

ORIGINAL RESEARCH

Low-cost solid-state LiDAR/inertial-based localization with prior map for autonomous systems in urban scenarios

Yihan Zhong¹ | Feng Huang² | Jiachen Zhang² | Weisong Wen^{2,3} | Li-Ta Hsu²

¹Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

²Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

³The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China

Correspondence

Weisong Wen, Room PQ408, The Hong Kong Polytechnic University, Hong Kong
Email: welson.wen@polyu.edu.hk

Funding information

Guangdong Basic and Applied Basic Research Foundation, Grant/Award Number: 2021A1515110771; Riemann Laboratory, Huawei Technologies, Grant/Award Number: ZGD2; University Grants Committee of Hong Kong under the scheme Research Impact Fund, Grant/Award Number: R5009-21

Abstract

A low-cost and accurate positioning solution is significant for the massive deployment of fully autonomous driving vehicles (ADV). Conventional mechanical LiDAR has proven its performance, but its high cost hinders the massive production of autonomous vehicles. This paper proposes a low-cost LiDAR/inertial-based localization solution for autonomous systems with prior maps in urban areas. Instead of relying on the costly mechanical LiDAR, this paper proposes to utilize the solid-state LiDAR (SSL) with the prior map to estimate the position of the vehicle by matching the real-time point clouds from the SSL and the prior map using the normal distribution transformation (NDT) algorithm. However, the field of view (FOV) of the SSL is significantly smaller than the conventional mechanical LiDAR, which can easily lead to failure during the NDT map matching. To fill this gap, this paper proposes to exploit the complementarity of the inertial measurement unit (IMU) and the SSL, where the IMU pre-integration provides a coarse but high-frequency initial guess to the map matching. To evaluate the effectiveness of the proposed method in this paper, the authors collect the dataset in two typical urban scenarios through a pedestrian hand-hold and a vehicle driving condition. The results reveal that the SSL-only-based localization is significantly challenged in dynamic scenarios. With the help of the IMU, the robustness of the proposed method is significantly improved, achieving an accuracy of within 0.5 m. To show the sensitivity of the SSL-based map matching against the initial guess of the state, this paper also presents the convergence results of the map matching under different levels of accuracy in terms of the initial guess.

1 | INTRODUCTION

Autonomous driving vehicles (ADV) [1–4] have recently received increasing attention due to their high potential to improve dense traffic congestion and unexpected traffic accidents. Accurate and globally referenced positioning is one of the keys to its realization. Global navigation satellite systems (GNSS) can provide high accuracy positioning for autonomous vehicles in open areas up to 5 m [2, 5]. GNSS positioning errors can rise significantly in urban canyons due to multipath effects and non-line-of-sight (NLOS) [6, 7]. To fill this gap, the LiDAR matching-based localization [3, 4, 8] with the prior map is a promising solution to provide the globally referenced and accurate positioning. Currently, the existing LiDAR map matching-based localization solution mainly relies on the 360°

rotating mechanical LiDAR, such as the 64 channels Velodyne HDL 64 (about US\$75,000) [9]. Specifically, the map matching-based localization is achieved by associating the real-time 3D point clouds from LiDAR with the globally referenced prior map. However, the high cost of the mechanical rotating LiDAR was one of the major factors preventing its massive deployment in autonomous vehicles. Due to its cost-effectiveness, the recently developed solid-state LiDAR (SSL) (about US \$800) opens a new window for LiDAR matching-based localization [10]. However, the utilization of the SSL has a major drawback as the field of view (FOV) of the SSL is limited and significantly smaller than the one from conventional mechanical 3D LiDAR. A comparison of the FOV for the SSL and 16 channels mechanical LiDAR is shown in Figure 1. The mechanical LiDAR can capture the features from 360° of FOV, but only the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *IET Intelligent Transport Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

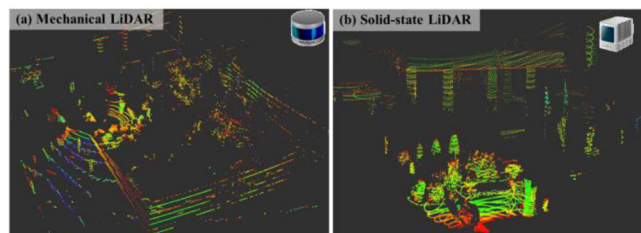


FIGURE 1 Comparison of the FOVs of conventional mechanical LiDAR versus solid-state LiDAR. (a) The 3D point clouds from the VLP-16 (about US\$4000). (b) Livox-Horizon (US\$800). FOV, field of view

front features are perceived via the SSL. As a result, the failure in SSL-based map matching occurs [11].

To exploit the potential of the SSL and relax the problems of smaller FOV, this paper proposes an SSL/inertial integrated localization solution with a prior map in urban canyons. The contributions of this paper are mainly threefold:

- a. This paper proposes an SSL-based LiDAR localization method for autonomous driving vehicle localization in urban canyon scenarios. Unlike the conventional Livox-based LiDAR SLAM methods that can only provide relative pose estimation and are subject to drift over time, this paper exploits the potential of the low-cost SSL-based drift-free localization with the prior map significant implications for reducing the cost of future self-driving vehicles.
- b. To alleviate the drawback of the small FOV of the SSL, this paper improves the SSL matching performance in dynamic scenarios by complementarily fusing a low-cost inertial measurement unit (IMU) sensor, which provides a better initial guess of the LiDAR matching.
- c. This paper verifies the effectiveness of the contributions of this paper step by step using hand-held and driving datasets. Meanwhile, we analyze the sensitivity of the initial pose against the LiDAR matching. We also discuss the effect of different grid sizes and initial guesses on the NDT-based LiDAR matching algorithm in the case of using the new SSL.

The rest of this paper is structured as follows: First, the related work will be presented in Section 2 before the overview of the proposed method is introduced in Section 3. Then, the proposed localization system based on LiDAR/inertial/prior map integration is presented in Section 4. Finally, the experimental verification is presented in Section 5 before the conclusion is illustrated in Section 6.

2 | RELATED WORKS

Scholarly works on localization for ADV are extensive and any attempts to give a full relevant review would be incomplete. In this connection, this section focuses on LiDAR matching-based localization and its advancement. The most pioneer work for LiDAR matching-based localization was conducted by the team from Stanford University [12], where the reflectivity-based

matching algorithm is developed to achieve accurate positioning. The work was improved [3] by integrating more sensors. However, only the 2D localization result was derived with the probabilistic map where the 3D geometry is not fully explored. As a result, the method can be sensitive to the change in road textures. To fill this gap, work in [13] developed a multi-resolution-based probabilistic map to increase the robustness of the system against unexpected changes in the road textures. However, these methods failed to fully explore the potential of the 3D pose estimation from LiDAR matching. Another map matching-based localization was based on geometry-based data association which could provide 3D pose estimation. Its primary principle of LiDAR matching is to match the point clouds collected from the 3D LiDAR and the prior map. The iterative closest point (ICP) [14] is the typical method to estimate the relative motion of two overlapping point clouds, which is the most classical matching method. Due to the sensitivity of the ICP on the local minima phenomenon, improvements are studied in [15–17], one of which is G-ICP [18], which uses the geometry distributions via the Gaussian modelling, resulting in a significant improvement in accuracy over the original ICP algorithm [18]. However, the nearest-neighbour search imposes a heavy computational load on the ICP algorithm, making the system unable to perform real-time matching when the number of point clouds is vast, which is a fatal flaw for autonomous vehicle systems [18]. V-GICP is a further improvement of the G-ICP algorithm [19]. It combines normal distribution transformation (NDT) to voxelized the point cloud and then estimates the voxel distribution. The experimental results show that V-GICP reduces the computational load of the system while ensuring accuracy with the G-ICP algorithm. Nevertheless, the various G-ICP algorithms cannot meet the requirements for system real-time in real scenarios [20, 21]. Therefore, the bottleneck of G-ICP algorithms and their variants is how to guarantee the matching accuracy and at the same time significantly reduce the computational load [22]. To fill this gap, the NDT algorithm was developed for efficient map matching [22, 23]. This was achieved by dividing the dense raw point clouds into representative voxels which guaranteed computational efficiency. Moreover, the modelling of the voxel representation enabled the exploration of the geometry features which increases the accuracy of the algorithm in terms of the data association. According to a recent evaluation in [24], better positioning accuracy and robustness are obtained based on the assessment using typical autonomous driving datasets collected by the 360° rotating mechanical LiDAR. As a result, the NDT-based map matching was employed in a wide range of autonomous driving applications [25–28]. Unfortunately, all these LiDAR matching-based localization mainly relied on the mechanical 3D LiDAR which is too expensive for massive deployment in ADVs. This is one of the key issues preventing the arrival of the ADV.

LOAM-Livox [29] utilizes the odometry from directly registering a new scan to a global map to deal with the small FOV of SSL. To reduce the cost of the localization system, the SSL-SLAM [30] proposed a novel lightweight SLAM framework for low-cost Realsense L515 SSL. High-rate IMU measurements can compensate for LiDAR distortion and increase the robust-

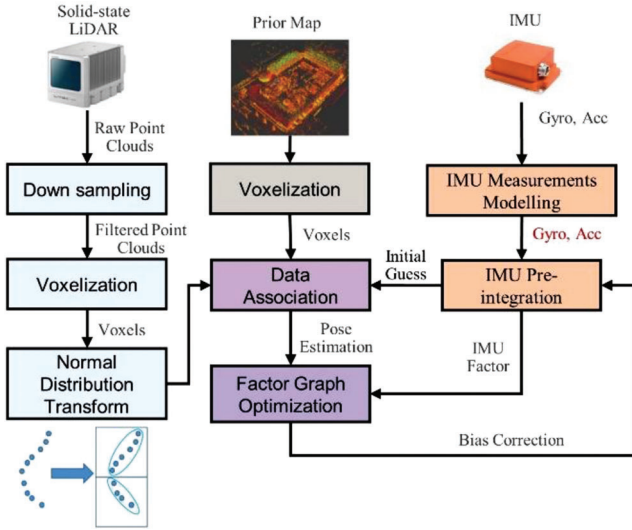


FIGURE 2 Overview of the proposed algorithm. The main inputs are the point cloud data collected by the LiDAR, the prior map, and IMU data. The output is the pose estimation of the system. IMU, inertial measurement unit

ness and accuracy of the LiDAR odometry by providing good initial guesses. The LiLi-OM [30] integrated the SSL point cloud and the IMU measurements by employing the sliding windows FGO in the system. Instead of using FGO, the FAST-LIO [31] adopted a tightly coupled iterated Kalman filter to fuse the measurements from LiDAR and IMU. As an extension of FAST-LIO, FAST-LIO2[32] achieved high accuracy with lower computational load by integrating an incremental k-d tree and direct raw point cloud registering into FAST-LIO. Taking the advantages of camera sensors, R²LIVE [33] fused the camera measurements and the LiDAR-inertial system shows its sufficient robustness in various environments. However, all these works focus on achieving low-drift relative motion estimation using the onboard SSL LiDAR, IMU, and camera, where the estimated pose is subjected to drift over time.

3 | OVERVIEW OF THE PROPOSED METHOD

3.1 | System overview

The overview of the proposed system is shown in Figure 2. The input of the system contains three parts: one is 3D point cloud data collected using low-cost SSL, the others are the measurement of the accelerators and gyroscopes in IMU, and the last one is the 3D point cloud from the prior map. High-frequency IMU measurements are obtained between two consecutive frames. To reduce the computational loads, the IMU pre-integration [34] technique is employed to integrate multiple IMU measurements into a pre-integration term. The IMU pre-integration provides high-frequency pose estimation (~ 200 Hz), serving as the initial guess of the LiDAR matching (~ 10 Hz). Meanwhile, the estimated pose from the LiDAR matching will be integrated with the IMU pre-integration term to correct the

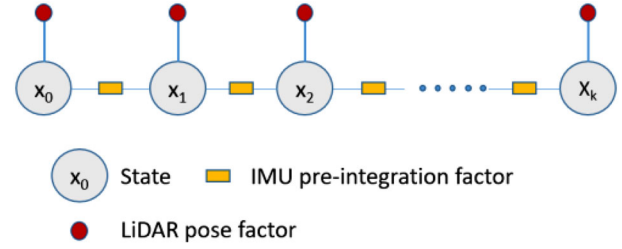


FIGURE 3 Factor graph structure of the proposed integration scheme

bias of IMU. To make the presentation clear, in this paper, matrices are denoted in uppercase with bold letters. Vectors are denoted in lowercase with bold letters. Variable scalars are represented as lowercase italic letters. Constant scalars are denoted as lowercase letters. The notations and frames used in the whole paper are defined as follows for clarifying the proposed pipeline.

4 | LIDAR/INERTIAL/PRIOR MAP INTEGRATION

4.1 | Factor graph structure of the proposed localization system

As shown in Figure 2, the key of the proposed system is the integration of the pose estimations from IMU pre-integration and the LiDAR scan matching based on the prior map. To achieve this, this paper adopts factor graph optimization (FGO) to integrate the two pieces of information. Figure 3 illustrates the factor graph of the integration. The grey circle represents the state of the IMU frame. Two types of factors constrain the states, one is the LiDAR pose factor and another one is the IMU pre-integration factor.

In this paper, the IMU body frame is represented as $\{\cdot\}^B$. The LiDAR frame is represented as $\{\cdot\}^S$. $\{\cdot\}^W$ represents the world frame which refers to the reference point of the prior point cloud map. The IMU state at k_{th} epoch can be expressed as

$$\chi \in \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n\} \quad (1)$$

with

$$\mathbf{x}_k = \left[\mathbf{p}_{B_k}^W, \mathbf{v}_{B_k}^W, \mathbf{q}_{B_k}^W, \mathbf{b}_{a,k}, \mathbf{b}_{g,k} \right] \quad (2)$$

where the subscript B_k is the IMU body frame while the k_{th} stage point cloud is obtained. The $\mathbf{p}_{B_k}^W$ denotes position and $\mathbf{v}_{B_k}^W$ represents velocity. The orientation is represented as $\mathbf{q}_{B_k}^W$ in the world frame. $\mathbf{b}_{a,k}$ and $\mathbf{b}_{g,k}$ denote acceleration and gyroscope bias, respectively. The transformation matrix of IMU concerning the world frame at the k_{th} epoch can be represented as $\mathbf{T}_{B_k}^W \in SE(3)$

$$\mathbf{T}_{B_k}^W = \begin{bmatrix} \mathbf{R}_{B_k}^W & \mathbf{p}_{B_k}^W \\ 0 & 1 \end{bmatrix} \quad (3)$$

where $\mathbf{R}_{B_k}^W \in SE(3)$ represents the rotation that corresponds to $\mathbf{q}_{B_k}^W$. The LiDAR point cloud at the k th frame is \mathcal{P}^{S_k} . Note that the extrinsic parameter between the IMU and then LiDAR which is denoted by the \mathbf{T}_B^S is calibrated beforehand [35]. Therefore, the transformation $\mathbf{T}_{B_k}^W$ can be defined which would be used in the processing of modelling of IMU measurements as follows:

$$\mathbf{T}_{B_k}^W = \mathbf{T}_B^S \mathbf{T}_{B_k}^{B_0} \quad (4)$$

4.2 | IMU pre-integration factor

This section presents the modelling of the IMU pre-integration factor. IMU measures the rotation rate and acceleration of the sensor itself concerning its body frame using accelerometers and gyroscopes.

The raw measurements include acceleration $\hat{\mathbf{a}}^B$ and angular velocity $\hat{\mathbf{v}}^B$ in IMU frame B . \mathbf{b}_g represents the bias of $\hat{\mathbf{v}}^B$, \mathbf{n}_ω denotes the additive noise of the acceleration of $\hat{\mathbf{v}}^B$. \mathbf{b}_a represents the bias of $\hat{\mathbf{a}}^B$, \mathbf{n}_a denotes the additive noise of the acceleration of $\hat{\mathbf{a}}^B$. Taking into account the measurement noise \mathbf{n} and bias \mathbf{b} , the measurement of IMU is defined as follows:

$$\hat{\omega}^B = \omega^B + \mathbf{b}_g + \mathbf{n}_\omega \quad (5)$$

$$\hat{\mathbf{a}}^B = \mathbf{R}^{BW} (\mathbf{a}^W - \mathbf{g}^W) + \mathbf{b}_a + \mathbf{n}_a \quad (6)$$

\mathbf{R}^{BW} is the rotation matrix from the world frame to the IMU body frame. \mathbf{g}^W is the constant gravity vector in the world frame. \mathbf{a}^W represents the noise-free acceleration of the system in the world frame.

Typically, the frequency of LiDAR is much smaller than the frequency of IMU, so there are many IMU measurements in the time it takes to receive two frames of LiDAR point clouds. IMU pre-integration accumulates a large number of IMU measurements and models them as a single relative motion constraint. Such processing avoids repeated calculations for the same pose and reduces the computational load on the system. The t and $t + 1$ are assumed to be two consecutive time instants between B_k and B_{k+1} . The angular velocity ω and the acceleration a between t and $t + \Delta t$ can be expressed as

$$\omega = \frac{1}{2} \left((\hat{\omega}^{B_t} - \mathbf{b}_{g_k}) + (\hat{\omega}^{B_{t+1}} - \mathbf{b}_{g_k}) \right) \quad (7)$$

$$\mathbf{a} = \frac{1}{2} \left(\mathbf{q}_{B_t}^{B_k} (\hat{\mathbf{a}}^{B_t} - \mathbf{b}_{a_k}) + \mathbf{q}_{B_{t+1}}^{B_k} (\hat{\mathbf{a}}^{B_{t+1}} - \mathbf{b}_{a_k}) \right) \quad (8)$$

where the pre-integrated translation $\alpha_{B_{k+1}}^{B_k}$, the velocity $\beta_{B_{k+1}}^{B_k}$, the rotation $\mathbf{q}_{B_{k+1}}^{B_k}$ between B_k and B_{k+1} can be expressed by the following equations [34, 36]:

$$\alpha_{B_{k+1}}^{B_k} = \alpha_{B_t}^{B_k} + \beta_{B_{k+1}}^{B_k} \Delta t + \frac{1}{2} a \Delta t^2 \quad (9)$$

$$\beta_{B_{k+1}}^{B_k} = \alpha_{B_t}^{B_k} + a \Delta t \quad (10)$$

$$\mathbf{q}_{B_{k+1}}^{B_k} = \mathbf{q}_{B_t}^{B_k} \otimes \left[\frac{0}{\frac{1}{2} \omega \Delta t} \right] \quad (11)$$

Therefore, the observation measurements $\mathbf{z}_{B_{k+1}}^{B_k}$ generated by pre-integration can be denoted as

$$\mathbf{z}_{B_{k+1}}^{B_k} \text{ (IMU)} = \left\{ \alpha_{B_{k+1}}^{B_k}, \beta_{B_{k+1}}^{B_k}, \mathbf{q}_{B_{k+1}}^{B_k}, \mathbf{b}_{a_{k+1}}, \mathbf{b}_{g_{k+1}} \right\} \quad (12)$$

the residual for IMU pre-integration between two IMU states can be defined as

$$\mathbf{r}_B \left(\mathbf{z}_{B_{k+1}}^{B_k} \text{ (IMU)}, \chi \right) = \begin{bmatrix} \mathbf{r}_p \\ \mathbf{r}_v \\ \mathbf{r}_q \\ \mathbf{r}_{ba} \\ \mathbf{r}_{bg} \end{bmatrix} = \begin{bmatrix} \mathbf{q}_{B_k}^{W-1} \left(\mathbf{p}_{B_{k+1}}^W - \mathbf{p}_{B_k}^W - \mathbf{v}_{B_k}^W \Delta t - \frac{1}{2} \mathbf{g}^W \Delta t^2 \right) - \alpha_{B_{k+1}}^{B_k} \\ \mathbf{q}_{B_k}^{W-1} \left(\mathbf{v}_{B_{k+1}}^W - \mathbf{v}_{B_k}^W - \mathbf{g}^W \Delta t \right) - \beta_{B_{k+1}}^{B_k} \\ 2 \left[\mathbf{q}_{B_{k+1}}^{B_k-1} \otimes \left(\mathbf{q}_{B_k}^{W-1} \otimes \mathbf{q}_{B_{k+1}}^W \right) \right]_{\Im \Im \Im} \\ \mathbf{b}_{a_{k+1}} - \mathbf{b}_{a_k} \\ \mathbf{b}_{g_{k+1}} - \mathbf{b}_{g_k} \end{bmatrix} \quad (13)$$

where $[\cdot]_{\Im \Im \Im}$ is used for extracting the imaginary part of a quaternion. Here the \otimes denotes multiplication between two different quaternions.

4.3 | LiDAR modelling

4.3.1 | LiDAR NDT matching

The principle of the LiDAR matching-based localization is to associate the real-time point cloud with the prior map. This paper refers to the prior map point cloud as the source point cloud and the point cloud scanned by the SSL as the matching point cloud. In this paper, the pose is first estimated by conducting the map matching based on the NDT before its integration with IMU using FGO. Unlike ICP mentioned in the previous section, the NDT algorithm first divides the point cloud into voxels instead of matching directly with the target points. If the voxel being divided contains more than five points, the method calculates the mean μ and the covariance matrix Σ to calculate the normal distribution [23]

$$\mu = \frac{1}{m} \sum_{i=1}^m \mathcal{P}_{(k,i)} \quad (14)$$

ALGORITHM 1 Proposed IMU Aided LiDAR matching

Input: The source point cloud \mathcal{P}_S and the matching point cloud \mathcal{P}_m , IMU measurements; Initial pose of transformation $\vec{\mathbf{p}}_{ini}$
Output: The pose estimation of the system ($\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k)$).
S1: $\vec{\mathbf{p}} \leftarrow \vec{\mathbf{p}}_{ini}$
S2: Matching \mathcal{P}_S and \mathcal{P}_m based Equations (13)–(17)
S3: While $\vec{\mathbf{p}}$ has been updated
 (Perform IMU pre-integration: solving Equations (5)–(13) with FGO)
S4: For all points $\mathbf{x}_n \in \mathbf{x}_m$ find the cell \mathbf{x}_f that contains ($\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k)$), then update $\vec{\mathbf{g}}\mathbf{r}$, \mathbf{H}
S5: Solve Equation (17), then $\vec{\mathbf{p}} \leftarrow \vec{\mathbf{p}} + \Delta\vec{\mathbf{p}}$
S6: Repeat Step 2 to Step 5

$$\Sigma = \frac{1}{n} \sum_{i=1}^n (\mathcal{P}_{(k,i)} - \mu) (\mathcal{P}_{(k,i)} - \mu)^T \quad (15)$$

the $\mathcal{P}_{(k,i)}$ denotes the probability of a point enclosed by the cell \mathcal{P}_i . When both the covariance matrix and the mean are derived, the D-dimensional normal distribution is represented as follows [22, 23]:

$$p(\mathcal{P}_{(k,i)}) = \frac{1}{(2\pi)^{\frac{D}{2}} \sqrt{|\Sigma|}} \exp \left(-\frac{(\mathcal{P}_{(k,i)} - \mu)^T \Sigma^{-1} (\mathcal{P}_{(k,i)} - \mu)}{2} \right) \quad (16)$$

The score function for a single-source point can be calculated from Equation (17). Taking the logarithmic likelihood of the entire expression can have the effect of correcting for the impact of outliers [22, 23]:

$$score(\vec{\mathbf{p}}) = - \sum_{k=1}^n \tilde{p}(\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k)) \quad (17)$$

$$\mathbf{H} \Delta \vec{\mathbf{p}} = -\vec{\mathbf{g}}\mathbf{r} \quad (18)$$

and Equations (18) expressed Newton's law. \mathbf{H} represents the Pseudo-Hessian matrix, and $\vec{\mathbf{g}}\mathbf{r}$ denotes the gradient vector. $\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k) \in SE(3)$ represents the transform \mathcal{P}_k by vector $\vec{\mathbf{p}}$:

$$\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k) = \begin{bmatrix} \mathbf{R}_{S_k}^W & \mathbf{p}_{S_k}^W \\ 0 & 1 \end{bmatrix} \quad (19)$$

where $\mathbf{R}_{S_k}^W \in SO(3)$ denotes the rotation matrix, $\mathbf{p}_{S_k}^W$ stands for the translation.

Once the source and matched point clouds are input into the system, the pose estimation is performed according to Algorithm 1. Specifically, the performance of the NDT matching relies heavily on the performance of the initial guess of the pose estimation. Different from the conventional LiDAR standalone-based matching scheme, we adopt the high-frequency pose prediction from the IMU prediction as to the initial guess of the LiDAR scan matching, to get rid of the potential local minimum of solving Equations (17).

Using the SSL for NDT matching differs from the conventional mechanical LiDAR for NDT matching. Theoretically, the pseudo-Hessian matrix (\mathbf{H}) is the bridge connecting the measurements (the raw 3D point clouds) and the states to be estimated. Therefore, the degree of constraints can be inferred by the pseudo-Hessian matrix. In our previous research [37], the eigenvalues of the pseudo-Hessian matrix \mathbf{H} were investigated to be used to measure the degree of the overall constraints based on visual measurements. If the degree of the overall constraint is low, the degeneration of the state estimation can occur. In other words, these eigenvalues represent the degree of constraints in six different directions (position and orientation). By comparing the six eigenvalues of two different sensors, conventional mechanical LiDAR, and SSL, we can illustrate the adverse effect of small FOV on NDT matching. For the matrix \mathbf{H} , the singular value decomposition (SVD) [38] can be represented:

$$\mathbf{H} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T \quad (20)$$

the matrixes \mathbf{U} and \mathbf{V} are both the real 6×6 orthogonal matrix, and the matrix $\mathbf{\Lambda}$ is a real 6×6 diagonal matrix with non-negative real numbers on the diagonal. The diagonal entries $\lambda_e = \Lambda_{ee}$ can be considered the eigenvalues. The subscript e denotes the index of all the six eigenvalues:

$$\boldsymbol{\lambda} = [\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5 \lambda_6] \quad (21)$$

where the $\lambda_1, \lambda_2, \lambda_3$ denote the eigenvalues associated with the position estimation. The $\lambda_4, \lambda_5, \lambda_6$ denote the eigenvalues associated with the orientation estimation. The smaller λ_e indicates a weaker constraint in the associated direction. Therefore, the comparison of all the six eigenvalues can show the differences in the degree of constraints from different LiDAR sensors while using NDT matching.

4.3.2 | LiDAR matching factor modelling

The transformation $\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k)$ estimation from LiDAR matching is estimated by NDT matching, then for the LiDAR matching factor, this transformation can provide an absolute constraint. The transformation from the LiDAR matching frame to the IMU body frame can be represented as follows:

$$\mathbf{T}_{B_k}^W = \mathbf{T}_{S_k}^W \mathbf{T}_B^S \quad (22)$$

then we can get observation from LiDAR scan matching as follows:

$$\mathbf{z}_{B_k}^{W(LiDAR)} = \begin{bmatrix} \tilde{\mathbf{p}}_{B_k}^W & \tilde{\mathbf{q}}_{B_k}^W \end{bmatrix} \quad (23)$$

where the $\mathbf{z}_{B_k}^{W(LiDAR)}$ denotes the observation concerning the LiDAR scan matching, which includes the position ($\tilde{\mathbf{p}}_{B_k}^W$) and orientation ($\tilde{\mathbf{q}}_{B_k}^W$). Therefore, the residual $\mathbf{r}_s(\mathbf{T}(\vec{\mathbf{p}}, \mathcal{P}_k), \chi)$ from

the NDT matching can be defined as

$$r_s \left(T(\vec{p}, \mathbf{P}_k), \chi \right) = \begin{bmatrix} \mathbf{p}_{B_k}^W - \tilde{\mathbf{p}}_{B_k}^W \\ 2 \left[\mathbf{q}_{B_k}^{W^{-1}} \otimes \tilde{\mathbf{q}}_{B_k}^W \right]_{\mathcal{N}(\vec{q})} \end{bmatrix} \quad (24)$$

4.4 | Factor graph optimization

FGO aims to minimize the sum of the Mahalanobis norm of all factors to obtain the best estimation for the state. Based on the derived factors from IMU pre-integration and LiDAR scan matching, the objective function can be formulated as follows:

$$\chi = \min_{\chi} \frac{1}{2} \left\{ \sum_{k \in \{0, \dots, m\}} \|r_s(T(\vec{p}, \mathbf{P}_k), \chi)\|_{\mathbf{m}_{S_k}}^2 + \|r_B(\vec{z}_{B_{k+1}}^{(IMU)}, \chi)\|_{\mathbf{m}_{B_{k+1}}}^2 \right\} \quad (25)$$

where \mathbf{m}_{S_k} and $\mathbf{m}_{B_{k+1}}$ represent the information matrix for LiDAR matching and IMU pre-integration, which are experimentally determined. In this study, the sliding window optimization technique [39] is adopted to guarantee the real-time performance where the window size is set to 100, and when the total number of states contained in the window exceeds 100, we keep the last state and use it as the first frame in the new window.

5 | EXPERIMENT EVALUATION

5.1 | Experiment setup

To verify the effectiveness of the proposed method, we first collect a hand-held dataset in a typical campus scenario with lots of moving pedestrians. To further verify the performance of the proposed method, we conducted the other driving experiment in an urban canyon in Hong Kong.

Experiment setup for experiment 1 : In this experiment, the prior map is generated by LiDAR/inertial integration with accurate loop closure, based on the work in [40]. We carefully check the generated map and centimeter-level accuracy can be guaranteed with the help of the loop closure. The sensors for map generation include the VLP-16 LiDAR and XSENS MTI-10 IMU which were hand-held.

The generated map is shown in Figure 4 where Figures 4b and 4c denote the trajectory and the hand-held sensor kit, respectively. During the real-time localization, the SSL/inertial sensor set is shown in Figure 4d which consists of XSENS MTI-10 IMU and Livox Horizon SSL. In Experiment 1, since it is hard to obtain the ground truth of the LiDAR matching using a hand-held data collection system, we adopt the residual of the LiDAR matching as a performance evaluation metric which is commonly used in LiDAR point cloud registration [41].

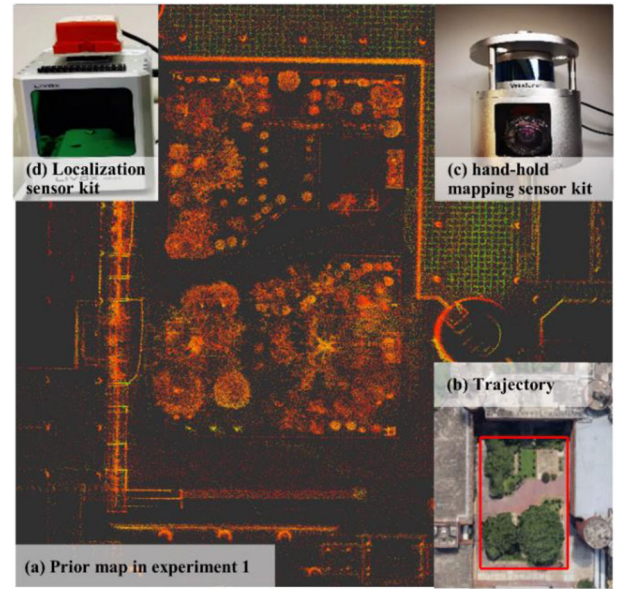


FIGURE 4 Prior map for the localization in Experiment 1, collected on our campus. (a) Illustration of the map in Experiment 1. (b) Tested trajectory. (c) Illustration of the hand-held sensor kit for map generation. (d) Illustration of the hand-held sensor kit for online localization based on SSL. SSL, solid-state LiDAR

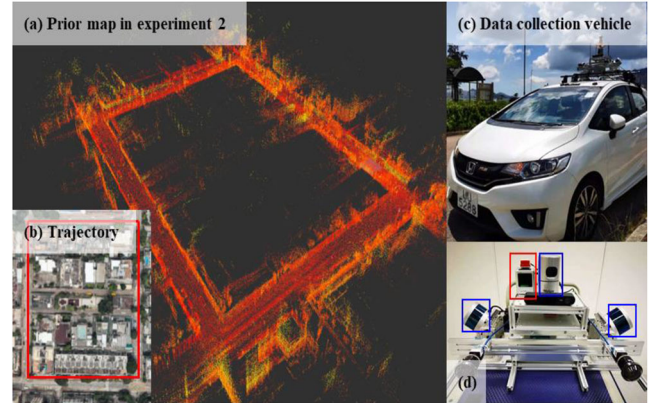


FIGURE 5 Prior map for the localization in Experiment 2, collected in an urban canyon of Hong Kong. (a) Illustration of the map in Experiment 2. (b) Tested trajectory. (c) Illustration of the data collection vehicle. (d) Illustration of the sensor kit for map generation and online localization where the three LiDARs annotated with blue boxes are used for prior map generation. The sensors annotated with a red box are used for online localization. Evaluated Methods: we compare the two localization methods as follows

Experiment setup for experiment 2 and experiment 3 : In those experiments, the point clouds are collected based on the three LiDARs (one HDL-32 LiDAR, one Leishen C-16 LiDAR, and one VLP-16 LiDAR) which are shown in Figures 5c to 5d. Moreover, the map is generated directly using the ground truth pose provided by the NovAtel SPAN-CPT, a GNSS (GPS, GLONASS, and BeiDou) RTK/INS (fibre-optic gyroscopes, FOG) integrated navigation system. The gyro bias in-run stability of the FOG is 1° per hour, and its random walk is 0.067° per hour. The baseline between the rover and

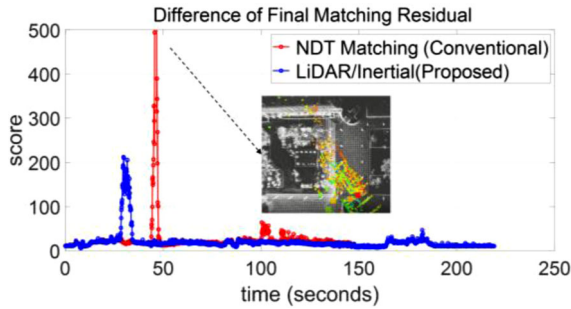


FIGURE 6 Comparison of the scores of the two algorithms. The blue line denotes the proposed method scores. The red line denotes the conventional method scores

the GNSS base station is within 7 km. Note that the extrinsic parameters between different sensors are already calibrated in our previously open-sourced UrbanNav dataset [35]. During the real-time localization, the SSL/inertial sensor set is the same as Figure 4d which consists of XSSENS MTI-10 IMU and Livox Horizon SSL. In Experiment 2, we adopt the NovAtel SPAN-CPT as the ground truth.

1. Conventional [24]: SSL LiDAR standalone matching-based localization.
2. Proposed: SSL/inertial matching-based localization proposed in this paper.

5.2 | Experiment 1 in campus scenario

The difference between the conventional method and the proposed method in the campus scenario is compared¹. The SSL is moved slightly faster than the typical adult walking speed during the experiment. And yielding at a higher angular velocity is also one of the features of this experiment. Figure 6 shows the score results of the matching. At about 50 s of the matching process, the matching score of the conventional method becomes very large. Meanwhile, the LiDAR localization using the conventional algorithm failed. The advantages of the method proposed in this paper are shown here. Where conventional methods fail, the proposed algorithm continues to provide a good initial guess so that the matching process can continue rather than fail.

We start the experiment in static mode, which ensures that a given initial pose facilitates the convergence of the NDT algorithm. Since hand-held is different from vehicle-mounted in the experiments, it will cause the LiDAR to reach a very high angular velocity. The motion in the Z -axis (vertical direction) will also be more intense than that of the vehicle-mounted LiDAR. This motion will have a significant impact on the positioning. As Figure 7 shows, the conventional methods cannot provide a good initial guess for point cloud matching in this case. An inaccurate initial guess often leads to poor quality of the computed results or even the computation fails to converge.

The proposed method can provide more trustworthy initial guesses than the conventional method in the scenario con-

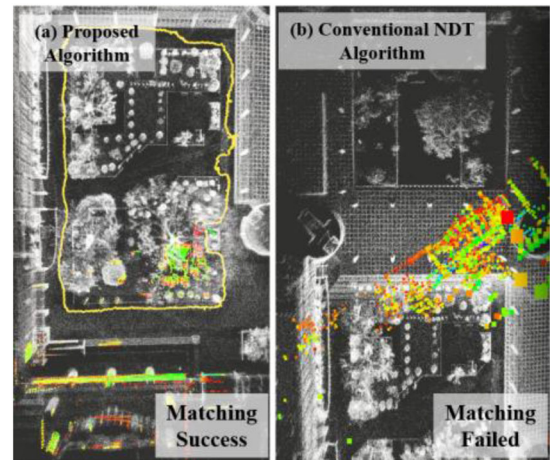


FIGURE 7 Comparison of the proposed algorithm with the conventional NDT algorithm. (a) The visualization result of the proposed method. (b) The visualization result of the conventional NDT algorithm. NDT, normal distribution transformation

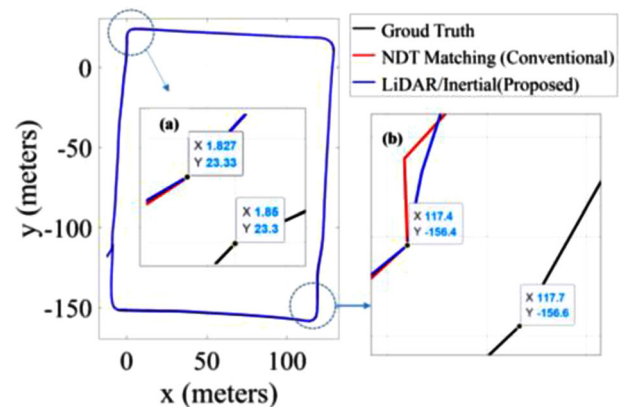


FIGURE 8 Comparison of the paths of the two algorithms with ground truth. The red and blue curves denote the trajectories from conventional NDT matching and the proposed method. (a) and (b) indicate the enlarged view in the corresponding area, respectively

structed for this experiment. Moreover, the proposed method can complete the entire matching and pose estimation process in the challenging scenario, which is an essential improvement over the conventional method.

5.3 | Experiment 2 in urban canyon experiment

Experiment 2 is conducted in a typical urban canyon in Hong Kong. The area is not too dense with buildings and moderate traffic flow. The results of the experiments are shown in Figures 8 and 9. The paths obtained by matching using a low-cost SSL are acceptable. Table 1 shows the positioning results of the two methods. The mean error was defined by the absolute pose error (APE) from the EVO toolkit [42].

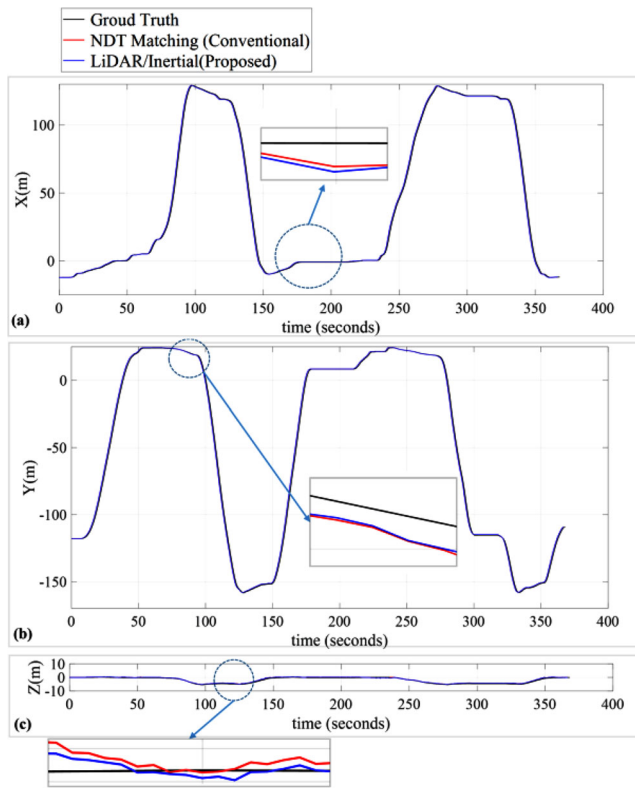


FIGURE 9 Comparison of values on XYZ-axis. (a) Trajectory in the x -direction. (b) Trajectory in the y -direction. (c) Trajectory in the z -direction. There is no significant difference in the XY direction; what is different is that both algorithms show some fluctuations in the Z -axis direction

TABLE 1 Performance of the proposed algorithm in urban Canyon I (in the unit of metre)

Results	Conventional	Proposed
Mean error	0.3831	0.3893
RMSE	0.4082	0.4140
Minimum error	0.0297	0.0378
Maximum error	0.7587	0.7575
Median error	0.4033	0.4092
STD	0.140896	0.140889

In Table 1, the mean value of the localization error using the proposed algorithm is 0.3893 m, while the value obtained by the conventional NDT algorithm is 0.3831 m, with a difference of 0.0062 m. The minimum error value using the conventional NDT algorithm is 0.0297 m, and the minimum value obtained by the new algorithm is 0.0378 m. The median error of the new algorithm is 0.4092 m, while the median error of the conventional NDT algorithm is 0.4033 m.

The above experimental data also show that the new algorithm proposed in this paper slightly differs from the conventional NDT algorithm in the actual urban canyon environment.

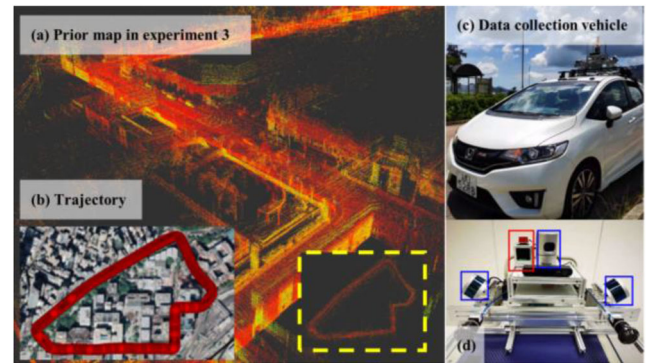


FIGURE 10 Demonstration of the prior map for the localization in Experiment 3 and the experimental scene in an urban canyon of Hong Kong. (a) Illustration of the map in Experiment 3. (b) Tested trajectory. (c) Illustration of the data collection vehicle. (d) Illustration of the sensor kit

5.4 | Experiment 3 for comparison of the difference between proposed method and horizon SLAM methods

Though both the proposed and SLAM methods can provide the localization results, there is a significant difference between our method and the existing SLAM methods. Our proposed method has the global HD prior map for the LiDAR point cloud matching to obtain the absolute pose for the user. The global map is one of the inputs of our system. Furthermore, we use the NDT algorithm with the global map and the scan from the SSL for the point cloud matching.

In order to evaluate the proposed approach more comprehensively, this section compares the proposed approach with the SSL SLAM method. We selected a new urban canyon area in Hong Kong for Experiment 3. The experiment setup is the same as Experiment 2 in the urban canyon. Figure 10 shows the prior map and the trajectory of Experiment 3.

Evaluated methods: We compare the two SLAM methods as follows:

1. LiLi-OM [10]: a real-time tightly coupled LiDAR-inertial odometry and mapping system for SSL (Livox Horizon) and conventional LiDARs
2. Horizon-LOAM1: an open-sourced package for Livox Horizon LiDARs in a low-speed scenario. The localization results are the LiDAR odometry results.
3. LIO-Livox2: an open-sourced package for Livox Horizon LiDARs in a high-speed scenario that integrated the IMU pre-integration and FGO for better estimation.
4. Proposed: SSL/inertial matching-based localization is proposed in this paper.

The performance between the LiLi-OM, Horizon-LOAM, LIO-Livox, and the Proposed method is shown in Table 2. We also use the EVO toolkit [42] for the evaluation. LiLi-OM will be failed while running with our dataset. And It can be seen there are significant improvements from the proposed method and LIO-Livox to the Horizon-LOAM like most LO methods

TABLE 2 Performance of the proposed algorithm in urban Canyon II (in the unit of metre)

Results	LiLi-OM	Horizon-LOAM	LIO-Livox	Proposed
Mean error	Fail	65.9767	1.2160	1.0462
RMSE	Fail	82.9049	1.3200	1.1863
Minimum error	Fail	14.8182	0.1062	0.1351
Maximum error	Fail	175.6583	3.9052	3.5909
Median error	Fail	37.1513	1.1618	0.9704
STD	Fail	50.2026	0.5135	0.5592

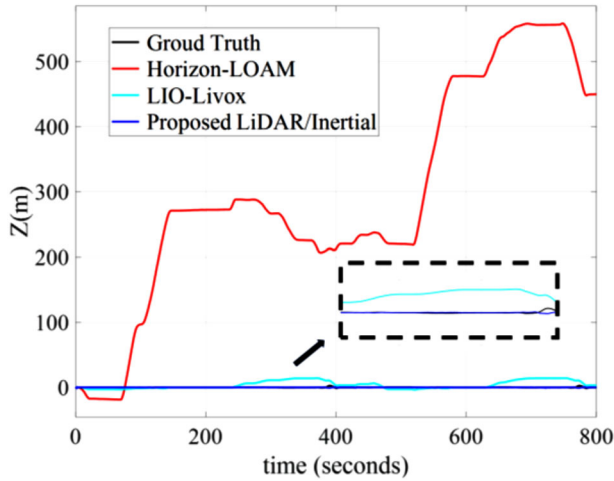


FIGURE 11 Comparison of values on Z-axis. The Horizon-LOAM occurs significant drift in Experiment 3. The red, cyan, and blue curves denote the trajectories from Horizon-LOAM, LIO-Livox, and the proposed method, respectively

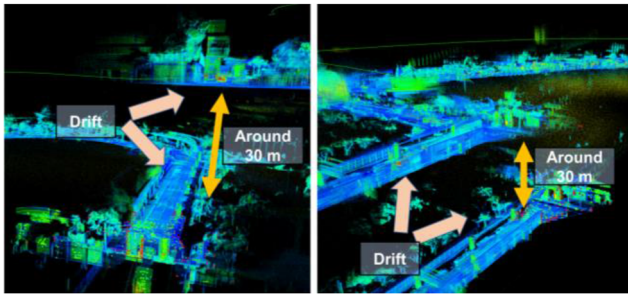


FIGURE 12 Typical drift occurs while using the Horizon-LOAM for mapping

would have a drift in the z -axis like Figures 11 and 12 show. The proposed method has better performance than LIO-Livox and the mean errors are 1.0462 and 1.1260, respectively.

5.5 | Discussion about the effect of different FOV

To experimentally show the differences between SSL laser scanning and conventional mechanical LiDAR laser scanning, we

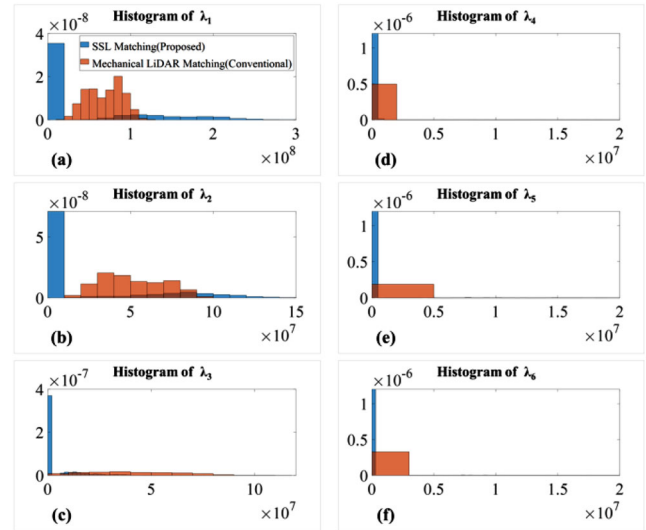


FIGURE 13 Histograms of eigenvalues of the SSL and mechanical LiDAR matching, associated with the position and orientation. The blue histogram shows the histogram obtained using solid-state LiDAR. The orange histogram denotes the histogram obtained using conventional mechanical LiDAR. (a) The histogram of λ_1 indicates the constraints in the x -direction. (b) The histogram of λ_2 indicates the constraints in the y -direction. (c) The histogram of λ_3 indicates the constraints in the z -direction. (d) The histogram of λ_4 indicates the constraints of the roll. (e) The histogram of λ_5 indicates the constraints of the pitch. (f) The histogram of λ_6 indicates the constraints of the yaw

analyzed the histogram of all eigenvalues in the NDT matching process for two different LiDARs in the same segment path. As shown in Figure 13, the X-axis is the eigenvalues and the Y-axis is the frequency density of occurrences of the values within a bin. Statistically, we find that the eigenvalues in the case of using SSL are generally smaller than those in the case of using conventional mechanical LiDAR. This is caused by the small FOV of SSL resulting in weaker constraints in the corresponding direction compared to the conventional mechanical LiDAR. Alternatively, using the novel SSL makes NDT matching more challenging due to the limited FOV.

5.6 | Discussion: impact of the grid size on the positioning performance of the NDT matching

Different grid sizes of the NDT can significantly impact the system solution. After adding the new SSL to the system, the impact of grid size on the system solution accuracy and computation time is also worth discussing. Figure 14 shows the object in the frame of the SSL point cloud and the grid size effectiveness illustration. The NDT algorithm would divide the point cloud into several voxels, as shown in Figure 14b. If the grid size is large, there would be only one voxel for the whole point cloud which means that the n in Equation (17) is only one in this case. And the small grid size means the n would be larger than the large grid size case which would be more computational load

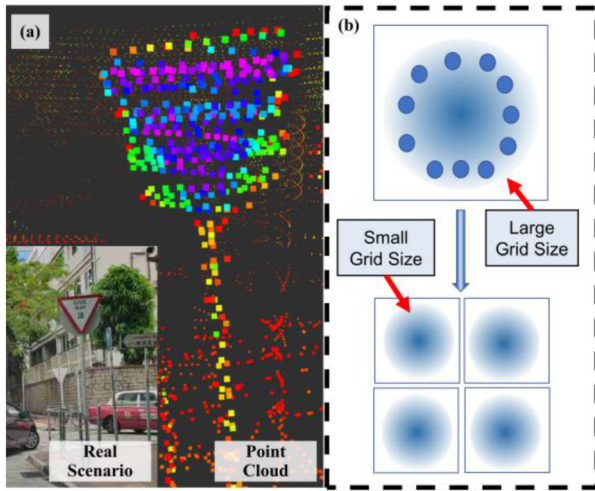


FIGURE 14 (a) The road sign and its point cloud from SSL. (b) The illustration of the different grid sizes in the NDT algorithm

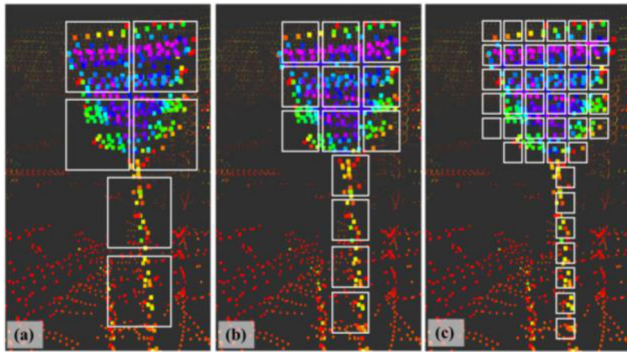


FIGURE 15 (a) Large grid size for the road sign. (b) Medium grid size for the road sign. (c) Small grid size for the road sign. A smaller grid size would add more computational load for the estimation

TABLE 3 The used time of the NDT algorithm under different grid sizes (in the unit of milliseconds)

Grid edge length	0.5 m	1 m	1.5 m	2 m
Mean	153.6406	222.2349	153.1312	187.9326
STD	67.9078	87.0368	46.9278	79.1065
Maximum	1215.20	1504.80	1019.10	1678.40
Minimum	84.862	93.999	88.821	92.028

while solving Equation (17). An illustration of different grid sizes for the voxelization of the road sign is shown in Figure 15.

Table 3 shows the align time of the NDT algorithm for different grid sizes, and the best performance is found for a grid edge length of 1.5 m. In other words, the least amount of aligning time is required under these conditions.

We also evaluated the accuracy of the NDT algorithm for different grid sizes. As shown in Table 4, the mean value of the localization error is the smallest for a grid edge length of

TABLE 4 Error value of NDT algorithm at different grid sizes (in the unit of metre)

Grid edge length	0.5 m	1 m	1.5 m	2 m
Mean error	0.3919	0.3812	0.3774	0.3946
RMSE	0.4274	0.4052	0.4016	0.4268
Maximum error	1.9189	0.7694	0.8171	1.5011
Minimum error	0.0704	0.0592	0.0809	0.0460

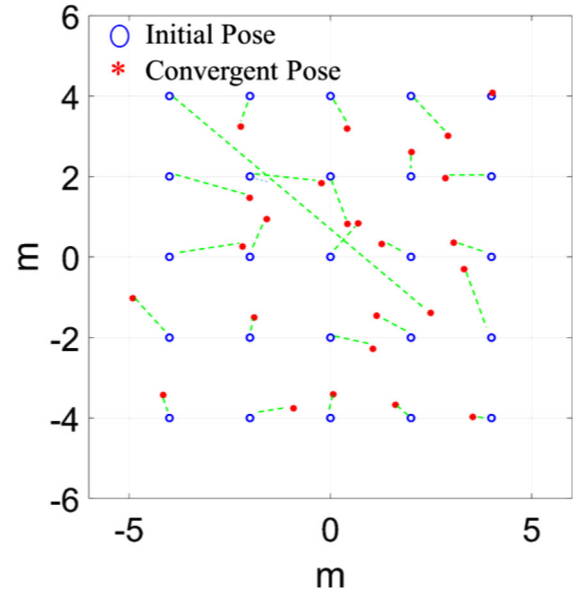


FIGURE 16 The results of the convergence of the NDT algorithm for different initial poses. The blue circle indicates the given initial pose, and the red star indicates the convergent poses

1.5 m. The maximum value of the localization error at a grid edge length of 0.5 m is larger than the other three sets of data.

5.7 | Discussion: sensitivity analysis of the NDT matching upon different initial guess

Current evaluations of the performance of new low-cost SSL are incomplete. The accuracy of the initial pose has a direct and significant impact on the results for LiDAR localization. The small FOV of the new SSL will undoubtedly make it more likely to encounter unknown problems than the 360° laser scan of a mechanical rotating LiDAR [43]. Therefore, it is necessary to explore the effect of different initial poses on the localization using low-cost SSL. As can be seen from Figures 16 and 17, the convergence points at different initial poses do not converge to the same point for the same scene as the mechanical rotating LiDAR. The point cloud of SSL converges in the system to obtain different solutions. Usually, the closer the solution obtained by convergence is to the (0, 0) point, it indicates the system's better convergence performance at this initial positional. The subsequent performance of the system in different initial positions is also an interesting point. In Figure 17, the

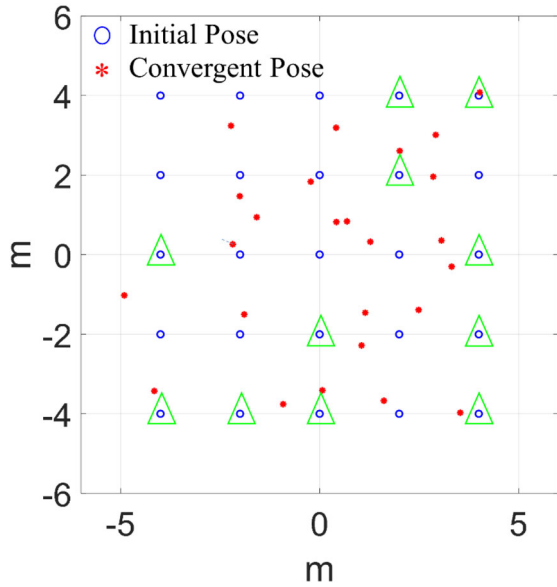


FIGURE 17 The points are surrounded by triangles indicate the initial poses that will eventually cause the matching to fail. Only the initial poses are changed from the 2D plane, that is, x, y, θ

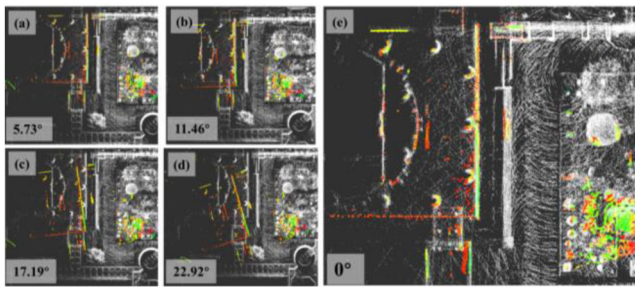


FIGURE 18 Comparison of the effect of angle changes between 0° and 22.92° on the convergence of the NDT algorithm. The NDT algorithm can complete the subsequent localization within this angle variation range. (a) The case where the initial angle is 5.73° . (b) The case where the initial angle is 11.46° . (c) The case where the initial angle is 17.19° . (d) The case where the initial angle is 22.92° . (e) The case where the initial angle is 0°

points circled by triangles are the points that will eventually make the system fail to converge, and the matching eventually fails. These points do not seem to have any pattern and cannot be judged simply from a distance between the initial posture and $(0, 0)$.

This experiment also explores the practical performance of the new SSL with varying angles. Figure 18 represents the initial angle that allows for successful subsequent localization. While Figure 19 represents the convergence of the NDT algorithm when the angle is negative, all the cases shown in the figure make the matching fail. Unlike the previous [43] in this respect, the SSL can only vary between $0^\circ < \theta < 22.92^\circ$. Otherwise, the system cannot converge. Figure 20 shows that the conventional mechanical rotating LiDAR can record all the surrounding features even if a certain angle rotates it. In contrast, once an angle offsets the SSL, it will result in the incomplete recording of the surrounding scene, which is a severe problem for

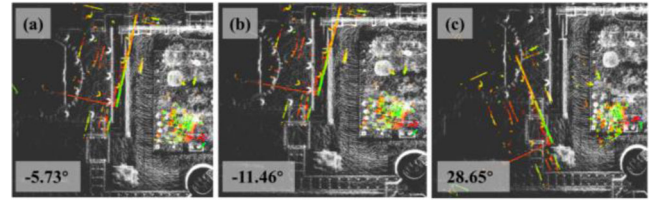


FIGURE 19 The angles that would make the convergence value of the NDT algorithm deteriorate. (a) The case where the initial angle is -5.73° . (b) The case where the initial angle is -11.46° . (c) The case where the initial angle is 28.65°

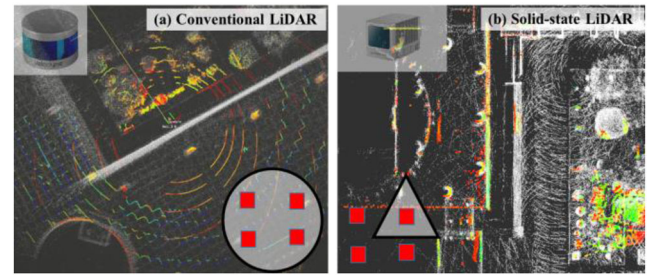


FIGURE 20 Comparison of the two LiDARs laser scanning. (a) The matching process of the conventional LiDAR. (b) The matching process of the solid-state LiDAR

autonomous vehicles, while the scene's complexity also has a significant impact on the new SSL.

6 | CONCLUSIONS

Achieving accurate and low-cost localization is highly required by ADV. In this paper, we investigate the potential of low-cost SSL-based localization with a prior map. To alleviate the problem of small FOV of low-cost SSL, this paper adopts the IMU to provide a high-frequency initial guess for the LiDAR matching. The evaluated datasets show that the proposed method can achieve ~ 0.5 m of localization accuracy. With the help of the IMU, the robustness is improved during high dynamic motions in urban scenarios. Moreover, we also provide the sensitivity analysis of the SSL-based matching upon the initial guess.

Considering the advantage of the low-cost SSL, installing multiple SSL on one autonomous driving vehicle does not have an unacceptable cost increase for the cost of production. We will explore the performance of multiple low-cost SSL in urban canyons in our future work. As shown in Figure 21, assuming an ADV with four low-cost SSL, there is an opportunity to compensate for their small FOV when used simultaneously to catch up with the performance of high-cost conventional LiDARs.

In the future, we will evaluate the proposed systems in more complex scenarios. Firstly, we will investigate the map update using multiple SSL. In practice, the environments are changed which causes the inconsistency between the real-time environment and the previously built point cloud map. Therefore, we will also explore the update of the HD map using the low-cost SSL to guarantee the freshness of the map. Furthermore, wider FOV is vital for map updates and will directly determine the

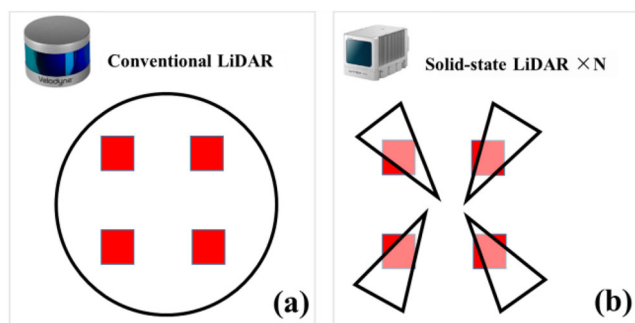


FIGURE 21 Idea of combining multi-solid-state LiDAR sensors. (a) Illustration of conventional LiDAR scanning. (b) Illustration of multi-solid-state LiDAR scanning

quality of the updated map. Secondly, the limited field-of-view (FOV) of the SSL can lead to the degeneration of the localization. Unlike a single SSL, the degree of FOV provided by multiple SSL will be close to that of a conventional mechanical LiDAR, which means that we can acquire more point clouds in a single-frame scan. As a result, the geometry constraints can be significantly improved because more point clouds can provide better constraints for point cloud matching. Therefore, we will investigate the potential of multiple SSL-based localization in more complex scenarios.

ACKNOWLEDGEMENT

This research was funded by the Guangdong Basic and Applied Basic Research Foundation (2021A1515110771). This research was also supported by Riemann Laboratory, Huawei Technologies with grant number ZGD2 and University Grants Committee of Hong Kong under the scheme Research Impact Fund (R5009-21).

CONFLICT OF INTEREST

The authors do not have a conflict of interest.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

REFERENCES

- Pettersson, I., Karlsson, M.: Setting the stage for autonomous cars: A pilot study of future autonomous driving experiences. *IET Intell. Transp. Syst.* 9(7), (2015)
- Wen, W., et al.: GNSS/LIDAR integration aided by self-adaptive Gaussian mixture models in urban scenarios: An approach robust to non-Gaussian noise. In: 2020 IEEE/ION Position, Location and Navigation Symposium (PLANS) (2020)
- Levinson, J., Thrun, S.: Robust vehicle localization in urban environments using probabilistic maps. In: 2010 IEEE International Conference on Robotics and Automation, IEEE (2010)
- Wang, L., Zhang, Y., Wang, J.: Map-based localization method for autonomous vehicles using 3D-LIDAR. *This Work Is Supported in Part by the National Natural Science Foundation of China under Grant No. 61473209. IFAC-PapersOnLine.* 50(1), 276–281 (2017)
- Brefler, J., et al.: GNSS positioning in non-line-of-sight context—A survey. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) (2016)
- Groves, P.D.: Principles of GNSS, inertial, and multisensor integrated navigation systems, 2nd ed. [Book Review]. *IEEE Aerospace Electron. Syst. Magazine* 30(2), 26–27 (2015)
- Groves, P.D.: Shadow matching: A new GNSS positioning technique for urban canyons. *J. Navigat.* 64(3), 417–430 (2011)
- Wen, W., et al.: Uncertainty estimation of LIDAR matching aided by dynamic vehicle detection and high definition map. *Electron. Lett.* 55(6), 348–349 (2019)
- Tomasello, P., et al.: Dscnet: Replicating Lidar Point Clouds with Deep Sensor Cloning (2019)
- Li, K., Li, M., Hanebeck, U.D.: Towards high-performance solid-state-LIDAR-inertial odometry and mapping. *IEEE Robot. Autom. Lett.* 6(3), 5167–5174 (2021)
- Broggi, A., et al.: Extensive tests of autonomous driving technologies. *IEEE Trans. Intell. Transp. Syst.* 14(3), 1403–1415 (2013)
- Levinson, J., Montemerlo, M., Thrun, S.: Map-based precision vehicle localization in urban environments. In: *Robotics: Science and Systems*, Citeseer (2007)
- Wan, G., et al.: Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE (2018)
- Mohammadzade, H., Hatzinakos, D.: Iterative closest normal point for 3D face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 381–397 (2013)
- Yang, J., et al.: Go-ICP: A globally optimal solution to 3D ICP point-set registration. *IEEE Trans. Pattern Anal. Mach. Intell.* 38, 2241–2254 (2016)
- Pavlov, A.L., et al.: AA-ICP: Iterative closest point with anderson acceleration. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 1–6 (2018)
- Muñoz, F.I.I., Comport, A.I.: Point-to-hyperplane ICP: Fusing different metric measurements for pose estimation. *Adv. Robot.* 32, 161–175 (2018)
- Segal, A.V., Hähnel, D., Thrun, S.: Generalized-ICP. In: *Robotics: Science and Systems* (2009)
- Koide, K., et al.: Voxelized GICP for fast and accurate 3D point cloud registration (2020)
- Wang, Y., Xiong, R., Li, Q.: EM-based point to plane ICP for 3D simultaneous localization and mapping. *Int. J. Robot. Autom.* 28 (2013)
- Luong, H., et al.: Consistent ICP for the registration of sparse and inhomogeneous point clouds. In: 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), pp. 262–267 (2016)
- Saarienen, J., et al.: 3D normal distributions transform occupancy maps: An efficient representation for mapping in dynamic environments. *Int. J. Robot. Res.* 32, 1627–1644 (2013)
- Magnusson, M.: The three-dimensional normal-distributions transform: An efficient representation for registration, surface analysis, and loop detection. (2009)
- Pang, S., et al.: 3D scan registration based localization for autonomous vehicles—A comparison of NDT and ICP under realistic conditions. In: 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), IEEE (2018)
- Akai, N., et al.: Autonomous driving based on accurate localization using multilayer LIDAR and dead reckoning. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), IEEE (2017)
- Wen, W., Hsu, L.-T., Zhang, G.: Performance analysis of NDT-based graph SLAM for autonomous vehicle in diverse typical driving scenarios of Hong Kong. *Sensors* 18(11), 3928 (2018)
- Kato, S., et al.: Autoware on board: Enabling autonomous vehicles with embedded systems. In: 2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPs), IEEE (2018)
- Huang, F., et al.: Point wise or feature wise? Benchmark comparison of public available LIDAR odometry algorithms in urban canyons. *ArXiv abs/2104.05203* (2021)
- Lin, J., Zhang, F.: Loam Livox: A fast, robust, high-precision LIDAR odometry and mapping package for LIDARs of small FoV. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE (2020)
- Wang, H., Wang, C., Xie, L.: Lightweight 3-D localization and mapping for solid-state LIDAR. *IEEE Robot. Autom. Lett.* 6(2), 1801–1807 (2021)

31. Xu, W., Zhang, F.: Fast-Lio: A fast, robust LIDAR-inertial odometry package by tightly-coupled iterated Kalman filter. *IEEE Robot. Autom. Lett.* 6(2), 3317–3324 (2021)
32. Xu, W., et al.: Fast-Lio2: Fast direct LIDAR-inertial odometry. *IEEE Trans. Robot.* 38(4), 2053–2073 (2022)
33. Lin, J., et al.: R2LIVE: A robust, real-time, LIDAR-inertial-visual tightly-coupled state estimator and mapping. *IEEE Robot. Automat. Lett.* 6(4), 7469–7476 (2021)
34. Forster, C., et al.: IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation. In: *Conference on Robotics: Science and Systems* (2015)
35. Hsu, L.-T., et al.: An open-sourced multisensory dataset for benchmarking positioning algorithms designed for urban areas. In: *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*. (2021)
36. Lupton, T., Sukkarieh, S.: Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions. *IEEE Trans. Robot.* 28(1), 61–76 (2012)
37. Bai, X., Wen, W., Hsu, L.T.: Degeneration-aware outlier mitigation for visual inertial integrated navigation system in urban canyons. *IEEE Trans. Instrum. Meas.* 70, 1–15 (2021)
38. Golub, G.H., Reinsch, C.: Singular value decomposition and least squares solutions. In: *Linear Algebra*. Springer, Heidelberg (1971)
39. Qin, T., Li, P., Shen, S.: Vins-Mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Trans. Robot.* 34(4), 1004–1020 (2018)
40. Shan, T., et al.: LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5135–5142 (2020)
41. Magnusson, M., et al.: Evaluation of 3D registration reliability and speed-a comparison of ICP and NDT. In: *2009 IEEE International Conference on Robotics and Automation*. IEEE (2009)
42. Grupp, M.: Evo: Python package for the evaluation of odometry and slam. <https://github.com/MichaelGrupp/evo> Cited by: Table, p. 7 (2017)
43. Akai, N., et al.: Robust localization using 3D NDT scan matching with experimentally determined uncertainty and road marker matching. (2017)

How to cite this article: Zhong, Y., Huang, F., Zhang, J., Wen, W., Hsu, L.-T.: Low-cost solid-state LiDAR/inertial-based localization with prior map for autonomous systems in urban scenarios. *IET Intell. Transp. Syst.* 17, 474–486 (2023).
<https://doi.org/10.1049/itr2.12273>