

# Efficient Mobile Robot Navigation using 3D Surfel Grid Maps

Jochen Kläß, Jörg Stückler, and Sven Behnke  
Autonomous Intelligent Systems, University of Bonn, Germany

## Abstract

We present robust and efficient means for mobile robot navigation using a 3D representation of the environment. We build 3D surfel grid maps and propose Monte Carlo localization with probabilistic observation models for 2D and 3D sensors in such maps. In contrast to localization methods that utilize a 2D laser scanner in a static horizontal mounting, our method takes advantage of the 3D structure in the environment. This is useful, for instance, to localize in crowds of people: The robot can focus on the static parts of the environment above the person's height. Finally, we extract navigation maps for 2D path planning from the 3D maps. Our approach avoids discretization effects and considers the complete height range of the robot to estimate traversability. In experiments, we demonstrate the accuracy and robustness of our approach for pose tracking and global localization – even in a crowded environment.

## 1 Introduction

Many approaches to indoor localization and mapping use 2D laser scanners to acquire 2D footprints of the environment. Occupancy grid maps are used to represent the map, because they provide dense information about free and occupied space for localization and path planning. One problem of such 2D maps appears in path planning, when obstacles cannot be perceived on the laser scanner's height. Localization with 2D laser scanners imposes further restrictions when dynamic objects occur, or the environment changes in the scan plane of the laser. Then, localization may fail since large parts of the measurements are not explained by the map.

We address these problems by building 3D maps of the environment and by localizing in these maps (see Fig. 1). We choose to represent the map in a 3D grid in which each voxel maintains the 3D distribution of points by mean and covariance (denoted as *surfel*). The robot acquires such maps from several 3D scans. Once the map has been obtained, we extract a 2D navigation map that considers obstacles in the complete height range of the robot. Our approach avoids discretization effects by making use of the continuous occupancy information in the map.

For localization, we developed a Monte Carlo method that can incorporate full 3D scans as well as 2D scans. When used with 3D scans, we extract surfels from the scans and evaluate their observation likelihood. From 2D scans, we extract line segments and associate them with surfels in the map. The localization in 3D maps is specifically useful in crowded environments. The robot can then take advantage of measurements above the height of people to localize at the static parts of the environment.

## 2 Related Work

Research on mobile robot localization in 2D maps has a long tradition (e. g., [1, 2, 3]). In recent years, approaches have been proposed that use 3D maps [10, 8, 7, 5, 6].

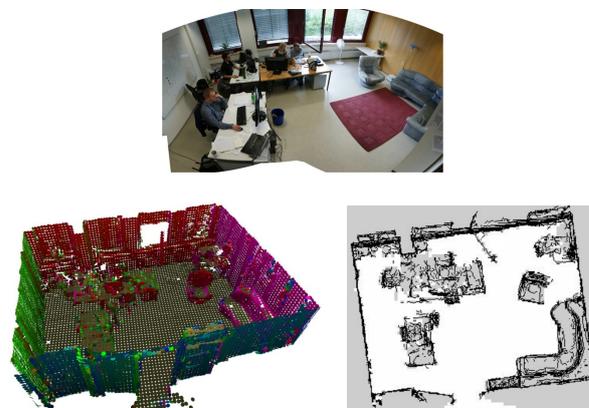


Figure 1: Top: Panorama image of an office environment. Bottom left: 3D surfel map acquired with our approach (surfel orientation coded by color). Bottom right: 2D navigation map extracted from the 3D surfel map.

Kuemmerle et al. [8] apply Monte Carlo localization in Multi-Level Surface (MLS) maps [11]. These maps extend elevation maps to store height intervals of occupied space in a 2D grid. They propose an approximate end-point sensor model for individual sensor beams in the MLS map. In order to gain efficiency, they sample 3D points from the segments in the MLS map. Expected measurements for beam end-points are then obtained by fast closest point queries in a k-d tree of the sample points. In our framework, an approximation by samples is not necessary.

Octrees constitute a further memory-efficient data structure for 3D map representation which has been applied for occupancy mapping [13]. For this kind of map, Hornung et al. [7] developed efficient means for 6-DoF localization of a humanoid robot. They integrate postural measurements from an IMU and joint feedback with 2D laser scans to localize within the OctoMap. For the laser measurements, they use the end-point model. In contrast to uniform grids, the octree data structure does not allow for precalculation of the sensor model in a likelihood field. Instead, closest point queries have to be executed in the octree. Re-

garding the map representation, traversability can hardly be estimated in low resolutions (e. g., 5 cm voxel size) due to step effects. This effect must be handled by choosing a higher resolution which is less memory-efficient. In our map representation, we model local surface patches in each voxel. Despite a relative coarse resolution of the voxel grid (10 cm), our maps provide an accurate estimate of surfaces like the floor plane.

For localization in dynamic environments, Fox et al. [3] handle misleading measurements on dynamic objects by incorporating them into the sensor model. Other approaches classify the environment into static and dynamic parts (e. g. [12]). Our approach avoids the use of misleading measurements and focuses on the measurements above the person’s height.

### 3 3D Surfel Grid Maps

We discretize the environment in a voxel map with a fixed resolution (10 cm in our experiments). In each voxel we maintain a local surface element (surfel) which we represent as a normal distribution of 3D point measurements within the voxel.

We acquire maps by integrating 3D scans obtained from several locations in the environment. A 2D SLAM approach provides us with a drift-free pose estimate for mapping. We apply gMapping [4] for this purpose. In order to obtain an accurate 6-DoF registration of the point clouds, we match the 3D scans using ICP.

For each surfel in the map, we determine the possible view directions onto the estimated surface. We estimate surface normals by the eigenvector to the smallest eigenvalue of the 3D sample covariance. If a surfel is viewed from a scan pose, we memorize the view direction along the surface normal.

Spurious noise or measurements on dynamic objects may be incorporated during mapping. We thus perform occupancy mapping. From each sensor pose, we trace sensor rays to update the occupancy belief in the swept voxels. In this process, we exploit the detailed surface information represented in our voxels (see Fig. 2). We consider a voxel as observed occupied, if a ray ends in the voxel. If a ray crosses a voxel, we distinguish two cases: The ray may either miss the surfel, or intersect the planar patch defined by the surfel mean and its normal. In the latter case, the voxel is observed free.

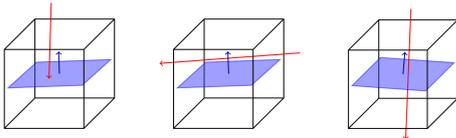


Figure 2: Possible measurements in a voxel containing a surfel. Left: Ray ends in voxel. Middle: Ray passes voxel but misses the surfel. Right: Ray passes a voxel and intersects the surfel.

## 4 Monte Carlo Localization in 3D Surfel Grid Maps

We developed efficient Monte Carlo localization in 3D surfel grid maps. Our method can incorporate full 3D scans from sensors such as 3D laser scanners or depth cameras as well as 2D scans in arbitrary scan position and orientation. By integrating 2D scans, our method can also be used for mobile robots that are only equipped with a 2D laser scanner. Furthermore, the robot could continuously sweep a tilting 3D laser scanner and integrate each 2D laser scan instantly.

### 4.1 Sensor Models

**Plane-to-Plane Model:** For the use of 3D scans, we discretize the scan  $z_t$  into a voxel grid and extract surfels  $z_{t,i}$  from the scan. We evaluate the observation likelihood of the scan with regard to the surfel map by the observation model

$$p(z_t|x_t, m) = \prod_i [\alpha_{\text{rand}} p_{\text{rand}}(z_{t,i}|x_t, m) + \alpha_{\text{hit}} p_{\text{hit}}(z_{t,i}|x_t, m)], \quad (1)$$

where  $x_t$  is the robot pose,  $m$  is the surfel map, and the observation model includes two components that explain the measurement. The first component  $p_{\text{rand}}$  models random false measurements, while

$$p_{\text{hit}}(z_{t,i}|x_t, m) = p_d(z_{t,i}|x_t, m) p_\alpha(z_{t,i}|x_t, m) \quad (2)$$

measures the observation likelihood, if the scan surfel corresponds to a surfel in the map. The likelihood  $p_d(z_{t,i}|x_t, m)$  compares the distance between the surfels by the deviation under their normal distributions, i. e.,

$$p_d(z_{t,i}|x_t, m) = \mathcal{N}(\mu_m - T(x_t) \mu_i; 0, \bar{\Sigma}_m + R(x_t) \bar{\Sigma}_i R(x_t)^T). \quad (3)$$

Similar to the work in [9], we assume local planar structures in our map and therefore flatten the surfel covariances  $\Sigma_m$  and  $\Sigma_i$  along their normals, i. e.,

$$\bar{\Sigma} := RDR^T, \quad R = \begin{pmatrix} n & v_1 & v_2 \end{pmatrix}, \quad D = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (4)$$

where  $n$  is the surfel normal,  $v_1$  and  $v_2$  are the first and second principal axes of the surfel covariance, and  $\epsilon$  is a small positive constant. By this, we only account for deviations along the surface normals.

In addition, the model incorporates the view direction onto the surfel and the surfel orientation by the observation of the surface normal

$$p_\alpha(z_{t,i}|x_t, m) = \mathcal{N}(\langle n_i, n_m \rangle; 0, \sigma_\alpha^2), \quad (5)$$

where we set  $\sigma_\alpha$  empirically.

**Line-to-Plane Model:** We integrate planar scans in an analogous way to the case of 3D scans. We find line elements  $z_{t,i}$  in the 2D scan and determine the likelihood of the line elements in the 3D surfel map using Eq. (1). Now, we modify the covariance  $\Sigma_i$  of the line element to not measure distance along the line, i.e.,

$$\bar{\Sigma} := RDR^T, \\ R = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{pmatrix}, \quad (6)$$

where  $v_j$  are the eigenvectors of the line element covariance sorted in descending order of the eigenvalues and  $\epsilon$  is a small positive constant. Note, that the line element is compared with a flattened surfel in the map, such that deviations are still not measured in the plane of the surfel. The orientation of the surface normal cannot be considered in the 2D case.

## 4.2 Data Association

For the efficient association of measurements to map surfels, we precalculate nearest surfel neighbors in the voxel map. The metric  $d(p, \mu, \Sigma) = (p - \mu)^T \bar{\Sigma}^{-1} (p - \mu)$  for a point  $p$  and a surfel  $(\mu, \Sigma)$  assumes planar structures in the map. Since the robot is moving in the horizontal plane, we only consider nearest neighbors within the same horizontal slice in the map.

## 5 Navigation Maps

Path planning requires maps that contain information about traversable, occupied, and unknown space. We extract 2D navigation maps from our surfel maps.

Our approach avoids many problems that one typically encounters in occupancy voxel maps (see Fig. 3). The discretization in such maps may introduce virtual untraversable steps. Furthermore, actual untraversable bumps could not be represented that are smaller than the resolution of the map. Depending on the characteristics of the mobile robot, this would require choosing a high resolution for mapping. High resolutions not only increase memory-requirements, but the resolution of the 3D sensor may not be large enough to sample the map densely. A further advantage of our representation is, that we can exploit the de-

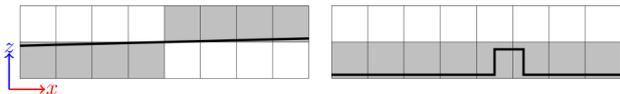


Figure 3: Typical discretization effects in occupancy voxel maps. (a) The discretization introduces steps and (b) small bumps are neglected.

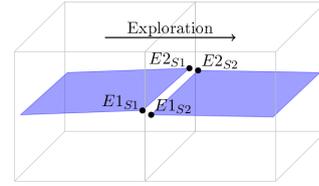


Figure 4: We explore surfel maps for traversable voxels. We examine the transition points ( $E k_{S1}$ ) between surfels ( $S1$  and  $S2$ ) to detect untraversable gaps.

tailed information in our voxel cells to generate navigation maps in high resolutions.

**Exploration of the Traversable Surface:** Starting from the scan poses, we explore the traversable surface by region growing in the map. We traverse between voxels, if

- the surfel in the next voxel is almost horizontal,
- the gap between the surfels at the common voxel border plane is small (see Fig. 4), and
- no voxel is either occupied nor unknown within the height range of the robot.

**Interpolation of High-Resolution Navigation Maps:**

We project the surfels into the horizontal plane and rasterize the projected lines in the 2D navigation grid map. We also include the information on the possible view directions onto the surfels. If one side of a surfel has not been viewed, we mark the corresponding cells unknown in the navigation map. An example map is shown in Fig. 1.

## 6 Experiments

We compare our approach (10 cm res.) to Monte Carlo localization in 2D occupancy grid maps (5 cm res.) which uses scans of a horizontally mounted laser. First, we assess our approach in a static environment. We then demonstrate our approach in a dynamic environment in which multiple persons walk randomly.

### 6.1 Localization in Static Environments

In a static office environment, we compare the pose tracking accuracy of the methods. In our approach, we continuously tilt a laser scanner in a height of 1.15 m to obtain laser scans in varying orientations. We integrate these scans at high update rates into the localization estimate.

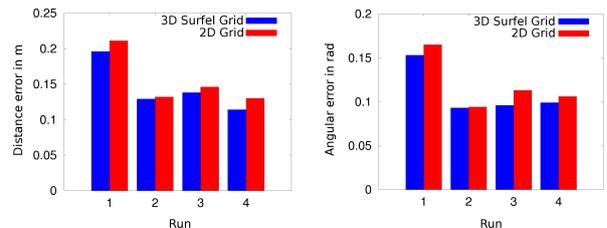


Figure 5: Comparison of the average tracking error in an office between our approach with a continuously tilting laser scanner (blue) and MCL with a horizontal laser (red).

Fig. 5 shows the accuracy of pose tracking in 4 runs within an office environment. Our approach yields slightly lower distance and angular errors. Our method is also similarly efficient like MCL and, hence, real-time capable, although the use of 3D maps and the extraction of surfels and line segments from scans introduces a small overhead. When integrating line segments from 2D scans, our approach performs updates for 500 particles within ca. 18 msec on a quadcore CPU (MCL: 10 msec).

## 6.2 Localization in Dynamic Environments

In a dynamic setting, we compare standard MCL with our approach which focuses on measurements above the persons' heights. Eight persons were randomly walking in the test environment. We quantify, how often and how accurately the localization methods estimate the final position of the trajectory. If with our approach the robot stands during a full sweep, the complete 3D scan is integrated. Otherwise, we use the 2D scans instantly.

For pose tracking, we initialize the methods with ground truth. For global localization, we initialize the methods with a uniform distribution of 5000 particles. We evaluated global localization at 45 starting points in 5 trajectories.

Fig. 6 shows results of this experiment. It can be seen that our approach localizes the robot more accurate. Global localization in the 2D map only succeeds in about 30% of the runs, whereas our approach achieves 97.5% success rate at a distance threshold of 0.5 m. While our approach yields superior results in accuracy, it still retains the efficiency of 2D localization.

## 7 Conclusion

We presented efficient means for navigation using 3D surfel grid maps. The map discretizes the environment and estimates occupancy in each cell. In contrast to standard occupancy grid maps, it maintains a continuous estimate of the occupied space. We exploit this to better judge traversability and to generate high-detail navigation maps. We proposed probabilistic sensor models of 2D and 3D scan measurements in our map. We integrated these sensor models in a Monte Carlo localization framework. In experiments, we could demonstrate that our method performs similarly well like a standard MCL approach with a horizontal 2D laser scanner in a static environment. In dynamic environments such as crowds of people, our approach yields superior results, since it can focus on the static parts of the environment.

The overhead in using our 3D approach amounts only in a small constant factor (ca. 2) towards 2D MCL for a measurement update in the particle filter.

In future work, we will extend our approach to localize a robot in six degrees of freedom (6-DoF). For the navigation in 6-DoF, we will also generate 3D navigation maps. Finally, we integrate 6-DoF SLAM into our system to make full use of our representation during incremental mapping.

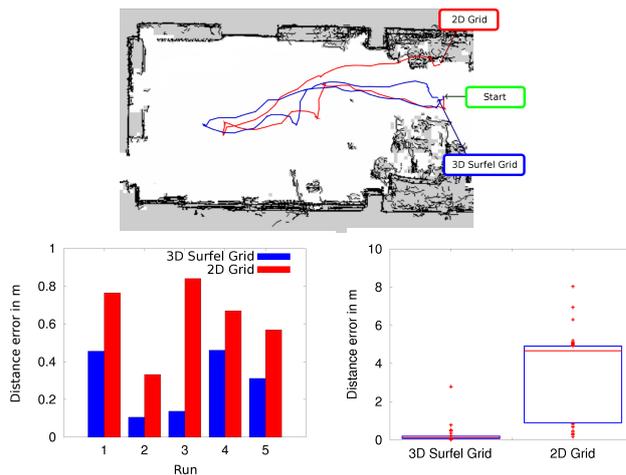


Figure 6: Localization in dynamic environments. Top: Example estimates of 2D MCL (red) and our 3D approach (blue). Bottom Accuracy for pose tracking (left) and after global localization (right).

## References

- [1] F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte Carlo localization for mobile robots. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 1999.
- [2] D. Fox. Adapting the sample size in particle filters through KLD-sampling. *Int. Journal of Robotics Research (IJRR)*, 2003.
- [3] D. Fox, W. Burgard, and S. Thrun. Markov localization for mobile robots in dynamic environments. *J. of Artificial Intelligence Research*, 1999.
- [4] G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Trans. on Robotics*, 2007.
- [5] J.-S. Gutmann, M. Fukuchi, and M. Fujita. A floor and obstacle height map for 3D navigation of a humanoid robot. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2005.
- [6] J.-S. Gutmann, M. Fukuchi, and M. Fujita. 3D perception and environment map generation for humanoid robot navigation. *Int. Journal of Robotics Research*, 2008.
- [7] A. Hornung, K. M. Wurm, and M. Bennewitz. Humanoid robot localization in complex indoor environments. In *Proc. of the IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, 2010.
- [8] R. Kümmerle, R. Triebel, P. Pfaff, and W. Burgard. Monte carlo localization in outdoor terrains using multi-level surface maps. In *Proc. of the Int. Conf. on Field and Service Robotics (FSR)*, 2007.
- [9] A. Segal, D. Haehnel, and S. Thrun. Generalized-ICP. In *Proc. of Robotics: Science and Systems*, 2009.
- [10] S. Thompson, S. Kagami, and K. Nishiwaki. Localisation for autonomous humanoid navigation. In *Proc. of the IEEE-RAS Int. Conf. on Humanoid Robots*, 2006.
- [11] R. Triebel, P. Pfaff, and W. Burgard. Multi-level surface maps for outdoor terrain mapping and loop closing. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2006.
- [12] D. Wolf and G.S. Sukhatme. Online simultaneous localization and mapping in dynamic environments. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2004.
- [13] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard. Octomap: A probabilistic, flexible, and compact 3D map representation for robotic systems. In *Proc. of the IEEE ICRA Workshop on Best Practice in 3D Perception and Modeling for Mobile Manipulation*, 2010.