

# T-61.183 Multimodal Systems

with Eye Movements

Kai Puolamäki

31 January 2005

# Eye movements

- The eye can move in three ways:
  - *Smooth pursuit movements*
  - *Convergent movements*
  - *Saccadic movements*
- Saccades and fixations: our perception is static snapshots of the world (in a rough approximation)
- Saccades take .02-.1 sec. We are practically blind during the saccades (*saccadic suppression*).
- Saccades are *ballistic* (pre-programmed)

# Fixations

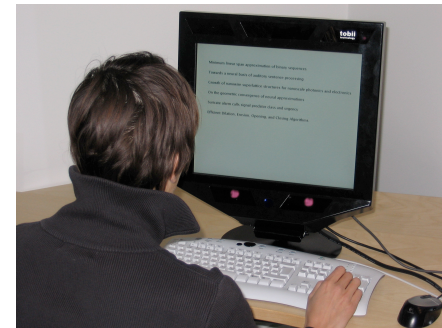
- Fixations last .2-.6 sec
- Fixations have internal structure (microsaccades)
- Most information is obtained from the central foveal area (1-2 deg)
- Information is also obtained from periphery (moving objects are effectively seen at 40 deg off the visual axis)
- Approach e.g. in reading models: eye is capable of taking in a constant flow of information. Rare and/or unexpected words are looked at longer.

# Measuring eye movements

- Experiments are not easy (same problem for many other modalities)
- The best equipment has sampling rate of 250+ Hz
- With 20 k€ you get a decent 50 Hz system
- You could get a coarse eye tracker (user is/is not looking at the monitor) cheaply with a webcam and some programming :)

# Shopping list

- SMI ([www.smi.de](http://www.smi.de))
  - Camera is mounted on the subject
  - Good if user needs to walk around freely
  - In e.g. reading studies the head must be held still, which might be uncomfortable
  - Problems with accuracy especially in reading application
- Tobii ([www.tobii.com](http://www.tobii.com))
  - Camera integrated into LCD display
  - Nice to use
  - Calibration is good
- Arrington Research ([www.arringtonresearch.com](http://www.arringtonresearch.com))

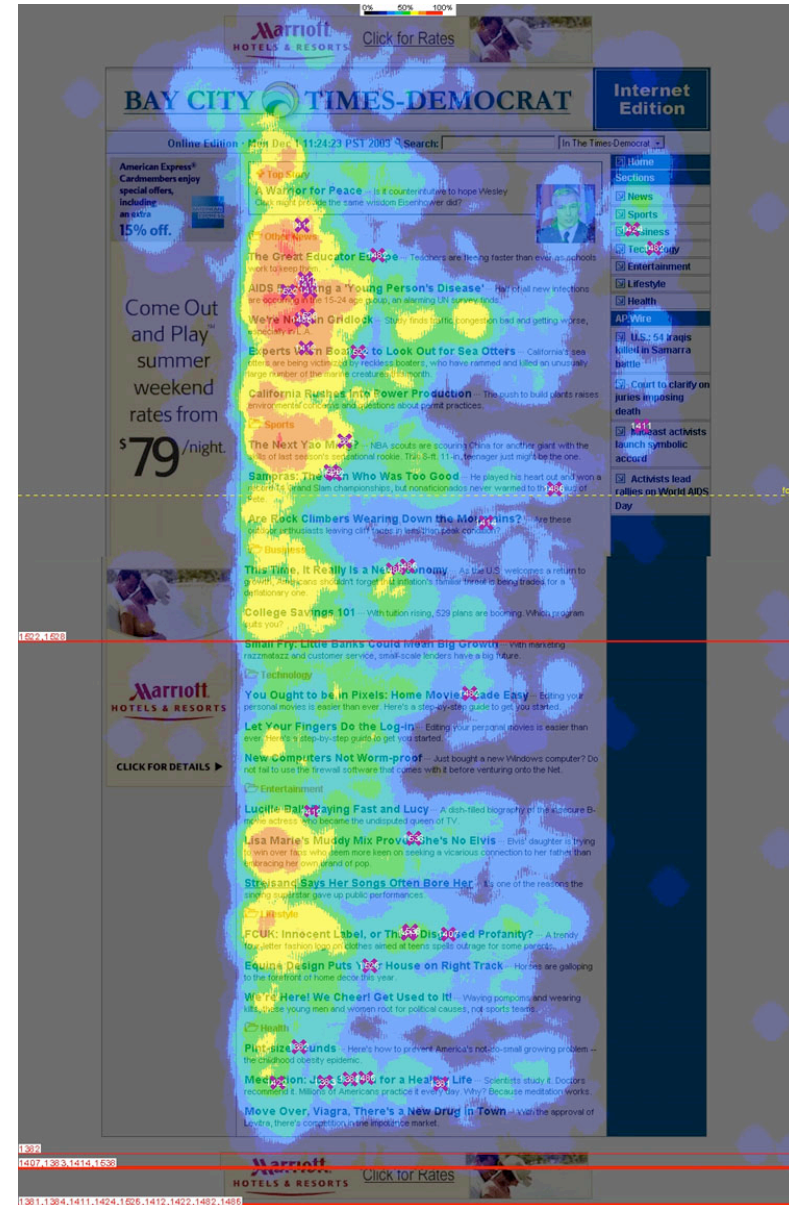


# Interpretation and methodology

- You can measure eye movements. Then what?
- There have been lots of psychological studies
- Assumption: if people look at something, then it is interesting/they take in information
  - Examples: Eyetrack III, iDict
- To really use the eye movements you must set up a controlled experiments where you know the ground truth
  - Example: relevance determination

# Eyetrack III

- Scanning pattern depends on the task
- In a web news site:
  - upper left corner is scanned first
  - pattern of vertical viewing: first few words are scanned most
  - people tend to look text first, then photos



# Eye typing

- Traditionally eye typing has some sort of a keyboard (not necessarily QWERTY): typing is done by looking at the keys
- DASHER:
  - letters flow from right to left
  - predicts the next word: most likely letters are larger
  - easy and fast to use after initial learning period





# DASHER

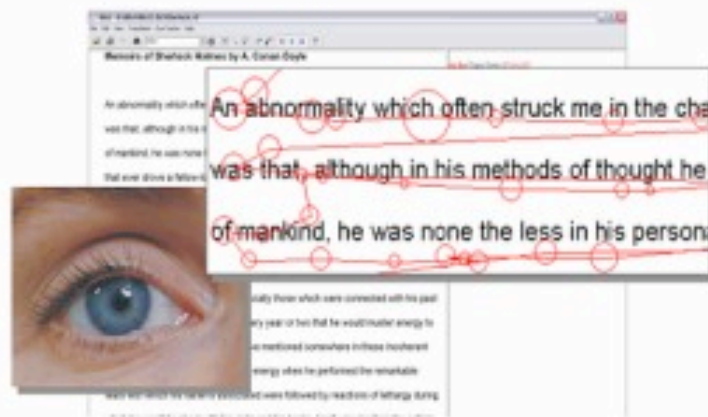




## iDict, a Gaze-Assisted Reading Aid

Text written in foreign language:

dictionary look-ups  
interrupt the normal  
flow of reading



### iDict

- monitors eye movements
- automatically detects irregularities in a reading process
- consults the embedded dictionaries and provides assistance

# Information retrieval

- Task: find answers to relevant questions
- Controlled setup (ground truth is known)
- A probabilistic time series model is used to find the relevant lines from eye movements only
- Scenario: computer learns the relevant articles from the user without implicit feedback, feedback information is combined to other sources, helping information retrieval (*proactive information retrieval*)
- [Movie]

# SUITOR

- An attentive information system by IBM
- Tracks user on multiple channels, including eye movements
- Finds potentially interesting information for the user

# References

# Topics

- Eye movements in...
  - reading
  - information retrieval (as a feedback channel)
  - (looking at images)
- ?