Local Multi-Resolution Representation for 6D Motion Estimation and Mapping with a Continuously Rotating 3D Laser Scanner

David Droeschel, Jörg Stückler, and Sven Behnke

Abstract— Micro aerial vehicles (MAV) pose a challenge in designing sensory systems and algorithms due to their size and weight constraints and limited computing power. We present an efficient 3D multi-resolution map that we use to aggregate measurements from a lightweight continuously rotating laser scanner. We estimate the robot's motion by means of visual odometry and scan registration, aligning consecutive 3D scans with an incrementally built map.

By using local multi-resolution, we gain computational efficiency by having a high resolution in the near vicinity of the robot and a lower resolution with increasing distance from the robot, which correlates with the sensor's characteristics in relative distance accuracy and measurement density. Compared to uniform grids, local multi-resolution leads to the use of fewer grid cells without loosing information and consequently results in lower computational costs. We efficiently and accurately register new 3D scans with the map in order to estimate the motion of the MAV and update the map in-flight.

In experiments, we demonstrate superior accuracy and efficiency of our registration approach compared to state-of-the-art methods such as GICP. Our approach builds an accurate 3D obstacle map and estimates the vehicle's trajectory in real-time.

I. INTRODUCTION

Micro aerial vehicles (MAV) such as quadrotors have attracted attention in the field of aerial robotics. Their size and weight limitations pose a challenge in designing sensory systems. Most of today's MAVs are equipped with ultra sound sensors and camera systems due to their minimal size and weight. While these small and lightweight sensors provide valuable information, they suffer from a limited field-of-view and are sensitive to illumination conditions. Only few systems [1], [2], [3], [4] are equipped with 2D laser range finders (LRF) that are used for navigation.

In contrast, we build a continuously rotating laser scanner that is minimalistic in terms of size and weight and thus particularly well suited for obstacle perception and localization on MAVs, allowing for environment perception in all directions.

We use a hybrid 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 distance accuracy and measurement density. Compared to
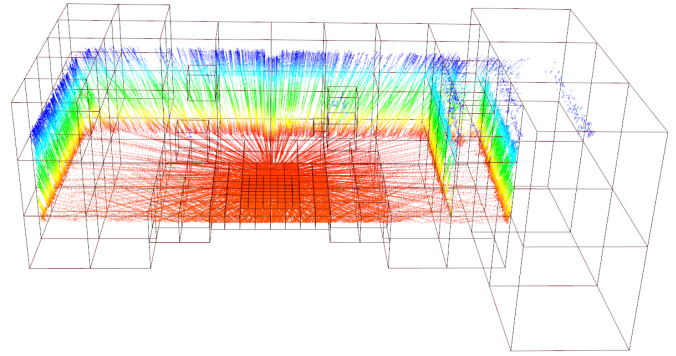
Fig. 1: Local multi-resolution grid-map with a higher resolution in the close proximity to the sensor and a lower resolution with increasing distance. Color encodes height.

uniform grids, local multi-resolution leads to the use of fewer grid cells without loosing information and consequently results in lower computational costs. Fig. 1 shows our local multi-resolution grid-map.

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 (denoted as surfels) between local multi-resolution grid-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). Since laser-based ego-motion estimation relies on structure in the scene, it works best in scenarios where GPS typically is not available, like in indoor or urban environments.

II. RELATED WORK

For mobile ground robots, 3D laser scanning sensors are widely used due to their accurate distance measurements even in bad lighting conditions, and due to their large field-of-view (FoV). For instance, autonomous cars often perceive obstacles by means of a rotating laser scanner with a $360°$ horizontal FoV, allowing for detection of obstacles in every direction [5], [6].

Up to now, such 3D laser scanners are rarely used on lightweight MAVs due to their payload limitations. Instead,
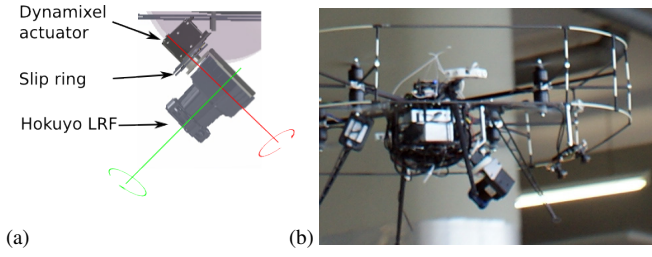
(a)   (b)

Fig. 2: (a) CAD drawing of the continuously rotating laser scanner with the two rotation axes. The Hokuyo 2D LRF is mounted on a bearing and rotated around the red axis. (b) The 3D laser scanner mounted on our multicopter.
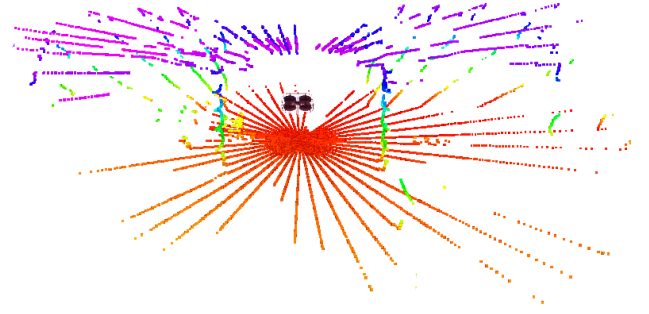

Fig. 3: Indoor 3D scan acquired with our continuously rotating laser scanner. Color encodes height.
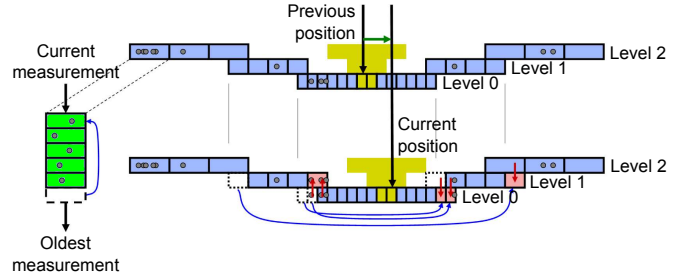

Fig. 4: One-dimensional illustration of the hybrid local multi-resolution map. Along with the occupancy information, every grid-cell (blue) maintains a circular buffer with its associated measurement points (green). The map is centered around the robot and in case of a robot motion, ring buffers are 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).

2D laser range finders [7], [1], [2], [3], [4], [8] are used, which restricts the FoV to the 2D measurement plane of the sensor.

A similar setup to ours is described by Scherer and Cover et al. [9], [10]. Their MAV is used to autonomously explore rivers using visual localization and laser-based 3D obstacle perception. In contrast to their work, we aggregate consecutive laser scans in our multi-resolution map by 3D scan registration.

For mobile ground robots, several approaches have been proposed to estimate the motion of a robot by means of 3D scan registration [11], [12], [13]. Most of these approaches are derived from the Iterative Closest Points (ICP) algorithm [14]. Generalized ICP (GICP) [13] unifies the ICP formulation for various error metrics such as point-to-point, point-to-plane, and plane-to-plane. The 3D-NDT [15] discretizes point clouds in 3D grids and aligns Gaussian statistics within grid cells to perform scan registration. Recently, multi-resolution surfel maps have been proposed that match Gaussian statistics in multi-resolution voxel representations to efficiently and accurately register RGB-D images [16] and 3D laser scans [17]. In this work, we extend surfel registration with probabilistic data association to better cope with sparse point clouds from fast spinning laser scanners.

We aim at perceiving as much of the surroundings as possible in order to obtain almost omnidirectional obstacle detection. Distance measurements are aggregated in a 3D grid-map at multiple resolutions and acquired sparse 3D scans are registered with the map.

III. 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 actuator to gain a 3D FoV. As shown in Fig. 2, 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. 2) 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 2D laser scanner has a size of $62\times62\times87.5$ mm and a weight of 210 g. Together with the actuator (72 g) and the slip ring, 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.

IV. LOCAL MULTI-RESOLUTION MAP

Distance measurements from the sensor are accumulated in a 3D multi-resolution map with increasing cell sizes from the robot's center. The representation consists of multiple robot-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 and doubled cell length.

We use a hybrid representation, storing 3D point measurements along with 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. 4 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 multiple 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 using a predictive potential field method to avoid occupied cells [18].

V. 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.

A. 3D Scan Assembly

We estimate the motion of the MAV on a short time scale using visual odometry [19] 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. 5).

B. Scan To Map Registration

We register a 3D scan $\mathcal{P} = \{p_1, \ldots, p_P\}$ with the points $\mathcal{Q} = \{q_1, \ldots, q_Q\}$ in the local grid map of the environment.
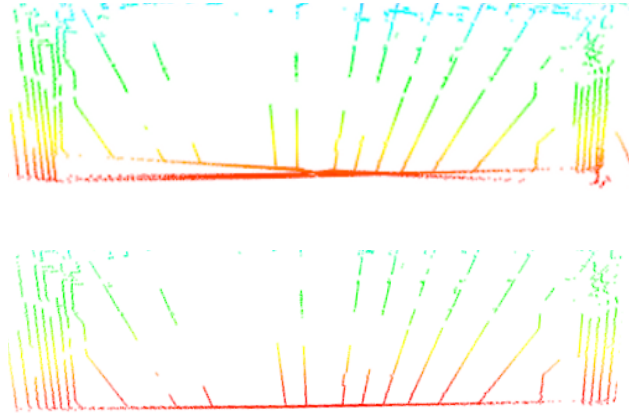

Fig. 5: Side view on an indoor 3D scan with flat ground. Top: assembled 3D scan without considering sensor movement during the scan acquisition. Bottom: We incorporate visual odometry to correct for the sensor movement.

We formulate the registration of the 3D scan with the local environment map as optimizing the joint data-likelihood

$$p(\mathcal{P} \mid \theta, \mathcal{Q}) = \prod_{k=1}^{P} p(p_k \mid \theta, \mathcal{Q}). \quad (1)$$

Instead of considering each point individually, we map the 3D scan into a local multi-resolution grid and match surfels, i.e.,

$$p(\mathcal{P} \mid \theta, \mathcal{Q}) \approx \prod_{i=1}^{N} p(x_i \mid \theta, Y)^{P_{x,i}}. \quad (2)$$

By this, several orders of magnitudes less map elements are used for registration. We denote the set of surfels in the scene (the 3D scan) by $X = \{x_1, \ldots, x_N\}$ and write $Y = \{y_1, \ldots, y_M\}$ 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. Our aim is to recover the relative pose θ of the scene towards the model. An exemplary surfel map together with its originating points is shown in Fig. 6.

C. 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 [20]. A surfel x_i is observed under the mixture defined by the model surfels and an additional uniform component that explains outliers, i.e.,

$$p(x_i \mid \theta, Y) = \sum_{j=1}^{M+1} p(c_{i,j}) \, p(x_i \mid c_{i,j}, \theta, Y), \quad (3)$$

where $c_{i,j}$ is a shorthand for the 1-of-(M+1) encoding binary variable $c_i \in \mathbb{B}^{M+1}$ with j-th entry set to 1. Naturally, c_i indicates the association of x_i to exactly one of the mixture components. The model is a mixture on Gaussian

components for the M model surfels through

$$p(x_i \mid c_{i,j}, \theta, Y) :=$$
$$\mathcal{N}\left[T(\theta)\mu_{x,i}; \mu_{y,j}, \Sigma_{y,j} + R(\theta)\Sigma_{x,i}R(\theta)^T + \sigma_j^2 I\right], \quad (4)$$

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 $p(c_{i,M+1}) = w$. For this uniform component, the data likelihood of a surfel x_i is

$$p(x_i \mid c_{i,M+1}, \theta) = \frac{P_{x,i}}{P} \mathcal{N}(0; 0, R(\theta)\Sigma_{x,i}R(\theta)^T + \sigma_j^2 I). \quad (5)$$

For the prior association likelihood, we assume the likelihood of x_i to be associated to one of the points in the model map to be equal. Hence, for each Gaussian mixture component $j \in \{1, \ldots, M\}$ we have $p(c_{i,j}) = (1-w)\frac{Q_{y,j}}{Q}$. By modeling the scene surfels as samples from a mixture on the model surfels, we do not make a hard association decision between the surfels sets, but a scene surfel is associated to many model surfels.

D. Registration through Expectation-Maximization

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

$$\ln p(\mathcal{P} \mid \theta, \mathcal{Q}) \approx \sum_{i=1}^{N} P_{x,i} \ln \sum_{j=1}^{M+1} p(c_{i,j}) \, p(x_i \mid c_{i,j}, \theta, Y). \quad (6)$$

We optimize this objective function through expectation-maximization (EM) [21]. The component associations $c = \{c_1, \ldots, c_N\}$ are treated as latent variables to yield the EM objective

$$L(q, \theta) := \sum_{i=1}^{N} P_{x,i} \sum_{j=1}^{M+1} q(c_{i,j}) \ln \frac{p(c_{i,j}) \, p(x_i \mid c_{i,j}, \theta, Y)}{q(c_{i,j})}, \quad (7)$$

for which we exploit $q(c) = \prod_{i=1}^{N} \prod_{j=1}^{M+1} q(c_{i,j})$. In the M-step, the latest estimate $\bar{q}$ for the distribution over component associations is held fixed to optimize for the pose θ

$$\widehat{\theta} = \underset{\theta}{\arg\max} \, L(\bar{q}, \theta) \quad (8)$$

with

$$L(\bar{q}, \theta) := const. + \sum_{i=1}^{N} P_{x,i} \sum_{j=1}^{M+1} \bar{q}(c_{i,j}) \ln p(x_i \mid c_{i,j}, \theta, Y). \quad (9)$$

This optimization is efficiently performed using the Levenberg-Marquardt method as in [16]. The E-step obtains a new optimum $\widehat{q}$ for the distribution q by the conditional likelihood of the cluster associations given the latest pose estimate $\bar{\theta}$

$$\widehat{q}(c_{i,j}) = \frac{p(c_{i,j}) \, p(x_i \mid c_{i,j}, \bar{\theta}, Y)}{\sum_{j'=1}^{M+1} p(c_{i,j'}) \, p(x_i \mid c_{i,j'}, \bar{\theta}, Y)}. \quad (10)$$

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
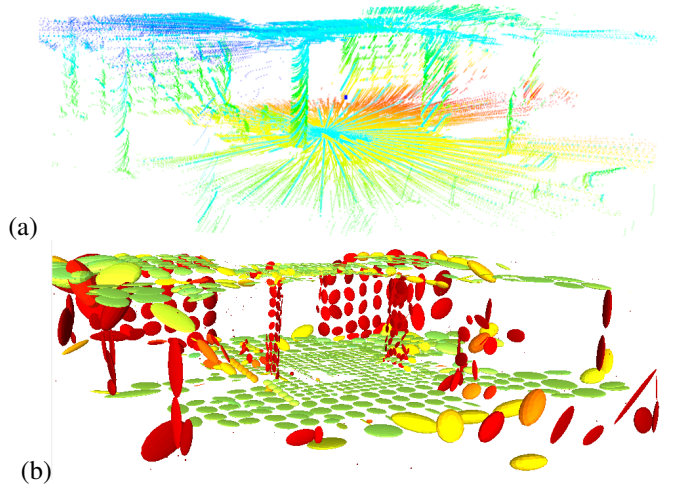
(a)

(b)

Fig. 6: The point-based representation (a) of our local environment map and corresponding surfels (b).

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.

VI. EXPERIMENTS

We assess our registration method using two datasets which have been acquired with our MAV in-flight and compare it to state-of-the-art registration methods. We register the point sets of 3D scan and local multi-resolution map using ICP and Generalized ICP (GICP) [13].

The first dataset provides ground-truth pose information from an indoor motion capture (MoCap) system. The MoCap system provides accurate pose information of the MAV at high frame rates (100 Hz), but is restricted to a small capture volume of approximately $2 \times 2 \times 3$ m. During the 46 s flight, visual odometry and laser data for 92 3D scans have been recorded. A second dataset has been acquired in a parking garage which allows for larger flight distances, but ground-truth data is not available.

For assessing pose accuracy without pose ground-truth, we calculate the *mean map entropy*, a quantitative measure which evaluates with the sharpness of a map. The entropy h for a map point q_k is calculated by

$$h(q_k) = \frac{1}{2} \ln |2\pi e \, \Sigma(q_k)|, \quad (11)$$

where $\Sigma(q_k)$ is the sample covariance of mapped points in a local radius r around q_k. We select $r = 0.3$ m in our evaluation. The mean map entropy $H(\mathcal{Q})$ is averaged over all map points

$$H(\mathcal{Q}) = \frac{1}{Q} \sum_{k=1}^{Q} h(q_k). \quad (12)$$

A. Motion Capture Dataset

Since ground-truth data from a MoCap system is available for the first dataset, we quantify mapping accuracy by the absolute trajectory error (ATE) [22] based on the estimated
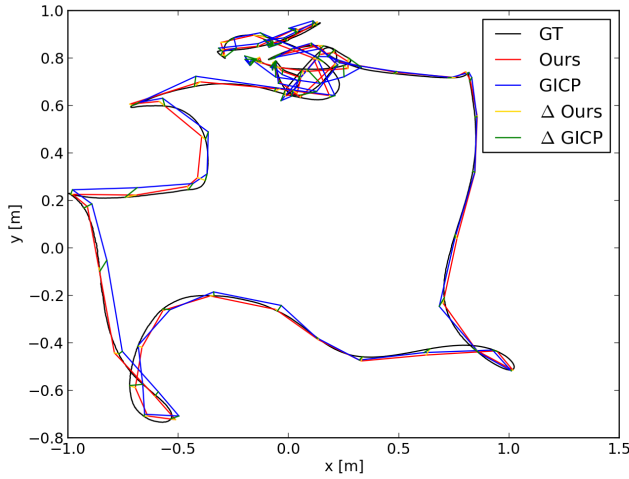
Fig. 7: Absolute trajectory error of the scan registration using the multi-resolution map. Points of the trajectory are projected on the xy-plane.

and the ground-truth trajectory. Table I summarizes the ATE of our method, comparing it to the estimates by visual odometry, ICP, and GICP registration. The results indicate that all scan registration methods improve the motion estimate produced by visual odometry. Our method results in a lower ATE compared to ICP and GICP. In addition, the run-times reported in Table I demonstrate that our method is computationally more efficient. The results in mean map entropy confirm improved accuracy by our registration method. In Fig. 7 we shows the trajectory estimate obtained with our registration method and GICP, and displays deviations of both estimates from the ground-truth trajectory.

Throughout the experiments, four resolution 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.

B. Garage Dataset

Since ground-truth data, e.g., from a MoCap system is not available for the second dataset, we evaluate accuracy by visually inspecting the sharpness of the map, as shown in Fig. 9 and comparing mean map entropy as for the previous dataset. Using GICP to estimate the motion and building the map results in an map entropy of −3.438, whereas using our method results in a lower entropy of −3.696. Fig. 8 illustrates the increase of measurement density through the aggregation of measurements.

VII. CONCLUSIONS

We presented an efficient 3D multi-resolution map that we use for obstacle avoidance and for estimating the motion of the robot. We aggregate measurements from a continuously rotating laser scanner that is particularly well suited for MAVs due to its size and weight.

By using local multi-resolution, we gain computational efficiency by having a high resolution in the near vicinity of the robot and a lower resolution with increasing distance from
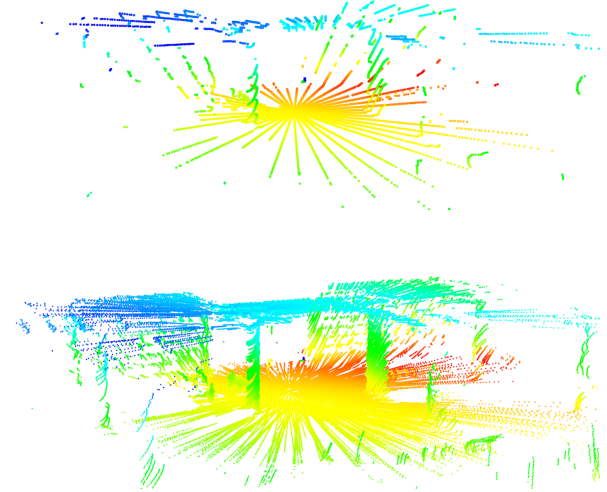

Fig. 8: Scene of the garage dataset (top). Aggregated 3D map after 1 scan (middle) and 10 scans (bottom) using our registration method (color encodes height).

the robot, which correlates with the sensor's characteristics in relative distance accuracy and measurement density.

Scan registration is used to estimate the motion of the robot by aligning consecutive 3D scans to the map. Hence, we are able to efficiently align new 3D scans with the map and aggregate distance measurements from consecutive 3D scans to increase the density of the map. We do not match individual scan points, but represent 3D scans also in local multi-resolution grids and condense the points into surface elements for each grid cell. These surface elements are aligned efficiently and at high accuracy in a registration framework which overcomes the discretization in a grid through probabilistic assignments.

In experiments, we compare the laser-based motion estimate with ground-truth from a motion capture system and the GICP, a state-of-the-art registration algorithm, as well as standard point-based ICP. Overall, our approach is more accurate and results in sharper maps as indicated by the lower ATE and map entropy. Besides that, our approach is computational more efficient, allowing to register scans and to build local 3D maps in-flight in real-time.

TABLE I: ATE, map entropy, and run-time of our surfel registration method, in comparison to visual odometry (VO), ICP, and GICP.

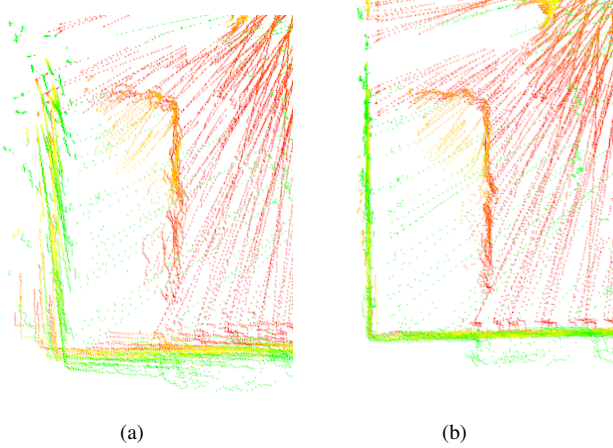|	ATE (m)						mean map entropy	run-time (ms)		
sequence	RMSE	mean	median	std	min	max		mean	std	max
VO	0.151	0.134	0.129	0.059	0.024	0.324	-3.112			
ICP	0.040	0.035	0.034	0.019	0.006	0.117	-3.411	290.31	108.72	521
GICP	0.034	0.031	0.030	0.014	**0.005**	0.088	-3.363	1769.52	813.92	5805
ours	**0.021**	**0.019**	**0.016**	**0.010**	**0.005**	**0.061**	**-3.572**	**51.06**	**27.30**	**121**

(a)　　　　　　　　　(b)

Fig. 9: Cut-out part of the map, generated by GICP (a) and our method (b). As indicated by the lower map entropy, the map generated by our method is sharper.

REFERENCES

[1] T. Tomić, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I. Grixa, F. Ruess, M. Suppa, and D. Burschka, "Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue," *Robotics Automation Magazine, IEEE*, vol. 19, no. 3, pp. 46–56, 2012.

[2] S. Grzonka, G. Grisetti, and W. Burgard, "Towards a navigation system for autonomous indoor flying," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2009.

[3] A. Bachrach, R. He, and N. Roy, "Autonomous flight in unstructured and unknown indoor environments," in *European Micro Aerial Vehicle Conf (EMAV)*, 2009, pp. 1–8.

[4] S. Shen, N. Michael, and V. Kumar, "Autonomous multi-floor indoor navigation with a computationally constrained micro aerial vehicle," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2011, pp. 2968 –2969.

[5] C. Urmson, J. Anhalt, H. Bae, J. A. D. Bagnell, C. R. Baker , R. E. Bittner, T. Brown, M. N. Clark, M. Darms, D. Demitrish, J. M. Dolan, D. Duggins, D. Ferguson, T. Galatali, C. M. Geyer, M. Gittleman, S. Harbaugh, M. Hebert, T. Howard, S. Kolski, M. Likhachev, B. Litkouhi, A. Kelly, M. McNaughton, N. Miller, J. Nickolaou, K. Peterson, B. Pilnick, R. Rajkumar, P. Rybski, V. Sadekar, B. Salesky, Y.-W. Seo, S. Singh, J. M. Snider, J. C. Struble, A. T. Stentz, M. Taylor , W. R. L. Whittaker, Z. Wolkowicki, W. Zhang, and J. Ziglar, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics Special Issue on the 2007 DARPA Urban Challenge, Part I*, vol. 25, no. 8, pp. 425–466, June 2008.

[6] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt, and S. Thrun, "Junior: The Stanford entry in the Urban Challenge," *Journal of Field Robotics*, vol. 25, no. 9, pp. 569–597, 2008.

[7] S. Grzonka, G. Grisetti, and W. Burgard, "A fully autonomous indoor quadrotor," *IEEE Trans. on Robotics*, vol. 28, no. 1, pp. 90–100, 2012.

[8] S. Huh, D. Shim, and J. Kim, "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, pp. 3158–3163.

[9] S. Scherer, J. Rehder, S. Achar, H. Cover, A. D. Chambers, S. T. Nuske, and S. Singh, "River mapping from a flying robot: State estimation, river detection, and obstacle mapping," *Autonomous Robots*, vol. 32, no. 5, pp. 1–26, May 2012.

[10] H. Cover, S. Choudhury, S. Scherer, and S. Singh, "Sparse tangential network (SPARTAN): Motion planning for micro aerial vehicles," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2013.

[11] A. Nuechter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM with approximate data association," in *Robotics and Automation (ICRA), IEEE International Conference on*, 2005, pp. 242–249.

[12] M. Magnusson, T. Duckett, and A. J. Lilienthal, "Scan registration for autonomous mining vehicles using 3D-NDT," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.

[13] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. of Robotics: Science and Systems (RSS)*, 2009.

[14] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 14, no. 2, pp. 239–256, 1992.

[15] T. Stoyanov, M. Magnusson, H. Andreasson, and A. J. Lilienthal, "Fast and accurate scan registration through minimization of the distance between compact 3D NDT representations," *The International Journal of Robotics Research*, vol. 31, no. 12, pp. 1377–1393, 2012.

[16] J. Stückler and S. Behnke, "Multi-resolution surfel maps for efficient dense 3D modeling and tracking," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 137–147, 2014.

[17] M. Schadler, J. Stückler, and S. Behnke, "Multi-resolution surfel mapping and real-time pose tracking using a continuously rotating 2D laser scanner," in *Proc. of IEEE Int. Symp. on Safety, Security, and Rescue Robotics (SSRR)*, 2013.

[18] M. Nieuwenhuisen, D. Droeschel, J. Schneider, D. Holz, T. Läbe, and S. Behnke, "Multimodal obstacle detection and collision avoidance for micro aerial vehicles," in *Proceedings of 6th European Conference on Mobile Robots (ECMR)*, 2013.

[19] J. Schneider, T. Läbe, and W. Förstner, "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.

[20] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[22] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, Oct. 2012.