2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Normal Distributions Transform Occupancy Map Fusion: Simultaneous Mapping and Tracking in Large Scale Dynamic Environments

Todor Stoyanov, Jari Saarinen, Henrik Andreasson and Achim J. Lilienthal
Center of Applied Autonomous Sensor Systems (AASS), Örebro University, Sweden

Abstract— **Autonomous vehicles operating in real-world industrial environments have to overcome numerous challenges, chief among which are the creation of consistent 3D world models and the simultaneous tracking of the vehicle pose with respect to the created maps. In this paper we integrate two recently proposed algorithms in an online, near-realtime mapping and tracking system. Using the Normal Distributions Transform (NDT), a sparse Gaussian Mixture Model, for representation of 3D range scan data, we propose a frame-to-model registration and data fusion algorithm — NDT Fusion. The proposed approach uses a submap indexing system to achieve operation in arbitrarily-sized environments. The approach is evaluated on a publicly available city-block sized data set, achieving accuracy and runtime performance significantly better than current state of the art. In addition, the system is evaluated on a data set covering ten hours of operation and a trajectory of 7.2km in a real-world industrial environment, achieving centimeter accuracy at update rates of 5-10 Hz.**

I. Introduction

In recent years, an increasingly large number of complex autonomous vehicles have been deployed in industrial environments. In order to operate successfully and efficiently, such vehicles need to be capable of determining their position and orientation with respect to the factory environment in a reliable, repeatable and accurate manner. This need has resulted in a large number of Automated Guided Vehicle (AGV) systems, which rely on additional infrastructure installed in the operational environment. While such systems operate reliably, there are several drawbacks: physically installing the reference beacons is time consuming and costly, modifies the environment and constrains vehicle operation to specific areas. Using alternative, map-based localization approaches is thus an important direction of development for future automated industrial vehicles.

The problem of map-based localization is well researched in the scientific community and offers a solid base for the next generation of industrial AGV systems. Simultaneous Localization and Mapping (SLAM) systems have matured to a state in which two-dimensional maps are ubiquitous in robotic research but have not yet gained acceptance in industrial systems. With the increasingly widespread availability of 3D range sensors and the deployment of AGV systems in more challenging scenarios, the development of accurate and reliable 3D mapping and localization algorithms is rapidly becoming an important industrially relevant research topic.

SLAM in large-scale three-dimensional industrial environments is a challenging problem in many respects. Vehicles may move at high speeds and cover long distances in diverse and large-scale environments, ranging from storage halls, through production areas, mining or construction sites. In addition, these environments are highly dynamic, featuring both slow changes when various goods are delivered for storage or towed away, as well as fast changes induced by numerous other moving vehicles. Thus, an industrially relevant 3D SLAM system needs to be capable of reliable performance in an environment with varying amount of dynamic events over long periods of time.

In this work we propose a mapping and tracking system[1] that aims at addressing the challenges of 3D mapping in industrial environments. We build upon the recently proposed Normal Distributions Transform Occupancy Map (NDT-OM) [1] representation to produce consistent maps in dynamic environments at real-time update rates. We propose to track the vehicle pose using a frame-to-model registration approach, based on the recent NDT Distribution-to-Distribution registration algorithm [2] and iteratively fuse the sensor data into the NDT-OM map. Using a submap indexing approach the system is capable of representing large scale environments, with combined registration and fusion update times ranging between 100 milliseconds and 2 seconds, depending on configuration parameters, the sensor used and the environment. The proposed system is evaluated on a large-scale public data set [3] from a city-block sized environment, yielding absolute trajectory errors (ATE) as low as 1.7 meters over a 1.5 kilometer trajectory. In addition, a ten-hour long data set from a real world deployment of an AGV in a milk production facility is used to evaluate the system, resulting in ATEs of under ten centimeters and update rates of about 5-10Hz.

In the rest of this article we first review the relevant contributions in the area of SLAM and overview the NDT-OM and NDT-D2D algorithms. Next, Section III describes the proposed mapping and tracking approach. Section IV presents the evaluation metrics and data sets used, before proceeding to analyze the system performance. Finally, section V conclude with a summary of the key findings and discussion of limitations and future work.

II. Background

A. Mapping and Tracking Systems

Simultaneous Localization and Mapping is one of the most prolific research areas in robotics and a comprehensive overview of the field is beyond the scope of this work.

[1]source code available at `http://code.google.com/p/oru-ros-pkg`

Instead, we briefly summarize some key contributions in two sub-classes of SLAM approaches relevant to this work.

Maximum-likelihood SLAM algorithms [4] take a simple two step approach to SLAM — first, the most likely range sensor pose is determined and then the information is fused into the map. While this class of approaches has been largely outperformed by more sophisticated recent algorithms, they offer a simple strategy to achieve locally consistent mapping. The obvious drawback of ML SLAM algorithms is that they do not handle loop closures and thus are not capable of correcting the accumulated pose error, resulting in a possibly unbounded divergence from the true trajectory. Nevertheless, this type of approaches have recently been re-examined by the scientific community, with prominent works such as the Kinect Fusion algorithm [5] producing impressive tracking performance in small-scale environments. In this work we refer to this type of SLAM approaches as "Mapping and Tracking", to distinguish them from globally consistent loop-closure SLAM. The approach proposed in this work falls in this category of mapping algorithms.

The second relevant class of SLAM systems also follow a two step approach: first, the sensor data is used to extract constraints on the vehicle motion, and then those constraints are used to minimize the global trajectory error. The first step is often referred to as the front-end and is performed using registration algorithms, such as ICP [6], [7], Generalized ICP (GICP) [8] or NDT registration [9], [2]. The second step — or back-end of the SLAM system, performs the trajectory optimization and is often based on pose graphs [10], [11]. In order to over-constrain the optimization problem, a method for explicitly detecting loop closures [12], [13] in an automatic fashion is also required. Fusing together all of the above components of a SLAM system can be a challenging task, with many free parameters. In order to avoid comparisons against a sub-optimally configured baseline SLAM approach, in this work we compare directly against results previously reported by the scientific community [14], [15] on a challenging standardized data set [3].

B. NDT-OM

The Normal Distributions Transform was originally developed in the context of 2D laser scan registration [16]. The central idea is to represent the observed range points using a set of Gaussian probability distributions. The application of NDT to the problem of 3D range data registration is discussed in more detail in the next subsection. In the context of the system presented in this work, it is important to first summarize the general properties of the NDT representation and the recently proposed NDT Occupancy Map [1] extension. Given a point set, its NDT model is created by discretising space using a regular grid and fitting a Gaussian probability density function $\mathcal{N} = \{C_i, \boldsymbol{\mu}_i\}$ to the samples in each voxel (the data points used in estimating the Gaussian pdf are then discarded). This formulation of NDT is feasible when modeling a single point cloud, but has several shortcomings — it does not explicitly model free space and cannot be used in an incremental fashion. The NDT-OM mapping approach [1], used also in this work, improves on the process by additionally tracking the probability of occupancy of each cell, the consistency of each Gaussian distribution and offering efficient incremental update procedures, maintaining numerical stability over an unbounded number of update range points. The NDT-OM approach assumes point clouds collected by a mobile range sensor and provides incremental, viewpoint- and dynamics-aware model updates. The approach has been demonstrated to produce consistent maps in the context of mapping with known poses in large scale dynamic industrial environments [1].

C. NDT-D2D

Magnusson et al. [9] applied the NDT representation to the domain of 3D scan registration. The central idea in their approach is to maximize the likelihood of points from one range scan, given the NDT model created from a previously known reference 3D scan. In a recent work Stoyanov et al. [2] propose an extension of the registration approach which operates solely on NDT models. The algorithm — NDT Distribution-to-Distribution (D2D) — minimizes the sum of L_2 distances between pairs of Gaussian distributions in two NDT models. Formally, the transformation between two point sets $\mathcal{M}$ and $\mathcal{F}$ is found by minimizing:

$$f(\boldsymbol{p}) = \sum_{i=1,j=i}^{n_{\mathcal{M}},n_{\mathcal{F}}} -d_1 \exp\left(-\frac{d_2}{2}\boldsymbol{\mu_{ij}}^T(R^T C_i R + C_j)^{-1}\boldsymbol{\mu_{ij}}\right) \tag{1}$$

over the transformation parameters $\boldsymbol{p}$, where: $n_{\mathcal{M}}$ and $n_{\mathcal{F}}$ are the number of Gaussian components in the NDT models of $\mathcal{M}$ and $\mathcal{F}$; R and $\boldsymbol{t}$ are the rotation and translation components of $\boldsymbol{p}$; $\boldsymbol{\mu}_i, C_i$ are the mean and covariance of each Gaussian component; $\boldsymbol{\mu_{ij}} = R\boldsymbol{\mu_i} + \boldsymbol{t} - \boldsymbol{\mu_j}$ is the transformed mean vector distance; and d_1, d_2 are regularization factors (fixed values of $d_1 = 1$ and $d_2 = 0.05$ were used). The optimization over $\boldsymbol{p}$ can be done efficiently using Newton method optimization with analytically computed derivatives. The NDT-D2D approach was chosen for use in this work, as it was shown to yield state of the art registration results at competitive runtimes and operates using only NDT models.

III. NDT-OM Mapping and Tracking

The algorithm presented and evaluated in this work makes direct use of the two previously discussed techniques: NDT-OM for map modeling and NDT-D2D for frame-to-model registration. As an initial step, the first sensor range scan is directly inserted in the map, using the standard NDT-OM update step. The initial pose of the vehicle $\boldsymbol{p}_0$ is then taken as the reference point for building the full map. In order to facilitate the subsequent evaluation, we set the initial vehicle pose to the known initial ground truth vehicle position and orientation, obtaining the same coordinate system for the estimated and ground truth trajectories. For each subsequent range scan, we iterate between two steps - track and fuse.

The track step of the algorithm performs an NDT-D2D registration between the acquired range scan $\mathcal{P}_i$ and the map $\mathcal{M}_{NDT-OM}$. First, a local NDT map is created by

inserting the point cloud $\mathcal{P}_i$ into a sensor-centric regular grid of the same spatial resolution as the map and computing the sample mean and covariance in each cell. The obtained local NDT model $\mathcal{M}_{NDT}(\mathcal{P}_i)$ is then offset to an initial guess pose, obtained using the previous estimated vehicle pose and the available on-board ego-motion sensors (IMU, odometry, etc.). Finally, the NDT-D2D algorithm is used to align the local model $\mathcal{M}_{NDT}(\mathcal{P}_i)$ to all the consistently occupied distributions from the global model $\mathcal{M}_{NDT-OM}$. As described in [1], selecting only the consistently occupied distributions from the global map reduces the effect of dynamic entities and thus improves the reliability of the registration procedure.

The registration algorithm used in the track step has been slightly modified from the original algorithm [2], to suit the needs of the proposed system. Unlike the original D2D registration algorithm, at this stage we only perform registration at the map resolution, without employing a coarse-to-fine iteration strategy. This reduces the registration runtime, but sacrifices on the convergence basin of the algorithm and makes a good initial guess more important. Finally, the original registration algorithm presented in [2] makes an approximation when evaluating Eq.(1) — namely, the double sum is only evaluated for pairs of closest neighbors from the two NDT models (a one-to-one data association strategy). In the current implementation we use a more accurate one-to-many approximation by associating a fixed size neighborhood $F(\mathcal{N}_i) = \{\mathcal{N}_j \in \mathcal{M}_{NDT-OM}\}$ for each distribution $\mathcal{N}_i$ from the reference scan. We then evaluate and sum all the L_2 distances between $\mathcal{N}_i$ and each $\mathcal{N}_j$.

Once the tracking step has converged to a candidate pose, the new point cloud is inserted in the NDT-OM. This step involves another local map creation, this time while keeping the local and global grids aligned in order to avoid discretisation effects. The local NDT cells are then used to recursively update the distributions in the NDT-OM, while the sensor pose in the global map is used in a ray-tracing step to update the map occupancy. This approach would already produce good tracking performance in smaller-scale environments, but runs into memory size limitations as the size of the mapped area grows. In order to relax the assumption of an environment of limited size, we use a submap-based indexing procedure for the global map $\mathcal{M}_{NDT-OM}$.

Fig. 1 illustrates the proposed submap tiling approach. Instead of maintaining one big map, we use at all times an active grid of three by three smaller size submaps. Raytracing and distribution update are then accomplished by tracking through multiple submaps. As long as the sensor range (blue circles in Fig. 1) is not larger than the submap size, data is not lost and this approach results in maps identical to a single large map. As the vehicle moves through the environment, the estimated pose may move out of the central submap (such cases occur at the red crosses in Fig. 1). When this happens, a new grid center is set and the set of currently active submaps (thick black square in Fig. 1) changes. Submaps that no longer fall in the currently active grid are then saved to disk, while the newly active submaps are either re-loaded

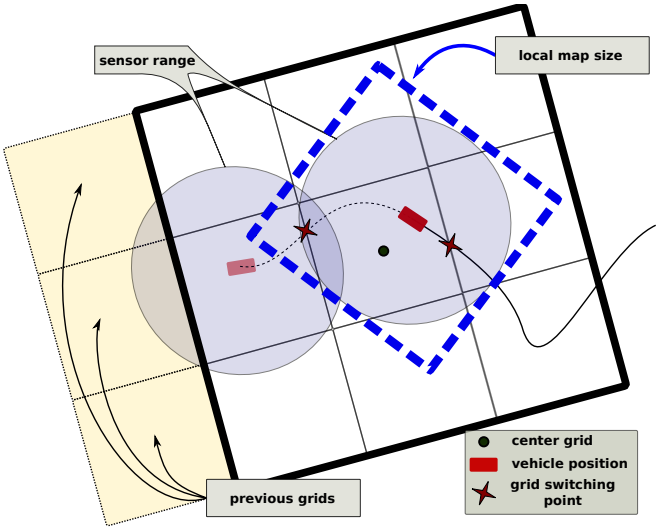


Fig. 1. Illustration of the submap tiling approach. The vehicle (red rectangle) tracks a path through the environment. The grids are allocated in a vehicle-aligned coordinate frame. As the estimated sensor pose moves into the next submap, the grid center switches and old maps (yellow shaded tiles) are saved to disk. New maps are loaded or allocated to keep the grid balanced. A local NDT map is created at each step in a sensor-centric coordinate system, as indicated by the dashed blue line square.

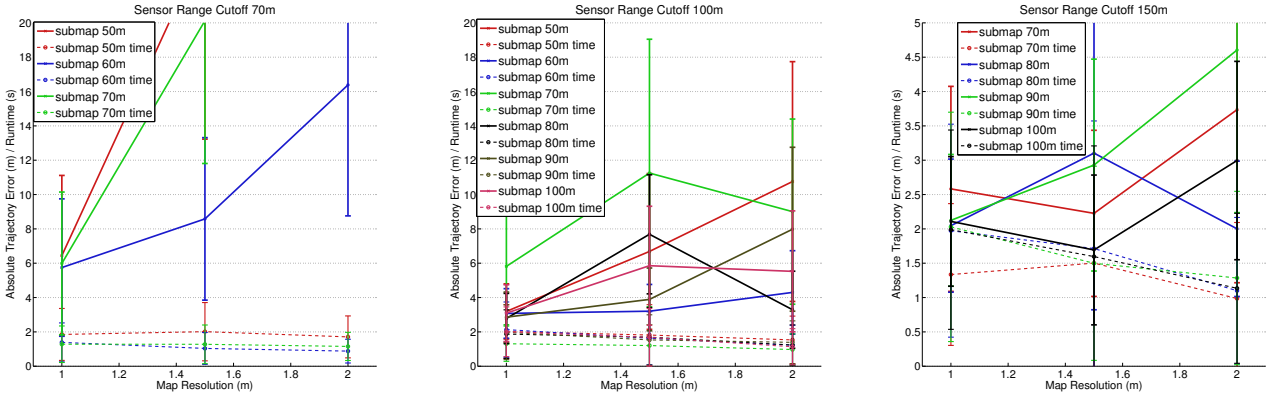
from disk or freshly allocated in cases when the area has not been explored previously. This approach may induce time penalties for map loading and saving, but the effects can be minimized with double buffering and asynchronous read/writes. In this work, we adopt a naive direct saving and loading approach and the extra times are factored into the average update times reported.

Several parameters can be used to influence the accuracy and runtime performance of the proposed mapping and tracking approach. A first important parameter is naturally the cell size of the regular grids used. As shown previously [17], the resolution can affect the accuracy of representation and should be chosen with respect to the sensor data point density. We note however, that the resolutions used by NDT-OM are much larger than typical occupancy map resolutions, as the estimated distributions inside each cell provide an additional representation flexibility. A second important parameter is the size of the local NDT maps used in registration and fusion of new data (illustrated as a blue square in Fig. 1). This size provides an effective sensor cutoff range — measurements that fall outside the local maps are not used in registration or data fusion. Finally, the size of each submap tile governs the extent of the currently visible map and influences registration and fusion runtimes and accuracy. In the subsequent evaluation we will analyze the performance of the proposed approach when varying these three parameters — cell size (or map resolution), local NDT map size (or sensor cutoff) and submap size.

IV. Evaluation and Results

A. Data Sets

The proposed algorithm is evaluated on two challenging large-scale data sets. The first data set [3] was collected



(a) (b) (c)

Fig. 2. Combined absolute trajectory error and runtime plots for different parameters of the mapping and tracking algorithm in the FORD data set. Sensor cutoff ranges of 70 meters (2(a)), 100 meters (2(b)) and 150 meters(2(c))

by a car equipped with a Velodyne HDL-64E lidar scanner, driving a 1.6km trajectory in downtown Dearborn, Michigan USA. The vehicle was equipped with several additional sensors, of which we use two — an Xsens MTi-G IMU for providing an initial guess and an Applanix POS-LV 420 INS for ground truth comparison. The vehicle drives at speeds of up to 15 m/s and provides laser measurements and pose estimates at 10Hz and 100Hz respectively. In the subsequent evaluation, we refer to this data set as the FORD data set.

The second data set is collected by an AGV deployed in a milk production facility located in southern Sweden. The data was collected on top of an AGV in production use and consists of Velodyne HDL-32 measurements at 10Hz together with localization data from the commercial reflector-based navigation system and the on-board odometry at 15Hz. The full data set consists of 10 hours of operation and a total of 7.2km of vehicle track at speeds of up to 2.5m/s. We use a representative 10 minute sample run with a trajectory length of 245 meters in the initial evaluation and then select two parameter configurations for a full 10 hour run test. We refer to this data set as the FACTORY data set. Both the data sets feature many dynamic entities — other vehicles, pedestrians, and moving goods; thus providing a challenging mapping in dynamic environments test case.

In order to evaluate the quality of the obtained estimated sensor trajectories, we compute the absolute trajectory error measure as described in [18]. In essence, the two trajectories are first aligned and then the difference in the 3D pose estimates of corresponding nodes in the two trajectories is computed. We report mean values and standard deviation for the ATE, instead of the relative pose errors as ATE is more suitable for evaluating the overall consistency of the global trajectories. In addition, for the FORD data set we compute the final position offset from ground truth, in order to compare to previously reported results.

B. Tracking from a Moving Car (FORD data set)

Fig. 2(a)-2(c) show the ATE and runtime plots for different configurations of the proposed mapping and tracking algorithm. Each figure shows the mean ATE and standard deviation at three distinct map resolutions - 1, 1.5 and 2 meters. These resolutions were selected as suitable, due to the nature of the available lidar data — namely, sparse samples with high distances from the sensor origin. Each figure also shows the average runtimes for the combined registration and map update cycle, obtained on two cores of a system with an Intel(R) Core(TM) i7-3770K CPU @ 3.50GHz processor and 16GB of memory. Fig. 2(a) shows the performance of the algorithm when a local NDT map of 70x70 meters is used for registration and fusion, at three different submap tile sizes — 50, 60 and 70 meters. Evidently, the estimated trajectory diverges from the ground truth at bigger cell sizes. The only case in which tracking is maintained reasonably is at 1m cell sizes, resulting in mean errors of 6-8 meters. This behavior can be explained by the fact that in some portions of the data set few features are available in the local 70 meter map and, in this range, the large cell sizes don't offer enough expressive representation to maintain tracking. Fig. 2(b) shows considerably more stable performance at 100 meter sensor cutoff across all tested submap sizes and resolutions. Finally, the most accurate and most stable results are obtained when the sensor cut-off is set to 150 meters, thus using almost all of the available sensor readings for registration and map update. Fig. 2(c) shows the tracking performance in this case, at a notably different error scale — all mean ATEs remain below 5 meters, with best performance achieved at 1.5m grid size and 100m submap size — 1.7 meters mean ATE.

At an initial glance, the reported errors might seem larger than expected. Some understanding in the nature of the tracking performance and an explanation of these errors can be gained by examining the actual estimated trajectories. Fig. 3 shows some example reconstructed NDT-OM maps, as well as a comparison of two of the estimated trajectories to the ground truth. From these representative trajectory plots, we can conclude that there is a systematic drift on the z component of the pose, which results in significant ATEs. In some cases (for example, the best performance trajectory with 1.5m grid and 100m submap), the tracked pose re-

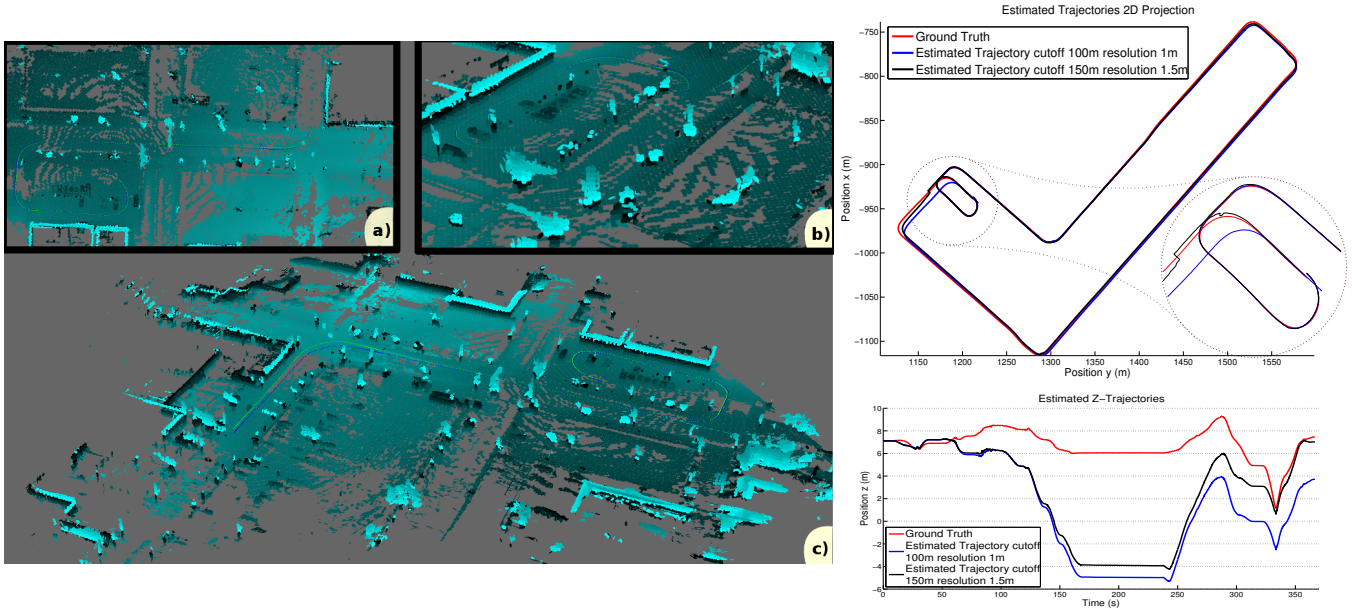


Fig. 3. Mapping and Tracking results on the FORD data set. Left: Maps produced by the system while tracking a) top view, b) zoomed view of the start of the trajectory c) overview. The ellipsoids represent height-coded scaled covariance matrices in each map cell from a map at 1 m resolution. Right column: trajectory plots, at the top x-y trajectory for the 100 and 150m cutoff settings, bottom estimated z position over time. Note the zoomed in detail and the re-entry into previously mapped area.

enters the starting area with a sufficiently small z-component error in order to re-localize with respect to the first submaps and successfully close the trajectory loop. Although the z component exhibits a significant drift, the results are largely locally consistent with the ground truth. According to the specifications of the Applanix INS sensor used for recording the ground truth trajectory, an RMS error of 0.5 meters on the z component can be expected. Other possible sources of error include a sensor-to-vehicle pose miscalibration and intrinsic calibration issues for the Velodyne lidar.

Finally, we compare the obtained results to the performance on the same data set reported by two previously published articles — by Pandey et al. [14] and Tamjidi and Ye [15]. In the first article, Pandey et al. introduce a method to use in addition camera images in order to bootstrap Generalized ICP (GICP). They obtain good frame-to-frame registration accuracy and use the estimates as constraints in an iSAM [10] graph SLAM optimization framework. Unfortunately, the authors do not report any performance measures on the obtained trajectories, other than visually comparing them to the ground truth (with a 2D projection overlaid on satellite images). The results look qualitatively similar to the ones shown in Fig. 3. Tamjidi and Ye [15] also propose to use fused lidar and camera data for a visual-feature based tracking system. They use the correspondence results in a similar fashion to Pandey et al. and obtain globally optimized pose trajectories using iSAM. The authors also report a comparison to an open loop GICP track and a combined GICP and iSAM result. The authors report the error to ground truth only for the final trajectory position — 16.75 meters for GICP, 26.59 meters for the globally optimized GICP trajectory and 27.93 meters for the proposed system. We compare these values to the final position error obtained for several selected NDT-OM parameter configurations in Table I. The final position error is significantly lower for the approach proposed here, with a best-case performance of 1.49 meters. Lastly, we note that the average runtime for the combined tracking and map update step for our system in the configurations tested on this data set is consistently below 2 seconds, compared to an average runtime of 22 seconds reported for the GICP registration algorithm [15].

TABLE I

FORD: FINAL POSITION ERROR, 100M SUBMAP

Sensor Cutoff	1m cell	1.5m cell	2m cell
100m	8.396m	15.576m	14.992m
150m	3.623m	1.498m	7.061m

C. Operation in Industrial Environments(FACTORY data set)

In this subsection we evaluate the performance of the proposed mapping and tracking system in a real-world industrial scenario. Fig. 4(a)-4(f) show the absolute trajectory errors and runtimes for the system, under different parameter configurations. For this data set, we test at grid resolutions of 0.2, 0.4, 0.6, 1 and 1.5 meters. The resolutions are lower than the ones used for the FORD data set, due to the significantly reduced environment size and the higher average sample point density. We test three different submap tile sizes — 30, 40 and 50 meters; at three different sensor cutoff ranges — 30 meters in Fig. 4(a) and 4(d), 50 meters in 4(b) and 4(e), and 70 meters in Fig. 4(c) and 4(f). In addition to testing the presented 3D tracking system, we also perform a test under a flat floor assumption by constraining the registration transformation search to a plane.

The results for the full 3D tracking system in Fig. 4(a)-4(c) show a clearly identifiable increase in accuracy as

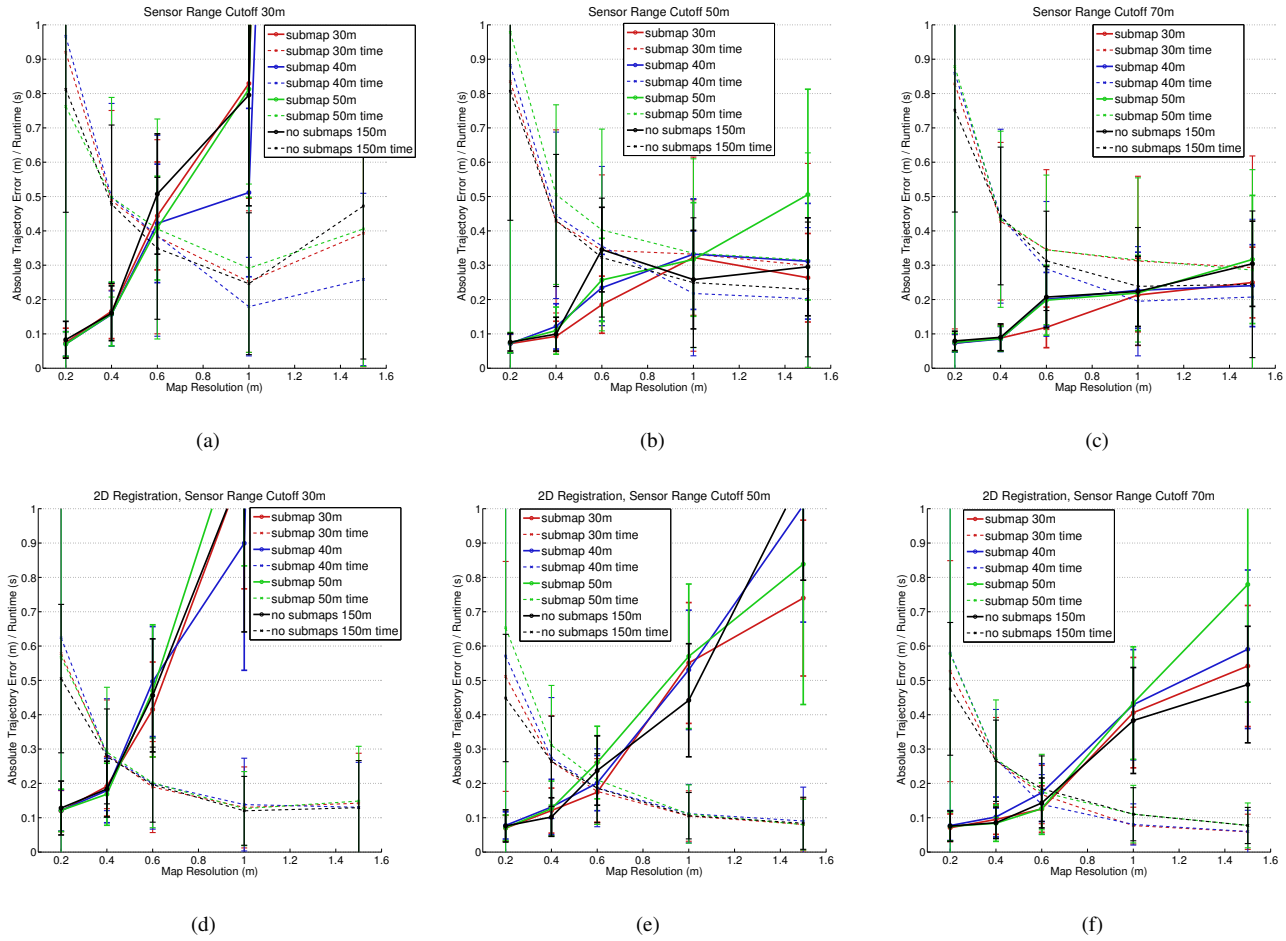


Fig. 4. Combined absolute trajectory error and runtime plots for different parameters of the mapping and tracking algorithm in the FACTORY data set. Fig. 4(a)-4(c): sensor range cutoffs of 30, 50 and 70 meters. Fig. 4(d)-4(f): results using registration constrained to a flat floor assumption.

the sensor range cutoff increases. This mirrors the trend in the FORD data set, suggesting that far-range data is important for obtaining good tracking performance and that limiting the local maps for registration and fusion is not desirable. A second important observation is that the update runtime decreases significantly for larger cell sizes, while the accuracy of the approach deteriorates only slightly. When using the full 70 meter sensor data (Fig. 4(c)) the ATE remains below 30 cm even for cell sizes of 1.5 meters. We found the best performance trade-off in this environment between 0.4 and 0.6 meter cell sizes, resulting in ATEs under 10 cm with update times of about 350ms. Finally, we note that the performance is relatively unaffected by the size of the submap tiles, and in fact seems compatible with a reference monolithic map implementation (labeled no "submaps" in the figures). Fig. 4(d)-4(f) show the corresponding results when using the 2D registration approximation. The results follow a similar pattern as in the 3D case, with accuracy deteriorating slightly faster at larger cell sizes, but at notably faster run times. Again, the best performance is obtained at resolutions about 0.4-0.6 meters and sensor range cutoff of 70 meters — accuracy of about 10-15 cm at runtimes of about 150 ms.

Based on the previously discussed results, we select two parameter configurations — one for 2D and one for 3D tracking and evaluate the system performance on the full FACTORY data set. Snapshots from the obtained NDT-OM maps, as well as a sample full data set trajectory are shown in Fig. 5. The absolute trajectory errors obtained over the full 10 hours of vehicle operation are shown in Table II. The final ATE for a reference no-submap implementation operating on a 150x150 meter map at 0.4m resolution is 10 cm over the full 7.2 km trajectory. When using 40x40 meter tiles at 0.4 m resolution and 70 meter sensor range, we obtain ATEs of 7.6 and 7 cm for the 2D and 3D registration respectively. The slightly better performance of the 3D tracking variation is likely due to a light slope on the factory floor, which can be identified in the final 3D registered map. The obtained trajectory precision is on par with that of the commercial reflector-based localization system, which has an accuracy in the range of 1-10 cm. Thus, the proposed mapping and tracking approach is capable of repeatably obtaining

TABLE II
FACTORY: ABSOLUTE TRAJECTORY ERROR IN 10 HOUR RUN

Method	ATE (m)	σ(ATE) (m)
No Submaps	0.1043	0.0627
2D Registration	0.0761	0.0534
3D Registration	0.0704	0.0348

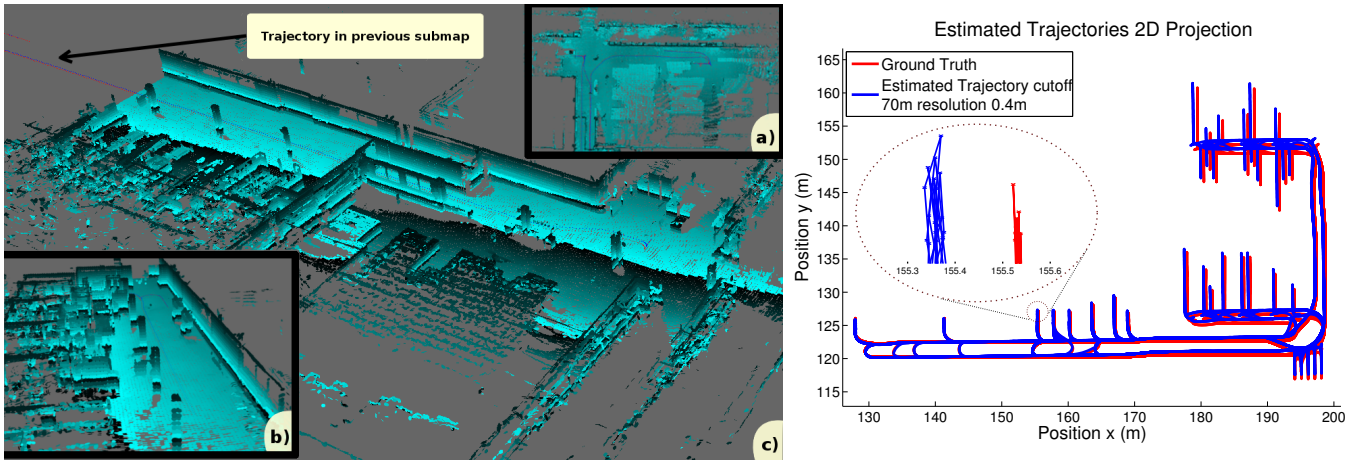


Fig. 5. Mapping and tracking results on the FACTORY data set. Left: maps obtained by the algorithm during operation — a) product delivery part of the facility, b) zoom-in on the production area and c) top-level view from the middle of the vehicle trajectory. The ellipsoids represent height-coded scaled covariance matrices in each map cell from a map at 0.2 m resolution. Right: Trajectories obtained wtih 3D registration compared to ground truth.

accurate pose estimates over long periods of time in a highly challenging dynamic real-world production environment.

V. Summary and Discussion

This article proposed a mapping and tracking approach, based on two recent algorithms — NDT-D2D registration and NDT-OM map building. We use a simple two step maximum-likelihood approach by registering each new measurement to the current map and then updating the model. By leveraging the accuracy and stability of the NDT-OM representation, we are able to maintain accurate tracking performance in large scale dynamic industrial environments. We evaluated the proposed approach on two data sets from different large-scale environments and obtain accuracy and update rates an order of magnitude better than previously reported.

Although the results obtained in the industrial environment are very accurate, the evaluation in the larger city-scale data set strongly suggests that some drift can accumulate over longer trajectories. This is a clear limitation due to the maximum-likelihood SLAM approach used in this work. In order to address this limitation and provide loop closing capabilities, the natural future work direction for the proposed approach is to incorporate the obtained submaps in a hybrid metric-topological framework, similar to the work proposed by Blanco et al. [19], and implement loop closing and error relaxation procedures.

References

[1] J. Saarinen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "Normal Distribution Transforms Occupancy Maps: Application to Large-Scale Online 3D Mapping," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[2] 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.

[3] G. Pandey, J. R. McBride, and R. M. Eustice, "Ford campus vision and lidar data set," *The International Journal of Robotics Research*, vol. 30, no. 13, pp. 1543–1552, 2011.

[4] S. Thrun, W. Burgard, and D. Fox, "Probabilistic robotics," 2006.

[5] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking," in *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. IEEE, 2011, pp. 127–136.

[6] P. Besl and N. McKay, "A Method for Registration of 3D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 239–256, 1992.

[7] 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.

[8] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.

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

[10] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *Robotics, IEEE Transactions on*, vol. 24, no. 6, pp. 1365–1378, 2008.

[11] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A General Framework for Graph Optimization," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, May 2011.

[12] M. Magnusson, H. Andreasson, A. Nüchter, and A. J. Lilienthal, "Automatic Appearance-Based Loop Detection from 3D Laser Data Using the Normal Distributions Transform," *Journal of Field Robotics*, vol. 26, no. 11–12, pp. 892–914, Nov. 2009.

[13] B. Steder, M. Ruhnke, S. Grzonka, and W. Burgard, "Place recognition in 3D scans using a combination of bag of words and point feature based relative pose estimation," in *Proc. of the Int. Conf. on Intelligent Robot Systems (IROS)*. IEEE, 2011, pp. 1249–1255.

[14] G. Pandey, J. McBride, S. Savarese, and R. M. Eustice, "Visually Bootstrapped Generalized ICP," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 2660–2667.

[15] A. Tamjidi and C. Ye, "6-DOF Pose Estimation of an Autonomous Car by Visual Feature Correspondence and Tracking," *Int. Journal of Intelligent Control and Systems*, vol. 17, no. 3, pp. 94–101, 2012.

[16] P. Biber and W. Straßer, "The normal distributions transform: A new approach to laser scan matching," in *Proc. of the Int. Conf. on Intelligent Robot Systems (IROS)*, 2003, pp. 2743–2748.

[17] T. Stoyanov, M. Magnusson, and A. J. Lilienthal, "Comparative Evaluation of the Consistency of Three-Dimensional Spatial Representations used in Autonomous Robot Navigation," *Journal of Field Robotics*, vol. 30, no. 2, pp. 216–236, 2013.

[18] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A Benchmark for the Evaluation of RGB-D SLAM Systems," in *Proc. of the Int. Conf. on Intelligent Robot Systems (IROS)*, Oct. 2012.

[19] J.-L. Blanco, J. González-Jiménez, and J.-A. Fernández-Madrigal, "Subjective Local Maps for Hybrid Metric-Topological SLAM," *Robotics and Autonomous Systems*, vol. 57, no. 1, pp. 64–74, 2009.