Article

Clouds-Based Collaborative and Multi-Modal Mixed Reality for Virtual Heritage

Mafkereseb Kassahun Bekele

School of Media, Creative Arts and Social Inquiry (MCASI), Curtin University, Perth, WA 6102, Australia; mafkereseb.bekele@postgrad.curtin.edu.au or mafkereseb@hotmail.com

Abstract: Recent technological advancements in immersive reality technologies have become a focus area in the virtual heritage (VH) domain. In this regard, this paper attempts to design and implement clouds-based collaborative and multi-modal MR application aiming at enhancing cultural learning in VH. The design and implementation can be adopted by the VH domain for various application themes. The application utilises cloud computing and immersive reality technologies. The use of cloud computing, collaborative, and multi-modal interaction methods is influenced by the following three issues. First, studies show that users' interaction with immersive reality technologies and virtual environments determines their learning outcome and the overall experience. Second, studies also demonstrate that collaborative and multi-modal interaction methods enable engagement in immersive reality environments. Third, the integration of immersive reality technologies with traditional museums and cultural heritage sites is getting significant attention in the domain. However, a robust approach, development platforms (frameworks) and easily adopted design and implementation approaches, or guidelines are not commonly available to the VH community. This paper, therefore, will attempt to achieve two major goals. First, it attempts to design and implement a novel application that integrates cloud computing, immersive reality technology and VH. Second, it attempts to apply the proposed application to enhance cultural learning. From the perspective of cultural learning and users' experience, the assumption is that the proposed approach (clouds-based collaborative and multi-modal MR) can enhance cultural learning by (1) establishing a contextual relationship and engagement between users, virtual environments and cultural context in museums and heritage sites, and (2) by enabling collaboration between users.

Keywords: mixed reality; cloud computing; virtual heritage; collaborative interaction; multi-modal interaction

1. Introduction

Interactive digital technologies, alongside various interpretive multimedia approaches, have recently become a common appearance in traditional museums and cultural heritage sites. These technologies are enabling museums to disseminate cultural knowledge and enrich visitors' experiences with engaging and interactive learning. Studies have also shown that collaboration, interaction, engagement, and contextual relationship are the key aspects to determine the effectiveness of virtual reality applications from a cultural learning perspective [1–5].

Specifically, immersive reality technologies, such as augmented reality (AR), virtual reality (VR), and mixed reality (MR) enable the creation of interactive, engaging, and immersive environments where user-centered presentation of digitally preserved heritage can be realised. Cultural heritage, more specifically the VH domain, has been utilising these technologies for various application themes [6]. For instance, the ARCHEOGUIDE is a typical example of one of the earliest adoptions of the technology with a well-defined goal of enhancing visitors' experience at heritage sites [7].

Interaction with virtual content presented in VH applications is an essential aspect of immersive reality that has a defining impact on the meaningfulness of the virtual en-

vironment. In this regard, studies in the VH domain have demonstrated how interaction methods play a role in terms of enhancing engagement, contextual immersivity, and meaningfulness of virtual environments. These characteristics are crucial aspects of interaction in VH for enhancing cultural learning. There are six interaction methods that VH commonly adopt: tangible, collaborative, device-based, sensor-based, hybrid, and multi-modal interfaces [8]. Collaborative interfaces often use a combination of complementary interaction methods, sensors, and devices to enable co-located and/or remote collaboration among users. Multi-modal interfaces are a fusion of two and more sensors, devices, and interaction techniques that sense and understand humans' natural interaction modalities. This interface group allows gestural, gaze-based, and speech-based interaction with virtual content. Combining collaborative and multi-modal interaction methods with MR allows multiple users to interact with each other (social presence) and with a shared real-virtual space (virtual presence). This combination, therefore, results in a space that enables collaboration and multi-modal interaction with the real-virtual environment, thereby resulting in a scenario where both social and virtual presence can be achieved. The interaction method proposed in this paper, therefore, attempts to bring collaborative and multi-modal MR to museums and heritage sites so that users will be able to interact with virtual content at the natural location of the heritage assets that are partially or fully represented in the virtual environment.

Some of the technical obstacles that immersive reality applications face include the cost associated with owning the technology, computational and rendering resources, and the level of expertise required to implement and maintain the technology and its underlying infrastructure. This paper, therefore, proposes to utilise cloud computing to tackle these difficulties.

The application proposed in this paper "clouds-based collaborative and multi-modal MR" introduces a novel approach to the VH community, museum curators, and cultural heritage professionals to inform the practical benefits of cloud computing to function as a platform for the implementation of immersive reality technologies. The novelty of this approach is that it integrates cloud computing, multiple interaction methods and MR while aiming at cultural learning in VH. Here, we would like to note the deliberate use of the term "clouds" in a plural form to signify that the application attempts to utilise cloud services from multiple providers. The proposed application is motivated by: (1) cloud computing technology's ability for fast development/deployment and elasticity of resources and services [9], (2) the ability of collaborative and multi-modal interaction methods to enhance cultural learning in VH as demonstrated by the domain's existing studies, for instance [10], and (3) the continuous improvement of natural interaction modalities and MR devices. The contributions of the application are, therefore:

Ultimately, the success of a VH application is determined by its effectiveness to communicate the cultural significance and values of heritage assets. Enhancing this knowledge communication/dissemination process is the primary motivation behind the proposed application. Hence, this paper contributes to VH applications, especially where cultural learning is at the centre of the application design and implementation process.

Cloud computing is a relatively new area in computing. As a result, it is not common to find cloud-based systems and applications in the cultural heritage domain. Similarly, cloud-based immersive reality applications are rare in VH. This paper, therefore, serves as one of the few early adoptions of cloud computing as a platform for immersive reality implementations in VH.

Studies show that VH applications and their virtual environments are not often preserved after their implementation [11]. Cloud computing will play a major role in preserving VH applications and their virtual environments for a longer period if cloud resources are maintained for this purpose. The proposed approach will attempt to preserve both the application and digital resources via an institutional repository.

Interaction and engagement with a given virtual environment in VH determine whether users can acquire knowledge and understand the significance of cultural her-

itage assets that the application is attempting to communicate. To this end, the proposed application will attempt to balance interaction and engagement with the technology and cultural context through collaborative and multi-modal interaction methods.

The remainder of this paper is organised as follows. Section 2 will discuss existing literature and exemplar VH applications that mainly utilise immersive reality technology, cloud computing, and adopt collaborative and multi-modal interaction interfaces. Section 3 will provide a detailed discussion on the system architecture proposed in this paper. Following that, Section 4 will explain the implementation phase in detail from a technical perspective. Section 5 will present a discussion on the built prototype focusing on the expected impact of the application on cultural learning in museum settings and provides discussion on identified limitations of the application. Finally, Section 6 will summarise the paper and will discuss future works and provides suggestions on parts of the system architecture that need improvement.

2. Related Works

With the advent of MR, recent developments in the presentation aspect of VH show that this technology has the potential of becoming the dominant member of immersive reality technologies, especially, when the main goal of the applications under consideration is delivering an engaging and interactive real-virtual environment [8,12]. However, several technical difficulties associated with the technology are preventing it from a wider adoption across domains and application themes. One of these technical challenges is the computational resources that immersive reality devices are required to be equipped with. For instance, the resources required for mobile augmented reality applications are often available on the same mobile device. As such, mobile augmented reality applications are widely available [13–17]. On the other hand, fully immersive and interactive MR applications are difficult to find. This is because MR applications are resource-intensive. Such applications often involve heavy graphical computations, rendering, and very low latency to deliver an engaging and interactive experience to the end-user [18–21]. In this regard, serval studies have demonstrated the potential of cloud computing to meet the computational demands of various application themes in the cultural heritage domain. For instance, a recent study by Abdelrazeq and Kohlschein [9] proposed a modular cloud-based augmented reality development framework that is aimed at enabling developers to utilise existing augmented reality functionalities and shared resources. Moreover, the proposed framework supports content and context sharing to enable collaboration between clients connected to the framework. Another study by Fanini and Pescarin [22] presented a cloud-based platform for processing, management and dissemination of 3D landscape datasets online. The platform supports desktops and mobile devices and allows collaborative interaction with the landscape and 3D reconstructions of archaeological assets.

Similarly, a study by Malliri and Siountri [23] proposed an augmented reality application that utilises 5G and cloud computing technologies aimed at presenting underwater archaeological sites, submerged settlements and shipwrecks to the public in a form of virtual content. Yang, Hou [24] also proposed a cloud platform for cultural knowledge construction, management, sharing, and dissemination.

A recent study by Tocźe and Lindqvist [18] presented an MR prototype leveraging edge computing to offload the creation of point cloud and graphic rendering to computing resources located at the network edge. As noted in the study, edge computing enables placing computing resources closer to the display devices (such as MR devices), at the edge of the network. This feature enables applications that are too resource-intensive to be run closer to the end device and streamed with low latency. Besides computing and graphical rendering resources that enable effective MR experience, engaging interaction with virtual environments is equally significant.

Collaborative interaction methods enable a multiuser interaction with a shared and synchronised virtual and/or real-virtual environment. As a result, this interaction method can easily establish a contextual relationship between users and cultural context by adding

a social dimension to the experience. In this regard, a study by Šašinka and Stachoň [25] indicates the importance of adding a social dimension to the knowledge acquisition process in collaborative and interactive visualisation environments. Multi-modal interaction methods integrate multiple modes of interaction, such as speech, gaze, gesture, touch, and movement. This interaction method resembles how we interact with our physical environment and enables users to establish a contextual relationship and collaboratively interact with the real virtual environment. As a result, enhanced engagement with virtual environments and cultural context can be realised due to the method's ease of use and resemblance to natural interaction modalities.

3. System Architecture and Components of the Clouds-Based Collaborative and Multi-Modal MR

This work is a continuation of a previous design and implementation of a "Walkable MR Map" that employed a map-based interaction method to enable engagement and contextual relationship in VH applications [26]. The Walkable MR Map is an interaction method designed and built to use interactive, immersive, and walkable maps as interaction interfaces in a MR environment that allows users to interact with cultural content, 3D models, and different multimedia content at museums and heritage sites. The clouds-based collaborative and multi-modal MR application utilises this map-based approach as a base interaction method and extends the interactivity aspect with a collaborative and multi-modal characteristic. The resulting virtual environment allows a multiuser interaction.

The proposed application has five major components: MR device, Collaborative and Multi-modal Interaction Framework (CMIF), Walkable MR Map Framework, Cultural and Multimedia Content Manager (CMCM), and Shared Location and Session Manager (SLSM). Figure 1 shows the overall system architecture and its major components. A detailed discussion on the architecture is provided below. The discussion will focus mainly on the components that are newly added to the "Walkable MR Map" framework. However, a brief introduction to some components of this existing framework will be provided, when possible, to make the reading smooth.
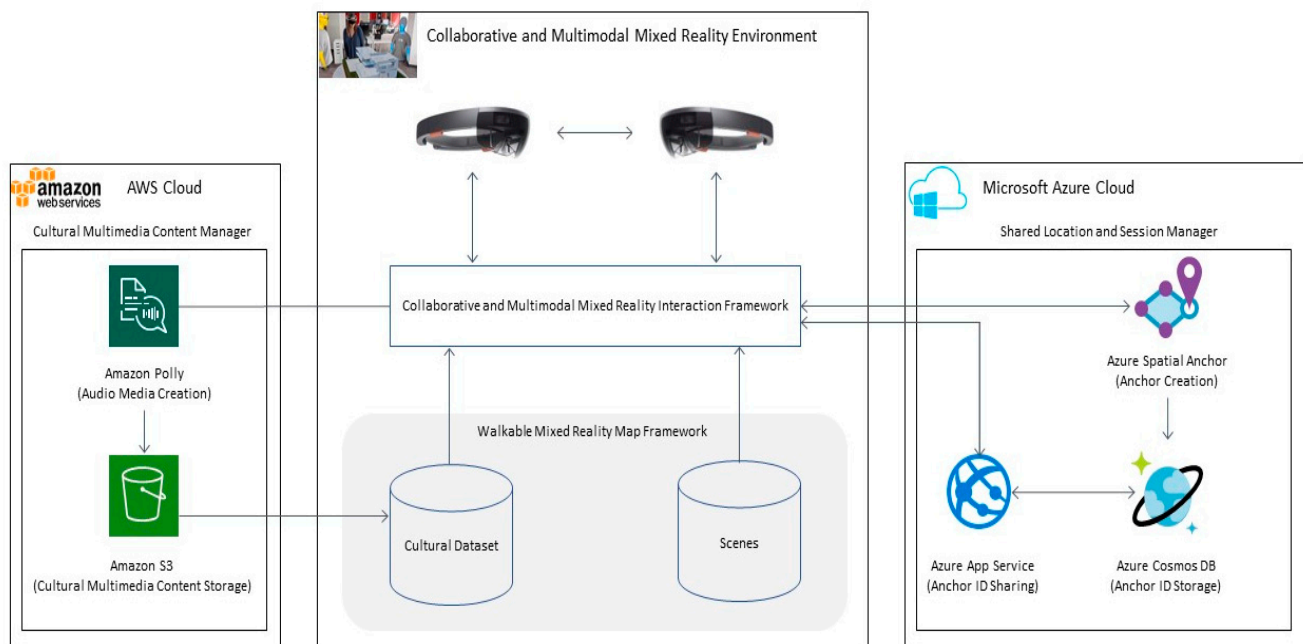

Figure 1. System architecture of the clouds-based collaborative and multi-modal mixed reality (MR).

3.1. MR Devices (Microsoft HoloLens)

The proposed architecture uses a minimum of two Microsoft HoloLens devices to enable collaborative and multi-modal interaction with a shared real-virtual environment.

HoloLens is preferred over other immersive reality devices available in the market because it has inbuilt processing units, tracking, interaction (gaze, gesture, and speech), and rendering capabilities. It also can be integrated with Microsoft Azure Mixed Reality Services[1], such as Azure Spatial Anchors. These features are the enablers of the collaborative and multi-modal aspects of the proposed framework. A detailed discussion of the technical specifications of HoloLens and a complete workflow of application development and deployment are presented in [26,27].

3.2. Collaborative and Multi-Modal Interaction Framework (CMIF)

This component is central to the overall objective of the framework. It builds upon and extends the map-based interaction in the "Walkable MxR Map" architecture. In addition to the interactivity provided, this component introduces collaborative and multi-modal interaction to the experience. To this end, the component integrates with other parts of the framework, namely, SLSM and CMCM.

3.3. Walkable MxR Map Framework

The Walkable MxR Map has five major components: Head-Mounted Display, Geospatial Information and Event Cue, Interaction Inputs and MR (MxR) Framework, Event and Spatial Query Handler, and Cultural Dataset containing historical and cultural context (3D models, multimedia content and event spatiotemporal information). Figure 2 shows the overall system architecture and its major components. A detailed discussion on the overall architecture and on each component is presented in [26].
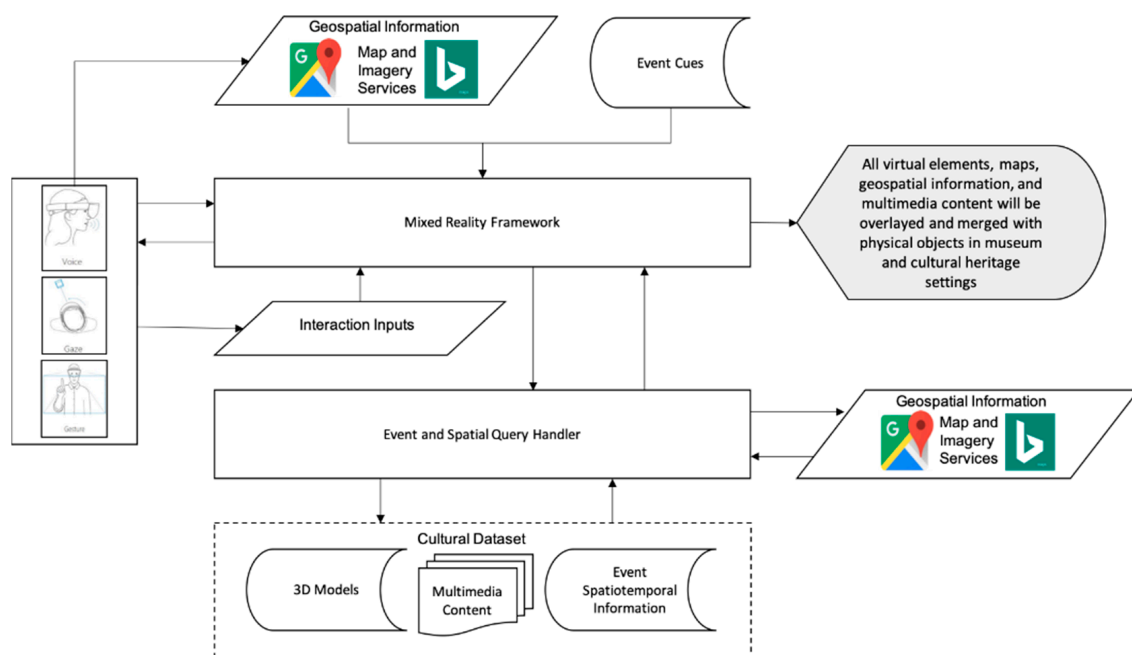

Figure 2. System architecture of the Walkable MR Map [26].

3.4. Cultural Multimedia Content Manager (CMCM)

Cultural Multimedia Content Manager (CMCM) is responsible for the creation, storage, and dissemination of all cultural multimedia content, such as audio and video content. It has two subcomponents, namely, Audio Media Creation (AMC) and Cultural Multimedia Content Storage (CMCS). This component plays a significant role in terms of reducing the MR application size that will be deployed to the HoloLens devices. The deployable application size often gets bigger given that the device is untethered, and developers tend to utilise the storage, processing and rendering capability onboard of the device. As such, their performance will be impacted as the deployed applications need to load all

content at run time. The architecture proposed in this paper uses the CMCM to move the storage of such content into cloud-based storage and load specific content at run time to the application using API calls. Hence, the deployable application size can be reduced greatly and allow for more multimedia content sharing as the application will not be limited to the storage size onboard the device. This provides opportunity and flexibility to the application in terms of sharing a wider range of content.

The CMCM relies mainly on two cloud services from Amazon Web Services—Amazon Polly and Amazon S3. Amazon Polly[2] is a cloud service that turns text into lifelike speech. Amazon Simple Storage Service[3] (Amazon S3) is an object storage service that offers scalability, data availability, security, and performance. This cloud storage is used to store content generated by the AMC subcomponent and other multimedia content from external sources. In addition, these cloud services are utilised for two main purposes. First, Amazon Polly is used to generate audio files that enable the proposed framework to include speech-enabled interaction with content and the MR environment. This is achieved by converting textual information sourced from different historical collections and instructions into lifelike speech using Amazon Polly. The converted audio media files are then stored in Amazon S3 and made available for the CMCM to load to a scene at run time. As a result, users are guided through the MR experience by speech-enabled instructions and interactive content. Second, Amazon Polly is also used to generate audio media files in a form of narration about specific cultural contexts presented through the application. The multi-modal aspect of the interaction method relies on this feature.

3.5. Shared Location and Session Manager (SLSM)

This is a crucial component to enable sharing the MR environment across multiple devices, thereby achieving a collaborative experience. To this end, the SLSM component relies on Azure Spatial Anchors[4], Azure Cosmos DB[5], and Azure App Service[6].

Azure Spatial Anchors is a cross-platform service that allows developers to create multi-user and shared experiences using objects that persist their location across devices over time. For example, two people can start a MR application by placing the virtual content on a table. Then, by pointing their device at the table, they can view and interact with the virtual content together. Spatial Anchors can also connect to create spatial relationships between them. For example, a VH application used in museums as a virtual tour guide or wayfinding assistant may include an experience that has two or more points of interest that a user must interact with to complete a predefined visit route in the museum. Those points of interest can be stored and shared across sessions and devices. Later, when the user is completing the visiting experience, the application can retrieve anchors that are nearby the current one to direct the user towards the next visiting experience at the museum.

In addition, Spatial Anchors can enable persisting virtual content in the real world. For instance, museum curators can place virtual maps on the floor or wall, that people can see through a smartphone application or a HoloLens device to find their way around the museum. Hence, in a museum or cultural heritage setting, users could receive contextual information about heritage assets by pointing a supported device camera at the Spatial Anchors.

Azure Cosmos DB is a fully managed, scalable, and highly responsive NoSQL database for application development. The SLSM component uses Azure Cosmos DB as persistent storage for Spatial Anchors identifiers. This service is selected because it is easy to integrate with other Microsoft Azure cloud services. The stored identifiers are then accessed and shared across sessions via Azure App Service. Azure App Service is an HTTP-based service for hosting web applications, REST APIs, and mobile back ends.

4. Implementation of "Clouds-Based Collaborative and Multi-Modal MR"

The system architecture proposed above was implemented using commercial Head-mounted-display (HMD), proprietary software, cloud services, opensource development toolkits, and custom scripts.

4.1. Head-Mounted-Display–Microsft HoloLens

Microsoft HoloLens is a self-contained immersive reality device that runs on Windows MR operating system. The main features of immersive reality devices, including Microsoft HoloLens, can be categorised into four enabler groups: tracking, experience scale, interaction, and spatial awareness [26]. This device has all the features required for tracking, computation, and presenting virtual objects and audio-visual elements. Most HMDs that are currently available in the market, however, rely on high-end VR-ready computers for computation, meaning they always must be physically attached to computing resources or at least connect wirelessly. HoloLens, however, performs all computations using the processing units (CPU and GPU) onboard the device. This extends the applicability of the application to indoor and outdoor settings.

4.2. Cloud Services

As discussed in the previous section, there are five cloud services that have been utilised to build the prototype. A brief introduction to these could services and their overall role in the proposed application has been provided in Section 3. The discussion below focuses on specific features of these services used during the implementation (prototype) stage. Here, it is worth noting that the proposed application can be reproduced using services and platforms from any provider if the technical requirements discussed are met.

1. Amazon Polly is mainly used to convert textual content to lifelike speech. For instance, the built prototype has a scene that introduces the MR device (HoloLens) and interaction methods. This introductory scene informs on how to interact with virtual content and the device using gaze and gestures, such as Air tap and Bloom gestures[7].
2. Amazon S3 is used to store all types of media files the application requires. The files are then retrieved and loaded to the HoloLens device at runtime.
3. Azure Spatial Anchors is primarily used to enable sharing the MR experience across HoloLens devices. The sharing can potentially extend to a range of smartphone devices. The prototype presented in this article, however, targets HoloLens devices.
4. Azure Cosmos DB is used to store Azure Spatial Anchors identifiers for persistent sharing. This service was selected given that it can easily integrate with Azure App Service. Alternatively, any other database, such as Amazon DynamoDB[8], can be used to store the identifiers generated by Azure Spatial Anchors.
5. Azure App Service is used as a web application to post and retrieve anchor identifiers from Azure Cosmos DB.

4.3. Development Platform and Toolkits

The proposed system was prototyped using Unity 2019.4.x, Mixed Reality Toolkit 2.5, Azure Spatial Anchors SDK 2.7, and Microsoft Visual Studio 2019. The Unity game engine was selected because it enables multiplatform game development and deployment. Currently, Unity supports more than 25 platforms. Mixed Reality Toolkit (MRTK) is a Microsoft-driven open source and extensible framework that provides a set of foundational components and features to accelerate MR developments in Unity. The version of MRTK used in this implementation supports a wide range of platforms, such as Microsoft HoloLens 1 and HoloLens 2, Microsoft Immersive Headsets, Windows Mixed Reality Headsets, and OpenVR headsets (HTC Vive/Oculus Rift). Thanks to this toolkit, rapid prototyping via in-editor simulation that allows seeing changes immediately, was achieved. Once the development was ready to deploy onto target devices. Microsoft Visual Studio 2019 was used for debugging and deploying the code from Unity to HoloLens.

5. Discussion

The implementation of the clouds-based MR architecture proposed above was realised using the tools, technologies and cloud services discussed in the previous sections. This section will provide a detailed discussion on the built prototype, experiential aspects of the

application, limitations and areas identified for future improvement. It also provides a brief historical context of SS Xantho[9], the heritage asset used as a case study for the prototype.

5.1. The Story of SS Xantho: Western Australia's First Coastal Steamer (1848–1872)

Note: the following text is extracted and compiled from materials published by Western Australia Museum [28–30].

The paddle steamer Xantho, one of the world's first iron ships, was built in 1848 by Denny's of Dumbarton in Scotland. Like most 19th century steamships, Xantho was driven by both sails and steam. In 1871, after 23 years of Scottish coastal service, Xantho was sold to Robert Stewart, 'Metal Merchant' (scrap metal dealer) of Glasgow. Rather than cut it up for scrap, he removed the old paddlewheel machinery and replaced it with a ten-year-old propeller engine built by the famous naval engineers, John Penn and Sons of Greenwich. Stewart then offered the 'hybrid' ship for sale.

Xantho's new owner was the colonial entrepreneur Charles Edward Broadhurst, who visited Glasgow partly to purchase a steamer to navigate Australia's northwest. In November 1872 on her way south from the pearling grounds, Xantho called in to Port Gregory, and there, ignoring his captain's pleas, Broadhurst overloaded his ship with a cargo of lead ore. On the way south to Geraldton the worn-out SS Xantho began to sink. Soon after entering Port Gregory, they hit a sandbank and the water already in the ship tore through three supposedly watertight bulkheads, entered the engine room, and doused the boiler fires. This rendered the pumps inoperable, and the ship slowly sank, coming to rest in 5 metres of water, about 100 metres offshore.

In 1979, when searching for the Xantho for Graeme Henderson, the Western Australia Museum's head of the colonial wreck program who was researching the very late transition from sail to steam on the Western Australia coast, volunteer divers from the Maritime Archaeological Association of Western Australia (MAAWA) were led to what they knew as the 'boiler wreck' by Port Gregory identities Robin Cripps and Greg Horseman. A wreck report from the MAAWA team was filed, together with artist Ian Warne's impressions showing how the wreck had disintegrated over the years. Figure 3 presents impressions showing how the wreck of SS Xantho disintegrated.

In 1983, following reports of looting at the site, the task of examining and protecting the site was given to the Museum's Inspector of Wrecks, M. (Mack) McCarthy, who has coordinated all aspects of the project ever since.

5.2. Interacting with the Clouds-Based Collaborative and Multi-Modal MR Application

The prototype built as part of the clouds-based MR application has two different flavours or versions—Xantho-Curator and Xantho-Visitor (See Table 1). The features and modes of interaction are slightly different between these two versions of the prototype.

Table 1. Features of Xantho-Curator and Xantho-Visitor versions of the clouds-based MR application.

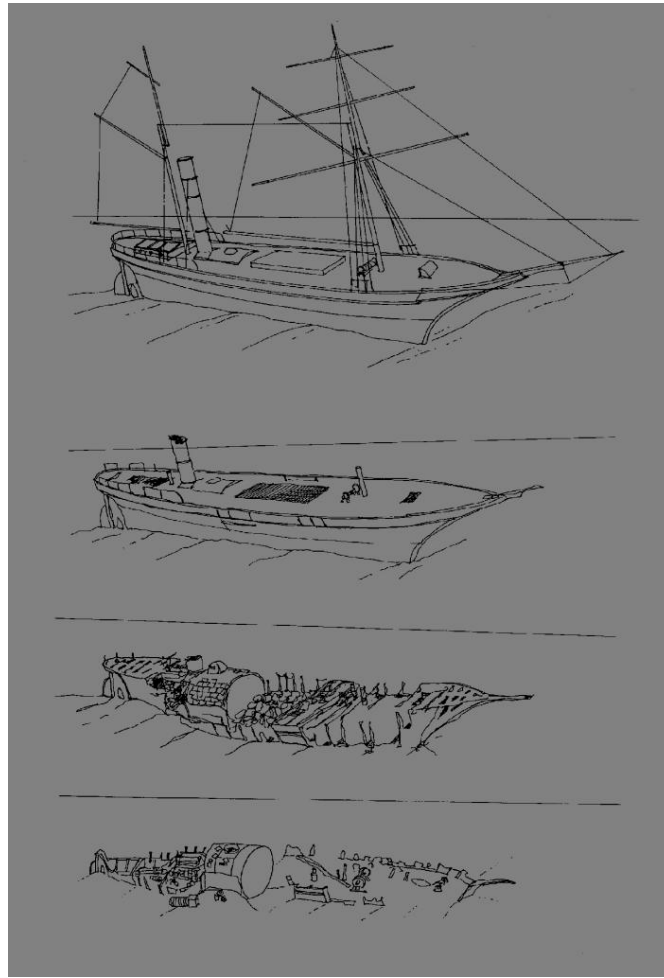| Feature	Xantho-Curator	Xantho-Visitor
Create Spatial Anchor object (Stage ID)	Xantho-Curator can create, store, and retrieve Spatial Anchor objects and their identifiers	Xantho-Visitor cannot create Spatial Anchor objects
Locate Spatial Anchor object (Stage ID)	Xantho-Curator can locate previously created and stored Spatial Anchor objects	Same as Xantho-Curator
Modes of interaction	Xantho-Curator can use gaze, gesture, and speech to interact with the virtual environment	Same as Xantho-Curator
Introductory Scene	The assumption is that Xantho-Curator users are not new to the HoloLens device. Hence, this version does not include an introductory scene.	Xantho-Visitor loads the introductory scene (focusing on modes of interaction) when the experience begins. This is assuming users are new to the HoloLens device.

Figure 3. Artist Ian Warne's impressions showing how the wreck of SS Xantho disintegrated. Image courtesy (Ian Warne, MAAWA).

Both Xantho-Curator and Xantho-Visitor allow users to interact with the display device (HoloLens) via a combination of gaze, gesture, and speech. In terms of functionality, Xantho-Curator provides a unique feature to create Spatial Anchor objects that will serve as a stage for the collaborative MR environment.

Interaction with the collaborative MR environment always begins with the Xantho-Curator version user either retrieving or creating the Spatial Anchor object that servers as a shared stage for the experience. Once a stage is identified (Spatial Anchor object created and/or located), an identical MR environment will be loaded to the HoloLens device for both users (Xantho-Curator and Xantho-visitor) to interact with. This plays a significant role to augment users' interactive experience with a sense of collaboration and engagement. Figures 4 and 5 show each step of Xantho-Curator and Xantho-Visitor, respectively. Interested readers can visit the link provided under Supplemental Materials section to access video file that shows how the built prototype functions.

*5.3. Expected Impact of the Clouds-Based MR Application on Cultural Learning*

The key theoretical background that led to designing and building a clouds-based collaborative and multi-modal MR application is the assumption that contextual relationship and engagement lead to enhanced cultural learning in VH environments. It is further assumed that VH environments can be augmented with these properties via a combination of immersive reality technology and collaborative and multi-modal interaction methods.
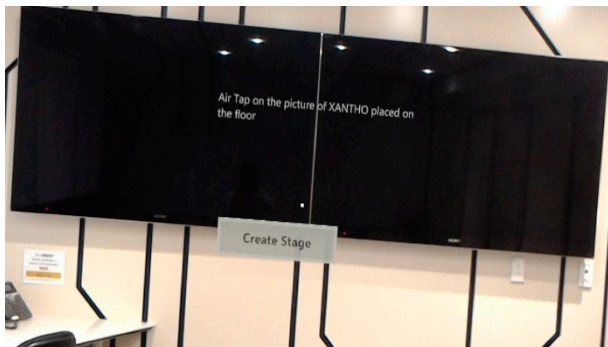
(**a**) Xantho-Curator version asks the user to air tap on the picture of SS Xantho placed on the floor. This location will be used as a stage to load the MxR environment.
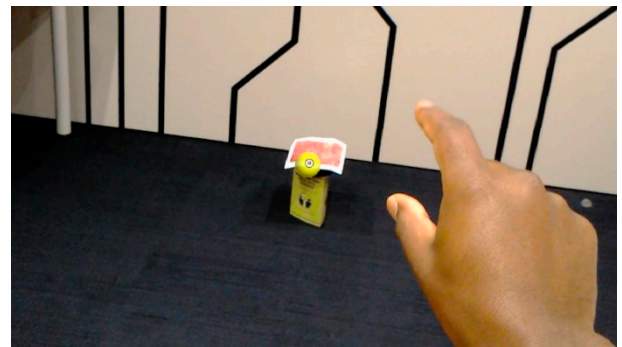
(**b**) User performs an air tap on the picture of Xantho and the application creates a Spatial Anchor object.
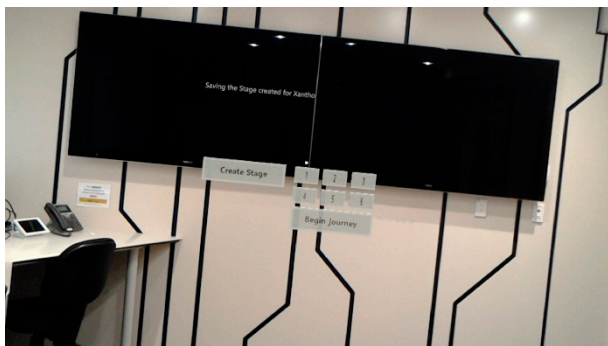
(**c**) The Spatial Anchor object created in the previous step is stored in the cloud and its identifier is displayed to the user. This identifier is then used as a Stage ID for the MR experience.
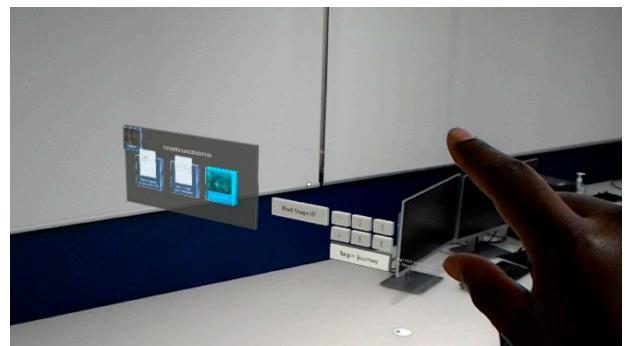
(**d**) The user shares the Stage ID with their collaborator and begins the MR experience, which will take place at the exact location where a Spatial Anchor object is created.

Figure 4. User interacting with the Xantho-Curator version of the clouds-based MR application. (**a**) user is asked to air tap on the picture of Xantho placed on the floor; (**b**) Spatial Anchor object is created and stored into the cloud Azure Spatial Anchor Service; (**c**) the application will display the identifier of the stored Spatial Anchor object, and; (**d**) user shares the identifier to their collaborator (user interacting with the Xantho-Visitor version) and begins the shared and collaborative MR experience together with their collaborator.

The clouds-based MR application is, therefore, expected to provide a shared real-virtual space that enhances cultural learning and enriches users' experience by (1) establishing a contextual relationship and engagement between users, virtual environments and cultural context in museums and heritage sites, and (2) by enabling collaboration between users.

5.4. Applicabilty of the Clouds-Based Mixed Realtiy—Museums and Heritage Sites

Conventional museums and heritage sites are known for preventing physical manipulation of artefacts. Visitors acquire knowledge about the artefacts from curators, guides, printed media and digital multimedia content available in museums and heritage sites. The clouds-based MR application, however, enables users to collaboratively manipulate and interact with the digital representations of the artefacts (3D models) via an interactive and immersive virtual environment, thereby resulting in a real virtual space where visiting experience can be augmented with a social and virtual presence. This enhances the learning experience since the cultural context presented to users is dynamic and interactive, rather than a linear information presentation format pre-determined by curators and cultural heritage professionals. The application can also be used to enable collaborative experience between curators located in museums and remote visitors/users wandering at heritage

sites or vice versa. Curators in the museum can communicate and collaborate with remote visitors to provide guidance.
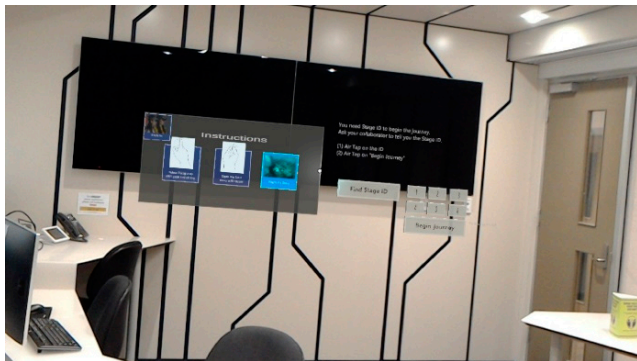
(**a**) Xantho-Visitor version asks the user to obtain a Stage ID from their collaborator who interacts with the Xantho-Curator version to set up the Stage ID.
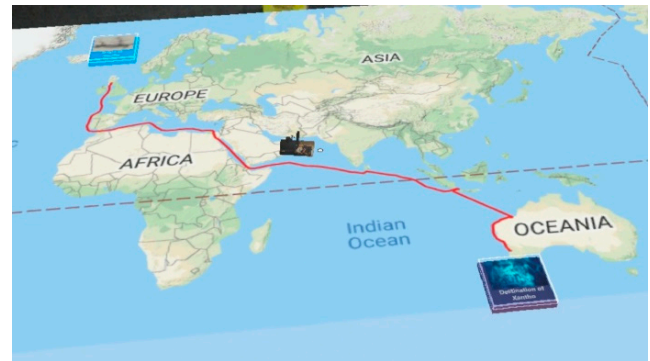

(**b**) Once the Spatial Anchor object is located for the Stage ID provided, the Xantho-Visitor loads the relevant scene. At this stage, both users of Xantho-Curator and Xantho-Visitor start interacting with the identical virtual environment.
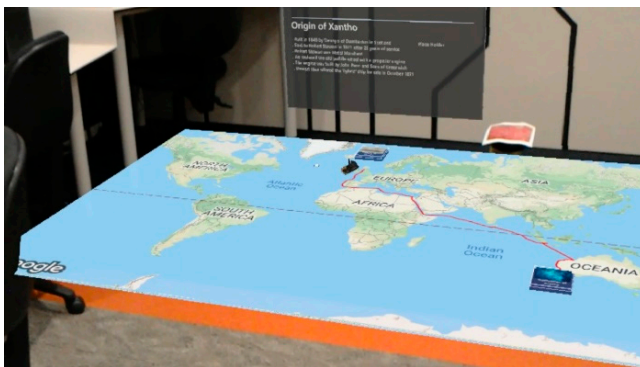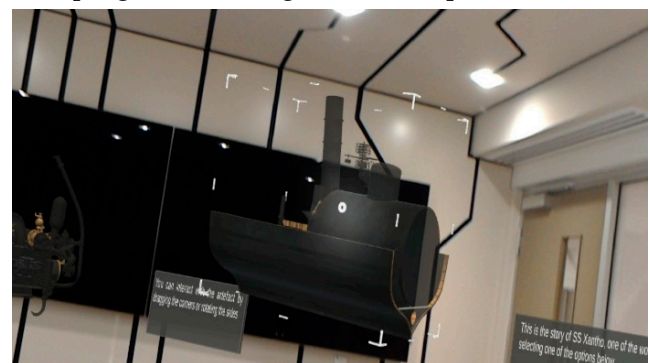

(**c**) User can interact with specific virtual objects and content.

(**d**) More content will be loaded to the scene as the user progresses through the MR experience
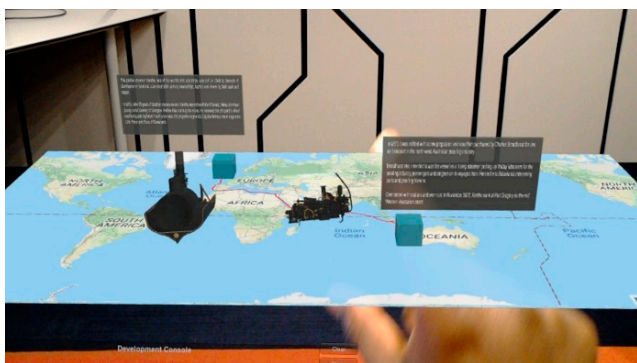
(**e**) User can interact with 3D models at different scales.

(**f**) User can manipulate and rotate 3D models.

Figure 5. User interacting with the Xantho-Visitor version of the clouds-based MR application. (**a**) user is asked to provide stage ID, a Spatial Anchor identifier they received from their collaborator; (**b–d**) the MR application will locate the shared Spatial Anchor object and loads the relevant scene at the exact location that was stored by the Xantho-Curator version, and; (**e**,**f**) both the Xantho-Curator and Xantho-Visitor version user of the application will interact with the MR environment displayed at a location identical to both users.

6. Conclusions

In this paper, we have presented a clouds-based collaborative and multi-modal MR application designed and built for VH. The primary objective of the application is enhancing cultural learning through engagement, a contextual relationship, and an interactive virtual environment. The implementation of the proposed application has utilised cloud services, such as Amazon Polly, Amazon S3, Azure Spatial Anchors, Azure Cosmos DB, and Azure App Service. In addition to cloud services, the implementation phase has exploited immersive reality technology (specifically, Microsoft HoloLens) and development tools and platforms, such as Unity, Mixed Reality Toolkit, Azure Spatial Anchors SDK, and Microsoft Visual Studio. SS Xantho, one of the world's first iron ships and Western Australia's first coastal steamer, was used as a case study for the prototype. Future works will attempt to evaluate the clouds-based MR application with curators, VH and cultural heritage professionals to validate whether the application enhances cultural learning in VH applications.

Supplementary Materials: Interested readers can visit the link provided to access video file that shows how the built prototype functions. https://cultural-dataset.s3.ap-southeast-2.amazonaws.com/xantho/video/XanthoMR.mp4.

Notes

1. https://azure.microsoft.com/en-us/topic/mixed-reality/ (accessed on 1 June 2021).
2. https://aws.amazon.com/polly/ (accessed on 1 June 2021).
3. https://aws.amazon.com/s3/?nc2=h_ql_prod_st_s3 (accessed on 1 June 2021).
4. https://azure.microsoft.com/en-au/services/spatial-anchors/ (accessed on 1 June 2021).
5. https://azure.microsoft.com/en-au/services/cosmos-db/ (accessed on 1 June 2021).
6. https://azure.microsoft.com/en-au/services/app-service/ (accessed on 1 June 2021).
7. https://docs.microsoft.com/en-us/dynamics365/mixed-reality/guides/authoring-gestures (accessed on 1 June 2021).
8. https://aws.amazon.com/dynamodb/ (accessed on 1 June 2021).
9. http://museum.wa.gov.au/explore/broadhurst/ss-xantho (accessed on 1 June 2021).

References

1. Tost, L.P.; Economou, M. Worth a thousand words? The usefulness of immersive virtual reality for learning in cultural heritage settings. *Int. J. Archit. Comput.* **2009**, *7*, 157–176. [CrossRef]
2. Champion, E.; Bishop, I.; Dave, B. The Palenque project: Evaluating interaction in an online virtual archaeology site. *Virtual Real.* **2012**, *16*, 121–139. [CrossRef]
3. Caputo, F.M.; Ciortan, I.M.; Corsi, D.; De Stefani, M.; Giachetti, A. *Gestural Interaction and Navigation Techniques for Virtual Museum Experiences*; Zenodo: Bari, Italy, 2016; pp. 32–35.
4. Economou, M.; Pujol, L. Evaluating the impact of new technologies on cultural heritage visitors. *Technol. Strategy Manag. Socio-Econ. Impact Herit. Manag. Ser.* **2007**, *2*, 109–121.
5. Maye, L.A.; Bouchard, D.; Avram, G.; Ciolfi, L. Supporting Cultural Heritage Professionals Adopting and Shaping Interactive Technologies in Museums. In Proceedings of the 2017 Conference on Designing Interactive Systems, Edinburgh, UK, 10–14 June 2017; pp. 221–232.
6. Bekele, M.K.; Pierdicca, R.; Frontoni, E.; Malinverni, E.S.; Gain, J. A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. *J. Comput. Cult. Herit. JOCCH* **2018**, *11*, 7. [CrossRef]
7. Vlahakis, V.; Karigiannis, J.; Tsotros, M.; Gounaris, M.; Almeida, L.; Stricker, D.; Gleue, T.; Christou, I.T.; Carlucci, R.; Ioannidis, N. Archeoguide: First results of an augmented reality, mobile computing system in cultural heritage sites. *Virtual Real. Archeol. Cult. Herit.* **2001**, *9*, 584993–585015.

8. Bekele, M.K.; Champion, E. A Comparison of Immersive Realities and Interaction Methods: Cultural Learning in Virtual Heritage. *Front. Robot. AI* **2019**, *6*, 91. [CrossRef] [PubMed]

9. Abdelrazeq, A.; Kohlschein, C.; Hees, F. A Cloud Based Augmented Reality Framework-Enabling User-Centered Interactive Systems Development. In Proceedings of the International Conference on Human Systems Engineering and Design: Future Trends and Applications, Dubrovnik, Croatia, 23–25 September 2021; pp. 417–422.

10. Kaufmann, H. *Collaborative Augmented Reality in Education*; Institute of Software Technology and Interactive Systems, Vienna University of Technology: Monaco Mediax, Monaco, 2003.

11. Thwaites, H. Digital Heritage: What Happens When We Digitize Everything? In *Visual Heritage in the Digital Age*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 327–348.

12. Bekele, M.K.; Champion, E. Redefining mixed reality: User-reality-virtuality and virtual heritage perspectives. In *Intelligent & Informed, Proceedings of the 24th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Wellington, New Zealand, 15–18 April 2019*; Victoria University of Wellington: Wellington, New Zealand, 2019.

13. Deliyiannis, I.; Papaioannou, G. Augmented reality for archaeological environments on mobile devices: A novel open framework. *Mediterr. Archaeol. Archaeom.* **2014**, *14*, 1–10.

14. Liarokapis, F.; Greatbatch, I.; Mountain, D.; Gunesh, A.; Brujic-Okretic, V.; Raper, J. Mobile augmented reality techniques for geovisualisation. In Proceedings of the Ninth International Conference on Information Visualisation, London, UK, 6–8 July 2005; pp. 745–751.

15. Marto, A.; Gonçalves, A. Mobile AR: User Evaluation in a Cultural Heritage Context. *Appl. Sci.* **2019**, *9*, 5454. [CrossRef]

16. You, S.; Thompson, C.K. Mobile collaborative mixed reality for supporting scientific inquiry and visualization of earth science data. In Proceedings of the 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 18–22 March 2017; pp. 241–242.

17. Zimmer, C.; Bertram, M.; Büntig, F.; Drochtert, D.; Geiger, C. Mobile Augmented Reality Illustrations that entertain and inform: Design and Implementation issues with the Hololens. In Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, Bangkok, Thailand, 27–30 November 2017; p. 23.

18. Toczé, K.; Lindqvist, J.; Nadjm-Tehrani, S. Performance study of mixed reality for edge computing. In Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing, Auckland, New Zealand, 2–5 December 2019; pp. 285–294.

19. Ali, G.; Le, H.-Q.; Kim, J.; Hwang, S.-W.; Hwang, J.-I. Design of seamless multi-modal interaction framework for intelligent virtual agents in wearable mixed reality environment. In Proceedings of the 32nd International Conference on Computer Animation and Social Agents, Paris, France, 1–3 July 2019; pp. 47–52.

20. Dermawan, K.F.; Yusuf, R. Moving Mixed Reality Game to the Cloud: A Survey on Feasibility. In Proceedings of the 2020 6th International Conference on Interactive Digital Media (ICIDM), Bandung, Indonesia, 14–15 December 2020; pp. 1–4.

21. Al-Barhamtoshy, H.M.; Al-Ghamdi, A. Toward cloud-based mixed reality e-learning system. In Proceedings of the 2017 International Conference on Informatics, Health & Technology (ICIHT), Riyadh, Saudi Arabia, 21–23 February 2017; pp. 1–6.

22. Fanini, B.; Pescarin, S.; Palombini, A. A cloud-based architecture for processing and dissemination of 3D landscapes online. *Digit. Appl. Archaeol. Cult. Herit.* **2019**, *14*, e00100. [CrossRef]

23. Malliri, A.; Siountri, K.; Skondras, E.; Vergados, D.; Anagnostopoulos, C.-N. The enhancement of underwater cultural heritage assets using augmented reality (AR). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**. [CrossRef]

24. Yang, S.; Hou, M.; Huo, P.; Li, A.; Jiang, L. A Cloud Computing Platform to Support Digital Heritage Application Using a Service-Oriented Approach. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 1497–1504. [CrossRef]

25. Šašinka, Č.; Stachoň, Z.; Sedlák, M.; Chmelík, J.; Herman, L.; Kubíček, P.; Šašinková, A.; Doležal, M.; Tejkl, H.; Urbánek, T. Collaborative Immersive Virtual Environments for Education in Geography. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 3. [CrossRef]

26. Bekele, M.K. Walkable Mixed Reality Map as interaction interface for Virtual Heritage. *Digit. Appl. Archaeol. Cult. Herit.* **2019**, *15*, e00127. [CrossRef]

27. Rahaman, H.; Champion, E.; Bekele, M. From photo to 3D to mixed reality: A complete workflow for cultural heritage visualisation and experience. *Digit. Appl. Archaeol. Cult. Herit.* **2019**, *13*, e00102. [CrossRef]

28. Museum, W.A. Broadhurst Familiy an Extraordinary Group. Available online: http://museum.wa.gov.au/explore/broadhurst/ss-xantho (accessed on 18 July 2021).

29. McCarthy, M. Broadhurst Familiy an Extraordinary Group. Available online: http://museum.wa.gov.au/explore/articles/story-ss-xantho-wa-s-first-coastal-steamer (accessed on 18 July 2021).

30. McCarthy, M. Charles Edward Broadhurst (1826–1905): A Remarkable Nineteenth Century Failure. Ph.D. Thesis, Murdoch University, Murdoch, Australia, 1989.