

## FISHEYE MULTI-CAMERA SYSTEM CALIBRATION FOR SURVEYING NARROW AND COMPLEX ARCHITECTURES

Luca Perfetti <sup>1,\*</sup>, Carlo Polari <sup>1</sup>, Francesco Fassi <sup>1</sup>

<sup>1</sup> 3D Survey Group, Politecnico di Milano, Dept. of Architecture, Built environment and Construction engineering (ABC), Italy – (luca.perfetti, carlo.polari, francesco.fassi)@polimi.it

Commission II, WG II/8

**KEY WORDS:** Fisheye, Photogrammetry, Multi-camera, Calibration, Accuracy, Cultural Heritage Survey, Narrow Spaces

### ABSTRACT:

Narrow spaces and passages are not a rare encounter in cultural heritage, the shape and extension of those areas place a serious challenge on any techniques one may choose to survey their 3D geometry. Especially on techniques that make use of stationary instrumentation like terrestrial laser scanning. The ratio between space extension and cross section width of many corridors and staircases can easily lead to distortions/drift of the 3D reconstruction because of the problem of propagation of uncertainty. This paper investigates the use of fisheye photogrammetry to produce the 3D reconstruction of such spaces and presents some tests to contain the degree of freedom of the photogrammetric network, thereby containing the drift of long data set as well. The idea is that of employing a multi-camera system composed of several fisheye cameras and to implement distances and relative orientation constraints, as well as the pre-calibration of the internal parameters for each camera, within the bundle adjustment. For the beginning of this investigation, we used the NCTech iSTAR panoramic camera as a rigid multi-camera system. The case study of the Amedeo Spire of the Milan Cathedral, that encloses a spiral staircase, is the stage for all the tests. Comparisons have been made between the results obtained with the multi-camera configuration, the auto-stitched equirectangular images and a data set obtained with a monocular fisheye configuration using a full frame DSLR. Results show improved accuracy, down to millimetres, using a rigidly constrained multi-camera.

### 1. INTRODUCTION

This Paper tackles the problem of surveying narrow spaces in cultural heritage (CH) using image-based techniques.

Though it is a small portion of CH 3D mapping, narrow spaces surveying is a key obstacle to overcome to reconstruct a complete 3D model of many CH. Indeed, narrow and meandering spaces like corridors, passages, tunnels, stairwells etc... are far from being a rare encounter in restoration yards. because of their shapes and extension, tackle the problem with traditional instruments would be a very burdensome process. Although there are some examples described in the literature (Roncat 2011, Bonacini 2012, Rodríguez-González 2015), to survey a narrow tunnel with terrestrial laser scanners (TLSs) require a great number of scan stations that commonly results in a very time-consuming process.

Both range- and image-based solutions usually employed in CH are not fine-tuned to address the task of mapping narrow areas. The main obstacles being the quantity of data required, the scarce mobility, and the propagation of uncertainty. The narrow environment of many passages hampers the operator ability to perform the survey e.g. placing a topographic tripod and carrying heavy instruments. The shape of those areas is usually characterised by a high ratio between space extension and width, bringing to a configuration that suffers greatly from the problem of propagation of uncertainty. Moreover, the lack of natural illumination in indoor environments adds the problematic of carrying artificial illuminators alongside the camera sensor, and the lack of strong texture can make the employability of image-based techniques not feasible at all.

Potential solutions to the problem can come from both instruments categories, active and passive sensors: fisheye photogrammetry, on the image-based side, can reduce the quantity of data to be collected. On the other hand, the relatively new indoor range-based mobile mapping systems (MMSs) and

especially the handheld MMSs like the GeoSlam ZEB series maps the surrounding at the pace of a walk solving the time-effectiveness problem. The only question left before welcoming them into day-to-day practice is whether they can guaranty the accuracy level required by CH applications. Comparisons of different solutions have been presented by Mandelli et al. 2017. Although the indoor MMSs approach may be just as promising, this paper focuses on the image-based solution.

Time- and consequently cost-effectiveness are over all others aspects, the most discriminating parameters that decide if a particular technique will be used on restoration yards for survey activities. Photogrammetry successfully found its way in many different applications concerning cultural heritage preservation and valorisation, however, for indoor applications of complete 3D reconstructions, other techniques like terrestrial laser-scanning are far more predominant. The feasibility of employing photogrammetric techniques for the 3D mapping of complex closed architectonic spaces has always been limited by different factors, the main one being the number of photographs to be captured that those case studies require and therefore the number of tie points to be extracted (manually until a few years ago). Moreover, the narrower the environment is, the more impractical collecting all those data becomes.

The ultra-wide field of view (FOV) that characterise fisheye lenses, carries a potential solution to the problem. At least for some application, the use of fisheye photogrammetry has already been proven effective for shortening the acquisition phase compared to the time needed by photogrammetry using normal lenses and, more importantly, by competing techniques. Troisi et al. 2017 shown a successful use of fisheye video-photogrammetry to speed up the acquisition phase, Covas et al. (2015); Strecha et al. (2015); Fiorillo et al. (2016); Marčič et al. (2016); Barazzetti et al. (2017a); Perfetti et al. (2017) shown the potential of fisheye lenses to survey narrow spaces. Some years

\* Corresponding author

earlier instead, other authors presented applications of photogrammetry for narrow areas but using regular rectilinear wide-angle lenses (Roncella et al., 2012; Arles et al., 2013).

### 1.1 Multi-camera Constraints:

However, though fisheye photogrammetry does speed up the acquisition phase and though it can produce high-quality results, the time required by the processing phase and its reliability still prevent this technique to spread out into common practice. The speed advantage, that this tool might have over others, is almost obliterated by the careful supervision over the alignment results and the number of ground control points (GCPs) that are needed to achieve those results (Perfetti et al., 2017).

This paper investigates the possibility of minimising the work needed in the processing phase. The idea is to reduce to the minimum the number of external constraints required to achieve the desired accuracy (i.e. GCPs), using "internal" or "within the capturing geometry" constraints that could be calculated a priori just once. Specifically, the paper evaluates the advantage in using a multi-camera system against a monocular system to contain the degree of freedom of the photogrammetric network thereby containing the dependency from external constraints as well.

Since many photogrammetric networks require the same capturing geometry to be repeated several times, conserving roughly the same distances and angles among different stereopairs or image blocks, using a multi-camera device could potentially unburden the processing phase and, at the same time, further speed up the acquisition phase by making the whole process more controlled and less subject to human errors e.g. wrong acquisition angles, distances and overlap.

For this investigation, we decided to exploit a commercial panoramic camera: the iSTAR from NCTech. This camera is able to produce 360° degree images by stitching together the pictures acquired by four fisheye cameras. The intent wasn't that of using the auto-processed equirectangular images but instead that of using the raw fisheye photos coming from each of the cameras that constitute the system. It is indeed common for the automatic stitching of these "off the shelf" panoramic cameras not to be suited for metrology, a problem that has been covered by Barazzeti et al. (2017b) for the Samsung Gear 360 throughout a self-calibration of both the lenses distortion and the stitching parameters. This is certainly a valid approach to turn commercial 360° degree cameras, designed for photography, into metric imaging devices. Although in a different way, it provides additional "internal" constraints to a set of images (the stitching), ultimately reducing the degree of freedom of the photogrammetric network.

The research presented in this paper, on the other hand, only starts by investigating panoramic cameras but only seeing them as a rigid pre-calibrated multi-camera system. The idea is to use the original raw fisheye photos that the iSTAR record for each acquisition, and therefore the additional "internal" constraints are provided as rigid distances between the four optical centres, the relative orientation of the cameras and their internal parameters.

### 1.2 Manuscript Structure:

The first step was that of precisely estimate those parameters; this was done through a process of self-calibration of the multi-camera system at ones. The calibration procedure along with the results obtained is described in section 2. On section 3 we present the challenging case study of the Amadeo Spire of the Milan Cathedral, the 3D survey of the spire was used as a real-world

test to evaluate different approaches. In Section 4 we describe the different tests that were performed using the equirectangular auto-stretched images, the single fisheye pictures with and without additional internal constraints as well as the same survey carried out with a Nikon D810 DSLR and the full frame fisheye lens Samyang 12mm. In section 5 we discuss and draw some conclusion over the results and section 6 highlights possible future direction on this research.

## 2. CALIBRATION OF THE NCTECH ISTAR

### 2.1 The iSTAR Panoramic Camera:

The target scenario of this research is the 3D survey of architectonic narrow spaces. The capturing geometry best suited for this specific application is arguably difficult to be defined, however, we can safely state that a stereoscopic configuration would function better than a panoramic configuration will do. Nevertheless, the panoramic configuration is the one we decided to use as a starting point of this investigation. The reasons are: the convenience of using the same capturing geometry for all the application, even if it is not optimized for any of them; the relative affordability of today's panoramic cameras based on low-cost sensors and fisheye lenses; and the fact that those cameras come as a rigid system in which singles sensors position never changes relative to others.

For the aforementioned reasons, we conduct our tests using the iSTAR panoramic camera from NCTech, a 360° degree imaging system that hosts 4 fisheye lenses in a square configuration slightly tilted up with respect to the horizontal plane where the lenses lie. NCTech offers a software that automatically processes the raw data captured by the camera and produces an equirectangular image as output. The iSTAR can also automatically bracketing different shots (camera must always be steadily mounted on a tripod) at different exposure levels in order to process high dynamic range (HDR) images. Also, the automatic panoramic stitching offers the possibility to process HDR data if they were acquired in the first place. The ability to capture HDR images is certainly a convenience whether the light conditions are not optimal likewise those of dark narrow areas are, nevertheless, this functionality of the camera was exploited only for the processing of the equirectangular images and never for the single fisheye shots.



Figure 1. NCTech iSTAR during the surveying operations.

## 2.2 The Calibration Polygon:

The idea of using the original raw fisheye data coming from the single cameras that compose the iSTAR, as the aim of find a general approach suitable for each hand-carriable multi-camera rig that allows to exploit a wide ground coverage (FOV) by combing multiple views and, at the same time, to contain the degree of freedom of the photogrammetric network. The aim of the calibration is, therefore, that of deriving both the internal distortion coefficient that characterises each camera and the relative spatial relation between them, namely: the rigid distances between the four optical centres and the relative orientations.

To accommodate for the large FOV of the iSTAR, we decided to set up a tri-dimensional polygon consisting of a great number of photogrammetric circular coded targets displaced on three vertical walls as well as on the floor and ceiling of a room (Figure 2). Secondly, a redundant photogrammetric network was designed to perform the survey of the environment/polygon making sure to rotate the iSTAR at different angles with respect to the vertical axis as well as around it. For each position of the iSTAR (Figure 3) four shots have been acquired (each composed of four raw fisheye photos) rotating the camera by 90° degrees horizontally. This was done in order to acquire enough data to first calibrate the single cameras of the systems one by one and only secondly to calibrate the relative relationship between them. all the targets centres were measured with the aid of a total station to be used as GCPs lately in the process. The calibration of the camera was obtained twice following two different pipelines using two software: Agisoft Photoscan and Micmac.

**2.2.1 Calibration with Photoscan:** The first calibration using Photoscan were obtained as follows: As mentioned above, four projects were made in order to calibrate each of the cameras of the multi-camera panoramic system. Markers were used as GCPs and the tie points computed by the software were heavily filtered. After the optimization process, the derived camera calibration has been exported.

A comprehensive project has then been made using pictures from all the singles cameras. For this project, the distortion parameters of each camera have been fixed on the values coming from the individual projects. At this step, for each position of the iSTAR only one shot as been considered (each composed by four pictures) to avoid possible errors on the reconstruction that might have occurred from multiple images with the same position.

As Figure 3 shows, all the pictures pointing the sixth side of the room, the only one without coded markers, were discarded. The image to be removed came from a different camera at each position of iSTAR. At this point, the development of a python script to be used inside Agisoft Photoscan was necessary to extract the distances between the four optical centres for each camera. The distances were averaged together and stored to be used in future projects as constraints. A second script allows the user to implement such constraints as the first step: one can choose the type of the used camera and load a calibration for the distortion, the number of cameras in a general multi-camera system and the number of pre-calibrated connections among them. The script implements the rigid distances as scalebars in the project to be considered in the bundle adjustment.

[mm]	O1-O2	O1-O3	O1-O4	O2-O3	O2-O4	O3-O4
mean	59.41	83.85	59.33	59.33	83.84	59.45
StD	0.22	0.18	0.22	0.21	0.21	0.22

Table 1. Average distances between cameras optical centres and relative standard deviation.

The strength of constraints can also be implemented giving the script the standard deviation of the original distances used to obtain the average (Table 1). The calibration obtained with Photoscan can be used also to derive the relative orientations of the cameras within the multi-camera system, however, to authors' knowledge there is no way to implement such constraints in the bundle adjustment using Photoscan yet. In order to do so the use of the software Micmac has been necessary.



Figure 2. Photogrammetric calibration room.

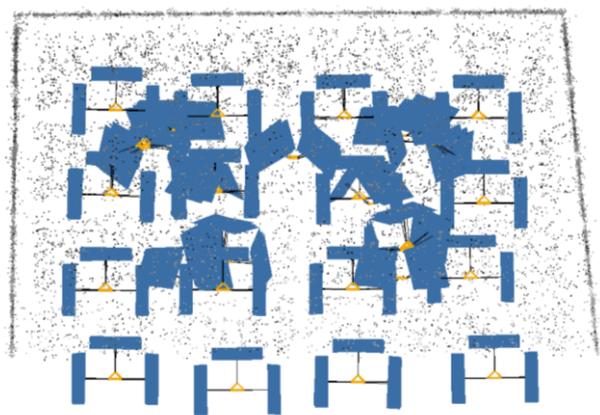


Figure 3. Plan view of the photogrammetric network used for the calibration.

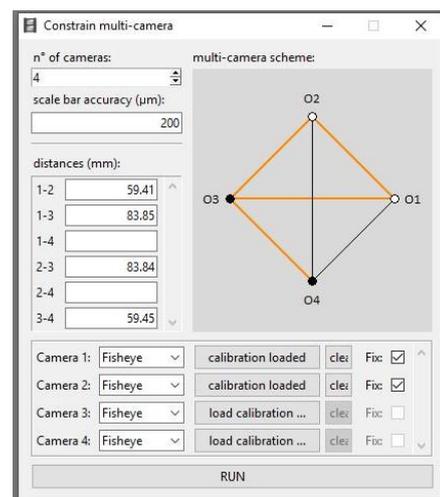


Figure 4. interface of the Photoscan script to constrain camera distances between rigid multi-cameras.

**2.2.2 Calibration with Micmac:** The software Micmac has been used both to test different calibration model specifically designed for fisheye lenses and because it offers a tool to calculate and implement relative translation and rotation between cameras inside a rigid rig (the block structure), namely: the "Blinis" tool.

The steps followed with Micmac were different from those followed with Photoscan to better accommodate the features these software offers. Micmac does not consider GCP observations as tie points and therefore is not able to align the set of photos on markers only as Photoscan does. Tie points are mandatory to compute the alignment and since the texture quality of the used calibration polygon/room was quite poor, it follows that the tie points computed from the image set were not robust enough to compute the calibration of the fisheye cameras from scratch. To overcome this problem another image set has been acquired using only one of the cameras of the system and framing a highly textured wall with convergent images. The calibration obtained from this set was then used as the initial value to compute the four calibrations of the four cameras in the polygon dataset. Tie points used in this set were exported from Photoscan using a script able to write them following the Micmac format. After the cameras were successfully aligned, we run an optimization and computed the block structure using the Blinis tool.

### 3. THE AMADEO SPIRE CASE STUDY

#### 3.1 A Spiral Staircase Inside the Spire:

A test of the proposed solution has been carried out in a challenging case study; the Amadeo Spire of the Milan Cathedral (Figure 5). The Amadeo Spire is a highly decorated architectonic element that encloses a very narrow spiral staircase that measures only 70-80cm in width and that is about 8m tall. The core of the spire is a pillar, octagonal in shape, made of marble blocks around which the steps revolves. A "shell" made of many highly detailed small pillars enclose the spiral stair composing a "filter" with the outside. The Amadeo Spire is located at the North-East corner of the cathedral's lantern while similar spires are located at the others. It connects the higher level of the roofs with the area of the dome's "sordine". A series of relatively small spaces that cover the dome of the cathedral and that are home to four bells from the 16<sup>th</sup> century.

This case study was chosen firstly because of the spatial characteristic of the staircase: the revolving stair makes it impossible to imaging the top end of the staircase from a viewpoint at the bottom, actually from each position along the stair the view is very limited and thereby the propagation of uncertainty become a serious problem. Secondly, for the illuminance conditions, the location on the outside allows to temporally avoid the problem of carrying artificial illuminators to light up the environment that would be instead of paramount importance for "true" indoor applications. The drawback is that the external "filter", the "skin" of the spire, is characterised by many openings toward the outside that take away a lot of the surface area for tie points detection.

The inside environment of the spiral staircase has been surveyed using the iSTAR panoramic camera mounted on a tripod (Figure 1), a data set of 55 multi-images were acquired, one for each of the steps and a few more at the top, for a total number of 220 individual fisheye photos. The pictures cover the area of the stair from the bottom to the top, a total length of 8m.

At the time of the surveying operations, the Veneranda Fabbrica of Milan's Cathedral was preparing for the restoration activity of the Amadeo Spire, the subjects of the restoration being the marble blocks and their conditions. The 3D Survey Group was

asked to produce high-resolution orthoimages of the elevation of each of the sides of the spire: eight for the core pillar and eight for the exterior. While the 3D models of the exterior elevations were obtained using a 12mm rectilinear lens mounted on the Nikon D810, the inside environment as surveyed using the same camera body but coupled with a 12mm stereographic fisheye lens. The comparisons between the full frame DSLR and the one obtained using the iSTAR is discussed in section 5.



Figure 5. Orthoimages of the Amadeo Spire: the complete (left) and a zoom of the top portion (centre) of the outside elevation. On the right, the inside elevation obtained with monocular photogrammetric network performed with Nikon D810 and 12mm fisheye lens.

### 4. DIFFERENT APPROACHES TO COMPARISON

#### 4.1 The method:

The aim of the tests is to evaluate whether by adding multi-camera constraints, therefore reducing the degree of freedom of the photogrammetric network, results in improved accuracy in comparison to what can be obtained with the monocular approach.

The methodology used to conduct the comparisons is that of processing the same data set, 202 images, acquired with the NCTech iSTAR, differently several times. Starting from the plain images aligned without constraints as if they were acquired by a

monocular system and adding test after test more and more constraints: the self-calibration on internal parameters, the fixed distances between the optical centres, and the constraints of relative orientations.

Once the different alignments were obtained, they were georeferenced in the same coordinates system throughout the aid of photogrammetric targets used as GCPs of known coordinates. The resulting RMSs on control points were then computed. To better get an understanding of the drift problem and to evaluate whether to use such constraints helps reducing it, we run three optimizations for each alignment: for the first one, few GCPs were conserved only at the top of the staircase and CPs, therefore, shows the error along the stair and at the bottom; for the second one we repeated the same process but adding only one GCP at the very bottom of the stair in order to check the improvements; for optimization number three, instead, GCPs were placed all over the extension of the staircase alongside the CPs.

Since the aim is to reduce to the minimum the number of GCPs that are required to achieve accurate high-quality results, the goal has been that of get low errors – within the 1cm threshold that corresponds to error of 1:50 scale – on CPs, when GCPs are picked only at one of the extremities of the narrow staircase.

## 4.2 Raw Images Multi-camera:

**4.2.1 No Constraints:** The first test sees the raw fisheye images from iSTAR aligned without any kind of multi-camera constraints. In this case, the pictures were processed with Agisoft Photoscan (Table 2, C). This test represents the worst possible scenario. A more sensible test was conducted adding only the information on the pre-calibration of the lenses internal coefficients; The test was conducted with Photoscan first (Table 2, D), and then repeated in Micmac (Table 2, F): in this case initial value of camera locations have been considered as well by exporting the coordinates from Photoscan (from elaboration E in table 2) in Micmac format using a script.

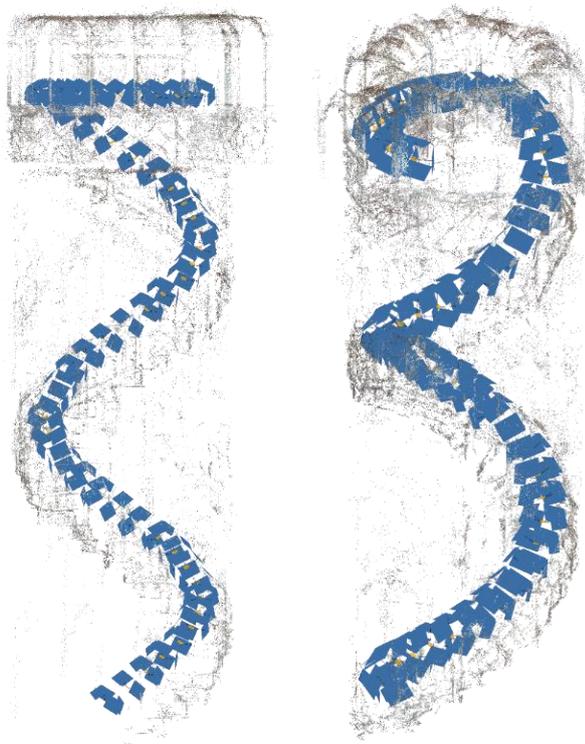


Figure 6. Sparse point cloud of the test with distances constraints (Table 2, E) computed in Photoscan.

**4.2.2 Constraints on Distances:** The second Step sees the constraints of the distances between the optical centres of cameras as well as the fixed self-calibration of the lenses distortion as before (Table 2, E). The distances were implemented in Photoscan using the script described in section 2 (Table 1, Figure 4) before running the orientation process. For the iSTAR, the constrained distances were six, the four sides of the square configuration and the two diagonals. The standard deviation of the data averaged to derive these measurements was around 0.2mm and this same value became the accuracy on the scalebars. Figure 6 shows the alignment results achieved by this test.

**4.2.3 Constraints on Relative Orientation:** The third test sees the constraining of relative orientation of cameras as well, it has been performed using the software Micmac and the tool "Blinis" – to compute the average values – and the tool "Campari" to optimize the alignment considering those measures (Table 2, G). The "Blinis" tool was run on the calibration room data set, and the idea was to use the results on the Amadeo data set by aligning first only the photos coming from one of the lenses and then consequently being able to locate all the other thanks to the calibration of relative orientation. However, this turned out not to be allowed by Micmac since the tool "Campari", that perform the optimization, require all the camera to be pre-aligned. Since the iSTAR has a panoramic configuration and very short base distances (relative to the environment) between the pictures, to pre-align the images together without any given constraints turned out to be too difficult for Micmac. We therefore choose, to use as initial location of the cameras the alignment results obtained in Photoscan with the previous test.

## 4.3 Equirectangular Images:

Secondly to the tests on the multi-camera configuration, we decided to check the results achievable using the auto-stitched equirectangular panoramic images (Table 2, B). Indeed, panoramic images reduce the degree of freedom of the network by themselves, even without the need of implementing additional constraints. The stitching of the panoramas entails already that rigid relationship between cameras is known and that the lenses distortions are to be corrected. However, we decided not to cover the calibration of the equirectangular stitching in this paper, as Barazzetti et al. (2017b) did, the reasons being the will of keeping the procedure suitable for all kind of multi-camera system. We therefore tested the potential of the pre-calibrated stitching of the iSTAR, already knowing its limits (misalignments of images composing the panorama can be clearly seen).

## 4.4 Monocular DSLR:

The last test is the survey performed with the full frame DSLR Nikon D810 and a 12mm stereographic fisheye from Samyang (Table 2, A). The survey carried out with this camera is quite different from the others: while with the iSTAR the aim was to obtain the best results possible in terms of accuracy of the alignment, not bothering about the completeness of the data, the DSLR survey was focused of obtaining complete and high-resolution orthoimages (Figure 5). As a consequence, the number of pictures acquired was much more (about 1600). Operations were carried out the same way as presented in Perfetti et al. (2017), all the pictures were masked along the circumference of maximum ground sampling distance (GSD). This test is interesting to weight the importance of image resolution against the degree of freedom of the network, aside from the density of data, the DSLR acquisitions offer much more resolution as well as a much higher magnitude of liberty.

## 5. RESULTS

In Table 2 we report the RMSEs of both the GCPs and the CPs in the three configurations described in 4.1 for each of the tests introduced in section 4. Also, the worst errors on the observations are reported.

Starting from the data set acquired with the Nikon D810 we can see how the errors are not satisfying (greater than the 1cm threshold), especially for elaboration A.1 where GCPs were picked only at the top of the staircase, the error on CPs exceed the tolerance of 1:50 scale (2-3 cm). It becomes just acceptable for elaborations A.2 and especially A.3 where GCPs were picked along the entire extension of the stair, confirming what previously found in Perfetti et al. (2017): monocular fisheye photogrammetry can produce satisfying accuracy, but it requires a great number of GCPs observations.

Moving on to the auto-stitched panoramas test, we can see errors getting way greater than our threshold for any trials. Alignment drift still gets reduced by adding more constraints but average RMSE on CPs never get down 10cm. It is worth noticing though that all the equirectangular images were aligned without problems.

	GCPs		CPs	
	Average	Worst	Average	Worst
<b>A</b>	Nikon D810 - monocular no constraints			
A.1	0,0077	0,0091	0,0332	0,0795
A.2	0,0083	0,0104	0,0183	0,0359
A.3	0,0153	0,0242	0,0136	0,0214
<b>B</b>	iSTAR - auto-stitched equirectangular images			
B.1	0,0089	0,0123	0,1636	0,3818
B.2	0,0364	0,0584	0,1126	0,1674
B.3	0,0954	0,1796	0,1015	0,1775
<b>C</b>	iSTAR - multi-camera no constraints			
C.1	0,0178	0,0280	0,4738	1,4868
C.2	0,0609	0,0875	0,1306	0,3435
C.3	0,1553	0,3784	0,0866	0,1483
<b>D</b>	iSTAR - multi-camera pre-calibration			
D.1	0,0071	0,0095	0,1036	0,3008
D.2	0,0139	0,0213	0,0329	0,0737
D.3	0,0154	0,0271	0,0204	0,0431
<b>E</b>	iSTAR - multi-camera pre-calibration & distances			
E.1	0,0268	0,0463	0,0272	0,0384
E.2	0,0250	0,0470	0,0293	0,0449
E.3	0,0205	0,0174	0,0251	0,0335
<b>F</b>	iSTAR - multi-camera pre-calibration with Micmac			
F.1	0,0071	0,0084	0,0116	0,0193
F.2	0,0070	0,0082	<b>0,0104</b>	0,0169
F.3	0,0068	0,0108	<b>0,0071</b>	0,0116
<b>G</b>	iSTAR - pre-calibration & relative orientations			
G.1	0,0054	0,0063	<b>0,0107</b>	0,0201
G.2	0,0053	0,0077	<b>0,0092</b>	0,0172
G.3	0,0057	0,0103	<b>0,0057</b>	0,0117

Table 2. The RMSEs of GCPs and Cps are listed in the table: letters (A-G) refers to the test categories, numbers (1-3) refers to the three configurations of GCPs and CPs used for each test as described in 4.1. all measures are reported in metres.

Tests C and D and F show the behaviour of the iSTAR data set before any multi-camera constraints were implemented: in test C no pre-calibration of the cameras were considered as well and errors are significantly high, probably due to poor shooting condition for camera calibration to be correctly determined; errors get down for test D and are now comparable with the ones obtained with from the DSLR data set, we can see the strong influence of the propagation of uncertainty with average CPs error going from 10cm (D.1) to 2-3cm (D.2 and D.3); Test F shows instead very differ errors, much lower from the beginning. This can result from the camera calibration being better defined by the Fisheye model of Micmac (FishEyeEqui model in "Tapas" tool) and/or by the initial orientation of the cameras coming from test E, where rigid distances were constrained.

Test E is the first one for which some multi-camera constraints were implemented, at it is the first one of which results were not suffering too much the problem of propagation of uncertainty, as already mentioned, the test F was based on the orientation resulting from E. E.1 shows RMSEs just within the limit of 1:50 scale tolerance but the jump from test D.1 is especially relevant since their share the same distortion calibration. Results improve by considering GCPs speeded all along the stair but not that much showing that those errors may be due to the accuracy on the calibration of lens distortion resulting from Photoscan.

Finally, test G shows by far the best values, here the orientation coming from test E has been used in Micmac just as for test F, the calibration of the cameras are the same of test F obtained in Micmac using the "FishEyeEqui" model and following the methodology described in 2.2.2. Moreover, constraints on relative orientations has been added in the bundle adjustment and results are remarkable: average CPs errors are always within the threshold of 1cm and errors on GCPs are just around 5mm.

## 6. CONCLUSIONS AND FUTURE WORKS

The results of the test discussed in this paper shows clearly the advantage of a constrained multi-camera system over a monocular one, tests E, F and G against test B. However, this is not surprising, a more accurate result from a more well-thought processing was to be expected. What is worth noticing is the possibility to contain the alignment drift within very low value even in such extreme scenario like the survey of a spiral staircase is. Errors are even more promising given the fact that the iSTAR panoramic camera host very low-resolution sensors.

Limits in the process of constraining the relative position of cameras and distortions can be appointed to the fact that a pre-alignment of cameras is mandatory for the "Campari" tool to implement the block structure constraints in the bundle. Here we overcame the problem starting from the alignment obtained with Photoscan that is also way faster than Micmac but for future works it may be thought otherwise considering also that Micmac's fisheye calibration models appear to be more effective. Future works will consider testing similar procedures on even more challenging case studies like the one of underground tunnels. The integration of artificial illuminators will be tested and, related to that and to the fact that panoramic configuration of the multi-camera system may not be optimal, a system based on stereoscopy will be used instead.

## ACKNOWLEDGEMENTS

Veneranda Fabbrica of the Milan's Cathedral supported this work. Special thanks to the director Eng. Francesco Canali and to all staff especially Francesco Aquilano and Massimiliano Regis.

## REFERENCES

- Agisoft, 2018. [www.agisoft.com](http://www.agisoft.com)
- NCTech, 2018. [www.nctechimaging.com](http://www.nctechimaging.com)
- Arles, A., Clerc, P., Sarah, G., Téreygeol, F., Bonnamour, G., Heckes, J. and Klein, A., 2013, 3d Reconstruction And Modeling Of Subterranean Landscapes In Collaborative Mining Archeology Projects: Techniques Applications And Experiences, In: International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, Vol. XL-5/W2, pp. 61–66, doi:10.5194/isprsarchives-XL-5-W2-61-2013.
- Barazzetti, L., Previtali, M., Roncoroni, F., 2017a. Fisheye lenses for 3d modeling: evaluations and considerations. In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Nafplio, Greece, Vol. XLII-2/W3, pp. 79-84, doi:10.5194/isprs-archives-XLII-2-W3-79-2017.
- Barazzetti, L., Previtali, M., Roncoroni, F., 2017b. 3D Modelling with the Samsung Gear 360. In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Nafplio, Greece, Vol. XLII-2/W3, pp. 85-90, doi: :10.5194/isprs-archives-XLII-2-W3-85-2017.
- Bonaccini, E., D'Agostino, G., Galizia, M., Santagati, C. and Sgarlata, M., 2012. The Catacombs of San Giovanni in Syracuse: Surveying Digital Enhancement and Revitalization of an Archaeological Landmark. In: Cultural Heritage Preservation, Springer Berlin Heidelberg, pp. 396–403. doi:10.1007/978-3-642-34234-9\_40.
- Covas, J., Ferreira, V. and Mateus, L., 2015. 3D reconstruction with fisheye images strategies to survey complex heritage buildings. In: 2015 Digital Heritage, Granada, pp. 123- 126, doi: 10.1109/DigitalHeritage.2015.7413850.
- Fiorillo, F., Limongiello, M. and Fernández-Palacios, B., 2016, Testing GoPro for 3D Model Reconstruction in Narrow Spaces, In: ACTA IMEKO, Vol. 5 no. 2, pp. 64-70, doi:10.21014/acta\_imeko.v5i2.372.
- Kannala, J, Brandt, SS., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. In: IEEE transactions pattern analysis and machine intelligence, Aug; 28(8):1335-40.
- Mandelli, A., Fassi, F., Perfetti, L. and Polari, C., 2017. Testing Different Survey Techniques To Model Architectonic Narrow Spaces. In: 26th International CIPA Symposium 2017, Ottawa, Canada, Vol. XLII-2/W5, pp. 505–11, doi:10.5194/isprs-archives-xlii-2-w5-505-2017.
- Marčič, M., Barták, P., Valáška, D., Fraštia, M., and Trhan, O., 2016. Use of image based modelling for documentation of intricately shaped objects. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B5, 327-334, doi:10.5194/isprs-archives-XLI-B5-327-2016.
- Perfetti, L., Polari, C., Fassi, F., 2017. Fisheye photogrammetry: tests and methodologies for the survey of narrow spaces. In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Nafplio, Greece, Vol. XLII-2/W3, pp. 573-580, doi:10.5194/isprs-archives-XLII-2-W3-573-2017.
- Rodríguez-Gonzálvez, P., Nocerino E., Menna F., Minto S. and Remondino F., 2015. 3D Surveying & Modeling Of Underground Passages In WWI Fortifications. In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. Vol. XL-5/W4, pp. 17–24. doi:10.5194/isprsarchives-XL-5-W4-17-2015.
- Roncat, A., Dublyansky, Y., Spötl, C. and Dorninger, P., 2011. Full-3D Surveying of Caves: A Case Study of Märchenhöhle (Austria). In: IAMG 2011 - Annual Conference of the International Association for Mathematical Geosciences, Salzburg. doi:10.5242/iamg.2011.0074.
- Roncella, R., Umili, G. and Forlani, G., 2012, A Novel Image Acquisition And Processing Procedure For Fast Tunnel Dsm Production. In: International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, Vol. XXXIX-B5, pp. 297–302, doi:10.5194/isprsarchives-XXXIX-B5-297-2012.
- Strecha, C., Zoller, R., Rutishauser, S., Brot, B., Schneider-Zapp, K., Chovancova, V., Krull, M. and Glassey, L., 2015, Quality Assessment Of 3d Reconstruction Using Fisheye And Perspective Sensors. ISPRS Annals of Photogrammetry Remote Sensing and Spatial Information Science, Vol. II-3/W4, pp. 215–22, doi:10.5194/isprsannals-II-3-W4-215-2015.
- Troisi, S., Baiocchi, V., Del Pizzo, S. and Giannone, F., 2017. A Prompt Methodology To Georeference Complex Hypogea Environments. In: Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Nafplio, Greece, Vol. XLII-2/W3, pp. 639–44, doi:10.5194/isprs-archives-XLII-2-W3-639-2017.