Eye movements and tracking

Miroslav Živković
The retina is a light sensitive structure inside of the eye responsible for transforming light into signals, which are later converted into an image by the visual cortex in the brain.

The fovea is a section of the retina that contains a high density of both kinds of light receptor cells found in the eye, i.e. cones and rods.
Events and movements

• Fixation
  – The most reported event
  – Not actually a movement, but the state when eye remains still over a period of time
  – When we measure a fixation, we actually measure attention to a particular position (Eye-mind hypothesis)
  – Accompanied by tremors (90 Hz), drifts and micro-saccades
  – Duration: anything from tens of msecs to seconds
Events and movements

• Saccades
  – Motion from one fixation to another
  – Very fast movements, the fastest a human body could make
  – No processing of information possible
  – Duration: typically in the range 30-80 msecs
The pupil

• Not only the opening through which the light enters
• Two types of muscles in the pupil are controlled by the autonomic nervous system
• Excellent source of information about stress, cognitive load
• However, measurement requires tight control
• Large pupils => more variability
• Small pupils => more stability
An example

Instructions
1: examine at will
2: estimate wealth
3: estimate ages
4: guess previous activity
5: remember clothing
6: remember position
7: time since last visit
The Eye Tracker
The Eye Tracker

• monitors eye-movements

• provides information about where people look and for how long

• Does it with a certain sampling rate
Websites

- Most people view websites in a “F” shaped flow.
- First they scan the page at the top, from left to right.
- Then the eyes go back to the left and down the page.
- They again scan to the right and back along the same pattern.
Accuracy and precision

• Accuracy is defined as the average difference between the real stimuli position and the measured gaze position.
• Precision is defined as the ability of the eye tracker to reliably reproduce the same gaze point measurement.
Accuracy and precision errors
Eye-Tracking While Reading
• A gaze replay, of a participant in a reading study
• Recorded at 300Hz using the Tobii TX300 eye tracker
Some Facts About Reading

• Readers **fixate** a particular word (200 - 300ms) and jump to the next (no “smooth” reading)
• Readers often make a jump (or **saccade**) that covers 7-9 letter spaces
• During a saccade visual input is reduced (we are practically blind)
• Readers skip short words and words that are highly predictable
• Readers regress (look back)
• Readers often undershoot on return sweeps (going from the end of a line to the next line)
Since Jay always jogs a mile seems like a short distance to him.

- Green circle = fixation after progressive saccade (first-pass)
- Red circle = fixation after regressive saccade
- Blue circle = fixation after progressive saccade (second-pass)
• If readers experience some sort of trouble they may fixate the difficult region longer and the may even regress to earlier parts of the sentence/text.
Different Measures

- **First fixation duration**: duration of first fixation in a region
- **First-pass duration**: time spent in a region before moving on or looking back
- **Regression path duration**: time from first entering a region until moving the eyes beyond that region, includes regression time
- **Second-pass duration**: duration of re-fixations
- **Total duration**: the sum of all fixations in a region
- **Probability of a regression**: the percentage of regressive eye-movements out of a region
Different Measures

Bart annoyed *Homer* because…

1 2 5 3 4 6 7

**Reading Times for word 3 (Homer)**

- First fixation duration = 3
- First-pass duration = 3 + 4
- Regression Path duration = 3 + 4 + 5
- Second-pass duration = 6
- Total duration = 3 + 4 + 6
Eye Mind Assumption

- The Mind-Eye hypothesis
  - Relationship between eye fixations and the meaning of concurrently spoken sentence
  - Using this relationship as a research tool in cognitive psychology and psycholinguistics
  - Applications:
    - Speech perception and memory
    - Language processing


(a) Raw samples at 50 Hz.

(b) Raw samples at 50 Hz.

(c) Raw samples at 1250 Hz.
Eye-tracking signal

Raw eye tracking signal, participant 9, text 2725

Cleaned eye tracking signal, participant 9, text 2725
Fixations

- Dots aggregate to form a large blob of many dots
- How closely the raw sample dots are is directly related to sampling frequency of the eye-tracker
- How smooth the raw data appears is proportional to the accuracy of the eye-tracker
- Usually calculated by an algorithm from a raw data => parameters impact a lot
- Oculomotor, algorithmically derived and cognitive fixations overlap but are not identical
- Word could be processed before being read
Fixations

• Relatively stable eye-in-head position:
  – 90% viewing time is devoted to fixations
  – duration: 100 ms +
  – Spatial dispersion : < 2°
  – Threshold velocity : < 15 – 100° /msec
Saccades

- Rapid eye movements between fixations used to reposition fovea
- Involuntary, abrupt, rapid, small movements of both eyes simultaneously in changing the point of fixation
- Effectively blind during transition
Saccades: parameters

- Amplitude - the size of the saccade, usually measured in degrees or mins. of arc
- Peak Velocity - the highest velocity reached during the saccade. Velocity profile is symmetric
- Duration - the time taken to complete the saccade
- Latency - the time taken from the appearance of a target to the beginning of a saccade in response to that target
Saccades: parameters

• Duration: 10 ms – 100 ms
• Velocity: 400 - 600°/sec
• Velocity threshold: 30-100°/sec
• Skipping words: 2-3-letter words are skipped 75%, 8-letter words are never skipped;
• Regressions: 10-letter spaces happen because of problems in understanding text
Blinks

- Rapid bilateral eyelid closure and co-occurring eye movement
Salvucci & Goldberg, 2000

- Identifying Fixations and Saccades in Eye-Tracking Protocols
- Event detection algorithm can have a significant impact on the results of higher-level analyses
- Badly designed detection algorithm can detect too little or a lot of fixations
- The authors propose taxonomy to facilitate the comparison of different algorithms
## Taxonomy

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Spatial</th>
<th>Temporal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Velocity-based</td>
<td>Duration sensitive</td>
</tr>
<tr>
<td></td>
<td>Dispersion-based</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Area-based</td>
<td>Locally adaptive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I-VT</td>
<td>I-HMM</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
Taxonomy

- Velocity-based algorithms emphasize the velocity information in the eye-tracking protocols, taking advantage of the fact that fixation points have low velocities and saccade points have high velocities.

- Assuming a constant sampling rate, velocities are simply distances between sampled points and thus we can ignore the temporal component implicit in velocities.
Taxonomy

• Dispersion-based algorithms emphasize the dispersion (i.e. spread distance) of fixation points, under the assumption that fixation points generally occur near one another.

• Area-based algorithms identify points within given areas of interest (AOIs) that represent relevant visual targets.
Velocity-threshold Identification (I-VT)

- Based on **point-to-point velocities**:
  - Threshold: 20°/sec, Sen & Megaw, 1984
  - When only point-to-point velocities are known, infer the velocity threshold based on sampling frequency along with some exploratory data analysis
  - Fixations (< 100°/sec)
  - Saccades (> 300°/sec)

- **Pseudo Code Algorithm** (protocol, velocity threshold)
  - Calculate point-to-point velocities for each point in the protocol
  - Label each point below velocity threshold as a fixation point, otherwise as a saccade point
  - Collapse consecutive fixation points into fixation groups, removing saccade points
  - Map each fixation group to a fixation at the centroid of its points
  - Return fixations in format <x,y,t,d>
Dispersion-based algorithm (I-DT)

- Based on groups of consecutive points within a maximum separation
  - Moving window initially spans minimum consecutive data points based on duration threshold and sampling frequency
  - Checks dispersion of points in window by summing differences between the points’ max and min x and y values
  - \( D = [\max(x) - \min(x)] + [\max(y) - \min(y)] \). Note that alternative dispersion metrics could be based upon spatial variance or area of samples.
  - When dispersion is above the dispersion threshold, the window does not represent a fixation, and the window moves one point to the right
  - If the dispersion is below the dispersion threshold, the window represents a fixation. Expand the window to the right until the dispersion is above threshold.
- Output is \( <x,y,t,d> \)
- Requires two parameters, the dispersion threshold and the duration threshold
Dispersion-based algorithm (I-DT)

I-DT (protocol, dispersion threshold, duration threshold)

While there are still points

  Initialize window over first points to cover the duration threshold

  If dispersion of window points <= threshold

    Add additional points to the window until dispersion > threshold
    Note a fixation at the centroid of the window points.
    Remove window points from points

  Else

    Remove first point from points

Return fixations
R package: emov (github)

• An R package for fixation and saccade detection in eye tracking recordings
• This package implements a dispersion-based algorithm (I-DT)
• Detects fixations in the first place, compared to the velocity threshold algorithms which detect saccades.