EVALUATION OF A COMPACT HELMET-BASED LASER SCANNING SYSTEM FOR ABOVEGROUND AND UNDERGROUND 3D MAPPING

Jianping Li 1, Bisheng Yang 1, * , Yiping Chen 2, Weitong Wu 3, Yandi Yang 3, Xin Zhao 1, Ruibo Chen 3

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China – (lijianping; bshyang) @whu.edu.cn
2 Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China – chenyiping@xmu.edu.cn
3 Guangxi Zhuang Autonomous Region Remote Sensing Institute of Nature Resources, Nanning 53000, China – ruiboil@163.com

Commission II, WG II/3

KEY WORDS: HLS (helmet-based laser scanning); ubiquitous point cloud; 3D mapping; SLAM; urban underground space;

ABSTRACT:
As a strategic resource, urban underground space can be used for rail transportation, commercial streets, which has high economic and social benefits, and is of great significance to sustainable city development. Due to denied Global Navigation Satellite System (GNSS) signal, traditional mobile mapping systems have difficulty collecting accurate 3D point clouds in urban underground space. Thus, a helmet-based laser scanning system, named "WHU-Helmet", is integrated in this paper to make up for the shortcomings of the existing traditional mobile mapping systems. "WHU-Helmet" is mainly equipped with four types of sensors: a GNSS receiver (optional), an IMU, a laser scanner, and a global shutter camera. "WHU-Helmet" is not relying on GNSS signal and has the advantages of low cost, small volume and easy operation. Using "WHU-Helmet", a multi-scale Normal Distributions Transform (NDT) based LiDAR-IMU SLAM is implemented to collect underground 3D point cloud in real-time. To validate the performance of "WHU-Helmet" in aboveground and underground 3D mapping experiments, were conducted in a typical urban metro station. The experiments show that the average and RMSE of HLS point errors of "WHU-Helmet" are 0.44 meters and 0.23 meters, respectively, showing great potential of "WHU-Helmet" in the application of aboveground and underground 3D mapping.

1. INTRODUCTION
With the improvement of urbanization, the contradiction between rapid urban development and limited land resources is becoming prominent (Von der Tann et al., 2020). As a strategic resource, urban underground space can be used for rail transportation, commercial streets, and other public infrastructures. Thus, urban underground space has high economic and social benefits, which is of great significance to sustainable city development. Rational utilization of urban underground space, promoting the collective development of aboveground and underground, is a promising solution for improving efficiency use of urban land, reducing urban population density, and expanding capacity of public infrastructure (Qiao et al., 2019). Nevertheless, 3D mapping of the urban underground space is the premise of reasonable planning and maintenance of urban underground resources.

Mobile mapping system (MMS) is one of the most advanced 3D mapping technologies in the field of photogrammetry, and has been widely used in urban infrastructure digitalization (Dong et al., 2018; Mi et al., 2021). Traditional mobile mapping systems (e.g., UAV laser scanning system (Li et al., 2019), car-based laser scanning system (Jaakola et al., 2010), and et al.) are mainly equipped with two types of sensors, namely, position and orientation system (POS), and laser scanner. Using direct-georeferencing technology (Skaloud and Legat, 2008), the observations of laser scanner could be transformed to the mapping system. However, the POS is relying on the Global Navigation Satellite System (GNSS), and could not be applied in the urban underground space. Long-time and accurate 3D mapping in large-scale GNSS denied underground environment is a research hotspot in both academia and industry (Rouček et al., 2019).

In recent years, a lot of wheeled robot-based systems for 3D mapping in GNSS-denied environments using simultaneous localization and mapping (SLAM) are developed (Chang et al., 2019; Zhang and Singh, 2018). However, the large-weight and high-cost limits the wheeled robot-based mapping system in complex urban underground environments. Wearable mapping systems have the advantages of low cost, small volume and easy operation, which attract attention of the field of photogrammetry (Karam et al., 2020) and robotics (Alliez et al., 2020). Su et al. (2020) developed a backpack laser scanning system, and applied the system in forest inventory successfully. The handheld laser scanning system ZEB developed by GEOSLAM 1 has been applied in several applications, including forest inventory (Camarretta et al., 2021), building information system (Perviti et al., 2019), and protection of ancient buildings (Di Stefano et al., 2021). However, there is still no helmet-based laser scanning system. Thus, a compact helmet-based laser scanning (HLS) system, named WHU-Helmet, is integrated for aboveground and underground 3D mapping of a metro station in this paper.

The remainder of this paper is organized as follows: the hardware description of the HLS system is elaborated in Section 2. A multi-scale Normal Distributions Transform (NDT) based LiDAR-IMU SLAM is implemented in Section 3. In Section 4, the experimental studies are undertaken to evaluate the point cloud accuracy collected by the HLS, after which conclusions are drawn at the end.

* Corresponding author
1 https://geoslam.com/
2. HARDWARE DESCRIPTION OF THE HELMET-BASED LASER SCANNING SYSTEM

2.1 Sensor configuration

The WHU-Helmet is composed of four types of sensors: GNSS receiver, MEMS-based IMU, global shutter camera, and solid-state LiDAR. Each sensor is integrated in the HLS as illustrated in Figure 1. The GNSS receiver is used to obtain absolute geolocations and reference time for the whole system. A micro-electro-mechanical system (MEMS)-based inertial measurement unit (IMU) is used to propagate the system initial position and orientation continuously. A solid-state LiDAR and a global shutter camera are integrated to collect geometry and optical information from the underground environment. Besides, visual and laser features are extracted to constrain the positioning drift of the IMU using SLAM. All the sensors are time-synchronized electronically referencing to the GNSS time according to our previous solution (Li et al., 2019). The total weight of the HLS is about 1.5 kg, which is compact and easy to operate.

2.2 3D mapping system definitions

The 3D coordinate systems involved in the HLS include mapping frame, body frame, and LiDAR frame, which are illustrated in Figure 2. As for a LiDAR observation $r_p^m$ in the LiDAR frame, the corresponding point in the mapping frame could be obtained by:

$$
r_p^m = R_m^b r_p^b + R_m^b r_v^b + r_v^b (t)
$$

where, $R_m^b$ and $r_v^b (t)$ are the system orientation and position at time $t$ obtained by the SLAM algorithm. $R_m^b$ and $r_v^b$ are the calibration parameters between IMU and the solid-state LiDAR, which are pre-calibrated using similar strategy proposed in our previous work (Li et al., 2020).

3. MULTI-SCALE NDT BASED LIDAR-IMU SLAM

The workflow of the proposed multi-scale NDT based LiDAR-IMU SLAM is illustrated in Figure 3. After receiving time-synchronized IMU and LiDAR observations, three steps are involved: (1) IMU pre-integration and correction of motion distortion, (2) Multi-scale NDT based matching and (3) LiDAR and IMU fused optimization, which are detailed as follow:

3.1 IMU pre-integration and correction of motion distortion

As the motion distortion caused by the continuous observation mode of the laser scanner, the IMU measurements collected within the $k^{th}$ laser frame are used to correct the LiDAR motion distortion. The IMU pre-integration model (Qin et al., 2018) is used to calculate the relative motion as follow:

$$
r_v^b (t_{k+1}) = r_v^b (t_k) + v_v^b (t_k) \Delta t - \frac{1}{2} g \Delta t^2 + R_v^b (t_k) \int_{t_k}^{t_{k+1}} (\alpha(t) - b_v(t)) dt
$$

Figure 2. 3D mapping frame involved in the helmet-based laser scanning system.

Figure 3 Workflow of multi-scale NDT based LIDAR-IMU SLAM
\[ v^n_i(t_{i+1}) = v^n_i(t_i) - g^n \Delta t + R^n_{i+1}(t_i) \int_{t_i}^{t_{i+1}} [R^n_i(\dot{t})(\dot{a}(t) - b^n(a(t)))] dt \]

\[ q^n_i(t_{i+1}) = q^n_i(t_i) \otimes \int_{t_i}^{t_{i+1}} [\frac{1}{2} \Omega(\dot{a}(t) - b^n(a(t)))] dt \]

where, \( \dot{a}(t) \) and \( \dot{\omega}(t) \) are the accelerator and gyroscopic measurements, respectively. \( g^n \) is the gravity vector in the mapping frame. \( a^n_{i+1}, p^n_{i+1}, \) and \( \lambda^n_{i+1} \) are the pre-integration parts, which reflect the relative motion of the short time period during the \( k^n \) laser frame.

### 3.2 Multi-scale NDT based matching

NDT is a well-known LiDAR SLAM technology (Magnusson et al., 2007), which transform the point cloud registration problem to optimization the probability density function \( f(p) \) as follow:

\[ f(p) = \frac{1}{2(2\pi)^{\frac{p}{2}}} \exp \left( \frac{-\|p\|^2}{2} \right) \]

where, \( \mu \) is the mean value, \( \Sigma \) is the covariance matrix. As for the solid-state laser scanner equipped in the helmet-based laser scanning system, there is great differences in the point density of the one laser. The point density near the scanning center is large, on the contrary, the point density far away from the scanning center is small. It is hard to decide the voxel size used for NDT. In order to ensure that there are enough points in the distant voxels to accurately calculate the covariance, the voxel size needs to be set large, resulting in the low resolution of the near grid and the loss of the details. Therefore, it is difficult to select the appropriate voxel size to balance the details of near and far voxels.

In this paper, multi-scale normal distribution transformation is adopted to overcome the above problem, as follows: (1) the voxel size is set to \( S_{\text{oxel}} \) (0.5 m used in the experiment), then calculate the mean value, covariance matrix, point size, eigen vector, and the geometric attributes (linear or planar or irregular)(Magnusson, 2009); (2) Carry out iterative merging of voxels according to the merging conditions listed in Table 1; (3) all parameters in the merged voxels are updated.

### Table 1. Voxel merging conditions.

| Voxel size after merging is less than \( S_{\text{oxel}} \) (2 m used in the experiment); and |
| Two voxels are linear before merging, they are still linear after merging; or |
| Two voxels are planar before merging, they are still planar after merging; or |
| Two voxels are irregular before merging; or |
| One voxel is linear, another one is irregular, they are linear after merging; or |
| One voxel is planar, another one is planar, they are planar after merging; or |

### 3.3 State estimation

In this paper, two kinds of data are used to estimate the state parameters of helmet-based laser scanning system in real time, namely, IMU pre-integration constraint \( e^{k+1}_{\text{Pre-int}} \) and NDT constraint \( e^{k+1}_{\text{NDT}} \). NDT constraint \( e^{k+1}_{\text{NDT}} \) is similar with existing works (Magnusson et al., 2007).

As for the IMU pre-integration constraints, which are derived according to Eq. (2-4) as follow:

\[ e^{k+1}_{\text{Pre-int}} = \begin{bmatrix} \delta x_{k+1} \\ \delta \beta_{k+1} \\ \delta \delta_{k+1} \end{bmatrix} = \begin{bmatrix} R^n_i(t_i) (v^n_{i+1}(t_{i+1}) - v^n_i(t_i) + \frac{1}{2} g^n \Delta t^2) - \lambda^n_{i+1} \\ R^n_i(t_i) (v^n_{i+1}(t_{i+1}) - v^n_i(t_i) + \frac{1}{2} g^n \Delta t^2) - \lambda^n_{i+1} \\ 2[\delta q^n_i(t_{i+1}) - \delta q^n_i(t_i) \otimes \lambda^n_{i+1}] \end{bmatrix} \]

![Image](https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-215-2022)

Considering the real-time 3D mapping requirements of the HLS system, sliding window (Huang et al., 2011) is used. The historical laser frame and IMU information are marginalized and transformed into a priori constraints \( e_{\text{Margin}} \). These three kinds of constraints constitute the energy function \( E \), as follow:

\[ E = \sum e^{k+1}_{\text{Pre-int}} e^{k+1}_{\text{Pre-int}} + \sum e^k_{\text{NDT}} e^k_{\text{NDT}} + e^k_{\text{Margin}} e^k_{\text{Margin}} \]

Specifically, to solve \( E \), the general least square equation is as follow:

\[ H \delta x = b \]

where, \( H \) is the Hessian matrix, \( \delta x \) is correction value for system states, \( b \) is the error terms. \( \delta x \) is consisted of two parts: \( \delta x = [\delta x_{\text{Margin}}, \delta x_{\text{remain}}, \delta x_{\text{Marginal}}] \). \( \delta x_{\text{Marginal}} \) is the historical state parameters to be deleted, \( \delta x_{\text{remain}} \) is the current state parameters to be estimated. Then we could rewrite Eq. (8):

\[ \left[ \begin{array}{c} H_{\text{Marg}} \\ H_{\text{Remain}} \\ H_{\text{Corr}} \end{array} \right] \left[ \begin{array}{c} \delta x_{\text{Marg}} \\ \delta x_{\text{Remain}} \end{array} \right] = \left[ \begin{array}{c} b_{\text{Marg}} \\ b_{\text{Remain}} \end{array} \right] \]

According to Schul complementation, Eq. (9) is rewritten as:

\[ \left[ \begin{array}{cc} H_{\text{Marg}} & 0 \\ H_{\text{Remain}} & H_{\text{Corr}} \end{array} \right] \left[ \begin{array}{c} \delta x_{\text{Marg}} \\ \delta x_{\text{Remain}} \end{array} \right] = \left[ \begin{array}{c} b_{\text{Marg}} \\ b_{\text{Remain}} - H_{\text{Corr}} H_{\text{Marg}}^{-1} b_{\text{Marg}} \end{array} \right] \]

Without solving \( \delta x_{\text{Marg}} \), \( \delta x_{\text{remain}} \) is solved with considering historical constraints. The marginalization term \( e_{\text{Margin}} \) is obtained as follow:

\[ e_{\text{Margin}} \simeq H_{\text{Corr}} \delta x_{\text{Remain}} + b_{\text{Remain}} - H_{\text{Corr}} H_{\text{Marg}}^{-1} b_{\text{Marg}} \]

### 4. EXPERIMENTS

#### 4.1 Study area and data collection

The study area as shown in Figure 4 is located in Computer Town Metro Station (30.53 N, 114.36 E), Wuhan City, Hubei Province, China. The operator carried the HLS, and walked at the speed of 1.3 m/s to scan the environment from aboveground to the underground using about 20 minutes. The real-time 3D mapping results are illustrated in Figure 5, Figure 6, Figure 7, and Figure 8.
8, which are rendered according to intensity in the RVIZ ². First the operator walked along the roadside to the metro station entrance 1. Then the operator walked into the metro station through the elevator. It took about 10 minutes to collect underground data in the metro station. At last, the operator walked outside the metro station through the metro station entrance 2.

Figure 4. Study area in Wuhan. (a) Location of Computer Town metro station in satellite image; (b) Snapshot of the metro station aboveground; (c) Snapshot of the metro station underground;

Figure 5. Overview of the aboveground and underground point clouds collected by the helmet-based laser scanning system.

² http://wiki.ros.org/rviz
4.2 Accuracy evaluation

To validate the geometry accuracy of the HLS in aboveground and underground 3D mapping, point clouds were collected using terrestrial laser scanning (TLS) too. The registered multiple scans from TLS are served as references. 30 evenly distributed corresponding corner points are selected from both TLS point clouds and HLS point clouds. The error distributions of the 30 corresponding corner points are plotted in Figure 9 and listed in Table 2. The average and RMSE of the corresponding point errors are 0.44 meters and 0.23 meters, which has shown a good potential of the HLS for the accurate digitalization of the urban underground space.

| Study area    | MIN | MAX | AVERAGE | RMSE |
|---------------|-----|-----|---------|------|
| Metro station | 0.19| 0.81| 0.44    | 0.23 |
This contribution has been peer-reviewed.

https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-215-2022 | © Author(s) 2022. CC BY 4.0 License.