Methods for fusing text content and eye movements for information retrieval

A PASCAL pump-priming project

Kai Puolamäki
Helsinki University of Technology (TKK), Finland
Kai.Puolamaki@tkk.fi

29 January 2008
Partners

Helsinki University of Technology (TKK) [coordinator]
   Samuel Kaski, Kai Puolamäki (Antti Ajanki)

University of Helsinki
   Wray Buntine, Petri Myllymäki (Janne Kataja)

University of Southampton / University College London
   John Shawe-Taylor (David Hardoon)
• Introduction
• Is the user interested in the text she is reading?
• What is the user interested of?
• Integration into IR application
• Results
People can infer a lot from gaze

- Gaze is one of the modalities (feedback channels) that give information of the mental state of the person
- So far, machines have been unable to interpret it

Kai Puolamäki / Methods for fusing text content and eye movements for information retrieval
New field

• Measuring eye movements non-intrusively in practical applications is becoming also commercially feasible
• Quantitative relation between eye movements and the mental state of an user in a typical IR task has been unknown (until us)
• Lots of appealing things:
  – Interaction with the theory and application:
    ▶ Connection to many machine learning problems
    ▶ We can design the experiments and make the measurements; we are the application area experts
  – Psychological interpretation of the results
  – An important future application
Brief history

• 1950’s onwards: rigorous psychological research on eye movements. (Quite orthogonal to our work, however.)
• 2000-2003: SUITOR, agent that monitors eye movements when user is viewing web pages. No quantitative conclusions. (Maglio, Barret, Campbell)
• 2003: iDict, proactive translator (Hyrsykari, Majaranta, Räihä). A translation is provided if user stares at the word long enough.
• Before and after 2000: Eye typewriters - an important application, for example, for people with disabilities. Example: Dasher; Ward, MacKay, 2002.
Relevance from eye movements

• 2003: Initial studies
• 2004-2005: Does the user find the text relevant?
• 2006-2007: What is the user interested in?
• 2008-: Integration to search engine/information navigator, images, other modalities
• Projects, events:
  - PRIMA under the Academy’s Proactive Computing Research Program (PROACT) in 2002-2005
  - PASCAL Pump Priming project in 2005-2007
  - FP7 STREP project PinView in 2008-2010 (direct consequence of the Pump Priming project)
Our tracker

Looks like a normal LCD monitor, good for reading text etc.

Infrared lights (invisible to bare eye)

Allows normal head movement, tracking accuracy $\frac{1}{2}^\circ$

Infrared camera

Currently expensive, consumer versions maybe in 5-10 years(?)

Tobii 1750 eye tracker
Our tracker

Looks like a normal LCD monitor, good for reading text etc.

Tobii 1750 eye tracker

Infrared lights (invisible to bare eye)

Allows normal head movement, tracking accuracy $\frac{1}{2}^\circ$

Infrared camera

Currently expensive, consumer versions maybe in 5-10 years(?)

Kai Puolamäki / Methods for fusing text content and eye movements for information retrieval
• Introduction
• **Is the user interested in the text she is reading?**
• What is the user interested of?
• Integration into IR application
• Results
Is the user interested in text she is reading?

Task

- Task: predict relevance for sentences, given eye movement trajectory, that is, estimate $p(\text{relevance} | \text{eye movements})$.

- Textual content was not taken into account at all!
PASCAL Eye Challenge 2005

Is the user interested in text she is reading?

Helle vei kuljettaalta tajun
Lomaltapaluu lisäsi stressiä
Lisää hellettä luvassa
Helteet alkoivat vihdoin
Papukaija nähtiin Lahdessa
Kulttuuriministeri ei saapunut kesätapahtumaan
Kuolonuhrien määrää selvitettiin
Helle pilasi viljasadot
Ennätyskuuma helle mitattiin Suomessa
Finlandia-vodkan muotoilu vaihtui


Kai Puolamäki / Methods for fusing text content and eye movements for information retrieval
PASCAL Eye Workshop 2005

• In NIPS workshop programme, Whistler, BC, Canada
• Best performing solutions of the PASCAL Eye Challenge
• Eye feedback and eye movements

Is the user interested in text she is reading?

• Introduction
• Is the user interested in the text she is reading?
• **What is the user interested of?**
• Integration into IR application
• Results
Task

• User tries to find documents of a specific topic, for example, of dinosaurs
• The computer does not know the topic
• Task: guess topic, or user’s query, from the way the user is reading the documents shown to her
• Hardoon, Ajanki, Puolamäki, Shawe-Taylor, Kaski. AISTATS’07.
Xenotarsosaurus ("strange-ankle lizard") is a little-understood theropod of the late Cretaceous (~83 – 73 mya). It probably weighed 0.7 – 1.0 tons. The only fossil evidence consists of a small number of vertebrae and leg bones, retrieved from the Bajo Barreal Formation, Chubut, Argentina. From these samples, Martinez, Giminez, Rodriguez and Bochatey named the type species, X. bonapartei, in 1986. It was probably an allosaurid.
The history of film is one of the most rapidly moving of any artistic or communications medium ever, as befits perhaps the first great mass medium of the modern era. Film has gone through a remarkable array of changes and developed a remarkable variety and sophistication in barely more than one hundred years of existence.
A Post Office box is a uniquely-addressable lockable box located on the premises of a Post Office station. Generally, Post Office boxes are rented from the post office either by individuals or by businesses on a basis ranging from monthly to annual, and the cost of rent varies depending on the box size. CBD PO boxes are usually more expensive than a rural PO Box. In the United States, the rental rate used to be uniform across the country. Now, however, a postal facility can be in any of seven fee groups by location; in addition, certain postal patrons qualify for free box rental.
Among the recognizable Olympic symbols:

The Olympic flag: A white flag with the Olympic Rings on it in five colours.

The Olympic Flame: A flame burning day and night for the duration of the Olympic Games.

The Olympic Fanfare and Theme: A musical composition by John Williams.

The Kotinos: A crown made from an olive branch, which can seen atop many statues of ancient Olympic victors.
Assumptions

- There is a link between relevance of a word and eye movements related to it
- There is data for which the relevance and eye movements are known (but for different topics)
- The link is independent of actual topic
- The link can be learned and used on new topics
Assumptions

End result (simplified):
weight of the word in a query = f(saccade length before first fixation to the word, did regression initiate from the following word, duration of the first fixation to the word divided by the total fixation durations on the document, relative position of the word on the document, ... | λ)

Xenotarsosaurus ("strange–ankle lizard") is a little–understood theropod of the late Cretaceous (~83 – 73 mya). It probably weighed 0.7 – 1.0 tons.

The only fossil evidence consists of a small number of vertebrae and leg bones, retrieved from the Bajo Barreal Formation, Chubut, Argentina. From these samples, Martinez, Giminez, Rodriguez and Bochatey named the type species, X. bonapartei, in 1986. It was probably an allosaurid.

Xenotarsosaurus ("strange–ankle lizard") is a little–understood theropod of the late Cretaceous (~83 – 73 mya). It probably weighed 0.7 – 1.0 tons.

The only fossil evidence consists of a small number of vertebrae and leg bones, retrieved from the Bajo Barreal Formation, Chubut, Argentina. From these samples, Martinez, Giminez, Rodriguez and Bochatey named the type species, X. bonapartei, in 1986. It was probably an allosaurid.
• Introduction
• Is the user interested in the text she is reading?
• What is the user interested of?
• **Integration into IR application**
• Results
Gazillion Eye Browser

- Petri Myllymäki, Janne Kataja of University of Helsinki
- An eye browser
- Published under an open source (MIT) license at http://www.cs.helsinki.fi/u/jkataja/gazillion/
- Integrated into University of Helsinki Wikipedia search engine

Kinetic sculptures are sculptures that are designed to move. Additionally, sound sculpture is considered kinetic sculpture. The name is derived from the Greek "kinesis" meaning motion, and related to the scientific term kinetic energy, which is the energy contained in a mass due to its motion.

Alexander Calder (inventor of the mobile) and George Rickey pioneered kinetic sculpture. Other leading exponents include Yaacov Agam, Marcel Duchamp, Naum Gabo, Ronald Mallory, Jean Tinguely, Zero group and Arthur Ganson.
Gazillion Eye Browser

- The first general purpose eye browser
- We have also a Python web server
- Easy to make interactive experiments and applications: our current primary research platform
- Further releases forthcoming

Kinetic sculptures are sculptures that are designed to move. Additionally, sound sculpture is considered kinetic sculpture.

The name is derived from the Greek "kinesis" meaning motion, and related to the scientific term kinetic energy, which is the energy contained in a mass due to its motion.

Alexander Calder (inventor of the mobile) and George Rickey pioneered kinetic sculpture. Other leading exponents include Yaacov Agam, Marcel Duchamp, Naum Gabo, Ronald Mallory, Jean Tinguely, Zero group and Arthur Ganson.

Kai Puolamäki / Methods for fusing text content and eye movements for information retrieval
• Introduction
• Is the user interested in the text she is reading?
• What is the user interested of?
• Integration into IR application
• Results
Results of the research

• We performed eye movement experiments and organized a PASCAL challenge and workshop.
• We showed that implicit feedback can be extracted from eye movements:
  - When used together, eye movements and textual content improved information retrieval results.
  - We can construct implicit queries from a word-specific movements.
    ▸ Interesting machine learning problems
• We have an eye browser application, integrated with a web server.
  ▸ Test bed, makes large(r) scale data collection possible, usable application
Successful collaboration

- Sharing of data
  - Purchase of the eye tracker
  - The Wikipedia corpus from University of Helsinki
  - Eye movement measurements
- Numerous visits, discussions
Continuing research, new collaborations

- **FP7 STREP PinView**, with 6 partners, 2008-2010
  - TKK, U Southampton, UCL, U Leoben, Xerox, celum
  - **Goal:**
    develop the necessary principles to replace or extend textual IR queries using multi-modal feedback;
    make a prototype IR application
- Other projects to be announced
- New data usable by other researchers
- The Pump Priming project had a crucial role in making a new research area possible
Methods for fusing text content and eye movements for information retrieval

A PASCAL pump-priming project