

Neuromorphic Control of Gaze Direction

Omid Sadat Rezai, Bryan Tripp

University of Waterloo

200 University Ave., Waterloo, Canada

omid.srezai@uwaterloo.ca; bptripp@uwaterloo.ca

Abstract - Neurobiological control systems are of practical interest because animals interact proficiently with novel, unstructured environments. This study investigates a model of neurobiological control of eye movements. We re-implement an abstract control model of smooth-pursuit (tracking) eye movements as a network of spiking neurons. We find that the behaviour of the neural model is similar to that of the abstract model, but there are several differences that seem to be inherent in the neural implementation, including greater lag and the introduction of point attractors. We also describe a preliminary integration of the model with a pan-tilt camera. Performance of different variations of the model suggests the importance of covert visual attention in this system.

Keywords: tracking, smooth pursuit, neural engineering framework, vision.

1. Introduction

The control systems that guide animal motion are quite different from those that guide robots, and are more effective in natural environments. Neurobiological systems have been studied and recorded intensively for several decades, but so far there has been limited success in reverse-engineering superior control strategies from these systems. Detailed investigation of specific brain areas often reveals surprisingly simple computations, suggesting that the robust and adaptive behaviour of the brain arises in part from its high-level architecture. This presents a barrier to the application of brain-inspired approaches to engineering control problems, in that in order to understand why neurobiological control works so well, it may be necessary to model extensive control networks. In this context, the network that controls eye movements in humans and other primates is of particular interest, because all of its parts are relatively well-studied. Eye movements are also important for vision, in that they allow most computational resources of the visual cortex to be concentrated on a series of small regions, whereas parallel analysis of the whole scene in such detail would require a much larger brain. In this work, we focus on smooth pursuit eye movement, which involves visual tracking of a moving object in the environment (in contrast with saccades, which are rapid step-like shifts in gaze direction). Our main contribution is the re-implementation of an existing high-level control model of smooth pursuit eye movements (Churchland & Lisberger, 2001), which has been validated against monkeys. Our implementation is composed of spiking neurons, i.e. neurons that communicate through 1ms pulses (spikes) in cell-membrane potential. Also, the original model was developed with scalar retinal velocity as input, which avoids the issue of obtaining this signal from the eye. As a step in addressing this point, we present initial results from embedding the model in a loop with a pan-tilt camera.

2. Methods

Our starting point is the smooth-pursuit model of Churchland & Lisberger (2001), which focuses on the dynamics of feedback control in this system. Specifically, it assumes that the visual cortex accurately

encodes the velocity of the pursuit target, and fits a system of transfer functions and static nonlinearities to the monkeys' eye velocities in response to various moving stimuli. The structure of this control model is based loosely on the anatomy of the neural system.

A key limitation of the control model is that it does not contain any neurons, but only abstract dynamic blocks. Our main goal was therefore to approximate this abstract control model as a system of spiking neurons, and to verify that such an approximation behaves similarly to the abstract model (and the monkeys). For this purpose, we used the Neural Engineering Framework (NEF; Eliasmith & Anderson, 2003), the modelling framework that was recently used to build the world's most functionally sophisticated large-scale brain model (Eliasmith et al., 2012). The NEF provides a way of translating a high-level system description, in the form of a system of explicit nonlinear ordinary differential equations, into a network of spiking neurons that has similar behaviour. The dynamics of the resulting neural model arise from low-pass filtering properties of connections between neurons (synapses). Biological synapses are low-pass because an incoming spike opens a large number of ion channels in the cell membrane of the receiving neuron, which then close at random with exponential continuum dynamics. An ensemble of neurons in an NEF model has correlated spike rates, which are nonlinear functions of a state vector x . The ensemble is said to "encode" x . A connection from one ensemble to another communicates some function $y = f(x)$, which corresponds to the synaptic weights of connections between individual neurons. These weights are determined analytically from $f(x)$, as $w_{ij} = e_i^T d_j$, where w_{ij} is the weight of the synaptic connection from neuron i in the sending ensemble to neuron j in the receiving ensemble, d_j (the "decoder") are coefficients in an optimal linear reconstruction of $f(x)$ from the spike rates, and e_i (the "encoder") projects x into the space of the net synaptic current flowing into post-synaptic neuron j .

Implementation of the control model in neurons involved decoding nonlinear functions from neural ensemble activity, and constructing feedback networks to approximate low-pass and band-pass filters and integrator components of the control model. A low-pass filter occurs naturally in the connection from one ensemble to another, but it has the time constant of synaptic dynamics, a molecular property that is most typically in the 5-10ms range. We added feedback connections in order to match a 55ms time constant and a pure integrator in the control model, using methods described by Eliasmith & Anderson (2003). We used a network of two interconnected ensembles to approximate a second-order band-pass filter with a .4ms time constant, based on a network in Tripp & Eliasmith (2012).

We also tested the model as a controller for a pan-tilt camera (Figure 1B). We captured frames from the camera at 25Hz, estimated optical flow, extracted average velocity from a central region, interpolated to obtain a .001s time step, and used this as input to the controller. The controller output was then used to drive pan and tilt servos. We calculated optical flow using a real-time GPU implementation of pyramidal algorithm of Lukas & Kanade (Marzat et al., 2009). This algorithm solves the aperture problem (i.e., the direction of motion of an edge is ambiguous given only local information). We used this velocity field as an approximation of neural activity in area MT of the primate cortex, which also solves the aperture problem (Pack & Born, 2001) through as-yet unclear mechanisms.

3. Results

Figure 2 shows example simulations that compare the Churchland et al. (2001) control model with our spiking neural model. We implemented our neural model so it had the same low-pass and band-pass filter time constants and the same nonlinear saturations. However, three major factors contributed to some differences (see Figure 2). First, the decoded output of the neural model contains high-frequency noise that is related to the brevity of spikes relative to other timescales in the system (Eliasmith & Anderson, 2003). We have some control over the amplitude of this noise, as it decreases with increasing numbers of neurons in the model. To make the simulations tractable our model included far fewer neurons (i.e., 9 thousand) than

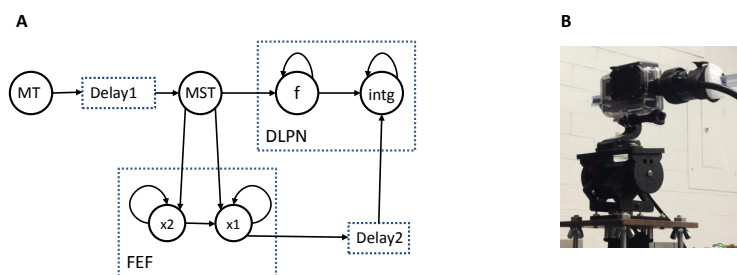


Fig. 1. A, High-level structure of the neural model, showing connections between neural ensembles and brain areas to which they are meant to correspond (see Mustari et al., 2009, for more information about the roles of these areas). MT = middle temporal area of the visual cortex; MST = middle superior temporal area; FEF = frontal eye fields; DLPN = dorsolateral pontine nucleus. FEF and DLPN are subdivided into multiple ensembles in order to approximate the required dynamics: x_1 and x_2 encode different state variables in a band-pass filter; f encodes the state of a low-pass filter; $intg$ encodes the state of an integrator. B, The pan-tilt camera.

the corresponding biological systems (i.e., millions), but we did not find that this noise substantially affected control. Second, there are small differences between the mean output of each ensemble for a given input and the ideal output. In a network with feedback, these differences can create attractors that produce steady-state errors in the output (Eliasmith & Anderson, 2003). We observed this effect in our models as steady-state error which depended on randomly-selected neuron parameters (including maximum firing rates, etc.). These errors were particularly pronounced with signals that were much smaller than the neurons' operating ranges (see example in Figure 2C). Thirdly, our neural model contains an additional pole that corresponds to synaptic dynamics in the connection between the MT and MST ensembles. Nothing in the original control model reflected these inevitable dynamics. This difference increases phase lag and causes our model to oscillate at somewhat lower gains relative to the control model (not shown).

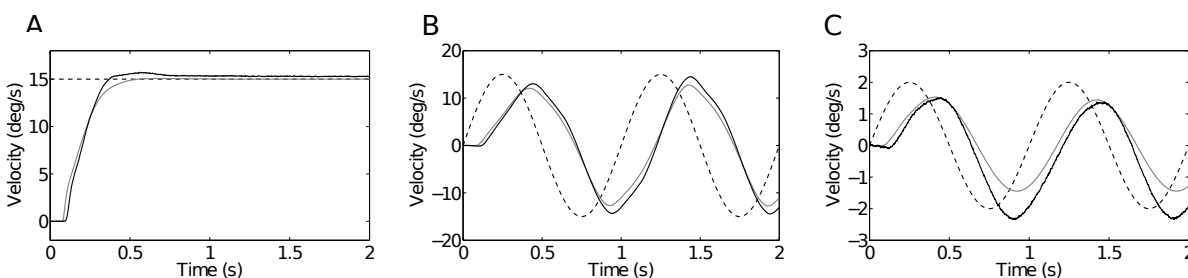


Fig. 2. Comparison of Churchland & Lisberger (2001) control model with our neural model. Dashed lines are input (visual velocity). Light solid lines are the output of the control model, and dark solid lines are the output of the neural model, which is an optimal linear estimate of the state of the integrator ensemble from the ensemble's spike responses. A, Step input. B, Sinusoidal input. C, An example of smaller sinusoidal input that illustrates non-ideal behaviour of the neural model (note difference in scale vs. B).

3.1. Integration with Pan-Tilt Camera

We integrated the control model with a pan-tilt camera and tested the system's ability to track a pendulum, with minimal additional processing of the velocity field. The system clearly distinguished pendulum from background, and moved in the same direction as the pendulum, partially tracking it. However unsurprisingly,

tracking was not accurate or robust because whenever the pendulum was not well centred in the image due to (realistic) lag between target and camera, the image-velocity estimate averaged over parts of both the pendulum and the background. We then added an attention-like mask (based on colour) so that velocity was averaged only over the pendulum (excluding the background). This greatly improved the results. This seems to be an essential feature that has typically been omitted from neurobiological models of smooth pursuit.

4. Conclusion

Animals operate much more effectively in natural environments than current robots, so it makes sense to examine neurobiological mechanisms of control. Notably, monkeys do not track visual motion very precisely, and they exhibit large lags. This suggests that accurate visual pursuit is not critically important for excellent visual perception. However, there may be other features of the dynamics that are important. Our main current result is that the abstract model of smooth pursuit dynamics proposed by Churchland & Lisberger (2001) is essentially consistent with a spiking neural implementation. This is a first step toward developing higher-fidelity neural models that are validated more extensively against neural activity in the corresponding brain regions. In the future, we hope to understand the important dynamic features in more detail by developing increasingly realistic neurobiological models. There are several obvious directions for future expansion. It is critical for practical operation of a camera to add a mechanism to ensure that only the image velocity of the target (excluding the background) is fed into the controller. We used colour information for this purpose. This could also be done by clustering velocities, finding the target by matching intensity histograms, etc. but more work is needed to explore biologically plausible mechanisms. This is an interesting topic that is likely to be related to neural mechanisms of covert attention (i.e., preferential processing that is not strictly dependent on retinal location), which is intimately related to eye movements (Moore & Fallah, 2001). Robust smooth pursuit may be an important application of covert attention. Another important extension would be to add catch-up saccades, which animals use when smooth pursuit fails, e.g. because the target is moving too fast (De Brouwer et al, 2002).

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

- Churchland, M. M., & Lisberger, S. G. (2001). Experimental and computational analysis of monkey smooth pursuit eye movements. *J Neurophysiol*, 86, 741759.
- De Brouwer, S., Missal, M., Barnes, G., & Lefvre, P. (2002). Quantitative analysis of catch-up saccades during sustained pursuit. *J Neurophysiol*, 87(4), 177280.
- Eliasmith, C., & Anderson, C. (2003). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. MIT Press.
- Eliasmith, Chris, Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science (New York, N.Y.)*, 338(6111), 12025.
- Marzat, J., Dumortier, Y., & Ducrot, A. (2009). Real-time dense and accurate parallel optical flow using CUDA. In *WSCG*.
- Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *PNAS*, 98(3), 12736.
- Mustari, M. J., Ono, S., & Das, V. E. (2009). Signal processing and distribution in cortical-brainstem pathways for smooth pursuit eye movements. *Ann N Y Acad Sci*, 1164, 147154.
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 409(6823), 10402.

Tripp, B. P., & Eliasmith, C. (2010). Population models of temporal differentiation. *Neural computation*, 22(3), 62159.