

Systems (IVIS) for intuitive and non-intrusive information display to the driver.

The design of AR in-vehicle systems for infotainment is a challenging task. Rao *et al.* [49] performed an analysis of design methods on different use cases aiming to identify the difficulties in implementation aspects. Despite the vast amount of requirements for these systems to work reliably, such as latency, bandwidth, weather conditions etc, they concluded that the integration of augmented reality in vehicles will help drivers navigate their environment better, and thus will be more widely adopted.

While IVIS existing in many modern vehicles with touch Liquid-Crystal Display (LCD) displays and voice commands may seem to offer most of the utilities of an AR infotainment system, they may actually be distracting to the driver. David *et al.* [50] showed in a recent study that some IVIS require a high cognitive demand or complex command sequences to be handled, and this can in turn lower the awareness of the operator. This is perpetuated by the fact that most IVIS are placed on the dashboard and usually demand their operation to avert (even momentarily) the driver's gaze from the road. In contrast, AR HUDs perform information rendering on top of the environment and thus the driver does not need to share focus in multiple locations.

The distraction potential of AR HUDs was assessed by Kim *et al.* [51]. An AR-enabled windshield was used in a simulated environment with a real-life driving video feed to test various methods of pedestrian visualization. The gaze behavior and cognitive processes were measured and it was found that the visual and cognitive distraction potential of AR depends on the perceptual forms of graphical elements presented on the displays. Specifically, in some cases visualizations, e.g., in the form of a "virtual transparent shadow" indicating the pedestrian's anticipated path, improved the driver's attention without degrading awareness of other objects or scene elements. On the other hand, the use of bounding boxes localizing pedestrians showed to have negative effects, because this approach either overloaded (visually) the scene or degraded the driver's attention on other – not highlighted but possibly critical – scene elements. These outcomes indicate that, while the potential of AR for improving situational awareness is tangible, a lot of attention must be paid for the AR design to not end up cluttering and obstructing the driver's attention.

The research on augmented reality displays on windshields for improving driver awareness also extends to fully Autonomous Vehicles (AV). Such informative human-machine interfaces may help to form a mental model of the vehicle's sensory and planning system, thereby enhancing trust in AV, which is currently quite low in the general public [52], [53], [54]. Lindemann *et al.* [55] conducted a user study on urban environments for evaluating the situational awareness of the driver in various scenarios. They found that their explanatory windshield display had positive results and improved the operator's trust. Yontem *et al.* [56] also designed an AR windshield interface targeting future vehicles. Their main focus was also to increase driver awareness by presenting graphical cues in a non-intrusive way based on a human-centric design and taking into account the human peripheral vision.

While the above methods provide essential feedback on the assessment of such interfaces' design, a significant limitation is that most studies were based on basic or non-interactive simulations, with the steering wheel and pedals not influencing the simulated environment and thus restricting the feeling of immersiveness of the simulations during the evaluation study. A more realistic, experimental study on the benefits of AR in driver's behavior was performed by Kim *et al.* [57] outdoors in a parking lot. It focused on pedestrian collision warning based on visual depth cues delivered in a conformal manner through a monocular display seated above the dashboard, or a volumetric display providing binocular disparity. A limitation of this study, which we address through our AR visualization component (subsection A of section IV), is the limited field of view of the display used in the experiments, potentially creating a tunneling effect of the human vision.

III. OBSTACLE DETECTION

This section presents the proposed methodology on obstacle detection and is followed by section IV on visualization and communication aspects. The main components of the methodology are illustrated in the schematic diagram in Fig. 1 and can be encapsulated in the next steps:

- *Extraction of saliency map:* A saliency value is estimated for any point of the point cloud scene based on its local geometry, as well as the local geometry of its neighboring points.
- *Scene segmentation:* The estimated saliency map is then used as a feature to segment the point cloud into areas characterizing (i) the safe area of the road, (ii) be-aware or dangerous areas within the range of the road, and (iii) areas out of the range of the road.
- *Static object recognition:* Static objects (i.e., potholes and bumps) can be identified and their point coordinates are stored and then used for the AR-based visualization and communication to other nearby vehicles.

In this work, we assume the existence of two or more vehicles (referred as *ego1* and *ego2* vehicles in this paper) that are moving on the same map of a town but not necessarily at the same time, i.e., they are in spatial proximity but possibly not in temporal proximity. Fig. 6 presents an example of two registered point clouds, as received by the LiDAR devices of *ego1* and *ego2* vehicles, showing also their starting points (in arrows). We would like to mention here that all the following analysis is applied to each vehicle separately.

A. Notations

Before presenting details on the individual steps, we provide here the necessary definitions and notations. The input data constitute a sequence of point clouds \mathbf{P}_i , $i = 1, \dots, l$ that represents a set of l consecutive frames acquired by a LiDAR device. Each point cloud \mathbf{P}_i consists of m_i vertices \mathbf{v} , where the value of m_i may be different from frame to frame. The j -th vertex (\mathbf{v}_j) of a point cloud \mathbf{P}_i is represented by the Cartesian coordinates, denoted $\mathbf{v}_j = [x_j, y_j, z_j]^T$, $\forall j = 1, \dots, m_i$, where the index i of the point cloud is omitted for simplification. Thus, all the vertices can be represented as a matrix