GM-PHD: Estimation of Dynamic Detection Probability in Video Data using YOLO

Michal Seibert
Faculty of Information Technology
Czech Technical University in Prague
Prague, Czech Republic
seibemic@fit.cvut.cz

Kamil Dedecius
Faculty of Information Technology
Czech Technical University in Prague
Prague, Czech Republic
&
Institute of Inf. Theory and Automation
Czech Academy of Sciences
Prague, Czech Republic
kamil.dedecius@fit.cvut.cz

***Abstract*—In standard target tracking algorithms such as the Gaussian mixture probability hypothesis density (GM-PHD) filter, it is required to manually preset the target detection probability *a priori*. The filters are relatively sensitive to correctness of this value and its inappropriate setting can lead to filter degeneracy. However, the estimation of the detection probability is tightly connected with the particular application. In this paper, we focus on video data, and introduce a novel method for estimation of dynamic time- and state-varying detection probability in the GM-PHD filter. It is primarily based on statistics and analysis of colors of a given frame with regard to a preceding frame sequence. To further enhance the performance of the filter in challenging real-life scenarios, we also propose a modified method for hypotheses pruning.**

***Index Terms*—Bayesian modeling, target tracking, detection probability, PHD filter**

I. Introduction

The objective of multi-target tracking (MTT) is to concurrently estimate the time-varying number of targets and their states from given sets of observations. Observations from various sensors come with the presence of data association uncertainty, detection uncertainty, and false detections (clutter) not originating from any target. That is, not all targets are detected by given sensors, and false alarms are present.

There are two paradigms in MTT. The first one considers explicit associations among measurements and targets [1]. The generic algorithms in this branch are the Joint Probabilistic Data Association (JPDA), Joint Integrated Probabilistic Data Association (IPDA), and the Multiple Hypothesis Tracking (MHT) filter, see, e.g., the overview in [2], [3]. The second paradigm is based on the more general Random Finite Set (RFS) approach of Mahler [3]. It provides a more flexible and robust framework to model uncertainties associated with data association, detection, and clutter. The Probability Hypothesis Density (PHD), Multi-Bernoulli Mixture (MBM), and Poisson MBM filters belong to this paradigm [3], [4].

The probabilistic formulation of the RFS-based filters and the inherent Bayesian processing of available information allow to accommodate the uncertainty arising from the presence of false detections, missed detections, and data association ambiguities. Nevertheless, a fundamental challenge persists: The performance of the filters is highly sensitive to the accurate setting of the target detection probability. This quantity, representing the likelihood of correctly identifying and associating observations with actual targets, is a critical parameter. It has a substantial impact on the Bayesian updating of the prior information. However, in real-world scenarios, the sensor performance is susceptible to various environmental conditions. Adverse weather, occlusions, or just the nature of the current scenario can lead to variations in detection probabilities. A mismatch between assumed and actual detection probabilities can result in a suboptimal tracking performance, leading to missed detections, false alarms, or inaccurate target state estimates [5].

The RFS-based formulation of the PHD or (P)MBM filters naturally takes the uncertainty about the detection probability into account [5]. The update formulae involve it as a function of the target state. In the figurative sense, this allows to model it as a function of the spatial and temporal properties of the environment. Still, two difficulties arise. First, the (Gaussian) filters are analytically tractable only if the detection probabilities are scalar numbers. Second, the nature of the detection probabilities differs from scenario to scenario.

In general, several methods have been proposed to deal with unknown detection probabilities. A Gaussian-beta modeling of a slowly-varying detection profile in the Cardinalized PHD filter is reported in [6]; its alternative for MBM filters follows in [7], and for PMBM filter in [8]. In [9], the authors propose to overcome some deficiencies in the CPHD filter [6] by different clutter/detection probability models. Another variant was recently proposed in [10]. A track-state augmentation with an amplitude offset-based prediction of the detection probability appeared in [11], however, this method suffers difficulties in multistatic fields. An automatic identification system-based sensor performance assessment for clutter-free environments is developed in [12]. A recent paper [13] deals with the unknown detection profile in the trajectory PHD/CPHD filters. There, the algorithm learns from the history of the unknown target detection probability.

This paper focuses on tracking targets in video data. This allows us to avoid the generic solutions and focus on the peculiarities associated with this specific data type. In particular, we take advantage of YOLO (You Only Look Once), offering high-performance real-time object detection with high accuracy and efficiency [14]. We use its ability to simultaneously predict multiple bounding boxes and class probabilities within an image, providing a streamlined and efficient approach to object detection. However, YOLO-based multi-target tracking systems can still be compromised under conditions such as adverse weather, low lighting, or scenarios with occlusions. Sensors may encounter difficulties in accurately detecting and localizing targets, leading to gaps or errors in the tracking process.

Our contribution is as follows: We propose a novel framework for a straightforward evaluation of the object detection probability based on bounding boxes provided by YOLO. In particular, YOLO detects objects and the framework predicts the location of the associated bounding boxes, and checks the color statistics. The method is independent of a particular MTT algorithm, and we demonstrate it in connection with the PHD filter. In order to improve the PHD hypotheses management, a modified method for hypotheses pruning is proposed too.

II. Background

This section outlines the PHD recursions with the dynamic time-varying detection probability and its implementation for image streams using object detection and segmentation. The development is given on a background of the standard PHD filter [3] and its Gaussian implementation [15].

In principle, the PHD filter employs a first-moment approximation: Instead of tracking the full posterior density over time, it focuses on propagating the posterior intensity. This approach resembles the constant-gain Kalman filter, which propagates the mean of the single-target state [15]. Due to the page limit, we do not delve into details of the PHD filter, they can be found in the cited literature.

A. The PHD recursions

Let us assume a target state x described by an intensity function $\nu(x)$. At time k, the prediction of the prior intensity $\nu_{k-1}(x)$ is given by

$$\nu_{k|k-1}(x) = \int p_{S,k}(\xi)\phi_{k|k-1}(x|\xi)\nu_{k-1}(\xi)d\xi + \nu_{\gamma,k}(x), \quad (1)$$

where $p_{S,k}(\cdot)$ is the probability of target survival, $\phi_{k|k-1}(\cdot|\cdot)$ is the target state transition density, and $\nu_{\gamma,k}(\cdot)$ denotes the prior PHD of the targets birth at time k. The predicted intensity $\nu_{k|k-1}$ is then updated by the measurement set Z_k given by sensors at time k according to the Bayesian update

$$\begin{aligned}\nu_k(x) = {} & [1 - p_{D,k}(x)]\nu_{k|k-1}(x) \\ & + \sum_{z\in Z_k} \frac{p_{D,k}(x)g_k(z|x)\nu_{k|k-1}(x)}{\kappa_k(z) + \int p_{D,k}(\xi)g_k(z|\xi)\nu_{k|k-1}(\xi)d\xi},\end{aligned} \quad (2)$$

where $g_k(\cdot|\cdot)$ is the likelihood function, $p_{D,k}(\cdot)$ is the probability of detection, and $\kappa_k(\cdot)$ is the clutter density.

B. Modeling detection probability from frame

As stated in the Introduction, the correct setting of the detection probability $p_D(x)$ can be a crucial step for satisfying results in target tracking algorithms. In this paper, we suggest a method to estimate the detection probability of a tracked object using the object detector YOLO and an object segmentation algorithm. YOLO itself is able to produce the object segmentation in real time, but for better results, more advanced algorithms such as Segment Anything from Meta are favorable. We propose to base the current point estimate of $p_D(x)$ at time k on the similarity of the subsequent frame properties,

$$p_{D,k}(x) = S_c\left[h\left(M(x;k|k-1)\circ D_k\right), h\left(M(x;k-1)\circ D_k\right)\right], \quad (3)$$

where D_k is the frame in the given color spectrum, $h(\cdot)$ is color histogram made from given color spectrum, $A \circ B$ is the Hadamard product, $M(\cdot;\cdot)$ is the object binary mask, and $S_c[\cdot,\cdot]$ is the cosine similarity of two vectors defined by

$$S_c[A,B] = \frac{A\cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2}\cdot\sqrt{\sum_{i=1}^n B_i^2}}, \quad (4)$$

where the sums run over all elements of the arguments.

C. The PHD recursion for linear Gaussian models

The linear Gaussian multiple-target model covers certain assumptions on the birth, death, and detection of targets. Summarization of these assumptions is as follows:

- Each target follows a linear Gaussian dynamical model and the sensor has a linear Gaussian measurement model

$$f_{k|k-1}(x|\zeta) = \mathcal{N}(x; F_{k-1}\zeta, Q_{k-1}), \quad (5)$$

$$g_k(z|x) = \mathcal{N}(z; H_k x, R_k), \quad (6)$$

  where $\mathcal{N}(x; m, P)$ denotes a Gaussian density with mean m and covariance P, F_{k-1} is the state transition matrix, Q_{k-1} is the process noise covariance, H_k is the observation matrix and R_k is the observation noise covariance.
- The intensity of the birth RFS is a Gaussian mixture of the form

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)}\mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}), \quad (7)$$

where $\gamma_k(x)$, $w_{\gamma,k}^{(i)}$, $P_{\gamma,k}^{(i)}$, $i = 1, ..., J_{\gamma,k}$ are given model parameters that determine the shape of the birth intensity. These assumptions are commonly used in many tracking algorithms. Formula (7) gives peaks of the spontaneous birth intensity. These points represent places with the highest probability of targets birth.

D. The GM-PHD recursion with dynamic detection probability

In the context of the linear Gaussian multiple-target model, the PHD recursion equations (1) and (2) attain analytical

solution [15]. Suppose that the posterior intensity at time $k-1$ is a Gaussian mixture of the form

$$\nu(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}). \tag{8}$$

The predicted intensity for time k is also a Gaussian mixture of the form

$$\nu_{k|k-1}(x) = \nu_{S,k|k-1}(x) + \gamma_k(x), \tag{9}$$

where $\gamma_k(x)$ is given in Formula (7). This yields

$$\nu_{S,k|k-1}(x) = p_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \mathcal{N}(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}), \tag{10}$$

$$m_{S,k|k-1}^{(j)} = F_{k-1} m_{k-1}^{(j)}, \tag{11}$$

$$P_{S,k|k-1}^{(j)} = Q_{k-1} + F_{k-1} P_{k-1}^{(j)} F_{k-1}^T. \tag{12}$$

Thus the predicted intensity at the time k is a Gaussian mixture

$$\nu_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}), \tag{13}$$

and the posterior intensity at time k is also Gaussian mixture,

$$\nu_k(x) = \sum_{i=1}^{J_{k|k-1}} (1 - p_{D,k}^{(i)}(x)) w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}) + \sum_{z \in Z_k} \nu_{D,k}(x; z), \tag{14}$$

where

$$\nu_{D,k}(x; z) = \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) \mathcal{N}(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}), \tag{15}$$

$$w_k^{(j)}(z) = \frac{p_{D,k}^{(j)}(x; z) w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k}^{(j)}(x; z) \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)}, \tag{16}$$

$$m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}(z - H_k m_{k|k-1}^{(j)}), \tag{17}$$

$$P_{k|k}^{(j)}(z) = [I - K_k^{(j)} H_k] P_{k|k-1}^{(j)}, \tag{18}$$

$$K_k^{(j)}(z) = P_k^{(j)} H_k^T (H_k P_{k|k-1}^{(j)} H_k^T + R_k)^{-1}. \tag{19}$$

The dynamically estimated detection probability in (14) follows from one of two possible scenarios. First, there is no current measurement at time k, hence the current mask results from its predicted position,

$$p_{D,k}^{(i)}(x) = \begin{cases} S_c^{(i)} \left[h(M_{k|k-1}^{(i)}(x) \circ D_k), h_{0,k}^{(i)} \right] & \text{if } h_{0,k}^{(i)} \text{ exists,} \\ p_{D,k}^{(i)} & \text{otherwise,} \end{cases} \tag{20}$$

and where the histogram from the last detection $h_{0,k}^{(i)} \equiv h_{0,k-1}^{(i)}$ is used in the first scenario. A user-preset value $p_{D,k}^{(i)}$ is used for targets undetected so far. The mask is obtained via

$$M_{k|k-1}^{(i)}(x) = \begin{cases} M_{k-1}^{(i)}(x)[m - v_x, n - v_y] & \text{if } m \geq v_x, n \geq v_y, \\ \mathbf{0} & \text{otherwise,} \end{cases} \tag{21}$$

where m, n are indices of the binary mask M and v_x, v_y are x- and y-direction velocities of the target. $\mathbf{0}$ is a zero vector of the same size as the mask. The second scenario takes the current measurement z into account. The detection probability is modified according to

$$p_{D,k}^{(j)}(x; z) = S_c \left[h(M_k^{(j)}(z) \circ D_k), h_{0,k}^{(j)} \right], \tag{22}$$

where $h_{0,k}^{(j)}$ is the histogram resulting from the last detection,

$$h_{0,k}^{(j)} = \begin{cases} h(M_{k-1}^{(j)}(z) \circ D_{k-1}) & \text{if } h(M_{k-1}^{(j)}(z)) \text{ exists,} \\ h_{0,k-1} & \text{otherwise.} \end{cases} \tag{23}$$

The histogram $h(M_{k-1}^{(j)}(z) \circ D_{k-1})$ stands for the previously detected target, and $h_{0,k-1}$ is the histogram from the time step when the target was detected the last time.

E. Modified pruning for GM-PHD filter

As the PHD filter propagates the posterior intensity, the number of potential hypotheses exponentially increases. This leads to an enormous increase in memory and computational demands. Pruning techniques become indispensable in mitigating this computational load by selectively discarding less likely hypotheses, allowing for a more focused and efficient tracking process. These pruning mechanisms, guided by predefined thresholds or heuristics, ensure that computational resources are allocated judiciously, striking a balance between accuracy and computational efficiency in multi-target tracking applications.

Recall that sensors in our work are represented by cameras. Naturally, the detecting algorithms are unable to recognize an object hidden by an obstacle. These obstacles may differ not only in size but also in color. In situations where the color of an obstacle closely matches the surrounding scene, the likelihood of detection remains sufficiently high, increasing the risk that the target may not survive even though it is only hidden. In order to suppress this phenomenon, we introduce a method for modified pruning along with standard pruning and merging techniques [15].

The most deployed generic method for pruning consists of removing targets with weights below some predefined threshold. Nonetheless, as mentioned before, targets with some history may only be hidden, and it is not desired to remove these targets from the scene. To overcome this issue, we assign each target a tag that represents the state in which the target is likely to occur,

$$S = \{\text{detected, hidden, dead}\}.$$

This tag is modeled by a discrete-time Markov chain with the transition matrix

$$P = \begin{bmatrix} p_{D,k} & 1 - p_{D,k} & 0 \\ p_{D,k} & (1 - p_{H,k}^{T_H}) \cdot (1 - p_{D,k}) & p_{H,k}^{T_H} \cdot (1 - p_{D,k}) \\ p_{D,k} & (1 - p_{H,k}^{T_H}) \cdot (1 - p_{D,k}) & p_{H,k}^{T_H} \cdot (1 - p_{D,k}) \end{bmatrix}, \tag{24}$$

where $p_{D,k}$ is a generic detection probability, and T_H is an exponent for controlling probabilities that is set higher in

scenarios, where the target color mask is similar to the color mask of the obstacle, or may be set lower otherwise. $p_{H,k}$ is the probability that the target is removed. This probability is the result of (27), where the bounding boxes of the object detector are used. If we denote by $B^{(j)}_{P,k}$ the bounding box predicted T_c steps ahead and by $B^{(j)}_{S,k}$ the bounding box of the last step k when the object was detected, then

$$B^{(j)}_{P,k} = B^{(j)}_{P,k-1} + n^{(j)}_k \cdot [v^{(j)}_{x,k}, v^{(j)}_{y,k}, {}^{(j)}_{x,k}, v^{(j)}_{y,k}], \quad (25)$$

$$B^{(j)}_{S,k} = B^{(j)}_{S,k-1}, \quad (26)$$

$$p_{H,k} = S_c \left[D_k(B^{(j)}_{P,k}), D_k(B^{(j)}_{S,k}) \right], \quad (27)$$

where

$$n^{(j)}_k = \begin{cases} T_c \text{ if } k \ mod \ T_c = 0, \\ 0 \quad \text{otherwise.} \end{cases}, \quad (28)$$

$D_k(\cdot)$ is a part of a frame within the bounding box given by $B(\cdot)$. With this approach, when a target is most likely in the *hidden* state, the pruning threshold is heuristically lowered to prevent the target removal.

Experiment results

To test the proposed method, a video of a challenging traffic environment was recorded. The state $x_k = [p_{x,k}, p_{y,k}, v_{x,k}, v_{y,k}]^T$ of each target consists of two-dimensional position $(p_{x,k}, p_{y,k})$ and velocity $(v_{x,k}, v_{y,k})$. The measurement is the center of the mask of the detected object, which is typically noisy. The survival probability of the targets $p_{S,k} = 0.99$. A state evolution model (5) is employed with

$$F_k = \begin{bmatrix} I_2 & \Delta I_2 \\ 0_2 & I_2 \end{bmatrix}, \quad Q_k = \sigma^2_v \left[\Delta I_4 \right], \quad (29)$$

where I_n and 0_n denote the $n \times n$ identity and zero matrices respectively, $\Delta = 1$ is the sampling period and $\sigma_v = 0.1(px/s^2)$ is the standard deviation of the process noise. The mean of the birth place is roughly in the middle of the traffic lane and its confidence ellipse of size of three standard deviations is bounded by a blue circle in Figures 1-6. The covariance matrix of spontaneous birth is $P = diag([800, 800, 800, 800])$.

The measurement follows the observation model (6) with $H_k = [I_2, 0_2]$, $R_k = \sigma^2_\epsilon I_2$, where $\sigma_\epsilon = 15px$ is the standard deviation of the measurement noise. I_2 and 0_2 stand for the 2×2 identity and zero matrices, respectively. The detection probability for a previously undetected target from equation (20) is set to 0.3, T_H for the transition matrix to 1.3 and T_c to 3. Only the YOLO segmentation model has been used.

The video is recorded at 10 fps. The red bounding boxes in Figures 1-6 represent bounding boxes given by YOLO object detection model. The red ellipses are the covariance matrices of the targets and the labels above the targets are the numerical representations of the state targets: $0 =$ detected, $1 =$ hidden, $2 =$ dead. A traffic light pole obstructs the camera's view, making it impossible for YOLO to detect objects and thus obtain measurements. In Fig. 1 (frame number 36), the last detection of the first car before an obstacle is obtained. In Fig. 2 the first car is hidden but still identified due to the GM-PHD filter with the time-varying detection probability. It can also be seen that the first car label is $1 =$ hidden. The second car entered the birth area. In Fig. 3, the first car receives a new measurement given by YOLO. It is important to note that the car was not recognized for 9 consecutive time steps, and the target representing the car was able to survive, which would not be possible with a constant high detection probability with a reasonable pruning threshold. The same scenario applies for the second car as well, which can be seen in Figures 4-6. In the Figure 6, both cars continue their paths with the correct label $0 =$ detected.

	Pd=0.7, 1 SP	Pd=0.95, 1 SP	Pd=0.95, 2 SP	dynamic Pd
MSE	0.8316	0.8811	0.40594	0.0297

TABLE I
MSE OF NUMBER OF TARGETS ACROSS DIFFERENT SETTINGS

Figure 7 and Figure 8 depict the number of identified targets across 100 time steps. In the former figure, three approaches are compared – the proposed one, and fixed detection probabilities equal to 0.7 and 0.95, respectively. Under the proposed dynamic detection probability, the targets are not only identified when hidden, but also when YOLO is not able to detect them. On the other hand, constant detection probabilities lead to track losses. This can be seen at around time steps 38, 52 and 83. The standard way to resolve this issue is to add a spawning point right after the obstacle. The result of this procedure is shown in Fig. 8. Even though the target is then again identified when detected by YOLO, another problem arises: The target is considered as a totally new target. To demonstrate this issue, we assign each new target an ID – a natural number with the increasing order. In Fig. 9, we see that the proposed method tracks three IDs, while the standard PHD filter initialized new targets with new IDs.

Table I shows the mean squared error (MSE) of the number of targets. With only one or two targets appearing in a scene at once, with constant detection probability 0.7 and 0.95, we got error 0.83 and 0.88, respectively. With one extra spawn point right after the obstacle, the error is 0.40 with $p_d = 0.95$. Clearly, in the presented scenario, the dynamically set detection probability outperforms the constant one.

Conclusion

The proposed dynamic detection probability estimation and the modified pruning technique for the GM-PHD filter address challenges in multi-target tracking. Real-time adaptation to varying detection probabilities enhances robustness, while the modified pruning mitigates the computational complexity without prematurely discarding potentially hidden targets. Experimental results demostrate the efficacy of the proposed approach even in a challenging traffic environment.

References

[1] Y. BarShalom and T. E. Fortmann, *Tracking and Data Association*. San Diego, CA: Academic, 1988.

[2] S. Blackman, *Multiple Target Tracking with Radar Applications.*.Norwood, MA: Artech House, 1986.

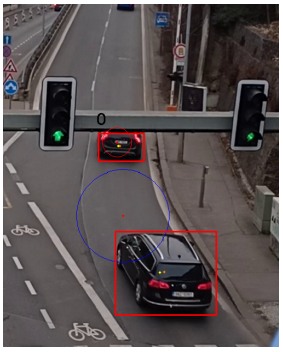

Fig. 1. Frame number 36.

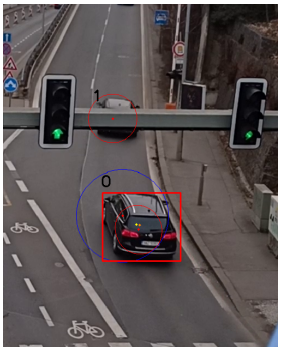

Fig. 2. Frame number 40.

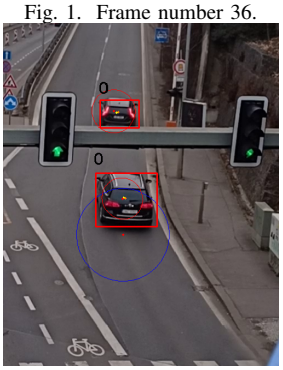

Fig. 3. Frame number 45.

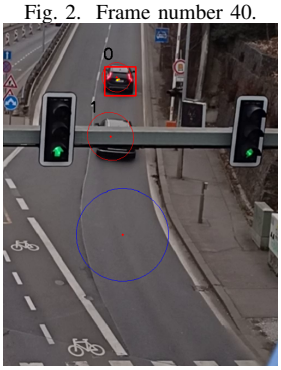

Fig. 4. Frame number 55.

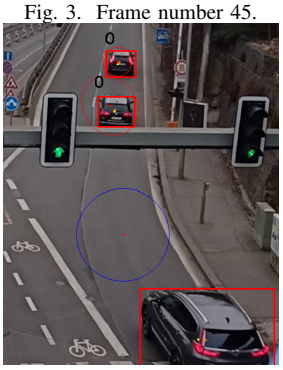

Fig. 5. Frame number 61.

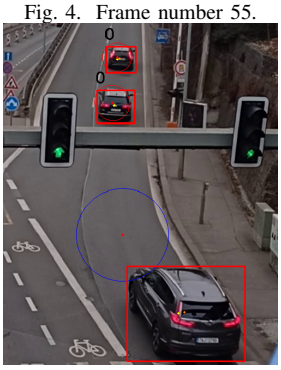

Fig. 6. Frame number 63.

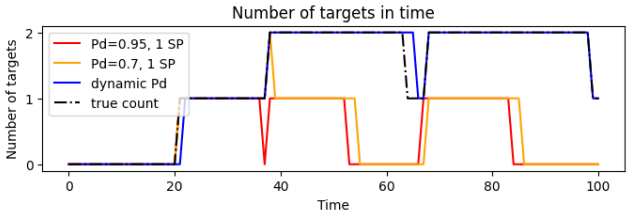

Fig. 7. The number of targets when using the proposed dynamic detection probability and the constant detection probabilities with values 0.7 and 0.95, respectively. One spawn point is set before the obstacle.

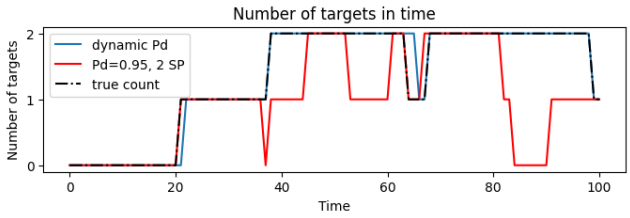

Fig. 8. The number of targets when using the proposed dynamic detection probability and the constant detection probability with value 0.95. One spawn point is set before the obstacle and one spawn point is set right after the obstacle.

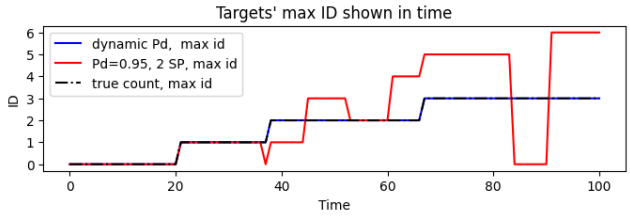

Fig. 9. Maximum of unique ID detected using dynamically set and constant detection probability (0.95). One spawn point is set before the obstacle and one spawn point is set right after the obstacle.

[3] R. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[4] A. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson Multi-Bernoulli Mixture Filter: Direct Derivation and Implementation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1883–1901, 2018.

[5] G. Hendeby and R. Karlsson, "Gaussian mixture PHD filtering with variable probability of detection," *Proc. 17th International Conference on Information Fusion (FUSION)*, 2014.

[6] R. P. S. Mahler, B.-T. Vo and B.-N. Vo, "CPHD Filtering With Unknown Clutter Rate and Detection Profile," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3497-3513, Aug. 2011,

[7] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad and R. Mahler, "Robust multi-Bernoulli filtering", *IEEE Journal of Select Topics in Signal Processing*, vol. 7, no. 3, pp. 399-409, Jun. 2013.

[8] G. Li, L. Kong, W. Yi and X. Li, "Robust Poisson Multi-Bernoulli Mixture Filter With Unknown Detection Probability," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 886-899, Jan. 2021

[9] C. Li, W. Wang, T. Kirubarajan, J. Sun and P. Lei, "PHD and CPHD Filtering With Unknown Detection Probability," *IEEE Transactions on Signal Processing*, vol. 66, no. 14, pp. 3784-3798, 2018.

[10] J. Wei, F. Luo, J. Qi and L. Ruan, "A Modified BGM-PHD Filter with Unknown Detection Probability," *2023 6th International Conference on Information Communication and Signal Processing (ICICSP)*, Xi'an, China, 2023, pp. 492-496, doi: 10.1109/ICICSP59554.2023.10390615.

[11] E. Hanusa and D. Krout, "Track state augmentation for estimation of probability of detection in multistatic sonar data", *Proc. Asilomar Conference on Signals Systems and Computers*, pp. 1733–1737, Nov. 2013.

[12] S. Horn, "Near real time estimation of surveillance gaps", *Proc. 16th International Conference on Information Fusion (FUSION)*, pp. 1871–1877, Jul. 2013.

[13] S. Wei, B. Zhang and W. Yi, "Trajectory PHD and CPHD Filters With Unknown Detection Profile," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8042-8058, Aug. 2022.

[14] S. Yan, Y. Fu, W. Zhang, W. Yang, R. Yu and F. Zhang, "Multi-Target Instance Segmentation and Tracking Using YOLOV8 and BoT-SORT for Video SAR," *2023 5th International Conference on Electronic Engineering and Informatics (EEI)*, Wuhan, China, 2023, pp. 506-510, doi: 10.1109/EEI59236.2023.10212903.

[15] Vo, B.-N. and Ma, W.-K., "The Gaussian Mixture Probability Hypothesis Density Filter," *in IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091-4104, Nov. 2006, doi: 10.1109/TSP.2006.881190.