Semantic-assisted 3D Normal Distributions Transform for scan registration in environments with limited structure

Anestis Zaganidis*   Martin Magnusson†   Tom Duckett*   Grzegorz Cielniak*

Abstract—Point cloud registration is a core problem of many robotic applications, including simultaneous localization and mapping. The *Normal Distributions Transform* (NDT) is a method that fits a number of Gaussian distributions to the data points, and then uses this transform as an approximation of the real data, registering a relatively small number of distributions as opposed to the full point cloud. This approach contributes to NDT's registration robustness and speed but leaves room for improvement in environments of limited structure.

To address this limitation we propose a method for the introduction of semantic information extracted from the point clouds into the registration process. The paper presents a large scale experimental evaluation of the algorithm against NDT on two publicly available benchmark data sets. For the purpose of this test a measure of smoothness is used for the semantic partitioning of the point clouds. The results indicate that the proposed method improves the accuracy, robustness and speed of NDT registration, especially in unstructured environments, making NDT suitable for a wider range of applications.

I. INTRODUCTION

Point cloud registration is the estimation of the pose transform between two three-dimensional scans of the environment, originating from sensors such as a lidar or an RGB-D camera. By minimizing a distance function between two point clouds, the relative movement of the sensor is obtained.

Point cloud registration is a crucial part of many applications of robotics including tracking of dynamic objects, 3D modeling and position estimation. Simultaneous localization and mapping (SLAM) is the problem of constructing a map of an unknown environment while tracking the robot's location, using the constructed map to bound the odometry error. Sensors providing global positioning, odometry, visual and depth information are typically integrated. With lidar scanners and RGB-D cameras becoming cost-competitive, point cloud registration has become a major component of 3D SLAM systems.

Normal Distributions Transform (NDT) registration [1], later extended to 3D in [2] is a well-established registration method, which has been used for map creation, loop closure and complete SLAM. By transforming the data points into a set of normal distributions representing the local structure of points, and registering distributions instead of matching
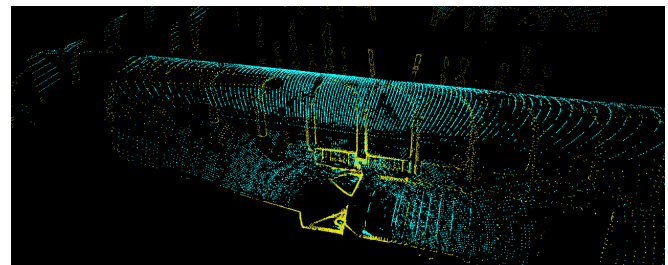
*Lincoln Centre for Autonomous Systems (LCAS), University of Lincoln, UK.

†Centre for Applied Autonomous Sensor Systems (AASS), Örebro University, Sweden.
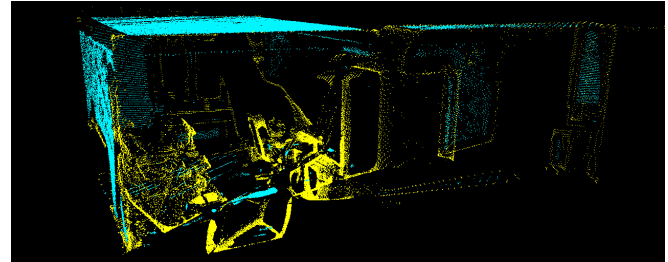
(a)

(b)

Fig. 1. *Challenging data sets* partitioned point clouds. Yellow points high smoothness, blue points low. (a) *Apartment*, (b) *ETH*.

individual scan points, NDT reduces dimensionality and increases registration speed and robustness when compared to point-to-point registration methods [3]. Due to the data reduction NDT is an efficient representation of the environment, desirable in mapping applications. However the smoothing of local features and the loss of detail results in degradation of performance in environments with limited structure.

We propose a method that assists registration using semantic labels for each point. The point cloud is partitioned into disjoint sets according to the points' labels and the Normal Distributions Transform is estimated for each partition separately. As a proof of concept, in the experiments we use labels that describe the local geometry, applying a smoothness measure to partition the point cloud into the categories of *edge* and *plane* points. Any number of criteria can potentially be used to segment the cloud, resulting in a proportional number of transforms. Instead of matching all Gaussians, correspondences are searched only on the NDT representation constructed from the same type of partition. As a further benefit of our approach, we are able to discard the *unassigned* points with no semantic label (75% of the points in our experiments), resulting in faster registration if a computationally efficient method for semantic labelling is

applied, or the semantic information is readily available from another source.

The proposed algorithm is tested and compared with NDT on two data sets, *Challenging data sets* as introduced by Pomerleau et al. [4] and the *KITTI* odometry benchmark suite, introduced by Geiger et al. [5]. The smoothness measure used for labeling is similar to [6]. The algorithm is not used for full SLAM in the experiments, but tested only on its registration performance, thus no loop closure or global map construction is applied.

The experiments validate that the algorithm improves the registration accuracy and robustness when the overlap of the scans is large, which is the general case in real world localization applications, and when there is limited structure.

II. 3D-NDT

NDT is a method for the representation of the environment and registration of data, proposed by Biber and Straßer [1] for 2D scan registration, and later extended for the registration of three-dimensional data by Magnusson et al. [2]. As opposed to other methods that use either full point cloud or feature points to perform the registration, 3D-NDT assumes a local Gaussian distribution of points and uses the probability density function as a representation. Space is segmented into voxels, and for each voxel a Gaussian model is fitted to the data. In the point-to-distribution (P2D) variant, the registration of a new scan then becomes a problem of fitting its points to the distribution, and is solved as a least-squares problem. Several proposals have been made regarding the segmentation of space [2], the representation of orientation [7], its use in loop closure [8], and the use of features instead of the full point cloud [9]. Important extensions include the distribution to distribution registration (D2D-NDT) presented in [10], the fusion of NDT with occupancy maps [11]–[14] its use with Monte-Carlo localization [15], [16], and the integration of color information in the registration [17].

The distribution-to-distribution variant of NDT [10], [3] consists of the following steps to register a point cloud $\mathcal{M}$ to a point cloud $\mathcal{F}$. First, the space is discretized into voxels (for other structures see [2], [14]). Let $\mathcal{S}_i$ be the set of points $\mathbf{v}$ of $\mathcal{F}$ in voxel i. A Gaussian distribution of points is assumed for every voxel, resulting in the sets of distributions $\mathcal{G}_F, \mathcal{G}_M$. The mean vector $\boldsymbol{\mu}_i$ and the covariance matrix $\mathbf{C}_i$ for each distribution are estimated according to

$$\boldsymbol{\mu}_i = \frac{1}{|\mathcal{S}_i|} \sum_{\mathbf{v} \in \mathcal{S}_i} \mathbf{v}, \tag{1}$$

$$\mathbf{C}_i = \frac{1}{|\mathcal{S}_i| - 1} \sum_{\mathbf{v} \in \mathcal{S}_i} (\mathbf{v} - \boldsymbol{\mu}_i)(\mathbf{v} - \boldsymbol{\mu}_i)^T. \tag{2}$$

The probability density function at voxel i is then given by

$$p_i(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^3 |\mathbf{C}_i|}} exp\left(-\frac{(\mathbf{x} - \boldsymbol{\mu}_i)^T \mathbf{C}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)}{2}\right). \tag{3}$$

Let $\mathbf{T}$ be the 6-DOF transformation matrix from $\mathcal{M}$ to $\mathcal{F}$, with $\mathbf{R}$ and $\mathbf{t}$ the rotation and translation components

respectively. The distance between two transforms i, j is defined as

$$dist(i, j) = -d_1 exp\left(-\frac{d_2}{2} \boldsymbol{\mu}_{ij}^T \left(\mathbf{R}^T \mathbf{C}_i \mathbf{R} + \mathbf{C}_j\right)^{-1} \boldsymbol{\mu}_{ij}\right) \tag{4}$$

and the transformation from $\mathcal{M}$ to $\mathcal{F}$ is found by minimizing the function

$$f(\mathbf{T}) = \sum_{i=1, j=1}^{|\mathcal{G}_M|, |\mathcal{G}_F|} dist(i, j), \tag{5}$$

where $\boldsymbol{\mu}_{ij} = \mathbf{R}\boldsymbol{\mu}_i + \mathbf{t} - \boldsymbol{\mu}_j$ and d_1, d_2 are regularization factors. Newton optimization is used to obtain the transformation $\mathbf{T}$ with analytically computed derivatives.

III. SEMANTIC-ASSISTED NDT

We propose a general method for semantic-assisted NDT registration (SE-NDT), where the point cloud is partitioned according to per-point semantic labels, and sets of Normal Distributions Transforms are estimated and registered for each partition independently. Inspired by [6], a method that uses partitioning based on local geometric structure, and [18], [19], which use semantic labels to aid Iterative Closest Point [20] (ICP), we attempt to apply the same principle on D2D-NDT registration and generalize it as a measure-agnostic approach.

Instead of computing a single NDT using a full point cloud, the point cloud is partitioned and multiple NDTs are constructed. To partition the point cloud, a measure is used to calculate a value for each point, which ideally has high consistency between scans, i.e. produces approximately the same value for the same region, independent of the viewpoint. Two NDTs are created for each measure used, for the lower and upper tails of the distribution of their values, corresponding to *edge* and *plane* points respectively in our approach. We define an optional threshold for rejecting the points with values within an interval around the median of the distribution. The effect of the threshold depends on the quality and the number of measures used. It is required to preserve sufficient points for the creation of the NDTs, and at the same time reject enough points to make the labels distinct. By considering each point on the NDT where its value is more distant to the median of the distribution, the probability of re-observing a point of the same type in the next scan is higher. Each point is assigned to exactly one NDT according to its measure value. To register a point cloud the same procedure is followed to generate a set of NDTs. The two sets are then registered, by finding correspondences only among NDTs of the same type.

More formally, let $\mathcal{P}$ represent the set of points in a point cloud and $\mathbf{P}$ its matrix representation. Suppose we have n measures with functions $\mathbf{f}_n : \mathbf{R}^{|\mathcal{P}| \times 3} \to \mathbf{R}^{|\mathcal{P}|}$ and the point cloud partitions $\mathcal{P}_n$. Let $F_n(c)$ be the cumulative distribution function of the measure, $\bar{F}_n(c)$ the complementary distribution and r the rejection threshold. The points are assigned to $2n$ clouds according to:

$$G(\mathbf{P}, \mathbf{v}) = \min_{\forall n}\left(\min\left(F_n(f_n(\mathbf{P})_{\mathbf{v}}), \bar{F}_n(f_n(\mathbf{P})_{\mathbf{v}})\right)\right), \tag{6}$$

4065

$$\mathcal{P}_{n,1} = \left\{ \mathbf{v} \in \mathcal{P} : \bar{F}_n \left(f_n(\mathbf{P})_{\mathbf{v}} \right) \equiv G(\mathbf{P}, \mathbf{v}) < r \right\}, \quad (7)$$

$$\mathcal{P}_{n,2} = \left\{ \mathbf{v} \in \mathcal{P} : F_n \left(f_n(\mathbf{P})_{\mathbf{v}} \right) \equiv G(\mathbf{P}, \mathbf{v}) < r \right\}. \quad (8)$$

Examples of the partitioning of point clouds can be seen in Fig. 1. The NDT sets are then estimated as in Section II, resulting in $G_{\mathcal{M}}^n$ and $G_{\mathcal{F}}^n$, and the function to optimize becomes

$$f(\mathbf{T}) = \sum_{\forall n} \sum_{i=1, j=1}^{|\mathcal{G}_M^n|, |\mathcal{G}_F^n|} dist(i, j). \quad (9)$$

For the purpose of this paper a smoothness measure was used to label the points, as presented in [6]. In their work the authors assumed that the points are partially ordered, i.e. the scan-time of each point is known. As this is not true for *Challenging data sets*, we use a modified measure to search for nearest neighbors in a defined radius. Another difference is that the original implementation was designed for a rotating laser scanner and only considered points from the same scan (a line of points), while here we consider points in three dimensions, belonging to a surface. The smoothness measure in our work is defined as

$$c_{\mathbf{v}} = \frac{1}{|\mathcal{K}_{\mathbf{v}}| \cdot \|\mathbf{v}\|} \| \sum_{\mathbf{u} \in \mathcal{K}_{\mathbf{v}}} (\mathbf{v} - \mathbf{u}) \|, \quad (10)$$

where $\mathcal{K}_{\mathbf{v}}$ is the set of neighbors of point $\mathbf{v}$. The assumption is that points of the same region will have approximately the same smoothness value across scans.

IV. DATA SETS

Two data sets were used for the evaluation of the algorithm; the *challenging data sets for point cloud registration algorithms* [4], and the autonomous driving data set *KITTI* [5]. The distribution-to-distribution variants of NDT and the proposed algorithm were used for comparison.

A. Challenging data sets for point cloud registration algorithms

The data set consists of six sequences of laser scans collected in different environments, and was previously used by [21] and [3] to evaluate the performance of registration algorithms. The scans were acquired with a Hokuyo UTM-30LX mounted on a tilting device. Six environments are included;

- Apartment, 45 scans in a single floor with 5 rooms.
- Stairs, 31 scans of a small staircase transitioning from indoor to outdoor.
- ETH Hauptgebaude, 36 scans of a large hallway with pillars and arches.
- Gazebo, 32 scans of wine trees covering a gazebo in a public park.
- Wood, 37 scans of dense vegetation around a small paved way.
- Plain, 31 scans of a small concave basin with alpine vegetations.

For each of these environments 6720 combinations of reference frame, target frame and initial transformation estimation are provided. The initial pose transforms are categorized as *Easy* (standard deviation $0.1\,\text{m}$ and 10°), *Medium* (standard deviation $0.5\,\text{m}$ and 20°) and *Hard* (standard deviation $1.0\,\text{m}$ and 45°). The overlap ratio is also provided. The accuracy of the estimated transform is measured based on the translation error and rotation error, as suggested in [21].

B. KITTI odometry benchmark

The data set contains 22 sequences of lidar and stereo camera data. The data are collected outdoors by a moving vehicle, moving at up to $88\,\text{km/h}$. For 11 of the sequences ground truth is provided. The scans originate from a Velodyne HDL-64, so unlike the *Challenging data sets*, *KITTI* point clouds consist of parallel lines. There is *Large* overlap between consecutive scans, and there is no categorization of the difficulty of the initial pose transform. The accuracy of the algorithms is estimated by the total translation and rotation error as in [5] and for all sub sequences of lengths $100\,\text{m}$ – $800\,\text{m}$, following the benchmark guideline. We evaluate the registration performance of the algorithms, without map construction and loop closure detection, on 11 sequences of the rectified data set with the available ground truth.

C. Parameters

1) Challenging data sets: For NDT the results obtained by [3] were used as reference. The registration parameters used were 5 iterations, with grid resolutions $\{1, 2, 1, 0.5\}$ m, and regularization factors $d_1 = 1$, $d_2 = 0.05$. Instead of using all distributions in the map, as in (5), for each distribution only the 8 closest distributions were considered.

SE-NDT was tested using the same parameters as NDT with maximum number of iterations 100. The point rejection threshold, r, was set to 0.125 (using 25% of all points). For the estimation of the smoothness measure, neighbors within a radius of $0.2\,\text{m}$ were considered for each point.

2) KITTI: We optimized the parameters for both algorithms on a validation set for the tests with *KITTI*, instead of using default parameters as above. An initial estimation of the transform was used for both methods, assuming constant linear and angular velocity, which improved registration in sequences where the vehicle was moving with high speed.

Using the same parameters for NDT as in IV-C.1 resulted in considerably higher registration error, not used in the comparison but presented in Table II. The optimized parameters were set as follows; the resolution of the NDT grid was set to $0.4\,\text{m}$, with maximum number of iterations 15. The registration was performed with two resolutions, the first in 2D with resolution $1.2\,\text{m}$, which speeded up the registration and increased the accuracy (as noted in [11]). The 27 closest distributions were considered.

For SE-NDT, to speed up the registration an approximation of the smoothness measure was used, taking into account only the 10 closest neighbors instead of all neighbors within a radius. The resolution was set to $1.7\,\text{m}$ for planar points and $0.4\,\text{m}$ for edge points, which favors points with high smoothness and improves registration when limited structures are present, such as the highway of Sequence 01. The maximum number of iterations was set to 50. In the validation set, due
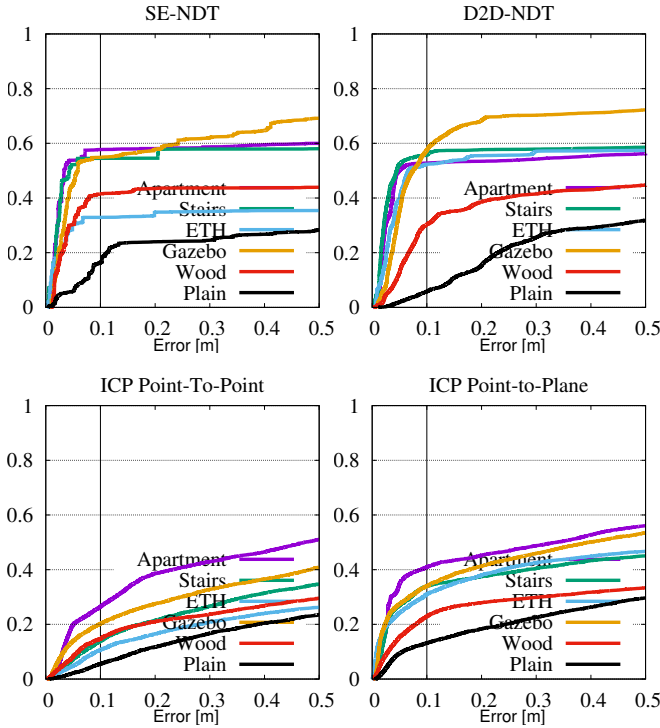
Fig. 2. Cumulative probabilities of translation errors for all poses and overlaps.

Fig. 3. Cumulative probabilities of translation errors for *Large* overlaps over 75%, for all poses.

to the use of the approximate measure, the upper 5% of the distribution consisted of noise and was discarded. The measure function was less consistent on planar points, so the rejection thresholds were set as $0.00 - 0.05$ for planar and $0.05 - 0.25$ for edge points. The remaining parameters were set the same as for NDT.

V. RESULTS

A. Challenging data sets

The *Challenging data sets* have several categories of difficulty, being divided into *Easy* and *Hard* poses, *Low* and *Large* overlap, and structured and unstructured environments. Closer to a realistic scenario of a moving robot is the subset of *Easy* poses with *Large* overlap, due to the high frame rate of the sensors. In this subset, the proposed algorithm performs significantly better than NDT, especially in environments with limited structure.

TABLE I
NULL HYPOTHESIS TESTING FOR CHALLENGING DATA SETS.

	Apartment	Stairs	ETH	Gazebo	Wood	Plain	All
All	•				•	•	
Easy					•	•	•
High Ov.	•	•	•	•			
Easy&High Ov.	•				•	•	•

• indicates cases where SE-NDT significantly outperforms D2D-NDT

Table I presents the results of a Z-test. At the 95% confidence interval, SE-NDT is significantly more robust than NDT for *Easy* poses and *Large* overlap, but over all poses and overlaps the null hypothesis cannot be rejected.
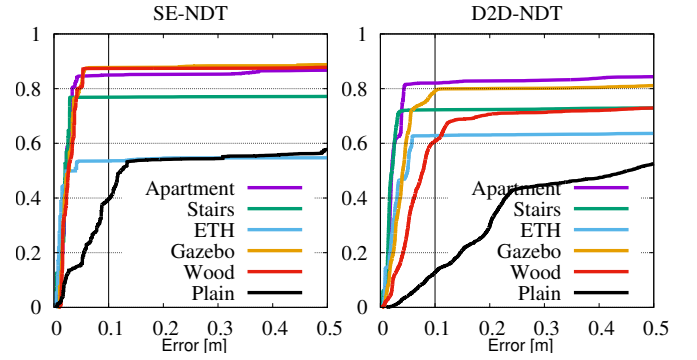
Testing SE-NDT with the same number of iterations as standard NDT resulted in comparable results that were not included due to space limitation.

To estimate the Z statistic only registrations that are accurate, i.e. have translation error less than $0.1\,\mathrm{m}$ and rotation error less than 2.5^o are considered as successful. From the registrations that succeeded, SE-NDT had an average translation error $0.029\,\mathrm{m}$ while NDT had $0.036\,\mathrm{m}$. Rotation error was 0.50^o for SE-NDT and 0.49^o for NDT.

The cumulative probabilities of registration errors for different categories of difficulty are shown in Figs. 2–6. Curves that cross the marked vertical line higher are better, signifying higher registration accuracy. They can be interpreted as the probability (vertical axis) that the registration error is less than the corresponding value on the horizontal axis. Vertical lines at $0.1\,\mathrm{m}$ of Fig. 2–5 and at 2.5^o of Fig. 6 mark the upper limit of translation and rotation error, respectively, where a registration is considered successful. To ease the comparison of accuracy, we quantify those results using the quantiles of the cumulative distribution of translation errors, so that Q50 is the median of the distribution.

In Fig. 2 additional results are included for point-to-point ICP and point-to-plane ICP [22]. We note that both SE-NDT and D2D-NDT outperform the ICP variants, with the exception of the *Plain* sequence, where point-to-plane ICP is more accurate than D2D-NDT, but not SE-NDT. For a detailed comparison of ICP vs NDT, please refer to [3].

The registration of sequences *Wood* and *Plain* had the largest improvement in accuracy with SE-NDT, where the lack of man-made structures affects the robustness of D2D-NDT. For *Wood*, NDT was accurate up to Q30, Q60, Q44 of translation error, and SE-NDT up to Q41, Q87, Q50 for All, *Large* overlaps and *Easy* poses, respectively. For *Plain*, NDT was accurate up to Q5, Q12, Q8 and SE-NDT up to Q16, Q39, Q17 for All, *Large* overlaps and *Easy* poses.

Apartment and *Stairs* are structured environments, and *Gazebo* contains man-made structures, so D2D-NDT is able to operate. Using SE-NDT the improvement in accuracy is evident when the overlap is *Large*.

ETH is the only sequence where SE-NDT fails. The reason for this failure is that NDT relies on specific landmarks (i.e. columns, see Fig. 1b), points that are removed by the
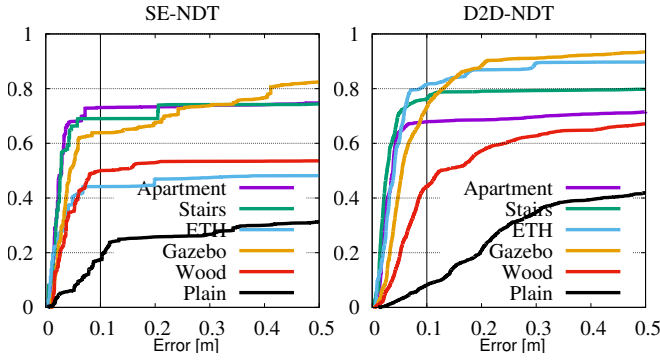
Fig. 4. Cumulative probabilities of translation errors for *Easy* poses and all overlaps.
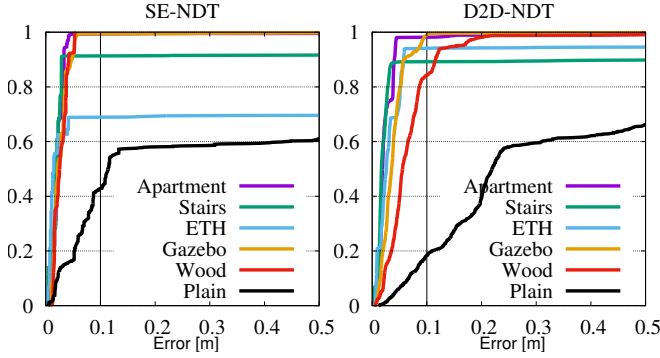

Fig. 5. Cumulative probabilities of translation errors for *Easy* poses and *Large* overlaps.
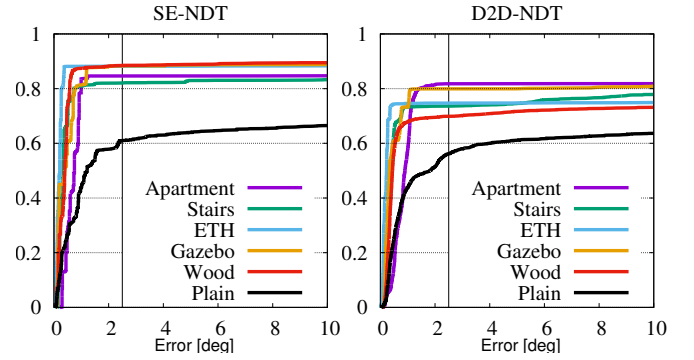

Fig. 6. Cumulative probabilities of rotation errors for *Large* overlaps over 75%, for all poses.

TABLE II

KITTI TRANSLATION AND ROTATION ERROR

Method	Translation (m/m)		Rotation $(10^{-4}deg/m)$	
	Mean	Variance	Mean	Variance
SE-NDT	0.0260	0.0003	2	0.0002
D2D-NDT	0.0409	0.0087	2	0.0003
D2D-NDT (IV-C.1)	0.1256	0.0312	6	0.0149

TABLE III

NULL HYPOTHESIS TESTING FOR KITTI.

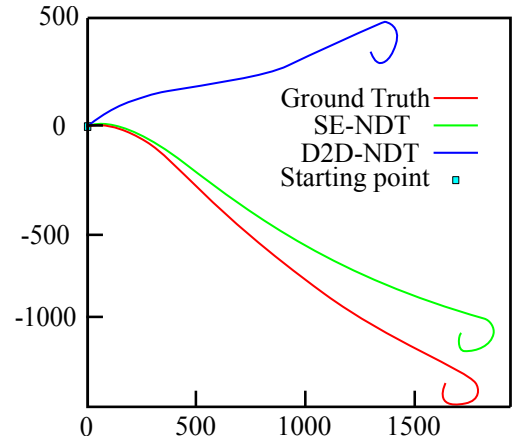|		00	01	02	03	04	05	06	07	08	09	10	All
T	SE-NDT	●	●	●	●				●		●	●	●
	D2D-NDT						●						
R	SE-NDT		●		●				●		●	●	
	D2D-NDT	●				●	●	●		●			●

Significantly better:●. Translation: T. Rotation: R.

Fig. 7. KITTI, Sequence 01 estimated path.

preprocessing step of SE-NDT. The smoothness measure can not capture cylindrical surface characteristics, while at the same time it misclassifies boundaries of cylindrical objects. As these boundaries are viewpoint-dependent the algorithm's robustness is affected significantly.

Regarding execution time, the preprocessing step for estimating smoothness is not efficient, using up to 30 seconds (Intel Core i7, 6 threads) for scenes where the points are concentrated close to the sensor (e.g. *Apartment* and *Wood*), due to the iteration over all points and the density of points in close ranges. After the estimation of smoothness, the registration step is faster than NDT. The average time per registration for all sequences and all difficulty categories was 0.37 seconds for SE-NDT and 0.94 seconds for NDT.

B. KITTI

Our method performs significantly better than NDT due to the high overlap between scans and the close initial estimation of the transform (analogous to *Large* overlap and *Easy* poses of *Challenging data sets*). Furthermore, the data set consists mainly of environments with limited structure. For instance, in the highway of sequence 01 (Fig. 7), due to the lack of consistent objects in a rapidly changing environment, NDT registers the clouds according to other vehicles, moving with the same speed. This is demonstrated by the peak translation error after 80 km/h (see Fig. 8). However, NDT's rotation error declines when the speed is over 65 km/h (Fig. 9) due to the small variation of orientation while moving along highways.

The running time of the algorithms is evaluated by the CPU time used for the processing of each frame. Preprocessing and registration are both included in this test. The proposed method needed 2.82 seconds for each registration for the most demanding data set, while NDT needed 4.15 seconds (7.26 for the most demanding data set). The observed rate, due to parallelization, was 0.98 Hz and 0.27 Hz for the proposed method and for NDT respectively.
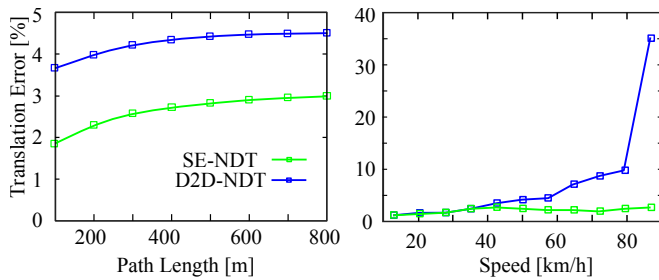
Fig. 8. KITTI, Translation error versus (a) path length and (b) speed for all sequences.
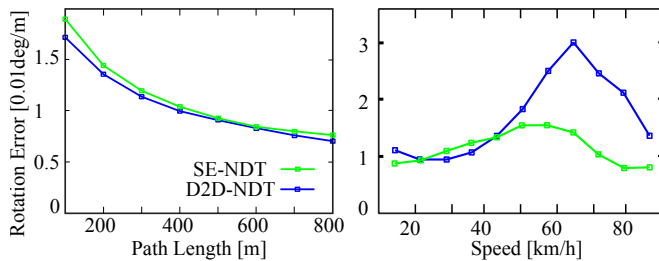

Fig. 9. KITTI, Translation error versus (a) path length and (b) speed for all sequences.

VI. CONCLUSION

This paper presented a method for enhancing NDT registration of 3D point clouds with semantic information. To achieve that we proposed the partitioning of the point cloud according to per-point semantic labels, and the independent registration of the partitions. In the experiments we evaluated the performance of the algorithm against D2D-NDT on two publicly available data sets, a general data set for point cloud registration algorithms, and an autonomous driving benchmark. The experiments validated the improvement of registration in environments with limited structure and when the overlap between the clouds is large, making it suitable for outdoor mapping and localization applications.

When the overlap is low, only few distribution correspondences are found as many of the voxels do not contain a distribution due to down-sampling and filtering of the point cloud. This negatively affects the performance on the registration of point clouds with low overlap. More tests are needed to investigate whether combining several different measures to construct more than two semantic regions can overcome this limitation and further improve registration. An efficiently computed measure to partition the point cloud is necessary to improve the performance of SE-NDT in runtime. It was also observed that the measure used in this paper performs poorly in environments where cylindrical structures are dominant, indicating the importance of viewpoint-invariant measures.

Future work will include partitioning of the point cloud into more categories such as cylinders, and investigation of other measures that capture more useful semantic information. The susceptibility of the method to the rejection threshold is to be investigated. It will also include testing if point-to-distribution NDT can benefit from the same method, and evaluation of the algorithm on other data sets.

REFERENCES

[1] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst.*, vol. 3, Oct 2003, pp. 2743–2748.

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

[3] M. Magnusson, N. Vaskevicius, T. Stoyanov, K. Pathak, and A. Birk, "Beyond points: Evaluating recent 3D scan-matching algorithms," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2015, pp. 3631–3637.

[4] F. Pomerleau, M. Liu, F. Colas, and R. Siegwart, "Challenging Data Sets for Point Cloud Registration Algorithms," *Int. J. Robotics Research*, vol. 31, no. 14, pp. 1705–1711, 2012.

[5] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012, pp. 3354–3361.

[6] J. Zhang and S. Singh, "LOAM : Lidar Odometry and Mapping in Real-time," *Robotics: Sci. and Syst.*, 2014.

[7] M. Magnusson, "The Three-Dimensional Normal-Distributions Transform," Ph.D. dissertation, Örebro University, 2009.

[8] M. Magnusson, H. Andreasson, A. Nüchter, and A. J. Lilienthal, "Automatic appearance-based loop detection from three-dimensional laser data using the normal distributions transform," *J. Field Robotics*, vol. 26, no. 11-12, pp. 892–914, Nov 2009.

[9] T. Schmiedel, E. Einhorn, and H. M. Gross, "IRON: A fast interest point descriptor for robust NDT-map matching and its application to robot localization," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst.*, Sept 2015, pp. 3144–3151.

[10] T. Stoyanov, M. Magnusson, and A. J. Lilienthal, "Point set registration through minimization of the L2 distance between 3D-NDT models," in *Proc. IEEE Int. Conf. Robotics and Automation*, Minnesota, USA, 2012, pp. 5196–5201.

[11] J. Saarinen, T. Stoyanov, H. Andreasson, and A. J. Lilienthal, "3D normal distributions transform occupancy maps: an efficient representation for mapping in dynamic environments," *Int. J. Robotics Research*, pp. 1–18, 2013.

[12] J. Saarinen, T. Stoyanov, H. Andreasson, and A. Lilienthal, "Fast 3D mapping in highly dynamic environments using normal distributions transform occupancy maps," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst.*, vol. 32, no. 14, Nov 2013, pp. 4694–4701.

[13] J. Saarinen, H. Andreasson, T. Stoyanov, J. Ala-Luhtala, and A. J. Lilienthal, "Normal Distributions Transform Occupancy Maps: Application to large-scale online 3D mapping," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2013, pp. 2233–2238.

[14] E. Einhorn and H. M. Gross, "Generic NDT mapping in dynamic environments and its application for lifelong SLAM," *Robotics and Autonomous Syst.*, vol. 69, pp. 28–39, 2015.

[15] J. Saarinen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "Normal distributions transform Monte-Carlo localization (NDT-MCL)," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Syst.*, 2013, pp. 382–389.

[16] R. Valencia, J. Saarinen, H. Andreasson, J. Vallve, J. Andrade-Cetto, and A. J. Lilienthal, "Localization in highly dynamic environments using dual-timescale NDT-MCL," in *Proc. IEEE/RSJ Int. Conf. Robotics and Automation*, 2014, pp. 3956–3962.

[17] B. Huhle, M. Magnusson, W. Strasser, and A. J. Lilienthal, "Registration of colored 3D point clouds with a Kernel-based extension to the normal distributions transform," in *Proc. IEEE Int. Conf. Robotics and Automation*, May 2008, pp. 4025–4030.

[18] A. Nüchter, O. Wulf, K. Lingemann, J. Hertzberg, B. Wagner, and H. Surmann, "3D mapping with semantic knowledge," in *Robot Soccer World Cup*. Springer, 2005, pp. 335–346.

[19] G. Pandey, S. Savarese, J. R. McBride, and R. M. Eustice, "Visually bootstrapped generalized ICP," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2011, pp. 2660–2667.

[20] P. Besl and N. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 14, no. 2, pp. 239–256, Feb 1992.

[21] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP variants on real-world data sets: Open-source library and experimental protocol," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, 2013.

[22] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 3, Apr 1991, pp. 2724–2729.