

Article

Metric Rectification of Spherical Images

Luigi Barazzetti 

Department of Architecture, Built Environment and Construction Engineering (ABCE), Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy; luigi.barazzetti@polimi.it; Tel.: +39-022399-8779

Abstract: This paper describes a method for metric recording based on spherical images, which are rectified to document planar surfaces. The proposed method is a multistep workflow in which multiple rectilinear images are (i) extracted from a single spherical projection and (ii) used to recover metric properties. The workflow is suitable for documenting buildings with small and narrow rooms, i.e., documentation projects where the acquisition of 360 images is faster than the traditional acquisition of several photographs. Two different rectification procedures were integrated into the current implementation: (i) an analytical method based on control points and (ii) a geometric procedure based on two sets of parallel lines. Constraints based on line parallelism can be coupled with the focal length of the rectified image to estimate the rectifying transformation. The calculation of the focal length does not require specific calibrations projects. It can be derived from the spherical image used during the documentation project, obtaining a rectified image with just an overall scale ambiguity. Examples and accuracy evaluation are illustrated and discussed to show the pros and cons of the proposed method.

Keywords: 360° image; digital recording; homography; rectification; spherical cameras



Citation: Barazzetti, L. Metric Rectification of Spherical Images. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 248. <https://doi.org/10.3390/ijgi11040248>

Academic Editor: Wolfgang Kainz

Received: 12 February 2022

Accepted: 8 April 2022

Published: 11 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Digital metric documentation with images and laser scans is fundamental in projects requiring accurate metric deliverables. Traditional metric deliverables are plans, sections, orthophotos, digital maps, or different types of 3D models such as BIM, surface or solid models, surface meshes, and finite elements models, among others [1]. Nowadays, the use of digital photogrammetry and laser scanning is one of the most common digital surveying methods able to generate accurate geometric records in different formats.

Several papers and textbooks describe both theoretical concepts and practical issues in digital documentation projects of three-dimensional objects carried out using images [2] and laser scans [3]. The reader is referred to [4–8] for some examples related to digital workflows using a combination of instruments and measuring techniques.

This paper aims to show a relatively fast alternative solution that can be used with 360° low-cost cameras available on the commercial market. Low-cost sensors are becoming more popular within the framework of digital photogrammetry because a digital camera can be turned into a measuring tool [9], obtaining 3D models from blocks and sequences of digital images.

The proposed method is not based on traditional frame (also called central perspectives or pinhole) cameras. Here, 360 images (also called spherical or equirectangular projections) are used, and the considered objects are planar surfaces, i.e., two-dimensional objects.

Metric documentation of 2D objects is typical in several projects, such as the digital documentation of building facades, mosaics, and paintings, among others. Here, rectified photographs are still a popular solution in several practical applications.

The spherical camera model is already available in commercial software for image-based 3D modeling, such as Agisoft Metashape or Pix4Dmapper. However, the reconstruction workflow is mainly developed to deal with three-dimensional geometry objects. In fact, commercial software requires multiple spherical images acquired from different

locations, which are then processed to generate 3D textured models or additional outputs (orthophotos, digital elevation models, etc.).

Different authors described the use of spherical photogrammetry for 3D modeling, which is a powerful tool to document small and narrow spaces [10–19]. Aspects related to camera calibration are described in [20]. Accuracy evaluation are reported in [21–26]. In [27], a set of spherical images is used to merge photogrammetric projects carried out with other camera models.

Other examples, including both metric and non-metric projects with spherical images, were discussed by different authors. For instance, heritage documentation is a field in which spherical images allow rapid mapping of buildings and monuments [28–31] including sites at risk [32]. In [33], spherical images are used as a base for virtual reality applications, whereas an immersive tool is described by [34]. The use of such images and deep learning is discussed in [35]. The work illustrated in [36] uses a low-cost 360° sensor for crime scene documentation. Ref. [37] uses spherical images for biomass estimation in forestry. Integrated use of drones and spherical images is presented in [38].

Spherical images are also available on the web. Different sharing services (e.g., 360Cities.net, Mapillary, Facebook, Kuula, Roundme, and Theta360) offer the opportunity to upload this type of image. Some of them allow data sharing and the creation of virtual tours.

The method presented in this manuscript differs from traditional projects based on spherical photogrammetry. As mentioned, only planar objects are considered. Although 2D metric documentation can be carried out with the traditional photogrammetric workflow (planar objects are sub-cases of 3D projects), there is a lack of software for metric rectification of spherical images.

Metric rectification [39] is a well-known digital recording solution that can be carried out using just a photograph, removing the perspective deformation using external constraints such as control points or information about the geometry (e.g., sets of parallel lines and known aspect ratio). Digital orthophotos created with 3D photogrammetry can compensate for the lack of flat geometry. However, this would require a photogrammetric project with several images featuring good overlap.

This paper tries to avoid the multi-step processing workflow (image matching, bundle adjustment, generation of a dense point cloud, extraction of a mesh or a DSM, and orthophoto production) available in commercial software. The idea is to use a single spherical image and a direct metric rectification approach.

In the case of metric rectification projects carried out with images based on the pinhole camera model, lens distortion is removed beforehand using distortion parameters derived from a specific calibration project. Different commercial packages offer predefined calibration coefficient sets for several camera bodies and lens configurations.

Images acquired with a spherical camera (also called 360° or equirectangular projections) feature a 360° × 180° field of view, capturing the entire scene around the camera. Low-cost and professional cameras are available on the commercial market and allow the photographer to capture a spherical image, usually generated from a set of pinhole or fisheye lenses mounted on the 360° camera. Although the acquired data are not 360° images, stitching is automatically carried out to provide the final image in real-time.

This paper extends what was briefly proposed by [40], in which different procedures for condition mapping in restoration projects were illustrated and discussed. This paper focuses on the concept of metric rectification from 360° images, describing the algorithms with more details and some additional metric evaluation experiments in order to provide information about the achievable metric accuracy.

The paper is structured as follows. Section 2 illustrates the general workflow of the proposed method for metric rectification of spherical images. The following Sections 3–5 describe the different steps of the implemented solution. Experiments were carried out with low-cost cameras available on the commercial market at a cost of about EUR 500.

2. The Proposed Method for Metric Rectification of Spherical Images

The procedure for metric rectification is based on a multi-step workflow in which different products are generated from a single spherical image. The workflow starts with acquiring a spherical image, which can be captured with low-cost or professional 360° cameras, or using a rotating camera and stitching software capable of mosaicking multiple shots [41,42]. In the case of a rotating camera, the rotation point must be perspective center to avoid parallax errors during the stitching phase.

Figure 1 illustrates the flowchart of the proposed workflow for the extraction of the metric images.

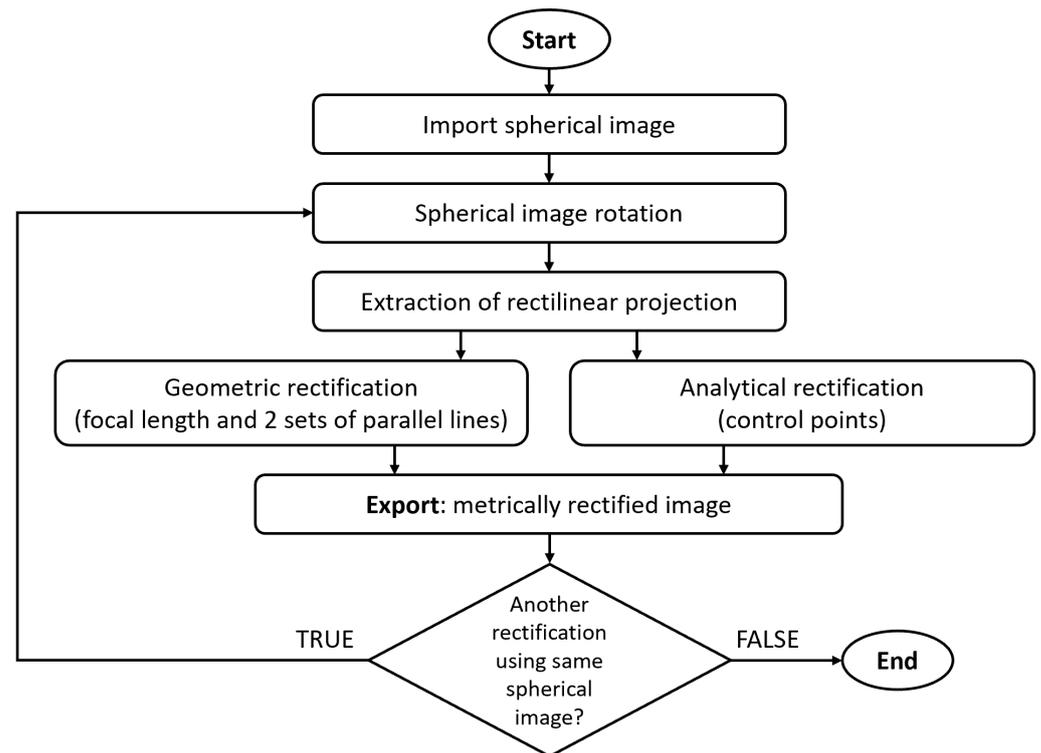


Figure 1. The developed workflow for metric rectification of several images from a single spherical projection.

An example is proposed here to clarify the implemented solution and the different outputs that can be produced. Figure 2 shows the results achievable with the proposed metric rectification approach using a spherical image retrieved from Google Maps. The image was acquired in a room of the Alcázar of Seville (Spain). Image resolution is $13,312 \times 6656$ pixels and the covered field of view is $360^\circ \times 180^\circ$. The following sections will prove that a complete $360^\circ \times 180^\circ$ (i.e., no cropping to limit the field of view) provides all required information to calculate internal orientation parameters, namely the center of the sphere in the camera reference system and the focal length (in pixels) f .

The spherical image generated six metrically rectified images: four images for the lateral walls, ceiling, and floor. As can be seen in Figures 3 and 4, a single spherical image acquired inside a room allows the user to record all the different flat surfaces. It also provides an immersive visualization that allows the user to understand the relative position of the different walls. A recording specialist could avoid taking notes in the case of a more traditional rectification project based on standard frame images.

Recovering metric properties means generating a new (metrically) rectified image that provides angles and ratios of distances. The correct scale can be recovered only if some (metric) information is available, such as a known distance measured with a measuring tape or a set of control points collected with a total station. In other words, if there is no

external information about the considered surface, the rectification process is affected by an overall scale ambiguity, which is another advantage of the proposed approach that encapsulates the focal length. A more traditional geometric rectification would require instead a known width–height ratio, i.e., two distances.



Figure 2. A spherical image acquired in the Alcázar of Seville (Spain).

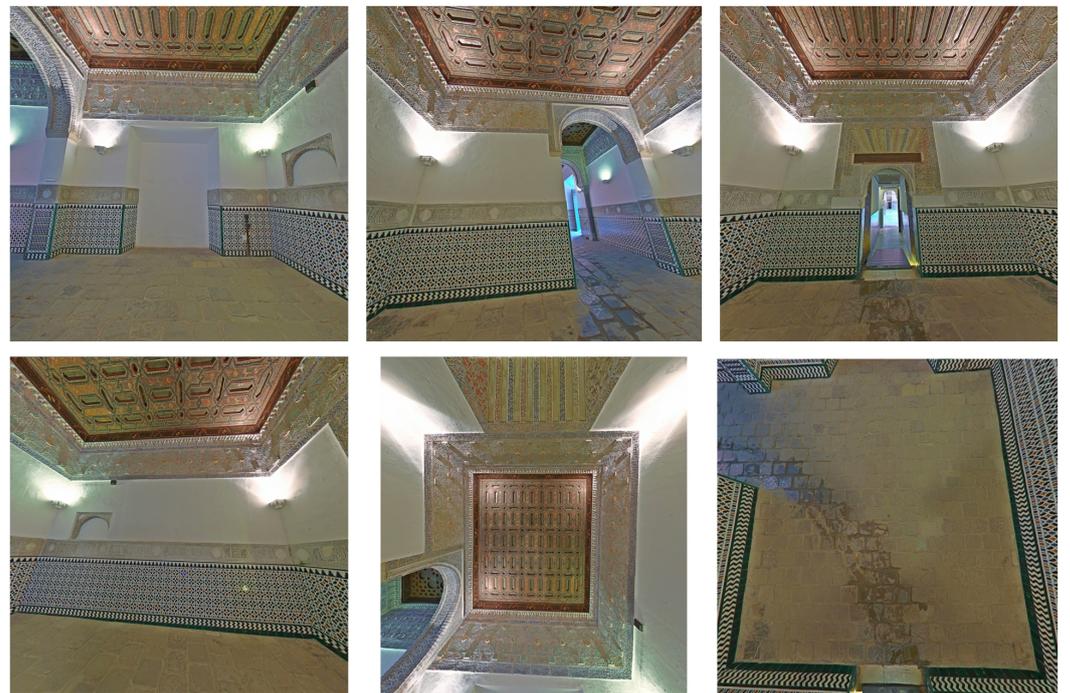


Figure 3. Six rectilinear images generated from a single spherical image using a gnomonic projection coupled with rotations.

A spherical image can be synthetically rotated to change the viewing direction towards specific areas. The user is probably familiar with the traditional “bubble” visualization, in which the viewing direction can be interactively changed by dragging the point of view (POV). The idea is here replicated considering the planar surface requiring rectification.

We can define the following three angles for a leveled 360° image:

- h = heading, rotation around the vertical direction;
- r = roll, rotation around the line of sight;

- p = pitch, rotation around the transverse axis.



Figure 4. The six metrically rectified images generated for the planar surfaces.

Changing the heading h without altering pitch and roll ($r = 0^\circ, p = 0^\circ$) provides the same effect of a user looking left or right. This is the most common way to point the viewing directions towards the center of the vertical wall to be rectified. A variation of the pitch instead would rotate the camera up or down, allowing the user to also capture the floor and ceiling.

The user must be aware that the order of rotations (in this case $h \rightarrow r \rightarrow p$) is fundamental. A positive pitch variation corresponds to point the POV up if $h = 0^\circ$. The opposite, i.e., POV pointing down, will be achieved for $h = 180^\circ$. Roll variations are helpful in refining the captured area, mainly when the spherical image was not acquired in the center of the room.

The image used to rectify the first vertical wall (top-left in Figure 3 was extracted by setting the values $\text{POV} = (h, r, p) = (0^\circ, -10^\circ, 0^\circ)$. The extraction is carried out using a gnomonic projection, which allows projecting the points of the spherical image onto a plane tangent to the sphere. The projection center is the center of the sphere.

The user has to define the field of view (FOV) of the new image based on the pinhole camera model (also called central perspective or rectilinear projection). The FOV can be interactively enlarged or reduced to capture the entire wall, and it was set to $100^\circ \times 100^\circ$.

The second image (top-middle) was set to $\text{POV} = (85^\circ, 0^\circ, -20^\circ)$ and $\text{FOV} = 110^\circ \times 110^\circ$ to capture the wall on the right. The third wall (top-right) has parameters $\text{POV} = (180^\circ, 15^\circ, 0^\circ)$ and $\text{FOV} = 120^\circ \times 120^\circ$. Finally, the fourth wall (bottom-left) has $\text{POV} = (-85^\circ, 0^\circ, 15^\circ)$ and $\text{FOV} = 120^\circ \times 120^\circ$. As can be seen, FOV values were changed to cover the entire walls, which have different sizes.

The rectilinear image for the ceiling (bottom-middle) has parameters $\text{POV} = (90^\circ, -90^\circ, -90^\circ)$ and $\text{FOV} = 120^\circ \times 125^\circ$. The floor is the flat surface requiring the largest field of view because the camera is relatively closed to the subject. Parameters were set as $\text{POV} = (0^\circ, 85^\circ, -2^\circ)$ and $\text{FOV} = 135^\circ \times 135^\circ$. More details about the extraction of the rectilinear images are discussed in Section 3. The implementation used relies on the Panorama Tool Files, which is the library used to develop Hugin software. Files can be downloaded from <https://sourceforge.net/projects/panotools/files/> (last accessed on 10 March 2022).

The last step is metric rectification, which is carried out on the different rectilinear images. Two different approaches are illustrated and discussed in the manuscript. The first method is the typical solution for accurate metric rectification projects to record flat surfaces. Control points are measured with an external instrument, usually a total station. The same points are measured in the rectilinear images, obtaining pixel coordinates and calculating the parameters of the homographic transformation. The details of such an approach and an accuracy evaluation are discussed in Section 4.

The solution used to generate the metrically rectified images relies on images with no external information. Section 5 will show that a homography able to rectify a rectilinear image can be estimated from the vanishing line of the image and the focal length of the camera used.

Flat surfaces with at least two sets of parallel lines allow calculating the vanishing line. In contrast, the focal length of the rectilinear image can be derived from the pixel size of the spherical image.

The six surfaces of the considered room were therefore rectified without taking extra on-site measurements, obtaining six images with an overall scale ambiguity (Figure 4). As surfaces share common discontinuity lines, they can be scaled to reduce the number of known distances to just 1 for the entire room.

The bottom-left facade in Figure 4 is a clear example of the possibility to compensate for just deformations on the chosen rectification plane. The wall has an internal niche that cannot be rectified using the proposed approach, which only operates on the chosen object plane.

3. Extraction of Rectilinear Images from a Spherical Projection

This section describes the procedure to extract a rectilinear image from a spherical projection, which can be used during the metric rectification step.

We define a spherical (equirectangular) projection as an $L_x \times L_y$ image obtained by mapping latitude and longitude (spherical) coordinates (λ, ϕ) onto the (image) plane (u, v) using the following Equations (Figure 5):

$$u = r\lambda \quad ; \quad v = r\left(\frac{\pi}{2} - \phi\right) \quad (1)$$

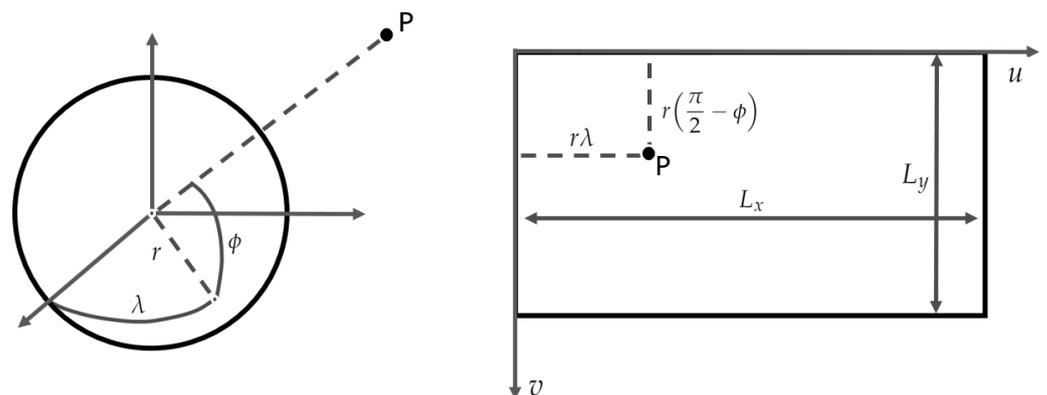


Figure 5. Conversion from spherical coordinates (λ, ϕ) to pixel coordinates (u, v) based on an equirectangular projection.

Equirectangular projections cover 360° horizontally and 180° vertically so that the aspect ratio is 2:1 ($L_x = 2L_y$). Let us consider a sphere with center $(0, 0, 0)$ and radius $r = L_x / (2\pi)$. The plane tangent to the sphere at point $(r, 0, 0)$ represents the new rectilinear image. The point has latitude and longitude $(\lambda, \phi) = (0, 0)$ and lies on the equator. Mapping is based on a gnomonic projection, i.e., from the center of a sphere to a plane tangential to the sphere.

Horizontal $\Delta\lambda$ and vertical $\Delta\phi$ angles define the field of view and the size (in pixels) $w \times h$ of the new rectilinear image:

$$\begin{aligned} w &= 2r \tan \frac{\Delta\lambda}{2} \\ h &= 2r \tan \frac{\Delta\phi}{2} \end{aligned} \quad (2)$$

Vertical and horizontal field of views (FOVs) can be estimated as $(2\Delta\lambda, 2\Delta\phi)$.

An essential property of the rectilinear image is the relationship between its focal length and the sphere's radius. Indeed, the focal length of the rectilinear image is equal to r if the mapping is carried out, preserving the maximum image resolution. Moreover, no barrel or cushion distortion needs to be corrected, and the rectilinear image can be considered a novel distortion-free image.

Figure 6 shows the extracted rectilinear images changing the field of view, the values of the angles were chosen as $2\Delta\lambda = 2\Delta\phi = 15^\circ$, progressively increasing resolution: 30° , 50° , 80° , 110° , and 150° . It is recommended to keep the angles under 110° – 120° , to reduce deformations at the edges. Very large fields of view result in stretched elements close to image edges, making the resulting rectilinear projection unusable for metric application.

The resolution of the final image is also proportional to the chosen FOV because the original level of detail encapsulated in the spherical projection is preserved in the current implementation. This means that each image can be considered as a part extracted from the next one. It is up to the user to find a specific balance between the FOV and the area to be rectified.

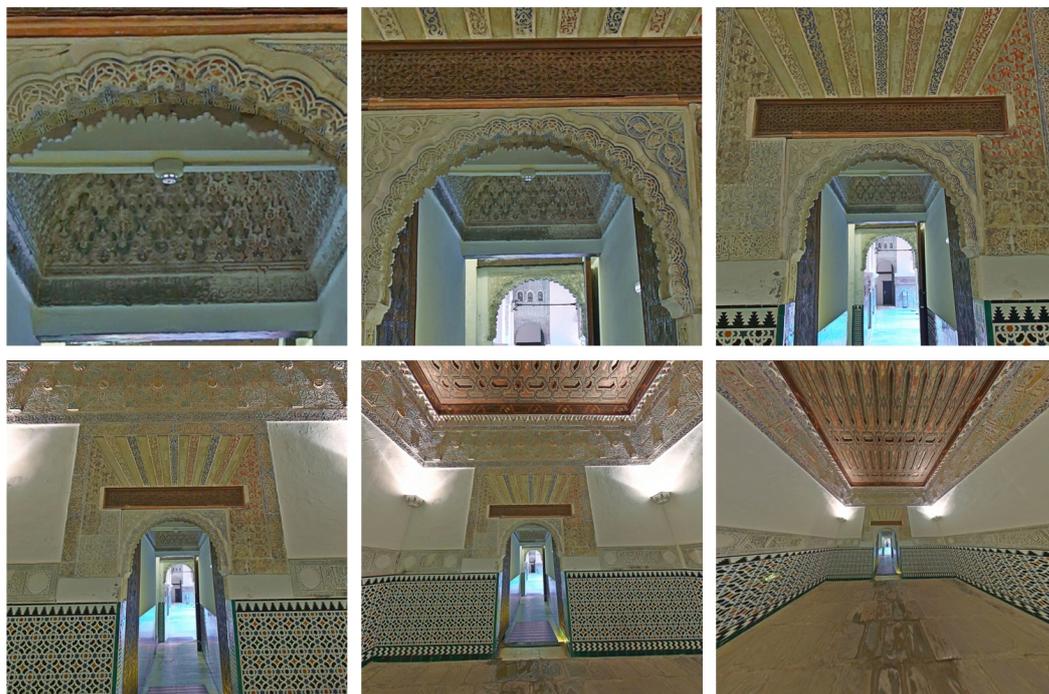


Figure 6. The same rectilinear projection extracted from a spherical image. The figure shows the effect of a variable FOV.

The extraction of multiple rectilinear images is based on selecting alternative tangent planes to the sphere at a generic point (λ, ϕ) . Moving the plane extends the previous case and projection equations must be modified accordingly. However, the alternative solution implemented in this work is the synthetic rotation of the original equirectangular projection using heading, roll, and pitch in sequential order. Changing the point of view allows to orient the sphere so that the tangency point becomes $(\lambda, \phi) = (0, 0)$, without requiring a modification of the current implementation.

4. Spherical Image Rectification with Control Points

Metric rectification using control points is the typical solution for the accurate digital documentation of planar surfaces. Control points are usually measured with instruments able to provide object coordinates in a metric system. The typical case is based on control points measured with a reflector-less total station. Then, the user must select the corresponding image points and calculate transformation parameters from image to object space. The next section introduces a possible solution for the estimation problem.

4.1. Homography Estimation Using Image-to-Object Correspondences

Homography is the transformation between a planar object and the corresponding rectilinear image extracted from the spherical projection. Parameter estimation can be carried out using projective geometry.

A point in the Euclidean 2-space has inhomogeneous coordinates $\mathbf{x} = (x, y)^T$. Adding an extra coordinate to the pair provides a new triplet $\mathbf{x} = (\lambda x, \lambda y, \lambda)^T$. We say that this 3-vector is the same point in homogeneous coordinates (for any non-zero λ). A homogeneous vector $\mathbf{x} = (x_1, x_2, x_3)^T$ represents the point $\mathbf{x} = (\frac{x_1}{x_3}, \frac{x_2}{x_3})^T$ in \mathbb{R}^2 .

A planar homography (also called projective transformation) is represented by a 3×3 matrix \mathbf{H} with 8 degrees of freedom:

$$\mathbf{x}' = \begin{bmatrix} X'_1 \\ X'_2 \\ X'_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{H}\mathbf{x} \quad (3)$$

which can be cast in the form:

$$\begin{aligned} X' &= \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \\ Y' &= \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \end{aligned} \quad (4)$$

We may multiply the numerator by the denominator to obtain two linear equations:

$$\begin{aligned} X'(h_{31}x + h_{32}y + h_{33}) &= h_{11}x + h_{12}y + h_{13} \\ Y'(h_{31}x + h_{32}y + h_{33}) &= h_{21}x + h_{22}y + h_{23} \end{aligned} \quad (5)$$

The values of parameters h_{ij} can be determined from $n \geq 4$ corresponding points $(X', Y') \leftrightarrow (x, y)$. The last element of \mathbf{H} can be set $h_9 = 1$ to take into consideration scale ambiguity.

A system of equations can be written as:

$$\begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -X'x & -X'y \\ 0 & 0 & 0 & x & y & 1 & -Y'x & -Y'y \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ \vdots \end{bmatrix} = \begin{bmatrix} x' \\ Y' \\ \vdots \end{bmatrix} \quad (6)$$

or with the more compact notation $\mathbf{A}\mathbf{h} = \mathbf{b}$.

If more than four point correspondences are given (over-determined set of equations), the solution is not exact. The least squares solution is:

$$\mathbf{h} = (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{b}) \quad (7)$$

4.2. Alternative Solution via Singular Value Decomposition

An alternative solution for the computation of the rectifying homography is based on the use of homogeneous coordinates. The equation $\mathbf{x}' = \mathbf{H}\mathbf{x}$ can be cast in a more convenient form using the product $\mathbf{x}' \times \mathbf{H}\mathbf{x} = \mathbf{0}$. The explicit form is given by:

$$\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ X' & Y' & 1 \\ \mathbf{h}^{1T}\mathbf{x} & \mathbf{h}^{2T}\mathbf{x} & \mathbf{h}^{3T}\mathbf{x} \end{bmatrix} = 0 \quad (8)$$

where \mathbf{h}^{1T} , \mathbf{h}^{2T} , and \mathbf{h}^{3T} are the rows of the matrix \mathbf{H} . This yields two equations (the third one is not linearly independent):

$$\begin{bmatrix} \mathbf{0}^T & -\mathbf{x}^T & Y'\mathbf{x}^T \\ \mathbf{x}^T & \mathbf{0}^T & -X'\mathbf{x}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^2 \end{bmatrix} = \mathbf{0} \quad (9)$$

and the final system has the form $\mathbf{A}\mathbf{h} = \mathbf{0}$ (linear in the unknown \mathbf{h}). The trivial solution $\mathbf{h} = \mathbf{0}$ can be avoided using the constraint $\|\mathbf{h}\| = 1$. A solution to solve this system is to perform singular value decomposition (SVD) on the matrix \mathbf{A} . SVD factors the matrix \mathbf{A} into a diagonal matrix \mathbf{D} and two diagonal matrices \mathbf{U} and \mathbf{V} as follows:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T \quad (10)$$

The solution is given by the last column of \mathbf{V} (in ordered SVD).

4.3. Evaluation of Metric Quality

Three reference objects featuring planar geometry were created installing 31 targets on three perpendicular planes: two vertical walls W1 and W2 and the floor F. Reference coordinates were measured with a Leica TS30 total station, obtaining three-dimensional coordinates (X, Y, Z) with a precision of about ± 1 mm.

A single 360° image was acquired with an Insta 360 One R, which has two front- and rear-facing fisheye lenses [43]. The final image (after stitching) has a resolution of 6080×3040 pixels. A picture image with the locations of the three planes is shown in Figure 7.

The extraction of the three rectilinear images was carried out varying the values of heading, pitch, and roll to move the center of the equirectangular projection close to the center of each wall. The chosen angles were $(0, 0, 0)$, $(90, 0, 0)$, and $(0, 0, 90)$. The field of view of the rectilinear images was set to $100^\circ \times 100^\circ$.

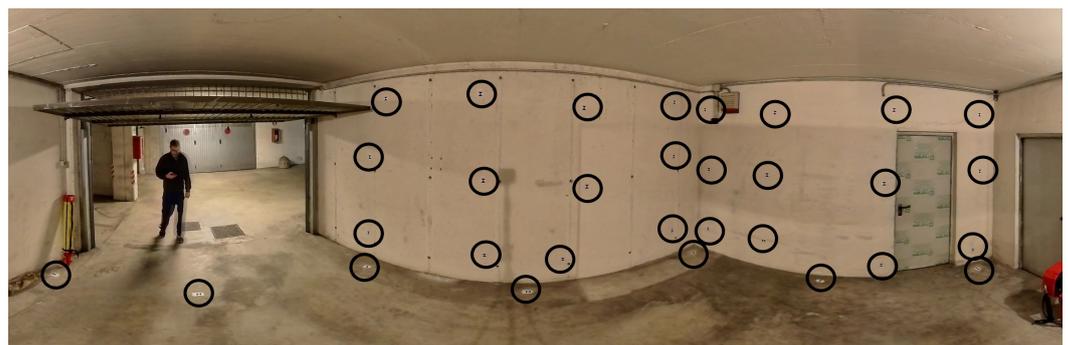


Figure 7. The reference object made up of three orthogonal surfaces measured with a single spherical image, which was partially cropped in this figure.

After creating the rectilinear images, image coordinates (x, y) were manually measured. The set of object coordinates (X, Y, Z) was instead split into three new groups using a rotation, placing the reference system in the plane of the walls, obtaining a set of ground truth coordinates (X', Y') , in which $Z' = 0$.

Three different systems of linear equations were written and solved, obtaining metric rectification parameters. After calculating the solution vector \mathbf{h} , images can be metrically rectified. Twelve points were used for the two vertical walls W1 and W2, whereas only seven targets were placed on the floor F.

Figure 8 shows the three rectilinear images extracted setting a squared field of view (on top). The metrically rectified images are shown in the bottom part of the figure. As can be seen, the reference system for object coordinates is not directed along the walls. The image of the floor is rotated depending on the horizontal orientation of the total station project.



Figure 8. (Top): The three rectilinear images extracted from a single spherical projection; (Bottom): the metrically rectified images.

The residual vector of least squares can be calculated as $\mathbf{v} = \mathbf{A}\mathbf{h} - \mathbf{b}$, and the posterior variance is $\sigma_0^2 = \mathbf{v}^T \mathbf{v} / (2n - 8)$. The formulation based on ordinary least squares (described in Section 4.1) was used.

A graph with the computed residuals for the different walls is shown in Figures 9 and 10. Wall W1 has smaller residuals than W2, notwithstanding that the camera–object distance is similar, leading to a ground sampling distance (GSD) of about 0.0015 m for both walls. The values for σ_0 were ± 0.002 mm and ± 0.011 mm, respectively.

The better results achieved for W1 can be explained considering the chosen rotation angles $(0, 0, 0)$, indicating that the fisheye sensor was pointing directly at the center of the wall. In the case of W2, the spherical image is the stitching result of part of both front- and rear-facing images. Similar considerations can be extended to the image of the floor, which shows even larger residuals.

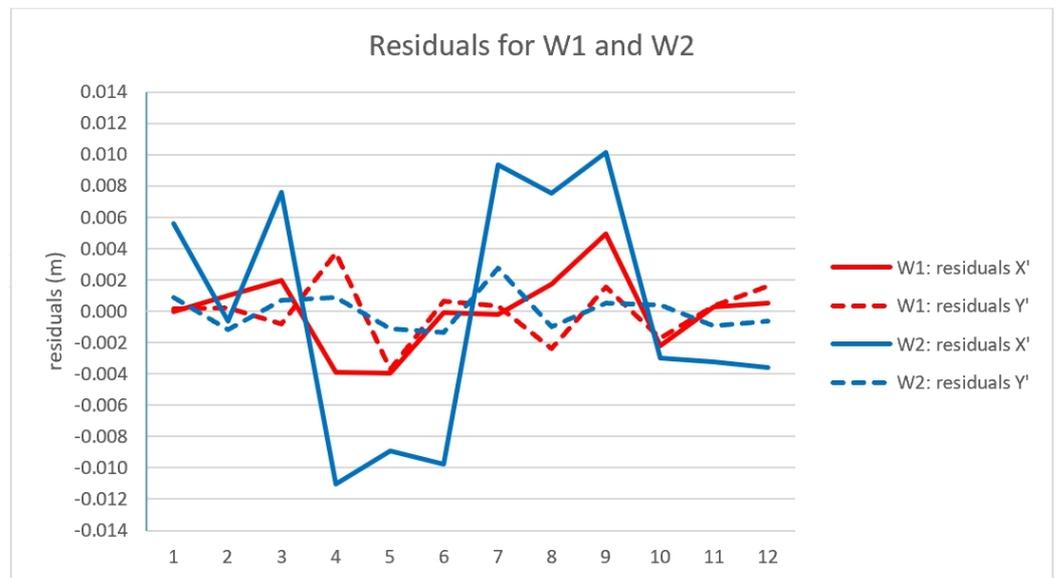


Figure 9. Residuals in meters for the different targets on the walls.

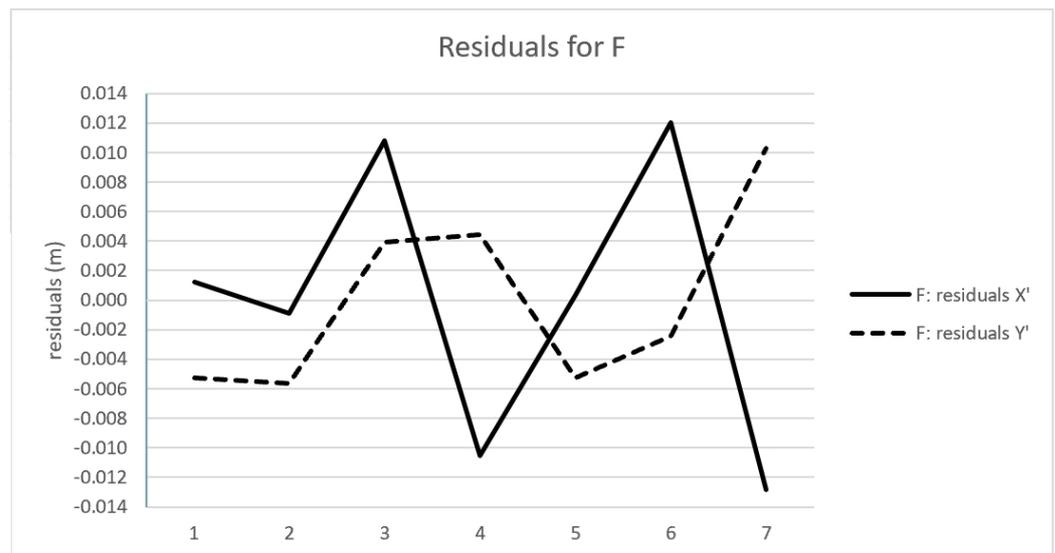


Figure 10. Residuals in meters for the different targets on the ceiling.

5. Geometric Rectification Based on the Focal Length and the Vanishing Line

5.1. Description of the Implemented Method

The focal length of the rectilinear projection coupled with two sets of parallel lines (not necessarily orthogonal) allows retrieval of metric properties up to a scale factor, which can then be recovered with a known distance.

As the focal length of the rectilinear image extracted from the spherical projection corresponds to the radius $f = r = L_x / (2\pi)$, geometric rectification based on sets of parallel lines can be carried out without a known aspect ratio, which is the more traditional approach for geometric rectification.

We consider the calibration matrix \mathbf{K} of the rectilinear image:

$$\mathbf{K} = \begin{bmatrix} f & 0 & w/2 \\ 0 & f & h/2 \\ 0 & 0 & 1 \end{bmatrix} \tag{11}$$

The identification of the rectifying homography requires the vanishing line of the plane. A generic line in the plane is represented by the equation $ax + by + c = 0$. A line can be identified by a vector $\mathbf{l} = (a, b, c)^T$. A point $\mathbf{x} = (x_1, x_2, x_3)^T$ lies on the line $\mathbf{l} = (a, b, c)^T$ if and only if $\mathbf{x}^T \mathbf{l} = 0$.

The vanishing line $\mathbf{l}^* = (l_1^*, l_2^*, l_3^*)^T$ can be computed from two sets of parallel lines [44]. First, a vanishing point \mathbf{v} can be estimated from the intersection of the lines \mathbf{l} and \mathbf{l}' using the product $\mathbf{v} = \mathbf{l} \wedge \mathbf{l}'^T$, as illustrated in Figure 11. Then, the calibration matrix and the vanishing line provide the orientation of the object plane with respect to the camera. The normal \mathbf{n} to the plane is $\mathbf{n} = \mathbf{K}^T \mathbf{l}^*$.

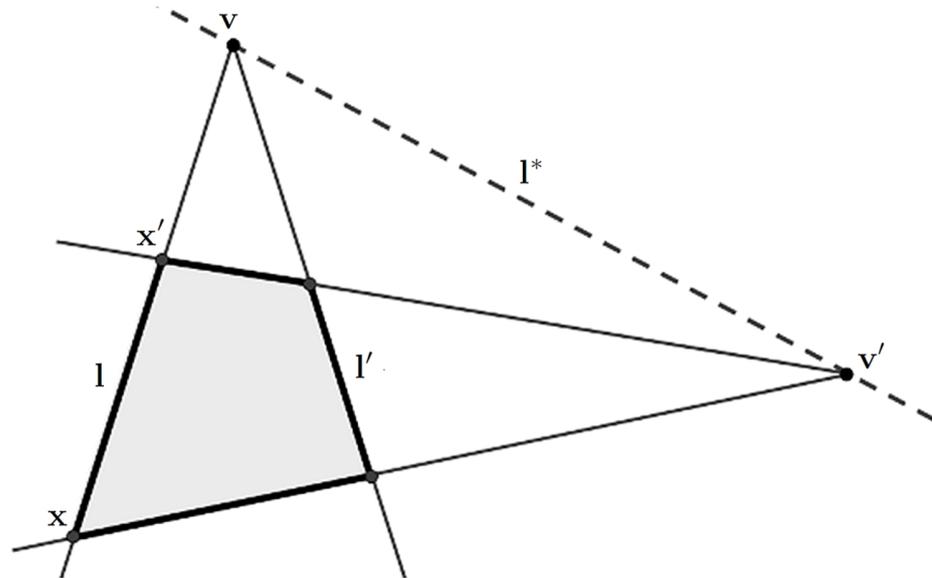


Figure 11. Residuals in meters for the different targets.

The image can be synthetically rotated to generate a new rectified image with a homography $\mathbf{H} = \mathbf{K}\mathbf{R}\mathbf{K}^{-1}$, implying that the unary vector $\mathbf{u}_n = \mathbf{n} / \|\mathbf{n}\|$ must lie along the the camera optical axis $\mathbf{R}\mathbf{u}_n = (0, 0, 1)^T$.

The matrix \mathbf{R} is made up of a set of vectors that forms an orthonormal set $\mathbf{R} = (\mathbf{u}_p, \mathbf{u}_m, \mathbf{u}_n)^T$.

The rectified image have an extra ambiguity due to a rotation because there exist an infinite number of vectors perpendicular to \mathbf{n} , resulting in an under-determined system of equations. Constraints must be applied to calculate the second vector \mathbf{u}_m . Finally, the last unary vector \mathbf{u}_n can be estimated with a cross product.

5.2. Evaluation of Metric Accuracy

Evaluation of metric accuracy of the implemented procedure was carried out using the same spherical image used in the previous section. The walls' vertical and horizontal lines were first inspected with the total station to verify their reciprocal orthogonality. Then, image rectification was carried out measuring the lines in the images and calculating two vanishing points and the vanishing line.

The experiment was performed using only the two vertical walls because the inspection with the total station revealed that the room's plan has a trapezoidal shape. The two vertical walls are sufficiently rectangular, with discrepancies of about $\pm 2\text{--}3$ mm.

Two metrically rectified images were generated for the two walls. Evaluation of metric accuracy was performed, estimating a similarity transformation between the pixel coordinates of the rectified image \mathbf{x}' and the corresponding object coordinates \mathbf{X}' measured with a total station after setting the reference system parallel to the considered wall.

The use of an extra similarity transformation applied to the metrically rectified image allows a direct comparison between the set of total station control points. The procedure

for metric rectification based on the focal length and vanishing line has a scale ambiguity as well as an extra rotation ambiguity due to the choice made for \mathbf{u}_m . An additional similarity transformation can recover the alignment of the two reference systems without altering the shape of the rectified images.

The transformation can be written as:

$$\begin{aligned} X' &= sx' \cos \alpha + sy' \sin \alpha + t_x \\ Y' &= -sx' \sin \alpha + s(-y') \cos \alpha + t_y \end{aligned} \quad (12)$$

where α is a rotation angle, s a scale factor, and (t_x, t_y) a translation vector. The component y' requires the negative sign to change the direction of the axis. Indeed, image coordinates are measured using a system with origin in the top-left corner of the images, and the y' axis is pointing downwards. In contrast, reference coordinates Y' measured with the total station are in a system pointing up. y' is therefore made negative to invert its direction.

The previous equations can be cast in a linear form:

$$\begin{aligned} X' &= ax' + by' + t_x \\ Y' &= -bx' - ay' + t_y \end{aligned} \quad (13)$$

with the substitution $a = s \cos \alpha$ and $b = s \sin \alpha$. Given a set of correspondences $(X', Y') \leftrightarrow (x', y')$, we may write a linear system of equations that can be solved via least squares:

$$\begin{bmatrix} x' & y' & 1 & 0 \\ -y' & -x' & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a \\ b \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} X' \\ Y' \\ \vdots \end{bmatrix} \quad (14)$$

The residuals are shown in Figure 12 and confirm that better results were still achieved for W1, like in the previous section. The discrepancy between the two walls is significantly more significant than in the case of using control points.

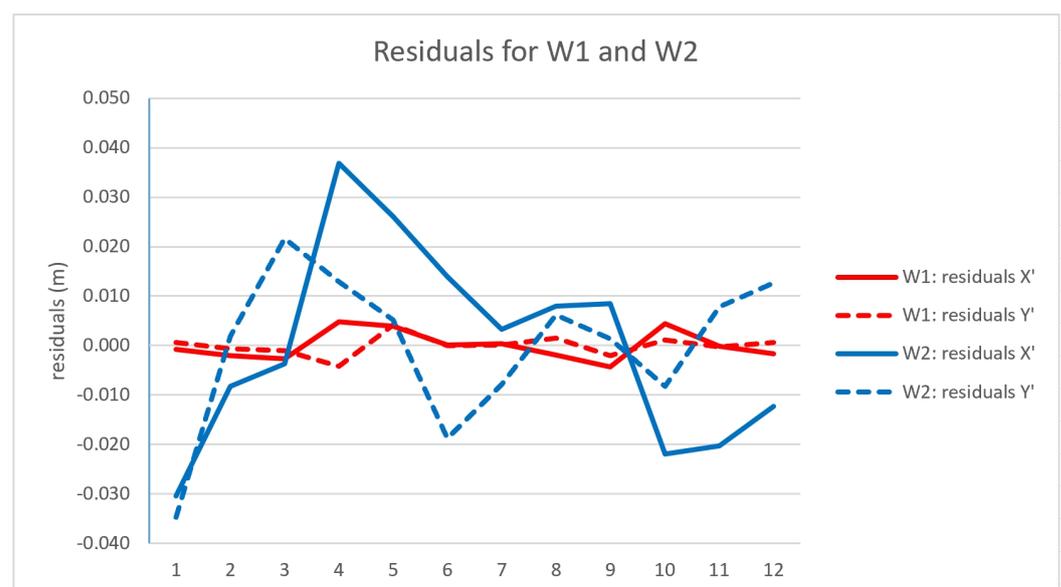


Figure 12. Residuals in meters for the different targets.

6. Considerations and Conclusions

The paper described a method for image rectification using spherical images. Recovering metric properties of planar objects is a common requirement in several technical disciplines requiring accurate digital documentation.

The difference between traditional metric rectification and the proposed method is the use of spherical images. Such images can be acquired with low-cost 360° cameras and allow rapid documentation because the camera can be pointed in any direction, capturing the whole scene around the photographer.

Because 360° low-cost cameras with better resolution are becoming more common among users, the use of their images for metric applications opens new opportunities for different specialists requiring metric documentation.

Photogrammetric packages that generate 3D models using sequences of spherical images are already available on the commercial market. However, there is a lack of a solution for 2D metric rectification based on the spherical camera model.

The 360° images used for metric applications feature several pros and cons compared to the traditional photogrammetric workflow with frame-based cameras. As mentioned in the introduction, different authors have already investigated the topic of 3D modeling with such images. Instead, the case of metric rectification with a single 360° image is new and practical applications are not available in the scientific literature.

The proposed method should be integrated in traditional rectification approaches for rapid documentation of flat surfaces in narrow spaces, especially for the interior of buildings with several rooms. The achievable image resolution (i.e., the ground sampling distance of the rectified image) and its metric accuracy cannot be compared to the results with a more traditional rectification approach. Although 360° images feature a high resolution (e.g., superior to 18 or 20 megapixels), the large field of view under an angle of 360° × 180° results in a GSD indicatively 5–6 times worse than the same rectification carried out with a frame camera.

However, it is the author's opinion that the rectification with 360° images can become suitable for those metric applications in which several internal surfaces of buildings have to be rapidly documented with a metric accuracy of ±2–4 cm, limiting the number of images collected, and providing an overall context using the 360° images themselves. As a 360° image provides an immersive (bubble) visualization of the entire space, it is simple for the user to recognize the room in which the image was acquired and the location of the walls in the room.

The implemented workflow relies on the preliminary conversion of the spherical image into a rectilinear image using a gnomonic mapping. The user's point of view can be oriented towards the planar surface requiring metric rectification. The field of view can also be interactively resized depending on the size of the area.

The rectilinear image can then be metrically rectified using two solutions: (i) the use of control points to estimate homography parameters or (ii) a geometrical approach based on the vanishing line coupled with the focal length of the rectilinear image, which can be derived from the radius of the sphere. This second method allows users to recover metric properties up to an overall scale ambiguity, which can be removed by measuring a single distance. The vanishing line of the plane can be instead calculated using at least two sets of parallel lines.

Metric evaluation experiments demonstrated that control points provide precision up to the ground sampling distance of the rectilinear image extracted from the spherical projection. The use of the geometrical method instead is simple and effective when the objects feature sets of parallel lines. In this second case, the user must know that the achieved metric image is affected by an overall scale ambiguity, which can be removed with a known distance.

Finally, an important consideration deserves to be mentioned. The authors acquired the spherical images used in this paper during metric evaluation with a low-cost Insta One R camera. Differing results were obtained for similar surfaces in similar project configurations, depending on the position of the surface and the front- and rear-facing fisheye lenses. Results for walls captured using the front-facing lens were better than the rear-facing lens, demonstrating that stitching of the two images introduces some deformations in the spherical image.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to thank Fabio Roncoroni and Mattia Previtali for their help in the preparation of some of the experiments described in the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Weigert, A.; Dhanda, A.; Cano, J.; Bayod, C.; Fai, S.; Santana Quintero, M. A review of recording technologies for digital fabrication in heritage conservation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 773–778. [\[CrossRef\]](#)
2. Luhmann, T.; Robson, S.; Kyle, S.; Boehm, J. *Close-Range Photogrammetry and 3D Imaging*; De Gruyter: Berlin, Germany, 2013. [\[CrossRef\]](#)
3. Vosselman, G.; Maas, H. *Airborne and Terrestrial Laser Scanning*; Whittles Publishing: Dunbeath, UK, 2010.
4. Santana Quintero, M. Harnessing digital workflows for conserving historic places. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 9–14. [\[CrossRef\]](#)
5. Smith, L.; Quintero, M.S. Digital workflows for the conservation and sustainability of historic places. *J. Cult. Herit. Manag. Sustain. Dev.* **2018**, *8*, 402–404. [\[CrossRef\]](#)
6. Lerma, J.L.; Cabrelles, M.; Navarro, S.; Fabado, S. From Digital Photography to Photogrammetry for Cultural Heritage Documentation and Dissemination. *Disegnarecon* **2013**, *6*, 1–8. [\[CrossRef\]](#)
7. Fassi, F.; Achille, C.; Fregonese, L. Surveying and modelling the main spire of Milan Cathedral using multiple data sources. *Photogramm. Rec.* **2011**, *26*, 462–487. [\[CrossRef\]](#)
8. Fassi, F.; Achille, C.; Mandelli, A.; Perfetti, L.; Polari, C. Recording the marble blocks of the Milan cathedral: Synergy between advanced survey techniques. In Proceedings of the International Masonry Society Conferences, Milan, Italy, 9–11 July 2018; pp. 2124–2134.
9. Shults, R. New opportunities of low-cost photogrammetry for culture heritage preservation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 481–486. [\[CrossRef\]](#)
10. Kim, H.; Hilton, A. 3D scene reconstruction from multiple spherical stereo pairs. *Int. J. Comput. Vis.* **2013**, *104*, 94–116. [\[CrossRef\]](#)
11. Fangi, G.; Nardinocchi, C. Photogrammetric processing of spherical panoramas. *Photogramm. Rec.* **2013**, *28*, 293–311. [\[CrossRef\]](#)
12. Barazzetti, L.; Previtali, M.; Roncoroni, F. 3D modelling with the samsung gear 360. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 85–90. [\[CrossRef\]](#)
13. Kwiatek, K.; Tokarczyk, R. Photogrammetric applications of immersive video cameras. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 211–218. [\[CrossRef\]](#)
14. Fangi, G.; Pierdicca, R.; Sturari, M.; Malinverni, E.S. Improving spherical photogrammetry using 360°OMNI-Cameras: Use cases and new applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2018**, *42*, 331–337. [\[CrossRef\]](#)
15. Matzen, K.; Cohen, M.F.; Evans, B.; Kopf, J.; Szeliski, R. Low-Cost 360 stereo photography and video capture. *ACM Trans. Graph.* **2017**, *36*, 148. [\[CrossRef\]](#)
16. Gottardi, C.; Guerra, F. Spherical images for cultural heritage: Survey and documentation with the NIKON KM360. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2018**, *42*, 385–390. [\[CrossRef\]](#)
17. Strecha, C.; Zoller, R.; Rutishauser, S.; Brot, B.; Schneider-Zapp, K.; Chovancova, V.; Krull, M.; Glassey, L. Quality assessment of 3D reconstruction using fisheye and perspective sensors. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2*, 215–222. [\[CrossRef\]](#)
18. Herban, S.; Costantino, D.; Alfio, V.S.; Pepe, M. Use of Low-Cost Spherical Cameras for the Digitisation of Cultural Heritage Structures into 3D Point Clouds. *J. Imaging* **2022**, *8*, 13. [\[CrossRef\]](#)
19. Escher, T. *Spherical Photogrammetry with the Panono Camera*; gis.Business: Berlin, Germany, 2018; pp. 26–28.
20. Aghayari, S.; Saadatseresht, M.; Omidalizand, M.; Neumann, I. Geometric calibration of full spherical panoramic Ricoh-Theta camera. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 237–245. [\[CrossRef\]](#)
21. Barazzetti, L.; Previtali, M.; Roncoroni, F. Can we use low-cost 360 degree cameras to create accurate 3D models? *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2018**, *42*, 69–75. [\[CrossRef\]](#)
22. Sun, Z.; Zhang, Y. Accuracy evaluation of videogrammetry using a low-cost spherical camera for narrow architectural heritage: An observational study with variable baselines and blur filters. *Sensors* **2019**, *19*, 496. [\[CrossRef\]](#)
23. Pramulyo, H.; Harto, A.B.; Mertotaroeno, S.H.; Murtiyoso, A. *Towards Better 3D Model Accuracy with Spherical Photogrammetry*; Lecture Notes in Geoinformation and Cartography; Springer: Cham, Switzerland, 2017; pp. 107–120.
24. Bruno, N.; Roncella, R. Accuracy assessment of 3d models generated from Google Street View imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 181–188. [\[CrossRef\]](#)

25. Xie, D.; Zhong, R.; Wu, Y.; Fu, H.; Huang, X.; Sun, Z. Relative Pose Estimation and Accuracy Verification of Spherical Panoramic Image. *Cehui Xuebao/Acta Geod. Cartogr. Sin.* **2017**, *46*, 1822–1829.
26. Cui, T.; Ji, S.; Shan, J.; Gong, J.; Liu, K. Line-based registration of panoramic images and LiDAR point clouds for mobile mapping. *Sensors* **2017**, *17*, 70. [[CrossRef](#)] [[PubMed](#)]
27. Barazzetti, L.; Previtali, M.; Roncoroni, F.; Valente, R. Connecting inside and outside through 360° imagery for close-range photogrammetry. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 87–92. [[CrossRef](#)]
28. Pisa, C.; Zeppa, F.; Fangi, G. Spherical photogrammetry for cultural heritage—San Galgano Abbey and the Roman theater, Sabratha. *J. Comput. Cult. Herit.* **2011**, *4*, 9. [[CrossRef](#)]
29. Prizeman, O.E.C.; Barazzetti, L. Low-cost 3d acquisition of geometric data for living heritage: Attempting to record the pudhu mandapam, madurai. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2021**, *46*, 555–562. [[CrossRef](#)]
30. Losè, L.T.; Chiabrando, F.; Tonolo, F.G. Documentation of complex environments using 360° cameras. The santa marta belltower in montanaro. *Remote Sens.* **2021**, *13*, 3633. [[CrossRef](#)]
31. Fangi, G. *Some Metric Documentation of Cultural Heritage in Poland by Spherical Photogrammetry*; Lecture Notes in Geoinformation and Cartography; Springer: Cham, Switzerland, 2018; pp. 189–203.
32. Fangi, G. Documentation of some cultural heritage emergencies in Syria in August 2010 by spherical photogrammetry. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2*, 401–408. [[CrossRef](#)]
33. Koeva, M.; Luleva, M.; Maldjanski, P. Integrating spherical panoramas and maps for visualization of cultural heritage objects using virtual reality technology. *Sensors* **2017**, *17*, 829. [[CrossRef](#)]
34. Argyriou, L.; Economou, D.; Bouki, V. Design methodology for 360° immersive video applications: The case study of a cultural heritage virtual tour. *Pers. Ubiquitous Comput.* **2020**, *24*, 843–859. [[CrossRef](#)]
35. Karara, G.; Hajji, R.; Poux, F. 3D point cloud semantic augmentation: Instance segmentation of 360° panoramas by deep learning techniques. *Remote Sens.* **2021**, *13*, 3647. [[CrossRef](#)]
36. Abate, D.; Toschi, I.; Sturdy-Colls, C.; Remondino, F. A low-cost panoramic camera for the 3d documentation of contaminated crime scenes. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 1–8. [[CrossRef](#)]
37. Dai, X.; Ducey, M.J.; Wang, H.; Yang, T.; Hsu, Y.; Ogilvie, J.; Kershaw, J.A. Biomass estimates derived from sector subsampling of 360° spherical images. *Forestry* **2021**, *94*, 565–575. [[CrossRef](#)]
38. Ihsanudin, T.; Affriani, A.R. The Combination of Spherical Photogrammetry and UAV for 3D Modeling. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Shanghai, China, 19–22 October 2017; Volume 98.
39. Liebowitz, D.; Zisserman, A. Metric rectification for perspective images of planes. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, USA, 23–25 June 1998; pp. 482–488.
40. Barazzetti, L.; Previtali, M.; Scaioni, M. Procedures for condition mapping using 360° images. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 34. [[CrossRef](#)]
41. Brown, M.; Lowe, D.G. Recognising panoramas. In Proceedings of the IEEE International Conference on Computer Vision, Washington, DC, USA, 23–26 October 2003; Volume 2, pp. 1218–1225.
42. Brown, M.; Lowe, D.G. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* **2007**, *74*, 59–73. [[CrossRef](#)]
43. Campos, M.B.; Tommaselli, A.M.G.; Marcato Junior, J.; Honkavaara, E. Geometric model and assessment of a dual-fisheye imaging system. *Photogramm. Rec.* **2018**, *33*, 243–263. [[CrossRef](#)]
44. Liu, T.; Xu, W.; Yin, X.; Zhao, X. Planar metric rectification from perspective image via vertical vanishing point. In Proceedings of the 2014 7th International Congress on Image and Signal Processing, CISP 2014, Dalian, China, 14–16 October 2014; pp. 498–502.