LIDAR-Based High Reflective Landmarks (HRL)s For Vehicle Localization in an HD Map

Farouk Ghallabi*,†, Marie-Anne MITTET‡, Ghayath EL-HAJ-SHHADE† and Fawzi Nashashibi*

*INRIA Paris-Rocquencourt, Paris, France (farouk.ghallabi, fawzi.nashashibi)@inria.fr, †Renault s.a.s, Guyancourt, France (ghayath.el-haj-shhaide, marie-anne.n.mittet}@renault.com

Abstract—Accurate localization is very important to ensure performance and safety of autonomous vehicles. In particular, with the appearance of High Definition (HD) sparse geometric road maps, many research works have been focusing on the deployment of accurate localization systems in a previously built map. In this paper, we solve a localization problem by matching road perceptions from a 3D LIDAR sensor with HD map elements. The perception system detects High Reflective Landmarks (HRL) such as: lane markings, road signs and guard rail reflectors (GRR) from a 3D point cloud. A particle filtering algorithm estimates the position of the vehicle by matching observed HRLs with HD map attributes. The proposed approach extends our work in [1] and [2] where a localization system based on lane markings and road signs has been developed. Experiments have been conducted on a highway-like test track using GNSS/INS with RTK corrections as a ground truth (GT). The third-party map † is built with centimetric precision (≈ 5 cm) and it contains high level road features such as lane markings, road signs, road intersections, etc.

I. INTRODUCTION

Recently, the development of High Definition (HD) maps has gained increasing interest in many research and industrial projects. These maps contain key elements for the development of level 4 and 5 of autonomous driving. In particular, one of the most important uses of HD maps is self-vehicle localization. The standard and ready-to-use localization systems rely mainly on global navigational satellite systems (GNSS). However, GNSS-based solutions do not provide better than 2-3 m in open sky environments [4]. A reason why, GNSS receivers are often integrated in a multi-sensor architecture in which a variety of sensors (active or passive) are fused to analyze the surrounding environment. This task is commonly known as: perception.

In the proposed approach, a multi-layer lidar is used to detect High Reflective Landmarks (HRL) that are highly reflective objects that characterize the structure of highway environments. Different sensors’ inputs are used to process lidar data and to perform a map-matching algorithm based on Markov localization (see Fig. 1). As an extension to our previous works [1] and [2] where the use of lane markings and road signs have been investigated, we propose to integrate guard rail reflectors (GRR) as new landmarks in our localization system. To evaluate our approach, experiments were conducted on a highway-like test track. The lidar used for our experiments is a Velodyne VLP-32C. A highly accurate GNSS/INS with RTK correction signals (ixblue: ATLANS-C) was used as ground truth (GT). The third-party map † is built with centimetric precision (≈ 5 cm) and it contains high level road features such as lane markings, road signs, road intersections, etc.

The remainder of this paper is structured as follows: section II is a brief survey on map-based localization methods, section III describes the proposed detection approach for guard rail reflectors. Section IV illustrates the developed map-matching algorithm. Experimental evaluation and discussion are presented in section V and VI and we conclude with perspectives and future work in section VII.

II. RELATED WORK

For map-based localization, state-of-the-art approaches often use two different map structures: dense maps [5], [6] and feature maps [7]. In both cases, the localization algorithm is the process of finding correspondences between data coming from sensors and map attributes. Typical sensors for this are cameras and laser scanners. For camera-based approaches, sparse techniques usually rely on low level features detected in the image like corners, edges and lines [8]. Feature descriptors such as

†The name of the map provider is not mentioned for confidential reasons

![Fig. 1. Global system architecture](image-url)
SIFT [9], SURF [10] and ORB [11] are used to find correspondences by computing similarity measures such as the sum of squared differences (SSDs) or the normalized cross correlation (NCC). On the other hand, other methods use dense techniques to register a query image to multiple synthetic views generated from the map using Normalized Mutual Information (NMI) [12].

For lidar sensors, most of state-of-the-art approaches project 3D point clouds to 2D representations such as 2D orthographic reflectivity grid [13], precise height grid or elevation map [14], [15], or a combination with other information such as colors, curvatures and normals [16]. In [13], laser raw data are accumulated to build a reflectivity grid that is matched by directly comparing grid cells to cells from the map. In [14], a scan-to-map matching is implemented by minimizing an energy function, the adopted grid representation is a combination of height and intensity maps. Despite promising results, the used maps for these approaches highly depend on the sensor used. In contrast, a sparse geometric map which is agnostic to the sensor could be matched to dense lidar grid representation (reflectivity & occupancy grids), as shown in a recent work [6].

In this paper, we adopted a feature-based localization technique. We used a third-party sparse geometric map and a multi-layer laser scanner. However, differently to [6], we have kept sparse representations in all the steps of our system. Thus, a laser point cloud preprocessing step is implemented to extract highly-reflective geometric landmarks of the environment.

### III. Detection of High Reflective Landmarks (HRL)s

![Fig. 2. Projection to a polar grid](image)

In this paper, we propose a detection of high reflective objects from a multi-layer laser scanner. Objects like lane markings, road signs and GRR are very distinguishable by using the lidar data reflectance. In a previous work [1], only lane markings have been studied for vehicle localization. As expected, results have shown good accuracy in the lateral direction but without having significant contribution to the longitudinal direction. An extension using road signs have also been investigated [2]. This time, improvements are obtained in the longitudinal direction but they are not sufficient for autonomous driving scenarios. This is because road signs can be absent for hundred of meters in highways. In this situation, the localization system is only relying on inertial data and GNSS data, thus, the longitudinal error is subject to drift.

![Fig. 3. Detection of guard rail reflectors](image)

Interestingly, on many highway roads, guard rails are usually equipped with reflector markers to indicate the location of guard rails in low light, night or poor conditions (fig. 3.a). In addition, they are repeatable landmarks with a regular spatial distance which make their detection very effective for a localization system. Since our prototype map does not contain guard rail reflectors, we manually constructed an extra map layer by adding the positions of GRR in our test track. Reflector detection and mapping are discussed in the following parts.

#### A. Coordinate Systems Overview

Many coordinate systems are involved in our approach. First, we have the global frame:

$$\mathcal{G} = (x_0, y_0, z_0)$$

It denotes the reference system of the map and for which we have the cartographic transformation to project
from WGS84 coordinates \((\text{latitude}, \text{longitude})\) to UTM\(^2\) coordinates. Then, we have the ego vehicle frame:

\[
\mathcal{V} = (x_{ego}, y_{ego}, z_{ego}).
\]

where the origin is the middle of the rear axle, the x-axis points forward and the y-axis points laterally from right to left (Fig 2). The transformation (rotation+translation) matrix:

\[
T^\mathcal{V}_c = [R|t],
\]

is computed by using inputs from GNSS/INS data with RTK corrections. Finally, two velodyne lidars (VLP-32C) are used. The first is mounted on the top of the vehicle \(L_{top}\) and the second is in the bumper \(L_{bump}\). Again, the transformations:

\[
[T^\mathcal{V}_{top}, T^\mathcal{V}_{bump}]
\]

are computed by performing an extrinsic calibration of the lidars.

### B. Mapping of guard rail reflectors

The mapping stage of GRR was performed manually. Here, we used the bumper velodyne since the sensor setup is very favorable to have more point density in guard rail areas. For each laser scan, reflector points are manually searched and selected. Then, the centroid point \(p_c\) is computed and transformed to the global frame. This is achievable since all data are synchronized to the same clock. Hence, by using the lidar timestamp, the position of the vehicle at the time of the scan can be interpolated from the set of GNSS/RTK positions. The point transformation is therefore given by:

\[
p^\mathcal{V}_c = (T^\mathcal{V}_c \cdot T^\mathcal{V}_{bump}) \times p_c \tag{1}
\]

Consequently, for each side (left and right), a GRR map layer is constructed from the corresponding set of points \(p^\mathcal{V}_c\).

### C. Projection to a Front Polar Grid

In this section, the point cloud of the lidar placed in the bumper is projected onto a polar grid (see Fig. 2). A polar grid is a discretization of the space parameter \(\langle \theta, \phi \rangle\), where \(\phi\) is the azimuth angle and \(\theta\) is the vertical angle. Grid cells \(m_{i,j}\) contain different information from the lidar. For example, in our case, the maximum intensity value, the minimum/maximum ranges and the vector of 3D points that fall within the cell \(m_{i,j}\) are retained.

### D. Detection of guard rail reflectors

Many image representations can be extracted from the polar grid. For instance, since we want to detect high reflective landmarks, a reflectivity image is constructed in the proposed approach. Moreover, our method limits the projection to a polar grid only within an angular field of view \([\theta_{min}, \theta_{max}]\) and \([\phi_{min}, \phi_{max}]\). An example of a front reflectivity image is shown in figure 3.c.

The detection process is implemented by matching an image patch using Normalized Cross Correlation (NCC)\(^1\). The image patch has been carefully chosen by cropping a local window from a clear and visible front reflectivity image 3.b. The result of NCC is a set of candidates (bounding boxes in figure 3.d) that may include false matches. False matches can be detected from other reflective objects such as road markings. As a result, it is mandatory to eliminate false detections and to only keep the most likely candidates. We propose a geometric interpretation of the set of detections in order to eliminate false matches. The 3D points of each bounding box are extracted from the polar grid. Two selection criteria have been set:

1) The flatness of the surface to which the patch belongs.

2) The orientation of the surface to which the patch belongs

A RANSAC plane estimation is applied to 3D points of each match. If more than 60% of the points are inliers then the first criterion is satisfied. Otherwise, the match is rejected. A retained flat surface is now checked according to its normal vector \(n\). If the patch is located in a guard rail area, then the surface normal is nearly parallel to the \(y\)-axis of the ego vehicle. Therefore, the second criterion can be verified by the following condition:

\[
(1 - n \cdot y_{ego}) < \epsilon
\]

where \(\epsilon\) is a close to zero value. In figure 3.d, red bounding boxes are considered as the most likely candidates and the green bounding boxes are false matches.

To extract the GRR from a valid match, a reflectivity threshold is set to identify 3D points belonging to GRR. Then the centroid point is calculated. Figure 3.e illustrates a 3D visualization of the reflectors in the point cloud as red boxes.

### IV. PARTICLE FILTERING

Our localization approach is based on the implementation of a particle filter which is a non-parametric implementation of a recursive Bayes filter. The posterior density distribution is approximated by a number of weighted samples or particles. Suppose that the set of samples at time \(k\) is:

\[
X_k := (x_k^1, w_k^1), ..., (x_k^N, w_k^N) \tag{2}
\]
The state space vector
\[ \mathbf{x} = [x, y, \gamma] \] (3)
consists of a 2D position \( (x, y) \) and a heading \( \gamma \). For our use case, we think that a Constant Velocity and Yaw Rate (CVYR) is sufficient to update the motion model [18]. To do that, an inertial measurement unit is used to feed the motion model with the velocity and yaw rate \([v, w]\). In order to calculate the weight \( w_k \), different measurements are considered: lane markings \( w_{\text{lanes}} \), road signs \( w_{\text{signs}} \), GRR \( w_{\text{reflectors}} \) and GNSS position \( w_{\text{GNSS}} \). For the rest of this paper, the time index \( k \) is omitted for the sake of simplicity.

A. Particle-to-map projection

Each particle \( i \) handles a copy of how observations should look like if the real ego vehicle position was \( x^{[i]} \). This is done by transforming the coordinates of map attributes from the global reference frame to the particle reference frame. In this way, map attributes and observation data can be integrated to update a particle weight.

Map elements \( \mathcal{E} \in \{\text{road markings}, \text{road signs}, \text{guard rail reflectors}\} \) undergo the following transformation:

\[
\mathcal{E}_c^{V_i} = (T^V_i) \times \mathcal{E}_c^G
\] (4)

Where \( \mathcal{C} \) stands for the class of a given map element and \( V_i \) is the reference frame related to a particle \( i \). Applying (4) to road signs and GRRs is straightforward since they are modeled by 3D points. However, this is not the case for road markings, since they are expressed in polar coordinate system \( \{r, \theta\} \) [1]. Thus, the same information should be calculated from map road markings (modeled by polylines). A particle position is projected to the nearest segment of a map polyline as depicted in figure 4. Then, the polar coordinate of the nearest segment with respect to the particle reference frame \( V_i \) is calculated.

B. Update from lane markings

Assume that the set of observed lane markings is:

\[
\Omega^{\text{lanes}} = \{r_1, \theta_1\}, \ldots, \{r_N, \theta_N\}
\] (5)

As previously stated, projecting a particle \( x^{[i]} \) to map road markings gives:

\[
\mathcal{E}_{\text{lanes}}^{V_i} = \{r_1^{V_i}, \theta_1^{V_i}\}, \ldots, \{r_M^{V_i}, \theta_M^{V_i}\}
\] (6)

The adopted likelihood function for road markings is [19]:

\[
p_{k,l} = \frac{-||(r_k - r_l^i)\|^2}{2\sigma^2} + \frac{-||(\theta_k - \theta_l^i)\|^2}{2\sigma^2^2}
\] (7)

Finally, the calculation of \( w_{\text{lanes}}^{[i]} \) is given by the following:

\[
w_{\text{lanes}}^{[i]} = \alpha \sum_{k=1}^{M} \sum_{l=1}^{N} p_{k,l}
\] (8)

where \( \alpha \) is a normalization factor. \( \sigma_r \) and \( \sigma_\theta \) are the lane measurement variances.

C. Update from GRR

Suppose that the set of detected reflector points are:

\[
\Omega^{\text{refl}} = \{x_1, y_1, z_1\}, \ldots, \{x_p, y_p, z_p\}
\] (9)

Similarly, suppose that by applying (4) to map GRR, we obtain:

\[
\mathcal{E}_{\text{refl}}^{V_i} = \{x_1^{V_i}, y_1^{V_i}, z_1^{V_i}\}, \ldots, \{x_N^{V_i}, y_N^{V_i}, z_N^{V_i}\}
\] (10)

An observed reflector \( o_k = (x_k, y_k, z_k) \) is associated to its nearest element in \( \mathcal{E}_{\text{refl}}^{V_i} \):

\[
\langle x_l, y_l, z_l \rangle = \text{arg min} \| e - o_k \| \in \mathcal{E}_{\text{refl}}^{V_i}
\] (11)

Then a likelihood function is calculated:

\[
p_{k,l} = \frac{-||(x_k - x_l)\|^2}{2\sigma^2} + \frac{-||(y_k - y_l)\|^2}{2\sigma^2^2}
\] (12)

Finally, the calculation of \( w_{\text{reflectors}}^{[i]} \) is given by the following:

\[
w_{\text{reflectors}}^{[i]} = \beta \sum_{k=1}^{M} \sum_{l=1}^{N} p_{k,l}
\] (13)

where \( \beta \) is a normalization factor. \( \sigma_x \) and \( \sigma_y \) are the reflector measurements variances.
D. Update from GNSS position

The idea behind using GNSS position is to define an upper limit of the filter estimate. In many approaches such as in [19], particle positions are compared to a confidence interval of GNSS data by adopting a euclidean distance. The problem of this approach is that both lateral and longitudinal directions are affected. In our approach we want to act only on the longitudinal direction as we already have a robust and reliable lateral localization (road markings) [1]. To do that, we implemented the approach proposed in [2]. The idea behind this approach is to make the lateral direction unaffected by changing the coordinate system from the absolute reference frame to the Frenet reference frame linked to the map. This change of system of coordinates is ensured by projecting the particles onto map lane markings (polylines) and obtaining as a consequence what is called map tracker points (MTPs). Then, instead of comparing euclidean distances (particles to GNSS position), the comparison is done by calculating a distance \( D_{\text{tracker}} \) between GNSS MTPs and particle MTPs. Finally, the calculation of \( w_{\text{GNSS}} \) is given by:

\[
\begin{align*}
  w_{\text{GNSS}}^{[i]} &= 1, & & \text{if } D_{\text{tracker}}^{[i]} \leq \rho_{\text{max}} \\
  w_{\text{GNSS}}^{[i]} &= 0, & & \text{otherwise}
\end{align*}
\]

Where \( \rho_{\text{max}} \) is a confidence value of the GNSS receiver [19].

E. Update from road signs

The calculation of \( w_{\text{signs}} \) is not different from that of \( w_{\text{reflector}} \) except that we only used \( x \)-coordinates in the formula (12) as follows:

\[
p_{k,l} = e^{-\frac{(x_{V_i} - x_l)^2}{2\sigma_x^2}}
\]

and

\[
w_{\text{signs}} = \beta \sum_{k} \sum_{l} p_{k,l}
\]

V. LOCALIZATION ACCURACY

Fig. 6. Renault test track (google earth)

Experiments were conducted on a highway-like test track of 5 km long (Fig 6). Our prototype vehicle «Melo» is equipped with an automotive-grade GNSS receiver u-blox, a highly accurate GNSS/IMU with RTK correction signals (ATLANS-C) and five velodynes VLP-32C lidars (four on top and one in the bumper). In our experiments, only two lidars were used: the bumper lidar and one roof lidar. The roof lidar is mainly used to detect road signs. The bumper lidar barely detects road signs, but it does detect road markings and GRR. All the sensors are time synchronized to the GNSS clock and data were recorded in ROS bag files.

A. Evaluation metrics

Error evaluations are given in terms of cross track (CT) and along track errors (AT) [3] and in terms of absolute errors as well. CT and AT errors are very relevant to map-based navigation. Absolute errors give an overview
on the overall error in the absolute frame. Details about the computation of CT and AT errors are explained in [2].

B. Accuracy results

We tested our system at two different vehicle speeds: 70 and 90 Kph. The use-cases studied here are related to the availability of high reflective landmarks. Since road markings are always present, we studied the availability/absence of road signs and GRR. A first case is the accuracy when only road signs are used. The second is when only GRR are used and finally we tested when both landmarks are available.

The time variation of the AT and CT errors are illustrated in figure 7. Tables I and II give a summary of the mean and standard deviation errors for AT and CT. An additional information about the absolute error is also given.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>ERROR EVALUATION AT 70 KPH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Along track error (m)</td>
</tr>
<tr>
<td>Road signs Only</td>
<td>mean: 0.74</td>
</tr>
<tr>
<td>Reflectors Only</td>
<td>mean: 0.28</td>
</tr>
<tr>
<td>Road signs + reflectors</td>
<td>mean: 0.32</td>
</tr>
<tr>
<td>GNSS Only</td>
<td>mean: 1.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>ERROR EVALUATION AT 90 KPH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Along track error (m)</td>
</tr>
<tr>
<td>Road signs Only</td>
<td>mean: 0.52</td>
</tr>
<tr>
<td>Reflectors Only</td>
<td>mean: 0.31</td>
</tr>
<tr>
<td>Road signs + reflectors</td>
<td>mean: 0.3</td>
</tr>
<tr>
<td>GNSS Only</td>
<td>mean: 2.67</td>
</tr>
</tbody>
</table>

VI. DISCUSSION

In the absence of road signs and GRR, the along-track error is bounded by GNSS data. The obtained error depends on the velocity: 1.27 m at 70 Km/h versus 2.67 m at 90 Km/h. However, this accuracy could not always be guaranteed if the GNSS data is very erroneous (signal loss for example). The integration of GRR improves significantly the AT error: from 1.04 m to 0.28 m at 70 Km/h. A slight improvement is also obtained in the CT error. On the other hand, when integrating both road signs and GRR, the obtained accuracy is very similar to GRR only. Combining perception data with map elements can deal with the case of false detections or occlusions. Indeed, according to the equation (13) only a detected reflector that is metrically close to a map reflector has a significant contribution to the calculation of \( w_{\text{reflectors}} \).

In addition to CT and AT errors, in table I and II, the absolute error is given to allow the comparison with state-of-the-art approaches. For instance, our method outperforms the accuracy achieved in [7] where the achieved absolute error is 0.93 m whereas the achieved absolute error in our approach is 0.60 m. However, the accuracy presented in [6] (overall error is 0.28 m) is better than ours. Notice that this comparison is not very rigorous since the driving conditions and the use-cases are not necessarily the same. Indeed, this paper tackles the problem of high-speed scenarios: up to 90-110 Km/h.

The shown results proved the robustness of a feature-based localization system on a third-party geometric map. This is important as the map-matching process is agnostic against the type of the sensor used. Indeed, by processing raw lidar data to extract high level objects such as road markings and road signs, many other sensors (such as a camera) can easily be integrated and fused in order to increase the robustness and redundancy of landmarks detection.

The use of high definition geometric map is also a contribution of our paper. Generally, dense maps have proven to be more accurate for localization systems. Real time execution is a challenge in this case. Sparse maps, as in our case, are more time efficient and agnostic against the type of sensor which proves the usefulness of these maps.

VII. CONCLUSION AND FUTURE WORK

In this paper, we presented a LiDAR-based localization approach by detecting high reflective landmarks (HRLs). Detected landmarks in this paper are: road markings, road signs and guard rail reflectors. Our system is able to achieve high accurate localization outputs. The obtained results are approximately 0.3 m in the along direction and 0.19 m in the cross direction. To ensure a real-time capable system, we showed that the use of a sparse geometry third-party map is sufficient for a good localization outputs. Further work will focus on more challenging use cases (heavy traffic) and adverse weather conditions.

REFERENCES

Fig. 7. Error evaluations at 70 & 90 Kph


[13] J. Levinson and S. Thrun, “Robust vehicle localization in urban...


