COMPUTATIONAL IMMERSIVE DISPLAYS

Daniel E. Novy, BFA, MA

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology
June 2013
© 2013 Massachusetts Institute of Technology. All Rights Reserved.

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUL 19 2013
LIBRARIES

Author

Daniel E. Novy
Program in Media Arts and Sciences

Certified By

Dr. V. Michael Bove, Jr.
Principal Research Scientist
Media Lab, Massachusetts Institute of Technology

Accepted By

Prof. Patricia Maes
Associate Academic Head
Media Lab, Massachusetts Institute of Technology

COMPUTATIONAL IMMERSIVE DISPLAYS

Daniel E. Novy, BFA, MA

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology
June 2013
© 2013 Massachusetts Institute of Technology. All Rights Reserved.

Abstract

Immersion is an oft-quoted but ill-defined term used to describe a viewer or participant's sense of engagement with a visual display system or participatory media. Traditionally, advances in immersive quality came at the high price of ever-escalating hardware requirements and computational budgets. But what if one could increase a participant's sense of immersion, instead, by taking advantage of perceptual cues, neuroprocessing, and emotional engagement while adding only a small, yet distinctly targeted, set of advancements to the display hardware? This thesis describes three systems that introduce small amounts of computation to the visual display of information in order to increase the viewer's sense of immersion and participation. It also describes the types of content used to evaluate the systems, as well as the results and conclusions gained from small user studies.

The first system, Infinity-by-Nine, takes advantage of the dropoff in peripheral visual acuity to surround the viewer with an extended lightfield generated in realtime from existing video content. The system analyzes an input video stream and outpaints a low-resolution, pattern-matched lightfield that simulates a fully immersive environment in a computationally efficient way.

The second system, the Narratarium, is a context-aware projector that applies pattern recognition and natural language processing to an input such as an audio stream or electronic text to generate images, colors, and textures appropriate to the narrative or emotional content. The system outputs interactive illustrations and audio projected into spaces such as children's rooms, retail settings, or entertainment venues.

The final system, the 3D Telepresence Chair, combines a 19th-century stage illusion known as Pepper's Ghost with an array of micro projectors and a holographic diffuser to create an autostereoscopic representation of a remote subject with full horizontal parallax. The 3D Telepresence Chair is a portable, self-contained apparatus meant to enhance the experience of teleconferencing.

Thesis Supervisor: V. Michael Bove, Jr.

Title: Principal Research Scientist, MIT Media Lab

COMPUTATIONAL IMMERSIVE DISPLAYS

Daniel E. Novy, BFA, MA

The following people served as readers for this thesis:

Thesis Reader:

Dr. V. Michael Bove, Jr.
Principal Research Scientist
Media Lab, Massachusetts Institute of Technology

Thesis Reader:

Pattie Maes
Alexander W. Dreyfoos Professor
Associate Academic Head
Media Lab, Massachusetts Institute of Technology

Thesis Reader:

Glorianna Davenport
Visiting Scientist
Media Lab, Massachusetts Institute of Technology

ACKNOWLEDGMENTS

Michael, thanks for the freedom to explore and the trust that I'd come back with something useful. Thanks for helping me navigate time and space when projects got popular. Thank you most of all for your belief in me even though I hadn't looked at Calculus in 20 years.

My readers, thank you for such sustaining encouragement from the start. Your challenging questions and gentle guidance have made these projects, this thesis, and myself as a student and researcher better than they ever could've been without you.

To my lovely, patient, and beatific wife Chia. Thank you for letting me uproot our lives from a successful career and travel into the unknown just as most couples are settling down to live out their years in comfort. None of this would have been possible nor successful without you.

My daughter Connor Kelly Novy, who has to put up with being in college at the same time as her old man and seems to be handling it much better.

Joi, for stressing the organism to make it grow.

OBMGíers, storytellers and wizards of light!

Neil G. and Joe P.

Linda Peterson, William Tilden, Kristin Hall, and Skylar Harvey-Forsythe, the therapeutic wonder dog.

Tom Lutz, John DiFrancesco, and Martin Seymour for generous and expert guidance in the shop.

Mom, Dad, and my Father-in-Law, Dr. Jack Evers.

The OpenFrameworks community.

Table of Contents

1. Introduction

"What the eye sees and ear hears, the mind believes."
—Harry Houdini

1.1 Why less is more.

Most traditional display technologies are just that—technologies for displaying recorded information in a two-dimensional, presentational format. The first Paleolithic artist placed his or her images on a flat wall because that was the technology available at the time, and the paradigm hasn't improved much since. From entertainment media, such as film and television, to business media, including specialized teleconferencing systems and Skype, the de facto interface for receiving and viewing information and interactions is still a vertical display, set perpendicular to the viewer's gaze and parallel to his or her body, which defines a set field of view. The tyranny of the screen locked its dominance in early and has held on tight since.

Recently, there have been a variety of efforts to increase the immersive quality of visual displays by manipulating the size of the display, as with large-format televisions and IMAX screens, or by simulating three-dimensional environments through the use of stereoscopic projection. While these systems attempt to increase a viewer's feeling of being present in the space depicted by the media, most lack the ability to adapt dynamically to the display environment and/or the narrative content. These forms are still guilty of presenting merely a bounded frustum's worth of information terminating in the lens of the eye, with no fundamental advances to the form or new affordances presented with the content. A song played louder is still the same song; scale does not equal innovation.

Immersion does not require increased screen real estate; it requires additional ways of deepening the engagement and attention of the user. Careful attention must also be paid to the way in which human beings perceive, process, and react to signals in the environment. There is much that we do not see or hear simply because we lack the biological hardware to sense energies above or below certain critical thresholds. Even within the spectrum of frequencies we can detect, rates of neural processing affect how much or how quickly we can make sense of what we perceive.

Added to this system of signal sensing and processing is the added layer of socially significant symbolic interpretation. Personally and culturally, we understand what we perceive through the filter of our past experiences. This triumvirate of domains—Signal, System, and Symbol—must all be present in some combination or ratio for immersion[1] to occur. Therefore, to define

[1] Staffan Björk and Jussi Holopainen, in *Patterns In Game Design*, divide immersion into similar

immersion for the purpose of this thesis, we will attempt to create spaces of emotional investment combined with a sense of cognitive and spatial engagement, which can be amplified through a number of finely tuned perceptual cues rather than crude increases in scale. Systems freed from linear progressions of narrative or that adjust to their surroundings and modify the story they tell have the potential to increase a user's sense of engagement in the storyworld itself.

This thesis centers on the introduction of small amounts of computation to increase the viewer's sense of immersion by overcoming the limitations of flat displays or more novel form factors. The thesis work includes three separate, but complementary, systems.

Infinity-by-Nine allows us to simulate a full-scale environment within a limited space, such as a home theater. It takes advantage of the different neural processing in the non-foveal region of the eye to augment the viewing experience by projecting a synthetic lightfield, generated in realtime from existing video content, into the viewer's peripheral awareness. Because visual acuity drops off quickly in the non-foveal region, the generated content can be low-resolution, so long as it is scene-consistent, allowing us to create an immersive experience in a

---

categories, but call them sensory-motoric immersion, cognitive immersion and emotional immersion, respectively. In addition to these, they add a new category: Spatial immersion – Spatial immersion occurs when a player feels the simulated world is perceptually convincing. The player feels that he or she is really "there" and that a simulated world looks and feels "real" [Björk].

computationally efficient way.

The Narratarium is a context-aware, 360-degree, catadiotropic projection system that enhances a user's interaction with a story, game, or other object. The system applies pattern recognition and natural language processing to gestural, ludic, or narrative input to create an interactive, immersive environment that can be tailored to a variety of applications.

The 3D Telepresence Chair adds an autostereoscopic 3D display, which is visible without special glasses or other equipment, to traditional teleconferencing. The system combines a Pepper's Ghost apparatus with a portable office chair setup, allowing a remote participant to attend meetings as a full-parallax "hologram."

All three of these projects support the objective of this thesis: to explore the potential for immersive gains through the addition of minimal computation and perceptual cues to traditional display technologies.

2. Background

2.1 Wherefore Immersion?

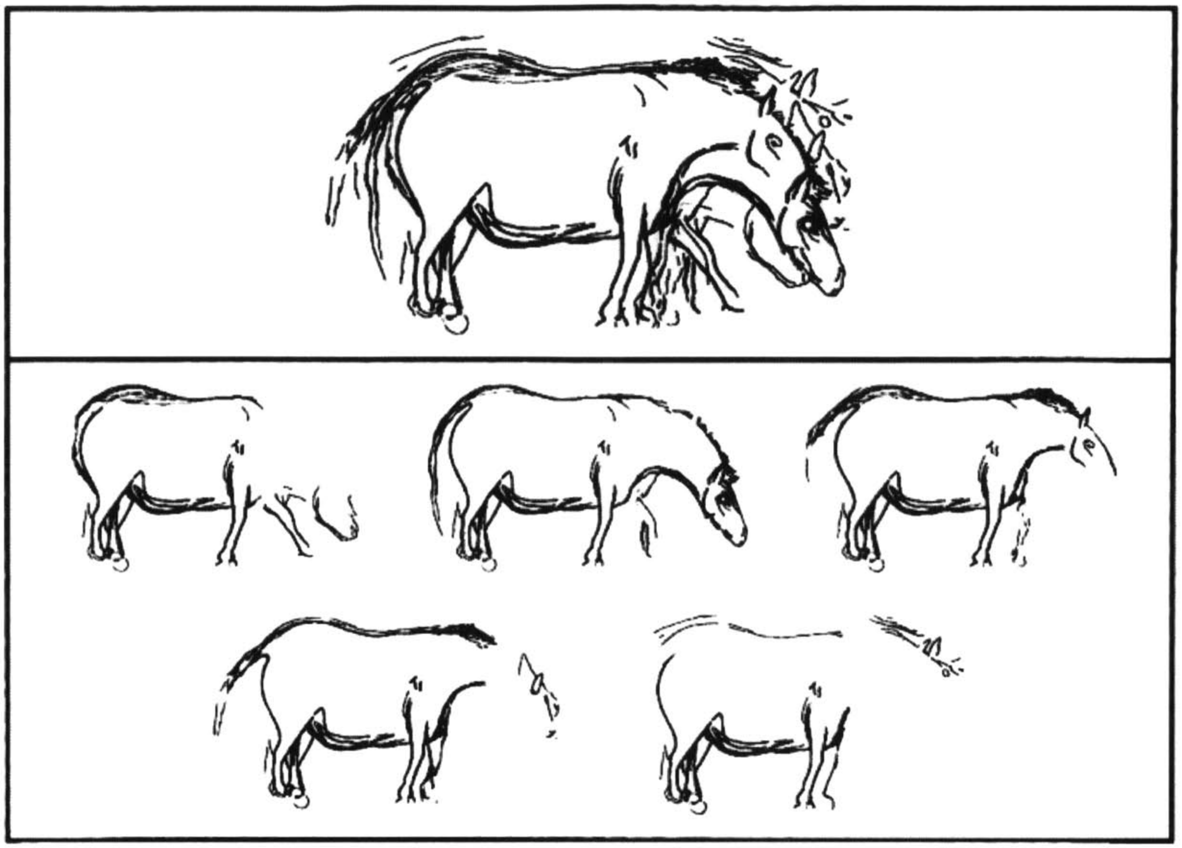
Cave painting, Lascaux, France, 15,000 to 10,000 BCE

Storytelling matters. One need only to look to any of the multimillion-dollar, visual effects-driven extravaganzas that disappear from the multiplex within a week to make room for the next attempt to see that increases in the scale and volume of spectacle that ignore the importance of symbolic processing—aka, the story—will fail to produce a sense of immersion. Stories are how we filter and understand reality. Stories allow us to compare that which has gone before, personally and culturally, with that which is happening now and make decisions based on this comparison.

Humans will apply complex pattern-recognition skills to find stories even when they are not obviously present. We structure information into narrative packages that can easily be compared or stored for later processing. We seek out immersive experiences to facilitate this process. Immersion is attenuation to provide more signal and less noise. Immersion provides a cleaner flow of information, allowing the pattern recognition centers of our brains to drill down to ever-deeper levels and find the story we need to fill our understanding.

In a crowded movie theater, dimming the lights signals the trail-head of the immersive experience. It attenuates the visual noise of the theater and its inhabitants, shuts out spurious information that may overload or confuse the pattern recognition needed for immersion, and focuses attention on the screen. Deep immersive experience allows the cleanest of signals to reach the deepest centers of our brains, where they will have the most significant narrative and aesthetic effect.

Humans have always sought out ways to make experiences more immersive, allowing them to glean greater understanding. Cave paintings dating from 15,000 BCE, found in the Chauvet Cave in Southern France, show examples of the artist employing "split action movement by superimposition of successive images" [Azéma].



A representation of the "split action movement by superimposition of successive images" found in Chauvet Cave.

In modern production terminology, this is "tweening," a part of the animation process. The superimposition seems merely stylistic until one takes into account the sources of illumination used deep within the cave. The flickering of a torch or fire pit would have triggered a persistence-of-vision effect and presented the viewer with animals that walked or roared or leapt, leading participants already immersed in the natural structure of the cave deeper into the immersive experience of the story of a hunt or the mythology of the tribe these images are believed to represent.

As human culture moved out of the Paleolithic into the Neolithic, labor specialization began to appear. One such specialist apropos to this conversation was the Shaman. The shaman operated within the tribe as doctor, lawyer, soothsayer, and entertainer. One of the aspects of shamanism world-over is an aspect of performance in which the audience, be they medical patients in need of cures or initiates engaged in rites of passage, are plunged into an immersive experience with the shaman acting as guide and stage manager. Northern shamans of the Tuva would often perform in darkened tents, using oral storytelling techniques to guide the audience through the performed journey. Shamans were experts in animal mimicry, birdcalls, and ventriloquism, which they used as auditory special effects to deepen the immersion of the participants and allow them to focus on the story being told or the spiritual cure being sought [Eliade]. Shamanic performance space was engineered to control lighting, sound, and the audience's sight lines, all to increase the immersive potential.

Moving forward through the Agricultural Revolution into the Bronze Age in Europe, storytelling and information display technology evolved as human cultures grew into larger City-States. Theater, from the Greek "theatron," means "the seeing place," and was specifically designed to maximize sight lines and audio resonance. Individual shamans could no longer perform the myths and stories reinforcing social cohesiveness for the increasing numbers of people living in the new cities. Smaller rites, such as the Oracle at Delphi, retained elements of shamanic performance, but the unified domain of the shaman in ancient Greece was split into the law courts, which share their form and placement with the theater; practicing doctors who began to study the body rather than use psychopomp to induce a placebo effect; and actors whose professional training in rhetoric allowed them a prestigious place in society. The shaman's roles had been divided, but the need for a performance space that served large numbers of people and could be engineered to increase immersive experiences continued to grow.

The Greek theatrical performance did not just include immersive technologies, such as sound-enhancing masks and synchronized choral speaking; it was also the culminating ritual of a yearly festival. The festival of Dionysus, held in the spring, was one of the many "year-king" rituals designed to celebrate the return of the vegetative spirit [Frazer]. It focused not only the space of Athens, but also the time. The "liminal" space/time of the festival would further immerse the participants, both actor and audience, in tales that defined what it meant to be Athenian or even human as the participants understood it at the time. The specialized affordances of the theater, and isolation from day-to-day time and space, allowed audience members to set aside their common concerns—the noise, as it were—and concentrate on the story being expressed—the signal—in a ratio much more conducive to immersion. The symbology of the data presented could then be processed by the human mind and turned into reality.

Following a Kurzweillian curve of technological innovation, the next great leap in immersive technology in the West occurred during the Renaissance, with the advent of Linear Perspective in the graphic arts. No longer was the relative size of persons rendered in a portrait meant to portray their importance, as it had been since Egypt. Rather, attempts were now made to capture a true sense of illusory depth on a canvas.

Through the use of mathematics, keen observation, and possibly even the use of early optics, artists began to apply 3D-to-2D image transformations of subjects and landscapes onto the image plane [Hockney]. Depth cues such as occlusion, foreshortening, and atmospheric effects became the tools for the visual effects extravaganzas of the day.

Large-format religious paintings, triptychs, and painted chapel ceilings were meant to immerse supplicants and overpower their perceptions with the majesty and awe of divine power, while functioning to reinforce the power of the divine's earthly representatives. But the Church was not the only one interested in the immersive and persuasive power of this new process in art. A growing merchant and middle class soon began to challenge the Church and landed gentry for positions of dominance and prestige. They commissioned the finest of artists—those employing the most advanced practices in optics, chemistry, color theory, and anatomy, and engaging in ever-increasing and ever-refined use of linear perspective—to present fantastic vistas and narratives that were no longer tied to the Catholic Church, but meant to show the power and possibility of the secular mind.

In this period, the immersive power of the new display technologies was no longer just to the "glory of God," but allowed for a human agency to derive its own path. Linear perspective and verisimilitude became the ways in which we created the world in our own image, without the need for the mysteries of the divine. Not long after, with the Protestant Reformation, the illusory powers of art and music began to be denigrated and removed from Calvinist churches, which replaced them with contemplative practices and a sober design palette designed to be specifically non-immersive; all symbol, with no signal and no supposed need for immersion to attenuate it.

One can argue that the austerity of the Calvinist church merely relocated the necessity of immersion to the predominant form of information burgeoning at the time—the printed word. With the advent of movable type and the following increases in literacy, the immersion necessary to clarify the signal to allow pattern recognition to happen more easily was occurring not in the public space of the church, but in the private space of reading. As more printed materials became available, literary forms such as the novel appeared, where the immersion takes place not spatially but emotionally and cognitively. To a Calvinist, a clean austere room and the word of God as represented in a printed copy of the Bible were all that was necessary to contemplate the mystery and will of the divine. But as these stories were being told in print, and then the mind, to create a sense of belonging to a particular sect or denomination, other stories were being told that began to forge a sense of nationalism, of belonging to a people or a state.

Near the end of the 16th Century, after Puritan rule had run its course, the English populace reclaimed its monarchy. With it came a return of the Theater and a headlong rush into pure, effects-driven, immersive spectacle. While immersive, illusory spectacle was no stranger to the theaters and court before Cromwell, its use was limited by the prevalent architecture at the time, as well as the size of the halls in which masques making use of spectacle were performed. In 1547, John Dee, future court philosopher and magus to Queen Elizabeth I, built what seems to have been a mechanical flying dung beetle for an undergraduate production of Aristophanes' Pax. During the action of the play Dee's artificial insect took to the air, inspiring "a great wondring, and many vaine reportes spread abroad of the meanes how that was effected" [Woolley]. But true advancements in architecture and stagecraft would be imported from the Continent by stage design practitioners such as Inigo Jones and Christopher Wren.

For the first time, buildings dedicated to immersive spectacle were constructed solely to execute non-religious narratives purely as a form of mass popular entertainment. Jones and Wren brought with them two- and three-point perspective scenic painting, moveable scenery, weather effects, controlled lighting, and a daunting number of flying machines. Theater had moved indoors, away from the foul weather that could cancel a performance, and shifted to later in the day, to take advantage of better lighting technology and an audience with ever-increasing leisure time. Restoration Theater was the pride of the English people, and while the literati preferred the more constrained and dialogue-driven works we associate with the period, the common people of the day delighted in mass spectacles like the Siege of Rhodes, which could last up to 13 days. The expense of mounting such a production could make or break a theater company. Thomas Shadwell's 1674 adaptation of the Tempest, with its mechanical marvels, began the national sanctification of Shakespeare, who had been admired previously but was now made into England's greatest playwright, as a people began to create the story of who they were. The opening stage directions read:

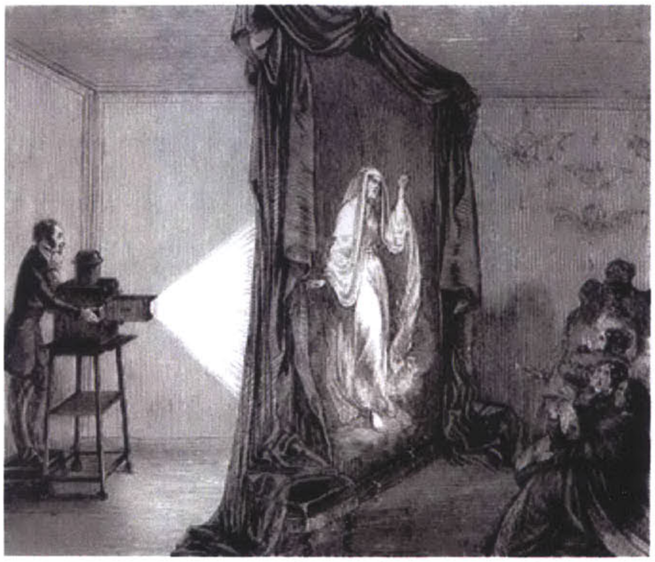
> *The Front of the Stage is open'd, and the Band of 24 Violins, with the Harpsicals and Theorbo's which accompany the Voices, are plac'd between the Pit and the Stage. While the Overture is playing, the Curtain rises, and discovers a new Frontispiece, joyn'd to the great Pylasters, on each side of the Stage... Behind this is the Scene, which represents a thick Cloudy Sky, a very Rocky Coast, and a Tempestuous Sea in perpetual Agitation. This Tempest (suppos'd to be rais'd by Magick) has many dreadful Objects in it, as several Spirits in horrid shapes flying down amongst the Sailers, then rising and crossing in the Air. And when the Ship is sinking, the whole House is darken'd, and a shower of Fire falls upon 'em. This is accompanied with Lightning, and several Claps of Thunder, to the end of the Storm.*

Stage illusion and the immersion it provided drew the 17^{th} Century into a moderning world, in which the story on the stage took a back seat to the story of modern technology and engineering that paved the way to the Industrial Revolution.

It is no surprise then that stage illusionists would be the first to adopt and adapt the new media of film and moving pictures when they arrived at the end of the 19^{th} Century. Following the era of the Restoration Spectacle, smaller forms of immersive theater began emerging along the boulevards of Paris and traveling fairs of Europe and America. Although the principles were known as far back as 1668[2], advances in optics technology and chemical illumination led to the Magic Lantern show becoming a preeminent form of popular entertainment. Light would be projected through painted glass slides, which could then be moved to create animation. Techniques we associate with film—dissolves, fades, superimpositions, cross-editing, different forms of lighting, and different camera angles—were used to heighten the engagement and immersive qualities of the production. Tales from the Bible, phantasmagorical horror shows depicting frightening ghosts and demons, travel and science lectures, all were common on the Magic Lantern circuit.

---

[2] *It produces Effects not only very delightful, but to such as know the contrivance, very wonderful; so that Spectators, not well versed in Opticks, that could see the various Apparitions and Disappearances, the Motions, Changes and Actions, that may this way be presented, would readily believe them super-natural and miraculous. So far our Inventor; who has not contented himself with the bare speculation, but put the same in practice some years since, in the presence of several members of the R. Society, among whom the publisher had the good fortune to see the successful performance of what is here delivered* [Hooke].



A rear-projection magic lantern show in a private home.

A detailed slide from a Phantasmagoria show.

Lantern shows could be set up in homes, fraternal lodges, fair grounds, and other temporary venues, attracting huge crowds from a populace not yet saturated by the mediated image. But the venue of most interest to the history of Immersive Display technologies is the permanent "stereopticon" lantern and stage illusion theater of Georges Méliès. Having been present at the Lumière brothers' first exhibition, Méliès immediately saw the value of the moving image to stage illusion and creating immersive experiences for his audiences. Méliès quickly reversed a projector bought in London to create his first camera, taught himself how to shoot and develop film stock, and quickly went on to create some of the first and most memorable examples of visual effects in history. The moving image had replaced the glass slide of the Magic Lantern.

While the new technology gave audiences the highest level of immersion and verisimilitude they had ever experienced, tales of people running from the screen upon seeing Lumière's locomotive speeding toward them are exaggerated for effect. As French scientist Henri de Parville wrote,

> *The animated photographs are small marvels. ...All is incredibly real. What a* ***power of illusion!*** *...The streetcars, the carriages are moving towards the audience. A carriage was galloping in our direction. One of my neighbors was so much captivated that she sprung to her feet... and waited until the car disappeared before she sat down again.*
>
> (Emphasis added.)

Here we see illusion offered to an audience so immersed that they worry for their safety, immersion so perfectly attenuating the signal that carries the symbolic story that the human fight-or-flight response is engaged. Vive le Cinéma.

Film remained the dominant form of popular, visual entertainment until the advent of television and its rise to power in the 1950s and '60s. While television would become a mass medium in its own right, it could not challenge the immersive affordances of the cinema. Film was being projected in full Technicolor long before an affordable color television could be offered to consumers. The size of the television screen negated even the grandest of epics when they were transmitted into the home for the first time. Radio at this time offered an immersive experience better suited for the home, and in its early days, television programs were akin to photographed radio with moving pictures. Radio shows such as "Lights Out" would issue instructions to the listeners to alter their environments to make the suspense drama more immersive. Television is only now beginning to catch up to cinema's immersive qualities, and sadly only through advances in scale and surround sound.

Film itself has followed this ill-advised quest for scale, starting early on with the advent of wide format Cinemascope, VistaVision, Cinerama, and finally IMAX. Current efforts to increase immersive quality include increased frame rates and stereoscopic projection. All of these advances come at the steep price of an exponentially escalating amount of computation, processing, and storage. Television and film have always been able to increase immersive engagement through the symbolic arena, emotionally and cognitively, but seem to rely on "summer blockbuster" or "tent pole" films filled with spectacle designed to immerse viewers in, as Shakespeare correctly predicted, "sound and fury signifying nothing." In our equation of Signal to System to Symbol, the immersive spectacle is necessary to clarify what little significant symbol is being presented within the signal.

VR Goggles and Pinch Glove

C.A.V.E. System

But film and the cinema was not the end all, be all of immersive display technology. Once computers became powerful enough to create synthetic images, the race was on to develop the ultimate in immersion. Virtual Reality (VR) systems were the hot property of the early 1990s [Rheingold]. Appearances of VR systems in news media and fiction implied that the technology would be commonplace within the foreseeable future. Helmet-mounted displays and pinch glove interfaces allowed VR participants to shut out the outside world completely for the first time and replace it with a computed synthetic reality of bits rather than atoms.

VR wasn't the first to try and break the fourth wall to take advantage of other senses. William Castle, showman extraordinaire, was the king of novelties such as electrified seats and floating skeletons timed to appear in the auditorium as a ghoul on screen flew toward the audience. Another entrepreneur of immersive experience, Morton Heeling, went on to create the Sensorama, a display device aimed at stimulating all of the senses, including smell and touch. While Castle and Heilig's work was novel, their creations never left the arena of the gimmick. Each of their systems remained solidly within the paradigm of passive spectator facing a flat screen, with a few cheap thrills thrown in. VR, on the other hand, seriously challenged the predominance of the seated interface.

VR allowed users to roam about a virtual world that appeared to envelop them, even if the display hardware was in actuality two small monitors in a helmet. Stereoscopic vision and the ability to move through space, either by walking in open physical space while the system tracked the user's movements, or walking on any number of contrived treadmills, were at last going to allow us to treat "Computers as Theater" [Laurel] and perform "Hamlet on the Holodeck" [Murray].

Helmet-mounted displays eventually gave way to less alienating display technologies, such as the recursively named C.A.V.E (CAVE automatic virtual environment) system, which projected a synthetic environment on rear projection flats and motion-tracked the user to derive and update his or her position in the simulated space [Cruz-Neira]. Humanity had come full circle; from the cave of the Paleolithic to the C.A.V.E of the 21st Century, we still watched our hopes and dreams flickering in the light of the torch or projector on the wall in front of us. Each step of the way, we developed new technologies to increase the immersion of the experience—immersion helping to clarify, helping to attenuate the signal, filtering the noise—so we could feel and process the messages passing from one human being to another, no matter the distance traversed or the time elapsed. Immersion allows us a deeper and more meaningful human understanding.

3. Related Work

Many of the immersive technologies discussed in the prior section altered the balance of signal to symbol with very little thought to the biological system that would process them. That biological system, us, is gifted with many more senses than many display technologies provide content for. Even within the normal senses that display technologies cater to, mainly sight and sound, there are physical, neurological, and cognitive affordance that can be exploited for deeper immersion. Display technologies that are willing to adjust to often-ignored portions of the sensory and pattern recognition portion of the triumvirate ratio have the potential to increase immersion more deeply and more cleanly without incurring the cost of increased computational load or overscaling. Some work has been done in the areas covered by the three prototypes discussed in the thesis, and has been used either as a starting point or as supporting material for the development of the immersive displays we are building.

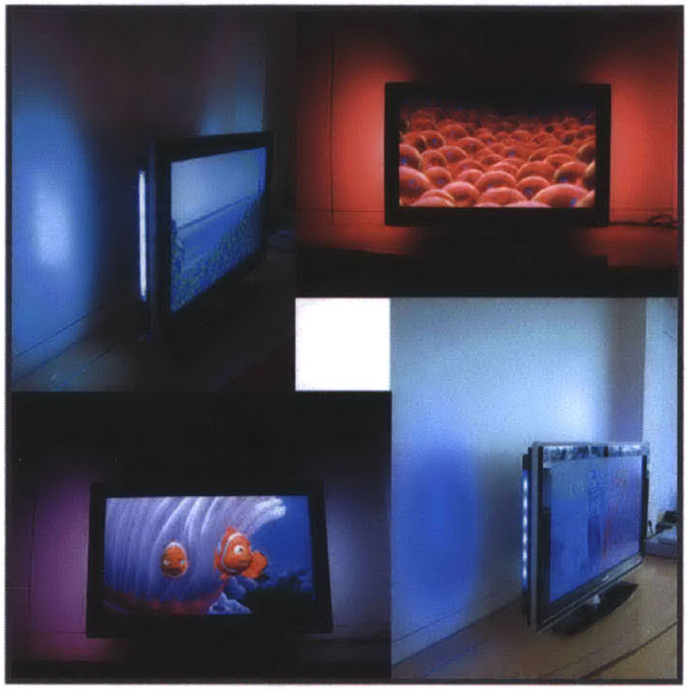
3.1 Infinity-by-Nine

3.1.1 Bruce

Infinity-by-Nine is in part inspired by results obtained in "Bruce," an unpublished 2002 project by Surj Patel of the MIT Media Lab's Object-Based Media Group. A grid of networked, color-controllable LED lights was placed on the ceiling above a large-screen television. Dynamic "lighting tracks," creating not just room lighting but also motion effects, were authored in the group's Isis scripting language for scenes from several films, including Blade Runner and Clue. Viewers reported that they perceived an increased sense of involvement with the films [V.

Michael Bove, Jr., private communication].

3.1.2 Ambilight



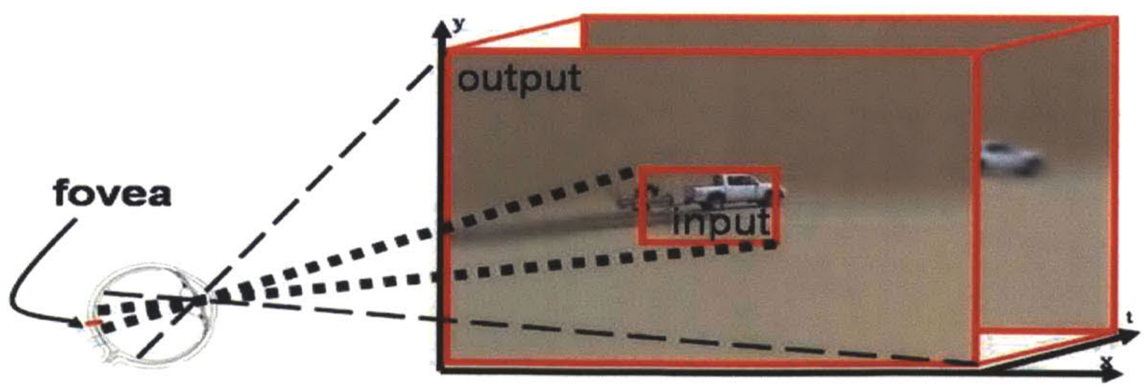
Philips Ambilight

During user studies and demonstrations of Infinity-by-Nine, the comparison heard most often is to a commercial product by Philips, marketed as the Ambilight system. Ambilight operates by taking the azimuthal distribution of colors of the top and side edges of the screen and passing these RGB values for each frame in realtime to a line of RGB LEDs mounted on the edges of the television. The LEDs then project the averaged color against the wall behind the television. The Ambilight system mimics a larger screen, but limits the experience to the same plane as the main viewing display and restricts it to the low resolution of a single, averaged color per screen edge [Philips Research].

The Ambilight system uses the eye's natural "brightness adaptation" to decrease or minimize awareness of the bezel, in addition to suggesting that the television screen is larger than it actually is. It does not collect or compute any knowledge of the source media other than the averaged color information. Ambilight in no way attempts to bring the storyworld into the confines of the user's space.

While the Ambilight system does bear a superficial resemblance to Infinity-by-Nine, it is not aimed at creating a spatially three dimensional, adaptive, and immersive experience in a computationally efficient way. Infinity-by-Nine holds per-scene state knowledge and computes a virtual camera to project the light field out into the user space at 30 frames per second. However, the Infinity-by-Nine system could easily be modified to incorporate a mode similar to Ambilight if the viewer desired less immersion than usual.

3.1.3 Ultrawide Foveated Video Extrapolation



An example of Avraham and Schechner's Ultrawide Foveated Video Extrapolation in use.

Tamar Avraham and Yoav Schechner's work, which was inspired by Ambilight, expands the scale of a two-dimensional video or still image. Like Infinity-by-Nine, their system also extrapolates or outpaints color samples from the edges of a raw video stream via "an improved completion algorithm that rejects peripheral distractions [and] significantly reduces attention-drawing artifacts." More importantly in relation to Infinity-by-Nine, they focus on "a foveated video extrapolation approach [which] exploits weaknesses of the human visual system, in order to enable efficient extrapolation of video, while further reducing attention-drawing artifacts" [Avraham].

Like Infinity-by-Nine, Ultrawide Foveated Video Extrapolation takes advantage of non-foveate information to provide the visual data that would surround a user standing in the narrative space. However, it extrapolates only in the original plane of the raw video, again in an attempt to fill scale rather than activate the user space. Avraham's algorithm uses a patch-based extrapolation approach with additional processing incurred to minimize distracting artifacts. This computational overhead unfortunately leads to a system that is less than realtime. They state, "wide extrapolations such as those described in the experiments section may take up to several minutes per frame." Our work is distinguished from Avraham and Schechner, and other video texture extrapolation research, in that we are processing the video in realtime on commodity hardware, such as might be found in a typical home setting.

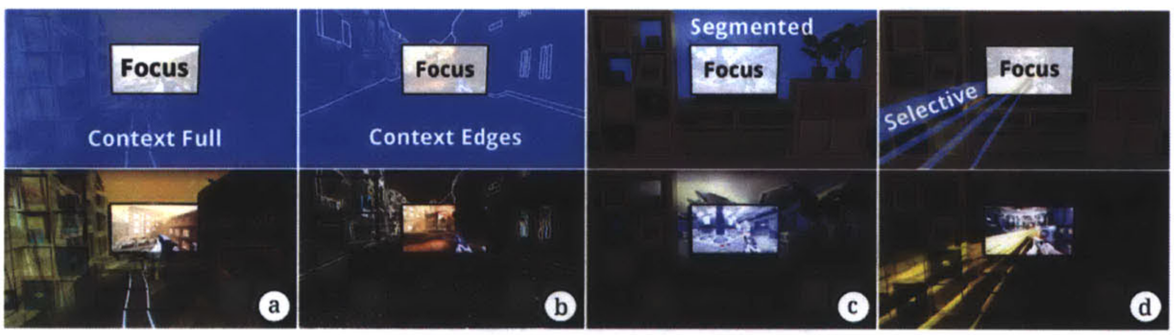
3.1.4 BBC Surround Video

A demonstration of Surround Video in a simulated living room setting.

Much closer to the intent of Infinity-by-Nine is the BBC's Surround Video system. Surround Video is "an immersive video technology which can be integrated into a domestic-scale viewing environment" [Mills et al.]. Surround Video operates by "surround image," taken by a wide-angle lens placed around a raw video stream on a conventional High Definition monitor. The image is aligned to create an extrapolation of the video stream.

The resolution of a projector needed to fill a user's 180-degree field of vision at the maximum level at which humans can resolve detail would be 21,600 x 12,150 [ibid.]. However, detail of this magnitude would be lost beyond 3 percent of the field of vision that corresponds to the fovea. This is true of any display system, including Avraham and Schechner's Ultrawide Foveated Video Extrapolation and Infinity-by-Nine. Like Infinity-by-Nine, Surround Video maps resolution per degree of viewing angle to the gradient of human visual acuity, as determined by photoreceptor density. The BBC system places lower-resolution content in the periphery, but does not extrapolate in realtime. It requires content to be produced using a double camera rig system, and the two videos must be carefully registered and synchronized for the projection to be effective. Surround Video cannot produce an immersive effect for non-Surround Video films, video games, television or other forms on the fly.

3.1.5 Microsoft's IllumiRoom



Different styles of IllumiRoom's Context projection.

IllumiRoom is also a system that augments the area surrounding a television with projected visualizations, but it is meant to enhance the traditional gaming experiences, not necessarily film or television. Patrick Baudisch and Nathaniel Good coined the term "Focus + Context" screens to describe systems such as IllumiRoom, Surround Video, and Infinity-by-Nine, which is a good summation of how these interfaces operate.

IllumiRoom further subdivides the types of context according to different rendering techniques used in the "Context" portion of the field of view. Full, Edges, Segments, or Selective areas of the "Context" area can be chosen to project different styles of "peripheral projected illusions" [B.

Jones, et al]. Using a depth map generated by a Microsoft Kinect ® and a calibration routine allows the projector and gaming monitor to be registered together. A 3D point cloud of the area surrounding the monitor or television is created and texture mapped with the RGB signal of the Kinect. This point cloud is then projected onto the original geometry of the room in a fashion similar to the art form known as Architectural Projection or Projection Mapping. This allows for interactive effects. For example, if a character takes damage in the game, the system could send a distortion wave effect rippling through the user's living room. Through the use of depth sensing, different areas of the field of vision can be segmented and differing effects projected there according to events in the game.

Currently, IllumiRoom is targeted at the gaming environment, but experiments with a dual camera setup similar to that of Surround Video have been attempted. A key factor of IllumiRoom versus Surround Video and Infinity-by-Nine is the majority of the "Context" projection happens on the computed geometry surrounding the television, utilizing little to none of the ceiling or side walls. The illusions do not project past the user's peripheral vision, but remain predominantly forward of the gamer.

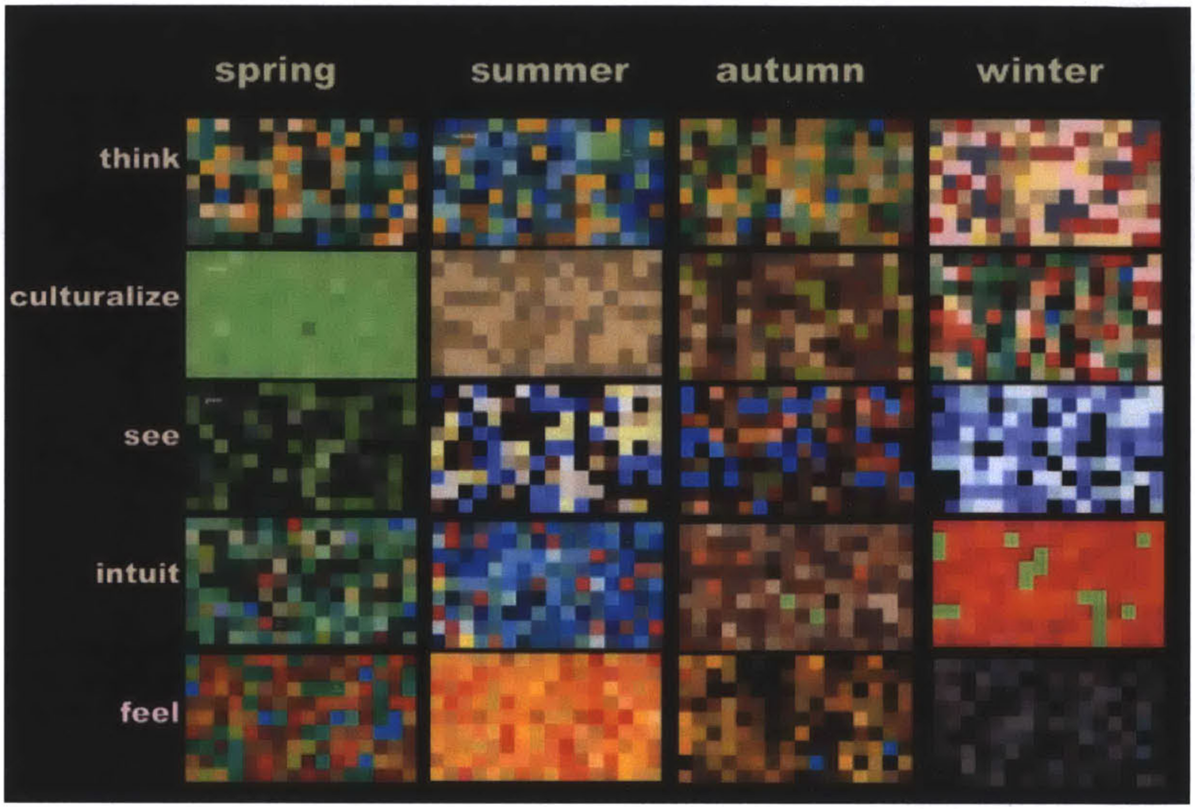
3.1.6 Virtual Reality

Infinity-by-Nine also solves some of the problems encountered in virtual reality (VR) research. VR systems—whether based on head-mounted displays or projector-equipped "caves"—are designed to be completely immersive, but they require custom-made content, and seem to be most effective when combined with physical environments specifically designed to enhance the virtual worlds they depict [Mine]. Although Infinity-by-Nine could be used to showcase films or television programs specifically tailored for it, it is primarily intended to enhance existing movies, sports programming, cooking shows, and other traditional content. In addition, cave VR systems require that every potential view and location in each scene be rendered at high resolution to maintain the illusion of a cohesive environment. Not only is this computationally expensive, it is also wasteful and unnecessary, something which VR research is beginning to accept [Lawton]. In our system, we do not intend for the viewer to look directly at the synthesized images, and thus we purposely render them with low spatial resolution.

3.2 Narratarium

3.2.1 The Aesthetiscope

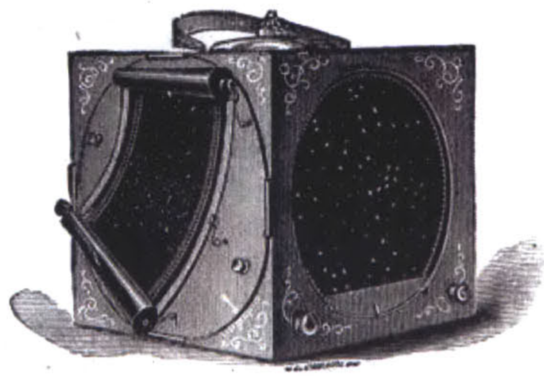
The Aesthetiscope, built by Patricia Maes and Hugo Liu, creates a colorful visual display based upon the aesthetic character of a written text [Liu et al. 2005]. The color palette is determined through psycho-physiological heuristics that apply a customized Jungian analytic framework to the text. While projected on a full-scale wall of the "Living Room of the Future," the display never attempts to envelop the user in an immersive experience. The Aesthetiscope could be an excellent addition to the modes available on the Narratarium and allow users to explore a spatially immersive experience of the text.



A visual result of the Aestheticscope's linguistic analysis.

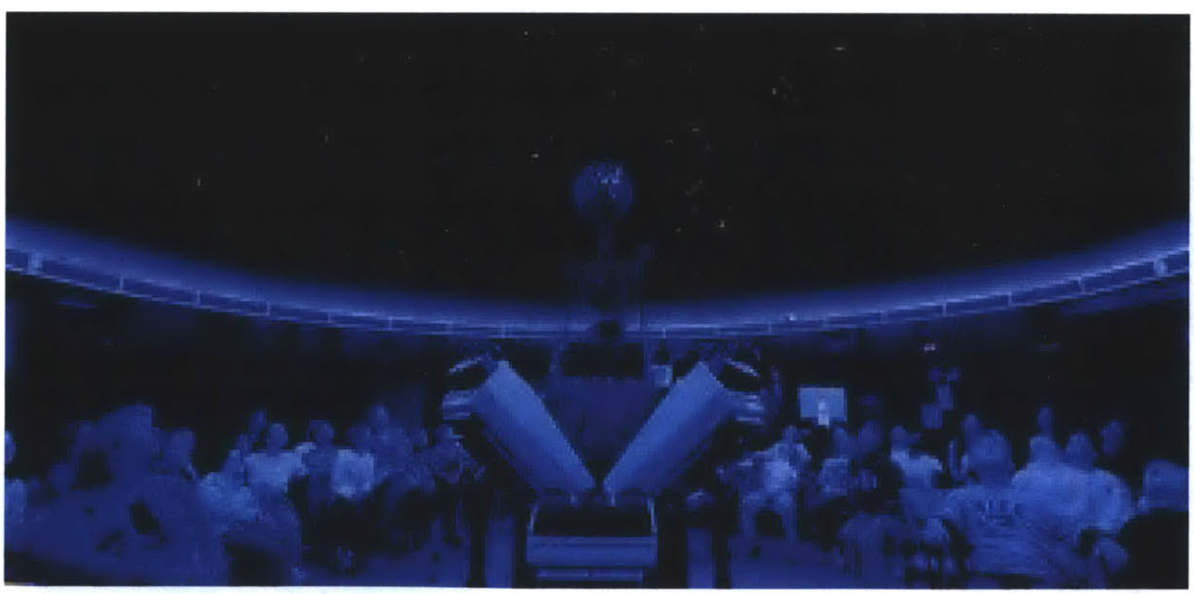
3.2.2 Other 360-Degree Projections and Displays

There have been many attempts to create a 360-degree viewing experience. In 1880, Bailey's Astral Lanterns placed a light source inside a box comprised of tin plates that had been pierced to project the position and relative luminance of stars and planets in the night sky onto the walls and ceilings of a schoolroom or museum.



Bailey's Astral Lantern

Current planetarium projection technology is well advanced, with precision lenses and powerful projectors grouped to achieve maximum image quality. These special-purpose installations are expensive and immutable. Often, content created for them is meant to be used for several years, to pay for the initial outlay of their production.



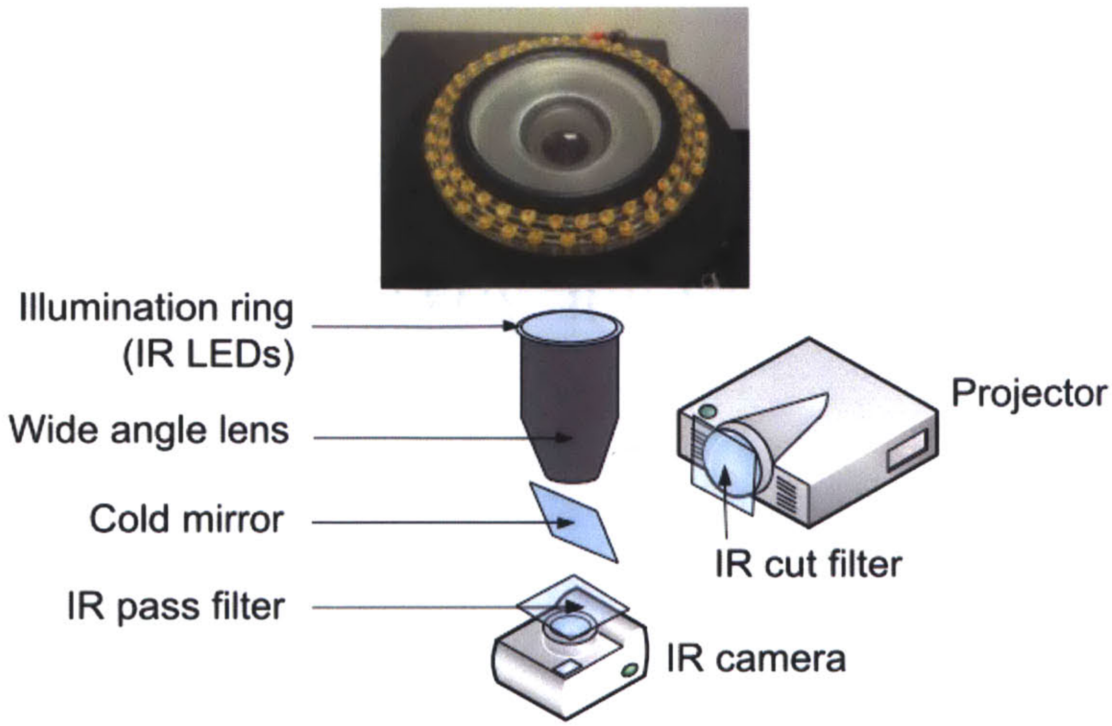
A modern Planetarium starshow.

Planetarium projection systems are often tied directly to the auditorium they are built in and solely to the subject matter of astronomy. While they are good at simulating the experience of standing and looking into the night sky, or can take viewers on a journey through the solar system, they are single-subject and do not react in a dynamic way to the input of the audience. Home versions of planetarium projectors often amount to little more than updated versions of the Bailey's device, with brighter light sources and better optics.

Cyclorama or diorama theaters were popular in the 19^{th} Century and consisted of continuous panoramic paintings that encircled a small audience and usually presented a simulacrum of famous battles from antiquity, or popular foreign travel destinations. The 1889 Cyclorama interpretation of Ben Hur on Broadway ran for over 20 years and involved real horses and chariots. One can only wonder whether the olfactory sensation this addition offered increased the sense of immersion or not.

3.2.3 Multi-Point Interactions with Immersive Omnidirectional Visualizations in a Dome

Closer in form to the Narratarium, Hrvoje Benko and Andrew D. Wilson of Microsoft Research have created an interactive immersive experience using gesture and speech recognition to interact with a projected dome. Their Pinch-the-Sky Dome utilizes a single centrally located omnidirectional projector-camera unit capable of projecting across the entire 360 degrees.



Pinch-the-Sky hardware rig.

An infrared (IR) camera is used to track gestures, while IR laser pointers provide additional input. The projector and the IR camera share the same optical axis. A ring of IR LEDs provides IR illumination, and an IR pass/cut filter combination ensures that the camera sees only IR, while the projection is limited to the range of visible light. Pinch-the-Sky offers several modes, including astronomical data exploration, social networking 3D graph visualizations, immersive panoramic images, 360-degree video conferencing, a drawing canvas, and a multi-user interactive game.

What all of these forms of immersive projection share, whether professional or at home, is that they are linear, passive, and non-interactive storytelling technologies that present a set of known images in a known sequence and remain blind and deaf to the input of the participants.

The Narratarium, on the other hand, requires the input of the audience or ludic agent to even begin to operate. It extends to multiple genres by applying natural language processing and pattern recognition not only to written texts, but also to spoken language, physical objects, free play, and other improvisational experiences. The output includes color samples, text, patterns, audio, animated textures, and moving images intended to augment the input in a context- and content-aware way.

3.3 3D Telepresence Chair

3.3.1 Pepper's Ghost

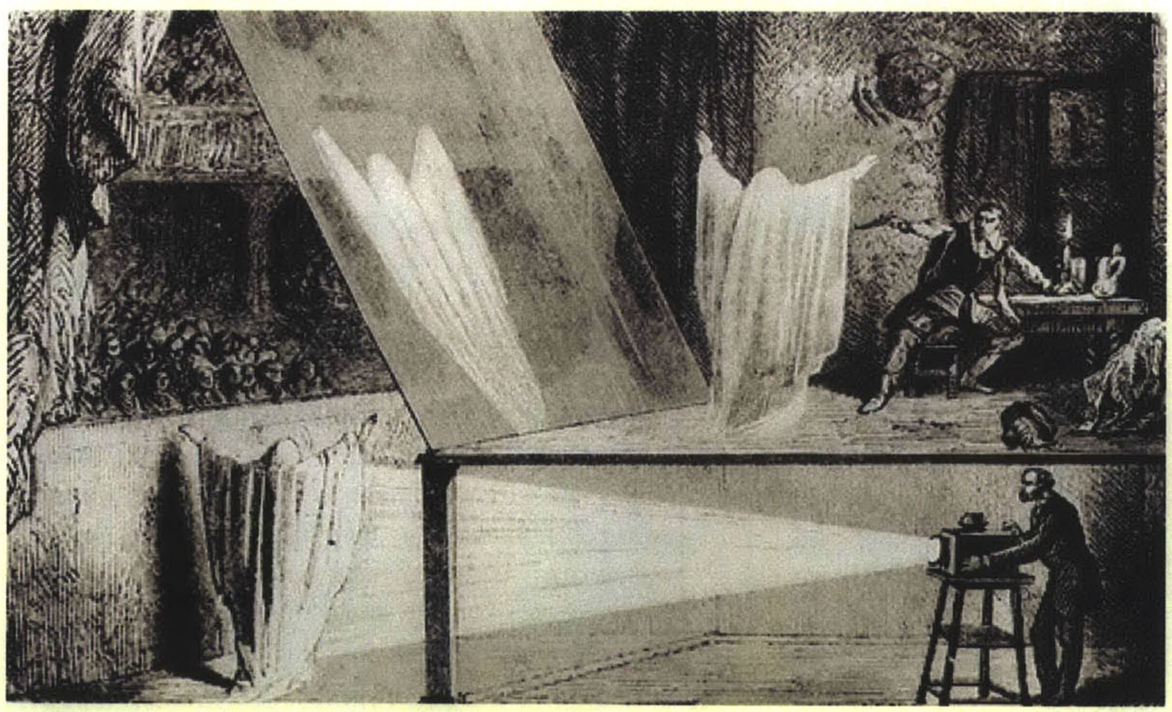


Illustration of Pepper's Ghost in front of a live audience.

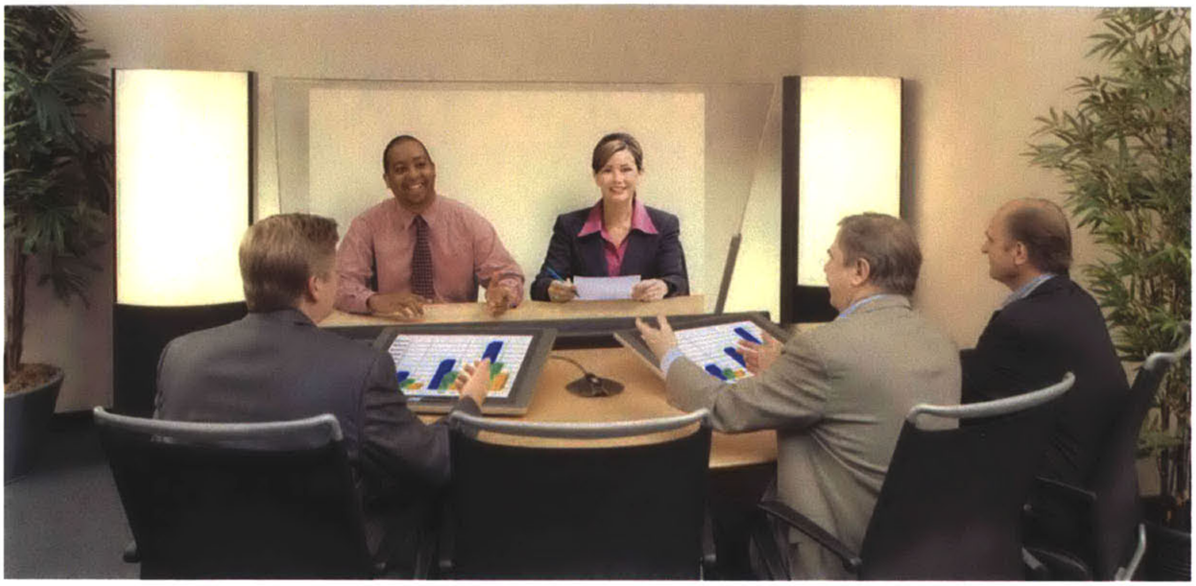
The Pepper's Ghost illusion upon which the 3D Telepresence Chair is based was first described by Giambattista della Porta, a 16th Century Neapolitan scientist and natural philosopher. In 1584's Magia Naturalis (Natural Magic), he describes an illusion entitled "How we may see in a Chamber things that are not."[3] Nearly 300 years later, in 1862, engineer and inventor Henry Dircks demonstrated a small-scale version of this illusion at a booth in the Royal Polytechnic Institute of London. Dircks was the proprietor of a Phantasmagoria, a stage show featuring fearsome displays of the supernatural created with the optical and special effects of the day. He may have been introduced to the ghost illusion by traveling Italian stage magicians familiar with della Porta's version, as these performers likely shared the same traveling circuit as the Phantasmagoria and Magic Lantern shows. However, Dircks' version of the effect was so cumbersome that it would have required the theater to be remodeled or purpose-built to house the effect. After seeing the demonstration, John Henry Pepper, director and "Professor" of the Royal Polytechnic Institute, developed the effect into a practical illusion that could be more easily deployed in preexisting theaters. He has been remembered ever since as the inventor of the effect. It has been used to create stage illusions for over 100 years.

---

[3] *Let there be a chamber wherein no other light comes, unless by the door or window where the spectator looks in. Let the whole window or part of it be of glass, as we used to do to keep out the cold. But let one part be polished, that there may be a Looking-glass on bothe sides, whence the spectator must look in. For the rest do nothing. Let pictures be set over against this window, marble statues and suchlike. For what is without will seem to be within, and what is behind the spectator's back, he will think to be in the middle of the house, as far from the glass inward, as they stand from it outwardly, and clearly and certainly, that he will think he sees nothing but truth. But lest the skill should be known, let the part be made so where the ornament is, that the spectator may not see it, as above his head, that a pavement may come between above his head. And if an ingenious man do this, it is impossible that he should suppose that he is deceived* [della Porta].

Current applications include the ballroom scene at the Haunted Mansion ride at Disneyland and Walt Disney World, and the recent "holographic" performance by Tupac Shakur at Coachella. There, AV Concepts of San Diego stretched a space age material known professionally as Musion Eyeliner™ across portions of the stage, and projected an animation in which Digital Domain—a visual effects house in Venice, California—had composited a 3D head of Tupac on top of a body double who was filmed performing several of Tupac's greatest hits. The shortcoming of these illusions is that they remain fundamentally two dimensional, and lack the horizontal parallax implemented by the 3D Telepresence Chair. Heads Up Displays (HUDs) used for targeting systems in the cockpits of modern fighter aircraft are also based on the Pepper's Ghost optical effect and overlay telemetry and reticules on the aircraft canopy to allow pilots to see the information they need without dropping their gazes to the instrument panel.

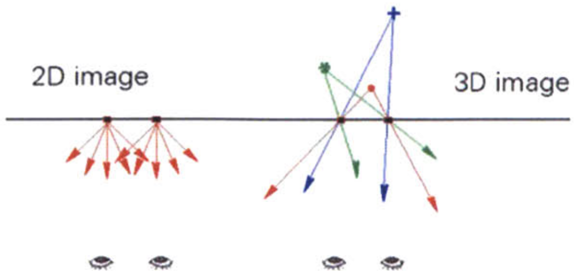
3.3.2 DVE Telepresence



A promotional image illustrating the DVE teleconferencing system.

DVE Telepresence uses the Pepper's Ghost illusion, housed in a purpose-built teleconferencing system. The system is 2D only and meant to be used in a permanent setting. DVE Telepresence has a whole product line based around the Pepper's Ghost illusion, but makes no effort to create a true 3D or holographic experience.

3.3.3 Holografika



Holografika's Voxel Rendering System

3D object displayed directly on diffuser.

The Holografika Corporation has a line of "holographic" displays that use a lenslet array to steer light into discreet view angles to achieve full motion parallax. The screen is within full view of the audience and is meant to be a replacement for traditional 2D displays. No attempt is made to place the on-screen content within the 3D space of the user's field of vision. It is simply a display for viewing from multiple angles. While effective, the display requires 20 Dual Link DVI inputs or 20 Cat6 inputs running from a pre-configured cluster and switch. The 3D Telepresence Chair attempts to achieve a similar level of quality and ease of use with only a single computer using five to seven 720p inputs. However, it could be scaled up to include more computers and achieve a broader field of view.

3.3.4 Teleconferencing and Multi-View

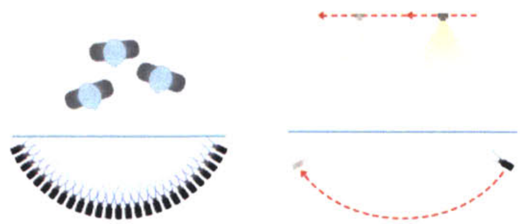
ICT's One-to-Many Video Teleconferencing System

Andrew Jones, Paul Debevec, and their colleagues at the Institute for Creative Technologies (ICT) in Los Angeles have developed a One-to-Many 3D Video Teleconferencing System that allows the remote participant to make eye contact with individual members of a group [A. Jones et al.]. However, the remote participant is displayed only as a floating head, which may be disorienting or alienating for the in situ participants. The system also requires a mirror spinning at over 10,000 RPM, which must be encased in safety glass, making it impractical for everyday use.

The 3D Telepresence Chair is self-contained, has no moving parts, and can easily be moved between teleconferencing locations. Subjects are displayed at life-size, from the waist up, as though they were physically seated at the table.

With Joel Jurik, Debevec has also created a theoretical prototype for a full-parallax light field display utilizing an array of micro projectors, each of which acts as a single pixel [Jurik et al.]. As micro projectors are not yet affordable enough to make a full array feasible, the original setup used a single micro projector mounted on a gantry to determine how many micro projectors would be required to build the complete system. The first experiments suggested that an array of 200 projectors spaced 4 mm apart would create an acceptable level of horizontal-only parallax. Full 3D parallax at a resolution of 640x480, however, would require 300,000 projectors; a resolution of 1920x1080 would require 3 million of them.

Single projector with lenticular lens

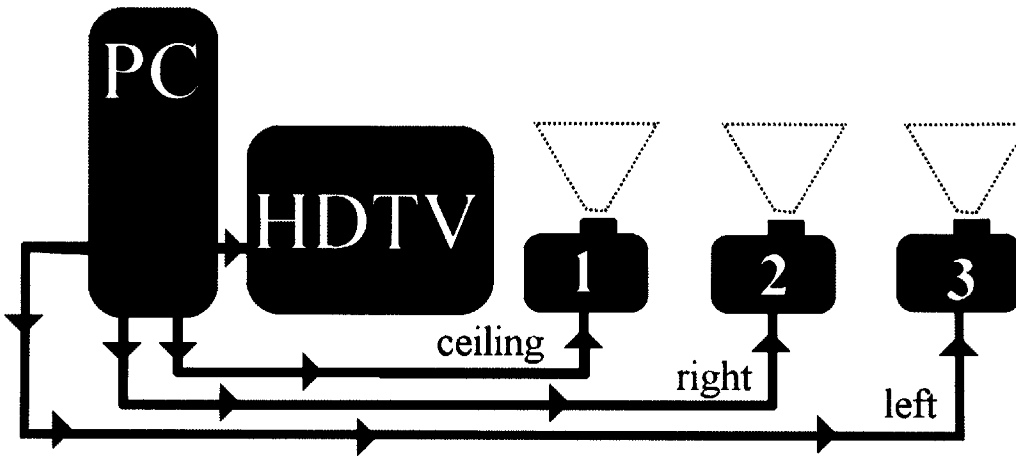


Projector array for full-parallax display

In addition, Debevec's work leaves the lenticular diffuser in full view of the audience, while the 3D Telepresence Chair conceals the mechanism within the Pepper's Ghost apparatus. Debevec's system is also currently being designed to project synthetic imagery, such as 3D models and data sets, rather than live streamed input for use in the telepresence domain.

4. System Development

4.1 Infinity-by-Nine



A functional flowchart of the Infinity-by-Nine system

Infinity-by-Nine currently comprises three projection screens and three projectors arranged around a flat-screen television in a simulated living room environment. The system uses open source computer vision toolkits to apply optical flow, color analysis, and pattern-aware outpainting algorithms to existing video content, such as movies and television programs. This lets us augment the traditional home theater or other viewing environment by creating a synthetic light field beyond the edge of the television screen and projecting it into the viewer's peripheral awareness. The system takes advantage of the lack of detail and different neural processing in the peripheral region of the eye.

Infinity-by-Nine creates a representational model for imagery by matching the gradient of human visual acuity. The fovea, the center of gaze, is the region of the human eye where color-sensitive cones and detail-processing ganglia are at their highest concentration; visual acuity falls off rapidly outside the center of gaze. Avraham and Schechner have shown that features falling outside the fovea are given less processing time by the brain and functionally can be abstracted and extrapolated without creating a break in the narrative space [Avraham et al.].

Other parts of the visual processing system are also coarser-grained in the periphery; a good overview of the subject is provided in "Picturing Peripheral Activity," by S. Anstis. While detail sensitivity falls off in the visual periphery, motion sensitivity remains accurate, and sensitivity to temporal flicker actually increases, so a system such as ours must maintain good temporal performance in the synthesized imagery [McKee et al.]. There is less sensitivity to vertical motion than horizontal axis motion [Duchowski et al.]. The human visual system also demonstrates a peripheral falloff in color perception. In a study by the United States Air Force, subjects were unable to identify the color of a circle beyond 74.3 degrees from the fovea for green, 76.3 degrees for red, and 83.3 degrees for blue [Ancman].

Infinity-by-Nine takes advantage of this non-foveated information to provide a simulation of the environment that would surround a user standing in the narrative space. By taking advantage of the brain's neural interpolation of sparse information in the non-foveated region of the eye, Infinity-by-Nine creates a sense of immersion and participation with very little additional information. The images displayed on the peripheral panels can be processed and rendered at less than half Standard Definition (SD), including motion blur and pattern matching operations, before being re-sized for projection. Great care is taken to maintain the minimal amount of information in the peripheral panels, not only for computational expediency, but also to maintain focus on the target media—the film or television program being experienced. The addition of the synthetic light fields should not distract or draw focus from the experience, but rather add to it by filling in spatial details, in the same way surround sound functions for audio.

A conceptual rendering of the peripheral content generated by Infinity-by-Nine.

Users perceive the scene-consistent, low-resolution color, light, and movement patterns projected into their peripheral vision as a seamless extension of the primary content. By wrapping the user in real-time, adaptive light fields, Infinity-by-Nine creates a low-cost alternative to traditional, processing-intensive Virtual Reality.

The content plays on a generic PC containing NVIDIA GPUs, custom software developed in C++, and ceiling-mounted projectors that direct the synthesized patterns onto screens positioned within the user's peripheral vision. Video is analyzed on a per frame basis using a FAST (Features from Accelerated Segment Test) tracking algorithm [Rosten et al.] to determine salient features for subsequent Lucas-Kanade Optical Flow analysis [Lucas et al.]. A realtime affine camera motion is then computed based on pixel position and velocity. A combination of luminance and pattern-aware histograms are generated and continually referenced to detect scene changes. Upon scene change, pixel color samples are taken from dynamically adjustable regions of the main screen and propagated to side screens using an extrapolated outpainting algorithm. The computed affine camera motion is then applied to the pixel arrays on each side screen to synchronize their flow with the camera motion of the original media content.

Infinity-by-Nine in use.

The overall goal is to create color and movement patterns on the side panels that match the main screen well enough to be interpreted by the brain as consistent in neuro-narrative space [Martinez-Conde]. Currently, transforms and filters are implemented in GLSL (Open GL

Shading Language). Subsequent versions will move the majority of operations, such as the *FAST* feature finding, optical flow analysis, and affine transformation calculation, to multiple shared GPUs, allowing full 1080p HD playback at acceptable frame rates and allowing the inclusion of further advanced features.

4.2 The Narratarium

The Narratarium is a context-aware projection system that creates an immersive storytelling environment to augment narrative activity or creative play using texture, color, and image. In the current iteration of the system, we are using natural language processing (NLP) to listen to and understand stories being told, and thematically augment the environment using color and images.

Currently, the Narratarium synchronizes to a children's book from Hallmark Publishing called *Under the Same Moon*. Illustrations from the book were animated in a visual effects compositing package and loaded onto the unit as QuickTime animations. The Narratarium listens for key phrases on each page to synchronize the environmental projections with the book. The user can also simply call out the animals presented in the book and be taken to the page showing that animal. An ambient soundscape is also triggered for each page to add to the immersive experience.

As development continues, the system will be trained to respond to additional content. For example, we envision a mode in which a parent reads *Where the Wild Things Are* aloud, while the Narratarium paints the walls with foliage patterns in greens and browns as the forest grows in Max's room, then shifts to cool blues, rippling waves, and seagull sounds when he boards the boat to travel to the isle of the wild things.

Hardware and Software

The physical prototype of the Narratarium combines two 500 lumen Optoma 500ML micro projectors, mounted above a hemispherical mirror on a transparent acrylic base to create a catadiotropic projector that extends the generated visual content 360 degrees horizontally and 180 degrees vertically.

The Narratarium's catadiatropic projector.

Under the Same Moon displayed with the Narratarium.

To ensure that the imagery viewed in the immersive environment of the Narratarium is perspectively correct, the Narratarium image engine induces a PreWarp image transform customized to the hemispherical mirror or lens. The software applies a modified and user-configurable Albers Equal-Area Conic Projection to high-definition content such as film clips, illustrations, and realtime animations. Where λ is the longitude, λ_0 the reference longitude, φ the latitude, and φ1 and φ2 the standard parallels:

$$x = \rho \sin \theta$$
$$y = -\rho \cos \theta$$

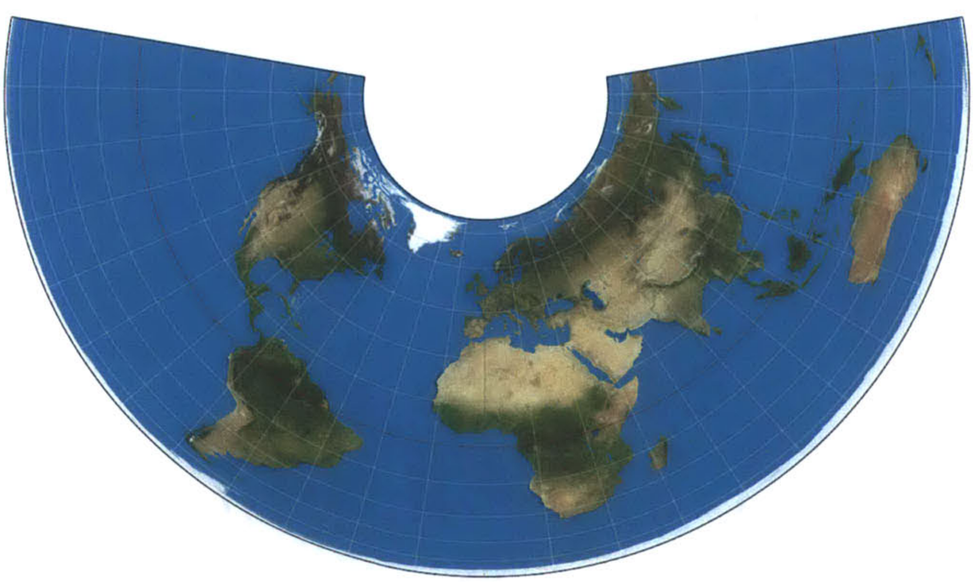
where

$$n = \tfrac{1}{2}(\sin \phi_1 + \sin \phi_2)$$

$$\theta = n(\lambda - \lambda_0)$$

$$C = \cos^2 \phi_1 + 2n \sin \phi_1$$

$$\rho = \frac{\sqrt{C - 2n \sin \phi}}{n}$$



An Albers Equal-Area Conic Projection

Depending on the projection mode, all of the imagery and the results of the GLSL shaders are written into a Frame Buffer Object (FBO) to utilize GPU computing speed and texture-mapped onto a subdivided OpenGL mesh. The mesh is then pre-transformed and projected from the light engine onto the hemispherical mirror, which acts as a lens physically warping the imagery so when it arrives on the wall or screen of a rectilinear or hemispherical display space, it appears to be undistorted. Users can modify the transform to fit any space and save these modifications into an xml file for later recall. The OpenGL mesh can be subdivided as much as necessary to achieve smooth borders, and individual control points can also be accessed and transformed to fit imagery to non-regular or non-rectilinear shapes.

Currently, all creative assets are designed to be 3840 x 1080. This "double wide" HD format is split in half and mapped to two OpenGL meshes, each output by its own projector and each responsible for one half of the immersive environment. Further iterations of the Narratarium may allow for four or more texture mapped meshes to achieve more precise and finite control over projection placement in any environment. A Mac Mini is embedded in the base of the unit to provide input sensing, computation and GPU rendering. The system is small enough to be placed on a bedside table.

At present, an echo-canceling microphone supplies an audio feed to the Mac Mini, which runs custom Speech Recognition software that parses the audio into discreet words. In one mode, the recognized word is matched against a dictionary supplied by the developer and an action, such as displaying a particular page of a book or playing a sound sample is triggered. Another mode uses ConceptNet5 from Catherine Havasi's Digital Intuition Group [Liu et al. 2004] to match emotional vocabulary or visual content with the text, and determines the proper output images, sound, and color. A third mode uses a 2.4GHz radio chip from Nordic Semiconductor to send a radio payload to the unit identifying a particular object being handled by a participant. The identity and action of the object then triggers output images and sounds to establish a deeper context. A perfume bottle can trigger video clips of the flowers to make the perfume being harvested, or a child's toy can trigger the projection of the landscape of that toy's story universe.

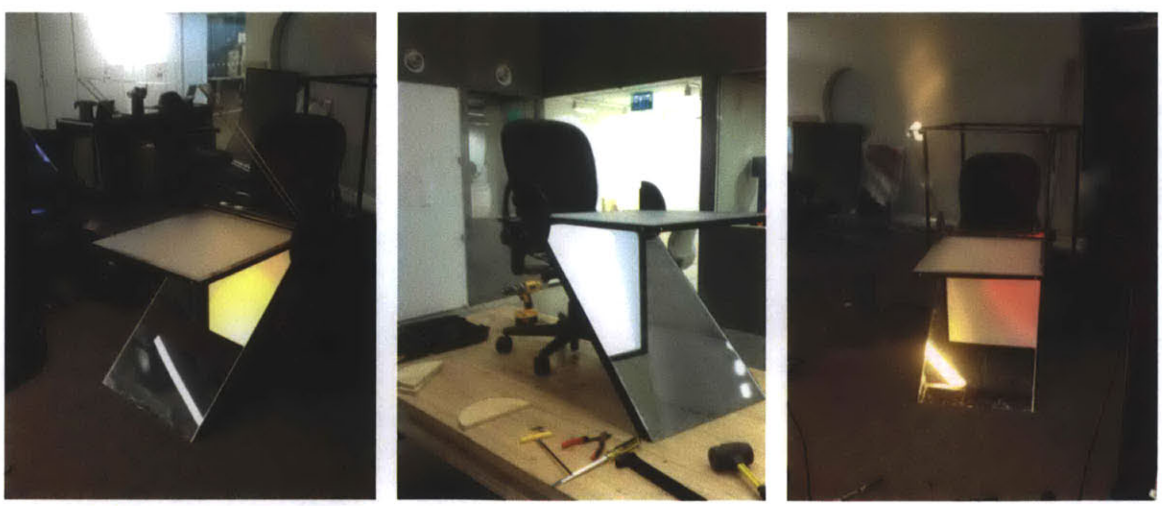
Future releases could add the ability to interface with other objects, including ebooks, using wireless data, infrared tracking, or non-audio information from sensors, such as position, tilt, and acceleration.

4.3 3D Telepresence Chair

The 3D Telepresence Chair combines an autostereoscopic (no glasses) 3D display engine with a Pepper's Ghost apparatus to create an office chair that appears to contain a remote meeting participant. The system geometry is also suitable for other applications such as tabletop or automotive heads-up displays.

The 3D Telepresence Chair allows a remote meeting participant using Skype to have his or her image projected onto a transparent surface that can be placed behind a conference table or in a similar setting. A webcam mounted on the chair and looking through the virtual image provides a visual feed to the remote participant, and two-way audio communication occurs through normal computer speakers and an echo-canceling microphone. The speakers and microphone are spatialized, to enhance the illusion that the remote participant is physically present in the meeting.

The system utilizes a PC with several graphics cards, each connected to a number of micro projectors. Multiple view angles of the same subject are projected through a Light Shaping Diffuser (LSD) and then reflected off of a beamsplitter in the Pepper's Ghost apparatus. Each individual view angle is controlled by the LSD. By blocking incorrect views, it allows each of the viewer's eyes to receive only the appropriate view angle for the viewer's current position, thus creating an autostereoscopic display. The viewer sees objects rendered with full horizontal parallax and occlusion. Unlike lenticular diffusers, such as those used on autostereoscopic televisions, the LSD contains no dead zones or flipped view angles. The number of view angles can be increased or decreased dynamically to strike a balance between detail and network bandwidth.



Design and construction of the 3D Telepresence Chair

The remote participant uses a multi-camera array or a Microsoft Kinect depth camera outputting a real-time mesh rendered by multiple virtual cameras and fed into the appropriate output projector in the array.

Hardware

The optical effect behind the Pepper's Ghost illusion can best be explained using the Fresnel Equations.

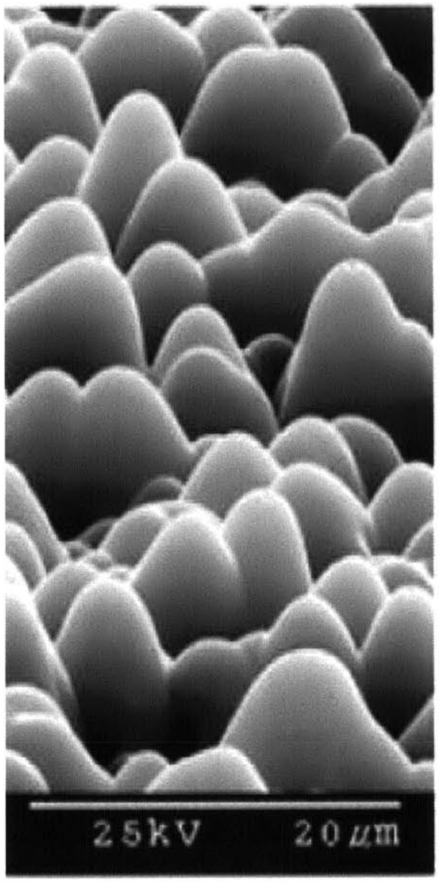
$$r_s = \frac{n_1 \cos\theta_\mathrm{i} - n_2 \cos\theta_\mathrm{t}}{n_1 \cos\theta_\mathrm{i} + n_2 \cos\theta_\mathrm{t}}$$

$$t_s = \frac{2 n_1 \cos\theta_\mathrm{i}}{n_1 \cos\theta_\mathrm{i} + n_2 \cos\theta_\mathrm{t}}$$

$$r_p = \frac{n_2 \cos\theta_\mathrm{i} - n_1 \cos\theta_\mathrm{t}}{n_1 \cos\theta_\mathrm{t} + n_2 \cos\theta_\mathrm{i}}$$

$$t_p = \frac{2 n_1 \cos\theta_\mathrm{i}}{n_1 \cos\theta_\mathrm{t} + n_2 \cos\theta_\mathrm{i}}$$

As non-polarized light from the projector bounces off the mirror and is transmitted through the Light Shaping Diffuser, finally striking the beamsplitter at an incident angle of 45 degrees, a certain portion of the light, depending on the index of refraction of the material comprising the beamsplitter and governed by the above equations, will reflect at the same angle and fall on the eye of the user. The rest of the energy will be bent and transmitted through the beamsplitter at an angle also governed by the above equations and Snell's Law.



Randomized, non-periodic microstructures make up the Light Shaping Diffuser

Several materials were evaluated for use as the beamsplitter during multiple iterations of the 3D Telepresence Chair. Acrylic was originally chosen for its light weight and clear optical properties as well as its self-rigidity. The 2D-only version of the chair used a fixed piece of 1/2" acrylic with a refraction index of 1.48, which according to the equations above results in ~8% of the energy passing through the Light Shaping Diffuser being reflected back to the viewer. However, 1/2" acrylic suffers from a visible double image, as the transmitted wave has been offset in space by the refraction angle and then reflects again off of the back surface of the material.

Although the second image is less bright, because of its lost energy at each reflection/refraction event, it is still bright enough to be seen and perceptually degrades the image quality.

One solution to the double image challenge is to use a thinner material, so the transmitted beam is less spatially offset before this second reflection occurs. The thinner the beamsplitter, the less perceptible the double image will be, because the amount it diverges from the first reflection dwindles. Therefore, two different forms of Mylar were used for the fully three-dimensional version of the chair.

Mylar typically has an index of refraction of 1.64 to 1.67, much higher than glass or acrylic. Mylar delivers back almost 13 percent of the energy to the viewer, resulting in a distinctly brighter image for the same amount of energy transmitted from the projector. The purpose-built Mylar known as Musion Eyeliner is slightly metalized to provide even greater reflectance.

While the Fresnel properties of the beamsplitter provide the Pepper's Ghost illusion, it is the unique properties of the Light Shaping Diffuser that elevate the 3D Telepresence Chair from 19^{th} Century parlor trick to a 21^{st} Century technology. LSD diffuser films are weakly diffracting optical components with non-periodic holographic surface relief patterns that redistribute light intensity along the 3D microstructure surface of the LSD film. Fabricated using holographic exposure techniques involving laser recording of sub-micron speckle features in photosensitive materials, the non-periodic microstructures change both the amplitude and phase of each local wavefront, creating a statistical ensemble of micro-diffraction profiles in the far-field. The size and orientation of the microstructures can be designed to create a specific angular redistribution of energy to create custom view angles.

In the case of the 3D Telepresence Chair, the Light Shaping Diffuser provides a new view every 10 degrees horizontally, while all views are expanded 60 degrees vertically. With 10-degree view angles, the projector array and interocular distance between virtual cameras can be tuned to show each of the viewer's eyes a spatially offset image, creating an autostereoscopic effect without the need for polarizing or shutter glasses. The autostereoscopic effect is maintained no matter where the viewer is standing within the arc of the projector array. The system can be configured to provide variable amounts of autostereoscopic effect and parallax depending on varying distances to the chair. Some autostereo can be sacrificed for a broader viewing angle zone, or maximized for a smaller area of interaction.

The mirror used by the 3D Telepresence Chair is simply a silver-backed acrylic chosen for its lightweight characteristics and ease of fabrication. The mirror is held at a 45-degree angle to the projector array and the Light Shaping Diffuser. This allows for a longer path of the projected material to be sized and focused properly. While the acrylic mirror operates well, it is not a first-surface mirror, so a small amount of image degradation, perceived as a softening, is induced by small multiple refractions occurring on the acrylic surface as the light travels through it while reflecting off the mirror. However, a first-surface mirror would add much more weight to the design of the system. The trade off between weight and image quality is currently acceptable.

The projector array consists of six Optoma 500 lumen micro-projectors sourced from three Nvidia GPU cards. The array is calibrated by projecting a registration grid rendered from all six virtual cameras in the same virtual location, and aligning the gridlines so each view angle completely overlaps. Each projector in the array can be adjusted in heading or pitch to register the views. In future iterations, a bracket with fine adjustment screws will house the projectors to make this process less arduous and protect the alignment from unintentional bumps.

As projectors move away from the center of the array, the incident angle increases and increasing distortion becomes apparent. Currently, this is compensated for roughly with the keystoning features native to the projectors, but in continuing iterations each projected view will be mapped to a deformable OpenGL mesh with finely adjustable vertex control points. This will allow small warps, similar to those used by the Narratarium, to "pre-correct" the image for better registration and quality. A seventh output from the GPU cards is a simple monitor which displays debugging information and application state. This monitor is optional and can be removed while the chair is in use during telepresence.

The 3D Telepresence Chair is designed to be mobile and easy to use. The chair retains the original wheels of the office chair from which it was born and additional wheels are mounted where the mirror meets the floor. This allows the whole unit to be easily rolled from one location to another or for small adjustments in framing and scale to be made relative to the projector array. The mirror-diffuser-beamsplitter unit is mounted to the arms of the chair with only two bolts, using holes originally designed to attach the arm rests. If necessary, the unit could be removed and the chair returned to its original state simply by reinstalling the padded seat and armrests.

Software

The software running the 3D Telepresence Chair is a custom-written OpenGL and C++ application utilizing open source libraries for Microsoft Kinect stream capture and point cloud generation. A point cloud of the remote user is generated from the RGB and Depth images provided by the Kinect. The RGB color information is mapped to the vertices of the point cloud to provide texture mapping, and the point cloud is subsequently rendered from six different view angles determined by a user-definable interocular offset distance. Each of these rendered views is fed in realtime to the projector array and then through the Light Shaping Diffuser. Controls in the software provide for tuning scale and interocular offset for the virtual renders, as well as near- and far-field clipping of the depth map to maximize eight bit range of the Kinect's VGA depth image.

Ideally, a depth camera with a narrower field of view would allow for a better image. The Kinect was originally designed as a gaming device to track multiple users in a room of unknown size, while its use in the 3D Telepresence Chair is almost the opposite need—a single user at a known distance. A software Running Average Filter is applied to the Kinect's signal to decrease noise in the image stream.

5. User Studies

5.1 Infinity-by-Nine

The user study conducted for Infinity-by-Nine took place in a simulated living room environment in the Object-Based Media Group area on the 4th floor of the Media Lab's classic building. The test living room included three small couches arrayed before a large screen television set at a comfortable distance. Two projection screens were hung from the ceiling to simulate the side walls of the room, while a third screen was suspended above the space to operate as a ceiling. Each screen was slightly angled inward to create a foreshortening or small "forced perspective" effect that emphasizes the relationship of the imagery on the screen to its origin on the main television. The side and top projection surfaces were luminance- and color-matched as closely as possible to the main video screen to help create a continuous light field.

Group consensus was used to decide which film from an offered list would be shown. A short entrance questionnaire establishing the viewer's normal film- and television-watching preferences was provided prior to an explanation of Infinity-by-Nine's operating principles. Test subjects were encouraged to verbalize or ask questions at any time, whether it be to comment on a particularly enjoyable moment created by the algorithm, or to criticize something that did not work well. Data from the entrance questionnaire indicated that almost all the test subjects normally watched film or television in groups, and almost never watched alone.

Great care was taken in the entrance questionnaire and in the system explanation to avoid the use of the word "immersion," in order to avoid biasing the experience for the users and skewing the results. During the study viewers were encouraged to switch seating positions on the couches, ensuring that everyone was given time in the center seat versus the edges, and to move their heads or bodies around the synthetic light field to explore the effect. Movement, emotional reaction, and intensity of focus were noted during the film.

The study continued for the entire length of the feature film and a short debriefing period after the film's conclusion. During the debriefing period, subjects were asked to fill out an exit questionnaire allowing user feedback and suggestions.

General questions the study endeavored to answer were whether the peripheral content would help create a state of immersion and whether this immersive state would allow greater focus or more enjoyment of the viewing experience. Observations were also taken to indicate whether the viewers appeared to be experiencing eye strain or fatigue; none were noted.

5.1.1 Entrance and Exit Questionnaires

Entrance Questionnaire

The questions on the Entrance Questionnaire were designed to establish the home television watching habits of the participants. The study was particularly interested in the size of the screen, which devices the participants tended to use, and the social behavior of the participants while engaged in viewing.

1. Do you enjoy watching films and/ or television? Both?

2. Do you enjoy films at home or the theater? Both?

3. If you enjoy watching films at home, what do you usually watch them on? Circle all that apply.

Television

Computer Monitor

Projected

Hand Held Device (Ipad, Tablet, Smart Phone)

Other (Please Explain)

4. For your favorite way to watch film or television:

A) About how big is the screen or projected image that you watch?

B) How close or far do you usually sit?

5. Do you enjoy watching stereoscopic (3D) films? If so, what do you like about the experience? If not, what do you dislike about it?

6. Do you watch films or television alone? If not, with whom do you watch?

Exit Questionnaire

The Exit Questionnaire was designed to allow the subjects to offer qualitative observations about their experience viewing a feature film on Infinity-by-Nine. Subjects were asked to offer suggestions on types of content they might enjoy on Infinity-by-Nine. In addition, they were asked to compare the experience offered by Infinity-by-Nine to the experience of watching a stereoscopic film, as we surmise that Infinity-by-Nine could enhance the theatrical stereoscopic experience.

1. What did you think about the experience? (Any or all thoughts are appreciated)

2. Did you find the peripheral content distracting?

3. If you have watched stereoscopic (3D) films, did you find this experience more or less enjoyable than watching a stereoscopic film?

4. Did you find your experience with Infinity-by-Nine more similar to watching a film at home, or in the theater?

5. The purpose of Infinity-by-Nine is to create a sense of "immersion," or emotional and cognitive engagement, with filmed content. Did you feel that watching a film with the peripheral content was more or less immersive or engaging than watching a film without it?

6. What other films do you think would work well on Infinity-by-Nine?

7. What other forms of entertainment do you think would work well on Infinity-by-Nine?

5.1.2 User Testing Conclusions

Usual Viewing Habits

The Entrance Questionnaire indicated that most of the subjects prefer to watch films with family or friends, and would rather watch films in a theater than at home. When they do watch television or films at home, they prefer large screens—several subjects said, "The bigger, the better." As Infinity-by-Nine is currently displayed on a 50-inch television screen, and the user studies were conducted with groups of two to six people, the experimental setup was consistent with user preferences.

Immersiveness and Attention

More than 80 percent of the subjects indicated that Infinity-by-Nine increased their sense of immersion. Representative statements included:

- "I find the peripheral inputs brought me into the film without being distracting."
- "More engaging, I loved the extension of the image."
- "This experience seemed to be very immersive like a film."

A surprising result was that several subjects found that this immersive quality helped them bring their attention back to the film if it wandered.

- "Everytime I lost my mind [sic], the ambient draws me back."
- "I often get distracted watching this particular film (Blade Runner), because it's beautiful, but slow. With [Infinity-by-Nine], it held my attention all the way through."
- "Sometimes helped me feel immersed and helped focus."

Many of the subjects offered observations on which shots within the films' vocabularies were most effective on Infinity-by-Nine. The overall consensus was that slow pans, establishing shots, and large, open scenes benefited most from the peripheral content. As one subject put it, "I definitely felt more 'there' in wider, more establishing shots." Several subjects also commented on the way Infinity-by-Nine worked with the films' lighting. For example, one said, "Scenes with light colors and bright lights were impressive and vivid," and another said, "Shots with off-camera light sources seemed to translate well with this presentation." Others were even more specific, stating that Infinity-by-Nine "worked really well in left/right lighting splits," and "the best use of peripheral content was... continuous moving action scenes with one light source."

Some users indicated that smaller, interior shots did not work well, although one found that, "The subtle effect of a light/color wash in a quiet scene are [sic] sometimes more arresting/less distracting than chase scenes and explosions."

Criticisms and Observations

Overall, the subjects did not find the peripheral content distracting. One of them wrote, "Most of the time I would not even notice the walls." However, the criticism that stood out most clearly was that many of the subjects were aware of moments when the optical flow tracking failed to find enough salient features to derive a global camera solve. When this occurs, movement on the peripheral screens does not match the movement of the main screen.

Some of the subjects found this discrepancy to be distracting, but not all of the time. They expressed this in a variety of ways, saying that the content projected with Infinity-by-Nine was "more engaging when it was projecting accurately, but distracting when it lagged," that there "were times the peripheral input did not quite seem to match the action on screen," or that they were distracted "when the main screen was moving and the side screens weren't." One said that "Sometimes [the peripheral content] was a really beautiful, obvious enhancement..., often a subtle enhancement, and sometimes a detriment (fast movement, diegetic flashing lights, etc.)."

Several subjects also noted that the discrepancy was more apparent during quick scene changes, saying that the peripheral content was distracting "only when it changed suddenly and didn't quite match up to the content on-screen, especially if there's a solid, bright bar across the bottom of the side screen," and that, "The thing to avoid seems to be cuts."

Other criticisms involved minor inconsistencies or personal preferences:

- "At some point the extended projection picked the minor color in that scene to render, which changed the whole visual impact of color. (Minor becomes major e.g. warm --> cold) It might not fit to the director's intention…"
- "I wished the peripheral beams were relative to a vanishing point, instead of just horizontal."
- "Wish it was dimmer, more subtle."

Some users also commented that the luminance and color calibration of the side projectors differed from that of the main screen:

- "The bright scenes were way too bright reflected on 4 surfaces. It seemed like just a projection horizontally of the edges of the screen, but I'm not sure it added anything for me."
- "Sometimes side-screens brighter than main image."
- "It would be better if the colors between all screens were calibrated—sometimes the peripheral screens were so bright, they distracted from the main screen."

A few users indicated their position in the lightfield affected their experience:

- "[It] felt like center seat would be best somehow. Like I was too much on the periphery to get the intended experience perhaps?"
- "When sitting in the center, it enhanced the experience. When viewing off-center, it was a bit distracting."
- "More [immersive] when viewed head-on and when the content on the peripheral screens wasn't too bright. Less when viewed off-center.

However, some subjects found that the more accustomed they became to the peripheral content, the less their positions mattered. As one said, "It took a few minutes for the effect to sink in... but once you're immersed it felt really nice. Toward the beginning, leaning forward (so that the edges of the peripheral screens could not be seen at all in my field of view) helped with the feeling of immersiveness. Later that didn't seem to matter as much."

Subjects suggested a number of other applications, such as games (the clear favorite), sports, video conferencing, music visualization and music videos. They also said that they would be interested in watching horror films, animation, documentaries, action movies, science fiction, and westerns. A few specific titles or directors were mentioned, including The Abyss, Top Gun, '90s-era Batman films, Otto Preminger, and Dario Argento.

Subjects were also asked to compare the Infinity-by-Nine experience with that of stereoscopic 3D. Some focused on the differences between Infinity-by-Nine and stereoscopic 3D, saying that "3D effects can be cool in 3D film," "[Infinity-by-Nine] added to experience, rather than transformed e.g. 3D is a complete immersion, this was more 'enhanced,'" "3D would bring me closer to the image. Your system enhances colors, the TV appeared bigger," or "Even less intense (good), but less seamless."

Some gave answers that were unclear or noncommittal. One said, "The depth of image is more compelling to me," which could mean either that stereoscopic is more compelling, or that Infinity-by-Nine is. Another stated that, "It's a wash at the moment. I think there is potential for this being more immersive with [illegible]."

Those whose Entrance Questionnaires indicated that they dislike wearing 3D glasses preferred the experience of Infinity-by-Nine. Others found the experience "more integrated," or said, "This was more comfortable than a 3D film. This brought a more immersive experience than a standard film, and I enjoyed it more than most 3D films, but the novelty might have played a role." One agreed with our theory that Infinity-by-Nine might be used to enhance stereoscopic 3D, saying that, "Any film that works in 3D would benefit from this. Combining this and 3D could potentially be even better."

5.2 The Narratarium

In its current configuration, the Narratarium is fully programmed to respond to *Under the Same Moon*, or to respond to extemporaneous speech with a limited palette of text, color, and shape. The current modalities provide a fundamental introduction to the possibilities of the system. However, at this time we feel that more modalities, including a deeper integration with ConceptNet5, need to be added to the Narratarium in order to properly perform a user study.

5.3 3D Telepresence Chair

5.3.1 Singtel Board Meeting

An early version of the Telepresence Chair was used to allow Media Lab Director Joichi Ito to appear as a "hologram" to the Board of Directors of Singtel at their annual meeting. The board was in residence on the 6^{th} floor of the MIT Media Lab in Cambridge, Massachusetts, while Mr. Ito was in Phoenix, Arizona. The Chair utilized Skype in full screen mode and was limited to two dimensions for this early iteration. Two-way communication went well, even considering that no specialized teleconferencing hardware or dedicated channels were being used. Mr. Ito was utilizing a wireless connection provided by his hotel and the Telepresence Chair was connected to MIT's standard wired network. The members of Singtel's board were delighted with the novelty of the approach and enjoyed communicating with Mr. Ito as if he were in the room.

In order to amplify Mr. Ito's voice so that it could be heard by a large group, his audio feed was input into the room's PA system. This, unfortunately, decreased the illusion that Mr. Ito's speech came from the chair. It also confused the echo-canceling microphone, which was no longer capable of matching its input to what it would normally have been projecting.

Many members of the board stated that the use of the Telepresence Chair was "just the sort of thing they had become Members of the Media Lab for."

5.3.2 Continuing "In Use" Studies

The 3D Telepresence Chair will continue to be evaluated "in use," with user feedback driving design iterations. The Chair will be used through a number of regularly scheduled meetings and demonstrations with experts in the teleconferencing field. Continual feedback will be solicited and incorporated into changes in the hardware and software environments. The goal of these tests is to determine if the affordance of full parallax 3D enhances or eases the teleconferencing experience by providing better communication and whether users prefer the experience to traditional teleconferencing.

6. Future Work

6.1 Infinity-by-Nine

In its current form, Infinity-by-Nine enhances the film and television watching experience, but could easily be expanded to desktop user interface design, user experience (UX) design, security applications, and teleconferencing. Infinity-by-Nine's visuotopic model is well suited to any environment where gaze-dependent computation can be used to create areas of detailed information only where and when needed, while information which is currently less important to the viewer may temporarily require fewer resources.

Expanding the technology to full cinema experience offers the possibility of bridging the divide between traditional film and the current trend of stereoscopic films. In addition to computational savings, the spatial relationships within the model maintain the continuity of a single environment space, deepening the intrinsic understanding of how to operate within this continuum.

Processing the video in a causal fashion restricts our ability to synthesize optimal side images for certain camera motions such as zooming out (where textures need to enter from the outer edges of the side screens); thus we are in the process of adding a pipeline delay to support look-ahead in the analysis and texture generation. Development will also continue in the area of better pattern matching algorithms and other texture synthesis techniques to replace the current outpainting algorithm employed at scene change [Barnes et al.]. These improvements will help to alleviate some of the criticisms expressed by the user study participants.

The system as it exists now offers viewers an enhanced capacity to enjoy and experience preexisting materials. However, the opportunity also arises to ask, "How would custom content created specifically for this system look? Could content providers, such as directors or cinematographers, purposefully place object patterns or colors near the screen edge to take advantage of Infinity-by-Nine's outpainting algorithms? Could genres such as horror use subtle cues in the extended visual field to amplify the audience's experience in moments of shock?" Art forms such as music videos, educational programming, and even advertising may find ways to utilize this technology to maximize their effectiveness with minimal changes to production and without creating additional content.

Future versions are envisioned that include short throw projectors built into the bezel of the television set and map the projection surfaces through the depth map sensing capabilities of range-finding cameras such as Microsoft's Kinect. By utilizing this device or a similar solution to create a scanned, depth-attenuated model of the viewing environment, the system could determine candidate projection surfaces and pre-transform the output of the synthesized material to be perspectively correct once it reaches the display surface, much as the IllumiRoom does, but for the entire space. We are also exploring other optical solutions, such as the use of hemispherical mirrors or other reflective surfaces, as developed by the BBC Surround Video, to reduce the number of projectors required, as well as guiding the image more easily to the best areas of view.

6.2 The Narratarium

As a context-aware projection system, the Narratarium has the potential to provide a new way of exploring narrative and displaying information. Providing this deeper immersion will allow parents or other family members who are traveling or who live some distance away to participate in creative play, rather than encouraging passive consumption of entertainment. This active, shared media will lead to a stronger sense of emotional engagement.

In a retail setting, the Narratarium can provide supporting information and immersive content to more fully inform and delight the customer. Providing an experience, telling the story of a product, or promoting ecologically sound instructions for the refilling or recycling of a product can create a better customer relationship and brand loyalty.

The Narratarium may also interest stage designers, DJs, festival planners, and others whose work involves storytelling. Further modalities can be added as new sensors and output actions are conceptualized.

6.3 3D Telepresence Chair

The 3D Telepresence Chair creates a more natural experience for meeting participants. The immersive quality of the system arises from the sense of a comfortable, emotional connection with a life-sized, interactive representation of a distant colleague, rather than a faceless voice emerging from a telephone or computer speaker. As such, it could result in a decrease in business travel, which would reduce both costs and environmental impact. One can imagine a time when multiple units could be placed around a conference table and a meeting could take place between a mixture of live and remote participants from multiple locations around the globe.

Like the Narratarium, the 3D Telepresence Chair also has the potential to enhance narrative and storytelling opportunities, as well as stage and festival technology, by allowing audience members to view a "holographic" character with full horizontal parallax. These applications would require the projector array, Light Shaping Diffuser, and beamsplitter to be scaled up to fit the venues in which they were installed. The physical prototypes of the system have demonstrated that the output scales linearly.

Media Lab Director Joichi Ito has requested a permanent chair to be ready in his office at the Lab when he travels. He wishes to use the chair to meet with staff and visitors at the Lab while he is traveling on business. Having a 3D Telepresence Chair in his office would allow the local participants to perceive Mr. Ito and interact with him in a more comfortable and lifelike manner.

Topper Carew, a Visiting Researcher/Scholar at the Media Lab, is planning to build a two-dimensional version of the Telepresence Chair as part of a shrine to Dr. Martin Luther King, Jr., which will travel with his documentary film about Dr. King as it circulates during its run. The Chair would show Dr. King seated and telling stories from his past.

As development continues, other narrative possibilities will be suggested and presented. British Telecom (BT) has expressed an interest in having a pair of Chairs separated by the Atlantic for use by their executives in London and liaisons at the Media Lab.

New materials and construction techniques will be sought out to clarify and increase output image quality. As this technology matures, a small camera array might be used, rather than a Kinect, so that H.264 or a novel form of compression could be implemented. H.264 compression is calculated by analyzing the differences between frames separated temporally. We will use this compression to calculate the difference between images separated both temporally and spatially. As the differences between successive frames are small, so are the differences between successive view angles. As H.264 compression operates by transmitting only the differences, this form of compression would be ideal for the 3D Telepresence Chair.

Calculations indicate that the amount of data needed to transmit multiple views increases by only 1.5 to 2 times, even for three or more views of the same subject. The current Multiview Extension to the H.264 AVC already allows for multi-view transmission. Ideally, newer depth camera models will allow the depth image used to create the 3D point cloud to be larger than 640 x 480, which is the current resolution of the Kinect's depth image. An HD depth map would be the number one improvement in quality for the 3D Telepresence Chair.

7. Conclusions

This thesis presented three examples of display technologies that augment or amplify immersive experience though economic computational, cognitive, or perceptional means. Infinity-by-Nine, the Narratarium, and the 3D Telepresence Chair all increase the immersive quality of the content they carry by adding to the users' sense of emotional, spatial, or cognitive awareness. They are designed to answer the problem of user engagement in a computationally economical way, rather than relying upon non-focused, large-scale hardware solutions.

A broad history of Immersive Display technologies was presented to frame these new technologies within the ever-evolving continuum of human perception and cognition. The theoretical argument was made that humans seek immersive experience to condition the Signal (the media being presented) so that the System (the human brain and sensory organs relaying information to it) can more easily apply pattern recognition and evaluation of the Symbol (the narrative or aesthetic message that the storyteller is attempting to pass to the audience).

Infinity-by-Nine demonstrated that scene-consistent, low-resolution color, light, and movement patterns projected into the participants' peripheral vision act as a seamless extension of the primary content and create an immersive experience with little computational overhead. The narrative space of the story can be made to feel as if it is taking place all around the users, rather than just on a screen in front of them. In a user study, most subjects indicated that Infinity-by-Nine created a more enjoyable and immersive experience and allowed them a deeper understanding of the films they were watching.

The Narratarium was introduced as a novel apparatus providing fresh narrative opportunities for storytelling spaces and formats. The Narratarium's inexpensively created immersive environment provides completely new ways of interacting with and creating story experience. The addition of artificial intelligence via Concept Net or other forms add to the potential to create truly unique interactive experiences that allow users to explore more deeply in a natural way using language and the projected environment. Creating opportunities for artists or parents to surprise and delight audiences and children in new ways is an exciting prospect.

Finally, the 3D Telepresence Chair was shown to offer a more naturally engaging and emotionally immersive experience than traditional teleconferencing, using a 19th Century stage illusion and a 21st Century optical material. Being able to interact with a business colleague or family member without the alienating changes induced by typical teleconferencing technologies will create a more conducive and deeper channel of communication without incurring the overhead of a purpose-built telepresence suite.

Each of these technologies has shown that the addition of a small amount of computation, combined with novel ways of taking advantage of human perception and cognition, can lead to greater immersion in an economically efficient way, and this immersion can help deepen and strength human communication and our understanding of story and each other.

References

Ancman, E. "Peripherally located CRTs: Color perception limitations." In *Aerospace and Electronics Conference, 1991. NAECON 1991., Proceedings of the IEEE 1991 National*, pp. 960-965. IEEE, 1991.

Anstis, Stuart. "Picturing peripheral acuity." *Perception* no. 27 (1998): 817-826.

Azéma, Marc, and Florent Rivère. "Animation in Palaeolithic art: a pre-echo of cinema." *Antiquity* 86, no. 332 (2012): 316-324.

Avraham, Tamar, and Yoav Y. Schechner. "Ultrawide foveated video extrapolation." *Selected Topics in Signal Processing, IEEE Journal of* 5, no. 2 (2011): 321-334.

Barnes, Connelly, Eli Shechtman, Adam Finkelstein, and Dan Goldman. "PatchMatch: a randomized correspondence algorithm for structural image editing." *ACM Transactions on Graphics-TOG* 28, no. 3 (2009): 24.

Baudisch, Patrick, Nathaniel Good, and Paul Stewart. "Focus plus context screens: combining display technology with visualization techniques." In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pp. 31-40. ACM, 2001.

Benko, Hrvoje, and Andrew D. Wilson. "Multi-point interactions with immersive omnidirectional visualizations in a dome." In *ACM International Conference on Interactive Tabletops and Surfaces*, pp. 19-28. ACM, 2010.

Björk, Staffan, and Jussi Holopainen. *Patterns in Game Design*. Charles River Media, 2005.

Cruz-Neira, Carolina, Daniel J. Sandin, and Thomas A. DeFanti. "Surround-screen projection-based virtual reality: the design and implementation of the CAVE." In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pp. 135-142. ACM, 1993.

della Porta, Giambattista, *Magia Naturalis (Natural Magick)*. Thomas Young and Samuel Speed, 1658. Accessed May 13, 2013.
http://lcweb2.loc.gov/service/rbc/rbc0001/2009/2009pre23451/2009pre23451.pdf

Duchowski, Andrew T. *Eye Tracking Methodology: Theory and Practice*. Vol. 373. Springer, 2007.

Eliade, Mircea. *Shamanism: Archaic Techniques of Ecstasy (Bollingen Series (General))*. Princeton: Princeton University Press, 2004. First published in French as *Le Chamanisme et les techniques archaïques de l'extase* by Librairie Payot, 1951.

Frazer, James George. *The Golden Bough*. Simon and Schuster, 1996. First published by Macmillan Publishing Company, 1922.

Hockney, David. *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters*. Viking Studio, 2006.

Hooke, Robert. "A Contrivance to Make the Picture of Any Thing Appear on a Wall, Cub-Board, or within a Picture-Frame, &c. in the Midst of a Light Room in the Day-Time; Or in the Night-Time in Any Room That is Enlightned with a Considerable Number of Candles; Devised and Communicated by the Ingenious Mr. Hook, as Follows." *Philosophical Transactions of the Royal Society*. no. 3, 17 Aug. (1668): 741-743.

Jones, Andrew, Magnus Lang, Graham Fyffe, Xueming Yu, Jay Busch, Ian McDowall, Mark Bolas, and Paul Debevec. "Achieving eye contact in a one-to-many 3D video teleconferencing system." *ACM Transactions on Graphics (TOG)* 28, no. 3 (2009): 64.

Jones, Brett R., Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. "IllumiRoom: peripheral projected illusions for interactive experiences." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 869-878. ACM, 2013.

Jurik, Joel, Andrew Jones, Mark Bolas, and Paul Debevec. "Prototyping a light field display involving direct observation of a video projector array." In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, pp. 15-20. IEEE, 2011.

Laurel, Brenda. *Computers as Theater*. Addison-Wesley Pub. Co., 1993. First published by the University of Minnesota in 1991.

Lawton, George. "Making virtual reality more accessible." *Computer* 39, no. 6 (2006): 12-15.

Liu, Hugo, and Push Singh. "ConceptNet—a practical commonsense reasoning tool-kit." *BT Technology Journal 22*, no. 4 (2004): 211-226.

Liu, Hugo, and Pattie Maes. "The aesthetiscope: Visualizing aesthetic readings of text in color space." *Proceedings of FLAIRS2005* (2005): 74-79.

Lucas, Bruce D., and Takeo Kanade. "An iterative image registration technique with an application to stereo vision." In *Proceedings of the 7th international joint conference on Artificial intelligence*. 1981.

Martinez-Conde, Susana, Stephen L. Macknik, Sandra Blakeslee, and Henry Holt. *Sleights of Mind: What the Neuroscience of Magic Reveals About Our Brains*. Profile Books, 2011.

McKee, Suzanne P., and Ken Nakayama. "The detection of motion in the peripheral visual field." *Vision research 24*, no. 1 (1984): 25-32.

Mills, Peter, Alia Sheik, Graham Thomas, and Paul Debenham. British Broadcasting Corporation, "Surround Video." Accessed May 10, 2013.
http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP208.pdf.

Mine, Mark. "Towards virtual reality for the masses: 10 years of research at Disney's VR studio." In *Proceedings of the workshop on Virtual environments 2003*, pp. 11-17. ACM, 2003.

Murray, Janet H. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Simon and Schuster, 1997.

Philips Research, "Ambilight." Accessed May 10, 2013. http://www.research.philips.com/technologies/ambilight/index.html.

Rheingold, Howard. *Virtual Reality*. Simon and Schuster, 1992.

Rosten, Edward, and Tom Drummond. "Fusing points and lines for high performance tracking." In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2, pp. 1508-1515. IEEE, 2005.

Shadwell, Thomas. *The Tempest (Opera)*. 1674. Based on the adaptation by William Davenant and John Dryden.

Woolley, Benjamin. *The Queen's Conjurer: The Science and Magic of Dr. John Dee, Advisor to Queen Elizabeth I*. Holt Paperbacks, 2002.