



<http://www.diva-portal.org>

Preprint

This is the submitted version of a paper presented at *European Conference on Mobile Robotics (ECMR), Prague, Czech Republic, September 4 - 6, 2019*.

Citation for the original published paper:

Adolfsson, D., Lowry, S., Magnusson, M., Lilienthal, A., Andreasson, H. (2019)  
A Submap per Perspective: Selecting Subsets for SuPer Mapping that Afford Superior  
Localization Quality  
In: *2019 European Conference on Mobile Robots (ECMR)* IEEE  
<https://doi.org/10.1109/ECMR.2019.8870941>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-79739>

# A Submap per Perspective - Selecting Subsets for SuPer Mapping that Afford Superior Localization Quality

Daniel Adolfsson, Stephanie Lowry, Martin Magnusson, Achim Lilienthal, Henrik Andreasson

**Abstract**—This paper targets high-precision robot localization. We address a general problem for voxel-based map representations that the expressiveness of the map is fundamentally limited by the resolution since integration of measurements taken from different perspectives introduces imprecisions, and thus reduces localization accuracy. We propose SuPer maps that contain one Submap per Perspective representing a particular view of the environment. For localization, a robot then selects the submap that best explains the environment from its perspective. We propose SuPer mapping as an offline refinement step between initial SLAM and deploying autonomous robots for navigation. We evaluate the proposed method on simulated and real-world data that represent an important use case of an industrial scenario with high accuracy requirements in an repetitive environment. Our results demonstrate a significantly improved localization accuracy, up to 46% better compared to localization in global maps, and up to 25% better compared to alternative submapping approaches.

## I. INTRODUCTION

This paper is concerned with high-precision robot localization. High-precision localization requires highly precise maps, and a challenge in mapping is that the environment can appear different from different perspectives. Integrating these different appearances into a single representation introduces imprecisions, and thus reduces localization accuracy. This paper propose “SuPer”, which instead to keep multiple submaps, each representing a single more clean view of the environment. The robot localizes against the submap that best explains the environment from the current point of view, thereby reducing pose error introduced by uncertainty in the map representation.

As an example, consider an environment with two corridors separated by a wall (see Fig. 1). The environment has been discretised into a voxel grid and the surfaces inside each voxel have been estimated by a normal distribution (visualized as ellipsoids). The robot’s observations of the wall from the left-hand corridor (represented by the red ellipsoids) and the observations of the wall from the right-hand corridor (represented by the blue ellipsoids) are very precise: the ellipsoids are nearly flat and possess virtually no uncertainty along the thickness of the wall. However, if these observations are merged into a single global map (represented by the green ellipsoids) there is high uncertainty introduced by the finite thickness of the walls. A robot

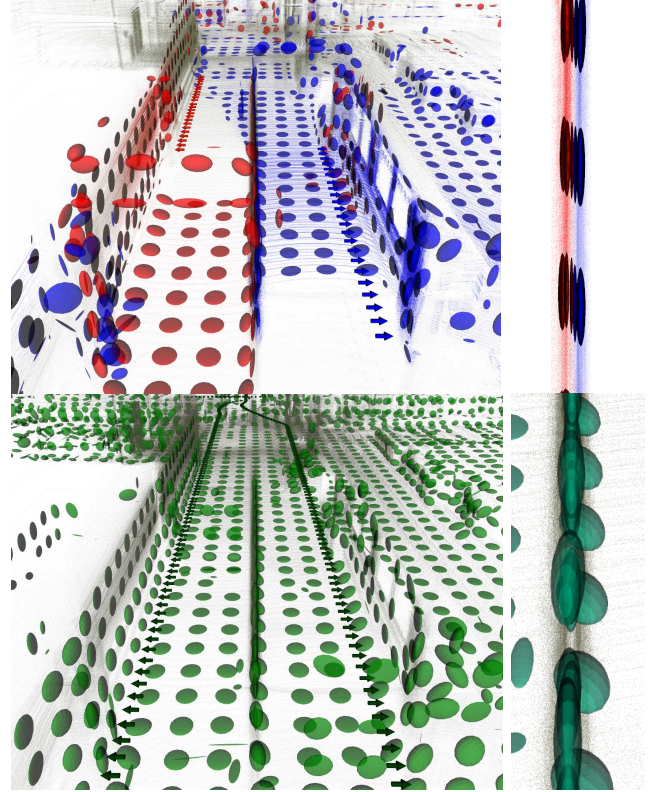


Fig. 1: Submaps can eliminate uncertainty caused by mapping with different robot perspectives. The colored arrows visualize where the corresponding map was updated from. Top row uses submaps whereas bottom row uses a single map. Top right: Observations of the wall from the left-hand corridor (red flat ellipsoids) and the right-hand corridor (blue flat ellipsoids) contain little horizontal uncertainty. Bottom right: Merging observations from both corridors (green large ellipsoids) introduces horizontal uncertainty due to the finite thickness of the wall, explained by a single voxel.

localizing against the global map would do so with less precision than what is achievable with the two submaps individually. This is a general problem for voxel based map representations where the expressiveness in a map and localization accuracy is fundamentally limited by the map resolution.

SuPer takes as input a set of scans with corresponding sensor poses which have been correctly aligned and can be obtained from an accurate SLAM or ground truth system. It performs a global partitioning of all scans into submaps by applying a spectral clustering method with a similarity score that combines both the appearance of the scans and the distance between the origins of the scans. The scans are thus clustered with other observations captured from similar perspectives so as to decrease the variation in representations within a particular submap.

The authors are with the MRO lab of the AASS research centre at Örebro University, Sweden. E-mail: Daniel.Adolfsson@oru.se

This work has received funding from the Swedish Knowledge Foundation (KKS) project “Semantic Robots” and European Union’s Horizon 2020 research and innovation programme under grant agreement No 732737 (ILIAD).

SuPer operates offline as an intermediate refinement step between mapping the environment using a SLAM system and deploying autonomous robots for navigation in repetitive industrial scenarios with high accuracy requirements. The paper evaluates our method on a simulated and two real-world industrial environments and demonstrates that our method significantly improves localization accuracy.

## II. RELATED WORK

This paper is related to the use of submaps for robot navigation and localization. Submapping was originally used to manage computational time and memory requirements [1]–[3] and to enable mapping in real-time in filtering approaches. In large-scale environments, submaps have enabled localization in maps impaired by high drift [4] and were used to close loops and reconstruct a global map [5], [6]. More recently, submaps have been used to refine a subset of previously aligned scans [7] or to accumulate and accurately register the most recent set of scans with the map [8].

This paper uses submaps with a different motivation: We aim to increase the expressiveness of the map and precision of localization within the map by allowing the map to maintain multiple independent representations of the environment. Instead of attempting to jointly explain all measurements, a single submap specializes in explaining the environment from a certain perspective. Most related to our work is research in topological mapping which uses clustering to partition space into topologically independent maps. Brunskill et al. achieve this by grouping spatially correlated locations into 2D submaps [9]. Locations are clustered using a co-visibility metric, finding which pairs of landmarks which are within line of sight. In contrast to our system, the maps are partitioned on map location rather than scan similarity. Accordingly their method does not provide the benefit of modeling surfaces as observed from different perspectives. Observed features are added to a submap only, even if their presence would be important in multiple submaps. This diminishes the usefulness of the submap for localization.

A similar approach to this paper is the work of Blanco et al. [10] who partition scans into 2D submaps using a similarity measure based on the average amount of overlapping measurement points between scans, where points are considered as overlapping if their distance is less than a threshold. The measure does not consider the geometry of the observed surface, and does not penalize measurements where the observed structures appear different.

In our previous work [11], we demonstrated that localization using submaps can improve the accuracy by up to 40% compared to localization in a global map. We used a well-established incremental submapping method that groups scans based on the Euclidean distance metric between the submap origins and the sensor pose. The robot continuously updates only the closest map. If no submap origin exist within a distance from the robot’s sensor location, a new submap is created. In this paper we refer to this method as “Incremental submapping” and it is not part of SuPer but

used as a baseline in our evaluation. This paper extends our previous work by also taking into account the appearance of the environment, and by globally grouping scans based on the similarity of the observations as well as the sensor location.

The rest of the paper is organized as follows: In Sec. III we describe our partitioning method and discuss various similarity measures and their ability to produce clusters with similar perspectives. While the framework is independent of map representation, we demonstrate how to use our method with Normal Distributions Transform (NDT-OM) representation. In Sec. IV, we evaluate our method in one simulated and two real-world industrial scenarios.

## III. METHOD

In this paper we propose a localization system using a submap per perspective (SuPer). An overview of the method is depicted in Fig. 2. As input we assume that we have a set of previously correctly aligned scans  $\mathcal{S} = \{\mathcal{P}_{1..n}\}$  together with the corresponding transformations  $\mathcal{T} = \{T_{s,1..n}^w\}$  containing the location of the sensor for each scan. Each scan is given in a fixed world frame and has been adjusted to compensate for the robot’s movement, meaning that all points in the scan have been projected to the time of the latest point measurement, as described by Zhang and Singh [8].

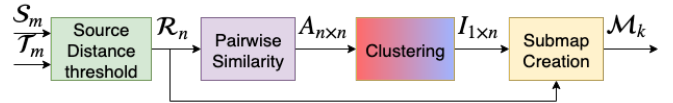


Fig. 2: Block diagram of the SuPer system. The system uses a set of scans  $\mathcal{S}_m$  with associated sensor locations  $\mathcal{T}_m$ . First, scans in the input set are removed such that the distance between subsequent scans is greater than a threshold (0.1 m). The subset of scans  $\mathcal{R}$  is then used to create the affinity matrix  $\mathbf{A}_{n \times n}$  by calculating the similarity between all combinations of scan pairs. The affinity matrix is used to partition the scan set into  $k$  clusters which then is used to build  $k$  submaps.

### A. Source distance threshold

In a first step, a subset  $\mathcal{R}_n \subseteq \mathcal{S}_m$  is obtained by imposing a minimum distance between scans. A scan  $\mathcal{P}_i$  is only added to  $\mathcal{R}$  if the Euclidean distance between the origin of the current and the previously added scan exceeds a threshold. This is simply to limit the density of point clouds when the robot is not moving. We fixed the distance threshold to 0.1 m.

### B. Pairwise similarity

For all combinations  $\mathcal{P}_i, \mathcal{P}_j \in \mathcal{R}_n$ , a similarity score is calculated and stored in the affinity matrix  $A_{i,j} = f_{sim}(\mathcal{P}_i, \mathcal{P}_j)$ . The similarity score should ideally have the following properties:

- 1) The score between scans should be high only if the scans have a common perspective of the environment. Specifically, a common perspective means that observed entities in the environment, appear geometrically similar.
- 2) The scale of the environment (the total amount of observed volume) should not impact the similarity

measure as this would tend to assign higher score between scans acquired in a large rooms.

- 3) The similarity score needs to be symmetric:  $f(\mathcal{P}_i, \mathcal{P}_j) = f(\mathcal{P}_j, \mathcal{P}_i)$ , which is a requirement for the clustering method used.

In the rest of the section we describe similarity scores based on scan similarity and distance between scan locations. The goal is to assess intuitive similarity measures to get a better understanding of the problem as a whole. Non symmetric measures were replaced by the average of the bidirectional measure  $(f_{i,j} + f_{j,i})/2$ . All affinity matrices are normalized to the range  $[0, 1]$ , with negative values rounded to 0.

1) *Sensor source distance*: The first similarity measure is based purely on the distance  $d = |\mathbf{t}_i - \mathbf{t}_j|$  between scan origins. In contrast to the incremental version, we employ global partitioning which jointly consider all scan origins. The intuition is that scans originating at similar locations typically have a similar perspective. A Gaussian kernel was used to map the distance to  $[0, 1]$ :  $f(d_{i,j}) = e^{-d_{i,j}^2/(2\sigma^2)}$ .

2) *D2D-NDT score*: Another approach is to look at the point clouds directly to assess the similarity between the scans. A measure which captures the similarity of two scans is the D2D-NDT score [12] which provides a fast way of measuring the overlap between point clouds. In contrast to the score based on overlap described in it provides a probabilistic measure between distributions that takes underlying surface structure into account. Thus, two distributions on each side of a wall would produce a significantly lower similarity score compared to two distributions on the same side. The D2D-NDT score is dependent on the number of voxels in the environment and we therefore normalize by the total amount of distribution used to calculate the score. While the D2D-NDT score takes the structure of the environment into account, it still rewards overlap between any distributions, regardless of where they were observed.

3) *Point Normals*: The Point Normals measure assign similarity of point clouds by comparing normals of nearby points. This penalizes the similarity between scans that do not observe the environment from the same direction. For example, the wall in the center in Fig. 1 will provide normals with different directions if observed from different sides. The normal of a point is obtained by computing the sample covariance of all points within a radius  $r$  around  $p_k$  and using the smallest eigenvector as the normal of  $p_k$ , aligned towards pose  $\mathcal{T}_k$ . All points with an insufficient number of neighbors to compute sample covariance are discarded. In our implementation, the point clouds are down sampled by subdividing the world into voxels, computing the centroid and average normal for each voxel containing points. The normals are then used when computing the similarity as described in algorithm 1.

4) *Point Normals & Distance*: The last similarity score uses a combination of score (1) and (3), combining appearance and the location of the scans. Practically this is done by element-wise multiplication of the two individual affinity matrices. Scans spaced with a distance  $> 3\sigma$  are set to zero.

**Data:** Points & normals  $\mathcal{P}_i, \mathcal{N}_i$  compared with  $\mathcal{P}_j, \mathcal{N}_j$

$score \leftarrow 0$

**foreach**  $p_k \in \mathcal{P}_i$  **do**

    find all points  $\mathcal{L}_j \subset \mathcal{P}_j$  within radius  $r$  of  $p_k$

$a_j \leftarrow$  average point normal in the set  $\mathcal{L}_j$

$score = score \leftarrow n_k \cdot a_j$ ,

    where  $n_k$  is the normal of  $p_k$

**end**

$score \leftarrow score/|\mathcal{P}_i|$

**Algorithm 1:** Computes the average dot product of all overlapping points.

### C. Clustering

To partition the scans (to create submaps) we apply the spectral clustering method by Ng et al. [13]. The method uses as input the affinity matrix  $\mathbf{A}_{n \times n}$  described above to partition the scans into  $k$  clusters. Spectral clustering is a relaxation of the NP-hard graph-partitioning problem which tries to find the minimal cut by clustering the eigenvectors of the affinity matrix [14] using k-means with randomly initialized cluster locations. The desired minimal cuts in this context correspond to locations where the appearance of objects as observed by the scans drastically changes.

### D. Creating local maps

After the set of scans  $\mathcal{R}_n$  has been partitioned into  $k$  clusters  $\mathcal{C}_1.. \mathcal{C}_k$ , the point clouds  $\mathcal{P}_i \in \mathcal{C}_j$  are fused into a separate local map  $\mathcal{M}_j$  per cluster. Ideally this fusing step should handle both dynamic entities and filter out spurious readings. In this work we utilize the NDT-OM [15] framework, which combines the NDT map representation with occupancy grid maps. The NDT representation has been shown to afford robust and accurate localization in industrial scenarios, even with relatively large voxels. The proposed techniques are however not specific to NDT-OM and could be used with other map representations and localization frameworks. Our proposed simply finds how to group the data in a way which will reduce imprecisions in voxel based map representations. Our code with a tutorial on how to integrate any map representation into our framework can be downloaded from our ROS package<sup>1</sup>.

## IV. EVALUATION

We evaluated SuPer in one simulated environment and two real-world industrial environments. We compared it to a global map and the commonly used incremental partitioning method based on the distance from sensor to nearest submap.

### A. Localization and map selection

The evaluation used D2D-NDT registration for scan-to-map registration as a base for our localization system. While there exist more robust localization methods using the NDT map representation such as NDT-D2D with soft constraints [16] or NDT-MCL [17], their robustness is obtained by fusing the pose estimates and odometry. As this fusion

<sup>1</sup>[https://git.vn-nt.oru.se/dlan/graph\\_map-release](https://git.vn-nt.oru.se/dlan/graph_map-release)



would bias the localization largely on odometry instead of the localization accuracy achieved by SuPer map partitioning compared to incremental submaps and global maps, we opted to not use these methods in the evaluation. Instead, the odometry was used as a first guess in the D2D-NDT scan-to-map optimization procedure.

To select the submap that best explains the environment from the robot's current pose, the locations from where the submaps were updated are stored. The most frequently updated submap close to the current estimated robot pose is then selected as described in [11].

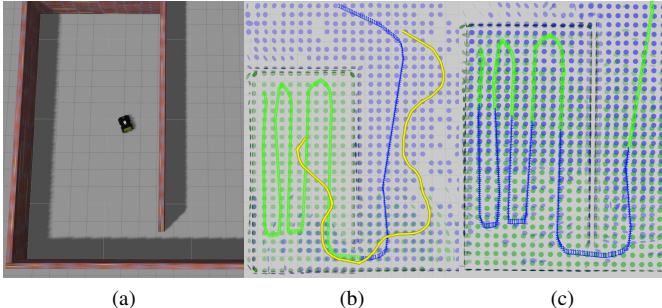


Fig. 3: (a) Simulation of 4-wheeled robot equipped with 3d laser range finder. (b) Sequences partitioned using Point Normals into  $k = 2$  clusters (shown in green and blue). The edge between the clusters is located at the entrance of the enclosed room. The performance of the map is evaluated by localization using the yellow sequence. (c) Partitioning based on distance.

### B. Simulation

To illustrate the behavior of our SuPer, a dataset in a simulated world was created (see Fig. 3). A robot equipped with a 3d laser range finder was driven twice through an environment enclosed by 4 walls. Fig. 3b depicts two sequences: The first sequence, (shown in blue/green), was used to map the environment using the ground truth pose of the sensor location. In the figure, the mapping trajectory has been clustered into  $k = 2$  clusters using “Point Normals”. Cluster membership is indicated by the color.

The second sequence (shown in yellow) was used to evaluate the performance of the submaps by localization. The majority of the scans in the mapping sequence were acquired from within the enclosed room to the left in Fig. 3a while the majority of the scans in the localization sequence was acquired from outside the enclosed area. Hence, there is potential bias in modelling the walls from within the enclosed area. Unless this bias is resolved, the robot is effectively always localizing against a map explaining the inner walls of the enclosed area. We evaluated the similarity measures “Point Normals”, “Sensor Distance” and “Point Normals & Distance” and compared with localization in a global map. The results (depicted in Fig. 5) show that the global map affords high localization error. Submaps clustered by distance performed slightly worse as the method finds a suboptimal partitioning as seen in Fig. 3c where the two submaps are assigned scans both from inside and outside the enclosed area and are thus not specialized in explaining the walls from any perspective. Using “Point Normals &

Distance”, on the other hand, produces submaps similar to the result in Fig. 3b.

### C. Real world experiments

SuPer was also evaluated on two different real-world industrial environments: a dairy production site and a warehouse. These evaluation datasets were selected as both had very high ground truth accuracy: the ground truth system has an absolute accuracy of  $< 0.02$  m with an angular error of  $< 0.1$ deg. The ground truth system was used to align the scans and the maps were built directly from these scans. This ensured the system did not require a refinement process to align the scans which might bias the results.

The data was collected using an already installed automatic guided vehicle (AGV), which was additionally equipped with a Velodyne HDL-32E lidar. The lidar data was logged along with wheel encoders (steer and drive). Ground truth pose estimates were provided by a commercial ground truth reflector system. In the dairy production dataset, the AGV was operating autonomously during the data collection, so the paths driven were very similar, whereas in the warehouse data set the vehicle was manually driven and therefore the paths demonstrated more variation. The data sequences used can be seen in Fig. 6. Mapping and localization data is depicted in red and yellow respectively.

The datasets pose challenges on different levels. The dairy production site (see Fig. 4a) is a highly structured indoor environment with a high amount of important landmarks. The environment is dynamic since it is shared with human workers and other trucks and the AGV navigates through automatic opening doors. The warehouse (see Fig. 4) is characterized by narrow aisles between the warehouse shelves, storing pallets. The view between the aisles is partially open causing view-point dependent modeling challenges when observing the aisles from different sides. The walls of the facility are far from the center of the aisles and the aisles have few unique features making localization a challenging task. During collection of datasets, the operator and additional staff were moving behind the truck, adding difficulties by introducing dynamics and blocking line-of-sight to the walls.

### D. Similarity measures in warehouse dataset

In order to demonstrate how the similarity measures behave in a real-world environment, we show a comparison of the affinity matrices in Fig. 7 from the warehouse dataset. The matrices were obtained from the sequence that originates from the top right of Fig. 6a. We aim to find a similarity measure such that scans originating from different aisles are not considered very similar, considering that the intermediate shelves and pallets are seen from different perspectives. The distance score (using  $\sigma=10$ m) is shown in Fig. 7a. Most scan pairs in the sequence are assigned a high similarity, decreasing by the distance, scans are assigned low similarity only when spaced by approximately 4 aisles. The D2D-NDT similarity seen in Fig. 7b finds similar scan pairs along the lanes and across adjacent lanes. The Point Normals score seen in Fig. 7c finds high similarity between pairs at the



Fig. 4: Environments. (a) Dairy production environment, with production area (left) and fridge storage area (middle). The production site is shared with trucks and humans. (b) Warehouse environment (right) with high-storage shelves and long corridors. The robot also navigates in a human-robot shared environment between narrow aisles where the view to important landmarks is often obscured.

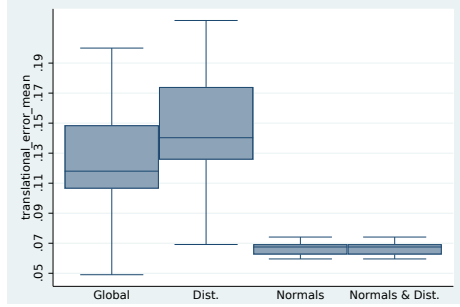


Fig. 5: Simulated dataset. Overall translation error [meters]. Partitioning using a distance based similarity measure is not sufficient for small clusters.

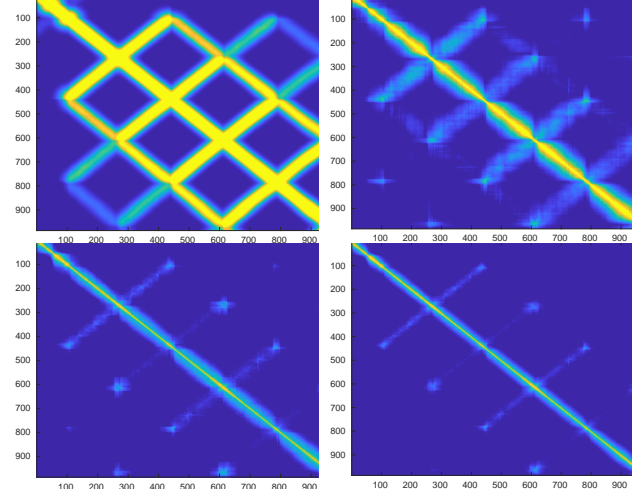
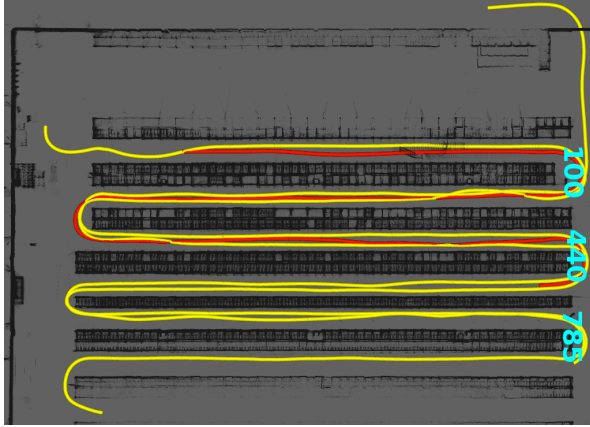
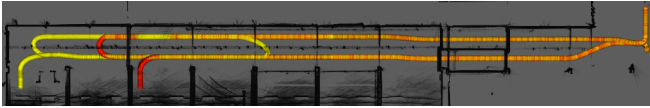


Fig. 7: Affinity matrices using various similarity functions. From top left to bottom right: (a) scan position distance using  $\sigma=10m$ , (b)  $D2D - NDT$ , (c) Point Normals, (d) scan position and Point Normals using  $r = 0.4m$ . The measure (c) and (d) are consistent with our aim as they assign low similarities between scans acquired from different sides of shelves walls.



(a) Warehouse dataset. Certain scans are labeled by their index.



(b) Dairy production site.

Fig. 6: Overview of the trajectories for the data sets. Red and yellow depict the paths used for building the map and localizing in the map respectively.

larger open areas locations e.g. between the scans: (100, 440, 758), highlighted in Fig. 6b, but lower scores along the narrow aisles, this is natural as the perspective from which the sensor observes objects changes more rapidly when objects are close. Fig. 7d combines the Point Normals (Fig. 7c) and the Distance (Fig. 7a) to produce a smoother score than the Point Normal solely. We observed that clustering based on this combination can lead to cleaner cuts between clusters.

### E. Dairy production site localization

The overall localization error from the dairy production site can be seen in Fig. 8a. It can be seen that any submap approach improves the accuracy significantly compared to a global map. In Fig. 9 the error is plotted wrt. to the number of clusters. Between 8-20 clusters, similarity measures based on Point Normals slightly improves localization compared to the distance similarity and incremental submap. At the point of 20 clusters, the size of the clusters is small enough that the similarity measure itself does not seem to significantly affect the results. SuPer outperforms incremental submap regardless of similarity measure.

### F. Warehouse localization

The results from the warehouse experiments can be seen in Fig. 8b and Fig. 10. Regardless of similarity, SuPer outperformed the global map and incremental submap. Clustering based on point normals was superior to clustering on distance up to 20 clusters. In that case, the clusters are large enough to require a spatially logical partitioning, otherwise important landmarks are explained from different perspectives. For a large number of clusters ( $k>50$ ) the distance similarity provides the best results, The measure produces submaps

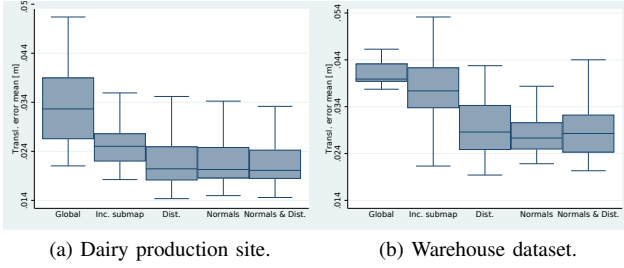


Fig. 8: Overall translation using a global map (Global), incremental distance submapping (Inc. submap) and SuPer with similarity measure (1), (3) and (4). In the dairy set, incremental submaps reduces error by 24% compared to a global map. SuPer reduces the error by 33% compared to a global map, regardless of similarity measure. SuPer reduces the error by 12% compared to Incremental submap. In the warehouse dataset, Inc.submap reduced the error by 9% compared to a global map. SuPer reduced error by 29% using distance similarity and 32% using Point Normals, compared to a global map. Point Normals reduced the error by 4% compared to Distance. In general, SuPer reduced the error by 25% compared to Incremental submap.

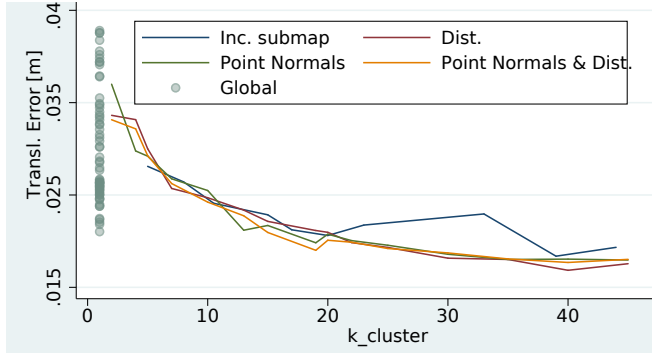


Fig. 9: Dairy production dataset. Translational error with respect to number of clusters. SuPer with various similarity measures is compared to incremental submap and global map.

of tightly spaced sensor origins which itself narrows the perspective of the explained environment.

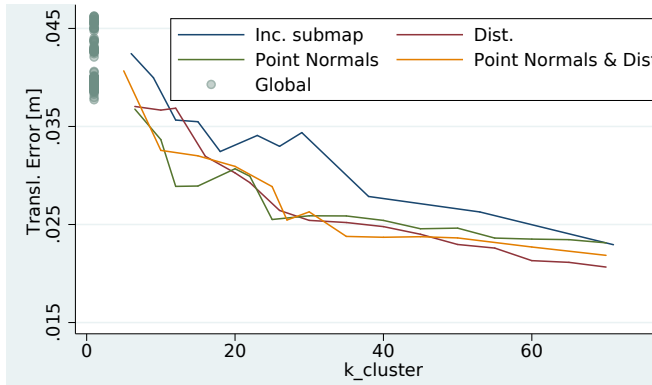


Fig. 10: Warehouse dataset. Translational error with respect to number of clusters. SuPer outperform the global and incremental approaches regardless of the number of clusters. For a small amount of clusters ( $k=10-20$ ), Point Normals reduced the error by 8% compared to Distance similarity. For a large amount of clusters ( $k=50-60$ ), the following error reduction wrt. a global map was found: Inc submap 37%, Distance 46%, Point Normals 46%, Point Normals & Distance 44%.

## V. CONCLUSIONS

Creating accurate and precise voxel-based maps for robot localization is a hard task as the expressiveness is fundamentally limited by the map resolution. We found that

creating submaps by globally cluster scans could address this problem and produce accurate maps for localization. We found a 46% improvement compared to a global map and a 25% improvement compared to a previous submapping approach. We also found that when partitioning scans into a small number of submaps, using a similarity measure based on point normals can lead to better results compared to using a measure based on scan source location. However, partitioning scans into a large number of submaps was best done using a similarity based on sensor origin distance. In the future we will evaluate the framework with different map representations. We will also investigate sharing data between submaps e.g. by jointly filter dynamics or allowing scans to belong to more than one submap.

## REFERENCES

- [1] J. L. Blanco, J. A. Fernández-Madriral, and J. González, "Toward a unified bayesian approach to hybrid metric-topological slam," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 259–270, April 2008.
- [2] M. Bosse, P. Newman, J. Leonard, M. Soika, W. Feiten, and S. Teller, "An atlas framework for scalable mapping," in *ICRA*, vol. 2, Sept 2003, pp. 1899–1906 vol.2.
- [3] M. Bosse, P. Newman, J. Leonard, and S. Teller, "Simultaneous localization and map building in large-scale cyclic environments using the atlas framework," *The International Journal of Robotics Research*, vol. 23, no. 12, pp. 1113–1139, 2004.
- [4] J. Marshall, T. Barfoot, and J. Larsson, "Autonomous underground tramming for center-articulated vehicles," *Journal of Field Robotics*, vol. 25, no. 67, pp. 400–421, 2008.
- [5] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2d lidar slam," in *ICRA*, May 2016, pp. 1271–1278.
- [6] J. Blanco, J. Fernández-Madriral, and J. González, "A new approach for large-scale localization and mapping: Hybrid metric-topological slam," in *ICRA*, April 2007, pp. 2061–2067.
- [7] D. Droschel and S. Behnke, "Efficient continuous-time slam for 3d lidar-based online mapping," in *ICRA*, May 2018, pp. 1–9.
- [8] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," in *Robotics: Science and Systems Conference*, July 2014.
- [9] E. Brunskill, T. Kollar, and N. Roy, "Topological mapping using spectral clustering and classification," in *IROS*, Oct 2007, p. 3491.
- [10] J. L. Blanco, J. González, and J. A. Fernández-Madriral, "Consistent observation grouping for generating metric-topological maps that improves robot localization," in *ICRA*, May 2006, pp. 818–823.
- [11] D. Adolfsson, S. Lowry, and H. Andreasson, "Improving localisation accuracy using submaps in warehouses," in *(IROS), Workshop on Robotics for Logistics in Warehouses and Environments Shared with Humans*, 2018.
- [12] 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.
- [13] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in neural information processing systems*, 2002, pp. 849–856.
- [14] F. R. K. Chung, *Spectral Graph Theory*, ser. CBMS Regional Conference Series in Mathematics. American Mathematical Society, 1997, vol. 92, no. 92.
- [15] J. P. Saarienen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "3d normal distributions transform occupancy maps: An efficient representation for mapping in dynamic environments," *The International Journal of Robotics Research*, 2013.
- [16] H. Andreasson, D. Adolfsson, T. Stoyanov, M. Magnusson, and A. Lilienthal, "Incorporating ego-motion uncertainty estimates in range data registration," 09 2017, pp. 1389–1395.
- [17] J. Saarienen, H. Andreasson, T. Stoyanov, and A. Lilienthal, "Normal distributions transform monte-carlo localization (ndt-mcl)," 11 2013, pp. 382–389.