

知能化空間での移動ロボットによる 自己位置推定と自動カメラキャリブレーションの同時実行

池 勇勲^{*1}, 山下 淳^{*1}, 浅間 一^{*1}

Simultaneous Localization and Automatic Camera Calibration for Intelligent Space by Mobile Robot

Yonghoon JI^{*1}, Atsushi YAMASHITA^{*1}, and Hajime ASAMA^{*1}

^{*1} Department of Precision Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

To operate mobile robots in an intelligent space, pre-calibration of whole environmental sensors is very important task which would be extremely tedious. Here, we propose a fully automatic simultaneous method to estimate both a robot and environmental camera poses. To simultaneously estimate the whole states which are the 3DOF robot pose and the camera poses, the particle filter is utilized with encoder data from the dead-reckoning sensor and the image data from the environmental cameras. By designing a novel probabilistic measurement model exploited for the particle filter algorithm, the framework is able to accommodate several errors such as odometry and measurements. The experimental result in real environments shows that the propose framework can estimate the whole states including the robot pose and the camera parameters successfully.

Key Words : Mobile robot, localization, camera calibration, particle filter, intelligent space

1. 序論

最近, ロボット分野では産業用ロボットだけでなく, 日常生活支援のためのサービスロボットを実用化するための研究が活発にされている. これらのサービスロボットを屋内環境内で運用するためには環境に対するモデル, すなわちロボットのための地図情報が必要である. このため図 1 に示すように距離センサ情報などを処理し, ロボットのための地図生成に対す様々な SLAM (simultaneous localization and mapping) 技術が提案されている⁽¹³⁾. しかし, このように生成された地図情報は静的なモデルであるため, 環境の変化に対応不可能であり, この場合新しく地図を生成し直す作業が不可欠である.

一方, 図 2 に示すような人間を取り巻く環境内に分散センサインフラを構築する知能化空間 (iSpace: intelligent space) 概念が提案されている⁽⁴⁻⁶⁾. このような知能化空間では環境に対するモデルがなくとも, 分散的なセンサインフラを使用し, 環境をモニタリングすることにより, 人間の行動パターンに応じたサービスを提供するなど, 動的な環境の変化にも対応可能なより優れたロボットの運用が可能であると考えられる.

しかし, このような知能化環境内で信頼性のある情報を取得するためには, すべての分散的なセンサに対する非常に面倒なキャリブレーション作業を事前に厳密に行う必要があり, 手動での測定による位置および姿勢関係の調整は限界があるため, 非常に不正確である.

分散的なセンサに対するキャリブレーション作業を自動的に行うため SLAC (simultaneous localization and auto-calibration) と呼ばれる技術が提案されている⁽⁷⁾. しかし, この技術は距離・方向 (range and bearing) センサ情報のみに基づいた理論のみ成立されているため, 環境のモニタリングのため広く使用されている CCTV カメラなどへの適用は不可能であり, 実際システムへの応用も困難である. そこで本論文は環境内に分散設置されているカメラ (以下, 分散カメラ) 間の位置お

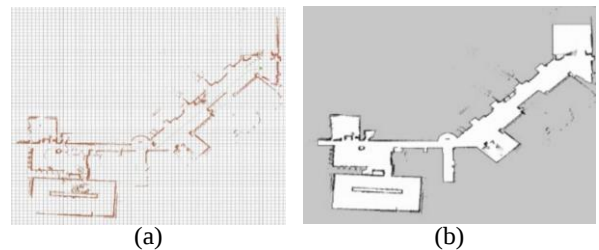


Fig. 1 Example of SLAM: (a) raw range data and (b) occupancy grid map for robot operation.

^{*1} 東京大学大学院工学系研究科精密工学専攻 (〒113-8656 東京都文京区本郷 7-3-1) {ji, yamashita, asama}@robot.t.u-tokyo.ac.jp

よび姿勢関係の手動調整にもかかわらず（すなわち、事前キャリブレーション作業を厳密に行わず）、ロボット自ら自己位置推定を行う同時に分散カメラの外部パラメータを自動的に同定する手法を提案する。本論文においては、ロボットに対する3自由度状態(x_r, y_r, φ_r)とすべての分散カメラに対する3自由度状態($x_c^k, y_c^k, \varphi_c^k$)を同時に推定するため、パーティクルフィルタに基づく状態推定アルゴリズムを用いる。パーティクルフィルタの適用において、移動ロボットの制御入力とカメラ観測値に対する不確実性を考慮した観測モデルを定義することにより、高次元状態の安定的な推定が可能であり、地図情報なしでロボットの運用が可能になると考えられる。

本論文の構成は、以下の通りである。まず第2章では推定対象になる知能化空間での状態変数を定義する。第3章では運動モデルおよび観測モデルの定義を含むパーティクルフィルタに基づく提案手法の詳細を説明する。また第4章では、提案した手法を用いて、実際の環境での実験結果を示す。そして最後に、第5章において結論を述べる。

2. 知能化空間の状態定義

本論文ではロボットおよび知能化空間の状態(すべての分散カメラに対する外部パラメータ)を同時に推定するため、推定対象になる状態ベクトル $\mathbf{X}$ を次式のように定義する。

$$\mathbf{X} = \left[(x_r)^T \quad (x_c^1)^T \quad (x_c^2)^T \quad \cdots \quad (x_c^k)^T \quad \cdots \quad (x_c^n)^T \right]^T \quad (1)$$

$$\mathbf{x}_r = [x_r \quad y_r \quad \varphi_r]^T \quad (2)$$

$$\mathbf{x}_c^k = [x_c^k \quad y_c^k \quad \varphi_c^k]^T \quad (3)$$

ここで、 $\mathbf{x}_r, \mathbf{x}_c^k$ はそれぞれ時間 t でのロボットに対する3自由度位置と姿勢、 k 番目の分散カメラに対する3自由度位置と姿勢を意味する状態ベクトルであり、図2に示すようにすべて世界座標系に対する状態変数

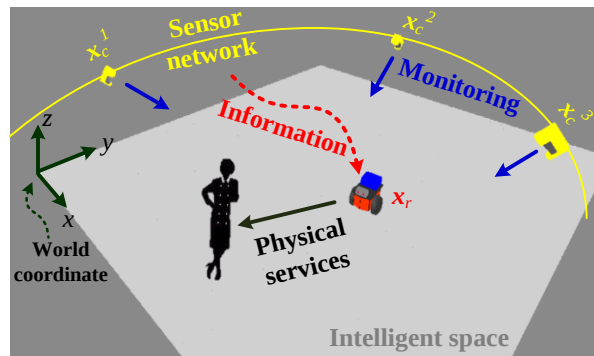


Fig. 2. Concept of intelligent space.

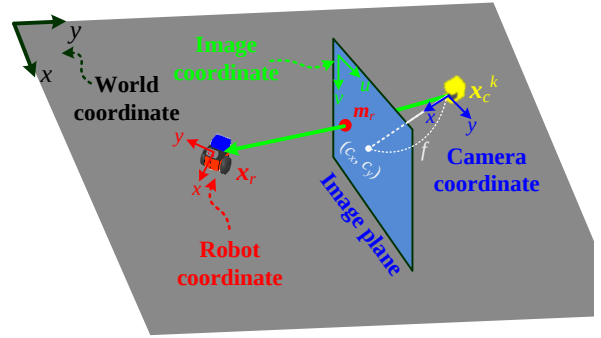


Fig. 3. Relationship between world coordinate and image coordinate according to camera calibration.

である（現在の状態を表す下付き文字 t は省略）。また本論文では、ロボットおよび分散カメラの状態は2次元空間上で定義されると仮定するため、状態ベクトルでの z 成分はすべて考慮しない。 n は知能化空間内に配置されている分散カメラの数である。すなわち本論文での提案手法は、 $3+3n$ 個の要素を含む高次元状態ベクトルを移動ロボットの制御入力と分散カメラからの画像情報のみに基づいて推定する手法である。

次式は空間座標系と画像座標系との関係を定義するカメラキャリブレーションモデルである。

$$\mathbf{m} = \mathbf{A}\mathbf{T}^k \mathbf{x} \quad (4)$$

$$\mathbf{A} = \begin{bmatrix} f & f_{\text{skew}} & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$\mathbf{T}^k = \begin{bmatrix} \sin \varphi_c^k & -\cos \varphi_c^k & 0 & y_c^k \cos \varphi_c^k - x_c^k \sin \varphi_c^k \\ 0 & 0 & -\cos \varphi_c^k & 0 \\ \cos \varphi_c^k & \sin \varphi_c^k & 0 & -x_c^k \cos \varphi_c^k - y_c^k \sin \varphi_c^k \end{bmatrix} \quad (6)$$

ここで、内部パラメータ(c_x, c_y)と f はそれぞれ画像の主点 (principle point) とカメラの焦点距離 (focal length) であり、 f_{skew} は非対称係数 (skew coefficient) を意味する。式(3)の分散カメラに対する状態が正しく推定可能であれば、それに対する同次変換行列の逆行列を計算することで式(6)の外部パラメータ行列 $\mathbf{T}^k$ が求まる。したがって、式(4)に示すように内部パラメータ行列 $\mathbf{A}$ が既知であれば、空間座標 $\mathbf{x} = [x \ y \ z \ 1]^T$ を画像座標 $\mathbf{m} = [u \ v \ 1]^T$ に変換することができ、分散カメラに対するキャリブレーション作業が可能となる。たとえば、図3に示すように $\mathbf{x}_r$ の状態であるロボットが $\mathbf{x}_c^k$ の状態である分散カメラにより観測されると、式(4)により k 番目の分散カメラでの画像内の観測値は $\mathbf{m}_r = [u_r \ v_r]^T$ になる。

3. パーティクルフィルタによる状態推定

本研究での推定対象になるシステムの全体状態（ロボットと各分散カメラの位置と姿勢） $\mathbf{X}$ は移動ロボット車輪のエンコーダからの制御入力と、各分散カメラからの画像情報を用いて推定され、この際、パーティクルフィルタに基づく状態推定が行われる。パーティクルフィルタはベイズフィルタに基づく状態推定手法であり、推定対象に対する状態の信念（確率分布）を有限個のパーティクルの分布から推定する手法である。すなわち、ノンパラメトリックな手法であるため正規分布などの仮定は不必要であり、任意の確率分布の表現が可能である。したがって、初期の分布が未知である場合など広きに渡り適用可能な手法である。本章ではベイズフィルタを簡単に説明した後、その実装方法の1つであるパーティクルフィルタを本論文での推定問題に適用し詳細に説明する。

3・1 ベイズフィルタ ベイズフィルタはロボットシステムの状態を制御入力およびセンサの観測値を使用し、再帰的に推定を行う手法である。本論文において、時間 t まで累積されたロボットの制御入力を $\mathbf{u}_{0:t} = \{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_t\}$ 、そして分散カメラの観測値を $\mathbf{z}_{0:t} = \{\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_t\}$ とするとき、時間 t でのシステムの状態 $\mathbf{X}$ に対する事後信念は次式のように表現される。

$$Bel(\mathbf{X}) = p(\mathbf{X} | \mathbf{z}_{0:t}, \mathbf{u}_{0:t}) \quad (7)$$

ここで、式(7)をマルコフ過程 (Markov process) を用いて展開すると、次式のように時間 t での制御入力 $\mathbf{u}$ と観測値 $\mathbf{z}$ のみを適用し、時間 $t-1$ での状態に対する信念 $Bel(\mathbf{X}_{t-1})$ から時間 t での状態に対する信念 $Bel(\mathbf{X})$ の計算が可能である。

$$Bel^-(\mathbf{X}) = \int p(\mathbf{X} | \mathbf{X}_{t-1}, \mathbf{u}) Bel(\mathbf{X}_{t-1}) d\mathbf{X}_{t-1} \quad (8)$$

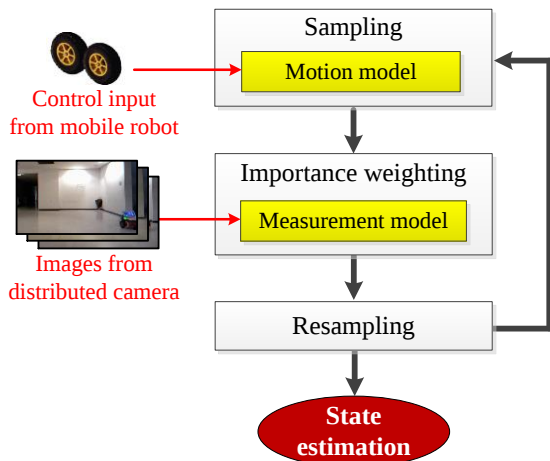


Fig. 4. SIR Particle filter process.

$$Bel(\mathbf{X}) = \eta p(\mathbf{z} | \mathbf{X}) Bel^-(\mathbf{X}) \quad (9)$$

ここで、式(8)は予測プロセスであり、状態遷移確率 $p(\mathbf{X} | \mathbf{X}_{t-1}, \mathbf{u})$ に対する運動モデルを定義することで、状態 $\mathbf{X}$ に対する信念を予測する。また、式(9)は更新プロセスであり、観測確率 $p(\mathbf{z} | \mathbf{X})$ に対する観測モデルを定義することで状態 $\mathbf{X}$ に対する信念を更新する。このようにベイズフィルタに基づくパーティクルフィルタでは、運動モデル $p(\mathbf{X} | \mathbf{X}_{t-1}, \mathbf{u})$ と観測モデル $p(\mathbf{z} | \mathbf{X})$ に対する定義によって性能が強く左右される。これに対して次節で詳細に説明する。

3・2 パーティクルフィルタ パーティクルフィルタは前述したベイズフィルタに対する信念 $Bel(\mathbf{X})$ を次式のように重み w を持つ N 個のパーティクルの集合 $\mathbf{S}$ として表現する。

$$\mathbf{S} = \{ \langle \mathbf{X}^{(i)}, w^{(i)} \rangle | i=1 \dots N \} \quad (10)$$

$$Bel(\mathbf{X}) \approx \sum_{i=1}^N w^{(i)} \delta(\mathbf{X} - \mathbf{X}^{(i)}) \quad (11)$$

ここで、 $\mathbf{X}^{(i)}$ と $w^{(i)}$ はそれぞれ時間 t での i 番目のパーティクルの状態とその重みであり、 N はパーティクルの数である。したがって、式(11)に示すように、重み付けられたパーティクルで、信念 $Bel(\mathbf{X})$ の近似が可能である。ここで、 $\delta(\cdot)$ はディラックのデルタ関数である。

本論文では最も一般的なパーティクルフィルタである SIR (sequential importance resampling) パーティクルフィルタを用いる⁸⁾。SIR パーティクルフィルタのプロセスは図4に示すように3つステップ（サンプリング、重み付け、リサンプリング）に分けられ、ロボットの制御入力はサンプリングステップでの運動モデルに、各分散カメラからの画像情報は重み付けステップでの観測モデルの定義に用いられる。以下、詳細に説明する。

A. サンプリング (Sampling)

サンプリングステップでは次式のようにすべてのパーティクル $i=1, \dots, N$ に対して状態遷移確率 $p(\mathbf{X} | \mathbf{X}_{t-1}, \mathbf{u})$ に基づき、新しいパーティクルの集合を生成する。

$$\mathbf{X}^{(i)} \sim p(\mathbf{X} | \mathbf{X}_{t-1}^{(i)}, \mathbf{u}) \quad (12)$$

このため、次式のような運動モデルを適用し、各パーティクルを遷移させることで、サンプリングステップの実装が可能である。

$$\mathbf{X}^{(i)} = \mathbf{f}(\mathbf{X}_{t-1}^{(i)}, \mathbf{u}) + \mathbf{e} \quad (13)$$

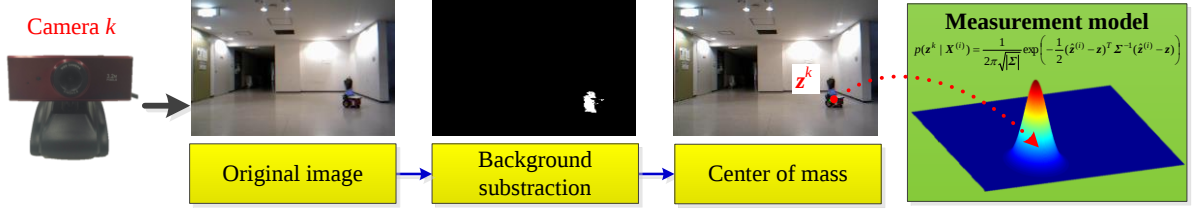


Fig. 5. Definition of measurement information $\mathbf{z}^k$ and measurement model $p(\mathbf{z}^k | \mathbf{X})$.

ただし, $\mathbf{u}=[\Delta x, \Delta y, \Delta \phi]^T$ と $\mathbf{e}$ はそれぞれ時間 t でのロボットの制御入力とランダムノイズを意味する. ここで, 各カメラの状態 $\mathbf{x}_c^k$ に対する運動モデルの定義は困難であるため, ランダムノイズのみ適用する.

B. 重み付け (Importance Weighting)

重み付けステップでは 次式のようにすべてのパーティクル $i=1, \dots, N$ に対して観測確率 $p(\mathbf{z} | \mathbf{X})$ に基づき, 各パーティクルの重みを付与する.

$$w^{(i)} = p(\mathbf{z} | \mathbf{X}^{(i)}) \quad (14)$$

このため, 次式のようなパーティクルの状態 $\mathbf{X}^{(i)}$ で, どのような観測値が観測されるかに関する観測モデルを適用し, 各パーティクルの重みを更新することが可能である.

$$p(\mathbf{z} | \mathbf{X}^{(i)}) = \prod_{k=1}^n p(\mathbf{z}^k | \mathbf{X}^{(i)}) \quad (15)$$

$$p(\mathbf{z}^k | \mathbf{X}^{(i)}) = p(\mathbf{z}^k | \hat{\mathbf{z}}^{(i)}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\hat{\mathbf{z}}^{(i)} - \mathbf{z}^k)^T \Sigma^{-1}(\hat{\mathbf{z}}^{(i)} - \mathbf{z}^k)\right) \quad (16)$$

ただし, Σ はノイズパラメータ行列であり, カメラの内部パラメータに対する誤差などに基づいて適切に設定する. $\mathbf{z}^k=[u_r^k \ v_r^k]^T$ は k 番目の分散カメラからの観測値であり, 図 5 に示すように原画像に背景差分法を適用した後, すべての画素による中心を計算することで求める. ここで, 本論文では環境内に動的物体としてロボットのみ存在する前提の下で, 混合正規分布に基づく背景差分法⁹⁾を使用した, ロボット以外に多数の動的物体が存在する環境においても画像内の物体追跡手法などを融合してロボットの検出が連続的に行われるのであれば, 移動物体の存在する環境にも対応可能である. また $\hat{\mathbf{z}}^{(i)}=[\hat{u}_r^i \ \hat{v}_r^i]^T$ は実際のシステムの状態がパーティクル $\mathbf{X}^{(i)}$ と等価であるとき, k 番目の分散カメラから予測される観測値を意味し, 式(4)のキャリブレーションモデルを適用することにより求まる. ここで, 個々の分散カメラの観測から生じるノイズの間に独立仮定が成り立つと考えられるため, 式(15)に示すよう

に分散カメラ全体に対する観測確率は, 個々の分散カメラに対する観測確率の積として表現される. すなわち, 結果的にロボットと各分散カメラの状態の組み合わせが最も良いパーティクル $\mathbf{X}^{(i)}$ に最も大きな重みが与えられることになる.

C. リサンプリング (Resampling)

リサンプリングステップでは, 各パーティクルの重みに従い, パーティクルを再配置する. すなわち, 高い重みを持つパーティクルからは, 次の段階のパーティクルを多く生成し, 低い重みを持つパーティクルからは, 少ない数のパーティクルを生成, あるいはパーティクルを生成しない. パーティクルの重みに従い, 再配置を行った後, 各パーティクルの重みを $w^{(i)}=1/N$ にリセットする. ここで, 全体システムの状態に対する推定値は次式のようにパーティクルの加重平均により求められる.

$$\hat{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^N w^{(i)} \mathbf{X}^{(i)} \quad (17)$$

4. 実験結果

提案した知能化空間での移動ロボットによる自己位置推定と自動カメラキャリブレーションの同時実行手法を検証するため, 図 6 に示すように MobileRobots 社の Pioneer 3-DX 移動ロボットと, 分散カメラとして解像度 640×480 のカメラ 3 つを環境内に配置し, 実験を行った. ロボットは約 0.4m/s の平均速度で環境内を走行し, 使用した計算機の CPU はクアッドコア 2.8GHz, 各分散カメラの画像取得周期は 200ms とした. 図 7 は提案したパーティクルフィルタによる, ロボットと分散カメラに対する状態推定過程を示す. ここで, 使用したパーティクルの最大個数は 200,000 個, 最小個数は 20,000 個である. 図 7(a)に示すように, パーティクルの初期分布は手動で調整した状態を中心に広く分布され, 不確実性が大きい. しかし, 図 7(b), (c)に示すように, 移動ロボットが走行し, 時間が経過するにつれ, 各カメラの観測値による確率的な更新が観測モデルにより繰り返されるため, ロボットおよび分散カメ

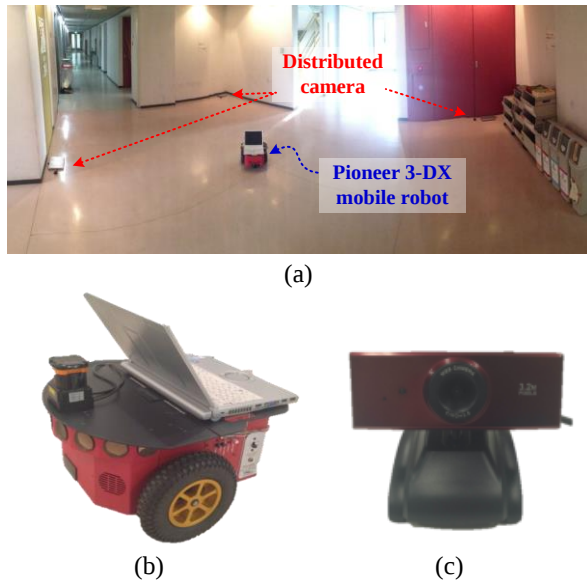


Fig. 6. Experimental setup : (a) experimental environment, (b) Pioneer 3-DX mobile robot, and (c) camera for distributed sensor.

ラのパーティクルが正しい状態に収束し、キャリブレーション(外部パラメータ同定)が自動的に行われた。

図8は図7の分散カメラ x_c^1 に対して、時間の経過に対する各状態変数の推定値の変化を示すグラフである。パーティクルフィルタのによる更新が約40回実行されると、各状態変数が正しい位置に収束し、外部パラメータ同定が完了された。

また、ロボットの自己位置推定性能に対しても、提案手法の有効性を検証するため、比較実験を行った。従来の地図情報に基づく自己位置推定結果と、提案手法により、地図情報なしで、分散カメラのキャリブレーションと同時にを行った自己位置推定結果を図9に示す。地図情報に基づく自己位置推定手法は、現在最も性能が良いと知られている MCL (Monte Carlo localization) と呼ばれるレーザー測距センサを用いたパーティクルフィルタベースの手法を使用した⁽¹⁰⁾。両方、少し精度の差はあるものの、デッドレコニングの

ような累積誤差は発生しなかった結果から、地図情報がない知能化空間でも十分安定的なロボットの運用が可能であると考えられる。

5. 結論

本論文では知能化空間でのサービスロボットの効率的な運用を実現することを目指し、移動ロボットの自己位置推定を行う同時に、知能化環境内に配置された分散カメラに対する外部パラメータを自動的にキャリブレーションする手法を確立した。実験結果により、以下のことが確認された。

- 知能化環境に配置された分散カメラの厳格な外部パラメータのキャリブレーション作業を事前に行わなくても、パーティクルフィルタに基づく提案手法を適用することで、ロボット自ら外部パラメータの同定が可能である。
- 地図情報がない環境でも、知能化空間に配置された分散カメラ情報を使用し、安定的な自己位置推定が可能である。
- 提案手法では上記の機能を同時に実行可能であるため、静的な地図情報に基づく既存のロボットの運用方法よりも、知能化空間に配置された分散カメラ情報を使用することで、動的な環境にも対応可能なより優れたロボットの運用が可能であると考えられる。

今後の展望としては、分散カメラに対する状態を3自由度から6自由度に拡張するとともに、それに対する精度を高めるため、位置情報だけではなく、姿勢情報の条件も考慮した、より完成度の高い分散カメラに対する観測モデルを定義する計画である。また、カメラの台数が増えるにつれ、要求されるパーティクルの最大個数も増加し、計算量も急激に増加する問題があるため、それに対応可能な手法を検討する計画である。

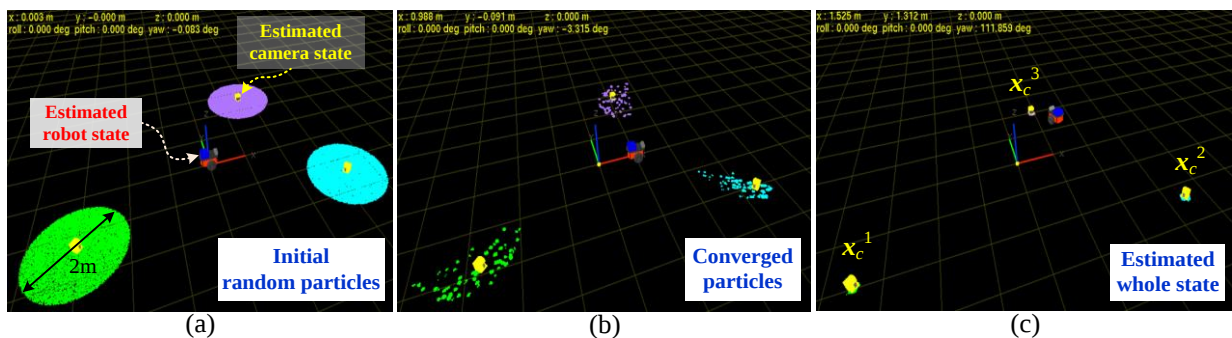


Fig. 7 Calibration result: (a) particle initialization, (b) converged particles, and (c) calibration completed.

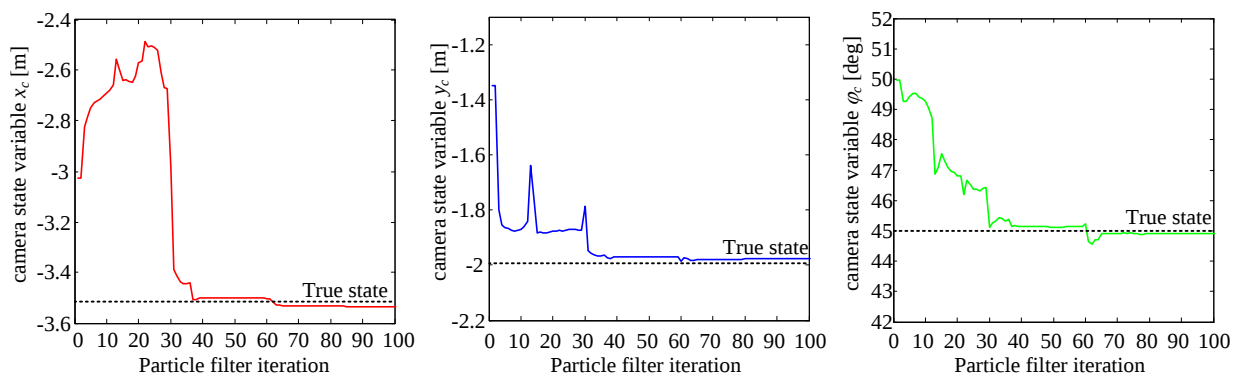


Fig. 8 Convergence process of the camera state variables (x_c , y_c , ϕ_c).

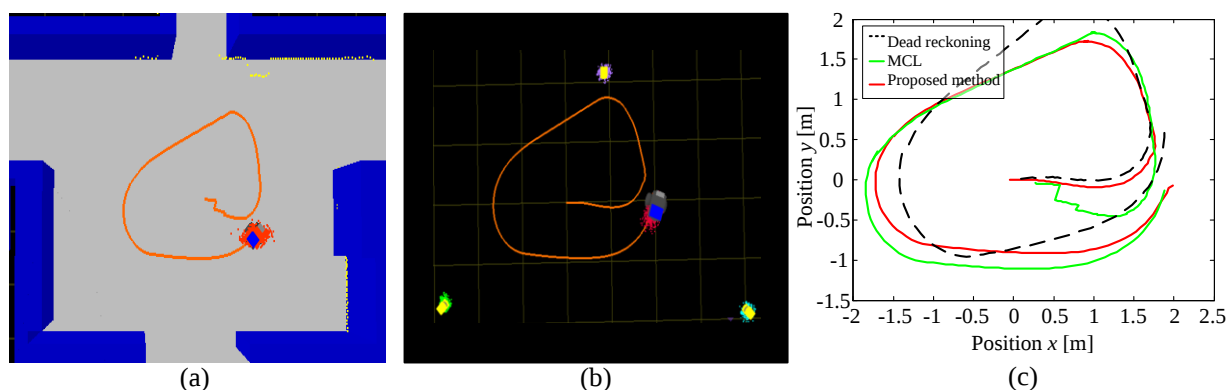


Fig. 9 Localization results: (a) MCL result using laser range data based on map information, (b) result of proposed method using distributed cameras, and (c) comparison result with dead-reckoning.

謝辞

本研究の一部は、総合科学技術・イノベーション会議により制度設計された革新的研究開発促進プログラム (ImPACT) 「タフ・ロボティクス・チャレンジ」の援助を受けた。

参考文献

- (1) Sebastian Thrun, Wolfram Burgard, and Dieter Fox, "Probabilistic Robotics", The MIT Press, (2005).
- (2) Hugh Durrant-Whyte and Tim Bailey, "Simultaneous localization and mapping: part I," *IEEE Robotics & Automation Magazine*, Vol. 13, No. 2, (2006), pp. 99-110.
- (3) Tim Bailey and Hugh Durrant-Whyte, "Simultaneous localization and mapping: part II," *IEEE Robotics & Automation Magazine*, Vol. 13, No. 3, (2006), pp. 108-117.
- (4) Joo-Ho Lee and Hideki Hashimoto, "Intelligent space-concept and contents," *Advanced Robotics*, Vol. 16, No. 3, (2002), pp. 265-280.
- (5) Tomomasa Sato, Yoshifumi Nishida, and Hiroshi Mizoguchi, "Robotic Room: Symbiosis with Human

through Behavior Media," *Robotics and Autonomous Systems*, Vol. 18, No. 3, (1996), pp. 185-194.

- (6) 水川 真, 山口 亨, "空間知能デザインに関する研究", 第7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講演論文集, (2009), pp. 534-537.
- (7) Eric M. Foxlin, "Generalized Architecture for Simultaneous Localization, Auto-Calibration, and Map-building", *Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, (2002), pp. 527-533.
- (8) Arnaud Doucet, Nando Freitas, and Neil Gordon, "Sequential Monte Carlo Methods in Practice", Springer, (2001).
- (9) Chris Stauffer and W.E.L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (1999), Vol. 2, pp. 246-252.
- (10) Dieter Fox, Wolfram Burgard, Frank Dellaert, and Sebastian Thrun, "Monte Carlo Localization: Efficient Position Estimation for Mobile Robots", *Proceeding of the National Conference on Artificial Intelligence*, (1999), pp. 343-349.