PEDESTRIAN INERTIAL NAVIGATION ALGORITHM BASED ON SCENE RECOGNITION

Tingting Lei [1], You Li [2,3] *

[1] School of Geodesy and Geomatics, Wuhan University, Luoyu Road, China - leitingting@whu.edu.cn
[2] Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Luoyu Road, China - liyou@whu.edu.cn
[3] Hubei Luojia Laboratory, Wuhan, China

Commission IV, WG IV/5

KEY WORDS: Pedestrian Navigation, Inertial Navigation, Scene Recognition, EKF

ABSTRACT:

As indoor location-based services become increasingly essential to people's daily life, it is necessary to build a stable and accurate indoor pedestrian positioning system. The foot-mounted inertial navigation system can provide a short-term robust solution but suffer from error accumulation over time. To alleviate this issue, this paper proposes a pedestrian inertial navigation algorithm based on scene recognition to reduce the heading drift. Based on the hypothesis that corridors in buildings are generally narrow and straight and pedestrians have a high probability to walk in straight lines in corridors, we use a scene recognition model to assist foot-mounted INS. When the scene recognition model determines that the pedestrian is walking in a corridor, a straight-line constraint will be implemented to reduce the heading drift and improve the observability of the vertical gyroscope bias. Experiments show that the algorithm can effectively improve the navigation performance and the observability of vertical gyroscope bias when the sensor biases are significant.

1. INTRODUCTION

With the rapid development of science and technology, location-based services in the indoor environment are increasingly indispensable to people's daily life. Social activities are no longer limited to familiar and fixed places. Especially in professional application fields such as fire rescue, electrical circuit, pipeline maintenance, etc., changes in the location of workers are important information, vital to the safety of life and property. Therefore, there is an urgent need for a stable and accurate indoor pedestrian navigation and positioning system.

In an indoor environment where satellite signals cannot be received, inertial navigation is a common method for pedestrian navigation. In recent years, pedestrian localization methods based on inertial sensors (i.e., gyroscopes and accelerometers) have been widely studied due to their strong anti-interference ability. Most of the current research is based on low-cost inertial sensors. A typical solution is to fix inertial sensors on the foot and use the periodic static state of the foot to suppress the divergence of system errors, which is essentially a Zero Velocity Update aided Inertial Navigation System (ZUPT Aided INS). However, the z-axis gyroscope bias is a weakly-observable element in the ZUPT Aided INS algorithm; thus, many studies on foot-mounted INS focus on reducing the heading error and improving the observability of unobservable elements.

To further improve the positioning accuracy and reduce the divergence of heading errors, other sources of information can be added to the foot-mounted inertial system. For example, wireless positioning methods are commonly used to assist inertial systems, such as WiFi (Biswas and Veloso, 2010), UWB (De Angelis et al., 2016), Bluetooth low energy (Faragher and Harle, 2015), etc. However, these methods need to deploy base stations in advance, increasing the workload and cost of the system. In contrast, vision sensors have lower costs and higher stability, which assist inertial

navigation systems well. Besides, with the rapid development of convolutional neural networks, researchers have begun to use vision sensors to perform scene recognition. One of the most famous datasets for scene recognition is the Places dataset (Zhou et al., 2018). With a large number of scene pictures and scene classifications, the Places dataset is the largest dataset with scenes as the core. Furthermore, with the rapid development of Convolutional Neural Networks (CNN), people have begun to use deep learning methods for scene recognition. Some CNN models trained on the Places dataset have already achieved good performance.

This paper proposes a pedestrian inertial navigation algorithm based on scene recognition. In the process of walking, the inertial sensor is fixed on the foot; meanwhile, the pedestrian holds the mobile phone camera to record a video, which is used to identify the scene. To process the data, firstly, through the INS mechanization, the original output of the foot-mounted IMU can be converted to the attitude, velocity, and position of the device. Then, the zero-velocity state is detected and the ZUPT is applied to suppress the accumulation of pure INS positioning errors. Also, Heuristic Drift Reduction (HDR) is used to reduce the heading error divergence. Next, through the ResNet trained on the Places dataset, the scene can be identified according to the video recorded by the mobile phone camera. When pedestrians walk in long and narrow corridors, pedestrians have a high probability to walk in a straight line, so heading constraints can be added. Finally, the Extended Kalman Filter (EKF) is used to estimate the attitude, velocity, and position of the pedestrian navigation system.

The pedestrian inertial navigation algorithm based on scene recognition has the following advantages: (1) Adding different constraints according to different scenarios can make inertial navigation constraints more reasonable and effective; (2)

---

* Corresponding author

Compared with wireless positioning methods such as UWB and WiFi which need to deploy base stations in advance, the scene recognition method only needs a camera and thus has a lower cost and higher flexibility.

The rest of the paper is organized as follows. In Section 2, related work about foot-mounted inertial navigation systems and scene recognition is provided. In Section 3, the algorithm of pedestrian inertial navigation based on scene recognition is proposed. In Section 4, the experimental results are presented. Section 5 summarizes the work and gives the conclusion.

2. RELATED WORK

2.1 Foot-Mounted Inertial Navigation

The foot-mounted inertial navigation system was first proposed by Foxlin (Foxlin, 2005), with the inertial sensors tied to the feet to observe and estimate the motion of pedestrians. Considering the portability, the inertial sensors on the foot usually have smaller sizes and lower weight. Therefore, the performance of the inertial sensors is lower and the drifts are more significant. To limit the drifts of the foot-mounted inertial system, we can use an EKF with the ZUPT strategy, which is called INS-EKF-ZUPT (Jiménez et al., 2010), or IEZ for short. Once the strapdown INS mechanization is performed, the navigation information (position, velocity, and attitude) of the pedestrian is estimated; then, according to the output of the inertial sensors, the zero-velocity state is detected every time the pedestrian's foot is on the floor, thus the ZUPT strategy can be introduced into the EKF to correct the navigation error.

Although the IEZ method can correct the navigation error after each stride, the heading and its bias are non-observable variables, thus the foot-mounted inertial system with ZUPT alone will gradually gain horizontal position errors. Several methods can be used to limit the heading drift, such as Zero Angular Rate Update (ZARU) (Rajagopal, 2008), HDR (Borenstein et al., 2009), etc. ZARU is based on the hypothesis that the angular rate of the IMU is zero when it is stationary, but the pedestrian's footstep is not entirely stationary when walking. HDR is based on the hypothesis that the pedestrian tends to walk in a straight line, so HDR will not work until the pedestrian walks in a fixed direction for a short while. Each method has its advantages and limitations.

2.2 Scene Recognition

Scene recognition is to define the place where the objects seat in the image, and assign semantic labels to the scene images. Different from object recognition, the task of scene recognition is more complex because the overall connection in the image needs to be taken into consideration, such as the background and spatial layout. Early scene recognition methods mainly used features to describe images, such as Scale-Invariant Feature Transform (SIFT), Generalized Search Trees (GiST), Histogram of Oriented Gridients (HOG), etc., which are simple to operate but lack semantic expression ability. Later, more recognition methods emerged and semantic information was added, such as Object Bank and Bag-of-Words. But it is difficult to further improve the recognition performance. Nowadays, with the rapid development of Convolutional Neural Networks (CNN), people have begun to use deep learning methods for scene recognition.

In terms of deep learning for scene recognition, several representative scene-centric databases are ImageNet, Places, SUN, etc. Compared with the ImageNet and SUN, the Places dataset has the largest number of images to feed deep learning algorithms, and it is the most diverse of the three dataset (Zhou et al., 2018). With rich scene classifications and larger quantities of scene images, the Places dataset is the largest scene-centric image dataset. Therefore, the Places dataset is a suitable training set for scene recognition.

According to Places dataset, four CNNs, AlexNet (Iandola et al., 2016), GoogLeNet (Zhong et al., 2015), VGG (Sengupta et al., 2019) and Residual Network (ResNet) (Tai et al., 2017), are trained on the Places dataset. The results show that VGG and ResNet have better performance in scene recognition, with a top-1 accuracy rate of over 50 % and a top-5 accuracy rate of over 80 %. Therefore, we can use the ResNet model trained on the Places dataset to perform scene recognition.

2.3 Integration

To further improve the positioning accuracy of the foot-mounted inertial system and reduce the divergence of heading errors, other navigation information can be integrated with foot-mounted INS. For example, satellite navigation can directly provide position information and reduce the divergence of INS errors, but satellite signals are blocked in indoor environments; some wireless location methods commonly used in indoor positioning, such as WiFi, UWB, etc., need a pre-installed infrastructure, which increases the workload and cost of the system. In contrast, vision-based navigation has low cost and high stability, which can assist inertial navigation systems well.

3. METHODOLOGY

As shown in Figure 1, the inertial navigation algorithm based on scene recognition is a fusion of an inertial module and a scene recognition module. The inertial module is comprised of an INS mechanization algorithm with ZUPT and HDR, while the scene recognition module determines whether the scene is a corridor. Information from two modules are fused by a filter.
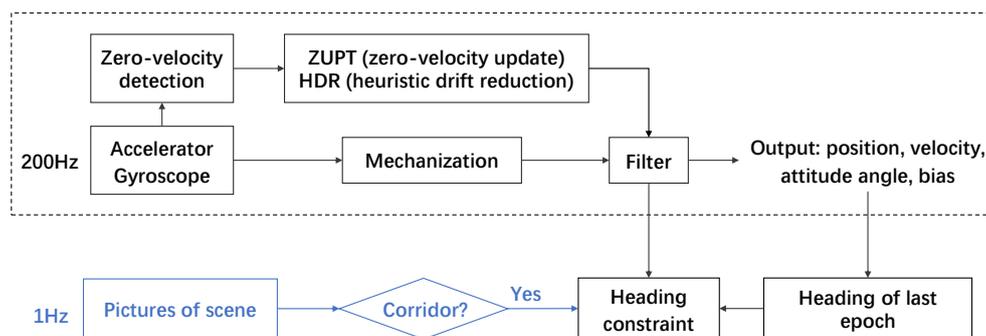

Figure 1. Main blocks in the inertial navigation algorithm based on scene recognition.

3.1 Foot-Mounted INS

The foot-mounted INS provides continuous positioning results for pedestrian navigation. The state vector of the foot-mounted INS is

$$\mathbf{X} = \begin{bmatrix} \mathbf{r} & \mathbf{v} & \boldsymbol{\varphi} & \mathbf{b}_g & \mathbf{b}_a \end{bmatrix}^T \tag{1}$$

where $\mathbf{r}$ is the position in North-East-Down (NED) geographic coordinate system; $\mathbf{v}$ is the velocity; $\boldsymbol{\varphi}$ is the attitude; $\mathbf{b}_g$ and $\mathbf{b}_a$ is the biases for the gyroscope and accelerometer, respectively. Each element in the state vector contains three elements, corresponding to three directions.

The error state estimation of continuous-time systems of the strapdown inertial navigation system (Titterton et al., 2004) is

$$\begin{cases} \delta\dot{\mathbf{r}}^n = -\boldsymbol{\omega}_{en}^n \times \delta\mathbf{r}^n + \delta\boldsymbol{\theta} \times \mathbf{v}^n + \delta\mathbf{v}^n \\ \delta\dot{\mathbf{v}}^n = \mathbf{C}_b^n \delta\mathbf{f}^b + \mathbf{C}_b^n \mathbf{f}^b \times \boldsymbol{\varphi} - (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \delta\mathbf{v}^n \\ \qquad + \mathbf{v}^n \times (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) + \delta\mathbf{g}^n \\ \dot{\boldsymbol{\varphi}} = -\boldsymbol{\omega}_{in}^n \times \boldsymbol{\varphi} + \delta\boldsymbol{\omega}_{in}^n - \mathbf{C}_b^n \delta\boldsymbol{\omega}_{ib}^b \\ \delta\dot{\mathbf{b}}_g = -\dfrac{1}{\boldsymbol{\tau}_g} \\ \delta\dot{\mathbf{b}}_a = -\dfrac{1}{\boldsymbol{\tau}_a} \end{cases} \tag{2}$$

or

$$\delta\dot{\mathbf{X}}(t) = \mathbf{F}(t)\delta\mathbf{X}(t) + \mathbf{G}(t)\mathbf{w}(t) \tag{3}$$

The discretized form of the above model is

$$\delta\mathbf{X}_k = \boldsymbol{\phi}_{k,k-1}\delta\mathbf{X}_{k-1} + \int_{t_{k-1}}^{t_k} \boldsymbol{\phi}_{k,\tau}\mathbf{G}(\tau)\mathbf{w}(\tau)d\tau \tag{4}$$

where

$$\boldsymbol{\phi}_{k,k-1} = e^{\int_{t_{k-1}}^{t_k} \mathbf{F}(\tau)d\tau} \approx \mathbf{I} + \mathbf{F}(t_{k-1})\Delta t \tag{5}$$

3.2 Zero Velocity Update (ZUPT)

ZUPT can assit foot-mounted INS to decrease the velocity divergence of the positioning error. In pedestrian navigation, the zero-velocity state is generally determined from the period when the footsteps are still on the ground. Zero-velocity detection is generally based on the output of the gyroscope and accelerometer, and there are three common methods: Acceleration-Moving Variance Detector (AMV), Angular Rate Energy Detector (ARE), Acceleration-Magnitude Detector (AM) (Skog et al., 2010). In this paper, we use the first two zero-velocity detection methods. When both AMV and ARE determine that the footsteps are stationary, the zero-velocity state can be determined, and then the ZUPT can be implemented, as shown in Figure 2.
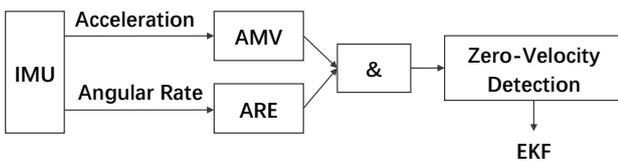

Figure 2. Main blocks in the zero-velocity detection.

The principle of AMV is as follows

$$T_{AMV} = \frac{1}{N}\sum_{k=1}^{N} \left\| \mathbf{f}_k^b - \bar{\mathbf{f}}^b \right\|^2 \tag{6}$$

The principle of ARE is as follows

$$T_{ARE} = \frac{1}{N}\sum_{k=1}^{N} \left\| \boldsymbol{\omega}_{ib,k}^b \right\|^2 \tag{7}$$

The measurement model of ZUPT is

$$\mathbf{z}_k = \mathbf{H}\delta\mathbf{x}_{k|k} + \mathbf{n}_k \tag{8}$$

$$\mathbf{R}_k = E\left(\mathbf{n}_k\mathbf{n}_k^T\right) \tag{9}$$

where $\mathbf{z}_k$ is the velocity error measurements, $\mathbf{H}$ is the measurement matrix, $\mathbf{H} = \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix}$. $\mathbf{n}_k$ is the measurement noise, and $\mathbf{R}_k$ is the covariance matrix.

3.3 Heuristic Drift Reduction (HDR)

HDR can assit foot-mounted INS to decrease the heading divergence of the positioning error and estimate the gyroscope bias.

As a matter of fact, in indoor buildings, corridors and aisles are generally straight, so pedestrians are more inclined to walk in straight lines in indoor environments. Based on this hypothesis, Borestein et al. proposed the HDR algorithm (Borenstein et al., 2009). When the pedestrian is walking in a straight line, the HDR algorithm can detect it, so that the gyroscope bias can be corrected to reduce the heading error.

Different from how traditional HDR determines whether a pedestrian is walking in a straight line, in this paper we take a different approach: the algorithm judges by the direction of the pedestrian's trajectory, rather than the heading of the footsteps, as shown in Figure 3. When a pedestrian is walking in a straight line, the direction of adjacent footsteps will also change (around 1~2°), so the trajectory can better reflect the real direction.
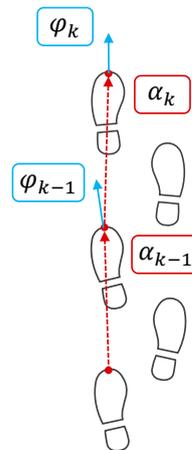

Figure 3. Comparison of heading orientation and trajectory orientation. Blue arrow is the heading orientation while red arrow is the trajectory orientation.

The orientation change is computed as:

$$\Delta\alpha_k = \alpha_k - \alpha_{k-1} \tag{10}$$

where α_k is the direction of the foot at current sample k computed as $\alpha_k = \frac{\mathbf{r}_k(2) - \mathbf{r}_{k-1}(2)}{\mathbf{r}_k(1) - \mathbf{r}_{k-1}(1)}$, $\mathbf{r}_k$ is the position in NED geographic coordinate system.

When the pedestrian is walking in a straight line, the orientation change will be small (below a given threshold), then a measurement m_k will be put into the EKF to correct the heading error; if the orientation change is larger than the given threshold, the pedestrian trajectory is not considered a straight line, and then no correction will be put into the EKF:

$$m_k = \begin{cases} \Delta\alpha_k & , |\Delta\alpha_k| \leq th \\ 0 & , otherwise \end{cases} \tag{11}$$

3.4 Scene Recognition

The scene recognition module determines whether the scene is a corridor, and then determines whether to add a heading constraint.

Zhou et al. (Zhou et al., 2018) chose several popular CNN architectures and then trained them on the Places dataset to create CNN models. In this paper, we use the Places-ResNet model trained by Zhou et al. to carry out the scene-recognition experiment. A test picture is shown in Figure 4.

During the data collection, the pedestrian needs to hold a cell phone for video recording. Then, the video is split into pictures frame-by-frame, and the pictures are put into the Places-ResNet model for scene recognition.

Top-1: corridor (0.874)
Top-2: hospital room (0.043)
Top-3: elevator lobby (0.020)
Top-4: clean room (0.016)
Top-5: basement (0.011)

Figure 4. A test picture for scene recognition. The top 5 predictions and their probabilities are shown on the right.

3.5 Integration of Foot-mounted INS and Scene Recognition

Figure 1 shows the main blocks in the pedestrian inertial navigation algorithm based on scene recognition. The parts in blue are how the constraints based on scene recognition work. When the scene recognition model determines that the environment where the pedestrian is walking is a corridor, a straight-line constraint will be implemented.

After integrating all the error correction methods into the EKF, the measurement vector will be:

$$\mathbf{m}_k = \begin{bmatrix} \Delta\mathbf{v}_k & \Delta\boldsymbol{\varphi}_k \end{bmatrix} \tag{12}$$

The measurement matrix will be a 4 by 15 matrix:

$$\mathbf{H} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & [0 \quad 0 \quad 1] & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} \end{bmatrix} \tag{13}$$

4. EXPERIMENTS AND RESULTS

4.1 Hardware

The IMU we use is the WHU-WearTrack system, and it is mounted on the left foot, as shown in Figure 5. The entire system is only 3.7*3.2*2.0 cm in size and 50 grams in weight, and the IMU is sampled at 200 Hz. The camera we use for scene recognition is an iPhone 13 camera.

Figure 5. IMU tied on the left foot.

4.2 Experiment

We test the algorithm inside the Xinghu Building of Wuhan University. When walking, the pedestrian needs to hold a cell phone for video recording, with the IMU mounted on the left foot. The real environment of the experiment is shown in Figure 6.

Figure 6. Experiment environment.

Since the algorithm is for a specific indoor environment with long corridors, the test environment is relatively simple. We select one set of data for display: the pedestrian walks through the narrow corridor four times, as well as going upstairs and downstairs. The

length of the corridor is approximately 40 meters, and the entire experiment lasts more than 5 minutes.

The whole experiment is divided into two parts: Experiment 1 and Experiment 2. For convenience, INS-EKF-ZUPT is abbreviated as IEZ, scene recognition is abbreviated as SR. In Experiment 1, we use the raw data of the gyroscope and accelerometer to implement the algorithm and make a comparison of the positioning results of three modes (real trajectory, IEZ+HDR, IEZ+HDR+SR). Because WHU-WearTrack already compensates for sensor biases according to the stationary time before pedestrian starts, different positioning modes have similar results. To better demonstrate the performance of the algorithm, we add biases to the sensor raw data in Experiment 2.

(1) Experiment 1

The horizontal positioning results of three modes (real trajectory, IEZ+HDR, IEZ+HDR+SR) are as shown in Figure 7. Vertical positioning results are ignored because the algorithm is for a specific indoor environment with long corridors, which only works horizontally.
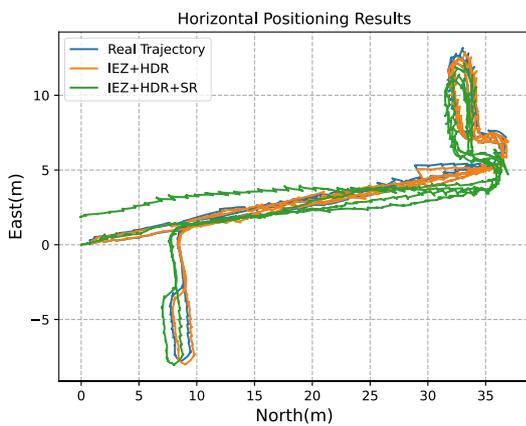

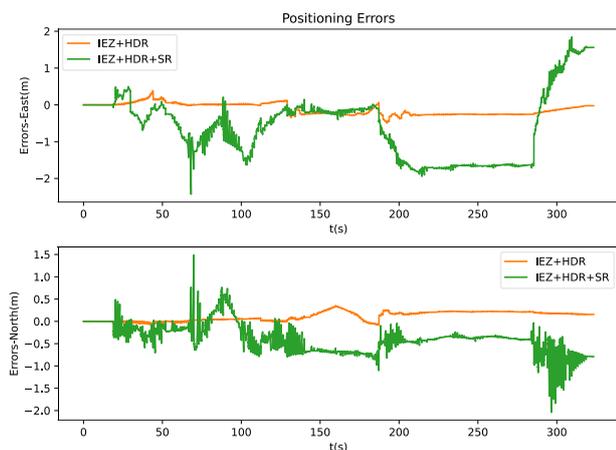**Figure 7**. Horizontal positioning results (Experiment 1).

Figure 8. Horizontal positioning errors (Experiment 1).

In Figure 7, the positioning result of IEZ+HDR is close to the real trajectory, while the result of IEZ+HDR+SR has larger deviation. The end point of IEZ+HDR+SR deviates from the real trajectory by around 2 meters. In Figure 8, the errors of IEZ+HDR are less

than 0.5 meters, while the maximum positioning errors of IEZ+HDR+SR in North and East directions are both around 2 meters.

	IEZ + HDR		IEZ + HDR + SR	
	East	North	East	North
Mean	-0.1153	0.1123	-0.5868	-0.3920
RMS	0.1893	0.1508	1.0558	0.5037
STD	0.1501	0.1006	0.8777	0.3163

Table 1. Horizontal positioning errors (Experiment 1).

After more than 300 seconds of navigation, the RMS of the positioning error of IEZ+HDR+SR is less than 1.1 meters, but the performance is better without SR, with the RMS less than 0.2 meters.

Figure 7, Figure 8, and Table 1 show that the algorithm based on scene recognition may not improve the results, and may even have a negative effect. However, WearTrack itself has already corrected the gyroscope bias, and it's not clear how the positioning results are affected by SR. Therefore, we carried out Experiment 2.
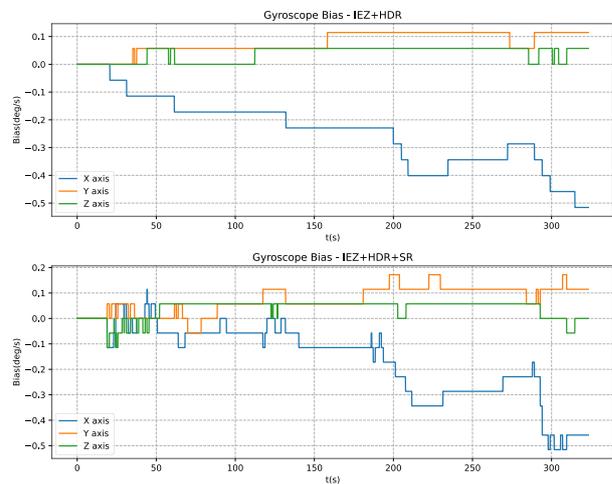

Figure 9. Estimation of gyro bias (Experiment 1)

The results in Figure 9 show that the algorithm based on scene recognition can improve the observability of vertical gyroscope bias. After using scene recognition, the correction of the vertical gyroscope bias (z-axis) is more frequent. Because WearTrack itself has already corrected the gyroscope bias, the effectiveness of scene recognition in correcting gyroscope bias are unclear. Results in Experiment 2 are more convincing.

(2) Experiment 2

In many cases, the IMU used for pedestrian navigation does not perform as well as WHU-WearTrack, and the pedestrians do not have the patience to stand still for a few seconds before starting to compensate the sensor biases. Therefore, we add biases to the sensor raw data to further test the performance of the algorithm based on scene recognition. The gyroscope biases in X, Y and Z axis was set to 0.5 rad/s.
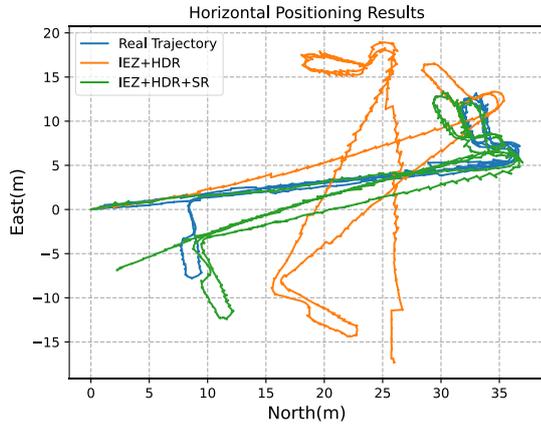
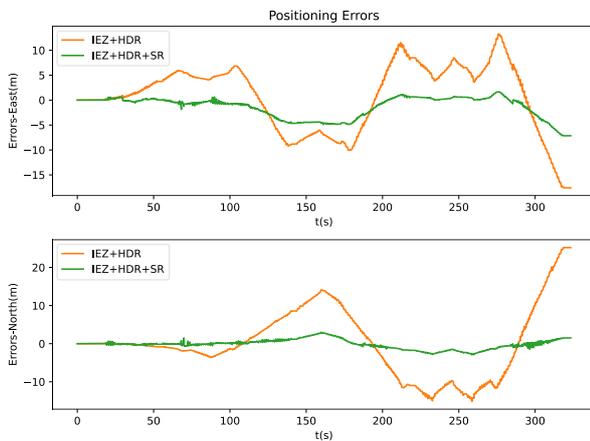**Figure 10**. Horizontal positioning results (Experiment 2).

Figure 11. Horizontal positioning errors (Experiment 2).

In Figure 10, the positioning result of IEZ+HDR has a larger deviation from the real trajectory, while the result of IEZ+HDR+SR is closer to the real trajectory. In Figure 11, the errors of IEZ+HDR are more than 20 meters, while the maximum positioning errors of IEZ+HDR+SR are less than 10 meters.

	IEZ + HDR		IEZ + HDR + SR	
	East	North	East	North
Mean	0.7297	0.2202	-1.3807	-0.1730
RMS	6.8157	9.0210	2.5507	1.2325
STD	6.7765	9.0183	2.1447	1.2203

Table 2. Horizontal positioning errors (Experiment 2).

After adding gyroscope biases, the RMS of the positioning errors of IEZ+HDR and IEZ+HDR+SR both increase, but the RMS of IEZ+HDR has greater increase. As for IEZ+HDR, the RMS of errors increases from less than 0.2 meters to more 6 meters; as for IEZ+HDR+SR, the RMS of errors increases from around 1 meter to around 2 meter. In Experiment 2, after using SR, the RMS and STD of horizontal directions are reduced by over 50 %.

Figure 10, Figure 11, and Table 2 show that the algorithm based on scene recognition can improve the positioning results when the IMU has significant sensor biases. After using the algorithm, the positioning errors of horizontal directions are reduced by over 50 %.
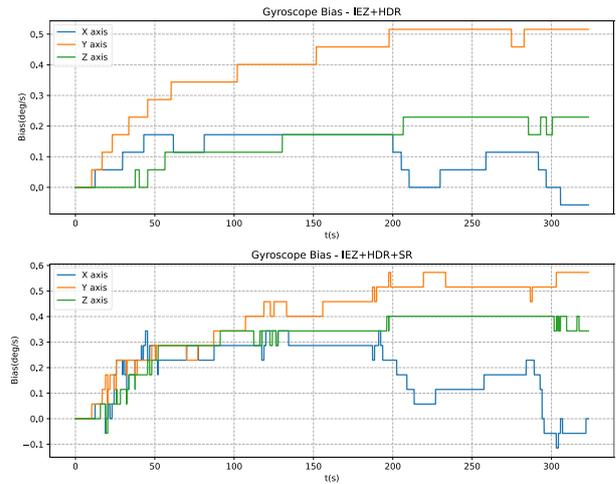

Figure 12. Estimation of gyro bias (Experiment 2)

Similar to Experiment 1, the results in Figure 12 show that the correction of the vertical gyroscope bias (z-axis) is more frequent after using scene recognition. After the gyroscope biases in X, Y and Z axis is set to 0.5 rad/s, the difference between the positioning results of IEZ+HDR and IEZ+HDR+SR is more obvious, which indicates that the algorithm can improve the observability of the vertical gyroscope bias.

5. CONCLUSION

We proposed a pedestrian inertial navigation algorithm based on scene recognition to reduce the heading drift and improve the observability of vertical gyroscope bias. The scene recognition module only needs to identify whether the scene is a corridor, so the studies are restricted to a horizontal plane. When the scene recognition module determines that the pedestrian is walking in a corridor, a straight-line constraint will be implemented.

We compared the positioning results before and after using the algorithm. Experiments show that the algorithm can better improve the positioning results when the IMU has significant sensor biases, and it can improve the observability of z-axis gyroscope bias. Especially when the IMU used for pedestrian navigation has sensors with low-performance, or when the pedestrians do not have the patience to stand still for a few seconds before starting to compensate the sensor biases, the algorithm can better improve the positioning results.

For future work, more scene options can be added to the scene recognition module, such as staircase, elevator, etc. In addition to heading constraints, height constraints can also be added to the foot-mounted INS. These approaches should help scene recognition better assist pedestrian inertial navigation.

REFERENCES

Biswas, J., Veloso, M., 2010. WiFi localization and navigation for autonomous indoor mobile robots, in: 2010 IEEE International Conference on Robotics and Automation. Presented at the 2010 IEEE International Conference on Robotics and Automation, pp. 4379–4384. https://doi.org/10.1109/ROBOT.2010.5509842

Borenstein, J., Ojeda, L., Kwanmuang, S., 2009. Heuristic reduction of gyro drift in IMU-based personnel tracking systems, in: Optics and Photonics in Global Homeland Security V and Biometric Technology for Human Identification VI. Presented at the Optics and Photonics in Global Homeland Security V and Biometric Technology for Human Identification VI, SPIE, pp. 244–254. https://doi.org/10.1117/12.816921

De Angelis, G., Moschitta, A., Carbone, P., 2016. Positioning Techniques in Indoor Environments Based on Stochastic Modeling of UWB Round-Trip-Time Measurements. IEEE Trans. Intell. Transp. Syst. 17, 2272–2281. https://doi.org/10.1109/TITS.2016.2516822

Faragher, R., Harle, R., 2015. Location Fingerprinting With Bluetooth Low Energy Beacons. IEEE J. Sel. Areas Commun. 33, 2418–2428. https://doi.org/10.1109/JSAC.2015.2430281

Foxlin, E., 2005. Pedestrian tracking with shoe-mounted inertial sensors. IEEE Comput. Graph. Appl. 25, 38–46. https://doi.org/10.1109/MCG.2005.140

Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K., 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. ArXiv160207360 Cs.

Jiménez, A.R., Seco, F., Prieto, J.C., Guevara, J., 2010. Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU, in: Navigation and Communication 2010 7th Workshop on Positioning. Presented at the Navigation and Communication 2010 7th Workshop on Positioning, pp. 135–143. https://doi.org/10.1109/WPNC.2010.5649300

Rajagopal, S., 2008. Personal dead reckoning system with shoe mounted inertial sensors. Master's Degree Proj. Stockh. Swed.

Sengupta, A., Ye, Y., Wang, R., Liu, C., Roy, K., 2019. Going Deeper in Spiking Neural Networks: VGG and Residual Architectures. Front. Neurosci. 13.

Skog, I., Handel, P., Nilsson, J.-O., Rantakokko, J., 2010. Zero-Velocity Detection—An Algorithm Evaluation. IEEE Trans. Biomed. Eng. 57, 2657–2666. https://doi.org/10.1109/TBME.2010.2060723

Tai, Y., Yang, J., Liu, X., 2017. Image Super-Resolution via Deep Recursive Residual Network. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3147–3155.

Titterton, D., Weston, J.L., Weston, J., 2004. Strapdown Inertial Navigation Technology. IET.

Zhong, Z., Jin, L., Xie, Z., 2015. High performance offline handwritten Chinese character recognition using GoogLeNet and directional feature maps, in: 2015 13th International Conference on Document Analysis and Recognition (ICDAR). Presented at the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 846–850. https://doi.org/10.1109/ICDAR.2015.7333881

Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A., 2018. Places: A 10 Million Image Database for Scene Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 40, 1452–1464. https://doi.org/10.1109/TPAMI.2017.2723009