EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2021)
A. Chalmers and V. Hulusic (Editors)

Triggering the Past: Cultural Heritage Interpretation Using Augmented and Virtual Reality at a Living History Museum

K. Shitut, [1] J. Geigel,[1] J. Decker,[2] G. Jacobs,[3] and A. Doherty,[3]

[1]Rochester Institute of Technology, Dept. of Computer Science, USA
[2]Rochester Institute of Technology, Dept. of History / Museum Studies, USA
[3] Rochester Institute of Technology, School of Design, USA

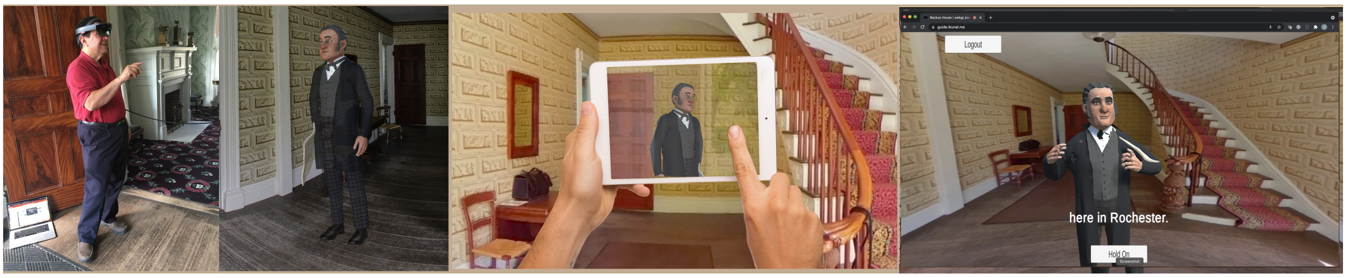

Figure 1: *Presentation of a digital character using three classes of mixed reality devices: Microsoft Hololens (left), Mobile device (center), WebXR enabled browser (right)*

Abstract

In this paper, we present a use case for the introduction of historical digital characters in the context of a living history museum. We describe a prototype system and framework that enables the use of augmented and virtual reality for placing a these characters in the museum space. The system uses a conversational interface for natural interaction, supports the scanning of objects in the museum space for guiding the conversation, and provides a common user experience on a variety of mixed reality devices including the Microsoft Hololens, mobile devices, and WebXR enabled Web browsers. We describe our character creation workflow, provide technical details on the implementation and discuss the user testing of the system. Findings from our testing suggest, that despite the analog, hands-on tradition of living history museums, the use of immersive technologies has the potential to greatly enhance the visitor experience while engaging users within the physical space of the museum.

CCS Concepts

• ***Computing methodologies*** → ***Mixed / augmented reality; Virtual reality;*** • ***Applied computing*** → *Computers in other domains;*

1. Introduction

The use of modern digital storytelling techniques can be a valuable tool to enhance the experience at cultural institutions. At the same time, this tool must be used carefully as not to disrupt the balance between the virtual assets and the physical spaces that they are designed to enhance [MPHH20].

In this work, we explore the use of locative digital storytelling to enrich the experience at a living history museum. Living history (or "open air") museums are venues that attempt to simulate life in another time, to present life as it once was [And82]. The museum experience is enhanced through costumed interpreters who may take on a particular role, with the premise that they are conveying what it was like to live in the past with the modern visitor encountering them in their daily life as they speak informatively, usually while demonstrating a skill or trade [Thi10, Rot98].

Our project extends this storytelling mechanisms by using augmented and virtual reality (AR/VR) to present digital characters, gleaned from history, who who engage the visitor through voice recognition, and a Q&A exchange. With personal narratives already being an essential part of the living history museum experience, we find it to be a perfect venue for introducing such digital characters.

At the same time, maintaining the balance between virtual and

DOI: 10.2312/gch.20211406

real presents a particular challenge as our project ultimately contradicts assumptions about the ways in which living history museums, which, by their very nature, are associated with historical tradition, hands-on,and analog experiences. We address this challenge by closely integrating the physical space as an integral part of the storytelling process. Not only is the character placed in a historically significant space with which they may have had a personal connection, but objects within a space are used as props to further progress the story.

In this paper, we describe a use case for and a prototype application that enables this kind of interaction. Major features of the prototype include:

1. The use of contextualized narratives, delivered digitally by a virtual character, to engage visitors in gaining historical knowledge about the past.
2. Integration of the use of trigger objects, comprising cultural heritage objects from the museum's collection, distributed around a physical space that can be scanned and used to unlock new conversation options for the character.
3. A common user experience on a variety of technical platforms (including the Hololens, mobile devices, and WebXR compatible browsers), for on-site and off-site interaction.

Working closely with Genesee Country Village and Museum (GCV&M) [Gen21], the third largest living history museum in the United States and the largest in New York state, our goal is to create a user experience that not only informs visitors and conveys information, but engages, affording visitors the opportunity to make personal connections with cultural heritage.

2. Background

The application of virtual and augmented reality technologies by cultural heritage institutions to attract and engage visitors has been a robust field of inquiry. A recent survey is presented in [BPF*18] providing an overview of the various technologies and applications that have contributed significantly in this domain.

A recent trend is the use of these immersive technologies in promoting storytelling to enhance cultural exhibits. This is particularly the case for augmented reality where digital content can be placed in a physical space and superimposed over physical artifacts.

Some exhibits focus on a single event, such as [Lie18], which describes a mobile application that allows users to interact with a digital recreation of the landings at Omaha beach, whereas others, such as the CHESS Project [KKK*14], enables the creation of AR enhanced stories designed to engage visitors as they are guided through a museum space.

Projects such as England Historic Cities [Eng20] and STORYTech 4 EVER [MSM*15] allow for storytelling amongst multiple cultural sites creating a common storytelling space that will enable cross institutional stories exploring common themes at related historical sites.

The use of SLAM (Simultaneous Localisation and Mapping) enabled mobile AR [DWB06] has been successfully demonstrated to develop a situated short story [KCE19]. In the spirit of Pokemon Go, objects in the environment are scanned, identified, and utilized to enhance the story narrative.

Research on the creation of virtual humans has advanced to the point where it is possible to create convincing avatars that combine realistic animation, facial expression generation, lip sync, natural language processing, speech synthesis, and non verbal behavior and that can naturally interact with human visitors [HTM*13,KHG*07, SGHJ*06].

These technologies have been employed in the creation of virtual tour guides for cultural heritage sites [MDC18]. The benefits of a virtual tour guide over the use of other media technologies such as video, audio or animations etc. is the ability to tailor the experience to the user's personal preferences. Use of digital docents has shown great promise in enhancing the museum experience; garnishing greater attention from museum goers and provide better content delivery leading to better learning in a cultural heritage setting [CCT*18]. For example, The Boston Museum of Science has made use of Embodied Conversational Agent as digital docents (Tinker [BVS13], Ada and Grace [TAA*12]), which are projected as robots on a large screens. The Speaking Celt [BS16], uses mobile AR to guide users throughout the Museum of Celtic Heritage in Hallein, Austria.

The goal and focus of our work is not to enhance any one of these individual technologies, but instead, to use current methods to explore the use of digital docents within the context of a living history museum using historical characters as the virtual guides. With such a strong focus on personal storytelling, the use of this kind of digital enhancement can be particularly advantageous in this context [CM12]. These museums provide a unique storytelling opportunity that melds the physical with the virtual that forges a more personal connection between museum artifacts, historical spaces and visitors.

3. Character Creation Workflow

As part of the development process for the application, we defined a protocol for creating digital characters who can present their story in first-person. Implementing this protocol for a given character is a truly interdisciplinary effort that involves collaboration between students and faculty from Museum Studies, 3D Digital Design, Computer Science and Performing Arts and has been integrated into the curricula in these disciplines. Our project design has been informed by pedagogy, as this project was conceived from the outset as a collaboration among faculty and student researchers across these disciplines. The pedagogical advantages afforded by the project are discussed in [DDGJ20]

In consultation with museum staff, Dr. Frederick F. Backus (1794-1858) was chosen as the inspiration for our first fully-developed character. Backus was an important and influential figure in Rochester history. In addition, the house in which Backus originally lived, the Livingston-Backus House, is resident in the museum village, making it possible to place the digital version of Dr. Backus in the actual space he once inhabited.

The workflow is illustrated in Figure 2. We use the creation of the Dr.Backus character as an example to explain the components of this workflow.

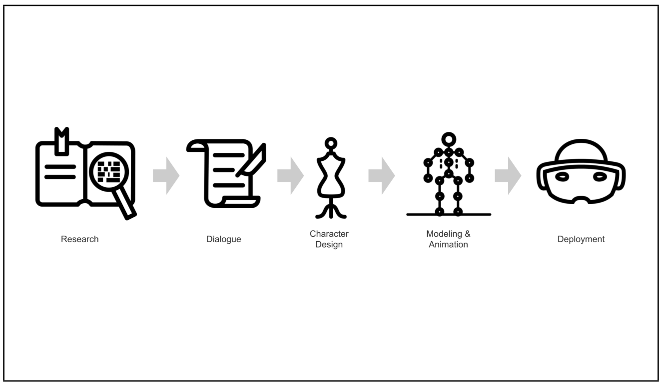

Figure 2: *Character Creation Workflow*

3.1. Research

The first step in the character creation process involves extensive research into the life of the character; basic detective work to determine *"Who exactly was Frederick Backus?"*. This exploration was achieved through examination of personal letters, journals, and personal narratives held in several archival collections, in addition to broader contextual documents including newspapers widely available now through digital repositories.

In addition to developing a personal profile of the character, this step results in a list of broad conversational topics that can later be discussed by the character.

For Dr. Backus, this portfolio of research focused on the Erie Canal, the cholera epidemic of 1832 in Rochester, local politics, philanthropic and benevolence projects, and the state of medicine and homeopathy and medicine at the time.

3.2. Dialog

Next, branching narratives for presentation of the content and context identified in the research step are defined.

For our framework, we employ a question-answering approach [LT11] to define dialog options and interaction with the digital docent. Major content areas are broken down into subtopics that allow for pauses and verbal prompts from the user enabling visitors to be able to direct the conversation with the character through their choices within the applications. Some of these topics will later be connected with physical artifacts in the museum space (as described in Section 6.2).

We engaged the concept of "thinking dispositions" as a mode of inquiry and visitor engagement. Developed by Harvard Graduate School of Education, "Thinking Dispositions" can be employed in museums through the use of storytelling, in order to present the content conversationally, in a first-person interactive narrative [Har16]. The "dispositions" encourage learners to ask questions of themselves and their world by tiering their understandings by asking learners to ponder the following questions:

1. Why might this [topic, question] matter to me?
2. Why might it matter to people around me [family, friends, city, nation]?
3. Why might it matter to the world?

In our application, through the dialogues we create for our characters, we engage visitors to motivate reflection upon the stories of our characters on a personal, local, and global level. For Backus, we structured the conversation into the branching narratives devised into topics, as illustrated in Figure 3.

3.3. Character Design

During the design phase, a 3D model of the character is created and rigged. Historical research and documents are used to guide the appearance of the character, both regarding physical and facial features as well as proper costuming that reflects the time period in which the character is presented (Figure 4). Interesting enough, there are no known photographs of Dr. Backus so his likeless had to be approximated from photos of his father, Azel Backus, from archives of Hamilton College at which he served as president. Other historical archives were used to guide the design in terms of costume and style.

We chose to build a stylized avatar, rather than a photorealistic 3D model, so as to avoid the "uncanny valley" [MMK12]; a feeling of unease and disconnect experienced when humans encounter robotic or audio/visual simulations that are too realistic.

Industry standard tools (e.g. Character Creator by Reallusion, for model creation and rigging, and Marvelous Designer 8, for cloth and clothing design) are used in the generation of the 3D model. In addition to modeling and rigging of the characters's body structure, facial blendshapes (Figure 5) are defined to assure proper lip-sync and facial expressiveness of the model when conveying his story.

3.4. Modeling and Animation

We use motion capture (for both full body and facial motion) to create the animations for the digital character. Trained actors act out the monologues whose gestures and facial movements are captured and applied on the rigged 3D model. Actors were chosen based on their vocal style to match the historic character. In order to preserve the legibility of the narrative in performance, the actor's voice-over track was recorded in advance allowing the performers to play with the inflection and vocal emphasis of segments of the script in a sound-isolated recording booth. The performers then recreated the character's movements in front of a motion capture system, using the audio playback as reference. These motions were then transferred to the avatar model and any jitter was removed.

3.5. Deployment

Once created, an application for presentation of, and interaction with, the character was created. One of the design goals for the application was to provide users the flexibility of viewing the character in a variety of ways, depending on the user's location and access to technology as illustrated in Figure 6. Technical details of this application are discussed in Section 5.

Visitors at the museum ("On Site"), will interact with the digital character using Augmented Reality (AR), meaning the character

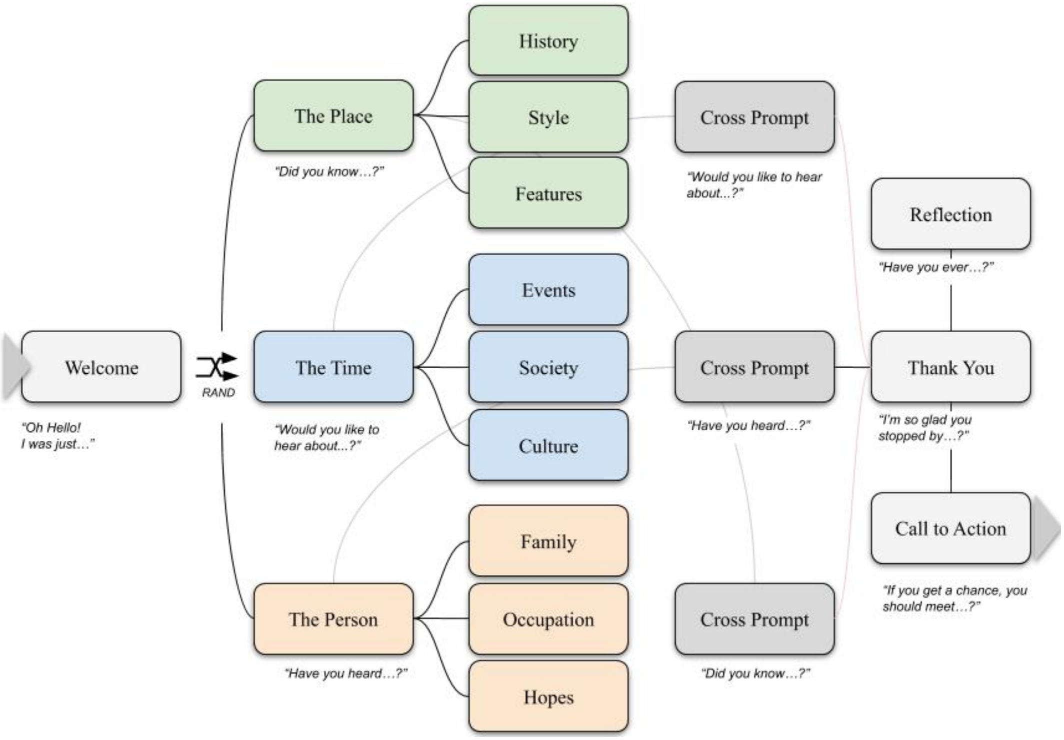

Figure 3: *Branched narratives created using "thinking dispositions".*

Figure 4: *Development of the design and costumes for the Backus character.*

Figure 5: *Facial blendshapes use for lip sync of spoken dialogues.*

would be superimposed on the real world utilizing a wireless device. The AR devices would either be provided by the museum, as is the case with the HoloLens, or a guest's personal device like a smartphone or tablet with a camera that would place the avatar in the surrounding environment.

Users outside of the museum ("Off Site"), could experience the character in a Virtual Reality (VR) context, where a VR device creates the environment background as well as the character and delivered via cross platform web-browsers that support WebXR [W3C20]. This will allow the museum to extend its interactive reach to homes and classrooms. For both deployment situations the same character content, model and dialog are presented as discussed in Section 4 below.

4. User Experience

In keeping with the spirit of interpretation and informal storytelling of a living history museum, when designing the user experience, our goal was to mimic a casual encounter with the digital character in his familiar space. It is hoped that this encounter will lead to an informal dialogue where the character will extend an invitation to explore his dwelling and take note of objects in the house. These "trigger objects" will then be used to prompt different threads of conversation.

Thus, the application, in addition to presenting the 3D character model, would need to support two main user activities: the initiation of conversation and the scanning of trigger objects to guide the

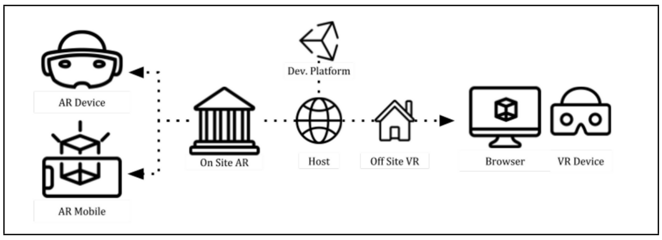

Figure 6: *Deployment strategy using augmented reality (AR) for onsite interaction and virtual reality (VR) for offsite interaction.*

conversation. Two modes for conversational interaction were defined: a menu based interface where GUI elements are used to pose questions to the character and a speech interface that will allow users to pose questions directly using voice input.

A common interface for deployment was designed for all three of the target device platforms with a few minor differences based on the capabilities of the device. Text captions are available for presentation using all three deployment platforms.

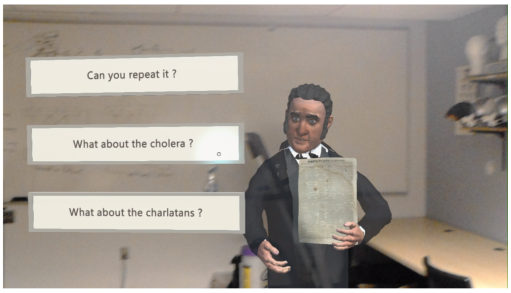

Figure 7: *The Hololens interface of the application. The character is placed in the 3D space where the application is initiated. (In this figure, the research lab where development of the app took place.)*

Figure 7 illustrates the Hololens interface. Initial model placement is performed based on the spatial scan of the environment performed by the Hololens device. The Hololens device is set up to always be listening for verbal input from the user, making conversational interaction intuitive and natural. For interaction with the GUI buttons, and for scanning trigger objects, the standard Hololens *gaze and commit* [Mic19] method, where users make an "air tap" gesture while gazing at an object, is employed.

For the mobile version, an access/user code feature is introduced to keep track of objects scanned by an individual user and to provide continuity between onsite and offsite interaction with the character. For initial placement of the character, the mobile version requires that users scan the surrounding environment upon startup to map out all the horizontal surfaces creating a mesh where the character can be placed (Figure 8). Users can tap on any visible part of the mesh and the digital model is then placed a that point.

Unlike the Hololens version, for the mobile version, users must explicitly indicate their desire to speak to the character and to scan objects using GUI buttons on the app. At any point users can press a "Press To Speak" button to verbally ask a question. Similarly, at

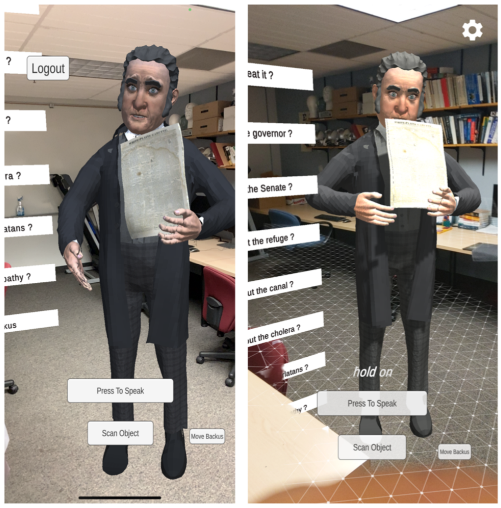

Figure 8: *Mobile version of the application implemented for both iOS and Android. The grid indicates a horizontal surface on which the character can be placed. Buttons are used to initiate conversation and to scan objects*

any point users can use the "Scan Object" button to scan an object and capture a photo of the object on which they wish to focus.

Figure 9: *Web version of the application viewed in a WebXR enabled browser.*

The Web version allows for off-site interaction with the character using a WebXR compatible browser. As the Web version is a VR as opposed to an AR experience, the character is placed in a virtual museum space using a 360 degree image of the interior of the museum location where the character would be placed physically.

Users start the experience by entering a url for the application in a compatible browser. Once opened, users are first asked to login to the experience using the same code that was used in the AR experience, thus creating continuity between the on-site and off-site interactions with the character. As the web version does not currently support navigation through the museum space, the scanning feature is also disabled, allowing users only to be able to use the dialog options that they unlocked in the AR version while onsite. One advantage of the web version is the seamless integration of transcription as users can, like with the Hololens version, ask questions without the need of a button press.

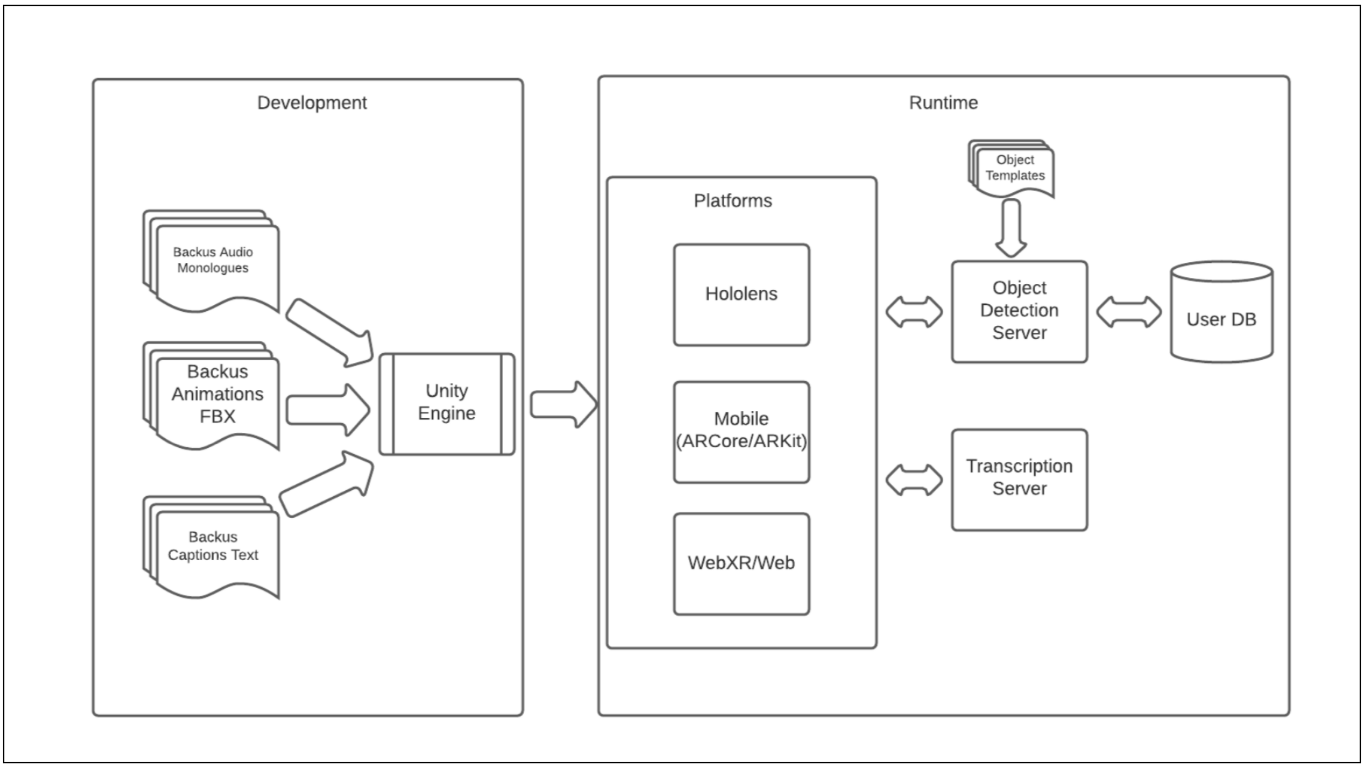

Figure 10: *Application system architecture.*

5. Technical Infrastructure

In designing the technical infrastructure for application development, one major goal was to create a common development pipeline for ease of deployment amongst all three target platforms, keeping in mind the disparity of processing capabilities amongst each of the possible target devices. We thus chose a client-server architecture where the client is responsible for providing immediate interaction with the user whereas major computational tasks (speech processing, target image identification) are performed by a more powerful server.

The overall architecture of the system is illustrated in Figure 10. Complete details of the implementation can be found in [Shi21].

5.1. Client Application

In developing the client, the content created for the character (as described in Section 3) are combined into a single application. Specifically the following sets of assets are integrated:

1. Character model (with animations included) – A 3D model of the digital docent with animations. The model and the captured (and cleaned) animations are stored in FBX files, an industry standard 3D asset exchange format that facilitates higher-fidelity data exchange between popular 3D applications.
2. Audio Recordings — pre-recorded audio of the docent's monologue, stored in MP3 format
3. Script – a textual transcript of the audio recording stored in a plain text file. These transcripts are used to generate captions displayed by the application.

The application is built using Unity, an industry standard 3D application development platform. Although it is not open source, it is freely available and used extensively for building 3D applications involving VR and AR. The choice of Unity was partially motivated by its "develop once, deploy many" philosophy, whereby a single spatial 3D application could be deployed on a variety of different devices and platforms, eliminating the need for individual code bases for each desired platform.

Though a few platform specific libraries specific are required, the bulk of the code in presenting the character can be shared between the applications for all output targets. Unity currently supports deployment on most popular output devices including the HoloLens, mobile device apps (both iOS and Android) and WebXR.

5.2. Back end server processing

Heavier computational tasks such as speech to text transcription and image matching of scanned target objects are performed on a Google Cloud server.

The transcription server uses Google's Cloud speech-to-text service [Goo21] to process audio clips into its text transcription. The audio data is uploaded to the server through a WebSocket connection [MDN21] which allows for two-way interactive communication session between the client and the server. The primary reason for using a server for transcription in contrast to integrating transcription into the on board application was to have a common transcription system, avoid security issues due to exposure of transcription API keys, and enable seamless switching of transcription libraries in the future if required.

With the object scanning component of our system being core to the user engagement, both accuracy and speed of the detection are essential for optimal user experience. For this, we have used SURF (Speeded Up Robust Features) [BETVG08] as the feature detector and descriptor which is part of OpenCV Computer Vision libraries [Ope20]. The server performs two inter-dependent tasks: First, It processes the images taken by a user and detects the relevant object if present in the scene; Secondly, it maintains a relevant database of all the objects processed by the user so that the unlocked options can be tracked across user sessions during onsite and offsite interaction.

6. User Testing

6.1. Initial Hololens Testing

Our first testing took place in April 2019 within the context of a museum studies course where students were familiar with the project as each student had contributed to it by writing monologues for characters. The purpose of this early study was to test the viability of the concept and suitability of the use of the Hololens in a museum context. 19 students took part in the testing with only 31.6% of the users indicating that the experience fulfilled their desire of a museum experience. This initial testing informed a number of interface and interaction modifications to the application.

Figure 11: *Testing of the Hololens version of the app in November 2019*

A second test of the updated Hololens app, complete with included content for the Dr. Backus character, was held in November 2019 as part of a demonstration at our university's annual AR/VR/XR symposium (Figure 11). Approximately 50 individuals took part in this informal test. In contrast to the first test session, users in November 2019 showed a greater sense of fulfilment from this experience.

This feedback suggested that users were indeed interested in the concept as part of the museum experience, prompting further development of the application for an onsite interaction.

6.2. On-site User Study

A more formal, two phase user study was conducted at the Livingston-Backus house in GCV&M, the original home of Dr. Frederick Backus. Volunteers were recruited with individuals comprising a diverse group ranging in factors such as age, familiarity with immersive technologies (especially AR, VR & MR), comfort with technology in daily usage, and familiarity with the museum itself.

Due to COVID-19 protocol restrictions, the Hololens version was not tested during this study, relying instead on the Mobile version of the client which was downloaded and installed on participants' mobile devices.

The user's interaction in the Livingston-Backus house begins at the main entrance of the house. As a visitor enters the main doorway, they find themselves in a large hallway that leads to the back of the house with a staircase on right that leads to the upper floors, and rooms on their left and right. This is where, after starting the application, Backus will describe the architecture of the house, referencing it's influence from the Greek revival style prominent in New York at the time. In this area of the house, volunteers can encounter the first set of objects in the hallway and the room on the left. These are a newspaper, medical bag, the inkwell and the family Bible (All the trigger objects are shown in Figure 12). As they move towards the kitchen, they will also find a pitcher in the dining room. The experience continues towards the kitchen of the house in the back, where a medicine bottle, herb and wall clock serve as triggers. The final trigger is the exit sign as volunteers exit the house through the back door into the garden. While moving through the house, Backus can be moved and placed into the desired position

By selecting trigger objects that are relevant to the history of the house and Dr. Backus, the layout and look of the house is preserved to be accurate to the conveyed time period. Thus, through optimal distribution of triggers objects within the space, exploration of the house is incentivized and an engaging user experience is created.

Figure 12: *Trigger objects used during onsite testing to unlock themes for conversation with Dr. Backus.*

6.2.1. Study Protocols

Volunteers were asked to fill a pre-visit survey and consent form to provide some idea of their familiarity with the museum, and consent to use their anonymous interaction data. Volunteers were suggested to complete these forms and install the apps to their devices before arriving at the museum. Upon arrival, all users were given an introductory overview of the system. They were shown a demonstration of the interactions with Backus and performing object scanning. After the introduction, they were allowed to look around the house, scan objects and unlock the desired options, and interact with Backus. While we were present onsite to help with any issues or questions that might arise regarding the system, face-to-face interactions were minimized to reduce the risk of COVID-19 transmission and also to enable the experience to happen as intended (Figure 13).

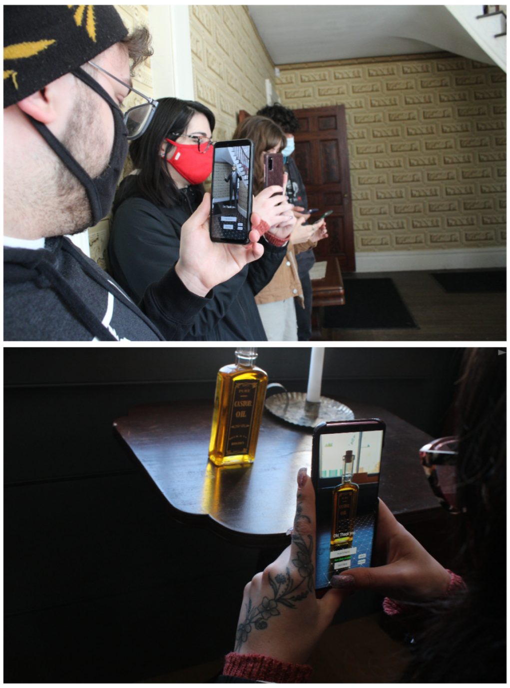

Figure 13: *Volunteers testing the mobile application on-site at the Livingston-Backus House on the museum grounds.*

After completing the experience, users were asked to complete a post-survey for the onsite component.

Users were also given a URL for the WebXR version and asked to interact with Backus using the same code as the one they had created for the onsite experience. They were also encouraged to share the experience with others who were not part of the onsite visit. Once users has completed this offsite component they were asked to fill out an offsite survey form. A window of 48 hours was given for completion of the offsite component from the completion of the onsite component. Once both the components of the study were completed the data from the study was collected and analyzed to determine the feedback from the system.

6.2.2. Findings

For the study, we tested the system with 9 GCV&M staff members and 17 volunteers who recruited via email and social media. The choice of including museum staff in the study was deliberate as we were interested in feedback from those very familiar with the living history museum experience to see how the introduction of digital characters affected that experience.

Out of the total 17 external volunteers, 15 completed the pre-visit survey. Out of the total 26 participants in the onsite and offsite study, 17 completed the onsite survey while 12 completed the off-site survey.

Based on the data gathered in the onsite version, a majority of respondents (88%) indicate that the experience did enhance their interest in the history of the museum. Volunteers were enthusiastic in making suggestions for integrating the system into other buildings or expanding with new features. Some specific written feedback is shared below:

- *"I see this as being more for younger generations (mainly kids or families). You could make a game/scavanger hunt/quiz for prizes. You could make interactive demos for (example) the tin smith"*

- *"I love the idea that visitors could meet Backus and have a better time learning more information when they step into the house. I would absolutely love to see additional characters developed, as well as additional houses represented"*

Regarding usability and intuitive interaction, the app met with a mixed response (only 46% found the app easy to use) with two avenues of improvement evident to us from onsite observations: better network speeds to speed up the transcription and object detection, and clear indicators for trigger objects; These issues are further discussed in Section 6.2.3. Despite this, most found the app to enhance the museum experience.

For the offsite version, while the amount data gathered is insufficient to draw any definite conclusions, responses seem to indicate a positive experience with 6 out of 10 users expressing preference for the WebXR version over mobile for offsite use. From this we see some potential of the application to enhance the museum's virtual outreach.

6.2.3. Technical Issues Encountered

A number of technical issues were identified from verbal feedback from users and the user surveys.

Network Connectivity – Given the design of our system, it became evident that access to a reliable Internet connection to be essential when running the application. While the primary solution to this problem would be to improve the network connectivity in the museum space, another would be to perform both object recognition and speech transcription on the device. Libraries such as Vosk [Alp21], an open source speech recognition library for on device and server side speech recognition, and OpenCV4Android SDK could be used to substitute or complement the server side APIs.

Object Detection – While the SURF-based object detector was successful with many 2D surfaces, matching 3D objects that did not have a high number of good points for the descriptor and feature detector and were susceptible to lighting changes in the environment, prevented some users from unlocking all the options. For future iterations, list of trigger objects should be more carefully chosen that perform better with the implemented detection algorithm. The list of images on which the detector is trained could also be expanded to include trigger object images taken in a variety of lighting conditions. Objects that have unique 2D faces, as opposed to 3D shapes, should be preferred.

Locating trigger objects – Another issue that was mentioned by users is lack of hints to signify which objects in the space were triggers. This can be easily addressed by providing photos of objects to find, indicate location of nearest objects within a range through GPS etc. or even employing a serious game approach (e.g. [MCB*14]) incorporating the the search for trigger objects as a sort of scavenger hunt. It is also necessary to ensure that indicators are either deployed through the app, or deployed in the real world in a manner that they blend effectively with the museum space.

7. Conclusions and Future Work

In this work, we explore the use of augmented and virtual reality to allow the placement of and interaction with a historical character to enhance the visitor experience at a living history museum. We described a prototype application that includes a conversation interface allowing for an intuitive means for interacting with the character, the ability to focus on objects with the space to guide the conversation, and that provides a common user experience amongst a variety of mixed reality devices and platforms.

Our onsite and offsite testing of the prototype suggests that the system has great potential in successfully increasing interest about the site's history even despite some of the technical flaws identified in Section 6.2.3. Incorporation of real world objects with the virtual tour guide seemed to further enhance interest in the site's history, creating an integrated story space where items from a cultural heritage collection define a rich set of conversational triggers. The system rewards exploration and creates an integrated visitor experience that complements, as opposed to eclipsing the physical cultural space. Link between the onsite and offsite experiences extend an onsite visit beyond the time spent in the museum.

Though we did not get the opportunity to test the Hololens onsite at the museum, informal offsite tests suggest that it produces a very natural user experience. That said, the device itself is somewhat bulky and unnatural to wear. Mobile devices, on the other hand, afford a much more lightweight experience at the cost of less natural interaction involving the use of onscreen GUI elements to initiate conversation and to scan objects. As development of AR devices continue to advance, we look forward to devices that combine the naturalness of the Hololens interface with the footprint of a mobile device as we feel these class of devices to be perfect for future explorations in this space.

Though effective, the question/answer paradigm utilized for conversing with our character may be overly simplistic to model actual conversation. The use of more comprehensive conversational models that mimics more natural human interaction (e.g. [RY10, TAA*12, BVS13]) along with tools for creating generative conversations (e.g. [SMG09]) could greatly improve the naturalness of the interface with our digital character. However, using these models may have space implications on the size of the resultant application making this approach unrealistic for mobile devices. One avenue of further exploration would be to apply existing compression or streaming algorithms to provide the natural conversation experience while constraining the application size.

Based on the success of Dr. Backus, we are currently in the process of developing four more characters which will be placed in other buildings in the museum. We plan on addressing some of the technical issues mentioned in Section 6.2.3 and exploring more sophisticated conversational models that could provide interconnected dialogues between characters resulting in integrated conversations as visitors traverse the buildings of the museum. We are also planning on further user testing, as the work could benefit from more extensive usability studies.

Acknowledgements

This system was developed over several years in collaboration with student contributors from several disciplines at RIT. We'd like thank all of the student who participated in the project throughout its development. In particular, special thanks to Brienna Johnson-Morris and Lizzy Carr who played a crucial role in organizing the test for Hololens version of the system and Koda Drake, for capturing and providing photographs of the onsite testing.

We would like to thank the team at Genesee Country Village & Museum, especially Becky Wehle, the President and CEO, and Peter A. Wisbey, the Curator of Collections, for their continuous and enthusiastic support for the project.

Finally, we would like to thank anonymous reviewers for their valuable feedback on the manuscript.

References

[Alp21] ALPHA CEPHEI: Vosk offline speech recognition, Jul 2021. URL: https://alphacephei.com/vosk/. 8

[And82] ANDERSON J.: Living history: Simulating everyday life in living museums. *American Quarterly 34*, 3 (1982), 290–306. 1

[BETVG08] BAY H., ESS A., TUYTELAARS T., VAN GOOL L.: Speeded-up robust features (surf). *Computer vision and image understanding 110*, 3 (2008), 346–359. 7

[BPF*18] BEKELE M. K., PIERDICCA R., FRONTONI E., MALINVERNI E. S., GAIN J.: A survey of augmented, virtual, and mixed reality for cultural heritage. *Journal on Computing and Cultural Heritage (JOCCH) 11*, 2 (2018), 1–36. 2

[BS16] BREUSS-SCHNEEWEIS P.: " the speaking celt" augmented reality avatars guide through a museum–case study. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (2016), pp. 1484–1491. 2

[BVS13] BICKMORE T. W., VARDOULAKIS L. M. P., SCHULMAN D.: Tinker: a relational agent museum guide. *Autonomous agents and multi-agent systems 27*, 2 (2013), 254–276. 2, 9

[CCT*18] CARROZZINO M., COLOMBO M., TECCHIA F., EVANGELISTA C., BERGAMASCO M.: Comparing different storytelling approaches for virtual guides in digital immersive museums. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (2018), Springer, pp. 292–302. 2

[CM12] CIOLFI L., MCLOUGHLIN M.: Designing for meaningful visitor engagement at a living history museum. In *Proceedings of the 7th nordic conference on human-computer interaction: Making sense through design* (2012), pp. 69–78. 2

[DDGJ20] DECKER J., DOHERTY A., GEIGEL J., JACOBS G. D.: Blending disciplines for a blended reality: Virtual guides for a living history museum. *Journal of Interactive Technology & Pedagogy 1* (2020), 17. 2

[DWB06] DURRANT-WHYTE H., BAILEY T.: Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine 13*, 2 (2006), 99–110. 2

[Eng20] ENGLAND ORIGINALS: England's historic cities, Dec 2020. URL: https://englandoriginals.com/ehc-app/. 2

[Gen21] GENESEE COUNTRY VILLAGE & MUSEUM: Gcv&m, Aug 2021. URL: https://www.gcv.org/. 2

[Goo21] GOOGLE: Speech-to-text, Jul 2021. URL: https://cloud.google.com/speech-to-text. 6

[Har16] HARVARD GRADUATE SCHOOL OF EDUCATION: Project zero: The 3 whys, 2016. URL: https://pz.harvard.edu/resources/the-3-whys. 3

[HTM*13] HARTHOLT A., TRAUM D., MARSELLA S. C., SHAPIRO A., STRATOU G., LEUSKI A., MORENCY L.-P., GRATCH J.: All together now. In *International Workshop on Intelligent Virtual Agents* (2013), Springer, pp. 368–381. 2

[KCE19] KETCHELL S., CHINTHAMMIT W., ENGELKE U.: Situated storytelling with slam enabled augmented reality. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry* (2019), pp. 1–9. 2

[KHG*07] KENNY P., HARTHOLT A., GRATCH J., SWARTOUT W., TRAUM D., MARSELLA S., PIEPOL D.: Building interactive virtual humans for training environments. In *Proceedings of i/itsec* (2007), vol. 174, pp. 911–916. 2

[KKK*14] KATIFORI A., KARVOUNIS M., KOURTIS V., KYRIAKIDI M., ROUSSOU M., TSANGARIS M., VAYANOU M., IOANNIDIS Y., BALET O., PRADOS T., ET AL.: Chess: personalized storytelling experiences in museums. In *International Conference on Interactive Digital Storytelling* (2014), Springer, pp. 232–235. 2

[Lie18] LIESTØL G.: Storytelling with mobile augmented reality on omaha beach: Design considerations when reconstructing an historical event in situ. *MW18: MW* (2018). 2

[LT11] LEUSKI A., TRAUM D.: Npceditor: Creating virtual human dialogue using information retrieval techniques. *Ai Magazine 32*, 2 (2011), 42–56. 3

[MCB*14] MORTARA M., CATALANO C. E., BELLOTTI F., FIUCCI G., HOURY-PANCHETTI M., PETRIDIS P.: Learning cultural heritage by serious games. *Journal of Cultural Heritage 15*, 3 (2014), 318–325. 9

[MDC18] MACHIDON O. M., DUGULEANA M., CARROZZINO M.: Virtual humans in cultural heritage ict applications: A review. *Journal of Cultural Heritage 33* (2018), 249–260. 2

[MDN21] MDN WEB DOCS: The websocket api (websockets), May 2021. URL: https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API. 6

[Mic19] MICROSOFT: Gaze and commit, 2019. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-commit. 5

[MMK12] MORI M., MACDORMAN K. F., KAGEKI N.: The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine 19*, 2 (2012), 98–100. 3

[MPHH20] MILLARD D. E., PACKER H., HOWARD Y., HARGOOD C.: The balance of attention: The challenges of creating locative cultural storytelling experiences. *Journal on Computing and Cultural Heritage (JOCCH) 13*, 4 (2020), 1–24. 1

[MSM*15] MONACI S., SANNA A., MORREALE D., CUNIBERTI G., EQUIZZI M.: Storytech 4 ever storytelling technologies for european values and heritage. In *2015 Digital Heritage* (2015), vol. 1, IEEE, pp. 415–416. 2

[Ope20] OPENCV - OPEN SOURCE COMPUTER VISION: Opencv modules, Dec 2020. URL: https://docs.opencv.org/4.5.1/. 7

[Rot98] ROTH S. F.: *Past into present: Effective techniques for first-person historical interpretation*. Univ of North Carolina Press, 1998. 1

[RY10] RIEDL M. O., YOUNG R. M.: Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research 39* (2010), 217–268. 9

[SGHJ*06] SWARTOUT W. R., GRATCH J., HILL JR R. W., HOVY E., MARSELLA S., RICKEL J., TRAUM D.: Toward virtual humans. *AI Magazine 27*, 2 (2006), 96–96. 2

[Shi21] SHITUT K. S.: *Using AR and VR characters for enhancing user experience in a museum*. Master's thesis, Rochester Institute of Technology, https://scholarworks.rit.edu/theses/10699, 5 2021. 6

[SMG09] STRONG C. R., MATEAS M., GROSSMAN D.: Generative conversation tool for game writers. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (2009), pp. 183–190. 9

[TAA*12] TRAUM D., AGGARWAL P., ARTSTEIN R., FOUTZ S., GERTEN J., KATSAMANIS A., LEUSKI A., NOREN D., SWARTOUT W.: Ada and grace: Direct interaction with museum visitors. In *International conference on intelligent virtual agents* (2012), Springer, pp. 245–251. 2, 9

[Thi10] THIERER J. M.: *Telling history: A manual for performers and presenters of first-person narratives*. Rowman & Littlefield, 2010. 1

[W3C20] W3C: Webxr device api, Jul 2020. URL: https://www.w3.org/TR/webxr/. 4