Multi-Modal User Interaction

Lecture 3: Eye Tracking and Applications

Zheng-Hua Tan

Department of Electronic Systems
Aalborg University, Denmark
zt@es.aau.dk

Part I: Eye tracking

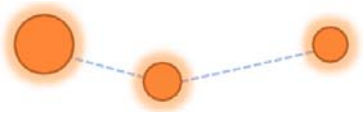
- Eye tracking
- Tobii eye tracker
- Applications
- Visual focus of attention

Eye tracking

- The process of measuring
  - the point of gaze ("where we are looking") or
  - the motion of an eye relative to the head.
- Eye trackers are used in
  - research on the visual system
  - human computer interaction
  - psychology
  - cognitive science
  - product design.

Eye movement: fixation

- Eye is a (relatively) still and “fixated” to the certain point. *E.g.* reading a single word.
- Duration varies from 120-1000 ms, typically 200-600 ms, during this stop the brain starts to process the visual information received from the eyes.
- Typical fixation frequency is < 3 Hz
- The length of a fixation is usually an indication of information processing or cognitive activities.
- All the information from the scene is (mainly) acquired during fixation.
- Interspersed with saccades ->



Eye movement: saccade

- Fast jumps from one fixation to the other
- Duration is typically only 40-120 ms
- The vision system is suppressed during the movement
- Saccades are used to move the fixation point
  - If larger than 30 degree movement is required, head moves along with eyes

Eye Tracking techniques

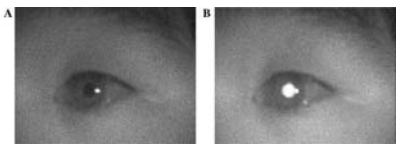
- Intrusive eye tracking
  - Contact lenses
- Remote eye tracking
  - Video-based techniques (the most widely used today)

Intrusive eye gaze trackers

- One of the most traditional methods is based on magnetic contact lenses worn by the subject.
- Each eye's movement leads to modifications of the magnetic field. These variations are recorded by captors placed on both sides of the subject's eyes, and allow knowing precisely the eye's position and movements. This technique is therefore extremely accurate (0.08º).
- However, the required equipment is very expensive and might be unsafe for the subject because of the magnetic contact lenses.

Video-based eye trackers

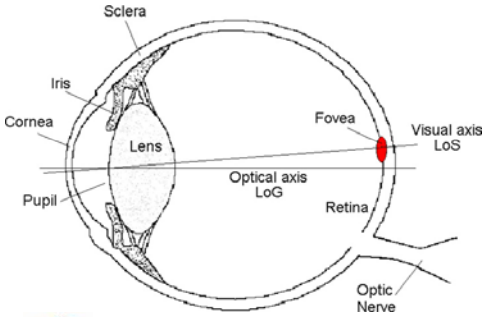
- A camera focuses on one or both eyes and records their movement as the viewer looks at some kind of stimulus.
- To enhance the contrast between the pupil and the iris, many eye trackers use an infrared (IR) light source. Because IR is not visible, the light does not distract the user. Sometimes, the IR source is placed near the optical axis of the camera which then sees a bright pupil.
- The light source also generates a corneal reflection (CR) or glint on the cornea surface near the pupil. This glint is used as a reference point in the pupil–corneal reflection technique.

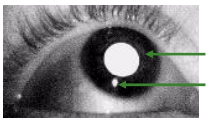


(Morimoto and Mimica, 2005)

Video-based eye trackers

- Use contrast to locate the center of the pupil and use infrared and near-infrared non-collimated light to create a corneal reflection (CR). The vector between these two features can be used to compute gaze intersection with a surface after a simple calibration for an individual.





Bright pupil

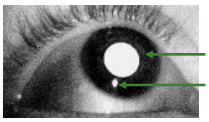
Corneal reflection

Single point video-based methods

- Tracking *one* visible feature of the eyeball, *e.g.:*
  - limbus (boundary of sclera and iris)
  - pupil
- A video camera observes one of the user's eyes
- Image processing software analyzes the video image and traces the tracked feature
- Based on calibration, the system determines where the user is currently looking
- Head movements not allowed
  - Bite bar or head rest is needed

Two point video-based method

- The same idea as in the single point method except now two features of eye are tracked – typically
  - corneal reflection
  - pupil
- Uses IR light (invisible to human eye) to
  - produce corneal reflection
  - cause bright or dark pupil, which helps the system to recognize pupil from video image



(Aaltonen, 2000)

Two point video-based methods

- The optics of the system can be mounted on
  - head
  - floor.
- If optics are floor mounted, the system is not in contact with the user
- Generally head movements are not restricted and they can be separated from eye movements, but
  - With floor mounted optics the system has to track the user's head in order to keep the eye in the field of view of camera, which limits the head movements.

Appearance-based eye gaze estimation

- Instead of using explicit geometric features such as the contours of the limbus or the pupile
- Treat an image as a point in a high-dimensional space: For a given image of an object, its viewing parameters can be estimated by finding the point on the object's appearance manifold that is nearest to the given image, and using the parameters for that point as the estimate.
- Easy to implement and more robust
- Accuracy of 0.38º

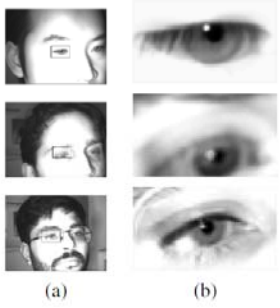


Figure 2. Samples of labeled data set collected from three subjects (from top) X, Y, and Z, who are at distances 18, 20, and 24 inches away from the display respectively. (a) Raw captured image. Cropped eye region is shown as an overlaid rectangle. (b) Cropped and scaled image of eye used as appearance sample.

(Tan et al. 2002)

Some terms

- Accuracy
  - The expected difference in degrees of visual angle between true eye position and mean computed eye position during a fixation.
  - Because of the vision system and physiology of eye the accuracy is usually 0.5-1º.
- Spatial Resolution
  - The smallest change in eye position that can be measured.
- Temporal Resolution (sampling rate)
  - Number of recorded eye positions per second.

Part II: Tobii eye tracker

- Eye tracking
- Tobii eye tracker
- Applications
- Visual focus of attention

Tobii X120 Eye Trackers

- The Tobii X120 Eye Trackers are standalone eye tracking units designed for eye tracking studies relative to any surface. They enable a variety of stimuli setups such as a TV or other displays, a projection screen or a physical object or scene.

Tobii X120 Eye Trackers

Models	X60/X120*	X120*
Data rate	60 Hz	120Hz
Accuracy	typical 0.5 degrees	typical 0.5 degrees
Drift	typical 0.1 degrees	typical 0.1 degrees
Spatial resolution	typical 0.2 degrees	typical 0.3 degrees
Head movement error	typical 0.2 degrees	typical 0.2 degrees
Head movement box	44 x 22 cm at 70 cm	30 x 22 cm at 70 cm
Tracking distance	50-80 cm	50-80 cm
Max gaze angles	35 degrees	35 degrees
Top head-motion speed	25 cm/second	25 cm/second
Latency	maximum 33 ms	maximum 33 ms
Blink tracking recovery	maximum 17 ms	maximum 8 ms
Time to tracking recovery	typical 300 ms	typical 300 ms
Weight (excluding case)	~ 3 kg / 7 lbs	
Eye tracking technique	both bright and dark pupil tracking	
Eye tracking server	Embedded	

17

A Tobii Text Writer

- Text Writer for Handicapped People – Using Eye-Gaze Tracker and Language Model [Bauduin, AAU 2008]
  - Google Web 1T 5-gram Corpus
    - The largest of the world.
    - Consists of English word N-grams and their corresponding frequencies, ranging from unigrams (one word) to 5-grams (five words).
    - Generated from ca. one trillion word tokens taken from accessible Web pages.
    - The corpus size is approximately 24 GB in compressed text files.

A Tobii Text Writer

Typing test results of the text writer

Input device		Language model			
		small	medium	large	w/o
Keyboard	Keystroke	101	69	65	274
	Time (sec)	274	298	365	85
	Errors	0	0	0	1
	sec/stroke	1.78	2.44	3.09	0.31
Mouse	Stroke (VK)	83	73	65	268
	Time (sec)	305	295	280	319
	Errors	0	0	0	1
	sec/stroke	2.24	2.34	2.37	1.19
Eye-gaze tracker	Stroke (VK)	85	77	69	297
	Time (sec)	395	385	370	592
	Errors	0	0	0	3
	sec/stroke	2.86	2.96	3.03	1.99

Part III: Applications

- Eye tracking
- Tobii eye tracker
- Applications
- Visual focus of attention

Applications

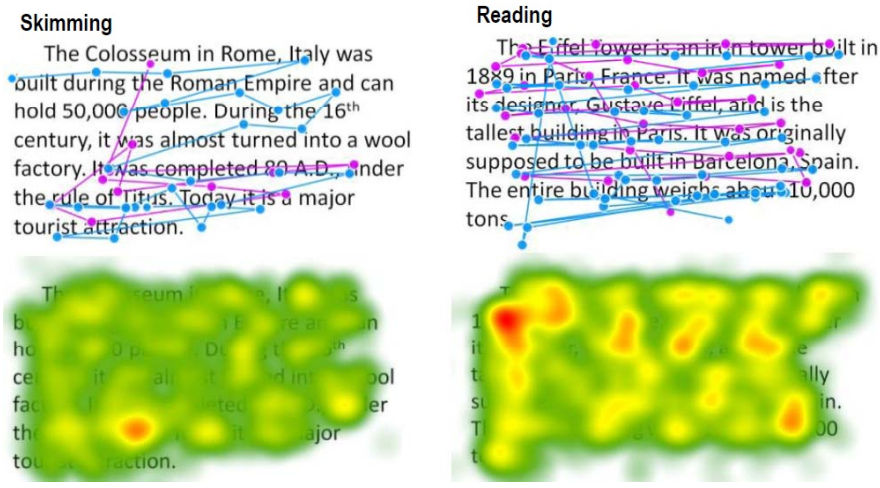
- Cognitive science
- Psychology (notably psycholinguistics, the visual world paradigm)
- Human-computer interaction (HCI)
- Marketing research
- Medical research (neurological diagnosis)
- Specific applications include the tracking eye movement in language reading, music reading, human activity recognition, the perception of advertising, and the playing of sport.

Commercial applications

- Web usability, advertising, sponsorship, package design and automotive engineering:
  - Presenting a target stimulus to a sample of consumers while an eye tracker is used to record the activity of the eye;
  - The resulting data can be statistically analyzed and graphically rendered to provide evidence of specific visual patterns.
  - By examining fixations, saccades, pupil dilation, blinks and a variety of other behaviors researchers can determine a great deal about the effectiveness of a given medium or product.
- Communication systems for disabled persons: allowing the user to speak, send e-mail, browse the Internet and perform other such activities, using only their eyes.

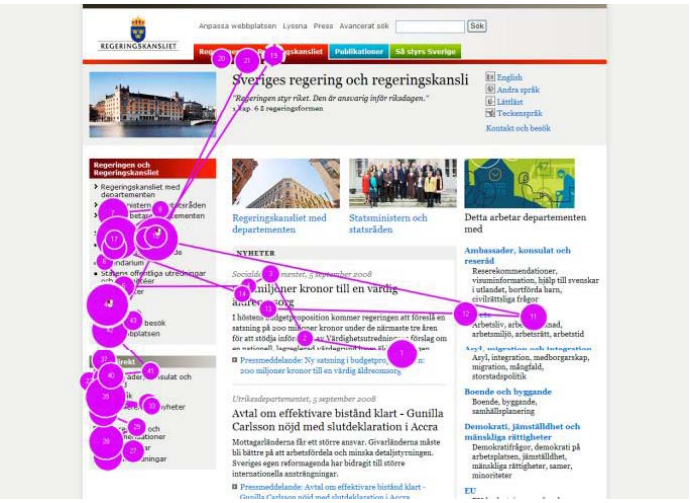
(wikipedia.org)

Reading research



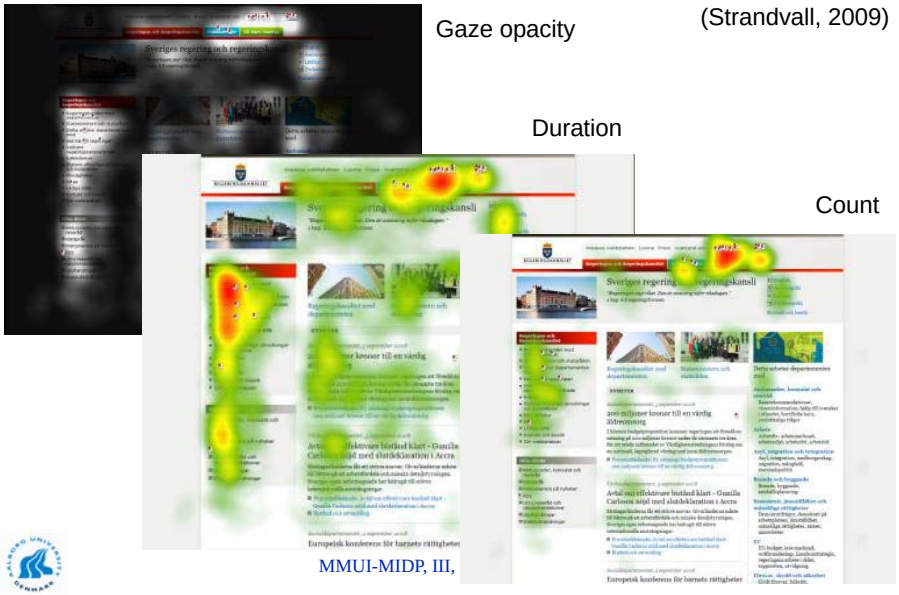
(Strandvall, 2009)

Visualizing behavior: Gaze plot



(Strandvall, 2009)

Visualizing behavior: Heat maps



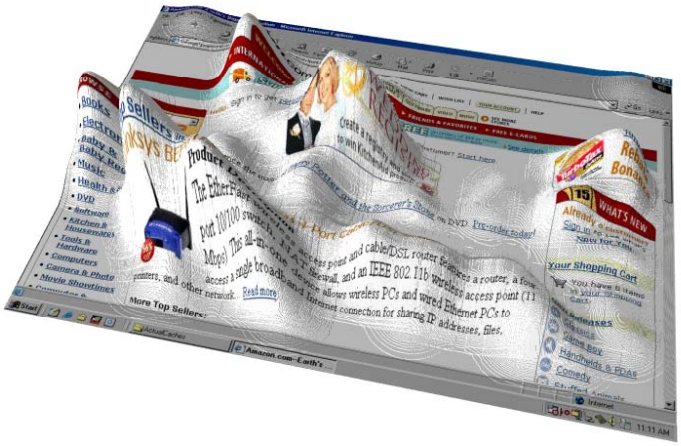
Gaze opacity

Duration

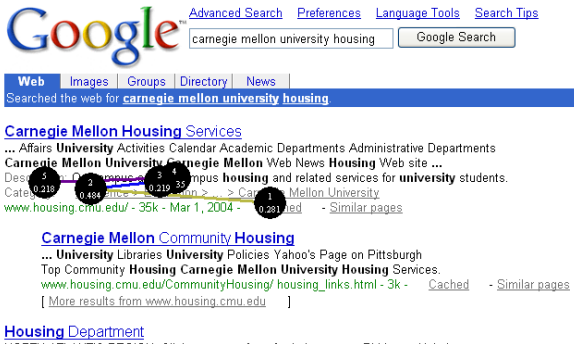
Count

(Strandvall, 2009)

Eye gazing for Web usability



(Nielsen and Coyne , 2006)



(Nielsen and Coyne , 2006)

Findings – home page

- Eye is drawn to standard (expected) navigation areas – top of page horizontal navigation bar
- Users ignore big images with top stories and images that look like ads
- Users expect standard info eg contact details (footer), search (top right hand corner) and privacy (footer) to be located in particular areas
- Online shoppers go straight for the navigation and ignore sales pitches especially those embedded in images that look like ads

Findings – home page

- Users are not interested in how fancy the home page looks. They navigate quickly to complete tasks, home page is just a ‘gateway’
- An indication of what happens in reality – people go to websites to find/do something so they are not open to promotional/marketing content
- Gimmicky/’marketese’ link names confuse users eg brand names eg ‘Sony Style Retail Store’

Text entry

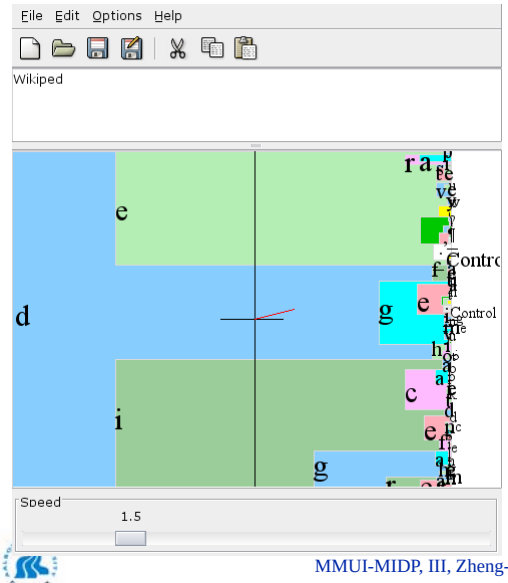
- Visual keyboards (many)
- Hierarchical key systems (Gazetalk)
- Dasher
- Pie menus
- Gestures

- Need to accommodate eye movements of user and inaccuracies of eye tracker

Dasher

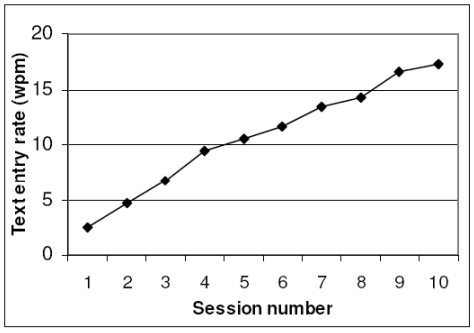
- A zooming interface displaying letters in alphabetical order. Users point, with the mouse, where they want to go and the application zooms in to this area. The interface is flowing and new letters appear in order to write words.
- Dasher also uses a language model predicting the next letters that help the user. Letters with high probabilities to follow fits inside a bigger box than low probability ones. Thus, the user first perceives the high probability letters although it is still possible to choose others by zooming in to the corresponding box.
- http://www.inference.phy.cam.ac.uk/dasher/

Dasher



Dasher

- 12 participants transcribed Finnish text with Dasher in ten 15-minute sessions using a Tobii 1750 eye tracker as a pointing device



(Tuisku et al. 2008)

Figure 6: Mean text entry rate per session.

Eye tracking open source

- Open source gaze tracking, freeware and low cost eye tracking
- openEyes, open-source open-hardware toolkit for low-cost real-time eye tracking
- Opengazer: open-source gaze tracker for ordinary webcams
- TrackEye: Real-Time Tracking Of Human Eyes Using a Webcam. Implemented in C++ using the OpenCV library

Part IV: Visual focus of attention

- Eye tracking
- Tobii eye tracker
- Applications
- Visual focus of attention

Visual focus of attention in open spaces

- Eye tracking technologies
  - Very accurage, but not appropriate for analyzing the ”visual focus of attention” of people in open spaces
  - Intrusive, specific equipment (infrared light sources to ease signal processing)
  - Head motion is limited, even chin rests or bite bars requried
  - Even eye-appearance vision-based tracking systems restrict the mobility to have high-resolution close-up eye images

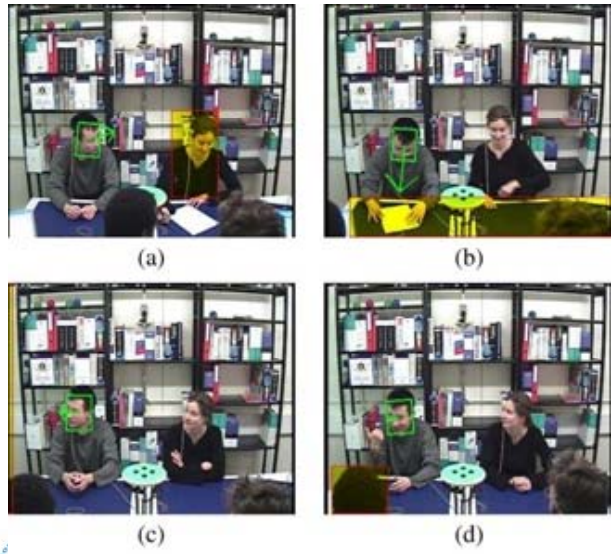
Visual focus of attention in open spaces

- Recognize the visual focus of attention based on head pose
  - Estimate a person’s gaze and visual focus of attention in open spaces
  - Motion and head orientation are unconstrained, high-resolution images of eyes are not available.

Tracking visual focus of attention

- Two generative models (head pose observations associated with each visual target are represented by Gaussion distributions)
  - Gaussian mixture model (GMM) separately handles each frame
  - Hidden Markov model (HMM) segments pose observation sequences into temoral segments of ”visual focus of attention”

Visual focus of attention

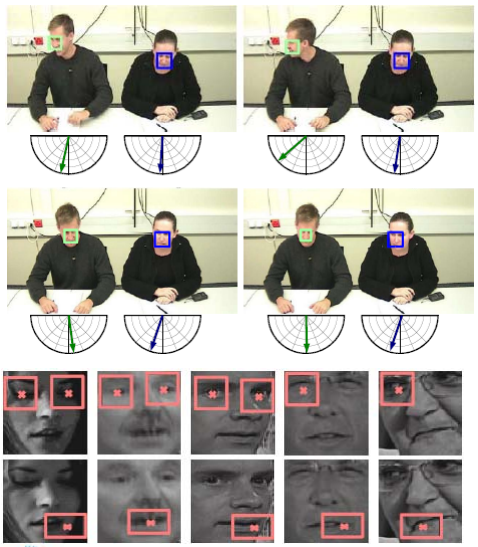


(Ba and Odobez, 2009)

Focus of Attention Tracking

- http://www.amiproject.org/showcase/still-and-moving-image-processing/focus-of-attention-tracking

Multimodal Tracking



(Wallhoff et al. 2006)

Summary

- Eye tracking
- Tobii eye tracker
- Applications
- Visual focus of attention