

# A Recurrent Neural Network Based Model of Predictive Smooth Pursuit Eye Movement in Primates

Hirak J Kashyap<sup>1</sup>, Georgios Detorakis<sup>2</sup>, Nikil Dutt<sup>1,2</sup>, Jeffrey L Krichmar<sup>1,2</sup>, Emre Nefci<sup>1,2</sup>

<sup>1</sup>Department of Computer Science

<sup>2</sup>Department of Cognitive Sciences

University of California, Irvine

Irvine, California, USA 92697

Email: {kashyaph, gdetorak, dutt, jkrichma, eneftci}@uci.edu

**Abstract**—A predictive mechanism in the brain enables primates to visually track a target with almost zero lag smooth pursuit eye movements, overcoming the delays in processing retinal inputs. Interestingly, it also allows pursuit of occluded targets with nonlinear motion patterns. We propose a recurrent neural network (RNN) model that rapidly learns the target velocity sequence and generates eye velocity signals to eliminate the initial lag between target and eye velocities, and to track occluded targets with nonlinear velocity. Moreover, the model is able to adapt to unpredictable perturbation and phase shift of target velocity and qualitatively reproduce the initial pursuit acceleration in experimentally observed timescales. We propose that the frontal eye field (FEF) region of the primate brain is homologous to the proposed RNN based on its persistent predictive activities during pursuit and location on the pursuit pathway.

## I. INTRODUCTION

Primates are incredibly proficient at tracking a visual target with their eyes. Since their foveal vision is narrow, they rotate their eyes continuously in the direction of the moving target to keep it centered on the fovea, which provides high acuity information. This type of eye movement is known as smooth pursuit, as eye velocity changes smoothly in response to target velocity. The pursuit system is able to track with almost zero lag between eye and non-linear target velocities [1]. This behavior is particularly impressive, since due to sensory and processing delays of 80-100 ms in the visual pathways [2], a system that solely relies on retinal error feedback for oculomotor movements cannot achieve near zero lag pursuit. Further, the primate eyes continue smooth pursuit of a target after its disappearance [3], [4]. These two pursuit behaviors indicate a predictive mechanism that is able to generate current and future eye velocities based on the target motion sequence in the past [1].

Early control-theoretic models of smooth pursuit [5], [2] did not consider the predictive capabilities of pursuit, rather they tried to mimic the experimentally observed typical initial acceleration, overshoot, and response latency, while tracking a

constant velocity stimulus. Later models used prior knowledge or memory from previous trials to eliminate sensory delays and to be able to continue pursuit during occlusions [6], [7], [8]. However, these memory-based models are not plausible since they rely on periodicity of target velocity, resulting in periodic improvement of pursuit lag [9]. They also cannot adapt to transient perturbations in target velocity, whereas humans adapt to perturbation and phase shift of a sinusoidal target within a cycle [10].

A fundamental problem associated with pursuit is to generate eye velocity that persists while target velocity on retina is zero. This takes place during perfect zero lag tracking and during target occlusion. The neural mechanism that generates predictive eye velocity signals in absence of visual inputs is not known from the existing pursuit models. The state-of-the-art Kalman filter based predictive pursuit models [6], [9] cannot generate persistent eye velocity during long occlusions of a target with nonlinear velocity, due to absence of the error feedback to correct filter predictions. In contrast, both humans and monkeys were found to continue pursuit of a target with sinusoidal velocity after disappearance [3], [4].

In this paper, we propose an RNN based smooth pursuit model that rapidly learns the target velocity sequence and generates self-sustained predictive eye velocity signals to track a moving target with near zero lag. The model is able to i) gradually eliminate the initial lag between eye and target velocities, ii) track an occluded target with nonlinear velocity, iii) adapt to unpredictable perturbation and phase shift in target velocity, and iv) qualitatively reproduce the typical initial pursuit acceleration observed in experiments [10], [4], [3], [11]. The RNN has a spontaneous baseline activity and is trained online using the FORCE learning procedure [12]. We also map our model to the pursuit pathway in the primate brain based on neuroanatomy and behavior of the brain regions [11]. We propose that the FEF region of the frontal cortex, which generates predictive eye velocity signals during pursuit of visible and hidden targets [13], [11], [14], is a plausible neural

correlate of our proposed RNN based on its activities during pursuit and location on the pursuit pathway.

## II. BACKGROUND

### A. Predictive smooth pursuit behavior

The velocity of a target on the retina, known as retinal slip ( $RS$ ), is relied upon to generate and maintain accurate smooth pursuit eye movements.  $RS$  can be considered as the error between eye and target velocities. Since  $RS$  becomes zero during zero lag tracking and occlusion, it cannot be used to drive pursuit eye movements. However, it may be used to correct them, as suggested in [15]. Moreover, the  $RS$  information is delayed by approximately 80-100 ms during early visual processing [15], [2]. As a result, a pursuit system based on a delayed  $RS$  is not stable [15]. The pursuit system needs to predict eye velocity 80-100 ms into the future using available  $RS$  information in order to eliminate lag between eye and target velocities.

If a target's motion is predictable, the primate vision system learns the spatio-temporal sequence of target velocity online [4], [3], [10]. This learned velocity pattern, also referred to as the internal model, gradually reduces lag between target and eye velocities during tracking. It also enables the vision system to continuously track a target during occlusion. Experiments with both humans and monkeys [4], [3] observed that when a sinusoidal target was suddenly turned off during an ongoing pursuit, the transition from target on to target off did not produce any transient deviation in pursuit velocity. Both studies concluded that the internal model was used to generate eye velocity even when the target was visible. Furthermore, Van den Berg [10] observed that humans adapted to unexpected perturbations and phase shifts of a sinusoidal target within a cycle.

### B. Previous models

Early pursuit models implemented a feedforward controller that canceled out efferent feedback to achieve high velocity gain [5], [16]. Krauzlis and Lisberger proposed a feedback control model using parallel velocity and acceleration pathways with second order filters to process  $RS$  [2]. This model used under-damped filters to generate the ringing behavior of the earlier models [5]. None of these models achieved zero lag pursuit of a periodic signal because they relied on a delayed  $RS$  information.

From control-theoretic perspective, the current target velocity information is required to predict current eye velocity without lag. To work around this, later models used prior knowledge or memory of target motion to estimate the current target velocity. Bahill and McDonald proposed a model for generating pursuit eye movements based on *a priori* knowledge of target trajectory [17]. Other memory based models proposed to generate pursuit eye movements based on stored patterns for periodic trajectories [8]. Orban de Xivry et al. [6] proposed a memory based model that used the target trajectory stored from prior trials to run a Kalman filter for eye velocity prediction. However, as indicated by [9], memory

based models are biologically not plausible, because i) they require a periodicity estimator in the brain, ii) the improvement in pursuit lag by memory-based models can only be periodic, whereas studies [10], [11] found a rather gradual decrease in lag between target and eye velocities, and iii) they cannot adapt to unpredictable perturbations of a periodic signal within a single cycle as observed in humans.

Shibata et al.'s model removed the requirement for prior knowledge of target trajectory and adapted the parameters of a Kalman filter online for eye movement predictions [9]. However, these Kalman filter based approaches have drawbacks, i) they cannot account for acceleration/deacceleration caused by the external target and operate using the negative feedback of prediction error and ii) filter state and prediction are static during occlusions. Both drawbacks arise since these models do not learn the target motion pattern.

As this paper focuses on the computations for the predictive pursuit behavior observed during experiments on primates [4], [3], [10], we will not discuss other models related to pursuit, such as target background separation for pursuit [18] and co-development of pursuit and motion perception [19], [20], as well as other types of eye movements.

## III. THE PROPOSED MODEL

Studies on primates have shown that the predictive pursuit system learns the spatio-temporal sequence of target velocity [4], [3], [10]. For this, the predictive system in the brain solves three challenges that arise from biology and the nature of the visual tracking task. First, the target velocity sequence is learned rapidly (e.g. within a few cycles for periodic target velocity [21]) to eliminate pursuit lag. Second, the learning occurs using an  $RS$  information that is 80-100 ms delayed during sensory processing. Third,  $RS$  vanishes during zero lag pursuit or occlusion, and therefore cannot be used to drive eye velocity.

Considering these challenges, we propose a smooth pursuit model using an RNN to rapidly learn the target velocity sequence and predict eye velocity during both presence and absence of retinal inputs. The model produces eye velocity prediction from spontaneous neural activations and uses the delayed  $RS$  as the error signal for learning. The eye velocity prediction is then converted to actual eye motor movements by a downstream Inverse Dynamics Controller (IDC). The processing by vision and attention systems to extract the target velocity from the visual field and the IDC are not modeled in this paper, as efficient models already exist for these tasks. Figure 1 depicts the complete predictive smooth pursuit model.

The target velocity on the retina, i.e.  $RS$ , results from actual movement of the target in the three-dimensional world and/or eye/head/body movements. Since target motion projected onto the retina is relative to eye motion, the resulting  $RS$  is the difference between head-centered target velocity ( $v_T$ ) and eye velocity ( $v_E$ ). When the eye perfectly tracks a moving target with foveal vision,  $RS$  will be zero.

$$RS(t) = v_T(t) - v_E(t) \quad (1)$$

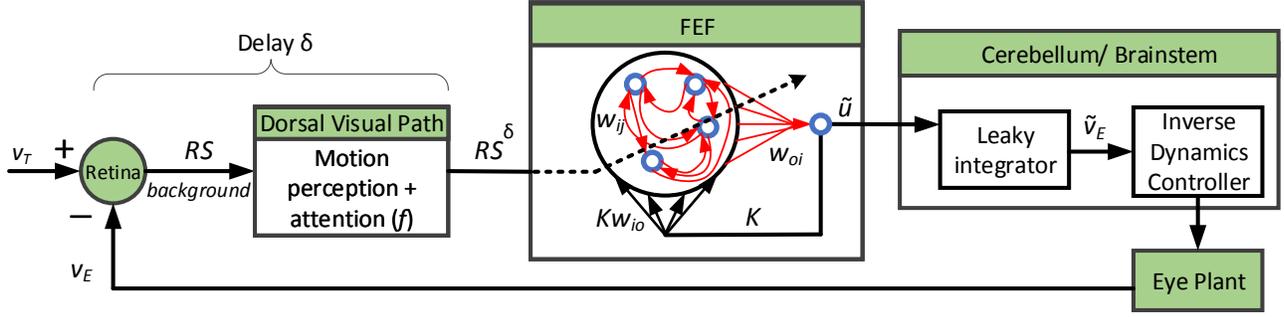


Fig. 1. The proposed model for predictive smooth pursuit eye movement generation in primates. The plausible brain regions performing the specific functions in the pursuit pathway are shown in green colored boxes. The retinotopic  $RS$  is extracted from visual field with a time delay of  $\delta$  by the dorsal visual pathway ( $RS^\delta$ ). A recurrent network of neurons (blue circles) in the FEF region uses  $RS^\delta$  to learn the target velocity sequence and generates  $\tilde{u}$ , which is then low pass filtered by a leaky integrator to obtain eye velocity predictions ( $\tilde{v}_E$ ). All red colored synaptic connections are modified during learning. Cerebellum and Brainstem together implement an inverse dynamic controller to generate the final eye velocity ( $v_E$ ) via oculomotor control.

The retinal output, where  $RS$  of the target is embedded in a visual scene background, is transmitted through two interconnected visual cortical pathways, one that recognizes objects (ventral pathway) and the other that extracts motion components in the visual field (dorsal pathway). In Figure 1, we only show the dorsal pathway as it processes motion information. The dorsal visual pathway extracts motion of all objects in visual field and the attention system selects the target motion component from background, in time  $\delta$ . We term the cumulative operations performed by the dorsal pathway as  $f$  and its target velocity output as  $RS^\delta$ .

$$RS^\delta(t) = f(RS(t - \delta)) \quad (2)$$

where,  $RS^\delta$  is delayed by time  $\delta$  since its projection on the retina.

The RNN, shown inside the box labeled FEF in Figure 1, uses  $RS^\delta$  as the error signal to learn target velocity sequence online and generates predictive eye velocity signals. Similar predictive activities during pursuit have been observed in the FEF region of frontal cortex, which receives outputs from the dorsal pathway [13], [11], [14].  $RS^\delta$  vanishes as the lag between target and eye velocities is reduced during learning, therefore the RNN needs to predict eye velocity in absence of visual inputs.

This RNN is a type of reservoir computer [22], [23] and is trained online using the FORCE learning procedure [12]. The RNN has 500 neurons connected all-to-all and operates in a chaotic regime. All neurons of the RNN connect to a single readout neuron ( $o$ ) via readout synapses. The recurrent and readout synapses are plastic and are modified online using  $RS^\delta$  as the error. The output of the readout neuron  $\tilde{u}$  is fed back to the RNN through fixed random weights with a gain  $K$ . The signal  $\tilde{u}$  is low pass filtered to obtain the eye velocity prediction  $\tilde{v}_E$ .

The neurons in the RNN follow the dynamics proposed

in [12]. Hence, the dynamics of neuron  $i$  can be written as:

$$\tau \frac{dx_i}{dt} = -x_i + \sum_{j \in Pre(i)} w_{ij} r_j + K w_{io} \tilde{u} \quad (3)$$

Where,  $\tau$  is the time constant,  $x_i$  is the neuron state,  $Pre(i)$  is the set of neurons that projects to post-synaptic neuron  $i$ ,  $w_{ij}$  is the weight of the synapse from neuron  $j$  to neuron  $i$ ,  $w_{io}$  is the weight of the synapse from the readout unit to neuron  $i$ , and  $r$  is the non-linear activation function given as,

$$r_i = \tanh(x_i) \quad (4)$$

The readout unit linearly combines neuron activations weighted by  $w_{oi}$ , which can be considered as an integrator with unit gain and time constant.

$$\tilde{u} = \sum_i w_{oi} r_i \quad (5)$$

The leaky integrator performs low pass filter on the output of the readout unit with gain  $K_l$  and time constant  $\tau_l$  to generate pursuit eye velocity prediction  $\tilde{v}_E$ , following the dynamics given by:

$$\tau_l \frac{d\tilde{v}_E}{dt} = -\tilde{v}_E + K_l \tilde{u} \quad (6)$$

The recurrent and the readout weights are updated periodically after every  $\delta$  time units, same as the visual processing delay. This is done because the effect of weight updates on  $RS$  is obtained after  $\delta$ . The FORCE learning procedure is applied to learn all recurrent and readout weights using the same delayed error signal  $RS^\delta$ . The weight update for the synapse from neuron  $j$  to neuron  $i$  is defined as,

$$w_{ij}(t) = w_{ij}(t - \delta) - RS^\delta(t) \sum_{k \in Pre(i)} P_{jk}(t) r_k(t) \quad (7)$$

Similarly, the readout weights are updated as,

$$w_{oi}(t) = w_{oi}(t - \delta) - RS^\delta(t) \sum_{k \in Pre(o)} P_{ik}(t)r_k(t) \quad (8)$$

where,  $P$  is a matrix containing individual learning rates for all synapses, updated regularly. It is initialized to  $I/\alpha$ , where  $\alpha$  is some constant and  $I$  is the identity matrix.  $P$  is updated as the inverse of the correlation matrix of neuron activations plus a regularization term  $\alpha I$  [12].

Biological interpretation of this type of neural dynamics has been suggested previously based on experimental data [24]. Cortical neural networks maintain a spontaneous baseline activity, which is chaotic and inherently unstable. However, short-term plasticity based on pre-synaptic firing dynamically tunes these networks to be stable and respond reliably to external stimuli. This self tuning principle allows these cortical networks to respond to external perturbations with characteristic transient response.

Similar to [9], we follow the findings and the theory that the cerebellum and the brainstem together implement an IDC, which cancels the dynamics of the eye plant [25]. As in [9], we assume that the IDC is ideal, and therefore we can write,

$$v_E = \tilde{v}_E \quad (9)$$

where,  $\tilde{v}_E$  is the low pass filtered eye velocity prediction. With this assumption, we will not model IDC dynamics in this paper.

#### IV. EXPERIMENTS AND RESULTS

We test the proposed predictive pursuit model on three characteristic pursuit tasks. First, we compare the pursuit initiation behavior of the model with experimental data from studies on primates [26], [27]. Second, a predictive pursuit task of a sinusoidal target, where we evaluate the capability of the model to eliminate lag between target and eye velocities caused by sensory delays and perform predictive pursuit of occluded objects [3], [4]. Third, we evaluate the ability of the model to adapt to unpredictable perturbations and phase shifts of target velocity in experimentally observed timescales [10].

In all the experiments, the RNN contains  $N = 500$  neurons and they are fully connected. The initial recurrent weights are drawn from a Gaussian distribution with mean 0 and standard deviation  $g/\sqrt{N}$  with  $g = 1.5$ , which results in a spontaneous chaotic behavior [12]. The readout weights are initialized to zeros. The feedback weights connecting the readout unit to the RNN neurons are drawn uniformly from the range -1 to 1 with gain  $K = 1$ . The integration timestep is 16 ms. The time constant  $\tau$  is set to 160 ms and like previous models, the sensory delay  $\delta$  is set to experimentally observed value of 80 ms [2], [6]. For the leaky integrator, the time constant  $\tau_l$  is equal to 128 ms. The values of  $\alpha$  and  $K_l$  are set to 1.25 and 0.5, respectively, for the initiation experiment, and to 100 and 1, respectively, for the predictive pursuit experiment. The Matlab code used in the experiments is available at <https://github.com/hkashyap/predictivePursuit>.

Similar to [6], pursuit onset is detected by fitting a piecewise linear function (0 before pursuit onset  $T$  and  $A(t - T)$  after  $T$ ) to  $\tilde{v}_E$  traces during an interval of 320 ms starting from stimulus onset. Similarly, for the initiation experiment, mean eye acceleration ( $B$ ) is calculated by fitting  $\tilde{v}_E$  traces during the interval 80-180 ms after pursuit onset to  $\tilde{v}_E(T + 0.08) + B(t - (T + 0.08))$ . The interval is selected to compare with experimental eye acceleration data [26], [6].

##### A. Pursuit initiation

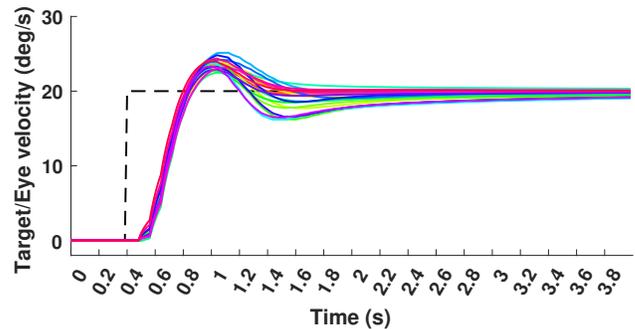


Fig. 2. Eye velocity during pursuit initiation in response to a ramp stimulus of constant velocity 20 deg/s. The black dashed line depicts the target velocity. The colored lines are the eye velocity responses generated by the proposed model in 20 trials.

The smooth pursuit observed during sudden movement of a target with constant velocity after fixation, known as a ramp stimulus, has a very characteristic initial acceleration profile, as observed in primate experiments [5], [2], [28], [29]. Figure 2 depicts the pursuit responses generated by our model for a ramp stimulus of velocity 20 deg/s that starts at 400 ms. In all trials, our model generates the typical initial acceleration and the following overshoot, comparable with experimental observations and outputs of the previous models [5], [2], [6]. The pursuit response latency of our model from the onset of the stimulus is  $146 \pm 13.7$  ms (mean $\pm$ SD), which is similar to the average pursuit response latency of 150 ms measured experimentally for the ramp stimulus [27].

Consistent with the recent studies [28], [29], [27], our model does not show an oscillatory behavior for the ramp stimulus after the overshoot. The recent predictive pursuit model by Orban de Xivry et al. [6] did not produce the oscillatory behavior either. Mainly early control theoretic pursuit models resulted in the oscillatory behavior [2], [5]. However, the pursuit model by Krauzlis and Lisberger [2] required a separate image-acceleration pathway, in addition to an image-velocity pathway, that generated the oscillatory behavior. Our model does not require separate mechanisms for pursuit acceleration and prediction.

We compare the mean eye acceleration during pursuit initiation generated by the proposed model versus experimental data of humans presented in [6], originally from a dataset by de Brouwer et al. [26]. Figure 3 depicts the comparison of mean eye acceleration profiles in response to targets velocities -50 deg/s to 50 deg/s at an increment of 5 deg/s. Similar to [6], eye

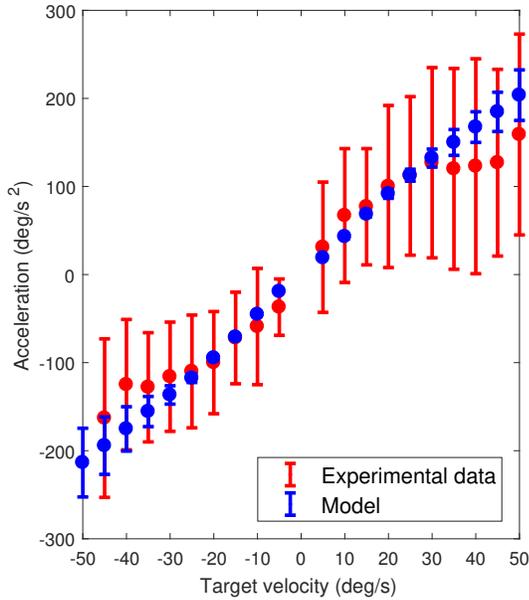


Fig. 3. Mean eye acceleration versus target velocity between 80 ms and 180 ms after pursuit onset. Blue circles correspond to predictions by the proposed model and red circles correspond to experimental data from de Brouwer et al. [26], reproduced from [6]. Vertical bars are the standard deviations from mean. Experimental data is not available for target velocity -50 deg/s.

acceleration is calculated during the interval 80 ms to 180 ms after pursuit onset. The plot shows that acceleration generated by the proposed model during pursuit initiation follows a comparable trend as the experimental data. Similar to the experiment, the standard deviation of acceleration generated by our model gradually increases for higher target velocities, which is caused by large weight updates due to higher  $RS$  error signals. The mean acceleration produced by our model matches the experimental data for target velocities up to 30 deg/s. Beyond this range, the experimental data shows the effect of physiological limits as the acceleration plateaus. Similar to the previous models [30], a saturation function for acceleration may be used to reproduce this behavior.

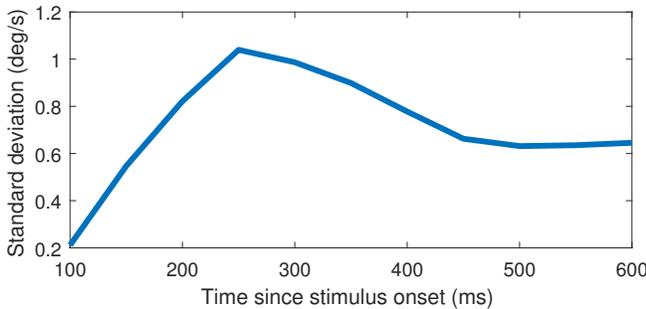


Fig. 4. The standard deviations of  $\tilde{v}_E$  traces at different time points after the stimulus onset with target velocity 20 deg/s, compares to Figure 1(b) of [27].

The variations in pursuit velocity observed between trials in Figure 2 is attributed to different initial states of the reservoir.

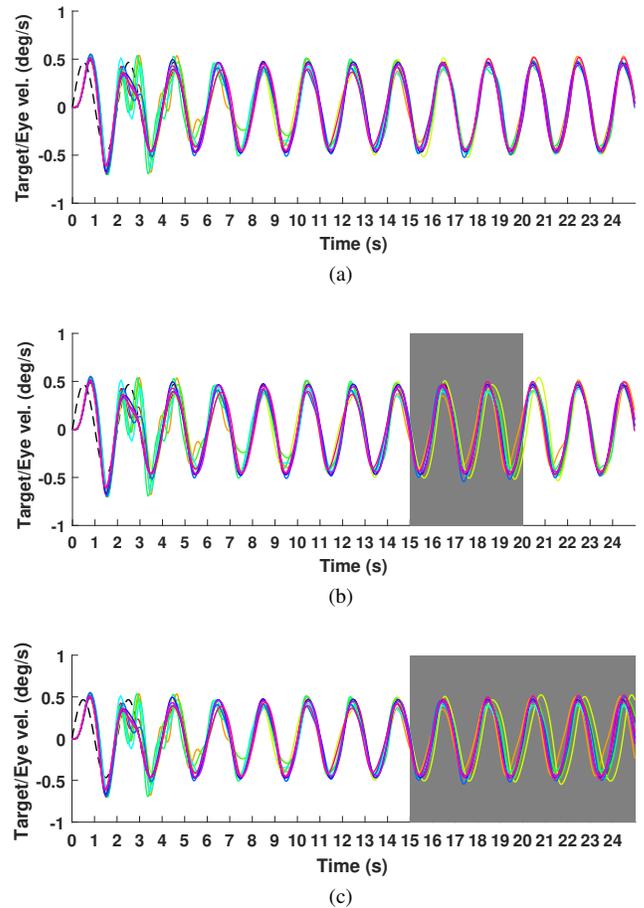


Fig. 5. The pursuit eye velocity generated by our model in response to sinusoidal target velocity pattern. The black dashed line is the target velocity and the colored lines are the eye velocity simulated using the proposed model. The grey areas are the time periods where the target is occluded. (a) The target is always visible, (b) the target is temporarily occluded and then reappears, and (c) the target is permanently occluded after 15 seconds.

Figure 4 depicts the standard deviation among the eye velocity traces at different time points after the onset of the stimulus with target velocity 20 deg/s. Similar to the experimental observations on humans and monkeys [27], [31], the standard deviation of eye velocity generated by our model in 20 trials jumps approximately 1 deg/s from its starting value (0 deg/s in our case) within 300 ms from the stimulus onset, which then linearly decreases and settles to approximately 0.64 deg/s between 500 ms and 600 ms from the stimulus onset.

### B. Predictive pursuit

Similar to the existing predictive pursuit models [9], [6] and experiments [10], [4], [3], [11], we test our model on a sinusoidally varying target velocity pattern. Primates track sinusoidal targets with little or no lag [11], [10]. Figure 5 demonstrates the results of the experiment, which shows the predictive capability of the proposed model in terms of almost zero lag pursuit and sustained tracking performance during occlusion. Figure 5a depicts the experiment where the target follows a sinusoidal velocity pattern with amplitude 0.47 deg/s and frequency 0.5 Hz. Since, the readout weights of the RNN

are initialized to zero, the initial eye velocity is 0 deg/s. The retinal slip or error signal for learning is not available to the RNN during first 80 ms after target onset (sensory delay) and therefore, the eye lags behinds the target (evident from the target and the eye velocity plots in Figure 5a). Learning process starts after 80 ms and despite using a delayed error signal, it is able to eliminate the phase lag between the target and eye velocities within the first cycle of the sinusoid. Within a few cycles of the sinusoid, the eye velocity closely follows the target velocity. Primate experiments using periodic stimuli also observe that the phase error between target and eye velocities becomes small within the third cycle of sinusoid [21], [10].

Figure 5b and Figure 5c illustrate the effect on pursuit performance due to temporary and permanent occlusion of the target, respectively, after the model has learned the target velocity pattern. In Figure 5b, when the target is occluded from 15 s to 20 s after target onset, the model continues to generate a eye velocity pattern that closely resembles the occluded target’s velocity. Although, a phase error develops between the eye velocity and the target velocity during occlusion. After the target reappears, the phase error is corrected within a single cycle, much faster than the initial learning. Figure 5c shows that when the target is occluded permanently at 15 s after target onset, the model continues to generate sinusoidal eye velocity pattern for many cycles. However, the phase error between the target and eye velocity gradually increases. Similar experiments on humans and monkeys [4], [3] report that pursuit movement continues for a few cycles after a sinusoidal target is turned off and then a phase error develops gradually.

### C. Unpredictable perturbation and phase shift

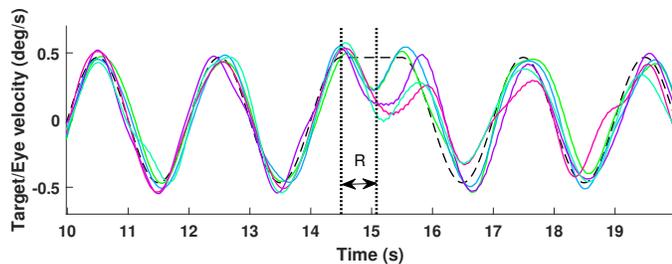


Fig. 6. Response of the predictive pursuit model to unpredictable perturbation and phase shift. Black dashed line is the target velocity and the colored lines are the eye velocity generated by the model in 5 trials.  $R = 0.58$  s is the experimental reaction time since perturbation calculated using Van den Berg’s formula [10]. Compares to Figure 8(a) of [10]. Pursuit starts at 0 s (not shown).

In human subject experiments, Van den Berg [10] studied the effect of unpredictable perturbation and phase shift on predictive pursuit by replacing a sinusoidal velocity stimulus with a ramp stimulus for half cycle. During perturbations, the eye initially accelerated following the original course of the sinusoidal target before reversing acceleration to match the modified velocity. During this transition, the time at which the acceleration becomes zero since the beginning of perturbation

is known as the reaction time (R). The study observed maximum reaction times when the target velocity was perturbed at the peak of the sinusoid, which were larger than one quarter of a cycle and approximated using the following formula.

$$R = \frac{1}{4 \times \text{frequency}} + 0.08s \quad (10)$$

Figure 6 depicts the eye velocity generated by our model during the same experiment, where perturbation occurs at the peak velocity of a 0.5 Hz sinusoidal target for half cycle. The reaction times achieved by our model are close to 0.58 s, the experimental reaction time obtained using Equation 10. Similar to the experimental data, our model adapts to the new phase within the first cycle after perturbation, and the phase error caused by the perturbation is corrected during the first two cycles after perturbation. The results from our model and Van den Berg’s experiment [10] show that the predictive pursuit system continuously learns the target velocity at a fast learning rate, which is not possible in memory based models.

### D. Unpredictable target velocity

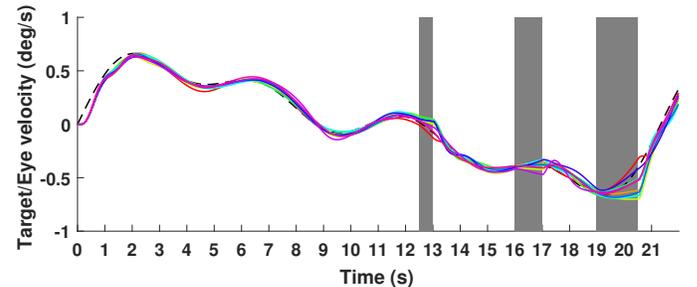


Fig. 7. Eye velocity prediction by our model in response to an unpredictable target velocity input. The black dashed line is the target velocity sequence and the colored lines are the model output during ten trials. The gray regions are occlusions.

We also test the response of our model when the target velocity is not predictable. Similar to the human pursuit experiment for unpredictable targets [32], we use a pseudo-random target sequence that is a sum of four sine waves with different frequency and amplitude. The target velocity is not predictable, as evidenced by the deviation in pursuit prediction during the three blank periods, shown in Figure 7. Similar to the experimental observations [32], [10] for unpredictable targets, our model is able to reduce the initial sensory lag using continuous prediction and then switches between phase lead and phase lag to maintain small prediction error. The third occlusion at 19 s illustrates the unpredictability of the target movement, as the model expected the target velocity to either plateau or increase or decrease. The deviations caused by target blanking are corrected after target reappearance.

## V. DISCUSSION

In this paper, we present a computational model of predictive smooth pursuit eye movement in primates. The model implements the predictive mechanisms using an RNN and is able to achieve almost zero lag tracking of sinusoidal targets by

eliminating sensory delays, pursuit of an occluded target with a non linear velocity profile, and adaptation to unpredictable perturbation and phase shift of target velocity in experimentally observed timescales. The model also qualitatively reproduces the experimentally observed initial pursuit acceleration. To the best of our knowledge, this is the first neural network model to achieve all of the above mentioned primate smooth pursuit behaviors. It demonstrates that a single neural network can generate pursuit initiation dynamics and persistent predictive pursuit signals. Although, pursuit experiments on primates suggested that an internal model of target motion may be used for pursuit prediction [21], [4], [3], [10], the neural mechanism to create and maintain the internal model was not known. Our work shows how the internal model is learned and updated rapidly by an RNN using a delayed  $RS$  signal as error, in order to reduce tracking lag, generate persistent pursuit during occlusions, and correct eye velocity during target perturbations.

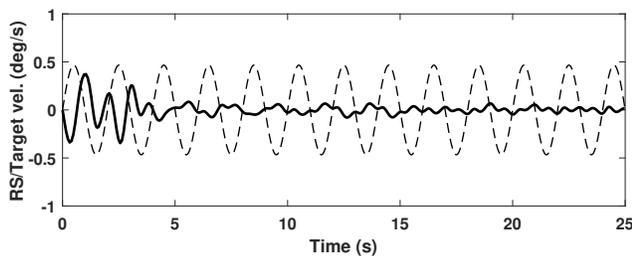


Fig. 8. Mean  $RS$  (the solid black trace) from 10 trials of the experiment shown in Figure 5a. The target velocity (the dashed line) is superimposed for reference. The  $RS$  signal is received by the predictive model after 80 ms to simulate sensory delays.

A puzzling aspect of smooth pursuit eye movement is that during zero lag tracking and occlusion, pursuit movement continues when  $RS$  is zero. Figure 8 depicts the  $RS$  signal at the retina (without delay) during pursuit of a sinusoidal target by our model. It can be seen that after the target motion pattern is learned, the  $RS$  is not exactly zero, but deviates by small amounts around zero. However, the small  $RS$  components cannot drive pursuit output during occlusion. We propose that the RNN is able to generate self-sustained eye velocity predictions. The small  $RS$  components are continuously used to correct the pursuit prediction. During occlusion, these corrective  $RS$  components are not available, and therefore, the pursuit eye velocity gradually lags behind the target, similar to experimental data [4]. In our model, the RNN operates in a chaotic regime and each neuron has its own baseline spontaneous activity. During occlusion, the learned neural activity pattern continues to produce the pursuit prediction.

For the control theoretic models without memory, a zero  $RS$  keeps the velocity prediction at a constant value. Therefore, these models can generate pursuit eye movement for a constant velocity ramp stimulus, but not for a sinusoidal target velocity stimulus during long occlusions. For memory based models [6], [8], occlusions larger than one cycle of a sinusoidal target velocity will result in incorrect pursuit

prediction. Moreover, [6] assumes that the brain knows the noise distribution of target velocity. The model by Shibata et al. [9] uses both  $RS$  and position error to run the Kalman filter based prediction, and the position error parameters are constantly updated to maintain almost zero  $RS$ . However, the model is not tested for long occlusions (maximum tested occlusion is 1/10th of a cycle).

#### *The role of FEF in predictive pursuit*

Many studies have observed a direct role of FEF in predictive smooth pursuit representation. Keating [11] found that lesions or ablations of FEF impaired monkeys' ability to conduct smooth pursuit of sinusoidal targets when the target was visible and during occlusions. Single unit recordings in FEF found neurons that continued to respond strongly after a sinusoidal target had been extinguished [14]. Fukushima et al. [13] found similar FEF neurons from recording studies on monkeys. Predictive pursuit eye movement signals in FEF were also observed during fMRI studies [33]. These studies suggest that FEF learns an internal model of the target velocity pattern to signal predictive pursuit eye movements, regardless of whether the target is visible or not.

FEF receives strong projections from the middle temporal area (MT) and the medial superior temporal area (MST), which are regions in the dorsal visual pathway that respond to object motion. Similarly, FEF connects to the lateral intraparietal area (LIP) with bidirectional connections [34]. LIP is upstream from MST in the visual pathway and has been associated with detection of object motion using bottom up attention signals [35]. On the output side, efferents from FEF are transmitted to dorsal pontine nuclei (PN) and reticularis tegmenti pontis (NRTP) regions in the brainstem. These brainstem regions relay information from FEF to cerebellum for oculomotor adaptation [36].

Based on the outcomes of lesion studies and anatomical connections, Keating [11] suggested that FEF is subsequent to the parietal areas of the visual pathway and prior to PN in controlling pursuit eye movements. This implies that FEF is a plausible neural correlate for our RNN, based on their common predictive activities during smooth pursuit and location on the pursuit pathway. Both FEF and our RNN use visual inputs to predict eye velocity commands for oculomotor controllers in cerebellum. Whereas, the leaky integrator can be realized in PN/NRTP as it relays FEF output to cerebellum.

## VI. CONCLUSION

Experiments on primates observed predictive eye velocity signals in FEF and suggested that an internal model of the target velocity is used to control eye movements instead of visual input [11], [14], [3]. However, it is not known how such an internal model is developed and updated rapidly by a neuron population in the brain.

Here we introduced a model of smooth pursuit using an RNN that rapidly learns an internal model of non-linear target velocity sequence and generates eye velocity for tracking with almost zero lag, when the target is visible and when the target

is occluded. Further, the proposed model is able to generate the typical acceleration pattern during pursuit initiation and adapt to unpredictable target velocity perturbation and phase shift in experimentally observed timescales [27], [10]. In this neural model of predictive smooth pursuit, we demonstrated that a population of recurrent neurons, continuously learning from a delayed retinal slip information, can generate both the pursuit initiation dynamics and the predictive pursuit behaviors observed in primates. The resulting model could have practical applications for tracking objects in video and autonomous systems.

#### ACKNOWLEDGMENT

This material is based upon work supported by National Science Foundation Award IIS-1302125, National Science Foundation under Grant No. 1652159 “CAREER: Scalable Neuromorphic Learning Machines”, Intel Corporation, and Korean Institute for Science and Technology.

#### REFERENCES

- [1] G. Barnes, S. Donnelly, and R. Eason, “Predictive velocity estimation in the pursuit reflex response to pseudo-random and step displacement stimuli in man.” *The Journal of Physiology*, vol. 389, no. 1, pp. 111–136, 1987.
- [2] R. J. Krauzlis and S. G. Lisberger, “A model of visually-guided smooth pursuit eye movements based on behavioral observations,” *Journal of computational neuroscience*, vol. 1, no. 4, pp. 265–283, 1994.
- [3] R. Eckmiller and M. Mackeben, “Pursuit eye movements and their neural control in the monkey,” *Pflügers Archiv European Journal of Physiology*, vol. 377, no. 1, pp. 15–23, 1978.
- [4] S. G. Whittaker and G. Eaholtz, “Learning patterns of eye motion for foveal pursuit.” *Investigative ophthalmology & visual science*, vol. 23, no. 3, pp. 393–397, 1982.
- [5] D. A. Robinson, J. Gordon, and S. Gordon, “A model of the smooth pursuit eye movement system,” *Biological cybernetics*, vol. 55, no. 1, pp. 43–57, 1986.
- [6] J.-J. O. de Xivry, S. Coppe, G. Blohm, and P. Lefevre, “Kalman filtering naturally accounts for visually guided and predictive smooth pursuit dynamics,” *Journal of Neuroscience*, vol. 33, no. 44, pp. 17301–17313, 2013.
- [7] D. Deno, W. Crandall, K. Sherman, and E. Keller, “Characterization of prediction in the primate visual smooth pursuit system,” *Biosystems*, vol. 34, no. 1-3, pp. 107–128, 1995.
- [8] G. Barnes and S. Wells, “Modelling prediction in ocular pursuit,” in *Current oculomotor research*. Springer, 1999, pp. 97–107.
- [9] T. Shibata, H. Tabata, S. Schaal, and M. Kawato, “A model of smooth pursuit in primates based on learning the target dynamics,” *Neural Networks*, vol. 18, no. 3, pp. 213–224, 2005.
- [10] A. Van den Berg, “Human smooth pursuit during transient perturbations of predictable and unpredictable target movement,” *Experimental Brain Research*, vol. 72, no. 1, p. 95, 1988.
- [11] E. Keating, “Frontal eye field lesions impair predictive and visually-guided pursuit eye movements,” *Experimental Brain Research*, vol. 86, no. 2, pp. 311–323, 1991.
- [12] D. Sussillo and L. F. Abbott, “Generating coherent patterns of activity from chaotic neural networks,” *Neuron*, vol. 63, no. 4, pp. 544–557, 2009.
- [13] K. Fukushima, T. Yamanobe, Y. Shinmei, and J. Fukushima, “Predictive responses of periarculate pursuit neurons to visual target motion,” *Experimental brain research*, vol. 145, no. 1, pp. 104–120, 2002.
- [14] M. G. MacAvoy, J. P. Gottlieb, and C. J. Bruce, “Smooth-pursuit eye movement representation in the primate frontal eye field,” *Cerebral Cortex*, vol. 1, no. 1, pp. 95–102, 1991.
- [15] S. G. Lisberger, E. Morris, and L. Tychsen, “Visual motion processing and sensory-motor integration for smooth pursuit eye movements,” *Annual review of neuroscience*, vol. 10, no. 1, pp. 97–129, 1987.
- [16] S. Yasui and L. Young, “On the predictive control of foveal eye tracking and slow phases of optokinetic and vestibular nystagmus.” *The Journal of physiology*, vol. 347, no. 1, pp. 17–33, 1984.
- [17] A. T. Bahill and J. D. McDonald, “Model emulates human smooth pursuit system producing zero-latency target tracking,” *Biological cybernetics*, vol. 48, no. 3, pp. 213–222, 1983.
- [18] C. Pack, S. Grossberg, and E. Mingolla, “A neural model of smooth pursuit control and motion perception by cortical area mst,” *Journal of Cognitive Neuroscience*, vol. 13, no. 1, pp. 102–120, 2001.
- [19] C. Zhang, Y. Zhao, J. Triesch, and B. E. Shi, “Intrinsically motivated learning of visual motion perception and smooth pursuit,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1902–1908.
- [20] C. Teulière, S. Forestier, L. Lonini, C. Zhang, Y. Zhao, B. Shi, and J. Triesch, “Self-calibrating smooth pursuit through active efficient coding,” *Robotics and Autonomous Systems*, vol. 71, pp. 3–12, 2015.
- [21] G. R. Barnes, “Cognitive processes involved in smooth pursuit eye movements,” *Brain and cognition*, vol. 68, no. 3, pp. 309–326, 2008.
- [22] H. Jaeger, “The echo state approach to analysing and training recurrent neural networks—with an erratum note,” *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, no. 34, p. 13, 2001.
- [23] W. Maass, T. Natschläger, and H. Markram, “Real-time computing without stable states: A new framework for neural computation based on perturbations,” *Neural computation*, vol. 14, no. 11, pp. 2531–2560, 2002.
- [24] D. Sussillo, T. Toyozumi, and W. Maass, “Self-tuning of neural circuits through short-term synaptic plasticity,” *Journal of neurophysiology*, vol. 97, no. 6, pp. 4079–4095, 2007.
- [25] M. Shidara, K. Kawano, H. Gomi, and M. Kawato, “Inverse-dynamics model eye movement control by purkinje cells in the cerebellum,” *Nature*, vol. 365, no. 6441, pp. 50–52, 1993.
- [26] S. de Brouwer, M. Missal, G. Barnes, and P. Lefèvre, “Quantitative analysis of catch-up saccades during sustained pursuit,” *Journal of neurophysiology*, vol. 87, no. 4, pp. 1772–1780, 2002.
- [27] C. Rasche and K. R. Gegenfurtner, “Precision of speed discrimination and smooth pursuit eye movements,” *Vision research*, vol. 49, no. 5, pp. 514–523, 2009.
- [28] M. Spering and K. R. Gegenfurtner, “Contextual effects on smooth-pursuit eye movements,” *Journal of Neurophysiology*, vol. 97, no. 2, pp. 1353–1367, 2007.
- [29] J. F. Medina and S. G. Lisberger, “Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum,” *Journal of neurophysiology*, vol. 102, no. 4, pp. 2039–2054, 2009.
- [30] S. J. Bennett and G. R. Barnes, “Smooth ocular pursuit during the transient disappearance of an accelerating visual target: the role of reflexive and voluntary control,” *Experimental brain research*, vol. 175, no. 1, pp. 1–10, 2006.
- [31] L. C. Osborne, S. G. Lisberger, and W. Bialek, “A sensory source for motor variation,” *Nature*, vol. 437, no. 7057, pp. 412–416, 2005.
- [32] H. Collewijn and E. P. Tamminga, “Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds.” *The Journal of Physiology*, vol. 351, no. 1, pp. 217–250, 1984.
- [33] R. Lencer, M. Nagel, A. Sprenger, S. Zapf, C. Erdmann, W. Heide, and F. Binkofski, “Cortical mechanisms of smooth pursuit eye movements with target blanking. an fmri study,” *European Journal of Neuroscience*, vol. 19, no. 5, pp. 1430–1436, 2004.
- [34] T. J. Buschman and E. K. Miller, “Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices,” *science*, vol. 315, no. 5820, pp. 1860–1862, 2007.
- [35] R. A. Andersen, “Visual and eye movement functions of the posterior parietal cortex,” *Annual review of neuroscience*, vol. 12, no. 1, pp. 377–403, 1989.
- [36] M. Takagi, D. S. Zee, and R. J. Tamargo, “Effects of lesions of the oculomotor cerebellar vermis on eye movements in primate: smooth pursuit,” *Journal of Neurophysiology*, vol. 83, no. 4, pp. 2047–2062, 2000.