

機械学習による産業用ロボットの自律的な動作生成に関する試験研究事業

平間 毅* 前島 崇弘** 飯村 修志*

1. はじめに

政府の「ロボット新戦略」策定に伴い、茨城県においても平成28年度中小企業IoT等自動化技術導入促進事業、平成29～30年度IoT・ロボット研究会、平成31年度IoT・ロボット分科会等を通じて、県内中小企業のロボット導入を促進するための取り組みを進めてきた。

また、平成28年経済センサスー活動調査産業別集計（製造業・概要版）（経済産業省）によると、茨城県は製造業における従業者数が26万1,726人で全国8位、製造品出荷額等が12兆376億円で全国8位に位置する工業県であり、ロボット活用による自動化をニーズに持つ企業も多い。

一般的に産業用ロボットを生産現場へ導入するためには、作業ごとにティーチングペダントを用いたロボットへの動作教示作業を行う必要があり、またロボットとカメラを組み合わせる場合には画像処理アルゴリズム等の構築が必要となる。

多品種変量生産が要求される中小企業にとって、ロボットシステム構築を都度システムインテグレーターに外注するのはコストや時間的に困難であり、できる限り自社での変更対応が求められる中、教示人材の不足やロボット操作の難しさが中小企業におけるロボット導入や活用を妨げる一つの要因¹⁾となっている。

このような背景のもと、当センターでは産業用ロボットを用いた自動化システムを構築する際の教示作業の負担を軽減するため、機械学習による作業対象ワークの認識及び認識結果をもとにした産業用ロボットの動作生成に関する研究開発に取り組んでいる。

2. 目的

本研究開発では、データ駆動型アプローチである機械学習を用いた作業対象ワークの認識及び作業対象ワークの位置姿勢に応じた適切なロボットの動作生成を目的としている。今年度はRGBDカメラで取得したカラー画像から機械学習技術により作業対象ワークの領域を矩形で検出し、検出した領域内の点群データ及び事前に用意した作業対象ワークのモデルデータを用いて作業対象ワークの位置姿勢を推定する手法について研究開発を行った。

3. 研究内容

3.1 実験環境

本研究開発の実験環境は、図1のとおり垂直多関節ロボット(Torobo Arm, 東京ロボティクス株式会社)、RGBDカメラ(RealSense D415^{2,3)}, Intel), ロボット用作業機で構成されており、既報⁴⁾のとおりRGBDカメラと垂直多関節ロボットの相対位置姿勢は既知となっている。

また、本研究開発で対象とする作業対象ワークは四角の穴が開いた四角柱(図2)とした。



図1 実験環境

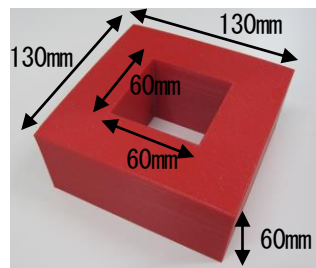


図2 作業対象ワーク

3.2 機械学習による作業対象ワークの認識

本研究開発では、ディープラーニングベースの一般物体検出アルゴリズムであるSSD(Single Shot multibox Detector)⁵⁾を適用することで、カラー画像中に写った作業対象ワークを矩形領域で検出することを検討した。

3.2.1 学習用データの作成

まず、作業対象ワークを画像の中心付近に写るようセットし、姿勢を変えながら45枚の画像を撮影した。次に、作業対象ワークに相当するピクセルの色相値をランダムに変更することで、色味の異なる作業対象ワークを表現する画像を作成した。さらに、PCA Color Augmentation⁶⁾と呼ばれる画像の色分布を考慮した色の加減を行うため、画像中の各ピクセルのRGBを3次元ベクトルの集合とみなして主成分分析を行い、平均0、標準偏差0.1のガウス分布から生成したノイズを固有ベクトル方向に付加することで合計1,890枚の学習用データを作成した。また、作成した学習用データのうち、元画像及び元画像にPCA Color Augmentation処理を行った画像中に写った作業対象ワークに対して、位置情報及び「作業対象ワーク」のラベルを付けた。

3.2.2 ディープラーニングによる物体検出

本研究開発では、ベースとなる検出モデルのフレームワークとしてSSDを用いた。SSDは物体候補領域と物体クラスを同時に推定するアルゴリズムであり、複

*研究推進グループ **IT・マテリアルグループ

数解像度の特徴マップを用いて、様々なスケールの物体を検出することができる。予測レイヤーではデフォルトボックスと呼ばれる矩形に対するオフセット及び物体のクラス尤度が特徴マップとして出力され、このクラス尤度を用いて物体のクラス識別を行う。また後処理として Non-maximum Suppression 処理を行うことで重複した検出矩形が削除される。

SSD の特徴抽出器には ResNet-50⁷⁾を用いた。ResNet は Microsoft research によって提案された画像のクラス分類を行うためのニューラルネットワークモデルであり、残差ブロックと呼ばれる機構を導入することにより深い層の学習が可能となっている。

また、本研究開発では COCO データセット⁸⁾を用いて事前に学習されたモデルをファインチューニングすることによりモデルの再学習を行った。学習時に用いた設定は表 1 のとおりである。

表 1 学習時の最適化手法とハイパーパラメータ設定

| | |
|--------|---|
| 最適化手法 | モーメンタム SGD |
| バッチサイズ | 32 |
| 学習率 | コサイン減衰 基本学習率 0.1 ウォームアップ学習率 0.001 ウォームアップステップ 2000 |
| 学習回数 | 10,000 回 |

3.2.3 学習したモデルの検証結果

学習したモデルの検証を行うため、赤の作業対象ワーク及び同一形状で色の異なる青のワークを黒、白、緑の作業機の中心及び四隅付近にランダムな姿勢で配置した状態で撮影を行い、60 枚の画像を作成した。これらの画像から作業対象ワークの検出を行った結果、青のワークを誤検出することなく、赤い作業対象ワークのすべて (30/30) を検出することができた。赤い作業対象ワークの検出結果の一部を図 3 に、青いワークを撮影した画像の一部を図 4 に示す。

また、作業対象ワークと同色の赤で異なる形状のワーク (図 5) を黒の作業機の中心及び四隅付近の 5 か所に配置した状態で撮影した画像で検証を行ったところ、これらの画像についても誤検出はみられなかった。検出結果の一部を図 6 に示す。

3.3 点群データを用いた作業対象ワークの位置姿勢推定

点群同士の位置合わせを行うために、事前準備として、作業対象ワークのモデル形状データを Cloud Compare⁹⁾を用いて点群データへ変換した。点群データへの変換結果は図 7 のとおりである。

なお、本研究開発ではモデル点群データ及び RGBD カメラで取得した点群データの位置合わせを行い、作業対象ワークの位置姿勢推定を行うため、図 8 に示す一連の処理フローを検討した。

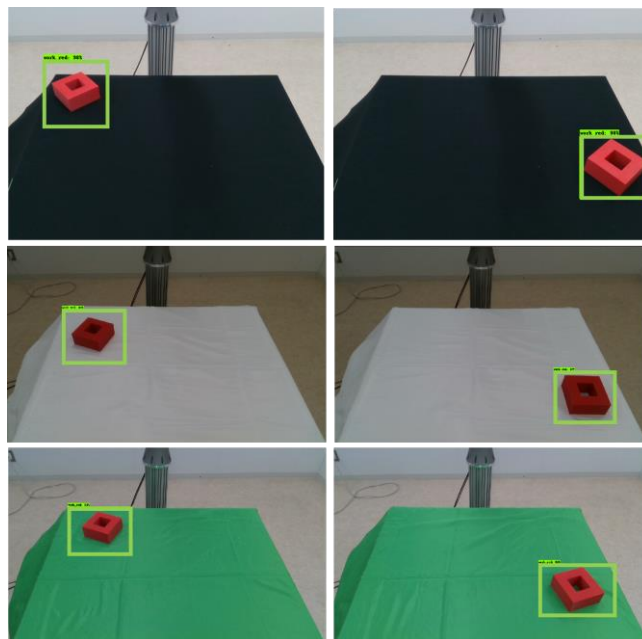


図 3 作業対象ワークの検出結果
(上段：黒の作業機の左奥，右手前に配置
中段：白の作業機の左奥，右手前に配置
下段：緑の作業機の左奥，右手前に配置)
緑の矩形は作業対象ワークの検出結果



図 4 青ワークの配置例と検出結果

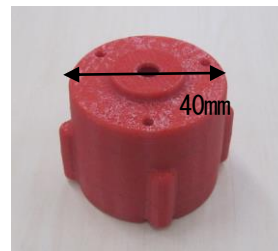


図 5 検証に用いた同色異形状ワーク

3.3.1 対象領域の点群データ抽出

3.2 で認識した作業対象ワークの矩形領域をもとに、対応する点群データの領域を抽出した。

RGBD カメラ内のカラーカメラとデプスカメラの相対的な位置姿勢関係が既知であるため、カラー画像の矩形領域に対応するデプス画像の領域を算出することができる。また、デプス画像及びカメラパラメータから点群データを構築することができるため、結果として対象領域に相当する点群データを抽出することができる。

また、既報⁴⁾から作業機に対応する平面が既知であることから、前処理として、抽出した点群データのうち作業機に相当する点群の除去を行った (図 9)。

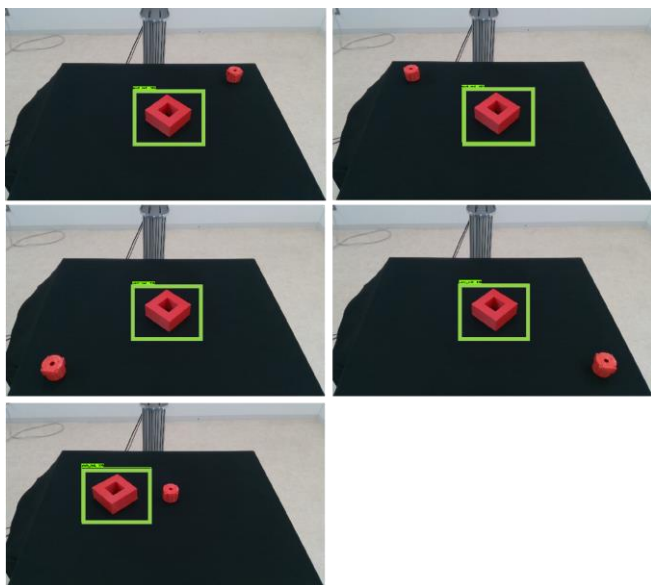


図 6 同色異形状ワークの検出結果
 矩形は作業対象ワークの検出結果

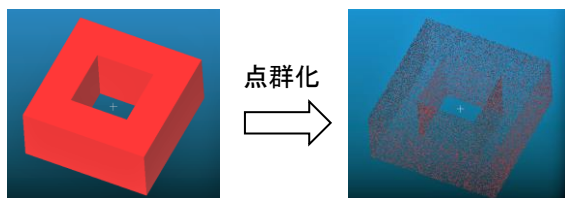


図 7 モデルの点群化
 (左:モデル形状データ, 右:点群データ)

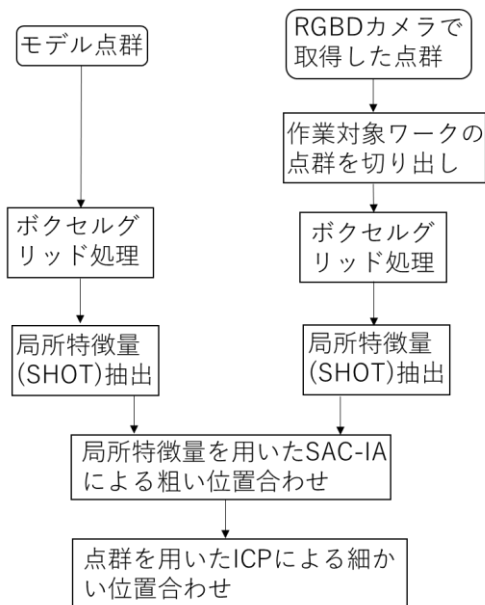


図 8 位置合わせ処理のフロー

3.3.2 粗い位置合わせ

モデル点群及び切り出した点群の位置合わせを行うため、まずそれぞれの点群に対して $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ サイズでボクセルグリッド処理を行った。ボクセルグリッド処理は、空間をボクセルグリッドで区切り、各グリッド内の点群をその重心で近似する処理であり点群の数を削減することができるため、その後の点群処理の計算コストを減らすことができる。また、モデル点群の上半分のみを用い位置合わせを行った。

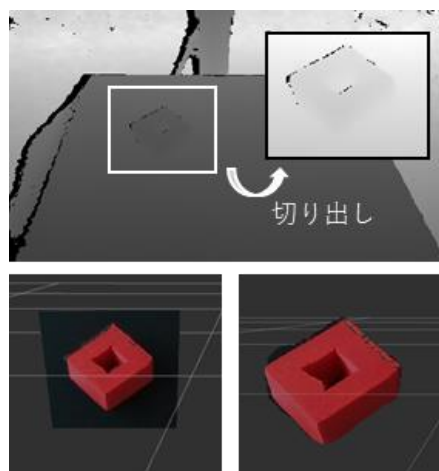


図 9 作業対象物に対応する点群の抽出
 (上:デプス画像の切り出し
 左下:切り出し領域の点群, 右下:作業機除去後)

次に、点群の局所的な特徴量を計算するためのキーポイントとして、 $10\text{mm} \times 10\text{mm} \times 10\text{mm}$ 間隔の一樣キーポイントを抽出し、各キーポイント(図 10)において SHOT 特徴量¹⁰⁾を算出した。SHOT 特徴量は各点 P で再現性と唯一性を持つローカル参照座標系を設定し、周辺の局所領域を $8 \times 2 \times 2$ に分割した各領域において、P の法線ベクトルと各領域に含まれる点の法線ベクトルの内積をヒストグラム化した特徴量である。この時、法線ベクトル及び SHOT 特徴量を算出する領域は半径 25mm の球とした。

この局所特徴量をもとに、SAC-IA (Sample Consensus Initial Alignment) を用いて点群同士の粗い位置合わせを行った。以下、本研究で用いた SAC-IA の大まかな処理の流れを示す。

1. 各特徴点の 2 点間の距離が 30mm より大きくなるようにモデル点群から 4 点をサンプリングする。
2. モデル点群からサンプリングされた各特徴点に対して、切り出した点群から近い特徴量を持つ特徴点を複数探索し、その中からランダムにひとつを選んで対応点とする。
3. 選択された対応点のペアから剛体変換を求め、モデル点群と切り出した点群の間の誤差を計算する。
4. 以上の処理を繰り返すことで、最も誤差が小さくなる剛体変換が粗い位置合わせの結果として得られる。

ここで、誤差の定義はモデル点群から 20mm 以下の距離にある切り出し点群の数とした。

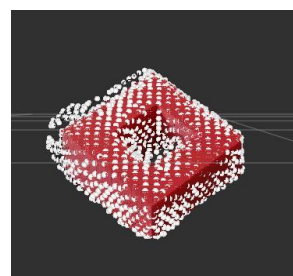


図 10 切り出し点群から抽出したキーポイント
 (白い点がキーポイントを表す)

3.3.3 細かい位置合わせ

3.3.2 で得た粗い位置合わせの結果をもとに、ICP(Iterative Closest Point)処理^{11,12)}により、細かい位置合わせを行った。ICP は点群の対応付け、剛体変換行列の導出と点群への適用、評価を繰り返し行うことで、最適な剛体変換行列を求めていく手法である。しかし、ICP では近傍探索により対応点を決定することから局所解に陥る可能性がある。そこで、本研究では局所解を避けるため、粗い位置合わせの結果を初期条件として用いた。また、位置合わせの精度を向上させるためのSAC-IAによる粗い位置合わせ及びICPによる細かい位置合わせを複数回繰り返し、対応点間の二乗誤差の和が最も小さくなる結果を最終的な位置姿勢推定結果とすることとした。

粗い位置合わせ及び細かい位置合わせによる結果の検証を行うため、作業対象ワークの配置2パターンにおいて、位置姿勢推定結果からモデルデータの位置姿勢を変換し、切り出し点群に重畳させることで可視化を行った。可視化の結果を図11に示す。ここで、配置1は作業対象ワークの頂点がRGBDカメラの方向を向く配置、配置2は作業対象ワークの辺がRGBDカメラの方向を向く配置とした。また、それぞれの配置における位置合わせ後の二乗誤差の和は約 5.2×10^{-5} 、約 4.3×10^{-5} であり、可視化の結果、本処理における位置合わせが概ね成功していることを確認できた。

なお、本結果及びRGBDカメラと垂直多関節ロボットの相対位置姿勢が既知であることから、垂直多関節ロボットから見た作業対象ワークの位置姿勢を算出することができ、引き続きロボットの動作生成につなげることができる。

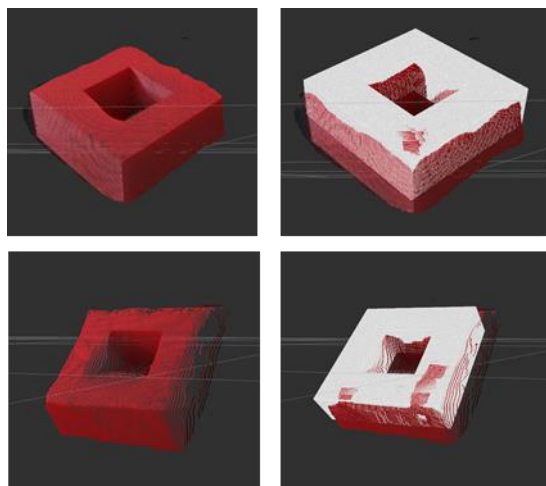


図11 細かな位置合わせの結果
 (上段左：配置1での取得点群
 上段右：配置1での位置合わせ結果
 下段左：配置2での取得点群
 下段右：配置2での位置合わせ結果)

4. まとめ

作業対象ワークの位置姿勢に応じた適切なロボット動作生成につなげるため、データ駆動型アプローチである機械学習を用いた作業対象ワークの認識及びデブスカメラで取得した点群データを用いた位置姿勢推定

を行った。

- ディープラーニングベースの物体検出アルゴリズムであるSSDを用いた作業対象ワークの領域検出モデルを構築し、ワークの色、形状、位置、背景色の条件を変更し検証実験を行った結果、誤検出なく作業対象ワークのみが検出できることを確認した。
- RGBDカメラで取得した点群データから作業対象ワークの領域に相当する点群データを切り出し、SHOT局所特微量及びSAC-IAアルゴリズムを用いた粗い位置合わせ、ICPアルゴリズムを用いた細かい位置合わせの一連の処理によりモデル点群データとの位置合わせを行い、結果を可視化することで、位置合わせ処理が概ね成功していることを確認できた。
- 位置合わせの結果及びRGBDカメラと垂直多関節ロボットの相対位置姿勢に関する情報から、垂直多関節ロボットのベース座標系から見た作業対象ワークの位置姿勢を算出することができた。

5. 今後の課題

引き続き、作業対象ワークの位置姿勢推定結果をもとに、シミュレータ上で垂直多関節ロボットによる挿入作業の動作計画を検討するとともに、実機を用いた検証に取り組む。

6. 参考文献等

- 1) 近畿経済産業局,平成27年度産業用ロボットの分野展開における導入阻害要因調査事業報告書,2016
- 2) <https://click.intel.com/intelr-realsensetm-depth-camera-d415.html>
- 3) <https://github.com/intel-ros/realsense/releases>
- 4) 平間毅他,茨城県産業技術イノベーションセンター研究報告第47号,2019
- 5) Wei Liu, et al., "SSD: Single Shot Multibox Detector", Computer Vision-ECCV2016pp21-37, 2016.
- 6) Krizhevsky, et al., "ImageNet classification with deep convolutional neural networks" Advances in Neural Information Processing Systems, 2012.
- 7) He, et al., "Deep residual learning for image recognition." In proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
- 8) Lin, T-Y. et al., Microsoft COCO: Common Objects in Context, arXiv:1405. 0312, 2014.
- 9) <https://www.danielgm.net/cc/>
- 10) Federico Tombari, et al., "unique signatures of histograms for local surface description", ECCV, 2010 (Crete Island, Greece)
- 11) Chen, Y., & Medioni, G., "Object modeling by registration of multiple range images, Image and vision computing, 10(3), 145-155, 1992
- 12) Rusinkiewicz, S., & Levoy, M., "Efficient variants of the ICP algorithm" In 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on (pp. 145-152). IEEE