

# 3D Surface Reconstruction from Voxel-based Lidar Data

Luis Roldão<sup>1,2</sup>, Raoul de Charette<sup>1</sup> and Anne Verroust-Blondet<sup>1</sup>

**Abstract**—To achieve fully autonomous navigation, vehicles need to compute an accurate model of their direct surrounding. In this paper, a 3D surface reconstruction algorithm from heterogeneous density 3D data is presented. The proposed method is based on a TSDF voxel-based representation, where an adaptive neighborhood kernel sourced on a Gaussian confidence evaluation is introduced. This enables to keep a good trade-off between the density of the reconstructed mesh and its accuracy. Experimental evaluations carried on both synthetic (CARLA) and real (KITTI) 3D data show a good performance compared to a state of the art method used for surface reconstruction.

## I. INTRODUCTION

A robust and accurate model of the environment is crucial for autonomous vehicles. In fact, imprecise representations of the vehicle’s surrounding may lead to unexpected situations that could endanger the passengers. Although many different geometrical representations have been proposed by both the robotics [1] and the graphics [2] communities, creating an accurate 3D model of the surroundings from mounted sensors remains a challenge.

In recent years, the constant evolution of Lidar sensors enables to obtain rich and accurate information, even in large and complex scenes. However, the size of the input data, along with the noise, misaligned scans, missing data and density variation, make the task a very challenging one. To overcome this, coarse 3D representations are often used and restrictive hypotheses are applied (i.e. planar assumptions for road detection). While these representations can be sufficient to perform path planning or obstacle avoidance, complex maneuvers might require a more accurate description of the scene’s geometry.

In this paper, we present an algorithm capable to perform a fine and accurate 3D surface reconstruction of the environment from depth sensors. From the statistics of the input point cloud sampled into a voxel grid, local approximations of the surface are performed using an adaptive neighborhood capable to cope with the heterogeneous density of the input scan. Then, a truncated signed distance field (TSDF) is estimated to obtain a continuous mesh that maintains a high level of detail and density in areas close to the vehicle. Our output mesh can be of special interest for both the robotics and the graphics community to perform different tasks, such as terrain traversability assessment or physical modeling.

<sup>1</sup>Robotics and Intelligent Transportation Systems (RITS) Team, INRIA Paris, 2 Rue Simone Iff, 75012 France. {raoul.de-charette, anne.verroust}@inria.fr.

<sup>2</sup>R&D Department of AKKA Technologies. 78280 Guyancourt, France. {luis.roldao@akka.eu}.

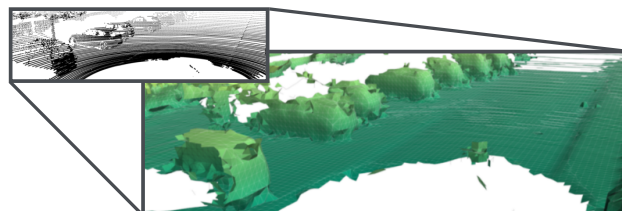


Fig. 1. Our pipeline reconstructs 3D surfaces from input Lidar point clouds (here KITTI dataset), keeping a good trade-off between accuracy and mesh density. We first rely on the approximation with explicit local surfaces, and then estimate a signed distance field from which the mesh is extracted at zero-crossing.

This paper is organized as follows: section II reviews existing works on 3D representations of the environment used for mobile robotics. A detailed description of our method is introduced in section III. An ablation study has been performed to evaluate the validity of our proposals, which is presented in section IV along with experimental results in both real and synthetic data. Conclusions and future work are presented in Sec. V.

## II. RELATED WORK

The 3D geometry of a scene described from range sensors can be represented in many different ways. The choice of the representation is often related to its final purpose: visualization, mapping, localization, terrain traversability, among many others.

Some methods use directly the 3D set of points received from the input sensor, which might be useful for visualization or obstacle detection tasks. However, the level of details of the representation depends on the amount of data used, that rapidly becomes prohibitive for outdoor scenes. Conversely, other works propose a regularly sampled grid as introduced in [3], where occupancy information is stored into each cell. This enables to handle big amounts of data more efficiently and reduce memory needs, specially by using recursive structures such as octrees [4], [5]. These approaches have become widely used for terrain traversability assessment, mapping and visualization. However, their discrete nature does not enable a continuous representation, which might be desirable for other tasks such as physical modeling.

Alternatively, the graphics community has explored different methods to create a triangular mesh from the 3D points of the scanned surface. These methods have a wide range of applications on completely different fields as described in [2]. For simplicity, we distinguish between explicit and implicit methods.

*Explicit methods:* These techniques are either parametric or triangulated with respect to how they represent the

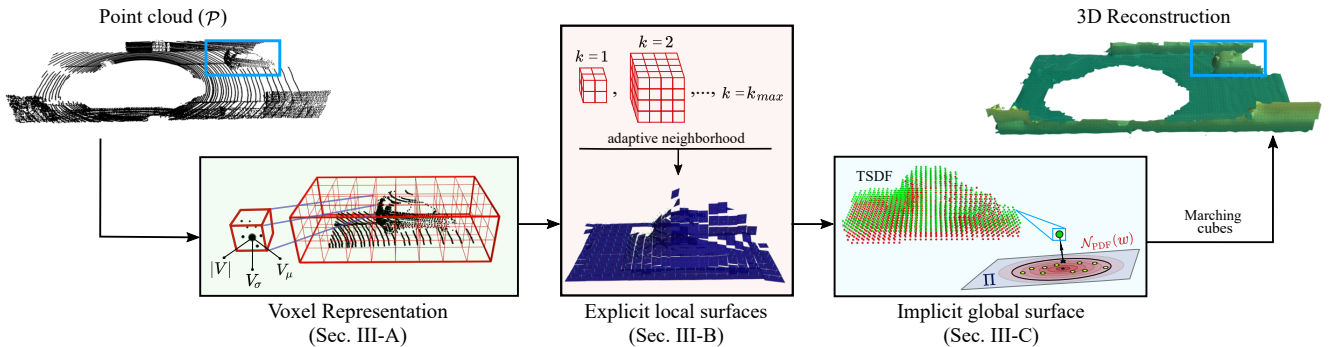


Fig. 2. Overview of our method. From left to right: the point cloud  $\mathcal{P}$ ; the voxel grid at where the statistical distribution of the points is updated (Sec. III-A); local surface approximations at neighborhoods  $k \in [1, k_{max}]$  of the grid vertices (Sec. III-B.1); TSDF calculated from the planes by considering its confidences from the statistics distribution (Sec. III-C); final 3D reconstruction.

reconstructed surface. For some applications, the continuity of the reconstruction is not required. Some researchers propose to fit local primitive shapes to the input point cloud and obtain a segmented mesh with simple primitives for the surface description [6], [7]. These approaches are not suitable for reconstructing complex scenes.

Other methods create local descriptions of the surface by a set of unoriented discs (*-surfels-*) calculated from the points distribution inside defined neighborhoods. Surfels has been used for traversability assessment [8], and more recently to perform simultaneous localization and mapping (SLAM) in outdoor urban environments [9].

*Implicit methods:* For other applications such as physical modeling or detailed terrain traversability, a more accurate and continuous representation might be preferred. To obtain this, some approaches commonly define a Truncated Signed Distance Field (TSDF) to represent the surface implicitly by a gradient field [2]. This implicit representation needs to be post-processed by meshification algorithms (e.g. Marching Cubes [10]) to obtain the final mesh.

In [11], range information across different viewpoints are integrated to average a TSDF from where the scalar field is obtained. By using this technique, 3D modeling has been performed in both small indoor [12] and large scale outdoor scenes [13], [14]. These methods typically require a large number of viewpoints to output a dense reconstruction and are susceptible to outliers.

Other implicit methods tend to perform local approximations of the surface from where the TSDF is obtained. Poisson reconstruction [15], [16] is a well-known technique for creating watertight surfaces from oriented point samples. Point-based methods such as Hoppe’s [17] or Implicit Moving Least Squares (IMLS) [18], [19], [20] locally fit the data to a lower degree polynomial by using a projection operator. For these methods, variable density inputs might represent a challenge since the data is treated equally along the complete space.

In contrast to most methods presented above, that are unable to accommodate to the heterogeneous density of the input data while keeping the accuracy of the reconstruction, we introduce a method that is able to handle this variable

density while keeping a good trade-off between the accuracy and the density of the outputted surface.

### III. METHODOLOGY

Let us consider an input point cloud  $\mathcal{P}$  obtained from any range sensor at a known viewpoint and sampled by a voxel grid, our aim is to reconstruct the underlying 3D mesh. From the statistical distributions of the points in each voxel (Sec. III-A), we first approximate local planar surfaces (Sec. III-B), and then compute the implicit surface representation (Sec. III-C) that encodes the distances to the local planar surfaces. To accommodate to the input data heterogeneous density, as well as to gain robustness to noise, we use an adaptive neighborhood kernel, resulting in a denser and smoother reconstruction. The overall overview of our methodology is shown in Fig. 2.

#### A. Voxel representation

We benefit of the work of [21] to update efficiently a regular voxel-wise representation of the point cloud  $\mathcal{P} \rightarrow \{V^1, V^2, \dots, V^n\}$ , with voxel size  $\omega$ . In addition to the number of points  $|V|$ , each voxel  $V$  stores the 3D statistical distribution of the points lying inside, that is: the mean  $V_\mu$  and the covariance  $V_\sigma$ . This enables a rich compact representation of the points inside each voxel, while being significantly lighter than storing all the points. The statistical distribution is computed incrementally upon insertion of new points.

#### B. Explicit local surfaces

It has been shown in [8], [9], that complex environments are efficiently approximated as a set of primitive local surfaces. This is especially suitable for mobile robotics, as robots usually evolve in well structured environments. Following this observation, we compute local planar surfaces using the 3D statistical distribution described in the previous section. While a naive implementation would fit planar surfaces to each voxel independently, this would inherently lead to a noisy reconstruction since some voxels may have very few points, if not none, due to the heterogeneous density of Lidar data.

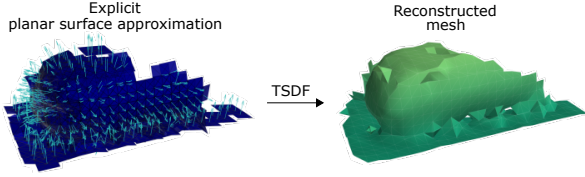


Fig. 3. Explicit local planar surfaces with their estimated normals for  $k = 1$  (left) and its corresponding mesh reconstruction from the TSDF (right). Note that applying  $k = 1$  neighborhood leads to a noisy surface estimation (e.g. car edges) that is overcome with our adaptive neighborhood strategy.

In our pipeline we propose using an adaptive neighborhood definition where local surfaces are estimated from multi-scale voxels statistics. Our neighborhood definition presents two main advantages: a) it increases the statistical robustness which improves large planar surface estimation (e.g. ground, walls), b) it counterbalances the lack of local data due to low density or occlusion.

1) *Neighborhood definition:* We define the multi-scale neighborhood at the vertices location rather than voxels, since implicit surfaces will later be estimated at each vertex of the voxel-grid representation. Let's consider  $v$  a vertex from the voxel-grid representation, its 8 adjacent voxels make up the first neighborhood level denoted  $\mathcal{H}^1(v)$ . Subsequently, the union of  $\mathcal{H}^1(v)$  and the voxels adjacent to  $\mathcal{H}^1(v)$  make up the neighborhood  $\mathcal{H}^2(v)$ . More formally, the neighborhood  $\mathcal{H}^k(v)$  is composed of the  $(2k)^3$  nearest voxels surrounding  $v$ . A two-dimensional illustration of  $\mathcal{H}^k(v)$  at levels  $k = 1$  and  $k = 2$  is presented in Fig. 5.

Following [8], for a given neighborhood  $\mathcal{H}$  (indices are dropped for clarity) we obtain the cardinal  $|\mathcal{H}|$ , statistical mean  $\mathcal{H}_\mu$ , and covariance  $\mathcal{H}_\sigma$  from the merging of statistical data of all voxels in  $\mathcal{H}$ .

2) *Planar estimation:* Having obtained the local statistics, we use the covariance  $\mathcal{H}_\sigma$  to estimate the local planar surface of  $\mathcal{H}$  through a Principal Component Analysis (PCA) if the following equation is satisfied:

$$|\mathcal{H}| \geq N_{min}, \quad (1)$$

where  $N_{min}$  is a hyper parameter. Suppose  $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$  the eigen vectors, and  $(\lambda_1, \lambda_2, \lambda_3)$  the eigen values with  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ , we define  $\vec{e}_3$  (the least dominant axis) as the unoriented normal of the planar estimation of the surface at neighborhood  $\mathcal{H}$ . Since normals need to be consistently oriented, the normal  $\vec{n}$  of the plane is oriented towards the sensor pose  $s_p$  as follows:

$$\vec{n} = \begin{cases} \vec{e}_3 & \text{if } \vec{e}_3 \cdot (s_p - \mathcal{H}_\mu) > 0, \\ -\vec{e}_3 & \text{otherwise.} \end{cases} \quad (2)$$

where  $\cdot$  stands for the dot product. We denote  $\Pi$  the plane formed by the pair of normal and center  $(\vec{n}, \mathcal{H}_\mu)$ . An example of the local planar surfaces estimation is visible in Fig. 3.

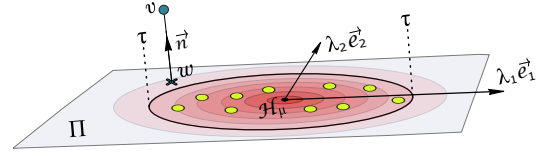


Fig. 4. The likelihood of  $w$  belonging to  $\Pi$  is estimated from  $\mathcal{N}_{PDF}(w | \mathcal{H}_\mu, \Sigma)$  shown in red. The likelihood must be higher than  $\tau$  in order to consider  $\Pi$  as a valid plane for  $w$ . ( $k$  indices dropped for clarity.)

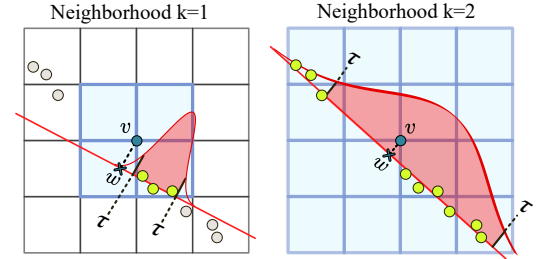


Fig. 5. Analogous 2D representation of the dynamic neighborhood. Planar surface approximations are performed at different neighborhood levels  $k \in [1, k_{max}]$ . The considered plane corresponds to the minimum neighborhood level that satisfies Eq. 3, where the schematic 1D Gaussian distribution is shown in red.

### C. Implicit global surface

To reconstruct the global continuous surface, we compute the Truncated Signed Distance Field (TSDF) for each vertex  $v \subseteq V$  close to the scanned surface.

To cope with the varying density of points in the point cloud, we first compute an optimal neighborhood level  $k'$  at each vertex, and then estimate the TSDF value given this optimal neighborhood.

1) *Adaptive neighborhood:* A naive implementation of our neighborhood definition would use a constant  $k$  throughout the scene. However, large  $k$  values will over smooth high density regions, while small values of  $k$  will lead to noisy estimation in low density regions. Instead, we compute the optimal neighborhood level  $k'$  for each vertex  $v$ , given the probability of  $v$  to belong to a multivariate Gaussian distribution projected on  $\Pi^{k'}$  (the neighborhood planar estimation).

For each vertex, we calculate the projection  $w^k$  of  $v$  onto the plane  $\Pi^k(v)$  and evaluate the likelihood of this projection to belong to the Gaussian  $\mathcal{N}^k$ , where  $\mathcal{N}^k$  is the 2D planar-Gaussian of the statistical distribution of  $\mathcal{H}^k$  projected onto  $\Pi^k$ , as illustrated in Fig. 4. The optimal neighborhood level  $k'$  is defined as the smallest level for which the projection of  $v$  onto  $\Pi^k$  has a probability to belong to the Gaussian  $\mathcal{N}^k$  greater than  $\tau$  (a hyper parameter). Formally, it is the smallest integer for which the probability density function  $\mathcal{N}_{PDF}$  satisfies

$$\mathcal{N}_{PDF}^k(w^k | \mathcal{H}_\mu^k, \Sigma) \geq \tau; \quad \Sigma = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \quad (3)$$

In practice, we compute the optimal  $k'$  iteratively starting at level 1 and stops when the above equation is satisfied as it is shown in Fig. 5. To avoid exponential computation time,  $k$  is bounded between  $[1, k_{max}]$ .

2) *TSDF computation*: We now compute the TSDF value for each vertex of the voxel grid, from the optimal local neighborhood level, while accounting for the normal estimation at the corresponding level. In other words, we compute  $\text{TSDF}(v)$  such that:

$$\text{TSDF}(v) = \vec{n}^{k'} \cdot (v - \mathcal{H}_\mu^{k'}). \quad (4)$$

Our adaptive neighborhood level selection can efficiently handle varying local density and fill gaps between missing data, such as those between two adjacent Lidar layers. Subsequently, the condition on the probability density function given in Eq. 3 avoids the extension of the surface at its borders which is visible in qualitative results.

After TSDF computation, we use the popular marching cubes [10] to extract the mesh from the zero-crossing level of the gradient field. Fig. 3 shows the reconstructed mesh with constant neighborhood ( $k = 1$ ).

#### IV. EXPERIMENTAL RESULTS

We evaluate our proposal on synthetic data from CARLA [22] and real data from KITTI [23], and compare our performance against the IMLS baseline [18]. For both synthetic and real data we used 100 frames equi-sampled from either an urban-like sequence for CARLA, or the residential sequence 0018 for the public KITTI dataset.

##### A. Methodology

Unless stated otherwise, our hyper-parameters remain unchanged in all experiments, with:  $\omega = 0.2m$ ,  $\tau = 0.2$ ,  $N_{min} = 10$ , and  $k_{max} = 5$ . For fair comparison, the spherical neighborhood radius of IMLS is set to  $\omega \times k_{max} = 1m$ , and its  $k$ -nearest neighbor search to  $N_{min}$ . The noise parameter of IMLS is set to  $h = 1/3(\omega \times k_{max}) = 0.33$ .

Noteworthy, it is impractical to have a real mesh ground-truth and KITTI can only be used for qualitative evaluation. Conversely, we use synthetic data from CARLA to evaluate the benefit of each of our contributions. The setup in CARLA replicates a top-mounted Velodyne HDL64E, similarly to the real KITTI setup. We also simulate a collocated noise-free lidar with abnormally high resolution (316 layers), which serves as ground-truth for the reconstruction. Consequently, we frame the evaluation as a set-to-set distance problem, and measure the quality as the distance of the predicted mesh vertices (P) to the ground-truth (GT) set.

We use two metrics derived from the literature: the Average Error (AE), and the Hausdorff Distance (HD). The former measures the average distance error from one point to its nearest point in the other set, that is:  $AE_{P \rightarrow GT} = \sum_{a \in P} \frac{1}{|P|} \min_{b \in GT} |a - b|$ . The Hausdorff distance [24] is classically used for point set distances and gives a sense of the largest minimum error, that is:  $HD_{P \rightarrow GT} = \max_{a \in P} \min_{b \in GT} |a - b|$ . As each of the two metrics are directed, we also report the symmetrical metric, as the average of both directed metrics. For Hausdorff,  $HD_{sym} = 0.5(HD_{P \rightarrow GT} + HD_{GT \rightarrow P})$ . We chose not to use the

TABLE I  
PERFORMANCE ON SYNTHETIC DATA

| Method     | Average Error (m) |      |      | Hausdorff Distance (m) |      |             |       |
|------------|-------------------|------|------|------------------------|------|-------------|-------|
|            | P→GT              | GT→P | Sym  | P→GT                   | GT→P | Sym         |       |
| IMLS [18]  | 0.37              | 0.09 | 0.23 | 4.54                   | 8.32 | <b>6.43</b> |       |
| Ours AN+GC | $k_{max} = 5$     | 0.14 | 0.13 | <b>0.14</b>            | 1.39 | 20.49       | 10.94 |
|            | $k_{max} = 3$     | 0.09 | 0.26 | 0.17                   | 0.69 | 30.84       | 15.77 |
| Ours AN    | $k_{max} = 5$     | 0.30 | 0.12 | 0.21                   | 1.69 | 20.36       | 11.02 |
|            | $k_{max} = 3$     | 0.14 | 0.25 | 0.19                   | 0.87 | 30.83       | 15.85 |
| Ours CN+GC | $k = 5$           | 0.15 | 0.16 | 0.15                   | 1.38 | 20.50       | 10.94 |
|            | $k = 3$           | 0.09 | 0.27 | 0.18                   | 0.69 | 30.85       | 15.77 |
|            | $k = 1$           | 0.03 | 3.44 | 1.73                   | 0.24 | 65.28       | 32.76 |
| Ours CN    | $k = 5$           | 0.30 | 0.14 | 0.22                   | 1.69 | 20.36       | 11.03 |
|            | $k = 3$           | 0.14 | 0.26 | 0.20                   | 0.87 | 30.83       | 15.85 |
|            | $k = 1$           | 0.03 | 3.44 | 1.73                   | 0.24 | 65.28       | 32.76 |

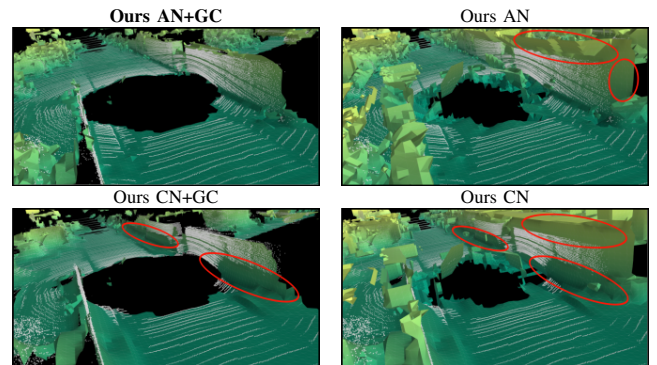


Fig. 6. Qualitative comparison of our method by removing the main components of our proposal. Images correspond to a single frame of a point cloud from CARLA. In shown images  $k$  or  $k_{max}$  equals 5.

Chamfer distance used in machine learning [25] because it isn't a metric-scale and is thus harder to interpret intuitively.

##### B. Ablation study

We assess the average performance on 100 frames from CARLA and subsequently report quantitative performance in Table I and qualitative results in Fig. 6. In the former, our pipeline is compared to IMLS [18].

To evaluate the importance of our contributions, we compare the benefits of our Adaptive Neighborhood (AN, Sec. III-C.1) with or without the Gaussian Confidence (GC, Eq. 3) and also with a naive Constant Neighborhood (CN) approach. More precisely, the neighborhood  $\mathcal{H}^{k'}$  considered for the plane estimation (cf. Sec III-B.2) is different:

**AN+GC**:  $\mathcal{H}^{k'}$  is the minimum neighborhood among  $\{\mathcal{H}^1, \dots, \mathcal{H}^{k_{max}}\}$  that satisfies Eq. 1 and 3.

**AN**:  $\mathcal{H}^{k'}$  is the minimum neighborhood among  $\{\mathcal{H}^1, \dots, \mathcal{H}^{k_{max}}\}$  that satisfies Eq. 1.

**CN+GC**:  $\mathcal{H}^{k'}$  with  $k' = k$ , is considered only if Eq. 1 and 3 are satisfied.

**CN**:  $\mathcal{H}^{k'}$  with  $k' = k$ , is considered only if Eq. 1 is satisfied.

From Table I, the accuracy on the reconstruction directly affects the P→GT distances while the GT→P distances are mostly influenced by the reconstruction density.

As expected, the benefit of our adaptive neighborhood strategy is noticeable when comparing AN+GC and CN+GC. Not only AN+GC exhibits higher accuracy (lower P→GT) but it also increases the reconstruction density (lower

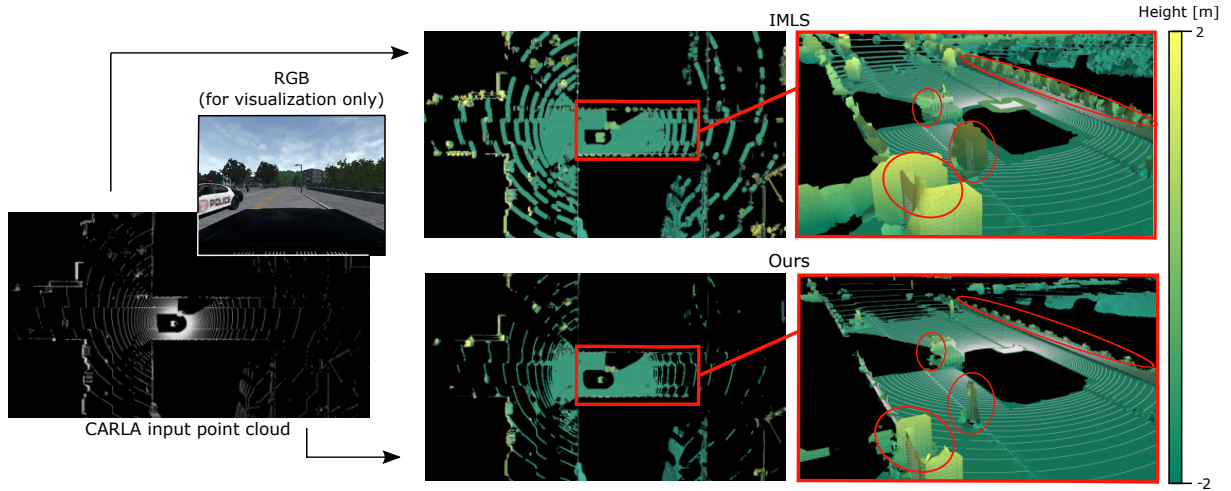


Fig. 7. Qualitative comparison on CARLA simulator. Notice that even though IMLS outputs a denser reconstruction, it also extend all surfaces at its border and creates a higher number of artifacts as it can be seen by the red circles at rightmost images. Our method is able to keep the structure of the surface and generates fewer artifacts, performing a more accurate reconstruction, while keeping a good density in areas near to the vehicle.

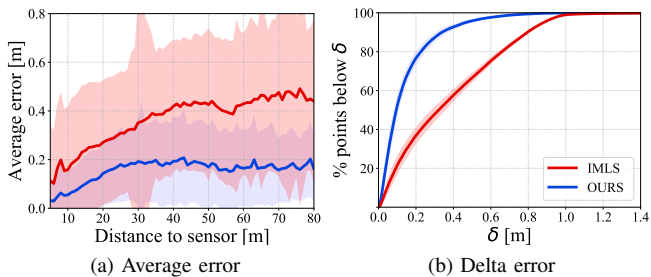


Fig. 8. Performance on the CARLA dataset from averaging of 100 frames evaluation. While the average error grows with the distance, at 30m distance our surface reconstruction error is  $\approx 0.2m$  whereas IMLS [18] is  $\approx 0.35m$ .

GT $\rightarrow$ P). When using our Gaussian Confidence (GC), there is a slight density loss (lower GT $\rightarrow$ P) but a significant accuracy increase. Qualitatively from Fig. 6, our adaptive neighborhood helps to maintain the details of the surface by not over-smoothing the data, while the Gaussian Confidence strategy avoids to extend the surface, keeping its structure and generating fewer artifacts.

Overall, we observe that our complete pipeline (AN+GC with  $k_{\max}=5$ ) keeps a good trade-off between accuracy and density, which is shown by the best result obtained  $AE_{\text{sym}}$  ( $0.14m$  vs.  $0.23m$  for IMLS). With larger neighborhood (bigger  $k_{\max}$ ), the density of the reconstruction increases but at the expense of a lower accuracy which is intuitive as there is a need to extrapolate more the data. Since IMLS performs a denser reconstruction, lower GT $\rightarrow$ P distances are obtained, which explains the best results on the symmetric Hausdorff Distance ( $10.94m$  vs.  $6.43m$  for IMLS). While our method is less dense, our predicted mesh is significantly more precise ( $1.39m$  vs.  $4.54m$  for IMLS).

### C. Synthetic data

To further evaluate the performance, we report the average error as a function of the sensor distance in Fig.8a, and the cumulative delta error that indicates the percentage of

vertices of the output meshes having an error lower than a given value in Fig. 8b.

For both metrics, the accuracy of our reconstruction is significantly higher than the one obtained by IMLS, as the average error of our reconstruction is often 50% lower (Fig. 8a). The percentage of vertices below a given error exhibits the same behavior (Fig. 8b), with a significant advantage for our method. Furthermore, almost 80% of the vertices of our mesh have an error lower than  $0.2m$ , while 40% of vertices lie below the same threshold with IMLS.

Finally, a qualitative comparison on CARLA is shown in Fig. 7. Again, one can see that IMLS outputs a more dense reconstruction of the scene but also that the method tends to extend all surfaces, generating artifacts and inaccuracies in the reconstruction. This can be observed in the circled areas where lampposts, traffic signs and walls are abnormally enlarged by IMLS. On the other hand, our method is able to keep these details on the reconstruction and maintains a high density in areas close to the vehicle, which confirms the quantitative results obtained in Fig. 8.

### D. Real data

Qualitative results on KITTI residential sequence 0018 are shown in Fig. 9. As for the results presented on the synthetic data, IMLS outputs a denser reconstruction but extends all surfaces out of the borders and generates artifacts. This can be observed at the circled areas on the rightmost images. Conversely, our method outputs a more accurate reconstruction, keeping the structure of the scanned surface and completing the surface at no data points areas close to the sensor. Even though our method outputs a less dense reconstruction, the density is still adequate for further robotics tasks such as trajectory planning or obstacle avoidance.

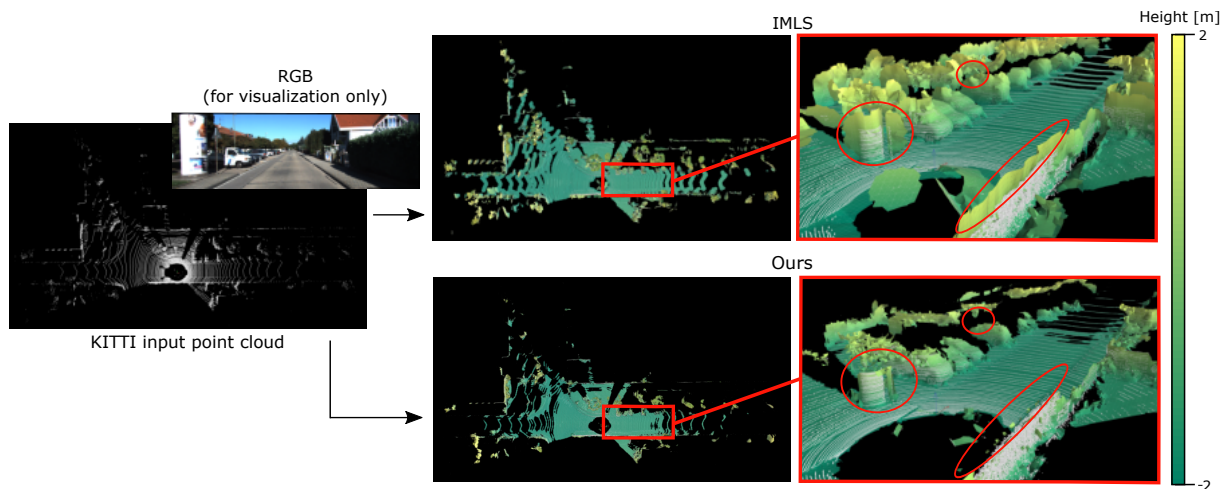


Fig. 9. Visual comparison on KITTI dataset. The results on both methods show the same behavior than seen in synthetic data. IMLS performs a denser reconstruction at expenses of accuracy. Our method keeps a good trade-off between density and accuracy on the reconstruction.

## V. CONCLUSION

In this paper we presented a surface reconstruction method based on a light voxel representation that benefits of statistical information. Our pipeline uses an adaptive neighborhood strategy coupled with a Gaussian confidence estimation, to best estimate local surfaces and the implicit surface estimation (TSDFs).

The performance of our method was demonstrated on both synthetic (CARLA) and real data (KITTI), significantly outperforming the classical IMLS though sparser reconstruction. For mobile robots, an accurate representation of the environment is a prerequisite. While our results show that local planar surfaces are sufficient for accurate reconstruction, we intend to extend this research to a more complex surface definition, for example with polynomials.

## REFERENCES

- [1] T. Stoyanov, M. Magnusson, and A. J. Lilienthal, "Comparative evaluation of the consistency of three-dimensional spatial representations used in autonomous robot navigation," *Journal of Field Robotics*, vol. 30, no. 2, pp. 216–236, 2013.
- [2] M. Berger, A. Tagliasacchi, L. M. Seversky, P. Alliez, G. Guennebaud, J. A. Levine, A. Sharf, and C. T. Silva, "A survey of surface reconstruction from point clouds," *Comput. Graph. Forum*, vol. 36, no. 1, pp. 301–329, Jan. 2017.
- [3] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *IEEE International Conference on Robotics and Automation*, vol. 2, Mar. 1985, pp. 116–121.
- [4] J. Bares, M. Hebert, T. Kanade, E. Krotkov, T. Mitchell, R. Simmons, and W. Whittaker, "Ambler: an autonomous rover for planetary exploration," *IEEE Computer*, vol. 22, no. 6, pp. 18–26, Jun. 1989.
- [5] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: an efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, Apr. 2013.
- [6] C. A. Vanegas, D. G. Aliaga, and B. Benes, "Automatic extraction of manhattan-world building masses from 3D laser range scans," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 10, pp. 1627–1637, Oct. 2012.
- [7] R. Schnabel, R. Wahl, and R. Klein, "Efficient ransac for point-cloud shape detection," *Comput. Graph. Forum*, vol. 26, no. 06, pp. 214–226, 2007.
- [8] J. Ryde, V. Dhiman, and R. Platt, "Voxel planes: Rapid visualization and meshification of point cloud ensembles," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013, pp. 3731–3737.
- [9] J. Behley and C. Stachniss, "Efficient surfel-based slam using 3D laser range data in urban environments," in *Robotics: Science and Systems (RSS)*, Pittsburgh, Pennsylvania, USA, Jun. 2018.
- [10] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *SIGGRAPH*, Jul. 1987, pp. 163–169.
- [11] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *SIGGRAPH*, 1996, pp. 303–312.
- [12] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *10th IEEE International Symposium on Mixed and Augmented Reality*, Oct. 2011, pp. 127–136.
- [13] T. Whelan, M. Kaess, M. F. Fallon, H. Johannsson, J. J. Leonard, and J. McDonald, "Kintinuous: Spatially extended kinectfusion," in *RSS Workshop on RGB-D: Advanced Reasoning with Depth Cameras*, Sydney, Jul. 2012.
- [14] F. Steinbrucker, C. Kerl, D. Cremers, and J. Sturm, "Large-scale multi-resolution surface reconstruction from rgb-d sequences," in *IEEE International Conference on Computer Vision (ICCV)*, Dec. 2013, pp. 3264–3271.
- [15] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Fourth Eurographics Symposium on Geometry Processing (SGP'06)*, Cagliari, Sardinia, Italy, Jun. 2006, pp. 61–70.
- [16] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Trans. Graph.*, vol. 32, no. 3, pp. 29:1–29:13, Jul. 2013.
- [17] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," in *SIGGRAPH*, vol. 26, no. 2, Jul. 1992, pp. 71–78.
- [18] C. Shen, J. F. O'Brien, and J. R. Shewchuk, "Interpolating and approximating implicit surfaces from polygon soup," in *SIGGRAPH*, Los Angeles, USA, Aug. 2004, pp. 896–904.
- [19] R. Kolluri, "Provably good moving least squares," *ACM Trans. Algorithms*, vol. 4, no. 2, pp. 18:1–18:25, May 2008.
- [20] H. Bouchiba, S. Santoso, J. Deschaud, L. Rocha-Da-Silva, F. Goulette, and T. Coupez, "Computational fluid dynamics on 3D point set surfaces," *CoRR*, vol. abs/1901.04944, 2019.
- [21] L. Roldao, R. de Charette, and A. Verroust-Blondet, "A statistical update of grid representations from range sensors," *arXiv preprint arXiv:1807.08483*, 2018.
- [22] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.
- [23] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Rob. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.

- [24] N. Aspert, D. Santa-Cruz, and T. Ebrahimi, "Mesh: measuring errors between surfaces using the hausdorff distance," in *Proceedings. IEEE International Conference on Multimedia and Expo*, vol. 1, Aug 2002, pp. 705–708 vol.1.
- [25] H. Fan, H. Su, and L. J. Guibas, "A point set generation network for 3d object reconstruction from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 605–613.