In Proceedings of 11th IEEE International Symposium on Safety, Security, and Rescue Robotics \\\(SSRR\\\), Linköping, Sweden, October 2013.

Multi-Resolution Surfel Mapping and Real-Time Pose Tracking using a Continuously Rotating 2D Laser Scanner

Mark Schadler Jörg Stückler Sven Behnke

Autonomous Intelligent Systems Group
Computer Science Institute VI, University of Bonn, Germany

***Abstract*—Mapping and real-time localization are prerequisites for autonomous robot navigation. They also facilitate situation awareness of remote operators in exploration or rescue missions. In this paper, we propose methods for efficient 3D mapping of environments and for tracking in real-time the 6D movement of autonomous robots using a continuously rotating 3D laser scanner. Multi-resolution surfel representations allow for compact storage and efficient registration of local maps. Real-time pose tracking is performed by a particle filter based on individual laser scan lines. We evaluate our approach using both data generated in simulation and measurements from challenging real environments.**

I. Introduction

Many challenges must be solved before a robot can successfully be deployed into an open environment—one of these being safe and reliable navigation. Search and rescue missions, in addition, involve time constraints and pose safety risks. Here, it is essential to provide a remote operator with sufficient situation awareness. The creation of 3D environment maps and real-time robot localization contribute to this end. These abilities are also a prerequisite for autonomous robot navigation, which can reduce operator workload. The search and rescue domain frequently poses additional constraints on mapping and localization, such as limited payload, limited onboard computing, and cost-effectiveness.

In this paper, we propose a method for using a continuously rotating small and lightweight 2D laser scanner as the sole sensor for 3D environment mapping and 6D tracking of a mobile robot. The Hokuyo UTM-30LX-EA laser scanner was used for experimentation; to allow for continuous rotation, the sensor was mounted on a slip-ring as pictured in figure 1. This affordable sensor has an omnidirectional field of view (FOV), is independent of lighting conditions, and does not require textured surfaces. When mapping, a trade-off must be considered between map quality and memory and performance requirements. We balance these by using multi-resolution surfel maps, which have been successfully used for real-time RGB-D SLAM on the CPU [1]. This map representation matches well to the density of measurements and noise characteristics of our sensor. It allows for memory efficient storage of the map, because not all measured 3D points are stored, but only local measurement statistics. The surface element representation also supports accurate reconstruction of smooth surfaces, which facilitates drivability assessment. Given a map, tracking is the ability to follow the changing pose of a mobile sensor or robot. The localization problem is the ability to determine both the initial pose of the sensor within a map followed by tracking. Within this work, we focus on the tracking and consider the global localization problem for future research.

The remainder of this paper is structured as follows. In the next section, we review related work on simultaneous localization and mapping (SLAM) using laser scanners and other applications of multi-resolution surfel maps. Sec. III gives an overview of the multi-resolution surfel map framework including global graph optimization for scan registration as well as detailed information concerning scan-to-scan registration. Real-time tracking using a particle filter is described in Sec. IV including a motion and observation model derivation. Sec. V reports evaluation results on the real and simulated experiments using the presented mapping framework.

II. Related Work

Most traditional research on mapping and localization using laser scanners in 3D environments focuses on the 2D sub-problem [2]–[4]. Even some more recent works have remained in two dimensions [5], [6]. Due to the availability of 3D laser sensors, research on mapping and localization in 3D has recently boomed [7]–[9]. Mueller at al. [10] compiled an overview on accuracy and limitations of 6D SLAM using laser scanners and iterative closest point algorithms (ICP).

Registering and storing all measured 3D points poses high computational demands. Consequently, multi-resolution maps have been proposed to maintain high performance and low memory consumption. Hornung et al. [11], for example, implement a multi-resolution approach based on octrees enabling the generation of maps at a user specified resolution. Ryde et al. [12] present a multi-resolution mapping solution using voxel lists for efficient look-up and localization. Both of these approaches considers mapping in three dimensions, however a voxel is the smallest surface unit available within the map. Our approach can model up to six surfels within a voxel based upon the view-direction to more accurately model non-cubic surfaces within voxels.

Thrun et al. [13] implement 2D localization in occupancy grid maps using Monte Carlo algorithms. Khoshelham proposes using solely planar objects for localization in 3D within indoor environments [14]. Kuemmerle et al. [15] apply Monte Carlo localization in Multi-Level Surface maps [16] which represent occupied height intervals on a 2D grid. Klaess et al. [17] model the environment in surfel maps in a fixed resolution, similar to the 3D-NDT [9]. They then localize in

978-1-4799-0880-6/13/$31.00

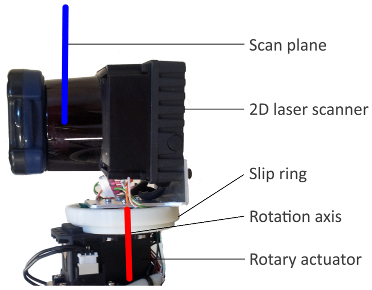

Fig. 1. Continously rotating Hokuyo UTM-30LX-EW laser scanner.

these maps using a tilting 2D laser by matching line elements extracted from the 2D scans in a particle filter framework—assuming motion of the robot in the horizontal plane. Our approach does not constrain the orientation of the robot and allows for 6 DoF motion.

The main contributions of this work include the integration of complementary methods to estimate the robot pose through particle filtering and global graph optimization with 3D scans. We also propose a fast and accurate method to determine the observation likelihood of individual laser scan lines against an allocentric surfel map.

III. 3D Multi-Resolution Surfel Maps

The data structure for representing the environment, the method for registration of local maps, and localization based on individual scan lines are the core elements of our approach.

A. Map Representation

We use multi-resolution surfel maps [1] to efficiently represent environments. Within voxels both surface shape parameters and surface reflectance distributions are stored. Octrees are the natural data structure for multiple-resolution information storage in 3D and thus form the foundation of our mapping system. Compared to a multi-resolution grid implementation, the octree data structure only represents space that has been observed and is thus more memory-efficient. Because storing all measured 3D points would be challenging in larger environments, we aggregate the measurements inside voxels and maintain only their statistical properties. Points $\mathcal{P}$ within a surfel are approximated by a sample mean μ and covariance Σ and are considered normally distributed. These statistical properties are stored through all resolutions in the octree, thus a non-leaf node maintains the statistical properties of all descendants allowing for quick map sampling at any resolution. The maximum resolution at a measured point is determined in dependency of the distance of the point from the sensor. This way, decreasing sampling density with distance from the sensor is captured which is caused by the constant angular resolution of our sensor.

As both shape and reflectance distributions are modeled by surfels, the mean and covariance represent a 4D normal distribution. Shape is simply the 3D spatial position in the map frame while reflectance is represented in the fourth dimension. We estimate the reflectance of the measured surface points from the intensity readings of our laser sensor. One must be careful when using direct reflectance values measured by the laser scanner as these readings are range and viewing-angle dependent. However as reflectances are compared using local contrast as opposed to absolute value, we consider these effects negligated.

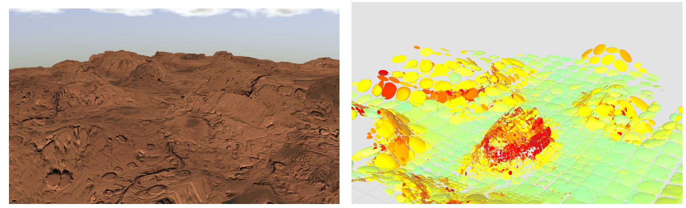

Fig. 2. The simulated terrain environment and accompanying multi-resolution surfel map generated during simulation trials.

Since complex surfaces may not be represented sufficiently well by a single Gaussian even within small volumes, the statistical properties of the measurements are maintained separately for the six orthogonal volume faces using the sensor's view direction to associate data with the appropriate surfel. This allows for a more accurate representation of detailed surfaces, improved surface reconstruction, and reduction of sensor view-angle effects. Figure 2 shows an example surfel map generated in simulation.

B. Efficient Map Registration

The registration of multi-resolution surfel maps is implemented in two main steps: surfel association and pose optimization. Both steps are repeated iteratively until the alignment accuracy reaches a threshold or a maximum number of iterations is reached.

1) Association: Surfel association must robustly associate surfels between a target and source map. Surfels are associated between maps from the finest resolution to coarser resolutions until associations have been determined for the entire map. Consequently, a surfel is only associated if no children have been associated. Using this scheme, associations within resolutions are independent and thus can be computed in parallel on multi-core machines. For performance results, please see section V-D.

For a surfel without a previous association, a volumetric query within the octree is performed to find association candidates. The query volume center is the surfel mean transformed by the current pose estimate with a cube edge size of twice the surfel resolution. Varying the query volume size from the surfel resolution implicitly causes the adaption of misalignments from coarse to fine resolutions. Surfels that have been previously associated are re-associated with the best matching surfel found within the direct voxel-grid neighbors of the previous association.

Associations are made between surfels having the closest Euclidean distance between shape-texture descriptors within the query volume. However these associations are only accepted when this distance $d_f(s_i, s_j)$ is below a threshold $\tau = 0.1$ where $d_f(s_i, s_j) := \sum_{c \in \{sh,r\}} d_c(s_i, s_j)$ [1] is a sum of the shape and texture descriptor distances.

2) 3D Laser Observation Model: Given a target map m_m and a 3D laser scan of an environment, we model pose optimization as finding the pose x that maximizes the likelihood

$p(z|x, m_m)$ of observing the laser measurements z at the pose x in the target map m_m. Poses $x = (q, t)^T$ are represented by a translational part $t \in \mathbb{R}^3$ and unit quaternion q. After creating a map m_s from the scan measurements z, we determine the observation likelihood between the source and target map given a pose x,

$$p(m_s|x, m_m) = \prod_{(i,j)\in A} p(s_{s,i}|s_{m,j}) \tag{1}$$

where A is the set of surfel associations discussed in III-B1 and $s_{u,v} = (\mu_{u,v}, \Sigma_{u,v})$ is the surfel v in map u. As we model surfels as normal distributions, we can easily calculate the observation likelihood of two associated surfels,

$$\begin{aligned} p(s_{s,i}|s_{m,j}) &= \mathcal{N}(d_{i,j}(x); 0, \Sigma_{i,j}(x)) \\ d_{i,j}(x) &:= \mu_{m,j} - T(x)\mu_{s,i} \\ \Sigma_{i,j} &:= \Sigma_{m,j} + R(x)\Sigma_{s,i}R(x)^T \end{aligned} \tag{2}$$

where $T(x)$ is the homogeneous transform matrix of pose estimate x and $R(x)$ is the corresponding rotation matrix.

3) Pose Optimization: To determine the map pose x of the observation, we optimize the logarithm of the observation likelihood from Eq. (2):

$$L(x) = \sum_{a\in A} log\left(|\Sigma_a(x)|\right) + d_a^T(x)\Sigma_a^{-1}(x)d_a(x) \tag{3}$$

in two stages. ICP uses a closed-form calculation to determine registration parameters. However this requires surfels to have a diagonal covariance matrix. As we have potentially non-diagonal covariances, first we calculate an initial transform using Levenberg-Marquardt (LM) optimization then using Newton's method for refinement.

4) Pose Uncertainty Estimation: An estimate of the observation pose uncertainty is calculated using the following closed-form approximation [1], [18]

$$\Sigma(x) \approx \left(\frac{\partial^2 L}{\partial x^2}\right)^{-1} \frac{\partial^2 L}{\partial s \partial x}\Sigma(s)\frac{\partial^2 L}{\partial s \partial x}^T \left(\frac{\partial^2 L}{\partial x^2}\right)^{-1} \tag{4}$$

where x is the pose estimate, s denotes associated surfels in both maps, and $\Sigma(s)$ is the covariance of the surfels.

C. Tracking with Dense 3D Laser Scans

Graph optimization is used to globally optimize the tracked pose from 3D laser scans. For each input 3D scan, a key-view v_i (pose visualized by a coordinate frame) is extracted along the sensor view trajectory and globally aligned to a reference key-view. This alignment implies a geometric constraint between the key-views and is thus maintained as an edge $e_{i,j} \in \mathcal{E}$ in a key-view graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. As an additional step, all key-views deemed "close" are registered against each other to add edges to the constraint graph. The key-views are optimized in a probabilistic pose graph using the g2o framework [19].

After each 3D scan has been converted into a local multi-resolution surfel map and inserted into the key-view graph, the updated poses of all key-views are used to create a global surfel map. This global map is used by the particle filter for observations.

IV. Monte Carlo Tracking using Scan lines

The 6-DoF pose of the laser scanner between key-views in the surfel map is tracked with a particle filter. This pose estimate is used to initialize the registration transform for the next key-view. Compared to other filtering methods that allow for non-linear motion and measurement models such as an Extended Kalman Filter (EKF), the particle filter allows for simple integration of a scan line to map measurement model and also has the ability to solve the global localization problem for future applications. The general idea of a particle filter is not discussed here; see [20] for a detailed introduction.

A. State Propagation Model

To compute a suitable state estimate, we model the state transition with a simple time-discrete linear dynamics model. The state estimation problem is posed as the estimation of the full 6-DoF configuration of the laser scanner $x_t = (r, t)^T$ represented by a translation part t and Euler angles r. Odometry $o_t = (\Delta r, \Delta t)^T$ is also given as input to the propagation model with assumed Gaussian noise in both translational and rotational parts. For convenience, we indicate the translational part as $t(x)$ and the Euler angles as $r(x)$ for a given pose x. We model the dynamics as a time-discrete linear dynamic system (LDS):

$$\begin{aligned} t(x_t) &= t(x_{t-1}) + t(o_t) + \mathcal{N}(n; 0, \Sigma_{o_t}) \\ r(x_t) &= r(x_{t-1}) + r(o_t) + \mathcal{N}(n; 0, \Sigma_{o_r}) \end{aligned} \tag{5}$$

where Σ_{o_t} represents the noise constant of odometry translation and similarly Σ_{o_r} the noise constant of odometry rotation.

B. Observation Model

The observation model of a scan line is similar to the model for 3D scan registration with modifications. As a single scan line does contain enough data to accurately register to a target surfel map; we circumvent this issue using the particle filter and simply determine the likelihood of a scan given a particle pose x. The observation model measures the alignment of the scan line $Z = \{z_i\}^n \in \mathbb{R}^3$ with n point measurements to the current target map m_t given a particle pose x,

$$\begin{aligned} p(Z|x, m_t) &= \prod_{i=1}^{n} p(z_i|x, A(z_i, x, m_t)) \\ A(z_i, x, m_t) &= \operatorname*{argmin}_{s_t} d_{plane}(s_t, T(x)z_i) \end{aligned} \tag{6}$$

where $A(z_i, x, m_t)$ associates the transformed measurement point $z_{transformed} = T(x)z_i$ to the surfel within the target map having the smallest distance between the surfel plane and the target point. The potential surfels to associate are found using a volumetric query around the transformed measurement with a volume size corresponding to the distance of the measurement to the sensor. If no association can be found within the region, the observation likelihood of the point is given a default no-association-likelihood corresponding to the sensor models false/random measurement probability. Given a surfel association, the observation likelihood of a measurement to surfel is given by

$$\begin{aligned} p(z_i|x, A(z_i, x, m_t)) &= p(z_i|x, s_t) \\ &= \mathcal{N}(d_{plane}(s_t, T(x)z_i; x); 0, \Sigma_{sz}(x)) \\ \Sigma_{sz}(x) &= \Sigma_{s_m} + R(x)\Sigma_{z_i}R(x)^T \end{aligned} \tag{7}$$

where Σ_{z_i} represents the point measurement covariance in $\mathbb{R}^3$ and $d_{plane}(s_t, T(x)z_i)$ represents the distance from the s_t surfel plane and the transformed measurement $T(x)\, z_i$.

C. Importance Weights

The importance weights $w^{(i)}$ in a particle filter compensate for the mismatch between the target and proposal distributions $w^{(i)} = \frac{\text{target distribution}}{\text{proposal distribution}}$. Since we use the motion model as the proposal distribution, the importance weights for each particle is given by its observation likelihood,

$$w^{(i)} = p(Z|x^{(i)}, m_t). \tag{8}$$

After importance weight calculations, the particles are resampled with probability proportional to the importance weight.

V. Experiments

Experiments were performed in real environments similar to those encountered during search and rescue missions. Additional experiments were performed in simulation. The accuracy of pose tracking is discussed for all experiments as well as the performance characteristics of mapping and tracking.

For all experiments, the laser-mounted vehicle was placed in an unknown pose and a complete 3D scan was accumulated defining the map frame. After initialization, as no exploration strategy is implemented, the vehicle is moved by manual control and brought to rest. A 3D laser scan is then taken at the new location and globally registered against the existing scans. This process is repeated until the end of the experiment.

A. Sensor Operation

Our research uses "stop-and-go" to accurately map an environment. This process consists of the robot maintaining a stationary pose as a full 3D scan is taken. After this "stop" 3D scan is completed, the sensor measurements are registered against the internal map and the resulting pose estimates updates the particle filter belief. The robot may then "go" to a region of interest for further mapping while tracking is enabled to initialize the next 3D scan registration. Ideally this process is repeated until an environment has been sufficiently explored.

In order to accurately map large environments sensor measurements must be sufficiently dense. To ensure dense point clouds from 3D laser scans, we rotate the laser at a slow speed ($\frac{1}{15}$ Hz) when creating full 3D scans. For tracking we wish to maximize visible space over time period and thus rotate the laser at a higher speed (1 Hz) during navigation.

B. Outdoor Environments

For all outdoor environment experiments, a laser-scanner was mounted on a mobile vehicle that recorded odometry and sensor data. This data was processed in real-time using an Intel Core i7-3610QM running at 2.3GHz with 16 GB RAM. An Intel Core i7-4770K (max 3.5 GHz) with 32 GB RAM is available on the robot, however was not used due to the current integration state.

Localization is performed using 250 particles with a constant noise variance of 0.025^2 for translation and 0.005^2 for orientation. Additional noise is added proportional to the movement delta with roll and pitch noise discounted by a factor of 0.02 since it was accurately measured with an IMU.

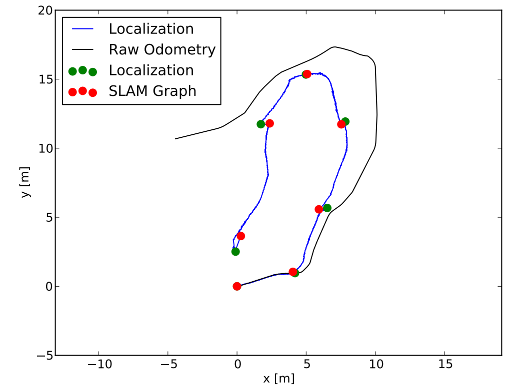

Fig. 3. Parking garage experiment's tracked vehicle movement viewed from above - blue path. Black path - raw odometry input. Green circles - the tracking estimate when starting a 3D registration. Red circles - the registration estimate.

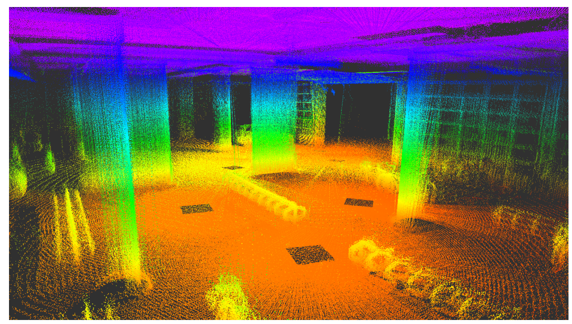

Fig. 4. Parking garage experiment registered point clouds coloured by height for visualization.

1) Parking Garage: The parking garage environment is relatively large for an enclosed space (approximately 25m x 60m) and contains various structures including vehicles, girders, support beams, and windows. Within this environment, 7 full 3D scans were taken with an average distance of 5.53 meters between key-views. Ground truth was measured by hand and consists of the relative distances between scanning poses. Through this metric, registration received a minimum error of 2 centimeters with a maximum error of 13 centimeters. Table I details the tracking accuracy compared to both ground truth measurements and the graph optimized registration pose estimate. Figure 3 shows the estimated path from scan line tracking including registration poses. Figure 4 shows the aligned point cloud generated during this experiment.

2) Disaster Courtyard: Testing within the open-air courtyard environment is more difficult for registration as less strong structural features such as floor/ceiling relationships, corners,

TABLE I. Parking Garage Distances Between 3D Scans

Movement (meters)	1	2	3	4	5	6	Avg. Error
Ground Truth	4.20	4.94	6.53	4.45	4.54	8.54	
Tracking	4.30	5.26	6.39	4.44	4.85	9.40	0.29
Slam Graph	4.18	4.90	6.37	4.40	4.49	8.43	0.07

TABLE II. COURTYARD DISTANCES BETWEEN 3D SCANS

Movement (meters)	1	2	3	4	5	6	Avg. Error
Ground Truth	6.60	8.25	6.59	6.69	5.69	4.37	
Tracking	6.59	8.26	6.47	7.27	5.97	4.57	0.20
Slam Graph	6.58	8.18	6.55	6.66	5.67	4.35	0.03

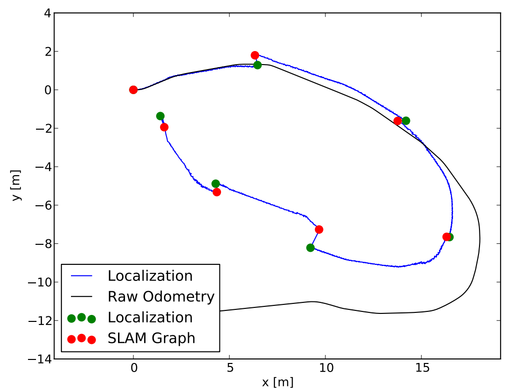

Fig. 5. Courtyard experiment's tracked vehicle movement viewed from above - blue path. Black path - raw odometry input. Green circles - the tracking estimate when starting a 3D registration. Red circles - the registration estimate.

etc. exist. The selected location was relatively feature-bare, thus boxes and large containers were arranged to create a more realistic test environment. Ground truth information is measured by relative distances between 3D scan locations. During the experiment, the average distance between key-views was 6.37 meters while the minimum and maximum errors from the SLAM-graph were 0.00 and 0.05 meters respectively. Table II details the relative distances between ground truth, tracking, and graph optimization. Figure 5 shows the estimated path from scan line tracking while figure 6 shows key-frame positions with the registered point clouds.

C. Simulation

Similar to the real-world experiments, the simulated experiment consisted of a mobile vehicle with a roof mounted laser scanner that is rotated to measure 3D scans. In simulation, odometry is generated from the true movement of the robot with systematic noise and used to update the particle filter. Ground truth information is known allowing for exact evaluation of the mapping and tracking systems. Figure 2 shows the

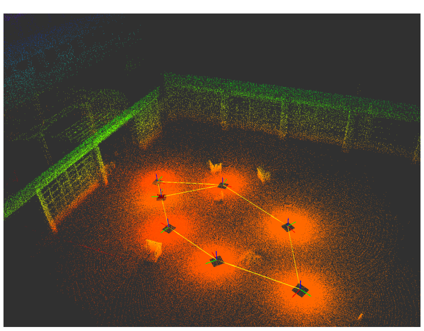

Fig. 6. Courtyard experiment registered point clouds coloured by height for visualization.

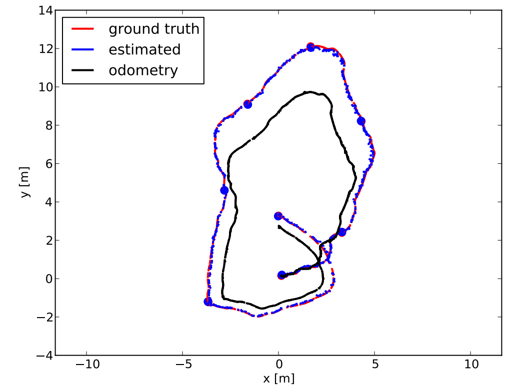

Fig. 7. Simulation experiment plots showing localization trajectory compared to odometry and ground truth. Blue scatter points show Graph-SLAM estimates, red scatter points (barely visible) are ground truth.

terrain environment used for experiments.

Sturm et al. suggest using absolute trajectory error (ATE) to evaluate SLAM trajectories [21]. ATE aligns the estimated and true path using timestamps for data-association and then calculates the absolute pose differences.

Table III shows ATE results from one simulation experiment for both localization and global registration estimates. A maximum localization error of ca. 0.4 meters is not ideal - however an average error of ca. 7 centimeters allows for excellent registration input. Graph-SLAM had an average error of ca. 3 centimeters with a maximum error of ca. 6 centimeters. For scale, figure 7 shows the trajectories for localization and Graph-SLAM.

D. Performance Characteristics

Figure 9 shows the time required to create the multi-resolution surfel map depending on point count. To account for timing error, each cloud was processed ten times and construction times were averaged. To frame these results, the insertion of ca. 1.23 million points took 2.792 seconds, while the accumulation of sensor data took 15 seconds.

Figure 8 shows the time required to perform graph SLAM during the garage experiments. Section III-C explains that all poses within a threshold distance are connected through the graph optimization, thus requiring map to map registration. This accounts for time growth after adding additional scans. However as a robot would not remain in a small area, the number of close key-views would eventually plateau and result in steady 3D scan addition times.

TABLE III. ATE ERROR METRICS FOR LOCALIZATION AND GRAPH-SLAM DURING SIMULATION.

Metric	Localization Error (m)	Graph-SLAM Error (m)
RMSE	0.084	0.037
Mean	0.069	0.034
Median	0.060	0.033
Std	0.072	0.015
Min	0.001	0.009
Max	0.397	0.061

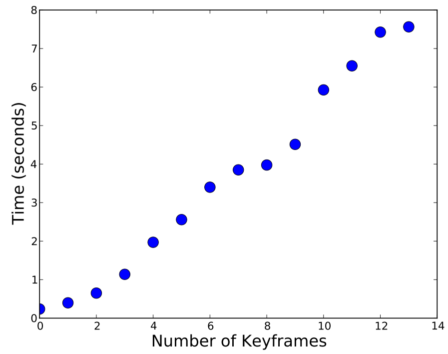

Fig. 8. Time for adding a 3D scan given previous key-view count.

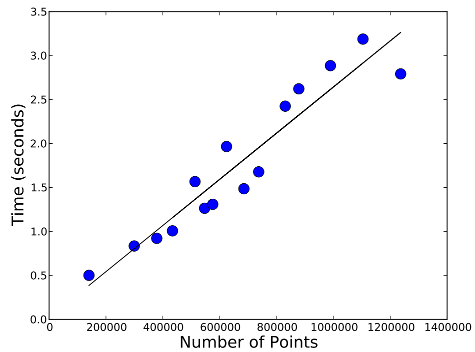

Fig. 9. Multi-resolution surfel map construction time.

Localization using the particle filter is limited by the laser scanner data acquisition rate. Using the Hokuyo laser scanner running at 40 Hz, the localization rate was observed between 30 Hz and 35 Hz depending upon the number of valid measurements from the scan lines.

VI. Conclusion

In this paper an approach for mapping and 6D pose tracking using a single laser scanner was presented. By combining multi-resolution surfel maps to efficiently represent the environment and a global optimizing SLAM graph used to align point cloud data, this method allows for robust and memory-efficient mapping suitable for path planning, object recognition, and other perception tasks. Real-time tracking using individual laser scans is possible through particle filtering and an allocentric multi-resolution surfel map.

Both real and simulated experiments have been demonstrated to successfully perform mapping and tracking tasks. During real-environment experiments, accuracy from SLAM-graph registration estimates had a maximum error of 13 centimeters, while maintaining an average of 7 centimeters and 3 centimeters in the garage and courtyard experiments respectively.

Accuracy within the simulated environment using noisy odometry and sensor readings indicates robustness within the framework, as an average SLAM estimate had an average of ca. 3 centimeters error.

Future work includes integrating occupancy mapping for an improved observation likelihood model and integrating further sensors into the framework. Additionally, while the current implementation is heavily parallelized, research on the applicability of GPUs to further reduce processing time is of interest. Similarly, the integration of additional sensors within the framework could lend itself well to future research topics.

References

[1] J. Stückler and S. Behnke, "Multi-resolution surfel maps for efficient dense 3D modeling and tracking," *Journal of Visual Communication and Image Representation*, 2013.

[2] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo localization for mobile robots." in *IRCA*, 1999.

[3] D. Fox, "Adapting the sample size in particle filters through kld-sampling." in *Int. Journal of Robotics Research (IJRR)*, 2003.

[4] D. Fox, W. Burgard, and S. Thrun, "Markov localization for mobile robots in dynamic environments." in *J. of Artifical Intelligence*, 1999.

[5] K.-H. Lin, C.-H. Chang, A. Dopfer, and C.-C. Wang, "Mapping and localiyation in 3d environments using a 2d laser scanner and a stereo camera," *Journal of Information Science and Engineering 28*, pp. 131–144, 2012.

[6] B. W. Christian Brenneke, Oliver Wulf, "Using laser range data for 3D SLAM in outdoor environments," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2006, pp. 1556–1563.

[7] J. Ryde and H. Hu, "Cooperative mutual 3D laser mapping and localization," in *Proc. IEEE Int. Conf. on Robotics and Biomimetics*, 2006.

[8] A. Nuechter, K. Lingemann, J. Hertzberg, and H. Surmann, "6d slam with approximate data association," in *Int. Conf. on Advanced Robotics*, 2005, pp. 242 – 249.

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

[10] M. Mueller, H. Surmann, K. Pervoelz, and S. May, "The accuracy of 6D SLAM using the AIS 3D laser scanner," in *Proc. IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, 2006.

[11] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: an efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, pp. 189–206, 2013.

[12] J. Ryde and H. Hu, "3D mapping with multi-resolution occupied voxel lists," *Autonomous Robots*, vol. 28, pp. 169 – 185, 2010.

[13] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust monte carlo localization for mobile robots," *Artificial Intelligence*, vol. 128, no. 1-2, pp. 99–141, 2000.

[14] K. Khoshelham, "Automated localization of a laser scanner in indoor environments using planar objects," in *Proc. Int. Conf. on Indoor Positioning and Indoor Navigation*, 2010.

[15] R. Kuemmerle, R. Triebel, P. Pfaff, and W. Burgard, "Monte carlo localization in outdoor terrains using multi-level surface maps," in *Proc. of the Int. Conf. on Field and Service Robotics (FSR)*, 2007.

[16] R. Triebel, P. Pfaff, and W. Burgard, "Multi-level surface maps for outdoor terrain mapping and loop closing," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2006.

[17] J. Kläß, J. Stückler, and S. Behnke, "Efficient mobile robot navigation using 3D surfel grid maps," in *Proc. German Conf. on Robotics (ROBOTIK)*, 2012.

[18] A. Censi, "An accurate closed-form estimate of ICP's covariance," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2007, pp. 3167–3172.

[19] R. Kuemmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A general framework for graph optimization," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2011, pp. 3607–3613.

[20] Z. Chen, "Bayesian filtering: From Kalman filters to particle filters, and beyond," *Statistics*, pp. 1–69, 2003.

[21] 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), 2012 IEEE/RSJ International Conference on*, 2012, pp. 573–580.