Chapter 5

CATEGORIZED LANE GRID MAP FOR HIGHER LEVEL REPRESENTATION

This chapter presents the last module constituting the perception framework proposed, the lane grid-based module. It seeks to provide a complementary high-level representation of (i) the environment based on the data of the DOG and (ii) the causes of unknown space, by using semantic labels and focusing on the road lanes.

In this thesis two grid maps are used: the low-level grid map presented in Chapter 3 (grid-based module), and the lane grid map introduced in this chapter (lane grid-based module). Therefore, remark that the term "grid map" is used to refer to the Cartesian low-level grid map ($\mathcal{G}$), while the term "lane grid map" is used to refer to a high-level grid map, focused on the area covered by the road and explained in the following sections. Note that the cells of the lane grid map will be denoted as sectors ζ to distinguish them from those in the grid map.

This chapter begins by introducing the motivation for using a CLG-based representation and a summary of the proposed approach in Section 5.1. Then, in Section 5.2, the review of the corresponding literature is provided, focusing on representation formats that facilitate the understanding of information and on strategies relying on the use of semantic descriptions. Subsequently, the necessary steps for the development of the proposed strategy are presented. As can be seen in Figure 5.1, these are distributed among three modules. First, Section 5.3 explains the construction of the lane grid using the information of digital maps. Then, leveraging the object-level module, Section 5.4 introduces a procedure to extract objects from the grid-based module and compute their theoretical occlusions. Lastly, Section 5.5 presents the calculation of the CG using the occupancy state of the grid-based module and the information of the extracted objects, followed by its transformation into the CLG. Afterward, in Section 5.6, various experiments are shown, evaluating the performance and utility of the proposed CLG. The chapter ends with Section 5.7 providing some conclusions.

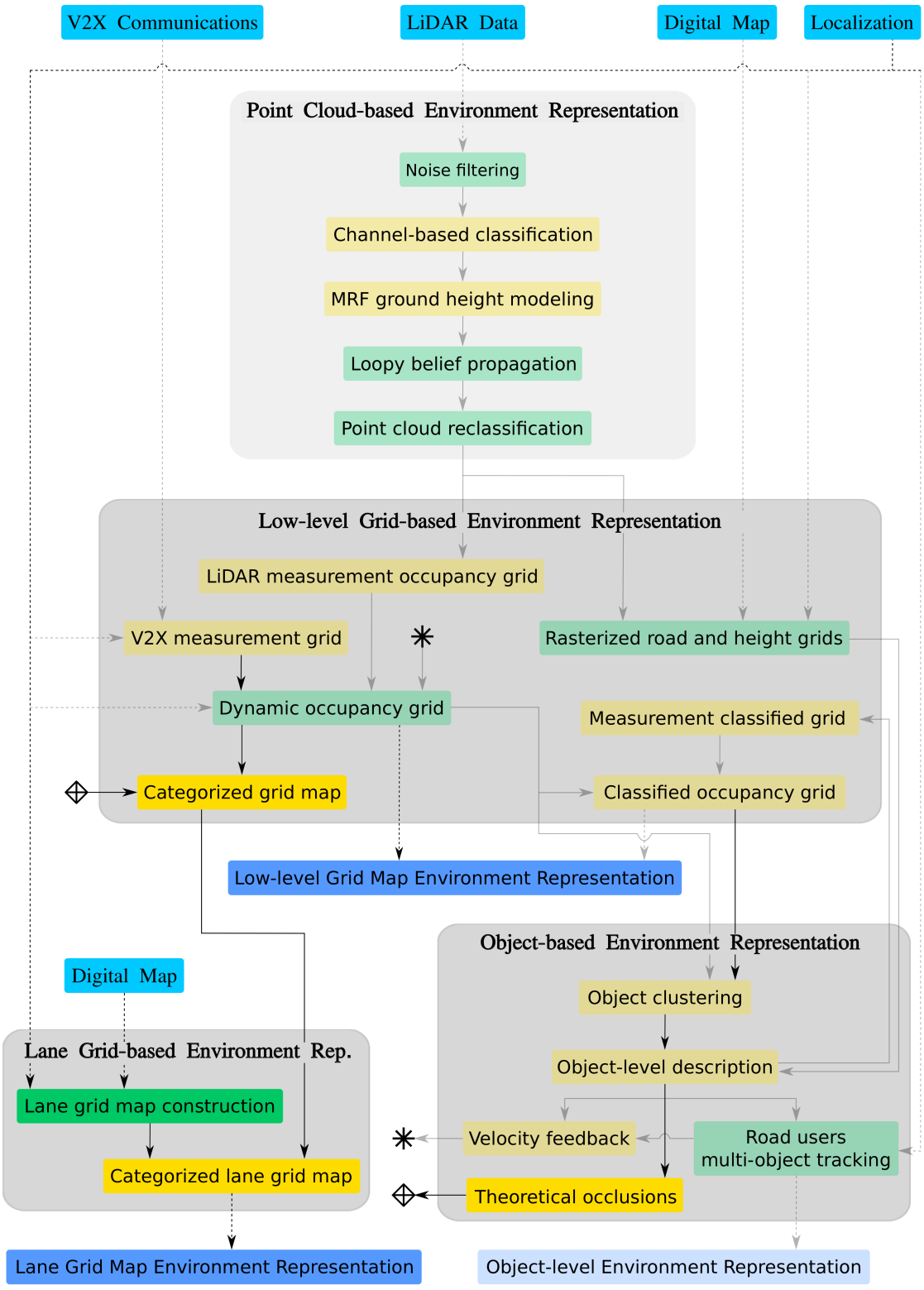


Figure 5.1: Perception framework scheme focused on the lane grid-based module. Modules constituting the framework are denoted in grey, input and output data in different shades of blue, commonly used tasks in green, and tasks including contributions in yellow. The tasks not addressed in this chapter are depicted in faded colors.

5.1. Introduction

The motivation to employ a CLG for representing the information of the DOG at a higher level is presented below. Also, a brief overview of the proposed approach is provided.

5.1.1. Categorized Lane Grid map motivation

As introduced in Section 1.1, in some circumstances, ease of interpretation and information management is preferable over highly precise details. In this regard, the grid-based module provides a complete representation of the environment; however, it is often composed of a large amount of data whose interpretation may not be straightforward. The object-level representation is a more commonly used representation of the environment and typically involves fewer data. Nevertheless, it only addresses obstacles, lacking the ability to correctly represent free and unknown space. Moreover, neither of them significantly exploits the road information available from digital maps. This chapter introduces a CLG with the aim of filling these gaps. The proposed lane grid serves as a complementary environmental representation that describes the surrounding environment at a high-level by focusing only on the road lanes and using semantic labels to describe the space occupancy, dynamics, and causes.

The motivation behind the proposed CLG is presented below with a simple use case, shown in Figure 5.2:

- **Lane grid:** The lane grid is a tesselation of the space, but, in contrast to the grid map, it only addresses the space covered by the road. This allows to reduce the irrelevant information to a minimum while enabling reasoning in terms of distance along the road. For example, it can be easily interpretable that if the ego-vehicle takes the right exit, it will find a static obstacle inside the road at approximately 20 meters in terms of road distance.
- **Semantic labels:** The low-level grid-based module provides an extensive description of the environment that is crucial to address multiple aspects of autonomous driving. However, it is not always easily interpretable. Describing the surroundings using labels enables the synthesis of information and facilitates the understanding of data. Figure 5.2 clearly shows that (i) the ego-vehicle's lane is occupied by a faster vehicle ahead, (ii) the adjacent lane is free and (iii) the right exit is blocked by a static obstacle.
- **Cell size:** The cells of the grid map are by design small in order to accurately describe the shape of the objects. Nevertheless, some tasks do not require such a detailed description. Indeed, it can make the analysis task more complex and lead to sub-optimal decisions due to irrelevant information or irregular cells' state estimations. A higher abstraction level can be directly achieved by choosing the size of sectors significantly larger than the cells' size. As a result, accuracy in the shape definition may be sacrificed at the expense of homogenizing the space representation and reducing the amount of data required.

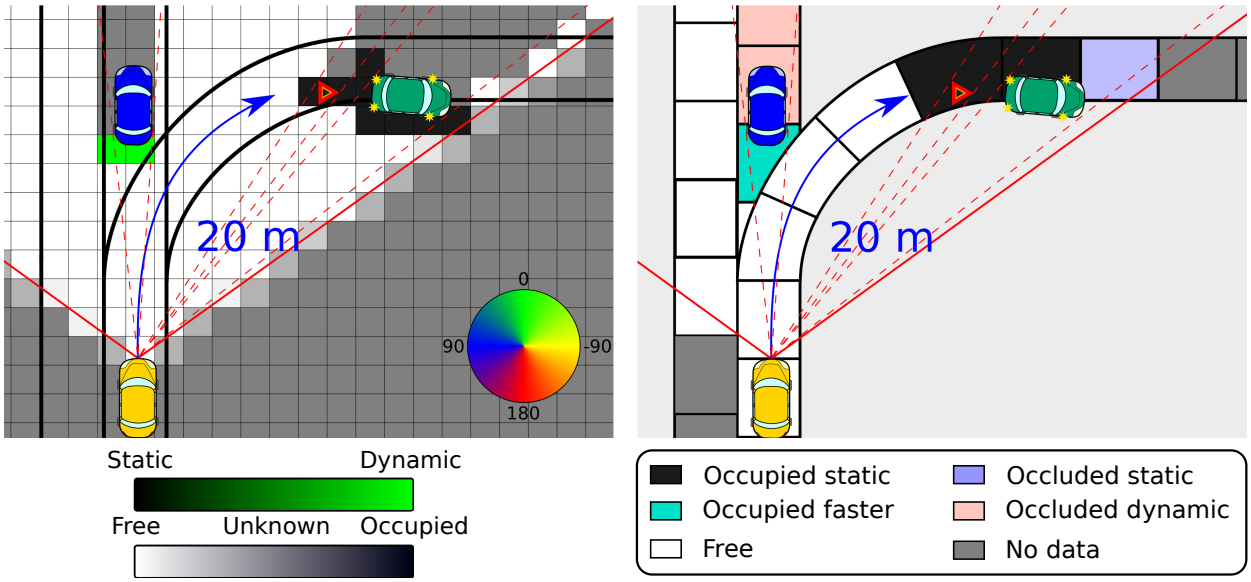


Figure 5.2: Illustrative example of the motivation for lane grid representation. **Left:** Dynamic Occupancy Grid. **Right:** Categorized Lane Grid.

5.1.2. Proposed Categorized Lane Grid overview

As introduced in the previous section, this chapter aims to provide a more easily understandable and manageable representation of the environment. This is accomplished by transforming the information of the grid-based module into a lane grid map representation with a higher level of abstraction.

The proposed semantic representation seeks to describe the surrounding environment using (i) the estimated occupancy and dynamic states of the grid-based module, as they are considered the most important features to describe a driving scene, and (ii) the causes of the unknown space, as it can add valuable environment information not already inferred by the grid-based module presented in Chapter 3.

As already presented during the grid map data description (Section 3.3), the space is primarily classified into three main categories, *occupied*, *free* and *unknown*:

$$\mathcal{L}^{\psi} = \{\mathcal{O},\ \mathcal{F},\ \mathcal{U}\} \tag{5.1}$$

However, the *occupied* and *unknown* categories are now further subdivided into:

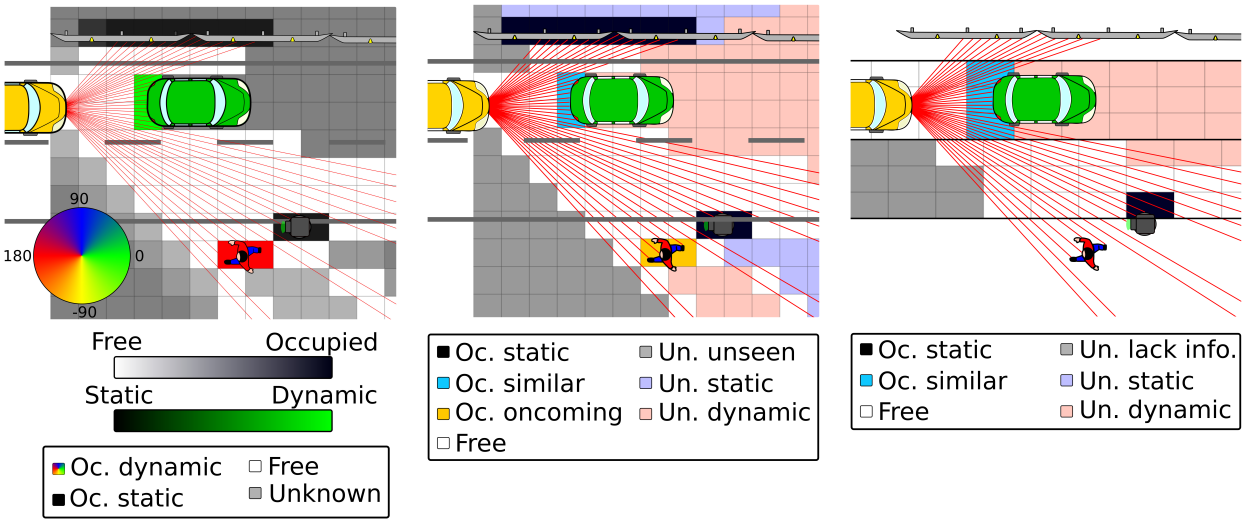
$$\mathcal{O}^{\psi} = \{uncertain,\ static,\ slower,\ similar,\ faster,\ oncoming\} \tag{5.2}$$

$$\mathcal{U}^{\psi} = \{occl.\ uncertain,\ occl.\ static,\ occl.\ dynamic,\ lack\ info.\} \tag{5.3}$$

The key idea behind this refined categorization is twofold: (i) to describe obstacles dynamics in relation to the ego-vehicle's state and (ii) to indicate whether an unknown area is caused by lack of

sensors data or occluded by an obstacle temporarily or permanently.

The generation of this environment representation can be summarized in two main steps: (i) the categorization of the cells of the low-level grid map, and (ii) the transformation of the obtained CG into a CLG. For the first step, the categorization of the occupied and free space can be directly obtained by analyzing the cells' occupancy and dynamic states. On the contrary, in order to categorize the unknown space, the occlusions caused by obstacles in the scene have to be estimated. For the second step, the lane grid map is constructed taking advantage of the available digital maps and an open-source map framework named *Lanelet2* [99]. Then, the labels of the grid cells are fused and translated into the sectors of the lane grid by applying a fusion method based on risk criteria, e.g. occupied space predominates over free and unknown space. Figure 5.3 shows an illustrative example of the desired results from both steps.



a) Input data: Dynamic Occupancy Grid.

b) Categorized Grid.

c) Categorized Lane Grid.

Figure 5.3: Illustrative examples of the proposed CG and CLG.

5.2. Related works

The environment can be modeled in multiple ways obtaining different advantages and accomplishing different objectives. Indeed, Chapters 3 and 4 have presented an environment representation that combines a grid-based representation—intended to provide a general and comprehensive description of the surroundings—and an object-based representation—intended to provide a more common and detailed description of road users.

This chapter proposes the use of a CLG in order to obtain a complementary environment representation to the grid-based module that facilitates the understanding of its outcome while also including new information regarding the drivable area and unknown space. Within the current

state-of-the-art, there exist several works that seek similar objectives. In the following, the works that motivated the development of the proposed strategy are presented. Figure 5.4 shows the results and concepts for some of these works.

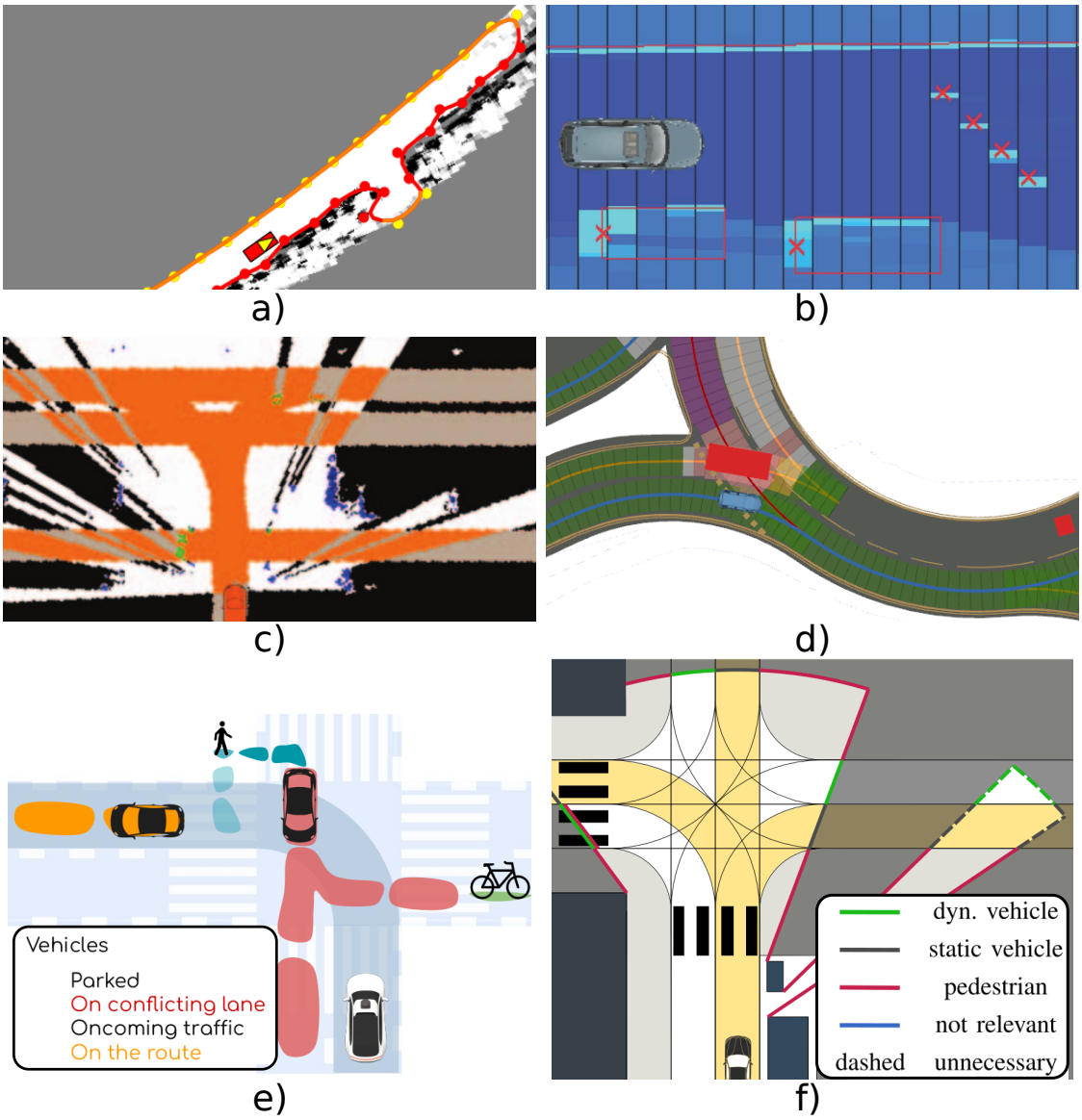


Figure 5.4: Example of four environment representation methods. a) Parametric Free Space map displayed in red and orange over an OG [132]. b) Interval map, occupancy denote by blue shades [160]. c) DOG and digital map, observable road area drawn in orange [51]. d) Semantic lane grid [122]. e) Semantic occupancy map [121]. f) Unknown space borders with phantom relevant traffic participants [66].

5.2.1. Representations for a more comprehensive environment modeling

Schreier et al. [131, 132] introduce a perception framework that combines an OG, an object-level tracking, and, additionally, proposes the calculation of a Parametric Free Space map. By means of the

latter representation, this work aims to enhance the free space representation of the OG and facilitate its interpretation. In contrast to the individual cells of the grid map, the parametric free space map constitutes a continuous and compact representation of the free space surrounding the ego-vehicle. In order to compute this representation, the reachable free space is estimated using the data of the OG as input by applying image processing techniques, such as morphological operations, flood fill, and clustering. Then, it is tracked over time using its boundaries as measurements. Figure 5.4a shows an example of the parametric free space map displayed over the input OG.

[74] also addresses the reachable free space estimation, but, instead of proposing a continuous representation, it samples interval spaces along the ego-vehicle's trajectory and computes the corresponding lateral free space boundaries. An OG is also used as initial environment representation, and steps such as clustering and boundary detection are performed in order to filter the OG data. The free space inside each sector is defined based on the grid map's estimation as the minimum distance from the trajectory to the border occupied areas. Using a similar representation, Weiherer et al. propose in [159, 160] to substitute the common grid map by a more easily interpretable 2D interval map that contains the occupancy state (Figure 5.4b). This decision is motivated by the fact that interval maps allow to obtain a compact representation with minor computation and memory requirements while still providing sufficient accuracy—the data is only discretized in the longitudinal vehicle direction, thus reducing the number of cells, whereas in the lateral component, continuous values can be stored. Moreover, as the map is always aligned with the ego-vehicle's heading, data can be more easily interpreted.

While both, the parametric free space map and the interval map, allow a more scalable representation of the environment and facilitate the interpretation of the obtained estimation, the parametric free space map only addresses free space and the interval map is limited to the ego-vehicle's current position or planned trajectory. Moreover, neither of them takes into account information about the road.

As denoted in [78], one way to translate the estimation of common perception algorithms, grid- and object-based approaches, into a representation format with a higher level of abstraction is to relate them with road map data, for example, associating obstacles to certain lanes. Indeed, the fusion of digital road maps with the outcome of perception modules is a common procedure in tasks such as motion and maneuver planning. For example, [81] combines digital maps and perception output data to compute several planning grids that allow the evaluation of possible trajectories and [51] rasterizes the information of a digital map into a DOG to assess potential hidden obstacles in unobservable regions, see Figure 5.4c. By using such approaches, part of the driving scene context can be included in the environment representation. Nevertheless, the interpretability of the data is not addressed.

The work presented by Sanchez et al. [123, 124, 125] represents the environment as a semantic lane grid (see Figure 5.4d). This representation consists of a tesselation of the space regarding the road lanes, where each cell is sampled along the lanes' centerline and is described by a semantic label.

In this way, the advantages of the interval map, reduced number of cells and easy-interpretability, are combined with the knowledge about the drivable area from digital maps. Moreover, the use of labels further facilitates data understanding by adding semantic context to the space described.

5.2.2. Semantic representation and unknown space categorization

As just briefly introduced in the previous section, the use of semantic labels allows the acquisition of a more easily interpretable representation. In works related to grid maps, semantic concepts are usually addressed by denoting road users' type and drivable space [33, 130]. However, in this thesis, the classification of road users is already achieved through the COG (Section 4.5) and the object-level tracking (Section 4.6.3), while the drivable space is directly obtained from digital maps. In contrast, the use of semantic data for interpretability is perhaps less common but also provides great value.

[69] proposes the use of evidential grids fed with LiDAR data and digital maps in order to categorize the space into different classes that allow a better scene understanding such as temporarily static obstacles, free non-navegable space, buildings, etc. With a similar intention, [121] presents a further categorization of the road users based on their potential interactions with the planned route of the ego-vehicle, see Figure 5.4 e). In this way, the planner module can assign different costs to each category, for instance, parked vehicles present a safer situation than oncoming vehicles.

The approach in [123, 125] aims to facilitate safe decision-making by inferring additional knowledge about unknown space based on detected obstacles and traffic rules. For example, assuming that vehicles keep safe distances with other vehicles, the free space ahead of them can be inferred (*safe* unknown space). Also, based on the lanes' driving rules and the presence of detected obstacles, it is able to infer if the traffic flow on a certain lane is blocked (*neutralized* unknown space).

[97] also addresses the categorization of the unknown space but from the point of view of which unknown areas are relevant for the ego-vehicle. Therefore it computes the boundaries of the visible space and intersects them with road lanes taken from digital maps. Taking into account traffic rules, it is able to identify whether an occlusion is relevant—e.g. an occlusion within a lane that merges into the ego-vehicle's lane—or not—e.g. an occluded lane with opposite direction. [66] proposes a similar approach but further includes potential obstacles that may affect the ego-vehicle trajectory. For example, a static obstacle ahead in the same lane or a pedestrian crossing the road, see Figure 5.4 f).

5.2.3. Progress beyond the state-of-the-art

As presented in the previous section, the use of a higher level environment representation can enhance the interpretability of the information.

Taking into account the reviewed works and the available resources, the semantic lane grid is considered the strategy that better fits the objectives of describing the environment at a higher level

while taking into account information about the road area. Moreover, the semantic categorization of objects' dynamics and unknown space causes provides the modeling with valuable information that is not included in the grid-based and object-based modules.

This chapter proposes a lane grid map representation of the environment that, instead of focusing only on the occupancy state or in the further characterization of the unknown space derived from vehicle interactions, addresses the categorization of the occupied and unknown space with respect to its dynamics and causes respectively.

5.3. Lane grid map

Representing the environment in terms of a lane grid presents two main challenges compared to the common and direct grid map space discretization. First, it involves the processing of digital road maps to obtain the set of lanes surrounding the ego-vehicle. Second, it requires the discretization of these lanes into a set of cells in a way that ensures easy accessibility to the information while preserving the shape of the road.

5.3.1. Lanelets

In order to handle digital road map information, the open-source map framework *Lanelet2* [8, 99] is used. This framework enables representing the drivable environment under geometrical and topological aspects. It is also the road map representation format chosen by other modules integrating the complete autonomous vehicle architecture [80, 81, 82, 153, 155].

Lanelet2 represents the drivable road as a set of interconnected road segments, named Lanelets. Each Lanelet is described by (i) left and right bounds which fit the road shape, (ii) a unique ID, and (iii) a set of relationships with the other Lanelets, e.g. the ID of the preceding Lanelet. Lanelets can also contain additional information, such as traffic rules, but this information is not used in this thesis. For further details, the reader may refer to the works [8, 99]. Figure 5.5 provides an illustrative example.

Therefore, the lane grid map $\mathcal{L}$ is constituted by the set of Lanelets ξ overlapping with the perception framework FOV. This is directly computed by exploring the graph of relationships between Lanelets and selecting all with at least one point within the maximum perception vision range. Other approaches focusing on more specific areas can be used, for example only the ego-vehicle's possible navigation corridors [81] or the set of lanes with which the ego-vehicle may interact [123].

5.3.2. Lanelets discretization

As explained in the previous section, the lane grid map is constituted by a set of Lanelets nearby to the ego-vehicle location. Nevertheless, the lane grid map is a tesselation of the space, thus these

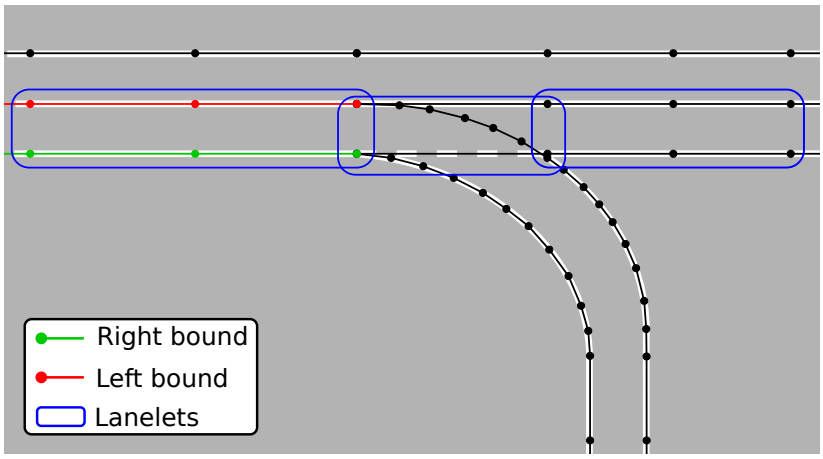


Figure 5.5: Lanelets road map representation example. Only three Lanelets are displayed.

Lanelet must be divided into a set of cells.

In order to obtain this discretization while preserving the road shape, the cells of the lane grid $\mathcal{L}$—named as sectors ζ in order to distinguish them from the low-level grid map cells c—are discretized longitudinally and laterally along each Lanelet. The longitudinal discretization is calculated by sampling along the Lanelet's centerline with a constant distance value. Since Lanelets have different lengths, the total number of longitudinal divisions depends varies depending on each Lanelet. On the other hand, lateral discretization is performed by dividing the lane's width into three sectors of equal lateral resolution. Other divisions can be used, for instance, defining sectors with a width equal to the road lane's [122] or small-sized sectors [154]. However, a division into three sections provides an adequate high-level representation while being more robust against possible obstacles slightly overlapping the lane, e.g. a car driving close to the centerline or vegetation at the sides of the road. As a result of this discretization process, sectors are not rectangular but quadrangular and their area is not fixed but variable depending on the road lane's width and curvature.

Figure 5.6 shows an example of (i) the input digital road map, (ii) the Lanelets selected for the current iteration, and (iii) the grids calculated for three Lanelets.

5.4. Object extraction and occlusion calculation

As introduced in Section 5.1.2, the semantic description of the surrounding space is obtained, in the first instance, by categorizing the cells of the low-level grid map. The categorization into *occupied*, *free*, and *unknown* space, as well as the further categorization of the *occupied* label can be directly performed by analyzing the cells' state individually. Nevertheless, the *unknown* space cannot be categorized in this way since it is necessary to identify occluded areas and the obstacles causing them. In order to accomplish these latter tasks, *occupied* cells are clustered into potential objects and their corresponding theoretical occlusions are calculated.

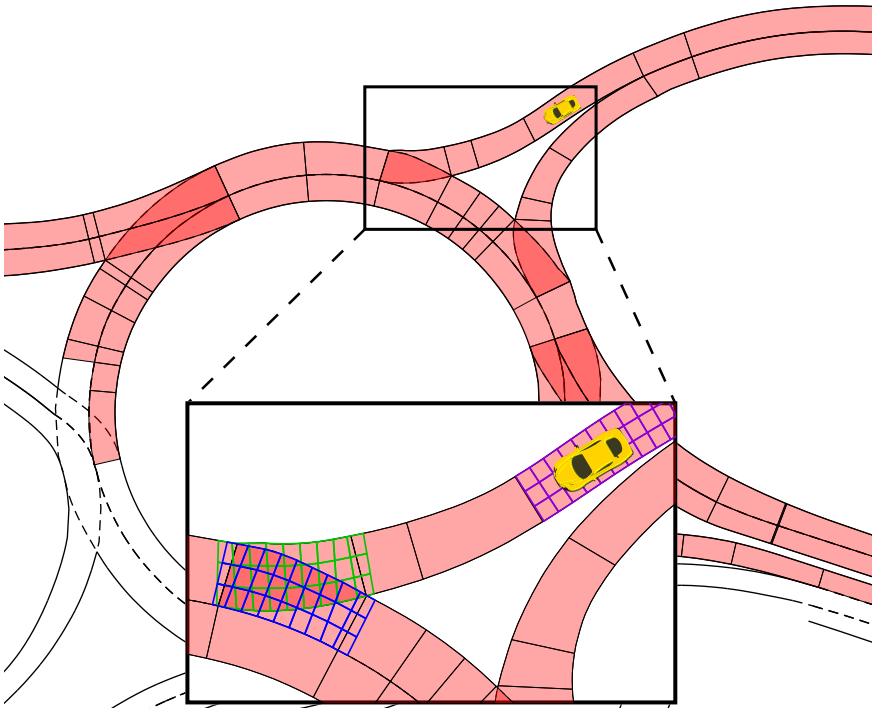


Figure 5.6: Lane grid construction example. The nearby Lanelets contained in the lane grid are denoted in red. Additionally, the discretization into sectors of three Lanelets is shown in the zoomed window.

Moreover, this clustering procedure serves a secondary purpose, the filtering of the data related to *occupied* cells. Since cells are independent of each other, two cells corresponding to the same obstacle may exhibit different dynamic states. The clustering step allows the homogenization of the space concerning obstacles and the filtering of noisy cells' states.

5.4.1. Clustering

Given the objectives of this clustering step, the clustering strategy selected is not intended to group all the cells corresponding to a single obstacle but focuses on obtaining compact clusters—which facilitates the computation of theoretical occlusion—and preventing object merging—to avoid incorrect filtering.

As presented more in detail in Section 4.3, related to object clustering, *occupied* cells are grouped by means of a Connected Components algorithm based on cells' proximity and velocity difference. In order to avoid merging and to obtain the desired compact clusters, an 8-connected kernel is applied. Moreover, the proposed categorization is conservative regarding occupied space, thus all clusters are accepted despite their size and the threshold that defines a cell as *occupied* is set with a low value, e.g. $\mathcal{T}_o^{cl} = 0.1$—recall that cells clustered will be categorized as *occupied*.

Each cluster ς obtained in this way is described by (i) the state $\boldsymbol{s}^{\varsigma}$, (ii) the set of cells constituting the cluster $\boldsymbol{\mathcal{C}}^{cl,\varsigma}$ and (iii) a set of cells $\boldsymbol{\mathcal{C}}_{border}^{occl,\varsigma}$ defining the boundary of its theoretical occlusion. The

state $\boldsymbol{s}^{\varsigma}$ is defined by:

$$\boldsymbol{s}^{\varsigma} = \left[x_{\mathrm{x}},\ x_{\mathrm{y}},\ v,\ \theta,\ n_t^{\rho}\right] \tag{5.4}$$

where x_{x} and x_{y} represent the gravity center, v and θ the velocity represented in terms of module and orientation, and n_t^{ρ} the average number of times that the particles gathered inside it have been resampled—see Section 3.6.3.5.

5.4.2. Theoretical occlusion calculation

The theoretical occlusion aims at approximating the space an obstacle is occluding based on its shape and position with respect to the sensors' reference system. The procedure consists in finding the lateral cells and then computing the visible boundary and the forward projection of the occlusion. Figure 5.7 displays an illustrative example and Algorithm 3 provides the pseudo-code of the calculation.

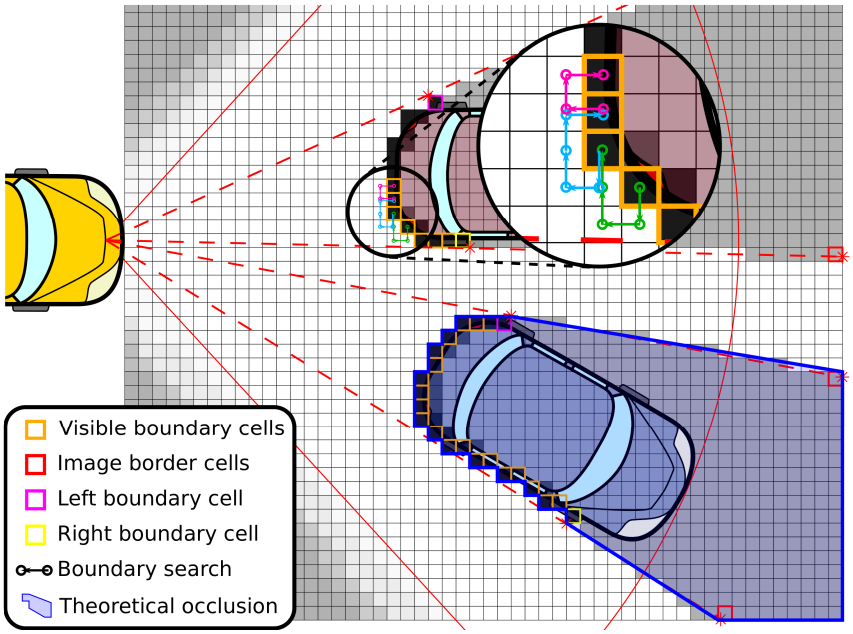


Figure 5.7: Illustrative example of occlusion calculation.

First, the lateral cells $c_{right}^{occl,\varsigma}$ and $c_{left}^{occl,\varsigma}$ are selected as the cells with the highest and lowest angular values with respect to the sensor's position. Once these cells are calculated, the theoretical lateral boundaries of the occlusion can be determined by projecting them forward using a virtual beam that traverses them and intersects with the grid limits, $c_{right}^{occl,\mathcal{G}}$ and $c_{left}^{occl,\mathcal{G}}$. Then, the visible boundary is computed as the set of cells in between the lateral cells $c_{right}^{occl,\varsigma}$ and $c_{left}^{occl,\varsigma}$ facing the ego-vehicle. Since the obtained clusters are compact, this set of cells is calculated using an implementation of the common boundary algorithm Moore-Neighbour Tracing algorithm [42]. This algorithm searches for the next border point by exploring the 8-neighboring cells in an ordered manner, as displayed in Figure 5.7.

Algorithm 3: Theoretical occlusion calculation.

Input: Set of cells $\mathcal{C}^{cl,\varsigma}$ constituting the cluster ς and grid map $\mathcal{G}$
Output: Set of cells $\mathcal{C}^{occl,\varsigma}_{border}$ constituting the theoretical occlusion

```
// Calculate the extreme lateral cells and project them
```
$[c^{occl,\varsigma}_{right}, c^{occl,\varsigma}_{left}] \leftarrow CalculateAngularEndCells(\mathcal{C}^{cl,\varsigma})$
$[c^{occl,\mathcal{G}}_{right}, c^{occl,\mathcal{G}}_{left}] \leftarrow CalculateOcclusionLimits(c^{occl,\varsigma}_{right}, c^{occl,\varsigma}_{left}, \mathcal{G})$

```
// Compute the visible boundary between extreme lateral cells
```
$c_{border} = c^{occl,\varsigma}_{right}$
while $c_{border} \neq c^{occl,\varsigma}_{left}$ **do**
  $c_{border} \leftarrow SearchNextBoundaryCell8PixelKernel(c_{border}, \mathcal{C}^{cl,\varsigma}, \mathcal{G})$
  $\mathcal{C}^{occl,\varsigma}_{visible} \leftarrow AddCell(c_{border}, \mathcal{C}^{occl,\varsigma}_{visible})$
end

```
// Combine into a single set of cells
```
$\mathcal{C}^{occl,\varsigma}_{border} \leftarrow Combine(c^{occl,\varsigma}_{right}, c^{occl,\varsigma}_{left}, c^{occl,\mathcal{G}}_{right}, c^{occl,\mathcal{G}}_{left}, \mathcal{C}^{occl,\varsigma}_{visible})$

Figure 5.8 shows an example of the clustering and theoretical occlusions obtained for a scene with real sensor data. Given that multiple obstacles are detected, in order to facilitate visibility, only the occlusions of the clusters corresponding to the vehicles in the scene are drawn. It is important to notice that theoretical occlusions can overlap, this can be caused by multiple reasons, e.g. objects with different heights, obstacles with parts that let laser beams pass through, etc. This fact needs to be taken into account when categorizing the space.

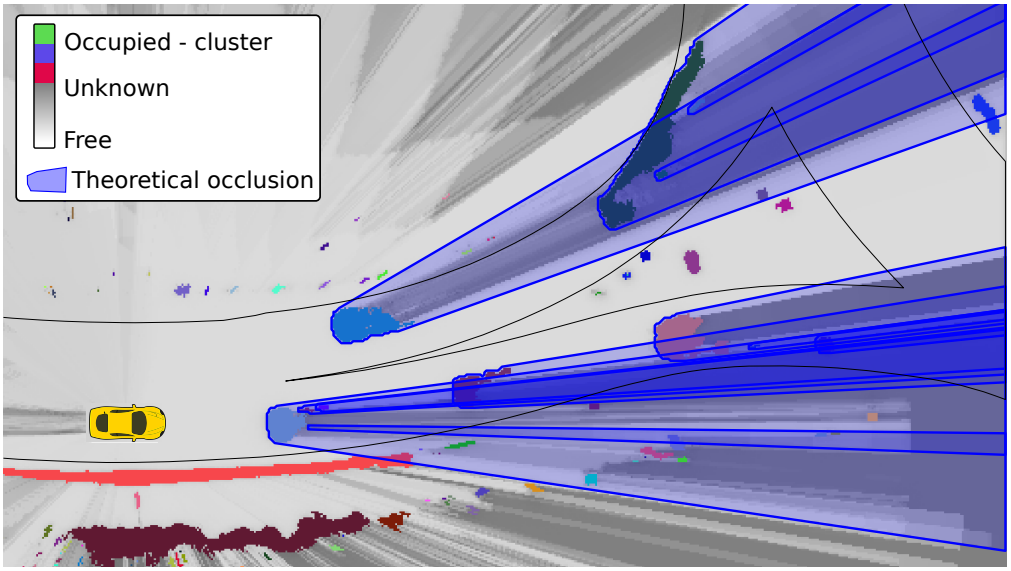


Figure 5.8: Clustering and theoretical occlusion calculation for the CG.

5.5. Space categorization

Once the clusters and their corresponding occlusions have been calculated, the space categorization stage is carried out. The process is illustrated in Algorithm 4 and explained in detail hereafter.

5.5.1. Cells free space categorization

As denoted in Algorithm 4, the free space is addressed first. The categorization strategy selected is conservative with respect to occupied and free space, thus, only the cells not considered as *occupied* and with a mass for free higher than a high threshold $\mathcal{T}_f^{\psi}$ are categorized as *free*. Consequently, and recalling the *occupied* categorization, the division into the three main classes is accomplished as follows:

$$\ell^{\psi} = \begin{cases} \mathcal{O}, & \text{if } m(O^c) \geq \mathcal{T}_o^{\psi} \\ \mathcal{F}, & \text{if } m(F^c) \geq \mathcal{T}_f^{\psi} \ \wedge \ m(O^c) < \mathcal{T}_o^{\psi} \\ \mathcal{U}, & \text{otherwise} \end{cases} \tag{5.5}$$

Small sets of *free* cells are considered irrelevant and not reliable, thus they are filtered using the morphological operation opening [111]. Once all cells are evaluated with respect to *free* category, the remaining cells are initialized as *lack info*. Note that, when categorizing the *occupied* and *unknown* space, if a cell labeled as *lack info.* is related to a cluster, either to its occupation or its occlusion, it is re-labeled accordingly.

5.5.2. Cells occupied space categorization

All the cells clusterized are considered *occupied* and are further categorized based on the cluster's features. This way, as all the cells corresponding to the same cluster are labeled equally, the filtering objective presented in Section 5.4 is accomplished.

The motivation and assignment of the labels concerning the categorization of the *occupied* space are explained in the following:

1. **Uncertain obstacle categorization:** The reliability of a cluster is measured by its approximated age. This is because cells' occupancy state is susceptible to noise due to possible misclassifications in LiDAR point cloud obstacle–ground classification and the dynamic state requires a minimum number of iterations to correctly estimate the velocity of the obstacles. In order to measure the age of a cluster, the age of the particles gathered inside its cells is assessed:

$$\ell^{\psi,\varsigma} = uncertain, \quad \text{if } n_t^{\rho,\varsigma} < \mathcal{T}_{age}^{uncertain,\psi} \tag{5.6}$$

where $n_t^{\rho,\varsigma}$ is the average number of times that the particles have been resampled (5.4) and $\mathcal{T}_{age}^{uncertain,\psi}$ a threshold denoting the minimum age to be considered as reliable.

Algorithm 4: Categorized Grid calculation.

Input: Grid map $\mathcal{G}$, lane grid $\mathcal{L}$, set of clusters $\boldsymbol{S}$ and ego-vehicle's global position $[\boldsymbol{x}^{ego,global}, \theta^{ego,global}]$ and velocity v^{ego}

Output: CG $\mathcal{G}^{\psi}$ and CLG $\mathcal{L}^{\psi}$

```
// Categorize free space
```
$\mathcal{G}^{\psi} \leftarrow CategorizeFreeCells(\mathcal{G}^{O}, \mathcal{T}_{f}^{\psi})$

$\mathcal{G}^{\psi} \leftarrow FreeSpaceImageOpening(\mathcal{G}^{\psi})$

```
// Initialize non-free cells as lack info.
```
$\mathcal{G}^{\psi} \leftarrow InitializeNonFreeCells(\mathcal{G}^{\psi}, lack\ info.)$

```
// Analyze clusters and categorize occupied and unknown space
```
foreach $\varsigma \in \boldsymbol{S}$ **do**

```
    // Categorize cluster ς and its corresponding occupied cells
```
    if $CheckUncertainObstacle(n_{t}^{\rho,\varsigma})$ **then**

        $\ell^{\psi,\varsigma} \leftarrow uncertain$

    else

        if $CheckOncomingObstacle(\boldsymbol{x}^{\varsigma}, v^{\varsigma}, \theta^{\varsigma})$ **then**

            $\ell^{\psi,\varsigma} \leftarrow oncoming$

        else

            $\ell^{\psi,\varsigma} \leftarrow EvaluateVelocity(v^{\varsigma}, v^{ego})$

        end

    end

    $\mathcal{G}^{\psi} \leftarrow CategorizeOccupiedCells(\boldsymbol{\mathcal{C}}^{cl,\varsigma}, \ell^{\psi,\varsigma})$

```
    // Select and update the categorization of unknown occluded cells
    // based on cluster ς
```
    $\boldsymbol{\mathcal{C}}^{occl,\varsigma} \leftarrow SelectUnknownOccludedCells(\mathcal{G}^{O}, \boldsymbol{\mathcal{C}}_{border}^{occl,\varsigma})$

    $\mathcal{G}^{\psi} \leftarrow CategorizeUnknownCellsBasedOnRisk(\mathcal{G}^{\psi}, \boldsymbol{\mathcal{C}}^{occl,\varsigma}, \ell^{\psi,\varsigma})$

end

```
// Transform the grid map into the lane grid map
```
foreach $\xi \in \mathcal{L}$ **do**

    foreach $\zeta \in \xi$ **do**

```
        // Select cells gathered by the sector ζ and fuse its data.
```
        $\boldsymbol{\mathcal{C}}^{\zeta} \leftarrow SelectSectorCells(\mathcal{G}, \zeta, \boldsymbol{x}^{ego,global}, \theta^{ego,global})$

        $\mathcal{L}^{\psi} \leftarrow FuseCellsData(\mathcal{G}^{\psi}, \boldsymbol{\mathcal{C}}^{\zeta})$

    end

end

2. **Oncoming obstacle categorization:** Oncoming obstacles constitute a higher risk than the rest of the dynamic obstacles. Therefore, dynamic obstacles whose current trajectory is

approximately heading towards the ego-vehicle are labeled as *oncoming*:

$$\ell^{\psi,\varsigma} = oncoming, \quad \text{if } \left|\theta^{\varsigma} - \text{atan}\left(\frac{x_y^{ego} - x_y^{\varsigma}}{x_x^{ego} - x_x^{\varsigma}}\right)\right| < \mathcal{T}_{\theta}^{oncoming,\psi} \ \wedge \ v^{\varsigma} > \mathcal{T}_{v}^{static,\psi} \tag{5.7}$$

where the superindex *ego* refers to the ego-vehicle, $\mathcal{T}_{v}^{static,\psi}$ is a velocity threshold and $\mathcal{T}_{\theta}^{oncoming,\psi} < 90^{\circ}$ is a angular threshold.

3. **Speed-based obstacle categorization:** The categorization based on the obstacles' speed ($\{static,\ slower,\ similar,\ faster\}$), seeks to provide a loose reference of the obstacles' dynamics with respect to the ego-vehicle in a way that facilitates decisions related with common interactions with traffic participants. For example, assume a vehicle driving ahead of the ego-vehicle, if its velocity is similar or greater there is no need to modify the current behavior. On the contrary, if its velocity is significantly slower or is a static obstacle, the decision to perform an overtaking or stopping maneuver should be considered. These labels are set as follows:

$$\ell^{\psi,\varsigma} = \begin{cases} static, & \text{if } v^{\varsigma} \leq \mathcal{T}_{v}^{static,\psi} \\ slower, & \text{if } \mathcal{T}_{v}^{static,\psi} < v^{\varsigma} < v^{ego} - \mathcal{T}_{v}^{similar,\psi} \\ faster, & \text{if } v^{\varsigma} > v^{ego} + \mathcal{T}_{v}^{similar,\psi} \\ similar, & \text{otherwise} \end{cases} \tag{5.8}$$

where $\mathcal{T}_{v}^{static,\psi}$ and $\mathcal{T}_{v}^{similar,\psi}$ are velocity difference thresholds.

Only one label can be assigned to each cell. Nevertheless, the extracted clusters can meet multiple of the above-stated conditions, e.g. an *uncertain* obstacle can be also categorized as *faster* and *oncoming*. Therefore, a decision procedure utilizing a risk criterion is applied. The label assigned to each cluster ς, and consequently to the set of the cells composing it, is determined by the following priority order: 1º *uncertain*, 2º *oncoming*, 3º speed-based.

5.5.3. Cells unknown space categorization

The further categorization of the *unknown* space is done by analyzing its possible causes, i.e., if it is occluded by an obstacle and the nature of this obstacle. In this way, the additional labels assigned to *unknown* space are also computed using the obtained clusters.

The set of cells $\boldsymbol{\mathcal{C}}^{occl,\varsigma}$ occluded by a cluster ς are calculated as those cells that fulfill the *unknown* condition stated in (5.5) and are located inside the area defined by the theoretical occlusion $\boldsymbol{\mathcal{C}}_{border}^{occl,\varsigma}$. The calculation of the cells within the theoretical occlusion can be addressed in several strategies, in this thesis, a flooding algorithm is employed.

The categorization of the *unknown* space and the motivation behind the labels are depicted in the following:

1. **Occluded space categorization:** The space theoretically occluded by a cluster is categorized according to the cluster's label $\ell^{\psi,\varsigma}$. However, for *unknown* space, only three labels are

considered: *occl. uncertain*, *occl. static* and *occl. dynamic*:

$$\ell^{\psi,c} = \begin{cases} occl.\,uncertain, & \ell^{\psi,\varsigma} = uncertain \\ occl.\,static, & \ell^{\psi,\varsigma} = static \\ occl.\,dynamic, & \text{otherwise} \end{cases} \tag{5.9}$$

The main objective of this division is to indicate whether the occlusion is temporary (caused by a dynamic obstacle) or constant over time (caused by a static obstacle). The label *occl. uncertain*, on the other hand, is included since *uncertain* clusters can be caused by an occupancy misclassification, and their dynamic state is not yet reliable.

2. **Space lacking information categorization:** The label *lack info.* denotes a space for which the sensors have not been able to provide a reliable occupancy estimation due to other reasons rather than occlusions. Several causes can lead to this situation, for example, the maximum FOV of the sensors, an occupied area for which the estimation is still converging, a free area where the laser beams passing through have an invalid height to model free space (see Section 3.4.2.1), etc. This categorization is accomplished by exclusion, i.e., all *unknown* cells that are not theoretically occluded by a cluster maintain the label *lack info.* assigned during initialization.

Similarly to the *occupied* categorization, each *unknown* cell can fulfill several of the above-stated conditions, but only one label can be assigned. This situation appears when multiple obstacles occlude the same area, as already denoted in Figure 5.8. Therefore, in order to tackle this issue, again, a risk-based decision procedure is performed: 1º *occl. static*, 2º *occl. dynamic*, 3º *occl. uncertain* and 4º *lack info.*.

Figure 5.9 illustrates an example of the obtained CG using real data. It can be seen that the labeling effectively represents the information gathered by the DOG: the ego-vehicle is surrounded by free space in its closest vicinity, but there are multiple obstacles with varying dynamics, e.g. the obstacles within the ego-vehicle's lane are moving forward, whereas those in the left lane are moving in the opposite direction. In addition to this semantic representation of the dynamic grid map, the categorization is able to provide insights about the unknown space. For instance, the unknown space located inside the road is primarily caused by the occlusions of dynamic obstacles, while on the left side of the ego-vehicle, the off-road space remains unknown due to a lack of information (*lack info.*).

5.5.4. Lane grid map categorization

The lane grid is directly categorized by translating the information of the CG cells into the lane grid sectors. This is accomplished by computing the cells located within each sector and fusing their label into a single one. Figure 5.10 provides an illustrative example. The set of cells $\boldsymbol{C}^{\zeta}$ corresponding to each sector ζ are computed by rasterizing its boundaries into the grid map using Bresenham's

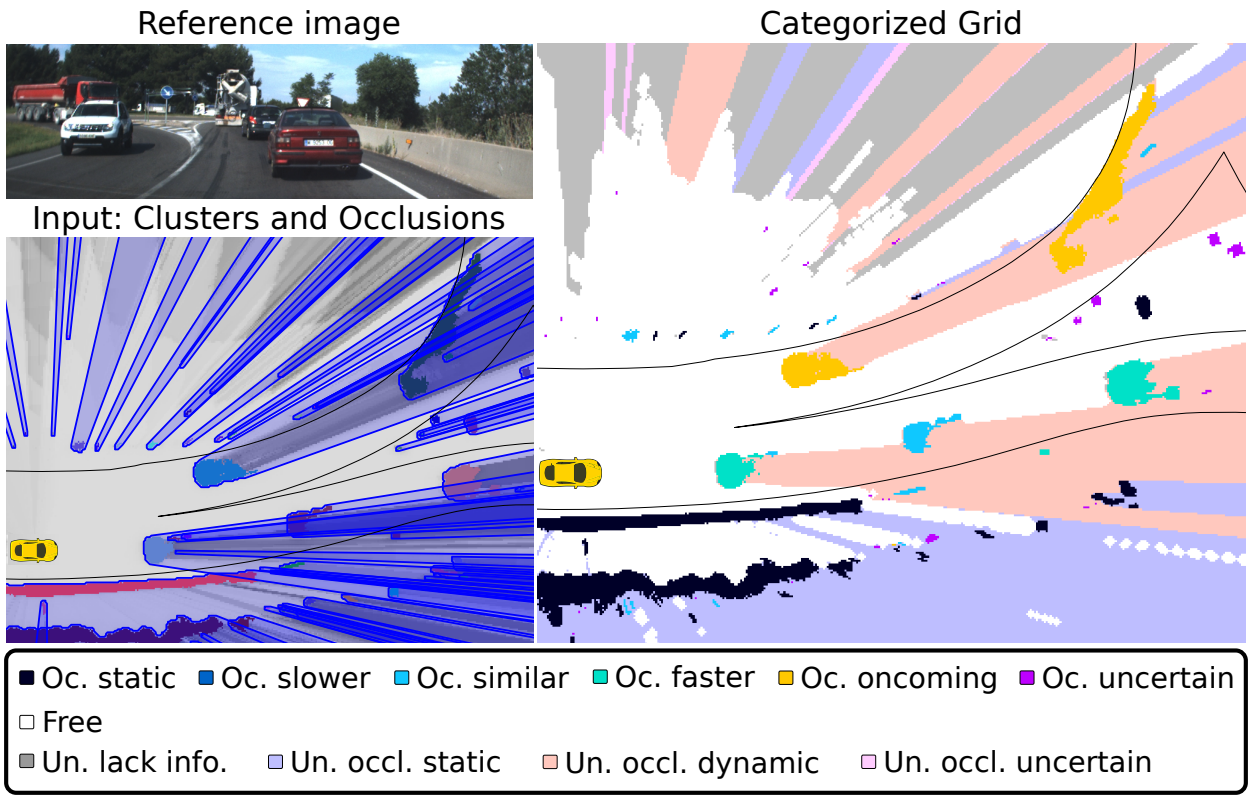


Figure 5.9: Categorized Grid example.

algorithm [12] and a general flooding algorithm. Sectors are slightly enlarged in order to gather partially overlapped cells, e.g. by half of the grid cell size.

The categorization into *occupied*, *free*, and *unknown* maintains the conservative approach that preserves *occupied* over the other labels and seeks to guarantee the space labeled as *free*. Therefore, similarly to (5.5), the division is calculated by:

$$\ell^{\psi,\zeta} = \begin{cases} \mathcal{O} & \text{if } (n_c^{\mathcal{O},\zeta}/n_c^{\zeta}) \geq \mathcal{T}_o^{\mathcal{L}} \\ \mathcal{F} & \text{if } (n_c^{\mathcal{F},\zeta}/n_c^{\zeta}) > \mathcal{T}_f^{\mathcal{L}} \ \wedge \ (n_c^{\mathcal{O},\zeta}/n_c^{\zeta}) < \mathcal{T}_o^{\mathcal{L}} \\ \mathcal{U} & \text{otherwise} \end{cases} \tag{5.10}$$

where n_c^{ζ} is the total number of cells in $\boldsymbol{\mathcal{C}}^{\zeta}$, $n_c^{\mathcal{O},\zeta}$ and $n_c^{\mathcal{F},\zeta}$ the number of cells labeled as *occupied* and *free* respectively and $\mathcal{T}_o^{\mathcal{L}}$ and $\mathcal{T}_f^{\mathcal{L}}$ are percentage thresholds. Once the occupancy categorization is performed, the further categorization of *occupied* and *unknown* space is accomplished by selecting the most frequent label.

Figure 5.11 shows an example of the CLG. It can be seen that a higher level of abstraction has been achieved while maintaining an appropriate level of description. The information of the CG has been significantly summarized and it is expressed in a way easily understandable from the point of view of the road layout. For example, (i) both lanes are occupied by dynamic obstacles moving in

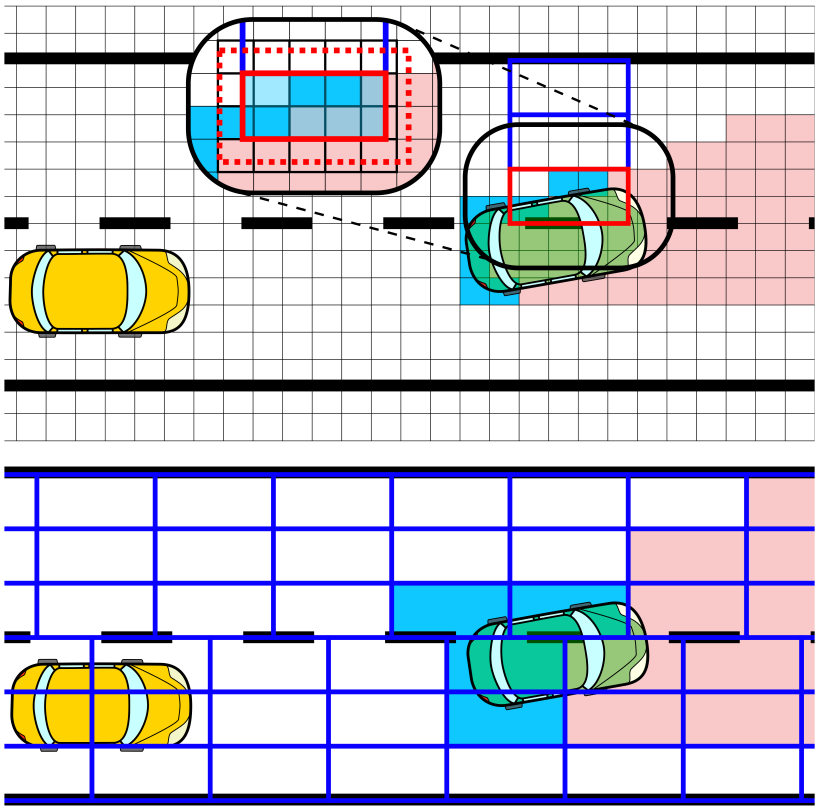


Figure 5.10: Illustrative example of the transformation from grid map to lane grid map. The lane grid sectors are categorized based on the cells gathered within them taking into account a safety threshold (dotted red rectangle). *Occupied* space is favored over *unknown* space, which, in turn, is favored over *free* space.

the direction of their corresponding lanes and close to the right edge and (ii) the ego-vehicle has approximately 3 free longitudinal sectors until the first obstacle in its lane, i.e., approximately 7.5 free meters.

5.6. Evaluation results

In the following the set of experiments conducted in order to validate the lane grid-based module and its usefulness are presented.

Figure 5.12 shows a common urban scenario, a two-way road with dynamic vehicles traveling along both lanes. It can be clearly seen that the process of translating the information from the DOG into the higher level representation of the lane grid map provides a more easily interpretable representation of the driving scene. Moreover, this case is particularly interesting to demonstrate how the categorized space is capable of conveying information that is not easily depicted by the common DOG representation (Section 3.6.4). On one hand, it is clearly denoted how most of the unknown space is caused by occlusions. Moreover, within the road, it is caused by dynamic occlusions, i.e. temporary occlusions. On the other hand, it is able to depict uncertain objects which,

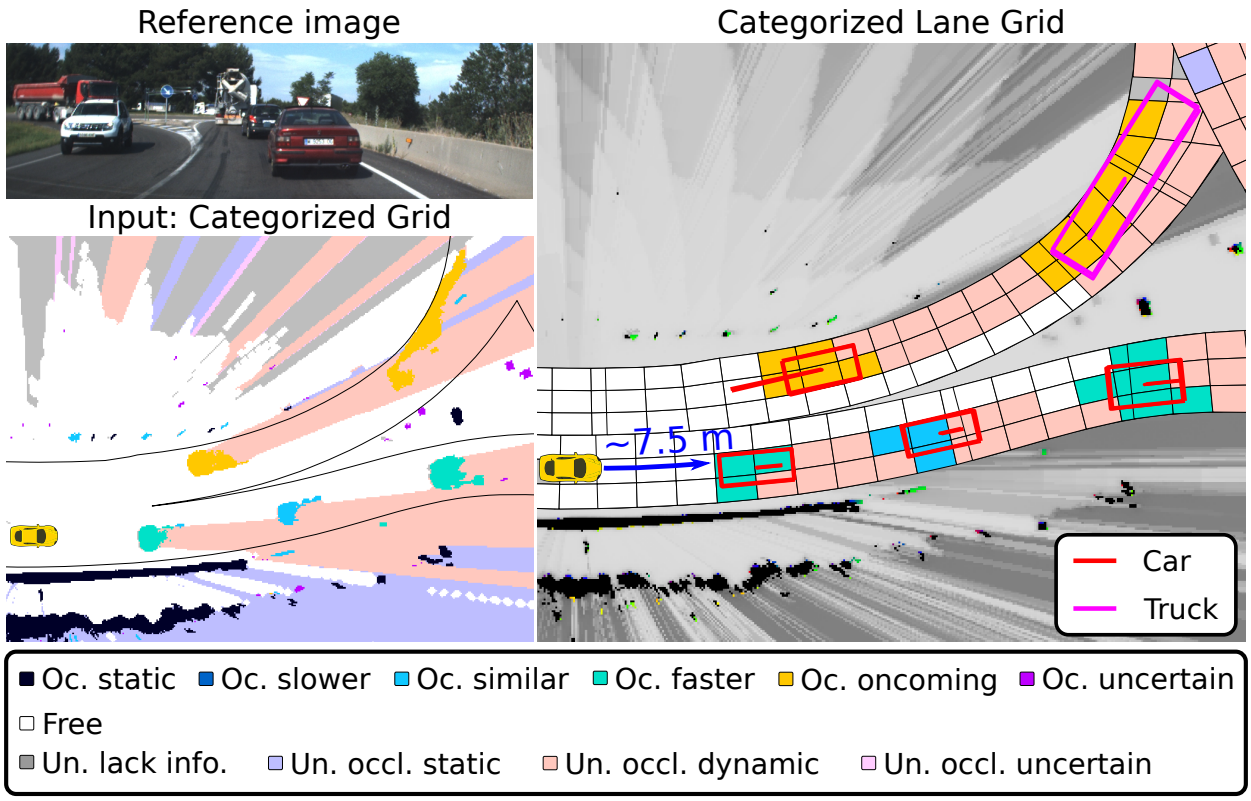


Figure 5.11: Categorized Lane Grid example. Object-level tracking is included in order to provide a reference of the vehicles in the scene.

in the classical representation shown in Figure 5.12a, are not distinguishable from the others.

This difference is highlighted by the dashed circles. The magenta dashed circle denotes a vehicle that has just entered the FOV of the sensors and, thus, its occupancy and velocity estimations have not yet converged. Indeed, the object-level tracking has not been able to initialize a new track for it. In this case, the CG and CLG are able to properly convey that this area is occupied by an *uncertain* obstacles. On the contrary, the DOG representation represents it indiscriminately from the rest of the obstacles. The red dashed circle shows a similar case where the detection of the vegetation surrounding the road is variable and noisy. Therefore, the DOG describes it as small dynamic obstacles, while the categorization process labels it as *uncertain*.

A more complex scene is shown in Figure 5.13 where the ego-vehicle navigates through a road with obstacles nearby to the roadway and enters a roundabout with dense traffic. As a result, most of the space inside the perception framework's FOV is unknown due to occlusions. If an object-level representation is used, as the one shown in Figure 5.13a, a complex environment description constituted by multiple polygons with overlapping areas is obtained. Moreover, the true unknown space is not represented but rather an upper theoretical boundary. When combining these object-level occlusions with the unknown cells of the grid map, a more accurate and interpretable representation is achieved: (i) the use of polygons is avoided, (ii) only the recursively estimated

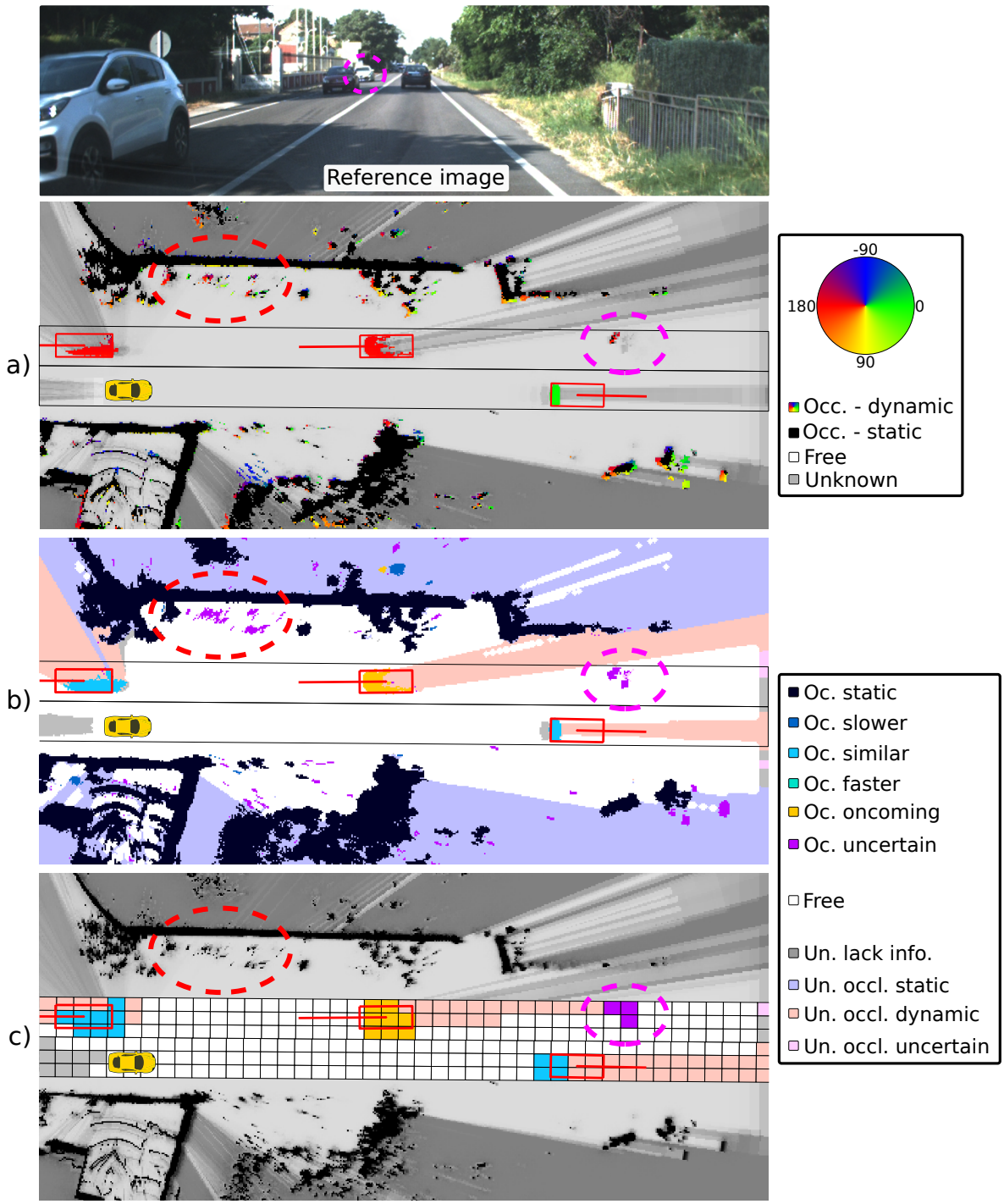


Figure 5.12: Comparison between a) DOG, b) CG and c) CLG. Object-level tracking is included in order to provide a reference of the road users in the scene.

unknown cells are selected, and (iii) the ambiguities due to overlapping occlusions are solved using risk criteria. These last two facts can be clearly appreciated in the occlusion caused by vehicle (1). The area covered by its theoretical occlusion (highlighted with a red dashed line) exceeds the actual estimated unknown space. In fact, within this area, two vehicles and some free space are detected

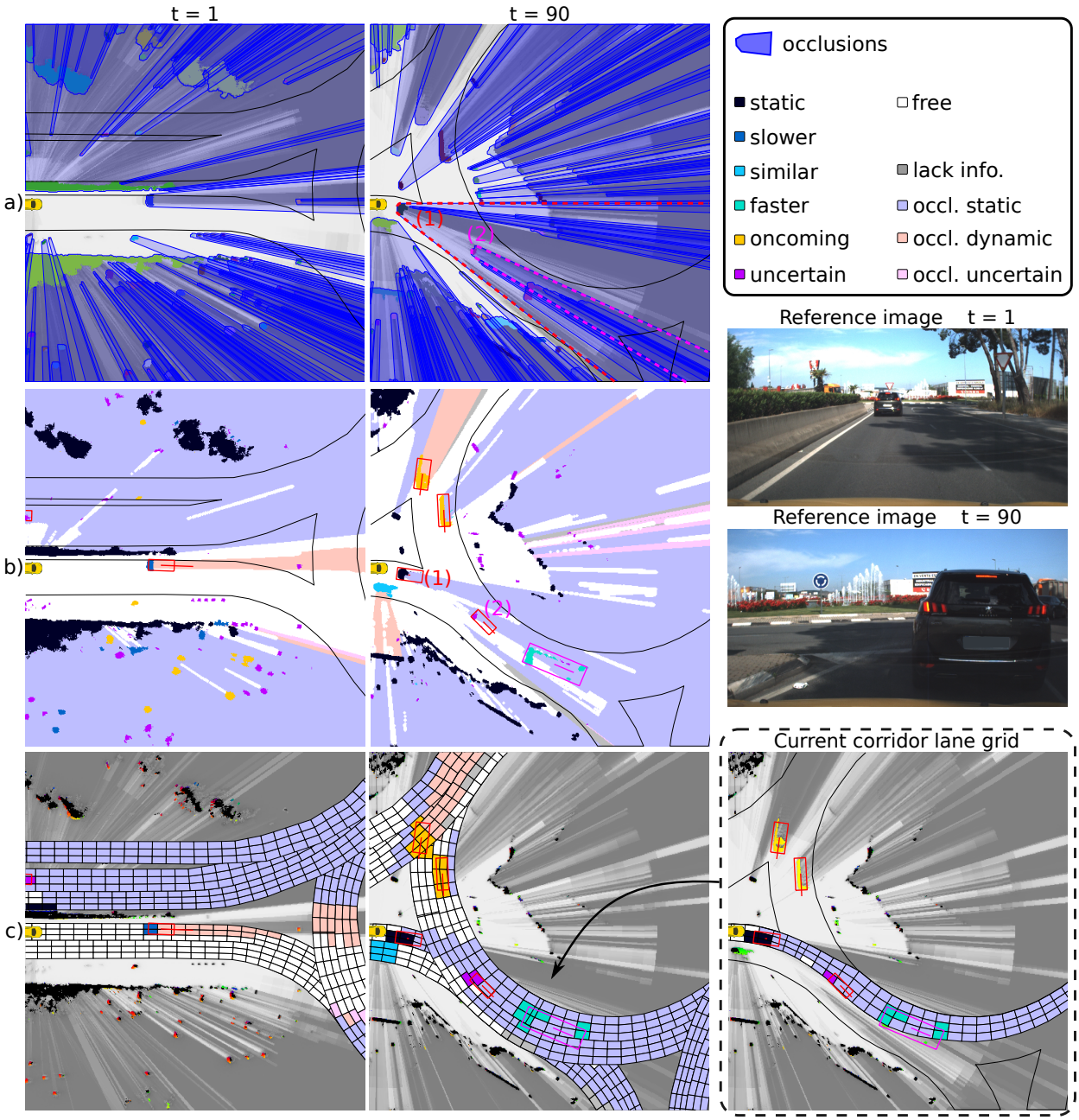


Figure 5.13: Example of the CLG in a complex scenario. a) Occlusions calculation. b) CG. c) CLG. Object-level tracking is included in order to provide a reference of the road users in the scene.

thanks to laser beams passing over and through the windows of the vehicle (1). Furthermore, this situation leads to the overlapping of two theoretical occlusions, the occlusion caused by the static vehicle (1) and the occlusion caused by the dynamic vehicle (2) (highlighted with a magenta dashed line). Since a static occlusion persists until the object causing it changes its dynamic state or the ego-vehicle moves, it is considered of greater importance. Therefore, the conflict between both semantic labels is resolved to *occl. static*.

The lane grid map further facilitates the interpretation of the estimated environment from a road-based perspective. As can be seen, the data used to describe the driving scene is significantly reduced in comparison to the grid map. A broader description is obtained, but it still conveys accurately the current situation: (i) an obstacle is located ahead the ego-vehicle, first with a dynamic state and then static, (ii) most of the space inside the roundabout is occluded due to static obstacles, and (iii) to the left of the ego-vehicle, inside the roundabout, there are oncoming obstacles traveling moving both lanes. Additionally, since the lane grid map is constituted by Lanelets with known IDs, the representation of specific areas is easily accessible and always expressed in the lane's direction. For example, in the dashed black rectangle, only the sectors corresponding to the current navigation corridor are displayed.

Finally, Figure 5.14 illustrates the evaluation of a basic lane-change maneuver from the standpoint of the detected obstacles and available space, illustrating how the CLG may ease the burden of the decision-making module. The objective is to obtain a preliminary evaluation using the information of the CLG instead of the grid-based and object-based representations—which are more complex and computationally expensive due to their high level of detail.

The lane-change maneuver is evaluated using two potential trajectories (denoted in green), one remaining at the current lane and the other changing to the adjacent lane. Each sector traversed assigns a cost c according to its semantic label, e.g. $c_{similar}$ defines the cost of traversing an occupied sector labeled as *similar*. This way, the corresponding cost of each trajectory is calculated based on the sectors traversed. If the obtained costs are favorable for the adjacent lane, the lane-change maneuver can be considered and should be further evaluated in a more comprehensive manner i.e., assessing more detailed environment estimation (object- and grid-based representations) and other relevant elements such as motion prediction, route planning, etc. The results of this preliminary evaluation are as follows:

- **Frame 1:** In the first frame, the left lane is completely free while the right lane is occupied and occluded by dynamic obstacles. Therefore, from the point of view of the perception of the environment, the left lane presents a lower risk than the right lane. Indeed the lane-change maneuver is performed.
- **Frame 2:** In the second frame, dynamic vehicles are found at both lanes. The right lane is circulated by a slower obstacle, but the left lane has a bigger area occluded. Therefore, both lanes receive similar scores. In this way, the lane-change can be further evaluated. In this case, it is dismissed.
- **Frames 3 and 4:** During frames 3-4, both lanes are traversed and occluded by dynamic vehicles. Nevertheless, slower vehicles are traveling through the right lane, thus favoring a lower cost when staying in the left lane.
- **Frame 5:** The left lane is clearly beneficial over the right lane since it is clearly visible.
- **Frame 6:** In the last frame, both lanes are visible. Thus, from the perspective of the perception

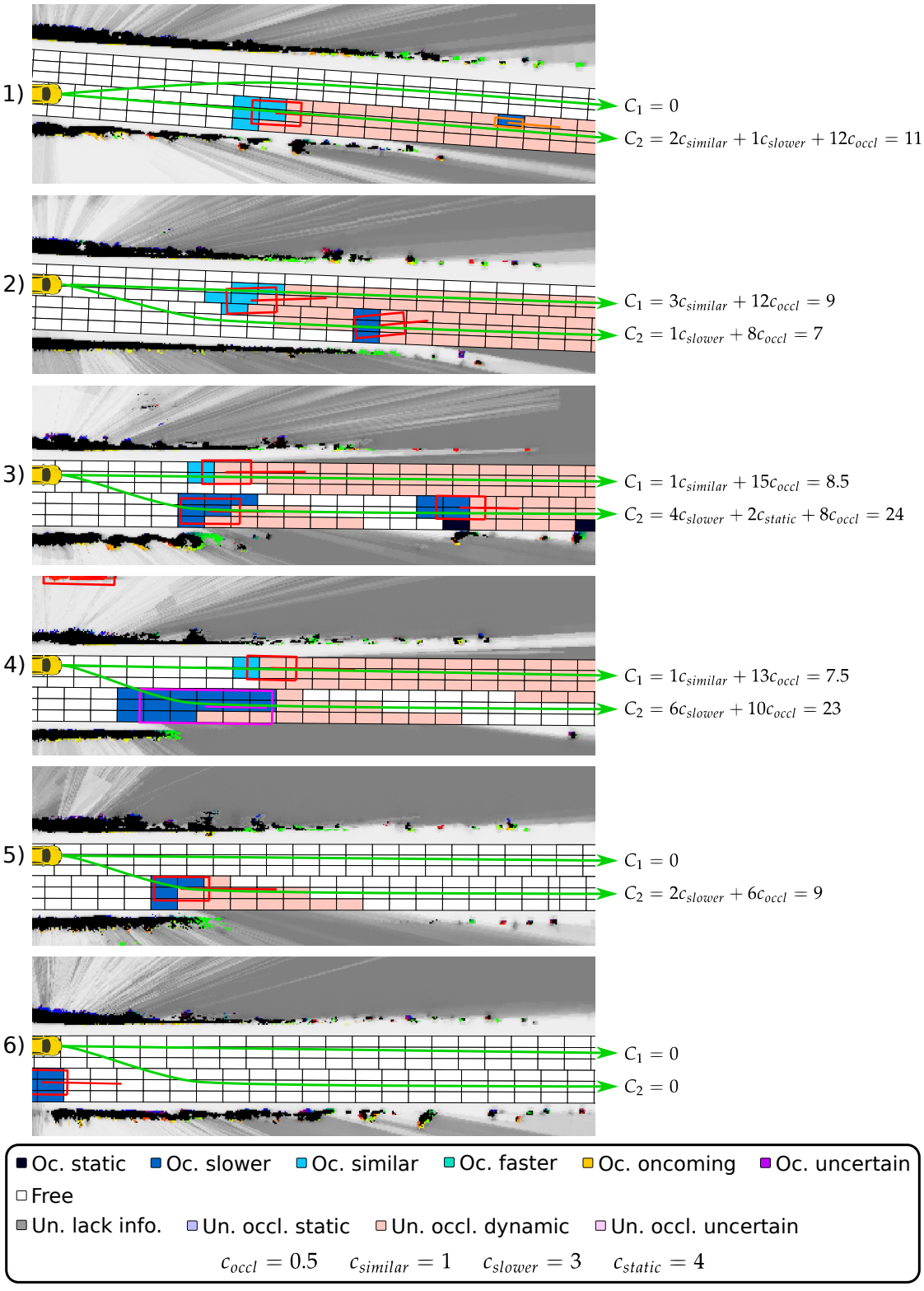


Figure 5.14: Simple example of CLG utility for preliminary decisions. The lane-changing maneuver is assessed based on the labels of the sectors traversed. Object-level tracking is included in order to provide a reference of the road users in the scene.

framework, both lanes have null risk.

Recall that this example is intended to provide a preliminary evaluation, a more elaborated evaluation should be performed when making the final decision. For example, in frame 6 the motion of the adjacent vehicle should be considered before lane changing.

5.7. Summary and conclusions

As stated in Section 1.2.3, the lane grid-based module is intended to provide a more easy-to-understand description of the estimated environment. This is achieved by introducing a complementary representation that (i) focuses only on the road area, (ii) uses a low-resolution discretization and (iii) describes the information using semantic labels.

The computation of the proposed CLG involves two main steps: (i) grid cells categorization and (ii) lane grid re-representation. The first step classifies the cells of the grid-based module into *occupied*, *free*, and *unknown*. Additionally, the *occupied* cells are further categorized with respect to their dynamic state and *unknown* cells with respect to their causes. In the second step, the categorized cells are transformed into a CLG that fits the shape of road lanes by taking advantage of digital maps and applying a risk-based fusion.

In comparison to similar approaches, the strategy presented in this chapter explores the use of the CLG to achieve a complementary re-representation of a DOG. The intention is to obtain a more understandable and manageable description, while also incorporating new insights into the causes of the unknown space.

Chapter 6

CONCLUDING REMARKS

This thesis presents a LiDAR-based perception framework for autonomous driving vehicles. It focuses on road user detection and occupancy space estimation, combining different representation strategies to provide a more comprehensive description. This chapter summarizes the developments presented in this manuscript, addresses the corresponding conclusions and outlines the directions for future research. The chapter concludes with the scientific dissemination resulting from this work.

6.1. General conclusions

Vehicles need to understand their surroundings in an accurate and comprehensive way in order to navigate safely and autonomously. In this connection, this thesis addresses the problem of perception of the environment by proposing a LiDAR-based framework primarily focused on describing the surrounding space in terms of occupancy while also thoroughly estimating the road users' dynamic state, shape, and type.

The proposed framework is composed of four modules, each one distinguished by the type of environment characterization used. Each representation type offers specific advantages and allows the pursuit of different objectives. Consequently, they are employed in a complementary manner, resulting in a more comprehensive description of the environment. Figure 6.1 visually summarizes the main steps of each module and their corresponding environment descriptions.

1. **Point cloud-based module:** The first module is in charge of processing the raw LiDAR data to distinguish obstacle points from the ground surface points. This information is crucial for the subsequent modules to accurately identify obstacles in the scene and model the available free space.
2. **Grid-based module:** This module serves as a central node that fuses, unifies and represents the input data (LiDAR data, V2X data and digital maps) as required by the object-based and

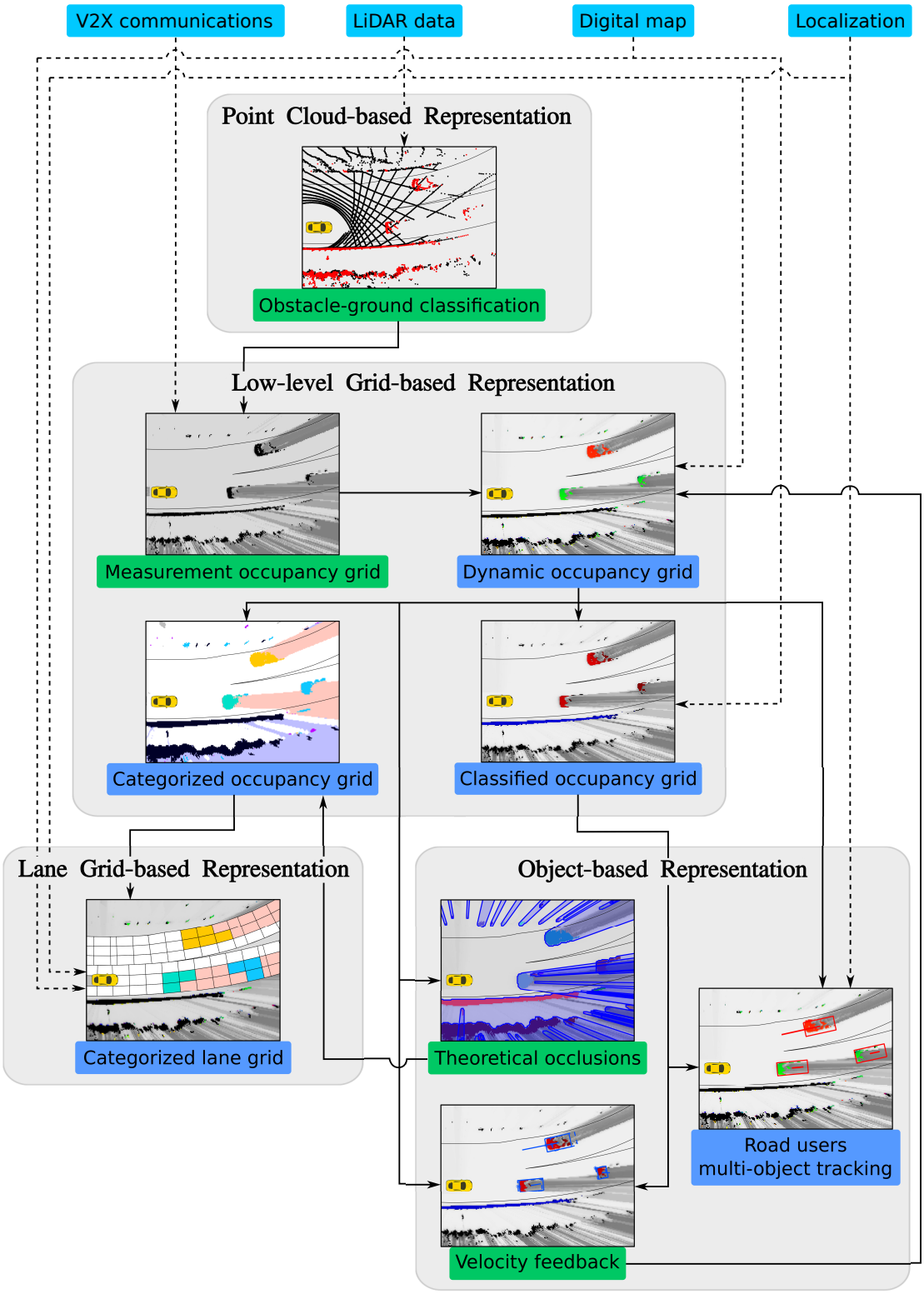


Figure 6.1: Schematic overview of the main environment representations estimated along the different steps of the complete perception framework. Input data is denoted in light blue. Outcome environment representations are denoted in dark blue, while internal use estimations are denoted in green.

lane grid-based modules. In addition, it is intended to provide a reliable and comprehensive low-level representation of all the surrounding environment in terms of occupancy, objects' dynamics and classification. Since it discretizes the space into small cells, it is capable of obtaining a precise delimitation of the boundaries between occupied, free, or unknown areas. Moreover, thanks to its low dependence on object-level assumptions, it is able to represent all types of obstacles.

3. **Object-based module:** It aims at identifying the road users and tracking them over time, thus enabling the use of IDs and data history. In contrast to the grid-based module, it addresses specific known types of objects, so it can take advantage of model assumptions to yield more accurate estimations. Additionally, by further exploiting these assumptions, it is responsible for inferring new input data for the grid-based module regarding the obstacles' velocity and class.

4. **Lane grid-base module:** The lane grid map seeks to provide a more easy-to-understand representation of the estimated driving scene. Specifically, it re-represents the estimation of the grid-based module using semantic labels and focusing only on the road area. The resultant higher level of abstraction facilitates the interpretability of the estimation. In addition to that, the road lane-based representation format makes the data more easily manageable from the perspective of the road layout.

The development of these four modules has led to various contributions. These are summarized hereunder:

- **Obstacle–ground classification:** A novel method for the classification of point clouds into *obstacle* or *ground* points in scenarios with sloped terrain and sparse data has been proposed. In order to achieve this goal, the presented strategy combines a heuristic-based initial classification with a posterior graph-based ground height estimation and point cloud reclassification. The first step leverages on the point cloud distribution to cluster the points into small groups and apply different classification heuristic rules. The second step takes the initial classification as input and models the surrounding ground height with an MRF solved using a message passing algorithm. Then, the point cloud is reclassified comparing the points with the estimated ground height.

- **Input data for the grid map:** The feasibility and use of different sources of information for the grid-based module have been analyzed. On the one hand, a method capable of fusing information from multiple laser beams and obstacles received through V2X communications has been presented. This method associates each occupancy measurement with a confidence value, thus enabling a more robust fusion that favors the most reliable data. On the other hand, in the absence of radar and camera sensors, the use of feedback data from the object-based module to the grid-based module has been explored. In each iteration, data about obstacles' velocity and class is inferred by grouping occupied cells and analyzing them based on object-level assumptions.

- **Classified-Dynamic-Occupancy Grid:** A state-of-the-art DOG has been extended to include semantic information concerning the obstacles' classes. The proposed extension takes advantage of the cell discretization and the PF inherent to the DOG and includes a naive Bayes classifier. Each particle is assigned a set of probabilities corresponding to each possible object class, thereby, the PF is in charge of performing the steps of prediction, hypotheses initialization and removal. The update step is in turn carried out using the naive Bayes classifier at the cell level.
- **Occupancy categorization:** In recent literature, the use of CLG to reduce the gap between perception and decision-making modules has been proposed. In this thesis, this concept is adopted aiming to re-represent the information of the grid-based module in a more comprehensible manner. First, the cells of the grid map are classified by occupancy state and further categorized based on their dynamics, if *occupied*, or their causes, if *unknown*. Then, the obtained CG is translated into a lane grid that fits the shape of the road lanes so the layout of the driving scene is taken into account.
- **Complete perception framework strategy:** A complete perception framework has been introduced, encompassing the processing of raw information to the final outcome. The strategy proposed employs a grid-based module as the central node to unify all incoming data and establish multiple interconnections between all the modules to achieve an overall enhanced outcome. Several minor contributions can be found throughout the developed modules.
- **Occupancy Grid evaluation framework:** Given the lack of a standardized method for OG evaluation, a new evaluation method from an object detection perspective has been proposed. This method measures how easy is to infer objects from the grid by clustering occupied cells and describing them at an object-level. The results are assessed using several metrics based on common clustering issues and object-level features.

Finally, besides progressing beyond the state-of-the-art, this thesis has behind it an important technical work. As introduced in Chapter 1, the present thesis is conducted under the activities of the AUTOPIA research group and the proposed perception framework is designed to be integrated within the architecture of one of its autonomous vehicles prototypes. Appendix B provides a concise overview of the efforts made to achieve this deployment and presents the results obtained during live demonstrations in different research projects.

6.2. Future work

The perception of the environment for autonomous vehicles is a wide field where there are still areas to explore and aspects to improve. Hereunder, some possible lines of future work are presented:

- **Point cloud-based module:** The proposed obstacle–ground classification algorithm has

proven to be successful, but it only considers spatial inter-relations. By incorporating temporal dependencies, a more robust ground height estimation and point classification could be achieved. This objective can be addressed, for example, by storing reduced versions of the previous iterations' point clouds or by including a temporal ground height estimation. Alternatively, there are currently neural network-based approaches capable of classifying not only *obstacle* and *ground* points, but also various other classes, such as *road*, *curb*, *car*, *pedestrian*, etc. This classification would be an interesting input for the COG.

- **Grid-based module:** The grid-based module used in this thesis is an extension of an already well-known approach in the literature. However, there are areas that can be further extended and modifications that can be incorporated. On the one hand, some works of the state-of-the-art have already proposed modifications in order to obtain a static–dynamic classification over time. In this thesis, dynamic cell segmentation is done by thresholding the Mahalanobis distance toward zero velocity, a method that can yield intermittent results. Similarly, instead of representing the height of the obstacles and of the ground using only the information measured at the current iteration, an estimation over time could be included. This would provide a more accurate description of the surrounding environment. Besides, an estimation of the space covered by the road using on-board sensors would be of great interest. Indeed, many algorithms rely nowadays on a perfect knowledge about the road, but digital maps are not always available, updated, or have the required accuracy.
- **Object-based module:** The proposed perception framework relies on the clustering algorithm at different stages. Aiming to obtain a robust performance, the current clustering method analyzes neighboring cells using multiple features and considering the predicted road users at object-level if possible. However, this approach is strongly parameter-dependent. Particularly, in densely populated environments, determining the radius of neighboring cells becomes challenging. Some works in the literature have proposed wise or dynamic parameter definitions, for example depending on the point cloud distribution, or new methods to further exploit the previous iteration information. However, in this regard, approaches using neural networks have shown exceptional results as they can model features that are difficult to explicitly consider in a heuristic way. With respect to the road users object-level tracking, multi-object tracking is a field that has already been extensively explored in the literature. Currently, there are approaches that could be directly integrated into the proposed perception framework and provide enhanced results. For example by including the estimation of more state variables, such as the yaw rate, or addressing the object initialization and matching steps in a more elaborated and robust manner.
- **Lane grid-based module**: The lane grid map aims at providing a more easy-to-understand environment representation. This is a field of research considered to be particularly interesting since it reduces the gap between the perception framework and the subsequent modules. In its present version, the space categorization only takes into account occupancy, dynamics, and

causes as features. However, as exposed in the literature, it is also possible to include other categories based on traffic rules and interactions with other road users. Moreover, including an estimation over time could yield more robust categorizations and allow knowledge about past iterations.

- **Sensors:** This thesis proposes a perception framework based solely on LiDAR data. Nevertheless, it is considered that the integration of more sensors including technologies such as radar and camera can significantly enhance the capacities and accuracy of the proposed strategy. Indeed, many autonomous vehicle prototypes embrace multiple sensor configurations. Moreover, it would also simplify the strategy workflow as, currently, the input data concerning velocity and object classification is obtained from the object-level module.
- **Real-time realization:** All the methods proposed in this thesis have been tested with real-world data, but not all of them have been implemented in real-time. Future work will address this aspect, with particular interest in accelerating the COG and the CLG.

6.3. Scientific dissemination

The work developed under this thesis has led to several publications in several journals and congresses related to automated driving. These and other dissemination activities are listed below.

6.3.1. Journal articles

- **Jiménez, V.**, Godoy, J., Artuñedo, A. and Villagra, J. (2023). "Object-level Semantic and Velocity Feedback for Dynamic Occupancy Grids". *IEEE Transactions on Intelligent Vehicles.* 1–18. doi: 10.1109/TIV.2023.3276578
- **Jiménez, V.**, Godoy, J., Artuñedo, A. and Villagra, J. (2023). "Object-wise comparison of LiDAR occupancy grid scan rendering methods". *Robotics and Autonomous Systems.* 161, 104363. doi: 10.1016/j.robot.2023.104363
- **Jiménez, V.**, Godoy, J., Artuñedo, A. and Villagra, J. (2021). "Ground Segmentation Algorithm for Sloped Terrain and Sparse LiDAR Point Cloud". *IEEE Access.* 9, 132914–132927. doi: 10.1109/access.2021.3115664
- Hortelano, J. L., Villagra, J., Godoy, J. and **Jiménez, V.** (2023). "Recent Developments on Drivable Area Estimation: A Survey and a Functional Analysis". *Sensors.* 23(17), 7633. doi: 10.3390/s23177633
- Godoy, J., **Jiménez, V.**, Artuñedo, A. and Villagra, J. (2021). "A grid-based framework for collective perception in autonomous vehicles". *Sensors.* 21(3), 744. doi: 10.3390/s21030744

6.3.2. Conferences

- **Jiménez, V.**, Godoy, J., Artuñedo, A. and Villagra, J. (2022). "Object-based velocity feedback for dynamic occupancy grids". *IEEE Intelligent Vehicles Symposium (IV).* doi: 10.1109/iv51971.2022.9827201
- Medina-Lee, J., **Jiménez, V.**, Godoy, J., and Villagra, J. (2022). "Maneuver Planner for Automated vehicles on urban scenarios.". *International Conference on Vehicular Electronics and Safety (ICVES).* 1–7. doi: 10.1109/icves56941.2022.9987128

6.3.3. Online & Open Access repositories

- **Jiménez, V.**, Godoy, J., Artuñedo, A. and Villagra, J. "Occupancy Grid Evaluation Framework". `https://git-autopia.car.upm-csic.es/open_source/occupancy_grid_object_detection_evaluation`