PFG 2013 / 4, 0309 – 0321
Stuttgart, August 2013

Article

Bundle Adjustment and System Calibration with Points at Infinity for Omnidirectional Camera Systems

JOHANNES SCHNEIDER & WOLFGANG FÖRSTNER, Bonn

Keywords: bundle adjustment, omnidirectional camera systems, multi-view cameras, calibration

Summary: We present a calibration method for multi-view cameras that provides a rigorous maximum likelihood estimation of the mutual orientation of the cameras within a rigid multi-camera system. No calibration targets are needed, just a movement of the multi-camera system taking synchronized images of a highly textured and static scene. Multi-camera systems with non-overlapping views have to be rotated within the scene so that corresponding points are visible in different cameras at different times of exposure. By using an extended version of the projective collinearity equation all estimates can be optimized in one bundle adjustment where we constrain the relative poses of the cameras to be fixed. For stabilizing camera orientations – especially rotations – one should generally use points at the horizon within the bundle adjustment, which classical bundle adjustment programs are not capable of. We use a minimal representation of homogeneous coordinates for image and scene points which allows us to use images of omnidirectional cameras with single viewpoint like fisheye cameras and scene points at a large distance from the camera or even at infinity.

We show results of our calibration method on (1) the omnidirectional multi-camera system Ladybug 3 from Point Grey, (2) a camera-rig with five cameras used for the acquisition of complex 3D structures and (3) a camera-rig mounted on a UAV consisting of four fisheye cameras which provide a large field of view and which is used for visual odometry and obstacle detection in the project MoD (DFG-Project FOR 1505 "Mapping on Demand").

Zusammenfassung: *Bündelausgleichung und Systemkalibrierung mit Punkten im Unendlichen für omnidirektionale Kamerasysteme.* In diesem Artikel stellen wir eine Kalibrierungsmethode für Multikamerasysteme vor, welche eine strenge Maximum-Likelihood-Schätzung der gegenseitigen Orientierungen der Kameras innerhalb eines starren Multikamerasystems ermöglicht. Zielmarken werden nicht benötigt. Das synchronisiert Bilder aufnehmende Kamerasystem muss lediglich in einer stark texturierten statischen Szene bewegt werden. Multikamerasysteme, deren Bilder sich nicht überlappen, werden innerhalb der Szene rotiert, so dass korrespondierende Punkte in jeder Kamera zu unterschiedlichen Aufnahmezeitpunkten sichtbar sind. Unter Verwendung einer erweiterten projektiven Kollinearitätsgleichung können alle zu schätzenden Größen in einer Bündelausgleichung optimiert werden. Zur Stabilisierung der Kameraorientierungen – besonders der Rotationen – sollten Punkte am Horizont in der Bündelausgleichung verwendet werden, wozu klassische Bündelausgleichungsprogramme nicht in der Lage sind. Wir benutzen eine minimale Repräsentation für homogene Koordinaten für Bild- und Objektpunkte, welche es uns ermöglicht, mit Bildern omnidirektionaler Kameras wie Fisheye-Kameras und mit Objektpunkten, welche weit entfernt oder im Unendlichen liegen, umzugehen.

Wir zeigen Ergebnisse unserer Kalibrierungsmethode für (1) das omnidirektionale Multikamerasystem Ladybug 3 von Point Grey, (2) ein Kamerasystem mit fünf Kameras zur Aufnahme komplexer 3D-Strukturen und (3) ein auf eine Drohne montiertes Kamerasystem mit vier Fisheye-Kameras, welches ein großes Sichtfeld besitzt und zur visuellen Odometrie und zur Hinderniserkennung im Projekt MoD (DFG-Projekt FOR 1505 „Mapping on Demand") verwendet wird.

www.schweizerbart.de
DOI: 10.1127/1432-8364/2013/0179
1432-8364/13/0179 $ 3.25

1 Introduction

1.1 *Motivation*

The paper presents a rigorous bundle adjustment for the estimation of the mutual camera orientations in a rigid multi-camera system. It is based on an extended version of the projective collinearity equation which constrains the relative poses of the cameras to be fixed, whereby all estimates can be optimized in one bundle adjustment. Further it enables the use of image and scene points at infinity like the bundle adjustment "BACS" (Bundle Adjustment for Camera Systems) presented in SCHNEIDER et al. (2012).

Bundle adjustment is the work horse for orienting cameras and determining 3D points. It has a number of favourable properties: It is statistically optimal in case all statistical tools are exploited, highly efficient in case sparse matrix operations are used, useful for test field free self calibration, and can be parallelized to a high degree. In this paper we want to extend bundle adjustment with the estimation of the parameters of the mutual camera orientation in a multi-camera system.

Multi-camera systems are used to increase the resolution, to combine cameras with different spectral sensitivities (Z/I DMC, Vexcel Ultracam) or – like *omnidirectional cameras* – to augment the effective aperture angle (Blom Pictometry, Rollei Panoscan Mark III). Following SCARAMUZZA (2008), omnidirectional cameras have a viewing range of more than a half-sphere, such as multi-cameras systems, catadioptric cameras including mirrors, or also special fisheye lenses, such as the Lensagon BF2M15520. Additionally, multi-camera systems gain importance for the acquisition of complex 3D structures.

Far or even ideal points, i.e. points at infinity, e.g. points at the horizon or luminous stars are effective in stabilizing the orientation of cameras, especially their rotations.

In order to exploit the power of a bundle adjustment, it therefore needs to be extended to handle multi-camera systems and image and scene points at infinity, see Fig. 1.

1.2 *Notation*

Vectors and matrices are typed slanted boldface, e.g. $\boldsymbol{x}$ and $\boldsymbol{R}$. Homogeneous vectors and matrices are typed upright, e.g. $\mathbf{x}$ and M. The skew 3×3-matrix $\boldsymbol{S}(\boldsymbol{a})$ induces the cross product, hence $\boldsymbol{S}(\boldsymbol{a})\boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}$, vertical vector concatenation e.g. of $\boldsymbol{x}$ and $\boldsymbol{y}$ is written in Matlab syntax $[\boldsymbol{x}; \boldsymbol{y}] = [\boldsymbol{x}^\mathsf{T}, \boldsymbol{y}^\mathsf{T}]^\mathsf{T}$.

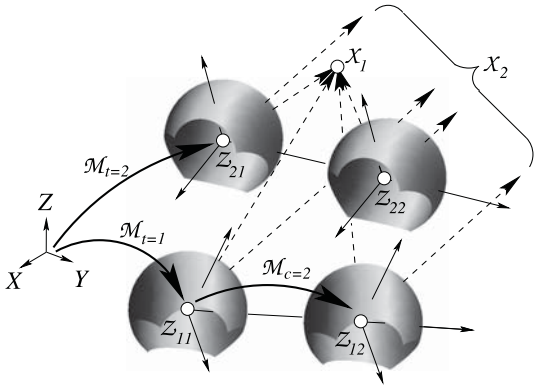


Fig. 1: A two-camera system with fisheye cameras $c = 1, 2$ with projection centers $\mathcal{Z}_{tc}$, rigid motion $\mathcal{M}_c$ and time-varying motion $\mathcal{M}_t$, having a field of view larger than $180°$ shown at two exposure times $t = 1, 2$ observing two points $\mathcal{X}_i, i = 1, 2$, one being close-by, the other at infinity. Already a block adjustment with a single camera moving over time will be stabilized by points at infinity.

1.3 *The Idea*

The classical collinearity equations for image points $\mathcal{x}'_{it}([x'_{it}; y'_{it}])$ of scene point $\mathcal{X}_i([X_i; Y_i; Z_i])$ in camera t with rotation matrix $\mathcal{R}_t([r_{kk'}])$ with k and $k' = 1, ..., 3$ and projection center $\mathcal{Z}_t([X_{0t}; Y_{0t}; Z_{0t}])$ read as

$$x'_{it} = \frac{r_{11}(X_i - X_{0t}) + r_{21}(Y_i - Y_{0t}) + r_{31}(Z_i - Z_{0t})}{r_{13}(X_i - X_{0t}) + r_{23}(Y_i - Y_{0t}) + r_{33}(Z_i - Z_{0t})} \quad (1)$$

$$y'_{it} = \frac{r_{12}(X_i - X_{0t}) + r_{22}(Y_i - Y_{0t}) + r_{32}(Z_i - Z_{0t})}{r_{13}(X_i - X_{0t}) + r_{23}(Y_i - Y_{0t}) + r_{33}(Z_i - Z_{0t})} \quad (2)$$

Obviously, these equations are not useful for far points or ideal points, as small angles between rays lead to numerical instabilities or singularities. They are not useful for bundles of rays of omnidirectional cameras, because rays perpendicular to the viewing direction, as they may occur with fisheye cameras, cannot be transformed into image coordinates. This would require different versions of the collinearity equation depending on the type of sensor as one would need to integrate the camera model into the bundle adjustment.

We can avoid these disadvantages by using homogeneous coordinates $\mathbf{x}'_{it}$ and $\mathbf{X}_i$ for im-

age and scene points, a calibration matrix K_t and the motion matrix M_t, containing the pose parameters of the camera system, in: $\mathbf{x}'_{it} = \lambda_{it}[\mathsf{K}_t \mid \mathbf{0}]\mathsf{M}_t^{-1}\mathbf{X}_i = \lambda_{it}\mathsf{P}_t\mathbf{X}_i$. Obviously, (a) homogeneous image coordinates allow for ideal image points, even directions opposite to the viewing direction, (b) homogeneous scene coordinates allow for far and ideal scene points, and including an additional motion is simply an additional factor.

However, this leads to two problems. As the covariance matrices $\mathbf{\Sigma}_{\mathbf{x}'_{it}\mathbf{x}'_{it}}$ of homogeneous vectors are singular, the optimization function of the maximum likelihood estimation $\sum_{it} |\mathbf{x}_{it} - \lambda_{it}\mathsf{P}_t\mathbf{X}_i|^2_{\mathbf{\Sigma}_{x'_{it}x'_{it}}}$ formally cannot be used. A minor, but practical problem is the increase of the number of unknown parameters, namely the Lagrangian multipliers, which are necessary when fixing the length of the vectors $\mathbf{X}_i$. In large bundle adjustments with more than a million scene points this prohibitively increases the number of unknowns by a factor 5/3.

1.4 *Task and Challenges*

The task is to model the projection process of a camera system as the basis for a bundle adjustment for a multi-view camera system, which (a) consists of mutually fixed single-view cameras, (b) allows the single cameras to be omnidirectional, requiring to explicitly model the camera rays and (c) which allows for far or ideal scene points for stabilizing the configuration. The model formally reads as

$$\mathcal{X}_{itc} = \mathcal{P}_c(\mathcal{M}_c^{-1}(\mathcal{M}_t^{-1}(\mathcal{X}_i))) \qquad (3)$$

with the I scene points $\mathcal{X}_i, i = 1, ..., I$, the T motions $\mathcal{M}_t, t = 1, ..., T$ of the camera system from the scene coordinate system, the C motions $\mathcal{M}_c, c = 1, ..., C$ of each single camera from the camera system, which makes the mutual orientation explicit, the projection $\mathcal{P}_c$ into the camera systems $c = 1, ..., C$, and the observed image points $\mathcal{X}_{itc}$ of scene point i in camera c at time/pose t.

In order to realize this we need to be able to represent bundles of rays together with their uncertainty, using uncertain direction vectors, to represent scene points at infinity using homogeneous coordinates, and minimize the number of parameters to be estimated. The main challenge lies in the inclusion of the statistics into an adequate minimal representation.

2 Related Work

Multi-camera systems are proposed by many authors. E. g. MOSTAFA & SCHWARZ (2001) present an approach to integrate a multi-camera system with GPS and INS. NISTÉR et al. (2004) discuss the advantage to use a stereo video rig in order to avoid the difficulty with the scale transfer. SAVOPOL et al. (2000) report on a multi-camera system for an aerial platform to increase the resolution. In all cases, the multi-view geometry is only used locally.

Orientation of a stereo rig is discussed in HARTLEY & ZISSERMAN (2000, p. 493). MOURAGNON et al. (2009) propose a bundle solution for stereo rigs working in terms of direction vectors, but they minimize the angular error without considering the covariance matrix of the observed rays. FRAHM et al. (2004) present an approach for orienting a multi-camera system, however not applying a statistically rigorous approach. MUHLE et al. (2011) discuss the ability to calibrate a multi-camera system in case the views of the individual cameras are not overlapping. IKEDA et al. (2003) describe a geometric and photometric camera calibration for omnidirectional multi-camera systems using a calibration board and a total station. CARRERA et al. (2011) calibrate a general multi-camera system by mapping each camera individually and applying a global bundle adjustment afterwards. ZOMET et al. (2001) discuss the problem of re-calibrating a rig of cameras due to changes of the internal parameters. Bundle adjustment of camera systems are extensively discussed in the thesis of KIM (2010).

Uncertain geometric reasoning using projective entities is extensively presented in KANATANI (1996), but only using Euclideanly normalized geometric entities and allowing the estimation for some single geometric entities only. HEUEL (2004), eliminating these deficiencies, proposes an estimation procedure which does not eliminate the redundancy of the representation and also cannot easily include elementary constraints between observations, see MEIDOW et al. (2009). The following developments are based on the minimal representation schemes proposed in FÖRSTNER (2012) which reviews previous work and generalizes e.g. BARTOLI (2002).

3 Concept

3.1 *Model for a moving single-view Camera*

3.1.1 Image coordinates as observations

Using homogeneous coordinates

$$\mathbf{x}'_{it} = \lambda_{it}\mathsf{P}_t\mathbf{X}_i = \lambda_{it}\mathsf{K}_t\boldsymbol{R}_t^\mathsf{T}[\boldsymbol{I}_3 \mid -\boldsymbol{Z}_t]\mathbf{X}_i \quad (4)$$

with a projection matrix

$$\mathsf{P}_t = [\mathsf{K}_t \mid \mathbf{0}_{3\times 1}]\mathsf{M}_t^{-1}, \quad \mathsf{M}_t = \begin{bmatrix} \boldsymbol{R}_t & \boldsymbol{Z}_t \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix}$$

makes the motion of the camera explicit. It contains for each pose t: the projection center $\boldsymbol{Z}_t$ and the rotation matrix $\boldsymbol{R}_t$, describing the translation and rotation between the scene coordinate system and the camera system, and the calibration matrix K_t, containing parameters for the principal point, the principal distance, the affinity, and possibly lens distortion, see McGlone et al. (2004, 3.149 ff.) and (10). In case of an ideal camera with principal distance c, thus $\mathsf{K}_t = \mathrm{Diag}([c, c, 1])$, and Euclidean normalization of the homogeneous image coordinates with the k-th row $\mathbf{A}_{t,k}^\mathsf{T}$ of the projection matrix P_t

$$\mathbf{x}'^{e}_{it} = \frac{\mathsf{P}_t\mathbf{X}_i}{\mathbf{A}_{t,3}^\mathsf{T}\mathbf{X}_i} = \begin{bmatrix} \mathbf{A}_{t,1}^\mathsf{T}\mathbf{X}_i/\mathbf{A}_{t,3}^\mathsf{T}\mathbf{X}_i \\ \mathbf{A}_{t,2}^\mathsf{T}\mathbf{X}_i/\mathbf{A}_{t,3}^\mathsf{T}\mathbf{X}_i \\ 1 \end{bmatrix} \quad (5)$$

we obtain (1) and (2), e. g. $x'_{it} = \mathbf{A}_{t,1}^\mathsf{T}\mathbf{X}_i/\mathbf{A}_{t,3}^\mathsf{T}\mathbf{X}_i$.

Observe the transposition of the rotation matrix in (4), which differs from Hartley & Zisserman (2000, (6.7)), but makes the motion of the camera from the scene coordinate system into the current camera system explicit, see Kraus (1997).

3.1.2 Ray directions as observations

Using the directions from the cameras to the scene points we obtain the collinearity equations

$$\begin{aligned} {}^k\mathbf{x}'_{it} &= \lambda_{it}\,{}^k\mathsf{P}_t\mathbf{X}_i = \lambda_{it}\boldsymbol{R}_t^\mathsf{T}(\boldsymbol{X}_i - \boldsymbol{Z}_t) \\ &= \lambda_{it}[\boldsymbol{I}_3 \mid \mathbf{0}]\mathsf{M}_t^{-1}\mathbf{X}_i\,. \end{aligned} \quad (6)$$

Instead of Euclidean normalization, we now perform spherical normalization $\mathbf{x}^s = \mathsf{N}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|$, where $|\mathbf{x}|$ is the length of $\mathbf{x}$, yielding the collinearity equations for camera bundles

$${}^k\mathbf{x}'^{s}_{it} = \mathsf{N}({}^k\mathsf{P}_t\mathbf{X}_i)\,. \quad (7)$$

We thus assume the camera bundles to be given as T sets $\{{}^k\mathbf{x}_{it}, i \in \mathcal{I}_t\}$ of normalized directions for each time t of exposure. The unknown parameters are the six parameters of the motion in ${}^k\mathsf{P}_t$ and the three parameters of each scene point. Care has to be taken with the sign: We assume the negative Z-coordinate of the camera system to be the viewing direction. The scene points then need to have non-negative homogeneous coordinate $X_{i,4}$, which in case they are derived from Euclidean coordinates via $\mathbf{X}_i = [\boldsymbol{X}_i; 1]$ always is fulfilled. In case of ideal points, we therefore need to distinguish between the scene points $[\boldsymbol{X}_i; 0]$ and $[-\boldsymbol{X}_i; 0]$ which are points at infinity in opposite directions.

As a first result we observe: The difference between the classical collinearity equations and the collinearity equations for camera bundles is twofold. 1.) The unknown scale factor is eliminated differently: Euclidean normalization leads to the classical form in (5), spherical normalization leads to the bundle form in (7). 2.) The calibration is handled differently: In the classical form it is made explicit, here we assume the image data to be transformed into camera rays taking the calibration into account. This will make a difference in modelling the individual cameras during self-calibration, a topic we will not discuss in this paper.

3.1.3 Handling far and ideal scene points

Handling far and ideal scene points can easily be realized by also using *spherically* normalized coordinates $\mathbf{X}_i^s$ for the scene points leading to

$${}^k\mathbf{x}'^{s}_{it} = \mathsf{N}({}^k\mathsf{P}_t\mathbf{X}_i^s)\,. \quad (8)$$

Again care has to be taken with points at infinity.

The confidence ellipsoid of 3D points can be used to visualize the achieved precision, in case the points are not too far. For a simultaneous visualization of confidence ellipsoids of 3D points which are close and far w.r.t. the origin one could perform a stereographic projection of the 3D-space into a unit sphere, i.e.

$\boldsymbol{X} \mapsto \boldsymbol{X}/(1+|\boldsymbol{X}|)$ together with the transformation of the confidence ellipsoids. The relative poses of points close to the origin then will be preserved, far points will sit close to the boundary of the sphere. Their uncertainty in distance to the origin then can be inferred using their distance to the boundary of the sphere.

3.2 *Model for Sets of Camera Systems*

With an additional motion $\mathsf{M}_c(\boldsymbol{R}_c, \boldsymbol{Z}_c)$ for each camera of the camera system we obtain the general model for camera bundles

$$ {}^k\mathbf{x}'^s_{itc} = \mathsf{N}\left([\boldsymbol{I}_3 \,|\, \mathbf{0}_{3\times1}]\; \mathsf{M}_c^{-1}\; \mathsf{M}_t^{-1}\; \mathbf{X}_i^s\right) \quad (9) $$

which makes all elements explicit: The observed directions $\chi'_{itc}({}^k\mathbf{x}'_{itc})$ represented by normalized 3-vectors, having two degrees of freedom, unknown or known scene point coordinates $\mathcal{X}_i(\mathbf{X}_i^s)$, represented by spherically normalized homogeneous 4-vectors, having 3 degrees of freedom, unknown pose $\mathcal{M}_t$ of camera system, having 6 parameters for each time a set of images was taken and unknown calibration M_c containing the relative pose of the cameras which are assumed to be rigid over time, having 6 parameters per camera. We refer relative poses to the first camera as reference camera with $\boldsymbol{R} = \boldsymbol{I}_3$ and $\boldsymbol{Z} = \mathbf{0}$.

3.3 *Generating Camera Directions from observed Image Coordinates*

In most cases the observations are made using a digital camera whose sensor is approximately planar. The transition to the directions of the camera rays needs to be performed before starting the bundle adjustment. As mentioned before, this requires the internal camera geometry to be known. Moreover, in order to arrive at a statistically optimal solution, one needs to transfer the uncertainty of the observed image coordinates to the uncertainty of the camera rays. As an example we discuss two cases.

3.3.1 Perspective cameras

In case of perspective cameras with small image distortions, we can use the camera-specific and maybe temporally varying calibration matrix

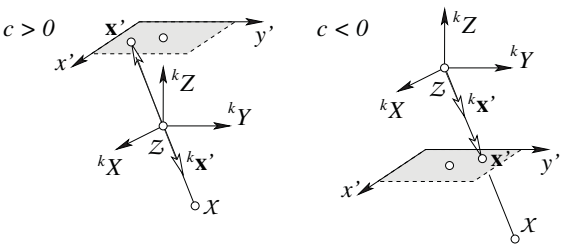


Fig. 2: The direction of the homogeneous image coordinate vector and the direction of the ray is different depending on the sign of the principal distance c.

$$ \mathsf{K}(\boldsymbol{x}', \boldsymbol{q}) = \begin{bmatrix} c & cs & x'_H + \Delta x(\boldsymbol{x}', \boldsymbol{q}) \\ 0 & c(1+m) & y'_H + \Delta y(\boldsymbol{x}', \boldsymbol{q}) \\ 0 & 0 & 1 \end{bmatrix} \quad (10) $$

for the forward transformation

$$ {}^g\mathbf{x}' = \mathsf{K}(\boldsymbol{x}', \boldsymbol{q})\; {}^k\mathbf{x}'^s \quad (11) $$

leading to the observable image coordinates ${}^g\mathbf{x}'$. The ^{g} indicates that the mapping can handle general distortions via additional parameters $\boldsymbol{q}$. Besides the basic parameters, namely the principal distance c with the image plane ${}^kZ = c$, the shear s, the scale difference m, and principal point $\boldsymbol{x}'_H$, the calibration matrix contains additive corrections for modelling lens distortion or other deviations, which depend on the additional parameters $\boldsymbol{q}$ and on $\boldsymbol{x}$, the position of the image point. In case of small deviations (11) can easily be inverted. However, one must take into account the different signs of the coordinate vector and the direction from the camera to the scene point (Fig. 2),

$$ {}^k\mathbf{x}'^s \approx s\; \mathsf{N}\left(\mathsf{K}^{-1}({}^g\boldsymbol{x}', \boldsymbol{q})\; {}^g\mathbf{x}'\right) \quad (12) $$

with $s \in \{-1, +1\}$ such that ${}^kx'^s_3 < 0$. This relation is independent of the sign of the third element of the calibration matrix. Given the covariance matrix $\Sigma_{{}^g x' {}^g x'}$ of the image coordinates, the covariance matrix of ${}^k\mathbf{x}'$ can be determined by variance propagation, omitting the dependency of the calibration matrix on the point coordinates $\boldsymbol{x}'$. Note that a point ${}^g\mathbf{x}'$ at infinity corresponds to the direction ${}^k\mathbf{x}'$ perpendicular to the viewing direction.

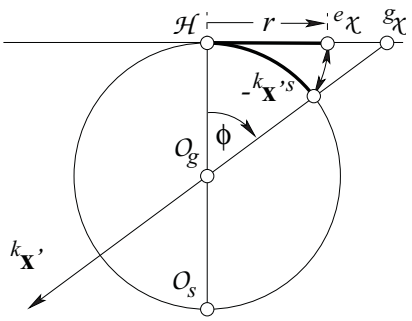


Fig. 3: Relation between sensor point, viewing direction and viewing ray.

3.3.2 Omnidirectional single-viewpoint cameras

As an example for an omnidirectional single-view camera we take a camera with a fisheye-lens. We model the fisheye objective with the equidistant-model described in ABRAHAM & FÖRSTNER (2005). The interior orientation of a camera is determined separately by camera calibration according to ABRAHAM & HAU (1997) using Chebyshev polynomials. Using the equidistant-projection and applying all corrections, we obtain image points ${}^e\mathcal{X}$ which lie closer to the principal point $\mathcal{H}$ than the gnomonic projections ${}^g\mathcal{X}$ of the scene points (Fig. 3). The ray direction ${}^k\mathbf{x}'^s$ can be derived from ${}^e\mathcal{X}$ by using the normalized radial distance $r' = |{}^e\boldsymbol{x}|$ growing with the angle ϕ between the viewing direction and the camera ray.

Again, the uncertainty of the image coordinates can be transformed to the uncertainty of the direction ${}^k\mathbf{x}'^s$ of the camera ray via variance propagation. In all cases the covariance matrix of the camera ray is singular, as the normalized 3-vector only depends on two observed image coordinates.

3.4 *The Estimation Procedure*

The collinearity equations in (9) contain three equations per observed camera ray and four parameters for each homogeneous scene point, though, both being unit vectors. Therefore, the corresponding covariance matrices are singular and more than the necessary parameters are contained in the equations. We therefore want to reduce the number of parameters to the necessary minimum. We do this after linearization.

3.4.1 Linearization and update for pose and relative pose parameters

Linearization of the non-linear model leads to a linear substitute model which yields correction parameters that allow to derive corrected approximate values. We start with approximate values for the poses of the reference camera given for every time of exposure $t = 1, ..., T$ by R_t^a for the rotation matrix and $\boldsymbol{Z}_t^a$ for the projection center, the relative poses from the reference camera to each other camera $c = 2, ..., C$ given by R_c^a and $\boldsymbol{Z}_c^a$, $\mathbf{X}_i^{sa}$ for the $i = 1, ..., I$ spherically normalized scene points, and $\mathbf{x}_{itc}^a = \mathsf{N}\left([I_3 \,|\, \mathbf{0}_{3\times1}]\ \mathsf{M}_c^{-1}\ \mathsf{M}_t^{-1}\ \mathbf{X}_i^s\right)$ for the normalized directions.

The Euclidean coordinates will be simply corrected by $\boldsymbol{Z} = \boldsymbol{Z}^a + \Delta\boldsymbol{Z}$, the three parameters $\Delta\boldsymbol{Z}$ are to be estimated. The rotation matrix will be corrected by pre-multiplication with a small rotation, thus by $R = R(\Delta\boldsymbol{R})\,R^a \approx (I_3 + S(\Delta\boldsymbol{R}))R^a$, where the small rotation $R(\Delta\boldsymbol{R})$ depends on a small rotation vector $\Delta\boldsymbol{R}$ that is to be estimated.

3.4.2 Reduced coordinates and update of coordinates

The correction of the unit vectors is performed using reduced coordinates (FÖRSTNER 2012). These are coordinates, say the 2-vector $\boldsymbol{x}_r$ of the direction $\mathbf{x}^s$, in the two-dimensional tangent space null $\left(\mathbf{x}^{sa\mathsf{T}}\right) = [\boldsymbol{r}, \boldsymbol{s}]$ of the unit sphere S^2 evaluated at the approximate values $\mathbf{x}^{sa}$

$$\boldsymbol{x}_r = \text{null}^{\mathsf{T}}\left(\mathbf{x}^{sa\mathsf{T}}\right)\mathbf{x}^s = \begin{bmatrix} \boldsymbol{r}^{\mathsf{T}}\mathbf{x}^s \\ \boldsymbol{s}^{\mathsf{T}}\mathbf{x}^s \end{bmatrix}. \quad (13)$$

The corrections $\Delta\boldsymbol{x}_r$ of these reduced coordinates are estimated. This leads to the following update rule

$$\mathbf{x}^s = \mathsf{N}\left(\mathbf{x}^{sa} + \text{null}\left(\mathbf{x}^{sa\mathsf{T}}\right)\Delta\boldsymbol{x}_r\right). \quad (14)$$

Obviously, the approximate vector $\mathbf{x}^{sa}$ is corrected by

$$\Delta\mathbf{x} = \text{null}\left(\mathbf{x}^{sa\mathsf{T}}\right)\Delta\boldsymbol{x}_r \quad (15)$$

and then spherically normalized to achieve the updated values $\mathbf{x}^s$.

Using (13) we now are able to reduce the number of equations per direction from three to two, making the two degrees of freedom of the

observed direction explicit. This results in premultiplication of all observation equations on (9) with $\text{null}^{\mathsf{T}}\left({}^{k}\mathbf{x}_{itc}^{sa\mathsf{T}}\right)$. Following (15) we use the substitution $\Delta\mathbf{X}_i^s = \text{null}\left(\mathbf{X}_i^{sa\mathsf{T}}\right)\Delta\boldsymbol{X}_{r,i}$ when linearizing the scene coordinates. Then we obtain the linearized model

$$
\begin{aligned}
&{}^{k}\boldsymbol{x}_{r,itc} + \widehat{\boldsymbol{v}}_{x_{r,itc}} \qquad (16)\\
&= \boldsymbol{J}^{\mathsf{T}}\boldsymbol{R}_c^{a\mathsf{T}}\boldsymbol{R}_t^{a\mathsf{T}}\boldsymbol{S}(\boldsymbol{X}_{i0}^a)\widehat{\Delta\boldsymbol{R}}_t\\
&\quad - X_{ih}^a\boldsymbol{J}^{\mathsf{T}}\boldsymbol{R}_c^{a\mathsf{T}}\boldsymbol{R}_t^{a\mathsf{T}}\widehat{\Delta\boldsymbol{Z}}_t\\
&\quad + \boldsymbol{J}^{\mathsf{T}}\boldsymbol{R}_c^{a\mathsf{T}}\boldsymbol{S}\left(\boldsymbol{R}_t^{a\mathsf{T}}(\boldsymbol{X}_{i0}^a - X_{ih}^a\boldsymbol{Z}_t^a)\right)\widehat{\Delta\boldsymbol{R}}_c\\
&\quad - X_{ih}^a\boldsymbol{J}^{\mathsf{T}}\boldsymbol{R}_c^{a\mathsf{T}}\widehat{\Delta\boldsymbol{Z}}_t\\
&\quad + \boldsymbol{J}^{\mathsf{T}}[\boldsymbol{I}_3 \,|\, \mathbf{0}_3](\mathsf{M}_c^a)^{-1}(\mathsf{M}_t^a)^{-1}\text{null}\left(\mathbf{X}_i^{a\mathsf{T}}\right)\widehat{\Delta\boldsymbol{X}}_{ri}
\end{aligned}
$$

with

$$
\boldsymbol{J} = \frac{1}{|\mathbf{x}|}\left(\boldsymbol{I}_3 - \frac{\mathbf{x}\mathbf{x}^{\mathsf{T}}}{\mathbf{x}^{\mathsf{T}}\mathbf{x}}\right)\text{null}(\mathbf{x}^{\mathsf{T}})\Big|_{\mathbf{x}={}^{k}\mathbf{x}_{itc}^a} \qquad (17)
$$

and the partitioned homogeneous vector $\mathbf{X}^s = [\boldsymbol{X}_0; X_h]$ depending on $\widehat{\Delta\boldsymbol{R}}_t$, $\widehat{\Delta\boldsymbol{Z}}_t$, $\widehat{\Delta\boldsymbol{R}}_c$, $\widehat{\Delta\boldsymbol{Z}}_c$ and $\widehat{\Delta\boldsymbol{X}}_{r,i}$.

We now arrive at a well-defined optimization problem: find $\widehat{\Delta\boldsymbol{X}}_{r,i}, \widehat{\Delta\boldsymbol{R}}_t, \widehat{\Delta\boldsymbol{Z}}_t, \widehat{\Delta\boldsymbol{R}}_c, \widehat{\Delta\boldsymbol{Z}}_c$ minimizing

$$
\begin{aligned}
&\Omega\left(\widehat{\Delta\boldsymbol{X}}_{r,i}, \widehat{\Delta\boldsymbol{R}}_t, \widehat{\Delta\boldsymbol{Z}}_t, \widehat{\Delta\boldsymbol{R}}_c, \widehat{\Delta\boldsymbol{Z}}_c\right) \qquad (18)\\
&= \sum_{itc}\widehat{\boldsymbol{v}}_{r,itc}^{\mathsf{T}}\boldsymbol{\Sigma}_{x_{r,itc}x_{r,itc}}^{-1}\widehat{\boldsymbol{v}}_{r,itc}
\end{aligned}
$$

with the regular 2×2-covariance matrices

$$
\begin{aligned}
&\boldsymbol{\Sigma}_{x_{r,itc}x_{r,itc}} \qquad (19)\\
&= {}^{k}\boldsymbol{J}_s^{\mathsf{T}}({}^{k}\mathbf{x}_{itc}^a)\begin{bmatrix}\boldsymbol{\Sigma}_{x_{itc}x_{itc}} & \mathbf{0}\\ \mathbf{0}^{\mathsf{T}} & 0\end{bmatrix}{}^{k}\boldsymbol{J}_s({}^{k}\mathbf{x}_{itc}^a).
\end{aligned}
$$

4 Experiments

4.1 *Implementation Details*

We have implemented the bundle adjustment as a Gauss-Markov model in Matlab. Observations are bundles of rays, the interior orientations of all cameras of the camera system are assumed to be known. To overcome the rank deficiency we define the gauge by introducing seven centroid constraints on the approximate values of the scene points. This results in a free bundle adjustment, where the trace of the covariance matrix of the estimated scene points is minimal. We can robustify the cost function by down-weighting measurements whose residual errors are too large by minimizing the robust Huber cost function HUBER (1981).

For the initialization sufficiently accurate approximate values for the scene point coordinates $\mathbf{X}_i^a$ and for the translation and the rotation of the relative poses M_c^a of each camera in the camera system of the reference camera as well as for the poses M_t^a of the reference camera in the scene coordinate system at the times of synchronized exposure are needed. Firstly, we determine the pose of each camera without considering the cameras as a rigid multi-camera rig using the bundle adjustment program Aurelo provided by LÄBE & FÖRSTNER (2006). With the first camera as the reference camera we then determine approximate values for all $c = 2, ..., C$ relative poses M_c^a robustly using the median quaternion and median translation over all $t = 1, ..., T$ unconstrained estimated relative poses. Scene points are triangulated by using all corresponding image points that are consistent with the approximated relative poses and the poses of the reference camera in the scene coordinate system. Ray directions with large residuals are discarded.

4.2 *Test on Correctness and Advantage*

We first check the correctness of the implemented model and then show the advantage of including far points or points with glancing intersections within the bundle adjustment based on a simulated scenario.

4.2.1 Simulated scenario

We simulated a multi-camera system moving on a radiused square, observing 50 close scene points and 10 scene points far away at the horizon, i.e. at infinity (Fig. 4). The multi-camera system contains three single-view cameras. Every scene point is observed by a camera ray from all 20 positions of the camera system. The simulated set-up provides a high redundancy of observations. Assuming the standard deviation of an image coordinate to be 0.3 pixel and

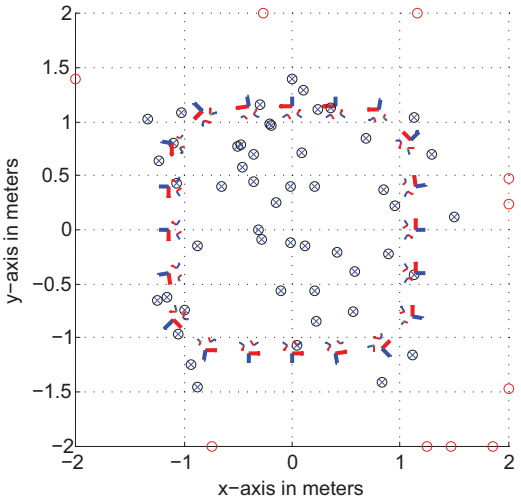


Fig. 4: Simulation of a moving multi-camera system (poses of reference camera shown as bold tripods) with loop closing. Scene points nearby (crossed dots) and at the horizon (empty dots) being numerically at infinity are observed.

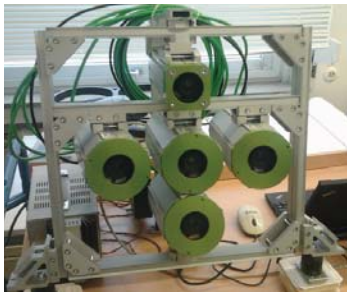


Fig. 5: Multi-camera system consisting of five overlapping perspective camera views: Infrared camera on top, RGB camera in the middle and three monochromatic cameras. The distances from the RGB camera to the others are about 10 cm.

a principal distance of 500 pixel, we add normally distributed noise with $\sigma_l = 0.3/500$ radiant on the spherically normalized camera rays to simulate the observation process (Fig. 4). As initial values for the bundle adjustment we randomly disturb both the generated spherical normalized homogeneous scene points $\mathbf{X}_i^s$, which are directions, by 6°, the generated motion parameters of the reference camera $\boldsymbol{R}_t$ and $\boldsymbol{Z}_t$ of M_t by 3° and 2 cm, and the relative pose parameters of the remaining cameras $\boldsymbol{R}_c$ and $\boldsymbol{Z}_c$ of M_c by 3° and 10 % of the relative distances between the projection centres.

The iterative estimation procedure stops after eight iterations, when the maximum normalized observation update is less than 10^{-6}. The residuals of the observed image rays in the tangent space of the adjusted camera rays, which are approximate angles between the rays in radiants, do not show any deviation from the normal distribution. The estimated a posteriori variance factor $\widehat{\sigma}_0^2 = 0.9925^2$ approves the a priori stochastic model with the variance factor $\sigma_0^2 = 1$. In order to test if the estimated orientation parameters and scene point coordinates represent the maximum likelihood estimates for normally distributed noise of the observations, we have generated the same simulation 2000 times with different random noise. The mean of the estimated variance factors is not significantly different from one, indicating an unbiased estimator with minimum variance. These results confirm the correctness of the approach and implementation.

4.2.2 Decrease of rotational precision excluding far points

Bundle adjustment programs, such as Aurelo, cannot handle scene points with glancing intersections, e.g. with maximal intersection angles lower than $\gamma = 1$ gon, which therefore are excluded in the estimation process to avoid numerical difficulties. Far scene points, however, can be observed over long periods of time and therefore should improve the quality of the rotation estimation significantly. We investigate the decrease of precision of the estimated rotation parameters of $\widehat{\boldsymbol{R}}_t$ and $\widehat{\boldsymbol{R}}_c$ when excluding scene points with glancing intersection angles. In detail, we will determine the average empirical standard deviation $\sigma_{\alpha_t} = \widehat{\sigma}_0 \sqrt{\operatorname{tr} \boldsymbol{\Sigma}_{\widehat{R}_t \widehat{R}_t}/3}$ and $\sigma_{\alpha_c} = \widehat{\sigma}_0 \sqrt{\operatorname{tr} \boldsymbol{\Sigma}_{\widehat{R}_c \widehat{R}_c}/3}$ for all estimated rotation parameters and report the average decrease of precision by excluding far points. They are determined by the geometric mean, namely $\exp\left[\sum_t^T \log(\sigma'_{\alpha_t}/\sigma_{\alpha_t})/T\right]$ and $\exp\left[\sum_c^{C-1} \log(\sigma'_{\alpha_c}/\sigma_{\alpha_c})/(C-1)\right]$, where σ'_{α_t} and σ'_{α_c} represent the resulting average empirical standard deviations when scene points whose maximal intersection angle are lower than a threshold γ are excluded.

We determine the decrease of precision for the estimated rotation parameters by excluding a varying number of scene points at infinity on the basis of the introduced simulation

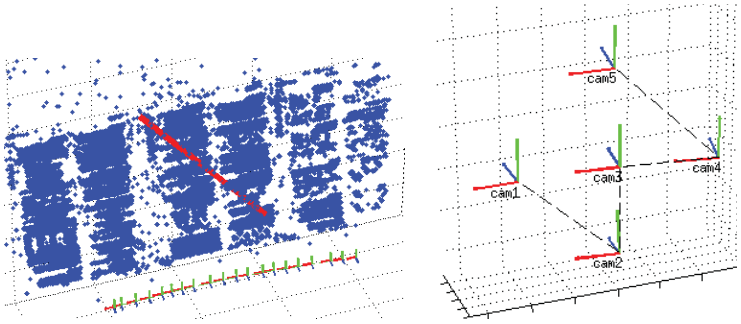


Fig. 6: Left: Illustration of the estimated scene points and poses of the reference camera. The red line denotes the known length on a poster for scale definition. Right: The estimated relative poses.

of a moving multi-camera system. Again we generate 50 scene points close to the multi-camera positions and vary the number of scene points at infinity to be 5, 10, 20, 50 and 100. The resulting average decrease in precision of the estimated rotations in $\widehat{\mathsf{M}}_c$ is 6.21 %, 8.98 %, 19.90 %, 42.29 % and 75.60 % and in $\widehat{\mathsf{M}}_t$ 7.15 %, 11.77 %, 27.67 %, 54.56 % and 91.28 %, respectively. This strongly proves the points at infinity to have a highly relevant positive influence on the rotational precision.

4.3 *Calibration of Multi-Camera Systems*

4.3.1 Calibration with overlapping views

We now describe the calibration of the camera system shown in Fig. 5 with highly overlapping views, which is used for 3D reconstruction of vines. In order to determine the relative poses of the multi-camera system we apply the bundle adjustment to 100 images of a wall draped with highly textured posters. The images were taken at 20 stations in a synchronized way. We use Aurelo without considering the known relative orientation between the cameras to obtain an initial solution for each camera and the scene points. The dataset contains 593,412 image points and 63,140 observed scene points.

Starting from an a priori standard deviation of the image coordinates of $\sigma_l = 1$ pixel, the a posteriori variance factor is estimated with $\widehat{\sigma}_0^2 = 0.11^2$ indicating the automatically extracted Lowe points to have an average precision of approximately 0.1 pixel. This high precision of the point detection results mainly from the good images and the calibration quality of the camera used. Fig. 6 illustrates the estimated scene points and poses as well as the estimated relative poses.

The estimated uncertainty of the estimated rotations of the cameras with regard to the reference camera is 0.1 mgrad – 0.2 mgrad around the viewing direction axis and 0.4 mgrad – 0.8 mgrad orthogonal to it. We scale the photogrammetric model by using a measured distance of 1.105 m with an error of about 0.1 %. The uncertainty of the estimated relative translations is 0.02 mm – 0.04 mm in viewing direction and 0.1 mm – 0.2 mm orthogonal to it.

4.3.2 Multi-camera system Ladybug 3

The omnidirectional multi-camera system Ladybug 3 consists of six cameras, five of which are mounted in a circular manner, one showing upwards, together covering 80 % of the full viewing sphere. Neighbouring images only have a very small overlap, which is too weak for system calibration without additional information. We have mounted the omnidirectional multi-camera system Ladybug 3 on a robot (Fig. 7a), which executes a circular movement with a radius of 50 cm in a highly textured room while the Ladybug is taking synchronized images. This ensures overlapping images of different cameras at different times of the exposure. Approximate values for this image sequence consisting of 150 images taken by the five horizontal cameras at 30 exposure times are obtained with Aurelo that provides 135,012 image points of 24,078 observed scene points. The resulting 150 camera poses are shown in Fig. 7b.

After applying our bundle adjustment the estimated a posteriori variance factor amounts to $\widehat{\sigma}_0^2 = 0.25^2$ using a priori stochastic model with $\sigma_l = 1$ pixel for the image points, indicating the

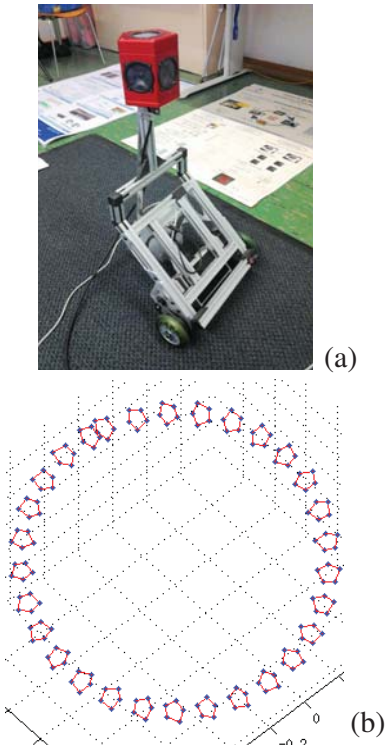


Fig. 7: (a) Ladybug 3 on robot, (b) Estimated camera poses.

automatically extracted Lowe points to have a quite good precision. Two parallel walls with a known distance of 7.01 m can be estimated out of the estimated scene points, which is used to define the scale between the estimated relative camera poses. The estimated rotation parameters show a very high precision. The maximal deviation to the manufacturer's calibration parameters is 0.6°. The estimated uncertainties of the rotations and translations between the cameras are in the order of 1.5 mgrad – 2.5 mgrad and 0.1 mm – 0.2 mm, respectively.

To compare the estimated poses with the ones provided by the manufacturer we apply a rigid transformation which minimizes the distances between the estimated and given projection centers. The resulting estimated relative poses in Fig. 8 show significant translational deviations in the order of 1 mm – 4 mm compared to the manufacturer's calibration parameters.

The interior angles differ from a regular pentagon where each interior angle is 108° by up to 13°. Possible reasons for the deviations are to few observed scene points near to the camera system and that we have used a different interior orientation for each camera from our own calibration, which is different from that of the manufacturer.

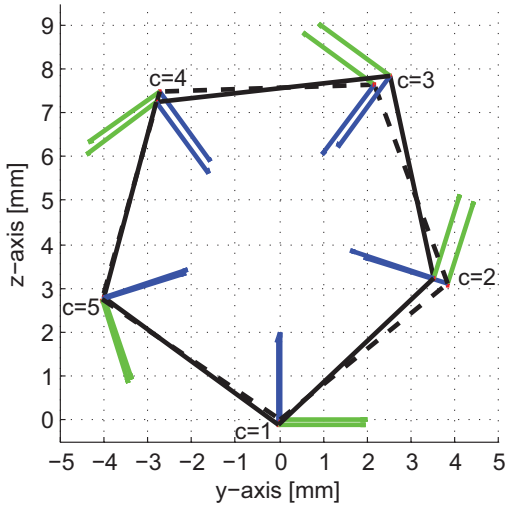


Fig. 8: Comparison of relative poses: estimated (solid) and manufacturer given (dashed).

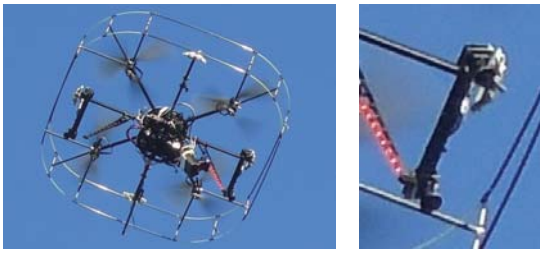


Fig. 9: Illustration of the UAV and the stereo cameras used in the MoD-project. One stereo pair is looking forward and one backwards giving a wide field of view.

4.3.3 Multi-camera system with fisheye lenses

We make the same investigation on an image sequence consisting of 96 images taken by four synchronized cameras with Lensagon BF2M15520 fisheye lenses having a field angle up to 185°. The cameras are mounted on an UAV (unmanned aerial vehicle) to generate two stereo pairs, one looking ahead and one looking backwards, providing a large field of view (Fig. 9). The UAV moves along a circle at a height of 5 m above a parking lot while rotating around its own axis, providing four overlapping images at each time of exposure.

Fisheye objectives cause severe distortions at the image boundary. Thus, in order to find corresponding points using the SIFT-operator we need to use a transformation between overlapping images which is very similar to a conformal projection, i.e. one that preserves angles because the SIFT operator is translation,

Fig. 10: Sample images of the Ladybug 3 dataset.

scale and rotation invariant. For this reason we transform the original images using the stereographic fisheye model. This ensures a conformal mapping between two different images when observing a scene at infinity as they themselves are conformal mappings of the spherical image of the scene. We obtain low deviations from a similarity transformation for locally planar points not too close to the cameras, fulfilling the preconditions for rotation and scale invariant SIFT-matching.

Aurelo provides approximate values for the 96 camera poses and 81,821 image points of 15,344 observed scene points which are transformed into image directions using (16). After the bundle adjustment the estimated variance factor is $\widehat{\sigma}_0^2 = 1.47^2$ using an a priori stochastic model with $\sigma_l = 1$ pixel for the image points, indicating a quite poor precision of the point detection. The cause for this low precision, which still needs to be analyzed, may be a lower image quality caused by both, the fisheye projection, or vibrations. The uncertainty of the estimated rotations and translations between the cameras within a stereo pair is 2 mgrad – 6 mgrad and 0.5 mm – 1.5 mm, respectively, and the uncertainty of the estimated rotations and translations between the forward and backward looking stereo camera systems is 5 mgrad – 9 mgrad and 1.5 mm – 2.5 mm.

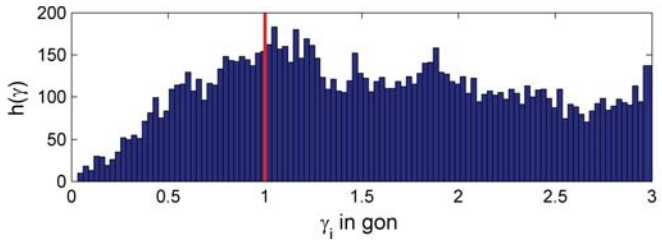
4.4 *Decrease of rotational Precision excluding far Points*

In order to examine the decrease of the rotational precision of the estimated camera system poses we apply the bundle adjustment to an image sequence consisting of 360 images taken by four of the six cameras of the multi-camera system Ladybug 3 (Fig. 10) excluding and including far points. The Ladybug 3 is mounted on a hand-guided platform and is triggered for one shot per meter with the help of an odometer. Approximate values are obtained with Aurelo by combining the individual cameras into a single virtual camera by adding distance-dependent corrections to the camera rays (SCHMEING et al. 2011).

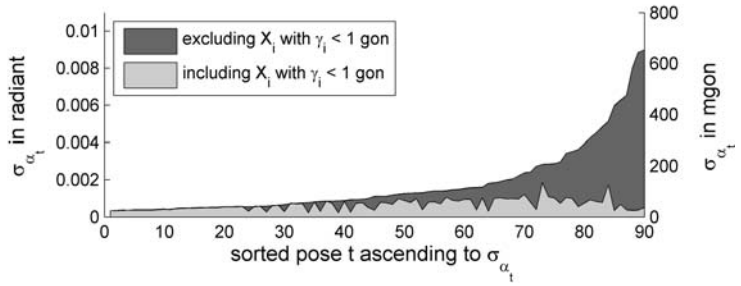
The dataset contains 10,891 of 26,890 scene points observed with maximal intersection angles per point significantly lower than $\gamma = 1$ gon (histogram in Fig. 11a). The average standard deviation of each estimated rotation parameter is shown in Fig. 11b showing the individual gain in precision, sorted according to ascending rotational standard deviation. For some images the gain obviously is very large. The gain is mainly obtained due to a higher number of observed scene points at the individual poses, which can be seen in the scatter plot in Fig. 11c. Some of the estimated rotations show very large differences in the precision, demonstrating the relevance of the far scene points in the Ladybug 3 dataset. The use of far points results in an almost constant precision of the rotation parameters over all camera stations, compared to the results of the bundle adjustment if far points are excluded. The estimated a posteriori variance factor is $\widehat{\sigma}_0^2 = 1.05^2$ using an a priori stochastic model with $\sigma_l = 1$ pixel for the image points, indicating a quite poor precision of the point detection which mainly results from the low image quality.

5 Conclusions and Future Work

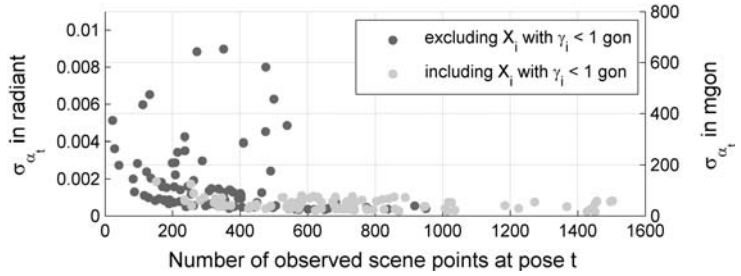
We proposed a rigorous bundle adjustment for omnidirectional and multi-view cameras which enables an efficient maximum likelihood estimation using image and scene points at infinity and which can be used to calibrate a general



(a) Number of scene points with small intersection angles.



(b) Average empirical standard deviation of estimated rotations.



(c) Scatter plot of σ_{α_t} against the number of observed scene points at t.

Fig. 11: The histogram in (a) shows the number of scene points in the multi-camera dataset with small intersection angles. The average precision σ_{α_t} determined by excluding and including scene points with $\gamma < 1$ gon for all poses $t = 1, ..., T$ is compared to each other in (b) and against the number of observed scene points in (c).

multi-camera system. Our experiments on simulated data show that scene points at the horizon can stabilize the orientation of the camera rotations significantly. Future work will focus on improving the precision of the relative poses by testing different image acquisition strategies. Furthermore, we are developing a fast C-implementation and eventually will extend our software by a self calibration part.

Software

The Matlab code of the proposed bundle adjustment will be made available at: `www.ipb.uni-bonn.de/bacs`.

6 Acknowledgements

This work was supported by the DFG-Project FOR 1505 “Mapping on Demand”. We thank THOMAS LÄBE for supporting the experiments, especially performing the camera calibration, and the reviewers for their helpful comments.

References

ABRAHAM, S. & FÖRSTNER, W., 2005: Fish-Eye-Stereo Calibration and Epipolar Rectification. – ISPRS Journal of Photogrammetry & Remote Sensing **59** (5): 278–288.

ABRAHAM, S. & HAU, T., 1997: Towards Autonomous High-Precision Calibration of Digital Cameras. – Videometrics V, SPIE Annual Meeting **3174**: 82–93.

BARTOLI, A., 2002: On the Non-linear Optimization of Projective Motion Using Minimal Parameters. – 7th European Conference on Computer Vision - Part II: 340–354.

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: 2652–2659.

FÖRSTNER, W., 2012: Minimal Representations for Testing and Estimation in Projective Spaces. – PFG – Photogrammetrie, Fernerkundung, Geoinformation **2012** (3): 209–220.

FRAHM, J.-M., KÖSER, K. & KOCH, R., 2004: Pose estimation for Multi-camera Systems. – Pattern Recognition, 26th DAGM-Symposium: 286–293.

HARTLEY, R.I. & ZISSERMAN, A., 2000: Multiple View Geometry in Computer Vision. – Cambridge University Press.

HEUEL, S., 2004: Uncertain Projective Geometry: Statistical Reasoning for Polyhedral Object Reconstruction. – LNCS **3008**, Springer.

HUBER, P.J., 1981: Robust Statistics. – John Wiley, New York, NY, USA.

IKEDA, S., SATO, T. & YOKOYA, N., 2003: Calibration Method for an Omnidirectional Multicamera System. – SPIE **5006**: 499–507.

KANATANI, K., 1996: Statistical Optimization for Geometric Computation: Theory and Practice. – Elsevier Science.

KIM, J.-H., 2010: Camera Motion Estimation for Multi-Camera Systems. – PhD thesis, School of Engineering, ANU College of Engineering and Computer Science, The Australian National University.

KRAUS, K., 1997: Photogrammetry. – Dümmler Verlag Bonn: Vol.1: Fundamentals and Standard Processes. Vol.2: Advanced Methods and Applications.

LÄBE, T. & FÖRSTNER, W., 2006: Automatic Relative Orientation of Images. – 5th Turkish-German Joint Geodetic Days.

MCGLONE, C.J., MIKHAIL, E.M. & BETHEL, J.S., 2004: Manual of Photogrammetry. – American Society of Photogrammetry and Remote Sensing.

MEIDOW, J., BEDER, C. & FÖRSTNER, W., 2009: Reasoning with Uncertain Points, Straight Lines, and Straight Line Segments in 2D. – International Journal of Photogrammetry and Remote Sensing **64**: 125–139.

MOSTAFA, M.M. & SCHWARZ, K.-P., 2001: Digital Image Georeferencing from a Multiple Camera System by GPS/INS. – ISPRS Journal of Photogrammetry and Remote Sensing **56** (1): 1–12.

MOURAGNON, E., LHUILLIER, M., DHOME, M. & DEKEYSER, F., 2009: Generic and Real-time Structure from Motion Using Local Bundle Adjustment. – Image and Vision Computing **27** (8): 1178–1193.

MUHLE, D., ABRAHAM, S., HEIPKE, C. & WIGGENHAGEN, M., 2011: Estimating the Mutual Orientation in a Multi-Camera System with a Non Overlapping Field of View. – 2011 ISPRS Conference on Photogrammetric Image Analysis.

NISTÉR, D., NARODITSKY, O. & BERGEN, J., 2004: Visual Odometry. – IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 652–659.

SAVOPOL, F., CHAPMAN, M. & BOULIANNE, M., 2000: A Digital Multi CCD Camera System for Near Real-Time Mapping. – International Archives of Photogrammetry and Remote Sensing, Volume XXXIII. ISPRS Congress, Amsterdam, The Netherlands.

SCARAMUZZA, D., 2008: Omnidirectional Vision: from Calibration to Robot Motion Estimation. – PhD thesis, ETH Zurich.

SCHMEING, B., LÄBE, T. & FÖRSTNER, W., 2011: Trajectory Reconstruction Using Long Sequences of Digital Images From an Omnidirectional Camera. – 31th DGPF Conference.

SCHNEIDER, J., SCHINDLER, F., LÄBE, T. & FÖRSTNER, W., 2012: Bundle Adjustment for Multi-camera Systems with Points at Infinity. – ISPRS Annals of Photogrammetry, Remote Sensing and the Spatial Information Sciences, 22nd Congress of the ISPRS: 75–80, Daniel Verlag, Bonn.

ZOMET, A., WOLF, L. & SHASHUA, A., 2001: Omni-rig: Linear Self-recalibration of a Rig with Varying Internal and External Parameters. – Eighth IEEE International Conference on Computer Vision: 135–141.

Address of the Authors:

M.Sc. JOHANNES SCHNEIDER, Rheinische-Friedrich-Wilhelms-Universität Bonn, Institut für Geodäsie und Geoinformation, Photogrammetrie, Nußallee 15, D-53115 Bonn, Tel.: +49-228-73-2901, e-mail: johannes.schneider@uni-bonn.de.

Prof. Dr.-Ing. Dr. h. c. mult. WOLFGANG FÖRSTNER, Josef-Schell-Str. 34, D-53121 Bonn, e-mail: wf@ipb.uni-bonn.de

Manuskript eingereicht: März 2013

Angenommen: Mai 2013