2025 IEEE Vehicular Networking Conference (VNC) | 979-8-3315-2437-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/VNC64509.2025.11054108

Demo: Integrating AR headsets with Mobile Object Detection to extend road sensing and VRU safety

Gonçalo Silva*†, André Clérigo *†, Pedro Rito*†, Susana Sargento*†
*Departamento de Electrónica, Telecomunicações e Informática, Universidade de Aveiro, Portugal
†Instituto de Telecomunicações, Universidade de Aveiro, Campus Universitário de Santiago, Aveiro, Portugal

***Abstract*—This work showcases a framework that integrates Augmented Reality (AR) headsets and mobile object detection to enhance the safety of road users in a smart city environment. The framework leverages AR devices, such as the Microsoft HoloLens 2, as mobile sensing hubs to capture real-time sensor data. By equipping Vulnerable Road Users (VRUs) with these devices, the system enables dynamic sensing and data contribution, addressing the limitations of fixed infrastructure sensors. The proposed architecture employs microservices to ensure modularity and real-time data processing, meeting the latency requirements for safety-related systems, of under 300 ms. The demonstration features real-world tests, showcasing its potential to improve urban safety and mobility applications.**

***Index Terms*—Vehicle-To-Everything, Augmented Reality, Vulnerable Road User, Collective Perception System, Sensor Fusion.**

I. Introduction

Smart Cities are at the forefront of communication and orchestration technologies to deliver scalable solutions tailored to the needs of their constituents. A prominent example is the city of Aveiro, which established the Aveiro Tech City Living Lab (ATCLL) platform. This platform integrates multiple communication, sensing and computing technologies into a unified platform capable of environmental and behavioral sensing that enhances traffic management, Intelligent Transportation Systems (ITSs), and road safety [1].

However, deploying Internet of Things (IoT) devices city-wide poses significant challenges. Existing approaches rely on sensors placed at fixed infrastructure or single-modality processing [1], which limits their adaptation to urban scenarios.

Industry leaders consider AR as the next evolutionary step beyond smartphones, potentially replacing them [2]. Studies have already investigated the benefits of VRUs using AR assistance for road navigation [3] [4] using smart city infrastructure and data; however, these devices can also contribute to the city by acting as a mobile sensor unit. Having multi-modal sensing capabilities, these devices can provide real-time data otherwise unreachable, allowing for smart cities to implement innovative ITS applications, especially when paired with Vehicle to Everything (V2X) communication technologies [5].

This work demonstrates a framework that utilizes an AR headset and computing devices to act as mobile sensing hubs. This integration enables VRUs to work as dynamic sensors, supplying data for urban safety and mobility applications, while also consuming the smart city's data to enhance road users' safety.

II. System Architecture

The proposed system architecture is designed to be part of ITS, integrating V2X communications, by issuing Collective Perception Message (CPM) through a Collective Perception Service (CPS). This architecture, displayed in Figure 1, leverages microservices and focuses on mobile computing on the VRU.

The system employs the Microsoft HoloLens 2 as the AR device, chosen for its advanced research mode capabilities [6], along with a GNSS device for localization data. The research mode on Hololens grants access to raw sensor data from its onboard components, including the depth camera, Inertial Measurement Unit (IMU) and RGB camera, among others.

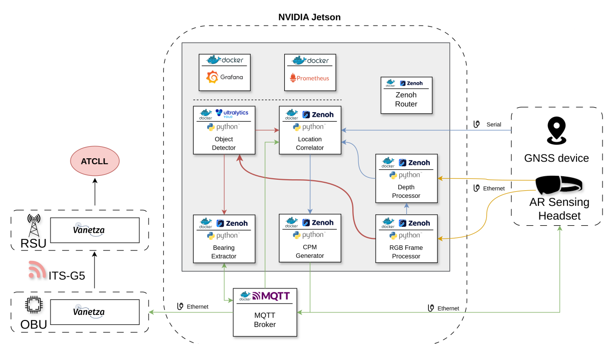

Fig. 1. Demo Architecture.

The system consists of several independent services. The RGB Frame Processor ensures consistent processing and distribution of RGB frames. The Depth Processor correlates depth frames with the nearest, in time, RGB frames, using both frames' calibration data. The Object Detector uses YOLO[1], to obtain the objects' bounding boxes from RGB frames. The Location Correlator performs sensor fusion on detected objects with correlated RGB-Depth (RGBD) frames, GNSS data, and the VRU's heading (through the AR headset) to determine the detected object's GPS location and bearing. The CPM Generator generates and publishes CPMs to an MQTT broker for forwarding to Vanetza-NAP [7]. The Bearing Extractor [4] determines the bearing to objects using detection results, and the user's heading, demonstrating its potential to enhance road user safety. Here, this service is displayed as an example of

[1]https://github.com/ultralytics/ultralytics

979-8-3315-2437-1/25/$31.00 ©2025 IEEE

what ITS applications can leverage from the mobile sensing of AR headsets.

The communication between services, illustrated in Figure 1, primarily occurs over Zenoh (blue arrows), in a peer-to-peer Pub/Sub manner. Additionally, some services have been developed to connect with external functions and use ZMQ sockets (red arrows), TCP (yellow arrows) or MQTT (green arrows).

III. Demo Description

This work demonstrates the capabilities of real-time object detection, V2X communications and their contribution to extend road safety, through testing in both laboratory and real-world conditions, particularly leveraging the existing ATCLL infrastructure. This demonstration will be shown live with an edge node, a real bicycle and an autonomous vehicle.

A. Scenario Setup

The Microsoft HoloLens 2 device is flashed with a Unity application that includes a library for remote sensor access[2] and connected to an NVIDIA Jetson Orin Nano Developer Kit[3] via Ethernet (Figure 2). The Jetson hosts the computationally heavy services that process all the information needed for object detection (see Figure 1). It connects to an On-board Unit (OBU), equipped with ITS-G5 communications. This setup allows messages to be broadcasted through the OBU's Vanetza-NAP instance and relayed to the ATCLL infrastructure.

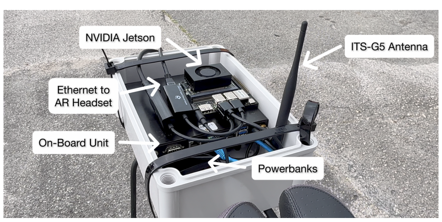

Fig. 2. Demo Setup.

B. Real-time CPM generation

In this demonstration, the user wears the Microsoft HoloLens 2 whilst on a bicycle, with the computing and communication devices attached to it. The scenario involving a cyclist was selected because bicycles are capable of carrying the extra hardware required for this mobile sensing unit. Additionally, future trends could integrate AR into cyclists' helmets, as this has already been done in research and commercial equipment aimed at motorcyclists[4].

As depicted in the Figure 3, the bicycle was stationary for recording purposes, but it had no power source nor connectivity cable attached, meaning that it could be ridden freely, as long as there is ITS-G5 connectivity. The demonstration begins with the autonomous vehicle, PIXKIT 2.0[5], passing by the user. The system recognizes the vehicle, retrieves its distance, location and bearing, and generates a CPM, sending it to Vanetza-NAP. Shortly after, a walking VRU passes by, demonstrating the system's ability to recognize different types of objects. The same metrics and messages are produced for both, and their attributes are displayed on the real-time ATCLL platform. The real demonstration can be visualized in this link[6].

[2]https://github.com/jdibenes/hl2ss
[3]https://developer.nvidia.com/embedded/jetson-nano
[4]https://www.garmin.com/en-US/p/721258/
[5]https://www.pixmoving.com/

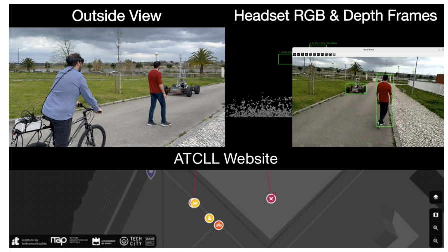

Fig. 3. Demonstration of the results.

IV. Conclusions

This demonstration integrates an AR headset for mobile object detection to enhance road safety. By leveraging AR devices as mobile sensing hubs, this framework addresses the limitations of fixed infrastructure sensors in existing approaches. The proposed architecture uses microservices for modularity and real-time data processing, achieving object detection within 300 ms, specified by ETSI for safety-related systems. The demonstration in real-world conditions showcases its potential in mobility applications. Future work will focus on addressing challenges such as depth camera range limitations (< 7 m) and integrating additional sensing technologies to enhance the system's capabilities.

Acknowledgment

This work is supported by the European Union / Next Generation EU, through Programa de Recuperação e Resiliência (PRR) Projects Nr. 15: AM2R (C644866475-00000012).

References

[1] P. Rito, et al., "Aveiro Tech City Living Lab: A Communication, Sensing, and Computing Platform for City Environments," *IEEE Internet of Things Journal*, vol. 10, pp. 13489–13510, Aug. 2023.
[2] A. Heath, "Why Mark Zuckerberg thinks AR glasses will replace your phone," Sept. 2024.
[3] A. Clérigo, et al., "SafeARCross: Augmented Reality Collision Warnings and Virtual Traffic Lights for Pedestrian Safety," in *16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, (Stanford CA USA), pp. 63–73, ACM, Sept. 2024.
[4] A. Clérigo, et al., "Cooperative augmented reality: Displaying occluded vehicles using v2x communications," in *2025 IEEE Vehicular Networking Conference (VNC)*, 2025, Accepted.
[5] F. Harthi and et al., "Implications of Augmented Reality Applications for Vehicle-to-Everything (V2X)," in *Immersive Virtual and Augmented Reality in Healthcare*, CRC Press, 2023. Num Pages: 25.
[6] J. Dibene, et al., "HoloLens 2 Sensor Streaming," Nov. 2022. arXiv:2211.02648 [cs].
[7] R. Rosmaninho, et al., "Vanetza-nap: Vehicular communications and services in microservices architectures," in *2024 IEEE Vehicular Networking Conference (VNC)*, pp. 297–304, 2024.

[6]https://youtu.be/SQbYsRSChRM