Visual Search as Active Sequential Hypothesis Testing

Rajesh Sundaresan

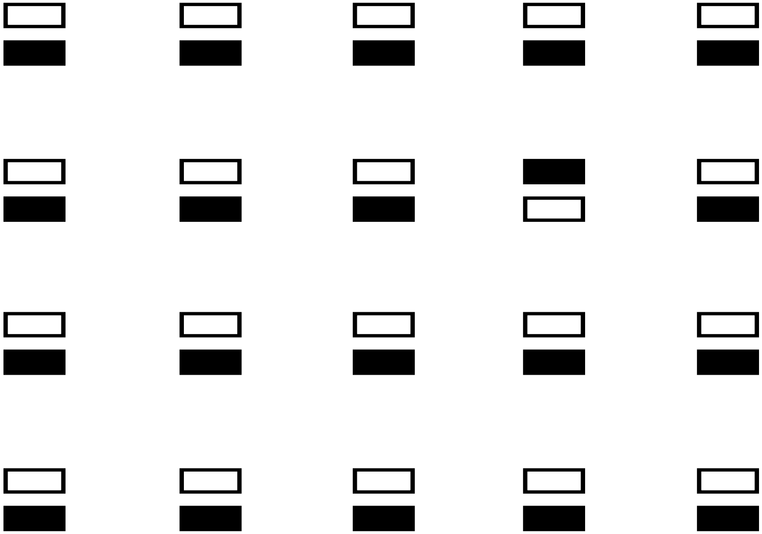
Indian Institute of Science

09 September 2014

Based on ongoing work with Nidhin Vaidhiyan and Arun Sripati

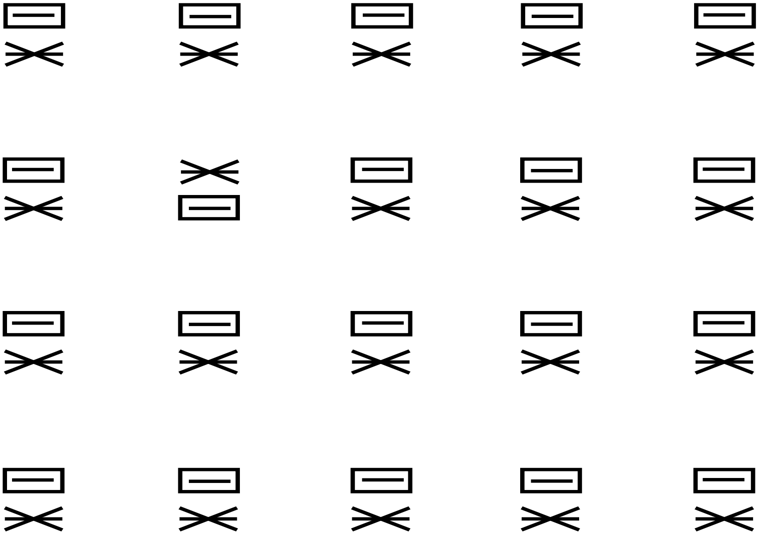
Find the odd image - 1

Find the odd image - 1



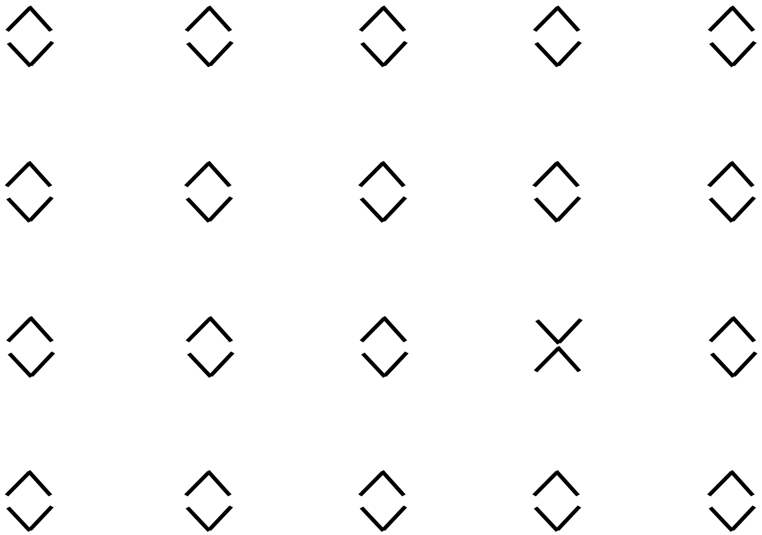
Find the odd image - 2

Find the odd image - 2



Find the odd image - 3

Find the odd image - 3



The hypothesis of Sripati and Olson 2010

Hypothesis

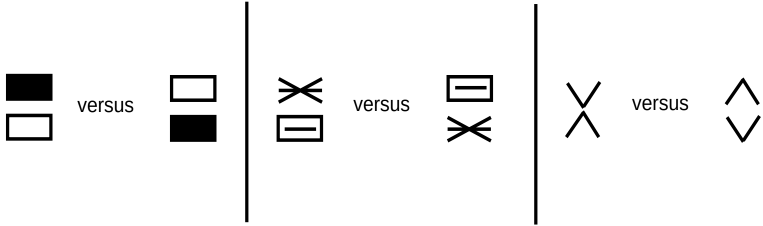
Visual search performance depends on the similarity of the two images. Similarity is measured in terms of neuronal activities in the visual cortex in response to each of the two images.

The hypothesis of Sripati and Olson 2010

Hypothesis

Visual search performance depends on the similarity of the two images. Similarity is measured in terms of neuronal activities in the visual cortex in response to each of the two images.

The time to identify the odd one depends on the similarity of the neuronal responses to

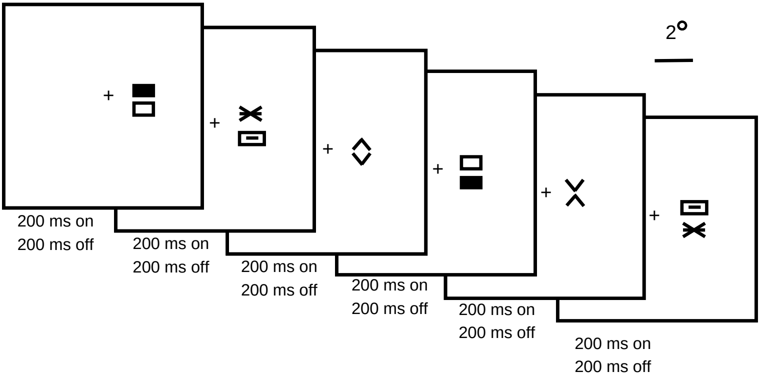


Experimental procedure on rhesus macaques and recording (done by Sripati and Olson)

- Cleared by CMU institutional animal care and use committee
- Two macaques were surgically fitted with:
  - a cranial implant for neuronal activity recording;
  - a scleral search coil for recording eye movements.
- Data was collected over several days. Before each day's experiment, an electrode was inserted so that the tip was 1 cm above the inferotemporal cortex.
- The electrodes were pushed, reproducably, along tracks forming a square grid with 1 mm spacing.
- Neuronal activity was recorded. Individual neurons' action potentials then isolated using a commercially available tool (Plexon).

The experiment on macaques

- The macaques were trained to fixate on the + while a series of stimuli appeared one after another.



- Images were randomly interleaved. Neuronal activity recorded as these images were shown over several 2 second rounds.
- A drop of juice after each round of 2 seconds.

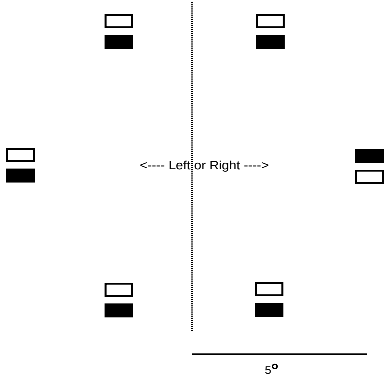
The neuronal data

- Firing rates of $N = 174$ neurons in response to these six images
- Data collected in a similar manner for a total of 24 images

- For each image i, the neuronal response is summarized by the firing rate vector $(\lambda^i(n), 1 \leq n \leq N)$.
- If i and j are pairs, then $d(i,j) := ||\lambda^i - \lambda^j||_1$.

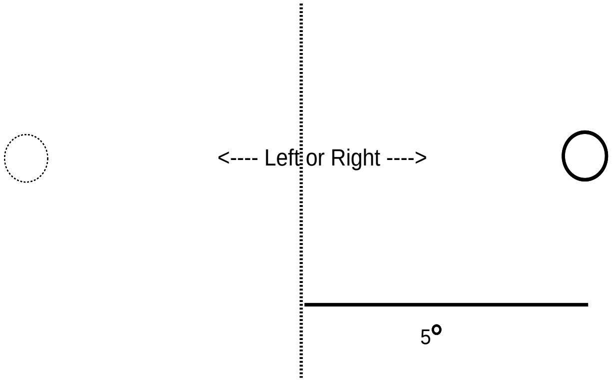
A reaction time study on humans

- Study conducted on six subjects (approved by CMU's review board)
- Identify the location of the oddball and hit a key to tell left or right



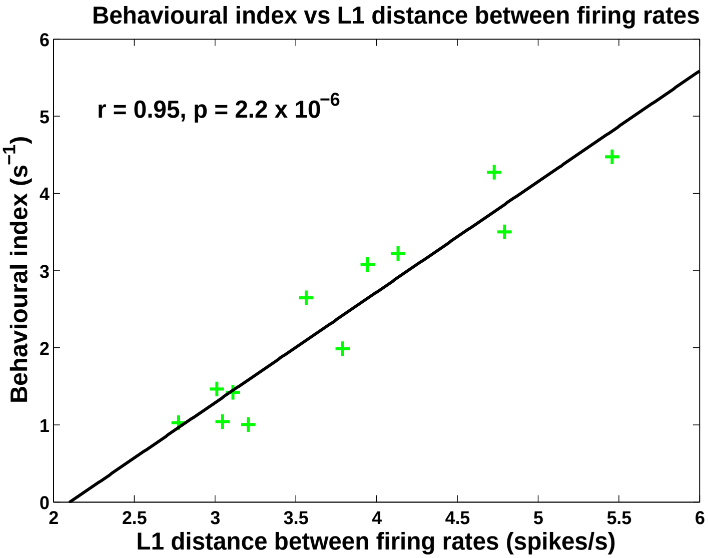
- Image displayed until reaction (which if correct, valid trial), or until 5 seconds (aborted)
  RT(i, j) = average reaction time
  Data averaged over both *oddball i among distractors j and vice versa*

Baseline reaction time



- RT_b = baseline reaction time
- $s(i,j) = \mathrm{RT}(i,j) - \mathrm{RT}_b$

Their punch line



Questions

- IT cortex response in macaques and behavioural experiments on humans.
  Single image response on macaques, multiple images in humans.

- Ad hoc subtraction of crudely estimated baseline reaction time.

- Why L^1?

- Why $1/s$ versus the neuronal metric?

- Is there a model that explains this behaviour?
  Sripati and Olson proposed a model that accumulates confidence, and stops when a threshold is reached.

A better model grounded in a theory

- What would the prefrontal cortex do if it got observations from the IT cortex and could control the eye?

A better model grounded in a theory

- What would the prefrontal cortex do if it got observations from the IT cortex and could control the eye?

- Active sequential hypothesis testing ingredients:

  - Hypothesis (ℓ, i): The oddball location is ℓ and its type i among distractors j.

  - Divide time into slots.

  - Control: Given observations and decisions in all previous slots,
    - decide to stop and declare the oddball, or
    - decide to continue, and direct the eye to focus on location b, one of the six locations.

  - Observation: If the object in location b is k, then N Poisson point processes with rates $(\lambda^k(n), 1 \leq n \leq N)$.

  - Performance: For any hypothesis, policy should satisfy $\Pr\{\text{Error}|\text{Hypothesis}\} \leq \varepsilon$.
    What is the delay versus ε tradeoff?

A more precise question

- Let $\Pi(\varepsilon) = \{\pi : P_h^\pi(\text{decision} \neq h) \leq \varepsilon \text{ for all } h\}$.

- Given hypothesis h, let τ denote the stopping time for a decision under $\pi \in \Pi(\varepsilon)$. What is

$$\lim_{\varepsilon \to 0} \inf_{\pi \in \Pi(\varepsilon)} E_h^\pi \left[\frac{\tau}{|\log \varepsilon|} \right]?$$

From the simple to the complex

- Fixed sample size: Suppose we observe $X_1, X_2, \cdots, X_T$ iid q_0 under H_0 and iid q_1 under H_1.

- Stein's lemma: Under H_0, probability that null hypothesis is rejected diminishes exponentially fast in T.

$$\Pr\{ \text{ Reject null hypothesis } | H_0 \} \approx e^{-TD(q_1||q_0)+o(T)}.$$

If we want this to be $\leq \varepsilon$, then $T \approx \frac{|\log \varepsilon|}{D(q_1||q_0)}$.

- Stopping problem: At each time step, stop and decide, or ask for another sample.

- Wald's approximation: $\frac{\mathbb{E}_0[\tau]}{|\log \varepsilon|} \to \frac{1}{D(q_0||q_1)}$ as $\varepsilon \to 0$.

- With controls: $\frac{\mathbb{E}_0[\tau]}{|\log \varepsilon|} \to \frac{1}{\max_\mu \sum_a \mu(a) D(q_0^a||q_1^a)}$ as $\varepsilon \to 0$.

Chernoff 1959

- Let $\Pi(\varepsilon) = \{\pi : P_h^\pi(\text{decision} \neq h) \leq \varepsilon \text{ for all } h\}$.

Theorem (Chernoff 1959)

Given hypothesis h, let τ denote the stopping time for a decision under $\pi \in \Pi(\varepsilon)$. Then

$$\lim_{\varepsilon \to 0} \inf_{\pi \in \Pi(\varepsilon)} E_h^\pi \left[\frac{\tau}{|\log \varepsilon|} \right] = \frac{1}{D_h}$$

where $D_h = \max_\mu \min_{h' \neq h} \sum_a \mu(a) D(q_h^a || q_{h'}^a)$*, where* q_h^a *is ...*

- Interpretation: Let μ_h^* denote the argmax. For each hypothesis h, the best strategy after initial searching is a randomized strategy (μ_h^*) designed for the worst alternative (among $h' \neq h$).

- The average delay is the average of these. (Delay is in slots).

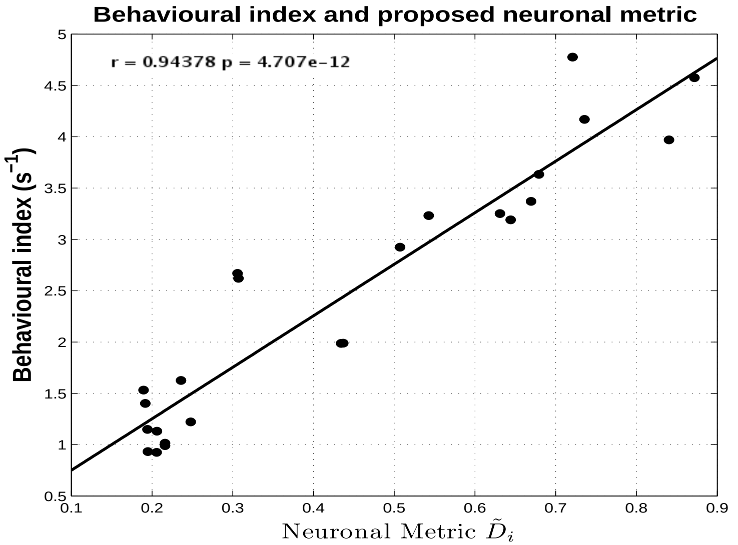
Cut to the chase

- Define $I(i,j) = \frac{1}{N}\sum_{n=1}^{N}\left(\lambda^i(n)\log\frac{\lambda^i(n)}{\lambda^j(n)} - \lambda^i(n) + \lambda^j(n)\right)$.

- Then average delay per unit time per neuron per $|\log\varepsilon|$, when the oddball image is i, is

$$s(i,j) = \frac{5I(i,j) + 3I(j,i)}{4I(i,j)I(j,i)} =: \frac{1}{\tilde{D}(i,j)}.$$

The modified punch line



How close is our model to a theory?

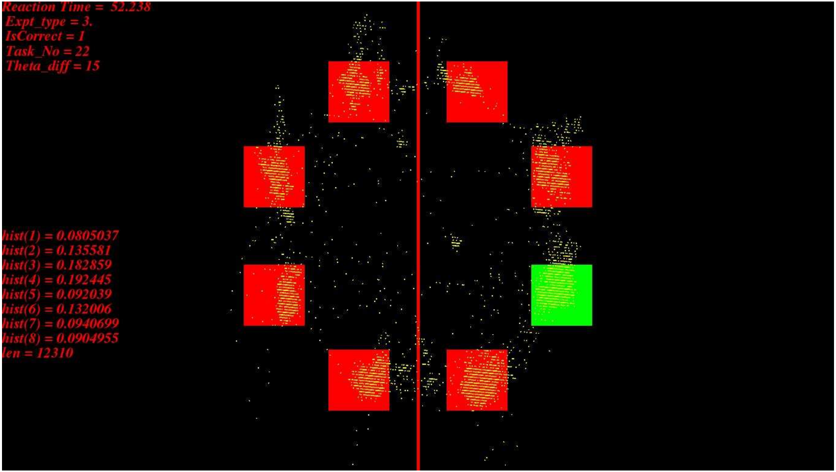
- A higher-order prediction : If hypothesis is h, then the actions are asymptotically distributed as $\mu_h^*(\cdot)$. Does it match with experiments?

- Let $h = (\ell, i)$, i.e., i be the oddball image at location ℓ. Then

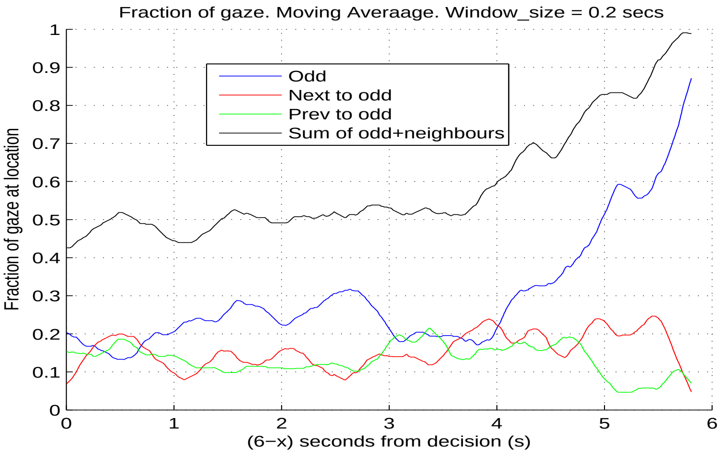
$$\begin{aligned} \mu^*_{(\ell,i)}(\ell) &= \frac{5I(j,i)}{5I(j,i)+7I(i,j)}, \\ \mu^*_{(\ell,i)}(\ell') &= \frac{I(i,j)}{5I(j,i)+7I(i,j)}, \quad \ell' \neq \ell. \end{aligned}$$

So, if $I(i,j) \approx I(j,i)$, then $5/12$ fraction of the time the oddball will be sampled, with the remaining time divided equally ($1/12$ fraction) among the distractors.

A crude experiment



A failed prediction



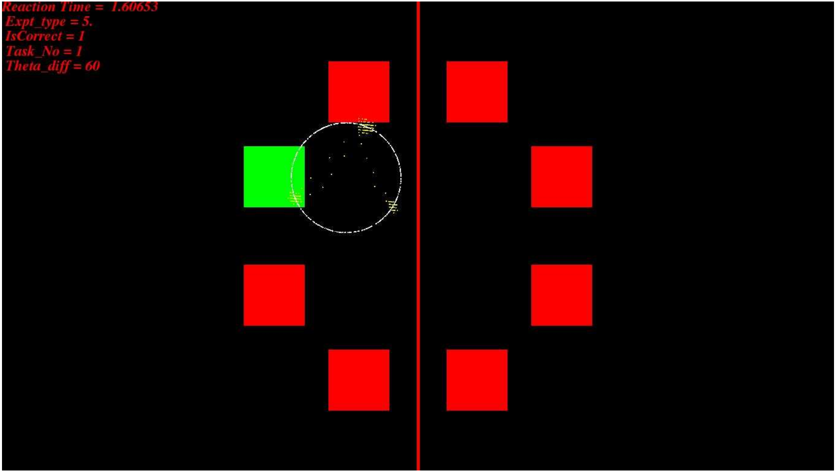
Challenges

- Cost of saccades

Challenges

- Cost of saccades
- Broad and shallow search versus narrow and deep

Broad and shallow versus narrow and deep



Challenges

- Cost of saccades
- Broad and shallow search versus narrow and deep

Challenges

- Cost of saccades
- Broad and shallow search versus narrow and deep
- Storage of beliefs in the course of the experiment
- Do not know $\lambda^i(n)$