

IVE WINTER SCHOOL

FUNDAMENTALS OF XR

Mark Billinghurst

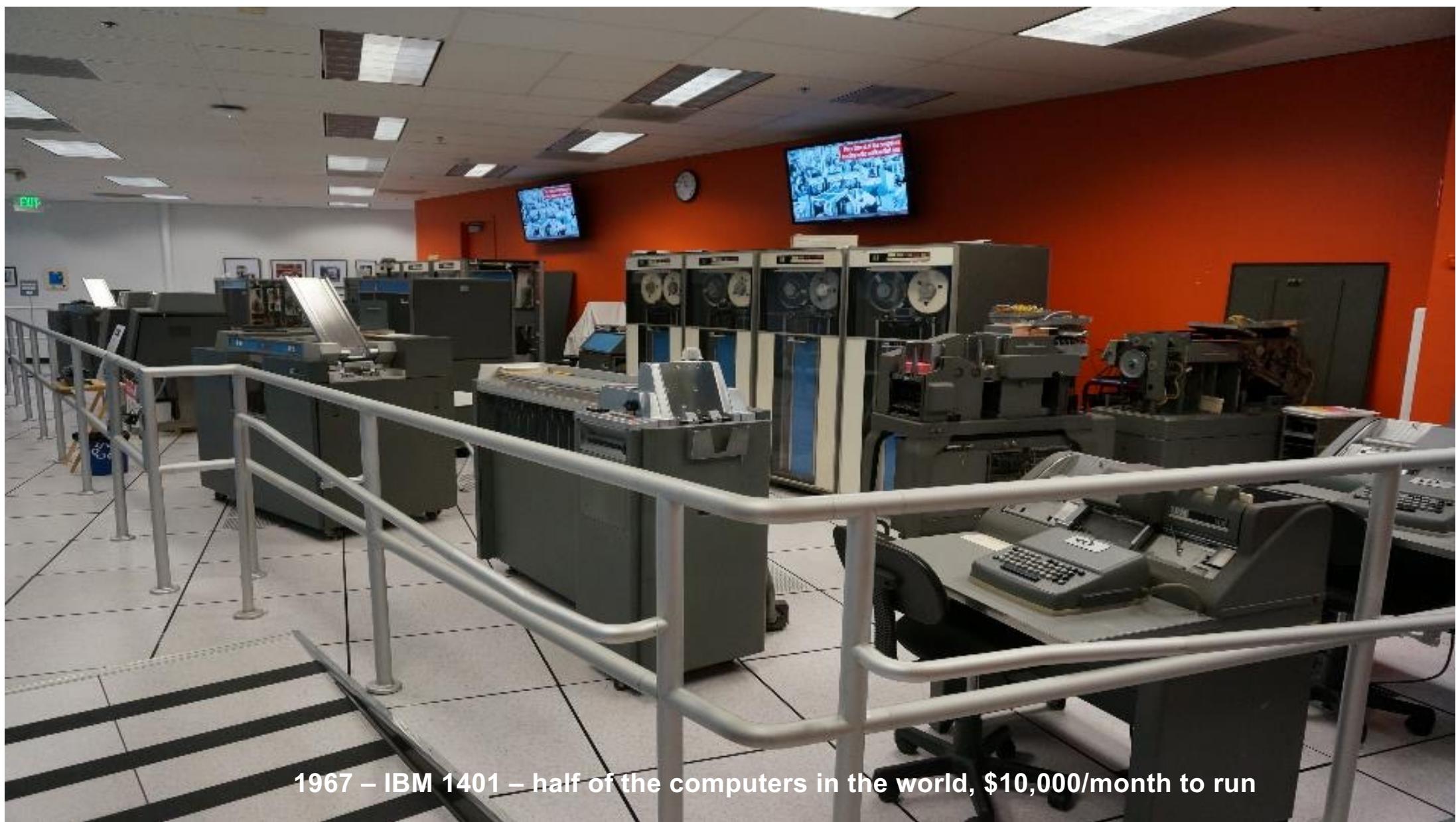
mark.billinghurst@unisa.edu.au

July 14th 2025



University of
South Australia

WHAT IS AR/VR/MR/XR?



1967 – IBM 1401 – half of the computers in the world, \$10,000/month to run

The Incredible Disappearing Computer



1960-70's

Room



1970-80's

Desk



1980-90's

Lap

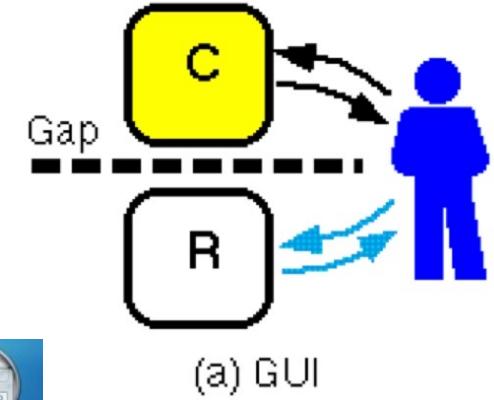


1990-2000's

Hand

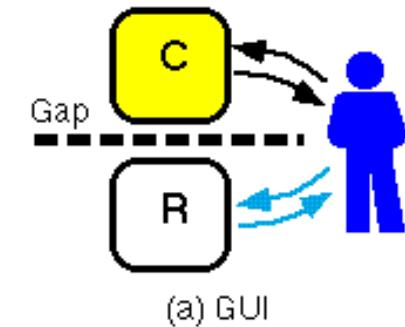


Graphical User Interfaces

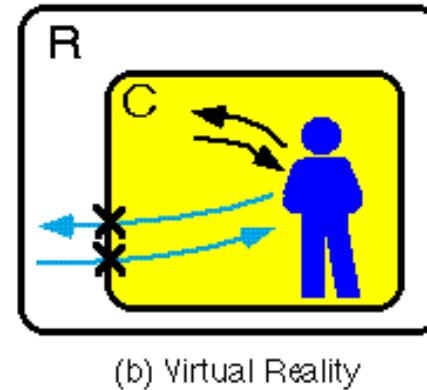


- Separation between real and digital worlds
 - WIMP (Windows, Icons, Menus, Pointer) metaphor

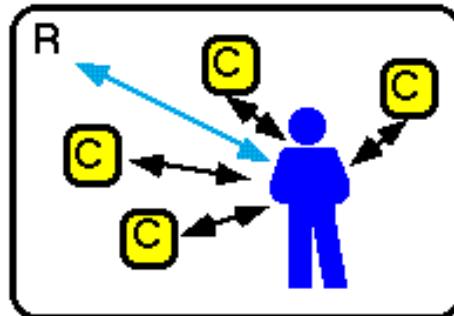
Making Interfaces Invisible



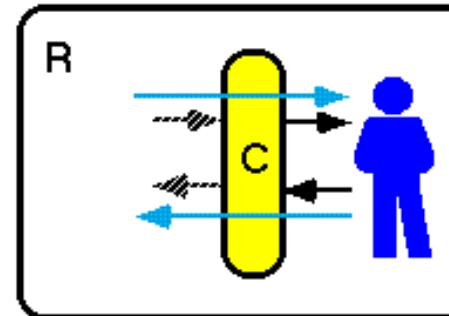
(a) GUI



(b) Virtual Reality



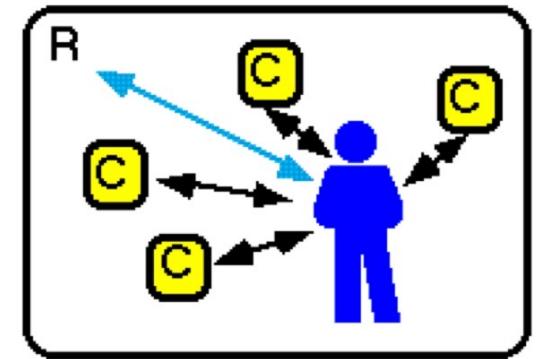
(c) Internet of Things



(d) Augmented Interaction

Rekimoto, J. and Nagao, K. 1995. The world through the computer: computer augmented interaction with real world environments.

Internet of Things (IoT)..

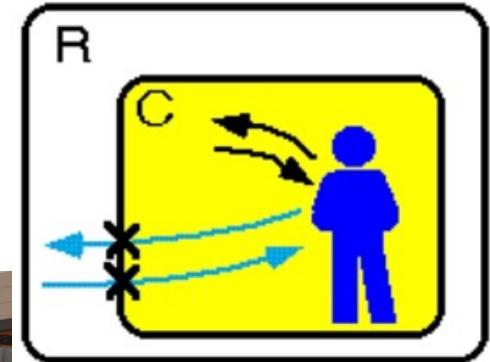


(c) Internet of Things



- Embed computing and sensing in real world
 - Smart objects, sensors, etc..

Virtual Reality (VR)



(b) Virtual Reality

- Users immersed in Computer Generated environment
 - HMD, gloves, 3D graphics, body tracking
- Key features: Immersion, Interaction, Imagination

HAMMERHEAD

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

Key Characteristics of VR

- Virtual Reality has three key characteristics
 - 3D stereoscopic display
 - Wide field of view display
 - Low latency head tracking
- When these three things are combined they provide a compelling immersive experience

VR and Presence ..

“The subjective experience of being in one place or environment even when physically situated in another”



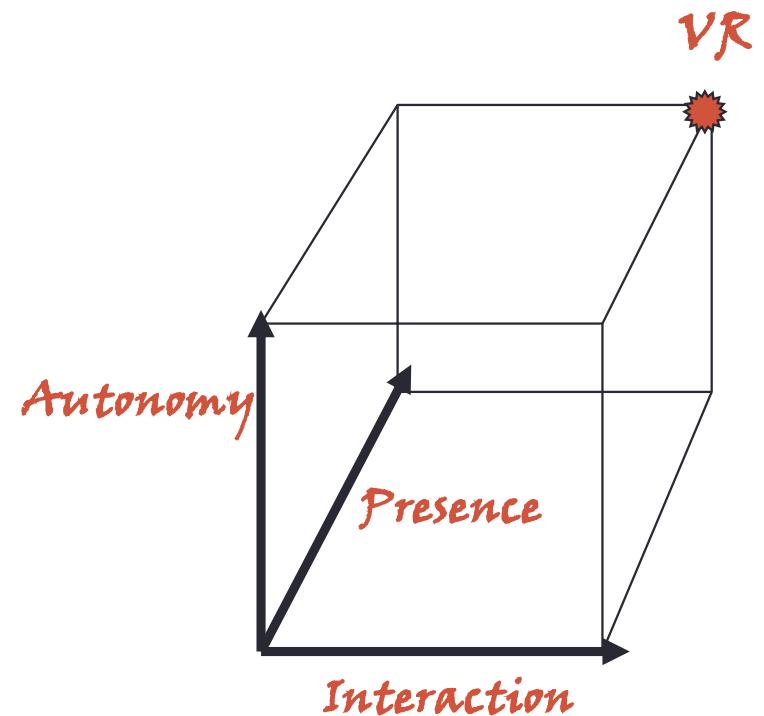
Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, 7(3), 225-240.

David Zeltzer's AIP Cube

Autonomy – User can react to events and stimuli.

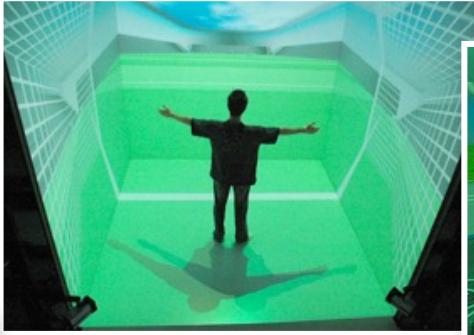
Interaction – User can interact with objects and environment.

Presence – User feels immersed through sensory input and output channels

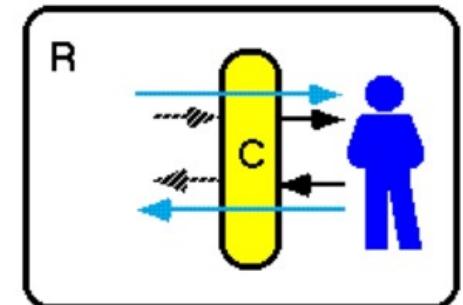


Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence: Teleoperators & Virtual Environments*, 1(1), 127-132.

Types of VR



Augmented Reality (AR)



(d) Augmented Interaction

- Virtual Images blended with the real world
 - See-through HMD, handheld display, viewpoint tracking, etc..



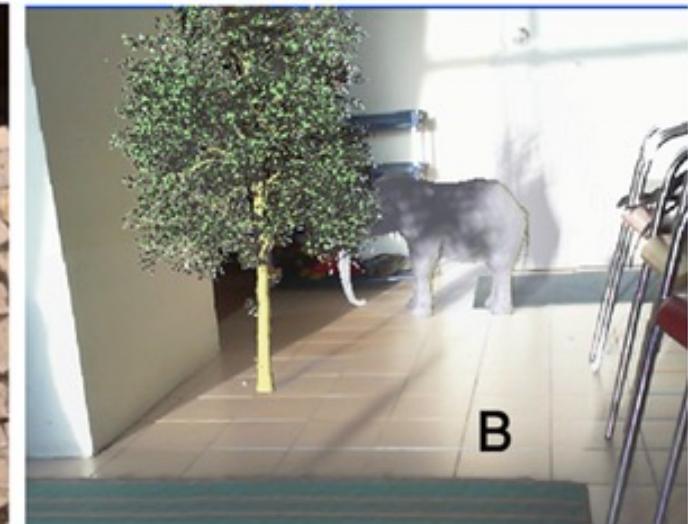
Augmented Reality Definition

- Defining Characteristics [Azuma 97]
 - Combines Real and Virtual Images
 - Both can be seen at the same time
 - Interactive in real-time
 - The virtual content can be interacted with
 - Registered in 3D
 - Virtual objects appear fixed in space

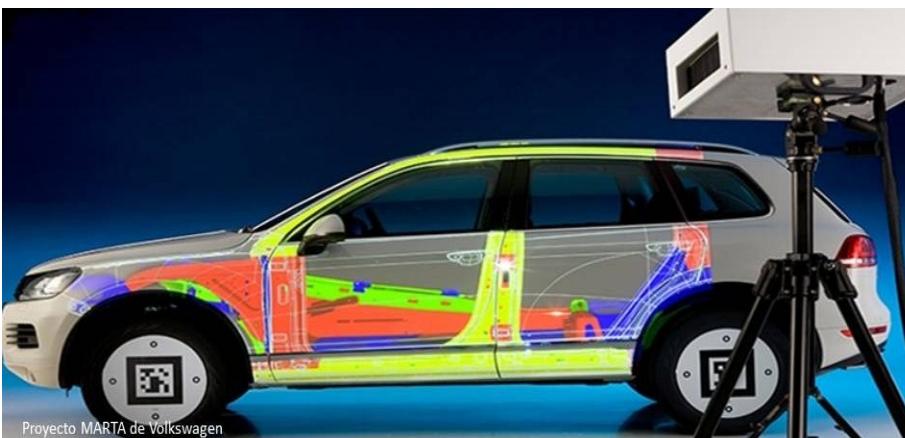
Azuma, R. T. (1997). A survey of augmented reality. *Presence*, 6(4), 355-385.

AR and Object Presence

- Object Presence
 - Feeling that an object is really in your space
- What makes an object appear real?
 - Touch/Haptic feedback
 - Appearance
 - Lighting
 - Audio cues
 - Occlusion
 - Etc..

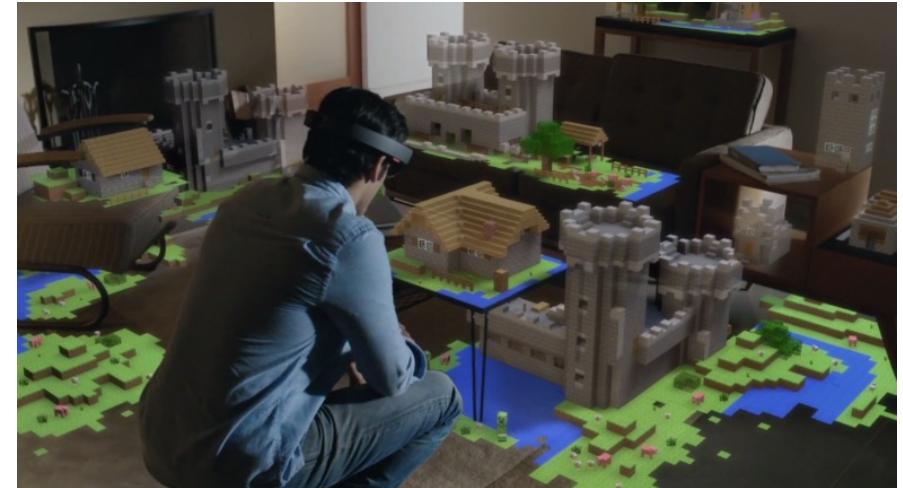


Types of AR



Strong vs. Weak AR

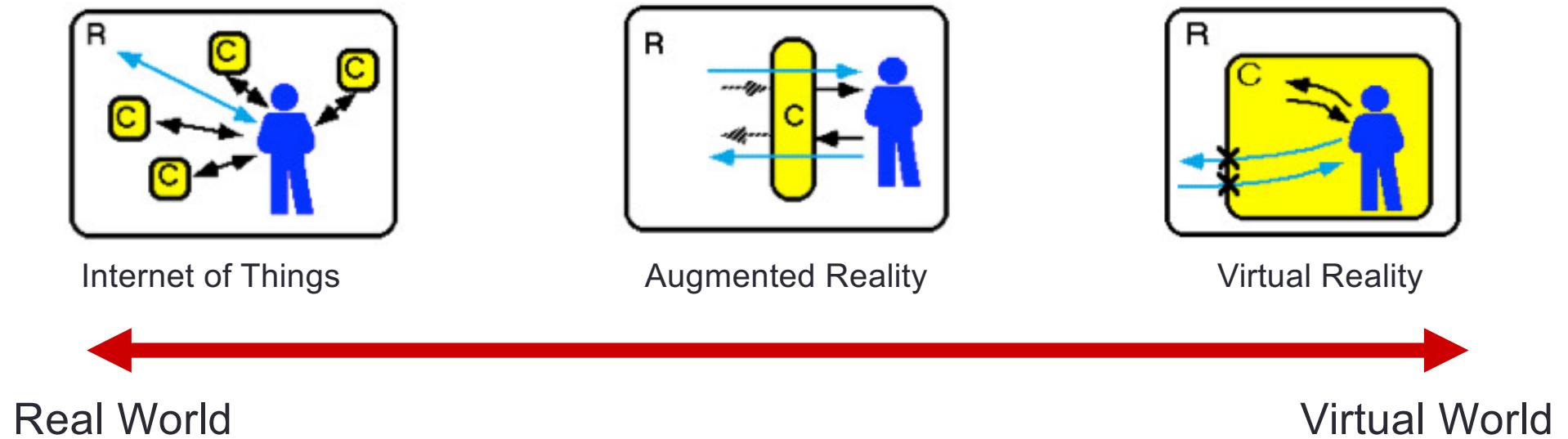
- **Weak AR**
 - Imprecise tracking
 - No knowledge of environment
 - Limited interactivity
 - Handheld AR
- **Strong AR**
 - Very accurate tracking
 - Seamless integration into real world
 - Natural interaction
 - Head mounted AR



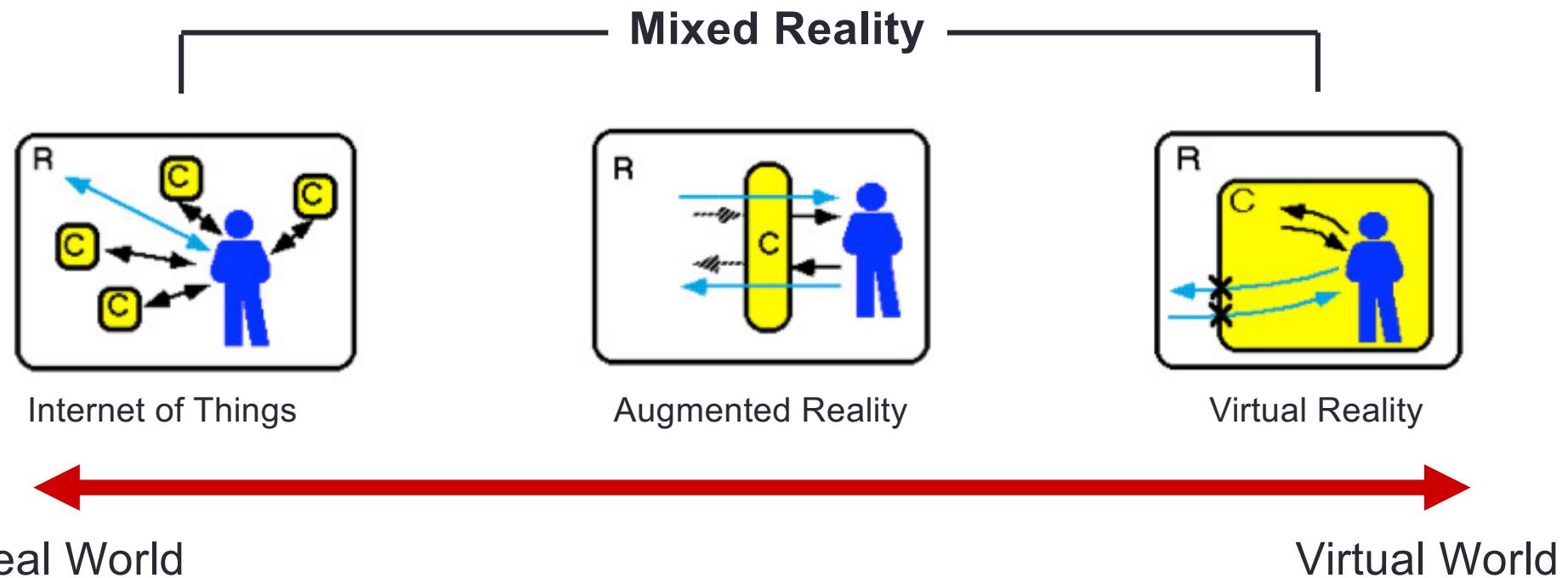
AR vs VR

	Virtual Reality <i>Replaces Reality</i>	Augmented Reality <i>Enhances Reality</i>
<i>Scene Generation</i>	Requires realistic images	Minimal rendering okay
<i>Display Device</i>	Fully immersive, wide field of view	Non-immersive, small field of view
<i>Tracking</i>	Low to medium accuracy is okay	The highest accuracy possible

From Reality to Virtual Reality



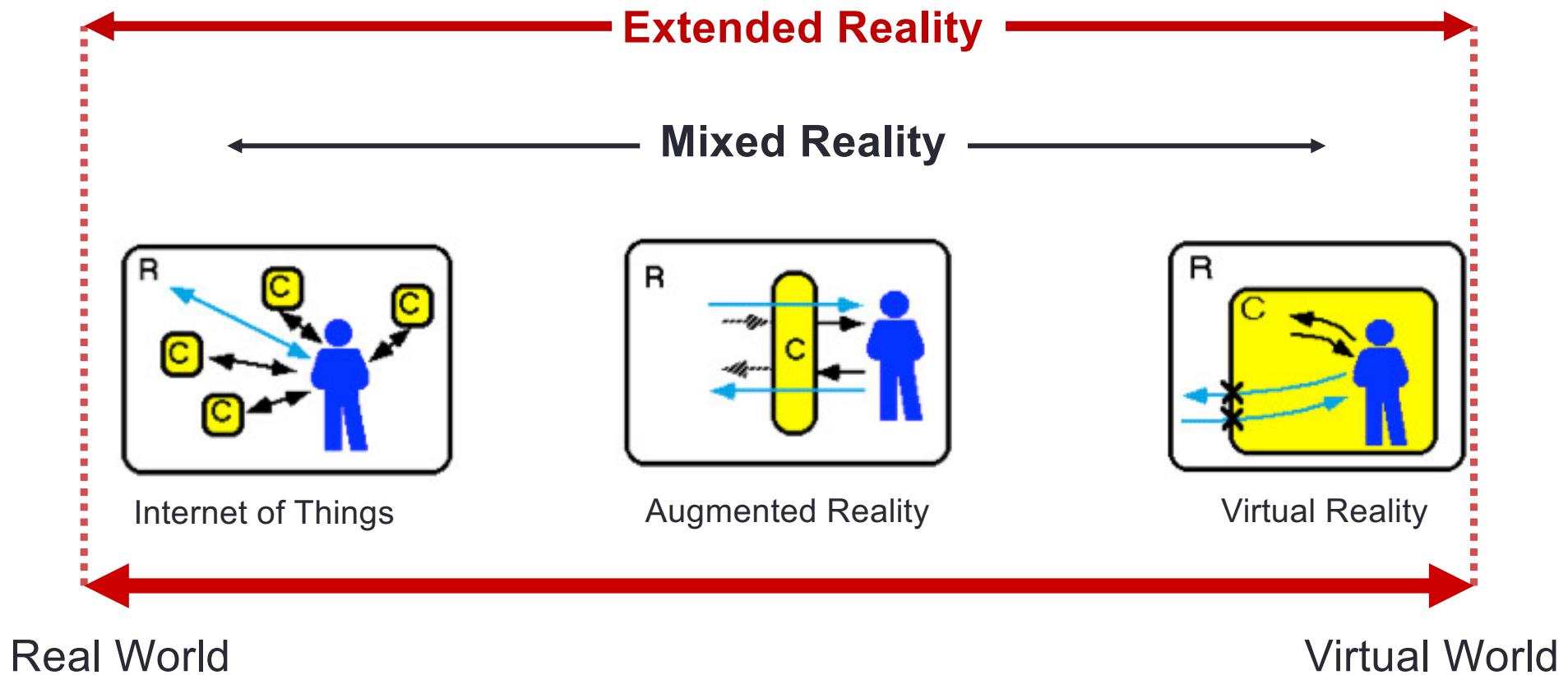
Milgram's Mixed Reality (MR) Continuum



"...anywhere between the extrema of the *virtuality continuum*."

P. Milgram and A. F. Kishino, (1994) A Taxonomy of Mixed Reality Visual Displays

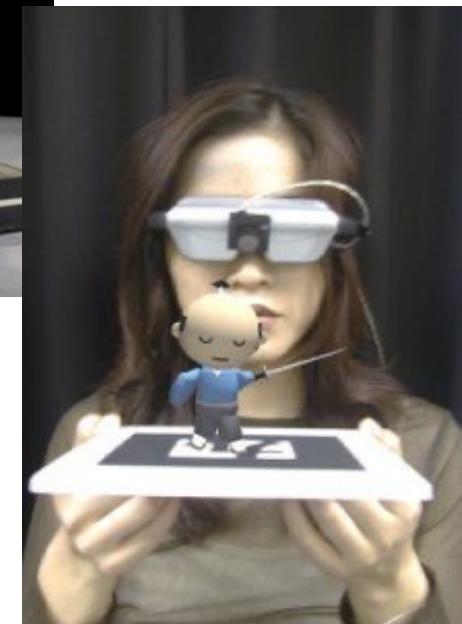
Extended Reality (XR)



TECHNOLOGY

Key Technologies

- **Display**
 - Visual - Stimulate visual sense
 - Audio/Haptic - Stimulate hearing/touch
- **Input Devices**
 - Supporting user interaction
- **Tracking**
 - Measuring changing viewpoint
 - Registering virtual content in real world



Augmented Reality Technology Requirements

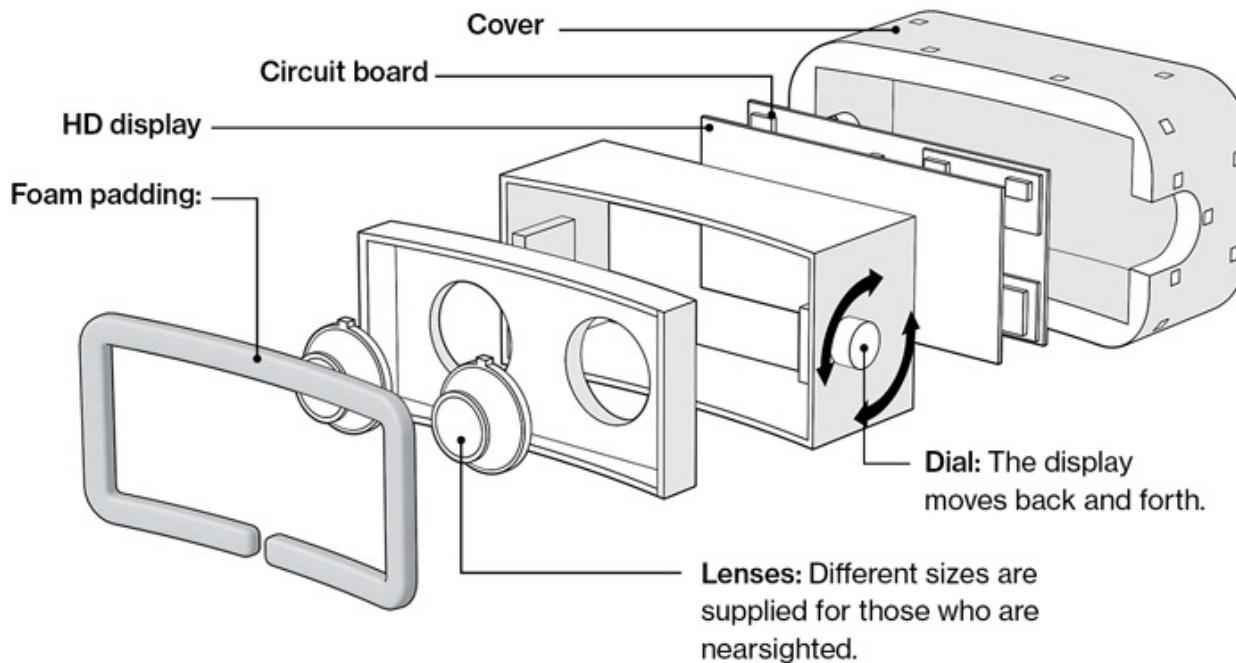
- Combines Real and Virtual Images
 - Needs: Display technology
- Interactive in real-time
 - Needs: Input and interaction technology
- Registered in 3D
 - Needs: Viewpoint tracking technology

AR/VR Head Mounted Displays

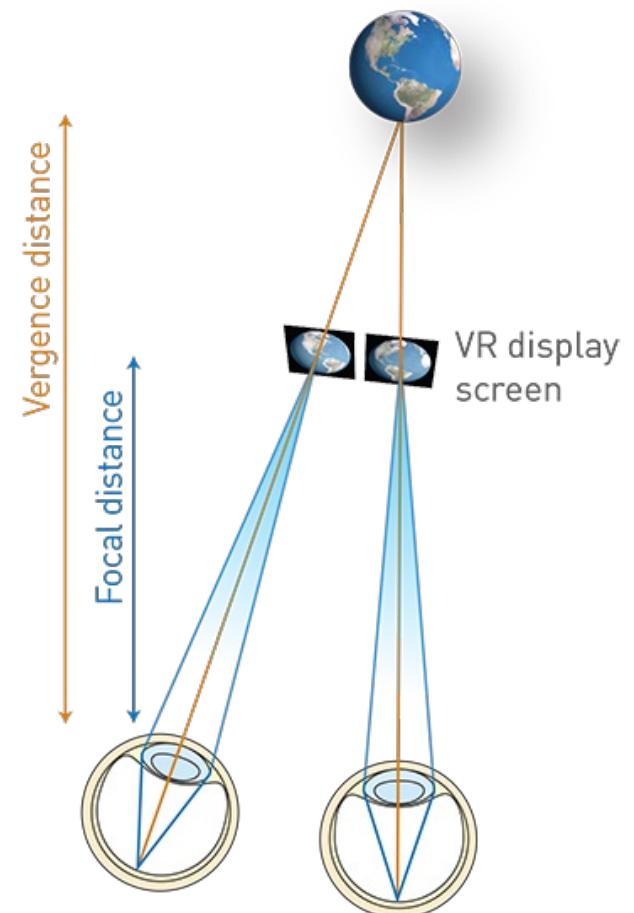
Inside View

Oculus VR's first commercial headset was built using just a few components, including off-the-shelf electronics and simple lenses.

THE BUILD



Virtual object as seen in head-mounted display



Key Properties of HMDs

- **Lens**
 - Focal length, Field of View
 - Ocularity, Interpupillary distance
 - Eye relief, Eye box
- **Display**
 - Resolution, contrast
 - Power, brightness
 - Refresh rate
- **Ergonomics**
 - Size, weight
 - Wearability

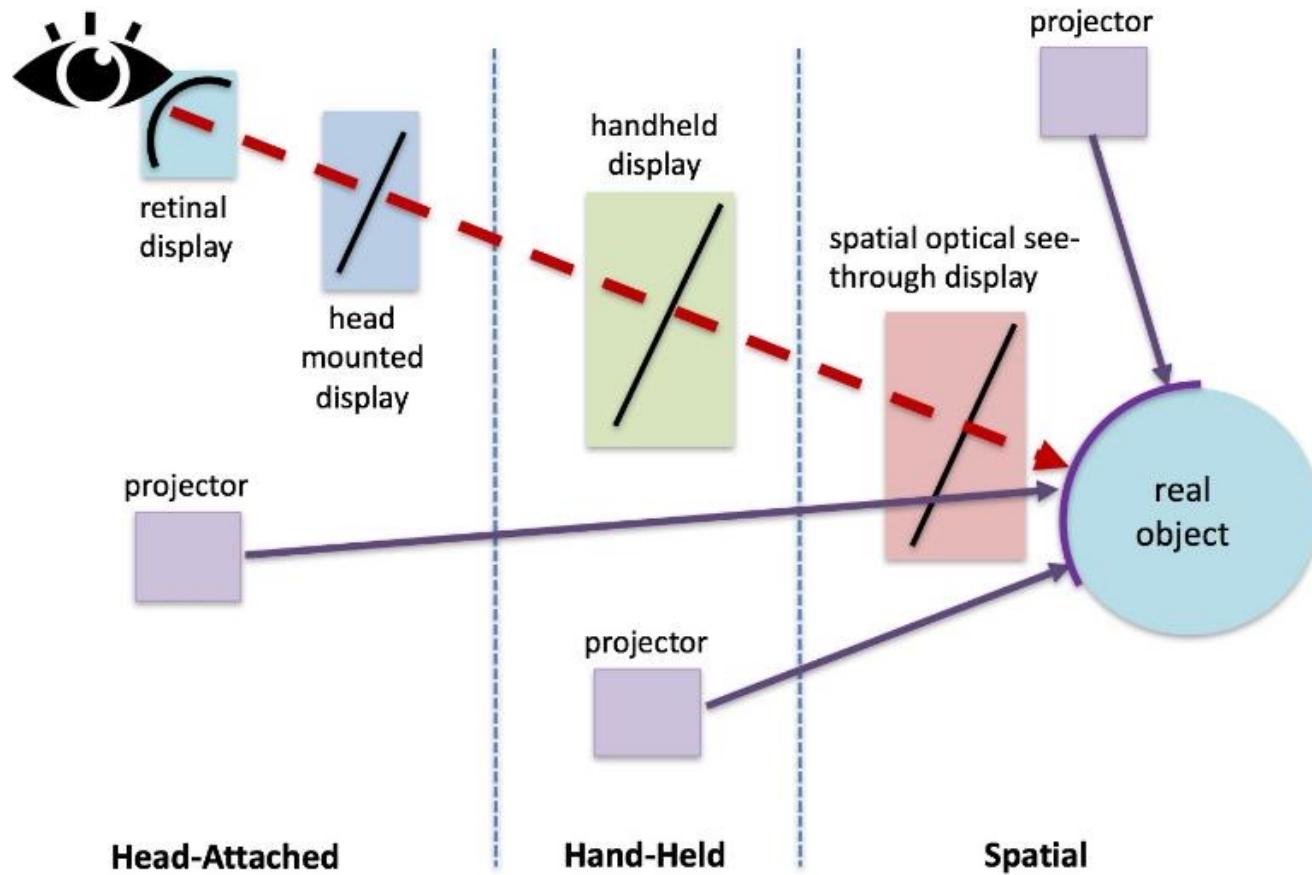


Display Technologies

- Head attached
 - Head mounted display/projector
- Body attached
 - Handheld display/projector
- Spatial
 - Spatially aligned projector/monitor



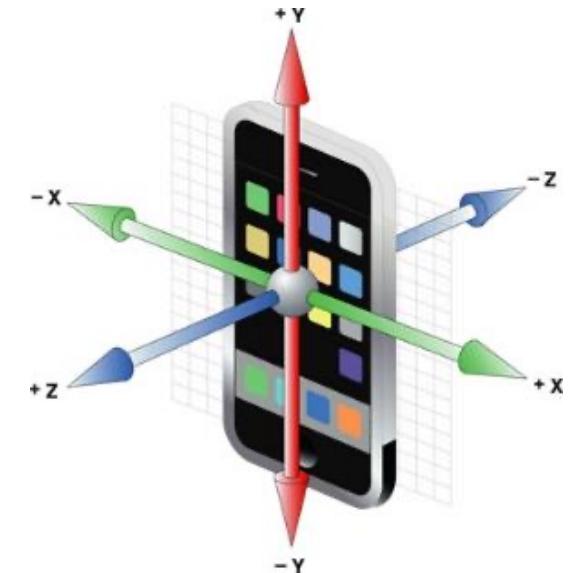
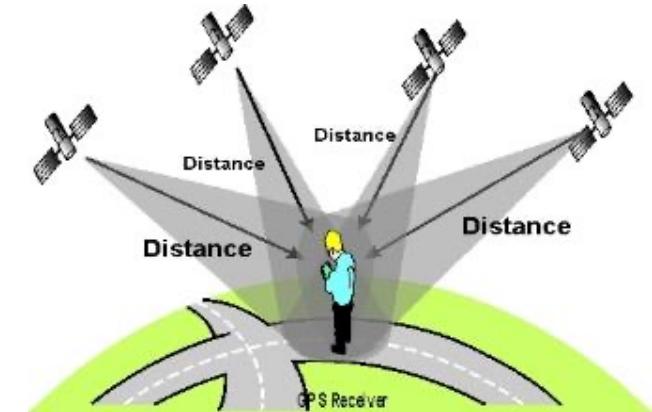
Display Taxonomy



Bimber, O., & Raskar, R. (2005). *Spatial augmented reality: merging real and virtual worlds*. CRC press.

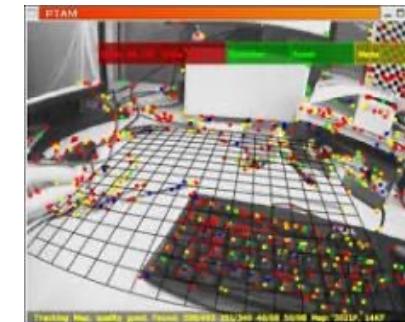
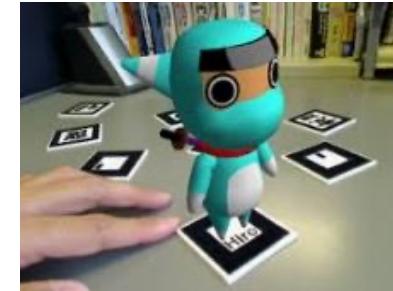
Tracking

- Continually locating user's viewpoint
- Active
 - Mechanical, Magnetic, Ultrasonic
 - GPS, Wifi, cell location
- Passive
 - Inertial sensors (compass, accelerometer, gyro)
 - Computer Vision
- Hybrid Tracking
 - Combined sensors (e.g. Vision + Inertial)



Computer Vision Tracking

- **Marker Tracking**
 - Tracking known artificial markers/images
 - e.g. ARToolKit square markers
- **Markerless Tracking**
 - Tracking from known features in real world
 - e.g. Vuforia image tracking
- **Unprepared Tracking**
 - Tracking in unknown environment
 - e.g. SLAM tracking



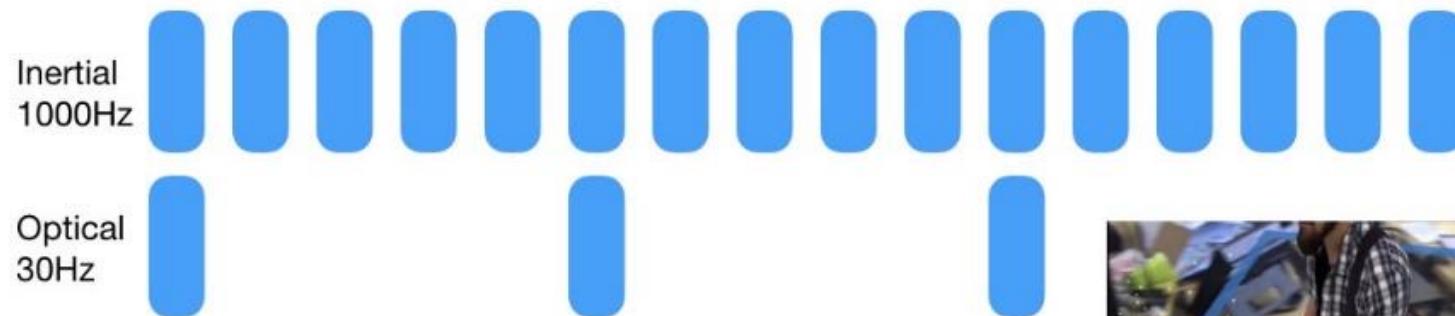


<https://www.youtube.com/watch?v=1Qf5Qew5zSU>

Hybrid Example - ARKit

Visual Inertial Odometry

- Visual Inertial Odometry - track where I am moving as I move



- Slow camera
- Fast IMU, but drifts
- If camera drops out IMU takes over
- Camera corrects IMU errors



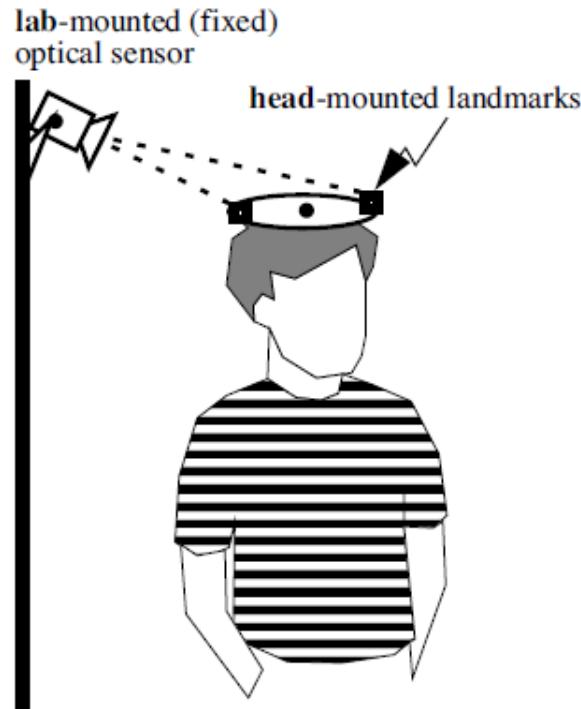


Apple ARKit Tracking Demo

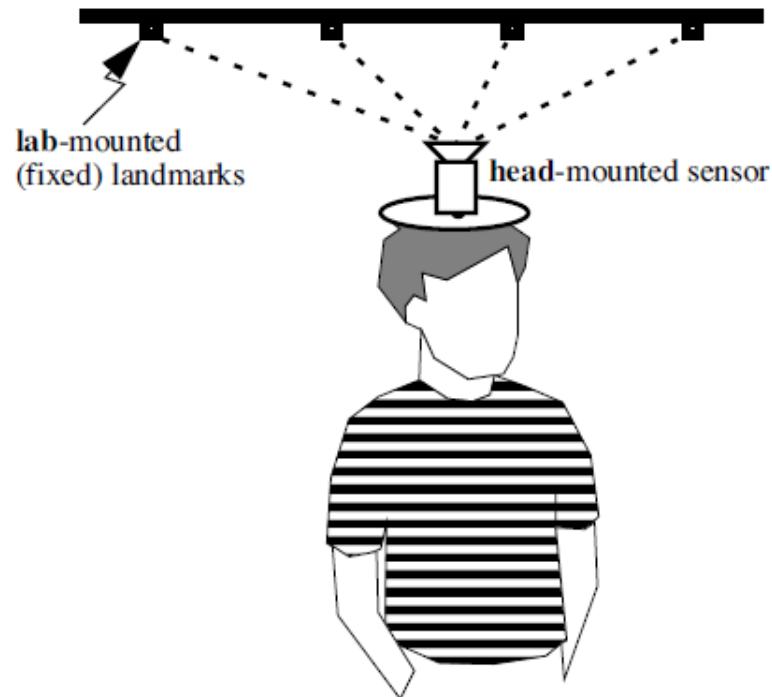


- <https://www.youtube.com/watch?v=dMEWp45WAUg>

Outside-In vs. Inside-Out Tracking



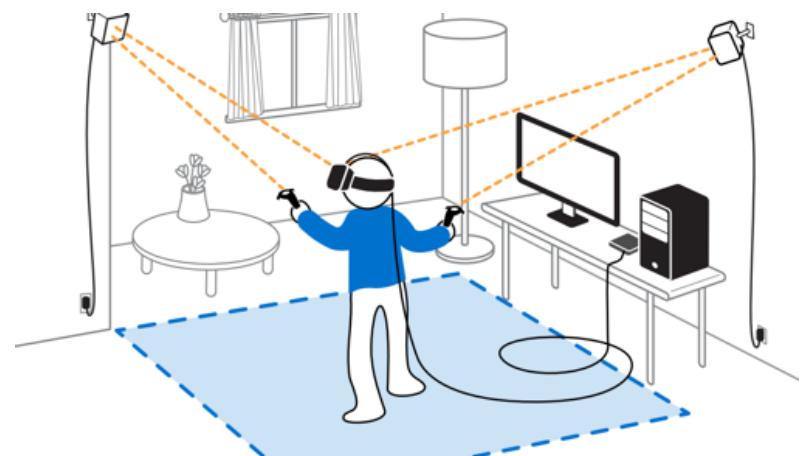
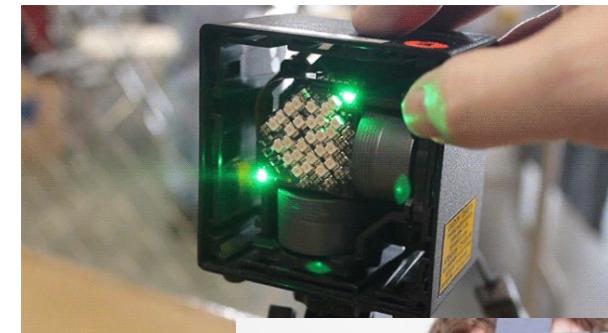
Outside-Looking-In



Inside-Looking-Out

Outside-In Tracking - Lighthouse

- 2 base stations
 - Each with 2 laser scanners, LED array
- Headworn/handheld sensors
 - 37 photo-sensors in HMD, 17 in hand
 - Additional IMU sensors (500 Hz)
- Performance
 - Tracking server fuses sensor samples
 - Sampling rate 250 Hz, 4 ms latency



Inside Out VR Tracking

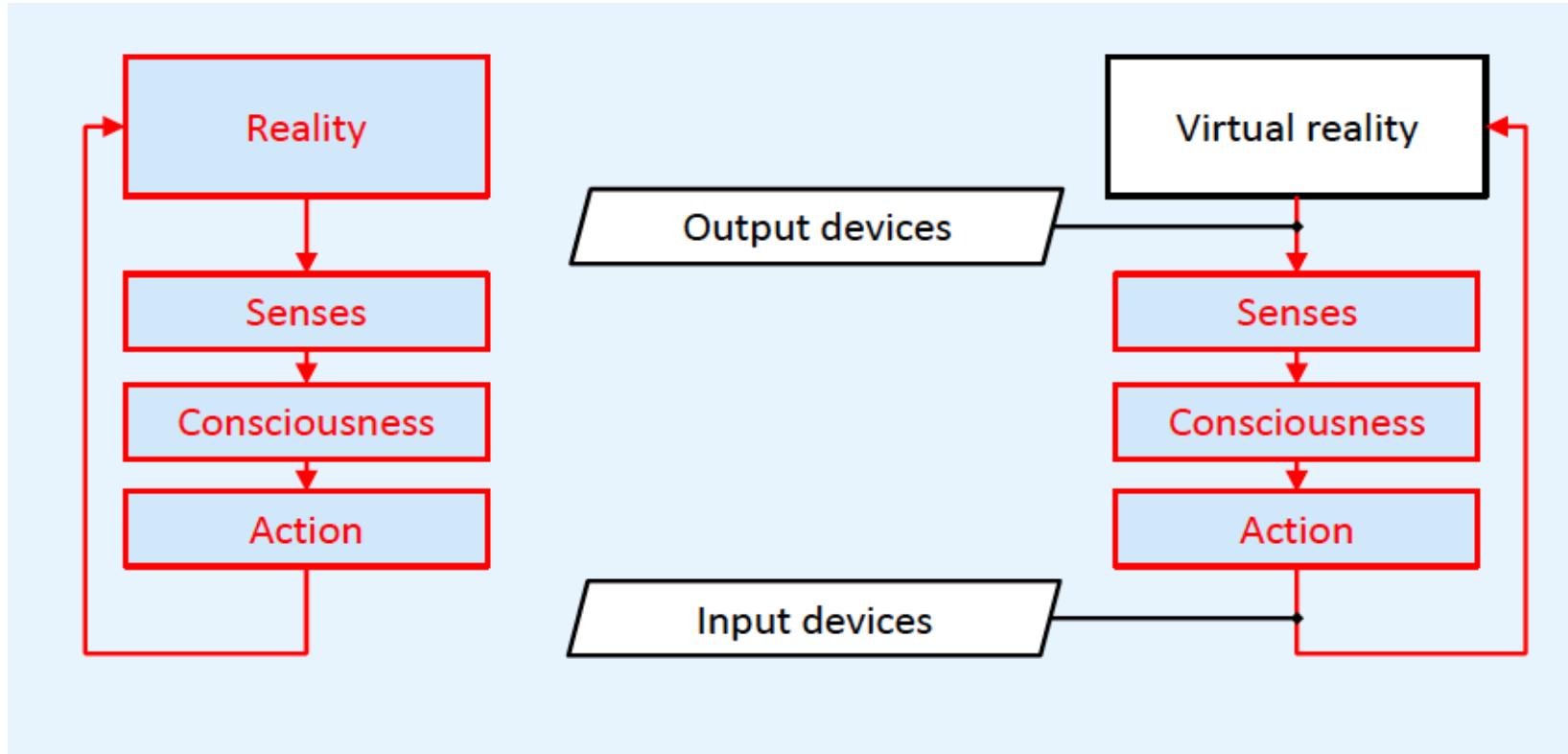
- Inside out tracking
- Four cameras on corner of display
- Searching for visual features
- On setup creates map of room



Oculus Quest
(4x corner cameras)



Input and Interaction



- In a VR system there are input and output devices between human perception and action

Tracked Handheld Controllers



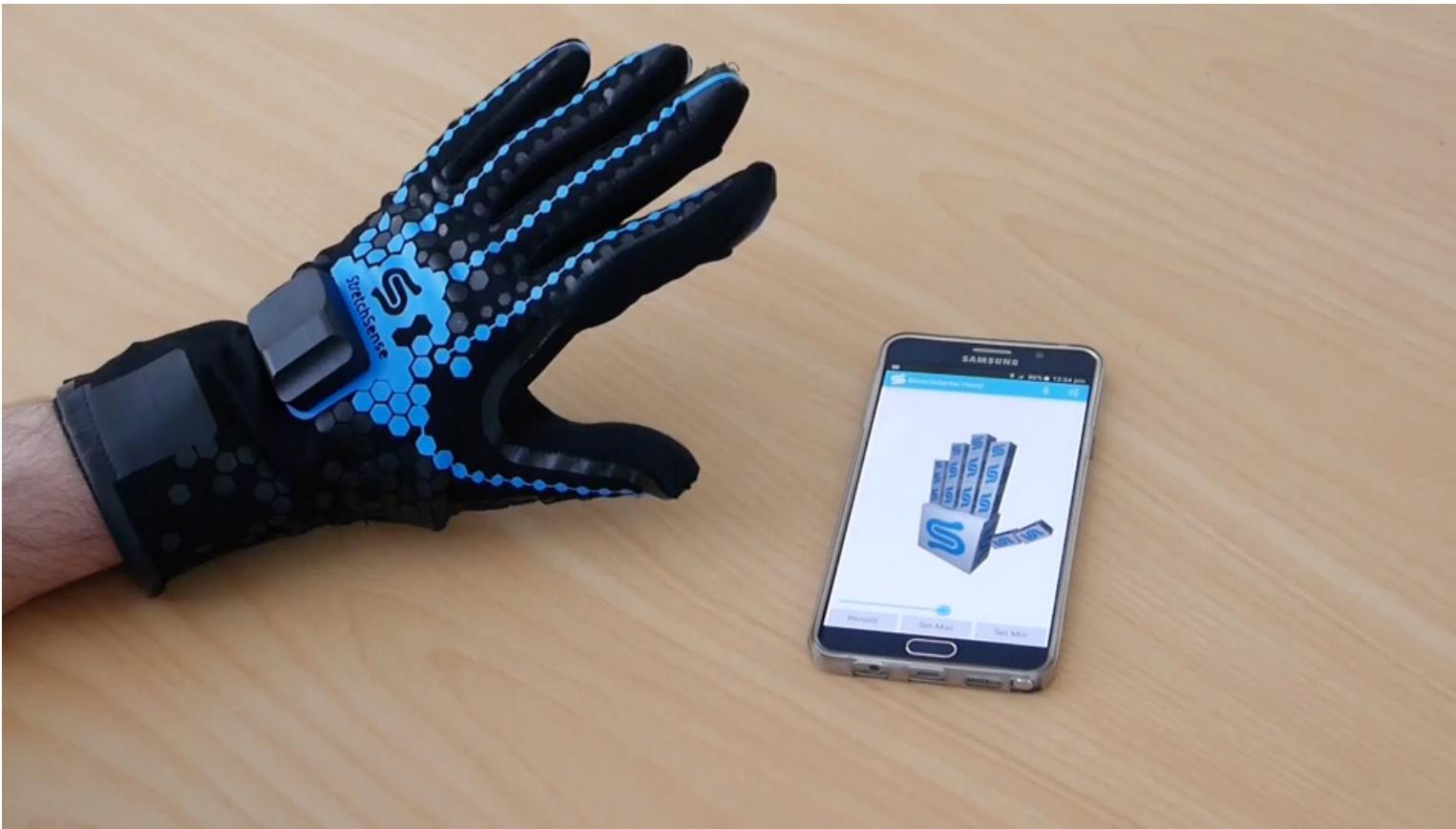
HTC Vive Controllers



Oculus Touch Controllers

- Handheld controller with 6 DOF tracking
 - Combines button/joystick input plus tracking
- One of the best options for VR applications
 - Providing proprioceptive, passive haptic touch cues
 - Direct mapping to real hand motion

StretchSense Glove Demo



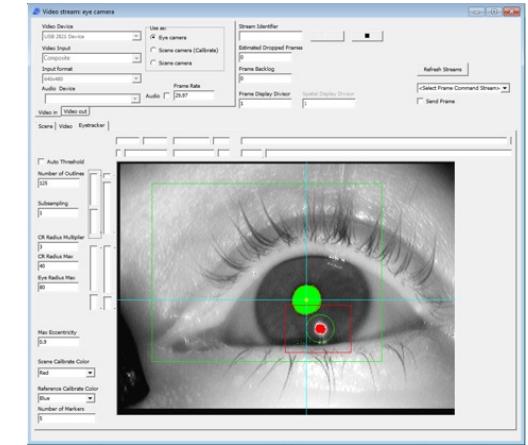
- <https://www.youtube.com/watch?v=wYsZS0p5uu8>

Bare Hands



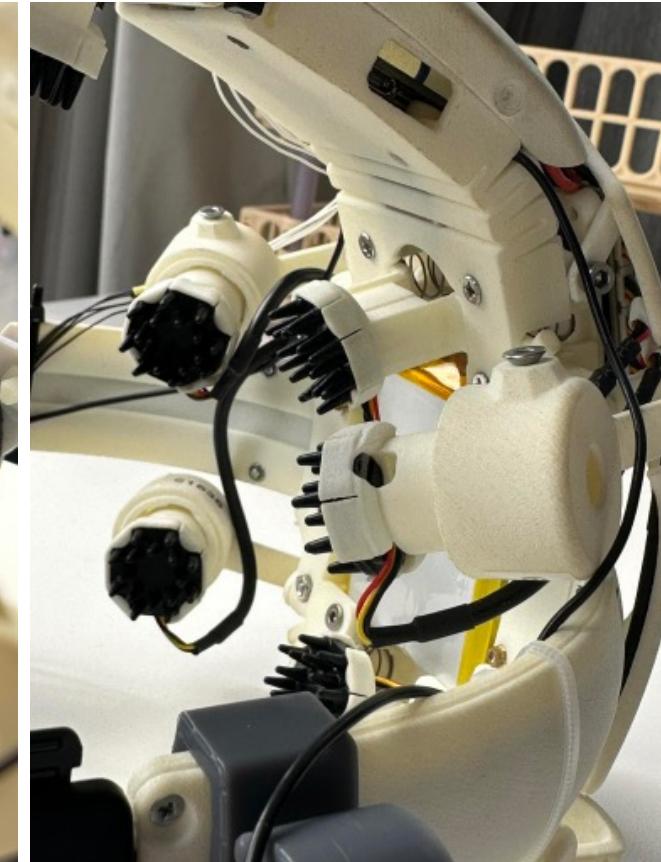
- Using computer vision to track bare hand input
- Creates compelling sense of Presence, natural interaction

Eye Tracking



- **Technology**
 - Shine IR light into eye and look for reflections
- **Advantages**
 - Provides natural hands-free input
 - Gaze provides cues as to user attention

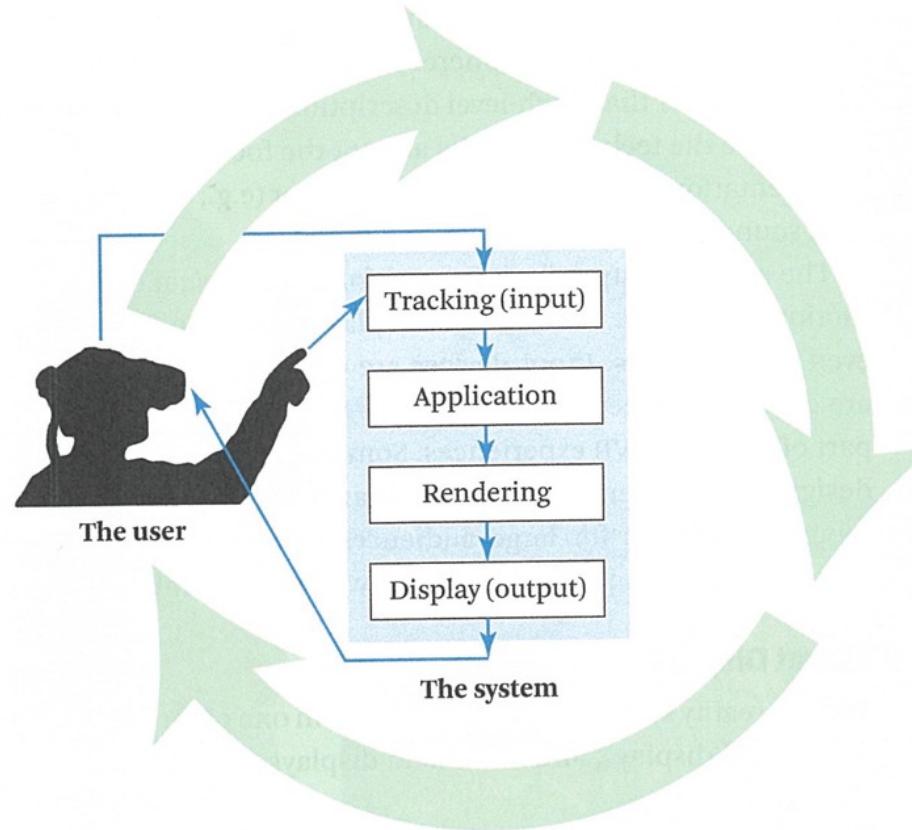
Sensor Enhanced HMDs



Project Galea

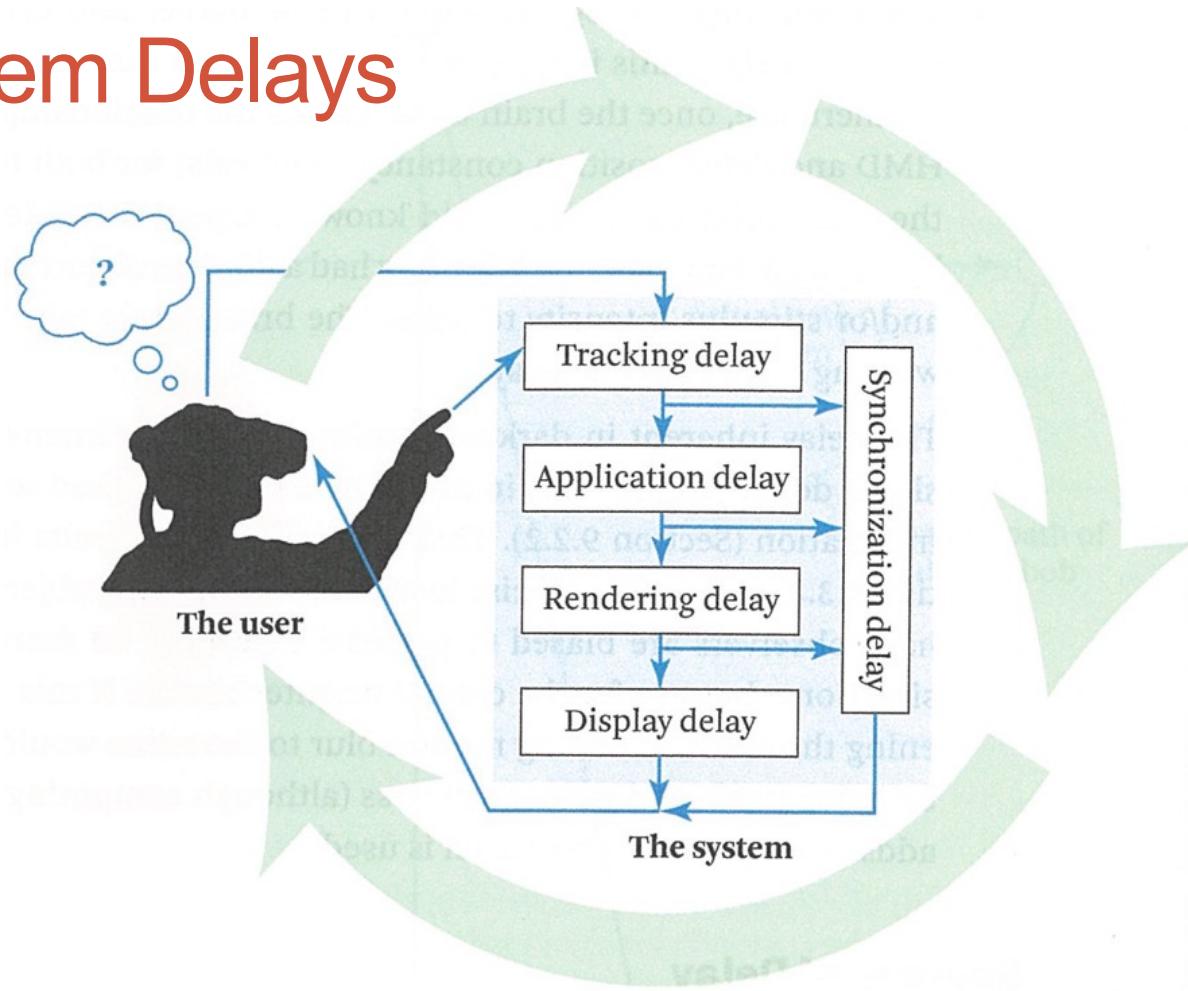
EEG, EMG, EDA, PPG,
EOG, eye gaze, etc.

Typical XR Simulation Loop



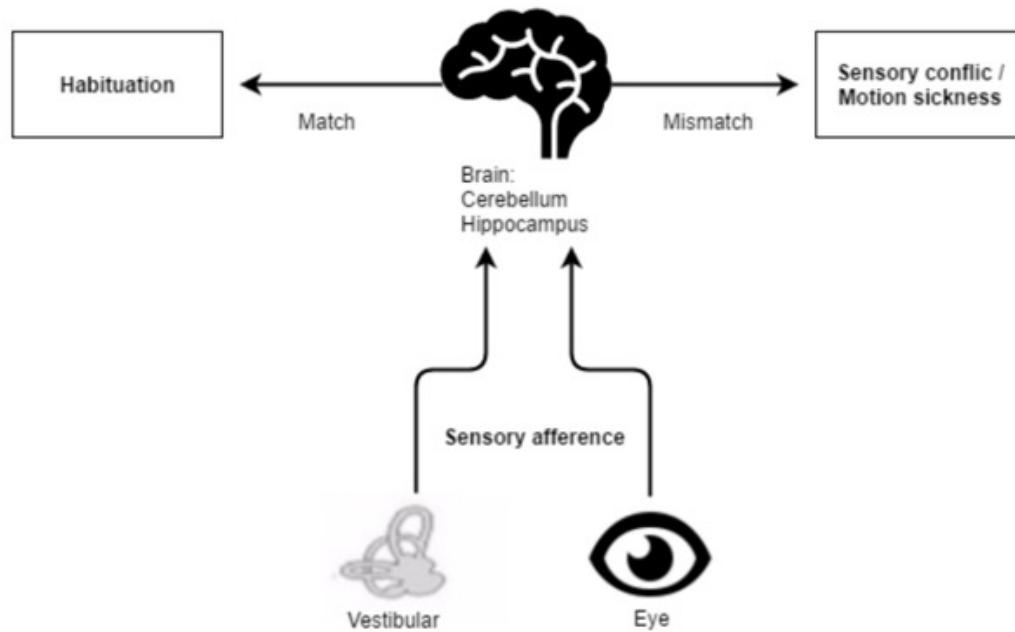
- User moves head, scene updates, displayed graphics change

System Delays



- Need to synchronize system to reduce delays

What Happens When Senses Don't Match?



- 20-30% VR users experience motion sickness
- **Sensory Conflict Theory**
 - Visual cues don't match vestibular cues
 - Eyes – “I’m moving!”, Vestibular – “No, you’re not!”

Many Causes of Simulator Sickness

- 25-40% of VR users get Simulator Sickness, due to:
- Latency
 - Major cause of simulator sickness
- Tracking accuracy/precision
 - Seeing world from incorrect position, viewpoint drift
- Field of View
 - Wide field of view creates more peripheryvection = sickness
- Refresh Rate/Flicker
 - Flicker/low refresh rate creates eye fatigue
- Vergence/Accommodation Conflict
 - Creates eye strain over time
- Eye separation
 - If IPD not matching to inter-image distance then discomfort

APPLICATIONS

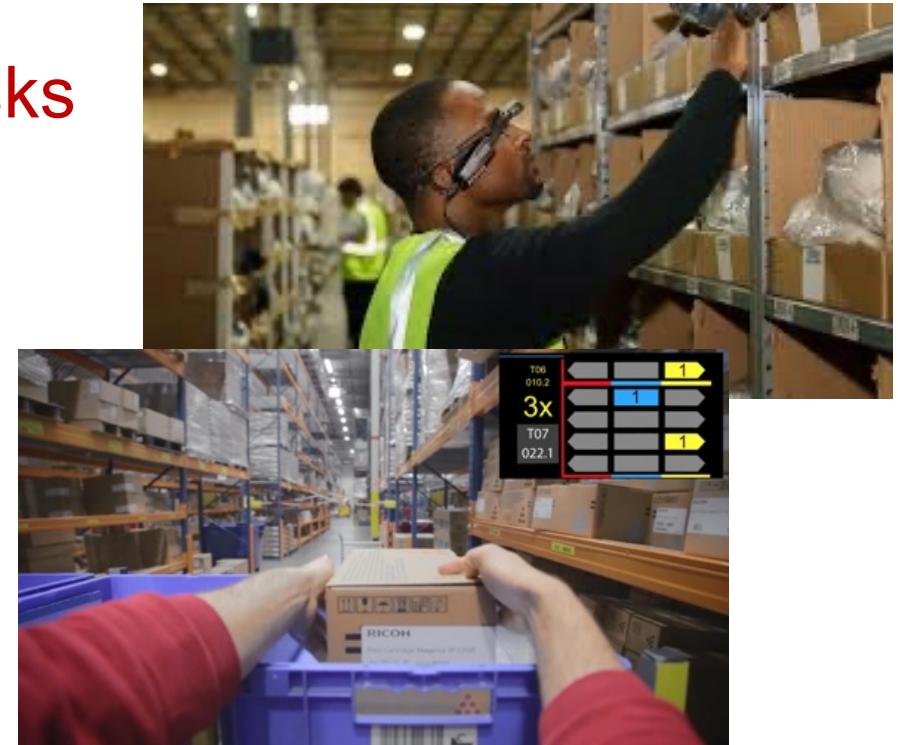
Using AR/VR/XR/MR in the Workplace

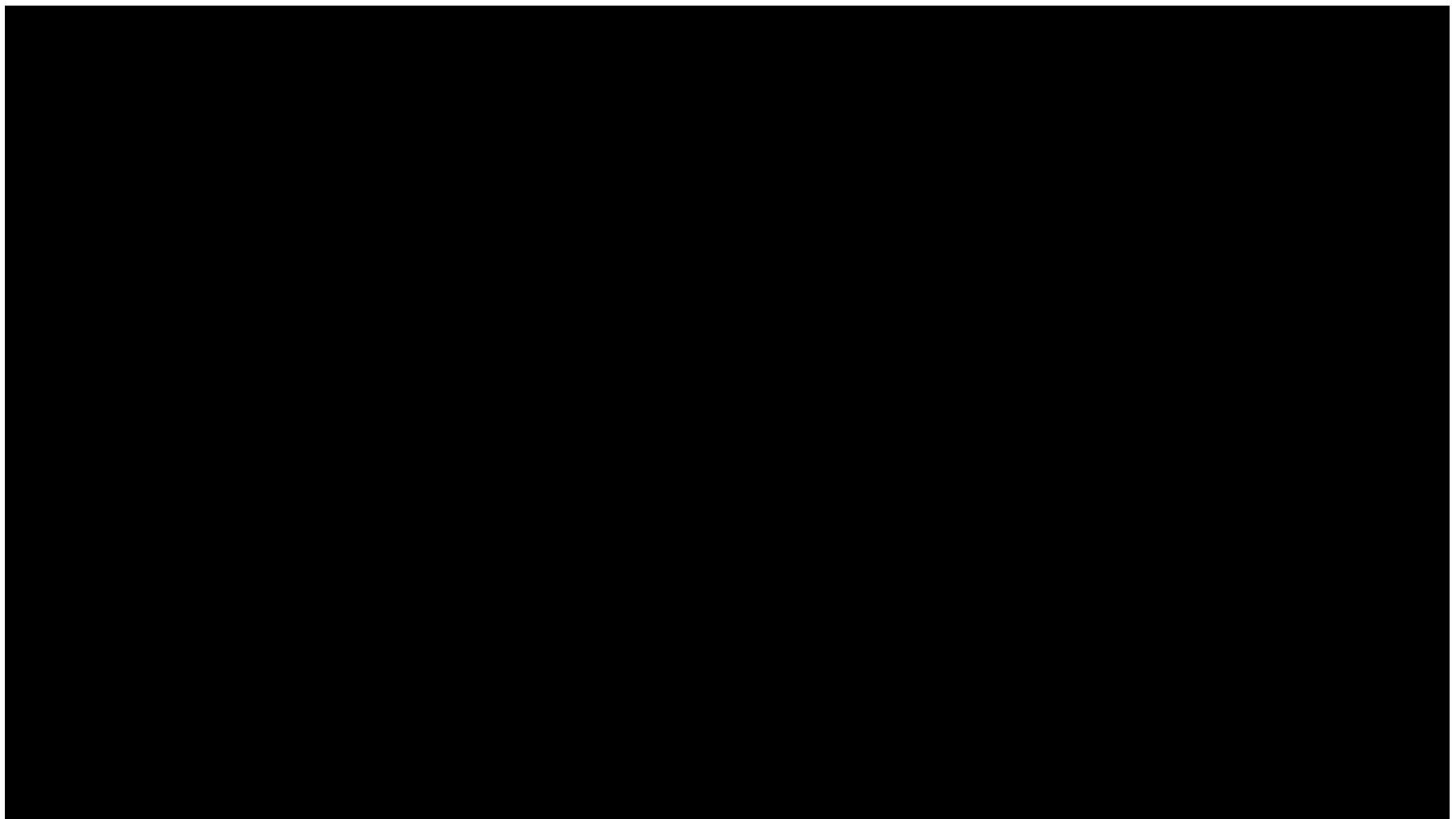


- Blending digital/physical worlds..
 - Productivity, Marketing, Training, Collaboration
- Ideal applications
 - Spatial data, connection to real world, mobile/hands busy

Example: Productivity

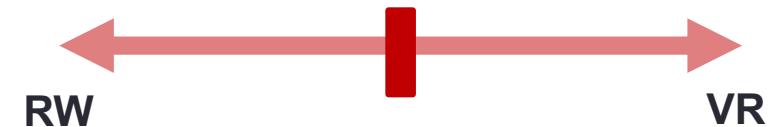
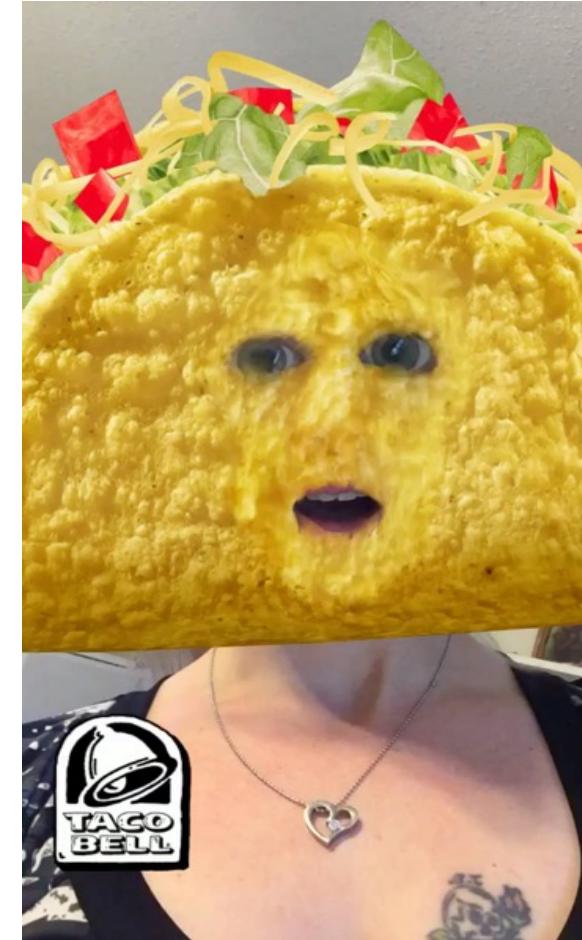
- AR technology for hands busy tasks
- Warehouse picking (DHL, SAP)
 - Smart glasses show pick location
 - Simple user interface
 - Hands-free interaction
 - Improved performance





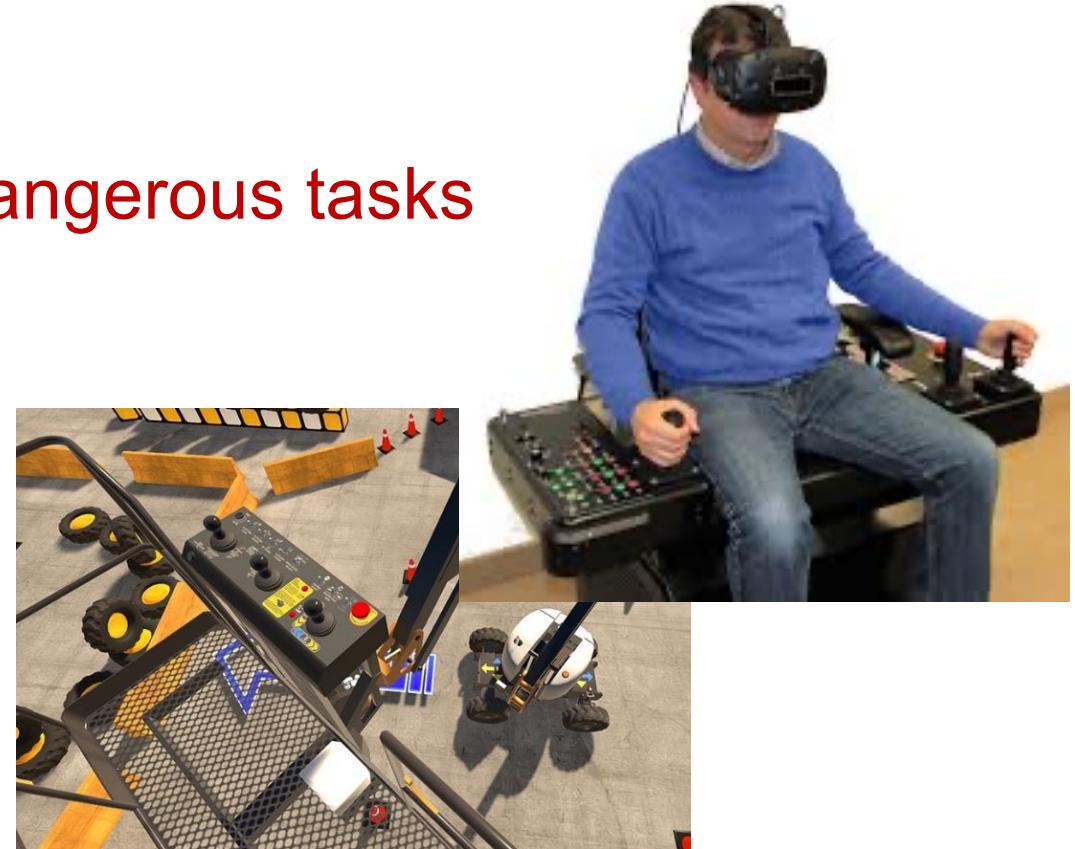
Example: Marketing

- Add magic to companies products
- Snapchat lenses
 - Face tracking technology
 - Simple interaction
 - Animated content
 - Significant ROI
 - \$750K for 224 million views
 - c.f. \$5 million for SuperBowl ad



Example: Training

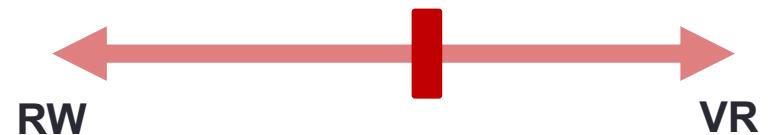
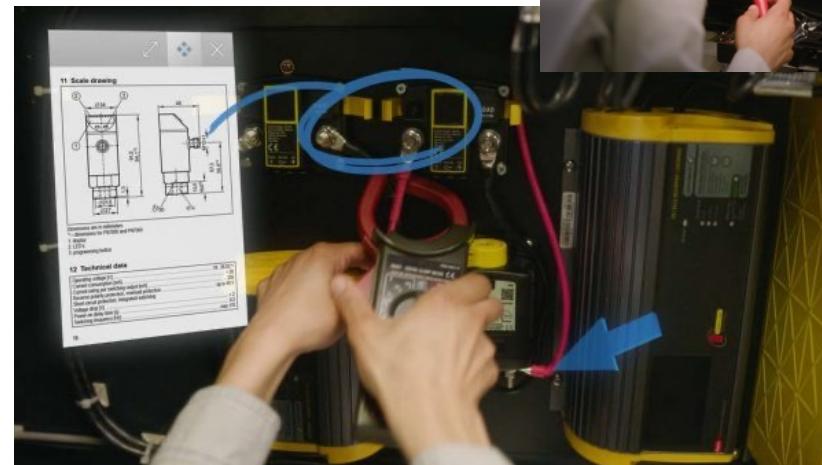
- AR/VR training for difficult/dangerous tasks
 - Train with real objects, controls
- GlobalSim Crane Simulator
 - Virtual Reality simulation
 - Train with real controls
 - Virtual body/hands
 - Realistic physics

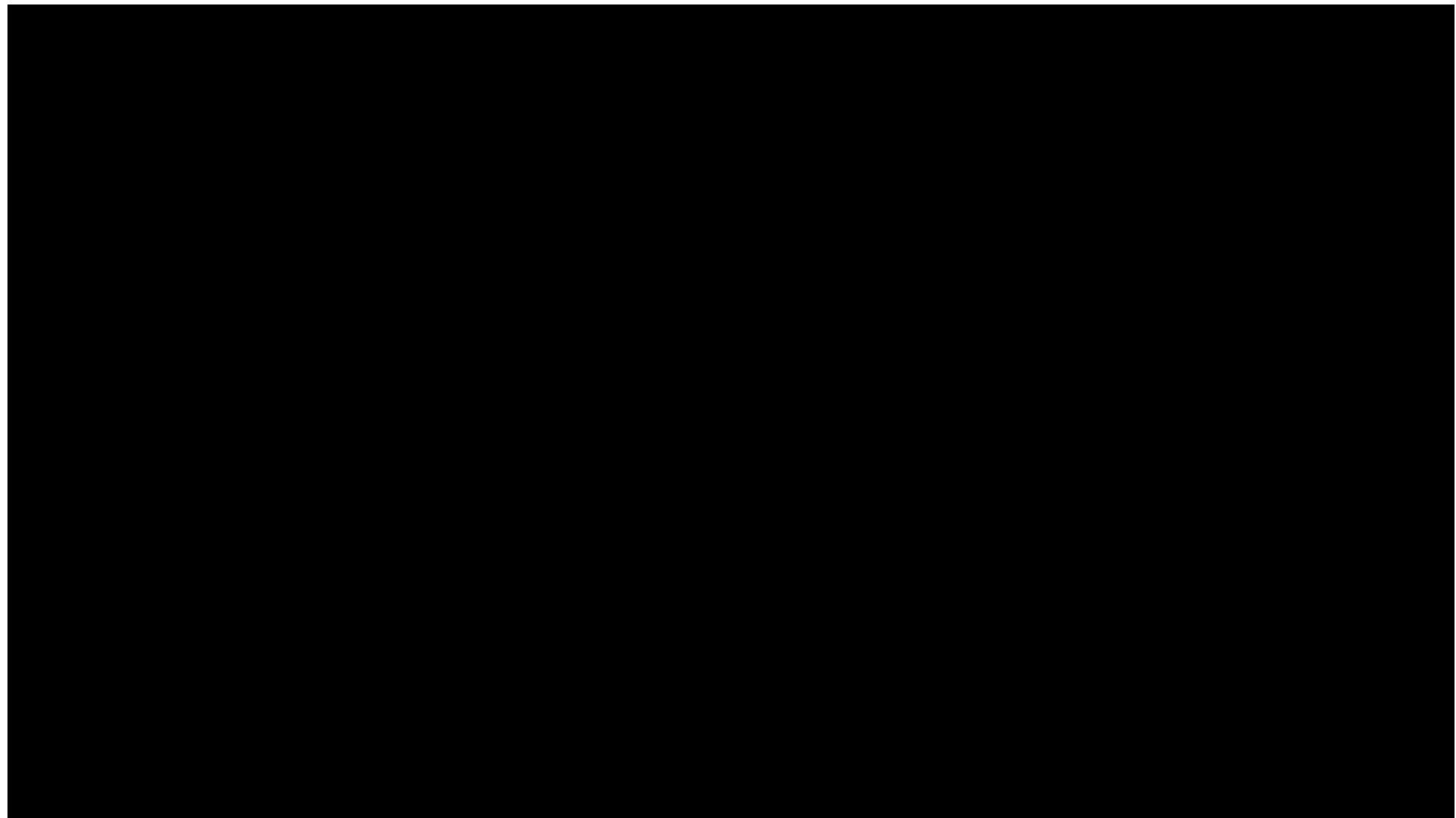




Example: Collaboration

- Use MR technology for remote collaboration
 - See through another's eyes, virtual avatars
- Microsoft Remote Assist
 - HoloLens see-through display
 - Remote video sharing
 - AR drawing/annotation
 - Document sharing
 - Group calling



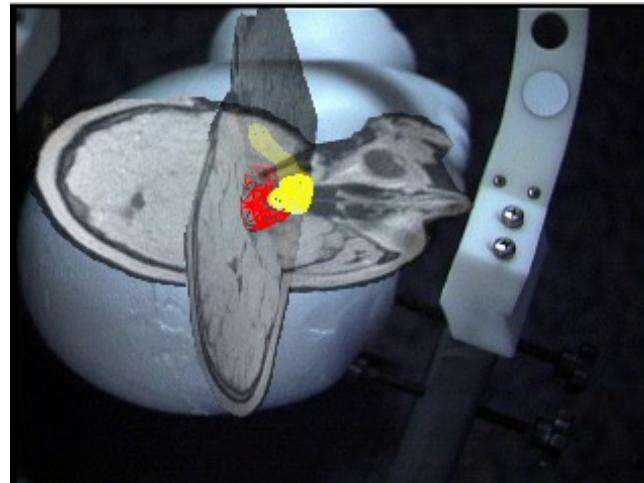
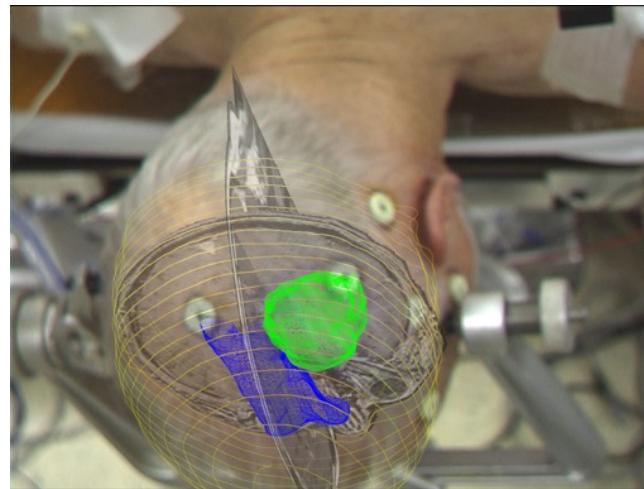


Medical AR Trials

- Sauer et al. 2000 at Siemens Corporate Research, NJ
- Stereo video see through



F. Sauer, Ali Khamene, S. Vogt: An Augmented Reality Navigation System with a Single-Camera Tracker: System Design and Needle Biopsy Phantom Trial, MICCAI 2002



Medical Demo



Gaming: Rock-em Sock-em



- Shared AR Demo
- Markerless tracking

Rockem Sockem Demo



USING AI FOR XR

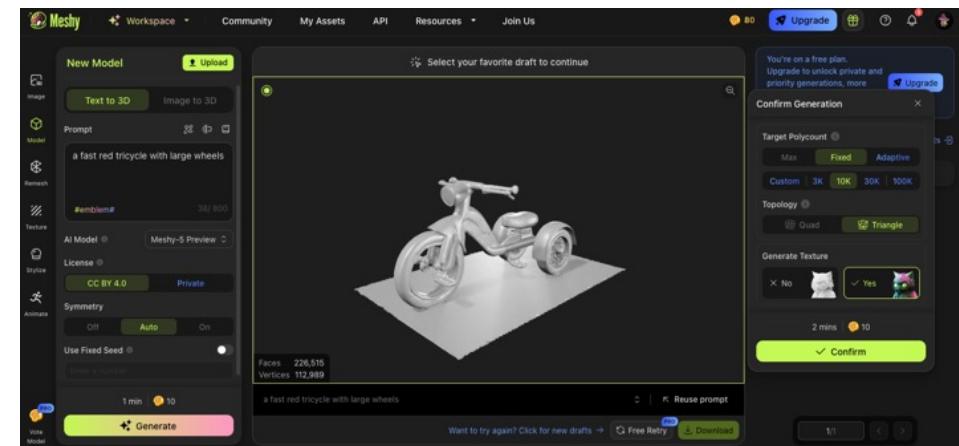
Using AI for XR

- Accelerated Content Creation
 - Gen. AI for 3D Assets, Procedural Content Generation, Optimization
- Intelligent Prototyping Tools and Workflows
 - AI-Assisted Interface Design, User simulation, Real-time Feedback
- Enhanced User Experience and Personalization
 - Adaptive and Context-Aware XR, Natural User Interfaces, Avatars
- Data Analysis and Insights
 - User Analytics, Emotion detection

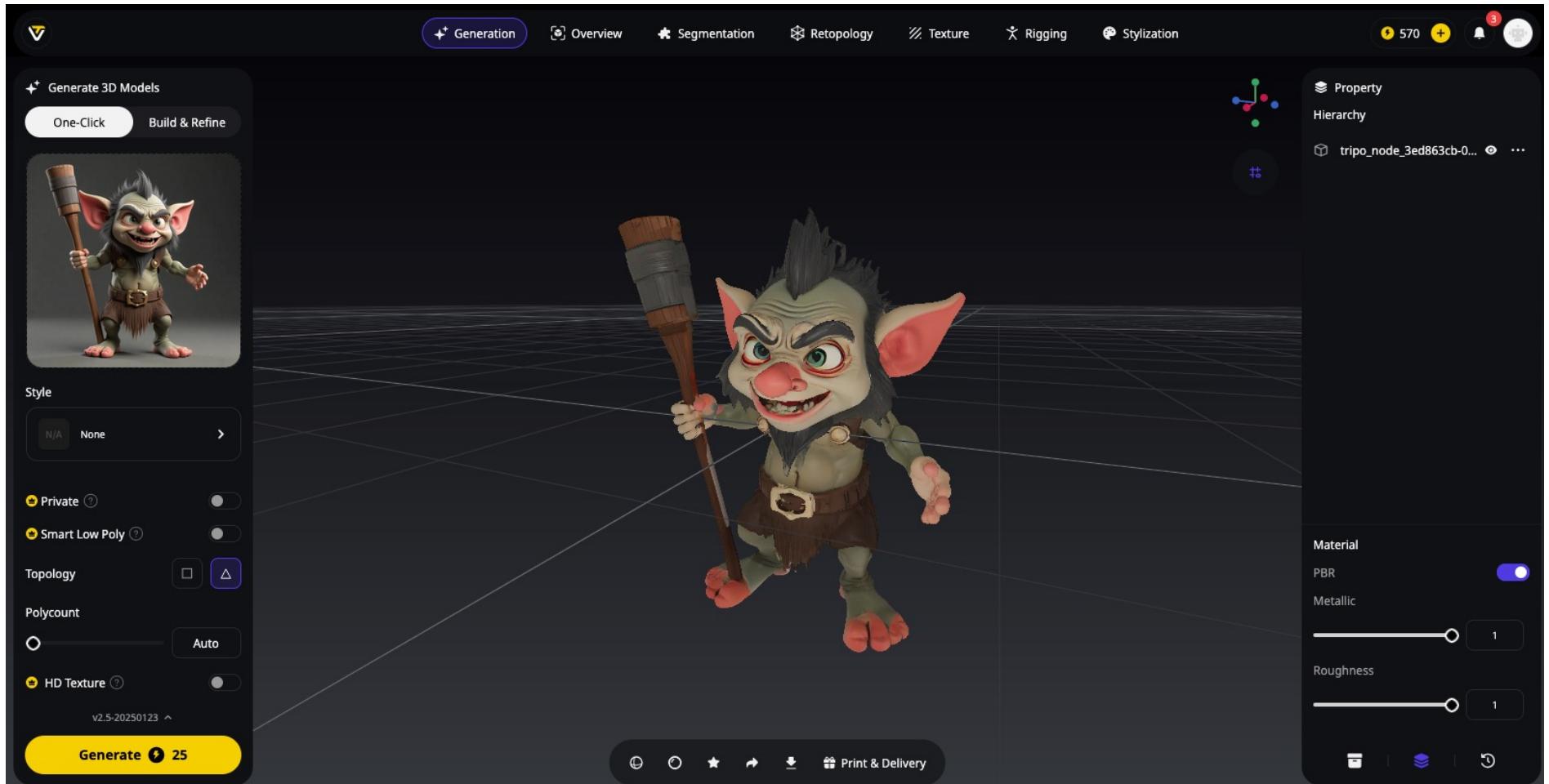
CONTENT CREATION

Generative AI and Content Creation

- Many content creation tools
- 2D images
 - Midjourney, DALL-E 3, Imagen 3, Stable Diffusion
- 360 images
 - Blockade Labs, Stable Diffusion, PromeAI
- 3D models
 - Text to 3D - Meshy, Tripo AI, Hunyuan3D
 - Image to 3D - Alpha3D, Rodin by Hyper 3D
- Audio
 - Suno AI, Udio, ElevenLabs, LoudMe
- Video
 - Luma Dream Machine, Runway ML, Sora



Tripo AI – www.trip03d.ai



“A cute troll carrying a giant club”

Rodin – Hyper3d - <https://hyper3d.ai/>



The screenshot shows the Hyper3d AI interface. At the top, there's a user profile icon and the name "Billinghur...". Below that are sharing options: "Share", "Eye", "Heart", and a lock icon. The main area displays a 3D rendering of the pink troll character with a blue dress, set against a dark background with a camera icon in the top left corner. To the right of the 3D model, there are sections for "Geometry Generation" and "Material Generation".
Geometry Generation: A search bar contains the query "pink troll character with blue dress." Below it are filter buttons: "Detailed" (which is selected), "smooth", "complex", "game-ready", and "character". A progress bar is shown below these filters.
Material Generation: This section shows a preview of the generated material and includes a "Confirm" button. There are also "Redo" and "x2" buttons.
Pack: This section allows for file export. Under "Geometry", "Base Model" is checked, while "LOD" and "High-poly" are unchecked. File formats listed are ".obj", ".fbx", ".gib", ".usdz", and ".stl". Under "Material", "Shaded" and "PBR" are unchecked, with resolution settings for "2K" and "4K". Buttons for "Publish" and "Download" are at the bottom.

- 2D image to 3D model

Procedural Content Generation



- Automatically generate virtual content
 - Landscapes, buildings, objects, and textures dynamically.
 - Enables creation of very large environments/experiences
 - E.g. Minecraft, No man's sky
 - Automatically generate world at run time

Automatic Asset Optimization



48,000 polygons



convrse.ai



2,800 polygons

- AI optimization of 3D models and textures
- Improve real time performance
- Creating smooth user experience

INTELLIGENT PROTOTYPING TOOLS

Intelligent Prototyping Tools and Workflows

- **AI-Assisted Interface Design:**
 - Use AI to automate repetitive design tasks,
 - Provide intelligent design suggestions,
- **Automated User Interaction and Behaviour:**
 - Simulate user behaviour within a prototype,
 - Test interaction patterns and identify potential usability issues.
- **Real-time Feedback and Iteration:**
 - Provide instant feedback on design choices,
 - Enabling rapid iteration and refinement of prototypes.

Google Gemini Assistance with Prototyping

I am developing an Augmented Reality interface for a head mounted display that uses gaze and gesture input. What are four ways that gaze and gesture can be used...

- ★ Here are four ways gaze and gesture can be used to select virtual objects at a distance (over 10 feet) in an Augmented Reality head-mounted display:

1. Gaze-Dwell with Hand Confirmation Gesture:

- **Gaze:** The user gazes at the virtual object for a predefined duration (e.g., 1-2 seconds). This "dwell" acts as an initial targeting mechanism, highlighting or providing feedback on the gazed object.
- **Gesture:** Once the object is highlighted by gaze-dwell, the user performs a simple, deliberate hand gesture to confirm selection. This could be a "pinch" gesture, a "thumbs

Which of these is most efficient and why?

- When considering efficiency in Augmented Reality interaction, it's crucial to balance speed, accuracy, and user comfort/cognitive load. Each of the four methods has its trade-offs: 
- 1. Gaze-Dwell with Hand Confirmation Gesture:
 - Efficiency: Generally less efficient for rapid, repeated selections due to the dwell time. The required fixation period can feel slow and unnatural, leading to what's known as the "Midas Touch" problem (unintentional selections from simply looking at something). However, it offers high accuracy and reduces accidental selections. 
 - Why: Dwell time introduces a deliberate delay. While this improves accuracy by requiring clear intent, it inherently slows down the interaction. Hand gestures for confirmation can also add a small amount of time and physical effort compared to a purely gaze-based

How would I implement Gaze-Dwell with Hand Confirmation Gesture in Unity for a Quest Pro display?



Implementing Gaze-Dwell with Hand Confirmation Gesture in Unity for a Quest Pro display involves leveraging Unity's XR Interaction Toolkit, the Quest Pro's eye tracking, and its robust hand tracking capabilities.

Here's a breakdown of the implementation steps, concepts, and key components you'd use:

Core Concepts & Components

- **XR Interaction Toolkit:** Unity's official framework for building interactive XR experiences. It provides powerful tools for gaze interaction, hand tracking, and various interactors/interactables.

Implementation Steps

1. Project Setup & XR Interaction Toolkit

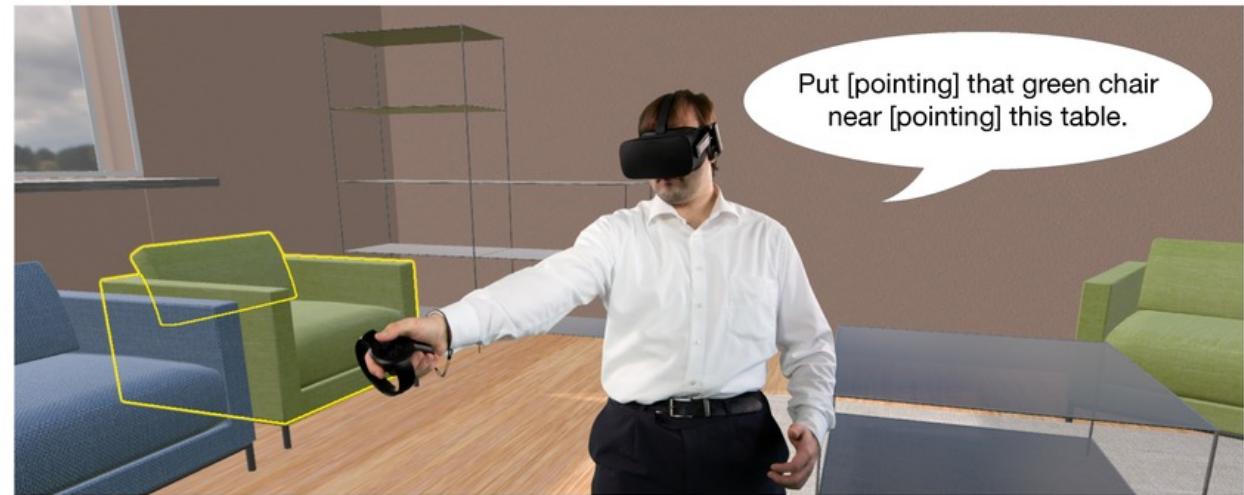
1. Create a New Unity Project: Use a 3D Core or URP template.
2. Install XR Plug-in Management: Go to `Window > XR > XR Plug-in Management`. Install it if not already.
3. Install Meta XR Plugin: In `XR Plug-in Management`, select the Android tab and tick `Meta XR Plugin`.
4. Install XR Interaction Toolkit: Go to `Window > Package Manager`. Select "Unity Registry" and install `XR Interaction Toolkit` (recommended latest verified version). Also, install the `XR Interaction Toolkit Samples` (especially "Starter Assets" and "XR Device Simulator" if you want them).
5. Configure Project for Hand & Eye Tracking:

ENHANCED USER EXPERIENCE

Enhanced User Experiences

- **Adaptive and Context-Aware XR:**
 - Analyze user behaviour, biometric signals, and environmental factors
 - Dynamically adjust the virtual world - creates personalized experiences.
- **Natural User Interfaces (NUIs):**
 - AI-powered gesture recognition, voice commands (NLP), and eye-tracking
 - More intuitive and natural XR interactions, mimicking real-world actions.
- **Intelligent NPCs and Avatars:**
 - Drive more realistic non-player characters and avatars
 - Enabling dynamic conversations, and human-like expressions and movements.
- **Personalized Learning Paths and Training:**
 - AI can analyze a trainee's performance and adapt the material
 - Provide targeted feedback, making learning more effective and personalized

Natural User Interfaces



- Using AI to understand user intent
- Gesture, gaze understanding
 - Using AI/ML for recognition
- Multimodal understanding
 - Fusing gesture + speech, gaze + gesture, etc.

Intelligent NPCs and Avatars



- Using AI to add intelligent conversational characters to XR
- Users can have free conversation with them
 - Speech recognition, dialogue understanding
- Characters can interact with one another
- Many options
 - conv.ai, inworld.ai, custom build, etc

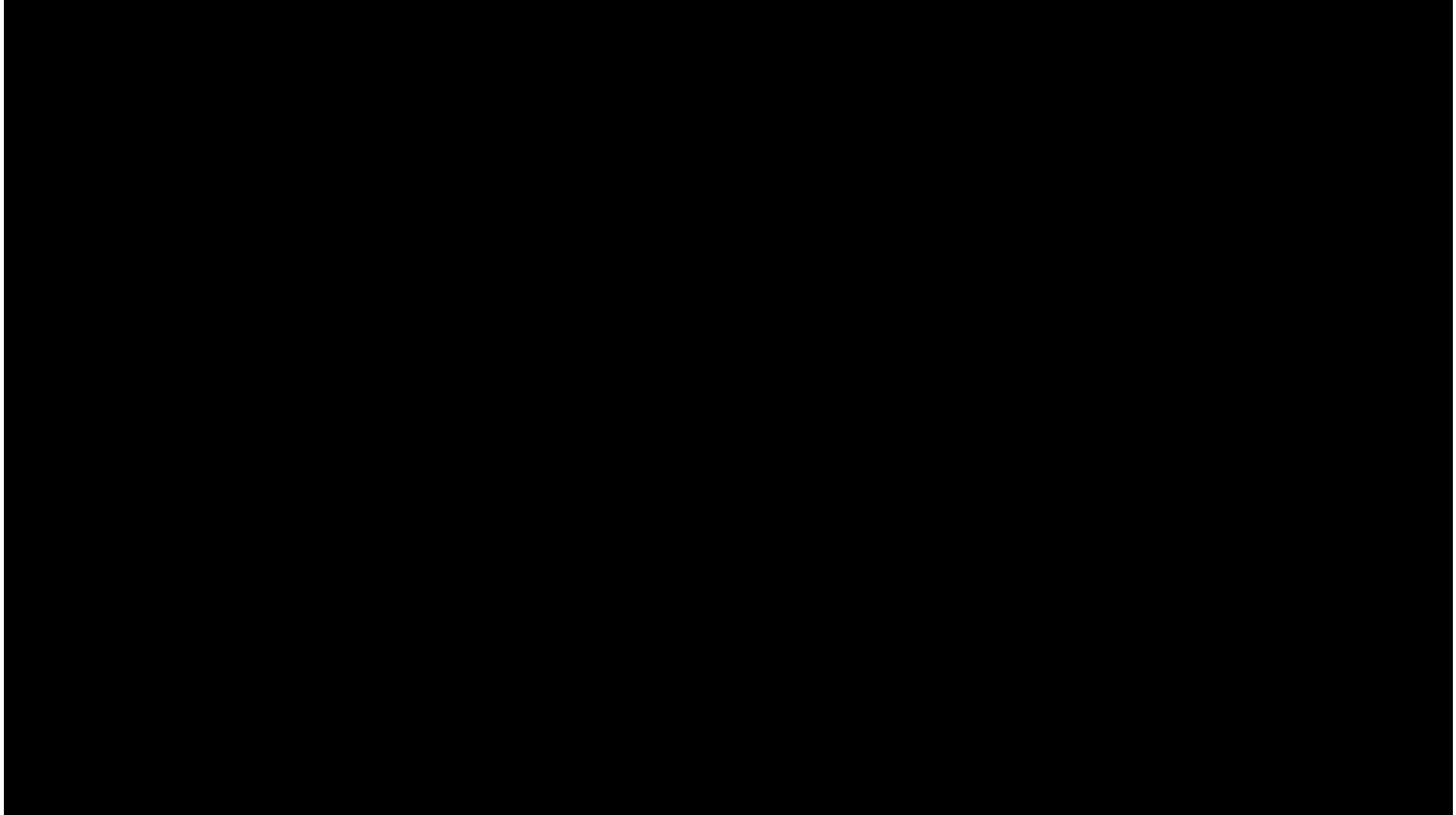
conv.ai - <https://convai.com/>



A screenshot of the Knowledge Bank section of the conv.ai website. The title "Knowledge Bank" is at the top. Below it are sections for "Character Description", "Language And Speech", and "Knowledge Bank" (which is highlighted in green). Other sections include "Personality & Style", "State Of Mind", "Memory", "Core AI Settings", "Actions", and "Narrative Design", "Visibility". A central area has a placeholder "Select a file or drag and drop here" with a "Select File" button. Below this is a table titled "Available files on your account" with two entries: "Detective ..." (2.17 KB, File available, Connect) and "c00446517...." (133.62 KB, File available, Connect).

- **AI driven conversational characters**

- **Actions:** Characters can understand the environment and perform actions
- **Knowledge Bank:** Customized domain knowledge for characters
- **Intelligent Animations:** Integrated with Lip-Sync and custom facial/body animations
- **Unity Integration:** Easily add into AR/VR/MR environments, Unity asset plug-in



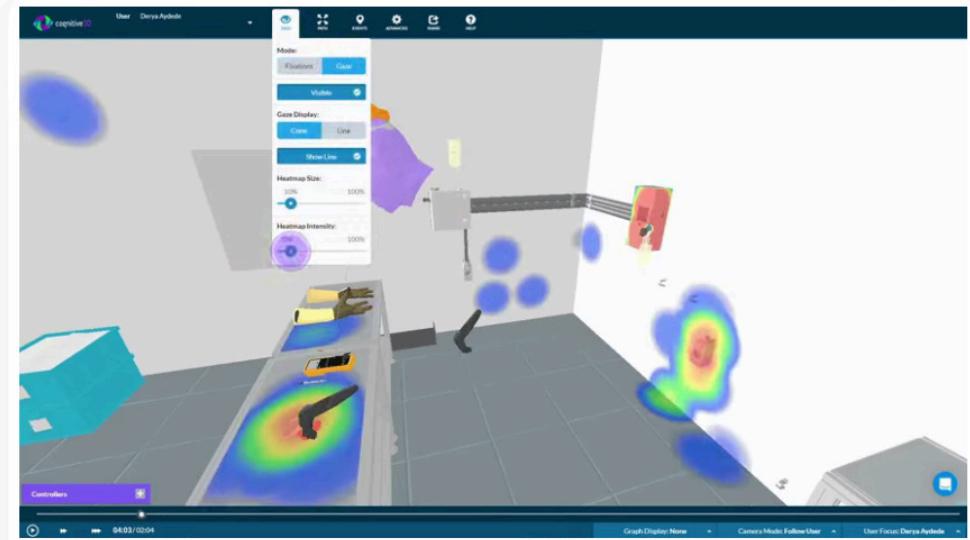
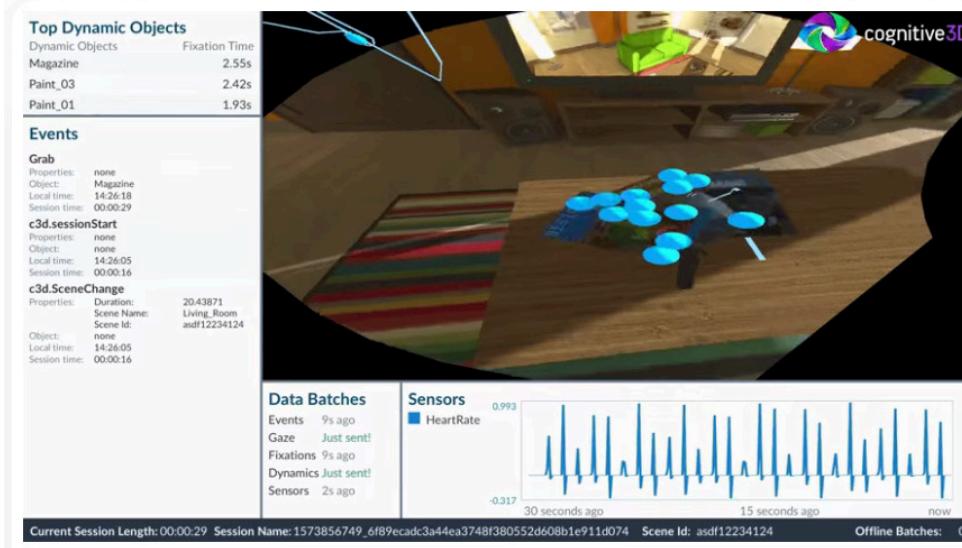
<https://www.youtube.com/watch?v=U0iupl4n3fs>

DATA ANALYSIS AND INSIGHTS

Data Analytics and Insights

- **User Analytics:**
 - AI can analyze data collected during user testing of XR prototypes
 - e.g., gaze tracking, dwell times, interaction patterns
 - Identify popular features, bottlenecks, and areas for improvement.
 - Data-driven approach helps optimize the user experience.
- **Emotion Detection:**
 - analyze facial microexpressions or voice tone to understand user sentiment
 - providing deeper insights into user engagement and satisfaction.

Cognitive3D - <https://cognitive3d.com/>



- Capture capture and analytics for XR
 - Multiple sensory input (eye tracking, HR, EEG, body movement, etc)
 - Use AI for data processing
- Integrate in with Unity/Unreal
 - Capture user performance during prototyping



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



www.empathiccomputing.org



mark.billinghurst@unisa.edu.au



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