

# ホームサービスロボットによる物体認識のための アクティブセンシング戦略

## アクティブセンシング戦略のためのデータセット構築

○坂巻新 Luis Contreras 武藤ゆみ子 水地良明 岡田浩之 (玉川大学)

### 1. はじめに

日常生活環境で人を支援する自律型ロボットの開発が盛んに行なわれている。そのようなロボットに求められる重要な機能の一つが物体認識である。これまで、ロボカップ@ホームや WRS などのロボット競技会において、多様で煩雑な環境下で利用可能な物体認識手法が提案されてきた [1-9]。しかし、複雑な環境下では、物体の置かれる場所や物体とカメラとの位置関係に依存して物体認識の精度が低下する問題がある。例えば、机の下に物体が置かれている場合、照明変化によって認識精度が低下する。我々の先行研究 [10] では、図 1(a), (b) のような視点で物体を認識を行うよりも、図 1(c) のように、ロボットが物体を把持し、カメラと物体の位置関係を一定にすることで認識精度が向上する結果を示した。このように、ロボットが物体を認識しやすい状況を作るアクティブセンシング戦略を取ることで、従来では認識することが難しかった状況において認識精度が向上することが期待できる。

物体認識を目的とした画像データセットの構築手法 [11,12] はこれまでも提案されている。しかしながら、これらは特定の実験環境や条件設定に焦点をあてたものであり、アクティブセンシング戦略に基づく物体認識に応用することは難しい。そこで、本研究ではアクティブセンシング戦略に基づく物体認識に必要なデータセットの種類について議論し、それらの構築方法を提案する。

### 2. 関連研究

ロボカップ@ホームなどの生活環境を対象とした物体認識では、YOLO [13] など、Convolutional Neural Network(CNN) をベースとした物体認識手法が主流となっている。YOLO はロボカップ@ホーム 2019 世界大会などで、11 チーム中 6 チームが使用している [1-9]。

我々の先行研究 [10] では、図 1 に示すように、対象から離れている距離から広範囲を見渡す random-distance、近い距離で物体を注視する close-distance、物体を検出し把持したあとに至近距離で認識する hand-distance の 3 つの観測戦略を挙げている。Random-distance では家具の前など大まかな位置から観測するため、全体を見渡すことができるメリットはあるが、小さい物体に対する認識精度が低下する。Close-distance では小さい物体でも安定して認識できるが、物体の見落としが起きやすい。Hand-distance ではロボットは最初に物体を検出し、把持したあとマニピュレータをロボットの上部カメラに近づけ固定距離で認識する。その結果、他の 2 条件よりも認識精度の向上するという結果が得られた。しかし、照明、物体のスケール、方

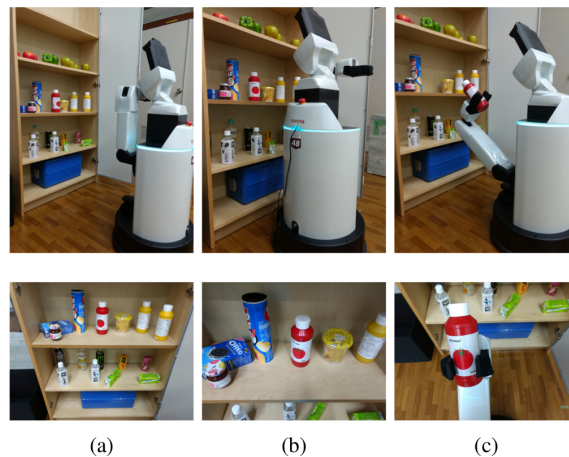


図 1: ロボットによる物体認識での観測戦略。

(a) Random-distance, (b) Close-distance, (c) Hand-distance.

向の変化、オクルージョンが大きい条件では依然として低い性能しか得られないことがわかった。

James ら [11] は、マニピュレーションシステムのためのテストプラットフォームを提案している。これはシミュレーターの中で固定設置されたロボットアーム、上部の RGBD カメラと、マニピュレータの中に埋め込まれた単眼カメラを使用しており、100 種類のタスクと、それぞれの視覚的な観測情報、ロボットの関節情報や速度情報を提供している。また、小野ら [12] は、TOYOTA 社製 Human Support Robot (HSR) を使用した物体認識を目的として、シミュレータを活用することで効率的にデータセットを構築する手法を提案している。このように、いくつかのタスクやベンチマーク手法が提案されているが、これらの殆どはタスクに特化したものであり、実験環境や設定に大きく依存している。これらに対し、本研究では、家庭内の複雑な環境下での汎用サービスロボットによる物体認識の精度向上を目的としたデータセットの構築に焦点を当てる。

### 3. アクティブセンシングのための画像データセット

CNNs をベースとした認識モデルの構築には数十万から数百万枚オーダーの画像データが必要となる。しかしながら、アクティブセンシング戦略に基づいて物体を認識する場合の画像を大量に取得するのは困難である。このような問題に対して、効率的に収集可能な画像データを用いて認識モデルを構築した後、目的に合

わせた追加学習のための画像データを用いてファインチューニングや転移学習し、精度向上を図るアプローチが一般的となっている。したがって、このような追加学習のための画像データセットを構築することも重要である。

アクティブセンシング戦略に基づく物体認識を実現するためには、上記の観点に基づいて構築した複数のデータセットを用い、適切な組み合わせを見つけ出す必要がある。そこで、YCB データセット [14–16] に含まれる 21 種類の物体を対象とした 5 種類の学習データセットを構築した。以下では、構築した各データセットについて詳述する。

### 3.1 3DCG を用いて構築した仮想画像データセット

膨大な画像データを収集する方法として、3DCG を活用する。ここでは、YCB データセットにて公開されている 3D オブジェクトモデルを用いて画像データを生成する。3D モデルの姿勢をそれぞれの各回転角について、一定間隔で変化させ、オブジェクトごとに全方位から撮影した 100 万枚の画像を記録した。この方法では、3D モデルを用いることで、人の手を借りずに自動で物体の配置から画像データの取得まで行うことができる。また、光源やカメラのパラメータなどを変えられる利点がある。一方で、物体のテクスチャが現実の物体とは違うため、実機ロボットを用いた認識では十分な精度が得られない可能性がある。

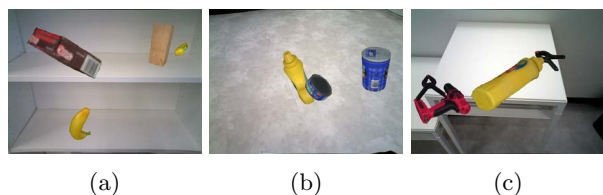


図 2: 3D モデルを用いて生成した仮想画像データ

### 3.2 ロボットシミュレータを用いて構築した仮想画像データセット

ロボットシミュレータを活用することで、物体の姿勢に制約を加えた条件で仮想画像データを収集する。図 3 に仮想ロボットを用いた物体画像データの記録の様子を示す。まず、仮想ロボットの RGBD カメラを用いて平面上に置かれた物体を観測する。次に、距離情報を用いて平面検出を行い、物体以外の背景領域を削除する。この方法では、実際にはありえない物体の姿勢で観測されたノイズとなりうる画像データを除外できる。一方で、ロボットの移動や距離情報に基づく物体領域のセグメンテーションを要するため、3DCG を利用した場合と比較して、膨大な画像データの生成のために時間がかかってしまう。

### 3.3 ロボットアームを用いて構築した実画像データセット

図 4 のように、ターンテーブルとロボットアームを用いて実画像データを収集する。まず、ターンテーブルの上で回転している物体を、ロボットアームの先端に取り付けた RGBD カメラを用いて撮影する。次に、距

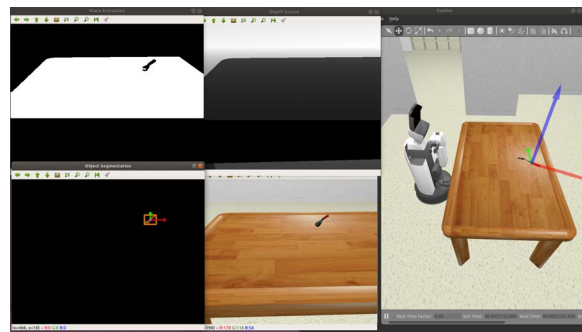


図 3: ロボットシミュレータ用いた画像データの取得

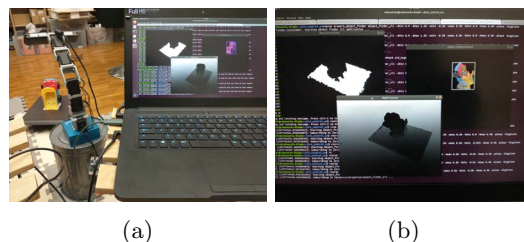


図 4: ターンテーブルとロボットアームを用いた実画像データの収集。(a) 撮影システムの構成、(b) 撮影された画像

離情報を用いて平面検出を行い、物体以外の背景領域を削除する。ターンテーブルが一周するごとに、ロボットアームを動かし、カメラの位置を変えながら、様々な角度から物体を撮影する。この方法では、実機ロボットでの認識を想定した実画像データを取得できるが、CG を利用した方法と比べて、画像データの取得に時間がかかる。

### 3.4 実機ロボットアームを用いた実画像データセット



図 5: 実機ロボットを用いて取得した把持物体画像

実機ロボットを用い、アクティブセンシングを目的とした物体の把持画像データを収集する。この方法では、ロボットアームで物体を把持した後、カメラの前に物体を移動させ、画像データを取得する。そのため、図 5 のように、物体とカメラの位置関係、認識する物体の姿勢に制約を設けた実画像データセットを構築することができる。この方法では実機ロボットを用いるため、膨大な画像データを収集することは難しいが、ファインチューニングなどの追加学習に利用するためのデータセットを構築できる。

### 3.5 仮想ロボットアームを用いた仮想画像データセット

仮想ロボットを用いて把持物体画像データを取得する。図 6 に、仮想ロボットと物体の 3D モデルを用い



て取得した把持物体画像を示す。この方法では、画像を自動生成できるため、実ロボットアームで生成する場合よりも、多くの画像データを効率的に収集できる利点がある。一方で、実機ロボットでの認識時に、実画像を利用して学習した場合よりも精度が低下する恐れがある。

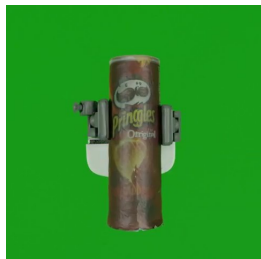


図 6: 仮想ロボットを用いて取得した把持物体画像

#### 4. 性能評価のためのテストデータセット

構築した画像データを組み合わせて学習した物体認識モデルの性能を評価、比較するためには、テストデータセットも必要となる。そこで、World Robot Summit<sup>1</sup>のTidy upタスクを想定した環境を用意し、実機ロボットを用いてテストデータセットの収集を行う。ここではロボットとしてHSRを用い、テストデータとなるRGBD画像、tf [17]、地図情報などをrosvbag<sup>2</sup>に記録する。

図7にロボットの動作を示す。ここでは、図7(a)にあるようにWRSのルールに則って、物体を設置する。その後、ロボットが対象物体を常に観測しながら物体の周りを動き回る。そこから物体を把持し、図7(b)のようにカメラと把持物体の距離に着目した位置にマニピュレータを動かす。

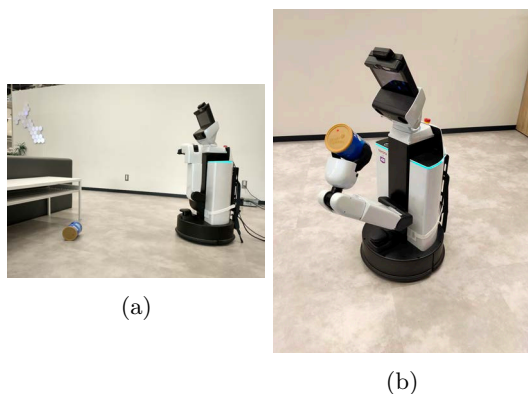


図 7: 観測時のロボットの動作

図8は物体の配置条件を示す。ここでは図8(a), (b)の2つの条件でrosvbagデータを記録する。1つめの条件では、対象物体のみを配置する。2つめの条件はさらに対象物体の周りに他の物体を配置した条件となっている。各条件で前述のようにロボットを動作させ、rosvbagを記録する。このデータセットは図1に示したすべての



図 8: 物体の配置条件

観測条件で取得した画像を含むものとなっている。このテストデータを利用することで、既存の手法を含め、構築した物体認識手法の精度を評価、比較することができる。

#### 5. 結論

本研究ではアクティブセンシング戦略に基づく物体認識手法の実現に必要なデータセットについて議論した。データの収集効率、ロボットによる物体の観測視点、画像のリアリティに着目し、5種類の学習データセットについて提案した。また、学習データセットを用いて学習した物体認識機の性能を評価、比較するために必要なテストデータセットについて提案した。

ここで構築したデータセットは、本研究のみならず開発者がシチュエーションや条件に合わせて様々な比較、検証ができるベンチマークデータセットとしても活用できるため、オープンデータセットとして公開する予定である<sup>3</sup>。

今後は、今回作成した学習データセットを活用し、アクティブセンシング戦略に基づく物体認識において、どのデータセットの組み合わせが有用であるか調査していく。

#### 参考文献

- [1] Raymond J. Mooney Scott Niekum Luis Sentis Peter Stone Justin W. Hart1, Yuqian Jiangt and Andrea Thomaz. Austin villa@home 2019 dspl team description paper, 2020.
- [2] Patrick Nagel Maximilian Schobel Minh Nguyen Roberto Cai Wu Rohan Haselo Paul G. Ploger Gerhard K. Kraetzschmar Alex Mitrevski, Argentina Ortega Sainz. The b-it-bots@home 2019 team description paper, 2020.
- [3] Yushi Abe Tomohiro Ono Kohei Kabashima Takuma Sakata Masashi Fukuyado Fuyuki Muto Takumi Yoshii Kazuki Kanamaru Daichi Kamimura Kentaro Nakamura Yuta Nishimura Takashi Morie Yuichiro Tanaka, Yutaro Ishida and Hakaru Tamukoh. Hibikino-musashi@home 2019 team description paper, 2020.
- [4] Lars Kunze Bruno Lacerda Chia-Man Hung Mark Finean Denis Koksall-Rivet Ioannis Havoutis, Nick Hawes and Stephen Kyberd. Team orion 2019 dspl team description, 2020.
- [5] Edgar Silva Jesus Hernandez Adrian Sarmiento Jesus Savage, Edgar Vazquez and Angelica Nakayama. Dspl pumas@home 2019 team description paper, 2020.

<sup>1</sup>World Robot Summit, <https://worldrobotsummit.org>

<sup>2</sup>rosvbag, <http://wiki.ros.org/rosvbag>

<sup>3</sup><https://trcp.gitlab.io/erasers/>

- [6] L. Beaufls M. Davis U. Visser P. Pena, J. Masterjohn and C. Lisetti. Robocanes-visage 2019 team description paper, 2020.
- [7] L.L.A.M. van Beek J. Geijsberts L.G.L. Janssen S. Aleksandrov-K. Dang H.W.A.M. van Rooy A.T. Hofkamp D. van Dinther A. Aggarwal M.F.B. van der Burgh, J.J.M. Lunenburg and M.J.G. van de Molengraft. Tech united eindhoven @home 2019 team description paper, 2020.
- [8] Jaebong Yi Hyunjoon Cho Youngbin Park Byoung-Tak Zhang-Jae-bok Song Il Hong Suh Seung-Joon Yi, Chung-Yeon Lee. Team tidyboy robocup@home domestic standard platform league team description paper, 2020.
- [9] P. Kydd M. Pagnucco D. Rajaratnam C. Sammut M. Thielscher M. Velonaki G. Castro, C. Flannagan. Unsw robocup@home spl team description paper, 2020.
- [10] Luis Contreras, Yumiko Muto, Hiroyuki Okada, and Arata Sakamaki. Robot-object interaction strategies for object recognition. *The 37th Annual Conference of the Robotics Society of Japan*, 2019.
- [11] Stephen James, Zicong Ma, David Rovick Arrojo, and Andrew J. Davison. Rlbench: The robot learning benchmark and learning environment, 2019.
- [12] 小野智寛, 阿部佑志, 石田裕太郎, 田向権. ホームサービスロボットののための gazebo シミュレータを用いた半自動学習データセット生成手法の開発. *The 38th Annual Conference of the Robotics Society of Japan*, 2020.
- [13] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.
- [14] Berk Calli, Aaron Walsman, Arjun Singh, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M. Dollar. Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set. *IEEE Robotics Automation Magazine*, Vol. 22, No. 3, pp. 36–52, 2015.
- [15] Berk Calli, Arjun Singh, Aaron Walsman, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M. Dollar. The ycb object and model set: Towards common benchmarks for manipulation research. In *2015 International Conference on Advanced Robotics (ICAR)*, pp. 510–517, 2015.
- [16] Berk Calli, Arjun Singh, James Bruce, Aaron Walsman, Kurt Konolige, Siddhartha Srinivasa, Pieter Abbeel, and Aaron M Dollar. Yale-cmu-berkeley dataset for robotic manipulation research. *The International Journal of Robotics Research*, Vol. 36, No. 3, pp. 261–268, 2017.
- [17] Tully Foote. tf: The transform library. In *Technologies for Practical Robot Applications (TePRA), 2013 IEEE International Conference on, Open-Source Software workshop*, pp. 1–6, April 2013.
- [18] A. Pacheco-Ortega, H. Estrada, E. Vazquez, R. Martell, J. Hernandez, J. Cruz, E. Silva, J. Savage, and L. Contreras. Intelligent flat-and-textureless object manipulation in service robots. *IEEE/RSJ IROS Workshops*, 2018.
- [19] Takashi Yamamoto, Koji Terada, Akiyoshi Ochiai, Fuminori Saito, Yoshiaki Asahara, and Kazuto Murase. Development of human support robot as the research platform of a domestic mobile manipulator. *ROBOMECH Journal*, Vol. 6, , 2019.