Intelligent feature selection method for accurate laser-based mapping and localisation in self-driving cars

Computer Engineering Department
University of Alcalá, Madrid (Spain)
Email: noelia.hernandez@uah.es, ivan.garciad@uah.es, carlota.salinasmaldo@uah.es

Abstract—Robust 3D mapping has become an attractive field of research with direct application in the booming domain of self-driving cars. In this paper, we propose a new method for feature selection in laser-based point clouds with a view to achieving robust and accurate 3D mapping. The proposed method follows a double stage approach to map building. In a first stage, the method compensates the point cloud distortion using a rough estimation of the 3-DOF vehicle motion, given that range measurements are received at different times during continuous LIDAR motion. In a second stage, the 6-DOF motion is accurately estimated and the point cloud is registered using a combination of distinctive point cloud features. We show and analyse the results obtained after testing the proposed method with a dataset collected in our own experiments on the Campus of the University of Alcalá (Spain) using the DRIVERTIVE vehicle equipped with a Velodyne-32 sensor. In addition, we evaluate the robustness and accuracy of the method for laser-based localisation in a self-driving application.

I. INTRODUCTION AND RELATED WORK

The booming field of self-driving cars has ushered in a new period of development in a number of scientific areas, being map building one of the most outstanding ones. A great deal of automotive companies are putting significant amounts of effort on building accurate 2D maps for automated driving purpose. Those maps contain information regarding the geo-referenced position and geometrical configuration of elements such as intersections, lane markers, road signs, road signals, etc. However, in order to achieve accurate localisation, self-driving cars need not only 2D maps but also 3D maps providing a distinct representation of the environment. Although some researchers have demonstrated the feasibility of using vision-only features for accurate localisation, such as the Daimler-KIT group did in the BERTHA route in Germany [1], 3D maps for accurate localisation are usually built using point clouds obtained with laser sensors. This is the case of Waymo [2] (formerly Google) self-driving cars, which have been performing automated driving missions in the Mountain View area in California for almost one decade already. A similar laser-based localisation approach has been followed by many researcher groups in the area of automated driving, such as the National Seoul University [3] or the SMART program (Singapore-MIT Alliance for Research and Technology) [4].

When it comes to automated driving, there is currently a debate in the scientific community regarding the trade-off between local perception capability and map dependence. It seems that further effort on the first topic is definitely needed, since self-driving cars have to make progress in their capability to better understand the world they see, very much in an attempt to mimic human driving style. However, map-based localisation is still a crucial, and necessary element in today’s automated driving systems. In this regard, the use of laser for accurate localisation provides a number of advantages with respect to other sensors, such as vision. It is well known that vision-based mapping and localisation is prone to failure at night-time, in low visibility conditions or under adverse weather conditions. Change of appearance is another relevant problem. For example, a road diversion can provoke a failure in a vision-based localisation system if the new features that the car is finding as it moves have not been previously stored in the system. A similar problem can occur during the fall season, when the leaves of trees can fall down and derive in a situation of strong change of appearance with respect to the time when the map was built (if it was built during the spring or summer time).

However, the use of LIDARs for mapping and localisation does not come without difficulties. Thus, motion estimation via moving LIDARs involves motion distortion in point clouds, as range measurements are received at different times...
during continuous LIDAR motion. Hence, the motion often has to be solved using a large number of variables and a computationally heavy optimization algorithm [5]. Scan matching also fails in degenerate scenes, such as those dominated by planar areas. Similarly, the localisation method can fail in situations in which a repetitive pattern is encountered, e.g. a park with symmetrically located trees, or in cases in which the amount of moving objects is largely predominant over the number of features provided by static elements. These difficulties require a further effort from the scientific community in order to develop really robust and fully operational laser-based mapping and localisation techniques for self-driving cars. In this line, Rohde [6] proposes a localisation method specifically designed to handle inconsistencies between map material and sensor measurements. This is achieved by means of a robust map matching procedure based on the Fourier-Mellin transformation (FMT) for global vehicle pose estimation. Consistency checks are then implemented for localisation integrity monitoring, leading to significantly increased pose estimation accuracy. Other approaches [7] segment moving parts in the sequence of point clouds, being capable of distinguishing rigid motions in dense point clouds and cluster those by applying segmentation schemes, such as graph-cuts into the ICP (Iterative Closest Point) algorithm. Consequently, the clustering process aims at segmenting moving vehicles out from the point cloud in an attempt to improve the accuracy of the laser-based localisation scheme. In [8], a vision and laser-based odometry system is presented in which an intelligent pre-selection of point-cloud features is carried out, using an appropriate distribution of edges and planes, with a view to increase the accuracy of the map while removing outliers.

In this paper, we propose a new method for feature selection in laser-based point clouds with a view to achieving robust and accurate 3D mapping. Pre-selection of features is absolutely necessary in order to achieve accurate map making capacity. Otherwise, the map incorporates artifacts that derive in loss of accuracy and, consequently, lack of localisation precision. The proposed method follows a double stage approach in which an intelligent pre-selection of point-cloud features is carried out, using an appropriate distribution of edges and planes, with a view to increase the accuracy of the map while removing outliers.

The rest of the paper is organized as follows. Section II provides a description of the mapping algorithm. In section III, a revision of the localisation method is carried out. Section IV presents and discusses the experimental results attained with the DRIVERTIVE automated car. Finally, section V analyzes the main conclusions and future work.

II. MAPPING

Outdoor scenes are characterized by being composed of objects which are placed within a wide range of distances. While mapping, small angular errors in the order of milliradians could lead to 1 meter errors for an object located at 50 meters. In addition, the point cloud is distorted by the vehicle egomotion during the acquisition. In order to achieve robust and accurate 3D mapping, a correction procedure of the 3D point clouds deformation is mandatory previous to the map creation.

A. 3-DOF sweep correction

Our experimental platform, DRIVERTIVE, consists on a commercial Citroën C4 modified for automated driving. DRIVERTIVE GPS-based localisation combines the information from an RTK-GPS, CAN bus and a low-cost IMU (Inertial Measurement Unit) in an 3-DOF EKF (Extended Kalman Filter) as explained in [9]. A Velodyne-32 sensor was attached approximately 50 cm over the vehicle’s roof to perform the mapping and localisation tasks based on LIDAR odometry. The final purpose is to provide an accurate and robust localisation not based on RTK-GPS which suffers from undesired blackouts due to urban canyons, tunnels, trees, etc.

Our Velodyne-32 delivers approximately 10 sweeps of 360 degrees per second. This means that, at normal driving speeds, the point cloud sustains considerable deformation. To correct this deformation, the angular and linear velocities provided by the EKF are considered constant between two consecutive Velodyne-32 firings. The motion undergone by the Velodyne-32 between two consecutive firings is compensated to create a single pose reference for a 360 degree sweep (Fig. 2). It is worth noticing that, as roll and pitch angles are not taken into account, this initial estimation introduces errors, specially during sharp turns such as roundabouts and speed bumps.

This sweep correction is used as an initial rough guess for the next stage.
B. 6-DOF LIDAR odometry

Using the 3-DOF initial guess, a registration technique will estimate the 6-DOF motion undergone by the vehicle to create a final corrected point cloud. This cloud will be used as input for an octomap-based mapping technique. For the registration process, distinctive point cloud features are extracted and selected using K-means and RANSAC. Then, an ICP will estimate the 6-DOF transformation on the features that will be used for the correction of the point clouds. Finally, the corrected point clouds will be used to create a 3D octomap-based representation of the environment.

Fig. 3 shows a block diagram of the algorithm.

![Block diagram of the LIDAR odometry](image)

Fig. 3. Block diagram of the LIDAR odometry.

1) Linear descriptors extraction: Let us define \( C_{k}^{sc} = \{p_{1}^{sc}, p_{2}^{sc}, \ldots, p_{N}^{sc}\} \) as the corrected point cloud using the 3-DOF sweep correction at time \( k \) composed of \( N \) world referenced 3D points \( p_{i}^{sc} = \{x_{i}, y_{i}, z_{i}\} \).

Initially, a filtering process is applied to \( C_{k}^{sc} \) to remove ground points. This will help the clustering process in the next step. Next, an iterative adaptive k-means algorithm is applied to the filtered cloud to extract cluster candidates for the following step. The L2 Euclidean distance used on each k-means iteration is adapted to account for the sparseness of far objects. In a similar process to [10], the best fitting linear descriptor is computed using RANSAC for each one of the clusters. Only descriptors with a fitting value above a predefined threshold and with an orientation and size corresponding to a vertical pole are selected. Finally, these descriptors will be used in an ICP to estimate the 6-DOF transformation undergone by the descriptors.

2) 6-DOF LIDAR odometry estimation: An ICP procedure will estimate the 6-DOF transformation on the descriptors based on the extracted vertical descriptors. Two sets of descriptors, new ones and tracked ones, are needed for the ICP computation. The new set is composed of the vertical descriptors detected at the current point cloud. The tracked set is composed of the vertical descriptors tracked or matched in the previous ICP iteration. Once the ICP transformation is computed, the vertical descriptors are matched in a brute force search analyzing the euclidean distance between them.

Those closer to a predefined threshold will be added to the tracked set.

The ICP obtains the linear and angular transformation needed to reach the minimum matching error. This transformation is only applied when two or more descriptors are available in both the new and the tracked descriptors sets. Otherwise, the ICP geometrical transformation will provide inaccurate results.

**Algorithm 1: LIDAR odometry correction**

**Input:** \( C_{k}^{sc} \) is the point cloud with a sweep correction.

**Result:** \( C_{k}^{lo} \) is the point cloud with 6-DOF LIDAR odometry correction.

**Data:** \( S \leftarrow \emptyset \) is the set of segmented point clouds.

**Data:** \( D \leftarrow \emptyset \) is the set of linear descriptors at current time. Where \( D_{l} = [x_{i}, y_{i}, z_{i}, \hat{x}_{i}, \hat{y}_{i}, \hat{z}_{i}] \).

**Data:** \( D_{map} \) is the set of linear descriptors in the mapped environment.

**Data:** \( D_{target} \leftarrow \emptyset \) is the set of target descriptors that will be used by the ICP procedure.

\[
\begin{align*}
\text{begin} & \quad \text{for } x = \{1, \ldots, 5\} \text{ L2 Euclidean distance values.} \\
& \quad \text{do} \quad S \leftarrow \text{EuclidianCluster}(x, C_{k}^{SW}) \\
& \quad \text{for } S_{i} \in S \text{ do} \\
& \quad \quad D_{aux} \leftarrow \text{RANSAC}(\text{inliers\_threshold} = 1m, S_{i}) \quad \text{Auxiliary descriptor.} \\
& \quad \quad \text{if } \hat{z}_{i} > 0.9997 \text{ then} \\
& \quad \quad \quad D \leftarrow D_{aux} \\
& \quad \text{if } \text{First\_iteration} \text{ then} \\
& \quad \quad D_{map} \leftarrow D \\
& \quad \text{else} \\
& \quad \quad \text{for } i \in \text{len}(D) \text{ do} \\
& \quad \quad \quad \text{for } j \in \text{len}(D_{map}) \text{ do} \\
& \quad \quad \quad \quad \text{if } \text{EuclidDistance}(D_{i}, D_{j}^{map}) > 1m \text{ then} \\
& \quad \quad \quad \quad \quad D_{map} \leftarrow D_{i} \\
& \quad \quad \quad \quad \text{else} \\
& \quad \quad \quad \quad \quad D_{source} \leftarrow D_{i} \quad D_{target} \leftarrow D_{j}^{map} \\
& \quad \quad \quad \text{6DOF\_correction} \leftarrow \text{ICP}(D_{source}, D_{target}) \\
& \quad \quad \quad C_{k}^{lo} \leftarrow 6DOF\_correction \cdot C_{k}^{sc} \\
\text{end}
\end{align*}
\]

C. Map creation

The geometrical transformation obtained in the previous step is now used over the \( C_{k}^{sc} \) point cloud. This way, a new corrected \( C_{k}^{lo} \) point cloud is obtained using the 6-DOF LIDAR odometry correction.

A high-resolution 3D map will be created using this new point cloud as the input for an Octomap algorithm [11]. Fig. 4 depicts the effect of the cloud correction on the map creation.
III. LOCALISATION

Indoor localisation has been traditionally solved by the use of 2D map representations and 2D sensors. This approach, which has been proved reliable and accurate indoors, is not enough to provide accurate outdoor localisation for autonomous vehicles. As a consequence, our objective is to obtain high definition 3D maps that are suitable to perform outdoor localisation.

In this paper, we propose to use a Monte Carlo Localisation method (MCL) [12], also called Particle Filter (PF), adapted to use high definition 3D maps and measurements collected from a high definition 3D sensor (Velodyne-32 LIDAR). A similar approach has been previously used in [13] using 3D maps and 2D sensors to perform indoor localisation. In this method a PF is used to obtain the 6D pose (3D position \((x, y, z)\) and the roll, pitch and yaw angles \((\phi, \theta, \psi)\)) of a humanoid robot carrying a Hokuyo laser on the head.

For the autonomous vehicle localisation we adapted these methods to obtain a 3D pose (2D position \((x, y)\) and the yaw angle \(\psi\)) using a ray casting model to evaluate the fitness of the point cloud over the 3D map. The main idea behind this method is to keep a set of particles that will represent possible locations of the vehicle. Each particle is scored according to the similarity between the real measurements collected using the Velodyne-32 and the measurements that should be obtained provided the vehicle was located exactly on a particle pose. Finally, the vehicle’s location can be estimated using the pose of the particles with the highest weights. This method consists of the following steps:

A. Initialisation

During the initialization step, particles have to be distributed over the map covering all possible vehicle poses. This area could be thousand square meters in outdoor environments making unfeasible the initial distribution of the particles over the whole map. Therefore, the initial distribution must be reduced to the poses around the initial position of the vehicle. This position is obtained using the rough location provided by a Garmin 18x LVC GPS (accuracy <15 meters, 95% typical), not enough for autonomous vehicles navigation but enough to initialise the filter. In order to reduce the initial error, a fixed number of equally weighted particles are randomly generated covering the area around the initial position.

B. Update (Weight computation)

At this step, a weight will be computed for each particle containing the probability of being the real location of the vehicle. The weight of each particle will be computed by scoring the similarity between the point cloud collected using the Velodyne-32 and the measurements that should be obtained provided the vehicle was located exactly on a particle pose. This similarity will be scored by using a ray casting algorithm computed from the position of each particle (Fig. 5). This way, one beam will be launched in the direction of each point \(p_i\) in the point cloud until it intersects with an object on the map. Then, a score \(\phi(p_i)\) (Eq. 1) will be calculated depending on the difference \(d\) between the distance to the real point \(d_r\) and the distance to the intersection with the map \(d_m\). An additional score is added if \(d_r\) is smaller than \(d_m\) (to cover occlusions, highly likely in driving outdoor scenes).

\[
\phi(p_i) = \begin{cases} 
\alpha \exp \left(-\frac{d^2}{2\sigma^2}\right) & , d_r \geq d_m \\
\alpha \exp \left(-\frac{\alpha^2}{2\sigma^2}\right) + \beta \left(\frac{\exp(\lambda d_r)}{1-\exp(\lambda d_m)}\right) & , d_r < d_m 
\end{cases} 
\]

where \(\sigma\) is the sensor noise and \(\alpha, \beta, \lambda\) are weighting factors.

![Fig. 5. Score given to a point depending on the difference between the real distance and the distance to the intersection on the map.](image)

The score \(\Phi\) assigned to each particle will be the sum of the scores of each point in the point cloud \(\Phi = \sum \phi(p_i)\). Finally, the particle weights are normalised so they sum up 1.

The point cloud is filtered to reduce the computational effort required to calculate the weights of the particles. To do so, an intelligent feature selection is applied: by eliminating points close to each other (with distances under the map resolution), by removing the points on the ground (there are no differential features on the ground plane) and selecting points on distinctive features such as corners or poles.
C. Pose estimation, resampling and propagation

Once the weight for all the particles is computed, the most likely pose of the vehicle is estimated as the mean pose of the particles with the highest weights. Then, during the resampling stage, the particles with high weights are replicated while the particles with low weights are removed to avoid the degeneration of the particle cloud. Finally, the particles are propagated using the vehicle’s motion model to continue with the next iteration of the PF.

IV. EXPERIMENTAL ANALYSIS

The final objective of creating a high quality map is to obtain accurate localisation of the vehicle. However, there are many factors to take into account in order to evaluate the map creation accuracy and the localisation performance.

Firstly, in our system, an RTK-GPS-based EKF is used as groundtruth, but it is not free of errors as we will show in the next sections. This implies that corrections made in the mapping or localisation phases over the EKF data will be considered as errors in the final results. As a consequence, and to give a more realistic figure of the mapping accuracy, street poles positions were manually measured using an RTK-GPS (Fig. 6). These poles positions were averaged over 500 samples. The mean distance and variance of the mapped poles positions to the real ones will be used as an indication of the mapping accuracy.

Secondly, the numeric results of the localisation stage should be considered as tentative and will have to be validated on autonomous driving experiments where the localisation outputs will be used for navigation tasks.

![Fig. 6. Poles position groundtruth represented as red dots.](image)

A. Experimental set-up

The experiments were performed on the Universidad de Alcalá (UAH) campus located at Alcalá de Henares (Madrid, Spain). The test area is a semi-industrial compound with wide open areas and large buildings connected by roundabouts. Data was collected driving in real traffic conditions for mapping and localisation on two consecutive runs (Fig. 7). Mapping data was collected at an approximate speed of 17 km/h in a naturalistic driving (Fig. 7(a)), while localisation test data was collected, first on a straight line, and then swerving (Fig. 7(b)).

Finally, the groundtruth was obtained using a 3-DOF EKF (based on RTK-GPS information) as explained in section II-A. For the localization stage, the mean euclidean distance error between the PF estimation and the groundtruth is used as performance indicator.

B. Mapping results

A reference map, based only on the EKF positions and the pointcloud without any further pre-processing, was created to compare the mapping results. In the creation of this map, the errors observed in the EKF positioning due to loss of coverage or loss of corrections were manually removed. The idea was to establish a baseline for comparison. It is worth reminding that the EKF estimation does not include pitch and roll angles, and thus, some improvement was expected to be gained with the 6-DOF LIDAR odometry correction. Table I shows the mean distance and variance of the poles position for both maps.

<table>
<thead>
<tr>
<th></th>
<th>Mean Euclidean Distance Error</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference map</td>
<td>22.30 cm</td>
<td>6.00 cm</td>
</tr>
<tr>
<td>6-DOF LIDAR odometry</td>
<td>9.74 cm</td>
<td>4.16 cm</td>
</tr>
</tbody>
</table>

As expected, the 6-DOF LIDAR odometry mapping reconstructs more accurately the poles position by a factor of almost two. This shows that the 6-DOF LIDAR odometry technique is able to correct for some of the errors introduced by the EKF and that pitch and roll angles estimation have a significant effect in the map creation accuracy.

C. Localisation results

The localisation system described in Section III was tested on both, the “error free” reference map and the 6-DOF LIDAR odometry based one. Our purpose was two-fold: First, to test the effect on localisation of introducing corrections of pitch and roll angles on the map creation. Second, to evaluate the performance of a map created relying on LIDAR odometry. Table II shows the mean localisation distance errors for both, the reference and the 6-DOF LIDAR odometry maps in the test trajectory.

![Fig. 7. Experimental environment and trajectories. (a) Trajectory for map creation. (b) Trajectory for localisation tests. In blue, straight line, in red roundabout and in green swerving.](image)
The performance of the localisation on the 6-DOF LIDAR odometry map is comparable to the “error free” reference map indicating that is possible achieve similar levels of accuracy using LIDAR odometry instead of RTK-GPS. Fig. 8 shows an example where the PF is correctly estimating the vehicle’s position, but the EKF groundtruth is off by about 1 metre. As explained before, some of the localisation error is accounted for EKF errors caused by RTK-GPS blackouts, meaning that the final localisation accuracy should be slightly higher.

![Fig. 8. PF (red arrow) and EKF groundtruth (green arrow) localisation results. The Velodyne-32 hits are represented as red dots.](image)

Although more environments and challenging situations (i.e. strong occlusions) should be tested, these preliminary results indicate that the mapping and localisation techniques are accurate enough for navigation tasks in autonomous vehicles. This method is a first approach towards the removal of RTK-GPS from the mapping and localisation stages.

### V. Conclusion

In this paper, we proposed a new method for feature selection in laser-based point clouds with a view to achieving robust and accurate 3D mapping. The proposed method follows a double stage approach to map building: a 3-DOF point cloud distortion compensation and a 6-DOF LIDAR odometry-based motion estimation. Experiments were performed while driving in real traffic conditions in a semi-industrial compound. The results show that our mapping technique increases the mapping accuracy by a factor of two, while maintaining performance on the localisation stage. Our approach is a first step towards the full removal of RTK-GPS from both mapping and localisation stages.

As future work, for the localisation stage, we plan to test on different environments with strong occlusions. For the mapping stage, we want to introduce additional linear features and to pitch the Velodyne-32 around 45° only for the map creation. This is expected to reduce sparseness of the maps and also some of the errors introduced by the furthest targets.

### Acknowledgment

This work was funded by Research Grants SEGV AUTO S2013/MIT-2713 (CAM), DPI2017-90035-R (Spanish Min. of Economy), and BRAVE, H2020, EC Contract #723021. This project has also received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737469 (AutoDrive Project). This Joint Undertaking receives support from the European Union Horizon 2020 research and innovation programme and Germany, Austria, Spain, Italy, Latvia, Belgium, Netherlands, Sweden, Finland, Lithuania, Czech Republic, Romania, Norway.

### References


