

Article

Robust Cooperative Multi-Vehicle Tracking with Inaccurate Self-Localization Based on On-Board Sensors and Inter-Vehicle Communication

Xiaobo Chen ^{1,*}, Jianyu Ji ² and Yanjun Wang ²

¹ Automotive Engineering Research Institute, Jiangsu University, Zhenjiang 212013, China

² School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China; 18852852561@163.com (J.J.); 18651403397@163.com (Y.W.)

* Correspondence: xbchen82@gmail.com; Tel.: +86-187-5296-5624

Received: 7 May 2020; Accepted: 1 June 2020; Published: 5 June 2020



Abstract: The fusion of on-board sensors and transmitted information via inter-vehicle communication has been proved to be an effective way to increase the perception accuracy and extend the perception range of connected intelligent vehicles. The current approaches rely heavily on the accurate self-localization of both host and cooperative vehicles. However, such information is not always available or accurate enough for effective cooperative sensing. In this paper, we propose a robust cooperative multi-vehicle tracking framework suitable for the situation where the self-localization information is inaccurate. Our framework consists of three stages. First, each vehicle perceives its surrounding environment based on the on-board sensors and exchanges the local tracks through inter-vehicle communication. Then, an algorithm based on Bayes inference is developed to match the tracks from host and cooperative vehicles and simultaneously optimize the relative pose. Finally, the tracks associated with the same target are fused by fast covariance intersection based on information theory. The simulation results based on both synthesized data and a high-quality physics-based platform show that our approach successfully implements cooperative tracking without the assistance of accurate self-localization.

Keywords: cooperative perception; multi-vehicle tracking; Bayes inference; track association

1. Introduction

Nowadays, intelligent vehicles equipped with advanced driver assistance systems (ADASs) can perceive other road participants and obstacles, including vehicles, pedestrians, etc., through on-board sensors. A variety of on-board sensors have been widely applied to achieve this goal, such as camera, Lidar, Radar, and so on. The perception system [1,2] of intelligent vehicles captures the measurements of surrounding targets through these sensors and build an environmental model which reflects the real states of different targets.

Multi-vehicle tracking (or more generally, multi-object tracking, MOT) is a crucial perception task since an accurate estimate of surrounding vehicles plays an important role in subsequent collision avoidance and route planning. The main challenge in MOT is the determination of the association between measurements and targets. In the literatures, extensive algorithms have been proposed to handle the data association problem. Multiple hypothesis tracking (MHT) [3] is known as a powerful algorithm to address the data association problem in MOT. Although MHT is Bayesian optimal in theory, the exact solution of MHT is computationally intractable, thus, requiring proper approximation. Joint probabilistic data

association (JPDA) [4] makes a soft association with the assumption that each measurement can originate from several candidate targets. JPDA can achieve reasonable results with lower computational burden. Recently, Mahler first used the Bayesian filter to derive a multi-target tracking algorithm based on random finite set (RFS) theory [5,6]. Under the RFS framework, the Bayes recursion formula for tracking single targets can be extended to multi-target tracking problems. The resulting probability hypothesis density (PHD) filter can propagate the PHD or the first-order moment of the multi-target posterior density. The integral of the PHD over a region in the state space gives the expected number of targets within this region, while the peaks in the PHD represents the state of the targets. To address the computationally intractable multiple integrals issue, Gaussian mixture PHD (GMPHD) [7] was further developed which led to a closed-form solution to the PHD filter recursion.

In reality, due to the physical mechanism, the on-board sensors of intelligent vehicles always suffer from inherent drawbacks, such as limited perception range or field-of-view (FOV). Moreover, in the crowd scenarios where frequent occlusion occurs, intelligent vehicles fail to detect the target that is occluded, thus, increasing the potential risk of traffic conflicts. To address the above issue, cooperative perception (or collaborative perception) [8–10] based on inter-vehicle communication has attracted much attention recently. For instance, dedicated short range communication (DSRC) has been introduced as a useful technique which allows vehicles to communicate with other neighboring vehicles through the on-board units (OBUs) installed in the vehicles. When the vehicles communicate successfully, they can exchange their respective local perception results. In fact, inter-vehicle communication (more strictly speaking, the wireless interface) can be viewed as a type of virtual sensor [11,12] in the sense that a host vehicle can combine its local estimate and a remote message transmitted from other cooperative vehicles to form a more accurate and complete environmental model. Due to the significant distance of inter-vehicle communication and the fusion of information from different viewpoints, cooperative perception not only increases the accuracy and perceptual range beyond line of sight, but also reduces blind spots caused by mutual occlusion, weather, and other external factors.

By introducing inter-vehicle communication into multi-object tracking, cooperative tracking has emerged as a promising technique. In [13], a collaborative sensor fusion algorithm was proposed for MOT by combining GMPHD filter and covariance intersection. The cooperative tracking algorithm was further applied to assist the overtaking system [14] and the results confirmed the advantage of cooperative sensing. In [15], a data association and fusion framework was proposed for multi-vehicle cooperative perception. An interactive multiple model (IMM) filter [16] was used to estimate the state of vehicles and Bhattacharyya distance was applied to measure the difference between local tracks from a host vehicle and communicated tracks from a cooperative vehicle. After association, fast covariance intersection (FCI) [17] which computed the weights directly without nonlinear optimization was employed to fuse together the tracks associated with the same target. This work relied on the assumption of on-board measurements available in global coordinate systems, which required the accurate self-localization of both host and cooperative vehicles. In [18], the temporal and spatial alignment between the local environment model of host and cooperative vehicles were reviewed, however, target association and track fusion were not discussed. Some works [19–21] studied the track-to-track association problem in the case of cooperative sensing under the assumption of accurate positioning information.

In summary, current cooperative multi-vehicle tracking methods typically assume the self-localization information, such as position and orientation, of both host and cooperative vehicles is accurate such that the local tracks perceived by two vehicles can be readily converted into either a global frame or the local frame of host vehicle by coordinate system transformation. Then, track association and fusion can be conducted and the information from cooperative vehicle can be used to enhance the perception performance of host vehicle. However, accurate self-localization through positioning systems, such as Global Positioning System (GPS), is not always accurate enough or even available [22]. For example, in dense urban environment where the

vehicles are surrounded by skyscrapers or tall buildings, the location information provided by GPS would not be reliable. In such cases, inaccurate or even unavailable self-localization information seriously affects the performance of cooperative tracking system.

To address the above problem and enhance the robustness of cooperative multi-vehicle tracking, we propose, in this paper, an integrated framework which can determine the relative pose (including translation and orientation) of host and cooperative vehicles relying on their respective local tracks, instead of global positioning information. Moreover, this work concentrates on the dynamic situation where the relative translation and orientation between host and cooperative vehicles are changing with time. This is more consistent with real traffic environment where vehicles are usually driving with different intentions. Consequently, our cooperative multi-vehicle tracking system still works without the assistance of accurate self-localization information.

It should be emphasized that in the literature of wireless sensor networks (WSNs), some target tracking algorithms [23–25] have been developed for simultaneous localization and tracking (SLAT). In [24], a Bayesian filtering framework was proposed to infer the joint posterior distribution of both target and multiple sensors. Variational method [26] was used to approximate the joint state during the measurements correction stage. In [25], a dynamic non-parametric belief propagation (DNBP) method was proposed for cooperative vehicle sensing. However, most works in SLAT tend to track one target by using multiple static or moving sensors, thus, restricting their application to more complex scenarios where the number of targets can vary at different times.

The rest of this paper is organized as follows: In Section 2, we briefly review the adaptive GMPHD filter which is the basic component for target tracking; in Section 3, we present our framework of cooperative tracking and propose a Bayes model for simultaneous track association and relative pose estimation; in Section 4 we report some simulation results based on both the synthesized data and PreScan-based system; and finally, we provide some conclusions and future works in Section 5.

2. Adaptive GMPHD Filter for MOT

2.1. PHD Filter Formulation

In PHD filter, both the multi-target state and the set of measurements at each time step are represented by RFS. The PHD filter approximates the multi-target Bayes filter through propagating the first-order moment. The recursive process of PHD filter includes two steps, i.e., prediction and correction (or update) as follows:

1. PHD prediction

$$v_{k|k-1}(x) = \int [P_{S,k}(\zeta) f_{k|k-1}(x|\zeta) + \beta_{k|k-1}(x|\zeta)] v_{k-1|k-1}(\zeta) d\zeta + \gamma_k(x) \quad (1)$$

where $v_{k|k-1}(x)$ and $v_{k-1|k-1}(\zeta)$ represent the predicted and posterior intensity of the target at time $k - 1$, $P_{S,k}(\zeta)$ is the survival probability when the target state is ζ , $f_{k|k-1}(\cdot|\cdot)$ is the single-target state transition density, $\beta_{k|k-1}(\cdot|\cdot)$ and $\gamma_k(x)$ denote the intensity of spawning target and newborn target, respectively.

2. PHD correction

$$v_{k|k}(x) = [1 - P_{D,k}(x)] v_{k|k-1}(x) + \sum_{z \in Z^{(k)}} \frac{P_{D,k}(x) g_k(z|x) v_{k|k-1}(x)}{\kappa_k(z) + \int P_{D,k}(\zeta) g_k(z|\zeta) v_{k|k-1}(\zeta) d\zeta} \quad (2)$$

where $Z^{(k)} = \{z_{k,1}, z_{k,2}, \dots, z_{k,M_k}\}$ denotes the measurement set at time k , M_k is the total number of measurements, $P_{D,k}(x)$ is the detection probability when the target state is x , $g_k(\cdot|\cdot)$ is the single-target measurement likelihood function, and $\kappa_k(\cdot)$ is the intensity of the clutter RFS.

The above two formulas are the basic recursive equations for PHD filtering. After correction at each time, the expected number of targets can be obtained by integrating the updated PHD intensity, and then taking the nearest integer value. The status of each target can be obtained from the state corresponding to the updated PHD peak point. It can be seen that the PHD filter can avoid the data association problem. However, the PHD filter includes multiple integration operations, which cannot obtain an analytical solution, and suffer from “dimensional disaster” in the calculation of numerical integration.

2.2. Gaussian Mixture Implementation

In order to give a closed-form solution for PHD recursion, Gaussian mixture PHD (GMPHD) filter uses a set of Gaussian components to approximate the posterior density where the weights, mean values, and covariances of each Gaussian component are continuously updated over time. Suppose that the posterior intensity at time $k - 1$ is expressed as follows:

$$v_{k-1|k-1}(x) = \sum_{i=1}^{J_{k-1|k-1}} w_{k-1}^{(i)} N(x; m_{k-1}^{(i)}, P_{k-1}^{(i)}) \quad (3)$$

where $J_{k-1|k-1}$ represents the number of Gaussian components at time $k - 1$, $N(x|a, B)$ stands for the multivariate Gaussian distribution with mean a and covariance B . It is assumed that transition probability density and observed likelihood function also follow Gaussian distribution as follows:

$$f_{k|k-1}(x|\zeta) = N(x; F_{k-1}\zeta, Q_{k-1}) \quad (4)$$

$$g_k(z|x) = N(z; H_k x, R_k) \quad (5)$$

where F_{k-1} and H_k are linear state transition matrix and linear observation matrix, respectively, and Q_{k-1} and R_k are the covariance matrices of the process noise and the measurement noise, respectively.

Substituting $v_{k-1|k-1}(x)$ in Equation (3) into the PHD prediction and correction equations, we can obtain the recursive form of the PHD in the Gaussian mixture form. Specifically, GMPHD performs the prediction and correction as follows:

1. GMPHD prediction

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} N(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}) \quad (6)$$

In this work, the spawning target is ignored and the prediction Formula (6) can be rewritten as

$$v_{k|k-1}(x) = v_{S,k|k-1}(x) + \gamma_k(x) \quad (7)$$

where

$$v_{S,k|k-1}(x) = P_{S,k} \sum_{i=1}^{J_{k|k-1}} w_{k-1}^{(i)} N(x; m_{S,k|k-1}^{(i)}, P_{S,k|k-1}^{(i)}) \quad (8)$$

$$m_{S,k|k-1}^{(i)} = F_{k-1} m_{k-1}^{(i)} \quad (9)$$

$$P_{S,k|k-1}^{(i)} = F_{k-1}P_{k-1}^{(i)}F_{k-1}^T + Q_{k-1} \quad (10)$$

2. GMPHD correction

$$v_{k|k}(x) = (1 - P_{D,k})v_{k|k-1}(x) + \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) N(x; m_{k|k}^{(j)}, P_{k|k}^{(j)}) \quad (11)$$

where

$$w_k^{(j)}(z) = \frac{P_{D,k} w_{k|k-1}^{(j)} N(x; H_k m_{k|k-1}^{(j)}, H_k P_{k|k-1}^{(j)} H_k^T + R_k)}{\kappa_k(z) + P_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} N(x; H_k m_{k|k-1}^{(l)}, H_k P_{k|k-1}^{(l)} H_k^T + R_k)} \quad (12)$$

$$m_{k|k}^{(j)} = m_{k|k-1}^{(j)} + K_k^{(j)} (z - H_k m_{k|k-1}^{(j)}) \quad (13)$$

$$P_{k|k}^{(j)} = (I - K_k^{(j)} H_k) P_{k|k-1}^{(j)} \quad (14)$$

$$K_k^{(j)} = P_{k|k-1}^{(j)} H_k^T (H_k P_{k|k-1}^{(j)} H_k^T + R_k)^{-1} \quad (15)$$

After the GMPHD correction is completed, the Gaussian components with small weight are pruned and the Gaussian components close to each other are merged. Finally, in order to extract tracks, the mean value of the Gaussian component with the weight greater than a certain threshold is used as the multi-object state estimation.

For the application of GMPHD, the intensity of newborn target $\gamma_k(x)$ should be defined at each time, indicating the possible state when new targets appear. In our cases, the target vehicles can encounter the FOV of host and cooperative vehicles at different positions and times. As a result, it is infeasible to define $\gamma_k(x)$ in advance. To address the problem, at each time, we let $\gamma_k(x)$ be driven by the observed measurements adaptively as follows:

$$\gamma_k(x) = \sum_{i=1}^{M_k} w^{(i)} N(x; \bar{z}_{k,i}, P_0) \quad (16)$$

where P_0 denotes the initial covariance matrix and $\bar{z}_{k,i}$ denotes the state constructed from the measurement $z_{k,i}$. In such a way, the resulting adaptive GMPHD filter is applicable to the cooperative tracking situations we concern.

3. Cooperative Tracking with Inaccurate Self-Localization

3.1. Framework of Cooperative Tracking

We assume that many vehicles are moving in the environment according to their maneuvers. Among these vehicles, some vehicles can independently sense the other vehicles within in certain range by using on-board sensors and exchange their local information via inter-vehicle communication. The vehicle transmitting message is called cooperative vehicle, whereas the vehicle receiving message and performing fusion is called host vehicle. Certainly, a vehicle can send or receive message depending on its role. The other vehicles not involved in cooperation are called target vehicles. This work concentrates on the situation where the relation translation and rotation between host and cooperative vehicles are dynamically changing and inaccurate (or even unknown), thus, preventing the direct application of existing technologies.

To address the above problem and achieve sensor fusion for cooperative multi-vehicle tracking, we propose a novel framework depicted in Figure 1.

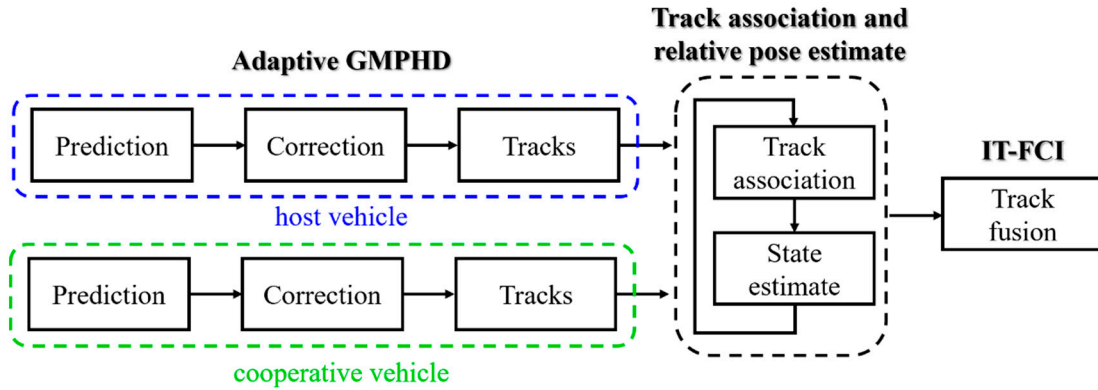


Figure 1. The proposed cooperative multi-vehicle tracking framework.

As shown in Figure 1, the host and cooperative vehicles first obtain respective local tracks by conducting the adaptive GMPHD algorithm presented in Section 2. Then, the cooperative vehicle transmits its local tracks to the host vehicle which attempts to estimate the relative translation and rotation between two vehicles and simultaneously associate the locals track from two vehicles by using the algorithm explained next. Finally, the matched tracks are fused following fast covariance intersection based on information theory (IT-FCI) [27].

3.2. Track Association and Relative Pose Estimation

3.2.1. Formulation

The relative pose estimation and track association problem in dynamic scenarios is shown in Figure 2. Here, we focus on two-dimensional space for brevity, however, our proposed method can be extended to higher-dimensional space with slight modification. Given two vehicles S1 and S2, here, S1 is assumed to be the host vehicle, and S2 is assumed to be the cooperative vehicle, indicating that S2 sends its local estimates to S1 where the information fusion is performed. At given time k , let $X_k^1 = \{x_{k,1}^1, x_{k,2}^1, \dots, x_{k,N_k^1}^1\}$ and $X_k^2 = \{x_{k,1}^2, x_{k,2}^2, \dots, x_{k,N_k^2}^2\}$ be the collection of local tracks of S1 and S2 which are obtained through MOT algorithm, such as adaptive GMPHD in Section 2. N_k^1 and N_k^2 denote the total number of tracks in S1 and S2, respectively. Moreover, the relative location and orientation of S2 with respect to S1 at time k is characterized by $s_k = [\xi_k, \dot{\xi}_k, \eta_k, \dot{\eta}_k, \theta_k, \dot{\theta}_k]^T$ where ξ_k and η_k , denote the translation of S2 in the Cartesian coordinate system of S1. Similarly, θ_k denotes the orientation angle. $\dot{\xi}_k$, $\dot{\eta}_k$, and $\dot{\theta}_k$ represent the corresponding velocities.

Any track $x_{k,j}^2$ in the coordinate of S2 can be exactly transformed to that of S1 as follows:

$$x_{k,j}^{2 \rightarrow 1} = R(\theta_k)x_{k,j}^2 + \begin{bmatrix} \xi_k \\ \eta_k \end{bmatrix} \quad (17)$$

where $R(\theta_k) = \begin{bmatrix} \cos \theta_k & -\sin \theta_k \\ \sin \theta_k & \cos \theta_k \end{bmatrix}$ denotes the rotation matrix. In the situation we consider, a major difficulty is that $(\xi_k, \eta_k, \theta_k)$ is inaccurate (or even totally unknown) and dynamically changing with time.

In addition, in the case of multiple targets, the association between tracks of different vehicles is also unknown.

Suppose the track association between two vehicles is denoted by the true but unknown $N_k^1 \times N_k^2$ association matrix $U_k = \{u_{ij}^k\}$ with each entry $u_{ij}^k \in \{0, 1\}$ representing the result of association between $x_{k,i}^1$ and $x_{k,j}^2$, $1 \leq i \leq N_k^1, 1 \leq j \leq N_k^2$. Formally, if $u_{ij}^k = 1$, it means that local tracks $x_{k,i}^1$ and $x_{k,j}^2$ refer to the same target; otherwise, it means they belong to different target. Since it is assumed that each local track in one sensor corresponds to one, and at most one local track in other sensor, we have the constraints applied to U_k below

$$\sum_{j=1}^{N_k^2} u_{ij}^k \leq 1, \sum_{i=1}^{N_k^1} u_{ij}^k \leq 1, \forall 1 \leq i \leq N_k^1, 1 \leq j \leq N_k^2. \quad (18)$$

In addition, in order to incorporate the prior information about s_k , it is supposed that s_k follows

$$P(s_k) = N(s_k|a, B). \quad (19)$$

For example, similar to Kalman filter, we let $a = F\hat{s}_{k-1}$, $B = FP_{k-1}F^T + Q$, \hat{s}_{k-1} , and P_{k-1} are the mean and covariance of s_{k-1} ; F is the state transition matrix of cooperative vehicle; and Q is covariance matrix of the process noise.

Similarly, according to Equation (17), we have the following likelihood function:

$$P(x_{k,i}^1|s_k, x_{k,j}^2) = N\left(x_{k,i}^1|R(\theta_k)x_{k,j}^2 + \begin{bmatrix} \xi_k \\ \eta_k \end{bmatrix}, \Sigma\right) \quad (20)$$

where Σ is the measurement noise covariance. The above assumption is reasonable, since $x_{k,j}^{2 \rightarrow 1}$ should be close to $x_{k,i}^1$, when $x_{k,i}^1$ and $x_{k,j}^2$ refer to the same target. We also assume that the local tracks are independent of each other. On the basis of the above discussion, we propose the following probabilistic model:

$$P(s_k, X_k^1, X_k^2|U_k) = P(s_k) \prod_j P(x_{k,j}^2) \prod_{i,j} P(x_{k,i}^1|s_k, x_{k,j}^2)^{u_{ij}^k} \quad (21)$$

Other prior knowledge, for example, given noisy measurement of partial entries of s_k , can be easily integrated into Equation (21) by introducing extra likelihood function.

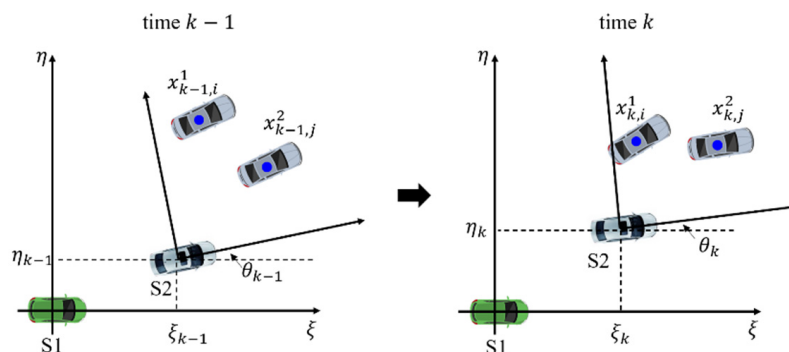


Figure 2. Relative translation and orientation between vehicle S1 and vehicle S2.

3.2.2. Expectation-Maximum (EM) Solution Algorithm

We treat s_k, X_k^1, X_k^2 , and U_k as complete data; X_k^1, X_k^2 as incomplete data; s_k as hidden variable; and U_k as unknown parameter. Then, we develop an effective solution in the spirit of the expectation-maximum (EM) solution algorithm, which attempts to jointly estimate the distribution of hidden variable s_k and track association matrix U_k in an iterative fashion. Specifically, the algorithm consists of the following two steps:

1. E-step

$$Q(U, U^{p-1}) = E\{\log P(s_k, X_k^1, X_k^2|U)|U^{p-1}\}; \tag{22}$$

2. M-step

$$U^p = \operatorname{argmax} Q(U, U^{p-1}). \tag{23}$$

where p refers to the p th iteration of the algorithm. The E-step and M-step repeat until certain convergence criteria is satisfied.

E-Step

Given the estimation of U_k at the $(p-1)$ -th iteration, we need to calculate the expectation of $\log P(s_k, X_k^1, X_k^2|U)$. Let $\Omega_k^{p-1} = \{(i, j)|u_{ij}^{k,p-1} = 1\}$, indicating those associated tracks between sensors. Then, according to Bayes theorem and Equation (21), we can have the posterior distribution of s_k as follows:

$$P(s_k|X_k^1, X_k^2, U^{p-1}) = \frac{P(s_k) \prod_{(i,j) \in \Omega_k^{p-1}} P(x_{k,i}^1|s_k, x_{k,j}^2)}{\int P(s_k) \prod_{(i,j) \in \Omega_k^{p-1}} P(x_{k,i}^1|s_k, x_{k,j}^2) ds_k} \tag{24}$$

Considering that $R(\theta_k)$ in the above distribution is nonlinear with respect to s_k , we apply Talyor series expansion to obtain the first-order linear approximation around current estimate θ_k^{l-1} as follows:

$$R(\theta_k)x_{k,j}^2 + \begin{bmatrix} \xi_k \\ \eta_k \end{bmatrix} \approx H(\theta_k^{l-1})s_k + R(\theta_k^{l-1})x_{k,j}^2 - \theta_k^{l-1}\bar{R}(\theta_k^{l-1})x_{k,j}^2 \tag{25}$$

where $\bar{R}(\theta_k) = \begin{bmatrix} -\sin \theta_k & -\cos \theta_k \\ \cos \theta_k & -\sin \theta_k \end{bmatrix}$ is the Jacobian matrix of $R(\theta_k)$ evaluated at θ_k , $H(\theta_k^{l-1})$ is defined as

$$H(\theta_k^{l-1}) = \begin{bmatrix} 1 & 0 & 0 & 0 & -x_{k,j}^2(1) \sin \theta_k^{l-1} - x_{k,j}^2(2) \cos \theta_k^{l-1} & 0 \\ 0 & 0 & 1 & 0 & x_{k,j}^2(1) \cos \theta_k^{l-1} - x_{k,j}^2(2) \sin \theta_k^{l-1} & 0 \end{bmatrix} \tag{26}$$

where $x_{k,j}^2 = [x_{k,j}^2(1), x_{k,j}^2(2)]^T$.

Incorporating Equation (25) into Equation (20), the likelihood can be approximated as

$$P(x_{k,i}^1|s_k, x_{k,j}^2) \approx N(x_{k,i}^1|H(\theta_k^{l-1})s_k + R(\theta_k^{l-1})x_{k,j}^2 - \theta_k^{l-1}\bar{R}(\theta_k^{l-1})x_{k,j}^2, \Sigma). \tag{27}$$

Noticing the cumulative product in the numerator of Equation (24), we can, thus, iteratively apply each likelihood function $P(x_{k,i}^1|s_k, x_{k,j}^2)$ to update the distribution of s_k . For instance, given any $(i, j) \in \Omega_k^{p-1}$, we have

$$N(s_k|a, B)N(x_{k,i}^1|R(\theta_k)x_{k,j}^2 + \begin{bmatrix} \xi_k \\ \eta_k \end{bmatrix}, \Sigma) \propto cN(s_k|a', B') \tag{28}$$

where

$$a' = a + K(x_{k,i}^1 - R(\theta_k^{l-1})x_{k,j}^2 + \theta_i \bar{R}(\theta_k^{l-1})x_{k,j}^2 - H(\theta_k^{l-1})a) \quad (29)$$

$$B' = (I - KH(\theta_k^{l-1}))B \quad (30)$$

where I is a identity matrix and K is the Kalman gain defined by

$$K = BH(\theta_k^{l-1})^T (H(\theta_k^{l-1})BH(\theta_k^{l-1})^T + \Sigma)^{-1}. \quad (31)$$

Since c is a constant irrelevant to s_k , the above correction equations mean that the posterior distribution of s_k after combing with likelihood $P(x_{k,i}^1 | s_k, x_{k,j}^2)$, also follows Gaussian distribution with updated mean a' and covariance B' . As a result, by replacing the original a and B in Equation (19) with the latest estimates a' and B' , the above correction procedure can repeat until all $(i, j) \in \Omega_k^{p-1}$ have been used to generate the final posterior distribution of s_k .

Finally, the conditional expectation of $\log P(s_k, X_k^1, X_k^2 | U)$ can be calculated as

$$Q(U, U^{p-1}) \propto \sum_{i,j} u_{ij} E \left\{ \log P(x_{k,i}^1 | s_k, x_{k,j}^2) | U^{p-1} \right\}. \quad (32)$$

From Equation (20), we can see that the conditional expectation in Equation (32) is difficult to evaluate because $P(x_{k,i}^1 | s_k, x_{k,j}^2)$ is nonlinear with respect to θ_k . Considering that s_k follows Gaussian distribution, a special case of Monte Carlo (MC) approximation method [28] which uses the mean of s_k is applied. Therefore, we obtain

$$Q(U, U^{p-1}) \propto \sum_{i,j} u_{ij} r_{ij}^T \Sigma^{-1} r_{ij} \quad (33)$$

where $r_{ij} = x_{k,i}^1 - \begin{bmatrix} \cos \hat{\theta}_k & -\sin \hat{\theta}_k \\ \sin \hat{\theta}_k & \cos \hat{\theta}_k \end{bmatrix} x_{k,j}^2 - \begin{bmatrix} \hat{\xi}_k \\ \hat{\eta}_k \end{bmatrix}$, $\hat{\theta}_k$, $\hat{\xi}_k$, and $\hat{\eta}_k$ denote the estimated entries of s_k .

M-Step

In the M-step, the association matrix U needs to be updated by solving the following constrained optimization problem:

$$\begin{aligned} U^p &= \operatorname{argmax}_U Q(U, U^{p-1}) \propto \sum_{i,j} u_{ij} d_{ij}^2 \\ \text{s.t. } &u_{ij} \in \{0, 1\}, \sum_{j=1}^{N_k^2} u_{ij} \leq 1, \sum_{i=1}^{N_k^1} u_{ij} \leq 1 \end{aligned} \quad (34)$$

where $d_{ij}^2 = r_{ij}^T \Sigma^{-1} r_{ij}$ is the Mahalanobis distance between local tracks $x_{k,i}^1$ and $x_{k,j}^{2 \rightarrow 1}$. We notice that Equation (34) is a linear sum assignment problem (LSAP) which can be readily solved by the Hungarian algorithm in polynomial time [29]. In addition, for local tracks corresponding to the same target, d_{ij}^2 should be small, otherwise d_{ij}^2 should be large. Taking these into account, an extra thresholding step is applied to d_{ij}^2 such that local tracks with large distance are removed from association.

3.3. Track Fusion

The last stage of our proposed framework is to combine different estimates (generated by host and cooperative vehicles, respectively) of the same target vehicle into one solution. Since it is difficult to calculate the cross-correlation among multiple estimates, especially for our distributed fusion architecture,

direct application of optimal Bayes fusion can lead to overconfidence [30]. To address this problem, we apply a special version of covariance intersection (CI), termed as information theory based fast CI (IT-FCI) [27] which is given as:

$$\hat{x}_{FCI} = P_{FCI}(\omega P_1^{-1} \hat{x}_1 + (1 - \omega) P_2^{-1} \hat{x}_2) \quad (35)$$

$$P_{FCI}^{-1} = \omega P_1^{-1} + (1 - \omega) P_2^{-1} \quad (36)$$

where (\hat{x}_1, P_1) and (\hat{x}_2, P_2) denote two estimates of state corresponding to the same target, ω is the weight. Let l denote the dimension of state, then ω is determined by

$$\omega = \frac{D(p_1, p_2)}{D(p_1, p_2) + D(p_2, p_1)} \quad (37)$$

$$D(p_i, p_j) = \frac{1}{2} \left[\ln \frac{|P_j|}{|P_i|} + \sqrt{(\hat{x}_i - \hat{x}_j)^T P_j^{-1} (\hat{x}_i - \hat{x}_j) + \text{tr}(P_i P_j^{-1}) - l} \right]. \quad (38)$$

4. Performance Evaluation and Results

Currently, it is not easy to test cooperative perceptions system using real vehicles due to the high cost and potential risk [31]. Therefore, following previous studies [12], in this section, we carry out two types of computer simulation to evaluate the performance of the proposed cooperative multi-vehicle tracking.

4.1. Simulation Based on Synthesized Data

A total of seven target vehicles are moving on the region $(-800 \text{ m}, 800 \text{ m}) \times (-800 \text{ m}, 800 \text{ m})$. In addition, there are two intelligent vehicles (Car-1 and Car-2) which are equipped with sensor, and thus can sense the target vehicles in the environment. Car-1 and Car-2 is treated as the host and cooperative vehicle, respectively. Therefore, the local tracks from Car-2 are sent and fused with local tracks from Car-1 to enhance the perception performance. The perception range for each sensor is 500 m, indicating that the target vehicles with distance larger than 500 from Car-1 and Car-2 cannot be tracked. The coordinate system of Car-1 is viewed as reference system and the relative motion between the target vehicles and Car-1 is assumed to follow the near constant velocity (NCV) [32] motion model

$$x_k = \text{diag}[F, F]x_{k-1} + \text{diag}[G, G]v_k \quad (39)$$

where diag represents a diagonal matrix, target state $x_k = [\xi_k, \dot{\xi}_k, \eta_k, \dot{\eta}_k]^T$, $F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$, $G = \begin{bmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix}$, and $v_k \sim N(0, \sigma^2)$ with $\sigma = 0.5$. For each target vehicle, the position ξ_k and η_k can be observed with measurement noise following Gaussian distribution with zero mean and covariance matrix $\text{diag}[1, 1]$. For Car-2, besides the above relative motion in position, the relative rotation angle (in radian) between Car-2 and Car-1 also changes according to nonlinear model $\theta_k = 0.3 + 0.1 \sin(0.1k)$ in order to simulate the dynamic variation of vehicle heading. False alarms at any scan time are generated by Poisson distribution with mean $\lambda = 3$. The probability of detection $P_D = 0.98$. Adaptive GMPHD filtering algorithm, presented in Section 2, is conducted on Car-1 and Car-2 independently so that the local tracks can be obtained. The simulation was performed for 50 Monte Carlo runs with randomly generated process noise and measurements, thus, changing the trajectory and measurement of each target at each run. The simulation length is set to 100 s. Figure 3 show one simulation run. Figure 3a shows the relative

trajectories of seven target vehicles and Car-2 in the coordinate system of Car-1. Figure 3b,c show the measurements from both the target vehicles and false alarms in the coordinate system of Car-1 and Car-2, respectively. As can be seen from Figure 3, at different times, Car-1 and Car-2 can track a different number of target vehicles because of vehicle appearance, disappearance, or out of perception range. The uncertain miss detection, i.e., false alarm, will influence the measurements observed by Car-1 and Car-2.

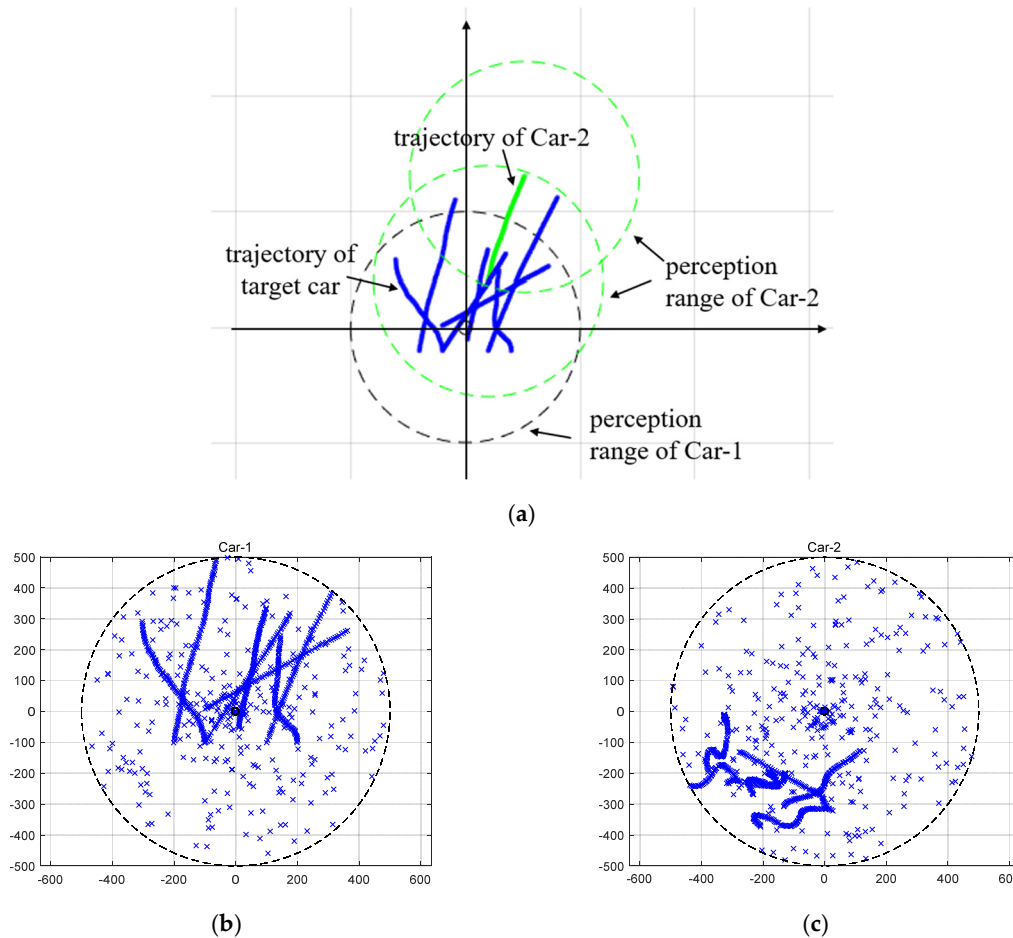


Figure 3. (a) Trajectories of target vehicles and Car-2 in the coordinate system of Car-1; (b) Measurements perceived by Car-1; and (c) Measurements perceived by Car-2.

Figure 4 shows the estimate of relative translation and orientation angle in a simulation run. As can be seen, the estimated results are rather close to the true state of Car-2 with respect to Car-1. The association between local tracks from Car-1 and Car-2 at time $k = 57$ is shown in Figure 5. To quantitatively measure the accuracy of state estimation of cooperative vehicle, we calculate the absolute error (AE) for the state estimation. For example, the AE for ξ is defined as follows:

$$E_{\xi} = \frac{1}{K} \sum_k |\xi_k - \hat{\xi}_k| \quad (40)$$

where $K = 100$ is the simulation length, ξ_k and $\hat{\xi}_k$ represent the estimated and true position of Car-2 at time k . Finally, we calculate the average, maximum, and minimum AE across all Monte Carlo runs. The overall results are shown in Table 1 showing that state estimation is rather accurate.

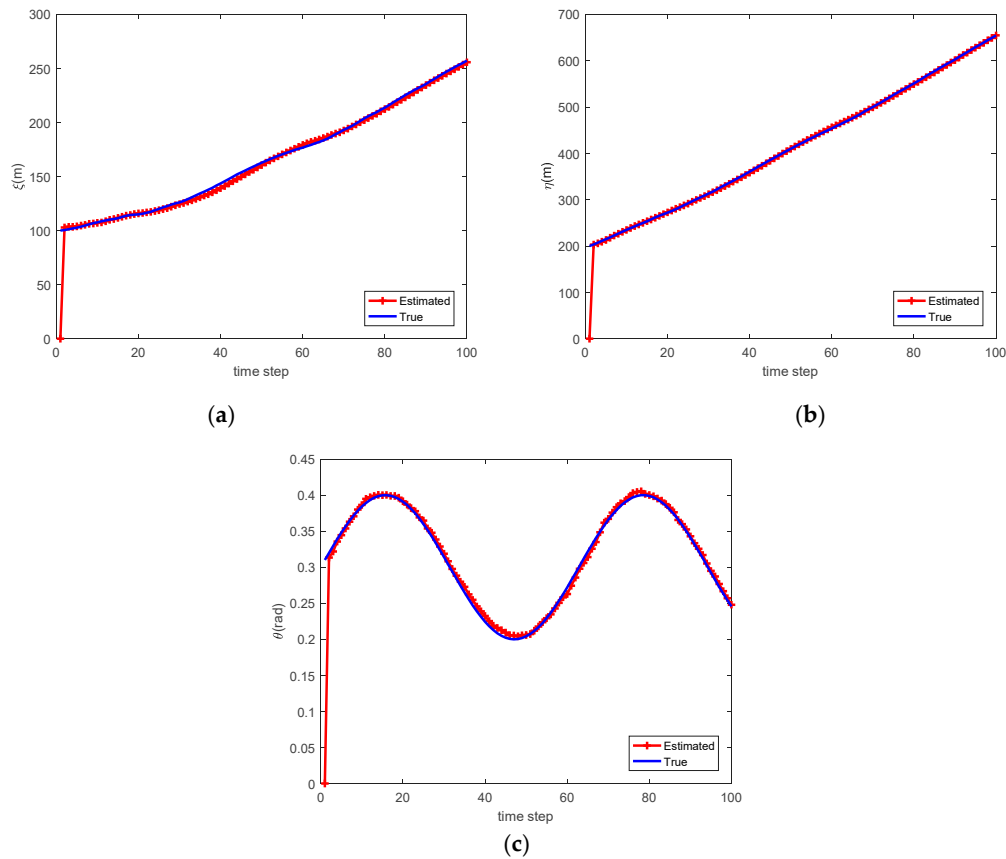


Figure 4. Estimate of (a) relative translation in x -axis; (b) relative translation in y -axis; (c) relative orientation.

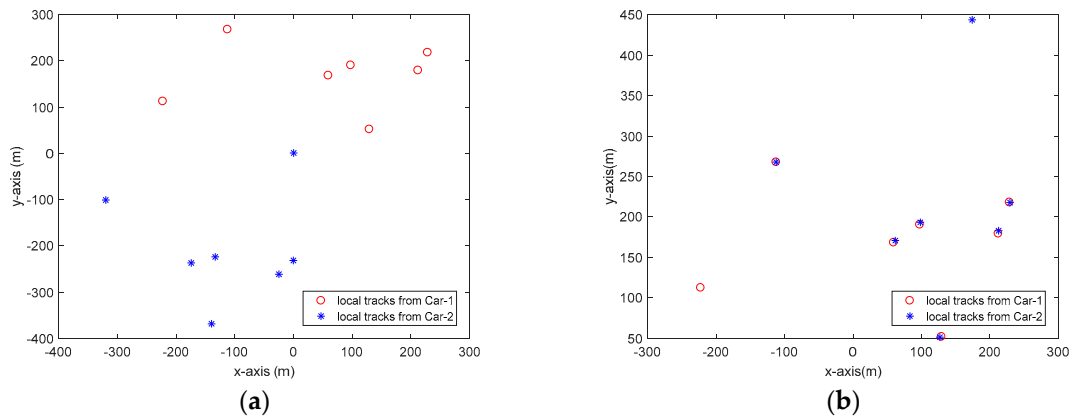


Figure 5. Illustration of local tracks from Car-1 and Car-2 at time step $k = 57$. (a) Before association; (b) After association.

Table 1. Absolute error (AE) of state estimation.

State	Average AE	Maximum AE	Minimum AE
ξ	2.8330	3.4499	2.2454
η	3.4710	4.2924	2.9480
θ	0.0071	0.0090	0.0059

To evaluate the effect of cooperative tracking, we use the criterion known as optimal subpattern assignment (OSPA) distance because it captures the difference in cardinality and individual elements between to finite sets [33,34]. The OSPA distance with order p and cut-off c is given by:

$$d_p^c[X_k, \hat{X}_k] = \left(\frac{1}{\hat{N}_k} \left(\min_{\pi(n) \in \Pi} \sum_{n=1}^{N_k} \min(c, \|x_k^n - \hat{x}_k^{\pi(n)}\|_2)^p + c^p |\hat{N}_k - N_k| \right) \right)^{1/p} \quad (41)$$

where $X_k = \{x_k^1, x_k^2, \dots, x_k^{N_k}\}$, $\hat{X}_k = \{\hat{x}_k^1, \hat{x}_k^2, \dots, \hat{x}_k^{\hat{N}_k}\}$, Π denotes all possible permutation of set $\{1, 2, \dots, \hat{N}_k\}$. The above definition is suitable when $N_k \leq \hat{N}_k$, otherwise we should use $d_p^c[\hat{X}_k, X_k]$.

In the simulation, we let $p = 1$ and $c = 50$. The Monte Carlo average of the OSPA distance obtained by Car-1, Car-2, and the fusion are shown in Table 2. In addition, we also show in Table 2 the optimal fusion result (fusion-opt), assuming the relative translation and orientation is accurately given. As can be seen, compared with independent perception, the average OSPA obtained by sensor fusion (cooperative tracking) has reduced by 27% and 43% with respect to Car-1 and Car-2, respectively. The variation of OSPA distance and the number of objects versus time in a simulation run is shown in Figure 6. As we can observe, the local tracks from Car-1 and Car-2 can be fused, thus, reducing the OSPA distance and improving the performance of MOT.

Table 2. Optimal subpattern assignment (OSPA) distance for multi-object tracking (MOT).

Method	Average OSPA	Maximum OSPA	Minimum OSPA
Car-1	3.992	5.926	2.641
Car-2	5.086	7.130	2.763
Fusion	2.896	3.937	1.930
Fusion-opt	2.063	2.662	1.583

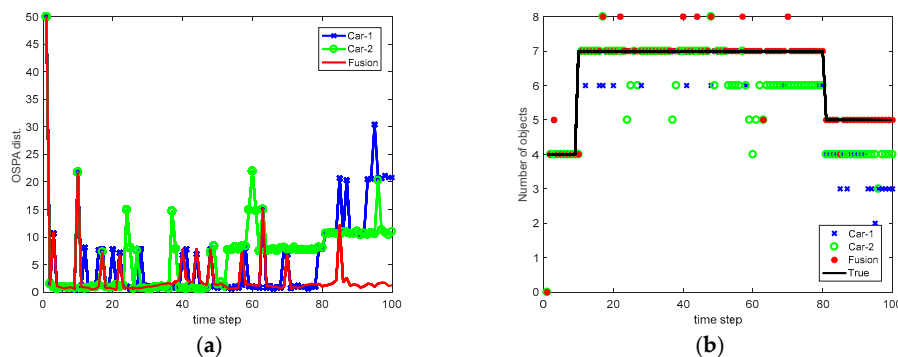


Figure 6. Results in a simulation run (a) variation of OSPA distance and (b) variation of the number of objects.

In order to compare the CPU time when dealing with an oncoming set of measurements, we show in Table 3 the average execution time (millisecond) required by Car-1, Car-2, and the fusion stage. Notice that the execution time of Car-1 and Car-2 is mainly consumed by adaptive GMPHD filtering, while fusion stage concentrates on local tracks association and covariance intersection fusion. We can see from Table 3 that the execution time consumed by sensor fusion is less than 10% of the adaptive GMPHD filtering. It indicates that the introduction of sensor fusion does not influence the efficiency, however, significantly improves the tracking performance of the whole tracking system.

Table 3. Execution time (ms) for MOT.

Method	Average Time	Maximum Time	Minimum Time
Car-1	32.86	35.01	31.43
Car-2	34.69	42.37	30.91
Fusion	2.64	3.19	2.39

4.2. Simulation Based on PreScan Platform

PreScan is a physics-based simulation platform which can be used to construct various driving environments for the design and verification of autonomous vehicles. In PreScan, a lot of elements, such as road, vehicles, sensors, vehicle-to-vehicle communication, weather etc., can be configured according to specific requirements. In order to evaluate our proposed cooperative multi-vehicle tracking system, we have built a driving scenario where 11 vehicles are deployed. Among these vehicles, two vehicles (called Car-1 and Car-2) are equipped with Radar sensor to percept surrounding vehicles. Table 4 shows the parameter of Radar. We can see from this table that for Car-1 and Car-2, only the leading vehicles with distance less than 150 m and azimuth in the range of -120° and 120° can be detected. The simulation length is 100. The simulation scenario at starting time and ending time is shown in Figure 7a,b, respectively. Car-1 and Car-2 are also marked in Figure 7 for clarity. Similar to the previous simulation, Car-1 is treated as the host vehicle, while Car-2 is treated as the cooperative vehicle. From Figure 7, we notice that both the relative translation and rotation between Car-1 and Car-2 are changing with time, especially when the two vehicles pass through the junction. In Figure 8, we illustrate the measurements observed by Car-1 and Car-2 in their respective coordinate system. For Car-1, we also show the true relative trajectories of the other vehicles. It can be observed that due to the limitation of perception range and possible occlusion between vehicles, some vehicles cannot always be detected during the simulation. Consequently, the measurements of certain vehicle cannot cover the corresponding trajectory completely.

Table 4. Radar sensor parameter configuration.

Parameter	Value
Scan pattern	Left to Right/Top to Bottom
Sweep rate	20 Hz
Beam range	150 m
Beam	120 deg
Beam	120 deg

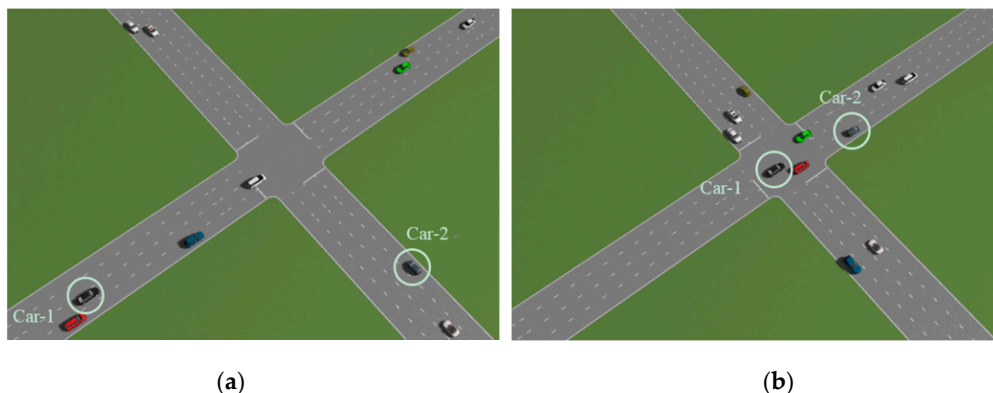


Figure 7. PreScan simulation. (a) Starting time; (b) Ending time.

The relative translation and orientation estimated based on our proposed approach are shown in Figure 9. We can see that in most cases, the estimated relative translation and orientation is rather close to the true value. Figure 10 shows the local tracks from two vehicles before and after association at time step $k = 50$. In this case, Car-1 and Car-2 can detect nine and seven vehicles, respectively. After association, local tracks from Car-1 and Car-2 can be correctly matched, thus, leading to a total of 10 vehicles being detected. The variation of OSPA distance and the number of targets versus the simulation time is shown in Figure 11. The mean OSPA distance for Car-1, Car-2, and Fusion is 12.13, 22.10, and 9.80, respectively. A comparison with the case where Car-1 and Car-2 perform tracking alone shows that the fusion of their perception results not only reduces OSPA distance but also leads to better estimation of the number of vehicles. In summary, we conclude that the cooperative tracking successfully extends the perception field of view, thus, achieving superior MOT performance.

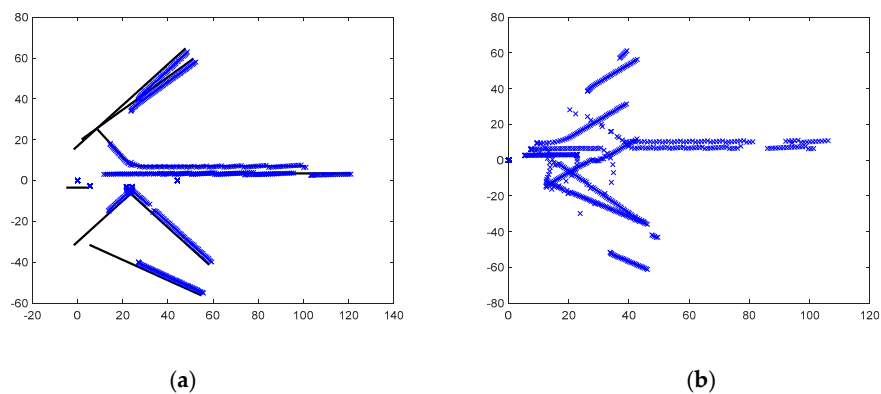


Figure 8. Illustration of observations. (a) Measurements and trajectory observed by Car-1; (b) Measurements observed by Car-2.

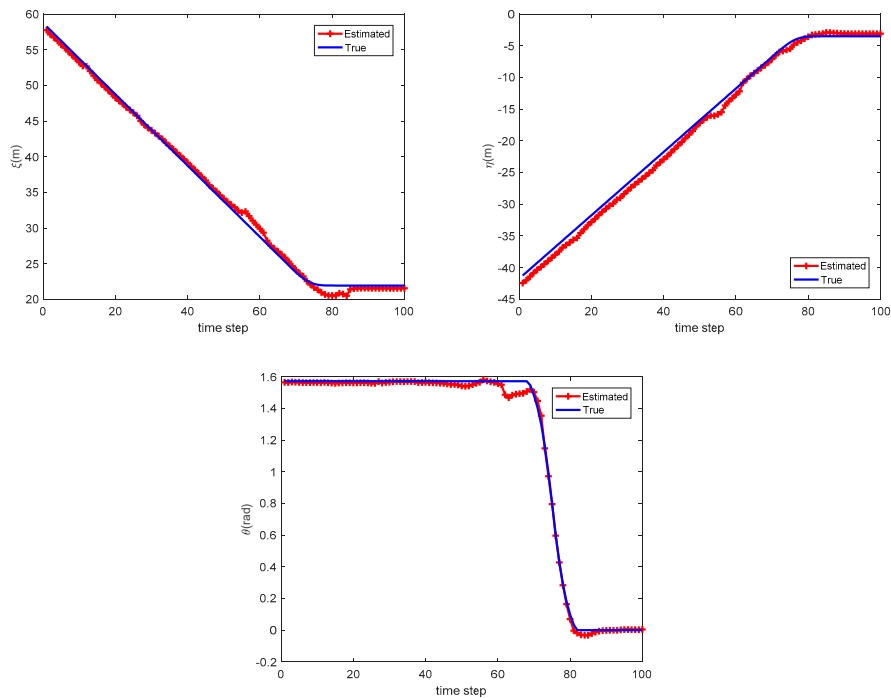


Figure 9. Estimate of relative translation and orientation.

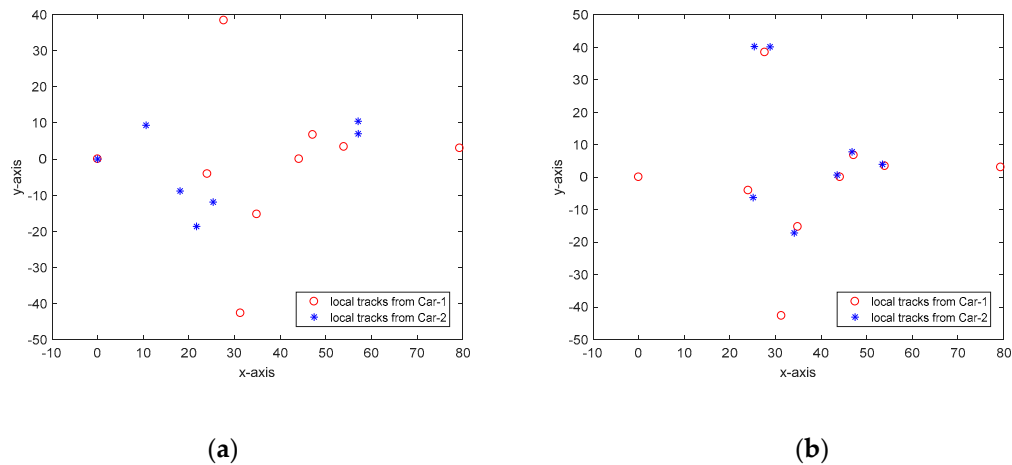


Figure 10. Illustration of local tracks from Car-1 and Car-2 at time step $k = 50$. (a) Before association; (b) After association.

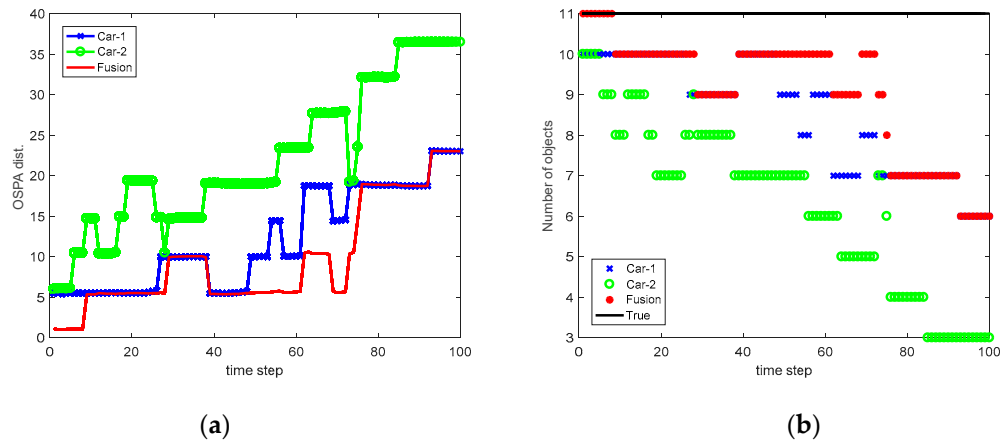


Figure 11. Results in a simulation run. (a) Variation of OSPA distance; (b) Variation of the number of objects.

5. Concluding Remarks and Future Work

In this paper, we present a novel framework for cooperative multi-vehicle tracking when the self-localization information is not available. The adaptive GMPHD filter is applied to implement effective vehicle tracking when the intensity of newborn target is infeasible to define in advance. A Bayes formulation for joint track association and relative pose estimation is developed and the solution is derived by following the EM algorithm. Finally, the tracks associated successfully are fused by fast covariance intersection based on theory information. The simulation results demonstrate the relative pose between host and cooperative vehicles can be inferred accurately in most cases. In addition, with slightly increased computational costs, the cooperative multi-vehicle tracking demonstrates clear advantage over the non-cooperative tracking algorithm in terms of perception performance.

In reality, communication delay is another important factor that affects the performance of cooperative perception [35], especially when the bandwidth of the wireless channel is limited. Therefore, suitable temporal alignment is necessary to account for the time bias caused by the communication delay and algorithm execution. The simplest approach is based on the prediction model for the compensation of communication delay. This work assumes the tracks from different vehicles have been synchronized to the same time instant before track association and relative pose estimation. In addition, loss of communication

links also hinders the application of our proposed algorithm. Actually, when the communication links interrupt, the host vehicle cannot receive the message from the cooperative vehicle, and thus fail to enhance the perception ability by fusion. After the communication links recover, the proposed algorithm can be performed by the host vehicle once the message from cooperative vehicle arrives. In future work, we plan to investigate the integration of temporal alignment into our framework to further enhance the performance of multi-vehicle tracking. Moreover, the extension of our approach to explicitly consider the effect of communication delays and failure is an interesting direction. One possible solution is to integrate these factors into our probabilistic model by introducing new variables. We would discuss these problems in future work.

Author Contributions: Methodology, X.C.; Writing—original draft preparation, J.J. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the National Key Research and Development Program of China (2018YFB0105000), the National Natural Science Foundation of China (grant nos. 61773184, 51875255, 6187444, U1564201, U1664258, U1762264, and 61601203), Six Talent Peaks project of Jiangsu Province (grant no. 2017-JXQC-007), the Key Research and Development Program of Jiangsu Province (BE2016149), the Natural Science Foundation of Jiangsu Province (BK2017153), the Key Project for the Development of Strategic Emerging Industries of Jiangsu Province (2016-1094, 2015-1084), and the Talent Foundation of Jiangsu University, China (no. 14JDG066).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kim, S.-W.; Qin, B.; Chong, Z.J.; Shen, X.; Liu, W.; Ang, M.H.; Frazzoli, E.; Rus, D. Multivehicle cooperative driving using cooperative perception: Design and experimental validation. *IEEE Trans. Intell. Transp. Syst.* **2014**, *16*, 663–680. [[CrossRef](#)]
2. Wang, H.; Yu, Y.; Cai, Y.; Chen, X.; Chen, L.; Liu, Q. A Comparative Study of State-of-the-Art Deep Learning Algorithms for Vehicle Detection. *IEEE Intell. Transp. Syst. Mag.* **2019**, *11*, 82–95. [[CrossRef](#)]
3. Kim, C.; Li, F.; Ciptadi, A.; Rehg, J.M. Multiple hypothesis tracking revisited. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4696–4704.
4. Hamid Rezatofighi, S.; Milan, A.; Zhang, Z.; Shi, Q.; Dick, A.; Reid, I. Joint probabilistic data association revisited. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3047–3055.
5. Mahler, R.P. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2007.
6. Vo, B.N.; Vo, B.T.; Clark, D.; Mallick, M.; Krishnamurthy, V. Bayesian multiple target filtering using random finite sets. *Integr. Track. Classif. Sens. Manag.* **2013**, 75–125. [[CrossRef](#)]
7. Vo, B.-N.; Ma, W.-K. The Gaussian mixture probability hypothesis density filter. *IEEE Trans. Signal Process.* **2006**, *54*, 4091–4104. [[CrossRef](#)]
8. Hult, R.; Campos, G.R.; Steinmetz, E.; Hammarstrand, L.; Falcone, P.; Wymeersch, H. Coordination of cooperative autonomous vehicles: Toward safer and more efficient road transportation. *IEEE Signal Process. Mag.* **2016**, *33*, 74–84. [[CrossRef](#)]
9. Kim, S.-W.; Liu, W.; Ang, M.H.; Frazzoli, E.; Rus, D. The impact of cooperative perception on decision making and planning of autonomous vehicles. *IEEE Intell. Transp. Syst. Mag.* **2015**, *7*, 39–50. [[CrossRef](#)]
10. Chen, Q.; Tang, S.; Yang, Q.; Fu, S. Cooper: Cooperative perception for connected autonomous vehicles based on 3d point clouds. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019; pp. 514–524.
11. Thomaidis, G.; Vassilis, K.; Lytrivis, P.; Tsogas, M.; Karaseitanidis, G.; Amditis, A. Target tracking and fusion in vehicular networks. In Proceedings of the 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, 5–9 June 2011; pp. 1080–1085.

12. Vasic, M.; Martinoli, A. A collaborative sensor fusion algorithm for multi-object tracking using a Gaussian mixture probability hypothesis density filter. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas, Spain, 15–18 September 2015; pp. 491–498.
13. Uhlmann, J.K. General data fusion for estimates with unknown cross covariances. In *Signal Processing, Sensor Fusion, and Target Recognition V*; International Society for Optics and Photonics: Orlando, FL, USA, 1996; pp. 536–547.
14. Vasic, M.; Lederrey, G.; Navarro, I.; Martinoli, A. An overtaking decision algorithm for networked intelligent vehicles based on cooperative perception. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 19–22 June 2016; pp. 1054–1059.
15. Yoon, D.D.; Ali, G.; Ayalew, B. Data Association and Fusion Framework for Decentralized Multi-Vehicle Cooperative Perception. In Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Anaheim, CA, USA, 18–21 August 2019.
16. Vasuhi, S.; Vaidehi, V. Target tracking using interactive multiple model for wireless sensor network. *Inf. Fusion* **2016**, *27*, 41–53. [[CrossRef](#)]
17. Niehsen, W. Information fusion based on fast covariance intersection filtering. In Proceedings of the Fifth International Conference on Information Fusion, FUSION 2002, (IEEE Cat. No. 02EX5997), Annapolis, MD, USA, 8–11 July 2002; pp. 901–904.
18. Allig, C.; Wanielik, G. Alignment of perception information for cooperative perception. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 1849–1854.
19. Yuan, T.; Krishnan, K.; Chen, Q.; Breu, J.; Roth, T.B.; Duraisamy, B.; Weiss, C.; Maile, M.; Gern, A. Object matching for inter-vehicle communication systems—An IMM-based track association approach with sequential multiple hypothesis test. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 3501–3512. [[CrossRef](#)]
20. Sakr, A.H.; Bansal, G. Cooperative localization via DSRC and multi-sensor multi-target track association. In Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 1–4 November 2016; pp. 66–71.
21. Rauch, A.; Klanner, F.; Rasshofer, R.; Dietmayer, K. Car2x-based perception in a high-level fusion architecture for cooperative perception systems. In Proceedings of the 2012 IEEE Intelligent Vehicles Symposium, Alcalá de Henares, Spain, 3–7 June 2012; pp. 270–275.
22. Amini, A.; Vaghefi, R.M.; Jesus, M.; Buehrer, R.M. GPS-free cooperative mobile tracking with the application in vehicular networks. In Proceedings of the 2014 11th Workshop on Positioning, Navigation and Communication (WPNC), Dresden, Germany, 12–13 March 2014; pp. 1–6.
23. Lyu, Y.; Pan, Q.; Lv, J. Unscented Transformation-Based Multi-Robot Collaborative Self-Localization and Distributed Target Tracking. *Appl. Sci.* **2019**, *9*, 903. [[CrossRef](#)]
24. Teng, J.; Snoussi, H.; Richard, C.; Zhou, R. Distributed variational filtering for simultaneous sensor localization and target tracking in wireless sensor networks. *IEEE Trans. Veh. Technol.* **2012**, *61*, 2305–2318. [[CrossRef](#)]
25. Ma, Y.; Zhang, T.; Tian, X. Cooperative Vehicle Sensing and Obstacle Avoidance for Intelligent Driving Based on Bayesian Frameworks. In *International Conference in Communications, Signal Processing, and Systems*; Springer: Singapore, 2017; pp. 566–573.
26. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
27. Hurley, M.B. An information theoretic justification for covariance intersection and its generalization. In Proceedings of the Fifth International Conference on Information Fusion, FUSION 2002, (IEEE Cat. No. 02EX5997), Annapolis, MD, USA, 8–11 July 2002; pp. 505–511.
28. Neath, R.C. On convergence properties of the Monte Carlo EM algorithm. In *Advances in Modern Statistical Theory and Applications: A Festschrift in Honor of Morris L. Eaton*; Institute of Mathematical Statistics: Washington, DC, USA, 2013; pp. 43–62.
29. Munkres, J. Algorithms for the assignment and transportation problems. *J. Soc. Ind. Appl. Math.* **1957**, *5*, 32–38. [[CrossRef](#)]
30. Battistelli, G.; Chisci, L.; Fantacci, C.; Farina, A.; Graziano, A. Consensus CPHD filter for distributed multitarget tracking. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 508–520. [[CrossRef](#)]

31. Bhat, A.; Aoki, S.; Rajkumar, R. Tools and methodologies for autonomous driving systems. *Proc. IEEE* **2018**, *106*, 1700–1716. [[CrossRef](#)]
32. Da, K.; Li, T.; Zhu, Y.; Fu, Q. A Computationally Efficient Approach for Distributed Sensor Localization and Multitarget Tracking. *IEEE Commun. Lett.* **2019**, *24*, 335–338. [[CrossRef](#)]
33. Schuhmacher, D.; Vo, B.-T.; Vo, B.-N. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [[CrossRef](#)]
34. Ristic, B.; Vo, B.-N.; Clark, D.; Vo, B.-T. A metric for performance evaluation of multi-target tracking algorithms. *IEEE Trans. Signal Process.* **2011**, *59*, 3452–3457. [[CrossRef](#)]
35. Yee, R.; Chan, E.; Cheng, B.; Bansal, G. Collaborative perception for automated vehicles leveraging vehicle-to-vehicle communications. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; pp. 1099–1106.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).