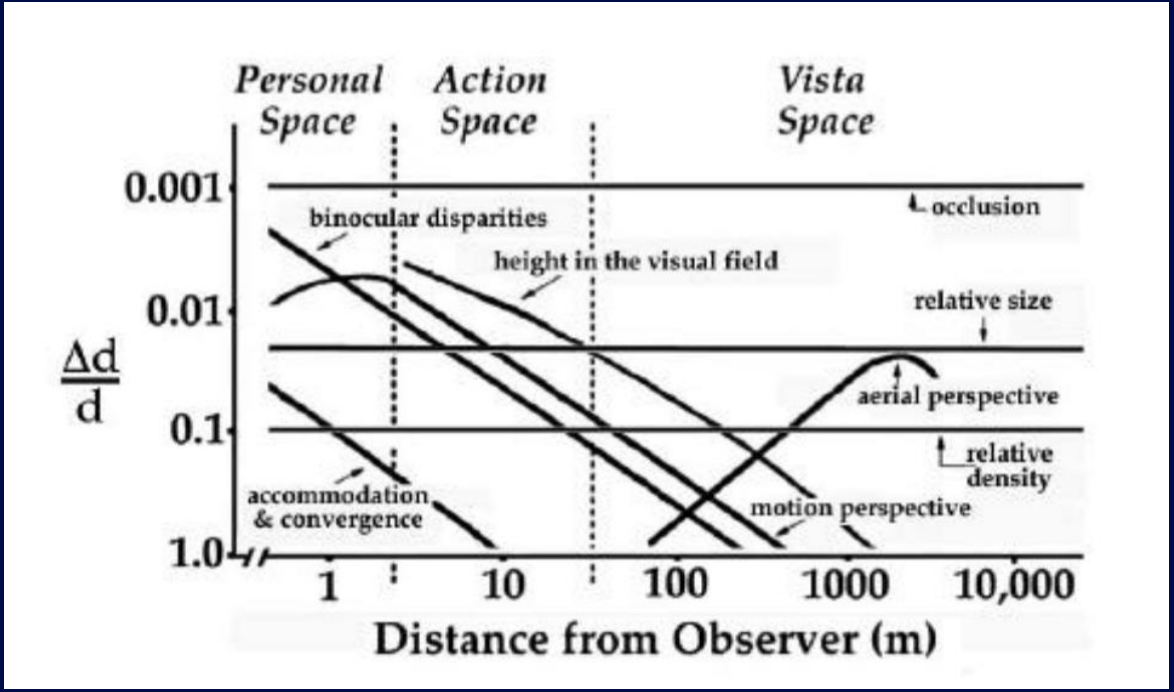
Microsoft

The Holoscene: Virtual and Mixed Reality

Reality is a Color

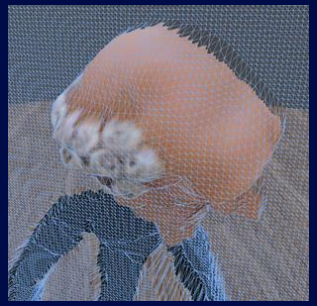
Microsoft Research
Faculty Summit
2015

Mark Bolas,
USC School of Cinematic Arts
& The Institute for Creative Technologies

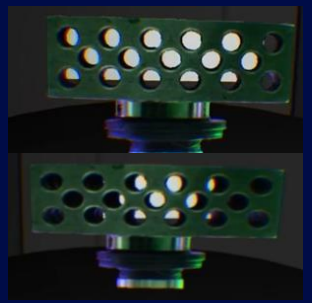


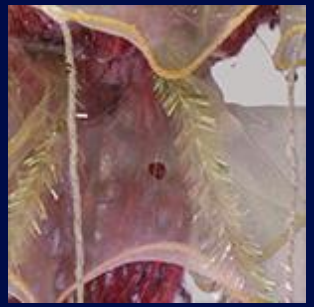
CUTTING, J.E., AND VISHTON, P.M. 1995. Perceiving layout: The integration, relative dominance, and contextual use of different information about depth. In Epstein, W., & S. Rogers (Eds.), Handbook of Perception and Cognition, vol. 5, Perception of Space and Motion, NY: Academic Press.

- Hand Interaction,
- Proprioceptive Placement,
- Fine Motor Control,
- Social and Emotional Reactions

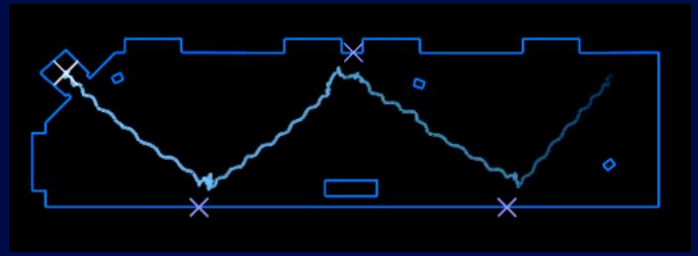



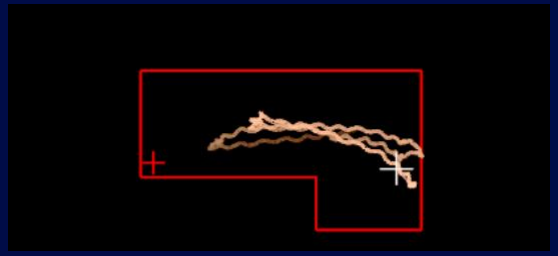




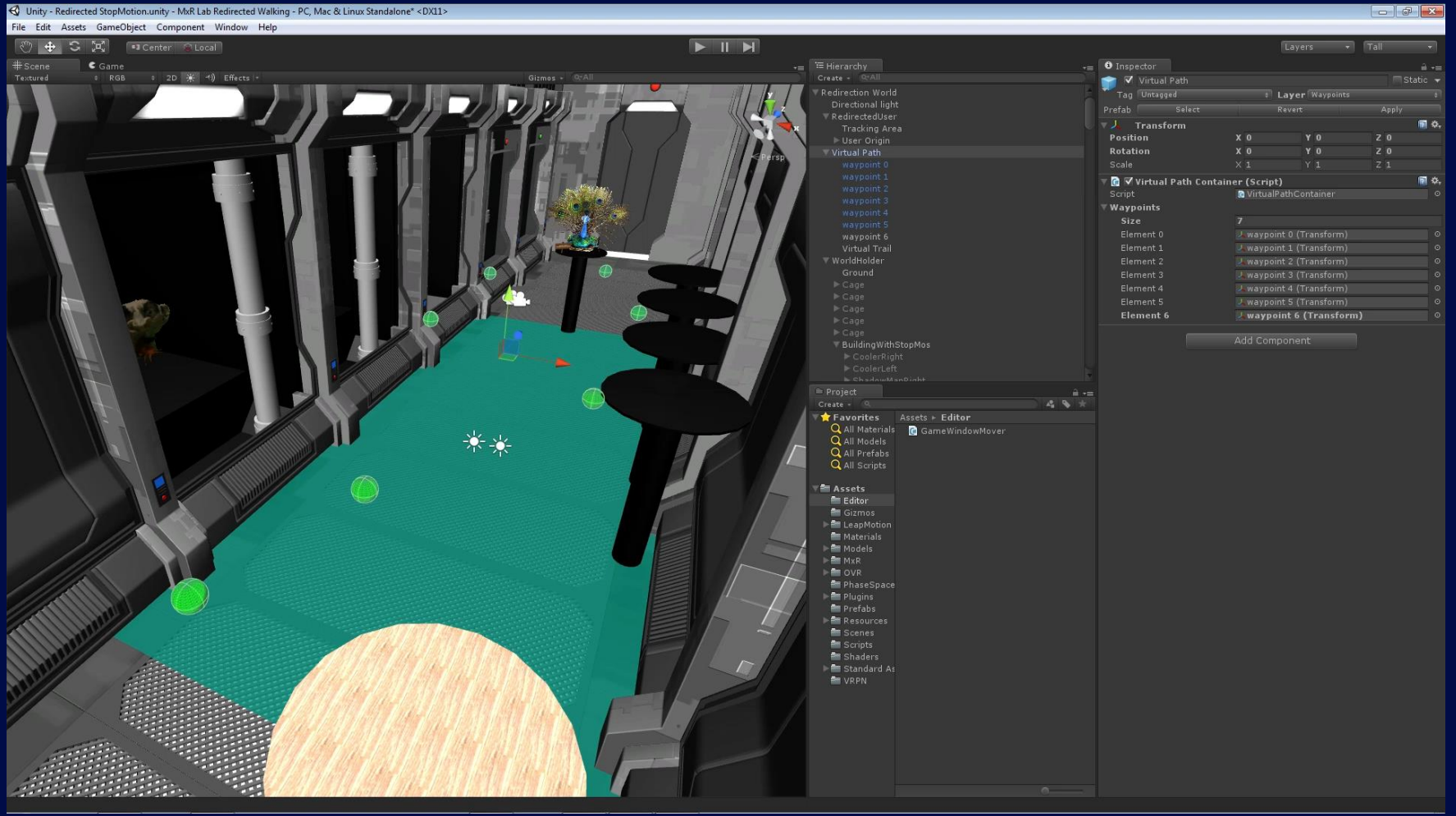



Suma, *Experimental Evaluation of the Cognitive Effects of Travel Technique in Immersive Virtual Environments*, Advisor: Dr. Larry F. Hodges

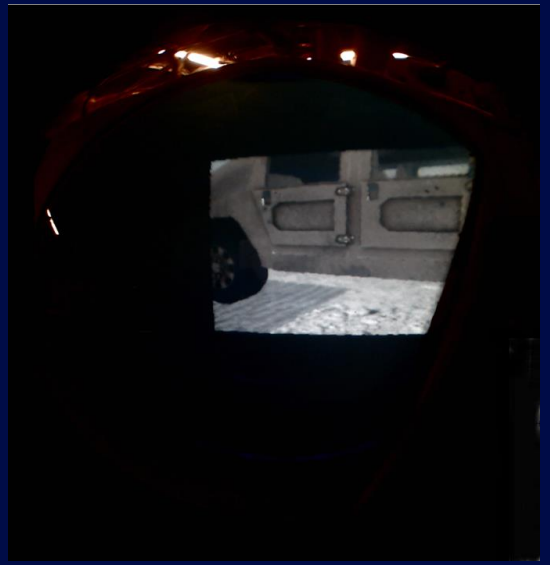


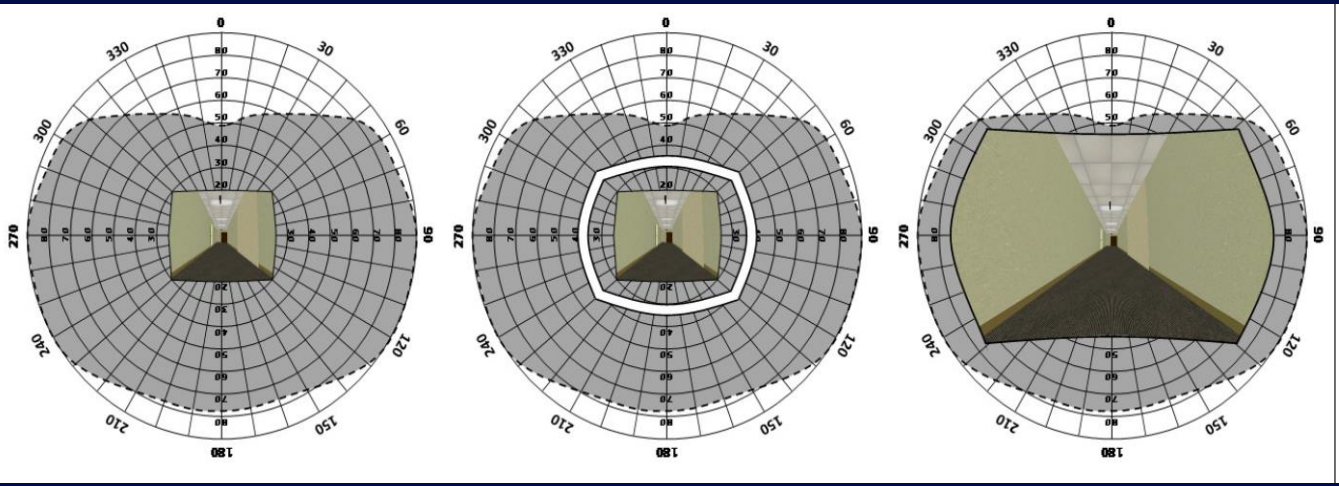


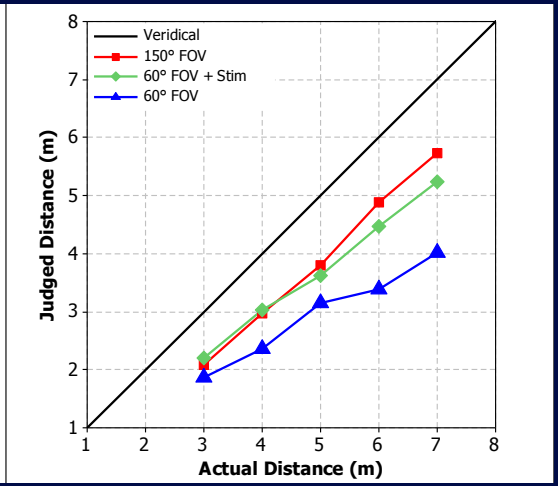
Sharif Razzaque, Zachariah Kohn, Mary C. Whitton, *Redirected Walking*, EUROGRAPHICS 2001


Unity - Redirected StopMotion.unity - MxR Lab Redirected Walking - PC, Mac & Linux Standalone* <DX11>
File Edit Assets GameObject Component Window Help
Scene
Game
Hierarchy
Redirection World
Directional light
RedirectedUser
Tracking Area
User Origin
Virtual Path
waypoint 0
waypoint 1
waypoint 2
waypoint 3
waypoint 4
waypoint 5
waypoint 6
Virtual Trail
WorldHolder
Ground
Cage
Cage
Cage
Cage
Cage
BuildingWithStopMos
CoolerRight
CoolerLeft
Inspector
Virtual Path
Static
Tag Untagged
Layer Waypoints
Prefab Select Revert Apply
Transform
Position X 0 Y 0 Z 0
Rotation X 0 Y 0 Z 0
Scale X 1 Y 1 Z 1
Virtual Path Container (Script)
Script VirtualPathContainer
Waypoints
Size 7
Element 0 waypoint 0 (Transform)
Element 1 waypoint 1 (Transform)
Element 2 waypoint 2 (Transform)
Element 3 waypoint 3 (Transform)
Element 4 waypoint 4 (Transform)
Element 5 waypoint 5 (Transform)
Element 6 waypoint 6 (Transform)
Add Component
Project
Favorites
All Materials
All Models
All Prefabs
All Scripts
Assets
Editor
Gizmos
LeapMotion
Materials
Models
MxR
OVR
PhaseSpace
Plugins
Prefabs
Resources
Scenes
Scripts
Shaders
VRPN
Assets ▸ Editor
GameWindowMover

Evan Suma






Veridical
150° FOV
60° FOV + Stim
60° FOV
Judged Distance (m)
Actual Distance (m)
1
2
3
4
5
6
7
8

Adam Jones

Ari Shapiro

Todd Richmond

Todd Richmond

Asymmetric reality
We bring our own offices with us
We do not see the same space

Reality is just a color on our palette

Interfaces go both ways
In vr it is clear that we are made to ‘feel’
This can make the interface use all of us
Joyful, satisfying, engaging, pleasurable

While industry is largely looking at vr as being one way
I ask that we see it as like the rope, two ways
Our whole bodies are involved

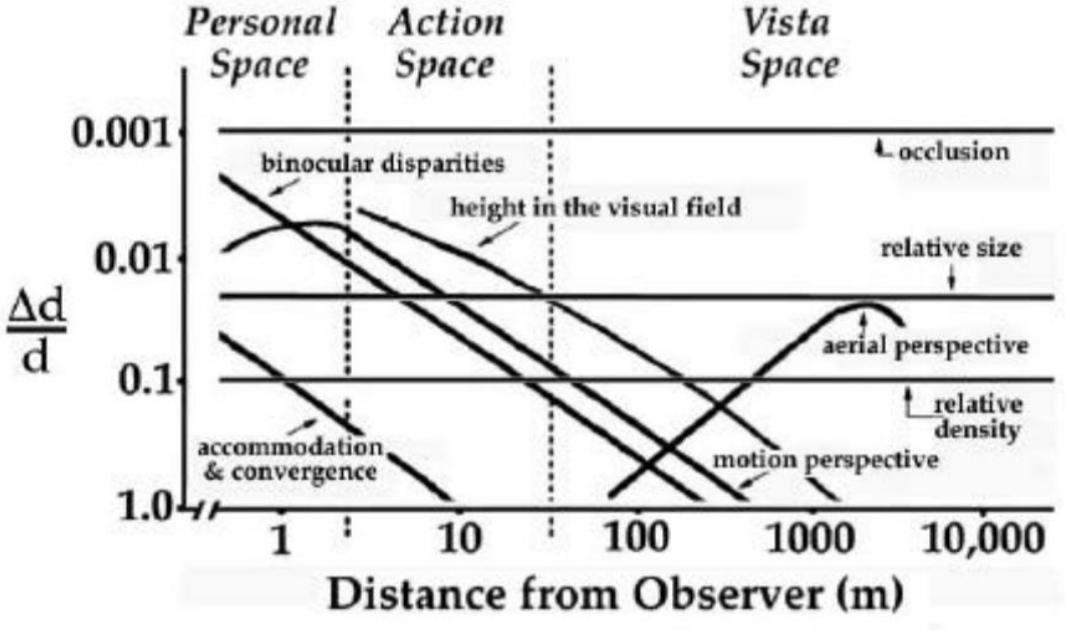
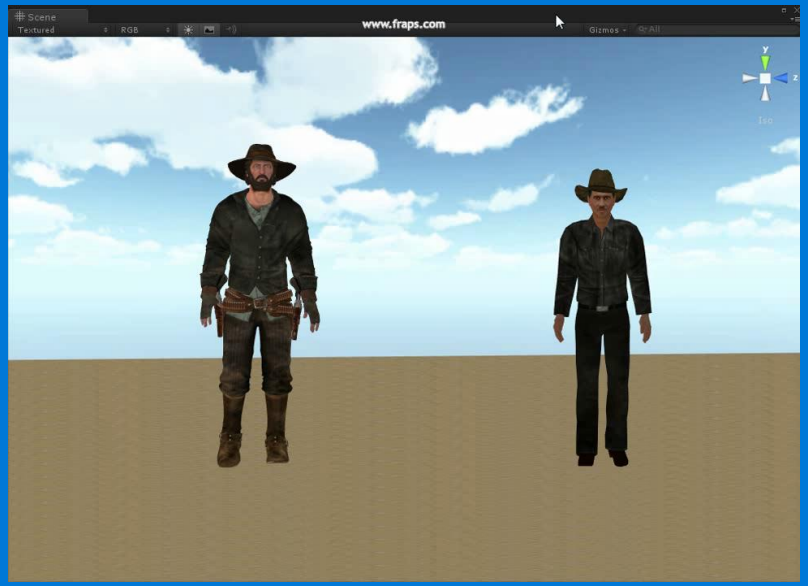
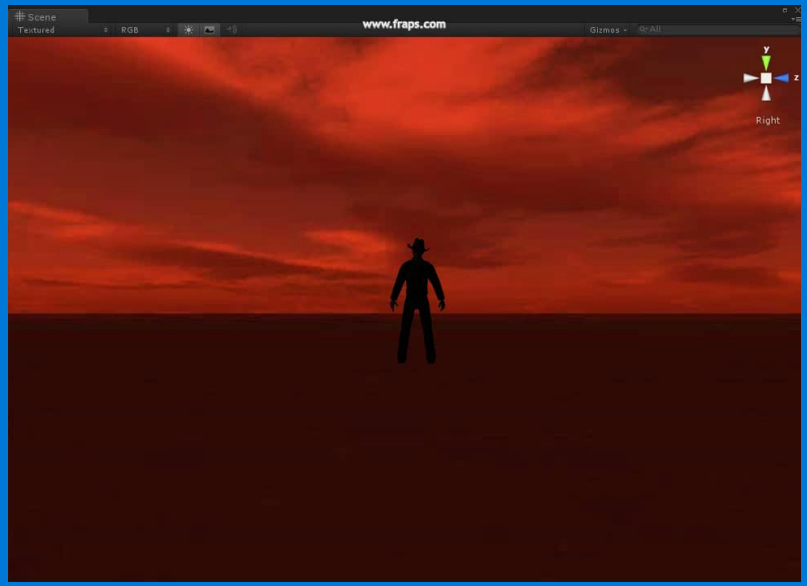


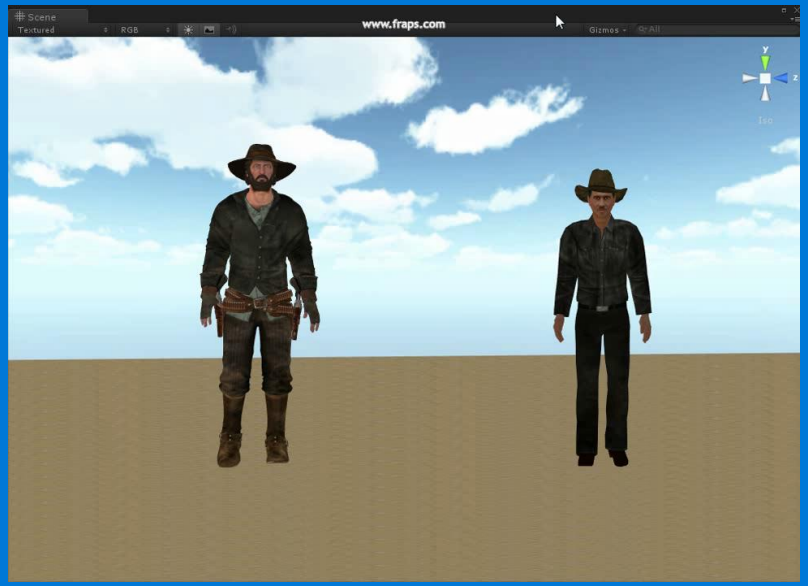
Figure 1: Just-discriminable depth thresholds as a function of the log of distance from the observer, from 0.5 to 5000 meters, for nine different sources of information about layout. Such plots were originated by Nagata (1981), and are extensively modified and elaborated here; they are plotted with analogy to contrast sensitivity functions. Our assumption is that more potent sources of information are associated with smaller depth-discrimination thresholds; and that these threshold functions reflect suprathreshold utility. These functions, in turn, delimit three types of space around the moving observer—personal space, action space, and vista space—each served by different sources of information, with different weights. This array of functions, however, is idealized; Figure 2 shows variations on the themes shown here.

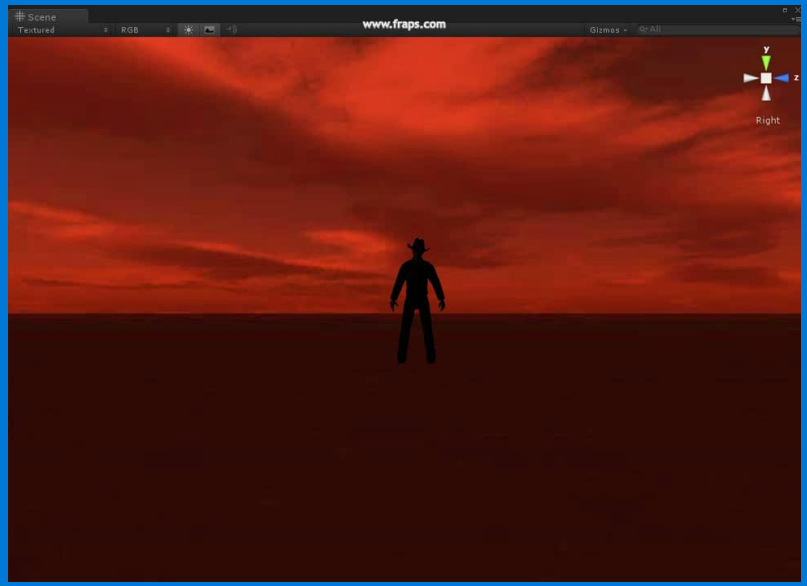
Body Language Matters



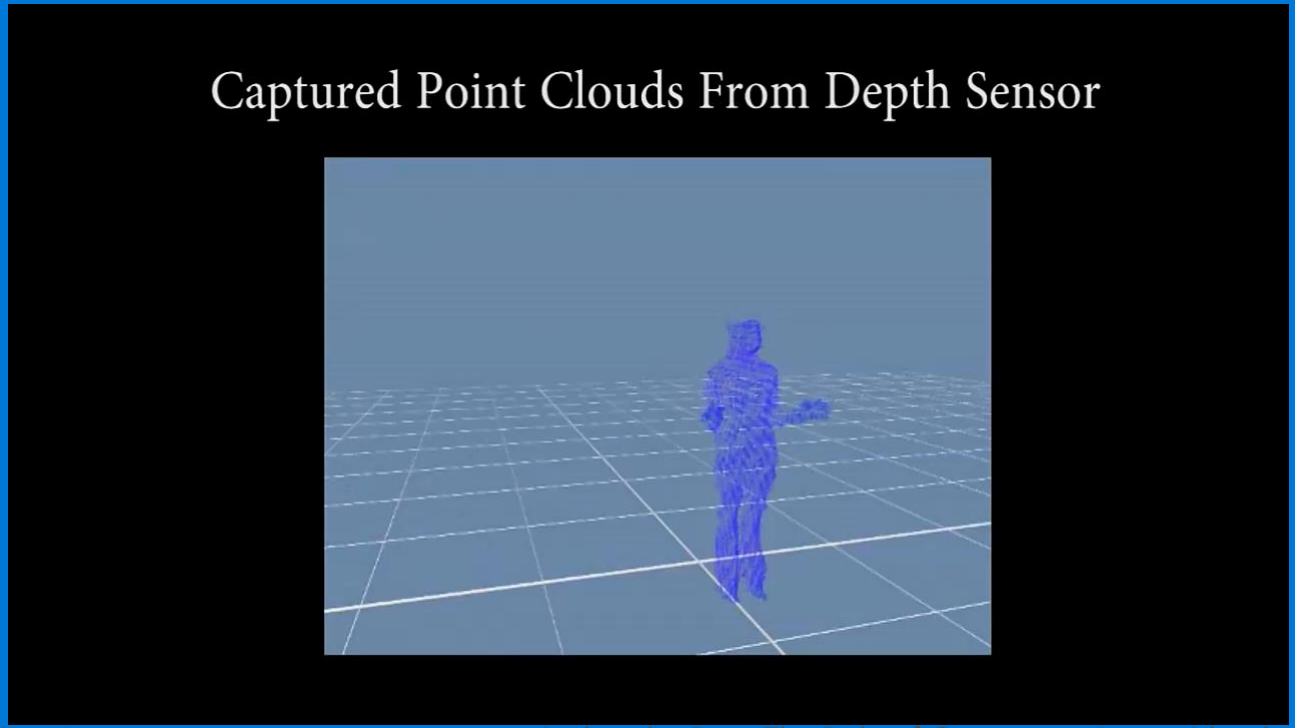


Body Language Matters





Rapid Capture, Control, and Display of Person-Specific Avatars



"I know who that is."

Acting the Part: The Role of Gesture on Avatar Identity. Feng, Shapiro.

Creating Near-Field VR Using Stop Motion Characters and a Touch of Light-Field Rendering

Mark Bolas*, Ashok Kuruvilla, Shravani Chintalapudi, Fernando Rabelo, Vangelis Lympouridis, Christine Barron, Evan Suma, Catalina Matamoros, Cristina Brous, Alicja Jasina, Yawen Zheng, Andrew Jones, Paul Debevec, David Krum
University of Southern California

Figure 1: *(a) Capture system. (b) Sequence of matted images. (c) Billboarded animation reacting to hand interaction.*

1 Introduction and Motivation

There is rapidly growing interest in the creation of rendered environments and content for tracked head-mounted stereoscopic displays for virtual reality. Currently, the most popular approaches include polygonal environments created with game engines, as well as 360 degree spherical cameras used to capture live action video. These tools were not originally designed to leverage the more complex visual cues available in VR when users laterally shift viewpoints, manually interact with models, and employ stereoscopic vision. There is a need for a fresh look at graphics techniques that can capitalize upon the unique affordances that make VR so compelling.

We have created a new type of VR experience and pipeline by using image-based rendering techniques to capture and render handcrafted stop motion puppets that are rich with micro geometry and detailed surfaces and materials. Our pipeline is computationally lightweight and integrated within the Unity3D game engine, thus allowing artists to quickly create lush models that can be plugged into a traditional virtual environment running in real time. Experientially, we found that these rich renderings create a magical place with a strong and immediate feeling of presence, which we call Near-Field VR. In particular, our ability to represent specular reflections, iridescence, sub-surface scattering, transparency, and fine geometric details correlates with findings in perceptual psychology regarding the hierarchy of cues that users employ to make sense of the Near-Field space [Cutting and Vishton 1995].

2 Approach

A custom-designed turntable captures a ring of images at 1° increments for each frame of an animated sequence of a stop motion puppet. Because the animator must now consider depth when manipulating poses, check images at 0 and 90 degrees are sent to Dragonframe, allowing time scrubbing and transparency. Primatte was used to matte images. As the user moves, we compute each eye's view angle and project the nearest captured view onto two co-axial billboarded polygons. Bounded by loading time, we only load and cache as-needed images in the DXT5 format from an SSD using 1K textures to match the HMD's resolution.

*e-mail: bolas@well.com

We also explored horizontally rebinned pixels to correct perspective from close viewpoints; projection onto proxy geometry derived from stereo correspondence for better stereoscopy; and flowed fields with 5¾° spacing for rebinned animation.

Interestingly, the perception of organic shapes proved quite tolerant of the geometric errors of our simplest billboarding technique, which we preferred due to its fast performance. We leveraged the game engine to minimize artifacts by setting default height to that of the user, placing geometry to control model proximity, and matching captured lighting. Our artists used complex materials to leverage the near-field cues retained by our image-based approach, for example, glints move from eye to eye with head translation.

Figure 2: *(a) While rebinning preserves the geometry of the puppet's nose and chin upon close inspection (b) users did not overtly notice the errors with the original image. (c) Organic shapes and rich surface materials used for Near-Field cues in later models.*

While we will more fully implement vertical parallax, flowed fields, projection onto geometry, and relighting, it is important to pause and present our current results. Our research proves that Near-Field VR is a viscerally engaging type of VR that needs to be further developed; that the Near-Field region must transcend typical polygonal techniques and move toward the image-based cues the human perceptual system demands; and that VR greatly benefits when it finds ways to leverage traditional craft.

References

CUTTING, J.E., AND VISHTON, P.M. 1995. Perceiving layout: The integration, relative dominance, and contextual use of different information about depth. In *Epstein, W., & S. Rogers (Eds.), Handbook of Perception and Cognition*, vol. 5, Perception of Space and Motion, NY: Academic Press.

In 2025, while flipping through the *elective courses* section of the Prospectus, we found something about *sound design* and *comedy*

The Prompt is: "In 2025, while flipping through the elective courses section of the Prospectus, we found something about sound design and comedy"
————————

The Prompt is also:
————————

```
<div class="phrase"><br />
In 2025, while flipping through the <br />
<span class="word">Elective Courses<br />
</span><br />
section of the Prospectus, we found something about<br />
<span class="word">Sound Design<br />
</span><br />
and <br />
<span class="word">Comedy<br />
</span> <br />
<br /><br /><br />
</div><br />
```

————————

Which are _words_, nice little _words_ that can fit in nice little _boxes_, which is what my reply, my feedback, my passion is supposed to find a way to _fit-in_ because if it does not _fit-in_ a _box_, a _text_box_ with !tone-of-voice and !body-language then it must-be the 9-5 time-card-punch-a-clock thinking of the 1900s and *surely* the way -forward- is to go *further back*, so lets skip this whole gotta-fit-in-a-box-with-logical-tags step, stop punching the box's cards, and get [back to people] and what makes [people], well, [people], and speaking in logical-tongues (ooh, thank you [logical-box] for correcting my spelling of tongues, so thought full of you (ha! I snuck :thought full: past the {bastard box}) of html-thinking, let us figure out how to be people again, but people that are surrounded by an accelerated flow of information that could, possibly, advance the lot of us, but instead, probably, certainly today, is simply distracting my eyes while confusing my heart.

In 2025, while flipping through the *elective courses* section of the Prospectus, we found something about *sound design* and *screenwriting*

The Prompt is: "In 2025, while flipping through the elective courses section of the Prospectus, we found something about sound design and comedy"

The Prompt is also:

In 2025, while flipping through the

elective courses

section of the Prospectus, we found something about

sound design

and

comedy

Which are _words_, nice little _words_ that can fit in nice little _boxes_, which is what my reply, my feedback, my passion is supposed to find a way to _fit-in_ because if it does not _fit-in_ a _box_, a _text_box_ with !tone-of-voice and !body-language then it must-be the 9-5 time-card-punch-a-clock thinking of the 1900s and *surely* the way -forward- is to go *further back*, so lets skip this whole gotta-fit-in-a-box-with-logical-tags step, stop punching the box's cards, and get [back to people] and what makes [people], well, [people], and speaking in logical-tongues (ooh, thank you [logical-box] for correcting my spelling of tongues, so thought full of you (ha! I snuck :thought full: past the (bastard box)) of html-thinking, let us figure out how to be people again, but people that are surrounded by an accelerated flow of information that could, possibly, advance the lot of us, but instead, probably, certainly today, is simply distracting my eyes while confusing my heart.

Preferred text layout (no bullets)

Main topic 1: Size 28 pt, Segoe UI Light

Size 20, Segoe UI for the subtopics

Size 20, Segoe UI for the subtopics

Main topic 2: Size 28 pt, Segoe UI Light

Size 20, Segoe UI for the subtopics

Size 20, Segoe UI for the subtopics

Main topic 3:Size 28 pt, Segoe UI Light

Microsoft

Title only – no graphics