### FISHEYE LENS IMAGE CAPTURE ANALYSIS FOR INDOOR 3D RECONSTRUCTION AND EVALUATION

H. A. León-Vega 1\*, M. I. Rodríguez-Laitón 2

<sup>1</sup> Cadastral and Geodesy Engineering, hotbed of research PENSANTE (Pensamiento Espacial y Territorial) Spatial and Territorial Thinking, Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia – haleonv@correo.udistrital.edu.co
<sup>2</sup> Cadastral and Geodesy Engineering, hotbed of research, GRSS-IEEE UD Group, Universidad Distrital Francisco José de Caldas, Bogotá D.C, Colombia – mirodriguezl@correo.udistrital.edu.co

#### **Commission II**

KEY WORDS: SfM Photogrammetry, 3D Reconstruction, Fisheye Lens, Indoor Spaces, Camera Calibration.

#### **ABSTRACT:**

The following article has as its purpose to solve a series of questions related to the acquisition of fisheye lens images made in the Benchmark FINE, making use of this data and images to generate a reconstruction to the tower of the chapel of San Vigilio in 3D using SFM photogrammetry and its application methodologies using low-cost instruments and sensors such as non-metric digital cameras. **The fisheye lens has a wide range of focus and field of view that makes it possible** to capture a scene with a limited number of images more **quickly and efficiently.** An analysis is intended to be carried out on the basis of the results obtained by assessing their accuracy and quality to determine the feasibility in the proposed initial use for the assessment of spaces difficult access by maintaining geometry without distances, scale, defined orientation in images, or ground control points (GPCs).

#### 1. INTRODUCTION

In recent years, sensors and cameras especially the low-cost cameras have been improving and developing a high degree of reliability, achieving greater performance at an affordable price. Recent advances in photogrammetry and computing have enabled the use and application of cameras for different purposes such as 3D reconstructions providing greater flexibility and efficiency at the time of choice of a sensor.

However, field data acquisition continues to be a major challenge in small, remote and difficult-to-reach sites with complicated lighting conditions thus generating high costs and portability of one scenario to others of these conventional photogrammetric equipment results in restricted use. The use of non-metric cameras equipped with fisheye lens facilitates the taking of photographs in small spaces that are predominant in the interior of cities, archaeological sites, among others (Sahin, 2016) . This type of fisheye lens has a wide range of focus and field of view that makes it possible to capture a scene with a limited number of images more efficiently. In this regard, camera improvements and methods for calibration of non-metric cameras have democratized access to photogrammetric modeling and the generation of final products (Westboy et al., 2012).

In this way, the fields of application for Photogrammetry are expanded to be more accessible and interdisciplinary, reducing costs and achieving high precision with high detail (Strecha et al., 2015). And for the case to work in this article opens the doors to research and application of the methodology to work in different areas of knowledge such as the valuation of cultural assets, which

have great historical value in societies. In this case, the availability of photogrammetric products with a Terrain Sample Distance (GSD) has a direct consequence on applicability across all necessary scale ranges ( Lambers et al., 2009) in indoor spaces especially.

#### 2. PHOTOGRAMETRIC STUDY METHODS

Since the emergence of short-range automated digital photogrammetry over the past decade, it has become a powerful and widely used tool for 3D optical modeling with image-based terrain extraction algorithms, which have radically improved the quality of terrain data, which can be derived from overlapping photographic pairs in addition to their methods for calibration of non-metric cameras (Schneider et al., 2009).

However, it differs fundamentally from conventional photogrammetry, in that scene geometry, positions and camera orientation are automatically resolved without the need to specify a priori, a network of lenses that have positions known three-three-three-day (José Luis et al., 2019). Instead, these are resolved simultaneously using a highly redundant iterative packet tuning procedure, based on a database of features automatically extracted from a set of multiple series of overlapping images (Snavely, 2011).

#### 2.1. SfM PHOTOGRAMMETRY

SFM (Structure from Motion) Photogrammetry is a low-cost photogrammetric method for high-resolution reconstruction for land, structures, monuments, among others, ideal for low-budget research. It works under the same principles of conventional

<sup>&</sup>lt;sup>1</sup> Author 1: <u>haleonv@correo.udistrital.edu.co</u>

photogrammetry where the 3D structure can be solved from a series of overlapping displaced images (Figure 1). However, scene geometry, positions, and camera orientation are resolved automatically without specifying ground control points (Caro, 2012). As a result, the camera positions lack scale and orientation provided by ground control, therefore, 3D point clouds are generated in a relative coordinate system that must be aligned with a coordinate system of the object in the real world. To do this, the transformation of the coordinates of the SfM image space into an absolute coordinate system can be achieved using a 3D similarity transform such as the Scale Invariant Feature Transform (SIFT) based on a small number of known known ground control points (GPCs) with object coordinates in reality identifying characteristics and obtaining their GPS (Barazzetti, et al., 2017).



Figure 1: Structure from Motion (SfM). Instead of a single stereo pair, the SFM technique requires multiple overlapping photos as input for feature extraction and 3D reconstruction algorithms.

#### **2.2. OBJECTIVES**

The objective of this work is to give a brief introduction to SfM photogrammetry and develop a methodology initially used in capturing photographs through the fisheye lens and processing them for the creation of a 3D model from the section of the lower dome located in the San Vigilio Chapel using the licensed software Pix4Dmapper Pro 2.0 analyzing the robustness and quality of the final texturization generated. The choice of this software is given by the processing capacity it has and for the short time in which it runs it comparatively with other free software's since the photographs to be worked on have a high resolution and size.

In addition, it is proposed to highlight two cases, where two patterns of geometry that make up the scenarios of the dome and the corridor of the staircase are predominantly performing an analysis on how it was carried out as the capture of the images and the results generated given the conditions of themselves, such as the homogeneity of the colors and the forms present, for finally evaluate the different processing such as the calibration of the camera performed in the software, cloud points, 3D mesh and the subsequent generation of textures that are adjusted to the physical and real conditions of the stilling, this being a reliable element in

<sup>3</sup> About Algorithms <u>https://support.pix4d.com/hc/en-</u>

conjunction with globally accepted precision standards, within those of the ASPRS (American Society of Photogrammetry and Remote Sensing).

#### 3. MATERIALS AND METHODOLOGY 3.1. SOFTWARE

All results were generated in the commercial Software Pix4Dmapper Pro 2.0 which implements a fully automated workflow for the different parts of the reconstruction from the initial processing, the calibration of the camera , the corresponding point cloud densification and the display of the final result of the 3D reconstruction with the corresponding textured mesh. It should be noted that since it is not open source software it does not reveal the precise nature of the algorithms used for the procedures described above. Machine Learning-based algorithms have been found<sup>3</sup>to be used in addition to the usual polygon modeling.

#### **3.2. DATA ACQUISITION**

For the development of the reconstruction, the reference data provided by the special session organized by the Benchmark FINE was acquired, which collected all the information from the tunnels of the Castle of San Vigilio found in the northern part of the Upper city of Bergamo in Italy and being this very important<sup>4</sup> in the city for its heritage and historical character.

The starting point of SfM Photogrammetry is the determination of the 3D location of the homologous characteristics in each image taken from different angles. To do this, the main characteristics of each image that are susceptible to correspondence through the already implemented Object Recognition System of Invariant Scale Features Transformation (SIFT) are initially identified. This transformation identifies key features or points in each image that are invariant in scale, rotation, and lighting conditions (Lowe, 2004).

#### 3.3. WORKFLOW AND SCENARIOS

For this article, we will work on the reconstruction of the lower dome of the Castle and the stairs that connect to the top of the structure. For the section of the corridor images were taken laterally pointing towards the walls (Figure 3) and front pointing towards the background of the corridor (Figure 4) In addition to the images of the two cupulas structure, including its exterior part. The aim is to answer the question included in the **Benchmark** of the event: **Are the number of images taken from different angles sufficient or redundant with the fisheye lens to achieve accurate and complete results ?** Performing a reconstruction of the corridor only with the side images and another with all the images captured from the area, analyzing the quality of the results yielded in the different procedures.

From the images it is evidenced that the fisheye lens with its wideangle lens generously encompasses the elements of the runner captured frontally and laterally accompanied by a great definition

<sup>4</sup> The FINE Benchmark <u>http://www.3d-arch.org/</u>

us/community/posts/204450483-About-algorithms

in the details product of the digital camera used for this data acquisition. For the inner dome (Figure 5) a good capture of images is shown showing in the foreground the geometric shape of the roof and the homogeneity of the colors, among other elements analyzed at detail later. The methodology proposed is illustrated below:

The methodology is a description of the steps carried out for the development of reconstruction and planning (Figure2). Integrates from the handling that was given to the data acquired in the Benchmark, to the information that is had as the Pix4D software carries out the generation of the cloud of points, mesh of triangles, and generation of textures which are the processes that They can be seen at first sight. This methodology can be used for later work as the first approach to the generation of reconstructions.



Figure 2: Workflow performed in Pix4d Rebuild software



Figure 3: Side capture of the wall in the corridor of the Castle

From the capture of the images such as those shown in figures 3, 4 and 5, a representation can be made where, taking into account (García, et al., 2011) the complex geometry that historical buildings usually present, it is necessary the correct planning of preliminary capture networks or routes which should allow obtaining enough points to define the basic structure of canvases, walls, vaults, etc.



Figure 4: Front capture of the runner at the Castle.



Figure 5: Front capture of the runner at the Castle.

## 3.4. RECONSTRUCTION OF THREE-DIMENSIONAL SCENES

For the execution of the workflow described above was developed from 3 proposed cases as follows:

IMAGE CAPTURE CASE	NUMBER OF IMAGES
1: Side Ladders	53
2: Side and Front Stairs	83
3: Dome around	25

Table 1: Description of the cases to be evaluated according to the reconstruction:

In addition, a reconstruction with all the photos. With an overlay (about 80% per image) presented between photos. This extraction was done automatically by establishing homologous points in each image and features that stand out that do not vary in scale such as colors, textures and other predominant elements in the photographs.

Obtaining a virtual model of buildings with real photographic textures is a product that provides enormous possibilities for any professional who may require representations of all or part of it.

#### 4. RESULTS

#### 4.1. CAMERA CALIBRATION ANALYSIS

From the processes previously calculated it is visualized that in the acquisition of pass points for a calibrated image have a greater number of points with a very successful distribution given the lighting conditions of the structure and the homogeneity of the colors present in the materials with which they were built at the time (Figure 6).



Figure 6: Capture the dome of the castle with the fisheye lens and its points of passage after modeling the software with the calibrated image.

However, in the images resulting from the uncalibrated camera (Figure 7), the acquisition of points from these was neither accurate nor well distributed with respect to the final reconstruction obtained since in these there are no points for their relationship and therefore cannot continue with the workflow proposed in Figure 2.



Figure 7: Capture a castle aisle with the fisheye lens and its passing points in an uncalibrated image.

#### 4.2. RECONSTUCTION ANALYSIS. CASE 1:

For this scenario, where the photographs were taken pointing to the walls (Figure 3) it is evidence that these walls have a good reconstruction around their detail and homogeneity, it must be borne in mind that being a flat surface, with homogeneous texture and colors and the overlap of each image is not less than 50% (Kumar et al., 2016), it facilitates and substantially improves the result as shown continued. However, the stairs did not have a good reconstruction since when handling a fisheye lens at the ends there is an aberration and it was not directly pointed at them. One problem that was evident in this methodology was the elements that present heights, in this case was the fence or handrail and the respective height comparatively with the wall that cannot be shown

with the naked eye. This denotes that more views are required relative to these objects to capture their depth and thus generate a good 3D reconstruction on that component, in addition to taking into account at the time of capture the height above the level of reference to consider for a better perspective and a more robust end result.



Figure 8: Result of 3D reconstruction for case 1 presenting a good texture and shape according to reality.



Figure 9: Result of 3D reconstruction in the handrail present, there are problems in given the lack of views in height and depth.

Below are the main results of the reconstruction developed result from reconstruction made on (Figure 8 & 9).

Description	Value
Medium error [pixels]	0.166893
Point density per m3	41.23
Number of calibrated images	53 out of 53

Table 2: Results values in case reconstruction case 1

#### 4.3. RECONSTUCTION ANALYSIS. CASE 2:

For this case there is an improvement in the reconstruction of the stairs since the images were captured front (Figure 4 and 5) but inconsistencies are presented in the colors of the walls. Stripes are presented given the overlap of the images and the lighting conditions present at the time of capture during the tour.



Figure 10: 3D Reconstruction result for case 2. Inconsistencies in the texture and colors present are evident.



Figure 11: Result of 3D reconstruction for scenario 2 in the fence, presenting similar inconsistencies in case 1.



Figure 12: 3D reconstruction result for the lower section of the stairs. Consistency occurs in the shape and edges of the stairs.

The stairs presented a good reconstruction at the end and in general the corridor is shown with consistent dimensions in addition to shape and color. (Figure 10 & 11). This means that the fisheye lens on its front shows no further distortion of the captured elements and contifys its captrua frontally to improve reconstruction in the processing of the final textured mesh. As in case 1,the fisheye lens has no problems in capturing in greater detail the elements that have depth or height although it makes a good approximation in terms of

distribution, color and shape, however taking into account analysis of the previous case presents very similar results for the fence that is anchored to the wall (Figure 12).

Description	Value
Medium error [pixels]	0.159326
Point density per m3	32.44
Number of calibrated images	83 out of 83

Table 3: Results values in case reconstruction case 2

#### 4.4. RECONSTUCTION ANALYSIS. CASE 3:

For the case it must be borne in mind that this has a defined geometric shape, so the results for this scenario are very consistent in all its aspects, in addition to the correlation plot thrown in the processing report in the Software is quite good in shape, frequency and overlay as below screenshot shown:



Figure 13: Result of the 3D reconstruction for the section of the dome at the entrance of the corridor analyzed above, a very favorable result is presented for evaluation of shape, colors and texture.



Figure 14: Result of 3D reconstruction for the top of the dome and its top

Description	Value
Medium error [pixels]	0.159326
Point density per m3	32.44
Number of calibrated images	83 out of 83

Table 4: Results values in case reconstruction case 3



Figures 15: Result of 3D reconstruction for the outside of the dome.

Reconstruction from the fisheye lens in this scenario presented the most favorable results (Figure 13 & 14). For the exterior display, the software performs a visualization of the interior elements on the surface of the dome, for orientation effects and highlights the level of detail presented by the final reconstruction on the inside as it is shown above, complementing a reconstruction to the outer facade of the dome (Figures 15).

#### 5. ADDITIONAL ANALYSIS

In addition to the analyses of the previous cases the reconstruction of the corridor was carried out taking only the front photographs, it was not included within the direct analysis since the results are very poor as shown below:



Figure 16: Result of 3D reconstruction of stairs using only frontcaptured photographs.

The textures presented on the walls are not soft and little in line with reality compared to cases 1 and 2 described. At first impression it is observed that the fence presents a good reconstruction (Figure 16), but that is given the view that the capture of the image was taken, however this is not how it is shown below (Figure 17)

An important aspect to evaluate is the distribution in the correlations presented in the reports of the reconstructions in the three cases, where in case 1 there is a greater correlation (darker sections) in a heterogeneous manner. For case 2, the correlations are correct but as discussed above, the lateral sections have a low quality. Finally, in the section of the dome (case 3) a strong coincidence is shown during the processing since the physical

characteristics of the structure allow a good reconstruction and a successful result according to reality.



Table 5: Correlation diagrams presented in each of the 3 reconstructions of each case



Figure 17: Result of 3D reconstruction from another point of view. Inconsistencies occur, but a substantial improvement in the shape of the fence.

It could be observed that in some areas there is no information and in other redundancies of it (Figure 19). This is due to the route taken by the staff at the time of capture. So in order not to fall into this error it is recommended to make a preliminary check of the study area and make a preliminary tour for the taking of the photographs, thus obtaining a good overlap between each photograph and a subsequent extraction of homogeneous information, which will be reflected in the reconstruction of the model.

As can be seen the reconstruction was better than in the previous cases, but does not meet the requirements in terms of quality. However, this view contributes more to a more successful reconstruction of elements that are deeply (Figure 18). This shows that more views should be taken for this type of element for future reconstructions in other spaces under this methodology. As in this case the upper pipe presented a good reconstruction given the point of view to which it is located, reinforcing one of the conclusions of this article on the need for more robust views to generate a satisfactory reconstruction (Figure 19).



Figure 18: 3D reconstruction result for the section of the top pipe. An appropriate result is observed.

However, for the evaluation of elements in enclosed spaces such as cave tunnels or mines, for example, it is an appropriate methodology to observe the details thanks to its great resolution present in the photographs, in addition to this great advantage it is proposed to articulate this methodology are GCP ground control points to perform some kind of measurement or calculation of areas or volumes, however it should be noted that the fisheye lens has a degree of aberration in the contours of the lens so that for wide spaces not it is so versaltil comparatively with scred spaces which is perfect where the maneuverability is reduced and only certain shots can be made from a few angles (Table 4).

3D modeling and photogrammetry are a tool that has come to play an important role in the generation of in a virtual product used by many places in order to document the goods and narrow spaces.



# Figure 19: 3D reconstruction result for the fence with front photographs, a coating result that the software tried to make is displayed.

Finally it shows for these elements that have a different height superior to the surface of the wall, above the bottom the software tries to fill the gaps given the absence of views, a factor that can be very useful in spaces where you don't need a high degree of detail, but a shape commensurate with reality while maintaining elements like scale and consistency in point cloud generation.

#### 6. DISCUSSION

Given the results of reconstruction in the different scenarios, It can be seen that the fisheye has a very wide versatility for capturing the most elements with a few shots, given the characteristics of the lens, however for elements that have some kind of height or depth to reference elements such as the bard or handrail among others, is not enough with a single point of view, are required between 1 or two viewpoints that cover all the characteristics of the same.

As an innovative aspect of this article is highlighted its low cost and the quality with which the photographs were obtained using the fisheye lens, result that is reflected in the 3D reconstruction of each of the spaces in the tower, since few studies have been carried out using this type of cameras and sensors democratizing access to this type of studies and analysis in addition that this type of studies and especially reconstructions had not previously been carried out in the castle or tower, so the documentation of this architectural building inconvenient. As part of the disadvantages of reconstructions from fisheye lenses, it can be seen that it is very practical for flat surfaces, but those that have elements with greater height is a problem that needs to be addressed by taking photographs from greater points of view, adding more time and work.

In addition, this type of reconstructions in indoor spaces have not been carried out by developing a methodology such as that described in this work, taking into account the degree of detail that was taken at the time of capture trying to address the greatest number of elements found in this structure. According to (Barazzetti, 2017) and his work done comparatively with the methodology worked on in this article and the results obtained, it is possible to demonstrate the potential of fisheye lenses for accurate 3D modeling. On this other problem have to be taken into consideration among those the importance of the Calibration of the camera. The cameras must be calibrated. Beforehand using an image block with adequate geometry, the network deformations for incorrect camera calibration can be much larger than typical deformations in an image. In addition, acute textures require limiting the wide field of view of fisheye lenses, which could be a difficulty as in Reconstructions based on fisheye images (Table 1, 2 & 3).

For the case of the dome it is observed that the path of taking the photographs is quite correct given that the pixel error presented in Table 2, 3 and 4 is very low, giving as such a good result. All this has to do with the high degree of correlation between the overlay of the images, this is partly thanks to the calibration method of the camera when detecting the software which is considered a fundamental step in the development of the reconstruction. Finally it is worth reviewing all additional functionalities such as measuring area calculation among others that can be developed with the final result as long as you get control points from the area for these respective calculations.

In addition to this great advantage it is proposed to articulate this methodology with GCP ground control points to perform some type of measurement or calculation of areas or volumes. However, there were places where it passed very quickly and few captures were made in the tower generating discrepancies and a low correlation in the reconstruction, but given the degree of size handled in the capture is negligible these cases.

Finally, the analysis that was carried out on the processing carried out by the software internally, taking into account the great performance that is required to make a reconstruction of this scale, although as a later work the comparison between different software's (commercial and free) to compare the results in the point cloud and the textured mesh since each of them implements a different type of algorithm based on the Transformed SIFT and in this case Pix4D is handled under the beginning of the box where it is not possible to know the algorithms and processing used in the generation of each part of the reconstruction accurately.

#### 7. CONCLUSIONS

This metodology and product, have a consistent result, this can be used both for heritage documentation, as virtual musealization, for architectural rehabilitation or restoration projects, archaeological studies of walls, in the same way that many more authors have reached a similar conjecture. 3D modeling is an effective way to respond to different studies, interdisciplinary that allows applications in various fields of research, for this case the use of non-metric cameras allows the capture of data to different scales as a local or to an object in a timely manner, where its elements such as texture, color, shape and dimensions are modeled according to reality responding to the questions initially raised in the Benchmark Within the framework of the objectives of this article concisely.

#### 8. ACKNOWLEDGMENTS

To Professor Erika Upegui PhD, for the acquisition of Benchmark data and their respective orientation in the development of this work.

#### 9. **BIBLIOGRAPHY**

Barazzetti, L., Previtali, M., & Roncoroni, F. (2017). FISHEYE LENSES FOR 3D MODELING: EVALUATIONS AND CONSIDERATIONS. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W3, 79-84

Caro, J. L. (2012). Fotogrametría y modelado 3D: Un caso práctico para la difusión del patrimonio y su promoción turística. Universidad de Malaga.

Caro J., Salvador H., 2015: De la fotogrametría a la difusión del patrimonio arqueológico mediante game engines: Menga un caso de estudio. Virtual Archaelogy Review, 58-68.

García, J. L. P., Calvache, A. T. M., Escarcena, F. J. C., & Arenas, A. L. (2011). Fotogrametría de bajo coste para la modelización de edificios históricos. Virtual Archaeology Review, 2(3), 121-125.

José Luis, P.-G., Antonio Tomás, M.-C., Vicente, B.-C., & Alejandro, J.-S. (2019). Photogrammetric studies of inaccessible sites in archaeology: Case study of burial chambers in Qubbet el-Hawa (Aswan, Egypt). Journal of Archaeological Sciences, 102.

Kumar, R., Muknahallipatna, S., & McInroy, J. (2016). An Approach to Parallelization of SIFT Algorithm on GPUs for Real-Time Applications. Journal of Computer and Communications, 04, 18-50. https://doi.org/10.4236/jcc.2016.417002

Lambers, K., & Remondino, F. (2009). Optical 3D Measurement Techniques in Archaeology: Recent Developments and Applications. First publ. in: Layers of perception. Proceedings of the 35th International conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2-6, 2007 / Ed. by Axel Posluschny ... Bonn : Habelt, 2008, pp. 27-35.

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2), 91-110.

Sahin, C. (2016). Comparison and Calibration of Mobile Phone Fisheye Lens and Regular Fisheye Lens via Equidistant Model. Journal of Sensors, 2016, 1-11.

Schneider, D., Schwalbe, E., & Maas, H.-G. (2009). Validation of geometric models for fisheye lenses. ISPRS Journal of Photogrammetry and Remote Sensing, 64(3), 259-266.

Strecha, C., Zoller, R., Rutishauser, S., Brot, B., Zapp, K. C., Chovancova, V., ... Glassey, L. M. (2015). Terrestrial 3d Mapping Using Fisheye and Perspective Sensors.

Snavely, N. (2011). Scene Reconstruction and Visualization from Internet Photo Collections: A Survey. IPSJ Transactions on Computer Vision and Applications, 3, 44-66.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314.