SAND2015-7135C

Exceptional service in the national interest

Sandia National Laboratories

Challenges in Eye Tracking Data Analysis: From the Laboratory to the Wild World of Information

Laura A. McNamara, David Stracuzzi and Kristina Czuchlewski

Sandia National Laboratories

Albuquerque, NM 87015

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Topical Outline

- **Visual Cognition at Sandia National Laboratories**
- **From Visual Inspection to Information Foraging**
- **Linguistic Inspiration: Behavioral Events as Texts**
- **A Concluding Thought**

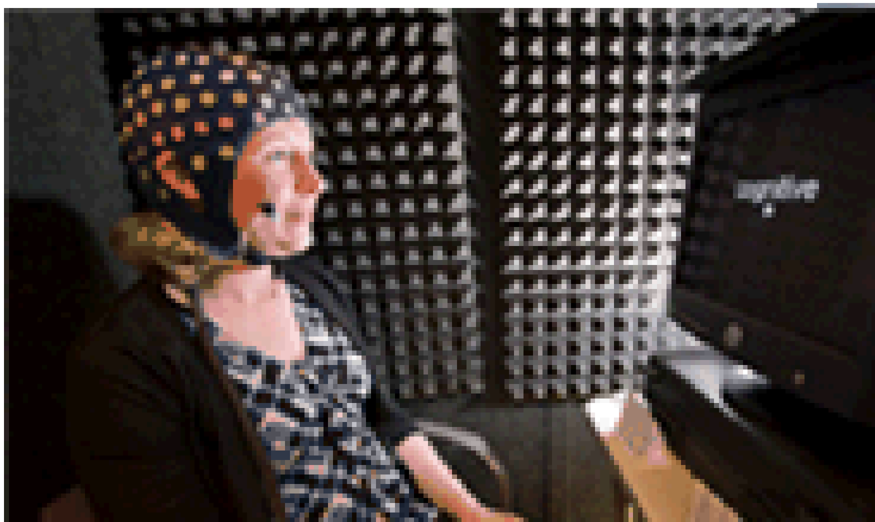
Topical Outline

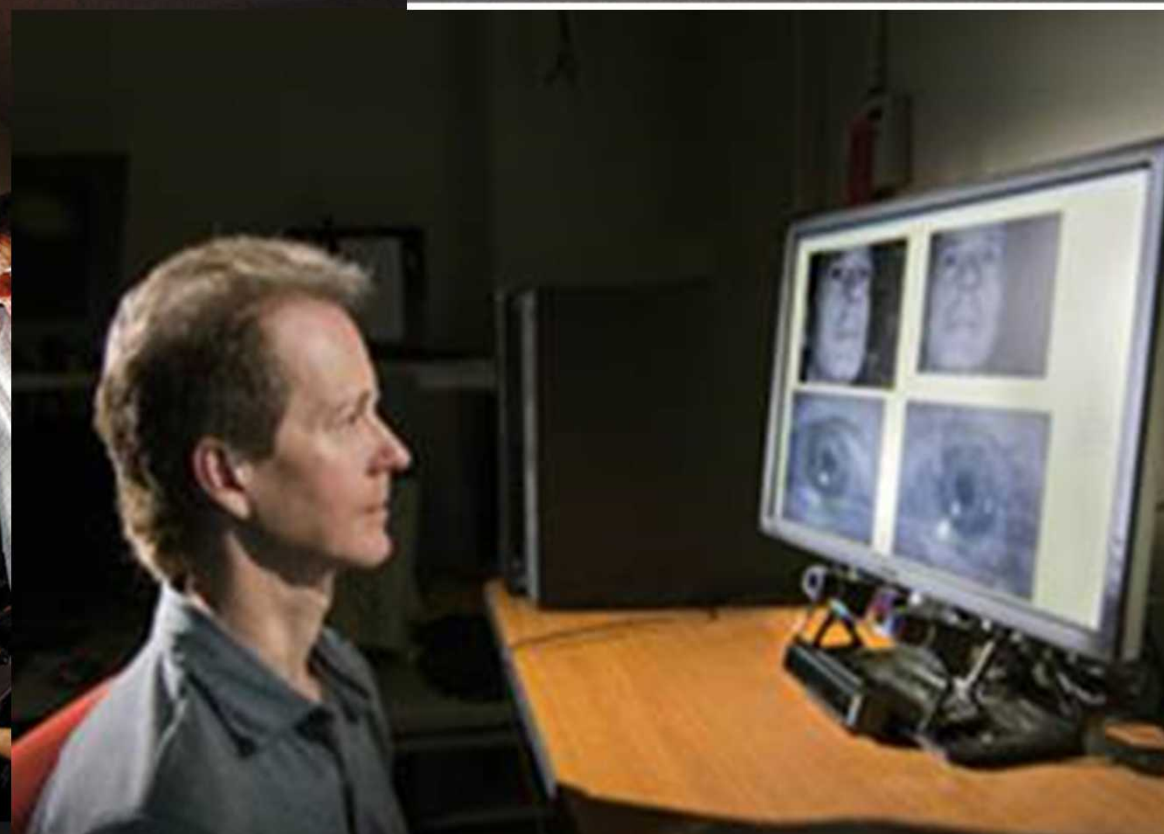
- **Visual Cognition at Sandia National Laboratories**
- From Visual Inspection to Information Foraging
- Linguistic Inspiration: Behavioral Events as Text
- A Concluding Thought

WordCloud from isymmetry.com, http://www.isymmetry.com/where-are-the-big-data-jobs last accessed 31 July 2015

Interdisciplinary Applied R&D Team

- Kerstan Cole, Human Factors
- Kristina Czuchlewski, Remote Sensing
- Kristin Divis, Cognitive psychology
- Michael Haass, Signal processing/physics
- Laura Matzen, Cognitive Neuroscience
- Laura McNamara, Organizational Anthropology
- J. Daniel Morrow, Computer science/analytics
- Susan Stevens-Adams, Cognitive Psychology
- David Stracuzzi, Computer science/mathematics





Visual workflows in diverse domains

How do top down and bottom up cues guide visual search?
What are the characteristics of domain-specific expertise in visual search?
How does domain expertise change domain-general cognition?

Image Analysis
EO, SAR, others

Base images and change products to enhance feature salience

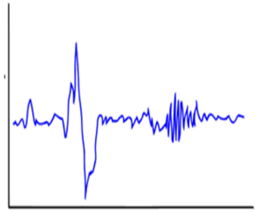
Security screeners
X-Ray image products

Enhance feature salience for visual detection

Novices

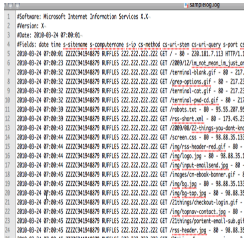
Natural scenes

Engineer Analysts
Waveforms



Visualizations of raw signal data

Cyber
Log Files



Raw log data, mostly text and hash

Topical Outline

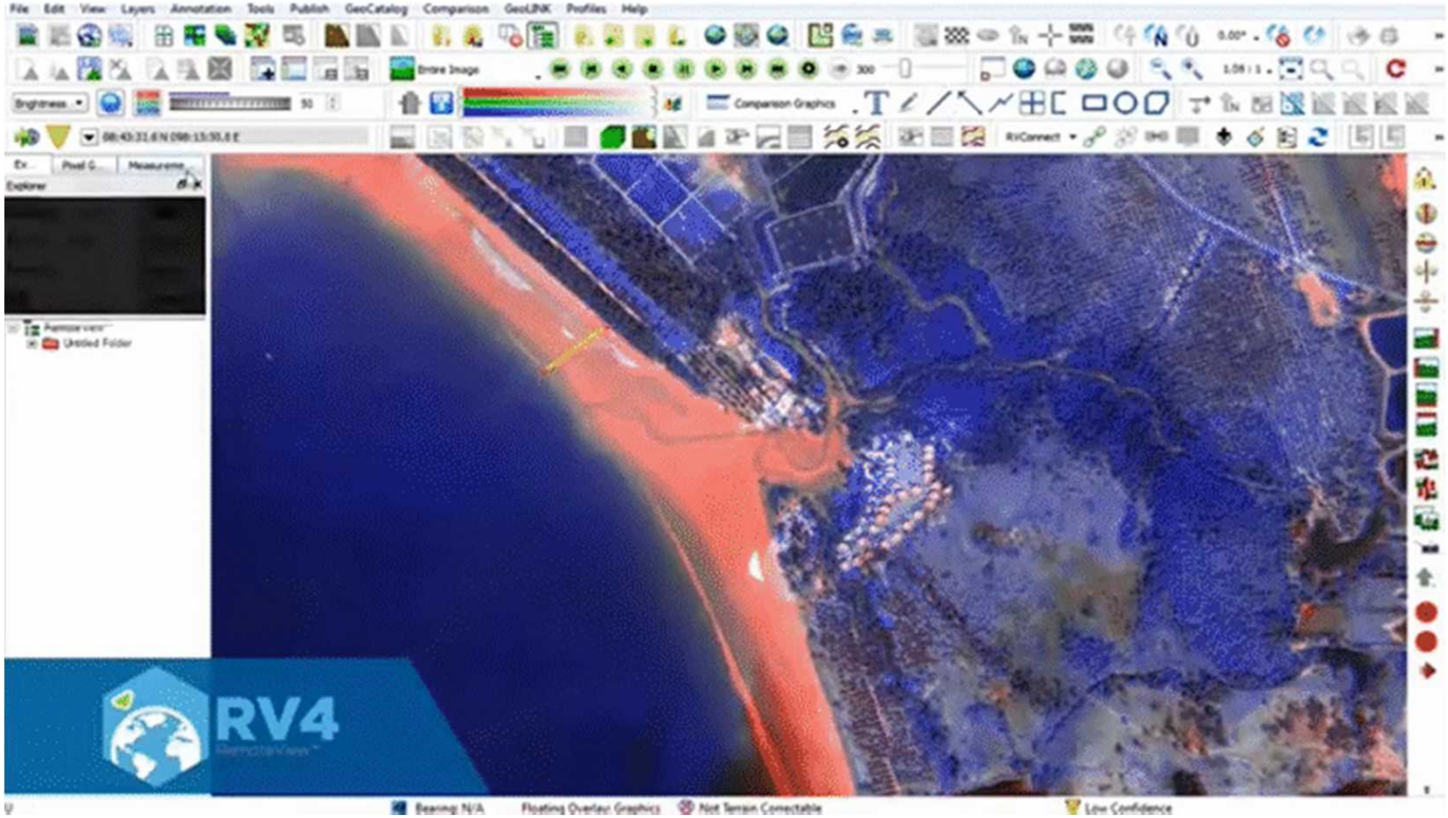
- Visual Cognition at Sandia National Laboratories
- **From Visual Inspection to Information Foraging**
- Linguistic Inspiration: Behavioral Events as Texts
- A Concluding Thought

Imagery Analysis as Visual Inspection

Models of Visual Inspection

- ***Select*** object to inspect
- ***Search*** object for anomalies
- ***Classify*** detected anomalies
  - What does this anomaly mean?
  - Does it merit concern?
- **Act** on anomaly
  - Do we invest resources in addressing the detected anomaly?

…Until the Age of Softcopy.



GIF excerpted from Textron Systems Geospatial Solutions RV4 Video, https://www.youtube.com/watch?v=y28nqAaAA_Y


It’s really visual inspection with foraging.

Hunting information in the wild

- Human informavores rely on learned semantic cues to navigate massive information spaces
- Valuable information isn't distributed regularly or evenly – instead, it's patchy
- We look for *proximal semantic cues*, or 'scents,' to navigate to patches of interest
- An effective foraging strategy should *maximize the value of information gained* in relation to *effort finding the information*
  - Time spent *between patches* vs. time spent *within patches*

Xerox PARC - Peter Pirolli, Stuart Card, Ed Chi

HUMAN–COMPUTER INTERACTION, 2007, Volume 22, pp. 355–412

SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web

Wai-Tat Fu
University of Illinois at Urbana-Champaign

Peter Pirolli
Palo Alto Research Center

Gaze behaviors:
What the individual is looking at, for how long, and in what order

Cognitive Workload Measures: Dual-task data, pupil-based measurements; possibly head position, facial expressions?

Geospatial imagery interaction: Movement across images, changes in image resolution, geospatial coordinates, time on screen, movement among available views, etc.

Manual operator actions
Mouse clicks, key presses, menu-based commands

Software application/tool usage: Opening, closing, moving among different functions within software and/or across different software applications; cutting and pasting data, creating products, etc.

Topical Outline

- Visual Cognition at Sandia National Laboratories
- From Visual Inspection to Information Foraging
- **Linguistic Inspiration: Gaze Events as Text**
- A Concluding Thought

Elements of a Gaze-Informed Foraging Model

Dynamically changing content →

Gaze tracked against display surface →

User operations (mouse, keyboard) →

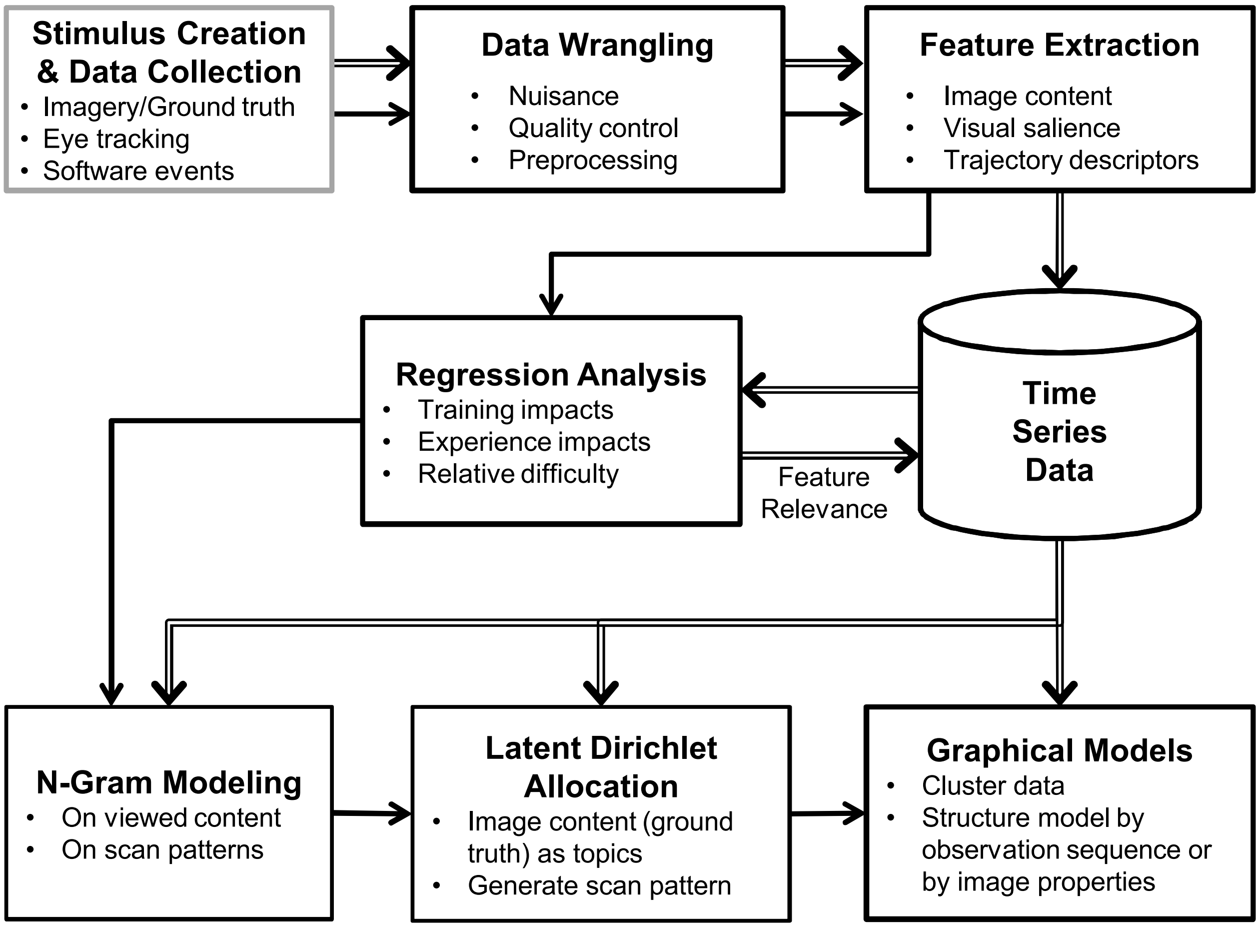
Application/content log files →

VECTOR OF EVENTS
integrated into a single timeline with minimal latency (~100-500 milliseconds).

T_0 → T_n

Developing models of visual cognitive strategies with geospatial temporal datasets will require integrated analysis of user operations, application logs, gaze patterns, and displayed content. Relevant events must be defined, identified, and synchronized with minimal temporal latency before we can develop the models needed to understand what makes a visual analytic system "effective" in supporting accurate signature search and characterization, particularly when data are noisy and/or ambiguous.

When a cognitive/AI scientist meets eye tracking data…



… This is the kind of thing you can expect. ☺

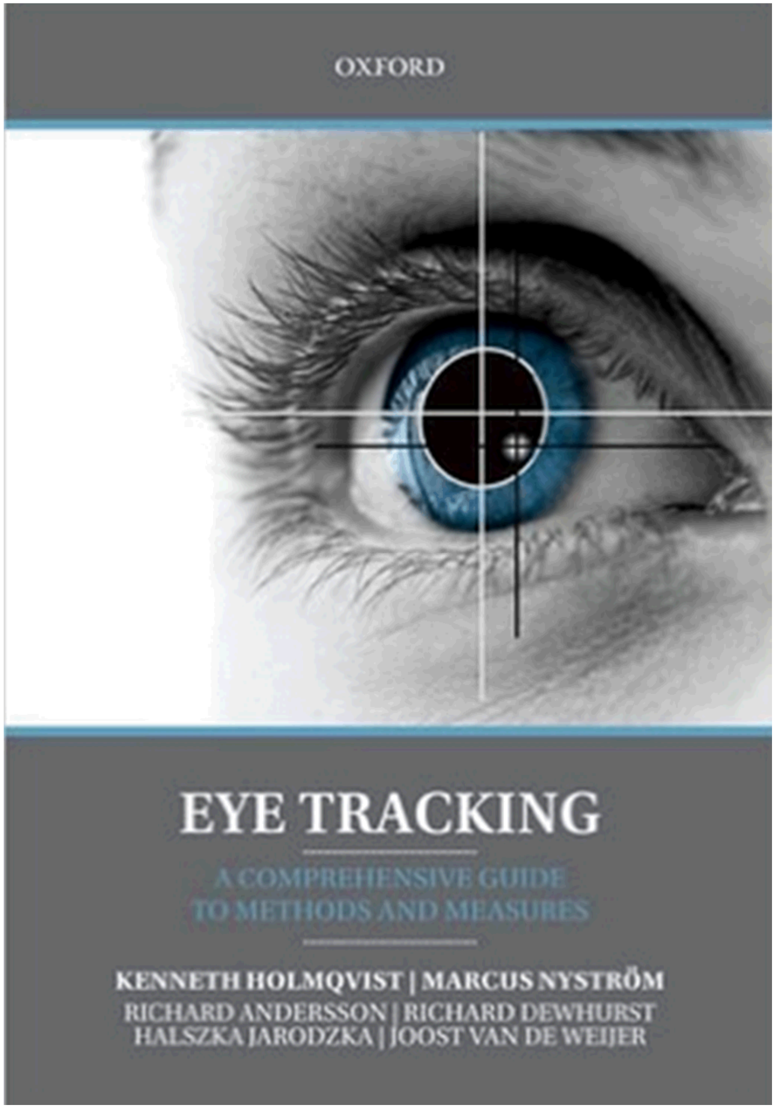
Human Foraging Behaviors as *Texts*

- How do we tease out structure from the sequences of events descriptive of user actions? Computational linguistics deals with a lot of the same problems.
- **N-Grams:** Probabilistic prediction of next item in a sequence of items, based on analysis of sequential patterns in previous *n - 1* observations
  - Rooted in Shannon's theory of information
  - Imagine putting a small window over a set of characters (a word or 'gram') so that only *n*-words are visible at one time.
  - Given observed events, what observation is most likely to occur next?
- Tease out local patterns within longer sequences of observations to examine structural regularities in data
  - Given a pattern of panning, zooming, and gaze behaviors, what's the probability that Imogene the Imagery Analyst will open a new image window?
- Enable comparison/contrast of imagery analytic strategies within and across analyst communities

Topical Outline

- Visual Cognition at Sandia National Laboratories
- From Visual Inspection to Information Foraging
- Linguistic Inspiration: Gaze Events as Text
- **A Concluding Thought**

Sandia National Laboratories



Eye-tracking data have a structure that lends itself to many different types of data analysis...

Analyzing eye-movement data is always a matter of reducing the amount of information in the data.

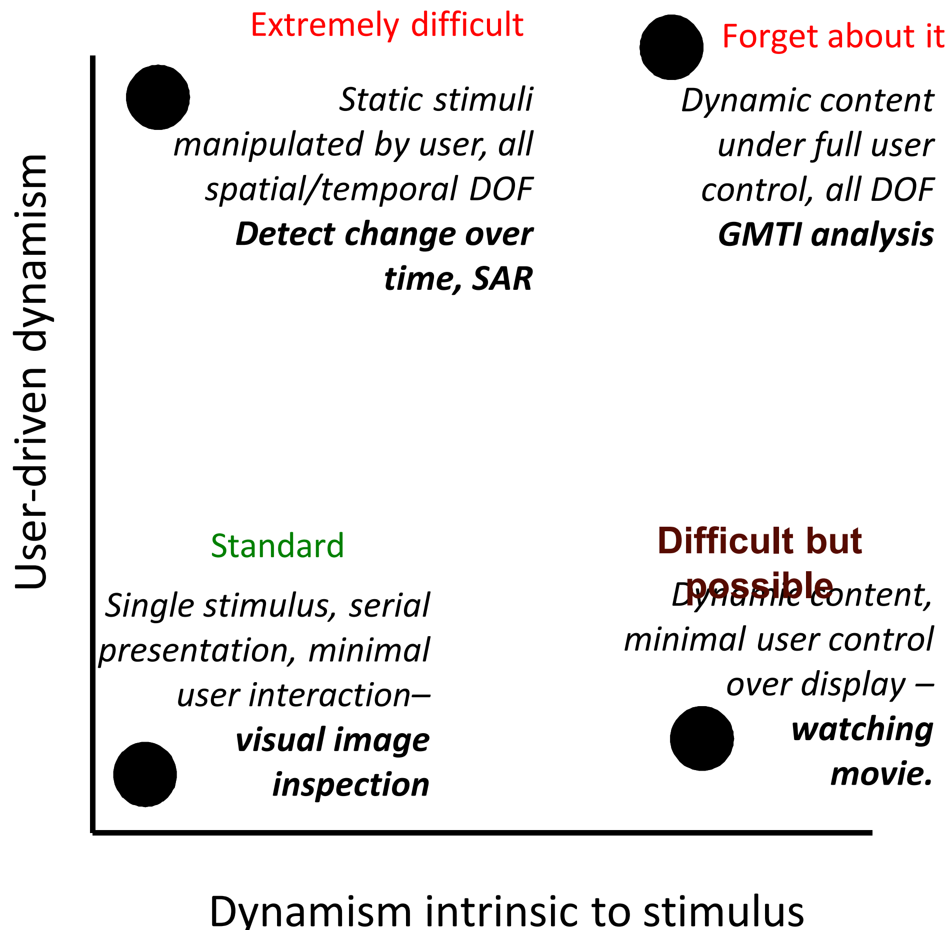
-Holmqvist, Nustrom, Andersoon, dewhurst, Jarodzka and van de Weijer 2011: 465

AUX SLIDES

The Problem

Gaze tracking technology originated in laboratories, where stimuli and tasks are highly structured and controlled. Infrastructure for naturalistic studies of user-driven behaviors doesn't exist – yet.

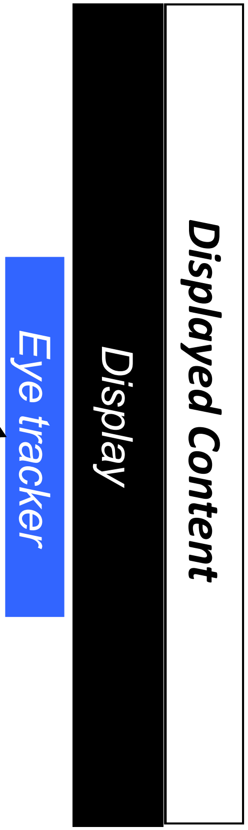
- Gaze tracking is a potentially powerful source of behavioral data for perception and cognition in complex information search, acquisition, and problem-solving
- We'd love to incorporate gaze tracking into studies of visual cognition as people interact with large, heterogeneous datasets.
  - User interaction drives change of stimulus content
  - Stimuli may be intrinsically dynamic – moving targets, for example
- However, gaze/eye tracking algorithms, software don't cope well with dynamism
  - Manageable when the stimulus is dynamic
  - The more the user is controlling what's displayed, the harder it gets
  - Even harder for capturing emergent search/problem solving behaviors – can't predefine "targets" nor contingent visual events

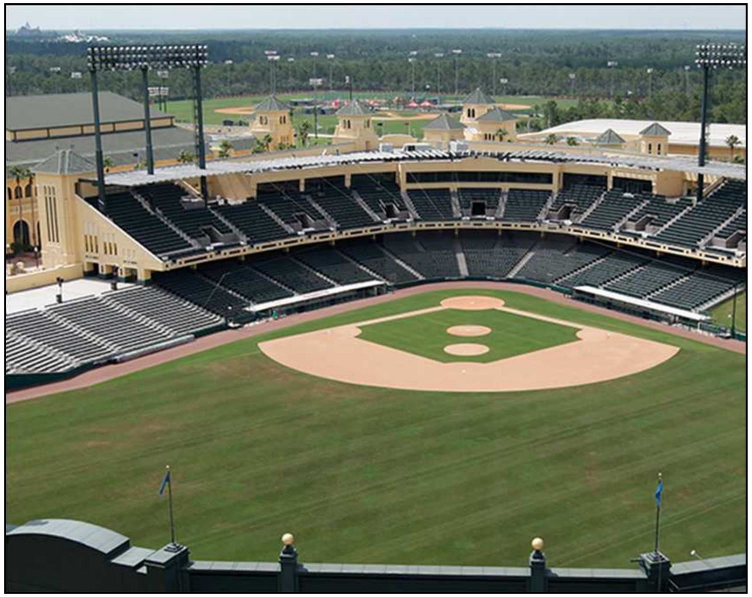


What features capture Charlie's attention as he inspects a baseball diamond?

Gaze sampling can vary from a few tens to ~1000 Hz, depending on system

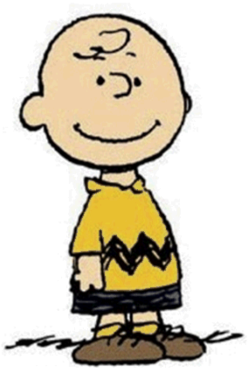
Charlie Brown, Imagery Analyst
Transnational Baseball Diamond Team



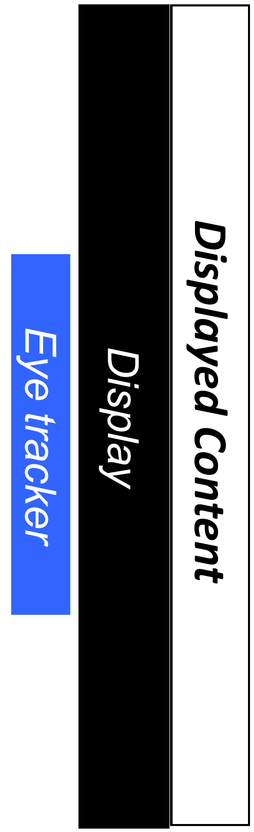


Content "slices" sampled @ 10 Hz throughout user session.
If content doesn't change, mapping gaze-to-display-to-content is straightforward.

But what if Charlie is looking for the perfect pitcher's mound?



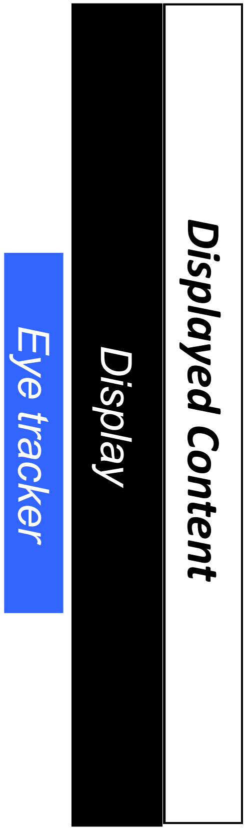
Charlie pulls up a few baseball diamonds, flickers across them to compare features. He briefly examines a close-up shot of home plate. He focuses tightly on one pitcher's mound, then another, and moves among these and home plate to compare ground characteristics.



When we can't predict what images Charlie will use, nor how he will interact with them, gaze tracking is very difficult – it usually requires laborious hand-coding.

Charlie has videos of teams playing at different fields.

Charlie watches a recorded game from one of the fields. He stops the video, rewinds to examine the pitcher's stance, stops, zooms in on the mound, toggles back and forth. He then starts a video from another game at a different field, toggling back and forth between the videos, rewinding and zooming and panning.



When content is dynamic, and when we can't predict which features will capture Charlie's attention, mapping gaze-to-display-to-content is *extremely* challenging.