2018 IEEE International Conference on Robotics and Automation (ICRA)
May 21-25, 2018, Brisbane, Australia

# DELIGHT: An Efficient Descriptor for Global Localisation using LiDAR Intensities

Konrad P Cop[1,2], Paulo V K Borges[1], and Renaud Dubé[2]

***Abstract*— Place recognition is a key element of mobile robotics. It can assist with the "wake-up" and "kidnapped robot" problems, where the robot position needs to be estimated without prior information. Among the different sensors that can be used for the task (e.g., camera, GPS, LiDAR), LiDAR has the advantage of operating in the dark and in GPS-denied areas. We propose a new method that uses solely the LiDAR data and that can be performed without robot motion. In contrast to other methods, our system leverages intensity information (as opposed to only range information) which is encoded into a novel descriptor of LiDAR intensities as a group of histograms, named DELIGHT. The descriptor encodes the distributed histograms of intensity of the surroundings which are compared using chi-squared tests. Our pipeline is a two-stage solution consisting of an intensity-based prior estimation and a geometry-based verification. For a map of 220k square meters, the method achieves localisation in around 3s with a success rate of 97%, illustrating the applicability of the method in real environments.**

## I. Introduction

Global localisation is an essential element for autonomous robotic navigation. It is used in tasks like the "robot wake-up" problem (lost robot problem) and loop closure in Simultaneous Localisation and Mapping (SLAM). The goal is to compute the pose of the robot by estimating the transformation of the local observations with respect to a global map.

Although there are scenarios in which the issue can be solved using external sensors (such as in GPS), there are several areas where this infrastructure is not available. Cameras have also been successfully applied for place recognition [1], but they suffer from lower performance in unfavourable lighting conditions. Addressing the shortcomings of other sensors, in this paper we propose a novel approach for global localisation using solely the information from 3D LiDAR sensors. The main novelty of the proposed approach is the fact that we exploit the intensity return (as opposed to only range).

The global localisation problem consists of two main elements: *i*) recognition of the previously visited place, and *ii*) pose estimation with respect to the existing map. In practice, it means comparing a small point cloud obtained by scanning a certain place in the environment (called in this paper a local scan) with the previously obtained data (global map), and finding the relative transformation. Although many works focus only on the recognition part without pose estimation [2]

[1] Robotics and Autonomous Systems Group, CSIRO, Australia paulo.borges@csiro.au, vini@ieee.org
[2] Autonomous Systems Lab, ETH Zurich, Switzerland kcop.wiel@gmail.com, renaudube@gmail.com

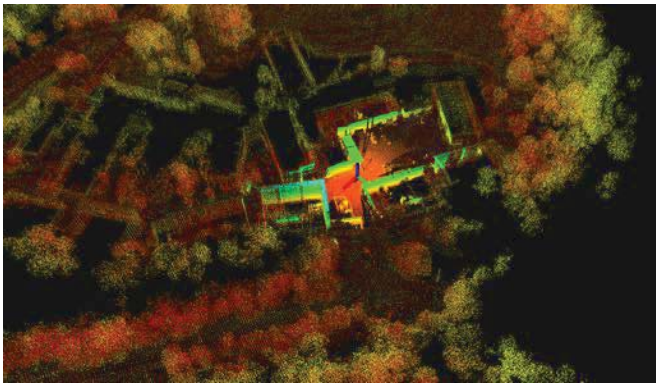


Fig. 1: A successful localisation based on our two-stage place recognition approach with the DELIGHT descriptor. The robot pose is shown as axes (red-green-blue), the point cloud is coloured based on intensity, and the local scan is depicted in bright colours.

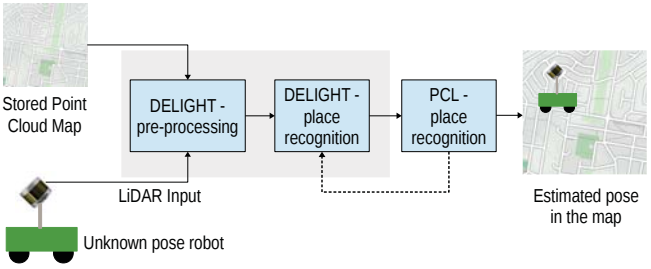



Fig. 2: Diagram illustrating the system structure, combining the global localisation (grey shaded area) with the local descriptors for the precise localisation (based on PCL). The grey shaded area also indicates the key contributions of this paper, based on DELIGHT.

[3], the two elements are necessary for achieving full global localisation. In our system we provide solutions to both aspects of the problem.

Typical approaches for place recognition in 3D LiDAR data are based on geometrical descriptors associated with certain features in the point clouds. Recognition comes down to finding the most similar descriptors between different clouds. Usually global and local descriptors are distinguished. The former compute single statistics for the whole local scans whereas, in case of the latter, multiple keypoints are selected from the cloud and local descriptors (of a significantly smaller radius than the point cloud) are calculated around them. While global descriptors are very efficient for dimensionality reduction and correspondence search, they, in gen-

978-1-5386-3081-5/18/$31.00 ©2018 IEEE

eral, cannot provide the relative transformation. There exist global descriptors that try to overcome this limitation [4], but any additional information requires further computational effort, e.g. calculating the viewpoint. Such additional steps impair the efficiency of global descriptors. Transformation can be also determined by using local descriptors but at the expense of increased processing time since the complexity of the problem grows significantly with the expansion of the map [5].

To simultaneously keep the advantages of local descriptors and to limit the processing time, the search space can be severely restricted using the descriptor proposed in this paper. Then, on this restricted area, local-descriptors-based recognition can provide the final precise position of the robot. These two steps form the pipeline of the framework proposed. There is additional information available in LiDAR data that is not only associated with the target's position in the space (range information) but also with the characteristics of its surface, namely the intensity (or reflectivity). This feature has only been partially investigated in literature, and as we show in this paper, it contains very meaningful information for place recognition. Hence, we propose a novel approach for place recognition based on DELIGHT: a DEscriptor of LiDAR Intensities as a Group of HisTograms.

In the first step of our system is to limit the map to the most likely area (prior) based on intensity similarity using the DELIGHT descriptor. Secondly, a geometrical recognition is performed between the local scan and the prior. Figure 2 summarizes the pipeline, also highlighting the specific contributions of this paper, represented by the shaded area. A video summarising the system is available online[1]. The reader is encouraged to view this video for an animated description of our approach.

To the best of our knowledge, this is the first work that proposes a sub-real-time solution to global localisation based on global intensity descriptors. The main contributions of this paper are:

- DELIGHT: a novel LiDAR intensity based global descriptor.
- An approach for performing place recognition based on the DELIGHT descriptor.
- A reliable and fast solution for global localisation using only a 3D LiDAR sensor, by combining DELIGHT-based recognition with keypoint-based geometry verification.
- A thorough validation of the proposed algorithm for a wake-up problem in large real-world environment.

This paper is structured as follows. In Section II, existing works leveraging LiDAR intensity and solutions for global localisation are reviewed. In Section III the DELIGHT descriptor is presented. In Section IV the whole pipeline for global localisation is outlined. Section V contains the evaluation of the system with experiments, followed by relevant conclusions in Section VI.

[1]Video demonstration of the system: https://youtu.be/nK2ylxLi8uQ

## II. Related Work

Several methods for global localisation in 3D point clouds have been proposed. Among techniques based on global descriptors where each point cloud is encoded using a single feature vector, Röhling et al. [2] proposed to describe places by a histogram of points' elevation. Another method is to create the descriptor by dividing the point cloud a into a cubic grid and calculating the density function which describes the shapes that are later gathered into histograms [3]. One issue with global descriptors is the fact that the map is interpreted as a discrete set of places along the trajectory, and each place has a defined origin forming a graph-like structure. In this case the only possible robot locations are in the graph vertices. When the robot is between the vertices or off the original trajectory the exact location cannot be retrieved.

Local descriptors can also be used for place recognition. Bosse and Zlot [6] used descriptors calculated around randomly selected keypoints in the global map and the local scan. Place recognition is based on the voting for the most probable places. After geometric consistency validation, the relative transformation is obtained using solution to the Absolute Orientation Problem [7]. In a similar approach, Tombari and Di Stefano [8] do not divide the global point cloud into a set of places but look for all correspondences instead. The correspondences are then grouped into actual instances of the place using Hough voting. Dubé et al. [10] recently proposed to recognize places on the basis of 3D segment extraction and matching, offering a trade-off between local and global descriptors. The authors have demonstrated that this strategy can be used in real-time multi-robot SLAM applications [11]. However, in order to extract segments which are robust to changes in point of view, this technique assumes that the robot moves in its environment, which is not always applicable or safe in the robot wake-up problem.

Closely related to the task of global localisation, the SLAM problem was approached by several works proposing to leverage the LiDAR intensity information. Khan et al. [12] focused on calibrating the intensity to obtain the reflectivity and applied the results to SLAM. A number of solutions also combine intensity with information from other sensors. Levinson et al. [13] [14] proposed a SLAM algorithm based on road reflectivity maps. Pandey et al. [21] introduced an interesting approach, which combines the information from an omnidirectional camera and LiDAR. Barfoot et al. [15] generate the intensity images from the LiDAR data, extract SURF descriptors and perform feature-based localisation similar to the visual SLAM. Their solution is an interesting attempt to use intensity only for localisation, however it is not always robust due to significant distortion arising from the robot motion and it does not solve the global localisation problem. Also, none of the above works propose to use the LiDAR intensity information in a global place recognition approach.

## III. DELIGHT Descriptor

In this section we outline the details of the DELIGHT descriptor, initially providing insights on the use of intensity,

followed by a presentation of the descriptor structure. The use of the intensity information encoded into the DELIGHT descriptor is the main novelty of the proposed system.

### A. Concept of intensity

In addition to range information, a LiDAR can also provide the intensity measure for each point, which is the strength of return after reflection from a surface [12]. Depending on the field, intensity can be referred to as *remission*, *brightness* or *reflectivity* (*reflectance*) [16] [17]. Additionally, *reflectivity* is a property of a surface obtained from intensity by compensating for intrinsic parameters of the sensor as well as extrinsic features such as distance to the object, incidence angle, and air density [12]. Once a sensor is calibrated, the data measured correspond to the unique characteristics of the surface. A typical approach for compensation (calibration) is to compare the intensity of the target object with the reflection from the reference surface [16].

### B. Descriptor structure

The DELIGHT descriptor consists of $m$ non-overlapping bins which are obtained by the following division approach: a spherical support structure centered around the sensor is divided radially into two concentric spheres with a predefined radius $r_1$ for the outer sphere and $r_2$ for the inner sphere. Additionally, the spheres are divided horizontally into upper and lower hemispheres. Finally, the azimuthal division is applied for hemispheres, every ninety degrees, resulting in $m = 16$, as illustrated in top picture of Figure 3. For each descriptor bin, an histogram of intensity of the $n$ points falling into it is computed. This histogram is composed of $b$ adjacent bins of equal size and can be computed in $O(n \log b)$. The proposed structure is embedded in the descriptor name: DEscriptor of LiDAR Intensities as a Group of HisTograms.

The descriptor can be computed around any arbitrarily selected keypoint. In case of the global descriptor, this keypoint is the origin of the LiDAR and the radius we define by this sensor's range $R$ which means that $R = r_1$. Example of the division of a typical local scan is depicted in the bottom picture of Figure 3, where $R = 100$ m, therefore the radius of the outer sphere equals to $r_1 = R = 100$ m. The inner sphere in this example has a radius $r_2 = 15$ m.

### C. Descriptor alignment

Due to the specific form of division, a repeatable reference frame must be established to determine in which bin the points are located (i.e., it must be clear where is the "top", "bottom", "left" and "right" of the cloud). To obtain the reference frame, a Principal Components Analysis (PCA) of all points located within the support is performed as proposed by Tombari et al. [9]. The method is based on the eigenvalue decomposition of the nearest neighbours covariance matrix and total least squares estimation. After obtaining the reference frame, bins boundaries are aligned with the axes and each point is given an additional feature depicting the unique bin identifier.

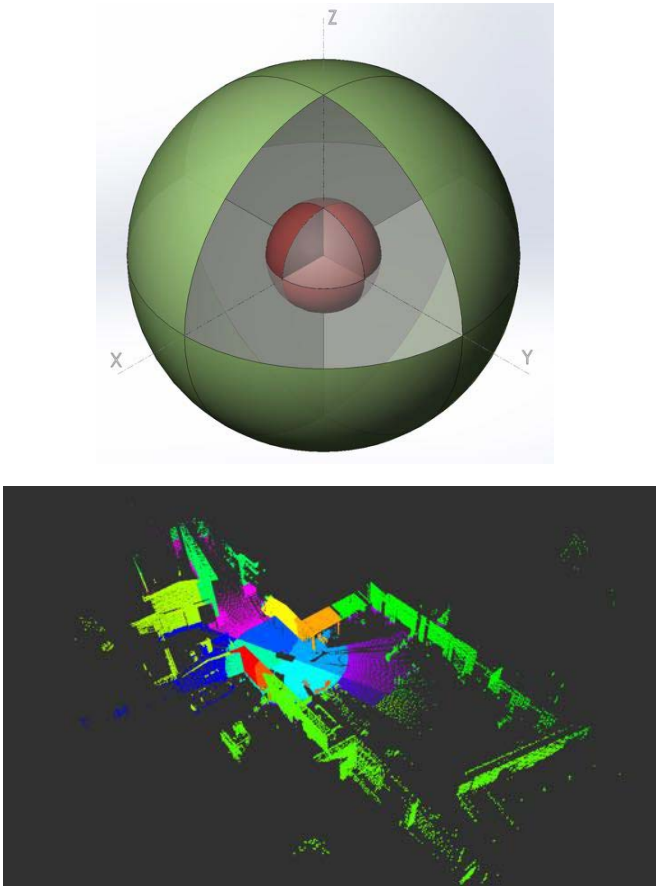



Fig. 3: The structure of the DELIGHT descriptor is shown in the top picture. The red colour depicts the inner sphere while the outer sphere is shown in green. Each sphere is divided in 8 bins which results in a total of 16 bins. The example of a point cloud division is shown in the bottom picture. In our system, DELIGHT is a global descriptor and for each cloud it contains an entire point cloud with the origin in the robot's position. The colours in the bottom picture correspond to the bins of the descriptor.

Next, for each bin of the descriptor, the histogram of intensity of all points located in it is computed. It is important to note that the intensity is the information directly obtained from the sensor (during scanning). This implies that no additional computation is necessary to create histograms apart from counting the occurrence of points with specific intensity, which results in $O(b)$ complexity and ensures high efficiency. Consequently the intensity descriptor consists of a tuple of spatially distributed intensity histograms.

### D. Matching

For the similarity assessment, the histogram of the $i$-th bin of descriptor $A$ is compared with histogram of the $i$-th bin

of the descriptor $B$ using the chi-squared test $S^i_{AB}$ defined as

$$S^i_{AB} = \sum_{k=1}^{b} \frac{2 \cdot (A(k) - B(k))^2}{A(k) + B(k)} \quad (1)$$

where $A(k)$ and $B(k)$ refer to the $k$-th bins of the histograms $A$ and $B$, respectively, and $b$ is the number of bins of each histogram. Hence, $S^i_{AB}$ corresponds to a similarity value. The similarity metric of the two descriptors $S_{AB}$ is obtained by the average of results of chi-squared tests of all bins. It is important to notice that other similarity tests, such as e.g. Euclidean distance, can be also used.

An important step for correctly matching the descriptors is the elimination of the ambiguity in axes directions obtained from the PCA. According to our experiments the disambiguation method proposed by Tombari et al. [9] is not reliable for clouds that are symmetric or have varying density. The approach proposed by Bosse and Zlot [6] requires an IMU to determine the vertical direction, while we aim to use exclusively the LiDAR data. For this reason we propose an alternative approach. As long as the directions $z$ and $x$ are defined, the direction of $y$ can be determined by the cross product of $z$ and $x$ unit vectors. Therefore, there are four possible combinations of the reference frame ($z$ "up" or "down" and $x$ "left" or "right"). We assume that all combinations of bins are equally probable and compare descriptor $A$ with four different "versions" of descriptor $B$. In this case, different versions mean different sequences of bins which depend on the orientation of the reference frame. The possible options are depicted in Figure 4, which shows the orientation of the axes and the numbering of bins. It is important to notice that it is only the unique identifier of the bin that changes but not the contents of the bin. The output of this comparison is a set of four similarity values. As a comparison result we select the minimum value:

$$S_{AB} = min\{{}^1S_{AB}, {}^2S_{AB}, {}^3S_{AB}, {}^4S_{AB}\}.$$

Due to the simplicity of the descriptor, the proposed approach adds no significant increase to the complexity and processing time, as illustrated in the experiments in Section V.

## IV. Global Localisation Pipeline

In this section the full global localisation pipeline is described. Since in our application the system is a module in an autonomous vehicle, the description is focused on solving the wake-up problem, although a similar approach can be used for other global localisation tasks.

### A. System overview

The main goal of the system is to provide full global localisation (as defined in the introduction) at the start-up of the operation (or in case the robot gets lost). Initially, the robot is given the global map of the environment, which was generated by this robot through SLAM [26], in the form of LiDAR-obtained point cloud and the corresponding trajectory. The procedure of map generation consists of two steps. First the robot drives through the environment and registers all data obtained by the LiDAR sensor. Next,

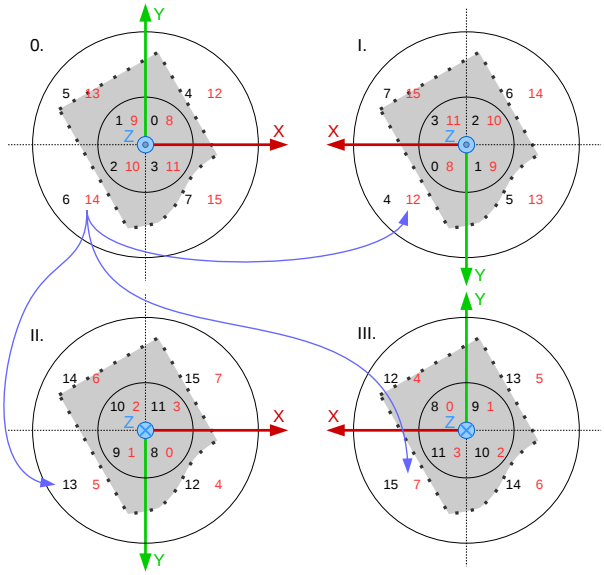



Fig. 4: Combination of bins depending on the direction of $x$ and $z$. The point cloud is represented by a shaded area within the descriptor. In all cases the position and orientation of the point cloud remain unaltered but the directions of the axes change. It is important to compare bins that contain the same parts of the point cloud. Case 0 shows the original reference frame, where the digits correspond to the unique bin identifiers. Cases I., II. and III. show the reference frames and bin sequences in case of different axes orientations. Red numbers correspond to the unique identifiers of bins in the bottom hemisphere and black in the upper. During the similarity check the corresponding bins should be compared. An example is shown using the blue arrows. Bin 0.6 should be compared with 0.6, I.4, II.13 and III.15, while bin 0.14 with 0.14, I.12, II.5 and III.7.

the collected data is processed by the mentioned SLAM algorithm which provides stitching of the consecutive scans and computes global alignment of the generated point cloud and the trajectory. The high-level scheme of the system is shown in Figure 5, with details in the following sections. The pipeline consists of two main stages: *i*) intensity-based prior extraction and *ii*) geometry-based verification.

### B. Preprocessing of the global map

The procedures described in this section only have to be performed once on a newly generated map. Once on a global map against which localisation is desired. This map can be obtained by driving the robot through the environment and by accumulating the LiDAR data. This data are fed to a SLAM algorithm which estimates the robot trajectory by registering successive scans. For the actual localisation pipeline, the computed elements are loaded form memory.

The first and key part is the division of the map, which is especially significant because each extracted place has to correspond to the scan that can be obtained by the robot being in that place and not moving. Extraction based on the sensor range (i.e., the descriptor of all points lying within a

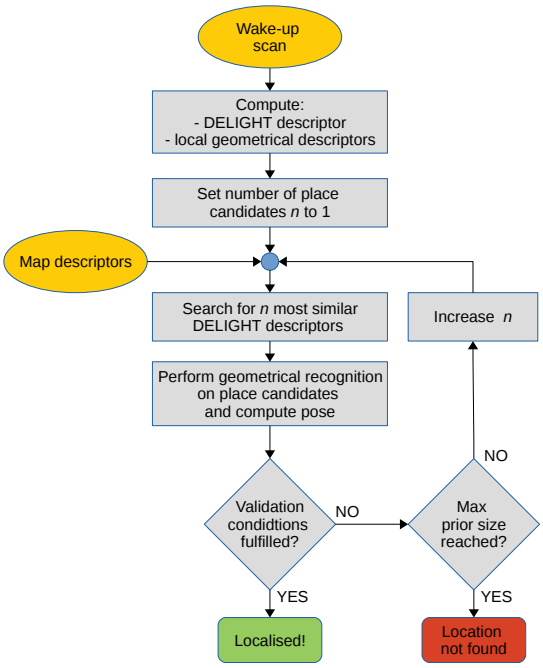



Fig. 5: High-level scheme of the localisation pipeline.

(a) Image of the vehicle (around the centre of the image) in the test environment.

(b) Corresponding intensity coloured point cloud as scanned from the vehicle.

Fig. 6: Example of the test environment.

sphere with radius equal to the sensor range) will not bring correct results. This is because the areas which cannot be seen by the robot, e.g. an object behind a wall, will fall into the descriptor sphere and therefore influence the descriptor.

Instead, we extract places based on the mapping trajectory. We first divide the trajectory into segments of length $k$ meters. For each segment we create a point cloud from all the points that were visible when the robot drove through this small part of the trajectory and assign the origin of the cloud to the middle of the segment. By partitioning the trajectory based on distances, the map is divided into places that are very similar to the scans which are generated by static robots. Therefore, it is important to keep the segments sufficiently short. In other words, with this approach we approximate that the map is divided into a set of static scans with origins located on the trajectory, and that have a distance of $k$ meters to each other (in our experiments, $k = 2$ m).

After extracting the places the DELIGHT descriptor is computed for each place, according to the description in Section III-B. Correspondingly, the local, geometrical descriptors are also calculated (for details see Section IV-E).

### C. Generating local scans

In our platform scanning is realised using a 3D LiDAR sensor Velodyne VLP-16. This sensor has only 16 rings which results in a low resolution for objects located far from the sensor. To increase robustness, the sensor is mechanically rotated around the $z$-axis and tilted 45° from the vertical position, as shown in Figure 7. The DELIGHT descriptor is composed of histograms which are the metrics of quantity and it is important to ensure that the distribution of points in the cloud is as uniform as possible (in case some areas are swept more times than others). For this purpose, after the scan is obtained, the points are downsampled with a high resolution using uniform sampling. Additionally, the rotating setup enables a direct control of the sweep and in our system the wake-up scan is generated from points obtained from two full rotations of the LiDAR. A typical local scan is depicted in Figure 6b.

### D. Intensity-based recognition

After computing the DELIGHT descriptor of the wake-up scan, a simple linear correspondence search is performed against all descriptors of the places extracted from the map (see Figure 5). From the comparison results (as explained in Section III-D) the most probable place is selected as the one with the highest similarity (smallest $S_{AB}$). For correctness validation and full localisation, the geometry-based recognition is introduced to the process as a second stage.

### E. Geometry-based verification

The transformation between two point clouds can be acquired using the solution to the Absolute Orientation Problem [7]. This algorithm however requires establishing the corresponding points between the clouds.For this purpose we perform recognition based on local descriptors. In our approach we follow the pipeline proposed by Aldoma et al. [19]. First keypoints are selected from the point clouds by downsampling, then the SHOT descriptors [9] are computed

for all keypoints. Next, the correspondences are found between the clouds and they are subsequently filtered based on the geometric consistency. For the correspondences found, the Absolute Orientation algorithm is applied followed by RANSAC [20] to eliminate inconsistent matches. Finally, ICP is used to refine the transformation and align the wake-up scan with the global map.

A secondary role of the geometrical stage is to verify the quality of the intensity-based recognition. In case of a significant change in the environment, close intensity similarity of the places in the environment or the robot being far off the trajectory, the initial intensity-based recognition, using one candidate, may not find a correct place. To consider this place a correct match two conditions must be fulfilled: *i*) the number of correspondences between local descriptors must be above a given threshold $T_1$ and *ii*) if the fitness score of the ICP must be below a second threshold $T_2$. These conditions provide the system with ability to discard false matches. Threshold parameters depend on the resolution of the point clouds and should be found experimentally. If the above requirements are not fulfilled the number of places candidates $n$ is increased and the procedure is repeated. The maximum value of $n$ can also be limited. If the amount of place candidates reaches the maximum and the conditions are not fulfilled, the place is considered to be located outside of the map.

## V. Experiments

In this section we describe our experimental platform and the method adopted for generating local scans, followed by a discussion of the numerical results.

The proposed system was evaluated in a large CSIRO site in Brisbane. The area in which the experiments took place is a industrial park with different characteristics - from structured buildings to unstructured bushland, as illustrated in Figure 8. This diversity enabled us to evaluate our algorithm in various conditions and with multiple surface materials. The environment was first mapped by driving the robot and places were extracted from the point cloud generated by the SLAM system, which is based on the continuous-time SLAM implementation of Bosse and Zlot [25][26]. The robot travelled approximately 4 km within the map, resulted in 2055 extracted places. The covered area has roughly 220,000 square meters (Figure 8). Test wake-up scans were generated in 101 locations by driving the robot to these locations and keeping it stationary during the scan acquisition.

### A. Platform overview

The considered platform is based on a commercially-available TE John Deere Gator, an utility vehicle that was transformed into an autonomous platform by the CSIRO team [23][24] (Figure 7). This robot is able to drive autonomously over all areas of the site, and has covered more than 200 km under unmanned operations [22]. The system proposed in this paper found direct and successful applicability to this platform, with efficient wake-up localisation.


Fig. 7: The John Deere Gator platform automated by the CSIRO. The LiDAR Velodyne sensor is mounted above the vehicle and is rotated by a motor at a 45° angle.

### B. Sensor calibration

As mentioned in Section III-A the sensor requires calibration for correct estimation of the intensity. The VLP-16 that we used for the experiments was calibrated by the producer so we assumed intrinsic and extrinsic factors to be compensated (e.g., laser power or distance to the object)[18] and we directly used the intensity that is output by the sensor. In this case, intensities are encoded in 8-bit values and we therefore set $b = 256$ (the number of bins per histogram). It is important to notice that in the case of the incidence angle, the surface reflection model is required for correct compensation. A universal model for various surfaces does not exist and it should be estimated for each surface type separately. To make the system more generic, we do not estimate the reflection model for each object due to the large variety of surfaces that exist in most environments. We also do not compensate for the incidence angle, although this could potentially be implemented. However, as the experiments illustrate (see next section), very good localisation performance was achieved without this compensation.

### C. Results and Discussion

There were 101 wake-up locations tested around the entire area. They are depicted by the red dots in the top image of Figure 8. In order to validate the robustness of the pipeline, these wake-up locations are located in both visually similar and dissimilar areas. The locations in the midst of the buildings (central part of the top picture of Figure 8) and in the semi-forested areas (bottom right quarter) are particularly similar in terms of materials and structures present. In addition, some of the locations were used several times in different lighting conditions to investigate the influence of sunlight. The experiments show that the performance is independent of the illumination. The results are shown in Figure 9a. The proposed localisation approach achieved an overall success rate of 97%, with wake-up scans that

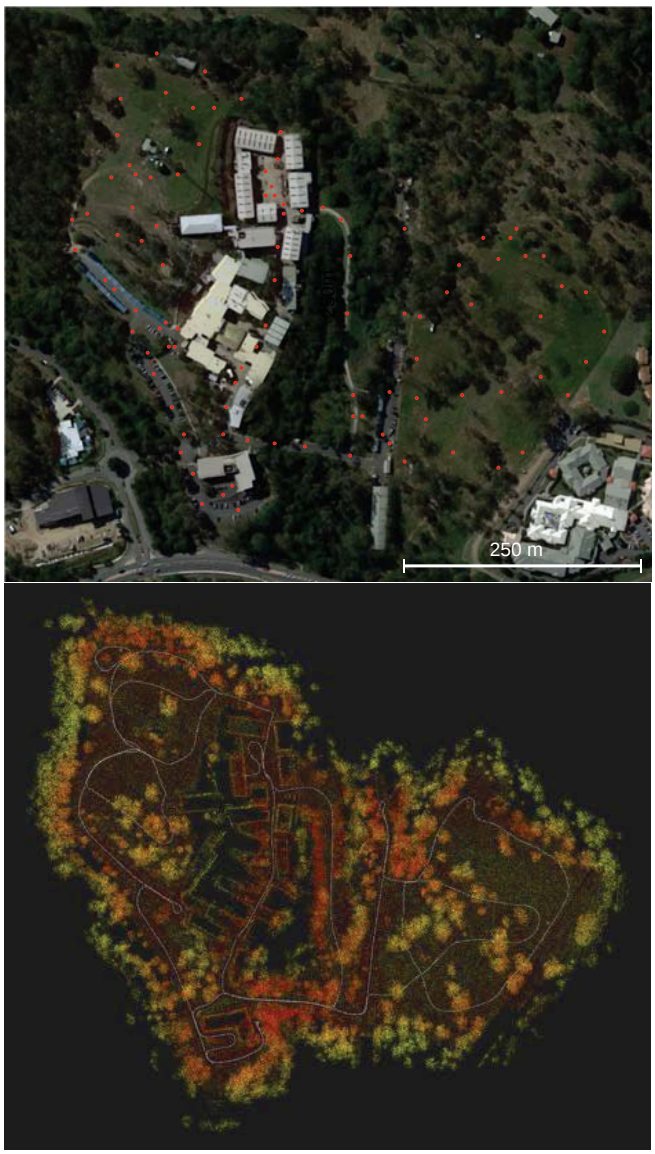
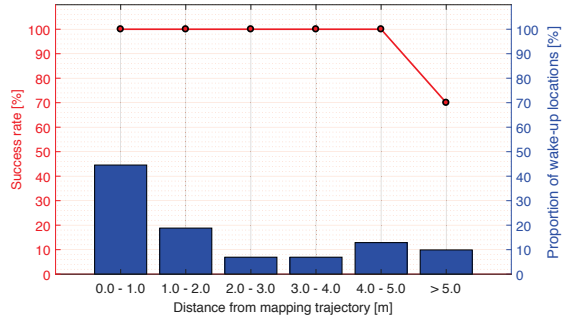
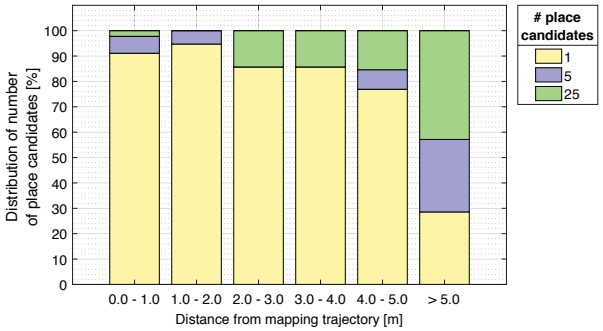


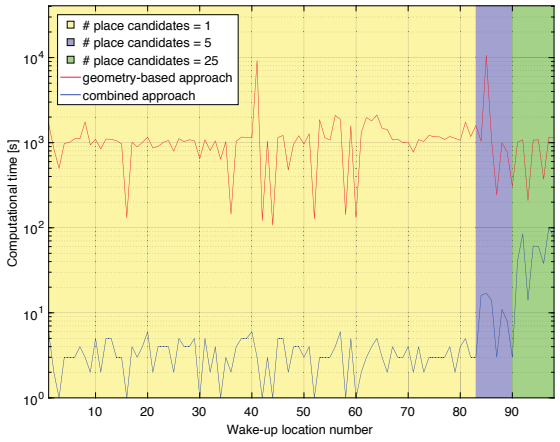

Fig. 8: Overview of the test environment. The top image shows the satellite picture with red dots indicating the wake-up locations in the tests. The bottom picture shows the corresponding point cloud with the mapping trajectory depicted in white.




(a) Success rate for different distances from trajectory depicted with the red curve. The blue bars show the proportion of wake-up locations at different distances from the mapping trajectory.




(b) Required number of place candidates to successfully localise at different distances from the mapping trajectory.




(c) Time reduction using the proposed system (combined approach) versus the geometry-based approach only. Note that the vertical axis is logarithmic.

Fig. 9: Success rate, statistics and computational timings resulting from 101 localisation attempts in a 220k square meters multi-characteristics environment.

were generated at various distances from the trajectory and different times of the day. Clearly, despite the fact that the places extracted from the map are located solely along the trajectory, the algorithm can perform successful localisation even for wake-up scans that lie far away from it.

The effect of noise and uncertainties can be reduced by increasing the amount of place candidates. Figure 9b shows that the further the robot is from the original trajectory, the less certain it is about its position. In that case it may happen that the geometrical validation fails because of an excessive difference between the section of the map's point cloud and the local scan. In this case, the number of candidates is increased. Nevertheless, in the majority of cases (82%) using a single place candidate was enough to successfully localise. In our experiments the furthest point from the mapping trajectory for which our system successfully managed to localize was located at a distance of 15 meters. The three failures (out of 101 tests) that occurred were noticed for places that lie closer to the trajectory, and these took place

in a mostly featureless environment (no characteristic objects around like trees or bushes).

In Figure 9c, the plot shows the processing times for two different approaches: *i*) classical geometry-based only approach (i.e., using the second step of the pipeline only, without the intensity-based candidates selection) where the search for the correspondences is performed between the wake-up scan and the whole, unlimited map, and *ii*) our approach, i.e., geometry-based recognition performed only on the places candidates selected based on intensity. Additionally, Figure 9c contains the information about the required amount of place candidates necessary for a successful localisation. This is shown by different background colours corresponding to the increased number of candidates. It is also illustrated that the number of candidate places directly correlates with the processing time of the algorithm. As it can be seen, for a single place candidate the average required processing time drops from approximately 1000 seconds to 3 seconds. Increasing the amount of candidates results in longer processing times which are still much shorter than in the traditional approach.

## VI. Conclusions

This paper presented a novel approach for global localisation based on the intensity information provided by LiDAR sensors, introducing the novel DELIGHT descriptor. Our experiments clearly show that the intensity contains valuable information that can be leveraged to achieve efficient place recognition. Additionally, when properly enhanced with geometrical information, it can provide a full and robust localisation algorithm. We showed that with our system we can achieve very fast localisation outperforming (in terms of processing time) our baseline approach using local descriptors. The algorithms have been implemented and tested in a robotic vehicle to assist in the wake-up problem and are now used in all operations of the robot.

## References


[1] S. Lowry et al., "Visual Place Recognition: A Survey," in *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 1-19, Feb. 2016.

[2] T. Röhling, J. Mack, and D. Schulz, "A fast histogram-based similarity measure for detecting loop closures in 3-d lidar data," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 736–741.

[3] M. Magnusson, H. Andreasson, A. Nuchter, and A. J. Lilienthal, "Appearance-based loop detection from 3d laser data using the normal distributions transform," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 23–28.

[4] A. Aldoma et al. , "OUR-CVFH–oriented, unique and repeatable clustered viewpoint feature histogram for object recognition and 6DOF pose estimation,", Joint DAGM (German Association for Pattern Recognition) and OAGM Symposium. Springer, 2012, pp. 113–122.

[5] A. Luis, "3d descriptors for object and category recognition: a comparative evaluation," in *International Conf. on Intelligent Robotic Systems - IROS*, 2012.

[6] M. Bosse and R. Zlot, "Place recognition using keypoint voting in large 3d lidar datasets," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2677–2684.

[7] B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *JOSA A*, vol. 4, no. 4, pp. 629–642, 1987.

[8] F. Tombari and L. Di Stefano, "Object recognition in 3d scenes with occlusions and clutter by hough voting," in *Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on*. IEEE, 2010, pp. 349–355.

[9] F. Tombari, S. Salti, and L. Di Stefano, "Unique signatures of histograms for local surface description," *Computer Vision and Image Understanding*, vol. 125, p. 251264, 2014.

[10] R. Dubé, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "Segmatch: Segment Based Place Recognition in 3D Point Clouds," *IEEE International Conference on Robotics and Automation*, 2017, pp. 5266–5272.

[11] R. Dubé, A. Gawel, H. Sommer, J. Nieto, R. Siegwart, and C. Cadena, "An Online Multi-Robot SLAM System for 3D LiDARs," *IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[12] S. Khan, D. Wollherr, and M. Buss, "Modeling laser intensities for simultaneous localization and mapping," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 692–699, 2016.

[13] J. Levinson, M. Montemerlo, and S. Thrun, "Map-based precision vehicle localization in urban environments." in *Robotics: Science and Systems*, vol. 4. Citeseer, 2007, p. 1.

[14] J. Levinson and S. Thrun, "Robust vehicle localization in urban environments using probabilistic maps," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4372–4378.

[15] T. D. Barfoot, C. McManus, S. Anderson, H. Dong, E. Beerepoot, C. H. Tong, P. Furgale, J. D. Gammell, and J. Enright, "Into darkness: Visual navigation based on a lidar-intensity-image pipeline," in *Robotics Research*. Springer, 2016, pp. 487–504.

[16] S. Kaasalainen, H. Hyyppa, A. Kukko, P. Litkey, E. Ahokas, J. Hyyppa, H. Lehner, A. Jaakkola, J. Suomalainen, A. Akujarvi *et al.*, "Radiometric calibration of lidar intensity with commercially available reference targets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 2, pp. 588–598, 2009.

[17] B. Steder, M. Ruhnke, R. Kümmerle, and W. Burgard, "Maximum likelihood remission calibration for groups of heterogeneous laser scanners," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2078–2083.

[18] *VLP-16 Velodyne LiDAR Puck. User's manual and programming guide.*, Velodyne, 3 2016, rev. B.

[19] A. Aldoma, M. Zoltan-Csaba, F. Tombari, W. Wohlkinger, C. Potthast, B. Zeisl, R. B. Rusu, S. Gedikli, and M. Vincze, "Tutorial: Point cloud library: Three-dimensional object recognition and 6 dof pose estimation," *IEEE Robotics & Automation Magazine*, vol. 19, pp. 80 – 91, 9 2012.

[20] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[21] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, "Toward mutual information based place recognition," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3185–3192.

[22] A. Pfrunder et al,"Real-Time Autonomous Ground Vehicle Navigation in Heterogeneous Environments Using a 3D LiDAR" to appear in *IEEE International Conference on Intelligent Robots and System (IROS)*, 2017.

[23] Romero, A.R., Borges, P.V.K., Elfes, A. and Pfrunder, A., 2016, September. "Environment-aware sensor fusion for obstacle detection", in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2016 (pp. 114-121).

[24] R. Aeschimann and P. V. K. Borges, "Ground or obstacles? Detecting clear paths in vehicle navigation," in *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, WA, 2015, pp. 3927-3934.

[25] M. Bosse and R. Zlot, "Continuous 3D scan-matching with a spinning2D laser,"in *IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, 2009, pp. 4312-4319.

[26] R. Zlot and M. Bosse, "Efficient large-scale three-dimensional mobile mapping for underground mines," in *Journal of Field Robotics*, vol. 31,no. 5, pp. 758779, 2014.