Pose Tracking of Moving Sensor using Monocular Camera and IMU Sensor

Sukwoo Jung, Seho Park, and KyungTaek Lee*
Contents Convergence Research Center, Korea Electronics Techonology Institute,
Seoul, South Korea
[e-mail: swjung@keti.re.kr, sehopark@keti.re.kr, ktechlee@keti.re.kr]
*Corresponding author: KyungTaek Lee

Abstract

Pose estimation of the sensor is important issue in many applications such as robotics, navigation, tracking, and Augmented Reality. This paper proposes visual-inertial integration system appropriate for dynamically moving condition of the sensor. The orientation estimated from Inertial Measurement Unit (IMU) sensor is used to calculate the essential matrix based on the intrinsic parameters of the camera. Using the epipolar geometry, the outliers of the feature point matching are eliminated in the image sequences. The pose of the sensor can be obtained from the feature point matching. The use of IMU sensor can help initially eliminate erroneous point matches in the image of dynamic scene. After the outliers are removed from the feature points, these selected feature points matching relations are used to calculate the precise fundamental matrix. Finally, with the feature point matching relation, the pose of the sensor is estimated. The proposed procedure was implemented and tested, comparing with the existing methods. Experimental results have shown the effectiveness of the technique proposed in this paper.

1. Introduction

The position and orientation information of an object is known as pose. Position represents the three-dimensional location of the object, while orientation can be expressed as a series of consecutive rotations. Estimating the pose of the sensor is important in navigation system. Pose estimation has been studied over the past few decades and various technologies have been researched. There are several approaches to estimate the pose of the sensor. The well-known method is using inertial or visual sensors.

Inertial Measurement Unit (IMU) sensors have been widely used for pose estimation in various applications including robots, aircrafts and navigation systems. IMU sensors normally contain three orthogonal accelerometer axis and three orthogonal gyroscope axis. IMU sensor can obtain the data in high-frequency rate and calculate the accurate results without complex operation. Therefore, IMU sensor can track the fast and abrupt movements. In addition, it can be small size, light weight, low cost, and adopt wireless communication technologies. Furthermore, they are robust to the illumination changes and visual occlusions. However, it has a drift problem accumulated from the position estimation [1, 28]. Therefore, additional sensor is normally used to aid the IMU sensor in the pose estimation. Some researches proposed to use Global Positioning System (GPS) [2, 25, 26] with IMU sensor to overcome the problem of a long-term drift. However, the major problem of the GPS is intermittent loss of signal, especially for indoor environment [3].

Vision-based tracking has high accuracy for pose estimation [4-6]. It can precisely track the relative motion between the camera and objects by measuring the movement of selected features in the consecutive image sequences. There are various methods in feature detection. Local detector such as SIFT (Scale-Invariant Feature Transform) [7] and SURF (Speeded-Up Robust Feature) [8] are invariant to the rotation and illumination change [9]. BREIF [10], ORB [11], AKAZE [12] are also well known feature detectors for tracking. Those feature points can be used to calculate the pose of the camera [13]. However, vision sensor normally suffers from lack of robustness, low frequency of data acquisition, and high computational cost. The main problem of vision-based tracking systems is instability with fast moving dynamic scene because of the loss of visual features. In addition, vision-based tracking systems are severely affected by occlusion matter. To overcome these problems, some researchers proposed approaches to take advantage of both the vision sensor and inertial sensor [14-16].

Typically, measurements of visual and inertial sensors are combined with filtering method [27]. Kalman filter methods are normally selected to perform sensor fusion by integrating measurements from various sensors. Kim and Park [17] developed a sensor fusion framework by combining Radar, Lidar and camera with Extended Kalman Filter (EKF) for pose estimation. Foxlin and Naimak [18] used both the inertial and vision data by a complementary Kalman filter for a pose estimation based on fiducials. Rehbinder and Ghosh [19] adopted gyroscope to measure the orientation, and proposed a vision-based system focusing on reducing the drift error of the IMU framework. Although several researches showed that the inertial data can improve the vision-based tracking, the optimal way to integrate the data from inertial sensor and vision sensor is yet to be brought up. Moreover, the integration of inertial data often causes instabilities in position estimation [20, 24].

The purpose of this paper is to overcome the instabilities of sensors and increase the accuracy of the pose estimation by using camera and IMU sensor. The proposed approach fused the data from the IMU sensor and the camera to estimate the pose of the sensor more precisely and more robustly. As a result, the proposed method can estimate the pose of the sensor in real time, and it can be applied for various applications including robot, autonomous driving, etc.

The remainder of this paper is organized as follows. Section 2 explains the approach to the proposed pose estimation method. Section 3 describes the methodology of the proposed pose estimation algorithm in details, and section 4 shows the results of the experimentation with indoor datasets. Finally, the summary of the paper is given in the last section.

2. Approach to The Camera-IMU Fusion

2.1 Feature Matching

To calculate the pose of the camera using the image sequences, visual features should be obtained from the image sequences. Feature points can be detected by the detector including SIFT, SURF, AKAZE, etc. Because of the lower computational cost and moderate performance, ORB [11] is widely used in feature tracking applications. Then, descriptors are used to find the matching relations in the given sets of feature points. **Fig. 1** shows the example of feature point matching result with ORB detector and descriptor. Using feature point matching relations and camera calibrated parameters, the camera extrinsic parameters can be calculated.
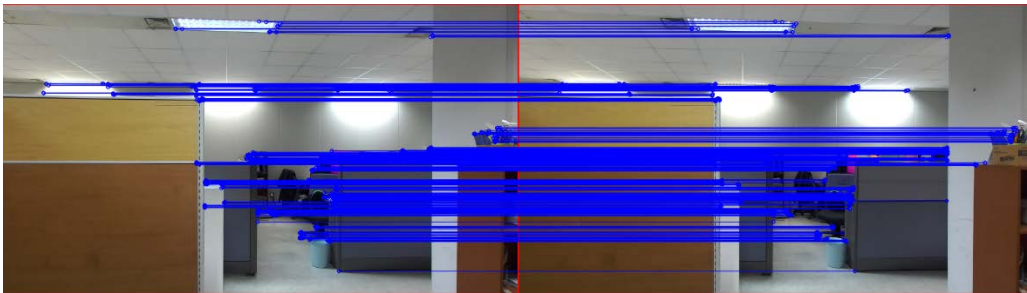

Fig.1. The example of ORB feature matching result.

However, as mentioned earlier, the pose estimation results can be unstable because of the dynamic image sequences. Illumination change, blur, and motion of the objects can adversely affect the position or matching relation of the feature points in the images. Consequently, the estimated pose results can be instable. To overcome this problem, we used epipolar geometry to remove the outliers of the feature points [13, 23].

If the camera is not moving, the same feature points must lie on the same position in the next image frame. Then, the distances of the corresponding feature points are calculated, and the feature points which have distances higher than the threshold can be removed from the pose estimation. On the other hand, if the camera is moving, the same feature points must lie on the epipolar line in the next image frame. The feature points which have distances higher than the threshold are considered as outliers. The epipolar line and the outliers are shown in **Fig. 2**.

(a)   (b)

Fig. 2. The example of outlier removal process with epipolar geometry (a) epipolar line example (b) outlier detected by epipolar geometry. Red dots are considered as outlier.

However, if the feature points have a large portion of outliers in an image (e.g. blur, occlusion), it is difficult to determine the outliers properly with epipolar geometry. The epipolar line is drawn with the fundamental matrix which is calculated from the matching feature points and camera parameters. If the matching of feature points is not accurate, the fundamental matrix and epipolar line become inaccurate. As a result, the outlier rejection becomes challenging. To overcome this instability of the vision sensor, we decided to use IMU sensor as the initial estimator of the camera pose rather than feature matching.

2.2 Camera-IMU Fusion

Normally, IMU sensor has 6 DoF (degree of freedom) and can obtain linear and angular acceleration with high-frequency rate. As we mentioned above, IMU sensors generate fast signals with high rate during dynamic motions. However, they are sensitive to accumulated drift error because of the double integration process in the pose estimation procedure. On the other hand, camera sensors can precisely estimation the ego-motion even with the long estimation time. However, they have weaknesses in blurred features under fast and unpredicted motions. **Fig. 3** shows example of the pose estimated from the vision and inertial sensor. It shows the translation change in x-axis over frame. We obtained the data from each sensor with the same frame rate (30fps). As the sensor captured the image in static state, the ground truth must be zero. Both errors from each sensor are low because the camera is in static state, but they show distinctively different tendency. IMU data has high rate of flipping error and because of the drift error, the error is slightly increasing. On the other hand, vision sensor has no. drift error, but sometimes error occur sharply because of the unpredicted motion of the image (e.g. illumination change, motion of the object). The aim of Camera-IMU sensor fusion is to overcome some fundamental limitations of visual-only tracking and IMU-only tracking by using the both sensor complementarily.
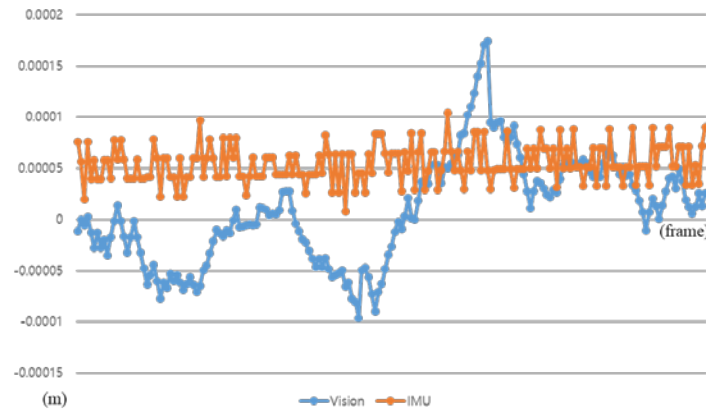
Fig. 3. The example of pose estimation from vision and IMU sensor.

As the IMU sensor can estimate the pose change of the consecutive image invariant to the abrupt motion, the transformation matrix obtained from IMU sensor can be used to the initial value for the fundamental matrix calculation. After removing the outliers with the initial fundamental matrix, the feature matching method will produce more precise transformation matrix. The overall procedure of the proposed method is shown in **Fig. 4**. Detail explanation will be given in the next section.
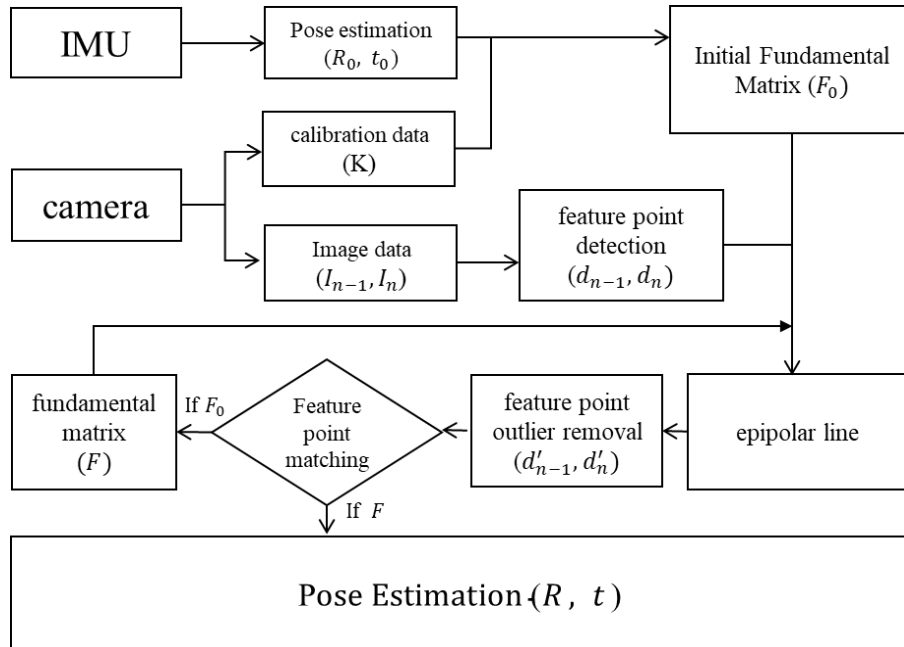

Fig. 4. Overall procedure of the proposed pose estimation method.

3. Proposed Pose Estimation Method

The goal of the proposed pose estimation is to find the corresponding feature points belonging to the same object in consecutive image sequences. Firstly, the connected pixels in the first image of the image sequences are marked as the target feature points. To obtain the connected pixels, the detector is used for the feature point detection. The extracted features can be colors, texture, edges, GLOH, SURF, SIFT, etc. The proposed method used ORB because of its fast and relatively robust feature detection property. Secondly, a tracking of feature points is performed to determine the location of the target feature points in the next image sequences. It is especially difficult to track feature points at the moving camera situation. Finally, with the obtained feature matching relations, the pose change of the sensor can be calculated. In the proposed method, the IMU estimates initial value of rotation (R_0) and translation (t_0). Since the raw data has a bias error, we applied Kalman filter [21] to reduce the error of the raw data. The initial fundamental matrix is calculated from the following equation.

$$F_0 = K^{-T} R_0 K^T [K R_0{}^T t_0]_X \quad (1)$$

where K are the calibration data of the camera. Calibration data is the intrinsic parameters of the camera. It can be calculated using the bundle adjustment.

The feature points (d'_{n-1}, d'_n) are detected from the consecutive images (I_{n-1}, I_n) captured by the camera. The well-known detector ORB [11] is used in the proposed method. Using the initial fundamental matrix F_0 and the feature points d_{n-1}, the epipolar line can be drawn in the consecutive image I_n. Then, the epipolar line and feature point distance is calculated with the equation (2).

$$d = \frac{|au + bv + c|}{\sqrt{a^2 + b^2}} \quad (2)$$

where u and v are the coordinates of the feature point, and a, b and c represent the coefficient of the epipolar line, $ax + by + c = 0$.

Feature points with distance above the threshold is considered outliers. The optimal threshold value is obtained by experiments. We applied a threshold value for 1.5 pixel, which means if the corresponding feature point is far from epipolar line more than 1.5 pixel, that feature point is considered as an outlier in the pose estimation procedure.

After the outliers are removed from the feature points (d'_{n-1}, d'_n), these feature points matching relations are used to calculate the precise fundamental matrix (F). Then, the above procedure is repeated again to remove a few outliers remaining in the feature points. Finally, with the feature point matching relation, the pose of the sensor is estimated.

Fig. 5 shows the error estimated from the feature point matching related to the iteration of outlier removal process. We tested with 100 dataset of indoor scene, and the average error is plotted according to the iterations. As it can be seen in **Fig. 5**, five times of iteration is enough for the outlier removal process.
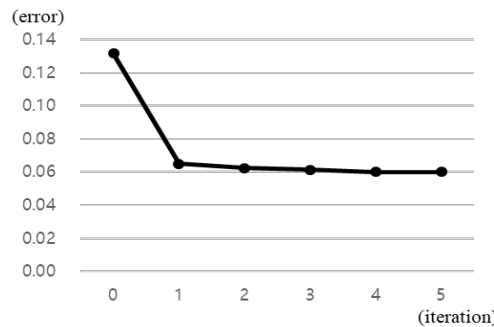
Fig. 5. Outlier removal result related to the iteration

In the next section, experimental results of the various pose estimation methods are demonstrated.

4. Experiments and Results

For the experiment, we used commercial sensor ZED 2 (StereoLabs, San Francisco, USA) which includes camera and IMU sensor in one platform. The raw IMU sensor data are transferred to the host computer at 100 Hz via the USB 3.0 port. Camera can capture image sequences in 30fps with Full-HD resolution. We captured the both data in 30fps to synchronize the data from camera and IMU sensor. The time elapsed, which is the difference between the time instant when the acquisition process starts and the time instant when a new image frame is available, is collected together with the sensor data. The drift of the IMU sensor in translational axis is 0.35% and the drift in rotational axis is 0.005 °/m. The configuration of the experiment is shown in **Fig. 6**. After collecting data with the sensor, the data were transferred to the personal computer and the various pose estimation methods were tested with an i7-4790K processor with 16GB memory and a Windows 10 operating system.


**Fig. 6.** Configuration of the experiment.

Four pose estimation methods were experimented. The proposed method and the pose estimation method using only feature point matching (ORB) were implemented using OpenCV library [22]. Also, the previous method using Kalman filter to integrate the result of ORB and IMU data and pose estimation method using only IMU were implemented for the comparison.

We collected the data of moving sensor from 1m to 5m for 10 times (50 datasets). The example of test results for each dataset is shown in **Fig. 7**. The results of four pose estimation methods were compared with the given dataset. (a) shows one test result with the dataset moving 1m. The position results of various pose estimation methods must converge to 1m. (b) shows one test result with the dataset moving 2m. The position results of each pose estimation methods must converge to 2m. (c) shows one test result with the dataset moving 3m, which the results of pose estimation must converge to 3m. (d) shows one test result with the dataset moving 4m, which results of pose estimation must converge to 4m. (e) shows one test result with the dataset moving 5m, which results of pose estimation must converge to 5m. 'proposed' is the pose estimation method proposed in this paper. 'IMU' is the method using only IMU to estimate the position of the sensor. 'ORB' is the method using feature matching to calculate the position of the sensor. 'IMU+ORB KF' is the previous method using both IMU and camera sensor with the Kalman Filter.

(a) 1m dataset   (b) 2m dataset
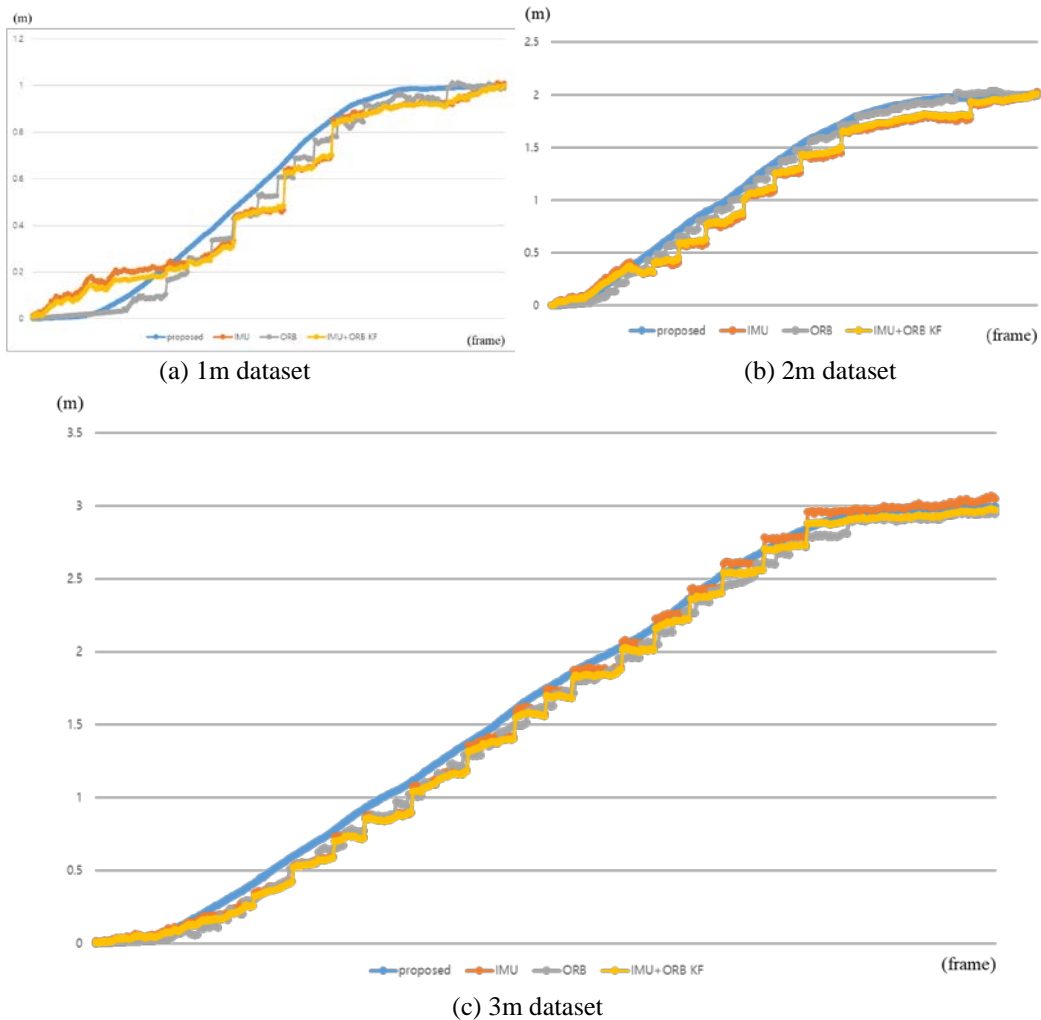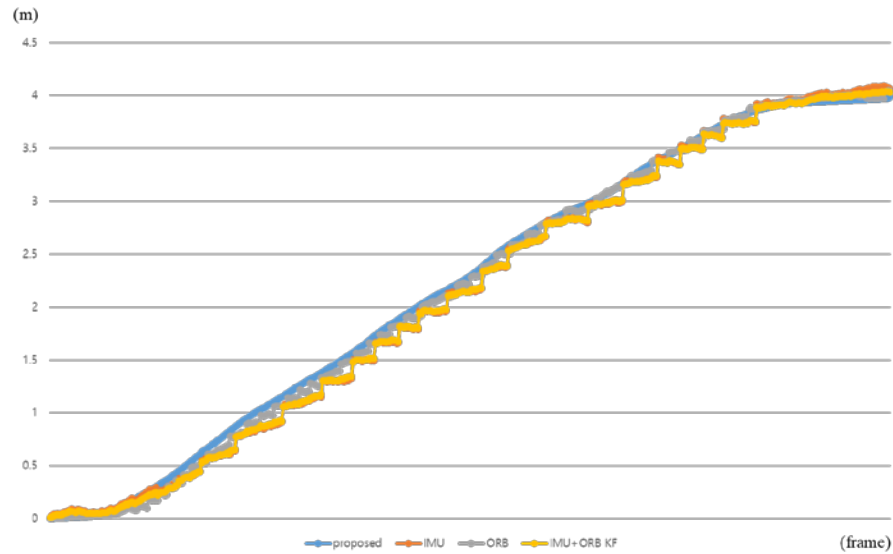
(c) 3m dataset
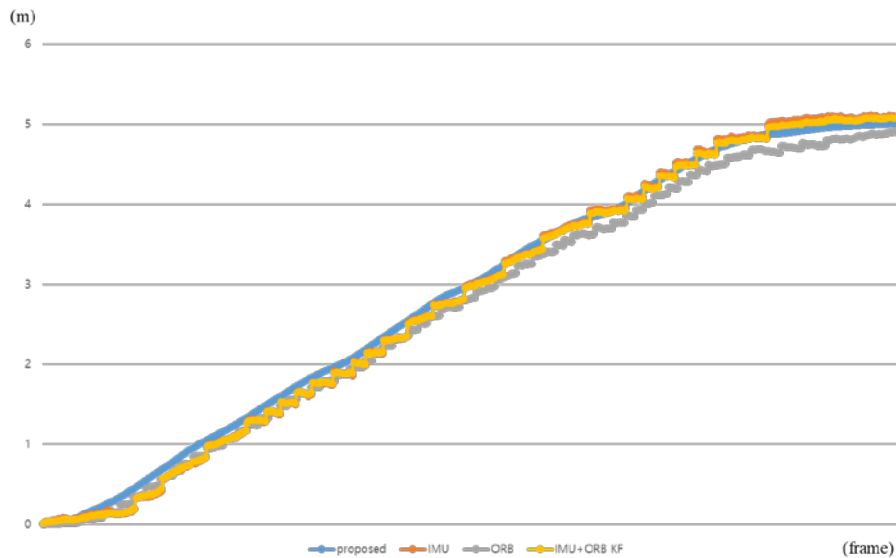
(d) 4m dataset

(e) 5m dataset

Fig. 7. Pose estimation result for the example dataset.

As it can be seen in the graph, the pose estimation result from the IMU sensor (orange line) has the largest error because of the drift. The value of IMU sensor is over estimated in the end of the measurement. Pose estimation from the ORB feature matching (grey line) fluctuates because of the instability of the image. In a complex scene, the error of feature matching might increase significantly. As the two sensor data were both used in the proposed method (blue line) and in the Kalman Filtering method (yellow line), they showed improved performance in the given experiment. In addition, the proposed method shows more stable tendency compared to the other methods. The average position of the various methods for each datasets are shown in **Table 1**.

Table 1. Average position of the tested algorithms.

Algorithm	Proposed	IMU	ORB	IMU+ORB KF
Data set1 (m)	0.99997	1.00986	0.98998	1.00101
Data set2 (m)	2.00021	2.02598	1.99001	2.00990
Data set3 (m)	2.99988	3.05017	2.94997	2.97474
Data set4 (m)	3.98997	4.06004	4.03996	4.04022
Data set5 (m)	5.00996	5.09877	4.89899	5.07997
Standard Deviation	0.00048	0.00896	0.02511	0.01170

The average error of each dataset can be shown in **Fig. 8**. The method using IMU shows the largest error, and the error increases when the data acquisition time increases. The method using feature matching algorithm shows randomized error. The error of Kalman Filtering method is slightly smaller than the feature matching or IMU method. The proposed method shows the smallest error compared to the other algorithms. The average error of the proposed method is 0.12%, which is better than the other methods, IMU 1.48%, ORB 1.23%, Kalman Filtering 0.81%.
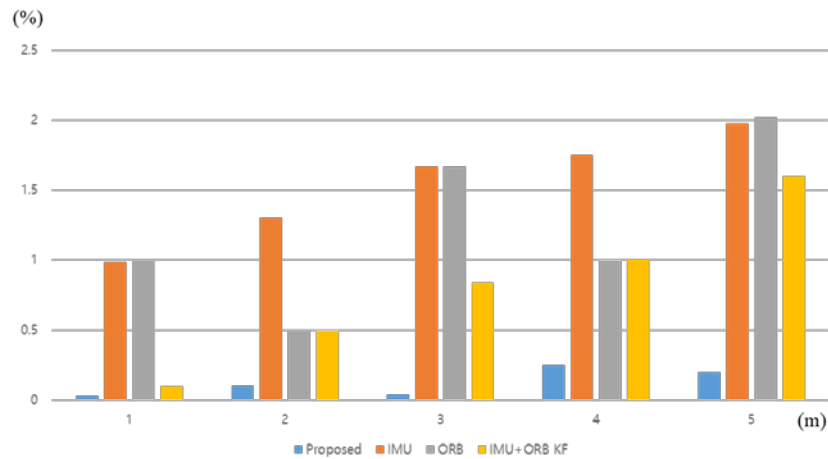
Fig. 8. Average error of various algorithms for each dataset.

We also tested each algorithm with the data which were collected by the cart moving in a circle inside the office. The **Fig. 9** shows the position tracking result with the four algorithms including the proposed method. The trajectory of the proposed method shows the smoothest curve which seems to be the real movement of the sensor. The Kalman Filtering method also shows smooth curve, but the end of the point doesn't match the beginning of the curve. Since the dataset shown in **Fig. 9** have no. ground truth, the error analysis is not appropriate in this situation. As it can be seen in **Fig. 7-9** and **Table 1**, the proposed method shows the best performance among the tested algorithms.
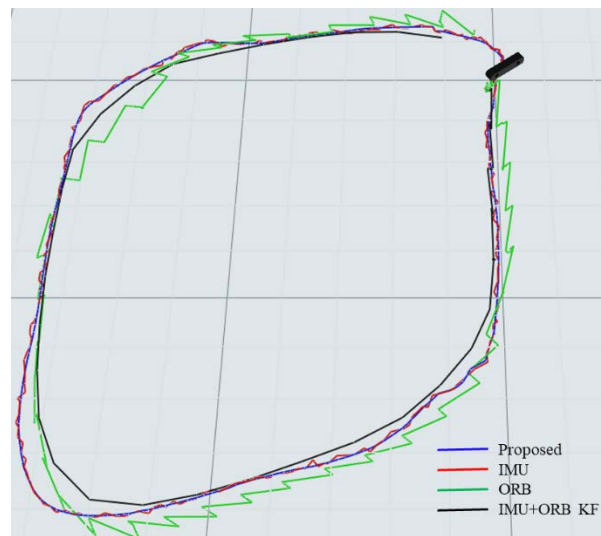
Fig. 8. Position tracking result of the indoor data.

5. Conclusion

In this Research, a new pose estimation method using image sequences and IMU sensor is proposed. The feature point matching method has a fundamental problem of instability in the dynamic scenes. The proposed method used IMU sensor to obtain the initial pose of the sensor. Then, the epipolar geometry is applied to eliminate the outliers in the feature points. With the feature points of outliers removed, matching relation generates the more accurate pose of the sensor.

One of the major attributes of the proposed method is that it improves the accuracy of feature-based pose estimation, which has weaknesses in the moving image sequence. The proposed method used IMU sensor to compensate the pose estimation procedure, and it reduces the instability error in the pose estimation.

The proposed method was implemented and was tested with the other methods. Experimentation shows that the proposed method provides better pose estimation results and is more robust compared to the other methods.

The limitation of the proposed method is the computational cost. Also, the abrupt movement will cause the algorithm to malfunction as it is the fundamental problem of the sensors. The proposed method needs optimization at abnormal situations through additional tests. More research must be taken using datasets with the ground truth data to analyze the error tendency of the pose estimation trajectory of various algorithms. Also, the research on the data from outdoor should be considered in the future work.

Acknowledgement

References

[1] A. R. Jiménez, F. Seco, J. C. Prieto, J. Guevara, "Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU," in *Proc. of 2010 7th Workshop on Positioning, Navigation and Communication*, pp. 135-143, 2010. [Article (CrossRef Link)](#)

[2] F. Caron, E. Duflos, D. Pomorski, P. Vanheeghe, "GPS/IMU data fusion using multisensor Kalman filtering: introduction of contextual aspects," *Information Fusion*, vol. 7, no. 2, pp. 221-230, 2006. [Article (CrossRef Link)](#)

[3] F. Subhan, S. Ahmed, S. Haider, S. Saleem, A. Khan, S. Ahmed, M. Numan, "Hybrid Indoor Position Estimation using K-NN and MinMax," *KSII Transactions on Internet and Information Systems,* vol. 13, no. 9, pp. 4408-4428, 2019. [Article (CrossRef Link)](#)

[4] M. Blösch, S. Weiss, D. Scaramuzza, R. Siegwart, "Vision based MAV navigation in unknown and unstructured environments," in *Proc. of 2010 IEEE International Conference on Robotics and Automation*, pp. 21-28, 2010. [Article (CrossRef Link)](#)

[5] M. E. Ragab, G. F. Elkabbany, "A Parallel Implementation of Multiple Non-overlapping Cameras for Robot Pose Estimation," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 11, pp. 4103-4117, 2014. [Article (CrossRef Link)](#)

[6] L. Li, Y. Liu, T. Jiang, K. Wang, M. Fang, "Adaptive Trajectory Tracking of Nonholonomic Mobile Robots Using Vision-Based Position and Velocity Estimation," *IEEE Transactions on Cybernetics,* vol. 48, no. 2, pp. 571-582, 2018. [Article (CrossRef Link)](#)

[7] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 1150-1157, 1999.
[Article (CrossRef Link)](#)

[8] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding*, pp. 404-417, 2006. [Article (CrossRef Link)](#)

[9] S. Jung, S. Song, M. Chang, S. Park, "Range Image Registration based on 2D Synthetic Images," *Computer-Aided Design,* vol. 94, pp. 16-27, 2018. [Article (CrossRef Link)](#)

[10] M. Calonder, V. Lepetit, C. Strecha, P. Fua, "BRIEF: Binary Robust Independent Elementary Features," in *Proc. of ECCV 2010: Proc. of the 11th European Conference on Computer Vision*, pp.778-792, 2010. [Article (CrossRef Link)](#)

[11] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. of 2011 International Conference on Computer Vision*, pp. 2564-2571, 2011. [Article (CrossRef Link)](#)

[12] P. Alcantarilla, J. Nuevo, A. Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces," in *Proc. of British Machine Vision Conference 2013*, vol. 13, pp. 1-11, 2013. [Article (CrossRef Link)](#)

[13] S. Jung, Y. Cho, D. Kim, M. Chang, "Moving Object Detection from Moving Camera Image Sequences Using an Inertial Measurement Unit Sensor," *Applied Sciences,* vol. 10, pp. 268, 2020. [Article (CrossRef Link)](#)

[14] J. L., J. A. Besada, A. M. Bernardos, P. Tarrío, J. R. Casar, "A novel system for object pose estimation using fused vision and inertial data," *Information Fusion*, vol. 33, pp. 15-28, 2017. [Article (CrossRef Link)](#)

[15] Y. Tian, Jie Zhang, J. Tan, "Adaptive-frame-rate monocular vision and IMU fusion for robust indoor positioning," in *Proc. of 2013 IEEE International Conference on Robotics and Automation*, pp. 2257-2262, 2013. [Article (CrossRef Link)](#)

[16] K. Kumar, A. Varghese, P. K. Reddy, N. Narendra, P. Swamy, M. G. Chandra, P. Balamuralidhar, "An Improved Tracking using IMU and Vision Fusion for Mobile Augmented Reality Applications," *International Journal of Multimedia and its Applications*, vol. 6, no. 5, 2014. [Article (CrossRef Link)](#)

[17] T. Kim, T.H. Park, "Extended Kalman Filter (EKF) Design for Vehicle Position Tracking Using Reliability Function of Radar and Lidar," *Sensors*, vol. 20, pp. 4126, 2020. [Article (CrossRef Link)](#)

[18] E. Foxlin, L. Naimak, "VIS-Tracker: A Wearable Vision-inertial Self-Tracker," in *Proc. of IEEE VR2003*, vol. 1, pp. 199, 2003. Article (CrossRef Link)

[19] H. Rehbinder, B. K. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 186-199, 2003. Article (CrossRef Link)

[20] G. Qian, R. Chellappa, Qinfen Zheng, "Bayesian structure from motion using inertial information," in *Proc. of International Conference on Image Processing*, 2002. Article (CrossRef Link)

[21] F. M. Mirzaei, S. I. Roumeliotis, "A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1143-1156, 2008. Article (CrossRef Link)

[22] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[23] S. Jung, Y. Cho, K. Lee, M. Chang, "Moving Object Detection with Single Moving Camera and IMU Sensor using Mask R-CNN Instance Image Segmentation," *International Journal of Precision Engineering and Manufacturing*, vol. 22, pp. 1049-1059, 2021. Article (CrossRef Link)

[24] H. Deilamsalehy, T. C. Havens, "Sensor fused three-dimensional localization using IMU, camera and LiDAR," *IEEE Sensors*, pp. 1-3, 2016. Article (CrossRef Link)

[25] E.D. Marti, D. Martin, J. Garcia, A. De la Escalera, J.M. Molina, J.M. Armingol, "Context-Aided Sensor Fusion for Enhanced Urban Navigation," *Sensors*, vol. 12, no. 12, pp. 16802-16837, 2012. Article (CrossRef Link)

[26] C. Yu, H. Lan, F. Gu, F. Yu, N. El-Sheimy, "A Map/INS/Wi-Fi Integrated System for Indoor Location-based Service Applications," *Sensors*, vol. 17, pp. 1272, 2017. Article (CrossRef Link)

[27] X. Li, Y. Wang, K. Khoshelham, "A Robust and Adaptive Complementary Kalman Filter based on Mahalanobis Distance for Ultra Wideband/Inertial Measurement Unit Fusion Positioning," *Sensors*, vol. 18, pp. 3435, 2018. Article (CrossRef Link)

[28] Y. Wu, X. Niu, J. Du, L. Chang, H. Tang, H. Zhang, "Artificial Marker and MEMS IMU-based Pose Estimation Method to Meet Multirotor UAV Landing Requirements," *Sensors*, vol. 19, no. 24, pp. 5428, 2019. Article (CrossRef Link)

Sukwoo Jung received the B.E., M.E., and Ph.D. degrees from the Department of Mechanical Engineering of Korea University, Seoul, South Korea, in 2013, 2015, and 2020, respectively. Since 2020, he has been working at Korea Electronics Technology Institute (KETI), Seoul, South Korea, where he is presently the Senior Researcher of Contents Convergence Research Center. His current research interests include image recognition, pose estimation, 3D reconstruction, and moving object detection.

Seho Park is a Principal Researcher at Korea Electronics Technology Institute (KETI), Seoul, Korea. Before joining KETI in 2005, he was with I&C Technology and Samsung Electronics, where he worked in projects relevant to SoC (System on Chip) for digital broadcasting and wireless networks. He received his B.S. and M.S. degrees in Electrical Engineering from Kyungpook National University in 1998 and 2000, respectively. His current research interests are in the areas of 2D/3D image processing, pattern recognition and AR/MR Interaction technology.

KyungTaek Lee received the B.S. and M.S. degrees in the Microelectronics and Device Engineering from Inha University, Inchen, Korea, in 1994 and 1996, respectively. He received Ph.D. degree in the School of Electrical & Electronic Engineering from Yonsei University, Seoul, Korea, in 2008. Since 2002, he has been working at Korea Electronics Technology Institute (KETI), Seoul, Korea, where he is presently the director of Contents Convergence Research Center. He was with the Human-Computer Interaction Institute(HCII) at Carnegie Mellon University as a visiting researcher in 2015 and 2016. His current research interests are in the areas of XR(eXtended Reality) and Artificial Intelligence(AI).