The ABC of Mixed Reality Interactions

Hrvoje Benko
Facebook Reality Labs


Waiting for command
Aug 29
2.45 pm
Center of city

Command Line Interfaces (text based)	**Graphical User Interfaces** (mouse + keyboard)	**Natural User Interfaces** (touch/gestures, tablets, smartphones)	**Mixed Reality Interfaces**
1960s	**1980s**	**2000s**	**2020s**

Command Line Interfaces
(text based)

1960s

Graphical User Interfaces
(mouse + keyboard)

1980s

Natural User Interfaces
(touch/gestures, tablets, smartphones)

2000s

Mixed Reality Interfaces

2020s

1968
Engelbart & English
MOAD

1983
Microsoft Mouse / Apple Mouse

Command Line Interfaces
(text based)

1960s

Graphical User Interfaces
(mouse + keyboard)

1980s

Natural User Interfaces
(touch/gestures, tablets, smartphones)

2000s

Mixed Reality Interfaces

2020s

1985
U of T - MultiTouch Tablet

2007
Apple iPhone

SK Lee, W. Buxton, and K. C. Smith. A multi-touch three dimensional touch-sensitive tablet. In Proc. of the CHI '85.

Command Line Interfaces
(text based)

1960s

Graphical User Interfaces
(mouse + keyboard)

1980s

Natural User Interfaces
(touch/gestures, tablets, smartphones)

2000s

Mixed Reality Interfaces

2020s

???


New York 2001

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine* - *ACM ISAR 2001*

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine* - ACM ISAR 2001

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine - ACM ISAR 2001*

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine - ACM ISAR 2001*

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine - ACM ISAR 2001*

Bell B., Feiner, S., and Hollerer, T. *Columbia Touring Machine - ACM ISAR 2001*

Command Line Interfaces
(text based)

1960s

Graphical User Interfaces
(mouse + keyboard)

1980s

Natural User Interfaces
(touch/gestures, tablets, smartphones)

2000s

Mixed Reality Interfaces

2020s

2001
MARS - Columbia Touring Machine @ ISAR 2001

Michael Abrash

"Imagine AR glasses that are socially acceptable and all-day wearable, that give you useful virtual objects like your phone, your TV and virtual work spaces, that give you perceptual super powers, a context-aware personal assistant, and above all the ability to connect, share and collaborate with others anywhere, any time. If those glasses existed today, we'd all be wearing them right now."

Might be “obvious”…

… why not available today?

2001 vs. 2018 - Same Challenges

Display

Optics

Battery

Tracking/localization

Compute power

Spatial understanding

Spatial audio

Input/Interactions

2001 vs. 2018 - Same Challenges

Display

Optics

Battery

Tracking/localization

Compute power

Spatial understanding

Spatial audio

Input/Interactions

What kind of interactions will define the MR era?

Compelling MR interactions will be *adaptive,* *believable,* and *computational*.

Compelling MR interactions will be ***adaptive***, *believable*, and *computational*.

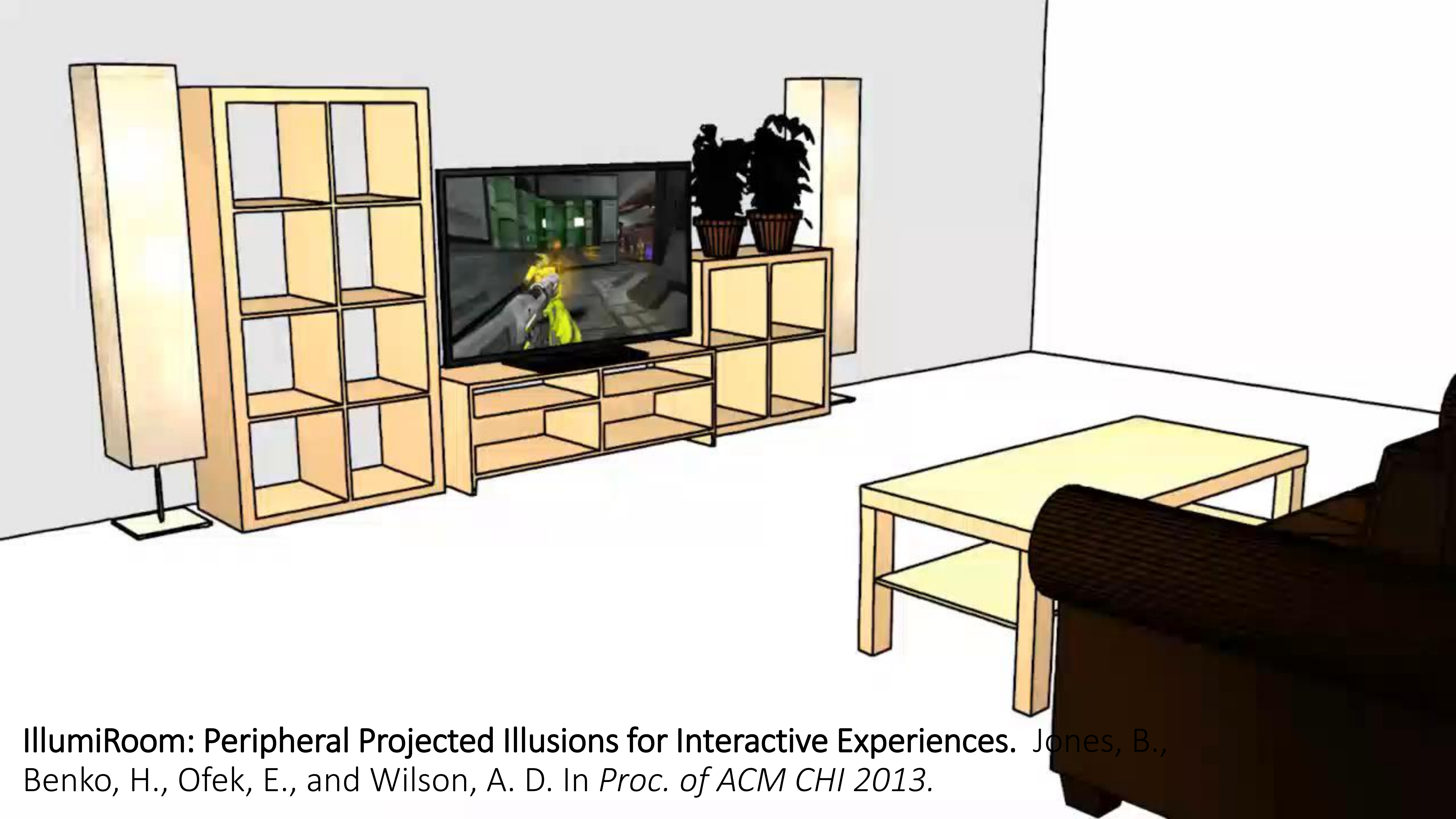
Magic of MR interactions happens when they are tightly coupled to the user’s environment.

Projection AR

+

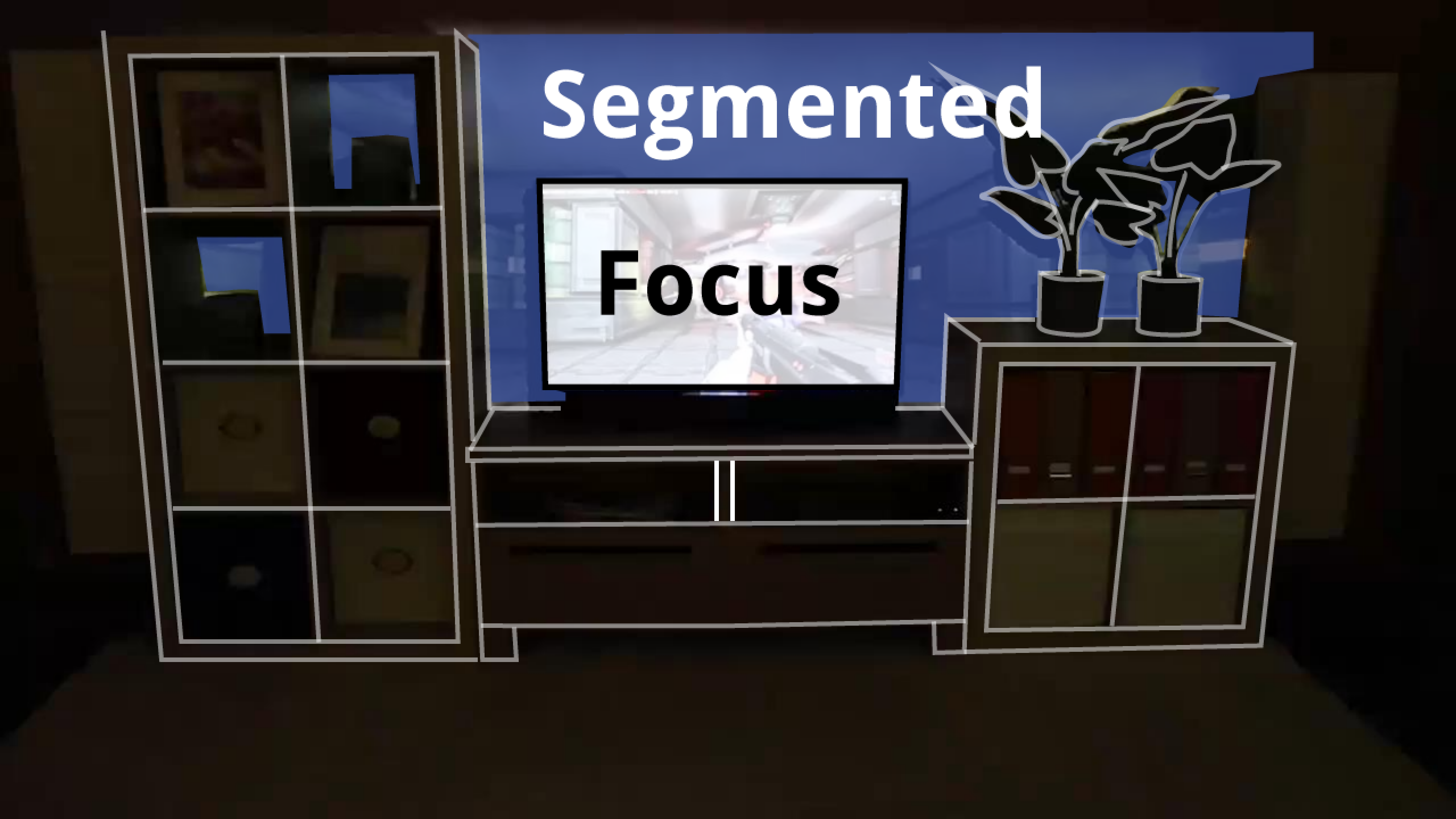
Projector

Depth Camera
(Kinect)

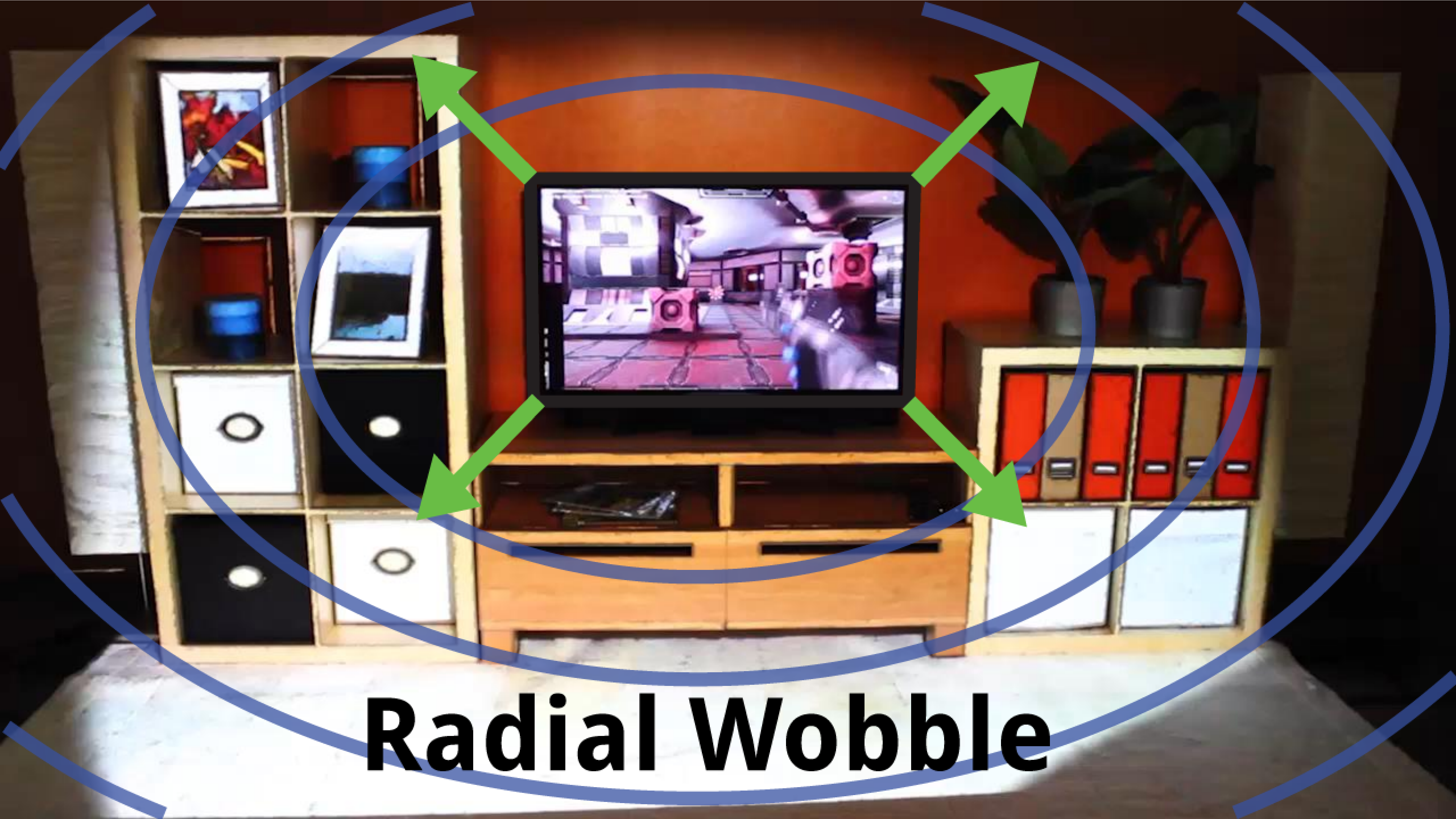


IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. Jones, B., Benko, H., Ofek, E., and Wilson, A. D. In *Proc. of ACM CHI 2013.*


Focus
Context Full


Segmented
Focus


Appearance


Radial Wobble


Lighting

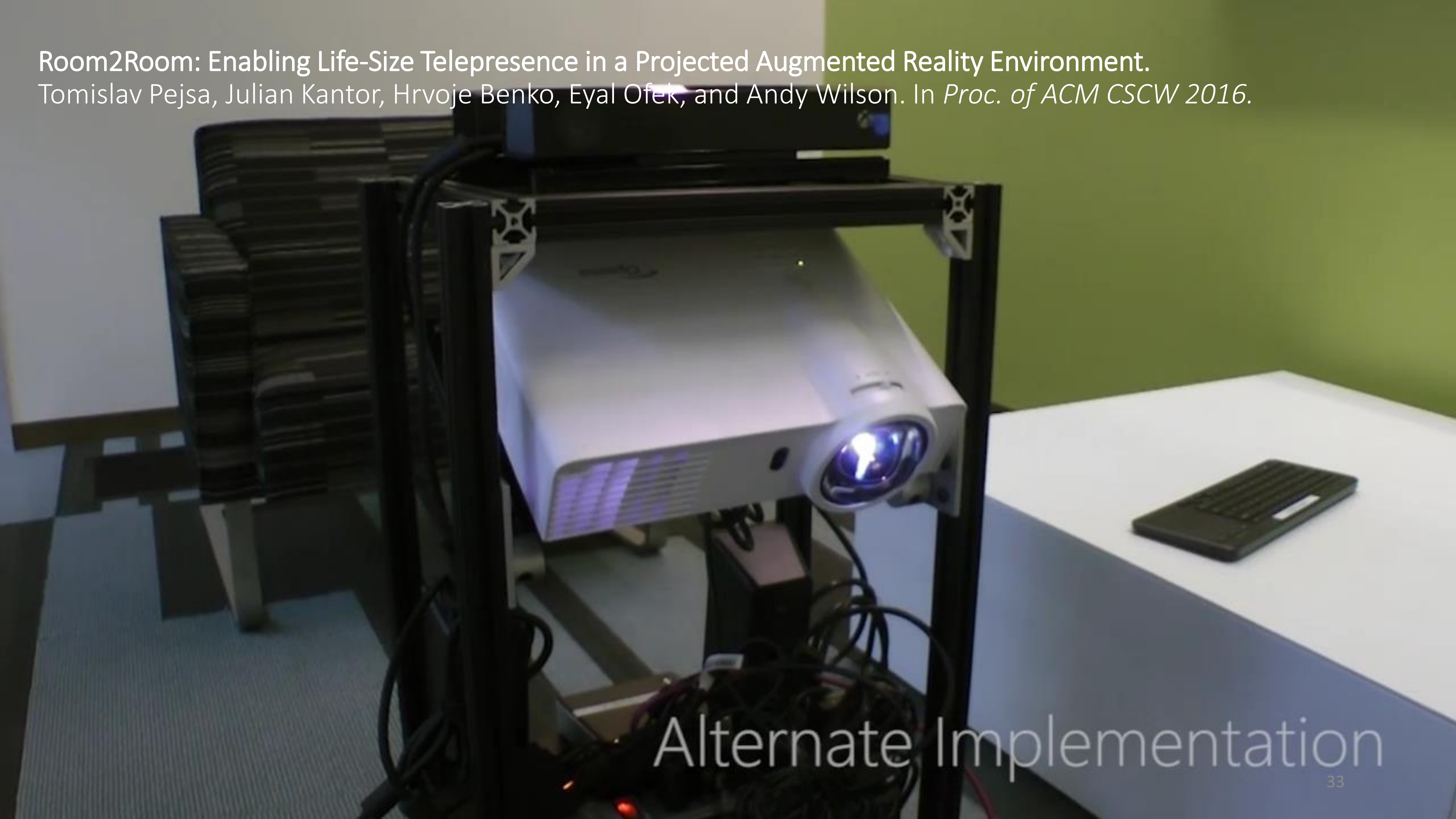
Steerable Augmented Reality with the Beamatron

Andy Wilson, Hrvoje Benko, Shahram Izadi and Otmar Hilliges
Microsoft Research

ACM UIST 2012


lenovo

Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment.
Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andy Wilson. In *Proc. of ACM CSCW 2016.*



Magic of MR interactions happens when they are tightly coupled to the user's ~~environment~~ context.

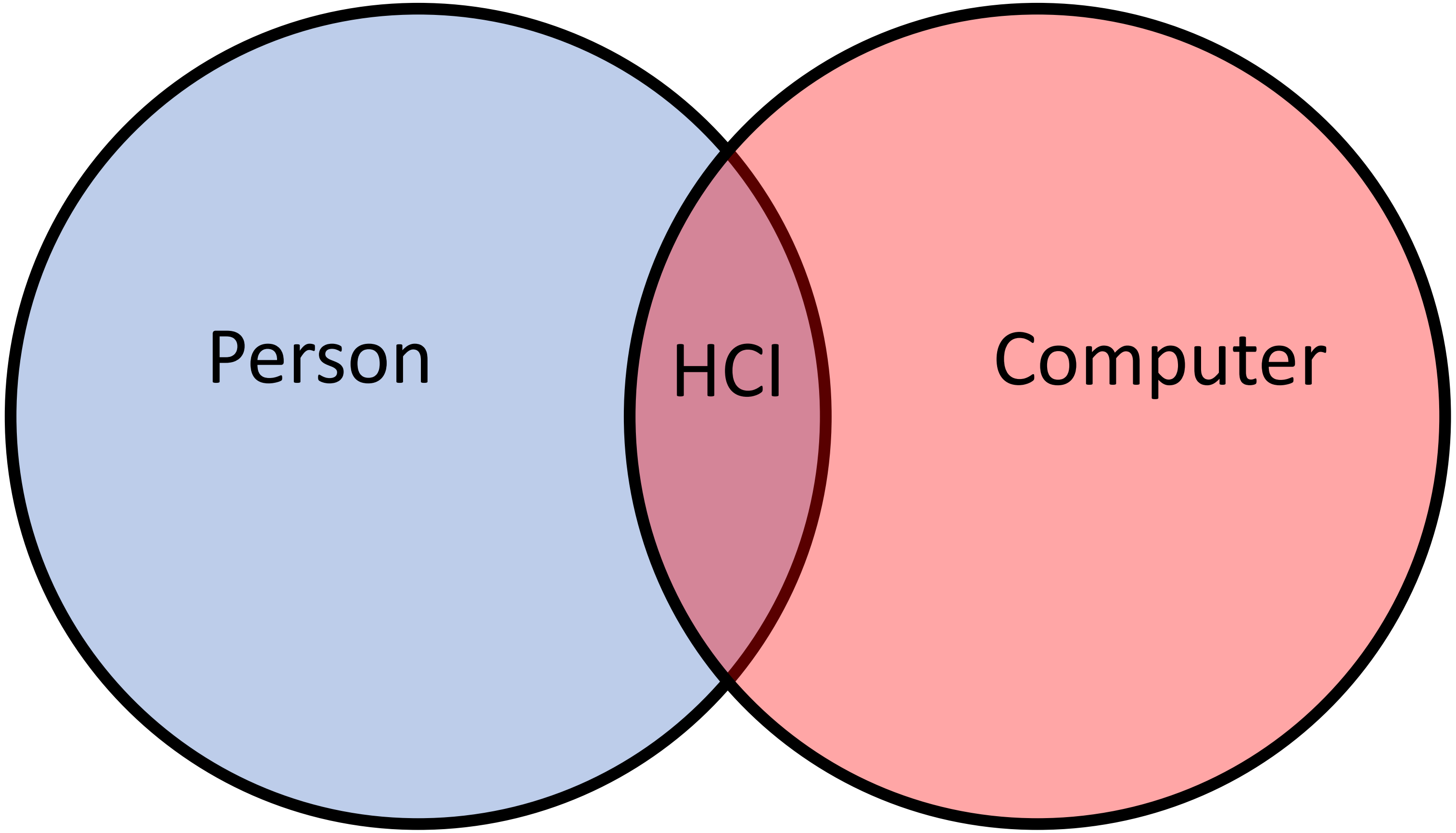
Context

environment
(e.g., space geometry, object semantics, people around)

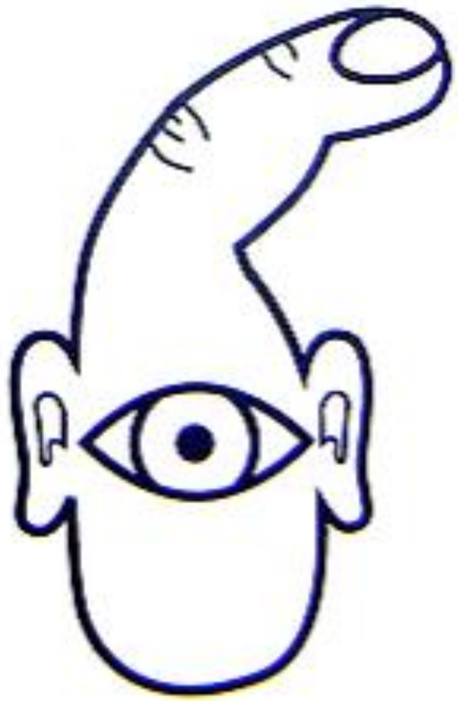
user actions
(e.g., gestures, body pose, bio-signals)

user's mental state
(e.g., emotional, mental load, cognitive, focus)

task
(e.g., communication, navigation, calendar)

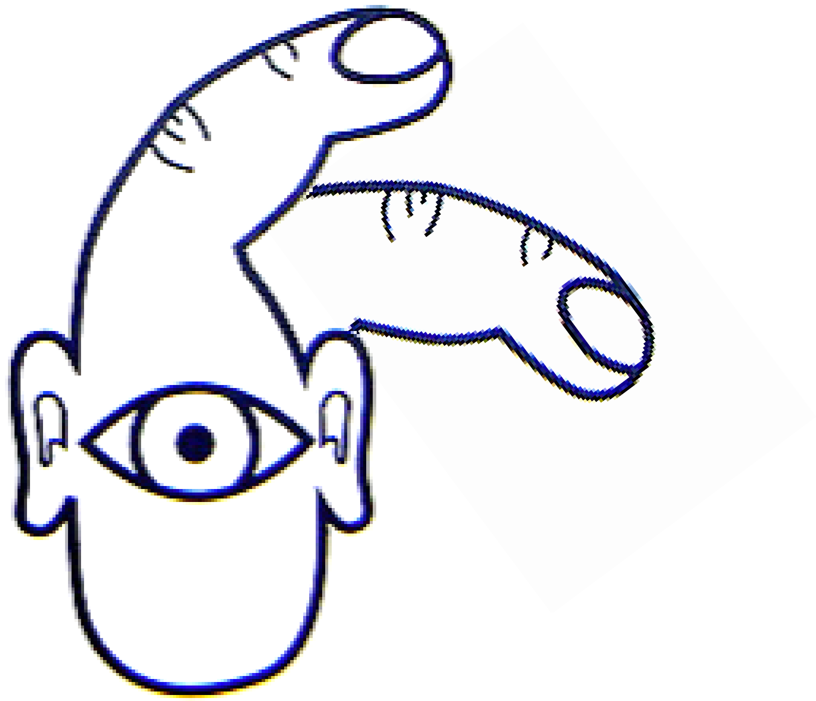

Person
HCI
Computer

How the computer sees us!

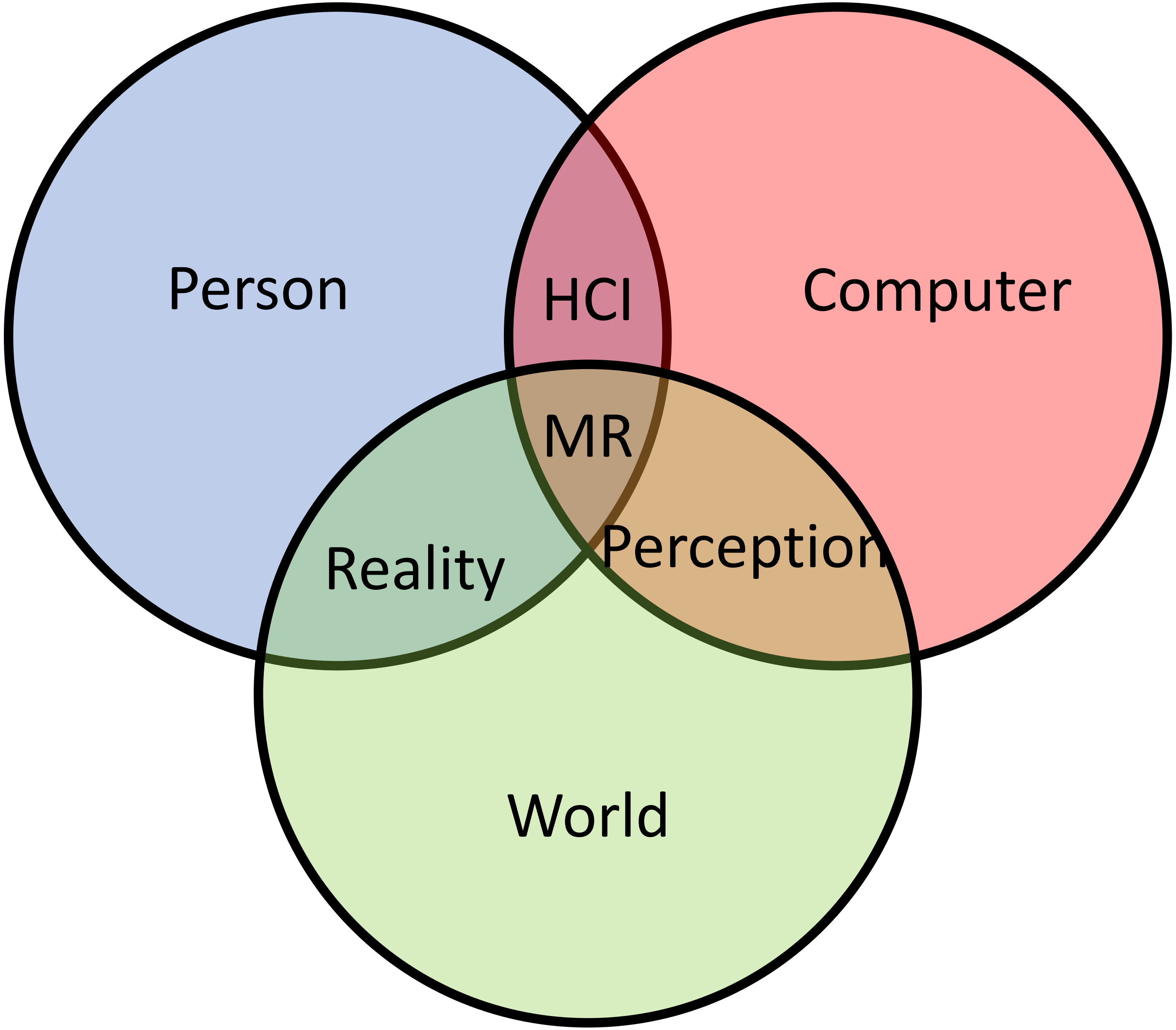


Tom Igoe and Dan O'Sullivan - *Physical Computing*. 2004.

How the phone sees us!

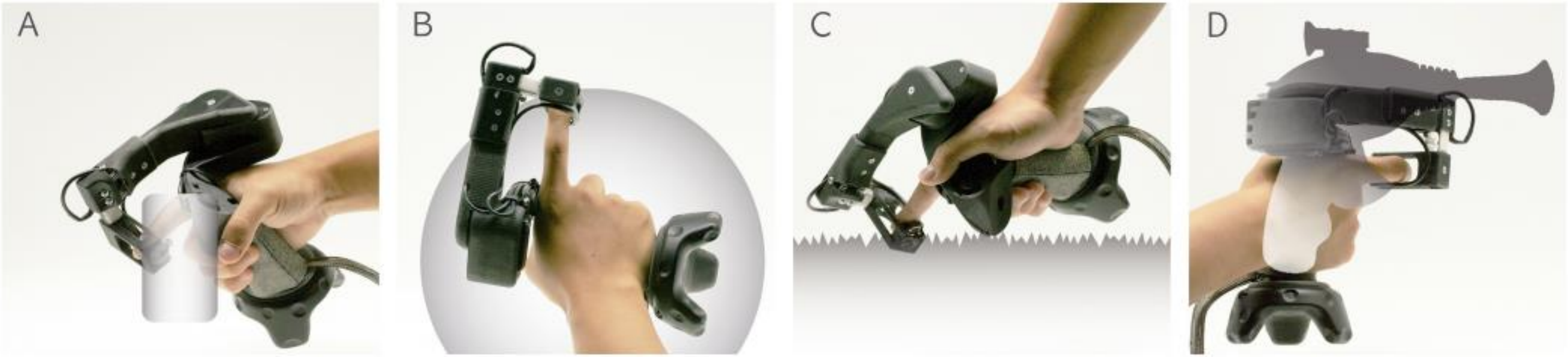


Modified Tom Igoe and Dan O’Sullivan - *Physical Computing*. 2004.


Person
HCI
Computer
MR
Reality
Perception
World

Example: Context-aware Tools

CLAW: *Multi-purpose* controller that *adapts* to the user's context of use.



Choi, I., Ofek, E., Benko, H., Sinclair, M. and Holz, C. CLAW: A Multifunctional Handheld Haptic Controller for Grasping, Touching, and Triggering in Virtual Reality. In *Proc. of ACM CHI '18*.

CLAW

Choi, I., Ofek, E., Benko, H., Sinclair, M. and Holz, C. CLAW: A Multifunctional Handheld Haptic Controller for Grasping, Touching, and Triggering in Virtual Reality. In *Proc. of ACM CHI '18*.


Servo Motor
Force Sensor
Voice Coil

Example: Hand interactions adapted to the environment

Shangchen Han, Beibei Liu, Robert Wang, Yuting Ye, Christopher D. Twigg, and Kenrick Kin. 2018. Online optical marker-based hand tracking with deep labels. ACM Trans. Graph. 37, 4, Article 166 (July 2018)

Accurate, Robust, and Flexible Real-time Hand Tracking

Toby Sharp[†] Cem Keskin[†] Duncan Robertson[†] Jonathan Taylor[†] Jamie Shotton[†]
David Kim Christoph Rhemann Ido Leichter Alon Vinnikov Yichen Wei
Daniel Freedman Pushmeet Kohli Eyal Krupka Andrew Fitzgibbon[*] Shahram Izadi[*]

Microsoft Research

Figure 1: We present a new system for tracking the detailed motion of a user's hand using only a commodity depth camera. Our system can accurately reconstruct the complex articulated pose of the hand, whilst being robust to tracking failure, and supporting flexible setups such as tracking at large distances and over-the-shoulder camera placement.

ABSTRACT

We present a new real-time hand tracking system based on a single depth camera. The system can *accurately* reconstruct complex hand poses across a variety of subjects. It also allows for *robust* tracking, rapidly recovering from any temporary failures. Most uniquely, our tracker is highly *flexible*, dra-

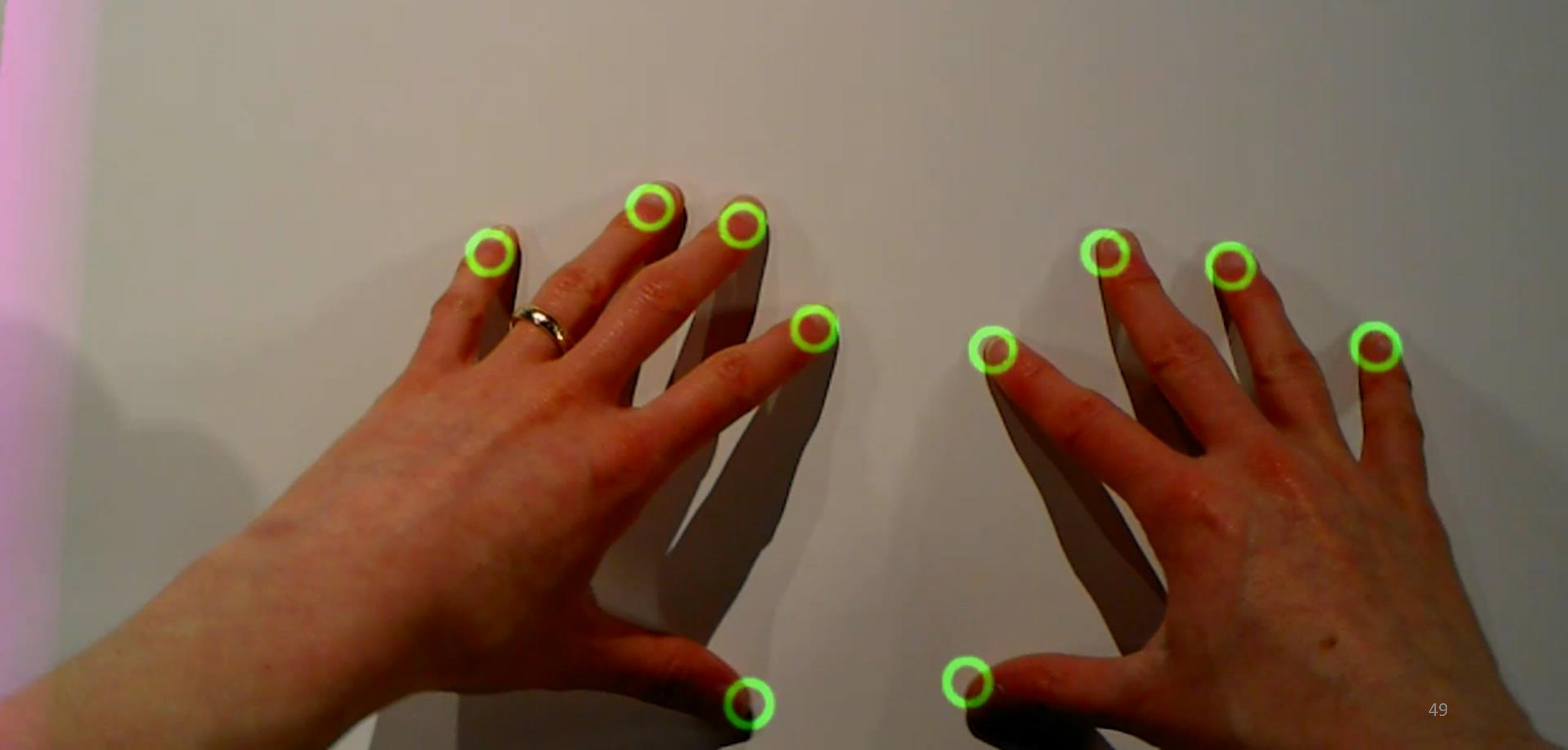
the user's hand with gloves or markers can be cumbersome and inaccurate. Much recent effort, including this work, has thus focused on camera-based systems. However, cameras, even modern consumer depth cameras, pose further difficulties: the fingers can be hard to disambiguate visually and are often occluded by other parts of the hand. Even state of the

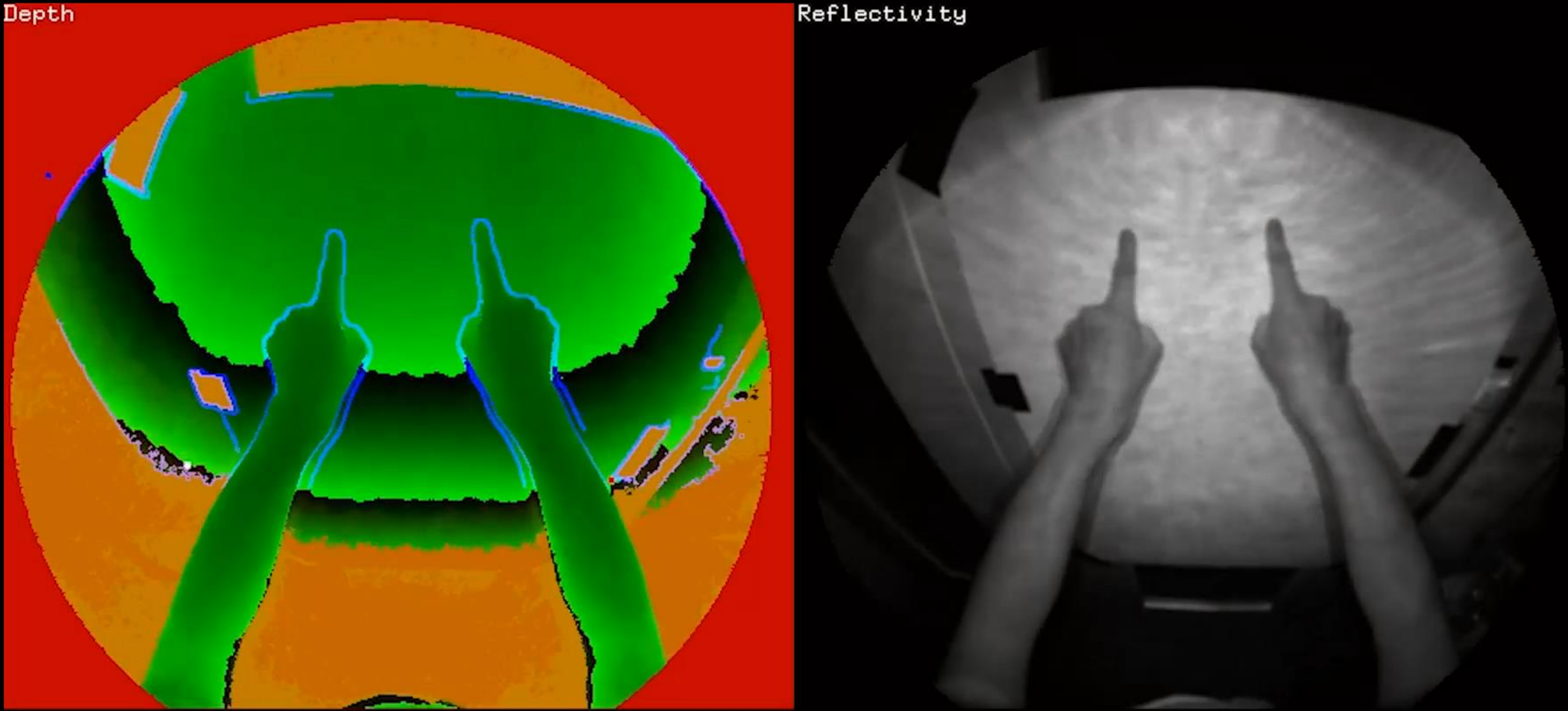
One small problem…

People don’t generally interact in mid air!

People interact with objects and on surfaces!

Hand interaction context = any available surface





MRTouch: Adding Touch to Head Mounted Mixed Reality.
Robert Xiao, Julia Schwarz, Nick Throm, Andrew D. Wilson, Hrvoje Benko. IEEE TVCG 2018 (Vol 24, No 4, April 2018)

Sample Interactions

(Continuous shot, first and only take)

FYI....

Hololens Research Mode API

Depth and reflectivity data now part of the public Research Mode API (developer mode only)

- https://github.com/Microsoft/HoloLensForCV
- Possible now to implement MRTouch using only public APIs

Compelling MR interactions will be *adaptive*, ***believable***, and *computational*.

In MR, we are obsessed with creating a rich sense of reality!

GENERATIONS / VANCOUVER 12-16 AUGUST
SIGGRAPH2018

Deep Appearance Models for Facial Rendering

STEPHEN LOMBARDI, JASON SARAGIH, TOMAS SIMON, YASER SHEIKH
Facebook Reality Labs

For interactions, realistic is not always better

Believable $\neq$ Realistic

Believable =
consistent with user’s
expectations + non surprising

In AR/VR we want to induce the suspension of disbelief, whereby users suspend their critical faculties (i.e., sacrifice logic and realism) to believe the unbelievable.

Haptic Revolver

Touch, Shear, Texture, and Shape Rendering on a Reconfigurable VR Controller

Eric Whitmire[1], Hrvoje Benko[2], Christian Holz[2],
Eyal Ofek[2], Mike Sinclair[2]

[1]Paul G. Allen School, University of Washington
[2]Microsoft Research, Redmond

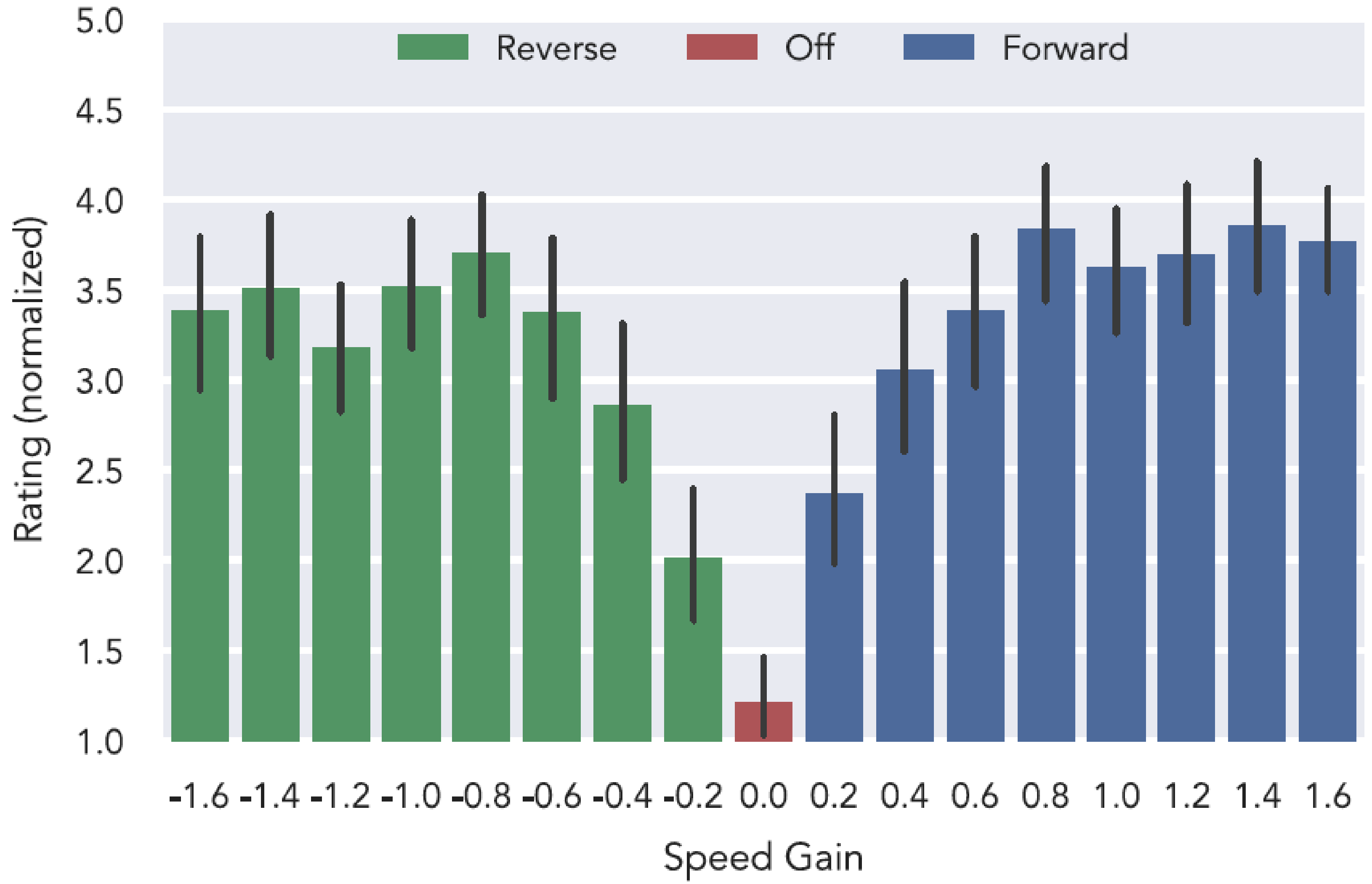
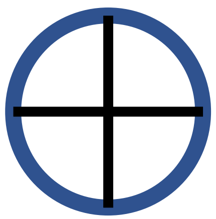


Figure 10. Results of the first user study showing mean realism ratings across participants as a function of the wheel speed gain. The error bars show a 95% confidence interval. A negative gain indicates the wheel was spun in the opposite direction.

Hand Movement

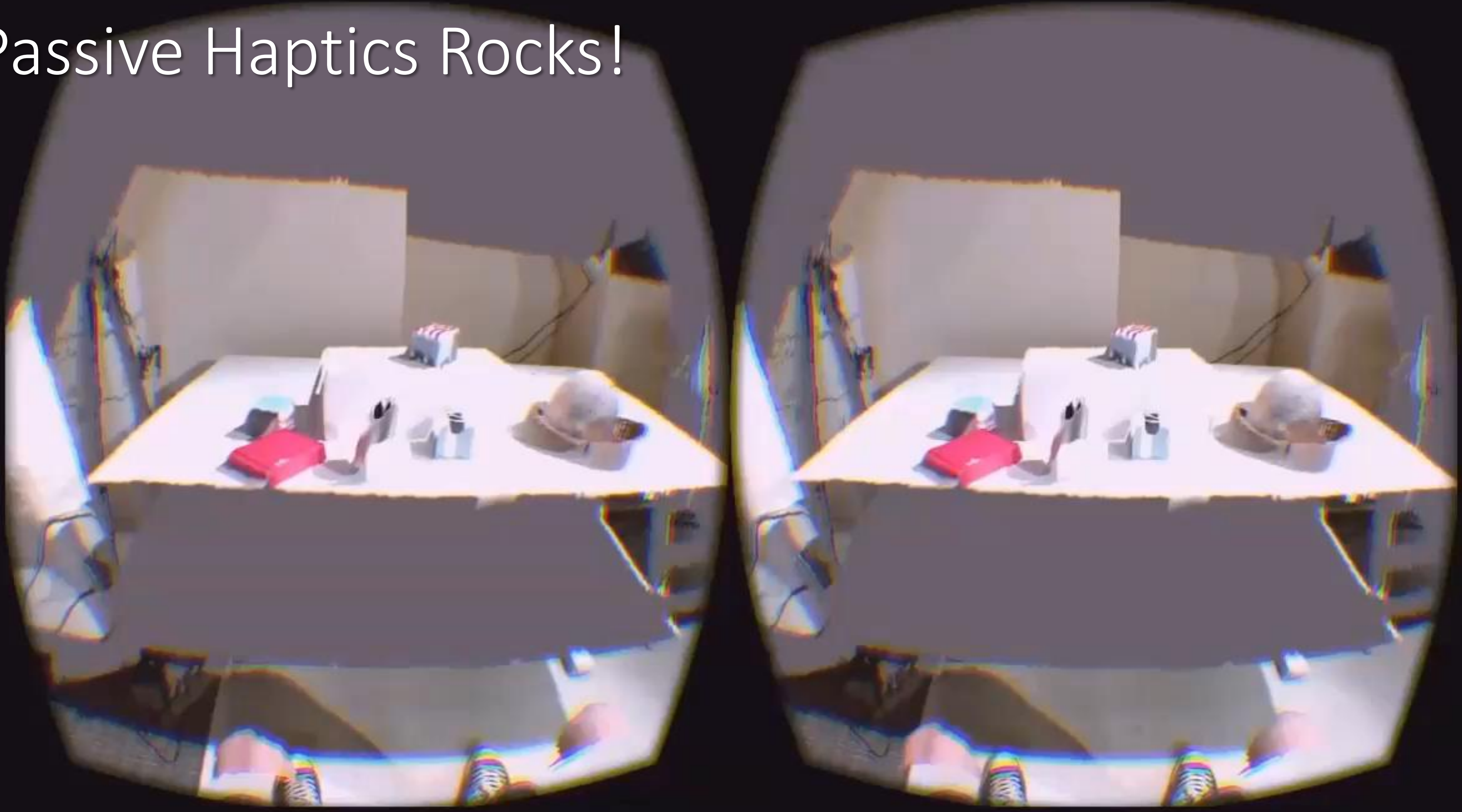
Wheel Movement



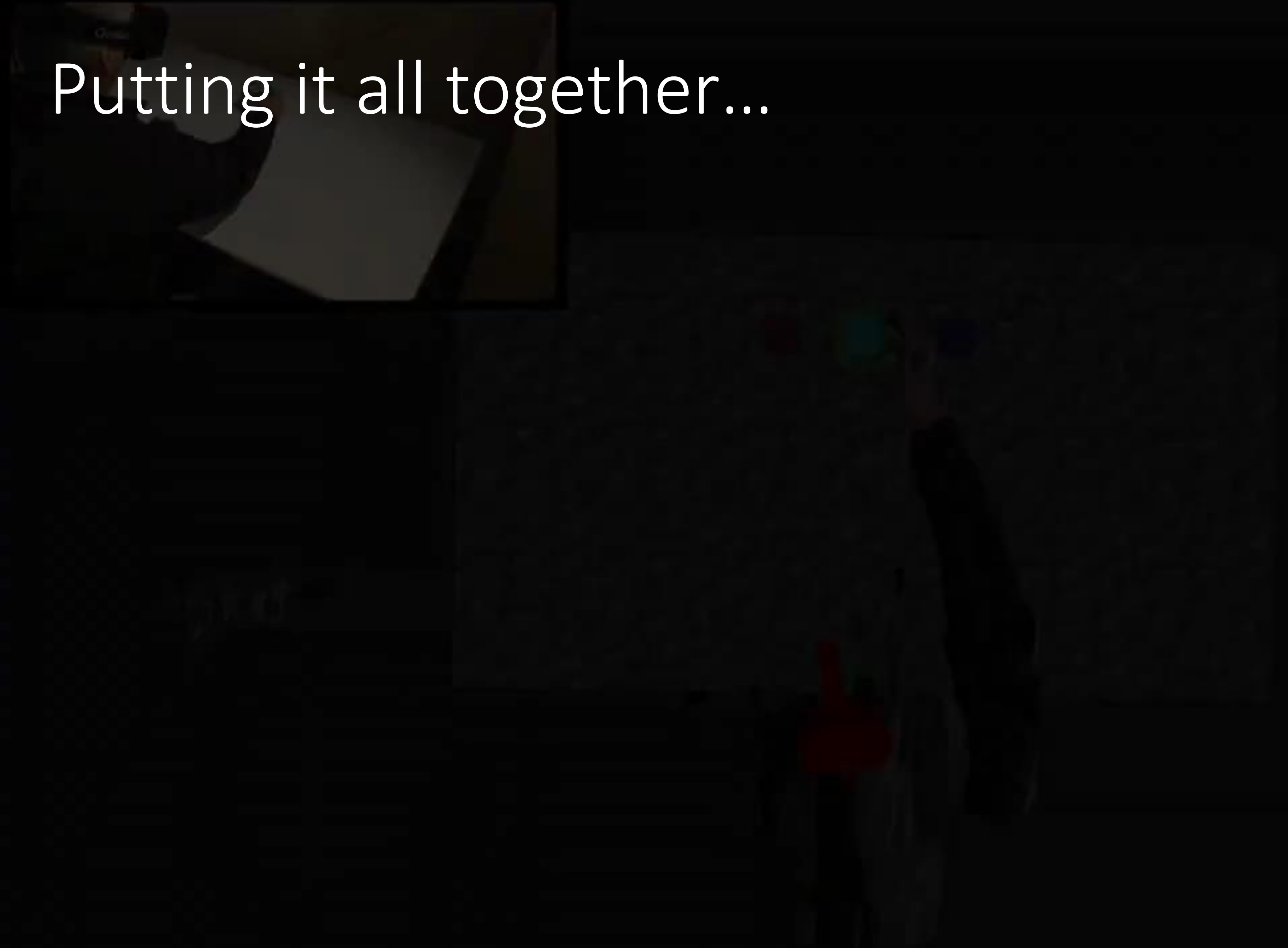
Haptic Retargeting

Azmandian, M., Hancock, M., Benko, H., Ofek, E., and Wilson, A. *Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In Proc. of* ACM CHI 2016.

Passive Haptics Rocks!



Putting it all together…



Focusing on “as real as possible” designs can lead to sub-optimal MR experience.

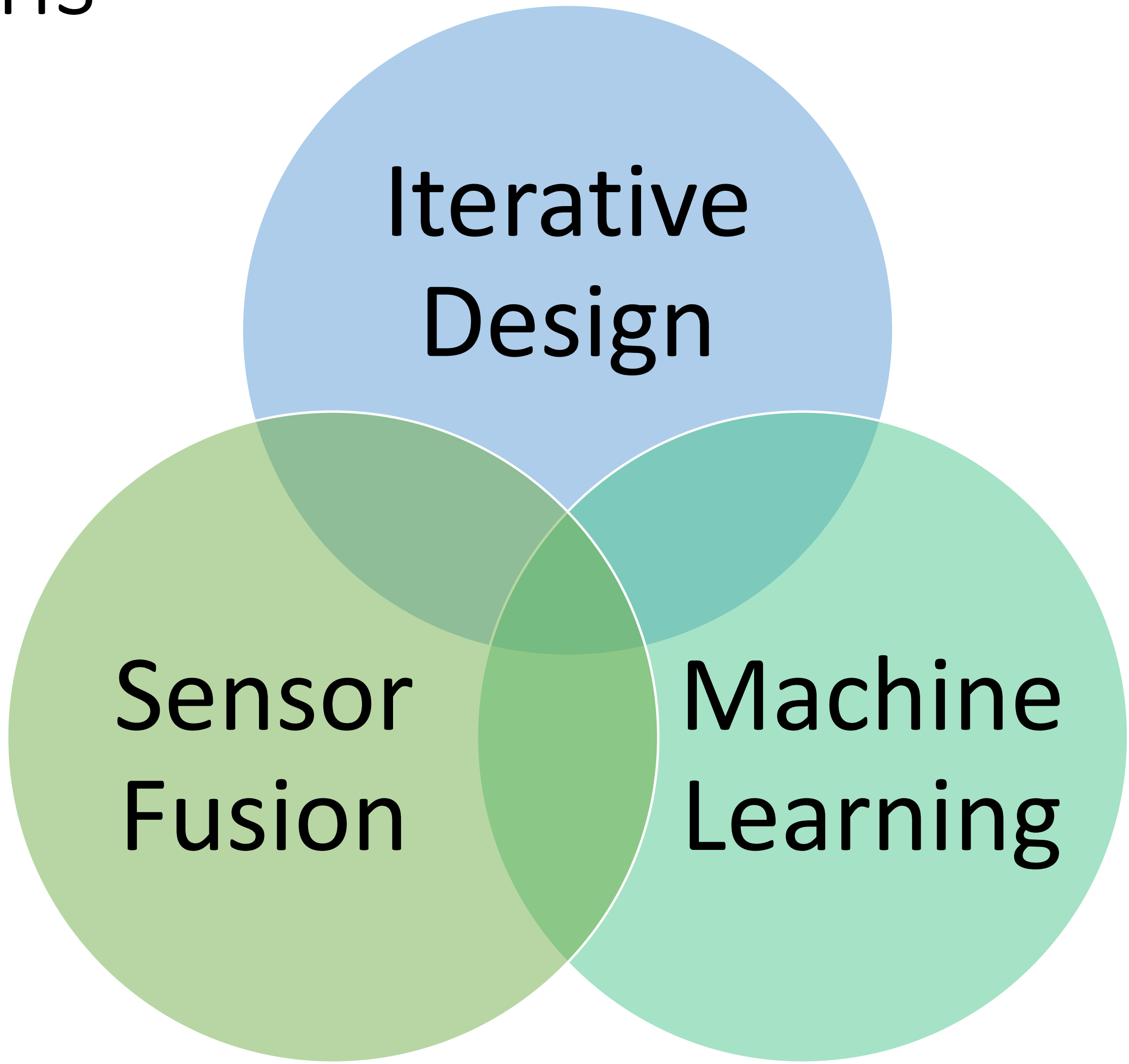
Design for BELIEVABILITY, not REALISM.

Compelling MR interactions will be *adaptive*, *believable*, and **computational**.

1960s	1980s	2000s	2020s
Command Line Interfaces (text based)	**Graphical User Interfaces** (mouse + keyboard)	**Natural User Interfaces** (touch/gestures, tablets, smartphones)	**Mixed Reality Interfaces**

Location fixed	⟷	Mobile
Explicit (command driven)	⟷	Implicit
Sensing poor	⟷	Sensing rich
Single modality	⟷	Multimodal
Dedicated/cognitively capturing	⟷	Glanceable/always-on
Precise and accurate inputs	⟷	Imprecise and noisy inputs

AR Interactions



Can you type on a phone touch keyboard?

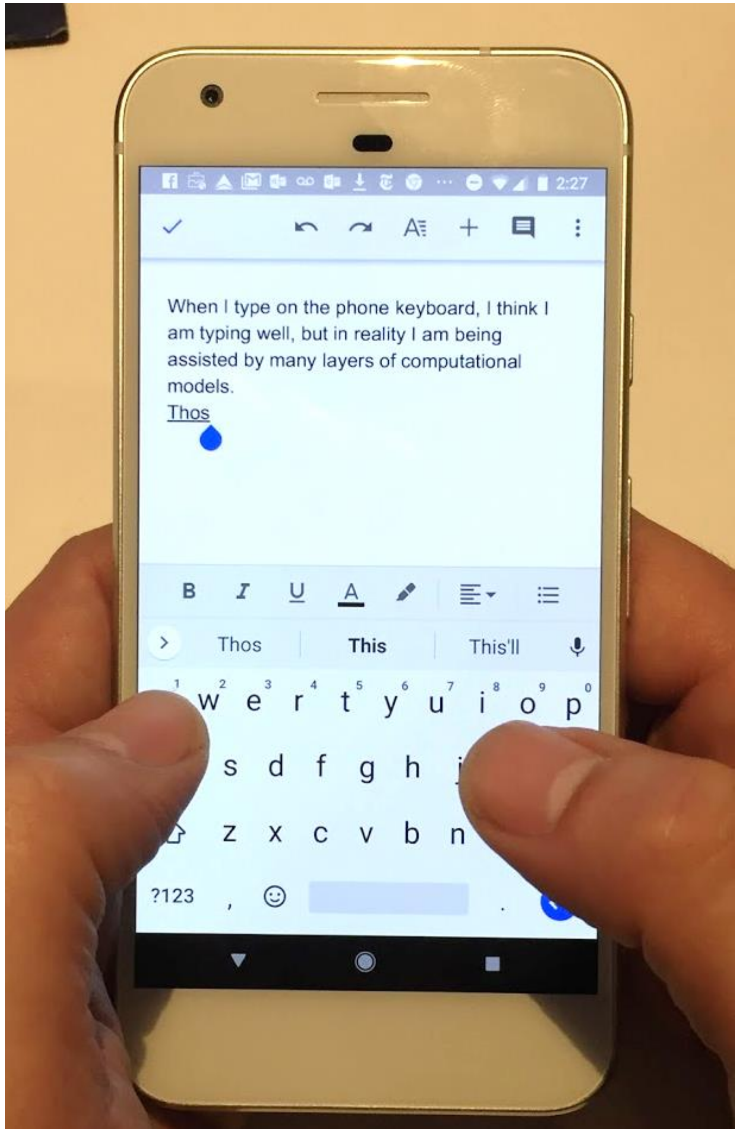
Modern Android touch keyboard (two thumbs)

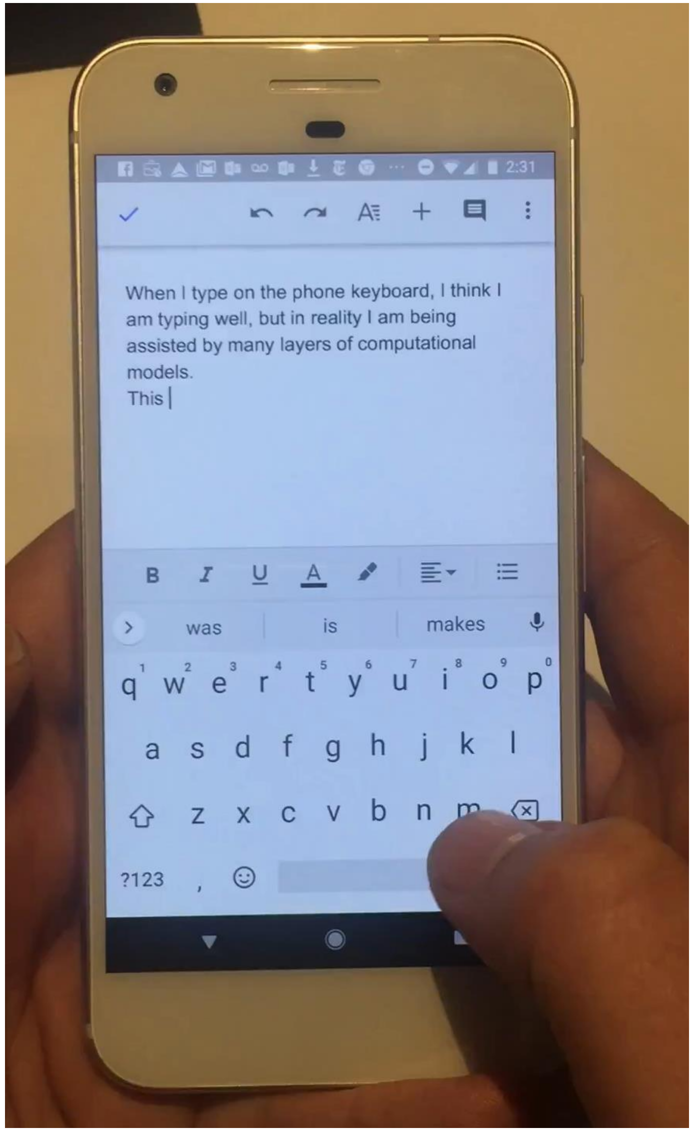
No correction – 27.5 WPM, error rate 6.5%

With correction – 31 WPM, error rate 1.1%

Reyal, S., Zhai, S. and Kristensson, P.O. Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild. In Proc. ACM CHI 2015. 679-688.

Probabilistic phone touch keyboard





Layers of probabilistic models:

- Touch precision model
- Dictionary model
- Language model

+ N-best list UI for error correction

+ Gesture model

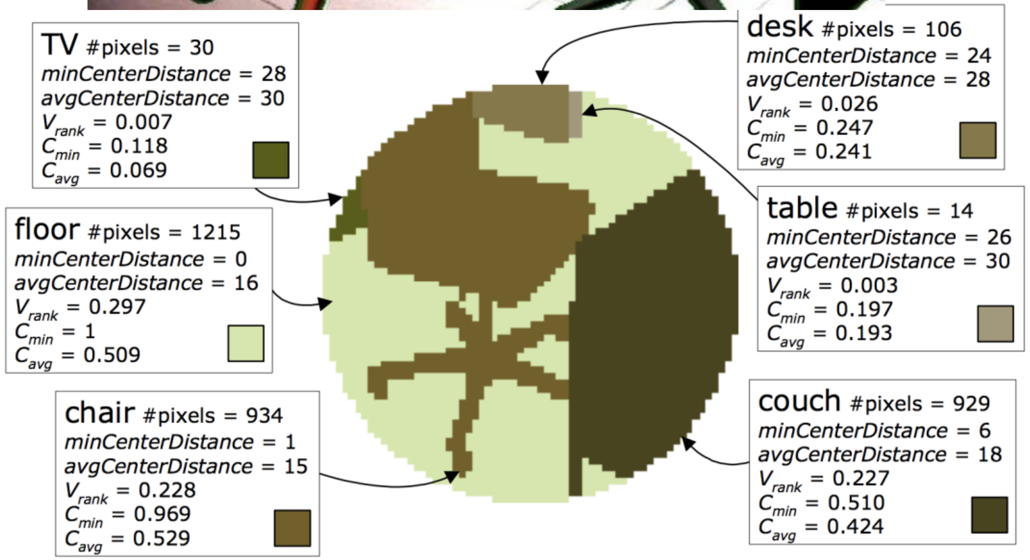
Similar optimizations needed in MR

- Text entry

Similar optimizations needed in MR

- Text entry

- Object selection



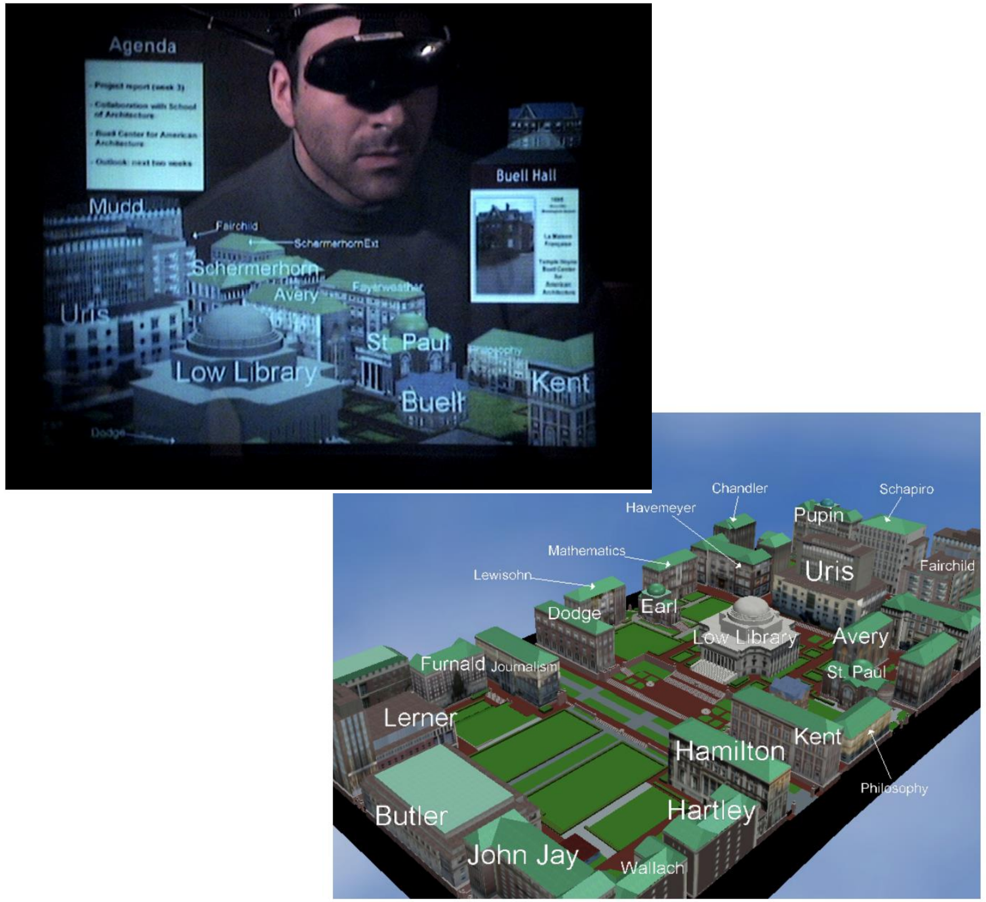
Olwal, A., Benko, H., Feiner, S. SenseShapes: Using Statistical Geometry for Object Selection in a Multimodal Augmented Reality System. *In Proc. ISMAR '03.* 300-301.

Similar optimizations needed in MR

- Text entry
- Object selection
- (Multimodal) input fusion

Similar optimizations needed in MR

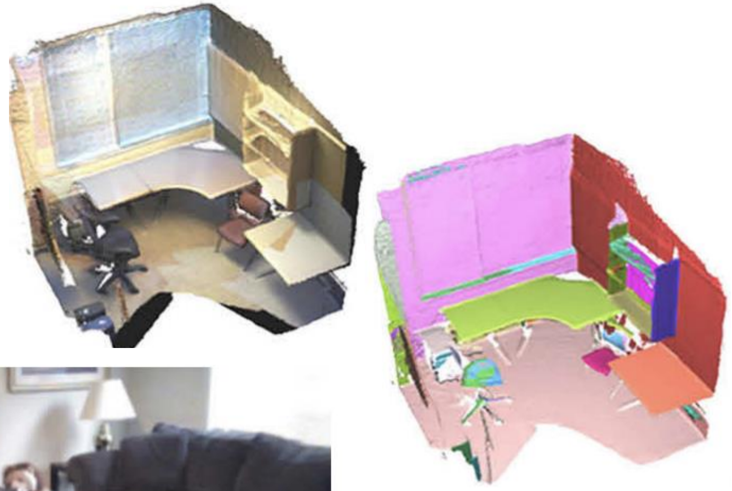
- Text entry
- Object selection
- Multimodal input fusion
- Output optimizations



Bell B., Feiner, S., and Hollerer, T. View Management for Virtual and Augmented Reality. In Proc. ACM UIST 2001.

Similar optimizations needed in MR

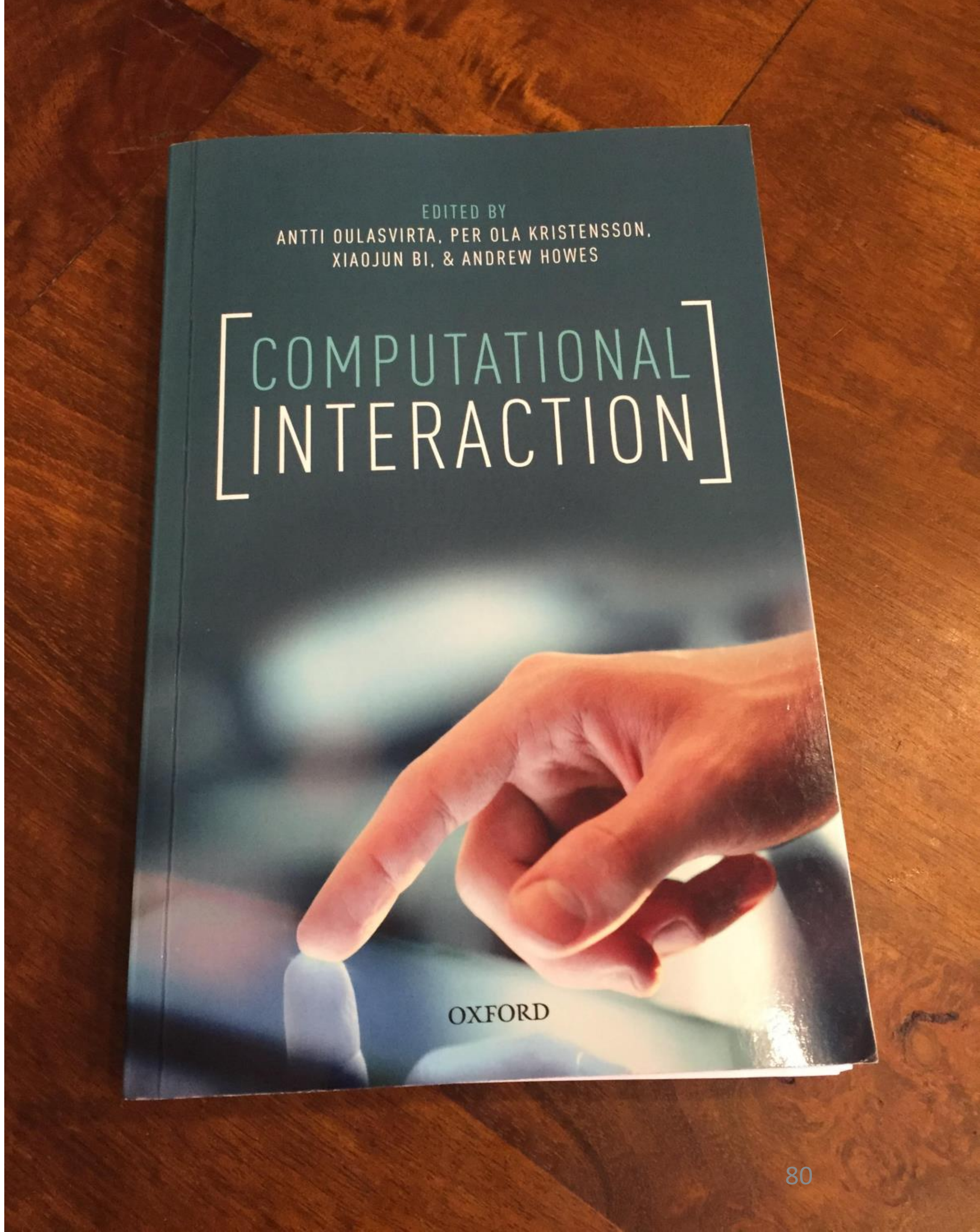
- Text entry
- Probabilistic object selection
- Multimodal input fusion
- Output optimizations



Gal R., Shapira, L., Ofek, E., Kohli, P. FLARE: Fast Layout for Augmented Reality Applications. In Proc. IEEE ISMAR 2014.

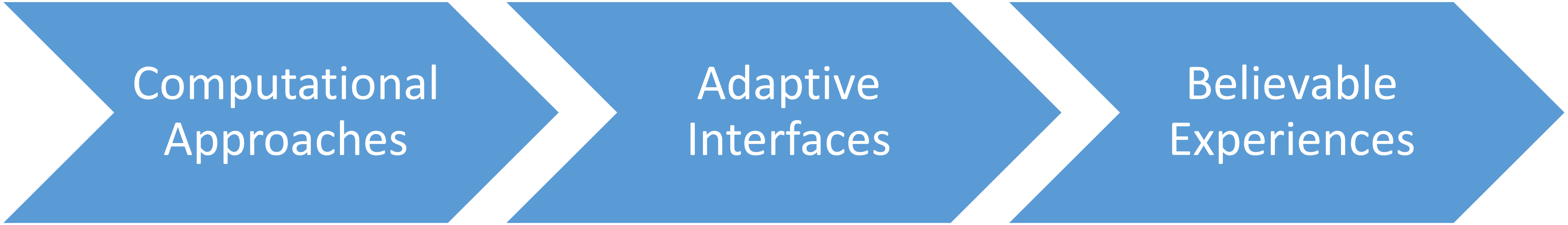
formulate the UI challenges as ***computational problems*** (inferring, sensor fusing, predicting, tolerating noise)

that adapt the interface depending on the user and world context



Compelling MR interactions will be *adaptive*, *believable,* and *computational*.

Maybe it should have been the CAB of MR Interactions?



Design interactions that adapt to the user's actions, the world around them, and the context of use.

Focus on believability. Reality is overrated!

Harness the computational methods to overcome uncertainty, scale, noise, and enable adaptivity.

Hrvoje Benko

benko@fb.com

Facebook Reality Labs

Thanks to all my collaborators!

Extra Slides

Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop

Lung-Pan Cheng[2,1], Eyal Ofek[1], Christian Holz[1], Hrvoje Benko[1], Andrew D. Wilson[1]

[1]Microsoft Research
Redmond, WA, USA
{eyalofek, cholz, benko, awilson}@microsoft.com

[2]Hasso Plattner Institute
Potsdam, Germany
lung-pan.cheng@hpi.de

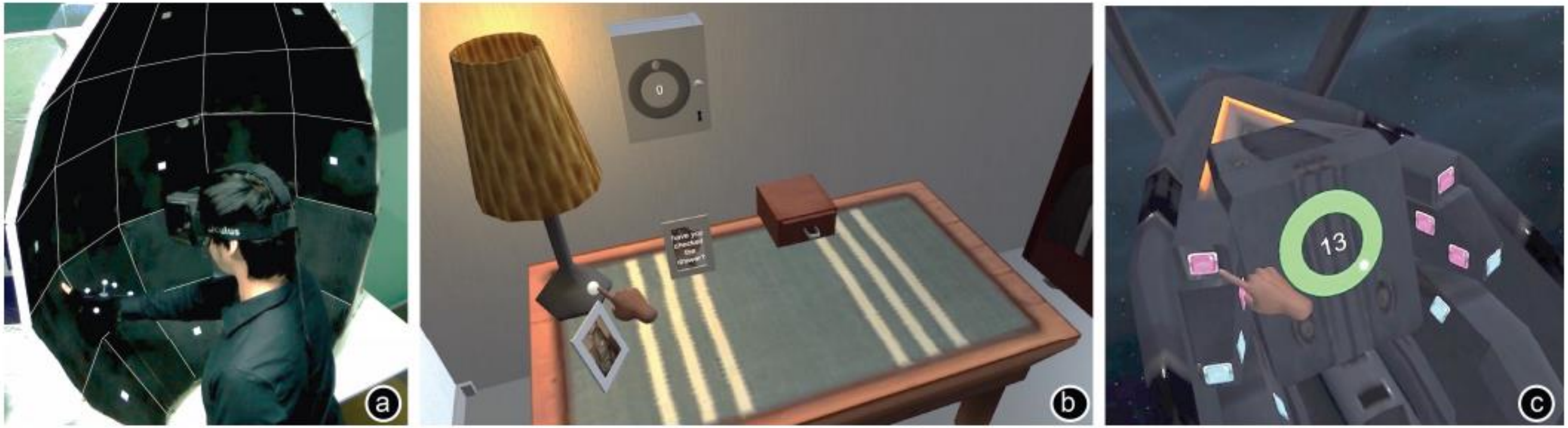
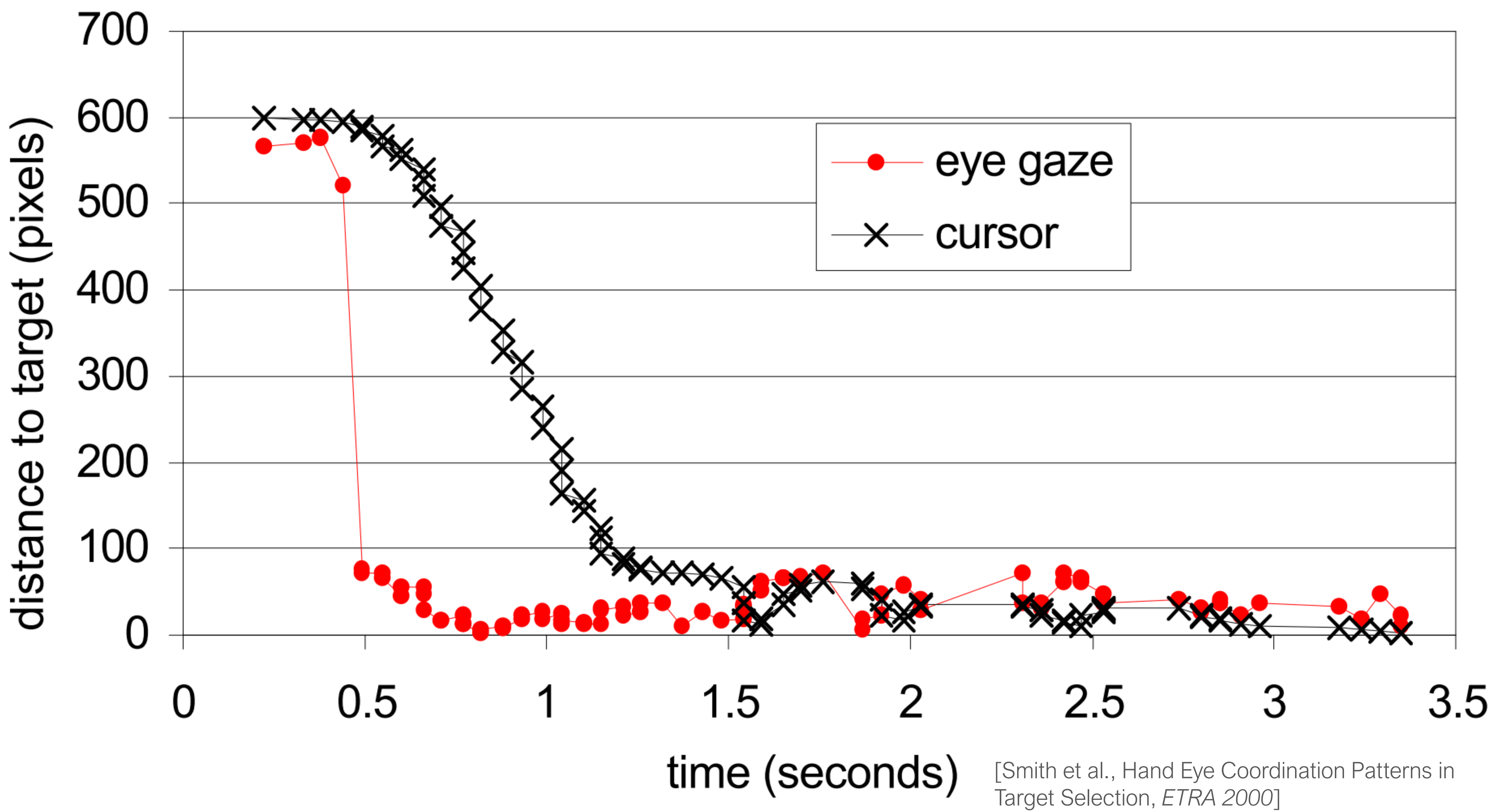


Figure 1. (a) Our hemispherical prop is an example of a *Sparse Haptic Proxy*. It simulates both, (b) a room and (c) a cockpit scene to provide physical touch feedback during interaction. (White lines on the prop added for visibility on the black background).

ABSTRACT

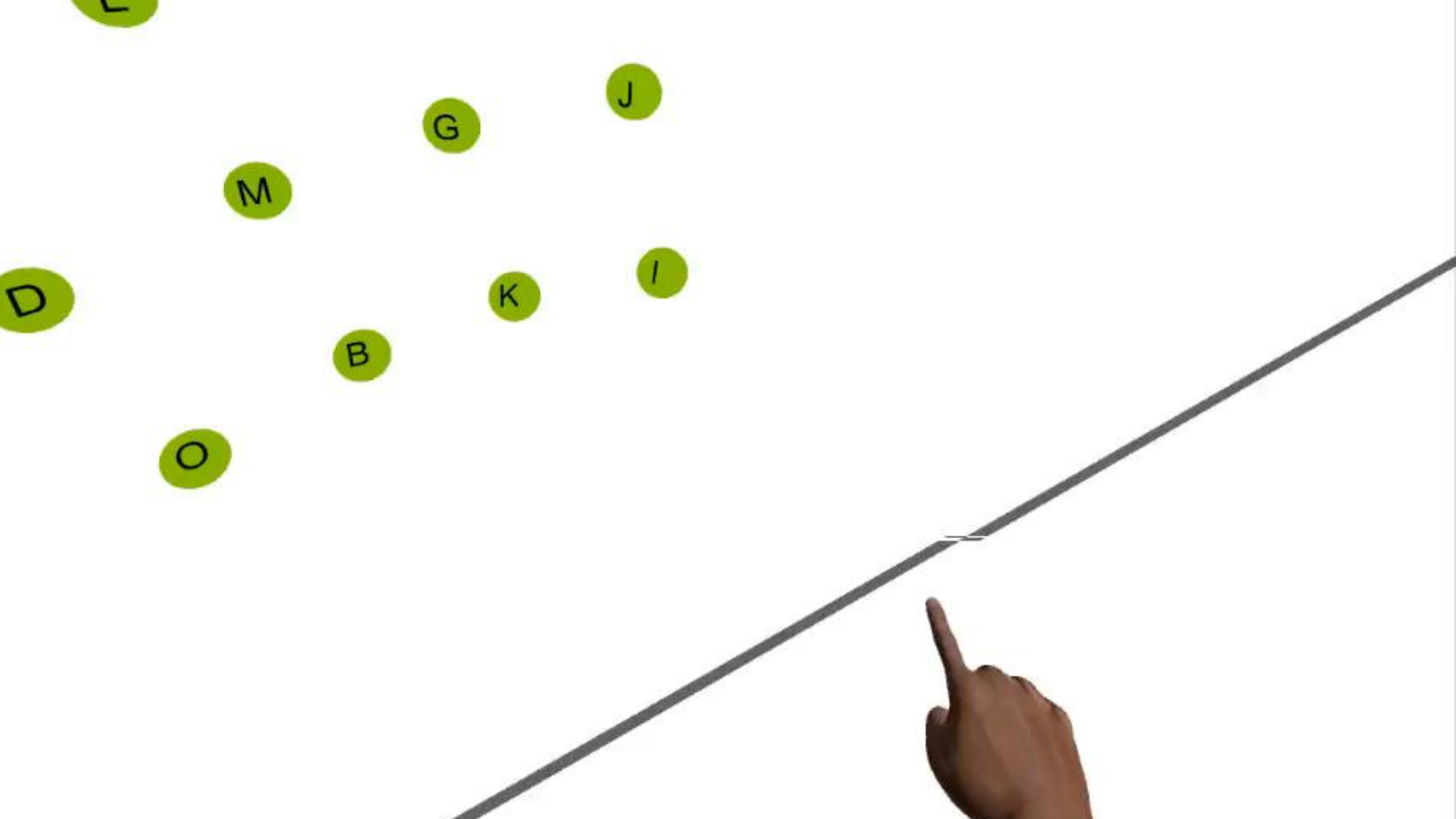
over better-matching surfaces of our Sparse Haptic Proxy



[Smith et al., Hand Eye Coordination Patterns in Target Selection, *ETRA 2000*]

SMI eye-tracking module

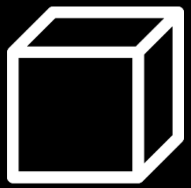
250 hz eye-tracker
on Oculus DK2

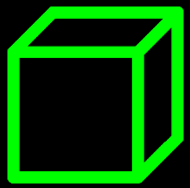

J
G
M
D
K
I
B
O

We can predict with 97.5% accuracy what is the user’s intended target 2 seconds before reaching the target!

Body Warping

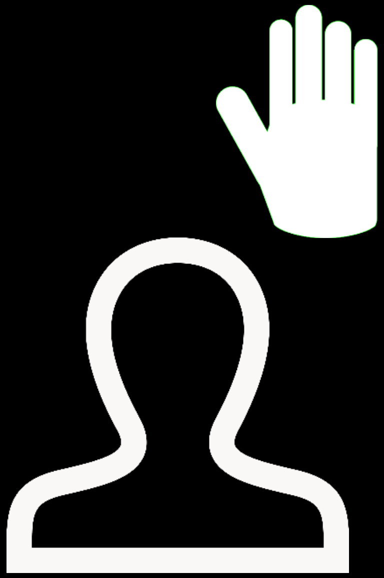
physical cube



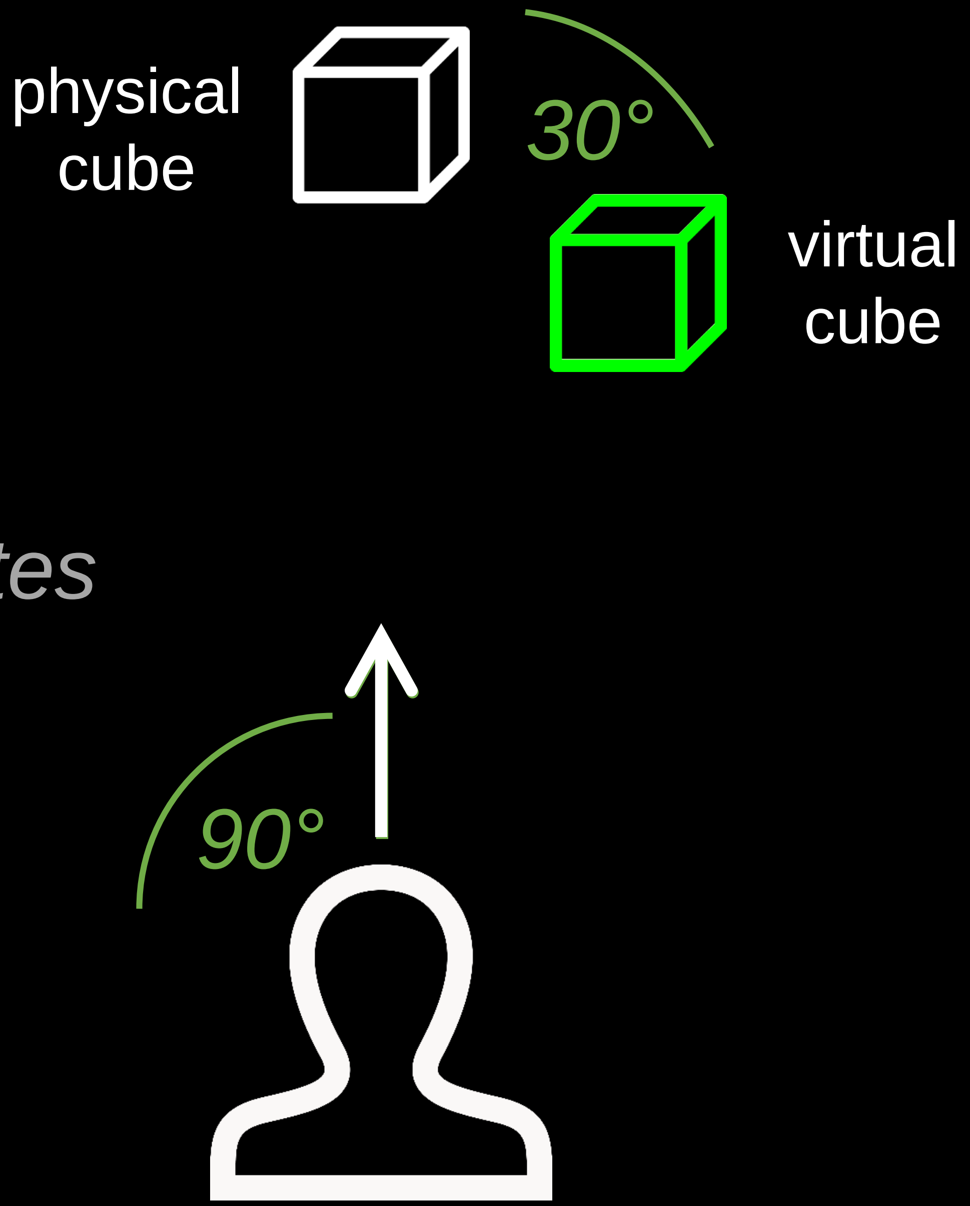


virtual cube

The Rendered Body Shifts to The Right



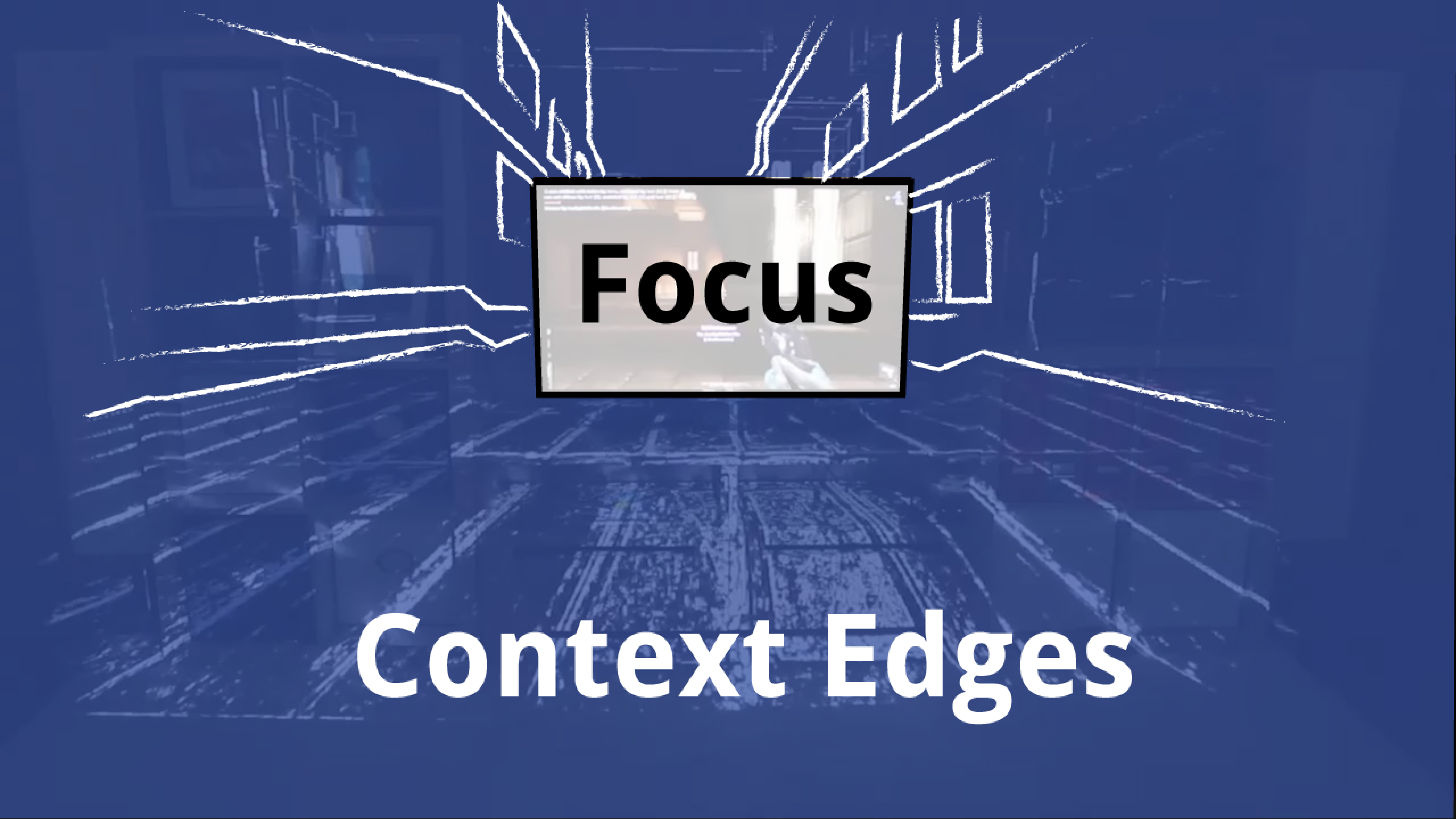
World Warping

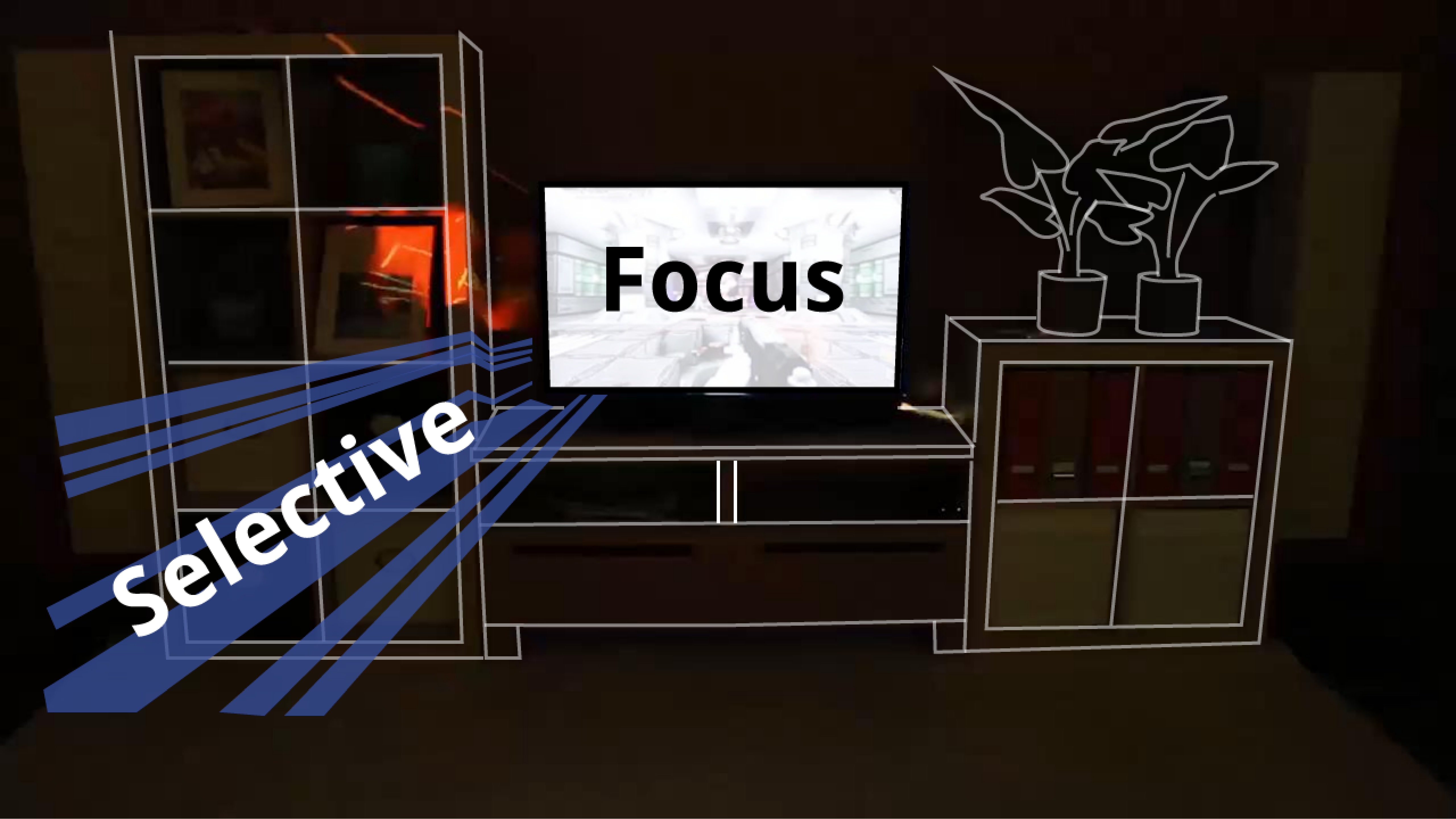


The World Also Rotates
(At Different Rate)

IllumiRoom

Peripheral Projected Illusions for Interactive Experiences.
Jones, B., Benko, H., Ofek, E., and Wilson, A. D. In *Proc. of ACM CHI 2013.*

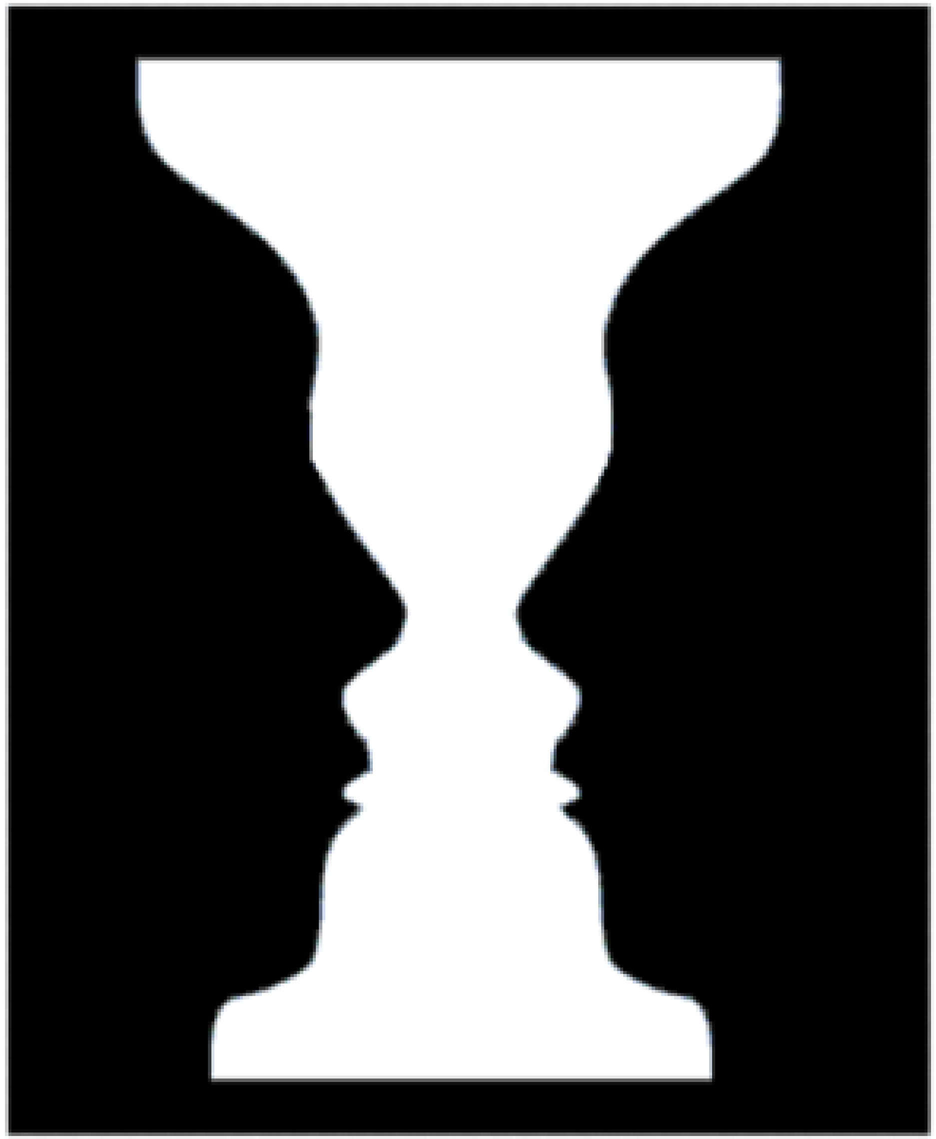

Focus
Context Edges


Focus
Selective

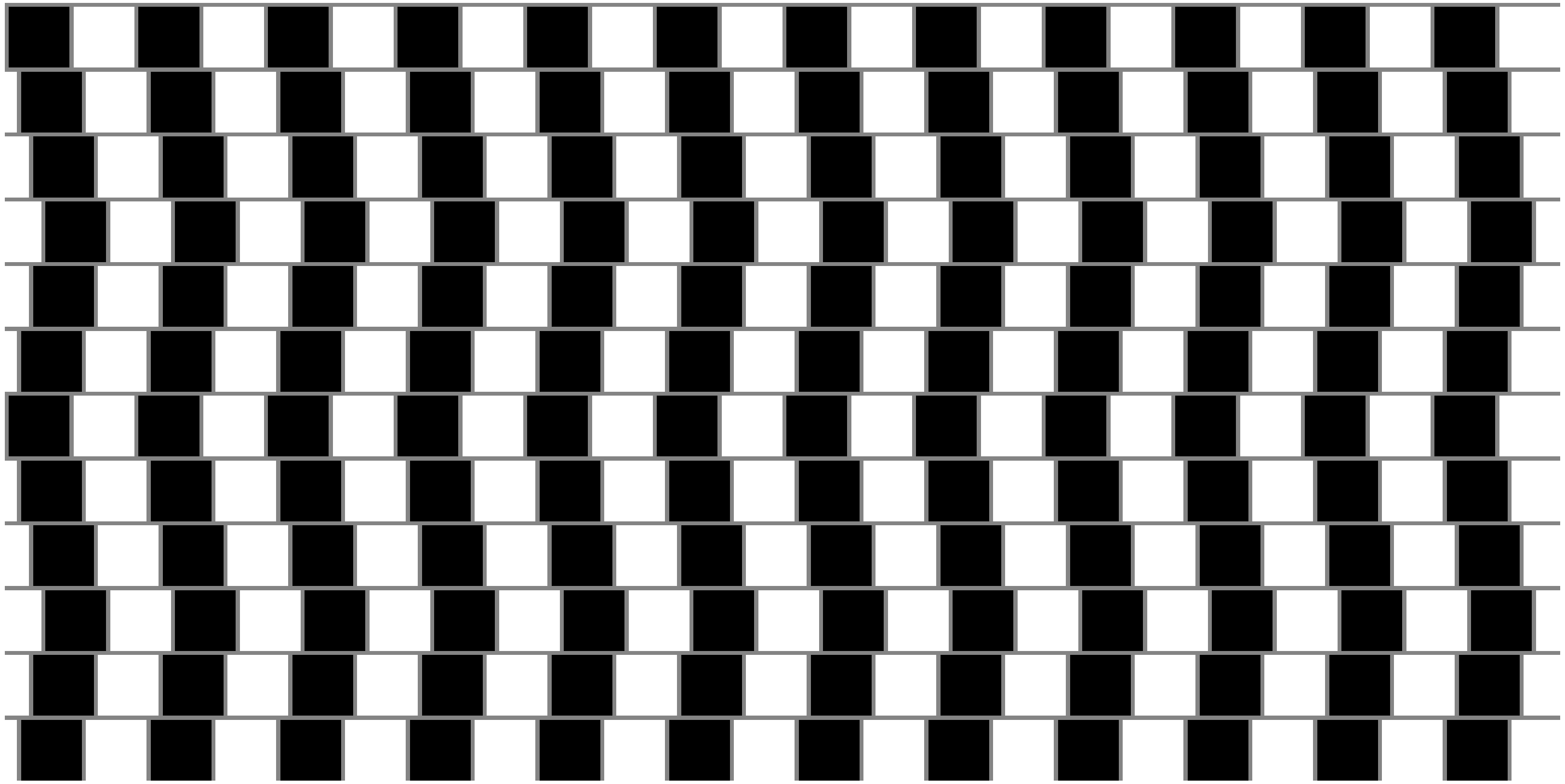

Snow

What is “real” vs. what we “perceive as real” is not necessarily the same!

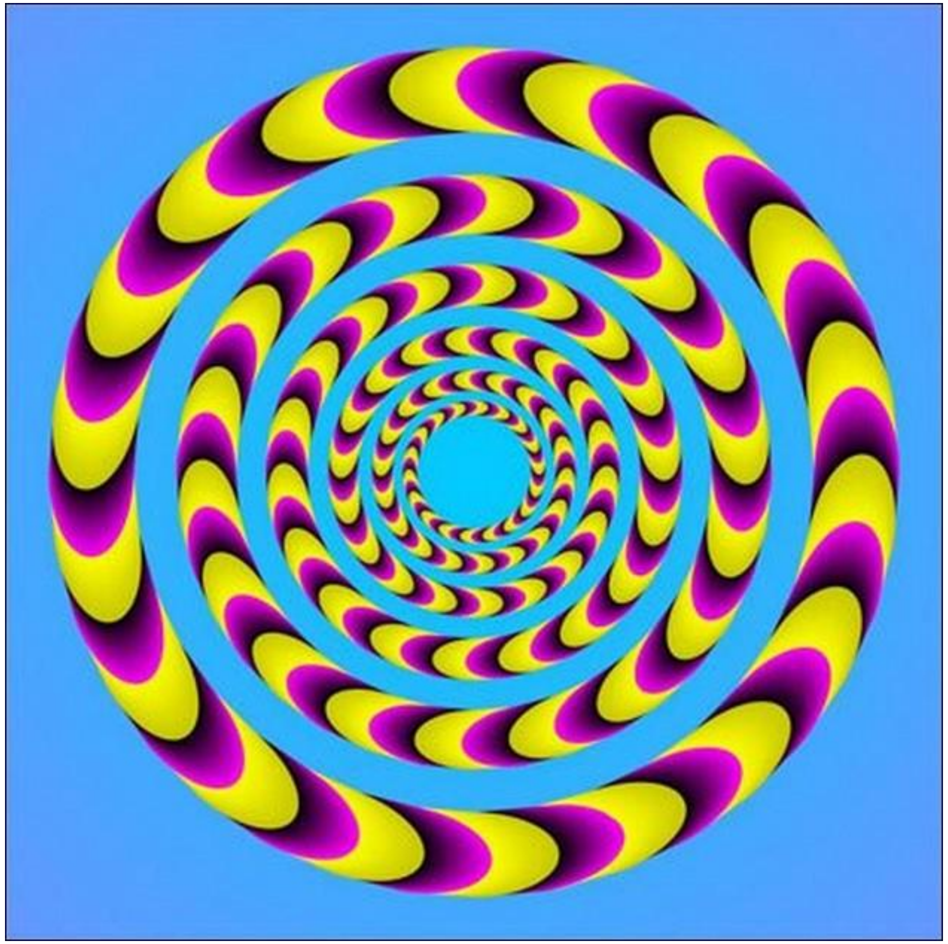
Faces or vase?



Straight or crooked?



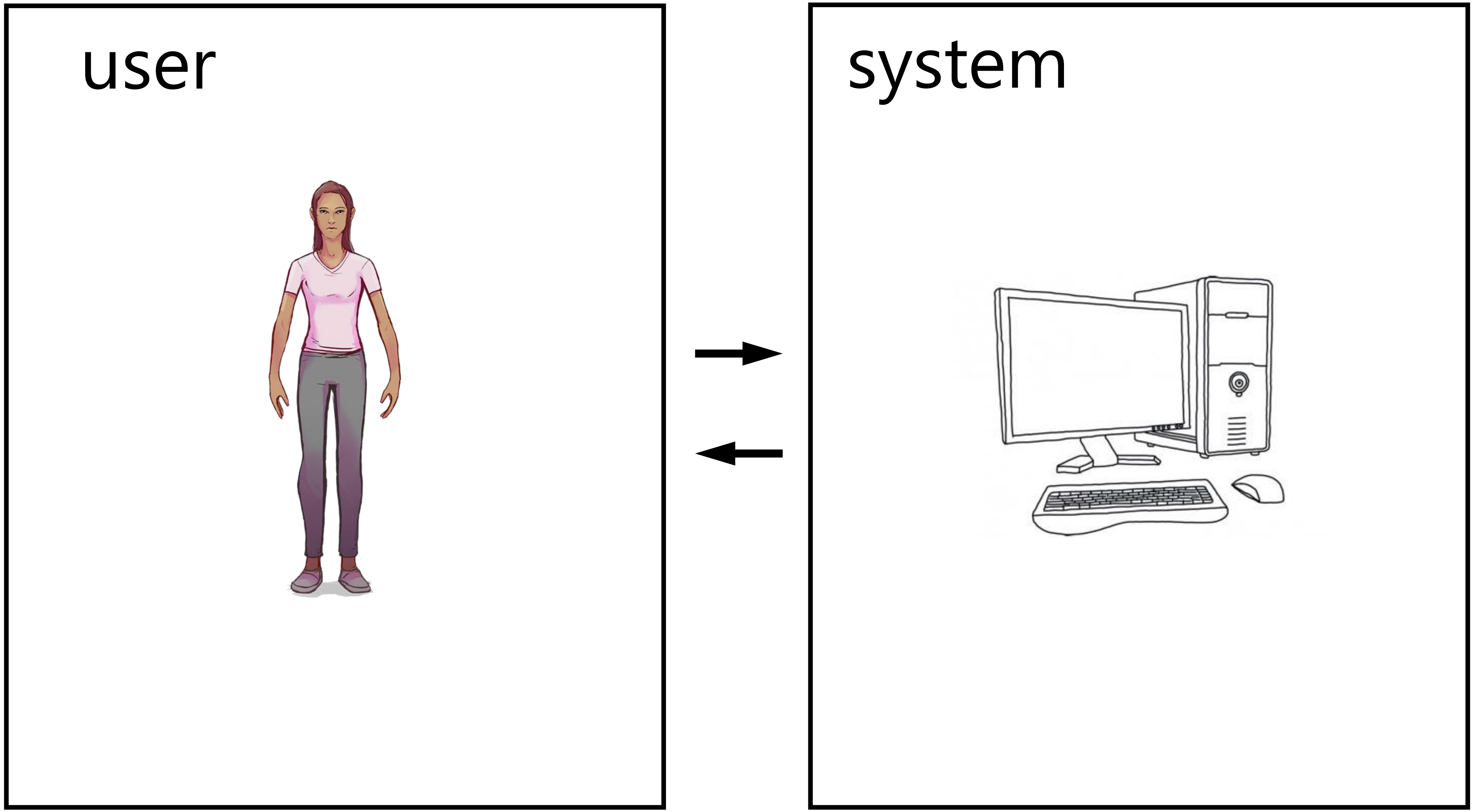
Moving or static?



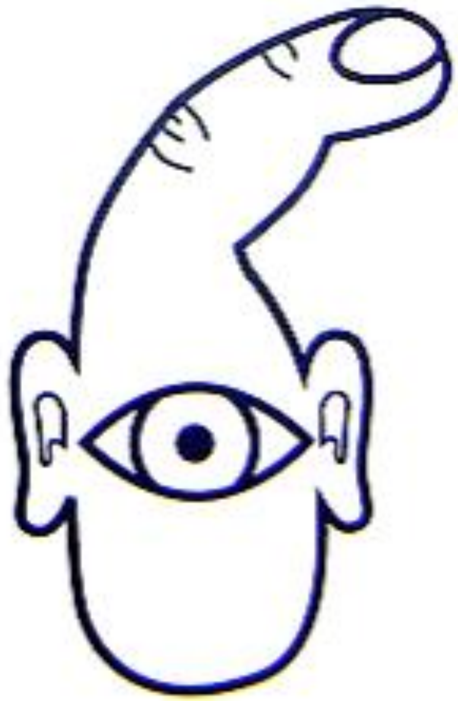
Insisting on “as real as possible” designs can lead to sub-optimal MR experience.

Think about MR interfaces as perceptual illusions that give the user a believable experience!

Typical HCI model

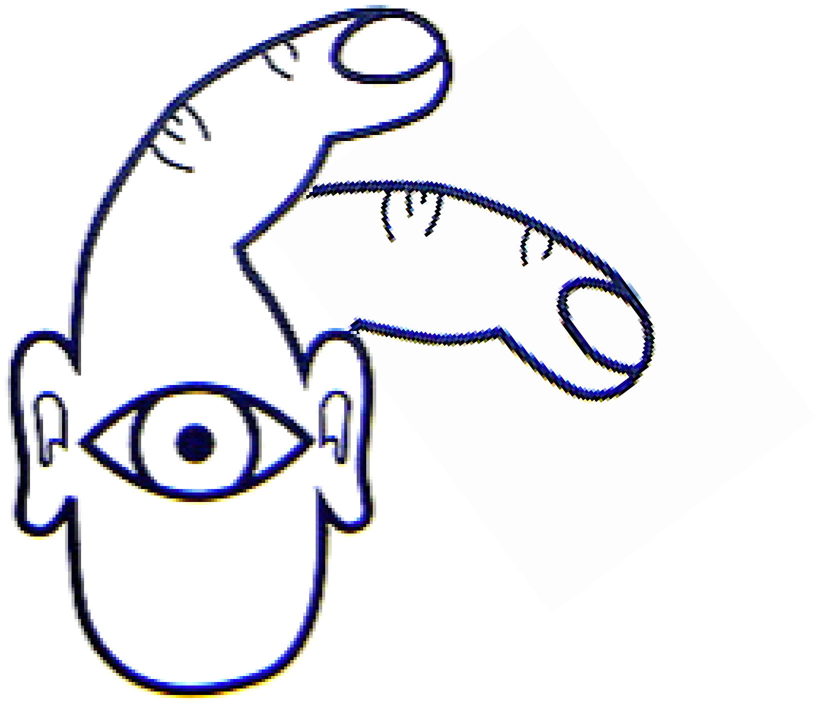


How the computer sees us!



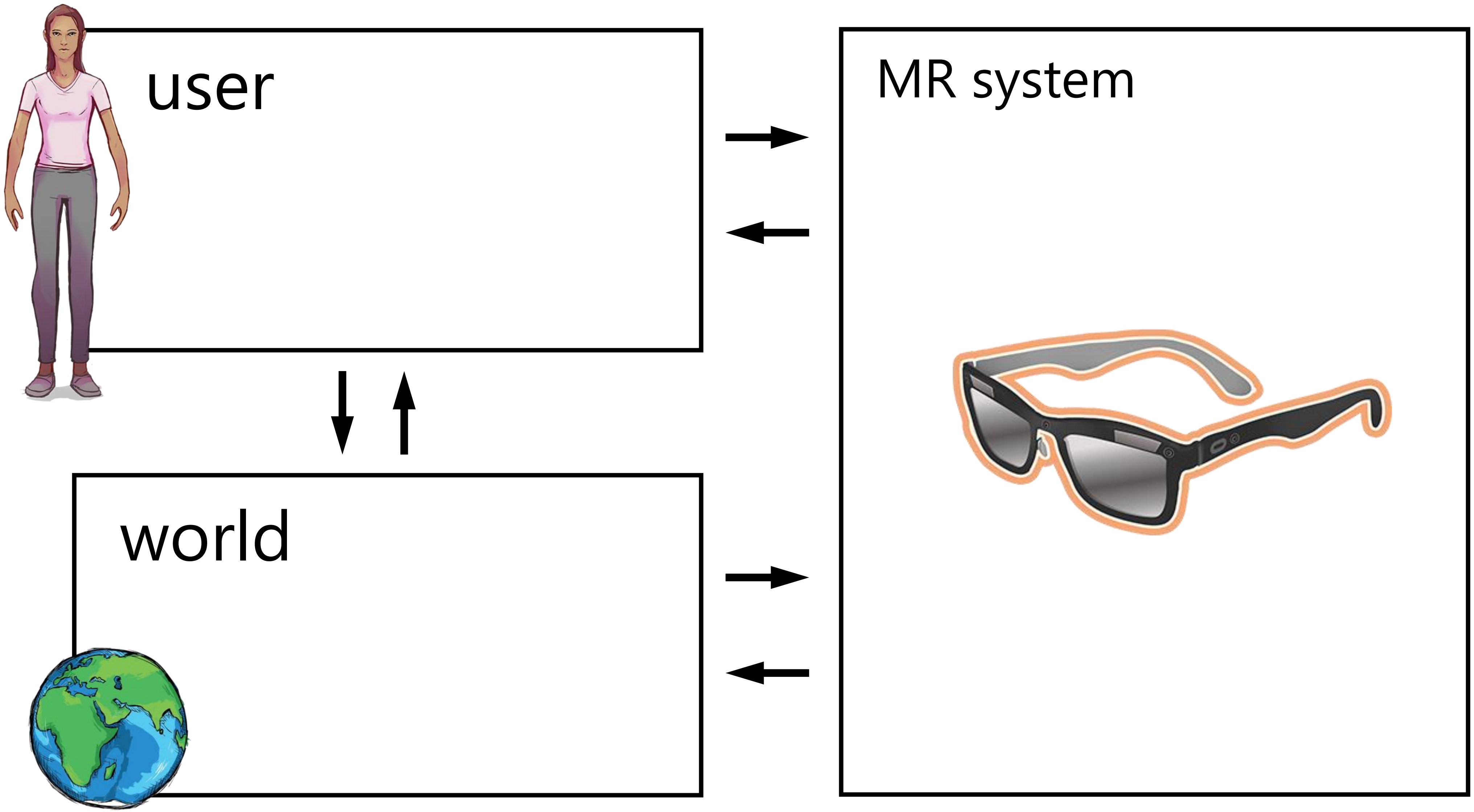
Tom Igoe and Dan O’Sullivan - *Physical Computing*. 2004.

How the phone sees us!



Modified Tom Igoe and Dan O'Sullivan - *Physical Computing*. 2004.

MR interaction model



MR interaction model