Scaling Multi-user Virtual and Augmented Reality

by

Sebastian Herscher

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

NEW YORK UNIVERSITY

JANUARY, 2020

Ken Perlin

Dedication

TO MY PUZZLE PIECE.

Acknowledgments

I WOULD LIKE TO THANK MY ADVISOR, PROF. KEN PERLIN, FOR HIS PHENOMENAL AND UNWAVERING SUPPORT THROUGHOUT THE PHD PROCESS.

I WOULD LIKE TO THANK THE MEMBERS AND ALUMNI OF THE NYU FUTURE REALITY LAB, WHO ARE TRUE PIONEERS IN THIS FANTASTIC DIGITAL AGE.

Abstract

The Virtual and Augmented Reality (XR) ecosystems have been gaining substantial momentum and traction within the gaming, entertainment, enterprise, and training markets in the past half-decade, but have been hampered by limitations in concurrent user count, throughput, and accessibility to mass audiences. Although a litany of XR devices have been made available for public purchase, most XR experiences have been developed for either a single user or a small set of users at a time. Few systems or experiments in co-located XR environments have expanded past a small set of users, leaving the paradigm of being part of a larger virtual audience relatively untested. This thesis presents a set of components, systems, and experiments that assist in the creation and scaling of multi-user virtual and augmented reality experiences, and outlines the strengths of techniques found in traditional co-located media for the design space of scaled co-located XR.

Contents

Listing of figures

List of Tables

1

Introduction

Virtual and augmented reality (VR and AR, often combined as XR) have been around since the beginnings of 3D rendering, with Sutherland et al. developing the fabled Sword of Damocles VR/AR headset[1] a full four years before Catmull et al. released their seminal work *A Computer Animated Hand*[2]. A digital world either sitting on top of or replacing our physical world has been a core trope of science fiction for more than half of a century, with a plethora of research being conducted to make said fiction a reality. Historically, these mediums have been locked behind both monetary and expertise barriers, with limited functionality due to a lack of processing power for realtime rendering. However, this most recent iteration of virtual and augmented reality hardwares have piggybacked on vast increases in compute power, and have been successfully delivered to the world at both consumer and enterprise price points.

The wide availability of affordable XR hardware has occurred in tandem with an explosion of interest, investment, and development of awesome immersive experiences and games. These have taken the form of Academy award winning immersive experiences[30], phenomenally successful at home VR games[52], and a new industry of "hyper reality" out of home location based experiences for small groups[25]. Companies have begun to empower their employees no longer develop 3D content using 2D interfaces, instead building and creating with the full fidelity of 3D input[50]. It is clear that our future includes both the use XR for both out of home and in home entertainment, as well as for everyday communication and collaboration in both professional and personal use.

1.1 Collective Audience Perception Spaces

However, for this future to come to fruition, two fundamental issues with the current wave of XR must be addressed. First, development of real-time immersive 3D content must be unified and standardized, such that when an immersive experience is created, it can be ubiquitously distributed across both the litany of current XR devices as well as those that will be released in the future. Second, the design space of XR must be taken past its current limit of 4-6 concurrent users at a time.

The social pattern of congregating together in large, collective groups is consistent across all layers of human social practice. In both professional and personal environments, we come together to experience, learn, and perceive en masse. In the immersive space, the ability to join into a large virtual congregation from a remote access point has been successfully built[56,35], but their colocated counterparts have only brought together a smattering of people in the same physical space.

We must endeavour to build and explore Collective Audience Perception Spaces, or CAPS - physically shared design spaces that include dozens if not hundreds of users that can see, hear, and share an immersive environment as a collective, through both virtual and augmented reality. Each audience member must have their own viewpoint and representation, while being synchronized with the aggregate. These environments mirror their non-immersive counterparts of lecture halls, cinema, and theater, with the added power and capability of spatial computing and immersive design. Whether through using virtual reality to create a shared virtual space, or augmented reality to add to the physical world, CAPS must enable an audience to congregate in a synchronous immersive experience, without need for participation outside of contributing to the collective knowledge and focus of the whole.

1.2 Research Questions

Many research questions arise out of the design space of Collective Audience Perception Spaces. First and foremost, the presence of a representation of an audience in an immersive environment changes both the design and implementation of an experience, as well as the reception and enjoyment of it by each viewer. Furthermore, there is a tension between the amount of agency that can be provided to each individual audience member, and the synchronicity of the experience to the collective. Limitations on both the design of content and systems to support managing the large scale shared system must also be investigated. Finally, the digital nature of an immersive space allows for the powering up of the user with tools or interfaces that are physically impossible. The impact of such tools or interfaces when put in the hands the collective digital audience is well worth investigating.

1.3 Research Contributions

The first chapter discusses the development and testing of a large scale shared VR platform, CAVRN, and a shared theatrical immersive short *CAVE* delivered on it. Based on our evaluation, we observed that a majority of participants greatly enjoyed being part of a live virtual audience, even when they arrived to the event on their own. Furthermore, while they were not afforded direct interaction with other audience members, the collective focus and knowledge of the live audience played an important role in their immersion in and enjoyment of the experience. Many participants wished for greater interactivity in the experience, but wished to interact with the environment or other audience members more than interacting with the narrative of the experience. Many participants noted the tension between allowing for interaction and maintaining a shared collective experience. Finally, many participants noted that a shared collective immersive experience felt like an amalgam of the traditional mediums, with the use of space, depth and format of theater mixed with the special effects and magic of cinema. Both sets of elements assisted in the audience feeling more immersed in the virtual environment.

The second chapter discusses the design and testing of an immersive experience built for both the VR and AR mediums called *MARY*, and an expansion on CAVRN to support greater audience representation in the virtual space. Furthermore, it describes the design, implementation, and testing of view-enablement tools that increase individual exploration without interruption of the collective immersive experience. It was observed that while the content of *MARY* did not change between the two mediums, the experience was well received and enjoyed as two unique and separate experiences by the participants. While the VR version of the experience was more immersive and engaging, the AR version was found to be more novel, theatrical, and accessible. The three view-enablement tools provided, while designed simply to increase visual accessibility of the content, were used in vastly different ways by participants. The addition of the view-enablement tools was found to have a positive impact on the viewing of *MARY* in both VR and AR, and the representation of said tools as digital totems allowed viewers to quickly learn and utilize them.

2

CAVRN: Scaling Up Virtual Reality

We first seek to investigate and explore the design space for a collective VR experience that evokes the feeling of receiving a story as part of an audience. We achieve this goal by i) establishing a set of design constraints for creating such an experience, ii) describing a system, CAVRN, which allows for the delivery of said experience to an audience, iii) describing *CAVE*, a 30-user VR experience built using CAVRN, and

iv) evaluating the strengths and weaknesses of CAVRN and *CAVE* through a qualitative analysis of user surveys and interviews.

The Collective Audience Virtual Reality Nexus (CAVRN) is a lightweight audience and experience management system that enables a physically co-located audience to both embody an avatar and view all other audience-embodied avatars during a synchronized experience. The design constraints used in defining CAVRN are derived from an analysis of the current state of the VR ecosystem and the effective practices and presentations of traditional media.

To evaluate the design constraints as implemented in the CAVRN system as well as an experience built on it, we conducted a mixed methods exploration of user experience which included a survey ($N = 317$) and a semi-structured interview ($N = 21$) during a demonstration of *CAVE* at a technical conference. Over the course of the conference, 1927 people attended *CAVE*, in groups of 30 audience members at a time. The survey asked participants to evaluate elements of the system and their overall experience, as well as gathering background details of their prior experience with VR. Our results provide a description of how co-location affected the respondents' experiences. We also found that among the responses we received familiarity with VR was associated with different judgments of the system and the experience, and that participants yearned for interaction between audience members and the environment over interaction with the story. We reflect on our findings and their implications on the proposed design constraints. Furthermore, we explore the unique aspects and design challenges of such an experience. We hope this work will help inform the design of future experiences and platforms that support multi-user VR.

2.1 RELATED WORK

This work builds on Storytelling in XR and telepresence.

There have been many efforts to bring ideas from theater into the world of VR. As Dixon[11] and McKenzie[6] have discussed, researchers and artists have explored immersive, head mounted display (HMD)-based theater since the early 1990's. Many of these efforts laid the groundwork for the exploration of user perspectives in a VR experience, user interaction in VR entertainment, and the role of VR in immersion. However, due to limitations in hardware, these projects had a limited impact on mainstream media.

Despite these limitations, the relationship between VR and theater has been discussed from a theoretical standpoint for many decades. Murray's book, "Hamlet on the Holodeck: The Future of Narrative in Cyberspace"[57], imagined what the future of theater and storytelling might look like in immersive, digital environments. While the scope of her discussion goes well beyond this work, some concepts, such as the evolution of immersion and perspective within a virtual world, were critical inspiration for our work.

Additionally, studies have compared the sense of immersion and emotional response in VR and traditional media. Some research has found that VR, in comparison to 3D film, enhances both immersion and emotional response[12]. Furthermore, room-scale VR has been evaluated as more immersive than 360 video due to a higher sense of spatial presence[39]. Others believe that this enhancement is even more extreme, comparing the sense of immersion experienced in VR to be more akin to dreaming than reading a book or watching a movie, due to the occlusion of the physical world[5] as well as the ability to physically interact with the virtual environment[10].

In recent years, studios have turned their attention to VR media productions. Some works, such as “Arden’s Wake”[32] and “Pearl”[21], have earned critical acclaim, demonstrating an expanding appreciation of the immersive medium. However, these productions focus on single-person experiences and do not put any emphasis on experiences for multiple viewers. McBurney’s “The Encounter”[19] explored the possibilities of virtual spatialized audio, but did not offer any virtual visual component. Other efforts, such as “Holojam in Wonderland”[36] and “The Apple”[38], have focused on co-located audiences of up to four people.

Many productions have turned to creating interactive storytelling VR experiences for small groups of people. The Void[45], Hologate[51], and Dreamscape[48] are all examples of experiences that enable rich interactions with the environment, closer to the experience of a game. However, these examples do not scale to large co-located audiences within the same concurrent experience.

While those experiences are currently focused on interactivity, the popularity of movie theaters demonstrate that there is a large audience for passive, shared experiences. The Movie Picture Association of America’s (MPAA) 2017 annual report found that at least 76% of the U.S. and Canadian population visited the movie theater at least once[41]. While the report details that home theater systems have been on the rise, theater attendance remains strong, which suggests that people still value the experience of the movie theater.

Others have focused on creating beam-in experiences, where many telepresent users can concurrently view the same piece of media. Systems like Oculus Venues[35] and Altspace[40] seek to bring users from different physical locations into the same virtual viewing experience. VRSync[42] allows for synchronized audiences within the same physical space, but focuses on 360 video experiences.

Other research has shown that emotional responses to experiences are amplified when they are shared with other people[15]. This set of studies indicated that individuals' levels of enjoyment and satisfaction with an experience increased when that experience was shared with others. This demonstrates the value of shared experiences, even when there is little to no interactivity between users.

2.2 Design Constraints

Given the limitations of throughput and accessibility found in the current VR ecosystem, we set out to form a set of design constraints for enabling a much larger VR experience which enables a unique rendered viewpoint, while maintaining as much of the fidelity of the medium as possible. We looked to traditional media and the prior works described above as a foundation for the following four constraints:

1. **Audience should be represented in the experience**: Prior work has shown that representation of audience members or collaborators affords a higher level of immersion and enjoyment of an experience.

2. **Local proximity between audience members should be preserved**: Co-location with neighbors and/or collaborators has been demonstrated in prior work.

3. **Audience should remain seated**: Due to the current limitations of VR tech-

nology, free roaming experiences do not scale well to a large audience. Furthermore, traditional media has shown success in having a seated audience for experiences, which simplifies logistical and technological constraints.

4. **Interaction should not be required to drive the experience**: Interaction within an experience often increases complexity with every added participant, which can hamper scaling and throughput. Furthermore, both traditional media and 360-degree VR experiences have had success with limited to no interactivity.

2.3 CAVRN

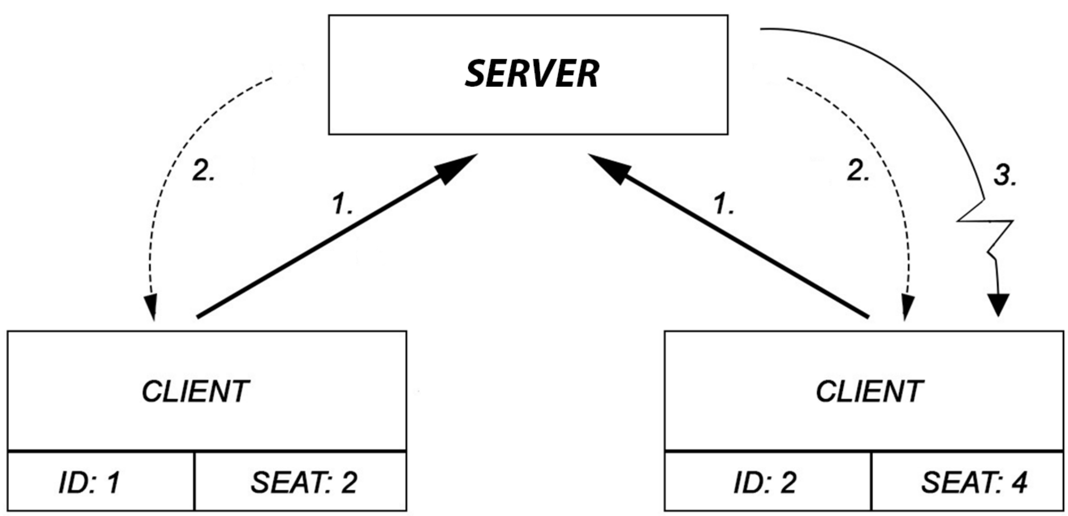

Figure 2.1: The software architecture relationship between the server and all clients. (1) All client applications report their position and orientation to the server. (2) The server sends experience and client state data to all client applications. (3) The server sends an event signal (e.g. seat assignment) to the client.

In order to explore the efficacy of the design constraints described above, we present the Collective Audience Virtual Reality Nexus (CAVRN). The CAVRN system uses a lightweight software architecture to synchronize both the experience and the head position and orientation for each connected client. This data is then used to render a virtual audience in a seated format (Figure 2.3), in which each virtual audience member is controlled by the viewer sitting in the corresponding physical seat. The implementation of CAVRN used for evaluation was built using the Unity game engine[44], and was executed on Google Lenovo Mirage Solo standalone headsets[37] for each audience member, and a Samsung Galaxy S8 smartphone[33] as a control unit. The CAVRN system uses a server to synchronize and maintain the state of the experience and its users, which is then distributed to all clients (Figure 2.1). In our implementation, the server and clients send data at a rate of 20Hz (*r=20*).

2.3.1 Client

Each client consists of a VR application deployed to a standalone VR headset. Each client is responsible for rendering the experience from an assigned virtual location, referred to as a "seat." The seat determines a unique viewpoint of the experience for each viewer, as each seat is placed in a different virtual location in the experience. Each client sends the following data to the server:

- Client Packet Identifier (1 byte)
- IP Address (4 bytes)
- MAC Address as ID (6 bytes)
- Local Position (3 floats, or 12 bytes)
- Local Orientation (4 floats, or 16 bytes)
- Seat Index (1 byte)
- Timecode (1 float, or 4 bytes)
- Logistic Info (e.g. battery, 4 bytes)

The client unicasts a total of 48 bytes to the server via a standard UDP protocol. In the authors' implementation, the send rate was set at *r=20*. While the timecode and logistic info are not necessary for reconstructing the client as a virtual audience member, these pieces of information were found to be critical in the deployment and maintenance of the CAVRN system.

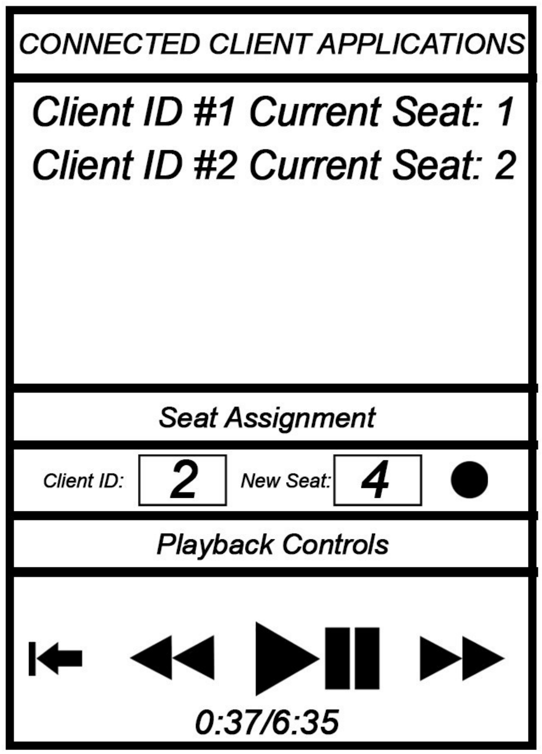

Figure 2.2: An example layout of the server. Note that the status of each client is visible, and that the server can assign a seat to a client, and control the playback of the experience.

2.3.2 Server

The server is a desktop application responsible for controlling the state of the experience, consisting of the current time location of the experience and the virtual seat assignment of each client. The server has the ability to reset, rewind, play/pause, and fast forward the experience, and maintains a master timecode. We found that a simple graphical user interface as seen in Figure 2.2 was sufficient for an operator to control an experience.

As the server receives information from each client, it maintains the most recently reported state of each client. The server is responsible for sending the following information to all clients:

- Server Packet Identifier (1 byte)
- All up-to-date client data (35 bytes * n active clients):
  - MAC Address as ID (6 bytes)
  - Local Position (3 floats, or 12 bytes)
  - Local Orientation (4 floats, or 16 bytes)
  - Seat Index (1 byte)
- The master timecode (1 float, or 4 bytes)

The server does not re-distribute the client IP addresses, timecodes, and logistic info, as these pieces of information are not necessary for reconstructing a client as an audience member. The server can either unicast or multicast a total of $35{*}n + 5$ bytes to all connected clients via a standard UDP protocol. In the authors' implementation, the server's send rate was set at *r=20*.

2.3.3 Server to Client Events

Some server-client interactions, such as assigning a seat to a specific client, are fixed events that only need to be sent once. These interactions are sent from the server to a client via a standard TCP protocol, and consist of the following structure:

- Event Packet Identifier (1 byte)
- Target Client MAC Address as ID (6 bytes)
- Event Info (e.g. seat index, max 4 bytes)

2.3.4 Synchronization

As the server sends a master timecode of the current point in the experience, each client adjusts its rendering of the experience to that moment in time. Rather than taking effect immediately, this synchronization signal can be used to slightly alter the rate of the clock that is running the experience. The difference *(dt = master timecode - local timecode)* is computed. The local game clock rate is then given as the local experience's internal clock rate multiplied by a factor of $(1+dt/N)$, where N is a constant factor for making the time adjustment more gradual. In the authors' implementation, any dt greater than five seconds were handled with setting the timecode directly to the master timecode, and $N = 2$. In practice, this technique ensures that there will not be any noticeable difference in perceived time between clients, and that there will not be any perceived sudden time shifts as a result of this synchronization process.

Figure 2.3: The representation of the audience in *CAVE*. Each audience avatar was driven by the position and orientation of the corresponding headset. Note that all audience members can see the position and orientation of the heads of all other audience members, which are calculated using the data received from all other users' VR headsets.

2.3.5 Audience Representation and Format

The CAVRN system currently requires the audience to be physically seated. There are two main reasons for this requirement: 1) it promotes audience safety by mitigating the risk of injury due to collision with objects or persons not fully represented in the virtual experience, and 2) the full virtual position and orientation of the viewer in the experience can be reconstructed using the local position and orientation of the VR headset, as well as the virtual position of the seat to which the viewer is assigned. With the viewer physically seated, their virtual representation is rendered as an avatar in a seated position. The spine, chest, and head of this avatar are connected via an inverse kinematic (IK) chain, and each arm is also rigged using a three-point

IK chain. As the viewer's head movement changes the position and rotation of their headset, the avatar that represents them in the virtual world updates to replicate their body movements. For example, if a user were to rotate and lean to the left, the avatar representing them in all other client applications would rotate their chest/head and lean their body in the same fashion. In the authors' implementation, the audience is reconstructed in the client application, after it has received the up-to-date data from the server. Only audience members in the view frustrum of the client are calcuated, to reduce computational overhead. Linear interpolation for position and orientation are used to smoothen out the movements of the audience, as the send rate (r=20) is lower than the refresh rate of the experience (60Hz).

2.3.6 Complexity and Technical Insight

In building CAVRN, considerations for both size complexity and wireless transmission packet sizes were used for defining client and server data structures, as well as data flow patterns. Best practice guidelines found in[28] were used, with an assumed size of 576 bytes for an unfragmented packet. With this size, all client packets are left unfragmented, and server packets are fragmented once per 14 users.

With n users and a send rate of r, the total amount of data can be calculated:

$$\mathbf{(48\boldsymbol{n} + (5+35\boldsymbol{n})\boldsymbol{n})\boldsymbol{r}}$$

Both size and connection complexity are $O(n^2r)$ if a unicast model is used for all client and server communications. Both complexities are reduced to $O(nr)$ if a broadcast protocol is used for sending data from the server to each client; however, both the hardware used in the authors' implementation and most mobile VR hardware that is available on the market provide poor broadcast support. In practice, while these devices can send and receive broadcast signals, the quality is low enough to thoroughly degrade the fidelity of an experience.

2.4 *CAVE*

The experience used for the evaluation of the CAVRN system consisted of a six minute VR narrative called *CAVE*. *CAVE* includes many dynamic visual moments, including the arrival of a woolly mammoth which was significantly larger than the viewers into the virtual space. The experience also includes a dynamic set of visual assets for the audience to look at and explore, such as moving cave paintings on all walls of the virtual set during a portion of this experience. These dynamic elements were incorporated into the experience in order to explore their impact in a VR environment. *CAVE* was designed as a way for dozens of audience members to share an immersive experience simultaneously, as an illusion that they were experiencing a live theater event together. This works as intended to be received as an experience that is fundamentally different from VR for an individual. Each audience member sees and hears the experience from their own viewpoint, while sharing the experience as a collective audience.

A 30-member virtual audience was implemented in *CAVE*, separated into two groups in a thrust stage theater winged format (Figure 2.5). While the CAVRN system could allow for any mapping between physical and virtual audience position, maintaining physical and virtual correspondence between audience members was

Figure 2.4: An example of how a piece of content in *CAVE*, the mammoth, is rendered in comparison to the virtual audience. Note that the front row of the audience avatars is rendered in between the virtual content and the back row of the audience avatars. This is vital in evoking a sense of audience presence in the virtual experience.

Design Constraint 2. The left group consisted of two rows of six avatars, totalling 12 seats. The right group consisted of three rows of six avatars, totalling 18 seats. Each row was physically and virtually consistent. In the virtual world, but not the physical world, the theater was raked: Each audience row was virtually raised 0.5 meters higher than the row in front of it, allowing members in back rows better viewing without the need of a physically raked audience. The two groups were significantly further separated from each other virtually than in their physical setup, as the virtual environment of the experience was significantly larger than the physical room in which it was shown. In sum, the physical and virtual distance between an audience member and their neighbors was consistent, while the virtual and physical distance between the rows in the groups and the groups themselves differed.

The audio of the experience was spatialized by placing the virtual audio sources onto the objects that were responsible for producing them. This allowed for the audio to sound as if it were coming from the correct direction in the virtual environment.

2.5 Evaluation of CAVRN and *CAVE*

2.5.1 Methods

We performed a mixed methodology study to capture participants' responses to *CAVE*, contrast *CAVE* with other media experiences, and qualitatively evaluate the effectiveness of the CAVRN system. The study consisted of a 43-item survey and a semi-structured interview conducted immediately following experience viewings. The full text of items can be found in the Appendix.

2.5.2 Survey

The survey was implemented in Google Forms and administered from six tablet computers immediately following viewings. A convenience sample of 374 participants began the survey (out of 1927 audience members, for a response rate of 19%). Due to technical difficulties, 57 did not complete the survey, leaving 317 complete responses. Participants reported the row in which they sat during experience. All rows were represented in the sample, with fewest responses from the right front row (57, or 18% of the sample) and the most responses from the left front row (76, or 24% of the sample).

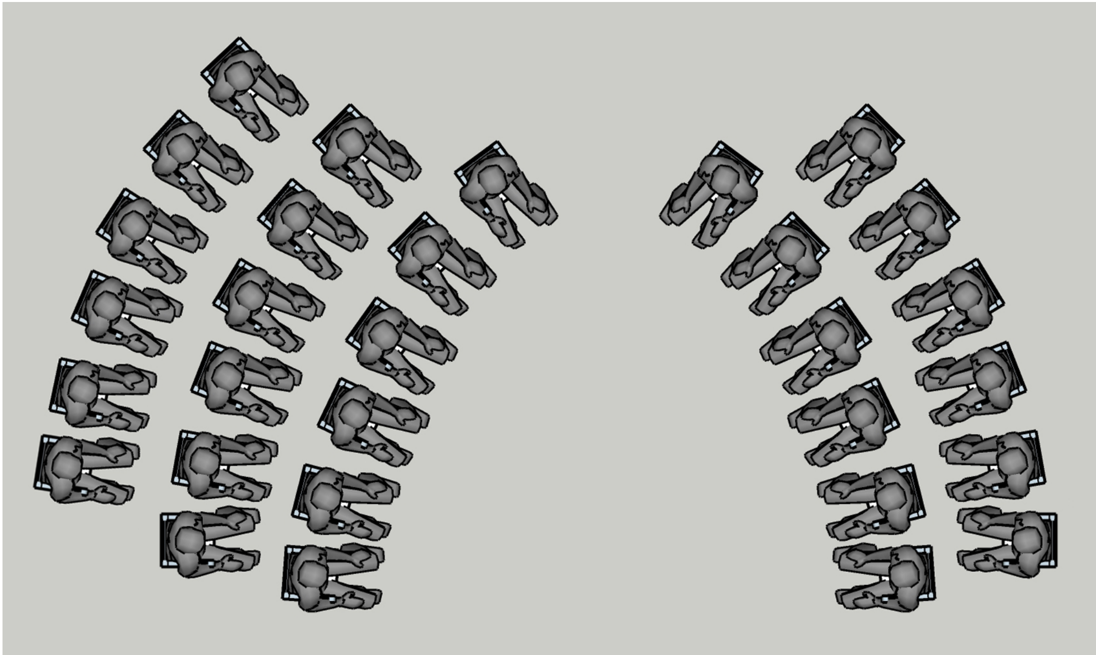

Figure 2.5: A rendering of the physical layout used for *CAVE*. The audience was split into five rows: front left, back left, front right, middle right, and back right.

2.5.3 INTERVIEW

Twenty-one participants additionally took part in a 5-15 minute interview, conducted in an adjacent screening area. The interviews were recorded and later transcribed prior to analysis.

2.6 RESULTS

The results have been organized based upon general themes that emerged from analysis of both the survey and the interviews. We first provide a profile of respondents and then an analysis of the expertise categories of the viewers, based on the expertise questions asked in the survey. We present a combination of the open ended responses

ID	*Title*
P1	VR Programmer
P2	VR Practitioner
P3	VR Master's Student
P4	VR Expert Media Designer
P5	Photographer
P6	Media Group Manager
P7	Computer Science Master's Student
P8	Chip Designer
P9	Director of University Lab
P10	Projection Research Engineer
P11	VR UX Designer
P12	Professor of Computer Science
P13	VFX Designer
P14	Programmer
P15	Professor of Computer Science
P16	VR Expert Programmer
P17-P21	Unknown

Table 2.1: Summary of interview participants.

from the survey, and responses from the interviews, to explain the findings in greater detail.

2.6.1 Profile of Participants

275 of the 317 participants provided a short description of their profession, which was coded independently by 2 of the authors with a very good inter-rater reliability (Cohen's Kappa) of 0.85, 89% agreement. The distribution of professions is presented in Table 2.2.

What is your professional role?

Profession	*Count*	*Percent*
Technical	94	34%
Academic	51	19%
Creative	49	18%
Administrative	47	17%
Students	34	12%

Table 2.2: Aggregate survey responses for reported professions.

How long have you used VR devices?

Time	*Count*	*Percent*
No experience	24	7.6%
0-6 months	47	14.8%
6-12 months	36	11.4%
12-36 months	130	41.0%
36+ months	80	25.2%

Table 2.3: Aggregate survey responses for prior experience with VR devices.

2.6.2 Familiarity with Virtual Reality

When asked about explicit prior experience, over half of those surveyed (210 participants, or 66%) reported having used VR for a year or longer. Of those reporting less than a year's experience, 24 (8% of total) reported this was their first ever time using VR. A complete breakdown of survey responses is given in Table 2.3.

Independent from prior experience, participants were asked to rate their expertise using VR devices on a scale from 1 (Not at all proficient) to 7 (Extremely proficient). The average rating in this sample was 4.79 (SD = 1.63). Most participants (202 participants, or 64%) gave themselves a rating of 5 or higher.

Based on the amount of time and self-reported expertise, we created three categories of familiarity with VR: High familiarity participants (86 participants, or 27%) had used VR for a year or longer, and rated themselves a 6 or 7 (Extremely proficient); Low familiarity participants (88 participants, or 28%) had used VR for less than a year, and rated themselves a 3, 2, or 1 (Not at all proficient); and Medium familiarity participants (143 participants, or 45%) provided any other combination of time and rating of expertise.

Figure 2.6: An audience watching *CAVE* at the technical conference.

2.6.3 Co-Location as a unique feature of the experience

The co-located nature of the VR theater experience provided both a satisfying experience and a challenge to participants' expectations for this hybrid format.

The majority of participants came to the performance alone (179 participants, or 57%), with the remainder coming with 1 other person (61 participants, or 19%), 2-5 people (63 participants, or 20%), or more than 5 people (14 participants, or 4%). Most participants felt they were "part of a live audience" (1 = Strongly disagree, 7 = Strongly agree; 217 or 68% gave ratings of 5 or above, $M = 4.96$, $SD = 1.97$), and that "being part of an audience in the experience was enjoyable" (223 or 70% gave ratings of 5 or above, $M = 5.24$, $SD = 1.58$).

Participants were also asked how specific elements of the experience compared to cinema and live theater experiences. While music, characters, and dialogue were the most frequent categories selected as possible to replicate in cinema or theater, the audience format used in this experience was seen as more difficult to replicate: this option was the least frequent response for cinema, and the bottom third for theater.

Open-ended responses to the question "Are there any elements not listed above that you believe could NOT be replicated?" highlighted the role of co-locating audience members. The audience was mentioned by 15% of the 87 respondents to this item ('immersion' was the most common, with 34% of responses). General comments included *"The feeling of the presence of the others, togetherness"* (S132). Participants highlighted the design choice to transform the appearance of audience members, although their position and movements reflected the individual: *"Seeing other audience members in a new form"* (S45) and *"The similarities between audience members and their ghostly appearance"* (S134). Participants also mentioned how the audience members affected their own experience, with comments like: *"The use of the helmets [in the content] to emphasize the orientation of others' heads in the experience encouraged a look around and exploratory experience"* (S210) and *"Feels a lot more like theater, but scale would not be replicated nor the audience as participating chorus"* (S27).

The comments from the interviewees also provided more insight into the experience of being part of an audience. For example, some reported being part of the co-located virtual audience improved the experience:

> (P1) It was the first thing I noticed. So [you] are in [the experience] but you don't seem alone. You don't think you're alone. There's nothing you

feel you can feel scared of.

Others reported that while being part of the audience was enjoyable, direct interaction with others was not necessary:

> (P3) I don't know if I want to like or talk to the other person you know. I don't want that, but I like being with everyone else. It's pretty cool. I like the fact that there are other audience because it really immerses me, and makes me think that I'm in a theater.

2.6.4 Interaction during the experience

Although the primary focus of the experience was for the audience to receive a story, many participants "felt [they] wanted more interactivity in the experience" (1 = Strongly disagree, 7 = Strongly agree), with most giving responses of 5 or higher (226 participants, or 71%); $M = 5.2$, $SD = 1.68$).

Participants did not feel they "could interact with neighboring audience member(s)" ($M = 3.16$, $SD = 1.84$), and most (240 participants, or 76%) did not attempt to do so. Of the 77 participants who reported in the affirmative, the most common modes of interaction were verbal (53 participants, or 69%) and virtual (i.e. avatar-to-avatar) (53 participants, or 77%), with the least common mode being direct physical interaction (14 participants, or 18%).

Although only a few elaborated on how to implement interactions, many participants expressed what they wanted to interact with: other audience members (24 participants, or 18%) and the characters at the center of the story (14 participants, or 10%), especially being able to speak to them; or interacting with the environment in general (43 participants, or 32%) and examining or adding to the content more specifically (14 participants, or 10%).

What elements do you believe could be replicated in traditional media?				
	Music	*Characters*	*Dialogue*	*VFX (Particles)*
Cinema	256	249	231	214
	80.8%	78.5%	72.9%	67.5%
Theater	225	211	213	**115**
	71.0%	66.6%	67.2%	**36.3%**
What elements stood out to you in *CAVE*?				
	82	51	123	149
	44.8%	25.9%	38.8%	47.0%

Table 2.4: First part of the aggregate survey responses for what users believe can be replicated in traditional forms of media, and what elements of *CAVE* stood out. Note that the *Scale* category refers to the scale of the content in relation to the scale of the audience in the experience.

What elements do you believe could be replicated in traditional media?				
	VFX (Cave Paintings)	*Environment*	*Scale*	*Audience Format*
Cinema	186	179	152	**123**
	58.7%	56.5%	47.9%	**38.8%**
Theater	186	185	**126**	163
	58.7%	58.4%	**39.7%**	51.4%
What elements stood out to you in *CAVE*?				
	160	163	**191**	**181**
	50.5%	51.4%	**60.3%**	**57.1%**

Table 2.5: Second part of aggregate survey responses.

Some participants expressed a desire to direct their own experience. This ranged from being able to change their visual perspective during the experience (13 participants, or 10%); to incorporating responsive features to existing elements of the story (11 participants, or 8%) such as the three participants who mentioned gaze-based interaction with some of the content; to a more autonomous experience (5 participants, or 4%) where participants described changing the course of the story, or as S178 said, to *"interact with the main story happening, and me moving through sub stories."*

A few participants struck on the tension between interacting with the material and the experience in its current design, with comments like *"More audience interactivity or e.g. drawing on the walls - but that would change the experience and I don't feel it needed interactivity necessarily"* (S11), although others placed a boundary on their desire for interactivity by focusing on how it would allow *"seeing things that wouldn't be possible in live theater"* (S275). Some were concerned with the effects of interactivity both in terms of story, and logistics: *"I wanted to maybe look around corners, but seating kept me stabilized. I feel being in a dark room might make it dangerous to walk around because of a potential loss of orientation"* (S246). But others rejected the need for interaction altogether, *"None. I want to sit back and watch the show."* (S314). In the experience, no character acknowledges the presence of the audience. One interviewee noted this, and expressed:

> (P15) I would have preferred if there was more dynamic somehow between the digital characters and the virtual audience.

2.6.5 Cinema and Theater as useful points of comparison

One of our intentions in the exploration of collective VR was to adapt the conventions of traditional media like cinema and live theater, and examine whether specific features of this performance offered a unique set of affordances for creators to use, resulting in a new experience for audiences. To examine whether these choices were effective, we asked participants whether elements of the system and experience could be replicated using other media.

To contextualize their responses, participants reported how often they attended a theater or movie theater. Participants were casual consumers of these experiences: most participants attended movies about once a month (84 participants, or 27%) or less than once a month (114 participants, or 36%), similar to industry reports that moviegoers purchase an average of 4.7 tickets each year[41]. Most participants (169 participants, or 53%) reported attending a theater event less than once a month, which is also similar to surveys of performing arts attendance (averages of non-musical plays 2.2 per year, or 4.8 per year for any 'benchmark' performance types as of 2012[58]).

Participants were asked whether specific elements of the experience could be replicated in cinema and live theater in two multiple selection items, and could select as many or as few elements as desired; frequency of responses and percent of participants providing that response are displayed in Table 2.4.

Music, characters, and dialogue were selected as elements of the experience which were possible to replicate in cinema or theater most frequently, with 73-80% of participants responding that it would be possible to replicate these elements in cinema, and 67 - 71% of participants responding that it would be possible to replicate these elements in theater.

A majority of participants (57 - 68%) also responded that it would be possible to replicate the special effects (57%) and environment (68%) in cinema. Representing the physical scale of the environment and the audience format were the least frequently selected as elements that could be replicated in a cinema (receiving 48 and 39% of responses, respectively).

In contrast, for theater, a majority of participants highlighted that a special effect (moving cave paintings), environment, and audience format (51 - 59%) could be replicated, with scale and effects (explosions and particles) receiving the fewest responses (40 and 36%, respectively).

To contrast with the items above, participants were also asked which elements of the experience stood out to them. The ranking of features is displayed in Table 2.4, and reverses the placement of scale and audience format from the previous items - these elements were seen as more difficult to replicate in traditional media.

Comparisons to cinema and theater were often used by participants to describe the experience, situate their responses, and highlight the combination of features making this experience unique. For example:

> (P1) I can see it as ... it feels more like a theater. More intimate experience. And you can actually have the audience just around you. I mean, with the effects that you can get in cinema.

Combining the audience format typical of live theater, with spatialized special effects was especially surprising, and frequently commented on by participants:

> (P2) I like how something could happen that couldn't have been real theater like the mammoth coming in for example. I think I would like to see more of that.

Participant 13 shared this feeling, and went into more detail about how the effects enhanced feeling transported from the physical space, and immersed in the virtual environment:

> (P13) The thing that was most different about this is when the mammoth appears and you have to look up at it. So being in that enclosed space and having something that large come in where you feel overwhelmed by this giant thing that then explodes and magic. That's something that you don't get sitting in a theater unless maybe if you're at an IMAX or something. But you don't get that feeling of there's something right there.

Others described feeling transported by the environment:

> (P4) It makes it where it take you away from the real life environment, and it actually puts you in there ... to more expand on to the captive audience. It puts the audience directly into the movie itself.

2.7 Discussion

We now reflect on our findings in how the CAVRN system enabled an audience of 30 co-located users to share a collective VR experience, *CAVE*, and evaluate the efficacy of our proposed design constraints (DC). We provide insights to inform future research and design in the HCI community.

2.7.1 Using Audience Format as a Design Choice (DC 1, 2)

Based on our analysis, we observed that the majority of participants both understood and enjoyed that they were part of a live audience. Since most of the participants came alone, it was significant that the presence of others was still an overall positive experience, and that virtual representations of other real audience members helped rather than hindered participants' enjoyment of the experience.

Furthermore, despite the lack of direct interaction between audience members, implicit interactions between audience members played an important role in their experience. Simply seeing the other audience members was a memorable aspect of the experience for some of the survey respondents and interviewees. Others mentioned that not only were the other members noticeable, but they improved experience immersion and provided valuable clues about where to look throughout the experience.

Consistently, participants noted that many aspects of being physically co-present with other audience members facilitated their enjoyment of *CAVE*. Being able to talk to their neighbors, reach out and hold hands, and share the experience with other people sitting next to them was repeatedly reported and discussed.

Finally, through the use of the CAVRN system, 1927 audience members were able to experience *CAVE* over the course of only 4 days at a technical conference. This level of throughput is not often seen in collective VR or AR exhibitions, and to our knowledge, has not yet been seen outside of our experience.

2.7.2 Supporting Richer Audience Interactions (DC 4)

Many viewers reported that they wished for more interactivity in the experience. Even with the representation of the audience as virtual avatars, some participants felt that they were unable to interact with neighboring audience members, and tended to only interact through speaking to one another. A common interaction requested by the participants was the ability to move around the virtual space. However, some participants noted that this type of interaction could change the nature of the experience, and might not be necessary.

Participants reported being more interested in interacting with the virtual environment or the other audience members, as opposed to directly interacting with the virtual characters or the narrative. Enabling richer audience-to-audience and environment interaction might satiate a participant's urge for interactivity while concurrently allowing that particular audience member to receive the narrative as a passive user. Furthermore, passive acknowledgement of the virtual audience by the prerecorded character, in the vein of a "wink from an actor," may satiate audience-content engagement without explicit interactivity.

2.7.3 The Medium as an Amalgam of Traditional Media (DC 3)

Participants reported that content scale and audience format stood out the most in the experience. They also reported that the special effects present in the experience were not necessarily replicable in a theater experience. The adaptation of both cinematic and theatrical elements offers a unique set of affordances for content designers and producers. Design decisions of audience format and content scale drawn from the theater can be leveraged to strengthen the fidelity of an experience. At the same time, computer graphic special effects commonly found in movies that would not be replicable in theater can be introduced. Both sets of elements assist in helping the audience member to feel more immersed in the virtual environment. In sum, the combination of the physicality of theater and the special effects possible in movies created a uniquely positive experience for our participants.

3

MARY: Experimenting Across Platforms

After the success of both CAVRN and *CAVE*, we set out to further investigate facilitating large scale multi-user immersive experiences in two ways: expand the agency of each user within a shared immersive experience without having that expansion detract or interrupt the collective experience, and investigate the enjoyment of an immersive experience across multiple viewing paradigms (VR, AR, and 2D).

Many single-user XR experiences to date have afforded the viewer to have great sway over the flow and pacing of the narrative, such as directly interacting with the main character and being required to interact with the environment to drive the story forward. These are utilized in attempt to increase a sense of immersion in the viewer, but can often lead to the breaking of immersion if the interaction fails. Furthermore, affording agency like direct interaction with a character can be prohibitive as an experience is expanded for concurrent viewing, and don't build on the established interaction spaces found in cinema and theater.

Furthermore, a majority of XR experiences released to date have been designed and implemented solely for one medium, to the extent that migrating the experience to another medium would thoroughly degrade and/or change the experience. This fact is completely orthogonal to the ubiquitous distribution of content made in the 2D format - a movie can be viewed in a multitude of ways and on a multitude of hardwares, without changing or degrading the fidelity of its content.

We wished to expand our investigation into large-scale collective XR by tackling the previously described problems. We tackled this goal through i) describing an expansion of both the design hypotheses and implementation CAVRN to support greater audience representation and augmented reality anchors, ii) describing the implementation of a new multi-user XR experience, *Mary and the Monster* (*MARY*), which is delivered synchronously to audience members in both VR and AR, and iii) describing a set of view-enablement tools, a pair of opera glasses in VR, a magnifying glass in AR, and a viewfinder in VR, that expand a viewers' agency in exploration during the experience.

The first chapter of *MARY* was demonstrated at a technical conference in 2019, where over 500 viewers went through the experience in groups of 8 at a time. Unlike *CAVE*, no surveys or interviews were conducted on-site due to logistic and physical constraints. To begin evaluation of the work described above, we performed an initial exploration of user experience using a semi-structured interview (N=8) with academic and industry professionals in XR. Our results provide a description of the effect of experiencing the content in different mediums, and the impact the view-enablement tools had on the viewing experience. We reflect on our findings and their implications, and explore the unique aspects of designing such an experience. We hope this work will help inform the design of future XR multi-user experiences and platforms.

3.1 Related Work

This work builds on the previous related work of CAVRN, as well as XR for multiple users and tools for XR.

The past several decades have provided a strong foundation for multi-user VR systems and collaborative experiences in VR. Earlier work in the 1990's, such as the Distributed Interactive Virtual Environment (DIVE)[4], created an initial foundation for collaborative VR environments. Other work throughout the next decade continued this line of thought in their outlines of collaborative virtual environments[8,9].

With more recent developments in VR and AR technology, extensive work has explored the use of VR to enhance the interaction of various tasks for multiple users[18], such as teleconferencing[16] and collaborative scene building[46]. The goal of these works was to design interaction techniques for tasks such as navigation, view sharing, and conflict prevention. While these works focus on rich-interaction and lean-forward experiences, our work, on the other hand, focuses on lean-back and co-located virtual experience.

Other studies have measured the impact of VR for collaboration. Prior research has outlined the power of VR and AR for telepresence and remote collaboration[20,43,17]. Several methods for enabling these forms of telepresent collaboration have been developed, such as recent advancements in pose reconstruction algorithms[23,27,34]. Other research has shown that the use of AR and VR technology improves remote collaboration as compared to previous methods, such as video conference calling[24] or no digital interface[29]. ShareVR explores asymmetric co-located VR gaming and found that users both inside and outside VR value the ability to physically engage with each other. When comparing remote presence to physical presence, research has demonstrated the importance of physical presence[22]. While this could be a consequence of the technological limitations of contemporary AR, this does demonstrate added value of the physical presence of others, even in a virtual experience.

Prior work has created virtual environments to support co-located, large audiences. One such work is the Cave Automated Virtual Environment (CAVE)[3,14,7,13]. The CAVE system can support a large number of users if the room is large enough, although only a single tracked user will have a perspective-correct view of the virtual world. In a similar vein, the global growth of premium large format movie theaters, such as IMAX[47] and ScreenX[26], indicates market interest in more immersive cinematic experiences[31]. However, CAVE, IMAX, and similar systems cannot produce a unique rendered viewpoint for each user, which HMD-based solutions can.

3.2 Design Constraints

Given the results found in *CAVRN* and the current presentations of cinematic content in the XR ecosystems, we set out to experiment further with design constraints for creating an immersive experiences across multiple platforms, while maintaining as much of the fidelity of each viewing medium as possible. Moreover, we looked to expand each viewers ability to explore the immersive content while maintaining an uninterrupted synchronous group experience. We looked to both traditional media and the results of prior work to develop the following three additional design constraints:

1. **Content of experience should be consistent across all mediums**: Traditional media distributes content consistently across multiple viewing paradigms.

2. **Audience should be afforded tools for greater exploration**: Many audience members of *CAVE* reported wanting to explore more of the environment. We should provide view-enablement tools that allow for the exploration of the digital world without physical or virtual locomotion.

3. **Audience should be as fully represented as possible in each medium***: Audience members of *CAVE* reported a wish for further representation in the experience. While this design constraint was used in the implementation in *MARY*, it was not investigated in our study, and will not be discussed as a result of the experiment.

3.3 Expanding CAVRN

In order to explore the design constraints described above, we expanded the capabilities of the CAVRN system. We increased the client data to include pose and state data for each hand, as well as state data for avatar customization. Our implementation of the augmented CAVRN was built using the Unity game engine, and was executed on Oculus Quest standalone or Oculus Rift headsets[49] for each VR audience member, Magic Leap One augmented reality headsets[53] for each AR audience member, and a Samsung Galaxy S8 smartphone[33] as a control unit. For the expanded implementation, as the fidelity of the user data was greatly increased, we increased the server and client send rate to 60Hz (r=*60*) to make reduce pose latency to a negligible level.

3.3.1 Adding Pose and State Data

As both the VR and AR headsets used in implementation support 6-dof hand controllers, we added the pose data of each hand to the data sent by each client. We also included one byte for each hand for state representation, and one byte for client state identification, which was used for avatar customization. Finally, the byte included for seat index is only used by VR clients, while it is used by AR clients as an AR anchor calibration status which will be discussed in the next section.

Including the pose and state data of each hand of the user would expand each client packet to 107 bytes. In order to reduce both packet size, all float representations were reduced to a signed short representation. For example, a position value of 23.14159 is converted to 23141, reducing the size of its representation from 4 bytes to 2 bytes, but also reducing granularity to millimeters within a roughly 60 meter cube. As the values of a quaternion are represented as a range of 0 to 1, their accuracy is reduced to steps of 1/65,535. In practice, this compression produces no visible reduction in quality, while allowing each client packet to be reduced to a total of 61 bytes:

- Client Packet Identifier (1 byte)
- IP Address (4 bytes)
- Hashed MAC Address as ushort ID (2 bytes)
- Local Position (3 shorts, or 6 bytes)
- Local Orientation (4 shorts, or 8 bytes)
- Seat Index / AR Calibration Status (1 byte)
- Timecode (1 float, or 4 bytes)
- Logistic Info (e.g. battery, device type, 4 bytes)
- Client State (1 byte)
- Left Hand Local Position (3 shorts, or 6 bytes)
- Left Hand Local Orientation (4 shorts, or 8 bytes)
- Left Hand State(1 byte)
- Right Hand Local Position (3 shorts, or 6 bytes)
- Right Hand Local Orientation (4 shorts, or 8 bytes)
- Right Hand State (1 byte)

As described in CAVRN, the server only redistributes data necessary for reconstructing a client as an audience member, and now sends a total of $46*n + 5$ bytes to all connected clients.

3.3.2 Reconstructing Hand Representation from State

Figure 3.1: An example of mapping an Oculus Touch controller to an audience avatar. The images show that the hand mapping occurs using button states, instead of physical hand poses.

In order to represent more than just the position and orientation of each of the users hands, we mapped the button states of the Oculus Touch controllers to finger or finger-group states. Each state had one of the following values, taking up 2 bits: NONE, TOUCH, or PRESS. The ABXY and thumbstick buttons were used for representing the thumb, under the presumption that only one button could be pressed at a given time. The Trigger button was mapped to the index finger, and the grip button was mapped to the remaining 3 fingers. This allowed us to represent each of these finger groups to 2 bits, and compress a hand representation into a single byte.

3.3.3 Utilizing AR Anchors

Both the Magic Leap[53] and Microsoft Hololens[55] platforms provide the ability to find and utilize unique world space anchors that are derived from the world mappings of the device. We implemented the AR version of *MARY* on the Magic Leap One, and relied upon Magic Leap anchors, called "Persistent Coordinate Frames"[54] (PCFs). PCFs allow for content to be placed in the physical world hand have it remain in said place over multiple sessions. As a space is mapped with high confidence, the Magic Leap platform derives PCFs from the 3D mapping of the space, which is deterministic across devices. This means that if one device is able to identify a PCF, other devices in the same space should also identify the same PCF without communication with the first device. Every PCF has a unique ID in the form of a string. Furthermore, the Magic Leap platform allows for the detection of horizontal and vertical planes in the real world, replicating them as virtual surfaces in the virtual world. Finally, the Magic Leap platform also allows for the detection and mapping of QR codes to create a position and orientation local to the headset.

We used the Magic Leap PCF, plane detection, and QR marker tracking systems to map the virtual content of *MARY* to the physical world. We established a calibration mode for a Magic Leap device running the experience, which could be enabled by pressing the home button twice on the Magic Leap One controller. When in calibration mode, the system places the root of the virtual content on top of a detected QR marker. The root is then moved to either the position and orientation of the QR marker if found, or moved to the position and orientation of the controller if not. Once the user has positioned the virtual content root, they press the bumper of the controller to bind the content to the closest PCF. If a horizontal plane is detected close to the proximity of the placed QR code, the upwards vector of the virtual root is adjusted to match the upwards vector of the plane. The client can then turn off the calibration mode by pressing the home button once more. The QR code can then be removed, as the mapping of virtual to physical is maintained by the Magic Leap PCF system.

In practice, the Magic Leap system requires high fidelity space mapping to create and restore PCFs. While PCFs persist across reboots of both the experience and the devices, we often noticed that changes in lighting condition or battery life could interfere with the creation and restoration of a binding. In order to detect whether or not an AR device was successfully calibrated, we utilized the Seat Index byte built in the CAVRN system. If a Magic Leap headset has successfully created or restored its calibration, the seat index of the device is reported to as a value of 255. In practice, this allowed the operators of the experience to monitor the calibration status of all connected AR devices from their smartphone control units.

3.4 *Mary and the Monster*

The experience developed using the described design constraints consisted of a two-chapter, eighteen minute immersive narrative called *Mary and the Monster*. *MARY* is the telling of the creation of the story of Frankenstein from the perspective of an 18 year old Mary Shelley. *MARY* includes many dynamic visual moments, including instantaneous and physically impossible scene transitions, and drastic visual and weather effects such as arc lightning over the audience. In an effort to meet the requirements of design constraint 1, we limited the design and implementation of *MARY* to a theatrically staged room. This would allow all of the content of the experience to be anchored to a single location and origin, previously referred to as the root of the virtual content, and all of the action and story of the experience to be contained within that single cohesive context.

Figure 3.2: An overhead view of the VR audience of *MARY*.

For the VR version of *MARY*, the content was presented at life scale, and included a 25-member virtual audience was implemented in a single rounded theater format within the space of the immersive experience. The audience consisted of 6 audience members in the first row, 8 in the second row, and 11 in the third row. Each audience row was virtually raised 0.5 meters higher than the row in front of it, but was not raised physically. Audience members are represented as a torso, hands, and head with a mask. At the beginning of the experience, each audience member is allowed to pick from up to 5 Venetian-style masks to customize their representation throughout the experience. Unlike *CAVE*, where we aimed to have the virtual audience look like the virtual world they were in, we reduced the uniqueness of each avatar to the customized masks. This was to not only reduce the complexity of the avatars for performance, but also to make the audience have less visual impact during the experience.

For the AR version of *MARY*, the content was presented at 1/6th scale on the top of a round table with a black table cloth. No avatar was used to reconstruct the AR audience, as the presentation of *MARY* in the Magic Leap One allowed each viewer to be represented by their physical body. However, each VR audience was reconstructed within the AR experience, as a 1/6th scale representation of themselves consistent with their position and orientation in the VR experience. This scaled down presentation of *MARY* was agreed upon due to the limited field of view available on AR devices during development.

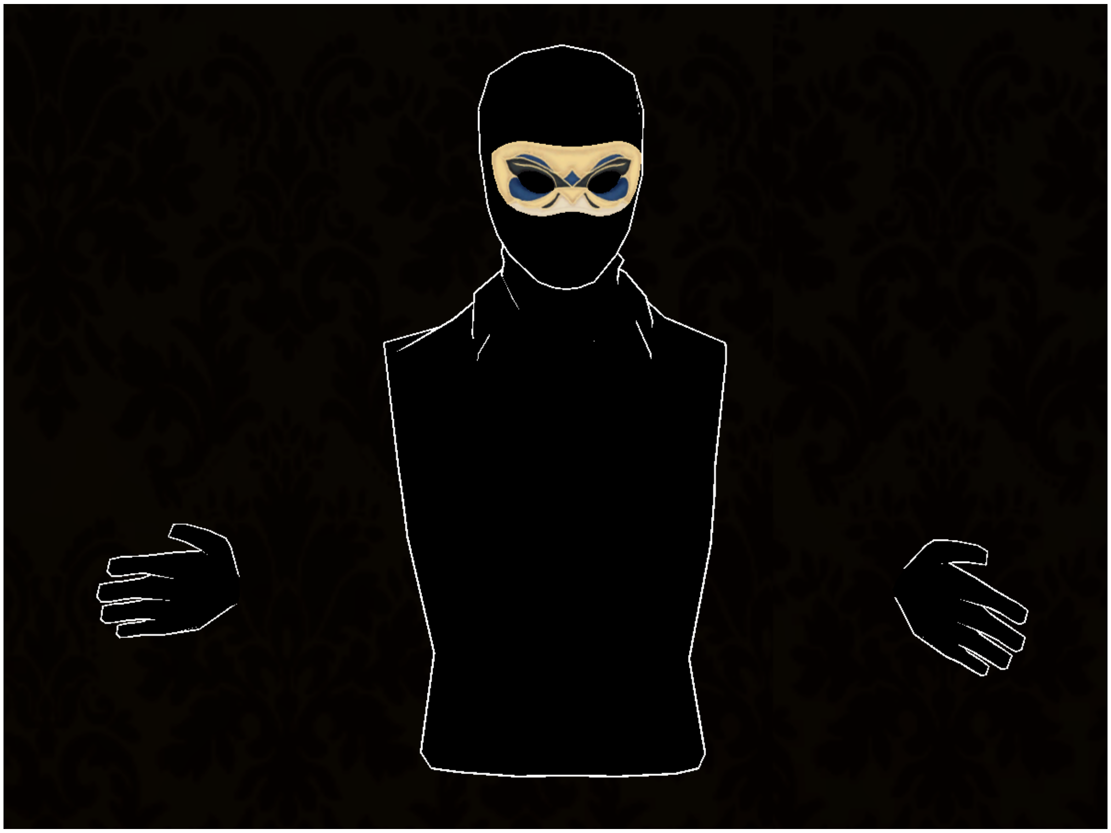

Figure 3.3: An example avatar of *MARY*.

3.4.1 Designing *MARY* for Both Mediums

Building on the design of *CAVE*, *MARY* was designed for a large audience to share an immersive experience together. However, instead of utilizing solely VR, we set out to build an experience that could also function well in the medium of AR. Each audience member in both mediums maintains a unique viewpoint, while sharing the immersive experience; however, the presentation of *MARY* occurred in fundamentally different ways in each medium. In VR, the content was presented to the audience at full scale, with them being placed directly in the room, or on the stage, of the experience. In AR,

Figure 3.4: A view of the VR version of *MARY* through the Oculus Quest.

the content was presented to the audience on top of a table, much like a dollhouse, and the audience was placed outside of the content during the experience.

In order to maintain the same content across the two viewing paradigms, we decided to limit the virtual space of *MARY* to a single stage. This allowed all of the action to be constrained and rooted to a single virtual space, allowing us to scale and the place the space in a manner that each medium best supported, without changing the content.

Unlike *CAVE*, we endeavoured to transport the viewers to multiple environments during the experience. We executed this by relying on the traditional theater technique

Figure 3.5: A view of the AR version of *MARY* through the Magic Leap One.

of a live "blackout" - dropping away the lighting on all of the environment except the characters, and shifting the environment in the darkness. As both mediums are fully digital, we were able double down by having these environment changes be physically impossible - in a matter of seconds, we could transition the environment from a fully set-dressed parlor, to a fully set-dressed laboratory.

Furthermore, we designed the full immersive space of *MARY* in as a circle, but implemented the immersive content in a traditional theatrical proscenium style. Half of the circular space of *MARY* was dedicated to the action of the narrative, with the virtual characters taking full advantage of the space, and the other half was reserved for the placement of the virtual audience (as seen in Figure 3.2. In practice, this created a sense of a staged performance for the audience, and gave them clear boundaries of where to expect action and story to occur.

3.5 View-Enablement Tools

The results of CAVRN indicated that a good portion of the audience wanted to further explore and examine the environment and characters of *CAVE*. However, many also noted that adding that level of agency and locomotion into an experience may detract or interrupt the experience for the other viewers. In lieu of this dichotomy, we built a set of digital view-enablement tools that assist in allowing a viewer to examine a virtual space and environment further without moving. These tools come in the form of i) a telescopic interface in VR, in the style of "opera glasses", ii) a low field-of-view viewing window in AR, in the style of a magnifying glass or "magnifier", and iii) a low field-of-view viewing window in VR, in the style of a viewfinder or "viewport".

A further design limitation we worked with was to not require any viewer to use and buttons in the experience. As such, the view-enablement tools were designed and implemented as totems - generally well understood objects that can be picked up and used immediately. While all of the view-enablement tools do not need to be mapped to their styles (such as the telescopic interface mapping to a pair of opera glasses), we found that providing the tools as such allowed for a majority of the viewers to use them with minimal to no training.

3.5.1 VR Telescopic Interface: The Opera Glasses

Figure 3.6: The telescopic interface tool implemented in *MARY*. The rendered style of the opera glasses was made to match that of the avatars. The third image shows a vignette rendered while the opera glasses are in use, in order to alleviate motion sickness.

The first view-enablement tool consists of a telescopic interface in VR. This interface worked directly like binoculars, magnifying the view of each eye of the virtual camera rig. When held up to the user's head, the tool directly magnified the view of the user based on their look direction. However, unlike physical binoculars, the user did not need to maintain the tool directly in front of their view in order to see through them - instead, the zoom factor of each virtual eye was applied whenever the tool was brought within a fixed distance of the head. In our implementation, we set the fixed distance to be 0.2m, and the zoom factor to be 2.5. Initial work on a custom implementation of the zoom occurred, but was quickly replaced with Oculus VR's built in zoom factor.

In the first implementation of the tool, head movement while using the tool was found to generate motion sickness with some of the testers. In order to ameliorate this, an opaque vignette was rendered on top of the view. The addition of this vignette greatly ameliorated the development of motion sickness in testers. However, it was found that the vignette obstructed a substantial portion the zoomed view. We connected the size and opacity of the vignette directly to the rotational velocity of the viewers head. In practice, this allowed for the vignette to scale and fade in when the viewer is moving their head (in moments such as looking between two characters), and then fade out as the user focuses on a subject.

In the first implementation of the tool, the user was required to hold their hand up next to their head in order to activate the zoom factor. In practice, this generated quite a bit of fatigue in the user, and prevented them from using the tool for extended periods of time. The opera glasses model was introduced as it both allowed the user to activate the tool primarily by rotating their wrist instead of holding up their arm, and as the totem was appropriate for the time period of *MARY*. Finally, the outline rendering style of the avatar was applied to the tool to denote the tool being part of the world space of the avatar, instead of the content.

3.5.2 AR Viewing Window: The Magnifier

The second view-enablement tool consists of a small 2D viewing window in AR. This interface worked much like the viewfinder of a camera, displaying whatever virtual objects are directly in front of it. In order to achieve a zooming or telescoping effect, a separate camera was used to render a new view to the plane of the viewing window. This camera was positioned along the forward vector of the center of the window - in our implementation, the camera was placed 0.25m forward. Furthermore, the field of view of the camera was reduced to be much lower than that of each virtual eye, in

Figure 3.7: The magnifier tool implemented in the AR version of *MARY*. The handle of the magnifier was removed, as the lens was positioned directly above the hand controller.

order to simulate both a higher resolution and tightened view of the viewing window. In our first implementation, we allowed testers to adjust the field of view by scrolling up and down on the touchpad of the Magic Leap One controller, but found that testers either were unable to utilize it effectively, or were thoroughly distracted by the feature. To address this, we removed the ability to adjust the field of view and set it to a fixed value of 15.

The style of a magnifier was appropriated for this tool, as it is a well understood totem that requires no input to be used effectively. It is worth noting that the design of the viewing window described above does not operate as a physical magnifier would - a physical magnifier magnifies objects that are along the ray passing along the viewing angle through the glass, whereas the described viewing window renders whats directly in front of the window. In practice, this affords the user to utilize the viewing magnifier normally when viewing a scene directly in front of them, and also "bend" a view of the scene towards them when moved and rotated.

3.5.3 VR Viewing Window: The Viewport

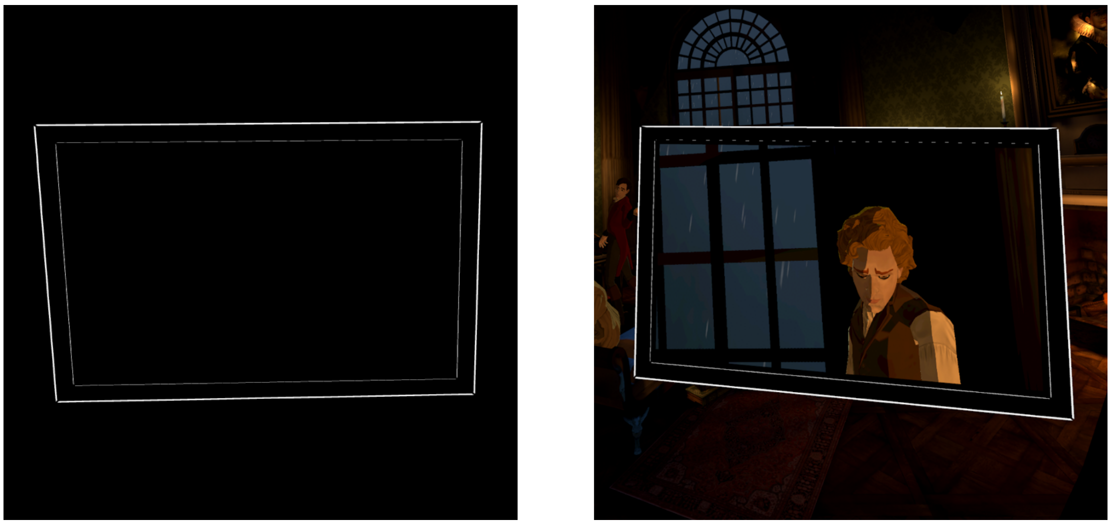

Figure 3.8: The viewing window tool implemented in the VR version of *MARY*.

The third view-enablement tool consists of a large 2D viewing window in VR. This interface functions similarly to its' AR counterpart, with a few differences. Instead of rendering a seperate view to a smaller window, the VR version was designed much larger, with a 16 by 9 aspect ratio, 0.5m on its longest side. In the first implementation of the tool, the position of the rendering camera was fixed along the forward of the window, but the natural jittery motion of a users' hand paired with the larger viewing window was reported to be sufficiently distracting for the user. Linear interpolation was applied to both the position and orientation of the rendering camera, creating the effect of the camera smoothly following the window. In early implementations, the field of view of the window could be adjusted using the thumbstick of the Oculus Touch controller, but due to the same issues found in the AR version of the tool, the capability was removed and the field of view was set to a fixed value of 15. A simple window frame style was used for this tool, and the outline rendering style of the avatar was applied tool to denote the tool being part of the world space of the avatar.

3.6 Initial Evaluation of *MARY*

Figure 3.9: An audience watching *MARY* at SIGGRAPH 2019 in Los Angeles.

The first chapter of *MARY* was presented at the SIGGRAPH 2019 conference in Los Angeles, California. Due to physical space constraints, each showing consisted of an audience of 6 AR viewers and 2 VR viewers. In this implementation, the VR avatars were removed from the AR version of the experience. Over the course of the four day conference, roughly 500 audience members experienced the first chapter. We received glowing responses from the majority of participants, with many participants returning a second time to view the experience through the other medium, but unfortunately no interviews were able to be conducted on-site due to logistic and space constraints.

In order to garner an initial understanding of the efficacy of *MARY* in VR vs. AR, and the efficacy of the proposed view-enablement tools, we presented both versions of the first chapter to XR industry and academic practitioners, in groups of two at a time (except for participant 7 8, who were chaperoned independently). In the VR version, the opera glasses were put in the left hand of the user, and the 2D viewport was placed in the right hand of the user. In the AR version, the magnifier was given to the user to hold in their dominant hand. A semi-structured interview was conducted immediately following the experience viewings.

3.6.1 METHODS

ID	*Title*
P1	VR Researcher
P2	VR PhD Student
P3	CG-ML PhD Student
P4	Animation Pipeline Engineer
P5	XR Producer
P6	VR HCI Engineer
P7	XR 3D Artist
P8	XR 3D Artist

Table 3.1: Summary of interview participants.

The participants (3.1) took part in a 15-20 minute semi-structured interview in order to facilitate discussion about both *MARY* and the investigation tools. The topics of i) the content and viewing experience of *MARY*, ii) viewing *MARY* in VR vs. AR, and iii) using the view-enablement tools were used as starting off points for conversation.

3.7 Results

The results have been organized based upon the general themes discussed in the interviews.

3.7.1 *MARY* in VR vs. AR

Both the VR and AR version of *MARY* were found to be a satisfying and engaging experience to the participants, with many of the participants noting that while both versions contained the same content, they stood apart as fundamentally different viewing experiences.

Participants consistently felt more immersion and engagement with the story during the VR viewing. Immersion was a key difference, and was noted by all of the participants.

> (P4) I thought that the VR version was definitely more immersive and I was able to connect with the story to a greater degree than the AR version.

> (P7) For the story, I really felt more immersed and that it felt more real when I was in VR. I got frightened watching the experience [in VR] ... I didn't think something on the table [in AR] could hurt me.

Participant 6 noted that even with the limitation of a seated experience, they felt greater immersion than in AR:

> (P6) You're a stationary viewer, but it feels like a very different medium in comparison than I expected it to be. I felt much more like I was in the experience [as apposed to AR].

While experiencing *MARY* in VR evoked a greater sense of immersion, experiencing *MARY* in AR evoked more of an ability to explore in many of the participants:

> (P1) I felt I had more of a need to explore the scene [in AR]. I'm less immersed, and therefore are more interested in [exploring].

Half of the participants noted without prompt that they felt the presentation of *MARY* in AR was more novel and theatrical. Participants identified that having a full view of the scene was especially engaging:

> (P5) The AR version gave me a lot more engaging experience because it was like an opera theater perspective.

> (P6) I was given this magical ability to watch something theatrical happen on it on an otherwise empty tabletop.

Furthermore, some participants noted that the AR version seemed more accessible and shareable:

> (P1) [We] could be watching this thing play out on the table, and I can be like, hey, check this part out.

> (P3) The AR version gave me the sense that we see people around us, we can invite them immediately [to join us], but in VR, someone would need to knock on my shoulder for me to [invite them].

Finally, participants noted the challenge of the limitations of current AR hardware, and understood the reasoning behind presenting *MARY* as a scaled down version on top of a table. While the AR version was well received, the sentiment of participant 4 was not unique:

> (P4) If there was a full scale [version] in AR, that would be a really cool way to view it, especially with an audience I could see.

3.7.2 VR Telescopic Interface: The Opera Glasses

The VR telescopic interface, or opera glasses, were the most consistently used and enjoyed tool of the provided view-enablement tools. Participants noted that learning and using the tool was instantaneous:

> (P4) I could quickly locate what I wanted [to look at]. At one point, I wanted to see the brain and I used them.

> (P5) If I wanted to look over somewhere, it was instantaneous [because] I'm just looking there.

The opera glasses were used by all participants to focus and expand on small facets and details, such as facial animation and handheld objects:

> (P1) I feel like your bit of a distant observer to the scene, and [when] someone starts talking ... I want to focus on them and look at their facial expressions.

Participant 8 included a nod to the implementation detail of not needing to re-position the opera glasses for them to continue using the tool. Overall, this view-enablement device was most universally accepted and used by the participants.

> (P8) I didn't find myself wanting to use the other tools as much as the opera glasses.

3.7.3 AR Viewing Window: The Magnifier

Participants found the AR viewing window, or magnifier, to have less efficacy in use than the opera glasses. Many of the participants noted the use of the magnifier for examining the environment and characters with greater detail:

> (P6) I found the magnifying glass is super helpful for me to zoom in or take a special angle on what [the characters] were doing.

> (P5) I understood the use of the magnifying glass, especially if I was in a theater and couldn't have a line of sight to a character. I could totally see using that to basically direct my own perception.

However, many participants noted that the magnifier often obstructed their view, or had difficulty aiming the tool by navigating the environment visible through its view window. Often, many participants abandoned the use of the tool until a specific high-detail moment, such as Frankenstein's handling of the creatures brain:

> (P7) I ended up leaning forward and in[to the experience], instead of using the magnifier. ... [It] was useful during the brain moment though.

One of the participants noted an unintended use of the magnifier as a view extension tool, operating like a viewfinder in the AR experience:

> (P5) I sometimes enjoyed using the magnifier to get a "perfect perspective" ... I can pick the angle I'd prefer to see. I think it would be very helpful for [students] who major in making movies.

Another noted that the tool would be more useful if it had a role in the experience, such finding clues or hidden views:

> (P4) I think it would be interesting if there was a reason to use it ... if I saw things through the magnifying glass that I wouldn't have seen otherwise, that would have made it slightly more interesting.

3.7.4 VR Viewing Window: The Viewport

While the functionality of the VR viewing window, or viewport, is comparable to that of the magnifier, the tool was used extensively in the VR version of *MARY* as a view extension tool, and was more well received by the participants than its AR counterpart. Many participants utilized the viewport as a way to keep track of other parts of the experience:

> (P1) I could focus on [the main character] without completely tuning out on the peripheral edges of [the experience]. You'd be much more likely to not lose part of the story.

> (P2) I would tilt it vertically and put it to the left so I could look at something else at the same time while the action is happening on the right.

Two participants even noted using the tool as a way of separating themselves from the space of the content:

> (P7) I could use it almost like a wall I could see through during the scary moments, but I could still see what was happening.

> (P3) I used it when I saw [Frankenstein] show up because he's too close to me, and I felt like I can hold the window in front of me so that he won't come to me. ... I felt very immersed in VR, so the viewport has a frame which separated me from the [content].

Participant 2 wished to use the viewport in two additional ways. First, they wished to be able to "pick-and-place" the viewport:

> (P2) I could want to temporarily place [the viewport] there without necessarily having to hold my hand there.

And furthermore, use it as a window to follow or lock on to a character:

> (P2) I want to track a characters perspective, without having to move around, like a target.

Overall, while having the same functionality as the AR viewing window, the VR version was both more well received by participants, and used in more novel fashions.

3.8 DISCUSSION

We now reflect on our initial findings in how the *MARY* experience was received in virtual and augmented reality, and the impact and efficacy of the view-enablement tools provided in each experience. We provide insights to inform future research and design in the HCI community.

3.8.1 SAME CONTENT, DIFFERENT MEDIUMS, UNIQUE EXPERIENCES (DC1)

We observed that while the content was not changed between the two versions outside the scale of presentation, *MARY* was well received and enjoyed as two fundamentally unique experiences by the participants.

It was consistently reported that the VR experience was more immersive, and that connecting with the story was easier in VR. Furthermore, emotional engagement such as being awed or frightened was more consistently reported in VR. However, participants found that AR was more novel, and was more accessible as a shared immersive experience. Simply being able to see the physical world while receiving the experience opened up other audience members for engagement and discussion. Others mentioned that the presentation of the AR version felt theatrical, and the ability to see the whole environment improved the experience.

The reception of the two versions as fundamentally different experiences is significant. The core content of *MARY* was not changed between the two versions, and was designed in a manner that was not excessively restrictive to the production of the experience - leaning on the design and implementation practices of the traditional medium of theater allowed for the fully digital experience to be delivered across two different and somewhat disparate mediums. This points to the potential of unlocking a great increase in distribution with little added work during implementation.

3.8.2 Different Uses of View-Enablement Tools (DC2)

While all three view-enablement tools were designed simply to increase the access and detail of the content to the viewer, the participants used each tool in vastly different ways.

The most enjoyed and utilized tool was the VR telescopic interface, or the opera glasses. Understanding and using the tool was instantaneous for all participants, and participants reported using the tool consistently throughout the experience. It was specifically used as intended, to focus and expand on the finer details of the experience such as facial animations and handheld objects. This tool required no button input from the user, and also activated when held up to the face regardless of held orientation. The quick and easy method of using this tool was greatly appreciated by the participants, and is an example of using the digital nature of the implementation to make the tool even more accessible than its physical counterpart. Furthermore, the addition of this tool did help mitigate the users wish to explore the space, as content and detail initially unavailable to the viewer was unlocked through use of the tool.

The least enjoyed and utilized tool was the AR viewing window, or the magnifier. While the tool was used to examine the environment and characters with greater detail, it often obstructed the view of the experience and was more difficult to navigate than the other provided tools. Participants found that as the experience was presented on a tabletop, they could achieve the magnification and exploration of the environment through slight physical movement in their seat, instead of utilizing the tool. Furthermore, participants wished for the tool to have some purpose in the experience outside of view-enablement. Enabling more interaction through this tool is consistent with the participants report that the AR version of the experience evoked more of a sense of exploration. Passive changes which serve to augment the narrative without interruption, such as seeing some additional parts of the environment through the tool, may increase the enjoyment of both the tool and the experience outright.

The VR viewing window, or viewport, was mostly used by participants in a manner not expected. The viewport was often used as a secondary viewing angle for tracking multiple parts of the narrative, or an attempt to separate the user from the the virtual content. Participants did report having some difficulty navigating and aiming the tool, but reported enjoyment in the tool when successfully used. This tool has the most room for expansion of the three tool, as the VR viewing window could easily be augmented to have significantly greater functionality during the experience. Zooming, placement, and character tracking could be added to the tool without minor increases its complexity, and could afford the user a sense of greater control and engagement during the experience.

In sum, the addition of view-enablement tools was found to have a positive impact on the viewing of the *MARY* experience. Incorporating intuitive digital totems can quickly power up a viewer and positively influence their engagement with and immersion in a VR or AR experience.

Conclusion

We took on the challenge of addressing i) the design and implementation space for Collective Audience Perception Spaces, ii) the design and delivery of experiences across both the VR and AR mediums, and iii) expanded the viewing capabilities of an audience member in such an experience. We built a series of components, systems, and experiments for this, and implemented two unique immersive experiences as testbeds for the design space. Overall, we found that audiences responded extremely well to our experiences, whether they came alone or in a group, whether they were in VR and AR, and whether they were brand new to the medium or an expert.

While we believe this work is a solid first exploration, we recognize a multitude of questions remain open for future work in Collective Audience Perception Spaces.

With regards to audience, we are keen to investigate the role of richer audience-audience and audience-environment interactions. We recognize that interactions affecting narrative content and subtle interactions (such as virtual actors maintaining gaze with each audience member separately) deserve focused attention. Furthermore, while we believe physical co-location between neighboring participants is vital to the fidelity of a co-located VR experience, we will investigate whether or not groupings of audience members must be physically co-located, or could be connected remotely. Telepresence between groupings of audience members may allow for larger virtual audience memberships with smaller physical installations. We are also keen to investigate asymmetric viewing experiences for each audience member, and solve technical and design details for supporting both non-linear and chaptered content.

Additionally, while we made the conscious design choice to maintain a strong positional correspondence between the physical and virtual positions of audience members, we note that changes to those distances may have logistical or narrative benefits. Still to be explored, for example, are investigations of what would be the narrative or emotional equivalent of a close-up or montage for this new theatrical medium. Modifying the virtual position of one or more audience members during an experience may also be a way to generate interesting new viewpoints for an audience, and this is a topic for future work. In a similar vein, we are curious to experiment with asymmetric physical/virtual seating topologies, such as would be achieved by rendering each user's view from an optimal virtual seat (while preserving local neighbors).

We are keen to investigate more deeply the ability to translate immersive experiences in between the various mediums that are available. Our experiments do not include the phone-based augmented reality medium, as well as the traditional mediums of both realtime games and 2D cinematics. An exciting ideal is the creation of one piece of content that can then be distributed ubiquitously across VR, AR, mobile, and standard cinematic platforms, with minimal impact on the fidelity of the experience on each platform. Furthermore, with the introduction and initial success of the view-enablement tools described, we are excited by their potential impact when ubiquitously distributed in a Collective Audience Perception Space. The capabilities of the three view-enablement tools proposed and examined could easily be augmented and expanded, or co-opted into new and more exciting tools for the viewer. We also recognize that this initial investigation of both *MARY* and the view-enablement tools served as just an initial exploration, and could be expanded upon further.

We're wildly excited by what's coming in XR. Amazing narratives and experiences are coming to life in the immersive space almost daily, and new headsets and tracking systems are being released at a breakneck pace. We are in the process of stepping into a fantastic shared immersive world, together.

References

[1] Ivan E Sutherland. A head-mounted three dimensional display. In Proceedings of the December 9-11, 1968, fall joint computer conference, part I, pages 757–764. ACM, 1968.

[2] Ed Catmull. "a computer animated hand". https://www.loc.gov/static/programs/national-film-preservation-board/documents/computer_hand2.pdf, 1972.

[3] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The cave: Audio visual experience automatic virtual environment. Commun. ACM, 35(6):64–72, June 1992.

[4] C. Carlsson and O. Hagsand. Dive a multi-user virtual reality system. In Proceedings of IEEE Virtual Reality Annual International Symposium, pages 394–400, Sept 1993.

[5] Joseph Psotka and Sharon Davison. Cognitive factors associated with immersion in virtual environments. NASA. Johnson Space Center, Proceedings of the 1993 Conference on Intelligent Computer-Aided Training and Virtual Environment Technology, 1993.

[6] Jon McKenzie. Virtual reality: Performance, immersion, and the thaw. TDR (1988-), 38(4):83–106, 1994.

[7] Jason Leigh, Andrew Johnson, and Thomas A DeFanti. Cavern: A distributed architecture for supporting scalable persistence and interoperability in collaborative virtual environments. Virtual Reality: Research, Development and Applications, 2(2):217–237, 1997.

[8] E. F. Churchill and D. Snowdon. Collaborative virtual environments: An introductory review of issues and systems. Virtual Reality, 3(1):3–15, Mar 1998.

[9] Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. Computer Supported Cooperative Work (CSCW), 11(3):411–446, Sep 2002.

[10] Kevin Brooks. There is nothing virtual about immersion: Narrative immersion for vr and other interfaces. alumni. media. mit. edu/˜ brooks/storybiz/immersiveNotVirtual. pdf (accessed May 2007), 2003.

[11] Steve Dixon. A history of virtual reality in performance. International Journal of Performance Arts and Digital Media, 2(1):23–54, 2006.

[12] Valentijn T. Visch, Ed S. Tan, and Dylan Molenaar. The emotional and cognitive effect of immersion in film viewing. Cognition and Emotion, 24(8):1439–1445, 2010.

[13] Thomas A. DeFanti, Daniel Acevedo, Richard A. Ainsworth, Maxine D. Brown, Steven Cutchin, Gregory Dawe, Kai-Uwe Doerr, Andrew Johnson, Chris Knox, Robert Kooima, Falko Kuester, Jason Leigh, Lance Long, Peter Otto, Vid Petrovic, Kevin Ponto, Andrew Prudhomme, Ramesh Rao, Luc Renambot, Daniel J. Sandin, Jurgen P. Schulze, Larry Smarr, Madhu Srinivasan, Philip Weber, and Gregory Wickham. The future of the cave. Central European Journal of Engineering, 1(1):16–37, Mar 2011.

[14] Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, J. D. Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, Dan Sandin, Luc Renambot, Andrew Johnson, and Jason Leigh. CAVE2: a hybrid reality environment for immersive simulation and information analysis. Proc.SPIE, 8649:8649 – 8649 – 12, 2013.

[15] Erica J. Boothby, Margaret S. Clark, and John A. Bargh. Shared experiences are amplified. volume 25, pages 2209–2216, 2014. PMID: 25274583.

[16] André Kunert, Alexander Kulik, Stephan Beck, and Bernd Froehlich. Photoportals: Shared references in space and time. In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW ’14, pages 1388–1399, New York, NY, USA, 2014. ACM.

[17] Daniel Leithinger, Sean Follmer, Alex Olwal, and Hiroshi Ishii. Physical telepresence: Shape capture and display for embodied, computer-mediated remote collaboration. In Proceedings of the 27th Annual ACM Symposium on

User Interface Software and Technology, UIST '14, pages 461–470, New York, NY, USA, 2014. ACM.

[18] Jerald Thomas, Raghav Bashyal, Samantha Goldstein, and Evan Suma. Muvr: A multi-user virtual reality platform. In Virtual Reality (VR), 2014 iEEE, pages 115–116. IEEE, 2014.

[19] Simon McBurney. "the encounter", Aug 2015.

[20] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. Virtual replicas for remote assistance in virtual and augmented reality. In Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology, UIST '15, pages 405–415, New York, NY, USA, 2015. ACM.

[21] Cassidy Curtis, David Eisenmann, Rachid El Guerrab, and Scot Stafford. The Making of Pearl, a 360° Google Spotlight Story. In ACM SIGGRAPH 2016 VR Village, SIGGRAPH '16, pages 21:1–21:1, New York, NY, USA, 2016. ACM.

[22] Dragos Datcu, Stephan Lukosch, and Heide Lukosch. A collaborative game to study presence and situational awareness in a physical and an augmented reality environment. j-jucs, 22(12):247–270, feb 2016.

[23] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. Holoportation: Virtual 3d teleportation in real-time. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16, pages 741–754, New York, NY, USA, 2016. ACM.

[24] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. Room2room: Enabling life-size telepresence in a projected augmented reality environment. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, CSCW '16, pages 1716–1725, New York, NY, USA, 2016. ACM.

[25] The Void. The Void: The Vision Of Infinite Dimensions. https://thevoid.com/, 2016.

[26] CJ CGV. Screenx, 2017. Online; accessed 4-April-2019.

[27] Mingsong Dou, Philip Davidson, Sean Ryan Fanello, Sameh Khamis, Adarsh Kowdle, Christoph Rhemann, Vladimir Tankovich, and Shahram Izadi. Motion2fusion: Real-time volumetric performance capture. ACM Trans. Graph., 36(6):246:1–246:16, November 2017.

[28] Fairhurst G. Shepherd G. Eggert, L. Udp usage guidelines, 2017.

[29] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, pages 4021–4033, New York, NY, USA, 2017. ACM.

[30] Alejandro G. Iñárritu. Carne y arena, 2017.

[31] Motion Picture Association of America. 2016 theatrical market statistics. Technical report, 2017.

[32] Devon Penney and Eugene Chung. Arden's Wake. In ACM SIGGRAPH 2017 Computer Animation Festival, SIGGRAPH '17, pages 28–28, New York, NY, USA, 2017. ACM.

[33] Samsung. Samsung galaxy s8 and s8+ - the official samsung galaxy site, 2017. Online; accessed 20-September-2018.

[34] Ruofei Du, Ming Chuang, Wayne Chang, Hugues Hoppe, and Amitabh Varshney. Montage4d: Interactive seamless fusion of multiview video textures. In Proceedings of ACM Interactive 3D Graphics (I3D) 2018, 2018.

[35] LLC Facebook Technologies. Oculus venues | oculus, 2018. Online; accessed 4-April-2019.

[36] David Gochfeld, Corinne Brenner, Kris Layng, Sebastian Herscher, Connor DeFanti, Marta Olko, David Shinn, Stephanie Riggs, Clara Fernández-Vara, and Ken Perlin. Holojam in wonderland: Immersive mixed reality theater. In ACM SIGGRAPH 2018 Art Gallery, SIGGRAPH '18, pages 362–367, New York, NY, USA, 2018. ACM.

[37] Google. Daydream – standalone vr, 2018. Online; accessed 20-September-2018.

[38] Naito K. Liu L. Gu, B. The apple, 2018. Raindance Film Festival.

[39] Beth Karlin, Hyun Tae Calvin Kim, Ruth Kelly, Johanna Blakley, Corinne Brenner, and Patricia Riley. Does medium matter? Technical report, The Norman Lear Center, University of Southern California, 2018.

[40] Microsoft. Altspacevr inc | be there, together, 2018. Online; accessed 4-April-2019.

[41] Motion Picture Association of America. 2017 theme report. Technical report, 2018.

[42] VR OWL. Vr sync | synchronized 360 vr for oculus go, gearvr, and android, 2018. Online; accessed 4-April-2019.

[43] Misha Sra, Aske Mottelson, and Pattie Maes. Your place and mine: Designing a shared vr experience for remotely located users. In Proceedings of the 2018 Designing Interactive Systems Conference, DIS '18, pages 85–97, New York, NY, USA, 2018. ACM.

[44] Unity Technologies. Unity. =https://unity3d.com/, 2018.

[45] The VOID. Step beyond reality, 2018. Online; accessed 19-September-2018.

[46] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. UIST, 2018.

[47] IMAX Corporation. Imax | the world's most immersive movie experience, 2019. Online; accessed 4-April-2019.

[48] Dreamscape Immersive. Dreamscape. A VR experience like no other. https://dreamscapeimmersive.com/, 2019.

[49] Facebook Technologies, LLC. Oculus Rift: VR Headset for VR Ready PCs | Oculus. https://www.oculus.com/rift/, 2019.

[50] Google. Tiltbrush by google. "https://www.tiltbrush.com/, 2019.

[51] Hologate. HOLOGATE - VR Virtual Reality + XR Extended Reality Solutions. http://hologate.com/, 2019.

[52] Hyperbolic Magnetism. Beat Saber. http://beatsaber.com/, 2019.

[53] Magic Leap, Inc. Home | Magic Leap. https://www.magicleap.com/, 2019.

[54] Magic Leap, Inc. PCFs | Magic Leap. https://developer.magicleap.com/learn/guides/persistent-coordinate-frames, 2019.

[55] Microsoft. Microsoft hololens | mixed reality technology for business. https://www.microsoft.com/en-us/hololens, 2019.

[56] VRChat Inc. Vrchat. https://www.vrchat.net/, 2019.

[57] Janet Horowitz Murray. Hamlet on the Holodeck: The Future of Narrative in Cyberspace. The Free Press, New York, NY, USA, 1997.

[58] Bohne Silber and Tim Triplett. A decade of arts engagement: findings from the Survey of Public Participation in the Arts, 2002-2012. National endowment for the arts, 2015.