

自動重畳表示のための 人工マーカを利用した物体認識に基づく物体姿勢推定手法の開発

○伊部 直樹（東京大学） 山下 淳（東京大学） 浅間 一（東京大学）

1. 序論

原子力発電プラントは数百万にも及ぶ部品や機器で構成されており、建設や保守において非常に高度な管理が不可欠である [1]。その中でも、点検員がプラント内を巡回し、各機器の安全を確認する巡視点検が、安全管理の面から重要となっている。原子力発電プラントにおける巡視点検では、計器データのチェックだけでなく、プラント内の機械音や匂いなど、五感を使った異常の察知が求められ、これには点検員の熟練した技術や経験が必要となる。しかし、点検員の高齢化による熟練点検員の減少や非熟練者への技術伝承の困難さから、安全管理の水準の維持が困難となっている。このため、点検員への作業支援システムの開発が求められている。特に安全管理の水準維持のために、配管や機器の腐食や変形、循環液の漏れなど、プラント内の異常の予兆を非熟練点検員へ視覚的に呈示するシステムの開発が重要とされているため、情報呈示の際に配管や機器の認識と情報表示位置合わせを行うことが不可欠である。

情報呈示技術の1つに拡張現実感（Augmented Reality：AR）技術がある。これは、実世界内に仮想物体や情報をユーザの視点位置変更などに追従しながら重畳表示し、あたかも実世界内に存在するように呈示する技術であり、これまでエンターテインメント分野や医療分野など様々な応用がなされてきた。ARシステムを実現する上で、重畳表示したものがあたかも実世界に存在するとユーザが感じるために、実世界内の物体と仮想物体の位置合わせの研究が数多く行われてきた。

AR技術の代表的な手法として、人工マーカを利用したマーカベースAR[2][3]と、人工マーカを必要としないマーカレスAR[4][5]がある。前者は、環境内にあらかじめパターンが決まったマーカを設置し、それを基準としてカメラの姿勢や仮想物体の位置合わせを行う手法である。この手法は、安定的なカメラトラッキングが実現でき、マーカに物体情報を与えることで物体認識を行うことも可能であるが、高精度な位置合わせをするためには全てのマーカについて物体上の設置位置を事前に設定し、厳密に位置を合わせてマーカを設置する必要がある。一方、後者は環境内の特徴点や3次元点群データを取得し、Simultaneous Localization and Mapping (SLAM) を応用して位置合わせを行う手法である。この手法では、事前に作業を行う必要がなく、実世界の物体の形状に合わせた高精度な位置合わせが可能であるが、物体の認識を行うことは困難である。

本研究で想定する発電プラントは、短時間で変化す

る物体のない静的な大規模環境であり、物体認識と高精度位置合わせを行う情報呈示システムを従来のマーカベースARやマーカレスARを適用して実現することは困難である。そこで本研究では、両者の利点を組み合わせることで物体認識と高精度位置合わせを同時に実現するため、人工マーカによる物体認識と3次元計測により取得した点群データの位置合わせを利用した新たな物体姿勢推定手法を開発する。

2. 提案手法

2.1 概要

本研究では、RGB画像と3次元点群データを利用して物体の姿勢推定を行うため、RGB-Dセンサを用いて環境のセンシングを行う。提案手法では物体姿勢推定を行うため、図1のように、カメラ姿勢トラッキングと点群データの位置合わせによる物体姿勢推定を行う。カメラ姿勢トラッキングでは、センサで取得したRGB画像から推定対象物体に設置した人工マーカを検知し、マーカの3次元姿勢からカメラの3次元姿勢を推定する。一方、点群データの位置合わせは計算量が大きく毎フレーム行うことが困難であるため、物体の姿勢推定はキーフレームを設定してバッチ処理を行う。事前に物体の種類ごとに3次元形状モデルをデータベースに登録しておき、マーカにより認識された物体の形状モデルを物体姿勢推定に利用する。キーフレームごとに取得された点群データと認識された物体の形状モデルに対して位置合わせを行うことで、物体の3次元姿勢推定を行う。

2.2 人工マーカを利用したカメラ姿勢トラッキングとマーカ姿勢推定

本研究では静的な環境を想定するため、人工マーカを設置した物体の位置は変わらない、という前提を利用してカメラの3次元姿勢を推定する。まず、ARライブラリを利用して、大きさが既知であるマーカの画像座標系での位置からカメラ座標系におけるマーカの3次元姿勢を求める。次に、マーカのワールド座標系での3次元姿勢は変化しないという前提を利用して、フレーム間のカメラ座標系におけるマーカ姿勢の関係からカメラの運動を推定する。 t フレーム、 $(t+1)$ フレームでのカメラ座標系におけるマーカの3次元姿勢をそれぞれ $\mathbf{P}_t^m, \mathbf{P}_{t+1}^m$ とすると、

$$\mathbf{P}_{t+1}^m = \mathbf{M}_{t+1}^c \mathbf{P}_t^m, \quad (1)$$

という関係が成り立つ。ここで、 \mathbf{M}_{t+1}^c はカメラの運動に対応する行列である。ワールド座標系の原点をカ

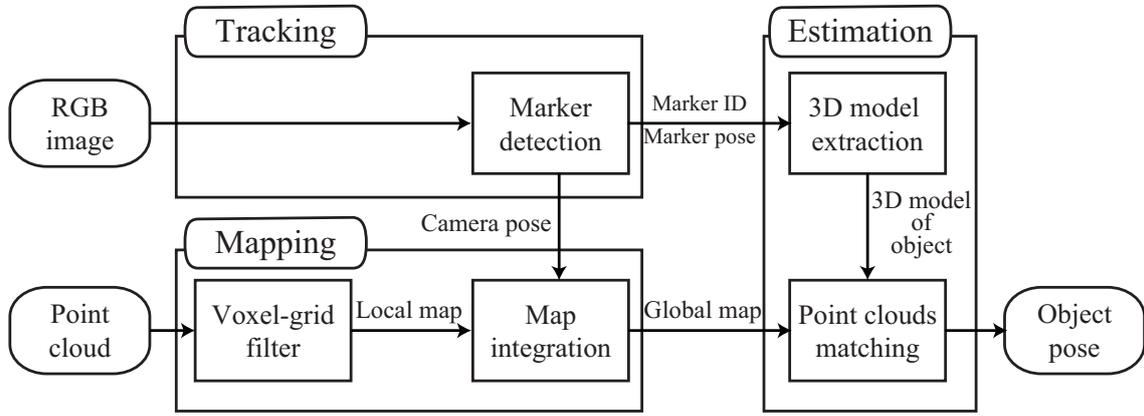


図1 提案手法

メラの初期位置とし、ワールド座標系の座標軸をカメラの初期姿勢でのカメラ座標軸と同じ方向に設定することで、ワールド座標系における t フレームでのカメラの姿勢 \mathbf{T}_t は、

$$\mathbf{T}_t = \prod_{\tau=1}^t \mathbf{M}_\tau \quad (2)$$

と推定される。

2.3 RGB-D センサによる 3次元点群の生成と点群データの統合

RGB-D センサから得られるのは、各画素に深度が格納されているデプス画像である。このデプス画像上における画素 (u, v) に対する深度 d から、後述の位置合わせに利用するために、カメラ座標系の 3次元点群を生成する。このとき、カメラ座標上での 3次元座標を (x, y, z) とすると、

$$x = \frac{u - c_x}{fd}, \quad (3)$$

$$y = \frac{v - c_y}{fd}, \quad (4)$$

$$z = d, \quad (5)$$

という関係に基づいて、3次元点群への変換を行う。ここで、 f はカメラの焦点距離、 (c_x, c_y) は主点を表す。このようにして生成した点群データを全て利用すると、計算負担が非常にかかる問題が生じる。そのため本研究では点群データに 3次元格子 (ボクセル) を導入し、ボクセルの中心点のみを点群データとして扱うボクセルグリッドフィルタ (図2) を用いる。これにより、均一に密度を減少させた点群データを利用し、計算量を大幅に減少させることが可能である。そして、フィルタリング処理によって生成されたカメラ座標系の 3次元点群 \mathbf{P}_t^{voxel} を以下の式のようにワールド座標系のカメラ姿勢 \mathbf{T}_t の逆行列を用いて、ワールド座標系の 3次元点群 ${}^G\mathbf{P}_t^{voxel}$ に変換する。

$${}^G\mathbf{P}_t^{voxel} = \mathbf{T}_t^{-1}\mathbf{P}_t^{voxel}, \quad (6)$$

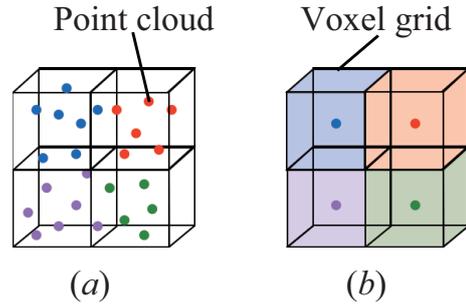


図2 ボクセルグリッドフィルタ：(a) 点群データから (b) ボクセルグリッドへフィルタリングを行い、点群を空間上に均一に分布させる

このように、キーフレームごとに生成された 3次元点群を統合することで、物体姿勢推定に利用する統合点群データを生成する。

2.4 点群位置合わせに基づく物体姿勢推定

推定対象物体上には物体の種類に応じて異なるマーカが設置されているため、マーカを認識することで物体の種類を特定する。データベースに事前に登録されている 3次元形状モデルの中から特定された物体の形状モデルを取得し、この形状モデルと前節で生成された統合点群データとの位置合わせを行う。本研究では、点群の位置合わせとして有効な手法とされる Iterative closest point (ICP) マッチング [6] を用いる。しかし、統合点群データと直接、ICP マッチングを行うと計算量が非常に大きくなり、物体姿勢推定を行うことが困難になる問題が生じる。そこで、マーカ位置を利用して、マッチングを行う点群を限定することで、計算時間を大幅に向上させる。また、マーカに基づいて実際の物体位置の近傍にマッチングの初期位置を設定することで ICP マッチングにおける繰り返し回数も減少すると考えられる。

具体的な位置合わせ手法について、以下で述べる。まず、RGB 画像からマーカを検知し、カメラ座標のマーカ姿勢 \mathbf{P}_t^m を取得する。このマーカ姿勢を以下の式のよ



図3 ASUS Xtion Pro Live

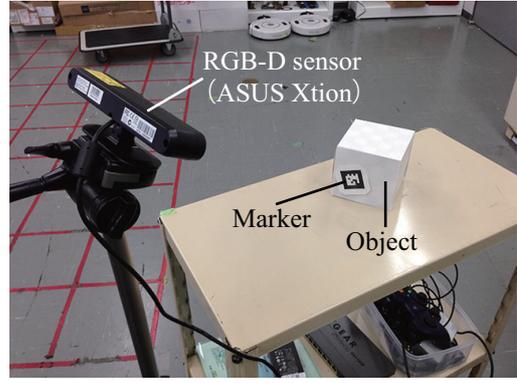


図4 実験環境

表1 Xtion PRO LIVE の主な仕様

Sensor range	0.8m ~ 3.5m
Field of view	58 deg horizontal 45 deg vertical
Frame rate	60 FPS (320 × 240)
Size	180 × 35 × 50 mm
Data interface	USB 2.0

うにカメラ姿勢を表す行列 \mathbf{T}_t の逆行列を用いて、ワールド座標系のマーカー姿勢 ${}^G\mathbf{P}_t^m$ に変換する、

$${}^G\mathbf{P}_t^m = \mathbf{T}_t^{-1}\mathbf{P}_t^m, \quad (7)$$

このようにして求められたマーカー姿勢 ${}^G\mathbf{P}_t^m$ から、ワールド座標系におけるマーカー位置 ${}^G\mathbf{x}_t^m$ を得る。

次に、3次元形状モデルとの位置合わせを行う点群を統合点群データから抽出する。推定対象物体に設置されたマーカーの近傍から抽出するために、ワールド座標系のマーカー位置 ${}^G\mathbf{x}_t^m$ を中心とした半径 r の球内に含まれる点群を抽出する。また、抽出された点群と ICP マッチングを行うために、推定対象物体の3次元形状モデルから点群を生成する。本研究では、形状モデルのメッシュ上から等間隔に点群をサンプリングして点群を生成する。そして、生成された2つの点群に対して ICP マッチングを適用し、3次元形状モデルと統合点群データの位置合わせを行う。このとき求まる3次元形状モデルの姿勢を、物体の姿勢として推定を行う。

3. 物体姿勢推定に基づく重畳表示実験

3.1 実験設定

本研究では提案した手法の実装のため、RGB-D センサとして図3に示す ASUS 社の Xtion PRO LIVE と公開されている OpenNI ライブラリを使用する。Xtion PRO LIVE の主な仕様を表1に示す。また、AR ライブラリとして、トラッキングの安定性とロバスト性の高さからオープンソースである Chilitags ライブラリ [3] を利用する。

本実験では、推定対象物体として図4のように表面に Chilitags マーカーを設置した立方体を使用する。立方

体の1辺は10.0cmであり、事前に3次元形状モデルを生成する。各立方体を RGB-D センサでセンシングして提案手法を適用することで、対象物体の姿勢が正しく推定し、センサで取得した RGB 画像に物体の3次元形状モデルを重畳表示する、という試行を行う。物体上のマーカーの位置は試行ごとに変化させ、試行1ではマーカーの各辺が立方体の各辺と平行になるように設置し、試行2では立方体面の法線回りに回転させて設置する。

3.2 実験結果

図5に、試行1の実験結果を示す。図5(a)は、試行中のある1フレームの RGB 画像であり、これに立方体の3次元形状モデルを重畳した画像が図5(b)である。マーカーの位置だけでなく、3次元点群を利用して立方体の形状に基づいた姿勢推定を行うことにより、マーカーの位置が不定であっても重畳表示が可能となった。

また、図6に、試行1の実験結果を示す。図6(a)は、試行中のある1フレームの RGB 画像であり、これに立方体の3次元形状モデルを重畳した画像が図6(b)である。試行1と異なる位置・姿勢でマーカーが設置された状態のときも、立方体の形状に合わせて重畳表示が行われることが示された。

4. 結論

本研究では、人工マーカーによる物体認識と3次元点群データの位置合わせを組み合わせて利用した物体姿勢推定手法を提案した。提案手法では、RGB-D センサにより RGB 画像と3次元点群を取得し、RGB 画像からは人工マーカーの検知を行い、物体認識を行った。また、人工マーカー近傍の3次元点群と事前に与えられた物体の3次元形状モデルに ICP マッチングを適用することで、物体表面上の任意の位置にマーカーを設置した状態でも位置合わせを行うことが可能であると示した。

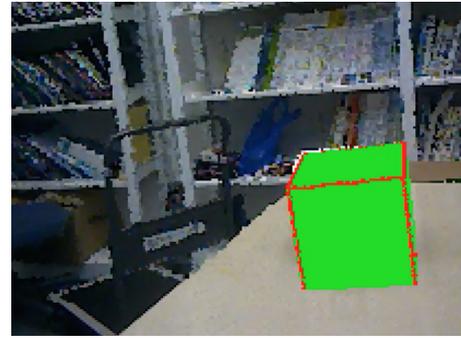
今後の課題としては、精度や計算時間の向上、プラント内での適用に向けて、過去画像との差分の取り扱いについての検討などが考えられる。

参考文献

- [1] 恩田公治, 室啓朗, 若林英祐: “原子力プラントへの RFID 高度応用システムの開発-生産管理技術, 建設技術, 保全



(a) RGB 画像

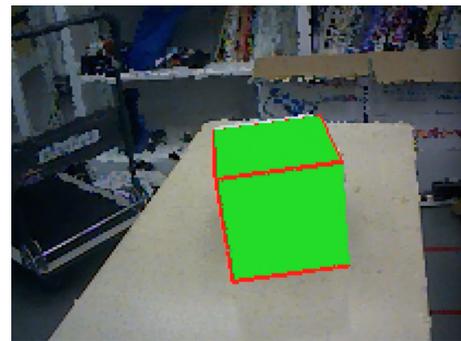


(b) 重畳表示

図 5 試行 1



(a) RGB 画像



(b) 重畳表示

図 6 試行 2

技術への展開 (特集 電力・エネルギー分野の最新開発技術)", 日立評論, vol.90, no.2, pp.156-161, 2008.

- [2] 加藤博一: "拡張現実感システム構築ツール ARToolKit の開発", 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, vol.101, no.652, pp.79-86, 2002.
- [3] Q. Bonnard, S. Lemaignan, G. Zufferey, A. Mazzei, S. Cuendet, N. Li, P. Dillenbourg: "Chilhtags2: Robust Fiducial Markers for Augmented Reality and Robotics," CHILI, EPFL, Switzerland, <http://chili.epfl.ch/software>, 2013.
- [4] G. Klein and D. Murray: "Parallel Tracking and Mapping for Small AR Workspaces," *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR2007)*, pp.225-234, 2007.
- [5] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon: "KinectFusion: Real-time Dense Surface Mapping and Tracking," *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR2011)*, pp.127-136, 2011.
- [6] S. Rusinkiewicz and M. Levoy: "Efficient Variants of the ICP Algorithm," *Proceeding of the International Conference on 3D Digital Imaging and Modeling*, pp.145-152, 2001.