Gaze-Based Filtering of Relevant Document Segments

Georg Buscher, Andreas Dengel
DFKI, Kaiserslautern, Germany

WSSP, 20.04.2009

Summary Snippets in Re-Finding Scenarios

Browser history

Classical Summary

***Honey bee* - Wikipedia, the free encyclopedia**
Honey bees (or honeybees) are a subset of bees, primarily distinguished by the production and storage of honey and the construction of perennial, **...**

?

Personalized Summary

***Honey bee* - Wikipedia, the free encyclopedia**
Visual Perception
Eye(s). Honey bees have two compound eyes that make a large part of the head surface. Each compound eye is composed of individual cells (ommatidium, **...**

!

Research Questions

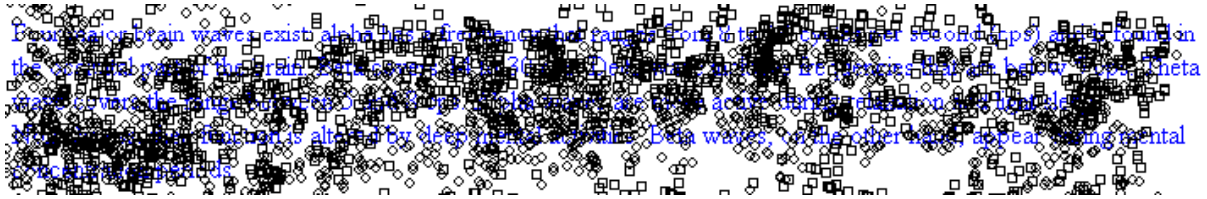
- When is the user reading?
- Which eye movements indicate relevance?

Outline

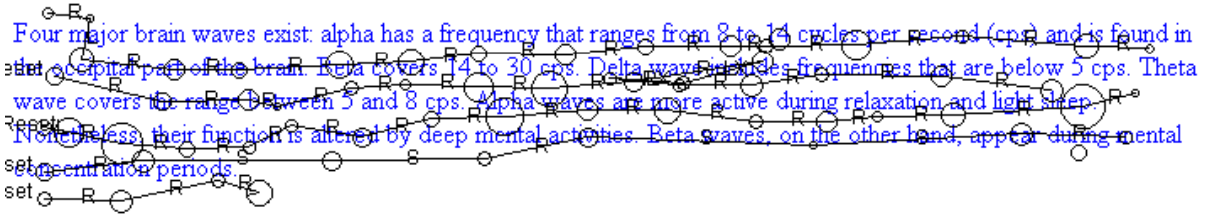
- Reading Detection Algorithm
  - Eye Movement Measures
  - Study Design
  - Results

Reading Detection – Overview

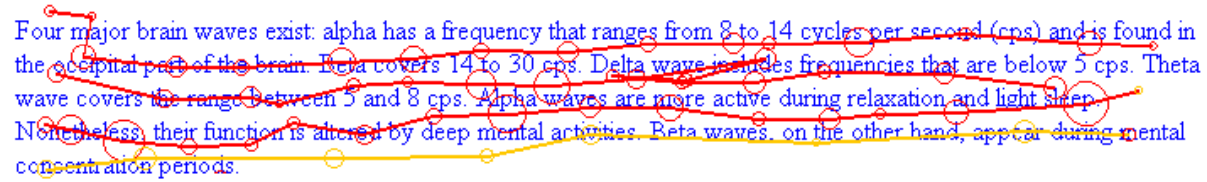
1. Starting point: Noisy gaze data from the eye tracker



2. Fixation detection and saccade classification

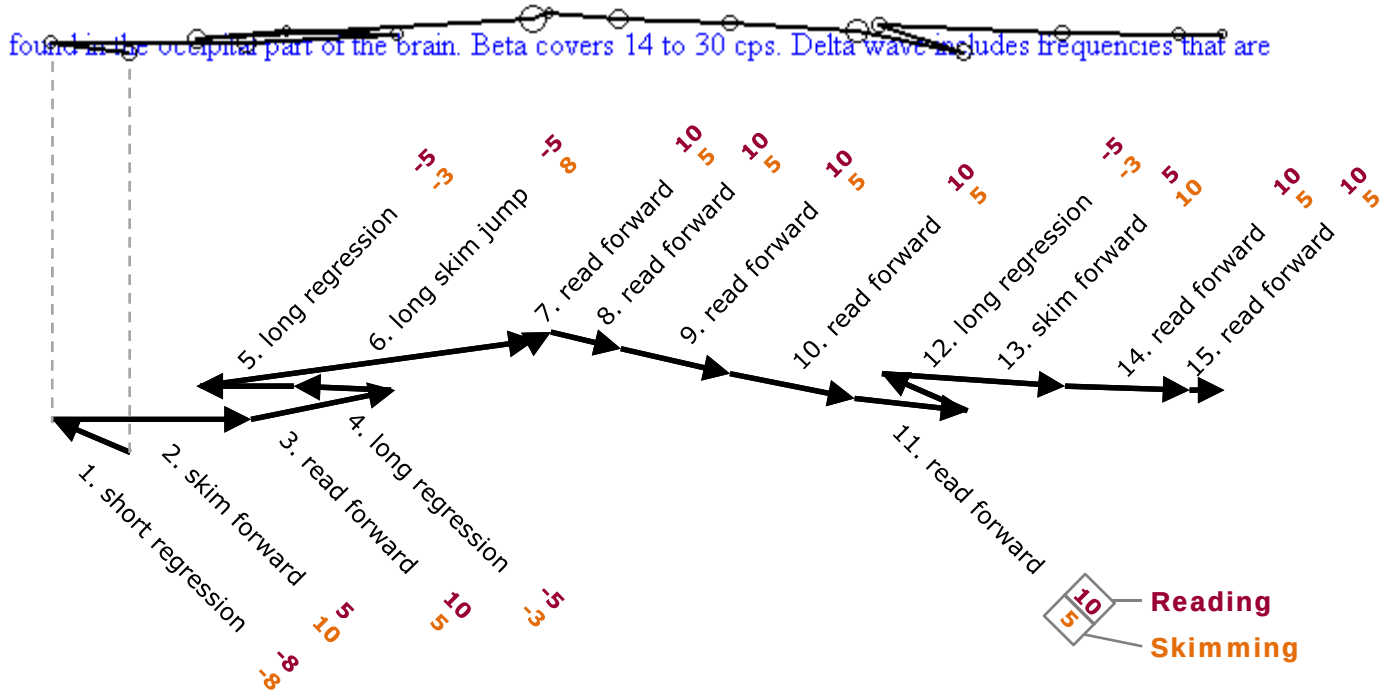


3. Reading (red) and skimming (yellow) detection line by line



See G. Buscher, A. Dengel, L. van Elst: “Eye Movements as Implicit Relevance Feedback”, in CHI '08

Reading Detection – Example



Plausibility of reading

S_r = -8 +5 +10 -5 -5 -5 +10 +10 +10 +10 +10 -5 +5 +10 +10 = **62**

Plausibility of skimming

S_s = -8 +10 +5 -3 -3 +8 +5 +5 +5 +5 +5 -3 +10 +5 +5 = **51**

62 > 51 → Reading behavior detected

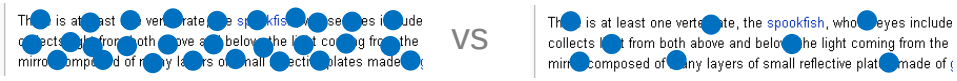
Outline

- Reading Detection Algorithm
- Eye Movement Measures
- Study Design
- Results

Eye Movement Measures

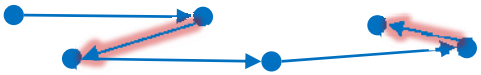
- Fixation duration

- Fixation count

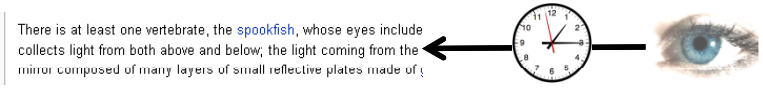


- Average saccade length

- Regression rate

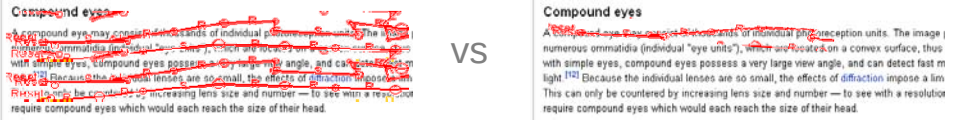


- Viewing time



- Reading vs. skimming behavior

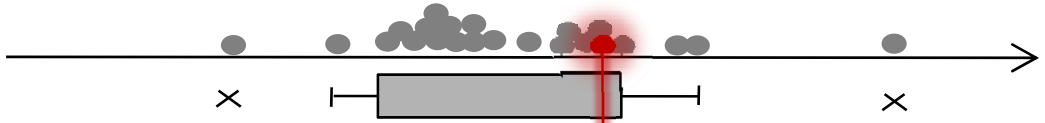
- Length of coherently read text



“Whisker”-Personalization of the Measures

- Absolute value range of a measure for a specific user

- Boxplot

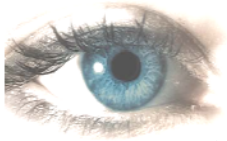


- “Whisker”-personalized interval for user

Outline

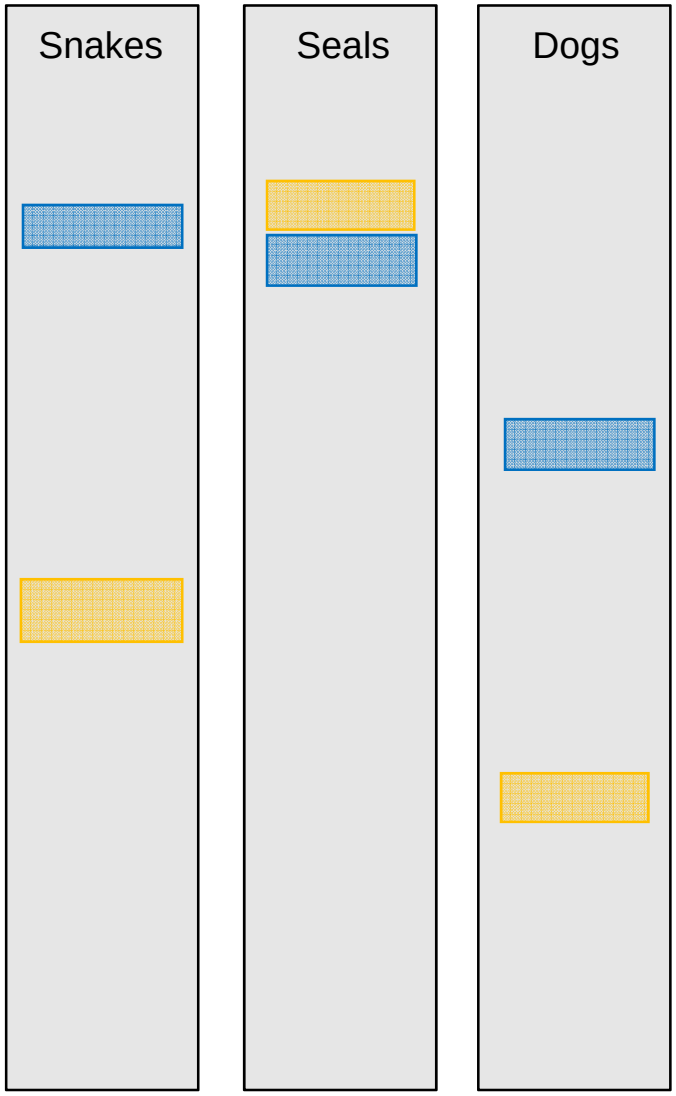
- Reading Detection Algorithm
- Eye Movement Measures
- Study Design
- Results

Study Design



Task: visual perception

Task: thermoregulation



- 4 lengthy Wikipedia articles
- 6% of each article highly relevant (for each task)
  → “ground truth”
- 25 participants

Outline

- Reading Detection Algorithm
- Eye Movement Measures
- Study Design
- Results

Results – Reading Detection

	unknown document	known document
precision	0.55	0.88
recall	0.88	0.84
f-measure	0.68	0.86

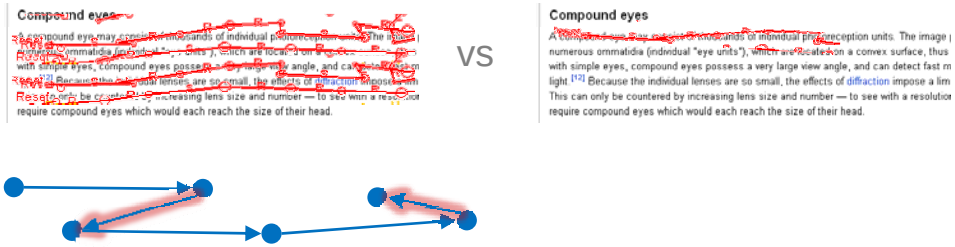
- High precision and recall for known documents
- Relatively low precision for unknown documents

Improvement by Gaze-Based Measures on Unknown Docs

	detected reading vs. skimming behavior >=						regression rate >=						average saccade length >=								length of coherently read text >=					
	0.1	0.2	0.3	0.4	0.5	0.6	0.1	0.2	0.3	0.4	0.5	0.6	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	50	100	150	200	250	300
precision	.11	.12	.16	.16	.16	.14	.13	.13	.21	.21	.14	.10	.02	.03	.05	.09	.13	.19	.22	.1	.02	.09	.11	.14	.14	.17
recall	-.06	-.12	-.18	-.30	-.41	-.53	-.03	-.15	-.32	-.55	-.75	-.87	-.02	-.03	-.04	-.09	-.20	-.44	-.73	-.90	-.01	-.01	-.02	-.03	-.07	-.11
f-measure	**.04**	.02	.00	-.08	-.15	-.27	**.06**	.0	-.07	-.27	-.52	-.71	.00	.01	**.01**	.01	-.02	-.17	-.48	-.78	.01	.05	.05	**.07**	.0	.04

- Best improvement of precision for average saccade length

- Best overall improvement (f-measure) for length of coherently read text
- And regression rate



- Fixation duration does not work well

Conclusion

- Difference between first-time- and second-time-reading
- Reading detection is an effective first relevance filter
- Best improvement for relevance estimation by
  - length of read text and
  - regression rate
- Fixation duration does not work well

Thank you for your attention!