UNIVERSITY OF CRETE
COMPUTER SCIENCE DEPARTMENT

COURSE CS-564 (OPTIONAL)
ADVANCED TOPICS
IN HUMAN – COMPUTER INTERACTION

Course Convenor: Constantine Stephanidis

Recent and emerging interaction techniques: Eye tracking, Brain-Computer interfaces

Eye Tracking

Introduction

- People use their eyes mainly for **observation** but gaze is also used to enhance **communication**
  - staring at somebody soon causes a reaction: “What? Do you want something?”
  - an intense look at a water jug may be enough to motivate someone at a dinner party to pour more water for you
- Thus the direction of gaze of a person not only allows **observation** of the world, but also reveals their **focus of visual attention**
- It is this ability to observe the visual attention of a person, by human or machine that allows eye gaze tracking for communication

Eye Tracking

- **Eye tracking** is a technique whereby an individual’s eye movements are measured to identify
  - where a person is looking at any given time
  - the sequence in which the eyes are shifting from one location to another
- Eye movements can be captured and used as
  - indicators of the user’s **visual focus of attention**
    - gaze is considered to be a good indicator of a person's attention on external objects [Stiefelhagen et al. 2001]
  - **control signals** to enable people to interact with interfaces directly without the need for mouse or keyboard input
    - this can be a major advantage for certain populations of users such as **disabled individuals**

Application Domains (1/6)

- **Neuroscience:** obtaining a complete understanding of human vision
  - Knowledge of the physiological organization of the optic tract, as well as of the cognitive and behavioral aspects of vision
  - Eye tracking is also used in research to acquire a better understanding of neurological functions and related diseases and impairments
    - For instance, investigating scan patterns and control mechanisms of eye movements can provide early indicators for such diseases as autism, Alzheimer’s and schizophrenia. Clinical areas of research include:
      - Autistic spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD)
      - Cognitive decline such as dementia, Alzheimer’s, Lewy body and Parkinson’s
      - Reading and learning difficulties such as dyslexia
      - Stroke and non-traumatic brain injury
      - Amyotrophic lateral sclerosis (ALS)
      - Schizophrenia
      - Receptive language, expressive language, or cognitive functioning

Application Domains (2/6)

- **Psychology**
  - Understanding eye-movement characteristics during:
    - Reading
    - Scene viewing
    - Search tasks
    - Natural tasks
    - Auditory language processing
    - Other information-processing tasks
  - Developmental psychology
    - Long before infants or young children can talk, eye tracking can provide detailed information about how they perceive the world. Applicability of eye tracking within the field of developmental research is broad, including studies of: developmental progression in infants' allocation of attention and interest, visual perception related to understanding and recall, ability to recognize motion signals, etc.

Application Domains (3/6)

- **Human-computer interaction**
  - Eye movements can be measured and used to enable an individual actually to interact with an interface
    - **Gaze pointing:** Users could position a cursor by **simply looking** at where they want it to go, or “click” an icon by gazing at it for a certain amount of time or by **blinking**
  - **Virtual reality** environments can also be controlled by the use of eye movements
    - The large three-dimensional spaces that users operate in often contain far-away objects that have to be manipulated
    - Eye movements seem to be the ideal tool in such a context, as moving the eyes to span long distances requires little effort compared with other control methods [Jacob & Karn, 2003]

Application Domains (4/6)

Heat map

Gaze plot

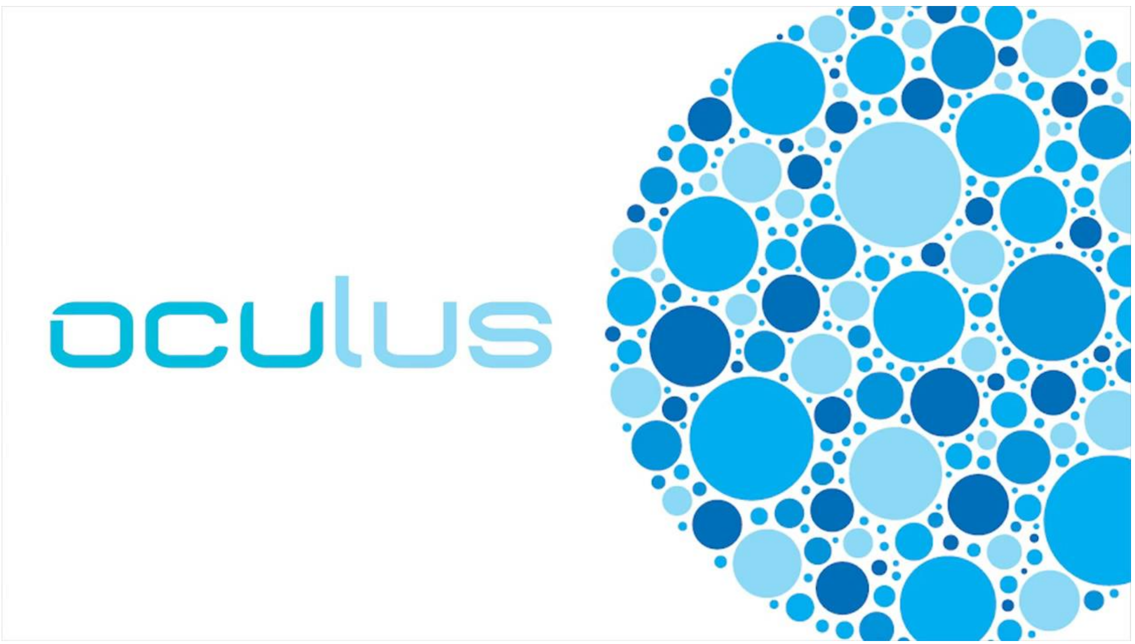
- **Research and Usability Evaluation**
  - Eye-movement recordings can provide a dynamic trace of **what** a person is looking at a visual display
  - Measuring other aspects of eye movements can reveal the **amount** of processing being applied to the focused objects
  - By defining "areas of interest" over certain parts of an interface
  - the visibility, meaningfulness and placement of specific interface elements can be objectively evaluated and the resulting findings can be used to improve the design of the interface

Application Domains (5/6)

- **Monitoring driver vigilance**
  - deficiencies in visual attention are responsible for a large proportion of road traffic accidents
  - Eye movement recording and analysis provide important techniques for understanding the **nature of the driving task** and are important for developing driver **training strategies** and accident **countermeasures**

Application Domains (6/6)

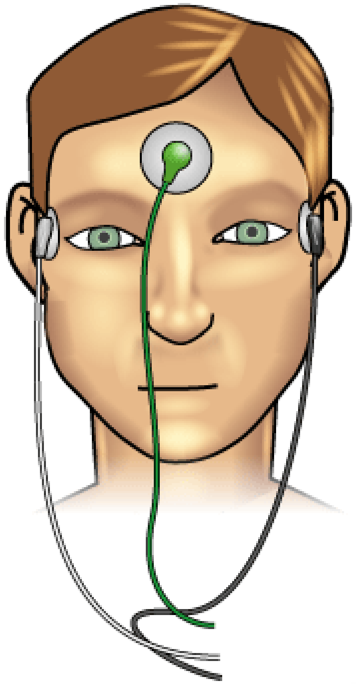
- **Marketing and advertising**
  - Eye tracking can provide insight into how the consumer disperses visual attention over different forms of advertising



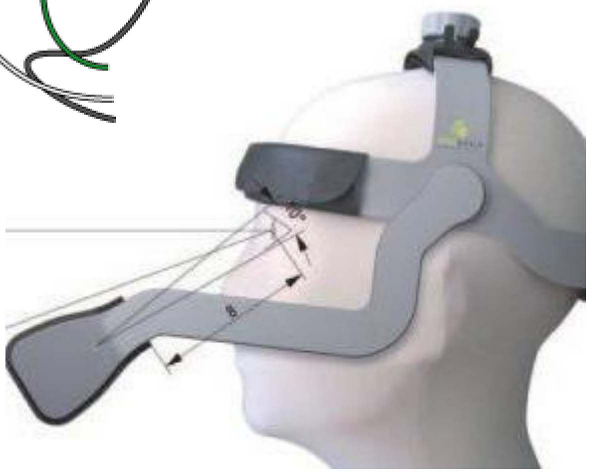
[Example of marketing eye-tracking analysis]

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

Eye-tracking Technologies



EOG

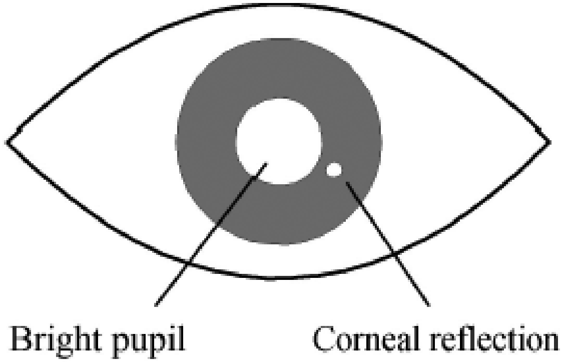


VOG

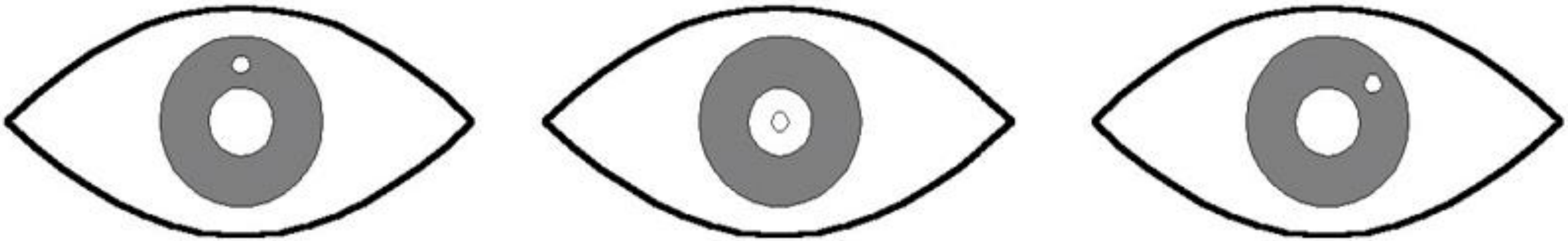
- electro-oculography (EOG)
  - the user wears small electrodes around the eye to detect the eye position
- scleral contact lens/search coil
  - the user wears a contact lens with a magnetic coil on the eye that is tracked by an external magnetic system
- video-oculography (VOG) or photo-oculography (POG)
  - still or moving images are taken of the eye to determine the position of the eye
- video-based combined pupil/corneal reflection techniques
  - extend VOG by artificially illuminating both the pupil and cornea of the eye for increased tracking accuracy
- Most of the currently available eye control systems are video based (VOG) with corneal reflection

How Does a video-based Eye Tracker Work? (1/3)

- Most commercial eye-tracking systems available today measure **point-of-regard** by the "corneal-reflection/pupil-centre" method [Goldberg & Wichansky, 2003]
- These kinds of trackers usually consist of:
  - an **infrared camera**
  - **image processing software** to locate and identify the features of the eye used for tracking
- Infrared light from a LED embedded in the infrared camera is first directed into the eye to create strong reflections in target eye features to make them easier to track
  - infrared light is used to avoid dazzling the user with visible light
- The light enters the retina and a large proportion of it is reflected back, making the pupil appear as a bright, well defined disc
- The corneal reflection is also generated by the infrared light, appearing as a small, but sharp, glint (see Figure)



How Does a video-based Eye Tracker Work? (2/3)



Directed below the camera — Directed at the camera — Directed down and to the right of the camera

- Once the image processing software has identified the centre of the pupil and the location of the corneal reflection
  - the vector between them is measured
  - and with further trigonometric calculations, point-of-regard can be found
- Although it is possible to determine approximate point-of-regard by the corneal reflection alone (see Figure), by tracking both features eye movements can, critically, be **disassociated from head movements**

How Does a video-based Eye Tracker Work? (3/3)

- Video-based eye trackers need to be fine-tuned to the particularities of each person's eye movements by a "**calibration**" process
- Calibration is accomplished by asking the user to fixate on specific markers, such as:
  - graphical dots displayed on a screen
  - printed markers for real-world setups

Eye-Movement Metrics (1/4)

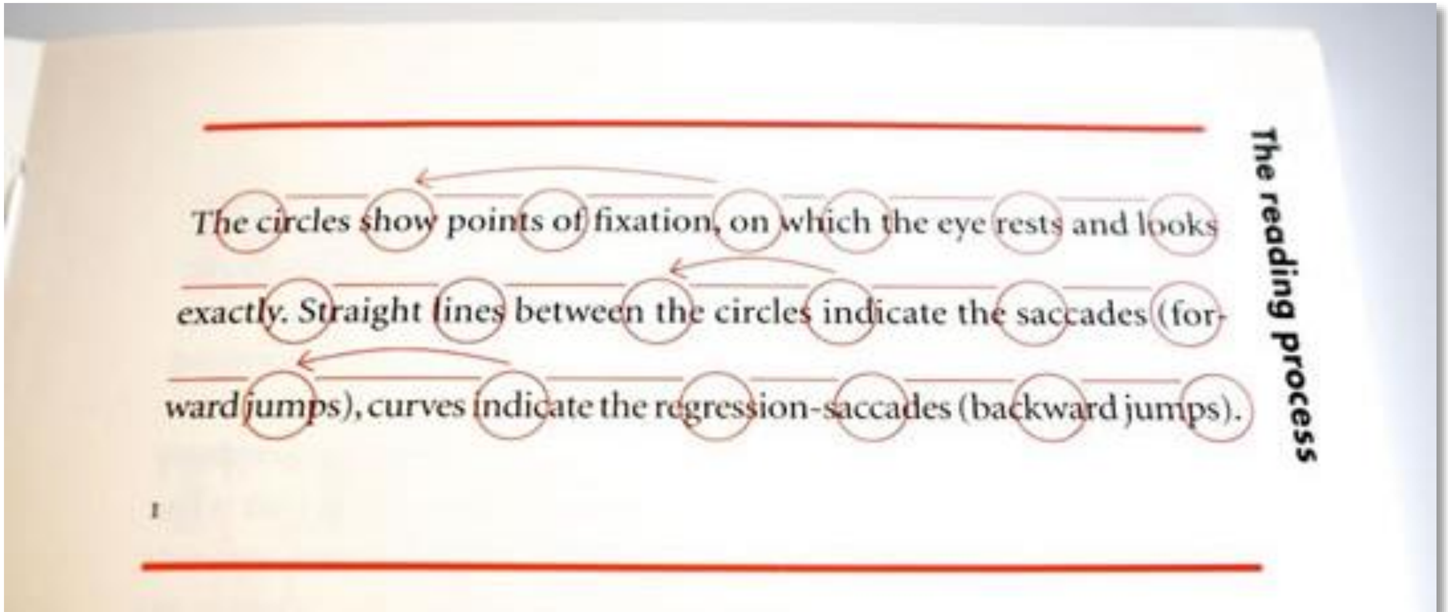
- **Fixations:** moments when the eyes are relatively stationary, taking in or “encoding” information
- Fixations can be interpreted quite differently depending on the context
  - In an encoding task (e.g., browsing a web page), higher *fixation frequency* on a particular area:
    - can be indicative of **greater interest** in the target
    - or it can be a sign that **the target is complex** in some way and more difficult to encode
  - However, these interpretations may be reversed in a search task
    - A higher number of single fixations, or clusters of fixations, are often an indicator of greater **uncertainty** in recognizing a target item
- The ***duration*** of a fixation is also linked to the processing-time applied to the object being fixated

Eye-Movement Metrics (2/4)

- **Saccades:** quick eye movements occurring between fixations
- No encoding takes place during saccades, so they cannot tell anything about the complexity or salience of an object in the interface
- However, **regressive saccades** (i.e., backtracking eye-movements) can act as a measure of **processing difficulty** during encoding
  - Most regressive saccades (or “regressions”) are very small, only skipping back two or three letters in reading tasks
  - Much larger phrase-length regressions can represent confusion in higher-level processing of the text
- Regressions could equally be used as a measure of recognition value, in that there should be an inverse relationship between the number of regressions and the salience of the phrase

Eye-Movement Metrics (3/4)

- **Scanpaths:** A scanpath describes a complete saccade-fixate-saccade sequence
- In a search task, an optimal scan path is viewed as being a straight line to a desired target, with relatively short fixation duration at the target



Eye-Movement Metrics (4/4)

- **Blink rate and pupil size:** Blink rate and pupil size can be used as an index of cognitive workload
- A lower blink rate is assumed to indicate a **higher workload**, and a higher blink rate may indicate **fatigue**
- Larger pupils may also indicate more cognitive effort
- However, pupil size and blink rate can be determined by many other factors, such as ambient light levels, so are open to contamination
  - For these reasons, pupil size and blink rate are less often used in eye tracking research

The Midas Touch Problem

- Interpreting a person's intentions from their eye movements is not a trivial task
- The eye is primarily a perceptual organ, not normally used for control, so the question arises of how **casual viewing** can be separated from **intended gaze** driven commands
- If all objects on the computer screen would react to the user's gaze, it would cause a so-called "Midas touch" (or perhaps "Midas gaze") problem:
  - **"everywhere you look something gets activated"**

Midas Touch and Selection Techniques (1/2)

- Systems that use eye tracking as an input device should be able to distinguish casual viewing from the desire to produce intentional commands, since the same communication modality (gaze) is used for both:
  - **perception** - viewing the information and objects
  - and **control** - manipulating those objects by gaze
- This way the system can avoid the Midas touch problem, where all objects viewed are unintentionally selected
- The obvious solution is to combine **gaze pointing** with **some other modality** for selection
- If the person is able to produce a separate “click”, then this click can be used to select the focused item
- This can be a separate switch, a blink, a wink, a wrinkle on the forehead or even smiling or any other muscle activity available to that person

Midas Touch and Selection Techniques (2/2)

- Using dwell time is a common approach for reducing false selections
  - **Dwell time:** a **prolonged** gaze, with a duration longer than a typical fixation
  - Most current eye control systems provide **adjustable** dwell time to avoid tiring the eyes and hinder concentration on the task
- Another solution for the Midas touch problem is to use a **special selection** area or an **on-screen button**
  - Thus, in addition to a long enough dwell time, it is beneficial to the user if eye control can be paused with for example an onscreen 'pause' command to allow free viewing of the screen without the fear of Midas touch

Other Limitations

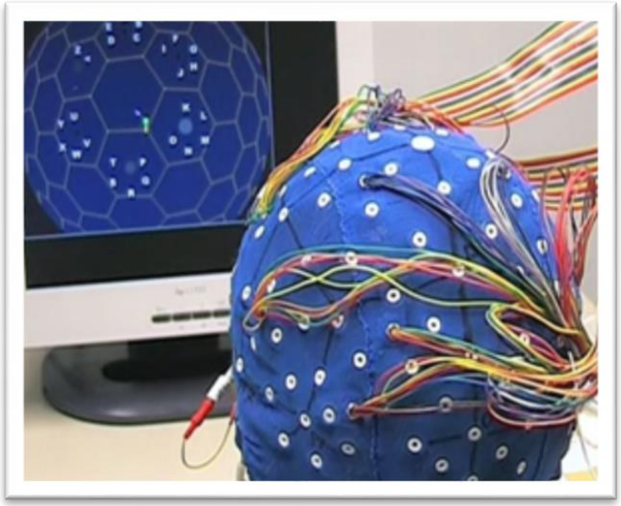
- Accuracy Limitations
  - On-screen interactive objects must be larger so that they become easier to “hit”
- Movement Freedom Limitations
  - Non-invasive solutions require that the user’s position is within a certain field of view so that the eye-tracking hardware has clear visibility of the eyes

Brain Interfaces

Introduction

- Brain-Computer Interfaces (BCIs) have been defined as a real-time communication system designed to allow a user to voluntarily send messages without using the brain's normal output pathways such as speech and gestures
  - but **only** using **bio-signals** from the **brain**
- Advances in cognitive neuroscience and brain imaging technologies provide us with the increasing ability to interface directly with activity in the brain
- In BCIs, humans intentionally manipulate their brain activity in order to directly **control a computer**
- The ability to communicate and control devices with thought alone has especially high impact for individuals with reduced capabilities for muscular response
  - In fact, applications for patients with severe motor disabilities have been the driving force of most brain-computer interface research
- The first feature that distinguishes BCIs is whether they utilize **invasive** or **non-invasive** methods of electrophysiological recordings

Non-invasive BCIs



- Non-invasive systems primarily exploit electroencephalograms (EEGs) to control computer cursors or other devices
  - Electrodes for EEGs recordings are placed **externally** to the surface of the head
- This approach has proved useful for helping paralyzed or ‘locked in’ patients develop ways of communication with the external world
- However, despite having the great advantage of not exposing the patient to the risks of brain surgery, EEG-based techniques provide communication channels of limited capacity
  - Their typical transfer rate is currently 5–25 bits s^{-1}
  - Such a transfer rate might not be sufficient to control the movements of prosthetic limps

Modern non-invasive BCIs

- In the past years, affordable non-invasive headsets have been made available to the public ($200-$500)
- The most well-known example is the EPOC headset by Emotiv (October 2014: EPOC+). Other examples include NeuroSky
  - No wires, simple bluetooth connection
  - Claims to reliably identify facial expressions, emotional states and mental commands
- These headsets are expected to accelerate dramatically the research results on BCIs

Invasive BCIs

- Invasive BCI approaches are based on:
  - recordings from ensembles of single brain cells (also known as single units)
  - or on the activity of multiple neurons (also known as multi-units)
- This methodology is based on signals obtained by **surgically** inserting electrodes inside the brain
- These approaches have their roots in experiments where monkeys learned to control the activity of their cortical neurons voluntarily
  - A few years after these experiments, Edward Schmidt raised the possibility that voluntary motor commands could be extracted from raw cortical neural activity and used to control a prosthetic device designed to restore motor functions in severely paralyzed patients

The Potential of Brain-Computer Interfaces

- The full potential of brain sensing technologies as an input mechanism lies in the extremely rich information it could provide about the state of the user
  - apart from controlling the movements of prosthetic limbs and other practical solutions such as computer operation
- Having access to this state, information is valuable to human-computer interaction (HCI) researchers and opens up at least three distinct areas of research:
  - Controlling computers with thought alone
  - Evaluating interfaces and systems
  - Building adaptive user interfaces

Controlling Computers with Thought Alone (1/2)

- Much of the current BCI work aims to **improve the lives** of patients with severe neuromuscular disorders in which many patients lose control of their bodies, including simple functions such as eye-gaze
- However, many of these patients retain full control of their higher level cognitive abilities
- These disorders cause extreme frustration or social isolation caused by having no way to communicate with the external world
- Providing these patients with brain-computer interfaces that allow them to control computers directly with their brain signals could dramatically increase their quality of life
  - The complexity of this control ranges from: simple **binary decisions**, to **moving a cursor** on the screen, to more ambitious **control of mechanical prosthetic devices**

Controlling Computers with Thought Alone (2/2)

- Nearly all current brain-computer interface research has been a logical extension of **assistive methods** in which one input modality is substituted for another
- However, now there is the need to start thinking about brain-computer interface applications for users with **no physical disabilities**
  - where brain activity can be seen as one of many of the possible input modalities that can be used sequentially or parallel with other input modalities
- Clearly, also able-bodied users can enter applications where they meet **situational impairments**
  - This includes applications in domains such as traditional communication and productivity tasks, as well as games and entertainment computing

Evaluating Interfaces and Systems

- The **cognitive** or **affective state** derived from brain imaging could be used as an **evaluation metric** for either the user or for computer systems
- Since we can measure the intensity of cognitive activity as a user performs certain tasks, we could potentially use brain imaging to assess cognitive aptitude based on how hard someone has to work on a particular set of tasks
- In addition to evaluating the human, we can understand how users and computers interact so that we can improve our computing systems
  - by learning from performance metrics such as task **completion times** and **error rates**
  - by using behavioral and physiological measures to infer **cognitive processes**, such as mouse movement and eye gaze as a measure of attention
- However, there remain many cognitive processes that are hard to measure externally
  - For example, it is still extremely difficult to ascertain cognitive workloads or particular cognitive strategies used, such as verbal versus spatial memory encoding

Building Adaptive User Interfaces (1/2)

- Tightening the iteration between measurement, evaluation, and redesign, could assist in designing interfaces that automatically adapt depending on the cognitive state of the user
- Interfaces that adapt themselves to available resources in order to provide pleasant and optimal user experiences are not a new concept
  - Researchers have been focusing on interfaces that dynamically adapt to best utilize:
    - display space
    - available input mechanisms
    - device processing capabilities
    - user task or context

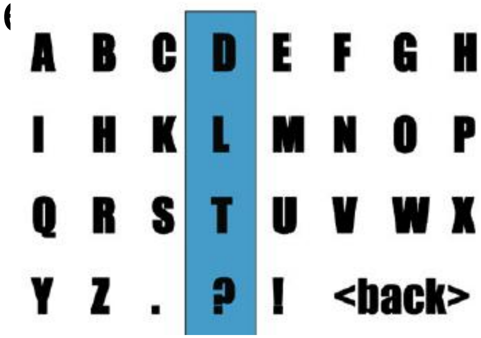
Building Adaptive User Interfaces (2/2)

- Adapting to users' **limited cognitive resources** is at least as important as adapting to specific computing affordances
- One simple way in which interfaces may adapt based on cognitive state is to **adjust information flow**
  - For example, using brain imaging, the system knows approximately how the user's attentional and cognitive resources are allocated, and could tailor information presentation to attain the largest communication bandwidth possible
  - If the user is verbally overloaded, additional information could be transformed and presented in a spatial modality, and vice versa
- Another way interfaces might adapt is to **manage interruptions** based on the user's cognitive state
  - For example, if a user is in deep thought, the system could detect this and manage pending interruptions such as e-mail alerts and phone calls accordingly
- Finally, the ability to sense higher level cognitive events like **confusion** and **frustration** or **satisfaction** and **realization** could lead to interfaces that provide feedback or guidance in training scenarios

Application Domains (1/7)

BCIs for Assistive Technology (1/2)

- Restoring the ability to communicate, controlling the environment, and providing mobility are critical concerns for people with severe physical disabilities
- **Communication**
  - One of the most critical needs for people with severe physical disabilities is restoring the ability to communicate
    - Yes/No Communication
    - Spellers
    - Web Browsers

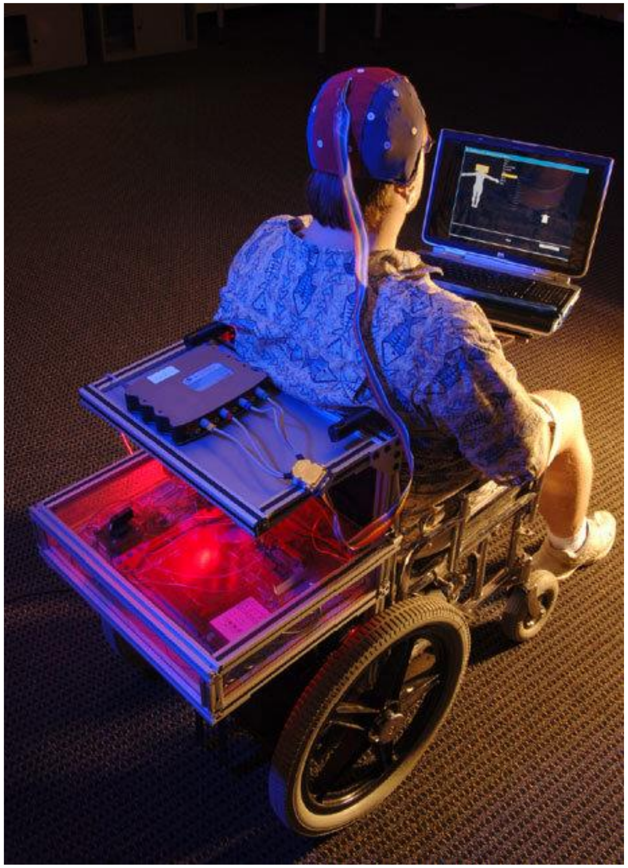


An example of matrix used for spelling

Application Domains (2/7)

BCIs for Assistive Technology (2/2)

- **Environmental Control**
  - Another important challenge for people with severe physical disabilities is controlling devices in the environment
- **Mobility**
  - One of the most profound assistive technologies BCI systems could provide to people with severe motor disability is restoring movement
    - Wheelchair control
    - Robotics ("helper" robots)



The Aware Chair [Adams et al. (2003)], integrating communication, navigation, and environmental control

Application Domains (3/7)

BCIs for Recreation (1/2)

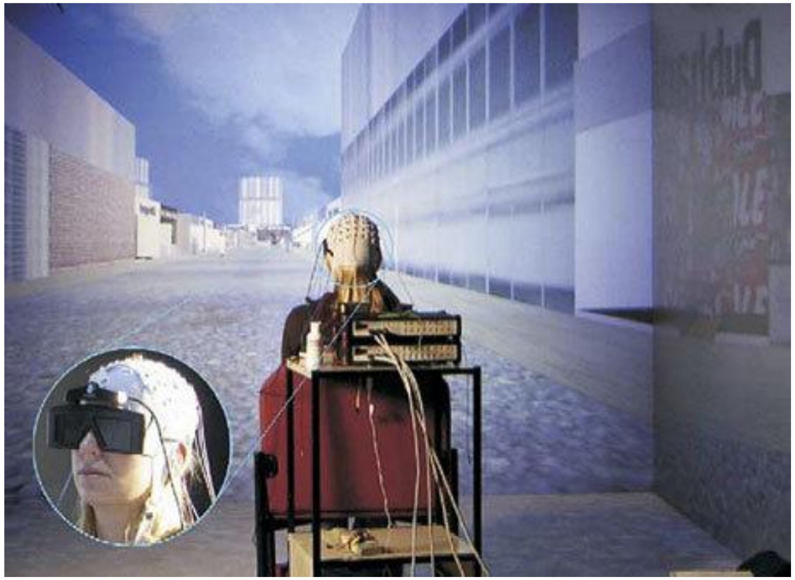
- As BCIs became more effective and new systems provided higher bandwidth, BCI control for mainstream applications such as games, virtual reality, and creative expression became possible
- **Games**
  - Early BCI controlled games focused on diagnostic brain signal detection, such as measuring a user's **attention** or **relaxation** to affect game components
  - More recently, games employing BCI in a more active control modality have gained popularity

The "Epoc" headset from Emotiv Inc. (2009) implementing "virtual telekenesis"

Application Domains (4/7)

BCIs for Recreation (2/2)

- **Virtual Reality**
  - The work of Pfurtscheller et al. (2006) introduces the ReaCTor "cave" environment



  - The subject can "walk" through the virtual world by imagining foot movement, and can "touch" things in the virtual world by imagining reaching and hand movement

Application Domains (5/7)

BCIs for Cognitive Diagnostics and Augmented Cognition (1/2)

- In addition to assistive technology and rehabilitation therapies, BCIs have been developed to aid in diagnosing, influencing, and augmenting cognitive function
- Often *neurofeedback*, representations of a user's brain signals or brain state, have been incorporated into applications to give insight into cognitive processes
- From detecting comas to monitoring attention for safety-critical tasks, diagnostic BCIs have played a role in recent research

Application Domains (6/7)

BCIs for Cognitive Diagnostics and Augmented Cognition (2/2)

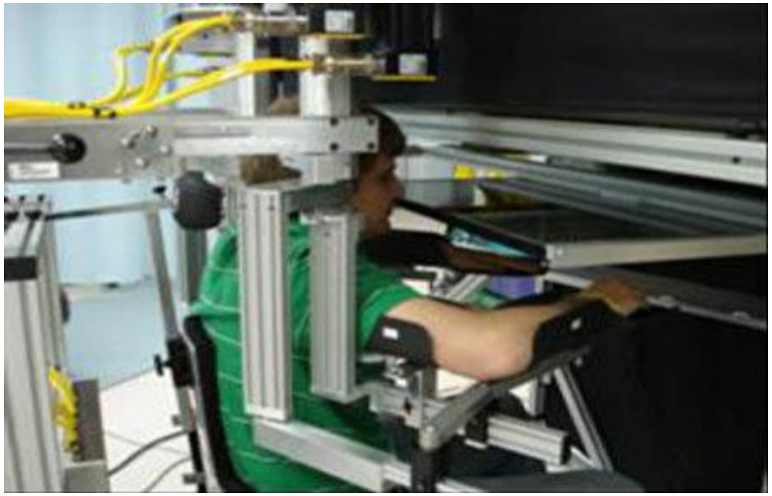
- **Coma Detection**
  - BCIs can be used as a diagnostic tool for people who appear to be in a vegetative state or coma
- **Meditation Training**
  - BCIs can be used to teach users to control changes in their EEG rhythms by meditating during mental task exercises
- **Visual Image Classification**
  - Automatically classifying visual images by measuring human brain signal responses to visual stimuli
- **Attention Monitoring**
  - Measuring alertness in safety-critical jobs requiring intense, prolonged concentration such as air traffic control and long-haul truck driving can prevent accidents and save lives

Application Domains (7/7)

Rehabilitation and Prosthetics

- One of the most significant and promising applications for BCIs currently under study involves creating therapies to regain motor control lost from diseases such as stroke
- Neural control of rehabilitation robots, for example, could provide treatments for people whose paralysis is too extensive for traditional therapies

The KINARM™ robot for upper limb movement controlled by a BCI as described in Moore Jackson (2008)



Limitations and considerations (1/3)

- Inevitable presence of artifacts traditionally deemed to be “noise” in traditional BCI explorations
- In HCI applications:
  - the environment cannot be controlled as tightly as in many medical applications
  - the user actions cannot be restricted
- Hence, new techniques must be introduced that either sidestep these issues, or better yet, that leverage the additional information

Limitations and considerations (2/3)

- All of the BCIs currently under development have limitations
- Issues of safety and the long-term stability of the recording electrodes used in invasive BCI systems remain to be resolved satisfactorily
- Demand on the user's attention
  - Rapid fatigue of users has been previously reported in certain of BCI control and inconsistent performance by individual users is characteristic of most methods
  - Ongoing changes in the user's performance (due to fatigue, distraction, disease progression, etc.) require continuing adaptation of the BCI system
- Problems of Dissemination and Support
  - All current BCIs require significant efforts to set up, calibrate, and operate
  - Problems transferring BCI technology from the laboratory to clinical setting
    - These include the ease and convenience of daily use, cosmesis, safety, reliability, usefulness of the BCI applications in the user's daily life, and the need for ongoing expert technical oversight

Limitations and considerations (3/3)

- The physical and social circumstances of potential BCI users are important
  - home situation, families, friends, and caregivers
- Unlike laboratory BCI systems, home BCI systems must be compact and able to fit into the user's environment with little or no inconvenience or disruption
  - In addition, home BCI systems must perform reliably in complex and unstable environments that often contain sources of electronic noise such as ventilators

References

- Alex Poole and Linden J. Ball. Eye tracking in Human-Computer interaction and usability research: Current status and future prospects. In Claude Ghaoui, editor, Encyclopedia of Human Computer Interaction. IGI Global, December 2005.*(Available through the course website)*
- Mikhail A. Lebedev, Miguel A.L. Nicolelis, Brain–machine interfaces: past, present and future, Trends in Neurosciences, Volume 29, Issue 9, September 2006, Pages 536-546, ISSN 0166-2236. *(Available through the course website)*
- Tan, Desney S., and Anton Nijholt. Brain-Computer Interfaces: applying our minds to human-computer interaction. Springer, 2010. *(Selected material available through the course website)*

Additional References

- Stiefelhagen, R., Yang, J., & Waibel, A. (2001). Estimating focus of attention based on gaze and sound. Proceedings of the 2001 workshop on Perceptive user interfaces - PUI '01, 1. New York, New York, USA: ACM Press.
- Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in Human-Computer Interaction and usability research: Ready to deliver the promises, In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605). Amsterdam: Elsevier.
- Plass-Oude Bos, D., Reuderink, B., De Laar, B. V., Gürkök, H., Mühl, C., Poel, M., Nijholt, A., et al. (2010). Brain-computer interfacing and games. In D. S. Tan & A. Nijholt (Eds.), *Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction (pp. 149-178). London: Springer London*
- Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A Practitioner's Guide. En: Hyona, J., Radach, R., Duebel, H (Eds.). The mind's eye: cognitive and applied aspects of eye movement research (pp. 573-605).
- Adams L, Hunt L, Moore M (2003) The aware system: Prototyping an augmentative communication interface. Paper presented at the Proceedings of the Rehabilitation Engineering Society of North America (RESNA)
- Pfurtscheller, G., Leeb, R., Keinrath, C., Friedman, D., Neuper, C., Guger, C., & Slater, M. (2006). Walking from thought. Brain research, 1071(1), 145-152.
- Jackson, M. M., & Mappus, R. (2010). Applications for brain-computer interfaces. In Brain-Computer Interfaces (pp. 89-103). Springer London.