Task Decoding Using Recurrent Quantification Analysis of Eye Movements

Daniel C. LaCombe, Jr.[1] & Elan Barenholtz[12]

1: Department of Psychology; 2: Center for Complex Systems and Brain Sciences

Introduction

- ❑ An age-old adage is that the eyes are the window to the soul; how much truth to this is there?
- ❑ Despite our experience of a richly-detailed visual environment, the most useful information we gain during visual perception is at the center of our gaze – an area called the fovea
- ❑ Because this area is very small, we must move our eyes around very quickly in order to sample from the environment
- ❑ Research suggests that eye movements contain a wealth of information about:
  - Who the observer is (biometrics; Rigas, Economou, & Fotopoulos, 2012)
  - What an observer is viewing (Borji, Lennartz, & Pomplun, 2015)
  - The interests and goals of a viewer (Borji & Itti, 2014)
- ❑ The present research aims to decode the task someone is engaged in from their eye movements alone using Recurrence Quantification Analysis in particular, which captures spatiotemporal information

Methods

- ❑ Participants:
  - 50 FAU undergraduate students
- ❑ Design & Procedure:
  - All participants engaged in both 20 visual search (VS) and 10 aesthetic judgment (AJ) tasks for 8 sec each
  - Only target-absent trials for the visual-search task were used (which constituted half of them), resulting in 10 trials total
  - All stimuli were used for both visual-search and aesthetic-judgment tasks
- ❑ Stimuli & Apparatus:
  - Natural images of indoor scenes (e.g., bedroom)
  - Targets for visual search were small arrows superimposed on the images (present for half of images as stated previously)
  - Eye movements were measured using a Tobii eye tracker
- ❑ Feature Extraction:
  - *Aggregate* (number of fixations, average fixation duration, average saccadic amplitude, area covered assuming 1° fovea)
  - *Fixation Map* (smoothed fixation map)
  - *RQA* (recurrence, determinism, laminarity, CORM)
  - *RQA2* (RQA + regression, LORM, size)
  - *RQAprob* (RQA using r_{prob}); *RQAprob2* (RQA2 using r_{prob}).
- ❑ Pattern Classification:
  - Each feature set trained a corresponding support vector machine with a linear kernel using WEKA (Hall et al., 2009)

Example Image with Target Shown in Blue Box

Aesthetic Judgment Instructions

Using the scale below, rate how pleasant the image was by pressing the corresponding number on the keyboard

[1] Very pleasant
[2] Somewhat pleasant
[3] Not sure
[4] Somewhat unpleasant
[5] Very unpleasant

Visual Search Instructions

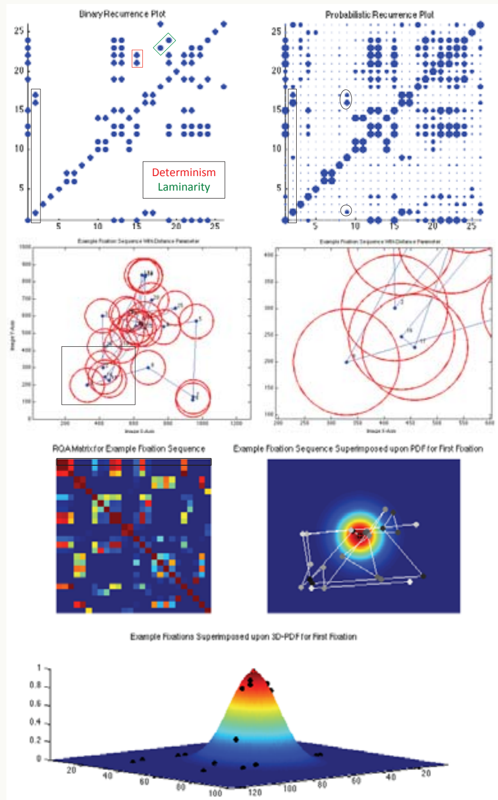
Which direction was the arrow pointing?

[left] – press the left arrow key
[right] – press the right arrow key
[up] – press the up arrow key
[down] – press the down arrow key
[no arrow] – press "n"

Recurrence Quantification Analysis

- ❑ Recurrence Quantification Analysis (RQA; Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013) is a means of deriving simple descriptive statistics from a recurrence plot that holds both spatial and temporal information

$$r_{ij} = \begin{cases} 1 & \text{if } d(f_i, f_j) \le \rho \\ 0 & \text{otherwise} \end{cases} \qquad r_{(prob)ij} = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-))$$



- ❑ From these recurrence plots, we then calculate:
  - Recurrence: the proportion of all fixation pairs that are recurrent with one another; measures how much the eyes go back to areas previously visited
  - Determinism: the proportion of all recurrent points that create diagonal lines; measures the degree to which spatiotemporal sequences were recapitulated
  - Laminarity: the proportion of all recurrent points that create diagonal lines; laminarity measures the degree to which the eyes stay in a particular area
  - Center of Recurrene Mass (CORM): measures the average temporal separation between all recurrent fixations

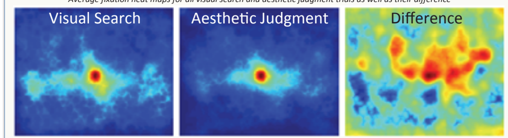
$$REC = 100 \times \frac{2\left(\sum_{j=2}^{N}\sum_{i=1}^{j-1} r_{ij}\right)}{(N \times N - 1)} \qquad DET = 100 \times \left(\frac{\sum_{j=2}^{N-1}\sum_{i=1}^{j-1}(r_{ij})(r_{i+1,j+1})}{\sum_{j=2}^{N}\sum_{i=1}^{j-1} r_{ij}}\right) \qquad LAM = 100 \times \left(\frac{\sum_{j=2}^{N-1}\sum_{i=1}^{j-1}(r_{ij})(r_{i+1,j})}{\sum_{j=2}^{N}\sum_{i=1}^{j-1} r_{ij}}\right) \qquad CORM = 100 \times \left(\frac{\sum_{j=2}^{N-1}\sum_{i=1}^{j-1}(r_{ij})(j-i)}{(N-1)\left(\sum_{j=2}^{N}\sum_{i=1}^{j-1} r_{ij}\right)}\right)$$

- ❑ Three other measures, regression, latency of recurrence mass (LORM), and size, were calculated for RQA2 and RQAprob2; these measures are novel and were developed based on their intuitive advantage in this decoding paradigm (equations not shown)

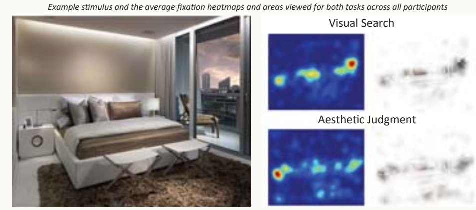
Results

- ❑ Data Cleaning Procedure: Participants' whose performance on the visual search task did not exceed chance levels (eight) or for whom the eye tracker was not confident it was tracking the eye for over a third of the trials (six) were removed from the analysis (14 total)

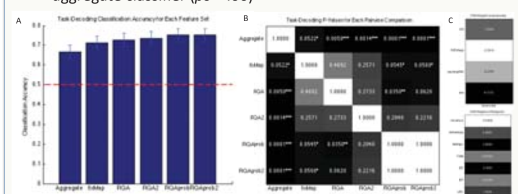
Average fixation heat maps for all visual search and aesthetic judgment trials as well as their difference



Example stimulus and the average fixation heatmaps and areas viewed for both tasks across all participants



- ❑ To evaluate the decoding performance of each classifier, we used an *n*-fold cross validation procedure
- ❑ Task-decoding performance of all classifiers were significantly above chance levels (chance = 50%; Binomial test; *ps* < .001; see panel A)
- ❑ We also compared the performance of each of the six classifiers to one another using paired-sample *t*-tests and Bonferonni-corrected α-values (.05/6 = .0083); *p*-values of all comparisons are in the contingency table in panel B and are summarized as follows:
  - All RQA classifiers performed significantly better than the aggregate measures (*ps* < .0083)
  - The two probabilistic RQA classifiers performed marginally better than the fixation map features (*ps* < .06)
  - RQAprob performed marginally better than RQA (*p* < .05)
  - The fixation map classifier performed marginally better than the aggregate classifier (*ps* < .06)



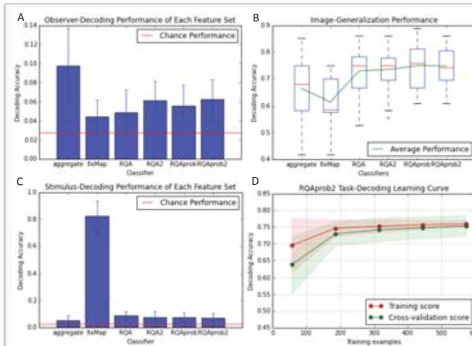
Note: The red line in panel A represents chance level at 50% in the task-decoding procedure. In panel B, ****p* <.0083, ***p* < .05, **p* < .06.

- ❑ We can interpret the weights of the SVM as follows (see panel C):
  - Negative and positive weights are feature dimensions in which aesthetic judgment and visual search scored higher, respectively
  - The size of the weight reflects its importance relative to the other features in the set
- ❑ Investigating these weights and their corresponding histograms reveals that some features were more informative than others
  - Aggregate measures:
    - Number of Fixations: VS > AJ
    - Mean Fixation Duration: VS < AJ
    - Mean Saccadic Amplitude: VS ≈ AJ
    - Area: VS > AJ

Stacked histograms for the normalized features in the aggregate feature set; red is visual search and blue is aesthetic judgment

  - RQA(prob)2 measures (roughly equal to that of RQA(prob)):
    - Recurrence: VS < AJ
    - Determinism: VS > AJ
    - Laminarity: VS > AJ
    - CORM: VS < AJ
    - Size: VS > AJ
    - Regression: VS < AJ
    - LORM: VS > AJ

Stacked histograms for the normalized features in the RQAprob2 feature set; red is visual search and blue is aesthetic judgment



- ❑ Because RQA measures were superior to the others in task-decoding we were also interested in how they performed in observer- and stimulus-decoding
  - These results revealed that aggregate measures are best for observer-decoding (panel A) and fixation maps were best for stimulus-decoding (panel C)
- ❑ Given that RQA retains topological information but looses all coordinate information, we were also interested in whether RQA would generalize well across images
  - This analysis revealed that RQA and aggregate measures generalized well whereas fixation map features performed worse when it was not trained on novel images (panel B)
- ❑ Analyses also revealed that further training data would likely not improve decoding performance for RQA feature sets (panel D)

Discussion

- ❑ In the present study, we asked whether we could decode an observers' task from eye movements alone using different features – in particular RQA – to train separate classifiers
- ❑ Our findings are consistent with those of Borji and Itti (2014) given that we can decode an observers' task significantly better than chance levels with all feature types
- ❑ The results suggest that RQA is better suited to task-decoding than aggregate or spatial features alone likely because it contains spatio-temporal information
- ❑ RQA also appears to generalize well to images that it is not trained on because it is invariant to spatial structure
- ❑ However, aggregate features excel at observer-identity decoding and spatial features are apt at stimulus decoding
- ❑ RQA's high performance on task-decoding and lower performance on observer- and stimulus-decoding suggests that it is relatively observer- and stimulus-invariant compared to the other features used here

Future Research

- ❑ Future research should explore alternative features and learning algorithms to improve decoding performance so that ubiquitous task-inference via computers and non-invasive eye-tracking can emerge
- ❑ We also intend to use optimization techniques (e.g., particle swarm optimization) in order to better choose parameters that were somewhat arbitrarily chosen in this experiment (i.e., distance and variance parameters)
- ❑ Automatic intention inference can enhance the way that we interact with computers!

References

Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavioral Research Methods*, *45*, 842-856.

Borji, A., & Itti, L. (2014). Defending Yarbus: Eye movements reveal observers' task. *Journal of Vision*, *14*, 1-22.

Borji, A., Lennartz, A., & Pomplun, M. (2015). What do eyes reveal about the mind? Algorithmic inference of search targets from fixations. *Neurocomputing, 149*, 788-799.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10-18.

Rigas, I., Economou, G., & Fotopoulos, S. (2012). Biometric identification based on the eye movements and graph matching techniques. *Pattern Recognition*, *33*, 768-792.