

INS-Camera Calibration without Ground Control Points

Daniel Bender*, Marek Schikora*, Jürgen Sturm[†] and Daniel Cremers[†]

*Department Sensor Data and Information Fusion, Fraunhofer FKIE, Wachtberg, Germany

Email: daniel.bender@fkie.fraunhofer.de, marek.schikora@fkie.fraunhofer.de

[†]Computer Vision Group, Technical University of Munich, Garching, Germany

Email: sturmju@in.tum.de, cremers@in.tum.de

Abstract—In this paper we present an approach for performing system calibration of a sensor suite consisting of a fixed mounted camera and an inertial navigation system without the usage of a photogrammetric calibration site. The aim of the presented work is to obtain accurate direct georeferencing of camera images collected with manned or unmanned aerial systems for search and rescue tasks. These time-critical missions require a straightforward calibration which can be performed without additional equipment. This induces an in-flight calibration of the intrinsic camera parameters in addition to the mounting offsets between the camera and the inertial navigation system. The optimization of these values can be done by introducing them as parameters into a bundle adjustment process. We show how to solve this by exploiting a graph optimization framework. The evaluation of the proposed approach with data from flight experiments leads to an improvement of roughly factor six compared to a terrestrial calibration.

I. INTRODUCTION

Aerial photogrammetry is widely used for orthophoto and digital terrain model creation. Nowadays these measuring flights are usually performed by an aircraft equipped with a digital large-format camera and a GPS corrected inertial navigation system (INS). This equipment enables the direct determination of the camera positions and orientations for the captured images. Therefore it provides, in contrast to the classical bundle adjustment, a very flexible flight planning and no necessity of ground control points (GCP) during the operation. Nevertheless, the essential system calibration is generally realized by a flight over a photogrammetric test site with a high number of signaled GCPs.

Time-critical surveillance and rescue tasks have the same demand for a flexible flight planning and the direct determination of object positions from image observations [1], [2]. In contrast to aerial photogrammetry, there is a need for a straightforward system calibration, which can be performed without the time-consuming and cost-intensive usage of large-scale calibration areas. Within the so called INS-camera calibration the static coordinate system transformation between the rigidly mounted sensors is determined. More precisely the translational offsets (lever-arm) and the angle misalignments (boresight) have to be estimated. The most accurate calibration procedures integrate these parameters as unknowns in a bundle adjustment (BA) [3], [4].

In this work, we formulate the system calibration as a graph optimization problem by representing the parameters to optimize as vertices and the observations between them as



Fig. 1. Our goal is to calibrate the static coordinate system offsets between a camera and an INS, which enables the utilization of the INS measurements as exterior camera orientations. For our flight experiments we integrated the sensors into a payload pod, which was mounted beneath one wing of a manned ultra light airplane.

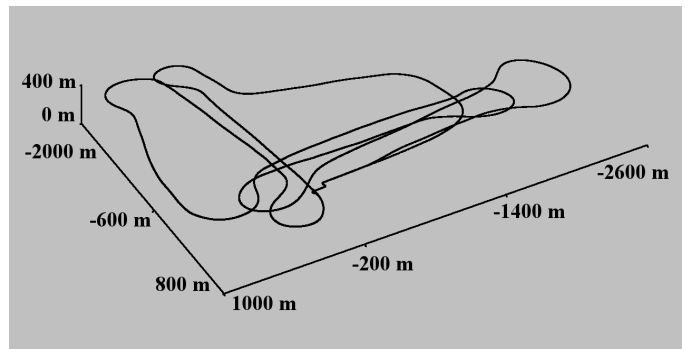


Fig. 2. Visualization of a flight course performed for the evaluation of our approach. The altitude was about 300 meters.

edges. We determine the optimal configuration of the state variables by using a graph optimization framework [5]. In contrast to previous work relying on BA [3], our approach is more intuitive and can easily be adapted when other calibration parameters become relevant.

Our previous work on simulated data indicated high accuracies for the graph based system calibration [6]. We extend the procedure to perform the calibration with real-world data from sensors mounted on an aerial vehicle. In our experiments, we used a manned ultra light airplane with a wingspan of 9.6 meters (Fig. 1) flying in heights of 300 to 800 meters (Fig. 2). Our evaluation shows that a calibration can be

performed without the usage of GCPs. The paper starts with an overview on related research areas. It follows the description of the realized system calibration. Finally experimental results achieved with this approach are shown and discussed.

II. RELATED WORK

A related process of the INS-camera calibration is the so called hand-eye calibration. Given a camera mounted on a robot arm, the rigid-body transformation between the coordinate systems of these devices is estimated. As a result, measurements from the acquired images can be transformed into the robot arm coordinate system to interact with objects recognized and located in the images. The calibration out of corresponding robot arm and camera poses was realized with a direct solution for the rotational part followed by solving the equations for the translation [7]. Alternatively a nonlinear optimization of the whole transformation at the same time was proposed in [8].

These approaches had a big influence on the calibration between an inertial measurement unit (IMU) and a camera. Measurements of the IMU in form of rotational velocities and linear accelerations can be integrated to determine the positions, velocities and orientations of the device. The small estimation errors which are summed up over time can be corrected by camera measurements, if the transformation between the two devices is known. In [9], the estimation was realized with a modified hand-eye calibration algorithm. Other approaches utilize a Kalman filter for the estimation of the pose transformation [10], [11]. This sensor combination is known as a vision-aided INS.

By contrast, GPS-aided INS exploit the GPS measurements to correct the IMU estimations. In conjunction with GPS correction signals from ground control stations, accuracies in the range of a few centimeters for the positions and a few hundredths of a degree for the orientations are achievable. Thus the INS provides a reliable stand-alone source describing its own movement. This leads to the estimation of the rigid-body transformation between the INS and the camera with methods similar to the hand-eye calibration. In a first step the camera movement is calculated with a structure from motion (SFM) approach and refined in a BA procedure. The observations of ground control points are used to scale the 3D model to real world coordinates. In a second step, the transformation between the two devices is estimated by relating these absolute camera poses to time synchronized measurements from the INS. This widely used approach is known as two-step procedure [12]. The advantage is that each bundle adjustment package can be used without modifications. On the other hand, the integration of the mounting parameters as variables to optimize in the BA is possible. This approach is known as single-step calibration and induces a simpler estimation of the mounting offsets due to more flexible flight courses [3]. The estimation of the INS-camera calibration has been examined widely within the OEEPE test [4]. They conclude that this approach is a serious alternative for many applications, even though it does not achieve the mapping accuracies of classical bundle adjustment.

The studies for the INS-camera calibration performed in the last decades were targeted at aircrafts equipped with high-precision INS and metric large-format cameras performing

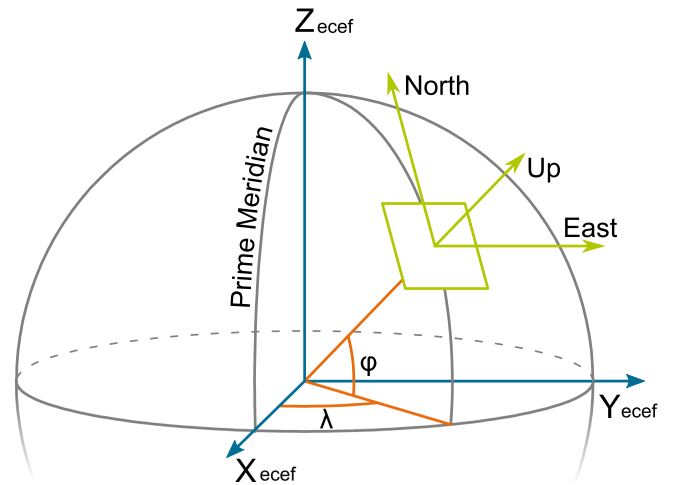


Fig. 3. Both, the navigation systems of the INS and our world reference frame are based on East, North, Up (ENU) coordinate systems. They are local Cartesian coordinate systems with the origin tangential to the earth ellipsoid. Further we use the global earth-centered, earth fixed (ECEF) coordinate system, to convert between ENU-coordinates and GPS measurements taken in latitude (φ) and longitude (λ) as polar coordinates.

flights at photogrammetric calibration sites [4]. In contrast to this our goal is to achieve direct georeferencing with small manned or unmanned aerial systems, without relying on time-consuming and cost-intensive GCPs. We phrase the calibration problem as a graph that we solve using a standard graph optimizer. This approach allows the intuitive integration of the mounting offsets as parameters to optimize and can easily be adapted when more parameters become relevant.

III. INS-CAMERA CALIBRATION

In this section, we describe the performed calibration of a sensor suite comprising a camera and an INS. First, we define the utilized coordinate systems and our parametrization of the rigid-body motions as well as the camera mapping. Then, we state the problem considered in this work and show how to solve it with a graph based approach.

A. Preliminaries

The INS measures the orientation as Euler angles with regard to local navigation systems. The latter are local East, North, Up (ENU) coordinate systems, each with the origin in the current device position described through a GPS measurement in form of latitude, longitude and altitude (Fig. 3). This implies tiny differences between the orientations of these coordinate systems and thus also in the measured angles for successive timestamps. Our INS-camera calibration will be performed in an area of less than one square kilometer, which allows to neglect these differences. Besides, we perform the INS-camera calibration in an ENU coordinate system W with the origin in the middle of the observed area. This prevents the need for earth curvature corrections which have to be taken into consideration for mapping frames like the Universal Transverse Mercator (UTM) coordinate system.

The rigid body motions $\mathbf{g}_{c,w}$ describe the camera poses at the exposure times $t_i, i = 1, 2, \dots, n$ with regard to the world frame W . Likewise, $\mathbf{g}_{i,w}$ specify the corresponding

configurations of the INS as rigid body motions. In general, a rigid body motion $\mathbf{g} \in SE(3)$ describes how the points of a rigid object change over time. Instead of considering the continuous path of the movement, we bring into focus the mapping between the initial and the final configuration of the rigid body motion. This movement can be described by a rotation matrix $\mathbf{R} \in SO(3)$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$. Consequently the rigid body displacement G of a 3D point $\mathbf{p} \in \mathbb{R}^3$ can be performed as

$$G : SE(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad G(\mathbf{g}, \mathbf{p}) = \mathbf{R}\mathbf{p} + \mathbf{t}. \quad (1)$$

The representation of the rotational part in form of the over-determined rotation matrix \mathbf{R} is not suitable for the optimization performed in this work. Thus we use a minimal representation in form of YXZ Euler angles $(\psi, \theta, \phi) \in \mathbb{R}^3$, called yaw, pitch and roll. These describe the rotational part of the movement as a mapping from \mathbb{R}^3 to $SO(3)$ by consecutive rotations around the principal axes as

$$\mathbf{R} = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix}. \quad (2)$$

The YXZ rotation order leads to singularities at $\theta = \pm\pi/2$. This corresponds to a pitch angle of $\pm 90^\circ$, which will never be achieved in practice by the considered aerial platforms and therefore does not constitute a problem.

Further we define the set of intrinsic camera calibration parameters

$$\mathbf{k} = \{f_x, f_y, o_x, o_y, k_1, k_2, k_3, p_1, p_2\}, \quad (3)$$

whereby (f_x, f_y) describe the focal length and (o_x, o_y) the principal point of the camera. Further we express the radial distortion with the parameters (k_1, k_2, k_3) and the tangential distortion with (p_1, p_2) . The projection π performs the mapping from a transformed 3D point $G(\mathbf{g}, \mathbf{p}) = (x, y, z)^\top$ to pixel coordinates by

$$\pi(\mathbf{k}, G(\mathbf{g}, \mathbf{p})) = \left(\frac{rf_x x}{z} - o_x + t_x, \frac{rf_y y}{z} - o_y + t_y \right)^\top, \quad (4)$$

with the radial distortion factor r and the tangential distortion offsets t_x and t_y being defined as follows [13]:

$$r = 1 + k_1 \left(\frac{x^2 + y^2}{z^2} \right) + k_2 \left(\frac{x^2 + y^2}{z^2} \right)^2 + k_3 \left(\frac{x^2 + y^2}{z^2} \right)^3, \quad (5)$$

$$t_x = 2p_1 \frac{xy}{z} + p_2 \left(\frac{3x^2 + y^2}{z^2} \right), \quad (6)$$

$$t_y = p_1 \left(\frac{x^2 + 3y^2}{z^2} \right) + 2p_2 \frac{xy}{z}. \quad (7)$$

B. Problem Formulation

The objective of this research is the direct determination of the camera positions and orientations for the captured images by exploiting measurements from a high-precision INS. In order to describe the rigid body motion $\mathbf{g}_{c,w}$ of the camera using the measured rigid body motion $\mathbf{g}_{i,w}$ of the INS (Fig. 4), the devices have to be synchronized in time and rigidly mounted. This induces that the offsets between them are

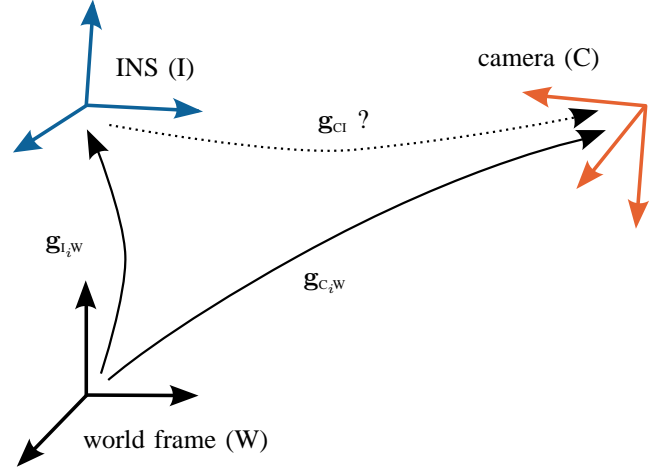


Fig. 4. The estimation of the static rigid body motion \mathbf{g}_{c_i} between the devices, enables by composition with the INS measurement $\mathbf{g}_{i,w}$ the description of the camera pose $\mathbf{g}_{c,w}$.

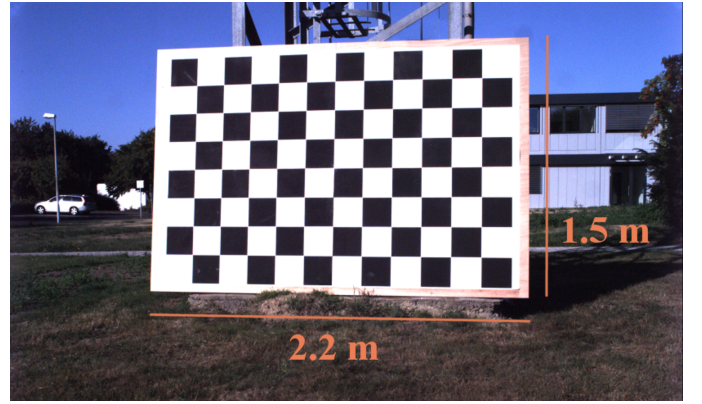


Fig. 5. The classical checkerboard calibration for a camera equipped with wide angle lens and focus at infinity requires bulky calibration patterns.

static and especially comprises that the rigid body motions describing the movements from the INS to the camera at various exposure times are constant

$$\forall k, l \in \{1, 2, \dots, n\} : \quad \mathbf{g}_{c_k l_k} \stackrel{!}{=} \mathbf{g}_{c_l l_l}, \quad (8)$$

where n is the number of images. Therefore we can simplify the notation by omitting the time indices for the rigid body motion $\mathbf{g}_{c_i} = (\psi_B, \theta_B, \phi_B, x_L, y_L, z_L)$ describing the mounting offsets as boresight and lever-arm. The composition with the measured INS movement leads to the camera motion

$$\mathbf{g}_{c,w} = \mathbf{g}_{c_i} \mathbf{g}_{i,w}. \quad (9)$$

Thus the knowledge of the mounting offsets is required to describe the camera poses with the INS measurements.

The use of a wide angle lens in conjunction with focus at infinity makes the use of standard camera calibration procedures laborious (Fig. 5). Furthermore, the estimations for the intrinsic camera parameters \mathbf{k} tend to differ slightly from a laboratory calibration, due to the climate and environmental conditions [14]. Thus it is an advantage to optimize these parameters by using images from a measurement flight.

Therefore the joint calibration of the intrinsic camera parameters \mathbf{k} in conjunction with the mounting offsets \mathbf{g}_{c_i} out of synchronized data from measurement flights is the problem considered in this work.

C. Algorithm Overview

We generate initial camera poses by combining the measured INS poses with the initial mounting offsets determined from construction drawings. The images and initial camera poses are the input for a SFM approach, which calculates a sparse 3D point cloud out of consistent image observations [15]. A refinement of these parameters is performed by the optimization of a non-linear function, which can be phrased as a graph. Our entire workflow is depicted in Fig. 6.

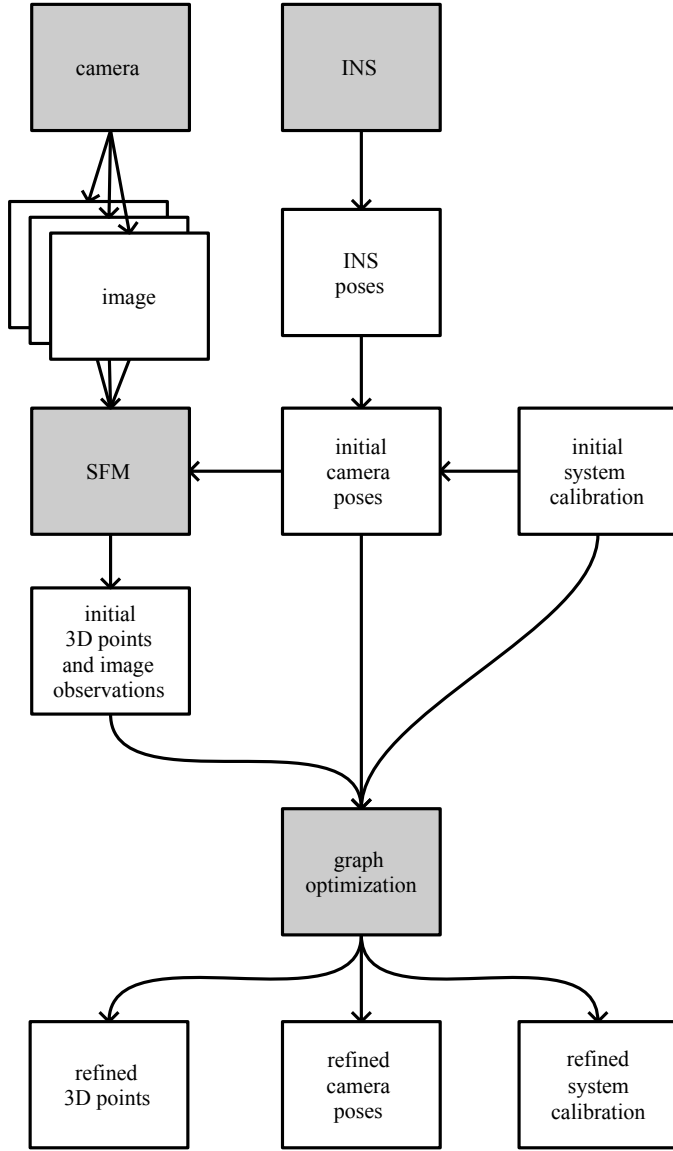


Fig. 6. Overview of the algorithm. Data from the INS and the camera are processed to determine the system calibration. The concatenation of the latter with the measured INS poses enables the very accurate determination of the camera poses.

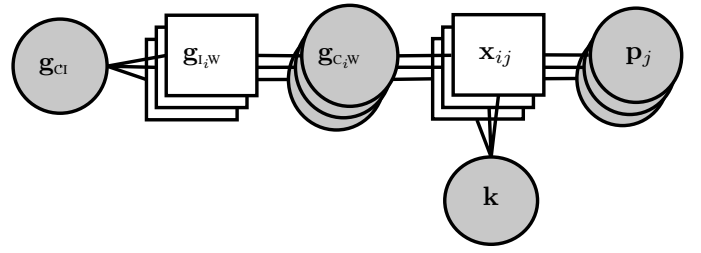


Fig. 7. The objective function of the stated problem can be illustrated by a hyper-graph. The measurements (boxes) are presented as links between the nodes concerning each multiple sets of variables (circles). For improved overview multiple state variables and measurements of the same type are visualized in a stacked view unrelated to their number of occurrence.

D. Graph Optimization

We optimize the system calibration by introducing constraints between the intrinsic camera parameters, camera poses and 3D points. The resulting non-linear error function can be represented as a graph by introducing the variables to optimize as nodes and the observations between them as edges. We determine the optimal configuration of the state variables with a graph optimizer [5]. In the following we will present our objective function and how to express it in the graph-based formulation.

The calibration parameters represented as the intrinsic camera parameters \mathbf{k} and the mounting offsets \mathbf{g}_{c_i} are added as nodes to the graph. Furthermore, we add each 3D point \mathbf{p}_j and each camera pose \mathbf{g}_{c_iw} as a node to the graph. The connection between these nodes is given by inserting observations as edges into the graph. A pixel measurement connects three different nodes, namely: a camera, a 3D point and the intrinsic camera parameters. This constraint can be realized with a hyperedge, which is able to connect an arbitrary number of nodes. The edge of an INS measurement connects the corresponding rigid body motion of the camera with the mounting offsets. A visualization of the graph is presented in Fig. 7.

Further, we have to define error functions, which measure how well measurements are described by the state variables they are connecting. The first constraint measures the error occurring from the reprojection of a 3D point into the image. The error function for this constraint can be expressed as

$$e_{ij}^x(\mathbf{k}, \mathbf{g}_{c_iw}, \mathbf{p}_j, \mathbf{x}_{ij}) = \pi(\mathbf{k}, G(\mathbf{g}_{c_iw}, \mathbf{p}_j)) - \mathbf{x}_{ij}. \quad (10)$$

The resulting error vector has dimension two and is $\mathbf{0}$ if the pixel observation \mathbf{x}_{ij} is perfectly described by the state variables. The second error function states how well the INS measurements can be described by the composition of the camera poses \mathbf{g}_{c_iw} and the mounting offsets \mathbf{g}_{c_i} as follows:

$$e_i^g(\mathbf{g}_{c_i}, \mathbf{g}_{c_iw}, \mathbf{g}_{i,w}) = (\mathbf{g}_{c_i}^{-1} \mathbf{g}_{c_iw}) \mathbf{g}_{i,w}^{-1}. \quad (11)$$

Using a minimal representation for the rigid body motions in form of three Euler angles and a three-dimensional translation, we receive a 6-dimensional error vector, which is $\mathbf{0}$ if the parameters perfectly satisfy the measurement.

Without limiting the generality, we refer to the whole state vector $[\mathbf{k}, \mathbf{g}_{c_iw}, \mathbf{p}_j, \mathbf{g}_{c_i}]$ as \mathbf{y} and formulate our objective

function as follows:

$$\min_{\mathbf{y}} \sum_i^n \sum_j^m e_{ij}^{\mathbf{x}}(\mathbf{y})^\top \Omega_{ij}^{\mathbf{x}} e_{ij}^{\mathbf{x}}(\mathbf{y}) + \sum_i^n e_i^{\mathbf{g}}(\mathbf{y})^\top \Omega_i^{\mathbf{g}} e_i^{\mathbf{g}}(\mathbf{y}), \quad (12)$$

where the information matrices $\Omega_{ij}^{\mathbf{x}}$ and $\Omega_i^{\mathbf{g}}$ respectively represent the inverse covariance of the pixel observations and INS measurements. A numerical solution of Equation (12) is computed with the Levenberg-Marquardt (LM) algorithm and therefore a good initial guess $\check{\mathbf{y}}$ of the state vector is needed. Iteratively, the first order Taylor expansion around the current guess $\check{\mathbf{y}}$ is used to approximate Equation (12) and optimize the local increments $\Delta\mathbf{y}$ by solving the resulting sparse linear system. The center for the next iteration is obtained by adding the optimized increments to the current guess. This is done by using the motion composition for the state variables represented by rigid body motions and a simple addition for the 3D points and intrinsic camera parameters. For a detailed description of the LM algorithm we refer the reader to [5], [16].

IV. EVALUATION

In this section, we present achieved results from flight experiments to validate the proposed approach. The equipped INS is based on fibre optic gyroscopes, which have a stability up to some hundredths of a degree per hour. In combination with real time kinematic enhanced GPS measurements very accurate pose information are generated. The optical system consist of a downward looking camera with 8 megapixels and a wide angle lens with a field of view of 54° in the horizontal and 42° in the vertical direction. The sensors were integrated into a payload pod, which was mounted beneath the wing of a manned ultra light airplane (Fig. 1). The performed flight course has to introduce measurements that constrain all dimensions of the calibration parameters. For our platform small movements in all axes occur even for straight and level flights, which aim at a constant heading and altitude by accomplishing immediate corrections to unintentional movements. We performed a total of four flights within two days. At an altitude of 300 meters and above, we captured two images per second at a speed of approximately 125 km/h. The flights were performed as crossing straight lines (Fig. 2). To achieve a high image overlap, we use only images within a circle of a radius of 600 meters around the central point. This results in a total number of nearly 700 images for the first two flights and 300 images for the other two.

These images were used to calculate an initial 3D point cloud of the observed area with a SFM approach under consideration of the camera poses [17]. The latter were generated by a concatenation of the INS measurements with the initial mounting offsets determined through terrestrial measurements. The output of the SFM (Fig. 8) was used as input for the graph optimization. Thereby the pixel observations were introduced as measurements with an accuracy of 1 pixel. Since no quality log files of the INS were available we considered the manufacturer information of an accuracy of 2 cm in the position, 0.04° for the yaw and 0.01° for the pitch and roll angles. Our previous work [6] reveals that the optimization of the translational part of the mounting offsets leads to a lower accuracy compared to usual terrestrial measurements. As further stated in [18], at least one GCP is needed for the

TABLE I. CALIBRATION RESULTS FROM THE CHECKERBOARD METHOD AND THE PRESENTED APPROACH

	ψ_B [$^\circ$]	θ_B [$^\circ$]	ϕ_B [$^\circ$]	f_x [pel]	f_y [pel]	o_x [pel]	o_y [pel]
checker	0.0	0.0	0.0	3334.68	3343.5	1744.32	1238.06
flight 1	0.846	0.215	-0.072	3342.89	3334.88	1730.6	1227.9
flight 2	0.816	0.205	-0.068	3343.4	3335.44	1724.04	1231.19
flight 3	0.795	0.205	-0.074	3343.73	3335.95	1725.39	1230.74
flight 4	0.805	0.193	-0.078	3346.28	3338.36	1730.62	1234.04

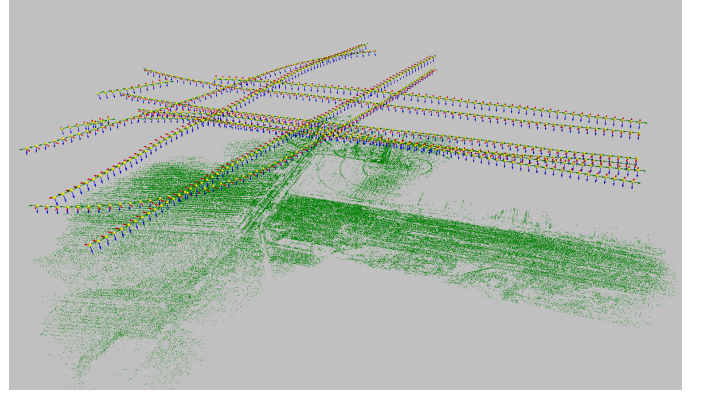


Fig. 8. Visualization of the camera poses and 3D points introduced as vertices in our graph optimization.

estimation of the vertical lever-arm and all other mounting and internal camera parameters can be decoupled from each other. Due to these observations, we fixed the lever-arm to the terrestrial measurements in our optimization.

The resulting calibration parameters show a high stability for different flights and differs from the initial checkerboard calibration (Table I). This holds especially for the distortion parameters, that we omitted here for the sake of readability. The differences between the optimized boresight angles ψ_B , θ_B and ϕ_B are very close to the stated accuracy of the INS. Small variations for the intrinsic camera parameters occur most likely due to different climate conditions during flight execution. To evaluate the accuracy of the achieved results we performed a least-square forward intersection for the pixel observations of five GCP. The image coordinates of these

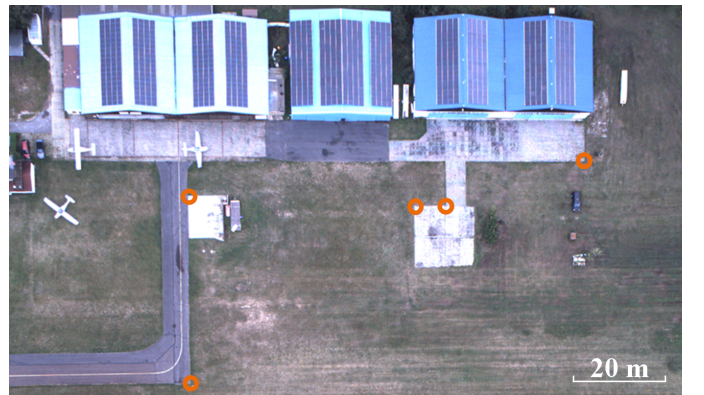


Fig. 9. We used five ground control points to verify the accuracy of our approach. These were placed at dominant image corners to allow easy manual measurement of their image coordinates (orange circles).

TABLE II. MEAN EUCLIDIAN DISTANCE BETWEEN FIVE GROUND CONTROL POINTS AND THE FORWARD INTERSECTION FOR THE INITIAL AND THE OPTIMIZED CALIBRATION (LEFT) AS WELL AS THE CALIBRATION FROM FLIGHT 1 FOR ALL FLIGHTS (RIGHT)

	init. [m]	opt. [m]	gain [factor]		opt. [m]	opt. flight 1 [m]	gain [factor]
flight 1	3.17	0.53	5.98	flight 1	0.53	0.53	1.0
flight 2	2.94	0.47	6.26	flight 2	0.47	0.55	0.85
flight 3	2.02	0.37	5.46	flight 3	0.37	0.66	0.56
flight 4	4.45	0.44	10.11	flight 4	0.44	0.58	0.76

points were measured manually (Fig. 9) and used to perform a forward intersection with the initial and optimized camera poses. This leads to 3D coordinates, which were compared to values measured with a mobile GPS-receiver. The latter stated a horizontal accuracy of about 30 cm and a vertical accuracy of about 50 cm. Our results from the forward intersection are in the same range, which shows the performance of our approach (Table II). We assume that the larger initial error of flight 4 occurs due to the range from 300 to 800 meters for the altitude, which was smaller for the other flights. Nevertheless, the results using the optimized calibration parameters for flight 4 are in the same range as for the other flights. The generation of the camera poses out of calibration results from flight 1 for all flights leads to a slightly decreasing performance (Table II). Given an altitude of 300 meters and above, the accuracies are high and clearly outperform our terrestrial calibration.

V. CONCLUSION AND FUTURE WORK

In this paper we presented a graph-based approach for the system calibration of a sensor suite consisting of a fixed mounted camera and an INS. We showed how to phrase the optimization problem as a graph and estimated the mounting offsets between the devices and the intrinsic camera parameters with a graph optimization framework. Our evaluation points out that a straightforward system calibration without the usage of GCP leads to results which show high potential for cost-saving in flight calibrations. It was shown that the calibration results can be used for consecutive flights, but the highest precision will be obtained by performing the INS-camera calibration during the mission. Compared to our terrestrial calibration, we achieved in our experiments an improvement of roughly factor six, which can be even higher for other setups due to larger angle misalignments between the devices.

Future work will investigate the proposed procedure in more detail. Furthermore, we will perform experiments with cost-efficient MEMS INS and evaluate if the approach is also usable to perform the INS-camera calibration if no RTK corrections are performed.

REFERENCES

- [1] M. Schikora, D. Bender, W. Koch, and D. Cremers, "Multi-target multi-sensor localization and tracking using passive antennas and optical sensors on UAVs," *Proc. SPIE Security + Defence*, vol. 7833, pp. 1–9, 2010.
- [2] M. Schikora, D. Bender, and W. Koch, "Airborne emitter tracking by fusing heterogeneous bearing data," in *Proc. of the 17th International Conference on Information Fusion (FUSION)*, 2014.
- [3] L. Pinto and G. Forlani, "A single step calibration procedure for IMU/GPS in aerial photogrammetry," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34 Part B3, pp. 210–219, 2002.
- [4] C. Heipke, K. Jacobsen, and H. Wegmann, "Analysis of the results of the OEEPE test 'Integrated sensor orientation'," *Integrated sensor orientation - Test report and workshop proceedings, OEEPE Official Publications No.*, vol. 43, pp. 31–49, 2002.
- [5] R. Kuemmerle and G. Grisetti, "g2o: A General Framework for Graph Optimization," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 3607–3613.
- [6] D. Bender, M. Schikora, J. Sturm, and D. Cremers, "Graph-based bundle adjustment for ins-camera calibration," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-1/W2, pp. 39–44, 2013.
- [7] R. Tsai and R. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," in *Proc. of the 4th International Symposium on Robotics Research*, 1989, pp. 287–297.
- [8] R. Horaud and F. Dornaika, "Hand-Eye Calibration," *International Journal of Robotics Research*, vol. 14, no. 3, pp. 195–210, 1995.
- [9] J. Lobo and J. Dias, "Relative Pose Calibration Between Visual and Inertial Sensors," *International Journal of Robotics Research*, vol. 26, no. 6, pp. 561–575, 2007.
- [10] F. Mirzaei and S. 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.
- [11] S. Weiss and M. Achtelik, "Versatile distributed pose estimation and sensor self-calibration for an autonomous mav," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 31–38.
- [12] M. Cramer, D. Stallmann, and N. Haala, "Direct georeferencing using GPS/inertial exterior orientations for photogrammetric applications," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 33 Part B3, pp. 198–205, 2000.
- [13] D. C. Brown, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.
- [14] K. Jacobsen, "Aspects of handling image orientation by direct sensor orientation," in *Proc. of the ASPRS Annual Convention*, 2001.
- [15] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [16] M. Lourakis and A. Argyros, "SBA: A software package for generic sparse bundle adjustment," *ACM Transactions on Mathematical Software*, vol. 36, no. 1, pp. 1–30, 2009.
- [17] C. Wu, "Towards linear-time incremental structure from motion," in *International Conference on 3D Vision*, 2013, pp. 127–134.
- [18] A. Kersting, A. Habib, and K. Bang, "Mounting Parameters Calibration of GPS/INS-Assisted Photogrammetric Systems," in *International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM)*, 2011, pp. 1–6.