European Conference on Mobile Robots (ECMR), Lincoln, UK, September 2015.

# Evaluation of Registration Methods for Sparse 3D Laser Scans

Jan Razlaw, David Droeschel, Dirk Holz, and Sven Behnke

***Abstract*—The registration of 3D laser scans is an important task in mapping applications. For the task of mapping with autonomous micro aerial vehicles (MAVs), we have developed a light-weight 3D laser scanner. Since the laser scanner is rotated quickly for fast omnidirectional obstacle perception, the acquired point clouds are particularly sparse and registration becomes challenging. In this paper, we present a thorough experimental evaluation of registration algorithms in order to determine the applicability of both the scanner and the registration algorithms. Using the estimated poses of the MAV, we aim at building local egocentric maps for both collision avoidance and 3D mapping. We use multiple metrics for assessing the quality of the different pose estimates and the quality of the resulting maps. In addition, we determine for all algorithms optimal sets of parameters for the challenging data. We make the recorded datasets publicly available and present results showing both the best suitable registration algorithm and the best parameter sets as well as the quality of the estimated poses and maps.**

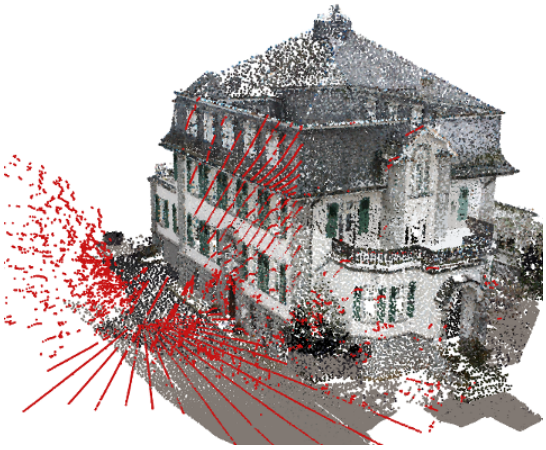


Fig. 1. Typical 3D scan (red) acquired with our continuously rotating laser scanner during a mapping mission aligned to a ground truth colored 3D model of a building.

## I. Introduction

Micro aerial vehicles (MAVs) such as quadrotors are used in an increasing number of research projects and application domains. Their size and weight limitations, however, pose a problem in designing sensory systems for environment perception. Most of today's MAVs are equipped with ultrasonic 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 cameras are sensitive to illumination conditions. Only few MAVs [1], [2], [3], [4] are equipped with 2D laser range finders (LRF) that are used for navigation. These provide accurate distance measurements to objects in the surroundings but are limited to the two-dimensional scanning plane of the sensor. Objects below or above that plane are not perceived.

3D laser scanners provide robots with distance measurements in all directions, allowing them to detect obstacles omnidirectionally, build 3D maps, and localize in 6D. For the task of mapping inaccessible areas with autonomous micro aerial vehicles, we have developed a lightweight 3D scanner [5] specifically suited for the application on MAVs. It consists of a Hokuyo 2D laser range scanner, a rotary actuator and a slip ring to allow continuous rotation. Just as with other rotated scanners, the acquired point clouds (aggregated over one full or half rotation) show the particular characteristic of having non-uniform point densities: usually a high density within each scan line and a larger angle between scan lines (see Fig. 1). Since we use the laser scanner for omnidirectional obstacle detection and collision avoidance, we rotate the scanner quickly with 1 Hz, resulting in a particularly low angular resolution of roughly 9°. These non-uniform point densities affect neighborhood searches and cause problems in local feature estimation and registration when keeping track of the MAV movement and building allocentric 3D maps. To compensate for the non-uniform point densities, there are two solutions: 1) addressing the non-uniform densities by non-uniform neighborhood relations to allow the robust registration of such 3D point clouds as in [6], or 2) aggregating acquired scans over short periods of time in egocentric maps so as to increase the point density as in [7]. In this paper, we focus on the latter and build local egocentric maps that are used for both collision avoidance while flying and registration (with higher point densities) for 3D mapping. The main question being investigated is which registration algorithm performs best for building the egocentric maps and how it needs to be parameterized to show this performance. Still, we include a variant of [6] in the presented comparative experimental evaluation. We make the following contributions:

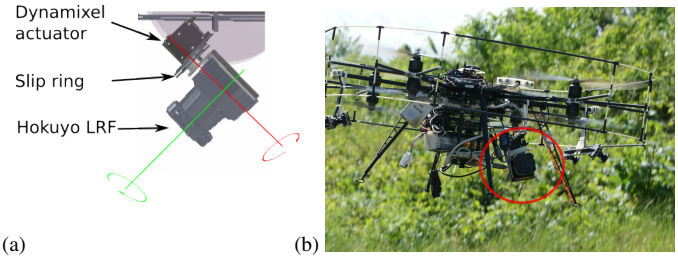
1) We design a set of metrics for assessing the quality of pose estimates and the quality of the resulting maps and aligned point clouds, respectively.
2) Using Hyperopt [8], we determine optimal sets of parameters for all registration methods.
3) We recorded several datasets both in a smaller motion capture volume to obtain ground truth pose estimates and during real missions for mapping buildings with the MAV. We make these datasets publicly available[1].
4) We present the results of a thorough experimental evaluation using the datasets and report gained insights and lessons learned.

This work has been supported partially by the German Federal Ministry for Economic Affairs and Energy (BMWi).

All authors are with the Autonomous Intelligent Systems Group, Computer Science Institute VI, University of Bonn, 53113 Bonn, Germany. Email: jan.razlaw@gmx.de, {droeschel, holz, behnke}@ais.uni-bonn.de

9781-1-4673-9163-4/15/$31.00 

[1] http://www.ais.uni-bonn.de/mav_mapping




(a) (b)

Fig. 2. Lightweight 3D laser scanner. (a) CAD drawing. The Hokuyo 2D LRF rotates its mirror around the green axis. It is continuously rotated around the red axis. (b) The 3D laser scanner mounted on our MAV (marked with a red circle).

## 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 [9], [10]. Up to now, such 3D laser scanners are rarely used on lightweight MAVs due to their payload limitations.

A similar setup to ours is described by Scherer and Cover et al. [11], [12]. 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 a local egocentric map by 3D scan registration and use the resulting maps for both collision avoidance and environment mapping.

For mobile ground robots, several approaches have been proposed to estimate the motion of a robot by means of 3D scan registration [13], [14], [15]. Most of these approaches are derived from the Iterative Closest Points (ICP) algorithm [16]. Generalized ICP (GICP) [15] unifies the ICP formulation for various error metrics such as point-to-point, point-to-plane, and plane-to-plane. 3D Normal Distributions Transform (NDT) [14], [17] discretizes point clouds in 3D grids and aligns Gaussian statistics within grid cells to perform scan registration. Recently, multiresolution surfel maps (MRSMaps) have been proposed that match Gaussian statistics in multiresolution voxel representations to efficiently and accurately register RGB-D images [18] and 3D laser scans [19]. In [7], we extend the latter to use probabilistic data associations in order to better cope with non-uniform density point clouds from fast spinning laser scanners. We include ICP [16], GICP [15], 3D-NDT [17], and the last-named surfel-based registration [7] in our experimental evaluation. For an extensive survey on these and other registration algorithms for 3D point clouds we refer to the recent work of Pomerleau et al. [20].

Several works address the benchmarking of registration methods. A comparative evaluation of ICP and NDT in terms of registration accuracy was presented by Magnusson et al. [21]. Wulf et al. [22] compare ICP-based pairwise and incremental registration [23] and a 3D-variant [24] of Lu-Milios-style graph SLAM [25]. A recent effort for benchmarking SLAM algorithms for RGB-D cameras including datasets and performance metrics for pose accuracy is the RGB-D SLAM Dataset and Benchmark[2] by Sturm et al. [26]. In order to compare constructed maps based on their quality, Schwertfeger et al. proposed a fiducial map metric [27]. In [7], we propose an entropy-based metric to measure the quality of maps and the alignment of point clouds. In this paper, we use metrics from Sturm et al. [26] to measure pose accuracy, and the entropy-based metric from [7] to measure map quality. In addition, we introduce another metric based on mean plane variance to evaluate the quality of alignment and map. As a final measurement of map quality, we propose an ICP-like fitness score derived from first aligning a built map with a ground truth map acquired by a statically mounted 3D laser scanner, and then searching for corresponding points and computing the root mean square error of the distances between the found point matches. The more the built map deviates from the ground truth map, the higher the error in this fitness score is.

## 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(b)) 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 62×62×87.5 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.

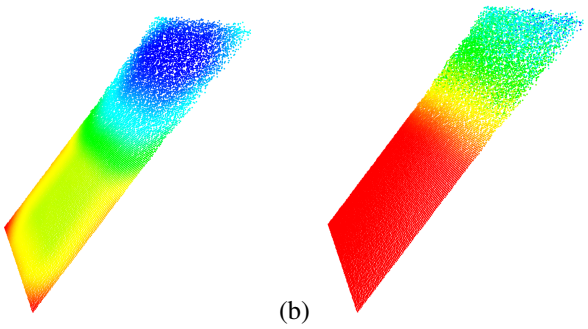
## IV. Method

We asses the accuracy of the different registration methods by three different measures. For datasets where ground truth data is available, e.g., from a motion capture system, we quantify mapping accuracy by the absolute trajectory error (ATE) [26] based on the estimated and the ground truth trajectory.

For assessing pose accuracy without pose ground truth, we calculate two quantitative measures which evaluate the sharpness of a map. These measures are the *mean map entropy* and *mean plane variance*.

For completeness, we also report the measured runtimes, i.e., the average processing time of one 3D laser scan for the

[2]http://vision.in.tum.de/data/datasets/rgbd-dataset



(a) (b)

Fig. 3. Visualization of the MME (a) and the MPV (b). The resulting entropy on an synthetic, wedge-shaped dataset. The color of a point indicates the value of the metric, from red/yellow (low) to green/blue (high).

different approaches and the optimized parameter sets. The runtime is not used in the following parameter optimization. Instead, we focus on the accuracy of the estimated trajectory and optimize the parameters w.r.t. the absolute trajectory error (ATE) [26]. Consequently, some approaches show larger processing times than others simply because of using parameters that achieve (slightly) better ATE scores at the cost of considerably longer runtimes.

### A. Mean Map Entropy

Following [7], the entropy $h$ for a map point $q_k$ is calculated by

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

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

$$H(\mathcal{Q}) = \frac{1}{Q} \sum_{k=1}^{Q} h(q_k). \tag{2}$$

### B. Mean Plane Variance

For the *Mean Plane Variance* (MPV) we make the assumption that most parts of the environment are planar surfaces. Therefore, we approximate a plane from the 3D points in a given radius and calculate the distance of every point to this plane.

The MPV $H(\mathcal{V})$ is averaged over all map points

$$V(\mathcal{Q}) = \frac{1}{Q} \sum_{k=1}^{Q} v(q_k), \tag{3}$$

where $v$ is the upper quartile of the distances in the radius.

Fig. 3 shows a visualization of the MME and MPV on a synthetic dataset.

### C. Parameter Optimization

The registration methods that are included in our evaluation all have different parameters that affect the behaviour of the method. Depending on the scenario or the data set, different parameter values sometimes lead to a significantly different registration result. In order to have a fair comparison between the different registration methods, we optimize these parameters using Hyperopt and the Tree-of-Parzen-Estimators [8]. The parameters of the registration methods are described in the following.

As parameters of the Iterative Closest Points (ICP) [16] algorithm, we use $\boldsymbol{d_{max}}$ as general distance threshold and $\boldsymbol{ICP_{recp}}$ to assure symmetric correspondences. Correspondences are rejected if the point-to-point distance exceeds this threshold $\boldsymbol{d_{max}}$.

The second algorithm in our evaluation is the Generalized Iterative Closest Points (GICP) [15]. We optimize the parameters $\boldsymbol{d_{max}}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{GICP_{inner}}$. As for the ICP algorithm, $\boldsymbol{d_{max}}$ is a correspondence threshold. The parameter $\boldsymbol{\epsilon}$ models the noise in the data and the $\boldsymbol{GICP_{inner}}$ determines the maximum number of inner iterations to optimize the transformation.

The parameters of the Normal Distributions Transform (NDT) [14] are $\boldsymbol{d_{max}}$, $\boldsymbol{\epsilon}$, $\boldsymbol{NDT_{res}}$, and $\boldsymbol{NDT_{step}}$. While the first two parameters are similar to the parameters of GICP, $\boldsymbol{NDT_{res}}$ controls the cell size in which the normal distribution is computed and $\boldsymbol{NDT_{step}}$ is the step size used for transformation optimization.

For the surfel-based soft assignment registration method (Surfel) [7] we also use a parameter $\boldsymbol{S_{prior}}$ to model the noise in the data. Besides that, we use three parameters to weight the soft assignments. First, $\boldsymbol{S_{size}}$ weights assignments by the surfel size. Second, $\boldsymbol{S_{points}}$ weights soft assignments by the number of points they encompass. And lastly, $\boldsymbol{S_{neighbor}}$ weights by the number of surfels in the vicinity.

In addition to parameters of the registration methods, we also optimize for the parameters of the underlying multiresolution grid map: the resolution of the map, the number of map levels, and the maximum number of points that are stored in each grid cell (cell capacity).

The mesh-based registration [6] is used with the default parameters. Here, no parameter optimization has been conducted. In [6], it is distinguished between initial pairwise registration of the raw non-uniform density scans and global optimization of the complete trajectory and map. In order to emulate the behavior of incremental registration against a local egocentric map, we do not use the original pairwise registration as presented in [6] but a variant [28] in which a newly acquired 3D scan is aligned in a local window of multiple 3D scans (five in our experiments). Consequently, the behavior of the resulting approach can be compared to aligning against a local egocentric map containing the points of five 3D laser scans.

## V. Experiments

In order to assess the performance of the registration methods in terms of pose accuracy and map quality, we have recorded different datasets with our MAV in flight. The first two datasets are acquired in a Motion Capture (MoCap) system that provides ground truth pose information. For these datasets, the Absolute Trajectory Error (ATE) can be calculated based on the MoCap poses. We use one dataset from the MoCap system to optimize the parameters of the different registration

methods (training dataset) and another dataset of the MoCap system to evaluate the different methods (test dataset).

Furthermore, we use this test dataset to evaluate our quantitative in-map measures by comparing the MME and the MPV to the ATE of the ground-truth poses. As a fourth metric for the evaluation, we estimate the registration accuracy by computing the point-to-point root mean square error (RMSE) between the constructed map and a ground truth map (GT-RMSE). For this metric, we first align the first point cloud of the dataset manually to the ground truth map and apply the determined transformations of the registration methods. We then search for closest points in the ground truth map for all points in the aligned 3D laser scans and compute the RMSE of the distances between the found point pairs. In the best case, the constructed map and the ground truth map overlap perfectly in the end and the GT-RMSE is close to zero.

We also use an outdoor dataset (Frankenforst dataset) recorded during a mission where the MAV is mapping a building (shown in the form of a colored ground truth point cloud in Fig. 1). Since ground truth pose information is not available in this dataset, we use the GT-RMSE as well as the MME and the MPV of the aligned point clouds to evaluate the registration methods. For both indoor and outdoor datasets, we also compare the runtime needed for every method to process the whole dataset.

In this section, we discuss the parameter optimization for the different methods together with the found parameters, as well as the results of the final evaluation (using the found parameter sets). Furthermore, we investigate the correlation between the MAP-MME and the MAP-MPV, and discuss insights gained during the evaluation.

### A. Parameter Optimization and Parameter Sets

Before comparing the different registration methods, we search, for each method individually, for the best parameters that minimize the ATE in the training dataset. Parameter search is performed using Hyperopt as previously described. We report the found parameter sets in Table I.

To illustrate the applied parameter estimation, we first optimize a single parameter for a single method (the $S_{size}$ factor of the surfel registration). We show the results of this optimization in Fig. 4. The minimum ATE can be achieved for a $S_{size}$ factor of roughly $0.375$. In the actual parameter optimization, it is not only a single parameter being optimized but the set of all parameters the registration method has. In comparison, in this particular example, the best found parameter set includes $S_{size}$ = $0.45$. This clearly shows how the different parameters influence each other and the achievable result.

In addition to the parameter sets, we evaluated three different types of registration:

1) pairwise registration, where scans are sequentially registered against the last acquired scan,
2) incremental registration, where all so far aligned scans form a map that is used as the target point cloud for aligned newly acquired scans, and
3) incremental registration using multiresolution surfel maps (with point lists for methods not using surfels).

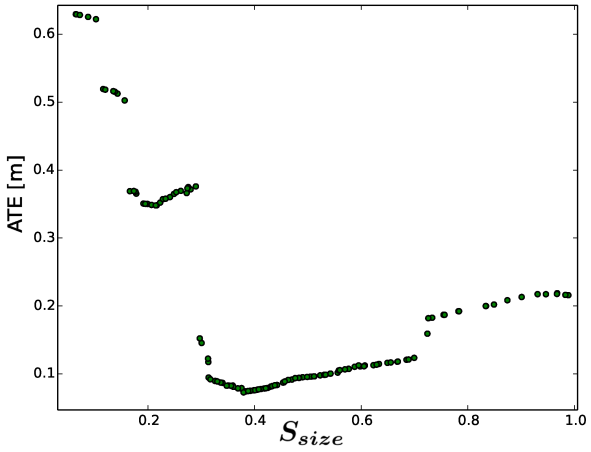



Fig. 4. Optimized $S_{size}$ parameter of the surfel registration method. The plot shows the characteristics of the parameter in relation to the ATE (other parameters fixed).

TABLE I. PARAMETER VALUES OF THE EVALUATED REGISTRATION METHODS AFTER OPTIMIZATION.

| | Parameter | Value | Range |
|---|---|---|---|
| ICP | $d_{max}$: | 2.5206 | [2.5 - 30] |
| | $ICP_{recp}$ | False | {True, False} |
| GICP | $d_{max}$ | 12.3845 | [2.5 - 30] |
| | $\epsilon$ | 0.4169 | [0 - 1] |
| | $GICP_{inner}$ | 13 | [5 - 30] |
| NDT | $d_{max}$ | 9.7048 | [2.5 - 30] |
| | $\epsilon$ | 0.8568 | [0.1 - 0.9] |
| | $NDT_{res}$ | 0.1431 | [0.05 - 1] |
| | $NDT_{step}$ | 0.6596 | [0.05 - 0.95] |
| Surfel | $S_{prior}$ | 0.25 | [0.05 - 0.95] |
| | $S_{size}$ | 0.45 | [0.05 - 1] |
| | $S_{points}$ | 0.9 | [0.1 - 1] |
| | $S_{neighbor}$ | 10 | [1 - 10] |

For all registration algorithms individually, the method of incremental registration using multiresolution surfel maps achieved the best results. We report the different parameters for the multiresolution surfel map for each registration algorithm in Table II. The results reported in the following have all been obtained by using a multiresolution surfel map as the map for incremental registration. The map is updated after every laser scan registration. The surfel-based registration directly uses the computed surfel statistics, the other methods use point lists stored in ring buffers in the multiresolution grid cells.

### B. Evaluation of Registration Algorithms

After optimizing the parameters on the training dataset, we evaluate the different registration methods on two test datasets. The first test dataset was also acquired in our MoCap volume. We report the results obtained using the different metrics in Table III.

In terms of pose accuracy, both ICP and GICP could not improve the initial pose estimates from visual odometry. We

TABLE III. RESULTS FOR THE TWO DATASETS (POSE ACCURACY, MAP QUALITY, AND RUNTIMES).

| | Method | ATE-RMSE [m] | GT-RMSE [m] | MAP-MME [] | MAP-MPV [m] | Runtime* [s] |
|---|---|---|---|---|---|---|
| Motion capture | VO | 0.029977 | 0.0058409 | -2.52013 | 0.1193130 | — |
| | ICP | 0.033751 | 0.0036989 | -3.65248 | 0.0495859 | 1.6659 ± 0.3866 |
| | GICP | 0.039058 | 0.0031005 | -3.41114 | 0.0576033 | 1.3534 ± 1.1927 |
| | NDT | 0.025246 | 0.0021494 | -3.74142 | 0.0460792 | 5.2497 ± 1.7360 |
| | Mesh | 0.024969 | **0.0019273** | -3.80589 | 0.0455877 | 0.1097 ± 0.0382 |
| | Surfel | **0.024224** | 0.0020391 | **-3.80874** | **0.0446906** | **0.0466 ± 0.0140** |
| Frankenforst | VO | — | 0.2722540 | -2.33082 | 0.199625 | — |
| | ICP | — | 0.0660561 | -2.65460 | 0.136387 | 1.0264 ± 0.3828 |
| | GICP | — | 0.0640025 | -2.55505 | 0.151351 | 0.3196 ± 0.2355 |
| | NDT | — | 0.0697013 | -2.63396 | 0.148460 | 2.0627 ± 0.7125 |
| | Mesh | — | **0.0421434** | -2.72391 | 0.137231 | **0.1314 ± 0.0415** |
| | Surfel | — | 0.0937490 | **-2.81387** | **0.121531** | 0.3508 ± 0.1522 |

* Runtimes are measured per 3D scan being registered and given with mean and standard deviation.

TABLE II. OPTIMAL PARAMETERS FOR THE MRSMAPS

| Method | Resolution | Levels | Cell capacity |
|---|---|---|---|
| ICP | 18 | 2 | 2950 |
| GICP | 6 | 3 | 1150 |
| NDT | 11 | 2 | 4050 |
| Surfel | 20 | 2 | 300 |

note that in the smaller MoCap volume, the visual odometry produces already very accurate relative pose estimates and a globally consistent trajectory. Hence, it is not surprising that plain registration methods could not considerably improve the absolute trajectory error (ATE). However, by accurately aligning the acquired laser scans, they achieve better values in the MAP-MME and MAP-MPV scores, compared to visual odometry. This is caused by the fact that the acquired laser scans are locally better aligned than in case of the visual odometry while the global trajectory is slightly more inaccurate, e.g., due to small drifts.

The surfel registration could achieve very good values in almost all metrics. It produces a considerably more accurate trajectory estimate, compared to visual odometry solely. Furthermore, the laser scans are very well aligned. The surfel registration is also the fastest of the compared algorithms. Note, however, that especially the parameter optimization focused only on the accuracies of trajectory and map rather than runtime. The map obtained from this best trajectory estimate is shown in Fig. 5. Only in the GT-RMSE map quality metric, the mesh-based registration achieves a better score than the surfel registration. It follows a very similar mechanism of aligning Gaussian statistics to perform robust scan registration, but on approximate surface reconstructions of the 3D scans.

The mesh-based registration is considerably slower than the surfel registration, but achieves both a comparable ATE and comparable map quality metrics. Overall, both approaches achieve very similar results although they follow two completely different approaches. An interesting fact is that the local window alignment applied in the mesh-based registration achieves a considerably better trajectory estimate (without any drifts) compared to the pairwise registration as reported in [6].

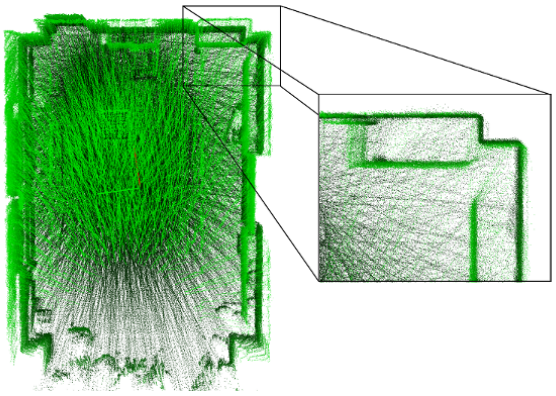


Fig. 5. Topview of the resulting map of the MoCap volume. The map has been registered with the surfel registration method (lowest ATE). The thin walls indicate accurate registration.

In fact, the obtained scores in both absolute trajectory estimate and MAP-MME do not rank behind the globally optimized trajectory in [6]. A likely cause is the aforementioned simplicity of the environment allowing for robustly optimizing the trajectory although it is only optimized in local windows.

The second test dataset—the Frankenforst outdoor dataset—has been recorded during a flight along the facade of a building of *Gut Frankenforst*—a research station operated by the Institute for Veterinary Research at the University of Bonn (see Fig. 6). In the acquired laser scans the building, surrounding vegetation with trees and the ground are visible. This dataset is far more challenging since the surroundings of the building are cluttered and the scene contains fewer visible distinct environmental structures compared to the small indoor motion capture volume. Due to missing ground truth pose estimates in this dataset, we cannot compute the ATE and only report the other performance measures in Table III.

For this dataset, all registration methods improved the initial visual odometry estimates. Here, the trajectory estimate by the visual odometry shows a significant drift. The map obtained from the mesh-based registration achieves the best GT-RMSE. Furthermore, by using a constant number of edges

Fig. 6. Resulting map of the Frankenforst outdoor dataset obtained by the surfel registration (lowest MAP-MME).

between scans being aligned and a constant size of the local window, it achieves nearly constant-time updates and the best runtime for this dataset. In terms of MAP-MME and MAP-MPV, the surfel-based registration shows the best results. The map obtained by the surfel-based registration is shown in Fig. 6.

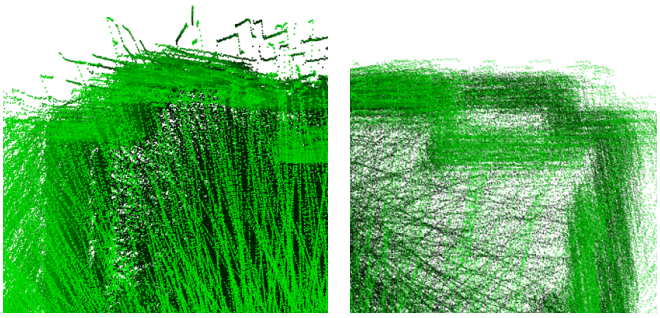
### C. Evaluation of Mean Map Entropy and Plane Variance

In order to evaluate the two proposed in-map measures (MME and MPV), we did a direct comparison. Since we have all four measures available for the indoor MoCap datasets, we can compare MME and MPV with the ATE for all parameter sets tested during optimization. The correlation between the two in-map measures and the ATE of the ground truth poses is shown in Fig. 7. One can see that both relate to the ATE. If the ATE is low both MME and MPV are also low.

A particular shortcoming of both metrics, however, is that they can only be applied in case of globally consistent trajectory estimates. Large registration errors can cause the map of all aligned laser scans to contain regions where points are scattered. In these regions, both the neighborhood searches for the computation of the metrics and our assumption of roughly aligned environmental structures fail. Consequently, both metrics may yield smaller values than those obtained from correct trajectory estimates. Moreover, in case of similar ATEs, MME and MPV can in the aforementioned cases suggest that a solution actually being worse is better. In order to demonstrate this effect, we present a particular example of two parameter sets in Fig. 8. They both achieved similar ATEs in the evaluation with their (not very good) trajectory estimates. The slightly better trajectory, however, obtains a higher MME. In this example, the worse trajectory estimate causes larger volumes around the actual room where single points are scattered. They negatively effect the overall metric since ill-formed neighborhoods are not explicitly handled. Hence, we suggest to only use metrics such as the MME and the MPV to evaluate the laser scan alignments in combination with trajectory errors such as the ATE, or together with visually inspecting the resulting trajectory and map to guarantee global consistency.

## VI. Conclusions

In this paper, we evaluated registration algorithms on data of a light-weight 3D laser scanner mounted on a micro aerial vehicle. The registration of this data is particularly challenging due to the sparsity of the data. Besides a comparison of our surfel-based registration method to state-of-the-art registration methods, we presented different metrics for assessing the quality of the resulting maps. In addition, we determined optimal sets of parameters for all algorithms—except for the mesh-based registration—in a hyper parameter optimization.



(a) Low MME in scattered regions (b) Better, but higher MME

Fig. 8. Snapshots of two maps (similar ATEs) where the obviously worse alignment (a) achieves a lower/better MME than the better alignment in (b).

For our experimental evaluation, we have recorded datasets in a motion capture volume with ground truth pose information and in an outdoor scenario. The datasets are made publicly available. Our evaluation shows that the surfel registration achieved the best values in the Motion Capture dataset (except for the GT-RMSE) and produces a considerably more accurate trajectory estimate compared to visual odometry. Only the mesh-based registration achieves a slightly better score in the comparison of the aligned 3D scans to a ground truth map of the environment (GT-RMSE).

By evaluating the different performance measures, we showed that the introduced MAP-MME and MAP-MPV map metrics correlate to the ATE of the ground truth poses.

## 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 *IEEE International Conference on Robotics and Automation (ICRA)*, 2009.

[3] A. Bachrach, R. He, and N. Roy, "Autonomous flight in unstructured and unknown indoor environments," in *European Micro Aerial Vehicle Conference (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 *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 2968 –2969.

[5] D. Droeschel, D. Holz, and S. Behnke, "Omnidirectional perception for lightweight MAVs using a continously rotationg 3D laser scanner," *Photogrammetrie Fernerkundung Geoinformation (PFG)*, vol. 5, pp. 451–464, 2014.

[6] D. Holz and S. Behnke, "Mapping with micro aerial vehicles by registration of sparse 3D laser scans," in *Proc. of the International Conference on Intelligent Autonomous Systems (IAS)*, 2014.

[7] D. Droeschel, J. Stückler, and S. Behnke, "Local multi-resolution representation for 6D motion estimation and mapping with a continuously rotating 3D laser scanner," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 5221–5226.

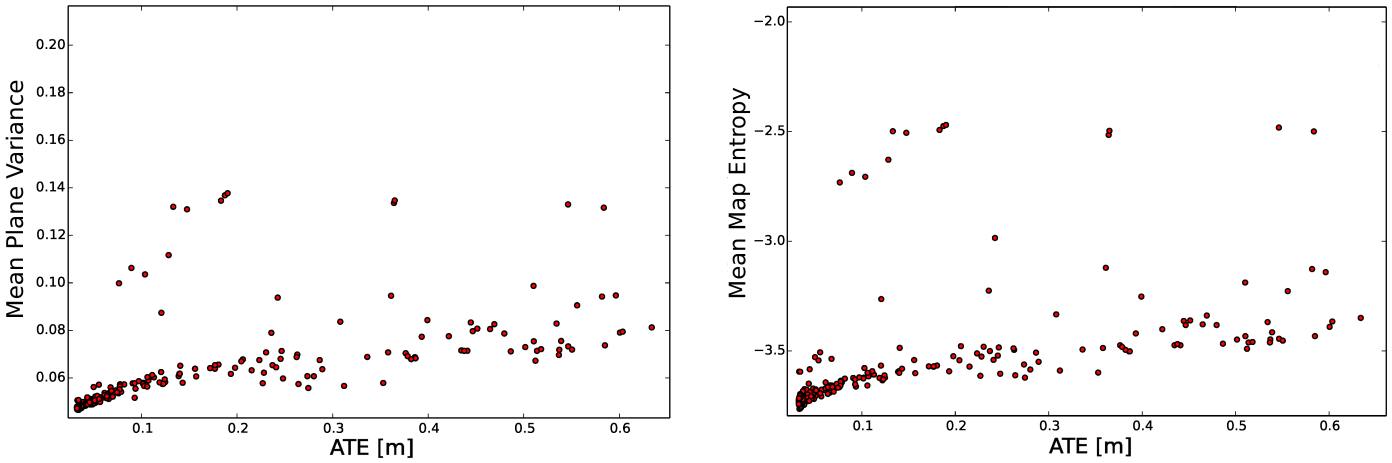



Fig. 7. Correlation between the in-map measures MPV (left) and MME (right) and the ATE of the poses estimated by the surfel registration.


[8] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in Neural Information Processing Systems*, 2011, pp. 2546–2554.

[9] 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.

[10] 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.

[11] 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.

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

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

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

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

[20] F. Pomerleau, F. Colas, and R. Siegwart, "A review of point cloud registration algorithms for mobile robotics," *Foundations and Trends in Robotics*, vol. 4, no. 1, pp. 1–104, 2013.

[21] M. Magnusson, A. Nüchter, C. Lörken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3D registration reliability and speed — a comparison of ICP and NDT," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2009, pp. 3907–3912.

[22] O. Wulf, A. Nüchter, J. Hertzberg, and B. Wagner, "Ground truth evaluation of large urban 6d slam," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2007, pp. 650–657.

[23] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM—3D mapping outdoor environments," *Journal of Field Robotics*, vol. 24, no. 8-9, pp. 699–722, 2007.

[24] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg, "Globally consistent 3D mapping with scan matching," *Robotics and Autonomous Systems*, vol. 56, no. 2, pp. 130–142, 2008.

[25] F. Lu and E. Milios, "Globally consistent range scan alignment for environment mapping," *Auton. Robots*, vol. 4, no. 4, pp. 333–349, 1997.

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

[27] S. Schwertfeger, A. Jacoff, C. Scrapper, J. Pellenz, and A. Kleiner, "Evaluation of maps using fixed shapes: The fiducial map metric," in *10th Performance Metrics for Intelligent Systems Workshop (PerMIS)*, 2010, pp. 339–346.

[28] D. Holz and S. Behnke, "Approximate surface reconstruction and registration for RGB-D SLAM," in *Proceedings of the European Conference on Mobile Robotics (ECMR)*, 2015.