Uncertainty estimation of LiDAR matching aided by dynamic vehicle detection and high definition map

W. Wen, X. Bai, W. Zhan, M. Tomizuka and L.-T. Hsu✉

LiDAR matching between real-time point clouds and pre-built points map is a popular approach to provide accurate localisation service for autonomous vehicles. However, the performance is severely deteriorated in dense traffic scenes. Unavoidably, dynamic vehicles introduce additional uncertainty to the matching result. The main cause is that the pre-built map can be blocked by the surrounding dynamic vehicles from the view of LiDAR of ego vehicle. A novel uncertainty of LiDAR matching (ULM) estimation method aided by the dynamic vehicle (DV) detection and high definition map is proposed in this Letter. Compared to the conventional Hessian matrix-based ULM estimation approach, the proposed method innovatively estimates the ULM by modelling surrounding DV. Then the authors propose to correlate the ULM with the detected DV and convergence feature of matching algorithm. From the evaluated real-data in an intersection area with dense traffic, the proposed method has exhibited the feasibility of estimating the ULM accurately.

Introduction: LiDAR is a prevalent sensor for providing autonomous vehicle (AV) [1] localisation service, using matching the real-time point clouds with pre-built points map [2]. The uncertainty of LiDAR matching (ULM) can be well estimated in traffic friendly scenarios. However, excessive dynamic vehicles that are not included in the pre-built point cloud map can cause occlusion of points map. This occlusion increases the difficulty of ULM estimation. Regarding the current ULM estimation methods, the literature review shows that approaches mainly include the environment feature-based [3] and Hessian matrix-based [4]. The main disadvantage of the previous researches is their assumption on modelling the all the point clouds as static ones during the matching process. As a result, this can introduce erroneous ULM estimation in scenarios with dense traffic. The other stream of ULM estimation is to model the existence of dynamic objects by leveraging conditional observation model (COM) [5]. However, the high computational cost is excessive. Accordingly, this Letter presents a novel dynamic vehicle (DV) detection aided approach to estimate the ULM in the dense traffic scene. Fig. 1 shows a case with six vehicles (marked as a green rectangle) surrounding the ego vehicle (marked as a purple rectangle). As shown in Fig. 1, α represents the masking elevation angle (MEA) of the line connecting LiDAR (atop the ego-vehicle) and roof of the dynamic objects. Essentially, MEA can be a clue of ULM estimation. The main reason for this is that many of the points in the pre-built points map can be occluded by the DV from the view of LiDAR of ego vehicle. Besides, the geometry distribution of surrounding vehicles relative to ego vehicle can also affect the ULM. Inspired by this, we propose to detect the car roof points that have the largest MEA (α_i) to represent the occlusion at each azimuth angle θ_i. Then the ULM is correlated with three components: (i) The mean MEA across all the azimuth angles. (ii) The geometry distribution is of DV relative to the ego-vehicle. (iii) The convergence feature of the matching algorithm (matching between the real-time point clouds with pre-built points map). The proposed method can accurately estimate the ULM with a small standard deviation. Real-time inference speed is obtained, which betters the superiority that our method generates. The listed strengths imply undoubted popularity of the proposed method in ULM estimation for AV.

DV detection using real-time 3D point clouds: The objective of vehicle detection [6] is to categorise and locate road objects of interests. However, only the occlusions caused by road DV in each azimuth angle are needed and significant in this Letter. Thus, we propose to detect the DV via two steps: (i) Refine the 3D point cloud of interest (point clouds inside the road) via curbs information in high definition (HD) map (shown in Fig. 2). (ii) Detect the highest MEA (α_i) at each azimuth angle θ_i. The received 3D point clouds can be represented as $P_{\mathrm{raw}} = \{p_1, p_2, \ldots, p_i, \ldots p_n, t\}$ at a given time t, where $p_i = (x_i, y_i, z_i)$ represents a single point in the coordinate system of LiDAR. Then, the refined point clouds are denoted as $P_r = \{p_1, p_2, \ldots, p_i, \ldots p_m, t\}$. Accordingly, we transform the point clouds P_r to S_{sp} in the spherical coordinate system via Algorithm 1. S_{sp} is indicated by $S_{\mathrm{sp}} = \{s_1, s_2, \ldots, s_i, \ldots s_k, t\}$ at a given time t, where $s_i = (\theta_i, \alpha_i, D_i)$ represents a point. θ_i and α_i indicate the azimuth angle and MEA, respectively. D_i indicates the Euclidean distance

from a point to LiDAR. Points can have overlap across a certain azimuth angle. We proposed to obtain the largest MEAs across all the azimuth angles. The input of Algorithm 1 is the point clouds P_{raw} and high definition (HD) map. The outputs are the MEA in each azimuth angle and the Euclidean distance between the point and the centre of LiDAR. The resolution of the azimuth angle is $1°$ in Algorithm 1.
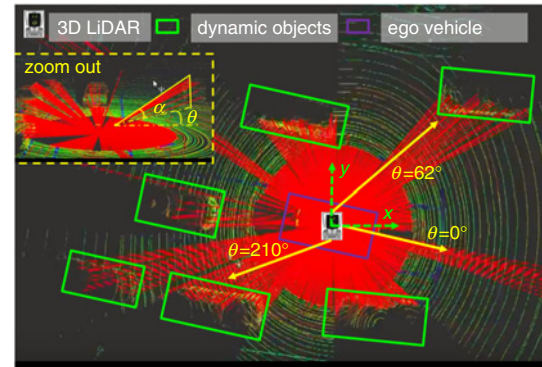

Fig. 1 *Illustration of traffic detection. The green points represent the real-time point clouds from 3D LiDAR. The red line represents the connection between the 3D LiDAR and detected roof (S_{sp}) of DVs. θ and α indicate the azimuth angle and MEA, respectively*
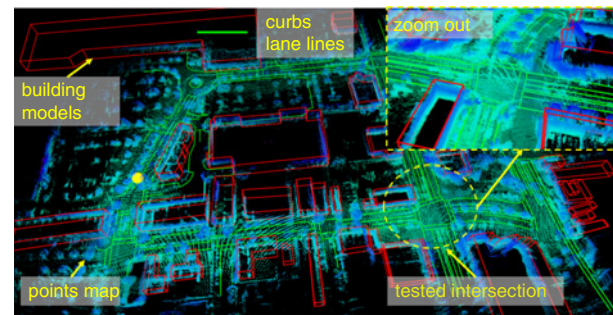

Fig. 2 *Demonstration of the high definition map. The red boxes indicate the 3D building models. Coloured points indicate the points map. The green curves represent the curbs and lane lines. The circled area is the tested intersection. The curbs information is used to refine the point clouds of interest*

Algorithm 1: MEAs identification

Input: Point clouds $P_{\mathrm{raw}} = \{p_1, p_2, \ldots, p_i, \ldots, p_n, t\}$, HD maps M
Output: $S_{\mathrm{sp}} = \{s_1, s_2, \ldots, s_i, \ldots, s_k, t\}$, where $k = 360$
1. Initialise S_{sp}, set $w = 0$
2. $P_r = refine(P_{\mathrm{raw}}, M)$
3. **while** ($w \leq 360°$)
4. $w = w + 1$
5.   **for** all point p_i in P_{sp} do
6. $D_i = \sqrt{(x_i^2 + y_i^2 + z_i^2)}$
7. $azi_i = \mathrm{atan}(y_i/x_i)$//azimuth angle
8. $ele_i = \mathrm{atan}\left(\dfrac{z_i}{\sqrt{(x_i^2 + y_i^2)}}\right)$ //mask elevation angle
9.   **if** ($ele_i \geq s_w(\alpha_w)$)
10. $s_w \leftarrow (azi_i, ele_i, D_i)$
11.   **end if**
12. **end for** P_{sp}
13. **end while**

Based on Algorithm 1, the MEAs are saved in S_{Polar}. Higher MEA can contribute more on the ULM as more of the pre-built point cloud map blocked by the DVs. Therefore, we propose to calculate the mean elevation mask angle as a quantitative indicator of occlusion as follows:

$$\beth_{\mathrm{ele}} = \sum_{i=1}^{k} s_i(\alpha_i)/k \tag{1}$$

where $s_i(\cdot)$ is an operator to get the MEA α_i. Each s_i indicates the detected points belonging to DVs. In this case, occlusion caused by DVs regarding the ego-vehicle is modelled.

Occlusion-based ULM estimation: As illustrated above, the quantity $\beth_{\text{ele}}$ is an indicator of occlusion from DVs. Thus, we propose to model the total position uncertainty u_{traffic} of LiDAR matching as follows:

$$u_{\text{traffic}} = K_{\text{scale}} \cdot \beth_{\text{ele}} \cdot \delta \qquad (2)$$

K_{scale} denotes the scaling factor that will be tuned heuristically. δ is the summation [7] of errors between the original point clouds and matched point clouds after the matching process. The u_{traffic} denotes the total ULM. The geometry distribution of DVs across the different azimuth angles can contribute differently in the ULM in the x-, y- and z-axis of LiDAR coordinate. Geometric dilution of precision (GDOP) is an effective parameter to model the geometry distribution of satellite in GNSS navigation. Drawing the inspiration from this, we obtain the geometry distribution matrix $T(3 \times 3)$ of DVs relative to the ego-vehicle as

$$T = \left(G^T G\right)^{-1} \qquad (3)$$

where G denotes the unit line of sight (LOS) vector between 3D LiDAR and points in S_{sp}. The matrix $G(k \times 3)$ can be calculated as

$$G = \begin{bmatrix} \cos \alpha_1 \cos \theta_1 & \cos \alpha_1 \sin \theta_1 & \sin \alpha_1 \\ \cdots & \cdots & \cdots \\ \cos \alpha_i \cos \theta_i & \cos \alpha_i \sin \theta_i & \sin \alpha_i \\ \cdots & \cdots & \cdots \\ \cos \alpha_k \cos \theta_k & \cos \alpha_k \sin \theta_k & \sin \alpha_k \end{bmatrix} \qquad (4)$$

where α_i and θ_i indicate the MEA and azimuth angles of points, respectively. Then the ULM in three different directions (u_{traffic}^x, u_{traffic}^y and u_{traffic}^z) can be calculated as

$$u_{\text{traffic}}^x = u_{\text{traffic}} \frac{T_{11}}{\sqrt{\left(T_{11}^2 + T_{22}^2 + T_{33}^2\right)}} \qquad (5)$$

$$u_{\text{traffic}}^y = u_{\text{traffic}} \frac{T_{22}}{\sqrt{\left(T_{11}^2 + T_{22}^2 + T_{33}^2\right)}} \qquad (6)$$

$$u_{\text{traffic}}^z = u_{\text{traffic}} \frac{T_{33}}{\sqrt{\left(T_{11}^2 + T_{22}^2 + T_{33}^2\right)}} \qquad (7)$$

Experimental results:: The performance of the proposed ULM estimation method is evaluated in an intersection area with dense traffic in Berkeley, California, USA (shown in Fig. 2). The HD map of the test area is generated beforehand. During the test, the Velodyne 64 LiDAR is employed to capture real-time point clouds. The matching between real-time point cloud and pre-built points map is implemented based on the normal distribution transform [8]. The differential GNSS is employed to provide the ground truth of localisation. The actual positioning error (summation of positioning error in the x, y and z directions) of matching is the ground truth of uncertainty estimation, which is the value that u_{traffic} aims to approach.
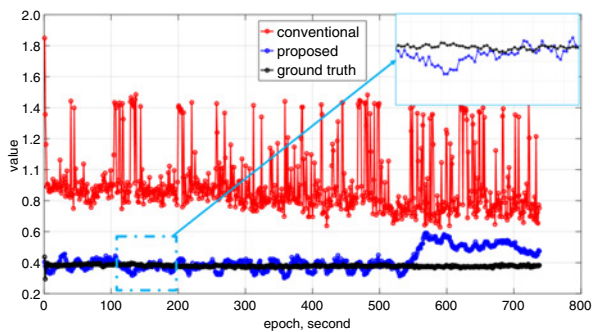

Fig. 3 *Comparison of the conventional [4] and proposed the ULM estimation method*

The result of ULM estimation is shown in Fig. 3. The black curve and red curve indicate the ground truth and conventional Hessian matrix-based method of ULM estimation, respectively. We can see from Fig. 3 that the ground truth (black curve) for ULM is about 0.4 m

during the 730 epochs. However, the Hessian matrix-based method (red curve) overestimates the uncertainty all through the test with a mean of about 0.85 m. More importantly, the ULM estimation fluctuates dramatically over the test. With the aid of the proposed ULM estimation method, the estimated uncertainty (blue curve) can well track the actual positioning error with smaller fluctuation. The uncertainty estimation errors are listed and compared in Table 1. It is shown that the accuracy of the proposed method is 0.10 m with the small standard deviation (0.12 m), while it is 0.51 m of uncertainty estimation error using the conventional Hessian matrix-based method [4]. More importantly, the standard deviation decreases from 0.45 to 0.12 m. This result shows that the proposed method can effectively track the trend of ULM.

Table 1: Uncertainty estimation errors of LiDAR matching using the conventional Hessian matrix-based and the proposed methods

Results	Hessian matrix-based method [7]	Proposed method (u_{traffic})
mean error	0.51 m	0.10 m
std	0.45 m	0.12 m

Conclusion: An ULM estimation method is presented in this Letter. The results in Table 1 show that the proposed method can effectively model the uncertainty caused by the surrounding DVs. The proposed method outperforms the traditional Hessian matrix-based method since the proposed method can effectively capture and model the effects of DVs with a small standard deviation. Centimetre-level positioning is required for AV, accurate ULM estimation is of equal importance for safe AV. The experimental results illustrate the effectiveness of the proposed method with the uncertainty estimation error of 0.1 m and the standard deviation of 0.12 m which is significantly smaller than the Hessian matrix-based method.

W. Wen, W. Zhan and M. Tomizuka (*Department of Mechanical Engineering, University of California at Berkeley, Berkeley, CA, USA*)

X. Bai and L.-T. Hsu (*Interdisciplinary Division of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*)

✉ E-mail: lt.hsu@polyu.edu.hk

W. Wen: Also with Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong

References

1 Campbell, M., Egerstedt, M., How, J.P., *et al.*: 'Autonomous driving in urban environments: approaches, lessons and challenges', *Philos. Trans. R. Soc. Lond. A, Math. Phys. Eng. Sci.*, 2010, **368**, (1928), pp. 4649–4672
2 Wan, G., Yang, X., Cai, R., *et al.*: 'Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes', arXiv preprint arXiv:1711.05805, 2017
3 Shetty, A.P.: 'Gps-Lidar Sensor Fusion Aided by 3d City Models for Uavs', 2017
4 Akai, N., Morales, L.Y., Hirayama, T., *et al.*: 'Toward Localization-Based Automated Driving in Highly Dynamic Environments: Comparison and Discussion of Observation Models', The 21st IEEE International Conference on Intelligent Transportation Systems, Maui, HI, USA, November 2017
5 Akai, N., Morales, L.Y., and Murase, H.: 'Mobile robot localization considering class of sensor observations'. Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), Madrid, Spain, October 2018
6 Wu, B., Wan, A., Yue, X., *et al.*: 'Squeezeseg: convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud'. 2018 IEEE Int. Conf. on Robotics and Automation (ICRA), 2018
7 Akai, N., Morales, L.Y., Takeuchi, E., *et al.*: 'Robust localization using 3d Ndt scan matching with experimentally determined uncertainty and road marker matching'. Intelligent Vehicles Symp. (IV), 2017, 2017
8 Merten, H.: 'The three-dimensional normal-distributions transform', *Threshold*, 2008, **10**, p. 3