GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 9 • Number 2 • 2015

http://dx.doi.org/10.7494/geom.2015.9.2.51

Karol Kwiatek*, Regina Tokarczyk**

Immersive Photogrammetry in 3D Modelling***

1. Introduction

Panoramic photography, on a massive scale, is widely used by Google Street View (exterior panoramas), Google Business View (company interiors) and Google Photo Sphere (Internet users create their own panoramic virtual tours using tablets and smartphones). These applications enable navigation outside and inside buildings. The navigation is possible by clicking arrows. The user clicks an arrow when they want to move in a specific direction. This method of moving, however, does not create a sense of continuity of movement which can be achieved when panoramas are presented as a video. However, in immersive video, when presenting a specific trajectory, it is not possible to change directions.

A sequence of panoramas presented as a 360° video is defined as an immersive video or 360° video [7]. This form of presentation has not become very popular mainly due to large sizes of files (e.g. 1-minute raw video is a file whose size is about 1–2 GB; the file needs to be compressed in order to be presented on-line) [13]. Immersive video typically consists of 15–30 panoramas/second. So far, such a large amount of data has not been widely explored for photogrammetric purposes. Furthermore, no metric tools have been added to Google panoramic virtual tours and to immersive video.

Immersive video is created by combining images generated by digital cameras that record in various directions. This video is becoming a new form of presenting art projects on interactive exhibitions, in film clips and also in advertising [18].

3D modelling from panoramic images is well know and described in [4, 14, 17, 19]. However, the potential of photogrammetric immersive video has not been examined so far and it is the subject of our experiments.

---

* Cracow University of Economics, Institute of Regional Economics and Spatial Engineering, Chair for Regional Economy, Poland

** AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Department of Geoinformation, Photogrammetry and Remote Sensing, Krakow, Poland

*** Work was done as part of the statutory research of the AGH University No. 11.11.150.949/14 and CUE No. 011/WF-KGR/2014/S/4207

Photogrammetric measurements that benefit from these kinds of images can refer to finding specific measures (e.g. lengths) and do not always apply to the creation of spatial models of the measured objects. However, the product of the photogrammetric measurement is a collection of selected points with designated coordinates that are arranged in a certain way, often with additional information (combined with vectors, point colouring) and this product is called a 3D model. Research presented herein refers to the elaboration of this product.

The images included in the video sequences are of a relatively low resolution. Hence, this factor reduces the accuracy of photogrammetric calculations. Another negative factor is the way in which the immersive video frame is created. It is not a perfect spherical panorama, because images that comprise it do not have a common projection centre. However, the advantage of stitched full-spherical images is their number that make the video and the way in which they are recorded – densely along a trajectory. Central perspective original images have limited field of view which causes problems in finding tie points between images, especially in rooms without much details (e.g. blank walls painted in one colour).

The research on the photogrammetric potential of the immersive video sequence described in this paper focused on the following factors that affected the modelling quality:

- the way in which images are stitched into panoramas,
- the choice of the sphere radius which was declared when creating video frames,
- image file format,
- density of video sequence images.

Experiments have been done on a test field designed in the AGH assembly hall. Photogrammetric measurements and modelling were performed with Agisoft Photoscan Professional software.

2. 3D Modelling with Panoramic Images

Spherical panoramas are created using different methods. Panoramic cameras with the rotating CCD line are quite expensive [14, 17, 19, 20]. These cameras provide an image which corresponds to the assumed projection method, unlike less expensive ways to get panoramas are based on image stitching. Spherical panorama can be recorded because digital SLRs with a wide-angle lens or a fisheye lens are mounted on a rotating panoramic head that rotates around the projection centre. The sequence of images is stitched on the virtual sphere [23]. Another way of taking panoramas is to use dioptric and catadioptric mirrors.

Currently, multi-station geolocalized panoramas are the basis for spatial information systems [2, 9] and for 3D reconstruction [1, 4–6, 14, 21]. Previous studies raise the question whether increasing the number of stations can raise the accuracy and automate the creation of 3D models.

The main accuracy problem of measurements with panoramas depends on the offset of the projection centres of images constituting the panorama. In the real panoramic camera with multiple lenses (Fig. 1a) projection centres do not coincide, which hinders stitching images. In the ideal panoramic camera (Fig. 1b) all projection centres are located exactly at one point which is the centre of the sphere [22], which is impossible to achieve in practice.

Spherical photogrammetry developed by Fangi [4] uses spherical panoramas where no offset of perspective centre is taken into account. The solution is based on the principles of classical photogrammetry: knowing the orientation of many spherical images it is possible to perform a space intersection to measured points that are depicted in at least two panoramas. Spherical photogrammetry serves mainly for architectural inventory and it was used in many international projects [1, 5, 6].

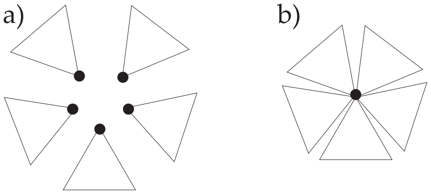

Fig. 1. Projection centres in the real (a) and (b) ideal panoramic camera

In recent years, efforts have been made to generate models from spherical panoramas automatically. These processes were applied in Agisoft Photoscan Professional which, from January 2013 (version 0.9.1), enables spherical panoramas to be imported. From April 2014, this program enables point clouds with spherical images to be generated automatically. Photoscan uses Structure from Motion (SfM) as the algorithm to determine the orientation of the cameras and to position object points. Photos are treated as spherical panoramas and no calibration parameters of spherical cameras can be imported or edited.

3. SfM and Panoramas

For modelling on the basis of images using the SfM many programs were developed, including: PMVS2, CMVS, VisualSFM, RunSFM, Insight3D, Agisoft Photoscan, Pix4D, Microsoft Photosynth, Arc3D, Autodesk 123D Catch, but only Agisoft Photoscan and Photosynth accept panoramas. The use of automatic orientation and restitution of spherical panoramas in Agisoft Photoscan is possible thanks to the research on epipolar lines conducted by [6, 14]. Research on the application of SfM to panoramas are presented, inter alia, in [16]. Chang and Hebert [3] notice that panoramic SfM tracks feature points for longer distance than traditional cameras which have smaller field of view than panoramas. The above-mentioned research provides the basis for the application of immersive video in SfM.

Mobile vehicles with spherical video cameras have been already used for filming purposes [13] but, only recently, omnidirectional systems have been applied to record interiors [15]. 3D modelling from panoramas registered with mobile systems is discussed in [12, 22], where GPS and IMU were used for georeference of images.

In recent years new programmes for SfM were introduced (e.g. 3DF Zephyr Pro, Acute3D). As they develop they may not only use traditional images but also spherical panoramas. Agisoft Photoscan is able to process spherical images. However, this feature has not been described in the literature so far, and there are no guidelines on how to use spherical panoramas for 3D modelling, either.

4. Study on the Accuracy of 3D Modelling from Immersive Video

The Ladybug®3 video camera and its dedicated software (LadybugCapPro) were used to generate the video. The camera consists of six Sony ICX274 cameras with CCD 1/1.8" matrix, with lenses of 3.3 mm focal length, that records the 360-degree view. The sixth camera records the upward view. LadybugCapPro software uses six images to generate a single spherical panorama whose resolution is 5400 × 2700 pixels. Owing to the fact that cameras do not have a common projection centre, it is not an ideal spherical panorama. As the mutual orientation of cameras is known, it is possible to transform images mathematically to a common, centrally located projection centre, by assuming the value of the added, third coordinate of transformed points, called the sphere radius. The third coordinate is equal to sphere radius.

The accuracy of photogrammetric measurements was verified on the test field established in the AGH assembly hall. The field consisted of 102 points distributed spatially, whose coordinates were determined through geodetic measurements using a total station. Some points were marked with code signals recommended by Agisoft developers, others were marked with circular signals and remaining ones were elements of the hall to be measured (Fig. 2).

In the assembly hall recordings were made with a mobile camera that was moved by means of trolley (Fig. 3a). The trajectory of the movement is presented in Figure 3b. 406 spherical panoramas were recorded. The total length of the recordings is approx. 5 minutes (some trajectories include common sections), while the whole recording process with setting relevant parameters (white balance, shutter etc.) took about approx. 15 minutes.

The main problem with processing panoramas in Agisoft is the RAM memory of the computer used for processing data, which is recommended by the software manufacturer. It is recommended to have 64 GB of RAM. Using a computer with 8 GB RAM equipped only with one graphic card (NVidia Quadro FX 880 M), a model was created out of 104 selected video frames, which corresponds to an average 1 frame per second.

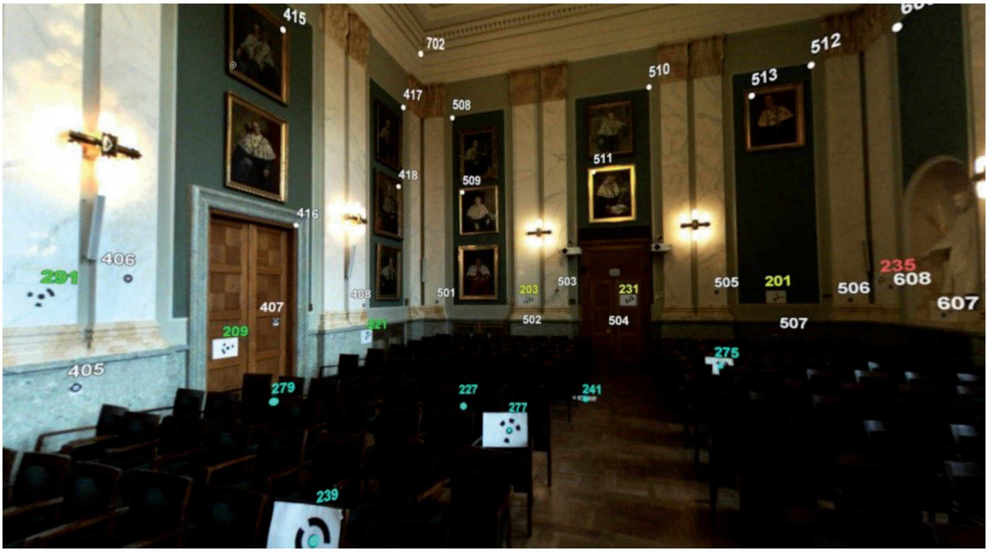

Fig. 2. Natural and signalized (coded and uncoded) targets in the assembly hall

Source: http://kwiatek.krakow.pl/3D_modelling_from_immersive_video/aula_agh

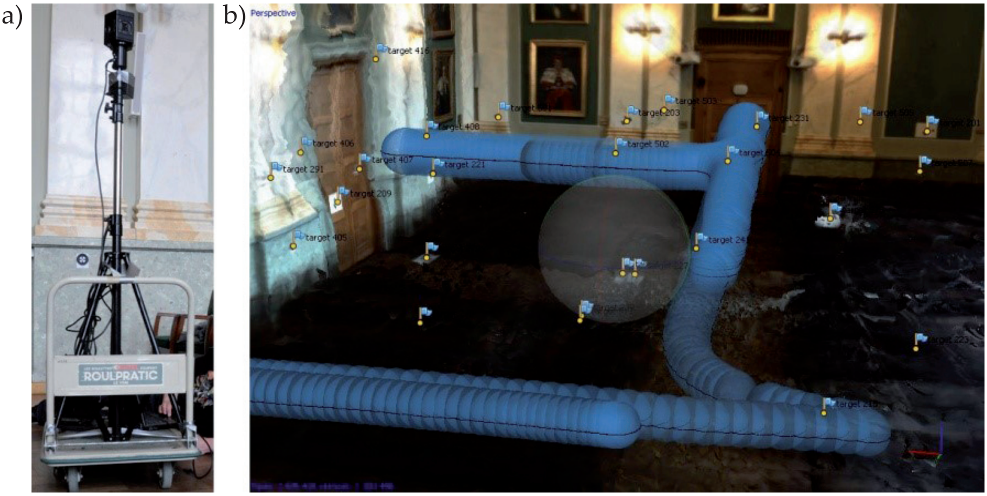

Fig. 3. Panoramic video camera Ladybug®3 was placed on a trolley (a) and then pushed through the hall. The spheres (b) represent recorded individual panoramic frames from the immersive video

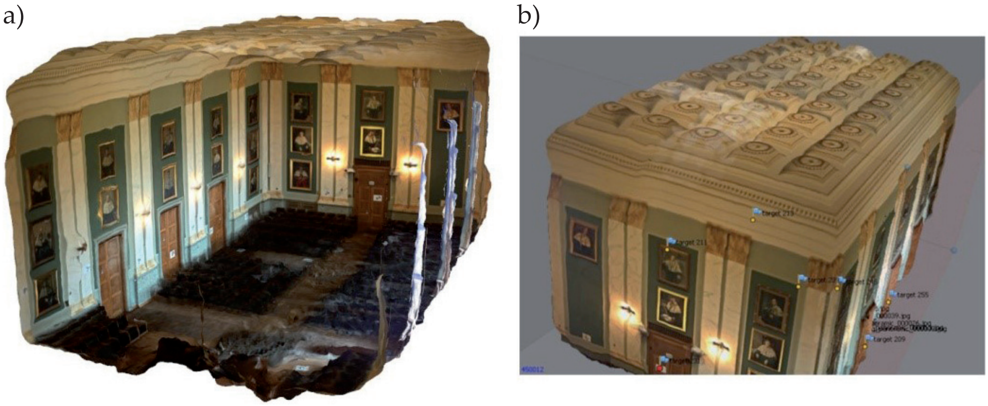

Fig. 4. 3D textured models of the assembly hall generated with Agisoft Photoscan: a) the interior model; b) model "twisted" in order to show the ceiling

Figure 4 illustrates a textured 3D model rendered in Agisoft Photoscan out of 104 panoramas.

After recording the immersive video individual frames were exported with LadybugCapPro, which enables conversion of PGR format (format developed by the producer of Ladybug cameras) into many different video formats (AVI, MP4, MOV) and image formats (TIF, BMP, HDR, PNG, JPG etc.). To find useful settings for the SfM to use panoramic images, different settings for a number of parameters were tried out (for example, the sphere size, feather and image format).

5. Investigations on the Impact of Different Factors

5.1. Impact of the Method for Creating Panoramas

Generation of a single spherical panorama (recorded with a hypothetical virtual camera) out of several images causes problems with stitching them together. This issue has been examined along with its impact on 3D modelling. In the first scenario stitching was done in PTGUI with parameters described in [12] and, although the panoramas were visually correct, 3D modelling from these panoramas provided the poorest results (Tab. 1).

Table 1. Errors in the panoramic bundle adjustment and their relation with the

Method for stitching images	Deviation between reference and calculated coordinates in spatial direction [m]
Photos stitched in PTGUI	0.129
Photos stitched in LadybugCapPro with the feather option	0.012
Photos stitched in LadybugCapPro without feather	0.009

This is obvious, since PTGUI treats single images recorded with a camera as if they were taken from one projection centre. Then, panoramas were stitched in LadybugCapPro with two options: with and without a feather. Table 1 indicates that the option without feather improves 3D modelling error slightly although the stitching line is visible between images that create panorama, which may be relevant when colouring the point cloud or when texturing the model. The difference with the use of feather increases the modelling error by about 0.3cm. The precision of the reference is 5 mm.

5.2. The Influence of the Selection of the Sphere Radius

The second parameter, which has an impact on the 3D modelling is the sphere radius that is selected in the LadybugCapPro menu. The choice of different radiuses mean that panoramas with objects at the chosen distance will be correctly stitched. Errors on control points comprised an additional indicator of accuracy. The calculations were done for three panoramas located along the Y axis, which is parallel to the long side of the hall. The results (Tab. 2, Fig. 5) indicate that best results are obtained when 10 m radius and the "dynamic" option are adopted.

Table 2. The influence of different sphere sizes of stitching immersive video on 3D modelling

The radius of the sphere [m]	Delta *X* [m]	Delta *Y* [m]	Delta *Z* [m]	Delta *XYZ* [m]
1	0.393	0.928	0.113	1.014
2	0.216	0.486	0.118	0.545
5	0.057	0.167	0.064	0.188
10	0.027	0.103	0.030	0.111
20	0.030	0.121	0.030	0.128
100	0.044	0.149	0.025	0.157
dynamic	0.025	0.111	0.025	0.116
automatic	0.032	0.111	0.027	0.119

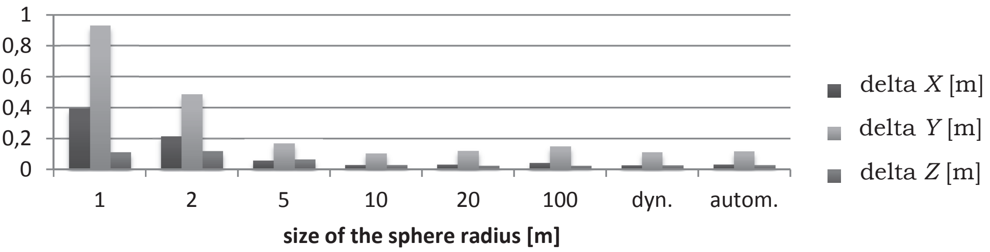

Fig. 5. The impact of the size of the sphere radius on 3D modelling

The adoption of 10 m radius is justified because the control points are located mainly on the walls of the assembly hall, whose size is 21 × 12 m, and pictures were taken from the middle of the hall. “Dynamic” selection of the radius is based on finding a common optimal radius for the whole immersive video, whereas “auto” radius searches for the best stitching radius for every panorama and in case of exporting hundreds of panoramas, it is very time consuming.

5.3. The Impact of the Format of the Panoramas

The third set of parameters that influences 3D modelling is the file format. Research by Guidi [8], indicates that HDR files enables more tie points to be found. Agisoft does not allow the importing of HDR files, as generated by LadybugCapPro, so they need to be converted to the EXR format in Adobe Photoshop. Table 3 presents a comparison of the accuracy on control points obtained on the network of 104 pictures in various formats (EXR, BMP and BMP tone mapped). Visually the best were the images in the BMP tone mapped format, BMP were too dark, and EXR too bright. It is worth noting that slightly more tie points (0.4%) were determined in the case of EXR/HDR files because they could be measured in the areas of images, however the time for aligning panoramas is better for BMP files. The best accuracy results were achieved for the EXR format, comparing both the average displacements on control points and average errors on images.

However, the time spent to align the network of panoramas argues in favour of choosing the BMP format. When the necessity of converting HDR to EXR is also taken into account, the BMP format is the best option.

Table 3. Comparison of file formats in aligning and cloud point generation

Picture format	EXR (HDR)	BMP (normal)	BMP (tone mapped)
Visual assessment	too bright	too dark	visually correct
File size [MB]	25	43	43
The number of determined points	37 505	37 357	37 178
The average deviation on control points [m]	0.007	0.007	0.012
The average error on images [px]	0.5	0.7	0.4
Time of aligning panoramas	7'33"	6'48"	6′23′′

5.4. The Impact of the Panorama Numbers

The next aspect that was tested is the optimal number of images necessary for proper 3D modelling. In the calculations, four versions of the network were used, the first one with all (71) panoramas on a selected section of the trajectory, the second version which uses every third panorama (24 panoramas in total), the third version which uses every sixth panorama (12 panoramas) and the fourth version which takes every ninth panorama (8 panoramas). Table 4 demonstrates that, as the number of panoramas grows, the average error on images is reduced while the average deviation on images remains at the same level, or even increases slightly. It is worth noting the number of points measured by the software. Within less than 1 hour, the program generated 116 056 points out of 71 panoramas, but when using only 8 panoramas, the matching time was approx. 1.5 minute, but only 17 083 points were generated (Fig. 6).

Table 4. The impact of the number of panoramas

Number of panoramas	The amount of determined points	The number of control points	The average deviation on control points [mm]				Single standard deviation [px]	Matching time
			delta *X*	delta *Y*	delta *Z*	point error		
71 (all)	116 056	29	12	7	8	16	0.4	57'3"
24 (every third)	52 564	29	11	8	7	16	0.5	08'24"
12 (every sixth)	26 855	29	10	8	7	14	0.5	03'5"
8 (every ninth)	17 083	29	10	7	6	13	0.6	01'32"

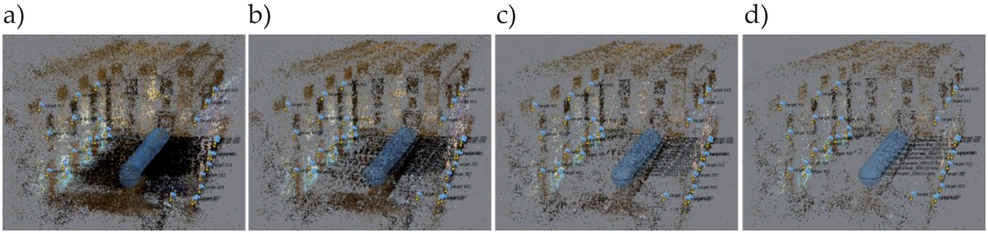

Fig. 6. The impact of the number of panoramas on generating points with 71 (a), 24 (b), 12 (c) and 8 (d) panoramas

6. Summary

In this study, the current photogrammetric methods that relies on spherical panoramas were transferred to immersive video. The main differences between the traditional spherical panoramas, which are successfully used for modelling, and immersive video are: relatively small resolution and a greater number of images. The advantage of video is the ability to visualize the object quickly by means of mobile camera movement.

Agisoft Photoscan was used to create a spatial model as it enables implementation of the automatic orientation and dense matching from panoramic images was used. The use of an automated program is necessary for video images, because the manual processing of such a large number of film frames would be too laborious.

The outcomes of this research suggest that:

1. The process of stitching images into panoramas supported by the immersive camera software leads to the creation of a panorama that is geometrically more correct than a panorama created with the dedicated software for traditional spherical panoramas, because the manufacturer has the possibility to use previous knowledge about interior orientation of each camera as well as the relative orientation between all cameras.
2. The choice of the sphere radius is very important and has a crucial impact on the accuracy. The radius value should be adopted on the basis of *a priori* estimated average distance between measured points, or rely on a selection option that is available in the camera software.
3. The study does not indicate any preference for the HDR format of image files as the necessity to convert them reduces the negligible accuracy benefits.
4. The number of panoramas used in modelling did not significantly affect the alignment accuracy on control points. However, an increase in the number of images is critical to the density of the point cloud of the model being created; therefore it causes an increase in the resolution of the created model.

The research conducted so far indicates that immersive video cameras are an imaging sensor with the potential for photogrammetric measurements and further studies need to be undertaken, especially on the optimisation of measurement accuracy and its economics.

References

[1] Barazzetti L., Fangi G., Remondino F., Scaioni M.: *Automation in Multi-image Spherical Photogrammetry for 3D Architectural Reconstructions*. VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage-Short and Project Papers, The Eurographics Association, 2010, pp. 75–81.

[2] Bednarczyk M., Pelc-Mieczkowska R.: *Propozycje wykorzystania zdjęć panoramicznych w GIS i geodezji.* Roczniki Geomatyki, t. 12, z. 64, 2014, pp. 163–176.

[3] Chang P., Hebert M.: *Omni-directional Structure from Motion.* Proceedings on Omnidirectional Vision, IEEE Workshop, 2000, pp. 127–133.

[4] Fangi G.: *The Multi-image Spherical Panoramas as a Tool for Architectural Survey*. XXI International CIPA Symposium, Athens, XXXVI-5/C53, 2007, pp. 311–316.

[5] Fangi G.: *Further Developments of the Spherical Photogrammetry for Cultural Heritage*. XXII CIPA Symposium, Kyoto, Japan, 2009, pp. 11–15.

[6] Fangi G., Nardinocchi C.: *Photogrammetric Processing of Spherical Panoramas.* The Photogrammetric Record, vol. 28, no.143, 2013, pp. 293–311.

[7] Griffiths A.: *Shivers Down Your Spine: Cinema, Museums, and the Immersive View*. Columbia University Press, New York 2008.

[8] Guidi G., Gonizzi S., Micoli L.: *Image Pre-processing for Optimizing Automated Photogrammetry Performances*. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. II-5, 2014, pp. 114–125.

[9] Guo J., Zhong R., Zeng F.: *Panoramic Images Mapping Tools Integrated within the ESRI ArcGIS Software*. IOP Conference Series: Earth and Environmental Science, vol. 17, no. 1, 2014, pp. 1–5.

[10] Kaess M., Dellaert F.: *Probabilistic Structure Matching for Visual SLAM with a Multi-camera Rig*. Computer Vision and Image Understanding, vol. 114, no. 2, 2010, pp. 286–296.

[11] Kwiatek K.: *360° Film Brings Bombed Church to Life*. ISPRS –International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII-5/W16, 2011, pp. 69–76.

[12] Kwiatek K., Tokarczyk R.: *Photogrammetric Applications of Immersive Video Cameras*. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 1, 2014, pp. 211–218.

[13] Kwiatek K., Woolner M.: *Embedding Interactive Storytelling within Still and Video Panoramas for Cultural Heritage Sites*. XV International Conference on Virtual Systems and Multimedia, VSMM 2009, IEEE, Vienna, 2009, pp. 197–202.

[14] Luhmann T., Tecklenburg W.: *3-D Object Reconstruction from Multiple-station Panorama Imagery*. Panoramic Photogrammetry Workshops, ISPRS Archives, vol. XXXIV-5/W16, Dresden 2004.

[15] Nakagawa M., Kataoka K., Yamamoto T., Shiozaki M., Ohhashi T.: *Panoramic rendering-based polygon extraction from indoor mobile LiDAR data*. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XL-4, 2014, pp. 181–186.

[16] Pagani A., Stricker D.: *Structure from Motion Using Full Spherical Panoramic Cameras*. IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2011, pp. 375–382.

[17] Parian J.A., Gruen A.: *A refined sensor model for panoramic* cameras. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 34, no. 2, 2004, pp. 1–12.

[18] Scheer E., Sewell S., *Scenario*. University of New South Wales Press, Sydney 2011.

[19] Schneider D., Maas H.-G.: *Geometric Modelling and Calibration of a High Resolution Panoramic Camera*. VI Conference on Optical 3D Measurement Techniques, Zurich, Switzerland, 2003, pp. 122–129.

[20] Schneider D., Maas H.-G.: *Application and Accuracy Potential of a Strict Geometric Model for Rotating Line Cameras*. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 34, 2004, pp. 1–5.

[21] Schneider D., Maas H.-G.: *Combined Bundle Adjustment of Panoramic and Central Perspective Images.* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36, no. 5, 2005, pp. 181–186.

[22] Shi Y., Ji S., Shi Z., Duan Y., Shibasaki R.: *GPS-supported Visual SLAM with a Rigorous Sensor Model for a Panoramic Camera in Outdoor Environments*. Sensors, vol. 13, no. 1, 2012, pp. 119–136.

[23] Szeliski R., Shum H.-Y.: *Creating Full View Panoramic Image Mosaics and Environment Maps*. Proceedings of 24th Annual Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York 1997, pp. 251–258.