V-SLAM Enhanced INS/GNSS Fusion Scheme for Lane Level Vehicular Navigation Applications in Dynamic Environment

Tung-Hua Yeh [1] *, Kai-wei Chiang [1], Pei-Ru Lu [1], Pei-Ling Li [2], You-Shyan Lin [3], Chih-Yuan Hsu [3]

[1] Dept. of Geomatics, National Cheng Kung University, Tainan, Taiwan – p66101061@gs.ncku.edu.tw,
kwchiang@geomatics.ncku.edu.tw, petty891206@gmail.com
[2] High Definition Map Research Center, Dept. of Geomatics, National Cheng Kung University, Tainan, Taiwan -
pointlpl@geomatics.ncku.edu.tw
[3] Automotive Research & Testing Center, Changhua, Taiwan. - jasonlin@artc.org.tw, evan@artc.org.tw

KEY WORDS: V-SLAM, Refreshed-SLAM, INS/GNSS Integration, Extended Kalman Filter, Dynamic Environment.

ABSTRACT:

With the development of different sensors, such as global navigation satellite system (GNSS), inertial measurement unit (IMU), LiDAR, radar and camera, more localization information is available for autonomous vehicular applications. However, each sensor has its limitations in different circumstances. For example, visual Simultaneous Localization and Mapping (SLAM) easily loses tracking in an open sky area where accurate GNSS measurements can be obtained. Sensors can complement each other by integrated their information in a multi-sensor fusion scheme. In this study, we proposed a visual-SLAM enhanced INS/GNSS localization fusion scheme for a high dynamic environment. Oriented FAST and rotated BRIEF (ORB) SLAM are used to pre-process image sequences from monocular camera, rescaled and refreshed after applying GNSS measurements, and convert to position and velocity information, which can provide updates to the system. The performance of the fusion system was verified through two field tests at different speed ranges (about 30~60 km/s), using a reliable reference system as ground-truth to assess the accuracy of the proposed localization fusion scheme. The results indicated that the proposed system could improve the navigation accuracy compared to INS/GNSS integration scheme and achieve which-lane level or even where-in-lane level.

1. INTRODUCTION

Navigation plays a crucial role for autonomous vehicular applications. To ensure the safety of self-driving, high accuracy of navigation states estimation should be promised. inertial navigation system (INS) and global navigation satellite system (GNSS) integration scheme is a typical example for accurate navigation. Nevertheless, the error of INS will accumulate over time; meanwhile, GNSS can't provide accurate position in challenging environment, such as urban area. Hence, many researchers have utilized the optical sensors which can collect the information from the surrounding environment, such as camera (Mourikis & Roumeliotis, 2007; Leutenegger et al., 2015; Cao et al., 2020) and LiDAR (Surachet & Chiu, 2022), to improve system accuracy and robustness.

Based on navigation accuracy requirements for intelligent transportation systems (ITSs), accuracy levels can be divided into the which-road, which-lane, and where-in-lane levels (Stephenson et al., 2011), as shown in Figure 1. At the which-road level, navigation must be accurate to within 5 m. To increase the level of navigation accuracy to the which-lane level, navigation must be accurate to within 1.5 m or less. At this level, the lane that a vehicle is on must be identified. To achieve fully autonomous driving, the exact position of a vehicle on a lane must be determined; therefore, navigation must be accurate to within <1 or even 0.5 m. Considering a normal speed range (about 30~60 km/s), it's difficult to achieve the accuracy of which lane level or where in lane level for a low-cost micro-electro-mechanical system (MEMS) inertial measurement unit (IMU) and GNSS chipset module. Hence, this paper proposed a visually enhanced INS/GNSS fusion scheme with GNSS-based refreshed Simultaneous Localization and Mapping (SLAM). The proposed method not only increased the navigation accuracy but

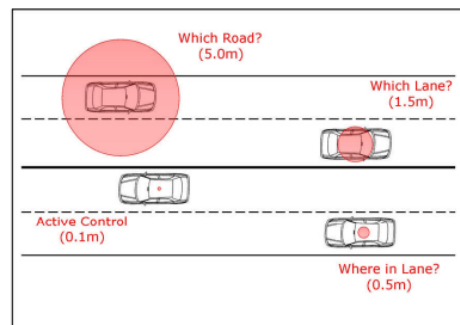also enhanced the relationship between the system and the real world.

Figure 1 Navigation accuracy classification (Stephenson et al., 2011)

A conventional vehicle positioning algorithm integrates an INS and a GNSS system through Extended Kalman Filter (EKF) with loosely coupled (LC) or tightly coupled (TC) schemes (Noureldin et al., 2013). Compared to LC scheme, TC scheme is more reliable and robust for a navigation system. However, LC scheme not only incurs a lower computation expense than TC scheme but also applicable for transforming feature points into reference frame when perception sensor like camera or LiDAR is applied (Surachet & Chiu, 2022).

The techniques of applying monocular camera to a navigation system like visual odometry (VO) and visual Simultaneous Localization and Mapping (V-SLAM) have been developed for many years. The main issue of monocular camera is that monocular camera can't estimate the depth of images, which

---

* Corresponding author

causes a wrong scale estimation for the trajectory. Many researchers utilized a stereo or RGB-D camera to resolve the scale issue. However, the baseline of stereo camera must be long enough for outdoor applications and RGB-D camera is not suitable for outdoor environments. Moreover, pure visual methods have drift problem with an accumulated errors over distance and lost tracking issue. Other researchers employed inertial measurement unit (IMU) to estimate the scale and enhance the system, such as VIO and the state-of-the-art VINS. Though VIO and VINS is robust enough to avoid lost tracking, they still have drift problems since IMU has an accumulated errors over time. Hence, the GNSS is employed to resolve the drift problem. In the study of INS/GNSS/V-SLAM, the authors in (Cao et al., 2020) employed a GNSS measurement as an anchor point and integrate inertial factors, visual factors, and Doppler factors with factor graph optimization. They employed RTKLIB as the ground-truth, but the solution is not reliable in a GNSS-challenging environment. The authors in (Chiang et al, 2020) analyzed the performance of INS/GNSS/V-SLAM fusion scheme using smartphone sensors and (Chiang et al, 2023) proposed refreshed-SLAM and an INS/GNSS/GCPs/V-SLAM fusion scheme using smartphone sensors for land vehicular navigation applications. They both evaluated their results with a highly precise reference system. However, the smartphone sensors hardly achieve the which-lane level in outdoor environments.

The present paper proposes a visually enhanced fusion scheme with a multi-sensor framework. Based on the advantages and disadvantages of multi-sensors, an integrated algorithm was developed and designed to solve the problem of changing environmental conditions. The core of the system is an EKF estimation algorithm using INS mechanism, and GNSS and V-SLAM measurements are used to update the system. GNSS can provide precise position measurements in an open-sky area. For GNSS-challenged environment, V-SLAM can augment position measurements and provide additional velocity measurements. To verify the performance of the fusion system, various field test data were collected in the outdoor environment and a trustable reference system for ground-truth is used to verify the accuracy of proposed localization fusion scheme for improving the navigation performance. We highlight the contributions of this paper as follows: First, stable states estimation provided by a loosely coupled visually enhanced INS/GNSS fusion scheme which has low computational cost. Second, a highly precise reference system is employed to evaluate our system in a dynamic environment. Third, our system achieves which-lane level or where-in-lane level at normal driving speed range.

Structure of this research is arranged in following sequence. Section 2 presents the navigation structure and the methodologies applied to fulfill multi-sensors integration scheme. Section 3 introduces the experiment setup and test fields selection. The experiment results and discussions will be shown in Section 4. Finally, the conclusion is made in Section 5.

2. METHODOLOGY

2.1 Multi-Sensor Fusion Scheme

The block diagram of proposed loosely coupled scheme of the INS/GNSS/V-SLAM integrated system is illustrated in Figure 2. The IMU measurements were processed using INS mechanization to provide a navigation solution that accounts for position, velocity, and attitude with an unbounded error in the navigation frame. The GNSS of the proposed system provides the absolute position as the predominant measurement update to limit

the accumulated drift error of the system. For the camera of the system, oriented FAST and rotated BRIEF (ORB) SLAM were selected as the visual SLAM (V-SLAM) method for processing data and providing measurement updates. Furthermore, various vehicular constraints were considered, including the zero-velocity update (ZUPT) and zero-integrated heading rate (ZIHR) which are observed in stationary vehicles and velocity and heading rate are set to zero, as well as the non-holonomic constraint (NHC) which restricts lateral and vertical velocities of moving vehicles. All sensor data and constraints were then integrated using an EKF estimation algorithm.
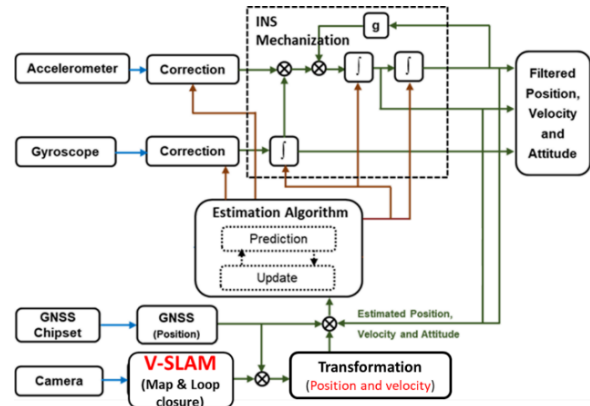

Figure 2 The Proposed Loosely Coupled V-SLAM enhanced INS/GNSS Integration Scheme

Regarding the INS/GNSS integration (Shin, 2005), the navigation state applied in the present study can be described using the following equation:

$$x_k = \begin{bmatrix} r & v & \psi & b_a & b_g & s_a & s_g \end{bmatrix}^T_{21 \times 1} \qquad (1)$$

where
r = position
v = velocity
ψ = attitude
b_a, s_a = bias and scale factor for the accelerometer
b_g, s_g = bias and scale factor for the gyroscope

The initial values for biases and scale factors can be estimated during the initial alignment.

The EKF system model can be simply expressed by dividing into two parts, prediction and update. The prediction part (also called error propagation part) estimates the state and noise of epoch k from the observations at epoch k-1, it can be simplified to the following equations:

$$\delta \hat{x}_k^- = \Phi_{k-1} \delta \hat{x}_{k-1} + w_{k-1} \qquad (2)$$

$$P_k^- = \Phi_{k-1} P_{k-1} \Phi_{k-1}^T + Q_{k-1} \qquad (3)$$

where
$\delta \hat{x}$ = error vector of the system state
Φ_{k-1} = state transition model from epoch k-1 to k
P = covariance of the error state vector
Q = covariance of the process noise

$(-)$ denotes the estimated state after prediction but before update process. When a new measurement z_k observed at epoch k, the measurement updates the system in the update process. The state estimate equations incorporate the new measurement are expressed as:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \qquad (4)$$

$$\delta\hat{x}_k^+ = \delta\hat{x}_k^- + \mathbf{K}_k(\delta z_k - \mathbf{H}_k \delta\hat{x}_k^-) \qquad (5)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\,\mathbf{P}_k^-\,(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{H}_k \mathbf{R}_k \mathbf{H}_k^T \qquad (6)$$

where $\mathbf{K}$ = Kalman gain
  $\mathbf{H}$ = measurement model
  $\mathbf{R}$ = covariance of the measurement noise
  $\mathbf{I}$ = identity matrix

$(+)$ denotes the estimated state after update. For the position update as obtained using GNSS and V-SLAM measurements, the EKF measurement model is described by the following equation:

$$z_{GNSS,SLAM} = (r_{INS} - r_{GNSS,SLAM})$$
$$= H_{GNSS,SLAM}\,\delta x_k + v_k \qquad (7)$$

where $z_{GNSS,SLAM}$ = GNSS or V-SLAM measurement
  r_{INS} = INS position vectors
  $r_{GNSS,SLAM}$ = GNSS or V-SLAM position vectors
  $H_{GNSS,SLAM} = [0_{3\times3} \quad I_{3\times3} \quad 0_{3\times15}]$
  v_k = noise in the measurements

2.2 Direct Geo-referencing of Visual Measurements

While each sensor is mounted at different places on the experiment platform and located under different frames (e.g., the GNSS measurements collected by antenna is located under global frame; measurements from IMU and camera are located under their own local frame), the transformation between sensors must be fulfilled by translation and rotation in three axes (i.e., lever arms and boresight angles), the process is called direct geo-referencing (DG) (Yoshimura & Hasegawa, 2004; Surachet & Chiu, 2022). Figure 3 shows the geometric relationship between the component from each sensor. Our aim is to transform the coordinates of V-SLAM solution from the local camera frame (c-frame) to the global mapping frame (m-frame) through the DG process. The equation of DG is given as:

$$(r_c^m)_k = (r_b^m)_k + (R_b^m)_k(r_c^b) \qquad (8)$$

where $(r_c^m)_k$ = position vector of the camera relative to m-frame at time epoch k
  $(r_b^m)_k$, $(R_b^m)_k$ = translation vector and rotation matrix for the b-frame relative to m-frame at time epoch k
  r_c^b , R_c^b = translation vector and rotation matrix measured in c-frame with respect to b-frame

The translation vector and rotation matrix from camera to IMU are also called lever arm and boresight, respectively. These translation vectors and rotation matrices can be determined by the calibration procedures beforehand.
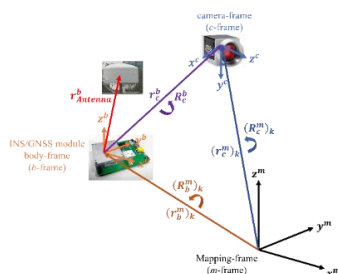

Figure 3 Geometric relationship between each component

2.3 GNSS-based Refreshed SLAM

To rescale and align the V-SLAM trajectory using monocular camera to the global coordinate system, we adopted the concept of GNSS-based refreshed SLAM proposed in (Chiang, 2023) but implemented the process continuously. The V-SLAM solution integrated with GNSS measurements first in this process and then the drift problem over distance of V-SLAM could be greatly reduced. Then the refreshed SLAM solution could provide updates to the system.

2.3.1 Visual SLAM

SLAM is a method for performing mapping in an unknown environment through sensors and for estimating the sensors' locations simultaneously. In the field of robotics and computer vision, SLAM has been extensively researched and applied to vehicle navigation (Mur-Artal & Tardos, 2017; Chiang et al, 2020; Chiang et al, 2023). SLAM can be implemented through various types of sensors, such as visual sensors (cameras), laser sensors (LiDARs), and sonar sensors (radars). Relative to the other types of sensors, cameras are inexpensive, provide abundant environmental details, and utilize passive sensing (i.e., they are unaffected by interference). A SLAM system that uses cameras as its primary sensors is a V-SLAM (V-SLAM) system. Given the advantages of a camera, numerous researchers have focused on camera-based SLAM applications and have proposed various V-SLAM methods (Afia, 2017). Figure 4 presents the strategic framework for implementing V-SLAM.
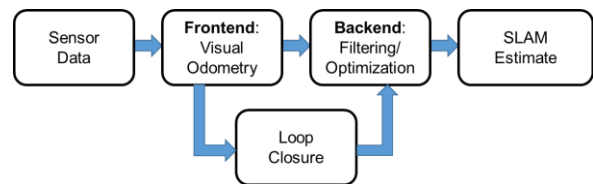

Figure 4 Classic V-SLAM framework (Chiang et al, 2020; Chiang et al, 2023)

ORB-SLAM is a keyframe-based V-SLAM system that uses ORB features (Rublee et al., 2011) to perform all SLAM tasks. This method facilitates robust vision-based navigation. ORB-SLAM performs three main tasks and uses three threads in its algorithms, namely the tracking, mapping, and loop closure threads (Mur-Artal et al., 2015; Mur-Artal & Tardos, 2017). The advanced version of ORB-SLAM provides three modes for processing SLAM using monocular, stereo, and RGB-D cameras (Mur-Artal & Tardos, 2017). In the present study, a monocular camera was used as the visual sensor; therefore, the monocular mode of ORB-SLAM was selected.

2.3.2 Refreshed SLAM

Since monocular cameras cannot estimate the depth of an image, the scale of a monocular V-SLAM translation is undefined. To integrate the visual solution of this system with other sensors, the following transformation processes are implemented. Figure 5 presents the V-SLAM transformation process. First, the scale of V-SLAM translation is recovered with the assistance of GNSS solutions. One short distance from the GNSS solutions in the world frame and one short distance from the V-SLAM solution in the camera frame are used to obtain the scale coefficient, which is calculated by applying the following equation (Chiang et al, 2020; Chiang et al, 2023):

$$\lambda = \frac{\|t_{GNSS}^w\|}{\|t_{SLAM}^c\|} \qquad (9)$$

$$t_{SLAM}^w = \lambda \cdot t_{SLAM}^c \qquad (10)$$

where $\quad \lambda$ = scale coefficient of the translation
$\|t_{GNSS}^w\|$ = distance derived from GNSS solutions located in the world frame
$\|t_{SLAM}^c\|$ = corresponding V-SLAM distance in the camera frame
t_{SLAM}^w = translation vectors of the continuous images with scale recovery
t_{SLAM}^c = translation vectors with original scale

This method of scale recovery is highly dependent on the quality of GNSS measurements. Typically, a section in an open sky area is selected to ensure the reliability and accuracy of the GNSS. However, this approach is unsuitable for real-time applications because the section for scale recovery cannot be manually selected. In this situation, continuous scale recovery was employed using optimal GNSS measurements. A threshold of the standard deviation (STD) of the GNSS position (θ_p) is set in accordance with requirements. During the scale recovery process, the STD of the GNSS measurements, (P_{STD}), is continuously compared with θ_p. If the measurements satisfy the condition $P_{STD} \leq \theta_p$, the section between the previous satisfied position and the current position is used to calculate the scale coefficient for scale recovery.

The V-SLAM solution has a higher sampling rate relative to the GNSS solution; however, it accumulates drift over distance (i.e., travel distance). Therefore, refreshed-SLAM, which takes advantage of GNSS solutions to reset the V-SLAM position, was employed to ensure that the accumulated errors of drift were removed. The GNSS measurements were also filtered by the standard deviation threshold and used to refresh the V-SLAM position. Furthermore, we derived the velocity in the body frame from the V-SLAM translation vector by applying the following equation (Chiang et al, 2020; Chiang et al, 2023):

$$v_{SLAM}^b = C_c^b \frac{t_{SLAM}^c}{\Delta t} \qquad (11)$$

where $\quad v_{SLAM}^b$ = V-SLAM velocity in the body frame
Δt = time period
C_c^b = rotation matrix from the camera frame to the body frame

The V-SLAM velocity update for the EKF system is formed in the measurement model together with the NHC through the equation as follows (Chiang et al, 2020; Chiang et al, 2023):

$$z_{SLAM} = \begin{pmatrix} v_{x-INS}^b - v_{x-SLAM}^b \\ v_{y-INS}^b - 0 \\ v_{z-INS}^b - 0 \end{pmatrix} = H_{SLAM} x_k + u_k \qquad (12)$$

where $\quad z_{SLAM}$ = V-SLAM velocity measurement
$v_{x-INS}^b, v_{y-INS}^b, v_{z-INS}^b$ = velocities derived from the INS mechanization
v_{x-SLAM}^b = V-SLAM velocity in the body frame
$H_{SLAM} = [0_{3\times3} \quad I_{3\times3} \quad 0_{3\times15}]$
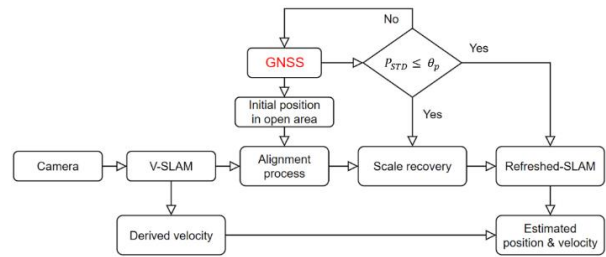u_k = noise in the V-SLAM velocity measurements

Figure 5 V-SLAM transformation for monocular camera

3. EXPERIMENT

3.1 Configuration and Environmental Description

Field tests were conducted using a land vehicle equipped with a reference system and an experimental system. Notably, the experimental system incorporated a self-assembling platform, a MEMS tactical-grade IMU (ADIS 16495-1), a GNSS receiver (ZED-F9P), and a monocular camera (Basler ACE acA1300-75gc). The reference system was used to provide the reference trajectory as the ground truth for field tests. The reference system comprised a navigation-grade IMU (iNAV-RQH) and a GNSS receiver (PwrPak7D-E2). Accurate reference trajectories were obtained using the commercial INS/GNSS processing software package Inertial Explorer (IE) 8.90 by applying a tightly coupled strategy. Figure 6 shows the configuration of our test platform. The IMU technical characteristics is presented in Table 1.

Figure 6 Configuration of test platform

	Accelerometer	Gyroscope
Gyroscope	$< 15\ \mu g$	$< 0.002°/hr$
Random Walk Noise	$8\ \mu g/\sqrt{Hz}$	$0.0018°/\sqrt{hr}$

TABLE 1 Technical characteristics of the iNAV-RQH IMU

To evaluate the proposed algorithm, we designed two tests that were conducted in Tainan City and Changhua City, Taiwan. Both experimental sites presented various types of scenes, including open sky areas and GNSS-challenging environments, as shown in Figure 7. Test 1 was conducted in Tainan City and involved an approximately 12-km-long route. The speed of the vehicle traveling this route was controlled at a maximum speed of 40 km/h. Test 2 was conducted in Changhua City and involved an approximately 23-km-long route. The speed of the vehicle traveling this route was controlled at a maximum speed of 60 km/h. Some trajectories were repeated in both tests to provide the V-SLAM system with the opportunities to perform loop closures.

Figure 7 Scenes in experiment field

3.2 Sensor Availability

On our experiment platform, the IMU collected the acceleration and angular velocity at the sampling rate of 125 Hz; the GNSS receiver provided positioning solutions at the sampling rate of 1 Hz; the camera collected the image sequence at the sampling rate of 20 Hz; and the reference system provided reference navigation solutions at the sampling rate of 200 Hz. Since ORB-SLAM trajectory only considers keyframe poses at the monocular mode (Mur-Artal & Tardos, 2017) due to the ambiguity of the image depth, the V-SLAM solutions cannot reach the sampling rate as camera's frame rate. Moreover, the solution of V-SLAM is not continuous and affected by the environment. Table 2 shows the sampling rate of each sensor for test 1 and test 2, the availability of V-SLAM is expressed by the average value. The navigation solutions of fusion scheme were aligned with the sampling rate of IMU (i.e., 125 Hz) and the reference solutions was also aligned with the same sampling rate of each comparison for statical analysis.

	Test 1	Test 2
IMU (Hz)	125	125
Camera (Hz)	20	20
GNSS (Hz)	1	1
V-SLAM (Hz)	1.32	1.65
Refreshed-SLAM (Hz)	20	20
Proposed fusion scheme	125	125
Reference (Hz)	200	200

TABLE 2 Sampling rate of each sensor

3.3 Results of test 1

The experiment implemented at lower speed range in test 1. Figure 8 shows the velocity distribution of our experimental vehicle in test 1, the maximum velocity is 12.2098 m/s, which approximately equals to 44.3 km/h. Figure 9 (a) displays the V-SLAM trajectories before (red) and after (blue) rescaled with GNSS and Figure 9 (b) presents the trajectories of the reference (red), V-SLAM after rescaled (green), refreshed-SLAM (blue), INS/GNSS (orange), and INS/GNSS/refreshed-SLAM (purple) solutions tested in test 1. The solutions of refreshed-SLAM, INS/GNSS, and INS/GNSS/refreshed-SLAM appeared to overlap with the reference solution. Nevertheless, the trajectory of V-SLAM had an obvious drift compared to the reference solution. As shown in the diagram of height, the V-SLAM solutions in Up-direction couldn't be rescaled accurately.
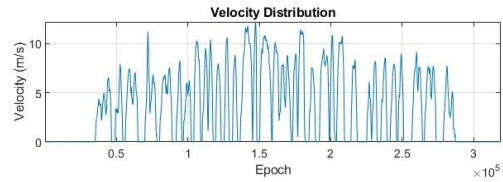

Figure 8 The velocity distribution of vehicle in test 1

The position errors and error statistics of the solutions in test 1 are shown in Figure 10 and Table 3, respectively. The V-SLAM solution exhibits significant position error compared to other solutions, especially in North-direction. The 3D error of the rescaled V-SLAM is approximately 30 m. Even though we performed loop closing in this test, the drift over distance is still significant. Table 3 indicates that refreshed SLAM can significantly improve the performance of system. The errors of INS/GNSS and INS/GNSS/Refreshed-SLAM are greater than refreshed-SLAM, which means our noise processing for IMU may be incomplete. However, the results in test 1 also shows that integrated with refreshed-SLAM can improve the accuracy of INS/GNSS fusion scheme to which-lane level (error < 1.5 m) to where-in-lane level (error < 0.5 m). Figure 10 also indicates that the INS/GNSS/Refreshed-SLAM solution provided by the proposed scheme has a similar trend with the INS/GNSS scheme but with smaller errors.
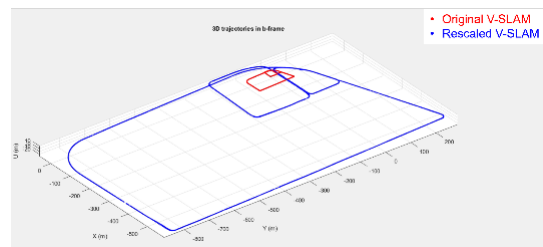
(a)
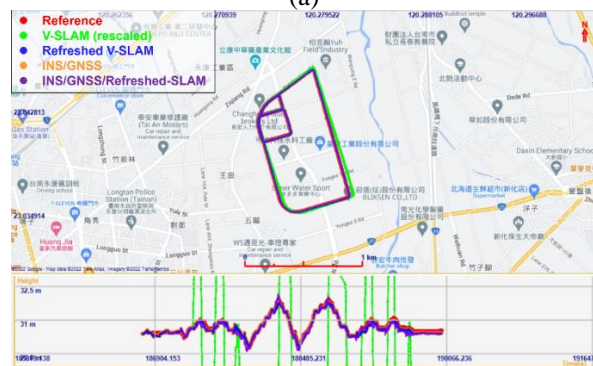
(b)

Figure 9 Comparison of (a) V-SLAM trajectories before and after rescaled and (b) 2D trajectories and heights in test 1.
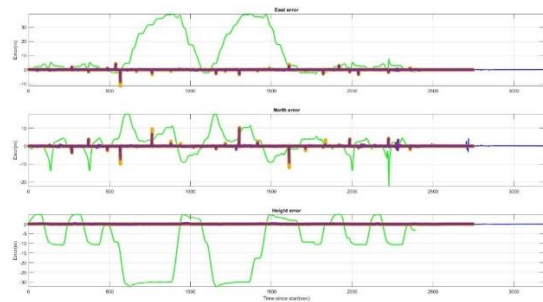

Figure 10 Comparison of position errors with respect to reference solution in test 1

RMSE (m)	V-SLAM (1.32 Hz)	Refreshed-SLAM (20 Hz)	INS /GNSS (125 Hz)	INS /GNSS /Refreshed-SLAM (125 Hz)
North	20.324	0.120	0.322	0.268
East	6.952	0.190	0.438	0.366
Up	17.412	0.115	0.036	0.075
2D	21.481	0.225	0.543	0.453
3D	27.651	0.253	0.545	0.459

Table 3. Statistical analysis of position errors of field test 1.

3.4 Results of test 2

Figure 11 shows our driving velocity distribution in test 2. The maximum velocity is 18.1037 m/s, which approximately equals to 65.2 km/h. In test 2, we implemented a more difficult experiment with a longer route, and larger velocity change, which lead to a highly dynamic environment for the system. Figure 12 (a) displays the V-SLAM trajectories before (red) and after (blue) rescaled with GNSS and Figure 12 (b) presents the trajectories of the reference (red), V-SLAM after rescaled (green), refreshed-SLAM (blue), INS/GNSS (orange), and INS/GNSS/refreshed-SLAM (purple) solutions tested in test 2. Figure 10 shows a similar result to test 1: the solutions of refreshed-SLAM, INS/GNSS, and INS/GNSS/refreshed-SLAM appeared to overlap with the reference solution but V-SLAM had an obvious drift, as well as the V-SLAM solutions in Up-direction couldn't be rescaled accurately. The rescaled V-SLAM estimated the elevation change of approximate 100 m, as shown in Figure 12 (a), while the real elevation change was less than 3 m.
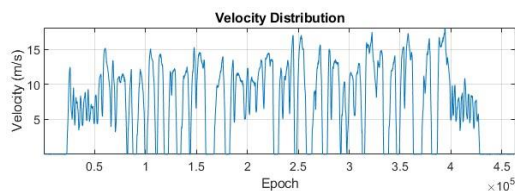

Figure 11 The velocity distribution of vehicle in test 2

The position errors and error statistics of the solutions in test 2 are shown in Figure 13 and Table 4, respectively. As presented in Figure 13, the error trends of INS/GNSS and proposed fusion scheme is relatively stable and have some larger errors at specific places unlike V-SLAM. Contrasted to the trajectory in Figure 12 (b), there errors usually occurred at the turning point. Even with the enhance of GNSS, V-SLAM still has a chance to lose tracking during a sharp turn or U-turn. As a result, we slowed down in order to keep the V-SLAM system not to lost tracking while we meet a sharp turn. The error increased due to the accumulate error of IMU in this period, especially our noise processing is incomplete as mentioned before. However, the proposed visually enhanced scheme is still helpful to the INS/GNSS navigation. We improve the system accuracy to which-lane level this time. It's worth mentioning that the IMU is still necessary for increasing the navigation sampling rate even though it has a drift issue. Higher sampling rate not only can lead to better resolution and robustness of the system's state, but also improve the ability of the navigation system to handle fast motions, such as sudden changes in velocity or rapid changes in direction.
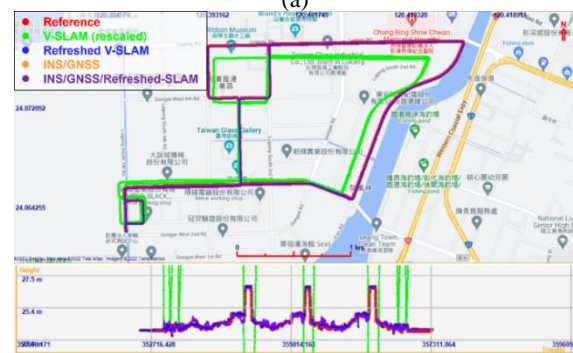

(a)

(b)

Figure 12 Comparison of (a) V-SLAM trajectories before and after rescaled and (b) 2D trajectories and heights in test 2.
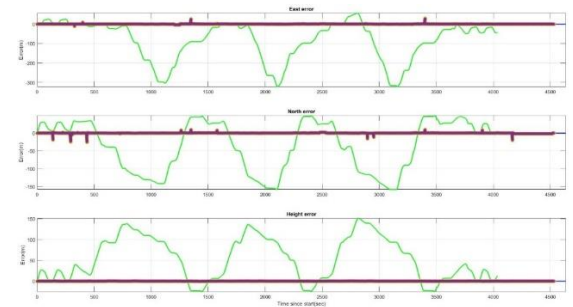

Figure 13 Comparison of position errors with respect to reference solution in test 2

RMSE (m)	V-SLAM (1.65 Hz)	Refreshed-SLAM (20 Hz)	INS /GNSS (125 Hz)	INS /GNSS /Refreshed-SLAM (125 Hz)
North	120.523	0.093	1.076	0.769
East	77.0270	0.130	1.248	0.863
Up	78.685	0.238	0.031	0.156
2D	143.168	0.238	1.648	1.155
3D	163.366	0.286	1.648	1.166

Table 4. Statistical analysis of position errors of field test 2.

4. CONCLUSION

The present paper proposed a visually enhanced INS/GNSS fusion scheme for lane-level navigation applications. Based on the advantages and disadvantages of multi-sensors, a fusion algorithm was developed and designed to solve the problems associated with various environmental conditions. The image sequence collected by a monocular camera is firstly proceed through ORB-SLAM; rescaled and refreshed using reliable GNSS measurements, then the reliable refreshed-SLAM solution is used to update the system. Due to the characteristics of keyframe-based SLAM, the monocular V-SLAM solution is not continuous, and the sampling rate cannot reach the frame rate of

camera. Hence, IMU is necessary for its stability and high frequency. The results indicate that the proposed fusion scheme is advantageous in a highly dynamic environment. Test 1 is implemented at a lower speed range for a 12-km-long route. The proposed method can improve the navigation accuracy from which-lane level to where-in-lane level. In test 2, the experiment implemented at a higher speed range and dramatically speed change for a 23-km-long route. Though the turning point issue and IMU noise increase the position error, the proposed fusion scheme can still achieve which-lane level. For the future work of this research, the IMU noise should be proceed completely and the turning point issue also needs to be resolved. Although we designed our system for real-time applications, it was tested in postprocessing. Future studies should integrate our framework into an embedded system and test it in real-time.

ACKNOWLEDGMENTS

REFERENCES

A. I. Mourikis and S. I. Roumeliotis, 2007. A multi-state constraint Kalman filter for vision-aided inertial navigation. *Proc. IEEE Int. Conf. Robot. Autom., Rome, Italy,* 3565–3572. https://doi.org/10.1109/ROBOT.2007.364024

Afia, A.B., 2017. Development of GNSS/INS/SLAM Algorithms for Navigation in Constrained Environments.

Cao, S., Lu, X., & Shen, S., 2022. Gvins: Tightly coupled gnss-visual-inertial fusion for smooth and consistent state estimation. *IEEE Trans. Rob. 38,* 2004–2021. https://doi.org/10.1109/TRO.2021.3133730

E. Rublee, V. Rabaud, K. Konolige, G. Bradski, 2011. ORB: an efficient alternative to SIFT or SURF. *2011 International conference on computer vision, IEEE,* 2564–2571. https://doi.org/10.1109/ICCV.2011.6126544.

Gleason, S. & Gebre-Egziabher, D., 2009. GNSS applications and methods.

K.-. W. Chiang, D. T. Le, K.-. Y Lin, M.-. L Tsai, 2023. Multi-fusion schemes of INS/GNSS/GCPs/V-SLAM applied using data from smartphone sensors for land vehicular navigation applications. *Information Fusion* 89, 305-319. https://doi.org/10.1016/j.inffus.2022.08.012

K.-. W. Chiang, D.T. Le, T.T. Duong, R. Sun, 2020. The performance analysis of INS/ GNSS/V-SLAM integration scheme using smartphone sensors for land vehicle navigation applications in GNSS-challenging environments. *Remote Sens.* 2020, 12(11), 1732. https://doi.org/10.3390/rs12111732.

Leick, A., Rapoport, L., Tatarnikov, D., 2015. GPS satellite surveying. *John Wiley & Sons.*

Noureldin, A., Karamat, T. B., Georgy, J., 2013. Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration. *Springer Berlin, Heidelberg.*

R. Mur-Artal, J.M.M. Montiel, J.D. Tardos, 2015. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans. Rob.* 31, 1147–1163. https://doi.org/10.1109/TRO.2015.2463671.

R. Mur-Artal, J.D. Tardos, 2017. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans. Rob.* 33, 1255–1262. https://doi.org/10.1109/TRO.2017.2705103.

S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, 2015. Keyframe-based visual-inertial odometry using nonlinear optimization. *Int. J. Robot. Res.,* vol. 34, no. 3, pp. 314–334. https://doi.org/10.1177/0278364914554813.

Stephenson, S., Meng, X., Moore, T., Baxendale, A., & Ford, T., 2011. Accuracy requirements and benchmarking position solutions for intelligent transportation location-based services. *In Proceedings of the 8th international symposium on location-based services.*

Surachet S. & Y.-T. Chiu, 2022. Adaptive Covariance Estimation of LiDAR-based Positioning Error for Multi-Sensor Fusion Scheme with Autonomous Vehicular Navigation System. *Proc. of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022), Denver, Colorado,* 1650-1664. https://doi.org/10.33012/2022.18473

Yoshimura, T., & Hasegawa, H., 2004. Effects of post-processed and real-time DGPS on the precision and accuracy of GPS positioning in forested areas. *Journal of The Japan Forest Engineering Society, 19(2),* 135-140. https://doi.org/10.18945/jjfes.KJ00007485202