

## ORIGINAL PAPER

## Low-cost 3D reconstruction of cultural heritage artifacts

Pedro O. Raimundo<sup>1</sup>, Karl Apaza-Agüero<sup>1</sup> and Antonio L. Apolinário Jr.<sup>1</sup><sup>1</sup>Department of Computer Science, Federal University of Bahia (UFBA)

\*praimundo@ufba.br; kaguero@ufba.br ; antonio.apolinario@ufba.br

Received: 12/31/2017. Revised: 03/23/2018. Accepted: 04/16/2018.

### Abstract

3D reconstruction methods can be used to obtain digital models by capturing the shape and appearance of real objects. Due to both scientific and technological advances, low-cost 3D reconstruction is now widely used to reconstruct objects of reasonable geometric complexity using cheap acquisition hardware; however, it is unclear whether or not these devices produce suitable results to preserve cultural heritage artifacts. First, this paper goes over the general concept of 3D reconstruction and the main differences between low-cost and general-purpose 3D reconstruction pipelines, these differences are illustrated using a prototypical low-cost pipeline that has been used to reconstruct heritage artifacts. Then, we present the results of a survey conducted in order to investigate how existing low-cost 3D reconstruction approaches attempted to improve the results of their reconstructions, with an emphasis on the realistic rendering techniques they have used. Throughout the survey process, a categorization for low-cost 3D methodologies was proposed based on the scanning apparatus used by the approaches and other similarities between them. Finally, based on the surveyed studies and the results obtained with the prototypical pipeline, we conclude that it is possible to preserve heritage artifacts using low-cost approaches, granted that modern realism techniques are used to improve the appearance of the models.

**Key words:** Low-cost 3D reconstruction; cultural heritage; digital heritage; survey.

### Resumo

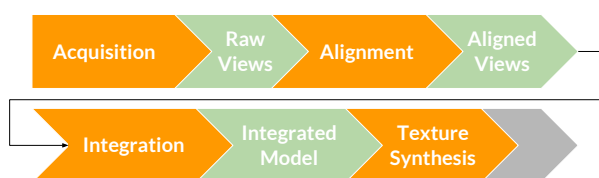
Os métodos de reconstrução 3D permitem a obtenção de modelos digitais através da captura da informação de forma e aparência dos objetos. Devido a avanços científicos e tecnológicos, técnicas de reconstrução 3D de baixo custo se tornaram amplamente utilizadas para reconstruir objetos com geometria razoavelmente complexa usando dispositivos de aquisição de baixo custo; no entanto, ainda é discutível a aplicabilidade desses dispositivos de baixo custo para preservação de artefatos culturais. Primeiramente, esse trabalho introduz a área de reconstrução 3D e discute as principais diferenças entre metodologias de reconstrução de baixo custo e metodologias de reconstrução de propósito geral, essas diferenças são ilustradas ao ser apresentado um protótipo de metodologia de baixo custo que foi utilizado para reconstrução de artefatos culturais. Posteriormente, são apresentados os resultados de um levantamento realizado a fim de investigar como as propostas existentes de reconstrução de baixo custo visaram melhorar a qualidade das reconstruções geradas, com ênfase nas técnicas de renderização realista utilizadas. Ao longo da realização do levantamento, foi proposta uma categorização dos estudos investigados levando em consideração o aparato de escaneamento utilizado e outras similaridades. Ao fim, com base nos estudos levantados e nos resultados obtidos usando a metodologia apresentada, conclui-se que é possível a utilização de metodologias de reconstrução de baixo custo para preservação de artefatos culturais, uma vez que sejam utilizadas técnicas de realismo para melhorar a fidelidade visual dos modelos obtidos.

**Palavras-Chave:** Levantamento; patrimônio cultural; preservação digital; reconstrução 3D de baixo custo.

## 1 Introduction

3D reconstruction is the process of capturing geometry and appearance from real objects to obtain a digital model. The current 3D reconstruction approaches usually employ a well-defined pipeline to obtain a polygonal mesh from the data collected by one or more acquisition methods (Pavlidis et al.; 2007).

Many different areas such as engineering, architecture, medical imaging and cinema take advantage of 3D reconstruction techniques; however, one of the oldest and most common applications of 3D reconstruction is the preservation of cultural heritage (Levoy et al.; 2000; El-Hakim et al.; 2004; Pavlidis et al.; 2007; Gomes et al.; 2014). The benefits of digitally preserving important pieces in this manner are manifold: a) digitally-preserved artifacts are immune to the effects of time, weather conditions and other agents; b) scholars are able to study these artifacts from up close without incurring the risk of damaging the original pieces; and c) digitally preserved artifacts can easily be shared with other scholars, universities and the general public (Gomes et al.; 2014; Wu-Wei; 2016). Wu-Wei (Wu-Wei; 2016) argues that in addition to preserving the *ground truth* of the artifacts, digital preservation of heritage helps transmit the culture and values of our communities to future generations; this, in turn, drives up the accuracy and realism requirements. The broad field that studies how to digitally preserve heritage artifacts and sites is often called *digital heritage*.



**Figure 1:** Steps of the commonplace 3D reconstruction pipeline. The orange blocks depict steps of the pipeline, the light green blocks depict the output of the preceding step of the pipeline, which serves as input for the next step, the gray block denotes the end of the pipeline where a textured 3D model is obtained.

The most common 3D reconstruction pipeline comprises four steps (Figure 1): a) acquisition of information, b) alignment of the information obtained from different angles, c) integration of the information into a single model and d) generation of the textured 3D model (Bernardini and Rushmeier; 2002; Gomes et al.; 2014). Given this general pipeline, several works add, remove or specialize some steps to better fit their domain (Berrier et al.; 2015; Nöll et al.; 2015; Zeppelzauer et al.; 2015). This survey focuses on several existing 3D reconstruction pipelines that utilize acquisition devices which usually cost up to several hundred dollars (Newcombe et al.; 2011; Dias et al.; 2006), instead of the traditional high-end 3D scanners which are much more expensive. Some advantages of these approaches are: the possibility

of real-time reconstruction, compatibility with GPU acceleration, and ease of acquisition due to the decreased bulk and weight of low-cost 3D scanners. For example, the KinectFusion (Newcombe et al.; 2011; Izadi et al.; 2011) system provides 3D surface reconstruction in real-time using only a Microsoft Kinect sensor.

This work presents low-cost reconstruction methods in the context of digital heritage. These low-cost approaches have been surveyed and categorized, and the advanced realism techniques that they used to increase the realism of the generated 3D reconstructions were tallied.

This work is structured as follows: Section 2 presents the area of 3D reconstruction from the perspective of several important digital heritage projects, and a prototypical low-cost 3D reconstruction pipeline used to digitize heritage artifacts; Section 3 presents and discusses the low-cost 3D reconstruction studies surveyed in the literature and their efforts to increase the realism of their reconstructions; and Section 4 presents our conclusions.

## 2 3D Reconstruction

Although there are older works, the Digital Michelangelo project (Levoy et al.; 2000) is considered one of the most important works in the field of digital heritage. This is a key work in this field because of its technical accomplishments, quality of reconstruction and cultural importance of the reconstructed artifacts. In the corresponding paper, the authors discussed their reconstruction workflow and pipeline, the challenges faced during the reconstruction procedure, and the solutions they devised to scan a grand total of ten Michelangelo statues, two building interiors and 1163 fragments of an ancient marble map.

Two years later, Bernardini and Rushmeier (Bernardini and Rushmeier; 2002) published a study where they survey several 3D reconstruction efforts and formalize a generalized pipeline consisting of acquisition, alignment, integration and texturing; which is the basis of even modern reconstruction efforts. Their pipeline considers range images as the input and 3D models as the output of the 3D reconstruction process; in this context, a 3D model is defined as a “numerical description of an object that can be used to render images of the object from arbitrary viewpoints and under arbitrary lighting conditions”.

While Bernardini and Rushmeier focused on 3D reconstruction using either triangulation laser scanners or time-of-flight systems, these are not the only options to acquire range information. Although laser scanners and time-of-flight systems are still the norm in digital heritage (Gomes et al.; 2014), some works use different acquisition technology, as seen on the reconstruction of the Great Buddha of Bamiyan, Afghanistan. Grün et al. (Grün et al.; 2004) used only still images to perform the reconstruction of the Great Buddha of Bamiyan, a 53-meter representation of Buddha which was completely destroyed in 2004. The 3D reconstructed model was then used to build a museum statue at one-tenth

scale of the original, making this work a testament to the importance and impact of digital heritage.

Various properties of regular pictures can be leveraged to obtain structural information for 3D reconstruction (Pavlidis et al.; 2007). Some popular approaches include using pairwise stereoscopy (Koutsoudis et al.; 2007), the silhouettes of the objects (Brusco et al.; 2005) or the estimate of motion between pictures (Lari et al.; 2015; Manferdini; 2012). Image-based 3D reconstruction is usually more accessible in terms of necessary hardware than scanner-based approaches, due to the wide availability and low price of digital cameras. However, image-based reconstructions usually are not as accurate as their scanner-based counterparts, and require a lot of pre- and post-processing to produce good results.

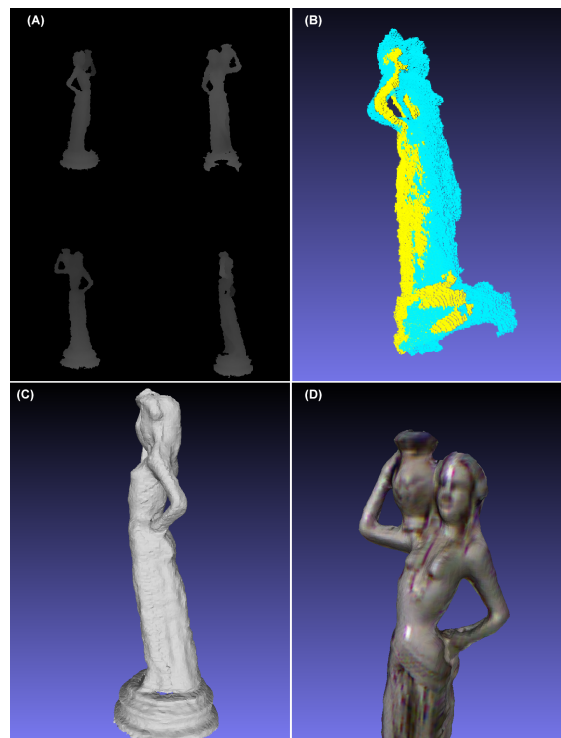
In this work, *low-cost* refers to the Kinect and other consumer-grade sensors which usually cost several hundred dollars, which belong in a separate category than that of the traditional laser-triangulation 3D scanners commonly used in digital heritage, with prices ranging from several thousand dollars to over one hundred thousand dollars<sup>1</sup>.

In 2010, Microsoft launched the Kinect sensor, a cheap motion-tracking device that works similarly to a time-of-flight rangefinder, but uses structured infrared light to acquire range information. Despite being initially intended for games, the Kinect has been widely used in research and other areas (Zhang; 2012). One of the most groundbreaking applications of the Kinect was the KinectFusion system (Newcombe et al.; 2011), a low-cost 3D reconstruction application that spurred a trend of 3D reconstruction research using it and other low-cost sensors (Zhang; 2012). To highlight the differences between low-cost and general purpose 3D reconstruction projects, the next subsection discusses the phases of a pipeline for low-cost digital heritage, its application on the reconstruction of heritage artifacts and some of its discerning characteristics.

### Low-cost 3D Reconstruction

Before the reconstruction process itself, low-cost reconstruction projects have to consider the limitations of the scanning equipment utilized and how they impact the possibilities and the objects that can be reconstructed. In the proposed pipeline, the main limitations of the acquisition device (Microsoft Kinect) are its operational range (0.3 m to 10 m<sup>2</sup>) and resolution (640x480 pixels). Coupled with the quality requirements of digital heritage, these limitations have caused the proposed pipeline to be most applicable to medium-scale objects such as pots, pans, figurines, busts and small statues.

The acquisition phase of the reconstructions consisted of using a simple application to retrieve the raw depth images from the Kinect scanner, which contrasts with the specialized software



**Figure 2:** Visual representations of each phase of the proposed low-cost 3D reconstruction pipeline. (a) Depth images acquired with the Microsoft Kinect sensor. (b) Alignment of depth information after converting the images to point clouds. (c) Integrated mesh with shape information from multiple views. (d) Final reconstructed model with texture information acquired with the Kinect RGB camera.

usually needed by high-end 3D scanning apparatus. Capturing the raw data from the Kinect provides the pipeline with more flexibility in terms of how this data is processed and handled by the other steps of the pipeline. Figure 2.A shows examples of depth images captured in this step, the statue depicted was rotated in steps of approximately 30 degrees between the captures, using a rotary base.

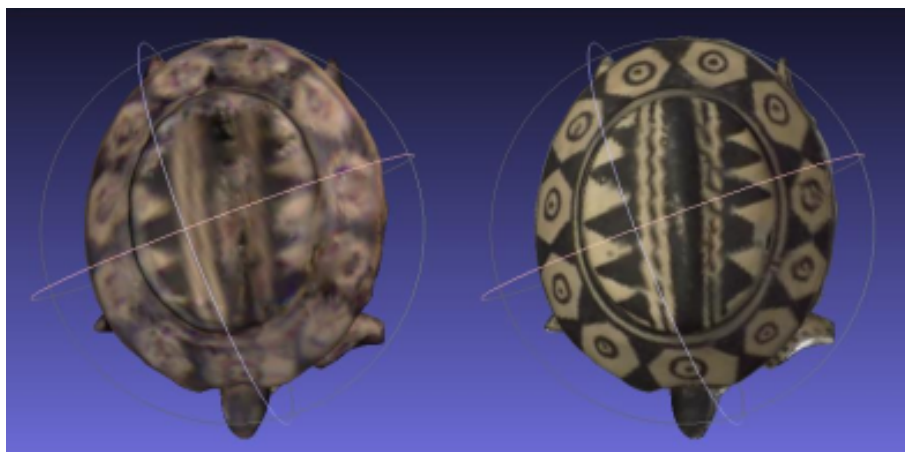
Figure 2.B shows part of the alignment phase of the proposed pipeline. Meshlab (Cignoni et al.; 2008), a popular tool within the computer graphics research community, was used to align the information from multiple views. The integration of multiple views and the generation of a complete 3D model were also performed in MeshLab (Cignoni et al.; 2008), using the screened Poisson surface reconstruction (Kazhdan and Hoppe; 2013) technique. The results of the reconstruction with and without the color information obtained from the Kinect can be seen in Figures 2.C and 2.D. The alignment and integration steps of the proposed pipeline do not initially deviate from what is common in general purpose pipelines. Nonetheless, only a subset of the commonly used techniques for these steps yield good results with low-cost acquisition devices, because fewer assumptions can be made with regards to the quality and other characteristics of the data provided by these devices.

The color information captured by the Kinect aids the manual part of the alignment phase

<sup>1</sup>An updated price list with current high-end scanners can be found at: [http://www.dirdim.com/prod\\_laserscanners.htm](http://www.dirdim.com/prod_laserscanners.htm)

<sup>2</sup>While the official values for minimum and maximum range provided by microsoft are 800 mm and 4000 mm, we have determined that the sensor operates outside these limits, although with a loss of accuracy





**Figure 3:** Indigenous turtle-shaped pan, black ink on natural-colored clay. The leftmost render shows the model and texture information captured with a Kinect; the rightmost render presents the same model, but with higher-resolution color information. This artifact is part of the Aristóteles Barcelos collection of the UFBA Museum of Archeology and Ethnology, it originated in the Xingu region and was made by members of the Waujá ethnicity around the year 2000.

of the pipeline; however, as shown in Figure 3 (left) the results obtained from using only this color information are somewhat blurry and do not accurately represent the appearance of the artifact. In contrast, Figure 3 (right) shows the increase in visual fidelity obtained by acquiring higher-resolution pictures and mapping their color information on the model. Thus, it is safe to assume that capturing and using high-resolution color information is important in low-cost digital heritage, because it helps convey the appearance of details that might not be present in the actual geometry of the reconstruction.

### 3 A Survey of Low-cost 3D Reconstruction Approaches

This study surveyed 3D reconstruction studies that adopted a low-cost reconstruction pipeline and either directly or indirectly relate to digital heritage. The studies were organized in three broad categories based on the methods and equipment used to perform the reconstruction: scanner-based approaches, image-based approaches, and topographical approaches. This categorization is specific to the domain of low-cost reconstruction and does not aim to encompass every 3D reconstruction study. Works that combine characteristics from one or more categories are discussed individually.

Provided that most low-cost 3D reconstruction studies focus their efforts on producing a final model that is accurate and detailed enough for their application, few works present advanced rendering and realism techniques. Notwithstanding, for the studies that present or suggest such techniques, they are highlighted and briefly explained in addition to the reconstruction approach proposed therein.

#### Scanner-Based

Several 3D reconstruction approaches use active devices, that is, those that interfere with the object via a beam of light and capture the result of this

interaction to calculate depth or other information. A common trait of this category of approaches is that the acquisition equipment (cameras, emitters, etc.) must be calibrated before the reconstruction and their main benefit is that the reconstructions are generally accurate. Three studies utilized this kind of approach (Rocchini et al.; 2001; Silva et al.; 2013; Banerjee et al.; 2013), with varying levels of quality and performance as the processing power available evolved.

Rocchini et al. (Rocchini et al.; 2001) laid the cornerstone of approaches based on beams of structured light, with the low-cost of the reconstruction process as one of their main concerns. Using only consumer technology available in 2001 (i.e. a regular projector, and a digital camera), they managed to reconstruct a bronze statue with a resolution of 0.7 and 1.4 millimeters in the X and Y axis respectively, 0.3 millimeters of accuracy and RGB color information with 1 byte per channel. However, this reconstruction system has some limitations, the main one being that the captured object must remain static between the capture of multiple views.

Other studies expand upon this approach by utilizing light beams outside of the visible spectrum. The KinectFusion system (Newcombe et al.; 2011) uses the Microsoft Kinect sensor to provide real-time 3D reconstructions. Viewing reconstruction results in real-time improves the reconstruction pipeline considerably; for example: the person operating the scanner becomes can quickly adjust for shadowed and poorly-captured portions of the object during a real-time reconstruction. Silva et al. (Silva et al.; 2013) expanded upon an open-source implementation of KinectFusion to provide users with a guided reconstruction process, one of their contributions was the use of super-resolution techniques (Richardt et al.; 2012) to increase the quality of the low-resolution data captured with the Kinect. A more recent study (Banerjee et al.; 2013) built a fully-automatic angular laser scanner for less than a thousand dollars which provides sub-millimeter resolution. Both works achieved promising results

in terms of accuracy and visual fidelity.

### Image-Based

We categorize the approaches that use acquired images of the objects without strictly relying on a patterned beam of light as *image-based*. These approaches became more popular as image-capturing equipment and processing power became more accessible; however, they face several additional issues due to the nature of the captured images; one such issue is separating the object of interest from the background of the images, which is much harder to do for image-based approaches than for scanner-based approaches, where it is easier to threshold the range information. Several properties of the captured images can be utilized to retrieve the information of the object as highlighted by the variation in techniques between the studies that belong to this category.

The most popular approach to obtain shape information of objects from still images is through stereoscopy (or shape from stereo (Pavlidis et al.; 2007)). Pedersini et al. (Pedersini et al.; 2000) developed a self-calibrating stereo reconstruction system based on a multiple-camera setup, which also automatically integrated the pieces obtained from the multiple views; the reconstructed model preserved even small details of the object. On the other hand, Koutsoudis et al. (Koutsoudis et al.; 2007) adopted an approach that sacrificed the details of the reconstruction in order to easily capture large-scale structures such as buildings and ancient sites, arguing that even lower-quality reconstructions are satisfactory to promote cultural artifacts. Boochs et al. (Boochs et al.; 2007) attempted to make the reconstruction and survey process easier by taking advantage of the newly available (at the time) high-resolution cameras, while preserving accuracy. Redmondino (Remondino; 2007) presented experimental results which indicate that the deviation from ground truth of image-based approaches in general varies between 1 and 3 millimeters.

Several authors proposed improvements to the shape from stereo reconstruction pipeline. Volgiatzis and Hernández (Vogiatzis and Hernández; 2010) lit the objects with three colored light sources and employed stereo-vision techniques to ultimately present a passive 3D reconstruction system that is robust to shadowing, deformation of the captured object, and movement of both scene elements and acquisition equipment. Their experimental results show that, with their improvements, it was possible to reconstruct Qing-Dynasty vases with details as small as 0.5 millimeters; however, this approach depends on actively projecting light from three sources on the object, which is not always possible. Taking a different road, Stentoumis et al. (Stentoumis et al.; 2013) devised a user-assisted method to separate the object of interest from the background and improve the accuracy of the reconstructions, but did not provide numerical measures of accuracy, simply stating that the results were “highly detailed”.

Sormann et al. (Sormann et al.; 2005) attempted to retrieve silhouette information from the captured

images, this approach is commonly called shape from silhouette (Pavlidis et al.; 2007). This study obtained interesting results with considerable robustness to noise and variations in the background of the image; however, the accuracy of the reconstruction was slightly compromised. This work also proposed an extension to image-based reconstruction pipelines, making texture synthesis more robust to noise and occlusion in the source images, specially those taken outside of controlled environments, thus improving the realism of the reconstruction.

Both shape from stereo and shape from silhouette techniques were used by Brusco et al. (Brusco et al.; 2005) in an image-based 3D reconstruction system; in doing so, they hoped to improve the reconstruction quality of image-based systems, and ultimately achieved 0.8mm accuracy using only low-resolution (1024x768) pictures. In addition, they proposed a novel technique for texture synthesis – DWT (discrete wavelet transform) stitching. DWT stitching uses transformations in the discrete wavelet domain to allow multiple textures to be mapped onto a model, while avoiding both the blurring that regular interpolation techniques incur, and the discontinuities (seams) that are observed in commonplace multi-texture mapping techniques.

With the even greater advances in computing power of the last years, a new approach became popular: shape from motion (Pavlidis et al.; 2007), which is present in several works (Manferdini; 2012; Makantasis et al.; 2014; Lari et al.; 2015). This approach generally relies on a large amount of still images with considerable overlap (or a video) and works by tracking points of interest across different images of a set.

Manferdini (Manferdini; 2012) set out to digitally preserve cultural artifacts in the city of Veleia by using shape from motion techniques; despite arguing that accuracy and detail are secondary for promotional applications of 3D reconstruction, her work presented good results due to the realism techniques utilized. One such technique comprised of processing the pictures taken in an uncontrolled environment to accurately calibrate their colors, which is easier than adjusting the real lighting of the scene in most scenarios. Other realism techniques used in (Manferdini; 2012) were the capture and synthesis of the surrounding scene (as initially discussed by El-Hakim et al. (El-Hakim et al.; 2004)), and the addition of *level of detail* (Lindstrom et al.; 1996), which allows applications to balance visual fidelity and performance as needed.

Makantasis (Makantasis et al.; 2014) et al. took advantage of shape from motion techniques, together with their novel data-filtering algorithm, to reconstruct a large heritage building using only images obtained from the Internet; however, they present only a point cloud, instead of producing a polygonal mesh. With the recent availability of consumer-grade unmanned aerial vehicles (UAV), one study (Lari et al.; 2015) has already set out to use images from UAVs to reconstruct an entire building using aerial videos and pictures, which is a good application of shape from motion techniques because of the large overlap that invariably happens between images taken from great heights.

Some works are close to being purely

photogrammetric; that is, they rely almost exclusively on inferring real-world measurements from images of the reconstructed object instead of directly using image-processing routines (Mikhail et al.; 2001). Despite photogrammetry being an old discipline, used for purposes other than 3D reconstruction, some studies (Covas et al.; 2015; Louiset et al.; 2016; Utomo and Wibowo; 2017) that fall under this category present interesting advances. For example, Covas et al. (Covas et al.; 2015) used fish-eye images to reconstruct hard-to-reach parts of two medieval castles in Portugal. The authors reported that the usage of fish-eye images made the acquisition process quicker, that it was easier to connect different parts of the surveyed location, and that it was possible to acquire images that would otherwise demand unmanned aerial vehicles.

Conversely, Louiset et al. (Louiset et al.; 2016) devised a reconstruction method using a miniature unmanned aerial autonomously flying in close proximity of the reconstructed object. In this approach, the vehicle flies twice around the object once in a circular trajectory and a second time with a trajectory that approximates the shape of the object, with the results of the second flight being used to improve the realism of the initial reconstruction. The presented results show that this approach performed better than those with a human guided acquisition phase in terms of surface coverage, but significantly worse in terms of accuracy, with a deviation of about 2 mm from ground truth. More recently, Utomo and Wibowo (Utomo and Wibowo; 2017) attempted to implement a photogrammetry solution using only open-source algorithms and software. This implementation produced results comparable to commercial photogrammetry solutions in terms of accuracy, whilst being completely free, although its performance w.

### Topographical

This category of approaches comprises the works that used information acquired from maps, total stations, treatises and other topographical assets to guide the reconstruction of their objects of interest. Isoda et al. (Isoda et al.; 2009) and Ragia et al. (Ragia et al.; 2015) utilized topographic information to obtain a rough outline of the objects from a set of few, but very accurate, points of their surface. Isoda et al. (Isoda et al.; 2009) used several maps and *a priori* knowledge of the architecture of Kyoto during the Edo period to produce a satisfactory representation of what the city looked like in the past. The city was reconstructed by customizing pre-established parametric models of the common buildings of the era in accordance with the map data.

More specifically, Ragia et al. (Ragia et al.; 2015) captured about 780 points of the large Neoria historical buildings in Crete, but ultimately produced good models via human-guided refinement, and an interactive visualization using state-of-art realism techniques. In their interactive tool, a user could simulate restoration efforts on the historical building and see the results in real time. This is a good example of how visualization techniques can drastically improve the appearance and usefulness of models with simple geometry. Because it is

more recent, this study utilized several realism techniques, including normal mapping and global illumination. Normal mapping (Blinn; 1978) is a popular technique to increase the realism of 3D models by slightly perturbing the normals of a surface before using it in lighting calculations, while global illumination (Dorsey; 1995) consists of calculating how light interacts with all the objects of a scene, instead of using simpler approximations

### Hybrid

The last category of approaches contains the studies that rely on the combination of techniques belonging to two of more categories. These works aimed to combine techniques in a way that improved the reconstruction results, while working around the intrinsic limitations of each approach.

El-Hakim et al. (El-Hakim et al.; 2004) discussed reconstruction methods in great length and presented a hybrid approach, combining laser scanning and photogrammetric techniques to reconstruct a large-scale structure, the Abbey of Pomposa. They used photogrammetry techniques to acquire the overall structure of the building and laser scanners to acquire fine details, making theirs a multi-resolution approach. This paper also suggests that immediate and distant surroundings can be captured to produce realistic scenes together with the reconstructed models.

An extension of (El-Hakim et al.; 2004) was published by El-Hakim et al. just a year later (El-Hakim et al.; 2005). The new approach expanded upon the multi-resolution trait of the original by using information from computer-assisted drawings, stereo-vision techniques, photogrammetry and laser scanning in the reconstructions. A hierarchy in terms of detail, accuracy and reliability of these sources is proposed. This hierarchy is then used to choose how and where each capture will be used in the reconstruction. The results presented are very good in terms of image quality and accuracy.

A low-cost 3D reconstruction workflow based on open-source software was proposed by Koutsoudis et al. (Koutsoudis et al.; 2008). They employed shape from stereo techniques, empirical measurements of the structures and topographical maps to reconstruct the city of Kavala, located in North-Eastern Greece. Even though they agree that lower-quality reconstructions are sufficient for the promotion of historical localities, they pose that good results were achieved by using realism techniques such as normal mapping and per-vertex illumination. Koutsoudis et al. also pointed out that ambient occlusion (Tarini et al.; 2006) and precomputed shadowing (Zhou et al.; 2005) are viable techniques to increase the realism of low-cost 3D reconstructions; however, they did not use those techniques in their approach.

With the advances in processing power, Prokos et al. (Prokos et al.; 2009) proposed a hybrid approach that combines shape from stereo (with additional geometric constraints) and laser-scanning principles. They managed to mimic the behavior of the high-end laser scanners using two web-cams and a hand-held laser beam, ultimately achieving 0.3 mm of accuracy on their reconstructions, which is as good as some high-end laser triangulation scanners.



**Table 1:** Summarized characteristics of each category of low-cost reconstruction methods, according to the papers surveyed in this work.

Category	Reported Accuracy	Reconstruction Rate			Degree of Automation
		offline	interactive	real-time	
Scanner-Based	Up to 0.3 mm	x	x	x	medium-to-high
Image-Based	Up to 0.5 mm	x			low-to-medium
Topographical	N/A	x			low
Hybrid	Up to 0.3 mm	x		x	low-to-medium

As laser scanners became cheaper, Caprioli and Scognamiglio (Caprioli and Scognamiglio; 2009) combined the use of both a topographical survey station and a long-range laser scanner to reconstruct a road bridge that looms over the Lama S. Giorgio river, although it is debatable whether or not the laser scanner they used can be considered low-cost, their setup is described as such.

Finally, several studies utilized the Microsoft Kinect sensor in their hybrid approaches (Dupuis et al.; 2014; Zollhöfer et al.; 2014), each with its own goals. Zollhöfer et al. (Zollhöfer et al.; 2014) used a single Microsoft Kinect sensor, together with topographical information to accurately reconstruct a large excavation site, the topographical information was used to perform a global warp that adjusted the obtained models to match the ground truth; this global warp, together with the super-resolution techniques they applied to the raw Kinect scans significantly increased the fidelity and realism of their results. Another contribution of Zollhöfer et al. is that their reconstruction pipeline does not need costly preprocessing operations before the generated models can be used.

With the advent of other cheap scanners, Dupuis et al. (Dupuis et al.; 2014) attempted to increase the realism of their 3D reconstructions by combining the scans from one such scanner with room-scale captures performed with the Microsoft Kinect sensor, much like the hierarchical approach of El-Hakim et al. (El-Hakim et al.; 2005). However, in this case the correspondence between the two captures is performed autonomously, via support vector machines (SVM) and the iterative-closest point (ICP) algorithm; this dramatically lowers the amount of human interaction needed during the alignment, if compared to previous approaches.

### 3.1 Discussion

Each low-cost reconstruction method has its strengths and weaknesses, knowing which one to use is one of the keys to a successful 3D reconstruction project. With the categorization presented in this survey, it is easier to highlight traits shared by a set of approaches and discuss their advantages.

In terms of reconstruction quality, scanner-based methods provide the best reconstructions in terms of accuracy, detail and reconstruction speed; sometimes allowing real-time construction. However, each scanning device is optimal for a certain type of reconstruction. For example: the Microsoft Kinect sensor can be used for room-scale reconstructions, as shown by Dupuis et al. (Dupuis et al.; 2014), while the scanner proposed by Banerjee et al. (Banerjee

et al.; 2013) was validated with smaller objects, with up to 30 cm in any dimension.

The works of Rocchini et al. (Rocchini et al.; 2001) and Banerjee et al. (Banerjee et al.; 2013) clearly presented the strengths and limitations of scanner-based approaches. Both works achieved results with good outward appearance and up to 0.3 mm of accuracy. The first approach was used to reconstruct large statues from distances of more than 1 meter, while the second one was validated with smaller objects located at approximately 1 meter of the scanner.

Image-based approaches are highly flexible and work well with regular digital cameras, or even previously available images of the subjects. The latter was employed by Makantasis et al. (Makantasis et al.; 2014) who used data-filtering technology to reconstruct a heritage building using only Internet images. The main limitation of image-based approaches is that they need either highly controlled environments or heavy processing to generate accurate reconstructions. For example, the pipeline proposed in (Brusco et al.; 2005) requires additional processing specifically to separate the object of interest from the background of the image, and even then manual intervention is sometimes needed to obtain good results.

Topographical approaches were used to create 3D models with very little input data; however, all of them required significant human interaction. This was the case in the study of Ragia et al. (Ragia et al.; 2015), where they reconstructed a large building using only 780 points as input; their good results were only possible with extensive post-processing and several steps of manual adjustments and refinements of the reconstructed models.

Hybrid reconstruction methods combine characteristics of several others to generate 3D models that are well-suited to a specific purpose. However, in doing so they often become too specialized to be reused in different domains. Zollhöfer et al. (Zollhöfer et al.; 2014) adopted a hybrid approach that combines both scanner-based and topographical principles to quickly reconstruct a complete excavation site. They used a single Kinect sensor to capture the site and took advantage of a topographical map to undistort the reconstructed model according to the ground truth, greatly increasing the quality of the final model; however, this kind of information is not always available.

Table 1 summarizes the main characteristics of the surveyed categories of low-cost 3D reconstruction methods, reiterating several points presented here. It should be noted that while the most accurate image-based approach reported 0.5 mm of accuracy,

**Table 2:** Realism features presented or discussed in the surveyed studies. *None* denotes the number of studies that did not discuss or present any realism features.

Realism Feature	References	Number of Studies
Texture Mapping	(Brusco et al.; 2005), (El-Hakim et al.; 2004), (El-Hakim et al.; 2005), (Isoda et al.; 2009), (Koutsoudis et al.; 2007), (Koutsoudis et al.; 2008), (Lari et al.; 2015), (Manferdini; 2012), (Ragia et al.; 2015), (Remondino; 2007), (Rocchini et al.; 2001), (Silva et al.; 2013), (Sormann et al.; 2005), (Utomo and Wibowo; 2017)	14
None	(Banerjee et al.; 2013), (Boochs et al.; 2007), (Caprioli and Scognamiglio; 2009), (Covas et al.; 2015), (Makantasis et al.; 2014), (Pedersini et al.; 2000), (Prokos et al.; 2009), (Stentoumis et al.; 2013), (Vogiatzis and Hernández; 2010)	9
Multi-resolution Scans	(El-Hakim et al.; 2005), (Dupuis et al.; 2014), (Louiset et al.; 2016)	3
Surroundings	(El-Hakim et al.; 2004), (Manferdini; 2012)	2
Super-resolution	(Silva et al.; 2013), (Zollhöfer et al.; 2014)	2
Normal Mapping	(Koutsoudis et al.; 2008), (Ragia et al.; 2015)	2
Light Calibration	(Manferdini; 2012), (Rocchini et al.; 2001)	2
Shadow Baking	(Koutsoudis et al.; 2008)	1
Ambient Occlusion	(Koutsoudis et al.; 2008)	1
Per-vertex Lighting	(Koutsoudis et al.; 2008)	1
Level of detail	(Manferdini; 2012)	1
Global Illumination	(Ragia et al.; 2015)	1

a previous survey of the area (Remondino; 2007) revealed that these approaches usually provide between 1 and 3 millimeters of precision. It is also noteworthy that none of the topographical approaches discussed the accuracy of their final reconstructed models; however, one work (Ragia et al.; 2015) reported that the data was acquired with a varying accuracy of  $\pm 2$  to  $\pm 5$  mm, which gives an idea of how accurate the reconstructions turned out.

Table 1 also presents the typical reconstruction rate and degree of automation of the surveyed reconstruction methods. *Reconstruction rate* refers to the speed at which 3D models are generated using a given approach; this is more notable in scanner-based and hybrid methods, where even real-time reconstruction is sometimes possible. *Degree of automation* is an inverse scale of how much human interaction is required during the reconstruction process and varies between *low* – where most steps of the pipeline require human intervention or refinement, and *high* – where a significant part of the reconstruction is performed automatically.

A secondary objective of this survey was to uncover which realism techniques have been used to increase the presentation of low-cost 3D reconstructions. Table 2 summarizes our findings in this regard. The main conclusion we draw is that while a considerable number of studies focused solely on the reconstruction itself, more than half the studies use at least texture mapping to improve the appearance of the models, and some studies successfully used advanced rendering techniques – such as global illumination, ambient occlusion, normal mapping, and level of detail – to greatly improve the final presentation of their reconstructions.

## 4 Conclusion

This survey presented studies that help bridge the gap between low-cost 3D reconstruction methods and the

realism requirements of digital heritage applications. We showed the dissonance between these two areas, but also pointed out several successful applications of realism techniques to improve the image quality of low-cost 3D reconstructions. This provides a positive outlook towards the usage of more recent and sophisticated realism techniques to finally allow 3D models obtained by low-cost methods to be used for digital heritage and cultural learning.

## Acknowledgments

The authors would like to thank the National Council for Scientific and Technological Development (CNPq) for partially funding this research via grant 372433/2017-0 under the project 406719/2016-0 (“*Replicação de baixo custo de objetos culturais combinando técnicas de digitalização 3D, impressão 3D e impressão hidrográfica*”).

The authors would also like to thank the entire MAE/UFBA staff for their help in choosing interesting pieces, for providing us with interesting heritage artifacts to work with, and all their support in terms of infrastructure, knowledge and manipulation of the artifacts.

## References

- Banerjee, S., Dutta, S., Biswas, P. K. and Bhowmick, P. (2013). A low-cost portable 3d laser scanning system with aptness from acquisition to visualization, *2013 Digital Heritage International Congress (DigitalHeritage)* 1: 185–188.
- Bernardini, F. and Rushmeier, H. (2002). The 3d model acquisition pipeline, *Computer graphics forum*, Vol. 21, Wiley Online Library, pp. 149–172.
- Berrier, S., Tetzlaff, M., Ludwig, M. and Meyer, G. (2015). Improved appearance rendering for



- photogrammetrically acquired 3d models, *Digital Heritage*, 2015, Vol. 1, IEEE, pp. 255–262.
- Blinn, J. F. (1978). Simulation of wrinkled surfaces, *ACM SIGGRAPH computer graphics*, Vol. 12, ACM, pp. 286–292.
- Boochs, F., Heinz, G., Huxhagen, U. and Müller, H. (2007). Low-cost image based system for nontechnical experts in cultural heritage documentation and analysis, *Proc. of XXI International CIPA Symposium*.
- Brusco, N., Ballan, L. and Cortelazzo, G. M. (2005). Passive reconstruction of high quality textured 3d models of works of art., *VAST*, pp. 21–28.
- Caprioli, M. and Scognamiglio, A. (2009). Low cost methodology for 3d modeling and metric description in architectural heritage, *Proceedings of the 3rd ISPRS International Workshop* 38.
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. and Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool, in V. Scarano, R. D. Chiara and U. Erra (eds), *Eurographics Italian Chapter Conference*, The Eurographics Association. <http://dx.doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136>.
- Covas, J., Ferreira, V. and Mateus, L. (2015). 3d reconstruction with fisheye images strategies to survey complex heritage buildings, *Digital Heritage*, 2015, Vol. 1, IEEE, pp. 123–126.
- Dias, P., Matos, M. and Santos, V. (2006). 3d reconstruction of real world scenes using a low-cost 3d range scanner, *Comp.-Aided Civil and Infrastruct. Engineering* 21: 486–497.
- Dorsey, J. (1995). Radiosity and global illumination, *The Visual Computer* 11(7): 397–398.
- Dupuis, J., Paulus, S., Behmann, J., Plümer, L. and Kuhlmann, H. (2014). A multi-resolution approach for an automated fusion of different low-cost 3d sensors, *Sensors*.
- El-Hakim, S., Beraldin, J., Gonzo, L., Whiting, E., Jemtrud, M. and Valzano, V. (2005). A hierarchical 3d reconstruction approach for complex heritage sites, *Proceedings of the 20th International CIPA Symposium, International Scientific Committee for Documentation of Cultural Heritage (CIPA)*.
- El-Hakim, S. F., Beraldin, J.-A., Picard, M. and Godin, G. (2004). Detailed 3d reconstruction of large-scale heritage sites with integrated techniques, *IEEE Computer Graphics and Applications* 24(3): 21–29.
- Gomes, L., Bellon, O. R. P. and Silva, L. (2014). 3d reconstruction methods for digital preservation of cultural heritage: A survey, *Pattern Recognition Letters* 50: 3–14.
- Grün, A., Remondino, F. and Zhang, L. (2004). Photogrammetric reconstruction of the great buddha of bamiyan, afghanistan, *The Photogrammetric Record* 19(107): 177–199.
- Isoda, Y., Tsukamoto, A., Kosaka, Y., Okumura, T., Sawai, M., Yano, K., Nakata, S. and Tanaka, S. (2009). Reconstruction of kyoto of the edo era based on arts and historical documents: 3d urban model based on historical gis data, *IJHAC* 3: 21–38.
- Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A. and Fitzgibbon, A. (2011). Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera, *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, ACM, New York, NY, USA, pp. 559–568. <http://doi.acm.org/10.1145/2047196.2047270>.
- Kazhdan, M. and Hoppe, H. (2013). Screened poisson surface reconstruction, *ACM Transactions on Graphics (TOG)* 32(3): 29.
- Koutsoudis, A., Arnaoutoglou, F. and Chamzas, C. (2007). On 3d reconstruction of the old city of xanthi. a minimum budget approach to virtual touring based on photogrammetry, *Journal of Cultural Heritage* 8(1): 26–31.
- Koutsoudis, A., Arnaoutoglou, F., Pavlidis, G., Tsiafakis, D. and Chamzas, C. (2008). A versatile workflow for 3d reconstructions and modelling of cultural heritage sites based on open source software, *Digital Heritage - Proceedings of the 14th International Conference on Virtual Systems and Multimedia*.
- Lari, Z., Al-Rawabdeh, A., He, F., Habib, A. and El-Sheimy, N. (2015). Region-based 3d surface reconstruction using images acquired by low-cost unmanned aerial systems, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40(1): 167.
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J. et al. (2000). The digital michelangelo project: 3d scanning of large statues, *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., pp. 131–144.
- Lindstrom, P., Koller, D., Ribarsky, W., Hodges, L. F., Faust, N. and Turner, G. A. (1996). Real-time, continuous level of detail rendering of height fields, *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, pp. 109–118.
- Louiset, T., Pamart, A., Gattet, E., Raharijaona, T., De Luca, L. and Ruffier, F. (2016). A shape-adjusted tridimensional reconstruction of cultural heritage artifacts using a miniature quadrotor, *Remote Sensing* 8(10): 858.
- Makantasis, K., Doulamis, A. D., Doulamis, N. D. and Ioannides, M. (2014). In the wild image retrieval and clustering for 3d cultural heritage landmarks reconstruction, *Multimedia Tools and Applications* 75: 3593–3629.
- Manferdini, A. M. (2012). The use of low-cost technologies for the promotion of cultural heritage sites: The case study of veleia, *VSMM*.

- Mikhail, E. M., Bethel, J. S. and McGlone, J. C. (2001). Introduction to modern photogrammetry, New York .
- Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., Kohi, P., Shotton, J., Hodges, S. and Fitzgibbon, A. (2011). Kinectfusion: Real-time dense surface mapping and tracking, *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, IEEE, pp. 127–136.
- Nöll, T., Köhler, J., Reis, G. and Stricker, D. (2015). Fully automatic, omnidirectional acquisition of geometry and appearance in the context of cultural heritage preservation, *Journal on Computing and Cultural Heritage (JOCCH)* 8(1): 2.
- Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V. and Chamzas, C. (2007). Methods for 3d digitization of cultural heritage, *Journal of cultural heritage* 8(1): 93–98.
- Pedersini, F., Sarti, A. and Tubaro, S. (2000). Automatic monitoring and 3d reconstruction applied to cultural heritage, *Journal of Cultural Heritage* 1(3): 301–313.
- Prokos, A., Karras, G. and Grammatikopoulos, L. (2009). Design and evaluation of a photogrammetric 3d surface scanner, *XXII CIPA Symposium on Digital Documentation, Interpretation and Presentation of Cultural Heritage* 2: 2.
- Ragia, L., Sarri, F. and Mania, K. (2015). 3d reconstruction and visualization of alternatives for restoration of historic buildings: A new approach, *2015 1st International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM)* pp. 1–9.
- Remondino, F. (2007). Detailed image-based 3d geometric reconstruction of heritage objects, *Trinational Conference of SGPBF, DGPF, OVG (Muttentz, Switzerland, 2007)*.
- Richardt, C., Stoll, C., Dodgson, N. A., Seidel, H.-P. and Theobalt, C. (2012). Coherent spatiotemporal filtering, upsampling and rendering of rgbz videos, *Computer Graphics Forum*, Vol. 31, Wiley Online Library, pp. 247–256.
- Rocchini, C., Cignoni, P., Montani, C., Pingi, P. and Scopigno, R. (2001). A low cost 3 d scanner based on structured light, *Computer Graphics Forum*, Vol. 20, pp. 299–308.
- Silva, J. W., Gomes, L., Agüero, K. A., Bellon, O. R. and Silva, L. (2013). Real-time acquisition and super-resolution techniques on 3d reconstruction, *Image Processing (ICIP), 2013 20th IEEE International Conference on*, IEEE, pp. 2135–2139.
- Sormann, M., Zach, C., Zebedin, L. and Karner, K. (2005). High quality 3d reconstruction of complex cultural objects, *Archives of CIPA XXth International Symposium, Torino*, Citeseer.
- Stentoumis, C., Livanos, G., Doulamis, A., Protopapadakis, E., Grammatikopoulos, L. and Zervakis, M. (2013). Precise 3d reconstruction of cultural objects using combined multi-component image matching and active contours segmentation, *International Symposium on Visual Computing*, Springer, pp. 148–157.
- Tarini, M., Cignoni, P. and Montani, C. (2006). Ambient occlusion and edge cueing for enhancing real time molecular visualization, *IEEE transactions on visualization and computer graphics* 12(5).
- Utomo, A. P. and Wibowo, C. P. (2017). 3d reconstruction of temples in the special region of yogyakarta by using close-range photogrammetry, *SEMNASTEKNOMEDIA ONLINE* 5(1): 4–3.
- Vogiatzis, G. and Hernández, C. (2010). Practical 3d reconstruction based on photometric stereo, *Computer vision*, Springer, pp. 313–345.
- Wu-Wei, C. (2016). Digital heritage: Digital sculpting, cyber-archiving and education, *SIGGRAPHASIA 2016 Crafting Spirituality: A Pedagogic Project for Digital heritage: Digital Sculpting, Projection Mapping and Beyond*, ACM, p. 1.
- Zeppelzauer, M., Poier, G., Seidl, M., Reinbacher, C., Breiteneder, C., Bischof, H. and Schuster, S. (2015). Interactive segmentation of rock-art in high-resolution 3d reconstructions, *Digital Heritage, 2015*, Vol. 2, IEEE, pp. 37–44.
- Zhang, Z. (2012). Microsoft kinect sensor and its effect, *IEEE multimedia* 19(2): 4–10.
- Zhou, K., Hu, Y., Lin, S., Guo, B. and Shum, H.-Y. (2005). Precomputed shadow fields for dynamic scenes, *ACM Transactions on Graphics (TOG)*, Vol. 24, ACM, pp. 1196–1201.
- Zollhöfer, M., Siegl, C., Riffelmacher, B., Vetter, M., Dreyer, B., Stamminger, M. and Bauer, F. (2014). Low-cost real-time 3d reconstruction of large-scale excavation sites using an rgb-d camera, *JOCCH* 9: 2:1–2:20.