Rotational Alignment of IMU-camera Systems with 1-point RANSAC

Banglei Guan[1,2], Ang Su[1,2(✉)], Zhang Li[1,2], and Friedrich Fraundorfer[3]

[1] College of Aerospace Science and Engineering, National University of Defense Technology, China
[2] Hunan Provincial Key Laboratory of Image Measurement and Vision Navigation, China
[3] Institute for Computer Graphics and Vision, Graz University of Technology, Austria
suang@nudt.edu.cn

Abstract. In this paper we present a minimal solution for the rotational alignment of IMU-camera systems based on a homography formulation. The image correspondences between two views are related by homography when the motion of the camera can be effectively approximated as a pure rotation. By exploiting the rotational angles of the features obtained by *e.g.* the SIFT detector, we compute the rotational alignment of IMU-camera systems with only 1 feature correspondence. The novel minimal case solution allows us to cope with feature mismatches efficiently and robustly within a random sample consensus (RANSAC) scheme. Our method is evaluated on both synthetic and real scene data, demonstrating that our method is suited for the rotational alignment of IMU-camera systems.

Keywords: Rotational alignment · Minimal Solution · Pure rotation · IMU-camera calibration.

1 Introduction

The fusion of vision and IMU data have been applied to a wide variety of applications, such as structure from motion (SfM) [21] and simultaneous localization and mapping (SLAM) [14]. The accuracy of these applications highly depends on the axis alignment between the IMU and the camera coordinate system [5,6,21]. This paper investigates the problem of IMU-camera calibration. In particular, we are interested in the minimal case, i.e. to compute the rotational alignment of IMU-camera systems exploiting one point correspondence together with rotational angles obtained by, *e.g.* SIFT detector [17]. The novel minimal case solution is significant within a RANSAC scheme, to cope with the outliers of feature matches efficiently and robustly.

The IMU-camera calibration problem has already been addressed by various researchers. A class of approaches use IMU measurements directly and estimate the calibration parameters as part of visual-inertial sensor fusion by adopting a filter-based approach [13,18,24]. Due to the large number of DOFs, these approaches require a high camera frame rate. Considering that common IMUs output the complete rotation information with respect to the IMU reference coordinate system, IMU-camera calibration is generally regarded as hand-eye calibration regarding the IMU as the hand

[3,10,16,19,23]. Hence, the IMU-camera calibration problem can be represented as the hand-eye calibration equation $\mathbf{AX} = \mathbf{XB}$, where $\mathbf{X}$ is the transformation between the IMU coordinate system and the camera coordinate system which consists of a rotational matrix and a translational vector, $\mathbf{A}$ and $\mathbf{B}$ are the relative rigid motions of the camera and the IMU, respectively. Hand-eye calibration problem has already been addressed by many researchers in the past. The traditional methods need recover the camera poses in advance by using a calibration device or a SfM approach [10,25,27]. Recently, some methods avoid requiring prior knowledge of the camera poses and compute the hand-eye calibration directly from feature matches. Heller *et al.*[9] and Ruland *et al.*[20] employ the branch-and-bound algorithm to obtain globally optimal hand-eye calibration by minimizing the residuals in image space. Bender *et al.*[2] perform an in-flight IMU-camera calibration by exploiting a graph optimization framework.

A class of methods are proposed to perform IMU-camera calibration when the motion of the calibrated camera is a pure rotation or can be effectively approximated as a pure rotation. Seo *et al.*[22] solve the rotational matrix between the IMU coordinate system and the camera coordinate system by assuming all the translations to be zero. Hwangbo *et al.*[11] propose a calibration method based on homography transformation of image correspondences. Karpenko *et al.*[12] calibrate the camera and gyroscope system by quickly shaking the camera while pointing it at a far-away object. Guan *et al.*[7] propose minimal case solutions to the rotational alignment of IMU-camera systems using homography constraints, especially only 1.5 point correspondences are required for the pure rotation case. In fact, the assumption that the motion of the camera is a pure rotation is not restrictive in practical environment, because we rotate the camera outside where the scene is far away, the parallax-shift of most objects is hardly noticeable and the calibration method for a pure rotation case can be directly applied to such data. Thus IMU-camera calibration in the pure rotation case has practical relevance.

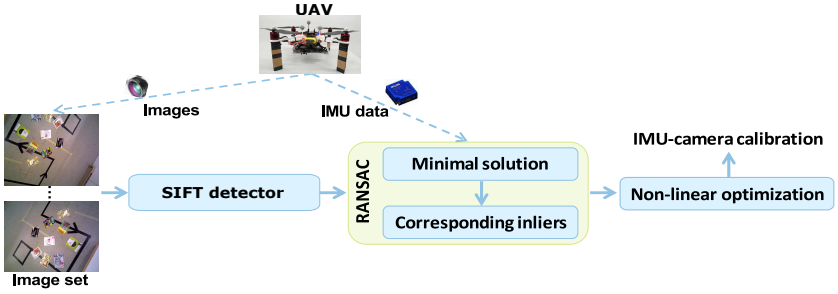


Fig. 1. Overview of the proposed IMU-camera calibration method.

The presented work is an extension of [7] which explores the different minimal case solutions to the rotational alignment of IMU-camera systems. Here, we extend [7] with a novel minimal case solution for the pure rotation case and achieve more accurate calibration result. Figure 1 illustrates the proposed IMU-camera calibration method. Our contributions can be summarized in the following way: (i) We propose to compute the rotational alignment of IMU-camera systems with 1 feature correspondence and the cor-

responding rotational angles obtained by *e.g.* the SIFT detector. (ii) Our method adopt the RANSAC [4] to cope with feature mismatches. The proposed minimal case solution is efficient within a RANSAC scheme, because the number of random samples that must be taken to find one outlier free sample depends exponentially on the number of parameters to instantiate one hypothesis. (iii) A non-linear parameter optimization over all image pairs is proposed. Our method not only can compute the rotational alignment of IMU-camera systems using a single image pair, but also can achieve more robust calibration results with multiple image pairs.

The remainder of the paper is structured as follows. We establish basics and notations for homography constraints for a pure rotation case in Section 2. In Section 3, we derive the minimal case solution by exploiting the rotational angles of the features and describe the non-linear parameter optimization over all image pairs. In Section 4, we validate the proposed method experimentally using both synthetic and real scene data.. Finally, concluding remarks are given in Section 5.

2 Homography constraints

Assume the intrinsic parameters of camera to be known, a general homography relation between two different views is represented as follows [8]:

$$\lambda \mathbf{x}_j = \mathbf{H}\mathbf{x}_i, \tag{1}$$

where $\mathbf{x}_i = [x_i, y_i, 1]^T$ and $\mathbf{x}_j = [x_j, y_j, 1]^T$ are the normalized homogeneous coordinates of the ideally projected image points in views i and j. $\mathbf{H}$ is the homography matrix and λ is a scale factor.

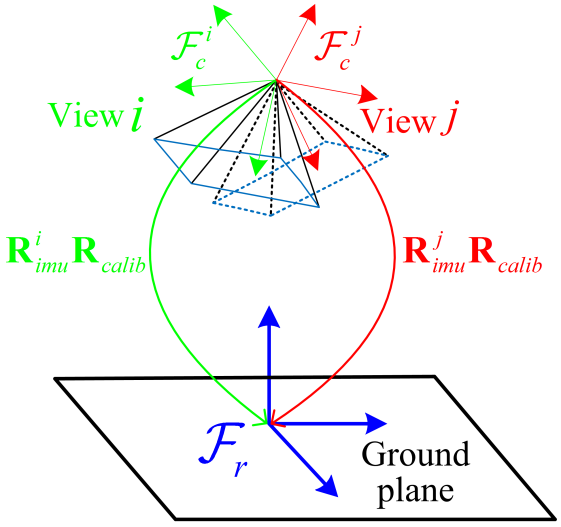


Fig. 2. The relationship between the views i and j.

As shown in Figure 2, the motion of the camera between views i and j is a pure rotation. $\mathcal{F}_r$ denotes IMU reference coordinate system. The camera coordinate systems of the views i and j are expressed with $\mathcal{F}_c^i$ and $\mathcal{F}_c^j$, respectively. The rotations of $\mathcal{F}_c^i$ and $\mathcal{F}_c^j$ in $\mathcal{F}_r$ can be expressed as $\mathbf{R}_{imu}^i\mathbf{R}_{calib}$ and $\mathbf{R}_{imu}^j\mathbf{R}_{calib}$, respectively. The rotational alignment difference between camera coordinate system and IMU coordinate system is expressed with $\mathbf{R}_{calib}$. The image correspondences between views i and j are related by homography and the homography can be written as:

$$\mathbf{H} = \mathbf{R}_{calib}^T(\mathbf{R}_{imu}^j)^T\mathbf{R}_{imu}^i\mathbf{R}_{calib}. \tag{2}$$

The skew-symmetric matrix $[\mathbf{x}_j]_\times$ is multiplied in both sides of Eq. 1 to further eliminate the unknown scale factor λ:

$$[\mathbf{x}_j]_\times\mathbf{H}\mathbf{x}_i = \mathbf{0}. \tag{3}$$

Since the skew-symmetric matrix $[\mathbf{x}_j]_\times$ is only of rank 2, Eq. 3 only imposes two independent constraints on $\mathbf{H}$. In many situations, the approximate installation relationship between the IMU and the camera can be obtained from hand measurements or device layouts [7]. Thus the rotational relationship between the IMU and the camera $\mathbf{R}_{calib}$ can be represented:

$$\mathbf{R}_{calib} = \hat{\mathbf{R}}_{calib}\mathbf{R}_A, \tag{4}$$

where $\mathbf{R}_A$ is the approximate installation relationship between the IMU and the camera, and $\hat{\mathbf{R}}_{calib}$ is the remaining rotation between the IMU and the camera. Since the remaining rotation angles are small, $\hat{\mathbf{R}}_{calib}$ can be expressed by its first-order expansion:

$$\hat{\mathbf{R}}_{calib} = \mathbf{I}_{3\times3} + [\hat{\mathbf{r}}]_\times, \tag{5}$$

where $\hat{\mathbf{r}} = [\hat{r}_x, \hat{r}_y, \hat{r}_z]^T$ is a three-dimensional vector. Thus Eq. 2 can be reformulated as follows:

$$\mathbf{H} = \mathbf{R}_A^T\hat{\mathbf{R}}_{calib}^T(\mathbf{R}_{imu}^j)^T\mathbf{R}_{imu}^i\hat{\mathbf{R}}_{calib}\mathbf{R}_A. \tag{6}$$

3 IMU-camera Calibration

3.1 1pt-RANSAC calibration method

The widely-used SIFT detector not only provides the image coordinates of point correspondence, but also provides the rotational angles and scales of features. Our paper aims at involving the rotational angle of feature into the process to reduce the size of the minimal sample required for IMU-camera calibration. The affine correspondence can be described as a triplet: $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{A})$. The local affine transformation $\mathbf{A}$ is defined as follows[1]:

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} s_x & w \\ 0 & s_y \end{bmatrix} \\ &= \begin{bmatrix} s_x\cos(\alpha) & w\cos(\alpha) - s_y\sin(\alpha) \\ s_x\sin(\alpha) & w\sin(\alpha) + s_y\cos(\alpha) \end{bmatrix}.\end{aligned} \tag{7}$$

Where the rotational angle α is computed by $(\alpha_j - \alpha_i)$, note that the rotational angles of point correspondence α_i and α_j can be obtained directly from the SIFT detector. s_x and s_y are the scales along axes x and y, respectively. w is the shear parameter. $\mathbf{A}$ is given as the first-order approximation of the related homography matrix for perspective cameras:

$$\begin{aligned} a_{11} &= \frac{\partial x_j}{\partial x_i} = \frac{h_{11} - h_{31}x_j}{s}, \qquad a_{12} = \frac{\partial x_j}{\partial y_i} = \frac{h_{12} - h_{32}x_j}{s}, \\ a_{21} &= \frac{\partial y_j}{\partial x_i} = \frac{h_{21} - h_{31}y_j}{s}, \qquad a_{22} = \frac{\partial y_j}{\partial y_i} = \frac{h_{22} - h_{32}y_j}{s}, \end{aligned} \tag{8}$$

where h_{ij} is the element from the ith row and the jth column of the homography matrix $\mathbf{H}$, $s = x_i h_{31} + y_i h_{32} + h_{33}$ is the projective depth.

Based on Eqs. 7 and 8, we obtain the relationship between the rotational angle of the feature and the corresponding homography matrix:

$$\frac{a_{11}}{a_{21}} = \frac{\cos(\alpha)}{\sin(\alpha)} = \frac{h_{11} - h_{31}x_j}{h_{21} - h_{31}y_j}, \tag{9}$$

We further expand Eq. 9 as follows:

$$\sin(\alpha)(h_{11} - h_{31}x_j) - \cos(\alpha)(h_{21} - h_{31}y_j) = 0. \tag{10}$$

Assume one point correspondence $\mathbf{x}_i = [x_i, y_i, 1]^T$, $\mathbf{x}_j = [x_j, y_j, 1]^T$ and the corresponding rotational angle α, obtained by *e.g.* SIFT detector, to be known. Combining Eqs. 3 and 10, we attain 3 polynomial equations in 3 unknowns $\hat{\mathbf{r}} = [\hat{r}_x, \hat{r}_y, \hat{r}_z]^T$:

$$f_w(\hat{r}_x, \hat{r}_y, \hat{r}_z) = 0, \;\; w = 1, 2, 3. \tag{11}$$

The automatic Gröbner basis solver [15] is used to solve the above polynomial equation system. The maximum polynomial degree of Eq. (11) is 2 and there is at most 8 solutions for $\hat{\mathbf{r}}$. This polynomial equation system only needs 140 lines to print out, which leads to an extremely short run-time for the solver. Thus this solver is suitable to perform IMU-camera calibration on smart devices with limited computational power.

In the 1-point RANSAC loop, we obtain the remaining rotation $\hat{\mathbf{R}}_{calib}$ from each solution $\hat{\mathbf{r}} = [\hat{r}_x, \hat{r}_y, \hat{r}_z]^T$ by Eq. 5. The corresponding exact rotation matrix can be retrieved by projecting the matrix to the closest rotation matrix. Then the homography $\mathbf{H}$ for the image features is composed with Eq. 6, and the solution with the maximum number of inliers is selected as the final solution. Finally, the rotational alignment of IMU-camera system $\mathbf{R}_{calib}$ is calculated by Eq. 4.

3.2 Non-linear parameter optimization

For each image pair, $\mathbf{R}_{calib}$ and the corresponding inliers can be obtained by 1-point RANSAC calibration method. The rotational alignment between the IMU and the camera is further optimized based on all the inliers in M image pairs. We minimize the total transfer errors of the inliers and the cost function is defined as follows:

$$\begin{aligned}\varepsilon &= \min_{\bar{\mathbf{R}}} \sum_{p=1}^{M} \sum_{k=1}^{N_p} \left\| \mathbf{x}_j^k - \mathbf{H}_p \mathbf{x}_i^k \right\| \\ &= \min_{\bar{\mathbf{R}}} \sum_{p=1}^{M} \sum_{k=1}^{N_p} \left\| \mathbf{x}_j^k - g(\bar{\mathbf{R}}, \mathbf{R}_{imu}^p) \mathbf{x}_i^k \right\|, \end{aligned} \tag{12}$$

where $\bar{\mathbf{R}}$ is a three-vector used for optimization which is represented in Euler angles. The initial value of $\bar{\mathbf{R}}$ is set to the mean or median angles of M calibration results. Each image pair p is composed of views i and j. N_p represents the number of inliers and k is the index of the inliers within each image pair. $\mathbf{x}_i^k$ and $\mathbf{x}_j^k$ are the homogeneous image coordinates of the inlier k. $\mathbf{R}_{imu}^p$ denotes the IMU rotation matrices of views i and j. The homography $g(\bar{\mathbf{R}}, \mathbf{R}_{imu}^p)$ is the transformation model within each image pair, which transfers the image coordinate $\mathbf{x}_i^k$ in view i to the corresponding image coordinate $\mathbf{x}_j^k$ in view j.

Considering that there may still be a few outliers existed in the image correspondences, the robust cost function created by Cauchy function is used to reduce the influence of outliers:

$$\rho(\varepsilon) = \frac{\sigma^2}{2} \log(1 + \frac{\varepsilon^2}{\sigma^2}), \tag{13}$$

where the σ parameter of the Cauchy function can be set to the inlier threshold of the RANSAC loop.

4 Experiments

The performance of the proposed IMU-camera calibration method is validated using both synthetic and real scene data. To obtain expressive results, we also compare the proposed calibration method to 1.5pt-GB calibration method and 1.5pt-3Q3 calibration method [7]. These methods are suitable for the rotational alignment of IMU-camera systems in the pure rotation case. For 1.5pt-GB and 1.5pt-3Q3 calibration methods, even though only one of the two available equations from the second point is used, both methods still need sample 2 feature correspondences in the RANSAC loop. When a RANSAC scheme is used to cope with feature mismatches, the necessary number of samples to get an outlier free sample with a chance of 99% and an outlier ratio of 50% is 17 for the 1.5pt calibration method, but the 1pt calibration method only need 7 samples.

4.1 Experiments with synthetic data

In the simulation experiments, we assess the calibration error by the root mean square (RMS) of the errors of all trials. The calibration error compares the angle difference between the true rotation and estimated rotation:

$$\xi_{\mathbf{R}} = \arccos((Tr(\mathbf{R}_{gt}\mathbf{R}_{calib}^T) - 1)/2), \tag{14}$$

where $\mathbf{R}_{gt}$ denotes the ground-truth rotation and $\mathbf{R}_{calib}$ is the corresponding estimated rotation.

Accuracy with increasing rotation Since the remaining rotation matrix is approximated to the first-order and the higher-order terms are truncated, the proposed method is evaluated with respect to increasing magnitudes of remaining rotation. We choose three approximate installation angles between the IMU and the camera randomly from $-180°$ to $180°$. Three remaining angles between the IMU and the camera ranges from $0°$ to $10°$ at an interval of $1°$. At each remaining rotation magnitude, 10000 independent trials are conducted, and for each test, one image feature correspondence is generated randomly. We report the results on the data points within the first interval of a 5-quantile partitioning[4] (Quintile) of 10000 trials. As shown in Figure 3, the calibration error of the proposed method increases slowly with increasing magnitudes of remaining rotation. Since all of these methods have utilized a first-order rotation approximation for the remaining rotation matrix, our method has similar accuracy with 1.5pt-GB and 1.5pt-3Q3 calibration methods.

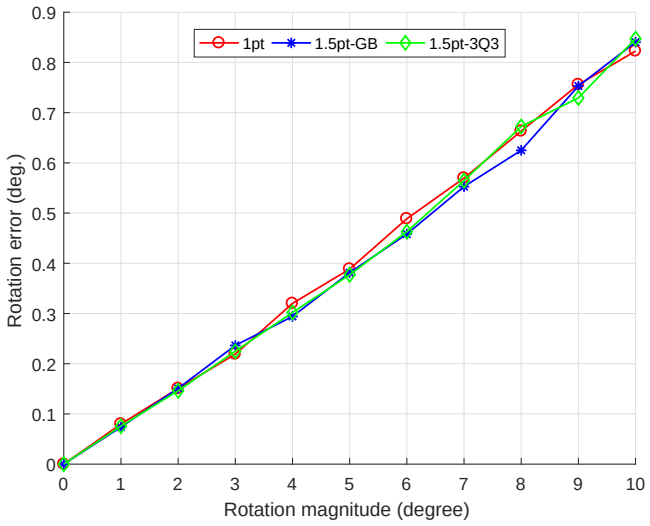


Fig. 3. RMSE for the calibration error with increasing magnitudes of remaining rotation.

Accuracy with increasing image noise We synthesize a pinhole camera with zero skew and an unit aspect ratio. The resolution is 800×640 pixels and the principle point is assumed to be at the image center. The focal length is chosen as 600 pixels, so that one pixel corresponds to about $0.1°$. The approximate installation angles between the IMU and the camera are set to $(180°, 0°, -90°)$, and the remaining rotation angles are set to $(1°, 1°, -1°)$. We add a different level of Gaussian noise to the image feature observations. The standard deviation of Gaussian noise is ranging from 0 to 2 pixels at an interval of 0.1 pixel. At each noise level, 10000 independent trials are conducted, and for each test, one image feature correspondence is generated randomly. We also report the results on the data points within the first interval of a 5-quantile partitioning of 10000 trials. As shown in Figure 4, the calibration error of the proposed method

[4] k-quantiles divide an ordered data set into k regular intervals

increases almost linearly with the increase of image noise. For image noise of more than 0.4 pixel, our calibration method is slightly more accurate than 1.5pt-GB and 1.5pt-3Q3 calibration methods.

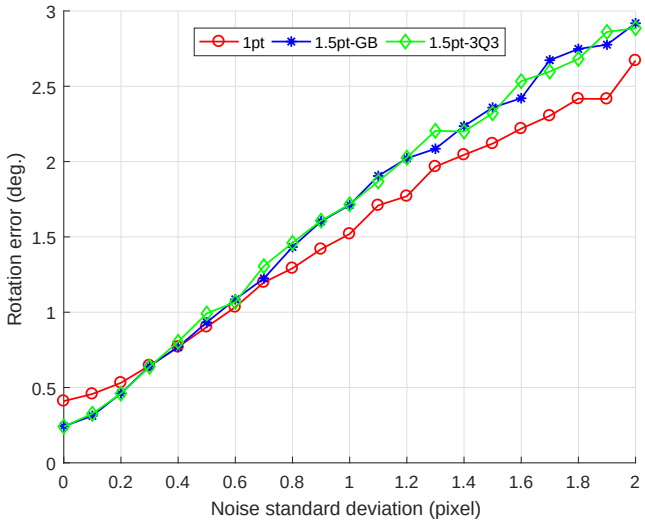
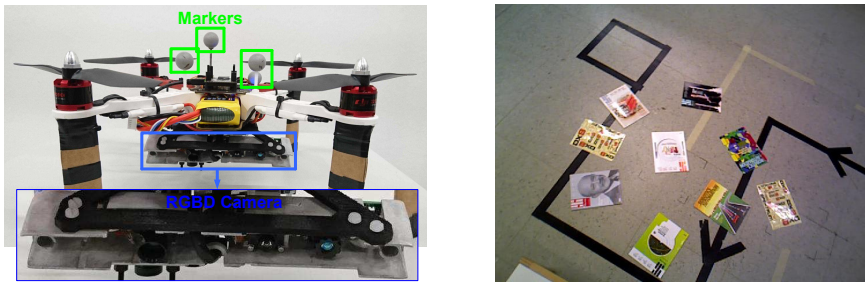


Fig. 4. RMSE for the calibration error with increasing image noise

4.2 Real scene data experiment

We demonstrate the proposed method using a real scene data set under pure rotation, which is acquired with the Pixhawk drone [7], see Figure 5. The markers are attached to the camera mount and the pose is tracked by a motion capture system consisting of 10 cameras. The marker poses are used as IMU data in the experiments. On the basis of the design of the 3D printed mount, the approximate installation angles between the IMU and the RGB camera are $(113°, 0°, 90°)$. The resolution of camera is 640×480 pixels and the intrinsic parameters are calibrated in advance. The camera is typically looking towards the ground and 81 images under pure rotation are captured.



(a) Pixhawk drone (b) Sample image

Fig. 5. Pure rotation data set.

In the 1pt-RANSAC calibration step, we consider feasible image pairs for image matching and feature matches are created using SIFT feature matching [17] for each image pair. The inlier threshold is set to 2 pixels and the maximum number of iterations is set to 1000 in the RANSAC procedure. In the subsequent optimization step, the median and mean angle values of the calibration results of all image pairs are chosen as the initial values for non-linear parameter optimization, respectively. However, the optimization using the inliers of all the image pairs converges to the same result for both initializations. The calibration results of the different calibration methods are shown in Table 1. The calibration result of the proposed method is quite consistent with 1.5pt-GB and 1.5pt-3Q3 calibration methods.

Table 1. The calibration results for the pure rotation data set.

Method	Calibration results (degree)
Approximate installation angle	(113.0, 0.0, 90.0)
1pt	(114.3300, 1.1364, 88.6720)
1.5pt-GB	(114.4211, 1.2609, 88.7395)
1.5pt-3Q3	(114.4241, 1.2845, 88.74310)

To evaluate the accuracy of the calibration results as shown in Table 1, a data set of images for a checkerboard is acquired by the Pixhawk MAV. 49 images are randomly taken around the checkerboard and the image poses are computed by OPnP algorithm [26]. The coordinates of the checkerboard corners are measured by the motion capture system. The ground truth of the relationship between the IMU and the camera can be determined directly by combining with the corresponding IMU data: rotational matrix is $(114.1497°, 1.1152°, 88.7120°)$ and translational vector is $(0.0316m, 0.0222m, -0.0638m)$[7].

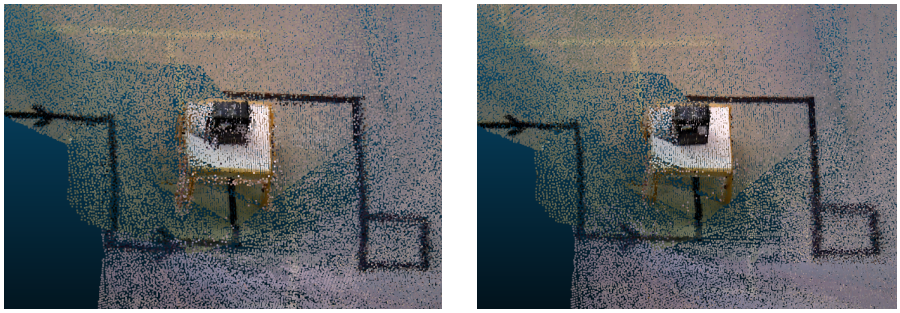
Then the accuracy of the calibration results is evaluated using the reprojection error, which is the mean distance between the measured image corners and the reprojection of the 3D corner. For comparison, the translational vector between the IMU and the camera is fixed as $(0.0316m, 0.0222m, -0.0638m)$. The results of the accuracy evaluation are shown in Table 2. The table shows that the proposed method produces obviously lower reprojection errors than using the approximate installation angles directly, and outperforms 1.5pt-GB and 1.5pt-3Q3 calibration methods in terms of accuracy.

Table 2. The results of the accuracy evaluation.

Calibration results	Ground Truth	3D printer	1pt	1.5pt-GB	1.5pt-3Q3
Reprojection error (pixel)	1.3495	11.5408	2.5066	2.7785	2.8244

The pose of the RGBD camera can be obtained directly from the IMU data and the calibration result. Thus we reconstruct a common scene using the RGBD camera to verify the calibration result intuitively. The offset of the RGB camera and the depth camera has been calibrated beforehand, which is a pure translation $(0.00m, -0.02m, 0.00m)$. The 3D reconstruction results based on the approximate installation angles and the calibration result of the proposed method are shown in Figure 6. There are many false

point clouds around the table and the deviation of the reconstructed line is quite large in the Figure 6(a). The 3D reconstruction result using our calibration result is significantly better than the 3D reconstruction result using the initial values of 3D printing. This experiment successfully demonstrates the practicability of the proposed calibration method. It also means that it is necessary to calibrate the rotational alignment of IMU-camera systems even though the approximate installation angles is known.



(a) Approximate installation angles (b) Calibration result of our method

Fig. 6. 3D reconstruction results.

5 Conclusion

In this paper, we show that by exploiting the rotational angles of the features obtained by *e.g.* the SIFT detector, it is possible to calibrate IMU-camera systems with only 1 feature correspondence in the pure rotation case. Our method need fewer point correspondences for IMU-camera calibration as compared to other calibration methods. The novel minimal case solution is useful to reduce the computation time and increase the calibration robustness, when using Random Sample Consensus (RANSAC) to cope with feature mismatches. Furthermore, a non-linear parameter optimization over all image pairs is performed for a more accurate calibration result. The experimental results of both synthetic and real experiments have demonstrated that our method is suited for the rotational alignment of IMU-camera systems.

References

1. Barath, D.: Five-point fundamental matrix estimation for uncalibrated cameras. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 235–243 (2018)
2. Bender, D., Schikora, M., Sturm, J., Greniers, D.: Ins-camera calibration without ground control points. In: Sensor Data Fusion: Trends, Solutions, Applications (SDF), 2014. pp. 1–6 (2014)
3. Daniilidis, K.: Hand-eye calibration using dual quaternions. The International Journal of Robotics Research **18**(3), 286–298 (1999)

4. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**(6), 381–395 (1981)
5. Fraundorfer, F., Tanskanen, P., Pollefeys, M.: A minimal case solution to the calibrated relative pose problem for the case of two known orientation angles. In: European Conference on Computer Vision. pp. 269–282. Springer (2010)
6. Guan, B., Vasseur, P., Demonceaux, C., Fraundorfer, F.: Visual odometry using a homography formulation with decoupled rotation and translation estimation using minimal solutions. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 2320–2327 (2018)
7. Guan, B., Yu, Q., Fraundorfer, F.: Minimal solutions for the rotational alignment of imu-camera systems using homography constraints. Computer vision and image understanding **170**, 79–91 (2018)
8. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge university press (2003)
9. Heller, J., Havlena, M., Pajdla, T.: Globally optimal hand-eye calibration using branch-and-bound. IEEE Transactions on Pattern Analysis and Machine Intelligence **38**(5), 1027–1033 (2016)
10. Horaud, R., Dornaika, F.: Hand-eye Calibration. International Journal of Robotics Research **14**(3), 195–210 (1995)
11. Hwangbo, M., Kim, J.S., Kanade, T.: Gyro-aided feature tracking for a moving camera: fusion, auto-calibration and gpu implementation. The International Journal of Robotics Research **30**(14), 1755–1774 (2011)
12. Karpenko, A., Jacobs, D., Baek, J., Levoy, M.: Digital video stabilization and rolling shutter correction using gyroscopes. CSTR **1**, 2 (2011)
13. Kelly, J., Sukhatme, G.S.: Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. The International Journal of Robotics Research **30**(1), 56–79 (2011)
14. Kneip, L., Chli, M., Siegwart, R.Y.: Robust real-time visual odometry with a single camera and an imu. In: Proceedings of the British Machine Vision Conference (2011)
15. Kukelova, Z., Bujnak, M., Pajdla, T.: Automatic generator of minimal problem solvers. In: European Conference on Computer Vision. pp. 302–315. Springer (2008)
16. Kukelova, Z., Heller, J., Pajdla, T.: Hand-eye calibration without hand orientation measurement using minimal solution. In: Asian Conference on Computer Vision. pp. 576–589. Springer (2012)
17. Lowe, D.G., Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
18. Mirzaei, F.M., Roumeliotis, S.I.: A kalman filter-based algorithm for imu-camera calibration: Observability analysis and performance evaluation. IEEE Transactions on Robotics **24**(5), 1143–1156 (2008)
19. Park, F.C., Martin, B.J.: Robot sensor calibration: solving ax=xb on the euclidean group. IEEE Transactions on Robotics and Automation **10**(5), 717–721 (1994)
20. Ruland, T., Pajdla, T., Krüger, L.: Globally optimal hand-eye calibration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1035–1042 (2012)
21. Saurer, O., Vasseur, P., Boutteau, R., Demonceaux, C., Pollefeys, M., Fraundorfer, F.: Homography based egomotion estimation with a common direction. IEEE Transactions on Pattern Analysis and Machine Intelligence **PP**(99), 1–1 (2016)
22. Seo, Y., Choi, Y.J., Lee, S.W.: A branch-and-bound algorithm for globally optimal calibration of a camera-and-rotation-sensor system. In: Proceedings of the International Conference on Computer Vision. pp. 1173–1178 (2009)
23. Tsai, R.Y., Lenz, R.K.: A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. IEEE Transactions on Robotics and Automation **5**(3), 345–358 (1989)

24. Weiss, S., Achtelik, M.W., Chli, M., Siegwart, R.: Versatile distributed pose estimation and sensor self-calibration for an autonomous mav. In: 2012 IEEE International Conference on Robotics and Automation. pp. 31–38 (2012)
25. Zhang, Z.Q.: Cameras and inertial/magnetic sensor units alignment calibration. IEEE Transactions on Instrumentation and Measurement **65**(6), 1495–1502 (2016)
26. Zheng, Y., Kuang, Y., Sugimoto, S., Astrom, K., Okutomi, M.: Revisiting the pnp problem: A fast, general and optimal solution. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2344–2351 (2013)
27. Zhuang, H., Shiu, Y.C.: A noise tolerant algorithm for wrist-mounted robotic sensor calibration with or without sensor orientation measurement. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. vol. 2, pp. 1095–1100 (1992)