A Modified Sequential Monte Carlo Bayesian Occupancy Filter using Linear Opinion Pool for Grid Mapping

Sang-Il Oh
Dept. of Digital Media
The Catholic University of Korea
nicolas0@catholic.ac.kr

Hang-Bong Kang
Dept. of Digital Media
The Catholic University of Korea
hbkang@catholic.ac.kr

Abstract

Occupancy grid state mapping is a key process in robotics and autonomous driving systems. It divides the environment into grid cells that contain information states. In this paper, we propose a modified SMC-BOF method to map and predict occupancy grids. The original SMC-BOF has been widely used in the occupancy grid mapping because it has lower computational costs than the BOF method. However, there are some issues related to conflicting information in dynamic situations. The original SMC-BOF cannot completely control an elongated vehicle that has conflicting information caused by the height difference between backward of vehicle and ground. To overcome this problem, we add confidence weights onto a part of the grid mapping process of the original SMC-BOF using the Linear Opinion Pool. We evaluate our method by LIDAR and stereo vision data in the KITTI dataset.

1. Introduction

Mapping and predicting occupancy grid probability are important tasks in autonomous driving systems or any robotic systems. A grid map has been widely used to estimate stationary indoor environments in combination with mapping and object tracking algorithms. Occupancy grid mapping was first proposed in the work of [6]. The occupancy grid mapping process is to represent the environment around an ego vehicle via grid cells. All grid cells contain a fusion of a variety of information from several different sensor measurements.

Moravec *et al*. [13] and Elfes [7] combined the likelihood function of several sensors using Bayesian reasoning to map and filter grid states. The latest algorithm that is used more often is the Bayesian occupancy filter (BOF) [2]. It uses a Bayesian framework that estimates the dynamic state of each cell with the occupancy probability. While early object tracking algorithm can track the shapes of predefined objects, the BOF can track any arbitrarily objects.

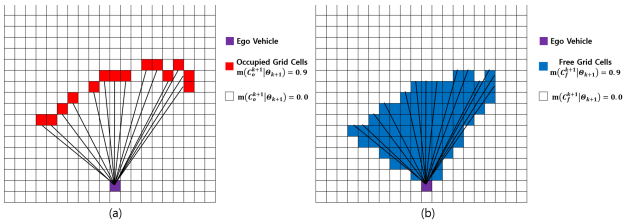

Figure 1. A simple overview of represented states via evidence theory. (a) Occupied state grids, (b) Free state grids.

Both algorithms mentioned above use the Bayes theory to map environments and to track objects using the information obtained from several sensors. However, the BOF has some problems. High computational cost occurs in the BOF because it applies a histogram filter to the dynamic state. To overcome this cost issue, Danescu *et al*. [3] used the sequential Monte Carlo Bayesian occupancy filter (SMC-BOF). Because the SMC-BOF uses particles to represent the dynamic state of grid cells that are treated independently, the computational cost can be reduced compared to the original BOF [5]. Danescu *et al*. [3] used stereo vision data for the SMC-BOF model. Negre *et al*. [14] and Tanzmeister *et al*. [19] applied multi-layer laser scanner data into the SMC-BOF. However, previous works related to the SMC-BOF cannot completely control an elongated vehicle, such as a trailer or a bus, or a stationary object. Due to the height difference between backward of a vehicle and the ground, it is very difficult to handle the information. Furthermore, a stationary object is often misjudged as a moving object with the same speed as that of the ego vehicle.

To solve this problem, Pathak *et al*. [16] used the Super-bayesian Independent Opinion Pool formula, and Moras *et al*. [12] used the Dempster-Shafer theory. Adarve *et al*. [1] presented the linear opinion pool for the occupancy grid mapping. The Linear Opinion Pool creates created more complex models by considering the interaction between

multi-sensor modalities via modifying the linear pooling of opinions [4].

In this paper, we proposed a modified SMC-BOF to handle conflicting information such as an elongated car. We use the Linear Opinion Pool for grid mapping instead of using the original mapping process of the SMC-BOF [5, 13, 16, 17]. We tested our model by using LIDAR and stereo vision data from KITTI dataset. The reason why we use LIDAR-stereo fusion data is that a stereo vision and a LIDAR can accurately measure the nearby object and a distant object, respectively.

This paper is organized as follows. Section 2 explains the SMC-BOF. Section 3 describes our proposed method, a modified SMC-BOF with the Linear Opinion Pool for the complex model. Section 4 shows the experimental results using the KITTI dataset and the conclusion is presented in Section 5.

2. The Sequential Monte Carlo Bayesian Occupancy Filter

This section explains an overview of the SMC-BOF implementation. The SMC-BOF is used to filter the occupancy probability of each grid cell. Most occupancy filters in object tracking algorithms follow similar process. First, information obtained from the multi-sensor is presented. Based on the presented data, we can predict the state of next frame. Finally, we estimate the posterior state by updating step which jointly computes the state of the predicted state and new observed states.

2.1. Representation of a grid state

The measured environment from equipped sensors is divided into 2D discrete grid cells. That is projected in birds eye view space. The size of the grid cell influences the computational time and the performance of the model. For a grid cell C, at time k, $S(C^k)$ is composed of the occupancy state $S(C_O^k)$ and the dynamic state $S(C_D^k)$. The occupancy state of a grid cell C considers free or occupied probability $S(C_O^K) \in \{P(C_o^k), P(C_f^k)\}$. The dynamic state of a grid cell C is comprises the 2D position and velocity $S(C_D^k) = [x, y, x_v, y_v]$.

The occupied probability of grid cell C at time k is estimated as the ratio between the number of particles in a cell C and the number of particles reflected by the obstacles:

$$P(C_o^k) = \frac{\sum_{i=1}^{N_C} w_i^k}{N_C}, w_i^k \in \Upsilon_C^k \tag{1}$$

$$w_i^k = \begin{cases} 1, \text{Reflected particles} \\ 0, \text{otherwise} \end{cases}$$

$$P(C_f^k) = 1 - P(C_o^k) \tag{2}$$

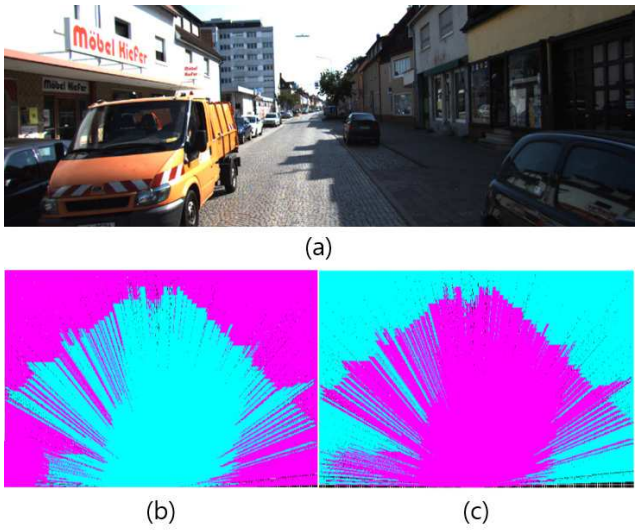

Figure 2. The occupancy grid state and the confidence grids are binarized. (a) Original input image, (b) occupancy grid state, (c) confidence grids.

, where N_C means the number of particles in grid cell C and Υ_C^k is a set of particles in a grid cell C at time k.

Most particles in grid C become a part of the unknown area, because of the occlusion area. In this area, we cannot define the occupancy probability and this area cannot have a dynamic state. To deal with this problem, we consider the grid with unknown state separately from other grids using the Dempster-Shafer theory of evidence [18]. We represent occupancy probability by the mass function based on a basic belief assignment (BBA). Therefore, we can represent the occupancy state of a grid cell C by $[m(C_o^k), m(C_f^k)]$. $m(C_o^k)$ is the mass of the occupied grid cell and $m(C_f^k)$ is the mass of the free grid cell at the time k.

In the occupancy state of an unknown area, we set $P(C_o^k) = 0.5$, such that $m(C_o^k) = m(C_f^k) = 0$. While an unknown area cannot be defined as either occupied or free state, can have potentials that can be either occupied or free state.

2.2. Prediction of the next observation state

This step provides the grid states at next time step using recent information. In this model, we individually estimate the posterior states of each grid cell $S(C^{k+1|k})$.

To predict the grid states at time $k+1$ from time k, we have to consider speed v and yaw rate ψ' information of the ego-vehicle in time interval T. This information is required to be compensated. The ego-vehicles yaw rate is ψ and the drive distance d is estimated using

$$\psi = \psi' T \tag{3}$$

$$d = \frac{2vTsin(\frac{\psi}{2})}{\psi} \tag{4}$$

Two coordinate axes D_x and D_z displace the origin of grid C [3]

$$D_x = dsin(\frac{\psi}{2})/x_{size} \tag{5}$$

$$D_z = dcos(\frac{\psi}{2})/y_{size} \tag{6}$$

, where x_{size} and y_{size} represent the size of the cell C. Position of the grid cell (x, y) is determined by following

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} cos\psi & -sin\psi \\ sin\psi & cos\psi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} D_x \\ D_y \end{bmatrix} \tag{7}$$

Finally, the following equation is used for the prediction

$$S(C_D^{k+1}) = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \\ x_v \\ y_v \end{bmatrix} + G_k \tag{8}$$

G_k represents randomly drawn discrete zero-mean white Gaussian noise with covariance matrix Q.

To predict the occupancy probability of grid cell C $S(C_O^{k+1})$, we use a transition parameter; that is, the occupancy state of cell a was moved to cell c, when the time changed T, $(\Delta_{c|a}^{k+1|k})$. At this point, we can also consider that $a = c$, which indicates that some particles remain in the same grid in the next time step.

The probability of occupancy is approximated by the mass function:

$$m(\Delta_{c|a}^{k+1|k}) = \sum_{w_{a,i}^{k+1|k} \in \Upsilon_{c|a}^{k+1|k}} w_{a,i}^{k+1|k} \tag{9}$$

Then, we can predict the occupancy mass as:

$$m(C_o^{k+1|k}) = 1 - \prod_{1 \le a \le C} (1 - m(\Delta_{c|a}^{k+1|k})) \tag{10}$$

To predict the mass for free space, we use another method that differs from the original SMC-BOF method in that the existing SMC-BOF method considers the absence of particles coming from an unobserved region as free space. However, the absence of particles does not indicate free space; it indicates that there is lack of information for the definition of the state. Therefore, a new method is required to solve this problem. Figure 1 is a simple example of grid states.

We can discount the mass for free of the last time step using nave free space prediction [5]:

$$m(C_f^{k+1|k}) = \alpha(F) m(C_f^k), \alpha(F) \in [0, 1] \tag{11}$$

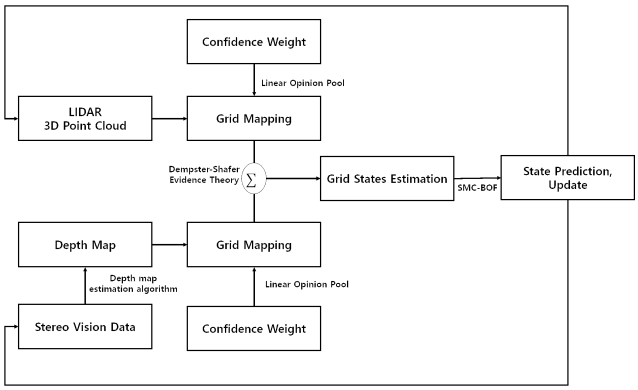

Figure 3. The model overview.

2.3. Update for new observation

This process provides the posterior cell states from the predicted cell states and a new observed state $S(C^{k+1|k+1})$. In the field of the occupancy grid filter, there are two sensor models. The first one is the forward sensor model. It is estimated by calculating the likelihood of the measurements obtained from each sensor [20]. The forward sensor model considers all inter-cell as dependent cells $P(measurements|Map)$. The second one is the inverse sensor model. The inverse model maps measurements to their causes and estimates all grid cells $P(Occupied|measurements)$ independently [10]. Most occupancy grid filters are designed for an inverse sensor model.

Therefore, we cannot apply the following in our model:

$$P(C_o^{k+1|k+1}) = \frac{P(\Theta_{k+1}|C_o^{k+1|k}) P(C_o^{k+1|k})}{h} \tag{12}$$

$$\begin{aligned} h = P(\Theta_{k+1}|C_o^{k+1|k}) P(C_o^{k+1|k}) \\ + P(\Theta_{k+1}|C_f^{k+1|k}) P(C_f^{k+1|k}) \end{aligned} \tag{13}$$

, where Θ means an observation.

We use the Dempster-Shafer theory to estimate the posterior occupancy probability. Here, we use a basic belief assignment (BBA) $m(C_o^{k+1}|\Theta_{k+1})$. This BBA is comprised of two mass functions: the evidence for occupied state $m(C_o^{k+1}|\Theta_{k+1})$ and the evidence for free state $m(C_f^{k+1}|\Theta_{k+1})$. To estimate the posterior occupancy BBA $m(C_o^{k+1|k+1})$, we use the Dempster-Shafer combination rule [15] to combine the posterior occupancy BBA and predicted BBA as follows:

$$m(S_O^{k+1|k+1}) = m(S_O^{k+1}|\Theta_{k+1}) \oplus m(S_O^{k+1|k}) \tag{14}$$

3. A Modified SMC-BOF with The Linear Opinion Pool

Conflicting information is common in autonomous driving systems and robotics caused by the feature difference between multi-sensors or dynamic environments.This causes grid mapping to be difficult. We apply the Linear Opinion Pool [4, 11] in the SMC-BOF to lessen the problem of the conflicting information [1].

The aim of this task is to estimate the confidence of the mapped grid cell C obtained from multi-sensors. Each sensor has an additional state that corresponds to a measure of confidence. Using the confidence measurement data containing irrelevant information can be discarded. We can easily model this method with an addition of the confidence parameter.

$$S(C_O^k)' = w^k(C)S(C_O^k) \tag{15}$$

, where $w^k(C)$ denotes the confidence weight of a cell C at time k. The confidence weights are designed differently based on the characteristics of each sensor. Figure 2 (b) and (c) are differences between the occupancy state map and the confidence map.

Our proposed method computes the confidence of input LIDAR and the depth map measured by stereo vision to represent grid states. Each grid states merged using the Dempster-Shafer evidence theory are filtered through the SMC-BOF. Figure 3 shows an overview of our proposed method.

3.1. Confidence of the LIDAR sensor model

For the confidence weight of the LIDAR sensor, we consider feature corresponds to the reflection by the angle of inclination of the road. A LIDAR is mounted at the top of a vehicle and a reflection similar to a contour is generated on the road. Therefore, the reflection data obtained from hit the road of beams is not reliable. This characteristic is modeled as follows:

$$w_{lidar}^{road}(\Theta) = max\{1 + \frac{\Theta tan(\Phi)}{h_0}, 0\} \tag{16}$$

, where h_0 is the height of the reflection at $\Theta = 0$, Φ indicates the angle between the reflection and the road surface.

3.2. Confidence of the stereo sensor model

The confidence weight for the stereo sensor model has two confidence components.

The first one considers the visibility of regions w_{stereo}^{vis}. This confidence includes a partially occluded region in the space. We can simply obtain the visibility confidence weight as the ratio of the number of possible P_p and visible pixels P_v.

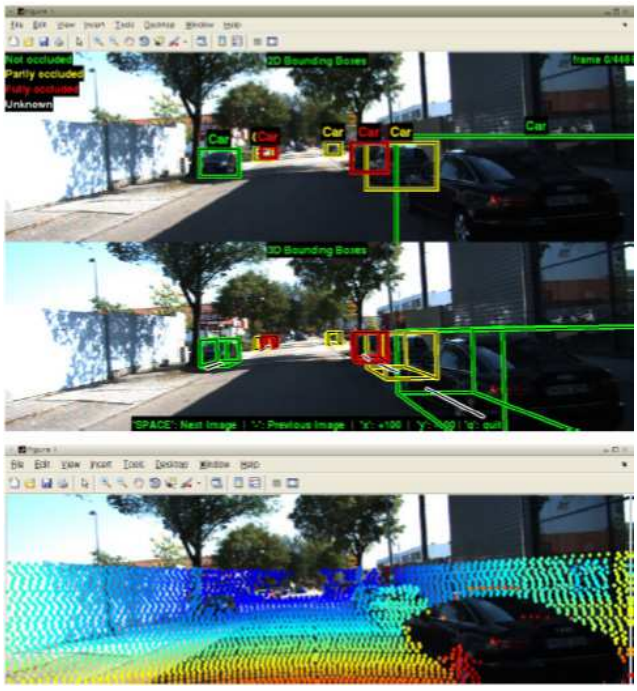

Figure 4. KITTI dataset. This dataset not only provides raw data measured by LIDAR, stereo cameras, GPS, but also includes other information such as synced data, objects label and so on.

$$w_{stereo}^{vis} = \frac{P_v}{P_p} \tag{17}$$

The second component is modeled as stereo vision works well at short distance range than greater distance. This confidence weight is estimated as follows:

$$w_{stereo}^{dist} = 1 - \frac{d_{max}^2}{d^2} \tag{18}$$

, where d and d_{max} denote the disparity values and the maximum possible disparity value, respectively.

Finally we can estimate the confidence weight of the stereo sensor model as follows:

$$w_{stereo}^{k} = w_{stereo}^{vis} \times w_{stereo}^{dist} \tag{19}$$

4. Experimental Results

The model was implemented on an Intel Core i5-4570 at dual 3.20 GHz with 8 GB RAM and NVIDIA GeForce GTX 650 graphics card with 3.7 GB of memory.

4.1. Dataset

We use the KITTI dataset for the evaluation of our model [8] (see Figure 4). The KITTI dataset is published by Andreas Geiger *et al.* in the Karlsruhe Institute of Technology for free. This dataset was captured from a Volkswagen

LIDAR	Velodyne HDL-64E, 10Hz, 0.09 degree angular resolution, collecting 1.3 million points per second, field of view: $[360]^\circ$ horizontal, $[26.8^\circ]$ vertical, range 120m
Color stereo	PointGray Flea2 (FL2-14S3C-C), 1.4 Megapixels
GPS	OXTS RT3003 inertial and GPS navigation system,6 axis, 100Hz, resolution: 0.02m per $[0.1]^\circ$

Table 1. The setting of equipped sensors.

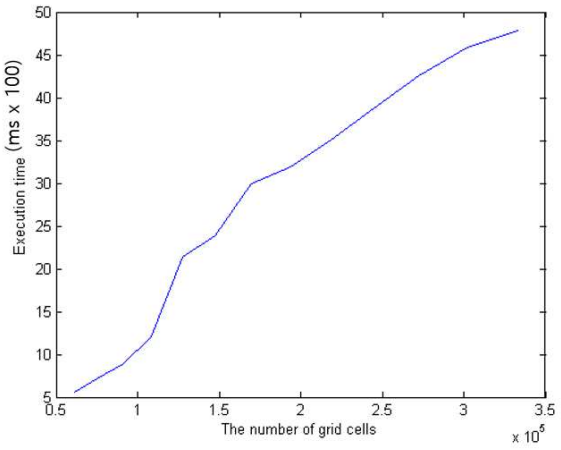

Figure 5. Graph showing the execution time according to the number of grid cells. Increasing of the number of grid cells results in an exponential increase in the computational times.

station wagon for mobile robotics and autonomous driving research. It has a variety of sensor modalities including color and grayscale stereo cameras, a Velodyne 3D LIDAR, and a GPS/IMU inertial navigation system. These sensors recorded various diverse scenarios captured in real-world traffic situations that range from freeways over rural areas to inner-city scenes with many objects. In addition, this dataset provides synchronized, timestamped data and a calibration matrix. Estimation of the disparity map via stereo vision is generated using the ARW method described by Lee *et al*. [9].

In this paper, we use LIDAR data, color stereo vision data, and GPS data from the KITTI dataset. Table 1 is the setup for each sensor.

4.2. Results

For the analysis of the execution time, we changed the grid resolution. The resolution is changed from 450x136 to 1030x311. Figure 5 is the execution time based on the grid resolutions.

To evaluate our proposed method, we performed a qualitative comparison between the original SMC-BOF and our modified SMC-BOF method, which includes the Linear Opinion Pool for the grid mapping process. Firstly, Figure 6 (a) is using a LIDAR-only data and the right side of bottom is using a LIDAR-Stereo vision data. The LIDAR-only data cannot completely measure grid states of nearby objects. Figure 7 (a) also used a LIDAR-only data. Similar to Figure 6, the LIDAR-only data cannot catch the nearby objects. Since a stereo vision is robust in detecting nearby objects, we can overcome this problem by using a LIDAR-stereo vision fusion data.

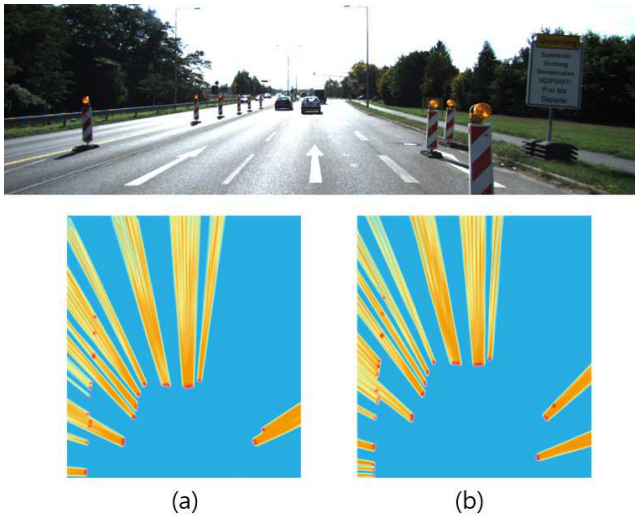

Figure 6. Result of the our modified SMC-BOF after 1s of observation. (a) LIDAR-only data, (b) LIDAR-stereo fusion data.

In particular, Figure 7 (b) and (c) are differences between the original SMC-BOF and the SMC-BOF with the linear opinion pool. Figure 7 (b) shows a situation in which the original SMC-BOF does not perform well in situations consisting of trucks or buses. Because of the elongated vehicle in the right side of scene, conflicting information is observed. The truck is disappeared in the original SMC-BOF method. On the other hand, our modified SMC-BOF method can solve these information conflicts.

5. Conclusion

In this paper, we proposed a modified SMC-BOF with linear opinion pool. We used the combination of LIDAR 3D point cloud data and stereo vision data obtained from KITTI dataset. Because a stereo vision is robust on nearby objects, it can improve the measurement ability of LIDAR. The SMC-BOF is used for the occupancy filter. To deal with complex objects, we used Linear Opinion Pool. The Linear Opinion Pool can solve the problem related to conflicting information.

There are drawbacks in using the combination of data.

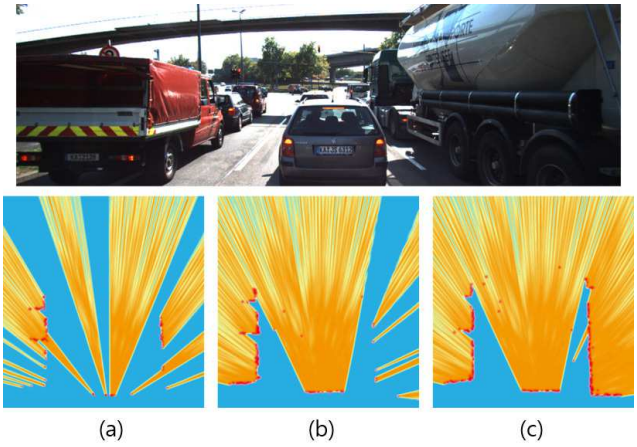

Figure 7. Result of the occupancy filter after 1s of observation. (a) our method on LIDAR-only data, (b) the original SMC-BOF on LIDAR-stereo fusion data, (c) our method on LIDAR-stereo fusion data.

The occupancy state mapping via stereo vision data depends on the depth estimation algorithm. Because the grid map is generated after the estimation of a depth map from the stereo vision data, the performance of the depth map estimation algorithm influences the grid mapping process. Therefore, we need an accurate depth estimation method. Another drawback is a hand designed metric for sensor confidences. A confidence metric in the linear opinion pool should rather be designed according to equipped sensor environments. Thus, this could have influenced the performance of our proposed model.

The SMC-BOF is definitely faster than the original BOF at the same system environment [5]. In the future work, we will implement a method with real-time capabilities by using GPU implementations.

Acknowledgements.

This research was supported by a grant from Agency for Defense Development, under contract #UD150016ID.

References

[1] J. D. Adarve, M. Perrollaz, A. Makris, and C. Laugier. Computing occupancy grids from multiple sensors using linear opinion pools. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 4074–4079. IEEE, 2012.

[2] C. Coué, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessière. Bayesian occupancy filtering for multitarget tracking: an automotive application. *The International Journal of Robotics Research*, 25(1):19–30, 2006.

[3] R. Danescu, F. Oniga, and S. Nedevschi. Modeling and tracking the driving environment with a particle-based occupancy grid. *Intelligent Transportation Systems, IEEE Transactions on*, 12(4):1331–1342, 2011.

[4] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[5] D.Nuss, T. Yuan, G. Krehl, M. Stuebler, S. Reuter, and K. Dietmayer. Fusion of laser and radar sensor data with a sequential monte carlo bayesian occupancy filter. In *Intelligent Vehicles Symposium (IV), 2015 IEEE*, pages 1074–1081. IEEE, 2015.

[6] A. Elfes. Occupancy grids: a probabilistic framework for robot perception and navigation. 1989.

[7] A. Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, 1989.

[8] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, page 0278364913491297, 2013.

[9] S. Lee, J. H. Lee, J. Lim, and I. H. Suh. Robust stereo matching using adaptive random walk with restart algorithm. *Image and Vision Computing*, 37:1–11, 2015.

[10] Y. Li. *Stereo vision and Lidar based dynamic occupancy grid mapping: Application to scenes analysis for intelligent vehicles*. PhD thesis, Université de Technologie de Belfort-Montbeliard, 2013.

[11] K. J. McConway. Marginalization and linear opinion pools. *Journal of the American Statistical Association*, 76(374):410–414, 1981.

[12] J. Moras, V. Cherfaoui, and P. Bonnifait. Credibilist occupancy grids for vehicle perception in dynamic environments. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 84–89. IEEE, 2011.

[13] H. P. Moravec. Sensor fusion in certainty grids for mobile robots. *AI magazine*, 9(2):61, 1988.

[14] A. Negre, L. Rummelhard, and C. Laugier. Hybrid sampling bayesian occupancy filter. In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, pages 1307–1312. IEEE, 2014.

[15] D. Nuss, B. Wilking, J. Wiest, H. Deusch, S. Reuter, and K. Dietmayer. Decision-free true positive estimation with grid maps for multi-object tracking. In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, pages 28–34. IEEE, 2013.

[16] K. Pathak, A. Birk, J. Poppinga, and S. Schwertfeger. 3d forward sensor modeling and application to occupancy grid based sensor fusion. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 2059–2064. IEEE, 2007.

[17] M. Perrollaz, A. Spalanzani, and D. Aubert. Probabilistic representation of the uncertainty of stereo-vision and application to obstacle detection. In *Intelligent Vehicles Symposium (IV), 2010 IEEE*, pages 313–318. IEEE, 2010.

[18] P. Smets. The combination of evidence in the transferable belief model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(5):447–458, 1990.

[19] G. Tanzmeister, J. Thomas, D. Wollherr, and M. Buss. Grid-based mapping and tracking in dynamic environments using a uniform evidential environment representation. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 6090–6095. IEEE, 2014.

[20] S. Thrun. Learning occupancy grid maps with forward sensor models. *Autonomous robots*, 15(2):111–127, 2003.