

POSSIBLE APPROACHES FOR PROCESSING OF SPHERICAL IMAGES USING SFM

David Zahradník and Jakub Vynikal

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Geomatics, Thákurova 7, Praha 6, Czech Republic; david.zahradnik@fsv.cvut.cz, jakub.vynikal@fsv.cvut.cz

ABSTRACT

Spherical cameras are being used more frequently in surveying because of their low cost and possibility to process spherical images with conventional Structure from Motion (SfM) method. Fish-eye lenses are now frequently found on modern 360° cameras, allowing for the capture of the entire scene and subsequent model reconstitution by a non-photogrammetrist specialist. Gyrospherics are a feature of cameras that ensure image stabilization to reduce blur when the camera is moving. This feature allows it to capture moving scenes in time-lapse mode. However, two main factors - hardware parameters and software algorithms - affect the quality of the images that are captured. The 360° camera design allows for a variety of data processing methods by SfM. These images were created using multiple image sensors and lenses from each 360° camera. These images can be processed individually or by applying rules to define relationships between images. Also, images from cameras can be stitched and processed with a spherical camera model. In this paper is proposed processing methods of data from 360° cameras and estimated accuracy of each method.

KEYWORDS

Spherical camera, 360 cameras, SfM, Agisoft

INTRODUCTION

Spherical cameras are mainly used for virtual tours to visualize the first-person perspective from a computer. These applications are used both indoors and outdoors for navigation. The advantage of this type of camera is that it captures the overall view by stitching together multiple images with a large field of view (FOV). Most low-cost 360° cameras on the market contain two image sensors with fisheye lenses.

Spherical images can also be used for photogrammetry. Their capabilities in this field are currently tested in indoor modelling, heritage documentation, street mapping and many more use-cases. While the quality of low-cost spherical camera footage can't compete with DSLR cameras, the omni-directional field of view is quite useful for alignment or connecting areas modelled with conventional frame cameras.

Some problems arise from the use of spherical images, the construction of the cameras, the non-identical centre of projection and the excessive distortion of the fisheye lenses, which leads to blurring at the edges of images. The distortion of composite panoramic images cannot be characterized with a standard distortion model (Brown etc.) [1] for planar/fisheye lenses for SfM. The inability to correct for lens distortion can be a serious problem in accurate reconstruction. One solution to mitigate this problem is to split the source images between the two original fisheye cameras and evaluate the alignment along with their internal parameters. In this paper, a comparison of different processing methods of data from spherical cameras is proposed.

The reason for the comparison is the increasing use of spherical cameras in surveying [3], [4], [5]. It is also a cheaper alternative to mobile mapping using simultaneous localisation and mapping (SLAM) technology [6], [7]. The accuracy of the results of SfM from spherical images is comparable to mobile mapping with SLAM in places with good light conditions and abundant features with distinct texture (unlike white walls). The advantages of using spherical imagery, such as low cost, fast data collection and guaranteed overlap, provide an interesting alternative for mobile mapping. [3], [8], [9], [10].

Video capture is the optimal option to achieve the same mobility as SLAM technology. Conventional photography is the better option for image quality, but as time intervals between photos are too long, better overlap can be more desirable than higher image quality. To achieve sufficient overlap, it is better to set the frequency of the images to 0.5s. Previous experience shows that images with an interval larger than 1s can be hard to process [7]. Insta360 cameras also offer Interval camera mode, which takes regular photos every fixed time interval. However, the shortest possible interval is 3s, which as mentioned is way too long.

Next problem with spherical images is the stitching step. Stitching has an impact on results of SfM. Some software does not differentiate the project centre of images, this causes inaccuracies of SfM methods [11]. Other possibilities allow calibrating spherical cameras to achieve interior orientation parameters (IOPs) and exterior orientation parameters (EOPs) of each individual's lens [12]. In this paper, we explore some of the possible reconstruction approaches with regards to stitching, evaluating their accuracy, reliability and convenience. One of the questions answered is whether to process the footage already stitched as a spherical photo (with discontinuity and resulting error along the seamline), or to process them unstitched as fisheye cameras.

METHODS

Material

Two devices were used for data collection. The first device, made of inexpensive components, will be referred to as the camera rig later in the text. Four Raspberry Pi HQ cameras with fisheye lenses were used for the camera rig. The camera sensor used is an IMX477R with 12.3 MP resolution. The pixel size of the camera sensor is 1.55 μm . The lens used has a variable focal length from 2.8 mm to 12 mm. The corresponding field of view ranges from 43° to 125°. During the measurement, the lenses were set to 3.4 mm focal length. The cameras are controlled sequentially by the Raspberry Pi4B to capture one image at a time. The device is not able to capture all four images at the same time because the port on the Raspberry control unit only communicates with one camera at a time. To take a picture, one camera must be active and the others must be set off.

The second device was a commercial 360° camera Insta360 One RS with a 360° module that covers the entire 360x180° scene and consists of two fisheye sensors. The spherical images of the 360° module have resolution of 18 MPx and size of 6080x3040 px. The pixel size of the camera sensor is 1.2 μm , lens with 7.2 mm equiv. focal length and F2.0. The camera is equipped with a gyroscope to achieve image stabilization in all conditions.



Fig. 1 – Camera rig (left) Insta360 One RS (right)

Case study

Testing took place in the laboratory of calibrations of surveying equipment. The laboratory is part of the Faculty of Civil Engineering at CTU in Prague and the laboratory is 27m long and 7.5m wide. Coded targets were set around the laboratory on the walls to be used as control and check points.

First operation was to define the reference system of coded targets. Concrete columns with joints for total stations are placed in line at the centre of the laboratory. One of the columns was set as the origin of the reference system and the link to the farthest column was set as the interlaced horizontal axis. Coded targets were measured with Leica TCR 403 from columns to obtain coordinates with accuracy of 3 mm+2 ppm in distance measurement and 3' in angle measurement. Final spatial accuracy for a target was 3mm. Three targets were used as control points, remaining 27 targets were used as check points for estimations of accuracy

Measurement methods

The measurement method utilized in this paper contradicts the photogrammetry recommendation for SfM. Data collected primarily along the centreline of the laboratory are unsuitable for SfM processing. SfM requires high overlap between images and, for improved accuracy, images captured from different angles. Methods in this paper presume high image overlaps, but images are in the same perspective with quite large leaps between them. Another drawback is that the target (the room) is not surrounded by images and images are not captured from different angles. This can lead to insufficient data for successful alignment. Reconstructed objects are floor, walls and ceiling.

Two approaches were used for data acquisition by camera rig with larger and smaller distances between rig positions.

Positions with larger gaps were predefined by columns in the laboratory. Distance between columns is from 1 to 1.5 meters. Camera images were captured from two angles of the camera rig rotated by 45°, thus the total number of images from one position was 8. The reason for this is that only 4 images with such a small overlap were not sufficient to be processed by SfM. Ten positions with eight images each were captured during measurement.

Second approach used positions 0.5 meters apart with each position containing 4 images. Direction of the images aimed to the middle between the axes. Main quadrant is defined as the direction of measurement and as the axis of the reference system. A total of twenty-two positions were measured in the line defined by the columns.

Data from the Insta360 camera was captured using a time-lapse function every 0.5 seconds, with all recording flow stabilization turned on. Each image was corrected by pitch and roll values from the gyroscope. Camera was mounted on a helmet and the two approaches were used: the same line track was used as for the camera rig and it was used as the maximum allowed track in the

laboratory for capturing time-lapse video. Camera lenses were heading perpendicular to the capturing direction to minimize lens distortion issue on closed objects in the scene.

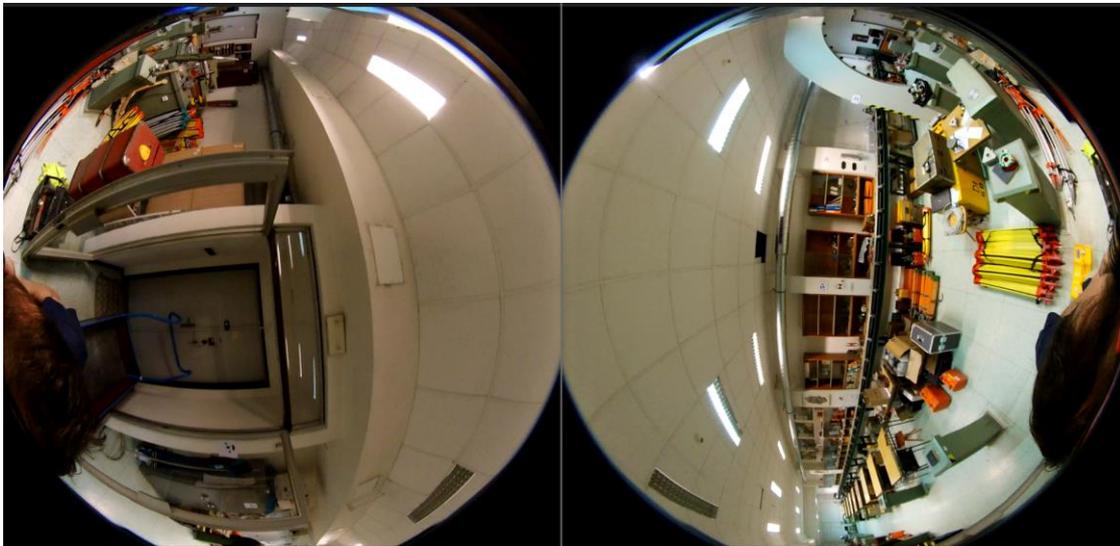


Fig. 2 – unprocessed Insta360 fisheye images

Processing methods

Data was processed using four methods to estimate its accuracy. Two image datasets were captured and each dataset was processed using these methods. Methods differ in the choice of camera settings: free, station, master-slave camera and spherical. Each method is standardly processed in the Agisoft Metashape workflow.

Video-sequence from time-lapse mode on Insta360 One RS is in insv format, which is mp4 video format with metadata from gyroscope and camera information for stitching. To extract spherical frames, the files must be stitched in Insta360 Studio software. After stitching, users are able to export 360 time-lapse videos to mp4 format to extract the individual frames. Time-lapse mode does not allow saving images like images, instead images are saved to video sequence with 29.97 fps. To use the individual images from each fisheye sensor, insv files must be renamed to mp4 format. After renaming, it is possible to extract individual frames from each sensor.

Photos from the Raspberry HQ camera are processed like jpg images. Agisoft Metashape was used for spherical images stitching and panorama export, with images from the same position in one folder, which is set as a station.

Individual (free) camera images are processed with the “Align photos” function in Agisoft Metashape. Each image is individually represented by position and orientation. Alignment accuracy was set to High, detected points were set to 40 000 key points per image, filtered to 5 000 tie points. Station cameras contain images taken from the same position. Images from the same station have the same image center. Orientation is independent and is estimated by aligned photos. To use the station option, images from the same position must be separated into subfolders. Subfolder in the project is marked as a station, so in the alignment process camera positions from the subfolder are the same.

Master-slave camera approach works with images taken at the same time, but compared to the station option, image centres are not the same. Master-slave option allows to predefine relative position and orientation between images on one camera rig. This option is suitable for multiple camera systems with two and more cameras with different offsets. One of the cameras is set as master and the others as slave. Orientation and position of slave cameras is related to the master camera. With calibrated relative position and orientation, the processes of Structure from Motion

should be faster. To use the master-slave approach, images of each camera must be separated into subfolders and must be in alphabetical order. Agisoft Metashape automatically allows importing photos in such folders as a multi-camera system.

Spherical images are the product of stitching multiple individual images. The mathematical model for spherical images of Agisoft Metashape [21] consists only of:

$$u = \frac{1}{2}w + \frac{w}{2\pi} \frac{X}{Z} \quad v = \frac{1}{2}w + \frac{w}{2\pi} \frac{Y}{\sqrt{X^2+Y^2}} \quad (1)$$

where (X, Y, Z) are the point coordinates in the local camera coordinate system. As the spherical model doesn't take in account any other variable, distortion and dispositions of the image centre can only be eliminated before stitching in the fisheye model.

All images were set to fisheye type, except for spherical images. The lens on the Insta360 One RS and the camera rig have high distortion, which shows up as chromatic aberration and other lens flaws. Images taken with the Raspberry HQ Camera captured in the lab at 4 meters distance result in sampling distance of 1.6mm. At the same distance, the Insta360 One sampling distance was 3.6mm.

Three control points from measured coded targets were used as references for each approach. Control points were evenly spaced across the lab, and two of them were positioned at the beginning and at the end of the laboratory wall, which is parallel to the capturing line track. Due to lens distortion, the perpendicular wall on the capturing line track was not chosen. The third control point was placed in the middle of the wall opposite the wall with the other two control points.

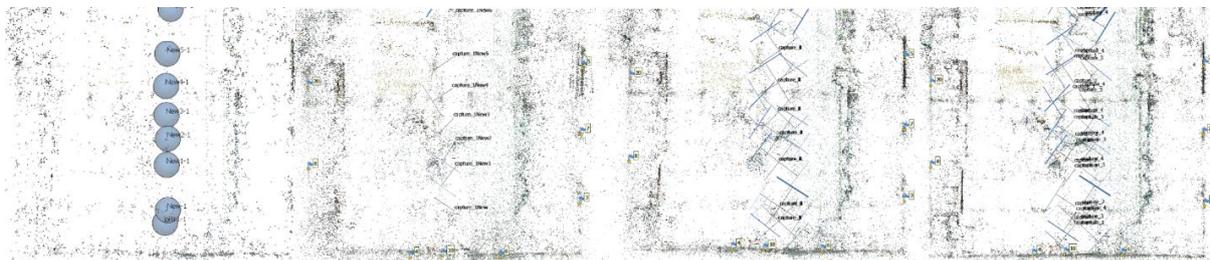


Fig. 3 – Camera rig (spherical, master-slave, station, free, left to right)

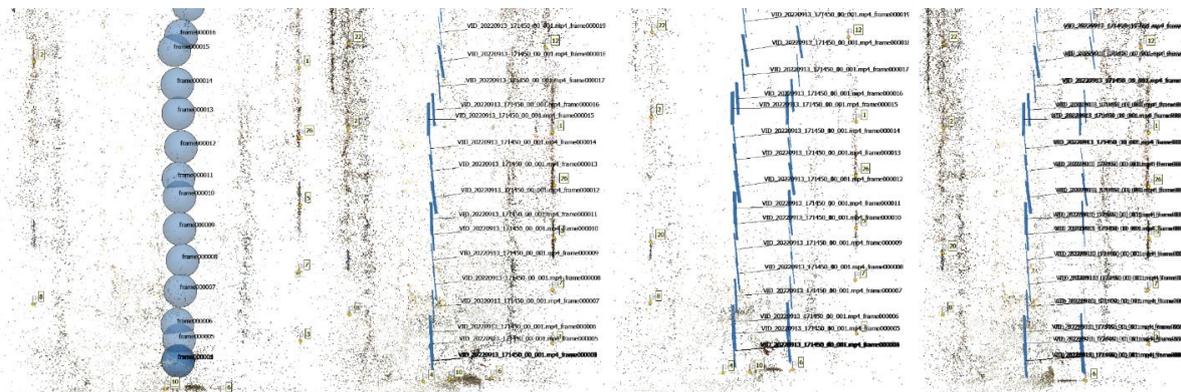


Fig. 4 – Insta360 (spherical, master-slave, station, free, left to right)

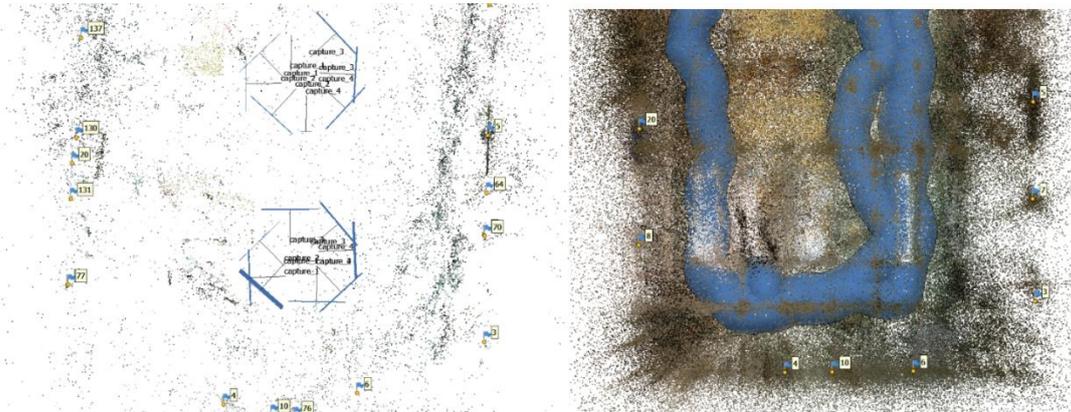


Fig. 5 – Raspberry HQ camera - 8 images station (left), Insta360 one rs - long track (right)

RESULTS

The data obtained from Raspberry PI HQ cameras and Insta360 One RS were examined and validated by position differences on checkpoints. Check points were always marked at least on four close images. Coded maskers had to be manually targeted on some images because of small image resolution. Large lens distortion and small image resolution make it difficult to automatically detect coded markers.

Data was processed on the Lenovo Legion 5 Pro laptop. The laptop contains AMD Ryzen 7 5800H with Radeon Graphics with 32GB ram memory and Nvidia GeForce RTX 3070 Laptop GPU. Process time is related to the working station and can be analysed only relatively. Spherical images from raspberry HQ Camera are 16000x8000 with 18 MB storage size. Original pictures are 4056x3040 with 24MB storage size. Spherical images from Insta360 One RS are 5760x2880 with 12 MB store size. Individual images from each sensor are 2880x2880 with 7 MB storage size. For a small dataset the results are only informative and in the future on a larger dataset the time-consuming data will be rediscovered.

Tab. 1 - Methods processing time

	Methods	Matching time [s]	Alignment time [s]	Memory usage [MB]
Raspberry HQ	spherical	78	3	5500
	master-slave	43	25	530
	station	40	32	530
	free	43	32	650
Insta360	spherical	46	24	400
	master-slave	54	23	720
	station	51	31	450
	free	46	26	400
Raspberry HQ	8x long distance	29	60	550
Insta360	long track	2160	480	1300
	upscale	450	5	12000

The phases of image matching, tie point (TP) extraction, and camera position estimation are primarily responsible for the accuracy on checkpoints.

Dataset from Raspberry HQ camera with eight images per station with longer distance between positions was not properly processed. Cameras closer to walls were discarded from position estimation. Long distance between positions degraded accuracy to more than 1 meter. This method with high overlap from one position is not suitable for further processing and usage.

Dataset from Insta360 with long track and spherical camera methods is comparable with spherical methods from capturing line track. More photos from the same spherical camera did not increase the accuracy of testing objects.

Master-slave and station methods have better accuracy than the spherical method from the Raspberry HQ camera. The cause lies in the stitching phase, where the stitching process in Agisoft Metashape assumes the same image centre. As a result, spherical pictures are incorrectly joined from different angles. The issue is depicted in the figure below.



Fig. 6 – Bad stitching of images from camera rig

Free camera option from Insta360 One RS dataset was not able to process all at one. Small overlap between images from the same position at part of image with high lens distortion excluded one sensor from the Estimation position phase.

Worst results of Raspberry HQ Camera were caused by lenses with changeable focal length. The cameras were not secured with the precisely same focal length. Small variance of focal lens degraded four same images sensors and lenses to four cameras with a bit different camera IOPs.

Best results are shown by the dataset which was processed with a free camera option. Accuracy of the free camera method is 2 times better than other methods from Raspberry HQ Camera. In the case of Insta360, there are not very high differences between processing methods except for the station option. Spherical stitching in insta360 studio is at a good level, so the camera model must count with different images centres and lens distortion. The station option groups images to the same image centre and causes worse accuracy.

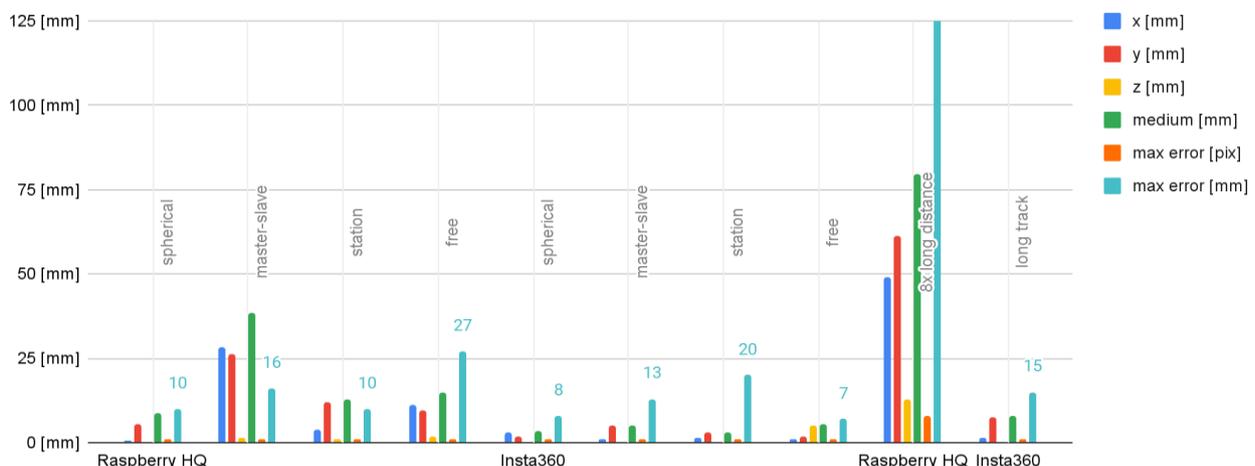


Fig. 7 – Accuracy test – check points

Values x, y and z in Figure 7 and 8 describe errors in individual axes. Medium and max values correspond with x, y and z values.

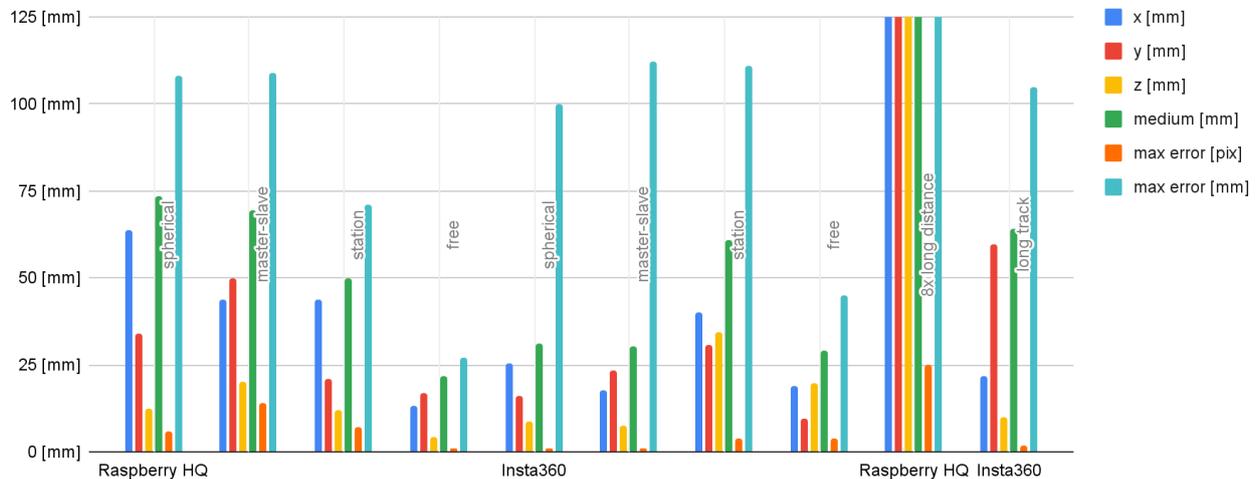


Fig. 8 – Accuracy test – control points

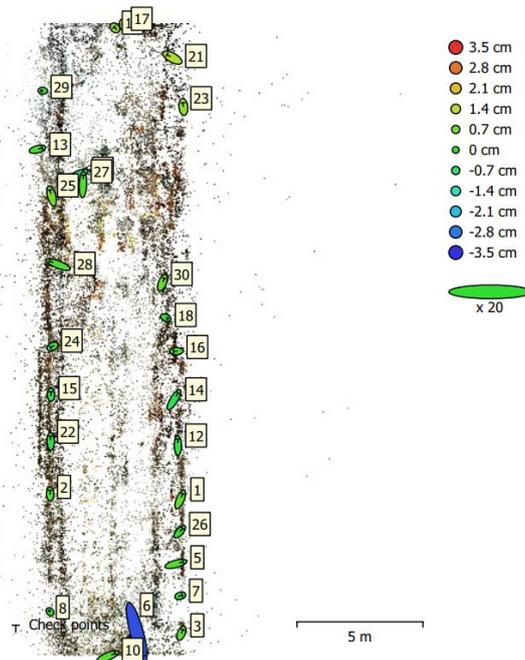


Fig. 9 – Checks and control points position

Check points of each method have the same relative value to corresponding accuracy. The worst accuracy was reached by the checkpoint on the perpendicular wall to capture line track. Hedging of lenses was on a parallel wall with a captured line track, so the checkpoint on this wall has better accuracy (Figure 9).

CONCLUSION

Images were processed with several approaches: free, station, master-slave and spherical option. Free option estimates the position and orientation of images individually. The station option assigns one image centre to images taken from a single position. Master-slave option defines relative position and orientation of images from one camera (master) and applies it to other cameras (slaves). Spherical images created or directly captured were processed as spherical in Agisoft Metashape software.

Best results are achieved with the free method when image position and orientation is independently estimated. Maximal error from camera rig dataset was 27 mm. Better results are achieved by the Insta 360 dataset with a maximum error of 45 mm. Other methods achieve worse results. For the Insta360 camera of the other methods, the maximum error at the control points is twice as large due to the failure to process a series of photos from one side of the camera. For the master-slave method with the camera rig, the maximum error on the control points is three times larger due to differences in IOP's. The individual lenses only had an approximate set focal length. Insta360 has median control point errors of around 30 mm for all methods except station, where the error is doubled. This is caused by the non-identical center of projection of images. Camera rig median control point error for free method is 22 mm. Other methods have more than twice the error due to errors in focal length evaluation.

Check points placed on the edge of images with high lens distortion show worse accuracy than primarily focused check points in image centers.

Mobile mapping with a 360° camera is possible in good light conditions and scenes without homogenous texture, i.e. white walls. 360° cameras with high resolution in good conditions can achieve 3 cm spatial accuracy on a 5–6-meter distance. As the distance increases, the sampling distance decreases and so does the accuracy.

ACKNOWLEDGEMENTS

This work was supported by the Grant Agency of the Czech Technical University in Prague, grants SGS23/052/OHK1/1T/11 and SGS23/050/OHK1/1T/11

REFERENCES

- [1] Luhmann, Thomas, et al. *Close-range photogrammetry and 3D imaging*. Walter de Gruyter, 2013.
- [2] Pavelka, Karel, Cinzia Pappi, and Karel Pavelka jr. (2021) "Modern Possibilities of Documentation and Replication of Archaeological Finds." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 46: 531-538.
- [3] Castanheiro, L. F., et al. "3D Reconstruction of Citrus Trees Using an Omnidirectional Optical System." 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS). Ieee, 2020.
- [4] Abate, D., Toschi, I., Sturdy-Colls, C., Remondino, F., 2017. A low-cost panoramic camera for the 3d documentation of contaminated crime scenes. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W8, pp. 1–8.
- [5] Zhao, C. (2021). Creating point clouds, textured meshed models and orthophotos with Weiss AG Civetta - 230 magapixels 360° HDR camera. Technical Report available on the Internet, <https://docplayer.net/209971609-Civetta-agisoft-metashape.html>, last accessed December 2021, 5 pages
- [6] Teppati Losè, Lorenzo, Filiberto Chiabrando, and Fabio Giulio Tonolo. "Documentation of complex environments using 360 cameras. The Santa Marta Belltower in Montanaro." *Remote Sensing* 13.18 (2021): 3633.
- [7] Pavelka, K., et al. (2022) "Analysis of data joining from different instruments for object modelling." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 43: 853-860.
- [8] Barazzetti, L., et al. "Connecting inside and outside through 360 imagery for close-range photogrammetry." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 42 (2019): 87-92.

- [9] Barazzetti, L., M. Previtali, and F. Roncoroni. "3D modelling with the Samsung Gear 360." 2017 TC II and CIPA-3D Virtual Reconstruction and Visualization of Complex Architectures. Vol. 42. No. 2W3. International Society for Photogrammetry and Remote Sensing, 2017.
- [10] Barazzetti, L., M. Previtali, and F. Roncoroni. "3D MODELING WITH 5K 360° VIDEOS." 9th International Workshop on 3D Virtual Reconstruction and Visualization of Complex Architectures, 3D-ARCH 2022. Vol. 46. No. 2. International Society for Photogrammetry and Remote Sensing, 2022.
- [11] Kwiatek, K., Tokarczyk, R., 2015. Immersive Photogrammetry in 3D Modelling. Geomatics and Environmental Engineering, Volume 9, Number 2, pp. 51-62.
- [12] Aghayaria, S., Saadatsereshta, M., Omidalizarandi, M., Neumann, I., 2017. Geometric Calibration of Full Spherical Panoramic Ricoh Theta Camera. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 4(1/W1), pp. 237-245.
- [13] Fangi, G. (2017). The book of spherical photogrammetry: Theory and experiences. Edizioni Accademiche Italiane.
- [14] Fangi, G., Pierdicca, R., Sturari, M., & Malinverni, E. S. (2018). IMPROVING SPHERICAL PHOTOGRAMMETRY USING 360° OMNI-CAMERAS: USE CASES AND NEW APPLICATIONS. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(2).
- [15] Schneider, D., & Maas, H. G. (2003). Geometric modelling and calibration of a high-resolution panoramic camera. Optical 3-D Measurement Techniques VI, 2, 122-129.
- [16] Barazzetti, L., Previtali, M., & Roncoroni, F. (2018). Can we use low-cost 360 degree cameras to create accurate 3D models?. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(2).
- [17] Pisa, C., Zeppa, F., & Fangi, G. (2010, October). Spherical photogrammetry for cultural heritage: san galgano abbey, siena, italy and roman theatre, sabratha, libya. In Proceedings of the second workshop on eHeritage and digital art preservation (pp. 3-6).
- [18] Suman, S., Rastogi, U., & Tiwari, R. (2016). Image Stitching Algorithms-A Review. Circulation in Computer Science Vol.1, No.2, pp: (14-18),
- [19] Szeliski, R. (2007). Image alignment and stitching: A tutorial. Foundations and Trends® in Computer Graphics and Vision, 2(1), 1-104.
- [20] Otero, R., Lagüela, S., Garrido, I., & Arias, P. (2020). Mobile indoor mapping technologies: A review. Automation in Construction, 120, 103399.
- [21] Agisoft PhotoScan <http://www.agisoft.com>
- [22] Insta 360 One rs twin edition <https://www.insta360cam.cz/insta360-one-r-twin-edition/>
- [23] Pagani, A., & Stricker, D. (2011, November). Structure from motion using full spherical panoramic cameras. In 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops) (pp. 375-382). IEEE.
- [24] Kwiatek, K., & Tokarczyk, R. (2014). Photogrammetric applications of immersive video cameras. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(5), 211.
- [25] Dlesk, Adam, Karel Vach, and Karel Pavelka. (2020) "Structure from motion processing of analogue images captured by rollei metric camera digitized with various scanning resolution."
- [26] Dlesk, A., K. Vach, and P. Holubec. (2019) "Analysis of Possibilities of Low-Cost Photogrammetry for Interior Mapping." The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 42: 27-31.
- [27] Barazzetti, L. (2022) Metric Rectification of Spherical Images. ISPRS Int. J. Geo-Inf., 11, 248. <https://doi.org/10.3390/ijgi11040248>
- [28] Pagani, Alain, and Didier Stricker. (2011) "Structure from motion using full spherical panoramic cameras." 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)
- [29] Pavelka, K., and E. Matoušková. (2019) "THE CONTRIBUTION OF GEOMATIC TECHNOLOGIES TO BIM." International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.
- [30] Pavelka, Karel, David Zahradník, and Jaroslav Sedina. (2021) "New Measurement Methods for Structure Deformation and Objects Exact Dimension Determination." IOP Conference Series: Earth and Environmental Science. Vol. 906. No. 1. IOP Publishing

APPENDIX A

Table of test accuracy follows in Figure 7 and 8.

	Methods	points	x [mm]	y [mm]	z [mm]	medium [mm]	max error [pix]	max error [mm]
Raspberry	spherical	control	1	6	0	9	1	10
		checks	64	34	13	73	6	108
	master-	control	28	26	1	38	1	16
	slave	checks	44	50	20	69	14	109
	station	control	4	12	1	13	1	10
		checks	44	21	12	50	7	71
	free	control	11	10	2	15	1	27
		checks	13	17	4	22	1	27
Insta360	spherical	control	3	2	0	4	1	8
		checks	26	16	9	31	1	100
	master-	control	1	5	0	5	1	13
	slave	checks	18	23	7	30	1	112
	station	control	1	3	0	3	1	20
		checks	40	31	34	61	4	111
	free	control	1	2	5	5	1	7
		checks	19	10	20	29	4	27

Raspberry	8x long -	control	49	61	13	80	8	130
	distance	checks	1173	528	133	1294	25	2318
Insta360	long -	control	1	8	0	8	1	15
	track	checks	22	60	10	64	2	105