FRANKFURT UNIVERSITY OF APPLIED SCIENCES

Input and Output (I/O)

Human-Computer Interaction Lecture

Slides adapted from hci-lecture.org (A. Schmidt, N. Henze, K. Wolf, V. Schwind), Image from: https://pxhere.com/de/photo/956874

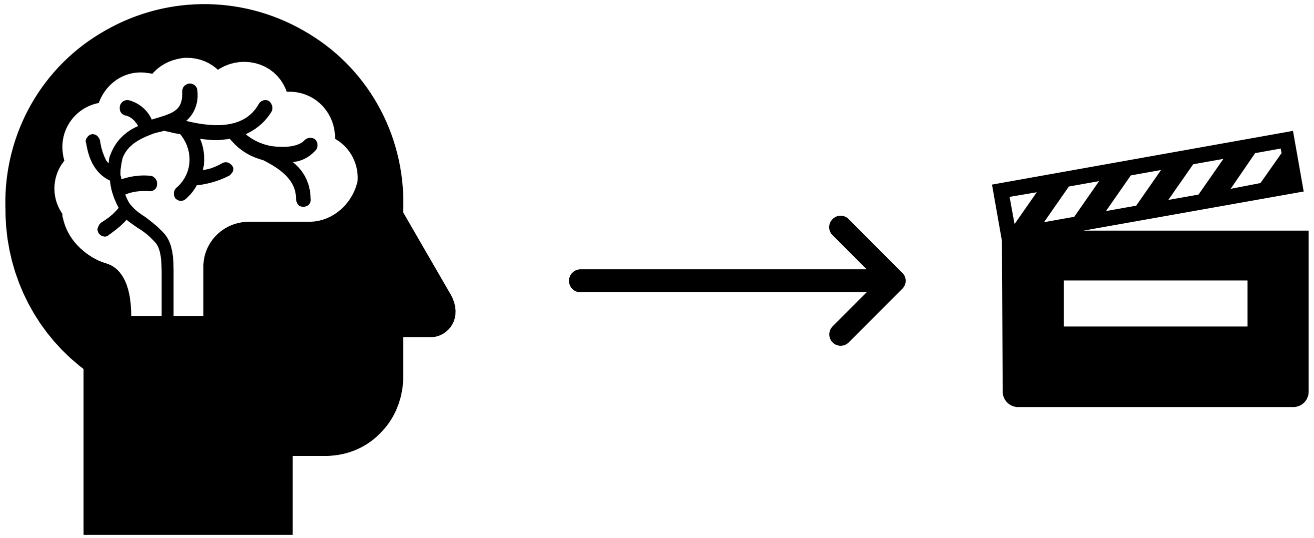
Learning Goals

- Understand …
  - The concept of bandwidth between human and system
  - The basic pointing, text entry, eye tracking, gesture and speech recognition techniques
  - The basic display and output techniques
  - What parameters can be tracked and rendered in HCI
  - The Midas touch problem and implications for user interfaces
- Know
  - How users can make input
  - How users can receive output
  - About different technologies that can be used to classify input

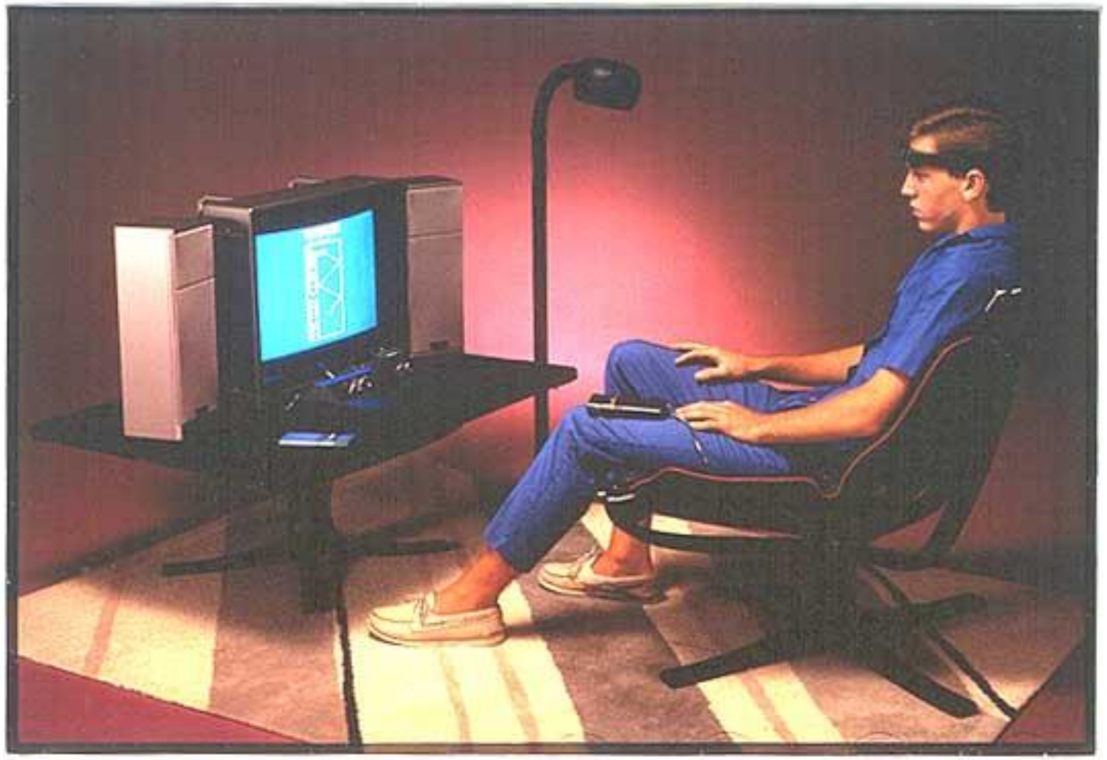
What is the ultimate user interface?

Discussion

The Perfect System



Atari Mindlink (1984, never released)



http://www.atarimuseum.com/videogames/consoles/2600/mindlink.html


esc
delete
tab
caps lock
enter
return
shift
shift
fn

Computing Systems

- A computer system for interactive use comprises various elements
  - Input
    - **speech & text entry**: traditional keyboard, phone text entry, speech and handwriting
    - **pointing**: the mouse, touchpad, stylus, etc.
    - **sensors/scanners**: real world registration, AR/VR/3D interaction devices, biosignals
    - ***memory***: recognition, contextual data, meta-data, distributed knowledge
  - Output
    - **interactive**: screens, displays, speakers (haptic, smell, sound, vision)
    - **non-interactive**: physical printers (2D, 3D)
    - ***memory***: files & folders, databases, networks

→ Memory can be both

In- and Output

Application Domain

input → output

In- and Output

Application Domain

Pointing (Gestures)
Textentry (Keyboard)
Drawing (Pen)
Direct Control
Speech, Audio
Video, Sensors
Context, Meta Data

input → **output**

Visual
Audio
Tactile
Smell
Taste
Temperature

In- and Output

Application Domain

Pointing (Gestures)
Textentry (Keyboard)
Drawing (Pen)
Direct Control
Speech, Audio
Video, Sensors
Context, Meta Data

input → **output**

bandwidth

Visual
Audio
Tactile
Smell
Taste
Temperature

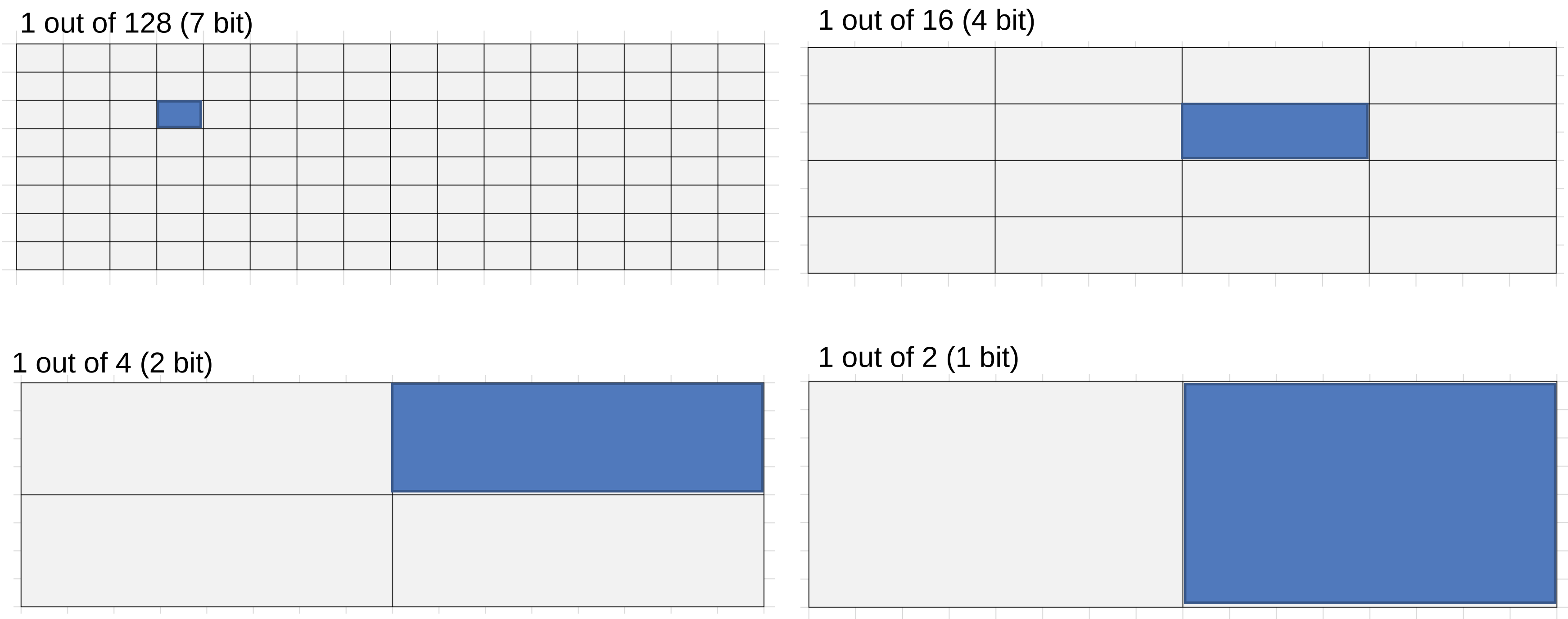
In- and Output

Application Domain

Input		Output
Pointing (Gestures) Textentry (Keyboard) Drawing (Pen) Direct Control Speech, Audio Video, Sensors Context, Meta Data	**input** ⇄ **output** bandwidth?	Visual Audio Tactile Smell Taste Temperature

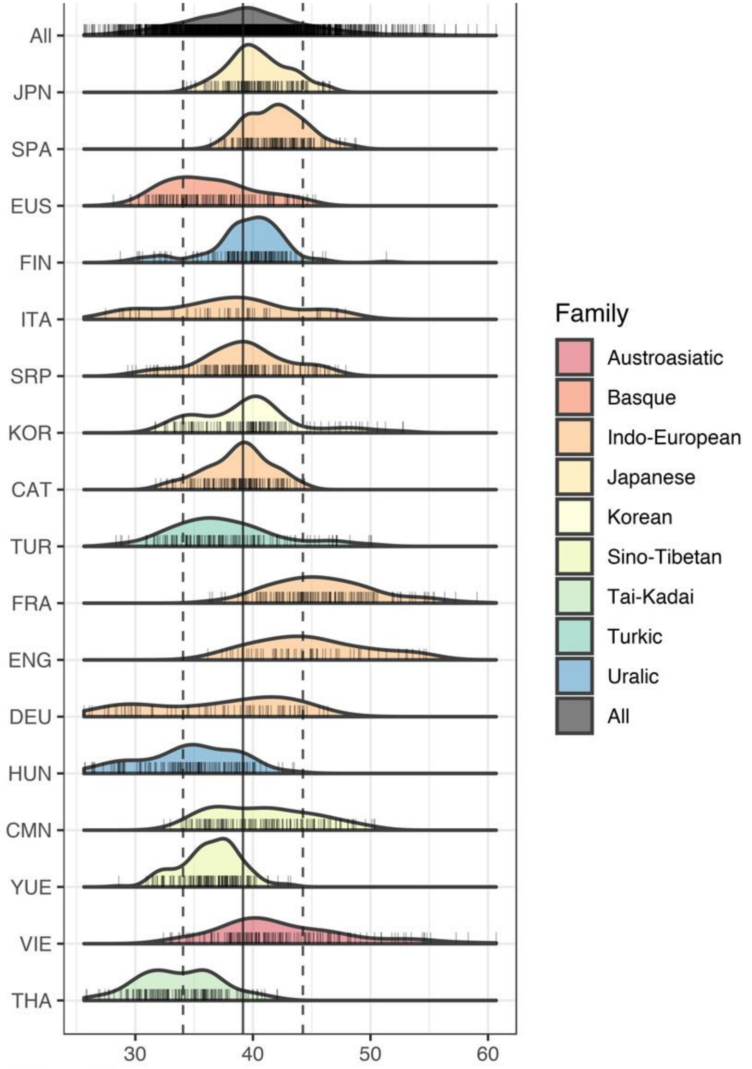
Bandwidth Simplified

- How difficult is it to click the highlighted field?
- How fast can you do it?



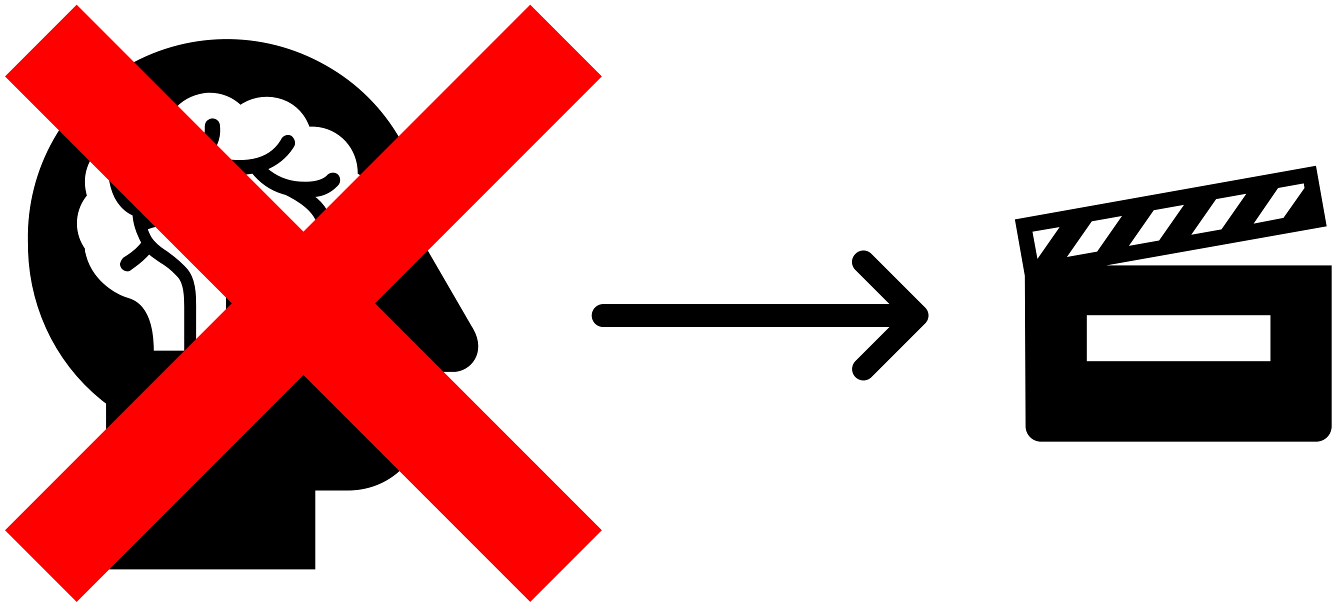
Reflections about Bandwidth

- **Communication bandwidth is information in bits/second**
- **Human input bandwidth?**
  - Seeing < 6Mbit/s (conscious visual perception may be as low as 100
  - Listening / hearing / touch / smell / taste
- **Human output bandwidth?**
  - Speech: ~ 39 bits/second
  - Writing: ~ 200 characters/minute ~ 3 characters/second < 24 bits/s
  - Pointing, thinking,…
- **Bandwidth is hart to tell!**
  - 4K TV has 8,000,000 pixels x 3-bit x 100 Hz ~ 20 Gbit/s
  - Low bandwidth for information intake (e.g., reading about 5 words/s
  - But we see/hear if things are wrong (e.g., music, movie, …)
- **Your brain is a limiting factor!**



Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, *5*(9), eaaw2594.

The Perfect System



Discussion

What input device has the highest bandwidth?

Pointing Devices

Target Selection & Referential Pointing

Image from: https://pxhere.com/de/photo/478405

Target Selection

- Central to most modern computing systems **is the ability to point and select a target** and **to manipulate the target** (by performing a function)
  - Particular in computer-aided design (CAD) where positioning and drawing are the major activities
- **Pointing devices** allow the user to **point**, **position** and **select items** either directly or by manipulating a referential pointer on the screen
  - Some pointing devices can also be used for hand drawing although the skill of drawing with a mouse is very different from using a pencil
  - The mouse is still most common for desktop computers but is facing challenges as laptop and hand-held computing increase their market share

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

The Mouse

- **Hand-held referential pointing device** that detects two-dimensional motion relative to the underlying surface
- **Mechanical**
  - The German company Telefunken published on their early ball *Rollkugelsteuerung* (RKS) mouse on 2 October 1968
- **Optical**
  - LED mice rely on one or more light-emitting diodes (LEDs) and an imaging array of photodiodes to detect movement relative to the underlying surface
  - A laser mouse is an optical mouse that uses coherent (laser) light for motion translation

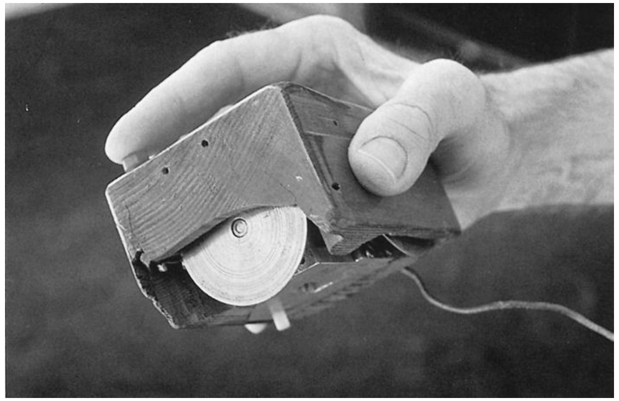


Image from Douglas Engelbart and the Bootstrap Institute

Image from: Computermuseum, Fakultät Informatik, Universität Stuttgart

Image from https://en.wikipedia.org/wiki/Optical_mouse#/media/File:Optical_mouse_shining.jpg

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

Touchpads

- Touchpads are pointing devices that use the electrical capacitance or resistance to **determine the position of an object** (e.g., a finger) **on the surface of a pad**
  - Extensively used in portable computers
  - Can now be obtained separately
- **Operated by stroking a finger** over their surface, rather like using a simulated trackball
  - May require several strokes to move the cursor
  - Require acceleration (see Transfer Functions) as the size of the touchpad does not match the size of the screen
  - Some support taps (clicks) and/or multi-finger touch

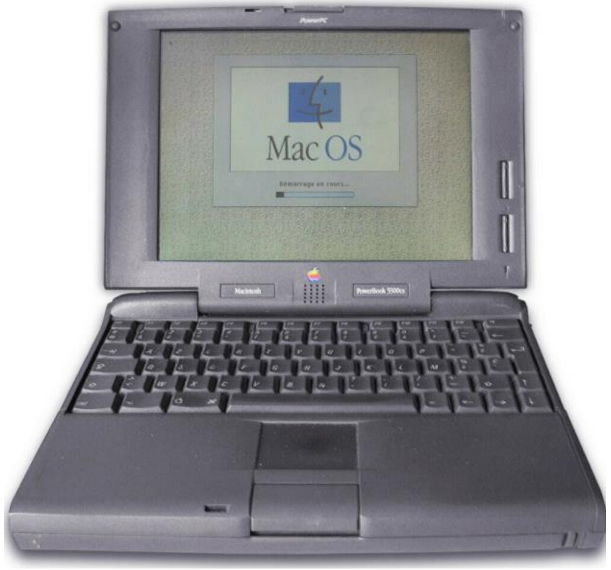


Image: Apple PowerBook 5300cs

Synaptics-Touchpad © Raimond Spekking / CC BY-SA 4.0 (via Wikimedia Commons)

Trackballs and Wheels

- The **trackball is an upside-down mouse**
  - A weighted ball faces upwards and is rotated inside a static housing, the motion being detected in the same way as with a mouse and a trackball facing downwards
- **Thumb and Scroll Wheels have dials to control the cursor position**
  - Single thumbwheels are more often included on a standard mouse in order to offer an alternative means to scroll documents (scrolling requires you to grab a small scroll bar with a cursor and drag it down)

Logitech MX Ergo Trackball Wireless Mouse

Joysticks and Trackpoints

- **Joystick** is actually a cross lever for controlling devices and vehicles and used for computers to simulate them
  - The orientation is detected either via four electrical currents, optical signal transmitters (digital), or potentiometers (analogue stick)
  - Typically used to control locomotion when the simulated representation and referential device coincident in 3D space
- **Trackpoints** (sometimes called „nipples“) measure the force exerted on it using strain gauges, which are thereby elastically deformed

Images from NilsSchneider - TrackPoint auf einer IBM Tastatur
https://de.wikipedia.org/wiki/Trackpoint#/media/Datei:Trackpoint.jpg
CC BY-SA 2.0 de, https://pxhere.com/de/photo/663324, and https://pxhere.com/de/photo/1445601

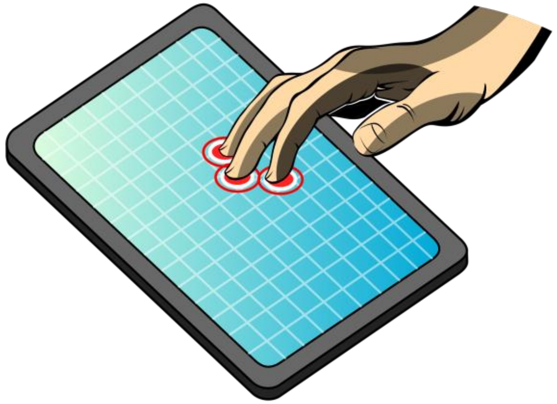
Touchscreens

- Actually, **the assembly of both an input** (the touch panel) **and output** (the display) device
  - A **resistive touchscreen** panel has one conductive layer along its sides, and another along top and bottom. When an object, such as a fingertip or stylus tip presses the upper layer down they become connected and a voltage is applied
  - A **capacitive touchscreen** consists of an insulator (glass), coated with a transparent conductor. As the human body is also an electrical conductor, touching the surface of the screen results in a distortion of the screen's electrostatic field, measurable as a change in capacita
    - Allow multi-touch technology

A virtual keyboard before iOS 7 on an iPad
matt buchanan - originally posted to Flickr as Apple iPad Event CC BY 2.0

Multi-Touch Devices

- A **multi-touch device** has a surface (a touchpad or touchscreen) to r**ecognize the presence of more than one point** of contact with the surface at the same time
  - Firstly developed in 1972 by Bent Stumpe at CERN [1] and used to develop a new type of human machine interface (HMI) for the control room of the Super Proton Synchrotron particle accelerator [2]
- **Plural-point awareness** are often be used to implement gestures, such as pinch to zoom



Multi-touch. Willtron. CC SA 1.0 (2022, September 8). In Wikipedia. https://en.wikipedia.org/wiki/Multi-touch

[1] Stumpe, Bent (16 March 1977), A new principle for x-y touch system (PDF), CERN, retrieved 2010-05-25
[2] Brian Merchant (22 June 2017). The One Device: The Secret History of the iPhone. Transworld. ISBN 978-1-4735-4254-9.

Stylus and Digital Pen

- A **stylus** is an input tool for touchscreen devices
  - A **passive stylus** acts just like any other object when touching a device resistive screen (such as a finger)
  - An **active stylus** includes electronic components that communicate with a device's touchscreen controller
    - The components increase the precision of the device and help to prevent the problem accidentally contacting the screen with fingers or hands
- A **digital pen** captures the handwriting or brush strokes of a user and converts analog information created using "pen and paper" into digital data

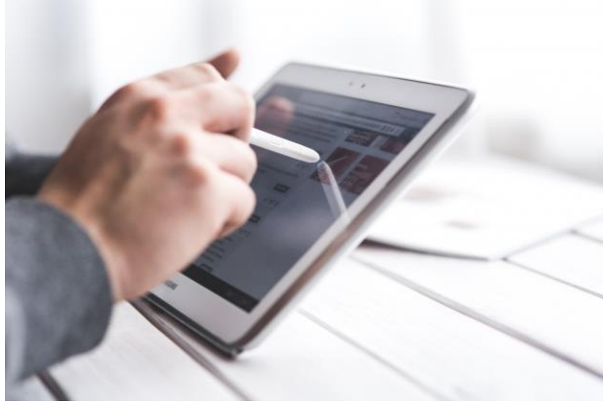


Image from
https://pxhere.com/de/photo/722543

Digital pen. (2022, September 1) Dominary - Own work
neo smartpen captures what is written on paper CC BY-SA 4.0. In *Wikipedia*. https://en.wikipedia.org/wiki/Digital_pen

Digitizing Tablet

- The first electronic handwriting device was the Telautograph, patented by Elisha Gray in 1888 [1].

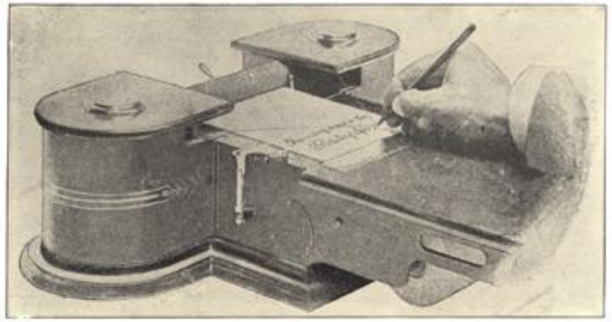


Fig. 34. THE TELAUTOGRAPH—TRANSMITTING INSTRUMENT.

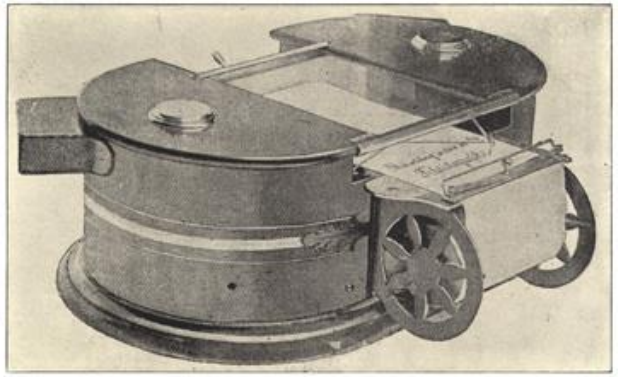


Fig. 35. THE TELAUTOGRAPH—RECEIVING INSTRUMENT.

- A **digitizer tablets are** referential pointing devices (also called graphic tablets, pen tablets, drawing tablets, or digital art boards) enables a user to hand-draw images, animations and graphics on a surface and with a special pen-like stylus
  - No screen but often support pressure
  - Work similar to touchscreens (actively and passively)

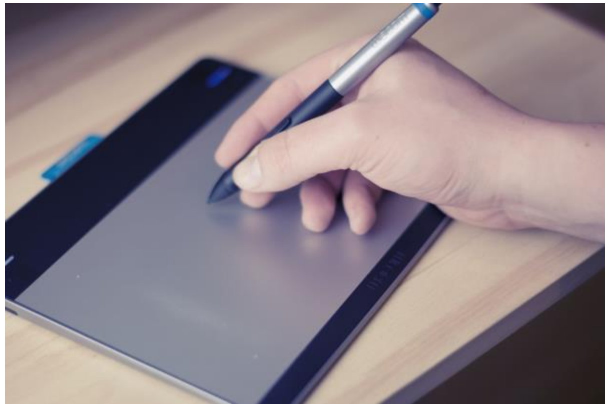


Image from https://pxhere.com/de/photo/1083778

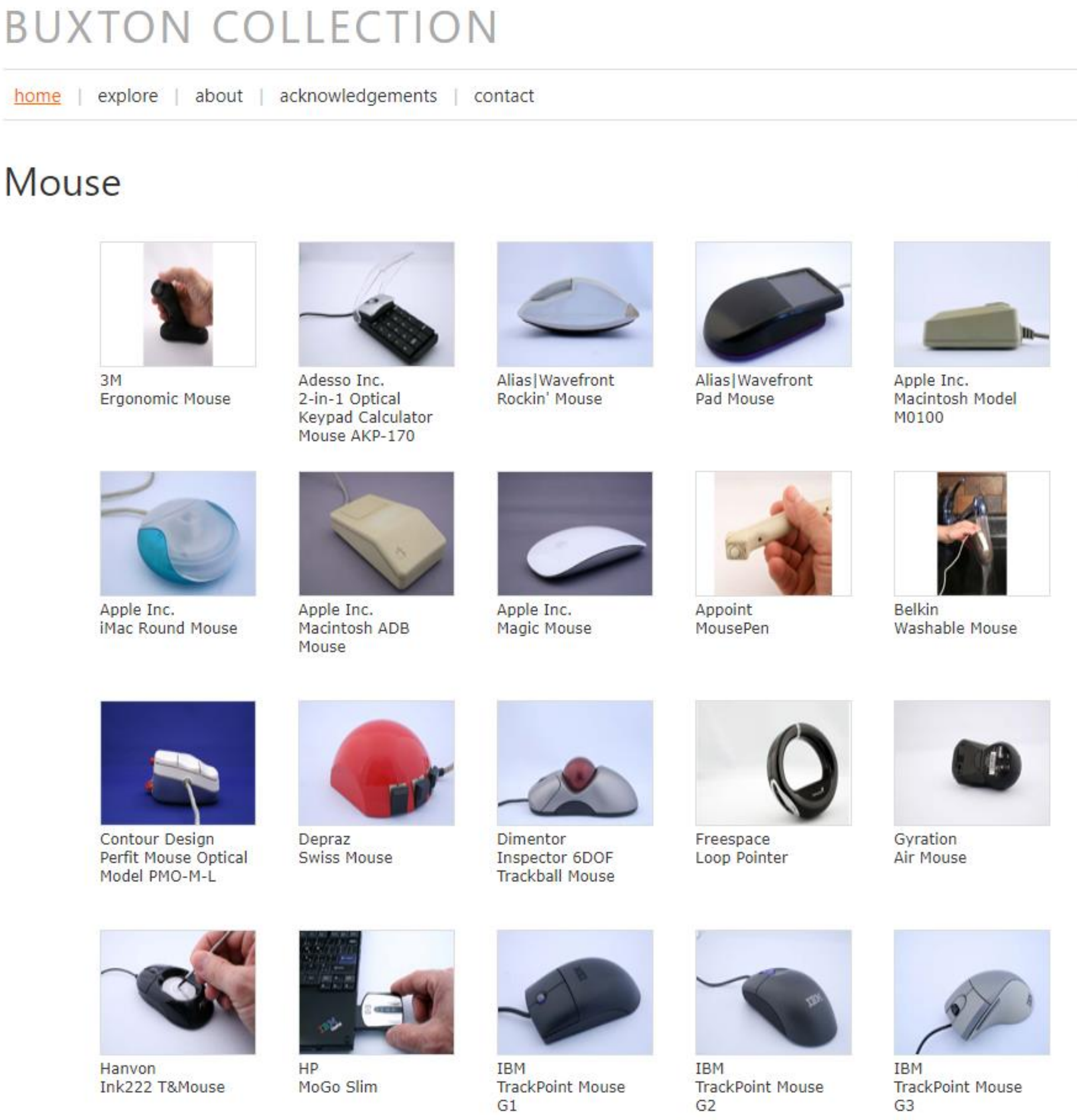
[1] Gray, Elisha (1888-07-31), Telautograph, United States Patent 386,815

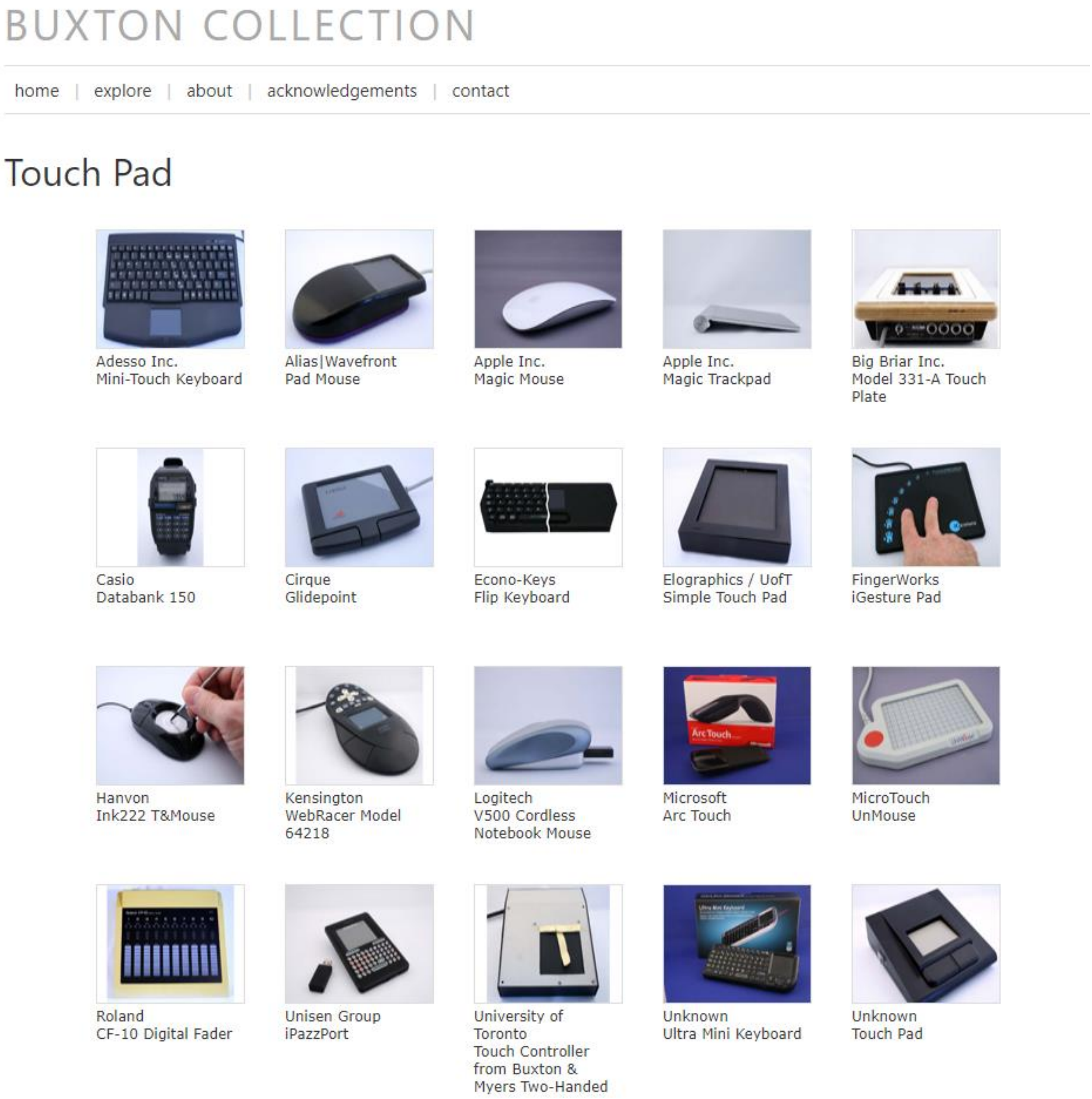
Cursors Key

- **Cursor keys** are mainly used **not for pointing but positioning**
  - Fast for moving within discrete text/cells using up/down left/right in one direction
- The **WASD key combination** brings the cursor keys to the left side of the desktop and allow 2D position changes with the left hand, while the right hand can still be used for right-hand target selection (target aiming).
  - John Carmack integrated WASD as an alternate layout in Quake II after a player called Dennis Fong had big success with WASD in tournaments 1997.

Image from
https://pxhere.com/de/photo/1256374

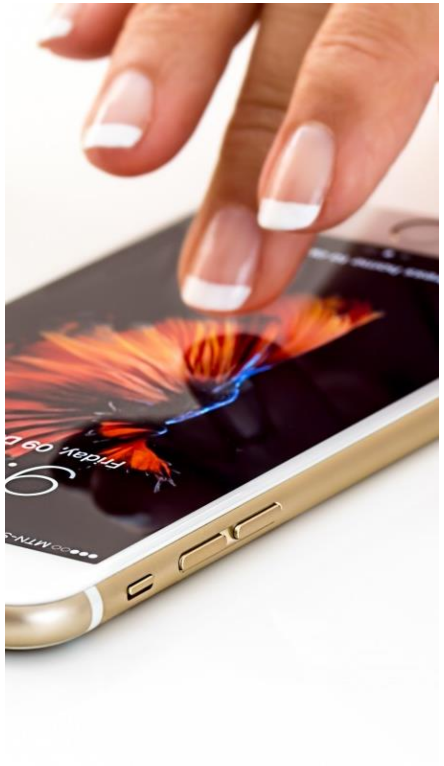
Image from
https://pxhere.com/de/photo/414231

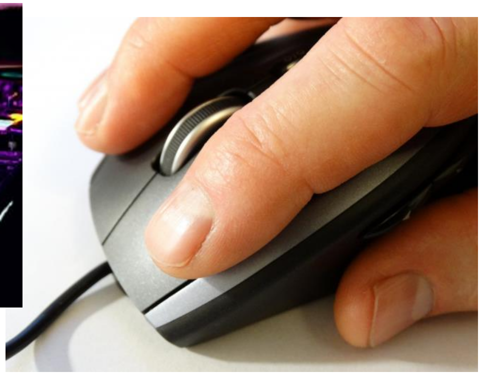


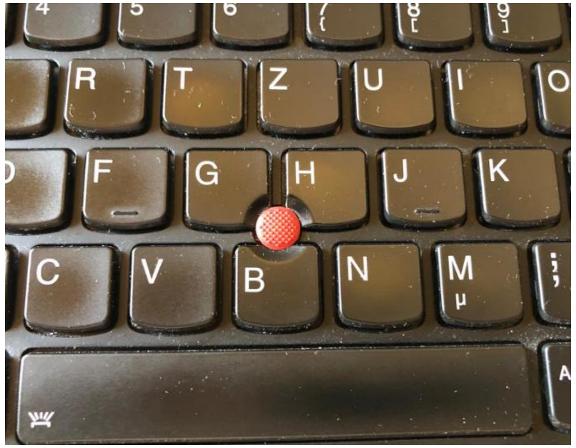


https://www.microsoft.com/buxtoncollection/

More Pointing Devices







Which Manipulation Operator?

- Example: mapping a rotary controller to linear movement

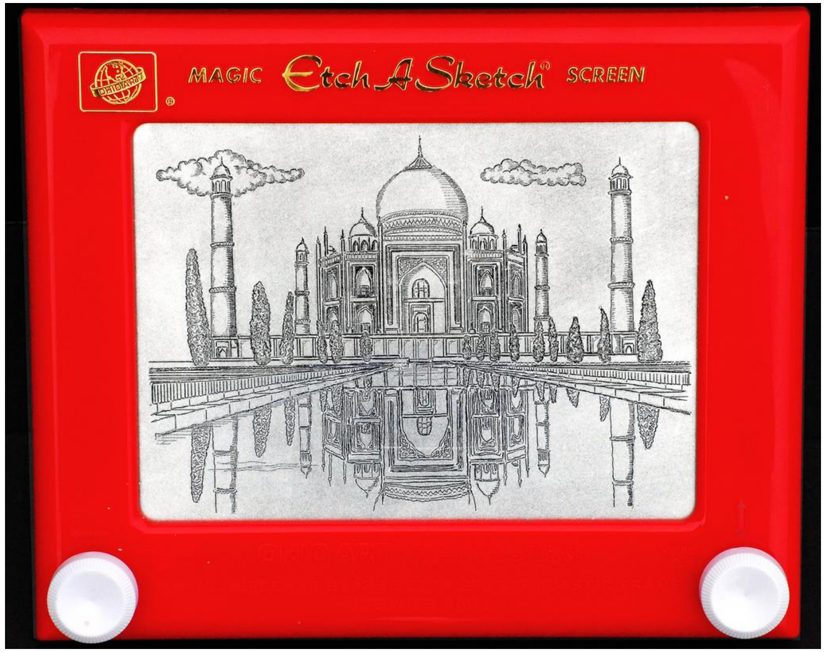


Foto by: Etcha (CC BY-SA)
https://commons.wikimedia.org/wiki/File:Taj_Mahal_drawing_on_an_Etch-A-Sketch.jpg

Foto by: casers jean (CC BY 2.0)
https://www.flickr.com/photos/casers/125482678

So many devices. Can we classify them?

- **Criteria** to assess input devices
  - **continuous vs discrete**?
  - **agent of control** (hand, foot, voice, eyes ...)?
- **Dimensions in the Taxonomy**
  - **what is being sensed** (position, motion or pressure), and
  - the **number of dimensions** being sensed (1, 2 or 3)
  - **motor skills** to operate (similar motor skills are in sub-columns)
  - **touch vs. mechanical intermediary** (directly touched vs devices that require a mechanical intermediary between the hand and the sensing mechanism (sub-rows))

A Taxonomy of Pointing Devices

- **Degrees of Freedom** (DOF) / Dimensions
  - 2 DOF, 6 DOF
  - 1D / 2D / 3D
- **Direct vs. Indirect Control** (integration with the visual representation)
  - Touch screen is direct
  - Mouse, trackpad, trackball are indirect
- **Discrete vs. Continuous Sensing** (resolution of the sensing)
  - Touch screen is discrete
  - Mouse, trackpad, trackball are continuous
- **Absolute vs. Relative Transfer** (movement/position used as input)
  - Touch screen is absolute
  - Mouse is relative

Degrees of Freedom (DOF)

- **1 DOF**
  - Slider or Knob to control the volue
- **2 DOF**
  - Mouse - you can move it in X and Y
  - Touchpad – you move on it in X and Y
- **3 DOF**
  - Mobile phone to rotate an object on the screen
  - Hypothetical device: a mouse that also registers rotation
- **6 DOF**
  - Input devices that can control translation (x, y, z) as well as rotation (pitch, yaw, roll) in 3D space

Katzakis, N., & Hori, M. (2010, March). Mobile devices as multi-DOF controllers. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)* (pp. 139-140). IEEE.

Controller Resistance

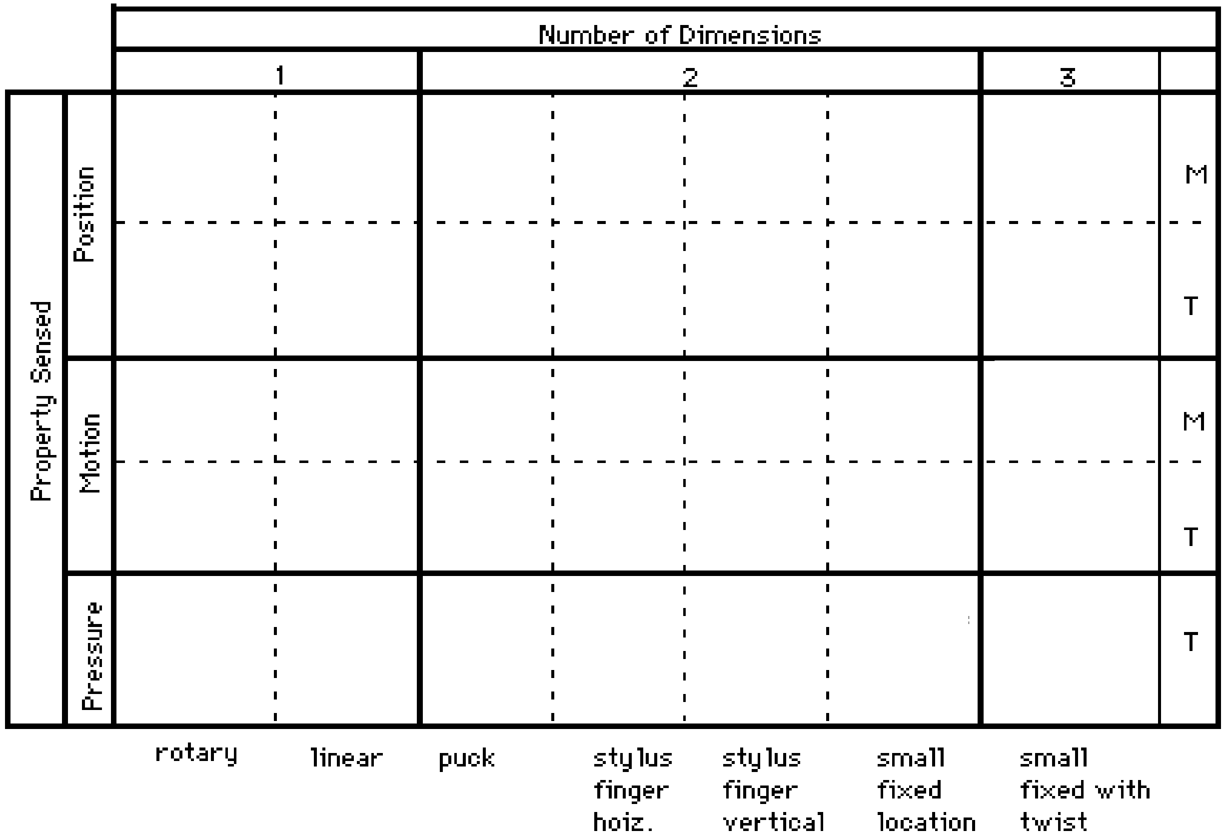
- **Isometric** (infinite resistance)
  - Device/handle is not moved
  - Pressure devices / force devices
  - Infinite resistance
  - Device that senses force but does not perceptibly move
  - Force is mapped to rate control of the cursor (typical) or to absolute position
- **Isotonic** (free moving)
  - Device/handle is moved
  - Displacement devices, free moving devices or unloaded devices
  - Zero or constant resistance (resistance stays the same)
  - Displacement of device is mapped to displacement of the cursor



Controller Resistance

- **Elastic**
  - Device/handle is moved
  - Device’s resistive force increases with displacement (also called spring-loaded)
  - Device can sense displacement or force
  - Force/displacement is mapped to rate control of the cursor (typical) or to absolute position
- **Viscous**
  - Resistance increases with velocity of movement
- **Inertial**
  - Resistance increases with acceleration

A Taxonomy of Pointing Devices



Buxton, W. (1983). Lexical and Pragmatic Considerations of Input Structures. Computer Graphics, 17 (1), 31-37. http://www.billbuxton.com/lexical.html
http://www.billbuxton.com/input04.Taxonomies.pdf

A Taxonomy of Pointing Devices

		Number of Dimensions							
		1		2				3	
Property Sensed	Position	Rotary Pot	Sliding Pot	Tablet & Puck	Tablet & Stylus	Light Pen	Isotonic Joystick	3D Joystick	M
					Touch Tablet	Touch Screen			T
	Motion	Continuous Rotary Pot	Treadmill	Mouse			Sprung Joystick Trackball	3D Trackball	M
			Ferinstat				X/Y Pad		T
	Pressure	Torque Sensor					Isometric Joystick		T
		rotary	linear	puck	stylus finger hoiz.	stylus finger vertical	small fixed location	small fixed with twist	

Buxton, W. (1983). Lexical and Pragmatic Considerations of Input Structures. Computer Graphics, 17 (1), 31-37. http://www.billbuxton.com/lexical.html
http://www.billbuxton.com/input04.Taxonomies.pdf

Physical Properties of Input Devices

	Linear	Rotary
Position		
Absolute	P (Position)	R (Rotation)
Relative	dP	dR
Force		
Absolute	F (Force)	T (Torque)
Relative	dF	dT

Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

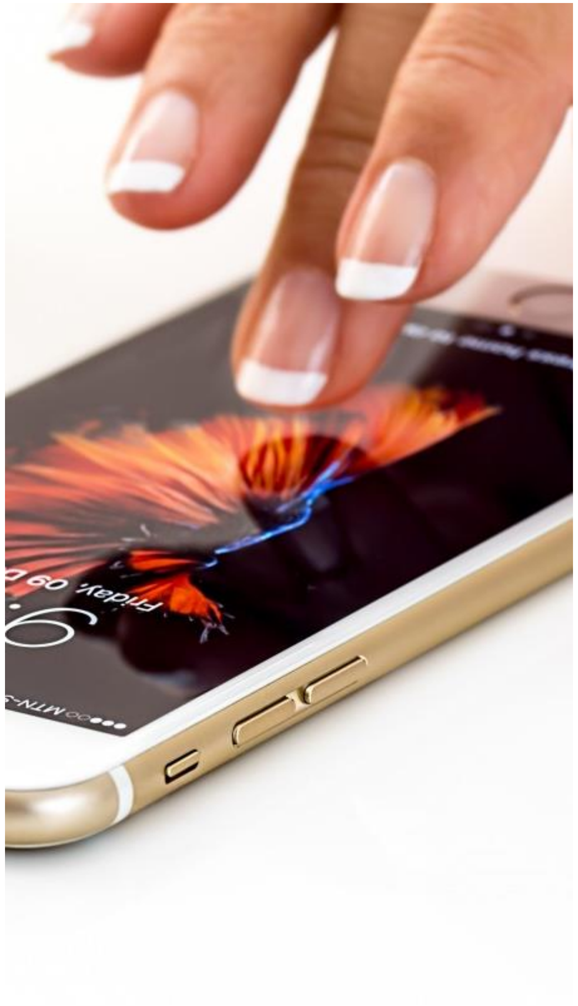
Design Space of Input Devices

	Linear			Rotary			
	X	Y	Z	rX	rY	rZ	
P							R
dP							dR
F							T
dF							dT
	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	

Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

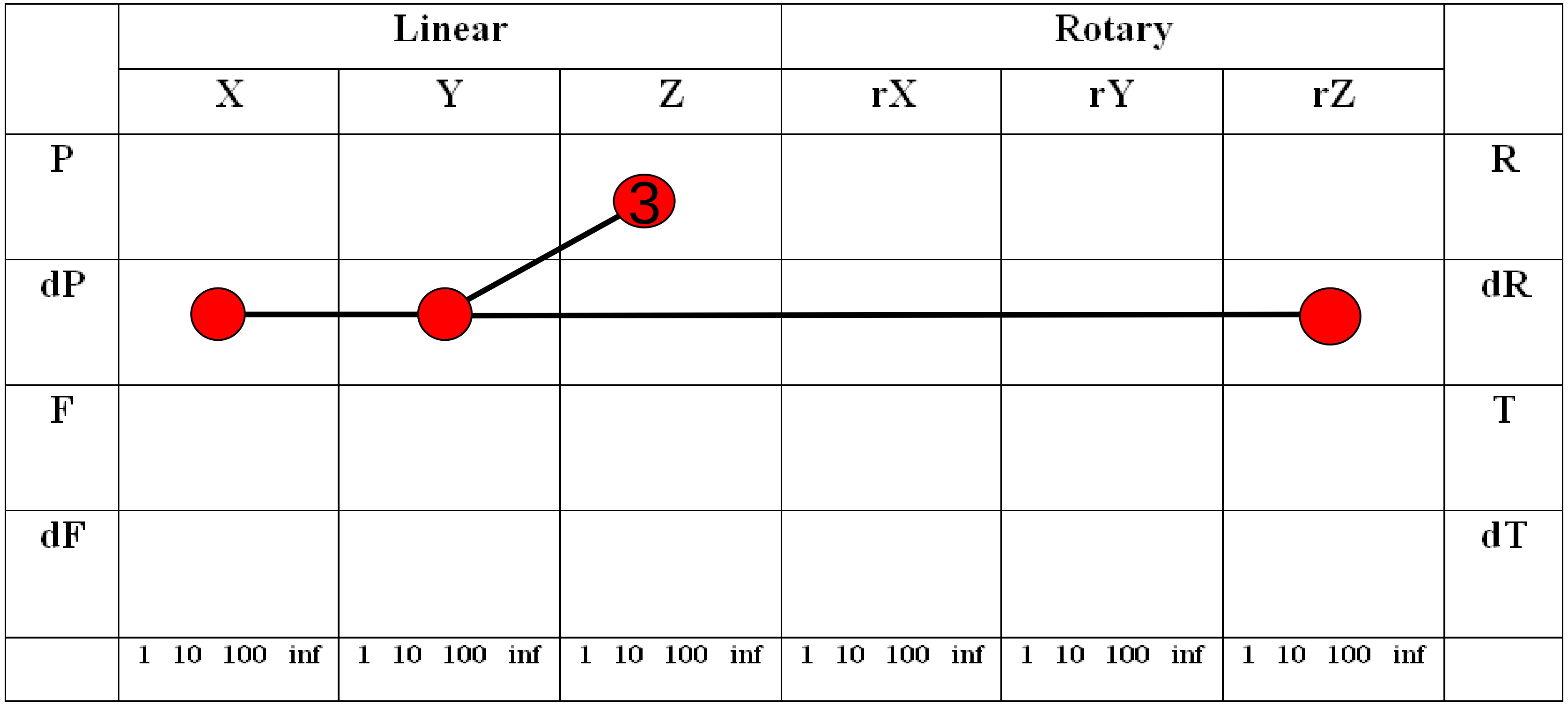
Design Space of Input Devices

	Linear			Rotary			
	X	Y	Z	rX	rY	rZ	
P							R
dP							dR
F							T
dF							dT
	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	



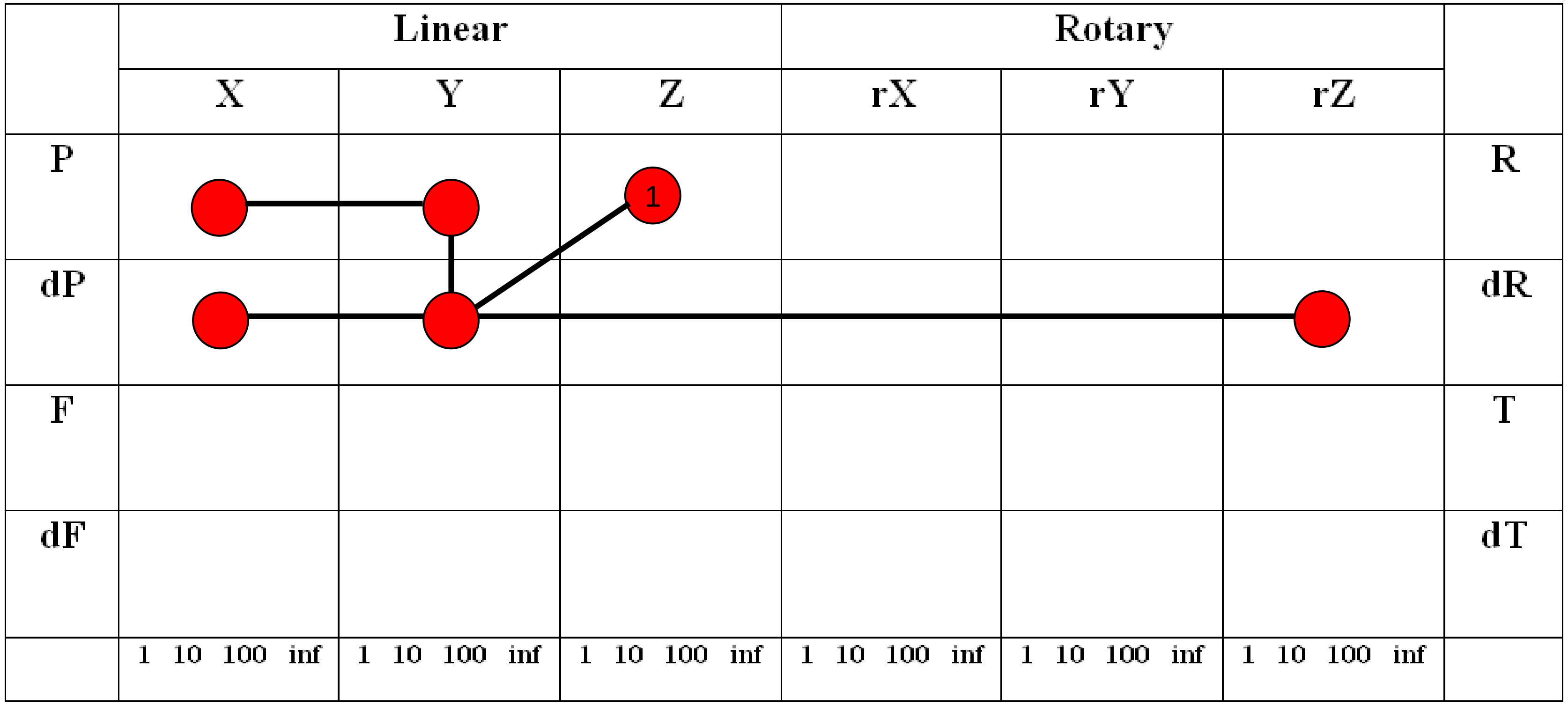
Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

Design Space of Input Devices



Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

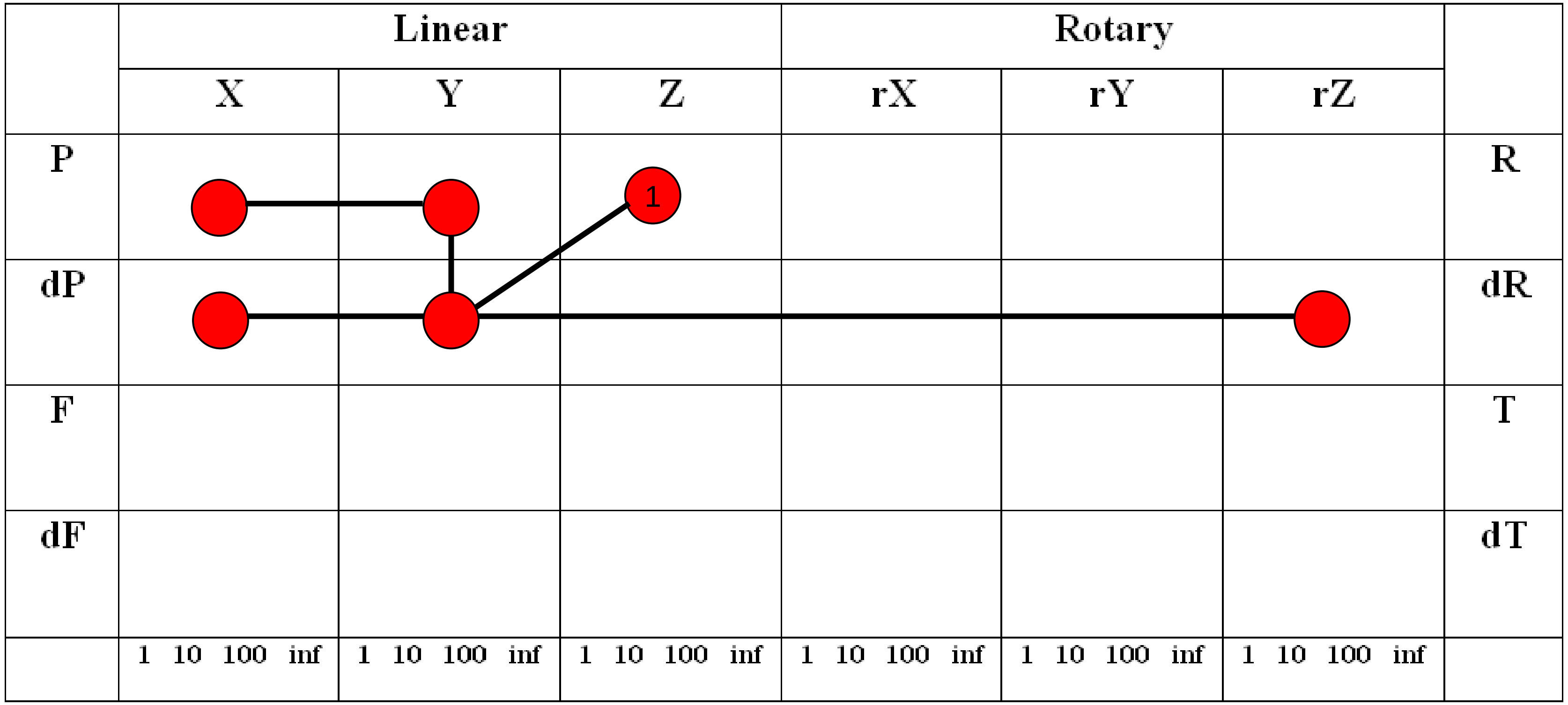
Design Space of Input Devices: Mini Exercise



?

Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

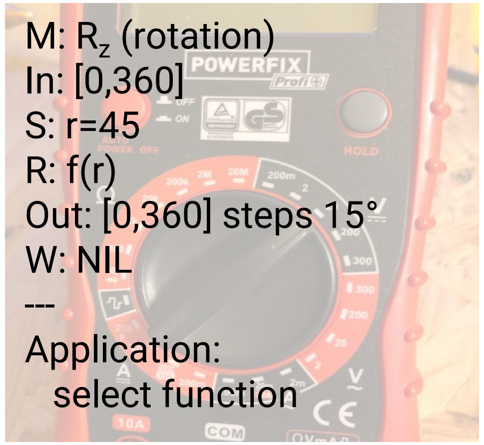
Design Space of Input Devices: Mini Exercise

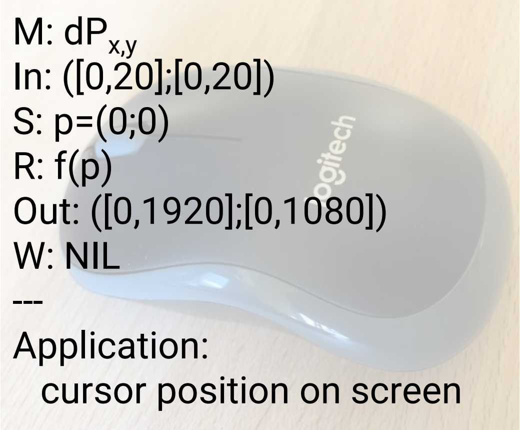


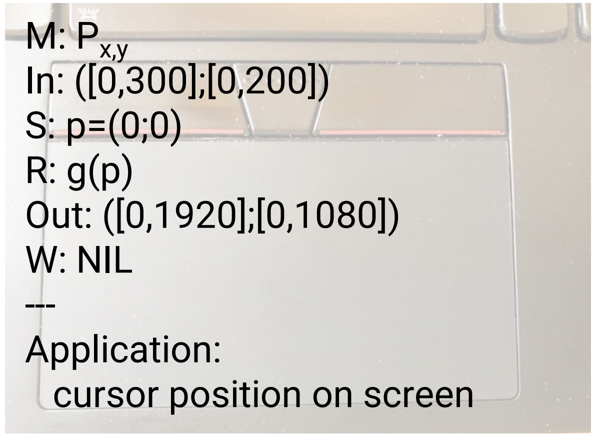
Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726

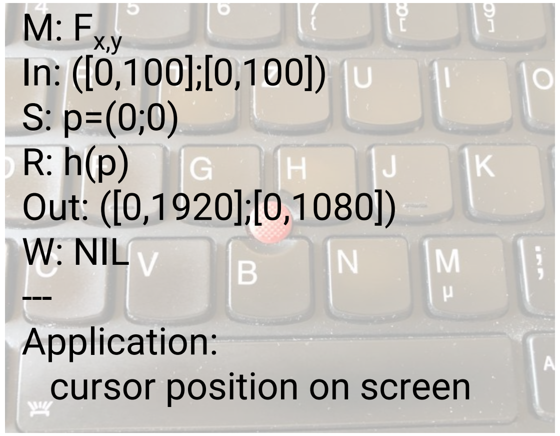
Input Devices as a Six-Tuple

- **M** is the manipulation operator
- **In** is the input domain
- **S** is the current state of the device
- **R** is a resolution function mapping from the input to the output domain
- **Out** is the output domain set







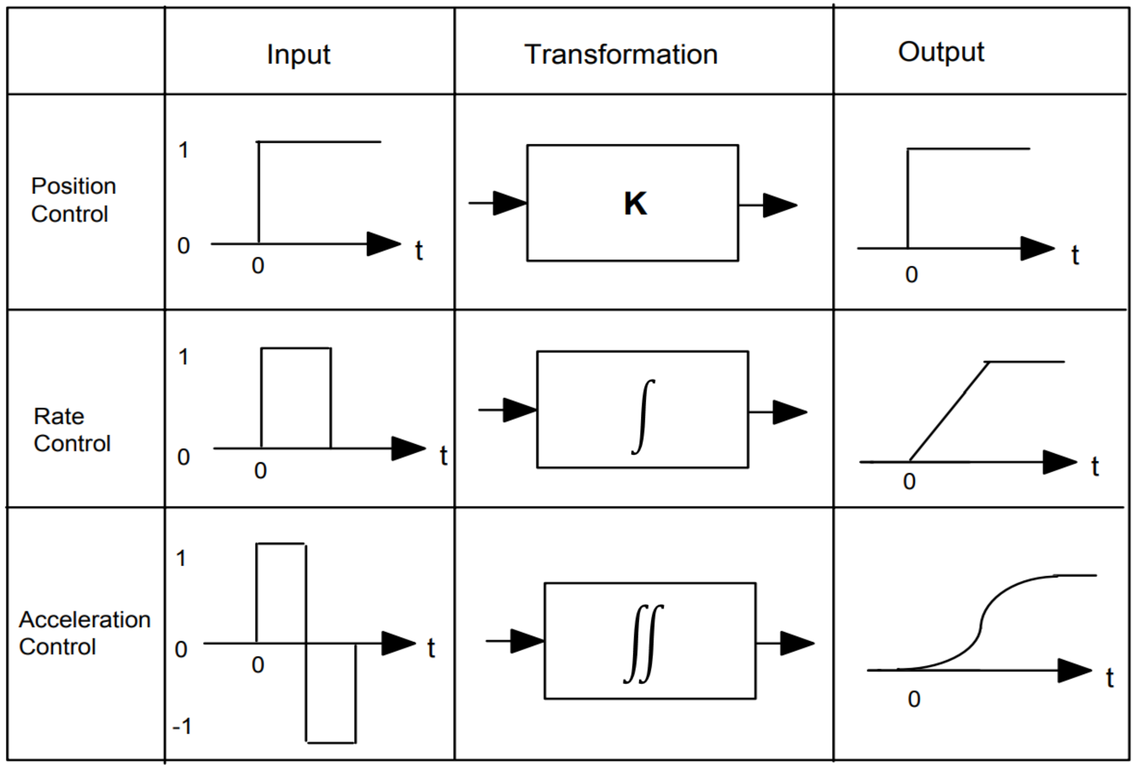


Transfer Functions

- A referential pointing device is **moved across a surface in X and Y**
  - The device reports the movement as changes to x and y (mostly dx and dy)
  - The resolution of the updates relate to the resolution of the sensor in the mouse (DPI)
- How often the changes are reported relates to the **polling rate**
  - e.g., 100Hz means you get an update every 10 ms, 500Hz gives an update every 2 ms
- We assume your cursor is at a certain position on the screen (Sx, Sy)
  - Your transfer function updates the screen position of the cursor based on the received values:
    - **(Sx,Sy) = f(Sx, Sy, dx, dy)**

Transfer Functions

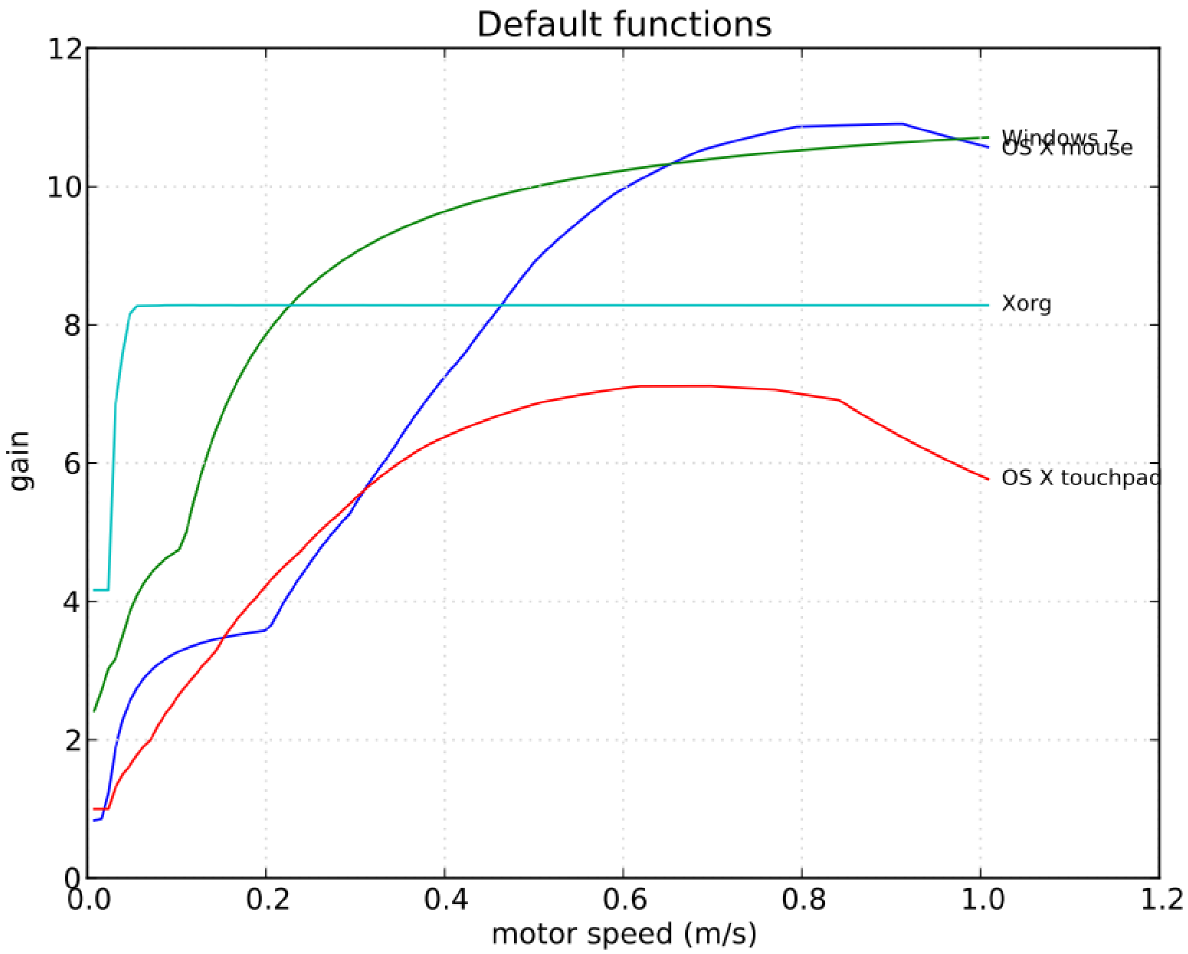
- **Position control** to map translation
  - device displacement
  - absolute force
- **Rate control** to cursor velocity
  - force
  - displacement
- **Acceleration control** of the cursor
  - force
  - displacement



Zhai, Shumin. Human performance in six degree of freedom input control. PhD Thesis. University of Toronto, 1996. p18 https://www.talisman.org/~erlkonig/misc/shumin-zhai%5Ehuman-perf-w-6dof-control.pdf

Transfer Functions

- The **physical movement** of the pointing device **is translated into the movement of a UI cursor**
  - There are different functions, default values, and parameters for different operating systems (mainly contribute to the look-and-feel of the OS)
  - There are also ideas for user-defined and adaptive transfer functions

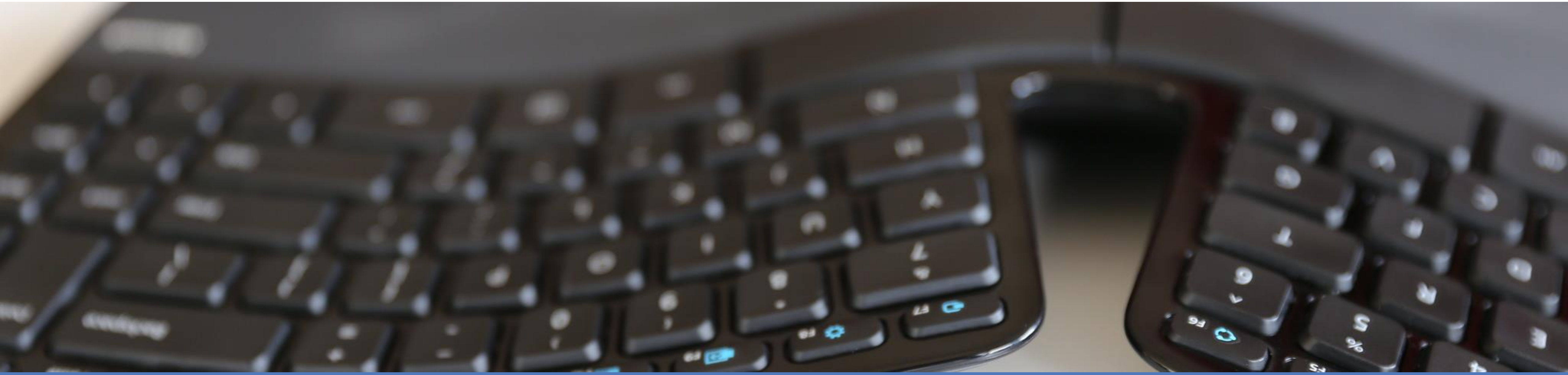


Géry Casiez and Nicolas Roussel. No more bricolage!: methods and tools to characterize, rep
symposium on User interface software and technology, pages 603–614. ACM, 2011.

Gestures

Human-Computer Interaction Lecture

Image from: https://pixabay.com/de/photos/blondine-m%C3%A4dchen-jung-frau-person-768681/



Text Entry

Keyboards, Handwriting & Speech

Slides adapted from hci-lecture.org (A. Schmidt, N. Henze, K. Wolf, V. Schwind), Image from: https://pxhere.com/de/photo/956874

Keyboards

- Are used to enter **textual signs** for human-readable language
  - depend on the language
- A **keypress** on a board **closes a current connection**, causing a character code to be sent to the computer
  - Training required, but people can achieve high input rates
- One aspect of keyboards that is important to users is the **‘feel’ of the keys**
  - Some keyboards require a very hard press, whilst others are lighter
  - The distance that the keys travel also affects the tactile nature of the keyboard
  - Most notebook computers are ‘half-travel’ keyboards

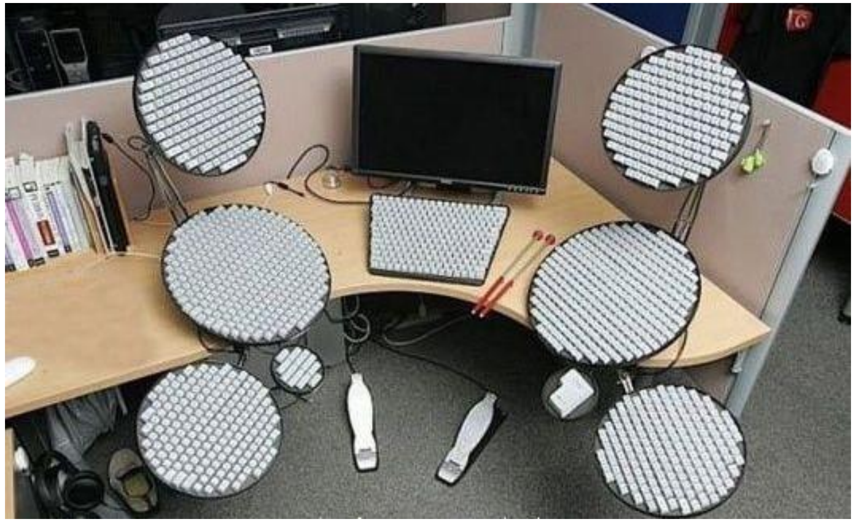
Alphanumeric Keyboard

- **Emerged from mechanical typewriters**
  - Hitting a key caused a mechanical letter arm to shoot towards the carriage
  - If two arms flew towards the paper quickly, they often jammed → QWERTY reduced to probability of jamming
- Electric typewriters and now the computer keyboard are not subject to the original mechanical constraints, **but the QWERTY keyboard remains the dominant layout**
  - Social reasons: the vast base of trained typists would be reluctant to relearn their craft

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

Dayi Method and Chinese Character Input

- A chinese character is built by combining up to four of 46 characters
- Dayi is **a system for entering Chinese characters on a standard QWERTY** keyboard using the set of 46 character components
  - needs the correct sequence of keys to enter
  - Chinese writing stroke order: "top first, then bottom", "left first, then right".
    - e.g., 壽 → 士乛工舟 (FBR.)



A Chinese keyboard with Dayi hints by Cangjie6 CC BY-SA 4.0 https://en.wikipedia.org/wiki/Dayi_method
https://www.digmandarin.com/how-to-type-in-chinese.html

The DVORAK Keyboard

- The **DVORAK keyboard** (1930) uses a similar layout of keys to the QWERTY system, but assigns the letters to different keys
- The keyboard is designed to help people reach faster typing speeds (later studies are inconclusive, today studies show no advantage)
  - Biased towards right-handed people
  - The layout of the keys also attempts to ensure that the majority of keystrokes alternate between hands, thereby increasing the potential speed
  - Supposed to reduce repetitive strain injury (RSI)

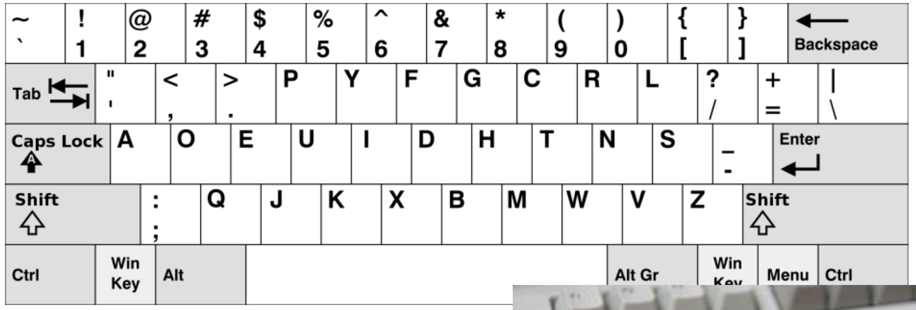
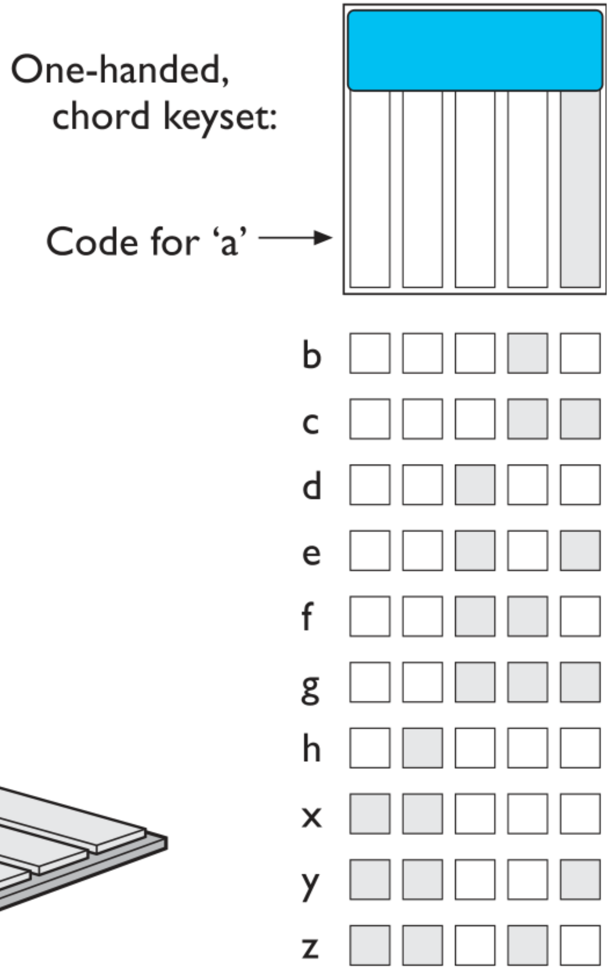
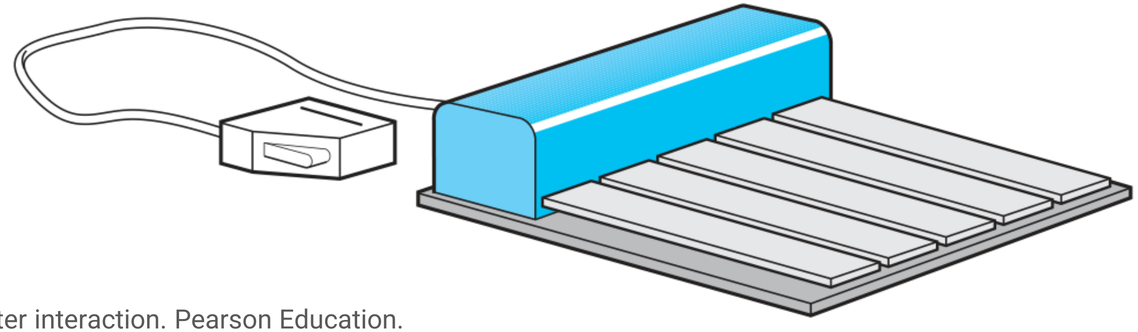


Image from Nayu Kim - https://www.flickr.com/photos/41506934@N03/4039885305/in/photolist-79Zt9V
IBM Model M keyboard in Dvorak CC BY 2.0

Chord Keyboards

- **Emerged by stenographers** (two handed)
  - Incredible typing speeds with trained users
- Takes some days for training
- Letters (and even words) can be produced by pressing one or more keys at once
- Extremely compact, fast, and suitable for one-handed operations
- Social resistance





Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

Braille Keyboards

Braille Keyboard Cover (with keyboard) Brand: A2i, Image from https://www.amazon.co.uk/Braille-Keyboard-Cover-with-keyboard/dp/B008OI60P4

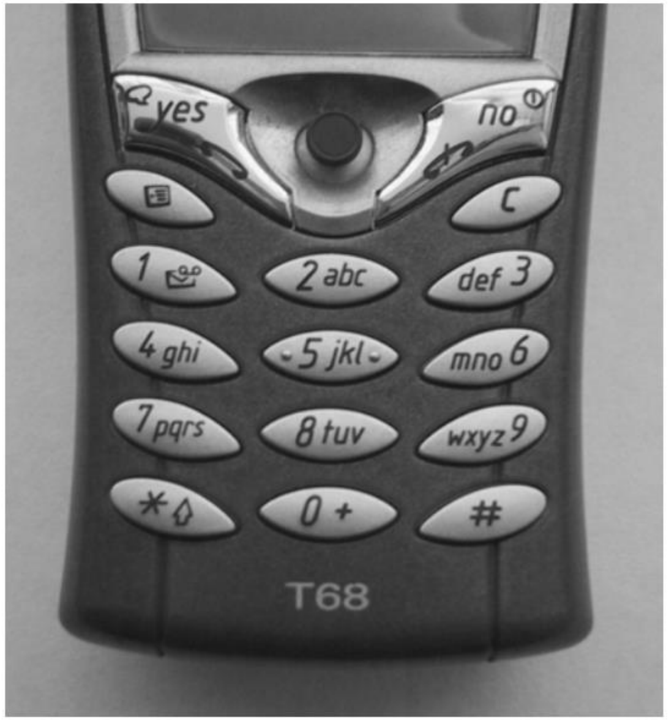
ATM keypad with braille. Redspotted from London, UK - ATM keypad 2/4 CC BY 2.0 https://commons.wikimedia.org/wiki/File:ATM_keypad_with_braille.jpg#/media/File:ATM_keypad_with_braille.jpg

Chord Braille computer keyboard, DASA Museum, Germany. Marcin Wichary - https://www.flickr.com/photos/mwichary/2251355897/sizes/l/in/photostream/ CC BY 2.0

Phone Pad and T9

- With mobile phones used for SMS messaging (see and WAP, the phone keypad has become an important form of text input
- A phone only has digits 0–9, not a full alphanumeric keyboard.
  - For text input the numeric keys are usually pressed several times
  - The T9 algorithm uses a dictionary to disambiguate words by simply typing the relevant letters once

  → '3926753' becomes 'example'

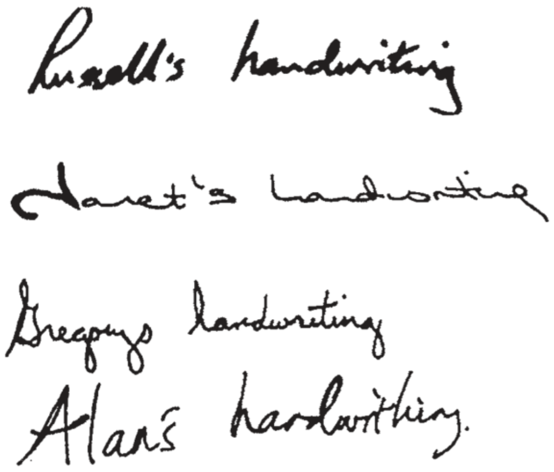


Typical key mapping:
1 – space, comma, etc. (varies)
2 – a b c
3 – d e f
4 – g h i
5 – j k l
6 – m n o
7 – p q r s
8 – t u v
9 – w x y z
0 – +, &, etc.

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

Handwriting Recognition

- We consider handwriting (and speech) as very **natural forms of communication** and text input
  - Recognition problem: the variation between the handwriting of individuals is large and letters are written differently according to the preceding and successive ones
    - No systems in use today that are good at cursive script recognition



cursive script recognition

  - When letters are individually written (small separation) the success of handwriting recognition becomes better but the user must adapt

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). Human-computer interaction. Pearson Education.

Speech Recognition

- Computational linguistics develop methodologies and technologies that enable **the recognition and translation of spoken language** into text by computer with **voice-based interfaces** and mostly require or support also speech synthesis
  - automatic speech recognition (ASR)
  - speech to text (STT)
- Most ASR systems today are **based on neural networks** and require extensive training for different language models
  - Attention-based models such as, „Listen, Attend, and Spell“ (LAS), do not have conditional-independence assumptions and can learn all the components of a speech recognizer including the pronunciation, acoustic and language model directly (end-to-end)

Chorowski, Jan; Jaitly, Navdeep (8 December 2016). "Towards better decoding and language model integration in sequence to sequence models". arXiv:1612.02695 [cs.NE].

Natural Language Processing (NLP)

- NLP combines computational linguistics—rule-based modeling of human language—with statistical, machine learning, and deep learning models.
- There are four key objectives in NLP:
  - **Speech Recognition**: Converting spoken language into text (text-entry)
  - **Language Understanding**: Involves parsing and understanding human language in a way that a computer can process.
    - Sentiment analysis, language detection, and contextual understanding
  - **Language Generation**: Generating human-like text based on input data
    - Text abstraction, text summarization, language translation, and content generation
  - **Text-to-Speech**: Speech synthesis and converting written text into spoken words
    - Virtual assistants and read-aloud features

Gesture and Activity Recognition

Sensors & Context Awareness

Image from: https://pxhere.com/de/photo/1058439

Gestures

- Gestures are essential part of **non-verbal communication** between humans and supports people in transporting **commands**, **metaphors**, **emotions**, and **rhythm**.
- **Physical movements** by a user …
  - must be **detected** using **a device**
  - must be **interpreted** using **a classifier**
- Different classes of gestures [1,2]:
  - **Online**/**Continuous** manipulations like moving, scaling, rotating (e.g., pointing on a screen)
  - **Offline**/**Discrete** events processed after the interaction has completed (e.g., pinching to open a menu)

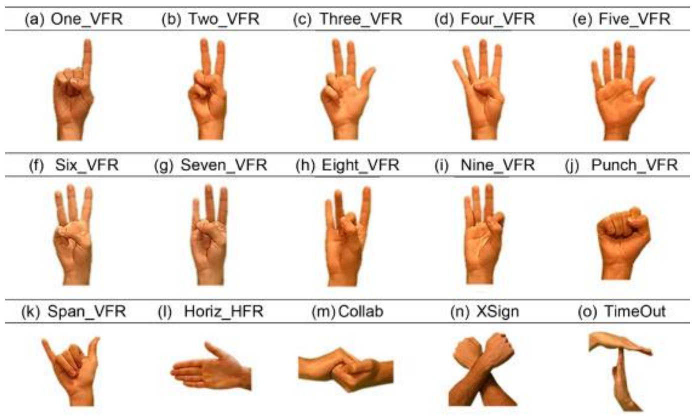
[1] Dietrich Kammer, Mandy Keck, Georg Freitag, Markus Wacker, Taxonomy and Overview of Multi-touch Frameworks: Architecture, Scope, and Features Archived 2011-01-25 at the Wayback Machine
[2] "touchless user interface Definition from PC Magazine Encyclopedia". pcmag.com. Retrieved 2017-07-28.

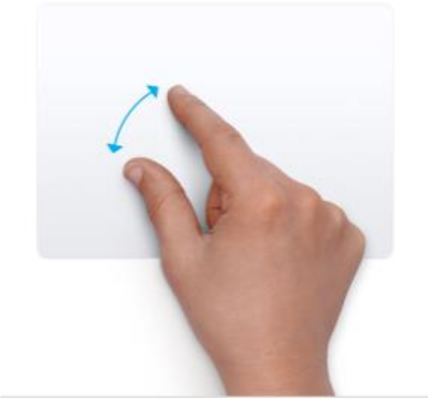
Gestures of Body Parts

- **Hand Gestures**: Common in touch-based interfaces (like smartphones and tablets) and in virtual or augmented reality environments
  - **Touchless**: without physical touching a device (often called **mid-air gestures**)
  - **Touch-based**: with touching or holding a device
  - Examples include pinching, swiping, and tapping
- **Arm Gestures**: Often used in immersive environments, like VR, where broad movements are tracked for interactive experiences
- **Facial Gestures / Mimics**: Facial expressions or head movements used for interaction, especially in accessibility contexts where traditional input methods might not be feasible
- **Body Gestures**: Full-body movements, often used in gaming or simulation environments

Hand Gestures

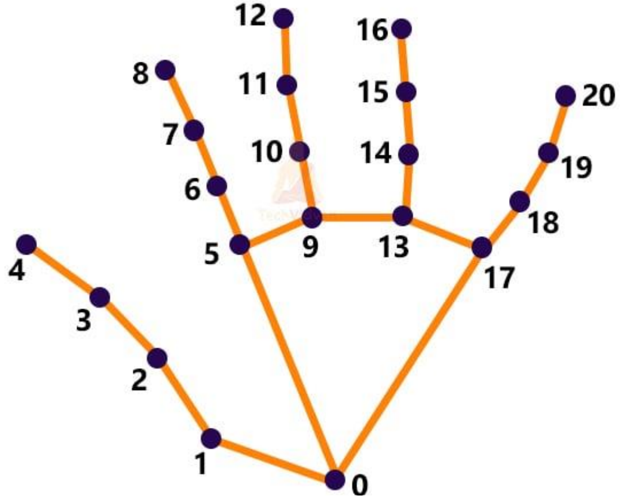
- **Specific movements or positions of the hands and fingers** that are used to communicate **commands or inputs** to a computer system
  - **Static Gestures** involve a specific hand shape or position, with little to no movement.
    - E.g., holding fingers in a certain configuration
  - **Dynamic Gestures** involve movement of the hands or fingers
    - E.g., swiping, pinching, or waving
- Technologies for detecting hand gestures
  - **Mid-Air**: e.g., camera-based systems, IMUs with accelerometers or gyroscopes
  - **Touch-based**: e.g., resistive and capacitive touchscreens



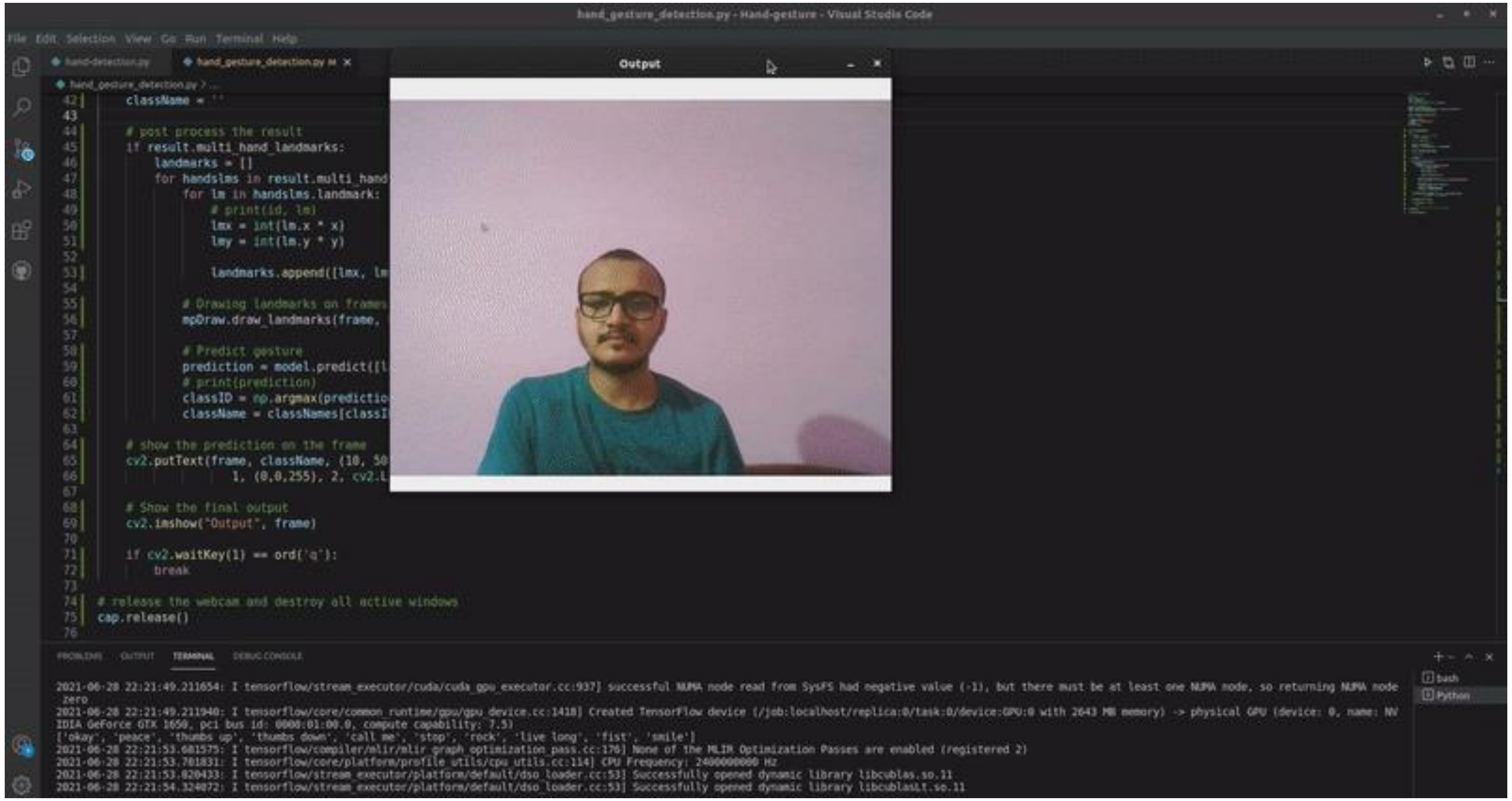


Images from: Nuzzi, C., Pasinetti, S., Pagani, R., Coffetti, G., & Sansoni, G. (2021). HANDS: an RGB-D dataset of static hand-gestures for human-robot interaction. Data in Brief, 35, 106791. doi: 10.1016/j.dib.2021.106791, https://techpp.com/2022/05/04/mac-trackpad-gestures-guide/,

Example: 2D Hand Recognition



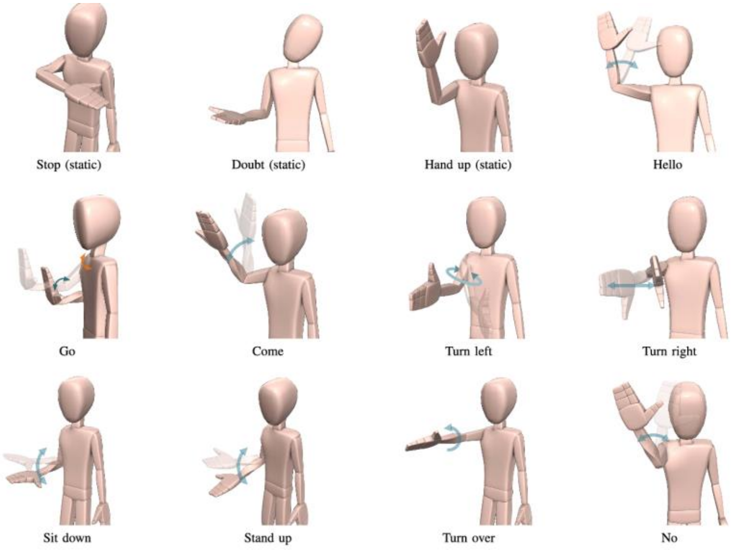
0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP



https://techvidvan.com/tutorials/hand-gesture-recognition-tensorflow-opencv/

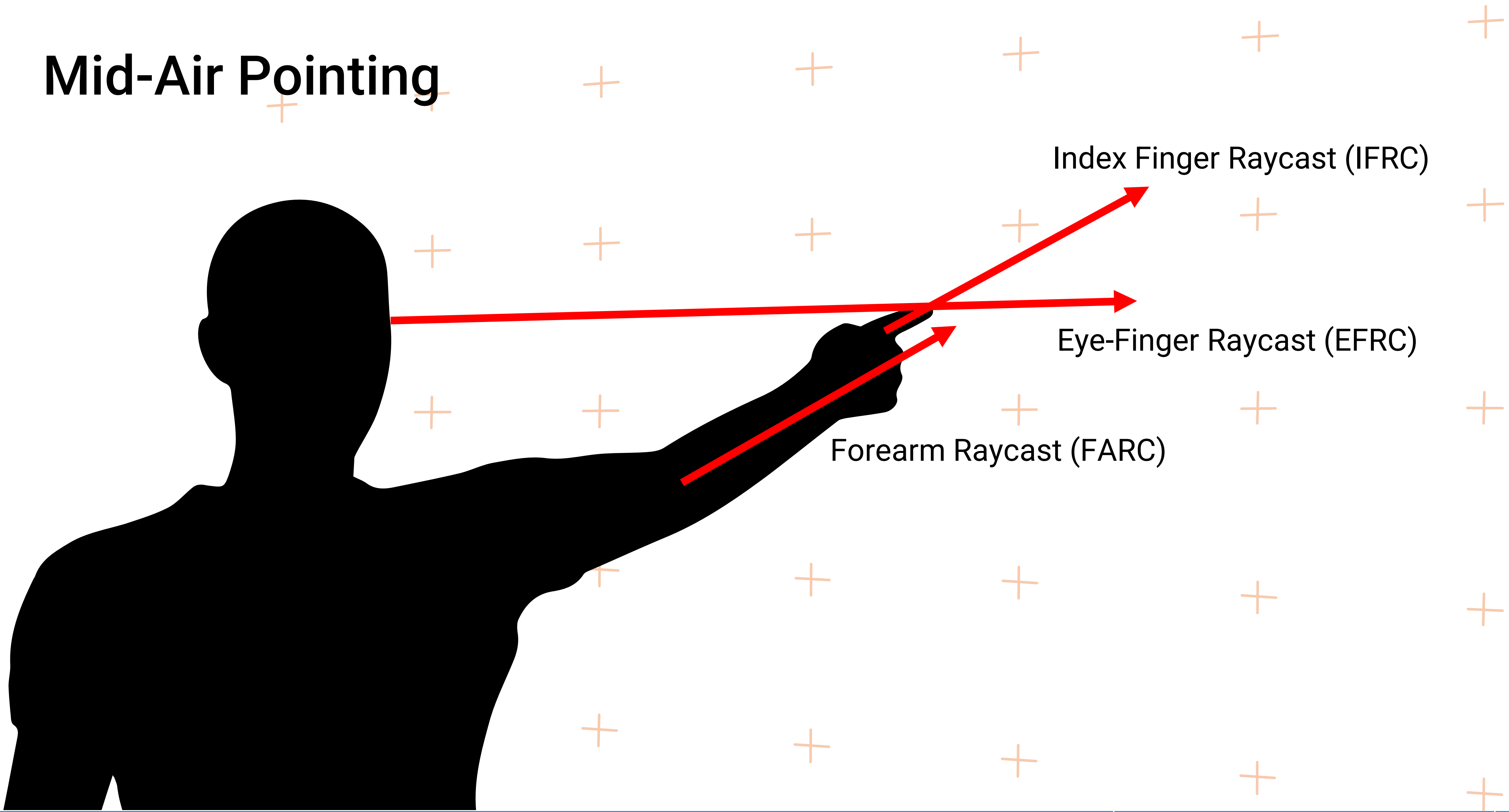
Arm Gestures

- **Movements or poses of the arm and hand** that are used to communicate with or control a computer system, particularly in spatially oriented or physically engaging applications
- Allow **natural forms of interaction**, where the user's physical actions are integrated into the experience
- Characteristics of arm gestures
  - **Spatial orientation**: required for almost all actions
  - **Expressiveness**: suitable for actions requiring larger motion
  - **Intensity**: indicating physical effort



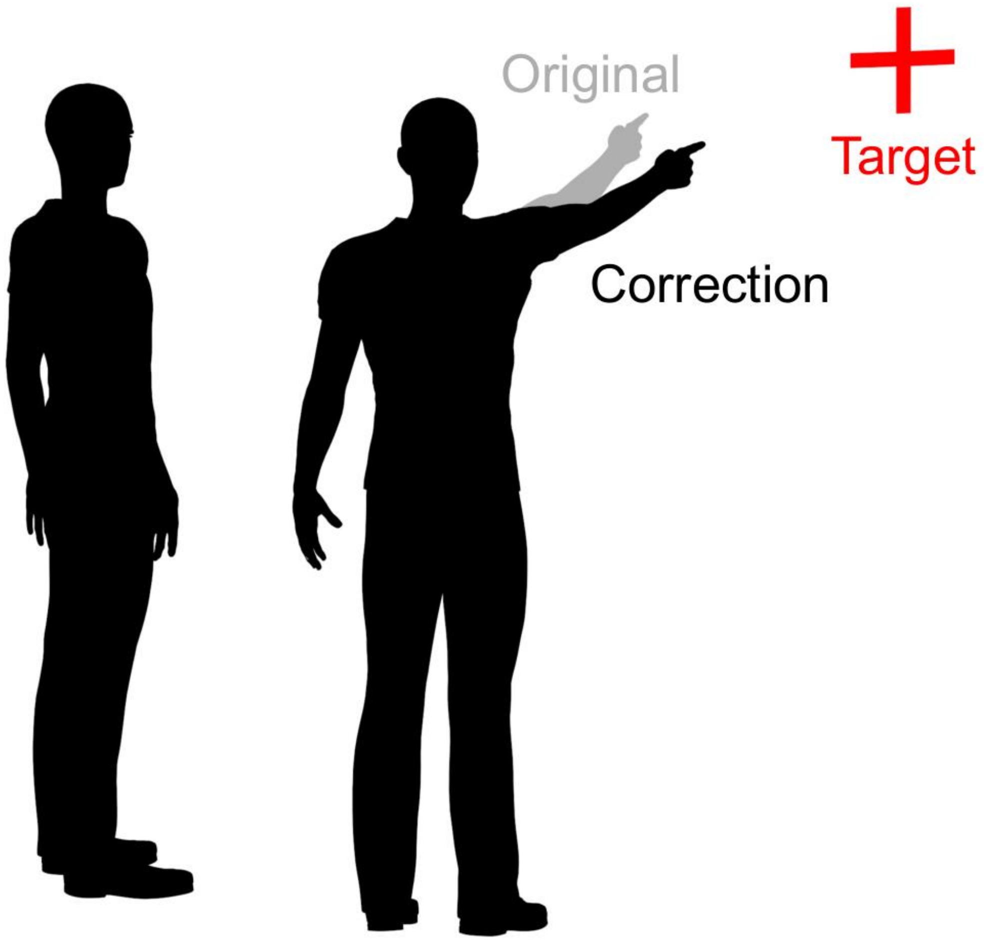
Wu, Y., Wu, Z., & Fu, C. (2018). Continuous Arm Gesture Recognition Based on Natural Features and Logistic Regression. IEEE Sens. J. Retrieved from https://www.semanticscholar.org/paper/Continuous-Arm-Gesture-Recognition-Based-on-Natural-Wu-Wu/3968dd5cf0ea78a66897423b87de68984a06c4d4

Mid-Air Pointing

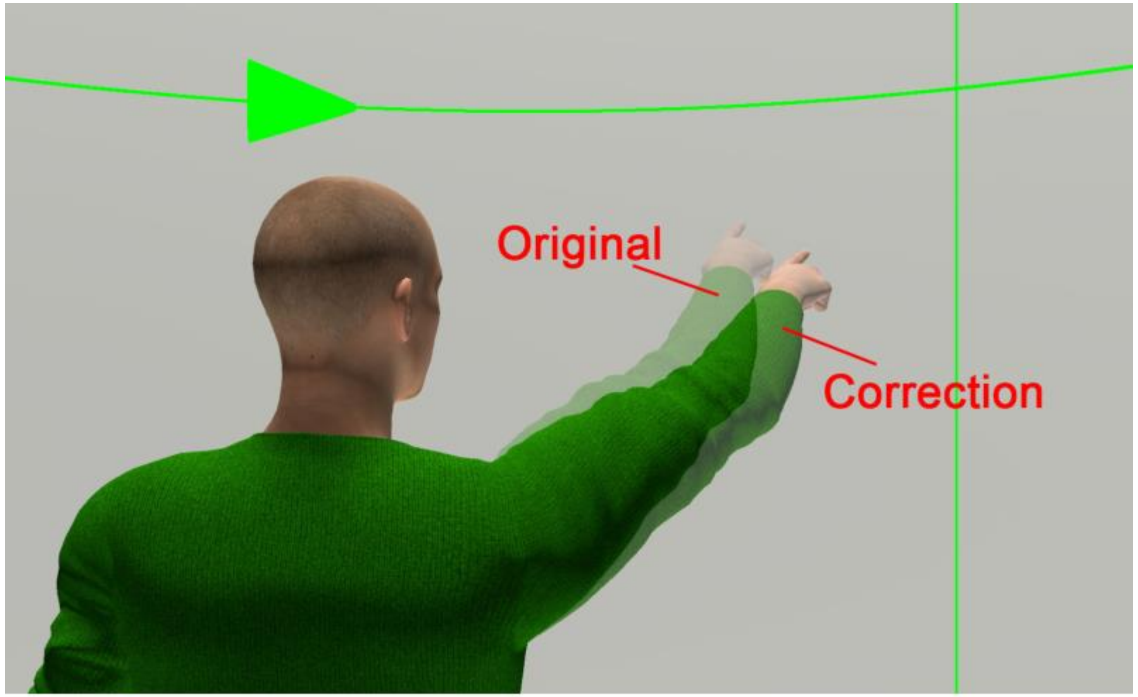


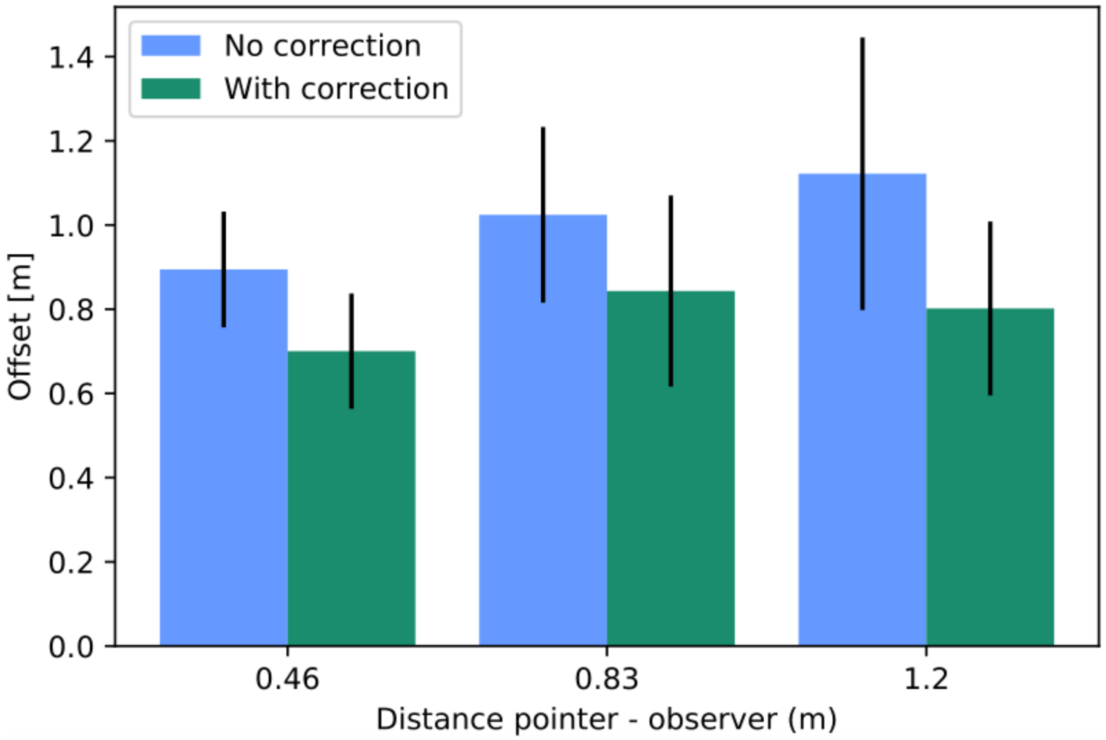

DistantPointing StudyGrid
Punkt anvisieren -> Klicken

Deictic Pointing



Deictic Pointing in VR



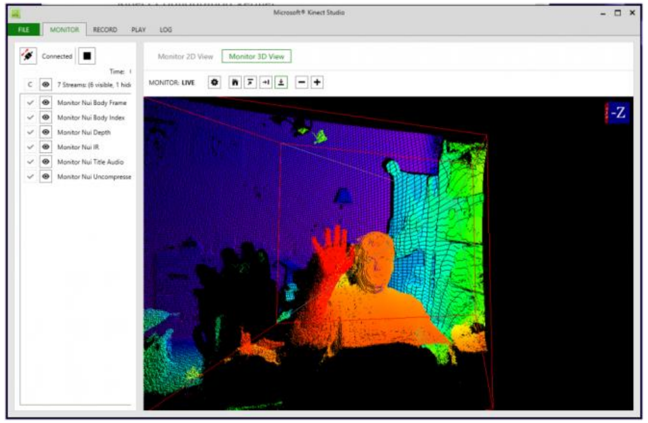


Sven Mayer, Jens Reinhardt, Robin Schweigert, Brighten Jelke, Valentin Schwind, Katrin Wolf, and Niels Henze. 2020. Improving Humans' Ability to Interpret Deictic Gestures in Virtual Reality. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14. DOI:https://doi.org/10.1145/3313831.3376340

https://de.m.wikipedia.org/wiki/Microsoft_PixelSense Microsoft PixelSense-Tisch Ergonomidesign • CC BY-SA 3.0

Gesture Recognition Algorithms

- Most techniques rely on pointers represented through coordinates
  - **3D-based**
    - Volumetric (3D model approximation)
    - Skeleton (joint angle parameters and segment lengths)
  - **Image/Video/Appearance-based**
    - 2D pattern matching based on deformable templates and feature detection (e.g., touchscreens)
  - **Electromyography (EMG)**
    - Classification of data received from the muscles
- Features from those algorithms serve as input for classification. **Hidden Markov models** or **artificial neural networks** are often used for this purpose.



Images from: https://thenewstack.io/kinect-for-app-development
Eadicicco, L. (2016). This Futuristic Armband Lets You Control Your Computer Like Magic. Time. Retrieved from https://time.com/4173507/myo-armband-review

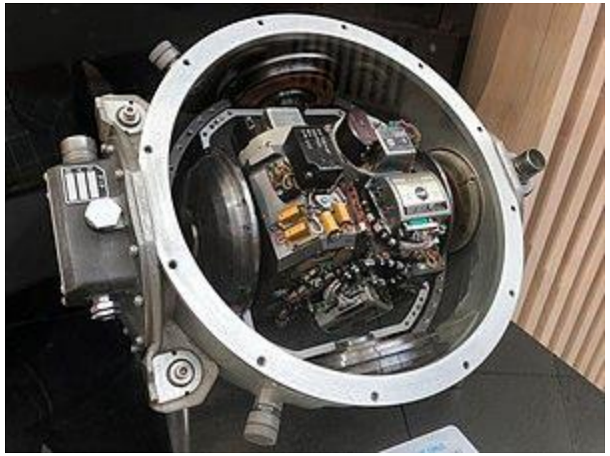
Example: EMG-based Gesture Detection using MyoBand

- The MyoBand also has an IMU (next slide)

https://gfycat.com/gifs/search/myo+armband

Inertial Measurement Units (IMU)

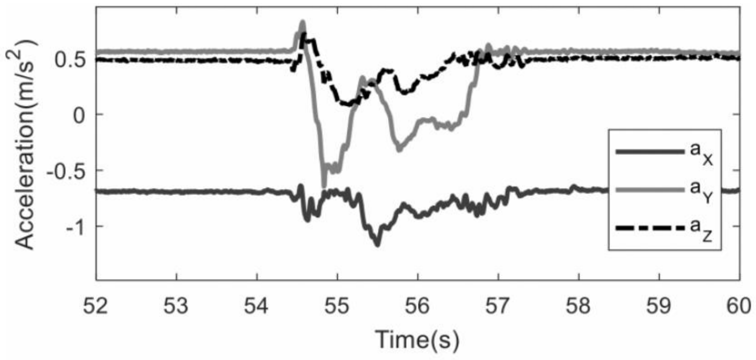
- **An electronic device (today just a microchip)** that can measure and report
  - a body's specific accelerating force using an **accelerometer**
  - orientation of the body using a **gyroscope**
  - surrounding magnetic fields using **magnetometers** (compass)
  - locate the body on earth with a **GPS sensor**
- Often **suffer from drifts** and ever-increasing differences between where the system thinks it is located and the actual one



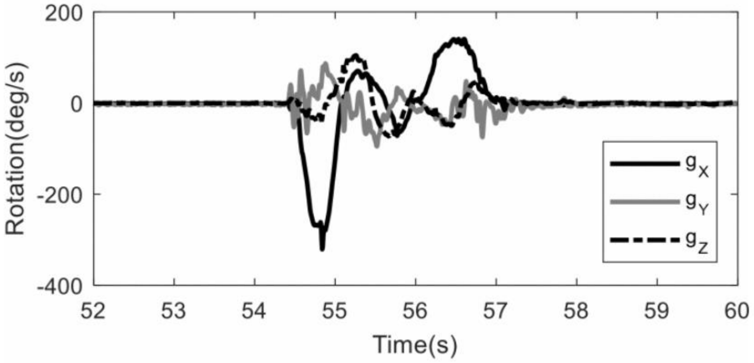
Apollo Inertial Measurement Unit by ArnoldReinhold - CC BY-SA 4.0 https://en.wikipedia.org/wiki/Inertial_measurement_unit#/media/File:Apollo_IMU_at_Draper_Hack_the_Moon_exhibit.agr.jpg

9-DoF IMU SiP breakout board, Sold for $16 by SparkFun in 2019. CC BY 2.0 https://en.wikipedia.org/wiki/Inertial_measurement_unit#/media/File:SparkFun_9DoF-IMU-Breakout_LSM9DS1_13284-03.jpg

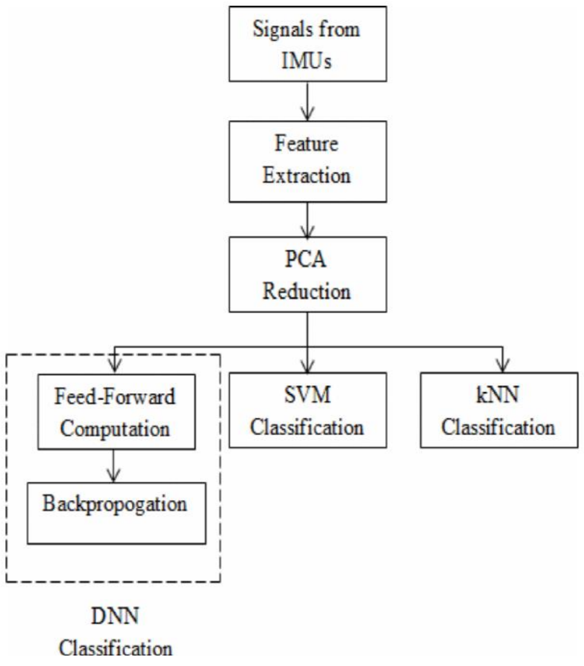
Gesture Recognition using IMUs



(a) Accelerometer signals for X, Y and Z coordinates



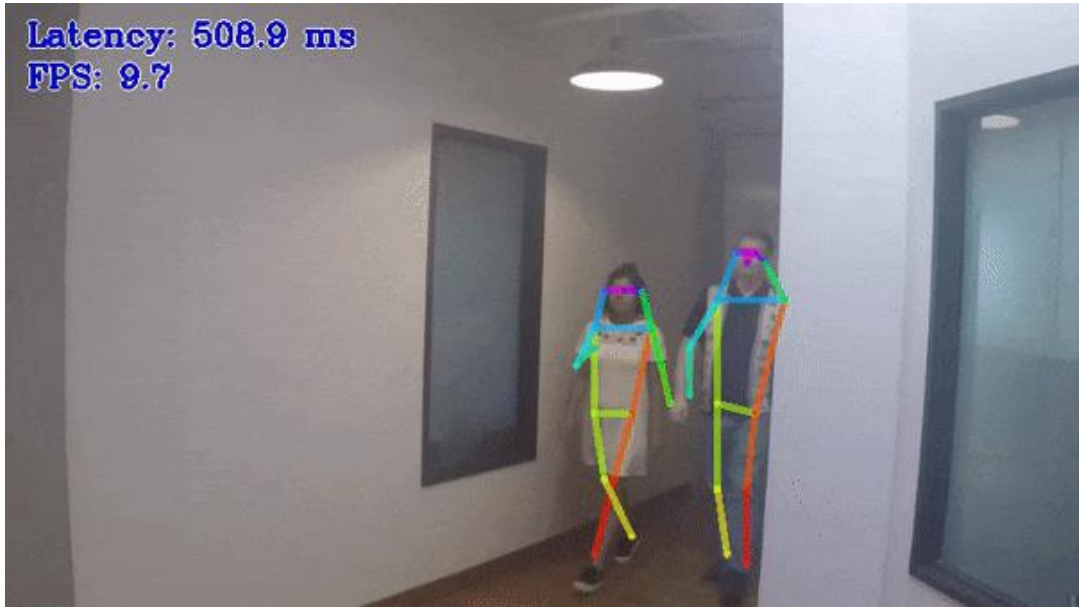
(b) Gyroscope signals for X, Y and Z coordinates



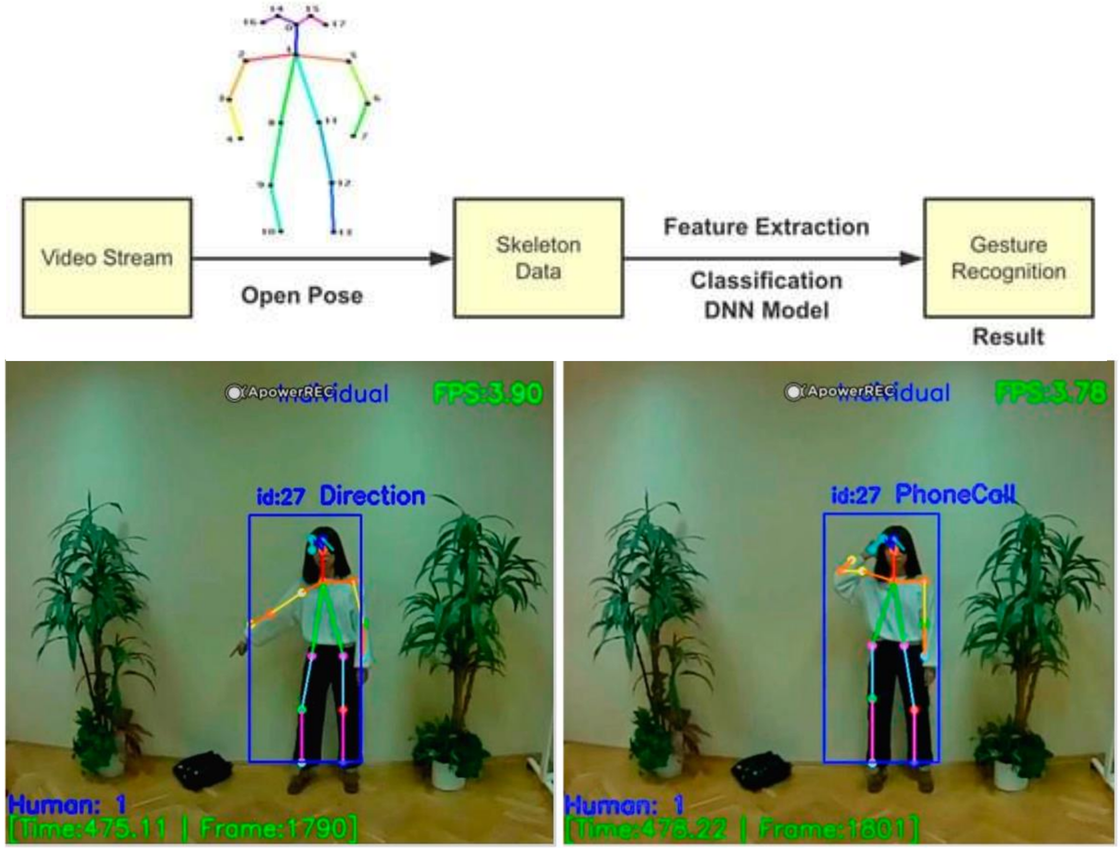
Public Domain. Image from https://de.wikipedia.org/wiki/Eingabeger%C3%A4t#/media/Datei:Wiimote-in-Hands.jpg

K. Suri and R. Gupta, "Classification of Hand Gestures from Wearable IMUs using Deep Neural Network," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, pp. 45-50, doi: 10.1109/ICICCT.2018.8473301.

Example: 3D Body/Skeleton Gesture Recognition



https://docs.openvino.ai/



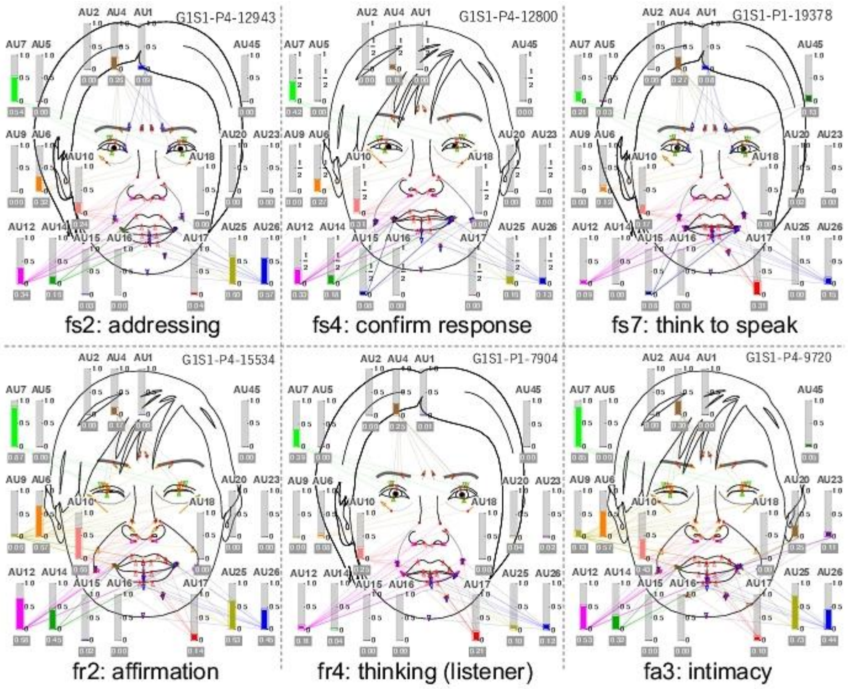
(a) (b)

https://docs.openvino.ai/latest/omz_demos_human_pose_estimation_demo_python.html

Liu, C.; Szirányi, T. Real-Time Human Detection and Gesture Recognition for On-Board UAV Rescue. Sensors 2021, 21, 2180. https://doi.org/10.3390/s21062180

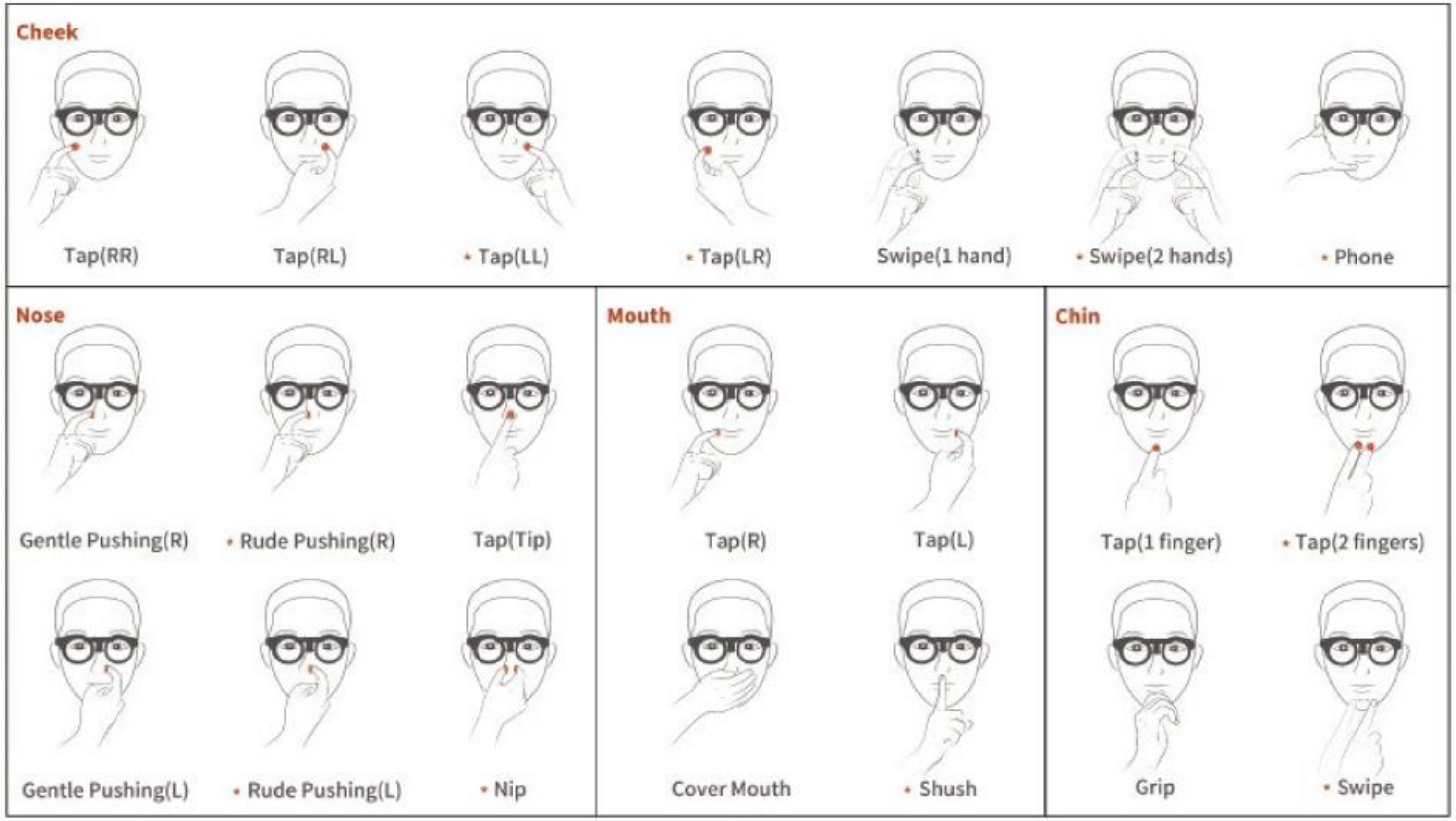
Face Gestures

- Facial gestures, often referred to as **facial expressions or facial mimics**, are a form of non-verbal communication where facial movements and expressions are used to interact with computer systems
- Leverages the natural and nuanced capability of the human face to **convey a wide range of emotions and intentions**
- **Face recognition raises significant privacy concerns** as a face can be identified; users must be informed and consent obtained.

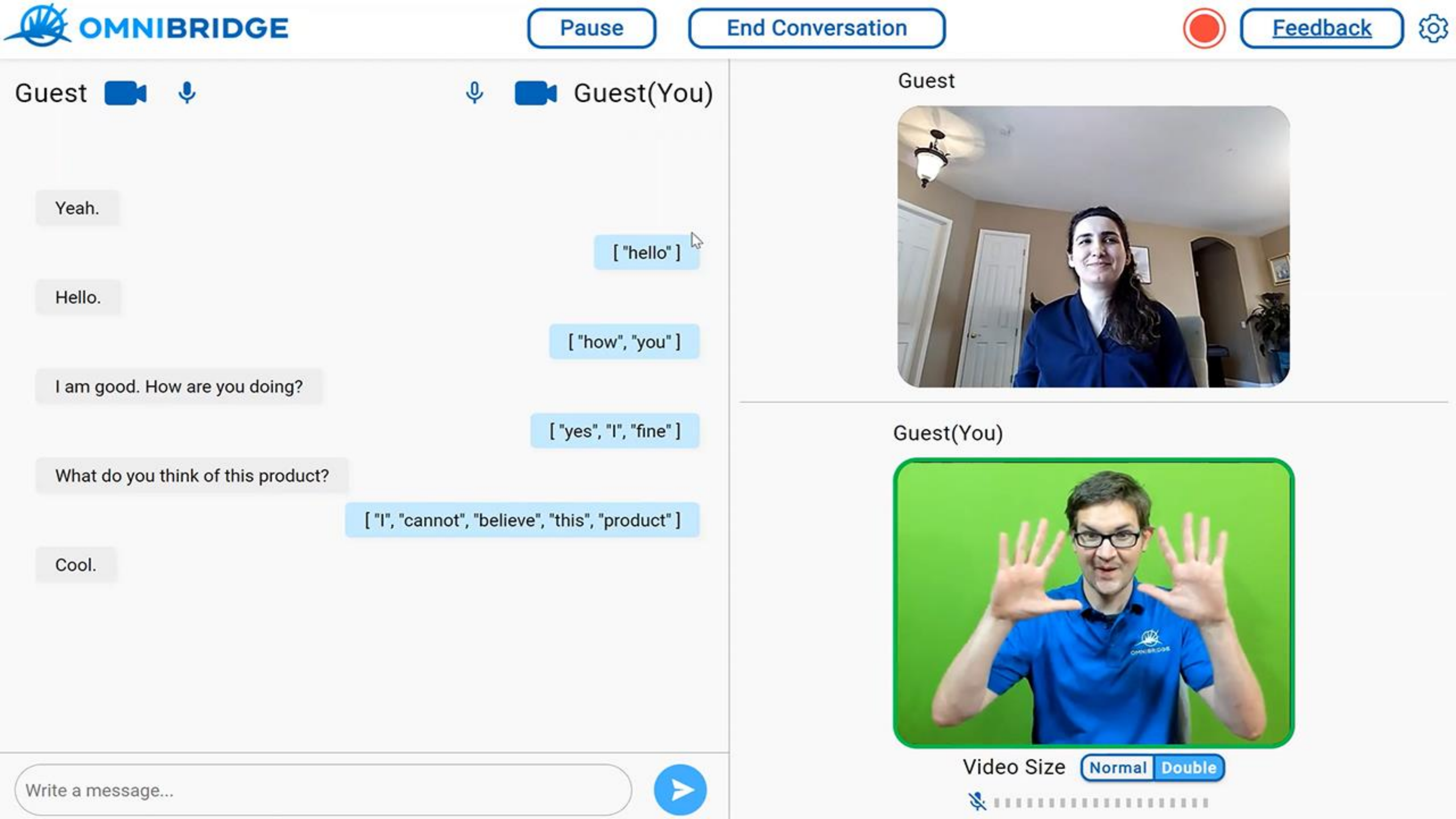


Imamura, M., Tashiro, A., Kumano, S., & Otsuka, K. (2023). Analyzing Synergetic Functional Spectrum from Head Movements and Facial Expressions in Conversations. ICMI '23: Proceedings of the 25th International Conference on Multimodal Interaction. Association for Computing Machinery. doi: 10.1145/3577190.3614153

Example: Controlling AR with Hand and Face Gestures

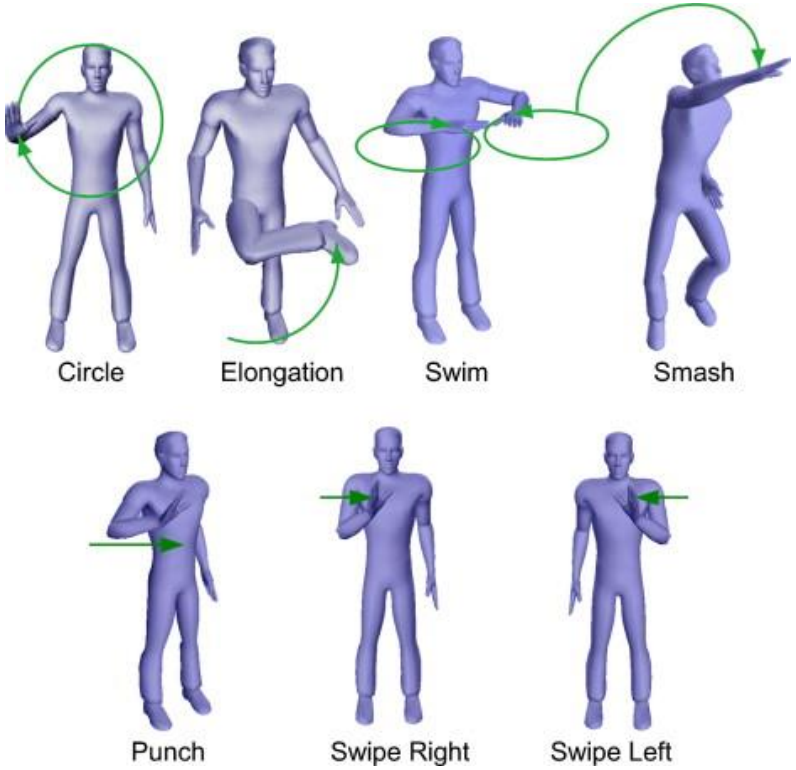


Weng, Y., Yu, C., Shi, Y., Zhao, Y., Yan, Y., & Shi, Y. (2021). FaceSight: Enabling Hand-to-Face Gesture Interaction on AR Glasses with a Downward-Facing Camera Vision. CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. doi: 10.1145/3411764.3445484


OMNIBRIDGE
Pause
End Conversation
Feedback
Guest
Guest(You)
Yeah.
["hello"]
Hello.
["how", "you"]
I am good. How are you doing?
["yes", "I", "fine"]
What do you think of this product?
["I", "cannot", "believe", "this", "product"]
Cool.
Write a message...
Guest
Guest(You)
Video Size
Normal
Double

Full-body Gestures

- Full-body gestures **involve the entire body** or large parts of it
- Types of Full-body Gestures:
  - **Posture-Based Gestures**: Assuming certain postures or stances as a form of input.
  - **Movement-Based Gestures:** Walking, jumping, turning, and movements that involve changing position or orientation of the body
  - **Complex Sequences**: Choreographed or patterned movements for specific interactions (e.g., dancing).



Ibañez, R., Soria, Á., Teyseyre, A., & Campo, M. (2014). Easy gesture recognition for Kinect. Adv. Eng. Software, 76, 171–180. doi: 10.1016/j.advengsoft.2014.07.005

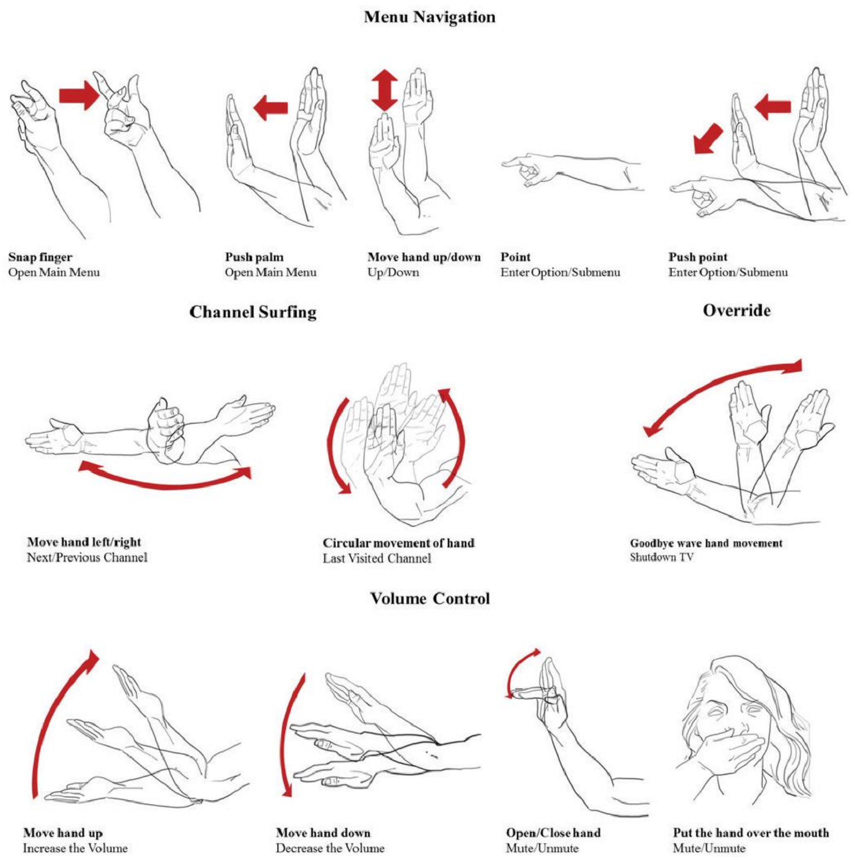
Sign Language

- The **Gestural Theory** states that language evolved from manual gestures
- **Manual articulation** (often combined with non-manual markers) **to convey meaning** instead of spoken words
  - Not universal (thgere are ~150 sign languages in the world)
- **A gesture** can represent a **word** or **phrase**.
- Mostly used by and emerged from deaf people
  - Also used by hearing individuals
- Used as benchmark for gesture recognition systems

Chirogram from Chirologia, 1644

Gesture Elicitation

- **Gathering potential gestures from end-users** for specific actions or commands in a system
  - **Researchers first identify a set of actions or commands** for which gestures are needed
  - A **group of participants is asked to propose gestures** for each action or command
  - **Researchers observe and document that gestures** using video recordings, notes, or sensor data
  - The **most suitable gestures** are selected based on the **agreement analysis**



Dong, H., Danesh, A., Figueroa, N., & Saddik, A. E. (2015). An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. IEEE Access. Retrieved from https://www.semanticscholar.org/paper/An-Elicitation-Study-on-Gesture-Preferences-and-a-Dong-Danesh/54617c296e4a2ba5a41d4dd9a65898aed7b30618

Human Activity Recognition (HAR)

- The problem of **identifying events performed by humans** given an input
  - Plan recognition
  - Goal recognition
  - Intent recognition
  - Behavior recognition
  - Location estimation
- Part of many **societal applications** including
  - smart surveillance
  - video search/retrieval
  - intelligent robots
  - monitoring systems

Sensing and Reasoning

- **Sensor** recognition
  - e.g., cameras, IMUs, microphone, WiFi, Bluetooth, GPS, etc…
- **Single-** or **multi-user** or **group behavior** recognition
  - e.g., Kinect, OpitTrack, etc…
- **Reasoning Chain**
  - Logic: keep track of all logically consistent explanations of the observed actions
  - Probabilistic: Hidden Markov Model (HMM), Dynamic Bayesian Networks (DBN), and Artificial Neural Networks (ANN) are popular choices in modelling activities from sensor data
- Example: by automatically monitoring human activities, home-based rehabilitation can be provided for people suffering from traumatic brain injuries or falling

Activity Recognition Chain (ARC)

- Usually done using the **activity recognition chain (ARC)**
  1. Sensor Data Acquisition (mostly time-sampled in regular intervals)
  2. Pre-Processing (event synchronization, noise removal)
  3. Data Segmentation (time-based segments of relevant information, e.g., sliding window)
  4. Feature Extraction and Selection (reducing signals into features)
  5. Training and Classification (algorithms to predict a label given a feature input vector)
  6. Evaluation (correctness, true/false-positive/negative rates)
- **Main challenges of HAR**: definition and diversity of human activities, imbalance between the classes, and ground truth annotation

Context Awareness

- Originated from ubiquitous computing and describes the capability to take into account **the situation of entities**
  - e.g., location is only the most obvious element of this situation
- Context-aware systems are concerned with the 3As
  - **Acquisition** of the context (e.g., using GPS sensors to perceive a situation)
  - **Abstraction** of the context (e.g., matching a perceived sensory stimulus to a context)
  - **Application** based on the recognized context (e.g., triggering actions based on context)
- Context awareness has particularly been applied to the area of computer-supported cooperative work (CSCW) to help individuals work and collaborate more efficiently with each other

Activity Recognition Context through Location

Eye Tracking

Gaze, pupil and blinking interaction

Image from: https://pxhere.com/de/photo/1058439

Eye Tracking

- The technology and methods used to **measure where and how a person's eyes are looking at**
  - Measuring the gaze point (where someone is looking at)
    - Through **determining eye rotation using eye tracker**
  - Detects the motion of the eyes in space and relatively to the head
    - Thus, eye tracking typically **includes head tracking**
- Used to **measure attention using fixations**
- Used to as **input technique** for interactive applications

Image from: SMI Eye Tracking Glasses Neuromarketing | by SMI Eye Tracking

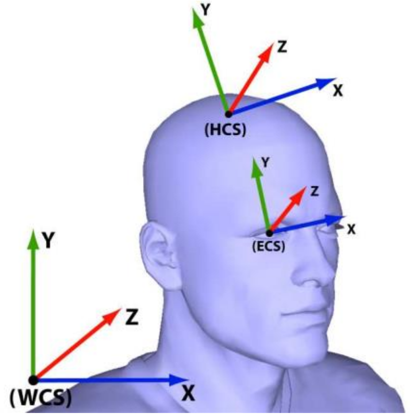
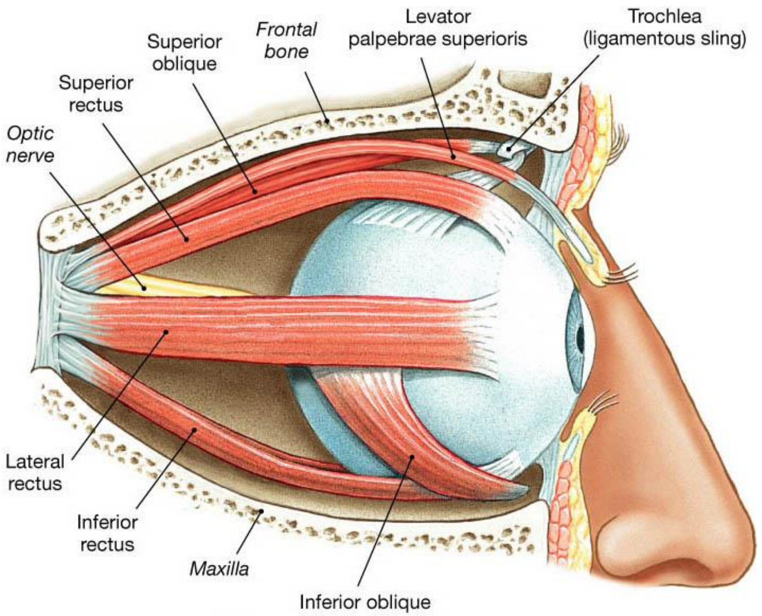
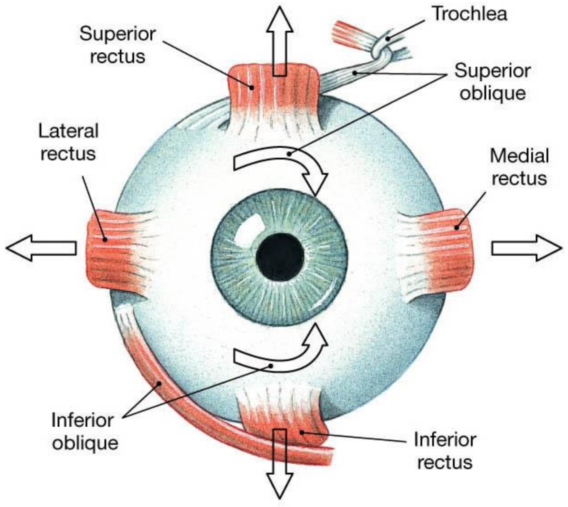


Image from: LibGaze: Real-time gaze-tracking of freely moving observers for wall-sized displays
Herholz, S., L. L. Chuang, T. G. Tanner, H. H. Bülthoff and R. W. Fleming

Eye Muscles



(a) Lateral surface, right eye

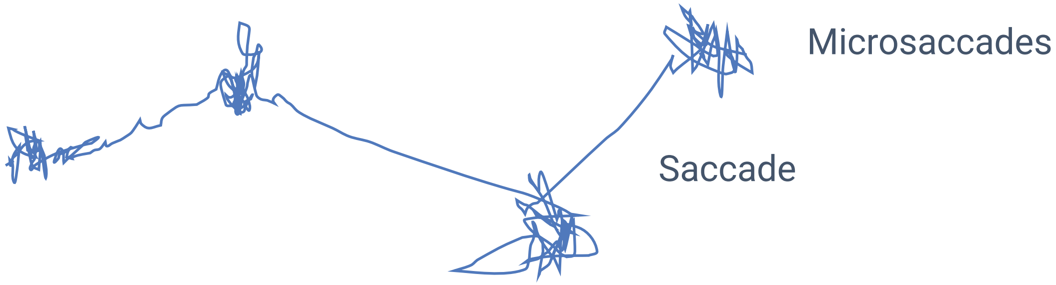


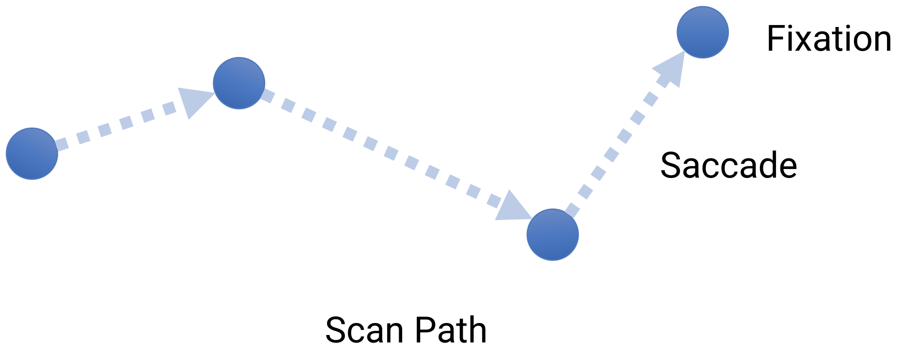
(c) Anterior view, right eye

Images from: https://droualb.faculty.mjc.edu/Lecture%20Notes/Unit%203/muscles%20with%20figures.htm and https://de.wikipedia.org/wiki/Augenmuskeln#Entwicklungsgeschichte

Types of Movements

- **Vergence movements**
  - Cooperation of both eyes to focus a single object
- **Pursuit movements**
  - Eyes follow a moving target smoothly
- **Saccadic movements**
  - The movement of eye to rapidly scan a scene





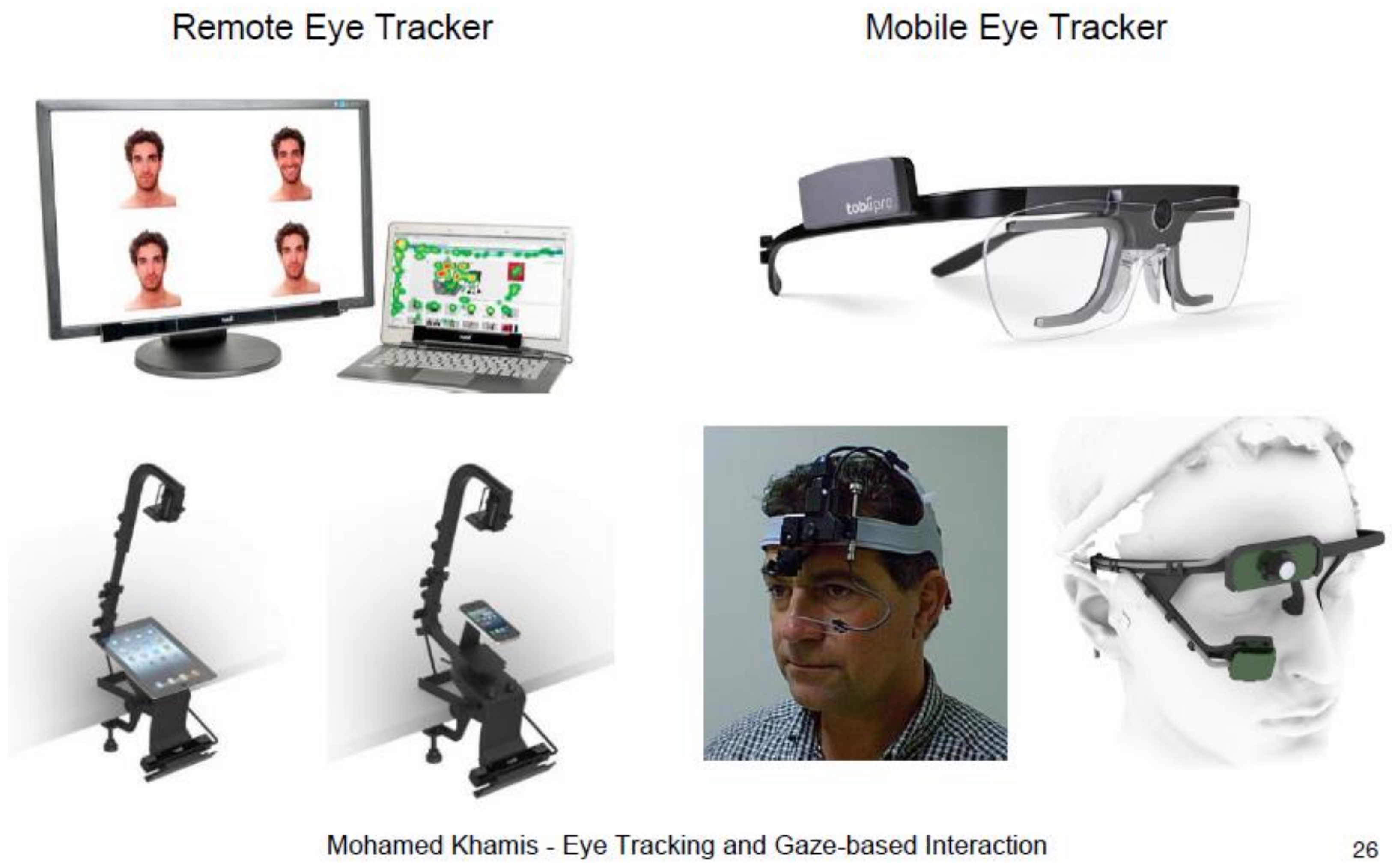
Eye Tracker Techniques

- **Video-based tracking**
  - Video-based/RGB
  - Infrared Corneal Reflection
- **Electrooculagraphy-based tracking**
  - Measures the electrical potential between the front (cornea) and the back (retina) of the human eye
  - Often used together with EEG devices
- **Contact lenses with coils**
  - The eye movement is calculated by inducted voltage of a magnetic field.

Non-contact, well suited for usability studies and experiments with several participants.

Work with closed eyes (eg. for REM-sleep measurements). Very precise and very expensive.

Remote and Mobile Eye Tracker

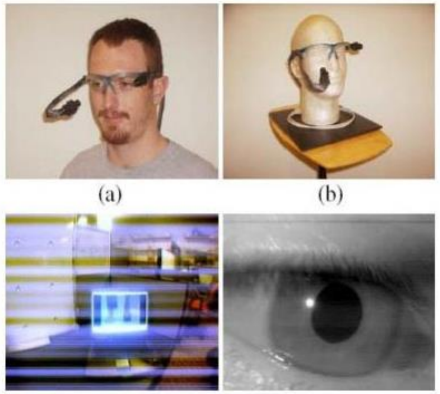


Eye Tracking Techniques

	Remote	Mobile
Gaze Tracking	On Displays	Natural / Real Environments
Requires	Software using a SDK	Labeling the surroundings (e.g. using QR-Markers)
Setup	Easy	Cumbersome to wear
Flexibility	Limited Movements in a Volume	Free Head / Body Movement
See Gaze Data?	Yes	No
Head Tracking	No	Yes

Commercial and Open Source Eye Trackers

Open Source



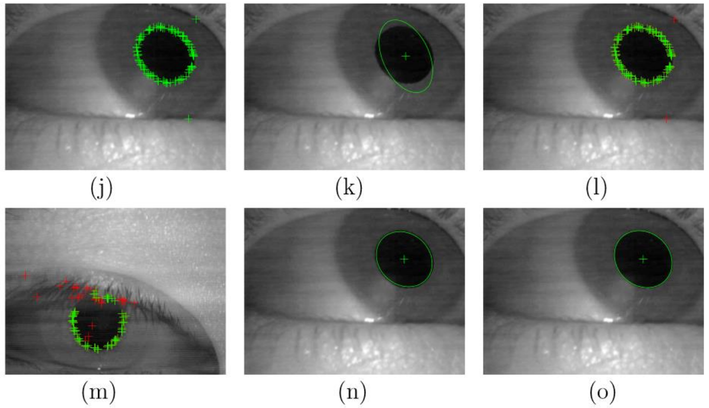
Open Eyes project
http://thirtysixthspan.com/openEyes/

Pupil Eye Tracker (open source software)

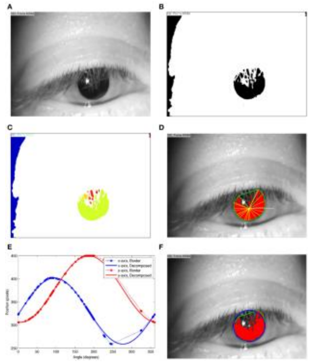
Commercial

Video-Based Eye Tracking

- **RGB**
- **Infrared**
  - Video based eye tracking uses the retina in our eyes that reflects infrared light (this caused red-eye effect from photo flashes).
  - A camera without infrared filter can increase the contrast
- **Pupil Detection**
  - Starburst
  - SET
  - Circular Hough transform



Li, D., & Parkhurst, D. J. (2005). Starburst : A robust algorithm for video-based eye tracking. *Image (Rochester, N.Y.)*, (September 2005), 22.



Javadi, A.-H., Hakimi, Z., Barati, M., Walsh, V., & Tcheang, L. (2015). SET: a pupil detection method using sinusoidal approximation. *Frontiers in Neuroengineering*, *8*(April), 1–10. http://doi.org/10.3389/fneng.2015.00004

Eye Tracker Calibration

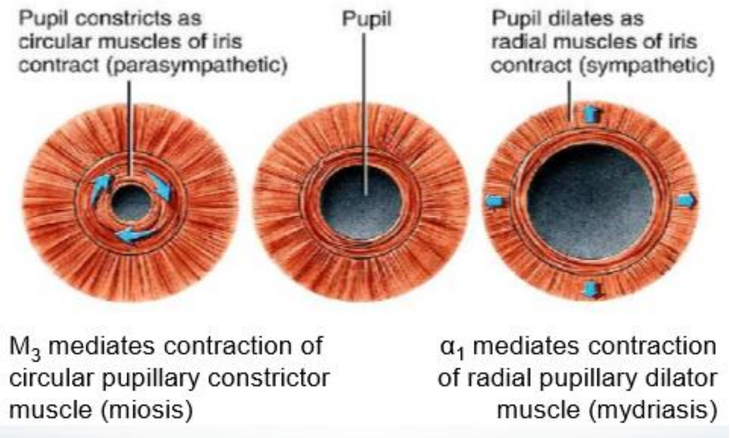
- Necessary to map eye gaze data to target area
- Difficult and tedious
- Might break if a users moves in relation to the device
- Almost impossible to determine the exact pixel a user is gazing at

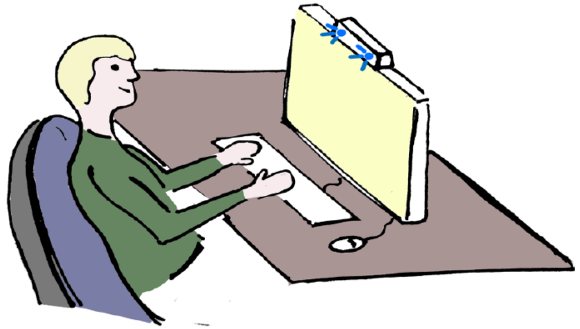
Calibration using markers

On-screen calibration using
the Eye-tribe

The Pupil Diameter

- The pupil is surrounded by two sets of muscles: a circular set and a radial set
  - The **circular muscles react to the presence** or absence of light
  - The **radial muscles react when a person** exerts mental effort.
- The **Index of Cognitive Activity (ICA)** measures pupil changes resulting from
  - mental effort (radial muscles)
  - factoring out the light reflex (circular muscles)





	Light (large impact)		Task (small impact)	
	low	high	difficult	easy
Pupil diameter	↗	↘	↗	↘

Model: $PD = PD_{light} + PD_{task}$

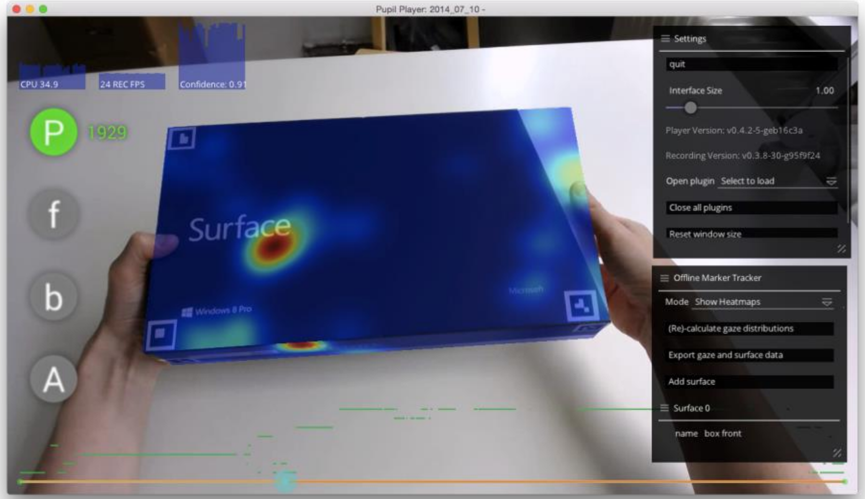
$PD_{task} = PD - PD_{light}$

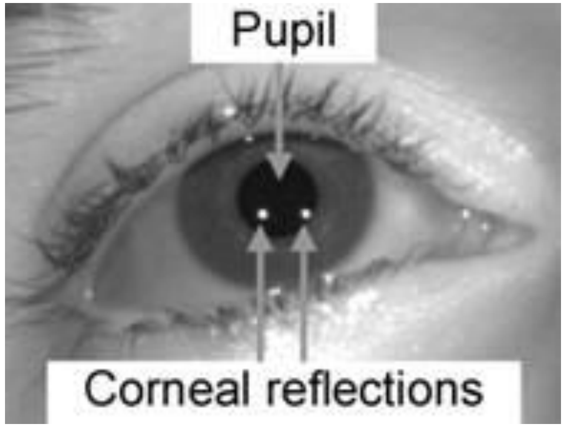
Estimate: $mental\ workload \approx f\,(PD_{task})$

A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions
Bastian Pfleging, Drea K. Fekety, Albrecht Schmidt, Andrew L. Kun, in Proc CHI `16

Head Tracking

- **Head-mounted**
  - AR-Mapping of the eye and world camera (e.g. with QR-Markers)
- **Remote**
  - **Pan-tilt**: Camera optics are moving according to the head movements
  - **Tilting-Mirror**: Camera and lens are fixed, servo-driven mirrors track eyes during head movements
  - **Fixed Camera(s)**: Image processing (e.g. using Reflections) without mechanical components
- **Chin rest**





Images from: http://www.ultimate3dheaven.com/cypchetr.html
https://github.com/pupil-labs/pupil/wiki/Basic-Workflow

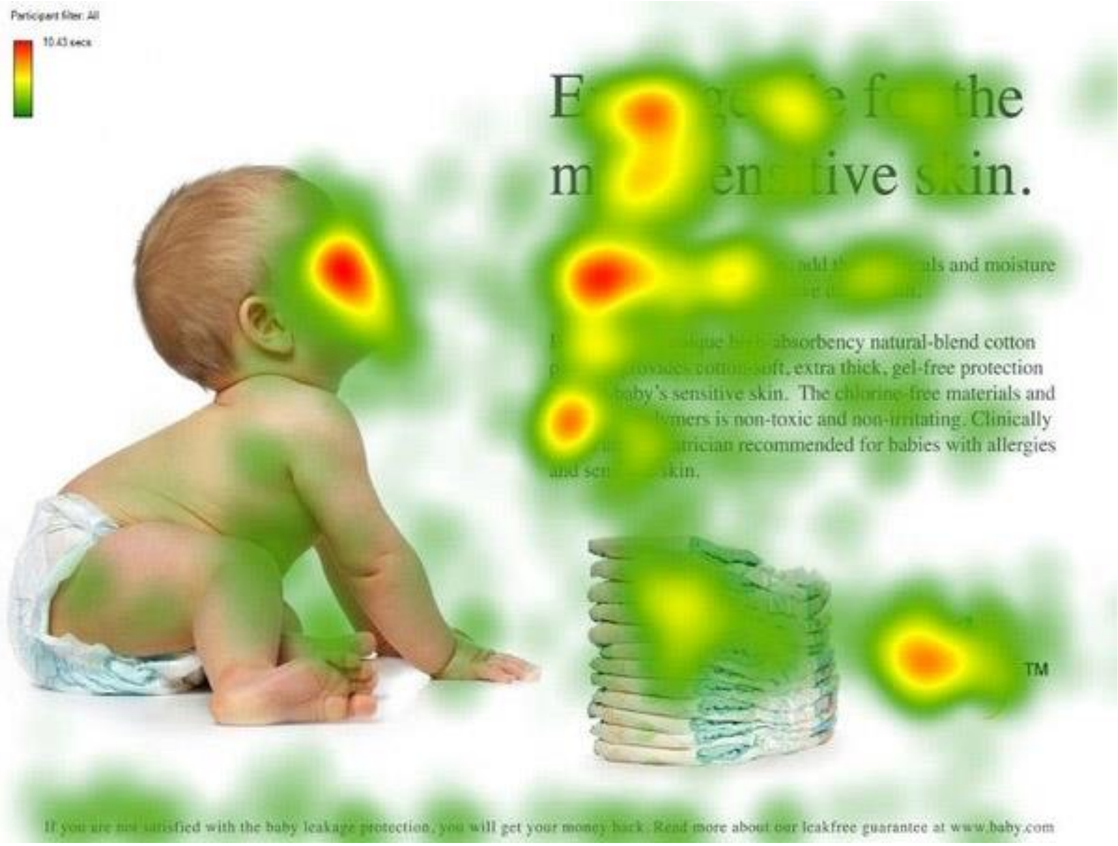
What you can learn through Eye Tracking?

- **Gaze Points/Fixations**
  - Attention (e.g. with websites, in supermarkets)
- **Pupil Diameter**
  - Cognitive Load/Stress, Health, Drug Consumption, Kidney Distortions
- **Blinks**
  - Fatigue (e.g., while driving a car)
- **Additional understanding** through
  - Inputs (e.g. response times, mouse moves, mouse clicks)
  - e.g. EEG, fMRI (brain activity)
  - Galvanic Skin Response (Stress)
- **No conclusions about what somebody saw or not saw** (!) and think or not think (!)
- Used to as **input technique** for interactive applications

Where is Eye Tracking used?

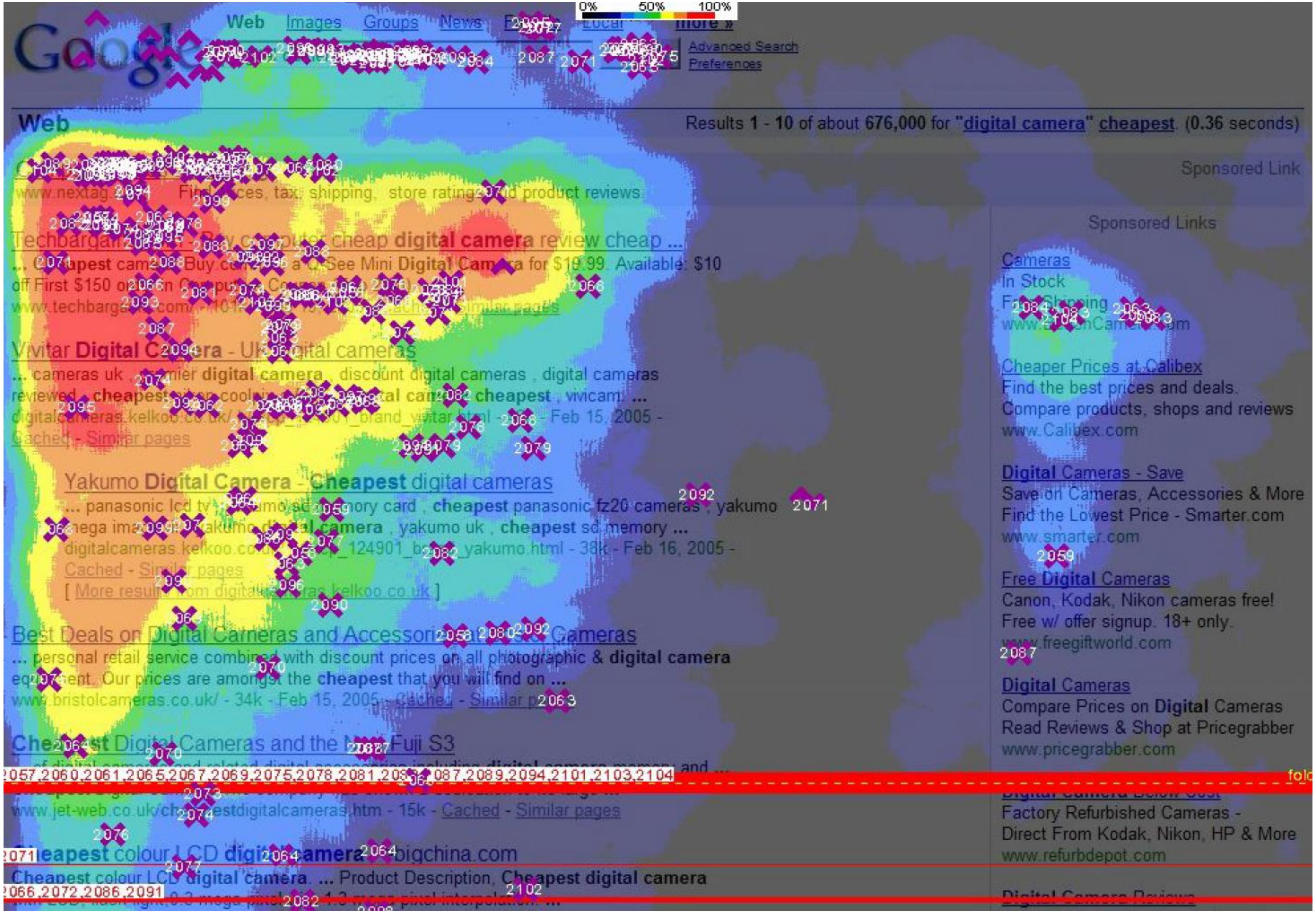
- Visual Analytics
- Psychology
- Neurosciences
- Advertising/Marketing Research
- Usability Research
- Linguistcs
- Medicine/Health
- Human-Computer Interaction


Media: Diapers-01.jpg
Participant filter: All
21.75 secs
tested and pediatrician recommended for babies with allergies
and sensitive skin.
TM
If you are not satisfied with the baby leakage protection, you will get your money back. Read more about our leakfree guarantee at www.baby.com


Participant filter: All
10.43 secs
TM


Wake up
the red
in your hair.
SHAMPOO
SUNSILK
Red
SUNSILK


Wake up
the red
in your hair.
SHAMPOO
SUNSILK
Red
SUNSILK



Gaze for Interaction

- Gaze is fast (faster than pointing) [Sibert and Jacob 2000]
- Natural: Concious / Subconcious
- Precedes Action
- Can be detected from a distance
- Reflects some insights into human states
  - Subconciouss Processes
  - Cognitive Work Load (Stress)
  - Fatigue

The Midas Touch Problem

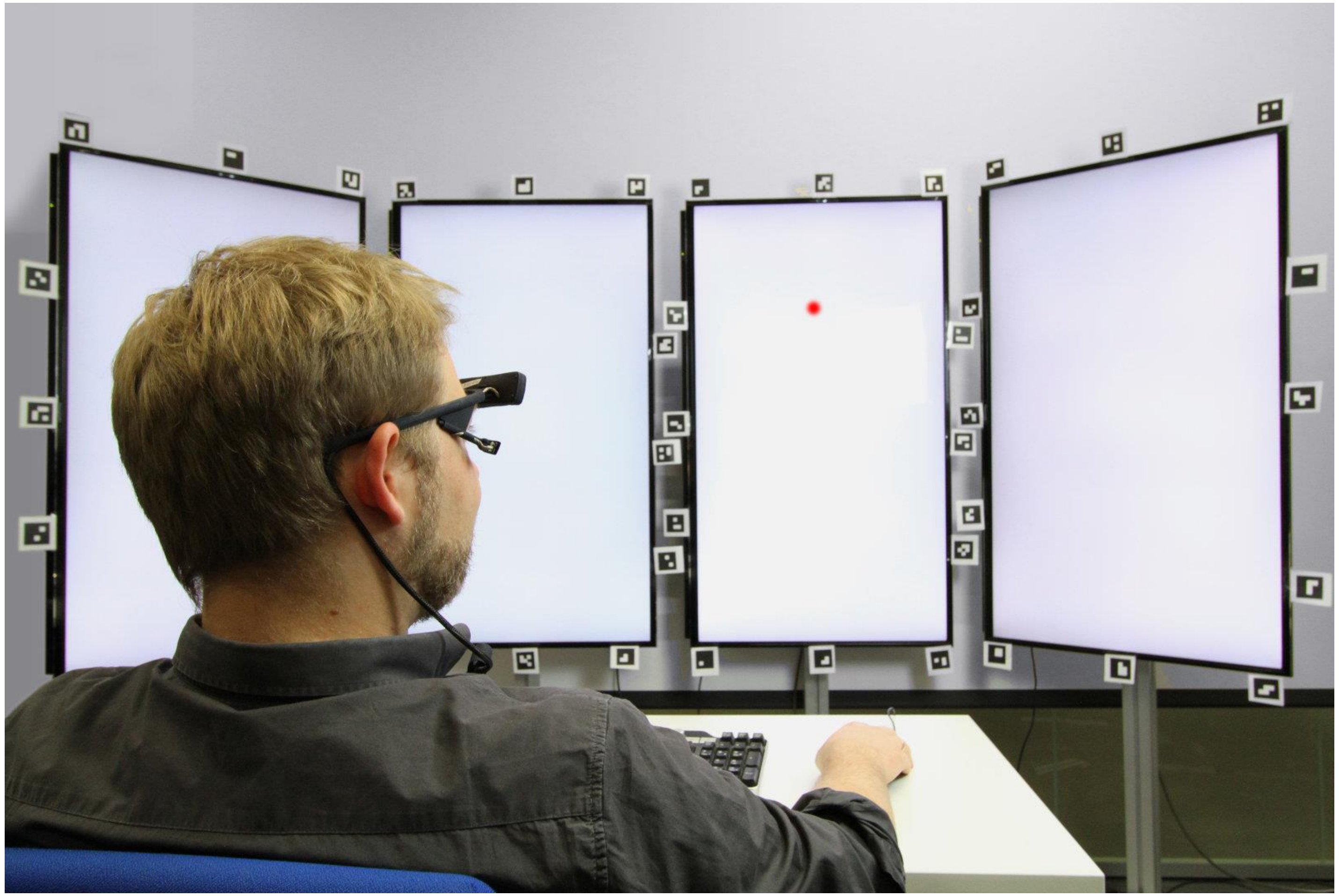
- Named after the mythical King Midas, whose touch turned everything to gold
- When eye gaze is used for interaction and as pointing device, **the basic function of eye to look** and to perceive visual information, **should be distinguished from the deliberate interaction**
- **When realized straightforwardly**, **every fixation** on an interface element **will lead to its activation of the user interface element** even when the user has no such intention.
  - This is called “the Midas touch problem”

King Midas with his daughter, from A Wonder Book for Boys and Girls by Nathaniel Hawthorne (text author), Walter Crane, 1893

Addressing the Midas Touch Problem

- **Dwell Time**: A common solution is to use 'dwell time,' where a selection is only made if the user's gaze remains on an element for a predetermined amount of time.
- **Gaze Gestures**: Implementing specific gaze patterns or movements as intentional commands, distinguishing them from regular viewing behavior.
- **Multimodal Interaction**: Combining eye tracking with other forms of input, like voice commands or hand gestures, to confirm selections or actions.
- **Contextual Awareness**: Designing systems that can interpret the context of the gaze, understanding whether a selection is likely to be intentional based on the user's current task or focus.
- **No solution for the one-way direction** in eye tracking: the ability to go back to the previous level (i.e. zooming out after zooming in)

Example: Eye Tracking on Large Displays



Example: Text Entry with Gestures and Eye Tracking

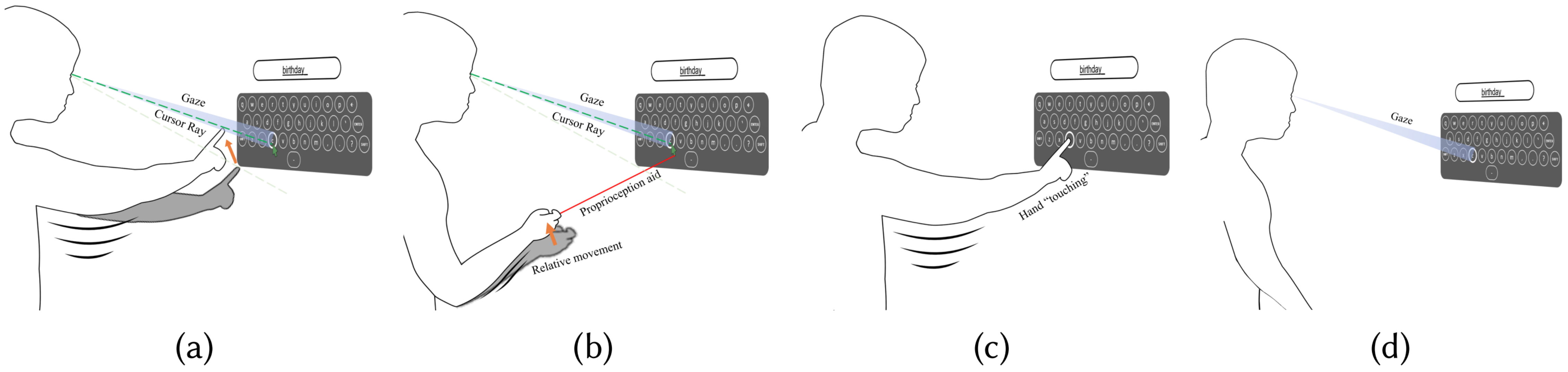
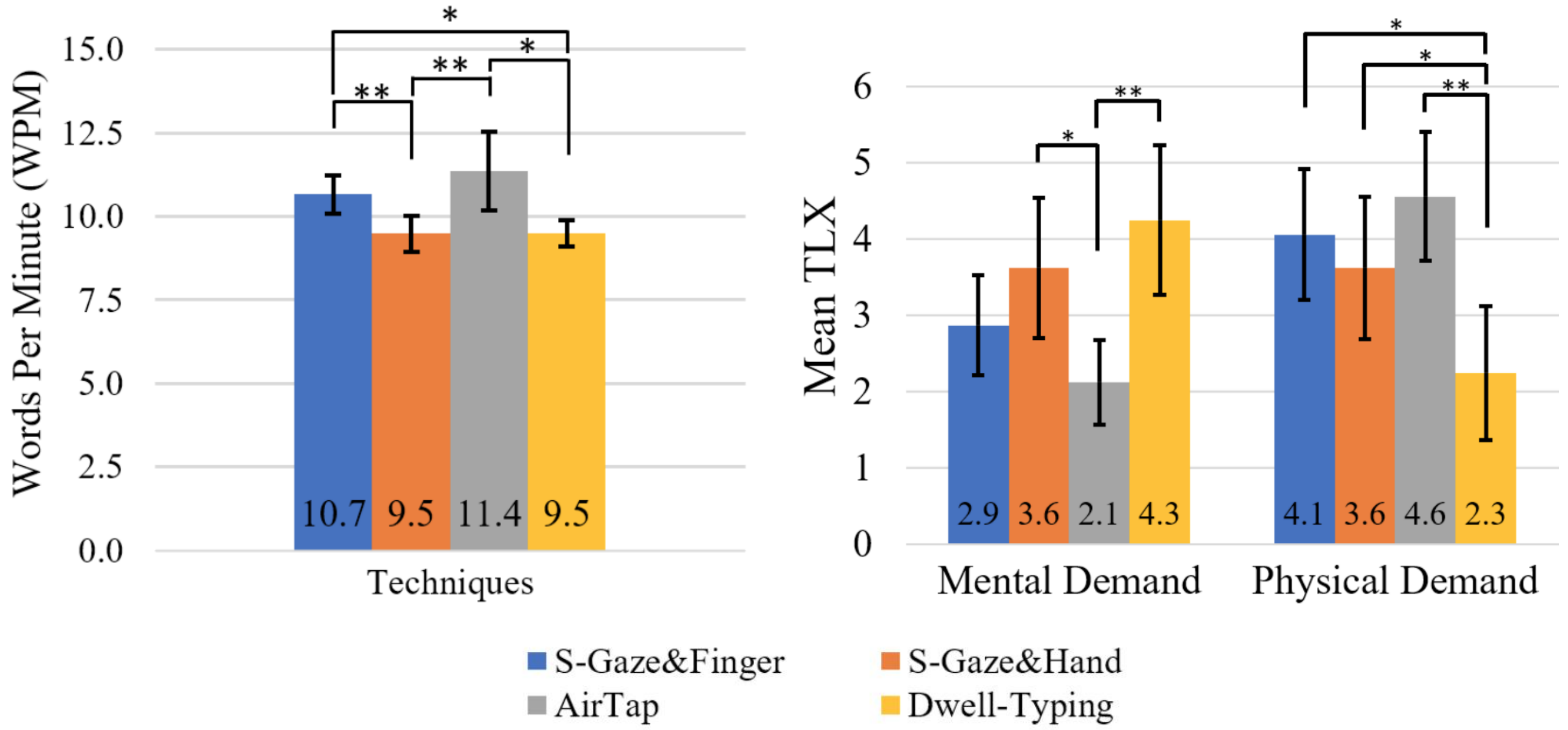


Fig. 1. Text entry using *S-Gaze&Finger* (a), *S-Gaze&Hand* (b), *AirTap* (c), and *Dwell-Typing* (d).

Mathias N. Lystbæk, Ken Pfeuffer, Jens Emil Sloth Grønbæk, and Hans Gellersen. 2022. Exploring Gaze for Assisting Freehand Selection-based Text Entry in AR. Proc. ACM Hum.-Comput. Interact. 6, ETRA, Article 141 (May 2022), 16 pages. https://doi.org/10.1145/3530882

Example: Text Entry with Gestures and Eye Tracking



Mathias N. Lystbæk, Ken Pfeuffer, Jens Emil Sloth Grønbæk, and Hans Gellersen. 2022. Exploring Gaze for Assisting Freehand Selection-based Text Entry in AR. Proc. ACM Hum.-Comput. Interact. 6, ETRA, Article 141 (May 2022), 16 pages. https://doi.org/10.1145/3530882

Input and Output (I/O)

Application Domain

Pointing (Gestures)
Text (Pen, Keyboard)
Direct Control
Speech, Audio
Video, Sensors
Context, Meta Data

input → **output**

bandwidth

Visual
Audio
Tactile
Smell
Taste
Temperature

Displays

Visual Output Technologies

Image from: https://pxhere.com/en/photo/893775

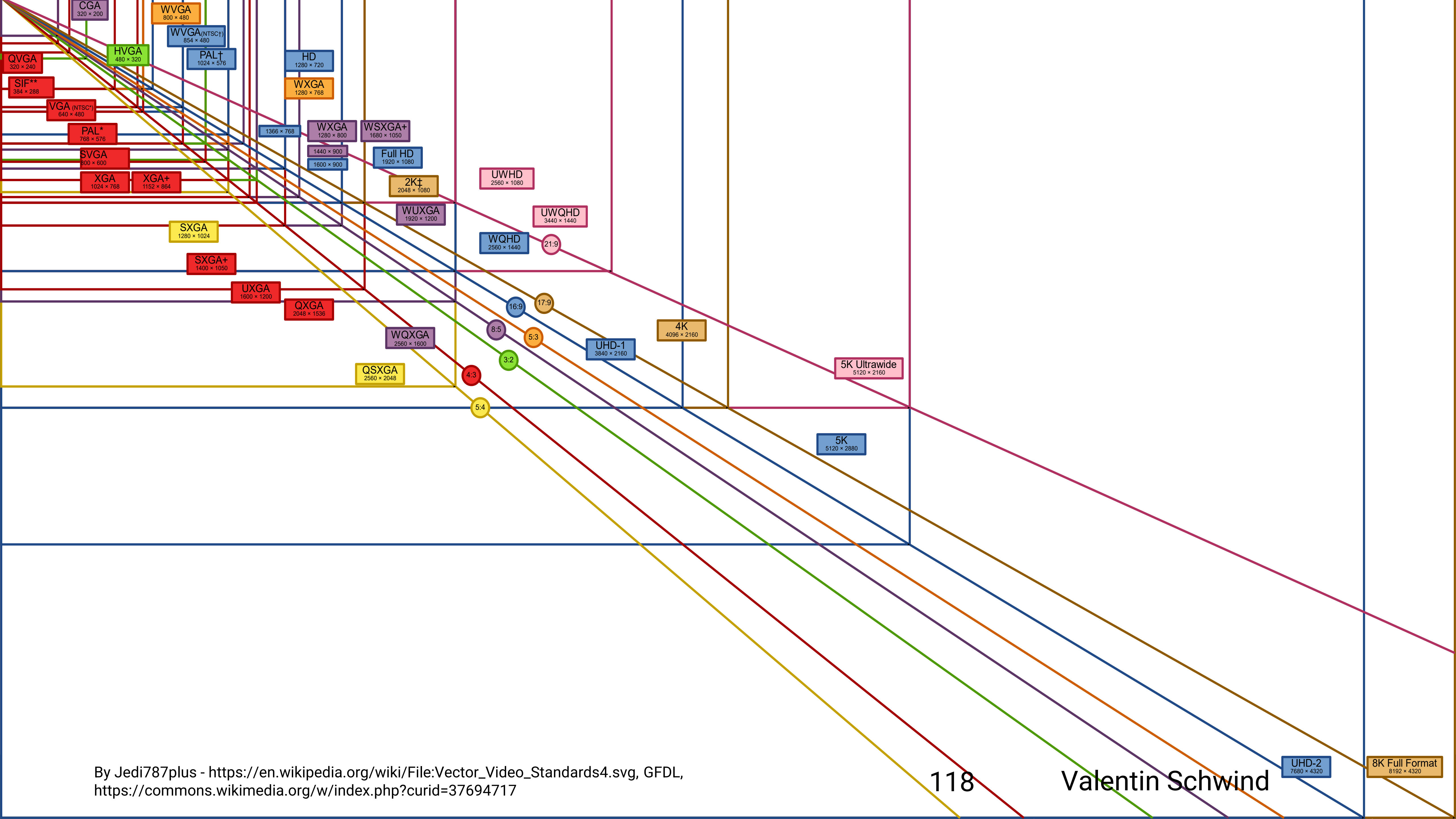
Screens? Displays? Monitors?

- **Electronic visual display = "screen"**: a device for **presentation of images, text, or video** transmitted electronically, without producing a permanent record
- **Monitor**: Any device composed by a (screen/display) and its case (plastic case, embedded in a panel) as **a peripheral device**.
- **Television Set (TV)**: Any monitor working as **a standalone device**.
- **Screen Technologies:**
  - **Active emit** light: LCD + backlight, TFT, LED, PDP, CRT
  - **Passive modulate** light: LCDs that use a simple grid to supply the charge: LC, LCD, E-Paper, E-Ink, IMODs
    - OLEDs emit light and can be both: PMOLED (passive matrix) or AMOLED (active matrix)
- Not all displays are visual displays!

Displays/Screens

- Screen Properties
  - **Size**: physical size, often the diagonal in inch, e.g., 65”
  - **Aspect ratio**: width/height, e.g., 4:3, 16:9, or 21:9
  - **Resolution**: number of pixel, width x height, e.g., 1920x1080
  - **Half- or full images**: progessive (p) or interlaced (i)
  - **Pixel density**: how close are pixels together, size of pixels, pixels per inch, dots per inch, e.g., 320ppi)
  - **Color depth**: how many colors, per color, e.g., 8-bit / 10-bit
  - **Color gamut**: which colors
  - **Refresh rate**: related to images per second, 60 Hz, 100Hz, 144 Hz, 200Hz, …

A macro shot of pixels on a screen. Image by maurobeltran. https://www.shutterstock.com/blog/inches-to-pixels-resize-image-quality



By Jedi787plus - https://en.wikipedia.org/wiki/File:Vector_Video_Standards4.svg, GFDL, https://commons.wikimedia.org/w/index.php?curid=37694717

“Retina Displays”

- Term coined by Apple: Example iPad 2 vs. iPad 3
  - 1024 × 768 pixel (132 ppi) vs. 2048 × 1536 pixel (264 ppi)
- Angular resolution of the eye is about 1 arcminute ~ 0.02°
- Assume the following viewing angle:
  - 60° ~ requires 3.000 pixel
  - 120° ~ requires 6.000 pixel

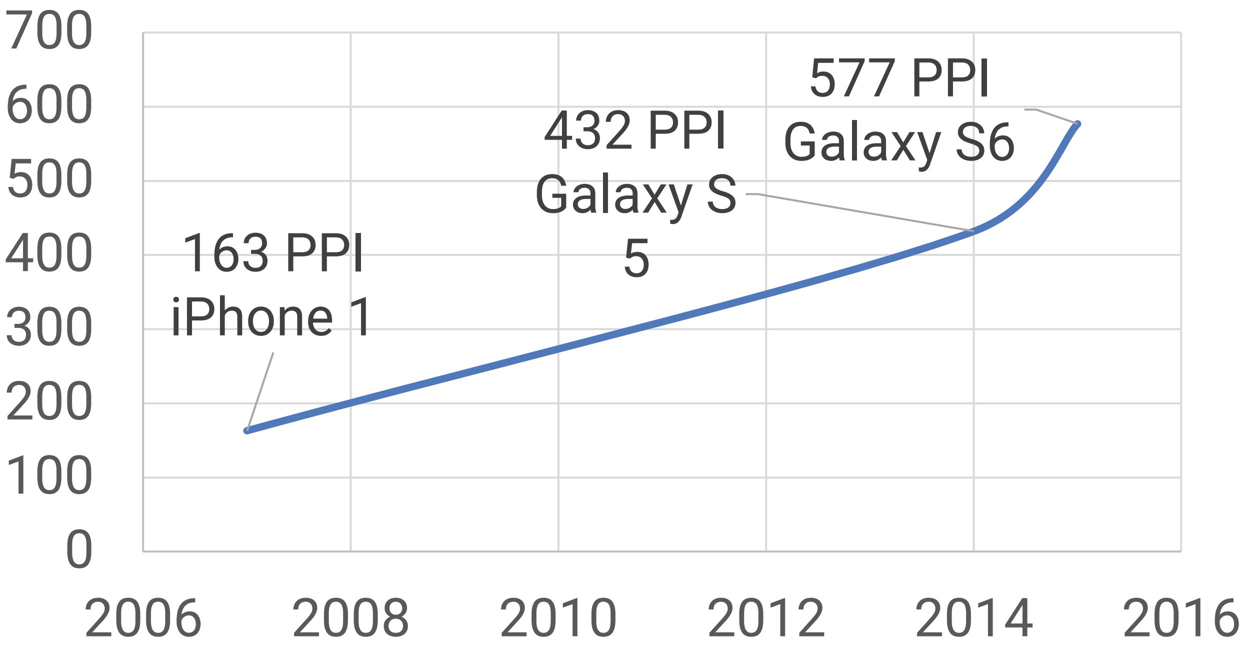
… hence 8K will be enough (with a reasonable viewing distance).

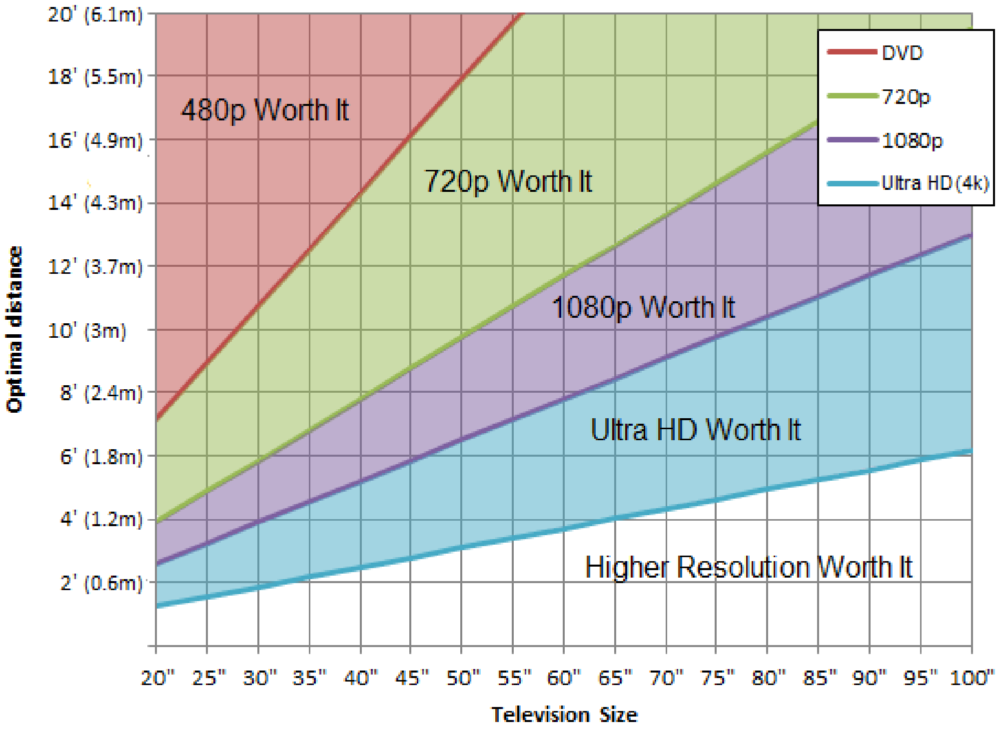
iPad2

iPad3

Pixel Density

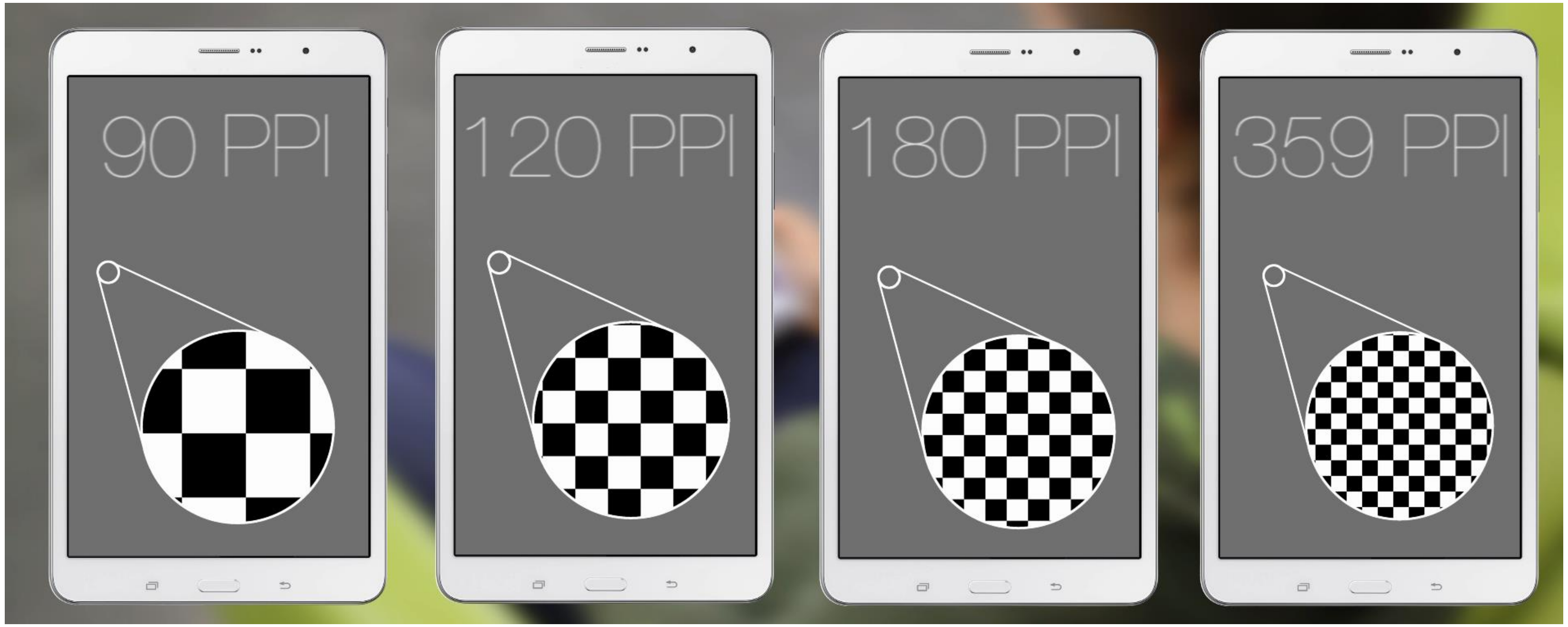
- Traditionally 72dpi (Apple) and 96dpi (Windows), rapid increased the last years
- Depends on viewing distance





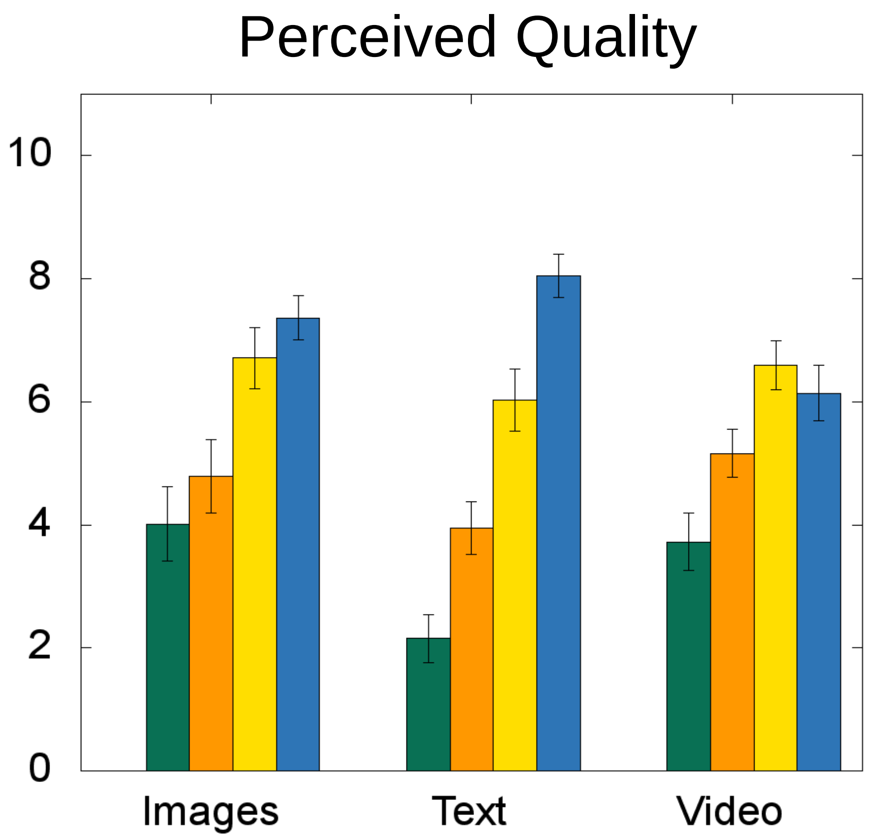
Pixel Density: An HCI Experiment

- Three Tasks: image search, word counting, video analysis
- 16 participants, four Screen Resolutions



Lischke, L., Mayer, S., Wolf, K., Sahami Shirazi, A., & Henze, N. (2015, April). Subjective and objective effects of tablet's pixel density. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2769-2772).

Pixel Density: An HCI Experiment

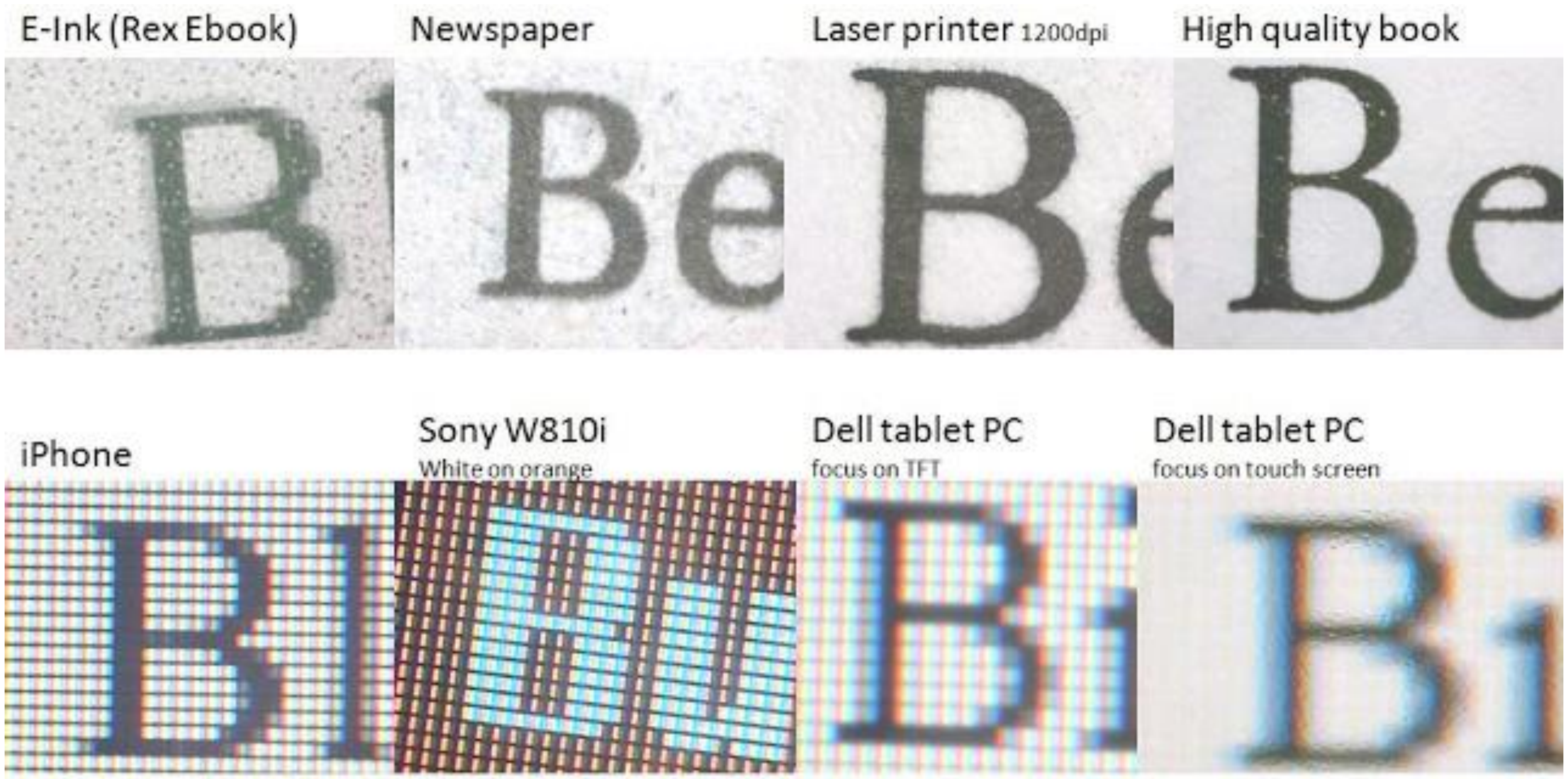


But…

- No effects of distance!
- No effects on task completion time!
- No effects on error rate!

Lischke, L., Mayer, S., Wolf, K., Sahami Shirazi, A., & Henze, N. (2015, April). Subjective and objective effects of tablet's pixel density. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2769-2772).

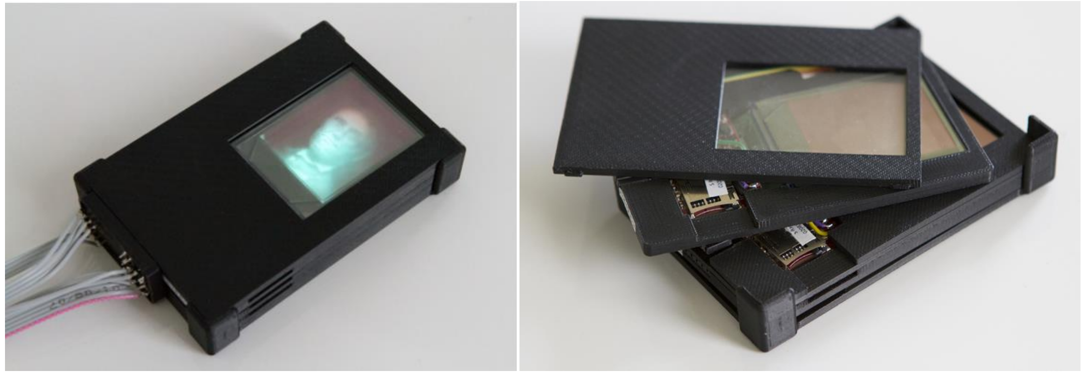
E-Paper vs. Paper Displays, other Displays



Visual Display Technologies

- **E-Paper:** slower update rate, black and white or few colors, readable outdoors, require light (like paper)
- **Projection**: resolution, brightness (daylight, outdoors), noisen projection distance, lens and image correction, connectivity, size and weight
- **Head-Mounted Displays:** stereoscopic, resolution, (video) see-through, light field
- **Volumetric**: layered, laser, mid-air, drones

Pascal Knierim, Funk, Markus, Tamara Müller, Marc Weise, Benjamin Schopf, Albrecht Schmidt "HMD++: Extending the Field of View of Head-Mounted Displays using Projection“, 2016.

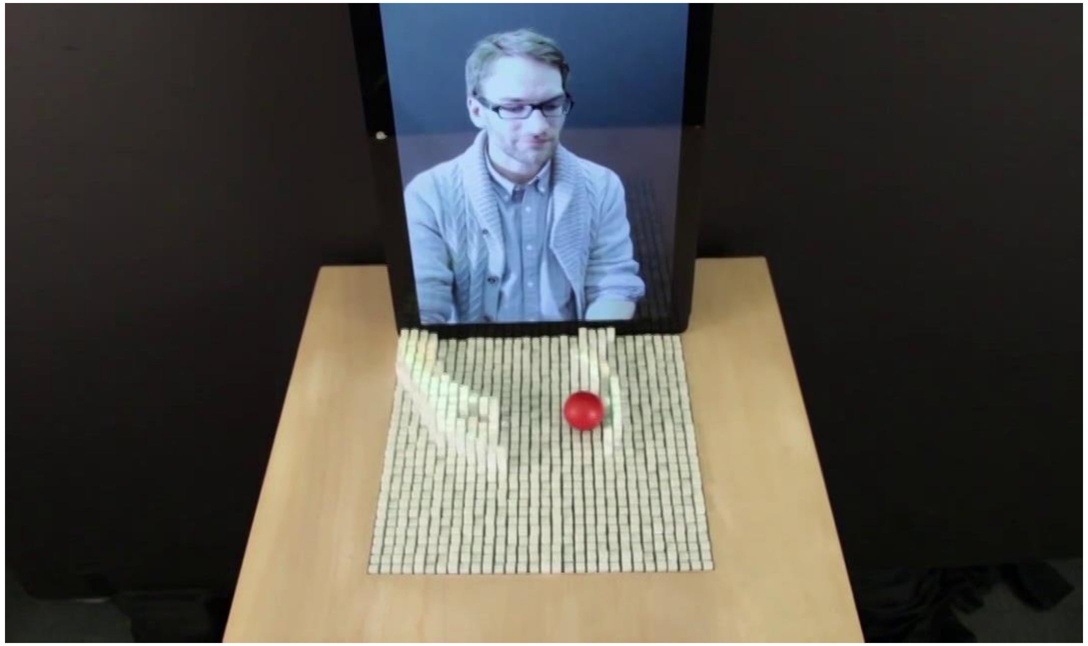


Patrick Bader, Stefan Schneegass, Niels Henze, Valentin Schwind, and Katrin Wolf. 2014. A mobile see-through 3D display with front- and back-touch. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14). Association for Computing Machinery, New York, NY, USA, 1063–1066. DOI:https://doi.org/10.1145/2639189.2670276

3D Displays

- Real physical objects

Projection on sand. Image from: https://www.youtube.com/watch?v=YdxupWDNYy4



inFORM - Interacting With a Dynamic Shape Display, so users can interact with digital information in a tangible way. inFORM can also interact with the physical world around it https://laughingsquid.com/inform-3d-display-can-form-moving-shapes-and-interact-with-physical-objects/

Stereoscopic Displays

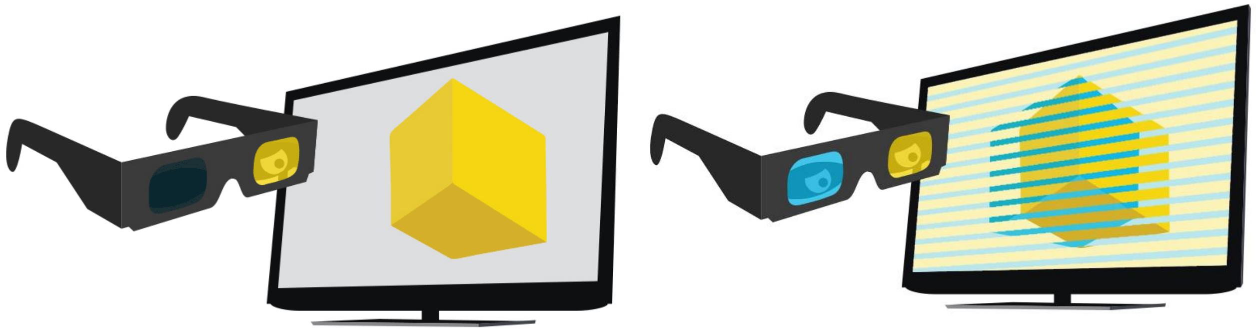
- Providing a display for each eye (headset)

By Alessandro Nassiri (CC BY-SA) wikimedia.org

User Psoreilly on en.wikipedia (CC BY-SA)

Stereoscopic Displays

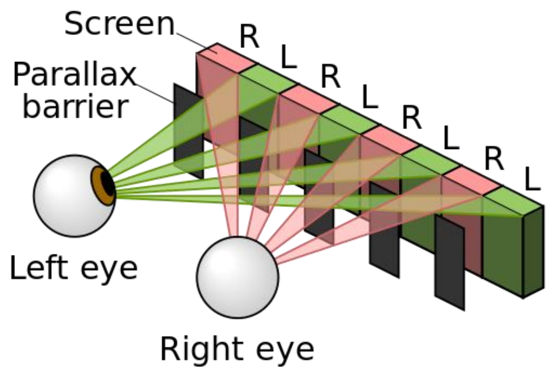
- Overlaying images, a separating them for each eye
  - Time synchronized, e.g., shutter glasses
  - Polarization filter on projector and glasses
  - Color anaglyph systems

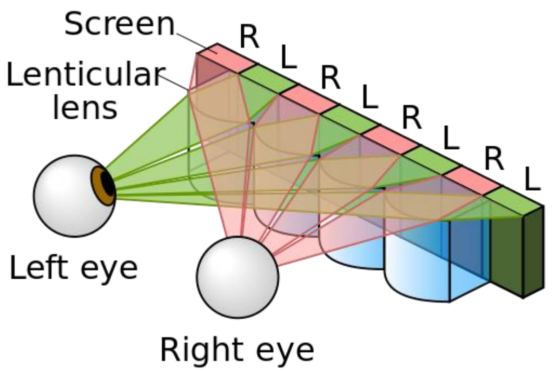


By Locafox.de (CC BY-SA) wikimedia.org

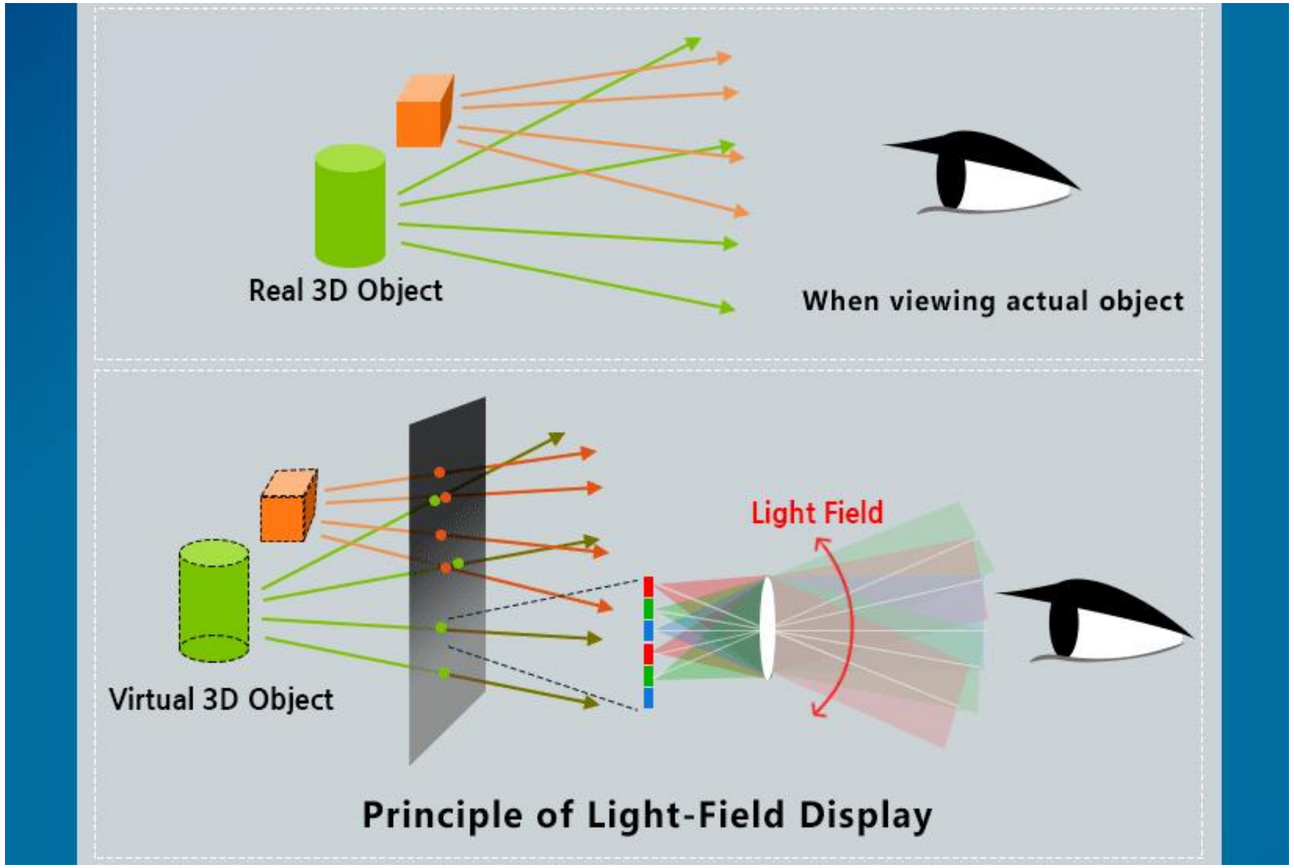
Snaily (CC BY-SA)

Autostereoscopic Displays





By Cmglee (CC BY-SA) wikimedia.org



Light-Field Display https://global.samsungdisplay.com/28567/

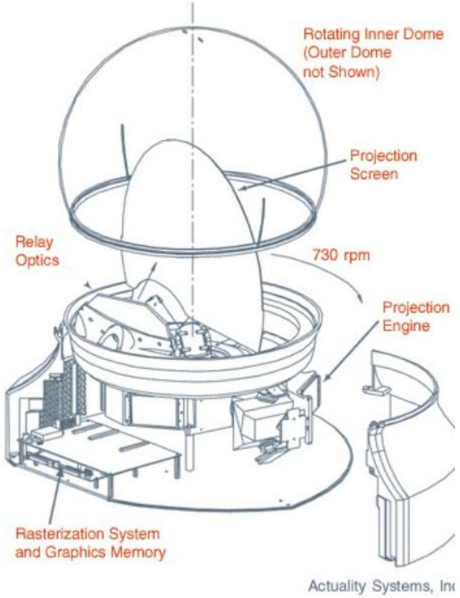
Eyeglasses-free Displays

F. Huang, G. Wetzstein, B. Barsky, R. Raskar. Eyeglasses-free Display: Towards Correcting Visual Aberrations with Computational Light Field Displays. ACM Proc. of SIGGRAPH (Transactions on Graphics 33, 4), 2014. Images from https://web.media.mit.edu/~gordonw/VisionCorrectingDisplay/

Volumetric Displays

Voxon VX1 3D Volumetric Display – Demonstration
https://www.youtube.com/watch?v=FVYoWsxqK8g

Tovi Grossman, Dhttps://youtu.be/FVYoWsxqK8ganiel Wigdor, and Ravin Balakrishnan. 2004. Multi-finger gestural interaction with 3d volumetric displays. In Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST '04). ACM, New York, NY, USA, 61-70.



Swept-screen multiplanar volumetric display

198 2-D slices

768 x 768 pixel slice resolution

100 million voxels

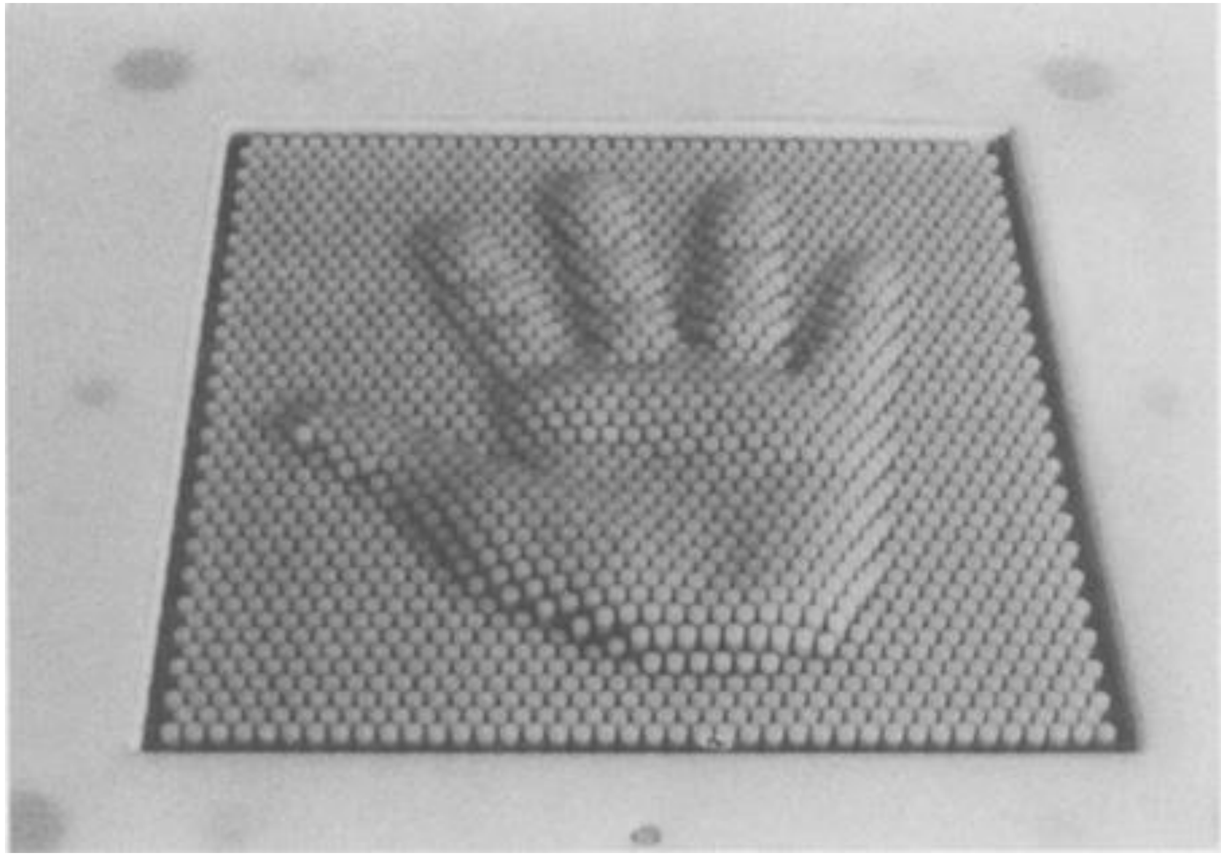
24 Hz volume refresh

Viewing Angle: 360º horizontal, 270º vertical

Volumetric Touchscreens

Patrick Bader, Stefan Schneegass, Niels Henze, Valentin Schwind, and Katrin Wolf. 2014. A mobile see-through 3D display with front- and back-touch. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14). Association for Computing Machinery, New York, NY, USA, 1063–1066. DOI:https://doi.org/10.1145/2639189.2670276

Tactile Displays

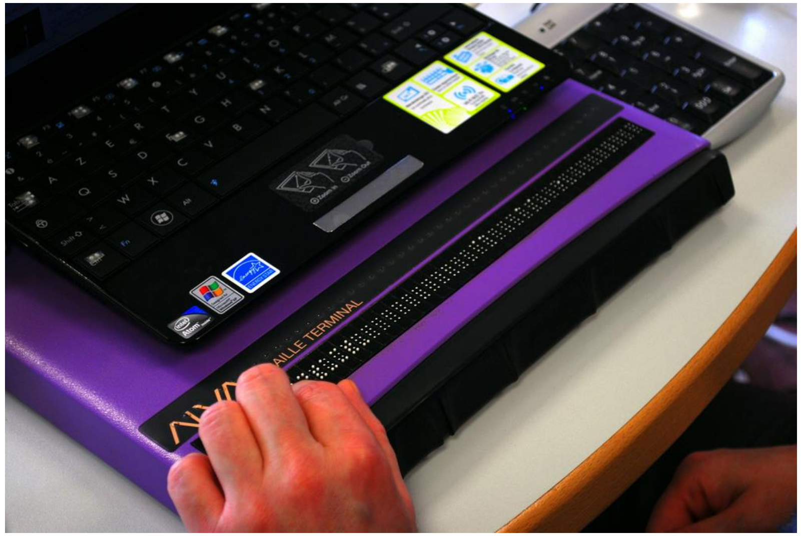


Shimizu, Yutaka & Saida, Shinya & Shimura, Hiroshi. (1993). Tactile pattern recognition by graphic display: Importance of 3-D information for haptic perception of familiar objects. Perception & psychophysics. 53. 43-8. 10.3758/BF03211714.

Informationszugang für Blinde und Sehbehinderte
GRANT - GRafische ANwendungen auf Taktilen Displays Image from: https://www.uni-potsdam.de/de/multimedia/projekte/anwendungen/grant

Refreshable Braille Displays

- **Electro-mechanical device for displaying braille characters**, usually with pins raised through holes in a flat surface
  - **For visually impaired computer users** who cannot use a standard computer monitor can use it to read text output
  - Often combined with a traditional braille keyboards
  - **Requires a screen reader software (or support from the OS)** to output the content of the screen and to convert it into braille characters



Refreshable braille display, Sebastien.delorme, CC BY-SA 3.0 https://en.wikipedia.org/wiki/Refreshable_braille_display#/media/File:Plage-braille.jpg

Image from https://www.linzy.ru/otvety-na-voposy/kakie-ustrojstva-dlya-slepyh-i-slabovidyashhih-lyudej-sushhestvuyut.htm

Olfactory & Gustatory Displays

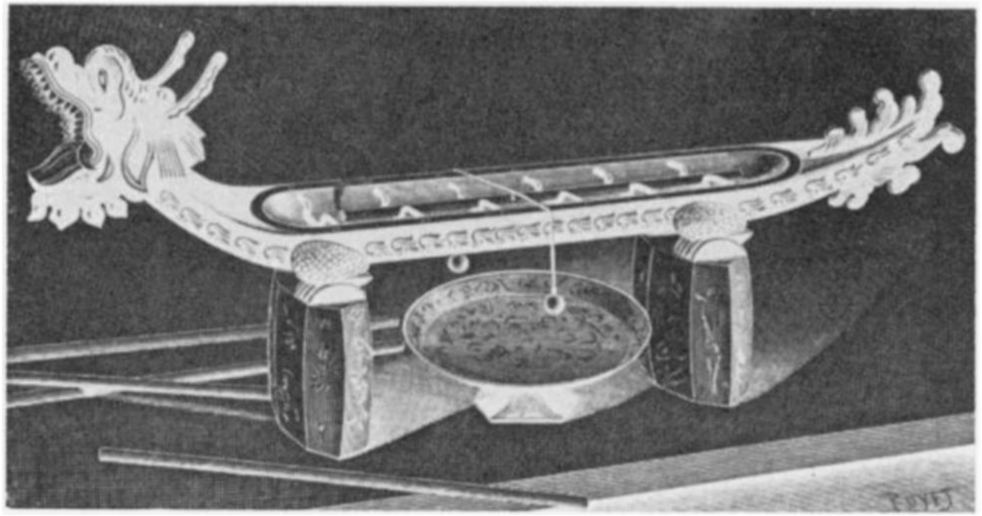
- **An olfactory or gustatory display uses scents** and is useful for slowly-moving, medium-duration information or information
  - Parameters of scent delivery: intensity, distance, volume, speed
- Key challenges
  - Creating a certain smell
  - Storing and providing the smell
  - Timed delivery of the smell to the user
  - Neutralizing / replacing the smell

"Meta Cookie" is a gustatory display to change the perceived taste of a cookie by overlaying visual and olfactory information onto a real cookie, The system induces cross-modal effect among vision, olfaction and gustation to let humans perceive various tastes without changing chemical substances

Dmitrenko, D., Vi, C. T., & Obrist, M. (2016, October). A comparison of scent-delivery devices and their meaningful use for in-car olfactory interaction. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 23-26)

Incense Clocks

- Timekeeping device (China, East Asia)
- Clock body holds incense sticks or powdered incense
- Different incenses along the body
- Calibrated brining rate to measure time
- Can include bells (auditory signal)
- Used in homes and temples
- New smell / other incense signals passage of time
- …you can smell the time



Bedini, S. A. (1963). The scent of time. A study of the use of fire and incense for time measurement in oriental countries. Transactions of the American Philosophical Society, 53(5), 1-51.

Tongue Displays

- Tongue Display Units (TDUs) are programmable pulse generators that delivers DC-balanced voltage pulses suitable for electrotactile (electrocutaneous) stimulation of the anterior-dorsal tongue
- Use a matrix of surface electrodes for sensory substitution and neurorehabilitation

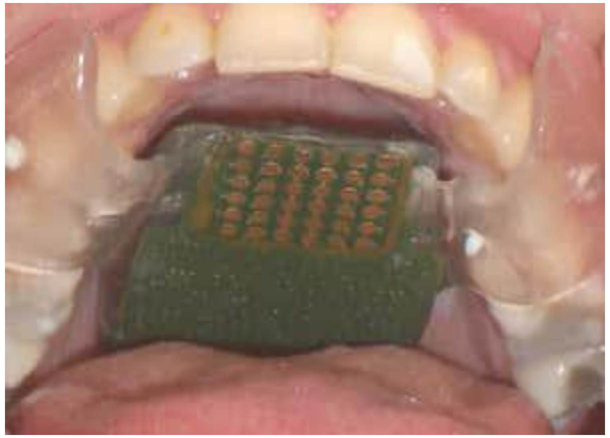
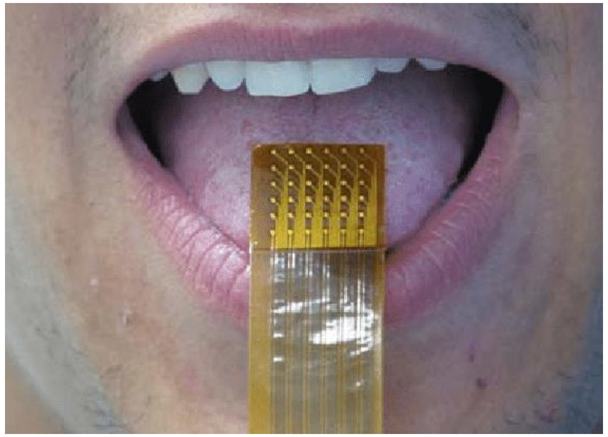


Image from: https://www.newscientist.com/article/dn14058-tongue-display-helps-you-keep-your-balance/



Vuillerme, N., Boisgontier, M., Chenu, O., Demongeot, J., & Payan, Y. (2007). Tongue-placed tactile biofeedback suppresses the deleterious effects of muscle fatigue on joint position sense at the ankle. Experimental Brain Research, 183(2), 235-240.

K.A. Kaczmarek, The tongue display unit (TDU) for electrotactile spatiotemporal pattern presentation, Scientia Iranica, Volume 18, Issue 6, 2011, Pages 1476-1485, ISSN 1026-3098, https://doi.org/10.1016/j.scient.2011.08.020.

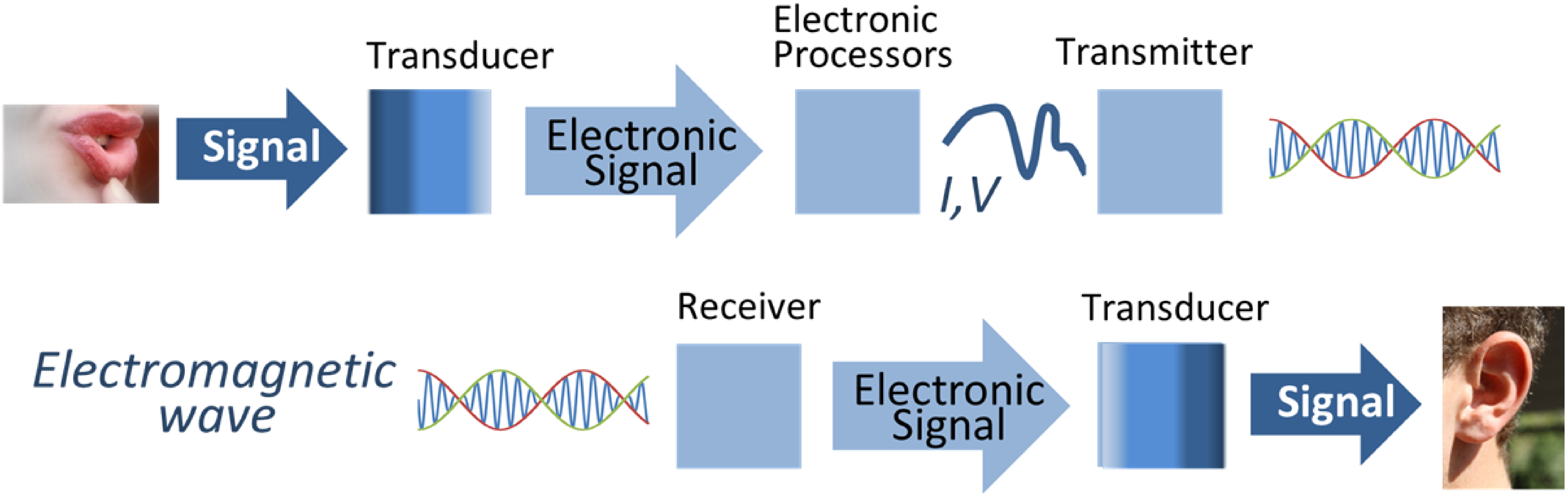
Audio

Speakers and Auditory Cues

Image from: https://pxhere.com/de/photo/598757

Loudspeaker

- **An electroacoustic transducer** converts an audio signal into a sound



Transducers are used in electronic communications systems to convert signals of various physical forms to electronic signals, and vice versa. In this example, the first transducer could be a microphone, and the second transducer could be a speaker. Brews ohare - Own work;... Uses File:EAR.jpg and File:Sexy-Red-Lips Happy-Girl.jpg Steps in a signal communications system CC BY-SA 3.0 https://en.wikipedia.org/wiki/Transducer#/media/File:Signal_processing_system.png

Audio Output Technologies

- Speakers: Mono, stereo, multi-speaker system (5.1, 7.1,…)
- Headphones: Over-ear/On-ear, In-ear/Earbuds, Bone Conduction
- Important Parameters
  - Frequency range and Frequency response
  - Volume range
  - Directionality
  - Physical setup and size
  - Connections and data transmission

https://pixabay.com/de/photos/menschen-frau-kopfh%C3%B6rer-musik-ton-2572972/
https://pixabay.com/de/photos/techland-audio-audiovector-stereo-1822627/

Audio Output Types

- **Simple sounds**
  - Beeps, e.g., warnings
  - Single frequency
- **Designed sounds**
  - Auditory icons
  - Earcons
  - Composed / Music
- **Sonification**
  - Turning other media into sound
- **Existing sounds**
  - Speech, nature, music

https://pixabay.com/de/photos/noten-musik-notenschl%C3%BCssel-4768600/

Auditory Icons vs. Earcons

- **Auditory Icons**
  - “informative sounds as an auditory icon” (Gaver, 1986)
  - “Gaver’s auditory icons have been used in several systems […] These use environmental sounds that have a semantic link with the object or action they represent.” (Brewster et al.,1994)
- **Earcons**
  - “We call such structured sounds earcons, which are defined as nonverbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction. Examples of computer objects are files, menus, and prompts. Editing, compiling, and executing are examples of operations. An example of an interaction between an object and an operation is editing a file.” (Blattner et al. 1989)

Auditory Icons

	TIME	SPACE
SOUND	**Sound exists <u>in</u> time.** • Good for display of changing events. • Available for a limited time.	**Sound exists <u>over</u> space.** • Need not face source. • A limited number of messages can be displayed at once.
VISION	**Visual objects exist <u>over</u> time.** • Good for display of static objects. • Can be sampled over time.	**Visual objects exist <u>in</u> space.** • Must face source. • Messages can be spatially distributed.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human-computer interaction, 2(2), 167-177.

Auditory Icons

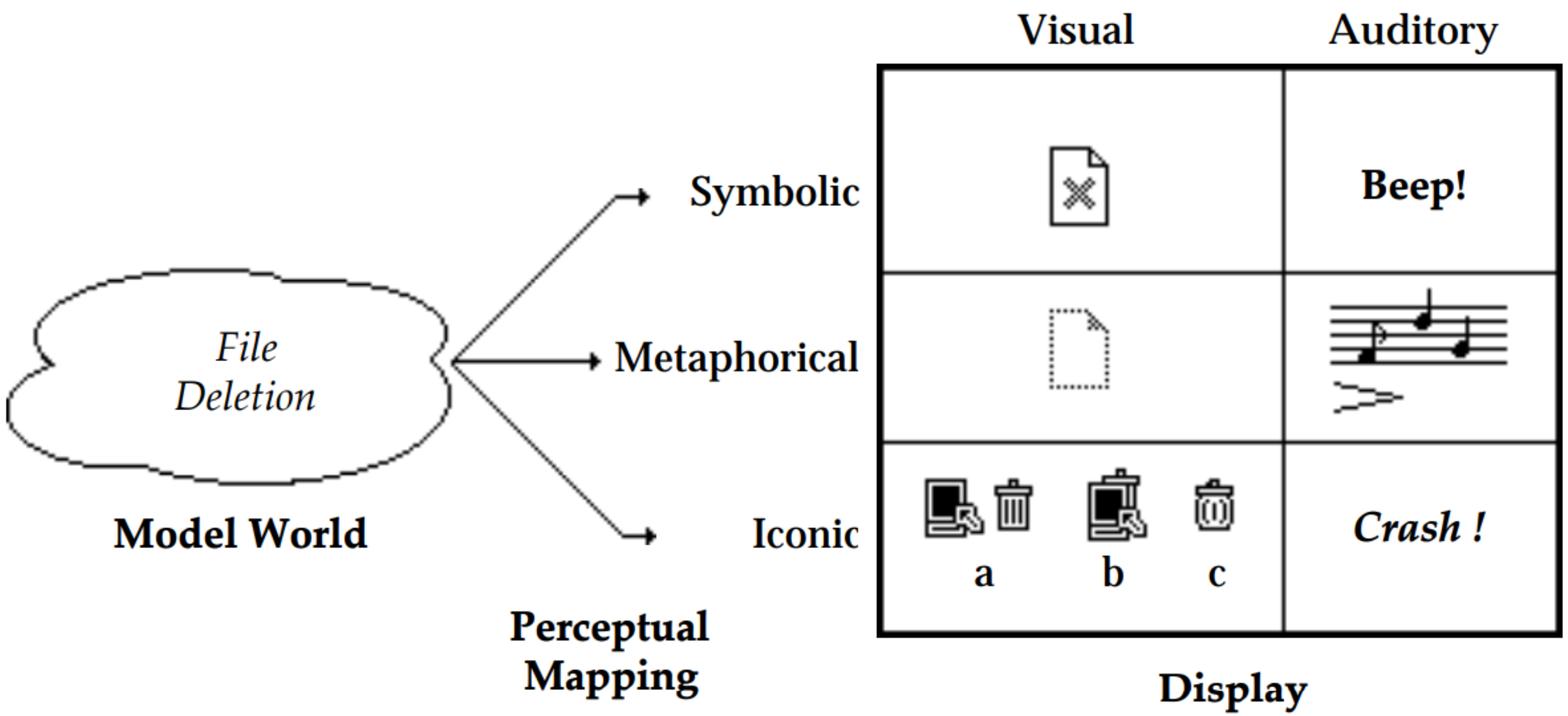
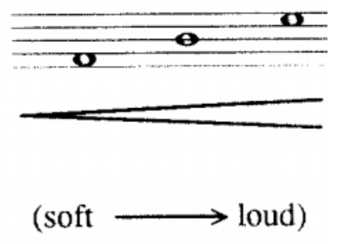


Figure 4. A conceptual file deletion may be mapped to a display in many different ways. Here six possibilities are shown, one visual and one auditory example each of symbolic, metaphorical, and iconic mappings between the event and the display.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human-computer interaction, 2(2), 167-177.

Earcons

- “One-element earcons may be **digitized sounds, a sound created by a synthesizer, a single note, or a motive**. An element may be compared to a word, whereas a note may be compared to a letter of the alphabet.”
- “A single-motive earcon has the attributes **of rhythm, pitch, timbre, register, and dynamics**. Because single-motive earcons are relatively simple, they can represent basic, common computer entities such as certain error messages, system information, windows, and files.”
- “The user hears the click each time a character is deleted.”



An earcon with dynamics that change from soft to loud.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. Human–Computer Interaction, 4(1), 11-44.

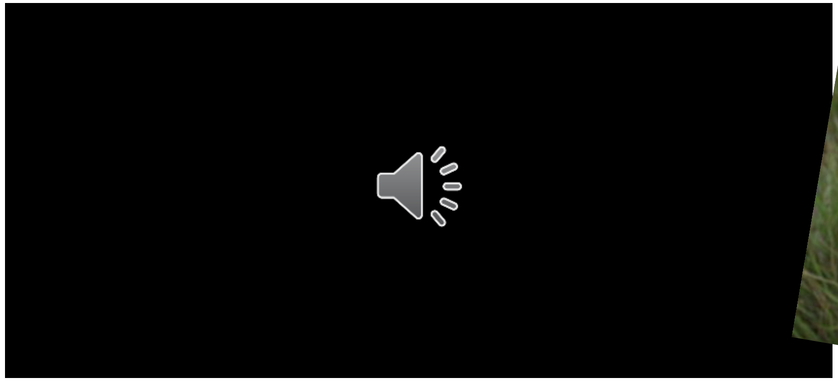
Sonification

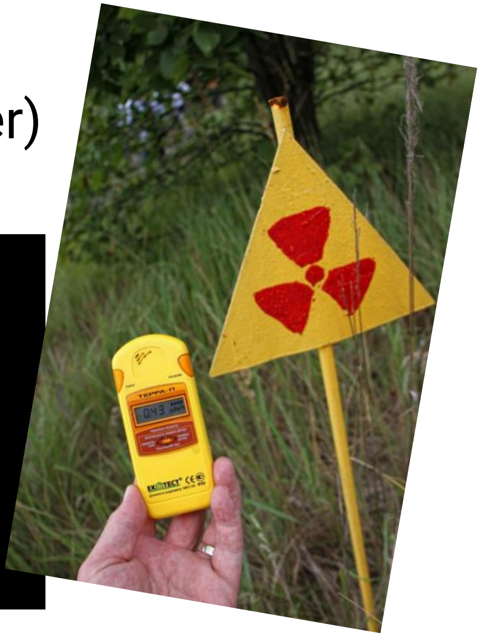
- **Maps non-speech audio data to sounds** where a sound dimension is chosen to represent a given data dimension.
  - Makes data and information audible, alternative or complement to visualizations
- **Typical acoustic dimensions**: pitch/frequency, amplitude/volume, tempo/duration, spatial arrangements
- Used for
  - Alerts, notifications, status and progress indicating functions
  - Data exploration functions
  - Auditory graphs and interactive maps
  - Art and entertainment

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. Human–Computer Interaction, 4(1), 11-44.

Sonicfinder

- Creating sounds that can be linked to events
- Sound helps to discriminate
  - Objects
  - Activities
  - Properties (e.g. size, number)





Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human-computer interaction, 2(2), 167-177.

MAPPING EVENTS TO SOUND IN THE SONICFINDER

FINDER EVENTS	AUDITORY ICONS
Objects	
Selection	Hitting sound
Type (file, application, folder, disk, trash)	*Sound source (wood, metal, etc.)*
Size	*Frequency*
Opening	Whooshing sound
Size of opened object	*Frequency*
Dragging	Scraping sound
Size	*Frequency*
Where (windows or desk)	*Sound type (bandwidth)*
Possible Drop-In?	*Selection sound of disk, folder, or trashcan*
Drop-In	Noise of object landing
Amount in destination	*Frequency*
Copying	Pouring sound
Amount completed	*Frequency*
Windows	
Selection	Clink
Dragging	Scraping
Growing	Clink
Window size	*Frequency*
Scrolling	Tick sound
Underlying surface size	*Frequency*
Trashcan	
Drop-in	Crash
Empty	Crunch

Biofeedback

Human-Computer Interaction Lecture

Image from: https://pxhere.com/de/photo/598757

Biofeedback

- Use of sensors and displays to provide users with real-time information about their physiological state
- Grounded in medical research to **train individuals to improve and gain control about bodily processes** that normally happen involuntarily, such as heart rate, muscle tension, and skin temperature
- Key Components of Biofeedback in HCI:
  - **Input Sensors**: Devices with sensors that measure physiological signals
  - **Output Feedback**: Real-time feedback of the input signal informing about the physiological state
  - **User Interaction**: The user interacts with the system, often using the feedback to try and control or alter their physiological responses

Input Channels

- **Heart Rate Monitors**: Measure the heart rate, often using photoplethysmography (PPG) sensors found in smartwatches or chest straps.
- **Electroencephalography (EEG) Devices**: Record brainwave patterns, used in neurofeedback.
- **Electromyography (EMG) Sensors**: Measure muscle activity and tension.
- **Galvanic Skin Response (GSR) Sensors**: Detect changes in skin conductance related to sweat gland activity, often used as a stress indicator.
- **Respiration Sensors**: Monitor breathing rate and pattern.
- **Temperature Sensors**: Measure skin temperature, which can change with stress and blood flow.
- **Eye Trackers**: Track eye movements and blinks, sometimes used in relaxation and focus training.

Output Channels

- **Visual Feedback**: Graphs, colors, or animations on a screen that change in response to the physiological data. For example, a color gradient that shifts from red to blue as the user's stress level decreases.
- **Auditory Feedback**: Sounds or music that vary with the physiological measurements. For instance, the tempo or pitch of music or sounds might change to reflect relaxation or stress levels.
- **Haptic Feedback**: Physical sensations such as vibrations that respond to the user's physiological state. For example, a wearable device that vibrates more intensely as stress levels rise.
- **Textual or Numeric Feedback**: Providing numerical readings or text messages explaining the physiological state or giving suggestions for improvement.

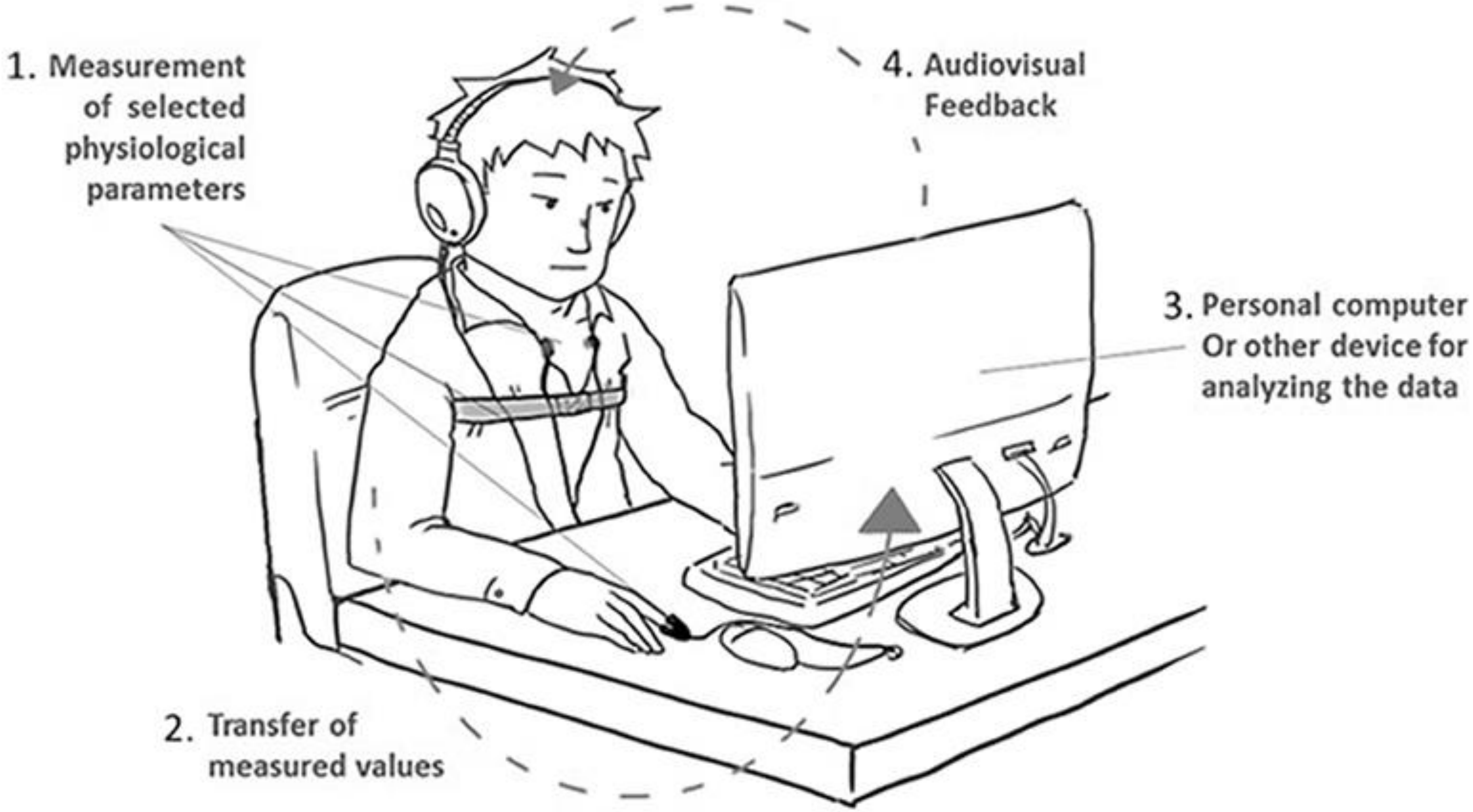
Neuroplasticity and Biofeedback

- Neuroplasticity means **the brain can rewire itself**
  - It is a fundamental aspect of how the brain adapts and evolves through learning and memory
- **Biofeedback can** be a helpful tool to **harness and support neuroplasticity,** particularly in the following ways:
  - By consciously observing and responding biofeedback, individuals can learn to control or influence these processes (e.g., after paralysis)
  - Can be helpful to regain control of motor functions (e.g., after stroke, brain injuries)
  - Can train individuals to modify their brainwave patterns and improve cognitive functions and emotional regulation (e.g., ADHD, anxiety, and depression)

Applications of Biofeedback in HCI:

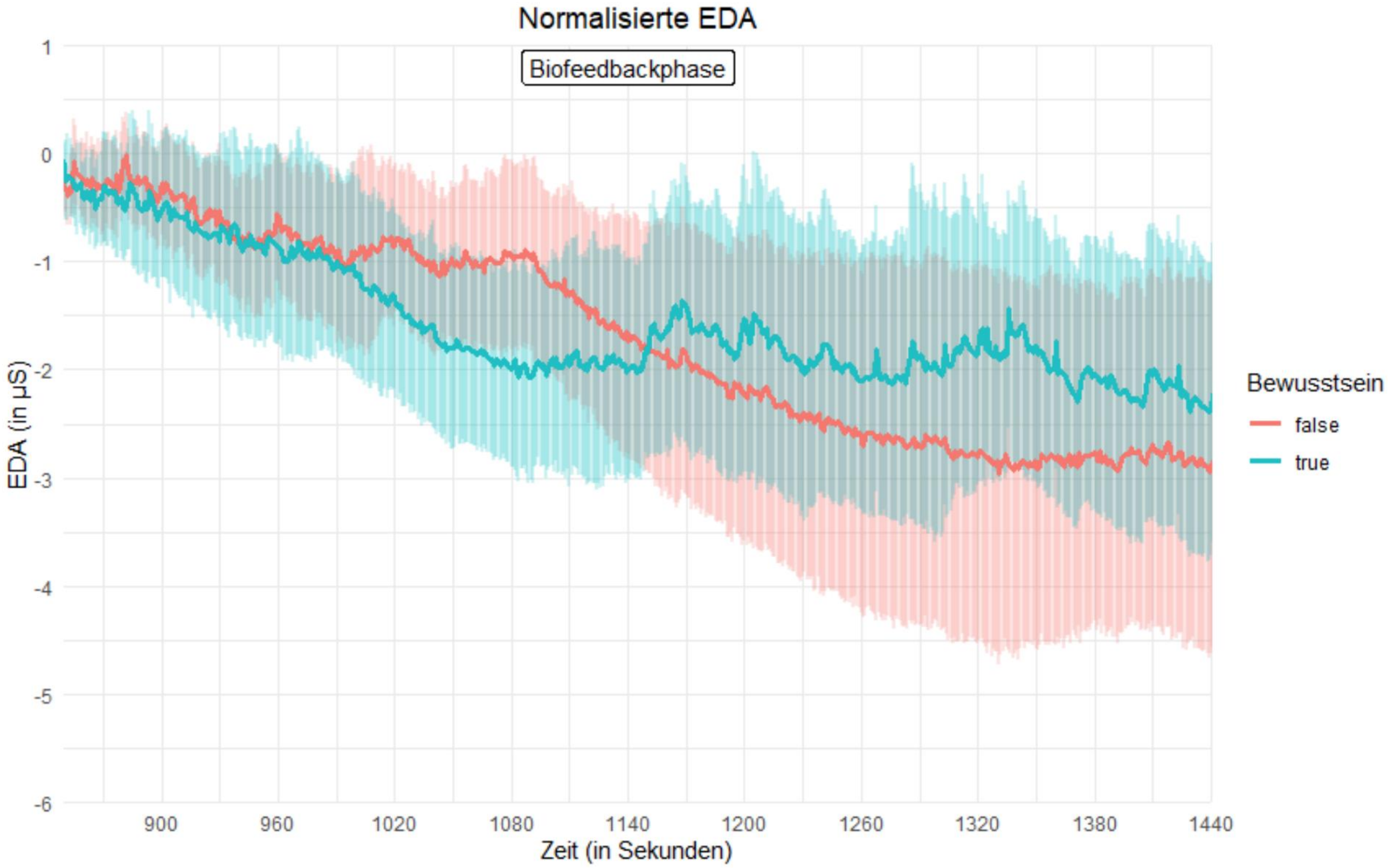
- **Health and Wellness**: Stress management, relaxation training, and treating various health issues like hypertension, chronic pain, and anxiety disorders
- **Gaming and Virtual Reality**: Biofeedback can enhance gaming or VR experiences by adjusting the environment or difficulty in response to the user's physiological state
- **Performance Enhancement**: Athletes and performers use biofeedback to improve focus, control anxiety, and optimize performance
- **Accessibility**: Biofeedback provides alternative interaction methods for individuals with disabilities
  - For example, muscle tension or eye movement-based systems can enable communication for people with severe motor impairments.

Example: Biofeedback for Stress Management

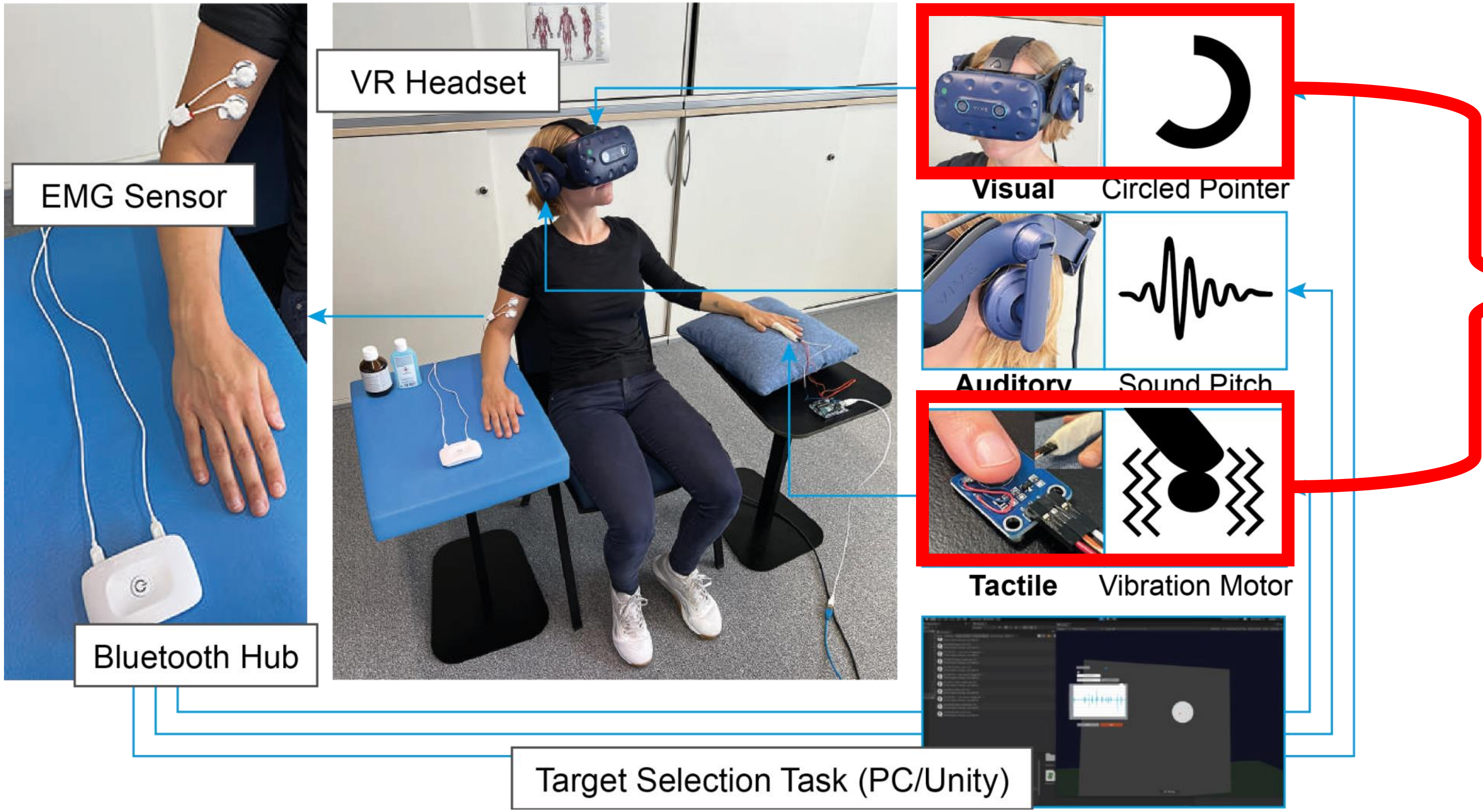


Yu, B., Funk, M., Hu, J., Wang, Q., & Feijs, L. (2018). Biofeedback for Everyday Stress Management: A Systematic Review. Front. ICT, 5, 297761. doi: 10.3389/fict.2018.00023

Biofeedback Awareness



Biofeedback Modalities and EMG Interaction



Combined input increases the bandwidth

Shared Biofeedback

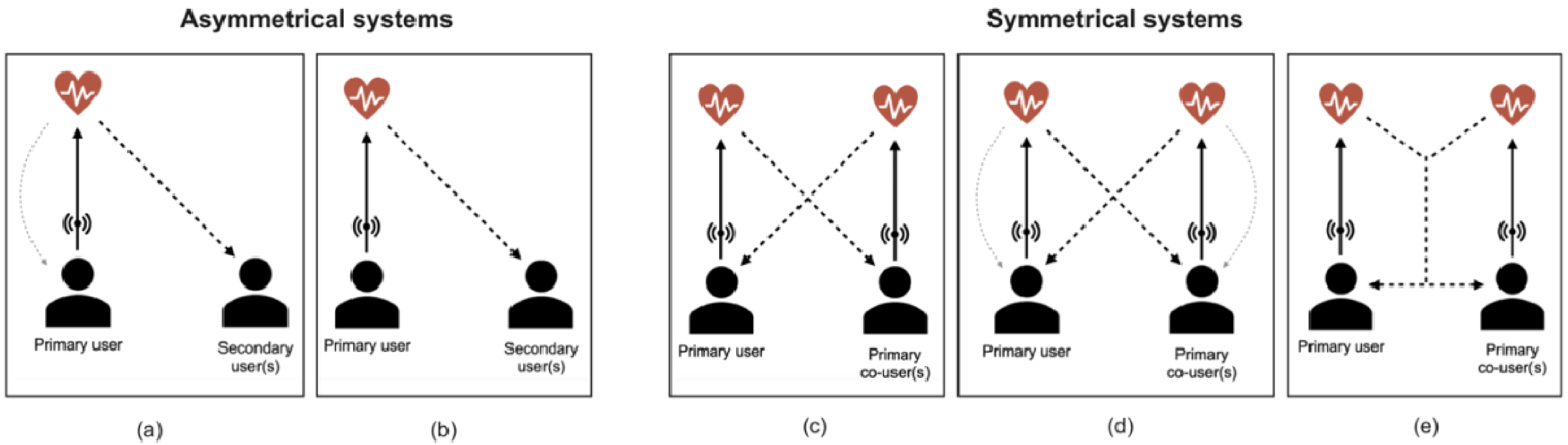


Figure 2. Social biofeedback data flows in (a, b) asymmetrical and (c - e) symmetrical systems. The heart icon

References

- Card, S. K., Mackinlay, J. D. and Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. ACM Transactions on Information Systems 9(2 April): 99-122
https://dl.acm.org/doi/pdf/10.1145/123078.128726
- Buxton, W. (1983). Lexical and Pragmatic Considerations of Input Structures. Computer Graphics, 17 (1), 31-37. http://www.billbuxton.com/lexical.html
http://www.billbuxton.com/input04.Taxonomies.pdf
- Card, S. K., Mackinlay, J. D., & Robertson, G. G. (1990, March). The design space of input devices. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 117-124).
https://www.cc.gatech.edu/classes/AY2009/cs4470_fall/readings/input-design-space.pdf
- MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001, March). Accuracy measures for evaluating computer pointing devices. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 9-16).