IMU-Camera Self-Calibration Using Planar Mirror Reflection

Ghazaleh Panahandeh and Magnus Jansson
KTH-Royal Institute of Technology, ACCESS Linnaeus Center, Stockholm, Sweden
Email:{ghazaleh.panahandeh, magnus.jansson}@ee.kth.se

***Abstract*—In this paper, first we look at the problem of estimating the transformation between an inertial measurement unit (IMU) and a calibrated camera, based on images of planar mirror reflection (IPMR) of arbitrary feature points with unknown positions. Assuming that only the reflection of the feature points are observable by the camera, the IMU-camera calibration parameters and the position of the feature points in the camera frame are estimated using the Sigma-Point Kalman filter framework. In the next step, we consider the case of estimating varying camera intrinsic parameters using the estimated static parameters from the previous stage. Therefore, the estimated parameters are used as initial values in the state space model of the system to estimate the camera intrinsic parameters together with the rest of the parameters. The proposed method does not rely on using a fixed calibration pattern whose feature points' positions are known relative to the navigation frame, additionally, motion of the camera which is mounted on the IMU is not limited to be planar with respect to the mirror. Instead, the reflection of the feature points with unknown positions in the camera body frame are tracked over time. Simulation results show subcentimeter and subdegree accuracy for both IMU-camera translation and rotation parameters as well as submillimeter and subpixel accuracy for the position of the feature points and camera intrinsic parameters, respectively.**

***Index Terms*—IMU-Camera calibration, IPMR, Sigma-Point Kalman filter, camera intrinsic parameters.**

I. Introduction

Inertial navigation systems (INS) use a combination of motion sensors and rotational sensors to provide the position, orientation, and velocity of moving objects. At the core of each INS, there is an inertial measurement unit (IMU) that provides angular velocity and specific force, which yields attitude and position. Due to the integration drift, the position and the attitude must be periodically corrected by input from complementary sensors. One of the most common alternative sensors is radio receivers, such as a global positioning system (GPS) for outdoor positioning. Integration of vision sensors with INS is another alternative that can be used for outdoor as well as indoor positioning. However, aiding the INS with a vision sensor such as a camera requires sensor-to-sensor relative transformation to be known; disregarding such an offset in the system will introduce un-modeled biases that may grow over time. Currently, several in-lab calibration techniques have been proposed to determine the 6 degrees-of-freedom (DoF) transformation between the IMU and the camera coordinate frame; however, they are mainly based on using a static checkerboard calibration pattern with a known position in the navigation frame. For instance, in [1] IMU-camera relative rotation and translation have been estimated separately; however, the correlations between the rotation and translation parameters are discarded by the separate calibration. In [2], calibration parameters are estimated by the extended Kalman filter through tracking static points with known positions in a calibration pattern; furthermore, the observability of the nonlinear system is studied by employing the observability rank condition. A similar nonlinear system is studied by [3], to derive the relative translation and orientation between the IMU and a spherical camera. Using the standard camera calibration pattern, a weighted quadratic cost function has been minimized within the standard gray-box system identification framework.

The main contribution of this paper is to present a novel algorithm for estimating the 6-DoF IMU-camera coordinate transformation via a Sigma-Point Kalman filtering framework; unlike the current calibration approaches, our method does not rely on a direct view of a static calibration pattern with known feature points position. Arbitrary feature points are selected in the camera body where no prior knowledge is assumed on the pose of the feature points relative to the camera optical center. Then, the calibration procedure tracks virtual views of the arbitrary feature points in a planar mirror. For this reason, the camera along with the IMU is moving in front of the planar mirror where the IMU is used to support the camera by providing angular velocity and specific force. Contrary to the existing approach [4] [5], no restriction in the IMU-camera movement is considered except the existence of the feature points in IPMRs.

A plane mirror is a mirror with a planar reflective surface. In this case, an image of an object in front of it appears to be behind the mirror plane. In fact, this image is equivalent to one being captured by a virtual camera that is located behind the mirror; additionally, the position of the virtual camera is symmetric to the position of the real camera. In our case, the reflection of the camera in the plane mirror is used for tracking virtual feature points in the IPMRs.

Using IPMR in [6], the transformation between the camera and the body frame of a robot is determined through the maximum-likelihood estimation framework. However, their method is based on tracking feature points whose position relative to the body frame is assumed to be known in advance. Additionally, it is claimed that no prior knowledge about the robot motion or the mirror configuration is considered; since their method is not based on the epipolar geometry [7],

considering such assumptions can be applicable. Considering the epipolar geometry, which relates a pair of images through the fundamental matrix, for the IPMRs is also possible under a certain constraint that limits the camera movement to be planar. In the planar motion, the rotational axis should be normal to the plane containing the direction of translation [4]. As a result the fundamental matrix between IPMRs has 6-DoF [7]. For instance in [8], the epipolar geometry of IPMRs is studied to estimate the pose of a real camera moving in front of the mirror. However, the rigid displacements between the virtual views are restricted to be planar.

The above methods all require intrinsically calibrated cameras, which can be achieved using a fixed calibration pattern. However, the availability of the IMU signal is one advantage of such a system that can be addressing the problem of camera self-calibration. For instance, in [9], these parameters are estimated using homographies, on the assumption of known relative rotation from an external sensor. Due to the camera intrinsic parameter estimation, in the second part of this paper, a camera-self calibration approach is introduced which uses the same IMU-camera and planar mirror structure. Using the estimated IMU-camera calibration parameters and pose of the target feature points in the camera body frame, the camera intrinsic parameters are estimated by IPMRs. The paper is organized as follows. The process and measurement models, which lead to the state-space equations of the system, is derived in Section II. The structure of the used Sigma-point Kalman filter algorithm is presented in Section II. The camera intrinsic parameter estimation is studied in Section III. In Section IV, the performance of the proposed method and simulation results in different scenarios are examined. Finally, the conclusion of the study is summarized in Section V.

In the following sections scalars are denoted by lowercase letters (s), vectors by bold letters ($\mathbf{f}$), and matrices by bold capitals ($\mathbf{K}$).

II. System Description

The goal of our proposed algorithm in this section is to estimate the position and orientation of the camera coordinate frame $\{c\}$ relative to the IMU frame $\{b\}$, where the camera is rigidly mounted in the IMU body frame. In order to simplify the treatment with different coordinate frames, we assume that the navigation frame $\{n\}$ is located in the center of the mirror coordinate frame.

A. Time evolution of the system

In order to estimate the parameters in the Sigma-Point Kalman filter framework, we first describe the total system state vector as

$$\mathbf{x} = [\mathbf{x}^{ins\top} \quad \mathbf{x}^{imuc\top} \quad \mathbf{x}^{f\top}]^\top \in \mathbb{R}^{21+3M}, \tag{1}$$

with

$$\mathbf{x}^{ins} = [\mathbf{p}_b^{n\top} \quad \mathbf{v}_b^{n\top} \quad \mathbf{q}_b^{n\top} \quad \mathbf{b}_a{}^\top \quad \mathbf{b}_g{}^\top]^\top,$$
$$\mathbf{x}^{imuc} = [\mathbf{p}_c^{b\top} \quad \mathbf{q}_c^{b\top}]^\top, \text{ and } \quad \mathbf{x}^{f} = [\pi_1^{c\top}, \ldots, \pi_M^{c\top}]^\top.$$

The considered state variables in the INS system are represented by $\mathbf{x}^{ins}$, where $\mathbf{p}_b^n$ and $\mathbf{v}_b^n$ are describing the position and velocity of the IMU in the navigation frame, respectively. $\mathbf{q}_b^n$ is the unit quaternion representing the rotation from body frame to navigation frame, and $\mathbf{b}_a$ and $\mathbf{b}_g$ are the bias vectors affecting the accelerometer and gyroscope measurements, respectively. $\mathbf{x}^{imuc}$ depicts IMU-camera calibration parameters, containing the translation from camera to the body frame $\mathbf{p}_c^b$, and the unit quaternion representing the rotation from camera frame to body frame $\mathbf{q}_c^b$. Finally, $\mathbf{x}^f$ contains pose of the arbitrary static feature points located in the camera body frame $\{\pi_i^c\}_{i=1}^M \in \mathbb{R}^3$. The time evolution of the INS state and the IMU-camera transformation [2] can be described by

$$\begin{aligned}
&\dot{\mathbf{q}}_n^b(t) = \frac{1}{2}\Omega(\omega(t))\mathbf{q}_n^b(t), \quad \text{where} \quad \Omega(\omega) = \begin{bmatrix} -[\omega]_\times & \omega \\ -\omega^\top & 0 \end{bmatrix} \\
&\dot{\mathbf{p}}_b^n(t) = \mathbf{v}_b^n(t) && \dot{\mathbf{v}}_b^n(t) = \mathbf{a}^n(t) \\
&\dot{\mathbf{f}}^b(t) = \mathbf{n}_{\delta\mathbf{f}}(t) && \dot{\omega}^b(t) = \mathbf{n}_{\delta\omega}(t) \\
&\dot{\mathbf{p}}_c^b(t) = \mathbf{0}_{3\times1} && \dot{\mathbf{q}}_c^b(t) = \mathbf{0}_{3\times1}
\end{aligned} \tag{2}$$

where $\omega(t)$ is the rotational velocity of the body frame and $[\cdot]_\times$ denotes the skew-symmetric matrix representation of the cross product operation [10]. The IMU and gyroscope bias increments, $\mathbf{n}_{\delta\mathbf{f}}$ and $\mathbf{n}_{\delta\omega}$, are modeled as white Gaussian noises. The output measurement signals of the accelerometer $\mathbf{f}_m$ and the gyroscope ω_m are modeled as

$$\begin{aligned}
\mathbf{f}_m(t) &= \mathbf{R}_n^b(t)(\mathbf{a}^n(t) - \mathbf{g}^n) + \mathbf{b}_a(t) + \mathbf{n}_\mathbf{f}(t) \\
\omega_m(t) &= \omega(t) + \mathbf{b}_g(t) + \mathbf{n}_\omega(t)
\end{aligned} \tag{3}$$

where $\mathbf{R}_n^b$ is the direction-cosine matrix [10], $\mathbf{n}_\mathbf{f}$, $\mathbf{n}_\omega$ are modeled with Gaussian distribution, and $\mathbf{g}^n$ is the gravitational acceleration expressed in the navigation frame.

B. Discrete-time process model

The discrete-time error state space model is described by

$$\begin{aligned}
\delta\mathbf{p}_{b,k+1}^n &= \delta\mathbf{p}_{b,k}^n + dt\,\delta\mathbf{v}_{b,k}^n \\
\delta\mathbf{v}_{b,k+1}^n &= \delta\mathbf{v}_{b,k}^n + dt[\hat{\mathbf{R}}_{b,k}^n\mathbf{f}_{m,k}]_\times\delta\theta_k + dt\hat{\mathbf{R}}_{b,k}^n(\delta\mathbf{f}_k^b + \mathbf{n}_{\mathbf{f},k}) \\
\delta\theta_{k+1} &= \delta\theta_k - dt\hat{\mathbf{R}}_{b,k}^n(\delta\omega_k^b + \mathbf{n}_{\omega,k})
\end{aligned} \tag{4}$$

$$\begin{aligned}
\delta\mathbf{f}_{k+1}^b &= \delta\mathbf{f}_k^b + dt\mathbf{n}_{\delta\mathbf{f},k} && \delta\omega_{k+1}^b = \delta\omega_k^b + dt\mathbf{n}_{\delta\omega,k} \\
\delta\mathbf{x}_{k+1}^{imuc} &= \delta\mathbf{x}_k^{imuc} && \delta\mathbf{x}_{k+1}^f = \delta\mathbf{x}_k^f
\end{aligned}$$

where $\hat{\mathbf{R}}_b^n$ is the measurement rotation matrix, and the equations have been derived based on the standard additive error definition for the position, velocity, and biases ($\hat{x} \simeq x + \delta x$) and quaternion error for the rotational angles ($\delta\mathbf{q} \simeq [1 \quad \frac{\delta\psi}{2}]^\top$); the advantage of quaternion error definition is the direct use of error angle vectors $\delta\theta$ and $\delta\varphi$ for the $\mathbf{q}_b^n$ and $\mathbf{q}_c^b$, respectively. According to the state vector model (1) and the discrete-time error state space model (4), the total error state vector is written as

$$\delta\mathbf{x} = [\delta\mathbf{x}^{ins\top} \quad \delta\mathbf{x}^{imuc\top} \quad \delta\mathbf{x}^{f\top}]^\top \tag{5}$$

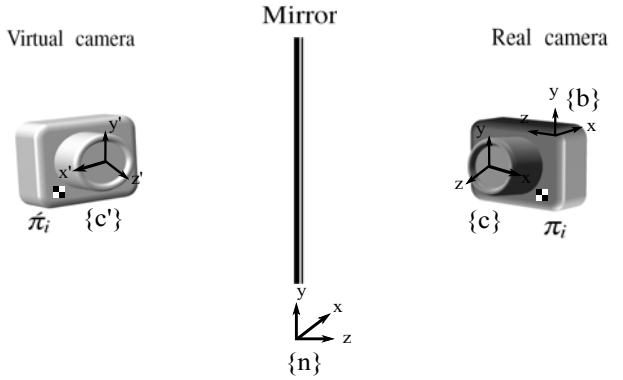

Fig. 1: System sensor structure in front of the planar mirror

with

$$\delta\mathbf{x}^{ins} = [\delta\mathbf{p}_b^{n\top} \quad \delta\mathbf{v}_b^{n\top} \quad \delta\theta^\top \quad \delta\mathbf{b}_a{}^\top \quad \delta\mathbf{b}_g{}^\top]^\top,$$
$$\delta\mathbf{x}^{imuc} = [\delta\mathbf{p}_c^{b\top} \quad \delta\varphi^\top]^\top, \text{ and } \delta\mathbf{x}^f = [\delta\pi_1^{c\top}, \ldots, \delta\pi_M^{c\top}]^\top.$$

Hence, the state-space model of the discrete time process is given by

$$\delta\mathbf{x}_{k+1} = f_k(\delta\mathbf{x}_k, \mathbf{n}_k) \in \mathbb{R}^{21+3M} \tag{6}$$

where the process noise $\mathbf{n}_k = [\mathbf{n}_{\mathbf{f},k}^\top \quad \mathbf{n}_{\omega,k}^\top \quad \mathbf{n}_{\delta\mathbf{f},k}^\top \quad \mathbf{n}_{\delta\omega,k}^\top]^\top$ is assumed to be time invariant.

C. Mirror reflection transformation

To define the discrete time measurement model of the system, first of all, we need to derive the geometry of the reflected point in the planar mirror. Without loss of generality, a standard right-handed Cartesian coordinate system is used for all the coordinate frames. The navigation frame, $\{n\}$, is located in the center of the mirror coordinate frame, see Figure 1, such that its *xy* plane is the reflective surface. For the sake of simplicity, equations are derived for only one feature point π_i which is selected arbitrary in the camera body frame. Defining the reflection of this point in the planar mirror by $\acute{\pi}_i$, they can be related in the navigation frame via the reflection matrix $\mathbf{A}$ as

$$\acute{\pi}_i^n = \mathbf{A}\pi_i^n \in \mathbb{R}^3, \quad \text{where} \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{7}$$

The geometric relation between the pose of π_i and $\acute{\pi}_i$ can also be described in the camera coordinate frame by considering the trasformation between the camera and the navigation frame as

$$\pi_i^c = \mathbf{R}_n^c\pi_i^n + \mathbf{p}_n^c \tag{8a}$$
$$\acute{\pi}_i^c = \mathbf{R}_n^c\acute{\pi}_i^n + \mathbf{p}_n^c. \tag{8b}$$

By substituting (8a) and (7) into (8b), we have

$$\acute{\pi}_i^c = \mathbf{R}_n^c\mathbf{A}\mathbf{R}_n^{c\top}(\pi_i^c - \mathbf{p}_n^c) + \mathbf{p}_n^c. \tag{9}$$

Inserting the IMU-camera transformation, $\mathbf{R}_n^c = \mathbf{R}_c^{b\top}\mathbf{R}_n^b$ and $\mathbf{p}_c^n = \mathbf{p}_b^n + \mathbf{R}_n^{b\top}\mathbf{p}_c^b$, in (9) and using the fact that $\mathbf{p}_c^n = -\mathbf{R}_n^{c\top}\mathbf{p}_n^c$, the vector $\acute{\pi}_i^c$ can be rewritten as

$$\begin{aligned}\acute{\pi}_i^c =& \mathbf{R}_c^{b\top}\mathbf{R}_n^b\mathbf{A}(\mathbf{R}_c^{b\top}\mathbf{R}_n^b)^\top\pi_i^c - \mathbf{R}_c^{b\top}\mathbf{R}_n^b(\mathbf{I}_3 - \mathbf{A})\mathbf{p}_b^n \\ & - (\mathbf{I}_3 - \mathbf{R}_c^{b\top}\mathbf{R}_n^b\mathbf{A}(\mathbf{R}_c^{b\top}\mathbf{R}_n^b)^\top)\mathbf{R}_c^{b\top}\mathbf{p}_c^b.\end{aligned} \tag{10}$$

Using the unit vector along the *z* axis, $e_z = [0,0,1]^\top$, the matrix $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{I}_3 - 2e_ze_z^\top$, where $\mathbf{I}_3$ is the identity matrix. By replacing $\mathbf{A}$, (10) can be simplified to

$$\acute{\pi}_i^c = \pi_i^c - 2\mathbf{R}_c^{b\top}\mathbf{R}_n^b e_z e_z^\top(\mathbf{R}_n^{b\top}\mathbf{R}_c^b\pi_i^c + \mathbf{p}_b^n + \mathbf{R}_n^{b\top}\mathbf{p}_c^b). \tag{11}$$

Finally, the position of the reflected feature points in the camera coordinate frame, which will be used in the camera projection model II-D, is described as a function of our defined states (1). The main advantage of this model is that instead of using $\acute{\pi}_i^c$ that maybe change from image to image while camera is moving, π_i^c is used which is fixed in the camera body frame and can be estimated efficiently over time by the Kalman filter framework. This implies, although no prior knowledge about the pose of π_i is assumed, it should be static relative to the camera and visible in each captured image.

D. Measurement model

When the camera along with the IMU is moved in front of a plane mirror, the body frame angular velocity and specific force are measured by the IMU. Meanwhile, the camera records IPMRs. The projection of the reflected feature point $\acute{\pi}_i^c$ to the image plane, based on the pinhole camera model [7], can be represented by

$$\mathbf{z}_i = h_i(\mathbf{x}^{ins}, \mathbf{x}^{imuc}, \acute{\pi}_i^c) + \mathbf{v}_i = [\mathbf{I}_2 \quad \mathbf{0}]\mathbf{K}\begin{bmatrix} \acute{\pi}_{i,x}^c \\ \acute{\pi}_{i,y}^c \\ \acute{\pi}_{i,z}^c \end{bmatrix}\frac{1}{\acute{\pi}_{i,z}^c} + \mathbf{v}_i \in \mathbb{R}^2$$

where $\mathbf{v}_i$ is the feature-measurement noise with covariance matrix $\mathbf{R}_i = \sigma_v^2\mathbf{I}_2$, and $\mathbf{K}$ is the camera intrinsic matrix, assumed to be known. For M observed reflected feature points described by (11), the discrete-time measurement model of the system at the kth time instant is

$$\delta\bar{\mathbf{z}}_k = h_k(\delta\mathbf{x}_k^{ins}, \delta\mathbf{x}_k^{imuc}, \delta\mathbf{x}_k^f) + \mathbf{v}_k \in \mathbb{R}^{2M}. \tag{12}$$

Using the Sigma-Point Kalman Filter framework [11], the statistics of the random variables under the nonlinear process model (6) and measurement model (12) are calculated. Additionally, this technique alleviates the requirement to explicitly calculate Jacobians, which for complex functions, such as (11), can be a difficult task or numerically unstable. The high-rate IMU signal measurement propagates the state and the covariance matrix before a new measurement is received. From each new captured image the reflection of feature points are detected, then the state estimates and the covariance matrices of the system are updated.

III. Camera intrinsic parameter estimation

The proposed approach in Section II provides an accurate estimate of the static IMU-camera calibration parameters and the positions of the feature points for a calibrated camera.

However, there are cases in which the intrinsic camera parameters change (e.g, the focal length). In this section, the previous system structure has been extended to estimate the varying camera intrinsic parameters using the estimated IMU-camera calibration parameters and feature points positions as an initial value in the same Sigma-Point Kalman filter framework.

The camera intrinsic matrix $\mathbf{K}$ is represented as

$$\mathbf{K} = \begin{bmatrix} k_u & s & p_u \\ 0 & k_v & p_v \\ 0 & 0 & 1 \end{bmatrix}, \tag{13}$$

where k_u and k_v are the magnifications in the two coordinate directions of the image, s is a skew parameter corresponding to a skewing of the coordinate axes, and p_u and p_v are the coordinates of the principal point [7]. Concatenating the camera intrinsic parameters $\mathbf{x}^{cc} = [k_u \quad k_v \quad p_u \quad p_v \quad s]^\top$ to the system state vector (1), the extended system state vector can be written as

$$\mathbf{x} = [\mathbf{x}^{ins\top} \quad \mathbf{x}^{imuc\top} \quad \mathbf{x}^{f\top} \quad \mathbf{x}^{cc\top}]^\top \in \mathbb{R}^{26+3M}. \tag{14}$$

Considering $\delta\mathbf{x}^{cc}_{k+1} = \delta\mathbf{x}^{cc}_{k}$ in the discrete time error state space model, the total error state vector is expanded to

$$\delta\mathbf{x} = [\delta\mathbf{x}^{ins\top} \quad \delta\mathbf{x}^{imuc\top} \quad \delta\mathbf{x}^{f\top} \quad \delta\mathbf{x}^{cc\top}]^\top. \tag{15}$$

Hence, the discrete-time measurement model of the system (12) will be a function of $\delta\mathbf{x}^{cc}$ as well as the rest of the parameters (5).

IV. Performance Evaluation

The proposed calibration approach has been evaluated by Monte-Carlo simulations for three selected feature points, which is the minimum number of required points for defining a plane. Table I and II summarize the final estimated values and the standard deviation of the error (σ) for IMU-camera 6-DoF and the position of feature points. Simulation results show that the proposed estimation method is able to reach subcentimeter and subdegree accuracy for the IMU-camera rotation and translation as well as subcentimeter errors for the position of feature points in the camera body frame. The initial and final estimates of the camera intrinsic parameters using the estimated value of the IMU-camera 6-DoF and feature point positions as the initial values in the extended Sigma-Point Kalman filter are shown in Table III.

TABLE I: Initial, final, and error statistics of the IMU-camera transformation, for 100 Monte Carlo simulations and calibrated camera.

	$\mathbf{p}^b_{c,x} \pm \sigma$[cm]	$\mathbf{p}^b_{c,y} \pm \sigma$ [cm]	$\mathbf{p}^b_{c,z} \pm \sigma$[cm]	$\varphi_\phi \pm \sigma[^\circ]$	$\varphi_\theta \pm \sigma[^\circ]$	$\varphi_\psi \pm \sigma[^\circ]$
Initial	1 ± 1	-5 ± 1	10 ± 1	-90 ± 2	0 ± 2	-90 ± 2
Final	0.69 ± 0.53	-5.08 ± 0.16	10.12 ± 0.31	-89.99 ± 0.01	0 ± 0.02	-89.98 ± 0.06

TABLE II: Initial, final, and error statistics of 3 selected feature points' positions, for 100 Monte Carlo simulations and calibrated camera.

	$\pi^c_{x,1} \pm \sigma$[cm]	$\pi^c_{y,1} \pm \sigma$ [cm]	$\pi^c_{z,1} \pm \sigma$[cm]	$\pi^c_{x,2} \pm \sigma$[cm]	$\pi^c_{y,2} \pm \sigma$[cm]	$\pi^c_{z,2} \pm \sigma$[cm]	$\pi^c_{x,3} \pm \sigma$[cm]	$\pi^c_{y,3} \pm \sigma$[cm]	$\pi^c_{z,3} \pm \sigma$[cm]
Initial	2 ± 10	-3 ± 10	-3 ± 10	5 ± 10	4 ± 10	-3 ± 10	-8 ± 10	6 ± 10	1 ± 10
Final	1.98 ± 0.03	-2.99 ± 0.03	-3.00 ± 0.03	4.98 ± 0.05	4.00 ± 0.04	-3.00 ± 0.03	-8.01 ± 0.08	6.00 ± 0.06	0.99 ± 0.01

TABLE III: Initial, final, and error statistics of the camera intrinsic parameters, for 100 Monte Carlo simulations.

	$k_u \pm \sigma$[pixel]	$k_v \pm \sigma$ [pixel]	$p_u \pm \sigma$[pixel]	$p_v \pm \sigma$[pixel]	$s \pm \sigma$[pixel]
Initial	833.33 ± 100	833.33 ± 100	2 ± 100	8 ± 100	3 ± 100
Final	833.93 ± 1.26	833.94 ± 1.33	2.73 ± 0.92	7.74 ± 0.95	3.31 ± 0.50

V. Conclusion

An approach for IMU-camera self-calibration has been proposed for estimating the 6-DoF IMU-camera transformation. The method does not require a fixed calibration pattern with known feature point positions. Instead, our calibration method is based on IPMR of the feature points that are fixed in the camera body frame where no prior knowledge of their positions is assumed. Relating the IPMR with the IMU measurements, we introduced a model which has been used in the Sigma-Point Kalman filter framework to estimate the IMU-camera calibration parameters as well as the position of the feature points. Moreover, no restriction for the IMU-camera movement in front of the planar mirror is considered; however, the reflection of the feature points in the plane mirror should be visible in the images.Additionally, using the estimated parameters, we show that it is possible to estimate the camera intrinsic parameters in the same Sigma-Point Kalman filter approach extended with the camera calibration matrix parameters.

References

[1] J. Lobo and J. Dias, "Relative pose calibration between visual and inertial sensor," *Int. Journal of Robotics Research*, vol. 26, pp. 561–575, Jun. 2007.

[2] F. Mirzaei and S. Roumeliotis, "A kalman filter-based algorithm for imu-camera calibration," in *Proceedings of Int. Conf. on Intelligent Robots and Systems,IEEE/RSJ*, pp. 2427–2434, Nov. 2007.

[3] J. Hol, T. Schön, and F. Gustafsson, "Relative pose calibration of a spherical camera and an imu," in *Proceedings of 7th IEEE/ACM Int. Symposium on Mixed and Augmented Reality, ISMAR.*, pp. 21–24, 2008.

[4] J. Gluckman and S. Nayar, "Planar catadioptric stereo: geometry and calibration," in *Proceedings of Computer Vision and Pattern Recognition. IEEE Computer Society Conference.*, 1999.

[5] J. Gluckman and S. K. Nayar, "Catadioptric stereo using planar mirrors," *Int. J. Comput. Vision*, vol. 44, pp. 65–79, Aug 2001.

[6] J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis, "Mirror-based extrinsic camera calibration," in *The Eighth Int. Workshop on the Algorithmic Foundations of Robotics, Mexico*, Dec. 2008.

[7] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[8] R. Rodrigues, J. P. Barreto, and U. Nunes, "Camera pose estimation using images of planar mirror reflections," in *Proceedings of the 11th European conference on Computer vision: Part IV*, ECCV'10, (Berlin, Heidelberg), pp. 382–395, Springer-Verlag, 2010.

[9] J.-M. Frahm and R. Koch, "Camera calibration with known rotation," in *Proceedings of IEEE Int. Conf. Computer Vision ICCV*, pp. 1418–1425, 2003.

[10] J. A. Farrell and M. Barth, *Global Positioning System, Inertial Navigation and Integration*. McGraw-Hill Companies, 1999.

[11] S. J. Julier and J. K. Uhlmann, "A new extension of the kalman filter to nonlinear systems," in *Proceedings of Signal Processing, Sensor fusion, and Target Recognition*, vol. 4, pp. 182–193, Apr. 1997.