16 channels Velodyne versus planar LiDARs based perception system for Large Scale 2D-SLAM

Nobili S.1 Dominguez S.2 Garcia G.3 Philippe M.4

Abstract—The ability of self-localization is a basic requirement for an autonomous vehicle, and a prior reconstruction of the environment is usually needed. This paper analyses the performances of two typical hardware architectures that we evaluate in our 2D Simultaneous Localization and Mapping (2D-SLAM) system for large scale scenarios. In particular, the selected configurations are supposed to guarantee the possibility of integrating at a later stage mobile objects tracking capabilities without modifying the hardware architecture. The choice of the perception system plays a vital role for building a reliable and simple architecture for SLAM. Therefore we analyse two common configurations: one based on three planar LiDARs Sick LMS151 and the other based on a Velodyne 3D LiDAR VLP-16. For each of the architectures we identify advantages and drawbacks related to system installation, calibration complexity and robustness, quantifying their respective accuracy for localization purposes. The conclusions obtained tip the balance to the side of using a Velodyne-like sensor facilitating the process of hardware implementation, keeping a lower cost and without compromising the accuracy of the localization. From the point of view of perception, additional advantages arise from the fact of having 3D information available on the system for other purposes.

I. INTRODUCTION

An efficient and accurate solution to the Simultaneous Localization and Mapping (SLAM) problem is the basic building block for an autonomously navigating platform. The perception system employed in the architecture plays a fundamental role in determining the quality of the performance. Indeed, the reliability of the map and consequently the accuracy of the localization are highly dependent on the measurements provided by the local sensors. Nowadays there are several options when choosing the hardware architecture that allows us to apply SLAM for positioning our vehicle. The question arises when we have to choose which hardware set-up is the most appropriate for our application given some constraints about budget, ease of installation, precision, reliability against changing environment, versatility, etc. Most commonly used architectures employ Light Detection and Ranging (LiDAR) technology measuring at different angles the distance to the surrounding environment. The SLAM problem is well-known for its increasing complexity in terms of accuracy, runtime and computational resources required while covering long outdoor distances and mapping. We regard these aspects as particularly relevant for precise localization of memory-restricted systems and we propose a multi-map LiDAR-based 2D-SLAM solution. We implemented an extended version of the GMapping Robot Operating System (ROS) package. In particular, we adapted it and integrated it into our architecture for being used in Large Scale multi-map 2D SLAM. Our version allows to start building a new map when required, saving previously the map under construction. A map-manager is in charge of deciding when a map must be stored and start building a new one. Later, during the localization phase, the sub-maps are loaded as they are required along the pre-recorded journey and Monte Carlo localization techniques using a probabilistic particle filter are applied to find the most likely position given the map, laser scan and odometry measurements. Connection points connect a sub-map with its neighbour and delimit when a sub-map ends and a new one starts. The local reference frame of a sub-map is normally positioned on a connection point (See figure 1).

Fig. 1. We represent the vehicle’s path (from right to left) in a chain of sub-maps. Each submap is connected to the previous and the next one through connection points.

In this paper, we present experimental results obtained in a urban context using two distinct laser-based hardware architectures typically used in SLAM but the same software for localization. In Table I, we summarize the main features of both types of sensors. Our main contribution focuses on a comparison between the performances of the two systems in terms of map quality, localization accuracy and robustness to temporarily static elements like parked vehicles.

The first experimental platform is an electric car Renault Zoe ZE equipped with three LiDARs (SICK LMS 151) placed at 50 cm from the ground level in a configuration

---

1 Nobili S., Dominguez S. and Philippe M. are with IRCCyN, Institut de Recherche en Communication et Cybernétique de Nantes, École Centrale de Nantes, 1 rue de la Noé, 44321 Nantes, France
2 salvador.dominguez@irccyn.ec-nantes.fr
3 Garcia G. is with ECN École Centrale de Nantes, 1 rue de la Noé, 44321 Nantes, France
1 simona.nobili@eleves.ec-nantes.fr
2 salvador.dominguez@irccyn.ec-nantes.fr
3 philippe.martinet@irccyn.ec-nantes.fr
that guarantees a 360° Field of View (FoV) about the car vertical axis. Specifically, two of the LiDARs are mounted on the two front corners of the car in order to cover straight and side views, whereas the third one covers the back side view (See figure 2 and 4). An extrinsic calibration process is performed to ensure that all the three LiDARs lie on a plane as closed as possible to a common horizontal plane such that the scans can be merged and given as input measurement to the SLAM solver.

Fig. 2. The LiDARs Sick LMS151 are installed in strategical positions to ensure a 360°FoV around the vehicle.

The second experimental platform is a vehicle Renault Fluence equipped with a 16 channels Velodyne LiDAR PUCK (VLP-16) placed some centimeters above the roof surface and scanning 360° about the car’s vertical axis (see Figure 3). This sensor is a 360° revolving-head 3D LiDAR with 16 laser beams vertically separated along a range of 30° with 2° of angular resolution. The 7th laser, after a full rotation, sweeps a cone in the 3D space. In this context, we take advantage of the vertical FoV (from -15° to 15° with respect to the sensor reference frame) to infer a 2D laser-scan information merging data belonging to a vertical range between 1.8 and 2.8 meters from the ground. In this way the scan measurements will not be influenced by the most common moving elements, which in an urban context are assumed to be cars, people or small objects. This height also provides robustness to slopes and defects of the road as the laser plane is less likely to intersect the ground. Moreover, note that the decision of working in two dimensions is justified by one main applicability reason. From a practical point of view, 2D information are sufficient for self-localization on a local flat map, as the car moves locally in two dimensions, and are manageable in the general case of restricted availability of computational resources.

TABLE I
SENSORS’ FEATURES

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Field View</th>
<th>Max Range</th>
<th>Layers</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sick LMS151</td>
<td>270°</td>
<td>50 m</td>
<td>1</td>
<td>3000 $</td>
</tr>
<tr>
<td>Velodyne VLP-16</td>
<td>360°</td>
<td>&gt; 100 m</td>
<td>16</td>
<td>8000 $</td>
</tr>
</tbody>
</table>

The remainder of this paper is organized as following. In the next section, we present some of the most effective perception systems currently employed for SLAM applications, along with some relevant state-of-the-art methods for large-scale SLAM. In section III we explain how we obtain a 360° planar laser scan with a certain angular resolution where the measurements are relative to the vehicle’s reference frame in both cases. Additionally, we explain how we generate the ground truth, as well as, how we perform the comparison between computed position and ground truth position. In section IV we present the experiments performed on this study and their purpose. In section V we present the results of the experiments performed giving some partial conclusions. And finally, in section VI, we summarize the main points of our results.

II. RELATED WORK

In the past, highly effective SLAM techniques have been developed and state-of-the-art SLAM solvers are now able to achieve good performances in terms of accuracy and real-time processing (e.g. GMapping [1] and Hector SLAM [2]). The first implementations of SLAM methods were based on combined motion control and features observations with an Extended Kalman Filter (EKF) [3], [4]. However, as reported in [5], the solution to the EKF-based SLAM is consistent in the linear-Gaussian case but diverges in general situations. Subsequently, Rao-Blackwellized particle filters have been introduced as effective SLAM problem solvers under conditions of non-linearity. The approach proposed in [1] and [6] uses a particle filter in which each particle carries an individual map of the environment and treats the consequent requirement of reducing the number of particles. This algorithm is open source for the community of researchers under the name of GMapping and is currently employed for many SLAM-based applications. However, the problem of computational complexity over large-scale environments, of the order of tens of kilometers, has not been directly addressed in this work.

Closely related to the solution we propose are hierarchical SLAM methods to deal with large-scale applications. Atlas [7] is a framework which builds a two-levels hierarchy combining a local Kalman Filter with global graph optimization. Similarly, Hierarchical SLAM [8] is a mapping technique which uses statistically independent local maps interconnected at an upper level through an
adjacency graph. Subsequent proposals employ independent local maps with robot-centred representation [9], local metric maps associated with topological places in a topological map [10], submapping methods in which the current submap summarizes all the information [11], local maps with feature positions and final robot pose associated with a global map without robot poses [12]. In [13] the authors present a SLAM technique requiring a small memory footprint. This feature makes this solution particularly suited for large-scale problems. In the case of this paper, we present an adaptation of the GMapping framework to deal with computational complexity problems while covering long distances.

Of fundamental importance for a SLAM-based architecture is the perception system. While quality and accuracy of the sensors are basic requirements for the reliability of the measurements, the costs in terms of system installation effort and finances have to be taken into account at conception time. During the DARPA Urban Grand Challenge in 2007 [14], fully equipped autonomous vehicles performed complex SLAM-based tasks to compete in a 96km course in a urban area. Boss, by the Carnegie Mellon University and General Motors Corporation team [15], took advantage of three 3D LiDARs (IBE0 Alasca XT) mounted at about 50cm from the ground assuming to cover a relatively flat road surface, processing each layer independently and assuming infrequent changes in the static environment. Junior, by the Stanford University team [16], used a Velodyne HDL-64 mounted on the car’s roof surface to perform 2D mapping and localization. Odin, by the Virginia Tech team [17], used the coupled information provided by three LiDARs (two IBE0 Alasca XT and a IBE0 Alasca A0) mounted at about 50cm from the ground level. In the last few years Google developed hardware and software architectures for a self-driving car [18]. The heart of the system is a Velodyne HDL-64 generating a detailed 3D map of the environment which requires high processing capabilities.

In this context, the aim of our work is to provide a comparison of the results obtained for the localization in an unknown environment using two differently equipped vehicles. In particular, we analyse the performances of our solution to a large-scale SLAM problem in the case of a (three) 2D LiDARs based versus a 16 planes Velodyne LiDAR based perception architecture.

III. METHODOLOGY

The required inputs to the SLAM system are odometry and planar laser scans, providing the information about the car’s motion and the surrounding environment respectively. Specifically, we generate 360° FoV scans with 0.5° of angular resolution @ 10 Hz from each of the laser-based perception architecture and odometry @ 100Hz from the data collected by the OBD-II connectors and IMUs integrated in the cars. For the sake of simplicity, we consider both odometry and scans data with respect to the car reference frame, placed at the center of the rear axis.

In order to provide an unbiased comparison, we run the same SLAM system on both our vehicles. However, given two distinct hardware architectures, the approaches adopted to generate the scans and their quality vary. In the following, we present the two different techniques developed to generate laser scan information, first from the three Sick LiDARs and then from the VLP-16.

A. Sick LiDARs

In the case of Zoe (Figure 4), we convert the individual scan measurements into the 360° output scan by knowing the pose of each of the sensors with respect to the car reference frame. In particular, we determine the exact pose of each sensor using a process of extrinsic calibration, fitting the overlapping parts of the individual scans and then we convert each of the measurements from the sensor’s local frame to the car’s frame by simple reference frame transformation (Equation 1).

\[ T_{\text{scan}} = T_{\text{sensor}} \ast T_{\text{point}} \]  \hspace{1cm} (1)

where \( T_{\text{point}} \) is the transformation matrix of a point with respect to the sensor’s frame, \( T_{\text{sensor}} \) is the fix transformation of the sensor with respect to the car’s reference frame and \( T_{\text{scan}} \) is the transformation of the measured point with respect to the car’s reference frame which is composed by a translation vector \( \{x, y, z\} \) and a rotation matrix. From the transformed position we can extract the angular position \( \alpha \) with the expression (2)

\[ \alpha = \arctan \frac{x}{-y} \]  \hspace{1cm} (2)

Depending on the scan’s angular resolution, for a given \( \alpha \), a unique corresponding index in the output scan vector is given by (Equation 3).

\[ i(\alpha) = \text{round}(\alpha / \Delta\alpha) \]  \hspace{1cm} (3)

where \( \Delta\alpha \) is the chosen angular resolution. In our case \( \Delta\alpha \) is 0.5° and \( \alpha \in \{0, 360\} \).

On the other hand, we obtain the output range as

\[ \text{range}_i = \sqrt{x^2 + y^2} \]  \hspace{1cm} (4)

Notice that, for the same \( \alpha \) there can be more than one measurement. In this case, we select the nearer point, i.e. the one with smaller range.

B. 16 Planes Velodyne

In the case of Fluence (Figure 5), the VLP-16 is placed above the roof surface. In particular, we compute the exact position of the VLP-16 as 1.457 meters from the car’s rear axis and 1.565 from the ground after a process of extrinsic calibration. Similar to the three Sick LiDARs’ case, we convert each of the measurements from the sensor’s local frame to the car’s frame by simple reference frame transformation (Equation 1). Notice that in this case the raw data provided by the VLP-16 correspond to 3D measurements. Therefore,
so as to obtain a planar scan information, we project the points belonging to the vertical range \{1.8, 2.8\} meters from the ground to a plane passing through the VLP-16 reference frame center and perpendicular to its vertical axis. Once this transformation is performed, the scan data type can be identified by an index and a corresponding range value as explained in the case of Zoe (Equation 3 and 4).

C. Comparison with ground truth

The ground truth is generated using a Proflex 800 RTK-GPS receiver that applies the differential corrections obtained from a DGPS station located on the roof of IRCCyN building. The measurements provided by the RTK-GPS have an error of less than 1 cm in position when in Fixed mode. We estimate the orientation (heading), by computing the direction of the movement.

We compare the position obtained by the SLAM system at time \( t \) with the ground truth interpolated to that time. The interpolation method used is through splines of position coordinates with the time as independent variable.

IV. EXPERIMENTAL SET-UP

A couple of experiments have been performed to quantify the localization accuracy and robustness of both the systems. The precision of each of the results is evaluated comparing the position of the car with a ground-truth generated using the measurements provided by the high precision RTK-GPS on the points where it is available. During both the experiments, the vehicles travel in convoy, that is, one following the other, in order to ensure the same environmental conditions.

A. First Experimental Set-up

For the first experiment, the vehicles travel for about 3 kilometers. The chosen trajectory covers the area around the campus of École Centrale de Nantes, which is a standard urban sector. Using the data recorded by both the vehicles, we perform the task of map building. Later Monte Carlo localization in these maps is performed. The aim of this experiment is to compare the positioning accuracy of both systems under the same conditions.

B. Second Experimental Set-up

A second experiment has been conducted in a highly changing environment such as the parking area of the campus of École Centrale de Nantes, where the position and number of parked vehicles constantly changes. We performed recordings over two different days in order to ensure changes in the temporarily-static elements (e.g. parked vehicles) met along the path. In particular, the first day we recorded in the early morning (when the parking was almost empty) and the second day during the day-time (when the parking was crowded). In this context, on the first day the vehicles are asked to perform a task of localization and mapping. Subsequently, on a second day, the vehicles have to localize themselves along the same trajectory but with the maps previously built. The aim of this experiment is to check the robustness of each set-up against environmental changes.

V. RESULTS

A. First Experiment. Same trajectory, same environment

In figure 6 we show the accumulative and differential histogram of the error resulting from localization under the conditions detailed in Section IV-A, for both the cars. As we can see the precision of Zoe’s localization (equipped with planar LiDARs) is slightly better than Fluence’s localization (equipped with VLP-16). For Zoe 95% of the measurements have less than 0.6 meters of error while for Fluence 95% have less than 0.7 meters. In this case, Zoe performs better because, since the LiDARS detections originally belong to a plane, the measurements are more stable and less noisy. Instead, the fact that each VLP-16 scan...
results from the projection of a collection of points onto a plane, causes the map not to be overall as well shaped as in the case of the planar LiDARs.

B. Second Experiment. Same trajectory, highly changing environment

Figure 7 shows part of the map built on the same area by both systems. In the left image, the map built by the VLP-16-based system doesn’t show short objects like the cars parked, but the walls of the surrounding buildings and trees. In the right one, the map build by Sick LMS151-based system show all the objects that can be seen at a height of 50 centimeter from the floor level. Just by looking at both maps we can expect significant differences on the analysis of the positioning accuracy.

Figure 8 shows the accumulative and differential positioning error histograms for the case of the planar LiDAR-based system. As expected the precision obtained for the second day using the maps of the first day is worst than in the case of the first day. However, an acceptable precision even under highly changing conditions was not an obvious result. In the first day 95% of the measurements lie under 0.4 meters while in the second that percentage moves to 0.8 meters. In this case we can assess that the common environmental changes affect the precision but still more than 97% of the measurements lie under 1 meter of error.

VI. CONCLUSIONS

In this paper we perform a quantitative comparison, using our 2D-SLAM localization system for large scale scenarios,
of two typical LiDAR-based hardware configurations: one based on several LiDARs strategically installed around the car at a low height, and the other one based on a single 3D-LiDAR installed on the roof of the car. Both systems generate a single 360° scan centered on the car’s reference frame which is used for map-building and/or map-localization. The qualitative results from studies conducted over our datasets are summarized in Table II. In the same conditions and with an static environment, the set-up using planar LiDARs performs slightly better than with the 3D sensor as the scans are less noisy on the range dimension. The positioning error obtained is, in general, 10 cm. smaller in the first case. The reason for being more noisy in the 3D sensor’s case is because we are projecting onto the same plane a range of heights (1 meter of range in our case), so that, the probability of variability is higher than in the case of one single height. Another conclusion is that the maps built with VLP-16 data provide more robustness against common medium term environment changes, like cars parked on the sides of the road, describing mostly the static details, like the walls of the buildings, trees or urban structures. This is a desirable feature on map localization. Furthermore, there are other advantages of using a single sensor regarding the simplicity of installation, extrinsic calibration and general set-up, speeding up considerably the set-up process. Regarding the cost-effectiveness, the newly released sensor VLP-16 is affordable for most of research projects in comparison with using 3 outdoors planar LiDARs. Finally, in the case of the 3D-sensor the fact of having 3D information available on the system opens a wide range of possibilities from the point of view of the perception.

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>COMPARISON TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy same environment</td>
<td>Planar LiDARs</td>
</tr>
<tr>
<td>Robustness env. changes</td>
<td>↑</td>
</tr>
<tr>
<td>Ease installation</td>
<td>↓</td>
</tr>
<tr>
<td>Ease extrinsic calibration</td>
<td>↓</td>
</tr>
<tr>
<td>Total cost</td>
<td>↑</td>
</tr>
<tr>
<td>Provides 3D data</td>
<td>No</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENT

This work has been funded by the European Fund in Regional Development FEDER-ROBOTEX trough the Pays de la Loire regional development from June 2013 to October 2015. Thanks also to the ICARS team of IRCCyN.

REFERENCES


