

# MOUNTING CALIBRATION OF A MULTI-VIEW CAMERA SYSTEM ON A UAV PLATFORM

M. Mohammadi<sup>1,\*</sup>, A. Khami<sup>2</sup>, F. Rottensteiner<sup>1</sup>, I. Neumann<sup>2</sup>, C. Heipke<sup>1</sup>

<sup>1</sup> Institute of Photogrammetry and GeoInformation, Leibniz University Hannover, Germany  
(mohammadi, rottensteiner, heipke)@ipi.uni-hannover.de

<sup>2</sup> Geodetic Institute Hannover, Leibniz University Hannover, Germany  
(khami, neumann)@gih.uni-hannover.de

Commission I, WG I/9

**KEY WORDS:** Mounting calibration, multi-camera system, non-overlapping views, bundle block adjustment, UAV

## ABSTRACT:

Multi-view camera systems are used more and more frequently for applications in close-range photogrammetry, engineering geodesy and autonomous navigation, since they can cover a large portion of the environment and are considerably cheaper than alternative sensors such as laser scanners. In many cases, the cameras do not have overlapping fields of view. In this paper, we report on the development of such a system mounted on a rigid aluminium platform, and focus on its geometric system calibration. We present an approach for estimating the exterior orientation of such a multi-camera system based on bundle adjustment. We use a static environment with ground control points, which are related to the platform via a laser tracker. In the experimental part, the precision and partly accuracy that can be achieved in different scenarios is investigated. While we show that the accuracy potential of the platform is very high, the mounting calibration parameters are not necessarily precise enough to be used as constant values after calibration. However, this disadvantage can be mitigated by using those parameters as observations and refining them on-the-job.

## 1. INTRODUCTION

### 1.1 Motivation and Goal

In many fields, such as autonomous driving or crash tests, a large field of view is necessary for cameras observing the environment, which can be achieved by using a multi-view camera system. Often, it is not possible to guarantee overlapping views due to cost, energy or physical capacity reasons. In order to be able to capture accurate and reliable measurements using a multi-sensor system (MSS), a calibration of the whole system is needed as a first step. Assuming a rigid platform, this exterior sensor calibration is achieved by geometrically referencing and synchronizing the sensors with respect to each other. In other words, the 6 degrees of freedom (6DoF) transformation, consisting of three translations and three rotations, and information about time synchronization for all sensors with respect to each other must be determined.

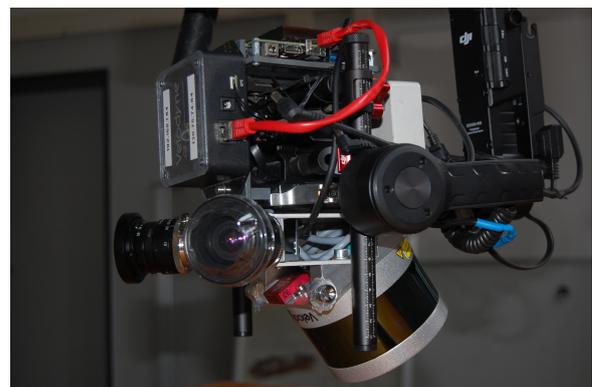
For the calibration, we define a so called platform coordinate system (CS) and find the 6DoF of the sensors with respect to this CS. A least squares estimation of the unknowns is performed based on a Gauß-Markov model (GMM). The adjustment model for various scenarios is described in detail and is tested in a set of experiments. Specifically, we study the precision of the results in terms of the variance-covariance matrix of the unknowns, when varying the number of cameras, the number of different positions of the platform in the lab (we call those positions "stations" in this paper), and the effect of considering the platform CS in the adjustment.

### 1.2 Multi-Sensor System

Our MSS comprises two cameras, a laser scanner and a GPS/IMU positioning system; it is installed on a rigid aluminium platform

\*Corresponding author

to be mounted on a hexacopter UAV (unmanned aerial vehicle) for surveying applications. It is advantageous to carry out the calibration with the platform being attached to the gimbal of the UAV due to complicated installing and un-installing procedures. In this paper, we are only concerned with the geometric calibration of the cameras. However, a similar procedure can be applied to laser scanners as well, e.g. (Hartmann et al., 2017).



**Figure 1.** The multi-sensor platform attached to the gimbal.; the laser scanner and IMU are not used in this paper.

Fig. 1 shows the sensor platform. Note that the two cameras are attached with viewing directions differing by approximately  $100^\circ$ , which results in a very narrow overlap between the two viewing cones only. We use *Basler acA2500 – 60uc* cameras with a sensor having  $2592 \times 2048$  pixels and a pixel size of  $4.8 \mu m \times 4.8 \mu m$ . The lens has a focal length of  $6 mm$ . The interior orientation of the cameras is determined in a separate pre-processing step and is considered to be constant.

In this multi-sensor system, time synchronization is achieved using a hardware trigger based on the GPS time signal, which is

then multiplied and provided for the cameras with a frequency of 10 Hz. This, however, is not a criterion to be fulfilled in the calibration procedure described here, because the calibration is performed in a static environment and therefore does not require time-synchronized data between different sensors.

## 2. RELATED WORK

There exist numerous approaches in the literature to determine the mounting calibration of multi-camera systems with non-overlapping fields of view. Xia et al. (2018) summarize several of these methods and divide them into different categories. The most important ones are discussed in the following paragraphs.

One category uses measuring devices with superior accuracy such as a laser tracker or theodolite. Kitahara et al. (2001) performed a large-scale space camera calibration. In order to realize a large calibration board, a small calibration checkerboard was scanned at various locations in the room using a laser scanner. Ortega et al. (2009) calibrated an outdoor distributed camera network using a laser range finder (LRF). Firstly, by registering the LRF data to an aerial image of the site, the coarse location, orientation and field of view (fov) of the cameras were estimated. Also, having the camera fov, the corresponding LRF data is found and a plane segmentation is done. Then, 3D lines are computed as intersections of perpendicular planes. In a second step the initial camera parameters were refined in a semi-automatic way by matching 2D and 3D features in a non-linear optimization.

The second category uses large-scale calibration targets. In (Liu et al., 2011a), the calibration field is divided into sub-targets for the non-overlapping views. The 6DoF between the sub-targets need not be known as long as the whole system is rigid. A minimum of four images of the target are acquired from different angles and the transformation parameters between the cameras and the global coordinate system are estimated, where the global coordinate system is based on one camera being selected as the reference sensor. In (Liu et al., 2011b), a calibration method applicable in a large area or narrow space based on 1D targets is proposed. In this work, each camera is paired with a neighboring camera and their relative rotations are calculated. Then, based on the feature point distances on the 1D target, the translation vector is determined. Strauss et al. (2014) present an approach which combines the calibration of interior and exterior orientation parameters of the cameras using coded targets and image sequences. Non-linear least squares optimization is used in the estimation procedure.

In the third category planar mirrors are employed. In this way Lébraly et al. (2010b) create overlap between different camera views. Sturm and Bonfort (2006) use a planar mirror to compute the position and orientation of one camera without direct view of the target placed side by side to the camera. Xu et al. (2015) use a similar approach by setting up one or more mirrors and calibrating cameras by solving the transformations between the cameras, including the mirrored virtual camera.

The fourth category is based on motion models. It uses image sequences to track movements and establish corresponding relationships for different fields of view. As these methods do not need any external devices, they are more cost efficient and flexible. Caspi and Irani (2002) calibrate a set of two non-overlapping cameras with a short baseline installed rigidly with respect to each other. The two sets of image sequences are

then aligned using the assumption of similar changes over time within the two sequences. It is assumed that the cameras share the same center of projection, which restricts the achievable accuracy. Esquivel et al. (2007) process image sequences of non-overlapping cameras by a structure from motion (SfM; called structure and motion, SAM, in that paper) algorithm, which estimates the positions and orientations of the cameras. The camera rig is again assumed to be rigid; however, a more general case is solved, not requiring the cameras to share a common projection center. The authors calculate the orientation, position and scale of each camera separately in an iterative adjustment procedure. In a different and mobile application, Pagel (2012) uses an approach to solve challenges of hand-eye calibration. Lébraly et al. (2010a) propose a calibration method for non-overlapping cameras aboard a vehicle observing a static scene and determine all unknowns using a bundle adjustment. Micusik (2011) proposes a solution to the relative 6DoF problem of non-overlapping surveillance cameras observing a moving object (person). The lack of conjugate points between two images is compensated by assumptions that the person is moving linearly and to a known gravity vector for both cameras. Then gravity vector can be estimated e.g. by an inertial sensor. These assumptions lead to a quadratic eigenvalue problem that is solved to obtain the unknown parameters.

The other categories are based on laser projection, e.g. (Liu et al., 2012, 2013), and visual measuring instruments, e.g. (Dong et al., 2016; Birdal et al., 2016; Gong et al., 2017). Further literature can also be found on the subject. A SLAM-based automatic exterior calibration technique is presented by Carrera et al. (2011) in a two and four camera configuration, where the camera rig is fixed. A final relevant paper is (Borkar et al., 2011), which presents a method to find the positions and angles of two cameras relative to a reference point by transforming images using Inverse Perspective Mapping (IPM) and acquiring bird-eye views.

In our work we take advantage of an accurate external measuring device, namely a laser tracker, and a static environment, thus our method falls into category 1. As the calibration of the platform is part of the pre-processing stage of our research, for which a high accuracy is desired, it is established in a lab environment. While this paper deals only with the cameras, a joint system calibration of the cameras and laser scanner of our MSS is planned as one of the next steps, initially again in a controlled environment.

## 3. CALIBRATION PROCEDURE

In our work, the exterior calibration of the multi-sensor system, i.e. a 6DoF transformation for each sensor with respect to a common 3D platform CS, is estimated via non-linear least squares adjustment. The six parameters are also called mounting calibration parameters, or mounting calibration for short. In this paper, we restrict ourselves to estimating the mounting calibration of the two cameras of our system, for which the interior orientation (calibrated focal length, image coordinates of the principal point and lens distortion parameters) are assumed to be known. Considering the cameras separately, the task amounts to a spatial resection with respect to the platform CS. As mentioned before, the cameras only have a very small common field of view, therefore stereo models do not exist.

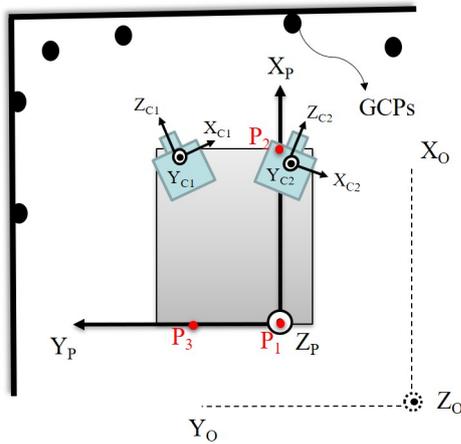
The platform CS is realized via three signalized points at three corners of the rigid platform. One point serves as origin of the

platform CS, the second point determines the direction of one of the coordinate axes (here  $x$ ). A plane is then defined containing this coordinate axis and the third point (and thus also the second axis ( $y$ ), which is perpendicular to the first axis), finally the plane normal points into the direction of the third coordinate axis. In order to determine the mounting calibration, we position the platform in a lab, where the cameras can see a number of targets placed in different heights on the lab walls. These targets, which were measured by the laser tracker, serve as ground control points (GCPs), cf. fig. 2.

### 3.1 Relationship between Coordinate Systems

The relationship between the coordinate systems used in the calibration is described in the following.

1. Placing a laser tracker in the lab, we can determine the coordinates of the GCPs and the three platform points ( $P_1, P_2, P_3$ ) in the CS of the laser tracker, which we call the *object CS*. As a result, the coordinates of all GCPs and of the three points are available in this system. Fig. 2 shows the GCPs, the three points on the platform, as well as the object, platform and camera coordinate systems.



**Figure 2.** Platform and GCPs in the object ( $X_O, Y_O, Z_O$ ), platform ( $X_P, Y_P, Z_P$ ) and camera coordinate systems ( $X_{C_i}, Y_{C_i}, Z_{C_i}$  with  $i \in \{1, 2\}$ ).

2. Using the three points on the platform as identical points in object and platform coordinate systems, the parameters of a 6DoF transformation between the two systems can be computed, and the object coordinates of the GCPs can be transformed into the platform CS as shown in eq. 1.

$$\mathbf{X}_{\text{platform}} = \mathbf{R}_{\text{OP}}^T (\mathbf{X}_{\text{object}} - \mathbf{X}_{\text{OP}}) \quad (1)$$

where

$\mathbf{X}_{\text{object}} = (X_O, Y_O, Z_O)^T$ : 3D coordinate vector of a point in the object CS

$\mathbf{X}_{\text{platform}} = (X_P, Y_P, Z_P)^T$ : 3D coordinate vector of a point in the platform CS

$\mathbf{R}_{\text{OP}}$ : Rotation matrix to rotate from the platform CS into the object CS

$\mathbf{X}_{\text{OP}}$ : 3D translation vector between the origins of the object and platform CS.

3. As mentioned earlier, the mounting calibration (exterior orientation of the cameras in the platform CS) can be determined via spatial resection, based on the GCP coordinates in the platform system. In order to do so, we can use

another 6DoF transformation and the standard collinearity equations, see eqs. 2 and 3. Here, the camera system is a right-handed system, where the  $y$ -axis points downwards and the  $z$ -axis in the viewing direction.

$$\mathbf{X}_{\text{camera}} = \mathbf{R}_{\text{PM}}^T (\mathbf{X}_{\text{platform}} - \mathbf{X}_{\text{PM}}) \quad (2)$$

where

$\mathbf{X}_{\text{camera}} = (X_C, Y_C, Z_C)^T$ : 3D coordinate vector of a point in the camera CS

$\mathbf{R}_{\text{PM}}$ : Rotation matrix to rotate from the camera CS into the platform CS (camera viewing direction, part of the mounting calibration)

$\mathbf{X}_{\text{PM}}$ : 3D translation vector between platform and camera CS (camera projection centre, part of mounting calibration).

$$\begin{aligned} x &= x_0 - c \frac{X_C}{Z_C} + \Delta x = f_1(\mathbf{X}_{\text{camera}}) \\ y &= y_0 - c \frac{Y_C}{Z_C} + \Delta y = f_2(\mathbf{X}_{\text{camera}}) \end{aligned} \quad (3)$$

where

$x, y$ : Image coordinates of a point

$X_C, Y_C, Z_C$ : Point coordinates in the camera CS

$x_0, y_0, c$ : Parameters of interior orientation of the camera (principal point and calibrated focal length)

$\Delta x = f_{\Delta x}(x, y, k_1, k_2, k_3, p_1, p_2)$

$\Delta y = f_{\Delta y}(x, y, k_1, k_2, k_3, p_1, p_2)$

$\Delta x$  and  $\Delta y$  are the lens distortion corrections,  $k_1, k_2, k_3$  are radial distortion parameters, and  $p_1, p_2$  are tangential distortion parameters of the camera. The distortion model of OpenCV is used in this context<sup>1</sup>.

For  $\mathbf{R}_{\text{OP}}$  the angles  $\omega, \phi, \kappa$  were used according to the representation of rotations often found in photogrammetry, however, in the rotation  $\mathbf{R}_{\text{PM}}$  of the mounting calibration, in order to avoid a possible gimbal lock the angles  $\alpha, \zeta, \kappa$  are used to represent the rotation from the platform to the camera CS. Here,  $\alpha$  is the rotation around the  $Z$  axis,  $\zeta$  around the (rotated)  $Y$  axis and  $\kappa$  is a second rotation around the (rotated)  $Z$  axis. These three angles define the camera viewing direction ( $\alpha, \zeta$ ) and rotation around the viewing direction ( $\kappa$ ) with respect to the platform CS.

Pixel coordinates of all visible GCPs are observed in the images. To ensure a geometrically stable solution, the targets should be well-distributed in the images. To determine initial values for the mounting parameters used in the subsequent adjustment, one can first determine approximate values for the pixel coordinates of four targets, preferably lying on the four corners of the image, either manually or automatically. In our case, due to the small number of points needed, this step was carried out manually. Spatial resection can then be carried out based on the Müller-Killian algorithm, discussed in (Müller, 1925) and (Killian, 1955).

Having thus determined initial mounting parameters for the camera and the 3D coordinates of all targets in the platform CS, their initial pixel coordinates can be determined by projecting the 3D coordinates into image space. Pixel positions can then be refined using any method for precise corner point location (we use the one described in (Förstner and Gülch, 1987)). Subsequently, a least squares

<sup>1</sup>[https://docs.opencv.org/4.5.2/d9/d0c/group\\_calib3d.html](https://docs.opencv.org/4.5.2/d9/d0c/group_calib3d.html)

adjustment is carried out to estimate the mounting parameters (see section 3.2 for details of the used functional model).

- Obviously, the whole procedure can be repeated various times by moving the platform to another station in the lab. Such additional measurements offer a means to check the computed mounting calibration and to increase the precision and reliability of the unknowns. We have carried out measurements from two different stations and have processed the data separately as well as simultaneously (see section 4 for details).

### 3.2 Adjustment Approach

The determination of the mounting parameters is based on an extended bundle adjustment in the form of a GMM. Non-linear observation equations are derived by introducing eq. 1 into eq. 2, and the result into eq. 3, giving eq. 4.

$$\begin{aligned} x &= f_1(\mathbf{X}_{\text{camera}}) = f_1(\mathbf{R}_{\text{PM}}^T(\mathbf{X}_{\text{platform}} - \mathbf{X}_{\text{PM}})) \\ &= f_1(\mathbf{R}_{\text{PM}}^T(\mathbf{R}_{\text{OP}}^T(\mathbf{X}_{\text{object}} - \mathbf{X}_{\text{OP}}) - \mathbf{X}_{\text{PM}})) \end{aligned} \quad (4)$$

$$\begin{aligned} y &= f_2(\mathbf{X}_{\text{camera}}) = f_2(\mathbf{R}_{\text{PM}}^T(\mathbf{X}_{\text{platform}} - \mathbf{X}_{\text{PM}})) \\ &= f_2(\mathbf{R}_{\text{PM}}^T(\mathbf{R}_{\text{OP}}^T(\mathbf{X}_{\text{object}} - \mathbf{X}_{\text{OP}}) - \mathbf{X}_{\text{PM}})) \end{aligned}$$

We use direct observations for the GCP coordinates in the object CS and for the object coordinates of the three points on the platform. For the latter, we also introduce their coordinates in the platform system in order to be able to position the platform CS with respect to the object CS. In order not to have to define the platform coordinates with superior precision, we select the weights for the observations of the three platform points in the platform CS to allow for a determination of the transformation without constraints: the three coordinates of Point  $P_1$  obtain very high weights, the Y and Z coordinates of  $P_2$  and the Z coordinate of  $P_3$  also, whereas the other coordinates obtain low weights. In this way the six transformation parameters can be determined and the other coordinates of these points do not affect the parameters in a negative way.

The resulting non-linear observation equations are shown in eq. 5 for the image coordinates, eqs. 6 and 7 for the 3D coordinates of the GCPs and the three platform points in the object CS, and eq. 8 for the points on the platform in the platform CS respectively.  $v$  denotes the residuals and  $k$  is the number of GCPs which are visible in the images.

$$\begin{aligned} \hat{v}_x &= f_1(\hat{\mathbf{R}}_{\text{PM}}^T(\hat{\mathbf{R}}_{\text{OP}}^T(\hat{\mathbf{X}}_{\text{object}} - \hat{\mathbf{X}}_{\text{OP}})) - \hat{\mathbf{X}}_{\text{PM}}) - x \\ \hat{v}_y &= f_2(\hat{\mathbf{R}}_{\text{PM}}^T(\hat{\mathbf{R}}_{\text{OP}}^T(\hat{\mathbf{X}}_{\text{object}} - \hat{\mathbf{X}}_{\text{OP}})) - \hat{\mathbf{X}}_{\text{PM}}) - y \end{aligned} \quad (5)$$

$$\hat{v}_{\text{GCP}_j, \text{object}} = \hat{\mathbf{X}}_{\text{GCP}_j, \text{object}} - \mathbf{X}_{\text{GCP}_j, \text{object}}; j \in \{1, \dots, k\} \quad (6)$$

$$\hat{v}_{\text{P}_i, \text{object}} = \hat{\mathbf{X}}_{\text{P}_i, \text{object}} - \mathbf{X}_{\text{P}_i, \text{object}}; i \in \{1, 2, 3\} \quad (7)$$

$$\begin{aligned} \hat{v}_{\text{P}_i, \text{platform}} &= \hat{\mathbf{R}}_{\text{OP}}^T(\hat{\mathbf{X}}_{\text{P}_i, \text{object}} - \hat{\mathbf{X}}_{\text{OP}}) - \mathbf{X}_{\text{P}_i, \text{platform}} \\ & \quad ; i \in \{1, 2, 3\} \end{aligned} \quad (8)$$

We thus use the following observations:

- Image coordinates  $(x, y)$  of GCPs
- 3D coordinates of GCP targets in the object CS  $(X_{\text{GCP}_j, \text{object}}, Y_{\text{GCP}_j, \text{object}}, Z_{\text{GCP}_j, \text{object}}); j \in \{1, \dots, k\}$
- 3D coordinates of platform points  $P_i$  in the object CS  $(X_{\text{P}_i, \text{object}}, Y_{\text{P}_i, \text{object}}, Z_{\text{P}_i, \text{object}}); i \in \{1, 2, 3\}$
- 3D coordinates of platform points  $P_i$  in the platform CS  $(X_{\text{P}_i, \text{platform}}, Y_{\text{P}_i, \text{platform}}, Z_{\text{P}_i, \text{platform}}); i \in \{1, 2, 3\}$

to determine the following unknowns:

- Camera mounting calibration  $(X_0, Y_0, Z_0, \alpha, \zeta, \kappa)$  as elements of  $\mathbf{X}_{\text{PM}}$  and  $\mathbf{R}_{\text{PM}}$  for each camera (see eq. 2)
- 6 parameters for the transformation between the object and the platform CS contained in  $\mathbf{X}_{\text{OP}}$  and  $\mathbf{R}_{\text{OP}}$  for each station (see eq. 1)
- 3D coordinates of GCP targets in the object CS  $(X_{\text{GCP}_j, \text{object}}, Y_{\text{GCP}_j, \text{object}}, Z_{\text{GCP}_j, \text{object}}); j \in \{1, \dots, k\}$
- 3D coordinates of platform points  $P_i$  in the object CS for each station  $(X_{\text{P}_i, \text{object}}, Y_{\text{P}_i, \text{object}}, Z_{\text{P}_i, \text{object}}); i \in \{1, 2, 3\}$

In the stochastic model, the variance-covariance matrix is assumed to be diagonal. For all observations, appropriate variances, and thus weights, are selected (numerical values are given in Section 4). The solution is based on the well known formulae of GMM in the nonlinear case, see eq. 9. Here,  $\hat{\Delta \mathbf{x}}$  is the difference between the vector of estimated unknowns in the last two iterations,  $\Delta \mathbf{l}$  is the difference between the observation vector and the observations which are acquired by the initial assumptions of the parameters. The Jacobi matrix  $\mathbf{A}$  contains the derivatives of the observation equations with respect to the unknowns.

$$\hat{\mathbf{v}} = \mathbf{A} \hat{\Delta \mathbf{x}} - \Delta \mathbf{l} \quad (9)$$

The adjustment is carried out iteratively according to the standard formulae. The stopping criterion is based on the relative change rate of the sum of the weighted squared residuals  $\Omega$  between the previous and the current iteration:

$$\frac{\Omega_{\text{prev}} - \Omega_{\text{curr}}}{\Omega_{\text{prev}} + \Omega_{\text{curr}}} < 1e-4 \quad (10)$$

## 4. EXPERIMENTS

### 4.1 Goals

In this section, we report the precision of the adjustment results, and in particular of the mounting calibration parameters, in different scenarios. Specifically, we vary the number of cameras, the number of stations, and the effect of introducing the platform CS into the adjustment.

The background of this study is the fact that the platform must be small to be flown on a UAV, and thus the three points defining the platform CS are located within a few centimeters of each other (10 - 15 *cm*). For the transformation between the object and the platform CS, these three points serve as identical points. While the platform coordinates of these points can be defined as mentioned above, the object coordinates are determined using a laser tracker from a distance of several meters. As a result, the elements of the rotation matrix  $\mathbf{R}_{OP}$  between object and platform system can only be determined with a limited precision.

### 4.2 Data Acquisition

Data acquisition for the exterior calibration was done in a lab with reference targets on the walls and ceiling which serve as GCPs. The object coordinates of these targets as well as those of the three platform points were accurately measured using a laser tracker. Initial values for the transformation parameters between object and platform CS were determined using the commercial software SpatialAnalyzer<sup>2</sup>. The platform coordinates of the three points on the platform are shown in table 1; the six coordinates relevant for defining the 6DoF of the platform system are all set to zero, the other coordinates are derived from distances computed from the object coordinates.

	$X_P$ [m]	$Y_P$ [m]	$Z_P$ [m]
$P_1$	0.000	0.000	0.000
$P_2$	0.149	0.000	0.000
$P_3$	0.024	0.127	0.000

**Table 1.** Platform coordinates of the three platform points.

The UAV platform was placed in the middle of the room with the two cameras tilted about 30° upwards. The focal length of 6 *mm* and the relatively long distance to the targets allow for a wide field of view and a minimum of 10 targets to be seen in each image (two of which appear in both cameras). After capturing two images of the scene in station 1, the platform was slightly rotated and moved (approximately 20° and 15 *cm*, respectively) and another two images were taken. The images can be seen in Fig. 3 for the left camera and in figure 4 for the right camera.

One pixel in an average distance for camera to objects of 5 *m* covers an area of 4x4 *mm*<sup>2</sup> in object space. The actual distance of the cameras to the targets varied from 3 to 6 *m*. Targets closer to the platform cover an area of up to 90x90 *px*, while the (few) targets further away were depicted in slanted view and with as few as 20x30 *px*. In addition, a number of targets were also viewed from a slanted direction, which might increase the standard deviation of the corresponding image coordinates.

<sup>2</sup><https://www.hexagonmi.com/en-GB/products/software/spatialanalyzer>



**Figure 3.** Images of the left camera in station 1 (left) and 2 (right).



**Figure 4.** Images of the right camera in station 1 (left) and 2 (right).

### 4.3 Assumptions about the Stochastic Model

Uncorrelated observations and a priori standard deviations for each observation type are assumed as follows.

$$\begin{aligned}\sigma_{xy} &= 0.8 \text{ px} \\ \sigma_{(X,Y,Z)_{GCP_j,object}} &= 0.3 \text{ mm} \\ \sigma_{(X,Y,Z)_{P_i,object}} &= 0.3 \text{ mm}\end{aligned}$$

For the platform coordinates of the three platform points, this high weight was only used for the six coordinates defining the transformation between the object and platform CS (i.e., those with a value of 0.000 in table 1, while a very low weight was chosen for the other three coordinates. Furthermore, the a priori standard deviation of the weight unit is set to 1.

### 4.4 Results

As mentioned before, we compare the precision of the adjustment obtained in different scenarios:

1. Cases 1a and 1b: the images of one camera only, but of both stations are used (1a refers to the left camera, 1b to right camera);
2. Case 2: the images of both cameras and both stations are employed simultaneously in one adjustment;
3. Case 3: same as case 2, but the transformation from the object to the platform system is assumed to be constant and error-free. For this experiment we transformed the object GCP coordinates to the platform CS using the parameters determined via SpatialAnalyser, and then considered the platform CS to be the object CS. Consequently, the three points on the platform were not needed any longer, and we dropped equations 7 and 8 from the adjustment.
4. Case 4: a special variant of case 2, where only one station is used.
5. Case 5: a special variant of case 3, where only one station is used. In other words, case 5 is similar to case 4, with the difference that the object CS is the platform CS.

Case	RMS val. of unit weight	RMS value of reprojection error per image [px]				Max. residual [px]	Precision of 3D coordinates [mm]		
		Station 1 (x, y)		Station 2 (x, y)			$\sigma_x$	$\sigma_y$	$\sigma_z$
1a	0.30	0.48	0.30	0.20	0.35	1.45	0.08	0.08	0.08
1b	0.83	1.00	0.72	1.23	0.89	2.32	0.22	0.21	0.23
2	0.63	0.51	0.32	0.19	0.41	2.26	0.15	0.14	0.15
		1.00	0.73	1.23	0.89				
3	1.08	0.64	0.55	0.59	0.72	2.39	0.32	0.32	0.32
		0.87	0.85	1.19	0.81				
4	0.55	0.42	0.37	-	-	2.16	0.13	0.13	0.13
		0.97	0.66	-	-				
5	0.91	0.41	0.36	-	-	2.15	0.27	0.27	0.27
		0.97	0.65	-	-				

**Table 2.** Adjustment results. The first columns show the image-based data, namely the unit weight error of the adjustment ( $\sqrt{\Omega_{curr}/(n-u)}$ ; where  $n$  is the number of observations and  $u$  the number of unknowns), the reprojection errors, for each image separately and for all targets, gathered as one value by the root mean square (RMS), and the maximum residual (highest reprojection error in image plane) seen in the adjustment (in  $x$  or  $y$ , however not specified in table for each case). Upper values in the reprojection errors of different stations for cases 2 to 5 show results of the left camera, whereas lower values denote those of the right camera.

The last three columns represent the 3D coordinate precisions of the GCPs and platform points as RMS of their uncertainties.

	Left camera		Right camera	
	$X_{PM}/Y_{PM}/Z_{PM}[m]$	$\alpha/\zeta/\kappa[grad]$	$X_{PM}/Y_{PM}/Z_{PM}[m]$	$\alpha/\zeta/\kappa[grad]$
Case 1	0.180±0.000	-164.000±0.041	0.156±0.001	145.003±0.114
	0.112±0.001	99.948±0.048	-0.046±0.001	99.397±0.097
	-0.015±0.000	99.713±0.036	-0.019±0.001	99.744±0.136
Case 2	0.180±0.001	-163.998±0.087	0.155±0.001	145.005±0.086
	0.112±0.001	99.947±0.101	-0.046±0.001	99.396±0.074
	-0.015±0.001	99.710±0.075	-0.019±0.001	99.744±0.103
Case 3	0.183±0.002	-163.941±0.034	0.156±0.002	144.984±0.026
	0.109±0.002	99.921±0.024	-0.045±0.001	99.385±0.029
	-0.015±0.002	99.680±0.016	-0.018±0.002	99.737±0.017
Case 4	0.181±0.001	-163.955±0.108	0.155±0.001	144.996±0.108
	0.111±0.001	99.908±0.127	-0.044±0.001	99.356±0.093
	-0.014±0.001	99.676±0.094	-0.018±0.001	99.781±0.129
Case 5	0.181±0.002	-163.955±0.035	0.155±0.002	144.996±0.035
	0.111±0.002	99.908±0.025	-0.044±0.001	99.356±0.037
	-0.014±0.002	99.676±0.017	-0.018±0.002	99.781±0.025

**Table 3.** Computed mounting calibration parameters and related precision.

The results are presented in tables 2 and 3, in which the precision values are determined on the basis of the estimated RMS error of the unit weight and the inverse of the normal equation matrix. They can be interpreted as follows:

1. The a posteriori RMS value of unit weight is smaller than 1 in all cases except case 3, where it is 1.08. Thus, the introduced standard deviations (see section 4.3) are somewhat too large. However, the results shown in table 2 are acceptable, and fine-tuning these standard deviations, e.g.

via variance-covariance estimation, is beyond the scope of this paper.

2. A similar behavior is found in the RMS values of the reprojection errors, which in most cases are smaller than the assumed value of 0.8 pixel. The general range of the reprojection errors is not surprising considering the employed tool for image coordinate determination; a better precision can be reached with dedicated measurement algorithms. The values for the left camera (case 1a and upper line in cases 2 - 5) are smaller than those for the right one

(case 1b and lower line in cases 2-5). A possible reason is that the interior orientation parameters of the right cameras might have changed slightly between interior calibration and mounting calibration procedures. The reported values are rather consistent across the different cases, perhaps with the exception of case 3, for which the left camera has slightly larger values. This can partly be explained by the fact that any inconsistencies of the adjustment can no longer be absorbed by the transformation parameters from object to platform CS, and are thus more visible in the image coordinates.

3. The maximum residuals amount to a little more than 2 pixels. The maximum residual is located, in all cases except case 1a, in the right image and belongs to the small, slanted target, which is not surprising. As the values discussed so far are image-based, major differences between the different cases were not expected, and are not visible in the results.
4. The estimated precision of the 3D coordinates varies (last column of table 2), but this variation is mainly due to the change of the variance factor of the unit weight. The largest part of it can be explained by a pure scaling of the inverse of the normal equation matrix by the RMS value of unit weight due to the high accuracy of the laser tracker measurements in comparison to the camera measurements in object space.
5. The estimated values of the mounting calibration parameters are nearly all consistent in all cases, taking into account their precision.
6. While the standard deviations for the components of the translation vector between platform and camera are nearly identical in all cases, differences can be found in the angular values. The difference of the standard deviations in case 1 between the left and right camera can be explained again by the difference of the RMS value of the unit weight (see also point 4). Comparing case 1 to case 2 we find contradictory results: for the right camera, case 2 delivers more precise results, which is expected, as the redundancy for the mounting calibration parameters is higher, the inverse is true for the left camera. At this stage the reason is not clear, but a comparison with table 2 suggests that, perhaps, the results of case 1a are a little too optimistic.
7. The potential of the MSS can be seen when looking at the precision for the rotation angles of case 3. Here, compared to cases 1 and 2, the achieved standard deviations are smaller by a factor of 3 to 5, demonstrating the high accuracy potential of the developed system. As a consequence, when using the system for high accuracy applications, the rotation angles of the mounting calibration should not be considered as constant values, but as observations with the precision obtained in the calibration phase, similar to a self-calibration scenario. Consequently, some scene information must be available to improve the precision of these mounting angles.
8. Case 4, where only a single station was used, shows a somewhat higher standard deviation compared to case 2; which is a consequence of having fewer observations and therefore a lower redundancy. However, the differences are not significant; therefore introducing more than two stations is not expected to yield much improvement in terms

of precision; however, in that case, detection of blunders can be improved and the setup will be more reliable.

In general, case 5 shows the same behavior to case 4 as case 3 does to case 2, which is expected.

## 5. CONCLUSIONS AND OUTLOOK

In this paper, we have presented an approach to determine the mounting calibration of an MSS with a focus on two cameras with non-overlapping fields of view. To do so, we have employed a static lab environment with GCPs and have computed the mounting calibration relative to a platform assumed to be rigid. Hence, the 6DoF of one sensor relative to another can be calculated easily, as long as it remains constant over time.

Being based on single camera resection, this approach works with any number of cameras. If cameras have a common field of view, the resulting ray intersections are automatically taken into account, as the approach is based on photogrammetric bundle adjustment.

The results show that while image coordinate determination still has a potential for improvement, the calibration procedure works well, and the obtained precision for the calibration parameters have reached the expected values. However, as they are determined indirectly using a laser tracker, additional object space information should be available for the mounting angles in order to be able to exploit the full accuracy potential of the multi-sensor system. A general benefit of using a laser tracker is the high accuracy with a validation of the results by reprojection. It can therefore be assumed that the estimated precision is close to the accuracy of the calibration.

Future work will address the common calibration of the cameras and the laser scanner in both, static and dynamic environments. Also, we will carry out georeferencing experiments of the platform based on building models similar to (Unger, 2020), but with input from the cameras and the laser scanner.

## ACKNOWLEDGEMENTS

This work was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) HE1822/37-1 for Institute of Photogrammetry and GeoInformation and NE1453/5-1 for Geodetic Institute Hannover. The platform was developed in cooperation with Prof. Dr. Stephan Schulz and his team at Dept. of Mechanical Engineering and Production Management, HAW Hamburg.

## References

- Birdal, T., Bala, E., Eren, T., Ilic, S., 2016. Online inspection of 3d parts via a locally overlapping camera network. *IEEE Winter Conference on Applications of Computer Vision (WACV)*.
- Borkar, A., Hayes, M., Smith, M. T., 2011. A non overlapping camera network: Calibration and application towards lane departure warning. *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*.
- Carrera, G., Angeli, A., Davison, A. J., 2011. SLAM-based automatic extrinsic calibration of a multi-camera rig. *IEEE International Conference on Robotics and Automation*.

- Caspi, Y., Irani, M., 2002. Aligning non-overlapping sequences. *International Journal of Computer Vision*, 48(1), 39–51.
- Dong, S., Shao, X., Kang, X., Yang, F., He, X., 2016. Extrinsic calibration of a non-overlapping camera network based on close-range photogrammetry. *Applied optics*, 55(23), 6363–6370.
- Esquivel, S., Woelk, F., Koch, R., 2007. Calibration of a multi-camera rig from non-overlapping views. *Joint Pattern Recognition Symposium*, 82–91.
- Förstner, W., Gülch, E., 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. *ISPRS intercommission conference on fast processing of photogrammetric data*.
- Gong, Z., Liu, Z., Zhang, G., 2017. Flexible global calibration of multiple cameras with nonoverlapping fields of view using circular targets. *Applied optics*, 56(11), 3122–3131.
- Hartmann, J., Paffenholz, A., J., Strübing, T., Neumann, I., 2017. Determination of position and orientation of LiDAR sensors on multisensor platforms. *Journal of Surveying Engineering*, 143(4). 04017012.
- Killian, K., 1955. Über das Rückwärtseinschneiden im Raum. *Österreichische Zeitschrift für Vermessungswesen*, 43(4), 97–104 and 171–179.
- Kitahara, I. T. A. R. U., Saito, H. I. D. E. O., Akimichi, S., Ono, T., Ohta, Y., Kanade, T., 2001. Large-scale virtualized reality. *Computer Vision and Pattern Recognition, Technical Sketches*. 2868352.
- Lébraly, P., Ait-Aider, O., Royer, E., Dhome, M., 2010a. Calibration of non-overlapping cameras-application to vision-based robotics. *British Machine Vision Conference (BMVC)*.
- Lébraly, P., Deymier, C., Ait-Aider, O., Royer, E., Dhome, M., 2010b. Flexible extrinsic calibration of non-overlapping cameras using a planar mirror: Application to vision-based robotics. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Liu, Q., Sun, J., Liu, Z., Zhang, G., 2012. Global calibration method of multi-sensor vision system using skew laser lines. *Chinese Journal of Mechanical Engineering*, 25(2), 405–410.
- Liu, Z., Wei, X., Zhang, G., 2013. External parameter calibration of widely distributed vision sensors with non-overlapping fields of view. *Optics and Lasers in Engineering*, 51(6), 643–650.
- Liu, Z., Zhang, G., Wei, Z., Sun, J., 2011a. A global calibration method for multiple vision sensors based on multiple targets. *Measurement Science and Technology*, 22(12). 125102.
- Liu, Z., Zhang, G., Wei, Z., Sun, J., 2011b. Novel calibration method for non-overlapping multiple vision sensors based on 1D target. *Optics and Lasers in Engineering*, 49(4), 570–577.
- Micusik, B., 2011. Relative pose problem for non-overlapping surveillance cameras with known gravity vector. *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Müller, F. J., 1925. Direkte (exakte) Lösung des einfachen Rückwärtseinschneidens im Raume. *Allgemeine Vermessungs-Nachrichten*, 37(16), 249–255.
- Ortega, A., Dias, B., Teniente, E., Bernardino, A., Gaspar, J., Andrade-Cetto, J., 2009. Calibrating an outdoor distributed camera network using laser range finder data. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Pagel, F., 2012. Motion adjustment for extrinsic calibration of cameras with non-overlapping views. *Ninth Conference on Computer and Robot Vision*.
- Strauss, T., Ziegler, J., Beck, J., 2014. Calibrating multiple cameras with non-overlapping views using coded checkerboard targets. *17th international IEEE conference on intelligent transportation systems (ITSC)*.
- Sturm, P., Bonfort, T., 2006. How to compute the pose of an object without a direct view? *Asian Conference on Computer Vision*, 21–31.
- Unger, J., 2020. Integrated estimation of UAV image orientation with a generalised building model. *Doctoral dissertation, Hannover: Leibniz University Hannover*.
- Xia, R., Hu, M., Zhao, J., Chen, S., Chen, Y., Fu, S., 2018. Global calibration of non-overlapping cameras: State of the art. *Optik*, 158, 951–961.
- Xu, Z., Wang, Y., Yang, C., 2015. Multi-camera global calibration for large-scale measurement based on plane mirror. *Optik*, 126(23), 4149–4154.