An Integrated Framework for the Interaction and 3D Visualization of Cultural Heritage

Abdelhak Belhi [1] (✉), Hosameldin Osman Ahmed [2], Taha Alfaqheri[2], Abdelaziz Bouras [3],
Abdul H. Sadka[2], Sebti Foufou[4]

[1] Joaan Bin Jassim Academy for Defence Studies, Qatar
abdelhakbelhi@gmail.com
[2] Brunel University, London, United Kingdom
{hosameldin.ahmed2, taha.alfaqheri, abdul.sadka}@brunel.ac.uk
abdelaziz.bouras@qu.edu.qa
[3] CSE, Qatar University, Doha, Qatar
[4] Department of Computer Science, The University of Sharjah, Sharjah, UAE
sfoufou@sharjah.ac.ae

Abstract

In this study, the aim is to design and develop a 3D acquisition, visualization, and interaction framework to preserve cultural heritage and provide new ways to enable museum visitors and cultural audiences to virtually interact with cultural objects. Indeed, cultural assets are nowadays at higher risk and most cultural institutions prohibit visitors from physically manipulating their collections. The main motivation behind our framework is to enable end-user interaction with high valuable cultural objects while addressing cost-effectiveness concerns as well as minimizing the time required to digitize and generate 3D models of cultural heritage objects. The design idea of our framework is to allow interaction with the protected assets' 3D representation using a real-world 3D screen equipped with a depth sensor namely the leap motion controller. Our framework is an end-to-end solution that optimizes all the stages of the 3D acquisition, pre-processing, visualization, and interaction pipeline while providing contributions to its stages. It achieves good quality results thanks to the use of machine learning in the acquisition and modeling stages. Indeed, we adapted a prior preprocessing work that performs super-resolution and motion interpolation on the acquired data. The preprocessed data is then used for the generation of the 3D models using photogrammetry, which optimizes the quality of the resulting 3D models. The created 3D models are then adapted for the visualization and interaction stages. A novel visualization and interaction paradigm is introduced to enable a real-world experience for museum visitors through a 3D screen called "the Looking Glass". The interaction with the 3D content is achieved through a motion sensor used to design our new interaction component of the framework. We propose two new interaction systems suitable for various user profiles focusing on their experience in dealing with motion sensors. The end-to-end framework tested in a museum environment was evaluated by cultural heritage curators and multimedia experts and found to provide an alternate reality tool for asset exhibition and a cost-effective alternative for asset exchange between cultural institutions. For the evaluation, we compared the end-user experience of our framework using various setups where users are visualizing the content through 2D screens and through the Looking glass while enabling and disabling motion interaction. The results of the evaluation suggest that the looking glass paired with the Leap motion sensor using our framework as a backend enables an alternate reality experience for museum visitors and new ways of interacting with cultural content, sharing of cultural knowledge, cultural education, and much more.

Keywords: Cultural Heritage, 3D Modeling, Human Computer Interaction, Alternate Realities, Digital twin, Leap Motion Controller, Machine Learning.

1 Introduction

Cultural heritage is the essence of humankind as it packs a massive amount of historical information that cannot be retrieved elsewhere. Cultural heritage assets are distinguished by their variety, shape, type, and value. Hence, the preservation of cultural heritage is an important process to curate and maintain assets along with their provenance information for current and future generations. However, their physical preservation is a tedious and delicate process that requires a long time, lots of resources, and has to be undertaken by highly skilled professional curators. As a cost-effective and reliable way for art and culture preservation, digital technologies offer additional ways to preserve and increase the value and excitement around cultural heritage [1-3]. A lot of effort was undertaken to provide Information Technology (IT) solutions in the cultural domain. Some of the existing applications are geared towards digital preservation and collections documentation to ease the management and the retrieval of assets information. Other applications focus mainly on the end-user experience with innovative ways to increase the value and attractiveness of assets using recent data acquisition and visualization technologies such as 3D, VR, AR and other immersive technologies to enable new ways of content consumption in the cultural heritage domain[4].

Not long ago, the British Museum visitors could interact physically with some assets such as the iconic Rosetta stone (See Figure 1). However, nowadays, such high value historical objects are protected from degradation and damage in controlled environments by glass shells preventing any kind of physical interaction. Nevertheless, it is rather more attractive to allow museum visitors, especially young guests, to interact physically with assets. But since it is impossible to do that anymore, several initiatives are leveraging the recent advances in technology to provide an alternate reality and a more immersive experience for museum visitors with the use of techniques such as 3D screens, VR, AR, projection, sound effects, mockups of original objects, and others. This new way of exhibition and interaction with tangible cultural heritage is set to induce drastic changes to the way that museum visitors, enthusiast, and cultural professionals deal with cultural heritage. In our study, this has been achieved through several consultations with cultural heritage curators and multimedia professionals from the MIA museum in Qatar.

Figure 1. Museum objects inside a glass shell (the Rosetta stone on the left and an iconic statue head from the MIA museum in Qatar, on the right)

In this paper, and in the context of digital heritage, we propose a novel framework for the 3D acquisition, processing, visualization and interaction for tangible cultural heritage. The proposed framework relies on multiple cutting-edge technologies such as a novel cultural heritage 3D capturing system which depends on a custom physical acquisition setup, machine learning with super resolution as well as motion interpolation for pre-processing, and photogrammetry. As far as we know, no existing framework matches the capabilities of our proposed one in terms of functionality and innovative ways to interact with cultural heritage assets.

In our case, we aim at tackling the challenge of 3D content acquisition, adaptation, and visualization using consumer-level hardware proposing a more cost-effective yet attractive framework to allow 3D visualization and interaction with high-value museum assets. In this framework, first, a Digital Single Lens Reflex (DSLR) camera, lighting setup and a turntable are used for the 3D content acquisition. In the subsequent stage, our framework uses photogrammetry as the main technology to model 3D cultural objects. Several contributions are introduced to this process where machine learning techniques such as super-resolution and motion interpolation are used to generate more high-resolution input images for the photogrammetry process. A state-of-the-art real-world 3D screen called "the Looking Glass" is used for the visualization part [5]. This screen uses a multiview light-field technology and is capable of projecting 3D content in real-time. Viewers can enjoy this 3D experience without having to wear glasses, as with traditional stereoscopic 3D screens, or headgear in case of VR headsets. Objects projected on this screen are viewed as they are in real-life. For the interaction part, this screen can be fitted with a touch screen digitizer layer but in our case, we wanted to allow visitors to interact with the cultural content as if they were holding it with their hand to provide them with an alternate reality experience close to naturally interacting with the object. We thus proposed a 3D interaction approach based on the Unity3D library leveraging a Leap Motion Controller (LMC) [6] used to capture end-users hand motion. This motion is then translated into controls to interact in real-time with the generated 3D model. A custom component is proposed based on the LMC API and allows various interaction paradigms not found in other approaches such as object 360° rotation, movements, zoom and other actions with the asset [7]. We propose two new interaction systems that are designed to serve users with different levels of experience when dealing with motions interaction modules.

The quality evaluation of the 3D models was undertaken with the help of museum professionals. The assessment of the proposed solution was undertaken using a surveyed audience where people were evaluating the quality, useability, adaptability and other modalities. The evaluation results of our system suggests that it is more responsive than comparable platforms while providing high-quality visualization. Therefore, our solution is suitable for real-world implementation and can be adopted by museums using screen fitted with the leap motion controller next to a protected asset.

These results have been confirmed by curators and cultural heritage multimedia professionals during our tests in the Museum of Islamic Art in Doha, Qatar where the main use case of this technology was the exchange of assets between cultural institutions. This exchange is considered extremely risky as it involves huge insurance costs which often are multiple times the price of the display for insuring a single asset. One of the real-world use cases of this technology is due to be tested during the Qatar-US cultural year where assets from museums such as the Metropolitan Museum of art (The MET) were set to be exchanged using 3D models and displayed using our framework.

The rest of this paper is organized as follows. In section two, we present works related to the 3D acquisition, visualization, and interaction with cultural heritage assets. In section three, we present the methodology, the implementation steps, and the evaluation of our framework focusing mostly on data acquisition, data preprocessing, photogrammetry, and motion interaction. Section four outlines our interaction system evaluation setup, the results as well as a discussion of the latter. In section five, we give our conclusions and set some perspectives for future work.

2 Related work

The digitization of cultural heritage plays an important role in long-term preservation as it is more reliable and less difficult to implement and maintain assets in a digital form. As an added challenge, museums and heritage institutions want to promote their collections using new digital content consumption techniques specially to attract visitors and give more value to their collections. In this regard, a lot of work has been undertaken for the design and implementation of 3D acquisition and visualization technologies for cultural heritage serving a wide range of use cases. In this work, our primary focus is geared towards applications dedicated to exhibitions and end-users, not for professional applications. The reason is that system requirements for the latter are usually strict about quality and do not necessarily focus on cost-effectiveness and interaction, the main drivers behind our study.

In the following, we present a review of works related to 3D acquisition and modeling as well as the motion interaction with 3D objects.

2.1 3D Acquisition and Modelling

The 3D acquisition and modeling of an object is an end-to-end process that starts from the physical object itself and ends with its three-dimensional representation. It usually involves the creation of a point cloud in space which is then used to create a triangulated network (mesh) or a textured surface. The representation of physical objects has been for a long time a top subject in computer science. Usually, two main approaches are used. The first one requires drawing the shape of the object in Computer-Aided Design software (CAD) which can be seen as the most difficult, time-consuming, and expensive 3D modeling approach [8]. But CAD is yet the most accurate one as all the details of an object are manually drawn. Furthermore, usually, it is only appropriate for a certain type of physical shape and works best for objects having uniform patterns, regular shapes, etc. A random rock or stone is almost impossible to reproduce in CAD at a reasonable cost. Nevertheless, CAD is still yet the most accurate approach in terms of representation, as for high-quality projects, designers often use high precision measuring tools to reproduce objects at the highest possible precision.

Another way of representing physical objects in virtual space is through scanning techniques. Usually, we can distinguish between two types of techniques: Laser scanning and photogrammetry. Laser scanning, which is the most accurate scanning technique, uses a distance-measuring laser beam to record the position of an object surface points in space. Laser scanners have a clear advantage in terms of the representation accuracy of the generated point cloud. Additionally, a digital camera is used to capture the object textures while the laser captures the shape of the object. The scanning would be accurate if the object is of a certain shape and has homogeneous textures. However, if the object has multiple levels of depth from a single point of view, it usually requires manual adjustments in addition to denoising and software-based cleanup to achieve accurate representation of the object. Laser scanners are expensive and some of the high precision laser units cost millions of dollars and are used only in certain environments. Consumer-level laser scanners, such as “Matter and Form” [9], are usually cheaper but unfortunately, only work with modest size objects in addition to not being accurate enough for applications requiring high accuracy representation. In the cultural context, the most iconic scanner used is the CultLab3D (see Figure 2) developed in Germany [10] (cost per scan are around 1000 USD as per 2020) [11].

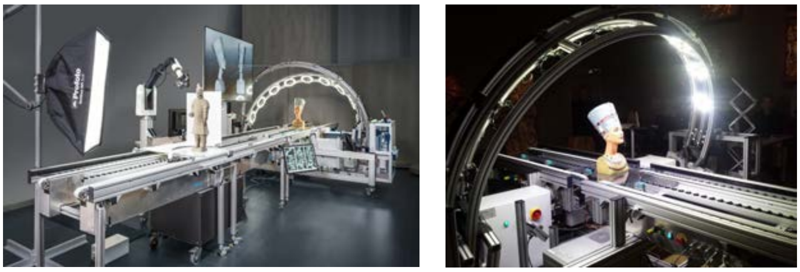


Figure 2. The CultLab3D scanner at Fraunhofer IGD [10]

Photogrammetry in comparison does not require fancy hardware. It uses consumer-level digital cameras configured in a certain setup and a software tool used to recreate the 3D shape from a combination of images and a depth estimation process. Some of the photogrammetry techniques are Shape from a stereo, Shape from motion, Shape from shading, and Shape from silhouette [12]. Photogrammetry is cost-effective, and can yield very accurate results, but it often fails as if one of the provided images is not well exposed, focused, or has a shift in perspective, the whole process is affected [13].

2.2 3D visualization

The main shortcoming of the current visualization technologies in most cultural heritage museums is the lack of 3D visual interactivity support for most of the museum's visitors. Most modern 3D visualization systems now include 3D models to generate an impressive 3D visualization experience[14]. Although 3D models are useful to preserve the information about historical artefacts, the potential of these digital contents are not fully accomplished until they are used to interactively communicate their significance to non-specialists[15]. Currently, there exists three commercially available 3D display technologies; Active glass, passive glass, and Multiview (Autostereoscopic) display, see Figure 3.

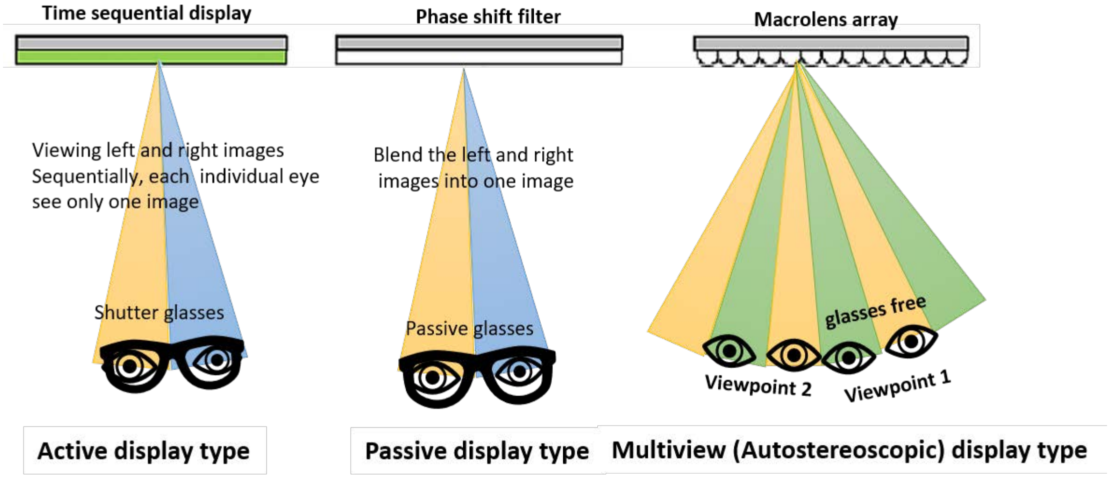


Figure 3. Commercially available 3D display types[13]

All these 3D display technologies share the same methodology to induce the perceived image depth. Both active and passive display types require a special glass for the display viewer. The active glass type displays the left and right image sequentially and in a high frequency shutter and synchronizes with the display light emitter [16]. In the passive glass display type, both the left and right images are visualized with different polarizations simultaneously by the viewer. The polarization state is fixed using a phase shift film glued on the display front surface that modifies the intrinsic linear polarization of the LCD to the required left or right circular state for each row of pixels alternatively. Thus, the circular polarization should be adjusted with its related emission angle.

In Multiview (Autostereoscopic) display, the image rays are emitted from parallax barrier placed in front of LCD panel. This type of 3D display does not require glasses. Additionally, this type of display provides more than two views to enable the viewers to provide true 3D depth perception [17]. Auto viewing angle adaptation is more advanced Multiview technique enables the viewers to track their head and adjust the displayed image with viewing angle accordingly [18]. The screen we are using in this work falls in the Multiview category. It uses a holographic technology and can project up to 45 different views. The screen is developed by the Looking Glass Factory company.

2.3 Motion Interaction

Human-computer interaction is a complicated field in computer science that focuses on the user experience and how to translate gestures into commands for computers to understand. Solutions such as VR headsets were introduced aiming at providing an immersive experience for users. These solutions try to reconstruct an immersive museum visit experience by modeling museum architecture as well as assets [19-21] . Other approaches use virtual reality headsets and motion controllers to provide a more immersive experience for end-users [15].

Since the introduction of consumer-level motion interaction sensors such as the Leap Motion Controller [6] and the Microsoft Kinect [22], a lot of applications targeting enhancement of user experience were proposed. These sensors can translate natural human gestures into commands without having to wear any type of device or headgear and without having to deal with sophisticated controllers. This is, in fact, suitable for public usage and beneficial for average users who may not know how to deal with VR headsets for example. In this work, we mostly focus on the tracking of hand motion. Kinect is thus not suitable for this scenario as it was designed to track the human posture. Leap motion controller (LMC) instead is a very accurate (0.01 mm) depth sensor which is designed to track human hands in space [23]. Due to its reasonable cost (99 $), the LMC was successfully used in multiple applications involving hand gestures.

In the cultural context, many approaches are leveraging depth sensors to provide new ways of enabling interaction with cultural objects. Among these approaches, in [24], the authors presented a framework that supports “kinesthetic interactions” where they used a Leap motion controller to capture hand motion. They tried to replicate a scenario where the user acts a sculptor and has to learn how to manipulate it for carving, smoothing, engraving and other manipulations during its conception and after it has cured. It is worth noting that this application is unique of a kind and shows the power of such a depth sensing device (see Figure 4).

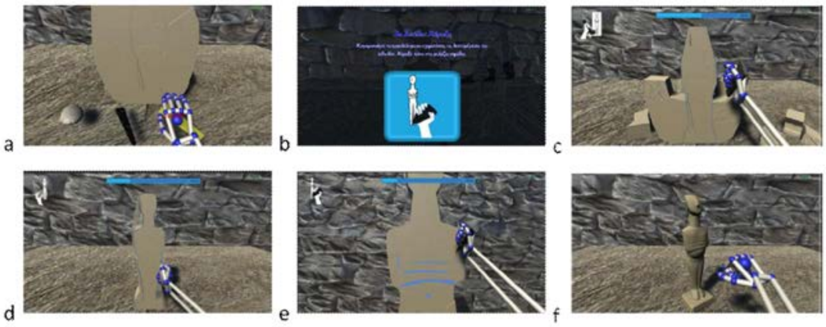


Figure 4. A Kinesthetic Approach to Digital Heritage using Leap Motion [24]

In the context of pottery, the authors of [25] provide a new three-dimensional prototype based on the LMC allowing users to solve virtual pottery puzzles using their hand as the main way of interaction. The authors of [26] highlight that in virtual museums, the LMC can be paired with virtual reality gear to provide an immersive experience when dealing with assets without having to manipulate sophisticated controls.

3 Methodology

In this section, we present the design and implementation details of our cost-effective cultural 3D acquisition, preprocessing, visualization, and interaction framework. Figure 5 presents the usage scenario of our 3D interaction and visualization platform in a museum. As shown in Figure 6, the proposed framework consists of four main stages. These include data acquisition, preprocessing, photogrammetry and 3D model adaptation, and motion-based 3D visualization. The following subsections discuss these stages in more details.

Figure 5. A typical usage scenario of our system where a screen is stationed next to a protected cultural object. By means of a leap motion controller, and our 3D interaction software, a museum visitor can interact with the asset in a virtual environment using his bare hands.

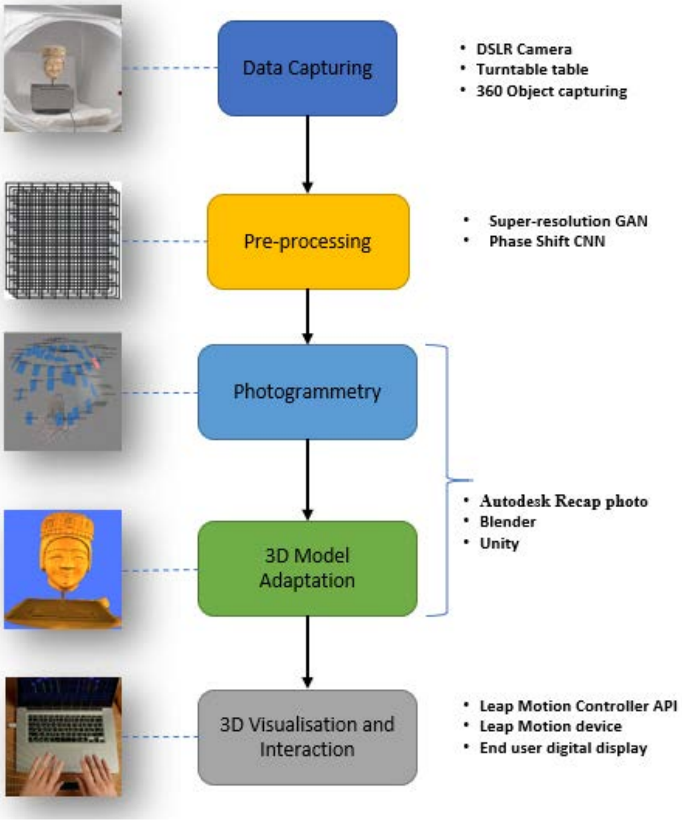


Figure 6. The main stages of our framework: from data acquisition to visualization and interaction

3.1 Data acquisition

The data acquisition stage is crucial for our framework as high-quality images are required for the process of photogrammetry in order to generate high quality 3D models assets for cultural assets. The registration accuracy of the object geometry usually depends on the object texture and size.

To achieve optimal results, we relied on a professional photography setup consisting of:

- Lighting: The lighting setup includes two light-emitting projectors. The light was diffused by a photography box and in some cases, for reflective objects, we used a foam board positioned next to the camera sensor to smooth harsh light (See Figure 7).
- Camera stands and tripods: The use of professional-grade camera stands is very important in our case to fully control the camera viewpoint.
- Measurement devices: We used an optical distance measuring device as well as a light intensity meter (LUX meter) to save the setup parameters to allow the reproduction of the optimal parameters depending on 3D modeling step.
- A constant speed self-rotating turntable: The turntable was used to automatically turn the object in a uniform pattern.
- Cameras: we used a Canon DSLR camera for our experiments.

To get good quality capturing results, all the capturing equipment is installed and configured in the best possible way with the help of professional photographers. The scenario is as follows: The camera is stationed at a certain distance and at a specific angle from the object. The object will automatically rotate, and the camera will uniformly capture it from various perspectives shifted by 5 to 6 degrees while the turntable is turning. After a trial-and-error session consisting of capturing a set of images, going to the photogrammetry software, looking at the result and then changing the setup, the best parameters for the acquisition setup were found as follows:

- The best range for lighting is ranging between 1500 and 2000 Lux for getting more details on the object surface.
- The turntable speed must be matched to the camera shutter speed (we used 1/250 of a second as the shutter speed in our tests).
- The camera was put in manual mode to avoid changes in the exposure when selecting shutter priority mode. The aperture selected was the lowest setting possible on the camera and the ISO range was between 100 and 400 depending on the object (for dark objects, a higher ISO is required to show the fine details).
- The camera was positioned into three different vertical positions shifted by 45 degrees (see Figure 8).

Following the assets capture, we used a machine learning visual enhancement framework to preprocess the acquired images before we perform the photogrammetry step. This was mainly to enhance the quality of the acquired pictures to achieve the best-looking 3D models possible.

Figure 7. Our Cultural Heritage Visual Data Acquisition Setup at the MIA museum

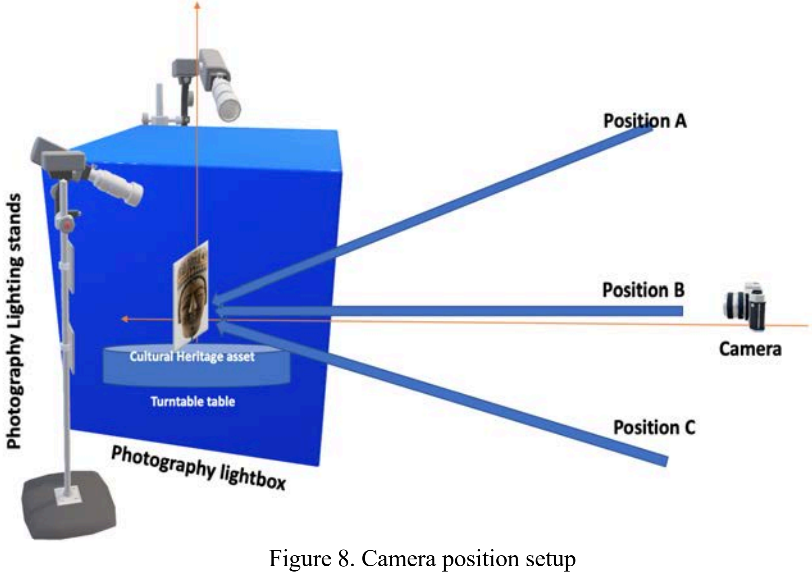


Figure 8. Camera position setup

3.2 Data preprocessing

After performing the data acquisition as described in the previous section, a preprocessing step is required in order to filter, enhance the quality and organize the captured visual content to be used for the 3D model generation. Photogrammetry has the main advantage of not requiring expensive 3D scanning hardware as it only works on single 2D images captured using a 360° shooting setup. However, if one or some of the captured pictures are in low quality due to bad camera positioning or environmental parameters, the whole 3D modeling process will most likely fail. The preprocessing step is thus just an optimization to the quantity and quality of the output.

In addition to the controlled environment setup, and to preserve cost-effectiveness while still having good quality results, we proposed a super-resolution and motion interpolation stages to enhance the quality of the acquired images and generate more frames for the photogrammetry process as shown in Figure 9 and Figure 10.

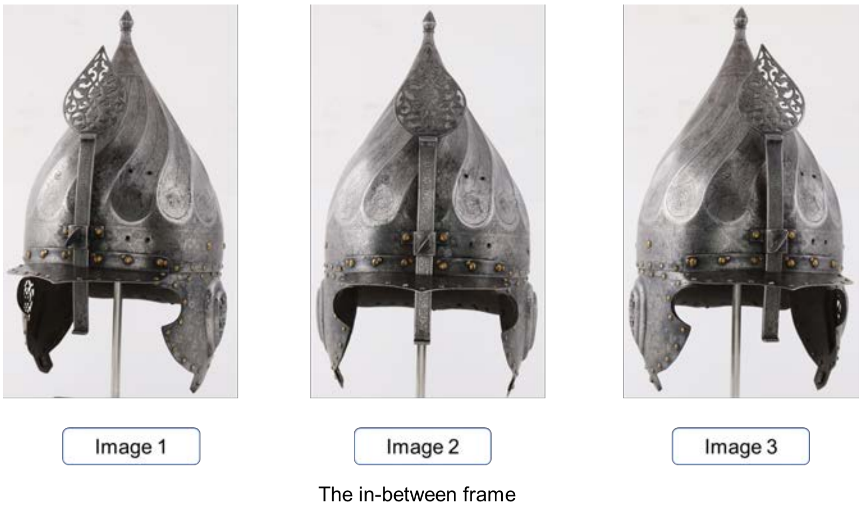


Figure 9. Computation of in-between frames (motion interpolation)

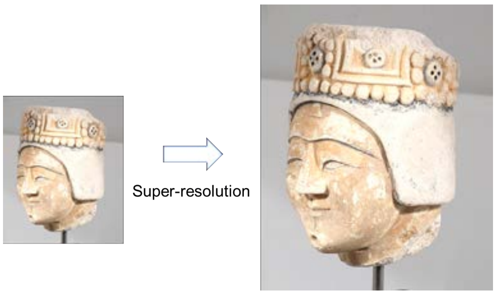


Figure 10. A super-resolution example

- **Super resolution**

Super-resolution refers to a set of machine learning techniques proposed in the context of computer vision to upscale and increase the resolution of images while trying to preserve details. Traditional techniques such as bicubic interpolation have major drawbacks in terms of quality. The Convolutional Neural Network (CNN) based super resolution techniques were proven to yield very good results. However, with the introduction of generative adversarial networks (GAN), super resolution approaches saw a boost in terms of performance and output quality. As a result, and as a cost-effective alternative to expensive high-end professional cameras, we applied our super-resolution and motion interpolation techniques to increase the spatial resolution of the images we captured. The framework we selected for our experiments is SRGAN known to have superior performance [27]. It is worth noting that our super resolution model was fine-tuned and retrained on a huge dataset of cultural heritage objects. The dataset we used for retraining was collected from various cultural institutions such as WikiArt, the Rijksmuseum, The Metropolitan Museum of Art, and the MIA museum of Qatar.

- **Motion interpolation**

Motion interpolation is a technique used to increase the framerate in video sequences by the computation of in-between frames using different techniques. In our case, since we are capturing our images using a 360° scenario with the help of our acquisition setup, we found that it is rather beneficial to use motion interpolation instead of having to slow down the turntable to record more frames. This is beneficial when dealing with a large volume of captures as often museums cannot afford to take their objects off display for a long period. The principle of motion interpolation is shown Figure 9. Usually, given two images of steps 1 and 3 of an arbitrary motion, applying interpolation to compute the in-between frame (image 2) of images 1 and 3 is called motion interpolation.

We thus experimented with two techniques of motion interpolation, the first one is based on "*miterpolate*" which is a pixel-level interpolation filter provided by the FFmpeg video encoding library [28] and the second one is based on deep learning (called PhaseNet) and was mainly intended for converting normal videos to slow-motion videos [29].

We found that the two motion interpolation approaches were similar, as in our case we do not have major changes of perspective in our images such that only the object rotates by 5° in each frame. The impact seen on the photogrammetry results was in some cases significant and resulted in more details specially in textures. For our final version, we selected the method based on deep learning.

3.3 Photogrammetry and 3D Model Adaptation

Photogrammetry is a technique for generating 3-dimensional shapes through the analysis, measurements, and interpretations from a group of images acquired using a set of strict guidelines. Compared to laser scanning, this technique has some disadvantages related to non-accurate measurements and some generated artifacts. However, its most appealing aspect is its easiness of use and cost-effectiveness.

For our framework, we compared two among the best available tools for photogrammetry in the context of cultural heritage: Autodesk Recap Photo (see Figure 11) and AgiSoft PhotoScan (see Figure 12). Both performed similarly but using our setup, Autodesk Recap Photo yielded consistent results with inputs coming from our preprocessing stage.

Figure **11**. Photogrammetry using Autodesk Recap Photo

Figure 12. Photogrammetry using AgiSoft PhotoScan

Once preprocessed, the images are imported into the Autodesk Recap photo software and the 3D model is then generated using the Autodesk Cloud service. After that, the 3D models are minimally tweaked using the provided tools as well as Blender 3D and then exported to either OBJ or FBX formats.

The following section of this paper describes the visualization and interaction techniques proposed within our framework to display and allow human interaction with the created 3D models. This is the most important stage of our framework regarding the user experience.

3.4 Visualization and Interaction

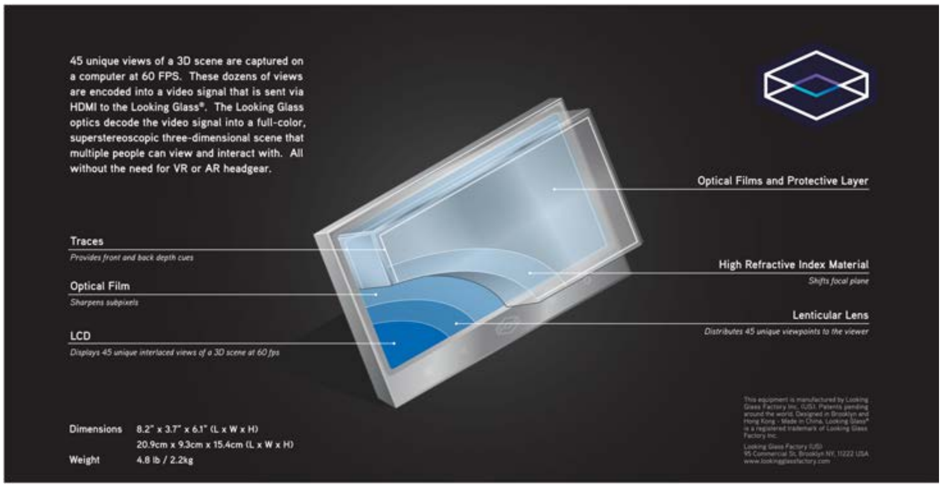
As pointed out in the introduction, our framework includes new approaches to visualize and allow human interaction with the created 3D models from the photogrammetry step. In the following, we discuss the 3D visualization stage as well as the two interaction modules we proposed to allow an alternate reality experience for museum visitors when it comes to interacting with cultural heritage 3D models using their hands. It is worth noting that the Unity 3D software library [30,31] was used to develop our visualization and interaction stage.

3.4.1 Visualization of the 3D models

One of the most important components of cultural heritage digital preservation is the preserved content visualization. Indeed, the digitally preserved content must be reproducible or representable in a form exactly similar to the original one with a high fidelity. This is the main aspect required to allow an alternate reality viewing experience and a digital twin of the physical asset. As a result, 3D visualization is an adequate technology as it allows the viewer to perceive a virtual representation of the original object in a three-dimensional space.

A wide range of 3D visualization technologies exist, and the majority of these technologies require the viewer to wear some kind of eye wear in the case of stereoscopic screens or headgear in case of virtual reality. These technologies are unfortunately not suitable to be used in an exhibition environment, especially during times where people are experiencing social distancing. For our solution, we are using the Looking Glass 15.6" screen which is a real-world 3D light field display capable of projecting 3D models. Viewers will be able to enjoy 3D content without having to wear any kind of headgear. The looking Glass uses a patented technology to display a Multiview representation of the 3D content using 45 different views. Figure 13 shows the technical highlights of this screen. The price of a single unit is around 3K USD, as per 2020, which makes it a compelling option for museums and cultural institutions in many use cases as it was highlighted in section 1.

For the integration of the 3D screen with 3D applications, Looking Glass provides an SDK which can be integrated with the Unity 3D library enabling the visualization of 3D content on the Looking Glass display [5]. A set of parameters controlling the rendering of the resulting images has been set depending on the environmental condition such as the distance from the screen, the 3D accelerator used (GPU) and the complexity of the 3D model. Indeed, to achieve good quality results, 4K resolution of the output at 60 frames per second with Ultra details is needed and this requires a very powerful graphics accelerator to be achieved. However, it turned out that for this task, consumer-level GPUs are not very suitable. We thus used a professional grade Nvidia Titan RTX GPU as our graphics accelerator.

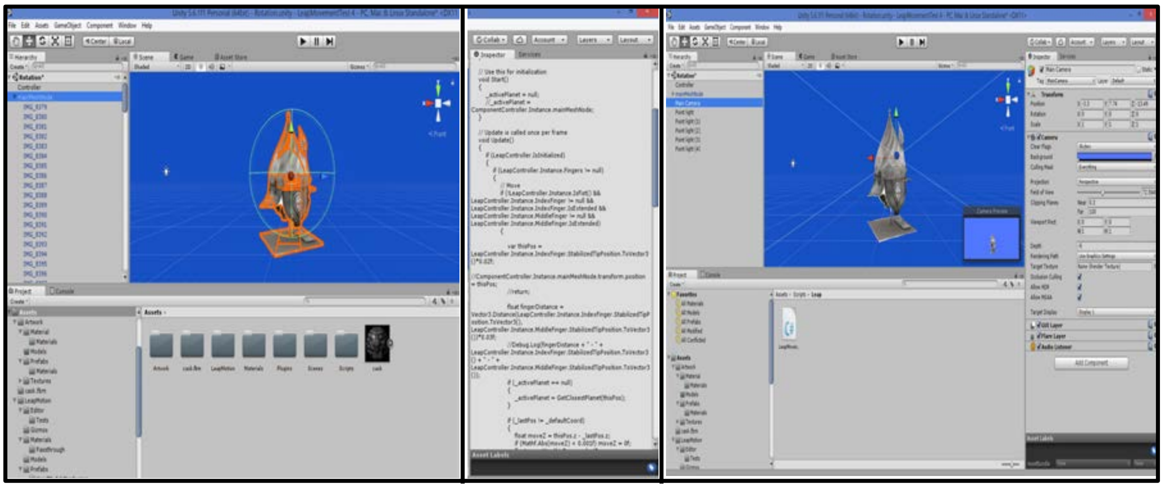


45 unique views of a 3D scene are captured on a computer at 60 FPS. These dozens of views are encoded into a video signal that is sent via HDMI to the Looking Glass®. The Looking Glass optics decode the video signal into a full-color, superstereoscopic three-dimensional scene that multiple people can view and interact with. All without the need for VR or AR headgear.

Figure 13. The Looking Glass 15.6" development kit

The visualization component of our framework was built using the Unity 3D library as it has the ability to be integrated with the HoloPlay service which is the Looking Glass Factory plugin to visualize the 3D output on the Looking Glass display and also because it can be integrated with a variety of motion controllers such as the leap motion controller. The Unity 3D software supports different platforms and operating systems, e.g. Windows, Mac, Linux, iPhone Operating system (iOS). Also, it can deal with different formats of 3D models such as OBJ, SBX, and FBX.

The visualization process starts by importing the targeted 3D object and its textures in the Unity 3D main window. Figure 14 shows visualization examples of the 3D objects in Unity 3D. Furthermore, as can be seen in the same figure, the Unity 3D provides an interface to manipulate the 3D models using programming scrips. We thus setup a set of object views and predefined perspectives to facilitate the viewing of objects on the user's interface. The solution can be applied on normal 2D and 3D screens thanks to the seamless integration of the HoloPlay service. Figure 15 presents visualization examples of the 3D objects in Unity 3D, whereas Figure 16 represents the visualization of the Head object from the MIA museum on the Looking Glass.



```
// Use this for initialization
void Start()
{
    _activePlanet = null;
    //_activePlanet =
ComponentController.Instance.mainMeshNode;
}

// Update is called once per frame
void Update()
{
    if (LeapController.IsInitialized)
    {
        if (LeapController.Instance.Fingers != null)
        {
            // Move
            if (!LeapController.Instance.IsFist() &&
LeapController.Instance.IndexFinger != null &&
LeapController.Instance.IndexFinger.IsExtended &&
LeapController.Instance.MiddleFinger != null &&
LeapController.Instance.MiddleFinger.IsExtended)
            {

                var thisPos =
LeapController.Instance.IndexFinger.StabilizedTipPosition.ToVector3
()*0.02f;

//ComponentController.Instance.mainMeshNode.transform.position
= thisPos;
                //return;

                float fingerDistance =
Vector3.Distance(LeapController.Instance.IndexFinger.StabilizedTipP
osition.ToVector3(),
LeapController.Instance.MiddleFinger.StabilizedTipPosition.ToVector3
()*0.02f;
                //Debug.Log(fingerDistance + " - " +
LeapController.Instance.IndexFinger.StabilizedTipPosition.ToVector3
() + " - " +
LeapController.Instance.MiddleFinger.StabilizedTipPosition.ToVector3
());
                if (_activePlanet == null)
                {
                    _activePlanet = GetClosestPlanet(thisPos);
                }

                if (_lastPos != _defaultCoord)
                {
                    float moveZ = thisPos.z - _lastPos.z;
                    if (Mathf.Abs(moveZ) < 0.002f) moveZ = 0f;
```

Figure 14. Unity 3D visualization and editing interface

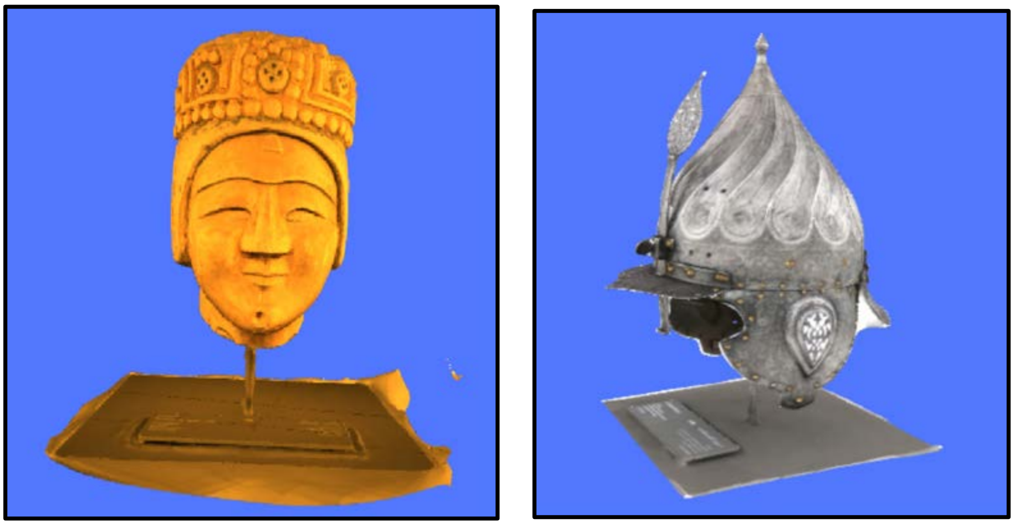


Figure 15. Visualization examples of the created 3D objects in Unity 3D

Figure 16. The head object Model on the Looking Glass

3.4.2 Interaction with the 3D models

Along with a state-of-the-art 3D visualization approach using the Looking Glass Factory 3D display, we have proposed a specific Human-computer interaction (HCI) using motion sensors and a 3D screen intended to allow end-user interaction with 3D cultural heritage assets. In this approach, users interact with the 3D content in a way considered almost natural similarly to manipulating a physical object using their own hands in space. This new solution does not require the user to have prior training and does not involve dealing with sophisticated controllers. This is possible thanks to the integration of the leap motion controller (LMC) device with our visualization method based on Unity 3D. The leap motion controller is a relatively cheap device (99 USD per unit as per 2020) which has the ability to track users' hands in space. In our solution, the hand positions are tracked in real-time using the LMC and based on certain gestures, actions are executed to manipulate the 3D object in the Unity 3D software and render the output to the Looking Glass display using the HoloPlay service.

The challenges then are two folds. On the one hand, users need to interact with the virtual object in the most natural way possible and need to get familiar with this interaction way quickly. This means that the gestures have to be fine-tuned in software with the help of multiple people during the development period. On the other hand, the action interpretation and classification need to be performed in real-time to ensure a perfect synchronization between the user-action, their interpretation and the object movement on the display.

In our solution, we have proposed two distinct interaction modules. The first interaction module is most suitable for users interacting with the framework for the first time and most likely not familiar with motion interaction. In this module, users can see a virtual representation of their hand on the screen which makes it easier to manipulate the 3D object. Indeed, users are able to accomplish different interaction tasks including turning, moving, grasping, zooming, pushing forward and backward, etc. The users' hand on the screen may however block some object details. The second proposed interaction module is geared towards more experiences and initiated users. The module enables the users to accomplish 14 different tasks and gestures. In the following subsections, we will discuss in detail these two proposed interaction modules.

- **The first interaction module**

As can be seen in Figure 17, our first interaction module consists of a laptop, a Leap motion controller, and the Unity 3D software accompanied by C# scripts to translate gestures into controls. The 3D object visualization comprises preprocessing object data by assigning different attributes through the Unity inspector tool and a C# script to control the location, rotation, and scale of the 3D objects in the Euclidean X/Y/Z space. For example, adding a collider, i.e., boundary, to the simulated 3D, controls how the model is displayed via mesh render, creating and applying the material to the model that contains the texture information. The leap motion controller is used to track the motion of bare hands to allow natural interactions with a 3D object in the controlled area of the Unity 3D visualization and interaction window. The Leap Motion controller includes three infrared emitters and two infrared cameras that can be used to track the hands and fingers positions. Then, based on this tracking, the controller extracts information and transmits it to the laptop to be used by the Unity 3D software for the interaction with the targeted object. In our first interaction module, the extracted information is used to present virtual hands that can interact with an asset's 3D model. The hand's interaction techniques can be categorized into pushing and grasping to achieve several motions translated to the 3D model, e.g., moving to the right side, moving to the left side, push forward, etc. (see Figure 17). For these tasks, we built specific scripts using the C# programming language. Figure 18 displays screenshots from various interaction types with a 3D head model using the first interaction module.

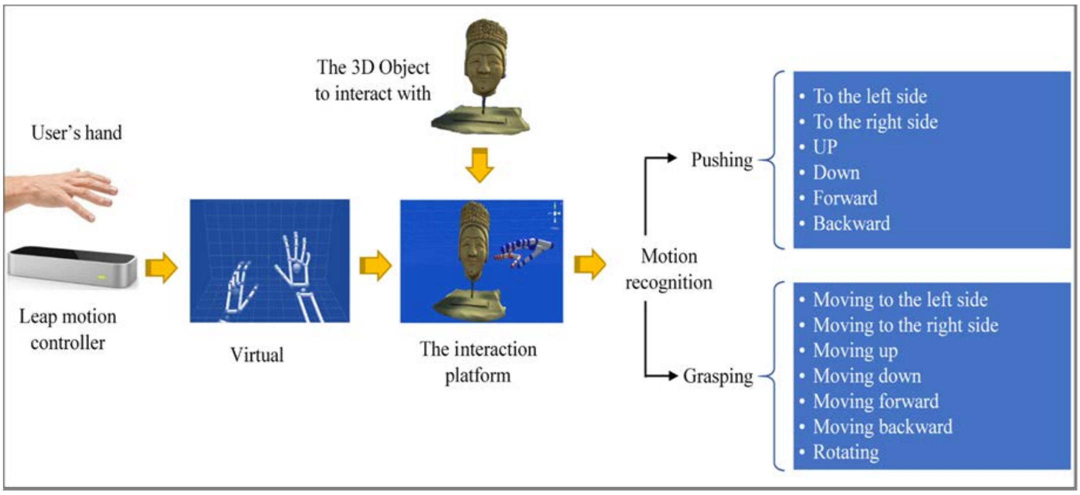


Figure 17. Diagram of the first interaction module

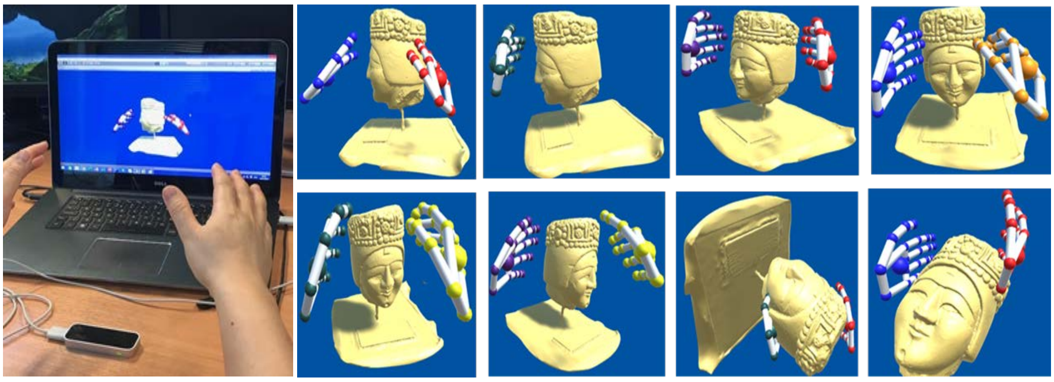


Figure 18. A user using the first proposed visualization and interaction module

- **The second interaction module**

The second interaction module uses three hand gestures, with no virtual hands appearing on the display, to interact with the created 3D objects, and to accomplish 14 movement tasks. As depicted in Figure 19, which shows the second interaction module's diagram, the leap motion controller is used to track the motion of the user's hand based on three hand gestures including an open hand with five fingers together, open hand with outstretched fingers, and closed hand. Each hand gesture is employed to accomplish specific motion tasks of the created 3D object. In this module, first, the leap motion controller extracts information of the three-hand gestures and transmits it to the interaction platform, i.e. the Unity 3D software, to be used by the implemented C# code scripts for the interaction with the targeted 3D object. As presented in Figure 19, the three-hand gestures are used to achieve 14 different motion tasks, six motion tasks for the open hand with five fingers together, seven tasks for the open hand with outstretching fingers, and one motion task for the closed hand gesture. Figure 20 displays screenshots from various interaction types with a 3D head model using the second interaction module.

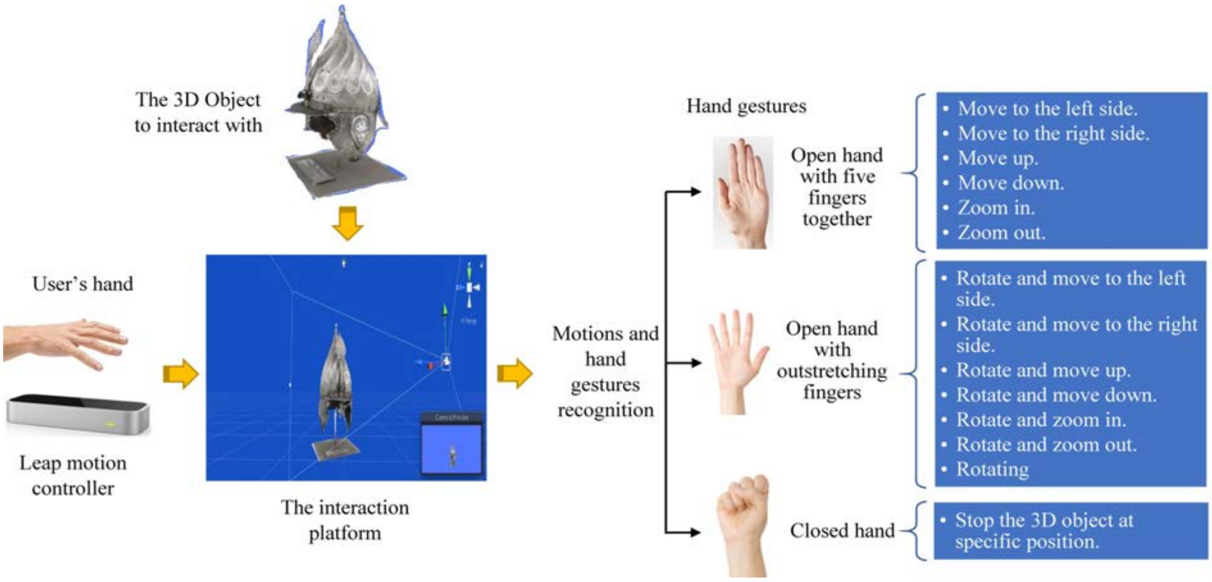


Figure 19. Diagram of the second interaction module

Closed hand (freeze the object movement at first position)

Open hand with five fingers together to Rotate and move to the left

Closed hand (freeze the object movement at a predefined position)

Zoom in and freeze the object movement with the closed hand gesture

Figure 20. A user using the second visualization and interaction module

4 Evaluation and Discussions

In this section, we present the evaluation of our system's visualization and interaction components. The HCI literature has highlighted several techniques for interaction systems evaluation [32,33]. Usability evaluation has been widely suggested to evaluate HCI systems in different application areas [34-37]. This kind of techniques are usually used to provide a degree of confidence in the system's design in the early stage of the development [32]. Unfortunately, to our knowledge, no objective methods exist to evaluate the usability and the quality of HCIs. Hence, to test our proposed 3D visualization and interaction modules, an audience-based testing technique is utilized. In this case, an audience-based technique is the most suitable to evaluate our approach. In this technique, the evaluation of the two proposed interaction modules is based on task analysis. In our evaluation, we mostly considered three aspects related to usability including effectiveness, efficiency, and user satisfaction [38,39].

In the earlier stages of development, we encountered several synchronization issues due to the use of software-based rendering and an improper hardware development platform for Unity3D. Most of these issues were addressed after the use of hardware acceleration as well as a state-of-the-art GPU accelerator.

Thus, in our experiments, a set of users aged between 16 – 58 years, with a gender mix, viewed the digitized cultural asset through the Looking Glass 3D screen and interacted with the 3D models using the Leap Motion controller through the two interaction modules described earlier. In these experiments, the users were tasked to perform different tasks, to test the abilities of both interaction modules. The results are recorded based on three types of measures: (1) the time (*ms*) required to successfully interact with the 3D model for each task corresponding to the two proposed interaction modules, which represents the response time, i.e. the time required by interaction module to react to a user's hand-gesture input; (2) A scaled interaction easiness score reflecting the interaction easiness or difficulty from 1 to 5 with 1 being the hardest and 5 being the easiest; and (3) the users overall satisfaction score with the interaction module using 1 – 100% rating scale with 1% being the lowest satisfaction percentage and 100% being the highest satisfaction percentage. It is worth noting that measure (1) and (2) are evaluated by the team not by user meaning that for measure (2), the team estimates the user's difficulty to accomplish the assigned task.

To evaluate the effectiveness of each interaction module, we evaluated each task on its own to better reflect the user feedback on each task and to evaluate the system functions one by one [40]. Table 1 and Figure 21 report a set of results from 10 users obtained from the task analysis based on the pushing gesture of the first interaction module. The first column refers to the users. The second column to the seventh column describe the time in *ms* to successfully interact with the 3D object for task 1 (T_1) to task 6 (T_6) respectively. The columns eighth to the thirteenth describe the accuracies of the accomplished task 1 (T_1) to task 6 (T_6) using 1 (hard) – 5 (easy) rating scales. The last column describes the overall user's satisfaction percentage. It can be seen from the data in Table 1 that the highest average time required to successfully interact with the 3D object is 4447.2 *ms*, which is the average time needed for T_4; and the minimum average time needed to successfully interact with the 3D object is 3583.6 *ms*, which is the average time required for T_5. The average interaction easiness to accomplish T_1 to T_6 is in the range of 3.1 to 3.5. Also, the average percentage of the overall user's satisfaction is 78%.

Table 1. Results of Pushing gesture of the first interaction module to perform T_1 to T_6

User	**Time (ms) to successfully interact with the 3D object**						**Interaction easiness (1 – 5)**						**Overall user's satisfaction (%)**
	T_1	T_2	T_3	T_4	T_5	T_6	T_1	T_2	T_3	T_4	T_5	T_6	
1	2352	2415	3112	2897	2986	3214	3	3	4	3	4	4	76%
2	3973	4103	3977	3231	4123	5207	2.5	3	3	2	4	2	70%
3	2657	2387	2996	3003	3482	3631	3	3	3	4	3	3	75%
4	3764	3930	4183	5748	3159	2837	3	4	4	3	4	4	78%
5	3372	4177	3917	3294	2896	3635	4	4	3	3	3	2	85%
6	6443	4849	3994	6968	3620	4927	2	3	3	4	3	3	70%
7	2274	3171	4087	3906	2788	4910	3.5	3.5	3	3	4	4	79%
8	3319	4289	4302	4114	3713	5048	4	4	3	3	4	4	80%
9	4106	3949	2909	5184	3690	3729	3	3	3	3	3	3	85%
10	5583	4291	3928	6127	5379	4981	3	3	4	4	3	3	85%
Average	**3784.3**	**3756.1**	**3740.5**	**4447.2**	**3583.6**	**4211.9**	**3.1**	**3.35**	**3.3**	**3.2**	**3.5**	**3.2**	**78%**

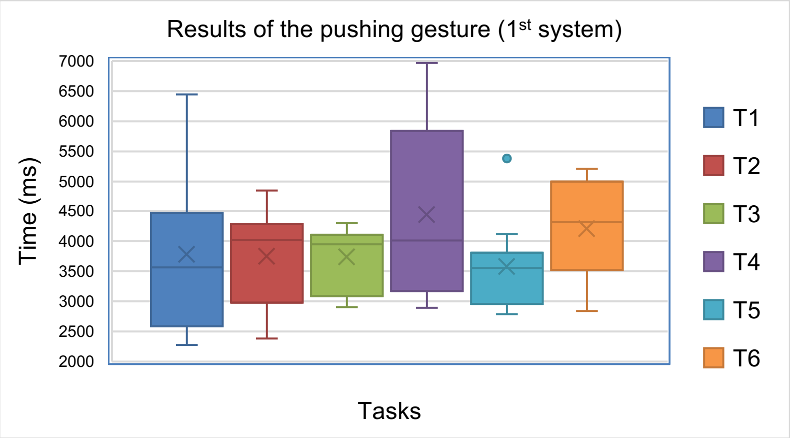


Figure 21. Results Box plot of the pushing gesture (1st system)

Table 2 and Figure 22 present the results achieved from the task analysis based on the grasping gesture of the first interaction module. The first column refers to the users. The second to the eighth columns describe the time in *ms* to successfully interact with the 3D object for task 1 (T_1) to task 7 (T_7) corresponding to the grasping gesture as shown in Figure 17. The ninth to the fifteenth columns describe the accuracies of the accomplished task 1 (T_1) to task 7 (T_7) relating to the grasping gesture using 1 – 5 rating scales. The sixteenth column describes the overall user's satisfaction percentage. The data in Table 2 shows that the maximum average time spent to successfully interact with the 3D object is 11342.8 *ms*, which is the average time needed for T_2; and the minimum average time needed to successfully interact with the 3D object is 10114.4 *ms*, which is the average time necessary for T_7. The average interaction easiness to accomplish T_1 to T_7 of the grasping gesture is in the range of 2.4 to 2.6. Also, the average percentage of the overall user's satisfaction is 71%.

Table 2. Results of the Grasping gesture of the first interaction module to perform T_1 to T_7

User	**Time (*ms*) to successfully interact with the 3D object**							**Interaction easiness (1 – 5)**							**Overall user's satisfaction (%)**
	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_1	T_2	T_3	T_4	T_5	T_6	T_7	
1	8634	9741	9467	8938	7928	8034	7928	2	2	2.5	2.5	2	2	2	65%
2	9378	10953	8329	11027	9510	9946	10137	3	3	2	2	2	2	2	70%
3	8741	8974	10933	10482	8837	8849	11386	2	2	2	2	2	2	2	60%
4	11784	10983	12036	9981	7955	8938	8737	2.5	2.5	3	3	2	2	3	65%
5	9566	10418	11245	12063	9762	10046	11410	2	2	3	3	2.5	2.5	3	75%
6	12079	11493	11382	12945	11137	12731	11199	3	3	3	3	3	3	3	80%
7	11029	10939	9961	9628	13003	11910	10526	3	3	2	2	3	3	3	78%
8	8658	11794	18274	12647	8567	9878	10738	2	2	3	3	2.5	2.5	2	75%
9	13353	15355	9873	8645	10662	12578	9361	3	3	2	2	2	2	3	75%
10	9397	12778	9232	8398	14287	12765	9722	2.5	2.5	3	3	3	3	2	70%
Average	**10261.9**	**11342.8**	**11073.2**	**10475.4**	**10164.8**	**10567.5**	**10114.4**	**2.5**	**2.5**	**2.6**	**2.6**	**2.4**	**2.4**	**2.5**	**71%**

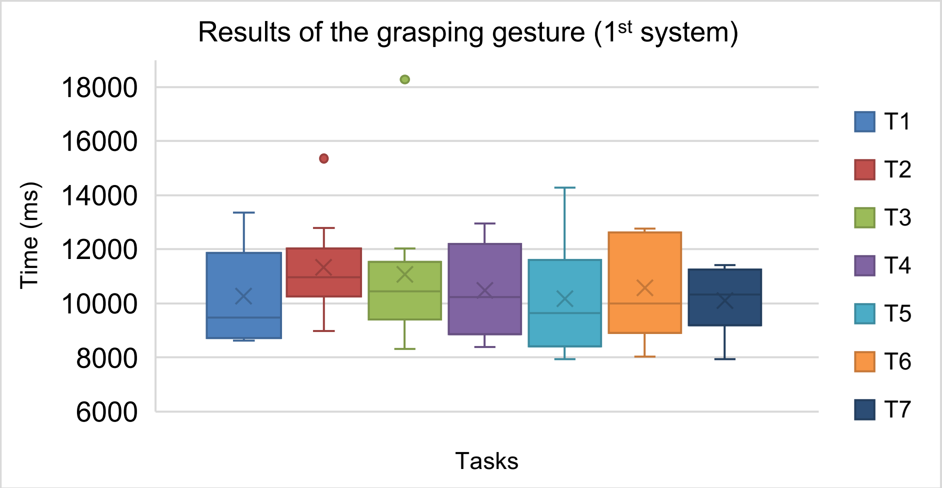


Figure 22. Results box plot of the grasping gesture (1st system)

In order to evaluate the second interaction module, the users performed the three hand gestures, an open hand with five fingers together, an open hand with outstretched fingers, and a closed hand, as shown in Figure 19. These three gestures were used to achieve fourteen tasks of movement of the created 3D object. Table 3 and Figure 23 introduce the results of the time (*ms*) required to effectively interact with the 3D object based on the performance of the users using the second interaction module. The first column refers to the users who took part in our experiments. The second to the fifteenth columns describe the time in *ms* to successfully interact with the 3D object for task 1 (T_1) to task 14 (T_{14}) as shown in Figure 19. The data in Table 3 shows that the maximum average time spent to successfully interact with the 3D object is 1555.2 *ms*, which is the average time needed for T_7 to rotate and move to the left side; and the minimum average time needed to successfully interact with the 3D object is 889.3 *ms*, which is the average time necessary for T_{14}, i.e., to stop the 3D object at a specific position task.

Table 3. Results of time (*ms*) required to effectively interact with the 3D object using the second interaction module

User	**Time (ms) to successfully interact with the 3D object**													
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
1	1128	1817	2111	1301	1318	1915	2709	1987	983	1723	1310	1062	1826	1007
2	1288	1782	1898	2003	1023	1481	2193	1104	1018	897	769	1117	927	1018
3	942	893	973	1873	868	924	1843	734	863	669	584	928	638	938
4	1738	1483	958	1378	896	1674	1745	987	764	929	723	1011	751	889
5	1993	934	773	1207	1782	1211	1398	879	1783	873	1033	1080	809	1033
6	1066	907	1190	1062	913	1564	1403	995	1208	998	1109	905	1123	996
7	1301	986	1176	984	1005	1372	933	884	1121	980	993	1055	1320	791
8	1211	890	1022	762	964	1031	984	1049	1204	996	973	1208	1023	655
9	1109	1029	897	881	979	1067	1128	965	1118	1101	948	1441	773	967
10	993	1405	593	775	894	1109	1216	689	1008	906	778	838	1063	599
Average	**1276.9**	**1212.6**	**1159.1**	**1222.6**	**1064.2**	**1334.8**	**1555.2**	**1027.3**	**1107**	**1007.2**	**922**	**1064.5**	**1025.3**	**889.3**

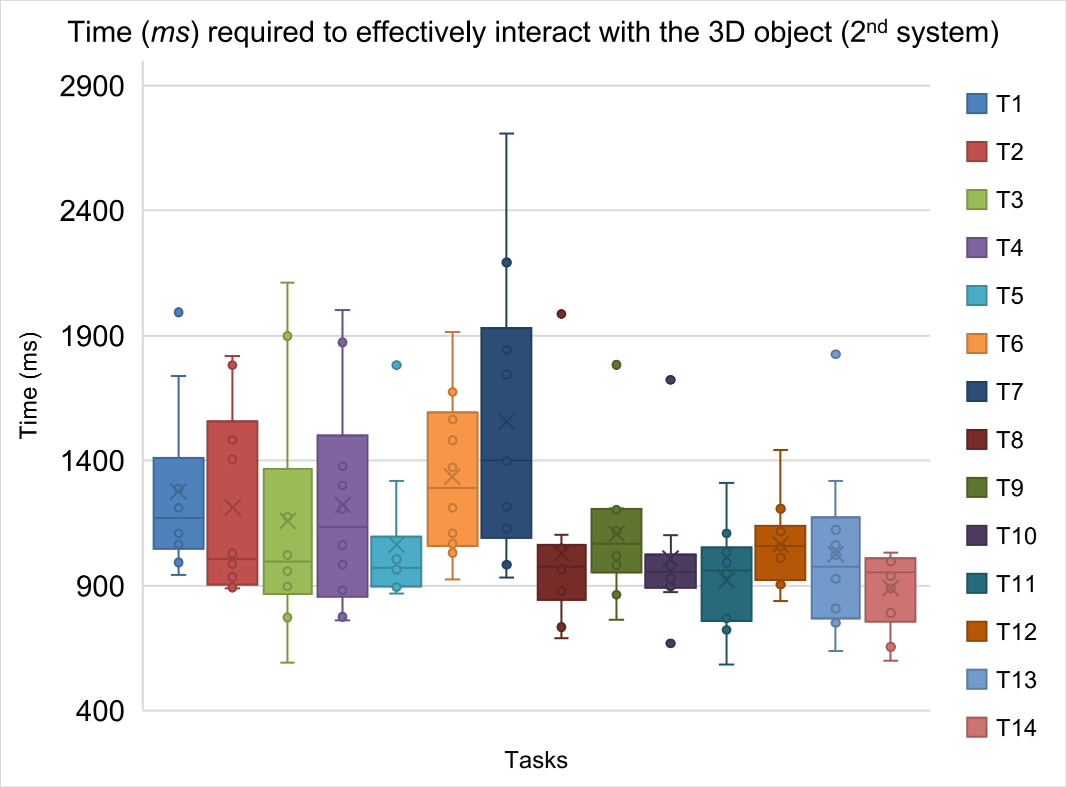


Figure 23. Results box plot of interaction effectiveness (2nd system)

Table 4 illustrates the results of interaction easiness (1 Hard – 5 Easy) for tasks T_1 to T_{14} and the overall user's satisfaction of the second interaction module given by the users in our experiments. Table 4 shows the average interaction easiness to accomplish T_1 to T_{14} of the second interaction module which is in the range of 2.6 to 4.4. Furthermore, the average percentage of the overall user's satisfaction is 92%.

Table 4. Interaction Easiness results (1 – 5) for tasks T_1 to T_{14} and the overall user's satisfaction percentage of the second interaction module

User	**Easiness of use (1 – 5)**														**Overall user's satisfaction (%)**
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	
1	4.5	4.5	4	4	4	4	3	3	3.5	3.5	3	3	3	5	90%
2	5	5	5	5	5	5	3	3	3	3	3	3	3	4	95%
3	4	4	4	4	4	4	4	4	4	4	3	3	3	4	90%
4	3.5	3.5	3.5	3.5	3	3	2	2	2	2	2	2	2	4	95%
5	4	4	4	4	3	3	3	3	3	3	3	3	3	4	90%
6	4	4	4	4	2	2	3	3	3	3	3	2	2	5	95%
7	5	5	5	5	3	3	3	3	3	3	3	3	3	5	95%
8	3	3	3	3	3	3	2	2	2	2	2	2	2	4	80%
9	5	5	5	5	4	4	3	3	3	3	3	2	2	4	90%
10	4	4	4	4	4	4	3	3	3	3	3	3	3	5	95%
Average	**4.2**	**4.2**	**4.15**	**4.15**	**3.5**	**3.5**	**2.9**	**2.9**	**2.95**	**2.95**	**2.8**	**2.6**	**2.6**	**4.4**	**92%**

Taken together, these results suggest that the second interaction module achieved better results, with 92% average overall user's satisfaction, 3.4 average score of all performed tasks, 1133.4 *ms* average time required to effectively interact with the 3D object based on all performed tasks; compared to both techniques of the first interaction module, i.e., pushing and grasping, with 78%, 3.2, 3920.6 *ms* and 71%, 2.5, 1057.4 *ms* respectively. Furthermore, it can be seen from Figure 24 that despite achieving fairly high user's satisfaction results using the pushing gesture of the first interaction module, the second interaction module is able to achieve higher user's satisfaction results, even though more motion tasks (14 tasks) were performed by the users in the experiments of the second interaction module compared to the pushing gesture with only six motion tasks. In summary, the test results indicate that the users found the practice positive and that it was easy for them to interact with the targeted 3D object using the second interaction module.

Furthermore, our two interaction modules offer good results with minimum response time 10.11s for the first interaction system and 0.89s for the second one compared to the fastest response time of 48.8s achieved using the method proposed in [24]. Also, our interaction modules based only on a leap motion controller are more cost-effective and yet effective compared to the existing methods that used a combination of Oculus and leap motion controller [15,41] which often causes nausea as reported by test users of the VR-based system proposed in [15].

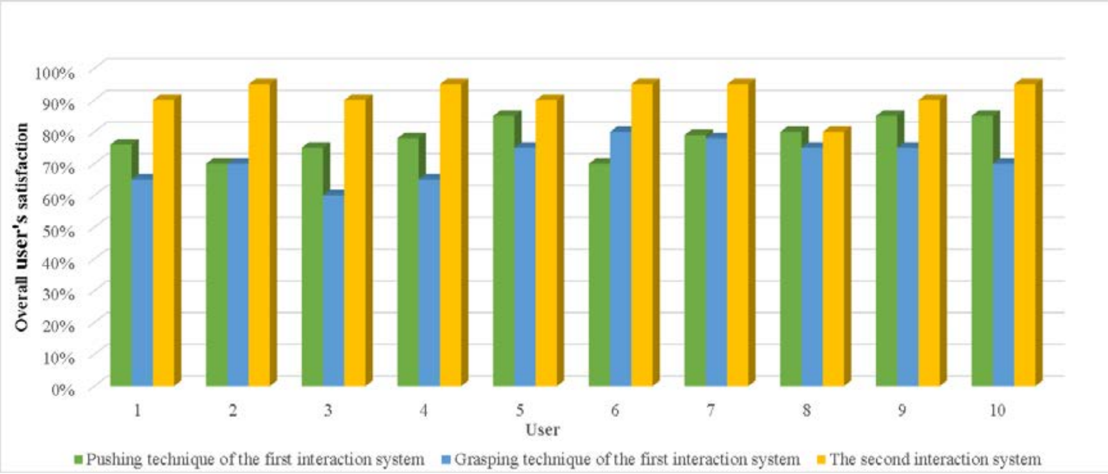


Figure 24. Overall user's satisfaction results using the first and second interaction modules

5 Conclusion

Through this paper, we presented a novel 3D acquisition, visualization, and interaction framework aimed at enabling alternate realities for cultural heritage objects. The framework is mainly geared towards museums where the most important studied use case is the lack of interaction with museum objects often exhibited and protected by glass shells. Through our framework, museum visitors can virtually interact with museum objects without having to deal with sophisticated controllers and without having to wear any kind of headgear.

The main contribution of this paper is the motion interaction and the 3D visualization of 3D cultural heritage in a museum environment. The framework we propose integrates also the adaptation of our prior work in the preprocessing stage where we use motion interpolation and super resolution to increase the quality of generated 3D content.

In our approach, the acquisition of 3D content is made through a custom-made capturing process involving DSLR cameras, a tripod, lighting equipment, and a turntable. With the adaptation of our prior work, the captured images are preprocessed through a deep learning-powered module to upscale their resolution and generate in-between frames minimizing the time required to record the object and increasing the output quality. Using the Unity 3D library, we proposed a novel custom visualization and interaction approach that imports the 3D models, translates users' hands position in space through a continuous tracking data stream coming from the leap motion controller motion interaction device, classifies the input hand tracking data into gestures, and commits the related actions as 3D model movements associated with the user gestures. The output is then rendered in real-time using the HoloPlay service which is the Looking Glass tool needed to convert the 3D content into 45 different views projected on the Looking Glass 3D screen. Experiments were conducted to find the best parameters to minimize delays between the leap motion inputs and the visualization of the result on the 3D display (input delay). Two new interaction modules were proposed. The first is dedicated to inexperienced visitors learning the basic gestures needed to interact with the 3D objects. The second interaction module is dedicated to a more experienced audience and allows more gestures to interact with the objects.

Tests of the framework were undertaken using real museum conditions with the help of museum professionals such as multimedia experts and curators. The framework is considered as a potential solution for future assets exchange between museums around the world thanks to its cost-effectiveness compared to physical assets exchange where insurance costs are far higher than the cost of the hardware and software for a single 3D visualization and interaction setup.

In the future, we aim at exploring the adaptation of motion interaction and real-world 3D visualization for 3D reconstruction to enable new paradigms for cultural heritage virtual reconstruction approaches.

Acknowledgement

This publication was made possible by NPRP grant 9-181-1-036 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors (www.ceproqha.qa).

The authors would also like to thank Mr. Marc Pelletreau, the MIA Multimedia team, the Art Curators and the management staff of the Museum of Islamic art, Doha Qatar for their help and contribution in the data acquisition.

References

1. Belhi A, Bouras A, Alfaqheri T, Aondoakaa AS, Sadka AH (2019) Investigating 3D holoscopic visual content upsampling using super-resolution for cultural heritage digitization. Signal Processing: Image Communication 75:188-198
2. Belhi A, Gasmi H, Bouras A, Alfaqheri T, Aondoakaa AS, Sadka AH, Foufou S Machine Learning and Digital Heritage: The CEPROQHA Project Perspective. In: Fourth International Congress on Information and Communication Technology, 2020. Springer, pp 363-374
3. Belhi A, Bouras A, Foufou S Towards a hierarchical multitask classification framework for cultural heritage. In: 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), 2018. IEEE, pp 1-7
4. Liang W (2019) Scene art design based on human-computer interaction and multimedia information system: an interactive perspective. Multimedia Tools and Applications 78 (4):4767-4785
5. Factory LG Looking Glass Factory. https://lookingglassfactory.com/. Accessed 13-08-2021 2021
6. Weichert F, Bachmann D, Rudak B, Fisseler D (2013) Analysis of the accuracy and robustness of the leap motion controller. Sensors 13 (5):6380-6393
7. Belhi A, Bouras A, Foufou S Digitization and preservation of cultural heritage: The CEPROQHA approach. In: Software, Knowledge, Information Management and Applications (SKIMA), 2017 11th International Conference on, 2017. IEEE, pp 1-7
8. Pieraccini M, Guidi G, Atzeni C (2001) 3D digitizing of cultural heritage. Journal of Cultural Heritage 2 (1):63-70
9. Matterandform.net The Matter and Form Scanner. https://matterandform.net/. Accessed 29-04-2020
10. Singh G (2014) CultLab3D: digitizing cultural heritage. IEEE Computer Graphics and Applications 34 (3):4-5
11. Belhi A, Bouras A, Al-Ali AK, Foufou S (2020) A machine learning framework for enhancing digital experiences in cultural heritage. Journal of Enterprise Information Management:1-13
12. Pavlidis G, Koutsoudis A, Arnaoutoglou F, Tsioukas V, Chamzas C (2007) Methods for 3D digitization of cultural heritage. Journal of cultural heritage 8 (1):93-98
13. Scopigno R, Callieri M, Cignoni P, Corsini M, Dellepiane M, Ponchio F, Ranzuglia G (2011) 3D models for cultural heritage: beyond plain visualization. Computer (7):48-55
14. Yang L, Dong H, Alelaiwi A, El Saddik A (2016) See in 3D: state of the art of 3D display technologies. Multimedia Tools and Applications 75 (24):17121-17155
15. Gonizzi Barsanti S, Caruso G, Micoli L, Covarrubias Rodriguez M, Guidi G 3D visualization of cultural heritage artefacts with virtual reality devices. In: 25th International CIPA Symposium 2015, 2015. vol 5W7. Copernicus Gesellschaft mbH, pp 165-172
16. Boher P, Leroux T, Bignon T, Collomb-Patton V Optical characterization of different types of 3D displays. In: Advances in Display Technologies II, 2012. International Society for Optics and Photonics, p 82800A
17. Agus M, Bettio F, Giachetti A, Gobbetti E, Guitián JAI, Marton F, Nilsson J, Pintore G (2009) An interactive 3D medical visualization system based on a light field display. The Visual Computer 25 (9):883-893
18. De la Barré R, Jurk S Improvements of aid free 3D Presentation. In: Electronic Displays Conference, Nuremberg, 2011.

19. Bastanlar Y, Grammalidis N, Zabulis X, Yilmaz E, Yardimci Y, Triantafyllidis G (2008) 3D reconstruction for a cultural heritage virtual tour system. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing 37:1023-1028
20. Rojas-Sola JI, Castro-Garcia M, del Pilar Carranza-Canadas M (2011) Content management system incorporated in a virtual museum hosting. Journal of Cultural Heritage 12 (1):74-81
21. Belhi A, Al-Ali AK, Bouras A, Foufou S, Yu X, Zhang H (2020) Investigating low-delay deep learning-based cultural image reconstruction. Journal of Real-Time Image Processing 17 (6):1911-1926
22. Han J, Shao L, Xu D, Shotton J (2013) Enhanced computer vision with microsoft kinect sensor: A review. IEEE transactions on cybernetics 43 (5):1318-1334
23. Bachmann D, Weichert F, Rinkenauer G (2018) Review of three-dimensional human-computer interaction with focus on the leap motion controller. Sensors 18 (7):2194
24. Vosinakis S, Koutsabasis P, Makris D, Sagia E A kinesthetic approach to digital heritage using leap motion: the Cycladic sculpture application. In: 2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES), 2016. IEEE, pp 1-8
25. Mustafa AW, Ismail AF 3D virtual pottery environment using hand gesture interaction. In: Proceedings International Conference on Virtual Systems and Multimedia, 2018. pp 1-6
26. Barbieri L, Bruno F, Muzzupappa M (2017) Virtual museum system evaluation through user studies. Journal of Cultural Heritage 26:101-108
27. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z (2017) Photo-realistic single image super-resolution using a generative adversarial network. arXiv preprint
28. Tomar S (2006) Converting video formats with FFmpeg. Linux Journal 2006 (146):10
29. Jiang H, Sun D, Jampani V, Yang M-H, Learned-Miller E, Kautz J Super slomo: High quality estimation of multiple intermediate frames for video interpolation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. pp 9000-9008
30. Buyuksalih I, Bayburt S, Buyuksalih G, Baskaraca A, Karim H, Rahman AA (2017) 3D Modelling and Visualization Based on the Unity Game Engine–Advantages and Challenges. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4:161
31. Spring AP, Peters C, Minns T (2010) Using mid-range laser scanners to digitize cultural-heritage sites. IEEE Comput Graph Appl 30 (3):15-19. doi:10.1109/MCG.2010.62
32. Ghaoui C (2005) Encyclopedia of human computer interaction. IGI Global,
33. Preece J, Rogers Y, Sharp H, Benyon D, Holland S, Carey T (1994) Human-computer interaction. Addison-Wesley Longman Ltd.,
34. Gray WD, Salzman MC (1998) Damaged merchandise? A review of experiments that compare usability evaluation methods. Human–computer interaction 13 (3):203-261
35. Rubin J, Chisnell D (2008) Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons,
36. Ssemugabi S, De Villiers R A comparative study of two usability evaluation methods using a web-based e-learning application. In: Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries, 2007. pp 132-142
37. Freire LL, Arezes PM, Campos JC (2012) A literature review about usability evaluation methods for e-learning platforms. Work 41 (Supplement 1):1038-1044
38. Yanez-Gomez R, Cascado-Caballero D, Sevillano J-L (2017) Academic methods for usability evaluation of serious games: a systematic review. Multimedia Tools and Applications 76 (4):5755-5784
39. Jacko JA (2012) Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications. CRC press,
40. Diaper D, Stanton N (2003) The handbook of task analysis for human-computer interaction. CRC Press,
41. Drossis G, Birliraki C, Stephanidis C Interaction with immersive cultural heritage environments using virtual reality technologies. In: International Conference on Human-Computer Interaction, 2018. Springer, pp 177-183