Radar Dynamic Object and Camera Detection Fusion

Junnan Shimizu

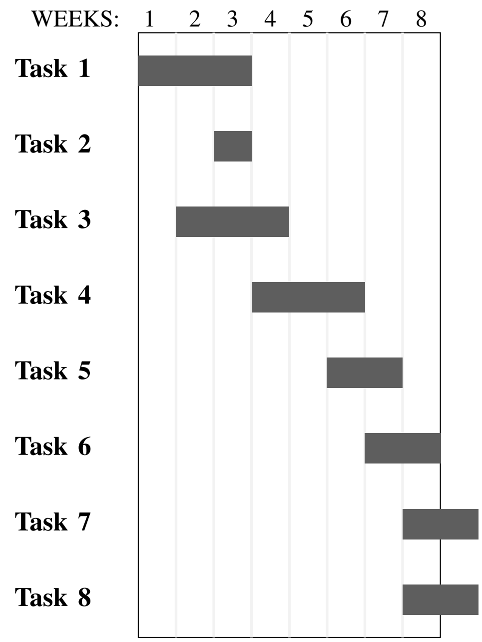
I. Research Question

Associate dynamic targets detected through radar imaging with objects detected through a camera image detection model.

II. Tasks

1) Utilize RVIZ to visualize radar and camera data, and experiment and learn about the data properties. (Week 1 - Week 3)
2) Develop a rough calibration of the camera and radar position in preparation for future data collection. (Week 3)
3) Get a pre-trained, real-time image detection model up and running (e.g. YOLOv3) and detect a subset of dynamic objects (e.g. Truck, Car, Bike, Person). (Week 2 - Week 4)
4) Develop a program that clusters radar dynamic target detections as radar dynamic targets. (Week 4 - Week 6)
5) Associate and identify radar dynamic targets with image detections — identify moving objects and categorize them. (Week 6 - Week 7)
6) Output in ROS a list of dynamic objects with the object class and position (range, azimuth). (Week 7 to Week 8)
7) Stretch goal: Calculate the another attribute of each dynamic object: object velocity.
8) Super stretch goal: Track the object's position and velocity over time.

National Science Foundation



III. Related Work

Radar technology improves substantially every year, and its development is only accelerated with the rapid growth of the autonomous and semi-autonomous vehicles industry [1] [13]. This growth has made radar systems smaller and cheaper, making it a very appealing option [10]. Mass market radars (MMRs) and high end radars (HERs) are in consideration for this project. Most modern vehicles use MMRs as it is cost effective, however HERs find benefits mainly in its ability to combine Dopper or micro-Doppler measurements with spatial resolution, as well as a higher resolution and greater dynamic range [1]. Initially, radar implementation in cars was used for crash avoidance, but quickly transitioned to functions such as: adaptive cruise control (ACC) [3], automatic emergency break (AEB), blind spot detection (BSD), and lane change assist (LCA) [13] with the aid of object detection and classification [5]. Radar has also been implemented in electronic equipment for the military (radar guidance for missiles, gunfire control, tracking of low-altitude targets) [15] as well as weather forecasting, environmental monitoring, and astronomical, atmospheric physics, ionospheric research, [7] intelligent robots, gesture recognition, biological sign recognition, [15], surveillance, among others [6]. It is important to note that with this increase of vehicles with some having up to

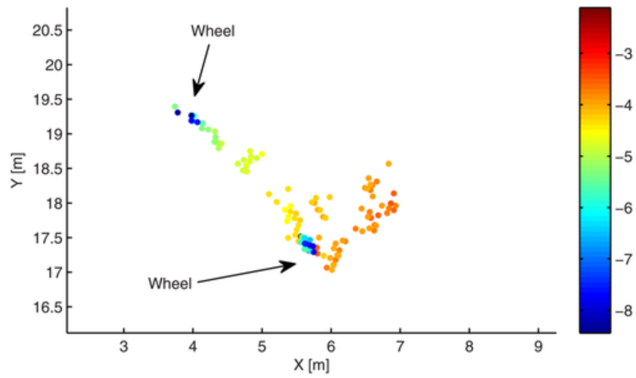


Fig. 1. Image of a passenger car, cropped from a high-resolution radar image. The colour indicates the radial velocity in m/s

12 radar sensors, interference between radar sensors is more likely causing moments of unreliable sensor measurement accuracy and sensor blindness [13]; however otherwise, it is advantageous compared to other imaging technologies because of its ability to function properly in low light and severe weather conditions [13] [12] [7]. Because of the reflection of radar off the road surface, vehicles will appear as laterally and longitudinally extended objects and the bottom side of vehicles can be seen which makes determining the class of the vehicle (cars, trucks, motorcycles, etc.) relatively easy compared to other imaging systems [5]. With the development of high resolution radar, a point-source assumption is not possible because an object will have more than one reflection, and therefore a clustering algorithm is required to track and determine the class of an object [3]. The most common clustering algorithm utilizes k-means which requires a predetermined number of clusters which is not valuable in an automotive scenario. On the contrary, DBSCAN is a density-based clustering method. The algorithm is given a threshold for the density of points within a certain degree, and the number of points must exceed for an object to be detected [5]. For this project, YOLOv3 will be utilized for clustering and tracking of objects [11] which avoids the issues experienced in [3].

The two main options to interpret the incoming radar data include synthetic aperture radar (SAR) and grid mapping. Within grid maps are amplitude and occupancy grid maps which are both useful for determining obstacles; however, occupancy grid maps provide a distinction between free space, obstacles, and unknown areas [13], so for this project, it is the most appropriate as it is highly transferable to an automotive context.

Another advantage that radar has over other imaging technologies is its ability to measure radial velocity via Doppler measurement. Examining this sample data of the radial velocity and its spatial data in Figure 1, the radial velocity of the points gradually increase from left to right with two exceptions. Given these unique clusters in the data, it allows for easier object identification, in this case, identifying the wheels of a vehicle [1]. Accuracy of [4]

In [13], by utilizing an analyze-before-measure approach, the receiver filters out the data depending on the operational parameters such as transmit direction or center frequency [13]. Within this project, data will be filtered using YOLOv3 to find the regions of interest (ROI) [11] [7] which will reduce the amount of data to process and therefore reduce processing times for dynamic object detection. By utilizing the YOLO algorithm, model classification performance saw great improvements. It is important to note that deep learning models with limited training data sets like radar samples will result in many over shooting scenarios and that is why there is great value in running a deep learning algorithm on camera data [7]. YOLO has also seen applications in similar veins such as the detection of marine ships. In [8], they fused LiraNet and YOLO and found strong performance benefits [8].

In [4], to detect moving objects and estimate their velocity, the project compared the optical flow of two consecutive frames. For points with no known depth, two or more consecutive frames are fused using Kalman filters [4]. For this project, in order to calculate velocity,

However, radar has its limitations. Although it has relatively high distance resolution [7] [12], it has low direction resolution and lower resolution in terms of orientation (azimuth/elevation) [15] [10] which can make detecting the objects' vertical position difficult [12]. Also, millimeter wave radar (MMW) is affected by electromagnetic waves and the detection is sparse so physical characteristics like sharp angles or curves are not easy, and in an automotive context, this means no line detection or easily detecting pedestrians [15]. On the contrary, a camera is able to fill the gaps that are left with radar as a camera provides high spatial resolution [12] with similar characteristics to human visual perception [14], but has poor depth perception. A fusion of these sensors fill each others shortcomings. For radar and camera fusion, the calibration of their locations is necessary [12]. For this project, a manual calibration of the radar and camera sensors will be established, but for future iterations, a dynamic calibration for the sensors will be implemented. In [12], their proposed method for calibration is to estimate the homography or the origin of data sets given by the radar and the camera planes. Many research projects for 3D object detection utilize the fusion of LiDAR and camera sensors [9]; however, LiDAR and camera still fail in high noise environments such as poor weather or fog [14]. In [9], they utilize CenterNet, a 3D object detection deep network with the use of a singular camera, and later fuse it with radar data called CenterFusion [9]. For this research, object detection will solely be done with the camera data and distance and velocity measurements will be done with the radar data. Another alternative for object detection, object size, and radial distance is by utilizing a multiple-camera system outlined in. This system fused with radar can obtain a fairly accurate 3D space. One thing to note from [10] is that a fixed y-coordinate can cause messy data and a rejection algorithm may be necessary for data filtration. Furthermore, when using multiple sensors, an efficient fusion rule is important to maintain the quality of the data from both sources [10]. In [2], they utilize a middle-fusion approach meaning that each set of sensor data is independently processed and merged later. From my understanding, [2] utilizes a similar process to [11], however [2] does not cluster the radar points and treats every radar point an an independent object detection.

REFERENCES

[1] Stefan Brisken, Florian Ruf, and Felix Höhne. Recent evolution of automotive imaging radar and its information content. *IET Radar, Sonar & Navigation*, 12(10):1078–1081, 2018.

[2] Siyang Han, Xiao Wang, Linhai Xu, Hongbin Sun, and Nanning Zheng. Frontal object perception for intelligent vehicles based on radar and camera fusion. In *2016 35th Chinese Control Conference (CCC)*, pages 4003–4008, 2016.

[3] Dominik Kellner, M. Barjenbruch, K. Dietmayer, Jens Klappstein, and Juergen Dickmann. Instantaneous lateral velocity estimation of a vehicle using doppler radar. pages 877–884, 01 2013.

[4] Dominik Kellner, Michael Barjenbruch, Jens Klappstein, Juergen Dickmann, and Klaus Dietmayer. Instantaneous full-motion estimation of arbitrary objects using dual doppler radar. pages 324–329, 06 2014.

[5] Dominik Kellner, Jens Klappstein, and Klaus Dietmayer. Grid-based dbscan for clustering extended objects in radar data. In *2012 IEEE Intelligent Vehicles Symposium*, pages 365–370, 2012.

[6] Du Yong Kim and Moongu Jeon. Data fusion of radar and image measurements for multi-object tracking via kalman filtering. *Information Sciences*, 278:641–652, 2014.

[7] Tong Lin, Xin Chen, Xiao Tang, Ling He, Song He, and Qiaolin Hu. Radar spectral maps classification based on deep learning. In *Proceedings of the 2020 International Conference on Computer Communication and Information Systems*, CCCIS 2020, page 29–33, New York, NY, USA, 2020. Association for Computing Machinery.

[8] Zhou Long, Wei Suyuan, Cui Zhongma, Fang Jiaqi, Yang Xiaoting, and Ding Wei. Lira-yolo: A lightweight model for ship detection in radar images. *Journal of Systems Engineering and Electronics*, 31(5):950–956, 2020.

[9] Ramin Nabati and Hairong Qi. CenterFusion: Center-based radar and camera fusion for 3d object detection. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, jan 2021.

[10] Felix Nobis, Maximilian Geisslinger, Markus Weber, Johannes Betz, and Markus Lienkamp. A deep learning-based radar and camera sensor fusion architecture for object detection. In *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–7, 2019.

[11] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement, 2018.

[12] S. Sugimoto, H. Tateda, H. Takahashi, and M. Okutomi. Obstacle detection using millimeter-wave radar and its visualization on image sequence. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 3, pages 342–345 Vol.3, 2004.

[13] Christian Waldschmidt, Juergen Hasch, and Wolfgang Menzel. Automotive radar — from first efforts to future systems. *IEEE Journal of Microwaves*, 1(1):135–148, 2021.

[14] Zizhang Wu, Guilian Chen, Yuanzhu Gan, Lei Wang, and Jian Pu. Mvfusion: Multi-view 3d object detection with semantic-aligned radar and camera fusion, 2023.

[15] Yong Zhou, Yanyan Dong, Fujin Hou, and Jianqing Wu. Review on millimeter-wave radar and camera fusion technology. *Sustainability*, 14(9), 2022.