

Local Multi-Resolution Surfel Grids for MAV Motion Estimation and 3D Mapping

David Droeschel, Jörg Stückler, and Sven Behnke

Computer Science Institute VI, University of Bonn, 53113 Bonn, Germany,
droeschel@ais.uni-bonn.de,
<http://www.ais.uni-bonn.de/Mod>

Abstract. For autonomous navigation in restricted environments, micro aerial vehicles (MAV) need to create 3D maps of their surroundings and must track their motion within these maps. In this paper, we propose an approach to simultaneous localization and mapping that is based on the measurements of a lightweight 3D laser-range finder. We aggregate laser-range measurements by registering sparse 3D scans with a local multiresolution surfel map that has high resolution in the vicinity of the MAV and coarser resolutions with increasing distance, which corresponds well to measurement density and accuracy of our sensor. Modeling measurement distributions within voxels by surface elements allows for efficient and accurate registration of 3D scans with the local map. The incrementally built local dense 3D maps of nearby key poses are registered globally by graph optimization. This yields a globally consistent dense 3D map of the environment. Continuous registration of local maps with the global map allows for tracking the 6D MAV pose in real time. In experiments, we demonstrate accuracy and efficiency of our approach.

1 Introduction

In recent years, micro aerial vehicles (MAV), such as quadrotors, have attracted much attention in the field of aerial robotics. In many of the current MAV applications, these vehicles are tracked by external motion capture systems, fly in obstacle-free altitudes based on global navigation satellite systems (GNSS), such as GPS, or are remotely controlled by a human operator. For autonomous navigation in restricted, GNSS-denied environments, MAVs need to create 3D maps of their surroundings from measurements of onboard sensors and must track their 6D motion within these maps in real time.

The size and weight limitations of MAVs pose a challenge for environment perception, however. Typical lightweight sensors, such as ultrasonic distance sensors, are restricted in measurement range and resolution. Cameras have a limited field-of-view (FoV) and depend on textured environments and lighting conditions. For navigation of ground vehicles, laser range finders (LRF) are popular sensors. Only few MAVs are equipped with 2D LRFs, however [28, 8, 2, 23].

We designed a small and lightweight 3D laser-range finder, making up to 40.000 measurements per second in almost all directions. Aggregation of these

range measurements to globally consistent maps is challenging, because the 6D MAV pose changes continuously during data acquisition.

For MAV-centric measurement aggregation, we use a local multi-resolution map that stores occupancy information and the respective distance measurements. Measurements are stored in grid cells with increasing cell size from the robot’s center. Thus, we gain computational efficiency by having a high resolution in the close proximity to the sensor and a lower resolution with increasing distance, which correlates with the sensor’s characteristics in relative accuracy and measurement density. Compared to uniform grids, local multi-resolution leads to the use of fewer grid cells without losing information and consequently results in lower computational costs.

Aggregating measurements from consecutive time steps necessitates a robust and reliable estimate of the sensor’s motion. Thus, we use the point-based representation in the map to gain an estimate of the sensor’s motion between consecutive 3D scans by scan registration. We propose a highly efficient and accurate registration method that matches Gaussian point statistics in grid cells (surfels) between local multi-resolution surfel maps. For registering 3D scans with a map, we also represent the scans in local multi-resolution grid maps. In order to achieve accuracy despite the sparsity of measurements and the discretization into grids, we assign surfels in a probabilistic way within a Gaussian mixture model (GMM).

The work presented here builds upon our previous work [6] by globally registering these local dense 3D maps by graph optimization. This yields a globally consistent dense 3D map of the environment. Continuous registration of local maps with the global map allows for tracking the 6D MAV pose in real time.

Since laser-based ego-motion estimation relies on structure in the scene, it works best in scenarios where GNNS typically are not available, like in indoor or urban environments.

2 Related Work

Building maps with 3D laser scanners has recently attracted attention in the field of autonomous navigation with mobile ground robots [17, 14]. Laser range sensors provide accurate distance measurements in a large FoV with only minor dependencies on lighting conditions.

A common research topic in SLAM with 3D laser scanners is how to maintain high run-time performance and low memory consumption simultaneously. Hornung et al. [10] implement a multi-resolution map based on octrees (OctoMap). Ryde et al. [18] use voxel lists for efficient neighbor queries. Both of these approaches consider mapping in 3D with a voxel being the smallest map element. Similar to our approach, the 3D-NDT [14] represents point clouds as Gaussian distributions in voxels at multiple resolutions. Our local multi-resolution surfel grids adapt the maximum resolution with distance to the sensor to incorporate measurement characteristics. Moreover, our registration method matches 3D scans on all resolutions concurrently, utilizing the finest common resolution

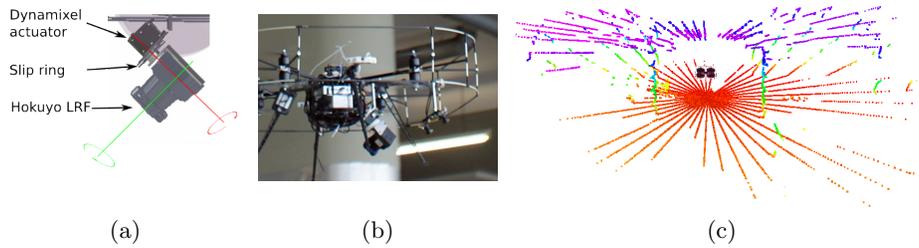


Fig. 1: 3D laser scanner obtained by rotating a 2D LRF. a) CAD drawing; b) mounting on the MAV; c) 3D scan of indoor environment (color encodes height).

available between both maps, which also makes registration efficient. In previous work [25, 19], we used this concept within an octree voxel representation. We model up to six view directions such that multiple 3D scans from different view points can be integrated in a single multi-resolution surfel map. Compared to the dense RGB-D images used in our previous work, the 3D scans obtained from our laser range finder are much sparser. We overcome this sparsity through probabilistic assignments of surfels during the registration process. While the many methods assume the robot to stand still during 3D scan acquisition, some approaches also integrate scan lines of a continuously rotating laser scanner into 3D maps while the robot is moving [4, 7, 24, 13, 1].

Up to now, such 3D laser scanners are rarely employed on lightweight MAVs due to their payload limitations and the difficulty of aggregating 3D scans in-flight. Instead, frequently MAVs are equipped with rigidly mounted two-dimensional laser range finders [9, 28, 8, 2, 23, 11], which restricts their FoV to a mostly horizontal plane at the height of the MAV.

Some MAVs have been equipped with 3D laser scanners. For instance, Scherer and Cover et al. [20, 5] use a 3D laser scanner for obstacle perception for autonomous river exploration. They approach MAV localization with a vision sensor. In contrast, we combine visual odometry with 3D scan registration in a 3D multi-resolution map to localize the MAV. Takahashi et al. [26] also build environment maps with a 3D laser scanner. They localize the robot using GPS and IMU sensors. Thrun et al. [27] propose a 3D mapping system with a rigidly mounted 2D laser scanner on a helicopter. The laser scanner measures in a vertical plane perpendicular to the flight direction. In order to localize the helicopter, measurements from GPS and IMU are fused and consecutive 2D scans are registered, assuming scan consistency in flight direction. In our approach, we do not make such an assumption on scan consistency.

3 Sensor Setup

Our continuously rotating 3D laser scanner consists of a Hokuyo UTM-30LX-EW 2D laser range finder (LRF) which is rotated by a Dynamixel MX-28 servo

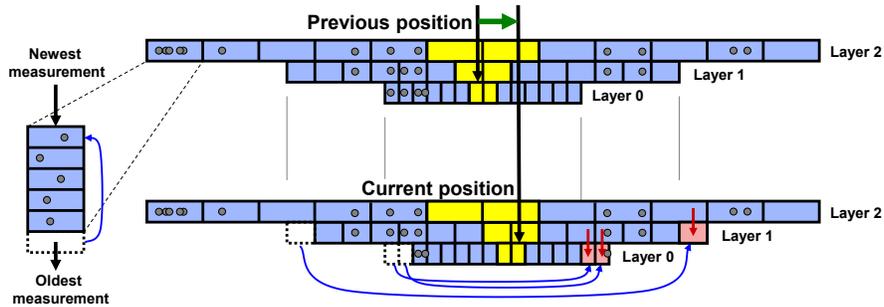


Fig. 2: One-dimensional illustration of the hybrid local multi-resolution map. Along with the occupancy information, every grid-cell (blue) maintains a circular buffer with its associated measurement points (green). The map is centered around the robot and in case of a robot motion, ring buffers are shifted according to the translational parts of the movement, obtaining the egocentric property of the map. Cells at coarser levels are used to retain points from vanishing cells at finer levels and to initialize newly added cells (red arrows).

actuator to gain a 3D FoV. As shown in Fig. 1a, the scanning plane is parallel to the axis of rotation, but the heading direction of the scanner is twisted slightly away from the direction of the axis—in order to enlarge its FoV. The 2D LRF is electrically connected by a slip ring, allowing for continuous rotation of the sensor. The sensor is mounted on our multicopter (Fig. 1b) pitched downward by 45° in forward direction, which places the core of the robot upwards behind the sensor. Hence, the sensor can measure in all directions, except for a conical blind spot pointing upwards behind the robot. The total weight of the 3D scanner is approximately 400 g.

The Hokuyo LRF has an apex angle of 270° and an angular resolution of 0.25° , resulting in 1080 distance measurements per 2D scan, called a *scan line*. The Dynamixel actuator rotates the 2D LRF at one rotation per second, resulting in 40 scan lines and 43,200 distance measurements per full rotation. Slower rotation is possible if a higher angular resolution is desired. For our setup, a half rotation leads to a full 3D scan of most of the environment. Hence, we can acquire 3D scans with up to 21,600 points with 2 Hz (Fig. 1c).

4 Local Multi-Resolution Surfel Map

Distance measurements from the sensor are accumulated in a 3D multi-resolution map with increasing cell sizes from the center of our MAV. The representation consists of multiple MAV-centered 3D grid-maps with different resolutions. On the finest resolution, we use a cell length of 0.25 m. Each grid-map is embedded in the next level with coarser resolution (doubled cell size).

We use a hybrid representation, storing 3D point measurements along with mean and covariance of the points and occupancy information in each cell. Point

measurements of consecutive 3D scans are stored in fixed-sized circular buffers, allowing for point-based data processing and facilitates efficient nearest-neighbor queries.

Fig. 2 shows a 1D schematic illustration of the map organization. We aim for efficient map management for translation and rotation. Therefore, individual grid cells are stored in a circular buffer to allow for shifting elements in constant time. We interlace three levels of circular buffers to obtain a map with three dimensions. The length of the circular buffers depends on the resolution and the size of the map. In case of a translation of the MAV, the circular buffers are shifted whenever necessary to maintain the egocentric property of the map. In case of a translation equal or larger than the cell size, the circular buffers for respective dimensions are shifted. For sub-cell-length translations, the translational parts are accumulated and shifted if they exceed the length of a cell.

Since we store 3D points for every cell for point-based processing, single points are transformed in the cell’s local coordinate frame when adding, and back to the map’s coordinate frame when accessing. Every cell in the map stores a list of 3D points from the current and previous 3D scans. This list is also implemented by a fixed-sized circular buffer. If the capacity of the circular buffer is exceeded, old measurements are discarded and replaced by new measurements.

Since rotating the map would necessitate to shuffle all cells, our map is oriented independent to the MAV’s orientation. We maintain the orientation between the map and the MAV and use it to rotate measurements when accessing the map.

Besides the scan registration described in the following section, the map is utilized by our obstacle avoidance control. Thus, measurements are integrated by ray-casting. For every measurement, the occupancy information for the cells along the ray and the end point are updated using a beam-based inverse sensor model. Occupied cells are avoided using a predictive potential field method [16].

5 Scan Registration

We register consecutive 3D laser scans with our local multi-resolution surfel grid map to estimate the motion of the MAV. We acquire 3D scans in each half rotation of the laser. Since the scans are taken in-flight in a sensor sweep, the motion of the MAV needs to be compensated for when assembling the scan measurements into 3D scans. We register 3D scans with the so far accumulated local map of the environment. The local map is then updated with the registered 3D scan.

5.1 3D Scan Assembly

We estimate the motion of the MAV on a short time scale using visual odometry [21] from two pairs of wide-angle stereo cameras. This 6D motion estimate is used to assemble the individual 2D scan lines of each half rotation to a 3D scan (see Fig. 3).

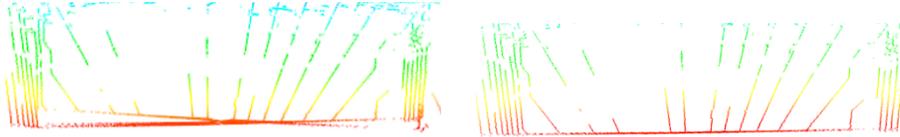


Fig. 3: Side view on an indoor 3D scan with flat ground. Left: assembled 3D scan without considering sensor movement during the scan acquisition. Right: We incorporate visual odometry to correct for the sensor movement.

5.2 Registration Approach

We register the points $\mathcal{P} = \{p_1, \dots, p_P\}$ in a 3D scan with the points $\mathcal{Q} = \{q_1, \dots, q_Q\}$ in the local grid map of the environment [6]. Instead of considering each point individually, we map the 3D scan into a local multi-resolution grid and match surfels, i.e.,

$$p(\mathcal{P} | \theta, \mathcal{Q}) \approx \prod_{i=1}^N p(x_i | \theta, Y)^{P_{x,i}}. \quad (1)$$

By this, several orders of magnitudes less map elements are used for registration. Similarly, the registration of two local maps is treated as the registration of their point sets. We denote the set of surfels in the scene (the 3D scan) by $X = \{x_1, \dots, x_N\}$ and write $Y = \{y_1, \dots, y_M\}$ for the set of model surfels in the environment map. E.g., a surfel x_i summarizes its attributed $P_{x,i}$ points by their sample mean $\mu_{x,i}$ and covariance $\Sigma_{x,i}$. We assume that scene and model can be aligned by a rigid 6 degree-of-freedom (DoF) transformation $T(\theta)$ from scene to model.

5.3 Gaussian Mixture Observation Model

We explain each transformed scene surfel as an observation from a mixture model, similar as in the coherent point drift (CPD) method [15]. A surfel x_i is observed under the mixture defined by the model surfels and an additional uniform component that explains outliers, i.e.,

$$p(x_i | \theta, Y) = \sum_{j=1}^{M+1} p(c_{i,j}) p(x_i | c_{i,j}, \theta, Y). \quad (2)$$

The binary variable c_i indicates the association of x_i to one of the mixture components. The model is a mixture on Gaussian components for the M model surfels that measure the matching likelihood between the surfels through

$$p(x_i | c_{i,j}, \theta, Y) := \mathcal{N} [T(\theta)\mu_{x,i}; \mu_{y,j}, \Sigma_{y,j} + R(\theta)\Sigma_{x,i}R(\theta)^T + \sigma_j^2 I], \quad (3)$$

where $\sigma_j = \frac{1}{2}\rho_{y,j}^{-1}$ is a standard deviation that we adapt to the resolution $\rho_{y,j}$ of the model surfel. We set the likelihood of the uniform mixture component to a

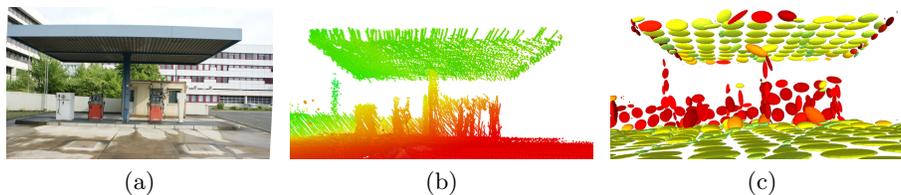


Fig. 4: a) photo of an outdoor scenario. b) the point-based representation of our local environment map. c) the corresponding surfel map.

constant. This way, we do not make a hard association decision for each surfel, but a scene surfel is associated to many model surfels.

5.4 Registration through Expectation-Maximization

The alignment pose θ is estimated through maximization of the logarithm of the joint data-likelihood

$$\ln p(\mathcal{P} \mid \theta, \mathcal{Q}) \approx \sum_{i=1}^N P_{x,i} \ln \sum_{j=1}^{M+1} p(c_{i,j}) p(x_i \mid c_{i,j}, \theta, Y). \quad (4)$$

We optimize this objective function through expectation-maximization (EM) [3]. In the M-step, the latest estimate \bar{q} for the distribution over component associations is held fixed to optimize for the pose θ

$$\hat{\theta} = \operatorname{argmax}_{\theta} \operatorname{const.} + \sum_{i=1}^N P_{x,i} \sum_{j=1}^{M+1} \bar{q}(c_{i,j}) \ln p(x_i \mid c_{i,j}, \theta, Y). \quad (5)$$

This optimization is efficiently performed using the Levenberg-Marquardt (LM) method as in [25]. The LM method is suitable for weighted non-linear least squares problems of the form $\operatorname{argmax}_x e^T(x) W e(x)$, where $e(x) = y - f(x)$ is a vector of residuals and W is a weighting matrix. We stack the residuals $\mu_{y,j} - T(\theta)\mu_{x,i}$ between associated surfels and neglect the effect of the pose on the covariance to obtain a constant block-diagonal weighting matrix. Each block for a surfel association is given by the inverse of the covariance $\Sigma_{y,j} + R(\theta)\Sigma_{x,i}R(\theta)^T$ of the surfel match. Additionally, according to Eq. (5), each association is weighted by a factor $w := P_{x,i}\bar{q}(c_{i,j})$ to the inverse covariance. The steps taken by LM optimization are

$$\Delta x := (J^T W J + \lambda I)^{-1} J^T W e(x), \quad (6)$$

where J is the Jacobian stacked from individual Jacobians per surfel association, and λ is adjusted by LM to trade between Gauss-Newton and gradient descent steps. Note that due to the block-diagonal structure of W , this update decomposes into sums over individual terms per association. The covariance of the LM estimate is readily obtained by $\Sigma(x) := (J^T W J)^{-1}$.

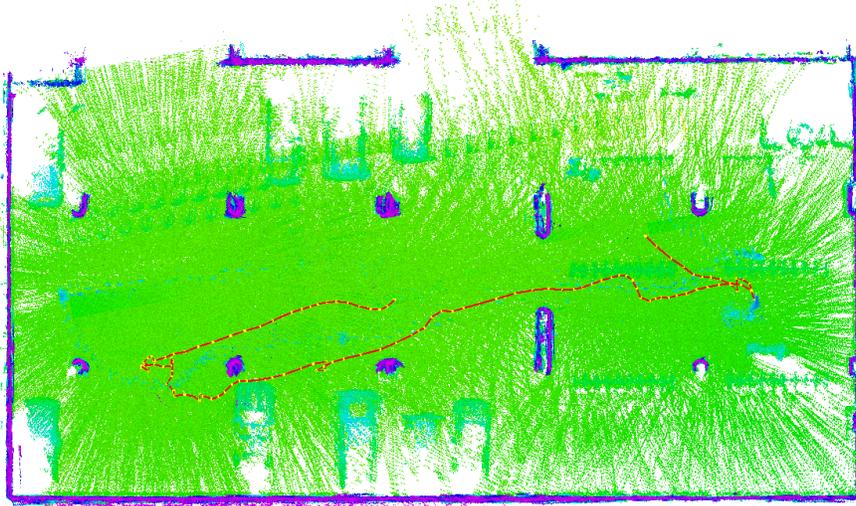


Fig. 5: A top-down view of the resulting map. The point color (green to purple) encodes the distance from the ground. The yellow points, connected by red lines, show the trajectory of the pose graph optimization.

The E-step obtains a new optimum \hat{q} for the distribution q by the conditional likelihood of the cluster associations given the latest pose estimate $\bar{\theta}$

$$\hat{q}(c_{i,j}) = \frac{p(c_{i,j}) p(x_i | c_{i,j}, \bar{\theta}, Y)}{\sum_{j'=1}^{M+1} p(c_{i,j'}) p(x_i | c_{i,j'}, \bar{\theta}, Y)}. \quad (7)$$

In order to evaluate these soft assignments, we perform a local search in the local multi-resolution surfel grid of the model. We first look up the grid cell with a surfel available on the finest resolution in the model map at the transformed mean position of the scene surfel. We consider the surfels in this cell and its direct neighbors for soft association.

6 Simultaneous Localization and Mapping

Our map representation and registration method is able to track the pose of the MAV in a local region, since we decrease the resolution in the map with distance to the MAV. In order to localize the robot in a fixed frame towards its environment and to concurrently build an allocentric map, we align local multi-resolution maps acquired from different view poses.

We register the current local multi-resolution map towards a reference key view to keep track of the MAV's motion. A new key view is generated for the current map, if the robot moved sufficiently far. The new key view is set as the reference for further tracking. The registration result x_i^j between a new key

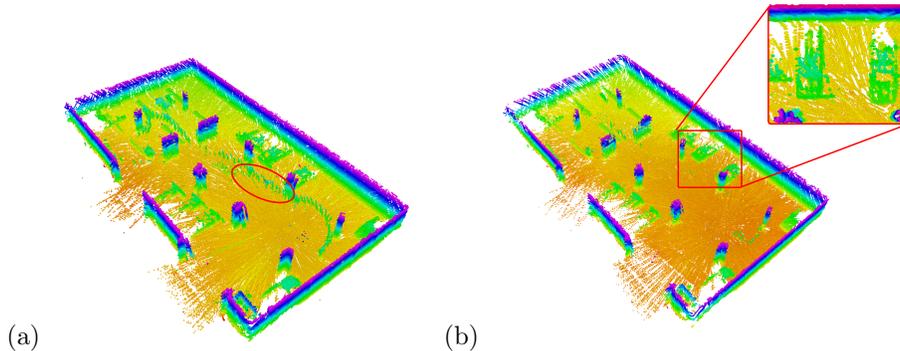


Fig. 6: The resulting map of the parking garage after pose graph optimization from different views. Points corresponding to the ceiling of the parking garage have been cut out for visualization. Color encodes the distance from the ground. a) vertical artifacts (red circle) are caused by a human safety pilot following the MAV; b) after filtering for dynamic obstacles.

view v_i and its reference v_j is a spatial constraint that we maintain as values of edges $e_{ij} \in \mathcal{E}$ in a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of key views.

To overcome pure time-sequential pose tracking by registration, we add spatial constraints between closely key views that are not in temporal sequence. On-line SLAM is enabled by establishing up to one spatial constraint per 3D scan update.

Constraint Detection On each scan update, we check for one new constraint between the current reference v_{ref} and other key views v_{cmp} . We determine a probability

$$p_{\text{chk}}(v_{\text{cmp}}) = \mathcal{N}(d(x_{\text{ref}}, x_{\text{cmp}}); 0, \sigma_d^2) \quad (8)$$

that depends on the linear distance $d(x_{\text{ref}}, x_{\text{cmp}})$ between the key view poses x_{ref} and x_{cmp} . We sample a key view v according to $p_{\text{chk}}(v)$ and determine a spatial constraint between the key views using our registration method.

Pose Graph Optimization From the graph of spatial constraints, we infer the probability of the trajectory estimate given all relative pose observations

$$p(\mathcal{V} | \mathcal{E}) \propto \prod_{e_{ij} \in \mathcal{E}} p(x_i^j | x_i, x_j). \quad (9)$$

Each spatial constraint is a normal distributed estimate with mean and covariance determined by our probabilistic registration method. This pose graph optimization is efficiently solved using the g^2o framework [12].

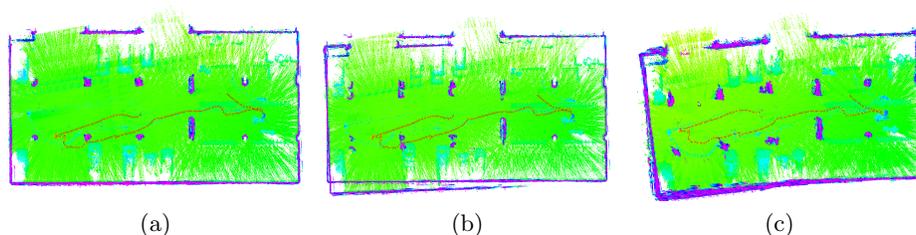


Fig. 7: Top-down views of resulting maps. Using our surfel registration method and global graph optimization (a), our surfel registration method without global graph optimization (b), and GICP registration with global graph optimization (c).

Filtering Dynamic Objects After global pose graph optimization we aggregate the local maps in one global grid map to filter for dynamic obstacles. Measurements are integrated by ray-casting to determine grid cells along the beam from the laser scanner to the end point grid cell. Spurious measurements from dynamic objects are filtered. An example of a dynamic object is shown in Fig. 6. Here, the human safety pilot was following the MAV.

7 Experiments

To evaluate our global registration method, we acquire data with our MAV during flight in a parking garage. The MAV was controlled by a human operator following it. The data sequence contains 200 3D scans and the overall trajectory length is 73 m, covering the complete parking garage. Throughout the experiments, four levels are used for the map with a cell length of 0.25 m at the finest level, which yields a cell length of 2 m at the coarsest level.

Fig. 5 shows the resulting allocentric map and trajectory after pose graph optimization. For this figure, we chose an orthogonal top-down perspective to get an indication about the consistency of the aligned 3D scans by the parallel walls. In contrast, Fig. 6 shows the resulting map from different perspectives, which allows for a better interpretation of the scene. Here, cars and pillars in the parking garage can be identified in the globally aligned 3D scans. Fig. 8 shows that even lamps hanging from the ceiling are modeled by the 3D point cloud.

To assess the improvement of the map accuracy that can be attributed to global pose graph optimization, Fig. 7 compares the resulting maps of different registration methods. Fig. 7b) shows that without pose graph optimization, the trajectory aggregates drift which results in inconsistencies indicated by the misalignment of the walls.

We compare our registration method to a state-of-the-art method by aligning the sparse 3D scans by the Generalized ICP (GICP) [22]. Note that, similar to our pipeline, scans are assembled by visual odometry and pose graph optimization is used to globally align the local dense 3D maps to have a fair comparison. Fig. 7c

shows that the resulting 3D map is less accurate and smeared. In terms of run-times our method is computationally more efficiency with 145 ± 50 ms compared to GICP with 1555 ± 613 ms. Here, mean run-times and standard deviation over the complete dataset are reported.

In a second experiment we acquire data in an outdoor scenario mapping a gas station during flight. The resulting global maps are shown in Fig. 9.



Fig. 8: The structure of the ceiling showing hanging lamps in a photo (left) and the resulting 3D map (right).

8 Conclusions

We propose an approach to simultaneous localization and mapping that is based on the measurements of a lightweight 3D laser-range finder mounted on a MAV.

Laser-range measurements are aggregated by registering sparse 3D scans with a local multiresolution surfel map that has high resolution in the vicinity of the MAV and coarser resolutions with increasing distance, which corresponds well to measurement density and accuracy of our sensor. Modeling measurement distributions within voxels by surface elements allows for efficient and accurate registration of 3D scans with the local map. The incrementally built local dense 3D maps of nearby key poses are registered globally by graph optimization. This yields a globally consistent dense 3D map of the environment. Continuous registration of local maps with the global map allows for tracking the 6D MAV pose in real time. We demonstrate accuracy and efficiency of our approach by showing consistent allocentric 3D maps, recorded by our MAV during flight in a parking garage and comparing them to results from a state-of-the-art registration method.

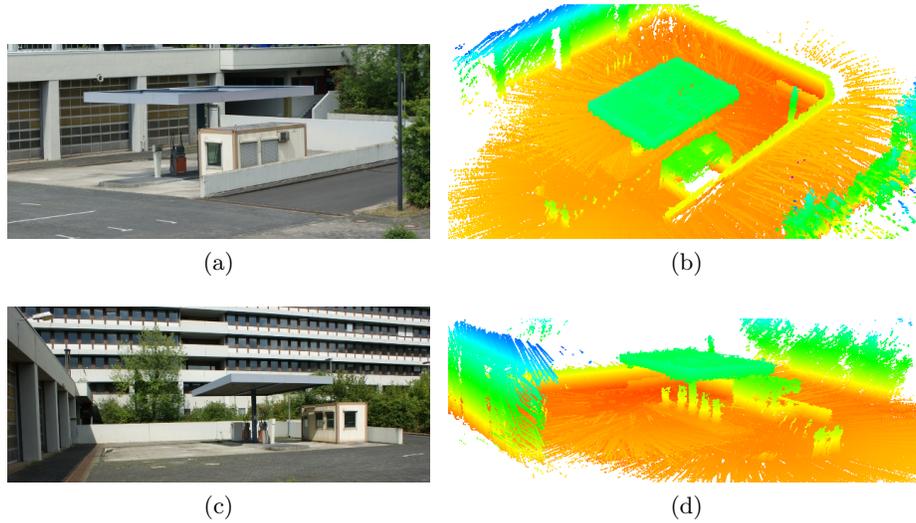


Fig. 9: (a+c) photo of the scene from different perspectives. (b+d) the resulting global maps acquired during flight with our MAV.

References

1. Anderson, S., Barfoot, T.D.: Towards relative continuous-time SLAM. In: Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA) (2013)
2. Bachrach, A., He, R., Roy, N.: Autonomous flight in unstructured and unknown indoor environments. In: European Micro Aerial Vehicle Conf (EMAV) (2009)
3. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
4. Bosse, M., Zlot, R.: Continuous 3D scan-matching with a spinning 2D laser. In: Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA) (2009)
5. Cover, H., Choudhury, S., Scherer, S., Singh, S.: Sparse tangential network (SPARTAN): Motion planning for micro aerial vehicles. In: Robotics and Automation (ICRA), IEEE International Conference on (2013)
6. Droschel, D., Stückler, J., Behnke, S.: Local multi-resolution representation for 6D motion estimation and mapping with a continuously rotating 3D laser scanner. In: Robotics and Automation (ICRA), IEEE International Conference on (2014)
7. Elseberg, J., Borrmann, D., Nuechter, A.: 6DOF semi-rigid SLAM for mobile scanning. In: Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on (2012)
8. Grzonka, S., Grisetti, G., Burgard, W.: Towards a navigation system for autonomous indoor flying. In: Robotics and Automation (ICRA), IEEE International Conference on (2009)
9. Grzonka, S., Grisetti, G., Burgard, W.: A fully autonomous indoor quadrotor. IEEE Trans. on Robotics 28(1), 90–100 (2012)
10. Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W.: OctoMap: an efficient probabilistic 3D mapping framework based on octrees. Autonomous Robots 34, 189–206 (2013)

11. Huh, S., Shim, D., Kim, J.: Integrated navigation system using camera and gimbaled laser scanner for indoor and outdoor autonomous flight of UAVs. In: *Intelligent Robots and Systems (IROS)*, IEEE/RSJ International Conference on (2013)
12. Kuemmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: G2o: A general framework for graph optimization. In: *Robotics and Automation (ICRA)*, IEEE International Conference on (2011)
13. Maddern, W., Harrison, A., Newman, P.: Lost in translation (and rotation): Fast extrinsic calibration for 2D and 3D LIDARs. In: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)* (May 2012)
14. Magnusson, M., Duckett, T., Lilienthal, A.J.: Scan registration for autonomous mining vehicles using 3D-NDT. *Journal of Field Robotics* 24(10), 803–827 (2007)
15. Myronenko, A., Song, X.: Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(12), 2262–2275 (2010)
16. Nieuwenhuisen, M., Droschel, D., Schneider, J., Holz, D., Läbe, T., Behnke, S.: Multimodal obstacle detection and collision avoidance for micro aerial vehicles. In: *Proceedings of 6th European Conference on Mobile Robots (ECMR)* (2013)
17. Nuechter, A., Lingemann, K., Hertzberg, J., Surmann, H.: 6D SLAM with approximate data association. In: *Int. Conf. on Advanced Robotics* (2005)
18. Ryde, J., Hu, H.: 3D mapping with multi-resolution occupied voxel lists. *Autonomous Robots* 28, 169 – 185 (2010)
19. Schadler, M., Stückler, J., Behnke, S.: Multi-resolution surfel mapping and real-time pose tracking using a continuously rotating 2D laser scanner. In: *Proc. of 11th IEEE Int. Symposium on Safety, Security, and Rescue Robotics (SSRR)* (2013)
20. Scherer, S., Rehder, J., Achar, S., Cover, H., Chambers, A.D., Nuske, S.T., Singh, S.: River mapping from a flying robot: State estimation, river detection, and obstacle mapping. *Autonomous Robots* 32(5), 1–26 (May 2012)
21. Schneider, J., Läbe, T., Förstner, W.: Incremental real-time bundle adjustment for multi-camera systems with points at infinity. In: *ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. vol. XL-1/W2 (2013)
22. Segal, A., Hähnel, D., Thrun, S.: Generalized-ICP. In: *Proc. of Robotics: Science and Systems (RSS)* (2009)
23. Shen, S., Michael, N., Kumar, V.: Autonomous multi-floor indoor navigation with a computationally constrained micro aerial vehicle. In: *Robotics and Automation (ICRA)*, IEEE International Conference on (2011)
24. Stoyanov, T., Lilienthal, A.: Maximum likelihood point cloud acquisition from a mobile platform. In: *Proc. of the Int. Conf. on Advanced Robotics (ICAR)* (2009)
25. Stückler, J., Behnke, S.: Multi-resolution surfel maps for efficient dense 3D modeling and tracking. *Journal of Visual Communication and Image Representation* 25(1), 137–147 (2014)
26. Takahashi, M., Schulein, G., Whalley, M.: Flight control law design and development for an autonomous rotorcraft. In: *Proceedings of the 64th Annual Forum of the American Helicopter Society* (2008)
27. Thrun, S., Diel, M., Hähnel, D.: Scan alignment and 3-d surface modeling with a helicopter platform. In: *FSR*. Springer Tracts in Advanced Robotics, vol. 24, pp. 287–297. Springer (2003)
28. Tomić, T., Schmid, K., Lutz, P., Domel, A., Kassecker, M., Mair, E., Grix, I., Ruess, F., Suppa, M., Burschka, D.: Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue. *Robotics Automation Magazine, IEEE* 19(3), 46–56 (2012)