An Active Sequential Hypothesis Testing model for Visual Search

Nidhin Koshy Vaidhiyan & Dr. Rajesh Sundaresan
Electrical Communication Engineering, IISc.
nidhin.koshy@gmail.com, rajeshs@ece.iisc.ernet.in

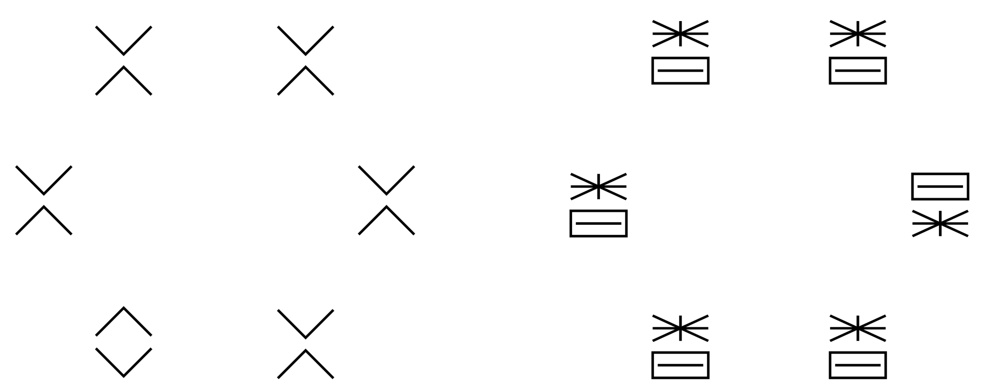
Introduction

We describe a decision-theoretic model for visual search. We first model visual search as an active sequential hypothesis testing problem. Our analysis suggests an appropriate neuronal dissimilarity index which correlates strongly with the reciprocal of search times. We will then consider a scenario where the subject has to find an oddball image, but without any prior knowledge of the oddball and distractor images. We model this scenario as one of detecting an odd Poisson point process having a rate different from the common rate of the others. The revised model suggests a new neuronal dissimilarity index. The new dissimilarity index is also strongly correlated with the behavioural data.

We thus propose a framework for connecting the perceptual distances in the neuronal and the behavioural spaces. Our framework can possibly be used to analyze the connection between the neuronal space and the behavioural space for various other behavioural tasks.

Prior Work

We invite the reader to participate in the following visual search tasks. There are two search tasks below. Find the oddball image in each of the two configurations. Based on the time taken for each of the tasks, identify which of the two is easier.



Among the two search tasks above, most subjects find the task on the left the easier, and the task on the right tougher. Visual search performance, as measured by the time taken to find the oddball image, should depend on the "similarity" of the two images. One has the natural hypothesis:

(H) The more "dissimilar" the two images, the shorter the time taken to find the oddball image.

How does one quantify the "dissimilarity"? Sripati and Olson [1] proposed the L^1 distance between the average neuronal firing rates elicited by the individual images as a measure of "dissimilarity". They found very high correlation between the inverse of the reaction times in the oddball detection task and their proposed L^1 distance metric. But why L^1? Can we come up with a model for the visual search task, which would suggest an appropriate neuronal dissimilarity index?

Components of the visual search task

- The task is a sequential hypothesis testing problem - hypotheses correspond to the odd image and the location of the odd image.
- Observation: The neuronal firing patterns generated in the brain.
- Actions: The subject has the ability to focus his attention on any location of his choice.
- Changing your focus of attention incurs a cost.
- Can be modelled as an active sequential hypothesis testing problem (ASHT).

Mathematical Modeling - ASHT

Known oddball and distractor images model

- $H_i, i = 1, 2, \ldots, M$ - the M hypotheses.
- $\mathcal{A}$ - the set of all possible actions, $|\mathcal{A}| = K < \infty$.
- $(X_n)_{n\geq 1}$, $(A_n)_{n\geq 1}$ - the observation and action processes. In our visual search scenario X_n, a vector, corresponds to the number of spikes observed in each tapped neuron at time slot n.
- q_i^a - conditional probability density function of observation, given action $A = a$ and hypothesis $H = H_i$. In our visual search scenario, q_i^a corresponds to a vector Poisson distribution with corresponding means.
- $D(q_i^a \| q_j^a) := E\left[\log\left(\frac{q_i^a}{q_j^a}\right) |H = i\right]$ - relative entropy between the observation densities.
- Policy - $\pi = (\pi_1, \pi_2, \ldots)$
  - $\pi_n : \mathcal{X}^n \times \mathcal{A}^n \to \{(stop, d), (continue, \lambda)\}$
    $d \in \{1, 2, \ldots, M\}$ and $\lambda \in \mathcal{P}(\mathcal{A})$

Let P_i and E_i denote the conditional distribution and conditional expectation, conditioned on hypothesis H_i.

- Policies of interest. For a given $0 < \alpha < 1$

$$\Pi(\alpha) := \{\pi : P_i^{\pi}(d \neq i) \leq \alpha, \ \forall\, i\}\,. \qquad (1)$$

- Stopping time under policy π - $\tau(\pi)$.
- Switching cost - $g(a, a') \geq 0 \quad \forall a, a' \in \mathcal{A}$, $g(a, a) = 0 \ \forall a \in \mathcal{A}$
- Total cost -

$$C(\pi) := \tau(\pi) + \sum_{l=1}^{\tau(\pi)-1} g(A_l, A_{l+1}).$$

ASHT - Asymptotically Optimal Policies

Let

$$\lambda_i := \arg\max_{\lambda\in\mathcal{P}(\mathcal{A})} \min_{j\neq i} \sum_{a\in\mathcal{A}} \lambda(a) D(q_i^a \| q_j^a)$$

$$D_i = \max_{\lambda\in\mathcal{P}(\mathcal{A})} \min_{j\neq i} \sum_{a\in\mathcal{A}} \lambda(a) D(q_i^a \| q_j^a)$$

Let $Z_{ij}(n)$ denote the log-likelihood ratio (LLR) process of hypothesis H_i with respect to hypothesis H_j. $Z_i(n) = \min_{j\neq i} Z_{ij}(n)$.

Policy Sluggish Procedure A: $\pi_{SA}(L, \eta)$
Fix $L > 0$, $0 < \eta \leq 1$.
At time n:

- Let $\theta(n) = \arg\max_i Z_i(n)$. Ties are resolved uniformly at random.
- If $Z_{\theta(n),j}(n) < \log((M-1)L)$ for some $j \neq \theta(n)$ then next action A_{n+1} is chosen as follows.
  - Generate U_{n+1}, a Bernoulli(η) random variable, independent of all other random variables.
  - If $U_{n+1} = 0$, then $A_{n+1} = A_n$.
  - If $U_{n+1} = 1$, then generate A_{n+1} according to distribution $\lambda_{\theta(n)}$.
- If $Z_{\theta(n),j}(n) \geq \log(M-1)L$, for all $j \neq \theta(n)$, then the test retires and declares $H_{\theta(n)}$ as the true hypothesis.

Theorem. *Consider the sequence of probability of false detection constraints $(\alpha^{(n)})_{n\geq 1}$, such that $\lim_{n\to\infty} \alpha^{(n)} = 0$. Then, for each n, the policy $\pi_{SA}(L_n, \eta)$ with $\log L_n = -\log alpha^{(n)}$ belongs to $\Pi(\alpha^{(n)})$. Furthermore, for each i,*

$$\lim_{n\uparrow\infty} \inf_{\pi\in\Pi(\alpha^{(n)})} \frac{E_i[C(\pi)]}{\log L_n} = \lim_{\eta\downarrow 0} \lim_{n\uparrow\infty} \frac{E_i[C(\pi_{SA}(L_n, \eta))]}{\log L_n} = \frac{1}{D_i}.$$

Thus ASHT suggests D_i as a neuronal dissimilarity index. D_i normalised by number of neurons is denoted by $\tilde{D}$.

Unknown oddball and distractor images model - Learning based model

- What if we did not know the images that are going to appear? What if the only information the subject has is that there is one odd image in the group?
- In the neuronal space, it translates to finding the image with a firing rate different from the common rate of the others.
- We cast this as one of detecting an multi-dimensional Poisson point process having a rate different from that of the others.

Basic Notation

- K - Number of Poisson point processes. One of which one has a rate different from the others.
- H - Index of the odd process.
- R_1 - Rate of the odd process (unknown).
- R_2 - Rate of the non-odd processes (unknown).
- $\Psi = (H, R_1, R_2)$ specifies the configuration.

Let

$$\hat{f}\,(X^n, A^n | H = j) := \max_{\Psi : H = j} f\,(X^n, A^n | \Psi)\,,$$

and let

$$f(X^n, A^n | H = i) := \int f(X^n, A^n | \Psi = (i, \theta_1, \theta_2))\, f_{1,1,1,1}((i, \theta_1, \theta_2)|H = i) d\theta_1 d\theta_2.$$

The modified GLR is defined as

$$Z_{ij}(n) := \log\left(\frac{f(X^n, A^n | H = i)}{\hat{f}\,(X^n, A^n | H = j)}\right)$$

Note that the numerator is an averaged likelihood under $H = i$, averaged with respect to an artificial prior, and denominator is a maximum likelihood under $H = j$. Let

$$Z_i(n) := \min_{j\neq i} Z_{ij}(n)$$

denote the modified GLR with respect to i for the nearest alternate.

Policy: Modified GLRT ($\pi_M(L)$)
Fix $L \geq 1$.
At time n (end of slot n):

- Let $i^*(n) = \arg\max_i Z_i(n)$, the index with the largest modified GLR after n time slots. Ties are resolved uniformly at random.
- If $Z_{i^*(n)}(n) < \log((K-1)L)$,
  then A_{n+1} is chosen according to $\lambda^*(i^*(n), \hat{\theta}^n_{i^*(n)1}, \hat{\theta}^n_{i^*(n)2})$, i.e.,

$$\Pr(A_{n+1} = j | X^n, A^n) = \lambda^*(i^*(n), \hat{\theta}^n_{i^*(n)1}, \hat{\theta}^n_{i^*(n)2})(j).$$

- If $Z_{i^*(n)}(n) \geq \log((K-1)L)$ then the test retires and declares $i^*(n)$ as the true hypothesis.

Theorem. *Consider K homogeneous Poisson point processes with configuration $\Psi = (i, R_1, R_2)$. Let $(\alpha^{(n)})_{n\geq 1}$ be the sequence of probability of false detection constraint, such that $\lim_{n\to\infty} \alpha^{(n)} = 0$. Then, for each n, the policy $\pi_M(L_n)$ with $\log(L_n) = -\log(\alpha^{(n)})$ belongs to $\Pi(\alpha^{(n)})$. Furthermore,*

$$\liminf_{n\to\infty} \inf_{\pi\in\Pi(\alpha^{(n)})} \frac{E\,[\tau(\pi)]|\Psi]}{\log(L_n)} = \limsup_{n\to\infty} \frac{E\,[\tau(\pi_M(L_n)))|\Psi]}{\log(L_n)} \qquad (2)$$

$$= \frac{1}{D^*(i, R_1, R_2)}. \qquad (3)$$

The new model suggests D^* as a possible neuronal dissimilarity index.

Results

An ideal neuronal dissimilarity index $diff(k, l)$ would satisfy

$$E[\text{Stopping Time}]diff(k, l) = constant$$

for all image pairs (k, l). Then, a performance measure for a dissimilarity index would be its performance in an equality of means test. We now provide the statistics obtained in two equality of means tests: One-sided ANOVA and $\log$(Arith. Mean/ Geo. Mean). Lesser the statistics, better the performance.

[ht]

diff	ANOVA statistic	ANOVA p-values	log(AM/GM)
$\tilde{D}$	06.30	9.35×10^{-19}	0.0200
KL	06.68	2.88×10^{-20}	0.0211
Chernoff	06.74	1.61×10^{-20}	0.0252
L^1	24.00	3.42×10^{-87}	0.0678
D^*	06.34	6.93×10^{-19}	0.0233

Conclusions

- Framed the visual search problem as an Active Sequential Hypothesis Testing problem.
- ASHT suggests a neuronal dissimilarity $\tilde{D}$ index which explains the behavioural data as good as or better than L^1.
- Obtained D^* as an index when there is no prior knowledge of the image pairs.
- $\tilde{D}$ outperform D^*, though marginally, in all tests.

References

[1] A. P. Sripati and C. R. Olson, "Global image dissimilarity in macaque inferotemporal cortex predicts human visual search efficiency," *Journal of Neuroscience*, vol. 30, no. 4, pp. 1258–1269, Jan 2010.

Acknowledgements

We would like to acknowledge Dr. S. P. Arun from Center for Neuroscience, IISc for all the experimental data and the discussions on the problems.