PART IV: Additional Topics

Interaction Devices

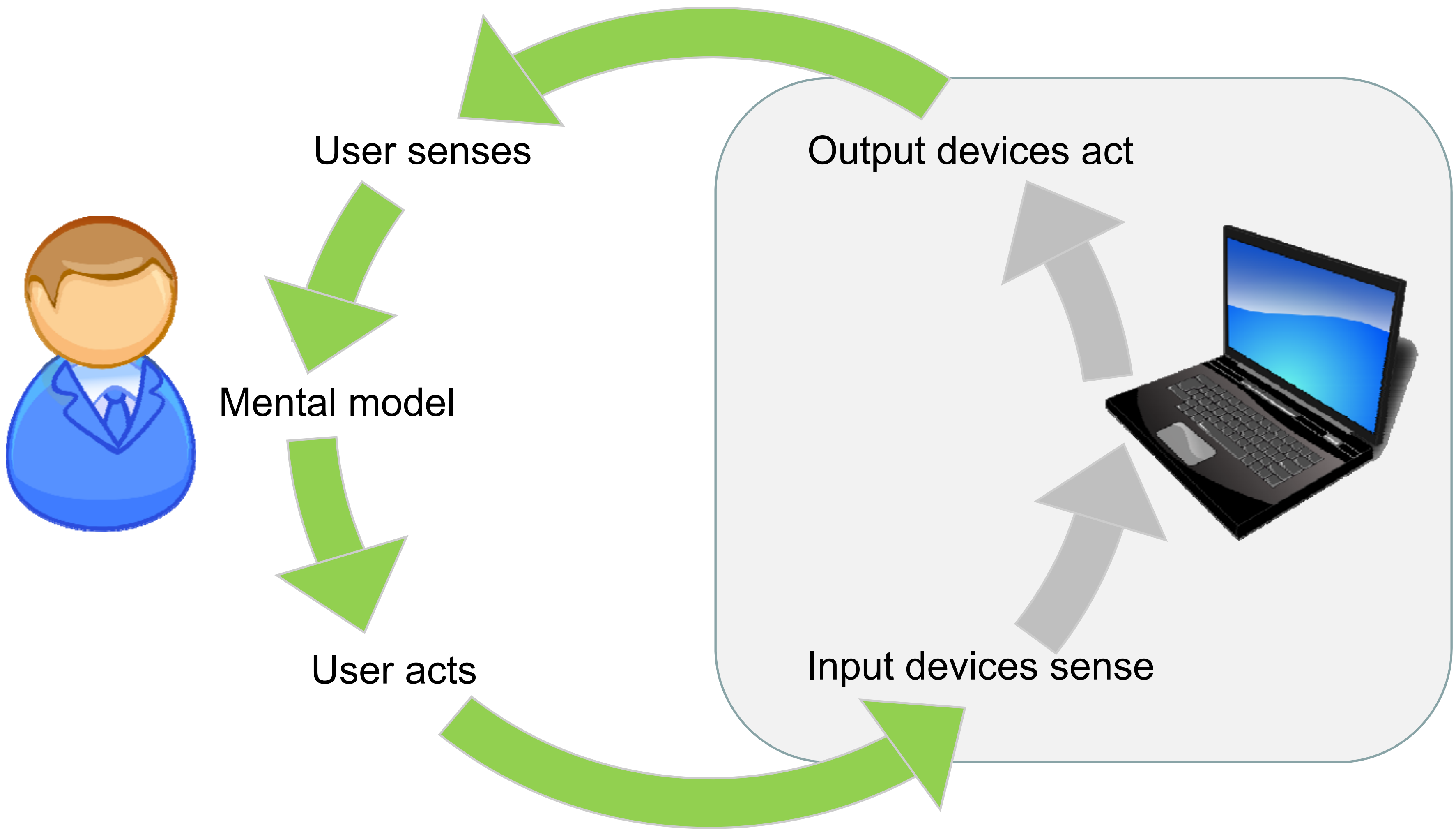
SEEM3510 Human-Computer Interaction

By Prof. Helen MENG & Prof. Philip FU

Course Outline

Part 1: Basics	Week 1 - Usability of Interactive Systems Week 2 - Universal Usability Week 3 - Guidelines, Principle and Theories	By Helen
Part 2: Development Process	Week 4 - Design Process (Assignment 2) Week 5 - Evaluation (Assignment 3)	By Philip
Part 3: Interaction Paradigms	Week 6 - Direct Manipulation	By Philip
	Week 7 Midterm Review and Exam (weeks 1-6 material)	
	Week 8 - Fluid Navigation: Menu and Form Fill-in	By Philip
	Week 9 Expressive Human and Command Languages	By Helen
Part 4: Additional Topics	Week 10 Interaction Devices Week 11 Communication and Collaboration	By Philip
	Week 12 Timely User Experience Week 13 Information Search and Data Visualization	By Helen

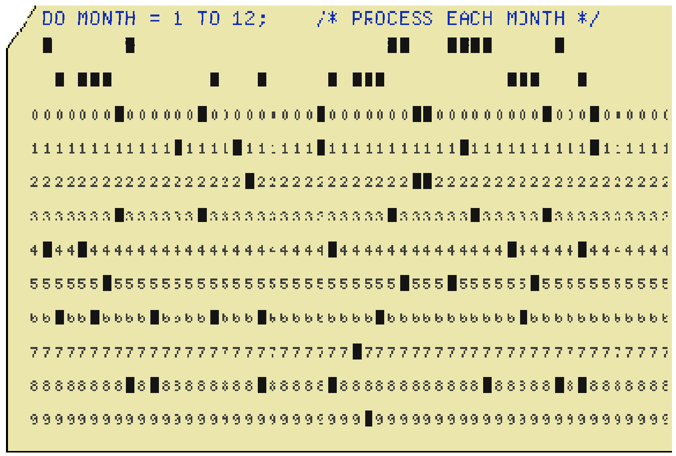
The Interaction Cycle

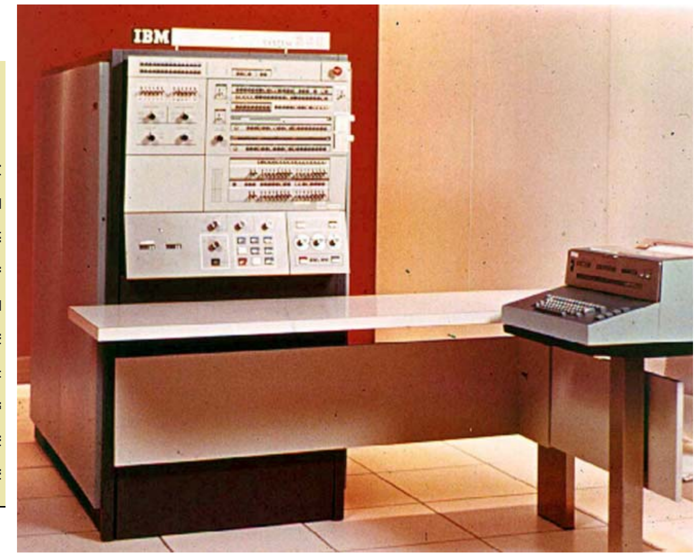


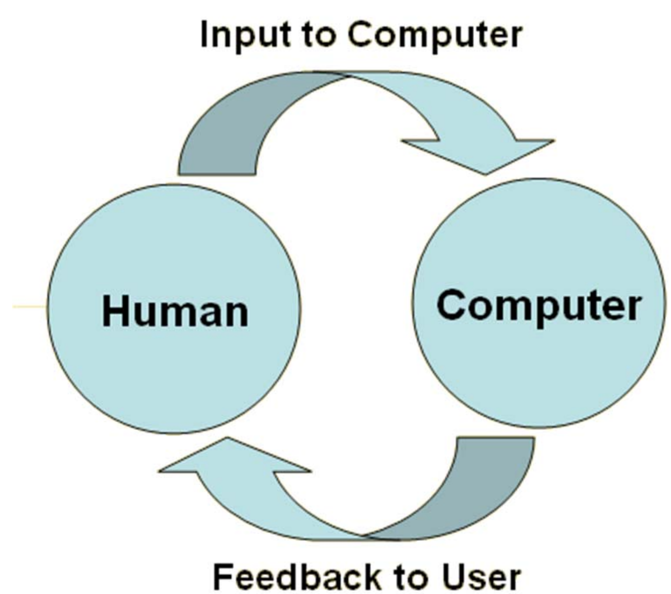
Interaction Performance

- 1960s vs. Today
  - Performance
    - Hz -> GHz
  - Memory
    - kb -> GB
  - Storage
    - kb -> TB
  - Input
    - Punch cards
    - Keyboards, multitouch screen, tablets, mobile phones, mouse, cameras, etc.
  - Output
    - 10 character/sec

    -> Retina displays, HMD VR, 3D display, surround sound, force feedback, etc.
- Substantial bandwidth increase!





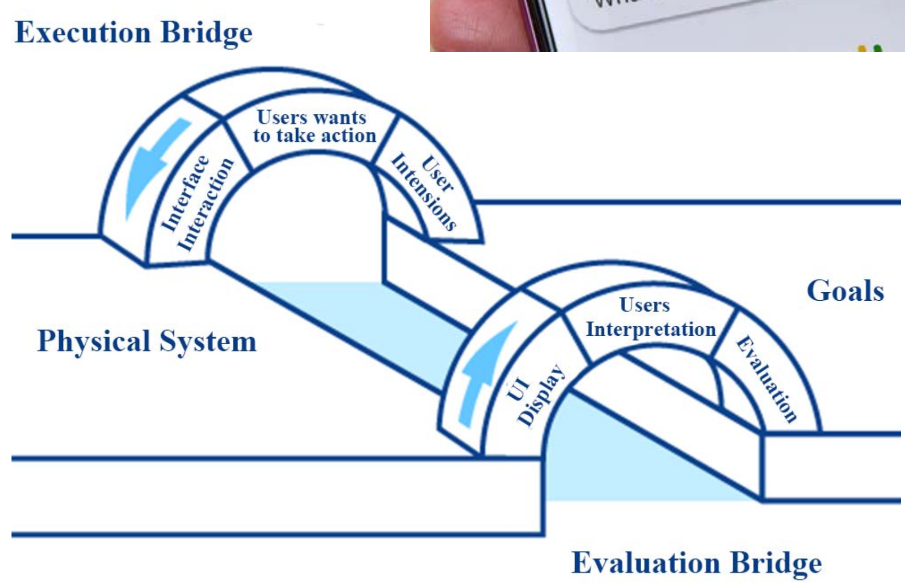


<u>IBM System/360</u>: 1964,
64 KB Memory, 75KHz, 8-bit per byte, etc.
http://www.youtube.com/watch?v=1OdRHImbPug
http://en.wikipedia.org/wiki/IBM_System/360

Interaction Performance

- Near Future?
  - AI & Machine Learning
  - Multi-modal: voice, wearable, whole body, eye trackers, hand gestures, etc.
  - Tangible (and Physical) -> TUI
  - Natural -> NUI & NLP
  - AI Assistant on “IoT”
  - Brain-Computer Interaction
  - Human-Robot Interaction

Aims to reduce the gulfs



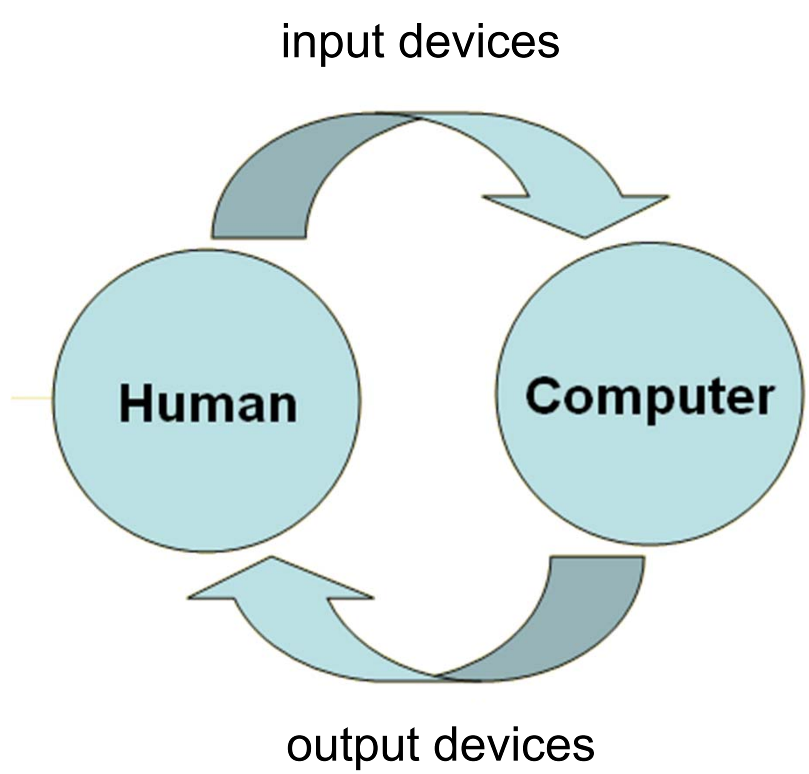
- Current trend:
  - Wearable technologies
  - VR and AR (XR)
  - Smart home / smart city
  - IoT (Internet of Things)
  - AI assistant & AI-enabled interaction

Consumer Electronic Show (CES)

https://www.ces.tech/

“Interaction Devices”

- Goals
  - Introduce to you various types of interaction (IN/OUT) devices
  - Introduce “Fitts’s Law” for motion modeling
- Topics
  - Introduction: past, current, and future
  - Input devices
  - Output devices
  - Other devices
  - AI-enabled Interactions



Topics:

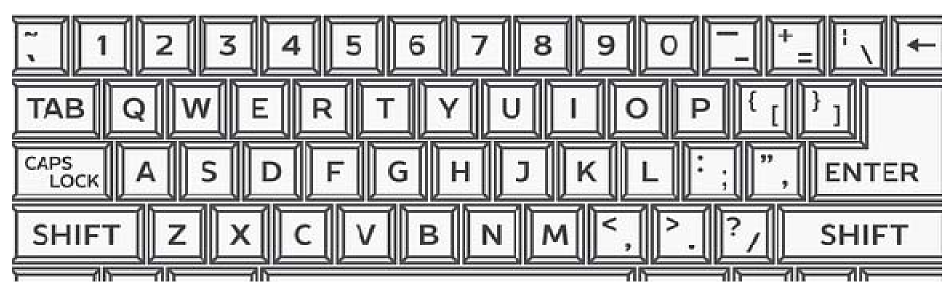
- Introduction: past, current, and future
- Input devices
  - Text entry: keyboards and keypads, etc.
  - Pointing: terminologies, devices, and Fitts's Law
  - Natural User Interface (NUI)
  - Other Input Devices
- Output devices
- Other devices
- AI-enabled interactions

How can we enter text?

#1 Keyboard

- QWERTY keyboards been around for a long time
  - **Cons:** Not easy to learn
  - **Pros:** Familiarity
  - Stats:
    - Beginners:
      1 keystroke per sec
    - Average office worker:
      5 keystrokes per sec
      (50 wpm)
    - Experts:
      15 keystrokes per sec
      (150 wpm)

- Is it possible to do better? Suggestions?



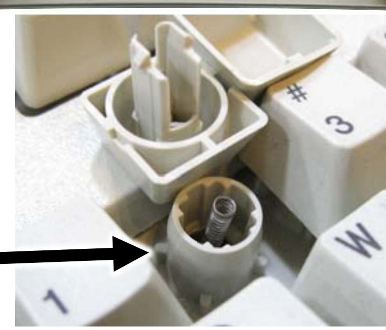
Why QWERTY?

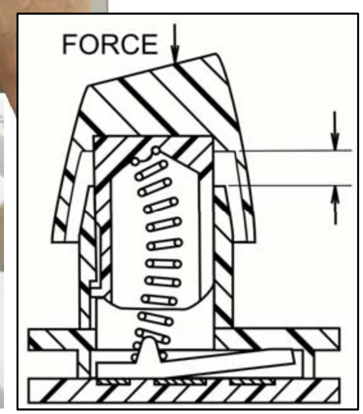
#1 Keyboard

How to do better?

- Dvorak layout
  - QWERTY (1870s): prevent jams in typewriters; this layout was carried over to electronic keyboard
  - DVORAK (1920s): another layout with better performance highest frequency keys in the middle row, etc.
- Ergonomic keyboards
  - Reshaped to reduce stress on wrists and fingers
- Tactile feedback
  - "Click" sensation when pressing a key (buckling springs)
  - Marks on some keys: "F" and "J" (eyes-free usage)

http://www.youtube.com/watch?v=FkUXn5bOwzk





http://www.mechtype.com/wp-content/uploads/2016/07/buckling-spring-working-switch.gif (good for **eyes-free interaction**)

#2 Keypad

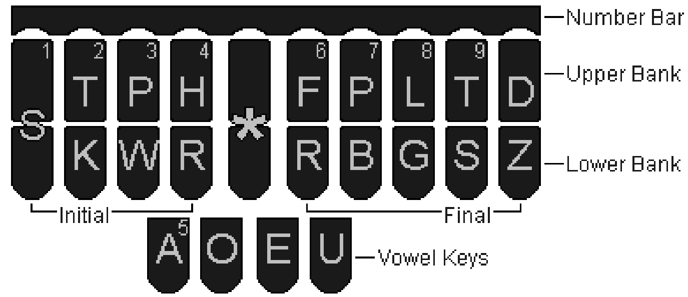
- Numeric keypads are very common
- Two common layouts!!!
  - 123 on top: phone, ATM…
  - 123 on bottom: calculator, keyboard-side keypad…
- Letter mapping was originally for phonewords as memory aid

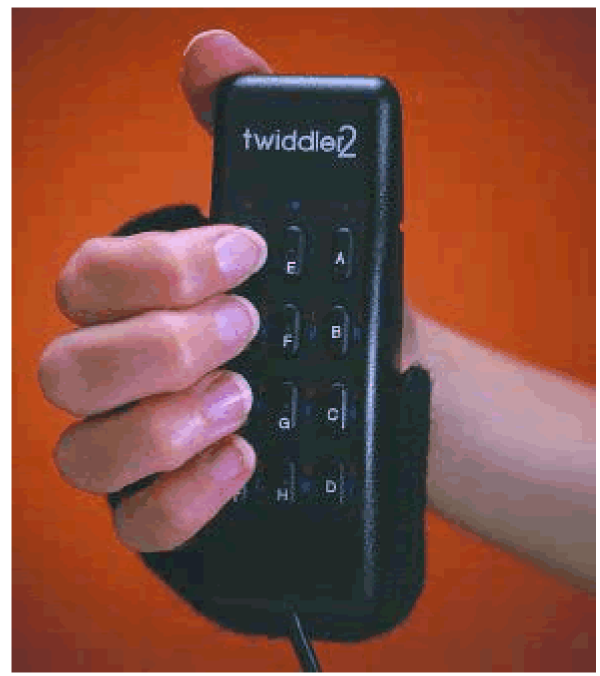
  e.g. 1800-FLOWERS

There are many theories or explanation behind this: http://www.vcalc.net/Keyboard.htm

#3 Chording Devices

- Chording == Pressing multiple keys simultaneously
- Stenotype (user = "Stenographer")
  - Very fast typing, e.g. for transcription in law courts: speed record 375 wpm!
  - Phonetic code (or shorthand system)
    - Writes syllables instead of normal spelling
    - 1 chord = 1 syllable
  - Processor converts syllable sequences into text: http://www.youtube.com/watch?v=6t7-IarTESc
- Chorded Keysets
  - For fast single-handed text entry (50 wpm)
  - 1 chord = 1 character
  - e.g. Twiddler



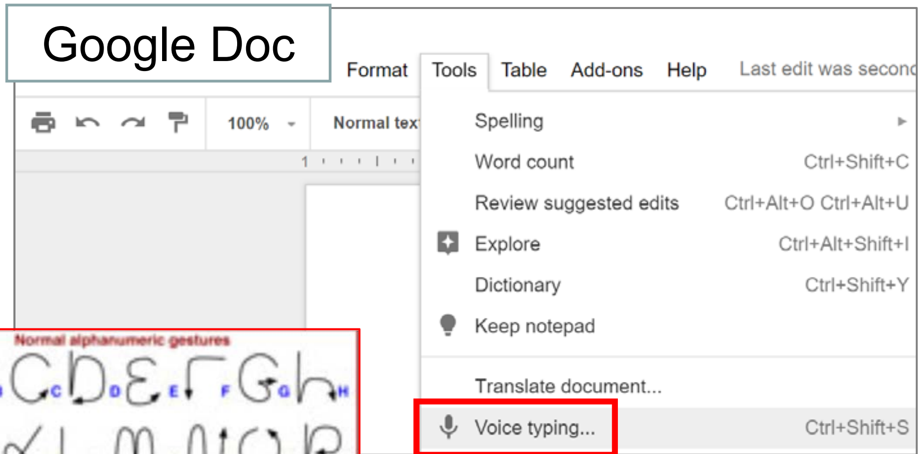


#4 Other Text Input Methods

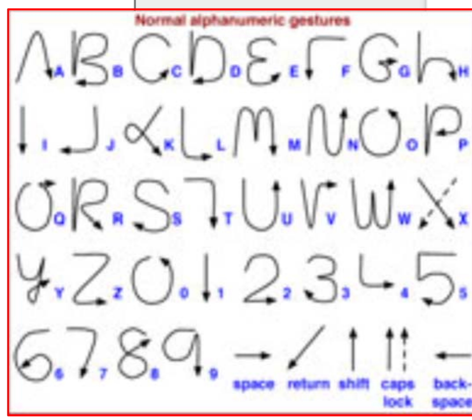
- Speech-to-Text (STT)
  - Continuous Speech Recognition

    https://docs.google.com

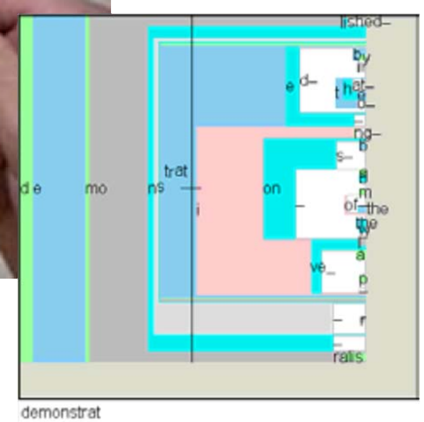
Google Doc



- Handwriting
  - Single letter recognition
    - stylus-based
    - Graffiti system
  - Handwriting recognition
    - Work for Chinese, etc.
    - Also for math equation



- Chording on Multi-Touch Devices
- Dasher software
  - Fast text entry by navigating
  - Steering a pointer
  - Demo: http://www.youtube.com/watch?v=0d6yIquOKQ0



Word Input Speeds

- Words per minute (WPM)
  http://en.wikipedia.org/wiki/Words_per_minute
- Keyboard speed tests online:
  http://speedtest.aoeu.nl/

World records:
- Speech: 637 wpm
- Stenotype: 360 wpm

Type	Experienced User WPM	Professional WPM
Stenotype	180	230
Speech	100-150	(250 for auctioneers, but software cannot yet process at this rate)
QWERTY keyboard	60	100
Twiddler	50	
Dasher	40	
Handwriting	30	
Graffiti	15	

Courtesy of Prof. CHAM Tat Jen

Topics:

- Introduction: past, current, and future
- Input devices
  - Text entry: keyboards and keypads, etc.
  - Pointing: terminologies, devices, and Fitts's Law
  - Natural User Interface (NUI)
  - Other Input Devices
- Output devices
- Other devices
- AI-enabled interactions

What is Pointing?

A pointing device is an input interface for a user to input a *continuous spatial position* to a computer

What can Pointing do?

- Select:
  - user chooses from a set of items., e.g., traditional menu selection
- Position:
  - user chooses a point in a one-, two-, or higher-dimensional space
- Orient:
  - user chooses a direction in a two-, three-, or higher-dimensional space
- Path:
  - user sketches a curving line, e.g., drawing a route on a map

Terminologies

(1) Direct vs. Indirect pointing

- Direct: human actions perform directly on the display surface
- Indirect: human actions perform away from the display surface; motions are mapped to the display

(2) Absolute movement vs. Relative movement

- Absolute movement: a consistent mapping between a point in the input space and a point in the output space (e.g., screen).
  - e.g., stylus, finger on touch screen, etc.
- Relative movement: maps displacement in the input space to displacement in the output space, so it moves the cursor relative to its initial position
  - e.g., mouse, joystick, etc.

#1 Relative Pointing

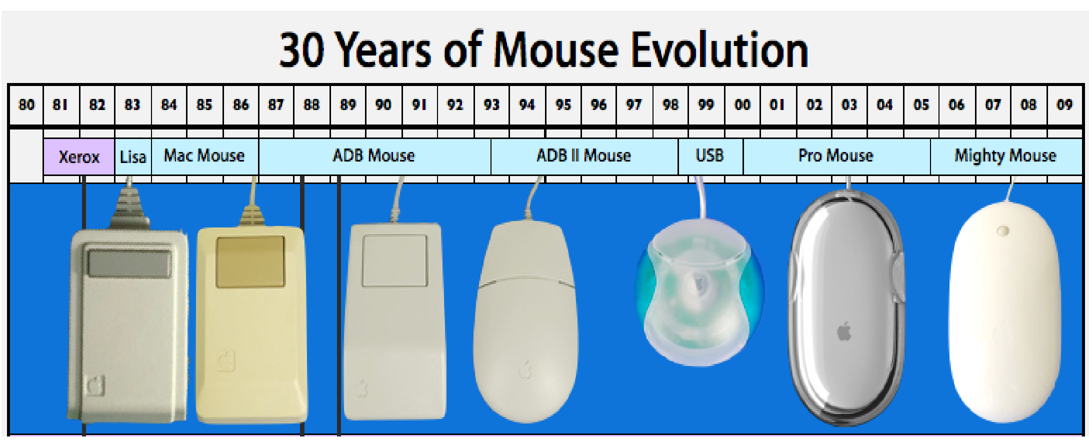
Control-Display Gain (CDG)

- The **scale factor** between movements in the **control space** to the cursor/pointer movements in the **display space**
- Can be tuned, e.g., mouse sensitivity

$$CDgain = V_{Display} / V_{Control}$$

Example #1: Mouse

- Pointing plus clicking (scrolling from 1997)
- Debate: one button vs. multi-button -> recent: multitouch



From: http://appleinsider.com/articles/13/08/10/xerox-parc-the-apple-inc-macintosh-innovator-duplicator-litigator

#1 Relative Pointing

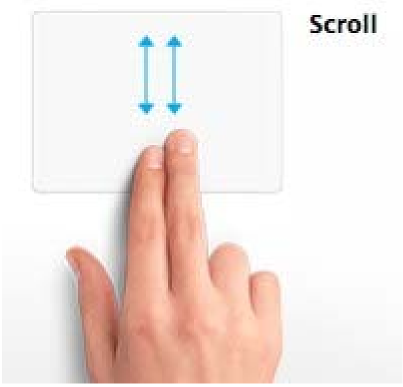
Example #2: Trackball

- Inverted ball mouse
- Compared to mouse:
  - No space issue

https://www.wikiwand.com/en/Computer_mouse

Example #3: Touchpad

- Finger replaces mouse
- Clicks by taps
- Some supports multitouch gestures
  - E.g., two-finger scrolling



#2 Absolute Pointing – 2D

Direct pointing to the actual position on screen

#1: Stylus-based Displays

- Great accuracy, but require an external device stylus

#2: Touch-based Displays

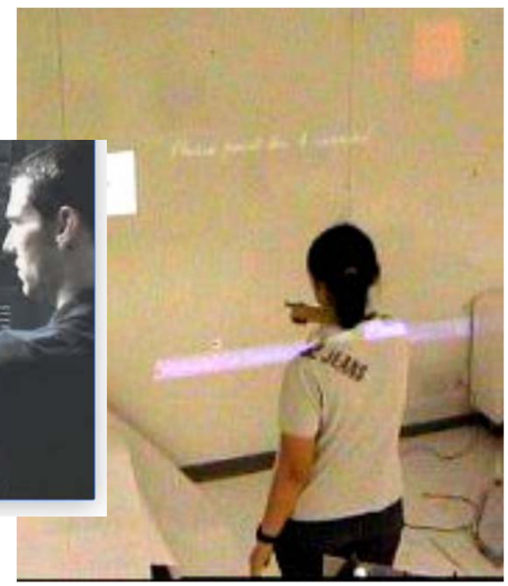
- More intuitive
- But…
  - Less accurate (why?)
  - May not rest wrist or arm
  - Occlusion problem: hand/finger
  - Hygiene issue (public kiosk)

Pointing in 3D?

Aimed Pointing

https://www.youtube.com/watch?v=xkbJC0dUrtc

- Large displays
- Interaction at a distance
- Lower accuracy / less steady
- Approach 1: Using a Device
  - Motion sensing, e.g., Wiipointing
- Approach 2: Natural pointing with arm/finger
  - Computer vision technique (multi-camera)
  - E.g., direction is line from eye to finger tip, not line of whole arm
  - But…
    - Tiring!
      Ergonomic concern
    - Precision



Minority report: https://www.youtube.com/watch?v=PJqbivkm0Ms

#3 Steering

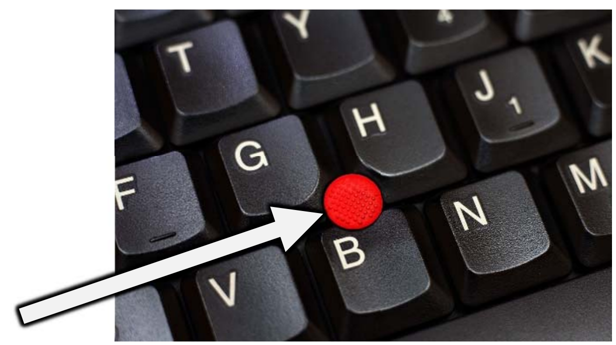
Map “**displacement**” of device to “**digital velocity**” of cursor

- While you keep holding the displacement, the cursor keeps moving accordingly
- Gaming input devices
  - Joysticks
  - Gamepads
- Pointing Stick
  (IBM’s TrackPoint, 1992)
  - Mini joystick; space constraint
  - Why? Motivation?

    Time to shift from keyboard to mouse

Extra reading:

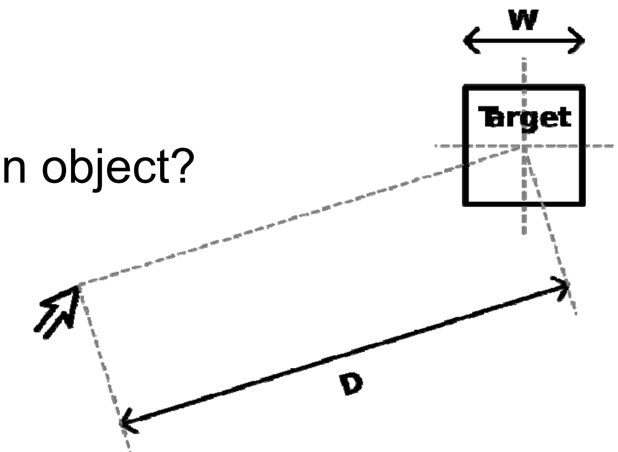
https://en.wikipedia.org/wiki/Pointing_stick



replaceable rubber cap
(pressure-sensitive nub)

Fitts's Law

- Paul Fitts (1954) developed a predictive model of human hand movement primarily used to "**predict time to point at an object**"



- What are the factors to determine the time to point to an object?
  - D – distance to target
  - W – size of target
- Just from your own experience, is this function linear?
  - No, since if Target A is *D* distance and Target B is *2D* distance, it doesn't take twice as long
  - What about target size? Not linear their either
- Mean (average) time to complete the movement, **MT = a + b** $\mathbf{\log_2(2D / W)}$
  - a = time to start/stop in milliseconds and b = inherent speed of the device
  - a & b are usually determined by experiments
  - Example: a = 300 ms , b = 200 ms, D = 16 cm, W = 2 cm
    - Ans: $300 + 200 \log_2(2*16/2) = 1100$ ms

https://en.wikipedia.org/wiki/Fitts%27s_law

Fitts's Law

Comparing Two Devices using Fitt's Law

- **Device 1**
  - A_1 = 500 ms (slower to get started)
  - B_1 = 100 ms (moves a bit faster)
- **Device 2**
  - A_2 = 300 ms (faster to get started)
  - B_2 = 200 ms (moves a bit slower)

- For Distance/Target D=16, W=4
  - Which is faster?

  MT = a + b log2(2D/W)

  Device 1 = 500 ms + 100 ms * $\log_2(32/4)$ = 800ms

  Device 2 = 300 ms + 200 ms * $\log_2(32/4)$ = 900ms

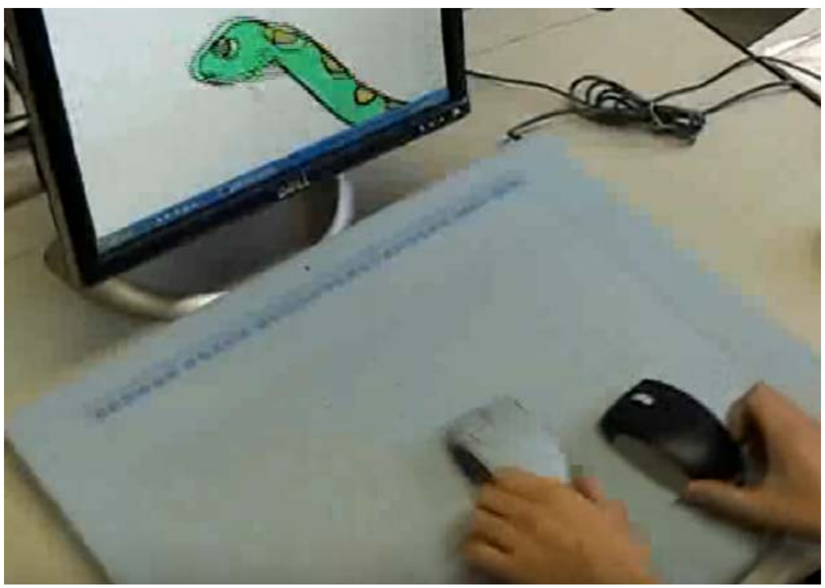
<u>COMMENTS:</u>

- Fitt's Law is proven to provide good timings for most age groups
  - DRAWBACKS
    - Fitt's Law considers 1D motion
    - Pixels at screen edges and corners are easier to select
- Any UI design related to Fitts's Law?
  - Top menu in Mac OS; Bottom-left corner in Windows; pie/radial menu

Others: Multitouch

Multitouch is direct pointing (direct manipulation)

- Allows multiple touch points simultaneously
- Obscuration problems: our hand occludes some of the graphics
- Multi-mouse (before multitouch becomes popular)
- Research demo “As-Rigid-As-Possible Shape Manipulation”: https://www.youtube.com/watch?v=1M_oyUEOHK8

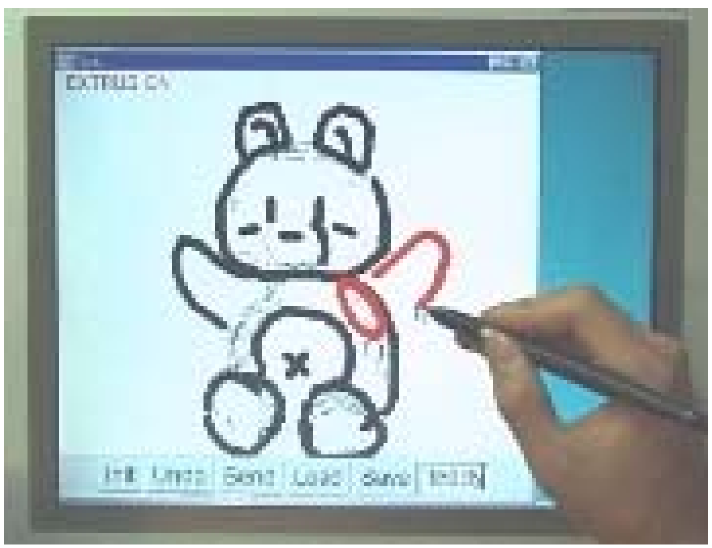


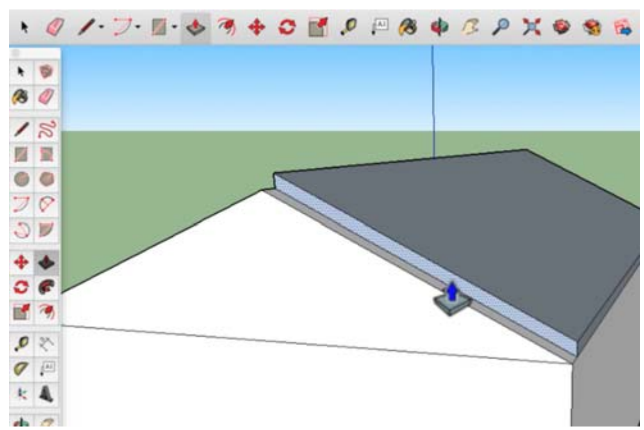
Others: Sketching as Inputs

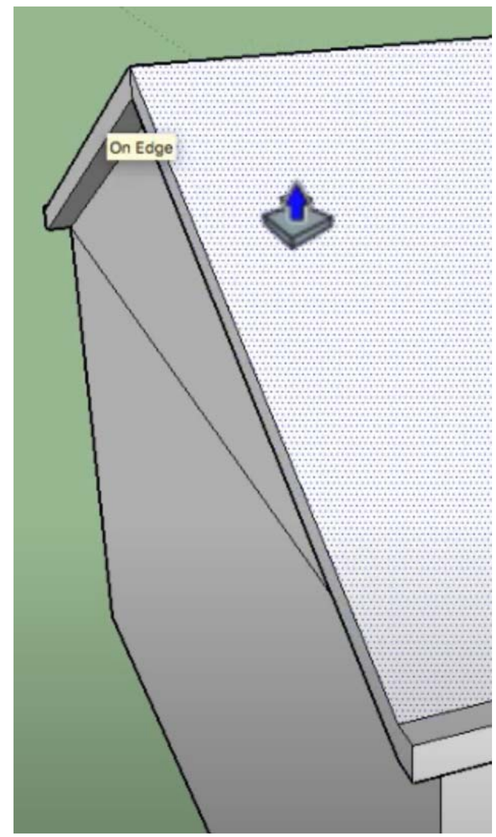
Intelligent systems interpret sketches into higher level concepts

Examples:

- 3D models from 2D sketches
  - Teddy: https://www-ui.is.s.u-tokyo.ac.jp/~takeo/teddy/teddy.htm
  - Google Sketch-up: https://www.sketchup.com/ & https://www.youtube.com/watch?v=RJtSeS5zcgI





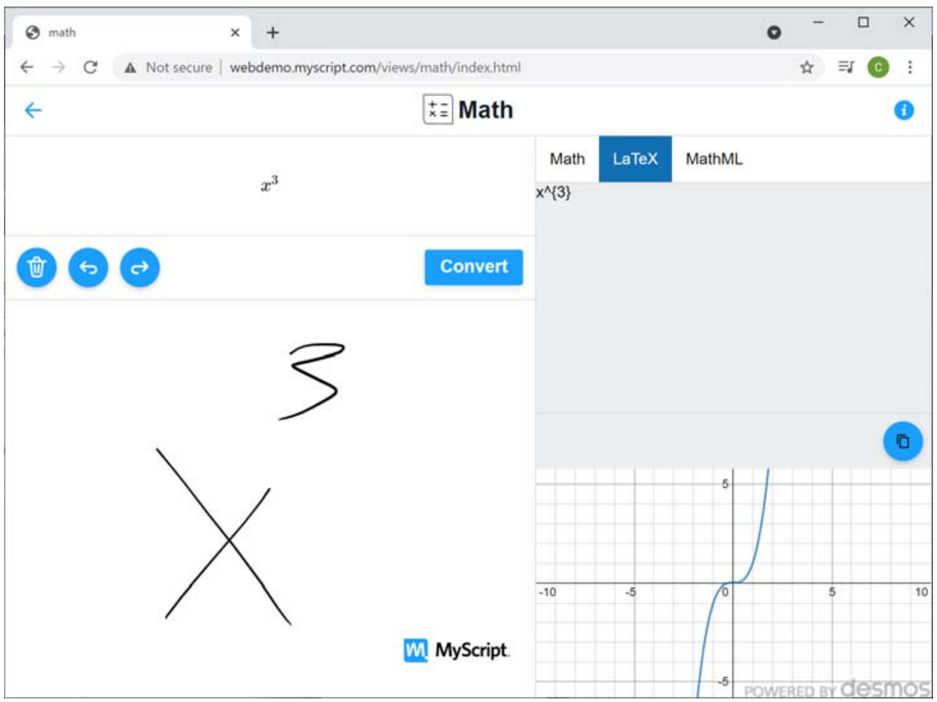


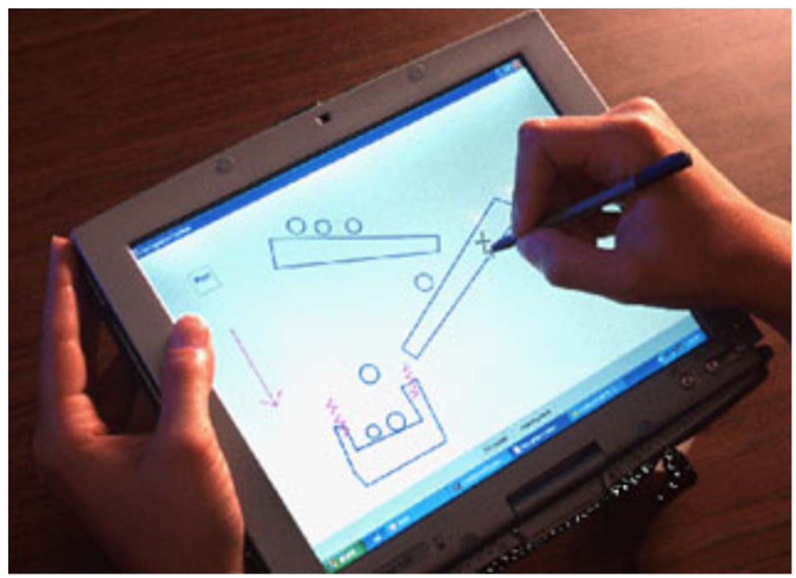
Others: Sketching as Inputs

Sketch recognition methods to convert sketches into higher level forms

Examples:

- MathPad: http://webdemo.myscript.com/
- Physics simulation
  - MIT / Microsoft Physics Illustrator (freeware) http://www.youtube.com/watch?v=1I2tDiyRWvw

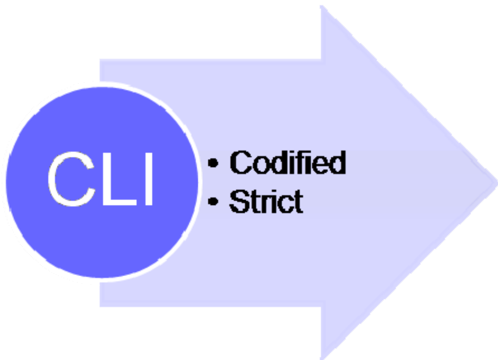




Topics:

- Introduction: past, current, and future
- Input devices
  - Text entry: keyboards and keypads, etc.
  - Pointing: terminologies, devices, and Fitts's Law
  - Natural User Interface (NUI)
  - Other Input Devices
- Output devices
- Other devices
- AI-enabled interactions

Before Natural User Interface (NUI)



C:\Windows\system32\cmd.exe - ftp

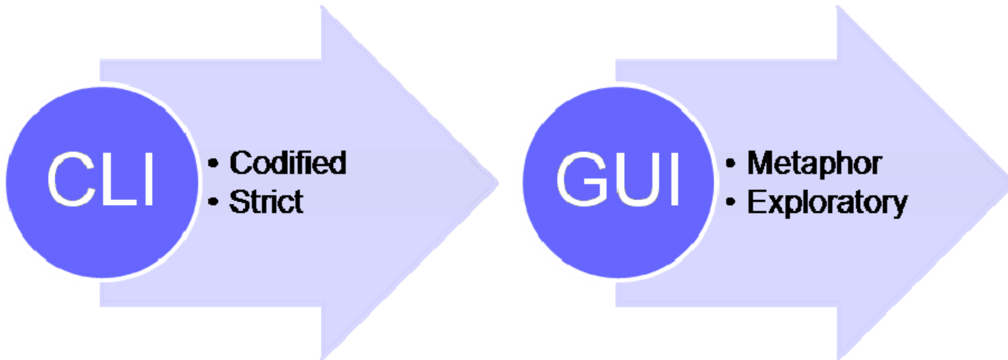
```
Microsoft Windows [Version 10.0.10586]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\Users\philipfu>ftp
ftp> help
Commands may be abbreviated.  Commands are:

!               delete          literal         prompt          send
?               debug           ls              put             status
append          dir             mdelete         pwd             trace
ascii           disconnect      mdir            quit            type
bell            get             mget            quote           user
binary          glob            mkdir           recv            verbose
bye             hash            mls             remotehelp
cd              help            mput            rename
close           lcd             open            rmdir
ftp> open ftp.cse.cuhk.edu.hk
```

```
(1) Wed May 25 19:18:53 H
```

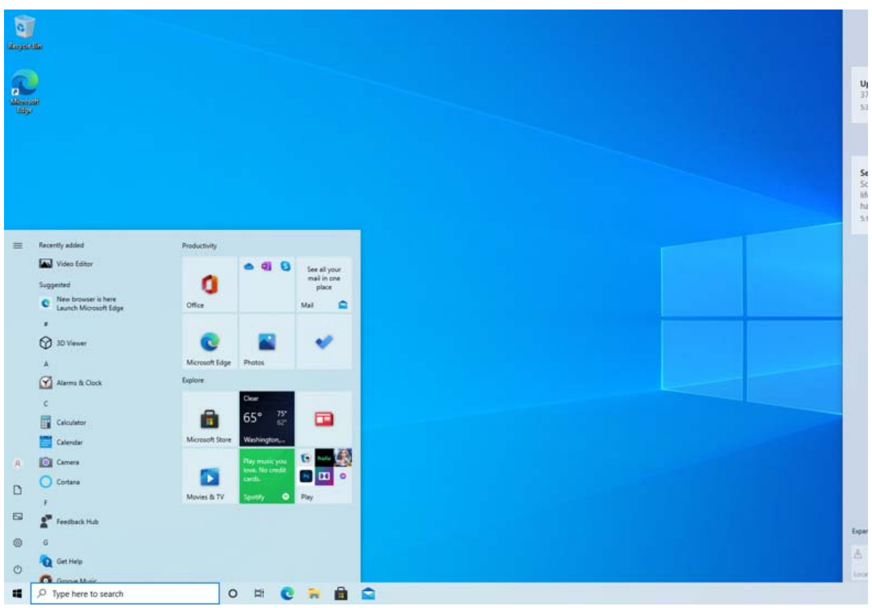
C:\Windows\System32\cmd.exe

```
Microsoft Windows [Version 10.0.10586]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\Windows\system32>cd c:\courses\SEEM3510

c:\courses\SEEM3510>delete midterm.doc
```

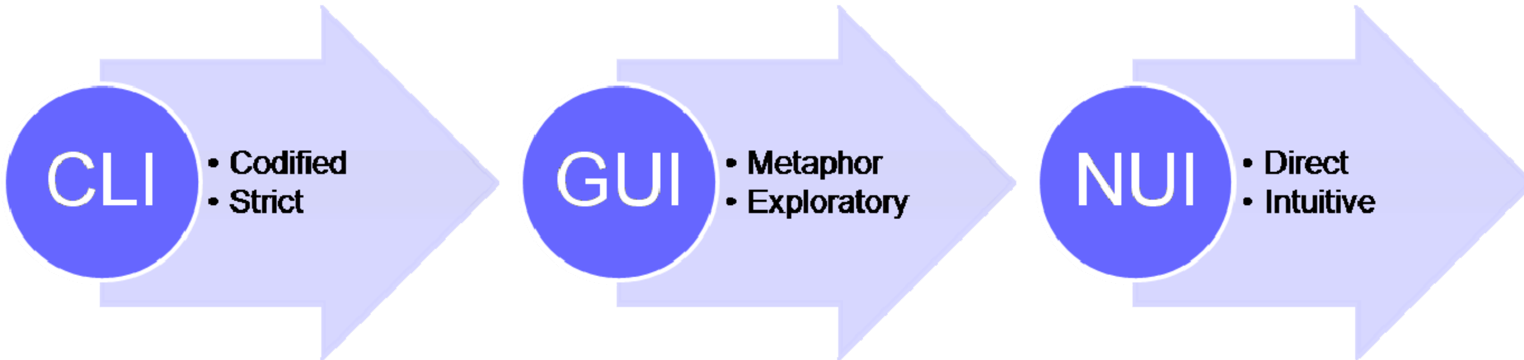
Before Natural User Interface (NUI)



Note:
Windows,
Icons,
Menus,
Pointer



Natural User Interface (NUI)



A user interface that is **effectively invisible** and **remains invisible** as the user progressively learns the interactions.

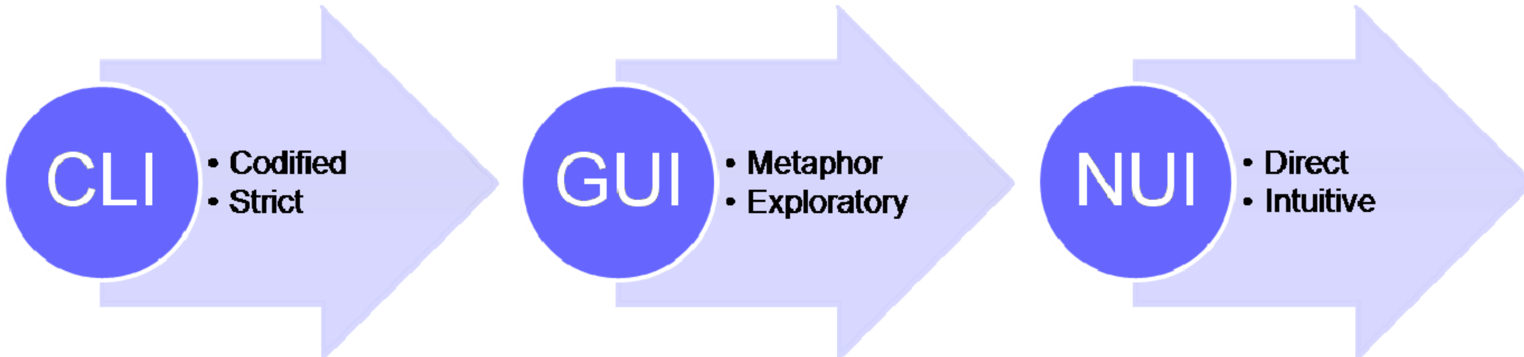
- Post-WIMP

The users are not aware of the existence of the UI

Related to which topics we covered?

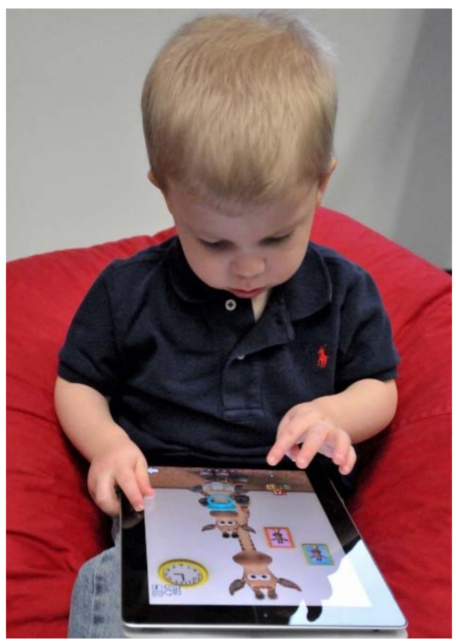
- Direct Manipulation
- Universal Usability

Natural User Interface (NUI)

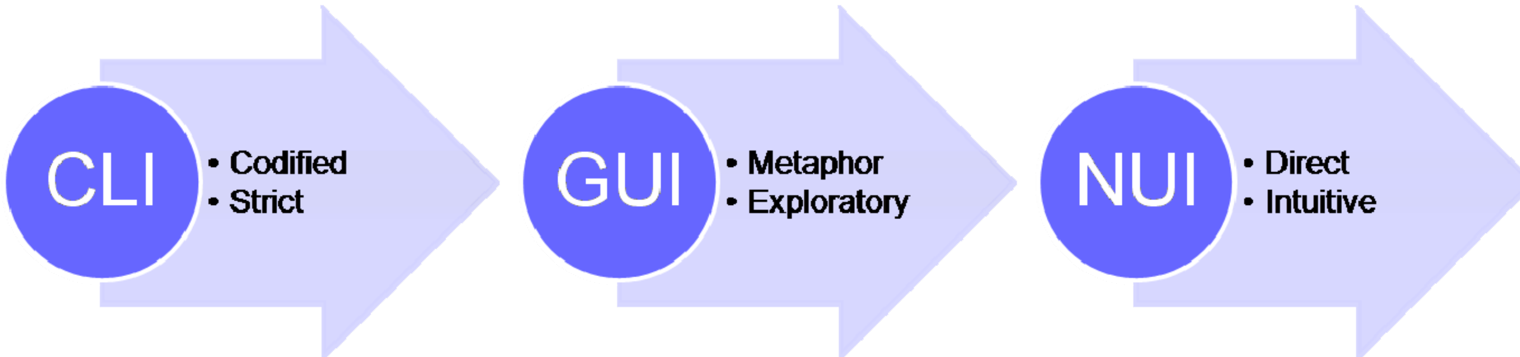


A user interface that is **effectively invisible** and **remains invisible** as the user progressively learns the interactions.

- Post-WIMP
- E.g.,
  - Multitouch
  - Voice



Natural User Interface (NUI)



A user interface that is **effectively invisible** and **remains invisible** as the user progressively learns the interactions.

- Post-WIMP
- E.g.,
  - Multitouch
  - Voice
  - Body posture and motion (Kinect)

  Mostly relate to gestures, motion, and voice (AI-enabled interaction)

Topics:

- Introduction: past, current, and future
- Input devices
  - Text entry: keyboards and keypads, etc.
  - Pointing: terminologies, devices, and Fitts's Law
  - Natural User Interface (NUI)
  - Other Input Devices
- Output devices
- Other devices
- AI-enabled interactions

Others Conventional Input Devices

Other common input devices (besides those we’ve discussed)

- Microphone
- Digital camera
- Steering wheel
- Paddle

etc. etc.

Mostly for non-interactive use:

- Barcode reader
- Fingerprint scanner and biometrics

Other Inputs Devices

What else we have? Or what else can computers sense?

- Device Direction Sensing
- Device Motion Sensing
- Body Posture and Motion Sensing
- Hand Tracking
- Eye Tracking
- Locomotion devices
- Affective Computing
- AI Assistant (Natural Language)
- Brain-Computer Interface

etc….

(1) Device Motion Sensing

Device Direction (Orientation)

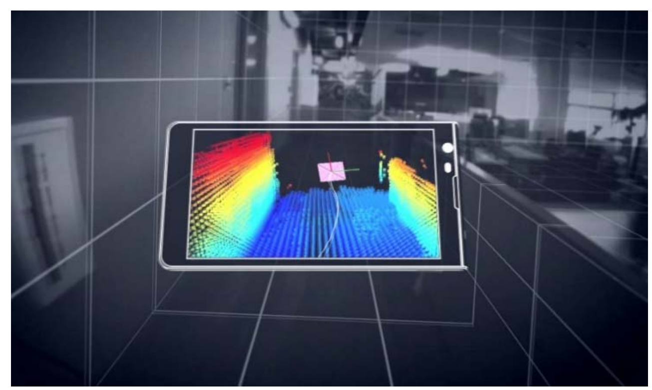
- Tilt sensor (gyroscope) and/or electronic compass
- OUT: vertical inclination, and/or horizontal bearing

https://www.youtube.com/watch?v=cquvA_IpEsA (self reading)

Device Motion Sensing (Inside-out)

- Accelerometers to sense motion
  - e.g., Nintendo Wii & Nintendo Switch

  https://www.youtube.com/watch?v=KZVgKu6v808 (self reading)



- Partial motion from cameras, e.g.,
  - AR phone games
  - Google Tango project

    (Depth sensing, motion tracking, etc.)

  https://www.youtube.com/watch?v=Qe10ExwzCqk

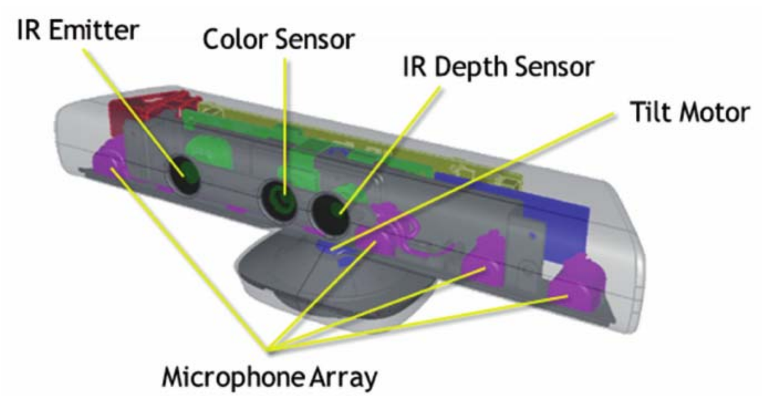
Note: * Inside-out means do not rely on external sensors

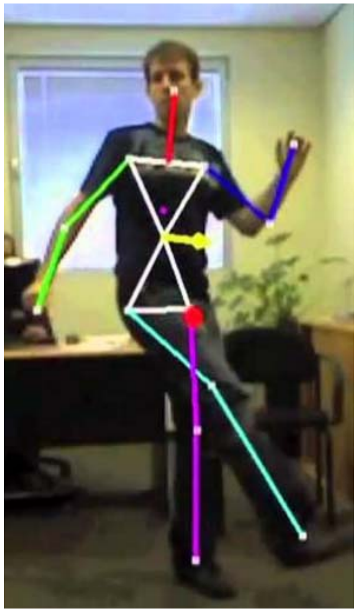
(2) Body Posture & Motion Sensing

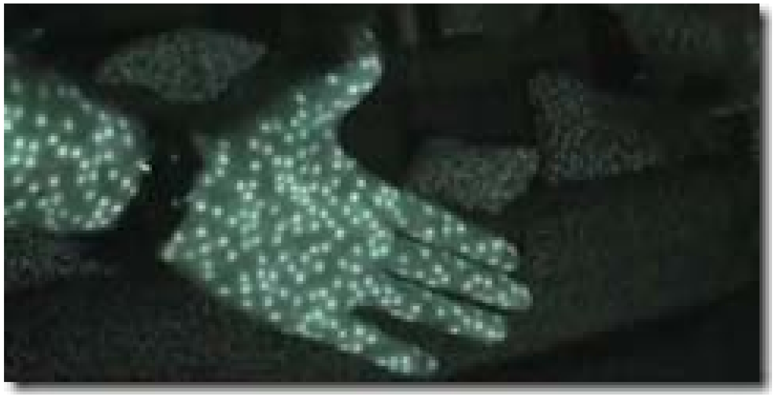
Microsoft Kinect (and many others)

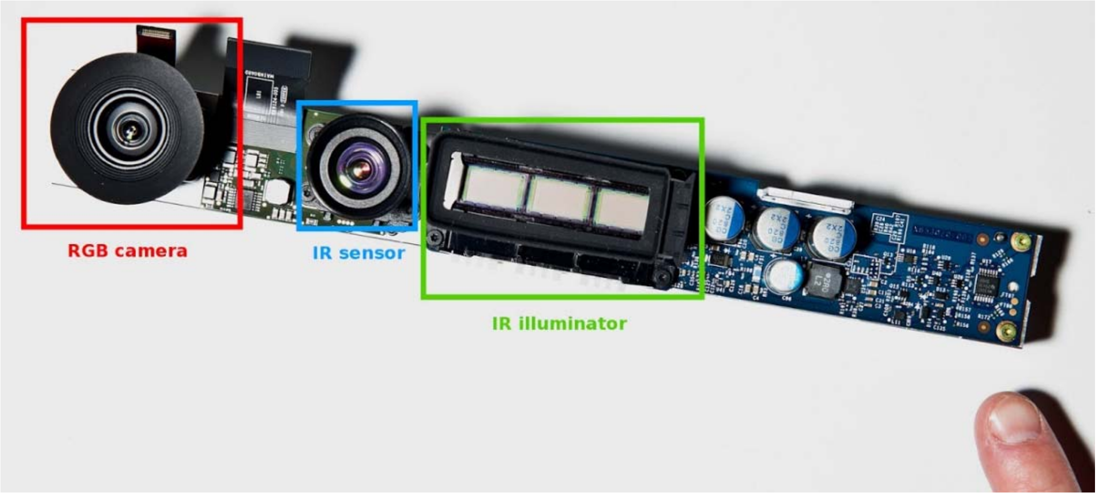
- OLD: structured light 3D sensing
- NEW: time of flight (TOF)

https://en.wikipedia.org/wiki/Kinect









(2) Body Posture & Motion Sensing

Recent research:

- AI-based
- Mostly data-driven & RGB-based

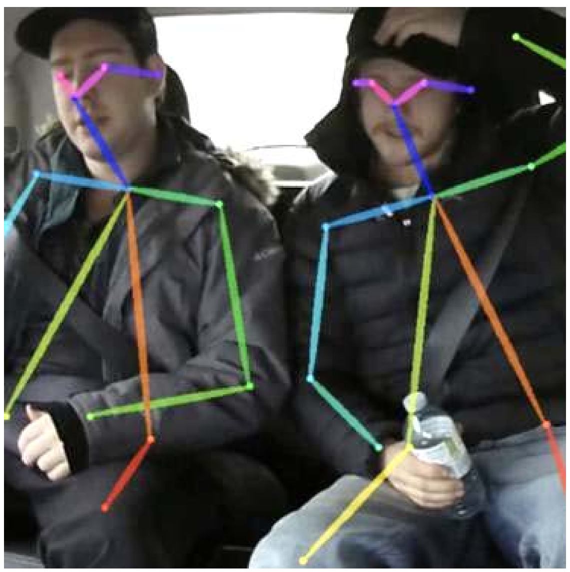
Pose Estimation

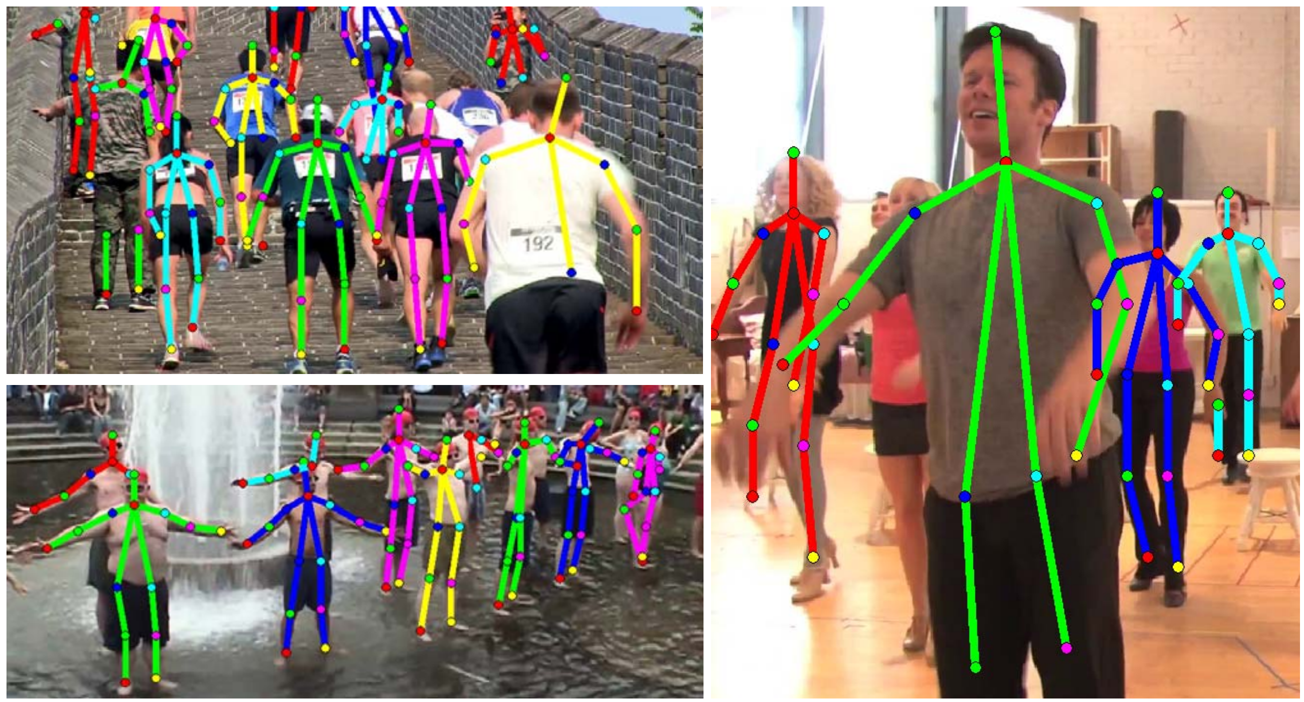
Edit Task

Computer Vision

502 papers with code **13** benchmarks **66** datasets

https://paperswithcode.com/task/pose-estimation



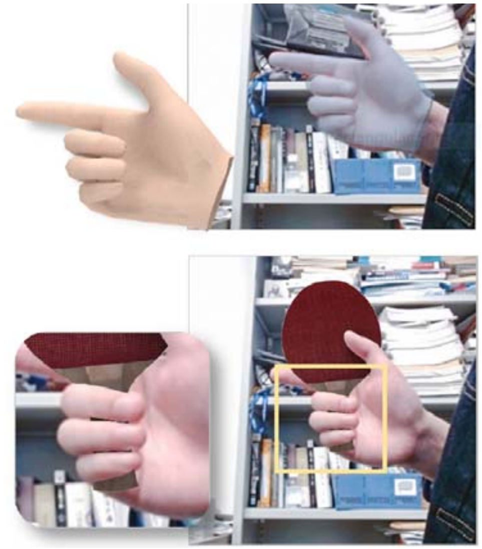


A benchmark dataset: MPII Human Pose Dataset
http://human-pose.mpi-inf.mpg.de/

(3) Hand Recognition & Tracking

Approaches (real-time):

- Data Glove / Cyberglove
  - Accurate + force feedback
- Marker-based
  - E.g., color glove
- Depth camera: RGBD
  - Markerless and low-cost
- Deep learning: RGB +/- D



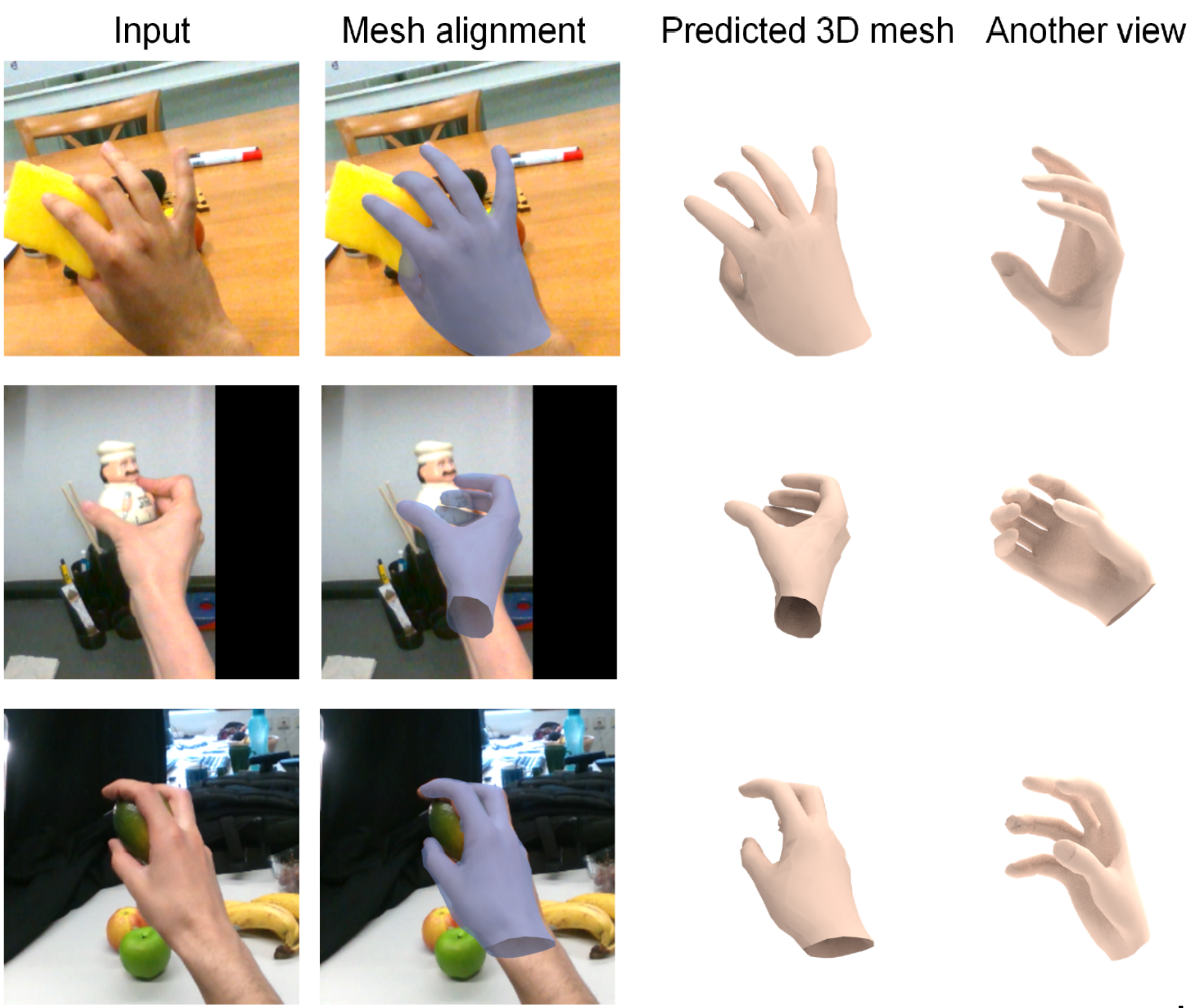
Our Recent work (RGB only)

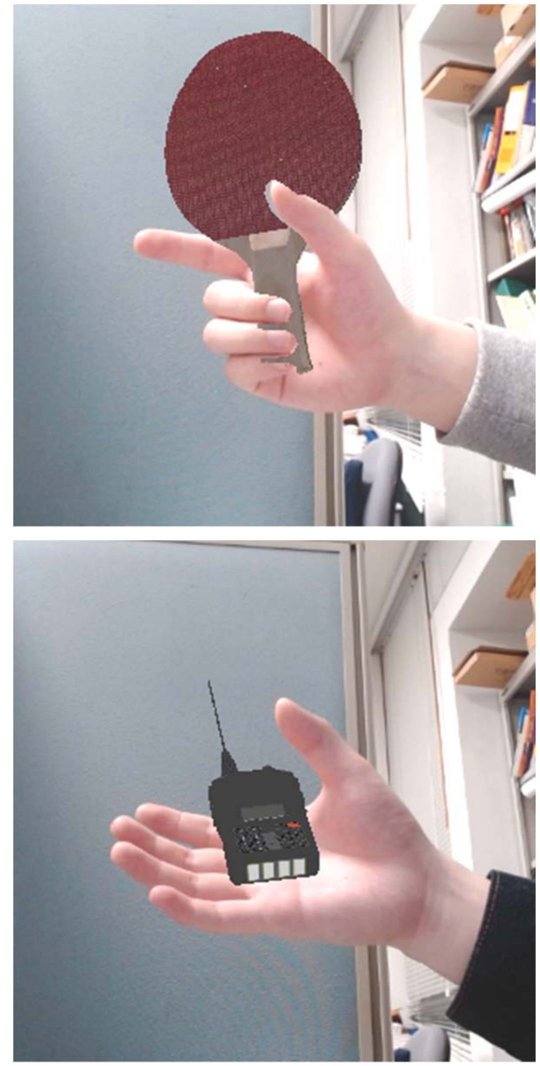
Real-Time Hand-Tracking with a Color Glove
http://people.csail.mit.edu/rywang/handtracking/

Fully Articulated Hand Tracking
https://www.youtube.com/watch?v=2W9x3zmKS_w&t=16s

(3) Hand Recognition & Tracking

Our recent work: Real-time 3D hand mesh from RGB image

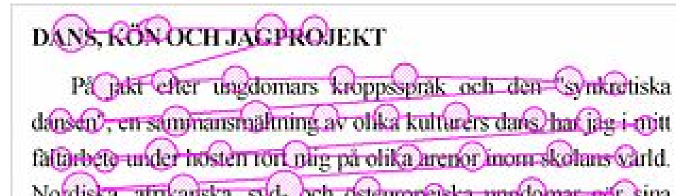




Interact with Virtual AR objects

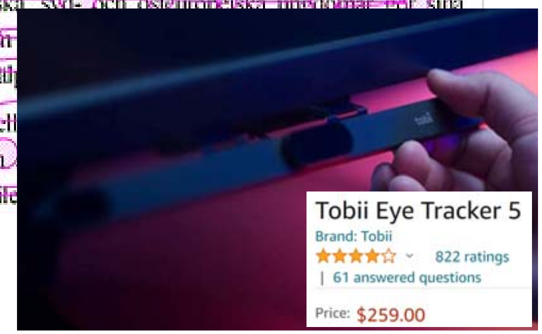
(4) Eye Tracking

Measure the point of **gaze** (where one is looking at) or the motion of an eye relative to the head



- **Fixations**: a series of short stops
- **Saccades**: quick movement

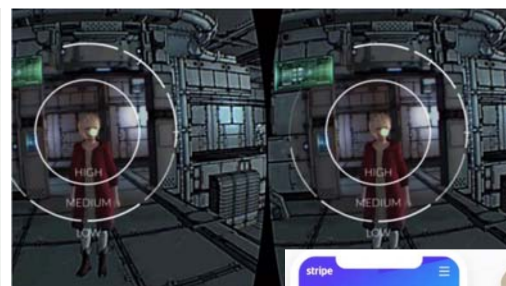
There are many commercial and free tools, both high-end and low-end:



Eye tracking in HMD VR: FOVE

https://www.getfove.com/ (released in early 2017)

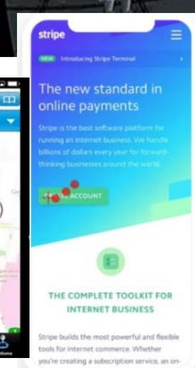
- HTC and Tobii are working together for another VR head set with eye-tracking capability

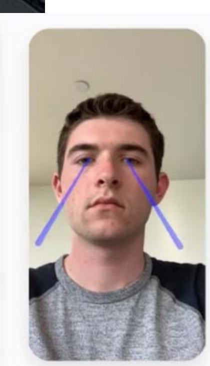


Eye tracking on phone

https://apps.apple.com/us/app/eyetracker/id1443854106

- There are some eye tracking app on phones but they are not yet popular





AI-based Eye tracking

https://paperswithcode.com/search?q_meta=&q_type=&q=eye+tracking

(5) Locomotion devices

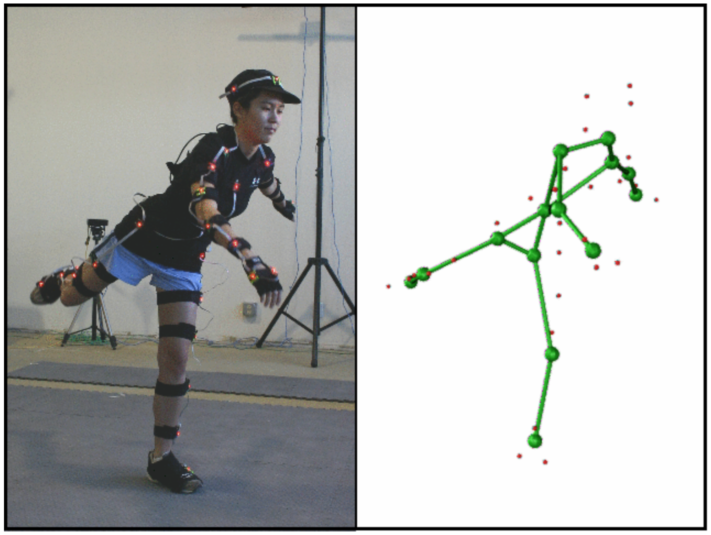
Sense our walking

- Wear HMD
- Physical steps map to virtual space

2000-2010 - Research

Utah's Treadport 2003

Tractor simulator [Iwata 2004]



Optical motion capture

Virtusphere

Torus treadmills

(5) Locomotion devices

Some recent examples:

- DIY Homemade
- Commercial products

Recent Locomotion Devices
https://www.youtube.com/watch?v=qh2UdRKNqH4

Homemade Omni-Directional Treadmill
https://www.youtube.com/watch?v=VeD2n7Y-xIQ

http://katvr.com
(walk, run, back, jump, sit)

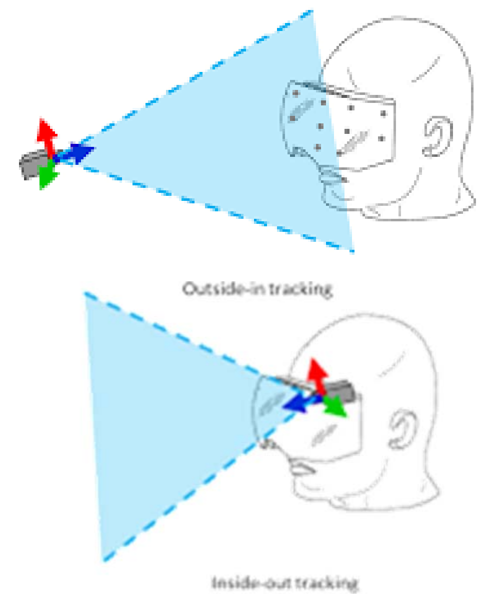
infinadeck.com
sliding floor

Q: possible to get rid of the wire?

Free walking in VR?

There are two types for positional tracking

- Outside-in: use external cameras to track user's position
- Inside-out: without any external cameras to track user

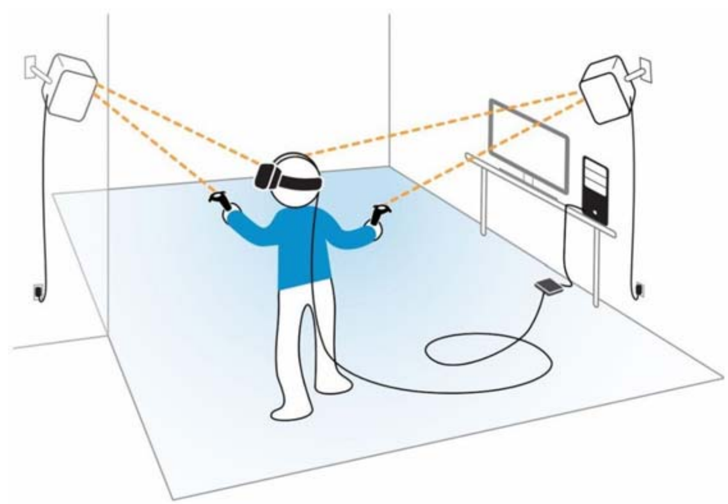


https://uploadvr.com/valve-third-vr/
https://metaversingdotcom.wordpress.com/2015/03/23/examining-the-valvehtc-vive-ecosystem-basic-lighthouse-operation/amp/

Type #1: Outside-in positional tracking

- Two or more external stationary cameras
- Configure the system: build a world reference system and you know the coordinates of the cameras
- Track the positions of reference points on headsets relative to positions of cameras
- Usually use infrared light

E.g., HTC Vive LightHouse



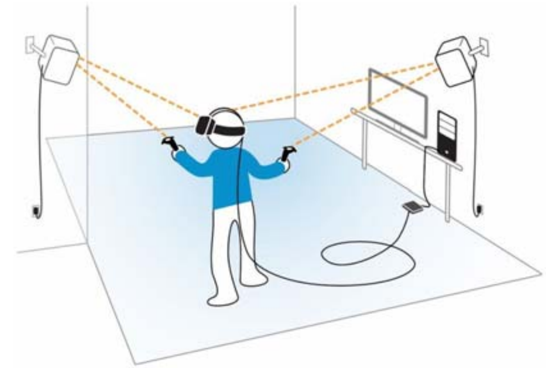
https://metaversingdotcom.wordpress.com/2015/03/23/examining-the-valvehtc-vive-ecosystem-basic-lighthouse-operation/amp/

Type #2: Inside-out positional tracking

- Two or more (usually RGB) cameras on headset's front
- The two cameras scan the real world around the user
- Try to find some special feature points (e.g., corners) in these images
- Possible to reconstruct the 3D position of such points to infer a rough 3d model of the world surrounding the user

Marker-based Positional tracking
SLAM (Simultaneous Localization And Mapping)

Comparison: Outside-in vs Inside-out



Outside-in advantages:

– **Ultra-precise and ultra-fast tracking**

– **Less need of powerful hardware**: Analyzing some reference points is easier than extracting feature points from RGB images

– **Better controllers tracking**: controllers can be tracked everywhere, even if users have them behind his/her back. With inside-out tracking this is not possible

– **It works in the dark**: thanks to IR lights emitted, Vive tracking can work even in the dark. Without visible light, it’s impossible to do it with inside-out tracking based on RGB cameras

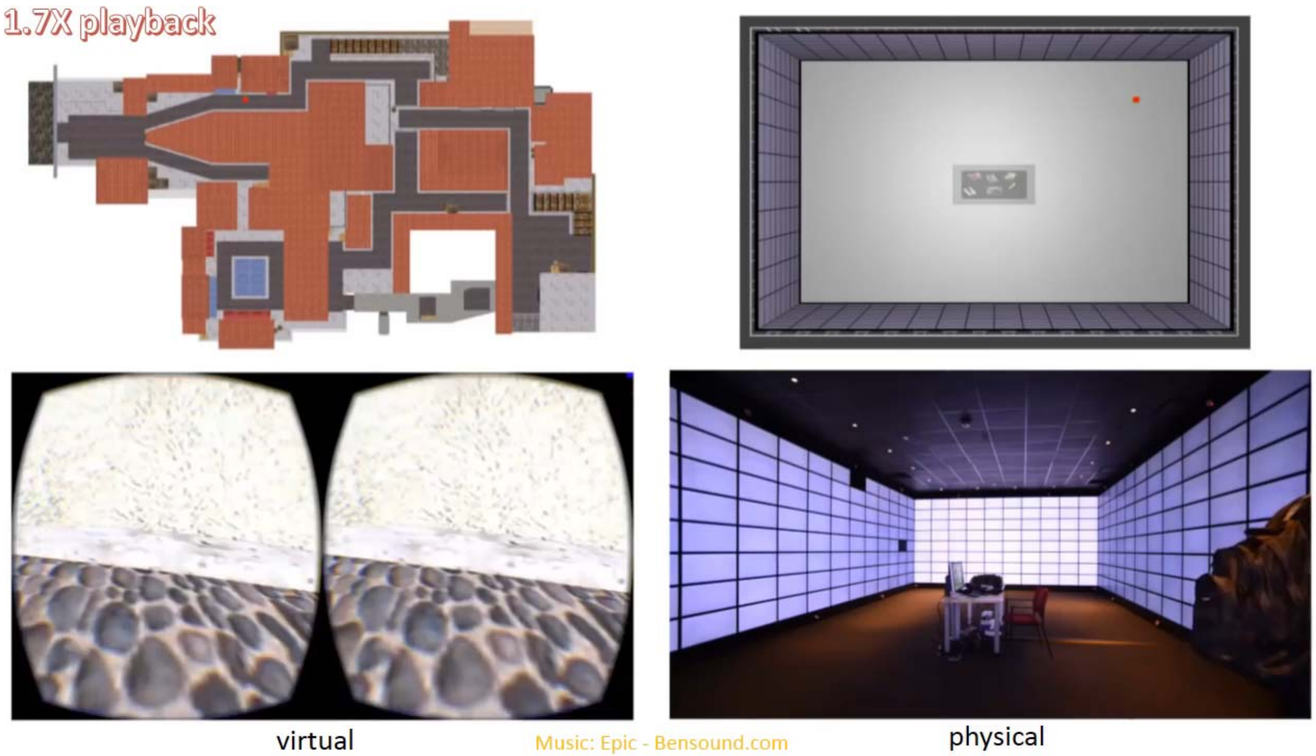
Comparison: Outside-in vs Inside-out

Inside-out advantages:

- **It just works out of the box**: no complex calibrations to be performed, no need to install cameras in the environment, no difficult setups
- **Less hardware**: no extra cameras & no mounts. Good for exhibition setting, you only carry the headset, and nothing more
- **Works everywhere**: Less restricted. Theoretically, you can wear it and walk through entire house. Microsoft calls this world-scale.

https://skarredghost.com/2017/08/09/vr-inside-vs-outside-tracking/

Space Problem? Redirected Walking



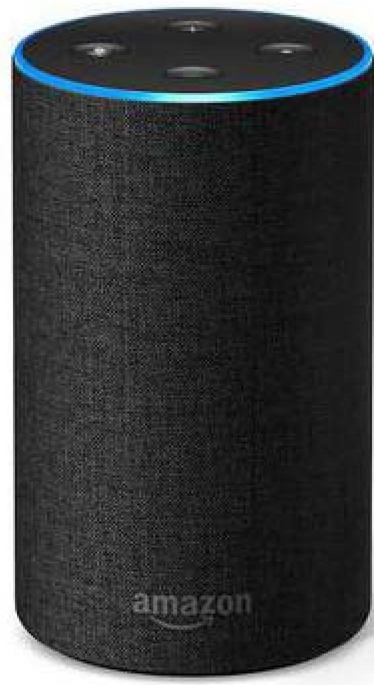
https://www.youtube.com/watch?v=YK9m0wkwjTE

- The rendering of the virtual world preserves the appearance of virtual world and fits the geometry of the physical world
- However, this method brings in visible scene distortion and is unable to handle moving obstacles

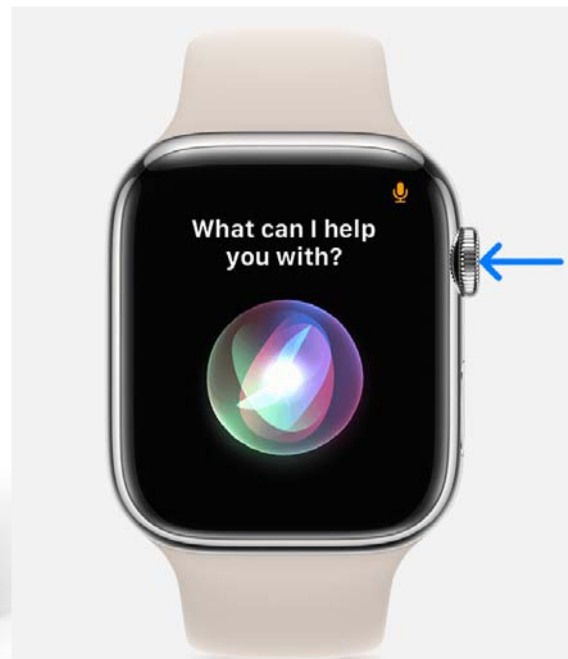
(6) AI Assistant & Virtual Agent

- AI-powered personal assistant
  - e.g., Amazon's Alexa, Google Assistant, Apple's Siri, etc.
- Operated by Natural Language!!!

[Google Assistant](https://www.youtube.com/watch?v=lXUQ-DdSDoE)
https://www.youtube.com/watch?v=lXUQ-DdSDoE



Alexa



Siri

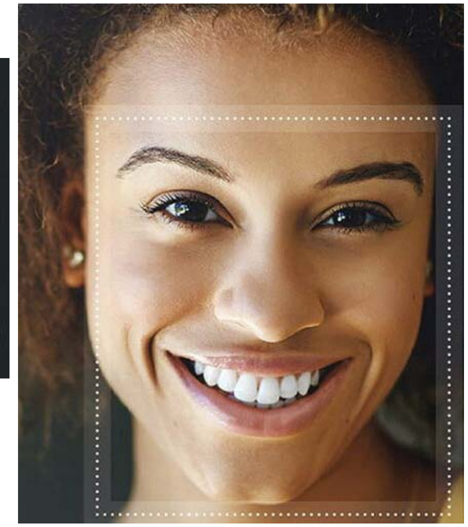
Issues:

- Personal privacy
- Human operators may not know that they are talking to a digital robot

(7) Affective Computing

- Goal: recognize user's emotional state:
  - Facial expression
  - Speech
  - body postures
  - Physiological indicators

    e.g., blood pressure, breathing rates, etc.



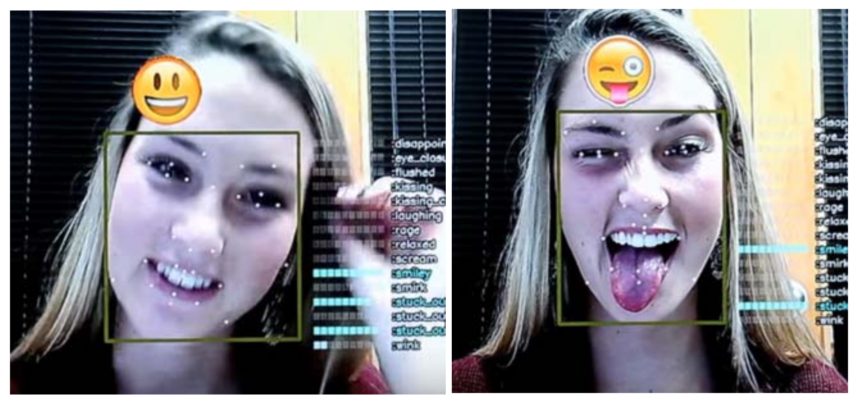
- Commercial Product: Affectiva Affdex

  https://www.affectiva.com/

- Applications?
  - E-learning
  - Monitoring, e.g., driver
  - Affectiva Face to Emoji Mapper

    https://www.youtube.com/watch?v=ABX5RNOalFs
  - Designing Emotionally Sentient Agents

    https://dl.acm.org/doi/10.1145/3186591; https://vimeo.com/297985150



Do you remember this slide?

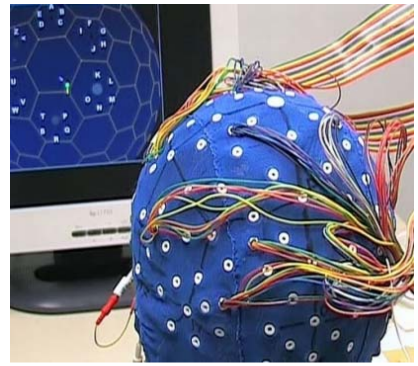
Another example application: from New York Times, Mar 4, 2019:

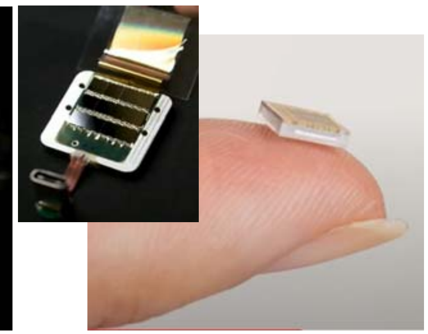
- A researcher at the Expedia Group lab used **emotion tracking** and **eye tracking** software to acquire participant feelings when they arrange trips on travel websites

https://www.nytimes.com/2019/03/04/business/ai-technology-travel-planning.html

(8) Brain-Computer Interface (BCI)

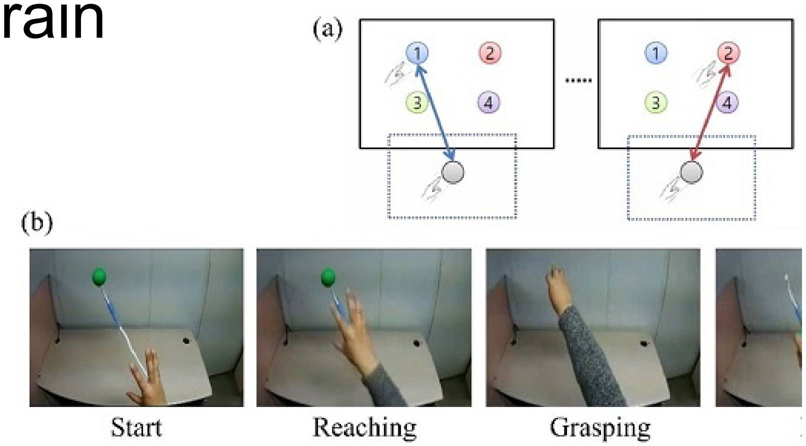
- Also known as mind-machine interface (MMI) or brain–machine interface (BMI).
- Active research area
- Most popular approach: measure EEG signals from the brain
    - Invasive BCIs (high-quality signal)
    - Partially-invasive BCIs
        - Inside skull but outside brain
    - Non-invasive BCIs (i.e., no gels)
        - lower signal quality





- Neuralink: implant tiny chips in brain
    - Goal: iPhone, cursor, typing, etc.
      https://www.youtube.com/watch?v=lA77zsJ31nA
- Decoding Brain Signals to Control a Robotic Arm

https://news.kaist.ac.kr/newsen/html/news/?mode=V&mng_no=19370



Summary: Input Devices

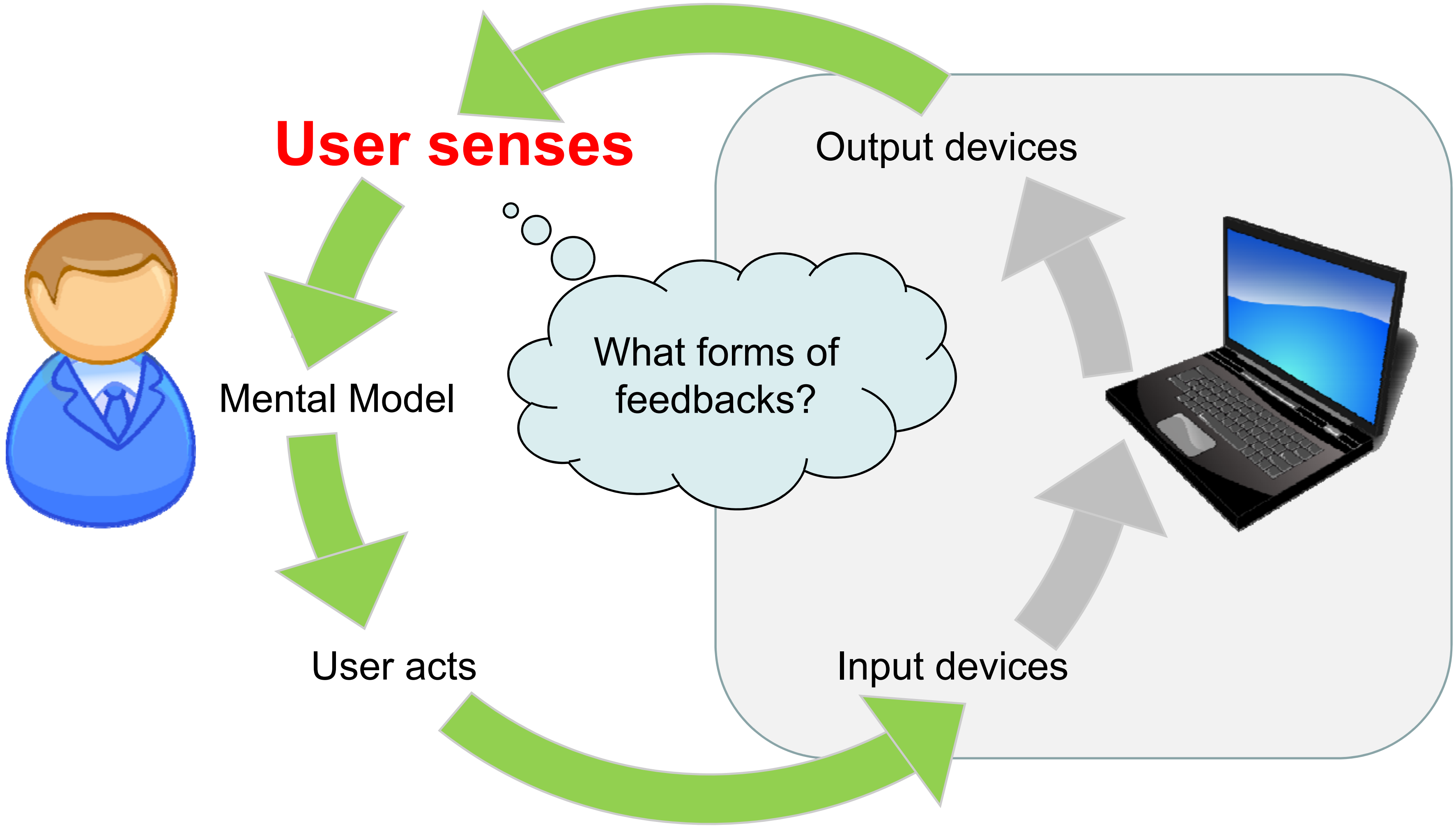
Various considerations and issues with input devices:

- **Eyes-free interaction**
- **Tactile feedback** and **audio feedback**
- **Ergonomic** (long usage, arm/wrist rest, comfortable, etc.)
- **Space constraint**: need how much physical space?
- **Obscuration** problems: hiding something?
- **Natural user interface (NUI)**: gesture, hand and body posture + motion, voice, etc. – physical actions; more natural, more acceptable by users

Topics:

- Introduction: past, current, and future
- Input devices
- Output devices
  - Human sensations
  - 2D & 3D Displays, audio and 3D sound, haptics, etc.
- Other devices
- AI-enabled interactions

Human Sensation



What can we sense?

1. Sight (vision)
2. Hearing (audition)
3. Smell (olfaction)
4. Taste (gustation)
5. Touch (tactition)
6. Balance (equilibrioception)
7. Personal body pose (proprioception)
8. Temperature
9. Pain
10. Others: thirst, hunger, need for breath, rest, etc.

Output Devices

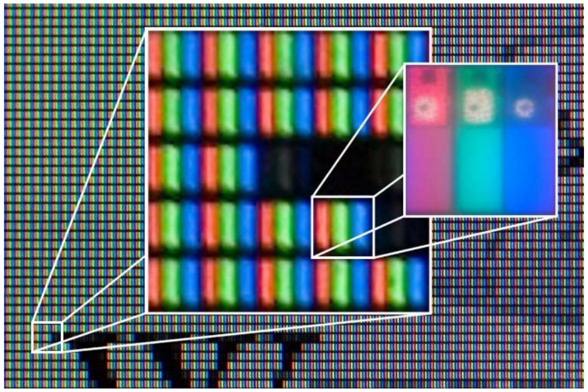
1. 2D / 3D Displays

- There are many different kinds of displays…

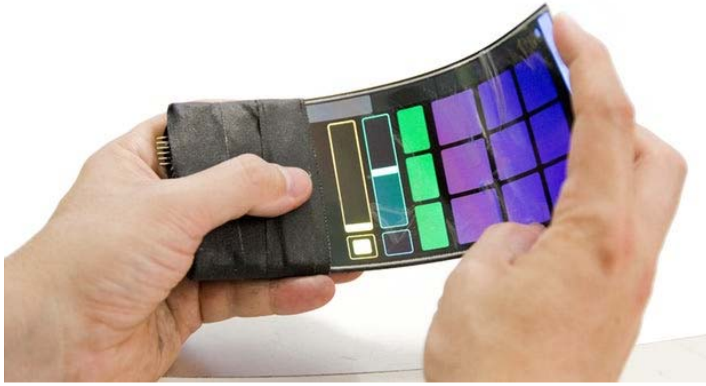
Amazon Kindle

iMax

Semi-transparency: AR



Common: LCD

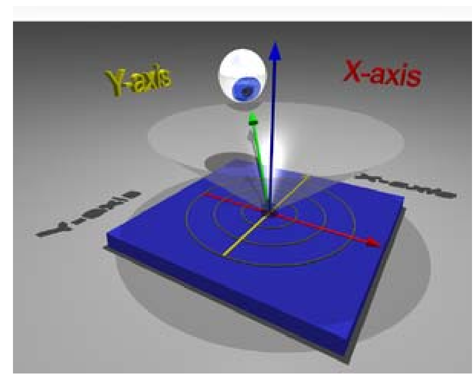


Bendable display & TV

https://www.youtube.com/watch?v=4qnLTQwxw1o

Spatial Characteristics of Displays

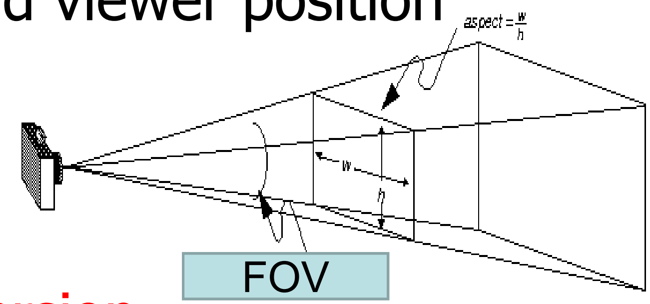
- Display Physical Sizes
  - typically defined by diagonal length
    e.g., 20" monitor, 47" LCD TV
- Aspect Ratios
  - Ratio: width to height
  - Typical: 4:3 or 1.33 (CRT, older LCD, iPad, business),
    1.5 (photo frames), 1.6 (LCD monitors),
    16:9 or 1.78 (TV's, and entertainment projectors),
    2.39 (widescreen cinema) * pixel aspect ratio (square pixel?)
- Display Pixel Resolution
  - Various standards, e.g. XGA (1024x768),
    UXGA (1600x1200); examples: iPad 1 = XGA,
    many 20" monitors roughly 1600x1000
  - HDTV: 1080p / 1080i (1920x1080)



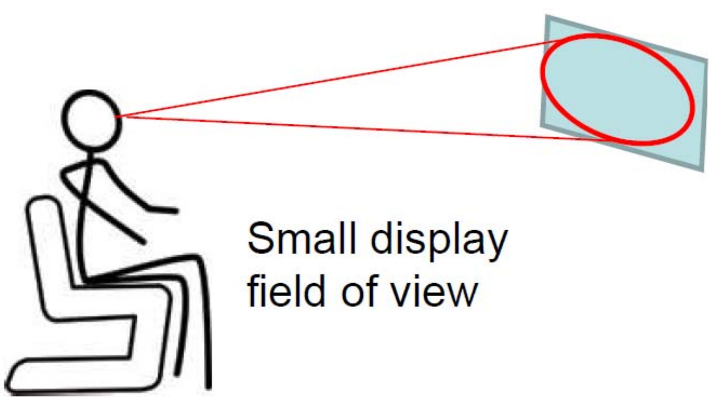
- Viewing Angle / Viewing Cone
  - Max off-center viewing angle where display quality is acceptable
  - Greater → better for sharing; smaller → better for privacy
  - Examples: 140° for desktop LCD, 175° for LCD HDTV

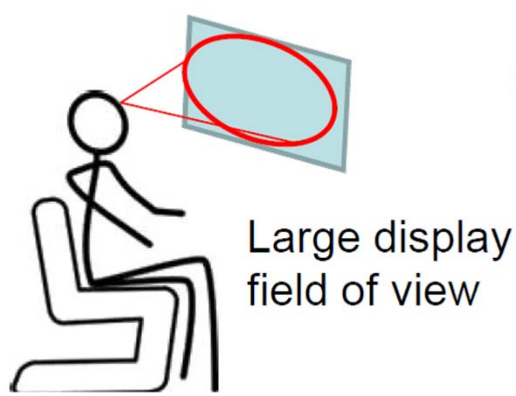
Display Field of View

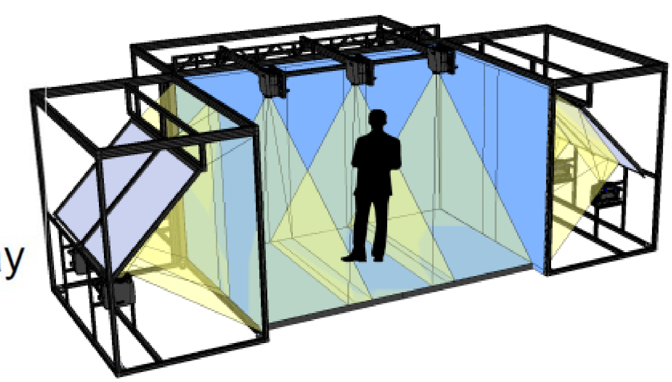
- Visual angle the display takes up from the viewer's position
  - Depends on display size, aspect ratio, and viewer position
- Typical examples (horizontal FOV):
  - 20″ monitor at 60cm distance ~ 35° FOV
  - 42″ HDTV at 3m distance ~ 15° FOV



- Larger field of view → greater viewer immersion
  - Recall: human visual field roughly 220° horizontal, 120° vertical
  - e.g., IMAX is full visual immersion, so are some CAVEs
- Q: *So why not just move closer to the display??*





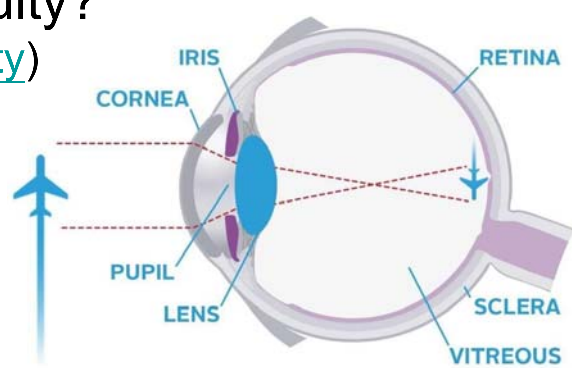


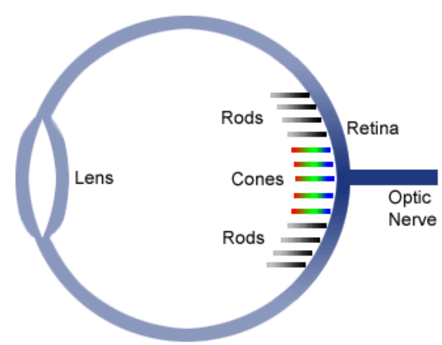
Pixel Density

- Pixel density measured in pixels-per-inch (PPI), e.g.,
  - 50” HDTV at 1080p = 45ppi
  - 20” LCD monitor with UXGA resolution = 100ppi
  - iPhone “Retina” display = 329ppi

Steve Jobs: *“It turns out there’s a magic number right around 300 pixels per inch, that when you hold something around 10 to 12 inches away from your eyes, is the limit of the human retina differentiate the pixels.”*

Homework: is this valid by geometric calculation and human visual acuity?
(https://en.wikipedia.org/wiki/Visual_acuity)





Intensity Characteristics of Displays

- Brightness
  - Monitors / TVs: specified by peak luminance
    - Luminance = luminous power per unit area
    - e.g. LCD HDTV ≈ 500 **candelas/m²**
  - Projectors, specified in luminous flux
    - Flux = power per unit 3D angle
    - e.g., portable projectors ≈ 2000 **lumens**
  - *High brightness more important in well-lit places*

- Black Level (black doesn't mean no light)
  - The black level is luminance of black pixels
    - Most displays always leak light
  - Low black level more important in dark places
    - e.g., virtual universe in planetarium / air traffic controller training simulators

- Contrast
  - Contrast ratio = peak luminance : black level
  - Not very useful – does not account for ambient light
  - In general: high luminance for bright places, low black levels for dark environ.

3D Displays Overview

- Two human visual cues used:
  - Stereopsis
    - Seeing 2 slightly different images in each eye
  - Motion parallax
    - Seeing slightly different images as you *move* around

- Terms used in classifying 3D display attributes:
  - "Stereoscopic"
    - Different image to each eye, viewer must wear special glasses
  - "Autostereoscopic"
    - Different image to each eye, does not require special glasses
  - "Multi-view"
    - Different images depending on viewer's position

  Note: "Multi-view" can be used independently of "(auto)stereoscopic"
    - Possible to have "multi-view 2D" (motion parallax only), "multi-view stereoscopic" or "multi-view autostereoscopic" (both motion parallax and stereopsis) displays

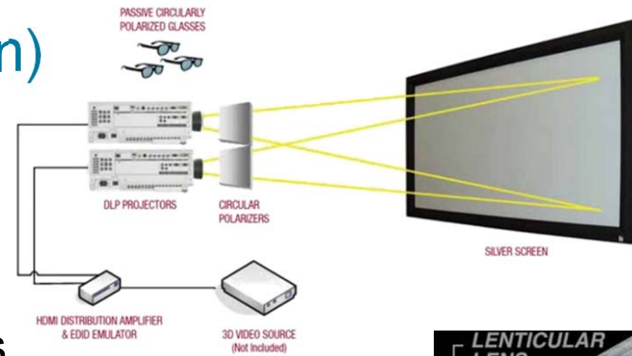
Stereoscopic and Autostereoscopic Displays

- Stereoscopic Displays (common)

  Two common approaches:

  1. Using circularly polarized glasses
     - 3D cinema and LG 3DTVs
  2. Using active shutter glasses
     - Panasonic / Samsung 3DTVs
     - Better quality, but require batteries

  - Good for more simultaneous users



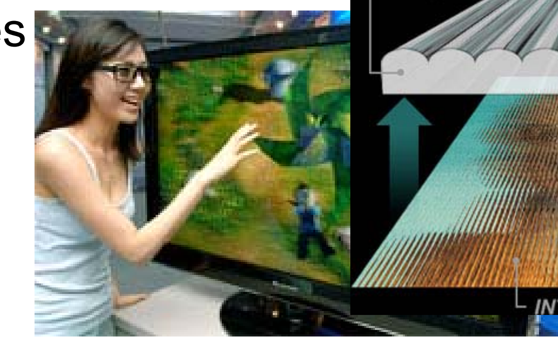
- Autostereoscopic Displays

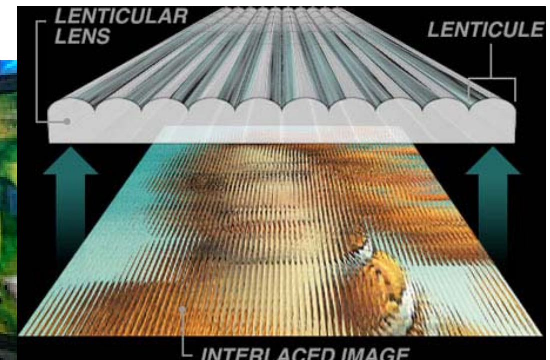
  Two common approaches:

  1. Lenticular lens (brighter but blurry)
     - Philips, some earlier 3D TVs
  2. Parallax barriers (darker but sharper)
     - Most other autostereoscopic displays
     - e.g. LG Optimus 3D phone

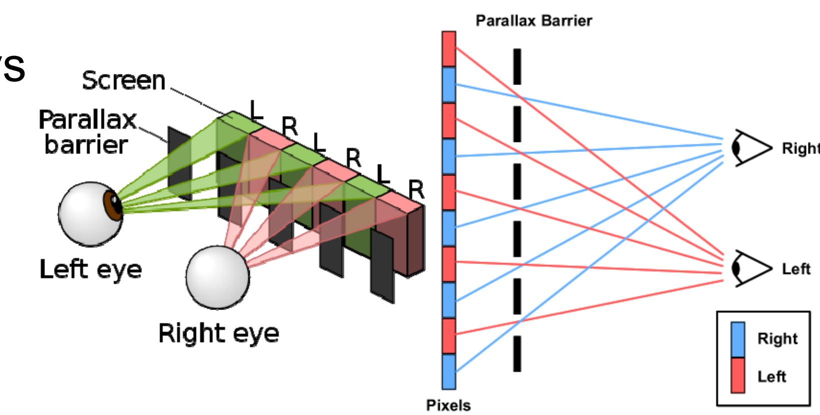
  Any limitations?

  - Typically limited to 1 (or very few) viewers
  - Narrow sweet-spot for viewing 3D





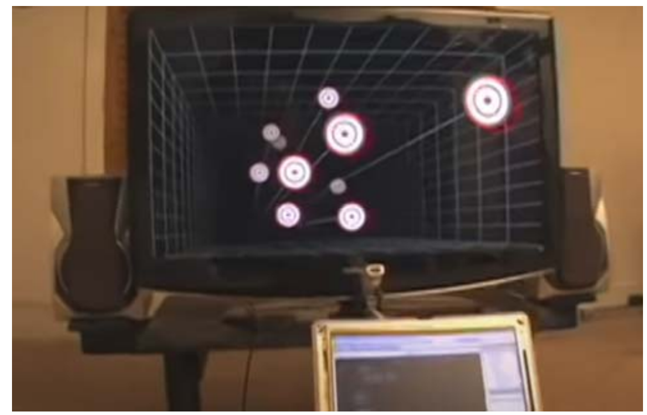
https://www.youtube.com/watch?v=VRNpFp2L8UY



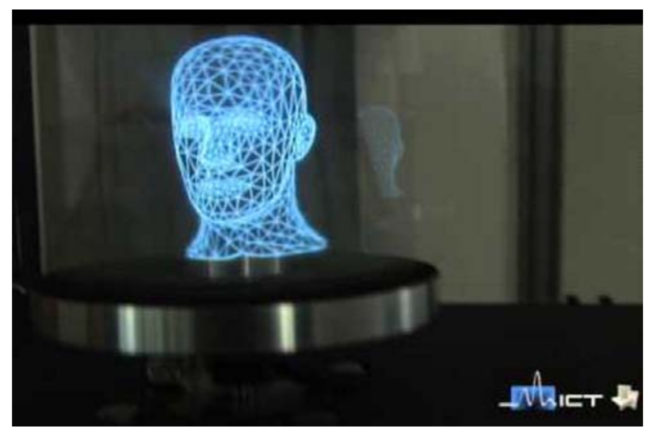
Multi-view Displays

Multi-view typically enabled by **tracking user's head**

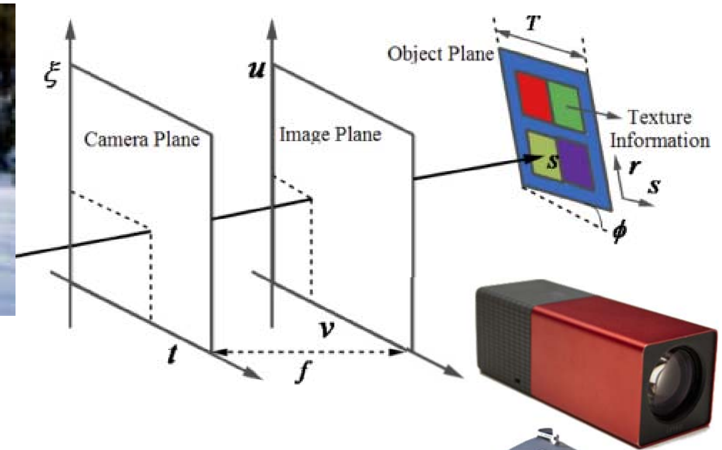
- Multi-view 2D Displays
  - Same image to both eyes, but changes with viewer's position
  - Johnny Lee's demo https://www.youtube.com/watch?v=Jd3-eiid-Uw#t=149s,
    iPad app i3D https://www.youtube.com/watch?v=bBQQEcfkHoE
  - Single user only, e.g., 3DS with tracking:
    https://www.youtube.com/watch?v=pWYgM1RGixM



- Multi-view Stereoscopic Displays
  - Normal stereoscopic system with user tracking
  - Typically single user only
- Multi-view Autostereoscopic Displays
  - There're various mechanisms:
    1. Sweep-volume displays
       - Fast speed 2D illusion
         https://www.youtube.com/watch?v=YKCUGQ-uo8c
         https://www.lumindustries.com/3d-vis
    2. Holographic displays (next page)

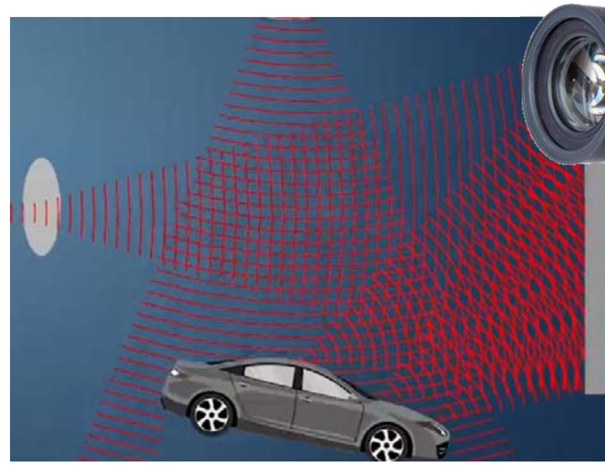


Other 3D Displays



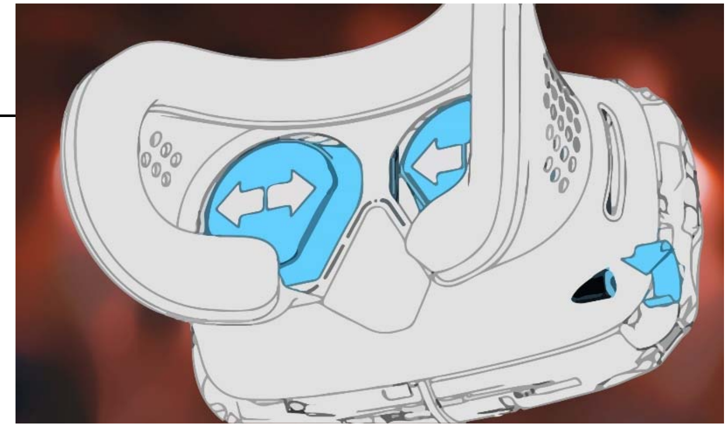
Goal: a 3D image in space

- Background concepts
  - What is 4D light field? All the light rays confined in a volume, e.g., between a pair of planes
  - Light field camera: https://en.wikipedia.org/wiki/Light-field_camera



- Holographic (Light field) display
  - Basic mechanism of hologram: https://www.youtube.com/watch?v=aTB2ryoWIFU
  - A recent product: https://www.theverge.com/2020/12/3/22150437/holographic-display-looking-glass-portrait-holograms-interactive-photos-videos; https://mymodernmet.com/looking-glass-portrait-holographic-display/

Head-mounted display (HMD)



- Show stereoscopic imagery
- Important parameters
  - Interpupillary distance (IPD)
    - Distance between pupil centers
    - Adjustable on some VR HMDs

      http://www.tomsguide.com/faq/id-3064385/set-ipd-correctly-htc-vive.html
  - Field of view (FOV) – the greater the FOV, the better the immersion
    - Human – slightly more than 180 deg.
    - Existing HMDs – around 100 to 110 deg.
  - Lens distance adjustment
    - PSVR & HTC Hive support this – good for those wearing glasses
  - Binocular overlap: scene region common to both eyes
    - Affects our sense of depth and stereo

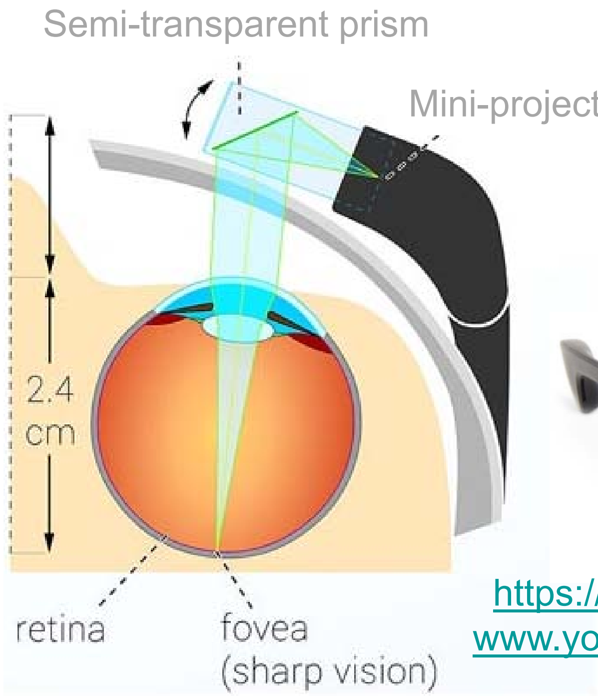
https://vr.google.com/cardboard/

DIY your own cardboard:

https://www.youtube.com/watch?v=asITXtq3iEg

See-through Glasses (AR)

- Displaying images on the visors in front of user's eyes
- Current issues:
  - Small field-of-view (FOV): virtual screen covers only a small region in user's view through the glass
  - Low resolution
  - Battery drain problem



Google Glass and Retina (eye)


HoloLens
https://www.microsoft.com/en-us/hololens
https://www.youtube.com/watch?v=eF2QhJjCaDc


https://www.varifocals.net/google-glass/
www.youtube.com/watch?v=d-y3bEjEVV8


Rumor: Apple AR glasses:
https://www.tomsguide.com/news/apple-glasses

2. Audio

- Purpose
  - Entertainment, communication, user feedback
    - Music, voice, feedback, etc.
    - Note: user feedback not just alerts, but can also be used to as cue

- Basic equipment
  - Earphones / headphones / loudspeakers
- Type: mono, stereo, surround…
- Surround Sound
  - Enhance sound reproduction quality with <u>additional</u> audio channels that surround the listener

2. Audio

Two typical surround sound approaches:

- Binaural audio
  - Dummy Head Recording:
    - Place microphones in ears of a dummy head, and synthesize sound via Head-Related Transfer Functions (HRTF)
  - Examples (close your eyes & use headphones!!!)

    Virtual Haircut <- SHOULD try it!!!

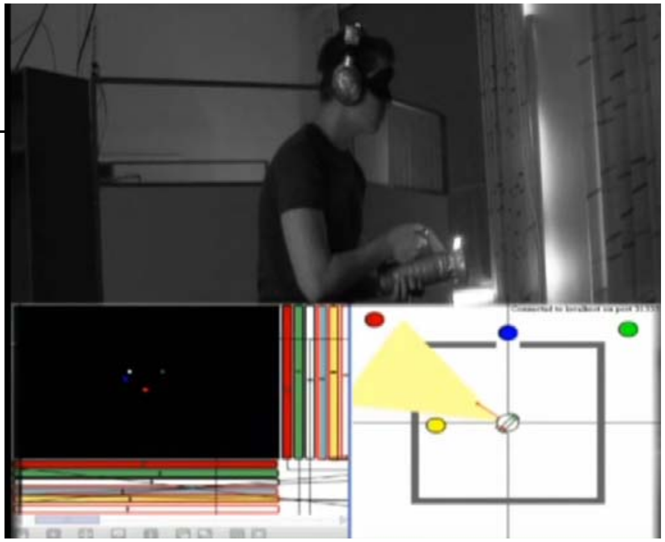
    www.youtube.com/watch?v=IUDTlvagjJA

    www.youtube.com/watch?v=LJwUVCXH-gM

- Loudspeakers
  - Standards: 5.1 or 7.1
    - Refers to number of audio channels
    - Full bandwidth channels feed to 5 (or 7) speakers
    - The ".1" is a low-frequency non-directional channel fed to the subwoofers

subwoofers

2. Audio



Interesting applications:

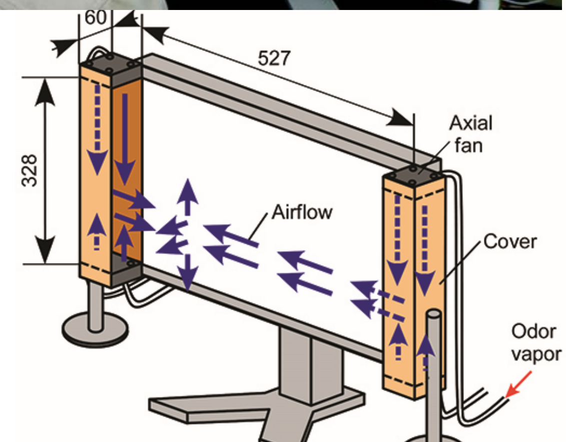
- Audio Based Games
  - Removes visual channel: https://www.youtube.com/watch?v=IJ7dCkWdPC0
  - Example: Virtual auditory based horror game (Medialogy) https://www.youtube.com/watch?v=Nlqfa2EValg
- Text-to-Speech (TTS)
  - Uses a speech synthesis engine
  - E.g., http://ttsreader.com/ and Win & Mac OS also support
  - Useful when text information needs to be conveyed to a user:
    - Visually-impaired and speech disabled
    - Or those whose visual attention focuses on something else
- Voice: Personal Assistant (see input devices)

3. Haptics

- Provide tactile feedback to users
  - Some even with temperature feedback
- Basic vibration cues ("force feedback")
  - Mobile phone vibration
  - Logitech iFeelmouse / Joysticks / Gamepads
- Sensation of virtual 3D shapes
  - Resistive feedback via air bladders (crude) or actuators (more accurate)
  - Data Gloves or Cyber Gloves
    https://www.youtube.com/watch?v=oYB_EJPMRHw
  - Pen-based devices, e.g. SensablePhantom
    - Less tangible but more accurate feedback
- Haptic Station (CyberGlove)

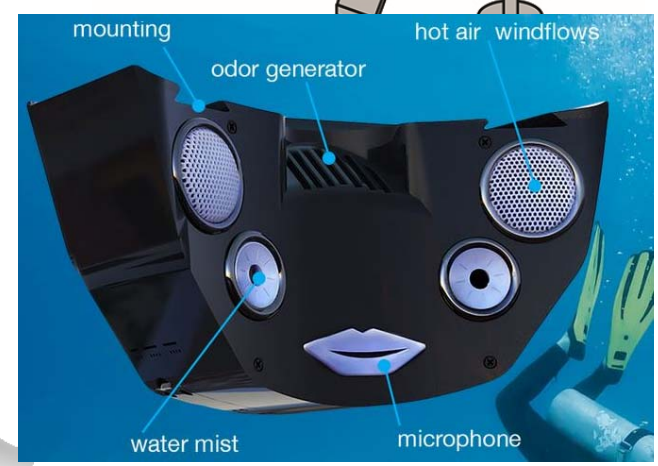
4. Digital Scent Synthesis

- Scent generation
  - aromatic compounds pre-dissolved in solvent
  - when needed, atomized into a fine spray
- How? Odor release systems
  - A fixed set of odors, e.g., tobacco, coffee
  - Scent synthesis from mixing basis scents



- Examples
  - Smell-O-Vision and AromaRama (1960s) in cinemas
  - Smelling screen (from Japan)

  https://en.wikipedia.org/wiki/Smelling_screen
  - VR Headset: Feelreal
    - work with existing HMD VR
    - http://feelreal.com/

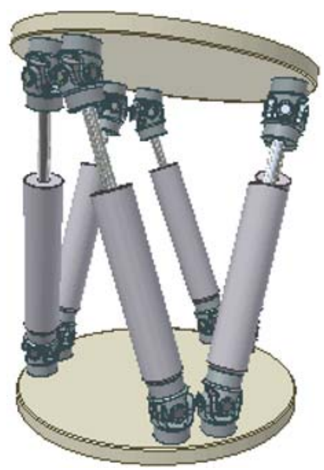


5. Motion Simulators

- Applications:
  - Originally for flight training
  - Now, widely used for entertainment
  - Also personal systems

- Advanced simulators
  - use a Stewart platform design with six prismatic actuators for full 3D movement
- How they trick the body?
  - motion cueing

https://en.wikipedia.org/wiki/Motion_simulator



6. “4D” Experiences

“4D” rides and theater stimulate multiple human senses:

- 3D visuals
- Surround sound
- Tactile input – e.g., wind, water drops, etc.
- Vestibular input – via motion simulation
- Temperature sensation – cold, heat (often via controlled flames)
- Olfactory input – via odor release

Topics:

- Introduction: past, current, and future
- Input devices
- Output devices
- Other kinds of devices
- AI-enabled interactions

Other Kinds of Devices…

- Calm Technology / Ambient Devices – 2 Design Principles:
  #1: Glanceable or pre-attentive (subconscious) feedback
  #2: Aim at minimizing user's mental effort (low attention)
  E.g., ambient umbrella, ambient orb & power aware cord

set to monitor a particular
stock market index or weather

show power or
network data transfer

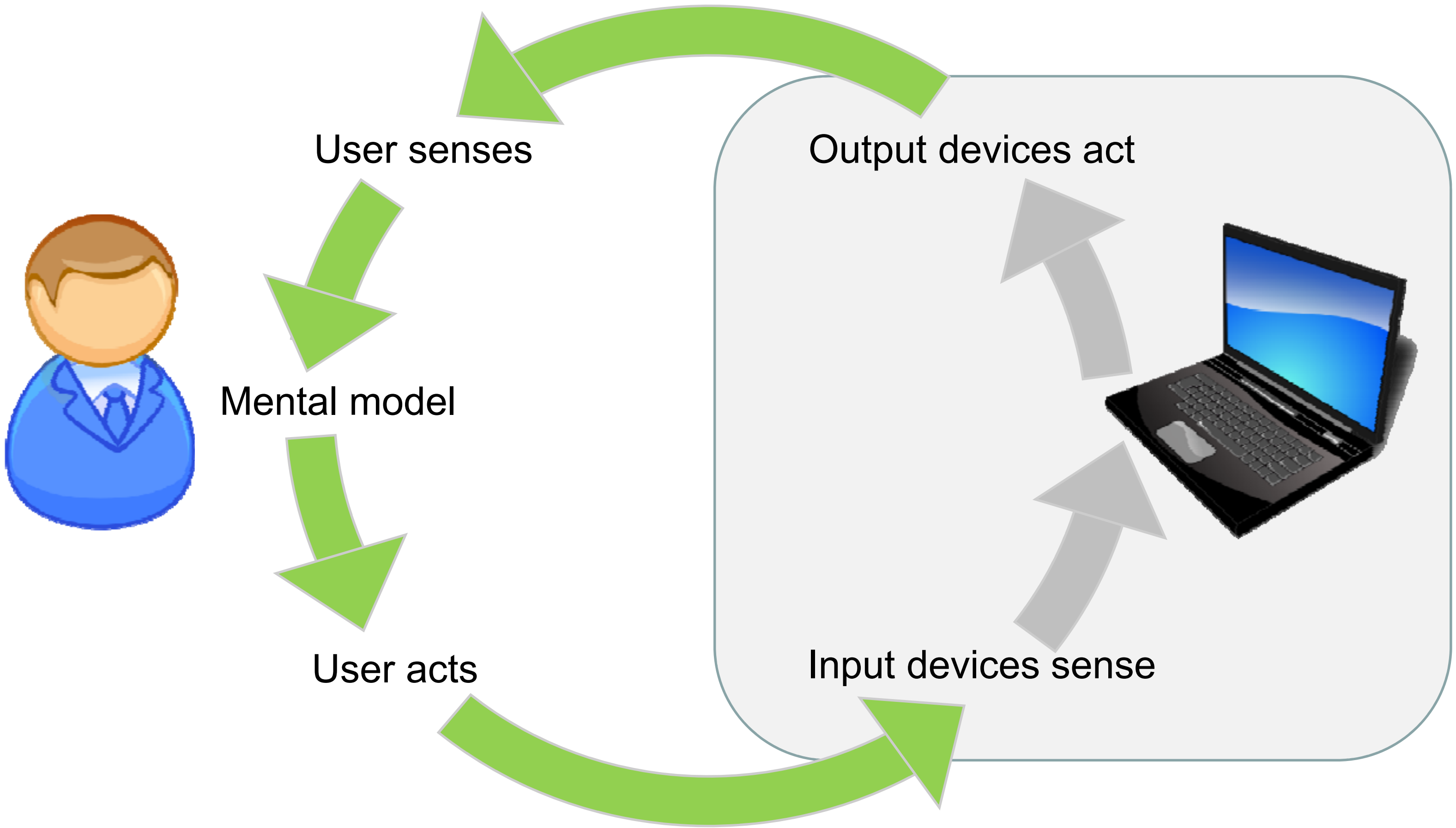
https://en.wikipedia.org/wiki/Ambient_device

Topics:

- Introduction: past, current, and future
- Input devices
- Output devices
- Other kinds of devices
- AI-enabled interactions

AI + Interactions

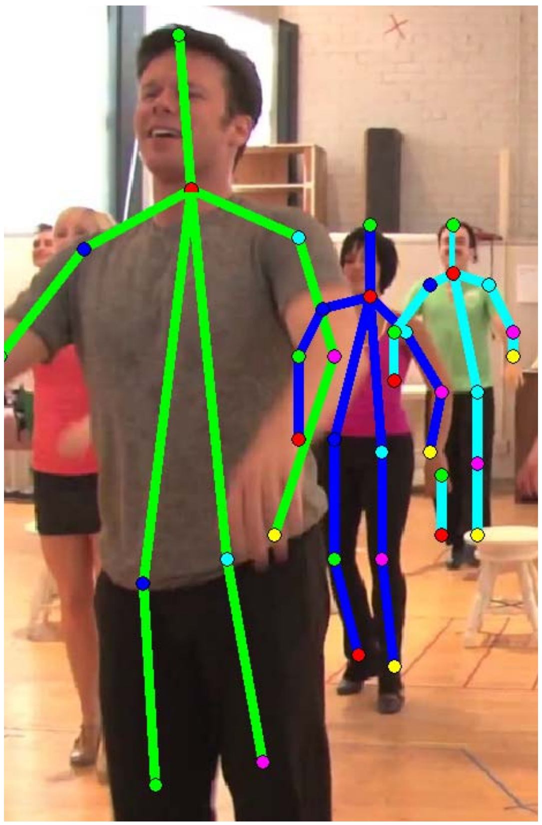
What can AI help in the interaction process?

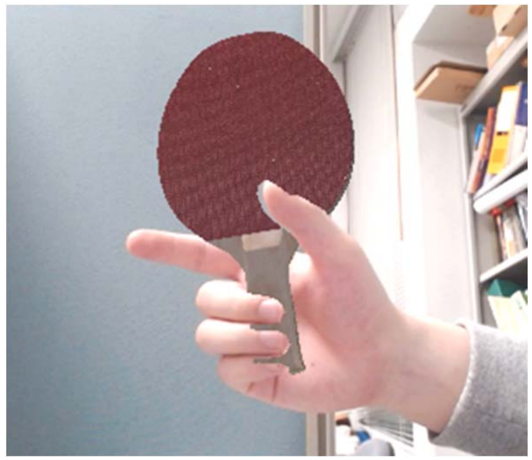


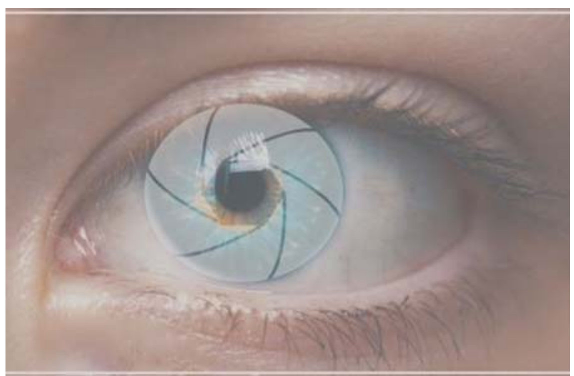
AI + Interactions

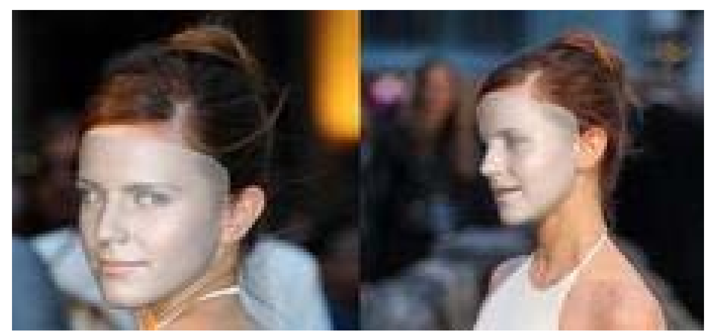
(1) What can AI help in acquiring user inputs?

Examples? Recognition, tracking, interpretation, etc.







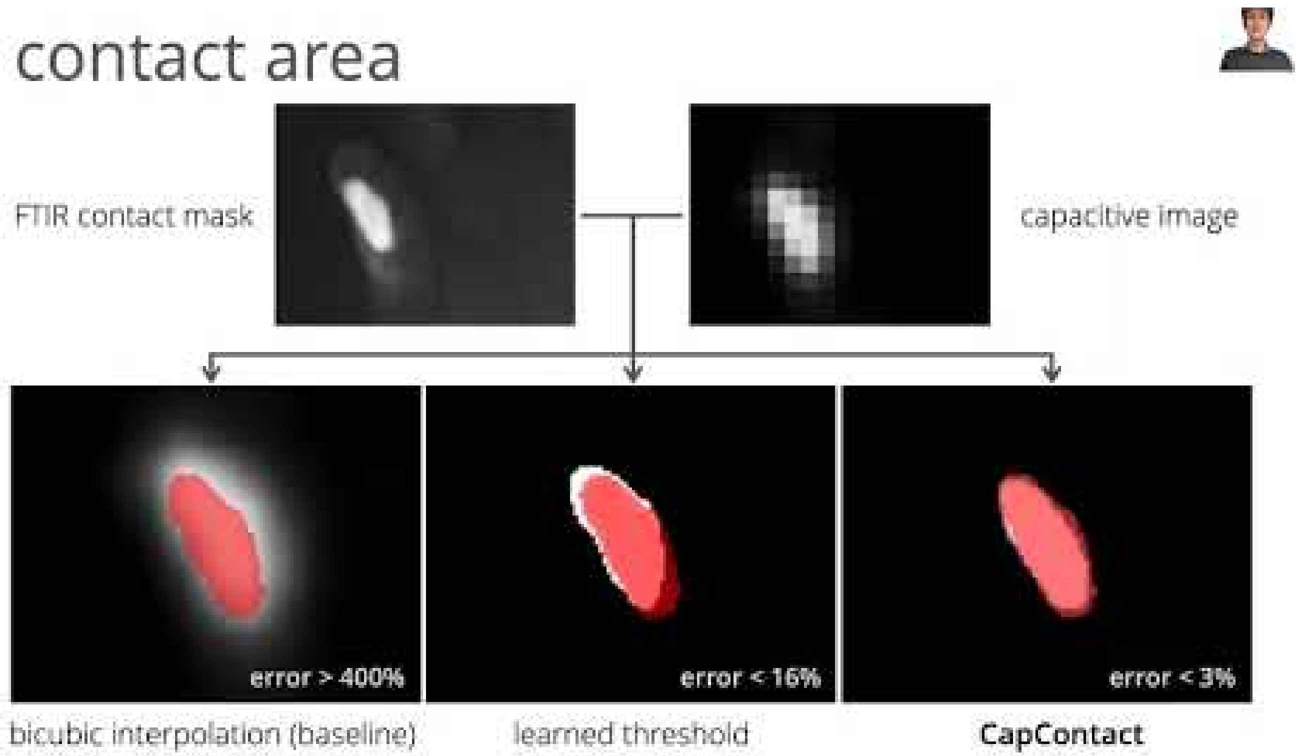


3D Face reconstruction
https://paperswithcode.com/task/3d-face-reconstruction

AI + Interactions

(2) What can AI help in enhancing user inputs?

Examples? Alleviating the fat-finger problem, etc.



Published in CHI 2021

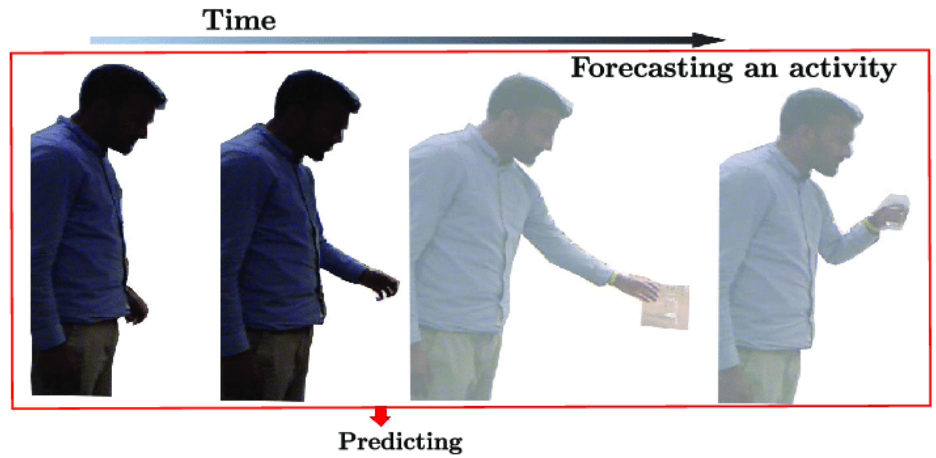
Source: https://ethz.ch/en/news-and-events/eth-news/news/2021/05/precise-touch-screens-thanks-to-ai.html

AI + Interactions

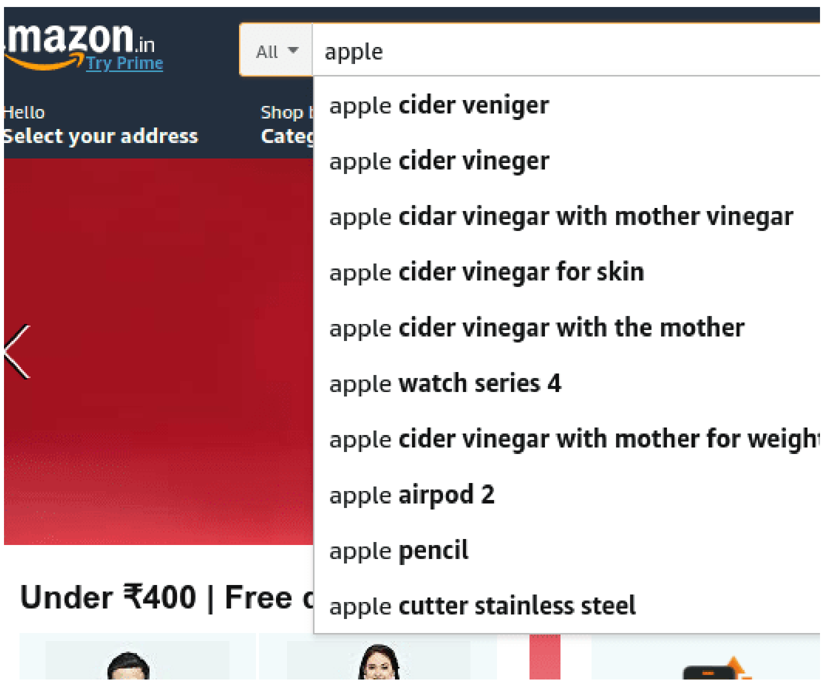
(3) AI can further help to predict user inputs?

Examples?

- Human activity prediction & motion prediction
- Autocomplete, e.g., contents-aware selection



https://www.youtube.com/watch?v=xMNYRcVH_oI



https://photoshoptrainingchannel.com/content-aware-fill-photoshop-cc-2019/

AI + Interactions

(4) AI can help to generate feedbacks to humans?

Examples?

- Synthesized human voice
- Synthesized light field for holographic displays

etc.

Summary

- Input devices
  - Keyboard and touch dominate for input
  - Fitt’s Law input model of device usage
  - Natural user interface
- Output devices
  - 2D/3D Displays and VR Displays
  - Audio and Surround Sound
  - Haptics & other feedbacks
- Ambient devices
- AI-enabled Interactions

Other Kinds of Interactions…

- Now, you know what is HCI…
- What are these interesting terms?
  - #1 HRI –
  - #2 ACI –
  - #3 SCI –

Note: don’t Google! Try to guess reasonably!