

Shared Situational Awareness with V2X Communication and Set-membership Estimation

Vandana Narri^{1,2}, Amr Alanwar³, Jonas Mårtensson², Christoffer Norén¹ and Karl Henrik Johansson²

Abstract—The ability to perceive and comprehend a traffic situation and to estimate the state of the vehicles and road-users in the surrounding of the ego-vehicle is known as situational awareness. Situational awareness for a heavy-duty autonomous vehicle is a critical part of the automation platform and depends on the ego-vehicle's field-of-view. But when it comes to the urban scenarios, the field-of-view of the ego-vehicle is likely to be affected by occlusions and blind spots caused by infrastructure, moving vehicles, and parked vehicles. This paper proposes a framework to improve situational awareness using set-membership estimation and Vehicle-to-Everything (V2X) communication. This framework provides safety guarantees and can adapt to dynamically changing scenarios, and is integrated into an existing complex autonomous platform. A detailed description of the framework implementation and real-time results are illustrated in this paper.

I. INTRODUCTION

When a Connected Automated Vehicle (CAV) is located in an urban scenario, the field-of-view (FOV) of the CAV is severely affected by occlusions which leads to limited information about the surroundings and affects situational awareness of the CAV. To be able to overcome the challenge of limited information in an urban setting, the plan is to use Vehicle-to-Everything (V2X) communication to share perception data between road-users.

V2X communication is synonymous with Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. A unit or a road user that is capable of transmitting and receiving information via V2X communication is referred to as a V2X unit. These communications use the ad-hoc wireless technology ITS-G5 (a.k.a. Dedicated Short-Range wireless Communication (DSRC) or IEEE 802.11p in the US). There are three different types of V2X communication today: ad-hoc (the focus of the present article), cloud-based, i.e., cloud-to-cloud communication, and cellular-assisted. The main strength of ad-hoc communication is the possibility to communicate to any V2X unit at any time without base station coverage, e.g., on rural roads and missing network coverage areas. In this paper, we propose a framework that can provide safety guarantees and increase

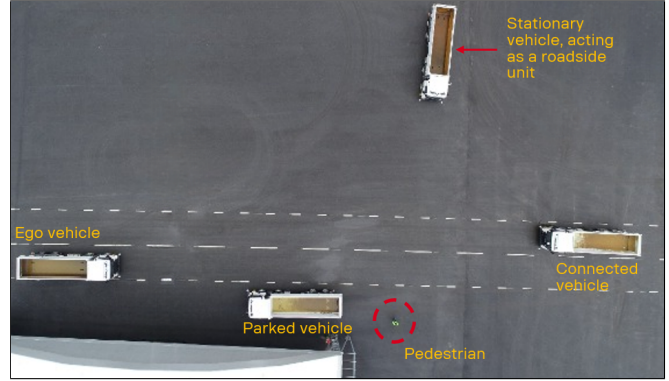


Fig. 1: This is a drone shot of the scenario with four vehicles and a pedestrian. The pedestrian is in the occluded area of the ego-vehicle. This picture is taken at the Scania test track.

situational awareness using set-membership estimation and ad-hoc V2X communication.

A. Literature Review

Situational awareness of the Ego-Vehicle (EV) depends on the perception module, which primarily relies on numerous onboard sensors, such as LiDAR, cameras, and millimeter-wave radar, but still lacks accuracy and complete information [1]. One way to overcome this problem is by using V2X communication. In [2], the authors provided a detailed review of cooperative driving by introducing current cooperative perception information fusion methods and information-sharing strategies. In [3], authors have provided a generalized cooperative perception framework based on decentralized V2V vehicular communication and presented evaluations of the framework based on randomized traffic simulations for multi-lane highway and roundabout intersection scenarios, considering both communication losses and sensor FOV resolution issues. Another related work in [4] introduces information entropy, which quantifies uncertain information in the blind area, into the motion planning module of autonomous vehicles. Then, the authors propose to plan collision-free trajectories using model predictive control.

In [5], a framework of vehicle-infrastructure cooperative perception and data fusion method is proposed along with an improved Kalman filter when there is a roadside unit failure. In [6], authors presented an architecture for coordinating an autonomous team of heterogeneous aerial and land robots that work together on collaborative mapping. Based on the presented concept, they proposed an IoE (Internet-of-Everything) architecture having heterogeneous support units for enhancing the situational awareness of autonomous vehicles in an unknown environment. A set-

¹The authors are with Research and Development, Scania CV AB, 151 87 Södertälje, Sweden. {vandana.narri, christoffer.noren}@scania.com.

²The authors are with Division of Decision and Control Systems at School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE100 44 Stockholm, Sweden. {narri, jonas1, kallej}@kth.se.

³The author is with the School of Computation, Information and Technology, Technical University of Munich, Heilbronn, Germany and the School of Computer Science and Engineering, Constructor University, Bremen, Germany. alanwar@tum.de.

membership state estimator for autonomous surface vehicles along with dynamical decomposition that decouples the estimation problem for the rotational and positional dynamics is proposed in [7]. In [8], authors have proposed a novel data fusion algorithm by implementing the latest deep learning techniques and aggregating object (road-user) detection information from multiple viewpoints to improve detection performance. The objective of [9] is to provide a preliminary study of edge-assisted collaborative perception in autonomous driving from a V2X communication design perspective.

B. Contributions

Situational awareness is the ability to perceive and comprehend a traffic situation and estimate the state of road-users in the surrounding of the EV. The traditional way of the situational awareness module is to rely entirely on onboard sensor perception, also known as local perception. This work aims to use V2X communication to receive external perception, i.e., perception from other CAV or Road-Side Units (RSUs), to create shared situational awareness, which helps improve the understanding of one's surroundings and guarantee safety. Inclusion of perception obtained via V2X communication, helps in widening EV's FOV and creating line-of-sight in occluded areas. But external perception can induce varying unknown uncertainties.

One of the challenging parts is to integrate the new functionality, i.e., shared situational awareness that relies on both local and external perception, into the existing CAV system architecture of the EV. This paper focuses on the fusion of local and external perceptions, which have varying uncertainties and involves different coordinate frames, to create a shared situational awareness for the EV using set-membership estimation. Set-membership estimation method estimates a set that contains the state with set containment guarantees. In this paper, we have considered a scenario in which the pedestrian is occluded by the parked vehicle in the EV's FOV, as presented in Fig. 1. This is an extended work, along with real implementation on the Scania test track, from our paper that highlighted the concept of shared situational awareness using V2X communication [10].

C. Outline

The paper is organized as follows: In Section II, the problem formulation is described. The necessary preliminaries are provided in Section III. The proposed framework is described in Section IV. Then, the evaluation section is provided in Section V, followed by Section VI to conclude the paper.

II. PROBLEM FORMULATION

In this paper, the problem of occlusions and blind spots caused by infrastructure, moving vehicles, and parked vehicles to the EV is considered. A specific scenario is considered for the experiment and is shown in Fig. 1. All the vehicles present in Fig. 1 are equipped with a V2X communication module, which allows us to share perception information between all the V2X units in the scenario except the parked vehicle and the pedestrian. It is also considered to be

distributed system, where the communication is set up as a peer-to-peer network and does not rely upon communication infrastructure such as access points or base stations. The objective is to improve situational awareness of the EV. The following research questions are formulated as follows to solve the problem in this paper:

- 1) How to obtain and fuse data from local and external sensors to improve the situational awareness of the EV?
- 2) How to provide safety guarantees of the fused data under varying measurement uncertainties?

III. PRELIMINARY

This section gives a detailed description of the system and measurement models for each road-user. Then, we introduce zonotopes, which are used to represent the reachable set mathematically and followed by the system architecture of a CAV.

A. System and Measurement Model

This paper considers a discrete-time linear system model for the road-user observed in the given scenario. The system state vector consists of the position and velocity of the road-user as described in (1) and (2). All the road-users' system state vectors are calculated with respect to the EV coordinate system. For $k = 1, 2, \dots, T$, where T is the time horizon, the system model can be described as follows.

$$x_{k+1}^j = F_k^j x_k^j + q_k^j, \quad (1)$$

with

$$x_k^j = [x_x \quad x_y \quad x_s]^T, \quad (2)$$

where $x_k^j \in \mathbb{R}^n$ is the system state vector of the j^{th} road-user with x_0^j as the initial state and $j = 1, \dots, n_r$, n_r is the number of road-users, x_x , x_y , and x_s are the x coordinate, y coordinate, the velocity of the j^{th} road-user respectively, q_k^j is the process noise, and $F_k^j \in \mathbb{R}^{n \times n}$ is the transition matrix at time k which is defined as follows:

$$F_k^j = \begin{bmatrix} 1 & 0 & \alpha_k^j \\ 0 & 1 & \gamma_k^j \\ 0 & 0 & 1 \end{bmatrix}$$

where $\alpha_k^j = \Delta t \cos \theta_k^j$, and $\gamma_k^j = \Delta t \sin \theta_k^j$. Here θ_k^j is the heading of the road-user at time k and Δt is the time step.

The observable discrete-time linear system is assumed to generate the sensor measurements. The measurement model of the sensor is described as the following:

$$y_k^j = H^j x_k^j + v_k^j, \quad (3)$$

with

$$y_k^j = [y_x \quad y_y \quad y_s]^T, \quad (4)$$

where $y_k^j \in \mathbb{R}^p$ is the measurement state vector, which consists of y_x , y_y and y_s i.e. x coordinate, y coordinate and velocity of the j^{th} road-user at time step k with respect to

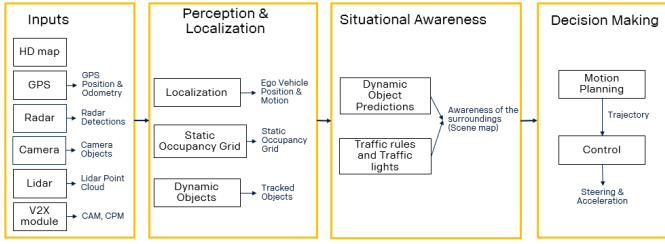


Fig. 2: Existing system architecture of a CAV.

the EV coordinate system and v_k^j the measurement noise. The measurement matrix is given by $H^j \in \mathbb{R}^{p \times n}$ and in this case it is an identity matrix. Since the processing of all the measurements will be done with respect to the local coordinate system of the EV, all the measurements obtained externally are converted to the local coordinate frame of the EV, which are denoted by x_E and y_E , where E subscript refers to the EV.

B. Zonotopes

There are various types of set representations, such as ellipsoids [11], polytopes [12], zonotopes [13], orthotopes [14] and intervals [15]. In this paper, zonotopes are considered to mathematically represent the set due to its special computational efficient properties. The definition and properties of zonotopes are stated below.

Definition 1 (Zonotope [16]): A zonotope $\mathcal{Z} = \langle c, G \rangle$ consists of a center $c \in \mathbb{R}^n$ and a generator matrix $G \in \mathbb{R}^{n \times e}$. The generator matrix G is composed of e generators $g_i \in \mathbb{R}^n$, $i = 1, \dots, e$, where $G = [g_1, \dots, g_e]$ and β_i is a scaling factor. Hence,

$$\mathcal{Z} = \left\{ c + \sum_{i=1}^e \beta_i g_i \mid -1 \leq \beta_i \leq 1 \right\}. \quad (5)$$

Given two zonotopes $\mathcal{Z}_1 = \langle c_1, G_1 \rangle$ and $\mathcal{Z}_2 = \langle c_2, G_2 \rangle$ and a scalar L , the following operations can be computed exactly [16]:

- Minkowski sum:

$$\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \langle c_1 + c_2, [G_1, G_2] \rangle. \quad (6)$$

- Scaling:

$$L\mathcal{Z}_1 = \langle Lc_1, LG_1 \rangle. \quad (7)$$

C. CAV System Architecture

A CAV has a complex system architecture consisting of multiple sub-systems handling various tasks that enable the CAV to drive with minimum human interaction. The term connected states that the vehicle is capable of communicating with roadside units or other CAVs. An overview of an existing system architecture of a CAV is presented in Fig. 2. The system architecture is divided into four sub-systems: input, perception and localization, situational awareness, and decision-making.

The input layer consists High-Definition (HD) map, GPS, lidars, radars, cameras, and a V2X module that enables the communication between other agents. This input layer is essential and supports the functionality between different

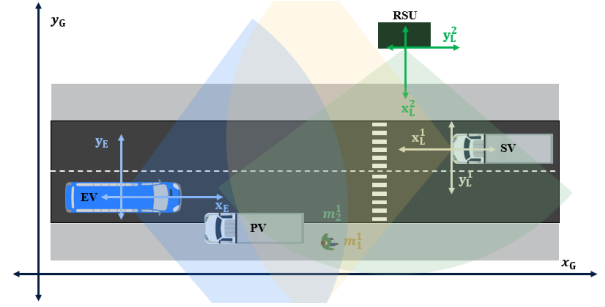


Fig. 3: Scenario with four road-users and their coordinate frames.

applications in the system. The perception module is responsible for detecting, identifying, classifying, and tracking different road-users in the near surrounding of the vehicle. This module is divided into two parts: static occupancy grid and dynamic objects. The static occupancy grid provides the FOV of the EV, and the dynamic objects provide the perception data, i.e., a list of objects detected in the FOV of the EV. The vehicle localization module is responsible for estimating the state of the vehicle (e.g., position, speed, orientation, and acceleration), both with respect to its surroundings (locally) and with respect to a map (globally). This module is essential for safe, comfortable, and precise vehicle motion.

The objective of the situation awareness module is to determine the predictions of the surrounding road-user's actions and provide scene understanding. The motion planning module's responsibility is to compute a trajectory while respecting vehicle dynamical constraints, being safe and smooth, and making the motion comfortable for passengers. The motion control module is responsible for stabilizing and guiding the vehicle toward a given reference path or trajectory. The vehicle controller has to handle disturbances to the vehicle, correcting its states (e.g., position, yaw, and velocity) back to the desired reference while maintaining the actuation limits and achieving a smooth and comfortable ride.

IV. METHODOLOGY

In this section, the architecture of the proposed framework and each module of the framework is discussed. The proposed framework extends the conventional control architecture of an automated vehicle [17] by using the set-membership method instead of statistical estimation for sensor fusion in order to have safety guarantees.

A. Scenario Description

The animated version of the considered scenario is presented in Fig. 3. This scenario consists of five road-users, which are an EV, a Side Vehicle (SV), a Parked Vehicle (PV), a pedestrian, and a Road-Side Unit (RSU). The FOVs for EV, SV, and RSU are presented in blue, yellow, and green circle segments, respectively. The PV is creating an occlusion in the FOV of the EV, and the pedestrian, also called the unprotected agent, is in the occluded region. In this case, the understanding of the surrounding environment is limited to the road elements that are observable within the FOV of the EV. Behavior learning and prediction are performed on

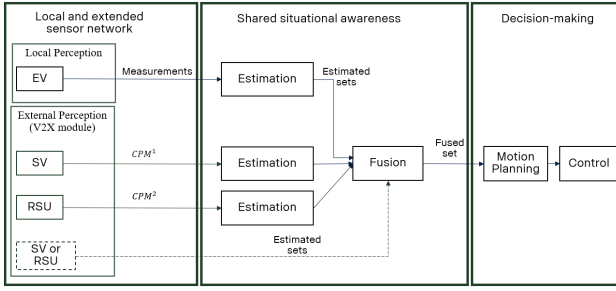


Fig. 4: Shared situational awareness internal architecture.

the detected road-users, which are time-sensitive. Hence, the risk assessment tasks for a short decision-making horizon can only be made through a proactive manner by lowering speed to be able to handle any road-user that might appear in the occluded area. The perception information is shared using Collaborative Perception Message (CPM) service. A detailed description of the CPM service is given in Section IV-C.

When it comes to real-time implementation, it can be observed that different coordinate frames are involved, as shown in Fig. 3. A global coordinate system is the earth-fixed coordinate frame, which is represented in meters in north and east directions from an agreed reference point, which is represented by x_G and y_G , respectively. Each agent has a local coordinate frame with respect to the reference point. These reference points are their global positions in the global coordinate frame. The EV has its local coordinate frame with axes x_E and y_E . The SV has its local coordinate frame with axes x_L^1 and y_L^1 . Similarly, for RSU, x_L^2 and y_L^2 are its local coordinate frame axes. The PV in the EV's FOV is detected as a static obstacle, creating an occlusion in EV's FOV as shown in Fig. 3. The pedestrian is in the occluded region of the EV's FOV. Therefore the EV can not detect the pedestrian. But whereas, SV and RSU can detect it in their FOVs.

B. Proposed Architecture

The proposed architecture consists of (i) Local and extended sensor networks, (ii) Algorithms for shared situational awareness, and (iii) Decision-making, as presented in Fig. 4. In the local and extended sensor network part, the blocks EV, SV, and RSU represent the road-users capable of providing perception data with respect to their local coordinate frame. We fuse the data on two levels: (i) measurement data and (ii) estimated sets. This two-level data fusion method helps us build a modular system, which means that if there is another SV or RSU in the scenario capable of sharing estimated sets, then these estimated sets can be fused directly with other estimated sets in the fusion module. Based on these fused sets, decisions can be made, and actions can be planned. In this paper, the main focus is on the sensor network and shared situational awareness.

C. Local and External Perception Data

The external perception data is obtained via the CPM service, which is implemented by using standards published by the European Telecommunications Standards Institute (ETSI) [18]. The general message structure for a CPM is

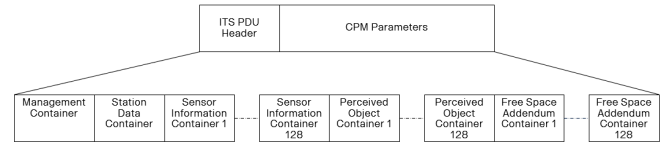


Fig. 5: General structure of CPM.

as specified in the standards, which is shown in Fig. 5. In this work, we use Management Container (MC), Station Data Container (SDC), and Perceived Object Container (POC). The MC provides basic information about the originating ITS-S (Intelligent Transport System-Station), which includes the station type and reference position in the global coordinates. The SDC provides more specific information about the originating ITS-S in addition to the common information provided by the MC. The data received from the perception module of SV or RSU is populated into the POC of the CPM. The POC enables a detailed description of the dynamic state and properties of a detected object.

The perception data from the SV or RSU in the scenario has different coordinate frames depending on their position and orientation with respect to global coordinates. For instance, when a road-user is detected in the FOV of SV, the perception data is denoted by $m_{i,k}^j$ as shown in Fig. 3, denoting that it is j^{th} road-user detected by i^{th} V2X unit. This data is obtained from the respective perception layer of the V2X unit, and information obtained from $m_{i,k}^j \in \mathbb{R}^{1 \times b}$ as shown below:

$$m_{i,k}^j = [m_x \quad m_y \quad m_s \quad \theta \quad l \quad w]^j_{i,k}, \quad (8)$$

where m_x and m_y are the relative distance from i^{th} V2X unit at a given time stamp indicated by k and the location of i^{th} V2X unit is given in terms of global coordinates. The m_s and θ are the velocity and heading of the detected j^{th} road-user. The size of the of road-user is indicated by length and width, i.e., l and w . This perception data, $m_{i,k}^j$ is converted to y_k^j with respect to the EV coordinate system with the help of the global coordinate of originating V2X unit and the global coordination of the EV.

In this work, it is considered to have all the local and external perception data with respect to the local coordinate frame of the EV to maintain uniformity. The local perception also has the same perception data structure as explained in (8). The local perception data is already given with respect to the EV's local coordinate frame, which means that it does not require conversion of the coordinate system and can be realized in the format of y_k^j . It is considered that the EV reference point is at the center of the vehicle, and the reference points of SV and RSU are at the front of the vehicles, which creates an offset between different perception measurements as shown in Fig. 3. This offset has to be handled systematically to be able to fuse local and external perceptions.

D. Algorithm for Shared Situational Awareness

Our shared situational awareness algorithm consists of two steps: (1) Set-based estimation and (2) Fusion. These steps are explained below and summarized in Algorithms 1 and 2.

1) Set-Based Estimation

The set-based estimation approach computes a set that contains the true state with guarantees. The prediction set is computed with the aid of the system model expressed in (1). The process and measurement noise for j^{th} road-user are assumed to be unknown but bounded by zonotopes: $q_k^j \in \mathcal{Z}_{Q,k}^j = \langle 0, Q_k \rangle$, and $v_k \in \mathcal{Z}_{R,k}^j = \langle 0, \text{diag}(r_k^j) \rangle$. We intersect the aforementioned predicted set with the set that aligns with the measurement set. In short, we have the following three sets and the aim is to compute the corrected state set $\hat{\mathcal{Z}}_k^j \subset \mathbb{R}^n$.

Definition 2 (Predicted State Set): Given the system in (1) and (3), the initial set $\mathcal{Z}_0^j = \langle c_0^j, G_0^j \rangle$, and the process noise zonotope $\mathcal{Z}_{Q,k}^j$, the predicted reachable set of states $\hat{\mathcal{Z}}_k^j \subset \mathbb{R}^n$ is:

$$\hat{\mathcal{Z}}_k^j = F_k^j \hat{\mathcal{Z}}_{k-1}^j \oplus \mathcal{Z}_{Q,k}^j. \quad (9)$$

Definition 3 (Measurement State Set): Given the system in (1) and (3), and the measurement noise zonotope $\mathcal{Z}_{R,k}^j = \langle 0, \text{diag}(r_k^j) \rangle$ where the measurement state set $\mathcal{S}_k^j \subset \mathbb{R}^p$ of the j^{th} road-user, is the set of all possible solutions x_k which can be reached given the measurement y_k^j . Note that when $y_k^j \in \mathbb{R}^p$ is a scalar, i.e., $p = 1$, this measurement set is a strip:

$$\mathcal{S}_k^j = \{x_k \mid |H^j x_k - y_k^j| \leq r_k^j\}. \quad (10)$$

Definition 4 (Corrected State Set): Given the system (1) and (3) with initial set $\mathcal{Z}_0^j = \langle c_0^j, G_0^j \rangle$, the reachable corrected state set $\hat{\mathcal{Z}}_k^j$ for the j^{th} road-user is defined as:

$$(\hat{\mathcal{Z}}_k^j \cap \mathcal{S}_k^j) \subseteq \hat{\mathcal{Z}}_k^j. \quad (11)$$

Set-based approaches intersect the set of states consistent with the model (predicted state set), denoted by $\hat{\mathcal{Z}}_{k-1}^j$, and the sets consistent with the measurements (measurement state set), denoted by \mathcal{S}_k^i , $i = 1, \dots, n_m$, to obtain the corrected state set, denoted by $\hat{\mathcal{Z}}_k^j$, which is an over-approximation of the resultant intersection. This corrected set is also known as estimated set. The strips represent measurements from the sensors and are fused together to get a corrected state set. We use the following proposition to fuse the measurements from the sensors and extended sensors.

Proposition 1 ([19]): Given are zonotope $\hat{\mathcal{Z}}_{k-1}^j = \langle \hat{c}_{k-1}^j, \hat{G}_{k-1}^j \rangle$, the family of n_m measurement sets \mathcal{S}_k^i in (10) and the design parameters $\lambda_k^i \in \mathbb{R}^{n \times p}$, $i = 1, \dots, n_m$. The intersection between the zonotope and measurement sets can be over-approximated by the zonotope $\hat{\mathcal{Z}}_k^j = \langle \hat{c}_k^j, \hat{G}_k^j \rangle$, where

$$\hat{c}_k^j = \hat{c}_{k-1}^j + \sum_{i=1}^{n_m} \lambda_k^i (y_k^i - H^i \hat{c}_{k-1}^j), \quad (12)$$

$$\hat{G}_k^j = \left[(I - \sum_{i=1}^{n_m} \lambda_k^i H^i) \hat{G}_{k-1}^j, \lambda_k^1 r_k^1, \dots, \lambda_k^{n_m} r_k^{n_m} \right]. \quad (13)$$

After fusing the measurements from the sensors, we fuse the estimated sets from multiple vehicles and infrastructure units. Set-membership estimation is performed on each measurement separately, and the algorithm is as shown in 1.

Algorithm 1: Set-based estimation for each road-user

Input: Initial set $\mathcal{Z}_0 = \langle c_0, G_0 \rangle$, transition matrix F_k^j , measurement matrix is H^j process noise zonotope $\mathcal{Z}_{Q,k} = \langle 0, Q_k \rangle$, measurement noise zonotope $\mathcal{Z}_{R,k} = \langle 0, \text{diag}(r_k^j) \rangle$, measurement y_k^j

Output: Corrected state set $\hat{\mathcal{Z}}_k^j = \langle \hat{c}_k^j, \hat{G}_k^j \rangle$

- 1 Compute the predicted set $\hat{\mathcal{Z}}_k^j = F_k^j \hat{\mathcal{Z}}_{k-1}^j \oplus \mathcal{Z}_{Q,k}^j$.
 - 2 Collect measurement set: $\mathcal{S}_k^j = \{H^j, y_k^j, r_k^j\}$.
 - 3 Compute the corrected set: $\hat{\mathcal{Z}}_k^j \supseteq (\hat{\mathcal{Z}}_k^j \cap \mathcal{S}_k^j)$.
-

2) Fusion

We propose to use a fusion function in order to fuse our internal estimated sets with the received estimated sets. More specifically, consider when a road-side unit or connected vehicle is reporting estimated sets of the same pedestrian. Then, we make use of the following proposition to fuse the estimated sets $\hat{\mathcal{Z}}_k^i$, $i = 1, \dots, n_e$, by finding their intersections.

Proposition 2 ([13]): The intersection between n_e zonotopes $\hat{\mathcal{Z}}_k^i = \langle \hat{c}_k^i, \hat{G}_k^i \rangle$ can be over-approximated using the zonotope $\hat{\mathcal{Z}}_k^j = \langle \hat{c}_k^j, \hat{G}_k^j \rangle \subset \mathbb{R}^n$ as follows:

$$\hat{c}_k^j = \frac{1}{\sum_{i=1}^{n_e} w_k^i} \sum_{i=1}^{n_e} w_k^i \hat{c}_k^i, \quad (14)$$

$$\hat{G}_k^j = \frac{1}{\sum_{i=1}^{n_e} w_k^i} [w_k^1 \hat{G}_k^1, \dots, w_k^{n_e} \hat{G}_k^{n_e}], \quad (15)$$

where w_k^i is a weight such that $\sum_{i=1}^{n_e} w_k^i \neq 0$.

The design parameters in Proposition 1 and 2, i.e. λ_k^i and w_k^i can be obtained by solving an optimization problem to minimize the size of the resultant zonotope as stated in [13]. The algorithm for set-based fusion is described in Algorithm 2. Estimation and fusion of the shared situational awareness algorithm are described with illustrations in Fig. 6. In Fig. 6, $\hat{\mathcal{Z}}_k^2$ can be estimated set from another V2X unit near the EV.

Algorithm 2: Set-based fusion for each road-user

Input: List of corrected sets $\hat{\mathcal{Z}}_k^i \forall i = 1, \dots, n_e$

Output: Fused set $\hat{\mathcal{Z}}_k^j = \langle \hat{c}_k^j, \hat{G}_k^j \rangle$

- 1 If a road-user has multiple corrected sets, i.e., $\hat{\mathcal{Z}}_k^i = \langle \hat{c}_k^i, \hat{G}_k^i \rangle, \forall i = 1, 2, \dots, n_e$, then compute:
 - 2 $\hat{c}_k^j = \frac{1}{\sum_{i=1}^{n_e} w_k^i} \sum_{i=1}^{n_e} w_k^i \hat{c}_k^i$;
 - 3 $\hat{G}_k^j = \frac{1}{\sum_{i=1}^{n_e} w_k^i} [w_k^1 \hat{G}_k^1, \dots, w_k^{n_e} \hat{G}_k^{n_e}]$;
 - 4 $\hat{\mathcal{Z}}_k^j = \langle \hat{c}_k^j, \hat{G}_k^j \rangle$;
-

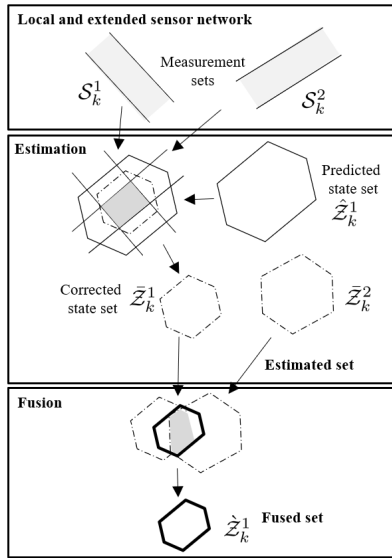


Fig. 6: Set-based estimation and fusion.

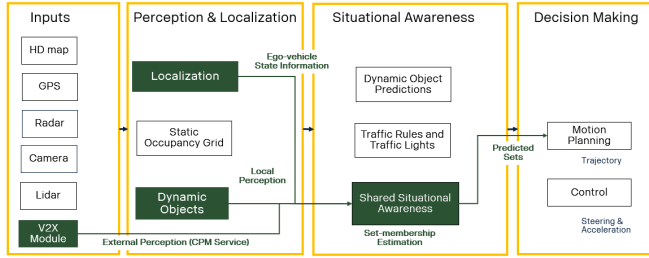


Fig. 7: Shared Situational Awareness integrated to the existing system architecture.

V. EVALUATION

In this section, we start by describing our experimental setup, then followed by presenting results.

A. Experimental Setup

The experiments were carried out at the Scania test track. For these tests, three CAVs and a parked vehicle were used. These three CAVs were equipped with various sensors and were capable of computing their own local perception. In Fig. 7, integration of the shared situational awareness algorithm with the existing system architecture of the EV is demonstrated. The EV state information and local perception data are extracted from the localization module and the dynamic object module. In order to receive CPM messages with external perception from other road-users, a V2X communication system was installed and configured.

In this project, Commsignia ITS-OB4, the fourth generation of vehicular connectivity system, enables vehicular communication, presented in Fig. 8. This system communicates to other connected agents in the scenario using Dedicated Short-Range Communications (DSRC), which is based on ITS-G5. ITS-G5 is a European standard for vehicular communication based on IEEE 802.11 standards for its lower layers.

For the test performed in this project, the frequency used for communication is 5.900 GHz. The range varies



Fig. 8: Commsignia ITS-OB4; vehicular connectivity system used in this project.

depending on the operating area. On an open field, the range is more than 1000 meters at the line of sight, but if the receiver is hiding behind something, it is possible the sender is not heard, even if they are close to each other. On the other hand, in an urban environment with many buildings, the signal will bounce on the buildings and therefore be available to receivers around corners and heard while there is no line of sight. This makes it powerful in crowded areas. The capacity is measured to receive around 500 msgs/sec; each message is at 10 Hz in this work.

B. Results

In this section, we will discuss the outcome of the test run of the scenario presented in Fig. 1. In Fig. 9, various snapshots of EV's scene map are presented. The EV's scene map consists of (i) the EV's location, which is represented by the yellow truck, (ii) a map with lane markings, (iii) the ground truth of SV and RSU, and (iii) estimated sets. The green arrows represent the lane center and direction, and the arrow's size is the lane's segmentation. This information about the lanes is obtained from the HD map. It can also be observed that the PV is not included in the local perception, as PV is considered as a static obstacle and is combined with the occupancy grid map, which contains information on infrastructure and occlusion in EV's FOV.

In Fig. 9a, the green polygons represent the estimated set of the detected objects by the local perception of EV. In this scene map, it can be seen that PV occludes the pedestrian. Two agents in the scenario are capable of sharing perception data using V2X communication, i.e., SV and RSU. The global truth of these agents is sent via the CPM message, which is used in coordinate conversion. In Fig. 9b, the red polygon represents the estimated set from the SV perception, and the blue polygon represents the estimated set from RSU perception with reference to the EV's coordinate frame. Both SV and RSU are able to detect all the road-users in the given scenario from their FOV.

After collecting all estimated sets from EV, SV, and RSU perception, these zonotopes are grouped together if they have an intersection indicating that they belong to the same road user and then fused using (2). The resulting fused sets, along with local and external perception, are presented in Fig. 9c. In Fig. 9d, we present the fused set along with the ground truth of SV and RSU. This information about the road-users in the occluded regions, along with the set size, can be used in the motion planning module of EV. The improvement in the scene map can be observed when we compare Fig. 9a and Fig. 9d.

To be able to analyze the sets, the results have to be plotted using the MATLAB plotting tool in Fig. 10. With

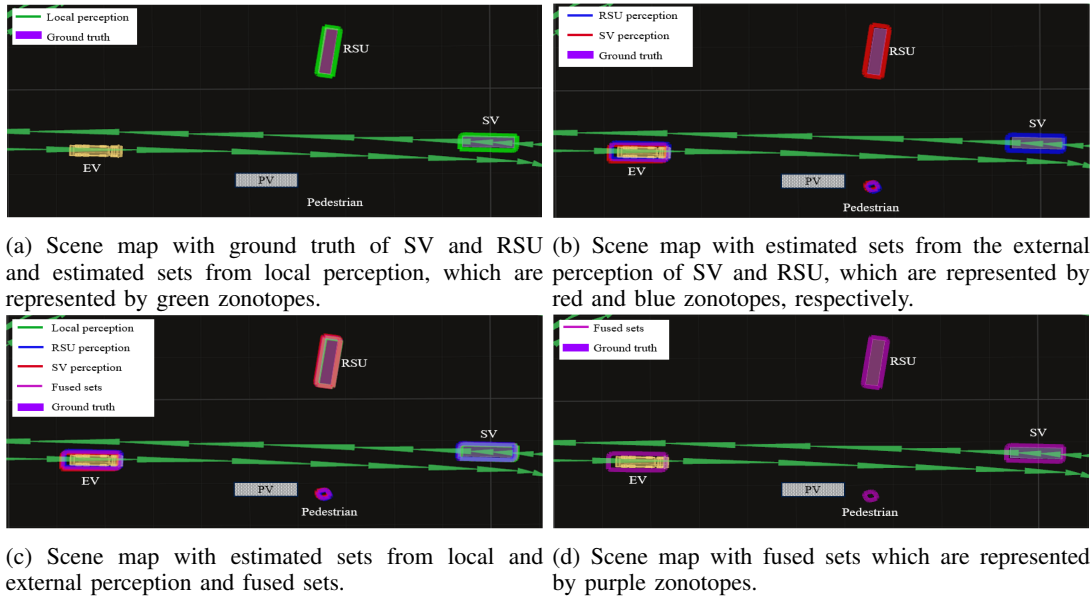


Fig. 9: Real time results from the test run of the given scenario.

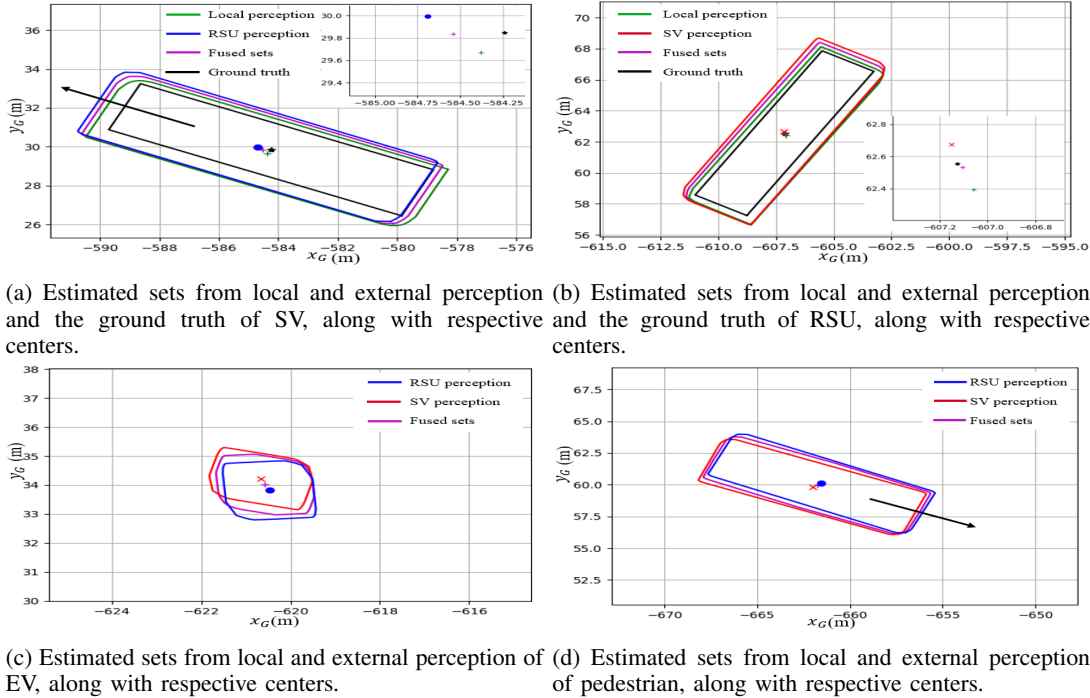


Fig. 10: Matlab plot of real time results of each actor.

TABLE I: Comparison of ground truth, estimated set and fused set of the SV and RSU in the given scenario.

Actor	Ground truth		Estimated set from local perception		Estimated set from external perception		Fused set	
	Center (m, m)	Area m^2	Center (m, m)	Area m^2	Center (m, m)	Area m^2	Center (m, m)	Area m^2
SV	(-584.24, 29.85)	28.08	(-584.38, 29.67)	40.80	(-584.70, 29.99)	42.92	(-584.54, 29.84)	42.04
RSU	(-607.13, 62.55)	28.08	(-607.06, 62.40)	38.18	(-607.15, 62.67)	43.64	(-607.10, 62.53)	41.13

the help of another V2X communication service, i.e., the Cooperative Awareness Message (CAM) service, we obtain the ground truth of both SV and RSU. In Fig. 10a, the estimated sets from local perception and RSU perception are presented with green and blue zonotopes, respectively, along with the fused set using purple zonotope. These estimations are compared with the ground truth of the SV, which is

presented with a black rectangle. On the top right corner of Fig. 10a, the centers of all estimated and fused sets, along with the ground truth, are presented. Similarly, in Fig. 10b, the estimated set from local perception and SV are presented with green and red zonotopes, respectively and along with the fused set presented using purple zonotope. These estimations are compared with the ground truth of the RSU, which is

presented by the black rectangle. The centers of the estimated and the fused sets are compared with the ground truth, which is presented in the lower right corner of Fig. 10b.

In the scenario in Fig. 1, the pedestrian is in the FOV's SV and RSU without an obstruction; therefore, they are detected by SV and RSU. In Fig. 10c, the estimated set from the external perception from SV and RSU are presented by red and blue zonotopes, respectively, and the purple zonotope represents the fused set. The intersecting area of both blue and red zonotopes ideally will have the pedestrian, given the uncertainties in the measurements, which is further over-approximated to the convex zonotopes presented by the purple zonotope. Similarly, estimated and fused sets of the EV can be observed in Fig. 10d.

In Table I, the numerical values of various zonotopes' centers of SV and RSU are compared with their respective ground truths. The area of each zonotope is also calculated and compared with each other. In the case of SV, the RMSE (Root Mean Square Error) between ground truth and estimated sets from local perception and external perception are 0.228 and 0.480, respectively. Similarly, in the case of RSU, the RMSE between ground truth and estimated sets from local perception and external perception are 0.165 and 0.122, respectively. The RMSE between ground truth and fused set for SV and RSU are 0.300 and 0.036, respectively.

VI. CONCLUSIONS

The proposed framework was successfully implemented on CAV's existing system architecture. In the given scenario with occlusion, it was assumed that there are two V2X units, i.e., a stationary and a dynamic V2X unit, close to the pedestrian crossing. These V2X units share external perceptions with EV, and this data is processed by using the proposed framework to obtain safety guarantees. Real-time perception data from both EV and V2X units were used to perform test runs. During the real-time tests, the level of uncertainty in the measurements was realized. The proposed framework can fuse multiple measurements from both local and external perceptions. Improvement in the scene map of EV can be observed by comparing Fig. 9a and Fig. 9d. The set-membership method provides state estimation along with guarantees. These guarantees are essential for safety-critical applications. An immediate next step is to find a way to quantitatively present the improvement of EV's situational awareness while using the shared situational algorithm.

ACKNOWLEDGMENT

This work was partially supported by the Wallenberg Artificial Intelligence, Autonomous Systems, and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. It was also partially supported by the Swedish Research Council, Swedish Research Council Distinguished Professor Grant 2017-01078 Knut and Alice Wallenberg Foundation Wallenberg Scholar Grant and Swedish Strategic Research Foundation CLAS Grant RIT17-0046. We would like to thank Henrik Pettersson from Scania CV AB for his valuable insights and support.

REFERENCES

- [1] Z. Wang, Y. Wu, and Q. Niu, "Multi-sensor fusion in automated driving: A survey," *Ieee Access*, vol. 8, pp. 2847–2868, 2019.
- [2] G. Cui, W. Zhang, Y. Xiao, L. Yao, and Z. Fang, "Cooperative perception technology of autonomous driving in the internet of vehicles environment: A review," *Sensors*, vol. 22, no. 15, p. 5535, 2022.
- [3] D. D. Yoon, B. Ayalew, and G. M. N. Ali, "Performance of decentralized cooperative perception in v2v connected traffic," *IEEE Transactions on Intelligent Transportation Systems*, 2021.
- [4] Z. Chen, L. Xiong, and C. Tang, "Perception-aware path planning for autonomous vehicles in uncertain environment," tech. rep., SAE Technical Paper, 2022.
- [5] Y. Mo, P. Zhang, Z. Chen, and B. Ran, "A method of vehicle-infrastructure cooperative perception based vehicle state information fusion using improved kalman filter," *Multimedia Tools and Applications*, vol. 81, no. 4, pp. 4603–4620, 2022.
- [6] J. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, "Collaborative mapping with ioe-based heterogeneous vehicles for enhanced situational awareness," in *2019 IEEE Sensors Applications Symposium (SAS)*, pp. 1–6, IEEE, 2019.
- [7] L. Orihuela, C. Combastel, and G. Bejarano, "Set-membership state estimation of autonomous surface vehicles with a partially decoupled extended observer," in *2022 European Control Conference (ECC)*, pp. 2024–2029, 2022.
- [8] R. Yee, E. Chan, B. Cheng, and G. Bansal, "Collaborative perception for automated vehicles leveraging vehicle-to-vehicle communications," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1099–1106, IEEE, 2018.
- [9] R. Yu, D. Yang, and H. Zhang, "Edge-assisted collaborative perception in autonomous driving: A reflection on communication design," in *2021 IEEE/ACM Symposium on Edge Computing (SEC)*, pp. 371–375, 2021.
- [10] V. Narri, A. Alanwar, J. Mårtensson, C. Norén, L. Dal Col, and K. H. Johansson, "Set-membership estimation in shared situational awareness for automated vehicles in occluded scenarios," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 385–392, 2021.
- [11] J. Bolting and S. Fergani, "The iterated extended set membership filter applied to relative localization between autonomous vehicles based on gnss and uwb ranging," *Asian Journal of Control*, vol. 21, no. 4, pp. 1556–1565, 2019.
- [12] J. Blesa, V. Puig, and J. Saludes, "Robust fault detection using polytope-based set-membership consistency test," *IET Control Theory & Applications*, vol. 6, no. 12, pp. 1767–1777, 2012.
- [13] A. Alanwar, J. J. Rath, H. Said, and M. Althoff, "Distributed set-based observers using diffusion strategy," *arXiv preprint arXiv:2003.10347*, 2020.
- [14] G. Belforte, B. Bona, and V. Cerone, "Parameter estimation algorithms for a set-membership description of uncertainty," *Automatica*, vol. 26, no. 5, pp. 887–898, 1990.
- [15] F. Xu, S. Yang, and B. Liang, "Interval set-membership estimation for continuous linear systems," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 14, pp. 5305–5321, 2020.
- [16] W. Kühn, "Rigorously computed orbits of dynamical systems without the wrapping effect," *Computing*, vol. 61, no. 1, pp. 47–67, 1998.
- [17] D. Gonzalez, J. Perez, V. Milanes, and F. Nashashibi, "A review of motion planning techniques for automated vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1135–1145, 2016.
- [18] ETSI, *Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Analysis of the Collective Perception Service (CPS); Release 2*. The European Telecommunications Standards Institute, ETSI TR 103 562 V2.1.1 (2019-12) ed., 2019.
- [19] V. T. H. Le, C. Stoica, T. Alamo, E. F. Camacho, and D. Dumur, "Zonotope-based set-membership estimation for multi-output uncertain systems," in *IEEE International Symposium on Intelligent Control*, pp. 212–217, 2013.