

Cooperative Multiple Dynamic Object Tracking on Moving Vehicles based on Sequential Monte Carlo Probability Hypothesis Density Filter

Jonathan Gan

Milos Vasic

Alcherio Martinoli

Abstract—This paper proposes a generalized method for tracking of multiple objects from moving, cooperative vehicles – bringing together an Unscented Kalman Filter for vehicle localization and extending a Sequential Monte Carlo Probability Hypothesis Density filter with a novel cooperative fusion algorithm for tracking. The latter ensures that the fusion of information from cooperating vehicles is not limited to a fully overlapping Field Of View (FOV), as usually assumed in popular distributed fusion literature, but also allows for a perceptual extension corresponding to the union of the vehicles’ FOV. Our method hence allows for an overall extended perception range for all cooperative vehicles involved, while preserving same or improving the accuracy in the overlapping FOV. This method also successfully mitigates noisy sensor measurement and clutter, as well as localization inaccuracies of tracking vehicles using Global Navigation Satellite Systems (GNSS). Finally, we extensively evaluate our method using a high-fidelity simulator for vehicles of varying speed and trajectories.

I. INTRODUCTION

The recent advent of commercial highly automated vehicles driven by advancements in autonomous vehicles research and the rise of trends in information technology such as big data and the “internet of things” have pushed automobile companies as well as supranational organizations to explore options for allowing these highly automated vehicles to cooperate with one another as well as with the surrounding infrastructure [1]. The goal of such a cooperation is to eventually increase road traffic safety and efficiency [2]. Despite the presence of advanced sensors on board, such measures are welcome because perceptual capabilities of a single highly automated vehicle often suffer from a limited FOV, occlusion, and an inherent sensor inaccuracy.

When considering vehicle-to-vehicle cooperation, the possible communication bottlenecks and the richness in data from advanced sensors imply that the naive communication of raw sensor data would not be favorable for such a task. Instead, processed and filtered data representing other vehicles or objects is preferred.

The effective and robust tracking of these targets also relies on the accurate localization of the tracking vehicle for a reliable representation of targets in a global reference frame.

This work has been carried out entirely at the Distributed Intelligent Systems and Algorithms Laboratory (DISAL), School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. J. Gan is currently with the Department of Mechanical and Process Engineering, Eidgenössische Technische Hochschule (ETH) Zürich, Switzerland, while M. Vasic and A. Martinoli are with DISAL. `jongan@student.ethz.ch`, `{milos.vasic, alcherio.martinoli}@epfl.ch`. M. Vasic is supported by PSA Groupe.

II. RELATED WORK

A solution that has gained attraction in the recent years in addressing the robust tracking of multiple targets is based on Random Finite Set (RFS) models, in which a set of objects of variable cardinality is modeled as a random finite set. Specifically, we focus on Probability Hypothesis Density (PHD) and Cardinalized PHD (CPHD) filters – both based on RFS models – as they deal with the *target-measurement association* implicitly. Prior approaches using explicit association [3], [4] are faced with the issues of clutter, where multiple measurements for a single target appear, as well as false associations. PHD filters are furthermore able to robustly and accurately track targets without having prior knowledge over the number of targets, unlike the Joint Probabilistic Data Association (JPDA) filter [4].

Common implementations of the PHD filter include the Gaussian Mixture (GM) and Sequential Monte Carlo (SMC) PHD filter. Vasic et al. [5], for example, extended the GM-PHD filter introduced in [6] to involve cooperation and fusion of information between two agents each running an instance of the GM-PHD filter.

While the case for the feasibility of a collaborative GM-PHD filter (C-GM-PHD) in tracking multiple targets was presented in [5], the C-GM-PHD approach has limitations in the fact that single-target transitional densities and likelihood functions must be Gaussian [7]. Transitional densities also have to be linear or approximately linear in order for the GM-PHD and approaches based on it to work.

As previously mentioned, SMC methods represent an alternative in approximating the PHD filter as shown in [8] and are able to handle nonlinear targets with relatively low complexity. Ristic et al. proposed an improved SMC approach in [9] that is able to reliably extract state estimates without the use of an explicit clustering algorithm. Clustering algorithms such as K -means or based on the Expectation-Maximization (EM) principle require a Gaussian assumption or an a priori knowledge of the number of targets, and hence are counterproductive to the advantages of the nonlinear SMC-PHD filter.

The distributed fusion of information from multiple instances of the SMC-PHD filter was explored in [10] by using Exponential Mixture Densities (EMD). However, it assumes that all agents share a common FOV in which targets are detected. This is a highly limiting assumption especially in the automotive domain where cooperation between vehicles should aim to increase the amount of information each vehicle has of its surroundings. Furthermore, applications

of the SMC-PHD filter and its distributed variation have been only experimentally verified with stationary sensors, emulating its application largely in the field of defense.

The global localization of moving vehicles, on the other hand, is often obtained by onboard consumer GNSS receivers which are characterized by significant uncertainty of a few meters on average, varying with atmospheric factors and surrounding environment [11]. An accurate localization of the tracking vehicle must be ensured in order for a feasible cooperation between vehicles to be possible.

One of the most common approaches for an accurate localization of targets with nonlinear models despite noisy sensor measurements is the Unscented Kalman Filter (UKF) [12], which resolves the main issue of the Extended Kalman Filter (EKF) being difficult to tune and yielding a large uncertainty [13]. The UKF has also been used for state estimation of land-based vehicles [14], [15].

We propose an integrated approach of localization with multiple target tracking, factoring uncertainties from the localization itself into tracking. Furthermore, we also propose an improvement to the distributed fusion of SMC-PHD filters by generalizing it for non-overlapping fields of views and hence increase the applicability of cooperative multi-target tracking. We refer to our generalized approach, from this point forth, as the Cooperative SMC-PHD filter. The goal of our approach is to improve *tracking accuracy* as well as provide each tracking vehicle with additional information beyond its own FOV.

The remainder of this paper is organized as follows. We start off with a brief overview of the SMC-PHD filter and the distributed fusion of multiple SMC-PHD filter instances as proposed in [10] and proceed to introduce our extension to the current approach. A short understanding of the UKF is also provided as well as an explanation as to how uncertainties from the UKF are factored into tracking. We conclude the paper with experimental results and their discussion.

III. COOPERATIVE SMC-PHD FILTER

In addressing the problem of multiple object tracking, we propose the use of a Cooperative SMC-PHD (C-SMC-PHD) filter that builds upon the SMC-PHD filter introduced in [9] and extends it with fusion of SMC-PHD densities. Our novel fusion approach generalizes the EMD method introduced in [10] by dealing with information acquired by a cooperating vehicle outside of the tracking vehicle's FOV.

The PHD overcomes the impracticability of the implementation of a multi-target Bayes nonlinear filter by propagating only the first-order moment of the multi-object posterior densities [16]. Like any recursive Bayesian filter, the PHD filter contains the prediction and update steps, both of which are carried out by a random sample approximation using an SMC implementation, such that the posterior of the PHD at time t given observations up to and including at time t is approximated by a random sample set as follows:

$$D_{t|t}(\mathbf{x}) \approx \sum_{n=1}^N w_{t|t}^n \delta_{\mathbf{x}_{t|t}^n}(\mathbf{x}) \quad (1)$$

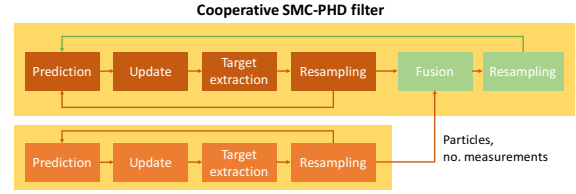


Fig. 1. The C-SMC-PHD filter illustrated. Fusion of intensities from cooperating agents is carried out only if information is available. Otherwise, the filter of the tracking agent continues without cooperation. The blocks in green represent our contributions while the other blocks represent the state-of-the-art approach.

where $\delta(\mathbf{x})$ is the Dirac delta function and $\{(w_{t|t}^n, \mathbf{x}_{t|t}^n)\}_{n=1}^N$ a weighted particle set comprising of N random samples $\mathbf{x}_{t|t}^n$ and its corresponding weights $w_{t|t}^n$. Furthermore, the number of targets can be extracted from the particle set by summing up the importance weights for large values of N . Ristic et al. also proposed in [9] that particles should be *spawned* around the measurement set \mathbf{Z}_t at every time step to avoid inefficient use of particles in estimating the posterior. The distinction between *newborn* and *persistent* particles is evident through the recursive implementation of the filter.

The general concept of the C-SMC-PHD is hence a five-step process illustrated in Fig. 1, where the prediction, update and target extraction steps are taken from [9] and are summarized below for completeness.

A. Prediction

In the prediction step, ρ newborn particles of states $\mathbf{x}_{t|t-1,b}^n$ are generated around each measurement $\mathbf{z} \in \mathbf{Z}_t$ with a uniform weight

$$w_{t|t-1,b}^n = \frac{\nu_{t|t-1}^b}{\rho m_t} \quad (2)$$

where m_t is the cardinality of measurements and $\nu_{t|t-1}^b$ is the expected number of newborn targets between time $t-1$ and t . Additionally, the states of persistent particles $\mathbf{x}_{t|t-1,p}^n$ are propagated according to the dynamics defined by the single-object transition density. Weights of persistent particles are also revised according to the survival probability p_S given by:

$$w_{t|t-1,p}^n = p_S(\mathbf{x}_{t-1}^n) w_{t-1,p}^n. \quad (3)$$

B. Update

In this step, the weights of the persistent ($w_{t|t-1,p}^n$) and newborn ($w_{t|t-1,b}^n$) particles representing both their respective PHDs are updated as such

$$w_{t|t,p}^n = \left[1 - p_D(\mathbf{x}_{t|t-1,p}^n) \right] w_{t|t-1,p}^n + \sum_{\mathbf{z} \in \mathbf{Z}_t} \frac{p_D(\mathbf{x}_{t|t-1,p}^n) g_t(\mathbf{z} | \mathbf{x}_{t|t-1,p}^n) w_{t|t-1,p}^n}{\mathcal{L}(\mathbf{z})} \quad (4)$$

for persistent particles and

$$w_{t|t,b}^n = \sum_{\mathbf{z} \in \mathbf{Z}_t} \frac{w_{t|t-1,b}^n}{\mathcal{L}(\mathbf{z})} \quad (5)$$

for newborn particles where

$$\mathcal{L}(\mathbf{z}) = \kappa_t(\mathbf{z}) + \sum_{n=1}^{\rho m_t} w_{t|t-1,p}^n + \sum_{n=1}^{N_{t-1}} p_D(\mathbf{x}_{t|t-1,p}^n) g_t(\mathbf{z}|\mathbf{x}_{t|t-1,p}^n) w_{t|t-1,p}^n. \quad (6)$$

$p_D(\mathbf{x})$ furthermore represents the detection probability of a target with state \mathbf{x} while $g_t(\mathbf{z}|\mathbf{x})$ represents the single-target measurement likelihood at time t . Measurement-associated clutter is also factored in with the term $\kappa_t(\mathbf{z})$.

C. State estimation extraction

It is noted that the second term in (4) can essentially be viewed as a sum of weights corresponding to each measurement $\mathbf{z} \in \mathbf{Z}_t$. If each measurement in the measurement set is indexed $\mathbf{Z}_t = \{\mathbf{z}_{t,j} | j = 1, 2, \dots, m_t\}$, an additional weight metric for each measurement of index j where $j = \{0, 1, \dots, m_t\}$ can be defined. The index $j = 0$ represents the measurement-independent first term of (4).

$$w_{t|t,p}^{n,j} = \begin{cases} [1 - p_D(\mathbf{x}_{t|t-1,p}^n)] w_{t|t-1,p}^n & j = 0 \\ \frac{p_D(\mathbf{x}_{t|t-1,p}^n) g_t(\mathbf{z}_{t,j}|\mathbf{x}_{t|t-1,p}^n) w_{t|t-1,p}^n}{\mathcal{L}(\mathbf{z})} & j = 1, \dots, m_t \end{cases}. \quad (7)$$

Therefore, each measurement index j can be given an associated weight $W_{t,p}^j = \sum_{n=1}^{N_{t-1}} w_{t|t,p}^{n,j}$ with $0 \leq W_{t,p}^j \leq 1$. Intuitively, $W_{t,p}^j$ would be high for a certain measurement if $\mathbf{z}_{t,j}$ resulted in non-zero likelihood for some particles, indicating that the measurement originated from a target, and close to zero otherwise. The state estimate to each measurement index is then only reported if the weight $W_{t,p}^j$ above a given reporting threshold c_R .

Correspondingly, the state estimate $\hat{\mathbf{x}}_{t,j}$ and its associated covariance matrix $\mathbf{P}_{t,j}$ can be obtained as:

$$\hat{\mathbf{x}}_{t,j} = \sum_{n=1}^{N_{t-1}} w_{t|t,p}^{n,j} \mathbf{x}_{t|t-1,p}^n$$

$$\mathbf{P}_{t,j} = \sum_{n=1}^{N_{t-1}} w_{t|t,p}^{n,j} (\mathbf{x}_{t|t-1,p}^n - \hat{\mathbf{x}}_{t,j}) (\mathbf{x}_{t|t-1,p}^n - \hat{\mathbf{x}}_{t,j})^T. \quad (8)$$

D. Resampling

The resampling step represents a core component of SMC methods and ensures that the filter contains more particles of higher weight and fewer of those with lower weights. In our implementation, we use the low variance resampling [17] (or more recently known as the resampling wheel). At the end of this step, the importance weights of all persisting particles are set to $1/N_{t-1}$ since weight-based likelihoods are replaced by frequencies when samples are drawn by weights.

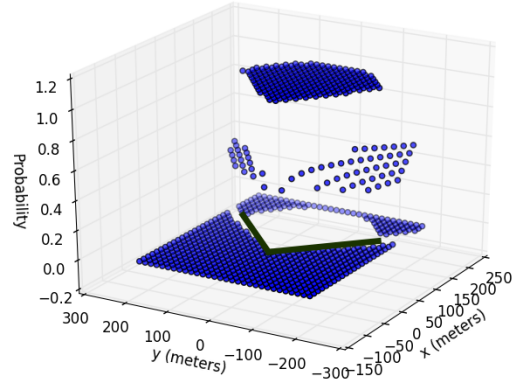


Fig. 2. The logistic function with $k = \pm 8$ with θ_{sup} being the extreme angles of the field of view of the LIDAR sensor.

E. Fusion

Given that each agent runs an instance of the SMC-PHD filter as shown in Fig. 1, the goal of fusion is to obtain the fused PHD posterior intensities of both PHD posteriors. In this case, our contribution involves providing a generalized approach to fuse PHD posterior intensities where FOVs of the cooperating agents are not completely overlapping. This is achieved by using the EMD approach introduced in [10] for completely overlapping segments of the field of view and a novel method for incorporating information from the cooperating agent that is outside of the tracking agent's FOV.

Upon receiving particles from the cooperating agent and transforming them to the global reference frame, we cluster particles over the whole state space using the DBSCAN algorithm [18] to obtain a set of R particle clusters $\{r_l | l = 1, 2, \dots, R\}$ and obtain the mean of each cluster \bar{p}_l . Instead of a binary assignment $\bar{p}_l \in FOV_{own}$ or $\bar{p}_l \notin FOV_{own}$, we employ the logistic function defined as:

$$f_p(x) = \frac{1}{1 + \exp(-k \cdot \text{ang}_{\theta_{sup}}(x))} \quad (9)$$

to obtain a continuous probability weighting in the immediate vicinity of FOV_{own} . $\text{ang}_{\theta_{sup}}(x)$ represents the angular difference from particle x to the extreme angles θ_{sup} of the field of view. The logistic function is hence calculated with respect to the angle of the FOV as better visualized in Fig. 2. A simple heuristic is used to decide whether to consider received particles as *external* to the tracking vehicle's FOV or as *common* to both FOVs is given in lines 1 – 14 of Algorithm 1. This non-binary classification of particles ensures that transitions between fields of view occur smoothly and target estimates are not lost.

Algorithm 1 Procedure for classifying received particles.

```

1:  $\xi_{ext} \leftarrow \emptyset$  ▷ ext. particle set
2:  $\xi_{cm} \leftarrow \emptyset$  ▷ common particle set
3: for  $r_l \in \{r_l | l = 1, 2, \dots, R\}$  do
4:    $p \leftarrow f_p(\bar{p}_l)$ 
5:   if  $p = 0$  then
6:     Insert  $r_l$  into  $\xi_{ext}$ .
7:   else
8:     Penalize  $r_l$  by  $p$ .
9:     Insert  $r_l$  into  $\xi_{cm}$ .
10:    if  $p < 1$  then
11:      Insert  $r_l$  into  $\xi_{ext}$ .
12:    end if
13:  end if
14: end for
15:  $\xi_{intersect} \leftarrow \emptyset$ 
16:  $\mathcal{P}_{own} \leftarrow \text{Polygon}(FOV_{own})$  ▷ Create polygon from FOV
17:  $\mathcal{P}_{other} \leftarrow \text{Polygon}(FOV_{other})$ 
18:  $\mathcal{P}_{intersect} \leftarrow \mathcal{P}_{own} \cap \mathcal{P}_{other}$ 
19: for  $(w^n, \mathbf{x}^n) \in \{(w^n, \mathbf{x}^n)\}_{n=1}^N$  do ▷ own particle set
20:   if  $\mathbf{x}^n \in \mathcal{P}_{intersect}$  then
21:     Insert  $(w^n, \mathbf{x}^n)$  into  $\xi_{intersect}$ 
22:   end if
23: end for
24:  $\xi_{cm} \leftarrow \xi_{cm} \cup \xi_{intersect}$ 
25: return  $\xi_{ext}, \xi_{cm}$ 

```

Bearing in mind that the fusion proposed in [10] require that two fields of view intersect, we are still left with removing particles that might be in $FOV_{internal} := FOV_{own} \setminus FOV_{other}$. This is achieved by determining the area defined by the intersection of two fields of view FOV_{own} and FOV_{other} by representing them as convex polygons. The intersection set of two convex polygons can be obtained in linear time [19] and is also a convex polygon itself. Lines 15 – 24 elaborate on the procedure for finding particles that belong to the common FOV.

1) *Fusion within the Common FOV:* Given two PHD posterior intensities of two i.i.d. cluster distributions f_i and f_j , the EMD of these two cluster distributions is also an i.i.d. cluster distribution defined as

$$f_{(\cdot)}(X) = n! \cdot \mu_{(\cdot)}(n) \prod_{x \in X} s_{(\cdot)}(x) \quad (10)$$

where

$$s_\omega(x) = \frac{s_i^{(1-\omega)}(x) s_j^\omega(x)}{Z_\omega(s_i, s_j)}, \quad (11)$$

$$\mu_\omega = \mu_i^{(1-\omega)} \cdot \mu_j^\omega \cdot Z_\omega(s_i, s_j) \quad (12)$$

$$Z_\omega(s_i, s_j) = \int s_i(x)^{(1-\omega)} s_j(x)^\omega dx. \quad (13)$$

In our case, we assume a Poisson distributed cardinality from both nodes given by the means $\mu_{(\cdot)}$. The resulting cardinality is hence also Poisson distributed.

The parameter $\omega \in [0, 1]$ determines the relative weight assigned to each distribution and needs to be extrinsically chosen. An optimal ω can be obtained by minimizing a cost function defined by the Renyi Divergence (RD) given its relevance in sensor management [20] and is given by

$$\omega^* = \arg \min_{\omega \in [0, 1]} (R_\alpha(f_\omega || f_i) - R_\alpha(f_\omega || f_j))^2 \quad (14)$$

where

$$R_\alpha(f_\omega || f_i) = \frac{1}{\alpha - 1} \log \sum_{n=0}^{k_{max}} \left[\mu_\omega^\alpha(n) \mu_i^{(1-\alpha)}(n) \left(\frac{Z_{\alpha\omega}(s_i, s_j)}{Z_\omega(s_i, s_j)^\alpha} \right)^n \right] \quad (15)$$

$$R_\alpha(f_\omega || f_j) = \frac{1}{\alpha - 1} \log \sum_{n=0}^{k_{max}} \left[\mu_\omega^\alpha(n) \mu_j^{(1-\alpha)}(n) \left(\frac{Z_{\alpha(1-\omega)}(s_i, s_j)}{Z_\omega(s_i, s_j)^\alpha} \right)^n \right] \quad (16)$$

and k_{max} representing the maximum number of targets that would be tracked by the filter at any time. $\mu_x(n)$ represents the probability mass function of the Poisson distribution over n with a mean λ_x being the cardinality of node x .

Bearing in mind that both PHD posteriors are implemented with SMC methods, the continuous approximations of the distributions have to be reconstructed from the particle sets of both distributions. This is achieved with the Kernel Density Estimation (KDE) method [21], in which the estimated density is obtained from the sum of kernel functions at particle points. An explicit particle to measurement association is also created by clustering particles using the DBSCAN [18] algorithm to construct separate KDEs over each cluster to ensure the closest approximation of the true posterior density instead of an oversmoothed density function.

For each cluster $k \in \mathcal{K} = \{k_1, \dots, k_K\}$, the covariance \mathbf{C}_k is computed to find the density estimate as such

$$\hat{s}(x | Z_{1:t}) = \frac{1}{M} \sum_{m=1}^M K_h(x, x^{(m)}; \mathbf{C}_{k^m}) \quad (17)$$

where the *bandwidth* $h = n^{-1/(d+4)}$ is calculated according to Scott's Rule [22], with n the number of data points and d the number of dimensions.

The union of particle sets P_U from both nodes with $M_\omega = M_i + M_j$ particles is obtained to sample from the fused localization density (11) in order to obtain estimates $\hat{s}_i(x)$ and $\hat{s}_j(x)$. Thereafter, the estimate of (13) is given as

$$\hat{Z}_\omega(s_i, s_j) := \sum_{x \in P_U} \frac{\hat{s}_i^{(1-\omega)}(x) \hat{s}_j^\omega(x)}{M_i \hat{s}_i(x) + M_j \hat{s}_j(x)}. \quad (18)$$

The weight of the m' -th particle in P_U is also estimated as

$$\hat{\zeta}^{(m')} \propto \frac{\hat{s}_i^{(1-\omega)}(x^{(m')}) \hat{s}_j^\omega(x^{(m')})}{M_i \hat{s}_i(x^{(m')}) + M_j \hat{s}_j(x^{(m')})}, \quad (19)$$

resulting in the union particle set $\{(\zeta^{(m')}, x^{(m')})\}_{m'=1}^{M_\omega}$.

Having obtained (18), (14) can be computed for a different values of ω using exhaustive search in equidistantly distributed points over the interval $[0, 1]$. Finally, the weights in (19) can be obtained for the optimal ω^* .

Thereafter, the particle set P_U and the set of particles in $FOV_{internal}$ are resampled separately. The union of P_U and the particles in $FOV_{internal}$ is created to be used in the prediction and update steps of the next iteration.¹

2) *Incorporating External Information:* Unlike particles in the common FOV that can be predicted and updated according to measurements, external particles can only be predicted by the transitional densities since the tracking agent receives no associated measurement. Algorithm 2 briefly explains the handling and propagation of external particles.

Algorithm 2 Procedure for handling particles representing external targets at a timestep t

- 1: **if** New information from cooperating agent available **then**
 - 2: Replace external particle set with incoming information from cooperating agent.
 - 3: **end if**
 - 4: External particles are propagated according to their dynamics and their weights reduced by the survival probability $p_{S_{ext}}$, analog to (3).
 - 5: In the target extraction step, an external target corresponding to an external measurement j is only reported if $W_{t,p}^j \geq c_R$.
 - 6: Sampling of external particles is carried out over each target estimate $\hat{\mathbf{x}}_{t,j}$ and its covariance $\mathbf{P}_{t,j}$ according to (8).
-

IV. CASCADING UKF AND C-SMC-PHD FILTERS

For the particular context of multiple target tracking on a moving vehicle, we propose a cascading of an UKF filter dedicated to the vehicle state estimation with a C-SMC-PHD filter as an approach to solve the full problem of localization of the tracking vehicle and the tracking of multiple objects. The UKF introduces a deterministic method for choosing a set of weighted *sigma points* approximating the distribution with a priori mean $\bar{\mathbf{x}}$ and covariance $\bar{\mathbf{P}}_{xx}$ which are subjected to a nonlinear transform. Thereafter, the a posteriori mean $\bar{\mathbf{y}}$ and covariance $\bar{\mathbf{P}}_{yy}$ of the distribution can be obtained by the weighted average and weighted outer product of the transformed points respectively.

The state and measurement are both defined by the position (x, y) , orientation $(\varphi)\mathbf{x} = \mathbf{z} = [x, y, \varphi, V]$, emulating GNSS data with an onboard compass with the following state

¹It should be noted that the number of particles chosen in the resampling step has to be identical in both agents when carrying out the fusion step. Otherwise, this might result in a problem due to unequal weights.

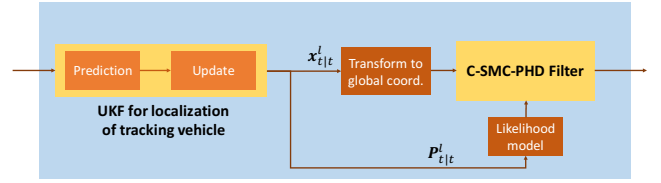


Fig. 3. An illustration of the cascading of an UKF and a C-SMC-PHD filter at time t .

transition function

$$x_t = x_{t-1} + V \cdot \sin(\varphi_{t-1}) \quad (20)$$

$$y_t = y_{t-1} + V \cdot \cos(\varphi_{t-1}) \quad (21)$$

$$\varphi_t = \varphi_{t-1} \quad (22)$$

$$V_t = V_{t-1}. \quad (23)$$

The a posteriori state estimate $\bar{\mathbf{x}}_{t|t}^l$ along with its error covariance $\bar{\mathbf{P}}_{t|t}^l$ (where superscript l denotes variables used in the localization filter) of the UKF is then used as input in the C-SMC-PHD filter in order to transform tracked targets from the local coordinate frame of the tracking vehicle into the global coordinate frame. The representation of tracked targets in global coordinate frame is necessary as cooperating vehicles might not possess relative positional information.

As the objects tracked by the C-SMC-PHD filter are assumed to be vehicles in this context, they are modeled in the same way as the state of the tracking vehicle. Hence, their single state transition model is also defined by (20) – (23). As such, we also assume that the measurement obtained contains the positional information as well as orientation of tracked targets by preprocessing data from the laser range sensor, as described in [5]. The likelihood model is given as a multivariate Gaussian distribution:

$$g_t(\mathbf{z}_t | \mathbf{x}_{t|t-1}) = \frac{1}{\sqrt{(2\pi)^3 |\Sigma|}} \exp(-0.5(\mathbf{z}_t - \mathbf{x}_{t|t-1})^T \Sigma^{-1} (\mathbf{z}_t - \mathbf{x}_{t|t-1})) \quad (24)$$

where Σ represents the measurement uncertainty.

Bearing in mind that the preprocessing of data from LIDAR sensors to extract the pose of tracked vehicles generally result in inaccuracies, especially at larger distances, the model provided in (24) has to take into account both the uncertainty in localization of the tracking vehicle as well as uncertainty of the tracked vehicles. We hence reconsider (24) such that

$$\Sigma = \bar{\mathbf{P}}_{t|t}^l + \text{diag}(\sigma_x^2, \sigma_y^2, \sigma_\varphi^2, \sigma_V^2), \quad (25)$$

where $\sigma_{(\cdot)}, (\cdot) = \{x, y, \varphi, V\}$ represent the standard deviation for x-y coordinates, the vehicle orientation in global coordinates and speed, respectively.

An illustration of this procedure is reported in Fig. 3.

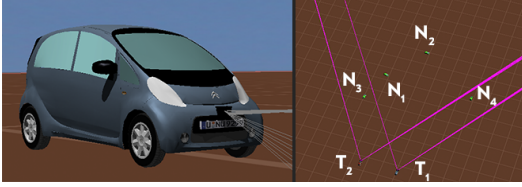


Fig. 4. The Citroën C-ZERO (left) car fitted with a Ibeo LUX LIDAR sensor simulated in the Webots environment (right).

V. EXPERIMENTAL RESULTS

Experiments were carried out in Webots [23], a high-fidelity robot simulator, with the RO2IVSim extension² which allows for simulation of real world vehicles with state-of-the-art sensors found on modern cars. Simulated Citroën C-ZERO cars were equipped with front-facing Ibeo LUX LIDAR sensors, a GNSS device, a wireless communication device and a compass. Sensor properties were calibrated from their respective data sheet. The simulation environment, along with the simulated car is provided in Fig. 4.

A. Scenario

A scenario involving two moving cars (T_1, T_2), each running an instance of the state estimation – C-SMC-PHD filter tracking up to four moving cars (N_1, N_2, N_3, N_4) in an open space is presented. Both T_1 and T_2 travel at constant velocity of 10 km/h. The LIDAR sensors in this scenario have an effective range of 160 meters, a horizontal field of view of 90 degrees, and operate at a scanning frequency of 25 Hz. The trajectories of the cars are shown in Fig. 5 with an initial speed of 20 km/h for $N_i, i = \{1, 2, 3, 4\}$ with varying speed following the given function

$$v(t) = \begin{cases} 20 & 0 \leq t \leq 4 \\ 20 - 4 \cdot (t - 4) & 4 < t \leq 8 \\ 4 + 4 \cdot (t - 8) & 8 < t \leq 18 \end{cases} \quad (26)$$

where $v(t)$ is measured in km/h and t in seconds. This scenario aims to model a typical slow-speed traffic scenario as close as possible while showing the general case where two vehicles running tracking algorithms do not perceive the exact same scene: in the beginning only cars N_1 and N_2 are found in the overlapped zone; towards the middle of the scenario, they leave the union of the two FOVs, about the same time the cars N_3 and N_4 enter the common FOV.

B. Parameters

The UKF used for localization of the tracking vehicles was initialized with the following parameters intrinsic to typical Kalman filter with the usual notations used here.

$$\mathbf{P}_0 = \mathbf{R} = \text{diag}(4.0^2, 4.0^2, 0.01^2, 0.01^2) \quad (27)$$

$$\mathbf{Q} = \text{diag}(0.005, 0.005, 0.01, 0.01). \quad (28)$$

\mathbf{P}_0 represents the initial estimate covariance while the mean of the initial distribution was chosen as the first measurement

²See <http://disal.epfl.ch/research/RO2IVSim> for more details.

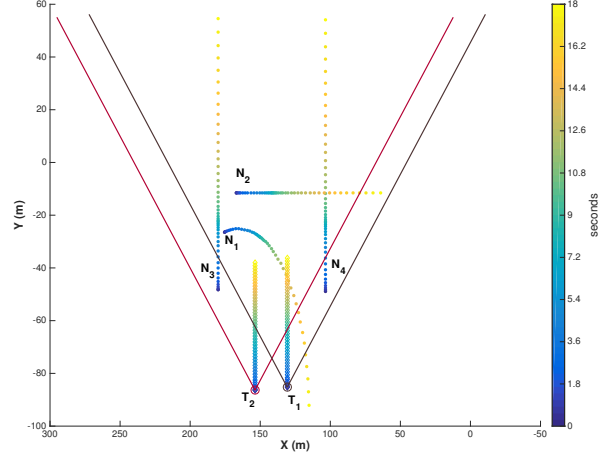


Fig. 5. The vehicle trajectories of the experimental scenario. The grey and red lines denote the field of views of T_1 and T_2 respectively.

of the simulation for each tracking vehicle while \mathbf{Q} represents the covariance of the process noise. The first two entries of the observation model covariance matrix \mathbf{R} correspond to the variance in GPS measurement while the orientation of the vehicle was assumed to be determined by a compass with insignificant inaccuracy. An insignificant inaccuracy in determining the speed of the vehicle was also assumed in this case. Furthermore, the choice of sigma points was taken from [24] as the state of the art for research and industry.

For the C-SMC-PHD filter, 100 particles per persisting measurement were used while $\rho = 40$ particles were birthed around each new measurement at every iteration. The “spawning” of newborn particles at time t around a measurement follows a Gaussian distribution with the standard deviation of the position corresponding to the first two diagonal entries of \mathbf{P}_t . A Poisson distribution with $\lambda_{clutter} = 1$ was used to simulate the occurrence of clutter resulting from multiple measurements from the LIDAR sensor at each iteration. The *cooperation* step of the C-SMC-PHD was carried out every third *sensing* step to realistically model real world conditions where network communications often occur at a lower rate than sensor readouts.

Additionally, we chose the second term of (25) to be $\text{diag}(1.3^2, 1.3^2, 0.1^2, 0.1^2)$. The standard deviation for the estimated position of tracked vehicles were experimentally obtained from a scenario where the tracking vehicles are stationary.

C. Results

We present the results of the scenario with the given parameters along with a discussion of certain noteworthy observations. The dataset collected in this experiment contains 30 distinct simulation runs where each vehicle starts at the same positions and runs an open-loop controller. We note that the fusion of data for the cooperative component of the C-SMC-PHD filter is done only on the vehicle T_1

unless otherwise mentioned. The following results are also discussed with respect to this vehicle.

The Optimal Sub-Pattern Assignment (OSPA) metric [25] was used for multi-object performance evaluation. It comprises of two components – one considering localization errors while the other considers the mismatch in cardinality. For our case, an important consideration here is the relative performance of the C-SMC-PHD filter to the non-cooperative variant. As no significant improvements in cardinality is expected through cooperation, we only focus on the comparison of localization errors. For computing the OSPA error, the sensitivity of the metric in penalizing estimated position $p = 1$ was chosen.

A comparison of the cooperative and non-cooperative filter’s OSPA localization errors averaged over 30 runs is provided in Fig. 6 and Fig. 7. The right diagram of Fig. 7 furthermore provides a comparison of the average difference in OSPA error between the cooperative and non-cooperative variant of the multiple vehicle tracking filter. To ensure a fair comparison, only targets in the FOV of the vehicle T_1 are considered when computing the OSPA error, since no targets outside of this area can be observed when the non-cooperative filter is used.

No statistically significant difference is observed in this comparison. We believe the absence of significant improvements could be attributed to the localization uncertainty of tracking vehicles, which correspondingly results in an uncertain transformation of the tracked vehicles from vehicle local frame to global coordinates. Furthermore, the fact that the two vehicles observe targets from a similar position could also contribute to this result. However, it is important to note that cooperation with a vehicle that tracks poorly due to localization error or other factors does not degrade the tracking vehicle’s tracking accuracy.

A comparison of the cooperative and non-cooperative filter’s tracking errors per tracked vehicle is shown in Table I, as observed from both tracking vehicles. Values here are averaged over all state estimates of a simulation run and over all simulation runs. Once again, only targets in the internal FOV of the corresponding tracking vehicle are considered. The C-SMC-PHD filter showed on average a slightly better accuracy than its non-cooperative counterpart.

It is furthermore worth noting that the significant tracking errors even at short distances (for vehicles N_3 and N_4 especially) could be partially attributed to the preprocessing of raw LIDAR data and the limited angular resolution of the LIDAR sensor, since a point cloud relating to a vehicle can vary significantly with vehicle’s pose in the sensing FOV.

An important aspect of our approach is also to provide a reliable estimate of external tracked vehicles (vehicles not in the tracking vehicle’s FOV), of which we have no additionally information except from that of cooperating vehicles. In the case of our scenario, T_1 and T_2 track N_3 and N_4 exclusively for a certain amount of time and provide this information to each other as described in Section III-E.2. The tracking accuracy of external vehicles for each $T_{(i)}$ is shown in Table II. Although it is observed that

TABLE I

A COMPARISON OF TRACKING ACCURACY OF COMMON VEHICLES BETWEEN COOPERATIVE AND NON-COOPERATIVE TRACKING.

Tracked vehicles	Tracking accuracy as perceived from vehicle (in meters):			
	T_1		T_2	
	Coop	No-coop	Coop	No-coop
N_1	2.001	2.133	1.854	1.842
N_2	1.920	2.060	1.776	1.832
N_3	0.994	1.042	1.467	1.499
N_4	1.377	1.392	1.048	1.116

TABLE II

A COMPARISON OF TRACKING ACCURACY OF EXTERNAL VEHICLES BETWEEN DIRECT AND COOPERATIVE TRACKING. THE ENTRIES IN BOLD CORRESPOND TO THE EXTERNAL VEHICLE OF T_i .

Tracked vehicles	Tracking accuracy as perceived from vehicle:	
	T_1	T_2
N_3	2.374	1.499
N_4	1.392	3.123

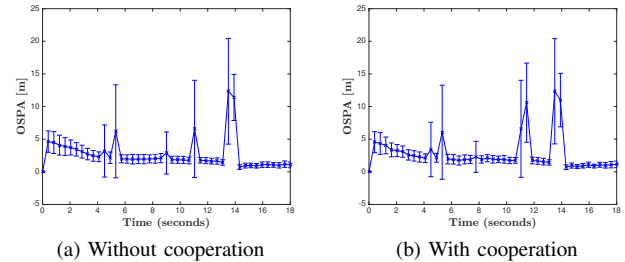


Fig. 6. Comparison of OSPA localization errors between non-cooperative tracking (left column) and the C-SMC-PHD filter (right column) as viewed from T_1 . The corresponding standard deviation is also plotted as error bars.

external vehicles are tracked worse compared to the tracking done with corresponding measurements, this is only to be expected owing to the reduced frequency of cooperation (once every three sensing steps, as noted in Section V-B) and localization uncertainty of the two tracking vehicles. When viewed in relation to the range and FOV of LIDAR sensors, the disparity observed is not very significant considering that information on external vehicles could be easily more than tens of meters away from the tracking vehicle itself.

VI. CONCLUSION AND FUTURE WORK

We presented an integrated approach for reliably tracking multiple mobile targets from a moving vehicle incorporating information from a neighboring vehicle partially perceiving the same scene. Our method takes into account vehicles’ localization inaccuracies and sensor noise. The localization of the tracking vehicle was based on an UKF while the tracking of multiple vehicles was accomplished with our proposed C-SMC-PHD filter. A simulation scenario aimed at closely emulating real-world low-speed driving conditions but including variable speeds and nonlinear trajectories was

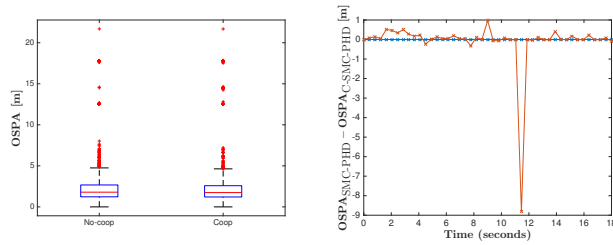


Fig. 7. Comparison of the statistics (left) of the total OSPA localization errors between non-cooperative tracking and the C-SMC-PHD filter as well as difference in OSPA error (right) between non-cooperative and the C-SMC-PHD filter as observed from T_1 .

used as a proof-of-concept of our approach cascading an UKF with a C-SMC-PHD filter.

While we have shown that our approach does accomplish the goal of provided extended visibility' for tracking vehicles, further work remains to be done in order for significant tracking accuracy improvements in the overlapping FOV to be observed. This work also provides a basis for many more issues to be addressed. For one, the reduction in the number of particles to be communicated to cooperating vehicles could be necessary if such a concept were to be scalable for use in a larger networks. Whiteley et al. [26] for example, propose the use of an auxiliary SMC-PHD filter to reduce the number of particles propagated in time.

Another aspect for future developments could deal with the increasing of the scale and complexity of the scenarios as well as introducing non-vehicle objects in the environment. In cases where one tracking vehicle also tracks another (cooperative) tracking vehicle, positional information from GNSS data and measurements from LIDAR data could be fused to provide a more reliable estimate of vehicles.

Needless to say, the real-world applicability of these proposed methods should be validated with real-world scenarios. We also aim to conduct real-world experiments involving real Citroën C-ZERO cars equipped with the real sensor counterparts currently reproduced only in our high-fidelity simulation environment.

REFERENCES

- [1] T. M. Corporation. Toyota to bring vehicle-infrastructure cooperative systems to new models in 2015. [Online]. Available: <http://newsroom.toyota.co.jp/en/detail/4228471/>
- [2] N. H. T. S. Administration, "Frequency of target crashes for intelligent safety systems," U.S. Department of Transportation, Tech. Rep., 2010.
- [3] R. MacLachlan and C. Mertz, "Tracking of moving objects from a moving vehicle using a scanning laser rangefinder," in *IEEE Intelligent Transportation Systems Conference*, 2006, pp. 301–306.
- [4] Y. Bar-Shalom, *Tracking and Data Association*. San Diego, CA, USA: Academic Press Professional, Inc., 1987.

- [5] M. Vasic and A. Martinoli, "A collaborative sensor fusion algorithm for multi-object tracking using a Gaussian mixture probability hypothesis density filter," in *IEEE Intelligent Transportation Systems Conference*, 2015, pp. 491–498.
- [6] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, Nov 2006.
- [7] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, Inc., 2007.
- [8] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo implementation of the PHD filter for multi-target tracking," in *Proceedings of the International Conference on Information Fusion*, vol. 2, 2003, pp. 792–799.
- [9] B. Ristic, D. Clark, and B.-N. Vo, "Improved SMC implementation of the PHD filter," in *13th Conference on Information Fusion*, July 2010. doi: 10.1109/ICIF.2010.5711922
- [10] M. Uney, D. Clark, and S. Julier, "Distributed fusion of PHD filters via exponential mixture densities," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 521–531, June 2013.
- [11] G. O. U. G. information about the Global Positioning System (GPS) and related topics. Gps accuracy. [Online]. Available: <http://www.gps.gov/systems/gps/performance/accuracy/>
- [12] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar 2004.
- [13] —, "New extension of the Kalman filter to nonlinear systems," in *AeroSense'97*. International Society for Optics and Photonics, 1997, pp. 182–193.
- [14] J. Andrade-Cetto, T. Vidal-Calleja, and A. Sanfeliu, "Unscented transformation of vehicle states in SLAM," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2005, pp. 323–328.
- [15] A. N. Ndjeng, A. Lambert, D. Gruyer, and S. Glaser, "Experimental comparison of Kalman filters for vehicle localization," in *IEEE Intelligent Vehicles Symposium*, June 2009, pp. 441–446.
- [16] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct 2003.
- [17] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2005, ch. Chapter 4: Nonparametric Filters.
- [18] M. Ester, H. Peter Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," AAAI Press, 1996, pp. 226–231.
- [19] G. T. Toussaint, "A simple linear algorithm for intersecting convex polygons," *The Visual Computer*, vol. 1, no. 2, pp. 118–123, 1985.
- [20] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, April 2011.
- [21] A. E. R. Chris Fraley, "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 611–631, 2002.
- [22] D. W. Scott, "Scott's rule," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 497–502, 2010.
- [23] O. Michel, "Webots: Professional mobile robot simulation," *International Journal of Advanced Robotic Systems*, vol. 1, pp. 39–42, 2004.
- [24] R. Van Der Merwe, "Sigma-point kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, Oregon Graduate Institute School of Science and Engineering, Beaverton, Oregon, 2004.
- [25] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, Aug 2008.
- [26] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary particle implementation of probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1437–1454, July 2010.