

# AUGMENTED REALITY AND ARTIFICIAL INTELLIGENCE

---

Mark Billinghurst

[mark.billinghurst@auckland.ac.nz](mailto:mark.billinghurst@auckland.ac.nz)



# Augmented Reality

## 1. Combines Real and Virtual Images

- Both can be seen at the same time

## 2. Interactive in real-time

- The virtual content can be interacted with

## 3. Registered in 3D

- Virtual objects appear fixed in space

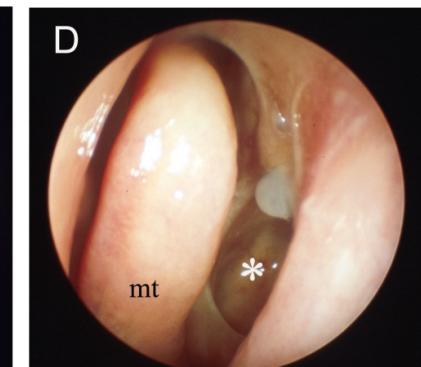
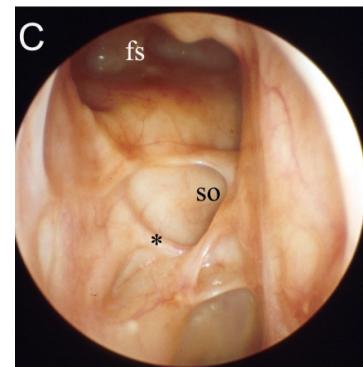
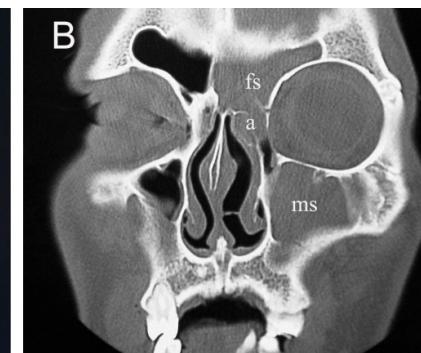
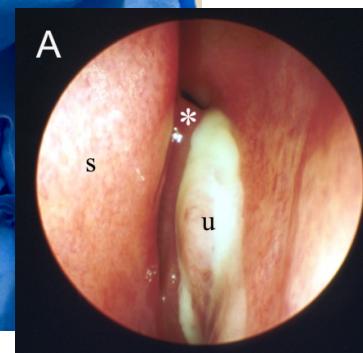
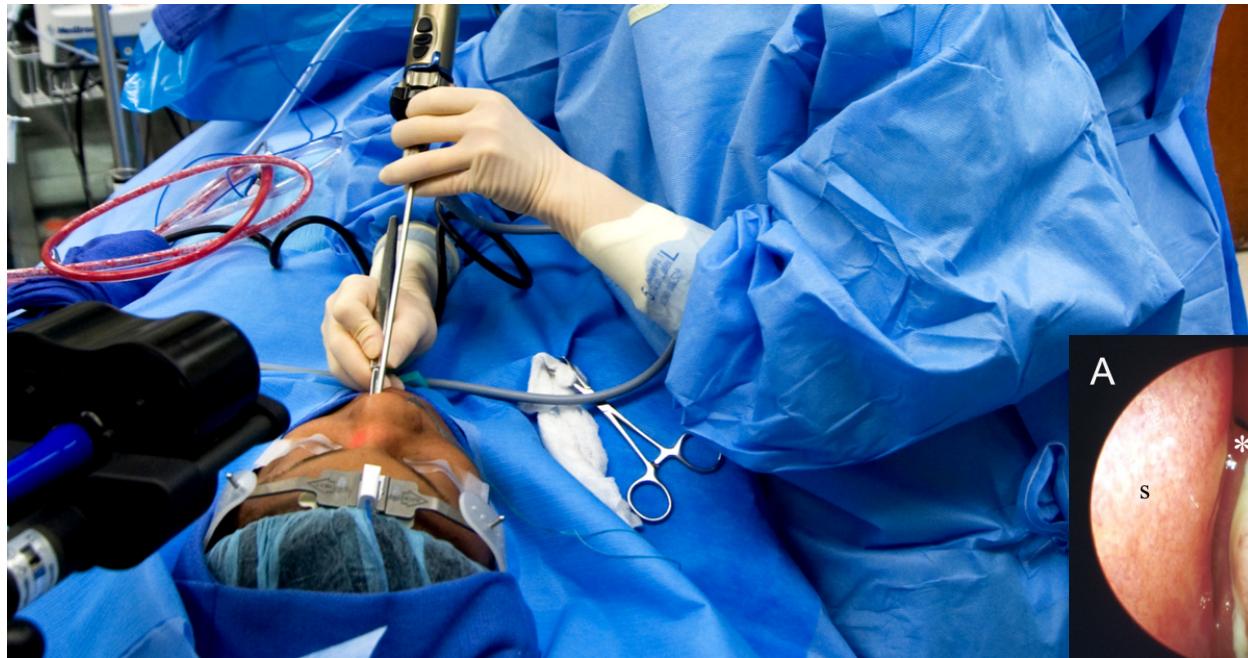
# 2008 - CNN



[https://www.youtube.com/watch?v=v7fQ\\_EsMJMs](https://www.youtube.com/watch?v=v7fQ_EsMJMs)

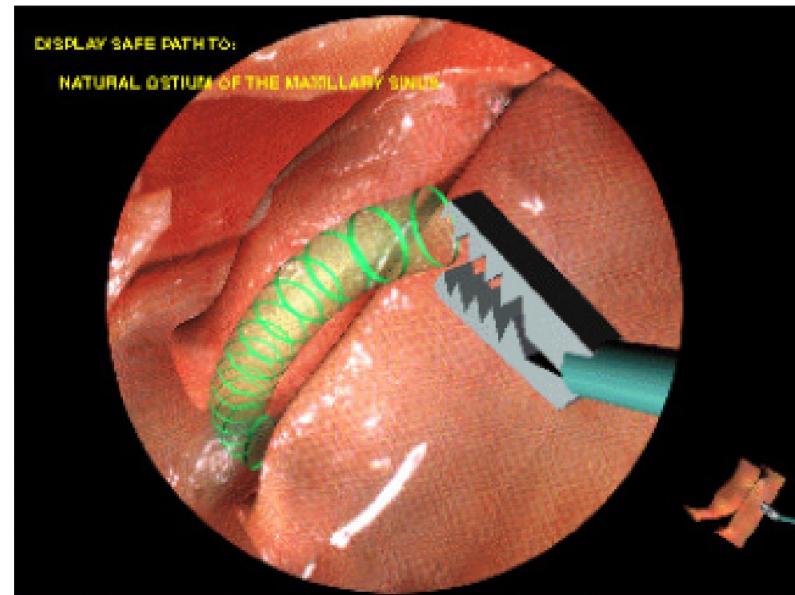
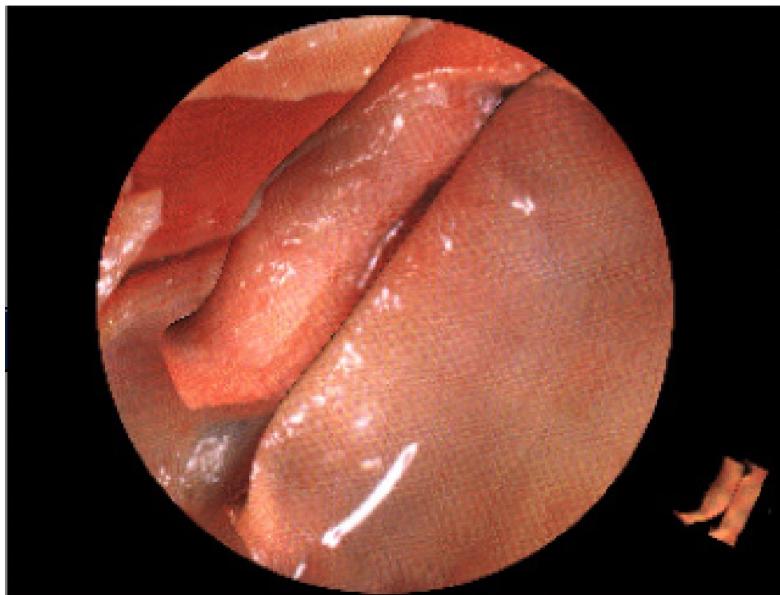


# Sinus Surgery



- Dangerous surgery
- Thin bones near brain, optic nerve

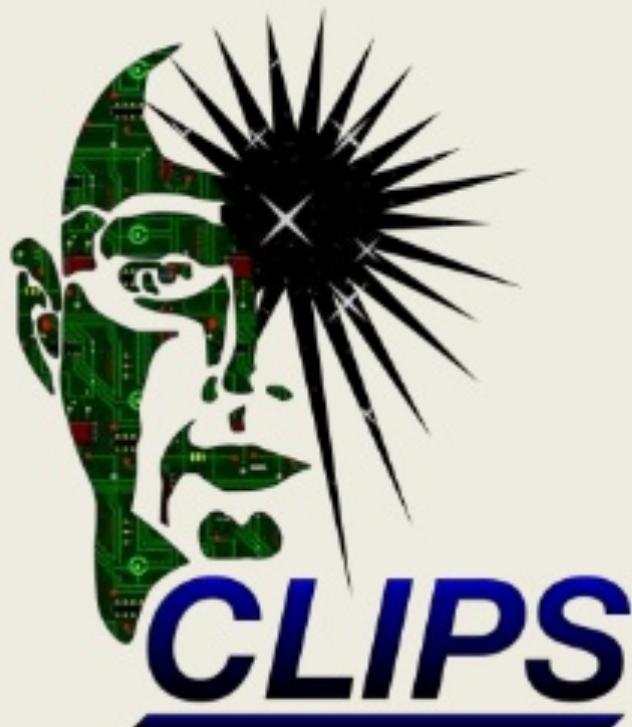
# Expert Surgical Assistant (1995)



- Desktop 3D training simulator, real instruments
- Rule-based expert system (CLIPS)
- Using speech and gesture to interact with surgical training

Billinghurst, M., Savage-Carmona, J., Oppenheimer, P. and Edmond, C. (1995). The Expert Surgical Assistant: An Intelligent Virtual Environment with Multimodal Input. In Weghorst, S., Sieberg, H.B. and Morgan, K.S. *Proceedings of Medicine Meets Virtual Reality IV*, pp. 590-607.

# CLIPS



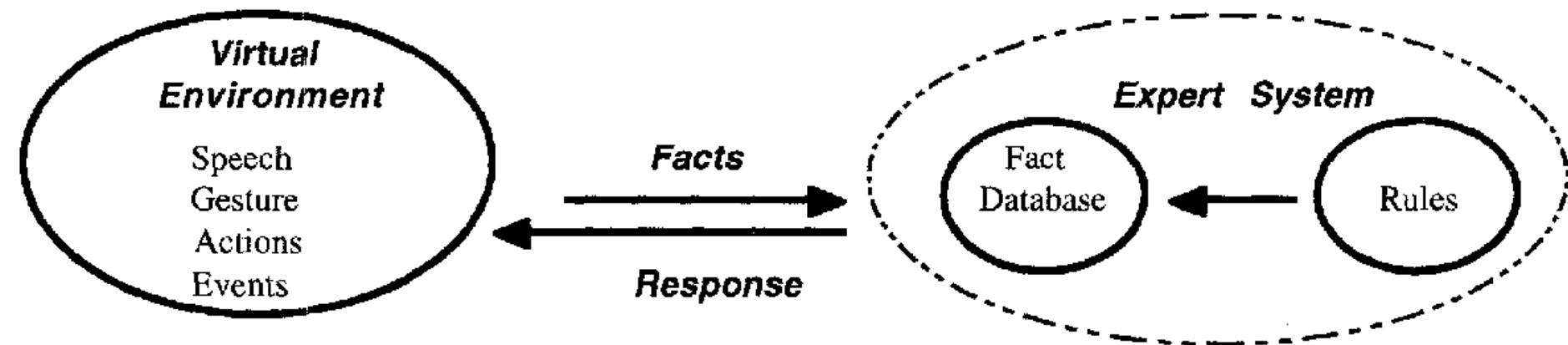
- ▶ The CLIPS were developed in 1985 at NASA-Johnson Space Center.
- ▶ CLIPS is a productive development and delivery expert system tool which provides a complete environment for the construction of rule and/or object based expert systems.
- ▶ This project is based on AI (artificial intelligence) language CLIPS.

See <http://www.clipsrules.net/>

# Example CLIPS Rule

```
(defrule sunday "Things to do on Sunday"
  (salience 0)      ; salience in the interval [-10000,
  10000]
  (today is Sunday)
  (weather is sunny)
  =>
  (assert (chore wash car))
  (assert (chore chop wood))
)
```

# CLIPS



- Fact and Rule database
- Networked applications
  - VR and CLIPS communicated via sockets

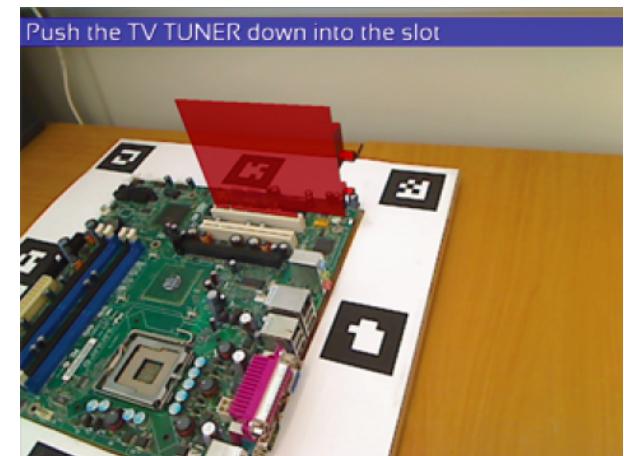
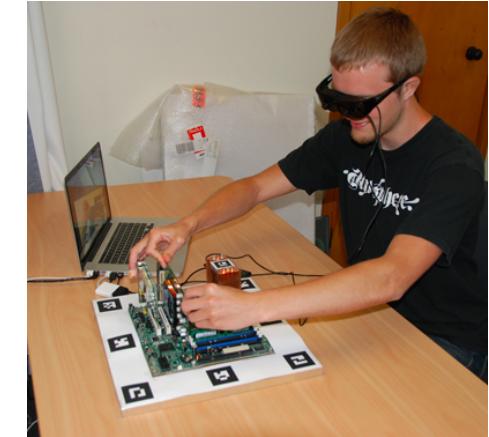
# Multiple Levels of Understanding

Understanding Level	Prior Knowledge Used	Types of Rules Matched
Pragmatic Understanding	Context, Semantic, Feature	Sequences of actions: <i>if user has made an incision they should remove the tumor next.</i>
Context Recognition	Context, Semantic, Feature	Sets of related actions: <i>if user is holding a scalpel and touching skin they must be making an incision.</i>
Semantic Understanding	Context, Feature	Single actions: <i>if user moves an object from the floor to the table then put it back on the floor.</i>
Feature Understanding	Context	Single Events: <i>if collision is detected sound a warning.</i>

Figure 2: The different levels of understanding and the types of rules matched at each level.

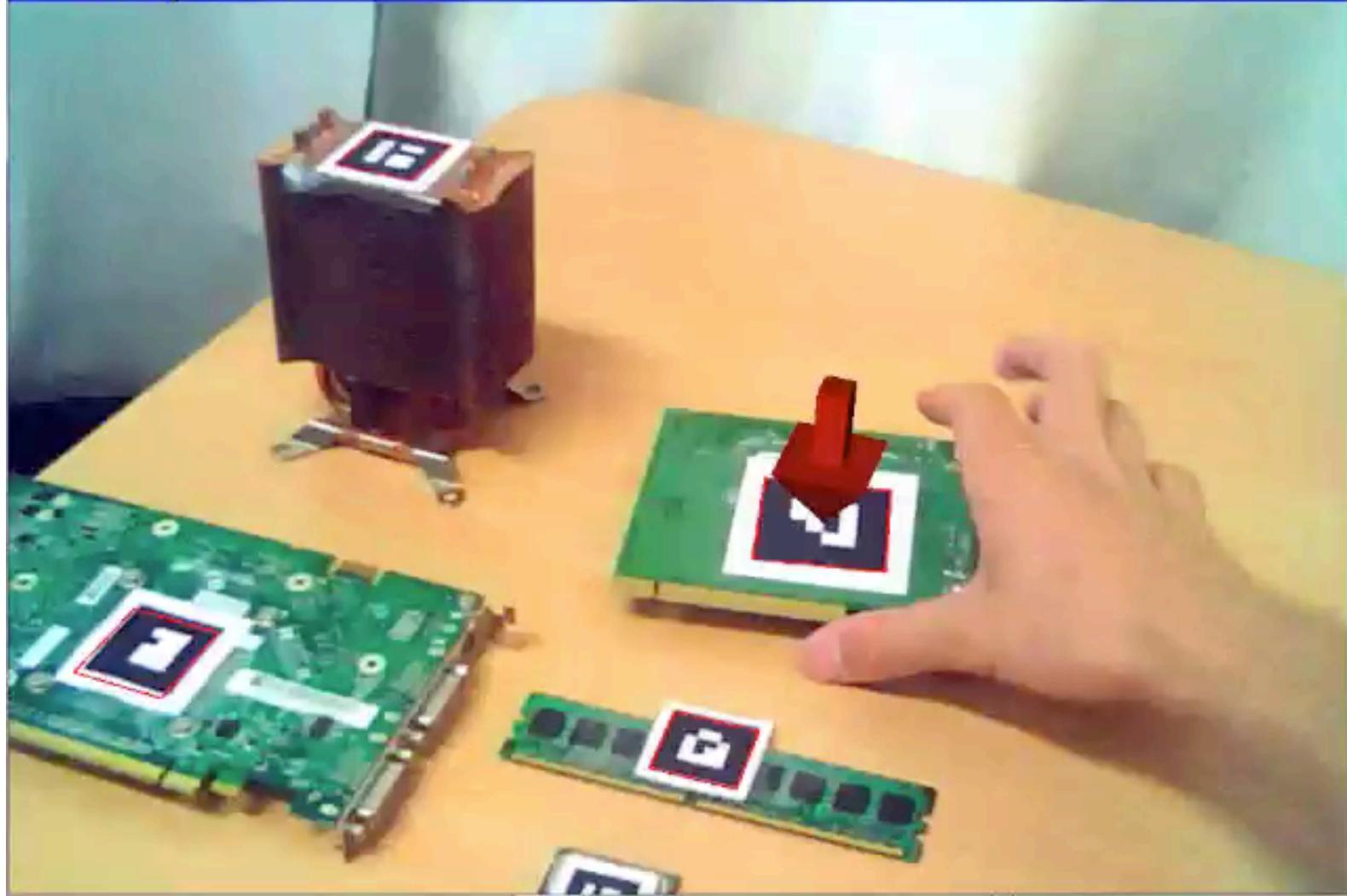
# Intelligent Tutoring

- Move to Implicit Input vs. Explicit
  - Recognize user behaviour
  - Provide adaptive feedback
  - Support scaffolded learning
  - Move beyond check-lists of actions
- Eg AR + Intelligent Tutoring
  - Constraint based ITS + AR
  - PC Assembly (Westerfield (2015))
    - 30% faster, 25% better retention

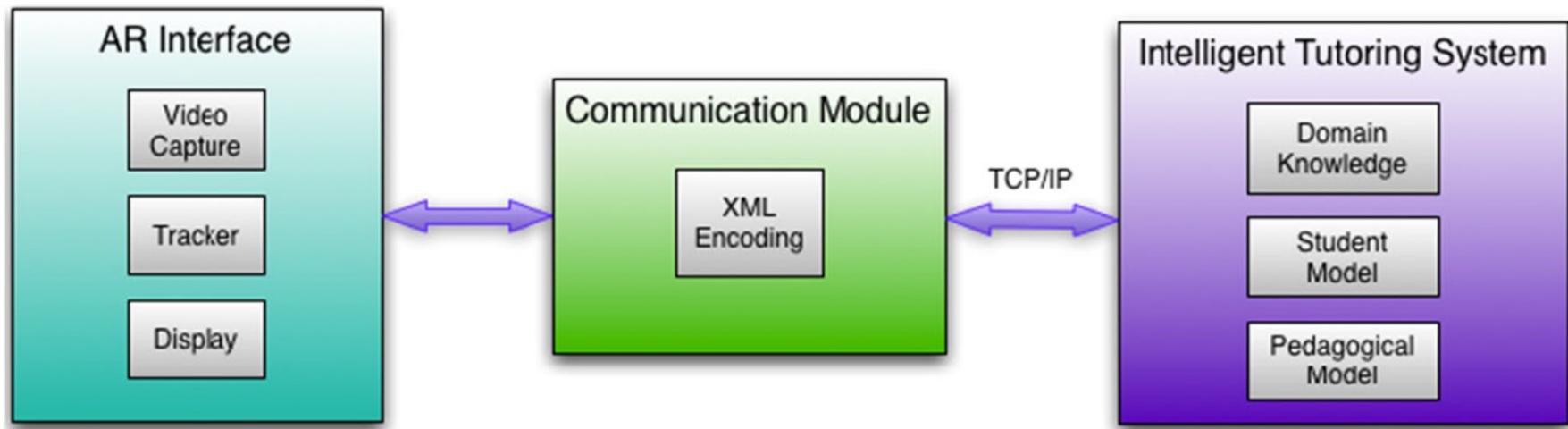


Westerfield, G., Mitrovic, A., & Billinghurst, M. (2015). Intelligent Augmented Reality Training for Motherboard Assembly. *International Journal of Artificial Intelligence in Education*, 25(1), 157-172.

Pick up the TV TUNER and move it near the motherboard

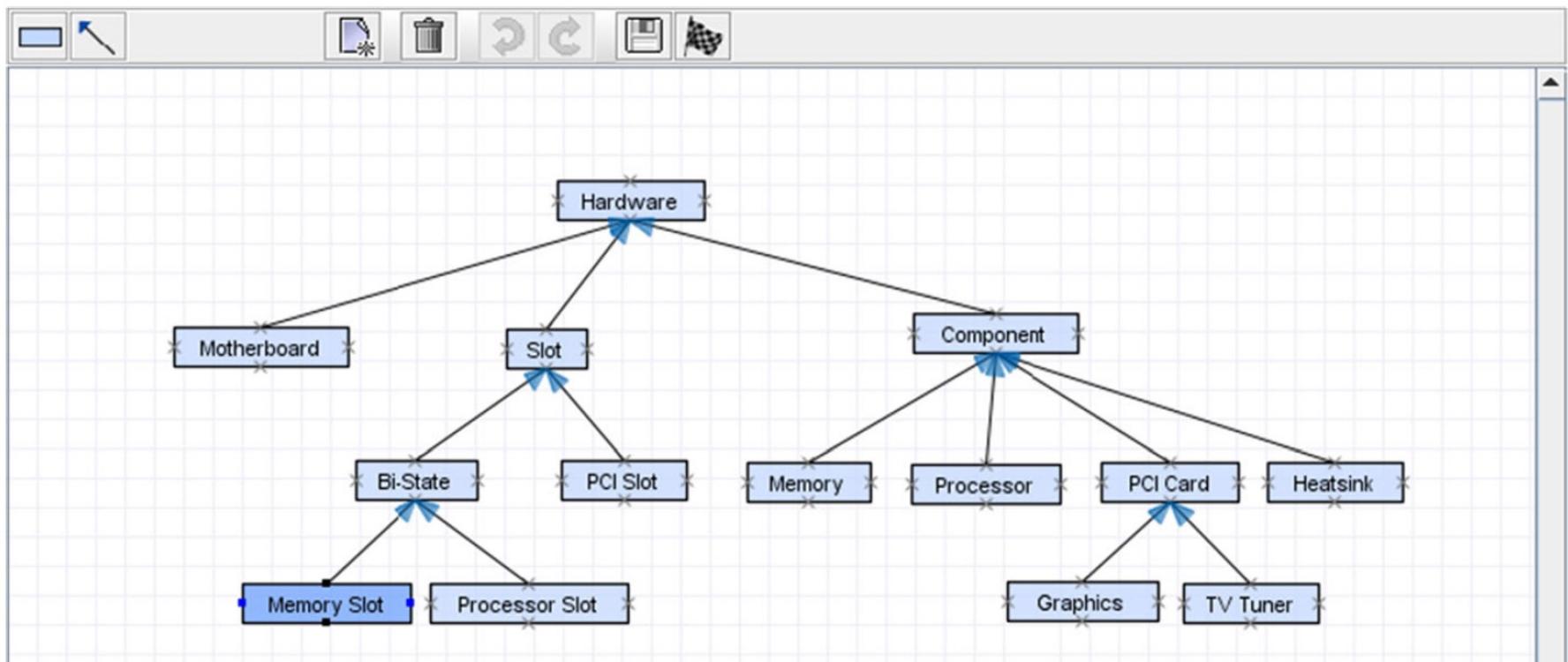


# ITS Architecture



- Networked module approach
- Separate AR interface from ITS module

# ASPIRE ITS



- <http://aspire.cosc.canterbury.ac.nz/>
- Constraint based intelligent tutoring
- Need to set domain ontology

# GIFT



- <https://www.gifttutoring.org/>
- Generalized Interactive Framework for Tutoring
- Open source ITS
- Integrated with Unity – support for AR/VR

# Intelligent Agents

- **AR characters**

- Virtual embodiment of system
- Multimodal input/output

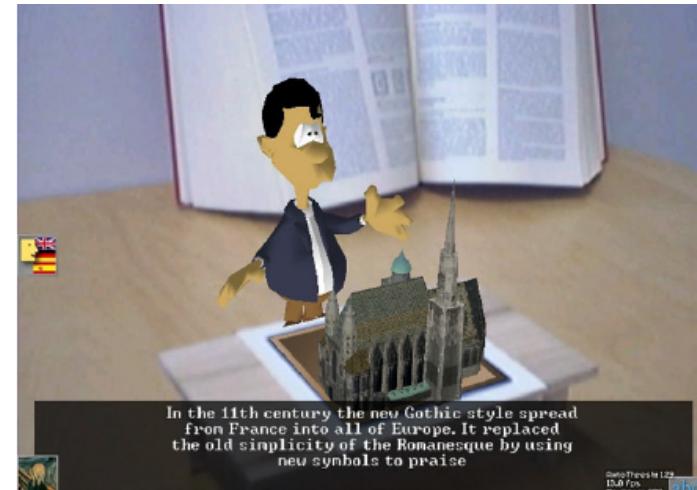
- **Examples**

- AR Lego, Welbo, etc

- Mr Virtuoso

- AR character more real, more fun

- On-screen 3D and AR similar in usefulness



Wagner, D., Billinghurst, M., & Schmalstieg, D. (2006). How real should virtual characters be?. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (p. 57). ACM.



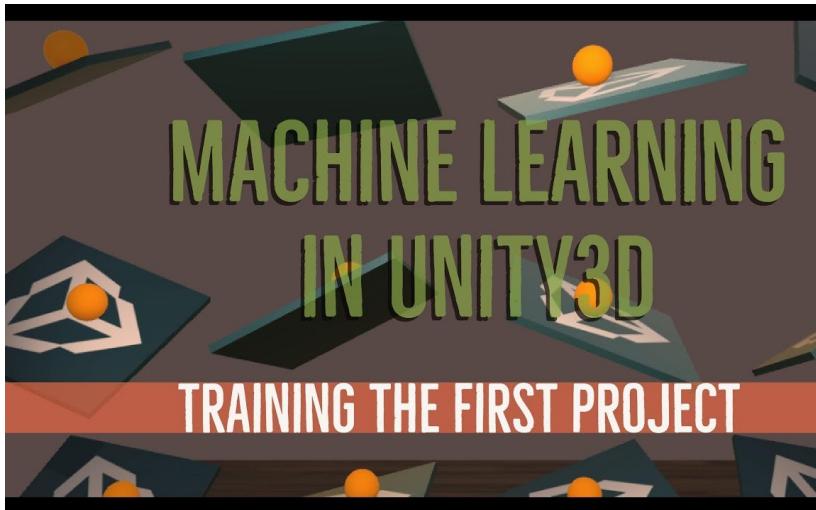
10

# Soul Machines



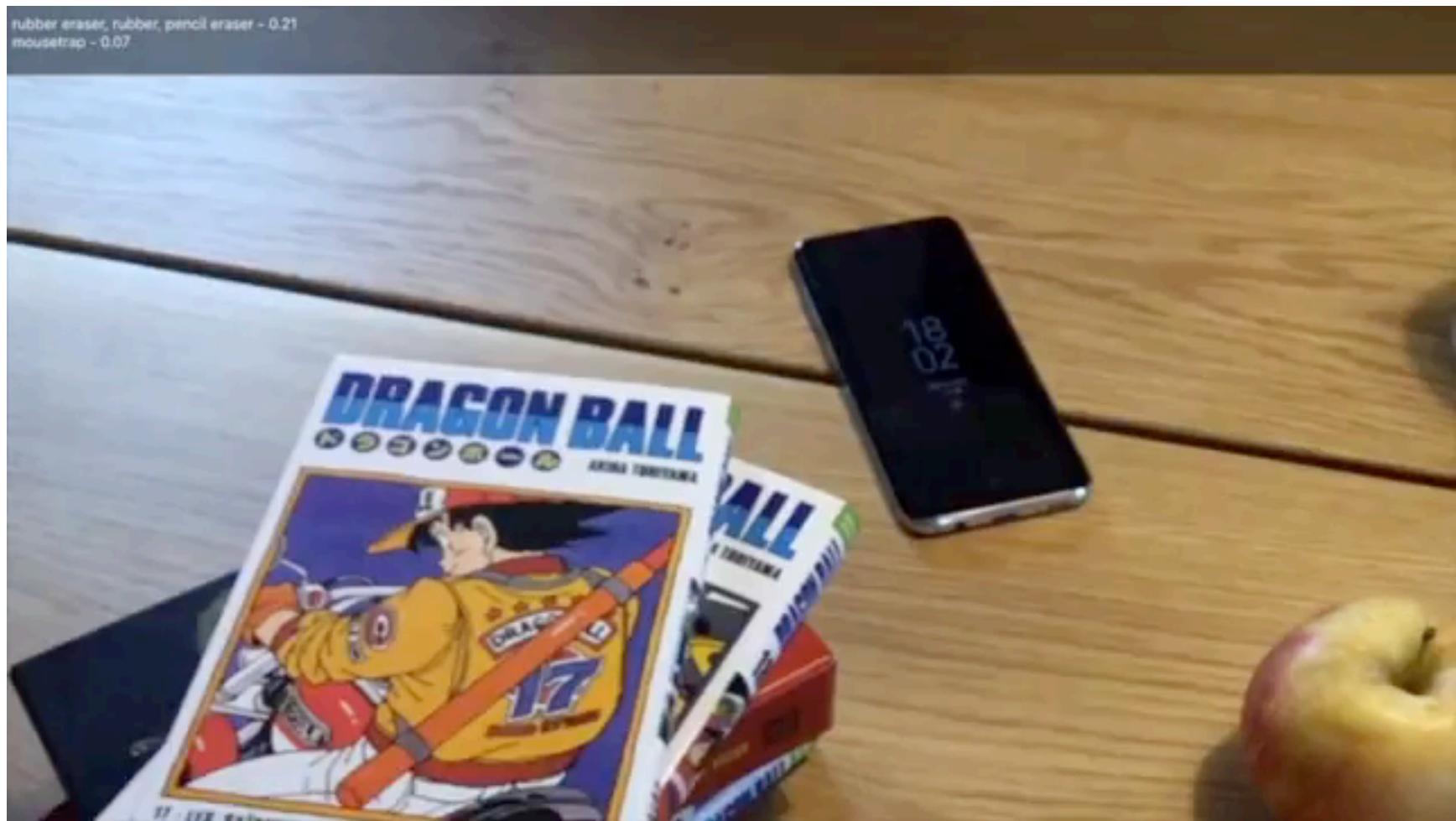
# What is a Digital Human?

# Machine Learning and AR



- Object/scene recognition
- Use Unity's existing Machine Learning Tools, TensorFlow, etc
- Combine with Vuforia and other AR tracking
- See <https://www.youtube.com/watch?v=Q6ERFwQNkzo>

rubber eraser, rubber, pencil eraser - 0.21  
mousetrap - 0.07

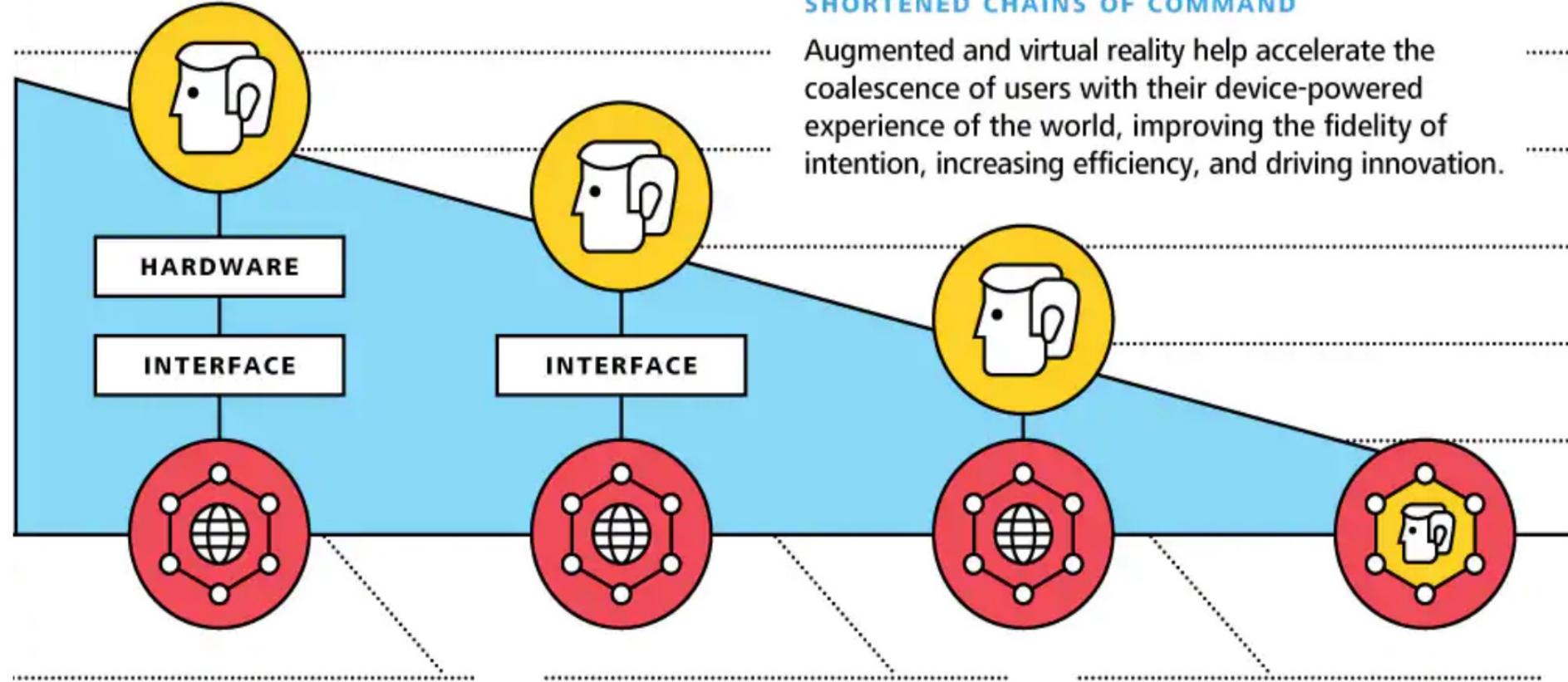


# SMART INTERFACES

---

## SHORTENED CHAINS OF COMMAND

Augmented and virtual reality help accelerate the coalescence of users with their device-powered experience of the world, improving the fidelity of intention, increasing efficiency, and driving innovation.

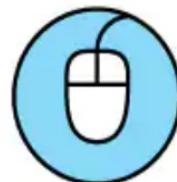


## MESSENGERS

Intermediate devices interact with interfaces; virtually all input occurs through a mouse or keyboard.



POINT



CLICK



TYPE

## SMART SCREENS

Screens manipulated based on environment facilitate direct physical or spoken interaction with displays.



TOUCH



SWIPE



TALK

## INTUITIVE INTERACTION

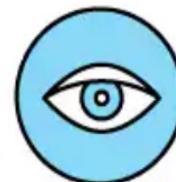
Devices respond to ambient cues and intentional movements to create empathetic, personalized experiences.



GESTURE



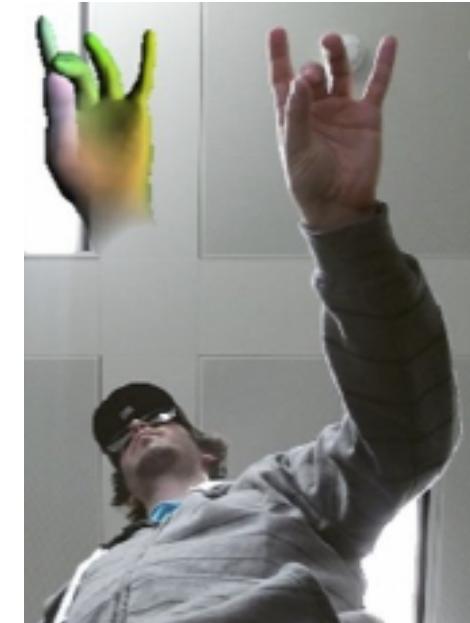
MOOD

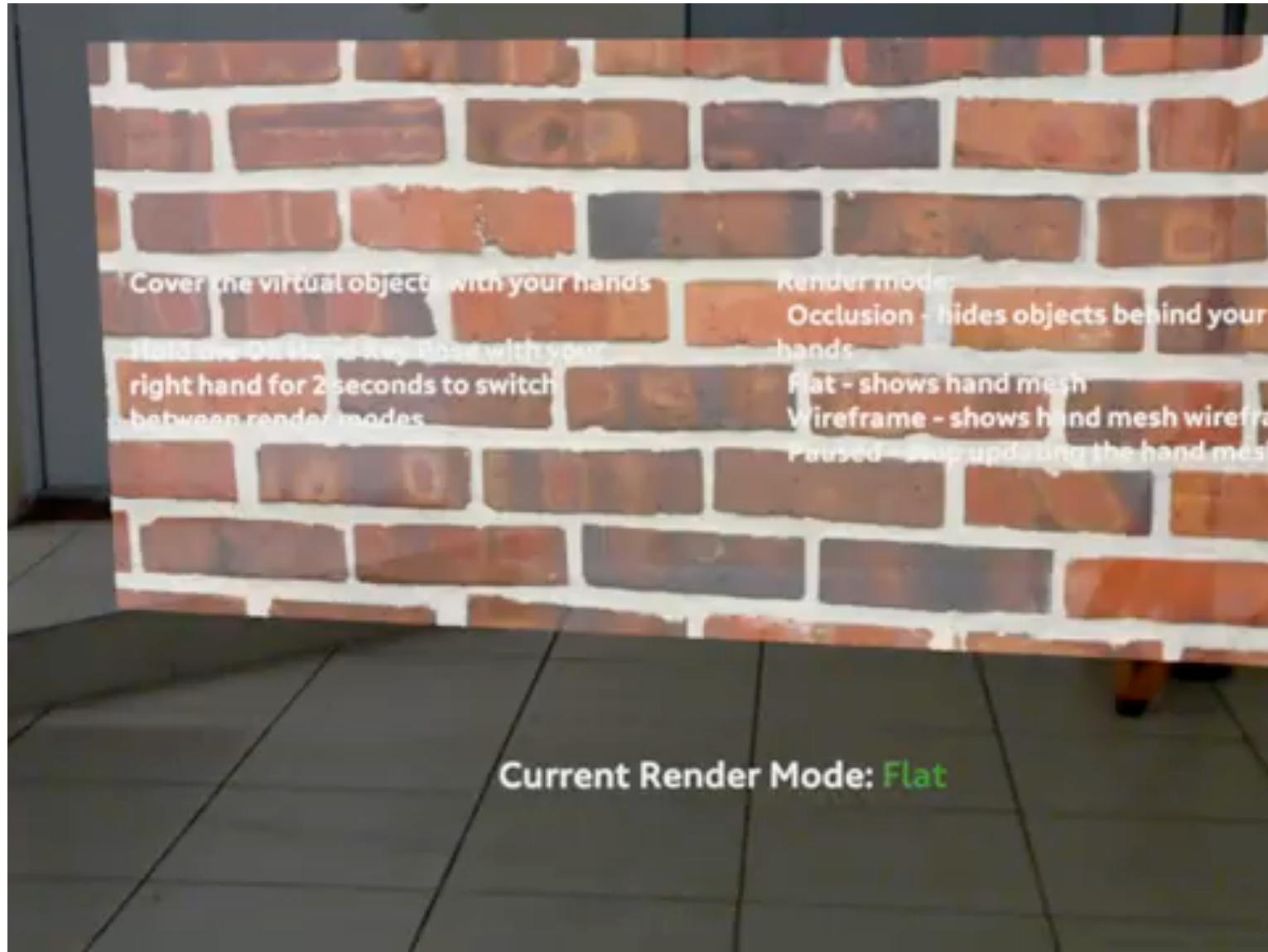


GAZE

# Natural Gesture

- Freehand gesture input
  - Depth sensors for gesture capture
  - Move beyond simple pointing
  - Rich two handed gestures
- Eg Microsoft Research Hand Tracker
  - 3D hand tracking, 30 fps, single sensor
- Commercial Systems
  - Meta, MS Hololens, Oculus, Intel, etc







<https://www.youtube.com/watch?v=LbIxKvbfEoo>

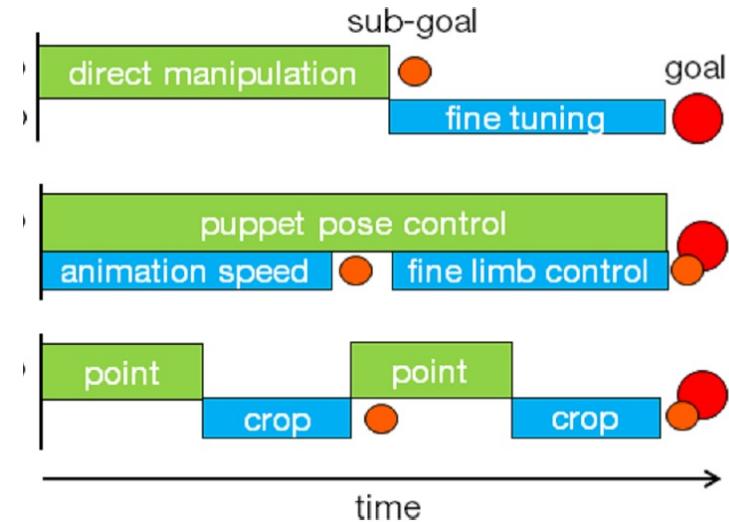
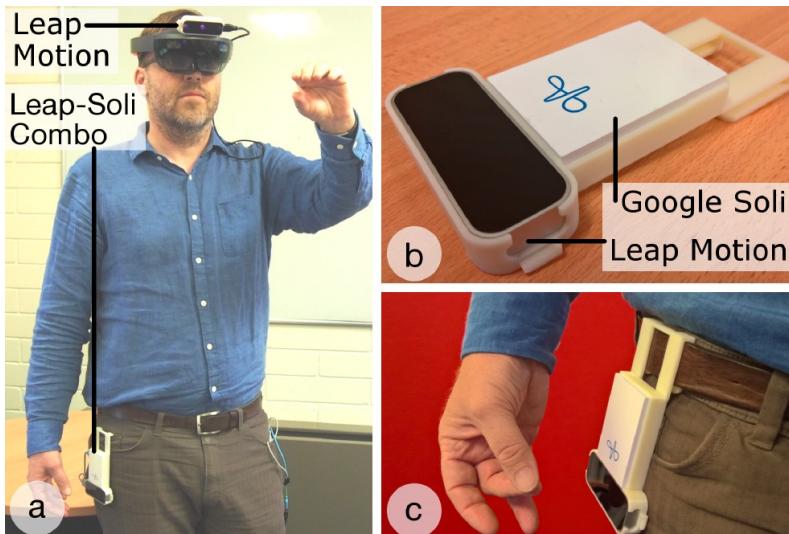
Microsoft  
Mixed Reality  
ToolKit



M R T K  
*MIXED REALITY  
TOOLKIT*

<http://aka.ms/MRTK>

# Multi-Scale Gesture



- Combine different gesture types
  - In-air gestures – natural but imprecise
  - Micro-gesture – fine scale gestures
- Gross motion + fine tuning interaction

Ens, B., Quigley, A., Yeo, H. S., Irani, P., Piumsomboon, T., & Billinghurst, M. (2018). Counterpoint: Exploring Mixed-Scale Gesture Interaction for AR Applications. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. LBW120). ACM.

# **Counterpoint: Exploring Mixed-Scale Gesture Interaction for AR Applications**

**CHI '18 Late Breaking Work**

**Barrett Ens**

**Aaron Quigley<sup>2</sup>**

**Hui-Shyong Yeo<sup>2</sup>**

**Pourang Irani<sup>3</sup>**

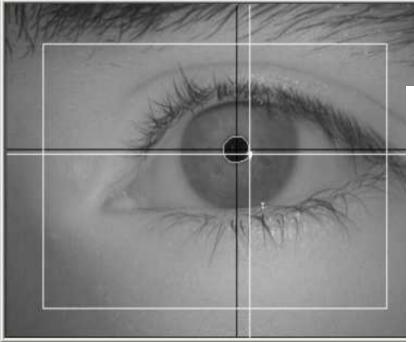
**Thammathip Piomsomboon<sup>1</sup>**

**Mark Billinghurst<sup>1</sup>**



<https://www.youtube.com/watch?v=TRfqNtt1VxY&t=23s>

# Eye Tracking Input



- Smaller/cheaper eye-tracking systems
- More HMDs with integrated eye-tracking
  - Hololens2, MagicLeap One
- Research questions
  - How can eye gaze be used for interaction?
  - What interaction metaphors are natural?
  - What technology can be used for eye-tracking?

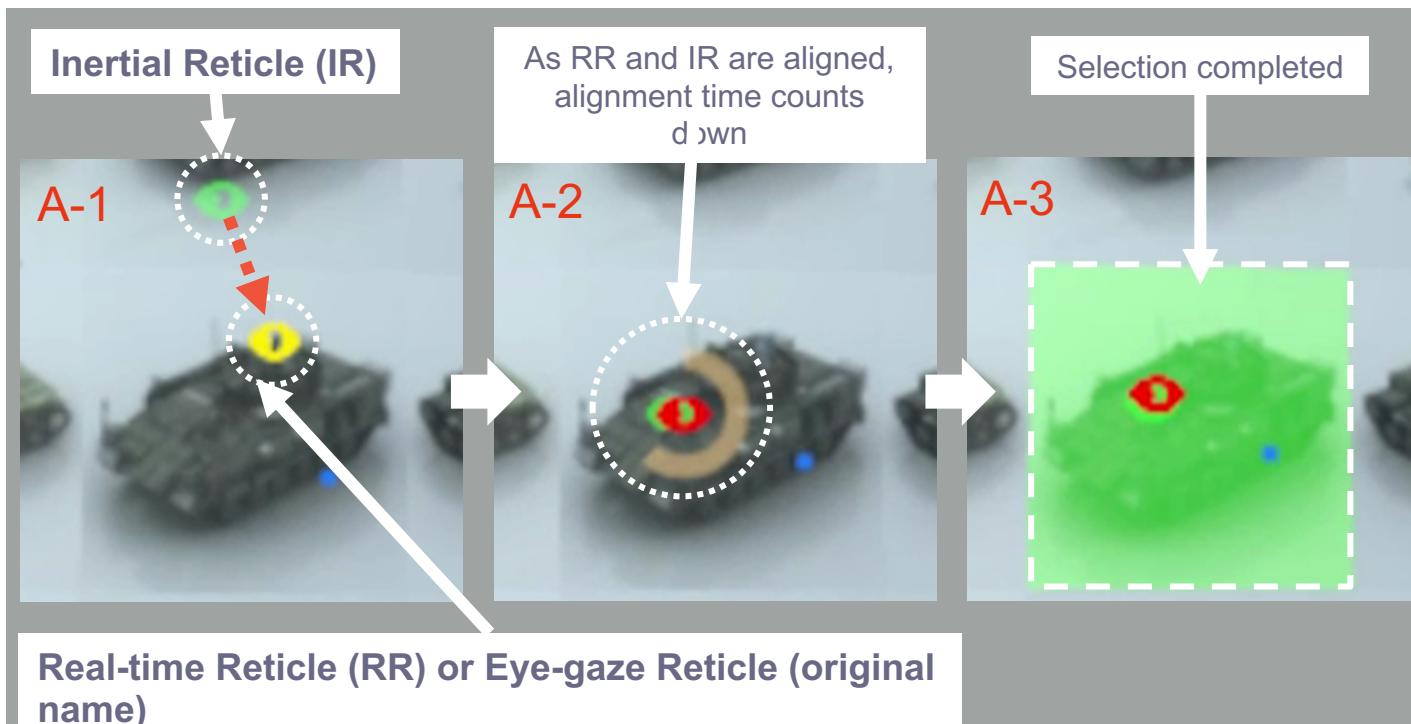
# Eye Gaze Interaction Methods

- Gaze for interaction
  - Implicit vs. explicit input
- Exploring different gaze interaction
  - Duo reticles – use eye saccade input
  - Radial pursuit – use smooth pursuit motion
  - Nod and roll – use the vestibular ocular reflex
- Hardware
  - HTC Vive + Pupil Labs integrated eye-tracking



Piumsomboon, T., Lee, G., Lindeman, R. W., & Billinghurst, M. (2017, March). Exploring natural eye-gaze-based interaction for immersive virtual reality. In *3D User Interfaces (3DUI), 2017 IEEE Symposium on* (pp. 36-39). IEEE.

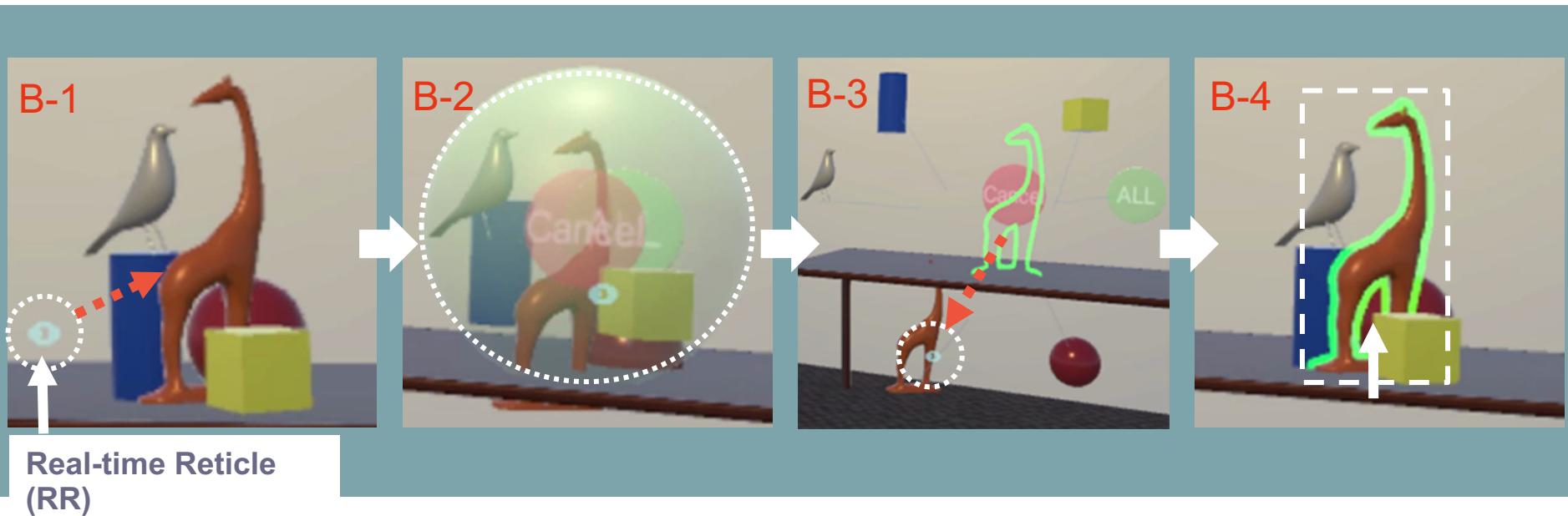
# Duo-Reticles (DR)



# Duo-Reticles (DR) – Video 1

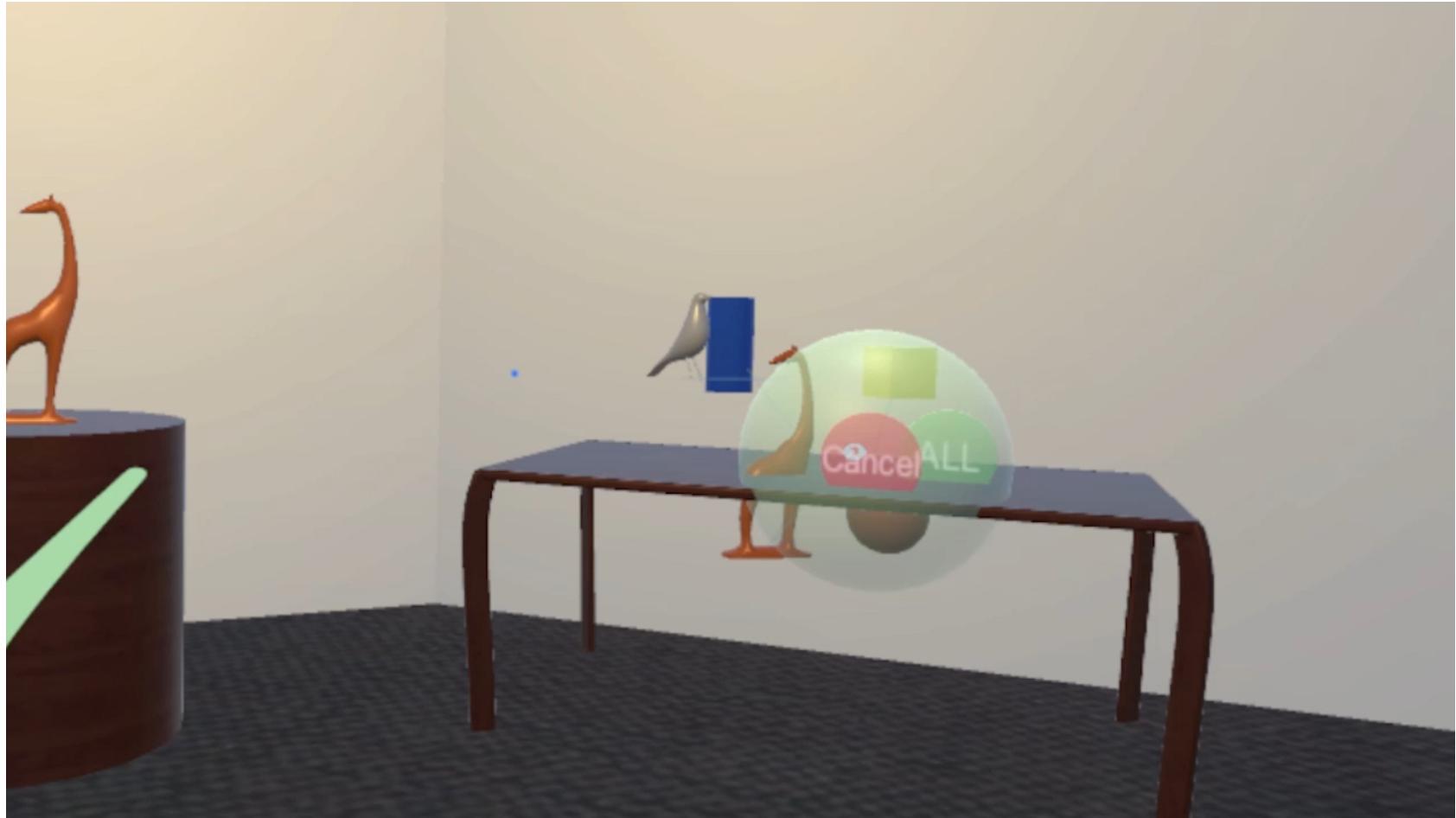


# Radial Pursuit (RP)

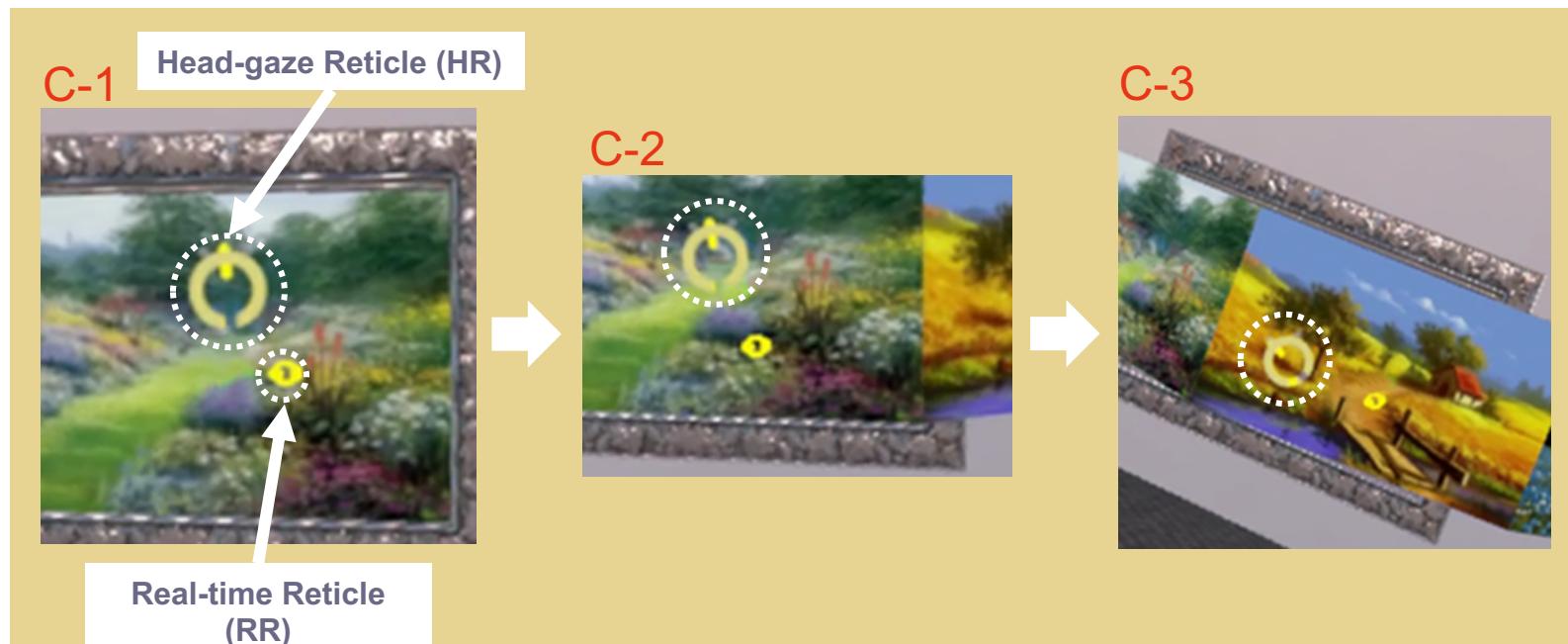


$$d_{min} = \min(d_0, d_1, \dots, d_k), d_i = \sum_{t=t_{initial}}^n |p(i)_t - p'_t|$$

# Radial Pursuit (RP) – Video 1



# Nod and Roll (NR)

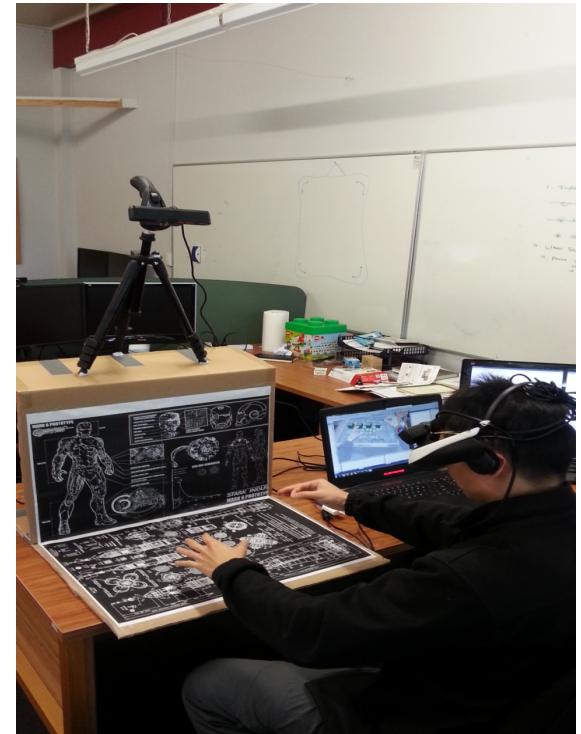


# Nod and Roll (NR) – Video 1



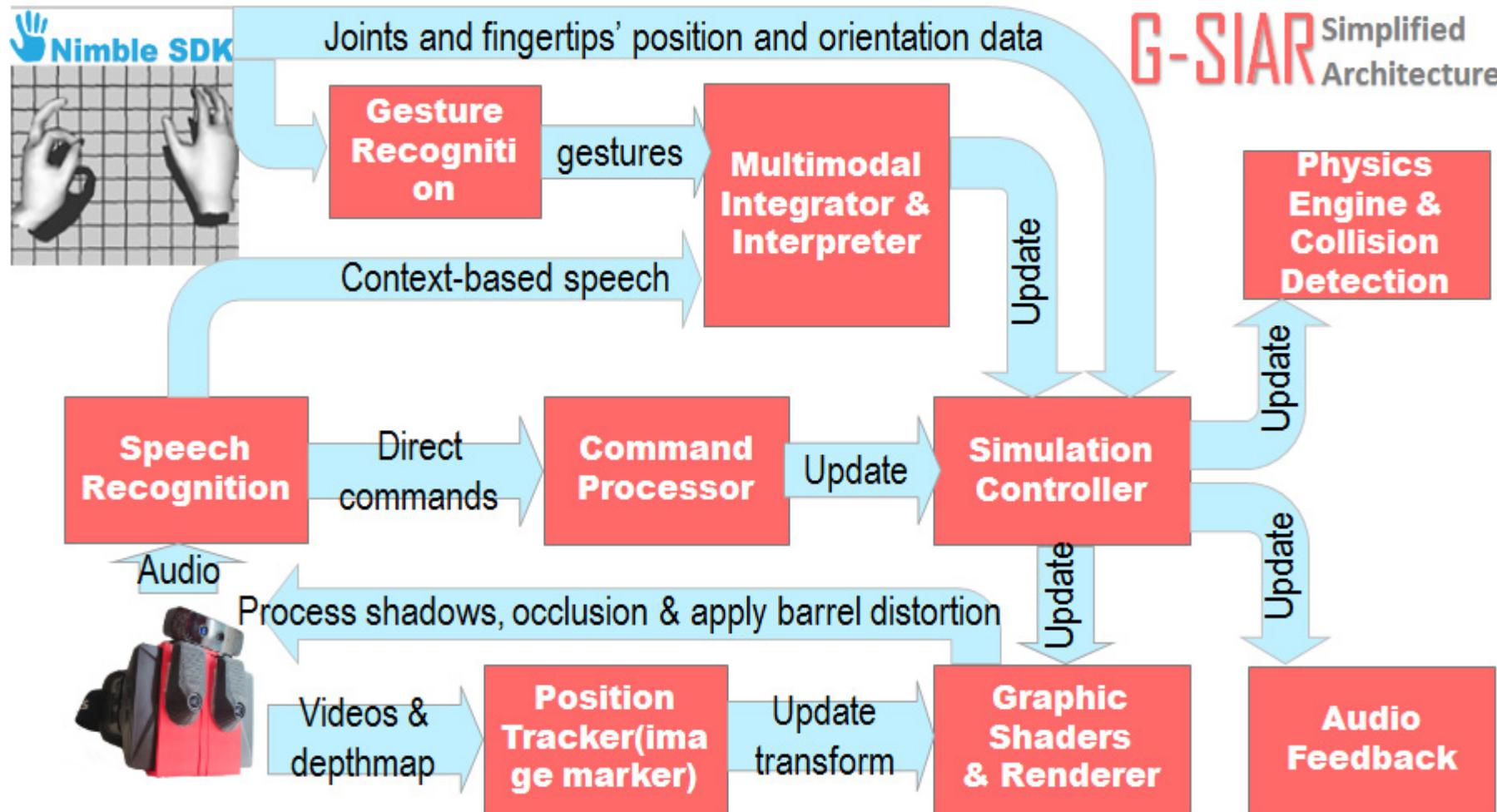
# Multimodal Input

- Combine gesture and speech input
  - Gesture good for qualitative input
  - Speech good for quantitative input
  - Support combined commands
    - “Put that there” + pointing
- E.g. HIT Lab NZ multimodal input
  - 3D hand tracking, speech
  - Multimodal fusion module
  - Complete tasks faster with MMI, less errors



Billinghurst, M., Piumsomboon, T., & Bai, H. (2014). Hands in Space: Gesture Interaction with Augmented-Reality Interfaces. *IEEE computer graphics and applications*, (1), 77-80.

# G-SIAR Simplified Architecture



# HIT Lab NZ Multimodal Input



# RAPID PROTOTYPING

---

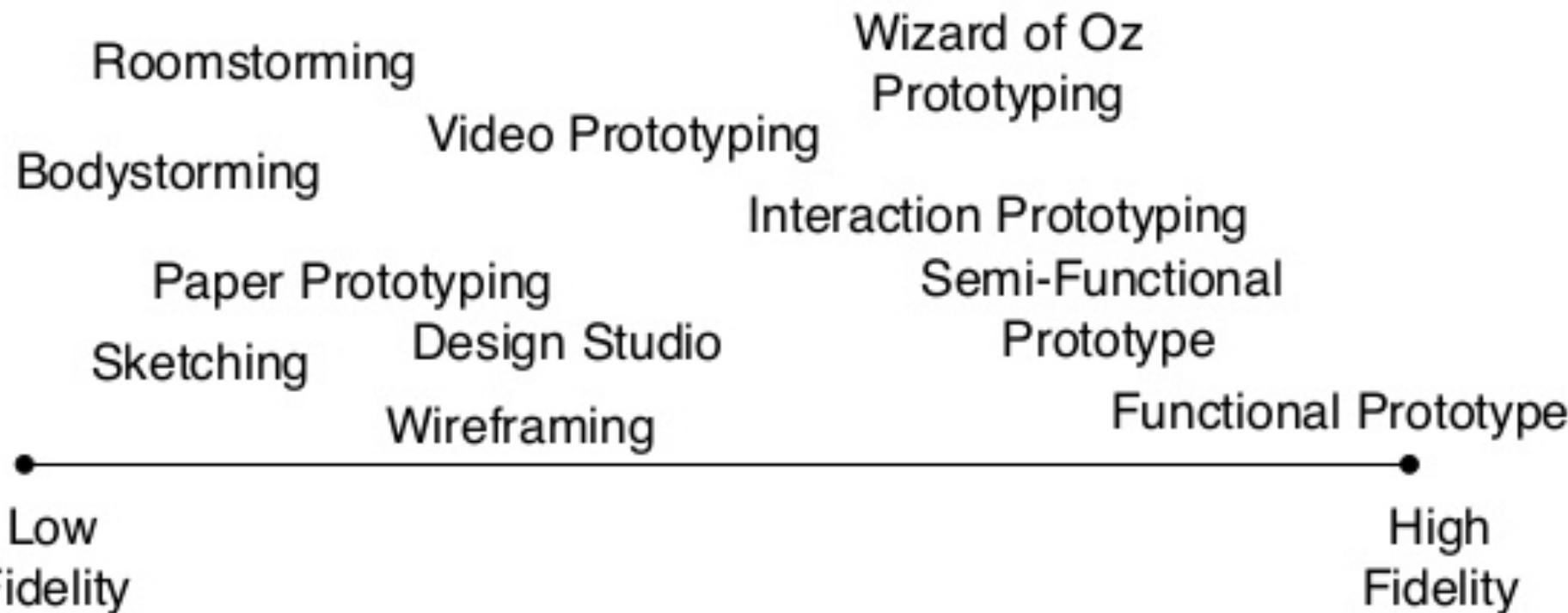
*How can we quickly  
prototype interactive  
AR and AI experiences  
with little or no coding?*



# Typical Prototyping Tools

- **Static/Low fidelity**
  - Sketching
  - User interface templates
  - Storyboards/Application flows
- **Interactive/High fidelity**
  - Wireframing tools
  - Mobile prototyping
  - Native Coding

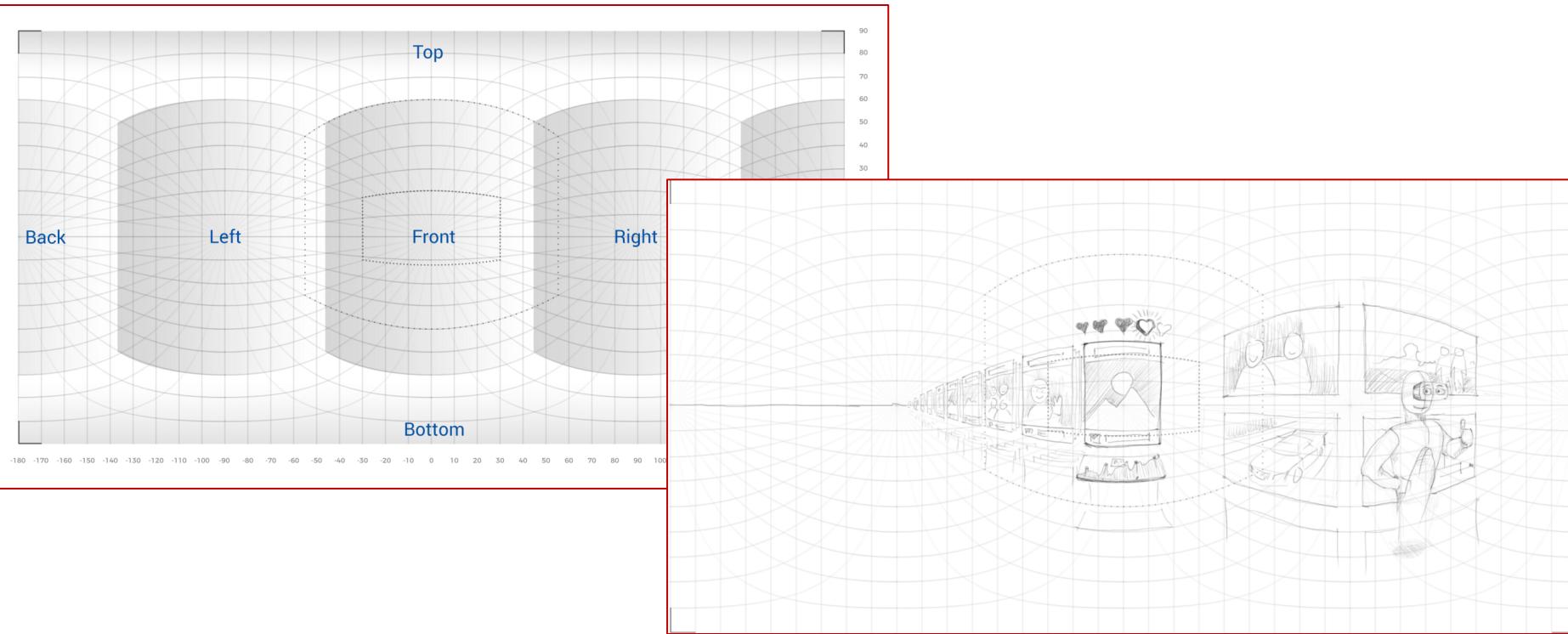
# Design/Prototyping Tools



# VR Prototyping Tools

- **Low Fidelity**
  - Sketched Paper Interfaces – pen/paper, non-interactive
  - Onride Photoshop tool – digital, non-interactive
  - InstaVR - 360 web based tool, simple interactivity
  - SketchBox – create VR interface inside VR
- **High Fidelity**
  - Entiti – template based VR with visual programming
  - A-Frame – web based VR tool using HTML
  - EditorVR – Unity wrapper inside VR
  - Unity/Unreal Game Engine – programming needed

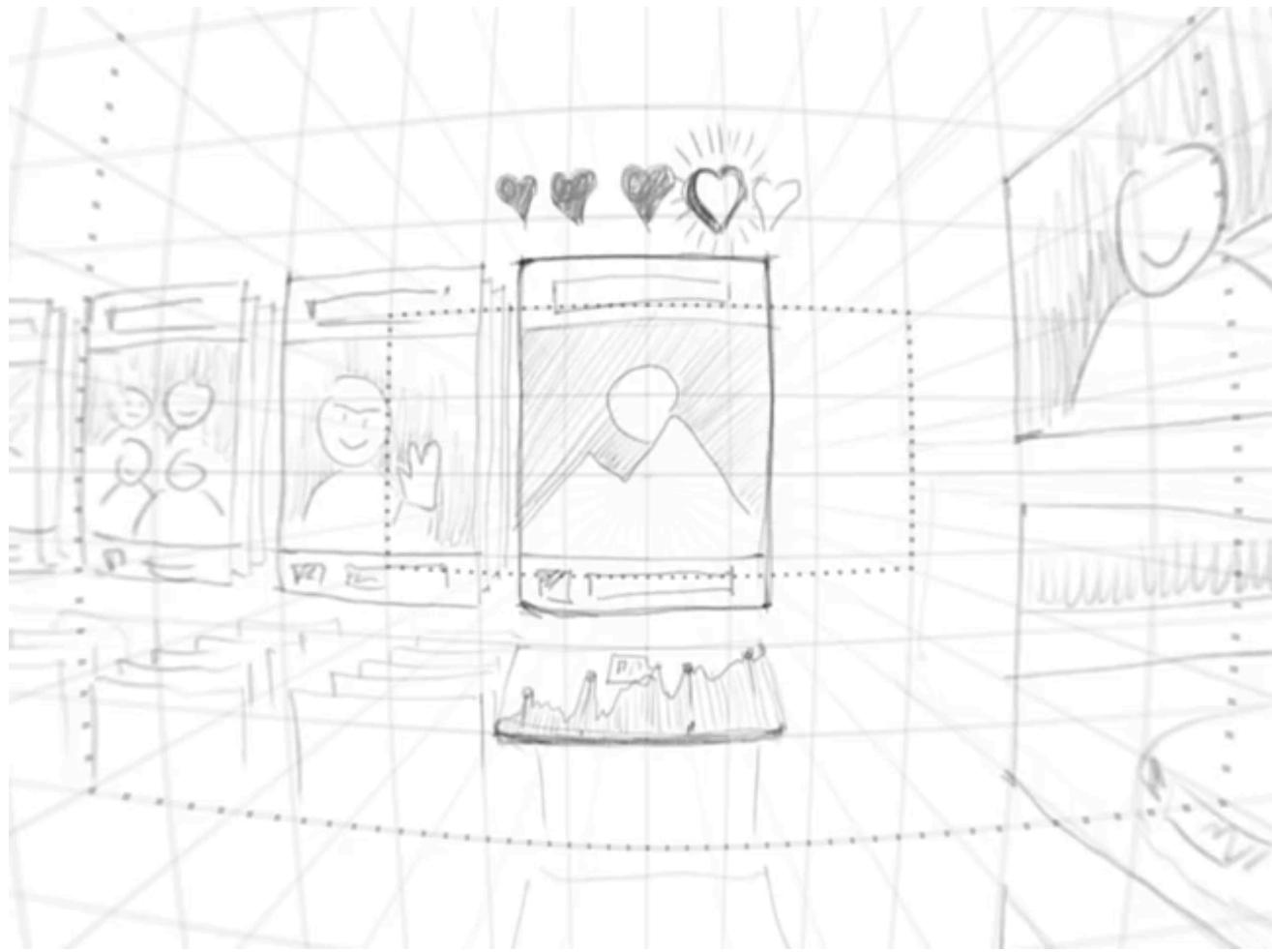
# Sketching VR Interfaces



- Download 360 panorama template grid
- Draw interface ideas into grid
- Scan into 360 photo viewer for VR HMD

See <https://virtualrealtypop.com/vr-sketches-56599f99b357>

# Example Sketched VR Interface



- <https://www.youtube.com/watch?v=BmMh6-jPWOc>

# ONIRIDE - 360° Art Plugin for Photoshop



- Draw 360 panorama's directly in Photoshop
- Preview in Photoshop, export to VR
- See <http://www.oniride.com/360art>

# OnRide Demo



- <https://www.youtube.com/watch?v=1P1EfGizal0>

# InstaVR

The screenshot shows the InstaVR website homepage. At the top, there's a purple header bar with social media icons (Facebook, Twitter, YouTube) and a "Sign In" button. Below the header, the InstaVR logo is on the left, followed by navigation links: Customer Stories, Industries, Blog, Pricing, and a prominent "TRY FOR FREE" button. To the right is a search icon. The main content area features three examples of VR content: a 360-degree interior view of a modern house, a collage of various outdoor scenes, and a multi-panel architectural rendering. Each example is accompanied by its respective VR platform logos: VIVE, Gear VR (Powered by Oculus), iOS, Android, and Web VR. To the right of these examples, a woman is shown wearing a VR headset and gesturing with her hands, illustrating the interactive nature of the content.

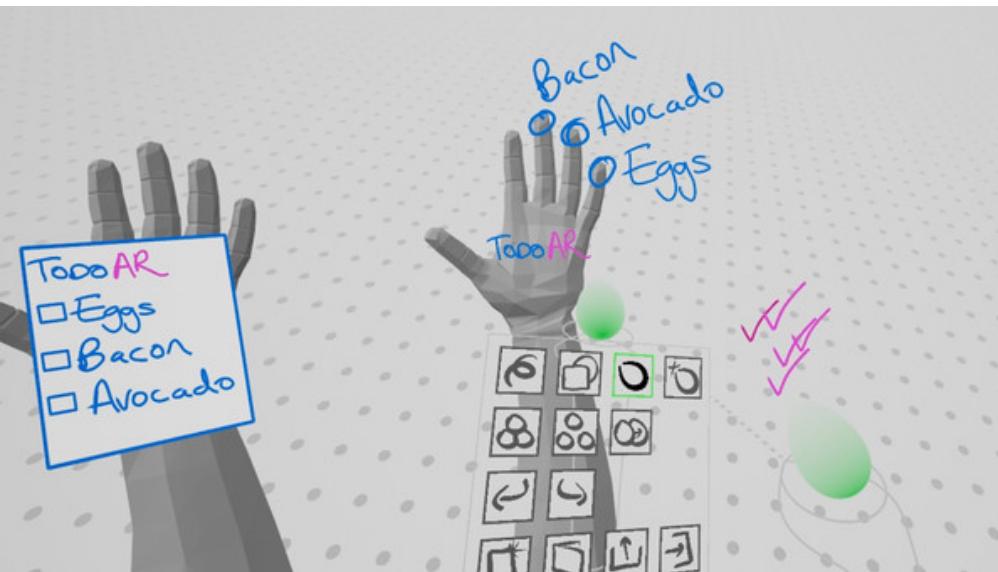
- <http://www.instavr.co/>
- Free, fast panorama VR, deploy to multi platforms

# Demo - Using InstaVR



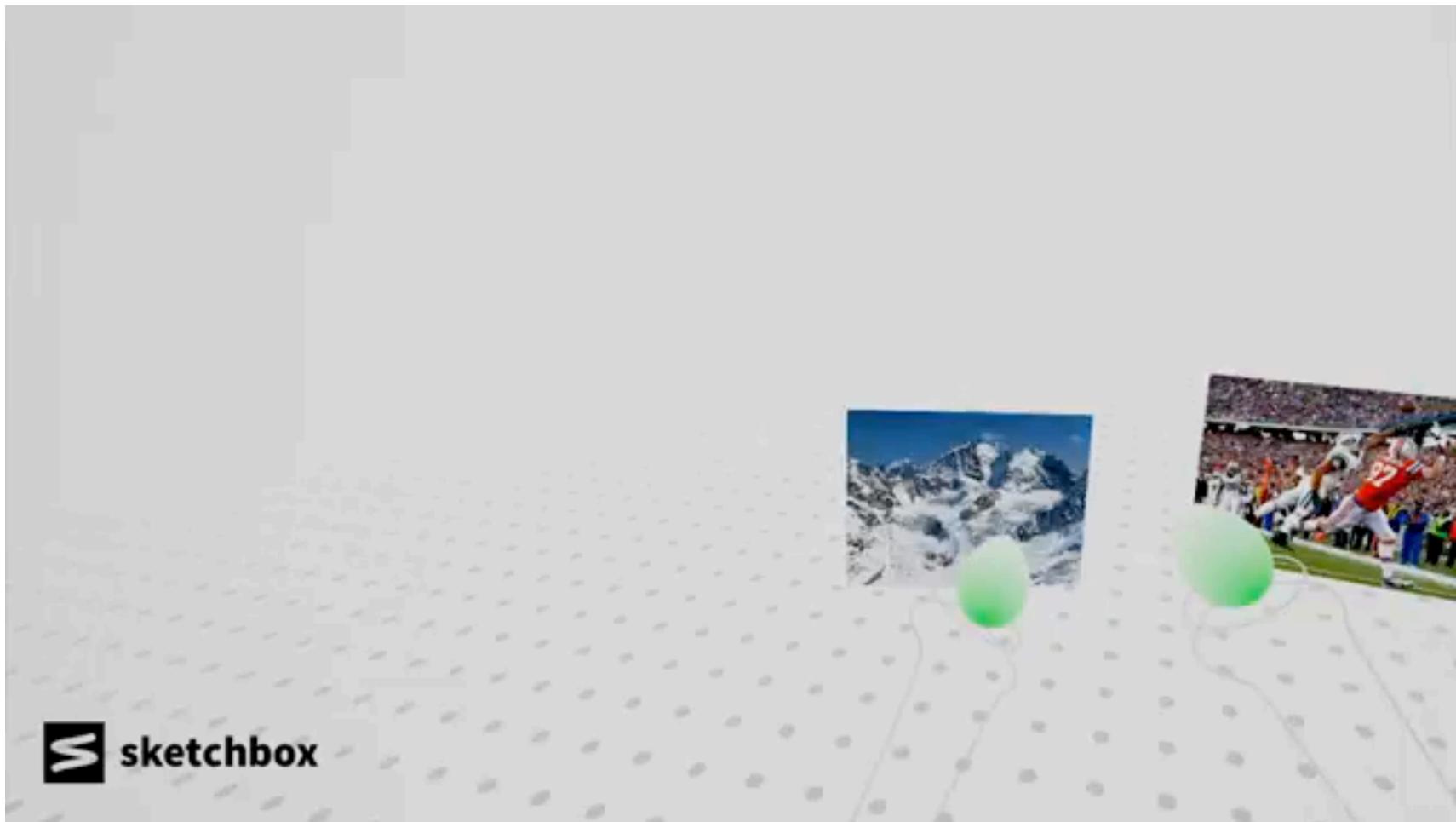
- <https://www.youtube.com/watch?v=M2C8vDL0YeA>

# Sketchbox



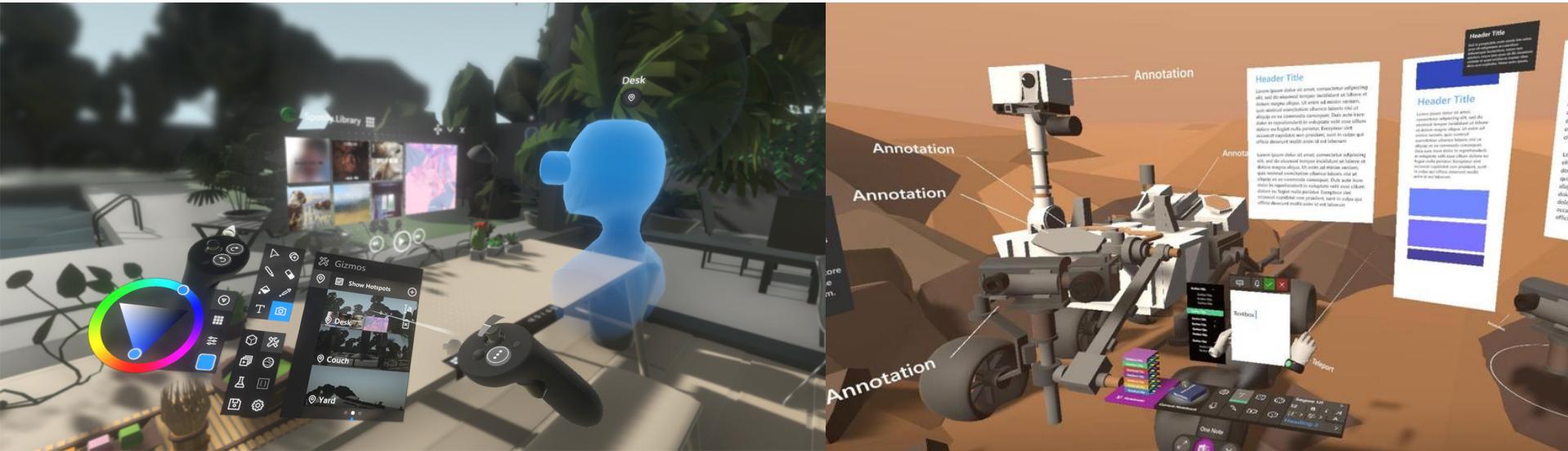
- VR design tool - create VR interface inside VR
  - Support for HTC Vive, Oculus Rift
- Easy to use VR sketching tool
- Available from SteamVR
- See <https://www.sketchboxvr.com/>

# Sketchbox Demo

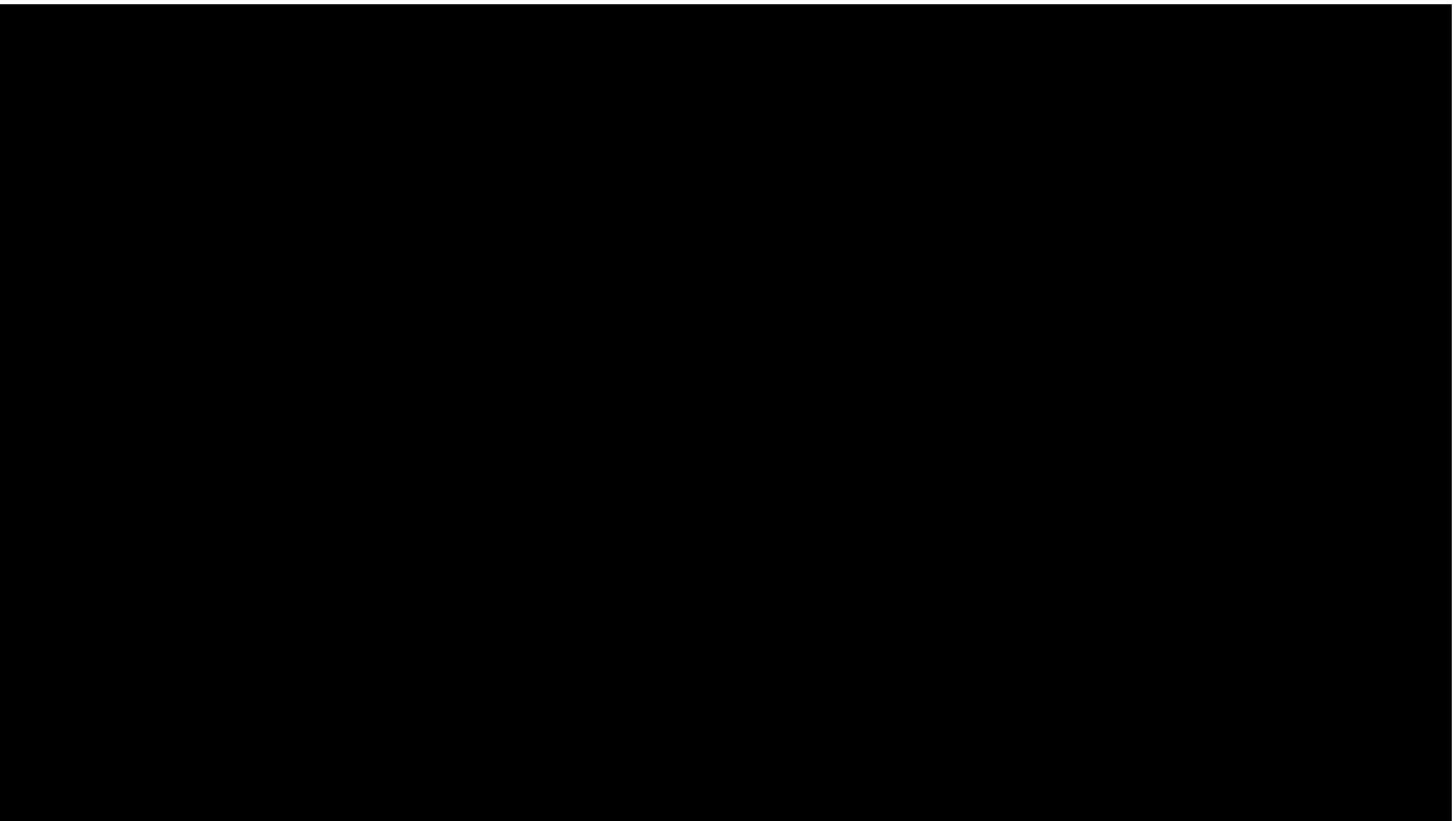


- <https://www.youtube.com/watch?v=gWfgewGzaEI>

# Microsoft Marquette

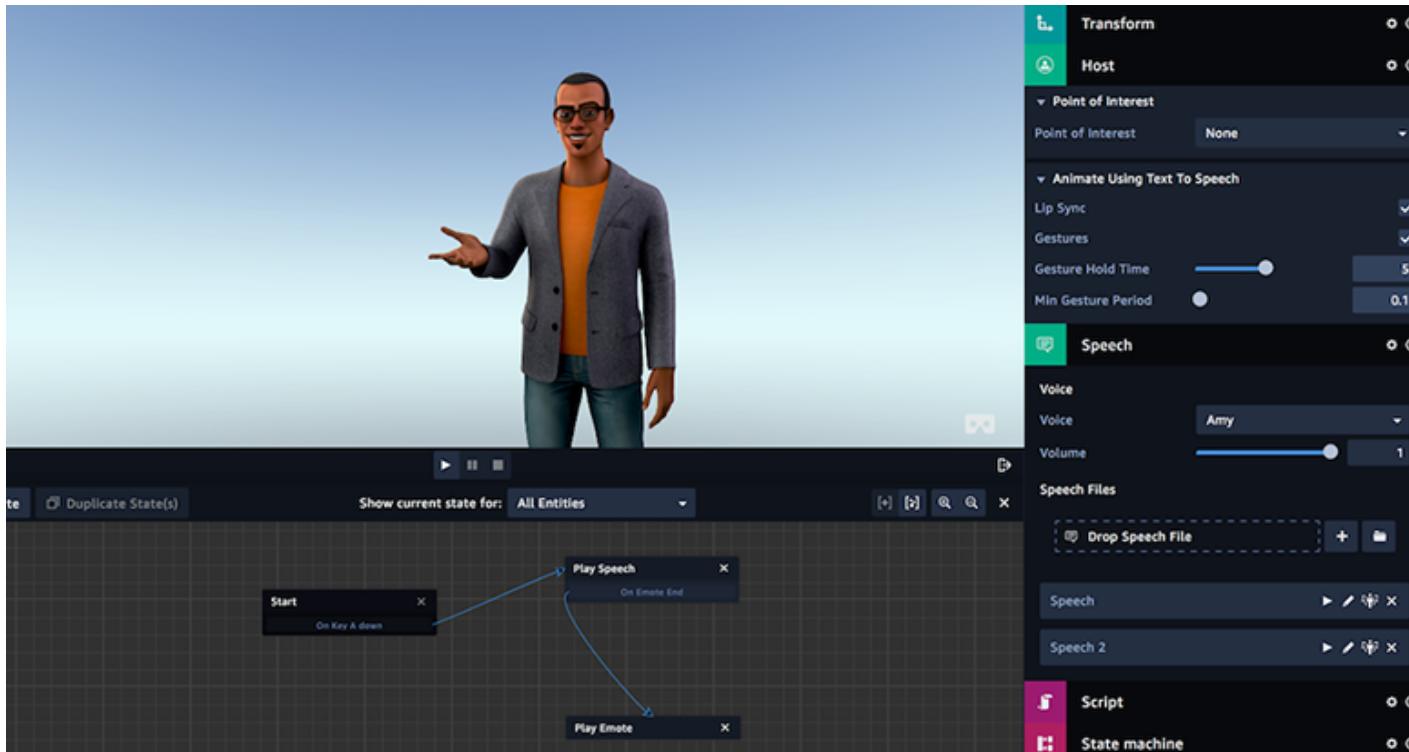


- Prototype VR interfaces from inside VR
- 3D UI for spatial prototyping
- Bring content into Unity with plug-in
- <https://www.maquette.ms/>

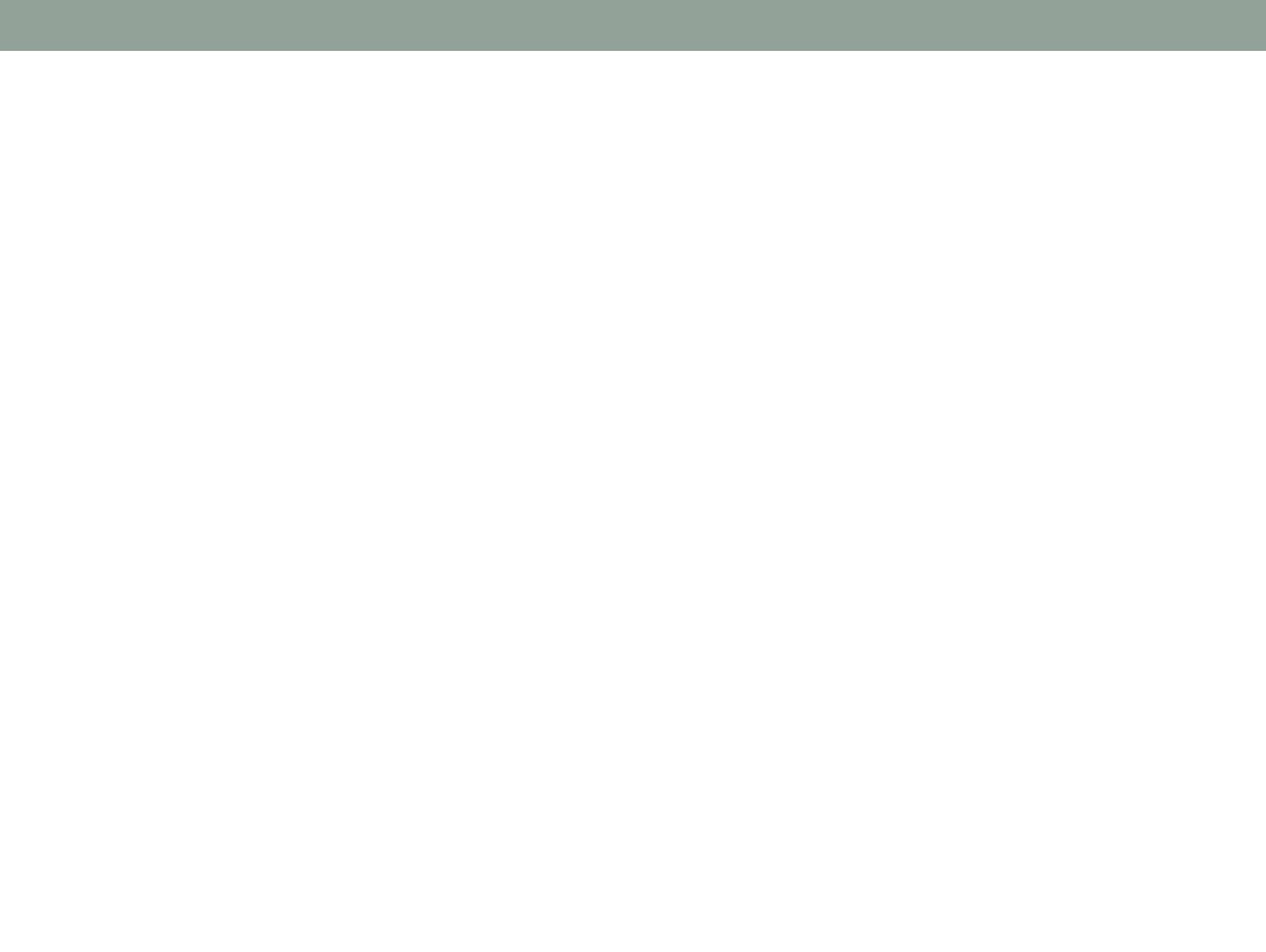


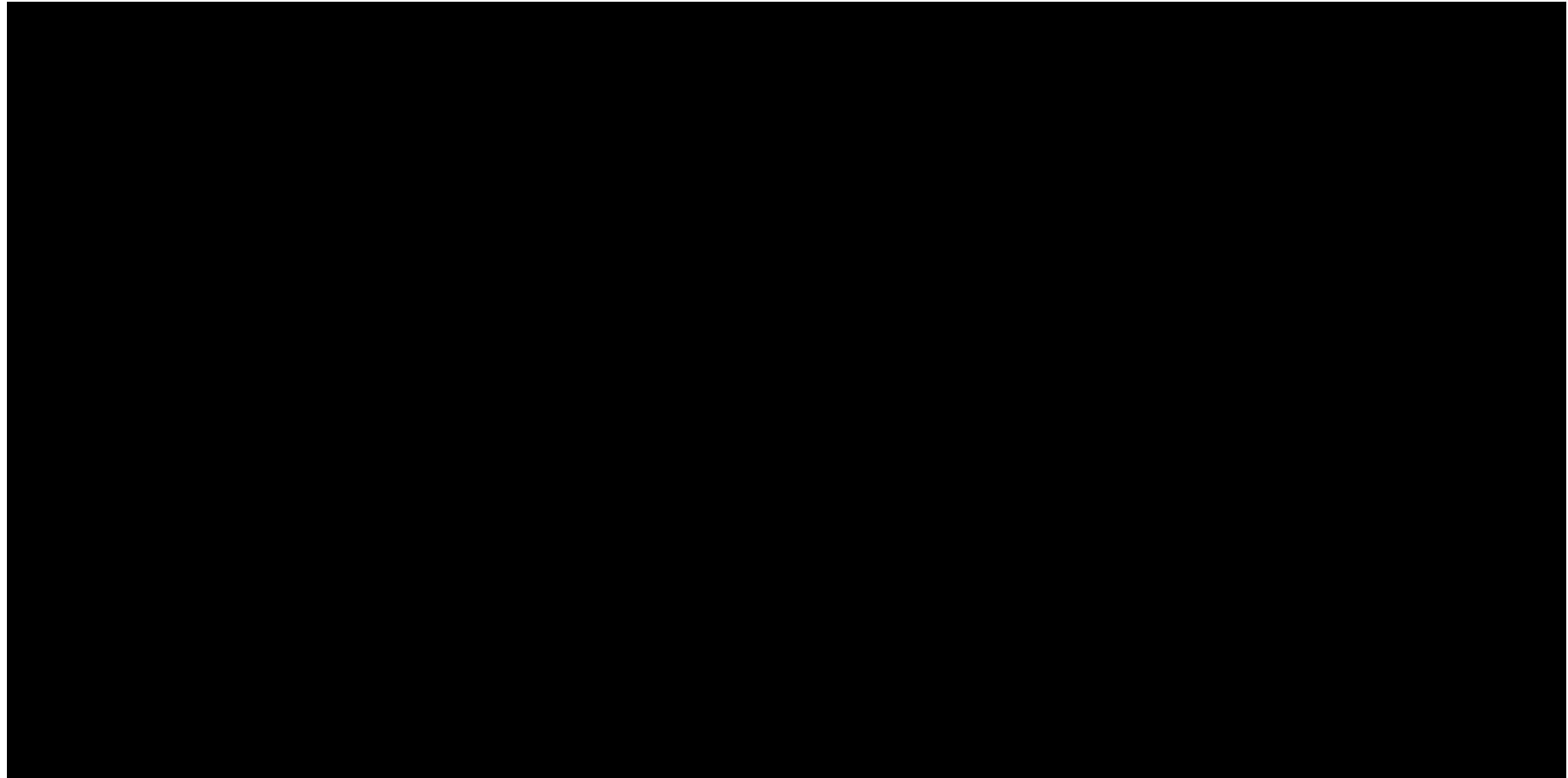
- <https://www.youtube.com/watch?v=capxS6C6ooY>

# Amazon Sumerian



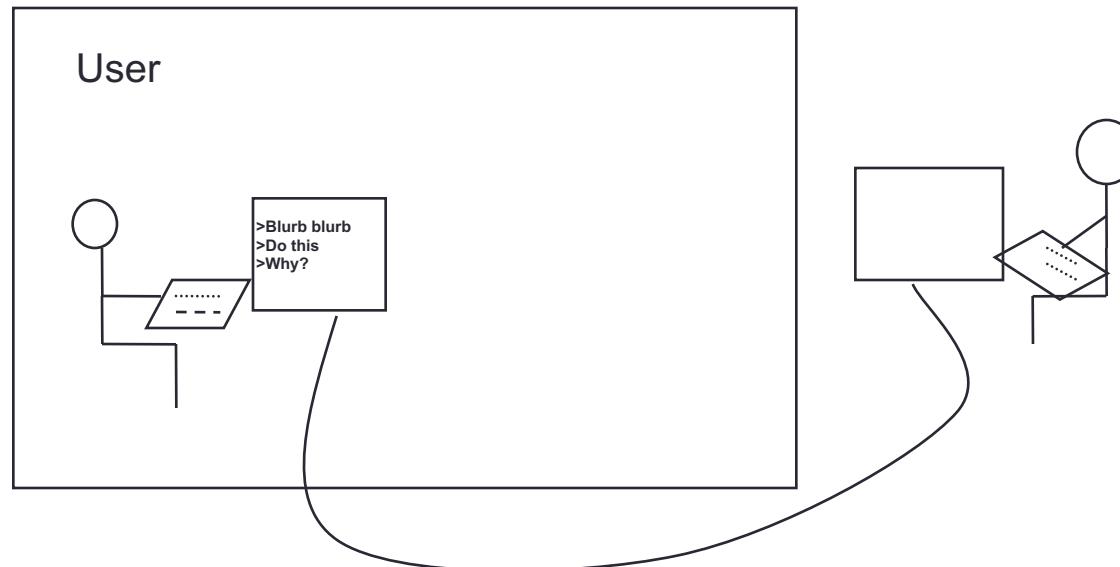
- <https://aws.amazon.com/sumerian/>
- Build AR/VR with no programming
- Great support for characters/conversation





# ‘Wizard-of-Oz’ Prototyping

- The user thinks they are interacting with a computer, but a developer is responding to output rather than the system.
- Usually done early in design to understand users’ expectations



# Role Playing



[http://www.youtube.com/watch?v=l8biQ8dfC\\_U&list=PL93019A1526FC1BE6&index=36&feature=plpp\\_video](http://www.youtube.com/watch?v=l8biQ8dfC_U&list=PL93019A1526FC1BE6&index=36&feature=plpp_video)

# Experience Prototyping

role playing

# RESEARCH DIRECTIONS

---

# Research Needed in Many Areas

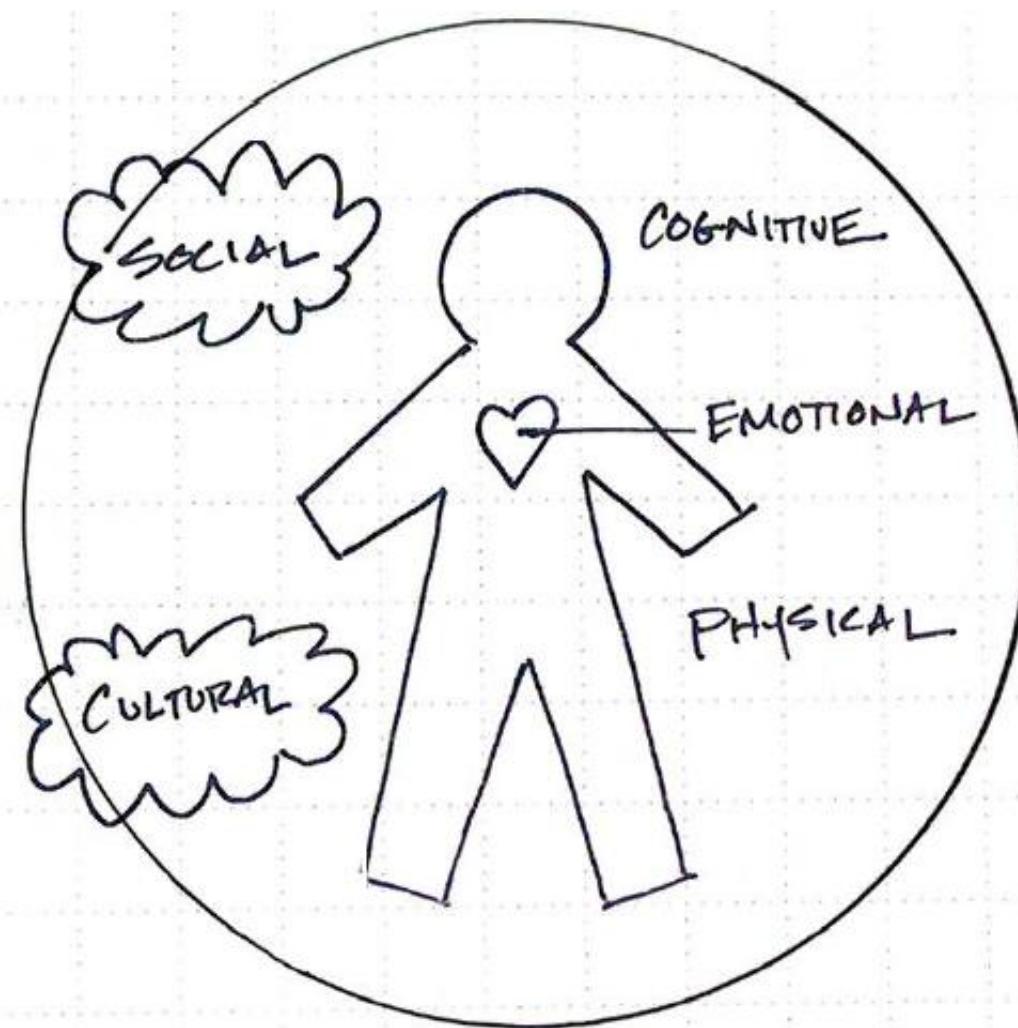
- Collaborative Experiences
  - AR teleconferencing
- Social Acceptance
  - Overcome social problems with AR
- Cloud Services
  - Cloud based storage/processing
- AR Authoring Tools
  - Easy content creation for non-experts
- Etc..

# Example: Social Acceptance

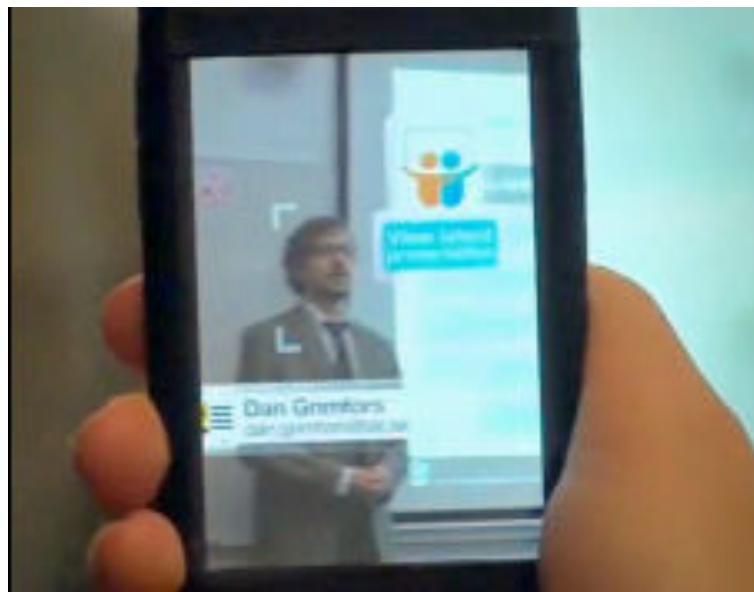


- People don't want to look silly
  - Only 12% of 4,600 adults would be willing to wear AR glasses
  - 20% of mobile AR browser users experience social issues
- Acceptance more due to Social than Technical issues
  - Needs further study (ethnographic, field tests, longitudinal)

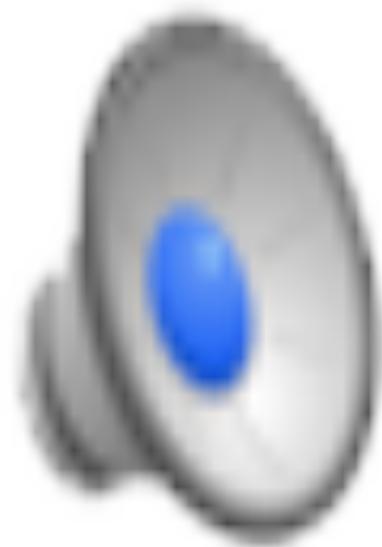
# Consider the Whole User



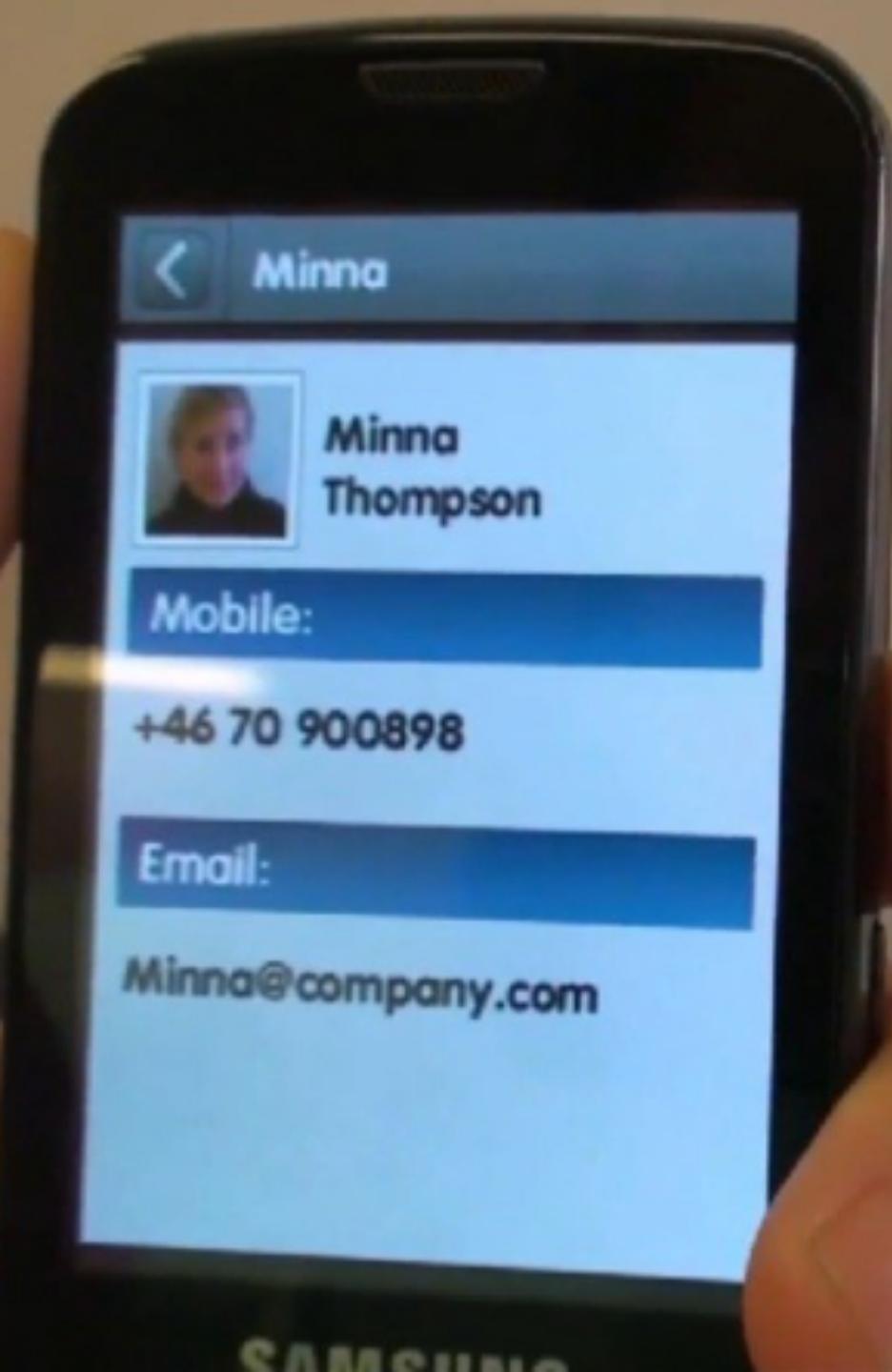
# TAT Augmented ID



# TAT AugmentedID



# Social Pattern



Like being stopped by  
the police for ID.

Or security scanned!

“Show me your papers.”

# Experience



Minna



Minna  
Thompson

Mobile:

+46 70 900898

Email:

Minna@company.com

“Anyone pointing a device in my direction to try to identify me better be prepared for either a **law suit**, or a **punch in the face.**”

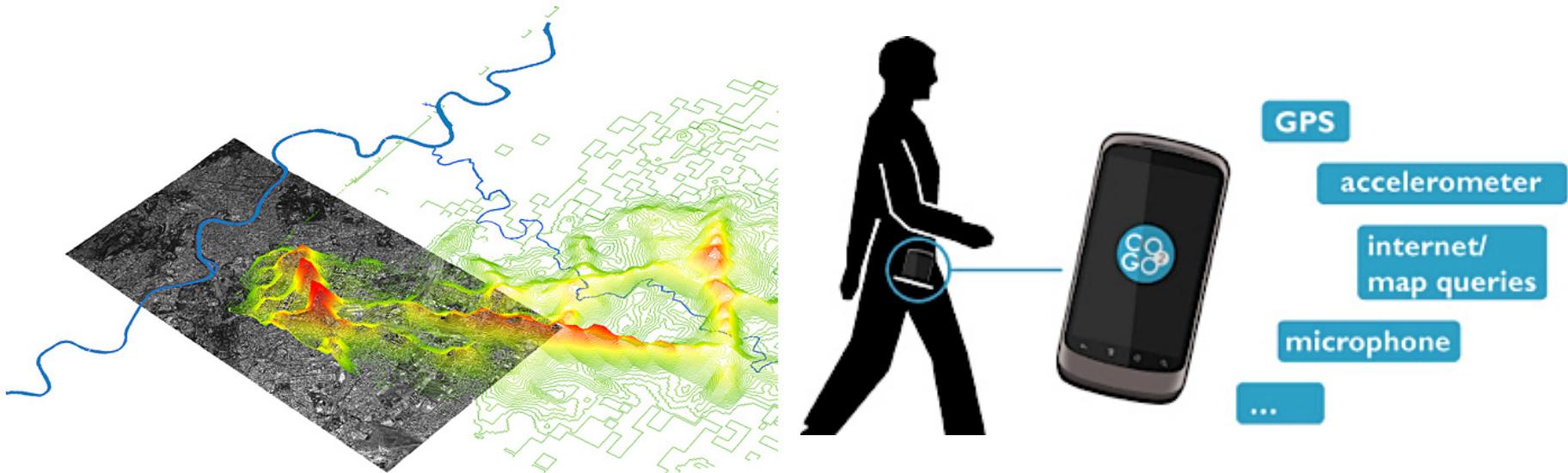
Anonymous Comment

# Scaling Up



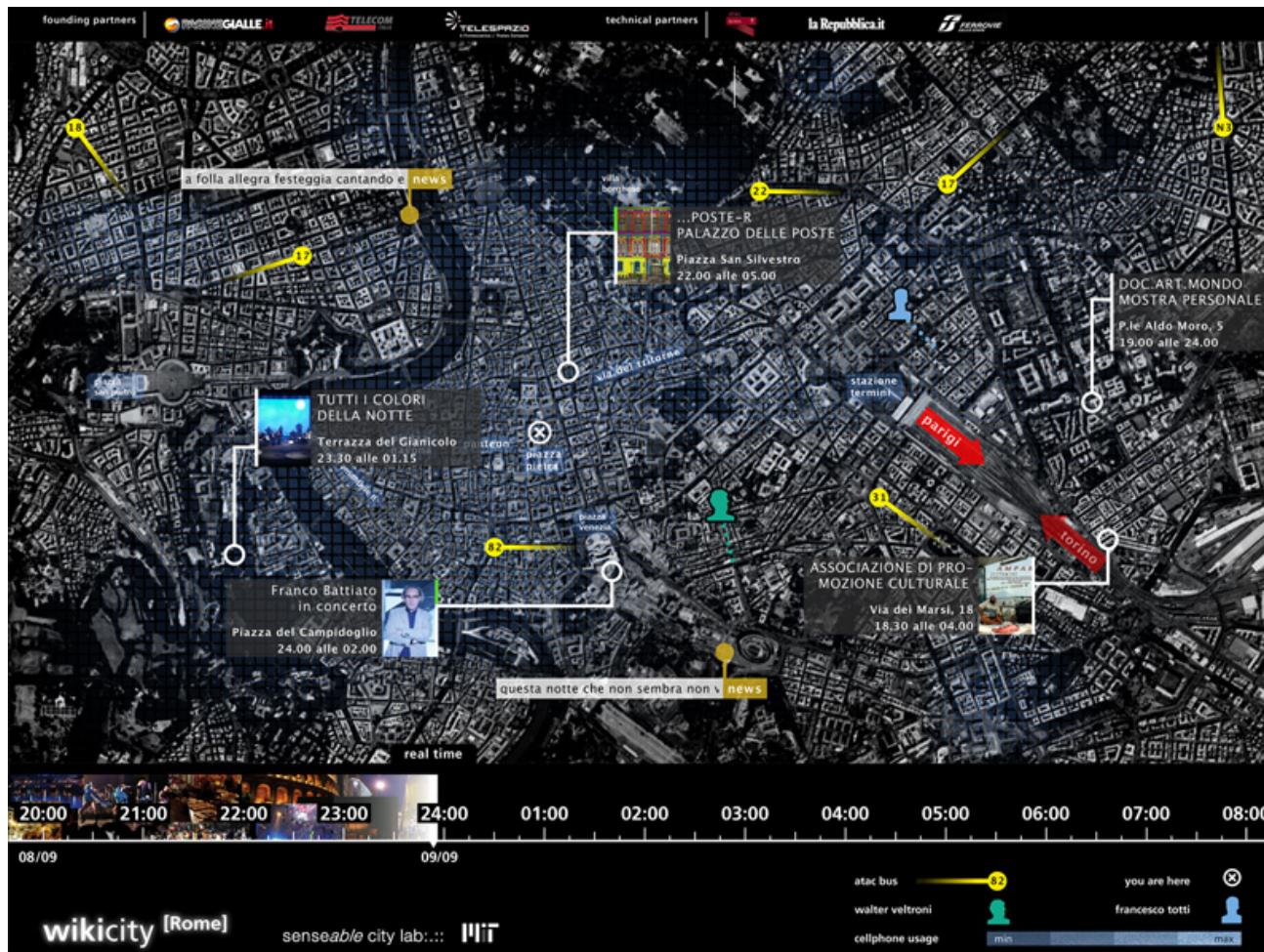
- Seeing actions of millions of users in the world
- Augmentation on city/country level

# AR + Smart Sensors + Social Networks



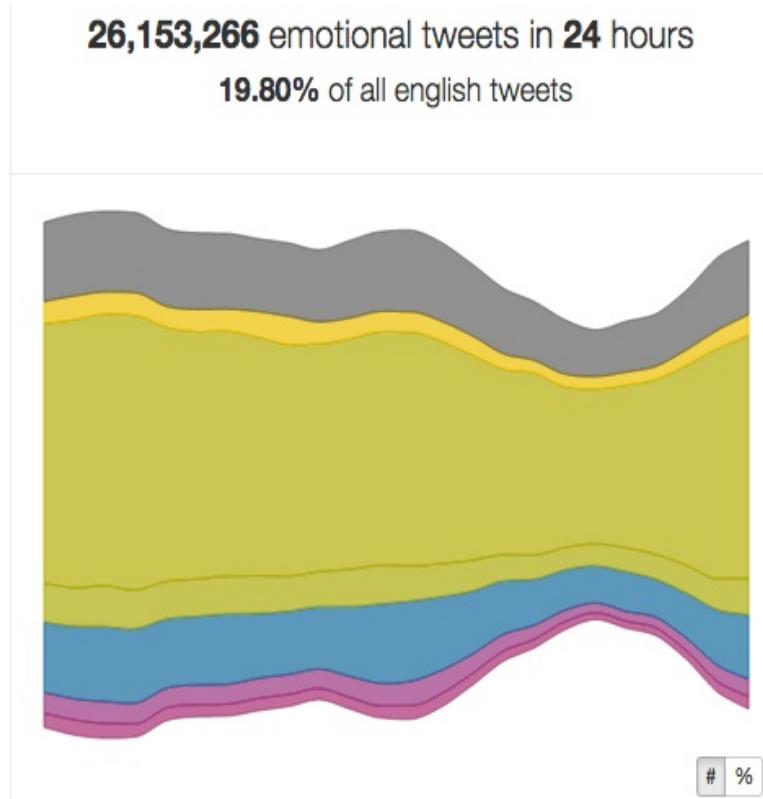
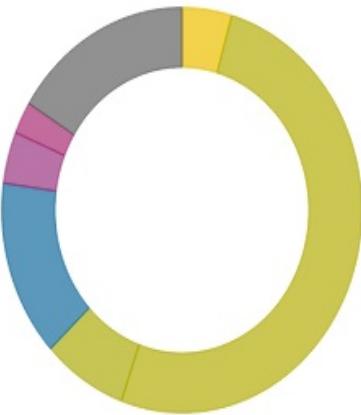
- Track population at city scale (mobile networks)
- Match population data to external sensor data
- Mine data for applications

# Example: MIT SENSEable City Lab



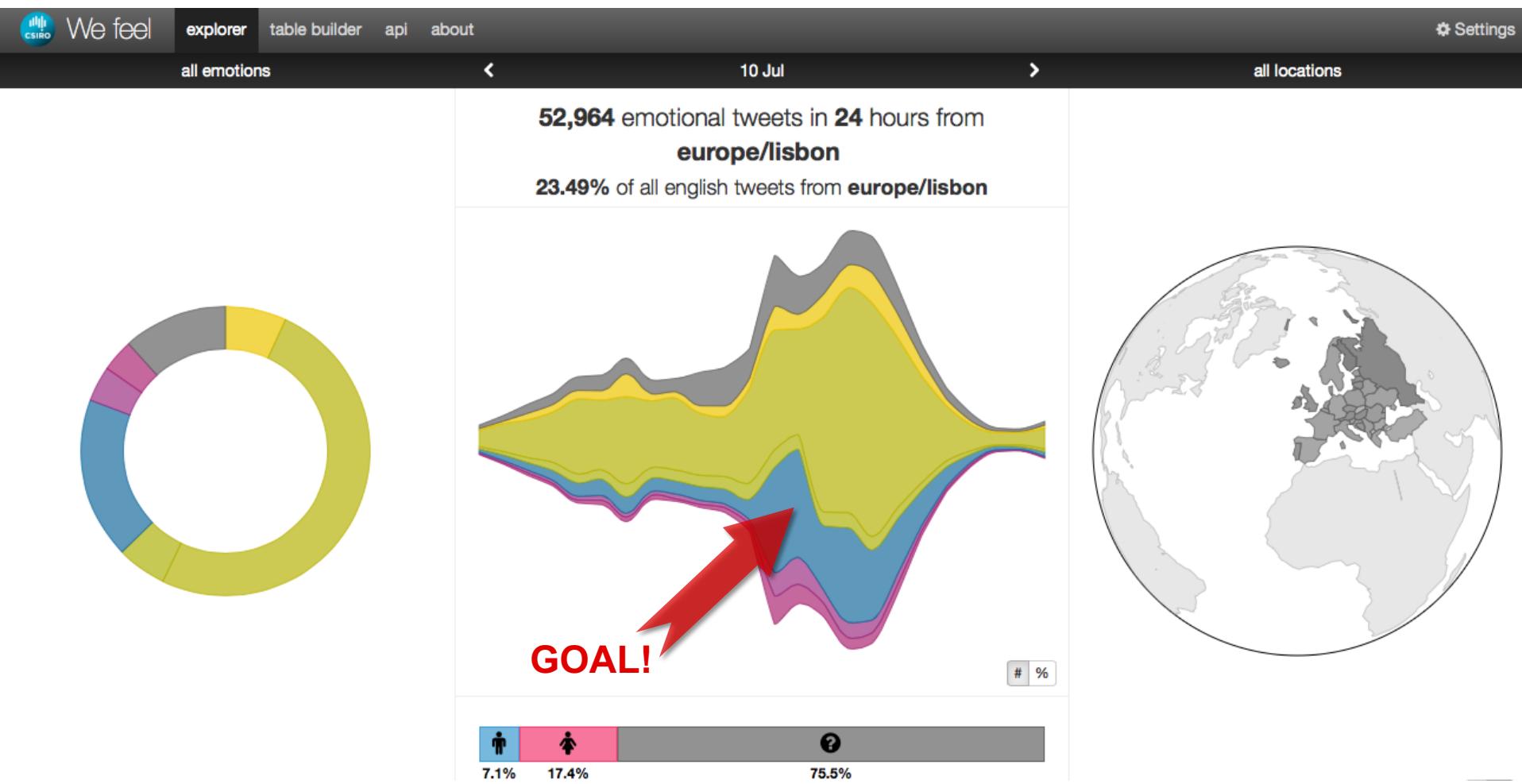
<http://senseable.mit.edu/wikicity/rome/>

# Example: CSIRO WeFeel Tool



- Emotionally mining global Twitter feeds
- <http://wefeel.csiro.au>

# European Cup – July 10th





---

# Empathic Computing Lab

Systems That Create Understanding



University of  
South Australia

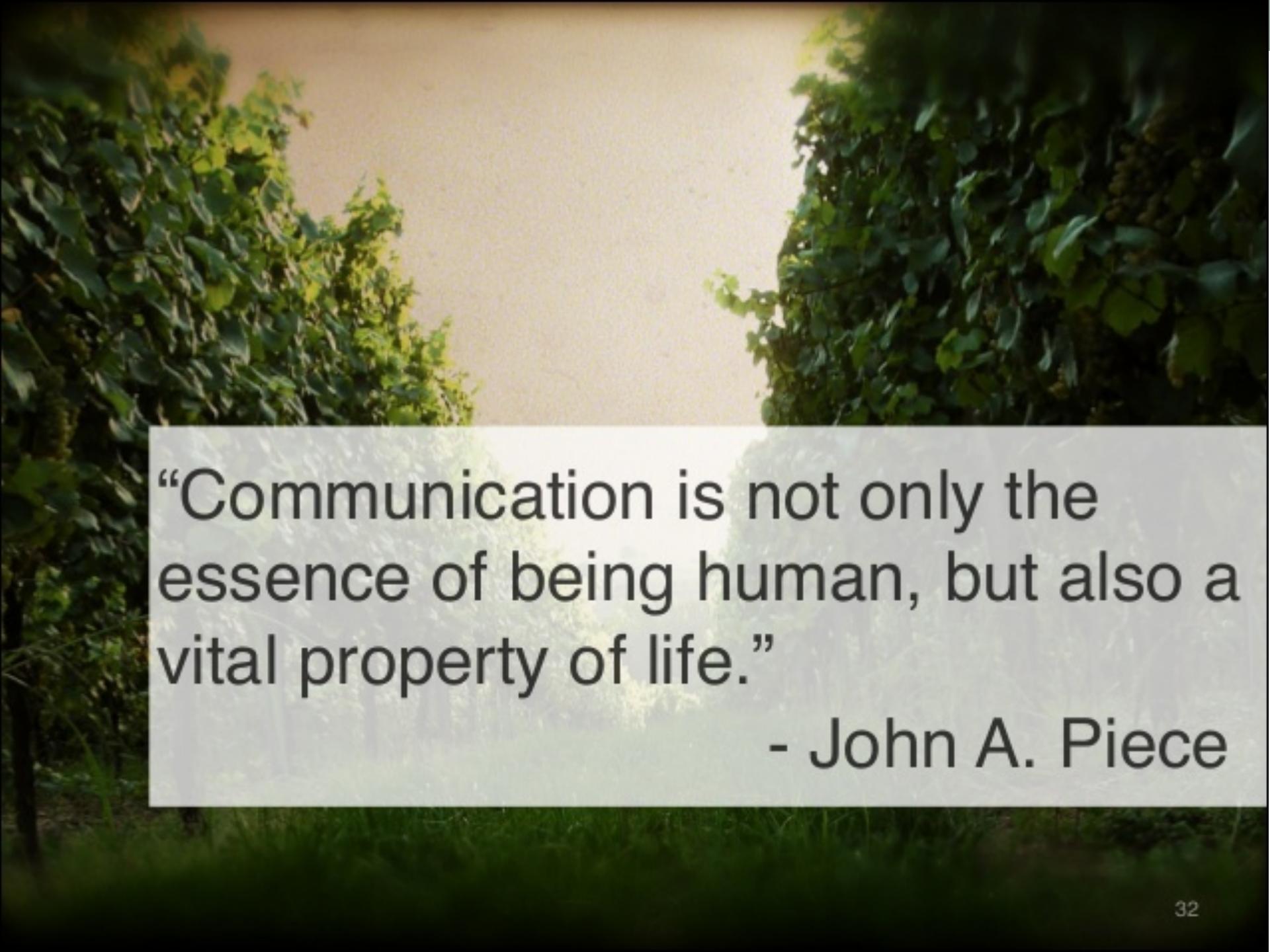
Mark Billinghurst

[mark.billinghurst@auckland.ac.nz](mailto:mark.billinghurst@auckland.ac.nz)



THE UNIVERSITY  
OF AUCKLAND  
NEW ZEALAND  
Te Whare Wānanga o Tāmaki Makaurau



A photograph of a vineyard. In the foreground, there are green vines with small leaves. The background shows a bright, clear sky. A white rectangular box is overlaid on the image, containing a quote.

“Communication is not only the essence of being human, but also a vital property of life.”

- John A. Piece

# Modern Technology Trends

## 1. Improved Content Capture

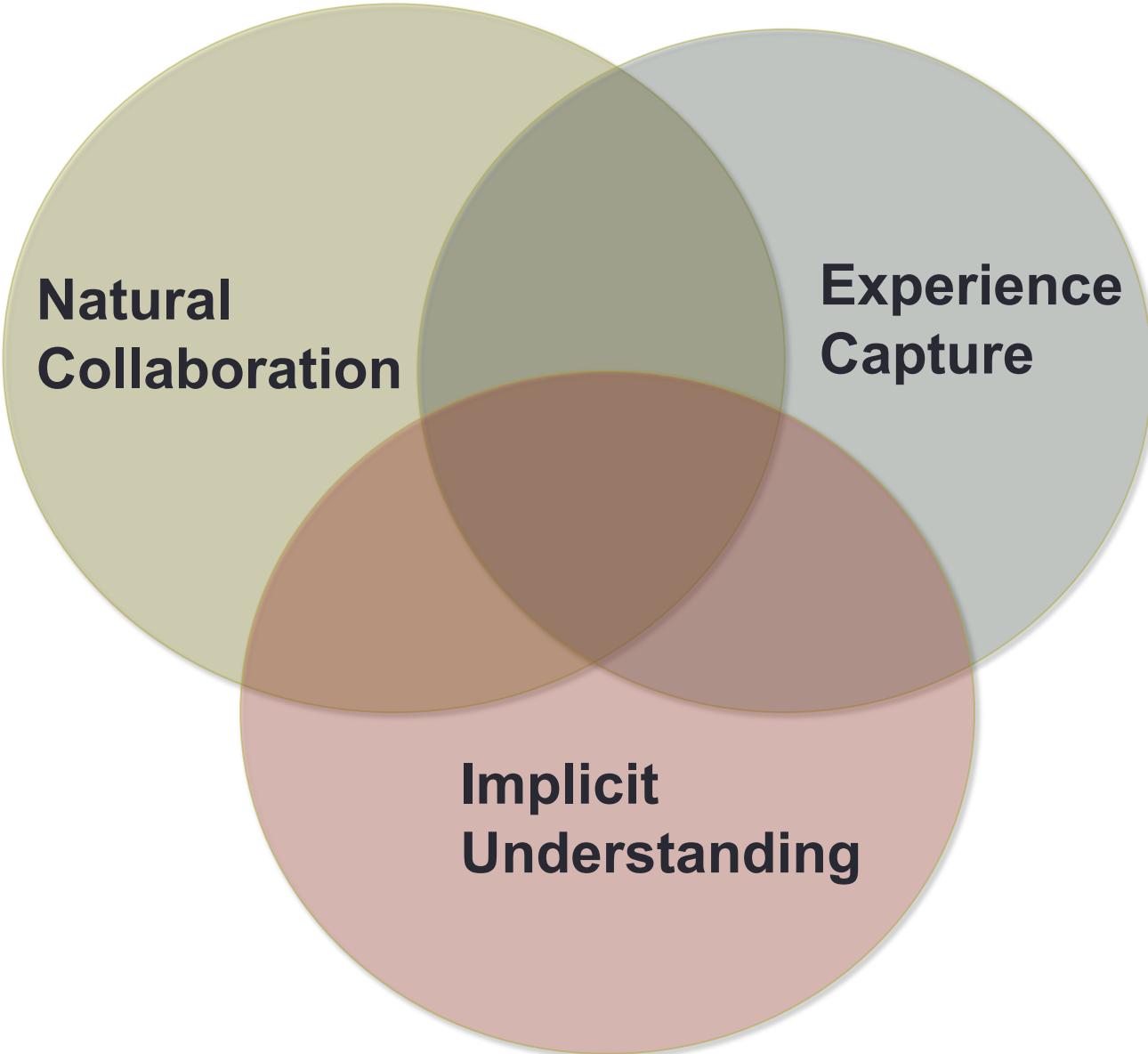
- Move from sharing faces to sharing places

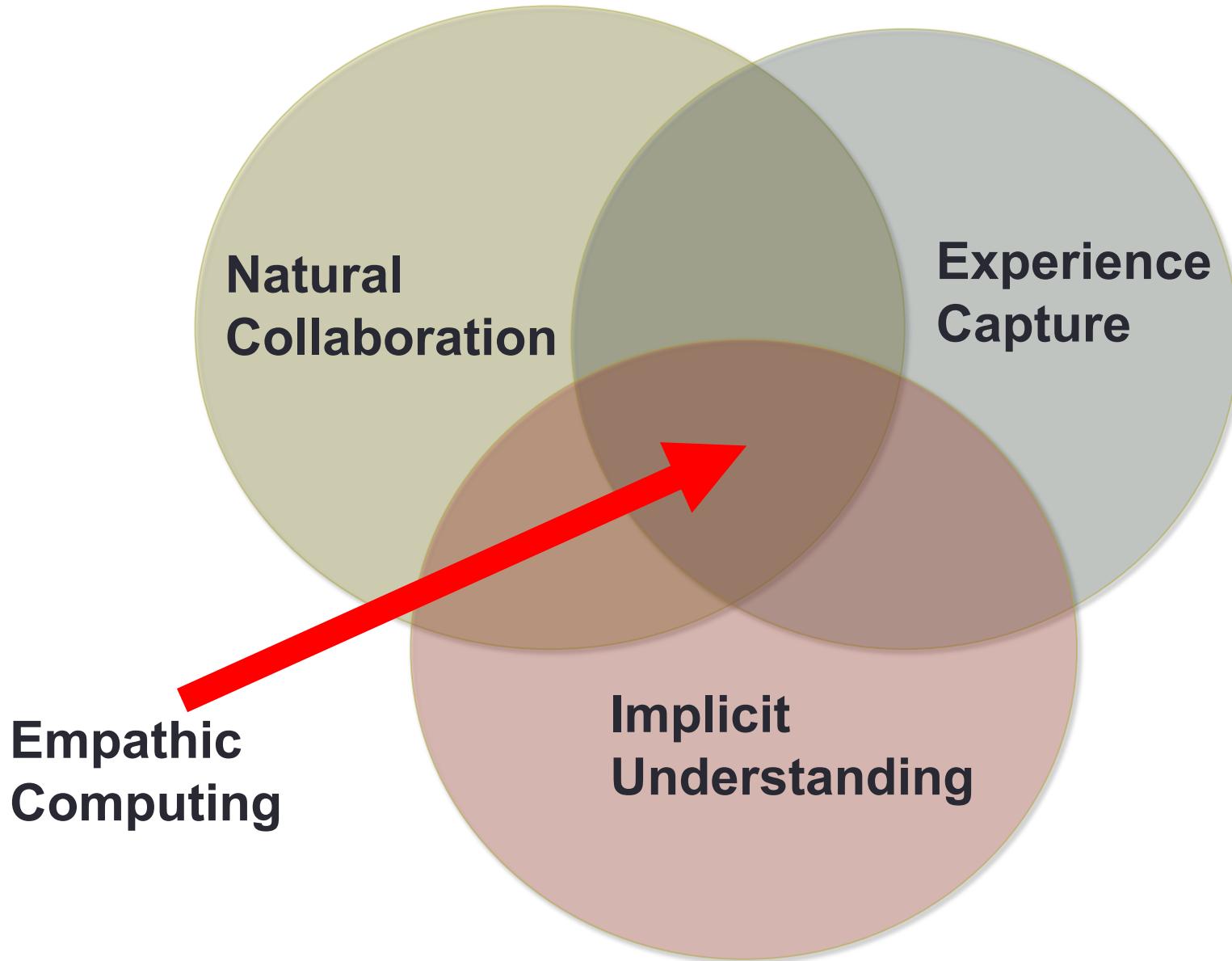
## 2. Increased Network Bandwidth

- Sharing natural communication cues

## 3. Implicit Understanding

- Recognizing behaviour and emotion





# EMPATHIC COMPUTING

---

# Empathy

“**Seeing** with the **Eyes** of another,  
**Listening** with the **Ears** of another,  
and **Feeling** with the **Heart** of another..”

*Alfred Adler*

# Research Focus

*Can we develop systems  
that allow us to share what  
we are seeing, hearing and  
feeling with others?*

# Empathic Computing

1. Understanding: Systems that can understand your feelings and emotions
2. Experiencing: Systems that help you better experience the world
3. Sharing: Systems that help you better share the experience with others

Sensors

VR

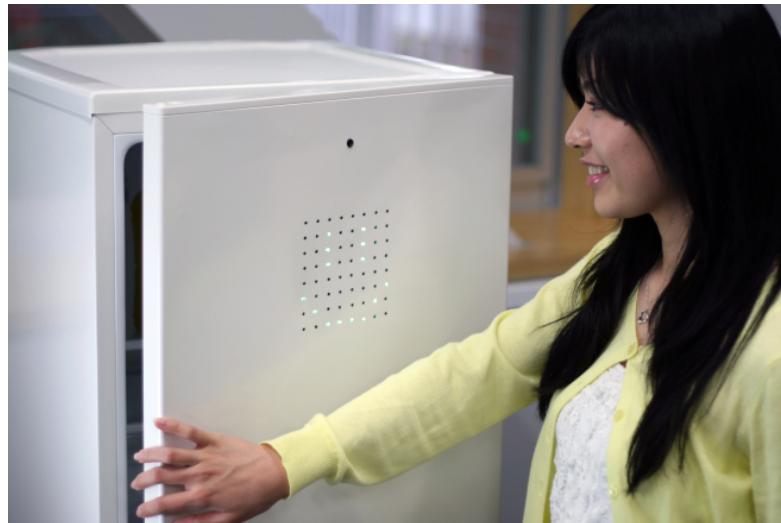
AR

# 1. Understanding: Affective Computing



- Ros Picard – MIT Media Lab
- Systems that recognize emotion

# Appliances That Make You Happy



- Jun Rekimoto – University of Tokyo/Sony CSL
- Smile detection + smart appliances



# **Smiling Makes Us Happier**

**H.Tsujita, J.Rekimoto**

**The University of Tokyo  
Sony CSL**

## 2. Experiencing: Virtual Reality

*"Virtual reality offers a whole different medium to tell stories that really connect people and create an empathic connection."*

Nonny de la Peña

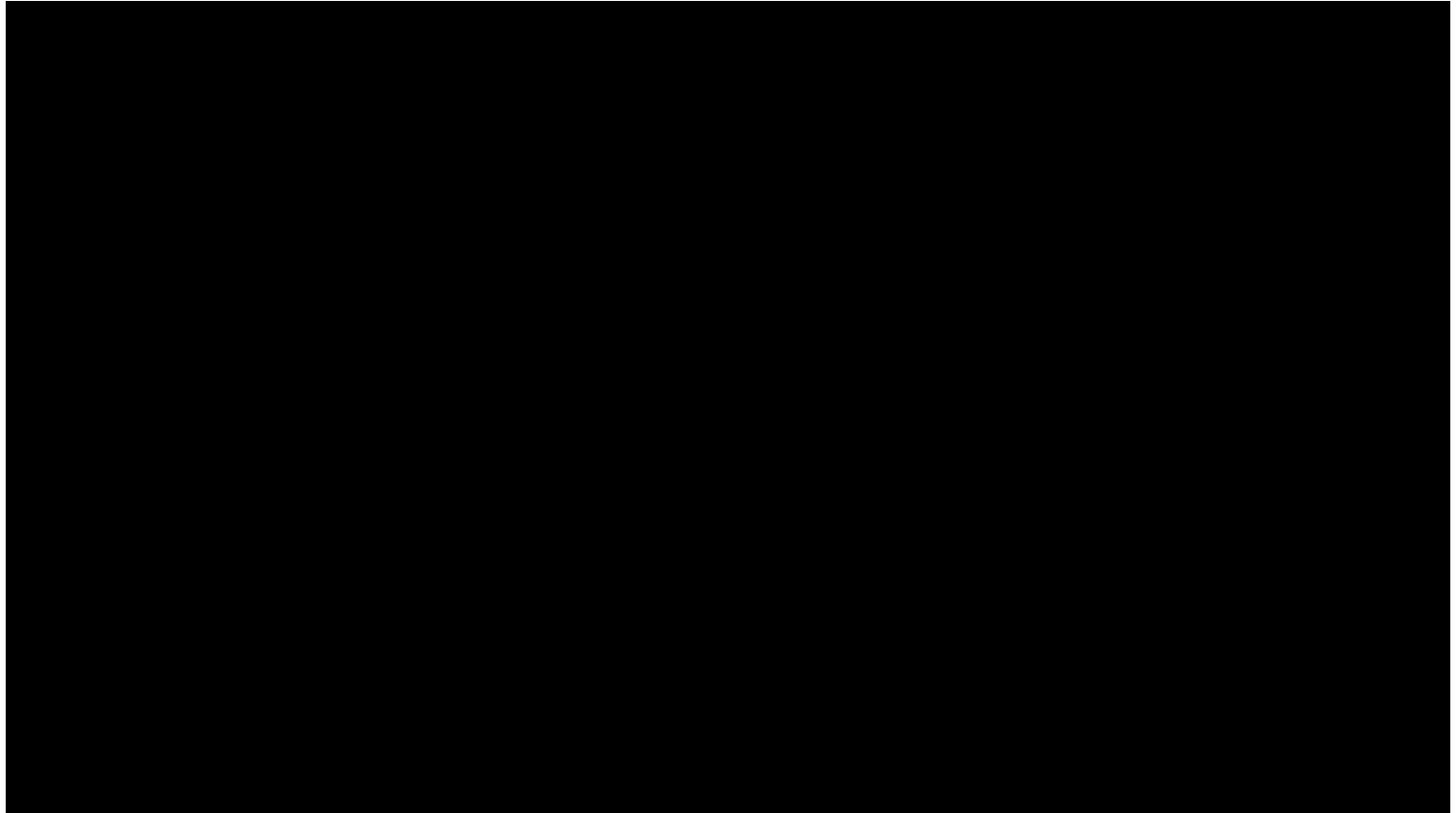
<http://www.emblematicgroup.com/>



# Using VR for Empathy



- USC Project Syria (2014)
- Experience of Terrorism
- Project Homeless (2015)
- Experience of Homelessness



### 3. Sharing: Augmented Reality

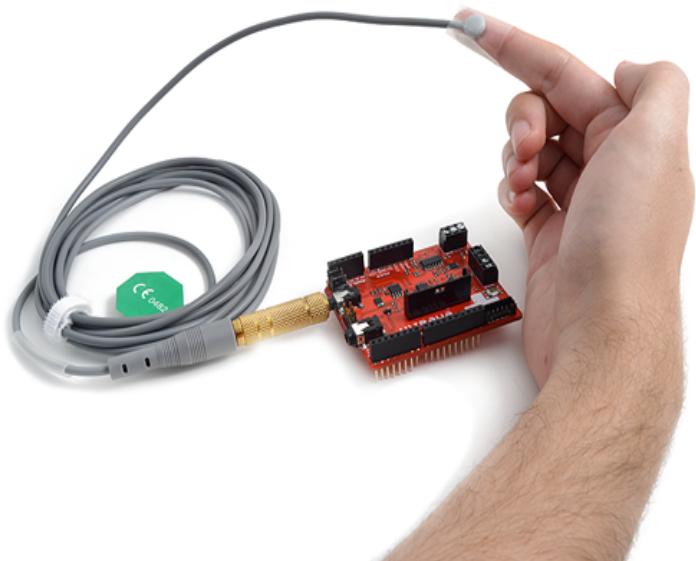
*Can we develop systems  
that allow us to share what  
we are seeing, hearing and  
feeling with others?*

# Example Projects



- Remote collaboration in Wearable AR
  - Sharing of non-verbal cues (gaze, pointing, face expression, emotion)
- Shared Empathic VR experiences
  - Use VR to put a viewer inside the players view
- Measuring emotion
  - Detecting emotion from heart rate, GSR, eye gaze, etc.

# CoSense (CHI 2015)



e-Health 2.0 board

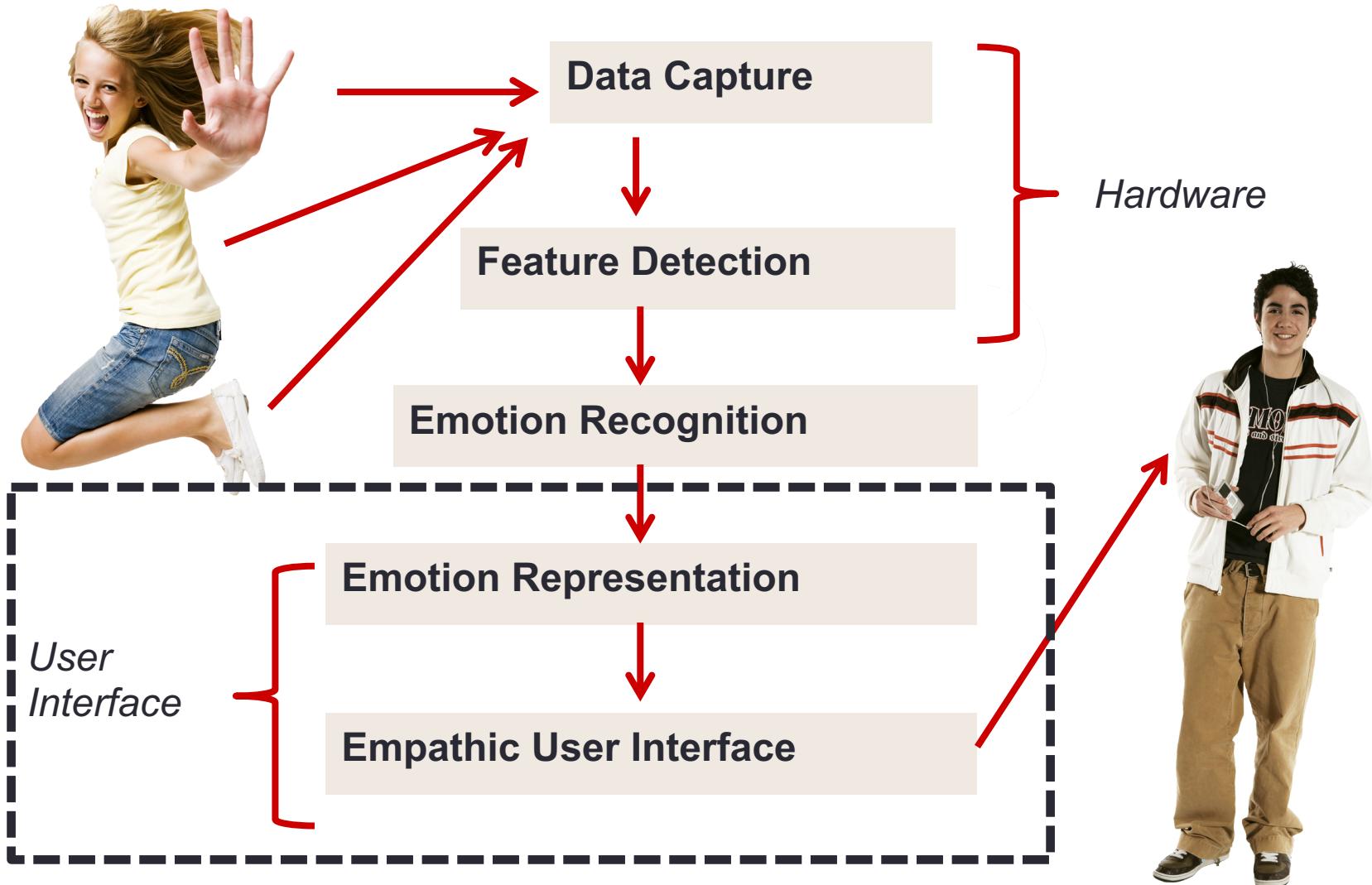


Google Glass

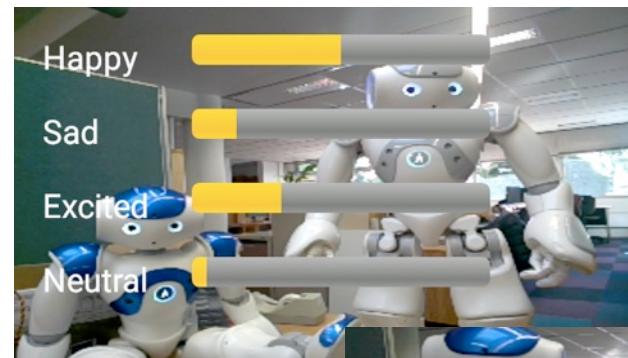
+

- Real time sharing - Emotion, video, and audio
- Wearable (send emotion) → Desktop (remote view)

# Implementation

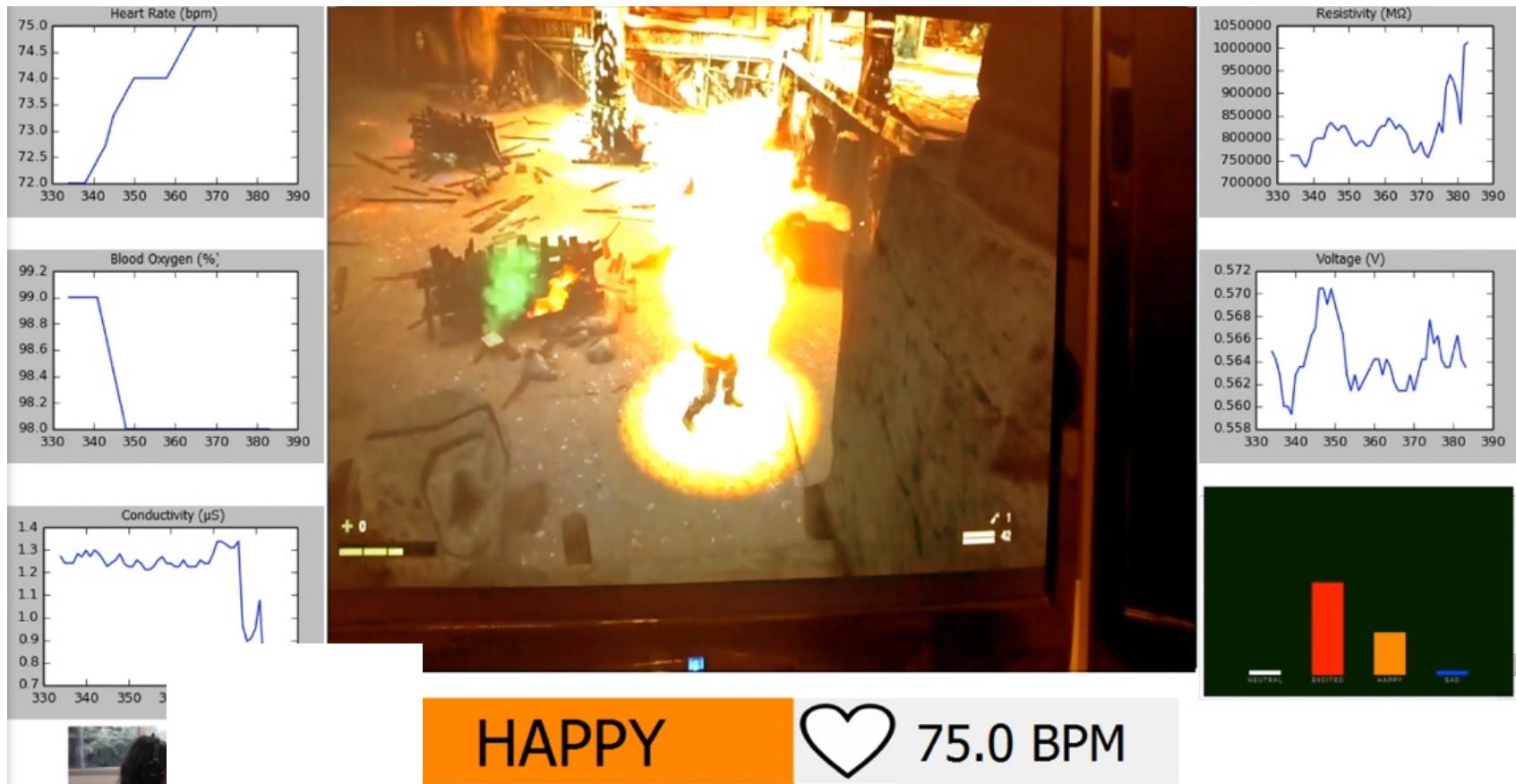


# Wearable Interface



- **Google Glass + e-Health + Spydroid + SSI**
  - Measure GSR, pulse oxygen, ECG, voice pitch
  - Share video and audio remotely
  - Representative emotions sent back to Glass user

# Desktop Interface



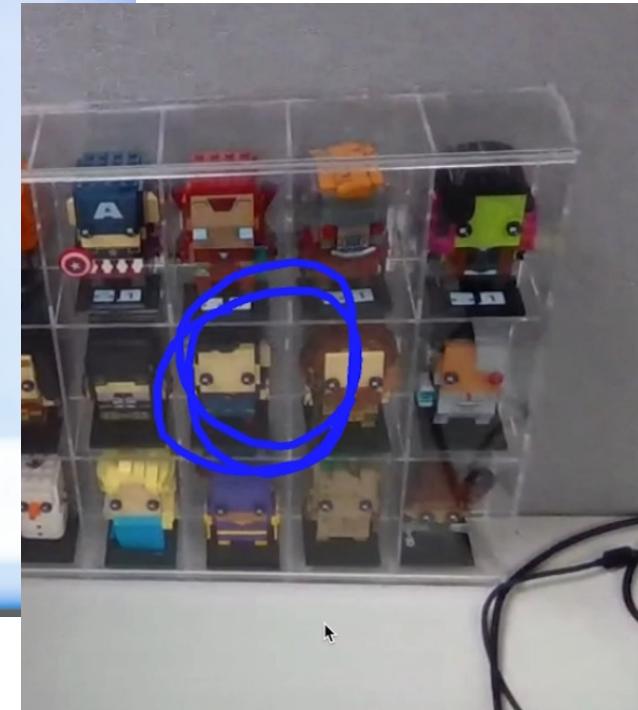
# CoSense Demo



Creating Shared Emotional Experiences  
Between Multiple Users

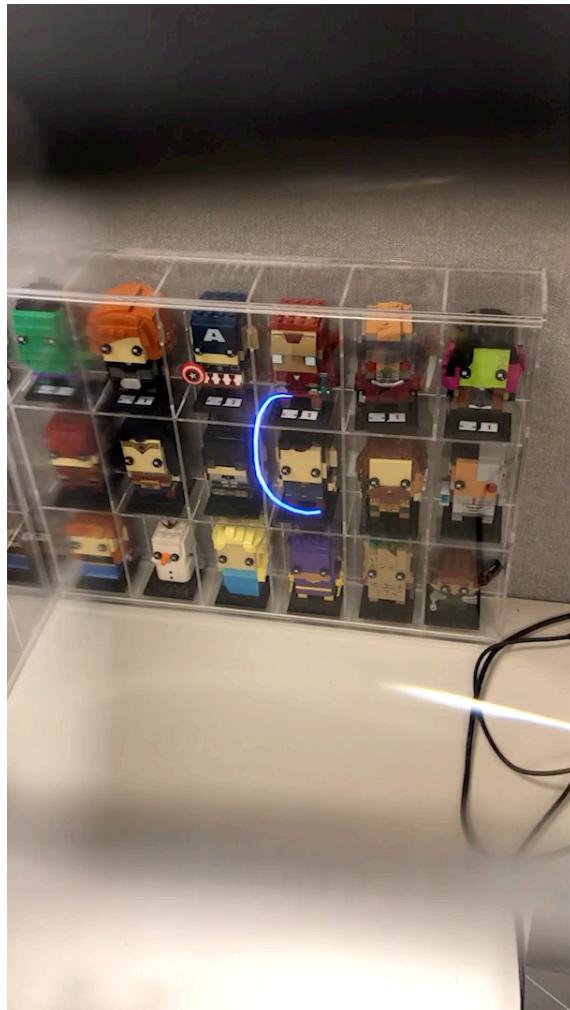


# AR for Remote Collaboration

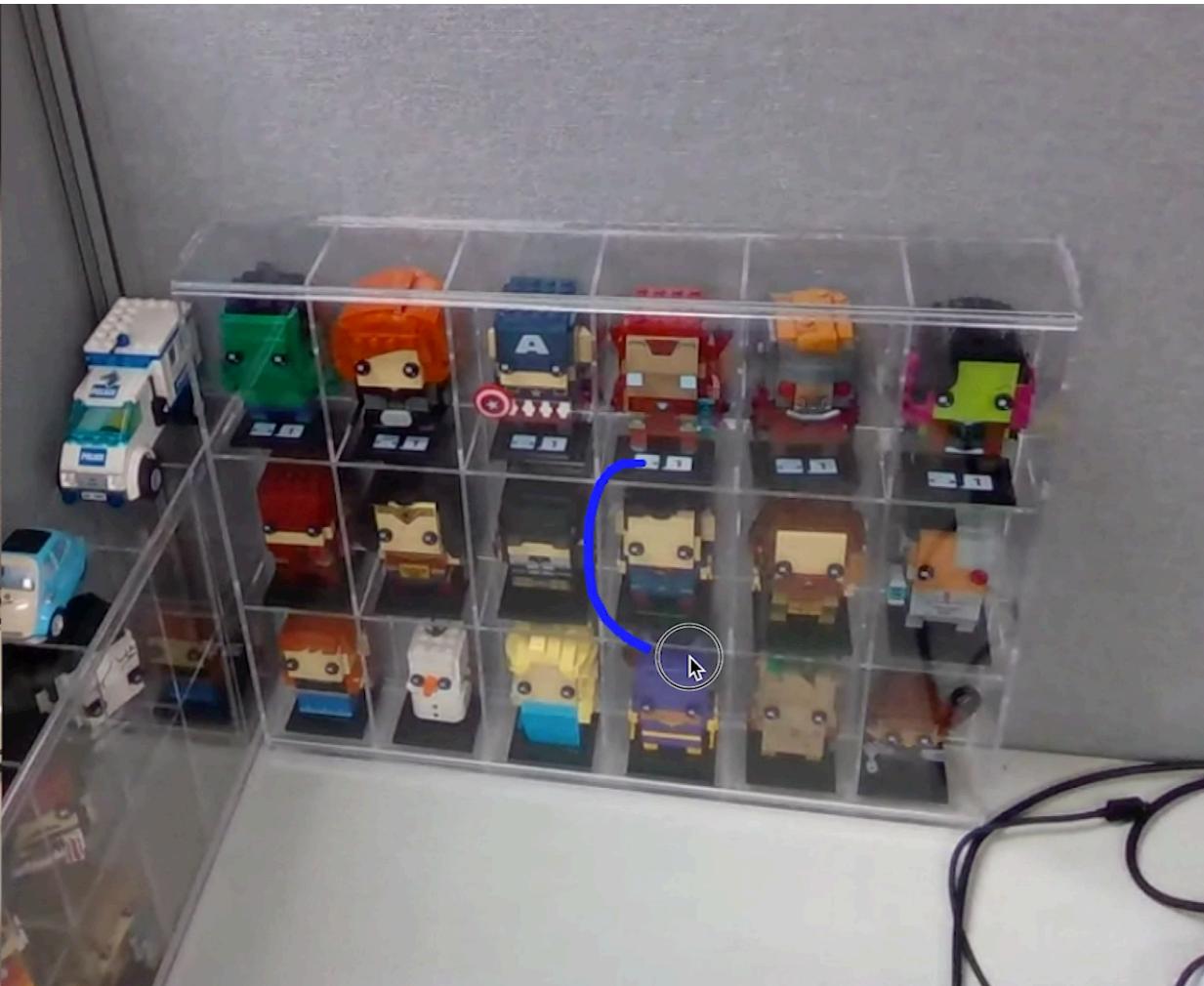


- Camera + Processing + AR Display + Connectivity
- First person Ego-Vision Collaboration

# AR View



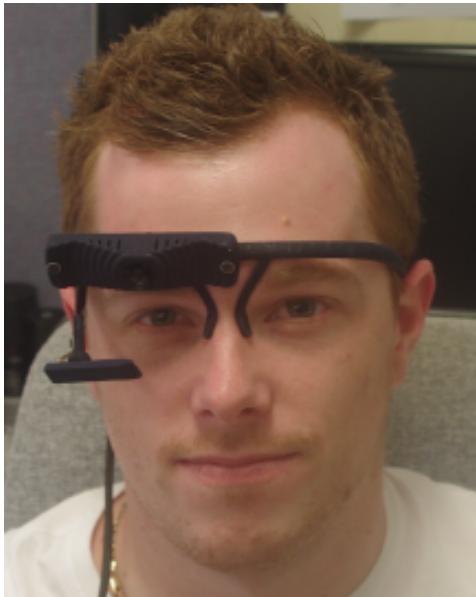
# Remote Expert View



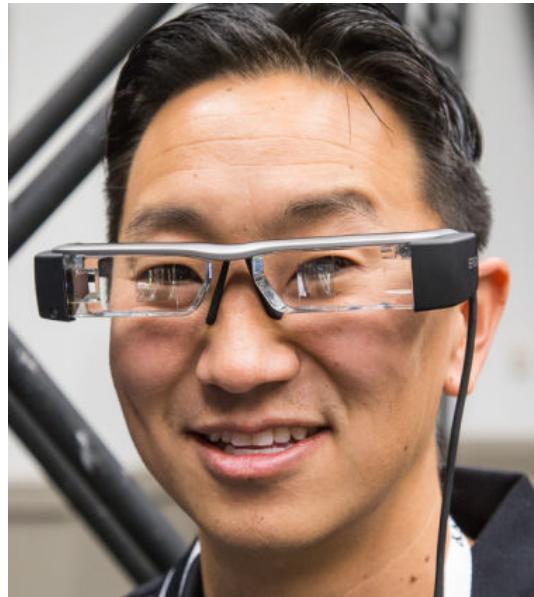
# An Expensive Mistake



# Empathy Glasses



Pupil Labs



Epson BT-200

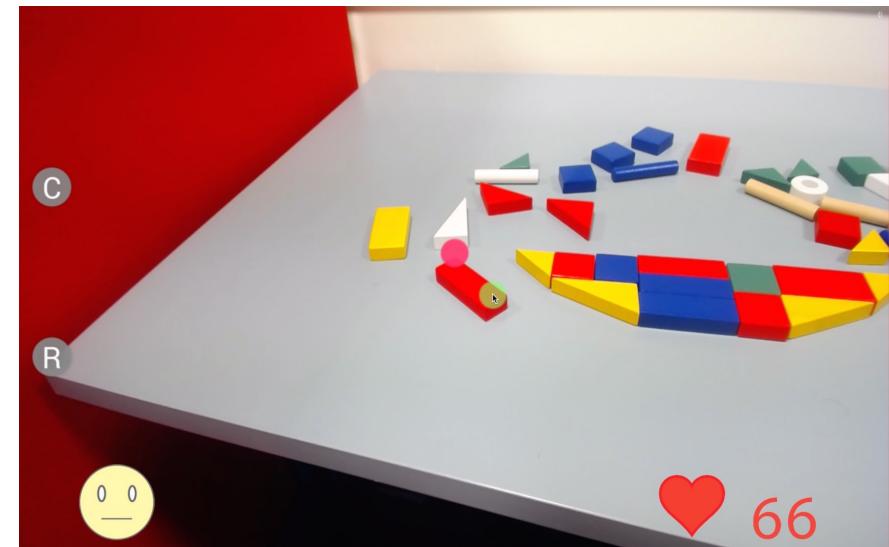
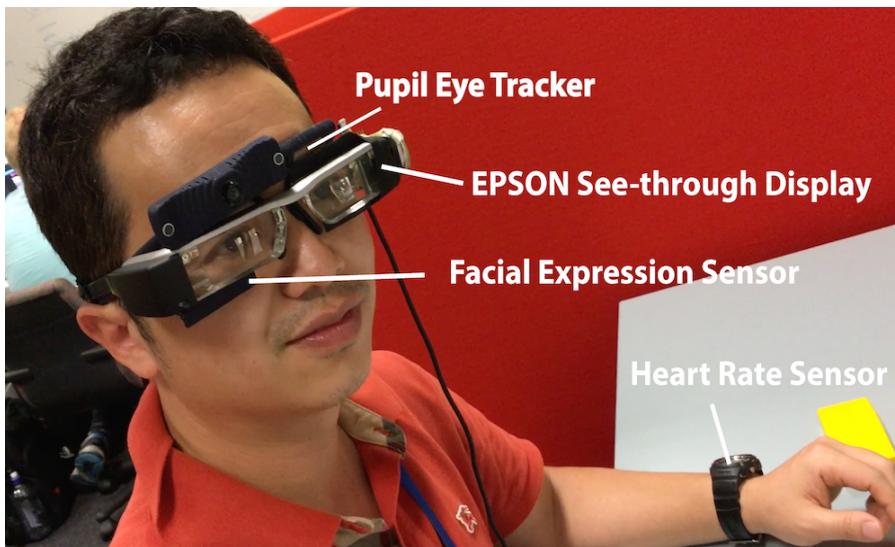


AffectiveWear



- Combine together eye-tracking, display, face tracking
- Implicit cues – eye gaze, face expression

# Remote Collaboration



- Eye gaze pointer and remote pointing
- Face expression display
- Implicit cues for remote collaboration



# Shared Sphere – 360 Video Sharing



Host User



Guest User



# Shared Sphere

– 360 Panorama-based Mixed Reality Collaboration

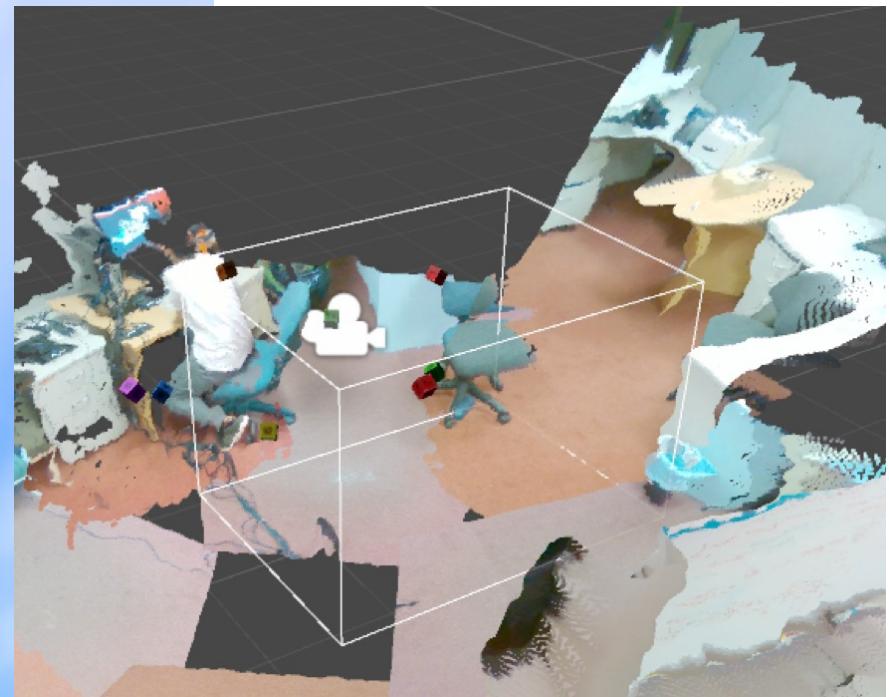
FRONTIER 4.0



University of  
South Australia



# 3D Live Scene Capture



- Use cluster of RGBD sensors
- Fuse together 3D point cloud

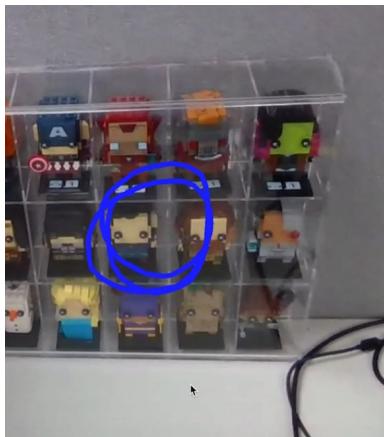
# AR View



# Remote Expert View



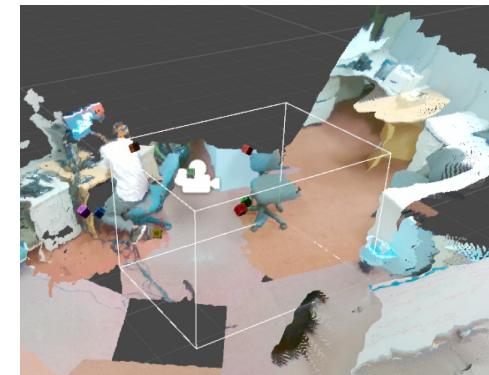
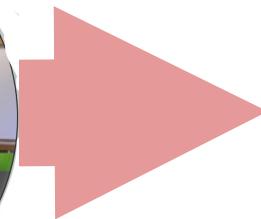
# View Sharing Evolution



2D



360



3D

- Increased immersion
- Improved scene understanding
- Better collaboration

# Sharing VR Experiences



# VR Environments



- **Butterfly World:** calm scene, collect butterflies
- **Zombie Attack:** scary scene, fighting zombies

# Demo Video

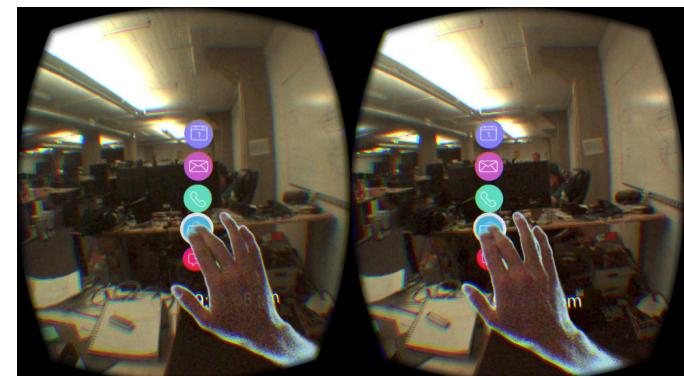
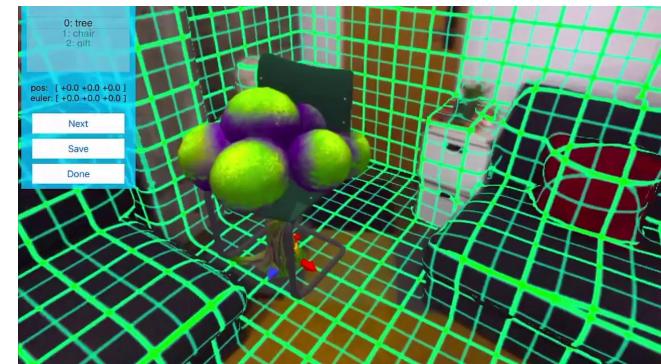
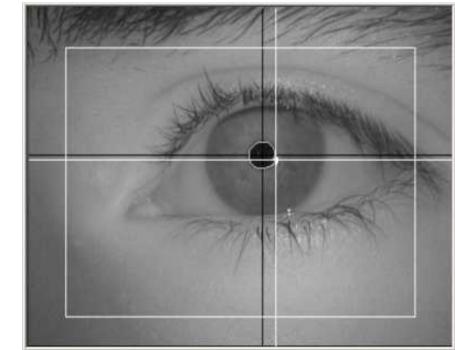


# AR and VR for Empathic Computing

- VR systems are ideal for trying experiences:
  - Strong story telling medium
  - Provide total immersion/3D experience
  - Easy to change virtual body scale and representation
- AR systems are idea for live sharing:
  - Allow overlay on real world view/can share viewpoints
  - Support remote annotation/communication
  - Enhance real world task

# Technology Trends

- Robust eye-tracking
  - Gaze points, focus depth
- Advanced displays
  - Wide FOV, high resolution
- Real time space capture
  - 3D scanning, stitching
- Natural gesture interaction
  - Hand tracking, pose recognition



# Emotion Sensing HMDs

- Looxid VR
  - EEG/Emotion sensing
  - Eye tracking
- Emteq
  - EMG sensing
- Neurable
  - EEG/BCI for VR
- PhysioHMD (MIT Media Lab)
  - GSR, PPG, Emotion Sensing



# Looxid VR HMD

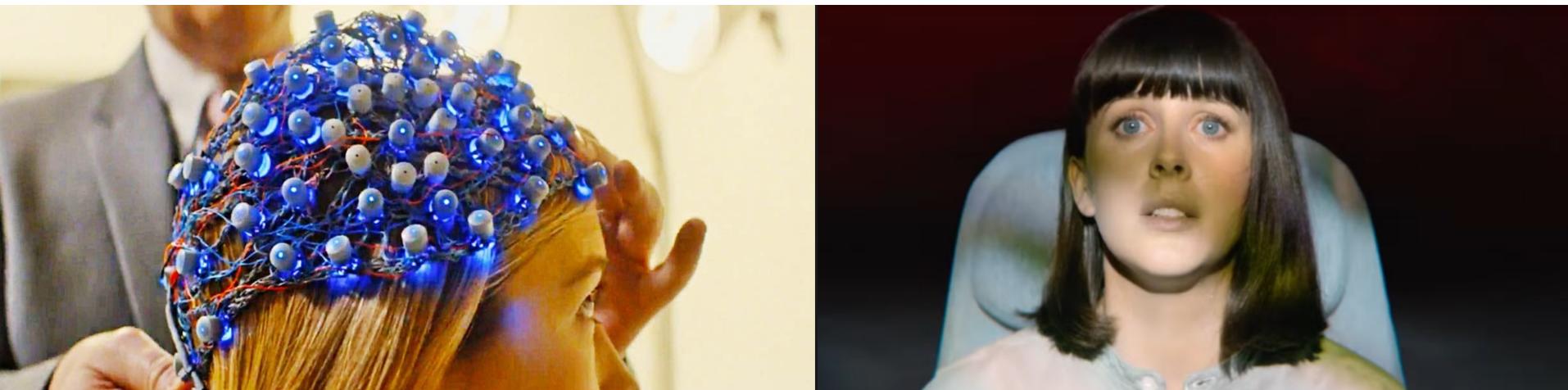


<https://www.youtube.com/watch?v=0jCGLUbJ800>

# Technology Trends

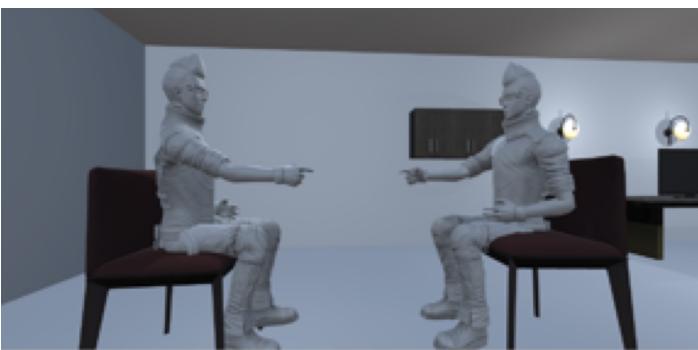
- Advanced displays
  - Real time space capture
  - Natural gesture interaction
  - Robust eye-tracking
  - Emotion sensing/sharing
- 
- Empathic  
Tele-Existence

# Empathic Tele-Existence

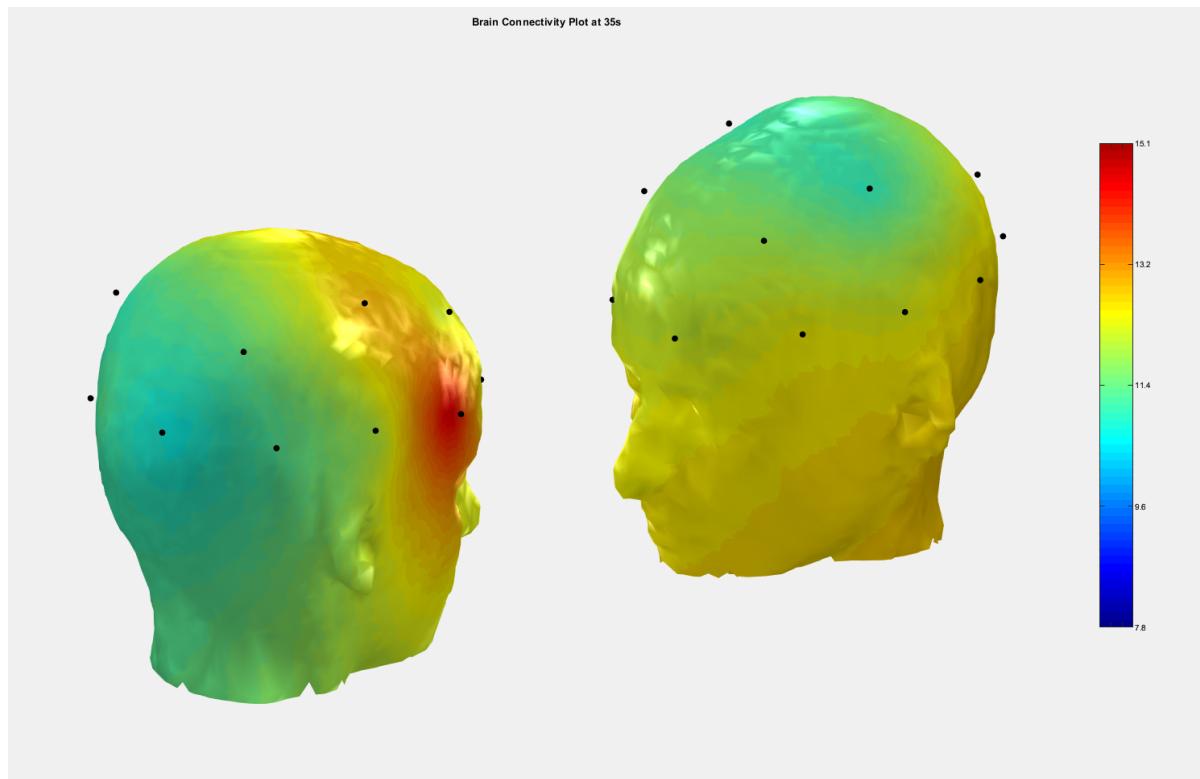


- Move from Observer to Participant
- Explicit to Implicit communication
- Experiential collaboration – doing together

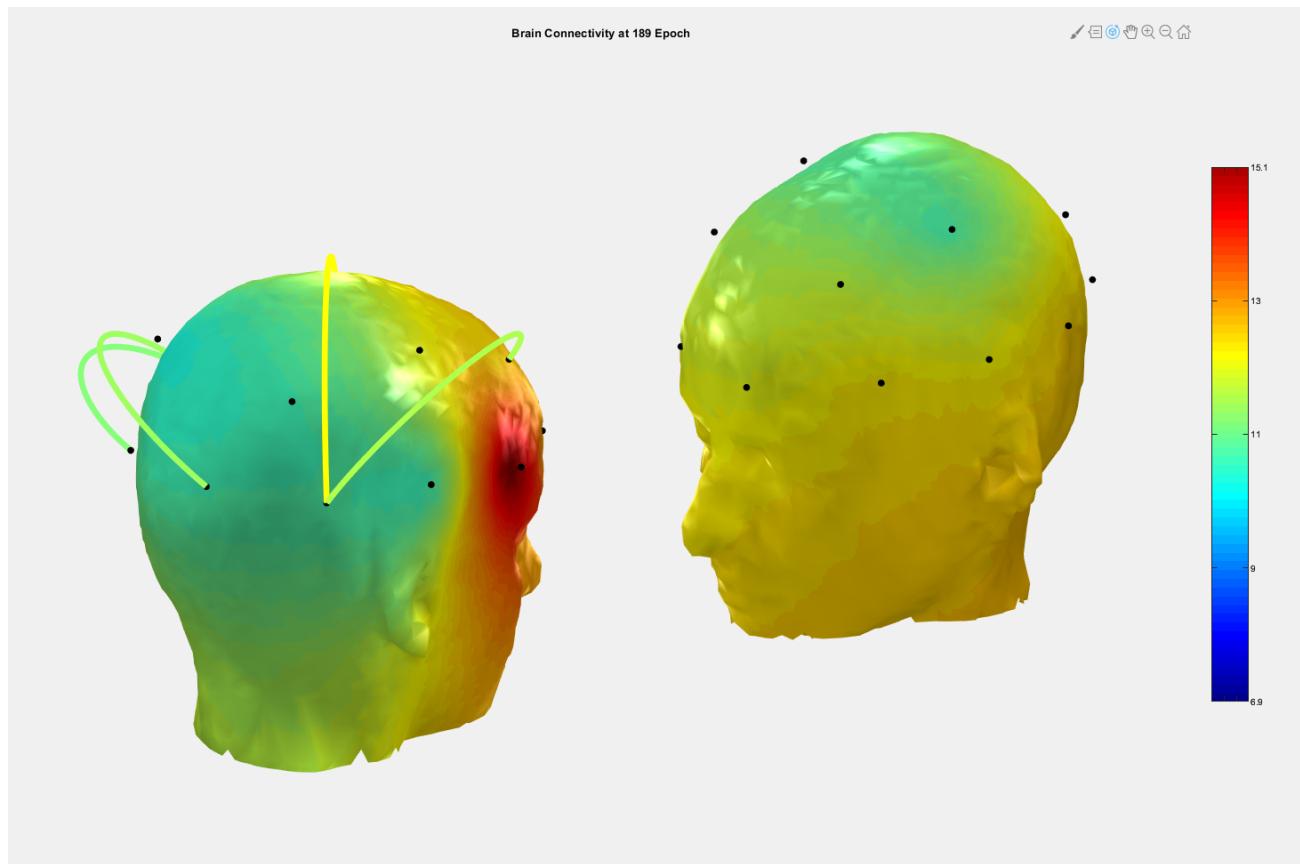
# Brain Synchronization



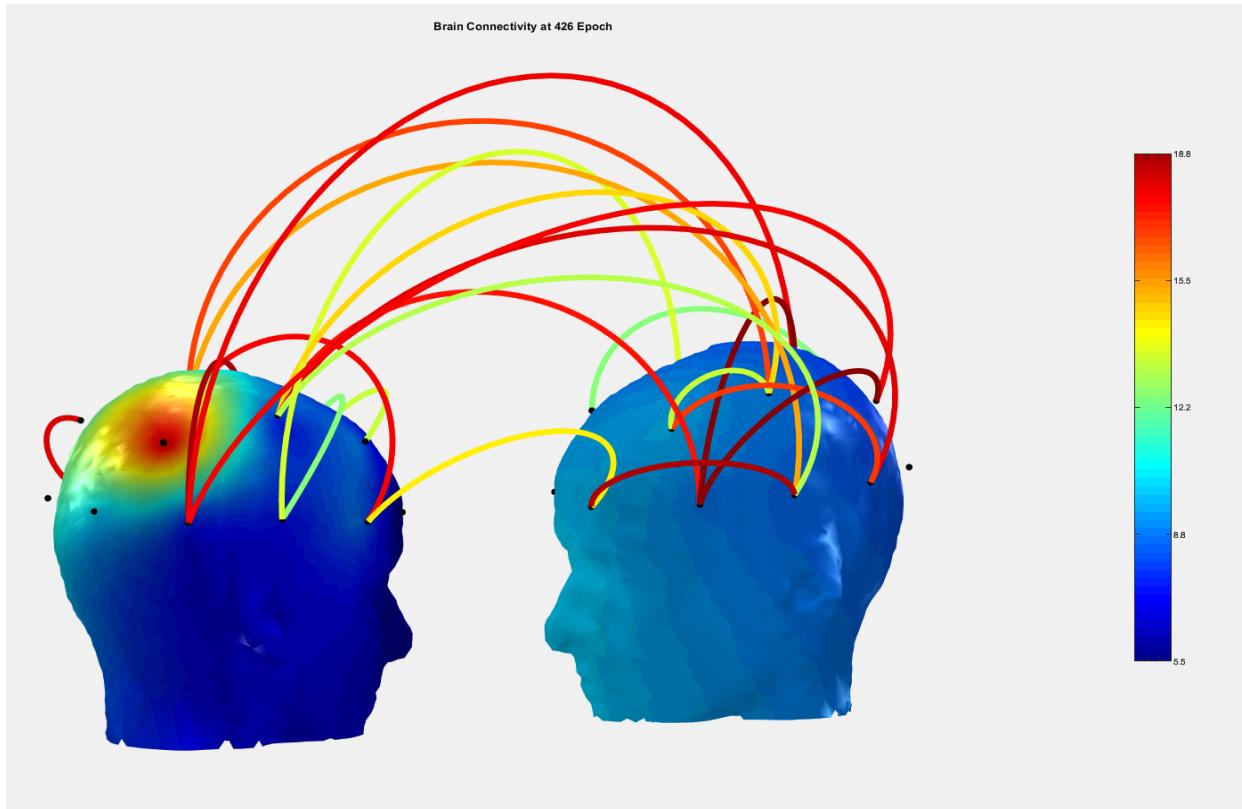
# Pre-training (Finger Pointing) Session Start



# Pre-training (Finger Pointing) Session End



# Post-Training (Finger Pointing) Session End



# Conclusions

- Trend towards Empathic Computing
  - Understanding, Experiencing, Sharing
- AR Enables Empathic Experiences
  - Changes perspective
  - Sharing space/experience
  - Supports annotation/communication
- Many directions for future research

# Come Play with Us..

- Industry collaboration
  - Government funded research
  - Research commercialization
- Academic collaboration
  - Student exchanges
  - Visiting researchers
  - Joint lab projects
- Funded positions
  - PhD scholarships
  - Postdoctoral research
  - Internships





[www.empathiccomputing.org](http://www.empathiccomputing.org)



[mark.billinghurst@auckland.ac.nz](mailto:mark.billinghurst@auckland.ac.nz)



[@marknb00](https://twitter.com/marknb00)