

3D MODELLING WITH THE SAMSUNG GEAR 360

L. Barazzetti¹, M. Previtali¹, F. Roncoroni²

¹Dept. of Architecture, Built environment and Construction engineering (ABC)
Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan, Italy

²Polo Territoriale di Lecco, via Previati 1/c, Lecco
(luigi.barazzetti, mattia.previtali, fabio.roncoroni)@polimi.it
<http://www.gicarus.polimi.it>

Commission II

KEY WORDS: 3D modelling, Accuracy, Automation, Bundle adjustment, Calibration, Fisheye, Surface extraction

ABSTRACT:

The Samsung Gear 360 is a consumer grade spherical camera able to capture photos and videos. The aim of this work is to test the metric accuracy and the level of detail achievable with the Samsung Gear 360 coupled with digital modelling techniques based on photogrammetry/computer vision algorithms. Results demonstrate that the direct use of the projection generated inside the mobile phone or with Gear 360 Action Direction (the desktop software for post-processing) have a relatively low metric accuracy. As results were in contrast with the accuracy achieved by using the original fisheye images (front and rear facing images) in photogrammetric reconstructions, an alternative solution to generate the equirectangular projections was developed. A calibration aimed at understanding the intrinsic parameters of the two lenses camera, as well as their relative orientation, allowed one to generate new equirectangular projections from which a significant improvement of geometric accuracy has been achieved.

1. INTRODUCTION

Consumer-grade cameras able to capture 360 photos and videos are becoming more popular for the opportunity to look in any direction, exploiting immersive visualization with virtual reality headsets. Different cameras are available on the market. Some examples are the Ricoh Theta S, 360fly 4K, LG 360 CAM, Kodak PIXPRO SP360 4K, Insta360, Kodak PIXPRO SP360, and the Samsung Gear 360. More professional (and expensive) cameras are the GoPro Odyssey, Sphercam V2, Nokia OZO, and GOPRO OMNI.

The camera considered in this paper is the Samsung Gear 360 (Fig. 1), which has a dual 15MP CMOS sensor with integrated f/2.0 fisheye lenses, dual cam video resolution of 2840×1920 pix, and dual cam photo resolution of 7776×3888 pix. The camera requires a Bluetooth connection with a Samsung mobile phone (such as the Samsung S6 or S7) to get real-time visualization and control acquisition parameters. Although the camera can be used as a standalone tool, the connection with the mobile phone is surely more user friendly and allows the user to check the quality of the acquired images, especially with bad illumination conditions.

The images can be downloaded from the camera as circular fisheye images or equirectangular projections for 360° visualizations (Fig. 1). The creation of the equirectangular project requires to stitch the two fisheye images and can be carried out with the mobile phone app Samsung 360 or with the desktop software Gear 360 Action Direction. Some parameters can be modified during image acquisition, such as ISO, extra sharpening, exposure and white balance, and HDR image acquisition options.

The aim of this paper was to try out the Samsung Gear 360 for 3D modelling. The interest in this kind of acquisition tools is motivated by the wide field of view available, which make spherical cameras very attractive for interior scenes that usually requires a large number of pinhole images.

The mathematical model for equirectangular (say spherical) image orientation is illustrated and discussed in Fangi and Nardinocchi (2013). The mathematical formulation is based on the preliminary conversion of pixel coordinates into horizontal and vertical angles, simulating a theodolite. Then, an additional correction is carried out to take into account the lack of verticality for the Z axis. Bundle adjustment for spherical images becomes similar to the adjustment of geodetic networks, without distance measurements and an additional corrections for the lack of verticality.



Fig. 1. The Samsung Gear 360 and its two circular images (top) turned into a single equirectangular projection (bottom).

Other examples of complete 3D modelling projects based on this mathematical formulation were described in Barazzetti et al. (2010), D'Annibale and Fangi (2009), Fangi (2007; 2009), Fangi and Pierdicca (2012), Pisa et al. (2010). In these papers, equirectangular projection were generated from a set of images stitched with a software for panoramic photography (such as PTGui, Autopano, etc.). Indeed, such projections can be

generated by rotating cameras (e.g. a camera that rotates around its perspective centres) or dioptric and catadioptric mirrors. The spherical camera model is also available in some 3D modelling software from images. Examples are PhotoScan and Pix4D, which allow one to extract texturized meshes and orthophotos from sets of unoriented images. Results were described in Kwiatek and Tokarczyk (2014), Kwiatek and Tokarczyk (2015), and Pérez Ramos and Robleda Prieto (2015). The work presented in this paper is divided into three different parts: section 2 describes the methodology for image acquisition and panoramic image generation, with the Samsung apps and an ad-hoc calibration (distortion and relative orientation) of the two fisheye images. In section 3, the different images were checked with a set of 3D points measured with a total station. Section 4 shows a 3D modelling project where the surface of a 3D object was extracted in a fully automated way. Results were then compared with laser scanning point clouds, revealing a significant improvement of metric accuracy for the images generated with the ad-hoc procedure.

2. FISHEYE IMAGE CALIBRATION AND STITCHING

The Samsung Gear 360 acquires a pair of fisheye images that are then stitched inside (a) the mobile phone or with (b) the Gear 360 Action Direction desktop application (Fig. 2). The quality of the result is not the same: the desktop app produces smaller images (in terms of memory occupation) with a strong degradation of radiometric quality.

In addition, the fisheye images were stitched with software for panoramic photography (PTGui (c) and Autopano Giga (d)). This method required the calibration of the Samsung Gear 360, including distortion correction and relative orientation parameters between the fisheye images. Such alternative approach provided better results in terms of image quality and metric accuracy, as illustrated in the next sections.

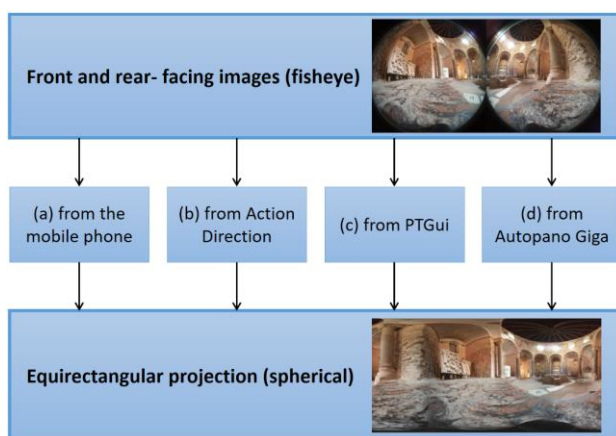


Fig. 2. The different options for the generation of equirectangular projections. (a) and (b) are default solutions, whereas (c) and (d) required a preliminary calibration project for the estimation of distortion coefficients and relative orientation between the fisheye image pair (more details are discussed in section 2.2).

2.1 Metric accuracy of single fisheye images

A set of 60 images of a wall was acquired with the Samsung Gear 360 placed on a tripod. Images have the typical configuration for camera calibration, i.e. several convergent images with roll variations. The aim was to run a markerless calibration procedure as described in Barazzetti et al. (2011) and

Stamatopoulos and Fraser (2014), in which calibration parameters are estimated from a block of target-less images. The used software is ContextCapture, which allows one to process fisheye images with a mathematical formulation based on the asymmetric camera model.

The estimated calibration parameters were then assumed as constant values for a 3D reconstruction project of a straight wall, on which a set of targets was installed and measured with a total station. The sequence was acquired only with the front-facing camera (Fig. 3). Images were oriented with ContextCapture, using 12 targets as ground control points and 7 targets as check points. The sequence is 42 m long and the camera object distance is 1.2 m. Statistics are shown in table 1 and reveal an error of about 5 mm, that confirms a good metric accuracy for the project carried out with front-facing camera.

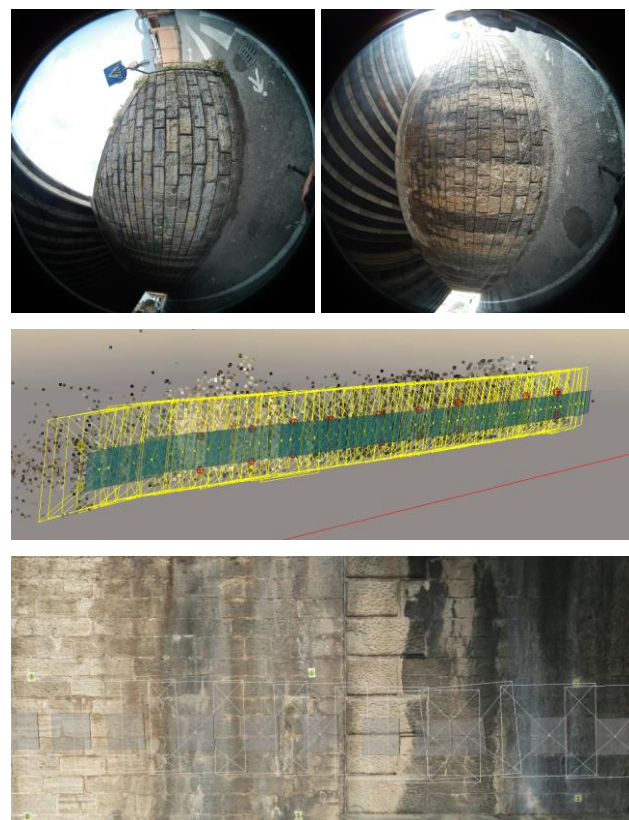


Fig. 3. Two images of the wall captured with the front facing camera (top), the sequence of images, control points and check points (middle), and a detail of the extracted mesh for a portion of the wall (bottom).

Number of GCPs	RMS of reprojection errors [pix]	RMS of horizontal errors [m]	RMS of vertical errors [m]
12	0.8	0.001	0.001
Number of Check Points	RMS of reprojection errors [pix]	RMS of horizontal errors [m]	RMS of vertical errors [m]
7	2.1	0.004	0.001

Table 1. Accuracy achieved with the front-facing camera.

Such results confirm the good metric quality of the Samsung Gear 360 when the original fisheye images are used for photogrammetric applications.

2.2 Relative orientation of front and rear-facing images

The calibration procedure aims at determining the relative position and attitude of front- and rear-facing images, as well as their distortion parameters. The Samsung Gear 360 was installed on a special support that allows the camera to rotate in a horizontal plane (Fig. 4). The system was installed on a tribrach previously levelled with a total station, so that the camera rotates around a vertical axis. A set of 12 front and rear facing images was acquired by rotating the Samsung Gear 360, obtaining 22 fisheye images. Images were then imported in PTGui and Autopano Giga to generate a single equirectangular projection. One may ask why several pairs of front and rear facing images were acquired using the basic technique of the rotating camera. Although the front and rear facing images have an overlap close to the border of the circular area, it was not possible to extract a good set of corresponding points from two images. The inclusion of other images allowed one to generate tie points and estimate the relative parameters between front and rear images with a better precision. These values are then assumed as a constant for a single pair of front and rear facing images.

The RMS of pixel coordinates achieved with PTGui was about ± 8.5 pixels, that is not an optimal result. The project was carried out also considering horizontal and vertical shifts, as well as horizontal and vertical shears. Distortion was also modelled with three coefficients. All these strategies were not able to reduce the final RMS.

Results with AutoPano Giga were instead better. The achieved RMS of image coordinates was ± 3.3 pixels. Fig. 4 shows the results with PTGui and the full set of 22 images. To generate the final projection, stitching between a single front and rear facing images was carried out. Indeed, only these images were included in the final panorama generation phase, removing the other “rotated” images but preserving the estimated relative orientation parameters.

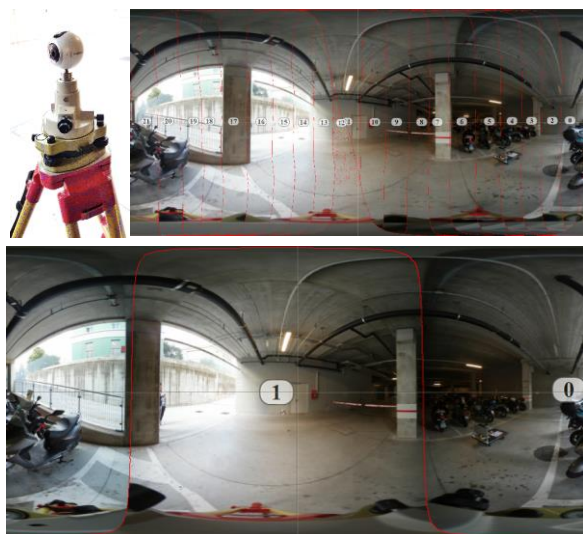


Fig. 4. The special calibration tool used to estimate the relative orientation between front and rear facing images. PTGui project with all images and the final result in which only a pair of front and rear facing images are used for generating the equirectangular projection.

3. EVALUATION OF METRIC ACCURACY WITH EQUIRECTANGULAR PROJECTIONS

The evaluation of the metric accuracy achievable with the equirectangular projections of the Samsung Gear 360 was carried out with a set of 3D points measured with a total station. 8 targets were placed around a station point and were measured with a total station (the expected accuracy is ± 2 mm). Their homogenous distribution (360° around the total station) allowed one to simulate the reconstruction project of a room. A set of 5 equirectangular projections was then acquired and processed. Targets are visible in at least 3 or 4 projections so that a good intersection of 3D rays is expected.

Images were acquired outdoor, so that a uniform illumination was easily achieved. Illumination conditions are an important issue for images acquired under an angle of 360° . Although the camera allows one to control some parameters for image acquisition, it is rather difficult to guarantee a uniform illumination in real projects where the full 360° scene is captured.

Equirectangular projections were generated with the methods illustrated in section 2: (a) from the mobile phone, (b) from the desktop app Action Direction, (c) from the parameters estimated with PTgui and (d) Autopano Giga.

The software used for photogrammetric processing is Agisoft PhotoScan, which integrates the spherical camera model based on the following equations given in a camera centred reference system:

$$\begin{aligned} x &= f \arctan \frac{X}{Z} \\ y &= f \arctan \frac{Y}{\sqrt{X^2 + Z^2}} \end{aligned} \quad (1)$$

where $f = \text{image_width} / (2\pi)$. The five images were oriented in a fully automated way and a 7-parameter transformation was used to rotate, scale and translate the project. 4 targets were used as control points and 4 as check points. Figure 5 shows one of the projections and the 8 targets, whereas the achieved metric results are shown in table 2.



Figure 5. One of the equirectangular projections and the 8 targets used as control and check points.

Results with the proposed calibration procedure “(c) and (d)” are better than those achieved with the images provided by Samsung software (a) and (b). Indeed, an overall improvement of a factor 3 was achieved with the ad-hoc calibration. Errors in the horizontal plane are worse than those along the vertical directions, for which the 4 methods gave similar results. The relative accuracy on check points (defined as RMS / scene_width) is about 1:200 for the better case (configuration “c”). The same value for control points is about 1:300 for the same configuration.

It is not clear why the accuracy along the vertical direction is better than a factor 2. One of the reasons could be the use of targets placed at the same height, instead of a uniform

distribution above and below the station point. Further tests are needed to clarify these aspects and will be taken into consideration in future work.

Control points (4)		Check points (4)	
RMS XY (m)	RMS Z (m)	RMS XY (m)	RMS Z (m)
(a) projections from the mobile phone			
0.061	0.011	0.088	0.009
(b) projections from Action Direction			
0.053	0.017	0.092	0.009
(c) projections from PTGui			
0.019	0.002	0.029	0.008
(d) projections from Autopano Giga			
0.011	0.001	0.037	0.009

Table 2. The statistics for control points and check points with the equirectangular projections generated with different solutions.

4. AUTOMATED 3D MODELING WITH THE SAMSUNG GEAR 360

4.1 Image acquisition with the Samsung Gear 360

Image acquisition with the Samsung Gear 360 can be carried out by placing the camera on a pole and taking the picture with the mobile phone (Fig. 6). This way to acquire images does not follow the classic photogrammetric approach, in which particular attention is paid to guarantee a good overlap between consecutive images. 360 cameras allow people with limited experience in photogrammetry to capture objects with a sufficient overlap. Illumination conditions are also extremely important because of the practical issues for uniform lighting conditions in a complete 360 scene.



Figure 6. Image acquisition can be carried out by placing the camera on a tripod; the camera can be pointed to any direction.

A dataset of 15 images was acquired inside an oratory with frescoes. The aim was to test the accuracy achievable from meshes extracted with the spherical camera model. The good texture of the object allowed one the use dense matching techniques for surface modelling. The analysed area has a vault that provides a reconstruction also for the area above the

camera, in which the equirectangular projection has significant deformations and the identification of tie points could be more difficult.

4.2 Orientation and surface extraction from equirectangular projections of the Samsung Gear 360

Image processing was carried out with Agisoft PhotoScan. Image orientation took less than 5 minutes, whereas dense matching for point cloud extraction took 40 minutes. Finally, a textured mesh was generated. The project was scaled with a known distance. The mesh was then compared with a laser scanning point cloud acquired with a Faro Focus 3D. The expected precision of the laser point cloud is 2-3 mm. The alignment between photogrammetric and laser scanning models was carried out with CloudCompare, which also provided a map of differences.

Figure 7 shows the results achieved with the equirectangular projection generated inside the mobile phone (procedure "a" in Fig. 2). The final mesh is very noisy and the overall error computed with CloudCompare is larger than 80 mm.

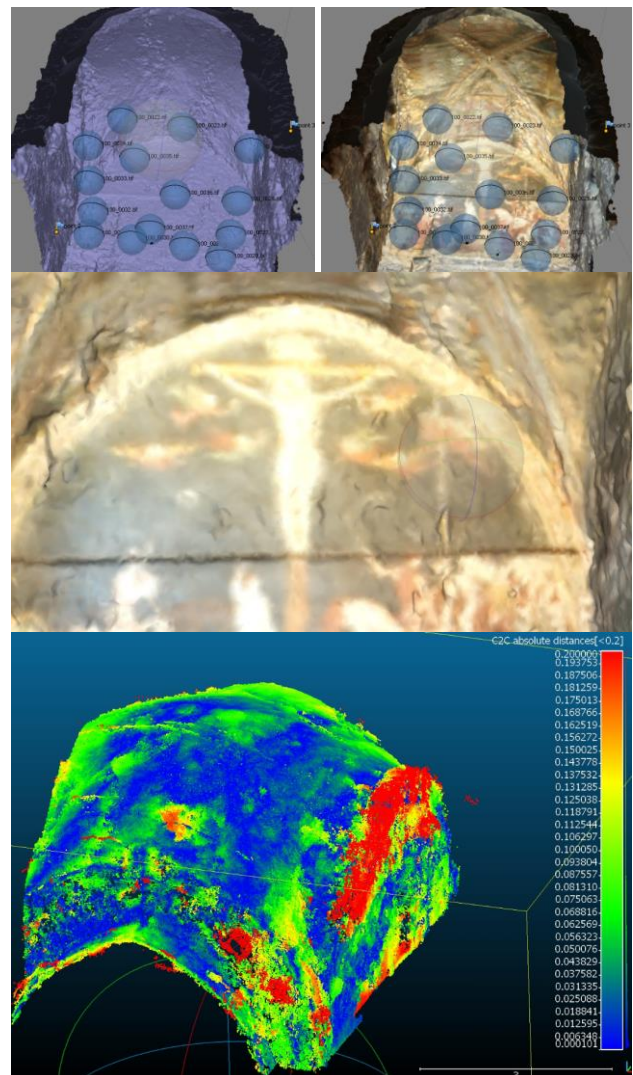


Figure 7. The discrepancy between laser scanning and image-based reconstruction with the equirectangular projections generated with the mobile phone app. The reconstruction is quite noisy and the overall error is about 80 mm.

Figure 8 shows the results achieved with the equirectangular projection created with the proposed workflow and PTGui (configuration (c) in Fig. 2). The mesh is smoother the previous result and the textured model has less discontinuities. The comparison revealed an accuracy of about 10 mm. Such accuracy is significantly better than the relative accuracy achieved in the previous section.

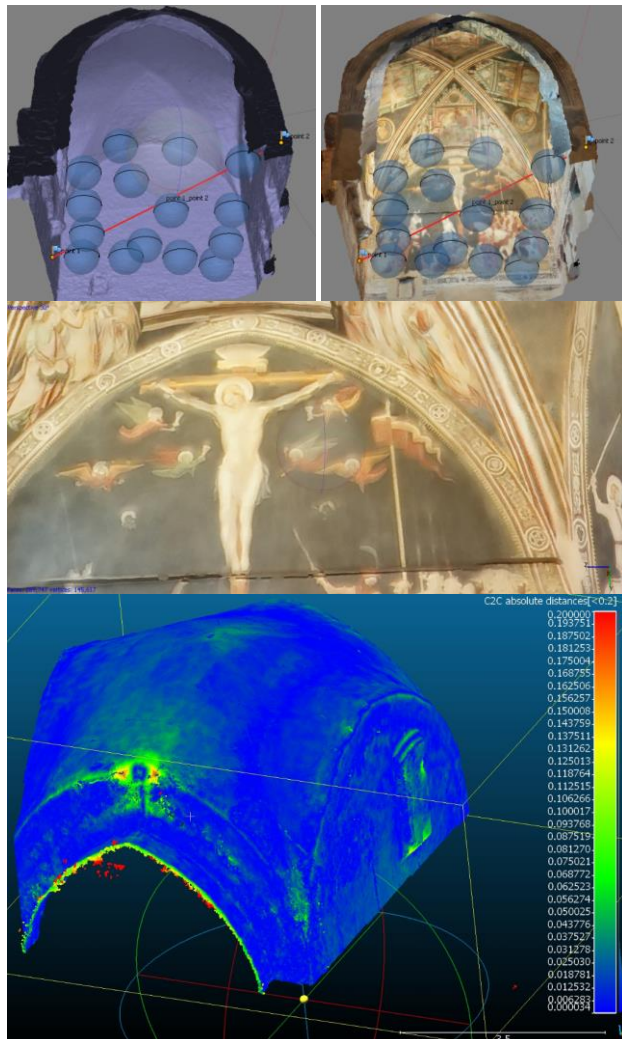


Figure 8. Results with the equirectangular projection generated with the workflow “c” (Fig. 2). The overall accuracy estimated with CloudCompare is 10 mm.

4.3 Single equirectangular projections or pairs of front- and rear-facing images?

As the results described in section 2.1 reveal a good metric accuracy when the single fisheye images are used, a reconstruction based on two fisheye images (front and rear) instead of a single equirectangular projection seems feasible. Figure 9 shows the image orientation results inside the oratory in the case of fisheye image processing (30 images). No constraint was used to fix the relative position of the images, which were instead processed as independent images. The used software is ContextCapture, in which camera the calibration parameters computed in section 2.1 were assumed as fixed for both front and rear facing images. The achieved mesh has a better quality than that generated by equirectangular image

processing. On the other hand, the reconstruction is partially incomplete, especially the area of the vault, which was instead modelled in the case of equirectangular projections.

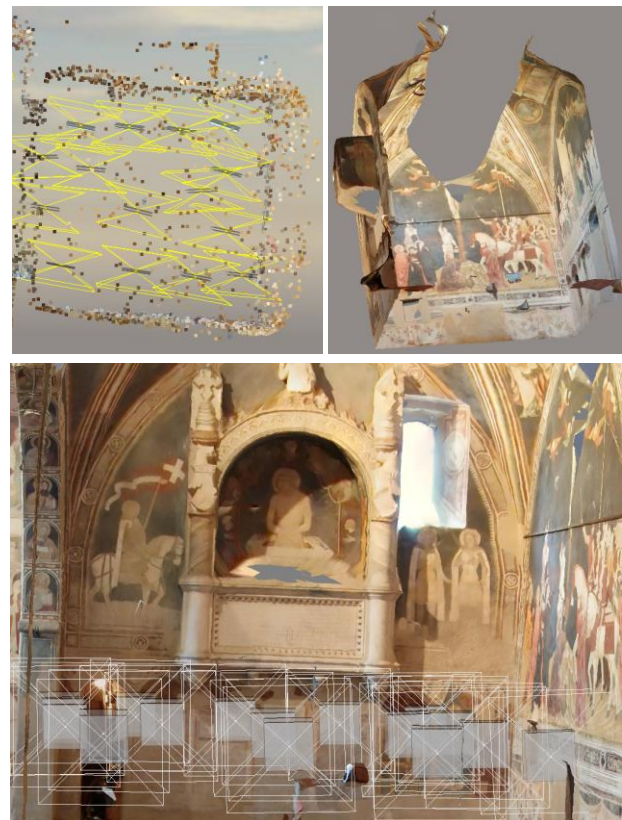


Figure 9. The reconstruction from front rear facing fisheye images has a better accuracy but is partially incomplete, especially the area above the camera.

5. CONSIDERATIONS AND CONCLUSION

The experiments described in this paper revealed that the Samsung Gear 360 can be used for metric reconstruction, although the achieved metric accuracy is not comparable with a traditional photogrammetric approach based on pinhole images. The achieved relative metric accuracy estimated with 3D points was about 1:300, that could be sufficient for applications aimed at determining the overall size or volume of a room. Better results were found for the 3D modelling project in section 4, in which the discrepancy between laser scanning and image-based models was only 10 mm. More experiments will be carried out in future work.

The proposed method is surely less expensive than a laser scanning and allows a rapid data acquisition of narrow rooms, in which a large number of pinhole images is needed.

On the other hand, there is limited control on camera parameters and images are acquired in an almost fully automated way. This makes the camera a photogrammetric tool also for people that have a limited experience in 3D modelling from images.

The variability of illumination conditions in rooms was also another problem. Uniform lighting conditions and nice colours are quite difficult to get when the field of the camera is 360°.

Another important consideration concerns the need of better solutions for stitching front and rear facing images in a unique equirectangular projection. The projections generated with the

mobile phone app or the desktop software Action Direction gave worse results than the proposed method, in which software for panoramic images were used after determining the relative position and attitude of the fisheye cameras.

The presented results highlighted that the Samsung Gear 360 has a limited use when a good metric accuracy is required. From this point of view, its usage is not recommended when the scale of the project required technical drawings at 1:10 – 1:20. The camera can be used for applications at coarser level (e.g. 1:100 – 1:200) could be carried out, especially when there is limited time for data acquisition.

An example of rapid processing is shown in Fig. 10, where the interior of the Basilica of San Pietro al Monte (Italy) was surveyed in only 10 minutes. Images were acquired in a very short time (less 5 minutes), obtaining a 3D model (data processing was about 3 hours) that could be suitable for applications with moderate metric accuracy.

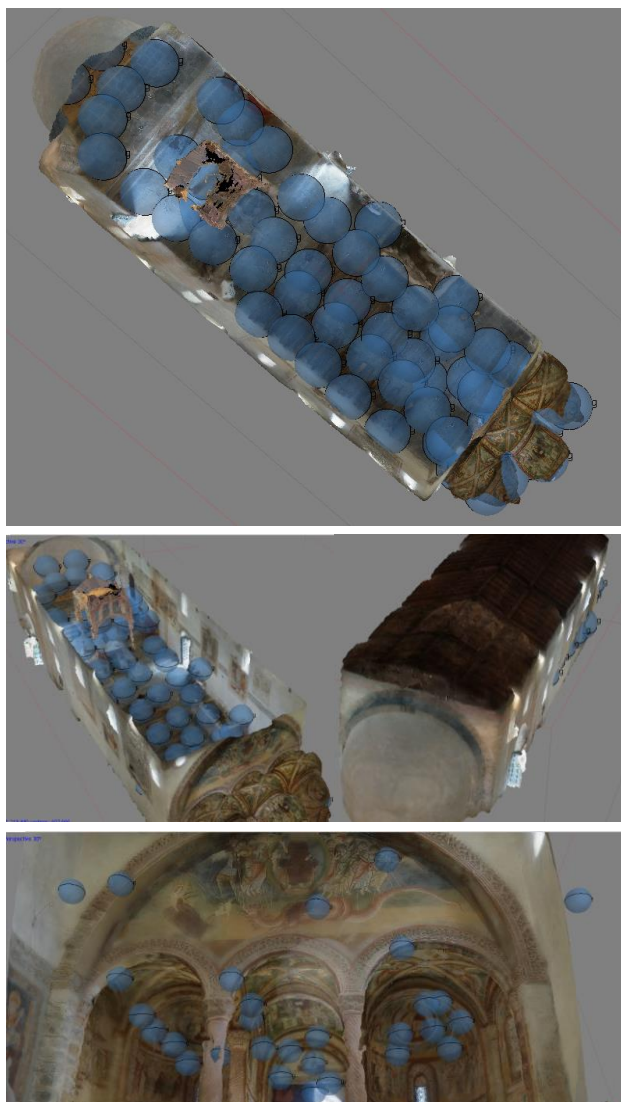


Figure 10. The interior of the Basilica of San Pietro al Monte (Italy) reconstructed with spherical images. Image acquisition required only 5 minutes, whereas data processing with PhotoScan took 3 hours.

REFERENCES

- Barazzetti, L., Fangi, G., Remondino, F. and Scaioni, M., 2010. Automation in multi image spherical photogrammetry for 3D architectural reconstructions. 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST), Paris, France. 75–81 (on CD-ROM).
- Barazzetti, L., Mussio, L., Remondino, F., Scaioni, M. (2011). Targetless Camera Calibration. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVIII-5/W16, pp.335-342.
- D'Annibale, E. and Fangi, G., 2009. Interactive modelling by projection of oriented spherical panorama. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(5/W1). 6 pages (on CD-ROM).
- Fangi, G., 2007. The multi-image spherical panoramas as a tool for architectural survey. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5/C53): 311–316.
- Fangi, G., 2009. Further developments of the spherical photogrammetry for cultural heritage. XXII International Committee for Cultural Heritage (CIPA), Kyoto, Japan. 6 pages (on CD-ROM).
- Fangi, G., 2010. Multi scale, multi resolution spherical photogrammetry with long focal lenses for architectural survey. ISPRS Midterm Symposium, Newcastle, UK. 228–233 (on CD-ROM).
- Fangi, G. and Pierdicca, R., 2012. Notre Dame du Haut by spherical photogrammetry integrated by point cloud generated by multi view software. International Journal of Heritage in the Digital Era, 1(3): 461–479.
- Kwiatek K., Tokarczyk R., 2014. Photogrammetric Applications of Immersive Video Cameras. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 1, pp. 211–218.
- Kwiatek, K., Tokarczyk, R., 2015. Immersive Photogrammetry in 3D Modelling. Geomatics and Environmental Engineering, Volume 9, Number 2, pp. 51-62.
- Pérez Ramos, A., Robleda Prieto, G., 2015. Only image based for the 3d metric survey of gothic structures by using frame cameras and panoramic cameras. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B5, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic, 363-370.
- Pisa, C., Zeppa, F. and Fangi, G., 2010. Spherical photogrammetry for cultural heritage. Proceeding of the Second Workshop on eHeritage and Digital Art Preservation, Florence, Italy. 3–6.
- Remondino, F., Fraser, C., 2006. Digital camera calibration methods: Considerations and comparisons. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.; 36: 266–272.
- Stamatopoulos, C., Fraser, C.S., 2014. Automated Target-Free Network Orientation and Camera Calibration. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2014;1: 339–346.
- Agisoft PhotoScan: <http://www.agisoft.com>
- ContextCapture: <https://www.acute3d.com/>
- CloudCompare: <http://www.danielgm.net/cc/>