。そのため、カメラ映像からリアルタイムで人物の検出・追跡を行う研究が多数行われている (Huchuan, et al., 2008)(高橋他, 2012)(大西, 依田, 2010)(Kurazume, et al., 2010)。

* 論文 No.13-00810 [DOI: 10.1299/transjsme.2014dr0052]

^{*1} 中央大学大学院理工学研究科精密工学専攻 (〒112-8551 東京都文京区春日 1-13-27)

^{*2} 学生員, 中央大学大学院理工学研究科精密工学専攻

^{*3} 中央大学理工学部精密機械工学科

^{*4} 正員, フェロー, 中央大学理工学部精密機械工学科

E-mail: kawashita@sensor.mech.chuo-u.ac.jp

監視カメラによりリアルタイムでの人物追跡を行うためには、映像内の各人物がどのように動いたかを表す時系列データの取得が重要となる。人物追跡の際に、不完全な観測値しか得られない状況で追跡をする場合には、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。時系列フィルタリングとは、過去の観測データから現状や未来の観測値を推定する手法である。代表的な手法として、カルマンフィルタ (Kalman, 1960) やパーティクルフィルタ (Neil, et al., 1995) が挙げられる。特に、近年では追跡対象の動きの変動や観測ノイズ、観測値が非ガウス型である場合においてもロバストな追跡が可能であることから、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。パーティクルフィルタによる人物追跡では、まず追跡対象の過去と現在の観測情報から予測される次状態を多数のパーティクルに見立てる。そして、全パーティクルの尤度より算出した重み付き平均を次状態として推定することで行われる。しかし、実環境での人物追跡では、人物同士のすれ違いや遮蔽物などに起因するオクルージョンにより、追跡が途切れてしまうという問題がある。

オクルージョン問題に対してセンサを複数設置し、様々な角度から観測情報を得ることで遮蔽領域を減らすアプローチがある。原口ら (原口他, 2010) 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。中村ら (中村他, 2006) は、レーザレンジファインダを最大 6 台用いて得られた足断面のレンジデータから歩行者の追跡を行うことで、駅構内において通常時 100%, 通勤ラッシュ時 80% の精度で追跡を可能にしている。曾我部ら (曾我部他, 2010) は、複数センサから得た色情報等の観測情報と JPDAF (Liebens, et al., 2006) とを組み合わせることで正確な複数移動物体の追跡を可能としているが、追跡対象の服の色が似ている状況であると、正確な追跡は困難になると考えられる。また、小林ら (小林他, 2007) は、Mixture Particle Filter の枠組みに基づいて、分散カメラとレーザ測域センサを相補的に統合することで高精度な人物追跡、頭部の向きの検出を可能にしている。単眼カメラによる追跡実験も行われているが、レーザ測域センサと組み合わせることが重要で、実環境による実験は行われていない。また、監視カメラの運用を想定すると、上記の手法では同じ箇所を複数台のセンサで計測するため、設置コスト等を考慮すると望ましくない。

カメラ 1 台の場合、人物を真上から俯瞰するように天井にカメラを設置することで、なるべく人物同士の重なりが生じないようにするといった対策がなされている。また、西濃ら (西濃他, 2010) は、サッカー場という距離等既知の環境で計測を行うことで、単眼カメラながらも 3 次元的に追跡を行うことで追跡精度を向上させている。しかし、これらの方法では、カメラの設置環境が限定されてしまうことが問題となる。その他にも池田ら (Ikoma, et al., 2007) は、PHD フィルタ (Ronald, 2002) と全方位カメラを組み合わせた人物追跡手法を提案しているが、3 次元的な追跡は行えていないため正確な人物の行動履歴を把握することは難しい。一方で、複雑な立体環境下でも 3 次元計測が簡易に行えるステレオカメラを用いた追跡も行われている (Kang-Il, et al., 2013) (Kovacevic, et al., 2011)。ステレオカメラの場合、1 台で 3 次元計測が行える上、設置場所も限定されない。また、3 次元的に追跡を行うことで、オクルージョンに対してもロバストな追跡が行える。

そこで本論文では、距離情報をロバストに取得できる差分ステレオ (Subtraction Stereo) (Umeda, et al., 2009) と人物の 3 次元位置と進行方向と色情報という評価値を用いたパーティクルフィルタを組み合わせることで、1 台のステレオカメラを用い設置場所を限定しない実用的な性能を持つシステムを構築する。

提案システムの構成を図 1 に示す。まず、差分ステレオにより距離情報を取得し、人物を検出する。次に、差分ステレオにより検出された人物にパーティクルフィルタを適用し、人物の追跡を行う。提案システムでは、ステレオカメラにより得られるワールド座標系, X, Y 平面上 (図 2) の位置情報と人物の色情報と人物の進行方向を用いて追跡対象とトラックの対応付けを行う。追跡対象判別に位置・進行方向・色の 3 つの評価値を組み合わせることでオクルージョンに対するロバスト性を向上させる。

本論文の構成は以下の通りである。まず、2 章で、差分ステレオを用いた人物検出手法について述べる。3 章では、3 次元距離情報と色情報と人物の進行方向を用いたパーティクルフィルタによる人物追跡手法について述べる。4 章では、提案システムの有用性を実証するため行った、監視カメラの運用を想定した人物追跡実験の結果を示す。最後に 5 章では、結論と今後の展望について述べる。

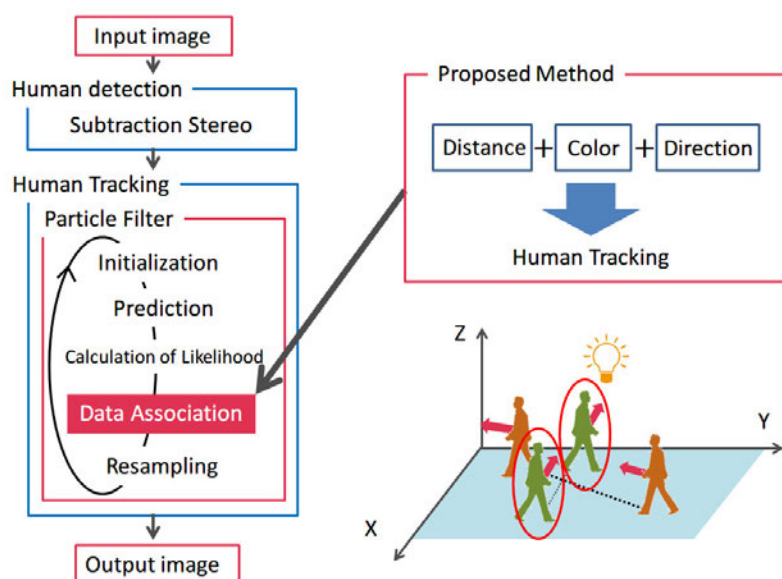


Fig. 1 Flow of proposed human tracking method

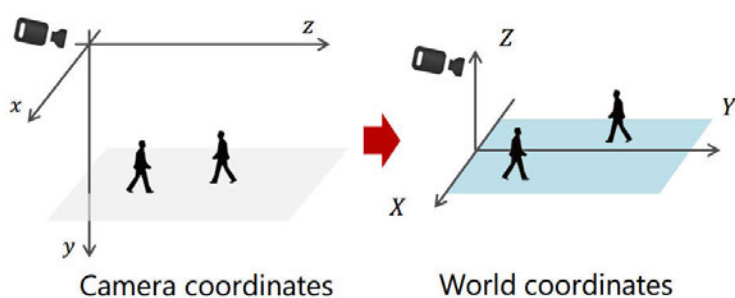


Fig. 2 Coordinate system setting

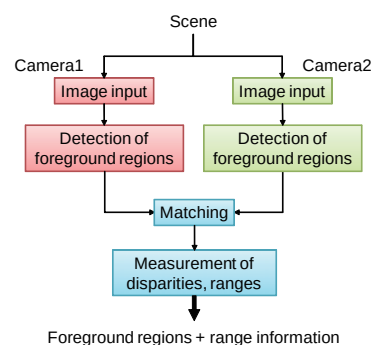


Fig. 3 Basic algorithm of subtraction stereo

2. 人物検出

2.1 差分ステレオ

提案システムでは、画像中の人物を検出する手法として、固定されたステレオカメラを用いた、差分ステレオ (Umeda, et al., 2009) と呼ばれる手法を用いる。差分ステレオの基本アルゴリズムを図3に示す。通常のステレオカメラでは、左右カメラの画像をマッチングすることで視差画像を得る。これに対し、差分ステレオでは左右カメラそれぞれで背景差分によって前景領域を抽出し、その後に抽出された前景領域の画像情報のみを用いてマッチングを行い、距離情報を得る。このように、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。差分ステレオによって図4(a)から検出された前景領域を図4(b)に青色で示す。なお、この前景領域は距離情報も持っている。

2.2 影検出

差分ステレオによって抽出された前景領域には、物体の影となる領域も含まれてしまう。これにより、人物検出の際に、人物領域に影が含まれる問題や影領域が人物として誤検出されるといった問題が生じる。そのため、影となる領域を前景領域から除去する必要がある。本論文では、文献 (Moro, et al., 2009) の手法により影領域を除去する。画像座標 (x, y) における輝度値を $I_{(x, y)}$ 、背景画像における同位置の輝度値を $I'_{(x, y)}$ とすると、影を判定する

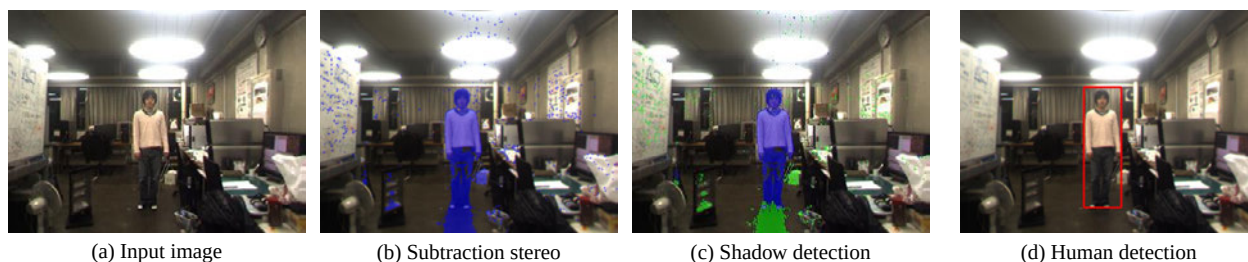


Fig. 4 Flow of foreground detection

評価関数は次式で表わされる.

$$\theta_{(t+1, x, y)} = \begin{cases} \alpha\Psi(x, y) + \beta\Lambda(x, y) + (1 - \alpha - \beta)\theta(t, x, y), & \left(\frac{I(x, y)}{\eta} < I'_{(x, y)}\right) \\ \infty, & (\text{otherwise}) \end{cases} \quad (1)$$

ここで, $\theta_{(t+1, x, y)}$ は影と判断するスコアを表し, $\theta_{(t+1, x, y)}$ 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 は前景の注目画素の近傍画素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 4(c) に示す. 図中の緑色の領域が影として検出されている. 検出された影領域を前景領域から除去し, 最終的に検出された人物を図 4(d) に示す. なお, 図 4(d) で出力される矩形は, 図 4(c) で抽出された前景領域を囲う形で出力される. 以後, 同図のように抽出される矩形の中心座標を人物の重心と定義する. また, 人物の重心は検出された前景領域 (追跡領域) の重心の画素の 3 次元位置より求める.

3. 人 物 追 跡

3.1 処理概要

提案システムでは, 2 章で述べた差分ステレオにより検出された人物の重心に対してパーティクルフィルタを適用し, ワールド座標系, X, Y 平面上にて追跡を行う. 主な処理は, 予測, 尤度計算, 対応付け, リサンプリングの 4 つである. 以上 4 つの処理を毎フレーム繰り返すことで追跡を行なう. 尤度は, 追跡対象の重心と各パーティクルの位置の誤差がどれだけ小さいかを表している. 対応付け処理は, 追跡中の人物とパーティクルを対応付ける. 提案システムでは, 対応付けの際に, 人物の進行方向を利用することで位置情報と色情報のみでは対応付けしきれない状況においても対応付けを可能にする. 図 5 にパーティクルフィルタの処理を図式化したものを示す. 図中の矩形は差分ステレオにより検出された人物領域を表している. 図 5(a) の白い点群は, 人物の重心周辺にランダムに散布されたパーティクル群を表している. なお, 図中の点群の数は, 説明のため実際に散布する数より少なく描かれている. 図 5(b) の青丸は予測処理を行う前のパーティクル, 橙色の丸は予測処理が終了した後のパーティクルを表している. 図 5(c), (d) の半透明の緑色の円は, 各パーティクルの尤度の大きさを視覚的に表している. 図 5(c) の赤点は, 尤度によって重み付けられたパーティクル群の重心位置を表している. また, 図 5(d) の赤いバツ印は, リサンプリング処理より削除されたパーティクルを表している. パーティクルフィルタの処理の詳細を以下に示す.

3.2 パーティクルフィルタ

(i) 初期化 | Initialization

人物検出時にパーティクルを人物重心周辺にランダムに散布する.

(ii) 予測 | Prediction

過去の追跡対象の位置情報から現フレームでの追跡対象の位置を予測し, 各パーティクルを遷移させる. その際, 前フレームで対応付けが成功したパーティクルと失敗したパーティクルでは, 異なる予測処理を行う.

- ・ 対応付け成功時

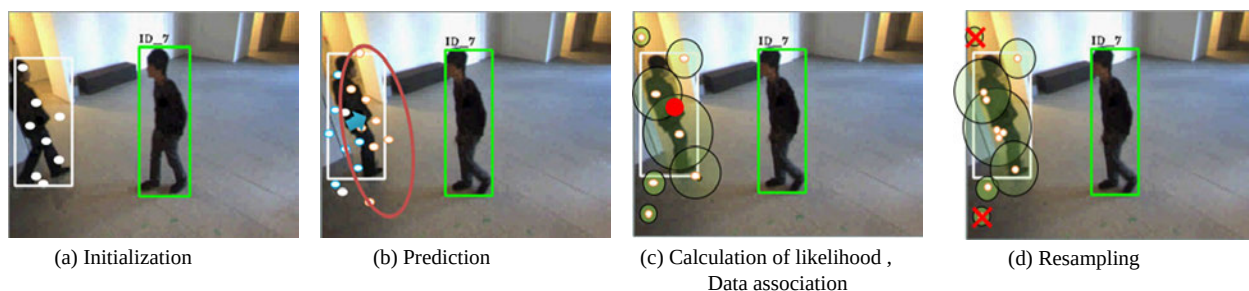


Fig. 5 Flow of particle filter

等速直線運動を仮定し、現フレームの追跡対象の位置を前フレームの位置情報から予測する。また、予測値には、追跡対象のランダムな動きにも対応するために乱数を加える。

・対応付け失敗時

パーティクルが保持している追跡対象の過去の移動量の平均から等速直線運動を仮定し、現フレームでの追跡対象の位置を予測する。対応付け失敗時には、乱数によってパーティクルが誤った方向へ遷移することを防ぐため、予測値に乱数は加えない。これにより、人物同士のすれ違い等で追跡対象が検出されない場合でも、平均速度で予測を行うことで追跡対象のおおよその位置を把握できる。そのため、すれ違いが終わった後に再び検出された場合でも、ロバストに再追跡が可能となる。

(iii) 尤度計算 | Calculation of likelihood

人物の重心と各パーティクルのワールド座標系, X, Y 平面上のユークリッド距離より各パーティクルの尤度を求める。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。なお、尤度計算はパーティクル群と最も距離が近い人物の重心より行う。

(iv) 対応付け | Data association

各フレームでパーティクルの追跡対象を探索し、人物とパーティクルの対応付けを行う。対応付けは、パーティクルの保持している位置情報・色情報・対象の進行方向と各フレームで人物から得られる上記3項目をそれぞれ比較することで行う。詳細は、3.3項にて後述する。

(v) リサンプリング | Resampling

各パーティクルの尤度に従い、追跡に有効なパーティクルの選択を行う。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する。尤度が小さいパーティクルは、その位置周辺に人物の重心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近傍に収束し、安定した人物の追跡が可能となる。

以上の処理を各フレームで繰り返すことで、人物の追跡を行う。また、追跡が一定期間続いた場合は、人物の識別番号である ID をパーティクルに付与する。

3.3 対応付け処理

対応付けにおける、追跡対象の判別方法について述べる。追跡対象の判別には、各フレームで取得した位置情報・色情報・進行方向の3つの評価値と過去の情報とを比較し、追跡中の人物を探索する。その後、追跡対象と判定された追跡領域とパーティクルの関連付けを行う。以後、対応付けに用いる位置情報・色情報・進行方向の評価方法について述べる。

a. 位置情報による評価方法

位置情報による対応付けの評価値は、式(2)により重み付けられたパーティクル群の重心位置と各人物の重心位置のユークリッド距離 D_d を比較することで算出する。評価値 D_d は、次式で求められる。

$$D_d = \sqrt{(X_p - X_h)^2 + (Y_p - Y_h)^2} \quad (3)$$

(X_p, Y_p) はパーティクル群の重心位置, (X_h, Y_h) は各フレームで検出された人物の重心位置を表している.

b. 色情報による評価方法

色情報による評価について述べる. 提案システムでは, HSV 色空間の Hue(色相) と Saturation(彩度) を用いて, 人物の色を表現する. 色情報の取得は, 差分ステレオによって抽出された人物領域 (図 3(d)) のうち, 彩度が閾値以上の画素の色相を取得し, ヒストグラムを作成することで行う. これは, 彩度の値が低いと, 色相の値が安定せず, 追跡対象の推定が困難であるためである. 色情報による評価値は, 各フレームで得られた色相ヒストグラムとパーティクルが保持している前フレームで対応づいた人物の色相ヒストグラムの類似度を計算することで取得する. 類似度の計算は Bhattacharyya 距離により行う. Bhattacharyya 距離 D_c は, 次式で求められる.

$$D_c = \sqrt{1 - \sum_{u=1}^m \sqrt{p_u q_u}} \quad (4)$$

ここで, p, q は比較する正規化された色ヒストグラム, u は色相番号, m は色相の成分数を表す.

c. 進行方向による評価方法

対象の進行方向による評価方法について述べる. 進行方向による評価は, 以下の 2 つの値の差を計算することで行う.

1. パーティクルの保持している移動履歴から算出した進行方向
2. 各フレームの各人物の進行方向

まず, 進行方向の求め方を述べる. 1. の進行方向を求める場合は, 各パーティクルの保持している対応付けられた人物の過去の位置情報を用いることで算出する. しかし, 進行方向を算出する際に必ずしも誤差の少ない位置情報を取得できるとは限らない. そのため, 実際の進行方向とは異なるものが算出される可能性がある. そこで, 各フレームで算出した進行方向と, 過去に算出した進行方向を比較する. この時, 一定値以上の差がある場合, 各フレームで算出した進行方向は有用でないと判定し, 過去の位置情報のみから進行方向を算出する. 図 6 に時系列データより進行方向の選択を行う例を示す. 図中の矢印は各フレームで算出された進行方向を表している. また, 赤いバツ印は, 算出された進行方向が過去に算出された進行方向と一定値以上の差があるため有用でないと判定されたことを表している.

2. の進行方向を求める際, 対応付け処理を行う前に進行方向の算出を行う. この時, 対応付けを処理を行う前なので, 各フレームでの人物の過去の位置情報は不明なため, 進行方向は算出できない. そこで, 各フレームの人物の重心位置と一定距離内にあるパーティクル群の過去の重心位置情報から進行方向を算出する.

図 7 に進行方向算出の例を示す. 図 7 では, 赤パーティクルに関連した進行方向のみを示している. 赤い矢印は, パーティクルから算出した進行方向 (1. の進行方向) である. 緑の点線矢印は, 緑の人物と赤線の位置情報を保持しているパーティクルが対応付いていたと想定し算出した進行方向である (2. の進行方向). 青い点線矢印は, 黒の人物と赤線の位置情報を保持しているパーティクルが対応付いていたと想定し算出した進行方向である (2. の進行方向). また, 右図は左図の俯瞰図であり, 矢印等は左図と対応している.

次に人物の歩行速度から算出した進行方向を用いるかの判定を行う. 例えば, 人がほぼ止まっている状態の時は, どちらの方向に動くか予想は難しい. 逆に人の動く速度が大きいときは, 大きく進行方向を変えることは少ないと考えられる. そのため, 速度が一定値以上の場合のみ, 進行方向を評価値として用いる. 進行方向による評価値 D_a を次式で求める.

$$D_a = \begin{cases} k^\alpha - 1, & (\alpha < \alpha_{thr}) \\ k, & (otherwise) \end{cases} \quad (5)$$

k は重みとして用いる定数, α はパーティクルの保持している位置情報から算出した進行方向 (1. の進行方向) と各フレームで算出した人物の進行方向 (2. の進行方向) の角度差 (図 7 中 α, α'), α_{thr} は閾値を表している. 本システムでは k は 1.5, α_{thr} は 1.22[rad] と設定している.

以上の位置・色・進行方向の評価値 D_d, D_c, D_a を足しあわせ, 最小の値を持つ人物を追跡中の人物とし, 対応付けを行う. 対応付け終了後, パーティクルの保持している色情報・位置情報を更新し, 対応付けられた人物と尤度計算を再び行う.

Table 1 Camera installation conditions

	Height [m]	Tilt angle [°]
Scene 1	2.4	30
Scene 2	6.4	60
Scene 3	5.9	40

3.4 情報の選択

実環境では、人物同士のすれ違い等でオクルージョンが頻繁に生じる。提案システムでは、そのような人物同士の重なりが生じた際には、距離画像をセグメンテーションする (Ubukata, et al., 2010) ことで人物を個々に検出している。しかし、ステレオカメラの特性上、対象との距離が大きくなるにつれて、距離の誤差も大きくなる。そのため、カメラと人物との距離が遠い場合、人物領域の切り分けが出来ず、適切に検出が行われない場合がある。図 8 の赤丸で囲まれている部分に、人物検出の失敗例を示す。適切に人物検出ができない場合、人物の重心がずれて算出される。これが原因でパーティクルが追跡すべき方向から外れて遷移し、誤追跡が発生してしまうことがある。図 8 のように複数人物が同時に検出された場合は、人物毎に検出された場合に比べ、矩形の高さ・幅が大きくなる。また、画像中の人物の大きさがカメラからの距離に反比例の関係にあることから検出された位置により検出矩形の最大の大きさが推定できる。そこで、推定される検出矩形の大きさより一定以上大きい矩形が出力された場合、そのデータは用いないこととする。

4. 評価実験

4.1 撮影条件

提案システムの有用性を検証するために、実環境での運用を想定した 3 シーンで実験を行った。それぞれのシーンの実験環境を図 9 に、カメラ設置時の高さ、仰角を表 1 に示す。具体的な実験の実施場所は、以下の通りである。

Scene 1 : 中央大学後楽園キャンパス新 2 号館 1 階ホール

Scene 2 : 中央大学後楽園キャンパス旧 2 号館前中庭

Scene 3 : AKIBA SQUARE

Scene 1 は、最大で 10 人程度が画面内に存在し、人物がランダムに歩行するといった内容で、動画の撮影時間は 150 秒である。Scene 2, 3 は、最大で 15 人程度が画面内に存在し、立ち止まる、引き返す等の行動をランダムに行ったもので、動画撮影時間はそれぞれ 135 秒、240 秒である。

ステレオカメラは Bumblebee2 (Point Grey Research, VGA) を使用し、20fps で保存した動画に対して処理を行った。使用した PC の CPU は Intel Core2 Extreme (2.93 GHz)、メモリは 3 GB (RAM) である。

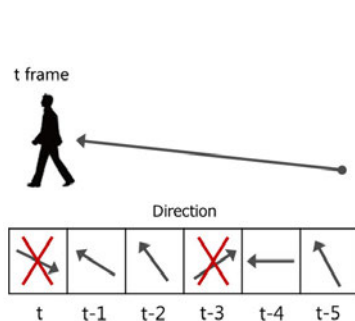


Fig. 6 Definition of direction from time series data of calculated directions

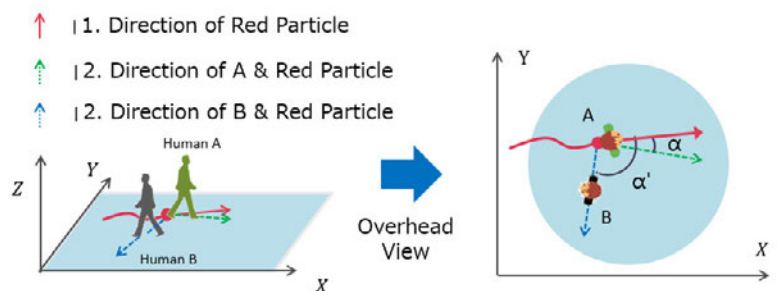


Fig. 7 Calculation of human directions and angular difference between particle's direction and human directions

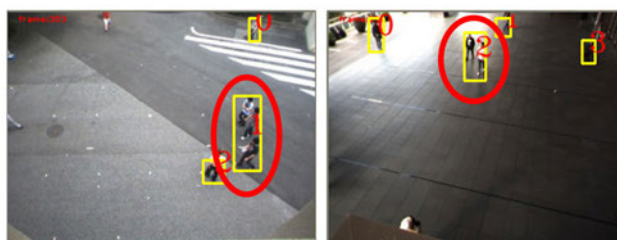


Fig. 8 Failure of human detection



Fig. 9 Experimental scenes

4.2 複数人物追跡実験

提案システムの有用性を検証するために、進行方向・距離・色情報の3つの評価値を用いた場合と距離・色情報のみを用いた場合での人物追跡精度を比較した。追跡には、各人物毎 500 個ずつパーティクルを散布した。それぞれのシーンでの実験の様子を図 10(a)~(c) に示す。各図の上図は入力画像、下図は出力結果を表している。下図の色の着いた矩形は、差分ステレオにより検出された人物領域を表し、視認性向上のため人物毎に振られている ID に応じて矩形の色分けを行っている。図中にある番号は、人物毎に割り振られた ID である。また、色のついた丸は、追跡を行っているパーティクル群の重心位置を表している。色付けについては、矩形と同様である。白丸は、まだ ID が振られていないパーティクルを表している。

実験結果を表 2 に示す。人物追跡精度の評価は、人物が検出され ID が与えられてから、画面外に出るまで ID が変化しない場合、人物追跡成功とすることで行った。また、複数人が重なった状態で画面内に入ってきた場合は、ひとつの追跡対象として追跡を行う。その後、追跡対象が別れた場合は、新たに検出された追跡対象に新たにパーティクルを散布し、そこから追跡を開始する。実験結果より、人物の進行方向を考慮して対応付けを行うことで、人物追跡精度が向上することがわかった。これは、進行方向を考慮することで、人物同士のすれ違いなどが生じる際に、追跡対象を判別できるためである。特に人物同士の服の色が似ている状況で有効であることがわかった。また、9~11fps で処理を行えるため、オンラインでの処理も可能である。Scene 1 において、色と距離情報のみではカメラ右奥の位置では人物がすれ違う際に ID の入れ替わりが起こることがあったが、進行方向を用いることで ID の入れ替わりなく追跡が可能となった。Scene 2, Scene 3 では、影の影響及び、人物同士の服装が似ていることから、色情報の評価値が対応付けに有用でない場合があった。しかし、提案システムでは進行方向を用いることで対応でき、正確な対応付けが可能となった。図 10(b) の $t = 35s$ から $t = 42s$ の間では、画面左へ歩く集団と上方向に歩く人物のすれ違い、 $t = 53s$ から $t = 63s$ の間では、左へ歩く集団と右へ歩く集団とのすれ違いが発生している。位置・色情報のみの場合、複数人の ID 変化が発生したが、進行方向を考慮することで ID の変化なく正確な追跡が可能となっている。

一方で、人物追跡が失敗してしまうシーンも見受けられた。特に、Scene 1 のように、カメラと人物の距離が近い場合には、オクルージョンの影響が大きく、人物検出できない期間が長くなる。そのため、その間の予測値と追跡対象の位置が大きくずれてしまい、進行方向を考慮しても対応付けできなくなるケースが生じる。追跡が失敗したシーン例を図 11 に示す。図の見方は図 10 と同様であり、赤丸で囲まれた人物に対しての追跡の失敗の過程

Table 2 Experimental results

	Number of detected person	Tracking Accuracy (Distance & Color) [%]	Tracking Accuracy (Distance & Color & Angle) [%]
Scene 1	40	88	90
Scene 2	35	88	94
Scene 3	51	86	92

を表している。t = 202s からすれ違いにより赤丸の人物の検出が失敗し、t = 205s で再検出された際には ID の変化が生じている。提案システムでは、人物の重心 (検出矩形の中心) を追跡しているが、Scene 1 では矩形の大きさの変動が大きく、それが原因で追跡が失敗してしまうことも考えられる。また、複数人物を同時に検出したことで人物の重心位置がずれてしまい、その情報を用いてしまうことによって追跡が失敗してしまうシーンも見られた。この点に関しては、人物検出が正確に行えているかの判定を、検出する際の矩形の大きさによる現状の判定だけでなく、他の判定基準の導入による拡張が必要であると考えられる。

5. お わ り に

本論文では、差分ステレオにより得られた距離情報と人の色情報と進行方向を利用したパーティクルフィルタによる複数人物追跡システムを構築した。例えば、服の色が似ている人物がすれ違う場合、人物間の距離も近いいため位置と色の評価値の差が出ないため、追跡中の人物の判別が難しい。しかし、すれ違い時には各人物の進行方向は差がでるため、提案システムを用いることで、このような状況においても正確な追跡対象の判別が可能となる。また、提案システムの有用性を実証するため、監視カメラの運用を想定した 3 シーンの実環境で実験を行った。結果より人物の進行方向を考慮することで人物追跡精度が向上し、提案システムがすれ違い等で生じるオクルージョンに対して有効であることを示した。

追跡失敗の原因として、人物同士のすれ違いなどで生じるオクルージョンにより、人物が検出されない期間が長くなることで予測が大きくずれてしまうが挙げられる。提案システムでは、人物が検出されないことで予測が多少ずれてしまった場合においても、進行方向を用いることで正確な対応付けが可能となった。しかし、人物が検出されない期間が長い場合、予測が大きくずれてしまい、人物再検出時に対応付けが正確に行えず ID が変化してしまう問題が生じた。そこで対策として、まず人物検出精度を向上させることが考えられる。そのために、背景差分の改良、または、近年人物検出手法として有効性が報告されている Joint HOG (三井他, 2009) を用い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を詳細に取得できるため、予測のずれが大きくなる前に対応付けが行え、情報を更新できるため追跡精度が向上すると考えられる。次に対策として、人物検出がされない間のパーティクル群の予測を改良することが考えられる。予測を正確に行うことで、オクルージョン等で人物が検出されないシーンが長くなっても、再検出時の予測値のずれが小さくなるため正確な対応付けが期待できる。また、距離情報・色情報・進行方向といった追跡対象判別の際に用いた評価値を、単純に足し合わせて用いるのではなく、状況によって評価値を重み付けして用いることが挙げられる。さらに、異なる実環境においての実験を行い、本システムの有用性の検証を行う。

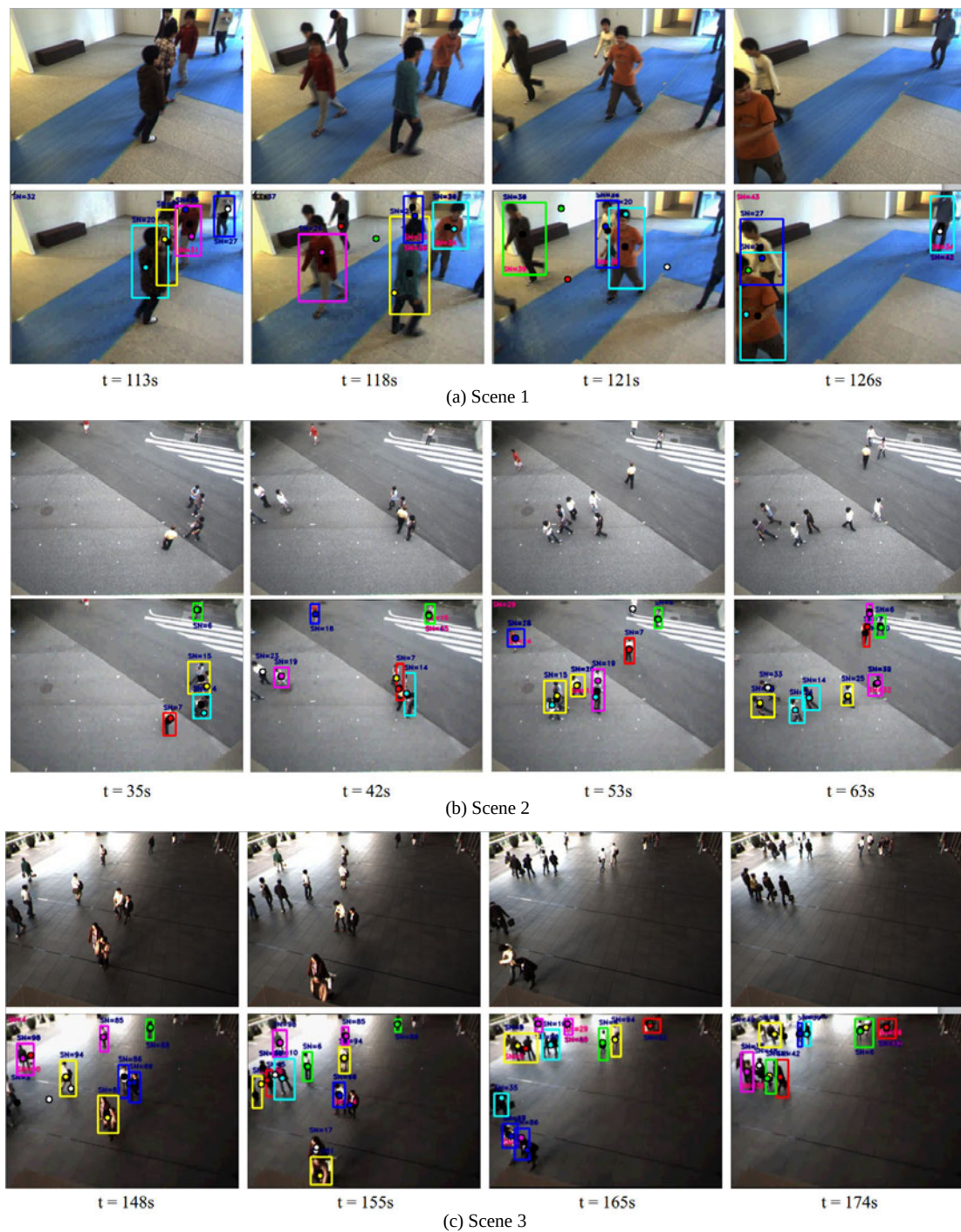


Fig. 10 Snapshots of human tracking

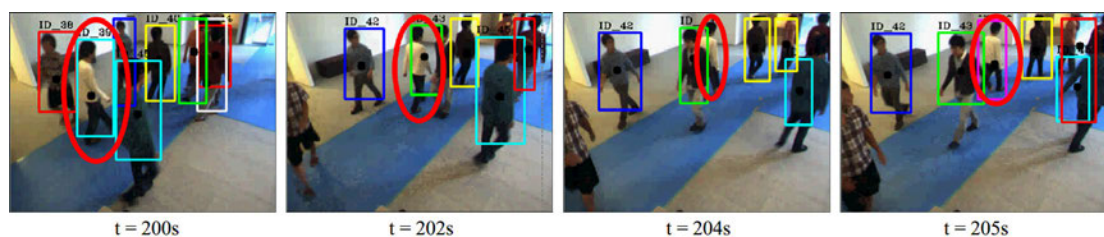


Fig. 11 Failure of human tracking: the system fails to track the human who is enclosed in a red ellipse.

文 献

- Breitenstein, M. D., Reichlin, F., Leibe, B., Koller-Meier, E. and Gool, V. L., Robust tracking-by-detection using a detector confidence particle filter, Proceedings of the 2009 IEEE International Conference on Computer Vision (2009), pp.1515-1522.
- Gordon, N., Salmond, D. and Ewing, C., Bayesian state estimation for tracking and guidance using the bootstrap filter, Journal of Guidance, Control, and Dynamics, Vol.18, No.6 (1995), pp.1434-1443.
- Gorji, A. A., Tharmarasa, R. and Kirubarajan, T., Performance measures for multiple target tracking problems, Information Fusion (FUSION) (2011), pp.1-8.
- 原口真悟, 福司謙一郎, 熊澤逸夫, 多数のステレオカメラと色情報を用いた人物追跡手法, 電子情報通信学会技術研究報告: 信学技報, Vol.109, No.471 (2010), pp.229-234.
- Huchuan, L., Ruijuan, Z. and Yen-Wei, C., Head detection and tracking by mean-shift and kalman Filter, Innovative Computing Information and Control (2008), p.357.
- Ikoma, N., Yamaguchi R., Kawano H. and Maeda, H., Tracking of multiple moving objects in dynamic image of omnidirectional camera for robot vision using PHD filter, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.12 (2008), pp.16-25.
- Kang-Il, P., Chan-Ik, P. and Jangmyung, L., Vision based people tracking using stereo camera and camshift, Intelligent Robotics and Applications, Vol.8102 (2013), pp.145-153.
- Kalman, R. E., A new approach to linear filtering and prediction problems, Transactions ASME - Journal of Basic Engineering, Series D, No.82 (1960), pp.35-45.
- 片岡裕雄, 田村仁優, 松井靖浩, 青木義満, 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), pp.1108-1115.
- 小林貴訓, 杉村大輔, 関真規人, 平澤宏祐, 鈴木直彦, 鹿毛裕史, 佐藤洋一, 杉本晃宏, 分散カメラとレーザ測域センサの統合によるエリア内人物追跡, 画像の認識・理解シンポジウム (MIRU) (2007), pp.IS-2-23.
- Kovacevic, J., Juric-Kavelj, S. and Petrovic, I., An improved camshift algorithm using stereo vision for object tracking, Proceedings of the 34th MIPRO International Convention (2011), pp.707-710.
- Kurazume, R., Yamada, H., Murakami, K., Iwashita, Y. and Hasegawa, T., Target tracking using SIR/MCMC particle filters by multiple cameras and laser range finders, Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (2008), pp.3838-3844.
- Liebens, M., Sakiyama, T. and Miura, J., Visual tracking of multiple persons in a heavy occluded space using person model and joint probabilistic data association, Multisensor Fusion and Integration for Intelligent Systems (2006), pp.547-552.
- Lu, W., Okuma, K. and Little, J. J., Tracking and recognizing actions of multiple hockey players using the boosted particle filter, Image and Vision Computing, Vol.27, No.1-2 (2009), pp.189-205.
- 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.J92-D, No.9 (2009), pp.1591-1601.
- Moro, A., Terabayashi, K., Umeda, K. and Mumolo, E., Auto-adaptive threshold and shadow detection approaches for pedestrians detection, Proceedings of the 2009 AWSVCI (2009), pp.9-12.
- 中村克行, 趙卉菁, 柴崎亮介, 坂本圭司, 大鋸朋生, 鈴川尚毅, 複数のレーザレンジスキャナを用いた歩行者トラッキングとその信頼性評価, 電子情報通信学会論文誌, Vol.37, No.7 (2006), pp.1-11.
- 西濃拓郎, 滝口哲也, 有木康雄, 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 (MIRU) (2010), IS1-39, pp.307-312.
- 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と可視化手法, 電子情報通信学会論文誌, Vol.J93-D, No.4 (2010), pp.486-493.
- Ronald, M., A theoretical foundation for the stein-winter "probability hypothesis density (PHD)" multitarget tracking approach, Proceedings of the 2002 MSS National Symposium on Sensor and Data Fusion (2002).
- 曾我部光司, 倉爪亮, 岩下友美, 長谷川勉, JPDAF とカラーヒストグラムを用いた複数センサによる移動体追跡, 第

11 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 (2010), 113-2.

圓田直樹, 福井真二, 倉橋 渉, 武知啓介, 岩堀祐之, 外観の類似する物体同士の交差に頑健な移動物体追跡手法, 電子情報通信学会論文誌, Vol.130, No.9 (2010), pp.1588-1596.

高橋和彦, 橋本雅文, 根田知樹, 白 志濤, 2 層レーザレンジセンサによる急加減速を伴う複数人物の検出・追跡, 日本機械学会論文集 C 編, Vol.78, No.785 (2012), pp.138-150.

多田和樹, 丁明, 竹村裕, 溝口博, Spatiograms を用いた平均値シフトと確率的予測による頑健な物体追跡手法, 日本機械学会論文集 C 編, Vol.78, No.787 (2012), pp.799-811.

Ubukata, T., Terabayashi, K., Moro, A. and Umeda, K., Multi-object segmentation in a projection plane using subtraction stereo, Proceedings of 20th International Conference on Pattern Recognition (2010), pp.3296-3299.

Umeda, K., Hashimoto, Y., Nakanishi, T., Irie, K. and Terabayashi, K., Subtraction stereo - a stereo camera system that focuses on moving regions -, Proceedings of the 2009 SPIE, Vol.7239 (2009), 723908.

References

Breitenstein, M. D., Reichlin, F., Leibe, B., Koller-Meier, E. and Gool, V. L., Robust tracking-by-detection using a detector confidence particle filter, Proceedings of the 2009 IEEE International Conference on Computer Vision (2009), pp.1515-1522.

Gordon, N., Salmond, D. and Ewing, C., Bayesian state estimation for tracking and guidance using the bootstrap filter, Journal of Guidance, Control, and Dynamics, Vol.18, No.6 (1995), pp.1434-1443.

Gorji, A. A., Tharmarasa, R. and Kirubarajan, T., Performance measures for multiple target tracking problems, Information Fusion (FUSION) (2011), pp.1-8.

Haraguchi, S., Hukushi, K. and Kumazawa, I., Person tracking using multiple stereo cameras with color information, The Institute of Electronics, Information and Communication Engineers, Vol.109, No.471 (2010), pp.229-234 (in Japanese).

Huchuan, L., Ruijuan, Z. and Yen-Wei, C., Head detection and tracking by mean-shift and kalman Filter, Innovative Computing Information and Control (2008), p.357.

Ikoma, N., Yamaguchi R., Kawano H. and Maeda, H., Tracking of multiple moving objects in dynamic image of omnidirectional camera for robot vision using PHD filter, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.12 (2008), pp.16-25.

Kang-Il, P., Chan-Ik, P. and Jangmyung, L., Vision based people tracking using stereo camera and camshift, Intelligent Robotics and Applications, Vol.8102 (2013), pp.145-153.

Kalman, R. E., A new approach to linear filtering and prediction problems, Transactions ASME - Journal of Basic Engineering, Series D, No.82 (1960), pp.35-45.

Kataoka, H., Tamura, K., Matsui, Y. and Aoki, Y., Improvement CoHOG detection and texture matching tracking for pedestrian's active safety, Transactions of the IEICE, Vol.132, No.7 (2012), pp.1108-1115 (in Japanese).

Kobayashi, Y., Sugimura D., Seki, M., Hirasawa, K., Suzuki, N., Kage, H., Sato, Y. and Sugimoto, A., People tracking by integrating observations from distributed cameras and laser range scanners, MIRU (2007), pp.IS-2-23 (in Japanese).

Kovacevic, J., Juric-Kavelj, S. and Petrovic, I., An improved camshift algorithm using stereo vision for object tracking, Proceedings of the 34th MIPRO International Convention (2011), pp.707-710.

Kurazume, R., Yamada, H., Murakami, K., Iwashita, Y. and Hasegawa, T., Target tracking using SIR/MCMC particle filters by multiple cameras and laser range finders, Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (2008), pp.3838-3844.

Liebens, M., Sakiyama, T. and Miura, J., Visual tracking of multiple persons in a heavy occluded space using person model and joint probabilistic data association, Multisensor Fusion and Integration for Intelligent Systems (2006), pp.547-552.

Lu, W., Okuma, K. and Little, J. J., Tracking and recognizing actions of multiple hockey players using the boosted particle filter, Image and Vision Computing, Vol.27, No.1-2 (2009), pp.189-205.

- Mitsui, T., Yamauchi, Y. and Fujiyoshi, H., Object detection by two-stage boosting with joint features, The Institute of Electronics, Information and Communication Engineers, Vol.J92-D, No.9 (2009), pp.1591-1601 (in Japanese).
- Moro, A., Terabayashi, K., Umeda, K. and Mumolo, E., Auto-adaptive threshold and shadow detection approaches for pedestrians detection, Proceedings of the 2009 AWSVCI (2009), pp.9-12.
- Nakamura, K., Zhao, H., Shibasaki, R., Sakamoto, K., Ohga, T. and Suzukawa, N., Tracking pedestrians using multiple single-row laser range scanners and its reliability evaluation, The Institute of Electronics, Information and Communication Engineers, Communications and Computer Sciences, Vol.37, No.7 (2006), pp.1-11 (in Japanese).
- Nishino, T., Takiguchi, T. and Ariki, Y., Tracking of multiple soccer players using 3D particle filter based on multiple likelihood, MIRU (2010), pp.307-312 (in Japanese).
- Onishi, M. and Yoda, I., Comparison and visualization method of pedestrian trajectories in a complex office for a long period, The Institute of Electronics, Information and Communication Engineers, Vol.93, No.4 (2010), pp.486-493 (in Japanese).
- Ronald, M., A theoretical foundation for the stein-winter "probability hypothesis density (PHD)" multitarget tracking approach, Proceedings of the 2002 MSS National Symposium on Sensor and Data Fusion (2002).
- Sokabe, K., Kurazume, R., Iwashita, Y. and Hasegawa, T., Target tracking by laser sensor and camera using JPDAF and color histogram, System Integration Division Annual Conference (2010), 113-2 (in Japanese).
- Sonoda, N., Hukui, S., Kurabashi, W., Takechi, K. and Iwahori, Y., Robust tracking method to intersections of objects with similar patterns, The Institute of Electronics, Information and Communication Engineers, Series C, Vol.130, No.9 (2010), pp.1588-1596 (in Japanese).
- Takahashi, K., Hashimoto, M., Konda, T. and Bai, Z., Detection and tracking of people with sudden changes in motion using two-layered laser range sensors, Transactions of the JSME, Series C, Vol.78, No.785 (2012), pp.138-150 (in Japanese).
- Tada, K., Ding, M., Takemura, H. and Mizoguchi, H., Robust tracking method by mean-shift using spatiograms and particle-filter, Transactions of the JSME, Series C, Vol.78, No.787 (2012), pp.799-811 (in Japanese).
- Ubukata, T., Terabayashi, K., Moro, A. and Umeda, K., Multi-object segmentation in a projection plane using subtraction stereo, Proceedings of 20th International Conference on Pattern Recognition (2010), pp.3296-3299.
- Umeda, K., Hashimoto, Y., Nakanishi, T., Irie, K. and Terabayashi, K., Subtraction stereo - a stereo camera system that focuses on moving regions -, Proceedings of the 2009 SPIE, Vol.7239 (2009), 723908.