

# RoboCup@Home DSPL

## Go Get It in Unknown Environment タスク用 ROS パッケージ

Hibikino-Musashi@Home, Kyushu Institute of Technology

### 1 提案手法の概要

Training Phase では、図 1 に示すように一般物体識別器である You Only Look Once v2(YOLOv2)[1] で識別した結果(図 1(b))に対して、音声情報により得た物体の呼称や特徴、場所などの情報(図 1(c))と、環境内のロボットの位置情報(図 1(d))を追加し、それらをデータベースに格納する。Test Phase では、コマンドとデータベースに格納された情報を比較し、目的のオブジェクトが何であるか判断する。以下、両 Phase の説明を記す。なお、本タスクで用いた音声認識、言語処理、Simultaneous Localization and Mapping (SLAM)、画像からの物体の切り出し等、ロボットの基本機能については Hibikino-Musashi@Home の TDP[2] を参照のこと。

### 2 Training Phase

#### 2.1 一般物体認識を用いたオブジェクトの識別

Training Phase では、YOLOv2 を用いて、オブジェクトの検出、識別を行なう。ここでは、COCO data set[3] を用いて学習を行なったパラメータを用いており、検出した物体をリアルタイムで 80 クラスに分類することができる。しかし、画像には他のオブジェクトや家具が写り込み、正しく対象を認識できないため。そこで、図 2 のように、画像上で認識対象とする部分を青色の矩形で指定しロボットのヘッドディスプレイに表示することで、目的のオブジェクトのみの記憶を可能にしている。

#### 2.2 音声認識による情報の付与

Training Phase では、YOLOv2 を用いて認識した物体にいくらかラベルが付与されるが、ロボットが未知物体の中からそれだけの情報をもとに目的のオブジェクトを特定するのは困難である。そこで、ロボットが目的のオブジェクトを特定しやすくするために、音声入力により情報を追加する。話者の英語音声を Google Chrome の Web Speech API [4] を用いて、文字データにして、オブジェクトの情報としてデータベースに格納する。図 1 に示す例では、tea というラベル(図 1(b))に追加して、bottle, shelf といった特徴や場所(図 1(c))などの情報を音声入力により追加し、さらに、SLAM から得た地図上の位置情報(図 1(d))を追加することでオブジェクトの特定を容易にしている。また、Test Phase で指定オブジェクトの確認のために用いるオブジェクトの画像(図 1(a))をここで撮影しておく。詳細は 3.2 章を参照のこと。

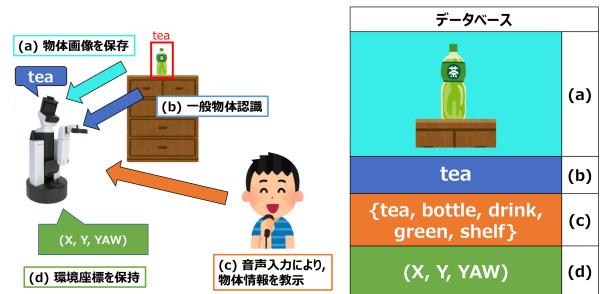


図 1: 音声認識による情報の付与

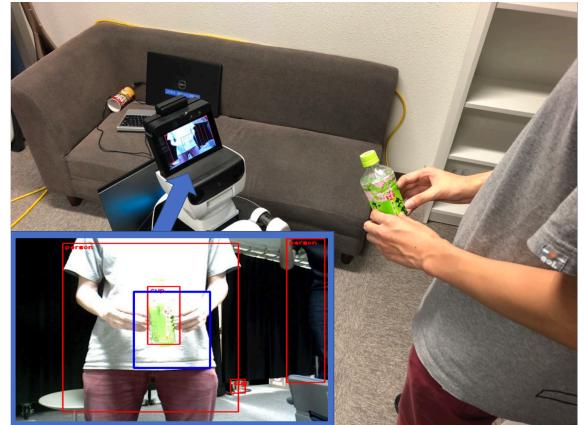


図 2: ヘッドディスプレイへの表示

### 3 Test Phase

#### 3.1 目的オブジェクトの特定

図 3 に音声情報による目的オブジェクトの特定方法を示す。オペレータにより入力された英語コマンドを Training Phase と同様に文字データ化して、それを単語ごとに品詞解析し、名詞と形容詞のみを抽出してデータベースに加える。次に、Training Phase で作成したデータベースと Test Phase で作成したデータベースを比較し、単語の一一致度が最も高いものを目的のオブジェクトとして判断する。

#### 3.2 ヘッドディスプレイを利用したオブジェクトの確認

目的のオブジェクトを判断した後、ロボットは判断したオブジェクトが本当に目的のオブジェクトであるかを確認する。確認作業を瞬時かつ正確に行うためにロボットのヘッドディスプレイを利用して視覚的に確認作業を行う。ヘッドディスプレイ



図 3: 目的オブジェクト特定の流れ

には、ロボットが判断した物体の写真 (Training Phase で撮影したオブジェクトの画像) を表示してオペレータに指示したものかどうか尋ねる。

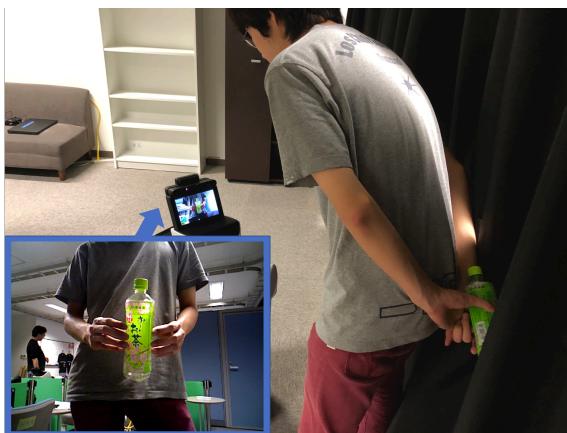


図 4: ヘッドディスプレイを用いた目的オブジェクトの確認

## 参考文献

- [1] J. Redmon, A. Farhadi. “YOLO9000:Better, Faster, Stronger,” 2017 IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, pp.6517-6525, 2017.
- [2] Y. Tanaka et al., “Hibikino-Musashi@Home 2019 Team Description Paper,” 2019.
- [3] COCO dataset. [http://ds.data.jma.go.jp/ghg/info\\_ghg.html](http://ds.data.jma.go.jp/ghg/info_ghg.html), (2019 年 8 月 11 日にアクセス).
- [4] Google Web Speech API. [https://developer.mozilla.org/ja/docs/Web/API/Web\\_Speech\\_API](https://developer.mozilla.org/ja/docs/Web/API/Web_Speech_API), (2019 年 8 月 11 日にアクセス).