

# Feature coordinate estimation by time differenced vision/GPS/INS

Yu dam Lee<sup>1</sup>, Kwang Ho Choi<sup>1</sup>, Won Jae Yoo<sup>1</sup>, La Woo Kim<sup>1</sup>, and Hyung Keun Lee<sup>2,\*</sup>

<sup>1</sup>Department of Electronics and Information Engineering, Korea Aerospace University, 10540 Gyeonggi-do, South Korea

<sup>2</sup>School of Electronics, Telecommunication and Computer Engineering, Korea Aerospace University, 10540 Gyeonggi-do, South Korea

**Abstract.** This paper proposes a multi-baseline method to estimate absolute coordinates of point clouds and the camera attitude parameter utilizing feature points in successive images. Conventionally, 3D map generation methodologies have been based on images acquired by aerial or land vehicles. Pixels corresponding to known landmarks are manually identified at first. Next, the coordinates are directly geo-referenced and automatically allocated to pixels with high-quality Global Positioning System (GPS) and Inertial Navigation System (INS). However, it is difficult to obtain accurate coordinates by the conventional methodology with low-cost GPS and INS. With camera positions and attitude parameters, image-based point clouds can be compensated accurately. A simulation was carried out to evaluate the performance of the proposed method.

## 1 Introduction

3D map generation methodologies have been studied for decades. They are generally based on images acquired by aerial or land vehicles. In the conventional methods, pixels corresponding to exact locations of landmarks appeared in images are identified manually to allocate them. Their coordinates can be directly geo-referenced and automatically processed with high-quality Global Positioning System (GPS) and Inertial Navigation System (INS). However, due to insufficient accuracy, it is hard to obtain their accurate coordinates [1, 2, 3, 5].

This paper proposes a method to estimate three-dimensional (3-D) coordinates of point clouds and attitude parameters simultaneously by two-dimensional (2-D) matched feature points in two successive images. To extract feature points and to match them in the two successive images, Speeded-Up Robust Feature (SURF) algorithm is utilized [4]. A loosely-coupled Kalman filter is applied to estimate absolute coordinates of the camera and its attitude parameters. With these parameters and error of them, image-based point clouds are compensated and converted to absolute coordinates of them.

Moreover, comparing to the conventional stereo method with static baseline, the proposed method can make it to time differenced multi-baseline; this makes depth estimation more accurate.

This paper is organized as follows. Firstly, overall system flow is explained. Then the proposed method to estimate absolute coordinates of point clouds and to compensate error of attitude parameters is briefly explained by pictures. Finally, a simulation result is analysed.

## 2 Absolute coordinates of point clouds with attitude error compensation

Fig. 1 shows the diagram of the proposed system. In the proposed system, error of attitude parameters from GPS/INS Kalman outputs can be compensated by 2-D feature points extracted from the images.

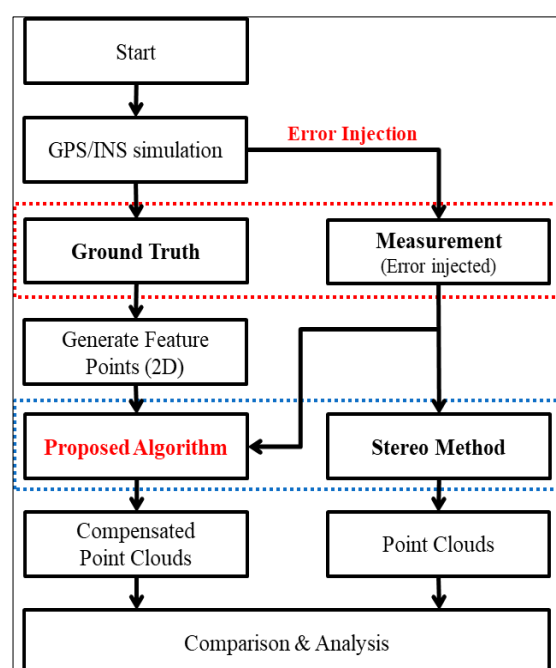
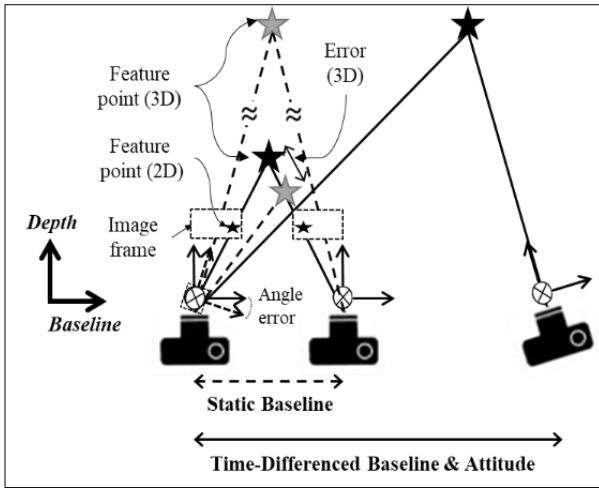
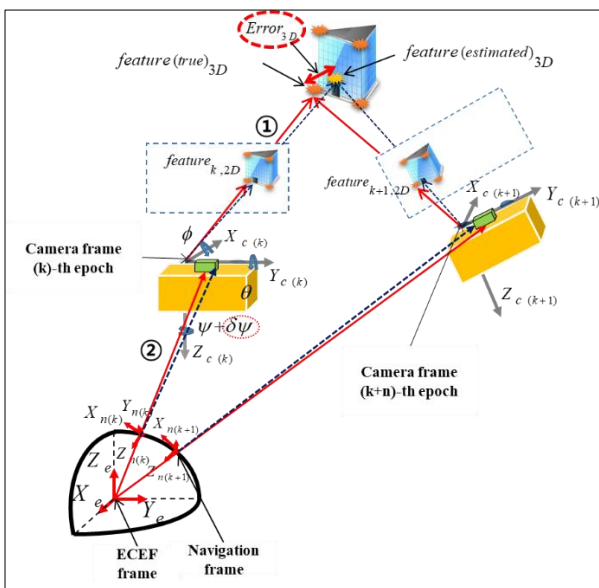


Fig. 1. Diagram of the proposed system.

\* Corresponding author: [hyknlee@kau.kr](mailto:hyknlee@kau.kr)



**Fig. 2.** Comparison between the conventional stereo method and the proposed method.



**Fig. 3.** Geometric relationship between different epochs in the proposed method.

In Fig. 2, the conventional and proposed methods are compared. In the figure, star symbols indicate feature points in 2-D and 3-D. The grey stars indicate 3-D feature points with errors. The direction to them are depicted in dotted lines. The estimation errors are partly due to the limitation of the stereo method for distant depth estimation and partly due to the insufficient accuracy when utilizing low-cost GPS and Inertial Measurement Unit (IMU) modules. On the other hand, the black stars indicate 3-D feature with reduced errors. The direction to them are depicted in solid lines.

The proposed method utilizes GPS and INS not only to overcome its limitation on static baseline but also to be applicable for different attitudes between cameras. Since the camera motion is directly geo-referenced at each epoch with GPS/INS, various baselines and attitude parameters can be calculated time-differentially.

Fig. 3 shows the geometric relationship between feature points extracted at different epochs. In the figure,

yaw angle is only considered for the compensation due to the fact that roll and pitch angles of GPS/INS Kalman outputs are fairly accurate enough to utilize.

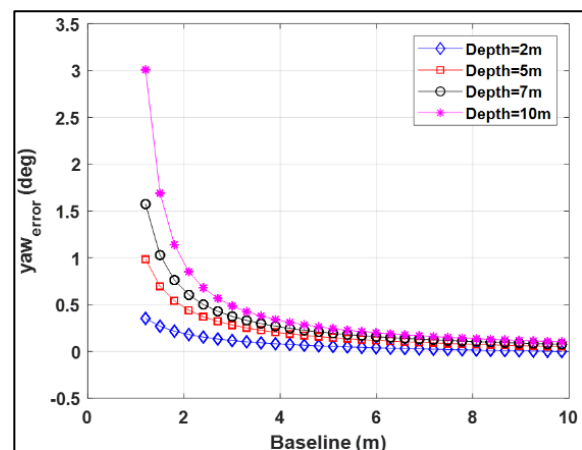
In Fig. 3, the line ① can be considered by the translation and transformation matrix from the Earth-Centered Earth-Fixed (ECEF) frame to the camera frame. The line ② can be considered by the relative coordinates of point clouds in the camera frame. If it is assumed that all of modules are aligned, the relative coordinates are converted to the absolute coordinates [5].

### 3 Simulation

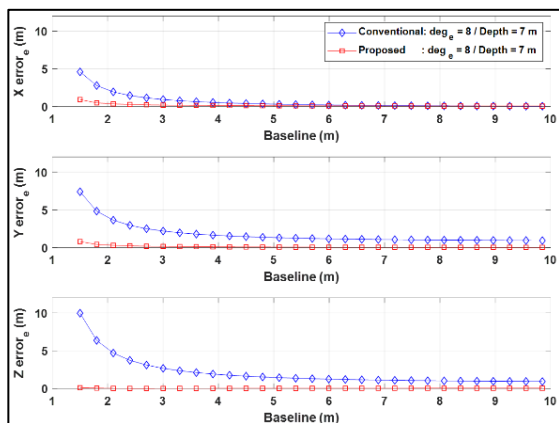
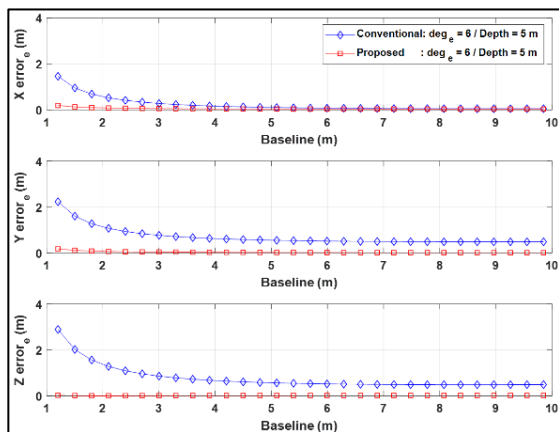
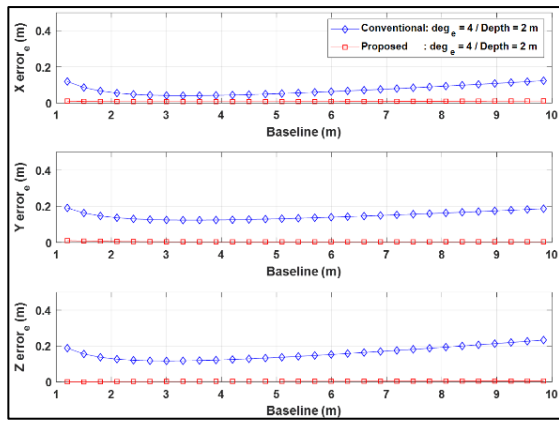
To evaluate the accuracy of the proposed method, a systematic simulation was carried out. For the system inputs, no error was added to provide the ground truth data. Then, errors were added to the measurements for simulation. 2-D feature points were generated with the ground truth data only. The 2-D feature points were generated for even distribution. Depths for the system input were set as 2, 5, 7, 10 m. To evaluate how accurate yaw angles can be compensated, large initial yaw errors were injected in measurements at the  $(k)$ -th and  $(k+n)$ -th epochs, respectively ( $n=1, 2, 3, \dots$ ). The effect of time-differenced baselines was also considered in this simulation. The main purpose of this simulation is to analyse the relationship between depth, length of baseline and injected yaw error.

Fig. 4 shows the relationship between the compensated yaw error, the length of baseline and the depth accuracy. In general, the longer the baselines are, the smaller the yaw errors are. It can be seen that, though the initial yaw error was set as 4 degree when length of baseline is more than 2 m, it was compensated below 1 degree by the proposed method.

In Fig. 5, the errors of the point clouds by the two methods are compared. It can be seen that errors increase when depth and injected yaw error increase. However, the errors are significantly larger by the conventional method than by the proposed method when the length of baseline is short. They are relatively small around cm-level in the proposed method.



**Fig. 4.** Characteristics of compensated yaw error with respect to time-differenced baselines and various depths (Injected yaw error: 4 degree).



**Fig. 5.** Comparison of absolute coordinate errors between the conventional method and the proposed method. Injected yaw errors are (4/6/8) degrees and depth setting: (2/5/7) m.

In summary, the simulation results show that the proposed method improves accuracy compared with the conventional stereo method. In more detail, the depth accuracy is similar by both methods if the length of baseline is long. However, the accuracy in Y-axis and Z-axis are considerably larger by the conventional method than the proposed method.

## 4 Conclusion

This paper proposed a multi-baseline method to estimate absolute coordinates of point clouds and the camera

attitude parameters simultaneously by 2-D matched feature points in two successive images. A simulation was carried out to evaluate its performance compared with the conventional stereo method. By the simulation, it was confirmed that the proposed method properly compensates important yaw error and improves the accuracy of the absolute coordinates of point clouds.

This research was supported by Basic Science Research Program (2016R1D1A1B01009881) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology, Korea.

## References

1. G. Zhou, C. Song, J. Simmers, P. Cheong, "Urban 3D GIS from LiDAR and digital aerial images," *C & G*, **30**, 345-353 (2004)
2. Y. Song, S. Nuske, S. Scherer, "A multi-sensor fusion MAV state estimation from long-range stereo, IMU, GPS and barometric sensors," *Sensors*, **17**, 1-26 (2017)
3. Y. B. Park, J. G. Lee, C. G. Park, "Performance analysis of INS/GPS integration system," *KIEE*, 2433-2435 (2000)
4. K. Mikołajczyk, C. Schmid, "A performance evaluation of local descriptors," *IEEE transactions on pattern analysis and machine intelligence*, **27**, 1615-1630 (2005)
5. Y. D. Lee, K. H. Choi, J. H. Lim, W. J. Yoo, L. W. Kim, H. K. Lee, "A System for Generating Spatial Information utilizing RTK GPS/INS/Stereo Camera for 3D Map-Based Hybrid Navigation," *JICRS*, **28**, 757-766 (2018)