Electronic Preprint for Journal of Information Processing Vol.29

Paper on Consumer Systems

Heterogeneous Sensor Fusion with GMPHD for Environmentally Adaptable Obstacle Detection in Mobility Systems

Chow Man Yiu[1,a)] Mitsuhiro Kitani[1,b)]

Received: October 30, 2020, Accepted: March 6, 2021

Abstract: Obstacle detection is an essential process in consumer's autonomous mobility systems such as autonomous vehicles inside the dedicated lane to acquire the location of obstacles, and it has become a popular topic in this decade with the blooming of various object detection algorithms and the enhancement of sensor quality. To maintain high accuracy of obstacles' detection in mobility systems outdoor, a sensor fusion system is required to essentially support environmental influence such as lousy weather as well as high moving speeds and adaptably deal with clutter and miss detection based on the incoming measurements from heterogenous sensors with Camera, LiDAR and Radar. Since no current literature about Gaussian mixture probability hypothesis density (GMPHD) handles the above low accuracy fusion problem due to environmental influence for heterogeneous sensors, we propose the concept of integrating GMPHD to heterogeneous sensor fusion with three architectures, Track-to-Track-Fusion (T2TF), Measurement-to-Track-Fusion (M2TF) and Track-to-Association-Fusion (T2AF) and further evaluate their performances respectively in terms of their fusion improvement abilities to determine their practicalities for mobility systems by using the simulation datasets which reproduce ordinary and poorer conditions with the degradation of sensors' performance in the assumption of environmental influences.

Keywords: obstacle detection, mobility systems, GMPHD, heterogeneous sensor fusion, T2TF, M2TF, T2AF

1. Introduction

For autonomous navigation in mobility systems, the full and precise comprehension of the obstacle position in obstacle detection is of paramount importance to make a proper driving decision to maneuver the car safely to a destination. To strengthen the obstacle detection dedicated for the autonomous mobility systems under the adverse environmental influence, several types of sensors including camera, Lidar, and millimeter-wave Radar were exploited to formalize the standard sensor system by previous researchers in a bid to compensate for the deficiencies from each of those sensors and to fulfill better detection quality accordingly [1], [2] Even though such system with heterogeneous sensors could well assist itself to acquire necessary measurements for detecting the surrounding obstacle, the perfect comprehension of the obstacle with high accuracy in such deployment for the fusion has become more challenging when the performance of sensors are still at the same time subject to the environmental influence along with the existence of clutter, scattering miss detection and other miscellaneous sensors' inabilities such as the measurement error and narrow Field of View (FOV). In other words, there are still several arduous issues for sensor fusion system to deal with for the enhancement of the sensor measurement reliability by adapting the different environments in mobility systems along with the problems of reducing the false alarm, estimating the accurate positions of the objects as well as interpolating the missed detection. Furthermore, the adoption of multi-heterogenous sensors with Camera, Radar, and LiDAR in the fusion system implies the special necessity to handle several types of sensors with different detection properties during fusion when the representations and qualities of the measurements from those types of sensors are usually dissimilar from each other. Therefore, sensor fusion with heterogeneous sensors in mobility system has become a prominent subject to resolve the problems of detection accuracy and the effective association between incoming measurements from diverse types of sensors for a good tolerance of the possible change of sensor properties due to environmental influence including system moving speed and poor weather with low illumination.

In this decade, several pieces of research regarding sensor fusion with Gaussian mixture probability hypothesis density (GMPHD) are published to solve the above detection reliability problems that we are concerned for an only homogenous type of sensor due to its original structure for a single type of sensor model. [3], [12] GMPHD is attributed to a new emerging paradigm of Random Finite Set (RFS) based on the rigorous mathematical foundation for stochastic multi-object problems—point process theory [4]. Among all prevalent RFS-based algorithms, GMPHD demands relatively low computational load without explicitly requiring additional data association to obtain a closed-form PHD

1 Center for Technology Innovation – Systems Engineering, Hitachi Ltd. Research and Development Group, Yokohama, Kanagawa 244–0817, Japan
a) manyiu.chow.dv@hitachi.com
b) mitsuhiro.kitani.jd@hitachi.com

recursion [5] for fair tracking performance by reducing the clutter and enhancing the detection accuracy and we think that it is suitable for our problem with outdoor mobility system. Thus, it was chosen as our primary approach to developing multi-object tracking-based sensor fusion. As the majority of the latest researches only focus on the GMPHD fusion problem of the stationary environment with a homogenous type of sensor, Our contribution in this paper is to specifically address heterogeneous sensor fusion problems due to environmental influence and propose new architectures which integrate GMPHD technique for the heterogeneous sensors in the application of consumer mobility systems with an improvement of detection accuracy by estimating the correct position of obstacle and counteracting the fault of each sensor under the adverse environmental influence. We further studied its applicability to heterogeneous sensor fusion in mobility environment moving at different speeds and situating in different illumination to determine whether the best architecture could address the need of the consumer autonomous mobility systems and enhance the detection performance in terms of improvement rate among our approaches of three architectures with GMPHD. The contribution of our research will be helpful for the development of obstacle detection algorithm in fundamental autonomous mobility systems such as autonomous driving vehicles and buses inside the dedicated lane which only requires the location of the nearest objects when it is challenging for the existing fusion system to comprehend the location of obstacle accurately based on the problems mentioned above.

This paper aims to present the architectures of heterogeneous sensor fusion integrated with GMPHD in the application of mobility systems such as autonomous driving vehicles and buses inside the dedicated lane for the consumers. In Section 2, the selection of the tracking algorithm for sensor fusion and the theory of GMPHD are presented. In Section 3, the general problems to the heterogeneous sensor fusion are illustrated to emphasize the challenges and the issues in which our sensor fusion in the mobility systems might suffer. In Section 4, three proposed architectures with GMPHD integration for heterogeneous sensor fusion are explained in terms of their implementations to unravel the fusion problem. In Section 5, the simulations of different environments for mobility systems are described. We examined the improvement effectiveness of three architectures for sensor fusion respectively under the assumption of environmental influence and further discuss the impact of our result on consumer products. In Section 6, the existing related works to the problem of sensor fusion are presented.

2. Gaussian Mixture Probability Hypothesis Density for Sensor Fusion

2.1 The Selection of Tracking Algorithms for Sensor Fusion

To estimate the object location with incoming measured object points from sundry sensors, tracking is a crucial step to associate all those points for the same object from different sensors. There are several popular data association approaches for tracking algorithms available in the past. The general ways include Global nearest neighbor (GNN), joint probability data association (JPDA), and Multiple hypothesis tracking (MHT). Despite the high availability of the existing data association approaches for tracking the measurement points, the computational cost for the real-time result and the tracking performance for the correct association is also one of the concerning issues that we are facing during tracking, and many researchers in the past tried to solve the problem by balancing these two aspects. For instance, classical MHT cannot run in real-time and demands enormous computational resources due to the accumulation of hypotheses from the pedigree of association history even though it has more superior performance than JPDA and GNN and comparable to state-of-the-art methods in recent years [16]. As a result, a considerable number of techniques to handle its computational problem in MHT had been proposed. For example, Fast MHT algorithm [17] endeavored to resolve the computational intractability issue in the original MHT by clustering and eliminating unlikely hypotheses. Another example of a solution to the problem in MHT, the roll-out algorithm [18], was exploited to maximize the measurement-to-track association likelihood and enhance the time efficiency of MHT, Tabu Search and Gibbs sampling [19] enhanced the tracking performance and further improved the computational efficiency in MHT. Compared to MHT, a similar intractable problem also happens in JPDA even though JPDA is considered as the approach which has relatively worse tracking performance but better computational efficiency due to its less combinatorial complexity. One of the papers in the past proposed a JPDA embedment with a simple tracking framework to reduce its processing time [20]. For GNN, it requires the least computational cost, but it could only perform well in a less cluttered environment. Therefore, the GNN data association is only adopted for a simple case with fewer clutters from the data measurements and some researchers proposed Suboptimal Nearest Neighbor (SNN) to improve the tracking performance of the GNN-based method [21].

Table 1 Performance comparison for autonomous Angles-Only Multi-Target Tracking.

Algorithm	Precision (%)	Runtime (millisecond /image)
GNN	98.24	40
JPDA	98.40	**33***
MHT	96.17	70
PHD	**99.27***	42

As data association involves the tradeoff between computational cost and implementation complexity, the tracking algorithm with good balance is, therefore, the critical criteria for our selection of sensor fusion algorithm to support a higher number of incoming measurements. In our study, PHD (probability hypothesis density) method has become popular in this decade and it was proven efficient. Based on **Table 1** which shows a performance comparison of the precision and runtime for tracking 3 targets with low clutter environment on a single image by different Multi-Target Tracking algorithms in MATLAB, [22] PHD (probability hypothesis density) method keeps fair computing time when it maintains higher precision among all the methods and this result means it has relatively strong ability to remove the false alarm measurement. Among existing implementations of PHD method, RFS-based GMPHD is prominent for its detection accuracy improvement with fair computational time and we therefore only focus on the algorithm as our main component for the sensor fusion in the application of mobility systems.

2.2 The Theory of Gaussian Mixture Probability Hypothesis Density for Sensor Fusion

GMPHD is an analytic solution to the PHD recursion under Gaussian assumption and PHD is an approximation to multitarget Bayes filter with the first-order statistical moment of the multi-target posterior density [22]. The derivation of PHD filter is at first provided to understand its fundamental concept before moving on to its approximation in the view of computational tractability.

Suppose the random finite set for multi-target (estimation target) set X_k and multi-target observation (measurement) set Z_k at time k are as follows [6],

$$X_k = \{x_{k,1}, x_{k,2}, \ldots, x_{k,M(k)}\} \in \mathcal{F}(\chi) \tag{1}$$

$$Z_k = \{z_{k,1}, z_{k,2}, \ldots, z_{k,N(k)}\} \in \mathcal{F}(\mathcal{Z}) \tag{2}$$

where $\mathcal{F}(\chi)$ is the collections of all finite subsets of target states χ with $M(k)$ states and $\mathcal{F}(\mathcal{Z})$ is the collections of all finite subsets of observation states $\mathcal{Z}$ with $N(k)$ states. Each target $x_{k-1,\cdot}$ in multi-target set X_{k-1} generates a Bernoulli RFS $S_{k|k-1}(x_{k-1})$ at time k with survival probability $p_{S,k}(x_{k-1})$ and new targets at time k are modeled by an RFS of spontaneous births Γ_k. Hence, the multi-target state X_k at time k according to the previous state X_{k-1} [5], [6],

$$X_k = \bigcup_{x_{k-1} \in X_{k-1}} S_{k|k-1}(x_{k-1}) \cup \Gamma_k \tag{3}$$

Similarly, each measurement $z_{k,\cdot}$ in observation set Z_k is generated by Bernoulli RFS $D_k(x_k)$ with detection probability $p_{D,k}(x_k)$ based on each target x_k in the set X_k at time k and spurious measurement set F_k [5], [6],

$$Z_k = \bigcup_{x_k \in X_k} D_k(x_k) \cup F_k \tag{4}$$

Based on the theory of Bayes recursion with multi-target set X_k and multi-target observation set Z_k at time k, the optimal multi-target Bayes filter is derived given by the recursion as follows,

$$\begin{aligned} &p_{k|k-1}(X_k \mid Z_{1:k-1}) \\ &= \int f_{k|k-1}(X_k \mid X) p_{k-1}(X \mid Z_{1:k-1}) \mu_s(dX) \end{aligned} \tag{5}$$

$$p_k(X_k \mid Z_{1:k}) = \frac{g_k(Z_k \mid X_k) p_{k|k-1}(X \mid Z_{1:k-1})}{\int g_k(Z_k \mid X) p_{k|k-1}(X \mid Z_{1:k-1}) \mu_s(dX)} \tag{6}$$

where $p_k(\cdot \mid Z_{1:k})$ is the multi-target posterior density, $p_{k|k-1}(\cdot \mid \cdot)$ is the multi-target prior density $f_{k|k-1}(\cdot \mid \cdot)$ is the multi-target transition density, $g_{k|k-1}(\cdot \mid \cdot)$ is the multi-target likelihood and μ_s is an appropriate reference measure on the subset F_k [6], [24].

However, multi-target Bayes filter is computationally intractable and it only works when the number of targets is small [6], [25], various approximations such as Sequential Monte Carlo (SMC), Cardinalized probability hypothesis density (CPHD), multi-Bernoulli, PHD, and Dynamic factorization have been proposed in the past [24], [26], [27]. As PHD is more mature, swifter, and more computationally efficient compared to the rest of other existing approximation tactics [5], [24], [26], [28], we only remark on this filter and further elaborate it under the linear Gaussian multi-target model.

The PHD filter propagates a first-order statistical moment of the multi-target posterior [22] with the theory of finite-set statistics (FISST) to approximate the optimal multitarget Bayes filtering in the recursion (5) and (6). FISST is a systematic, unified, and intuitive approach to multi-sensor-multi-target detection, tracking, and information fusion based on the mathematical foundation for stochastic multi-object problems, point process theory [4], [29]. Thus, the following approximated intensities ν_k and $\nu_{k|k-1}$ are approximated with the first moment of multi-target posterior density p_k from Eq. (6) and multi-target predicted density $p_{k|k-1}$ from Eq. (5) respectively through PHD recursion,

$$\begin{aligned} \upsilon_{k|k-1}(x) = &\int p_{S,k}(x_{k-1}) f_{k|k-1}(x_k \mid x_{k-1}) \upsilon_{k-1} dx_{k-1} \\ &+ \gamma_k(x_k) \end{aligned} \tag{7}$$

$$\begin{aligned} &\upsilon_k(x) \\ &= [1 - p_{D,k}(x_k)] \upsilon_{k|k-1}(x_k) \\ &\quad + \sum_{z \in Z_k} \frac{p_{D,k}(x_k) g_k(z \mid x_k) \upsilon_{k|k-1}(x_k)}{\kappa_k(z) + \int p_{D,k}(x_{k-1}) g_k(z \mid x_{k-1}) \upsilon_{k|k-1}(x_{k-1}) dx_{k-1}} \end{aligned} \tag{8}$$

where $f_{k|k-1}(\cdot \mid \cdot)$ is the multi-target transition density, $g_{k|k-1}(\cdot \mid \cdot)$ is the multi-target likelihood, $p_{S,k}(x_{k-1})$ is a survival probability at time k given that state x at previous time $k-1$, $p_{D,k}(x)$ is a detection probability given state x at time k, $\kappa_k(z)$ is an intensity of clutter RFS F_k and $\gamma_k(x)$ is an intensity of the birth RFS Γ_k [5], [6]. From Eqs. (7) and (8), its approximation based on FISST demonstrates the computationally cheaper approach without combinatorial computations from an unknown association of the Bernoulli RFS [6]. However, PHD filter does not offer any closed-form solution and suffers a curse of dimensionality due to the complexity of numerical integration [6], [30].

To obtain the closed-form solution, particle PHD filters such as Auxiliary particle PHD filter [31] and SMC-PHD filter had been developed in the past but they suffer from highly demanding computational cost even though they support highly nonlinear problems, and they are closer to the original PHD filter. For this reason, we ruled out the approach of particle PHD filters, and our study mainly adopted GMPHD which is another closed-form solution to PHD recursion under the linear Gaussian multitarget model. Equation (9) shows the Gaussian approximation of Eq. (7) with the form of $N(\cdot\,; m, P)$ which represents a Gaussian density with weight w, mean m, covariance P, and the number of components of the Gaussian intensity J.

$$\begin{aligned} &\upsilon_{k|k-1}(x) \\ &= \gamma_k(x) + p_{S,k} \sum_{i=1}^{J_{k-1}} w_{k|k-1}^{(i)} N(x; m_{k|k-1}^{i}, P_{k|k-1}^{(i)}) \end{aligned} \tag{9}$$

$$\gamma_k(x) = \sum_{i=1}^{J_{\Gamma,k}} w_{\Gamma,k}^{(i)} N(x; m_{\Gamma,k}^{i}, P_{\Gamma,k}^{(i)}) \tag{10}$$

Where, $\gamma_k(x)$ in Eq. (10) is the Gaussian mixture birth PHD with weights $w_{\Gamma,k}^{(i)}$, means $m_{\Gamma,k}^{i}$, and covariances $P_{\Gamma,k}^{(i)}$ of each Gaussian

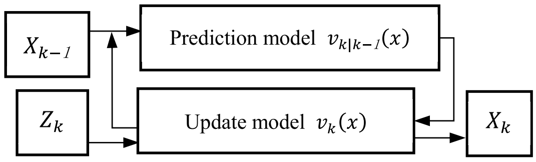

Fig. 1 The flow of Kalman filter for GMPHD estimation.

mixture hypothesis respectively. $v_{k|k-1}(x)$ in Eq. (9) is a predicted PHD at time k based on the estimated PHD at time $k-1$ with survival rate $p_{S,k}$, weights $w^{(i)}_{k|k-1}$, means $m^{i}_{k|k-1}$, and covariances $P^{(i)}_{k|k-1}$ of each Gaussian mixture hypothesis and combined with Gaussian mixture birth PHD $\gamma_k(x)$.

Meanwhile, the following Eq. (11) shows the Gaussian approximation of Eqs. (8).

$$\begin{aligned} &v_k(x) \\ &= (1-p_{D,k})v_{k|k-1}(x) \\ &\quad + p_{D,k}\sum_{z\in Z_k}\sum_{i=1}^{J_{k|k-1}} w^{(i)}_k(z)N(x; m^i_k(z), P^{(i)}_k) \end{aligned} \tag{11}$$

$$w^{(i)}_k(z) = \frac{w^{(i)}_{k|k-1}(z)q^{(i)}_k(z)}{\lambda_{F,k} + \sum_{l=1}^{J_{k|k-1}} w^{(l)}_{k|k-1}(z)q^{(l)}_k(z)} \tag{12}$$

$$q^{(i)}_k(z) = N(x; H_k m^{(i)}_{k|k-1}, S^{(i)}_{k|k-1})/p_{K,k}(z) \tag{13}$$

Where $v_k(x)$ in Eq. (11) is the posterior PHD with sensor detection rate $p_{D,k}$, the weight $w^{(i)}_k(z)$, mean $m^i_k(z)$ and the covariance $P^{(i)}_k$ of the estimated PHD from the incoming measurement set Z_k and the predicted PHD $v_{k|k-1}(x)$ in Eq. (9).

For simplicity, the technique of standard Kalman filter with both linear prediction model and linear update model for each type of sensor shown in **Fig. 1** is exploited in this paper to acquire the closed-form solution of Bayes filtering recursion under the assumption of linear Gaussian model for $v_{k|k-1}(x)$ in Eq. (9) and $v_k(x)$ in Eq. (11). In Fig. 1, X_{k-1} is the set of estimated object points x_{k-1} at time $k-1$, X_k is the set of estimated object points x_k at time k after filtering and Z_k is the set of sensor measurement points z_k at time k. The calibratable parameters including birth density $\gamma_k(x)$, survival rate $p_{S,k}$, detection rate $p_{D,k}$, clutter rate $p_{K,k}(z)$ in the models are changed based on the object trajectory presumption and sensor measurement properties.

To reduce the redundant computational resource, the Gaussian mixture hypotheses are purged and merged after each fusion at each time step. The criteria for hypothesis purging are to limit the number of hypothesis N_{purge} and to remove the hypothesis with weight below the purging threshold w_{purge}. For hypothesis merging, the criterion is to merge two hypotheses with the difference L less than merging threshold L_{merge}. The formula of difference L is as follows.

$$L(x^1_k, x^2_k) = (m^1_k - m^2_k)(P^1_k)^{-1}(m^1_k - m^2_k) \tag{14}$$

Where m^1_k and P^1_k are the mean and covariance of hypothesis x^1_k respectively and m^2_k is the mean of hypothesis x^2_k at time k.

Despite the credit of fairly good computational efficiency and tracking performance in GMPHD, the critical mission in this paper is the implementation and the practical integration of GMPHD tracking algorithm for heterogeneous sensor fusion architectures including M2TF, T2TF, and T2AF which the details are shown in Section 4 since this problem has rarely been fulfilled in the current works of literature due to the original GMPHD structure which is only customized to homogeneous sensor fusion and merely caters to the same prediction and update models. On the other hand, the uttermost issues are the strategies for the constituent of key models inside GMPHD for multiple dissimilar types of sensors and their performance of detection improvement.

3. Problem Formulation

This section presents the existing issues for the development of multi-target tracking-based heterogeneous sensor fusion (multi-sensor fusion). In Section 3.1, we define the system requirements of heterogeneous senor fusion in mobility systems by explaining three significant environmental influence which generally contributes to the sensor fusion issues in terms of low detection rate (False negative measurements) and clutter (False positive measurements) and they primarily cover most sensor detection issues in mobility systems. In Section 3.2, the problems in heterogeneous Sensor fusion architectures with GMPHD are explained in detail.

3.1 Problems of Sensor Fusion Performance under the Environmental Influence

3.1.1 Sensor Fusion Performance on Moving Mobility Systems

The tracking-based algorithm in the local filtering and a sensor fusion layer function as the role of removing the clutter and recovering objects' trajectories from sensor measurements [5] so as to enhance detection rate and distinguish the true position of object targets from a set of spurious measurements. As the first appearance object detected by sensors in mobility systems is usually moving at a higher speed when the mobility systems are operating, GMPHD implementation from the majority of existing methods in the literature such as Ref. [13] are no longer functioning well with its original default setups when they are always dedicated for the tracking problem in stationary environment and thus those methods easily mistake the incoming sensor measurements moving at higher speed as clutter. Focusing on GMPHD strategy with heterogeneous sensor fusion in mobility systems, the problem for sensor fusion is how the system could still maintain the fair improvement ability of any erroneous detection in the higher speed environment of the mobility systems.

3.1.2 Optical Sensor Detection Rate due to Environmental Illumination

The uncertain detection rate in sensors is always the issue of false-negative detection which miss the object measurements, and we found out that only color optical sensors are commonly susceptible to the problem of the unstable detection rate due to low amount of surrounding light source whereas radioactive sensors would not, and their detection rate are resiliently constant notwithstanding the kinds of environmental influences in most situations. In mobility systems, it is common that the illumination of the environment always fluctuates due to adverse weather conditions and insufficient lighting inside the operation area. This issue leads to the degradation of most color optical sensors' detection performance. As these sensors' properties cannot be stably controlled and change unpredictably in reality due to the influence of light source from the surrounding environment, it becomes our concerning problem for improving the ability of proposed heterogeneous sensor fusion architectures with GMPHD and pre-defined models for the sudden optical sensors' detection rate change. For example, the camera on mobility systems often cannot perform well in a dark environment but radioactive sensors such as Radar and LiDAR still function perfectly well. We would like that our proposed sensor fusion algorithm with GMPHD could still be able to improve the detection error effectively with sudden slight detection rate abatement under this adverse condition.

3.1.3 High Number of Radioactive Sensors' Clutters due to Surrounding Radioactive Noise

The sensor clutter is the sensor issue of false-positive alarms that falsely detect the existence of an object. Among all the sensors, Radar and LiDAR have the common issue of relatively higher random clutter due to the surrounding radiation noise regardless of its resilient detection rate and this is a tremendous issue to deal with and therefore provides accurate and useful object position measurements. An optical sensor with an object detection algorithm does not have such kind of problem as it would not be affected by radioactive noise by only recognizing the object with color and patterns. For this problem, we exploited our proposed sensor fusion algorithm with GMPHD, and our concerned problem is how much random noise/ clutter our different fusion architectures in the application of mobility systems could reduce while fusing with other sensors' measurements to improve the detection accuracy.

3.2 Problems in Heterogeneous Sensor Fusion Architectures

For multi-sensor fusion, the fusion architectures influence not only the communication efficiency owing to the bandwidth requirement but also error certainty estimation and information correlation between measurements. The decision of fusion architectures might therefore affect the ability of detection accuracy improvement due to environmental influence for mobility systems. There are two mainstream architectures of sensor fusion in the past, Centralized fusion architectures, and Hierarchical fusion architectures. Centralized fusion architectures transmit the object measurements from each deployed sensor directly to a global fusion node. (Shown in **Fig. 2**) They provide local stovepiped processing centers that limit network-centric development [15]. This architecture has its relatively simple structure with all the measurements from different types of sensors handled in global track but our concerned problem in this paper is whether this simple structure with GMPHD still provides fair detection accuracy improvement under the environmental influence in mobility systems.

Hierarchical architectures combine all the track estimates from each local centralized fusion processing node, forming a subordinate–superior relationship [15] (Shown in **Fig. 3**). This relationship forms the robust technique which further abates the estimation error and enhances the tracking performance. Even

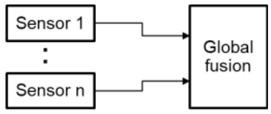

Fig. 2 The structure of centralized fusion architecture.

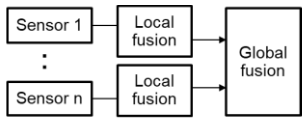

Fig. 3 The structure of hierarchical fusion architecture.

though this architecture has a more complex structure, we are also concerned whether this more complex structure with extra local fusion and the implementation of GMPHD still provides better detection accuracy improvement under the environmental influence in mobility systems.

Despite the availability of these two architectures' foundations, most of those with GMPHD integration in the literature only pay attention to the homogeneous sensor. In other words, the most challenging problem is to handle the choice of architecture which is more adaptive to the issues of high mobility system operational speed in Section 3.1.1, optical sensor issue mentioned in Section 3.1.2, and high radioactive sensor noise in Section 3.1.3 for heterogeneous sensors fusion with GMPHD in mobility systems.

For those architectures with GMPHD in which only homogenous sensors are required, the same prediction models and the same update models for each sensor are usually needed without any substantial change. As a result, no other special arrangement is indispensable in such architecture for better fusion performance as the same models could cater well for multiple same types of sensors with good consistency on local and global track estimation inside the architectures. However, the problem becomes more uncertain and arduous when it comes to heterogeneous sensors. For the case of heterogeneous sensors, each sensor has different properties from each other and the composition of architecture to satisfy different properties of sensors becomes uncertain because of the same global fusion track with incoming measurements in different local tracks from the dissimilar types of sensors with variant FOV, clutter rate, error covariance, and detection rate. Furthermore, the synchronization fusion order for those different prediction and update models based on dissimilar types of sensors' performances and abilities in the architecture is also an influential factor to affect the ability of detection performance improvement.

Therefore, opting for the most appropriate architecture and the techniques of fusion management with GMPHD integration are significant issues in this paper to address the enhancement issue on sensor fusion performance based on tracking estimation and also maintain its performance under the influence of environmental speed as well as the adverse scenario with high adaptability in our sensor fusion for consumers' mobility systems.

3.3 The Scope of Our Performance Requirements

Our research mainly focuses on the tracking performance of the sensor fusion result under the adverse environmental influence based on our chosen GMPHD. Since our chosen method has already been proven to have more efficient computational cost with fair tracking performance compared to other existing methods in Table 1, we expect that the same technique with different architectures should cost a similar amount of computation cost and it is not the main scope of this paper.

Therefore, our main scope of requirement in this paper is the performance comparison of the proposed fusion architectures with GMPHD techniques based on the problems in Section 3.2 with regard to the ability of detection accuracy improvement in terms of low detection rate (False negative measurements) and clutter (False positive measurements) under the adverse environmental influence explained in Section 3.1. Comparing our results, we targeted to find out the best architecture with GMPHD integration for the practical deployment in mobility systems.

4. Proposed Architectures for the Integration of GMPHD to Heterogeneous Sensor Fusion

In this section, the implementation details, and the concerns of three proposed architectures M2TF, T2TF, and T2AF based on the fundamental fusion architectures with the integration of GMPHD and the implementation of standard Kalman filter to heterogeneous sensors are presented. In each architecture with GMPHD, we especially focus on the novel design for the forming of Gaussian components in prediction and update models based on Eqs. (9) and (11) respectively with the implementation of Kalman filter in this section.

4.1 Measurement-to-track Heterogeneous Sensor Fusion (M2TF) with GMPHD Integration

The structure of measurement-to-track fusion (M2TF) is illustrated in **Fig. 4** based on Centralized fusion architectures with the integration of the GMPHD algorithm. In this architecture, the object point measurements from each type of sensor at each time step are sequentially fed into a single GMPHD globally.

Since the original GMPHD algorithm itself does not require any data association to group the corresponding measurements for each of the same target objects, the asynchronistic models are exploited, and object point measurement set from each type of sensor at each time step updates the global fusion track sequentially as the flow of the corresponding functions shown in **Fig. 5**.

In M2TF architecture, only a single GMPHD for global fusion is required. The challenging part of this architecture is to develop two essential models in global fusion with the implementation of GMPHD for heterogeneous sensors with different properties, which models are the global prediction model based on Eq. (9) and the global update model based on Eq. (11).

For global prediction modeling $v^t_{k|k-1}(x)$ of sensor type t, the composition of birth Gaussian components in $\gamma_k(x)$ based on the corresponding sensors' FOV are appended into a global track. The velocity of each birth component should be the same as the speed of mobility systems when the first appearance point is most likely to vary based on the current velocity of the mobility systems. The survival rate $p_{S,k}$ and prediction error covariance $P_{k|k-1}$ remain the same throughout the whole estimation. However, survival rate $p_{S,k}$ should not be too low, and the target states are supposed to survive for a certain period especially when other

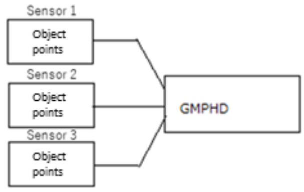

Fig. 4 The structure of measurement-to-track fusion with GMPHD integration.

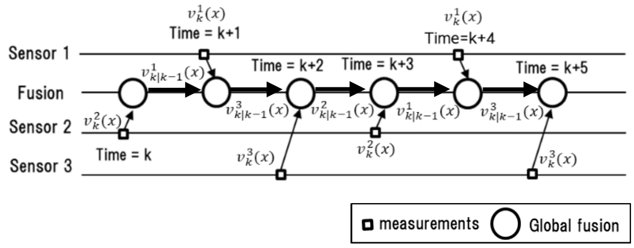

Fig. 5 The measurement management of measurement-to-track fusion with GMPHD integration.

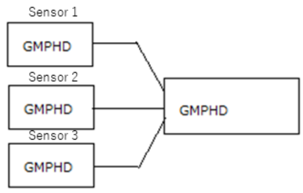

Fig. 6 The structure of track-to-track fusion with GMPHD integration.

types of sensors might not successfully capture the object due to the limited area of FOV. In other words, only the change of birth model $\gamma_k(x)$ is conducted based on the sensor type of incoming measurements for each prediction step in recursion.

For global update model $v^t_k(x)$ of sensor type t, the global track is retrieved for the fusion with incoming measurements Z_k. Based on the sensor type of incoming measurements, the corresponding detection rate $p_{D,k}$, the density of Poisson false alarm $p_{K,k}(z)$ and error covariance $S_{k|k-1}$ are applied to each measurement point and update target states correspondingly.

In this architecture, the number of computation steps is lower as only a single GMPHD is required to estimate all the incoming measurements from disparate sensors, and no extra data association and clustering are needed. Since the global fusion is shared among different sensors, we hypothesize that the slight change of single sensor properties due to the environmental influence would not largely be detrimental to the final fusion performance because of the normal operations from other sensors.

4.2 Track-to-Track Heterogeneous Sensor Fusion (T2TF) with GMPHD Integration

The structure of track-to-track fusion (T2TF) is illustrated in **Fig. 6** based on Hierarchical fusion architectures with the GMPHD algorithm. In this architecture, the raw measurements from each sensor at each time step are pre-filtered with local filtering in advance given that the raw measurements are the point object.

Since this architecture with GMPHD does not require data association to group the corresponding measurement for the same object as M2TF structure does, the asynchronistic models are exploited and the filtered measurements from each type of sensor

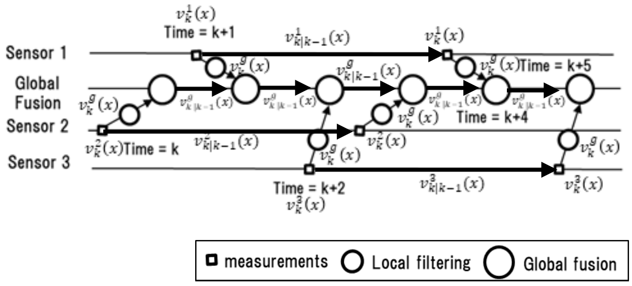

Fig. 7 The measurement management of track-to-track fusion with GMPHD integration.

at each time step sequentially update the global fusion track as shown in **Fig. 7**.

In T2TF architecture for heterogeneous sensors on the mobility systems, GMPHD for local filtering and global fusion are required. In other words, the challenging part of this architecture is not only to build up prediction and update models for accurate global fusion as M2TF illustrates but also the prediction and update models in local filtering for each sensor before fusion.

For each local prediction model $v^t_{k|k-1}(x)$ of sensor type t, the composition of birth Gaussian components in $\gamma_k(x)$ based on the corresponding sensors' FOV are appended into a local track for each sensor and the velocity of each component in the first appearance tend to be the same as the speed of mobility systems. This technique further distinguishes the clutters which potentially have unalike velocity. The survival rate $p_{S,k}$ and prediction error covariance $P_{k|k-1}$ remain the same.

For each local update model $v^t_k(x)$ of sensor type t, the local track from each sensor is retrieved for the fusion with incoming measurements. Based on the sensor type of incoming measurement, the corresponding detection rate $p_{D,k}$, the density of Poisson false alarm $p_{K,k}(z)$ and error covariance $S_{k|k-1}$ are applied to each measurement point and update the target state correspondingly.

For the global prediction model $v^g_{k|k-1}(x)$, all the parameters are the same as the local prediction model $v^t_{k|k-1}(x)$ based on the sensor type t of incoming measurements and the speed of the mobility systems. However, more different configurations are required in the global update model $v^g_k(x)$. As we noticed that the local filtering dedicated for each sensor has already enhanced the accuracy of the local track by reducing clutter and improved detection rate, the corresponding detection rate would be therefore higher, and the density of Poisson false alarm relatively decreases in the global update model to update the global track with an incoming filtered local track from the local filtering in respective heterogeneous sensors for fusion.

This architecture could further swiftly weed out the random clutter because of the strong involvement of two steps filtering and therefore we assume this architecture performs well in a high Signal Noise Ratio (SNR) environment. Regarding the sensor properties change due to the environmental influence, we hypothesize that the global fusion in this architecture would be more versatile to the abnormality in local GMPHD filtering in a bid to maintain a decent quality of measurement improvement under those influences.

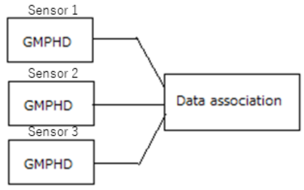

Fig. 8 The structure of track-to-association fusion with GMPHD integration.

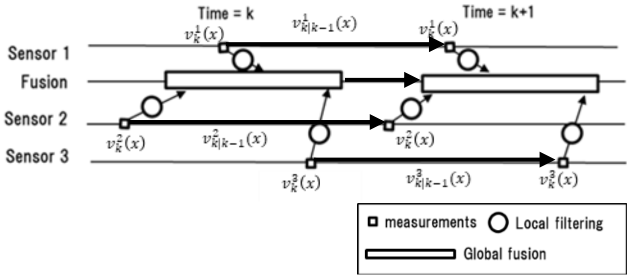

Fig. 9 The measurement management of track-to-association fusion with GMPHD integration.

4.3 Track-to-Association Heterogeneous Sensor Fusion (T2AF) with GMPHD Integration

The structure of track-to-association fusion (T2AF) is illustrated in **Fig. 8** based on Hierarchical fusion architectures with GMPHD and global fusion with data association. In this architecture, the raw measurements from each sensor at each time step are first filtered with local tracking given that the raw measurements are the point object which indicates a single point per object and then sequentially fed into the data association algorithm for the fusion step.

Since this architecture with GMPHD requires data association in the final fusion step to group the corresponding measurement for the same target object, the synchronistic data association models are exploited and the filtered measurements from each type of sensor at each time step collectively update the global fusion track with data association as shown in **Fig. 9**.

The challenging part is to set up two models, the local prediction model, and the local update model for local filtering in respective heterogeneous sensors with the implementation of GMPHD.

The local prediction models $v^t_{k|k-1}(x)$ are specifically built for each type of sensor t. For each prediction model in local filtering in each sensor, the birth models based on information of FOV of respective sensors and the velocity of mobility systems are appended to the existing local track. For each update model $v^t_k(x)$ in local tracking of each sensor t, the detection rate $p_{D,k}$, the clutter rate based on the density of Poisson false alarm $p_{K,k}(z)$ and error covariance $S_{k|k-1}$ for local tracking are set up to update the local track as per the sensor properties. As for the global fusion, GNN is the suggested data association approach to associate filtered points in a high Signal noise ratio environment and we assume local filtering with GMPHD has already efficiently tackled almost all clutters and miss-detection issues beforehand.

The architecture requires fewer computation steps provided that the cheaper data association is utilized. Nevertheless, GMPHD in local filtering from each sensor plays a key role in dealing with the problem of miss detection and clutter in this architecture. In this architecture, we hypothesize that the ability of random clutter removal is weak when data association in global fusion does not have such ability. All the handles of clutter measurements rely on local filtering. However, it might handle the miss detection well to interpolate the points with a single step of filtering during fusion.

5. Simulation and Performance Evaluation

5.1 Simulation Environment for Sensor Measurements

To determine the suitability of whether our three sensor fusion architectures with GMPHD are applicable to the consumer mobility systems product, various scenarios for detection performance evaluation are simulated based on the consideration of three aspects, 1. The nature of target obstacles from initial origins with different moving velocities, 2. Sensor properties (detection rate and clutter number) of Camera, Radar, and LiDAR which imitate the configuration of real hardware in most autonomous mobility system, and 3. Speeds of mobility systems from 0 km/h to 90 km/h based on the standard braking distance and the existing laws for the speed of large vehicles [32].

5.1.1 The Properties of Target Obstacles

For the simulation of the nature of target obstacles, we assume every single object is interpreted as a single measurement point by all the sensors, and three object points are set up to move according to the pre-defined initial position and predefined velocity. The number of objects in the setup is based on the frontest objects which possibly exist.

To determine the trajectory of each object, the constant velocity model with small random noise ε is created to simulate how the objects move in the scene. The constant velocity model is shown as follows in Eq. (15),

$$\begin{pmatrix} X_{1:k} \\ \dot{X}_{1:k} \\ Y_{1:k} \\ \dot{Y}_{1:k} \end{pmatrix} = \begin{pmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{0:k-1} \\ \dot{X}_{0:k-1} \\ Y_{0:k-1} \\ \dot{Y}_{0:k-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1:k} \\ 0 \\ \varepsilon_{1:k} \\ 0 \end{pmatrix} \tag{15}$$

where t is the time duration at each time step, $X_{0:k-1}$ and $Y_{0:k-1}$ are the set of previous position state of $X_{1:k}$ and $Y_{1:k}$ respectively and $\dot{X}_{0:k-1}$ and $\dot{Y}_{0:k-1}$ are the set of previous velocity state of $\dot{X}_{1:k}$ and $\dot{Y}_{1:k}$ respectively.

The assumptions of those objects' properties for all the simulation scenes are shown as follows in **Table 2**.

Based on Eq. (15), the trajectories of three objects are computed throughout the lifetime of sensor detection and their visualization with the moving directions in arrows are shown in **Fig. 10**.

5.1.2 Sensor Properties in Simulation

The sensor deficiency based on the property for its type of sensor is the crucial factor for our sensor measurement simulation when it could reflect the sensor problem which our fusion system needs to address. The simulation of the defects for each sensor includes detection rate, Field of View (FOV), sensor measurement error with error covariance, and clutter rate.

It is worth noting that our clutter rate in the sensors of our simulation is set up according to the number of random clutter appearances per frame. Furthermore, FOV in our simulation is defined by the distance range only for simplicity and each rectangular FOV area for each of the sensors has been illustrated as follows in **Fig. 11**.

Table 2 The assumption of Objects' properties.

	Initial position	Velocity
Object 1	(250, 1)	(-0.9, 0.5)
Object 2	(100, 4)	(-0.5, -1)
Object 3	(150, -6)	(-3, 0.7)

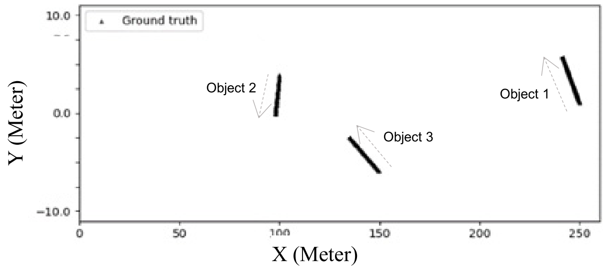

Fig. 10 The visualization of the trajectories of three objects.

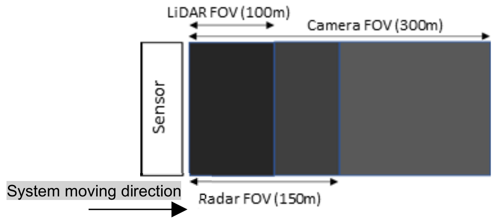

Fig. 11 The illustration of FOV for sensors in simulation.

In autonomous mobility systems, Camera, Radar, and Lidar are usually equipped as the essential sensors to comprehend the positions of obstacles in the vicinity of the system [1]. Therefore, we chose them as our target types of sensors and generate the simulation measurements for our evaluation. For camera measurements, we presume that stereo disparity and the convolutional deep neural network such as YOLO have been applied to decipher the color on raw RGB images and therefore each object and its depth on the image would be interpreted as a single point in the corresponding bird eye view location. For the measurements from LiDAR and Radar, we presume the measured points for each object were combined as a single position point through clustering in the corresponding bird eye view location.

To focus on the problem due to environmental influence in the simulation dataset, the detection problem in Camera based on the problem in Section 3.1.2 is determined by detection rate in the simulation and the simulation of clutter in Radar and LiDAR based on the problem in Section 3.1.3 is determined by the value of clutter rate for the corresponding sensor.

5.1.3 Speed for Mobility System in Simulation

We focus on the speed requirement of vehicles as our target autonomous mobility system in our simulation since it usually has the fastest speed among those systems based on the problem in Section 3.1.1. As the object detection algorithm and sensor in current technology are not able to stably measure the position of the faraway objects up to 300 meters, we would not further simulate the environment with more than 100 kilometers per hour since it requires a longer braking distance which is over 524 meters [32].

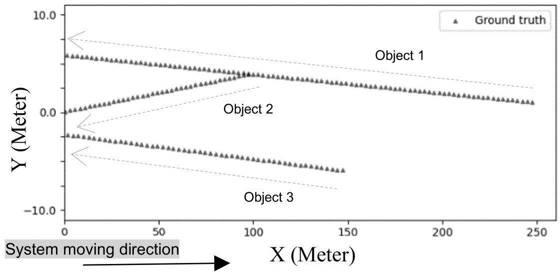

Fig. 12 The visualization of the trajectories of three objects under moving system with the speed 90 km/h.

For the speed of the simulation in our concerned mobility systems, the standard for a low-speed vehicle (LSV) is referred to in this paper because of their short braking distances. According to the Federal motor vehicle safety standards in the U.S., one of the criteria relevant to speed for LSV is "a speed attainable in 1.6 km (1 mile) is more than 32 kilometers per hour and not more than 40 kilometers per hour on a paved level surface [34]. Hence, 40 kilometers per hour is chosen as the simulation speed for our performance evaluation.

To further understand the case when the system runs over reasonable speeds, 60 kilometers per hour and 90 kilometers per hour are also considered in this paper to evaluate how the tracking performs. However, it is irrational to consider the case when the vehicle runs over 100 kilometers per hour due to braking distance and the inadequate FOV mentioned earlier and we, therefore, did not advance those cases in our evaluation. **Figure 12** illustrates the trajectories of 3 corresponding detected objects from Fig. 10 moving at relative velocity involving each original object speed and system speed 90 kilometers per hour. For the purpose of fair evaluation, all corresponding measurements of each object trajectory step in all cases would be also retained under higher system speeds. The trajectory for each object is the same and no measurement would be missed after the system speeds are applied into the simulation. This implies that all simulations captured the whole object trajectories from the actual starting point to ending point but different measurement positions due to the different system moving speeds.

5.2 Performance Evaluation for Sensor Fusion

In our evaluation, Generalized Optimal Sub-Pattern Assignment (GOSPA) was utilized as a performance indicator of the tracking algorithm. Compared to the traditional optimal subpattern assignment (OSPA) [33], GOSPA metric enables us to express the penalty in optimization over assignments rather than permutations between estimation set and ground truth. As missed targets and false targets are the most concerned attributes for the fusion performance of our algorithms, we took advantage of GOSPA to determine how our sensor fusion architectures improve the detection performance with its assignment's preference.

Let X be the ground truth set, $\hat{X}$ be the estimated set from sensor fusion and τ be the possible assignment set between X and $\hat{X}$ with combinatorial optimization algorithms such as Hungarian algorithm, the GOSPA metric for $\alpha = 2$ is formulated as follows,

$$d_p^{(c,2)}(X,\hat{X}) = \left[\min_{\gamma\in\tau}\left(\sum_{(i,j)\in\gamma} d(X,\hat{X})^p + \frac{c^p}{2}(|X| + |\hat{X}| - 2|\gamma|)\right)\right]^{\frac{1}{p}} \tag{16}$$

where α determines the error due to cardinality mismatch, p represents the dimension of the L-norm and c is the maximum allowable localization error [33]. In our performance evaluation of our sensor fusion algorithms, we assigned value p as 2 and value c as 10. A lower GOSPA value implies a lower error between X and $\hat{X}$.

To illustrate the improvement rate after sensor fusion, the least GOSPA with the least error among the measurements from all sensors before fusion and GOSPA from fusion tracking results are obtained to compute the difference for improvement comparison.

5.3 Simulation Dataset and Evaluation Results

In this section, four different speeds of mobility systems are evaluated based on the problem in Section 3.1.1 with only camera detection properties change for each case to examine the detection improvement ability of the proposed architectures for sensor fusion with GMPHD. We specially selected the cases with changing camera detection properties based on the problem in Section 3.1.2 because the detection property of optical sensors is always subject to illumination change when the mobility systems frequently encounter relatively dark environments while LiDAR and Radar are resilient to environmental influence. Meanwhile, we assume LiDAR and Radar have a relatively high number of clutters compared to Camera due to surrounding radioactive noise based on the problem in Section 3.1.3. For our simulation dataset, the environment is in fulfillment of three criteria, 1. speed of mobility system, 2. moving nature of target objects, and 3. sensor properties, and they are mentioned in Section 5.1. We ran our fusion scripts for all the evaluation cases on the computer with Intel(R) Core (TM) i7-9700F CPU @ 3.00 GHz and 32 GB RAM, the average fusion time of our evaluation result is around 82 milliseconds per each fusion step. The fusion time is longer than the processing time stated in Table 1 because of more sensor number and higher clutter number in our measurements but it is still acceptable for real-time application. In this evaluation, we used the same set of GMPHD model parameter setups for the architectures in all three cases for the purpose of testing its adaptability to the sudden sensor properties' change. Those parameter setups for prediction model $v^t_{k|k-1}(x)$ based on sensor t include survival rate $p_{S,k}$ with 98%, birth model $\gamma_k(x)$ based on t type sensor's FOV and resolution in normal case and prediction error covariance $P_{k|k-1}$ with diag(1, 1, 0.5, 0.5). Those parameter setups for update model $v^t_k(x)$ based on sensor t include detection rate $p_{D,k}$, the density of Poisson false alarm $p_{K,k}(z)$ and error covariance $S_{k|k-1}$ based on t type sensor's properties in normal case. For the rest of the GMPHD parameter setup, the Gaussian purging weight w_{purge} was set as 0.08, the number of hypothesis N_{purge} was set as 25 and the merging threshold L_{merge} was set as 5.

5.3.1 Evaluation of Normal Case

For the normal case, we assume that all the sensors perform normally and excel in their general functions without any sub-

Table 3 Configuration of sensor properties in normal case.

	Detection rate [%]	FOV distance [m]	Clutter rate (clutter/frame)	Error covariance (x [m], y [m], v_x [m/s], v_y [m/s])
Camera	95	300	0.1	(1, 1, 0.5, 0.5)
LiDAR	85	100	1	(0.5, 0.5, 0.5, 0.5)
Radar	90	150	3	(0.5, 0.5, 0.5, 0.5)

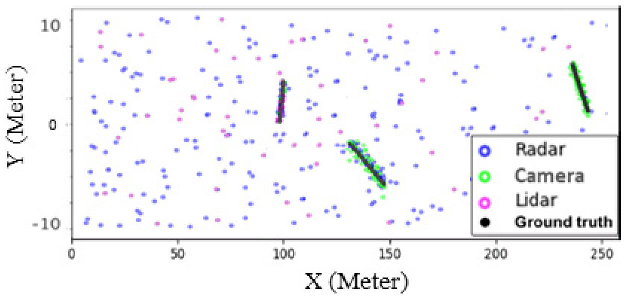

Fig. 13 The simulation measurements from all sensors at 0 km/h in normal case.

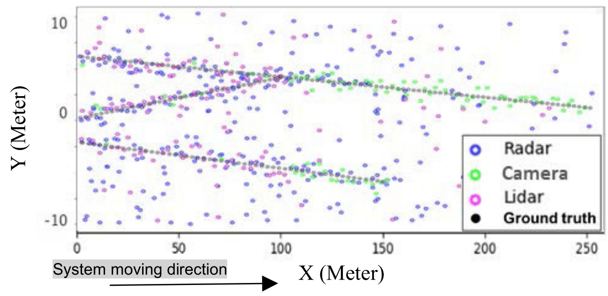

Fig. 14 The simulation measurements from all sensors at 90 km/h in normal case.

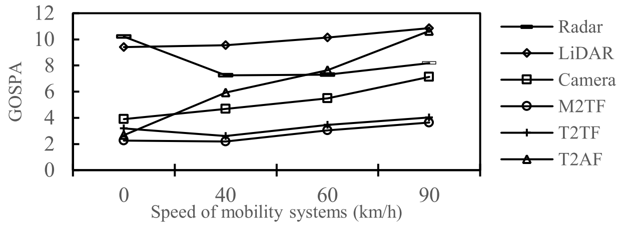

Fig. 15 GOSPA performance evaluation for tracking and fusion at different speeds in normal case.

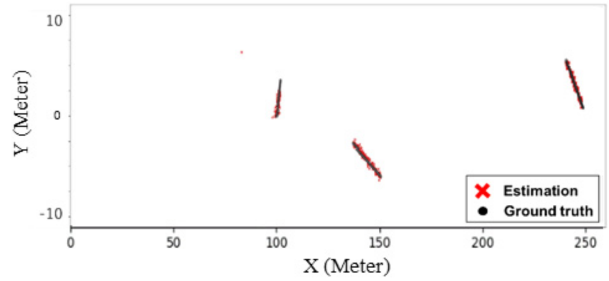

Fig. 16 Fusion result with M2TF architecture for mobility system at 0 km/h in normal case.

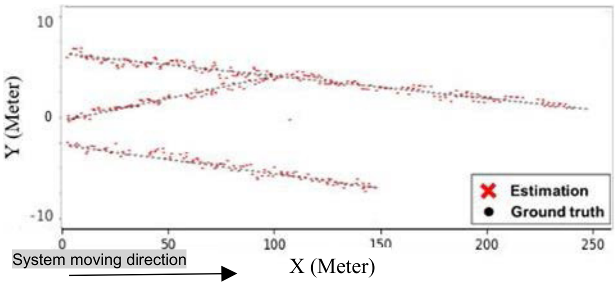

Fig. 17 Fusion result with M2TF architecture for mobility system at 90 km/h in normal case.

stantial influence from the surrounding environment. We suppose the fusion system performs its best detection improvement in this case and the system sensors are expected to be most of the time the same as the configuration of sensor properties in this case. The parameter configurations with detection rate, FOV distance, and clutter rate for every single sensor are detailed in **Table 3** shown below.

Based on the sensor properties detailed in Table 2, the simulation measurements from all sensor for the detected object are shown in **Fig. 13** and **Fig. 14** as the examples of the stationary environment and 90 km/h vehicle's speed environment respectively based on the objects' properties in Table 2 for the better result comparison.

In the results shown in **Fig. 15**, we found that the architecture M2TF performs superiorly compared to the rest of the two architectures and has a significant improvement in GOSPA. To focus on its improvement rate to the tracking result from the smallest GOSPA before fusion, the architecture GOPSA values in M2TF for the mobility systems speed at 0 km/h, 20 km/h, 60 km/h, and 90 km/h decreases by around 42.05%, 53.1%, 44.9%, and 48.95% respectively. In other words, the sensor fusion architecture M2TF remarkably reduces the localization error, the number of false detection, as well as the number of miss detection, compared to the ground truth data.

Overall, the architecture M2TF which performs better among all architectures has 47.5% GOSPA improvement on average from 5.30 to 2.78 to the relatively good tracking performance with all different speeds of mobility system when T2AF and T2AF have 37.32% improvement from 5.30 to 3.32 and 0% improvement on average respectively. It is worth noting that even though the architecture T2AF does not have an excellent improvement overall especially when the mobility system is moving, it has a 31.55% improvement from 3.9 to 2.67 for the stationary mobility system environment. To better understand the best performance of architecture M2TF in our result, **Fig. 16** and **Fig. 17** illustrate the exemplary fusion result of M2TF with the mobility system environment at stationary and fastest speed 90 km/h respectively.

5.3.2 Evaluation of Abnormal Case with Camera Detection Rate 75%

For the abnormal case with camera detection of 75% due to low illumination, we assume that all the sensors perform normally except the camera, and the detection rate of the camera is lowered from the original 95% to 75%, the parameter configurations for each sensor are detailed in the following table.

Based on the sensor properties detailed in **Table 4**, the simulation measurements from all sensor for the detected object are shown in **Fig. 18** and **Fig. 19** as the examples of the stationary environment and 90 km/h vehicle's speed environment respectively based on the objects' properties in Table 2 for the better result

Table 4 Configuration of sensor properties in abnormal case with camera detection rate 75%.

	Detection rate [%]	FOV distance [m]	Clutter rate (clutter/frame)	Error covariance (x [m], y [m], v_x [m/s], v_y [m/s])
Camera	75	300	0.1	(1, 1, 0.5, 0.5)
LiDAR	85	100	1	(0.5, 0.5, 0.5, 0.5)
Radar	90	150	3	(0.5, 0.5, 0.5, 0.5)

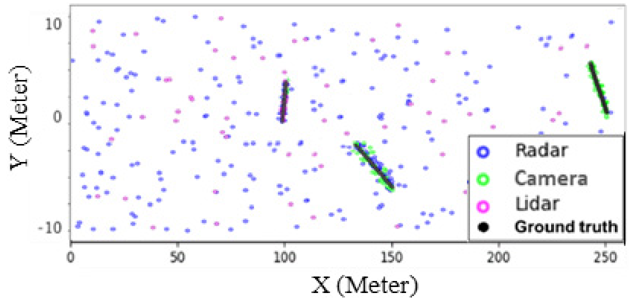

Fig. 18 The simulation measurement for mobility system at 0 km/h in abnormal case with camera detection 0.75.

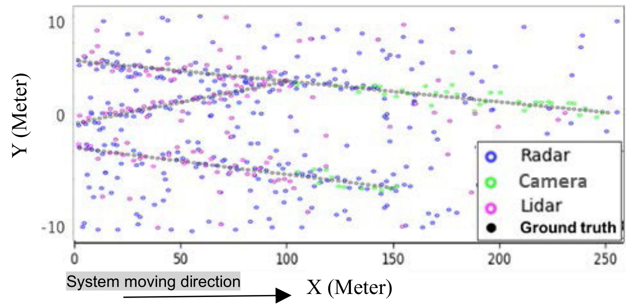

Fig. 19 The simulation measurement for mobility system at 90 km/h in abnormal case with camera detection 0.75.

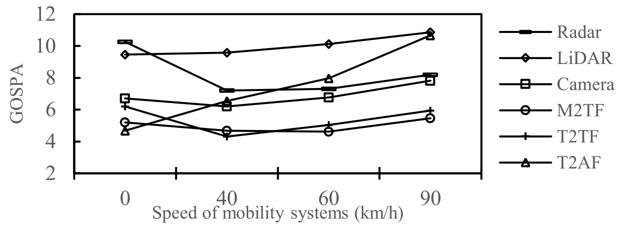

Fig. 20 GOSPA performance evaluation for tracking and fusion at different speeds in abnormal case with camera detection rate 75%.

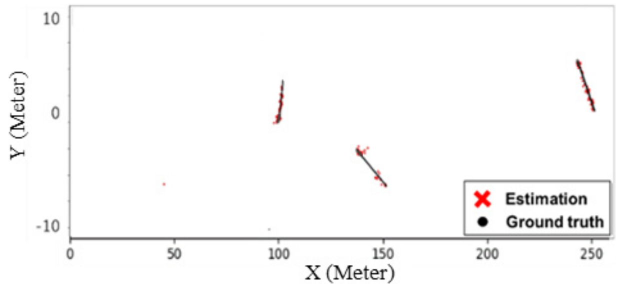

Fig. 21 Fusion result with T2AF architecture for mobility system at 0 km/h in the case with camera detection rate 0.75.

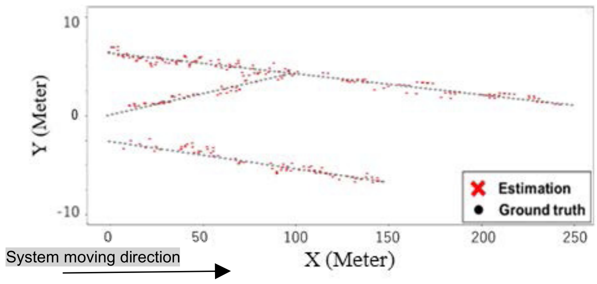

Fig. 22 Fusion result with M2TF architecture for mobility system at 90 km/h in the case with camera detection rate 0.75.

comparison.

In the result shown in **Fig. 20** below, we found that the architecture T2AF performs superiorly with an average 30.15% GOSPA improvement from 6.7 to 4.68 compared to the rest of the two architectures when the mobility system is stationary. However, it performs the worst with higher GOSPA when the mobility systems operate at higher speeds even though it has excellent fusion performance when the mobility system is stationary. This insinuates that the achievement of satisfactory fusion performance in T2AF architecture is contingent on the speed of the mobility systems. The rest of the architectures M2TF and T2TF demonstrate their improvement abilities and perform stably regardless of the mobility system speed. Both architectures M2TF and T2TF have similar performance overall with smaller GOSPA on average for all cases when the mobility systems are moving at different speeds.

On the whole, the architecture M2TF which perform better among all architectures has a 27.37% improvement on average from 6.87 to 4.99 to the relatively good tracking performance in Sensor 2 before fusion when T2TF has 20.82% on average from 6.87 to 5.37, It is worth noting that even though the architecture T2AF does not have an excellent improvement overall especially for the case when the mobility system is moving, it performs well in the stationary mobility system environment. To better understand the performance of the best fusion architectures for this abnormal case in the result, **Fig. 21** and **Fig. 22** illustrate architecture T2AF fusion result at stationary environment and architecture M2TF fusion result at the fastest speed 90 km/h respectively.

5.3.3 Evaluation of Abnormal Case with Camera Detection Rate 50%

For the abnormal case with camera detection of 50%, we assume that all the sensors perform normally except the camera, and the detection rate of the camera drops from the original 95% to 50%, the parameter configurations for each sensor are detailed in **Talbe 6** shown below.

Based on the sensor properties detailed in **Table 5**, the simulation measurements from all sensors for the detected object are shown in **Fig. 23** and **Fig. 24** as the examples of the stationary environment and 90 km/h vehicle's speed environment respectively based on the objects' properties in Table 2 for the better result comparison.

In **Fig. 25**, the GOSPA improvement tendency is similar to the normal case and the abnormal case with camera detection 75% although all the GOSPA values are relatively higher overall compared to the normal case and the case with a 75% camera detection rate change. On the other hand, both architectures M2TF and T2TF have similar performance regardless of the mobility system speed overall with smaller GOSPA with a 17.57% improvement rate from 7.23 to 5.96 and 18.67% improvement from 7.23 to

Table 5 Configuration of sensor properties in Abnormal case with camera detection rate 50%.

	Detection rate [%]	FOV distance [m]	Clutter rate (clutter/frame)	Error covariance (x [m], y [m], v_x [m/s], v_y [m/s])
Camera	**50**	300	0.1	(1, 1, 0.5, 0.5)
LiDAR	85	100	1	(0.5, 0.5, 0.5, 0.5)
Radar	90	150	3	(0.5, 0.5, 0.5, 0.5)

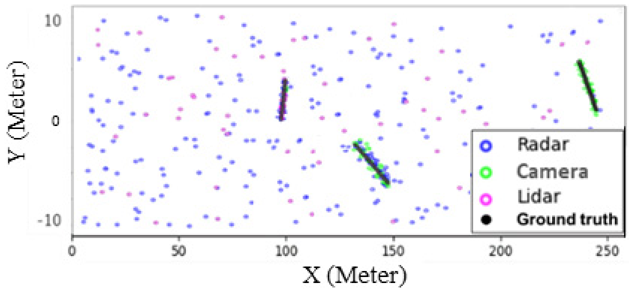

Fig. 23 The simulation measurement for mobility system at 0 km/h in abnormal case with camera detection rate 0.5.

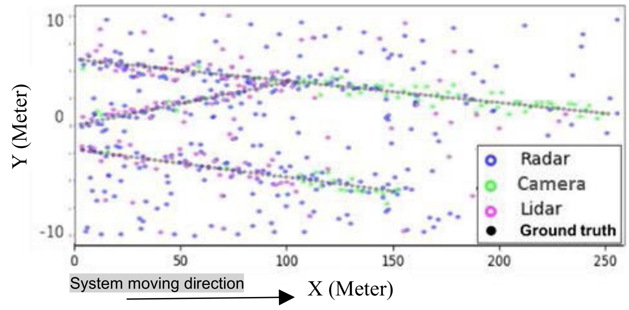

Fig. 24 The simulation measurement for mobility system at 90 km/h in abnormal case with camera detection rate 0.5.

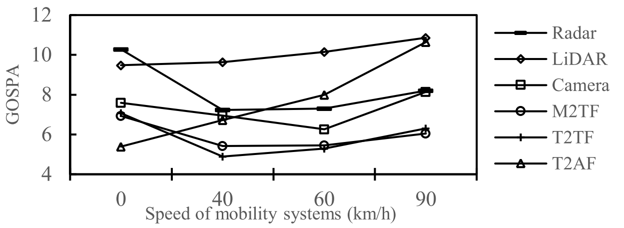

Fig. 25 GOSPA performance evaluation for tracking and fusion at different speeds in abnormal case with camera detection rate 50%.

5.88 on average respectively. We further found that the GOSPA of the architecture T2AF performs superiorly with 28.98% improvement from 7.59 to 5.39 compared to the rest of the two architectures and has a significant improvement in GOSPA when the mobility system is stationary. However, the architecture T2AF performs worst with higher GOSPA when the mobility systems operate at a higher speed. To better understand the performance of the best fusion architecture's improvement for this abnormal case in the result, **Fig. 26** and **Fig. 27** illustrate architecture T2AF fusion result at stationary environment and architecture M2TF fusion result at the fastest speed 90 km/h respective.

Overall, the architecture T2TF performs better among all architectures on average when M2TF has a similar improvement on average. Similar to the normal cases and other abnormal cases, it is worth noting that even though the architecture T2AF does not

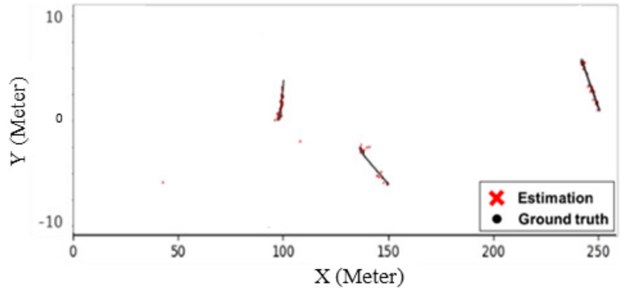

Fig. 26 Fusion result with T2AF architecture for mobility system at 0 km/h in abnormal case with camera detection 0.5.

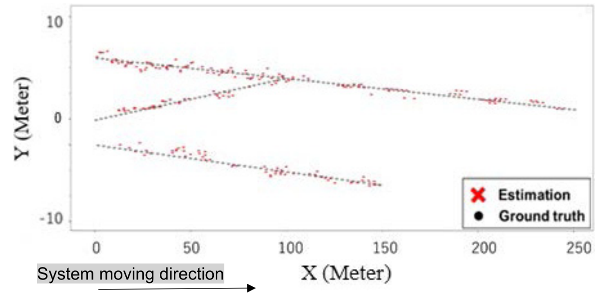

Fig. 27 Fusion result with M2TF architecture for mobility system at 90 km/h in abnormal case with camera detection 0.5.

have an excellent improvement overall, it performs well in the stationary mobility system environment.

5.4 Discussion of Our Research on the Consumer Products of Autonomous Mobility Systems

5.4.1 The Validity of Simulation Environment to the Consumer Products of Target Autonomous Mobility System

In our simulation environment for the performance evaluation, our control parameters for the simulation include the number of target objects, speed of mobility systems, and the sensor setup. Each of those setups in the simulation signifies the requirements of the detection performance under the environmental influence explained in Section 3.1 on the targeting consumer autonomous mobility systems and we presented each of their validities to the consumer product in this section.

For the number of target objects in the simulation environment, we set up 3 target objects because only the frontest objects concern the safety control of mobility systems all the time to decide an appropriate system operation. Specifically, we focus on a scenario similar to **Fig. 28** which does not require complicated path planning and the vehicles only move back and forth along the dedicated lane. Under this environment shown in Fig. 28, detecting only 3 objects at the front is more than enough to recognize the dangerous object in the environment and stop the vehicle safely. Therefore, this simulation property is useful for our target system which only operates along the dedicated lane without the need for complicated path planning.

Furthermore, it is more practical to set up a fewer target object based on the capability of our target setup with a low-resolution sensor to reduce the cost in our target mobility system. The mission of detecting a higher number of objects requires more expensive high-resolution sensors which make the consumer prod-

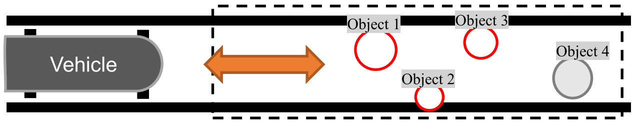

Fig. 28 Example of vehicle operation environment inside the dedicated lane.

ucts of autonomous mobility systems more unaffordable and it is unnecessary when our system in our goal only runs along the dedicated lane explained above which detecting 2 to 3 closest obstacles in the front are enough for the system safety control.

For the speed of mobility systems in the simulation environment based on Section 3.1.1, we targeted the speed of the ordinary vehicle from 0 km/h to 90 km/h based on the law explained in Section 5.1.3 when it could cover the wider spectrum of system operational speeds. The wider speed coverage also implies its support of the mobility systems with diverse levels of speeds and scales. Therefore, the improvement ability from all the speeds in the setup demonstrated the practicality of applying the sensor fusion system into most types of operating mobility systems as our target consumer product.

For the sensor setup in the simulation environment, we targeted Camera, LiDAR, and Radar and this setup follows the earlier research for the improvement of informative sensor measurements with environmental adaptability. We controlled the detection properties of the camera along with the clutter issue and system speed in the evaluation of our fusion system. In the setup, the Camera's performance always depends on the ambient light source and its detection rate (False negative) becomes more volatile in the real environment based on Section 3.1.2 whereas radiation sensors such as LiDAR and Radar can still maintain a fair object detection rate most of the time under the most adverse environmental influence. It is worthy of note that the least detection rate of the camera in our evaluation is 50% as we assume there is still room for the camera to detect with a reasonable light source. Although we only explained how environmental illumination aggravates camera performance in Section 3.1.2, other similar examples for the environmental influence of illumination in consumer products are sufficiently covered and they include overcast weather conditions such as rain and fog, unrecognizable objects due to the dazzling front light from the vehicles, and unrecognizable objects near the tunnel exit due to the backlight effect shown in **Fig. 29**. Those examples all have equivalent properties, and they are also applicable to the illumination influence with detection rate decrease in the evaluation of our paper. At the same time, Radar and LiDAR have significantly more clutter (False positive) than Camera due to surrounding radiation noise explained in Section 3.1.3 for any situation, and this simulation reflects the real situation in which 3 random clutters from the radioactive sensors are reasonable enough for interfering the detection of 3 objects in the fusion system of our consumer product to deal with and the system should ensure no clutter while accurately associating radioactive sensors' measurements with other sensor's during fusion. All these sensors' setups in the simulation can be considered as simulating the false detection problems consisting of false negative and false positive measurements which

Fig. 29 Unrecognizable object near the tunnel exit due to the backlight effect in exemplary optical sensor's image.

help us understand how best our fusion system can perform by improving them with GOSPA indicator in the evaluation.

5.4.2 The Impact of Our Result on the Consumer Products of Autonomous Mobility System

In the evaluation, we used all the architectures with the same GMPHD parameter configurations in all models based on the normal case for the rest of the two cases with camera properties' changes along with other environmental influences. This setting is significant to the consumer product when the systems cannot intelligently understand if the immediate sensor measurement quality is adversely affected by the environment in practice to automatically calibrate the models for detection accuracy enhancement.

In our result, we found out that M2TF architecture with GMPHD performs stably in general with an average 47.50% GOSPA improvement for normal case, 27.37% improvement for the abnormal case with 75% camera detection rate, and 21.83% GOSPA improvement for the abnormal case with 50% camera detection rate. Even though our result shows that the fusion system might not perform as ideally as the normal case does for the rest of the two cases, it can maintain relatively fair improvement ability with more than 20% under the environmental influence of higher system speed, lower camera detection rate as well as the clutter from radioactive sensors among all proposed sensor fusion architectures. Therefore, this evaluation result postulates M2TF fusion architecture with GMPHD is more suitable for the application of practical consumer mobility systems to improve the false detection problems compared to the rest and it positively impacts sensor measurement quality under the environmental influences by reducing the random noise and recovering false negative measurements in wider FOV area. Given the result, deploying this proposed fusion system in autonomous mobility systems could enable the efficient enhancement of the object detection safety in autonomous mobility systems under the environmental influence in contrast to the case when the system simply deploys those sensors without fusion system and suffers dangerously low accuracy.

5.4.3 The Practical Relation Between the Heterogenous Sensors and the Fusion System in Consumer Products

To practically deploy our proposed fusion system in the mobility system for high satisfaction of object detection performance under the environmental influence, we have to ensure that not only the proper model setting of sensor fusion system but the quality of each heterogeneous sensor with pre-fusion detection algorithm is also relatively resilient to the environmental change in object detection by confirming acceptable biased rate of detection rate and clutter rate in the model from the realistic performance to address the requirements of our fusion system. The reason for that is because the fusion accuracy of the proposed GMPHD architecture is relatively proportional to the original detection quality of sensor measurements.

Furthermore, it is essential to confirm that the obtained object measurement points from the chosen type and chosen number of sensors' deployment are informative and reasonable enough for the fusion improvement ability in a probabilistic manner. In the case of our heterogeneous sensor configuration, the fusion system will work practically well with the default setup in the paper for detecting 3 objects in the front at most 90 km/h with at least 50% camera detection rate with certain improvement according to our evaluation. For the consumer product in some actual extreme cases with the number of target object higher than 3, the mobility system speed higher than 90 km/h, the detection rate lower than 50%, and the number of clutter higher than 3 per detection time, our evaluation provides a valuable reference for developers. In other words, our works provide a good reference of sensor types and the number of sensors with using our fusion system for the customized reference configuration to achieve higher detection accuracy in most other cases. Therefore, setting up reasonable heterogeneous sensors based on their detection properties for the fusion system with appropriate architecture is of paramount importance in this sense.

5.4.4 The Future Research for Our Fusion System in the Consumer Products of Autonomous Mobility Systems

Since our case in consumer product only focuses on the fusion system working in a dedicated lane with low-resolution sensors in our current goal, we will exploit sensors of high resolution for our fusion system in the future and apply our system in a wider and more complicated environment when those kinds of sensors usually become more affordable in our consumer products as time goes by. Furthermore, we will focus more on the improvement of other insignificant environmental influences such as bumpy roads to advance our mobility systems with more real-world datasets and mature advancements of fusion techniques.

On the other hand, our fusion system tries to probabilistically recover the object detection measurement based on the incoming preprocessing measurements from several types of sensors and assumption model for those different sensor detection properties such as estimated detection rate and estimated clutter number. This means that the object in the scene cannot be detected if all the sensors in the system cannot obtain enough measurement for our fusion system to conduct statical recovery for the improvement of sensor detection accuracy in extraordinary cases. In our result, the occasional miss detection still happens due to the probabilistically unrecoverable detection which was mentioned above. Since the improvement in our fusion system still depends on the probabilistic assumption model of sensor properties as well as the detection quality of incoming sensor data, it implies that the limitation of our current fusion system is still yet to push the performance of sensor detection improvement ability to its finest state with probabilistic restriction in order to completely zero out the issues of all false positive and false negative measurements under the environmental influences. Therefore, we will strengthen the performance of our current probabilistic GMPHD fusion algorithm with the research direction of intelligently setting up a probabilistic model without relying on the original sensor quality and additionally pre-processing our heterogeneous sensor measurements to attain more adaptable object detection performance in the future.

6. Related Work

From the current literature, most of the proposed approaches that tackle the heterogeneous sensor fusion problem are based on track-to-track fusion which was originally proposed in Ref. [7]. Its variants include the refined data association with clustering [8] and integrating the non-kinematic information [9]. Their proposed techniques for the improvement of heterogeneous sensor fusion did not focus on our problems of missed detection and clutter in the fusion layer but only on the estimation of the positions of the target objects.

Furthermore, there are deep neural network-based fusion approaches such as Deep Multimodal Encoding [10] and Deep Fusion [11]. Nevertheless, this machine learning-based strategy requires enormous training data for creating the functional weighted model and it is still complicated at this point to ensure its faultlessness, safety, and high reliability due to its ambiguous inductive properties driven by restrictedly available training data. This ambiguity leads to impracticality and difficulty in creating an error-free model through debugging and re-training. Therefore, a deep neural network-based approach for sensor fusion is not our study target in this paper.

Whilst most of the existing literature focus on tracking problem of a homogeneous sensor with GMPHD such as Refs. [12], [13] and our result showed the feasibility of exploiting GMPHD for sensor fusion application, focusing on the fusion of homogeneous sensor is not applicable to practical tracking-based sensor fusion system for the consumer mobility system product due to incomprehensiveness of the environment in some abnormal circumstances with the detection ability constraint of a single type of sensor.

Even though there is a similar approach in literature [14] with one of our proposed GMPHD architectures for heterogeneous sensors, their target is different from ours and they only focus on the estimation of the contour shape of the object whereas our target is about point object tracking which locates the exact position of the object under the environmental influence. Therefore, our results are not comparable when we are solving two different problems.

7. Conclusion

This paper presents three architectures including T2TF, M2TF, and T2AF of heterogeneous sensor fusion with the integration of GMPHD in a bid to improve the detection ability in terms of the issues of clutter and miss-detection under the environmental influence. We have gone through the details of these three architectures and further evaluated their improvement ability with the same GMPHD parameter setup for all the cases. Our results have demonstrated that they all have significant improvement ability but performed differently during the speed change in mobility systems and the sudden change of one sensor properties due to environmental influence.

In our empirical results with GOSPA, it has been shown that the proposed sensor fusion architectures T2TF, M2TF, and T2AF with GMPHD can effectively improve the detection performance when the mobility system is stationary. However, the performance of T2AF starts to deteriorate when the mobility system speeds up from 0km/h even though it has an excellent 29.10% GOSPA improvement on average for all stationary situations compared to the rest. Among all architectures, M2TF architecture with GMPHD performs remarkably in general with an average 45.50% GOSPA improvement for normal case, 25.09% improvement for the abnormal case with 75% camera detection rate, and 21.83% GOSPA improvement for the abnormal case with 50% camera detection rate. This result also implies that M2TF is the best architecture with GMPHD integration which adapts to the environmental influence.

This evaluation is significantly critical for the adoption of appropriate sensor fusion architectures in consumer mobility systems since it reflects their abilities to keep the detection improvement performance by reducing the clutter and filling the miss detection especially when all the fusion setups remain unchanged in reality and there is a sudden change in sensor properties due to environmental influences and the moving speeds of systems. On the other hand, our evaluation result has a significant impact on helping the development of obstacle detection algorithm in fundamental consumer autonomous mobility systems when it is still challenging for other existing fusion systems to associate well sensor measurements to comprehend the obstacle accurately under the common environmental influence we have mentioned in this paper.

References

[1] Kawasaki, N. and Kiencke, U.: Standard platform for sensor fusion on advanced driver assistance system using Bayesian Network, *IEEE Intelligent Vehicles Symposium*, pp.250–255 (2004).

[2] Jia, X., Hu, Z. and Guan, H.: A new multi-sensor platform for adaptive driving assistance system (ADAS), *2011 9th World Congress on Intelligent Control and Automation*, pp.1224–1230, DOI: 10.1109/WCICA.2011.5970711 (2011).

[3] Hao, Y., Meng, F., Qiao, X., Zhao, Z., Zhang, C. and Cai, Y.: Extend Kalman Gaussian mixture probability hypothesis density filter based on radar and IR sensor fusion, *The 2010 IEEE International Conference on Information and Automation*, pp.1582–1586 (2010).

[4] Mahler, R.: "Statistics 102" for Multisource-Multitarget Detection and Tracking, *IEEE Journal of Selected Topics in Signal Processing*, Vol.7, No.3, pp.376–389 (June 2013).

[5] Vo, B.-N., Mallick, M., Bar-Shalom, Y., Coraluppi, S., Osborne, R., Mahler, R. and Vo, B.-T.: Multitarget Tracking, *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp.1–15 (2015).

[6] Vo, B. and Ma, W.: The Gaussian Mixture Probability Hypothesis Density Filter, *IEEE Trans. Signal Processing*, Vol.54, No.11, pp.4091–4104 (Nov. 2006).

[7] Yuan, T., Bar-Shalom, Y. and Tian, X.: Heterogeneous track-to-track fusion, *14th International Conference on Information Fusion*, pp.1–8 (2011).

[8] Houenou, A., Bonnifait, P., Cherfaoui, V. and Boissou, J.: A track-to-track association method for automotive perception systems, *2012 IEEE Intelligent Vehicles Symposium*, pp.704–710 (2012).

[9] Duraisamy, B., Schwarz, T. and Wöhler, C.: On track-to-track data association for automotive sensor fusion, *2015 18th International Conference on Information Fusion* (*Fusion*), pp.1213–1222 (2015).

[10] Liu, Z., Zhang, W., Lin, S. and Quek, T.Q.S.: Heterogeneous Sensor Data Fusion By Deep Multimodal Encoding, *IEEE Journal of Selected Topics in Signal Processing*, Vol.11, No.3, pp.479–491 (2017).

[11] Liang, P. et al.: Deep fusion of heterogeneous sensor modalities for the advancements of ADAS to autonomous vehicles, *2018 International Symposium on VLSI Design, Automation and Test* (*VLSI-DAT*), pp.1–4 (2018).

[12] Pham, N.T., Huang, W. and Ong, S.H.: Multiple sensor multiple object tracking with GMPHD filter, *2007 10th International Conference on Information Fusion*, pp.1–7 (2007).

[13] Sheng, X., Chen, Y., Guo, L., Yin, J. and Han, X.: Multitarget tracking algorithm using multiple GMPHD filter data fusion for sonar networks, *Sensors*, Vol.18, No.10, p.3193 (2018).

[14] Michaelis, M., Berthold, P., Meissner, D. and Wuensche, H.: Heterogeneous multi-sensor fusion for extended objects in automotive scenarios using Gaussian processes and a GMPHD-filter, *2017 Sensor Data Fusion: Trends, Solutions, Applications* (*SDF*), pp.1–6 (2017).

[15] Liggins, M.E. and Chang, K.-C.: Distributed fusion architectures, algorithms, and performance within a network-centric architecture, Ch.17, *Handbook of Multi-sensor Data Fusion: Theory and Practice*, Taylor & Francis, 2nd Edition (2009).

[16] Kim, C., Li, F., Ciptadi, A. and Rehg, J.M.: Multiple Hypothesis Tracking Revisited, *2015 IEEE International Conference on Computer Vision* (*ICCV*), pp.4696–4704 (2015).

[17] Bhattacharya, T., Premji, A., Nohara, T.J. and Weber, P.: Evaluation of fast MHT algorithms, *Proc. 1998 IEEE Radar Conference, RADARCON'98*, Challenges in Radar Systems and Solutions (Cat. No.98CH36197), pp.213–218 (1998).

[18] Özgen, S., Demirekler, M. and Orguner, U.: On the solution of data association problem using rollout algorithms, *2016 19th International Conference on Information Fusion* (*FUSION*), pp.695–702 (2016).

[19] He, S., Shin, H. and Tsourdos, A.: Track-Oriented Multiple Hypothesis Tracking Based on Tabu Search and Gibbs Sampling, *IEEE Sensors Journal*, Vol.18, No.1, pp.328–339 (2018).

[20] Rezatofighi, S.H., Milan, A., Zhang, Z., Shi, Q., Dick, A. and Reid, I.: Joint Probabilistic Data Association Revisited, *2015 IEEE International Conference on Computer Vision* (*ICCV*), pp.3047–3055 (2015).

[21] Konstantinova, P., Udvarev, A. and Semerdjiev, T.: A Study of a Target Tracking Algorithm Using Global Nearest Neighbor Approach 1, International Conference on Computer Systems and Technologies, *CompSysTech 2003* (June 2003).

[22] Kruger, J. and D'Amico, S.: Autonomous Angles-Only Multi-Target Tracking for Spacecraft Swarms *IEEE Trans. Aerospace and Electronic Systems*, Vol.1, No.1 (2020).

[23] Mahler, R.P.S.: Multitarget Bayes filtering via first-order multitarget moments, *IEEE Trans. Aerospace and Electronic Systems*, Vol.39, No.4, pp.1152–1178 (2003).

[24] Vo, B.-N., Singh, S. and Doucet, A.: Random finite sets and sequential Monte Carlo methods in multi-target tracking, *2003 Proc. International Conference on Radar* (*IEEE Cat. No.03EX695*), pp.486–491 (2003).

[25] Ma, W.-K., Vo, B.-N., Singh, S.S. and Baddeley, A.: Tracking an unknown time-varying number of speakers using TDOA measurements: A random finite set approach, *IEEE Trans. Signal Processing*, Vol.54, No.9, pp.3291–3304 (2006).

[26] Li, S., Yi, W., Kong, L. and Wang, B.: Dynamic factorization based multi-target Bayesian filter for multi-target detection and tracking, *2014 IEEE Radar Conference*, pp.1251–1256 (2014).

[27] Ristic, B., Vo, B., Vo, B. and Farina, A.: A Tutorial on Bernoulli Filters: Theory, Implementation and Applications, *IEEE Trans. Signal Processing*, Vol.61, No.13, pp.3406–3430 (2013).

[28] Vo, B., Vo, B. and Cantoni, A.: Performance of PHD Based Multi-Target Filters, *2006 9th International Conference on Information Fusion*, pp.1–8 (2006).

[29] Mahler, R.P.S.: "Statistics 101" for multisensor, multitarget data fusion, *IEEE Aerospace and Electronic Systems Magazine*, Vol.19, No.1, pp.53–64 (2004).

[30] Novak, E.: Some Results on the Complexity of Numerical Integration, *Mathematics & Statistics Monte Carlo and Quasi-Monte Carlo Methods*, pp.161–183, Springer (2016).

[31] Baser, E. and Efe, M.: A novel auxiliary particle PHD filter, *2012 15th International Conference on Information Fusion*, pp.165–172 (2012).

[32] Barney, D., Haley, D. and Nikandros, G.: Calculating Braking Distance, Conference: Safety Critical Systems and Software, *6th Australian Workshop on Safety-Related Programmable Systems*, pp.23–30 (2001).

[33] Rahmathullah, A.S., García-Fernández, Á.F. and Svensson, L.: Generalized optimal sub-pattern assignment metric, *2017 20th International Conference on Information Fusion* (*Fusion*), pp.1–8 (2017).

[34] Federal Motor Vehicle Safety Standards: Low Speed Vehicles 49, CFR 571 (Aug. 2005).

Chow Man Yiu received his Computer Science (M.Sc.) degree from The Chinese University of Hong Kong, Hong Kong, in 2018 and he joined Hitachi, Ltd. as an assistant researcher in the same year. He is currently working on the research of obstacle detection in the System productivity research department at Hitachi, Ltd. His research interests include sensor fusion, Machine learning, Data science, and Pattern Recognition.

Mitsuhiro Kitani received his M.Eng. degree in information and system engineering from Chuo University, Tokyo, Japan in 2004. He joined Hitachi, Ltd in same year. He is a member of IPSJ and is working on the system development process in mobility systems. His research interests include system development process using AI technique.