

逐次モンテカルロ動画像追跡： 速度パラメータ，検索範囲の自動調整による精度向上

中尾忠義

松井淳

中田洋平

松本隆

Tadayoshi Nakao

Atsushi Matsui

Yohei Nakada

Takashi Matsumoto

1 まえがき

動画像中のオブジェクト追跡は様々な応用範囲を持つ。セキュリティ面では防犯カメラ中の人物追跡に応用することができる。また走行中の車両追跡，車内からの歩行者追跡などによる事故の軽減も上げられる。オブジェクト追跡はその応用範囲の重要性から連続して得られるオブジェクトの位置を正確に追跡し続けることが求められる。オブジェクト追跡に対して逐次モンテカルロ法 [1] は非常に有効であることが様々な論文で示されている [2]。しかしながら，追跡対象の速度変化が激しい場合，正確に追跡し続けることが困難である場合が多い。

本稿では，逐次モンテカルロ法を用いたオブジェクト追跡において，追跡対象の速度変化と不確定性の変化を考慮したアルゴリズムを提案する。具体的には，運動モデルへの速度項の追加，および，運動モデルのノイズの分散（検索範囲）のオンライン学習則の導入によって，速度変化を伴う移動物体の追跡精度の改善を図る。屋内を歩行し往復する一人の人物を，固定カメラで撮影した動画像を用いた比較実験により，提案手法の有効性を検証する。

2 提案手法

2.1 観測モデル

時刻 t における追跡対象を囲う画像平面上の矩形領域を $\mathbf{X}_t = (x_t, y_t, w_t, h_t)^T$ で表す。ここで， x_t, y_t, w_t, h_t はそれぞれ矩形領域の中心座標，幅，高さを表す。

矩形領域の初期値 \mathbf{X}_0 が与えられたとき，第 1 フレームの入力画像の \mathbf{X}_0 内の RGB ヒストグラム $\mathbf{H}_0 = (h_0^R, h_0^G, h_0^B)$ を記録する。次に，時刻 $t > 0$ の入力画像 y_t が与えられたとき，矩形領域 \mathbf{X}_t 内の RGB ヒストグラム $\mathbf{H}(\mathbf{X}_t) = (h^R(\mathbf{X}_t), h^G(\mathbf{X}_t), h^B(\mathbf{X}_t))$ を計算する。このとき，矩形領域 \mathbf{X}_t の尤度を， \mathbf{H}_0 と $\mathbf{H}(\mathbf{X}_t)$ の RGB 成分ごとに計算した Bhattacharyya 距離 $D^c(h_0^c, h^c(\mathbf{X}_t))$ ($c \in \{R, G, B\}$) を用いて次式で定義する。

$$p(y_t | \mathbf{X}_t) \propto \exp\left(-\lambda \sum_{c \in \{R, G, B\}} (D^c(h_0^c, h^c(\mathbf{X}_t)))^2\right). \quad (1)$$

追跡対象の形状は必ずしも矩形で十分に近似できるとは限らないため，矩形領域の中心から離れた部分については背景画像が混入することが予想される。本稿では，各ヒストグラム $h^c(\mathbf{X}_t)$ を計算する際に，矩形領域と中心を同じとするガウス分布状の 2 次元窓関数を各画素値の頻度に乗じることによって，矩形領域の周辺部分に混入する背景画像の影響を抑制する [3][4]。

2.2 運動モデル

このようなオブジェクト追跡問題に逐次モンテカルロ法を適用する場合，矩形領域を定義するパラメータに対するダイナミクスが必要になる。最も一般的な方法としては，例えば正規ランダムウォークが仮定できる。

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \mathbf{D}\mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}\left(0, 1\frac{1}{\gamma}\right). \quad (2)$$

$\mathbf{D} = \text{diag}\{10, 10, 0, 0\}$ はスケーリング行列であり γ はノイズ項の分散の逆数である。 $\gamma = 1.0$ で固定する。

(2) 式は，分散固定の正規乱数を用いて直前の時刻の矩形領域の近傍を探索するため，追跡対象の位置や大きさが急激に変化する場合， γ^{-1} を非常に大きく設定する必要があり探索の効率が悪い。そこで本稿では，追跡対象の移動方向と移動量についての連続性を仮定し，2 フレーム前までの追跡結果から推定した速度成分の利用を考える。さらに，式 (2) で固定とした γ を時刻 t と共に変化するパラメータ γ_t と考え，オンライン学習する次式のようなダイナミクスを提案する。

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \alpha(\mathbf{X}_t - \mathbf{X}_{t-1}) + \mathbf{D}\mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}\left(0, 1\frac{1}{\gamma_t}\right) \quad (3)$$

$$\log \gamma_{t+1} = \log \gamma_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, 1). \quad (4)$$

ここで， α は速度項の慣性の強度をあらわすハイパーパラメータである。本稿では経験的に決定した固定値 ($\alpha = 0.5$) とする。

3 実験

3.1 実験条件

(2) 式の運動モデル（従来手法）と，(3),(4) 式の運動モデル（提案手法）を比較するため，移動速度が異なる 3 種類の評価データを用意した。それぞれの入力用画像は 1 人の人物が屋内を歩行し往復の様子を固定カメラで撮影した。正解データは，人物の胴体部（トルソー）を囲う矩形領域を手動で抽出した^{*1}。システムに与える矩形領域の初期値 \mathbf{H}_0 は，各々の評価データの第 1 フレームの正解データを用いた。それぞれの評価データの撮影時間と，追跡対象の画面上での移動速度を表 1 に示す。

3.2 実験結果

従来手法と提案手法の実験結果を累積推定誤差とサンプルの平均的中率の二つの指標を用いて評価した。ここで累積推定誤差は，逐次モンテカルロ法で抽出した \mathbf{X}_t の事後確率分布のサンプルの，正解データに対する中心座標の誤差を pixel 単位でフレームごとに積算した値を

^{*1} 正解データは本研究に精通していない 5 名が各々独立に抽出したデータを平均化したものである。

あらわす。また平均的中率は、中心位置が正解データの矩形領域内に存在するサンプルの全サンプルに対する割合をあらわす。ただし、平均的中率の計算において、正解データの矩形領域をその平均的な大きさに相当する半径 $r = 30\text{pixel}$ の円で近似した。

3種類の評価データについて、従来手法と提案手法それぞれの運動モデルを用いた場合の追跡結果の累積推定誤差を表2に、平均的中率を表3に示す。表2より、全ての評価データにおいて、従来手法に比べて提案手法での累積誤差が減少していることがわかる。特に、移動速度の最大値が比較的大きい評価データ(I)では、累積誤差の削減率が28.7%となり、提案手法による大幅な改善効果が得られた。また表3より、平均的中率が最大で約45%、全評価データの平均で約30%改善されていることから、従来手法に比べて提案手法ではより多くのサンプル点が追跡対象を捉えていることがわかる。提案手法による追跡結果の具体例を図1に示す。図1の矩形は評価データ(I)の正解データを、同図の十字のマークは追跡結果の中心位置のサンプルをそれぞれあらわす。図1より、推定位置が的確であることと、全てのフレーム数においてほとんどのサンプルが追跡対象を的確に捉えていることがわかる。

評価データ	撮影時間 [frame]	移動速度 [pixel/frame]		
		最大	最小	平均
(I)	250	53.15	0	2.46
(II)	250	17.80	0	2.93
(III)	180	16	0	3.38

表1 評価データ

評価データ	従来手法	提案手法	誤差削減率
(I)	1320.5	842.1	28.7%
(II)	1091.0	1014.5	7.0%
(III)	630.8	585.1	7.2%
平均	1014.1	813.9	19.7%

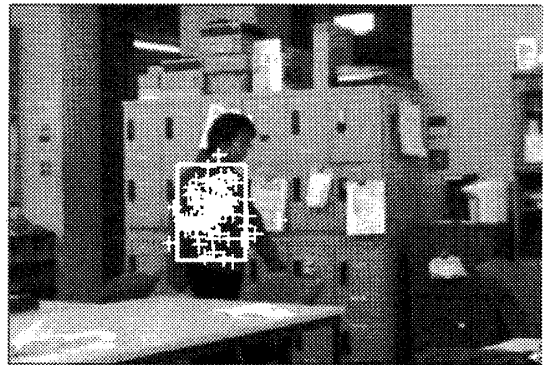
表2 追跡結果の累積推定誤差 [pixel]

評価データ	従来手法	提案手法	改善率
(I)	60.9	89.5	46.9%
(II)	70.1	89.4	27.5%
(III)	69.1	80.0	15.8%
平均	66.7	86.3	29.4%

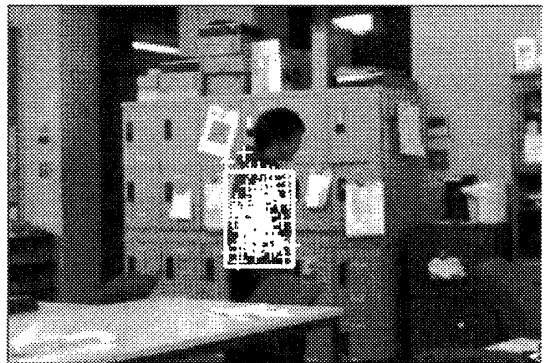
表3 追跡結果の平均的中率 [%]

4 まとめ

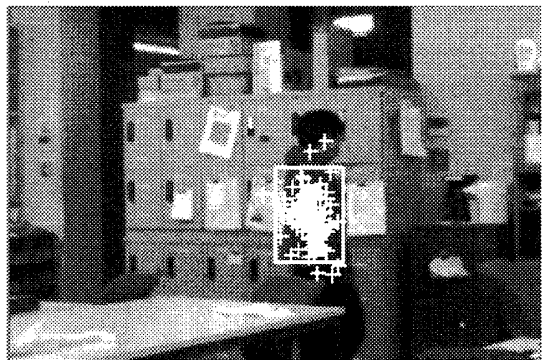
逐次モンテカルロ法を利用したオブジェクト追跡問題において、追跡対象の速度変化と不確定性の変化を考慮したオンライン学習アルゴリズムを提案した。移動速度が異なる評価データを用いて従来手法と提案手法を比較したところ、累積推定誤差、平均的中率ともに改善効果が見られた。なお、本稿では検索範囲を決めるハイパーパラメータの上限を設定しなかったが、追跡対象が停止している状態が長く続くとその検索範囲は徐々に小さくなっていく。停止した状態が長く続いた後に追跡対象



(a) 第125フレーム



(b) 第135フレーム



(c) 第145フレーム

図1 評価データ(I)の追跡対象(矩形)と追跡結果の中心位置(十字マーク)

が動き出したとき、迅速に捉えるために上限の設定や、検索範囲のリセット法などについてを今後の課題としたい。

参考文献

- [1] A. Doucet, S. Gotsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, Vol.10, No.3, pp.197-208 (2000).
- [2] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Tans. Signal Processing*, Vol.50, No.2, pp.174-188 (2002).
- [3] R. T. Collins and Y. Lui, "On-Line Selection of Discriminative Tracking Features," *Proc. of 9th International Conference on Computer Vision (ICCV 2003)*, Vol.2, pp.346-352 (2003).
- [4] H. T. Chen, T. L. Liu, and C. S. Fuh, "Probabilistic Tracking with Adaptive Feature Selection," *Proc. of 17th International conference on Pattern Recognition (ICPR 2004)*, Vol.2, pp.736-739 (2004).