

## Article

# Multi-Camera Workflow Applied to a Cultural Heritage Building: Alhambra's Torre de la Cautiva from the Inside

Anto J. Benítez <sup>1,2</sup> , Xose Prieto Souto <sup>1,2,\*</sup>, Manuel Armenteros <sup>1,3</sup>, Esteban M. Stepanian <sup>1</sup>, Rubén Cantos <sup>1</sup>, Miguel García-Villaraco <sup>3</sup>, Jaime Solano <sup>2</sup> and Ángela Gómez Manzanares <sup>4</sup>

<sup>1</sup> Department of Communication, Universidad Carlos III of Madrid, 28903 Madrid, Spain; abenitez@hum.uc3m.es (A.J.B.); marmente@hum.uc3m.es (M.A.); estepani@hum.uc3m.es (E.M.S.); rcantos@hum.uc3m.es (R.C.)

<sup>2</sup> Tecmerin Research Group, Universidad Carlos III of Madrid, 28903 Madrid, Spain; jsolano@hum.uc3m.es

<sup>3</sup> Miopía FX, 28006 Madrid, Spain; mike@miopiafx.es

<sup>4</sup> Department of Optics, University Complutense of Madrid, 28040 Madrid, Spain; angomez@ucm.es

\* Correspondence: japsouto@hum.uc3m.es

**Abstract:** It is increasingly necessary to generate accessible and navigable digital representations of historical or heritage buildings. This article explains the workflow that was applied to create such a digital component for one of the least accessible areas of the Alhambra palace in Granada, the so-called Torre de la Cautiva (Tower of the Captive). The main goal of this process was to create affordable, photorealistic 3D models that contribute to the dissemination of cultural heritage, the decision making for its conservation and restoration, and public engagement and entertainment. With enough preparation, the time spent gathering data following a Structure from Motion (SfM) approach can be significantly reduced by using a multi-camera (low cost DSLR) photogrammetric strategy. Without the possibility of artificial lighting, it was essential to use RAW images and calibrate the color in the scene for material and texture characterization. Through processing, the amount of data was reduced by optimizing the model's topology. Thus, a photorealistic result was obtained that could be managed and visualized in immersive Visual Reality (VR) environments, simulating different historical periods and environmental and lighting conditions. The potential of this method allows, with slight modifications, the creation of HBIMs and the adaptation to VR systems development, whose current visualization quality is below the resolution of actionable models in rendering engines.

**Keywords:** cultural heritage; photogrammetry; workflow; Alhambra's Tower of the Captive



**Citation:** Benítez, A.J.; Prieto Souto, X.; Armenteros, M.; Stepanian, E.M.; Cantos, R.; García-Villaraco, M.; Solano, J.; Gómez Manzanares, Á. Multi-Camera Workflow Applied to a Cultural Heritage Building: Alhambra's Torre de la Cautiva from the Inside. *Heritage* **2022**, *5*, 21–41. <https://doi.org/10.3390/heritage5010002>

Academic Editors:

Daniel Vázquez-Molini and Antonio Álvarez Fernández-Balbuena

Received: 28 October 2021

Accepted: 18 December 2021

Published: 22 December 2021

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



**Copyright:** © 2021 by the authors. 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

A fundamental concern in the field of Cultural Heritage (CH) is the need to document historical monuments as accurately as possible. Before the advent and popularization of digital culture, it was proposed to carry out extensive suitable inventories of each site [1] with photographs and drawings used to capture the nature of these environments, and to guide conservation, restoration, or excavation activities [2,3]. Technological advances have expanded the possibilities of documenting these CH sites, using not only natural light, but also monochromatic light, ultraviolet light, and infrared rays [4]. Subsequently, photogrammetry was specifically mentioned as a standard procedure to document the properties of cultural interest in the 1987 Charter for the Conservation of Historic Towns and Urban Areas [5].

From the time atoms began to be converted into bits [6], digitization has become increasingly widespread, allowing not only for assets to be documented for their conservation, preservation [7] or restoration [8], but also for digital versions to be created. With these, which constitute the Digital Cultural Heritage (DCH), dissemination actions are promoted, play is proposed [9] and access is allowed [7] through virtual tours even in complicated circumstances such as those experienced in recent times, marked by the COVID-19 pandemic [10].

Some arguments have been presented against the creation of DCH models. For example, their cost, the complexity of the processes or the specialization required to obtain them [11,12]. These are arguments that have been losing strength with the simplification of software platforms, the increase in the power of hardware equipment and the lowering of the cost of both elements [9,13].

At the same time, the need for a multidisciplinary approach has been recognized (The Norms of Quito, 1967). New approaches have emerged that make it possible to produce more complete documentation and expand the possible uses of the 3D models that were obtained. An increasing number of heritage buildings have an informative model that was obtained through Historic Building Information Modeling (HBIM) procedures [9,14]. Such models are based on a complex and precise methodology [15,16] originating from architectural functionalities. These approaches encourage the construction of parametric models with the integration of many disciplines [16].

Irrespective of their origin, 3D models have been employed in the interest of the conservation and restoration of CH sites, but also for other objectives. Education [12] or entertainment in the form of video game scenarios or Virtual Reality (VR) experiences [17] are examples. Their use with the holistic intention of gathering as much knowledge as possible about a building and reaching their recipients in the form of 'edutainment' is also remarkable [9]. Moreover, the data collected as a part of these processes can help in the restoration of irreplaceable sites after unexpected catastrophes, as recently highlighted by the fire at the Notre Dame cathedral in Paris [18].

Thus, this article is part of the exploration of the possibilities of digital modeling for historical heritage. Workflows (or pipelines) that are applicable to photogrammetry are being simplified, although they are still not fully standardized. Some of them are becoming clearer and simpler, making them accessible to the uninitiated, but they coexist with others that do require specialization. Some of the possible results can still be achieved in very different ways, sometimes by retracing routes that have already been experienced and commented on. The case that is discussed in this article has tried to avoid this as much as possible by using audiovisual procedures and using only digital photography cameras in the capture, with the aim of obtaining a useful model that works as a video game scenario as well as in a conservation and restoration environment. It was launched before some initiatives were published, several of whose elements we agree with [17], and with which we share the view that it will be necessary to continue looking for procedures that are easy to implement, effective, efficient, cheap, and that can be standardized or become universal.

In an attempt to gain a wider picture by drawing on a multidisciplinary environment, more than a few cinematography techniques are also suitable for projects of this type [19]. Minimum standards of precision and accuracy or fidelity must be achieved in the multi-spectral characterization of the surfaces in order to obtain the right data for HBIM [3,20]. During shooting, cinematographic photography is similarly concerned with the on-screen results, which are achieved through adequate attention to exposure and color calibration. As such, in both cases there is a need to take extreme care in the preparation procedures for shooting.

DHCs can and should be used for dissemination as they are part of a collective heritage [7,9,21]. The main goal of the audiovisual sector is that the stories engage the audience. Since the same is true for the dissemination of DCH, it would be logical to use the methods and procedures adopted by the former, i.e., defining and following a script prepared for that purpose [17]. Indeed, some of the processes that will be discussed in this text are specific to the audiovisual industry, such as: pre-production, multi-camera shooting in rigs, care in exposure, the use of RAW files, re-topology and the integration of lighting and environment.

In this case, the aim was to achieve immersion and presence, with full awareness of having discarded the real environment, in the sense attributed to it by [22]. The aim was to provide a simple viewing in VR with a level of quality that allows the user to experience photorealistic sensations [23].

There are ever more platforms available for this type of experience, although the systems still suffer from many limitations, as their manufacturers warn [17]. It is important to emphasize that at the time of starting this project, the most widespread VR devices were still unable to correctly reproduce the resolution and quality of the models that are being obtained through photogrammetry. However, given the vertiginous advances of the last few years, this team has no doubt that within a short time the performance of the systems that allow VR to be experienced will surpass that of most existing scenes and models.

The objective should be independent of the platform used. Rather, it will have to refer to the appearance of real materiality of the content, not to the predisposition to which the technology subjects the viewer. It will have more to do with contemplation, with the amazement of being in the middle of a story that leaves a feeling of photorealism. It will have to take advantage of the immersive capacity of the narrative device in order to overcome the resistance that the spectator may oppose to letting him or herself be carried away by the story, in the sense that whoever experiences it may reach a state of pathos such as that described by Sergei M. Eisenstein, whereby one is taken out of oneself and immersed in the reality presented on the screen [24]. To achieve this immersion, it was essential to work with lighting and the setting of scenarios with techniques that achieve an optimal degree of integration [12,17].

This article describes all of the processes that were carried out in order to obtain the model of the interior of the Torre de la Cautiva (The Tower of the Captive) of the Alhambra in Granada, to achieve a stereoscopic piece that can be experienced in VR 360°, according to a script that was set in simulations throughout different epochs. For a proper analysis of the experience, we will first present a brief approach to the selected case study and the challenges it posed. Subsequently, the methodology and workflows will be explained in greater detail. Finally, the conclusions of this study will be presented and a series of future lines of research will be proposed.

### *1.1. La Torre de la Cautiva (The Tower of the Captive)*

The Alhambra-Generalife complex in Granada, Spain, figures on the World Heritage List since 1984, according to UNESCO records.

The Alhambra has been the ideal setting for artists and writers of Romanticism, inspired on numerous occasions by folklore and medieval historical plots. According to legend, Doña Isabel de Solís, an Andalusian noblewoman, was kidnapped and fell in love with the king of Granada, Muley Hacén. She converted to Islam and adopted the name Zoraya. The outlines of this story were part of Washington Irving's book 'A Chronicle of the conquest of Granada' [25]. A few years later, Francisco Martínez de la Rosa published the historical novel 'Doña Isabel de Solís, Reina de Granada' [26], which delves into this story and is the one that justifies the name of the tower.

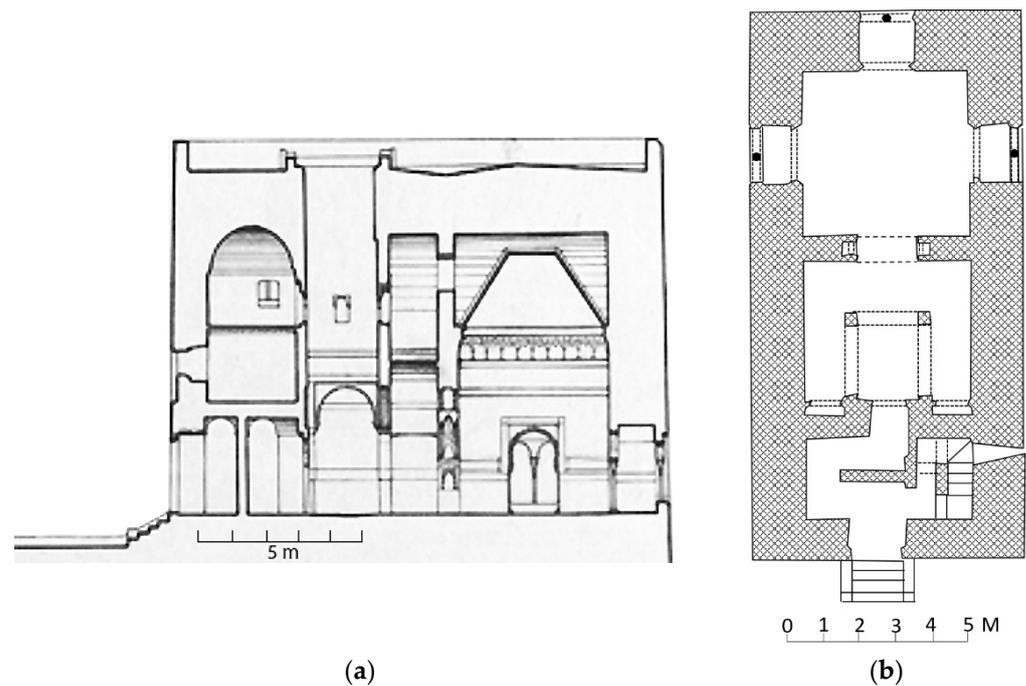
The tower is part of the north-northeast walled enclosure and is located between the Torre del Cadí (Tower of the Judge) and the Torre de las Infantas (Tower of the Princesses). It was built in the mid-fourteenth century, in the time of Yusuf I. Its exterior has a sober defensive aspect and a very simple geometry, but inside it is richly decorated. The main hall measures about 5 m × 5 m and its walls contain a good number of epigraphic poems composed by Ibn al-Yayyab (1274–1349).

In addition to the plans drawn up by architects or draftsmen, which are preserved in the Archives of The Alhambra and Generalife Trust, there is also a physical model of part of the interior of the tower, which was made by Enrique Linares in the last quarter of the 19th century, in the Victoria and Albert Museum in London [27]. It weighs 100 kg and is almost one meter high.

### *1.2. Case Study*

The Torre de la Cautiva is one of the parts of the Alhambra that is not usually open to the public. In the summer of 2018, the challenge was posed to produce a pilot piece of the interior of the tower. Its goal was to explore the possibilities of 3D models obtained through

affordable passive-sensor data-acquisition techniques, not only as a simulation platform for conservation and restoration, but also as a vehicle for dissemination and engagement. Two rooms were to be covered: the main room, with profuse decoration and ambient light, and the secondary room or courtyard, with simpler decoration and little natural light (see Figure 1b). The result would be used in stereoscopic 3D in the Head Mount Display (HMD). In Milgram and Kishino's Reproduction Fidelity dimension, the aim was to achieve what they call the "ultimate' graphic rendering", referred to as "real-time, high-fidelity 3D animation" [28].



**Figure 1.** (a,b) Surveys with delineation techniques of the profile and floor plans of the Torre de la Cautiva. The main room has balconies on the walls, while the courtyard or secondary room has no direct lateral exits to the exterior. Compiled by the authors based on documents and plans maintained in the collection of the Patronato de la Alhambra y El Generalife (The Alhambra and Generalife Trust): P-1158\_2 and P-4615.

A script was drafted to create a dramatic scene that would take place in different epochs and, therefore, different architectural and decorative moments would be shown in the same piece. The action was meant to observe a certain historical consistency. Some characters inhabited the stage, but they were not overly prominent. In this case, it was not so important to obtain data that were faultless from the architectural point of view, that is to say, the result did not have to consist of an informative HBIM model.

After becoming familiar with the first draft of the script (handling and division into temporal sequences), three strategies were combined, one for each line of work. The first consisted of the application of multi-camera photogrammetry using a vertical rig (see Figure 2) for interiors and walls in their current state of preservation. The second consisted of photogrammetry with a *giratutto* [29] to obtain the characters. Finally, in the third, conventional 3D modeling was used for the setting and assets.

Another of the project's goals, which was aimed at distinguishing this work from those that used similar approaches, was to achieve a look of veracity, or photorealism, by capturing images with low-cost DSLR cameras. The result would be experienced by the audience on a platform that allowed interaction with a freedom of observation on all three axes (pan, tilt, roll), but on principle, without the freedom of displacement. These interaction conditions are known as three degrees of freedom (3DoF) on the scale of

freedom of use [30]. It would also have a sound program to guide and complement what the viewer discovers.



**Figure 2.** Diagram of the vertical rig with three cameras. Image by the authors.

An additional challenge was the time that was available for data capture, only ten hours spread over one afternoon and the following morning, which was a key factor in choosing the method with which to carry out this task. To complicate matters further, surveying methods to secure control points could not be used, nor lighting supports for photography.

### *1.3. Choice of Strategy*

A wide repertoire of methods is available with which to obtain digital models [12,21], yet the unique nature of each project implies that the solutions employed in earlier strategies are not always the best approaches to adopt, and that the advances in hardware or software are not the only factors to consider [31]. Likely the only generalization that can be made for digital-model-acquisition projects is that there is not yet a methodology that is applicable to the majority of cases.

Among the known strategies, Structure from Motion (SfM) [7,11,13,29,32,33] using only digital camera photography is best adapted to the circumstances associated with the data collection in this study, and it has been used without a set order of shooting [7] or even without checkpoints. This feature constituted a good point of support, since we could not count on having active sensors or Total Stations in our work. Moreover, it was known that stereoscopically configured work reduces the need for a high number of shots [11,34].

Although positive results had been reported with mobile devices [35,36] (which is later than our data collection), much of the success of the work depended on the flexibility and reliability of the cameras in capturing, and therefore the cell-phone option was ruled out. We considered cell phones as offering insufficient and unreliable control over exposure in the conditions in which we had to work, which we would still argue today [36]. The aim was to interfere as little as possible with the physical environment, so the motto was “get in, do it and get out”, just as in photogrammetry projects that involve the police or forensics, which cannot alter the environment in any way whatsoever [37]. With this remit it was clear that the background preparation for the data capture had to be as thorough as possible.

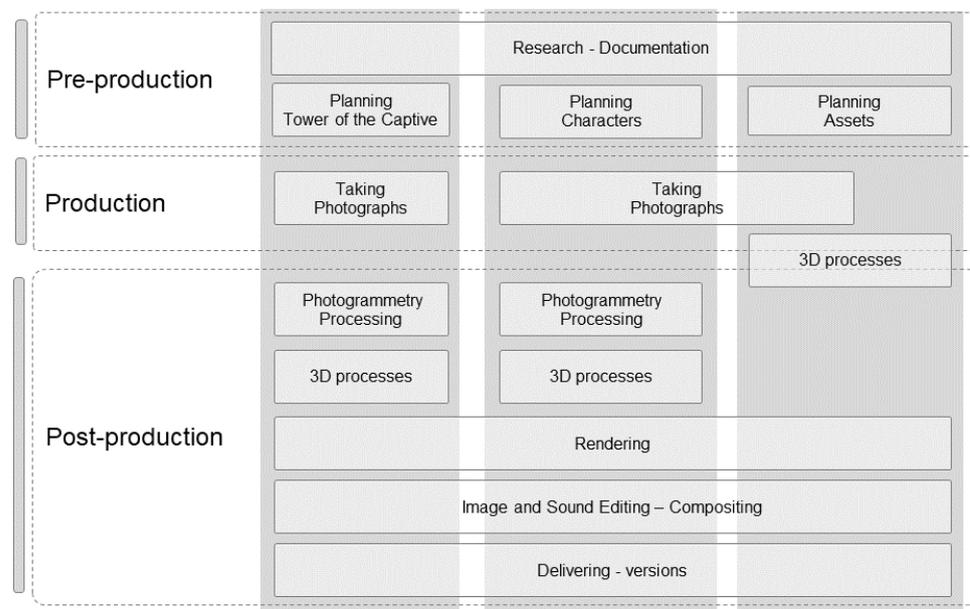
### 1.4. Photorealism

The great cinematographer Néstor Almendros used to argue that “we must learn to transcend the model, while respecting it” and that “the stylization of an image can sometimes be more important than history and logic” [38]. Obviously, this philosophy cannot be applied to obtaining data for HBIM, yet it is valid to immerse the public in the illusion of photorealism within an environment defined by a script. Similarly, it is not valid to simply confront the user with a 3D graphic model [23] via an HDM that fails to provide the user with any other immediate references to reality. Rather, it is necessary to go further, to create an impression of authenticity and accuracy in the user [20,39], but in this case one that is diachronic. It must be borne in mind that to do so implies moving with the imagination to the past, and the capacity to compare the result of a photograph of that time [40,41] with what is being experienced at that moment.

## 2. Methodology

This project was approached in three phases that are not necessarily sequential or consecutive: pre-production, production, and post-production. Based on approaches used in the audiovisual industry, these phases were applied to a script and focused on different aspects: inside the tower, obtaining the characters, and the creation of assets. While it is clear that careful preparation is required, as reflected elsewhere in the literature [42,43], phases such as *Data Capture*, *Data Processing* and *Output Presentation* can sometimes be found, which either do not consider shot preparation as a relevant step or do not include it in their first stage [17,29–31,33]. Due to the importance of the decisions that had to be taken in the preparation of the workflow, it is established here as a singular phase, while the presentation of the final result is included in post-production as one of the processes that usually make it up [44].

The general outline of the procedures (Figure 3) will be broken down below, with special attention to the fact that one of the objectives was specifically the development of the workflow.



**Figure 3.** Scheme of the processes employed in the tower project. All the preparative decisions are grouped together in the pre-production phase. Compiled by the authors.

### 2.1. Pre-Production Stage

The pre-production phase of the pipeline had to ensure that the restrictions that were imposed on the shooting inside the tower were not insurmountable, since the rest of the

lines of work could be readily controlled. After a process of documentation and deliberation about the different possibilities, it was concluded that the only certain method of covering the surfaces in such a short time and without lighting aids was to use a vertical rig of cameras shooting simultaneously (See Figure 2). The preparation of each shot would take more time, but if properly synchronized, the total time required could be divided by the number of cameras used. Initially, calculations were made for a vertical rig consisting of four mid-range DSLR cameras anchored to a rigid, telescoping pole on a tripod. The path of the rig would be parallel to the walls in different passes at different distances, and the relative position of the cameras on the rig would be changed to obtain segments at different heights. According to the specialist literature, this approach has rarely been used for photogrammetry projects and it is much more common to find stereoscopic approaches using two cameras [3,32].

The cameras that were used were four Canon EOS 1200D with 25–80 mm lenses for the main room, and one additional Canon EOS 60D with a 24–70 mm lens for the second room, bringing the total to five simultaneous cameras. The cameras were set to a resolution of  $5184 \times 3456$  pixels, about 18 MP (megapixels). The sensor of both rooms was identical, and according to the manufacturer had a total of 18.7 MP, but 18 effective MP and a size of  $22.3 \text{ mm} \times 14.9 \text{ mm}$  with an aspect ratio of 3:2. This means that the sensor captured 232.47 pixels per mm in width and 231.95 in height (dividing the number of pixels by the number of millimeters). In area:  $53,919.7 \text{ pixels}^2/\text{mm}^2$ . If photographs were taken at 80 cm from the walls, at a focal length of 25 mm, one would be working with just over three pixels/mm.

The problems of color characterization and exposure could be aggravated by the need to use natural light. The fact that the spaces were either diaphanous or almost diaphanous worked in our favor. The reference for our work was a color chart, such as Color Checker [45], which is often used in film shoots. In order to reduce the possibility of occlusions caused by over- or under-exposures, a fixed aperture was used to ensure a sufficient depth of field and low sensitivity in order to avoid problems with noise reading, and RAW files were also used in order to have the possibility of correcting the exposure values later [29,39,45].

Regarding the characters, the script conceived them as static figures, establishing situations and suggesting activities, but keeping them static so that the scenery would be the most prominent feature. The animation was applied to environmental or weather conditions, so that it would have an impact on the appearance of the surfaces. Research was done on the clothing, styling, and tools with which to dress the models, and they were asked to pose continuously in order to be captured by with photogrammetry methods, but on a *giratutto* and in controlled and constant lighting conditions.

In this phase, research was carried out on the assets that would be used to decorate the main hall. The initial documentation was collected precisely in the Alhambra's own museum, but a number of paintings set in the Nasrid period were also used.

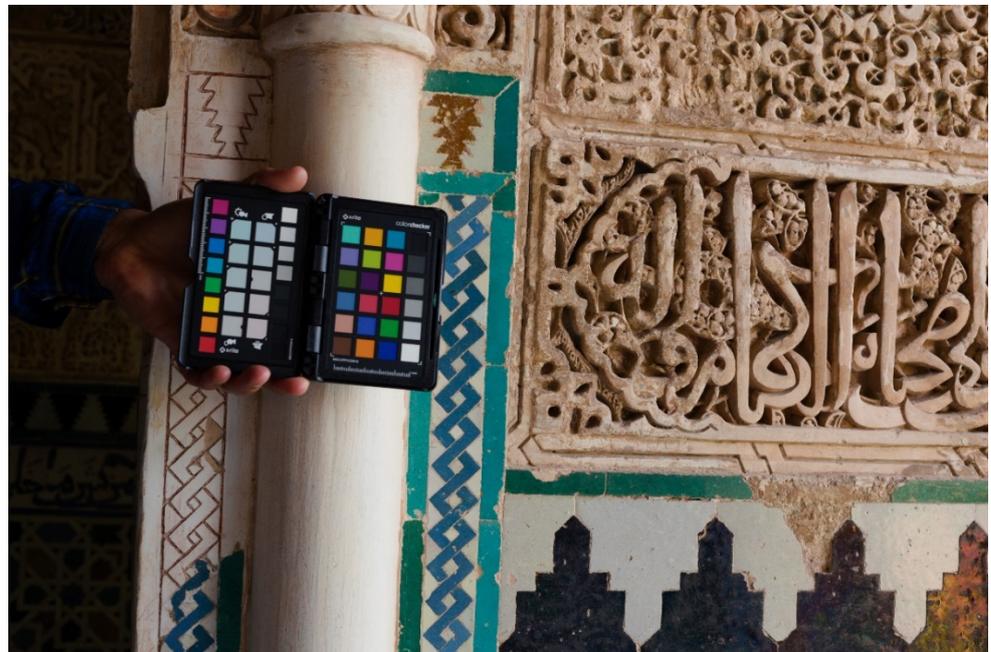
## 2.2. Production Phase. Shooting Inside the Torre de la Cautiva

Due to the prevailing weather conditions on the days of shooting, i.e., cloudy with clear spells, it was decided to use a vertical rig of three cameras for the main room and four cameras in the courtyard, given the more complicated geometry but simpler decoration of the latter (see Figure 1b). A route was established that ran parallel to the walls, with stops to take shots every 15 cm. At each height, four passes were made in both rooms, starting almost four meters from the parament that was being portrayed and moving closer. To compose the first horizontal band, the cameras were placed on a vertical pole on a tripod, with the camera axis parallel to the ground and perpendicular to the wall at approximately 80, 130 and 170 cm. In the second band, the cameras were settled at approximately two meters, 2.35 m and 2.70 m. Another band was not established due to the risk of excessive camera shaking causing a blurring of the pictures at the exposure speeds that were used. Furthermore, the user experience with 3DoF VR indicates that

the starting point for observation should be at the height of the person experiencing the room, which enables emphasis to be placed on the strong areas in the final result. This way of working requires coordination between the teams that are in charge of the script and photogrammetry, as demanded by some authors [17].

The courtyard needed more time to adjust the exposures and shots, and a Canon 5D camera was used to take cover photographs in hand-held mode at the brightest times of the day. These photographs focused on the most awkward details and the most difficult angles, considering that the main hall had arches with muqarnas and other filigree decorations. This camera was also used to take photographs of the coffered ceiling, which was particularly difficult due to its inverted boat hull shape, making it very difficult to achieve a sufficient depth of field. This situation was aggravated by the minimal relief and the homogeneity of the surface.

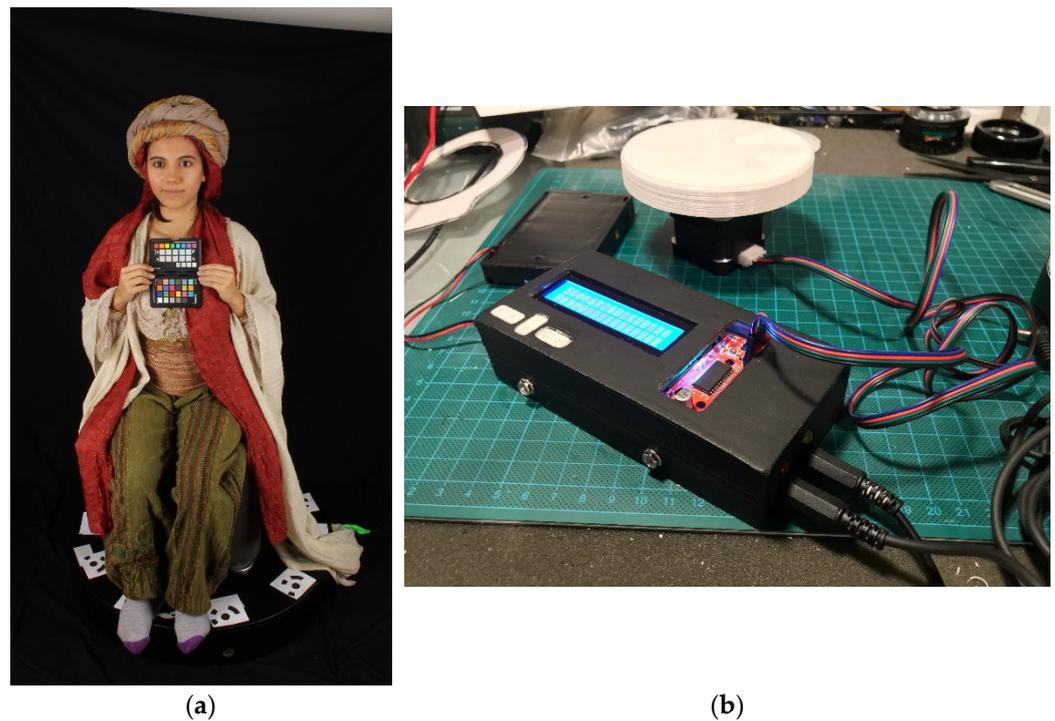
The shots of the main room were taken with a focal length of 25mm, at 100 ISO, an aperture of  $f/8$ , a manual exposure at  $1/6$  of a second, and were corrected according to the lighting conditions; manual white balance with presets for dense and light clouds and Color Checker images were taken in both situations (Figure 4).



**Figure 4.** A Color Checker shot near the epigraphic poems on the wall was used to quickly and automatically adjust the color, and to ensure fidelity of the textures and models.

In the inner room, being less accessible to natural light, 20 s exposures were taken. A fourth Canon 60D camera with a 25 mm focal length was attached. The material was simultaneously saved as 30 Mb RAW files and 8 Mb JPEGs.

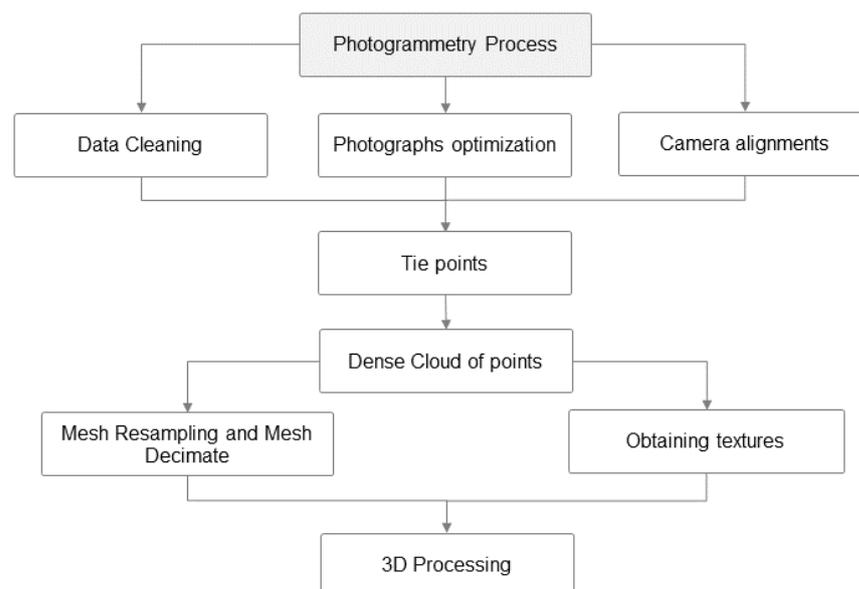
To obtain the images of the characters, a setup was made with a turntable with tickers to mark control points, a lighting installation with four flash-type fixtures (4 Linkstar DL-500D) and three vertical rigs of five Canon EOS 1200D cameras each [29]. Once characterized with period costumes and styling, the models were asked to pose as still as possible while the *giratutto* was set in motion. After photographing the Color Checker, a shot was taken every 30 degrees (Figure 5a), synchronized to the photographic flash lights through an Arduino ad hoc setup (Figure 5b) and triggering the 15 cameras simultaneously. This was done with two poses of the six different models, sometimes changing the characterization. Finally, the shots of three characters were used, one of them in duplicate.



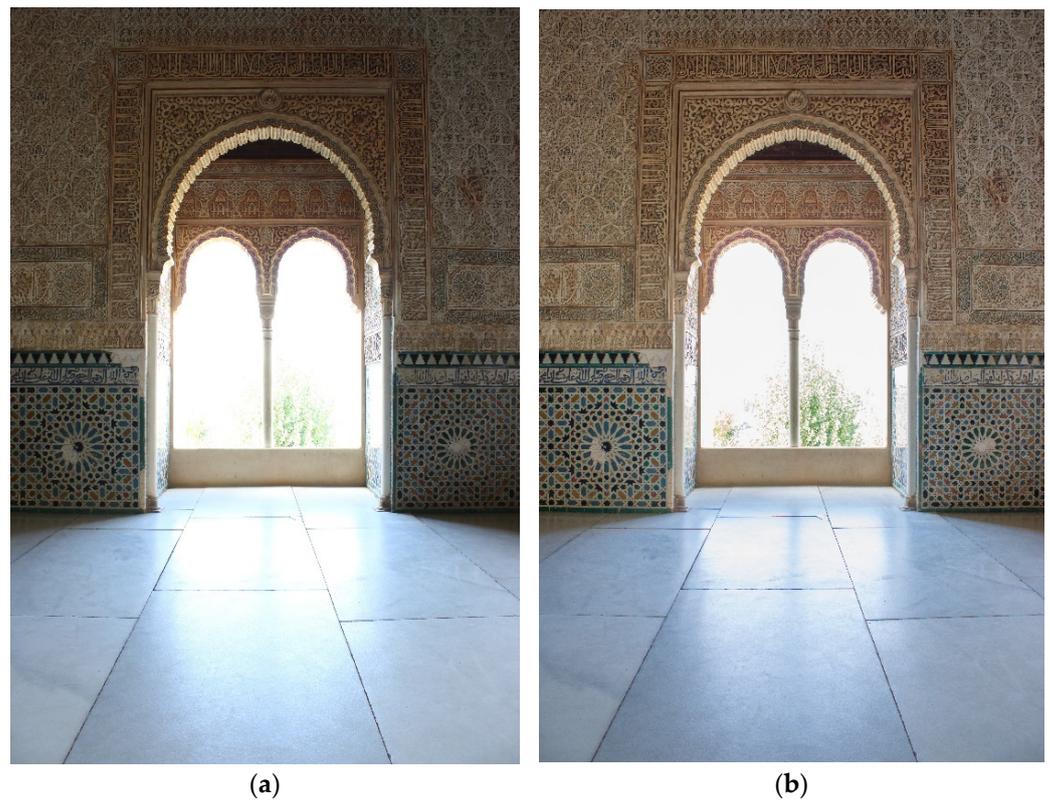
**Figure 5.** (a,b). Image of one of the models (a) on the giratutto with the control points and the color checker. In (b), the Arduino device used that is connected to a small turntable is shown.

### 2.3. Post-Production Phase

The first part of the processing (Figure 6), photogrammetry, required an initial step to prepare the photographs, using the RAW files [29], with two objectives: to adjust the color from the images that had been taken from the chart and to try to avoid occlusions due to incorrect exposures (Figure 7a,b).



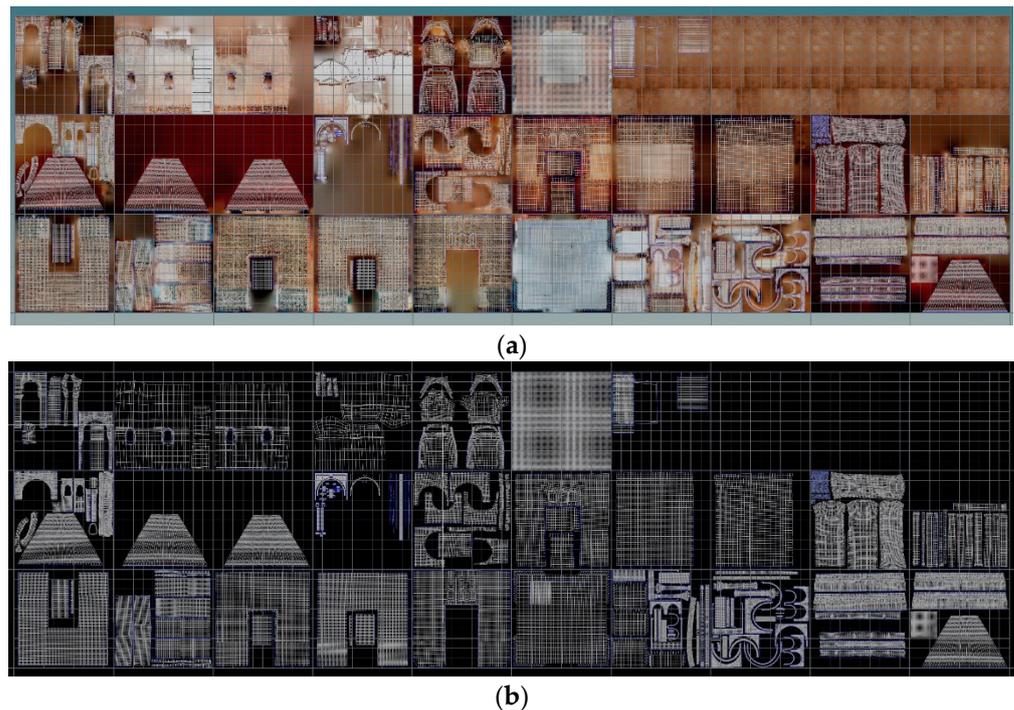
**Figure 6.** Pipeline for the photogrammetry phase. Compiled by the authors.



**Figure 7. (a,b).** Image before and after color correction from the RAW file. Note how the walls are much darker in the image on the left ((a), before applying corrections) owing to the contrast and how the edges of the column on the balcony are barely visible. When using a RAW file, the EVs can be modified to optimize the image before (b) copying it to the photogrammetry software.

Meanwhile, camera alignment tests were carried out to check if the software, Reality Capture 1.3, recognized the camera positions well and provided enough homologous points, which in this case was a number close to 22 million. Through adjustments and corrections, a corrected data set was obtained and redundant points were eliminated from the dense point cloud in order to achieve one that was as faithful as possible to the shots. This was necessary because the software is asked to propose a geometric model that is built with triangles, and this model is usually disproportionately large with a huge number of triangles and is very difficult to handle. In this case, about 50 million triangles were counted in the first version of the mesh.

With careful planning of the photographic sessions, the photogrammetry software can produce a geometry with sufficient detail and with textures that are true to the original surfaces. Once the cloud of points was obtained through photogrammetry, simplifying the mesh proved to be useful for the subsequent tasks, as recently noted elsewhere [17,29]. However, there are programs available that specialize in this type of task, such as Zbrush, which recreates the mesh in a much simpler form through a process known as re-topology. As a result, the mesh was eventually reduced to about 200,000 polygons. This process was necessary for two reasons: first to restore areas that have been left with little or no detail [17], which in our case affected the corners of the floor and some small interior sectors of the balcony arches; second, it offers the possibility of creating clean UVs in order to obtain masks that extend the resolution of the textures using multi-UDIM. If the program originally generated an 8K texture for each room, through this procedure it was possible to have that resolution for each of the maps that were generated, 27 in total (see Figure 8a).



**Figure 8.** (a,b). UDIMs with textures (a) and masks (b) prepared for back-projection on the optimized model once the re-topology process had been completed.

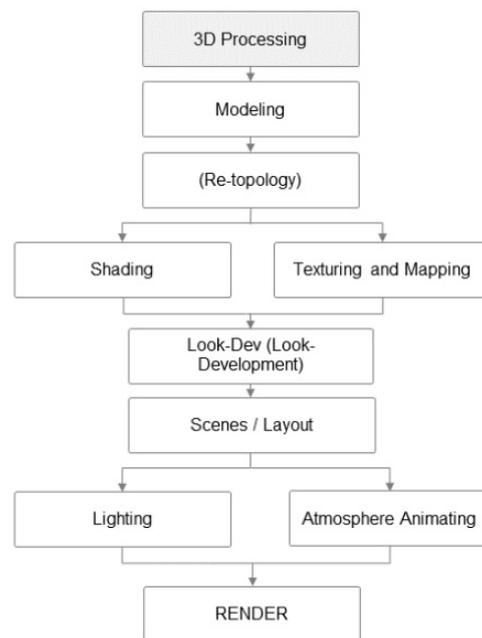
The textures that were obtained in this way were albedo (from photogrammetry with color and shadow or exposure correction), tile mask, wooden coffered ceiling mask, stone surfaces mask, floor mask, albedo for the artisan scene (because this scene is conceptualized at the time before the room decoration and therefore needs to be colorless), a displacement map, a normal map, a roughness map and a specular map [17,30].

A few days after the capture, the photogrammetry-generated 3D model visualization tests were carried out with the Clarisse iFX 3.5 program. Re-topology was a complex and laborious task, especially the geometry of the arches, in which there were many small details that required significant precision. It was also necessary to intervene in the wooden coffered ceiling, which was difficult to capture due to the lack of light and the lack of detail, which prevented the software from differentiating between sectors and locating the homologous points. Nowadays, in the same programs, there are new algorithms that perform very accurate automatic re-topology, especially for static models, but at the time of this project they did not exist, so everything was done by hand.

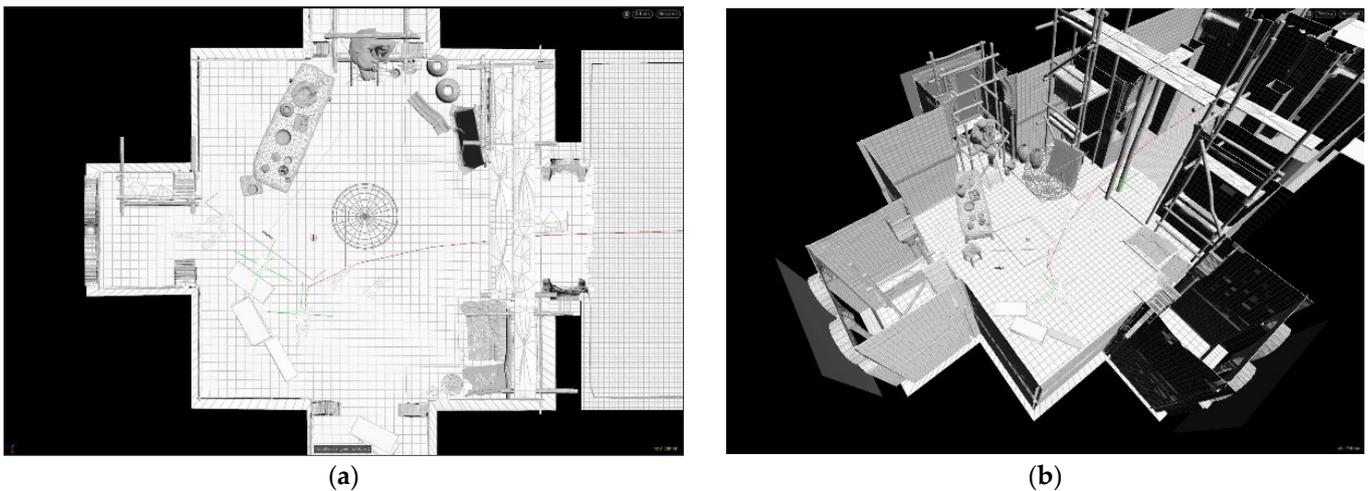
The shading and look-dev processes (Figure 9), which are common in the work of 3D artists in audiovisuals or video games, for example [17], were directly responsible for the desired photorealism in this study, since they controlled the perception of the surface of the objects and considered their reaction to the light illuminating the scene and the atmosphere.

The modeling and look-dev of the assets were also performed using Maya and Zbrush. In total about thirty were worked on, although not all of them were included in the final scenes. Carpets, curtains, cushions and pillows, the brazier, the jamuga (a type of medieval chair), the latticework, vases, the rebec, scaffolding, masonry objects and painter's tools, lamps, etc., were all inspired or taken directly from originals that were kept in the Alhambra Museum or the National Archaeological Museum of Madrid.

As mentioned above, these phases do not have to be carried out consecutively and thus, to save time, the integration of the characters and assets in the scenery can be done gradually, starting by obtaining a geometry of materials that are not yet fully worked on or are in wireframe. Positions, sizes, scales and points of view can be harmonized with enough experience to handle scene overlays in this way (see Figure 10a,b).

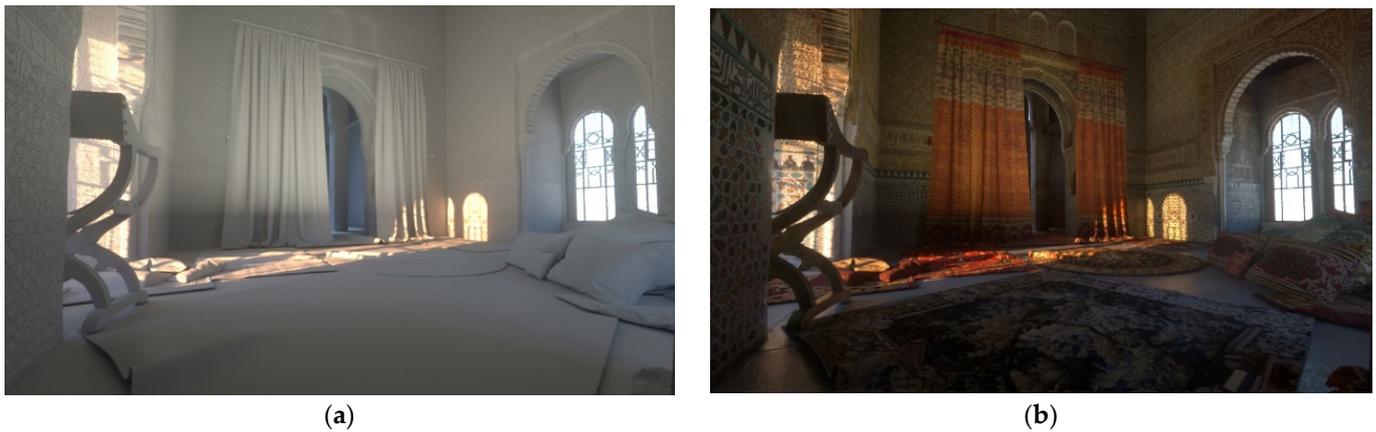


**Figure 9.** Diagram of the workflow of the last phases of post-production, just before obtaining the final artwork. Own elaboration.



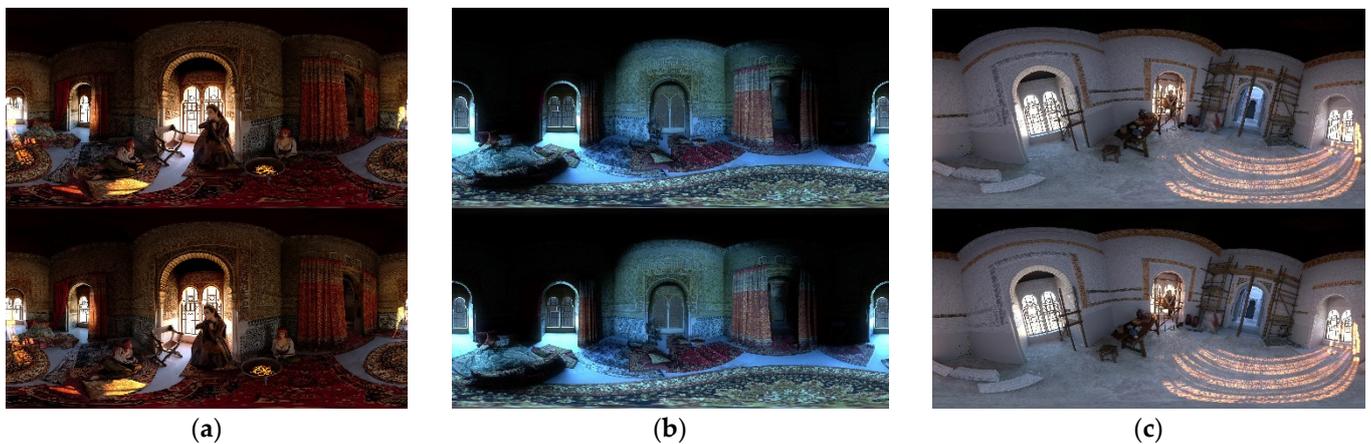
**Figure 10.** (a,b). Craftsman scene in plan (a) and perspective (b) views in wireframe render. It is crucial to arrange the elements to be able to work interactively with quick previews within the layout of the scenes.

However, in order to check if the scene works, it is necessary to test with more or less the final textures and a setting that brings the scene together properly, after ensuring that the chosen light interacts correctly with the surfaces and the camera positions and animations are well chosen. In addition to other qualities, the view of the integrated scene must not be disjointed. To achieve this, the first rendering tests were performed with Arnold. However, each image took between three and four hours to process at 4K with its optimal parameters (Figure 11), which was unacceptable, especially to observe the result of the animation of lights and cameras and make final decisions about its placement on screen. It was possible to reduce the time for each image at 4K to 25 min using Houdini 17.5 with Octane render 2019 1.2.0, which is a GPU-based engine, and a computer equipped with two NVIDIA 2080 TI GPUs.



**Figure 11.** (a,b) Light tests are essential to observe the integration of surfaces and environmental elements in the scene, as well as the suitability of the arrangement of elements and camera positions. For this, it is necessary to launch renders, which in conventional form can take up to four hours using powerful computers for processing.

Following the script, the scenes of the present day, the night storm, the artisan and the captive were composed, while fine-tuning the light and camera animations and creating transitions between them. With a previous low-quality rendering, the sound program was created, and the voiceover, music, effects and Foley were adjusted. This task, together with the sound synchronization, was performed using DaVinci Resolve 16. This program was also used to perform the final rendering at a resolution of  $4096 \times 4096$ , with Quick Time output format without YUV 4:2:2 compression. A conversion to the  $360^\circ$  3D stereoscopic format of the BT (Bottom-Top) type was performed so that it could be experienced with Oculus Rift glasses, with an appearance similar to that shown in Figure 12a–c.



**Figure 12.** (a–c) Scenes of the captive (a), the night storm (b) and the craftsman (c) in  $360^\circ$  3D Stereoscopic BT format.

### 3. Results and Discussion

#### 3.1. Multi-Camera Techniques with Rigs

Using a multi-camera data-capture setup ensures the overlapping of images so that each point appears in more than one, which is known since stereoscopy has been used for the same processes [3,11,32,46]. But it allows us to go a step further, covering a larger area in the same amount of time, relative to the number of cameras at each shooting position. A simple technique to prepare the multi-camera shot is to imagine the space in slices (which would be equivalent to the geographical longitude, the vertical sector to be covered), which are represented by each shooting position of a vertical rig, and segments (the geographical

latitude, the horizontal sector to be covered), which are represented by each camera anchor on the tripod.

There is concern about the minimum cost that needs to be addressed for a photogrammetry project [7,11,29,33,45]. While increasing the number of cameras clearly increases the budget, at current prices, it is still much less than covering the costs of using more photographers, a laser scanner or a Total Station [36]. In this project, mid-range DSLR cameras with about 18 Mpix resolutions and conventional optics were used, but mirrorless, medium-format cameras with higher resolution (thus providing more detail from the same number of shots) and faster speeds, because each shot does not involve mechanical displacement, are becoming more affordable. When a large number of shots is involved, time savings are multiplied, especially if bracketing is used to ensure correct exposure, which reduces the possibility of occlusion due to under- or over-exposure. In addition, access to better optics also reduces the exposure time, but raises the budget. Nevertheless, this project was less expensive overall as the costs associated with the number of cameras used were much lower than those that would have been imposed by increasing the number of photographers [45].

Another important issue to foresee is the shooting path. It is often necessary to resort to complex algorithms, such as those of the Traveling Salesman Problem for active sensors [47]. In this case, the chosen route was more or less simple and always parallel to the walls with sufficient shots. To generalize the optimization of the route, one should perhaps resort to artificial intelligence (AI), which considers factors such as the geometry of the object, the available equipment, the environmental conditions and the demands of the script [17]. Any software used must be able to interpret camera positions and locate sufficient homologous points, thereby avoiding occlusions. In our case, it was important to cover all possible angulations in order to obtain a sufficient number of overlapping points, but equal care must be taken not to oversize the shots, as it makes the project unmanageable. On the other hand, when using a multi-camera setup, coordination errors that may exist are multiplied, so extreme care must be taken in this regard. Other factors are focus and exposure, which are explained below.

Some 1900 photographs were taken in the tower (7 cameras  $\times$  4 passes  $\times$  2 heights  $\times$  29–35 rig positions), and some 800 without a tripod. Among them, about 2500 were used for the photogrammetric models. For the characters, about 2150 pictures were taken (6 models  $\times$  2 poses  $\times$  12 shots  $\times$  15 cameras) and about 720 were used.

The time required to make adjustments before each shot increases when using the multi-camera technique. The exposure and focus must be set manually, even if some parameters are automatic. Again, affordable technology facilitates the operation, as nowadays it is common to control camera viewfinders remotely (both the image displayed and the overlays) as they are the main operations for exposure, focus and shooting.

Another problem was camera shaking; in order to gain height and ensure angles that might otherwise cause occlusions, the cameras had to be raised as high as the pole would allow without swaying. Maintaining stability with light equipment three meters off the ground was difficult, but essential when the shots exceeded 1/30th of a second. This circumstance arose in virtually all of the cases in our work. However, the results were equally dramatic, enabling views from the coffered ceiling with the viewer positioned four meters above the floor to be included in the final experience.

There is no doubt that a robotic system that incorporates AI, which is capable of making some decisions regarding settings and routes and also of detecting and solving possible occlusions, would ultimately simplify and reduce the cost of projects, leading to optimal results.

### *3.2. The Importance of Preparation in Pre-Production*

Audiovisual projects such as film animation or video games are exceptionally complex, and the procedures they have been developing and turning into knowledge with their experience can certainly be beneficial to apply to a photogrammetry challenge that should

have a photorealistic final look. With this conviction, this project has assumed some of their techniques.

A key factor for the success of the project, given the restrictions that were to be experienced during data collection, was to be aware of the relevance of the pre-production-style decision-making process. What had been studied and agreed upon during this phase significantly determined the subsequent capacity for action, since it contained an element that was lacking in both the filming and post-production stages, i.e., time [48]. In the specialized literature, significant relevance is always given to the duration and intensity of this phase [49,50].

The first decision was to work on a draft script with a fixed number of sequences. It was prepared according to the capabilities that the project wanted to demonstrate: the creation of environments, the simulation of past scenarios with naturalness, and mastery over the scenery. This mastery was especially directed towards the fidelity of the geometry, the variety in the application of textures, and the animation of light sources.

Foresight and the careful preparation of the capture allows for decisions to be made ahead of time in terms of overcoming the main challenges, which allows more time to be dedicated to study and optimization of the workflow, thereby avoiding confusion. This factor will also help to define the areas where authenticity and accuracy are more or less important, so that the script does not conflict with the pre-production workflow of time optimization. On the contrary, storytelling should be another criterion by which to define technical parameters such as the number of polygons and their layout [17]. It is important to establish fluid communication between those who develop the script and the members of the team responsible for taking the shots and processing the digital models in order to optimize work times and effort.

### 3.3. Exposure and Color Correction

In cinema, a scientific approach to questions concerning the control of color and exposure when shooting has long been employed. It is common for all image- and video-editing programs to implement procedures with externally calibrated measurements [45], using very accurate meters and different procedures to control EVs, as well as employing Color Checker. These procedures, which are also described in the literature specialized in photogrammetry [29,45] were used in this project to solve the color-characterization problem and to avoid occlusions due to exposure problems. By obtaining RAW and compressed JPEG files simultaneously, a safety margin was established in the acquisition that neutralized the risks associated with working with natural light. In fact, during the days of data capture, the sky was very cloudy in Granada and at times quite dark. The use of RAW files allowed us to vary the exposure values and, to some extent, to correct the fluctuations and prevent over- or under-exposure.

In terms of focus, there is always a moment when the image is zoomed in when the circle of confusion is no longer small enough and, therefore, when further scaling up means losing sharpness. Obtaining the correct focus in the shot is very important both for the subsequent definition of the textures and for the photogrammetry program to determine where the points of interest are, and to be able to identify the homologous points. In this project, according to the provisions made in pre-production, we wanted to ensure a minimum aperture of  $f/8$  in order to obtain a sufficient depth of field. In this way, the details of the decorations (4 cm fretwork or longer muqarnas) could be captured without the need to correct the focus at each rig position. Such a decision implied a trade-off that increased the exposure time, which ran the risk of blurring the edges and consequently of increasing the time needed for stabilization and exposure. This time can be carefully calculated so that it is less than the time that would be invested in an eventual focus correction. An alternative solution would be to increase the sensitivity, which was kept at 100 ASA, but such an operation would have the consequence of increasing the noise in the images, which is detrimental to the unequivocal obtention of homologous points. Finally,

thanks to this set of decisions, it was ensured that the results surpassed what the playback or viewing platforms are currently capable of resolving in terms of fluidity and sharpness.

### 3.4. Photorealism

Shading and look-dev processes are crucial for making the scenery, assets and characters appear photorealistic on screen, but lighting and environmental effects are also equally important. In this project, a detailed study of each character in the environment, light and optics was carried out (see Figure 13a–c), and such a challenge requires a lot of time with conventional methods, so it is recommended, after the experience, to use a rendering engine for testing.



**Figure 13.** (a–c). Some of the lighting, framing and environment tests that were carried out with the captive character. Note the variations in the depth of field due to the use of different optics in the simulation.

Sometimes a large number of possibilities for configuration or modification can be counterproductive in achieving clarity in the final results, even making them less convincing. Nevertheless, such tools make it possible to obtain well-integrated scenes that give the illusion of coming close to what historical reality must have been like, or what could now be interpreted as such. In this sense, the audiovisual industry’s way of proceeding consists of trying to make the integration of the different elements invisible, so that even experts find it difficult to reverse engineer and extract the method that was used to achieve this result.

Photorealism aims to provide the audience with a product that seems natural, such that studies of user experience and acceptance should be considered in order to determine if this is achieved. Strictly speaking, by definition, no image can be photorealistic if it is intended to portray a past era, as it will resemble, at best, what other previous artists have imagined those realities to be. If 3D artists are given the opportunity to work closely with those in charge of the care of monuments and with specialists such as archeologists and architects, many possibilities open up for their conservation, or at least for the simulation of interventions and recreations.

In this project, the starting point was a narrative script that placed the story in the realm of fiction. However, under certain conditions and scales, the use of photorealism is not incompatible with the possibility of obtaining the necessary and sufficient data to make the HBIM of the heritage sites usable by conservators and restorers, from a data-collection method that is more or less accessible and rigorous with the color parameters and with elements of anchorage or location. For the objective of photorealism, accuracy and fidelity are not as important as the feeling of naturalness and coherence with the imagination of the audience.

It is true that the desire for photorealism pushes a project towards obtaining an excessive amount of data, thereby generating problems of manageability. For now, not all systems and formats can handle all of these models and sometimes the rendering capacity is insufficient. However, if the scenes have been set up with consistency and are ambitious, the renderings can be improved and adapted to systems with better performance in the near future.

### 3.5. Re-Topology

Coinciding in time with the work presented in this article, but more diligent in publishing their results, the text 'New realities for Canada's Parliament: A workflow for preparing Heritage BIM for game engines and Virtual Reality' [17] already explains clearly and sufficiently the importance of re-topology in the processes of building 3D models and, therefore, in photogrammetry projects. It does not deny that it could always have been photographed better and more effectively and, as is proposed in the mentioned research, that it would have been helpful to have a previous understanding of the script with respect to the points of most interest to the viewers, in order to make decisions during the data capture and processing that would have helped in the composition and the rendering.

Without going into the depths of the analysis, the 3D models that were directly produced by the data from photogrammetry platforms far exceeded the dimensions that were manageable by graphics cards and processors, however powerful they may have been, and yet such a profusion of data does not result in improved accuracy or fidelity, nor photorealism, as we have seen. An intermediate process, re-topology, is necessary to simplify the models, retouch possible occlusions or imperfections and produce textures that, once back-projected onto the new model, allow greater detail that is more similar to the real scenarios, and much greater fluidity in the interaction with the scenes.

The specific issue of viewer interaction with the scene also needs to be addressed, as there are specific methods that can improve the performance of the platforms. At the end of 2015, Nick Kraaman explained in his Headjack blog [51] the progress made by different professionals to develop scenes in which viewers would be immersed and equipped with HMD. One of the points made was the possibility of covering up areas of inactivity with static images instead of video, which would make interaction in the other areas much smoother and take up far fewer system resources.

There is no doubt that platforms will continue to improve, as will hardware devices, and it is clear that, by knowing the geography of the scene well and foreseeing the points of greatest interest, it will be possible to make better use of resources.

### 3.6. Integration and the Problem of Previsualization

Logic or common sense establishes an order when it comes to the integration of all the elements of the scenes. Once the geometries have been created, the textures made available, the look-dev has been worked on and the lighting, camera positions and eventual environmental effects and animations have been thought of, it is time to see what the scenes will look like.

But working according to a script has advantages: the challenges are already established, the range of possible decisions is narrower, a narrative structure that makes sense can be followed and the dimension of the result is more or less clear from the beginning. It is also possible to fit the pieces together without all the elements being finished, so that processes can be followed. For example, in this project it was possible to work in sequences, fitting elements into the layout and studying camera trajectories before the textures were applied.

As the time will ultimately arrive to check what the final result will look like, rendering must be performed. The usual and most sensible thing to do is to choose some frames first and process them, and then do the same with small sequences, using definitive textures (much more laborious for the machines) or materials depending on what you want to test.

Scenes with large amounts of information require a lot of calculation time and, therefore, necessary decisions must be postponed until the end of the computer processes. The experience of this project before and after using render engines for these tasks is clear: the entire scene must be passed to an engine of these characteristics in order to speed up the decision-making process as much as possible before the final calculations. The time required for pre-testing is drastically reduced in this way. This assertion is all the more valid when the final layout of the scenes is less studied, i.e., if there is no script, it is more essential to carry out tests with immediate visualizations.

#### 4. Conclusions

The Torre de la Cautiva is located in one of the walls of the Alhambra in Granada, a complex named a World Heritage Site in 1984, and is one of its enclaves to which the general public does not usually have access.

Bearing in mind its historic value and the fact that it can rarely be enjoyed in person, the challenge of creating a photorealistic 3D model of its interior was posed. This initiative was intended to serve both conservators and restorers, and as a DCH site for public dissemination and entertainment. Throughout this article we have explained the workflow that achieves this.

Regarding the execution of the project, it should be noted that the data collection was carried out with important restrictions: only ten hours to shoot, using only natural light, and with no cables or control points. The model was built with low-cost photogrammetry techniques, using conventional DSLR cameras as passive sensors. The display was planned on 360° Stereoscopic VR platforms with interactivity and 3DoF on the scale of freedom of action.

The script contemplated several scenes that predominantly focused on the past, although one of them was set in the present day. The others recreated a night storm, a typical scene of the period in which the interior of the tower was decorated, and a third evoking the routine atmosphere of the captive princess's lodging. In addition, three lines of work were defined for the script: that of the stage, which should be given maximum prominence, that of the characters with a narrative sense and that of the assets, fulfilling the function of ambience and decoration. A specific procedure was used to obtain data for each of these lines.

Without reaching the demands that HBIM informative models may have, a certain historical and traditional coherence with the period scenes was attempted, as well as a certain level of detail and fidelity to the original decoration.

Although in agreement with the academic literature in many points and starting from a script and a desire for photorealism, this project also adopted some procedures commonly employed in the audiovisual industry, which is accustomed to this type of complexity. The pre-production process is therefore analyzed in more detail within its workflow.

Since it is difficult to reach an agreement on the meaning of photorealism, especially if the past is to be shown, it was established as a principle of the project that it could be perceived naturally and coherently by the viewer. In this case, moreover, being accurate with geometry and color characterization was assumed as being of the utmost importance.

Another common process in the audiovisual industry that was an important part of the project workflow was the multi-camera work, which was carried out with vertical rigs and careful planning.

In the post-shooting processes, the importance of re-topology to obtain simpler meshes and of UDIM textures to retain great definition of detail was clear, as was the importance of the script to be able to work separately on the integration of the elements in the scenes, their lighting, environmental effects, and animation.

Two key ideas for future work stem from the experience of this project: the convenience of using robotic devices coupled to AI algorithms to optimize decisions during the data acquisition process, and the adaptation of rendering engines, not only for final viewing or creation of VR environments, but also for previsualization and tests, since they radically reduce the time required to perform different procedures.

**Author Contributions:** Conceptualization, A.J.B., X.P.S., M.A.; methodology, A.J.B., X.P.S., M.A.; validation, A.J.B., X.P.S., M.A., E.M.S., J.S., Á.G.M.; investigation, A.J.B., X.P.S., M.A., Á.G.M.; resources, M.A., M.G.-V.; data curation, M.G.-V., J.S., R.C.; writing—original draft preparation, A.J.B.; writing—review and editing, A.J.B., X.P.S., E.M.S.; visualization, A.J.B., M.A., M.G.-V., R.C.; supervision, A.J.B., X.P.S., M.A., E.M.S., J.S., Á.G.M.; funding acquisition, Á.G.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by of the Ministry of Science and Innovation of Spain, project number RTI2018-097633-A-I00, ‘Photonic restoration applied to cultural heritage: Application to Dali’s painting: Two Figures’.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The authors wish to thank Miopía FX, GVAM SL, Voxel School of Digital Arts and its students, as well as the Patronato de la Alhambra y El Generalife de Granada (Alhambra and Generalife Trust).

**Conflicts of Interest:** The authors have no conflict of interest to declare. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. The Athens Charter for the Restoration of Historic Monuments—1931—International Council on Monuments and Sites. Available online: <https://www.icomos.org/en/resources/charters-and-texts/179-articles-en-francais/ressources/charters-and-standards/167-the-athens-charter-for-the-restoration-of-historic-monuments> (accessed on 6 December 2021).
2. Icomos. The Venice Charter. 1964. Available online: <https://www.icomos.org/venicecharter2004/> (accessed on 12 December 2021).
3. Buill, F.; Jordana, J.J.R.; Andrés, M.A.N. *Fotogrametría Arquitectónica*; Universitat Politecnica de Catalunya: Barcelona, Spain, 2008.
4. Declaration of Rome (1983)—International Council on Monuments and Sites. Available online: <https://www.icomos.org/en/resources/charters-and-texts/179-articles-en-francais/ressources/charters-and-standards/185-declaration-of-rome> (accessed on 6 December 2021).
5. Icomos. Charter for the Conservation of Historic Towns and Urban Areas (The Washington Charter). 1987. Available online: [https://www.icomos.org/charters/towns\\_e.pdf](https://www.icomos.org/charters/towns_e.pdf) (accessed on 12 December 2021).
6. Negro Ponte, N. *El Mundo Digital*; Ediciones B: Barcelona, Spain, 1997.
7. Santagati, C.; Lo Turco, M.; Bocconcino, M.M.; Donato, V.; Galizia, M. 3D Models for All: Low-Cost Acquisition through Mobile Devices in Comparison with Image Based Techniques. Potentialities and Weaknesses in Cultural Heritage Domain. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W8*, 221. [[CrossRef](#)]
8. Remondino, F.; Rizzi, A. Reality-based 3D documentation of natural and cultural heritage sites—Techniques, problems, and examples. *Appl. Geomat.* **2010**, *2*, 85. [[CrossRef](#)]
9. Banfi, F.; Brumana, R.; Stanga, C. A Content-Based Immersive Experience of Basilica of Sant’ambrogio in Milan: From 3d Survey to Virtual Reality. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W11*, 159. [[CrossRef](#)]
10. El-Said, O.; Aziz, H. Virtual Tours a Means to an End: An Analysis of Virtual Tours’ Role in Tourism Recovery Post COVID-19. *J. Travel Res.* **2021**, 0047287521997567. [[CrossRef](#)]
11. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. ‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300. [[CrossRef](#)]
12. Bustillo, A.; Alaguero, M.; Miguel, I.; Saiz, J.M.; Iglesias, L.S. A flexible platform for the creation of 3D semi-immersive environments to teach Cultural Heritage. *Digit. Appl. Archaeol. Cult. Herit.* **2015**, *2*, 248. [[CrossRef](#)]
13. Inzerillo, L. smart SfM: Salinas Archaeological Museum. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W5*, 369. [[CrossRef](#)]
14. Zita Sampaio, A.; Gomes, A.M.; Sánchez-Lite, A.; Zulueta, P.; González-Gaya, C. Analysis of BIM Methodology Applied to Practical Cases in the Preservation of Heritage Buildings. *Sustainability* **2021**, *13*, 3129. [[CrossRef](#)]
15. Angulo-Fornos, R.; Castellano-Román, M. HBIM as Support of Preventive Conservation Actions in Heritage Architecture. Experience of the Renaissance Quadrant Façade of the Cathedral of Seville. *Appl. Sci.* **2020**, *10*, 2428. [[CrossRef](#)]
16. Murphy, M.; McGovern, E.; Pavia, S. Historic Building Information Modelling—Adding intelligence to laser and image based surveys of European classical architecture. *J. Photogramm. Remote Sens.* **2013**, *76*, 89. [[CrossRef](#)]
17. Pybus, C.; Graham, K.; Doherty, J.; Arellano, N.; Fai, S. New Realities for Canada’S Parliament: a Workflow for Preparing Heritage Bim for Game Engines and Virtual Reality. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W15*, 945. [[CrossRef](#)]
18. Al-Muqdad, F. Assessing the Potentials of Heritage Building Information Modelling (Hbim) in Damaged Heritage Reconstruction. In Proceedings of the 36th Annual ARCOM Conference, Leeds, UK, 7–8 September 2020; Scott, L., Neilson, C.J., Eds.; ARCOM: Leeds, UK, 2020; pp. 25–35.
19. Benitez, A.J.; Sánchez Cid, M.; Armenteros, M. Fotogrametría en el audiovisual. Perspectivas de evolución. In *Los Nuevos Retos en Ciencias Sociales, Artes y Humanidades*; Gedisa: Barcelona, Spain, 2018; pp. 31–45.
20. Roussou, M.; Drettakis, G. Photorealism and Non-Photorealism in Virtual Heritage Representation. In Proceedings of the 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, Brighton, UK, 5–7 November 2003. [[CrossRef](#)]
21. Pavlidis, G.; Tsiafakis, D.; Koutsoudis, A.; Arnaoutoglou, F.; Tsioukas, V.; Chamzas, C. Preservation of Architectural Heritage through 3D Digitization. *Int. J. Archit. Comput.* **2007**, *5*, 221. [[CrossRef](#)]

22. Slater, M.; Wilbur, S. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* **1997**, *6*, 603. [CrossRef]
23. Shashi, M.; Jain, K. Use of Photogrammetry in 3D Modeling and Visualization of Buildings. *ARPN J. Eng. Appl. Sci.* **2007**, *2*, 37.
24. Eisenstein, S.M. *Teoría y Técnica Cinematográficas*; Ediciones Rialp: Madrid, Spain, 1989.
25. Irving, W. Manuscript of The Chronicle of the Conquest of Granada: Isabella Stewart Gardner Museum. Available online: <https://www.gardnermuseum.org/experience/collection/28622> (accessed on 4 December 2021).
26. Martínez de la Rosa, F. Doña Isabel de Solís, Reina de Granada. 1837. Available online: <http://www.cervantesvirtual.com/obra-visor/dona-isabel-de-solis-reyna-de-granada-novela-historica--0/html/> (accessed on 5 November 2021).
27. Marques, S. Mixed media object: Large and fragile structure—Victoria and Albert Museum. Available online: <http://www.vam.ac.uk/content/journals/conservation-journal/issue-48/mixed-media-object-large-and-fragile-structure/> (accessed on 6 December 2021).
28. Milgram, P.; Kishino, F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Inf. Syst.* **1994**, *E77-D*, 1–15.
29. Apollonio, F.I.; Fantini, F.; Garagnani, S.; Gaiani, M. A Photogrammetry-Based Workflow for the Accurate 3D Construction and Visualization of Museums Assets. *Remote Sens.* **2021**, *13*, 486. [CrossRef]
30. Dhanda, A.; Reina Ortiz, M.; Weigert, A.; Paladini, A.; Min, A.; Gyi, M.; Su, S.; Fai, S.; Santana Quintero, M. Recreating Cultural Heritage Environments for Vr Using Photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W9*, 305. [CrossRef]
31. Tucci, G.; Bonora, V.; Conti, A.; Fiorini, L. Digital Workflow for the Acquisition and Elaboration of 3D Data in a Monumental Complex: The Fortress of Saint John the Baptist in Florence. In Proceedings of the 26th International CIPA Symposium 2017, Ottawa, ON, Canada; 2017; pp. 679–686.
32. Moyano, J.; Nieto-Julián, J.E.; Bienvenido-Huertas, D.; Marín-García, D. Validation of Close-Range Photogrammetry for Architectural and Archaeological Heritage: Analysis of Point Density and 3D Mesh Geometry. *Remote Sens.* **2020**, *12*, 3571. [CrossRef]
33. Rahaman, H.; Champion, E. To 3D or Not 3D: Choosing a Photogrammetry Workflow for Cultural Heritage Groups. *Heritage* **2019**, *2*, 112. [CrossRef]
34. Furukawa, Y.; Ponce, J. Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1362. [CrossRef]
35. Rehany, N.; Barsi, A.; Lovas, T. Capturing Fine Details Involving Low-Cost Sensors –a Comparative Study. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W8*, 213. [CrossRef]
36. Khalloufi, H.; Azough, A.; Ennahahi, N.; Kaghat, F.Z. Low-Cost Terrestrial Photogrammetry for 3D Modeling of Historic Sites: A Case Study of the Marinids’ Royal Necropolis City of Fez, Morocco. *Mediterr. Archaeol. Archaeom.* **2020**, *2*, 257–272. [CrossRef]
37. 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.* **2017**, *XLII-2/W8*, 1. [CrossRef]
38. Almendros, N. *Días de Una Cámara*; Seix Barral: Barcelona, Spain, 1990.
39. Mueller, P.; Vereenooghe, T.; Vergauwen, M.; Gool, L.V.; Waelkens, M. Photo-Realistic and Detailed 3D Modeling: The Antonine Nymphaeum at Sagalassos (Turkey). In Proceedings of the 32nd Conference on Computer Applications and Quantitative Methods to Archeology. CAA 2004, Prato, Italy, 13–17 April 2004; pp. 1–10.
40. Ferwerda, J.A. Three varieties of realism in computer graphics. In Proceedings of the Electronic Imaging VIII, 2003, Santa Clara, CA, USA, 17 June 2003. [CrossRef]
41. Tang, R.; Fritsch, D. Correlation Analysis of Camera Self-Calibration in Close Range Photogrammetry. *Photogramm. Rec.* **2013**, *28*, 86. [CrossRef]
42. Fraser, C.S. Network design considerations for non-topographic photogrammetry. *Photogramm. Eng. Remote Sens.* **1984**, *50*, 1115–1126.
43. James, M.R.; Chandler, J.H.; Eltner, A.; Fraser, C.; Miller, P.E.; Mills, J.P.; Noble, T.; Robson, S.; Lane, S.N. Guidelines on the use of structure-from-motion photogrammetry in geomorphic research. *Earth Surf. Processes Landf.* **2019**, *44*, 2081. [CrossRef]
44. Benítez, A.J.; Jiménez, I.; Sánchez Cid, M. Conceptos de Postproducción. In *Una Perspectiva Contemporánea*; Utray Armenteros, M.F., Benítez, A.J., Eds.; Postproducción digital: Madrid, Spain, 2015; p. 13.
45. Gaiani, M.; Apollonio, F.; Ballabeni, A.; Remondino, F. Securing Color Fidelity in 3D Architectural Heritage Scenarios. *Sensors* **2017**, *17*, 2437. [CrossRef] [PubMed]
46. Martínez Rodríguez, S. *Evaluación en Campo de Una Nueva Técnica Fotogramétrica de Rango Cercano y Bajo Coste Basada en la Correlación Automática de Imágenes*; U de Santiago de Compostela Campus de Lugo: Lugo, Spain, 2014. Available online: <https://minerva.usc.es/xmlui/handle/10347/12884> (accessed on 20 November 2021).
47. Diaz-Vilariño, L.; Frías, E.; Balado, J.; González-Jorge, H. Scan Planning and Route Optimization for Control of Execution of as-Designed Bim. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2018**, *XLII-4*, 143. [CrossRef]
48. Rea, P.W.; Irving, D.K. *Producción y Dirección de Cortometrajes y Vídeos*; Instituto Oficial de Radio y Televisión: Madrid, Spain, 2002.

- 
49. Millerson, G.; Owens, J. *Television Production*, 14th ed.; Focal Press/Elsevier: Amsterdam, The Netherlands; Boston, MA, USA, 2009; p. 399.
  50. Rabiger, M. *Dirección Cinematográfica: Técnica y Estética*; Ediciones Omega: Barcelona, Spain, 2009.
  51. Kraaman. The Best Encoding Settings for Your 4k 360 3D VR Videos. Available online: <https://headjack.io/blog/best-encoding-settings-resolution-for-4k-360-3d-vr-videos/> (accessed on 20 November 2021).