

LI-007 MCMC 法に基づく 3 次元環境情報を用いた複数人物追跡 MCMC based Multiple People Tracking using 3D Environmental Information

大澤 達哉[†] ウ 小軍[†] 数藤 恭子[†] 若林 佳織[†] 安野 貴之[†]
Tatsuya Osawa Xiaojun Wu Kyoko Sudo Kaoru Wakabayashi Takayuki Yasuno

1. はじめに

近年、高齢化社会の到来や社会不安の増大から監視カメラを用いた映像監視技術に大きな期待が寄せられている。映像監視技術の中でも人物追跡技術は人の動きに関する情報が人の行動およびその意図と関係していることから最も重要な技術の一つである。

実環境においても安定した人物追跡を可能とするには、監視エリアに存在する静止物体によるオクルージョン、人物の出入り、人物同士の接近による相互オクルージョンといった問題に対処する必要がある。これまでも、このような状況下において 2 次元画像上で人物を追跡する技術はすでに多く提案されている [1, 2]。しかしながら、我々の住む実世界と同じ 3 次元空間において人物を追跡することができれば、オクルージョンや人物の出入り、人物同士の接近といった問題により有効に対処できるというだけでなく、精度良く 3 次元空間中の人物位置を推定することで人物の行動意図理解へ活用できることが期待できる。

そこで本稿では、予め獲得しておいた 3 次元環境情報と 3 次元人物モデルを用いて監視エリア内の状態、すなわち監視エリア内に存在する人物の数、各人物の 3 次元空間中の位置およびその大きさをシミュレートすることで複数人物の追跡を行う手法を提案する。3 次元環境情報とは監視エリア内に存在する静止物体の 3 次元構造情報、追跡に用いる全てのカメラの外部・内部パラメータ、人物の出入りが行われるエリアに関する情報を指す。この 3 次元環境内に 3 次元人物モデルを配置して仮想のシーンを構築し、これを実カメラと同じパラメータを持つ仮想カメラで撮影することで実カメラから撮影された画像（実画像）をシミュレートした画像（シミュレーション画像）を生成する。実画像とシミュレーション画像の整合性と事前知識による状態の確からしさ（環境中の人物位置、人物同士の相対的な位置）を尤度として採用し、MCMC 法により監視エリア内の状態推定を逐次行う。

提案手法の有効性を確かめるために、2 台のカメラを用いたシステムで静止物体によるオクルージョンや人物の出入り、人物同士の接近があっても安定した追跡が可能であることを示す。

2. 3 次元環境情報

本研究では 3 次元環境情報を、追跡に用いるカメラの映像から取得する。追跡に用いる各カメラを所定の位置（追跡処理の際に実際に利用する位置）に設置する際に、カメラを移動させながら画像列を撮影する。撮影した画像列に対し、因子分解法とマルチビューステレオの融合法 [3] を適用することで、監視エリアの 3 次元点群およびカメラ運動情報を取得する。各カメラから独立に 3 次

元点群を復元するため、全てのカメラで統一された座標系に各カメラから復元された 3 次元点群を統合する必要がある。

ここではステレオ法で 3 次元点群を復元する際に、基準画像を各カメラ設置位置で撮影された画像とし、この基準画像上で少数の対応点を手動で与える。対応点に対応する各カメラより復元された 3 次元点同士を比較し、quaternions による手法 [4] を用いてカメラ間の座標変換行列を求め、統合を行った。

全 3 次元点群の統合後、3 次元 hough 変換 [5] を用いて、最大面積を持つ平面を床平面として検出し、この平面が XY 平面となるように 3 次元点群を変換し、人の出入りが行われる入退室エリアを手動で設定した。

図 1 に取得した画像列の一例、図 2 に復元した 3 次元環境情報を示す。



図 1: 3 次元環境情報復元のために取得した画像列の一例

3. 3 次元環境中での複数人物追跡

3.1 状態モデル

提案手法では監視エリア内の状態を複数の人物の状態の集合 S_t として定義する。各人物 i の状態 M_i は図 3 に示すように 3 次元楕円モデルを用いて、人物の位置を XY 平面上の座標 (x_i, y_i) 、人物の大きさを楕円体の半径および高さ (r_i, h_i) で表現した 4 次元ベクトル $M_i = (x_i, y_i, r_i, h_i)$ として人物状態を定義する。時刻 t において追跡している人物数を K とすれば、監視エリア内の時刻 t における状態 S_t は $S_t = (M_1, M_2, \dots, M_K)$ によって表される 4 K 次元ベクトルとして定義する。

このように推定する状態ベクトルの次元数は追跡する人物数によって動的に変化する。この次元数の動的変化に対応するため状態推定には Trans-Dimensional MCMC 法 [6] を用いる。

3.2 現時刻状態の予測

まず前時刻 $t-1$ で推定された状態 S_{t-1} に基づいて現時刻の状態の予測を行う。状態の予測には単純な線形予測を用いており、前時刻において K 人の人物を追跡されていた場合には、4 K 次元の速度ベクトル V を用いて式 (1) によって暫定的な現時刻の状態 $\hat{S}_{t,0}$ を予測する。

$$\hat{S}_{t,0} = S_{t-1} + V_{t-1} \quad (1)$$

[†]日本電信電話株式会社, NTT サイバースペース研究所

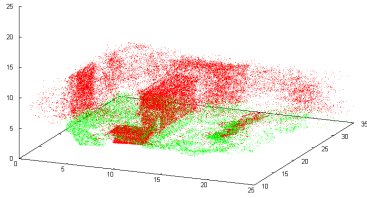


図 2: 復元された 3次元環境情報

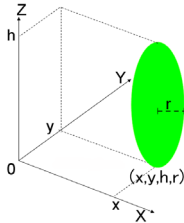


図 3: 3次元人物モデル

3.3 状態の逐次更新処理

予測された暫定的な状態 $\hat{S}_{t,0}$ を初期状態として、複数回の更新処理によって、現時刻の状態 S_t を推定する。図 4 にこの更新処理の処理フローを示す。この更新処理で

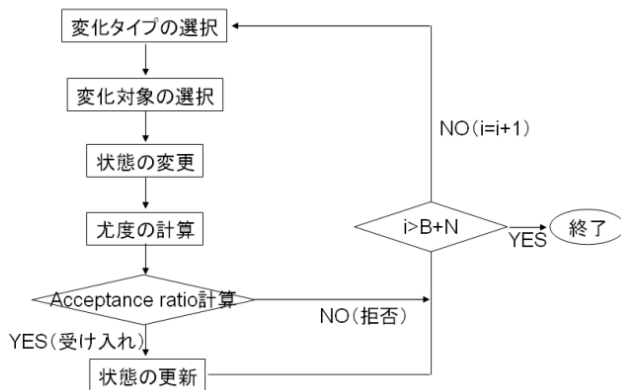


図 4: 更新処理による状態推定フロー

は人物の出現・消失、人物の位置、大きさの変化といった監視エリア内の状態を変化させて、その状態尤度を計算することで状態推定を行っていく。一度の更新処理で様々な状態を変化させるのは状態推定空間が広大であるため困難である。そこで各更新毎に状態を変化させるタイプと変化させる対象の選択を行う。更新回数は $B+N$ 回とし、最初の B 回の更新によって得た状態 $\hat{S}_{t,i < B}$ は推定精度が悪いと考えられるため利用せず、最後の N 回の更新によって得た状態 $\hat{S}_{t,B \leq i \leq N}$ を用いて現時刻の状態 S_t を計算する。

3.3.1 変化タイプおよび対象の選択

提案手法では変化タイプとして以下の4つを定義した。

1. ADD (現状態に新たな人物の追加)
2. DELETE (現状態中の人物の削除)
3. UPDATE-P (現状態中の人物の位置の変化)
4. UPDATE-S (現状態中の人物の大きさの変化)

これらの変化タイプから予め任意に決定した確率に従って、いずれかのタイプが選択される。ただし、現状態に誰も追跡している人物が存在しない場合には必ず ADD が選択される。また現状態で追跡中の人物が予め決めておいた入退室エリアに存在しない場合には DELETE が選択されることはない。

変化タイプが決定したら次に現状態中のどの対象を変化させるかランダムに選択を行う。ただし、ADD の場合には現状態の次元を増やして新たな対象を追加する。また DELETE の場合には入退室エリアに存在する人物の中からランダムに選択を行う。

3.3.2 状態の変化

ADD の場合 現状態に新たに人物状態を追加、すなわち新たな人物モデル M_n を生成し、現状態に追加する。人物の追加は予め決めておいた入退室エリアでのみ行う。人物を表す楕円体の平均的な大きさを μ_r, μ_h とすると、以下の式 2 によって M_n の生成を行う。

$$M_n = (\delta_x, \delta_y, N(\mu_r, \sigma_r), N(\mu_h, \sigma_h)) \quad (2)$$

ただし、 δ_x および δ_y は入退室エリアに限定した白色ノイズであり、 $N(\mu_r, \sigma_r)$ および $N(\mu_h, \sigma_h)$ はそれぞれ平均 μ_r, μ_h で分散 σ_r, σ_h のガウシアンノイズである。

DELETE の場合 現状態から選択された人物状態を一つ削除する。つまり人物 i が選択されたとすれば人物状態 M_i を現状態から削除する。

UPDATE-P の場合 現状態中から選択された人物状態 M_i 中の位置を表す (x_i, y_i) を以下の式 (4) により変化させる。

$$x_i = x_i + N(0, \sigma_x), y_i = y_i + N(0, \sigma_y) \quad (3)$$

ただし、 $N(0, \sigma_x), N(0, \sigma_y)$ はそれぞれ平均 0、分散 σ_x, σ_y のガウシアンノイズである。

UPDATE-S の場合 現状態中から選択された人物状態 M_i 中の大きさを表す (r_i, h_i) を以下の式 (4) により変化させる。

$$r_i = r_i + N(0, \sigma_r), h_i = h_i + N(0, \sigma_h) \quad (4)$$

ただし、 $N(0, \sigma_r), N(0, \sigma_h)$ はそれぞれ平均 0、分散 σ_r, σ_h のガウシアンノイズである。

3.3.3 状態尤度の計算

変化させた状態の尤度は、3次元環境情報と人物同士の相互位置関係に基づく事前知識と状態をシミュレートして仮想カメラによって撮影したシミュレーション画像と実画像の比較によって行う。

まず3次元環境情報によって、床面上でない位置（静止物体が存在）には人物が存在しないというペナルティ関数 $E(S)$ を式(5)のように定義することができる。

$$E(S) = \prod_{i=1}^K H(M_i) \quad (5)$$

ただし、 (x_i, y_i) が床面上の座標の場合 $H(M_i) = 1$ 、それ以外の場合には $H(M_i) = 0$ とする。

人物同士の相互位置関係に基づくペナルティ関数 $R(S)$ は人物同士が3次元空間上で重なることがないことから式(6)のように定義する。

$$R(S) = \prod_{i,j} \psi(M_i, M_j) \quad (6)$$

ただし、 $\psi(M_i, M_j)$ は人物 i と人物 j の距離 $D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ と任意に決定できる係数 λ を用いて、 $\psi(M_i, M_j) = 1 - \exp^{-\lambda D_{i,j}}$ とする。

最後に図5のように背景差分によって検出した人物領域を1 それ以外を0とした背景差分画像とシミュレーションによって生成した楕円体領域を1、それ以外を0とした背景差分シミュレーション画像を比較を行う。3次元環境情報を用いることで、シミュレーション画像は静止物体によるオクルージョンを予測できるため、オクルージョンの影響を排除することが可能である。画像の比較には以下の式(7)で表される人物領域と楕円体領域の重なり割合を全てのカメラ分足し合わせて、平均を取ったものを採用した。

$$V(S) = \frac{1}{C} \sum_{N=1}^C \frac{\sum_{k,l} Bg_N(k,l) \cap Sm_N(k,l)}{\sum_{k,l} Bg_N(k,l) \cup Sm_N(k,l)} \quad (7)$$

ただし、 C はカメラの台数、 $Bg_N(k,l)$ 、 $Sm_N(k,l)$ はそれぞれ、カメラ N の背景差分画像、シミュレーション画像の画素 (k,l) の値を表す。

最終的な尤度 L はこれらを掛け合わせた式(8)とする。

$$L(S) = E(S) \times R(S) \times V(S) \quad (8)$$

このようにすることで、人物の移動を床平面上に限定し、人物同士の3次元空間上での重なりを制限することができる。適切な状態推定を行うことができる。

3.3.4 Acceptance ratio の計算

計算された尤度を用いて変化させた状態を受け入れて新しい状態として更新するか、もしくは受け入れを拒否して前回更新された状態を今回の状態とするかの決定を行う。前回更新時の状態尤度を L_{old} 、今回の状態尤度を L_{new} とすると、確率 $\min(1, \frac{L_{new}}{L_{old}})$ で受け入れを行う。つまり今回の尤度の方が大きければ必ず受け入れ、そうでない場合にも $\frac{L_{new}}{L_{old}}$ の確率で受け入れを行う。

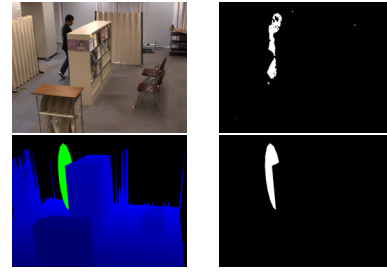


図5: 上段(左:実画像, 右:背景差分画像), 下段(左:シミュレーション画像, 右:背景差分シミュレーション画像)

3.4 MAP 計算

現時刻の状態 S_t を $B+N$ 回の繰り返し更新処理によって得られた状態 \hat{S} から MAP(maximum a posteriori) を持つ状態として計算する。初めの B 回の更新処理によって得られた状態 $\hat{S}_{t,i < B}$ は推定精度が悪いと考え、以下の式(9)によって計算する。

$$S_t = \frac{1}{N} \sum_{i=B}^{B+N} \hat{S}_{t,i} \quad (9)$$

4. 実験

4.1 実験システム

2台のカメラと PC (Athlon64x2 4800+) で構成されたシステムで実験を行った。図6は2台のカメラから撮影された画像の一例である。入退室エリアは2台のカメラの共通視野部分と単独視野部分の境界部分に設定した。状態推定の繰り返し回数 ($B+N$) を 200, MAP 計算には $B=100$ とし、最後の 100 個の状態を用いて行った。最適化されていない実装で処理速度は約 5fps であった。



図6: 2台のカメラから撮影した画像

4.2 複数人物追跡

本システムの基本的な性能を評価するために、2人の人物が時刻をずらして監視エリアへと入ってきて、2台のカメラ画像上で同時に完全に重なるようにしてすれ違うシーケンスに対して処理を行った。

図7に処理結果を示す。図中、上段が実画像シーケンスを示し、中段、下段はそれぞれ推定された状態をシミュレートしたシーケンスを実カメラと同じパラメータの仮想カメラ、仮想俯瞰カメラで撮影した画像を示している。まず画像左側より人物1が監視エリアに入ってきて、状態に追加される。次に少し遅れて人物2が監視エリアに入ってきて状態に追加された後にお互いがすれ違い、最

後に2人同時に監視エリアより退出する様子が安定して追跡できていることが確認できる。

図8はこの2人の人物の軌跡を俯瞰図としてグラフに表したものである。画像上では一時的に2人の人物が完全に重なってしまうが、3次元空間中では明確に分離されている。これは尤度計算の際の相互位置関係に基づくペナルティ関数が有効に働いているためである。

4.3 追跡位置精度

本システムの追跡位置精度の大きな評価を行うために、単独人物が予め決められたルートを歩いたシーケンスを処理し、本システムによって推定した軌跡と実際の軌跡の比較を行った。映像シーケンスと正解人物位置の時間的な同期が取れていないため、正解軌跡データからの推定人物位置の距離を平均することで、大まかな精度評価とすることとした。

図9に処理結果を示す。静止物体によって人物の大部分が画像上で隠れるようなシーンが存在するが、オクルージョンの予測により、安定した追跡ができていることが確認できる。

図10は本システムによって推定された軌跡と正解軌跡を俯瞰図としてグラフに表したものである。推定軌跡がほぼ正解軌跡と重なっていることが確認できる。推定軌跡の正解軌跡からの平均誤差は4.86cm、最大誤差は29.43cmであった。非常に大まかな評価ではあるが、本システムの追跡精度の高さが確認できる。

5. まとめと今後の予定

本稿では、3次元環境情報を用いることで、監視エリアに存在する人物数が動的に変化し、かつ人物同士が互いに接近、相互作用するような状況においてもオクルージョン問題を排除し、適切に追跡する手法を提案した。今後はより多人数の追跡実験とより詳細な評価を行う予定である。また人物の位置だけでなく、その姿勢も同時に推定できるシステムへと発展させていきたい。

参考文献

- [1] Z. Tao and N. Ram: "Tracking multiple humans in crowded environment", Proc. IEEE Conf. Computer Vision and Pattern Recognition, Vol. 2, pp. 406-413 (2004).
- [2] Z. Khan, B. Tucker and D. Frank: "Mcmc-based particle filtering for tracking a variable number of interacting targets", IEEE Trans. Pattern Analysis and Machine Intelligence, **27**, 11, pp. 1805-1918 (2005).
- [3] 大澤, 宮川, 若林, 荒川: "ロングシーケンス映像を用いたシームレスで密な三次元形状復元", 画像の認識・理解シンポジウム (MIRU2005), pp. 1224-1231 (2005).
- [4] B. K. P. Horn: "Closed-form solution of absolute orientation using unit quaternions", Journal of the Optical Society of America A, **4**, pp. 629-642 (1987).
- [5] K. Okada, S. Kagami, M. Inaba and H. Inoue: "Plane segment finder: Algorithm implementation and applications", International Conference on Robotics and Automation (ICRA'01), pp. 2120-2125 (2001).
- [6] P. J. Green: "Trans-dimensional Markov chain Monte Carlo, Highly Structured Stochastic Systems", Oxford Univ. Press (2003).

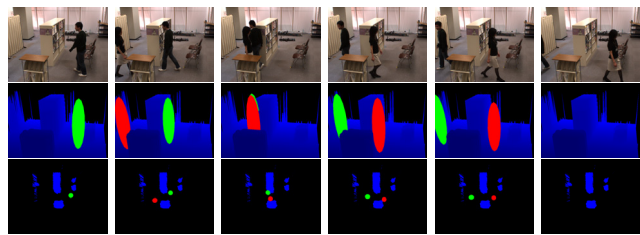


図7: 複数人物の追跡結果

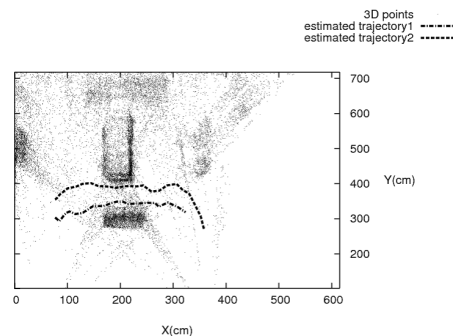


図8: 複数人物追跡実験の位置軌跡

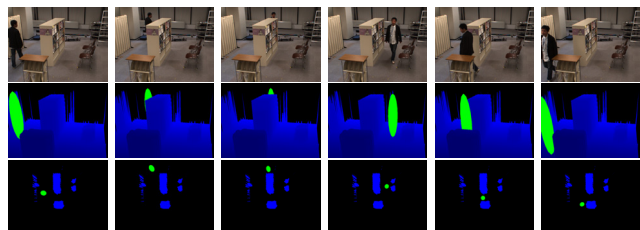


図9: 単独人物の追跡結果

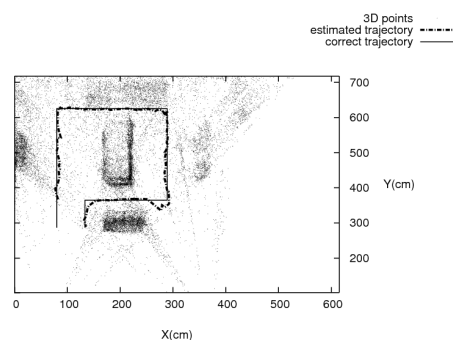


図10: 推定された位置の軌跡と正解軌跡