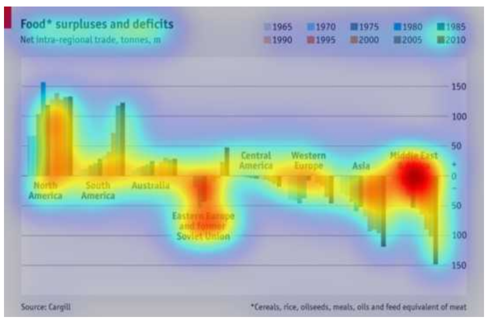
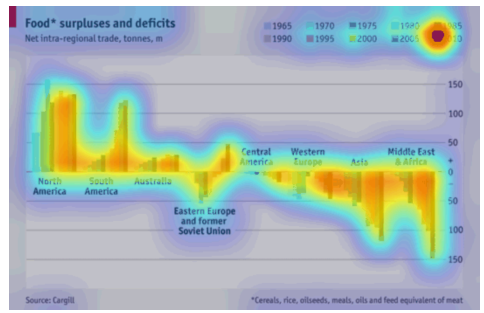
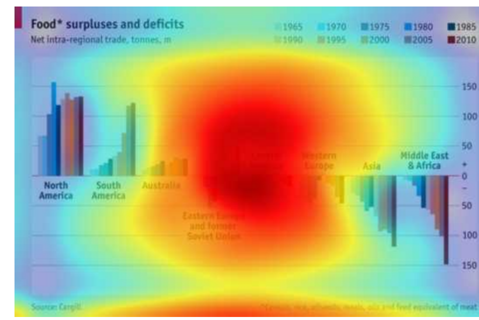
SAND2016-1008PE

Exceptional service in the national interest

Sandia National Laboratories







Modeling Human Comprehension of Data Visualizations

Michael Haass, Laura Matzen, Andy Wilson,
Kristin Divis, Mika Armenta, Laura McNamara

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

Outline

- My background
- Sandia National Laboratories
- Cognition at Sandia
- Visualization at Sandia
- Prior work – applied visual cognition & modeling
- Our current project – human comprehension of data visualizations

MS Physics

- *Optical instrumentation, thesis work at NIST, Gaithersburg, MD, Soft x-ray spectroscopy*

13 yrs. private industry, measurement solutions for life sciences

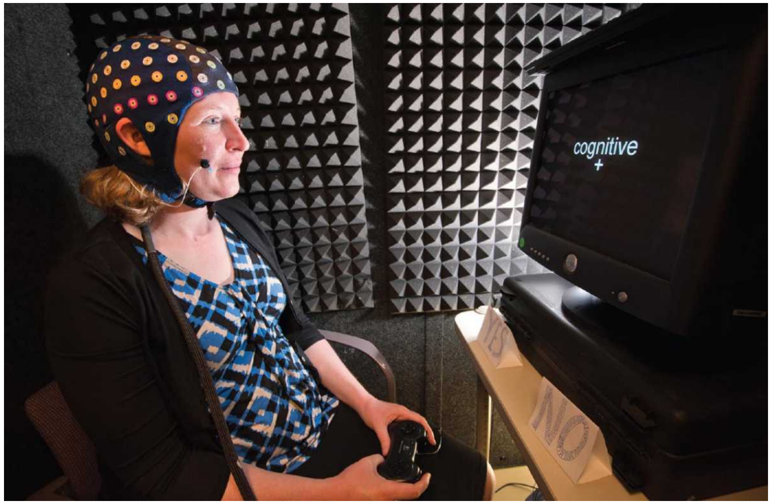
- *NIR spectroscopy, multivariate analysis, modeling and simulation, bus. dev., project/program management*

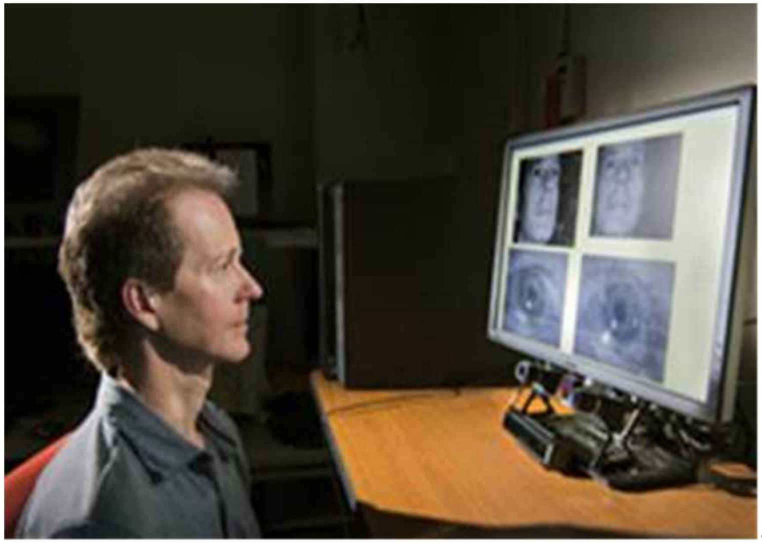
Joined SNL 2009

- *Knowledge capture, EEG signal analysis & BCI, automated team training technologies, pattern analytics*, …

Cognition at Sandia

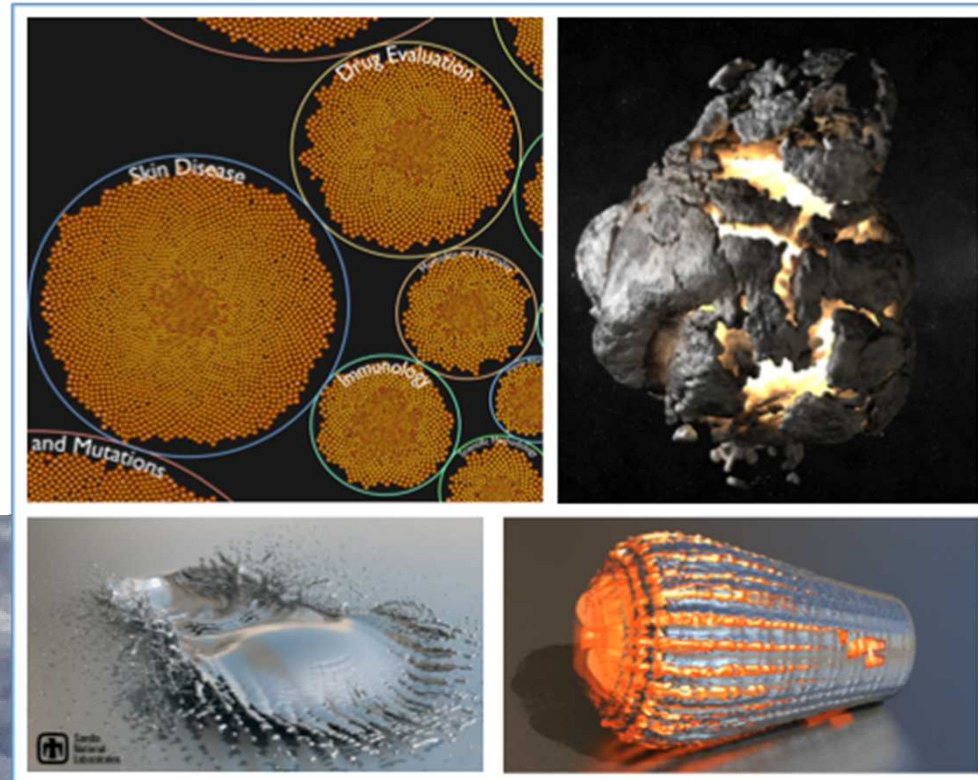
- Solutions that include both technology and human cognition aspects
  - understanding human decision making
  - improving human performance
  - human-centric data collection and analysis
  - advanced software development
  - ethics, legal, and social issues



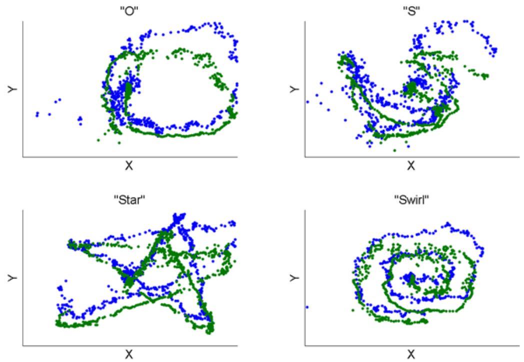


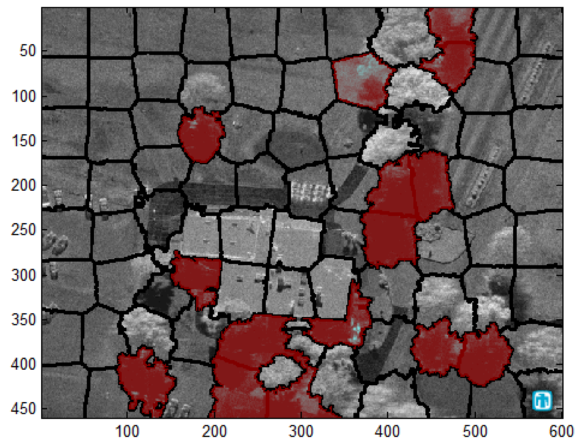
Visualization at Sandia

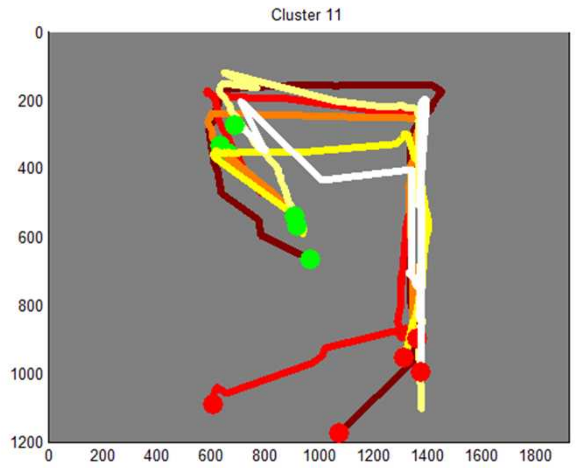
- Scalable techniques for providing decision support in national security challenges
  - advanced algorithms
  - innovative hardware architectures
  - scalable analysis components
  - state-of-the-art capabilities for critical data modeling and analysis applications



Exceptional service in the national interest





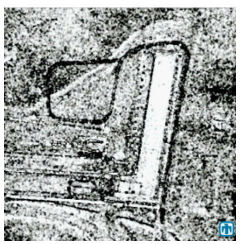


Applied Visual Cognition & Modeling

Visual Cognition: Core Scientific Questions

What features capture attention in non-optical imagery?
How does domain experience influence visual search/inspection?
How can top-down visual attention be modeled?
Do people with expertise in one domain perform differently on domain-general tasks?

SAR

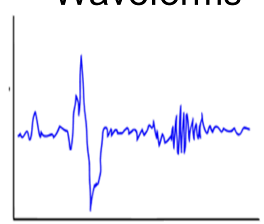


Intended to make important features more salient

False color X-rays

Intended to make important features more salient

Waveforms

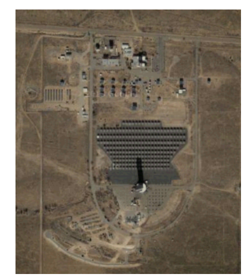


Visualizations of raw data

Log Files

Raw data

Satellite Imagery



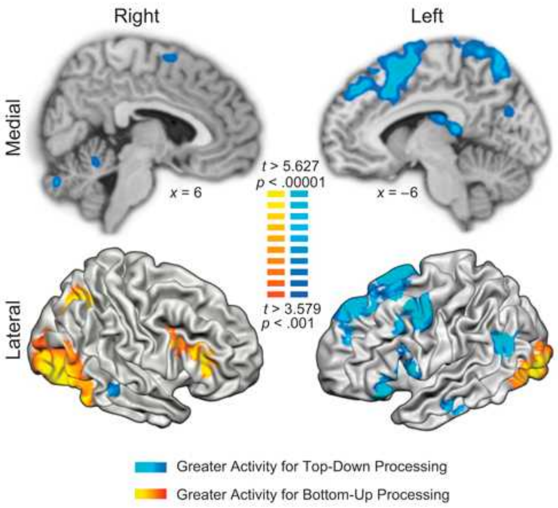
Similar to optical imagery

Novices

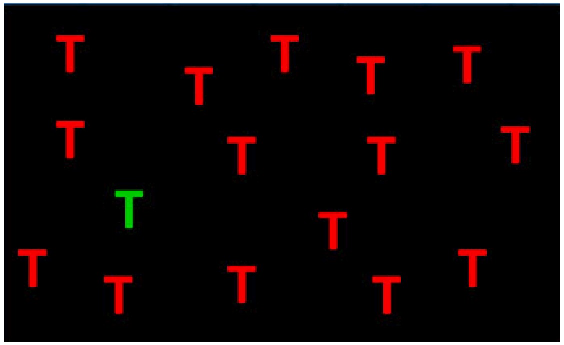
Experienced with optical imagery only

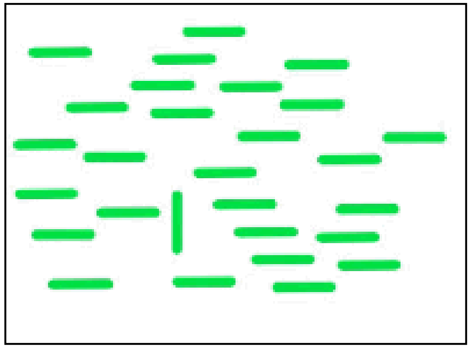
Bottom-up versus top-down visual processing

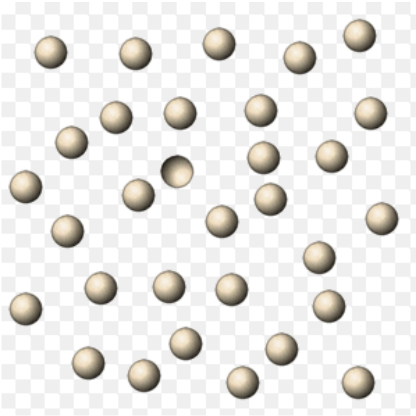
- Two parallel neural processes that guide visual processing
  - Bottom-up = stimulus-driven visual attention
  - Top-down = goal-oriented visual attention



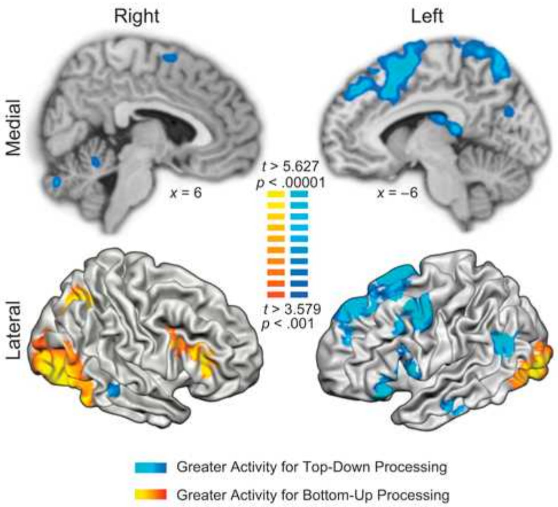
- Bottom-up attention is captured *automatically* by the physical properties of a stimulus
  - Color, shape, orientation, motion



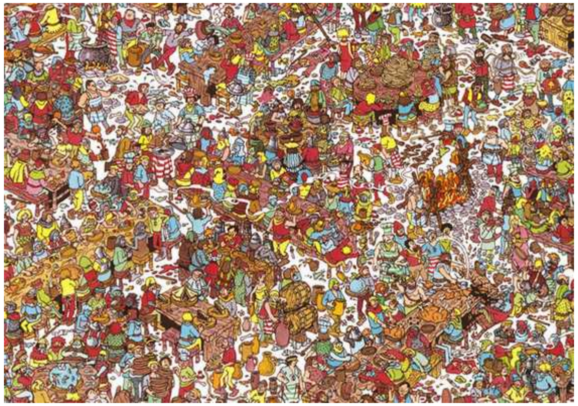


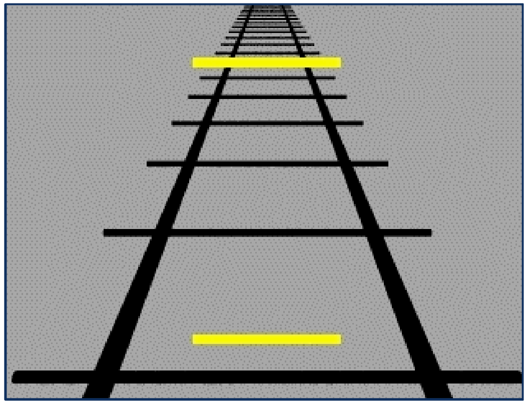


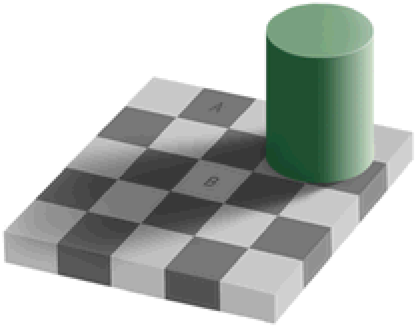
Bottom-up versus top-down visual processing



- Two parallel neural processes that guide visual processing
  - Bottom-up = stimulus-driven visual attention
  - Top-down = goal-oriented visual attention
- Top-down attention is allocated *voluntarily* according to the viewer's goals and expectations
  - Current goal, past experience, cognitive load

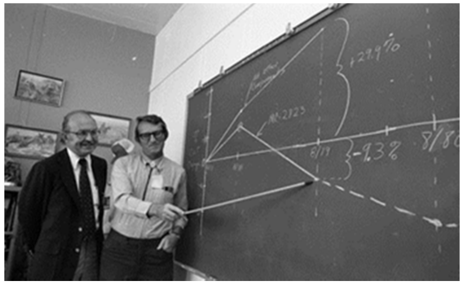






Exceptional service in the national interest

Sandia National Laboratories



Methodology for Knowledge Elicitation in Visual Abductive Reasoning Tasks

Michael J. Haass, Laura E. Matzen, Allen R. Roach ,
Susan M. Stevens-Adams
Sandia National Laboratories

HCII Conference, 5 Aug. 2015

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

Bias in Knowledge Elicitation

- Bias: misinterpretation or misrepresentation of expert knowledge or data

Motivational Bias

- Social pressure
  - Group or interviewer
  - Image of self
- Subtleties of language and mental models

Cognitive Bias

- Anchoring
- Inconsistency
- Actual-ideal discrepancies
- Availability
- Estimation of uncertainty

- The potential for bias to affect the results of knowledge elicitation studies is well recognized
  - Attempt to control for bias through careful selection of elicitation and analysis methods

For a thorough discussion of bias, see Booker, M.A.M.J.M., Eliciting and Analyzing Expert Judgment: A Practical Guide, in Eliciting and Analyzing Expert Judgment. 2001, ASA-SIAM. Chapter 3.

Motivation

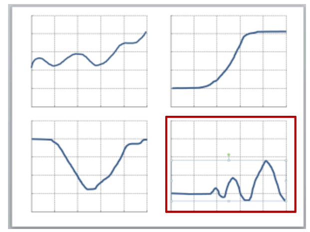
- New physiological sensors can provide additional dimensions of objective measurements, for example…

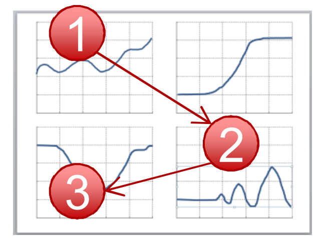
Sensor	Cognitive Attributes	Related Bias
 Eye Movements • Head mounted • Screen mounted	• Attention allocation • Workload	• Actual-ideal discrepancies
 Heart Rate • Chest strap • Wrist watch • “All-in-one” biophysical sensor systems	• Physical effort • Cognitive workload • Stress	• Social pressure
 EEG • Portable high fidelity • Gaming and neurofeedback headsets	• Error related negativity • Memory encoding • Drowsiness	• Inconsistency • Availability

Enhanced Knowledge Elicitation Methodology

- Incorporate one or more physiological sensors that provide cross referencing information for more traditional knowledge elicitation instruments
- Highlights actual-ideal discrepancies that can be missed during interviews and verbal walkthrough protocols
- Applied to a complex visual abductive reasoning task
  - Engineers who use multivariate time series data to diagnose the performance of devices throughout the production lifecycle

Engineer's Task





Pass/Fail
Cause?

Abductive Reasoning

What?

How?

Why?

Vision

Visual inputs → Encoding → Selection ⇄ Decoding → For motor action or decision making

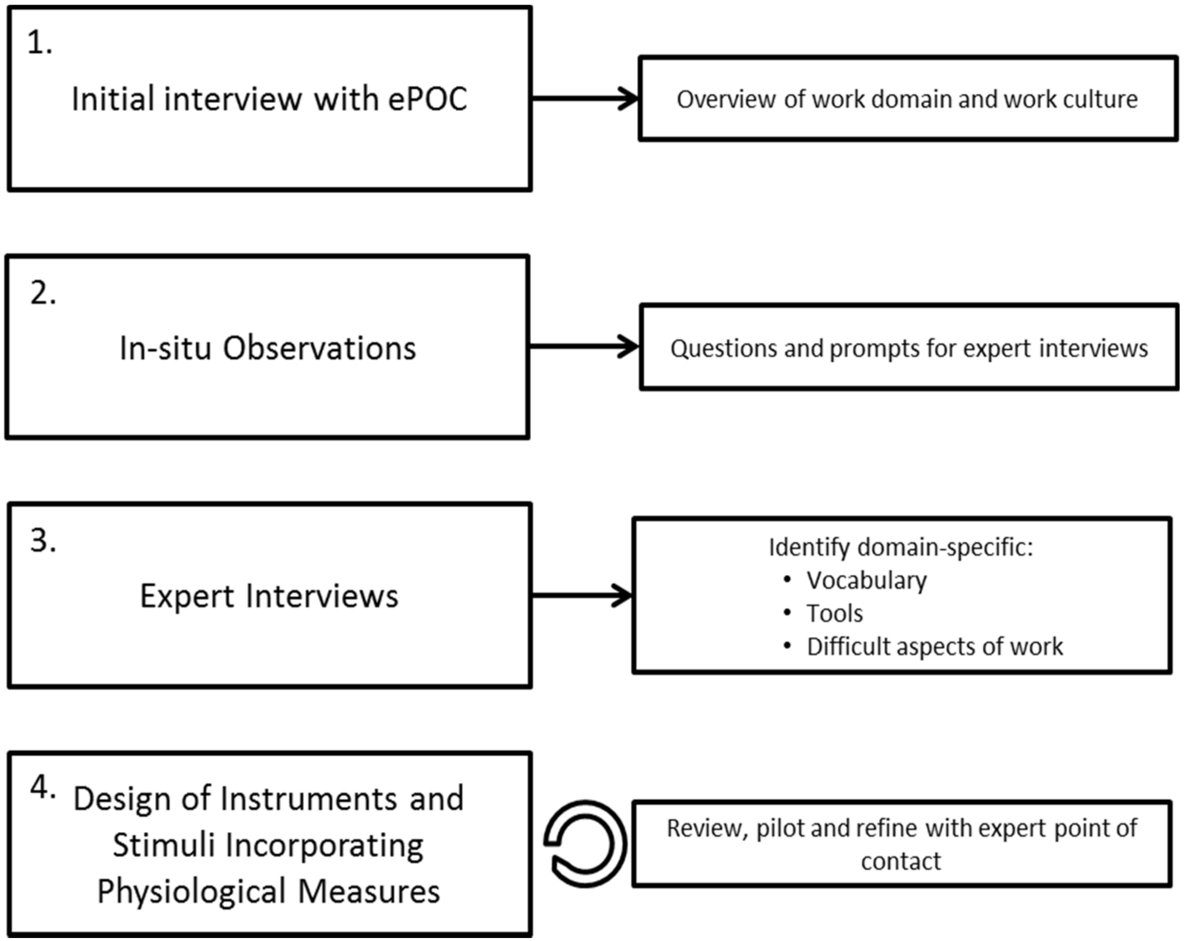
(e.g. by retinal neural activities)

(attentional selection, often by saccading to selected locations)

(e.g. face recognition)

Lower figure paraphrased from Understanding Vision; Theory, Models and Data, Li Zhaoping, Oxford University Press, 2014, Oxford, UK

Protocol Design Process

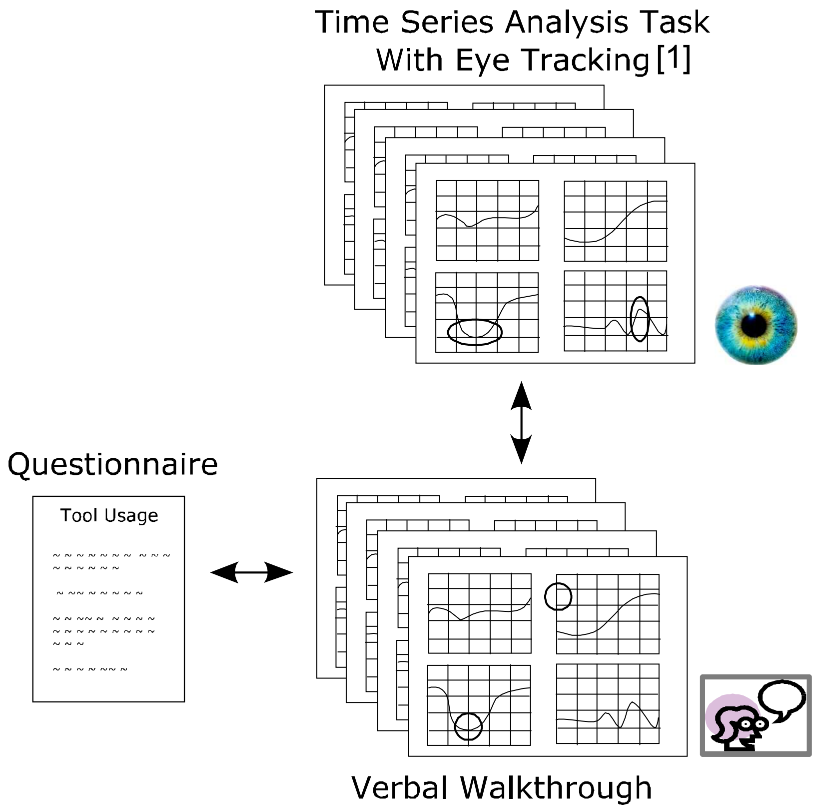


For a thorough discussion of design principles, see Booker, M.A.M.J.M., Eliciting and Analyzing Expert Judgment: A Practical Guide, in Eliciting and Analyzing Expert Judgment. 2001, ASA-SIAM. p. 3-16.

Study Overview

- 13 participants from highly specialized field
  - 3 highly experienced (“experts”)
  - 4 experienced (“practitioners”)
  - 6 without experience (“novices”)
    - For comparative performance baselines

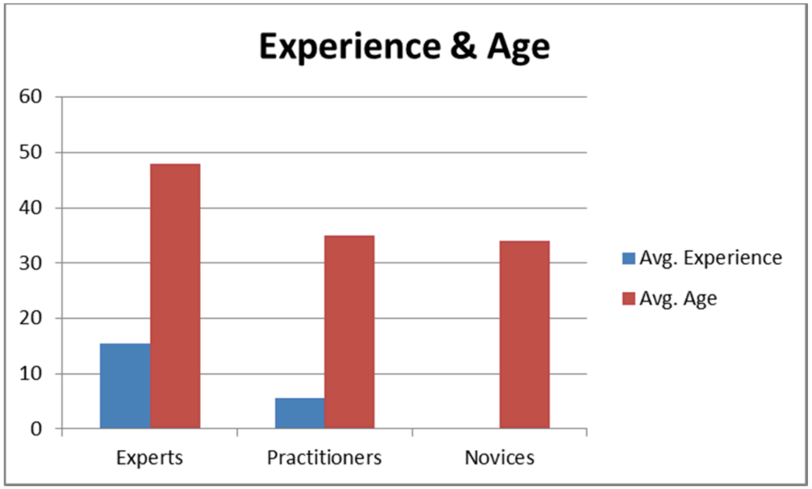
- Multivariate time series task and verbal walkthrough task
  - 15 trials for each subject

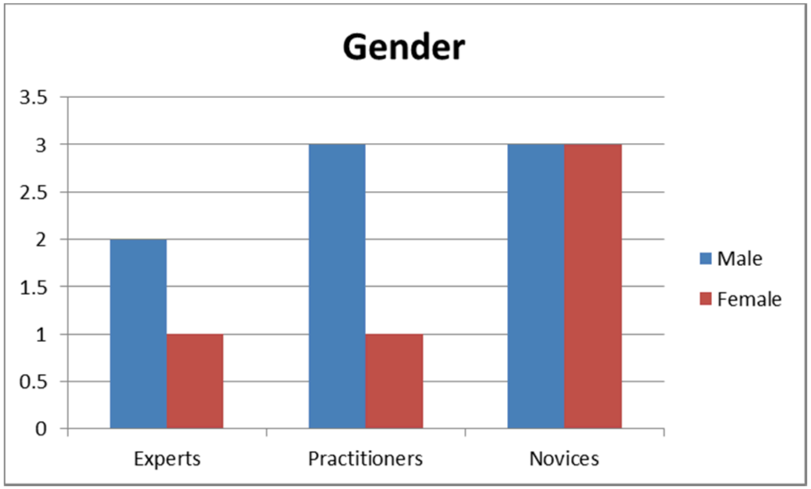


[1] FaceLAB 5 Standard System with two miniature digital cameras and one infrared illumination pod.

Subject Demographics

- All participants earned BA degree or higher
  - All but two earned graduate degrees



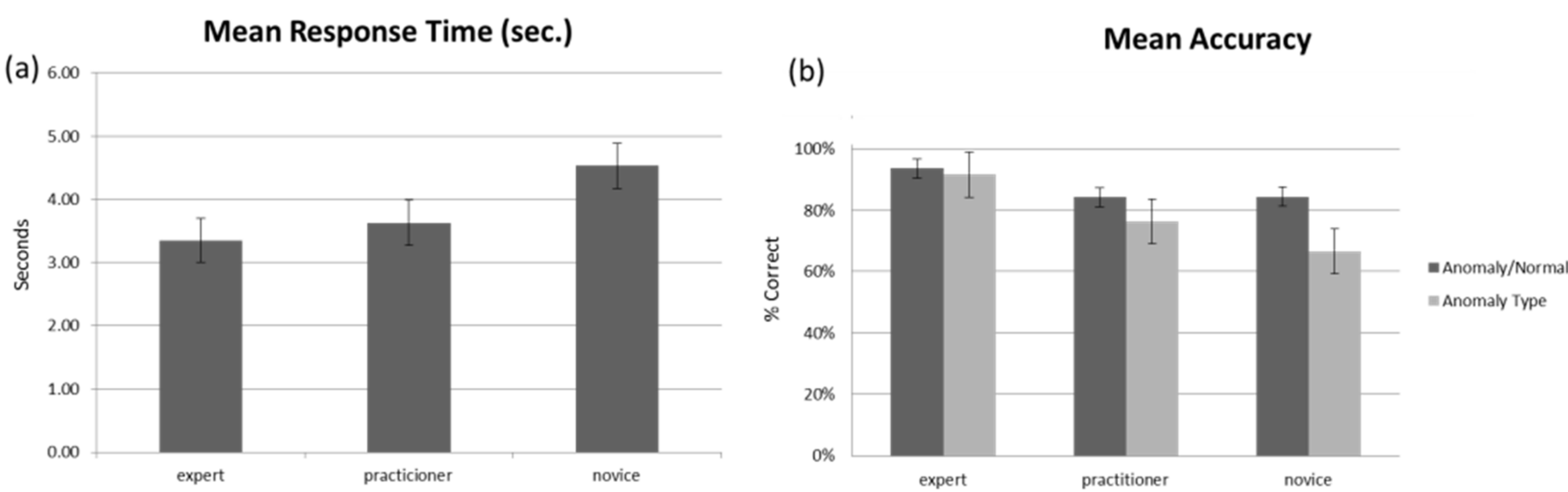


[1] FaceLAB 5 Standard System with two miniature digital cameras and one infrared illumination pod.

Analysis

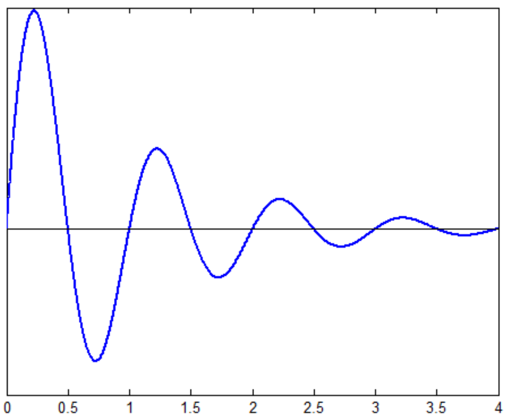
- Subject response times recorded by custom software written in Java
  - Subject responses for both the anomaly/normal decision and anomaly type were also recorded by this software
- Eye tracking fixation points and durations calculated using EyeWorks Analyze[1] software

✓ **Experts were faster and more accurate**

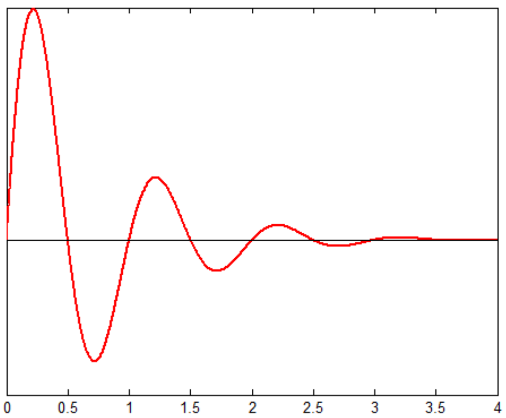


[1]EyeTracking Inc., 512 Via de la Valle, suite 200, Solana Beach, CA 92075, USA

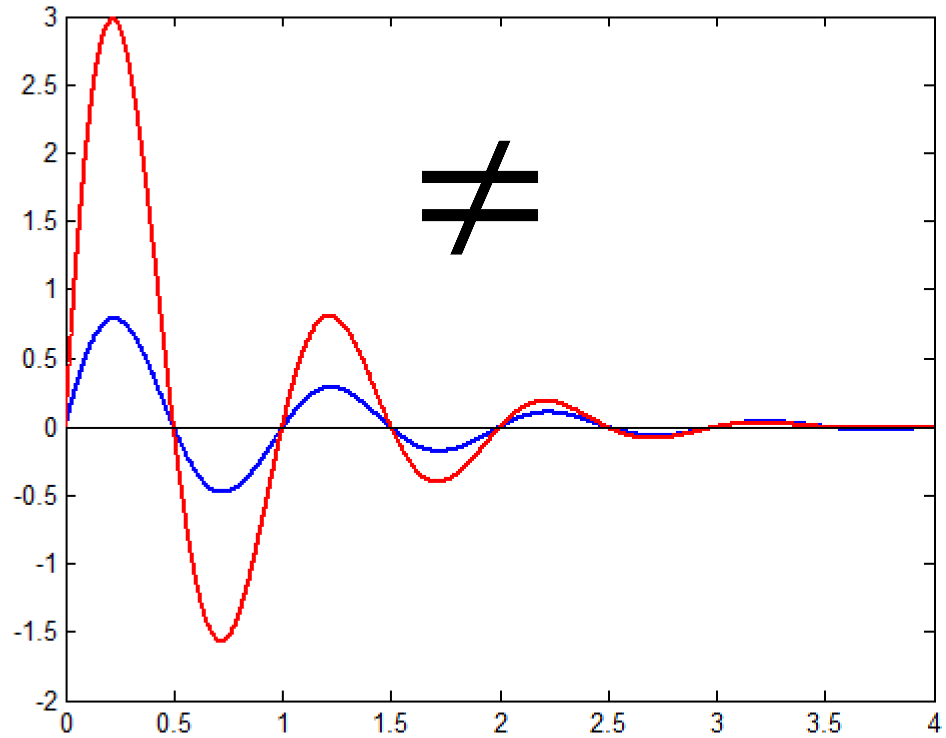
Shape Recognition Heuristic



$\stackrel{?}{=}$

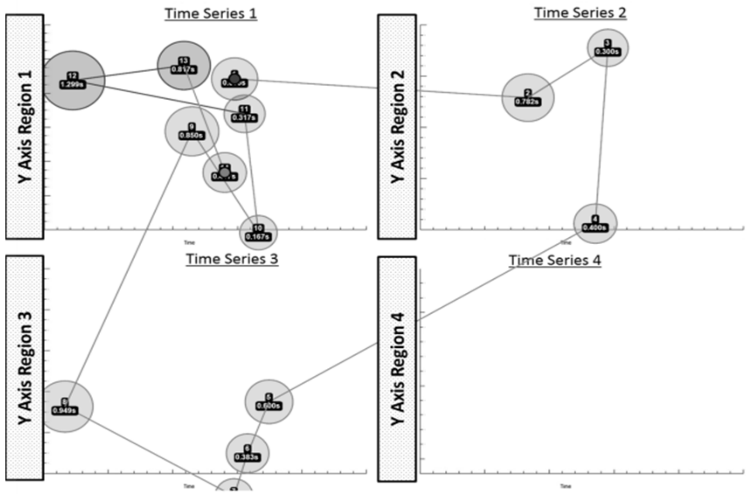


Shape Recognition Heuristic



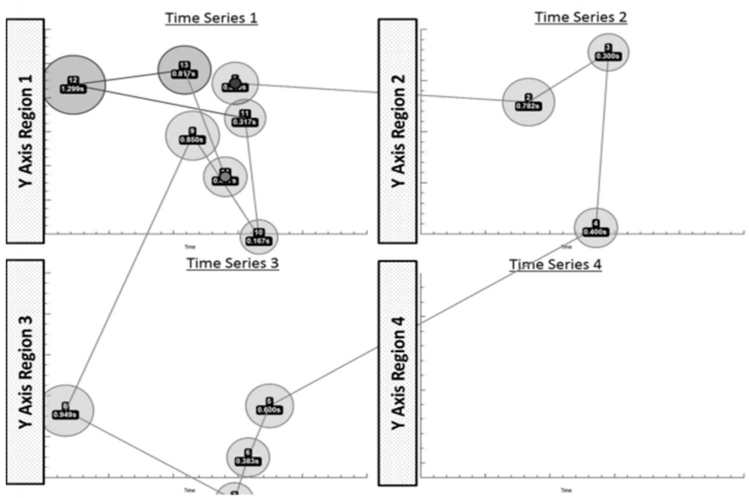
Results: actual-ideal discrepancies

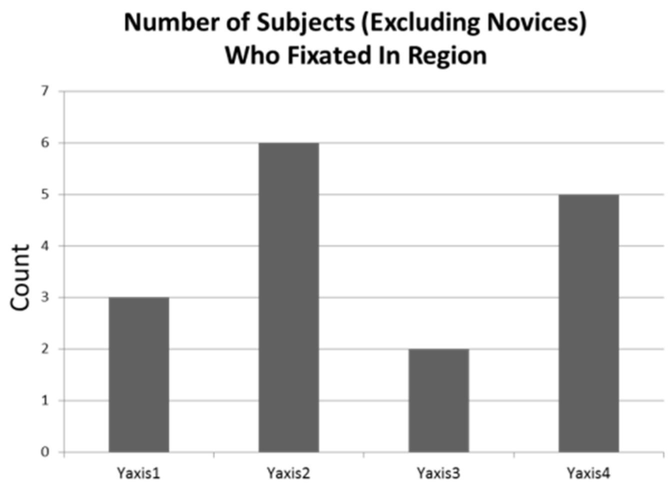
"Always check y-axis values"



Results: actual-ideal discrepancies

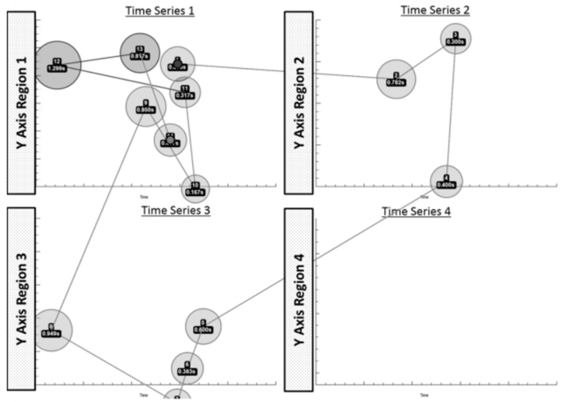
"Always check y-axis values"

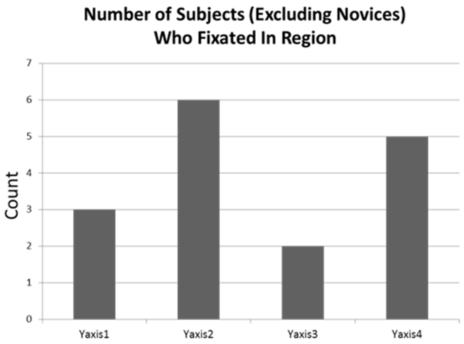




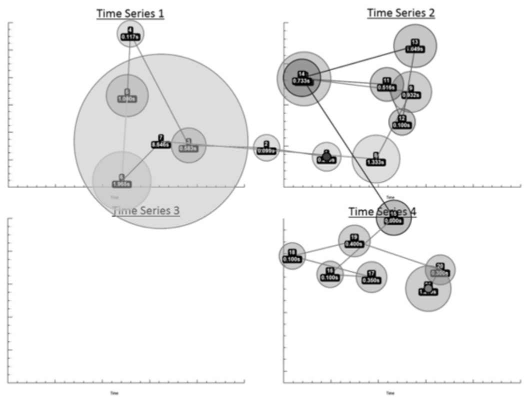
Results: actual-ideal discrepancies

"Always check y-axis values"





"Always check each data series"



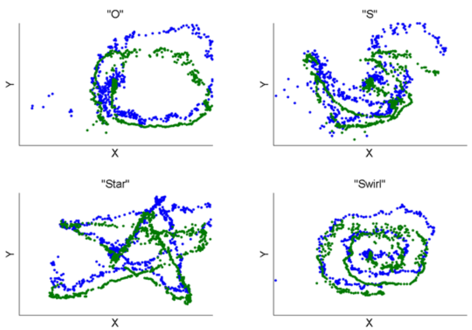
Experts seem to have developed fast shape recognition heuristic

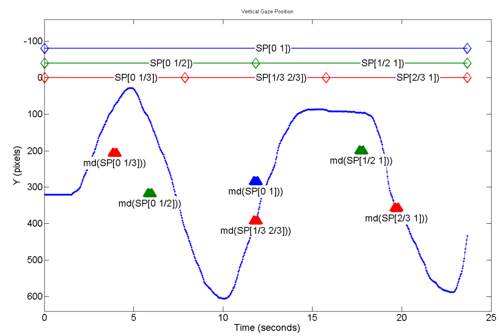
Summary and Conclusion

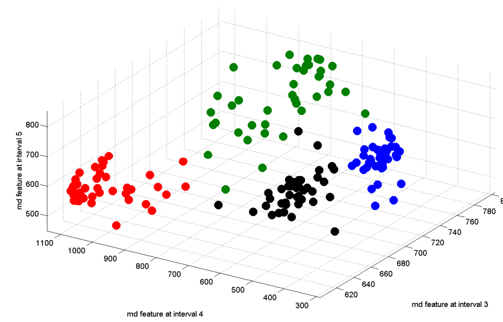
- Robustness is achieved through incorporation of one or more physiological sensors to provide cross referencing information for more traditional knowledge elicitation instruments.
- Eye tracking is effective at highlighting actual-ideal discrepancies that would not have been discovered by following a traditional verbal walkthrough protocol
- Future work
  - Apply to additional work domains and tasks
  - Develop detailed guidelines for selecting physiological sensors and metrics most appropriate for a given type of task or knowledge elicitation goal

Exceptional service in the national interest

Sandia National Laboratories





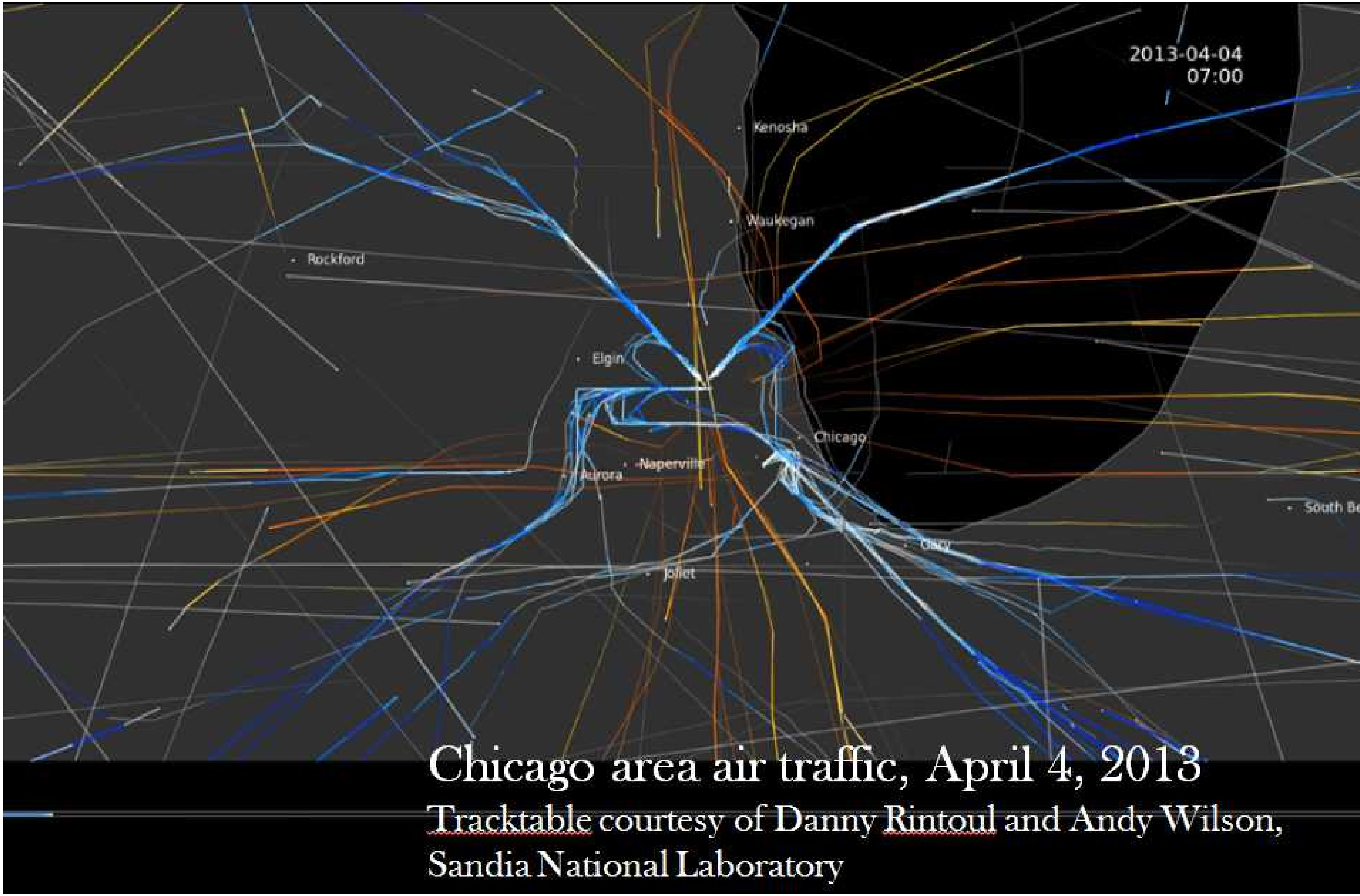


GazeAppraise*

Categorizing Gaze Trajectories

*Eye Tracking Research and Analysis, March 2016

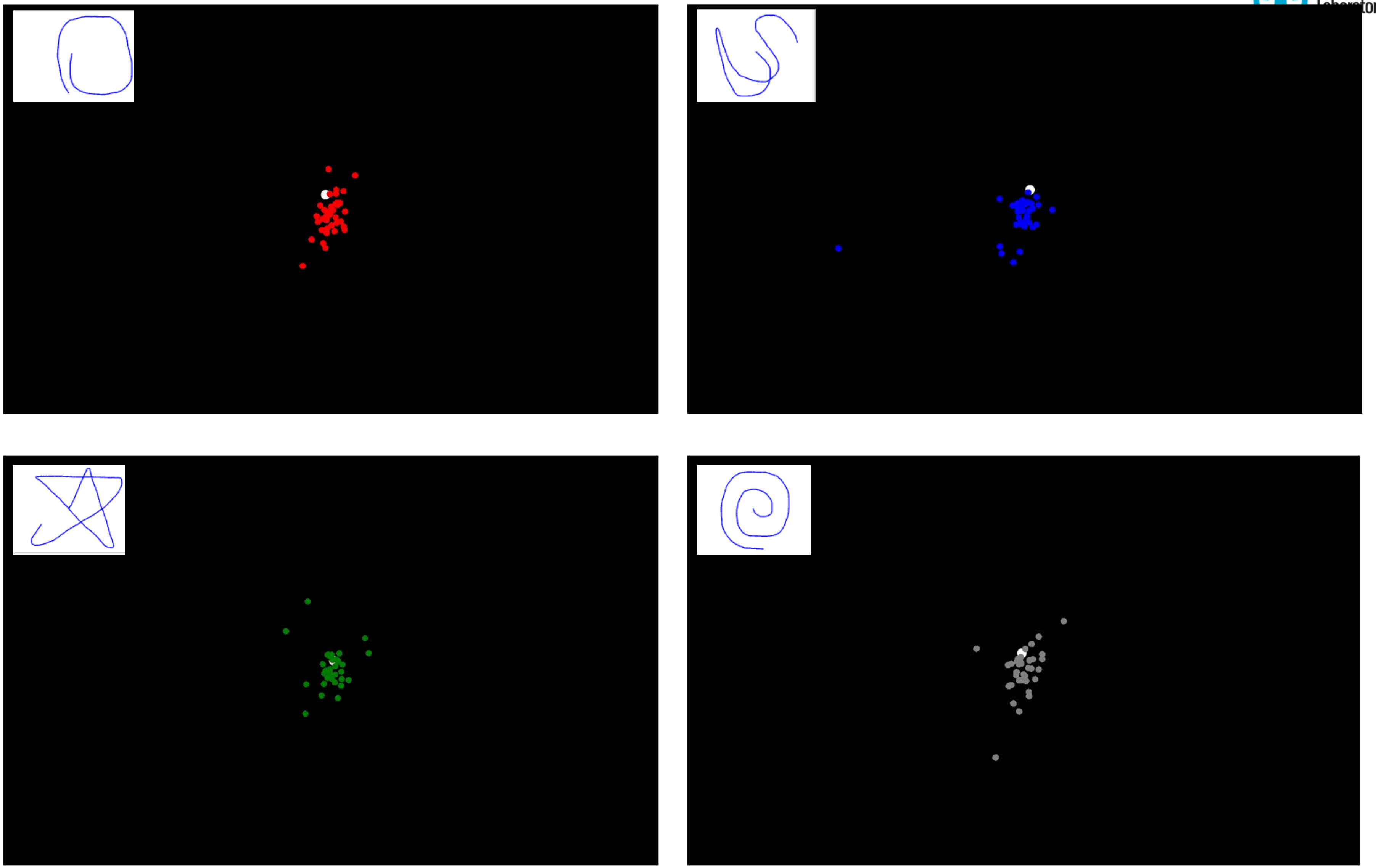
Flight Trajectories



Chicago area air traffic, April 4, 2013
Tracktable courtesy of Danny Rintoul and Andy Wilson, Sandia National Laboratory

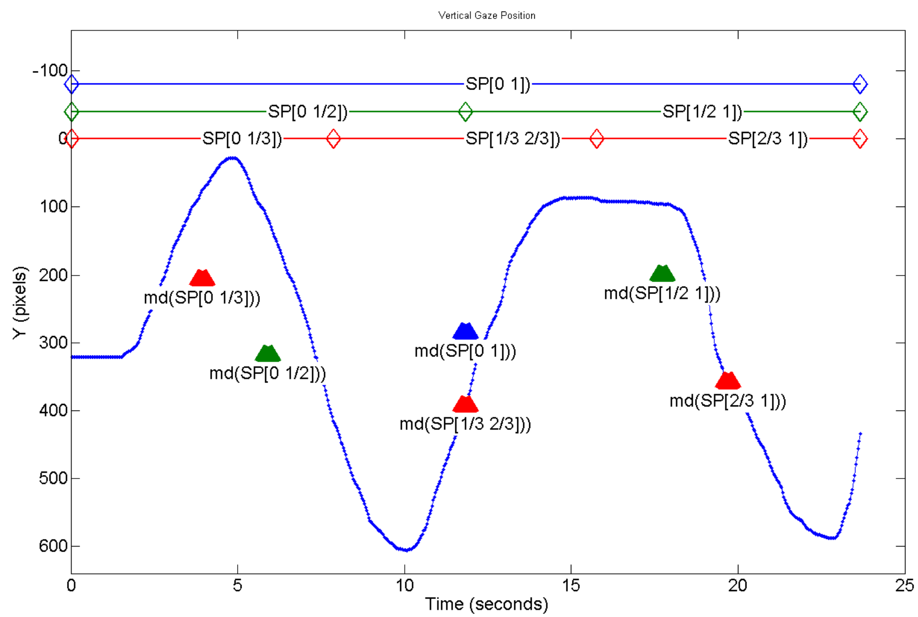
GazeAppraise

- Performs unsupervised cluster analysis on spatiotemporal sequences
- Requires zero-to-minimal preprocessing
- Does not require a priori specification of areas of interest
- Inspired and adapted from Tracktable [Rintoul et al., 2015], an application to cluster flight trajectories



GazeAppraise calculates geometric features at multiple scales

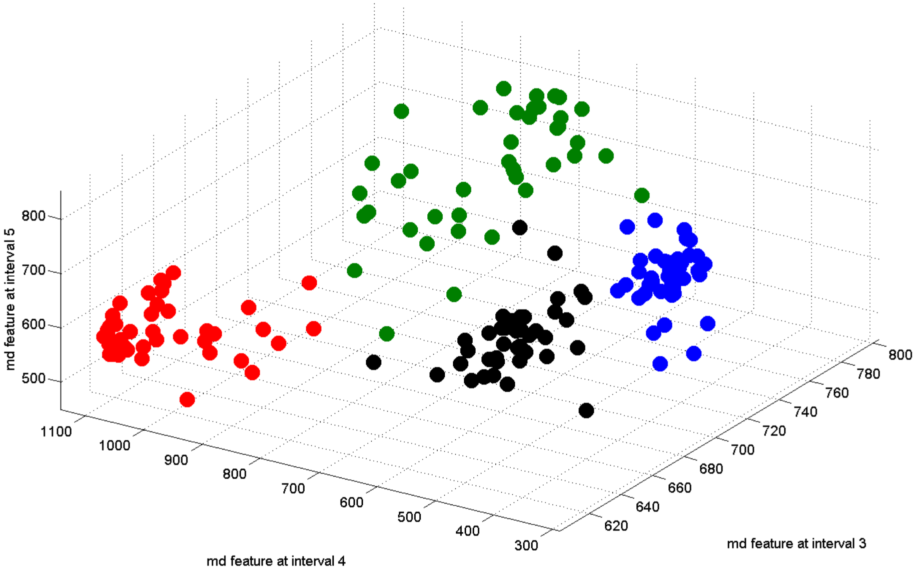
1	0-1 (whole scanpath)
2	0-.5
3	.5-1
4	0-.33
5	.33-.66
6	.66-.99
7	0-.25
8	.25-.5
9	.5-.75
10	.75-1



feature metrics can use any quantity calculable from samples in each interval

Trajectory Clustering

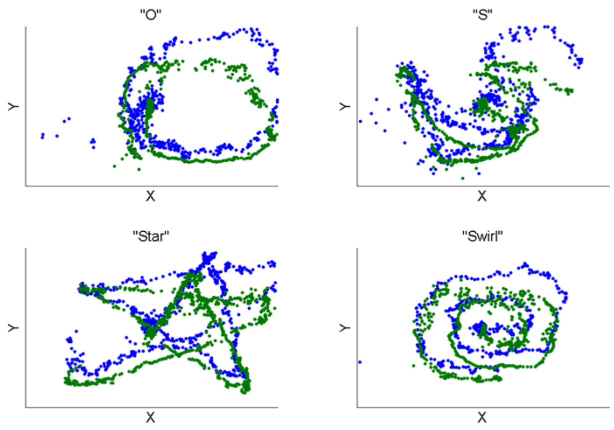
- GazeAppraise uses a density based clustering algorithm (DBSCAN) which does not require a priori knowledge of number of clusters
  - minPts: minimum # of members to form a cluster
  - Eps: neighborhood radius

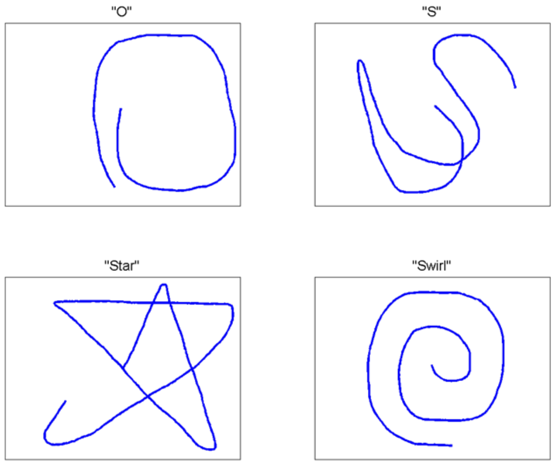


Results

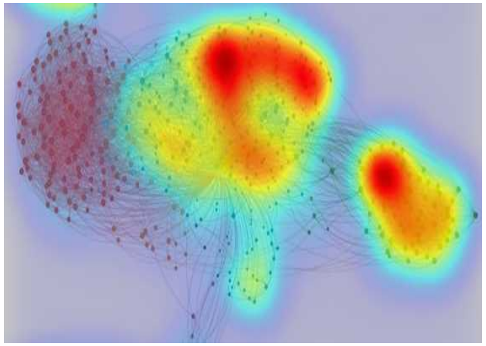
	Stimuli			
Cluster	O	S	Star	Swirl
1	40			
2	0	41		
3	0	0	40	
4	0	0	0	41
Outlier	1	0	1	0

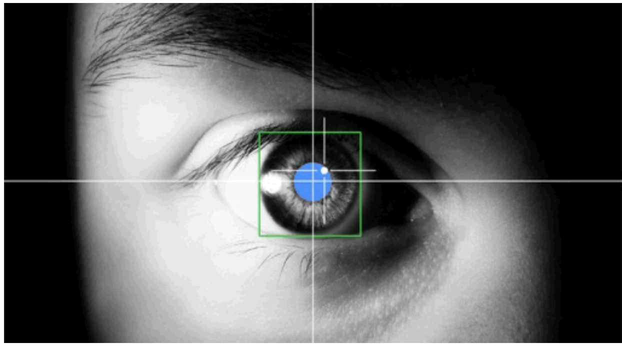
98.8% recall/sensitivity and 100% precision





Exceptional service in the national interest





Modeling Human Comprehension of Data Visualizations

Laura Matzen, Michael Haass, Andy Wilson
University of Illinois, at Urbana-Champaign
Georgia Tech

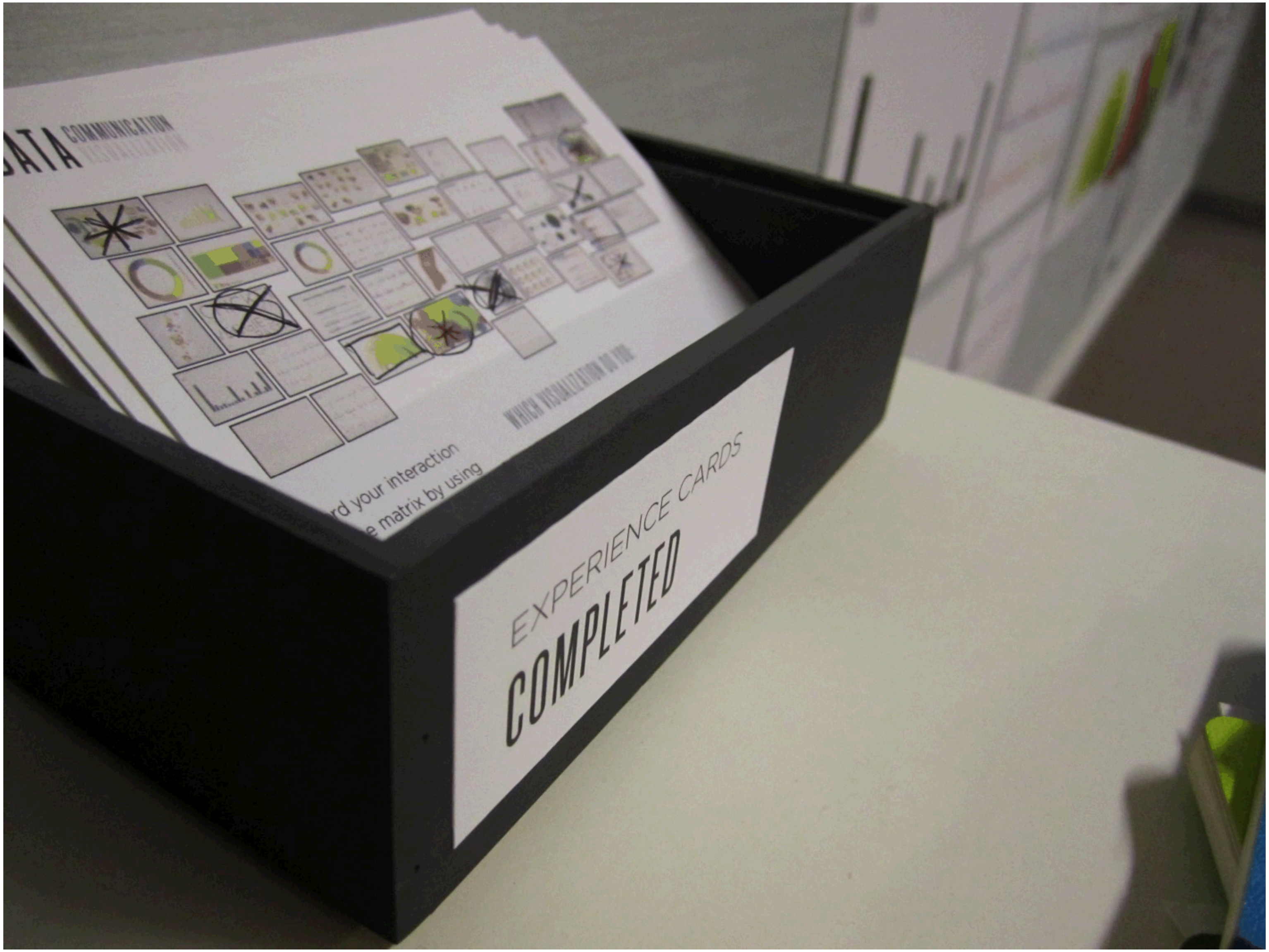
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Background

- Analysts often rely on data visualizations when making high-consequence decisions, but little is known about how to evaluate a visualization's effectiveness (value) for an end user

- The field of visual analytics is calling for the creation of models of human-computer cognitive processing that can address this gap and advance our understanding of how humans reason about data visualizations. Sandia is uniquely positioned to advance the state of the art in this area.

State of the art…



Goals

- The proposed research has three main goals:
  - **Develop models for assessing the bottom-up visual saliency of data visualizations**
    - Such models exist for images of natural scenes, but not for data visualizations.
    - Outcome will be open-source code that will create saliency maps for visualizations
  - **Conduct cognitive neuroscience experiments to characterize the top-down sensemaking strategies employed by users of visualizations**
    - Studies using eye tracking to map how analysts navigate abstract information spaces
    - The two outcomes:
      - Publications describing patterns of visual processing that are associated with sensemaking in vis
      - An empirical framework based on visual cognition to aid visualization designers in supporting their users' cognitive processing needs.
- Develop methods and metrics to establish value of a given data visualization
  - We want to collaborate with Georgia Tech HCI on this objective!

Scope

Past research:

- Saliency models that predict where viewers will look in natural scenes
- Models of top-down visual processing for scene-like visualizations
- One-off evaluations of specific visualizations
- Studies of navigation through simple visualizations (e.g. bar graphs)

Abstract representations of data

Visualizations designed by one person/group but analyzed and interpreted by someone else

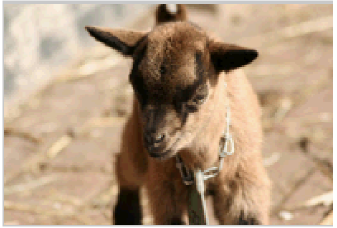
Visualizations with properties that require expanding the dimensionality of scene-based models

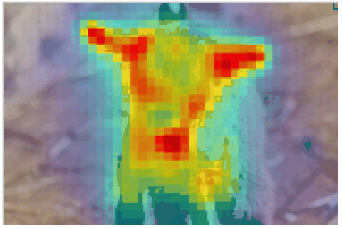
Out of scope for proposed project:

- Visualizations based on physical objects/scenes
- Dynamic visualizations*

Goal 1 - Develop models for assessing the bottom-up visual saliency of data visualizations

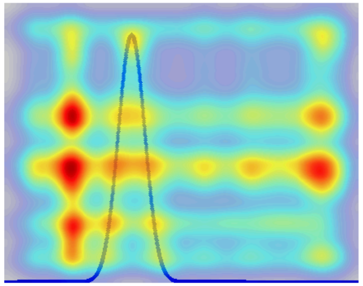
- Models of bottom-up saliency exist for natural scenes and can predict where people will look
  - A tool that could do the same for visualizations would be extremely useful for evaluations!



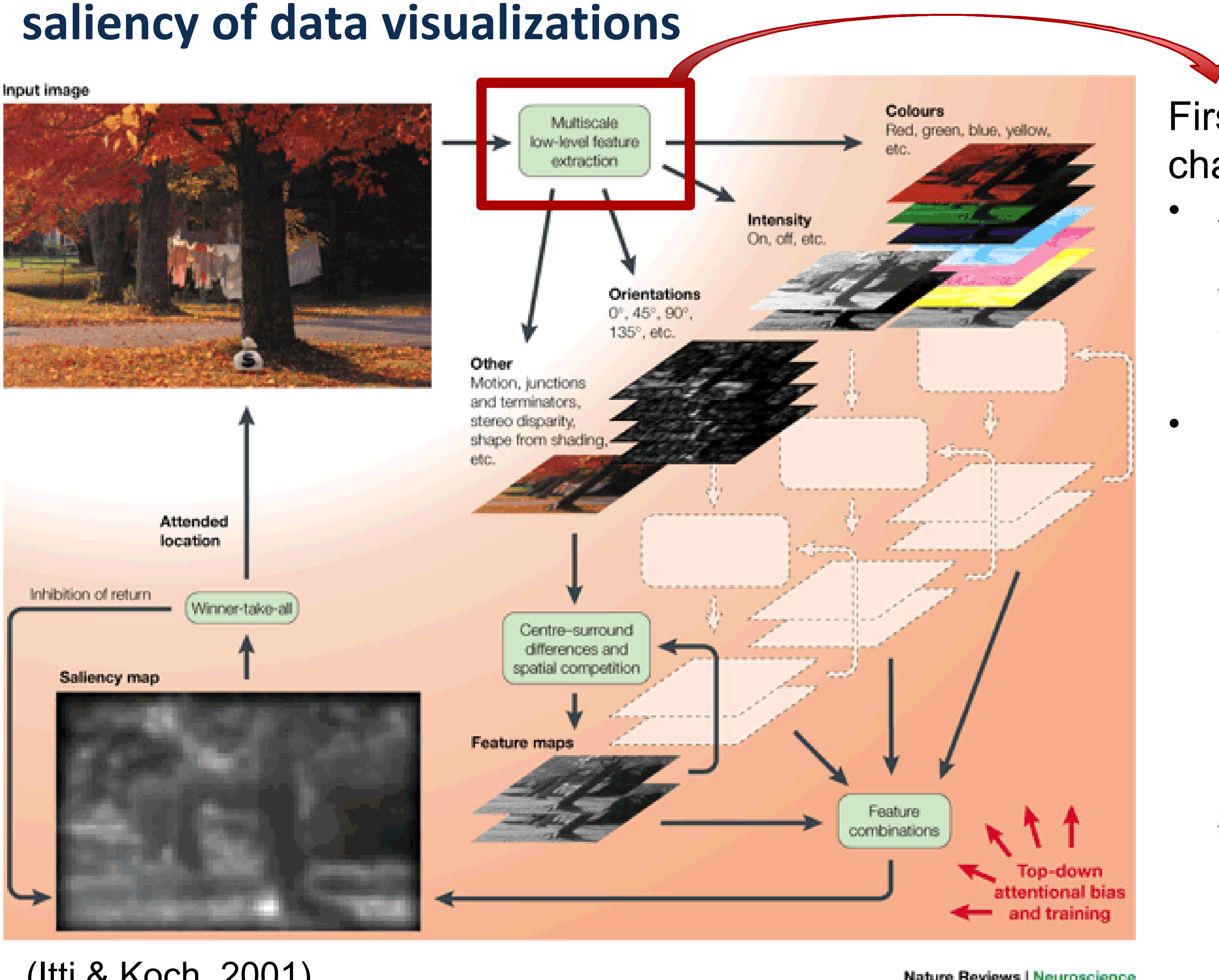


(Itti & Koch, 2001)

- These models fail for abstract visualizations
  - Inappropriate spatial scales and weighting
    - Visualization have features that are very small relative to the extent of the image
  - Inadequate feature sets
    - Features used for natural scenes (orientation, intensity, color) don't capture key contrasts used in data visualizations



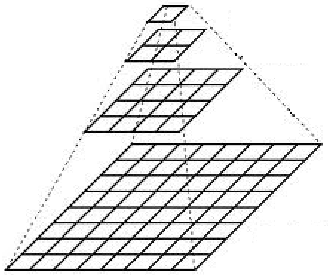
Goal 1 - Develop models for assessing the bottom-up visual saliency of data visualizations



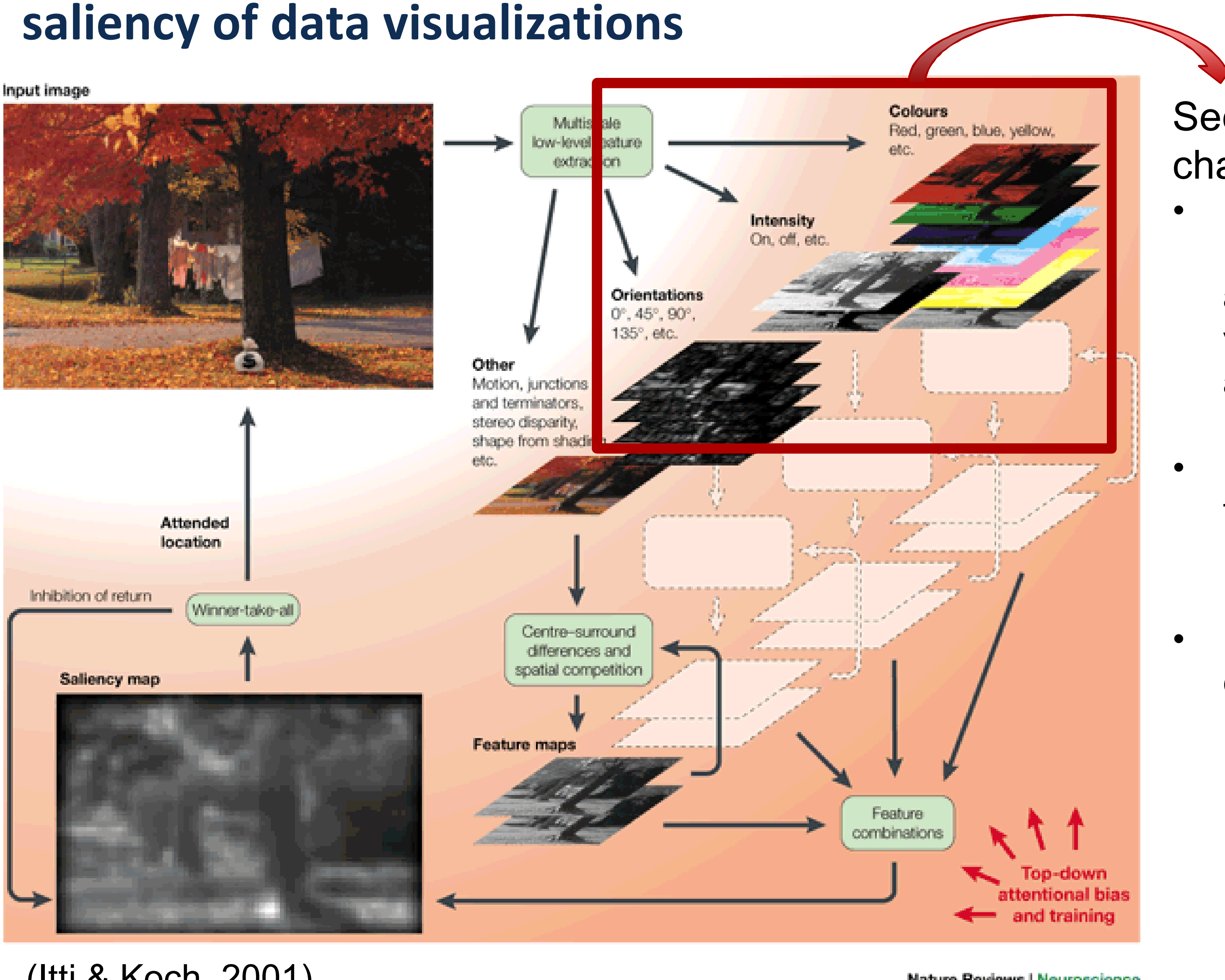
(Itti & Koch, 2001)

First technical challenge:

- Are there different choices of scale that will work for abstract vis?
- If not, a new spatial sampling approach is needed



Goal 1 - Develop models for assessing the bottom-up visual saliency of data visualizations



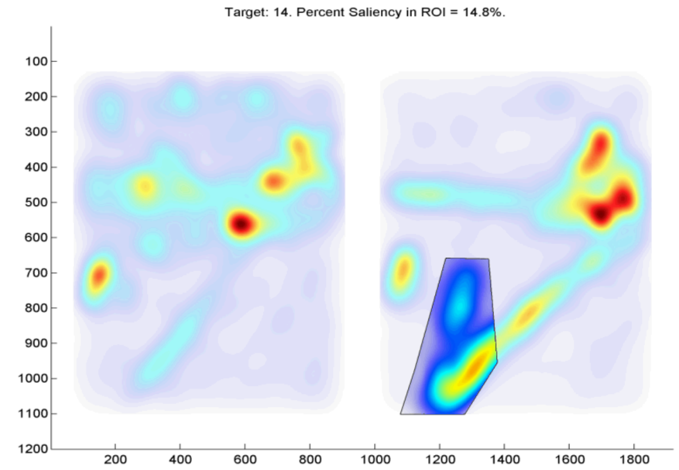
(Itti & Koch, 2001)

Second technical challenge:

- New feature sets must be developed to adequately capture visual properties of abstract vis
- Must be realistic in terms of neural processing
- Must be structured to capture the visual language of vis

Goal 2 - Characterize the top-down sensemaking strategies employed by users of visualizations

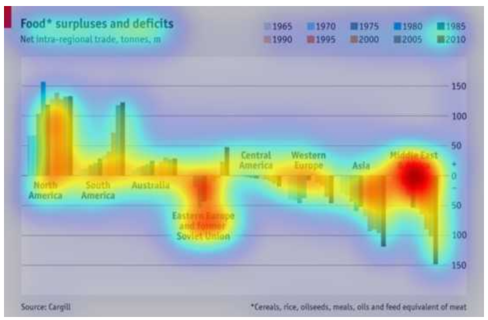
- **How do people navigate through abstract information spaces when making decisions?**
  - What information do they use and in what order?



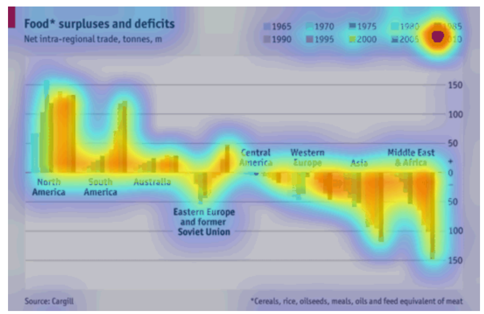
- New eye tracking methods allow us to map viewer's interactions with information and their cognitive biases
  - Comparisons of viewer gaze patterns to saliency maps can be used to model top-down visual cognitive strategies
  - Provide new methods for evaluating the effectiveness of visualizations
    - Support for data synthesis, sensemaking, communication of uncertainty, formation of insights

Currently testing “best in class*” saliency models on standard vis set to quantify performance and identify missing features of models

Itti & Koch



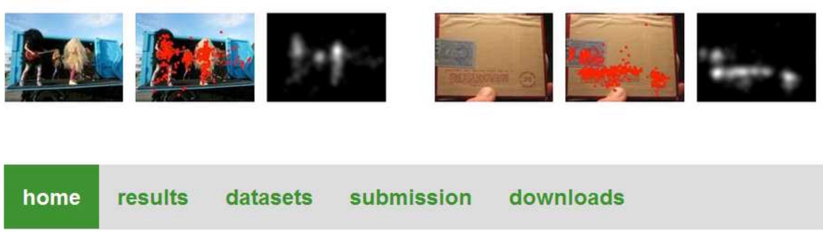
Binary Maps

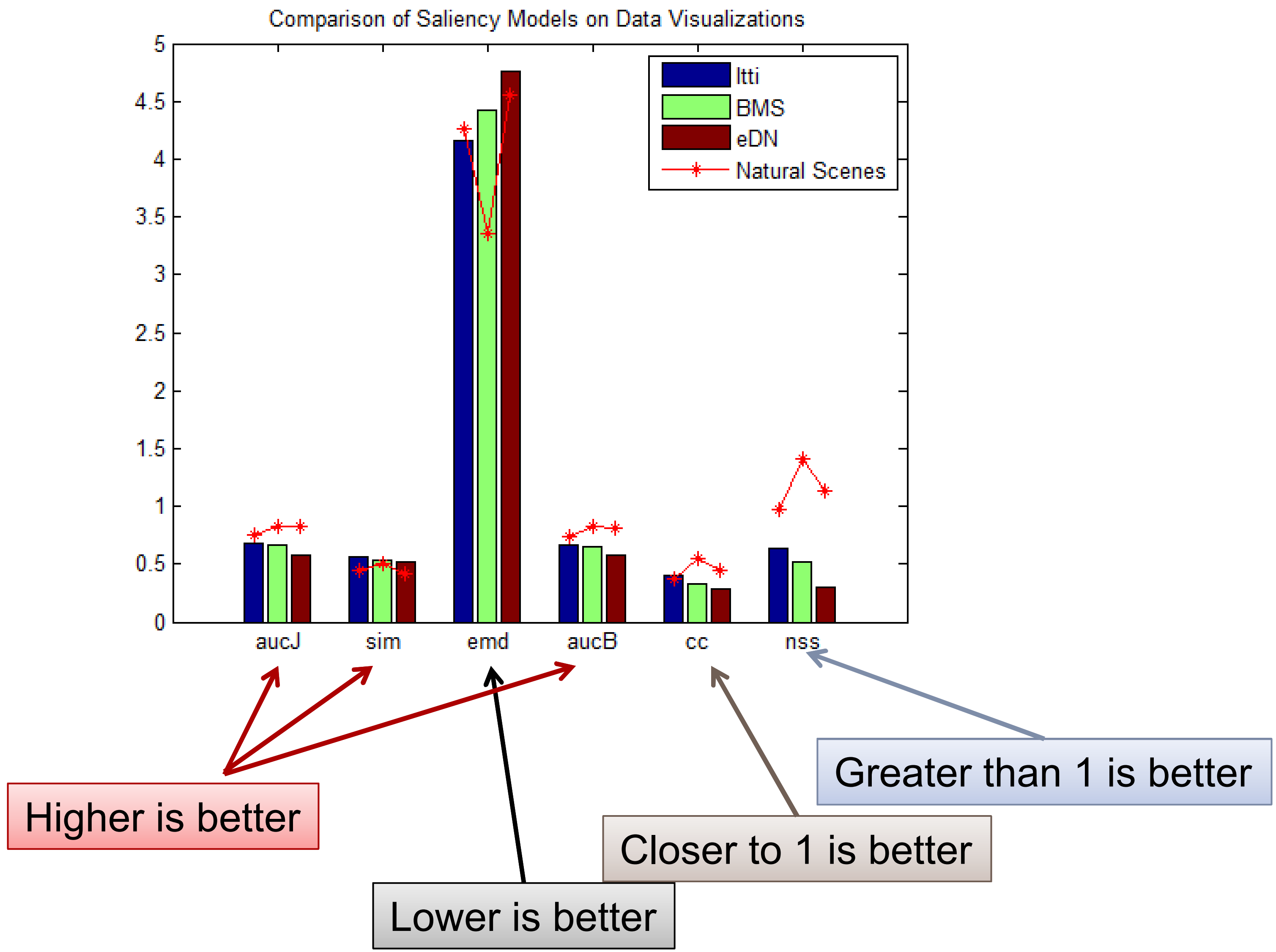


Ensembles of Deep Networks

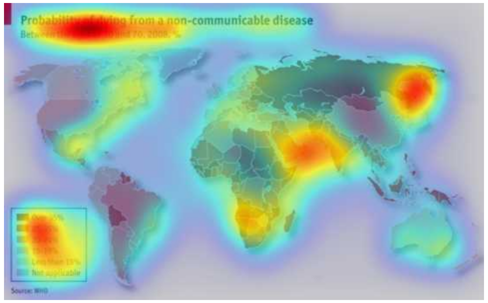
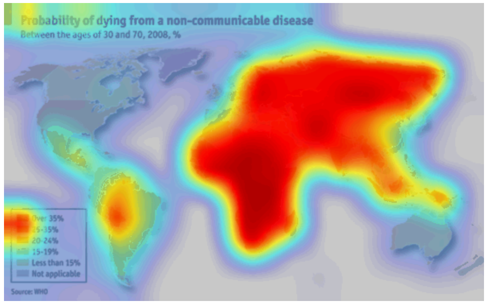
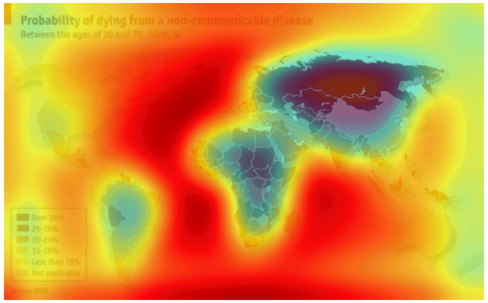
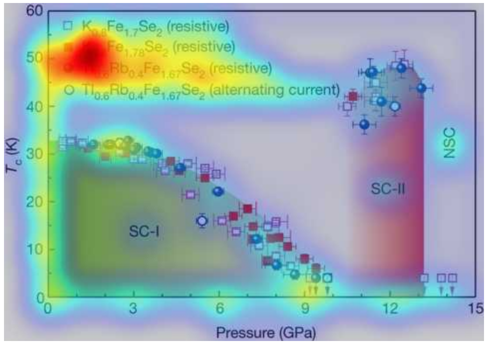
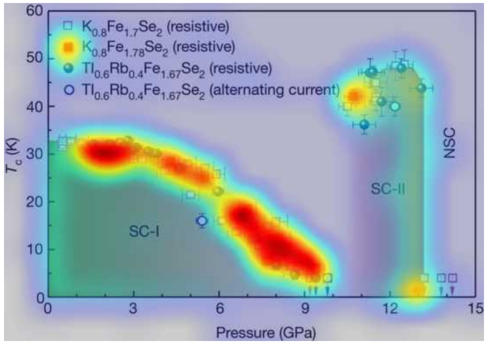
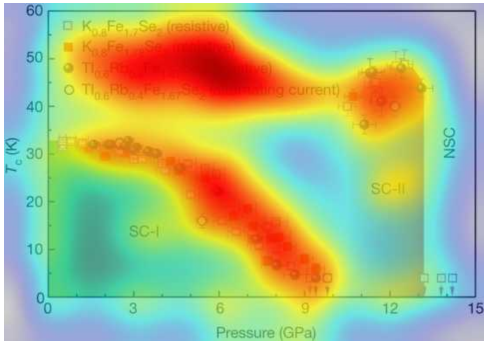
184 selected “info graphics”

***mit saliency benchmark**




Comparison of Saliency Models on Data Visualizations
Itti
BMS
eDN
Natural Scenes
aucJ
sim
emd
aucB
cc
nss
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Higher is better
Lower is better
Closer to 1 is better
Greater than 1 is better

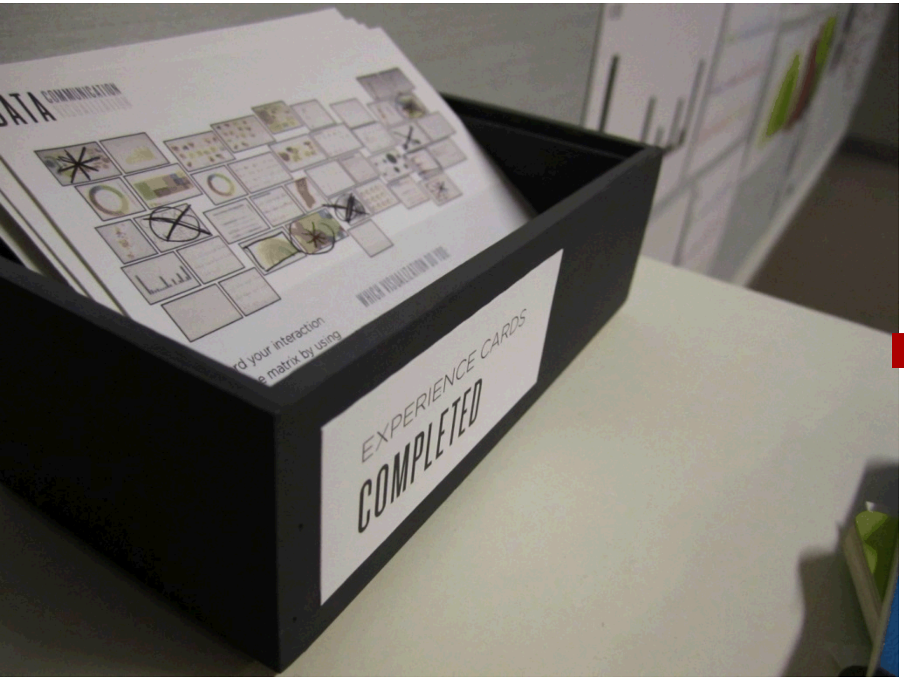
Some Additional Examples

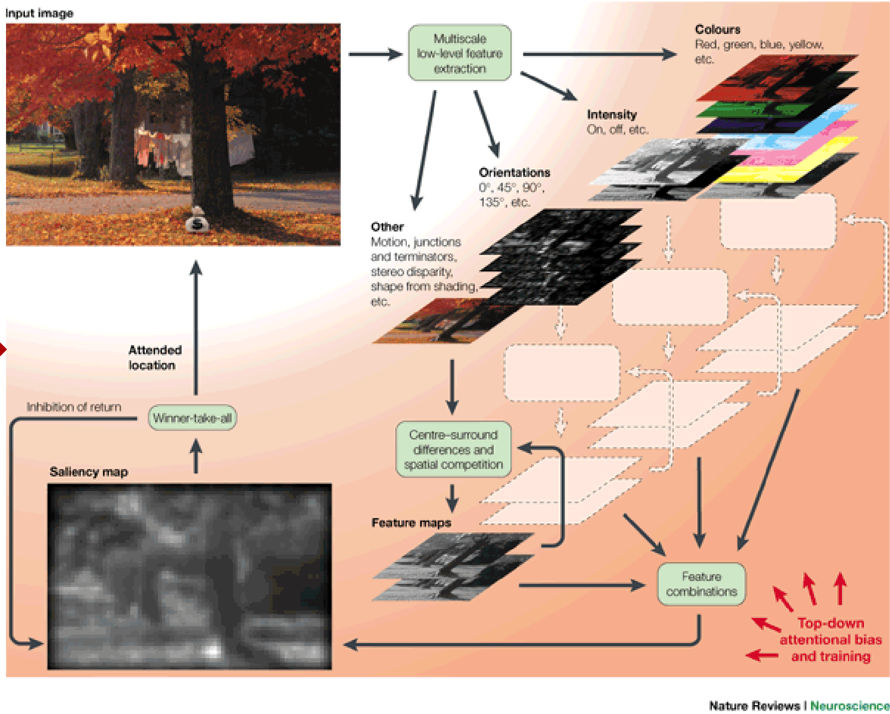
Itti & Koch	Binary Maps	Ensembles of Deep Networks
		
		

Next Steps

- Collect eye tracking data for “less curated” data visualizations @ UIUC
- Analysis & refine saliency models
- Collect additional eye tracking and top-down data
- Analysis & develop top-down guidelines for visualization designers

In a nutshell…





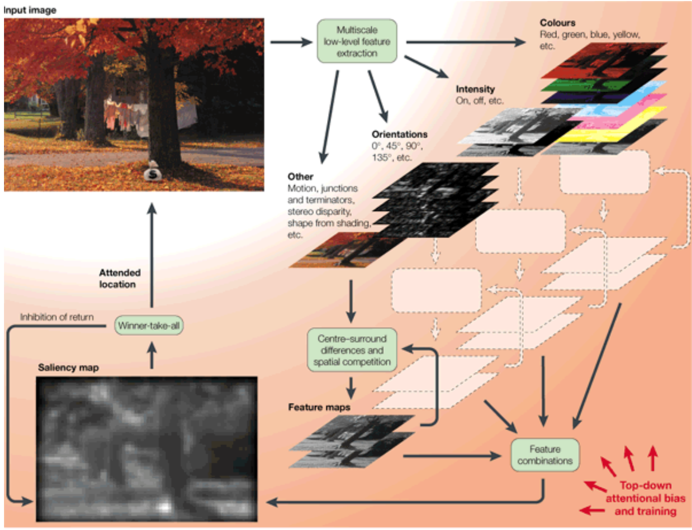
THANK YOU!

mjhaass@sandia.gov

BACKUPS

Saliency Estimation for Advanced Imaging Scenes Using Pixel Statistics*

- Large body of work on estimating visual saliency of natural scene imagery
- "Standard" models readily available for downloading
- Some efforts to continue making improvements
  - Large-scale Scene Understanding Challenge http://lsun.cs.princeton.edu/
  - MIT Saliency Benchmark http://saliency.mit.edu/index.html



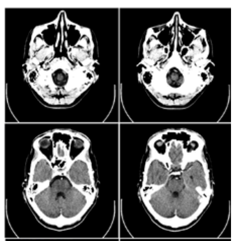
Itti & Koch, 2001 Nature Reviews | Neuroscience

*European Conference on Eye Movements, August-2015

Advanced Imaging Sensors

- But many of today's advanced sensors produce image products with novel visual characteristics

X-ray



- Shadowing
- Orientation

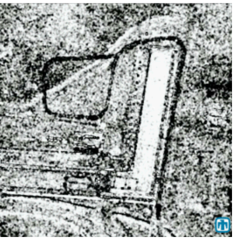
IR & Thermal

- Saturation
- Resolution
- False color

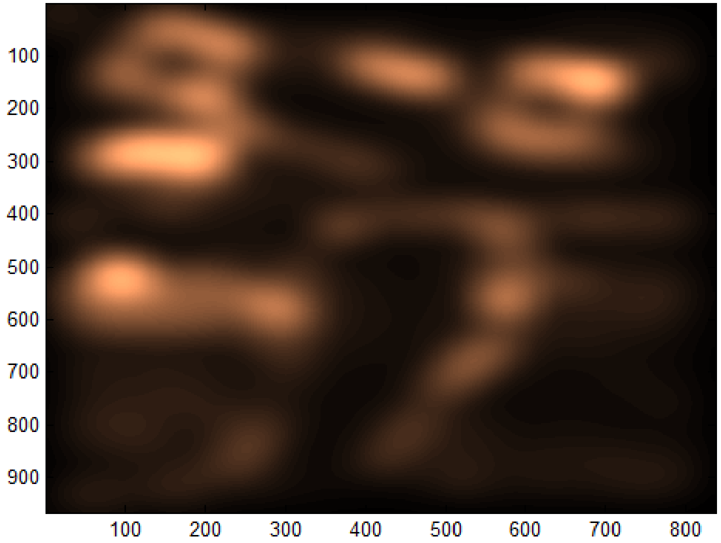
Radar

- Layover
- Shadowing
- Noise

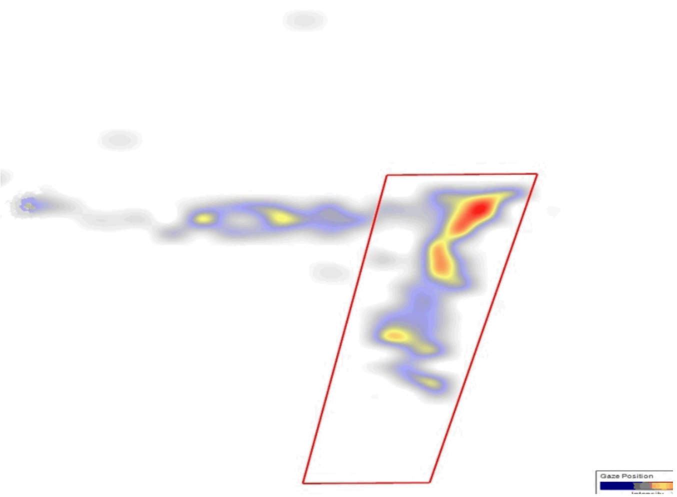
SAR Example – Saliency vs. Actual Gaze



Salience Map

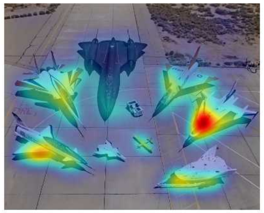


Gaze Map

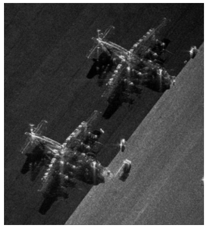


*European Conference on Eye Movements, August-2015

Challenge/Problem



- Data from advanced sensor systems are ultimately interpreted by human analysts - traditional saliency models will have some applicability
  - Information still encoded and displayed using standard visualization parameters such as contrast and color



- Developing technologies will continue to provide challenging imagery
  - *"While dual-energy imaging is now a reality in medical practice,* ***multienergy*** *is still in its early stage, but a promising research activity."*[1]

[1]Pacella, D., Reports in Medical Imaging, Vol. 8, 2015
Middle right SAR Image courtesy of Sandia National Laboratories, Airborne ISR

Study Overview

- How well does existing model (Itti & Koch) predict saliency in synthetic aperture radar (SAR) imagery?
- How can standard saliency estimation be improved to better predict gaze patterns of sensor-knowledgeable viewers?

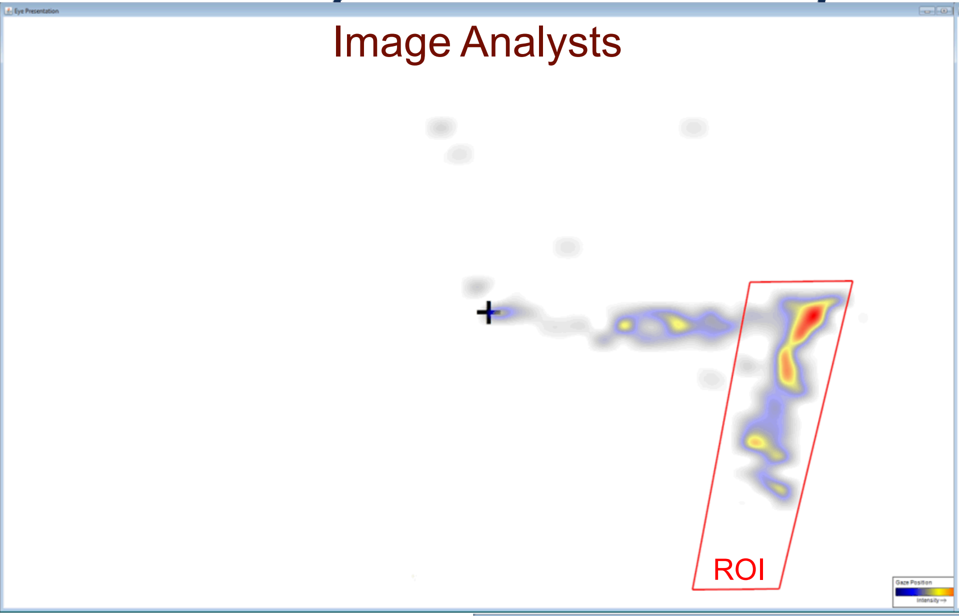
- Study task - change detection in SAR imagery
- Participants
  - 3 with no SAR experience ("novices")
  - 6 radar engineers familiar with SAR ("engineers")
  - 3 professional SAR imagery analysts ("experienced IAs")

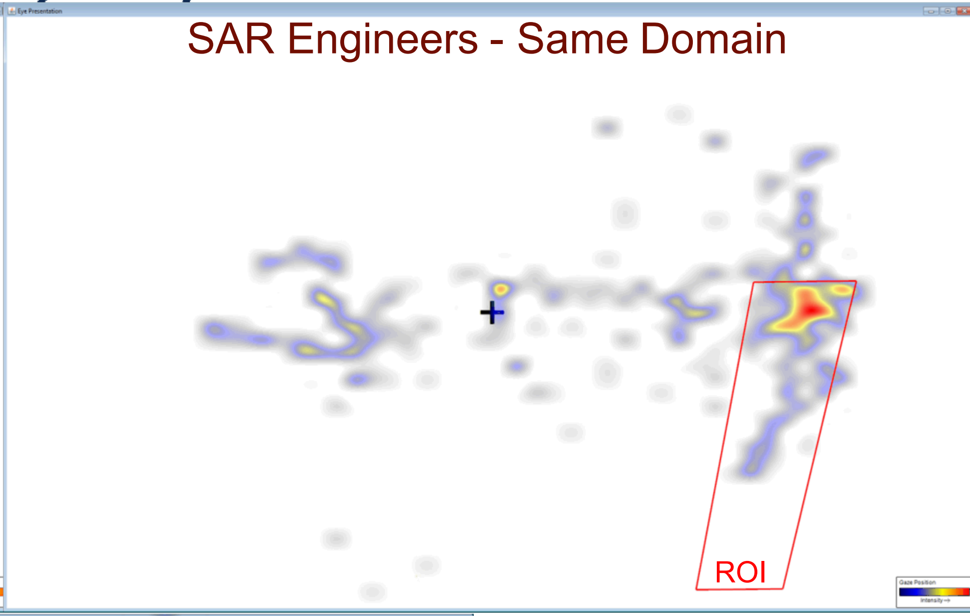
Saliency Comparison Metrics*

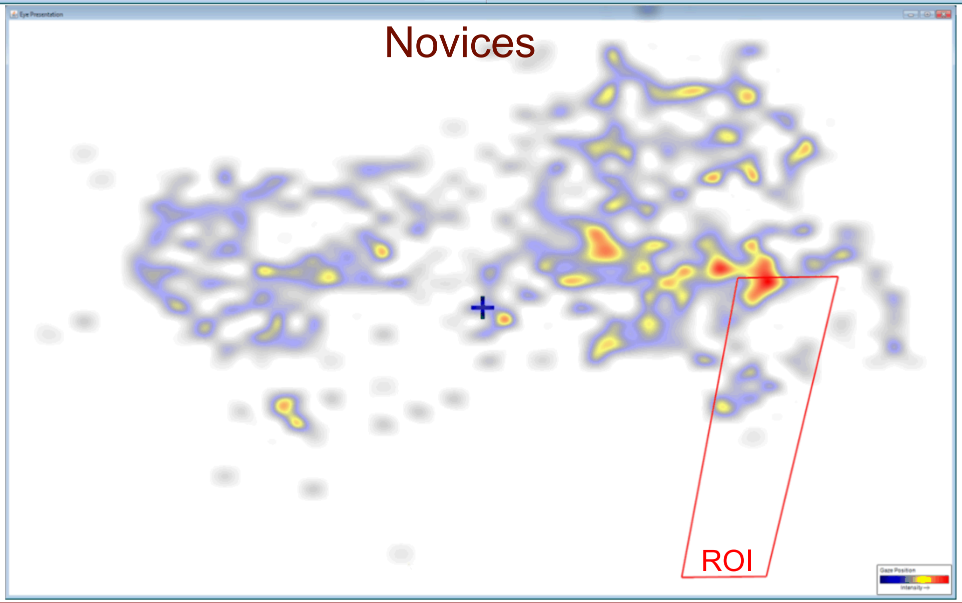
1. Linear Correlation Coefficient (CC)
   - Measure of the strength of a linear relationship between fixation map (G) and saliency map (S)
     - $CC(G,S) = \frac{cov(G,S)}{\sigma_G \sigma_S}$ When CC is close to ± 1, there is almost a perfectly linear relationship
2. Normalized Scanpath Saliency (NSS)
   - Average of saliency values at human gaze positions (saliency normalized to have zero mean and unit standard deviation)
     - NSS = 1 indicates that the subjects' gaze positions fall in a region whose predicted saliency is one standard deviation above average
     - When NSS ≥ 1, the saliency map exhibits significantly higher saliency values at human gaze locations compared to other locations
     - NSS ≤ 0 indicates the saliency model performs no better than picking a random position
3. Area Under Curve (AUC)
     - Human gaze positions are considered positive set, other points are negative set
     - Saliency map is treated as binary classifier to separate positive and negative sets

*Borji, A., et al. (2013). "Quantitative Analysis of Human-Model Agreement in Visual Saliency Modeling: A Comparative Study." IEEE Transactions on Image Processing **22(1): 55-69.**

Example Gaze Maps By Expertise





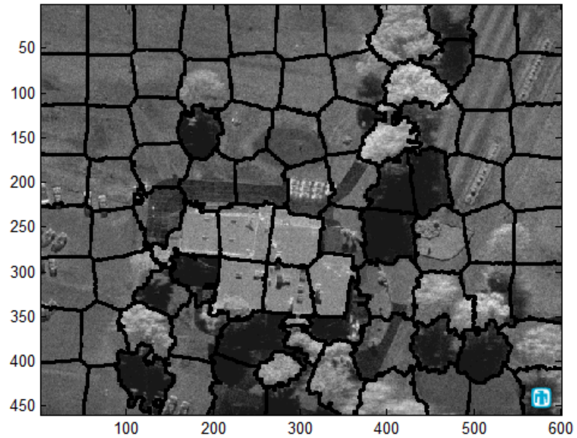


*European Conference on Eye Movements, August-2015

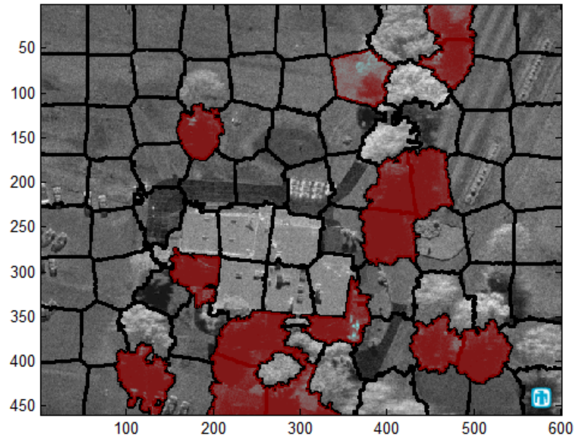
Reducing Salience Estimates in Shadow Regions

- Pixel-statistical methods used to segment[1] the scene and characterize the segment properties[2]
- These properties can serve as filters to modulate traditional saliency estimates
  - SAR Phenomenology - shadow regions have low coherence

Segment



Classify

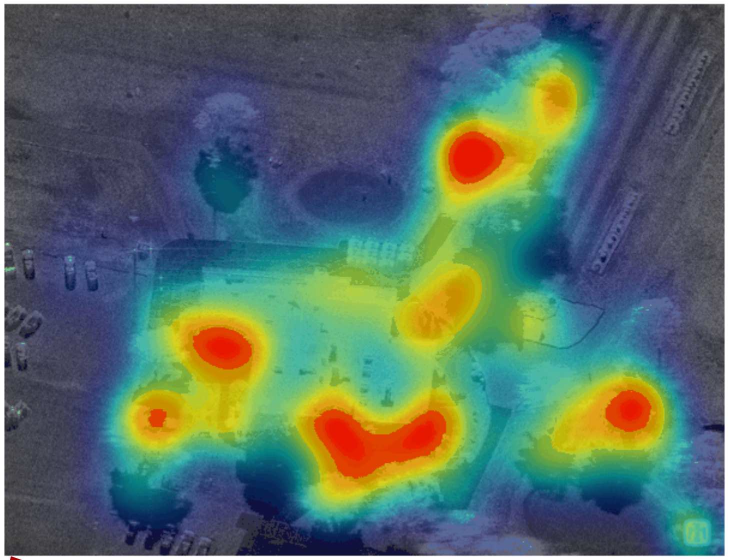


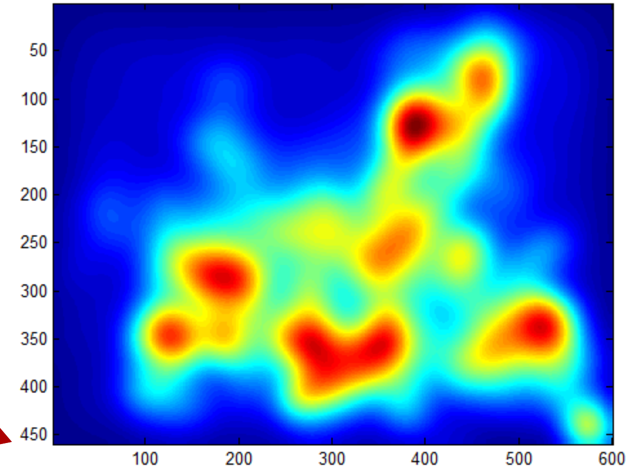
[1] M. M. Moya, et al, "Superpixel segmentation using multiple SAR image products" RADAR SENSOR TECHNOLOGY XVIII, Proceedings of SPIE VOL 9077, Conference on Radar Sensor Technology XVIII, MAY 05-07, 2014, Baltimore, MD

[2] M.M. Moya, et al., "Superpixel Classification for Signature Search in Synthetic Aperture Radar Imagery," Conference on Data Analysis (CoDA), March, 2014, Santa Fe, NM.

SAR Image courtesy of Sandia National Laboratories, Airborne ISR

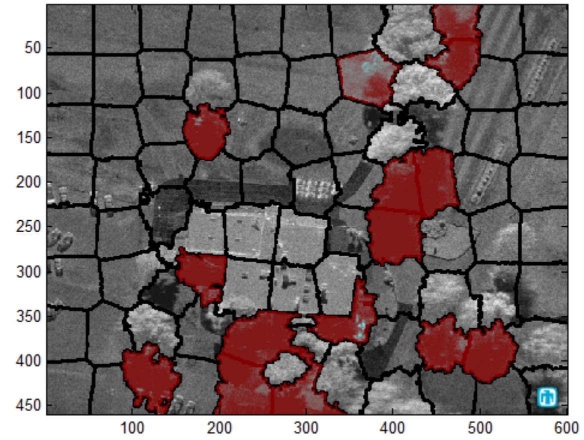
Method (1): Natural Scene Saliency Map



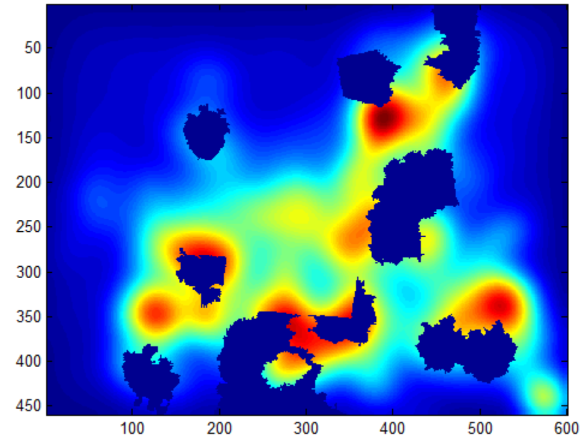


Method (2): Select and Filter Based on Superpixel Characteristics

- Select superpixels with certain characteristics (i.e. shadows)
  - Classify using pixel statistics within each superpixel

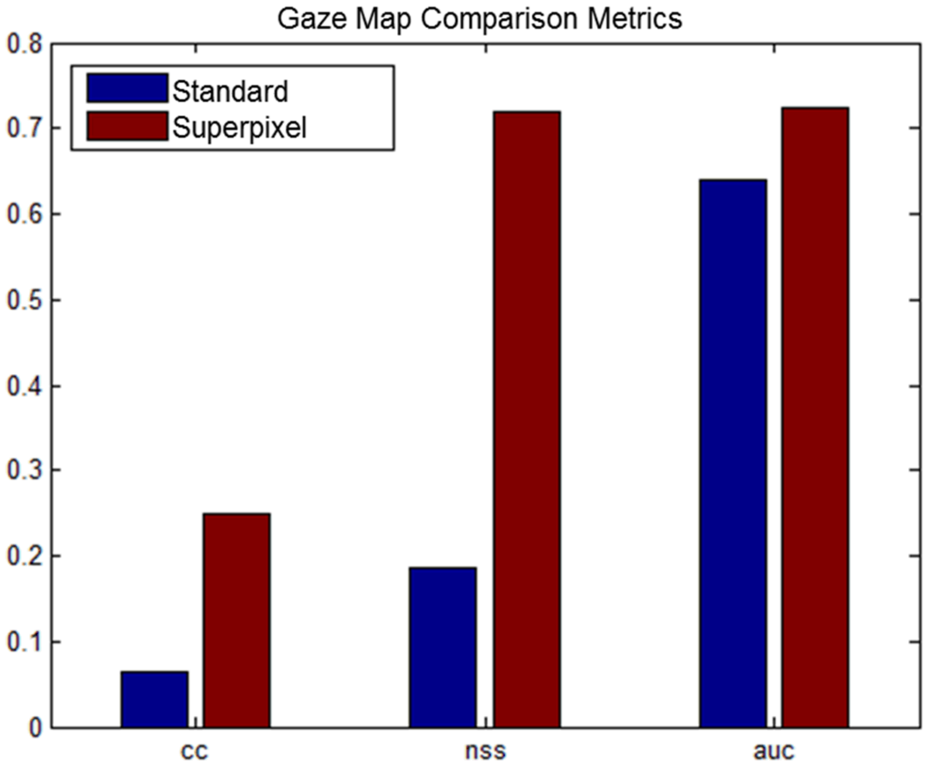


- Apply mask to original saliency map
  - Can add Gaussian, or other smoothing to reduce discontinuities



SAR image courtesy of Sandia National Laboratories, Airborne ISR

Study Results



Saliency map modulated by superpixel characteristics is more similar to analyst fixation maps

- Linear correlation (cc) improvement factor is 3.8X
- Normalized scan path saliency (nss) improvement factor is 3.9X
- Area under receiver-operator curve (auc) improvement factor is 1.1X

*European Conference on Eye Movements, August-2015

Conclusion

- Modulating standard model using superpixel segmentation and classification based on sensor phenomenology can improve salience – gaze agreement
- Using eye tracking technology to explore relationships between traditional saliency models and pixel-statistical properties we can understand eye movements of domain experts interacting with imagery from today's most advanced sensors