

Received April 25, 2018, accepted July 20, 2018, date of publication September 6, 2018, date of current version September 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2868244

# Factors to Evaluate Capability of Map for Vehicle Localization

**EHSAN JAVANMARDI<sup>1</sup>, (Member, IEEE), MAHDI JAVANMARDI<sup>2</sup>, (Member, IEEE), YANLEI GU<sup>1</sup>, (Member, IEEE), AND SHUNSUKE KAMIJO<sup>1</sup>, (Member, IEEE)**

<sup>1</sup>Department of Information & Communication Engineering, The University of Tokyo, Tokyo 153-8505, Japan

<sup>2</sup>The Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan

Corresponding author: Ehsan Javanmardi (ehsan@kmj.iis.u-tokyo.ac.jp)

**ABSTRACT** Recently, autonomous vehicle's self-localization based on the matching of laser scanner data to the high definition (HD) map become more popular due to the availability of HD map and price down of light detection and ranging technologies. Many types of research have been done to achieve locally and globally accurate HD map for accurate localization. However, the global accuracy of the map does not guarantee accurate self-localization within the map. To achieve accurate self-localization, the map should satisfy some requirements. In this paper, the focus is made on the map, as one of the high potential sources of error in localization. By investigating the erroneous scenarios in the map and comparing their characteristics, we introduced four criteria for the self-localization ability of the map. These criteria are feature sufficiency, layout, local similarity, and representation quality of the map. Then, in order to quantify these criteria, we introduce several factors for each criterion. Unlike evaluation criteria which are defined regardless of the map formats, factors are defined based on normal distribution map which is a map format of normal distribution transformation scan-matching. These factors are calculated for each position in the map, based on the map features within its local vicinity. By conducting the experiments in Shinjuku, Japan, we have evaluated these factors in a different part of the map with different scenarios by comparing them with the self-localization error.

**INDEX TERMS** Autonomous vehicles self-localization, scan matching, normal distribution transform, map evaluation, high definition map.

## I. INTRODUCTION

Accurate and robust self-localization is an essential task for autonomous vehicles (AVs). On the other hand, high precision self-localization solution when combined with a prebuilt map can simplify the difficult concept of perception and scene-understanding into a less complex positioning problem. Conventional approaches use global navigation satellite system (GNSS) for autonomous vehicle self-localization [1]–[4]. This technology is low cost and works very well in the open sky. However, in the urban environment, the accuracy degrades dramatically due to non-line of sight (NLOS) and multipath reception of the satellites' signals and make it difficult for AV applications [5]. According to [6], in an urban environment, even with the use of high-end IMU's localization error is more than a meter.

In recent years, use of light detection and ranging (LiDAR) for the perception of AVs become more popular due to its price down, miniaturization and density enhancement.

Compared to cameras, LiDARs are more reliable because problems caused by illumination change, light conditions and shadows do not affect it [7]. On the other hand, LiDARs can obtain more accurate distance information comparing to the stereo cameras which make it more suitable for the self-localization applications. LiDAR-based localization can be divided into two main categories. Map-based and without a map which is also known as simultaneous localization and mapping (SLAM). In the SLAM methods [8]–[11], there is no prebuilt map. Based on current position, the surrounding information is stored online as a map. In next time stamp, this stored map (information) are used to calculate the displacement of the vehicle. According to this displacement, the current position of the vehicle is calculated, and again the surrounding information is stored as a map to be used of next time stamp. SLAM methods work well in a short distance; however, due to its dependent nature, they are still suffering from accumulative error in long distances.

Therefore, in recent years, map-based methods gained more attention in most of the AV platforms [12]–[15]. In a map-based method, a raw point cloud of the environment is collected offline using high-end mobile mapping systems (MMSs). Then, based on the map formats, the map is generated from raw point cloud data. Map formats can be one of the point cloud (PCL) map, occupancy map [16], [17], Normal Distribution (ND) map, vector map [18], [19], feature map, planar surface map, etc. Later, in the self-localization phase, the scan acquired from LiDAR mounted on the top of the vehicle is matched to the map to obtain the position of the vehicle within the map. Matching algorithm are depending on the map formats. For example, if the map format is point cloud, for matching the scans to the map, one of the variants of iterative closest point (ICP) [20]–[24], can be used. And if the map format is occupancy map or normal distributions, the variants of grid-based approaches for matching such as [25]–[27] etc., can be used. In map-matching based self-localization, as each sequence are independent, there is no error accumulation.

In the map based categories, map plays a significant role in achieving high accuracy self-localization. For accurate self-localization, the global and local accuracy of the map is essential as well, and many types of research have been done to obtain such a highly accurate map [6], [28]–[30]. However, the highly accurate map does not guarantee the accuracy of the localization [31]. In other words, map accuracy is different than the ability of the map for localization. For example, in the case of the tunnel, no matter how much the map is locally and globally accurate, the lack of longitudinal features in the map causes localization error in moving direction.

To achieve accurate self-localization within a map, the map should satisfy some requirements. In other words, the map should meet some specific criteria which define the ability of the map for self-localization. To the best of this author's knowledge, there is no comprehensive study of the definition and formulation of these criteria. Therefore in this work, the required criteria regarding the ability of map for accurate self-localization are defined and formulated.

Some of these criteria highly related to the environment. For example, in the city environment, there are more structured artifacts than rural places or crop fields. Thus the features for the map-matching can be found easier, and as a result, the localization become more accurate [32]. Tunnels, urban canyons, and highways can be assumed as another example in which environment is not suitable for self-localization using map-matching. In these scenarios, the vehicle is surrounded by two long walls in each side. Thus, the features needed for longitudinal positioning are not enough, and as a result, there will be an error in moving direction. Surrounding environment is different from place to place in the map and should be evaluated by defined criteria.

Some other criteria are related to the quality of representation of the environment by the map. In addition to the sensor related parameters in the mapping phase such as frequency of laser scanner, layer count, the range of the beam, setup

parameters, etc., quality of representation is highly related to the format and abstraction ratio (resolution) of the map. Many types of research have been done to propose an abstracted map format to both reduce the map size and the computational complexity [20], [26], [33]–[38]. Quality of representation of the environment in each of the map formats are different. In other words, some map formats discards more details of the map comparing to others. This information loss of the map might change the quality of some of the criteria and lead to localization error [39]. However, in some part of the map, abstraction does not necessarily change the quality of the map, or rate of change is acceptable. Additionally, in some cases, other criteria might compensate the lack in one criterion. Therefore, in order to evaluate the self-localization ability of the map in a specific point, all criteria should be considered together.

In this paper, four general criteria for the map are defined. These criteria are feature sufficiency, layout, local similarity, and representation quality of the map. These criteria are defined regardless of the map format and can be applied to any other map formats. However, in this paper, to quantify each of these criteria, the focus is made on the ND map format, and several factors are defined. For each point in the map, these factors are calculated based on the features in surroundings called local vicinity. By obtaining the correlation of the map factors with localization error, the effectiveness of the factors is investigated. Additionally, by applying principal component regression (PCR), the predictability of the self-localization error based on to these factors are investigated. Figure 1 shows the overview of proposed map evaluation frameworks.

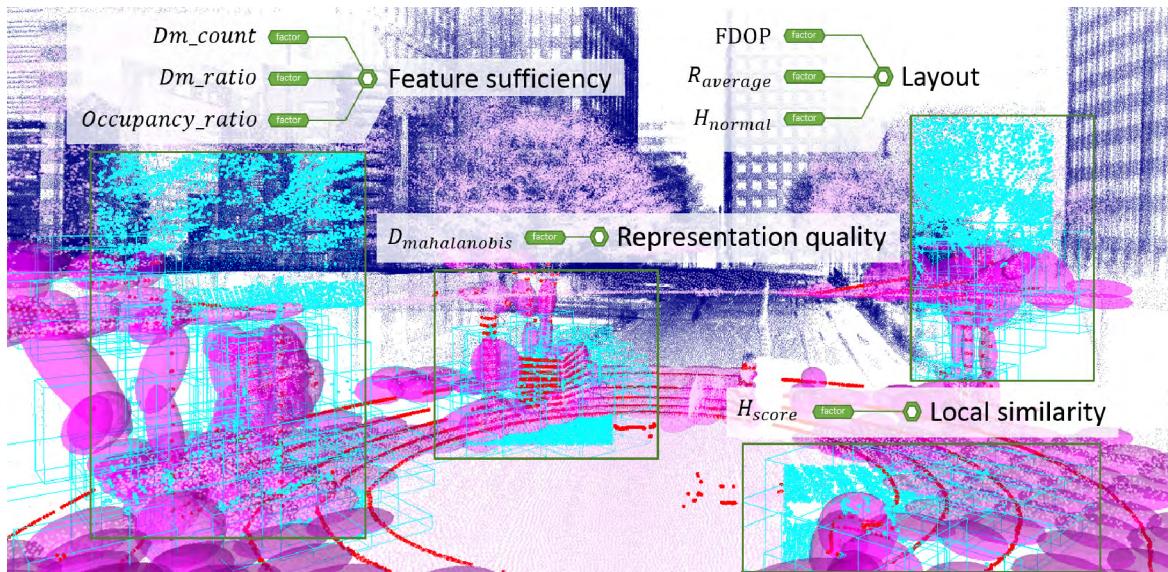
Outcomes of this study can be applied to the dynamic determination of the abstraction ratio of the map and sensor fusion for the upper layer applications as well.

The rest of this paper is organized as follows. Section II introduces related works. Section III describes the map evaluation criteria, and section IV formulates each of the criteria using map factors. Section V evaluates these factors and contribution of each factor to the self-localization error by conducting the experiments. Finally, section VI concludes this paper.

## II. RELATED WORKS

Methods without a prebuilt map are known as SLAM. These methods can be divided into feature-based [11], [40], [41] and scan-based [12], [42]–[44]. Typically, methods based on SLAM suffer from error accumulation. Im *et al.* [15] used a map of building corners to rectify the error accumulation of SLAM. Choi [43] proposed a hybrid map-based SLAM. In their work, the environment is represented by both a grid map and a feature map.

On the other hand, in map-based methods, a point cloud of the environment is stored as a map, and input scan from LiDAR are matched to it. One of the challenges of the map-matching methods is the map size. In point cloud map, for 300m to 300m area, around 20 million points should



**FIGURE 1.** Overview of the proposed map evaluation methods. For each target point in the map, based on features (ND) in local vicinity (purple ellipsoids), different factors for each of the four map evaluation criteria are calculated. Based on the values of the factors, localization error are estimated and based on this estimation ability of the map for self-localization are evaluated.

be stored. Tu *et al.* [45] proposed a method to compress continuous point cloud rather than static point cloud using image compression methods. The author evaluated their compressed point cloud in the self-localization application. One of the solutions for this big amount of data is a voxel-based abstraction. Biber and Strasser [19] first introduced NDT in which the idea was to represent the environment by normal distributions instead of the raw point cloud. In this method, map space is divided into fixed 2D cells, and for each cell, corresponding normal distribution is generated. Then, rather than matching Lidar data to the point cloud, it matches points to the set of normal distributions. Later Magnusson *et al.* extended the idea to the 3D domain and proposed many variants for NDT [33], [39]. The localization accuracy is highly related to the way of space discretization and the grid size. In our previous work [46], multilayer 2D vector map which consists of different heights of the building walls are used. The results show that in some part of the map more abstraction in the map results in more localization error.

One of the challenges of the optimization based map matching methods such as NDT is sensitivity to the initial guess. This sensitivity is different from place to place in the map. In [32], to predict the localization error in different place of the map from different initial guesses, experimentally determined uncertainty is used. In their method, a vehicle once traverses whole paths, and an error model is calculated empirically. Later in the localization phase, this error model is used for rectification. One of the drawbacks of this method is that it needs a high amount of field experiments. Contrary to this method, using proposed factors in this work, localization error can be modeled without a field test. However, error model in [32] directly consider the input scan quality such as density of the scan as well, while in this work, input scan

quality is inferred from map factors and for sure former can model more precisely.

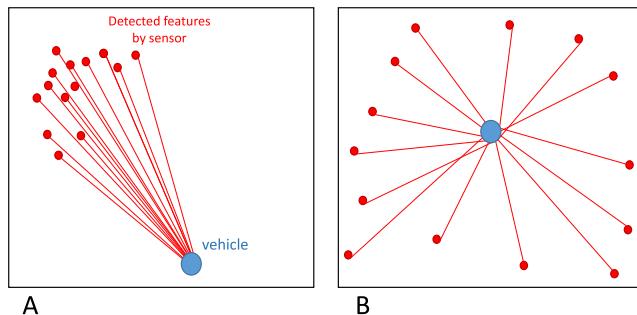
Vysotska and Stachniss [47] used the building information from the open street map (OSM) as background knowledge to predict the erroneous part of the paths to improve the graph-based SLAM. In this work, point cloud map are considered to be available. Thus, the factors are directly extracted from the point cloud. OSM does not contain detailed shape of the building or irregularity in the building walls which highly affect the map matching results. Merzić *et al.* [48] calculates the quality of map for visual odometry by heuristically estimating the entropy of the map.

### III. DEFINITION OF THE MAP EVALUATION CRITERIA

In this section criteria for a map evaluation are defined. These criteria are feature efficiency, layout, local similarity, and representation quality of the map and defined regardless of map formats.

#### A. FEATURE SUFFICIENCY OF THE MAP

In self-localization techniques based on map-matching, pre-built map is made up of features. The type of features varies depending on the map formats and matching algorithms. For example, in the point cloud map the features are points, and in the occupancy map, features are the occupied cells. Likewise, in the ND map the features are NDs, and in the case of vector map and planar surface map, the features are vectors and planar surfaces respectively. For better self-localization accuracy, plenty yet high-quality features are required. The quality and number of extracted features are related to the environment. For example, in the urban city, there are many buildings and structured environments that produce plentiful features in the map. However, in the crop fields, there are no



**FIGURE 2.** Different layout for the map. In (A) all features are lean to one corner. This layout can cause positioning error. In (B), the distribution of the features are even.

sufficient features for matching. Therefore generally in the urban area, the accuracy of localization using laser scanner is higher. In the case of a planar surface map, in the environment with many vegetation and trees, extraction of planar surfaces is difficult, and thus the features for matching are not sufficient.

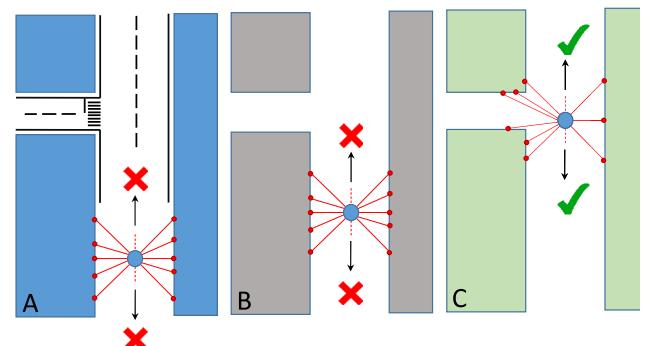
In addition to the number of features, the quality of the features is important as well. In some part of the map, there might be plenty of features; however, as the quality of them are not adequate, the localization cannot be performed accurately. Quality of the features varies in different map format as well. Consider ND map as an example. In order to make ND map, the space of the map are subdivided into static voxels. Then, each voxel is represented by the normal distribution calculated from the position of the points. In the matching phase, these normal distributions lead the input scan point cloud to its optimum location. In ND map, the shape of the normal distribution can be assumed as the quality of the features, and it might play an important role in optimization.

To formulate the feature sufficiency, three factors are proposed. These factors are  $feature_{count}$ ,  $dimension_{ratio}$ , and  $occupancy_{ratio}$ . These factors are described in section IV.

### B. LAYOUT OF THE MAP

In addition to the count and quality of the features, the layout or displacement of the features in the space might be important. In some part of the map, there might be plenty of high-quality features. However, as the features are all placed in one corner of the map, the quality of matching degrades (Figure 2(A)). In Figure 2, blue circle in the center shows the vehicle position and small red circles show the features of the map which are captured by a laser scanner. In Figure 2(A), captured features in the map are distributed evenly. However, in Figure 2(A) the features of the map are placed in the upper left corner of the vehicle. This uneven distribution of the map features can cause the localization error.

In fact, the concept of the layout of the features comes from a global positioning system (GPS). In the GPS-based localization, if satellites are not distributed evenly, in other words, satellites lean in one region, then the calculated position will be erroneous. This is because each of the pseudo-ranges calculated from each satellite has an error by itself.



**FIGURE 3.** Different situations (A and B) in that the localization in longitudinal direction has error.



**FIGURE 4.** Situation that the map does not have any lateral feature to get the longitudinal position.

If all satellites are lean in one small region, then they cannot compensate each of the pseudo range error, and the input error highly appear in the output. For GPS, the ratio that shows how much the input error affect the output localization result is formulated by geometrical dilution of precision (GDOP) [49]. We inspired from GDOP and defined a map factor called feature DOP ( $FDOP$ ) to formulate the distribution of the features in the map space.

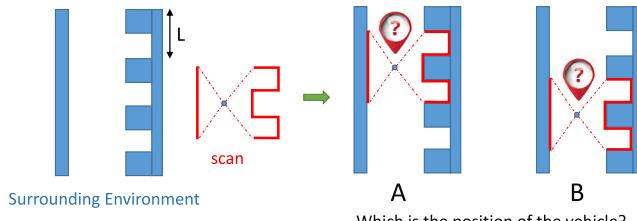
In addition to the distribution of the features, in some part of the map, there might be a situation such as tunnels, urban canyons, and highways in which the layout of the buildings is not suitable for longitudinal positioning (Figure 4). In these situations, the layout of the buildings is so that the lateral position of the vehicle can be obtained accurately but not longitudinal. As shown in Figure 3(A), when the vehicle is in the position A, the layout of the buildings is so that it cannot get the features for longitudinal positioning. In position B, the situation is same. However, in Position C, the layout of the buildings changed and the vehicle is able to observe the features for longitudinal positioning. In both A and B position, the previously defined criteria (feature sufficiency) are satisfied as the car is surrounded by plentiful high-quality features.

We introduce  $normal_{entropy}$  factor to formulate such kind of layout problem in the map.

To formulate the Layout criterion of the map, in addition to aforementioned factors,  $r_{average}$  and  $angular_{entropy}$  are proposed and described in section IV.

### C. LOCAL SIMILARITY OF THE MAP

There are some positions in the map that the environment has many similar features (Figure 5). These similar features make



**FIGURE 5.** Local similarity cause the positioning uncertainty. Longitudinal position cannot be obtained. The blue shapes are the buildings which surround the laser scanner. The red lines are scan on the specific time and shows the observed part of the building by laser scanner. In this scenario, vehicle cannot distinguish between position A and B.

the positioning difficult. In Figure 5, blue shapes are a map, and red lines show the observation of the laser scanner. In this case, as there are similar shapes in the map, it is impossible to detect which is the actual vehicle position; A or B. In other words, in both position A and B the observation of the laser scanner are same. In this situation, the map has local similarity, and this criterion is not satisfied. Again, in this scenario, there are plenty of features, so the feature sufficiency is met, the distribution of the features and layout of the buildings are satisfactory, so the layout criterion is satisfied as well. However, due to the presence of local similarity in this part of the map, localization is erroneous.

Environment and map abstraction both can cause local similarity. If the abstraction ratio become high, then some details of the map are eliminated, and consequently local similarity increases.

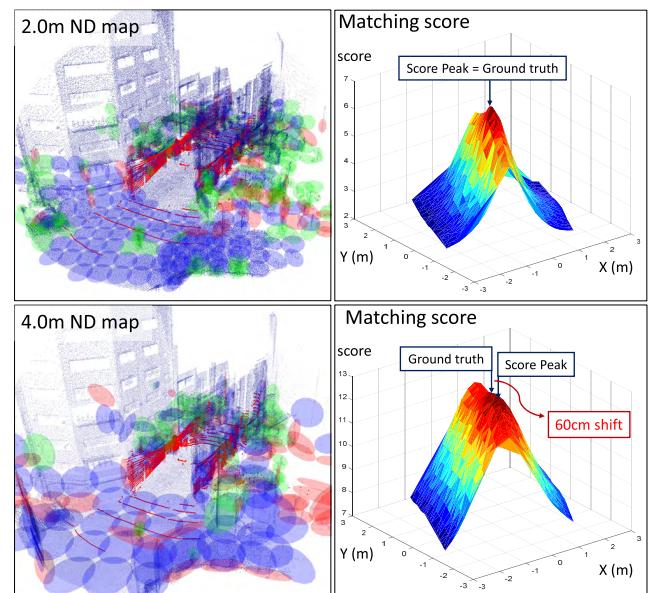
In order to formulate the local similarity,  $score_{entropy}$  is proposed. Formulation of  $score_{entropy}$  is described in section IV.

#### D. REPRESENTATION QUALITY OF THE MAP

One of the important criteria for the map which is directly related to the map format and abstraction ratio is the representation quality of the map. Representation quality shows how much the generated map is similar to the actual environment.

On the one hand, representation quality of the map is related to the type of laser scanner and sensor parameters in the mapping phase. These parameters are number of layers, the rage of beams, scanning frequency, number of sensors, and setup parameters such as pitch and roll angles. As an instance, raw point cloud collected by Velodyne HDL-64 which has 64 layers of laser beams can represent the environment more densely comparing to the VLP-16 which only has 16 layers. However, in this paper, effects of mapping phase to the quality of representation are ignored.

On the other hand, representation quality related to the map format and resolution of the map. In fact, the representation quality criterion shows how the final map is similar to the raw point cloud map collected in the mapping phase. Low representation quality of the environment might change the peak of the score function. If the peak is changed, correct answer from optimization process will be the wrong answer for



**FIGURE 6.** Shift of score peak due to abstraction of the map. Top images shows the 2.0m ND map and the matching score of scan (red points) on it. Score of 2.0m ND map does not have any shift. Bottom image shows the same pair for 4.0m ND map. In 4.0m ND map, due to higher abstraction of the map, peak of matching score is shifted by 60cm. This shift of peak at least cause 60cm localization error.

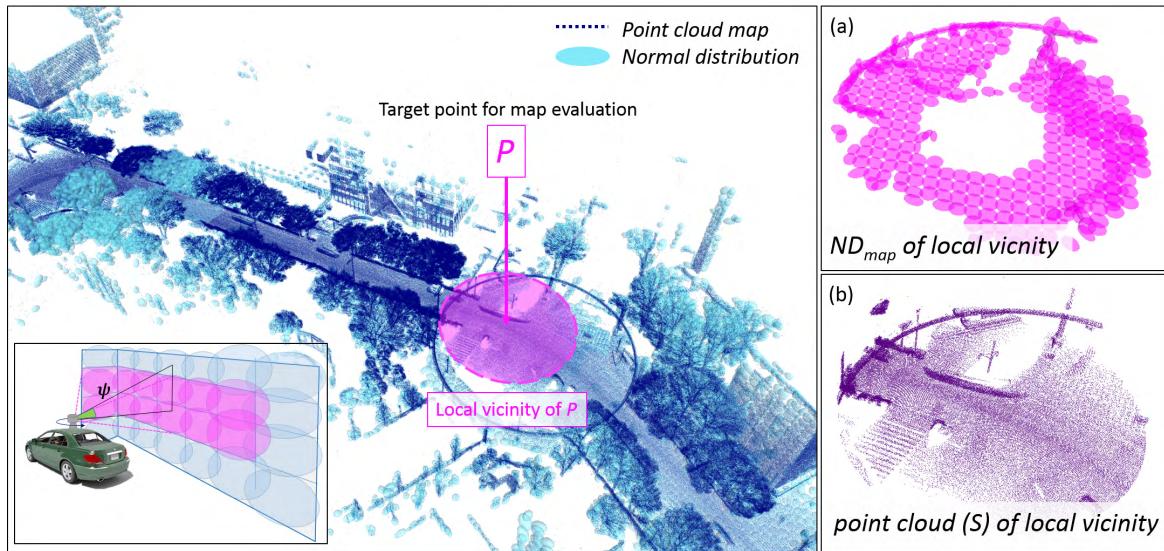
the localization. In Figure 6 change of peak after abstraction in ND map format is shown. Figure 6(top) shows the score of matching for 2.0m grids and below shows the matching for the same region with 4.0m grid size. Score peak should correspond to ground truth. However, in 4.0m grid size, as more abstraction applied to the map, the peak of scores are shifted 60cm. Consequently, even if the optimization process does not stick in local maximum, localization has a 60cm error.

Abstraction does not always change the score peak. Our investigation shows for 2.0m grids, about 77.3% and for 4.0m grid about 43.1% of the cases, peak remains unchanged. This is because abstraction always does not affect the representation quality. Consider a ground or a flat wall. No matter how much the size of grid expands, the representation quality remains same. Therefore, to evaluate the mapping capability for localization, there should be a factor to evaluate the map representation quality.  $mahanobis_{distance}$  factor can formulate the criterion for representation quality and explained in section IV.

#### IV. FORMULATION OF THE EVALUATION CRITERIA

In this section, to quantify each of the proposed criteria for the map, several factors are defined. Map evaluation criteria are defined regardless of map format. However, to formulate each factor, self-localization based on normal distribution transform (NDT) which is well-known in the field of vehicle self-localization is considered.

In NDT methods, instead of directly registering the Velodyne scan to the point cloud map, the scan is registered to the set of normal distributions called  $ND_{map}$  which is



**FIGURE 7.** Local vicinity of target point  $P$ . Local vicinity is extracted based on the range and vertical angle  $\psi$  of the laser scanner.  
(a) ND map of local vicinity. (b) point cloud map is the local vicinity.

made from point cloud map. In order to make  $ND_{map}$ , map space is subdivided into fixed size voxels called cells. Then, from the points inside each cell, normal distribution (ND) is calculated. To register the scan on this NDs, the scan is moved over the NDs with 6D transformation matrix  $T$  and for each pose, likelihood are calculated. The scan pose correspond to the maximum likelihood is obtained by Newton's optimization algorithm and considered as the position of the vehicle within a map. For further discussion readers are referred to [18] and [33].

To calculate the factors for position  $P$ , first, the map elements (normal distributions) in the range of laser scanner beam from  $P$  are extracted from the map. Second, the elevation angle  $\psi$  of each map elements to point  $P$  is calculated. In this paper for the localization, Velodyne VLP-16 is used. The elevation angle range of VLP-16 is  $-15$  to  $+15$  degree. Therefore, from the map elements extracted beforehand, those are considered with the elevation angle  $-15 < \psi < +15$ . These elements are called local vicinity of the position  $P$  in the rest of this paper and shown in Figure 7. Factors for position  $P$  is calculated based on the local vicinity of it.

Point cloud correspond to local vicinity is represented by  $S = \{x_1, x_2 \dots x_n\}$ , where  $n$  is the number of points. NDs correspond to local vicinity is represented by  $ND_{map} = \{N(\mu_1, \Sigma_1), N(\mu_2, \Sigma_2) \dots N(\mu_m, \Sigma_m)\}$ , where  $m$  is the number of NDs, and  $\mu$  and  $\Sigma$  are the mean and covariance matrix of the normal distributions respectively. Figure 7(b) shows the raw point cloud of the extracted region. Figure 7(a) shows the  $ND_{map}$  of the same region which is used for calculation of the factors.

#### A. FACTORS FOR FEATURE SUFFICIENCY CRITERION

The first criterion is feature sufficiency of the map. In order to formulate feature sufficiency, four factors are proposed.

These factors are feature count, dimension count, dimension ratio, and occupancy ratio.

##### 1) FEATURE COUNT, DIMENSION COUNT, AND DIMENSION RATIO

The first factor is for feature sufficiency criteria is feature count ( $feature_{cnt}$ ). It shows the number of map features in the local vicinity. In the case of ND map format, features are normal distributions. More feature count should lead to less localization error.

In order to evaluate the quality of the features in  $ND_{map}$ , the dimension value of the normal distributions are considered. The dimension value of each ND are calculated inspired by [50] as follows:

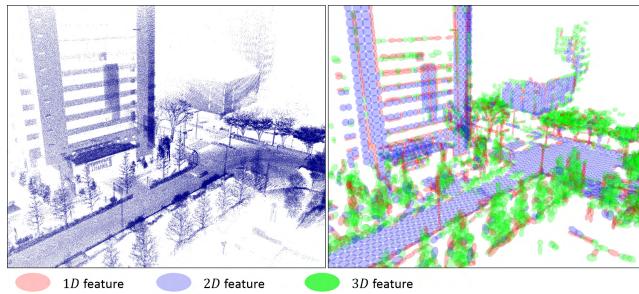
First, for each normal distribution, Eigenvalues are calculated. Consider that the Eigenvalues are  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$ . From Eigen values the standard deviations are calculated as follows:

$$\forall i \in [1, 3] \sigma_i = \sqrt{\lambda_i}, \quad (1)$$

where  $i$  is the index of Eigen values. Using standard deviation, three dimension behavior are defined as follows:

$$a_{1D} = \frac{\sigma_1 - \sigma_2}{\sigma_1}, \quad a_{2D} = \frac{\sigma_2 - \sigma_3}{\sigma_1}, \quad a_{3D} = \frac{\sigma_3}{\sigma_1}, \quad (2)$$

where  $a_{nD}$  is the  $n_{th}$  dimension behavior of the feature. If  $a_{1D} \gg a_{2D}, a_{3D}$ , the feature considered to be 1D feature which is pole-like features and if  $a_{2D} \gg a_{1D}, a_{3D}$ , the feature is 2D feature which is more like walls and planar surfaces. Finally, if  $a_{3D} \gg a_{1D}, a_{2D}$ , then the feature is 3D feature which is more like vegetation or scatters. Figure 8 shows different dimension property of the features with different colors. In Figure 8, red features are 1D, blue is 2D features, and green features are 3D features. As an instance ground and building walls form 2D features and lighting poles and



**FIGURE 8.** 1D (red), 2D(blue) and 3D (green) features in map. Left image is point cloud amp and right image is the ND map generated from left point cloud with 2.0m grid. Ground and building walls are 2D features and poles and trunk of trees are 1D and vegetation and trees has 3D features.

tree trunks are forming 1D features which is red. Using this definition and the property of the features, in addition to the  $feature_{cnt}$ , dimension count ( $Dm_{cnt}$ ) factor is defined.  $Dm_{cnt}$  factor ( $\forall m \in [1, 3]$ ), shows the number of  $Dm$  features in the space. As an instance,  $D1_{cnt}$  shows the number of 1D features in the space.

$$feature_{cnt} = D1_{cnt} + D2_{cnt} + D3_{cnt}, \quad (3)$$

In addition to the number of whole features in the local vicinity which can represent the sufficiency of the features, these factors can represent the quality of the features. In addition to the  $Dm_{cnt}$  factor ( $\forall m \in [1, 3]$ ), the ratio of each feature over all features are considered as another factor which is  $Dm_{ratio}$ . These factors are defined as follows:

$$Dm_{ratio} = \frac{Dm_{cnt}}{feature_{cnt}}, \quad (4)$$

Where  $m$  is the number of dimension and can be one of 1, 2, or 3. For instance,  $D1_{ratio}$  shows the ratio of the 1D features over the all other features.

## 2) OCCUPANCY RATIO

The next factor which can formulate the feature sufficiency criteria is occupancy ratio ( $occupancy_{ratio}$ ). This factor shows how much the surrounding environment which can be seen by laser scanner is occupied with the map features. This factor consider the feature sufficiency from the view of LiDAR instead of 3D space. In previously defined factors such as  $feature_{cnt}$ , some of the features counted on the factors might be redundant because they are not observed by the laser scanner. Laser scanners can only capture the closest object in the scene.

In order to consider this characteristic of the laser scanner, the local vicinity space is converted to the local vicinity depth image as shown in the bottom of Figure 9(b).

This depth image of the map can be seen as 2D histogram as well. Vertical and horizontal resolution of the depth image is set to the number of layers and horizontal resolution of the laser scanner respectively. In this paper, for localization, Velodyne VLP-16 is used. This laser scanner has 16 vertical

layers with resolution of  $2^\circ$ . Therefore the vertical resolution of the depth image is set to 2. Rotation frequency of the laser scanner directly related to its horizontal resolution. The rotation frequency of VLP-16 set to 20Hz (1200 rpm). Horizontal resolution of the VLP-16 in this rotation speed is  $0.4^\circ$ . Therefore the horizontal resolution of the depth image is set to 0.4 as well.

For each normal distribution in the local vicinity, the vertical angle  $\Psi$  and horizontal angle (azimuth)  $\Phi$  to point  $P$  are calculated. According to these two values, the depth image are filled with 0's and 1's. 0's in the depth image mean that there is no corresponding feature in that specified angle and 1's shows the presence of the feature. In Figure 9(b) black shows the absence and white shows the presence of the features.

$occupancy_{ratio}$  factor is the ratio of the occupied cell over all cells of the depth image and calculated as follows:

$$occupancy_{ratio} = \frac{\text{occupied\_cell}}{\text{all\_cells}}, \quad (5)$$

## B. FACTORS FOR THE LAYOUT OF THE FEATURES (MAP)

In order to formulate the layout criteria of the map, feature DOP, normal entropy, and angular entropy are defined.

### 1) FEATURE DOP

Inspiring from GDOP in global positioning systems, Feature DOP ( $FDOP$ ) is calculated.

Assume that the center of local vicinity is  $P = (x_p, y_p, z_p)$  and the position of the  $i_{th}$  ND feature is  $f_i = (x_i, y_i, z_i)$  and its distance to  $P$  is  $r_i$ . Matrix  $A$  is obtained as follows:

$$A = \begin{bmatrix} \frac{(x_1 - x)}{r_1} & \frac{(y_1 - y)}{r_1} & \frac{(z_1 - z)}{r_1} \\ \frac{(x_2 - x)}{r_2} & \frac{(y_2 - y)}{r_2} & \frac{(z_2 - z)}{r_2} \\ \vdots & \vdots & \vdots \\ \frac{(x_m - x)}{r_m} & \frac{(y_m - y)}{r_m} & \frac{(z_m - z)}{r_m} \end{bmatrix}, \quad (6)$$

From multiplication of  $A$  and its transpose,

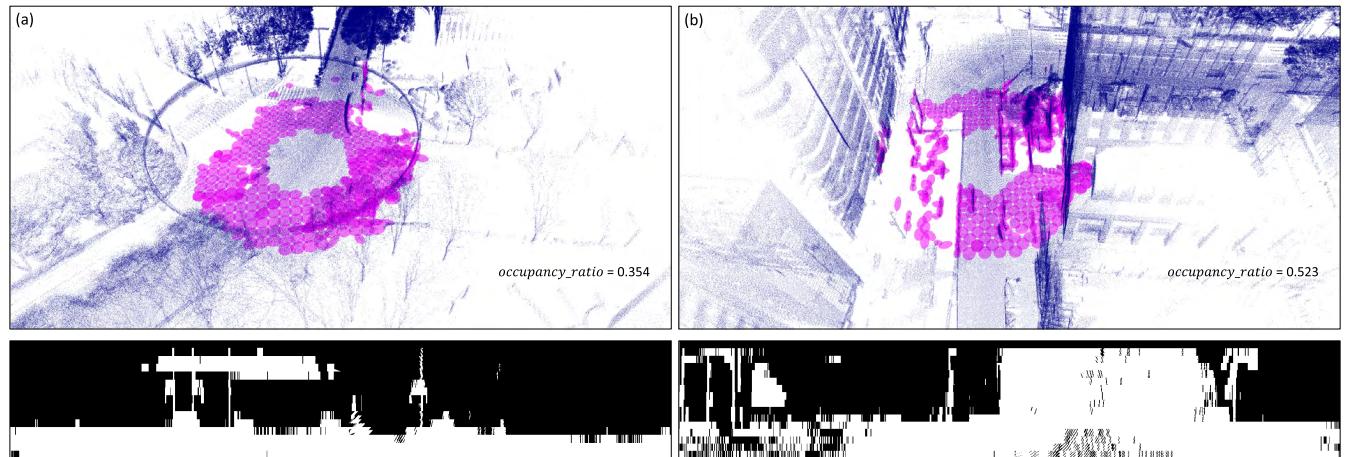
$$Q = A^T A = \begin{bmatrix} \tilde{\sigma}_x^2 & \tilde{\sigma}_{xy} & \tilde{\sigma}_{xz} \\ \tilde{\sigma}_{xy} & \tilde{\sigma}_y^2 & \tilde{\sigma}_{yz} \\ \tilde{\sigma}_{xz} & \tilde{\sigma}_{yz} & \tilde{\sigma}_z^2 \end{bmatrix}, \quad (7)$$

$Q$  is obtained which can be assumed as covariance matrix of the position of the features related to the vehicle. Finally,  $FDOP$  is calculated as follow:

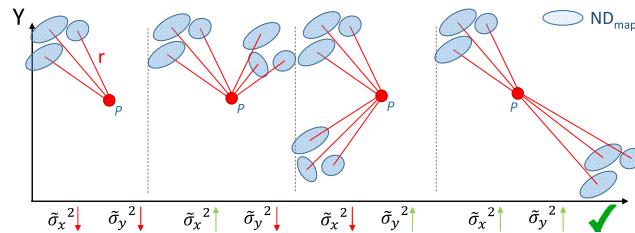
$$FDOP = 1/\sqrt{\tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + \tilde{\sigma}_z^2}, \quad (8)$$

If features are distributed uniformly in the space, the values of  $\tilde{\sigma}_x^2$ ,  $\tilde{\sigma}_y^2$ , and  $\tilde{\sigma}_z^2$  are high and thus  $FDOP$  is low.  $FDOP$  has a direct relation to the localization error.

In Figure 10 several scenarios for FDOP are shown. FDOP related to the environment and abstraction ratio does not affect it.



**FIGURE 9.** Depth images (bottom images) correspond to the features in local vicinity (purple ellipsoids). In depth image white shows presence of features and black shows the absence of features. Occupancy ratio for the left image is 0.354 and for the left is 0.523. In right scenario, local vicinity is occupied with more features thus the occupancy ratio is higher.



**FIGURE 10.** Comparison of the layout (distribution) of the feature with lateral and longitudinal FDOF.

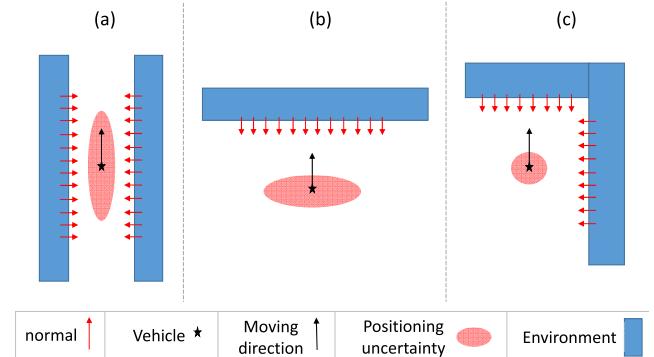
## 2) NORMAL ENTROPY

One of the important questions that should be answered in order to evaluate the layout criterion of the map is that, is the layout of the environment capable of providing features for both longitudinal and lateral positioning. Generally, in the matching based methods, the uncertainty of the positioning is related to the features normal.

In Figure 11 three scenarios with features with different normal are shown. Red arrows show normal of the features, and red ellipsoid around the car shows the positioning uncertainty. In Figure 11(a), all of the features are parallel to the moving direction of the vehicles. In this scenario, the longitudinal position cannot be calculated accurately, and the uncertainty in the moving direction is high. In Figure 11(b), the features are perpendicular to the moving direction. In this scenario, the longitudinal position can be obtained. However, the lateral position has a huge error. And finally in Figure 11(c), due to the presence of features in both direction, both longitudinal and lateral positioning can be obtained.

Consequently, if the features in the local vicinity face more directions, the positioning uncertainty decreases. Normal entropy factor ( $H_{normal}$ ) can show the degree of dispersity of the direction of the features.

To calculate  $H_{normal}$ , first, Eigenvalues and Eigen vectors of the NDs in the local vicinity are calculated. Second, for



**FIGURE 11.** Vehicle position uncertainty (red ellipsoid) in different layout of the building. (a) has high uncertainty in moving direction. (b) has high localization error in lateral direction. (c) layout of the building make the localization more accurate.

each ND, from Eigen values, normal is calculated. Corresponding Eigen vector to the smallest Eigen value are considered as the normal. Then the azimuth angle  $\phi$  and elevation angle  $\Psi$  of the normal are calculated. According to  $\phi$  and  $\Psi$ , the normal are stored in  $m \times m$  bins histogram. In this paper the  $H_{normal}$  is evaluated with  $m = 8 \times 8, 16 \times 16$ , and  $90 \times 90$  bins.

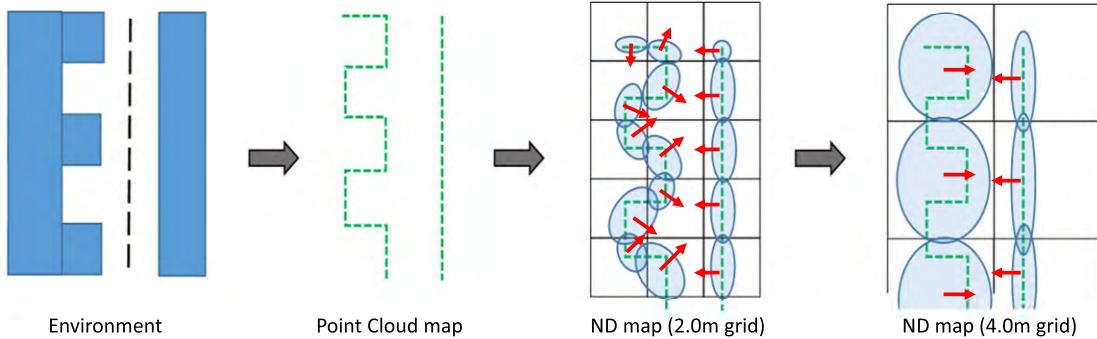
Finally, from this normal angle histogram,  $H_{normal}$  is calculated as follows:

$$H_{normal} = - \sum_{i=1}^b P_{normal}(i) (\log_2 P_{normal}(i)), \quad (9)$$

where  $b$  is the number of bins, and  $P_{normal}(i)$  is the probability of occurrence of  $i_{th}$  bin and calculated using following equation:

$$P_{normal}(i) = \frac{h_i}{\sum_{i=1}^b h_i}, \quad (10)$$

where  $h_i$  is the value of the  $i_{th}$  bin of the normal angle histogram.



**FIGURE 12.** Effect of map resolution on the normal of the features. Red arrows are the normal of each feature (ND).

If local vicinity has a higher degree of dispersity of normal, the histogram bins filled more evenly, therefore the entropy of the normal angle histogram increase. On the other hand, if the  $H_{normal}$  shows lower value, it means that the local vicinity has a low level of scatter, and localization might be erroneous. Therefore localization accuracy has direct relation with  $H_{normal}$  as shown in the followings.

$$\text{localization accuracy} \propto H_{normal}, \quad (11)$$

Value of  $H_{normal}$  is both related to the environment and abstraction ratio of the map. Figure 12 shows the effect of abstraction ratio in the  $H_{normal}$ . As shown in 4.0m ND map, the abstraction causes the normals to point to the same direction. In this case localization in the moving direction might have error.

### 3) ANGULAR ENTROPY

One of the factors which can help the quantification of the layout of the map is angular entropy ( $H_{angular}$ ). This factor shows a degree of uniformity of the feature distribution in the space. In order to calculate  $H_{angular}$ , the azimuth of all features are calculated. Then, based on azimuth value, angular histogram are filled. Angular histogram is one dimension and has 90 bins. From the angular histogram  $H_{angular}$  is calculated as follows:

$$H_{angular} = - \sum_{i=1}^b P_{angular}(i) (\log_2 P_{angular}(i)), \quad (12)$$

where b is the number of bins which is 90, and  $P_{normal}(i)$  is the probability of occurrence of  $i_{th}$  bin and calculated using following equation:

$$P_{angular}(i) = \frac{h_i}{\sum_{i=1}^b h_i}, \quad (13)$$

where  $h_i$  is the value of the  $i_{th}$  bin of the angular histogram.

The more features distribute uniformly around the vehicle, the higher  $H_{angular}$  achieved. Higher  $H_{angular}$  value might have positive effects on the localization accuracy. Abstraction ratio does not affect the  $H_{angular}$ .

### 4) AVERAGE R

Average r ( $r_{average}$ ) factor is the average of the distance of the features in the local vicinity from the point  $P$  (center of local vicinity) and calculated as follows:

$$r_{average} = \sum_{i=1}^m r_i, \quad (14)$$

where  $m$  is the number of features (NDs) in the local vicinity and  $r_i$  is the distance to the  $i_{th}$  feature from the center of the local vicinity.

### C. FACTOR FOR LOCAL SIMILARITY CRITERION

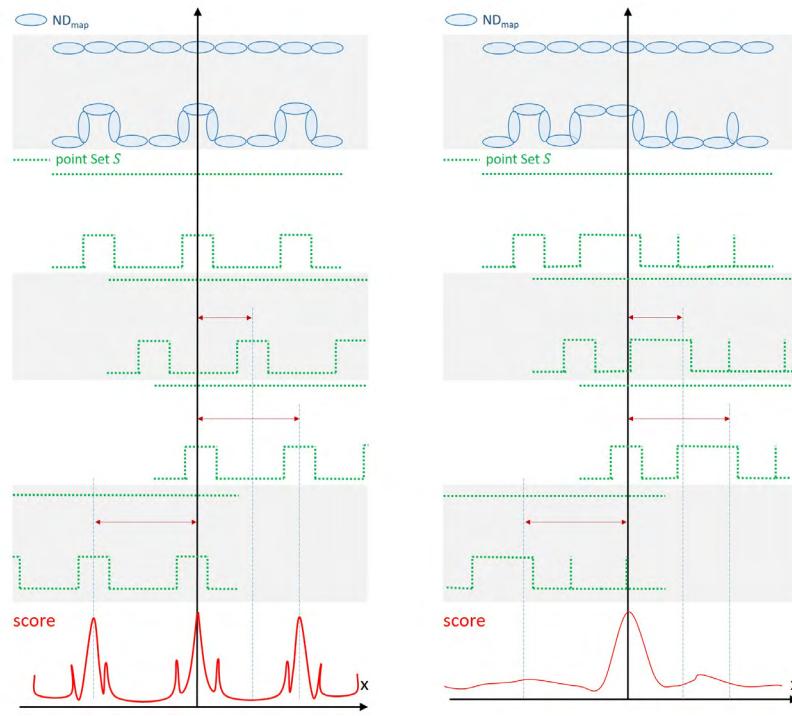
To formulate the local similarity of the map in position  $P$ , the score entropy ( $H_{score}$ ) is introduced.

The score of registration of local vicinity point cloud  $S$  on its  $ND_{map}$  is calculated as

$$S(\vec{v}) = - \sum_{k=1}^n \tilde{p}(T(\vec{v}, x_i)), \quad (15)$$

where  $\tilde{p}$  is the simplified log-likelihood function of the nearest normal distribution to the point  $x_i$  in the  $ND_{map}$  and  $T(\vec{v}, x_i)$  is the transformation function which transforms  $x_i$  with 2D transformation vector  $\vec{v}(x, y)$ . Here  $\vec{v}$  can only transform point  $x_i$  in two dimensions,  $x$  dimension and  $y$  dimension.

It is obvious that if the transformation vector is set to  $\vec{v}(0, 0)$ , the score function shows maximum value. As shown in Figure 13(b) if the point cloud is shifted  $\Delta$  to the right or left the score function decreases as the likelihood of  $S$  and  $ND_{map}$  are decreases. However, in Figure 13(a), as map has repetitive characteristics, by shifting  $S$  to the left and right, numerous peaks appear. These peaks usually known as local optimum. In the localization phase, where input scan is matched to the map, the optimization process might stuck in one of these peaks (local optimum) which is a wrong answer. Comparison of the score function  $S_a(\vec{v})$  with  $S_b(\vec{v})$ , which both are two dimension distributions, shows that in the presence of the repetitive objects in the map which is called local similarity here, the score function will distribute more evenly. Therefore if the score function is assumed as a probability distribution, and the entropy of the score function



**FIGURE 13.** Score for two different scenario. Left environment has local similarity and right does not. Left score has more peaks. Score is obtained by moving point cloud (green points) over ND map (blue ellipsoid) and calculating likelihood.

is calculated,  $H_{score}$  in the scenario such as Figure 13(a) shows higher value.

In order to calculate  $H_{score}$ , first the score is changed to the probability distribution as follows:

$$P_{score}(\vec{v}) = \frac{s(\vec{v})}{\sum s(\vec{v})}, \quad (16)$$

where  $s(\vec{v})$  is the score of registration for  $\vec{v}$  and  $P_{score}(\vec{v})$  is the corresponding probability. Using (16),  $H_{score}$  is defined as the entropy of the  $P_{score}$  as follows:

$$H_{score} = - \sum_{\vec{v}} P_{score}(\vec{v}) \log_2(P_{score}(\vec{v})), \quad (17)$$

Range and steps of the transformation vector  $\vec{v}$  is defined based on the initial guess requirements and here set to  $[-2m, 2m]$  for both direction  $x$  and  $y$  with steps of  $20cm$ .

As shown in the following, score entropy has a direct relation to the local similarity criterion.

$$\text{local similarity} \propto H_{score}, \quad (18)$$

#### D. FACTOR FOR REPRESENTATION QUALITY CRITERION

In order to quantify the representation criterion, Mahalanobis distance is proposed ( $D_{mahalanobis}$ ). This factor shows how much a  $ND_{map}$  of the local vicinity was able to preserve the details of the raw point cloud  $S$ .

To formulate this, for each point in the raw point cloud of local vicinity  $S$ , distance to corresponding  $ND$  is calculated. Distance of point to the  $ND$  is calculated by Mahalanobis

distance as it consider the scaling parameters. Mahalanobis distance  $d_{mah}$  of the point  $x_i$  to the  $ND_j$  is calculated as follows:

$$d_{mah}(x_i) = \sqrt{(x_i - \mu_j)^T \sum_j^{-1} (x_i - \mu_j)}, \quad (19)$$

where  $\mu_j$  and  $\sum_j$  are mean and covariance of the nearest  $ND_j$  to  $x_i$  respectively. Then average of these points for the local vicinity are considered as  $D_{mahalanobis}$  factor and calculated as follows:

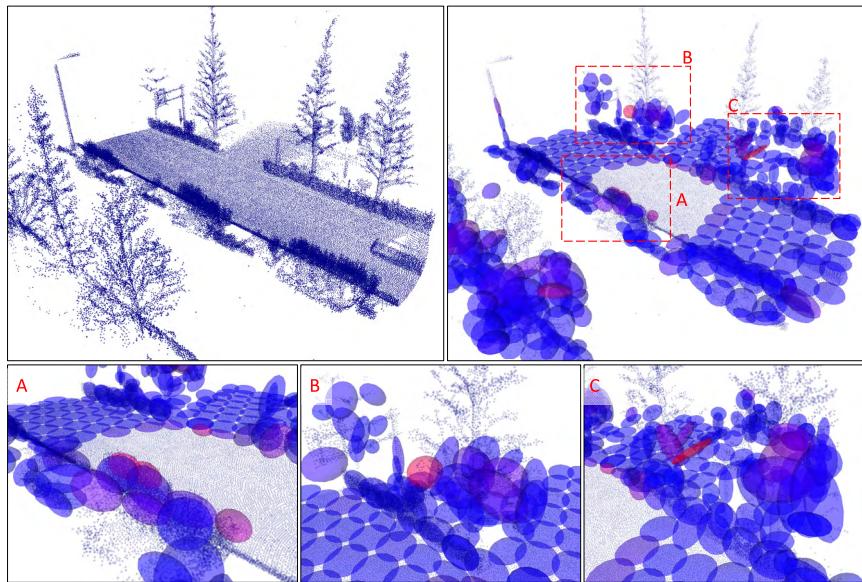
$$D_{mahalanobis} = \frac{\sum_{i=1}^n d_{mah}(x_i)}{n}, \quad (20)$$

where  $n$  is the total number of points in the local vicinity of  $P$ .

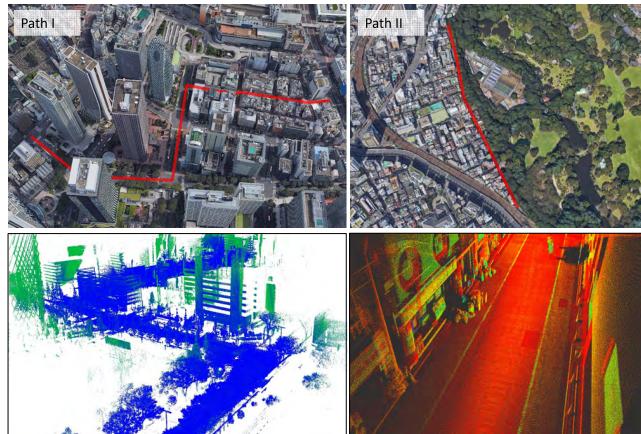
Figure 14 shows  $D_{mahalanobis}$  for different NDs. Blue shows lower  $D_{mahalanobis}$  values and red shows higher values. NDs correspond to ground shows lower Mahalanobis values, means the representation quality is high. However, vegetation and trees have a color near to red. This is because in these points representation quality is low. In fact, for the trees and vegetation, ND with  $2.0m$  grid cannot represent the details.

#### V. EVALUATION OF THE MAP FACTORS

So far, for each map evaluation criterion, several factors are introduced. In this section, the effectiveness of each factors is evaluated. In order to evaluate the contribution of each factor



**FIGURE 14.** Different Mahalanobis distance values for different NDs. Grid size is 2.0m. Blue shows lower mahalanobis distance and red shows higher values. Ground shows low value means the representation quality is high and vegetation and trees have higher Mahalanobis value means the representation quality is low.



**FIGURE 15.** The experimental area around Shinjuku, a dense urban area in Tokyo, Japan. Path I has around 1.2Km and Path II has 0.8Km. Bottom figures show some part of generated point cloud map.

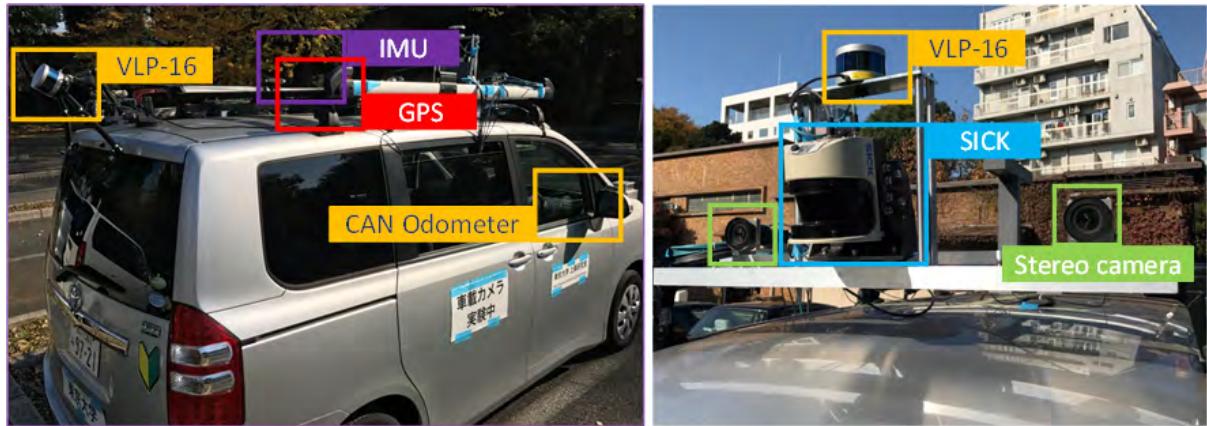
in the map matching-based localization error, an experiment was conducted in Shinjuku, a dense urban area of central Tokyo, Japan (Figure 15). Streets around Shinjuku is surrounded by skyscrapers, tall buildings, narrow streets, and trees. This scenario makes GPS-based positioning solution impractical.

The experiments was conducted in two paths (Figure 15). Path I and II that have 1.2Km and 0.8Km length respectively. Point cloud map is made by VLP-16 which is mounted on the back of our experimental vehicle (Figure 16) which is tilted 75° (pitch). Figure 15(bottom) shows a part of the generated map. The global accuracy of the map is achieved by applying automatic calibration of road markings to the airborne image [6].

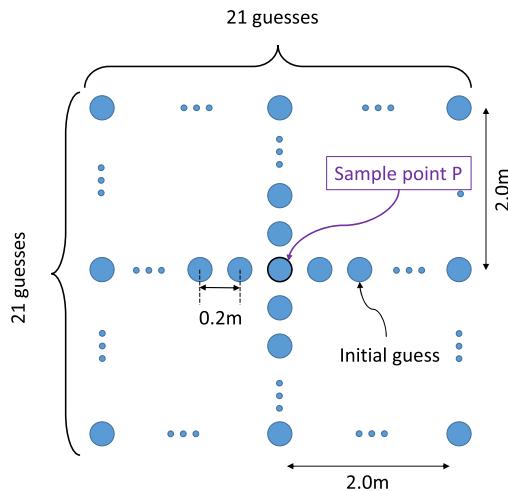
For each path, ND map with different grid size (2.0m ~ 5.0m) was generated from the point cloud map. Then for each ND map, self-localization based on NDT matching is performed, and localization error is obtained. For self-localization again VLP-16 with normal setup (0° pitch) is used (Figure 16). To avoid the scan distortion due to the motion of the vehicle, the vehicle's velocity was below 2 m/s while the frequency of the laser scanner was set to 20 Hz which limits the distortion in each scan to less than 10 cm.

In order to evaluate the map factors, along with the trajectory of the vehicle, sample points with 1.0m interval were selected. For each sample point, map factors were calculated. And for each sample point, NDT-based map matching was performed and mean, and max error was calculated. As initial guess affects the localization accuracy, for each sample point, errors for different initial guesses are obtained.

Initial guesses are distributed around the sample point within 2.0m to four sides as rectangle shapes with 0.2m intervals. Therefore for each sample point,  $21 \times 21 = 441$  times matching with different initial guesses are performed and mean, and max error are calculated. The distribution of initial guesses are shown in Figure 17. Ground truth is obtained in two steps. First, by matching the input scan (which down-sampled by the 1.0m grid) to the 1.0m ND map rough position is obtained. As the ND map is 1.0m, the result from the first stage is usually accurate, thus it can be assumed as ground truth. However, in order to not stuck in local optimum and confirm the result of the first stage, secondly, all the peaks around (within 50cm) the first result are calculated iteratively and compared to the previous peak. Then ground truth is updated with the position corresponding to the highest peak.



**FIGURE 16.** Mapping vehicle description: The front VLP-16 laser scanners was mounted horizontally to maximize the sensor range for SLAM. The rear VLP-16 was configured to point down (pitch  $-75^\circ$ ) to densely scan the buildings for the mapping.



**FIGURE 17.** Distribution of initial guesses around the sample point P.

Localization errors and factors related to the feature sufficiency criterion for a 2.0m grid of path II is shown in Figure 18. The first and second figures show the maximum and mean error of the localization respectively. In B-C and D-E periods, localization shows higher accuracy. However, in A-B and C-D high error appeared. These errors came from many factors which can be assessed from Figure 18, Figure 19, and Figure 20. Figure 18 shows that the  $occupancy_{ratio}$  and  $feature_{cnt}$  is highly related to the mean and max error. In A-B and C-D where error exist, these two values are lower. Among the  $Dm_{ratio}$  factor ( $\forall m \in [1, 3]$ ),  $D3_{ratio}$  contribute more to the localization accuracy. Whenever the ratio of  $D3$  features lowered the localization error increased.

Localization errors and factors related to the Layout criterion for a 4.0m grid of path II is shown in Figure 19. A-B period of Figure 19 shows,  $r_{average}$  and  $H_{angular}$  react better than other factors. It is shown that if the  $r_{average}$  is higher, localization become more accurate. In C-D,  $FDOP$ ,  $H_{angular}$ , and  $H_{normal}$  react well. For both A-B and C-D,

$H_{normal}$  reacts well. As shown in Figure 19,  $H_{normal}$  is calculated for  $90 \times 90$ ,  $16 \times 16$ , and  $8 \times 8$  bins. Among these three,  $90 \times 90$  bins can show error model more precisely.

Localization errors and factors related to both local similarity and representation quality criteria for a 2.0m grid of path II is shown in Figure 20. Form it can be seen that, errors are made of several factors, and each of factors reacts in a specific situation. In other words, each of the factors can model something that others cannot. Thus, for modeling the error, all of them are essential. Factors can overlap in some situations.

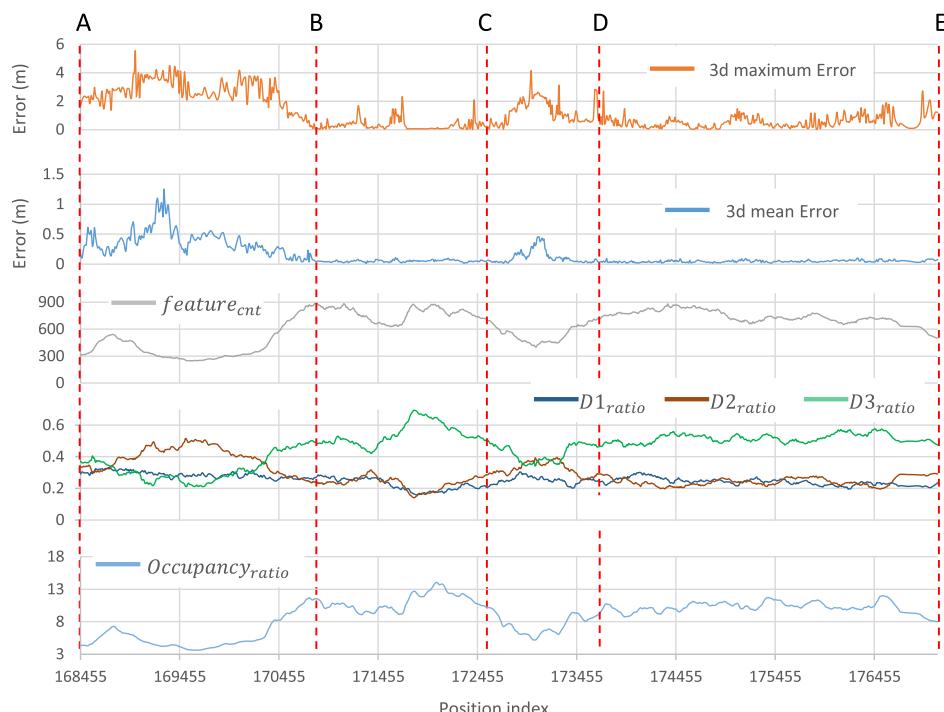
For example, in Figure 18, both  $occupancy_{ratio}$  and  $feature_{cnt}$  follows the similar trend. However, it is not always and at the same time they show different characteristics of the map as well. TABLE 1 shows a correlation of the factors related to the feature sufficiency with other factors and also with the mean and max error.

Likewise, TABLE 2 shows correlations for factors related to the layout, local similarity and representation quality. Correlation of the factors with errors can be assumed as effectiveness or contribution of each factor to the error. In other words, the higher the correlation value with error model, the more that factor contribute to the error. In TABLE 1 and TABLE 2, a value near to 1.0 (green cells) shows two variables are act highly related to each other and values near to -1.0 (red cells) shows they are act nearly opposite of each other. Values near the 0 (yellow cells) means they are not related at all. In TABLE 1,  $occupancy_{ratio}$  and  $feature_{cnt}$  has 93% of correlation. High correlation value could be expected from Figure 18 as well. Factors related to the feature sufficiency (TABLE 1), has high correlation values with localization error. Among them,  $feature_{cnt}$  has the highest correlation with 0.843. Contribution of each factor is different in mean and maximum error. Correlation of  $D2_{ratio}$  with error (0.761 for mean, 0.808 for max) shows that if the proportion of  $D2$  increase, the localization increase too. Among the features in TABLE 2,  $FDOP$  and  $D_{mahalanobis}$  has lower

**TABLE 1.** Correlation of factors related to the feature sufficiency with each other and mean and max error for path ii (2.0m nd).

| criteria                 | Feature sufficiency    |                   |                   |                   |                     |                     |                     |                        | Localization Error |                      |
|--------------------------|------------------------|-------------------|-------------------|-------------------|---------------------|---------------------|---------------------|------------------------|--------------------|----------------------|
|                          | feature <sub>cnt</sub> | D1 <sub>cnt</sub> | D2 <sub>cnt</sub> | D3 <sub>cnt</sub> | D1 <sub>ratio</sub> | D2 <sub>ratio</sub> | D3 <sub>ratio</sub> | occup <sub>ratio</sub> | $\mu_{3DErr}$      | max <sub>3DErr</sub> |
| feature <sub>cnt</sub>   | <b>1.000</b>           | 0.866             | 0.639             | 0.964             | -0.504              | -0.912              | 0.866               | 0.932                  | -0.764             | -0.843               |
| D1 <sub>cnt</sub>        | 0.866                  | <b>1.000</b>      | 0.722             | 0.721             | -0.017              | -0.716              | 0.552               | 0.679                  | -0.624             | -0.672               |
| D2 <sub>cnt</sub>        | 0.639                  | 0.722             | <b>1.000</b>      | 0.444             | -0.081              | -0.344              | 0.290               | 0.518                  | -0.517             | -0.515               |
| D3 <sub>cnt</sub>        | 0.964                  | 0.721             | 0.444             | <b>1.000</b>      | -0.670              | -0.950              | 0.952               | 0.955                  | -0.743             | -0.837               |
| D1 <sub>ratio</sub>      | -0.504                 | -0.017            | -0.081            | -0.670            | <b>1.000</b>        | 0.586               | -0.787              | -0.690                 | 0.509              | 0.557                |
| D2 <sub>ratio</sub>      | -0.912                 | -0.716            | -0.344            | -0.950            | 0.586               | <b>1.000</b>        | -0.961              | -0.902                 | 0.761              | 0.808                |
| D3 <sub>ratio</sub>      | 0.866                  | 0.552             | 0.290             | 0.952             | -0.787              | -0.961              | <b>1.000</b>        | 0.922                  | -0.753             | -0.805               |
| occup <sub>ratio</sub>   | 0.932                  | 0.679             | 0.518             | 0.955             | -0.690              | -0.902              | 0.922               | <b>1.000</b>           | -0.752             | -0.831               |
| r <sub>ave</sub>         | 0.712                  | 0.612             | 0.555             | 0.667             | -0.406              | -0.655              | 0.637               | 0.641                  | -0.631             | -0.674               |
| FDOP                     | 0.266                  | 0.256             | 0.145             | 0.254             | -0.042              | -0.217              | 0.180               | 0.256                  | -0.057             | -0.134               |
| H <sub>angular</sub>     | 0.923                  | 0.811             | 0.631             | 0.878             | -0.460              | -0.906              | 0.847               | 0.874                  | -0.825             | -0.809               |
| H <sub>normal (90)</sub> | 0.960                  | 0.804             | 0.588             | 0.939             | -0.542              | -0.950              | 0.908               | 0.924                  | -0.836             | -0.844               |
| H <sub>normal (16)</sub> | 0.902                  | 0.686             | 0.452             | 0.926             | -0.637              | -0.961              | 0.949               | 0.921                  | -0.824             | -0.827               |
| H <sub>normal (8)</sub>  | 0.727                  | 0.441             | 0.238             | 0.807             | -0.700              | -0.824              | 0.866               | 0.807                  | -0.763             | -0.717               |
| D <sub>mahalanobis</sub> | -0.149                 | -0.059            | -0.141            | -0.157            | 0.179               | 0.179               | -0.197              | -0.190                 | 0.026              | 0.164                |
| H <sub>score</sub>       | 0.646                  | 0.486             | 0.301             | 0.668             | -0.470              | -0.703              | 0.696               | 0.617                  | -0.592             | -0.654               |

Green cells show high positive relation and red cells show high negative relation. Yellow cells has no relation. In this table, occup<sub>ratio</sub> is occupancy<sub>ratio</sub> and r<sub>ave</sub> is r<sub>average</sub>.

**FIGURE 18.** Factors related to the feature sufficiency for path II (2.0m ND) compared to mean and max error of the same path.

correlation to other factors. Therefore these two factors can evaluate some characteristics of the map that cannot be evaluated by others and make them a key factor for modeling the error. On the other hand, FDOP and D<sub>mahalanobis</sub> factors has low correlation with the localization error (TABLE 2). This is because some of the map criteria does not appear frequently, thus, corresponding factors has lower correlation to the error. Low correlation in the TABLE 1 and TABLE 2

do not necessarily mean the inability of the factors for map evaluation.

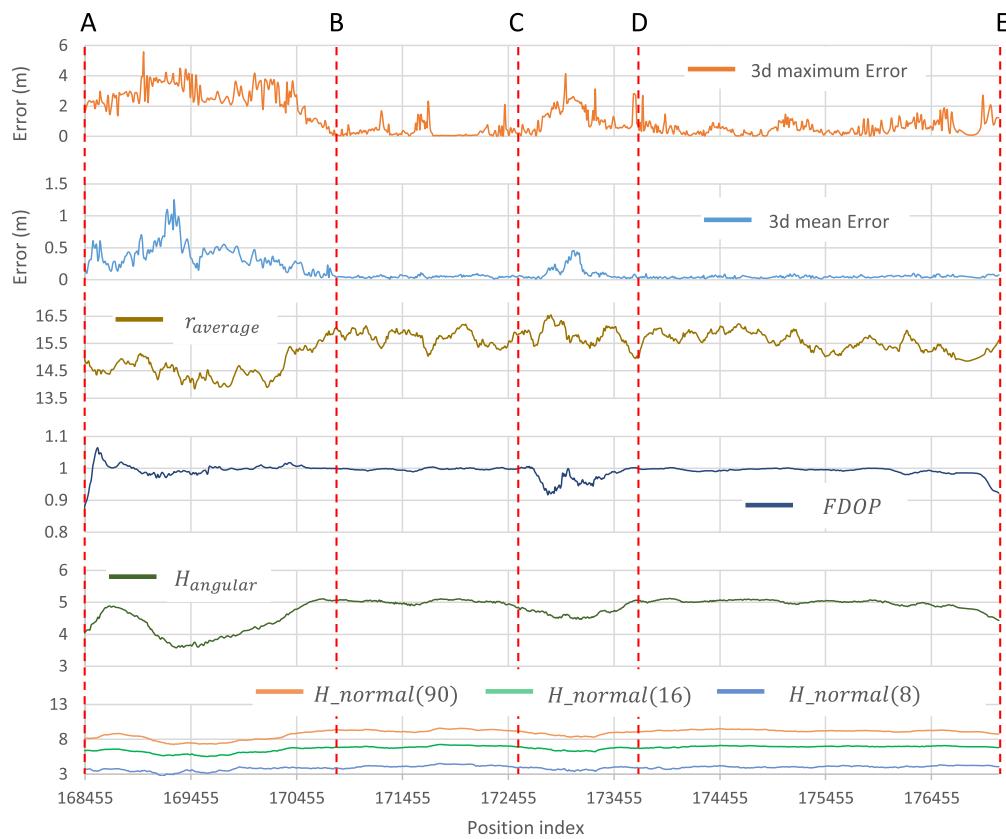
In order to model the localization error based on the affermentioned factors, we use principal component regression. Principal component regression (PCR) is a regression analysis technique that is based on the principal component analysis (PCA). PCR applied to a set of factors as explanatory variables and localization error as a response.

**TABLE 2.** Correlation of factors related to the layout, local similarity, and representation quality with each other and mean and max error for path ii (2.0m nd).

| criteria          | Layout       |              |               |                  |                  |                 | Representation quality | Local similarity | Localization Error |               |
|-------------------|--------------|--------------|---------------|------------------|------------------|-----------------|------------------------|------------------|--------------------|---------------|
| factors           | $r_{ave}$    | FDOP         | $H_{angular}$ | $H_{normal}(90)$ | $H_{normal}(16)$ | $H_{normal}(8)$ | $D_{mah}$              | $H_{score}$      | $\mu_{3DErr}$      | $max_{3DErr}$ |
| $r_{ave}$         | <b>1.000</b> | -0.190       | 0.684         | 0.727            | 0.654            | 0.463           | -0.372                 | 0.580            | -0.631             | -0.674        |
| FDOP              | -0.190       | <b>1.000</b> | 0.290         | 0.222            | 0.210            | 0.204           | 0.286                  | -0.032           | -0.057             | -0.134        |
| $H_{angular}$     | 0.684        | 0.290        | <b>1.000</b>  | 0.972            | 0.927            | 0.745           | -0.143                 | 0.640            | -0.825             | -0.809        |
| $H_{normal}(90)$  | 0.727        | 0.222        | 0.972         | <b>1.000</b>     | 0.971            | 0.811           | -0.189                 | 0.700            | -0.836             | -0.844        |
| $H_{normal}(16)$  | 0.654        | 0.210        | 0.927         | 0.971            | <b>1.000</b>     | 0.883           | -0.177                 | 0.696            | -0.824             | -0.827        |
| $H_{normal}(8)$   | 0.463        | 0.204        | 0.745         | 0.811            | 0.883            | <b>1.000</b>    | 0.022                  | 0.544            | -0.763             | -0.717        |
| $D_{mahalanobis}$ | -0.372       | 0.286        | -0.143        | -0.189           | -0.177           | 0.022           | <b>1.000</b>           | -0.421           | 0.026              | 0.164         |
| $H_{score}$       | 0.580        | -0.032       | 0.640         | 0.700            | 0.696            | 0.544           | -0.421                 | <b>1.000</b>     | -0.592             | -0.654        |
| $Feature_{cnt}$   | 0.712        | 0.266        | 0.923         | 0.960            | 0.902            | 0.727           | -0.149                 | 0.646            | -0.764             | -0.843        |
| $D1_{cnt}$        | 0.612        | 0.256        | 0.811         | 0.804            | 0.686            | 0.441           | -0.059                 | 0.486            | -0.624             | -0.672        |
| $D2_{cnt}$        | 0.555        | 0.145        | 0.631         | 0.588            | 0.452            | 0.238           | -0.141                 | 0.301            | -0.517             | -0.515        |
| $D3_{cnt}$        | 0.667        | 0.254        | 0.878         | 0.939            | 0.926            | 0.807           | -0.157                 | 0.668            | -0.743             | -0.837        |
| $D1_{ratio}$      | -0.406       | -0.042       | -0.460        | -0.542           | -0.637           | -0.700          | 0.179                  | -0.470           | 0.509              | 0.557         |
| $D2_{ratio}$      | -0.655       | -0.217       | -0.906        | -0.950           | -0.961           | -0.824          | 0.179                  | -0.703           | 0.761              | 0.808         |
| $D3_{ratio}$      | 0.637        | 0.180        | 0.847         | 0.908            | 0.949            | 0.866           | -0.197                 | 0.696            | -0.753             | -0.805        |
| $occup_{ratio}$   | 0.641        | 0.256        | 0.874         | 0.924            | 0.921            | 0.807           | -0.190                 | 0.617            | -0.752             | -0.831        |



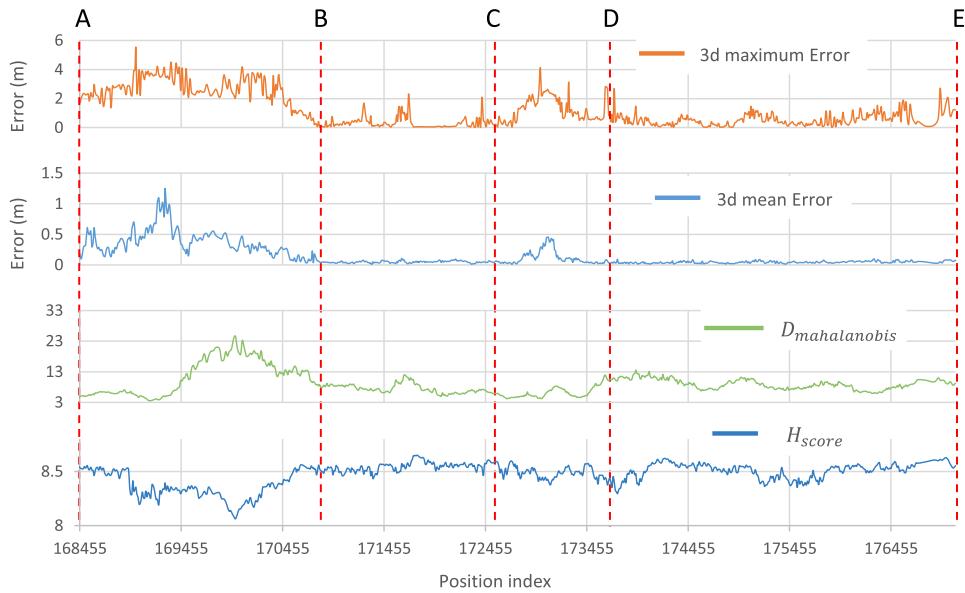
Green cells show high positive relation and red cells show high negative relation. Yellow cells has no relation. In this table,  $occup_{ratio}$  is  $occupancy_{ratio}$  and  $r_{ave}$  is  $r_{average}$ .



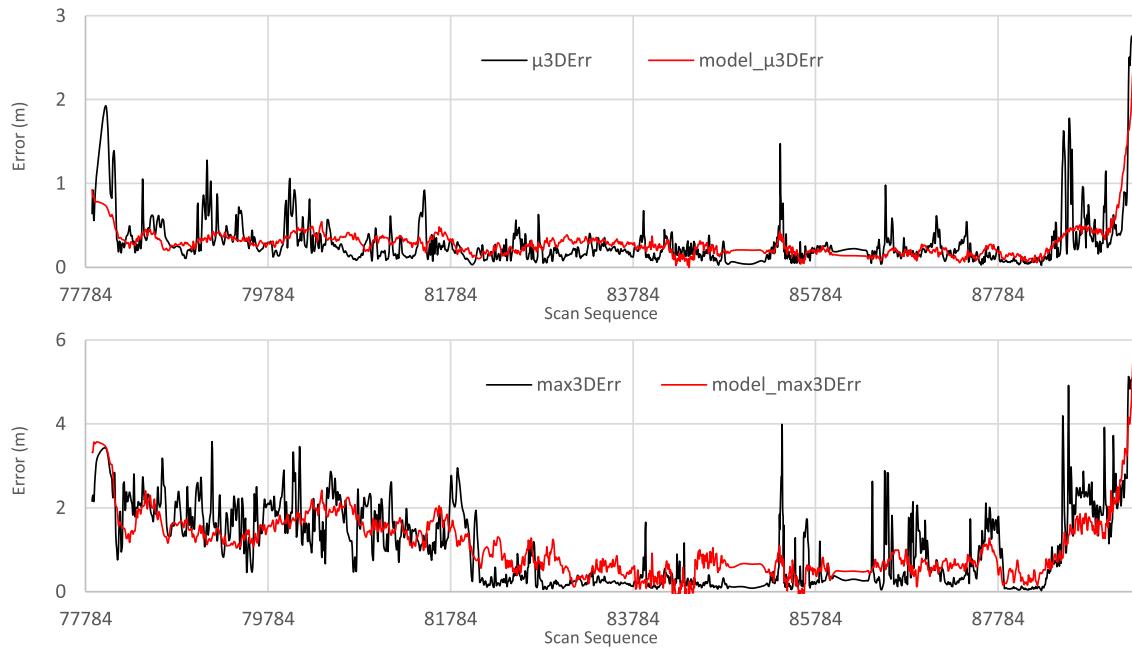
**FIGURE 19.** Factors related to the layout for path II (2.0m ND) compared to mean and max error of the same path.

Localization error model is obtained for different grid sizes and both mean and maximum error separately. As an instance, in order to model mean error of

localization within 5.0m ND map of path I, factors which are calculated based on a 5.0m grid in path I are used as explanatory variables. Generated model of mean and



**FIGURE 20.** Factors related to the representation quality and local similarity for path II (2.0m ND) compared to mean and max error of the same path.



**FIGURE 21.** Modeled mean and max error (red) for compared to actual error (black) for path I (5.0m grid).

max localization error for path I (5.0m ND) is shown in Figure 21.

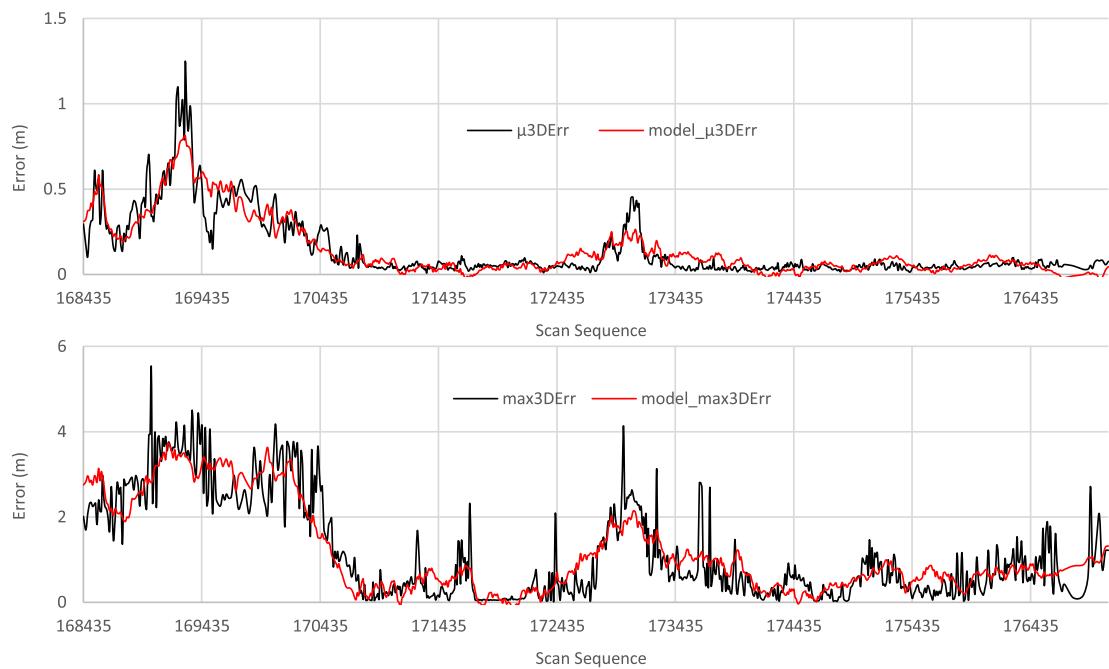
Figure 21 (top) shows the mean error model and bottom graph is for max error model. The black graph shows the actual error of the path, and red is the modeled error. As can be seen, introduced factors could model both mean and max error. In order to show how much the proposed factors could model the error, r-squared ( $R^2$ ) and root mean squared error (RMSE) are used.

These two parameters are usually used to evaluate a goodness of fit.  $R^2$  is ranges from 0 to 1, with 1 indicating

perfect prediction. For this specific case (Path I, 5.0m ND), RMSE and  $R^2$  for mean error are 0.206m and 0.551 respectively. Likewise, RMSE and  $R^2$  for max error are 0.607m and 0.6 respectively.

Figure 22 shows predicted error of localization within 2.0m ND map of path II. Top graphs show the actual (black) and predicted (red) mean error and bottom graphs show the actual (black) and predicted (red) max error of localization.

RMSE and  $R^2$  for mean is 0.07m and 0.828 respectively. Likewise, RMSE and  $R^2$  for mean is 0.52m and



**FIGURE 22.** Modeled mean and max error (red) for compared to actual error (black) for path II (2.0m grid).

**TABLE 3.** Goodness of fit statistics for modeled mean and max error.

|               | Path I           |       |                 |       | Path II          |       |                 |       |
|---------------|------------------|-------|-----------------|-------|------------------|-------|-----------------|-------|
|               | Mean error model |       | Max error model |       | Mean error model |       | Max error model |       |
| Grid size (m) | R <sup>2</sup>   | RMSE  | R <sup>2</sup>  | RMSE  | R <sup>2</sup>   | RMSE  | R <sup>2</sup>  | RMSE  |
| 2.0           | 0.770            | 0.09  | 0.401           | 0.77  | 0.828            | 0.07  | 0.788           | 0.52  |
| 3.0           | 0.789            | 0.09  | 0.600           | 0.56  | 0.538            | 0.08  | 0.724           | 0.4   |
| 4.0           | 0.757            | 0.12  | 0.612           | 0.56  | 0.450            | 0.11  | 0.677           | 0.45  |
| 5.0           | 0.551            | 0.206 | 0.600           | 0.607 | 0.376            | 0.214 | 0.648           | 0.473 |

R-squared and root mean squared error are shown for each path and different grid sizes.

0.788 respectively. Figure 22 shows that the factors could model both the mean and the max error well.

TABLE 3 shows the goodness of fit for other results. Best prediction is obtained for a 2.0m grid of Path II. And the worst prediction is for the same path with 5.0m grid size.

## VI. CONCLUSION

In this work, we have focused on the map, as one of the high potential sources of error in the map-matching based self-localization. By investigating the erroneous scenarios in the map and comparing their characteristics, we have introduced four criteria and requirements for the map to be able to perform self-localization with needed accuracy. These criteria are feature sufficiency, layout, local similarity, and representation quality.

Then in order to quantify these criteria, we have introduced several factors for each criteria. Unlike map criteria which are defined regardless of map format, factors have been proposed

based on map format of normal distribution transformation (NDT) matching. In NDT-based localization, map format is a normal distribution (ND). For feature sufficiency of the map, feature count, dimension count, dimension ratio, and occupancy ratio have been defined. For Layout, feature *DOP*, *r* average, angular entropy, and normal entropy have been proposed. Finally, for local similarity and representation quality, score entropy and Mahalanobis distance factor have been proposed respectively. Each of the map factors have been calculated for each position in the map, based on the map features within its local vicinity.

By conducting the experiments in Shinjuku, Japan, we have evaluated these four factors in different part of the map with different scenarios by comparing them with the self-localization error.

For evaluation, we have considered a mean and maximum error of localization. Experimental results have shown that most of the factors have a high correlation with the localization error. For instance, normal entropy factor which is used to formulate a layout criteria, has 0.836 and 0.844 correlation with mean and maximum error respectively. Among the factors, feature *DOP* and Mahalanobis distance have showed very low correlation with the errors. However, this does not necessarily show that these factors are not effective in order to evaluate the map. Because, some of the map criteria does not appear frequently, thus, corresponding factors has lower correlation to the error.

In order to model the localization error based on aforementioned factors, we have used Principal Component Regression. PCR have applied to a set of factors as explanatory variables and localization error as a response.

Experimental results have shown that, in the best case, the mean error can be modeled with RMSE and  $R^2$  of 0.07m and 0.828 respectively and max error with RMSE and  $R^2$  of 0.52m and 0.788 respectively. The result of this study can be applied to the dynamic determination of the abstraction ratio of the ND map and other formats of the map which be a potential future work of this study.

## REFERENCES

- [1] R. Yozevitch, B. Ben-Moshe, and A. Dvir, "GNSS accuracy improvement using rapid shadow transitions," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 3, pp. 1113–1122, Jun. 2014.
- [2] L. Heng, T. Walter, P. Enge, and G. X. Gao, "GNSS multipath and jamming mitigation using high-mask-angle antennas and multiple constellations," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 741–750, Apr. 2015.
- [3] U. Iqbal, J. Georgy, W. F. Abdelfatah, M. J. Korenberg, and A. Noureldin, "Pseudoranges error correction in partial GPS outages for a nonlinear tightly coupled integrated system," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1510–1525, Sep. 2013.
- [4] N. Alam, A. Kealy, and A. G. Dempster, "An INS-aided tight integration approach for relative positioning enhancement in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1992–1996, Dec. 2013.
- [5] S. Miura, L.-T. Hsu, F. Chen, and S. Kamijo, "GPS error correction with pseudorange evaluation using three-dimensional maps," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3104–3115, Dec. 2015.
- [6] M. Javanmardi, E. Javanmardi, Y. Gu, and S. Kamijo, "Towards high-definition 3D urban mapping: Road feature-based registration of mobile mapping systems and aerial imagery," *Remote Sens.*, vol. 9, no. 10, p. 975, Sep. 2017.
- [7] A. Hata and D. Wolf, "Road marking detection using LiDAR reflective intensity data and its application to vehicle localization," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 584–589.
- [8] S. Lagüela, I. Dorado, M. Gesto, P. Arias, D. González-Aguilera, and H. Lorenzo, "Behavior analysis of novel wearable indoor mapping system based on 3D-SLAM," *Sensors*, vol. 18, no. 3, p. 766, Mar. 2018.
- [9] J. Tang et al., "SLAM-aided stem mapping for forest inventory with small-footprint mobile LiDAR," *Forests*, vol. 6, no. 12, pp. 4588–4606, Dec. 2015.
- [10] O. Wulf, A. Nuchter, J. Hertzberg, and B. Wagner, "Ground truth evaluation of large urban 6D SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct./Nov. 2007, pp. 650–657.
- [11] F. Moosmann and C. Stiller, "Velodyne SLAM," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 393–398.
- [12] P. Biber, S. Fleck, and W. Strasser, "A probabilistic framework for robust and accurate matching of point clouds," in *Pattern Recognition*, C. E. Rasmussen, H. H. Bülthoff, B. Schölkopf, and M. A. Giese, Eds. Berlin, Germany: Springer, 2004, pp. 480–487.
- [13] H. Fu, L. Ye, R. Yu, and T. Wu, "An efficient scan-to-map matching approach for autonomous driving," in *Proc. IEEE Int. Conf. Mechatronics Automat.*, Aug. 2016, pp. 1649–1654.
- [14] T. Furukawa, L. Dantanarayana, J. Ziglar, R. Ranasinghe, and G. Dissanayake, "Fast global scan matching for high-speed vehicle navigation," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFIS)*, Sep. 2015, pp. 37–42.
- [15] J.-H. Im, S.-H. Im, and G.-I. Jee, "Vertical corner feature based precise vehicle localization using 3D LiDAR in urban area," *Sensors*, vol. 16, no. 8, p. 1268, Aug. 2016.
- [16] S. I. Oh and H.-B. Kang, "Fast occupancy grid filtering using grid cell clusters from LiDAR and stereo vision sensor data," *IEEE Sensors J.*, vol. 16, no. 19, pp. 7258–7266, Oct. 2016.
- [17] 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. Robot. Automat. (ICRA)*, May 2013, pp. 2233–2238.
- [18] M. Magnusson, "The three-dimensional normal-distributions transform—An efficient representation for registration, surface analysis, and loop detection," Ph.D. dissertation, School Sci. Technol., Center Appl. Auton. Sensor Syst., Örebro Univ., Örebro, Sweden, 2009.
- [19] P. Biber and W. Strasser, "The normal distributions transform: A new approach to laser scan matching," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, vol. 3, Oct. 2003, pp. 2743–2748.
- [20] A. Censi, "An ICP variant using a point-to-line metric," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2008, pp. 19–25.
- [21] O. Jež, "3D mapping and localization using leveled map accelerated ICP," in *Proc. Eur. Robot. Symp.*, H. Bruyninckx, L. Pfeučil, and M. Kulich, Eds. Berlin, Germany: Springer, 2008, pp. 343–353.
- [22] S. Hong, H. Ko, and J. Kim, "VICP: Velocity updating iterative closest point algorithm," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2010, pp. 1893–1898.
- [23] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. 3rd Int. Conf. 3-D Digit. Imag. Modeling*, May/Jun. 2001, pp. 145–152.
- [24] J. Serafin and G. Grisetti, "NCP: Dense normal based point cloud registration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep./Oct. 2015, pp. 742–749.
- [25] T. Furukawa et al., "Map-based navigation of an autonomous car using grid-based scan-to-map matching," in *Proc. 17th Int. Conf. Adv. Vehicle Technol.*, Aug. 2015, p. V003T01A005.
- [26] T. Kaminade, T. Takubo, Y. Mae, and T. Arai, "The generation of environmental map based on a NDT grid mapping: Proposal of convergence calculation corresponding to high resolution grid," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2008, pp. 1874–1879.
- [27] M. Rapp, M. Barjenbruch, M. Hahn, J. Dickmann, and K. Dietmayer, "Clustering improved grid map registration using the normal distribution transform," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun./Jul. 2015, pp. 249–254.
- [28] L. Cheng, Y. Wu, L. Tong, Y. Chen, and M. Li, "Hierarchical registration method for airborne and vehicle LiDAR point cloud," *Remote Sens.*, vol. 7, no. 10, pp. 13921–13944, Oct. 2015.
- [29] L. Yao et al., "Registration of vehicle-borne point clouds and panoramic images based on sensor constellations," *Sensors*, vol. 17, no. 4, p. 837, Apr. 2017.
- [30] H. González-Jorge, J. M. Sánchez, L. Díaz-Vilarín, I. Puente, and P. Arias, "Automatic registration of mobile LiDAR data using high-reflectivity traffic signs," *J. Construct. Eng. Manage.*, vol. 142, no. 8, p. 04016022, Aug. 2016.
- [31] P. Barrios, M. Adams, K. Leung, F. Inostroza, G. Naqvi, and M. E. Orchard, "Metrics for evaluating feature-based mapping performance," *IEEE Trans. Robot.*, vol. 33, no. 1, pp. 198–213, Feb. 2017.
- [32] N. Akai, L. Y. Morales, E. Takeuchi, Y. Yoshihara, and Y. Ninomiya, "Robust localization using 3D NDT scan matching with experimentally determined uncertainty and road marker matching," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 1356–1363.
- [33] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *J. Field Robot.*, vol. 24, no. 10, pp. 803–827, Oct. 2007.
- [34] E. Einhorn and H.-M. Gross, "Generic 2D/3D SLAM with NDT maps for lifelong application," in *Proc. Eur. Conf. Mobile Robots (ECMR)*, Sep. 2013, pp. 240–247.
- [35] L. Jun, L. Wei, D. Donglai, and S. Qiang, "Point cloud registration algorithm based on NDT with variable size voxel," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2015, pp. 3707–3712.
- [36] J. Saarinen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "Normal distributions transform Monte-Carlo localization (NDT-MCL)," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 382–389.
- [37] R. Cupec, E. K. Nyarko, D. Filko, A. Kitano, and I. Petrović, [Oct. 2013]. "Global localization based on 3D planar surface segments." [Online]. Available: <https://arxiv.org/abs/1310.0314>
- [38] K. Yoneda, C. Yang, S. Mita, T. Okuya, and K. Muto, "Urban road localization by using multiple layer map matching and line segment matching," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun./Jul. 2015, pp. 525–530.
- [39] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3D registration reliability and speed—A comparison of ICP and NDT," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2009, pp. 3907–3912.
- [40] L. Carbone, J. Du, M. K. Ng, B. Bona, and M. Indri, "Active SLAM and exploration with particle filters using Kullback–Leibler divergence," *J. Intell. Robot. Syst.*, vol. 75, no. 2, pp. 291–311, Aug. 2014.
- [41] R. Kümmel, B. Steder, C. Dornhege, A. Kleiner, G. Grisetti, and W. Burgard, "Large scale graph-based SLAM using aerial images as prior information," *Auton. Robots*, vol. 30, no. 1, pp. 25–39, Jan. 2011.
- [42] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time," in *Proc. Robot. Sci. Syst. Conf. (RSS)*, 2014, p. 9.

- [43] J. Choi, "Hybrid map-based SLAM using a Velodyne laser scanner," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 3082–3087.
- [44] D. Wang, H. Liang, T. Mei, H. Zhu, J. Fu, and X. Tao, "Lidar scan matching EKF-SLAM using the differential model of vehicle motion," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 908–912.
- [45] C. Tu, E. Takeuchi, C. Miyajima, and K. Takeda, "Compressing continuous point cloud data using image compression methods," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 1712–1719.
- [46] E. Javanmardi, M. Javanmardi, Y. Gu, and S. Kamijo, "Autonomous vehicle self-localization based on multilayer 2D vector map and multi-channel LiDAR," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 437–442.
- [47] O. Vysotska and C. Stachniss, "Improving SLAM by exploiting building information from publicly available maps and localization priors," *PFG—J. Photogramm., Remote Sens. Geoinf. Sci.*, vol. 85, no. 1, pp. 53–65, Feb. 2017.
- [48] H. Merzić, E. Stumm, M. Dymczyk, R. Siegwart, and I. Gilitschenski, "Map quality evaluation for visual localization," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May/Jun. 2017, pp. 3200–3206.
- [49] *Efficient DOP Calculation for GPS With and Without Altimeter Aiding | the Journal of Navigation | Cambridge Core*. Accessed: Feb. 22, 2018. [Online]. Available: <https://www.cambridge.org/core/journals/journal-of-navigation/article/efficient-dop-calculation-for-gps-with-and-without-altimeter-aiding/8FD31C94F27E922CD1401FD3EEA5BE70>
- [50] A. Gressin, C. Mallet, and N. David, "Improving 3D LIDAR point cloud registration using optimal neighborhood knowledge," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. I-3, pp. 111–116, Jul. 2012.



**EHSAN JAVANMARDI** (M'16) received the M.E. degree in computer architecture from the Amirkabir University of Technology, Iran, in 2012. He is currently pursuing the Ph.D. degree in information and communication engineering with The University of Tokyo, Tokyo, Japan. From 2014 to 2015, he was a Visiting Research Student with The University of Tokyo.

He was a Visiting Student Researcher at the University of California at Berkeley from 2016 to 2017. His research interests include the intelligent vehicles, autonomous vehicle's self-localization, mobile mapping systems, ADAS map, mapping, sensor fusion, and vehicle perception. He is a member of the IEEE Intelligent Transportation Systems Society.

Mr. Javanmardi was a recipient of the IEEE Intelligent Transportation Systems Society Best Student Paper Award in 2017.



**MAHDI JAVANMARDI** (M'16) received the M.Sc. degree from the Sharif University of Technology, Iran, in 2013, and the Ph.D. degree in information and communication engineering from The University of Tokyo, Japan, in 2017.

He was a Visiting Student Researcher at the University of California at Berkeley from 2016 to 2017. He has also completed the Graduate Program for Social ICT Global Creative Leaders at The University of Tokyo in 2017. He has been a Post-Doctoral Researcher at the Institute of Industrial Science, The University of Tokyo, since 2017. His research interests include localization and mapping for autonomous vehicle, autonomous vehicle's perception, and computer vision. He is a member of the IEEE ITS Society.



**YANLEI GU** (M'14) received the M.E. degree from the Harbin University of Science and Technology, China, in 2008, and the Ph.D. degree from Nagoya University, Japan, in 2012.

He has been a Post-Doctoral Researcher at the Institute of Industrial Science, The University of Tokyo, since 2013. His research interests include GNSS positioning, computer vision, and deep learning and their applications to ITS.

He is a member of the IEEE ITS Society. He has served as the organizing committee member for the IEEE ICVES2015 and ITSC2017.



**SHUNSUKE KAMIJO** (M'97) received the B.S. and M.S. degrees in physics and the Ph.D. degree in information engineering from The University of Tokyo, Tokyo, Japan, in 1990, 1992, and 2001, respectively. In 1992, he joined Fujitsu Ltd., as a Processor Design Engineer.

From 2001 to 2002, he was an Assistant Professor with the Department of Information and Communication Engineering, The University of Tokyo, where he has been an Associate Professor since 2002. His research interests are computer vision, wireless communication, and their applications to ITS. His research focuses are autonomous vehicles, traffic video surveillance, traffic signal control, V2X communications, pedestrian, and car navigations.

He is a member of the IEEE ITS Society, TRB, IEICE, and IATSS. He joined the International Program Committee of ITS World Congress in 2011. He has been a member of the Board of Governors of the IEEE ITS Society since 2015 and has been an Executive Committee Member of the Society since 2017. He has served as the Vice Chairman of the Program Committee for the ITS World Congress Tokyo 2013, the General Co-Chair for the IEEE ICVES2015 and ITSC2017, and the International Program Chair for the IEEE IV2017. He is an Editorial Board Member of the *International Journal of Intelligent Transportation Systems Research* (Springer) and *Multimedia Tools and Applications* (Springer).