Simultaneous fusion, classification, and traction of moving obstacles by LIDAR and camera using Bayesian algorithm

Fusión, clasificación y tracción simultáneas de obstáculos en movimiento mediante LIDAR y cámara usando algoritmo bayesiano

Author:

Masrour Dowlatabadi[1]

Ahmad Afshar[2*]

Ali Moarefianpour[3]

SCIENTIFIC RESEARCH

How to cite this paper:
Afshar A., Dowlatabadi M. and Moarefianpour A., Simultaneous fusion, classification, and traction of moving obstacles by LIDAR and camera using Bayesian algorithm, Iran. Innovaciencia 2020; 8 (1): 1-10. DOI: http://dx.doi.org/10.15649/2346075X.997

Reception date:
Received: 01 May 2020
Accepted: 01 August 2020
Published: 01 December 2020

Keywords:
Mahalonobis Distance; Bayesian Algorithm; Simultaneous classification and traction; LIDAR sensor and camera.

ABSTRACT

In the near future, preventing collisions with fixed or moving, alive, and inanimate obstacles will appear to be a severe challenge due to the increased use of Unmanned Ground Vehicles (UGVs). Light Detection and Ranging (LIDAR) sensors and cameras are usually used in UGV to detect obstacles. The definite tracing and classification of moving obstacles is a significant dimension in developed driver assistance systems. It is believed that the perceived model of the situation can be improved by incorporating the obstacle classification. The present study indicated a multi-hypotheses monitoring and classifying approach, which allows solving ambiguities rising with the last methods of associating and classifying targets and tracks in a highly volatile vehicular situation. This method was tested through real data from various driving scenarios and focusing on two obstacles of interest vehicle, pedestrian.

[1] Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran, Masror_d@yahoo.com

[2*] Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran. Corresponding author: aafshar@aut.ac.ir

[3] Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran, moarefian@srbiau.ac.ir

INTRODUCTION

Smart UGVs have been developed from being a robotic use of tomorrow to the current field of broad research and advancement. The most significant feature of a smart UGV system is that it should operate in increasingly unstructured situations being inherently uncertain and dynamic. Because of the increasing application of UGV, preventing the collisions with fixed or moving, alive and inanimate obstacles will be a serious challenge in the near future. Collision avoidance refers to a significant task in many uses, like ADAS (developed driver-assistance systems), industrial automation, and robotics. In an industrial automation setting, certain fields have to be off-limits to an UGV for protecting people and high-value assets. ADAS helps drivers to run intricate driving tasks in order to avoid dangerous condition.

Perceiving the situation includes the selection of various sensors for obtaining a detailed description of the situation and exact detection of the obstacles of interest. LIDAR sensors and cameras are applied in UGV in order to detect moving obstacles. LIDAR is a remote sensing method that is broadly used in many fields.

The management of imperfect information is a critical need for perception systems. The correct detection of moving obstacles is a significant aspect of a moving object tracking system. A lot of sensors are typically part of these systems. Tracking an obstacle like a car on the road is a three-step process with the stages: (1) synchronization, (2) association, and (3) fusion. The synchronization task predicts the development of the known obstacles to the current timestamp k, and knows information on their behavior at time k-1. Predicting obstacles is called tracks, while the observation of the sensors is called targets. The association step finds which track corresponds to which targets before they can be fused in the last step for obtaining a more precise description of the scene at time k. Multi-sensor fusion at track level needs a list of updated tracks from each sensor to fuse them into a mixed list of tracks. The works in [1], [2], solve this problem by focusing on the association problem.

The classification information of the obstacle was not used in estimating and predicting the obstacles tracking as we are aware of the class of obstacles that surround the vehicle provides a better perception of driving situations. Classification is regarded as a separate task in the DATMO (detecting and tracking the moving object) tasks or as aggregating information for the final perception output. Knowing the class of a moving object assists with learning and tracking the motion model. Classifying the obstacle's information by a camera improves the detection and tracking of the moving obstacles. We include an object's class as the critical component of a tracking technique, which provides uncertainty management from sensor detection. The goal is improving the results of the perception task. Thus, this study addressed the problem of sensor data association and tracking. The present study assumed that a rich list of tracked obstacles could enhance the future stages of an ADAS.

The rest of the paper was organized as follows. Section2 reviews the related work. In section3, the tracking process is described, and in Section4, the proposed method for classifying and tracking obstacles is expressed. Section 5 discusses the results of the proposed algorithm.

1. RELATED WORK

Data decision-making technology-based on the multi-sensor is highly valued by scholars at home and abroad. In addition, a lot of theorem and algorithms emerged in the field of data decision making. In this area, the traditional algorithms are statistical [3], empirical reasoning [4], a voting method [5], Bayesian inference [6], template method [7], and adaptive neural network [8]. Such regular methods can settle the decision fusion of multi-sensor information to some degree. Data association and track-to-track association, two vital problems in single-sensor and multi-sensor multi-target tracking, multi-object tracking is a central computer vision task with a wide variety of real-life applications which ranges from surveillance and monitoring to biomedical video analysis. Multi-object tracking is a challenging task because of complications caused by object appearance changes, complex object dynamics, clutter in the situation, and partial or full occlusions. Nguyen et al. [9] used a novel framework for the road estimation task through the incorporation of reliability into the multi-source fusion and the integration of an offline-trained knowledge base for the reliability assessment represented by Bayesian Network or Random Forests. Jing et al. [10] used a new algorithm for multi-sensor multi-target joint detection, tracking, and classification problems. A constitutional multi-sensor multi-target state estimator was derived, and the optimal solution was obtained based on the minimum Bayes risk criterion. Emami et al. [11] addressed the representation learning techniques for multi-sensor uses and concluded by presenting an overview of available multi-target tracking benchmarks. Fang et al. [12] suggested the Recurrent Autoregressive Network (RAN), which was a temporal generative modeling framework for characterizing the appearance and motion dynamics of multiple obstacles over time. The target detection and tracking fusion algorithm according to a minimum cost function was proposed to decrease the false alarm rate of the target in [13]. Demetriveski et al. [14] presented a novel 2D–3D pedestrian tracker designed for uses in autonomous vehicles. It employed Camera and LIDAR data fusion in order to solve the association problem in which the optimal solution was found by matching 2D and 3D detections to tracks through a joint log-likelihood observation model. Zhao et al. [15] searched for fundamental concepts, solution algorithms, and application guidance related to the use of infrastructure-based LIDAR sensors. Lee et al. [16] proposed the Permutation Matrix Track Association (PMTA) algorithm for supporting track-to-track, multi-sensor data fusion for multiple targets in an autonomous driving system. Zhang et al. [17] presented a Multi-Perspective Tracking (MPT) framework for smart vehicles. An iterative search procedure was proposed to relate detections and tracks from various perspectives. Yoon et al. [18] suggested a new deep neural network (DNN) architecture which could solve the data association problem with a variable number of both tracks and detections, involving false positives. Shakarji et al. [19] proposed a time-efficient detection-based multi-object tracking system through a three-step cascaded data association scheme that combined a fast spatial distance only short-term data association. Such researches focused on the multi-object tracking system.

A preference of the proposed method at the detection level was that describing the obstacles can be improved by adding knowledge from various sensor sources. For instance, LIDAR data can give a reasonable estimate of the distance to the object and its apparent size. Furthermore, classification information, typically obtained from a camera, lets making assumptions about the detected obstacles. An early enrichment of obstacles' descriptions could let the decrease of the number of false detections and integrate classification as a considerable component of the perception output instead of only an add-on. The problem of online multi-object tracking and classifying this study was to reliably relate obstacle trajectories with detections in each video frame and LIDAR signal according to their tracking and classifying information.

2. OBSTACLE TRACKING

Tracking obstacles refer to the process of connecting two detected obstacles in two consecutive frames. The relationship between two obstacles of i and j in two sequential frames is regarded as the Hij hypothesis. Each source (LIDAR, camera) has various attribute vectors for detecting obstacles. For instance, the camera cannot recognize the distance from obstacles. The camera can detect the obstacles using the image processing capability in terms of their geometric characteristics like width and transverse movement. The camera can run the segmentation through image processing algorithms and calculation of horizontal displacement and the horizontal velocity. LIDAR and Camera send displacement and velocity data as raw data to sensor fusion unit, and they should calculate the likelihood and confidence level of the probability of every hypothesis according to the raw data received.

Table1. raw data to sensor fusion unit

Obstacle Characteristics	Symbol
Horizontal displacement	Calculated by Camera
Vertical displacement	Calculated by LIDAR
Horizontal velocity	Calculated by the Camera
Vertical velocity	Calculated by LIDAR

The goal is to decide on a method that shows that how much the two feature vectors calculated in two consecutive frames are close to each other. We used Mahalanobis Distance in this study to examine the relationship between the obstacles. Two factors are calculated separately for Camera, and LIDAR and the mass function is estimated to relate the obstacles for camera and LIDAR. Each hypothesis is rejected or confirmed and calculated based on the values obtained in each current frame and having the values measured in the previous frame.

3. SIMULTANEOUS, CLASSIFICATION, AND TRACTION OF MOVING OBSTACLES

In this article, we are going to propose a method for detecting and tracking moving obstacles in an unmanned ground vehicle. In the unmanned ground vehicle, the LIDAR sensor and camera gather the obstacles information at any frame and send them to the data fusion unit. There is information in each frame of obstacles(Fig.1).

The information of the detected obstacles taken from two sources is different, and we intend to use the information on classification and tracking at the same time. The camera can classify obstacles with image processing. The HOG algorithm and the SVM classifier are the best methods used to detect humans from vehicles [20], which were used for the classification process in this paper .In such a situation, first, the sensor data fusion is done by the Mahalonobis distance, then, the Bayesian theory is used to classify and track obstacles.

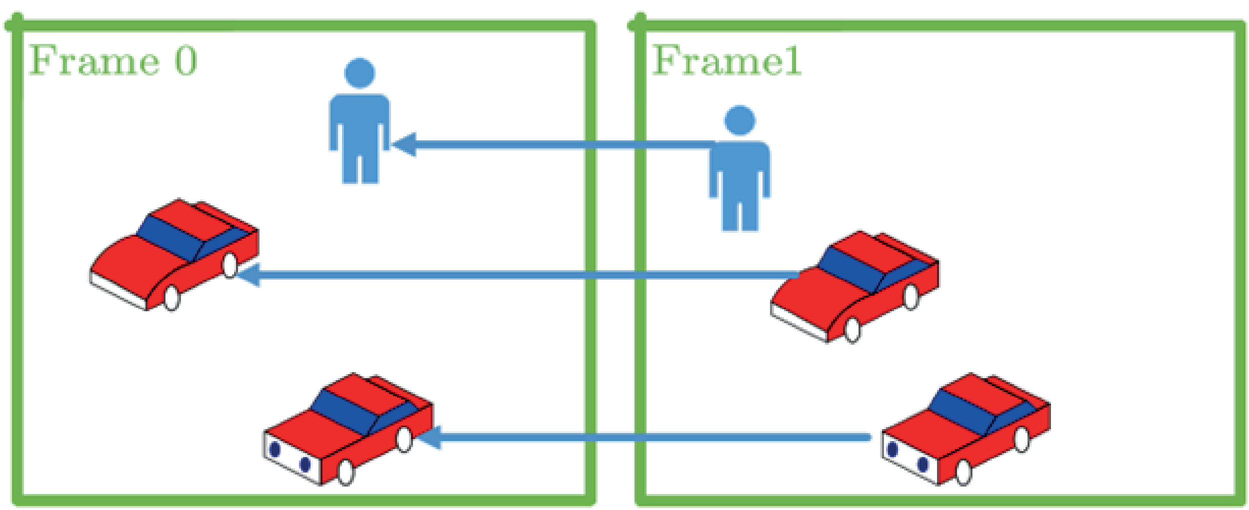

Fig.1. Example of Association problem

Step one: In frame zero (elementary frame), the obstacle classification and tracking object are independently performed by the camera and LIDAR sensor. Thus, for each object, the tracking process is done, and at each frame, the classification information obtained from the camera is expressed as C = {human, vehicle}. The camera and LIDAR sensor also assign a label O = {Hi1, Hi2, Hi3, ..., Hin} for each object in each frame. Hij is the connection hypothesis of the track i to target j.

In the proposed method, the camera and LIDAR sensor, in each frame, present their data for each object as a connection with the observed obstacles in the previous frame.

Step 2: At this step, in the new frame, the probability of tracking hypotheses space {Hi1, Hi2, Hi3, ..., Hin}is calculated with the Mahalonobis distance, and the probability of the class of each obstacle, in the same frame, is obtained in hypothesis space {human, vehicle}.

Step 3: At this step, the Bayesian algorithm is used in order to engage the obstacle classification information in the obstacle tracking information that is derived from the second step. The third step of the algorithm is divided into several sub-sections:

3-1: The algorithm uses two steps for updating classification and tracking probabilities. In the first step, the obstacle identification information is used for updating the anterior tracking obstacle probabilities. At this step, the information obtained from the previous frame is used to associate track to target. For each obstacle that is seen in the new frame, the combination of tracking and classifying probabilities.

To calculate each of the probabilities and update them, it is necessary to include the classification information in their probability value. The obstacle classification information is used to update posterior associating probabilities.

3-2: At this step, we intend to use the probabilities and use them in the Bayesian algorithm. We define the conditional mass function in theorem 1.

Theorem 1: The conditional mass function is:

$$\mu_t(H_{ij}|C_{ik}) = \frac{\mu_t(H_{ij}) * \mu_{t-1}(C_{ik}|H_{ij})}{\sum_{j=1}^{n} \mu_{t-1}(C_{ik}|H_{ij})\mu_t(H_{ij})} \quad \textbf{(1)}$$

In which the term is updated in each frame, and introduces the classification information of the obstacles in their tracking process. Its calculation is as follows:

$$\mu_t(C_{ik}|H_{ij}) = \frac{\mu_{t-1}(H_{ij}|C_{ik}) * \mu_t(C_{ik})}{\sum_{k=1}^{2} \mu_{t-1}\left((H_{ij})|C_{ik}\right) * \mu_t(C_{ik})} \quad \textbf{(2)}$$

3-3: In each frame, 2*n probability values must be calculated for each track where n is the number of detected obstacles that are seen in the previous frame. In this regard, the greatest probability of the obstacle's association and its class is considered simultaneously. For each obstacle in each frame, the highest possibility is regarded as the probability assigned to the option.

$$\mu_t(H_{ij} \mid C_k) = max(\mu_t^{l,c}(H_{ij}|C_k)) \quad \textit{(3)}$$

The Pseudo codes of the *proposed algorithm are as follows (Fig.2, Fig.3):*

tracking obstacles
Input: $\mu_{t-1}(C_{ik}\|H_{ij}), \mu_t^{l,c}(H_{ij})$ ***Output***: $\mu_t(H_{ij} \mid C_k), \mu total$ ***For (all existing tracks)*** *{* $\mu_t^l(H_{ij}\|C_{ik}) \leftarrow \frac{\mu_t(H_{ij}) * \mu_{t-1}(C_{ik}\|H_{ij})}{\sum_{J=1}^{n} \mu_{t-1}(C_{ik}\|H_{iJ})\mu_t(H_{iJ})}$ $\mu_t(H_{ij} \mid C_k) \leftarrow max(\mu_t^l((H_{ij})\|C_{ik}))$ *}* ***End procedure*** *}*

Fig.2. The proposed tracking obstacles algorithm

classifying obstacles
Input: $\mu_{t-1}(C_{ik}\|H_{ij}), \mu_t^{l}(H_{ij})$ ***Output***: $\mu_t(H_{ij} \mid C_k), \mu total$ ***For (all existing tracks)*** *{* $\mu_t(C_{ik}\|H_{ij}) \leftarrow \frac{\mu_{t-1}(H_{ij}\|C_{ik}) * \mu_t(C_{ik})}{\sum_{k=1}^{m} \mu_{t-1}\left((H_{iJ})\|C_{ik}\right) * \mu_t(C_{ik})}$ $\mu_t(C_k) \leftarrow max(\mu_t(C_{ik}\|H_{ij}))$ *}* ***End procedure*** *}*

Fig.3. The proposed classifying obstacles algorithm

In this article, we proposed a method to detect and track moving obstacles in an unmanned ground vehicle. In such a vehicle, the LIDAR sensor and camera gathered the obstacles information at any frame and sent them to the data fusion unit. There was information in each frame of the obstacles. The LIDAR sensor was able to detect the position of the obstacles. Yet, the camera was not able to detect the distance of the obstacles; however, it could classify them based on image processing techniques. Fig.4 shows the flowchart of the proposed method for tracking and classifying based on images and LIDAR signal when driving. The proposed method includes two processing steps: tracking, and classifying. Additionally, a preliminary step is required for this flowchart of the extracting the camera and LIDAR parameters.

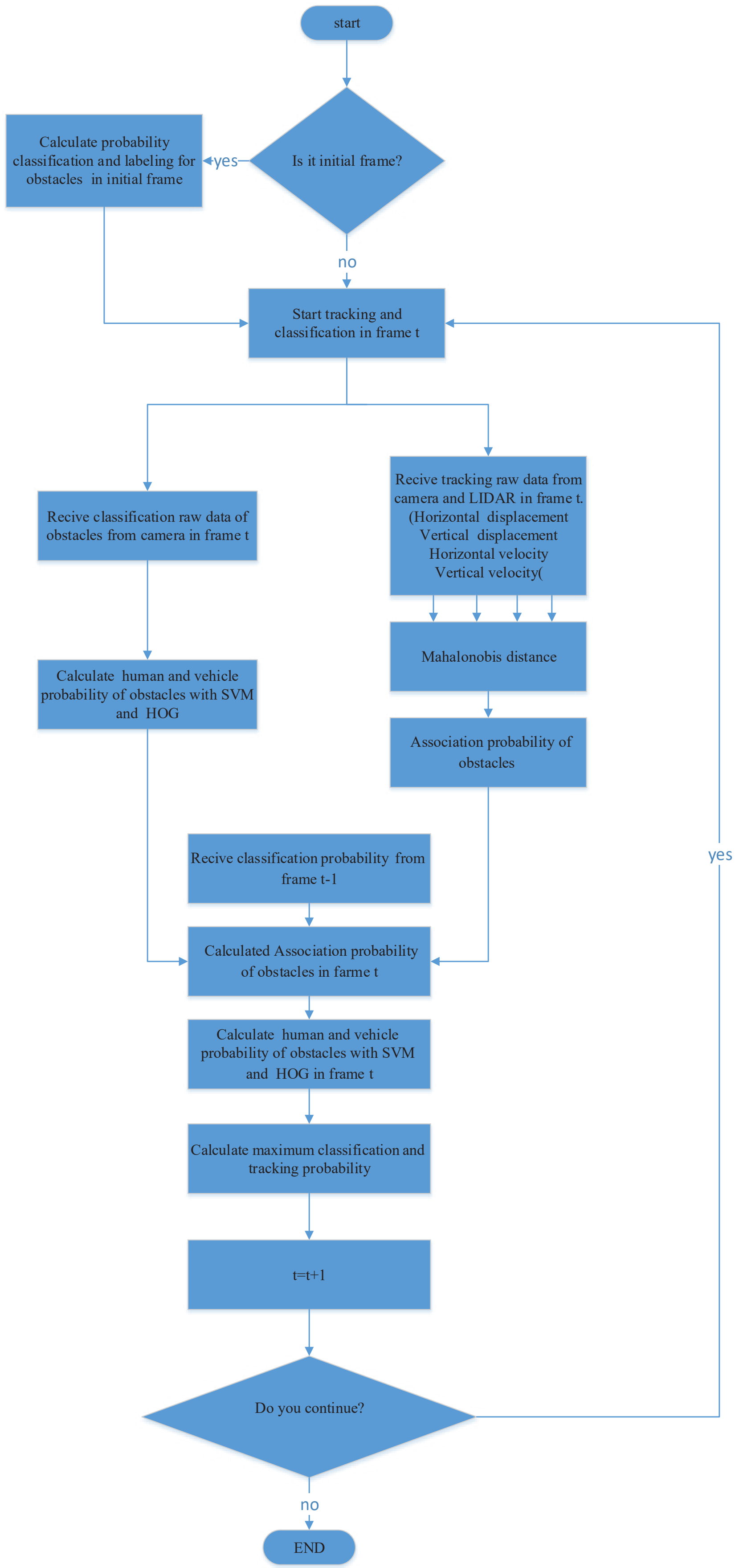

Fig.4. The proposed flowchart

4. EVALUATION

For the quantitative evaluation of the proposed method, we generated a benchmark set using seven different frames. To evaluate the results, the MATLAB software has been used. There are two obstacles in each frame, and for each object, there are two possibilities for classifying and tracking the obstacle. In each frame, the obstacle tracking probability is available in fig.5, , and classifying probability is shown in fig6.

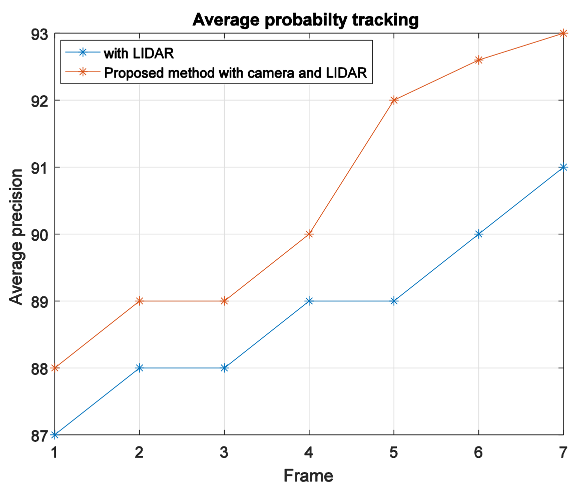

Fig.5. Probability tracking

The experimental results in Fig. 5, , show that the AP for the proposed method with LIDAR and camera increases with the iterations, while there are not significant changes with LIDAR. However, the results show an interesting trend, first increasing slowly until the 5th iteration, and then increasing well above any other combination. This final outcome shows the advantage of our multi-sensor system, which can eventually improve the online transfer learning process.

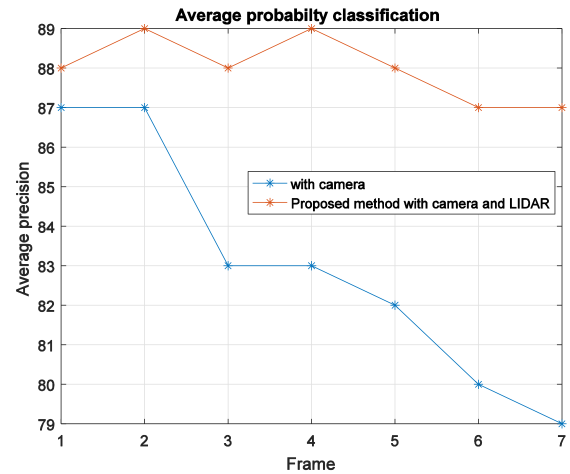

Fig.6. Probability classifying

This result demonstrates that the camera cannot completely measure obstacles located far from the UGV. Further, also the gap between scans of two laser beams is widely spread according to the distances, obstacles corresponding to pedestrians and vehicles could be missed.

5. CONCLUSION

In this paper, we proposed a new data association approach for multi-object tracking and classifying. The association probabilities are calculated by the Mahalonobis distance. For the quantitative evaluation of the proposed method, we generated a benchmark set using seven frames. The results show that by using the proposed method, the RMSE index has decreased. The probability of classifying is improved by decreasing the variance classifying signal. Simulation results of obstacle detection show the advantages of the proposed method in classifying and tracking obstacles.

REFERENCES

1. D. Gruyer, “Etude du traitement de données imparfaites pour le suivi multi-objets: Application aux situations routières,” Compiègne, 1999.
2. V. Schmidlin, “Poursuite multicible multicapteur à l’aide de réseaux neuronaux. Application à la poursuite de cibles aériennes,” Nice, 1994.
3. P. A. Samara, G. N. Fouskitakis, J. S. Sakellariou, and S. D. Fassois, “A statistical method for the detection of sensor abrupt faults in aircraft control systems,” IEEE Transactions on Control Systems Technology, vol. 16, pp. 789-798, 2008.
4. X. Zhu, “Fundamentals of applied information theory,” Beij ing: Tsinghua Universi ty Pres s, 2000.
5. M. Truchon, “Borda and the maximum likelihood approach to vote aggregation,” Mathematical Social Sciences, vol. 55, pp. 96-102, 2008.
6. Z.-J. Zhou, C.-H. Hu, D.-L. Xu, J.-B. Yang, and D.-H. Zhou, “Bayesian reasoning approach based recursive algorithm for online updating belief rule based expert system of pipeline leak detection,” Expert Systems with Applications, vol. 38, pp. 3937-3943, 2011.
7. S.-H. Oh, “Improving the error backpropagation algorithm with a modified error function,” IEEE Transactions on Neural Networks, vol. 8, pp. 799-803, 1997.
8. Y. Deng, “Generalized evidence theory,” Applied Intelligence, vol. 43, pp. 530-543, 2015.
9. T. T. Nguyen, J. Spehr, D. Vock, M. Baum, S. Zug, and R. Kruse, “A general reliability-aware fusion concept using DST and supervised learning with its applications in multi-source road estimation,” in 2018 IEEE Intelligent Vehicles Symposium (IV), 2018, pp. 597-604.
10. Z. Jing, M. Li, and H. Leung, “Multi-target joint detection, tracking and classification based on random finite set for aerospace applications,” Aerospace Systems, vol. 1, pp. 1-12, 2018.
11. P. Emami, P. M. Pardalos, L. Elefteriadou, and S. Ranka, “Machine learning methods for solving assignment problems in multi-target tracking,” arXiv preprint arXiv:1802.06897, 2018.
12. K. Fang, Y. Xiang, X. Li, and S. Savarese, “Recurrent autoregressive networks for online multi-object tracking,” in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 466-475.
13. S. Wang, Z. Ge, G. Lv, and K. Lu, “Research on Multi-Target Stable Tracking Algorithm Based on Detection and Tracking Fusion,” in Journal of Physics: Conference Series, 2018, p. 012175.
14. M. Dimitrievski, P. Veelaert, and W. Philips, “Behavioral pedestrian tracking using a camera and LiDAR sensors on a moving vehicle,” Sensors, vol. 19, p. 391, 2019.
15. J. Zhao, H. Xu, H. Liu, J. Wu, Y. Zheng, and D. Wu, “Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors,” Transportation research part C: emerging technologies, vol. 100, pp. 68-87, 2019.
16. K.-H. Lee, Y. Kanzawa, M. Derry, and M. R. James, “Multi-Target Track-to-Track Fusion Based on Permutation Matrix Track Association,” in 2018 IEEE Intelligent Vehicles Symposium (IV), 2018, pp. 465-470.
17. X. Ji, G. Zhang, X. Chen, and Q. Guo, “Multi-perspective tracking for intelligent vehicle,” IEEE Transactions on Intelligent Transportation Systems, vol. 19, pp. 518-529, 2018.
18. K. Yoon, D. Y. Kim, Y.-C. Yoon, and M. Jeon, “Data association for multi-object tracking via deep neural networks,” Sensors, vol. 19, p. 559, 2019.
19. N. M. Al-Shakarji, F. Bunyak, G. Seetharaman, and K. Palaniappan, “Multi-object Tracking Cascade with Multi-Step Data Association and Occlusion Handling,” in 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018, pp. 1-6.
20. P. Viola and M. Jones, “Robust real-time object detection,” International journal of computer vision, vol. 4, p. 4, 2001.

21. D. Z. Wang and I. Posner, "Voting for Voting in Online Point Cloud Object Detection," in Robotics: Science and Systems, 2015, p. 10.15607.
22. A. Geiger, C. Wojek, and R. Urtasun, "Joint 3d estimation of obstacles and scene layout," in Advances in Neural Information Processing Systems, 2011, pp. 1467-1475.
23. C. Premebida, J. Carreira, J. Batista, and U. Nunes, "Pedestrian detection combining rgb and dense lidar data," in 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014, pp. 4112-4117.
24. A. González, G. Villalonga, J. Xu, D. Vázquez, J. Amores, and A. M. López, "Multiview random forest of local experts combining rgb and lidar data for pedestrian detection," in 2015 IEEE Intelligent Vehicles Symposium (IV), 2015, pp. 356-361.
25. X. Chen, K. Kundu, Y. Zhu, A. G. Berneshawi, H. Ma, S. Fidler, et al., "3d object proposals for accurate object class detection," in Advances in Neural Information Processing Systems, 2015, pp. 424-432.