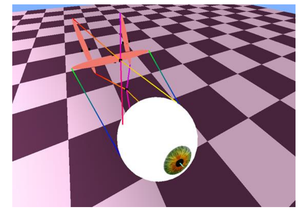
Learning open-loop saccadic control of a 3D biomimetic eye using the actor-critic algorithm

Henrique Granado[1], Reza J. Alitappeh[1], Akhil John[1], A. John van Opstal [2], and Alexandre Bernardino[1]

1. ISR, Instituto Superior Tecnico, Lisbon, Portugal
2. Neurophysics, DCN, Radboud University, Netherlands

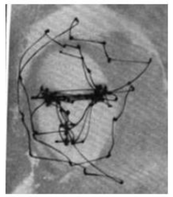


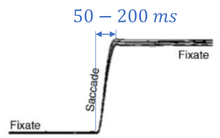
Objectives

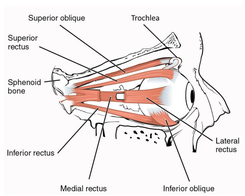
- **Understand how brain controls eye saccades by using a bio-inspired robotic eye.**
- **We utilize reinforcement learning with biologically motivated action and rewards.**
- **We aim to see stereotypical kinematic characteristics of primate saccades emerge from our control.**

Eye Saccades

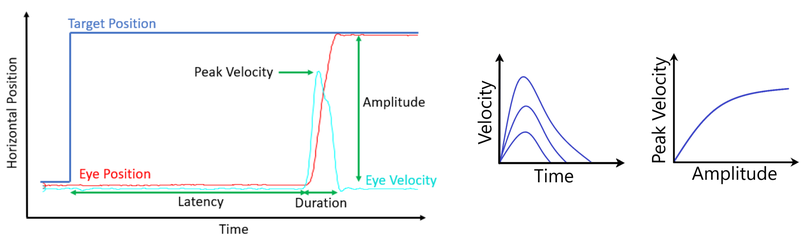
- **Rapid eye movements (~700º/sec) to change eye fixation.**



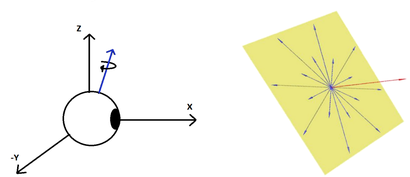




- **Duration too short for visual feedback.**
- **6 muscles control 3 DOF rotations**
- **High viscous friction due to surrounding fat tissues.**
- **Stereotypical velocity characteristics.**



- **Stereotypical orientation characteristic (Listing's Law)**

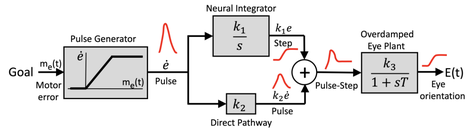


The rotation vectors for orientations from primary position all lie in a plane.

Existing Works

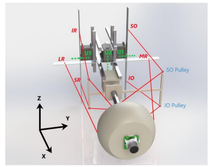
Works on understanding these characteristics:

1.(1973) Robinson 1D model control theory



2.(2010) Shadmehr et al. – Optimal control 1D model

3.(2021) John et al. – Optimal control 3D model

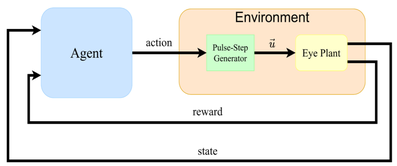


1.D. A. Robinson, "Models of the saccadic eye movement control system," Biologic Cybernetics, 1973.
2.R. Shadmehr, J. Orban, M. Xu-Wilson, and T.-Y. Shih, "Temporal discounting of reward and the cost of time in motor control," J. Neuroscience, vol. 30, pp. 10 507–16, 08 2010.
3.A. John, C. Aleluia, A. J. Van Opstal, and A. Bernardino, "Modelling 3d saccade generation by feedforward optimal control," PLOS Computational Biology, vol. 17, no. 5, pp. 1–35, 05 2021.

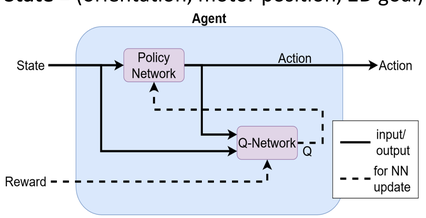
Our Approach

Agent – Actor-Critic Algorithm

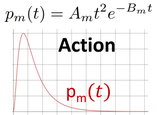
- **Policy Network** : Defines Action
- **Q Network** : Estimates reward for action.

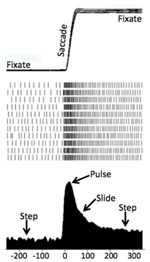


State = (orientation, motor position, 2D goal)

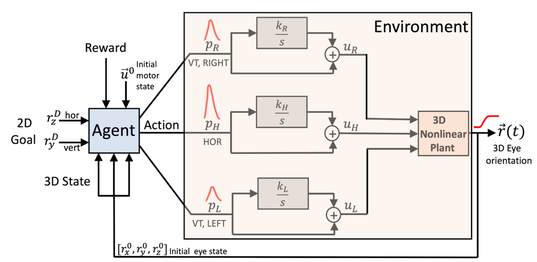


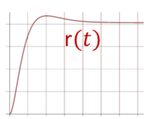
$$p_m(t) = A_m t^2 e^{-B_m t}$$



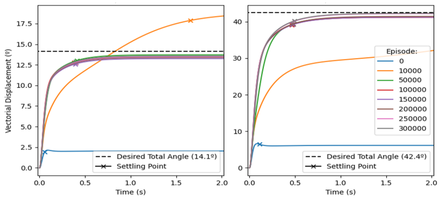


Reward = $\lambda_a R_{acc} + \lambda_e R_{en} + \lambda_d R_{dur} + \lambda_f R_{for} + \lambda_o R_{over}$

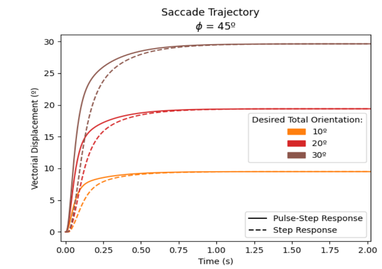




Results



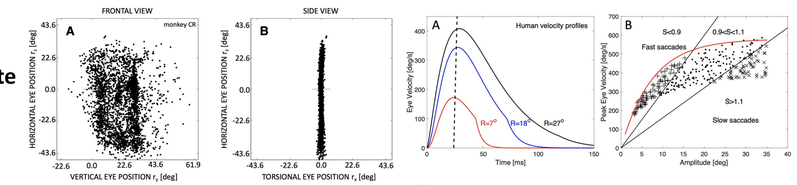
Amplitude responses for two different goals, [10º , 10º] and [30º , 30º], respectively, during different stages of the training (legend).



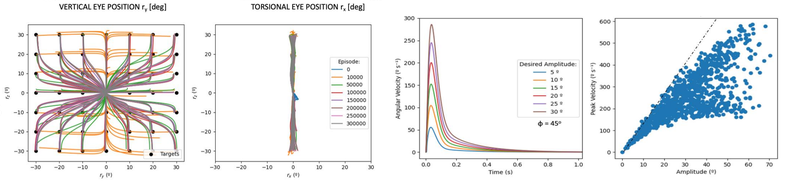
Showing difference between responses to the trained pulse-step signal (solid traces) vs a simple step signal (dashed), for 10º, 20º and 30º oblique saccades after training (300k episodes).

Listing's Plane | **Velocity Characteristics**

Primate Data



Our Results



Conclusions

- **Generated 3D saccades with human-like characteristics from optimizing accuracy, duration, energy, fixation force and overshoot.**
- **Our results emerged without imposing any constraints.**
- **In future, we wish to apply the method to a more realistic robotic eye with 6 independently actuated muscles.**

European Research Council
Established by the European Commission

LARSyS
Laboratory of Robotics and Engineering Systems