3D MOBILE MAPPING WITH A LOW COST UAV SYSTEM

J. Li [a, *], B. Yang [a, *], W. Wu [a], W. Dai [a], C. Chen [a], X. Zou [a], M. Tian[a]

[a] State Key Laboratory of Information Engineering in Survey, Mapping and Remote Sensing,
Wuhan University, No. 129, Luoyu Road, Wuhan, PR China -
(lijianping, bshyang)@whu.edu.cn

Commission II

KEY WORDS: imagery data, Inertial Measurement Unit, low cost UAV, mobile mapping, multi-sensor

ABSTRACT:

In order to accomplish the automatic mobile mapping task in a small area of interest, a low cost UAV system is proposed in this paper. Multiple sensors including a global shutter camera and an inertial measurement unit are calibrated and synchronized to collect data from the area of interest. First the images are matched by the chronological order and the SfM is utilized. Then the origin SfM result is integrated with the IMU data by adding the IMU constraints into the bundle adjustment. At last the photogrammetry point clouds are generated using PMVS according to the extrinsic parameters. Experiments are undertaken in a typical scene with photogrammetry point clouds generated. The trajectory estimated by the proposed integration method are compared with the method that relies on image only, showing that the proposed method has better performance.

1. INTRODUCTION

Using point clouds of a cultural domain or urban building has become the focus for different purposes, such as cultural relic protection, emergency & disaster relief and building of smart city (Adams and Friedland, 2011; Reiss et al., 2016; Yang et al., 2015). It is a common technique to capture images using an Unmanned Aerial Vehicles (UAV) system and generate point clouds for the region of interest. The price of a set of mobile mapping system based on a UAV is usually more than 500,000 dollars, which is expensive for the civilian users. In particular, for some private sectors and local government agencies, they have a strong demand for a mobile mapping system based on a UAV that has low cost, miniaturization and medium performance for collecting data from a small area of interest. Although the common commercial mobile mapping systems mounted on a UAV is powerful, but they can hardly meet these demands. As a result, the low-cost and small mobile mapping systems mounted on a UAV with medium performance attract a large number of researchers(Wallace et al., 2012; Yang and Chen, 2015).

Usually the mobile mapping systems are equipped with the integrated navigation system including a GNSS receivers (GPS) and an Inertial Measurement Unit (IMU), so their state (location, orientation, speed, etc.) can be estimated by fusing the GPS and the IMU data directly. However, a differential GPS system also increases the cost of the mobile mapping system and the difficulties in operation. Furthermore, in many scenarios, it would be difficult to ensure that the received GPS signals are reliable and accurate. For example, the GPS signals would be poor when the UAV system is operating between tall buildings. Therefore, it is essential to develop a solution that does not rely on the GPS information for the mobile mapping system, it would increase the application scope and reduce the cost efficiently.

In this paper, a strategy for the low-cost UAV based mobile mapping is developed by integration of imagery and IMU data. Image based matching and loop-closure detection can limit the error accumulation of the system effectively, while the accuracy of pose estimated by Structure from Motion (SfM) depends on the 3D feature points. Only when the baseline length is greater than a certain range, the 3D feature points can be obtained through triangulation accurately. Therefore, images are not suitable for the estimation of the extrinsic parameters in short time duration. Moreover, the absolute scale cannot be derived from the imagery data. In a short time duration, precise poses as well as absolute scale can be derived from IMU data efficiently, while the errors accumulate seriously after long time. So the imagery and IMU data are effectively complementary.

The remainder of this paper is organized as follows. The system description is detailed in Section 2. The proposed data integration method is elaborated in Section 3. In Section 4, the experimental studies are undertaken to evaluate the proposed method, after which conclusions are drawn at the end.
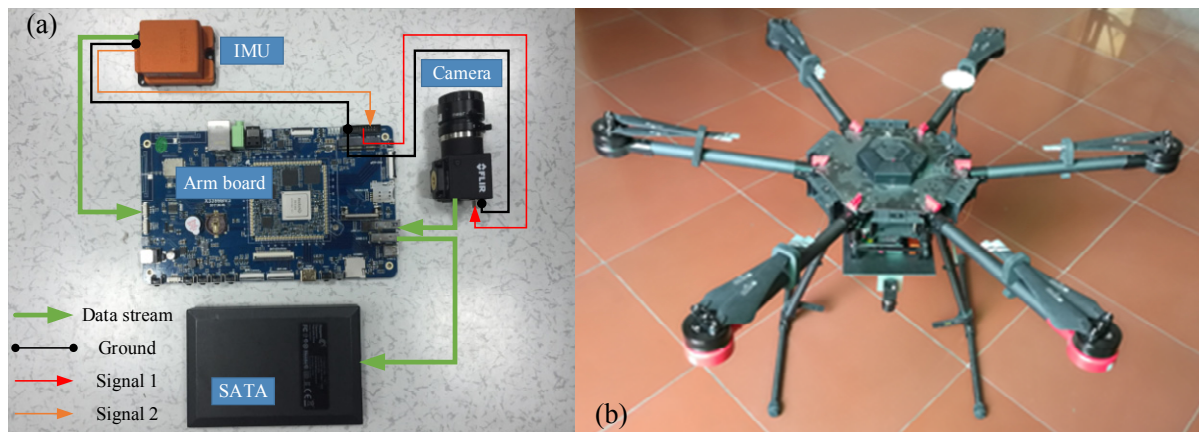
---

* Corresponding author

Figure 1. System description. (a)The green line represents the data stream. The black line is the ground wire. The red and orange wires represent the synchronization signals. (b) Our UAV platform.

2. SYSTEM DESCRIPTION

2.1 Sensors and UAV platform

Our system is consisting of two kind of measurement instruments including an IMU and a global shutter camera. The modularity of the system shows that new sensors can be easily added to the system. A typical configuration, including the synchronization signals is shown in Fig.1 (a).The black lines are the ground lines. The red line represents the synchronization signal send from arm board to the global shutter camera to trigger the camera. The orange wire represents the synchronization signal send from IMU to the arm board so the arm board can record the accurate time stamp once the IMU message comes. Table.1 is sensor specifications.

The IMU used in the system is Xsens MTI-300, which provides the measurements of the acceleration and the angular velocity. The noise density of the gyroscopes is 0.01 º/s/√Hz, and the in-run bias stability of the gyroscopes is 10 º/h. The noise density of the accelerometers is 60ug/√Hz, and the in-run bias stability of the accelerometers is 15ug. This IMU is a low-cost MEMS-IMU products can meet our needs.

The global shutter camera used in the system is Pointgrey Flea3. It has a 1280 × 1024 pixel sensor with the pixel size of 5.3 µm. It can collect at most 40 images per second and has no rolling shutter effect so it can be used to estimate motion of system. The wider the view of the camera is, the more reliable features can be detected in the image in the SfM algorithm. So the lens used in the system is a wide-angle lens (3.5mm/F1.4) made by Kowa.

The arm-board control unit contains a Cortex-A9 multi-core processor and 2 GB memory, which is the processing instrument of the system. The weight and power consumption of the instruments are severely limited in a UAV system. The arm-board is suitable for the UAV platform because it can work in the absence of a radiator and its power consumption is very low.

The UAV platform used in the system is the DJI m600 as shown in Fig.1 (b) with max takeoff weight of 15.1kg and self- weight 9.1kg which is enough to bring up all the sensors and the arm-board. With no payload, the hovering time of DJI m600 is about 35 min. With 6 kg payload, the hovering time of DJI m600 is about 16 min. So the DJI m600 is suitable for mobile mapping in a small area of interest.

Table 1. Sensor specifications.

Sensor	Description
IMU	Xsens MTI-300
Global shutter camera	Pointgrey Flea3
Lens	Kowa wide-angle lens (3.5mm/F1.4)
Onboard control unit	Cortex-A9 multi-core processor
UAV platform	DJI m600

2.2 Synchronization and calibration

The synchronization of the different sensors is fulfilled electronically. The two signals used for synchronization are the 200Hz signal from IMU and 40 Hz signal generated by arm-board. These signals were selected because the IMU cannot be synchronized accurately with an external signal while the camera can be triggered by an external signal easily. Whenever the IMU collects a data, it sends a pulse signal to the arm board. Then the arm board generate a timestamp when it receive the pulse signal. It can be assumed that there is no time delay during the pulse signal transmission. The arm-board will send a signal to the camera after receiving 5 signals from the IMU. When the system is collecting data, the arm board save all the data in the SATA disk. When post-processing, the data from different sensors is combined using the timestamp information recorded by the arm-board.

The intrinsic parameters of the global shutter camera are calibrated using the Matlab calibration toolbox(Bouguet, 2008). The calibration tool for Xsens IMU is offered by the MT-manager program. The bore-sight calibration of system is carried out using the method proposed by Furgale et al. (2013). Before collecting data, the intrinsic parameters of the global shutter camera, intrinsic parameters of the IMU and the bore-sight between IMU and camera are calibrated manually.

3. INTEGRATION OF IMAGERY AND INERTIAL DATA

We follow the notation introduced by (Barfoot, 2016) through this work. The definition of coordinate frames of the system is shown in Fig.2. $\overrightarrow{F_w}$ represents the world frame; $\overrightarrow{F_s}$ represents the inertial frame; $\overrightarrow{F_c}$ represents the image frame.

The variables to be estimated are the states $\{\mathbf{x}_s^k | k = 1,2 \ldots N\}$ at each time the kth image is captured and the feature points $\{\mathbf{x}_f^j | j = 1,2 \ldots M\}$. The state $\mathbf{x}_s^k$ is written as:

$$\mathbf{x}_s^k = \left[\boldsymbol{r}_w^{ws_k T}, \boldsymbol{v}_w^{k T}, \boldsymbol{q}_{ws_k}^T, \boldsymbol{b}_a^{k T}, \boldsymbol{b}_g^{k T} \right]^T$$
$$\in \mathrm{R}^6 \times \mathrm{SO}^3 \times \mathrm{R}^6 \qquad (1)$$

where SO^3 is the special orthogonal group. $\boldsymbol{r}_w^{ws_k}$ and $\boldsymbol{q}_{ws_k}$ are the position and the orientation at time of index k, which are represented in world frame. $\boldsymbol{v}_w^k$ is the velocity represented in world frame. $\boldsymbol{b}_a^k$ and $\boldsymbol{b}_g^k$ are the biases of accelerometers and gyroscopes.

All the variables will be estimated and optimized as one joint optimization through a cost function $J(\boldsymbol{x})$ including back-projection error $\mathbf{e}_r$ and IMU error term $\mathbf{e}_s$:

$$J(\boldsymbol{x}) \coloneqq \sum_{k=1}^{N} \sum_{j \in \mathcal{J}(k)} \mathbf{e}_r^{k,j^T} \mathbf{W}_r^{k,j} \mathbf{e}_r^{k,j} +$$
$$\sum_{k=1}^{N-1} \mathbf{e}_s^{k^T} \mathbf{W}_s^k \mathbf{e}_s^k \qquad (2)$$

where $\mathcal{J}(k)$ is the visibility set of the kth image. $\mathbf{W}_r^{k,j}$ and $\mathbf{W}_s^k$ are the information matrices. This non-linear optimization problem is optimized by the open source graph based optimization library G2O(Grisetti et al., 2011). The information matrices and error definitions are detailed as follow.
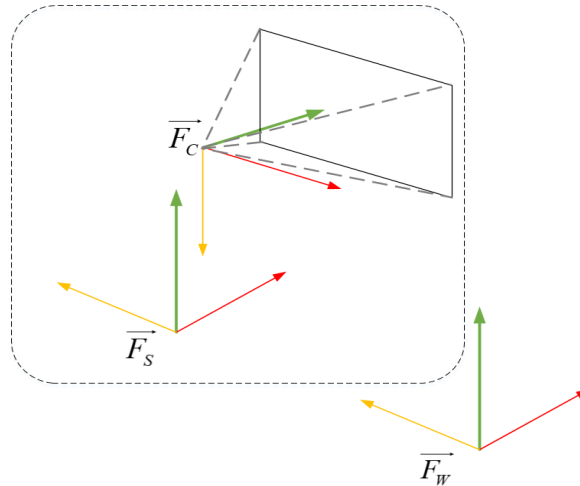

Figure 2. Coordinate frames involved in the system. Image data is acquired in $\overrightarrow{F_c}$, and IMU data is acquired in $\overrightarrow{F_s}$. The poses and feature points need to be estimated are represented in $\overrightarrow{F_w}$

The adjacent images are matched by the chronological order and the key frames are selected according to the parallax of the adjacent images, thereby increasing efficiency of bundle adjustment. Loop closures are detected by utilizing Bag of Words (BoWs)(Philbin et al., 2007) for suppressing cumulative error. The incremental SfM proposed by Schonberger and Frahm (2016) are to estimate the initial value of the scale free coordinates of feature points, meanwhile the false-matches are removed too.

Back-projection error is defined as follow:

$$\mathbf{e}_r^{k,j} = \mathbf{z}_r^k - \lambda K \left(\boldsymbol{C}_{cs} \left(\boldsymbol{C}_{ws_k}{}^T \mathbf{x}_f^j - \boldsymbol{C}_{ws_k}{}^T \boldsymbol{r}_w^{ws_k} \right) + \boldsymbol{r}_c^{cs} \right) \qquad (3)$$

where $\mathbf{z}_r^k$ is the measurement of feature point whose distortions are removed by using the calibrated camera distortion parameters. K is the calibrated projection matrix of the camera. $\boldsymbol{C}_{cs}$ and $\boldsymbol{r}_c^{cs}$ are the calibrated relative position and orientation between camera and IMU. information matrix $\mathbf{W}_r^{k,j}$ can be obtained as follow:

$$\mathbf{W}_r^{k,j} = \begin{bmatrix} \frac{1}{\delta x^2} & 0 \\ 0 & \frac{1}{\delta y^2} \end{bmatrix} \qquad (4)$$

where δx and δy are the error of feature points in image coordinates. In our experience, the error of feature points in image can be set to 2 pixels in each direction.

The IMU kinematics model used is the simplified model for low-cost MEMS-IMUs proposed by Shin and El-Sheimy (2004):

$$\dot{\boldsymbol{r}}_w^{ws} = \boldsymbol{v}_w \qquad (5)$$

$$\dot{\boldsymbol{v}}_w = \boldsymbol{C}_{ws}(\boldsymbol{a}_m - \boldsymbol{b}_a - \boldsymbol{a}_n) + \boldsymbol{g} \qquad (6)$$

$$\dot{\boldsymbol{q}}_{ws} = \frac{1}{2} \boldsymbol{q}_{ws} \otimes (\boldsymbol{\omega}_m - \boldsymbol{b}_g - \boldsymbol{\omega}_n) \qquad (7)$$

$$\dot{\boldsymbol{b}}_a = \boldsymbol{b}_a^n \qquad (8)$$

$$\dot{\boldsymbol{b}}_g = \boldsymbol{b}_g^n \qquad (9)$$

where $\boldsymbol{a}_m$ and $\boldsymbol{\omega}_m$ are the measurements from accelerometers and gyroscopes with White Gaussian Noise (WGN) $\boldsymbol{a}_n$ and $\boldsymbol{\omega}_n$. $\boldsymbol{g}$ is the gravity vector. The biases are modeled as random walk process with WGN $\boldsymbol{b}_a^n$ and $\boldsymbol{b}_g^n$. Therefore the error term of IMU can be find:

$$\mathbf{e}_s^k(\hat{\mathbf{x}}_s^k, \mathbf{x}_s^{k+1}, \mathbf{z}_s^k) = \begin{bmatrix} \hat{\boldsymbol{r}}_w^{ws_{k+1}} - \boldsymbol{r}_w^{ws_{k+1}} \\ 2[\hat{\boldsymbol{q}}_{ws_{k+1}} \otimes \boldsymbol{q}_{ws_{k+1}}^{-1}]_{1:3} \\ \hat{\boldsymbol{b}}_a^{k+1} - \boldsymbol{b}_a^{k+1} \\ \hat{\boldsymbol{b}}_g^{k+1} - \boldsymbol{b}_g^{k+1} \end{bmatrix} \qquad (10)$$

where $\mathbf{z}_s^k$ is the IMU measurement.

The error state kinematics can be obtained as follow:

$$\delta\dot{\mathbf{x}}_s = \boldsymbol{F}_s(\hat{\mathbf{x}}_s)\delta\mathbf{x}_s + \boldsymbol{F}_n(\hat{\mathbf{x}}_s)\boldsymbol{n} \tag{11}$$

$$\boldsymbol{F}_s(\hat{\mathbf{x}}_s) = \frac{\partial f}{\partial \delta\mathbf{x}_s} = \begin{bmatrix} \mathbf{0}_{3\times3} & \delta\boldsymbol{v}_w & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & -\widehat{\boldsymbol{C}}_{ws}[\boldsymbol{a}_m - \widehat{\boldsymbol{b}}_a]_\times & -\widehat{\boldsymbol{C}}_{ws} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -[\boldsymbol{\omega}_m - \widehat{\boldsymbol{b}}_g]_\times\delta\boldsymbol{\theta} & \mathbf{0}_{3\times3} & -\boldsymbol{I}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix} \tag{12}$$

$$\boldsymbol{n} := [\boldsymbol{a}_n{}^T, \boldsymbol{\omega}_n{}^T, \boldsymbol{b}_a^{n\,T}, \boldsymbol{b}_g^{n\,T}]^T \tag{13}$$

$$\boldsymbol{F}_n(\hat{\mathbf{x}}_s) = \frac{\partial f}{\partial \boldsymbol{n}} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ -\widehat{\boldsymbol{C}}_{ws} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & -\boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} \end{bmatrix} \tag{14}$$

The information matrix $\mathbf{W}_s^k$ of the IMU error term can be propagated according to the error state kinematics directly as follow:

$$\mathbf{W}_s^k = \left(\frac{\partial \mathbf{e}_s^k}{\partial \mathbf{x}_s^{k+1}}(\mathbf{W}_s^{k-1})^{-1}\left(\frac{\partial \mathbf{e}_s^k}{\partial \mathbf{x}_s^{k+1}}\right)^T\right)^{-1} \tag{15}$$

$$\frac{\partial \mathbf{e}_s^k}{\partial \mathbf{x}_s^{k+1}} = \begin{bmatrix} \boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times9} \\ \mathbf{0}_{3\times3} & [\widehat{\boldsymbol{q}}_{ws_{k+1}} \otimes \boldsymbol{q}_{ws_{k+1}}^{-1}]_{1:3,1:3}^{\oplus} & \mathbf{0}_{3\times9} \\ \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \boldsymbol{I}_{9\times9} \end{bmatrix} \tag{16}$$

The extrinsic parameters and sparse landmarks can be obtained after solving the above optimization problem, then we use the PMVS proposed by Furukawa and Ponce (2010) to generate dense photogrammetry point clouds.

4. EXPERIMENTS AND ANALYSIS

In order to verify the effectiveness of the platform, we conducted experiments in Wuhan University. We calibrate the intrinsic parameters of the global shutter camera, IMU and bore-sight between the IMU and camera first. Then it took about ten minutes to collect data for a teaching building using the integrated UAV system. As shown in Fig.3, we are operating our low-cost UAV system.

The state variables and extrinsic parameters are estimated by utilizing the proposed method, then the dense photogrammetry point clouds are generated by utilizing the PMVS algorithm as shown in Fig.4. In Fig.4, the yellow line is the estimated trajectory of the UAV system and the color point clouds are the dense photogrammetry point clouds.

To compare the proposed integration with SfM only using imagery data, results of the proposed method and results of the SfM only using imagery data are shown in Fig.5. The two trajectories are registered manually using the open source software CloudCompare (Girardeau-Montaut, 2015). The yellow line is the trajectory estimated by the proposed method

and the red line is estimated by imagery data only. From the white box in Fig.5, we can find that the estimated trajectory only using imagery data has more noise than the proposed method. Because the extrinsic parameters estimated only using images are not reliable in short time duration caused by short baseline. While the propose integration method has higher robustness.

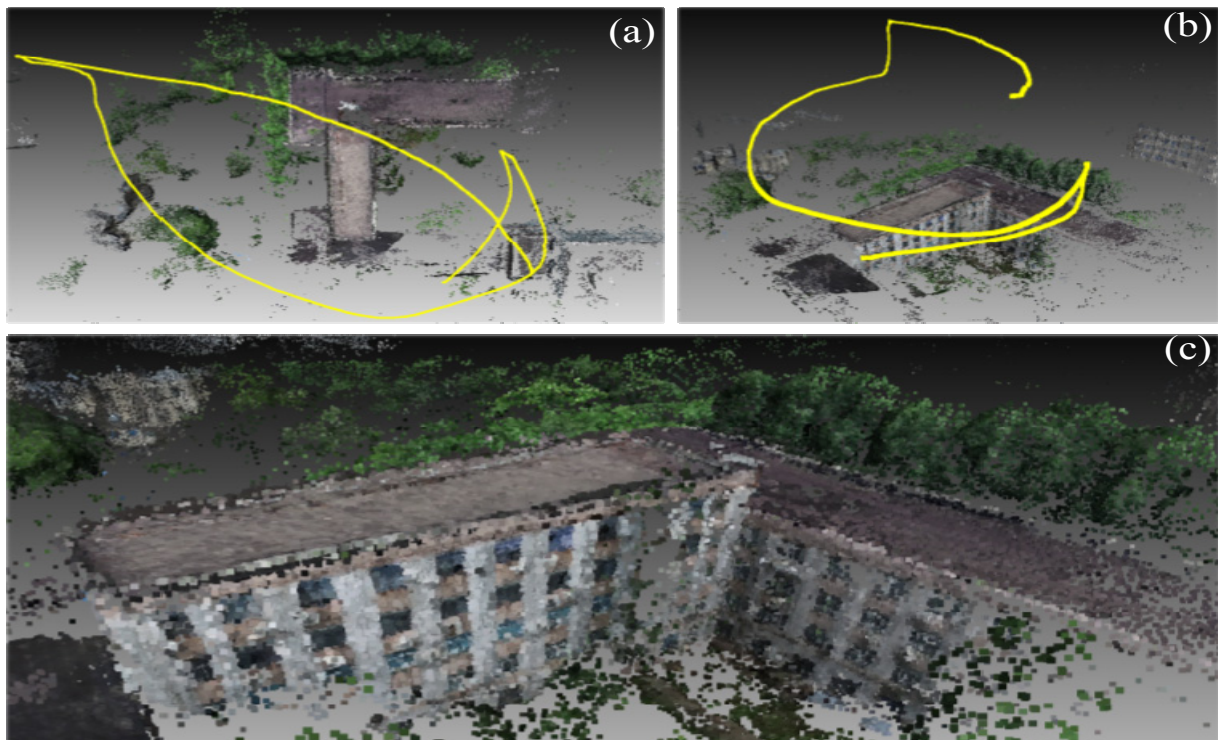Figure 3. The proposed low-cost mobile mapping system is operating under control.

Figure 4. Result of the proposed method for mapping of our teaching building in Wuhan University. The yellow line is the estimated trajectory of the UAV system. (a) and (b) show the top and the side view of result. (c) shows the details of the generated point cloud.
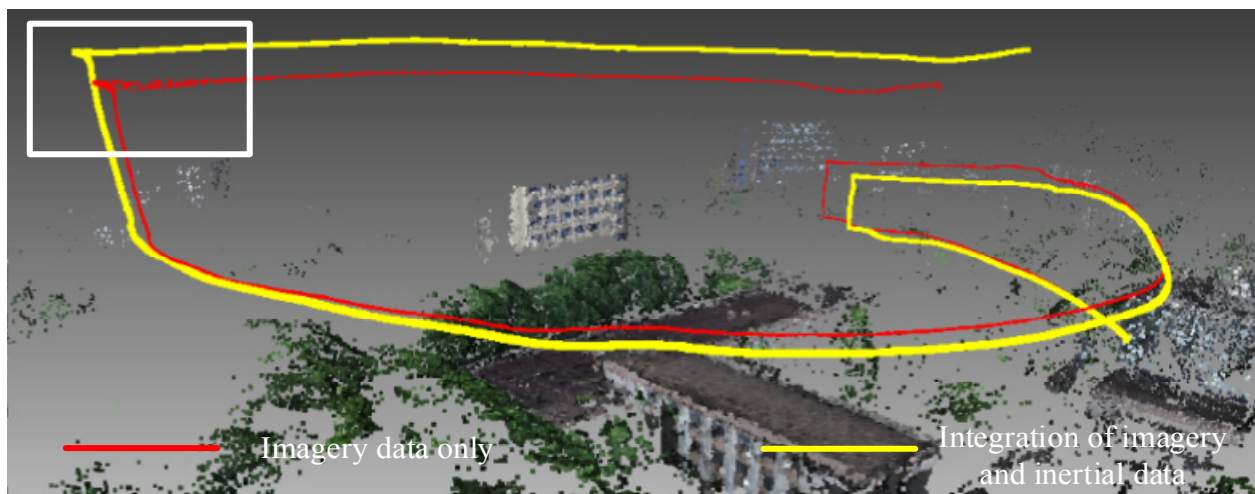

Figure 5. Comparison of the proposed method and SfM only using imagery data. Yellow line is the trajectory estimated by the proposed method and red line is estimated by imagery data only. The two trajectories are registered manually.

5. CONCLUSION

For the purpose of mobile mapping using low-cost UAV system, a low-cost UAV system integrating two kind of sensors are investigated in this paper. To estimate accurate extrinsic parameters, an integration method fusing the imagery data and the IMU is proposed in this paper. First the images are matched by the chronological order and the SfM is utilized. Then the origin SfM result is integrated with the IMU data by adding the IMU constraints into the bundle adjustment. The experimental results show that the designed low-cost UAV system can be used for mapping small area of interest by integrating imagery and inertial data effectively. Moreover, imagery and IMU data are complementary, so compared with SfM only using imagery data, more reliable extrinsic parameters and the 3D landmarks can be estimated by integrating IMU and imagery data.

7. REFERENCE

Adams, S.M.and Friedland, C.J., 2011. A survey of unmanned aerial vehicle (UAV) usage for imagery collection in disaster research and management. *9th International Workshop on Remote Sensing for Disaster Response*, p. 8.

Barfoot, T., 2016. State Estimation for Robotics: A Matrix Lie Group Approach. Draft in preparation for publication by Cambridge University Press, Cambridge.

Bouguet, J., 2008. Matlab calibration toolbox. URL: http://www.vision. caltech. edu/bouguetj/calib_doc.

Furgale, P., Rehder, J. and Siegwart, R., 2013. Unified temporal and spatial calibration for multi-sensor systems. *IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS), pp. 1280-1286.

Furukawa, Y. and Ponce, J., 2010. Accurate, dense, and robust multiview stereopsis. *IEEE transactions on pattern analysis and machine intelligence* 32, pp. 1362-1376.

Girardeau-Montaut, D., 2015. Cloud Compare—3D Point Cloud and Mesh Processing Software. Open Source Project.

Grisetti, G., Kümmerle, R., Strasdat, H. and Konolige, K., 2011. g2o: A general Framework for (Hyper) Graph Optimization. Technical report.

Philbin, J., Chum, O., Isard, M., Sivic, J. and Zisserman, A., 2007. Object retrieval with large vocabularies and fast spatial matching. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8.

Reiss, M., da Rocha, R., Ferraz, R., Cruz, V., Morador, L., Yamawaki, M., Rodrigues, E., Cole, J. and Mezzomo, W., 2016. Data Integration Acquired from Micro-Uav and Terrestrial Laser Scanner for the 3d Mapping of Jesuit Ruins of SÃO Miguel das MISSÕES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 315-321.

Schonberger, J.L. and Frahm, J.-M., 2016. Structure-from-motion revisited. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4104-4113.

Shin, E.-H. and El-Sheimy, N., 2004. An unscented Kalman filter for in-motion alignment of low-cost IMUs, Position Location and Navigation Symposium, 2004. PLANS 2004. IEEE, pp. 273-279.

Wallace, L., Lucieer, A., Watson, C. and Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. *Remote Sensing* 4, pp. 1519-1543.

Yang, B. and Chen, C., 2015. Automatic registration of UAV-borne sequent images and LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 101, pp. 262-274.

Yang, B., Dong, Z., Zhao, G. and Dai, W., 2015. Hierarchical extraction of urban objects from mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing* 99, pp. 45-57.