CAVE: Making Collective Virtual Narrative

Best Paper Award

Kris Layng, Ken Perlin, Sebastian Herscher, Corinne Brenner and Thomas Meduri

ABSTRACT

CAVE is a shared narrative six degrees of freedom (6DoF) virtual reality experience. In 3.5 days, 1,927 people attended its premiere at SIGGRAPH 2018. Thirty participants at a time each saw and heard the same narrative from their own individual location in the room, as they would when attending live theater. *CAVE* set out to disruptively change how audiences collectively experience immersive art and entertainment. Inspired by the social gatherings of theater and cinema, *CAVE* resonated with viewers in powerful and meaningful ways. Its specific pairing of colocated audiences and physically shared immersive narrative suggests a possible future path for shared cinematic experiences.

Kris Layng
Cocreator/Director
Future Reality Lab
Courant Institute Department of Computer Science
New York University
60 5th Ave, 3rd Floor
New York, NY 10011 U.S.A.
kel305@nyu.edu

Ken Perlin
Cocreator/Executive Producer
Future Reality Lab
Courant Institute Department of Computer Science
New York University
60 5th Ave, 3rd Floor
New York, NY 10011 U.S.A.
perlin@nyu.edu

Sebastian Herscher
Cocreator/Technical Director
Future Reality Lab
Courant Institute Department of Computer Science
New York University
60 5th Ave, 3rd Floor
New York, NY 10011 U.S.A.
herscher@nyu.edu

Corinne Brenner
Researcher
New York University
60 5th Ave, 3rd Floor
New York, NY 10011 U.S.A.
cjb399@nyu.edu

Thomas Meduri
Lead Technical Artist
Future Reality Lab
Courant Institute Department of Computer Science
New York University
60 5th Ave, 3rd Floor
New York, NY 10011 U.S.A.
medurithomas@gmail.com

Most virtual reality (VR) experiences today are delivered to one person at a time, creating an "isolating and lonely" experience that is inaccessible to the public [1,2]. Even critically acclaimed works like *Henry* [3], *Dispatch* [4] and *Allumette* [5] are designed and produced for an audience of one. The result is long lines at events, installations and specialized arcades, which breeds frustration and keeps VR out of the mainstream.

In contrast, *CAVE* was devised from the ground up as a way for dozens of people to share immersive content simultaneously (Fig. 1). *CAVE* uses the capability of VR to create the illusion, for an entire audience, that they are experiencing a live theater event together, even though the content is actually prerecorded. This creates an experience that is fundamentally different from VR for an individual viewer. Each audience member both sees and hears the story content from their own viewpoint in the room, as they would when attending traditional theater. In the shared virtual world, audience members also see each other as avatars, whose movements correspond to their own head movements.

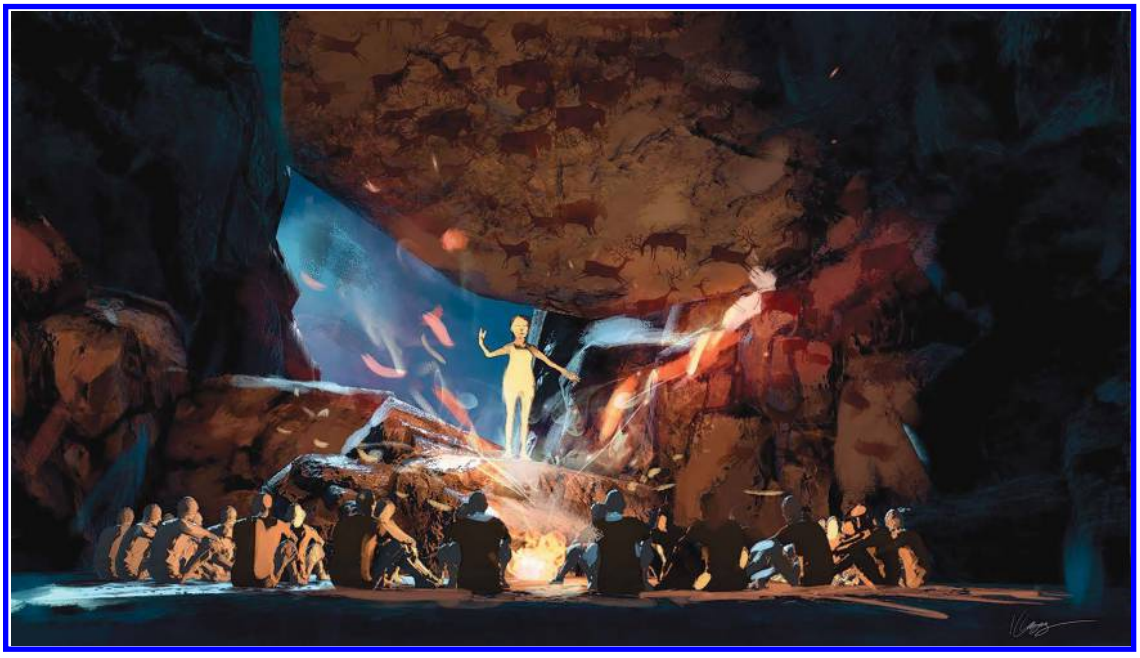

Fig. 1. Initial concept illustration for *CAVE* group experience. (© and Art: Kris Layng, 2018)

Many VR experiences for multiple colocated users are more game-like and "lean forward": Each user takes on an active role, responding to and perhaps altering the content of the experience. For example, *Life of Us* [6] is intended for up to four individuals to move together through a series of evolutionary transformations; experiences such as The VOID's *Ghostbusters Dimension*, *Star Wars: Secrets of the Empire* and *Ralph Breaks VR* allow groups of up to four individuals to pursue narrative-driven "missions," walking through a specially designed environment that tracks their movements [7]. Users share a common goal, although each user's focus of attention and specific actions may differ.

Holojam in Wonderland [8], a previous project by members of our team, was a multiuser, immersive narrative for six people at a time, using live actors in motion capture suits acting out a scene from *Alice in Wonderland*. However, due to the limitations of using live performers for each show and allowing participants to walk around the performance space, it was not scalable. The lack of scalability of *Holojam in Wonderland* motivated us to create *CAVE*.

CAVE represents multiple colocated users in a passive "lean back" virtual experience. As in traditional theater, audience members remain seated, each experiencing the same narrative but from their different physical locations in the theater (Fig. 2). Individual exploration is not the intention; audiences are simply asked to sit back and enjoy the show. We believe sharing a narrative in a virtual space while simultaneously sharing a physical space amplifies the emotional impact of the story and creates a new medium for collective experience. The combined effects of the story, design, sound and shared audience presence result in a collective experience similar to attending live theater, but with the ability to create stunning effects far beyond the physical limitations of production in a theater.

Story, Script, Set and Action

We set out to create a disruption in the way that people physically gather to experience immersive art and entertainment. We chose the subject matter—a fantasy story about cave paintings—to explore the features and consequences of physically gathering to experience art from many creative perspectives. Our concept began with putting audience members around a campfire—the place where storytelling originated (Fig. 3). From here, we imagined a cave setting covered in wondrous animal paintings, like those at

Fig. 2. *CAVE* at the ACM/SIGGRAPH 2018 conference. (Photo: Eric Chang, 2018)

Fig. 3. *CAVE* concept art. (© and Art: Kris Layng, 2018)

Lascaux and Chauvet in modern-day France. Under the flickering glow of the firelight, it's speculated that for ancient cave dwellers these paintings gave the illusion of animation and movement, creating a hallucinatory, movie-like experience. In some sense, it was the origin story of virtual reality.

Before any production began, we spent several months crafting the story and script. Once we made our content choice, our decisions were inspired by anthropological research to understand not just the appearance and clothing and tools of the late Paleolithic age, but also the mindset that a storyteller might have had in that era. Having made that choice, we realized we needed to make it clear that our story was situated, specifically, within the nomadic tribes of Northern Europe 12,000 years ago to avoid any inadvertent identification with current tribal and indigenous cultures.

As our story begins, a young woman, Ayara, climbs down from an opening in the cave above and lights a fire. In soliloquy, she explains that her community is suffering from a famine, and she seeks guidance from her unseen spirit ancestors. While her mother, a shaman, would have known what to do, Ayara must demonstrate her own magical abilities by painting her own images. As a result of her actions, a mammoth emerges from the back of the cave and transforms into the spirit of her mother. The cave paintings become animated, and Ayara and her mother embrace. Ayara then realizes she also has the power to channel the spirit world and save her tribe.

We wrote all dialogue in iambic pentameter to convey a sense of ritual speech for modern audiences. The intent was to set a tone of mythological significance, to suggest that the events taking place are larger than those that occur in everyday life. Sitting around the perimeter of the cave are the ephemeral ancestral spirits of past shamans (embodied by the audience members) (Fig. 4). The choice of audience representation as "spirit ancestors" was made so that audience members feel they are part of this world and part of the action, yet nothing is required of them. Their only role as spirit ancestors is to bear witness.

The narrative contains pauses and quiet moments that build in intensity as Ayara ritualistically prepares her paint, leading to the final climactic spectacle of the mammoth appearing and transforming into the spirit of her mother. We chose to keep the "stage magic" to a minimum until the climax to enhance the feeling of live theater. In a medium such as VR, where creative possibilities are overwhelming, our restraint was in

Fig. 4. Audience member avatars (spirit ancestors). (© and Art: Kris Layng, 2018)

service of keeping the audience members grounded in a believable physical reality. This way, when magic is introduced at the climax of the piece, the impact is more surprising and intense.

Elements Contributing to the Audience Experience

Costume Design

We brought on costume designer and researcher Kat Jeffery to advise us on the authenticity of the clothing from the place and time period of our story. She worked with images from nineteenth-century anthropologists who excavated and researched clothing from Russian and Siberian shamanic peoples who had been living in isolated tribal cultures. This was the closest primary research she could find to make informed choices about the clothing styles of people living in that region during the late Paleolithic period. After the director approved her designs, Jeffery worked closely with the animation team to translate the qualities of fabrics, furs and costume construction into an animated medium at the appropriate level of detail for the low-polygon aesthetic we chose (Fig. 5).

Fig. 5. Ayara costume design rendering. (© and Art: Kat Jeffery, 2018)

Each audience member, represented as a spirit ancestor, wears a ceremonial headpiece/mask made of the skull of an animal of the time period and region, such as a mountain lion or saber-toothed tiger. This mask serves two purposes: First, it acts as a nod to VR headsets and provides a visual way to link the virtual world to the real world. Second, it gives audience members a more pronounced head profile. This accentuates movement and makes other audience members' head movement and direction of gaze more clear. In a collective group experience such as *CAVE*, where it is vital to maintain synergy between audience members, it is essential to see where each person is looking.

Set Design

The virtual set of the cave is roughly dome-shaped, with a raised area toward the center. There is an opening at the top through which the audience can see the night sky. The rock walls are adorned with playfully stylized animal paintings (Fig. 6).

Fig. 6. *CAVE* set design. (© and Art: Kris Layng, 2018)

The director's background in theatrical and film set design provided a wealth of expertise for how to treat a physical space, and heavily informed qualities of the animation and digital design. The dimensions of the cave set were partially inspired by classical architecture, especially the single, round, overhead opening in the dome of the Pantheon, again to evoke a sense of significance and majesty.

Inside the cave, the audience is seated around a thrust stage, surrounding both sides of the performance space. We chose this arrangement, rather than a more traditional proscenium stage, to create intimacy between the audience and Ayara's performance. The performance was blocked as in traditional theater, using different "zones" of action around the space, so as to provide opportunities for all audience members to be close to the action. Ayara, the main character, moves through all three dimensions of the virtual environment, initially descending from above, then moving throughout the space as the story progresses. We thought it was essential that all audience members have a chance to be close to the character, especially since the low resolution of the headset display made it difficult to see her more subtle facial expressions from far away.

The virtual space deviated somewhat from the physical space, where such variations would better support the experience. For example, a central aisle between the two audience groups was about twice as wide in the virtual world as in real life, to allow the mammoth to pass through at the climactic moment (Fig. 7). The space between rows of audience seats was also wider and made to appear slightly raked or sloped upward in the virtual space to prevent obstructed views of the content.

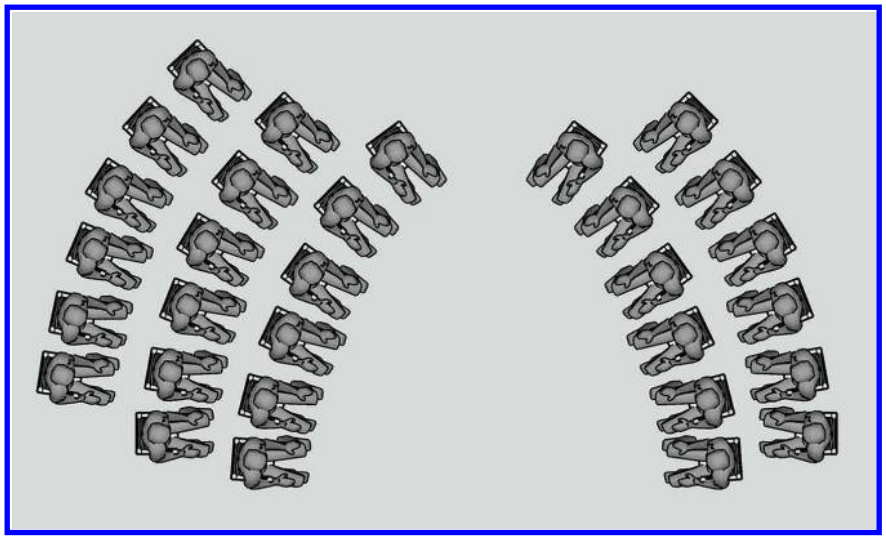

Fig. 7. Top view of physical audience layout, with narrow aisle. (Diagram: © Kris Layng, 2018)

We were concerned that the discrepancy between physical and virtual spaces would bother participants, but the audience seemed content to suspend their disbelief for these deviations from the physical environment. Indeed, these choices seemed to work as intended, providing a better vantage point for each audience member than a literal mapping would have allowed.

Sound Design and Music

Sound design and music were vital to creating a visceral sense of space by continually providing audience members with cues about both the physical environment of the cave and the emotional arc of the story. Elements of the sonic environment included ambient sound, as well as sounds associated with specific characters and objects. Sound for each audience member was created by combining off-ear audio headsets and speakers in the room, including selective use of a subwoofer. Each audience member heard an individualized synthesized soundfield from their own headset that corresponded to their position and head movements. This was supplemented by the ambient speakers and subwoofer, which added a background soundfield that everyone shared. For example, during the mammoth's entrance, the sound of massive footsteps rendered from the subwoofer intensified the climactic moment.

Ryan Shore, a composer with years of experience scoring Hollywood films, composed the original music score. Although we explored going a less conservative route by using experimental atmospheric music that would be diegetic to Ayara's world, we decided on a beautiful but conventional score that would help guide the emotional arc of the story and pull at the audience's heartstrings. The theme of the piece was based on an ancient Scandinavian lullaby that the director discovered online.

We decided against using over-ear headphones that would further isolate each participant and block communication between people. We used prototype off-ear headphones that Bose Corporation designed and produced specifically for *CAVE*, which allowed for an unobtrusive and comfortable experience for audience members. Because the Bose headphones projected a collimated beam of sound into each ear, audience members could still hear environmental sounds and communicate with each other (even when whispering), thereby intensifying the collective nature of the experience. Again, this helped cultivate a positive shared group experience.

Onboarding

The *CAVE* experience begins even before the audience steps into the virtual world, right as they enter the venue. Audience members walk into a room that is dimly lit with warm colored uplights and rows of seats, suggestive of a movie theater. The room arrangement and lighting were carefully designed to feel familiar and put the group at ease, as well as establish that this would be a passive experience.

After audience members take their seats and put on their headsets, they find themselves sitting together in a dark cave at sunset, where a mother is painting with her young daughter (Ayara as a child) while singing a beautiful, haunting melody. This onboarding builds anticipation and mystery, gradually inviting the audience into the dreamlike world of our story.

Audience Presence

An essential aspect to the *CAVE* experience is that audience members feel as though they are in a magical world together with other people, feeling the presence of those around them. The Google Lenovo Mirage Solo headset was used for this event because it was the only wireless headset on the market with 6DoF inside out tracking (Fig. 8). 6DoF was vital because we wanted the audience to feel present in the virtual world together with the characters and with each other. We feel that 6DoF is essential to a true feeling of shared physical copresence [9], because when audience members in 6DoF move their heads from side to side or lean forward and back, they experience true motion parallax.

To test the effect of physical presence in a 6DoF experience, we created an early prototype of a generic character reciting a Shakespeare soliloquy in Polish while walking around a virtual set. The words and the movements were entirely unrelated, and most of us did not know what the character was saying (since we don't speak Polish). Yet our sense of presence, our feeling of being together in a transformed space, while watching this scene was powerful and compelling. At that point, we knew our approach would work.

Fig. 8. Audience emotional reactions—participant gasps during mammoth entrance. (© and Photo: Eric Chang, 2018)

Research has shown that experiencing media collectively has distinct positive effects on audiences. For example, people enjoy humorous material more when they are together, whether in person or virtually [10,11]. When we showed *CAVE* at ACM/SIGGRAPH 2018, the majority of participants responding to a survey reported that the colocated experience was effective and enjoyable. 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, $M = 5.24$, $SD = 1.58$).

Adapting Tools and Workflows for a New Medium

CAVRN: Colocated Audience Virtual Reality Nexus

To create the collective audience experience, our team developed the CAVRN system, a lightweight software architecture that synchronized the experience for all viewers. CAVRN ensures time synchronization between all client headsets (so all audience members see and hear things at exactly the same time), as well as the ability to transmit the head position and orientation of each audience member to all other headsets (so each audience member can see all other audience members as avatars) (Fig. 9). The CAVRN system can be operated from a single smartphone or tablet. The content on the headsets themselves was rendered using the Unity game engine. We used the game networking service Photon to send signals between all devices.

Show Me in VR: Effective Use of Tools for Set Design in VR

The director primarily used the VR tools Google Blocks and Tilt Brush to visualize the set. With these tools, the director was able to interactively shift his point of view between (1) experiencing the set as a miniature and (2) experiencing it full size from the point of view of audience members. This dual approach to designing the virtual set turned out to be very productive and easy to use.

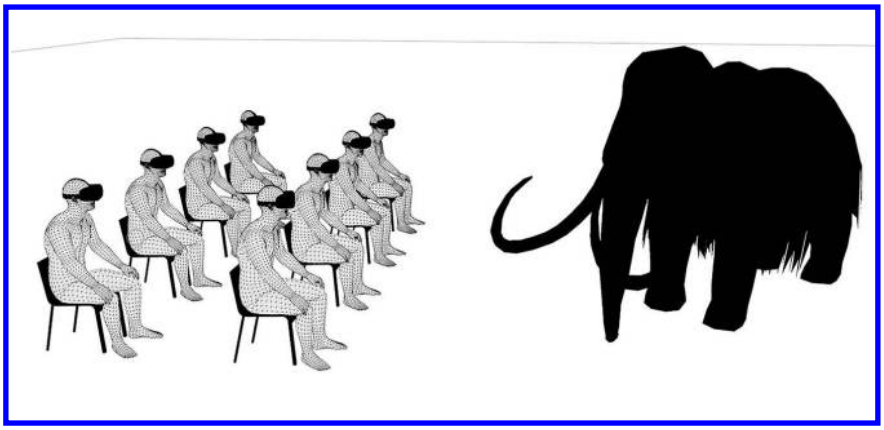

Fig. 9. With the CAVRN system, multitudes of people can see each other and the virtual world simultaneously. (Diagram: © Kris Layng, 2018)

The rallying cry for all design work was "show me in VR." Collaboration in VR was effective: Members of the team noted how easy it was

to understand the director's goals and share ideas, even when it was difficult to express those ideas verbally, since all were looking at the same scene playing in real time.

Conclusion

In 1895, the film industry reinvented itself by switching from kinetoscope to projection, thereby making cinema a truly collective experience for the masses. Similarly, *CAVE* demonstrates the promise of a new medium, defined by the transition of VR storytelling from an individual to a collective experience.

This new medium retains much of the collective audience experience of live theater, together with the imaginative potential and mass scalability of cinema (since the content is all prerendered and prerecorded). *CAVE* shows that narrative virtual reality is best experienced together with other people, where fellow audience members amplify emotion. When audiences share physical proximity, the ability to defy reality in immersive worlds with fantastic characters make moments of magic all the more impactful.

People being together was essential to the success of the piece. It has not escaped our notice that the specific pairing of colocated audiences and physically shared immersive virtual narrative immediately suggests a possible future path for the motion picture industry.

Acknowledgments

CAVE premiered at ACM/SIGGRAPH 2018 on August 12–16 in Vancouver, Canada. The production was created by a large team of people and supporters besides the authors: Jess Bass, Hanako Greensmith, MaYaa Boateng, Marta Gospodarek, Tatiana Turin, Dennis Dembeck, Kat Jeffery, Ryan Shore, Eric Chang, Paloma Dawkins, Mia Mamede, Susan Davishi, Marcus Z. Guimaraes, Pasan Dharmasena, Cynthia Allen, Lily Cushman, Connor DeFanti, Rowan Smith, Angela Yu, Dan Lazar, Brandon Buikema, Bose, Google Daydream, Lenovo, Future Tech and NYU Future Reality Lab.

References and Notes

1. J. Koetsier, "VR Needs More Social: 77% of Virtual Reality Users Want More Social Engagement," *Forbes*, 30 April 2018. Retrieved from www.forbes.com/sites/johnkoetsier/2018/04/30/virtual-reality-77-of-vr-users-want-more-social-engagement-67-use-weekly-28-use-daily/#3e0204bf18fc, accessed 3 May 2019.
2. R. Pringle, "Virtual reality is still too isolating to be 'the next big thing' in tech," *CBC News*, 14 February 2017. Retrieved from www.cbc.ca/news/opinion/vr-isolation-1.3980539, accessed 3 May 2019.
3. R. Lopez Dau (Director), *Henry* [Virtual Reality Experience], 28 July 2015. U.S.A.: Oculus Story Studio.
4. E. Robles (Developer), *Dispatch* [Virtual Reality Experience], 21 November 2017. U.S.A.: Here Be Dragons.
5. E. Chung (Director), *Allumette* [Virtual Reality Experience], 2016. U.S.A.: Penrose Studios.
6. C. Park et al. (Producers), C. Milk and A. Koblin (Directors), *Life of Us* [Virtual Reality Experience], 2017. U.S.A.: Annapurna Pictures, Unity.
7. S. Rogers, "Wreck-It Ralph Smashes into the VOID," *Forbes*, 16 November 2018. Retrieved from www.forbes.com/sites/solrogers/2018/11/16/wreck-it-ralph-smashes-into-the-void/#183e5708503d, accessed 3 May 2019.
8. D. Gochfeld et al., "*Holojam in Wonderland*: Immersive Mixed Reality Theater," *Leonardo* **51**, No. 4, 362–367 (2018).
9. W.A. IJsselsteijn and G. Riva, "Being There: The Experience of Presence in Mediated Environments," in *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments*, G. Riva, F. Davide and W.A. IJsselsteijn, eds. (Amsterdam: Ios Press, 2003).
10. J. Hofmann et al., "The influence of a virtual companion on amusement when watching funny films," *Motivation and Emotion* **39**, No. 3, 434–447 (2015).
11. D.J. Wimer and B.C. Beins, "Expectations and perceived humor," *Humor: International Journal of Humor Research* **21** (2008) pp. 347–363. doi: 10.1515/HUMOR.2008.

This article has been cited by:

1. John Desnoyers-Stewart, Ekaterina R. Stepanova, Bernhard E. Riecke, Patrick Pennefather. 2020. Body RemiXer: Extending Bodies to Stimulate Social Connection in an Immersive Installation. *Leonardo* **53**:4, 394-400. [Abstract] [PDF] [PDF Plus] [Supplemental Material]