# Simultaneous Calibration of Multiple Cameras and Generation of Omnidirectional Images

José M. Pacheco [1], Antonio M.G. Tommaselli [1]

[1] Department of Cartography, School of Sciences and Technology, São Paulo State University (UNESP), Presidente Prudente, São Paulo, 19060-900, Brazil, {jm.pacheco, a.tommaselli}@unesp.br

**Keywords:** Omnidirectional images, Multi-camera systems, Camera calibration, Stitching.

**Abstract**

Omnidirectional images are increasingly being used in various areas, such as urban mapping, virtual reality, agriculture, and robotics. These images can be generated by different acquisition systems, including multi-camera systems, which can acquire higher-resolution images. Stitching techniques are often used and can be suitable for non-metric applications, but rigorous photogrammetric processing is recommended when having more accurate requirements. The main challenges related to this kind of product are the system calibration and the generation of the final omnidirectional images. When using multi-camera systems, the displacement of the cameras' perspective centres can affect the generation of the omnidirectional images and the resulting accuracy. A common approach to minimising the resulting parallax error is to establish a value for the projection cylinder radius as close as possible to the object's depth. This work proposes a highly accurate simultaneous calibration technique for multiple camera systems using self-calibrating bundle adjustment with constraints of stability of the relative orientation parameters. These parameters are later used to generate a projecting cylindrical surface, maintaining the original camera perspective centres and relative orientation angles. The experiments show that using constraints improved both the calibration results and the final omnidirectional images. Residual mismatches between points in overlapping areas are subpixel.

## 1. Introduction

Omnidirectional images are becoming common in several applications, such as urban mapping, cadastre, robotics and agriculture. These images are also called panoramic and can provide an immersion impression. Omnidirectional images cover the entire viewing sphere or a partial vertical field of view and can be acquired by different sensors (Benosman and Kang, 2011). There has been a long tradition of developments and use of panoramic images in photogrammetry, with the development of acquisition devices, mathematical models, data processing techniques and applications (Luhmann, 2008).

Examples of omnidirectional imaging sensors are catadioptric systems, fisheye lens cameras, rotating scan cameras, and systems with multiple cameras (sets of perspective cameras with divergent optical axes) (Richter et al., 2013).

Fisheye lens cameras use lenses with a coverage angle of approximately 180°. One disadvantage of this type of sensor is the lower image resolution at the field of view limits (Schneider et al., 2009). Rotating scan cameras are also used for photogrammetric purposes. The sensor is based on a CCD line sensor mounted on a turntable parallel to the rotation axis (Scheibe et al., 2001).

Systems with multiple cameras allow the acquisition of images with higher resolution and less blurring than previously mentioned systems while maintaining geometric and radiometric quality (Khoramshahi et al., 2019). However, generating a single omnidirectional image requires several processing steps to stitch the individual images together. Image stitching is the process of combining multiple images into larger compositions covering a 3D scene (Capel (2004); Blaser et al. (2018); Wang and Yang (2020)). Due to assembling constraints, the camera's perspective centres are displaced, causing parallaxes in the overlapping areas between two cameras. This can result in discontinuities in the fused images, especially when objects are close to the camera.

Image acquisition systems for omnidirectional images can be installed on different platforms, either stationary or moving. A system with multiple cameras has the potential to be used in applications that require fast data capture rates, such as mobile mapping. An important issue when using multi-camera systems with mobile platforms is image georeferencing. Accurate determination of the orientation and position of each panorama is required for mapping applications. This orientation can be done by indirect or direct georeferencing techniques or integrating them. Direct georeferencing is the process of finding the position and orientation (Exterior Orientation Parameters - POEs) of the captured images in a global frame using additional sensors such as GNSS receivers and IMU sensors (Hassan et al. (2007); Cavegn et al. (2016)).

The generation of a single image from sequences of partially overlapping images is often known as mosaicking or stitching and requires several processes (Capel (2004); Wang and Yang (2020)): interest points extraction and matching in the overlapping areas, registration and blending. Image registration with 2D transformations, such as affine or perspective (homography), is often used for mosaicking image sequences. However, it is unsuitable for omnidirectional images since they do not cope with this complex imaging geometry. Instead, rigorous modelling with collinearity equations with correction of systematic errors provides more accurate results (Lichti et al. (2015); Rau et al. (2016); Detchev et al. (2018); Campos et al., (2018); Lichti et al. (2020)).

Several commercial systems are available and have been extensively used to acquire massive terrestrial datasets for several applications (Teledyne FLIR (2024a); Lynn (2023); EarthSense (2021)). A fundamental step in the generation of the omnidirectional image with rigorous modelling is the system calibration, which includes the calibration of the internal orientation parameters (IOPs) and the relative orientation parameters (ROP), also known as mounting parameters (Tommaselli et al. (2013); Habib et al. (2014)). Having rigorous and accurate calibration parameters will enable final accurate results in the ground space coordinates.

This paper presents an experimental essay on the calibration of a multi-camera system using stability-based constraints

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-1-2024-183-2024 | © Author(s) 2024. CC BY 4.0 License.

(Tommaselli et al., 2013). This rigorous calibration process estimates suitable parameters to be used in the process of omnidirectional image generation. The generation of the omnidirectional images relies on the estimated parameters that are used to project the individual images onto a cylindrical surface, showing that preserving the original geometry of the cameras improves the accuracy of the generated omnidirectional image.

## 2. Background

### 1.1 Calibration of multi-camera systems

Camera calibration is a technique for the estimation of the Interior Orientation Parameters (IOP) of a camera, usually the focal length, the principal point coordinates and lens distortion coefficients (Brown (1971); Clarke and Fryer (1998)). Nowadays, field calibration methods are frequently used, and they use common points observed in several images to simultaneously estimate the IOPs, the EOPs and ground coordinates with bundle adjustment. The collinearity equations (Equations 1), extended with the distortion models, are the most suitable models for this task.

$$x_f - x_0 + \delta x_r + \delta x_d + \delta x_a + f\frac{m_{11}(X-X_0)+m_{12}(Y-Y_0)+m_{13}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} = 0$$
$$y_f - y_0 + \delta y_r + \delta y_d + \delta y_a + f\frac{m_{21}(X-X_0)+m_{22}(Y-Y_0)+m_{23}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} = 0 \quad (1),$$

in which $x_f$, $y_f$ are the image coordinates of a point measured in the image; $X$, $Y$, $Z$ are the coordinates of the same point in the object space; $m_{ij}$ are the rotation matrix elements; $X_0$, $Y_0$, $Z_0$ are the coordinates of the camera perspective centre (PC); $x_0$, $y_0$ are the principal point coordinates; $f$ is the camera focal length and $\delta x_i$, $\delta y_i$ are the effects of radial, decentering lens distortion and affinity model (Habib and Morgan, 2003).

Equations (1) are linearised and used with the Least Squares Method to estimate the IOPs, the EOPs, and the ground coordinates of the points measured in the images simultaneously. The definition of the reference frame requires a minimum of seven constraints, which can be done by introducing the ground coordinates of some points or imposing absolute constraints in seven parameters (Clarke and Fryer, 1998). Correlations among parameters must be avoided with specific configurations (Brown (1971); Fraser (1997)).

Calibrating a set of rigidly attached cameras has been approached with several techniques. Usually, the IOPs and EOPs are computed with bundle adjustment, and then the Relative Orientation Parameters are estimated from the EOPs. Additional constraints on the ROPs have been suggested in several papers. He et al. (1993) proposed constraint equations with equal ROPs for different acquisitions. These absolute constraints, however, can cause an increase in the residual values since they do not consider the uncertainty in the EOPs indirectly determined. King (1995) developed two stereo-camera calibration approaches: with ROP constraints and with modified collinearity equations.

Tommaselli et al. (2009) proposed using constraints in the stability of ROPs in the calibration of dual-convergent camera systems, allowing random variations for these elements. The same research group used this concept for the generation of virtual images acquired by a dual-convergent aerial camera system (Tommaselli et al. (2010); Tommaselli et al. (2013)). The constraints are based on the hypothesis that a pair of cameras rigidly assembled remains stable. Thus, the ROPs should be the same, except for slight variations caused both by physical changes and the errors in the estimation of the EOPs. The Relative Rotation (RR) matrix can be estimated from the rotation matrix of both cameras with Equation 2:

$$R_{RR} = R^{C1}(R^{C2})^{-1} \quad (2)$$

In which $R_{RR}$ is the relative rotation matrix, $R_{C1}$ and $R_{C2}$ are rotation matrices for cameras 1 and 2, at a given instant.

The base components between the cameras' perspective centres can also be considered stable during the acquisition. Considering that these elements can be computed from the EOPs at an instant $t$ and, similarly, at the instant $t+1$, six constraints equations can be written. The first three constraints are a function of the lower triangular part of the resulting matrix from the difference between $R_{RR}^{(t)}$ and $R_{RR}^{(t+1)}$ (Equations 3, 4 and 5), and a second group of three stability constraints are written from the base components which are computed by Equations 6. The base components' stability constraints are written as Equations 7.

$$G_1 = \left(r_{21}^{c1}r_{11}^{c2} + r_{22}^{c1}r_{12}^{c2} + r_{23}^{c1}r_{13}^{c2}\right)^{(t)} - \left(r_{21}^{c1}r_{11}^{c2} + r_{22}^{c1}r_{12}^{c2} + r_{23}^{c1}r_{13}^{c2}\right)^{(t+1)} = 0 + v_{1c} \quad (3)$$

$$G_2 = \left(r_{31}^{c1}r_{11}^{c2} + r_{32}^{c1}r_{12}^{c2} + r_{33}^{c1}r_{13}^{c2}\right)^{(t)} - \left(r_{31}^{c1}r_{11}^{c2} + r_{32}^{c1}r_{12}^{c2} + r_{33}^{c1}r_{13}^{c2}\right)^{(t+1)} = 0 + v_{2c} \quad (4)$$

$$G_3 = \left(r_{31}^{c1}r_{21}^{c2} + r_{32}^{c1}r_{22}^{c2} + r_{33}^{c1}r_{23}^{c2}\right)^{(t)} - \left(r_{31}^{c1}r_{21}^{c2} + r_{32}^{c1}r_{22}^{c2} + r_{33}^{c1}r_{23}^{c2}\right)^{(t+1)} = 0 + v_{3c} \quad (5)$$

$$\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = R^{C1}\begin{bmatrix} X_0^{C2} - X_0^{C1} \\ Y_0^{C2} - Y_0^{C1} \\ Z_0^{C2} - Z_0^{C1} \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}^{(t)} - \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}^{(t+1)} = 0 + \begin{bmatrix} v_{4c} \\ v_{5c} \\ v_{6c} \end{bmatrix} \quad (7)$$

In Equations 3, 4, 5 and 7, the 0 value is considered a pseudo-observation with a variance calculated by covariance propagation from the values admitted for the variations in the ROPs, and $v_{ic}$ is a residual in the constraint equations.

The self-calibrating bundle adjustment with collinearity equations and the constraints of stability in the relative orientation (Equations 3, 4, 5 and 7) were implemented in C/C++ language on the CMC (Calibration of Multiple Cameras) program (Tommaselli et al., 2013), that uses the Least Squares combined model with constraints (Mikhail and Ackerman, 1976).

A similar approach was used by Lichti et al. (2015) for the calibration of a dual fluoroscopic imaging system. Detchev et al. (2018) developed a two-step approach with a conventional self-calibrating bundle adjustment in the first step, followed by the estimation of the camera mounting parameters from the EOPs estimated in the first step. Jarron et al. (2019) and Lichti et al. (2020) used a similar concept of constraints in the stability of relative orientation in calibrating a Ladybug5 multi-camera. Khoramshahi et al. (2019) developed a calibration technique for multiprojective cameras and applied the concept to a Panono camera with 36 individual cameras. Some commercial systems, like LadyBug, use their own patented calibration technique which details are not disclosure.

### 1.2 Generation of omnidirectional images from multi-camera images

Omnidirectional images can be generated from individual images acquired by multi-camera systems with a processing pipeline based on relative orientation parameters connecting the cameras and the IOPs of each camera. The relative orientation can be determined during assembly and refined during camera calibration when the IOPs are also determined.

A well-known system developed by Teledyne FLIR (2024a) is the Ladybug camera, which uses a processing sequence that can be summarised as (Teledyne FLIR (2024b)): (1) Image acquisition and colour processing, (2) Rectification of each image to compensate for the influence of lens distortion, (2) Stitching, which is the projection of the image coordinates of the rectified image to a 3D or 2D polygon mesh, according to the type of projection; this mesh is defined from calibration data (3) Blending, which is the adjustment of pixel values in the overlap areas to minimise the discontinuity effect in the seam-lines. According to Teledyne FLIR (2024b), depth variation cannot be treated, and a fixed range radius is used for calibration and omnidirectional image generation. Jarron et al. (2019) and Lichti et al. (2020) assessed the calibration and generation of omnidirectional images with a Ladybug5 camera. They used the proprietary pipeline and their in-house developed software, concluding that their rigorous calibration produced better results.

Omnidirectional images generated by multiple sensors will have some imperfections due to the displacement of the perspective centres of the individual images (Kwiatek and Tokarczyk, 2015). These offsets will generate parallaxes in the overlap area between images. In an ideal panoramic camera (Figure 6. a), all the projection centres should coincide with the centre of the projection sphere (Ji et al., 2014). However, the camera lenses must be displaced in practice, as shown in Figure 1. b. There are options to achieve a single perspective centre from multiple cameras but at the cost of using mirrors, such as the patented system from FullView (FullView, 2023).

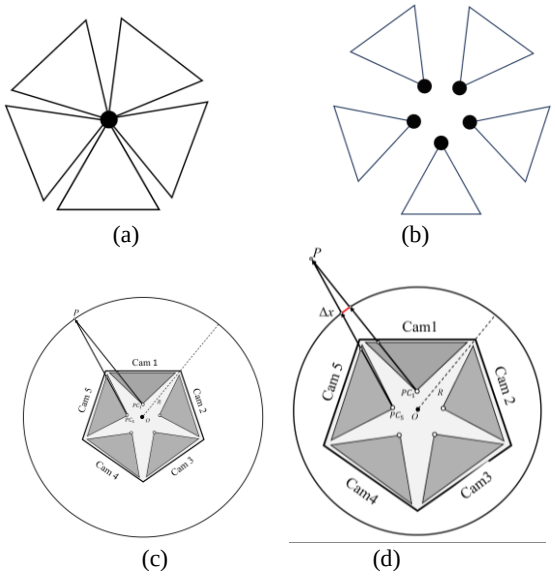



**Figure 1**. (a) An ideal camera geometry for omnidirectional viewing, (b) perspective centres in a real multi-camera arrangement (Kwiatek and Tokarczyk, 2015), (c) Top view of a multi-camera and the projection to a cylindric surface with radius R passing through a point P, (d) Parallax error when $P > R$.

There are several approaches for the generation of omnidirectional images from multiple images. Some approaches consider the projection centres as coincident and are used to create images for visualisation only, with the resulting errors corrected by stitching and blending (Ghosh and Kaabouch, 2016). These simplifications, however, will cause deviations in the projecting rays and should not be used in applications requiring high accuracy. Rigorous approaches aim to minimise the errors caused by multiple viewpoints by applying an algorithm that requires setting a radius value for the sphere ($R$) (or cylinder), which should be as close as possible to the operational object to camera distance (Kwiatek and Tokarczyk, 2018). The choice of a suitable radius has a crucial impact on final accuracy. In Figure 1. c, the radius ($R$) of the cylinder is equal to the distance from point $P$ to the cameras' perspective centres. Therefore, no parallax effect will appear when projecting both images to the cylinder surface. However, in Figure 1.d, point $P$ is located at a distance greater than the cylinder's radius, thus generating a parallax effect ($\Delta x$).

The omnidirectional image can be created by projecting the individual images to a common surface, according to a projection model, including Rectilinear, Cylindrical, Spherical or Stereographic.

## 3. Methodology

### 3.1 T-Map multi-camera system

T-Map is a mobile mapping system developed by Sensormap company as part of an industrial project funded by Fapesp (T-MAP Sensormap, 2016), comprising five optical cameras integrated with an Inertial Navigation System (INS). Figure 2.a depicts the system mounted over a car rooftop. The individual cameras were assembled with the INS, and thus, all mounting angles and offsets are previously known and will be refined in the camera calibration step. Figure 2.b shows the coordinate systems involved in the assembly and the initial angle values.

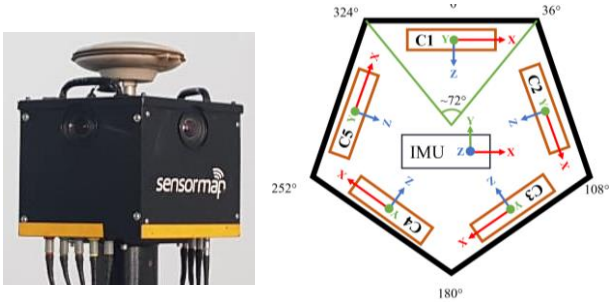



**Figure 2.** (a) The T-Map system mounted on a car rooftop; (b) The cameras and the INS assembly of the T-Map system with reference systems and initial angle values.

| | Offsets with respect to C1 (cm) | Rotations with respect to the INS (°) |
|---|---|---|
| C2 | 0.16 | -72° |
| C3 | 0.26 | -144° |
| C4 | 0.26 | 144° |
| C5 | 0.16 | 72° |

**Table 1.** Offsets between the cameras and the master camera (C1) and rotations of the cameras with respect to the Inertial Navigation System.

The multi-camera system is positioned over the vehicle's rooftop and integrates further electronics for data acquisition. The cameras were mounted with high precision in positions and

angular orientations at previously defined angles so that complete coverage of the scene in 360° is feasible. The offsets between each camera with the camera chosen as master (camera 1 – C1) and the rotations of the cameras with respect to the INS are shown in Table 1. These values are considered initial values and will be refined during the calibration process.

The T-Map system was mounted with five cameras, model JAI GO-5000C-PGE, from the JAI manufacturer (JAI, 2023). Table 2 describes the main specifications of this camera model.

| Camera model | Sensor width (mm) | Sensor height (mm) | Focal length (mm) | Photodetector Size (mm) |
|---|---|---|---|---|
| JAI GO-5000C-PGE | 12.8 | 10.24 | 6.2 | 0.005 |

**Table 2.** Specifications of the cameras used in the T-Map system.

### 3.2 Multi-camera calibration

Camera calibration was performed with images acquired on a terrestrial calibration field populated with ARUCO-type coded targets (Figure 3). Initial values for the 3D coordinates were obtained from previous surveys with a standard deviation of 2 mm.

**Figure 3.** Camera calibration field with coded targets and examples of images acquired by camera 1.

In order to avoid propagating these errors to the IOPs and ROPs being estimated, self-calibrating bundle adjustment was used with a minimum set of 7 constraints. The reference frame was defined by imposing absolute constraints on the ground coordinates of three points (targets' corners). Table 3 presents the 3 points used to set the reference system in the object space with the corresponding standard deviations used to compute the constraints. The horizontal distance between two points (6623-6522) was measured using a large precision calliper. Point 6623 was set as the origin; 3D coordinates of points 6623 and 6522 were set as absolute constraints, whilst only the Z component of point 6630 was fixed. For this third point, the X and Y components were weighted with a standard deviation of 3 cm, as presented in Table 3.

| Control Point | X (m) | Y (m) | Z (m) | σX (m) | σY (m) | σZ (m) |
|---|---|---|---|---|---|---|
| 6623 | 0 | 0 | 0 | 1 E-13 | 1 E-13 | 1 E-13 |
| 6522 | 1.155 | 0 | 0 | 1 E-13 | 1 E-13 | 1 E-13 |
| 6630 | 0 | 1.161 | 0 | 0.03 | 0.03 | 1E-13 |

**Table 3.** Coordinates of the control points and the respective standard deviations used to apply constraints in the self-calibrating bundle adjustment.

In each station, four images from each camera with different rotations were captured to reduce the correlation between the interior and exterior orientation parameters (See Figure 3). With these settings, 30 acquisition stations were selected, and 120 images from each camera were acquired, totalling 600 images. Images with few tie points from the sky, for instance, were eliminated, totalling 590 images for the bundle adjustment.

The acquired images were initially processed with Agisoft Metashape (Agisoft, 2023) to extract tie points automatically and estimate initial values for both IOPs and EOPs. The image coordinates and estimated values of IOPs and EOPs were exported with an in-house developed script. These exported values were then processed in the CMC software (Calibration with Multiple Cameras) (Tommaselli et al., 2013) using several configurations, including the stability constraints. This software uses self-calibrating bundle adjustment with the options of constraints in Ground Control Points, IOPs, EOPs and stability of the ROPs.

Camera 1 (C1) was selected as the reference camera (master) being used in all relative calibrations with the other cameras (C2, C3, C4, C5). The calibration trials were performed with pairs. A first set of trials with pair C1-C2 were used to establish the most reliable set of constraints, and the same values were used for the other pairs. The set of constraints will be detailed in the section of experiments.

### 3.3 Omnidirectional image generation

The generation of the omnidirectional image can be done by stitching the images generated by the sensor, which can produce some deviations in the original bundle of rays, depending on the technique used to project and merge the images. The approach proposed in this paper preserves the multi-camera geometry by rigorously projecting the perspective rays from each camera to a cylinder and then unwrapping them to an omnidirectional image (Figure 4) with later blending. The first step is the definition of the cylinder geometry and cell sizes. The cylinder radius is defined by the user, usually as the average distance between the camera to the objects of interest. From these values and the average cameras' focal lengths, an average scale and the pixel size in the resulting omnidirectional image are defined. The centre of the cylinder will be the perspective centre of the camera selected as master, as well as its focal length, which will be used for the calculations ($f_{ref}$). The cylinder axes are defined as right-handed, Z axis pointing upwards, and the X-axis coincident with the X-axis of the reference camera. The rotation matrices to be used in the projection step are defined from the relative orientation calculated in the calibration step. Equation 8 expresses the rotation matrix ($R_{cil}^{c_k}$) relating a camera *ck* to the cylinder coordinates system.

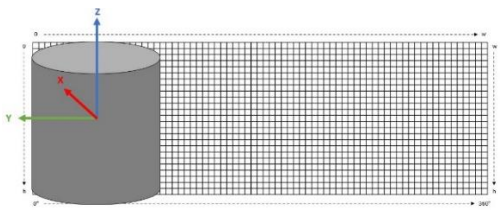



**Figure 4**. The omnidirectional image and the projecting cylindrical surface.

$$R_{cil}^{c_k} = (R_{c_k}^{c_{ref}})^{-1}.R_{cil}^{c_{ref}} \quad (8)$$

In which $R_{cil}^{c_{ref}}$ is the rotation matrix of the cylinder related to the reference camera (master); $R_{c_k}^{c_{ref}}$ is the rotation matrix between the camera and the reference camera (master).

An empty image is created, and each pixel is associated with a 3D coordinate at the cylindric surface. This 3D coordinate is

projected onto each image; the RGB values are interpolated with a bilinear transformation and assigned to the pixel cell in the omnidirectional image. These projections are performed with the collinearity equations using the IOPs of each camera, the rotation matrices defined by Equations 2, and the camera positions related to the cylinder reference system. The offset values (translations from the cylinder centre to each camera perspective centre) are defined as the base elements computed in the multi-camera calibration. For the master camera, the offsets are null. In the first step, the five images are generated independently, and later, they are fused with a blending technique. This strategy ensures that the original geometry of the forming bundle of rays are preserved, except in the overlap areas in which the blending technique can produce some ghost images. It is important to highlight that the use of these omnidirectional images for 3D measurements can be done with forward intersection. The coordinates of each camera perspective centre must be used to achieve a higher accuracy.

## 4. Experiments And Results

### 4.1 Calibration of the T-MAP Multi-camera

The determination of tie points with Metashape engine produced 3231 points in the object space for the C1-C2 camera pair. A similar number of tie points were extracted for the other camera pairs. The initial value for the standard deviation of the image observations was set as 0.5 pixel size, and the a priori sigma naught ($\sigma_0$) was set as the unit value. Several different configurations of the stability constraints were used in a set of trials with C1-C2 and C1-C5 camera pairs; the main features of these experiments are shown in Table 4. In experiment A, stability constraints were not applied, which means that a conventional self-calibrating bundle adjustment with two cameras and two sets of IOPs was considered. The other experiments (B, C, D and E) used constraints of stability in the relative orientation parameters. The ROPs are allowed to have random variations defined empirically and presented in Table 4 as the standard deviations for the relative rotation angles and base component offsets. The constraints applied in Experiment B allow variations of 100 mm and 1º for the ROPs, which can be considered very high variations and probably have no effect on the ROPs. The variations to be admitted in the ROP were then progressively reduced, up to 1 mm and 0.1º for Experiment E.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| Base **(mm)** | N | 100 | 4 | 2 | 1 |
| Angles (°) | N | 1 | 0.2 | 0.14 | 0.1 |

**Table 4.** Standard deviations admitted for the stability constraints in the self-calibrating bundle adjustment.

In Table 5, the estimated standard deviations of three IOPs (camera focal length and principal point coordinates) for both cameras (Camera 1 – CAM 1 and Camera 2 – CAM 2) and the *a posteriori* sigma naught (variance of the unit weight observation - $\hat{\sigma}_0$) are presented. It can be observed that the estimated sigma naught always remained lower than the a priori variance in all relative calibrations, which means that the standard deviations assigned to the observations (0.5 pixels) could be even lower. The variation of the *a posteriori* sigma naught in all experiments was very small, which means that imposing stability constraints did not significantly augment the image residuals. The magnitudes of the estimated standard deviations of the IOPs were also slightly reduced, which means that imposing ROP stability did not disturb the bundle adjustment with the values used. These values indicate that the IOPs are being estimated with high precision since the standard deviations are approximately 1/10 of the pixel size. This can be credited to the dense network of images with suitable geometry, their configurations and the high number of tie points used. Results for the other camera pairs were similar and will not be presented due to space limitations.

| | | A | B | C | D | E |
|---|---|---|---|---|---|---|
| *CAM 1* | $\sigma_f$ *(pixels)* | 0.1 | 0.1 | 0.09 | 0.09 | 0.08 |
| | $\sigma_{xo}$ *(pixels)* | 0.1 | 0.1 | 0.1 | 0.09 | 0.09 |
| | $\sigma_{yo}$ *(pixels)* | 0.07 | 0.06 | 0.06 | 0.06 | 0.05 |
| *CAM 2* | $\sigma_f$ *(pixels)* | 0.1 | 0.09 | 0.08 | 0.08 | 0.07 |
| | $\sigma_{xo}$ *(pixels)* | 0.09 | 0.09 | 0.09 | 0.08 | 0.08 |
| | $\sigma_{yo}$ *(pixels)* | 0.06 | 0.06 | 0.06 | 0.06 | 0.05 |
| | a posteriori sigma $\hat{\sigma}_0$ | 0.740 | 0.737 | 0.740 | 0.744 | 0.748 |

**Table 5.** Estimated standard deviations of the IOPs for Camera 1 and Camera 2 from the experiments with stability constraints and resulting a posteriori sigma naught.

The estimated standard deviations of the IOPs for all cameras were analysed, and the values are similar to those achieved for camera 1. Due to space limitations, these values will not be included in this paper. In all cases, the effects of the estimated IOP errors were lower than the nominal image measurement error. The calibration of the multi-camera system was performed in pairs due to the current limitation of the CMC program. Thus, the interior orientation parameters of Camera 1 (C1) were estimated four times since it was used as the master camera in all relative calibrations.

Table 6 presents the values of IOPs of Camera 1 and the standard deviations estimated in the bundle adjustment with Camera 2. Table 6 also shows the mean values of IOPs estimated with all pairs and the standard deviations with respect to the mean values. The mean value of the IOPs for Camera 1 and the standard deviations were computed to analyse the differences when using different camera pairs.

| **IOPs** | **Camera 1 IOPs** | **Mean value of the IOPs with all camera pairs** | **Standard deviation** |
|---|---|---|---|
| focal (mm) | 6.24241 | 6.24256 | 0.00028 |
| $x_0$ (mm) | -0.06569 | -0.06626 | 0.00052 |
| $y_0$ (mm) | -0.01080 | -0.01035 | 0.00035 |
| $k_1$ ($mm^{-2}$) | -3.303E-03 | -3.303E-03 | -3.301E-03 |
| $k_2$ ($mm^{-4}$) | 5.384E-05 | 5.384E-05 | 5.376E-05 |
| $k_3$ ($mm^{-6}$) | -2.393E-07 | -2.393E-07 | -2.383E-07 |
| $p_1$ ($mm^{-1}$) | -1.040E-04 | -1.040E-04 | -1.140E-04 |
| $p_2$ ($mm^{-1}$) | 1.900E-05 | 1.900E-05 | 2.400E-05 |

**Table 6.** IOPs of Camera 1 and the respective standard deviations estimated in the bundle adjustment; the mean values of IOPs estimated with all pairs and the standard deviations with respect to the mean values.

The analysis of Table 6 concludes that the IOPs of Camera 1 were estimated similarly in the four calibrations pairs. In this context, to determine which group of parameters will be used for Camera 1, the standard deviations of the mean parameters were analysed and compared with the standard deviations of the calibration that obtained the lowest value of these standard deviations, in this

case, the values of the calibration pair of Camera 1 with Camera 3.

In all calibration results, the correlations between the interior orientation parameters were below 50%, except for the correlation between $p_1$ and $x_0$, which varied between 85 % and 89 %. The correlations between the radial distortion parameters were high, which is common in all calibration processes. The correlations between the interior and exterior orientation parameters were low due to the variation of the distance from the camera stations to the targets of the calibration field in conjunction with the change in the rotation angles of the images of the same camera. The correlation values between the focal length and the exterior orientation parameters of all cameras were below 50%. The correlations between the principal point coordinates and the exterior orientation parameters for all cameras were below 65%.

The average values of estimated standard deviations of the cameras' perspective centres were below 1.6 mm for the pair C1-C2 and similar for the other pairs. The average value of the pixel size in object space units was, on average, 2.5 mm, which means that the camera perspective coordinates were estimated with subpixel precision. The standard deviations of the attitude parameters (ω, φ and κ) calculated among all pairs of cameras were approximately 0.02º, 0.012 º, and 0.016 º, respectively.

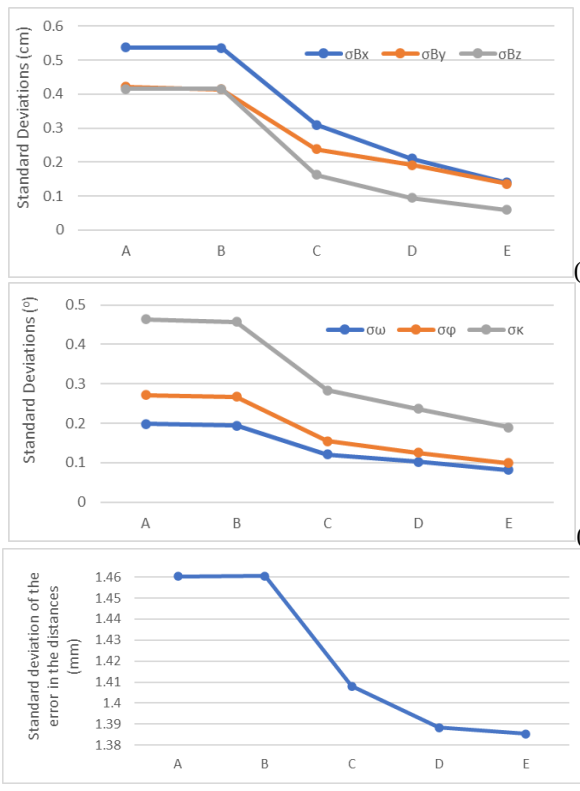



**Figure 5.** (a) The standard deviations of the base components; (b) the standard deviations of the relative rotation angles; (c) the standard deviation of discrepancies in control distances.

After bundle adjustment, the base components (camera offsets $b_x$ $b_y$, $b_z$ and relative rotation angles) were computed from the EOPs for each image pair. The average value of 42 pairs and the standard deviations were then computed. The standard deviations of the base components are presented in Figure 5.a, and the standard deviations of the relative rotation angles are presented in Figure 5.b. It can be seen that the first set of constraints (Experiment B) did not change significantly the variations of the ROP from pair to pair. Reducing the standard deviations allowed for the constraints caused the reduction of the ROPs standard deviations, which means that the final averaged value could be used later to produce omnidirectional images.

An external quality control was performed with distances between some existing control points. The distances between 34 control points (17 distances) were computed and compared with the reference values. The standard deviations of the average values of the discrepancies were computed for each experiment. These values are presented in Figure 5.c and reflect the improvement in the solution after imposing the stability constraints. The improvement is small (submillimetre) but shows that the constraints positively affect the self-calibrating bundle adjustment. A second quality control was performed by comparing the estimated base distances with the corresponding values directly measured during the mechanical mounting process. The discrepancies between the estimated base length and the directly measured values are presented in Table 7. It can be seen that the discrepancies are approximately at the millimetre accuracy.

| Camera Pair | C1/C2 | C1/C3 | C1/C4 | C1/C5 |
|---|---|---|---|---|
| Mean discrepancy (mm) | 1.6 | 1.8 | 1.1 | 1.1 |

**Table 7.** The discrepancies between the estimated base lengths and the directly measured values.

### 4.2 Omnidirectional image generation from calibration data

The omnidirectional images were generated according to the steps described in section 3.3. The IOPs and ROPs for each experiment were used to generate the individual projected images. The values used for the generation of omnidirectional images are presented in Table 8. For each of the five images from the cameras, empty images are created and populated with the procedure already described in section 3.3. Figure 6 depicts an example of original and projected images.

| Radius (mm) | Focal length (mm) | Width (pixels) | Height (pixels) |
|---|---|---|---|
| 5000 | 6.2433 | 7845 | 2048 |

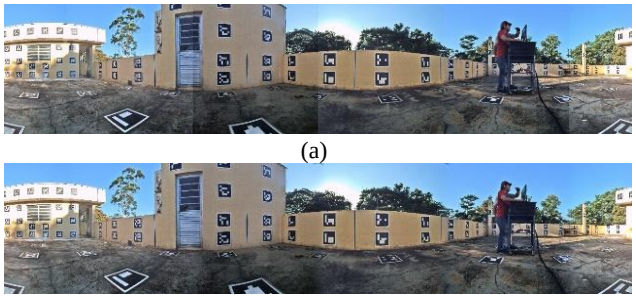
**Table 8.** Values used to generate the omnidirectional image.


**Figure 6.** Example of two images projected onto the cylindric surface (a) Camera 1 and (b) Camera 2.

The images were then fused and cropped to generate the final omnidirectional image with two options: (1) by defining a vertical seam line according to a predefined image area and (2) by merging them with a *feathering* effect on the overlap area.

Figure 7 illustrates an example of the two options. The *feathering* option can create smooth transitions, but some challenges still exist: depending on the object depth, parallax creates a double mapping (ghosts), which will cause problems when identifying

and measuring the object. On the other hand, without *feathering*, there is a clear seamline, but no errors are caused by feathering.



(a)

(b)

**Figure 7.** Omnidirectional images: generated without (a) and with *feathering* (b).

The effects of the multi-camera calibration with stability constraints in the generation of the omnidirectional images were assessed by analysing the coordinates of common points in the overlap area of images from Camera 1 and Camera 2 and for Camera 1 and Camera 5, for the five experiments. A small window with 131x420 pixels was selected in both images, and keypoints were extracted with SIFT operator. In both cases, more than 80 keypoints were defined and compared for both overlap areas. The discrepancies in the image coordinates and the resulting magnitude vector were computed. Figure 6 depicts the resulting magnitude of the discrepancies from corresponding points in the overlap areas after projection for each experiment. It is clear that the parameters generated with stability constraints produced smaller discrepancies and, thus, more accurate omnidirectional images. Preliminary results with forward intersection showed that accuracy in the object space is also increased. Details of photogrammetric intersection in this case, along with results are not given in this paper due to the lack of space and will be reported in a future work.

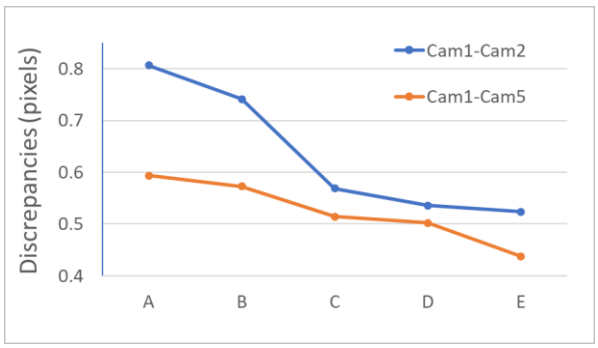



**Figure 8.** Average magnitude of discrepancies between corresponding points in the overlap areas com Camera 1 with Camera 2 and Camera 5 after projection for each experiment with stability constraints.

The effect of selecting the cylinder radius was not assessed in these experiments. It is expected that points close to the camera will produce parallax artefacts in the image, and the discrepancies in overlap areas will be higher than those observed in these experiments.

## 5. Conclusions

The main aim of this work was the development and assessment of processes for omnidirectional image generation from images of a multi-camera system. The rationale for this proposal is to preserve the original geometry of the projecting rays avoiding deviations caused by stitching and blending techniques. A fundamental step in this technique is the rigorous and simultaneous calibration of pairs of cameras. Additionally, it was shown that imposing constraints of stability in the relative orientation parameters improves the results.

Rigorous modelling of the projection of the original perspective images to a cylindrical surface was also presented and assessed. The results indicate that subpixel errors can be achieved in the fusion of overlapping images for depth distances similar to the cylinder radius. Depending on the point depth with respect to the camera, parallax effects can be noted in the overlapping areas. Nevertheless, using the omnidirectional images with original perspective centre positions preserves the geometry of the photogrammetric rays, with a potential for very accurate measurements. This technique can be used with any multicamera system, provided that the original images without stitching are available.

The assessment of the approach for the generation of 3D models by photogrammetric intersection will be presented in a future work, as well as the assessment of direct georeferencing as a way to determine the exterior orientation of the omnidirectional images. Comparing the proposed solution with other techniques based on stitching will also be performed in future work.

## Acknowledgements

This study was funded by the São Paulo Research Foundation FAPESP (Grants: 2016/07930-1 and 2021/06029-7) and by the National Council for Scientific and Technological Development CNPq (Grant: 303670/2018-5) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Code 88887.310313/2018-00 and 88881.310314/2018-01).

## References

Agisoft, 2023. MetaShape. URL https://www.metashape-la.com/en/home/ (accessed 2.11.23).

Benosman, R., Kang, S.B., 2011. *Panoramic vision: sensors, theory, and applications*. Springer, New York.

Blaser, S., Cavegn, S., Nebiker, S., 2018. Development of a portable high performance mobile mapping system using the robot operating system. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*. IV–1, 13–20. https://doi.org/10.5194/isprs-annals-IV-1-13-2018

Brown, D.C., 1971. Close-Range Camera Calibration. *Photogramm. Eng.* 37, 855–866.

Campos, M.B., Tommaselli, A.M.G., Marcato Junior, J., Honkavaara, E., 2018. Geometric model and assessment of a dual-fisheye imaging system. *Photogramm. Rec.* 33, 243–263. https://doi.org/10.1111/phor.12240

Capel, D., 2004. *Image Mosaicing and Super-resolution.* Springer London, London. https://doi.org/10.1007/978-0-85729-384-8

Cavegn, S., Nebiker, S., Haala, N., 2016. A systematic comparison of direct and image-based georeferencing in challenging urban areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLI-B1, 529–536. https://doi.org/10.5194/isprs-archives-XLI-B1-529-2016

Clarke, T.A., Fryer, J.G., 1998. The development of camera

calibration methods and models. *Photogramm. Rec.* 16, 51–66.

Detchev, I., Habib, A., Mazaheri, M., Lichti, D., 2018. Practical In Situ Implementation of a Multi-camera Multisystem Calibration. *J. Sens.* 2018, 12.

EarthSense, 2021. EarthSense. EarthSense. URL https://www.earthsense.co/home (accessed 1.4.24).

Fraser, C.S., 1997. Digital camera self-calibration. *ISPRS J. Photogramm. Remote Sens.* 52, 149–159.

FullView, 2023. FullView: Technology. URL http://www.fullview.com/technology.html (accessed 1.2.24).

Ghosh, D., Kaabouch, N., 2016. A survey on image mosaicing techniques. *J. Vis. Commun. Image Represent.* 34, 1–11. https://doi.org/10.1016/j.jvcir.2015.10.014

Habib, A., Detchev, I., Kwak, E., 2014. Stability Analysis for a Multi-Camera Photogrammetric System. *Sensors* 14, 15084–15112. https://doi.org/10.3390/s140815084

Habib, A.F., Morgan, M.F., 2003. Automatic calibration of low-cost digital cameras. *Opt. Eng.* 42, 948. https://doi.org/10.1117/1.1555732

Hassan, T., Ellum, C., Nassar, S., Wang, C., El-Sheimy, N., 2007. Photogrammetry for Mobile Mapping. *GPS Solut.* 18.

He, G., Novak, K., Feng, W., 1993. Stereo camera system calibration with relative orientation constraints, in: *Proceedings of SPIE Conference.* Presented at the SPIE Videometrics, Boston, EUA, pp. 2–8.

JAI, 2023. JAI | Products. URL https://www.jai.com/products (accessed 12.28.23).

Jarron, D., Lichti, D.D., Shahbazi, M.M., Radovanovic, R.S., 2019. Multi-camera panoramic imaging system calibration, In: *Proceedings of the 11th International Conference on Mobile Mapping*. Shenzhen, China.

Ji, S., Shi, Y., Shi, Z., Bao, A., Li, J., Yuan, X., Duan, Y., Shibasaki, R., 2014. Comparison of Two Panoramic Sensor Models for Precise 3D Measurements. *Photogramm. Eng. Remote Sens.* 80, 229–238.

Khoramshahi, E., Campos, M.B., Tommaselli, A.M.G., Vilijanen, N., Mielonen, T., Kaartinen, H., Kukko, A., Honkavaara, E., 2019. Accurate Calibration Scheme for a Multi-Camera Mobile Mapping System. *Remote Sens.* 11, 2778.

King, B.R., 1995. Bundle adjustment of constrained stereo pairs-Mathematical models. *Geomatics Res. Australasia*, 63, 67–92.

Kwiatek, K., Tokarczyk, R., 2018. Photogrammetric 3D Measurements Based on Immersive Panoramas. *Geomat. Environ. Eng.* 12, 55. https://doi.org/10.7494/geom.2018.12.4.55

Kwiatek, K., Tokarczyk, R., 2015. Immersive photogrammetry in 3D modelling: Fotogrametria immersyjna w modelowaniu 3D. *Geomat. Environ. Eng.* 9, 51. https://doi.org/10.7494/geom.2015.9.2.51

Lichti, D.D., Jarron, D., Tredoux, W., Shahbazi, M., Radovanovic, R., 2020. Geometric modelling and calibration of a spherical camera imaging system. *Photogramm. Rec.* 35, 123–142. https://doi.org/10.1111/phor.12315

Lichti, D.D., Sharma, G.B., Kuntze, G., Mund, B., Beveridge, J.E., Ronsky, J.L., 2015. Rigorous Geometric Self-Calibrating Bundle Adjustment for a Dual Fluoroscopic Imaging System. *IEEE Trans. Med. Imaging*, 34, 589–598. https://doi.org/10.1109/TMI.2014.2362993

Luhmann, T., 2008. A historical review on panorama photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 34.

Lynn, 2023. Revolutionizing Data Collection: Mosaic's 360º Cameras & SmartDelta in Prague - Mosaic51. Mosaic51.com. URL https://www.mosaic51.com/industry/revolutionizing-data-collection-mosaics-360º-cameras-smartdelta-in-prague/ (accessed 1.4.24).

Mikhail, E.M., Ackerman, f, 1976. *Observations and least squares.* University Press of America Inc, New York.

Rau, J.Y., Su, B.W., Hsiao, K.W., Jhan, J.P., 2016. Systematic calibration for a backpacked spherical photogrammetry imaging system. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLI-B1, 695–702. https://doi.org/10.5194/isprsarchives-XLI-B1-695-2016

Richter, K., Mader, D., Seidl, K., Maas, H.-G., 2013. Development of a geometric model for an all-reflective camera system. *ISPRS J. Photogramm. Remote Sens.* 86, 41–51. https://doi.org/10.1016/j.isprsjprs.2013.09.002

Scheibe, K., Korsitzky, H., Reulke, R., Scheele, M., Solbrig, M., 2001. EYESCAN - A High Resolution Digital Panoramic Camera, in: Klette, R., Peleg, S., Sommer, G. (Eds.), *Robot Vision, Lecture Notes in Computer Science.* Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 77–83. https://doi.org/10.1007/3-540-44690-7_10

Schneider, D., Schwalbe, E., Maas, H.-G., 2009. Validation of geometric models for fisheye lenses. *ISPRS J. Photogramm. Remote Sens.* 64, 259–266. https://doi.org/10.1016/j.isprsjprs.2009.01.001

Teledyne FLIR, 2024a. Spherical Vision Systems | Teledyne FLIR. URLhttps://www.flir.com/iis/machine-vision/spherical-vision-systems/ (accessed 1.2.24).

Teledyne FLIR, 2024b. Overview of the Ladybug Image Stitching Process. URL https://www.flir.com/support-center/iis/machine-vision/application-note/overview-of-the-ladybug-image-stitching-process/ (accessed 3.29.24).

T-MAP Sensormap, 2016. Desenvolvimento do sistema de mapeamento terrestre T-Map. URL https://bv.fapesp.br/pt/auxilios/98215/desenvolvimento-do-sistema-de-mapeamento-terrestre-t-map-e-analise-de-aplicacoes-do-projeto-mapkite/ (accessed 12.28.23).

Tommaselli, A., Galo, M., Bazan, W., Ruy, R., Junior, J.M., 2009. Simultaneous calibration of multiple camera heads with fixed base constraint. *Proc. 6th Int. Symp. Mob. Mapp. Technol.*

Tommaselli, A., Galo, M., Marcato, J.J., Ruy, R., Lopes, R., 2010. Registration and fusion of multiple images acquired with medium format cameras. *Can. Geomat. Conf.* 2010 Int. Symp. Photogramm. Remote Sens. Comm.

Tommaselli, A.M., Galo, M., De Moraes, M.V., Marcato, J., Caldeira, C.R., Lopes, R.F., 2013. Generating virtual images from oblique frames. *Remote Sens.* 5, 1875–1893.

Wang, Z., Yang, Z., 2020. Review on image-stitching techniques. *Multimed. Syst.* 26, 413–430. https://doi.org/10.1007/s00530-020-00651-y