Article information

Article title
Experimental Datasets and Processing Codes for the Semantic PHD Filter

Authors
Zhanteng Xie, Jun Chen and Philip Dames*

Affiliations
Department of Mechanical Engineering, Temple University, Philadelphia, PA 19122, USA

Corresponding author's email address and Twitter handle
pdames@temple.edu

Keywords
Multiple target tracking, PHD filter, Robot learning, Object detection

Abstract
This paper presents two datasets and their relevant processing codes for characterizing the range-bearing sensor used in semantic probability hypothesis density (SPHD) filters. One is called the water bottle detection dataset which contains a total of 4870 labeled water bottle images of four brands. The other is called the water bottle measurement-model dataset which contains a total of 3492 valid measurements with bottle semantic and location information. Both datasets were collected from different indoor environments around Temple University by a mobile robot equipped with a depth camera and a 2D lidar. First, the robot collected the water bottle detection dataset and then used its related processing codes to train a machine learning model to classify and detect four brands of mineral water bottles, namely Aquafina, Deer, Kirkland, and Nestle. Equipped with this trained water bottle detector, the robot was driven to collect the water bottle measurement-model dataset. Finally, the related codes of this water bottle measurement-model dataset were used to process the data and characterize three sensor models (*i.e.,* observation, detection, and clutter models) for the use of SPHD filters. Although both datasets are being combined to characterize the range-bearing sensor used in SPHD filters, they can also be used independently for some potential applications. For example, the water bottle detection dataset can be used for small object detection challenges, and the water bottle measurement-model dataset can be used for multiple object tracking and semantic mapping.

Specifications table

Subject	Mechanical Engineering
Specific subject area	Robotics and Artificial Intelligence
Type of data	• water bottle detection dataset: Images Labels Codes • water bottle measurement-model dataset: Rosbag Codes

How the data were acquired	• Hardware: A Turtlebot2 robot equipped with a NVIDIA Jeston TX2 embedded computer, a Stereolabs ZED stereo camera and a Hokuyo UTM-30LX lidar • Software: Ubuntu 16.04 system, ROS [1] Kinetic Kame, and LabelImg [1]
Data format	• water bottle detection dataset: Raw: JPG Labeled: PASCAL VOC, Darknet Code: Darknet[2] • water bottle measurement-model dataset: Raw: ROSBAG Code: Python, MATLAB
Description of data collection	• water bottle detection dataset: We drove the robot around the lobby and laboratory environments in the College of Engineering building at Temple University, and the living room and corridor environments in the Paseo Verde South Apartments for 30 min to collect 4870 RGB images of four brands of mineral water bottles at random positions. We then manually label these images using the LabelImg tool with the Pascal VOC and Darknet format. • water bottle measurement-model dataset: We drove the robot (running our YOLOv3 algorithm) around the lobby environment for 20 min to collect 3492 valid measurements of 12 targets (3 of each brand) at known positions.
Data source location	• Institution: Temple University, and Paseo Verde South Apartments • City/Town/Region: Philadelphia, PA 19122 • Country: USA
Data accessibility	• water bottle detection dataset: https://drive.google.com/file/d/1D5xOcHndxLHpf5r5d-pzmeli0itoHkuf/view?usp=sharing • water bottle detection processing codes: https://github.com/TempleRAIL/YOLOv3_bottle_detector.git • water bottle measurement-model dataset: https://drive.google.com/file/d/1CIkUx9RZVDSVsfqU2AfjtJ25OEFY1jbi/view?usp=sharing • water bottle measurement-model processing codes: https://github.com/TempleRAIL/sphd_sensor_models.git
Related research article	Jun Chen, Zhanteng Xie, and Philip Dames, The semantic phd filter for multi-class target tracking: From theory to practice, Robotics and Autonomous Systems 149 (2022) 103947.

Value of the Data

- The water bottle detection dataset can serve as a baseline for the tiny object detection (TOD) challenge. The water bottle measurement-model dataset and its relevant codes provide a template for experimentally characterizing sensor models for other researchers who wish to use the PHD/SPHD filter.
- Researchers with an interest in object classification and object detection can utilize the water bottle detection dataset and its relevant codes to gain further insights. Researchers with an interest in multiple

[1] https://github.com/tzutalin/labelImg
[2] https://github.com/pjreddie/darknet.git

object tracking and semantic mapping can utilize the water bottle measurement-model detection dataset and relevant codes to learn how to characterize sensor models for the use of Bayes state estimators.

- The water bottle detection dataset and its relevant codes can be used to train machine learning models that classify and detect four different brands of mineral water bottles. The water bottle measurement-model dataset and its relevant codes can be used directly to help researchers characterize sensor models for use with PHD/SPHD filters, possibly even extending to multiple object tracking or semantic mapping applications.

Data Description

In this data paper, we provide two types of datasets for semantic probability hypothesis density filters: 1) the water bottle detection dataset for training the water bottle detection detector, and 2) the water bottle measurement-model dataset for learning and modeling the observation, detection, and clutter models for SPHD filters. Note that both datasets contain their analysis and processing codes.

This water bottle detection dataset aims to train a YOLOv3 [2] detection network to classify and detect four brands of mineral water bottles, namely Aquafina, Deer, Kirkland, and Nestle. This detection dataset contains 4870 labeled images in total, where 4000 images are used for training and the rest of 870 images are used for validation. The samples of labeled images or detection result images are shown in Fig. 1. In addition to the collected dataset, we also provide the convolutional neural network (CNN) configure file and two pre-trained YOLOv3-tiny detection models, which are fine-tuned from this water bottle detection dataset. The confusion matrix calculated from our pre-trained detection models is shown in Table 3. The contents and files of the water bottle detection dataset and its related code are summarized in Table 2 and Table 4, respectively.

The purpose of the water bottle measurement-model dataset is to characterize the observation, detection, and clutter sensor models for the SPHD filters. This measurement-model dataset contains a 20 min raw Rosbag file that records five related topics. Specifically, it contains a total of 3585 raw individual measurements (i.e., YOLOv3 detection bounding boxes with accompanying point cloud data), but only 3492 measurements are valid. Table 5 summarizes this water bottle measurement-model dataset in detail. With this raw Rosbag dataset, we also provide the corresponding analysis and processing codes, written in Python and MATLAB, to process these raw data and develop observation, detection, and clutter models. The contents and files of these analysis and processing codes are summarized in Table 6.

Experimental design, materials and methods

This section describes the basic procedure and methods used to characterize a range-bearing sensor for the use of SPHD filters. The initial idea and procedure are from [4], which characterizes a bearing-only sensor for the use of PHD filter.

Figure summarizes the overall experimental design flow. First, we randomly placed water bottles in different poses in different environments (*i.e.,* living room, corridor, lobby, and laboratory), and then manually drove our Turtlebot 2 robot around these environments to collect the RGB images of these different brand water bottles. The randomly placing examples with different poses can be seen from Fig. 1. Then, we used the LabelImg tool to label these collecting images with four brand labels. We divided these labeled images into training data and validation data with a ratio of about 8:2 and got the final bottle detection dataset. After obtaining this bottle detection dataset, we used the Darknet framework and related configure files shown in Table 4 to train the YOLOv3-tiny detector which could detect and classify our four brands of water bottles. Based on this trained YOLOv3-tiny detector, we need to collect a measurement-model dataset to characterize our range-bearing sensor (*i.e.,* ZED camera). Before driving our robot to collect the data, we placed 12 bottle targets (3 of each brand) at specific positions in the lobby environment. The layout plan is shown in Fig. 3, where 12 specific positions were all selected in the corners of chairs, sofas, and walls. Since there was no motion capture system or other equipment to help us obtain the accurate ground truth positions of these 12 bottle targets, we proposed to use the lobby map information and these specific corner marks to help us get rough ground truth positions. Specifically, we used the RVIZ tool [3], a 3D visualization tool for ROS, to

[3] http://wiki.ros.org/rviz.

Table 2: Summary of the water bottle detection dataset.

Subject	**Description**
Name	bottle_detection_dataset.tar.gz
Collection environments	Paseo Verde South Apartments: living room and corridor College of Engineering building at Temple University: lobby and laboratory
Size	Total: 4870 images Train: 4000 images Validation: 870 images
Format	Data: JPG Label: PASCAL VOC and Darknet
Classes	Aquafina Deer Kirkland Nestle
Folder structure	./Annotations: contains the xml label files in PASCAL VOC format ./ImageSets: contains the training index files ./JPEGImages: contains the image data in JPG format ./Labels: contains the txt label files in Darknet format

Table 3: Confusion matrix of water bottles.

True \ Observed	Aquafina	Deer	Kirkland	Nestle
Aquafina	0.993	0	0	0.007
Deer	0	0.940	0.042	0.018
Kirkland	0	0	0.850	0.150
Nestle	0	0.252	0.032	0.716

visualize the lobby map where these specific corner marks could be located by their mouse clicks on the map. Then, the positions of these 12 specific corner marks can be used as our rough ground truth positions for these 12 bottle targets. After placing these 12 bottle targets in specific positions, we manually drove the robot around the lobby to collect an about 20 min measurement-model dataset. We then used the Python code shown in Table 6 to process the measurement-model data, determine the measurement-to-target association, and obtain the true positive, true negative, false positive, and false negative measurements.
Finally, we could use the MATLAB code shown in Table 6 to process the measurement-to-target association information, and fit the observation model, detection model, and clutter model which are used for characterizing our range-bearing sensor. The detailed sensor characterization procedure and modeling results are described in our related research paper [3].

CRediT author statement

Zhanteng Xie: Methodology, Software, Writing- Original draft preparation. **Jun Chen**: Conceptualization, Methodology, Writing- Reviewing and Editing. **Philip Dames**: Conceptualization, Methodology, Supervision, Writing- Reviewing and Editing.

Acknowledgments

This work was supported by NSF Grant IIS-1830419; and the Amazon Research Awards program. The authors would like to thank Olivia Chaves, Brandon Lutz, and Sacksit Vilayhong for their assistance in labeling the data used to train the water bottle detector.

Table 4: Summary of water bottle detection processing codes.

Subject	Description
Name	YOLOv3_bottle_detector
Framework	Darknet
Neural network architecture	YOLOv3-tiny network
Folder structure	./cfg: contains the YOLOv3-tiny network configure file ./weights: contains two pre-trained YOLOv3-tiny-sphd models, one for general purpose and one specially for our associated SPHD filter paper [3] ./demo: contains a detection video demonstration and several detection result image samples

Table 5: Summary of the water bottle measurement-model dataset.

Subject	Description
Name	botttle_measurement_model.bag
Collection environments	College of Engineering building at Temple University: lobby
Duration	19:59 s (1199 s)
Format	ROSBAG
Topics	/darknet_ros/detection_image: 3543 msgs /map: 1 msg /sphd_measurements: 3585 msgs /tf: 142727 msgs /tf_static: 1 msg
Message types	nav_msgs/OccupancyGrid sensor_msgs/Image sphd_msgs/SPHDMeasurements tf2_msgs/TFMessage

Declaration of Competing Interest

■ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Table 6: Summary of water bottle measurement-model processing codes.

Subject	**Description**
Name	sphd_sensor_models
Framework	ROS Kinetic
Format	Python MATLAB
Neural network architecture	YOLOv3-tiny network
Folder structure	./sphd_msgs: contains the files of our ROS custom message (*i.e.,* sphd_msgs) ./src/python/sphd_observation_model: contains the Python preprocessing codes and results for observation model ./src/python/sphd_detection_model: contains the Python preprocessing codes and results for detection model ./src/python/sphd_clutter_model: contains the Python preprocessing codes and results for clutter model ./src/matlab: contains three matlab analysis and processing codes for developing these three sensor models ./results: contains the final characterization and analysis result images

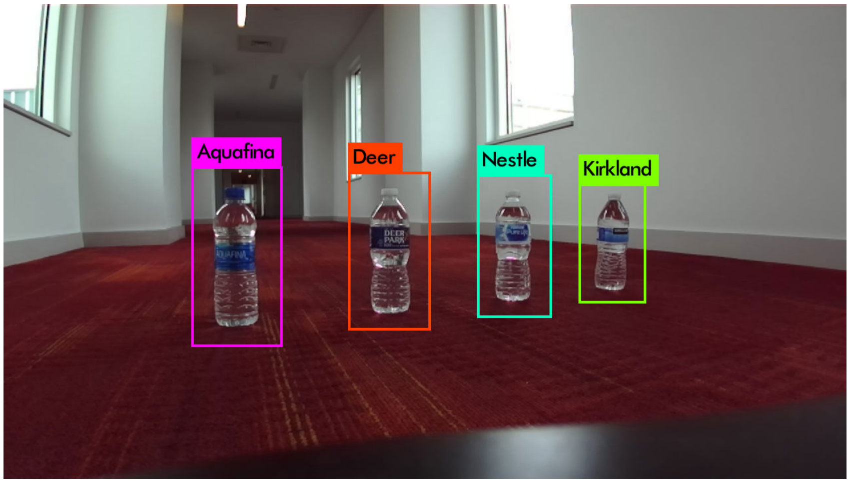
References

[1] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, et al., Ros: an open-source robot operating system, in: ICRA workshop on open source software, Vol. 3, Kobe, Japan, 2009, p. 5.

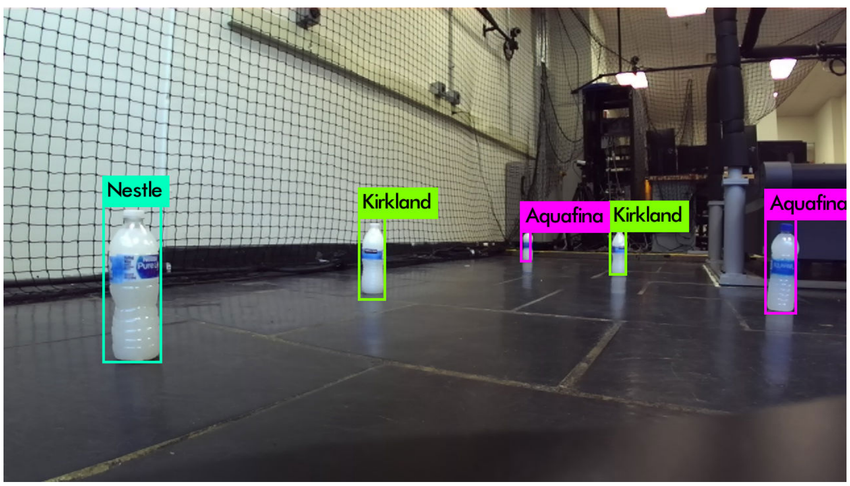
[2] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, arXiv preprint arXiv:1804.02767.

[3] J. Chen, Z. Xie, P. Dames, The semantic phd filter for multi-class target tracking: From theory to practice, Robotics and Autonomous Systems 149 (2022) 103947.

[4] P. Dames, V. Kumar, Experimental characterization of a bearing-only sensor for use with the phd filter, arXiv preprint arXiv:1502.04661.



(a) Corridor



(b) Laboratory

(c) Lobby

Figure 1: Samples of labeled images or detection result images at different environments.

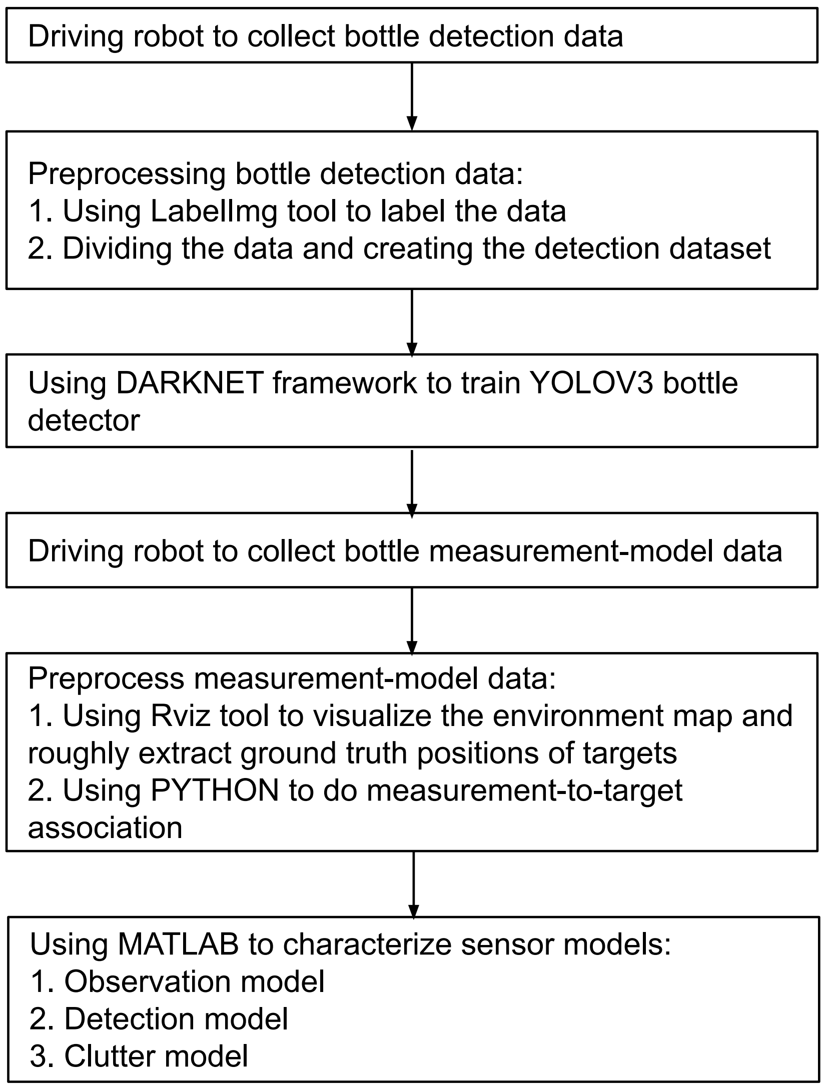


Figure 2: An overview of the experimental design process.

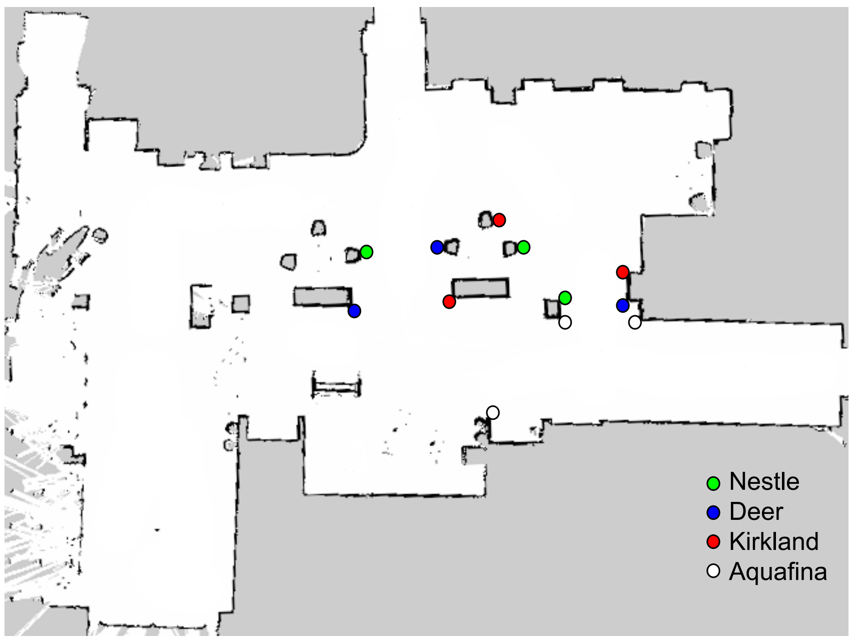


Figure 3: The water bottle layout plan for measurement-model data collection.