LiDAR Inertial Odometry Aided Robust LiDAR Localization System in Changing City Scenes

Wendong Ding, Shenhua Hou, Hang Gao, Guowei Wan, Shiyu Song[1]

Abstract— Environmental fluctuations pose crucial challenges to a localization system in autonomous driving. We present a robust LiDAR localization system that maintains its kinematic estimation in changing urban scenarios by using a dead reckoning solution implemented through a LiDAR inertial odometry. Our localization framework jointly uses information from complementary modalities such as global matching and LiDAR inertial odometry to achieve accurate and smooth localization estimation. To improve the performance of the LiDAR odometry, we incorporate inertial and LiDAR intensity cues into an occupancy grid based LiDAR odometry to enhance frame-to-frame motion and matching estimation. Multi-resolution occupancy grid is implemented yielding a coarse-to-fine approach to balance the odometry's precision and computational requirement. To fuse both the odometry and global matching results, we formulate a MAP estimation problem in a pose graph fusion framework that can be efficiently solved. An effective environmental change detection method is proposed that allows us to know exactly when and what portion of the map requires an update. We comprehensively validate the effectiveness of the proposed approaches using both the Apollo-SouthBay dataset and our internal dataset. The results confirm that our efforts lead to a more robust and accurate localization system, especially in dynamically changing urban scenarios.

I. Introduction

Transportation of people and goods in the last hundred years has changed at a drastic pace. This growth potential in the industry attracted the attention of technology experts who sought to solve this complex yet promising problem. More recently, disruptive technological concepts like ride-sharing and autonomous delivery are propelling this industry forward, like autonomous driving [1]. One of the primary requirements of an autonomous system is a defined mapped area and in order to navigate autonomously, the prevalent approach requires precise localization. But precise localization systems are not only complex but are also difficult to implement in a dynamically changing environment. Previous works [2], [3], [4], [5] have demonstrated that some specific changes in the environment, e.g. road repavements, puddles, snowdrifts, can be overcome using existing technologies. However, it is still one of the most challenging issues that causes the failure of the vehicle's localization module which is based on matching online sensor measurements.

In this work, we seek to integrate a LiDAR inertial odometry (LIO) together with our matching based global
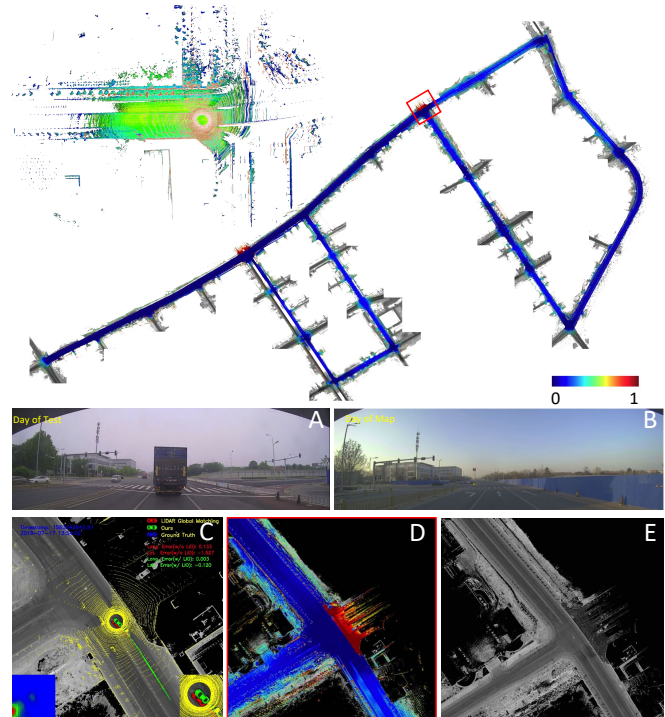
Fig. 1: Online LiDAR data (brown) and the submap (occupancy probability: blue/green/yellow) built by the LiDAR inertial odometry is shown on the left of the top panel. The prebuilt localization map marked with the probability of how likely change exists in each cell estimated by the environmental change detection module is shown at the bottom right of the top panel. Bottom panel: zoomed-in view of an intersection where metal walls were recently removed. (A) and (B) comparison of the scene at different times. (C) visual comparison of localization results: green car (w/ LIO) vs red (w/o LIO) vs blue (ground truth). (D) zoomed-in view of the change detection results. (E) Updated map by merging new LiDAR scans.

LiDAR localization module. Both the measurements from the two modules are jointly fused in a pose graph optimization framework. Given the fact that map matching and odometry are complementary methods, it allows us to deliver a robust localization system that can overcome temporary environmental changes or map errors in general as shown in Figure 1, and at the same time consistently provide precise global localization results.

To summarize, our main contributions are:

- A joint framework for vehicle localization that adaptively fuses global matching and local odometry cues, which effectively shields our system from their failure in changing urban scenes.
- A LiDAR inertial odometry that tightly couples LiDAR

and inertial measurements and incorporates both the occupancy and LiDAR intensity cues to provide real time accurate state estimation.
- A robust vehicle localization system that has been rigorously tested daily in crowded and busy urban streets demonstrating its robustness in challenging and dynamically changing environment.

II. RELATED WORK

Long-term Localization Building a 7 days 24 hours all-weather localization system is, by all means, a challenging task that has received significant attention in recent years. J. Levinson et al. [6] showed that the less reflectance caused by wet road surface can be adjusted by normalizing the brightness and standard deviation for each LiDAR scan. R. Wolcott et al. [2], [3] demonstrated a robust LiDAR localization system that can survive through road repavement and snowfall by introducing a multiresolution Gaussian mixture representation in the map. Wan et al. [5] showed that the LiDAR localization system successfully passed a challenging road section with newly built walls and repaved road by incorporating the altitude cues. M. Aldibaja et al. [4] enhance the robustness of their localization system by introducing principal component analysis (PCA) and edge profiles, especially during rainy or snowy days. While these works focus on solving specific problems using niche technologies, our work seeks to have a more general solution by adaptively fusing the complementary cues from the odometry and the global matching module. Other works [7], [8], [9], [10], [11] address a similar long-term localization problem but use vision sensors that are brittle to the scene's appearance change caused by time, light or weather.

LiDAR Inertial Odometry There is a lot of literature addressing the LiDAR odometry/SLAM problem [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]. Inertial measurements help solve the problem by providing prior estimation and compensating the motion distortion [14], [18], [17], [19], building tighter local motion constraints [18] or further establishing a tightly-coupled odometry [23], [24]. In this work, we follow Hess's work [13] and integrate an occupancy grid based LiDAR inertial odometry into our localization framework because of its similar map representation to our global matching module and its compatibility to multiple laser scanners. Inspired by previous works, the usage of inertial sensors and other extensions are introduced for performance improvement.

Localization Fusion Methods There are several methods related to the fusion of estimation from different sensors or methods. One important category is loosely-coupled fusion. Methods [25], [5], [26] leverage error-state Kalman filter and loosely fuse the pose estimation from different methods. Instead of using a Kalman filter, similar to [26], our method utilizes a graph-based fusion framework, which is known to outperform filtering methods with better accuracy per unit of computing time [27], [28]. A. Soloviev [29] demonstrates a tightly-coupled navigation system fusing GNSS, LiDAR and inertial measurements.
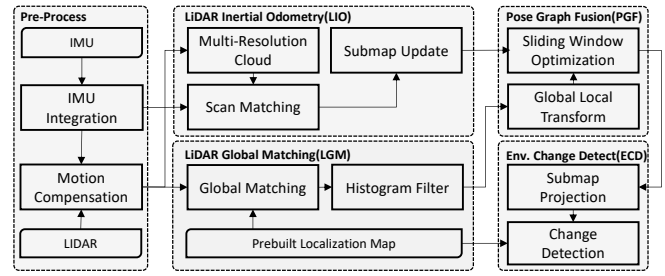

Fig. 2: The architecture of the proposed LiDAR inertial odometry aided LiDAR localization system with an environmental change detection module.

III. METHOD

This section describes the architecture of the proposed LiDAR localization framework designed in detail as shown in Figure 2. Our system consists of four modules: LiDAR inertial odometry (LIO), LiDAR global matching (LGM), pose graph based fusion (PGF) and environmental change detection (ECD). We build our system by following the latest LiDAR localization work by G. Wan et al. [5], and leveraging it as a submodule, the LGM, in our framework. It is a global localization method that matches online LiDAR scans against a pre-built map, and conducts a 3 DoF (x, y, yaw) estimation. The other 3 DoF (roll, pitch, altitude) can be estimated by reading IMU gravity measurements and a digital elevation model (DEM) map, once we successfully locate horizontally. The other two modules, LIO and PGF, are formulated in terms of solving different maximum a posteriori probability (MAP) estimation problems.

The MAP estimation in this work leads to a nonlinear optimization problem defined usually in a *sliding window* considering a window of the latest measurements. Let $\mathcal{K}$ denote the set of all frames in a window. The set of states and measurements in a sliding window are denoted as $\mathcal{X} = \{\mathbf{x}_k\}_{k \in \mathcal{K}}$ and $\mathcal{Z} = \{\mathbf{z}_k\}_{k \in \mathcal{K}}$, respectively. The state of frame k is represented by the orientation, position, velocity and IMU biases, $\mathbf{x}_k = [\boldsymbol{\omega}_k, \mathbf{t}_k, \mathbf{v}_k, \mathbf{b}_k]$. $\boldsymbol{\omega}_k$ is the *Lie algebra* usually denoted as $so(3)$. $R_k = \text{Exp}(\boldsymbol{\omega}_k)$ and the pose $(R_k, \mathbf{t}_k)$ belong to the Special Orthogonal Group SO(3) and the Special Euclidean Group SE(3) respectively. Exp() and Log() are the exponential and logarithmic map associating $so(3) \rightarrow \text{SO}(3)$ and SO(3) $\rightarrow so(3)$. Velocities are represented by vectors, i.e., $\mathbf{v}_k \in \mathbb{R}^3$. IMU biases can be further written as $\mathbf{b}_k = [\mathbf{b}_k^a, \mathbf{b}_k^g]$, where $\mathbf{b}_k^a, \mathbf{b}_k^g \in \mathbb{R}^3$ are the accelerometer and gyroscope bias, respectively.

A. LiDAR Inertial Odometry

The LiDAR inertial odometry plays an essential role in our system in promoting localization performance in challenging circumstances, for example, map expiration or environmental change due to road construction or severe weather. LiDAR odometry estimates relative poses between frames and simultaneously helps us build a local map, called a *submap*. This submap is always up-to-date, continuously updated with each new LiDAR scan. Our system takes advantage of the

submap, smoothes the estimated trajectory, and also ensures the system reliability in extreme circumstances.

Our LiDAR inertial odometry implementation follows W. Hess's work [13], but with a number of important extensions to improve its accuracy. Firstly, the method of the 3D occupancy grid is used instead of 2D to achieve a full 6 DoF odometry. This extension naturally allows its application in a three-dimensional environment, such as a parking structure or an overpass, and simplifies the under-mentioned IMU pre-integration. Secondly, important inertial cues are incorporated to provide motion prediction and relative constraints between frames. More importantly, the incorporation of the inertial cues allows us to implement the motion compensation for the distorted LiDAR scans caused by the moving platform. In order to make the computing time non-intractable, we adopt pre-integration of the inertial measurements introduced by [30] in our implementation. Thirdly, in consideration of the rich information from lane or road surface markers in these scenarios, the LiDAR intensity cues are incorporated during the occupancy grid registration as complementary to the occupancy probability of each grid cell. It provides valuable texture information of the environment. Finally, we apply a coarse-to-fine manner while solving the non-linear optimization problem by introducing multi-resolution into our occupancy grid implementation. It not only helps the grid registration converge but also keeps the computational requirement at bay, validated by experimental results in Section IV-C.

We formulate the LiDAR inertial odometry as a MAP estimation problem. The posterior probability of the state $\mathbf{x}_k^L$, given the previous state $\mathbf{x}_{k-1}^L$, the submap $\mathcal{S}_{k-1}$ updated until the recent frame $k-1$, and the measurements $\mathbf{z}_k$, is

$$P(\mathbf{x}_k^L|\mathbf{z}_k, \mathbf{x}_{k-1}^L, \mathcal{S}_{k-1}) \propto P(\mathbf{z}_k^P|\mathbf{x}_k^L, \mathcal{S}_{k-1}) \\ P(\mathbf{z}_k^I|\mathbf{x}_k^L, \mathbf{x}_{k-1}^L), \quad (1)$$

where $\mathbf{z}_k = \{\mathbf{z}_k^P, \mathbf{z}_k^I\}$, $\mathbf{z}_k^P$ and $\mathbf{z}_k^I$ are the point cloud and inertial measurement, respectively. The superscript L denotes that the state $\mathbf{x}_k^L$ is expressed in the local frame defined by the submap and odometry.

Under the assumption of zero-mean Gaussian-like probability, the likelihood of the measurements are defined by building the cost functions as:

$$P(\mathbf{z}_k^I|\mathbf{x}_k^L, \mathbf{x}_{k-1}^L) \propto \exp{-\frac{1}{2}\|\mathbf{r}_k^I\|_{\Lambda_k^I}^2}, \quad (2)$$

where $\|\mathbf{r}\|_\Lambda^2 = \mathbf{r}^T \Lambda^{-1} \mathbf{r}$ and

$$P(\mathbf{z}_k^P|\mathbf{x}_k^L, \mathcal{S}_{k-1}) \propto \prod_i \prod_j \exp{-\frac{1}{2\sigma_{o_i}^2}\|\text{SSOP}\|^2} \\ \prod_i \prod_j \exp{-\frac{1}{2\sigma_{r_i}^2}\|\text{SSID}\|^2}. \quad (3)$$

The Equation 2 is calculated according to the pre-integration method. Please refer to [30] for details. The Sum of Squared Occupancy Probability) SSOP and Sum of

Squared Intensity Difference) SSID represent the occupancy grid probability and the LiDAR intensity cost, respectively, defined as:

$$\begin{cases} \text{SSOP} = 1 - P(s) \\ \text{SSID} = \frac{u_s - \text{I}(\mathbf{p}_j)}{\sigma_s}. \end{cases} \quad (4)$$

Given a LiDAR point $\mathbf{p}_j \in \mathbb{R}^3$, a submap with resolution i, and a pose state $\mathbf{x}_k^L = [\text{R}_k, \mathbf{t}_k]$, the hit cell s in the submap can be found. $P(s)$ is the occupancy probability of the hit cell in the submap at the desired resolution i. This occupancy probability is maintained by keep inserting new LiDAR scans into the submap after we maximize the posterior $P(\mathbf{x}_k^L|\mathbf{z}_k, \mathbf{x}_{k-1}^L, \mathcal{S}_{k-1})$. This incremental upgrading problem is addressed by the *binary Bayesian filter* using the inverse measurement model and the *log odds ratio* introduced in [31]. $\text{I}(\mathbf{p}_j)$ is the LiDAR intensity of the point $\mathbf{p}_j$. u_s and σ_s are the mean and variance value of the LiDAR intensity of the hit cell, respectively. To better ensure the registration performance, we use cubic interpolation to obtain the probability and the intensity values from the submap. The variance σ_{o_i} and σ_{r_i} are used to weight the probability and intensity terms in the optimization at different resolutions of the submap.

The MAP estimate corresponds to the minimum of the negative log-posterior and the latter can be written as a sum of squared residual errors, yielding a non-linear least squares optimization problem. It can be minimized using iterative algorithms (e.g. Levenberg-Marquardt, Gauss-Newton), implemented in solvers, for example, Ceres [32].

B. Pose Graph Fusion

While a LiDAR inertial odometry can provide good relative constraints in the local frame, we still need global constraints to achieve global localization. The LGM module furnishes our system with global cues and our LGM implementation follows [5]. These complementary local and global cues are jointly optimized in a pose graph based fusion framework introduced in this section.

We formulate this fusion problem as a MAP estimation and factorize the posterior probability assuming uniform prior distribution as:

$$P(\mathcal{X}|\mathcal{Z}) \propto \prod_{k,s} P(\mathbf{z}_{ks}^O|\mathbf{x}_k^L, \mathbf{x}_s^S) \prod_k P(\mathbf{z}_k^I|\mathbf{x}_k^L, \mathbf{x}_{k-1}^L) \\ \prod_k P(\mathbf{z}_k^G|\mathbf{x}_k^L, \mathbf{x}_L^G). \quad (5)$$

The factorization can be visualized as a Bayesian network shown in Figure 3. If we assume zero-mean Gaussian-like probability, the likelihoods of the measurements can be defined as:

$$\begin{cases} P(\mathbf{z}_{ks}^O|\mathbf{x}_k^L, \mathbf{x}_s^S) \propto \exp{-\frac{1}{2}\|\mathbf{r}_{ks}^O\|_{\Lambda^O}^2} \\ P(\mathbf{z}_k^I|\mathbf{x}_k^L, \mathbf{x}_{k-1}^L) \propto \exp{-\frac{1}{2}\|\mathbf{r}_k^I\|_{\Lambda_k^I}^2} \\ P(\mathbf{z}_k^G|\mathbf{x}_k^L, \mathbf{x}_L^G) \propto \exp{-\frac{1}{2}\|\mathbf{r}_k^G\|_{\Lambda_k^G}^2}, \end{cases} \quad (6)$$
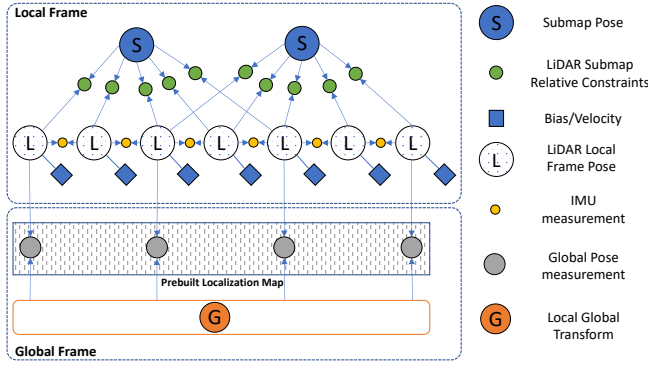
Fig. 3: The Bayesian network of the joint fusion problem. States and measurements are plotted in different colors. The factorization of the Bayesian network assumes uniform prior distribution.

where $\mathbf{r}_{ks}^O$, $\mathbf{r}_k^I$ and $\mathbf{r}_k^G$ are the odometry, inertial and global residuals, respectively.

In our fusion framework, we maintain a local frame $\mathbf{x}_k^L$ and a global state variable $\mathbf{x}_k^G$, which transfers $\mathbf{x}_k^L$ to the global frame. If we define $\mathbf{x}_k^L = [\mathrm{R}_k^L, \mathbf{t}_k^L]$, $\mathbf{x}_L^G = [\mathrm{R}_L^G, \mathbf{t}_L^G]$, and the global pose measurements $\mathbf{z}_k^G = [\mathrm{R}_k^G, \mathbf{t}_k^G]$ output by the LGM module [5], accordingly, we have the global residual cost as $(\mathbf{r}_k^G)^T = [\mathrm{Log}^T(\mathrm{R}_{rG}), \mathbf{t}_{rG}^T]$, where:

$$\begin{bmatrix} \mathrm{R}_{rG} & \mathbf{t}_{rG} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathrm{R}_k^G & \mathbf{t}_k^G \\ \mathbf{0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{R}_L^G & \mathbf{t}_L^G \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathrm{R}_k^L & \mathbf{t}_k^L \\ \mathbf{0} & 1 \end{bmatrix}. \quad (7)$$

With regard to the covariance Λ_k^G in the global residual, we write it as

$$\Lambda_k^G = \mathrm{diag}(\Lambda^{G_\omega}, \Lambda_k^{G_h}, \Lambda^{G_z}), \quad (8)$$

where the rotation and the altitude covariance, $\Lambda^{G_\omega} \in \mathbb{R}^{3\times3}$ and $\Lambda^{G_z} \in \mathbb{R}^{1\times1}$, respectively, are constant diagonal matrices since our LGM module only estimates the horizontal localization uncertainty using a 2D histogram filter as discussed in [5]. As shown by Equation 12 from [5], the horizontal matching covariance matrix $\Lambda_k^{G_h} \in \mathbb{R}^{2\times2}$ is calculated for each frame k accordingly. Note that the estimation uncertainty introduced here is critical to the performance of the localization system, yielding an adaptive fusion.

Concerning the odometry residual cost, we keep terminating old and creating new submaps as our LiDAR odometry propagates. The relative pose constraints $\mathbf{z}_{ks}^O$ between the local frame and the submap output by the LIO module are established and saved in the life cycle of the sliding window. Similarly, if we define the submap poses $\mathbf{x}_s^S = [\mathrm{R}_s^S, \mathbf{t}_s^S]$, and $\mathbf{z}_{ks}^O = [\mathrm{R}_{ks}^O, \mathbf{t}_{ks}^O]$, we have the odometry residual cost as $(\mathbf{r}_{ks}^O)^T = [\mathrm{Log}^T(\mathrm{R}_{rO}), \mathbf{t}_{rO}^T]$, where

$$\begin{bmatrix} \mathrm{R}_{rO} & \mathbf{t}_{rO} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathrm{R}_{ks}^O & \mathbf{t}_{ks}^O \\ \mathbf{0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{R}_s^S & \mathbf{t}_s^S \\ \mathbf{0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{R}_k^L & \mathbf{t}_k^L \\ \mathbf{0} & 1 \end{bmatrix}. \quad (9)$$

In terms of the covariance Λ^O in the odometry residual, we use a globally constant diagonal matrix across all frames and submaps assuming the estimation uncertainty is evenly distributed among all frames.

The pre-integration of the inertial constraints are treated in the same way as introduced in [30]. The non-linear least squares optimization is solved using the Ceres solver [32].

C. Environmental Change Detection

It is known that the LGM module relies on the freshness of the prebuilt localization map. We always seek to find out when it is time to update our localization map. The submaps generated in the LIO module not only can be used to estimate the relative pose between frames but also can be used to detect the environmental changes. This is crucial for a localization system, especially if we practically consider its commercial deployment. Therefore, we build an environmental change detection (ECD) module based on the submaps from our LIO. Firstly, we project our 3D multi-resolution submap onto the ground plane similar to the localization map building procedures in [6], [2], [5]. Then, a 2D submap in the same representation format to our prebuilt localization map is obtained. Given the localization outputs of the system, we can overlay the submap $\mathcal{S}$ onto the prebuilt map $\mathcal{M}$ and the occurrence of the environmental change can be determined by comparing them cell by cell. Let us denote u_s, σ_s, a_s, u_m, σ_m, and a_m as the intensity mean, intensity variance, and altitude mean of a pair of corresponding cells in the submap and the prebuilt map, respectively. We have

$$z_s(r) = \frac{(u_s - u_m)^2(\sigma_s^2 + \sigma_m^2)}{\sigma_s^2 \sigma_m^2} \quad (10)$$
$$z_s(a) = (a_s - a_m)^2,$$

where $z_s(r)$ and $z_s(a)$ evaluate how likely the environmental change exists in the cell.

We formulate the occurrence of the change within each cell as a binary state estimation problem addressed by the *binary Bayesian filter*. Similar to the occupancy grid update, if several vehicles have passed the same area multiple times generating more submaps, we can similarly update the probability of the change occurrence by using a *binary Bayesian filter* [31]. If we denote the occurrence of the change within a cell as a binary state variable d_s, the inverse measurement model for updating the binary Bayesian filter can be defined as:

$$P(d_s|z_s) = \eta P(d_s|z_s(r))^\gamma P(d_s|z_s(a))^{1-\gamma}, \quad (11)$$

where γ is a dynamic weighting parameter that balances the weights of the intensity and altitude cues. We can further define their inverse measurement model by normalizing $z_s(r)$ and $z_s(a)$ within the submap as:

$$\begin{cases} P^{-1}(d_s|z_s(r)) = 1 + \exp\left(-\beta_1\left(\frac{z_s(r)-u_s(r)}{\sigma_s(r)} - \theta_1\right)\right) \\ P^{-1}(d_s|z_s(a)) = 1 + \exp\left(-\beta_2\left(\frac{z_s(a)-u_s(a)}{\sigma_s(a)} - \theta_2\right)\right), \end{cases} \quad (12)$$

where $u_s(r)$, $u_s(a)$, $\sigma_s(r)$, $\sigma_s(a)$ are the mean and variance of $z_s(r)$ and $z_s(a)$, respectively, aggregated in a submap. β and θ are parameters set empirically.

IV. Experimental Results

A. Platforms and Datasets

Our system has been extensively tested in real-world driving scenarios primarily using two datasets, our internal dataset, and the Apollo-SouthBay dataset [33], [34]. While the Apollo-SouthBay dataset was collected in San Francisco Bay Area, United States, covering a driving distance of 380.5km, our internal dataset includes the various challenging urban scenarios in Beijing, China, especially the ones with map errors or environmental changes, such as map expiration (gradual environment change), road construction, closed lanes, dense traffic and so on. These challenges might be unusual or can be avoided by altering the operational routes in other parts of the world, but are actually quite common and happen daily in different parts of Beijing, the capital city of one of the largest developing countries - China.

Our vehicle platform is equipped with a Velodyne HDL-64E 360° LiDAR and a NovAtel PwrPak7D-E1 GNSS RTK receiver integrated with dual antennas and an Epson EG320N IMU. These sensors are all mounted on a Lincoln MKZ vehicle. The ground truth poses used in the evaluation are generated using offline LiDAR SLAM methods typically formulated as a large-scale global least-square optimization problem, which are beyond the scope of this work.

B. Localization Performance

The localization performance is compared against Wan et al. [5], in which 2-System mode is applied because LiDAR-based localization is our primary focus. Our quantitative analysis includes horizontal and heading errors with both RMS and maximum values. The horizontal errors are further decomposed to longitudinal (Long.) and lateral (Lat.) directions in Table II. The percentages of frames where the system achieves better than 0.1m, 0.2m, 0.3m or $0.1°$ in horizontal or heading estimations are shown in the tables.

Apollo-SouthBay Dataset The quantitative analysis using the Apollo-SouthBay dataset is shown in Table I. Overall, our system is comparable to [5] under all metrics. Note, that our system achieves better maximum horizontal and yaw RMS errors in most of the data sequences demonstrating the benefits of new design in challenging scenarios particularly.

Internal Dataset The quantitative analysis using the internal dataset is shown in Table II. Overall, note our performance improvement over [5] by having the new LiDAR inertial odometry aided framework.

To further illustrate the benefits obtained, Figure 4 shows sample outputs of the system from a few frames in the dataset. In Figure 4 (a) and (b) from Seq. DXH-02 and DXH-03, respectively, there are long-term seasonal changes in the environment. The driving test was taken in winter, while the map was built in summer as shown on the top panel. The bottom panel shows the localization status of the LGM module during these situations. It clearly demonstrates that the newly added LIO module can help localization to account for unstable global matching. The global matching uncertainty estimated by the histogram filters in the LGM

Area	Method	Horiz. RMS	Horiz. Max	< 0.1m Pct.	Yaw. RMS	< 0.1° Pct.
BaylandsToSeafood	[5]	**0.036**	0.203	98.87%	0.055	86.60%
	Ours	0.041	**0.150**	**99.05%**	**0.042**	**90.16%**
ColumbiaPark	[5]	0.046	0.160	**96.46%**	**0.081**	**67.27%**
	Ours	0.046	**0.143**	95.79%	0.095	55.88%
Highway237	[5]	**0.048**	0.196	93.18%	0.070	76.83%
	Ours	0.049	**0.129**	**96.86%**	**0.026**	**98.40%**
MathildaAVE	[5]	**0.040**	0.179	98.67%	**0.059**	**83.37%**
	Ours	0.048	**0.139**	**99.08%**	0.076	75.36%
SanJoseDowntown	[5]	0.058	0.290	87.47%	0.052	87.57%
	Ours	**0.056**	**0.264**	**90.46%**	**0.045**	**91.03%**
SunnyvaleBigLoop	[5]	**0.069**	0.368	**80.86%**	0.081	69.52%
	Ours	0.075	**0.306**	78.72%	**0.077**	**74.76%**

TABLE I: Quantitative comparison using the Apollo-SouthBay dataset. The lower maximum horizontal errors demonstrate that the proposed designs smooth the trajectories and make our system reliable in extreme circumstances.

module shown in the bottom left view of the bottom panel is the key to accomplish an adaptive LIO and LGM fusion. Figure 4 (c) (d) and (e) shows the benefits from LIO module in a similar way. In Figure 4 (c) from Seq. YF-01 shows metal walls were removed and crosswalks are repainted on the road. Figure 4 (d) from Seq. MXH, there are newly built walls on the left that are quite common when road construction happens. Figure 4 (e) from Seq. YF-04 shows a challenging case with many heavy trucks on a narrow road.

Run-time Analysis In our system, the LIO and PGF modules run on the CPU (Intel Xeon 3.5GHz CPU and 32GB RAM), while the LGM module runs on an FPGA card. It takes 44.9ms, 79.4ms, 2.8ms in LGM, LIO, and PGF module, respectively. Note, that the LGM and LIO modules actually run parallelly. Therefore, the entire system takes about 82ms per frame on average, yielding a real time system.

C. Ablations

In order to comprehensively verify the effectiveness of the proposed LIO module, we conduct several ablation experiments using our internal dataset.

LIO Drift To evaluate the performance of the LIO module individually, we manually disable the LGM module after system initialization. We adopt the widely used KITTI odometry benchmark metrics [35] evaluated by using the original toolkit, The LIO module obtains an average translation error of 0.9309% and rotation error of 0.0057deg/m, respectively.

Intensity Cue and Multi-Resolution In Section III, we introduce the intensity cue and the multi-resolution occupancy grid as our extensions to the LiDAR inertial odometry implementation. To demonstrate the effectiveness of each of the above contributions, we evaluate the localization accuracy with different settings in Table III. Evidently, the accuracy is improved when we incorporate the intensity cue or finer and more resolutions are used in our submap implementation.

D. Environmental Change Detection

Environmental changes can be detected by comparing the submaps generated by the LIO module against the prebuilt map. As shown in Figure 1, the likelihood environmental

Scenes	Dist. (km)	Method	Horiz. RMS	Horiz. Max	Long. RMS	Long. Max	Lat. RMS	Lat. Max	< 0.1m Pct.	< 0.2m Pct.	< 0.3m Pct.
DXH-01	11.03	[5]	0.056	0.803	0.037	0.311	0.034	0.740	91.42%	**99.08%**	**99.50%**
		Ours	**0.048**	**0.472**	**0.029**	**0.196**	**0.031**	**0.459**	**95.31%**	99.07%	99.44%
DXH-02	1.693	[5]	0.087	0.870	0.052	0.666	0.063	0.735	76.67%	94.02%	99.44%
		Ours	**0.074**	**0.260**	**0.041**	**0.142**	**0.054**	**0.256**	**81.49%**	**94.31%**	**100.0%**
DXH-03	1.615	[5]	0.093	0.402	0.080	0.401	0.036	0.202	61.54%	99.20%	99.93%
		Ours	**0.085**	**0.168**	**0.076**	**0.167**	**0.033**	**0.111**	**72.76%**	**100.0%**	**100.0%**
MXH	2.563	[5]	0.086	1.669	0.047	1.023	0.064	1.421	68.75%	99.15%	99.18%
		Ours	**0.069**	**0.223**	**0.036**	**0.126**	**0.052**	**0.223**	**89.52%**	**99.52%**	**100.0%**
YZ-01	7.263	[5]	0.070	2.841	**0.045**	0.233	0.046	2.836	85.43%	99.12%	99.65%
		Ours	**0.063**	**0.313**	0.047	**0.184**	**0.034**	**0.294**	**91.66%**	**99.26%**	**99.96%**
YZ-02	4.638	[5]	0.069	0.336	**0.039**	0.194	0.050	0.316	77.61%	**99.61%**	99.92%
		Ours	**0.062**	**0.274**	0.040	**0.169**	**0.041**	**0.251**	**86.43%**	99.60%	**100.0%**
YZ-03	2.567	[5]	0.064	**0.235**	0.046	**0.190**	0.037	0.228	83.88%	99.71%	100.0%
		Ours	**0.059**	0.250	**0.043**	0.250	**0.032**	**0.186**	**92.04%**	**99.92%**	100.0%
YZ-04	0.911	[5]	0.177	1.111	0.129	0.982	0.105	1.010	53.95%	80.16%	84.45%
		Ours	**0.085**	**0.365**	**0.065**	**0.186**	**0.042**	**0.361**	**68.63%**	**95.53%**	**98.44%**

TABLE II: Quantitative comparison using our internal dataset. The benefits of the LiDAR inertial odometry are clearly visible from the localization accuracy enhancement.
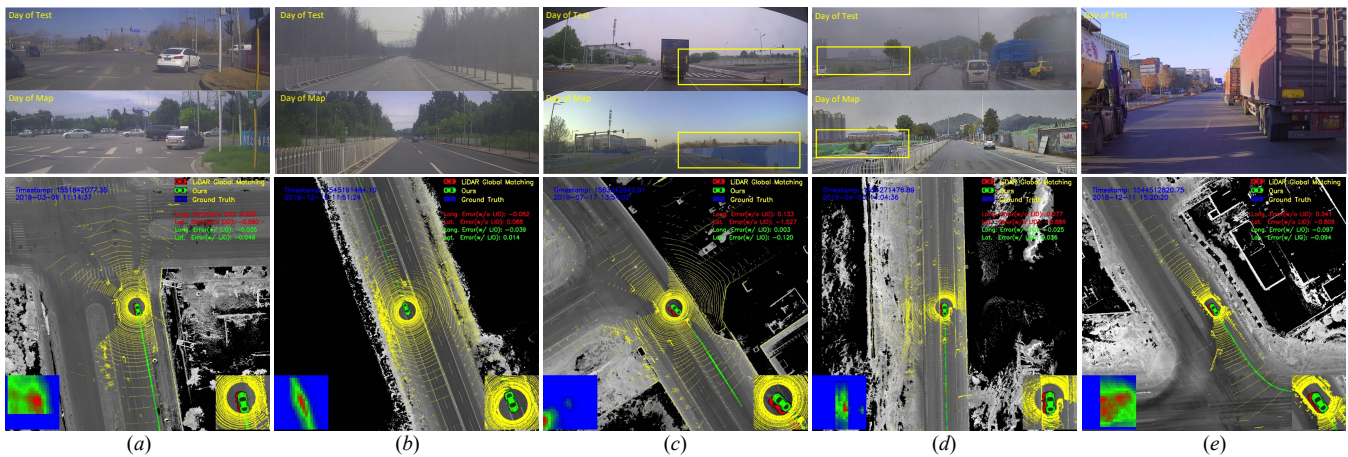

Fig. 4: Top panel: (a) and (b) long-term seaonal changes from DXH-02 and DXH-03, respectively. The driving test was taken in winter, while the map was built in summer. (c) metal walls were removed and crosswalks are repainted from YF-01. (d) new walls were built on the left from MXH. (e) dynamic objects, such as many heavy trucks, on a narrow road from YF-04. Bottom panel: visual comparison of localization results. Green cars are the fused results estimated in our joint optimization framework. Red cars are the results generated only by the LGM module. Blue cars are the ground truth. When the green cars perfectly overlap with the blue ones (can not be seen), the results are accurate. The bottom left view in the bottom panel depicts the histogram filter response in the LGM module.

Method	Translation (%)	Rotation (deg/m)
Ours w/ Intensity	**0.9309**	**0.0057**
{0.125, 0.25, 0.6, 1.2}	0.9498	0.0057
{0.125}	0.9520	0.0058
{0.6}	1.7637	0.0102

TABLE III: Comparison w/o the intensity cue or multi-resolution. Note that using coarser or single resolution grid gives worse results. Ours with the intensity cue yields the best odometry accuracy.

change exists in each map cell is marked on the map using a rainbow scale bar. An intersection where metal walls were recently removed is zoomed-in to further demonstrate the details of environmental change detection results.

V. CONCLUSION

We have presented a robust LiDAR localization framework, designed for autonomous driving applications, espe-cially countering the localization challenges in changing city scenes. The proposed method utilizes a pose graph based fusion framework adaptively incorporating results from both the LiDAR inertial odometry and global matching modules. It has been shown that the newly added LiDAR inertial odometry module can effectively aid the localization system in the prevention of localization errors caused by the global matching failure. We also propose an environmental change detection method to find out when and what portion of the map should be promptly updated in order to reliably support the daily operation of our large autonomous driving fleet in crowded city streets, despite road construction that may have occurred from time to time. These distinct advantages make our system quite attractive to real commercial deployment. In a further extension of this work, we plan to explore learning-based environmental change detection methods for comparison.

REFERENCES

[1] "Baidu Apollo open platform," http://apollo.auto/.

[2] R. W. Wolcott and R. M. Eustice, "Fast LiDAR localization using multiresolution gaussian mixture maps," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 2814–2821.

[3] ——, "Robust LiDAR localization using multiresolution gaussian mixture maps for autonomous driving," *The International Journal of Robotics Research (IJRR)*, vol. 36, no. 3, pp. 292–319, 2017.

[4] M. Aldibaja, N. Suganuma, and K. Yoneda, "Robust intensity-based localization method for autonomous driving on snow-wet road surface," *Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2369–2378, 2017.

[5] G. Wan, X. Yang, R. Cai, H. Li, Y. Zhou, H. Wang, and S. Song, "Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 4670–4677.

[6] J. Levinson, M. Montemerlo, and S. Thrun, "Map-based precision vehicle localization in urban environments." in *Proceedings of the Robotics: Science and Systems (RSS)*, vol. 4, 2007, p. 1.

[7] C. McManus, W. Churchill, W. Maddern, A. D. Stewart, and P. Newman, "Shady dealings: Robust, long-term visual localisation using illumination invariance," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 901–906.

[8] C. Linegar, W. Churchill, and P. Newman, "Work smart, not hard: Recalling relevant experiences for vast-scale but time-constrained localisation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 90–97.

[9] ——, "Made to measure: Bespoke landmarks for 24-hour, all-weather localisation with a camera," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 787–794.

[10] M. Bürki, I. Gilitschenski, E. Stumm, R. Siegwart, and J. Nieto, "Appearance-based landmark selection for efficient long-term visual localization," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 4137–4143.

[11] T. Sattler, W. Maddern, C. Toft, A. Torii, L. Hammarstrand, E. Stenborg, D. Safari, M. Okutomi, M. Pollefeys, J. Sivic *et al.*, "Benchmarking 6DOF outdoor visual localization in changing conditions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 8601–8610.

[12] J. Zhang and S. Singh, "LOAM: LiDAR odometry and mapping in real-time." in *Proceedings of the Robotics: Science and Systems (RSS)*, vol. 2, 2014, p. 9.

[13] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LiDAR SLAM," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 1271–1278.

[14] J. Zhang and S. Singh, "Low-drift and real-time LiDAR odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[15] C. Park, S. Kim, P. Moghadam, C. Fookes, and S. Sridharan, "Probabilistic surfel fusion for dense LiDAR mapping," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2418–2426.

[16] J. Behley and C. Stachniss, "Efficient surfel-based SLAM using 3D laser range data in urban environments," in *Proceedings of the Robotics: Science and Systems (RSS)*, 2018.

[17] D. Droeschel and S. Behnke, "Efficient continuous-time SLAM for 3D LiDAR-based online mapping," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 1–9.

[18] C. Park, P. Moghadam, S. Kim, A. Elfes, C. Fookes, and S. Sridharan, "Elastic LiDAR fusion: Dense map-centric continuous-time SLAM," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 1206–1213.

[19] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and ground-optimized LiDAR odometry and mapping on variable terrain," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4758–4765.

[20] J.-E. Deschaud, "IMLS-SLAM: Scan-to-model matching based on 3D data," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 2480–2485.

[21] Q. Li, S. Chen, C. Wang, X. Li, C. Wen, M. Cheng, and J. Li, "LO-Net: Deep real-time LiDAR odometry," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2019, pp. 8473–8482.

[22] Y. Cho, G. Kim, and A. Kim, "DeepLO: Geometry-Aware Deep LiDAR Odometry," *arXiv preprint arXiv:1902.10562*, 2019.

[23] H. Ye, Y. Chen, and M. Liu, "Tightly coupled 3D lidar inertial odometry and mapping," *arXiv preprint arXiv:1904.06993*, 2019.

[24] C. Qin, H. Ye, C. E. Pranata, J. Han, and M. Liu, "LINS: A LiDAR-Inerital state estimator for robust and fast navigation," *arXiv preprint arXiv:1907.02233*, 2019.

[25] Y. Gao, S. Liu, M. Atia, and A. Noureldin, "INS/GPS/LiDAR integrated navigation system for urban and indoor environments using hybrid scan matching algorithm," *Sensors*, vol. 15, no. 9, pp. 23 286–23 302, 2015.

[26] H. Liu, Q. Ye, H. Wang, L. Chen, and J. Yang, "A precise and robust segmentation-based LiDAR localization system for automated urban driving," *Remote Sensing*, vol. 11, no. 11, p. 1348, 2019.

[27] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *The International Journal of Robotics Research(IJRR)*, vol. 25, no. 12, pp. 1181–1203, 2006.

[28] H. Strasdat, J. Montiel, and A. J. Davison, "Real-time monocular SLAM: Why filter?" in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 2657–2664.

[29] A. Soloviev, "Tight coupling of GPS, laser scanner, and inertial measurements for navigation in urban environments," in *Proceedings of the IEEE/ION Position, Location and Navigation Symposium*, 2008, pp. 511–525.

[30] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.

[31] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[32] S. Agarwal, K. Mierle, and Others, "Ceres solver," http://ceres-solver.org.

[33] W. Lu, Y. Zhou, G. Wan, S. Hou, and S. Song, "L3-Net: Towards learning based LiDAR localization for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 6389–6398.

[34] W. Lu, G. Wan, Y. Zhou, X. Fu, P. Yuan, and S. Song, "DeepVCP: An end-to-end deep neural network for point cloud registration," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019.

[35] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 3354–3361.