Fisheye Photogrammetry to Survey Narrow Spaces in Architecture and a Hypogea Environment

Luca Perfetti [a], Carlo Polari [a], Francesco Fassi [a], Salvatore Troisi [b], Valerio Baiocchi [c], Silvio Del Pizzo [b], Francesca Giannone [d], Luigi Barazzetti [e], Mattia Previtali [e] and Fabio Roncoroni [f]

[a] 3D Survey Group, Politecnico di Milano, Architecture, Built environment and Construction engineering (ABC) Department, Milan, Italy; (luca.perfetti, carlo.polari, francesco.fassi)@polimi.it
[b] Centro Direzionale Isola C4, Parthenope University of Naples, Naples, Italy; (salvatore.troisi, silvio.delpizzo)@uniparthenope.it
[c] DICEA, Sapienza University of Rome, Rome, Italy; valerio.baiocchi@uniroma1.it
[d] Niccolò Cusano University, Rome, Italy; francesca.giannone@unicusano.it
[e] Gicarus, Politecnico di Milano, Architecture, Built environment and Construction engineering (ABC) Department, Milan, Italy; (luigi.barazzetti, mattia.previtali)@polimi.it
[f] Gicarus, Polo Territoriale di Lecco, Lecco, Italy; fabio.roncoroni@polimi.it

Abstract: Nowadays, the increasing computation power of commercial grade processors has actively led to a vast spreading of image-based reconstruction software as well as its application in different disciplines. As a result, new frontiers regarding the use of photogrammetry in a vast range of investigation activities are being explored. This paper investigates the implementation of fisheye lenses in non-classical survey activities along with the related problematics. Fisheye lenses are outstanding because of their large field of view. This characteristic alone can be a game changer in reducing the amount of data required, thus speeding up the photogrammetric process when needed. Although they come at a cost, field of view (FOV), speed and manoeuvrability are key to the success of those optics as shown by two of the presented case studies: the survey of a very narrow spiral staircase located in the Duomo di Milano and the survey of a very narrow hypogea structure in Rome. A third case study, which deals with low-cost sensors, shows the metric evaluation of a commercial spherical camera equipped with fisheye lenses.

Keywords: fisheye; photogrammetry; videogrammetry; narrow spaces; 3D modelling

1. Introduction

Fisheye camera models for photogrammetric applications were extensively studied, tested and validated in the first decade of the 2000s. Calibration procedures were presented by Abraham and Förstner (2005), Schwalbe (2005), Van den Heuvel et al. (2006) and Schneider et al. (2009), among others.

The recent introduction of the fisheye camera model in some commercial packages for automated image-based 3D modelling (such as Agisoft PhotoScan, Pix4D and Bentley ContextCapture) has allowed both professional and "less expert" users to generate 3D models in a fully automated way, starting from a set of digital images. Results presented in Strecha et al. (2015) confirm the new level of automation achievable for the different steps of the image modelling workflow: camera calibration, dense matching and surface generation.

Such a level of automation for fisheye cameras is quite similar to the automation already achievable in projects based on central perspective cameras (pinhole cameras). However, the risk of unreliable and "crude" digital reconstructions because of the lack of expertise in basic surveying concepts has already been described in Nocerino et al. (2014), in which the authors presented inaccurate reconstructions obtained from pin-hole (rectilinear projection) images.

In the case of a fisheye lens, the short focal length coupled with an extreme distortion makes automated 3D modelling more complicated. This could provide inaccurate 3D models without metric integrity.

The incorporation of the fisheye camera model in commercial software is a clear indicator of how users are becoming more familiar with such distorted projections, not only for photographic purposes but also for metric applications. Nowadays, automated fisheye image processing is possible without turning them into pinhole images.

This paper presents three different case studies with the aim of testing the current state-of-the-art regarding the usability of fisheye optics into different productive workflows.

The manuscript is divided as follows:

- The introduction opens with a small summary of the differences that exist among the available optical projections;
- in Section 2, the authors explain the motivations behind the three case studies presented here;
- Sections 3,4,5 deepen the topic by trying to evaluate whether it is currently possible, or not, to implement a functional workflow based on fisheye lenses, in particular: for 3D architectonic modelling, the localisation of narrow spaces and the extraction of fast shapes and volumes.
- Finally, Section 6 draws conclusions from the results.

1.1. Difference between Fisheye Lenses and Rectilinear Lenses

The main issue with regard to using fisheye lenses concerns the high probability of obtaining incomplete, weak and inconsistent results—a consequence of the considerable radial distortion. The main drawback appears to be the difficulty to take control of the variables that could invalidate the photogrammetry process: first of all, the questionable approximation of the radial distortion by the radial distortion coefficients (K1,K2,K3,…), and secondly, the unfamiliar correlation between the fisheye projection and the design of the capturing phase.

This happens when one tries to use the same consolidated pipeline designed for the rectilinear projection lenses with fisheyes. A fisheye lens is not a rectilinear lens: it is critical to understand the difference between the different optical projections in order to avoid incorrect survey planning in the first place.

A very wide FOV and a very short focal lens do not make a lens a fisheye. It is the particular interaction between the two, focal length and FOV, which makes the difference. The relation between the two parameters, which consists of the optical function, defines the characteristics of the lens. Each optical function maintains its own relation between focal length and FOV. For the same focal length, a different FOV can correspond to each available optical projection. The main advantage of fisheye lenses is that the incoming light beam converges on a circumference of a shorter radius on the sensor than a rectilinear lens at a given focal length.

The main type of available optical projections are the following (Ray, 2002; Kannala, 2006):

$$\text{Rectilinear:} \quad r = f \tan(\theta) \tag{1}$$

$$\text{Equidistant:} \quad r = f\,\theta \tag{2}$$

$$\text{Equisolid:} \quad r = 2f \sin\left(\frac{\theta}{2}\right) \tag{3}$$

$$\text{Stereographic:} \quad r = 2f \tan\left(\frac{\theta}{2}\right) \tag{4}$$

$$\text{Orthographic:} \quad r = f \sin\,(\theta) \tag{5}$$

where r = distance from the centre of the sensor

f = focal length
θ = angle of incidence of the light beams

The first one is the classical "pin-hole" projection also known as perspective projection (when the medium in which the light beams travel remains the same both outside and inside of the dark camera). The others are all different types of fisheye projections.

It must be noted that a FOV of 180° is impossible with the perspective projection, whereas with the fisheye projections a 180° FOV angle is always possible. The equidistant and equisolid angle projections can theoretically reach 360° (depending both on the focal length and the sensor size). The widest fisheye lens ever designed, though never pushed to mass production, is an equidistant-based projection which can reach a 270° field of view: Nikon 5.4 mm f/5.6 (U.S. Patent 3,524,697).

The advantage of a wider FOV alone can be crucial to obtain manageable data, where a rectilinear lenses approach would be prohibitive. This advantage comes with a price: the issues regarding the use of fisheye optics are numerous and linked to the fact that fisheyes follow a different type of optical projection.

1.2. The Ground Sampling Distance Degradation

The Ground Sampling Distance (GSD) is a fundamental concept to start planning the survey. It expresses the resolution, in object space, of the acquired images and, as a consequence, the potential accuracy of the 3D reconstruction.

The rectilinear optical projection (1) can be schematized in the pinhole projection scheme, where it is easy to recognise a similarity between triangles, and therefore the following ratio is instantly deduced:

f (focal length) : D (capturing distance) = pixel size : GSD, this straightforward and standard tool gives control over the results of the survey itself to the survey operator. The concept of GSD and the possibility of calculating/imposing it, provides the connection to the precision of the two- or three-dimensional representation to be extracted from the survey.

Though fisheye lenses follow a different optical projection, the abovementioned simple tool can no longer be employed in the survey designing phase. The GSD is indeed variable across the image frame; it goes from the minimum value, in the principal point, where the value is the same as the rectilinear projection at the same focal length, to a maximum value in the frame corner. The GSD can also reach infinity when the field of view is 180°.

Since the accuracy of the 3D reconstruction depends on the inferior resolution, the variable GSD hampers the operator's ability in designing the survey and, as a consequence, the control over the result. The risk of measurement's failure in the case of incorrect photogrammetric planning is very high and it leads to a warped model (bent, stretched, compressed, etc.) (Nocerino, 2014).

Figure 1. Two circular fisheye images of a straight wall; the same area was rendered at a very different resolution when placed in the centre or on the border of the frame.

2. Why Use Fisheye Lenses for Photogrammetry

2.1. Fisheye Lenses to Survey Narrow Spaces in a Complex Architectonic Environment

Nowadays, the use of Building Information Modelling (BIM) technologies and 3D representations in the architectural field has increased the demand for 3D survey techniques. For this reason, in the last few years, the world of Cultural Heritage has seen a strong and constant development of survey technology and practices aimed at acquiring complete, dense, and high-precision 3D information.

Dealing with a full 3D reconstruction of complex architectures, secondary spaces (such as staircase, corridors, passages and tunnels) play a major role. If the goal of the survey is to produce a BIM model of the building, they cannot be forgotten just because of their reduced accessibility.

Due to the lack of the "right instrument", surveying those areas results oddly in the most time-consuming and challenging part of the measurement campaign. The fisheye lenses approach could be the most promising solution to speed up the acquisition phase by reducing the total amount of pictures to be acquired, without compromising the achievable accuracy.

2.2. Fisheye Lenses for Quick Surveys in Critical Underground Environments

Very similar difficulties are encountered when surveying complex structures of subterranean spaces present in modern and historic city centres. Most of them are used to provide an efficient traffic-free transit system (i.e., subway, underpass, etc.), while others host a series of infrastructures such as several cable ducts or a

drainage system. Every underground space should be accurately mapped and georeferenced in order to simplify both the ordinary and extraordinary maintenance.

Actually, most parts of big cities do not have an accurate map of their underground spaces, although new evolutions of GIS (Geographic Information System) technology are starting to allow the management of such data. Moreover, while new infrastructures are accurately mapped, old ones often remain unsurveyed.

Generally, whether the metric survey is carried out in indoor or in hypogea environments, it is usually expressed in a Local Reference System (LSR), which in turn is linked to a global reference system, such as the geocentric coordinate system ETRS-2000. The relationship between the LRS and the global one is not always well known. The transformation parameters could be determined by a classical topographic survey performed with the use of a total station, where both the accuracy and the precision achievable are strictly related to the survey adopted techniques (i.e., open and/or close traverses). Furthermore, the specific technique adopted is conditioned by several factors: the morphology and position of the hypogea site, the extension of the site as well as the number of breaches that allow open spaces to be reached. In specific critical cases, when the extension of the hypogea site is huge, and the quality of the air can change quite rapidly even becoming unbreathable, a different kind of survey could be suitable: quicker than the topographic one, even if less rigorous. The solution suggested is to walk through the tunnel taking pictures (video frames) from a single camera, equipped with a fisheye lens to ensure a wide FOV, pointed forward while advancing in the tunnel. The reconstruction of camera movement as well as the 3D model of the hypogea environment will be successively obtained by the use of Structure from Motion (SfM) algorithms. Finally, in order to geolocate and to scale the obtained 3D model, several GCPs (Ground Control Points) have to be placed where possible; furthermore, several scale constraints have to be imposed in order to limit the model deformations.

2.3. Fisheye Lenses and Panoramic Cameras for Speeding up the Survey Phase of Closed Environments

Consumer-grade cameras able to capture 360 photos and videos are becoming more popular due to the opportunity to look in any direction, exploiting immersive visualisation 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, Sphericam V2, Nokia OZO, and GOPRO OMNI. Typically, these cameras are equipped with at least two fisheye lenses and provide a spherical 360°

panorama (Fangi, 2010). Due to the stable geometry of a cylindrical panorama, photogrammetric bundle adjustment can be performed with very few object points. Once each panorama is oriented in the global coordinate system, photogrammetric object reconstruction procedures such as space intersection can be applied (Luhmann, 2010). The idea behind the use of this kind of camera is to reduce the acquisition time using popular commercial instrumentation as well as simplifying the capture geometry approach so that even the less experienced operator has the possibility of performing a 3D reconstruction of close and complex environments. The camera's constructive and photographic qualities, the low-resolution acquisitions and the internal panorama stitching can negatively influence the quality of the results. Chapter 5 will investigate this aspect.

3. Fisheye Lenses for Cultural Heritage Survey

3.1. General Model to Calculate the GSD for Fisheye Lenses

The main drawback of fisheye lenses for photogrammetric use can be identified in the degradation of the GSD across the image frame. The GSD is the link to the chosen representation scale. It is commonly used as the primary and only reference on which to base the capturing phase design; losing the ability to calculate it a priori led to us losing control over the final results of the matching process, giving rise to the topic of fisheye lenses not being reliable enough for metric purposes.

Although an accurate calculation of the GSD is not required in the design phase, an approximate calculation is mandatory to have control over the image resolution.

A general mathematical model for describing the GSD is presented here; this theoretical approach allows one to evaluate different lenses and different optical projections a priori by deriving a specific GSD formula from the general one for each of them. One can say that the commonly used GSD definition is a specific formulation suitable only for rectilinear lenses, but not for fisheye lenses.

For all fisheye projections, unlike the pinhole scheme, it is true that the GSD varies also depending on the radial distance from the centre of the frame. Therefore, the new model has to express the GSD as a function of focal length, f, capturing distance, D, and pixel size, pix, as well as the parameter r.

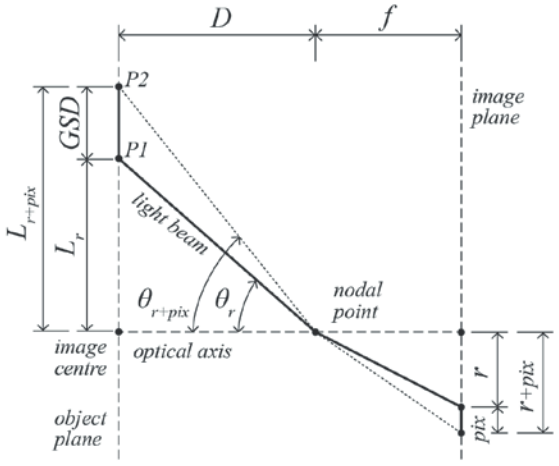

Figure 2. The relation between pixel size and GSD for each projection function.

Equation (6) expresses the GSD considering both the classical parameters usually considered in photogrammetry, D, f and pixel size with the addition of r to take into account the GSD degradation through the frame.

$$GSD = D \cdot \left\{\tan\left[2\,\text{arcsen}\left(\frac{r+pix}{2f}\right)\right] - \tan\left[2\,\text{arcsen}\left(\frac{r}{2f}\right)\right]\right\} \tag{6}$$

In Figure 2, θ_r and θ_{r+pix} depend on the optical projection of the lens, where theta is always related to r. It is possible to reverse the mapping function of each lens to make theta explicit in order to obtain the GSD formula for the chosen one. Below, a list of the formula for the equidistant (7), equisolid (8), stereographic (9) and orthographic (10) projection function is outlined.

$$GSD = D \cdot \left[\tan\left(\frac{r+pix}{f}\right) - \tan\left(\frac{r}{f}\right)\right] \tag{7}$$

$$GSD = D \cdot \left\{\tan\left[2\,\text{arcsen}\left(\frac{r+pix}{2f}\right)\right] - \tan\left[2\,\text{arcsen}\left(\frac{r}{2f}\right)\right]\right\} \tag{8}$$

$$GSD = D \cdot \left\{\tan\left[2\,\text{arctan}\left(\frac{r+pix}{2f}\right)\right] - \tan\left[2\,\text{arctan}\left(\frac{r}{2f}\right)\right]\right\} \tag{9}$$

$$GSD = D \cdot \left\{\tan\left[\text{arcsen}\left(\frac{r+pix}{f}\right)\right] - \tan\left[\text{arcsen}\left(\frac{r}{f}\right)\right]\right\} \tag{10}$$

The presented system allows the operator to calculate the resolution distribution across the images beforehand, and, as a consequence, to monitor the minimum value of GSD for the various optical projections; since the lower resolution will influence the outcome of the photogrammetric process, it is important to understand which parts of the images can be used in relation to the chosen representation scale. This method makes it possible to calculate the radius r

of the maximum circumference within which the resolution is compatible with the chosen scale; the remaining part, expressing a resolution lower than the acceptable minimum, can be discarded (Figure 3).

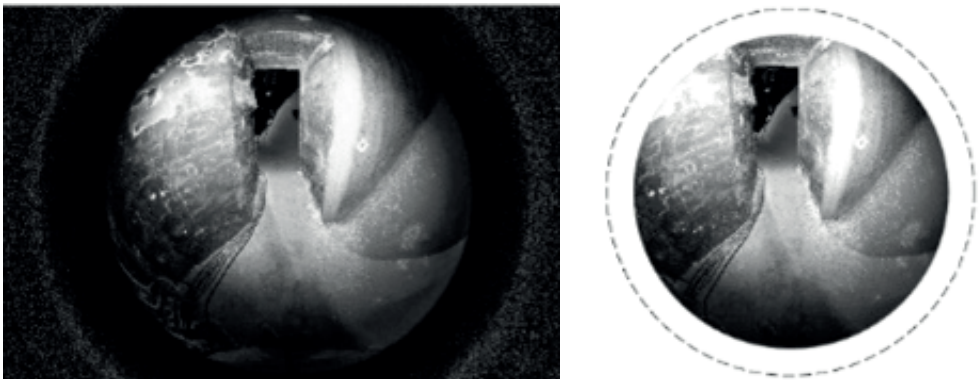

Figure 3. Original picture taken using a fisheye lens and the same image cropped to comply with the desired scale.

By drawing a parallel between the two charts below (Figure 4), the operator can predict in advance, for each case study, the behaviour of fisheye lenses as well as rectilinear lenses, managing to check, always in advance, if there is a significant advantage to using one over the other.

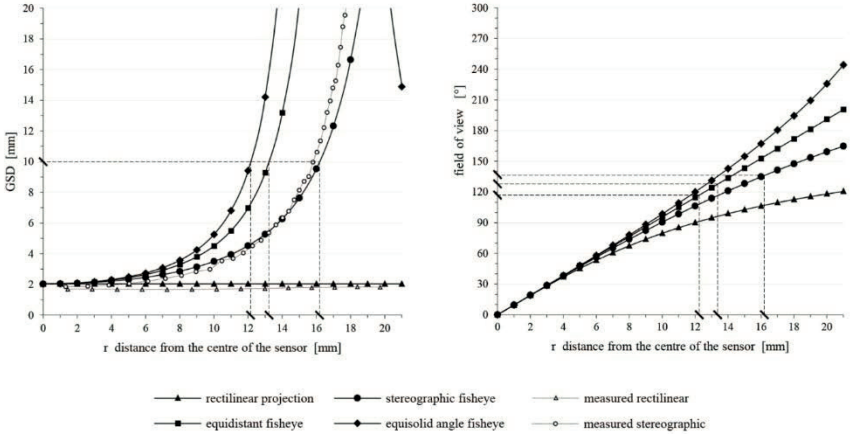

Figure 4. Comparison graph among different optical projection behaviours (left) and FOV that can be reached with each of them at a given radius distance (right). The simulation is done for a camera with a pixel size of circa 4.9 microns, a focal length of 12 mm and a taking distance from the object of 2.5 m.

In the scale 1:50, for instance, typically used in the survey of cultural heritage, the minimum resolution that the photographs must have is expressed by the value of the GSD, maximum 10 mm. At this point, according to the optical projection in use, it is possible to calculate the value of the radius r which defines the circumference that places the photographic area with sufficient resolution on the inside and the one with insufficient resolution on the outside. This information can be used to crop the images, discarding the marginal areas to keep only the portion on the inside of the aforementioned circumference.

3.2 Experimental and Validating Tests

3.2.1. Experimental Test on GSD Distribution

The graph in Figure 4 reveals the behaviour of the different optical projections that have to be taken into consideration in order to correctly design the survey and obtain the desired results. However, the actual distortion of any physical lens is characterised by some small differences when compared to their theoretical mathematical model. Therefore, in order to verify the theoretical model, a test was conducted, making it possible to measure the decrease in the GSD with two fisheye lenses: Samyang 12 mm fisheye stereographic and Sigma 8 mm fisheye equisolid.

The aim of this test was to obtain a sort of "scale factor" able to parameterize the fisheye lens behaviour, in terms of lens distortion and in relation to the rectilinear theoretical parameters. This will then take two aspects of the distortion into consideration at the same time: the difference between the two optical projections in question (rectilinear and fisheye), as well as manufacturing defects of the lens.

To obtain this "scale factor", a target was designed and attached horizontally to a wall, while the camera was positioned tilted, ensuring that the target was lying exactly on the diagonal of the sensor passing exactly into the centre of the frame (Figure 5). At this stage, a great deal of attention was paid to ensure the maximum parallelism and centring of the sensor with the target. Using this configuration, it was possible to measure the compression of the lengths on the metric scale of the target, and therefore, to compare it with the theoretical model.

Although the behaviour of the real lens (fisheye) differs slightly from the theoretical one (Figure 4), the variation is sufficiently limited to allow the theoretical model to be used for the design of the survey with real optics (Perfetti et al, 2017).

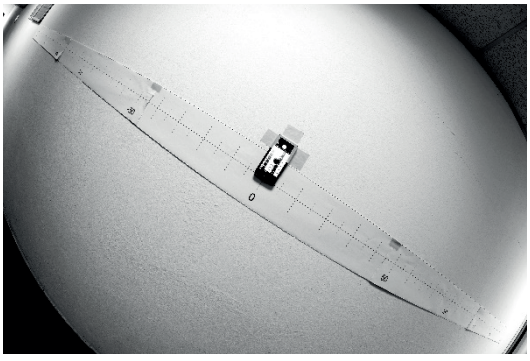

Figure 5. Photo of the metric target obtained with a diagonal fisheye, the Samyang 12 mm.

3.2.2. Experimental Tests on Ideal Capture Geometry: Base Distance/Capturing Distance Ratio

The next step is to define an optimal capturing geometry on which the photogrammetric survey can be based. Although the capturing geometry varies from case to case depending on the volumetric features of the survey object, it is important to define the *base distance/capturing distance* ratio which underlies the capturing geometry design.

While with rectilinear projection lenses the calculation of the *base distance/capturing distance* ratio comes from the percentage of overlap between adjacent images (about 70-80%), with fisheye lenses this approach fails. Due to the wider field of view of fisheye lenses, the overlap of two adjacent images lying on the same plane would be bigger, being even equal to infinity when the FOV is equal to 180° or more. It is clear that the percentage of overlap is not a useful parameter on which to base the capturing geometry. Many empirical tests were held in order to overcome this issue: for instance, a photogrammetric survey of a straight wall was performed using a high number of photographs precisely spaced; little by little, some images were removed with the aim of obtaining a correct survey with a minimum number images.

After these tests, a 1:1 *base distance/capturing distance* ratio turned out to be the best ratio to be used with fisheye lenses. This information makes it possible to design a suitable capturing geometry when the working distance is known.

3.3. The "Fisheye-Grammetric" Survey of the Minguzzi Spiral Staircase

The Minguzzi staircase is a marble stone spiral staircase located inside the right pylon of the main facade of the Milan cathedral. Along its extension of about 25 m in height, it connects three different levels of the building: at the upper

end, the lower level of the roofs; in the middle, the central balcony of the façade and, at the base, the floor of the church.

The staircase is extremely dark. The artificial illumination is of poor quality, and there are only a few openings towards the outside placed at regular intervals. From the inside, these openings are relatively large (85 cm) and deep (circa 2 m) but due to the considerable thickness of the wall, they narrow down noticeably towards the outside and end up as small vertical embrasures (30 cm width). At the centre of the staircase, there is a stone column with a diameter of about 40 cm around which the ramp rolls up; the space left for the passage is extremely narrow, about 70 cm. This extremely narrow space was complex enough to represent a serious test of our research topic. In this situation, there is simply no space to use the regular terrestrial laser scanner instrumentation. Moreover, many other factors would nullify the whole scan, making the job unnecessarily time-consuming and burdensome in terms of the amount of data.

Before moving on to the field, it was imperative to fully understand the geometry of the object in order to properly design the capturing geometry and to completely cover all the spaces useful for the re-design of the staircase.

Figure 6. Fisheye views of the internal winding of the staircase.

3.3.1. Acquisition Phase

The adopted shooting geometry was defined throughout some experimental tests aimed at finding out the right *base distance/capturing distance* ratio to be used with these optics (Perfetti et al, 2017).

The image acquisition phase was divided into five main placements of the camera. In addition, a series of integration acquisitions were taken to complete some complex areas.

The survey was carried out with two different cameras and three different lenses: Canon 5D mark III coupled with Sigma 8 mm circular fisheye and Nikon D810

coupled with Samyang 12 mm diagonal fisheye and a Sigma 12–24 mm rectilinear lens. The whole image acquisition process was accomplished with the aid of three portable synchronised photographic flashes, which made it possible to light up the otherwise too dark area of the staircase.

The different field of view and the GSD calculated value have determined the number of shots needed for each configuration.

Some scans using TLS Leica C10 were performed at the base and the top of the staircase. These scans were georeferenced in the topographic network built for the cathedral survey activity; some photogrammetric targets were measured at this stage as well. These measurements were performed in order to check the accuracy of the survey and in particular if any bending, stretching or compression distortion occurred to the final model. In this way, we had two check-stations at the base and the top of the staircase. In the interior of the staircase, where no additional instrumental check was possible, many reference distances among markers were taken in order to monitor the alignment manually.

3.3.2. Data Elaboration and Results

First, before proceeding with the software elaboration, it was important to apply the discussed methodology to crop the marginal areas of the pictures when not suitable for the desired scale.

For all the three tests, but in particular, for those with the fisheye lenses, it was necessary to manually intervene on the alignment after the automatic matching process, to provide additional manually selected constraints by manually picking the targets (Photoscan coded targets) that the software could not recognise due to fisheye distortion. This operation was necessary in order to optimise the calculation of the camera alignment.

Table 1. Residual errors RMSE on Ground Control Points (GCPs) for the different elaborations during the photogrammetric process.

Process	CANON 5D 8mm fisheye	NIKON D810 12mm fisheye	NIKON D810 12mm rectilinear
Allignment	0.032m	/	0.055m
Marker optimization	0.031m	0.013m	0.051m
Scale-bars optimization	0.010m	0.014m	0.051m
Topography optimization	0.009m	0.014m	0.051m

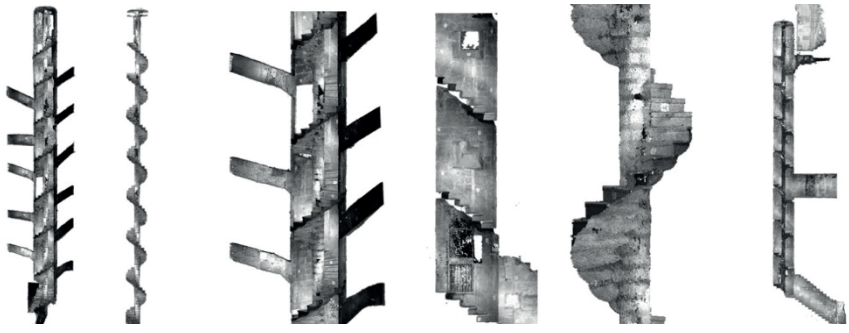

Figure 7. Dense point cloud obtained with Nikon D810 coupled with 12 mm stereographic fisheye; elaboration was done using Agisoft Photoscan Pro.

The results show (Figure 7) that using the GSD calculation to hold the resolution under control leads to a significant improvement of the alignment and matching process that produces more reliable results and less noisy dense point cloud; it could also help the camera calibration model by removing the most distorted portion of the frame. The best results were obtained using the Sigma 8 mm equisolid fisheye; in this case, the use of the cropping method was very successful (Table 1); the final Root Mean Square Error (RMSE) of 9 mm corresponds to the expected/designed accuracy.

The remaining issue is the very complex capturing geometry that has to be arranged case by case. No standards can be followed a priori, and the use of a large number of markers and a lot of manual refining work represented the key to success in the process: a complete autonomous elaboration when using fisheye lenses can only lead to a failure. Topographic data are necessary to improve the final accuracy.

4. Fisheye Lens Survey to Georeference Complex Hypogea Environments

4.1. Site Description

The ancient Romans mined the soft rock from underground to erect buildings (i.e., "pozzolana"); actually, this activity was carried out until the middle of the twentieth century. Such action has involved the appearance of a series of tunnels, about 10–15 m deep, realising an intricate labyrinth in Rome underground. There is no topographic map of this intricate labyrinth excavated more than 2000 years ago; furthermore, the total extension is still unknown. In this application, an accurate mapping of part of this complex system was needed to find the planimetric position of a specific point of one of these galleries located in the southern part of the city.

Several researchers focused their efforts on the survey of underground structure, such as necropolis (Remondino et al., 2011) or catacomb (Bonacini et al., 2012). Very low luminosity is a typical feature of a hypogea environment; therefore, the laser scanner is usually employed as the main instrument, while photogrammetry is generally used to obtain a high-quality texture. The results are very noteworthy, and the quality of the survey is very high, but these techniques are very time-consuming. For this specific application, a very agile methodology is required, because in every moment, due to a change of air current direction, some dumps present in the galleries itself could yield toxic air. Furthermore, for heritage conservation reasons, one is not allowed to leave permanent and invasive markers.

4.2. Photogrammetry Setup

The survey was carried out using a full frame DSLR camera Nikon D800E with a pre-calibrated Nikkor 16 mm fish-eye lens. Fisheye lenses allow a wide diagonal FOV of up to 180° to be achieved. Furthermore, the choice of fisheye lenses allows a substantial overlap between two consecutive frames to be achieved. The camera was set in video mode, and recording was performed in standard HD (High Definition) 1080 px on 30 fps (frame per second). Such camera settings allow to obtain a pixel size of 18.7 microns.

4.2.1. Camera Calibration

The camera calibration procedure is a fundamental task in the photogrammetric workflow. The well-known self-calibration method (Fraser, 1997) is generally used to determine the camera calibration parameters.

Specifically, this technique employs analytical calibration methods to derive the calibration parameters indirectly from photogrammetric image coordinate observations. The mathematical model is the classical Brown model (Brown, 1971) that is composed of classical internal orientation parameters (principal distance, coordinates of the principal point, pixel size) extended by the inclusion of additional parameters that model the image distortion effects. Fisheye lenses utilise a different optical design that departs from the central perspective imaging model to produce image circles of up to 180°. If the image format sensor is larger than the resultant image circle, the camera is termed as a fisheye system. Conversely, if the format is smaller than the image circle, such that the image diagonal covers about 180° of the field of view, a quasi-fisheye system is attained. It is well known that under planar perspective projection, images of straight lines in space have to be mapped into straight lines in the planar perspective image. However, such assertion, for fisheye cameras, is not true: the central perspective mapping is replaced by another model such as stereographic, equidistant and orthographic projection.

When modelling the distortions in a fisheye lens, conventional radial lens distortion corrections are mathematically unstable. In the peripheral region of the image sensor, the gradient of the distortion curve describing the departure from the central perspective case is high. In such a case, it is necessary to apply the appropriate fisheye lens model before using the conventional radial distortion model (Luhmann et al., 2014).

In this work, a quasi-fisheye system was realised; therefore, the peripheral portion of the image circle is not recorded on the sensor, and the fisheye distortions can be reasonably modelled using the conventional perspective camera model and its classical radial distortion coefficients. Such a model is more flexible than the fisheye model because it can be imported/exported in/from any photogrammetric software as well as integrated into any SfM algorithm.

The camera was preliminary calibrated in a controlled environment before starting any survey operations. Two types of calibration camera models were tested:

1 *Fisheye camera*: it is the more rigorous camera model; it combines two types of distortion: the fisheye distortion that the lens provides as well as the conventional pin-hole ones.
2 *Perspective camera*: it is the classical model used in photogrammetry, i.e., the Brown one. It uses only the conventional distortion model.

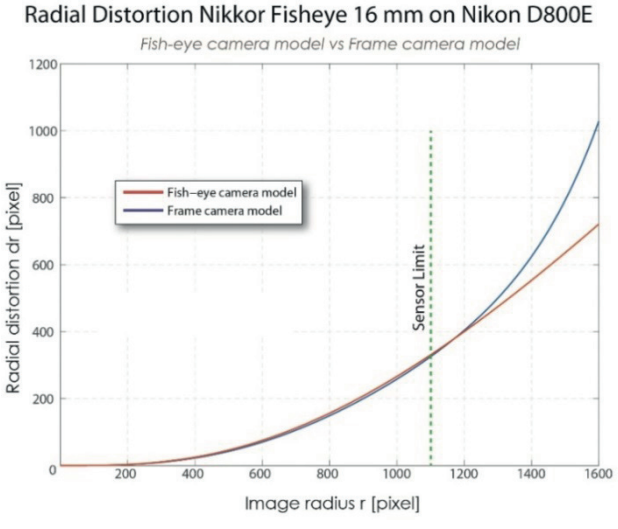

Figure 8. Radial distortion curves: in blue are the distortions obtained using the frame camera model, while in red are the distortions determined with the fisheye camera model.

In our tests, the two models provided the same values of image radial distortions in the central part of the sensor (close to the principal point), whereas on the peripheral parts the gap between two models increases progressively.

In order to quantify the differences between the two models, the components of computed radial distortion were plotted in Figure 8. Basically, the two curves are identical; just on the image periphery, the difference becomes approximately a few pixels.

In this case, study of the gap between the two camera models is primarily restricted because video recording automatically crops the frame and the image sensor is entirely included in the image circle (where the FOV is 180°) as shown clearly in Figure 9.

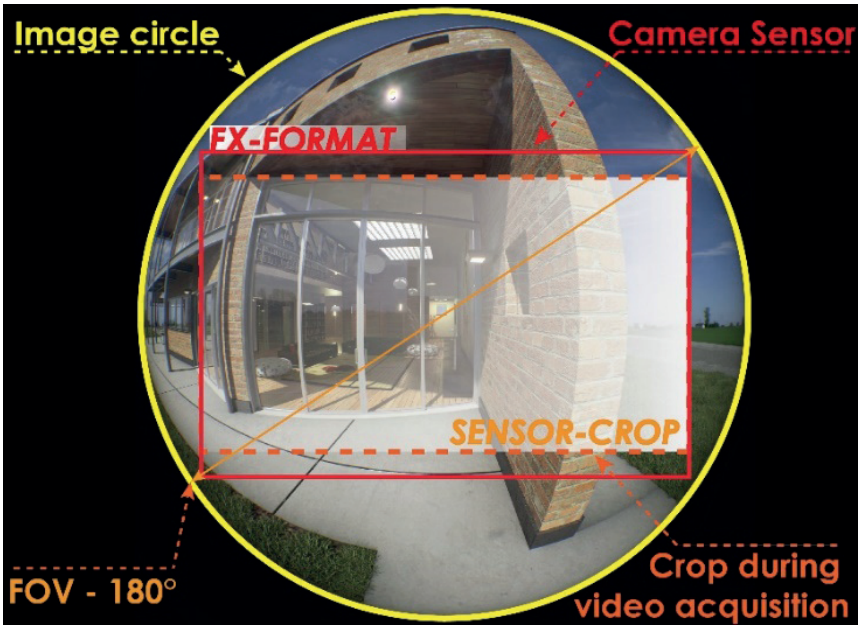

Figure 9. Representation of the full frame fisheye lens. In yellow is the image circle, in red is the coverage of the FX sensor; the orange dotted line shows the sensor active area during the acquisition.

4.2.2. Acquisition Phase

The acquisition phase started outside with the aim of acquiring sufficient targets in order to georeference the entire survey. The camera operator slowly brought it downhill to the tunnel labyrinth. The transition between the very bright environment (open-air) and the dark location (tunnels) was gradually performed. The focus camera was set to infinite, maximum aperture, while the sensitivity of the digital image system ISO automatically changed, to adjust the great bright

gradient. The data collection was performed holding the camera optical axis aligned with the path axis (Figure 10). The acquisition phase lasted less than two hours to cover about 1 km and it consisted of a simple video tape acquisition recorded along the path. For the duration of the acquisition, some caution was taken: the light source is never located in front of the camera and within the scene; no person or moving object was detected, except for the shadow of the camera operator projected by the light source located behind him.

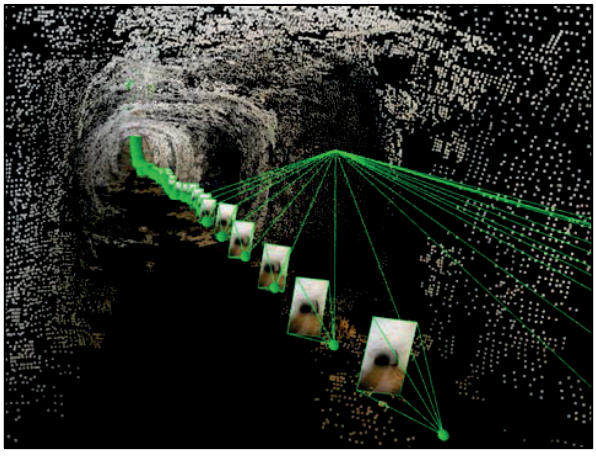

Figure 10. Camera geometry carried out during the acquisition phase, moving forward along the optical axis.

4.3. Elaboration and Results

4.3.1. Relative Image Orientation

The acquired video was processed using a standard procedure employed in 3D image-based modelling by sampling, on average, a frame every 1–3 seconds, for a total of 4425 camera stations.

In order to correctly position the light source, the video tape was not recorded continually; indeed, 33 video clips were acquired. Each clip was processed as a different sub-project and was oriented independently using internal camera parameters computed during the calibration task. Two following sub-projects have at least 20 images in common. Such overlap permitted the sub-projects to be joined among them. Figure 11 shows the number of extracted tie-points in the joint model. Such a parameter is very high, both for the great overlap assured by the fisheye lens and for the feature extractor employed during the orientation phase. A 3D sparse point cloud has been obtained by about 4200 images correctly oriented; indeed, due to imposed overlapping camera stations, the number of

unique images was reduced. The bundle result obtained with such camera stations provided about 2.5 million tie-points.

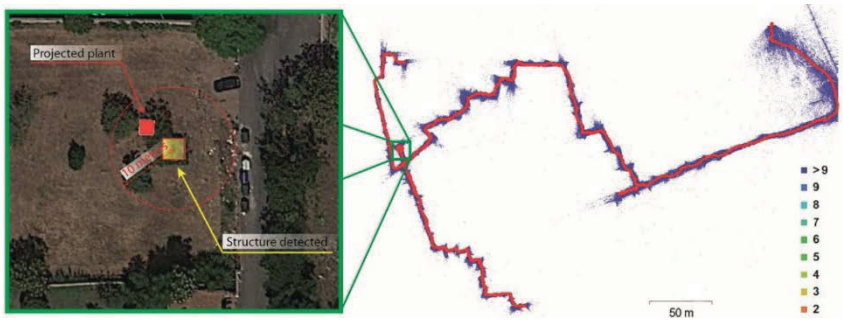

Figure 11. On the right: an orthographic projection of 3D tie-points, the colour shows the associated multiplicity whereas in red is the track of the moving camera. On the left: a detailed comparison between the structure inspected in the underground environment (in red) and the corresponding aboveground environment.

4.3.2. Scaling and Georeferencing

The photogrammetric model needed to be geo-referred using at least three GCPs (Ground Control Points). This step was performed at the beginning of the video recording. Specifically, the acquisition started outside the tunnel where three GCPs were signalised by targets. The GNSS observations, collected in fast-static mode, were processed and adjusted; the final 3D global accuracy is about 2–3 cm. Such achieved accuracy is more than enough for the goal of this work, where an accuracy of about 10 m is required. Further scale constraints were added to the project; indeed, as described in the previous section and during the acquisition phase, several measurements of some distances were carried out. The 3D model was scaled and georeferenced by three GCPs and two scale constraints. Initially, the bundle adjustment was conducted in free network mode; afterwards, adding GCP measurement, it was performed in minimal constraint mode. The total error reported on the GCP is about 1 cm. In order to limit these scale deformations, the bundle adjustment was carried out again, enhancing the constraints. The total error in the final three-dimensional reconstruction has grown up to about 6 m.

4.3.3. Obtained Results

The result is a geo-referred 3D point cloud of the explored tunnels. Finally, such a geo-referred sparse point cloud was further processed to obtain an orthoimage map. The hypogea site does not provide any reference point to control the solution, and the tunnels network is very complex; in any case, on the path,

several structures—probably well casing or building foundations—are present. Unfortunately, such structures are not easily recognisable on the surface. In order to verify the accuracy of the survey, every structure encountered along the underground pathway was inspected afterwards. By projecting the plant on Google Earth, a rectangular structure on the surface was discovered (Figure 11). Such a discovery verifies the 3D model obtained (only in planimetry). Comparing the positions of the barycentre, an estimated accuracy within 7 m was achieved.

4.4. First Conclusions on Complex Hypogea Environments Test

The complexity of this specific experimentation was further increased both by the environmental conditions (such as the low brightness, no electric energy, low level of oxygen, etc.) and by the morphology of the site. A generic camera HD and a simple measuring tape, further to a GNSS receiver, composed the equipment. Initially, no reference point was detected to check the solution or to add additional constraints. This type of approach allows one to conduct a prompt survey in a short time; furthermore, the essential equipment is easily reachable, and it does not require high profile skill to be managed. On the other hand, expertise and time are necessary for post-processing to obtain a reliable solution.

5. Photogrammetry with Fisheye Images from a Low-Cost Spherical Camera

The camera considered in this work is the Samsung Gear 360, which has a dual 15 MP CMOS sensor with integrated f/2.0 fisheye lenses, dual cam video resolution of 2840 × 1920 pixels, and dual cam photo resolution of 7776 × 3888 pixels. The camera requires a Bluetooth connection with a Samsung mobile phone (such as the Samsung S6 or S7) to obtain real-time visualisation and control acquisition parameters. The images can be downloaded from the camera as circular fisheye images or equirectangular projections for 360° visualisations.

The aim of this work was to try out the Samsung Gear 360 for 3D modelling, considering front and rear-facing fisheye images. The interest in this kind of acquisition tool is motivated by the wide field of view available, which makes spherical cameras very attractive for interior scenes that usually require a large number of pinhole images.

5.1. Evaluation of Metric Accuracy

The aim of this experiment was to test the metric accuracy of the Samsung 360 with a long image sequence. 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 fisheye images to be processed 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 (Figure 13). Images were oriented with ContextCapture, using 12 targets as ground control points and seven targets as check points. The sequence is 42 m long, and the camera object distance is 1.2 m. Statistics are shown in Table 2 and reveal an error of about 5 mm, that confirms a good metric accuracy for the project carried out 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.

5.2. 3D Modelling from Fisheye Images Acquired with the Samsung 360

The results described in the previous section revealed a good metric accuracy when the single fisheye images are used; a reconstruction based on two fisheye images (front and rear) seems feasible. Figure 14 shows the image orientation results inside the oratory of Lentate sul Seveso in the case of fisheye image processing (30 images, i.e., 15 + 15 image pairs). 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 calibration parameters were assumed as fixed for both front- and rear-facing images.

The achieved mesh is of better quality than that generated by equirectangular image processing presented in Barazzetti et al. (2017). 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 12. Example of front- and rear-facing fisheye images acquired with the Samsung 360.

Table 2. Accuracy achieved with the front-facing camera. The good metric accuracy is achieved by the short camera–object distance, as well as the short baseline between the images, i.e., the same point is visible on a large number of images providing multiple intersections.

Number of GCPs	RMS of reprojection errors [pixels]	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 [pixels]**	**RMS of horizontal errors [m]**	**RMS of vertical errors [m]**
7	2.1	0.004	0.001

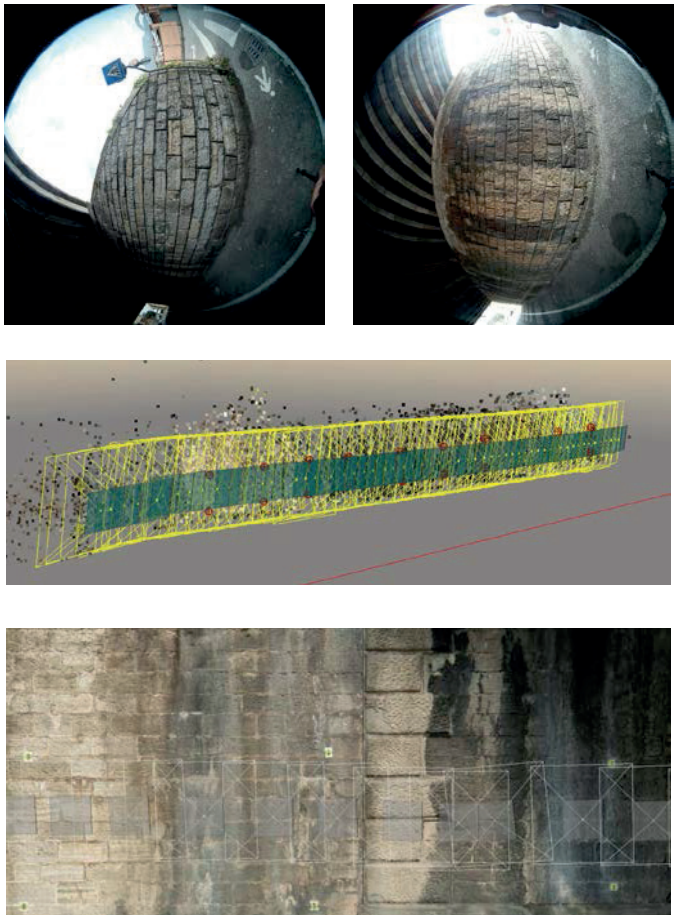

Figure 13. Two images of the wall captured with the front-facing camera (top), the sequence of images, control points and checkpoints (middle), and a detail of the extracted mesh for a portion of the wall (bottom).

Figure 14. The reconstruction from front- and rear-facing fisheye images is quite detailed but incomplete, especially the area above the camera; notwithstanding, the area is visible (with a very narrow angle) in the original fisheye images.

6. Conclusions

The three presented case studies aim to briefly describe the possible use of fisheye photogrammetry to solve some common environmental and architectural measurement problems. First, the conducted research tried to address the problem of designing a complete photogrammetric survey of narrow spaces, characterised by scarce illumination and intricate geometry. The second goal was to speed up the photogrammetric acquisition in this type of areas, especially when they are complicated and also huge in their extension, by reducing the number of images to be acquired. This can be obtained either by using extremely wide-angle lenses, or by reducing the number of images using low-cost commercial panoramic cameras. The idea for solving both problems was to use fisheye lenses and find rules that could standardise this type of survey and the ensuing elaboration process.

In the first case study, the architectonic reconstruction of narrow internal spaces, the goal was to obtain an accurate, high-resolution 3D reconstruction of the building. The theory of fisheye distortion and some preliminary tests suggest concentrating the attention on the non-uniform distribution of the image resolution across the frame. The idea was to use only the portion of the image that provides the GSD within the maximum acceptable limit for the desired restitution scale. The test applied to the Minguzzi staircase case study demonstrates the effectiveness of the approach and the possibility to obtain complete and accurate results at the architectonic representation scale of 1:50.

The second case study describes the use of alternative methods for a rapid reconstruction of underground environments. The aim was not to obtain a complete, high-resolution 3D reconstruction, but to quickly derive the position on the outside of an underground point. The research focused on testing the "videogrammetry approach" to speed up the acquisition and check the capabilities of automatic photogrammetric algorithms in the arduous task of orienting a high number of images with the purpose of substituting the classical topographic open

traverse. The tests showed the real possibility to conduct a prompt survey in a short period.

The third case study showed some preliminary tests of a ready-made instrument based on multiple fisheye lenses, able to speed up the acquisition phase and regularise the capturing geometry. The first results of the experiments revealed that the Samsung Gear 360 (the camera used in the test) is suitable for metric reconstruction, although the achieved metric accuracy is not comparable with a traditional photogrammetric approach based on pinhole images. The achieved accuracy is within the tolerance of the 1:300 representation scale, which could be sufficient for applications aimed at determining the overall size or volume of a room. The proposed setup is surely less expensive than a laser scanner and also allows rapid data acquisition where, instead, a large number of pinhole images would be needed. On the other hand, there is limited control on camera parameters and images are acquired in an almost entirely automated way. This makes the camera a low-cost photogrammetric tool for people that have limited experience in 3D modelling from images.

The presented case studies show how fisheye lenses can actually solve the problem of speeding up the acquisition and therefore, how their employment allows the reconstruction of complex narrow environments. Fisheye lenses are therefore a valid tool in practical applications where traditional lenses (central perspectives) would require an enormous number of images. The outcomes of the previous experiments allowed us to define a set of good practices for "fisheye photogrammetry":

- fisheye lenses mounted on SLR cameras are already a valid tool for metric applications;
- the large field of view of fisheye lenses allows a reduction of the number of images, but the use of the whole field of view produces outputs with a very variable resolution (very low close to the image edges); this is a significant limitation for some steps of the production workflow (e.g., camera orientation and texture mapping);
- camera calibration plays a vital role because the extreme image distortion could result in poor metric accuracy;
- good practices for traditional photogrammetry (e.g., inclusion of control points as pseudo-observations in the bundle adjustment instead of basic absolute orientation techniques) are still valid for fisheye photogrammetry;
- low-cost sensors (such as the Samsung 360°) are not ready for productive work, except for the case of the simplified model with limited metric integrity. On the other hand, significant technological improvement is expected (in terms of resolution and sensor quality), so they have remarkable potential;

The use of a multiple fisheye lenses instrument can significantly reduce the acquisition time but, in the future, it could be necessary to use higher-level instrumentation to improve the quality and the accuracy of the reconstruction in order to fit the architectonic scale requirements. Future works could develop a hand-held fisheye-based instrument capable of solving all the problems described here as well as increasing the resolution and the quality of the results.

References

1. Barazzetti, L.; Mussio, L.; Remondino, F.; Scaioni, M. Targetless Camera Calibration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2011**, *XXXVIII-5/W16*, 335–342.
2. Barazzetti, L.; Previtali, M.; Roncoroni, F. 3D modelling with the Samsung Gear 360. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2-W3*, 79–84, doi:10.5194/isprs-archives-XLII-2-W3-79-2017.
3. Bonacini, E.; D'Agostino, G.; Galizia, M.; Santagati, C.; Sgarlata, M. The catacombs of San Giovanni in Syracuse: Surveying, digital enhancement and revitalization of an archaeological landmark. In *Euro-Mediterranean Conference*; Springer: Berlin/Heidelberg, 2012; pp. 396–403.
4. Brown, D.C. Close-range camera calibration. *PE&RS* **1971**, *37*, 855–866.
5. Covas, J.; Ferreira, V.; Mateus, L. 3D reconstruction with fisheye images strategies to survey complex heritage buildings. *Digit. Herit.* **2015**, *1*, 123–126, doi:10.1109/DigitalHeritage.2015.7413850.
6. Fangi, G. Multi scale, multi resolution spherical photogrammetry with long focal lenses for architectural survey. In Proceedings of the ISPRS Midterm Symposium, Newcastle, UK, 22–24 June 2010; pp. 228–233.
7. Fassi, F.; Achille, C.; Mandelli, A.; Rechichi, F.; Parri, S. A new idea of BIM system for visualization, web sharing and using huge complex 3d models for facility management. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-5/W4*, 359–366, doi:10.5194/isprsarchives-XL-5-W4-359-2015.
8. Fraser, C.S. Digital camera self-calibration. *ISPRS J. Photogramm. Remote Sens.* **1997**, *52*, 149–159.
9. Kannala, J.; Brandt, S.S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1335–1340.
10. Luhmann T. Panorama Photogrammetry for Architectural Applications. *Mapping* **2010**, *139*, 40–45.
11. Luhmann, T.; Robson, S.; Kyle, S.; Boehm, J. *Close-Range Photogrammetry and 3D Imaging*; Walter de Gruyter: Berlin, Germany, 2014.
12. Marčiš, M.; Barták, P.; Valaška, D.; Fraštia, M.; Trhan, O. 2016. Use of image based modelling for documentation of intricately shaped objects. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B5*, 327–334, doi:10.5194/isprs-archives-XLI-B5–327–2016.

13. Nocerino, E.; Menna, F.; Remondino, F. Accuracy of typical photogrammetric networks in cultural heritage 3d modeling projects. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *XL-5*, 465–472, doi:10.5194/isprsarchives-XL-5-465-2014.
14. Perfetti, L.; Polari, C.; Fassi, F. Fisheye photogrammetry: tests and methodologies for the survey of narrow spaces. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W3*, 573–580, doi:10.5194/isprs-archives-XLII-2-W3-573-2017.
15. Pierrot-Deseilligny, M.; Clery, I. APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2011**, *5/W16*.
16. Ray, S.F. *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video and Electronic Imaging*; Focal Press: Oxford, UK, 2002.
17. Remondino, F.; Rizzi, A.; Jimenez, B.; Agugiaro, G.; Baratti, G.; De Amicis, R. The Etruscans in 3D: From space to underground. *Geoinformatics FCE CTU* **2011**, *6*, 283–290.
18. Stamatopoulos, C.; Fraser, C.S. Automated Target-Free Network Orientation and Camera Calibration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *1*, 339–346.

Perfetti, L.; Polari, C.; Fassi, F.; *et al.* Fisheye Photogrammetry to Survey Narrow Spaces in Architecture and a Hypogea Environment. In *Latest Developments in Reality-Based 3D Surveying and Modelling*; Remondino, F., Georgopoulos, A., González-Aguilera, D., Agrafiotis, P., Eds.; MDPI: Basel, Switzerland, 2018; pp. 3–28.