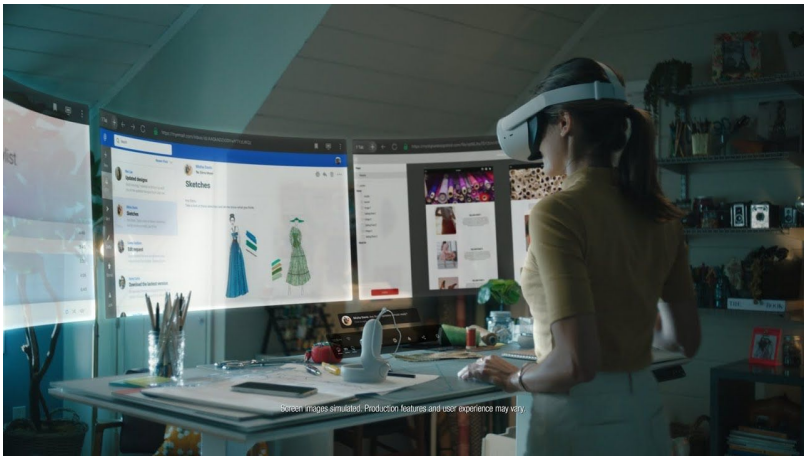
CMSC498F / CMSC731 / ENEE759N: Advances in XR

TuThu 12:30-1:45 cs.umd.edu/class/spring2026/cmsc731

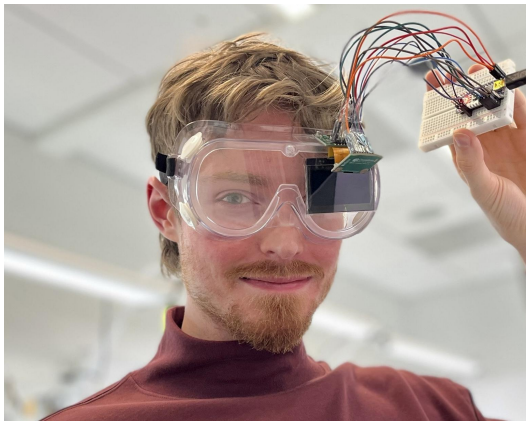


Ming C. Lin
http://www.cs.umd.edu/~lin/

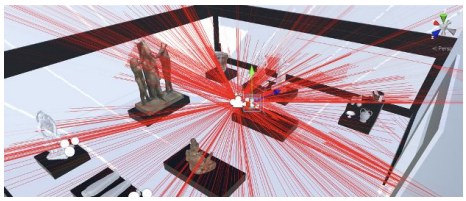
- BS, MS, PhD in Electrical Engineering & Computer Science
  University of California at Berkeley
- B. Mersky & CapitalOne E-Nnovate Endowed Professor 2021-
- UMD Distinguished University Professor, 2019-
- Former Elizabeth Iribe Chair of Computer Science @ UMD, 2018-2020
- J.R. & L.S. Parker Distinguished Prof. Emeritus @ UNC Chapel Hill
- ACM, IEEE & Eurographics Fellow; ACM SIGGRAPH Academy
- Areas of Research: **Virtual Reality, Robotics, AI/ML/Vision & Autonomy** with focuses on *physically-based modeling/simulation*, *multimodal interaction* (haptics & audio technology), *animation*, and *human-computer interaction*; applications in autonomous driving, virtual try-on, healthcare, digital design, rapid prototyping, and personalized fabrication/cybermanufacturing

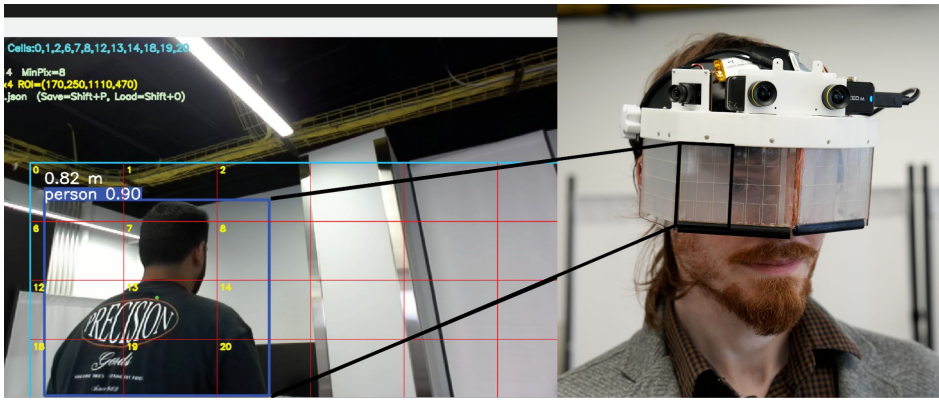
Logan Stevens

https://loganstevens.github.io/



- 2nd-year Ph.D. student in the Embodied Dynamics Laboratory (EmD Lab) advised by Dr. Jun Nishida.
- Undergrad from UMD: CS + Theatre
- Research interests: XR, HCI, psychophysics, cognitive enhancement, and Diminished Reality (DR).

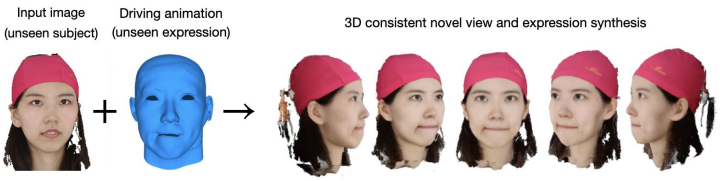


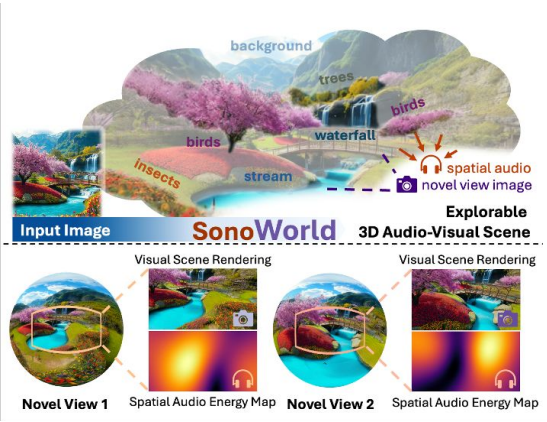


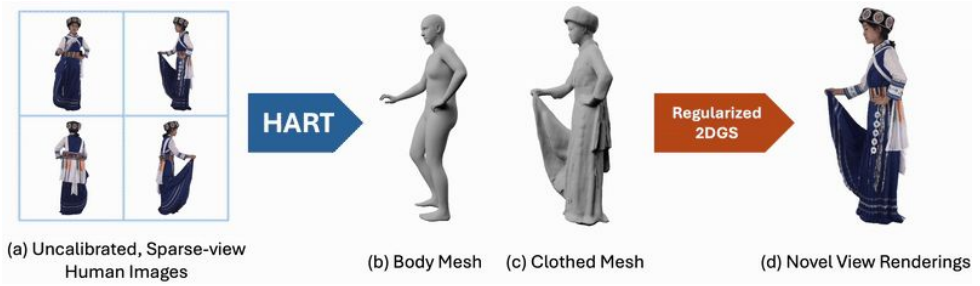
Xiyi Chen

https://xiyichen.github.io/

- 2nd year PhD student in the GAMMA lab
- BS from UMD, MS from ETH Zürich
- Research interests in digital humans, world/scene reconstruction

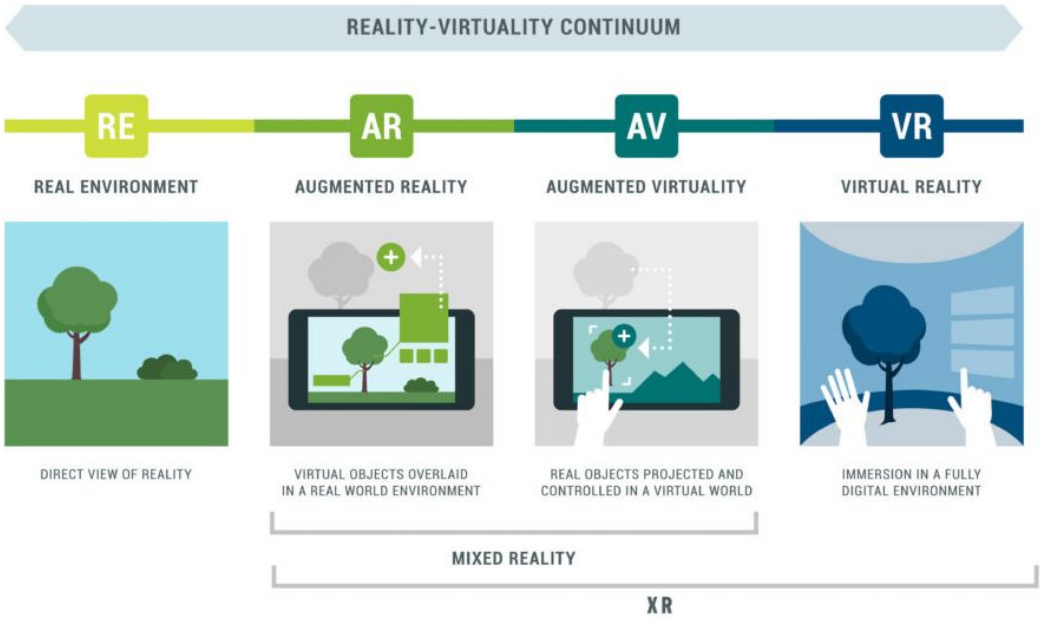






What is Extended Reality (XR)?

- Software & hardware that replaces or mixes real world stimuli with synthetic
- Different types of XR mostly differentiated by **display** and **tracking** methods



Sutherland 1965: The Ultimate Display

It (the Ultimate Display, referring to VR and AR) is a looking glass into a mathematical wonderland……. If the task of the display is to serve as a looking-glass into the mathematical wonderland constructed in computer memory, it should serve as many senses as possible. - Ivan Sutherland, 1965

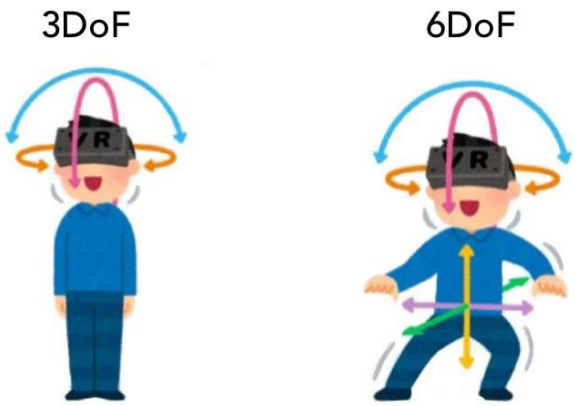
- A virtual world, through a HMD, appeared realistic through augmented 3D sound & tactile feedback
- Computer hardware to create the virtual word and maintain it in real time
- The ability users to interact with objects in the virtual world in a realistic way

Sutherland 1968: First XR headset

- “A head-mounted three dimensional display” Ivan Sutherland 1968
- Concept first introduced by a concept paper to DARPA in 1965

Degrees of Freedom (DoF)

- How many types of motion are tracked
  - Translational: x, y, z
  - Rotational: pitch, yaw, roll
- 3DoF in XR usually means rotational tracking only; 6DoF is rotation+translation
- Applies to display device (head-mounted display [HMD]/phone) and any other controllers



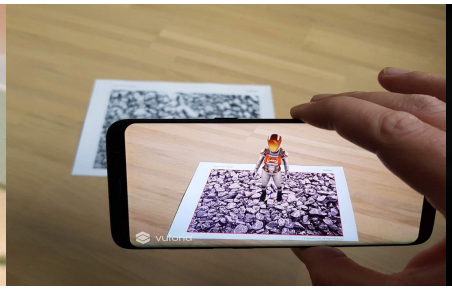
Virtual Reality (VR)

- Only handles virtual world and assumes real stimuli completely replaced with synthetic stimuli (e.g. Oculus Rift, HTC Vive, Google Cardboard)
  - If the synthetic stimuli not handled correctly, people can get simulator sickness, lack of “presence” or immersion, or worse (e.g. trauma)
- Main “illusions” needed for immersive experience (Mel Slater 2009):
  - **Place illusion**: feeling like you’re in the virtual world and not the real world
  - **Plausibility illusion**: feeling like what happens in the virtual world is really happening



Augmented Reality (AR)


- Overlays digital ‘elements’ onto real world, including graphics, images, video, sound, GIS data, text, animation, etc.
- Through head-mounted display, hologram, video-passthrough, etc.
- Focus on simple experiences or information easily available
- e.g. Google Glass, old Snap Spectacles, marker-based tracking–Nintendo 3DS, volumetric AR display


Mixed Reality (MR)

- Merger of real world and virtual world, co-existing & interacting in real time
- Focus on practicality, productivity, integration with day-to-day life
- e.g. tracking real world features & hands makes the **Oculus Quest** an MR headset

Diminished Reality (DR)

- Removing, obscuring, or diminishing real-world stimuli using XR hardware
- Could be used for searching, clutter management, distraction reduction, and focus
- Relatively recent variant of XR; does not fall onto the traditional continuum

Input environment | DR-enabled environment

(S2) Search

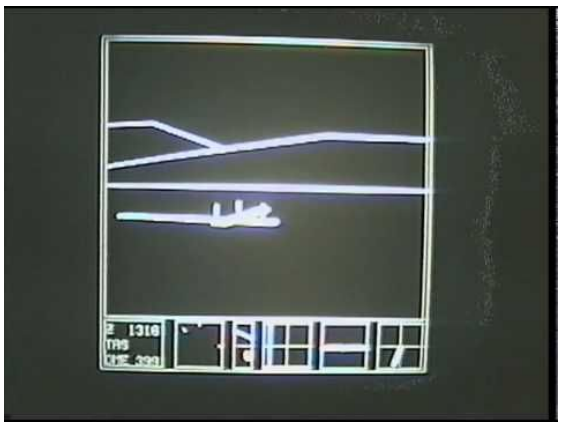
Input environment | DR-enabled environment

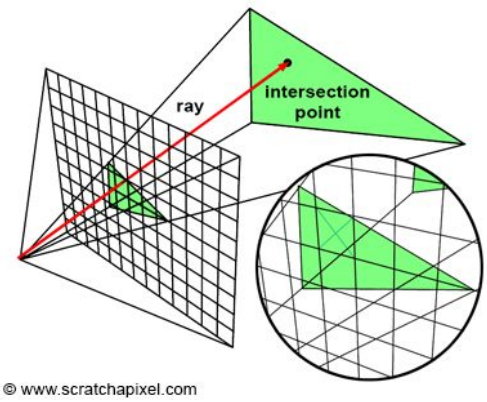
(S1) Clutter management

XR Trends Over Time

XR Trends Over Time: Pre-1970s

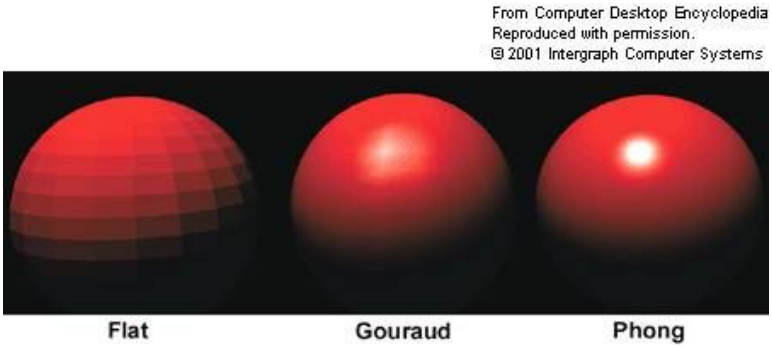
- Fundamental tracking, rendering, & display technology (esp. For simulators)
- First VR headsets (stereo headsets in the 1800s!)

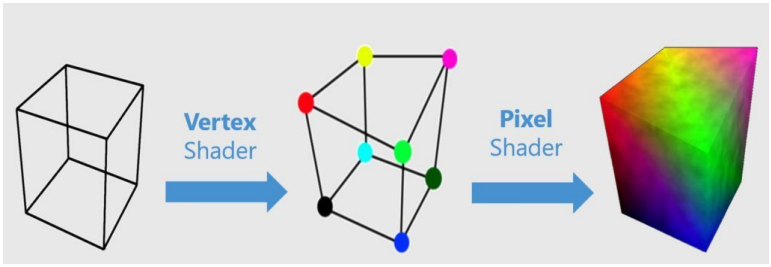


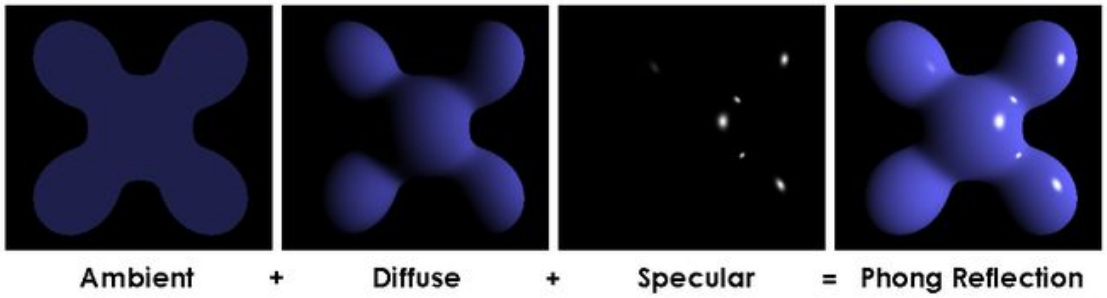


XR Trends Over Time: 1970s

- Fundamental 3D graphics technology (e.g. Phong shading)

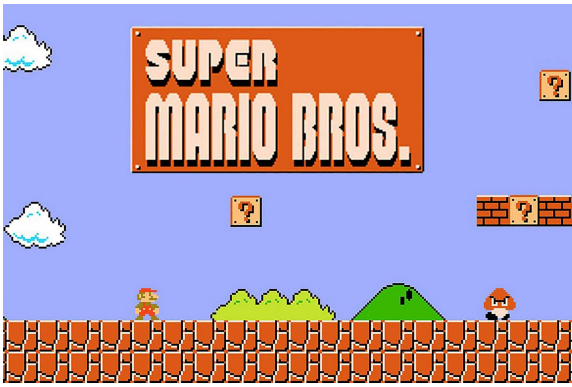


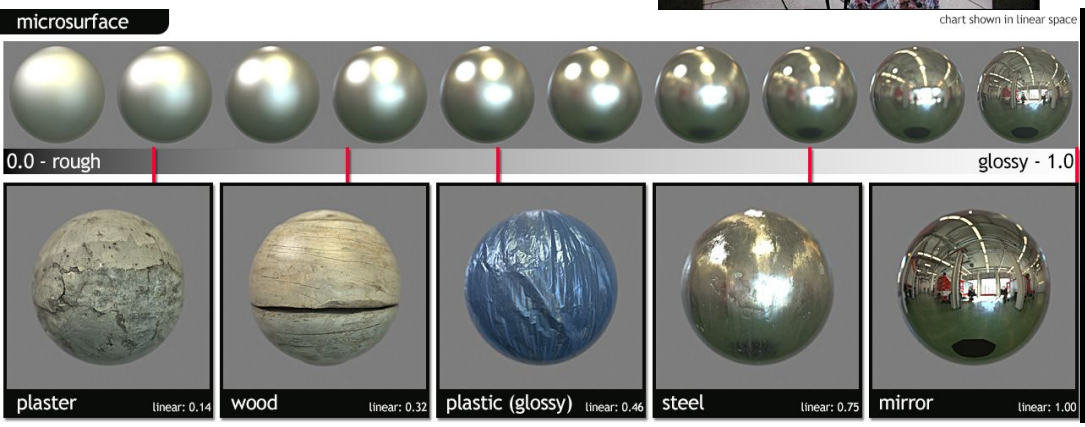




XR Trends Over Time: 1980s

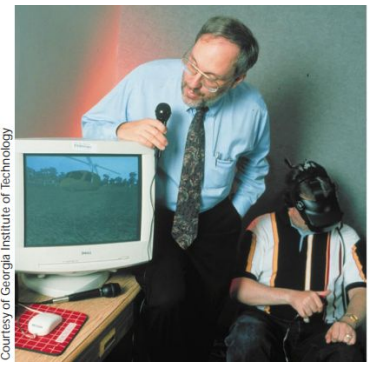
- Physically-Based Rendering (PBR)
- VR Simulators (esp. flight)
- Interactive Games
- Display adapters (old GPUs)
- Multimodal XR apps





XR Trends Over Time: 90s & Early 2000s

- Clumsy but effective XR headsets
- Boom of XR psychological studies
- 3D Game Engines
- Interactive 3D graphics
- Collision detection & interactive physics
- Boom in areas like haptics & locomotion



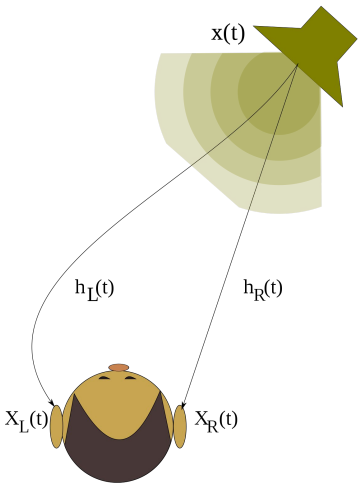
Courtesy of Georgia Institute of Technology

3 NASA-Houston's "Charlotte" Virtual Weightless Mass lets astronauts practice handling weightless massy objects.

Courtesy of NASA

XR Trends Over Time: Late 2000s

- Research on different modalities, audio, locomotion, rendering→all evolved close to what they are today
- Unity, Unreal 2/3, idEngine, Source Engine, Autodesk Maya, etc. established many interactive 3D graphics conventions

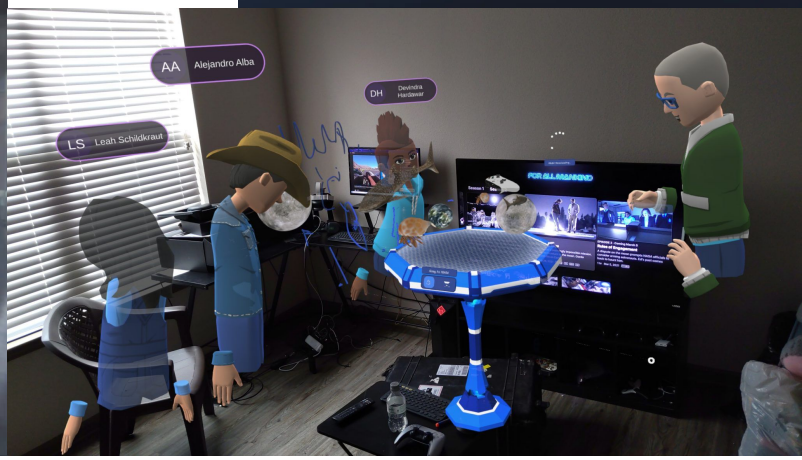


XR Trends Over Time: 2010s

- Good mobile technology
- Strong GPUs
- Decent commercial headsets
- Industry getting involved

XR Trends Over Time: Early 2020s

- Everyone is making XR headsets
- Metaverse hype
- Social XR



XR Trends Over Time: Predicting Late 2020s

- Focus on application development
- Procedurally-generated content
- Better, large-scale HRI
- Virtual assistants
- Attempts at neural interfaces

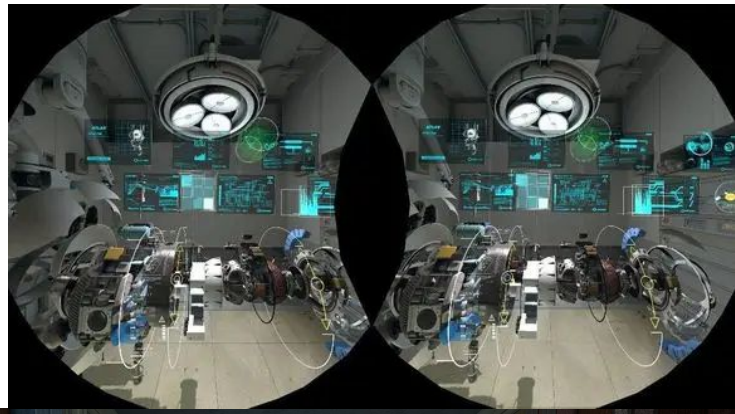
Challenges in XR

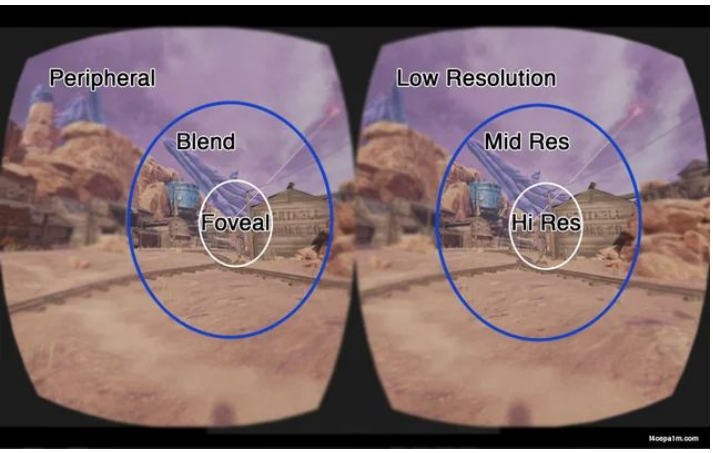
Challenges in XR: HMD Design

- Weight
- Physical dimensions & portability
- Field of view (FOV)
- Battery vs. computing power
- Pixel opacity
- Optics

Challenges in XR: Graphics

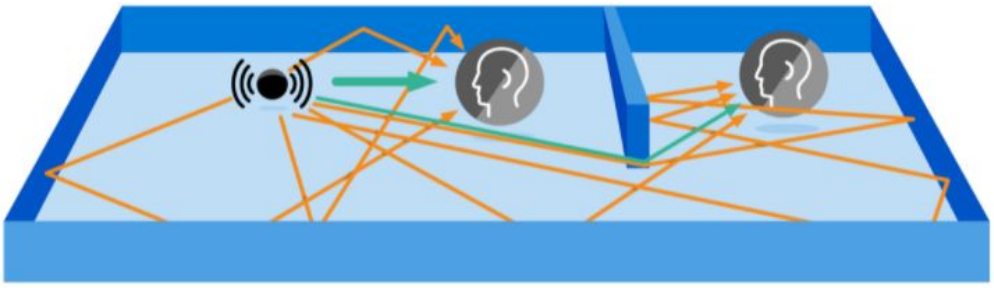
- Realistic materials
- Foveated rendering & optimization
- Physically-based interactions

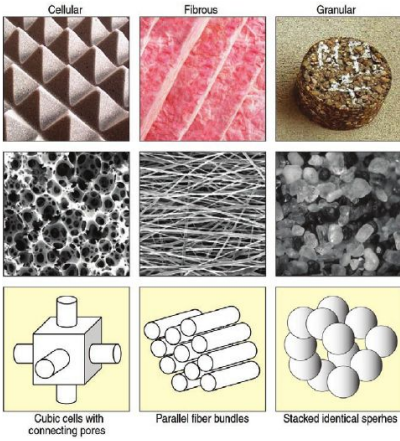


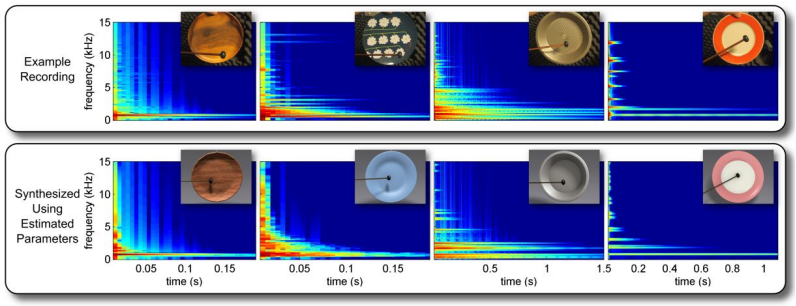


Challenges in XR: Audio

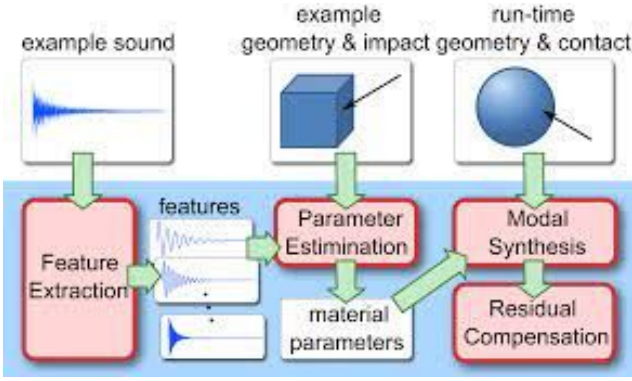
- Procedural generation vs. sound synthesis (esp. on mobile)
- Sound propagation & 3D acoustics (esp. on mobile)
- Personalized 3D audio display





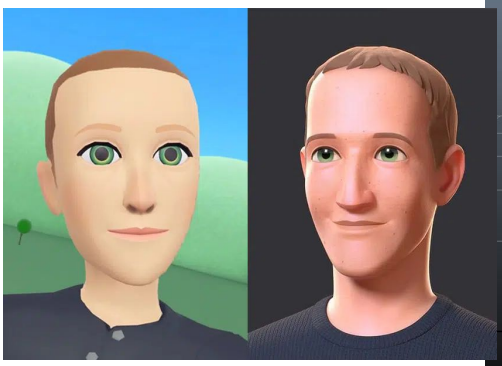


(a) *wood plate* (b) *plastic plate* (c) *metal plate* (d) *porcelain plate*



Challenges in XR: Virtual Humans

- Uncanny valley
- AI behavioral modelling
- Social cues
- Procedural animation & rigging
- Generative AI models
- UV mapping/texturing

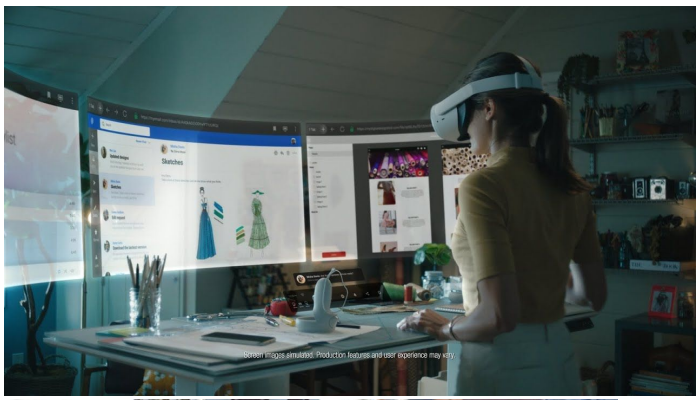


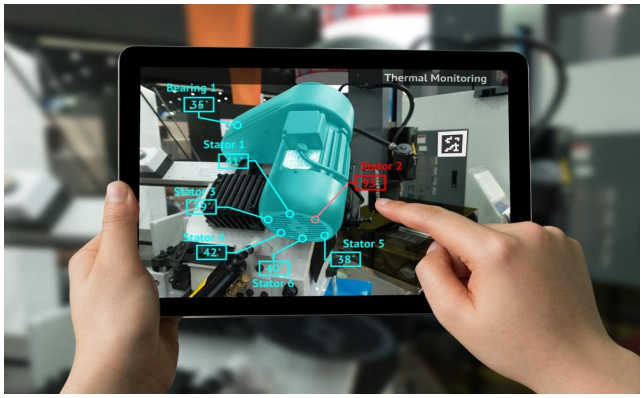
Challenges in XR: Content

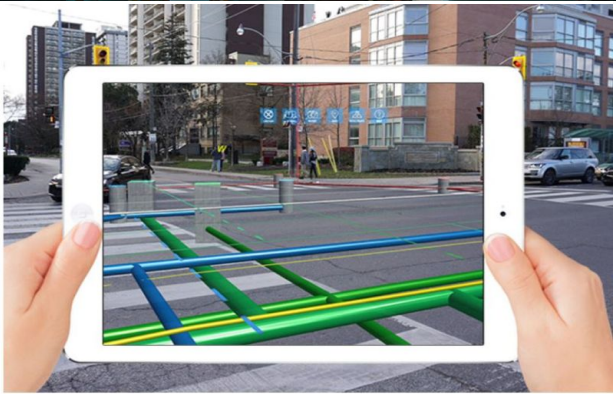
- Procedural generation
  - Doesn't only apply to environments! Used to speed up getting good textures, models, animation, audio, etc.
- Generative AI models
- User-generated vs. company-generated

Challenges in XR: Application Design & Utility

- How to keep people using it?
- How to reduce barrier of entry?
- How to integrate with people's lives?

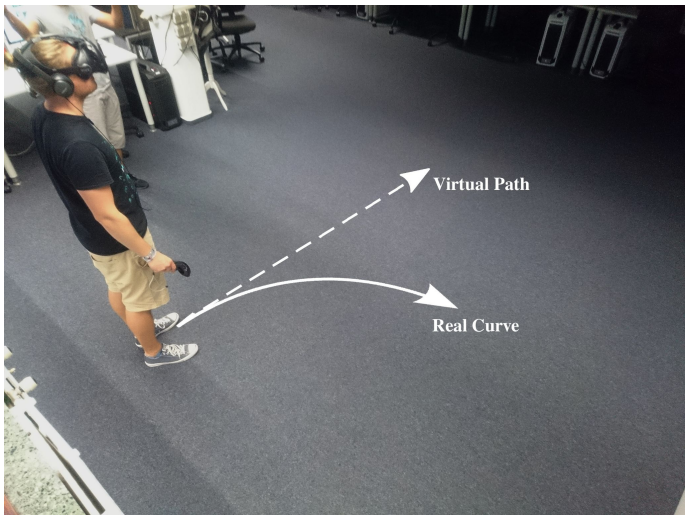






Challenges in XR: Natural Locomotion in VE

- Getting people to walk around naturally with limited space
- Handling sickness & perception (knowledge of the “illusion” changes its effect!)

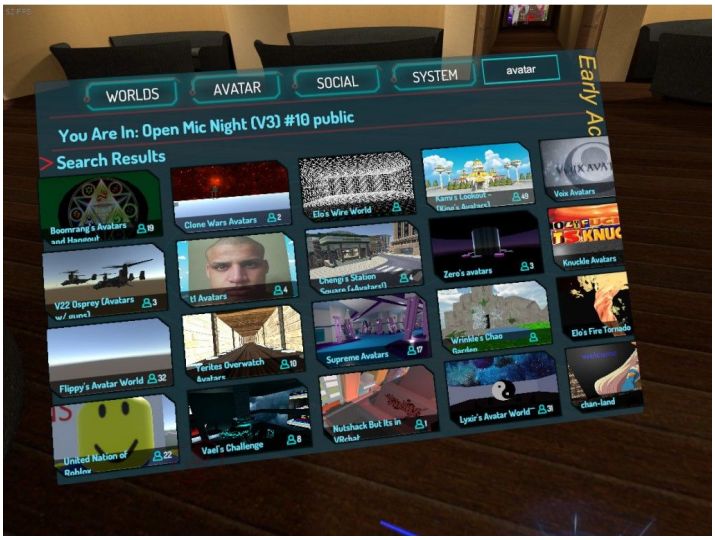


Challenges in XR: Tracking & Reconstruction

- **Local** tracking of headset more or less solved
- **Global** tracking still challenging
  - How to register virtual environments to real ones?
  - How to link the environments together?
  - How to overlay real and virtual worlds?
- Reconstructing people, real environments, etc. to make virtual world as convincing as real world
  - Photogrammetry is the state-of-the-art for high-fidelity asset creation but it’s incredibly slow & tedious

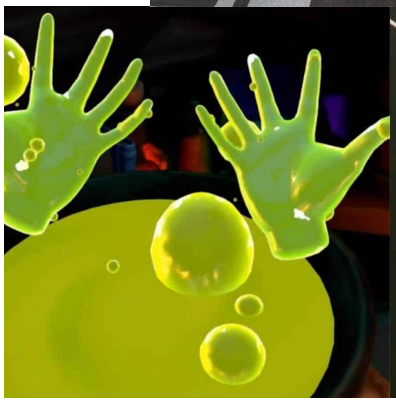
Challenges in XR: SW/HW Interoperability & Compatibility

- Huge issue with Metaverse
- How can assets trivially move between XR applications as if they are part of a unified metaverse?



Challenges in XR: Natural Interactions with Full-Body Tracking

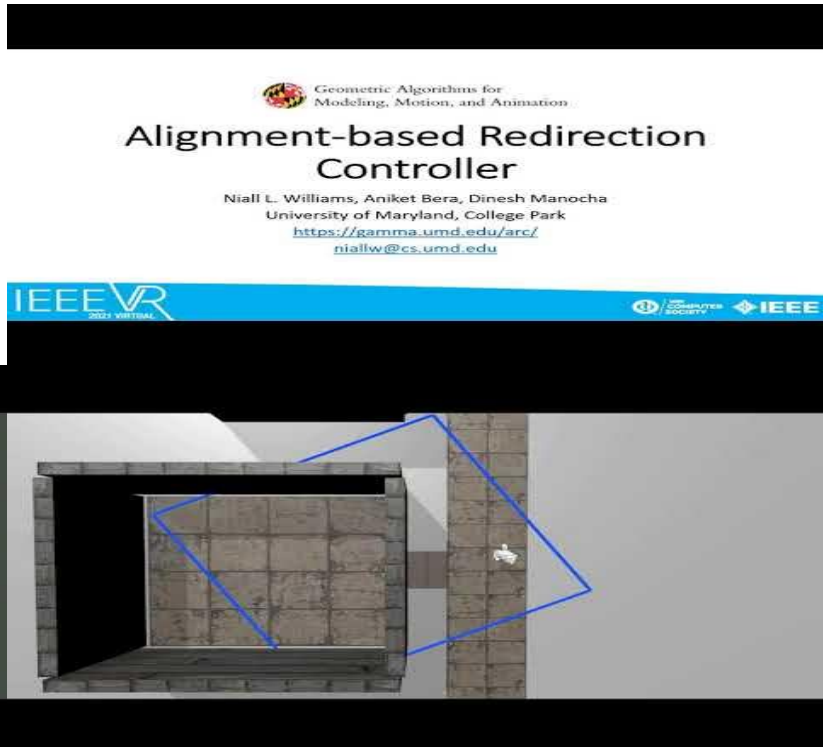
- Eyes, hands, etc.
- Predict user intent
- HMD's FOV of hands
- Gesture recognition
- Accessibility

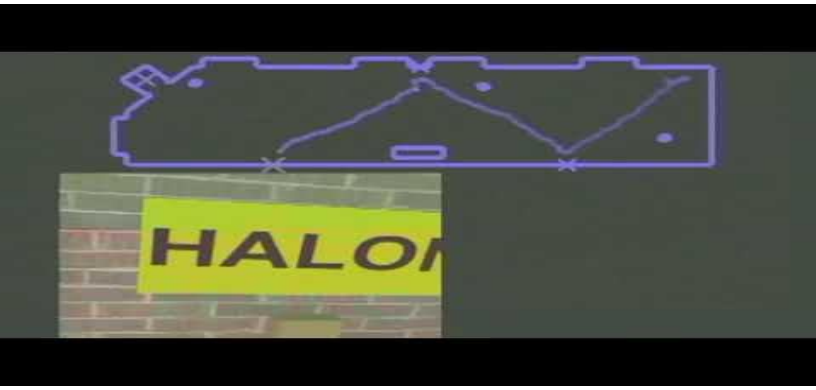


Current Research Directions

Research Directions: Natural Virtual Locomotion

- Use distortions to trick people into walking differently than they realize.
- Maximize real walkable space.
- Must support a range of users and environment shapes.



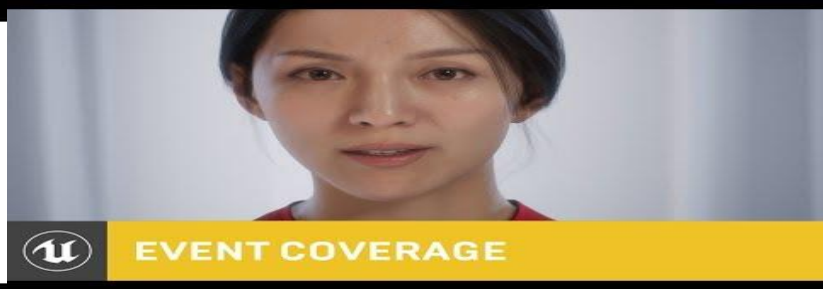


Research Directions: Intelligent Virtual Environments

- Future of work
- Context-aware interaction
- Integration with daily life

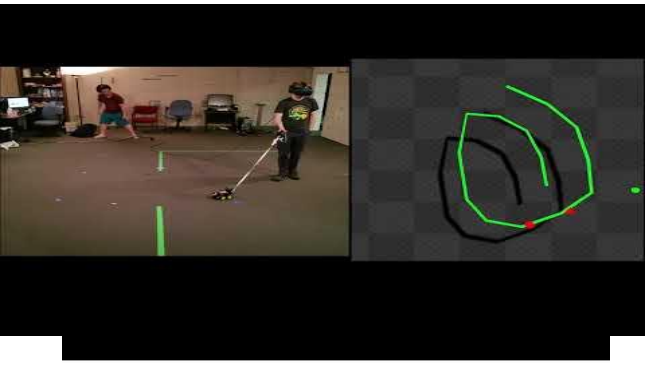
Research Directions: Emotive Virtual Humans

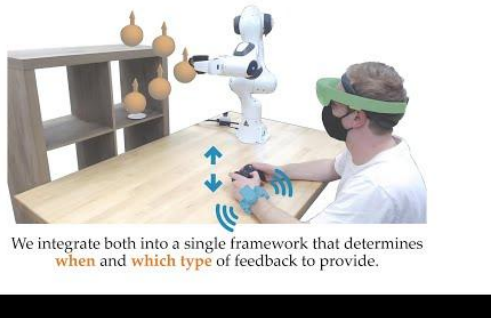
- Realistic virtual assistants w/ procedural animation
- Passing the uncanny valley



Research Directions: Human-Robot Interaction in XR

- Robot dogs to train the visually-impaired spatial mapping skills
- Robots in warehouses for collaborative XR





Research Directions: Metaverse

- (Seamlessly-Integrated Real & Virtual Environments)
- **Human-centric mapping between spaces**: pocket dimensions/ portals/ symbolic links, human-centric motion planning, context, behavior prediction
- **“Internet of XR”/ “IoX”**: sensor/device fusion, connected HMDs, maximizing knowledge of real world

Research Directions: Intelligent, Responsive XR

- Immersive, XR Interfaces for autonomous systems
  - Driverless vehicles
  - Autonomous drones
  - Intelligent sensor networks
- XR Systems for rapid design and personalized fabrication/manufacturing
- XR Systems for remote collaboration, tourisms, social events
- XR Systems for personalized healthcare and wellness
- XR Interface for virtual try-on
- XR Systems for social good
  - Immersive scenario replay for Police Training, Bias Training, AI for Fairness, etc.
- Audio-Visual Reconstruction

Course Information

Course Goals

- Understand **multimodal XR design**
  - Basic principles of audio, haptic, and visual rendering
  - How multiple modalities work together and interact with each other
  - Challenges of Metaverse & future of XR
- Understand “realism”, “illusion”, and “presence”
  - Roles of locomotion
  - Sense of ‘being there’
  - Avatar & object animation/simulation
  - Social interaction via virtual humans
- Application design & development
- Use the state-of-the-art APIs & XR tools

Course Components

- Lectures & Participation
- Homework Assignments [30%]
- Final Project [35%]
- Midterm & Final exams [35%]

Lectures

- **Some may be recorded, but not all**
- With emerging tech, the ***discussion is very beneficial***
- Rough sequence:
  - Basic 3D graphics
  - Basic game development
  - Immersion & presence
  - Virtual Locomotion
  - 3D Audio
  - 3D Animation
  - Virtual agents & AI
  - Tracking & Reconstruction
  - Displays, Optics, Lenses, etc.
  - Advanced Topics (light fields, haptics, olfactory, GANs, holography, etc.)
- Invited speakers from industry & academia

Game Engines

- **Game engines**: powerful, realtime, interactive, multimodal 3D applications
- De facto standard for XR development and most interactive 3D consumer programs
- We will provide support for Unity development, but not for Unreal Engine.
  - Use Unreal at your own risk

Headset Logistics

- The plan:
  - Oculus Quest rental system
  - We need to share!
  - Use your own device, if you own one
- Details being worked out, more info to come
- Anyone using personal devices?

Assignments

- Usually, a technical, programming part and some kind of small HCI part (analysis, mini-user study, etc.)
- **Shouldn't be that hard** in terms of programming, but practice skills you probably haven't used before (e.g. design, ANOVA analysis)
- Assignments are **mostly disjoint** extensions of the project you start in the beginning (so Unity setup assignment is very important)
- **Advice**: if something breaks and doesn't seem programming related, post about it on Piazza. Unity is buggy and has a lot of errors with easy solutions (e.g. delete cache). Unreal similarly has annoying linker errors (common in C++)

Assignments

- Address the major modalities & concepts of XR design
  - Goal is **breadth** of experience; final project is opportunity to dive deep
- Build on each other to result in complete multimodal XR application
- Planned sequence
  - **A1:** Introduction to XR and game engines
  - **A2:** Interaction with virtual environments
  - **A3:** 3D audio for virtual environments
  - **A4:** Natural virtual locomotion
  - **A5:** Virtual avatars and agents
  - **A6:** Introduction to augmented reality
- May include non-technical readings
- Grad section may have slightly ***extra requirements & higher expectations***
- Submission includes code/Git repo & video showing it works

Introducing Assignment 1

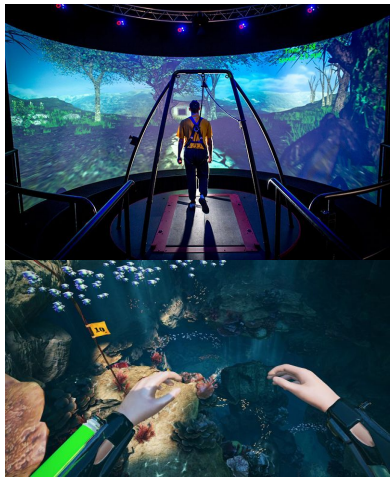
- Introduction to the Unity game engine
- 2 fundamental readings:
  - **“What’s Real About VR?”** Fred Brooks 1999
  - **“Recent advances in augmented reality”** Ron Azuma 2001
  - Will answer some short-answer questions about them

Final Project

- Will draw upon the concepts you learned in the course
- Planning a competition with prize winner(s)
- Details to be announced later

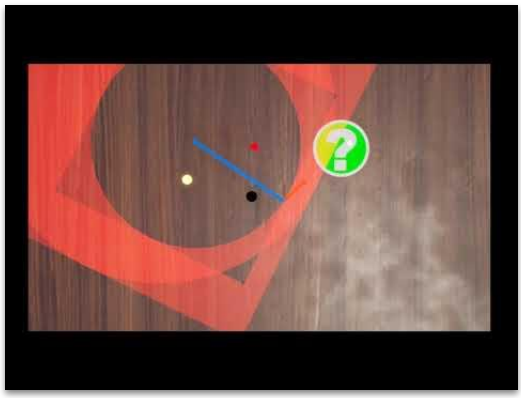
Virtual Reality

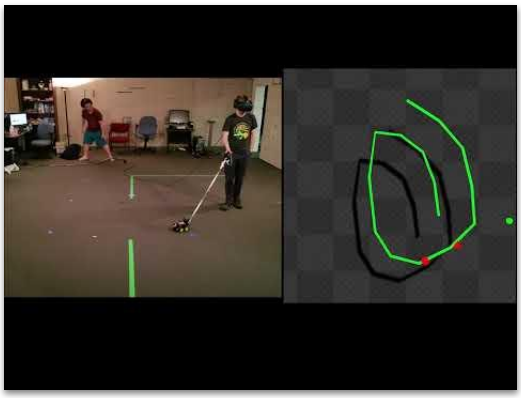
<-misleading!

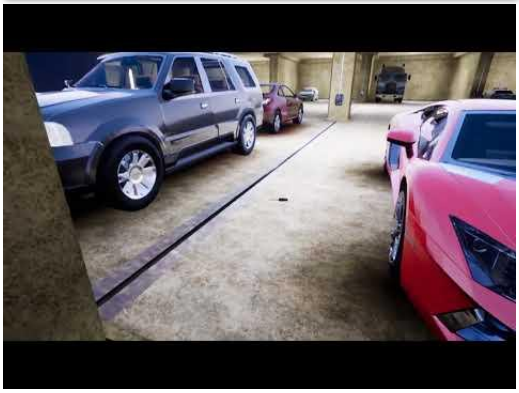


- As an idea:
  - (Ideally, completely) substitutes real stimuli with synthetic
    - E.g. real visuals replaced by 3D CG visuals
    - E.g. real audio replaced by 3D synthesized audio
    - As opposed to AR & MR, which blend real and CG
  - Attempts to evoke a sense of presence
    - A feeling of “being there,” in the 3D world **instead of** the real world (Minsky 1980)
- As a technology (most consumer devices..not all!)
  - **Stereoscopic Near-eye displays, engineering magic** to address vergence-accommodation conflict
  - Place constraints in the virtual world **based on** constraints in the real world, like walkable space or physical obstacles
  - **Track objects** like the HMD, controllers, etc. with (usually) 6 degrees of freedom (DOF) (**Location**: (x,y,z); **Rotation (aka Orientation)**: (pitch, yaw, roll))
  - Sense of presence is hard to perfect and some senses are difficult (if not impossible) to reproduce synthetically atm, so do the best we can

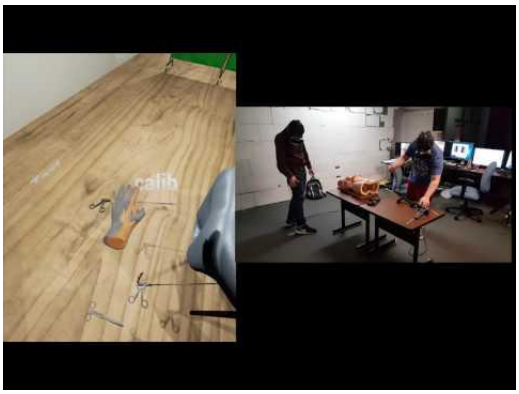
GAMMA/Oculus 3D audio, synthesized audio, AI, perception, VR

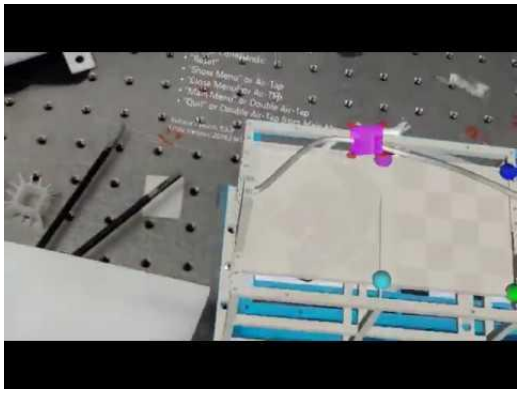


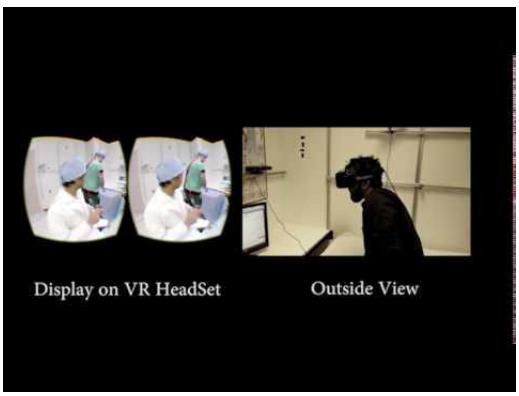




Telepresence 3D reconstruction, multi-user systems, AR, surgical apps

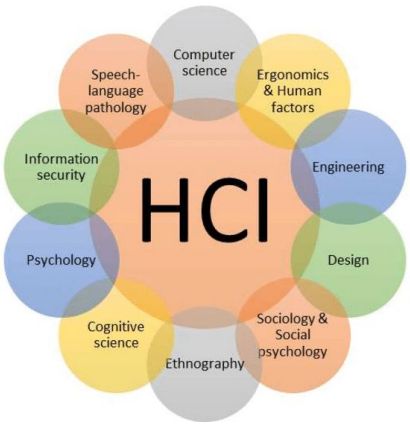


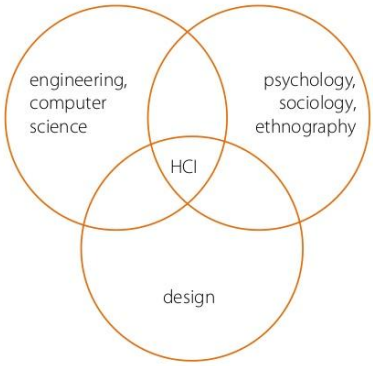




What is human-computer interaction (HCI)?

Class question



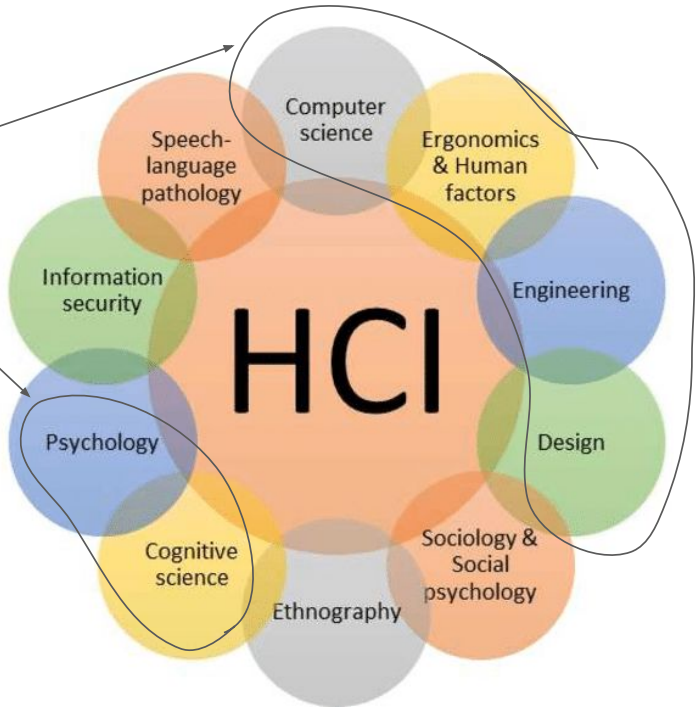


Human-computer interaction (HCI)

- Study of interaction between computers and their human users
- Similar, if not a subset of design thinking
  - From Interaction Design Foundation: “non-linear, iterative process which seeks to understand users, challenge assumptions, redefine problems and create innovative solutions to prototype and test”
- Create novel applications with computers to solve **human problems**
- Understand how and why computer apps work, iteratively improve them
- Focus on the human factor instead of **solely** optimization or design.
- If your app doesn’t work for its target users, then it **doesn’t work**. Even if it’s O(1)
- HCI ≠ UX. UX is often more business-focused (e.g. how to keep your users coming back. The FEEL of the app rather than applicability. How does Candy Crush get people to play 100s of levels? Or FB to get users to keep scrolling? Who cares why?)

HCI vs UX

VR HCI focus (relative to other HCI fields)



UX's focus

Some problems in XR addressed by HCI

- Sickness
  - We now know to **never use game controller input** to move!
  - "Scary" things in regular games could actually give users in VR heart attacks or trauma
- Applicability
  - Companies are learning to provide **multiple** locomotion **options**, graphical tweaks, etc.
- Novelty
  - Things that are usually boring in regular games can be more **addicting** in VR (e.g. rail shooters)
- Important questions to ask yourself in game and VR dev:
  - How do I know this works?
  - How do I go about testing practicality, sickness, user response, etc?
  - How do I improve this for the user?

Collaboration & General Policies

- 2-3 for each regular assignment
- 1-5 people in a group for final project
  - Final report must describe what each team member does
  - Each will be graded based on individual contributions specified
  - Group efforts should show higher levels of complexity and components
- Recommendation: try to stay with the same OS & platform combination
- ***No collaboration on in-class exams***
- Lateness
  - Total of 5 “late credits” (1 credit = extend deadline by 24 hours)
  - Linear late penalty, up to 3 days
  - Late assignments due to illness or unexpected events can be excused with doctor’s notes or other forms of written indication

Collaboration

- Assignments are not individually complex enough to warrant several partners. Groups of 2-3 max for assignments. Can discuss **high-level** concepts with others and on forum.
- Writeup can be done individually or with your partner/group.
- Github project link **must** be submitted with assignment, as well as Overleaf **EDIT** link if you did the writeup with partners
  - Overleaf has edit history that is only visible to editors
  - It will be more effort to cheat by faking the Overleaf edit history than to just do the assignment
  - Why: you should be doing both parts of assignments collaboratively, not splitting it
- Final project: Groups of 5 max. Bigger group=bigger expectations & better writeup.
- Any collaboration beyond this is an honor code violation

Exams

- Midterm
- Final
- Relatively high-level design questions
  - E.g. How would you design an XR app for a particular application? What are the important considerations?
  - E.g. Someone wants to make an XR app a certain way; what problems are they not addressing correctly?
  - E.g. Why do we need to do [a particular thing] for XR?
  - E.g. How did [this paper] likely accomplish [this feature]?
- No programming, but will ask high-level questions about technical problems
  - E.g. why are VR lenses designed the way they are
  - E.g. what do waveguides and lightboxes do in near-eye AR displays?

Office Hours

Instructor *Immediately after class or by appointment*

Graduate Teaching Assistants

Logan Stevens *TuThu 2-3pm or by request (E-Mail: lsteven7@umd.edu)*

Xiyi Chen *MW 3-4pm or by request (E-Mail: xiyichen@umd.edu)*

Undergraduate Teaching Assistants

Saunak Roy (E-Mail: saunakr@terpmail.umd.edu)

Suchismit Ghosh (E-Mail: sitghosh@terpmail.umd.edu)

Lab @ AVW 4120 Mon-Thur 3:30pm-6pm; Fri 12:30pm -2:30pm