Optimal prediction in natural gaze behavior

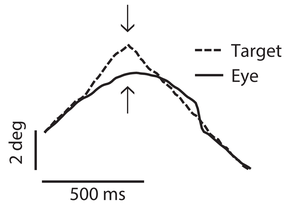
SangWook A. Lee[1], Stephanie E. Palmer[1,2], Leslie C. Osborne[1,2]

THE UNIVERSITY OF CHICAGO

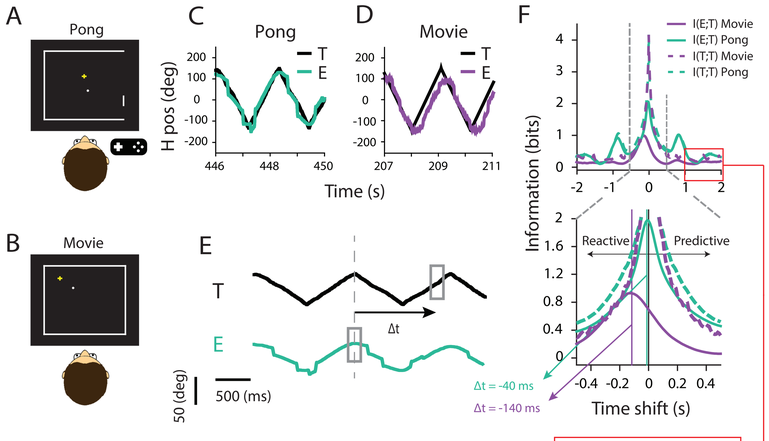
[1]Department of Neurobiology, [2]Department of Organismal Biology and Anatomy The University of Chicago, Chicago IL, USA.

1. Introduction

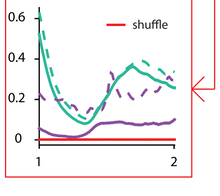
Delays in sensory processing give rise to a lag between a stimulus and the organism's reaction. This presents a particular challenge to tracking behaviors like smooth pursuit where a difference in eye and target motion creates image motion blur on the retina. One strategy that might compensate for processing delays is to extrapolate the future target position and make anticipatory eye movements. We recorded eye position in humans and monkeys engaged in tracking tasks to test for predictive information in eye movements. We used tasks with randomized jumps and direction changes (1D and 2D) and we created a Pong-like video game in which the subject moves a paddle to keep a ball bouncing within an arena. Using information theory, we show that in most tracking tasks, gaze behavior is predictive on short ~200ms timescales. But while playing Pong, prediction extends over several seconds and is very close to the bound imposed by the predictability of target motion. We also develop a model of short time scale prediction based on retinal inputs that accounts for decisions to saccade or pursuit during tracking.



2. Predictive information in human gaze behavior



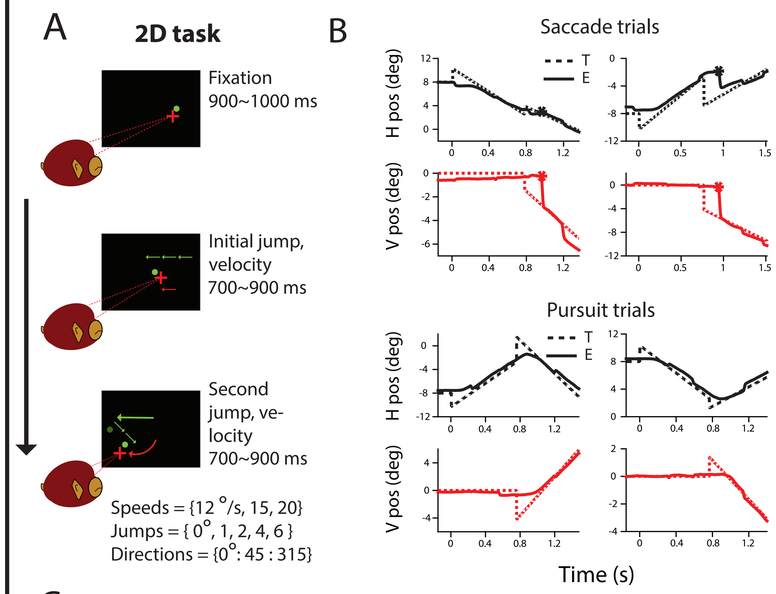
$$I(E(t),T(t+\Delta t)) = \left\langle -\sum_E \sum_T P(E(t),T(t+\Delta t)) \log_2 \left(\frac{P(E(t),T(t+\Delta t))}{P(E(t))P(T(t+\Delta t))} \right) \right\rangle_t$$

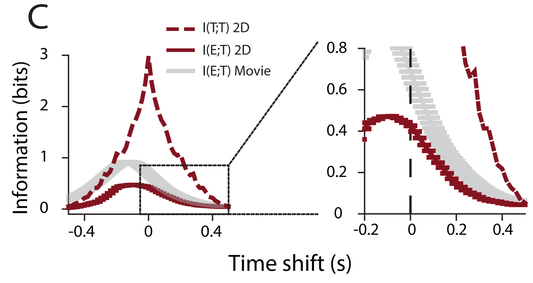


Context dependent predictive information in gaze behavior

- In active Pong, predictive information in eye position compensates for sensory delays, and is very near the bound determined by the predictive information in the target itself, thus all that is predictable about the target is incorporated into behavior and thus optimal.
- In contrast, watching a movie of the Pong game generates gaze that is more reactive than predictive.

3. Predictivie information in monkey gaze behavior

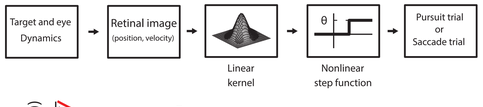


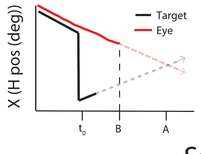


Predictive information in a randomized tracking task

- Predictive information peaks at the monkey's behavioral latency (120 ms) and extends over a similar range as human data in the Pong movie task replotted from Fig 2F (gray).

4. Gaze decision model - eye crossing velocity (VxE)





Eye crossing velocity (VxE)

$$\mathrm{VxE} = \frac{[\widehat{X_E}(t_0+\delta t) - \widehat{X_T}(t_0+\delta t)] - [\widehat{X_E}(A) - \widehat{X_T}(A)]}{A - (t_0+\delta t)}$$

$$\widehat{X_T}(A) = \left[\frac{\widehat{X_T}(t_0+\delta t) - \widehat{X_T}(t_0)}{(t_0+\delta t) - t_0}\right](A - (t_0+\delta t)) + \widehat{X_T}(t_0+\delta t)$$

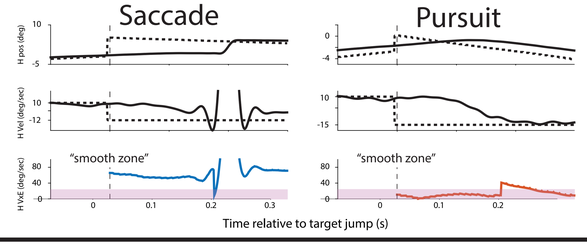
is extrapolated target position in the future,

$$\widehat{X_E}(A) = \widehat{V_E}(t_0+\delta t)(A - (t_0+\delta t)) + \widehat{X_E}(t_0+\delta t)$$

is an estimated eye position.

Modeling approach

- A model of gaze based on extrapolation of retinal inputs (image position, velocity). We find the linear kernel that best predicts saccades or pursuit after target jumps.
- A change in target motion is sampled after a delay, B. Target motion is extrapolated to time A in the future.
- If the difference in eye velocity needed to track the new target (VxE) vector is below a threshold, θ, pursuit is maintained, otherwise the subject saccades.



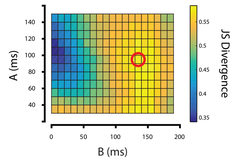
5. Model results

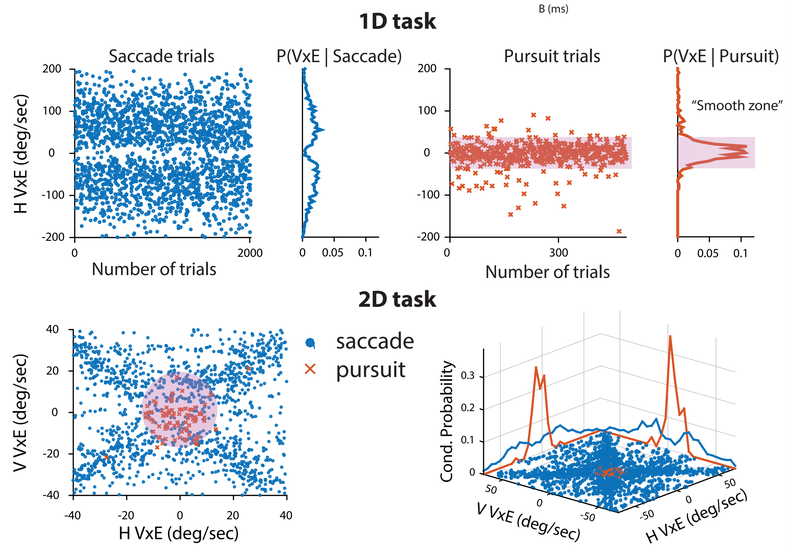
The VxE model predicts gaze across a range of times

We find the time A at which VxE is extrapolated, and the future time B at which VxE is evaluated to maximize Jensen-Shannon divergence (JS) of distributions P(VxE|Saccade) and P(VxE|Pursuit).

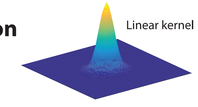
$$\gamma = \{A, B\}$$

$$\gamma = \underset{A,B}{\arg\max}\ JS(P(\mathrm{VxE}_{A,B}|\mathrm{Saccade}) \,||\, P(\mathrm{VxE}_{A,B}|\mathrm{Pursuit}))$$





Gaze prediction model

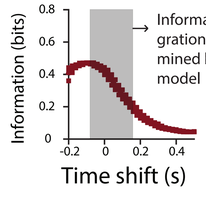


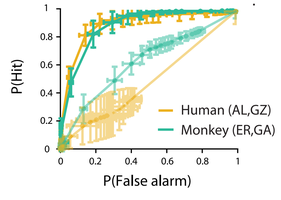
μ = [-5.51,4.19] °/s
σ = [7.84,6.21] °/s

$$f = P(\mathrm{Pursuit}|\mathrm{VxE}) = \frac{P(\mathrm{VxE}|\mathrm{Pursuit})P(\mathrm{Pursuit})}{P(\mathrm{VxE})} \sim \mathcal{N}(\mu, \sigma)$$

- We cross validate the kernel using ROC analysis to quantify the quality of its gaze predictions on novel data

Compare to predictive information





$AUC_{Human\_VxE} = 0.86$

$AUC_{Monkey\_VxE} = 0.83$

$AUC_{Human\_TxE} = 0.48$

$AUC_{Monkey\_TxE} = 0.61$

6. Summary

We find that the brain incorporates prediction in gaze behavior in a context dependent manner. Prediction on short time scales (~200-400ms) helps to compensate for visual delays. A model based on extrapolation of retinal inputs (eye crossing velocity) predicts gaze decisions on shorter time scales (i.e. decision to saccade or pursue). Active tasks like playing Pong produce predictive movements on 5-10x longer time scales, and are nearly optimal. The next step is to model longer timescale gaze patterns (i.e. continuous gaze behavior) during Pong, and to consider patterns of motion over time in the analysis of predictive information.

References :
[1] De Brouwer, Sophie, et al. "What triggers catch-up saccades during visual tracking?." Journal of Neurophysiology 87.3 (2002): 1646-1650.
[2] Palmer, Stephanie E., et al. "Predictive information in a sensory population." Proceedings of the National Academy of Sciences 112.22 (2015): 6908-6913.

Funding:
This research is supported by grants to LCO from Alfred P Sloan Foundation, Whitehall Foundation, Brain Research Foundation and NEI EY023371, and NSF IOS 1457024.