

# パーティクルフィルタと単眼視点固定カメラを用いた 3次元空間上での物体領域推定

吉本 勇介\* 岡村 健史郎\*\* 松村 遼\*\*\*

## Estimating Areas of Moving Objects on 3D Space Using Particle Filter and Single Fixed-Viewpoint Camera

Yusuke YOSHIMOTO, Kenshiro OKAMURA and Ryo MATSUMURA

### Abstract

In recent years, demands for video surveillance system have increased. Usually, a multi-cameras system is used to have some 3D information in video surveillance system. But such a system needs much cost and time to be put in place. For a low-cost system, we prefer to use only one fixed-view point camera which is calibrated using the projective geometry which projects 3D points onto 2D image. But using only one camera, infinite points on 3D space correspond to a single point on 2D image. In this paper, we propose a high performance method for estimating areas of moving objects on 3D space using one fixed-viewpoint camera and particle filters. Using some constraints on 3D size and position of objects, our method restricts the generation of particles which represent the candidate area of the object. Therefore, the distribution of the particles shows the appropriate estimation of the object.

Key words: particle filter, video surveillance, fixed-viewpoint camera, tracking moving object

### 1. はじめに

近年、社会不安が増大し、犯罪抑止として監視システムへの期待が高まっている。監視システムにおける物体検出と追跡技術は重要であり、関連する研究が多く報告されている。その中でもパーティクルフィルタを用いた物体追跡に関する研究が注目されている。パーティクルフィルタを用いた物体追跡では、パーティクルによって多数の仮説を保持できるため、頑健な物体追跡が可能であるという特徴がある<sup>[1],[2]</sup>。

松元ら<sup>[3]</sup>は複数のカメラ画像を用いて、実世界を表す3次元空間上での物体追跡を行っている。3次元空間における物体追跡は、2次元画像上での物体追跡に比べ、より頑健な追跡が可能であり、物体がカメラの視線上に重なることで発生するオクルージョンに対しても対応が可能である。しかし、この手法はカメラを複数用いるためカメラ間の同期が必要で、カメラの個数が増えるため導入コストが高くなる。また、実際の場面では、オクルージョンを防ぐ位置にカメラを設置することが困難なことが多く、単体の視点固定型カメラを用いて3次元空間上で頑健な物体追跡が可能なシステムの構築が望まれている。

単体の視点固定型カメラを用いて3次元空間上で物体追跡を行うと、観測される画像上の一つの2次元座標値に対して複数の3次元座標値が対応する一対多の関係が成り立つ。すなわち、3次元空間において物体領域外にあるパーティクルも2次元画像上では物体領域内に写像され、物体が存在しない3次元領域にパーティクルが分布するようになる。そのため、推定される物体の存在領域が広がり、隣接して進入する物体を精度よく追跡できない。この問題に対し、大澤ら<sup>[4]</sup>は人物を楕円体によりモデル化することで、高い精度での物体追跡に成功しているが、実世界では、様々な形状の物体が存在するため、それら全てに対応したモデルを作成することは難しい。

そこで本稿では、3次元空間上で真の物体位置にあるパーティクルはその位置を物体内部の範囲内であれば高さ方向に変えても2次元画像上では物体領域に投影される点に着目し、モデルの作成を行わずに物体存在位置を精度よく推定する手法を提案する。具体的には、各パーティクルの3次元座標値を高さ方向に変化させ、これらを2次元画像上に順次投影して尤度和を求め、3次元上で物体領域に含まれるパーティクルの尤度が高くなることを利用し、パーティクルを選択するようにした。これにより、単体の視点固定型カメラで撮影した動画像においても物体が存在する領域にパーティクルを集中させ、モデルを作ることなく隣接して進入した移動物体を検知、追跡することが可能となった。提案手法の有効性を、実画像を用いた実験により示す。

## 2. パーティクルフィルタを用いた物体追跡

パーティクルフィルタとは多数の粒子を用いて、前時刻の状態からの予測と現在の観測情報から現在の状態を推定する手法である。パーティクルフィルタでは予測、観測、リサンプリングの繰り返しによって物体追跡を行う。パーティクルフィルタアルゴリズムの概念図を図1に示す。図中の丸はパーティクルを表し、丸の大きさはそこに物体が存在する確からしさを表す。この確からしさを尤度と呼ぶ。図1のように、尤度が低いパーティクルは選択されずに消滅し、高いパーティクルは復元抽出により何度も選択され仲間を増やす。その結果、尤度の高いパーティクルが多く残る。

以下に3次元空間における物体追跡処理を詳しく説明する。ここで、時刻 $t$ における $i$ 番目のパーティクルを $x_t^{(i)} = (x_t^{(i)}, y_t^{(i)}, z_t^{(i)})$ と表し、そのパーティクルの尤度と重みをそれぞれ $\pi_t^{(i)}$ 、 $g_t^{(i)}$ と表す。また、 $x_t^{(i)}$ 、 $y_t^{(i)}$ 、 $z_t^{(i)}$ はパーティクルの3次元座標値を表し、パーティクルの数を $N$ とする。

「予測」

時刻 $t-1$ のパーティクルから時刻 $t$ におけるパーティクルの位置を予測する。本稿では、予測モデルとして式(1)を用いた。 $v_{t-1}^{(i)}$ は時刻 $t-1$ での各パーティクルの速度であり、これは時刻 $t-2$ から時刻 $t-1$ までの間にそのパーティクルがどのくらい移動したかを表す。また、 $w_t^{(i)} \sim N(0, \sigma^2)$ は3次元のガウシアンノイズを表す。

$$x_t^{(i)} = x_{t-1}^{(i)} + v_{t-1}^{(i)} + w_t^{(i)} \quad (1)$$

$$v_{t-1}^{(i)} = x_{t-1}^{(i)} - x_{t-2}^{(i)} \quad (2)$$

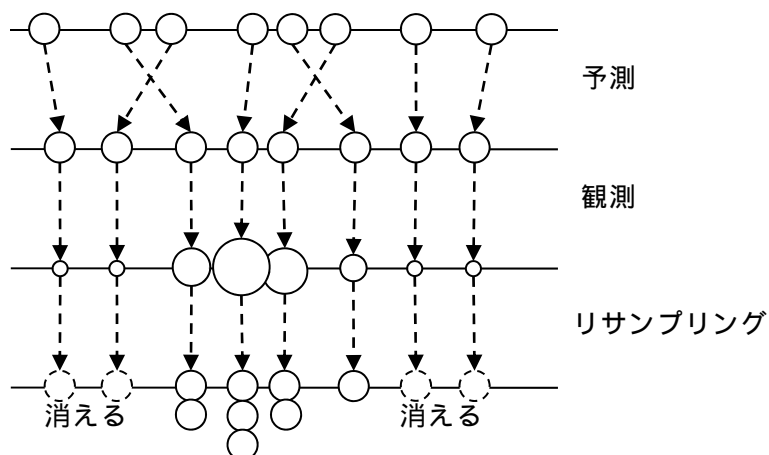


図1 パーティクルフィルタアルゴリズムの概念図

「観測」

予測された位置における各パーティクルの尤度 $\pi_t^{(i)}$ および重み $g_t^{(i)}$ を算出する。尤度 $\pi_t^{(i)}$ はパーティクルを2次元画像上に投影した点における背景差分値などにより求める。重み $g_t^{(i)}$ は全てのパーティクルの尤度合計で各パーティクルの尤度 $\pi_t^{(i)}$ を正規化したものとする。詳しくは3.1で述べる。

「リサンプリング」

観測により求めた各パーティクルの重み $g_t^{(i)}$ に比例した確率で復元抽出によりパーティクルを選択し、

時刻 $t$ でのパーティクル集合を求める。この処理により、物体が存在すると思われる領域にパーティクルが集中する。また物体中心位置はパーティクルの平均座標とし、物体の大きさは物体中心位置からの各パーティクルの分散より求める。

物体検出および物体追跡処理は、文献[5],[6]で提案されたアルゴリズムを用いる。まずパーティクルを予め定めた領域に一様に散布する。このパーティクルフィルタは観測のみを行い、物体の進入に備える。この状態にあるフィルタを検出フィルタと呼ぶ。この検出フィルタが観測のみを行う理由は、予測やリサンプリングの処理によってパーティクルが特定の領域に集中してしまうことや、検出領域外に出てしまうことで進入物体を検出できない状態になることを防ぐためである。検出フィルタは観測により物体が進入したと判断すると予測とリサンプリングの処理を開始する。この状態にあるフィルタを追跡フィルタと呼び、パーティクルフィルタは検出した物体の追跡を行う。また、これまでの説明ではパーティクルの集合は一つであることを前提としていたが、集合の個数は物体の進入・退出により変化する。

### 3. 世界座標系での物体の検出と追跡

#### 3.1 世界座標と画像座標の関係

3次元空間上で物体追跡を行う場合、2次元画像平面である画像座標と3次元空間である世界座標の対応関係をカメラキャリブレーションにより事前に求める必要がある。世界座標系の点を $P(x, y, z)$ 、それに対応する画像座標系の点を $p(u, v)$ とすると、 $P$ と $p$ の関係はそれぞれの同時座標表現 $P = [x, y, z, 1]^t$ と $p = [u, v, 1]^t$ を用いて以下の式で表現できる[2],[7]。

$$sp = A[R|T]P \quad (3)$$

ここで、 $A$ はカメラの内部パラメータ行列であり、 $[R|T]$ はカメラの外部パラメータ行列、 $s$ は画像のスケール係数である。内部パラメータ行列は焦点距離 $f$ 、 $u$ および $v$ 方向の画素サイズ( $s_u, s_v$ )、画像平面とカメラ光軸の交点である主点座標( $u_0, v_0$ )、画像平面上の $u$ および $v$ 軸の傾きを表すせん断係数 $k_s$ を用いて以下の式で表現される。

$$A = \begin{bmatrix} fs_u & fk_s & u_0 \\ 0 & fs_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

また、外部パラメータ行列はカメラの回転行列 $R$ とカメラの平行移動行列 $T$ で構成され、以下の式で表現される。

$$R = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \quad (5)$$

$$T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (6)$$

カメラの内部パラメータ $A$ 及び外部パラメータ $[R|T]$ は、Intel社の画像処理向けライブラリであるOpenCVライブラリを用いて求めた[7]。

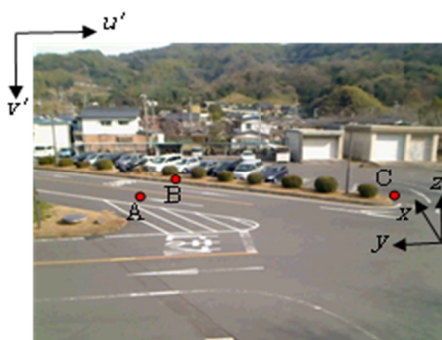


図2 カメラパラメータの検証に用いた点

表 1 カメラパラメータの検証結果

検証場所	3次元座標 (x, y, z)[cm]	画像より調べた 2次元座標 (u', v')[pixel]	式(3)で推定した 2次元座標 (u, v)[pixel]	誤差 (u' - u, v' - v) [pixel]
A	(823, 2580, 0)	(177, 248)	(175, 250)	(+2, -2)
B	(1123, 2387, 50)	(224, 222)	(225, 224)	(-1, -2)
C	(338, 108, 100)	(562, 250)	(566, 250)	(-4, 0)

図 2 に示す 3 点と式(3)を用い推定したパラメータの検証を行った。検証は 3 点の 3 次元座標(x, y, z)を実世界にて測定し、式(3)で求めた投影点の画像座標(u, v)と実際の画像座標(u', v')の比較により行った。検証結果を表 1 に示す。図 2 に示した 3 点における誤差は最大で 4pixel 程度と許容可能な範囲内の誤差であるといえる。

図 3 に実世界上の点が 2 次元画像平面へ投影される様子を示す<sup>[7]</sup>。3 次元空間での物体追跡では、まずパーティクルが存在する世界座標上の点P(x, y, z)を外部パラメータ行列[R|T]によりカメラ座標上の点P<sub>c</sub>(x<sub>c</sub>, y<sub>c</sub>, z<sub>c</sub>)に変換し、その座標を内部パラメータAにより画像座標に投影することで画像座標上の点p(u, v)を得る。そして求めた画像座標p(u, v)における尤度π<sub>t</sub><sup>(i)</sup>を求める。尤度π<sub>t</sub><sup>(i)</sup>は背景差分法を用いた場合、以下の式となる。

$$\pi_t^{(i)} = \sqrt{(R_I(u, v) - R_B(u, v))^2 + (G_I(u, v) - G_B(u, v))^2 + (B_I(u, v) - B_B(u, v))^2} \quad (7)$$

ここで R<sub>I</sub>(u, v)、G<sub>I</sub>(u, v)、B<sub>I</sub>(u, v) は入力画像上の点 (u, v) における画素値を表しており、R<sub>B</sub>(u, v)、G<sub>B</sub>(u, v)、B<sub>B</sub>(u, v) は背景画像上の点 (u, v) における画素値を表す。また、式(8)を用いて尤度π<sub>t</sub><sup>(i)</sup>の正規化を行い、各パーティクルの重みg<sub>t</sub><sup>(i)</sup>を求める。

$$g_t^{(i)} = \frac{\pi_t^{(i)}}{\sum_{i=1}^N \pi_t^{(i)}} \quad (8)$$

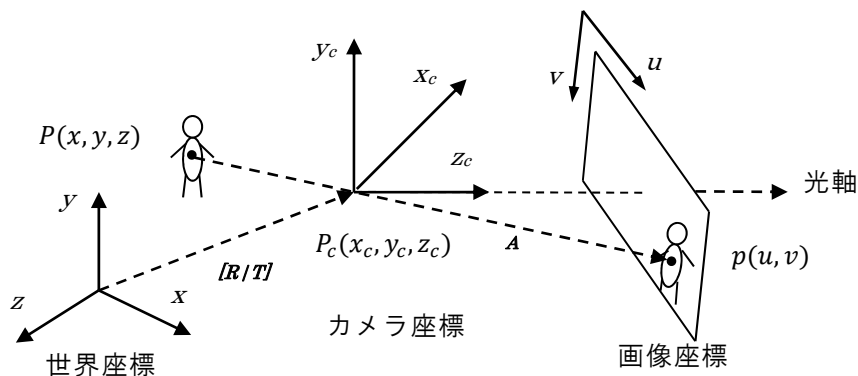


図 3 実世界上の点の画像座標への投影の様子

### 3.2 世界座標系での物体検出と追跡

パーティクルフィルタを用いて複数物体を追跡する場合、一つの物体に対し一つのフィルタを割り当てて物体追跡を行う。そのため、物体の進入検出は精度よく行う必要がある。そこで、文献[5],[6]に示された以下のような世界座標系での物体検出および追跡アルゴリズムを用いた。

まず、物体が進入すると思われる領域を検出領域と定義し、物体検出用にパーティクルを一様に散布する。散布したパーティクルにおいてある値以上の尤度を持つパーティクルの割合が閾値α以上になると、物体が進入したと判断し追跡フィルタとして物体追跡を開始する。同時に、次の進入物体に備えて新たに検出フィルタを設置する。

物体領域はリサンプリング後の各パーティクルの存在位置により推定を行い、図 4 に示すような直方体領域とした。物体中心位置を  $o = (x_o, y_o, z_o)$  と表し、式(9)により求める。また、物体領域の x 軸方向の大きさは推定された物体中心位置 o からの各パーティクルの偏差  $\sigma_x$  とし、式(10)より求める。y 軸および z

軸方向の物体領域の大きさも同様に求める。

以上の手法を用いることで、同時に複数の進入物体を個々に検出する。また、追跡フィルタは尤度を持つパーティクルの割合が閾値 $\beta$ 以下になると、物体が監視領域外に退出したとして追跡を終了する。

$$O = \frac{1}{N} \sum_{i=1}^N x_t^{(i)} \quad (9)$$

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_o - x_t^{(i)})^2 \quad (10)$$

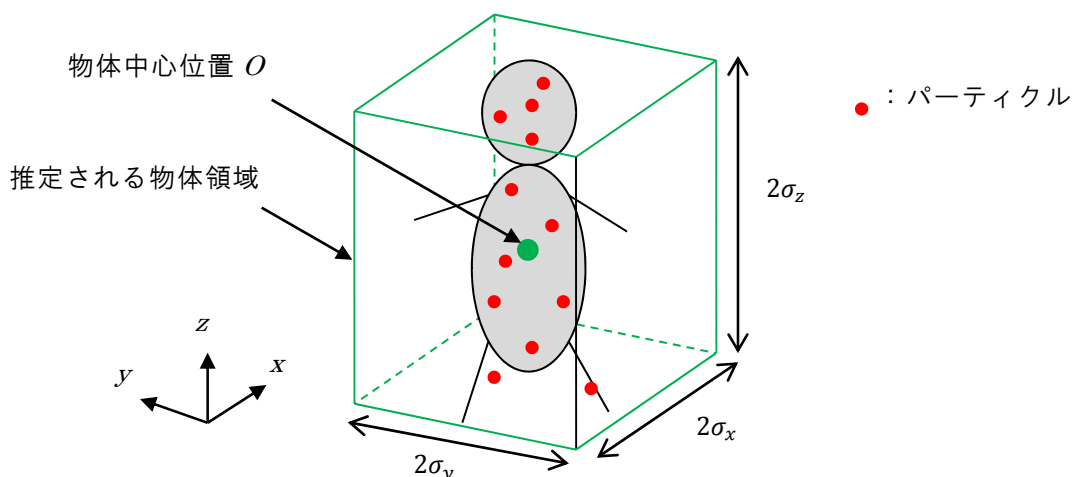


図4 推定される物体領域

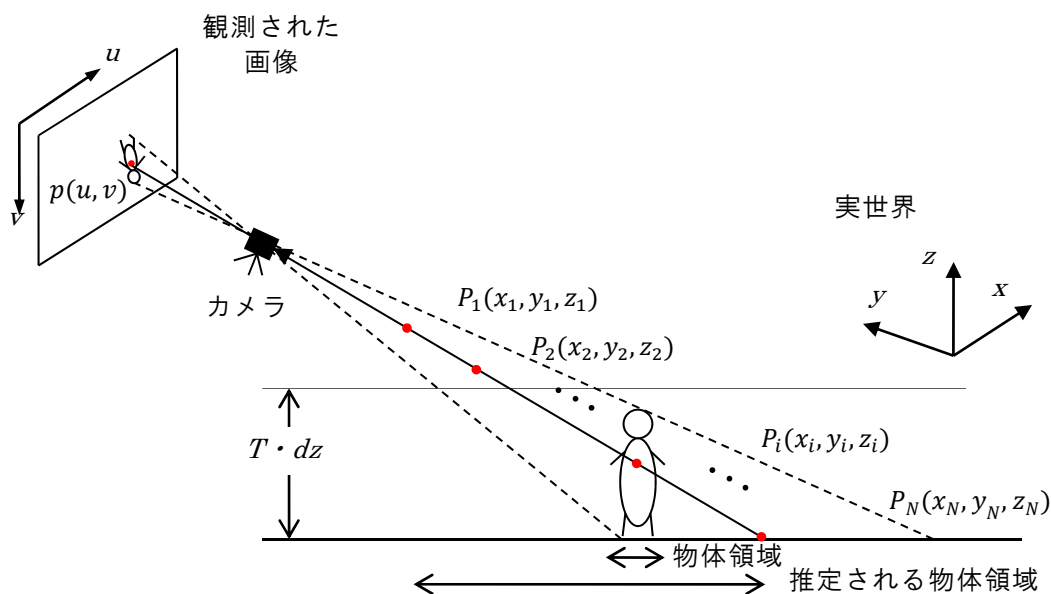


図5 パーティクルの投影の様子

#### 4. 物体の位置と大きさの推定

##### 4.1 単体の視点固定型カメラの問題点

本研究は単体の視点固定型カメラから得られる画像をもとに、3次元上で精度よく進入物体を検出し、追跡するシステムの構築を目的としている。しかし、図5に示すように単体の視点固定型カメラによる3次元空間上での物体追跡では、一つの画像座標 $p$ に対し複数の世界座標 $P_1 \sim P_N$ が対応する一対多の関係

が成り立つ。図 5 における二本の破線で囲んだ領域は 2 次元画像上で物体領域内に投影される 3 次元領域を表す。すなわち、実際には物体領域内に存在しない  $P_1$  や  $P_2$  にあるパーティクルが画像上では物体領域内に写像されることで大きな値となり選択される。その結果、推定される物体の存在領域が実際よりも大きなものとなり、隣接して進入する物体を精度よく追跡できない。そこで、以下のような方法で 3 次元上の物体領域にないパーティクルを削除することで物体の存在位置を精度よく推定する手法を提案する。

#### 4.2 物体領域の推定

本手法では、物体は 3 次元空間において基準となる水平面上に立ち、高さ方向に連続して存在し、その高さは一定値以下であると仮定した。仮定した環境を図 6 に示す。図 6 に示すように人物の場合、足元が高さ  $z = 0$  であり、 $z$  軸方向に沿って物体が存在する。この仮定のもと、3 次元空間上で真の物体位置にあるパーティクルは、その位置を高さ方向に変えても 2 次元画像上で物体領域内に投影されることに着目し、次のような物体領域推定方法を考える。

本追跡フィルタに対する観測において、 $i$  番目パーティクルの 3 次元座標値を 0 から高さ方向に  $dz$  ずつ変化させ、その座標を 2 次元画像に投影した点においてしきい値  $\gamma$  以上の背景差分値を持つ点の個数を尤度  $\pi_t^{(i)}$  とする。ここで、尤度  $\pi_t^{(i)}$  を差分値の合計ではなく座標点の個数としたのは、他と比べ非常に大きな差分値の影響を避けるためである。この尤度  $\pi_t^{(i)}$  は  $i$  番目のパーティクルの状態を  $x_i = (x_i, y_i, z_i)$  とした場合、以下の式により求める。

$$\pi_t^{(i)} = \sum_{k=0}^T \delta(x_i, y_i, k \cdot dz) \quad (11)$$

$$\delta(x_i, y_i, k \cdot dz) = \begin{cases} 1 & \text{if } F(x_i, y_i, k \cdot dz) > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

ここで、 $F(x_i, y_i, k \cdot dz)$  は 3 次元空間上で高さ方向を変化させた点  $(x_i, y_i, k \cdot dz)$  を 2 次元画像上に投影した点  $(u, v)$  における背景差分値を表し、以下の式により求める。ただし、背景の微細な明度変化の影響を除くために閾値  $\gamma$  以下の値は 0 とする。

$$F(x_i, y_i, k \cdot dz) = \sqrt{(R_I(u, v) - R_B(u, v))^2 + (G_I(u, v) - G_B(u, v))^2 + (B_I(u, v) - B_B(u, v))^2} \quad (13)$$

ここで  $R_I(u, v)$ 、 $G_I(u, v)$ 、 $B_I(u, v)$  は入力画像上の点  $(u, v)$  における輝度値を表しており、 $R_B(u, v)$ 、 $G_B(u, v)$ 、 $B_B(u, v)$  は背景画像上の点  $(u, v)$  における輝度値を表す。式(11)及び式(12)が意味するところは以下の通りである。3 次元上の点  $(x_i, y_i, z_i)$  に対し、このパーティクルが物体内に存在する場合は、高さ方向の座標  $z_i$  を 0、 $dz$ 、 $2 \cdot dz$ 、 $\dots$ 、 $T \cdot dz$  に変化させてもそれらの点の多くはその物体内に存在

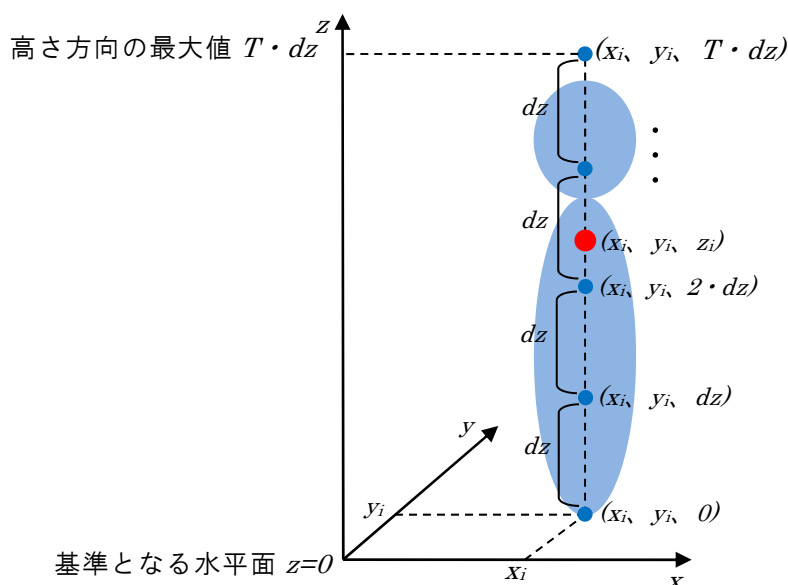


図 6 本手法における高さ方向を変化させる模式図

し、画像上では物体領域内に投影される。一方、3次元上において物体内に存在しない場合は、高さ方向に座標を変化させると2次元画像において物体領域外に投影される。例えば、図5の点 $P_N$ は高さ方向に $T \cdot dz$ まで移動すると破線で囲まれる領域から出てしまう。また、図5の $P_1$ や $P_2$ に対応する $0, dz, 2 \cdot dz$ といった点も、破線囲まれる領域から出てしまう。そのため、これらの点に存在するパーティクルの尤度は低いものとなり選択されにくくなる。一方、 $P_i$ のように真の物体領域内に存在するパーティクルは座標 $z_i$ を $0, dz, 2 \cdot dz, \dots, T \cdot dz$ と変化させても、ほとんどの点が破線内の領域に含まれ、これらは物体領域に投影されるため、高い尤度を持つようになる。このような理由から、3次元上の物体外 $P_1, P_2, P_N$ にあるパーティクルはサンプリングにより消滅し、物体上にある $P_i$ だけが残るようになる。

## 5. 実験

### 5.1 実験環境

単体の視点固定型カメラを用いて取得した動画像に対し、本手法を適用して物体の検出および追跡実験を行い、物体形状を考慮しない従来手法との比較を行った。観測対象を図7に示す。追跡対象は人物、自転車、自動車である。以下に実験環境の詳細を示す。

実験に使用したコンピュータは、OSがWindows 7 Professional、CPUがIntel Core-i7の2.80[GHz]である。撮影に使用したカメラはLogicool Qcam Orbit AFであり、動画像は $640 \times 480$ [pixel]の解像度、毎秒15フレームの設定で約10分間撮影したものである。観測対象の特徴として、多くの人物が冬用制服を着用しており色調が類似していること、人物が集団で行動している場面が多いこと、日が高い日中に撮影を行ったため太陽光の反射が強く自動車等の上部の輝度が他に比べ高いことが挙げられる。実世界を表す3次元座標軸と進入物体の検出を行う領域を図7に示す。 $y$ 軸を画像中央付近の橙色の線と並行にとり、 $z$ 軸は地面に対し垂直方向に、 $x$ 軸は二つの軸と垂直になるよう手前から奥方向にとった。実験に用いたパラメータを表2に示す。物体が進入したとして追跡を開始する閾値 $\alpha$ と物体が退出したとして追跡を終了する閾値 $\beta$ を、パーティクル数 $N$ を基準にそれぞれ2.5[%]と0.1[%]とし、観測における背景の明度変化によるノイズ除去のための閾値 $\gamma$ を15とした。また、提案手法における高さ方向の変化量 $dz$ を50[cm]、高さ方向の最大値 $T \cdot dz$ は200[cm]とし、検出領域の大きさは表2に示す通りである。

表2 実験パラメータ

一つフィルタに対するパーティクル数 $N$	2000
式(1)における正規乱数の標準偏差 $\sigma$ [cm]	15
物体追跡を開始する閾値 $\alpha$ [%]	2.5
物体追跡を終了する閾値 $\beta$ [%]	0.1
観測におけるノイズ除去のための閾値 $\gamma$	15
高さ方向の変化量 $dz$ [cm]	50
高さ方向の最大値 $T \cdot dz$ [cm]	200
検出領域(左) $\Delta x$ [cm] $\times$ $\Delta y$ [cm] $\times$ $\Delta z$ [cm]	$3000 \times 15 \times 200$
検出領域(右) $\Delta x$ [cm] $\times$ $\Delta y$ [cm] $\times$ $\Delta z$ [cm]	$4200 \times 15 \times 200$

検出フィルタを設置する領域

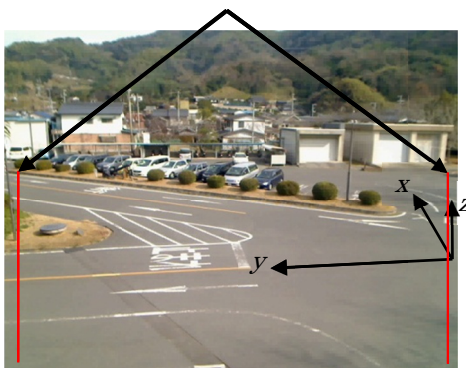
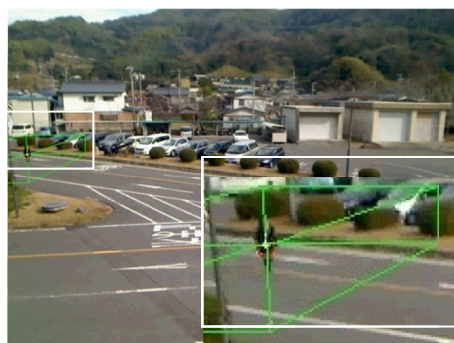
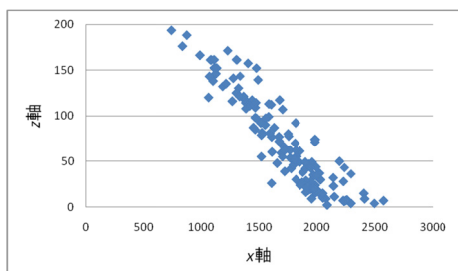


図7 観測対象

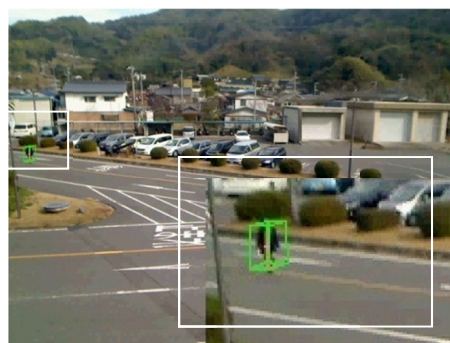


(a) 推定された物体領域

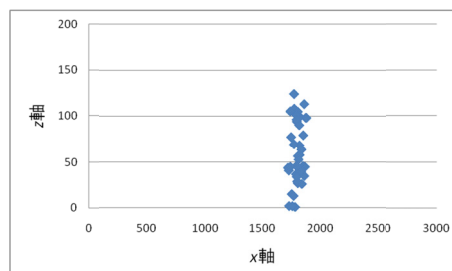


(b) x-z 平面でのパーティクル位置

図 8 従来手法



(a) 推定された物体領域



(b) x-z 平面でのパーティクル位置

図 9 提案手法

### 5.2 パーティクルの分布状況に関する考察

本手法の有効性を示すため、本手法と物体形状を考慮しない従来手法を用いて物体追跡を行い、3次元上におけるパーティクルの分布状況について調査した。図8に従来手法を用いた結果を示し、図9に本手法を用いた結果を示す。両図における(a)は推定された物体領域を画像上に直方形で表現した画像を示し、(b)はパーティクルの3次元座標上の点をx-z平面に投影した様子を示す。両図(a)において、画像の右下に物体領域の拡大画像を示し、推定領域を分かりやすくするためパーティクルは表示していない。両図(a)を比較すると、従来手法を用いた場合では実際の物体領域よりも非常に大きな領域が物体領域として推定されているのに対し、本手法を用いた場合ではより実際の物体領域に近い領域が推定されていることがわかる。これは従来手法の尤度算出が単純な背景差分法によるものであるため、カメラにより投影された2次元画像上で物体領域内に位置する全てのパーティクルが尤度を持ち、図8(b)に示すように、真に物体領域内に存在しないパーティクルも選択されるためである。一方、本手法ではパーティクルの3次元座標を高さ方向に変化させることで、図9(b)に示すように、3次元空間において真に物体領域内に存在するパーティクルのみが選択されていることがわかる。



(a) 従来手法



(b) 提案手法

図 10 推定された物体領域(自動車)



表 3 推定された物体領域の大きさ(自動車)

	$2\sigma_x$ [cm]	$2\sigma_y$ [cm]	$2\sigma_z$ [cm]
従来手法	226	88	74
提案手法	138	93	80

### 5.3 物体の大きさに関する考察

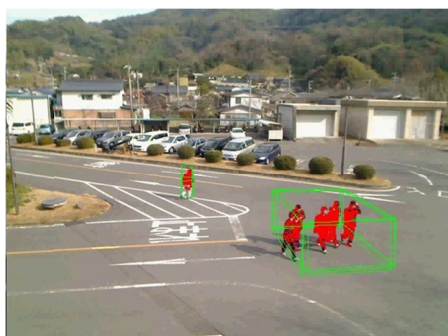
本手法の有効性を示すため、本手法と物体形状を考慮しない従来手法を用いて物体追跡を行い、3次元空間におけるパーティクルの分布状況について調査した。

#### 5.3.1 自動車の大きさに関する考察

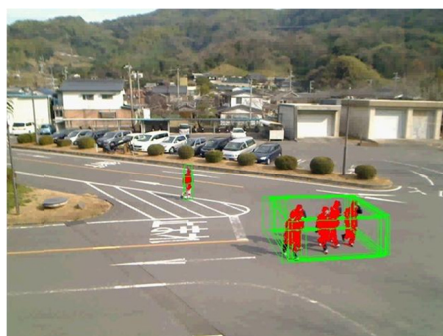
追跡中の自動車に対し推定された物体領域を図 10 に示し、推定された物体領域の大きさを表 3 に示す。図 10 における(a)は従来手法を用いた結果を示し、(b)は本手法を用いた結果を示す。図 10 の両図を比較すると、従来手法を用いた場合では画像奥行き方向に物体領域が広がっているのに対し、提案手法を用いた場合では物体の領域に沿った領域が推定されていることがわかる。また、表 3 における $2\sigma_x$ の値を比較すると、従来手法では 226[cm]であるのに対し、提案手法では 138[cm]となり、従来手法を用いた場合の方が 88[cm]も大きな領域を推定していることがわかる。

#### 5.3.2 複数人物の大きさに関する考察

追跡中の複数の人物に対し推定された物体領域を図 11 に示し、推定された物体領域の大きさを表 4 に示す。図 11 における(a)は従来手法を用いた結果を示し、(b)は本手法を用いた結果を示す。図 11 の両図を比較すると、従来手法を用いた場合では物体の存在領域よりも大きな領域が推定されているのに対し、提案手法を用いた場合では物体領域を精度よく推定されていることがわかる。また、表 4 における $2\sigma_x$ の値を比較すると、従来手法では 201[cm]であるのに対し、提案手法では 133[cm]と、従来手法を用いた場合の方が 68[cm]も大きな領域を推定していることがわかる。



(a) 従来手法



(b) 提案手法

図 11 推定された物体領域(複数人物)

表 4 推定された物体領域の大きさ(複数人物)

	$2\sigma_x$ [cm]	$2\sigma_y$ [cm]	$2\sigma_z$ [cm]
従来手法	201	31	53
提案手法	133	21	59

表 5 進入物体の追跡成功率

	従来手法	提案手法
進入した物体数[体]	69	69
追跡に成功した物体数[体]	51	59
追跡成功率[%]	73.9	85.5



図 12 追跡成功例

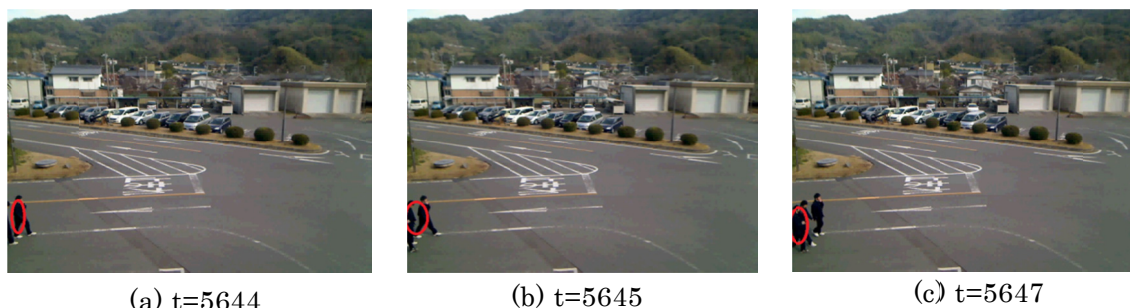


図 13 追跡失敗例

#### 5.4 追跡結果

提案手法の基本的な追跡性能を確かめるために、監視領域に進入する物体の追跡実験を行った。追跡結果を表 5 に示す。実験に用いた動画像全体の進入物体の数は 69 体で、従来手法を用いた場合には 51 体、提案手法を用いた場合には 59 体の検出および追跡に成功した。また、物体の追跡成功率は従来手法が 73.9%であるのに対し、提案手法では 85.5%と、従来手法と比較して 11.6%精度が向上した。

進入物体の追跡に成功した例を図 12 に示す。図 12 における赤い丸はパーティクルを表し、動画像の先頭からのフレーム数を時刻  $t$  とした。図 12 より、まず画面左上から単独で進入してきた人物の追跡を開始し、その後、隣接して進入する複数の人物を個々に検出し、追跡できていることがわかる。

複数の進入物体の追跡に失敗した例を図 13 に示す。図 13 における楕円はパーティクルフィルタの領域を表し、動画像の先頭からのフレーム数を時刻  $t$  とした。図 13(a) では画面左から進入した人物を検出することに成功しているが、次のフレーム ( $t=5645$ ) では続いて進入した人物を新たに検出せず、最初に進入した物体の一部として判定し、二人を一つの物体として扱っている。更に、(c) において進入した人物もこれらの一部として扱い、一人の人物にパーティクルが集中した結果、最初に進入した人物が追跡対象領域から外れた。これは式(1)におけるガウシアンノイズの分散  $w_t^{(i)}$  に対し、それよりも近接して物体が進入したためパーティクルが二つの物体領域に重複して分布してしまったことが原因と考えられる。しかし、分散  $w_t^{(i)}$  を小さくすると追跡中の物体を見失う恐れがある。そのため、このような進入に対しては一度、一つの物体として追跡を行い、その大きさや形状から複数物体として判定することが考えられる。

別の失敗例としては、非常に速い速度で進入する移動物体を検出できなかった。これはフレーム間隔が長すぎたため人物の領域が検出領域をすり抜け、移動物体の大部分が検出領域に存在しなかったことが理由である。今後、移動物体の移動速度、大きさ、および検出領域の大きさについて再検討が必要である。

## 6. まとめ

本研究では、単体の視点固定型カメラを用いた 3 次元上の物体追跡において、物体が基準となる水平面上に立ち、物体領域が垂直方向に連続的に存在するという仮定のもと、パーティクルフィルタの枠組みの中の「観測」において、パーティクルの 3 次元座標を高さ方向に変化させ尤度を算出する方法を提案した。この結果、本手法を用いた実験により、単体の視点固定型カメラでの 3 次元空間における物体の検出および追跡に対し有効であること、また推定される物体領域がより実際の物体領域に近いもので

あることを示した。

今後は、検出領域の大きさやフレーム間隔の再検討を行い、さらに自転車や幼児などの高さ方向に連続的に存在する領域が狭い物体に対し精度よく追跡する手法の検討、オクルージョンを伴って進入した複数の物体が追跡中に別々に行動するようになった場合に個々に追跡する手法の検討を行う必要がある。

## 謝辞

本研究は科研費(23510210)の助成を受けたものである。

## 参考文献

- [1] Isard, M. and Blake, A., "CONDENSATION Conditional Density Propagation for Visual Tracking," International Journal of Computer Vision, Vol.29, NO.1, PP.5-29(1998).
- [2] 加藤丈和, "パーティクルフィルタとその実装法", CVIM チュートリアルシリーズ, コンピュータビジョン 最先端ガイド 1, 八木康史, 斎藤英雄(編), アドコムメディア, 第3章, pp.75-102, 2009.
- [3] 松元郁佑, 加藤丈和, 和田俊和, 上田博唯, "複数カメラを用いた CONDENSATION による複数人物頭部の実時間検出・追跡", 電子情報通信学会技術研究報告, PRMU2005-277, Vol.105, No.674, pp.121-128, 2006.
- [4] 大澤達哉, 数藤恭子, 新井啓之, 小池秀樹, "単眼動画像を用いた近接する複数対象の三次元追跡", 電子情報通信学会技術研究報告, PRMU2008-38, Vol.108, No.94, pp.109-114, 2008.
- [5] 村北卓也, 石黒浩, "Sampled-Volume-based Data Association Filter による遮へい口バスタな実時間複数人物追跡", 電子情報通信学会論文誌 D, Vol.J90-D, No.10, pp.2868-2883, 2007.
- [6] 松村遼, 岡村健史郎, "パーティクルフィルタを用いた物体の進入検出", 電気・情報関連学会中国支部, 第59回連合大会講演論文集, pp.419-420, 2008.
- [7] 奈良先端科学技術大学院大学 OpenCV プログラミングブック制作チーム, "OpenCV プログラミングブック", 毎日コミュニケーションズ, 2007.

