

PERFORMANCE ANALYSIS OF SEMANTIC REFRESH INDOOR NAVIGATION FOR SMARTPHONE'S SENSORS USING INS/VINS INTEGRATION SCHEME

Cheng-Xian Lin ^{1*}, Jhih-Cing Zeng ¹, Mei-Chin Hung ¹, Meng-Lun Tsai ², Kai-Wei Chiang ³

¹ Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan - e9971217@gmail.com,

¹ Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan - jessica31017tngs@gmail.com,

¹ Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan - sylviahung85@gmail.com,

² Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan - taurusbryant@geomatics.ncku.edu.tw,

³ Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan - kwchiang@geomatics.ncku.edu.tw

Commission I, WG I/7

KEY WORDS: Semantic, YOLO-v3, INS, VINS, Sensor Integration, Indoor Navigation, GNSS-challenging, Smartphone.

ABSTRACT:

Positioning and Orientation System (POS), which integrates Inertial Navigation System (INS) and Global Navigation Satellite System (GNSS), is widely used to accomplish outdoor navigation missions for land vehicles. However, the positioning accuracy would become worse in GNSS-hostile environments (Chiang et al., 2013), which is quite challenging to accomplish indoor navigation environments. Nevertheless, smartphones are contained many embedded sensors, including GNSS, IMU, camera, which have the potential to be an ideal personal navigation device. In this research, we mainly propose an integrated scheme of INS/VINS/object detection refresh (ODR) for indoor challenging environments. The goal is to achieve indoor navigation for vehicular applications only using smartphones. The algorithm is developed based on the smartphone. By the conventional inertial navigation system, which is integrated with two designed processes to further improve the performance. First is assistance from the visual-inertial navigation system (VINS). The long-term drift caused by the INS could be decreased effectively, and complete the extended Kalman filter (EKF) composition. The second is to apply neural network, YOLO-v3 (Redmon et al., 2018), to detect objects and provide the object's describer information to refresh the proper position. Therefore, the proposed method uses visual estimation and recognition methods to assist the smartphone platform to obtain a more accurate solution.

Finally, we use the navigation-grade IMU as the reference system for accuracy verification. The accuracy comparisons of the three integration solutions are analysed reasonably. The position accuracy is reasonable. Compared with the original smartphone INS integration method, the proposed integration scheme improves the accuracy from the horizontal direction by 78.5%.

1. INTRODUCTION

Nowadays, the navigation estimation technology of land vehicles has been gradually innovated. This technology will integrate different platforms and a variety of sensors, equipped with a global navigation satellite system, inertial navigation system, integrated micro electro mechanical systems (MEMS), software engineering, and other components to assist the structure of navigation technology. Specifically, in GNSS-challenge environments like the concrete jungle of the urban city or indoors, the availability of GNSS would be degraded extremely (Chiang et al., 2013). Consequently, simultaneous localization and mapping (SLAM) using cameras or lidars in INS integration schemes is currently becoming more popular as well (Chu et al., 2012) (Li et al., 2019). The application of multi-sensors such as visual odometry (VO), lidar odometry (LO) enables the redundancy of the classical state estimation methods and improves the robustness of the current algorithms (Liang et al., 2020). Accurate indoor navigation is a coveted objective, which can be achieved through the strategy of multi-sensor integration. However, in indoor environments where GNSS signals are occluded, a variety of indoor positioning methods have been proposed. But most methods are expensive and can be difficult to implement. For example, Wi-Fi and Bluetooth based positioning and assistance from other sensors

(Huang et al., 2017) (Zhang et al., 2017) (Zhuang et al., 2016). However, these solutions are mostly designed based on pedestrian thinking and require the placement of additional equipment such as Bluetooth beacons or WiFi routers. Smartphones are the most prospective personal navigation devices, whether in pedestrian or in-vehicle mode. However, with the low-cost design of smartphones, the measurement quality is unsatisfactory. At present, several studies using smartphones related to navigation have been proposed.

For example, the smartphone fusion location method optimized by indoor/outdoor detection is proposed. Using the lightweight sensor, the magnetic sensor and the satellite signal are integrated to recognize the indoor/outdoor status (Zeng et al., 2017). And one completed the evaluation of sensors in modern smartphones (Forster et al., 2012). But the main purpose is vehicle traffic monitoring, not vehicle navigation applications. On the other hand, visual-inertial odometry (VIO) is proposed in the visual to assist the drift of the accumulated error of the IMU through visual estimation (Forster et al., 2012) (Mueggler et al., 2018) (Huai et al., 2018). Among those VIO algorithms, the VINS is an algorithm that integrates camera and IMU data to achieve visual odometry (Qin et al., 2018).

So far, there is no single technology that can provide reliable indoor positioning. In particular, several major problems need to be overcome. The first is that indoor navigation solutions

* Corresponding author

need to rely on additional equipment. The second is that most indoor navigation solutions using smartphones are specially developed for pedestrians. Applications need to be developed with vehicle navigation, even with indoor navigation. Therefore, two integration schemes are developed to overcome the above problems, namely INS/VINS with visual-inertial odometry assistance, and INS/VINS/ODR integration scheme with semantic recognition refresh algorithm. The purpose is to focus on indoor vehicle navigation applications using smartphones in this study. There is also no need to mount other additional equipment on the vehicle to achieve navigation.

2. METHODS

2.1 The Proposed VINS aided INS integration scheme

In this research, this loosely coupled INS/VINS integration scheme is shown in Figure 1. The EKF is integrated into the multi-sensor measurement of smartphones. Particularly, the IMU raw measurements include specific forces and angular velocities in three axes with six degrees of freedom. The IMU error model is based on the mechanization of the INS. The position (r), velocity (v), and attitude (ψ) can be derived from the INS mechanization based on the compensated observations of three axes gyroscopes (g) and accelerometers (a). Among them, the position (r_v) velocity (v_v) are developed from the VINS on the basis of fusing the advantages of INS and visual estimation. Furthermore, in the designed EKF, the biases (b) and scales (s) of IMU observations are estimated, and the complete navigation states (S) can be listed as the following equation. Instead of estimating the navigation, the state transition matrix predicts the state errors at the next epoch and calculates the uncertainties.

$$S = \begin{bmatrix} r & v & \psi & r_v & v_v & b_a & b_g & s_a & s_g \end{bmatrix}_{1 \times 27} \quad (1)$$

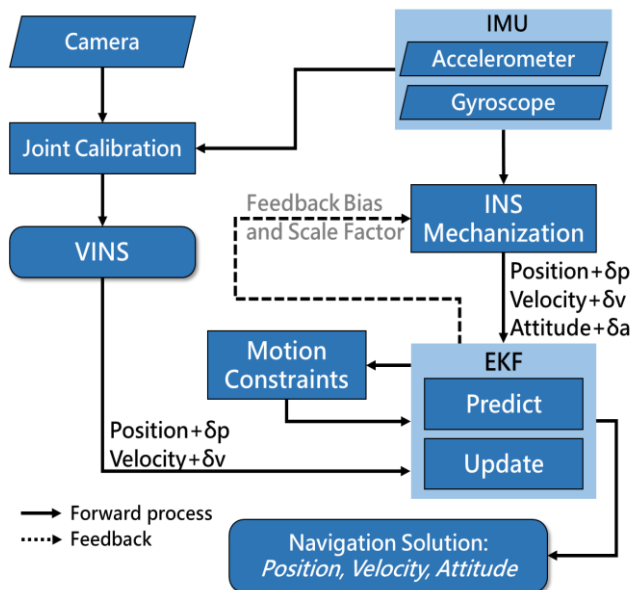


Figure 1. The proposed VINS aided INS integration scheme

Modern people rely on smartphones to provide sufficient location services for GNSS navigation support in outdoor scenarios. Due to the limitations of smartphone sensors, integrated navigation solutions relied on low-cost inertial

sensors within smartphones alone cannot retain sufficient accuracy in GNSS challenging areas or indoor scenarios. In this research, other vehicle motion constraints, such as NHC, ZUPT, and ZIHR, are also adopted. The details of these motion constraints will be explained in the following part.

In the applications of indoor vehicle navigation, GNSS signals will be completely blocked and shielded. However, in this case, the conventional pure INS integration is expected to have a long-term error accumulation even after applying those motion constraints. Therefore, the VINS is applied to improve accuracy over the driving distance. The features of most indoor parking lot scenarios can be easily extracted and matched by the VINS algorithm, which not only provides a robust position but also limits drift caused by long-running INS. This is a significant pre-eminence to acquire continuous relative orientations including position and attitude. Finally, VINS solutions are transformed into the position and velocity as the input in the designed integrated scheme.

2.1.1 Joint Calibration

In this part, we adopt an extension to the joint calibration tool Kalibr that allows for determining the extrinsic and intrinsic of multiple IMUs in a single estimator. We further demonstrated that it is feasible to infer the location of individual accelerometer axes to millimeter precision (Rehder et al., 2016). The method mainly calibrates a sensor suite that includes one or multiple IMUs and one or multiple external perception cameras. The goal is to improve the state estimation results of all fused sensors in our experiment smartphone. Using an EKF based framework was proposed to estimate the exterior orientation between the IMU and the camera from a sequence of calibration images recorded by moving the device in front of the target relationship transformation (Kelly et al., 2009). Consequently, we can obtain the configuration of the smartphone with accurate internal and external orientation parameters, which can provide the solution to good initialization of subsequent visual-inertial navigation. However, the smartphone joint calibration is calibrated with reference to a combination of a low-cost IMU and a rolling shutter camera, as shown in Figure 2 (Lee et al., 2018).

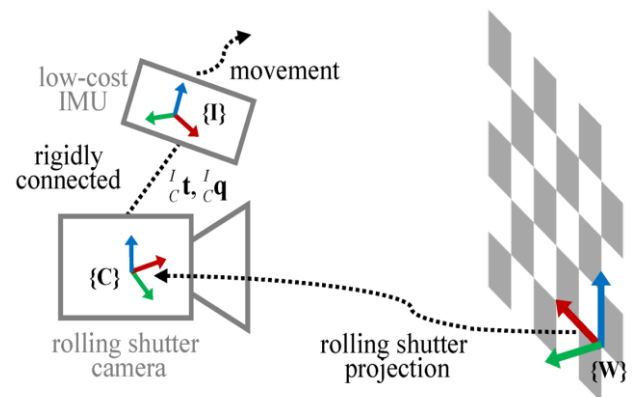


Figure 2. Schematic overview and calibration coordinate description of a simulated smartphone calibration. (Lee et al., 2018)

where $\{W\}$ = coordinate of the world
 $\{C\}$ = coordinate of the rolling shutter camera
 $\{I\}$ = coordinate of the IMU

smartphone to accomplish the pursuit of indoor navigation. Details concerning the proposed scheme are described below.

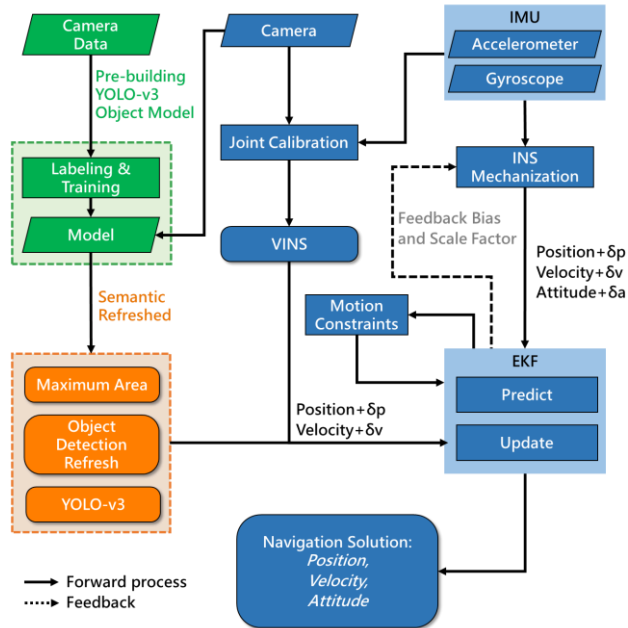


Figure 5. The proposed semantic refreshed INS/VINS/ODR integration scheme

2.2.1 Object Detection Refresh

Geo-referenced features acquired by the proposed Object Detection Refresh (ODR) algorithm are applied for the position refresh process for VINS. This is an innovative algorithm developed specifically for indoor vehicular parking applications, which can individually provide location information and provide updates on indoor environments shown in Figure 6. When the vehicle moves closed to certain objects in the indoor parking lots such as pillars, the proposed ODR algorithm detects and records relevant descriptions, including timestamp, serial number, image coordinates, bounding box size based on YOLO-v3 (Redmon et al., 2018), and some value of prior information, such as locating target (green point) built with an indoor mapping procedure.

The background of this geo-referenced feature includes inertial navigation and photogrammetry components. The location of each image is known, and the recorded images are used for classification and labeling. By appraising a variety of image detection algorithms, the most suitable one is YOLO-v3. This trained and predicted built model and the positions obtained by the mobile mapping system are translated into a pre-built dataset. As a result, the positioning target and its narrative content are obtained. With the above pre-built content, based on the condition that the vehicle must drive in the lane, the method of semantic detection can refresh the position by locating targets. Consequently, as long as the vehicle drives along the lane in that parking lot, location information is obtained whenever an object is detected. In conclusion, this approach can be achieved even without any pre-built location device.

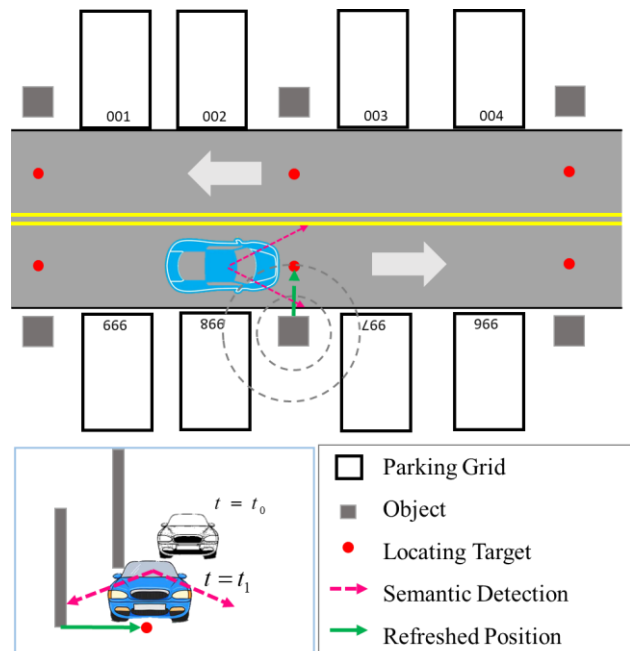


Figure 6. Schematic of proposed ODR

2.2.2 YOLO-v3

The process of establishing a model by correcting the weights of each layer through learning samples is called training. At present, most of the algorithms involve object detection based on artificial neural networks (ANN). In this research, the most common pillars in indoor scenarios are selected primarily as geo-referenced objects, which are easily identified. The current algorithm for object detection contains R-CNN, Fast R-CNN, Faster R-CNN (Ren et al., 2015) (Girshick et al., 2015), YOLO (Chen et al., 2021). YOLO has continuously updated and improved algorithms to achieve better performance and improve detection capabilities at different levels. Therefore, the YOLO-v3 (Redmon et al., 2018) of the supervised learning network is adopted for object detection in this study. The smartphone is installed above the vehicle dashboard, and the recorded content handles labeling and description of object details. The results of YOLOv3 object detection are shown and the pillars (yellow bounding box) are well detected in Figure7.

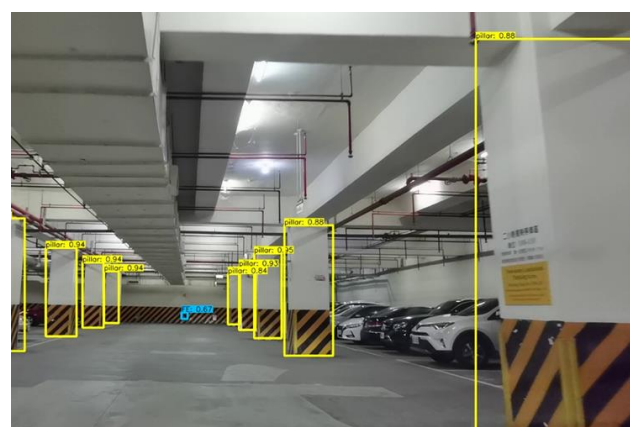


Figure 7. The results of YOLO-v3 object detection

2.2.3 Maximum Area Algorithm

The algorithm applied to determine the moment when the vehicle is about to drive past the pillar in this research is the proposed maximum area algorithm, as shown in Figure 8. The

images recorded by the smartphone camera are pre-processed first, and then the training model pre-trained by YOLO-v3 is used for prediction. During the process, the size and description of the bounding box of each frame will be continuously recorded. After the design time threshold is passed, it considers whether the area size of the bounding box is reduced from the maximum value to near zero. These updated measurements are also used to calculate the primary navigation states, which are the 10th to 15th values in the S state from equation 1.

As for the maximum area algorithm, it is designed based on the change of bounding size. The judgment method is composed of the threshold value of the area and duration time of the bounding box in which the image appears, both of which must exceed the threshold value at the same time. After the maximum area algorithm is established, the semantic refreshed INS/VINS/ODR integration scheme can be carried out. As a consequence, the algorithm will then update the position and velocity estimated by real-time VINS to correct the accumulated error. And record the maximum area value (red circle) of each frame, and provide information in combination with a semantic application, as shown in Figure 9.

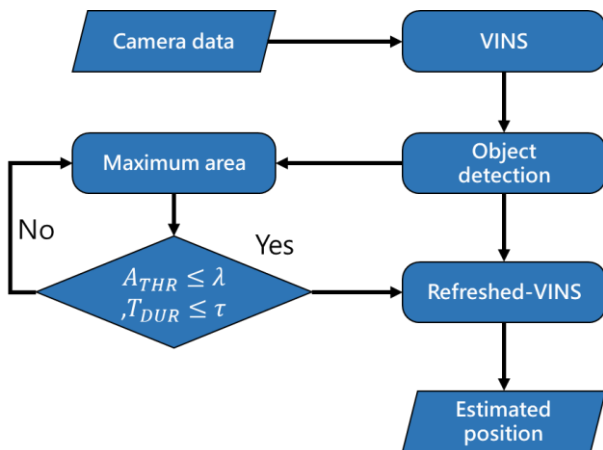


Figure 8. Flowchart of maximum area algorithm

Where A_{THR} = threshold of area
 T_{DUR} = threshold of duration time
 λ = area of bounding box
 τ = duration of detection maximum area

There are empirically A_{THR} based and T_{DUR} thresholds representing area and duration of time, respectively. Then, λ and τ represent the area and time measurement. Thus, whenever this method is available, the semantic refresh application is provided.

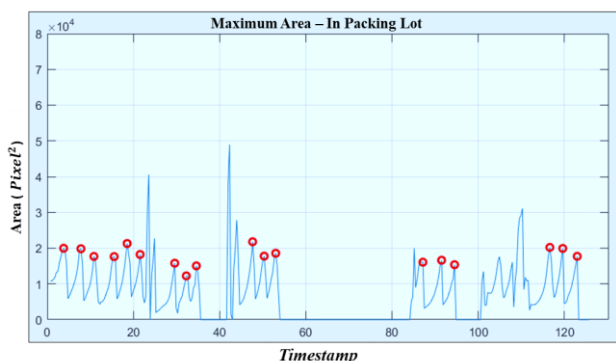


Figure 9. Illustration of the maximum area algorithm for detection

3. EXPERIMENT SETUP

The experiments are divided into two parts. The first part focuses on the indoor navigation solution to use the smartphone with INS/VINS integration. The second part emphasizes the performance of the semantic refreshed aided INS/VINS/ODR integration scheme specifically for indoor scenarios.

The experiments scenario starts from the open sky outset of Tzu Chiang Campus, Nation Cheng Kung University (NCKU). Drive the general road and verify indoor positioning through the library parking lot on the Cheng Kung Campus, and return to the starting point. There are two navigation systems equipped on the experimental vehicle, iMAR iNAV-RQH-10018, and NovAtel PwrPak7D-E2. The reference solution combining iNAV-RQH-10018 navigation-grade IMU with powerful OEM7 GNSS engine from PwrPak7D-E2 are processed by commercial software, which uses carrier phase measurements with DGNS and smoothed mode tightly coupled for INS/GNSS integration process are regarded as the true value. The antenna used in one of the systems was an Antcom 72GNSSA-XT-1 choke ring for PwrPak7D-E2, and the other system used a NovAtel GPS-703-GGG. In particular, the choke ring is resistant to multipath effects. The smartphone can be thought of as a multi-sensor combined system in this research. Which contains the low-cost MEMS-grade IMU, and rolling shutter camera. And the smartphone is mounted on the holder on the dashboard, as shown in Figure10. The configuration of the experimental reference system is shown in Figure 11.



Figure 10. Smartphone set above the dashboard.

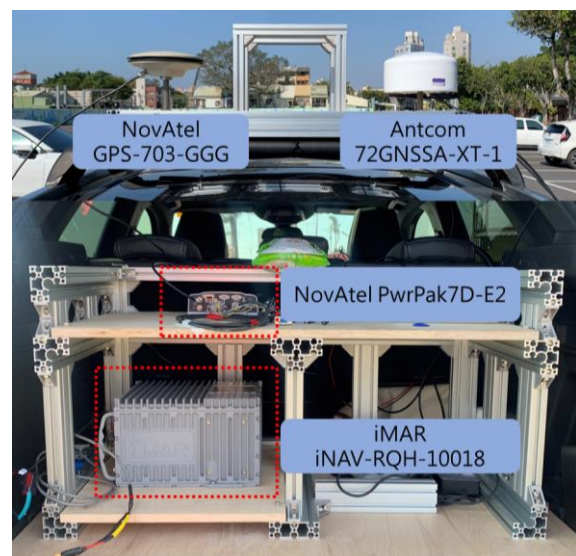


Figure 11. The experimental configuration of reference system.

4. RESULTS AND DISCUSSION

4.1 Performance Analysis of INS Integration Scheme

This is the analysis result of the INS integration solution of the smartphone compared with the reference system. Therein, the blue line means the smartphone pure INS integration result, and the red line means the reference, as shown in Figure 12. The size of this indoor parking lot is about 80x75 meters. And the experimental test indoor travel time is about 300 seconds. This shows that it is not ideal to only use the IMU of a smartphone for navigation estimation applications. Using it for navigation may cause the user to become disoriented. The position errors of the E and N directions are 6.96 meters and 2.87 meters, and the height direction is about 7.53 meters, as shown in Table 1.



Figure 12. The trajectory of INS integration method

4.2 Performance Analysis of INS/VINS Integration Scheme

The statistics with the smartphone for INS/VINS integrated filtered solutions are shown in Table 1. The horizontal position of the proposed INS/VINS scheme is shown in Figure 13 as well. The orange line is the reference system trajectory, and the pink line is the INS/VINS integration scheme. The results illustrate that with the addition of VINS assistance, the errors in the horizontal and vertical directions have been greatly improved and smoothed. In the RMSE position error analysis, the horizontal error is reduced by about half compared to the previous one, and the height error was reduced to 0.178 meters.



Figure 13. The trajectory of INS/VINS integration method

4.3 Performance Analysis of Semantic Refreshed INS/VINS/ODR Integration Scheme

The orange line is the reference system trajectory, and the green line is the proposed semantic refreshed INS/VINS/ODR integration scheme, as shown in Figure 14. When the vehicle passes through the pre-built model and detects objects, that is the described point (red square) to refresh the position in the proposed algorithm. Overall, the position accuracy analysis shows that the trajectory results have been greatly improved. Whether it is E and N direction errors or three-dimensional errors are drastically reduced.



Figure 14. The trajectory of proposed integration scheme

Unit: Meter	Position Error	East	North	Height
INS	Mean Error	-5.757	-0.271	6.851
	STD	3.885	2.861	26.530
	RMSE	6.955	2.874	7.525
INS/VINS	Mean Error	-1.286	-0.112	-0.161
	STD	3.268	1.676	0.075
	RMSE	3.512	1.680	0.178
INS/ VINS/ Object Detection Refresh	Mean Error	-0.470	-0.511	0.037
	STD	1.226	0.798	0.080
	RMSE	1.313	0.948	0.088

Table 1. The analysis of the positioning error, including INS, VINS, and proposed integration scheme.

The position error of the final scheme compared to the reference system is plotted in Figure 15. Whenever the vehicle records an image and detects this geo-referenced feature that has been built, the semantics of the object are refreshed. In other words, wherever the location and velocity information are refreshed, the position is going to be moved to the established locating targets (red circle). And continue to calculate the semantic refreshed INS/VINS/ODR integration navigation solution.

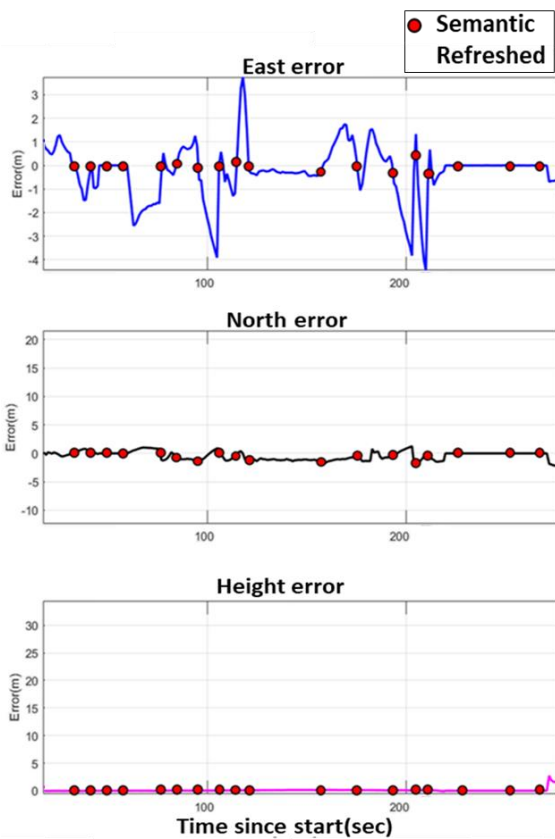


Figure 15. The position error of proposed INS/VINS/ODR integration scheme

To sum up, the RMSE of the horizontal position is less than 3 meters which is very accurate when a smartphone and a navigation grade IMU are applied as test and reference systems, respectively. Therefore, it can bring very good benefits to car navigation with smartphones in indoor scenarios. In conclusion, the final proposed fusion model has overcome more of the issues mentioned and achieved. The improvement in each direction is organized in Table 2.

Navigation Improvement				
Scheme	Error	East	North	Height
INS/VINS	RMSE	49.5%	41.5%	97.6%
Semantic Refreshed	RMSE	81.1%	67.0%	98.8%
INS/VINS/ODR				

Table 2. Summary of navigation improvement of the indoor test

5. CONCLUSIONS

This study proposed indoor navigation by smartphone sensors. One is the scheme for the VINS aided INS integration, and another scheme is semantic refreshed INS/VINS/ODR integration. The smartphone that collected the data was used for testing. Moreover, a system comprised of navigation-grade IMU by iMAR iNAV-RQH-10018, and differential GNSS receivers from PwrPak7D-E2 was used to generate reference solutions. The accuracy comparisons of these integration solutions are analyzed. As a result of the INS/VINS integrated solution, the horizontal position error is about 3.893 meters, which improves by 48%. Finally, the result of the INS/VINS/ODR integration scheme, the horizontal position error is about 1.620 meters,

which improves by 78%. For future work, different geo-referenced objects will be applied as recognition models in the integrated system to add new datasets and tests. And the objective is optimized for the semantic refreshed INS/VINS/ODR integration scheme, and it is expected that smartphone GNSS can also be added to form an integrated scheme for indoor/outdoor seamless positioning. Ultimately, this framework should be integrated with the smartphone environment and run directly on this low-cost platform.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge the technical advices from their supervisor, Dr. Kai-Wei Chiang and implementation assistance from, Jhih-Cing Zeng and Mei-Chin Hung. Authors acknowledge the financial support provided by the MOI in the Executive Yuan in Taiwan.

REFERENCES

- Chiang, K. W., Duong, T. T., & Liao, J. K. (2013). The performance analysis of a real-time integrated INS/GPS vehicle navigation system with abnormal GPS measurement elimination. *Sensors*, 13(8), 10599-10622.
- Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- Chu, T., Guo, N., Backén, S., & Akos, D. (2012). Monocular camera/IMU/GNSS integration for ground vehicle navigation in challenging GNSS environments. *Sensors*, 12(3), 3162-3185.
- Li, T., Zhang, H., Gao, Z., Niu, X., & El-Sheimy, N. (2019). Tight fusion of a monocular camera, MEMS-IMU, and single-frequency multi-GNSS RTK for precise navigation in GNSS-challenged environments. *Remote Sensing*, 11(6), 610.
- Liang, Y., Müller, S., Schwendner, D., Rolle, D., Ganesch, D., & Schaffer, I. (2020). A Scalable Framework for Robust Vehicle State Estimation with a Fusion of a Low-Cost IMU, the GNSS, Radar, a Camera and Lidar. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1661-1668). IEEE.
- Huang, X., Guo, S., Wu, Y., & Yang, Y. (2017). A fine-grained indoor fingerprinting localization based on magnetic field strength and channel state information. *Pervasive and mobile computing*, 41, 150-165.
- Zhang, W., Sengupta, R., Fodero, J., & Li, X. (2017, December). DeepPositioning: Intelligent fusion of pervasive magnetic field and WiFi fingerprinting for smartphone indoor localization via deep learning. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 7-13). IEEE.
- Zhuang, Y., Yang, J., Li, Y., Qi, L., & El-Sheimy, N. (2016). Smartphone-based indoor localization with bluetooth low energy beacons. *Sensors*, 16(5), 596.
- Zeng, Q., Wang, J., Meng, Q., Zhang, X., & Zeng, S. (2017). Seamless pedestrian navigation methodology optimized for indoor/outdoor detection. *IEEE Sensors Journal*, 18(1), 363-374.

- Forster, M., Frank, R., & Engel, T. (2012). Evaluation of sensors in modern smartphones for vehicular traffic monitoring. University of Luxembourg.
- Mueggler, E., Gallego, G., Rebecq, H., & Scaramuzza, D. (2018). Continuous-time visual-inertial odometry for event cameras. *IEEE Transactions on Robotics*, 34(6), 1425-1440.
- Huai, Z., & Huang, G. (2018, October). Robocentric visual-inertial odometry. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 6319-6326). IEEE.
- Qin, T., Li, P., & Shen, S. (2018). Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4), 1004-1020.
- Rehder, J., Nikolic, J., Schneider, T., Hinzmann, T., & Siegwart, R. (2016, May). Extending kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 4304-4311). IEEE.
- Kelly, J., & Sukhatme, G. S. (2009). Fast relative pose calibration for visual and inertial sensors. In *Experimental Robotics* (pp. 515-524). Springer, Berlin, Heidelberg.
- Lee, C. R., Yoon, J. H., & Yoon, K. J. (2018). Calibration and noise identification of a rolling shutter camera and a low-cost inertial measurement unit. *Sensors*, 18(7), 2345.
- Engel, J., Schöps, T., & Cremers, D. (2014, September). LSD-SLAM: Large-scale direct monocular SLAM. In *European conference on computer vision* (pp. 834-849). Springer, Cham.
- Mur-Artal, R., Montiel, J. M. M., & Tardos, J. D. (2015). ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE transactions on robotics*, 31(5), 1147-1163.
- Gao, X., Wang, R., Demmel, N., & Cremers, D. (2018, October). LDSO: Direct sparse odometry with loop closure. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 2198-2204). IEEE.
- Heo, S., Cha, J., & Park, C. G. (2018). EKF-based visual inertial navigation using sliding window nonlinear optimization. *IEEE Transactions on Intelligent Transportation Systems*, 20(7), 2470-2479.
- Sun, K., Mohta, K., Pfrommer, B., Watterson, M., Liu, S., Mulgaonkar, Y., ... & Kumar, V. (2018). Robust stereo visual inertial odometry for fast autonomous flight. *IEEE Robotics and Automation Letters*, 3(2), 965-972.
- Mur-Artal, R., & Tardós, J. D. (2017). Visual-inertial monocular SLAM with map reuse. *IEEE Robotics and Automation Letters*, 2(2), 796-803.
- Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R., & Furgale, P. (2015). Keyframe-based visual-inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 34(3), 314-334.
- Shin, E. H. (2005). Estimation techniques for low-cost inertial navigation. UCGE report, 20219.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 91-99.
- Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).
- Chen, W., Huang, H., Peng, S., Zhou, C., & Zhang, C. (2021). YOLO-face: a real-time face detector. *The Visual Computer*, 37(4), 805-813.