

サーマルカメラとパーティクルフィルタを用いた 3次元空間上での物体追跡と大きさ推定

岡村健史郎* 西郷佳和** 河合宏紀*

Object Tracking and Size Estimation in 3D Space with Thermal Camera and Particle Filter

Kenshiro OKAMURA, Yoshikazu SAIGO and Hironori KAWAI

Abstract

A coastal area is exposed to a great number of threats such as smuggling, poaching, stealing fishing implements and some other crimes. These crimes are mostly committed after dark and in places where there are no lights. A thermal camera can detect heat energy and it has higher detection accuracy in complete darkness. Until recently, the cost of using the thermal camera has been prohibitive for usual practical surveillance. But new materials and other improvements make its price more reasonable and provide good tools for all-day surveillance. Using the thermal camera, we show some methods that provide the automatic video surveillance system with the ability to detect moving objects and determine the 3D positions and the sizes of these objects.

Key words: Video Surveillance, Particle Filter, Thermal Camera, Object Tracking, Size Estimation

1. まえがき

近年、湾岸部などでは漁具の盗難や密猟などの犯罪が増加傾向にあり、これらの犯罪防止を目的とした物体検出・追跡ができる監視システムへの期待が高まっている。これらの監視システムにおいては、省力化や犯罪調査立証のため、カメラを用いて対象領域を長時間記録し、物体が進入してきた場合には、自動的に物体検出・追跡ができることが求められている。これまでにも、カメラを用いて湾岸部を対象に監視するシステムが報告されているが⁽¹⁾、これらは可視光領域を扱うカメラを利用したもので、夜間における観測には照明を必要とするため、安価な自動監視は実現が難しかった。

夜間に利用できるカメラとして赤外線サーマルカメラがある。検出波長が8~14 μm の赤外線サーマルカメラは、自然放射されている遠赤外線を検出できることから、赤外線投影機などの外部光源が不要で、自動車等のライトや照明の陰影などにも影響されないといった利点がある。従来、赤外線サーマルカメラは非常に高価で、工業製品の検査や軍用の監視システムに用いられることはあっても、一般利用を想定した夜間監視の研究にはあまり利用されていなかった。しかし、近年、非冷却型マイクロボロメータを検出素子とするサーマルカメラの性能が向上し、低価格で、耐久性にも優れ、監視カメラとして注目されている。

一方、湾岸領域を対象に物体の検出を行おうとすると、波浪により海上に浮かんだ船や海面が常に変化するため、これらが誤って進入物体として検出されることがある。これらの誤検出を防ぐためには複

数の進入物体を仮定しながら追跡する必要がある。このような複雑な背景のもとで追跡を行う手法としてパーティクルフィルタがある⁽²⁾⁽³⁾。パーティクルフィルタはパーティクルの分布によって同時に複数の仮説を保持でき、頑健な追跡が可能である。筆者らはパーティクルフィルタを用いて可視領域を撮影するビデオカメラによって撮影した高解像度フルカラー画像に対して、3次元上の物体の大きさと位置推定を精度良く行えることを示した⁽⁴⁾⁽⁵⁾。この手法を赤外線サーマルカメラにも応用できれば、昼夜を問わない自動監視システムが構築可能となる。しかしながら赤外線サーマルカメラは安価で性能が向上したとはいえ、ビデオカメラと比較すると解像度も低く、各画素に対してRGBの3次元情報として表現される可視光領域とは違い、赤外線サーマルカメラは1次元情報である。そのため文献4,5の手法において、各画素のRGB輝度を用いて算出しているパーティクルの尤度の算出や3次元世界座標と2次元画像座標を対応づけるキャリブレーションが精度良く処理できるか明らかでは無い。

そこで本研究では、近年、急速に利用が広まっている赤外線サーマルカメラを用いてカメラから遠く離れた湾岸領域を対象にした場合、昼夜を問わず進入物体の検出・追跡を行う自動監視が可能であることをいくつかのシミュレーション実験から明らかにする。

本稿では、まず、2.にて赤外線サーマルカメラの性能等について説明する。次に、3.にて検出領域を用いたパーティクルフィルタによる物体検出について説明した後、4. および5.にて赤外線サーマルカメラにおけるキャリブレーションと物体の大きさ推定の精度について考察する。最後に6.にて湾岸領域を長時間撮影したカメラ画像に対する物体検出と追跡精度について考察し、筆者らが文献4,5にて提案した手法が赤外線サーマルカメラを用いた湾岸領域の自動監視についても利用できることを示す。

2. 赤外線サーマルカメラ

近年、屋外監視に多く用いられるようになった非冷却型サーマルカメラは、マイクロボロメータと呼ばれる検出素子を持ち、測定波長はおよそ8~14 μm で、感度に波長選性が少ないため汎用性に富み、冷却装置が不要なため信頼性も高い。

本研究で用いたAXIS社製サーマルカメラQ1910の諸元を表1に示す。検出波長は8~14 μm で、自動車のライトに多く含まれる近赤外線の波長である0.8 μm より長い遠赤外線と呼ばれる波長の赤外線を主に検出する。そのため、夜間の自動車のライトなどに影響されずに撮影をすることが出来る。通常の監視用カメラと比べ赤外線サーマルカメラが夜間の撮影にどの程度有効か示すために同じAXIS社製のFull HD監視用カメラであるQ1755と本研究で用いたサーマルカメラの画像を比較した。監視カメラQ1755は、夜間利用のためのナイトモードを持ち、最低照度は白黒撮影時0.2ルクスである。図1(a)に示す場所を対象に、監視カメラQ1755と赤外線サーマルカメラQ1910を用いて夜間に撮影した画像をそれぞれ(b), (c)に示す。いずれも夜2時頃の同時刻に撮影した画像である。対象領域はカメラから湾岸部までは200m程度離れている。(b)のカメラでは外灯で照らされている部分しのみが確認可能で、車が停車していることが分からない。しかし、(c)の赤外線サーマルカメラで撮影した画像では車が停車していることが確認できる。このように、サーマルカメラは光源の無い暗い場所においても物体の観測が可能である。

しかしながら、Full HD監視用カメラの解像度1920×1080と比べるとサーマルカメラの解像度は低く、本研究に用いたサーマルカメラの場合160×128であり、これは一般的なサーマルカメラと同等である。このような解像度の低いカメラの場合、遠く離れた撮影対象がどの程度の大きさに写るかが問題となる。このような測定レンジを問題とする場合、その性能を表すものにジョンソンの判定基準(Johnson's criteria)がある。ジョンソンの判定基準は、検出、認識、識別にはそれぞれ1.5画素、6画素、12画素が必要と考え、この大きさに写る被写体からの最大距離を用いることで、撮影レンジを表している。従って、物体が大きくなるとこの距離は長くなる。

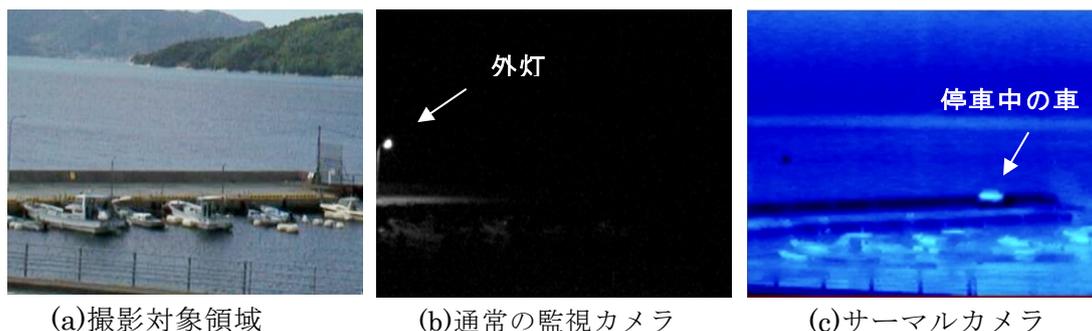


図1 夜間撮影時の画像の比較

表1 本研究に用いたサーマルカメラ Q1910 の諸元

検出器	マイクロボロメータ
検出波長[μm]	8~14
解像度[画素]	160×128
焦点距離[mm]	13
平均水平方向画角[°]	17
検出 (1.5 画素) 人/車[m]	260/800
認識 (6 画素) 人/車[m]	65/200
識別(12 画素) 人/車[m]	30/100

表1にこの基準に対するサーマルカメラ Q1910 の撮影レンジを示す. 例えば, 「認識」に必要な6画素を得ることが出来る最長距離は, 人(幅 0.5m)であれば 65m, 車(長さ 4m)であれば 200m であることをそれぞれ示している. 本実験では観測対象からおよそ 200m 離れた地点の 20m の高さにある場所にカメラを設置して漁港を撮影した. 表1から推測すると, このカメラを用いた場合, 人物の胴体はおよそ2画素, 自動車の側面は6画素程度にしか表示されないことが分かる. 本研究では3次元世界座標と画像の2次元座標との対応をとるためカメラのキャリブレーションを行う必要があるが, そのときの誤差が2画素以下であることが必要となる. このキャリブレーションの精度については, 改めて4.にて検討する.

3. パーティクルフィルタを用いた物体検出と追跡

パーティクルフィルタとは多数の粒子を用いて, 前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, 複数の仮説を同時に保持することができる. 本研究では文献4に述べられているトラップ型の領域を用いたパーティクルフィルタを用いて物体検出・追跡を行った. 文献4の手法を以下に簡単に説明する.

パーティクルフィルタを用いた移動物体の追跡は「予測」「観測」「リサンプリング」の処理を繰り返すことで行う. パーティクルはそれぞれが存在する確からしさを表す尤度をもつ. これら三つの処理を行うと, 尤度の低いパーティクルは消滅し, 尤度の高いパーティクルは復元抽出により何度も選択される. このようなアルゴリズムを用いて物体追跡を行う. 以下に三つの処理を説明する.

「予測」

時刻 t におけるパーティクルの位置を予測する. 予測モデルとして式(1)を用いた.

$$\mathbf{x}_t^{(i)} = \mathbf{x}_{t-1}^{(i)} + \mathbf{v}_{t-1}^{(i)} + \mathbf{w}_t^{(i)} \quad (1)$$

$$\mathbf{v}_{t-1}^{(i)} = \mathbf{x}_{t-1}^{(i)} - \mathbf{x}_{t-2}^{(i)} \quad (2)$$

式(1)の $\mathbf{x}_t^{(i)}$ は時刻 t における i 番目のパーティクルの3次元座標 $\mathbf{x}_t^{(i)} = (x_t^{(i)}, y_t^{(i)}, z_t^{(i)})$ であり、 $\mathbf{w}_t^{(i)}$ は3次元のガウシアンノイズを表している。式(2)の $\mathbf{v}_{t-1}^{(i)}$ は時刻 $t-1$ でのパーティクルの速度であり、時刻 $t-2$ から時刻 $t-1$ までのパーティクルの移動量を表す。

「観測」

予測された位置での各パーティクルの尤度を算出する。時刻 t における i 番目のパーティクルの尤度 $\pi_t^{(i)}$ は2次元画像上にパーティクルを投影した点における背景差分値などにより求める。具体的な尤度の算出方法については5. で述べる。

「リサンプリング」

各パーティクルの重みに比例した確率でパーティクルを復元抽出により選択し、パーティクルの集合を得る。重みは全パーティクルの尤度合計で観測の際に求めた各パーティクルの尤度 $\pi_t^{(i)}$ を正規化したものである。この処理を行うことで、重みの大きいパーティクルは何度も選択され、物体が存在すると推定される領域にパーティクルが集合する。

物体の進入検出と追跡は次のように行う。まずパーティクルを予め定められたトラップ型の領域に様に散布する。この領域を検出領域と呼び、この領域にあるパーティクルフィルタを特に検出フィルタと呼ぶ。検出フィルタは「観測」のみを行う。「観測」により物体が進入したと判断したフィルタは「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、このパーティクルフィルタは検出した物体の追跡を行う。進入物体の検出・追跡において、検出フィルタが進入物体を検出し、追跡を開始すると同時に別の物体の進入に備え、新たな検出フィルタを設置する。これにより、複数の物体に対する検出と追跡が可能となる。

パーティクルフィルタを用いた進入物体の追跡の様子を図2に示す。図2(a)は観測場所を通常のカメラで撮影した画像で、図2(b)~(d)はサーマルカメラで撮影し、濃淡画像として表示したものである。このうち、(b)は検出領域にあるパーティクルを可視化したもので、(c)は検出対象を含む入力画像で、(d)は追跡中のパーティクルを可視化した画像である。(b)の画像にみられる赤い帯状の部分は、設置した二つの検出領域にあるパーティクルを表しており、パーティクルは各検出領域に2000個ずつ散布されている。この設置した検出領域を(c)にみられる進入体である自転車が通ると、(d)の結果のように進入物体である自転車にパーティクルが集中し、物体の移動と共にパーティクルも追跡を行う。(d)は、連続して進入する二つの物体に対しそれぞれパーティクルの集団を割り当てて個別に追跡していることを示している。

4. サーマルカメラにおける世界座標と画像座標の対応関係の精度検証

4.1 カメラキャリブレーションによる2次元空間と3次元空間の対応付け

2次元の画像を観測して3次元空間(実世界)における物体の大きさや位置を知るために、カメラキャリブレーションを行い、3次元空間である世界座標と2次元空間である画像座標とを対応させる必要がある。しかし2. で述べたように低解像度のサーマルカメラにおいては、検出物体は非常に小さく写るためキャリブレーション誤差はより小さくなることが求められる。そこで、解像度の低いサーマルカメラを用いた場合でも、世界座標と画像座標との対応付けの誤差が許容範囲内に収まるか、その対応精度を調べた。

カメラキャリブレーションとは、カメラによって撮影された画像から、焦点距離などのカメラ固有の内部パラメータと世界座標系におけるカメラの位置や姿勢を表す外部パラメータを求める処理である。今回、カメラキャリブレーションはZangの手法⁽⁶⁾をもとに作成されたOpenCVライブラリ⁽⁷⁾を用いて行った。以下に概略を示す。

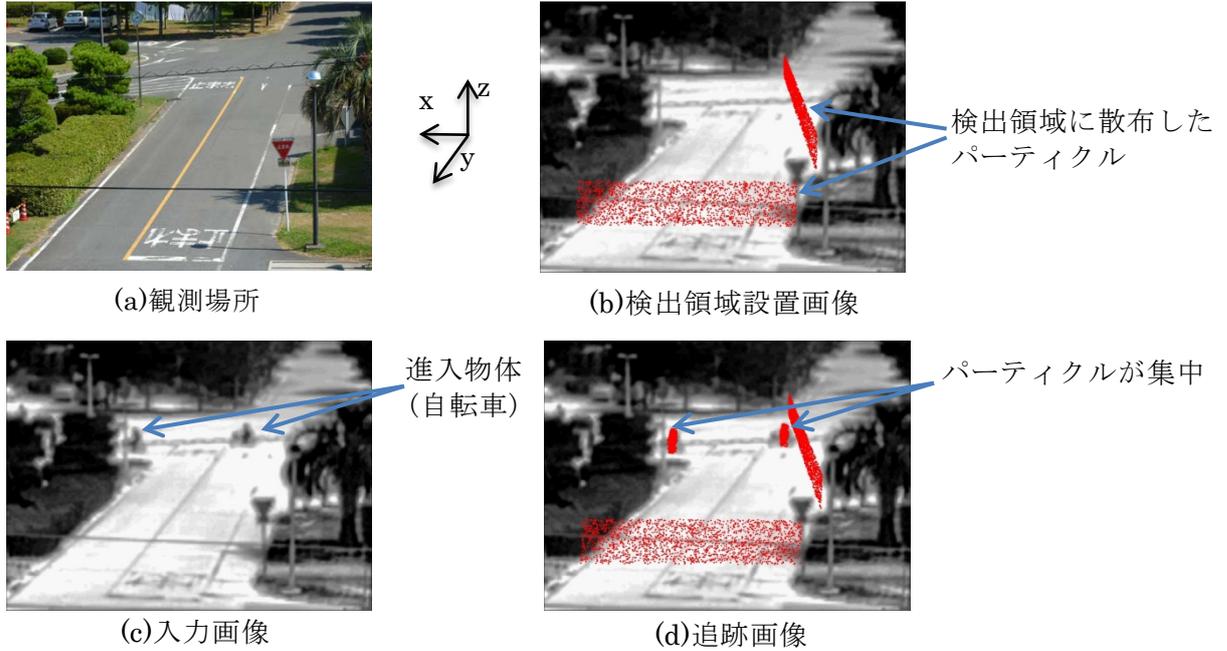


図2 パーティクルフィルタを用いた物体追跡

世界座標の点を $P = (x, y, z)$, P に対応する画像座標の点を $p = (u, v)$ とすると, これらの点の関係は, 点 P の同次座標表現 $P = (x, y, z, 1)^t$ と点 p の同次座標表現 $p = (u, v, 1)^t$ を用いて,

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

と表現できる. 式(3)における s は画像スケール係数, A はカメラの内部パラメータ行列, $[R|T]$ はカメラの外部パラメータ行列をそれぞれ表している. カメラの内部パラメータ行列 A 及び外部パラメータ行列 $[R|T]$ はそれぞれ式(4), (5), (6)で表現できる.

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

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

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

式(4)において, f は焦点距離, (s_u, s_v) は u および v 方向の画素サイズ, (u_0, v_0) は画像平面とカメラ光軸の交点である主点座標, k_s は画像平面上の u および v 軸の傾きを表すせん断係数を表し, 式(5)における $\theta_x, \theta_y, \theta_z$ はカメラの世界座標系に対する回転量, 式(6)の t_x, t_y, t_z は世界座標における原点からのカメラの平行移動量を表している.

4. 2 サーマルカメラに対するキャリブレーションの精度検証

カメラキャリブレーションは, まず内部パラメータ行列 A を推定し, この値を用いて外部パラメータ行列 $[R|T]$ を推定する. 内部パラメータ行列 A の推定は, チェッカーボードを様々な角度に変えて撮影した 15 枚の画像を用いて推定した. 外部パラメータ行列は, 推定した内部パラメータ行列と 3次元座標値 (x, y, z) , 画像座標値 (u', v') を用いて求める. 本研究では, 11 点の 3次元座標値 (x, y, z) および画像座標

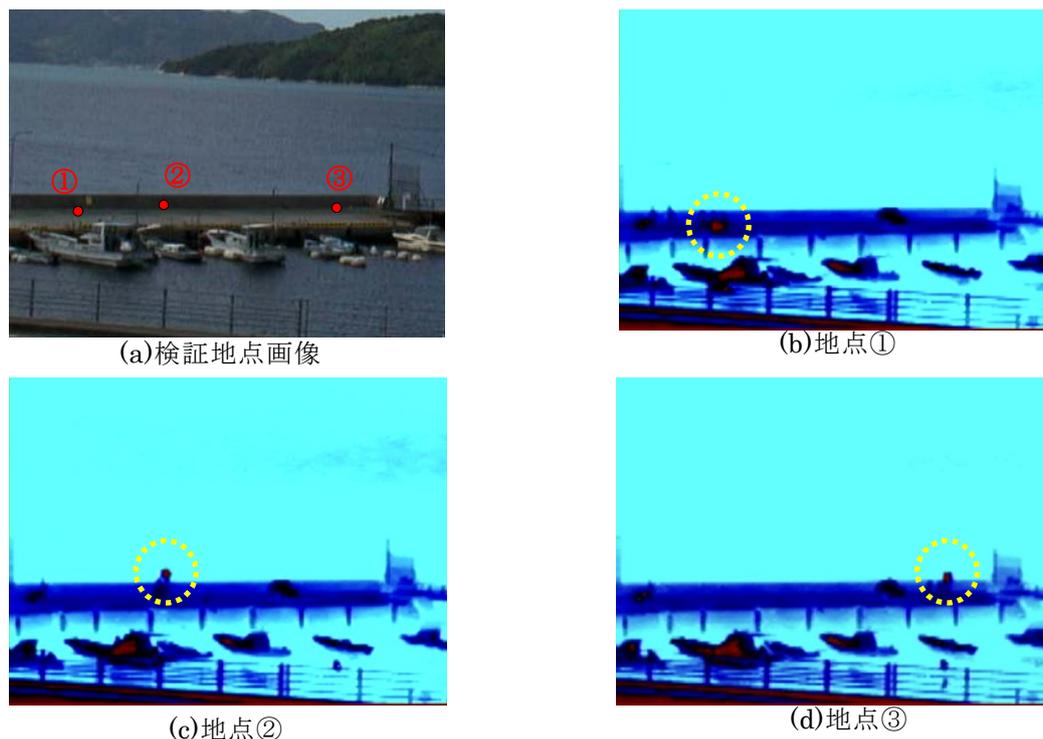


図 3 カメラパラメータの精度検証地点

表 2 カメラキャリブレーションの精度検証

検証地点	3次元座標値 (x, y, z)[cm]	2次元座標値 (u', v')[画素]	式(3)により推定 した2次元座標値 (u, v)[画素]	誤差 ($u' - u, v' - v$) [画素]
①	(82, 2468, 18)	(144, 323)	(146, 324)	(-2, -1)
②	(-771, 1644, 200)	(229, 287)	(232, 287)	(-3, 0)
③	(177, 305, 200)	(482, 292)	(481, 294)	(1, 2)

値(u', v')を用いて外部パラメータの推定を行った。

図 3(a)に観測対象領域を示す。キャリブレーションの精度検証は (a)の①～③の 3 地点にて行った。検証は次のような手順で行った。まず、各点の世界座標系での座標 $P = (x, y, z)$ を手作業にて測定する。次に式(3)を用いて、この三点が画像上のどこに投影されるか計算する。この値が式(3)における $p = (u, v)$ となる。最後に①～③の各点をキャリブレーションが終了したサーマルカメラで撮影し、各地点の画像上での座標 $p' = (u', v')$ を求め、投影点の画像座標値(u, v)と実際の画像座標値(u', v')との比較を行った。なお、①～③の各点を分かりやすくするために、撮影には各点に 1500W の電気ストーブ(発光部が 30cm × 30cm)を置いて測定した。

検証結果を表 2 に示す。各地点における誤差は平均すると 1.5 画素程度であり、2.で述べたように人物を撮影した画像サイズである 2 画素よりもわずかに小さく、また、自動車の長さ 6 画素よりも十分小さいため許容範囲であると考えられる。この結果から 160 × 128 程度の低解像度のサーマルカメラを用いて 200m ほど離れた地点でも、2次元空間である画像座標と 3次元空間である世界座標の対応付けが可能であることが分かった。

5. 物体に対する大きさ推定

単眼視点固定型カメラを用いた場合, 観測対象物体の大きさを精度良く推定することは難しい. この問題に対して, 筆者らは文献5において, 物体の高さ方向に対する存在位置を仮定することで, 精度良く物体の大きさを推定できることを示した. 本論文では文献5の手法が, 低解像度のサーマルカメラにおいても物体の大きさ推定に利用できることを確かめる. 以下に文献5で提案された手法を簡単に説明する.

5. 1 物体中心位置と物体の大きさ推定

物体は立方体であると仮定し, 物体の中心位置は, リサンプリング後のパーティクルの存在位置 $\mathbf{x}_t^{(i)} = (x_t^{(i)}, y_t^{(i)}, z_t^{(i)})$ をもとに, 式(7)により求めたパーティクルの平均位置 $\mathbf{O} = (x_o, y_o, z_o)$ とする. 次に, 物体の x 軸方向の大きさは推定された物体中心位置 \mathbf{O} を中心としたパーティクルの広がりを表す偏差 σ_x の2倍とし, 式(8)より求める. 式(8)において, N はパーティクルの個数を表す. 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.

$$\mathbf{O} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_t^{(i)} \quad (7)$$

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^N (x_o - x_t^{(i)})^2 \quad (8)$$

5. 2 パーティクルの存在範囲の限定

3次元空間上で進入物体検出・追跡を行うシステムにおいて, 視点固定型カメラを用いると, 図4に示すように一つの画像座標に対し複数の世界座標が対応する一対多の関係が成り立つ. このような場合, 実世界上では物体が存在しない P_1 や P_3 にあるパーティクルも高い尤度を持ち, 推定される物体領域が実際よりも大きなものとなる問題点がある.

この問題を解決するために文献5で提案された手法を以下に示す. この手法においては, 検出対象物体は3次元空間において同じ高さの水平面上に連続して存在し, その高さは一定値 H 以下であると仮定することにより, パーティクルの存在範囲の限定を行う. まず, 各パーティクルの3次元座標値を高さ方向に 0 から dz ずつ H まで変化させ, 各座標を2次元画像に投影し, $T+1$ 個の背景差分値を求める. ここで, $T = H/dz$ である. 求めた $T+1$ 個の背景差分値において, 閾値 γ より大きな背景差分値の数を尤度 $\pi_t^{(i)}$ とする.

図4を用いて物体上に存在しないパーティクルの尤度を小さくする方法について説明する. 図4において破線に囲まれる領域は物体内部の画像領域に投影される3次元上の領域を表している. 3次元上の点 $P_2 = (x_2, y_2, z_2)$ に対し, 高さ方向の座標 z_2 を $0, dz, 2 \cdot dz, \dots, T \cdot dz$ と変化させてもそれらの点の多くは破線で囲まれる領域内に存在し, 図4においては7点が物体領域内に投影されている. 3次元上の点 P_1 において, 同様に高さ方向の座標を変化させると3点が物体内に投影される. 同様に P_3 は4点となる. そのため物体外部の点である P_1 や P_3 は, P_2 に比べると尤度は低くなり, 物体外部にあるパーティクルは選択されにくくなる. ここで, 垂直方向に対するカメラの傾きが大きくなるに従い物体領域は正確に推定されるようになるが, 物体がカメラから遠くにある場合は実際よりも大きく推定されるようになることに注意しておく.

尤度 $\pi_t^{(i)}$ は i 番目のパーティクルの状態を $\mathbf{x}_t = (x_i, y_i, z_i)$ とした場合, 以下のように書くことができる.

$$\pi_t^{(i)} = \sum_{k=0}^T \delta(x_i, y_i, k \cdot dz) \quad (9)$$

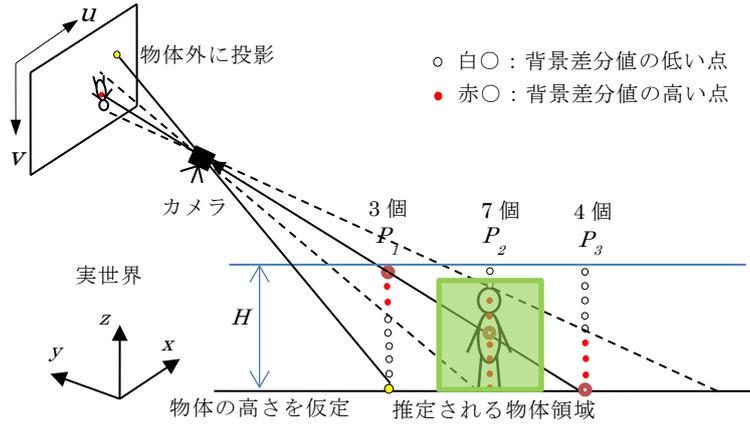


図4 パーティクルの存在範囲の限定の様子

$$\delta(x_i, y_i, k \cdot dz) = \begin{cases} 1 & \text{if } F(x_i, y_i, k \cdot dz) > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

式(10)の $F(x_i, y_i, k \cdot dz)$ は変化させた3次元座標値 $(x_i, y_i, k \cdot dz)$ を2次元画像投影した点 (u, v) における背景差分値であり，式(11)より求める．ただし，背景の明度変化の影響を除くために閾値 γ 以下の値は0とする．

$$F(x_i, y_i, k \cdot dz) = \sqrt{(r_I(u, v) - r_B(u, v))^2 + (g_I(u, v) - g_B(u, v))^2 + (b_I(u, v) - b_B(u, v))^2} \quad (11)$$

式(11)において $r_I(u, v)$ ， $g_I(u, v)$ ， $b_I(u, v)$ では入力画像上の点 (u, v) のRGB輝度値で， $r_B(u, v)$ ， $g_B(u, v)$ ， $b_B(u, v)$ は背景画像上の点 (u, v) におけるRGB輝度値をそれぞれ表す．本研究では背景変動を考慮した背景学習を，次式を用いて行った．ここで， I_t^b は時刻 t における背景画像， I_t は時刻 t における入力画像， ζ は背景画像更新パラメータを表している．なお，学習は推定された物体領域を除いた領域に限定して行っている．

$$I_{t+1}^b = \zeta I_t + (1 - \zeta) I_t^b \quad (12)$$

6. 物体のサイズ推定と追跡実験

6. 1 実験概要

本実験では，サーマルカメラを用いて湾岸部を撮影した動画像に対し，パーティクルフィルタを用いた物体検出・追跡実験を行い，検出と追跡成功率をそれぞれ求めた．観測対象は実験中に撮影された人物，自動車であり，以下に実験環境の詳細を示す．

使用したコンピュータはOSがWindows 7 Home Premium, CPUはIntel®Core™i7の2.93GHzで，メモリは6GB，使用した動画像の解像度は目視しやすくなる様デジタル的に拡大した640×480画素である．撮影に使用したカメラはAXIS社のQ1910であり，毎秒8フレームの設定で高さ約20mの位置から，およそ200m離れた湾岸部を夏期晴天時の昼夜に約18時間撮影した．実験に用いたパラメータ等を表3に示す．カメラから物体までの距離が約200mと離れているため，物体領域の大きさについては，奥行き方向であるx軸方向の大きさが実際よりも大きく推定されると予想し，5.2における dz の値を10cmと小さくとり，パーティクルの尤度を詳細に調べるようにした．

撮影した動画像は濃淡表示ではなくカメラに備わった疑似カラー表示にて処理をおこなった．これは事前に二つの表示方法を目視により比較した結果，人物や自動車などの物体がより明確に表示されると判断したためである．図5に疑似カラーモードにて撮影した画像を示す．(a)は昼間に撮影した画像，(b)は夜間に撮影した画像である．昼間の画像である(a)においては，漁港の地上部のアスファルトが太陽

表 3 実験条件

項目	値
カメラから検出対象までの距離[m]	200
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 $w_t^{(i)}$ [cm]	15
物体追跡を開始する閾値 α [%]	5
物体追跡を終了する閾値 β [%]	1
観測におけるノイズ除去のための閾値 γ	5
高さ方向の変化量 dz [cm]	10
高さ方向の最大値 H [cm]	200

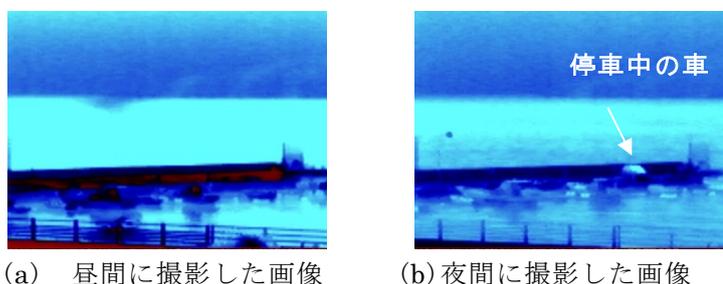


図 5 サーマルカメラを使って撮影した実験対象領域

光で高温となり赤く表示されている。(b)の夜間においては、昼間高温であった部分に車が駐車していることが分かる。

6. 2 物体の大きさ推定

撮影時間内に進入してきた物体は、人物と自動車の2種類である。2種類の物体の昼間、夜間における推定した大きさを表4に示す。まず、人物の領域に注目してみると、実際の y 軸方向(横方向)の大きさが50cm程度に対して、推定した大きさは、それぞれ、昼間が68cm、夜間が100cmであった。 y 軸方向に移動中の車では、実際が約400cmに対して、昼間が213cm、夜間が174cmであった。同様に、高さ方向である z 軸方向での推定は、人物が85cmと130cm、自動車が85cmと73cmとなり、やや小さいものの実際の大きさに近い値が推定されている。実際の大きさよりやや小さい理由としては、人物では下半身、自動車では車体の周辺部分が低温であったため、背景との差分が小さくなり、その結果物体領域が小さくなったためである。

表 4 推定された物体領域の大きさ

追跡対象 物体	時間	推定された物体領域[cm]		
		x 軸方向	y 軸方向	z 軸方向
人物	昼間	252	68	83
	夜間	278	100	130
自動車	昼間	322	213	85
	夜間	341	174	73

表 5 進入物体検出・追跡成功率
進入物体数 50 個

	成功数[個]	成功率[%]
検出	43	86
追跡	35	70

奥行き方向は、奥行き（ x 軸）方向に関しては、人物が 252cm と 278cm、自動車が 322cm と 341cm となっている。実際の高さは人物が 50cm、自動車が 200cm 程度であるから、実際よりも大きくなっている。これは、200m 離れた地点を 20m の高さから観測するため、垂直方向に対するカメラの水平線からのずれが少ないため、5. 2 で述べたパーティクルの存在範囲が大きくなるためである。物体がカメラから近い位置にある場合にはこの様な誤差は小さくなる。そこで、物体のカメラからの距離を用いて、大きさを補正するなどの対策が考えられる。

6. 3 物体の検出と追跡

検出と追跡成功率について、その性能を調査した。調査した結果を表 5 に示す。追跡成功の定義は、物体の進入から退出まで追跡を行っていることとした。

人物と自動車を合わせて 50 個の進入物体に対して、検出成功数は 43 個、追跡成功数は 35 個であり、検出成功率は 86% で追跡成功率は 70% となった。また、進入物体が無いにもかかわらず進入を検出したり追跡続けたりする例は無かった。

図 6 に検出領域と追跡中の画像を示す。図 6(a) は昼間の画像、(b) は夜間の画像である。検出領域をわかりやすくするために、検出領域にあるパーティクルを赤で表した。この領域に物体が入ると進入を検出し、追跡を始める。図 6(c)～(f) は追跡中の画像を示す。これらの画像においては物体をわかりやすくするため、パーティクルは表示していない。(c)～(f)において、推定した物体は、緑の直線で囲み、物体の中心を黄色の点で示した。物体部分を拡大してそれぞれの画像の右上隅に赤枠で囲んで表示している。撮影した時期が夏期であるため昼間の陸上のアスファルトは高温で、図 6(c)の陸上部分のほとんどが赤くなっているにもかかわらず小さく写った人物を昼間に追跡できている。夜間のコンクリート部分は周辺に比べ温度が低くなるため黒く、進入物体は温度が高く白く表示され、(d)(f)にあるように人物と自動車をそれぞれ追跡できている。陸上部は昼間と夜間の背景部分の温度差が激しく、同じ背景箇所でも昼と夜では全く違う色で表示される。このような背景画像の変動にもかかわらず昼夜を通して検出できたのは、式(12)による背景画像の学習が大きく寄与していると考えられる。

図 7 に人物を対象とした物体検出および追跡の失敗例を示す。図 7 の(a)～(c)は時刻 $t=761\sim 768$ (秒)における同一の人物に対する連続した検出・追跡結果の一部である。各画像において、人物周辺の画像を拡大し、人物領域付近を黄色の破線で囲んでいる。(a)は人物が検出領域に入る前の画像である。(b)は検出領域で物体が存在すると推定され物体を矩形で囲んだ状態の時の画像である。(c)は人物が存在するにもかかわらず追跡を終了した状態の時の画像である。このように物体領域が非常に小さく、地面の温度と人物の温度が近い場合、大きな尤度は得られず、(b)のように人物の体の上半身部分のみを検出し、(c)のように閾値処理によって追跡を途中で終了してしまっ。追跡の失敗は、このように進入物体が肉眼でも判別が難しい場合に多く発生している。

7. むすび

本研究では、昼夜を問わない進入物体の検出・追跡を行うことを目的としてサーマルカメラとパーティクルフィルタを用いた。カメラから 200m 程度離れた観測地点に対してキャリブレーション行った

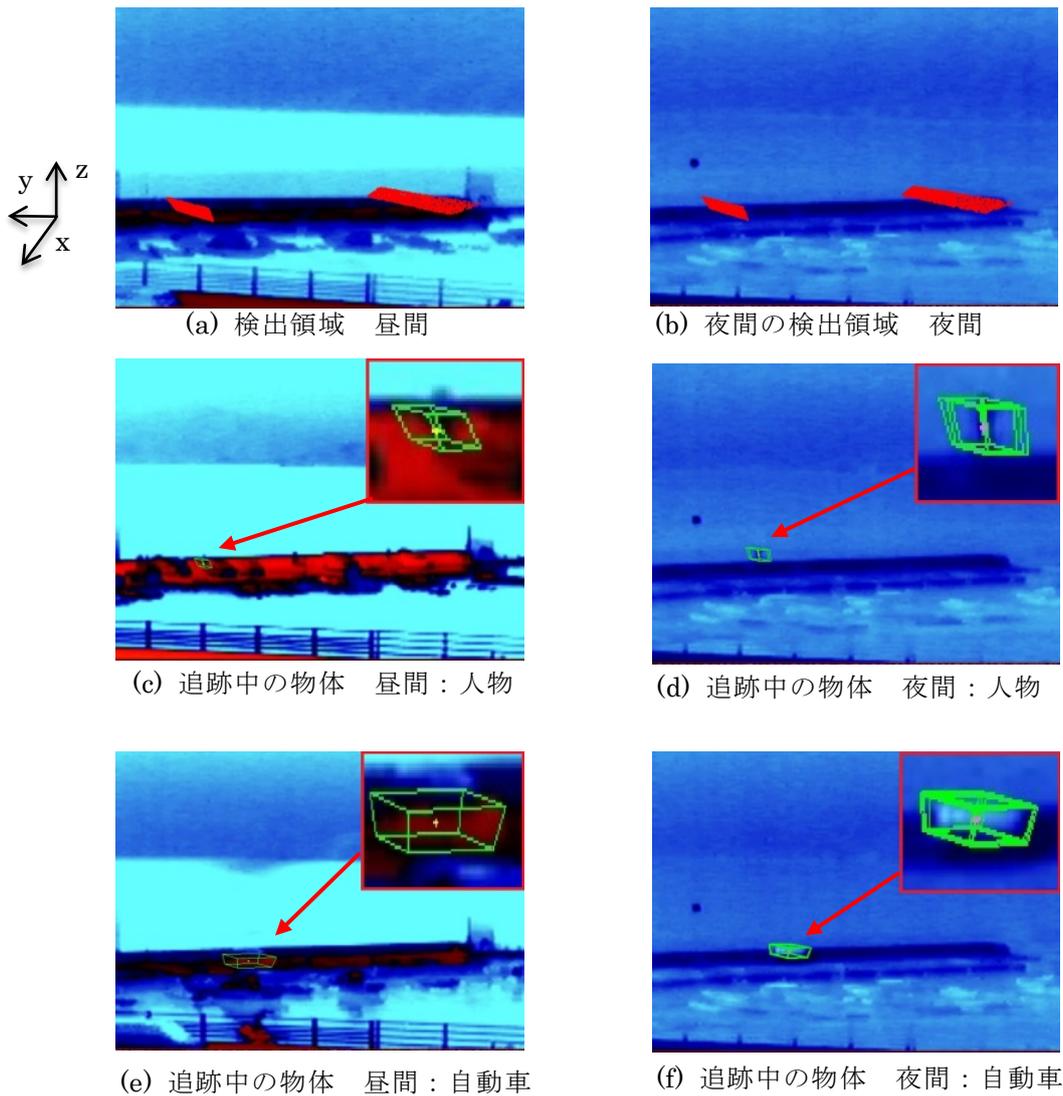


図 6 検出領域と追跡例

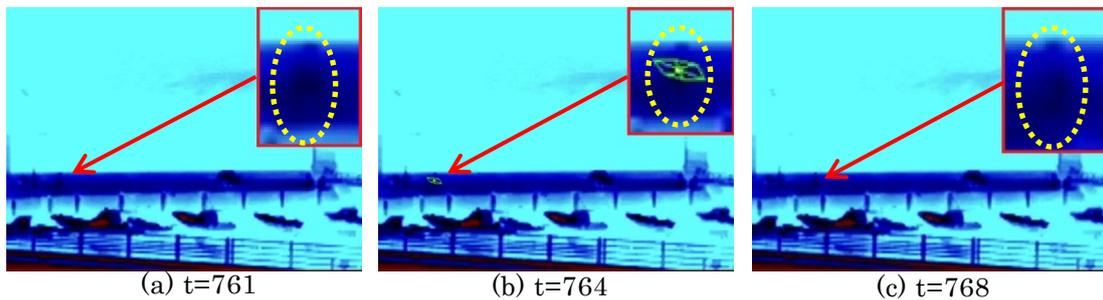


図 7 検出・追跡失敗例

結果, キャリブレーション誤差は小さく, 低解像度のサーマルカメラを利用しても画像座標と世界座標の対応付けが可能であることが分かった. 文献 5 で提案された手法をもとに物体領域の推定を行った結果, 実際の大きさに近い値を得ることができた. 奥行き方向の大きさに対しては実際よりも大きく推定されたが, カメラからの 3 次元上での距離を使うことにより誤差を小さくすることが出来ると考えられ

る。更に、サーマルカメラを利用した場合、昼夜の温度差が激しく背景部分の変動が大きく背景差分を利用した検出が難しくなるが、背景の学習を行うことにより対応出来ることが分かった。サーマルカメラは、現在でも急速に高解像度化と低価格化が進んでいる。そのため、高解像度のサーマルカメラを使えば更に良好な結果が得られると考えられる。

今後の課題として、雨、濃霧、晴天などの様々な天候下における長期的な観測、大きな波浪がある中での海上から進入してくる物体の検出・追跡実験などが必要である。

謝辞

本研究は日本学術振興会の科学研究費補助金（基盤（C）23510210）の助成を受けたもので有る。

参考文献

- [1] 浜村浩之，岡村健史郎，“湾岸監視を目的とした固有空間法による物体検出に関する考察”，電気・情報関連学会中国支部第 58 回連合大会,pp.211,2007
- [2] 樋口和之，“粒子フィルタ”，電子情報通信学会誌，Vol.88,No.12,pp.989-994,2005.
- [3] 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.
- [4] 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退出検出”，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
- [5] 吉本勇介，岡村健史郎，松村遼，“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”，大島商船高等専門学校紀要，第 44 号，pp.47-57，2011
- [6] Z.Zhang,” A flexible new technique for camera calibration”，IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22 (11), pp.1330-1334, 2000.
- [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.