Recurrent Neural Network Generates a Basis for Sensory Image Cancellation

Patrick D. Roberts

Neurological Sciences Institute
Oregon Health Sciences University
Portland, OR 97209
robertpa@ohsu.edu

Abstract

This study investigates the temporal dynamics of recurrent layers and their relevance to storage of temporal information. A recurrent layer is shown to generate a dynamical basis that allows cancellation of predictable sensory images via and adaptive mechanism based on spike-timing dependent plasticity.

Key words: recurrence, STDP, mormyrid, electrosensory

1 Introduction

Recurrence is ubiquitous in the nervous system, and many sensory processing structures adapt to the temporal pattern of repeated sensory signals to improve the detection of novel input. However, the contribution of recurrence to information coding of spike trains is not well understood. Without an understanding of how recurrence affects the storage and retrieval of temporal information will limit our exploration of the functional consequences of the nervous system to only simple, feed-forward dynamics.

Electrosensory system of mormyrid electric fish. The sensory processing systems that we have investigated is the primary electrosensory processing system of mormyrid electric fish. The mormyrid electric fish senses its environment by emitting an electric organ discharge (EOD) and detecting the perturbations that nearby objects cause in the self-generated electric field. The fish detects low-frequency, externally generated electric fields. Afferent fibers from electroreceptors respond to the fish's own EOD, and these responses are conveyed to the cortex of electrosensory lateral line lobe (ELL)

where the fibers terminate in separate mormyromast and ampullary zones [1]. The best-understood function of the ELL is to cancel predictable electrosensory signals so the the fish is exquisitely sensitive to novel electrical activity in its environment.

With each EOD, the ELL cortex is affected not only by input from the periphery but also by electric organ *corollary discharge* (CD) signals that originate centrally. These corollary discharge signals are time-locked with the EOD motor command, which, elicits the EOD. An advantage of studying the ELL of mormyrid fish is that the information carried by the centrally generated signals is well understood; the corollary discharge signals inform the ELL of *when* an EOD has taken place. The corollary discharge signals to the ELL arise from various central structures. One of these central structures is the eminentia granularis posterior (EGp) which gives rise to the parallel fibers of ELL (Fig. 1A). EGp receives corollary discharge signals at various delays out to about 80 ms [3]. Additional delays are probably added within EGp due to recurrence from inhibitory Golgi cells. The EGp also gives rise to parallel fibers that innervate the caudal lobe of the mormyrid cerebellum. Thus, understanding how EGp processes information will lead directly to hypotheses about information processing in the granule cell layer of the cerebellum.

Parallel fibers synapse onto apical dendrites of MG cells in the molecular layer of the ELL. MG cells are inhibitory (GABAergic) Purkinje-like interneurons, which synapse locally onto the efferent cells. In addition to the apical dendrites that receive parallel fiber input, MG cells also have basal dendrites that receive electrosensory input from the periphery. Extra- and intracellular recordings *in vivo* have shown that the responses of MG cells to corollary discharge are plastic [4]. Pairing the corollary discharge with an electrosensory stimulus for a few minutes results in the development of a response to the corollary discharge alone that is a negative image of previously paired responses to the sensory stimulus. The negative image is a kind of prediction about expected sensory input that is specific to the time course of the sensory response and its polarity (excitation or inhibition). The ELL acts as an adaptive sensory processor to predict or to expect certain sensory input. The negative image of expected sensory input is subtracted from actual sensory input, allowing unexpected input to stand out [5]. Anti-Hebbian plasticity at the synapses from parallel fibers onto MG cells has been demonstrated in the *in vitro* slice preparation [6], and such plasticity can explain the generation of negative images as observed *in vivo*. We show how the dynamics of recurrence in the EGp alone can be used by synaptic plasticity to generate the negative image.

Spike-timing dependent synaptic plasticity (STDP). The plasticity found at the synapses from parallel fibers onto MG cells is STDP [7,6]. This means that the size and direction of synaptic changes depend on the exact timing between pre- and postsynaptic spikes. MG cells have 2 types of spikes:

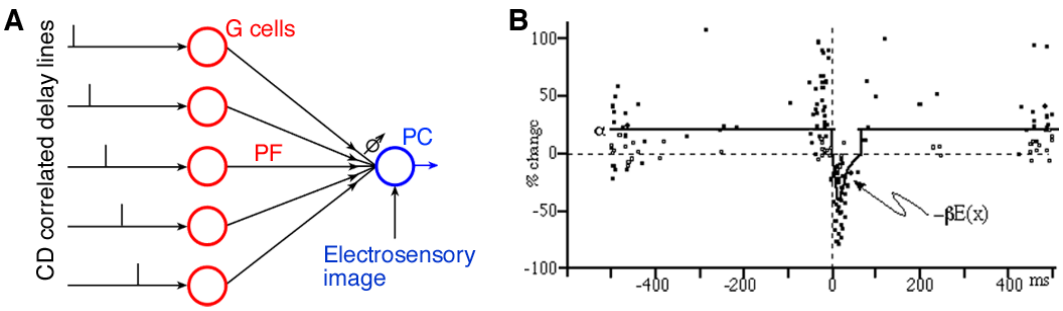

Fig. 1. *Local circuitry of the electrosensory lateral line lobe (ELL).* (A) The electric organ corollary discharge (CD) signals enter the ELL through the eminentia granularis posterior (EGp) that contains granule cells (G cells) giving rise to parallel fibers. The granule cells are assumed to respond to the correlated discharge at various delays so that the parallel fibers generate a series of delayed spikes correlated with each electric organ discharge. Parallel fibers (PF) synapses excite Purkinje-like cells (PC), the medium ganglion cells. Primary afferent fibers from electrosensory receptors enter the ELL and transmit temporal sensory pattern to PC. STDP at the PF synapse onto PC causes the output to adapt to the temporal pattern of the electrosensory image. (B) Data from [6] showing STDP at the PF synapse onto PCpercent change in excitatory postsynaptic potential (EPSP) amplitude as a function of delay between onset of EPSP and broad spike peak. If a dendritic spike in a MG cell follows the arrival of a parallel fiber spike within 40 msec, then the synapse is depressed proportionally to the excitatory postsynaptic potential; otherwise the synapse is enhanced.

(1) small, narrow axon spikes that do not invade the soma; and (2) large, broad soma-dendritic spikes that propagate into the apical dendrites. The threshold of the broad (dendritic) spike is 1.2 to 1.6 times higher than the threshold of the small spike [8]. It is the dendritic spike that participates in the STDP of the parallel fiber synapse [4,6]. Depression of the excitatory postsynaptic potential (EPSP) has been observed only after pairings in which the dendritic spike was evoked between 0 and 60 msec after the onset of the parallel fiber EPSP, while pairings at other delays yielded potentiation (Fig. 1).

Modeling studies have helped elucidate how the learning rule measured *in vitro* can account for the plasticity of ELL cells measured *in vivo* at the systems level [9,10,12]. Not only did the model generate adaptive responses that were similar to those observed *in vivo*, but the STDP learning rule measured *in vitro* appeared to be optimal for generating such adaptive responses.

However, a critical component of the model is that the the parallel fibers must act as delay-lines with respect to the CD. Otherwise, a negative image cannot be formed by the parallel fiber inputs. In addition, the spike activity of granule cell in EGp have never been recorded *in vivo* due to the difficulty of recording from such small cells. Thus, it is not known if this critical element of the model is present in this system. In addition, other cerebellum-like structures, and the cerebellum, have no known series of behaviorally important inputs,

but recurrence in their granule cell layer could generate temporal information about sensory stimuli. In the proposed research, we will extend our previous model to investigate how recurrence in the EGp could generate spike patterns in parallel fibers that allow the generation of a negative image.

2 Results

To test our hypothesis that recurrent connections in the layer of granule cells of ELL could lead electrosensory image cancellation, we conducted numerical studies of the network shown in Fig. 2A. The network consisted of a granule cell layer that were connected to the Purkinje-like cell. Each model neuron was non-spiking and interacted through continuous synapses, synapses that transmitted the spike probability weighted by the synaptic weight, w_{ij}. The spike probability was calculated from the membrane potential, $V_i(x_n, t)$, that was the sum of all inputs, and x_n denoted the time within each cycle and t denoted the cycle. The spike probability function was defined as in our spiking models [9,13], $P_i(x,t) = (1+\exp(-\mu(V_i(x,t)-\theta)))^{-1}$, where θ is the threshold and μ parametrizes the noise. This function has a sigmoid form and a value of $1/2$ at $\bar{V}(x,t) = \theta$.

The granule cell received a small depolarizing input to represent a synchronized spike that initiated the recurrent dynamics. The driver input was delivered between 10-30 time-steps. The 40 granule cells were interconnected by weights with a weight matrix defined by, $w_{i,i-1} = w_{i,i+1} = w_{i,i+2} = 3$, and all other weights were set to zero. These weights were not varied during adaptation.

Each granule cell was connected to the Purkinje-like cell by an adapting synapse. The Purkinje-like cell also received a non-adapting "sensory" input to mimic an electrosensory afferent modulation. This input was paired with the driver input during adaptation. The rule of adaptation was chosen to be a continuous analogy to the STDP learning rule shown in Fig. 1B. Since synaptic weight change only occurs in the STDP learning when there a presynaptic spike, we would increment the weight only if $P_{pre}(x,t) > 0.5$. The STDP learning rule depresses the synapses when there is a postsynaptic dendritic spike, and potentiates the synapse otherwise. Therefore, we chose to increment in the synaptic weights with the rule: $\triangle w = -1$ if $P_{post}(x,t) > 0.5$ and $\triangle w = +1$ if $P_{post}(x,t) < 0.5$.

We measured adaptation of the Purkinje-like cell to a temporal patterned input that is correlated with a single pulse to the granule cells by calculated the variance of the Purkinje-like cell's membrane potential, $\mathrm{var}(V_{pc}, t) = \sum_n (V_{pc}(x_n, t) - v(t))^2$, summed over all time-steps in each cycle, and $v(t)$ is

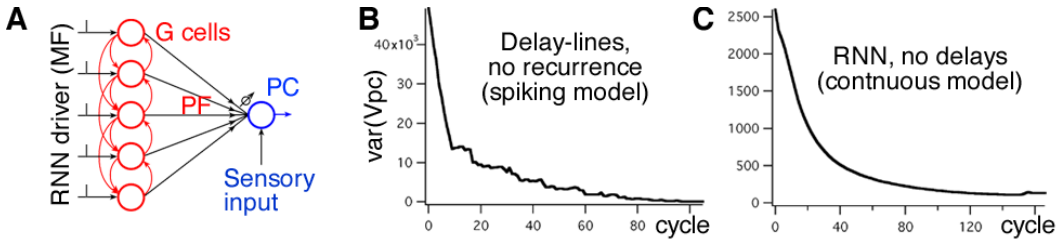

Fig. 2. *Adaptation of RNN output without delay-line input.* (A) RNN model of ELL where there are recurrent connections between the granule cells (G). The mossy fiber inputs (MF) are synchronized so that there are no delay lines, all RNN driver inputs are synchronized. (B) Variance of the membrane potential of the output neuron during adaptation shows cancelation of the sensory image for the model in Fig. 1A. At cycle=0, the Purkinje-like neuron is in the state portrayed in Fig. **??**A, and at cycle=100, the Purkinje-like neuron is in the state portrayed in Fig. **??**B. (C) Variance of the membrane potential for the model in (A) also shows adaptation, even though there are no delay-line inputs.

the average of the membrane potential during the cycle. The results are compared in Fig. 2 with are previous spiking model that relayed on the series of delay-lines for input. We found that delay lines were not necessary for sensory image cancellation if there is recurrent connections within the input layer.

3 Conclusion

Our results have forced us to reconsider our previous assumptions about how the ELL uses timing information and STDP to sculpt a negative image of predictable sensory inputs. Our previous requirement for delay-lines can now be relaxed, and this principle generalizes to other systems. The important aspect is that the parallel fibers do not individually carry a series of delayed spikes, but there is a linear combination of parallel fiber signals, that is a fixed point of the STDP learning rule, that combines the parallel fiber activity into temporal pattern that cancels the sensory input. Another important remark is that the details of the recurrent synaptic weights are not important for the result. The cancellation of the sensory image is robust under a wide range of synaptic weights within the recurrent layer. The weights do not even have to be fixed, as long as of their changes occur on a longer time-scale that the STDP learning rule at parallel fiber synapses onto the Purkinje-like cell.

Acknowledgments

I would like to thank Roberto Santiago for introducing me to recurrent learning networks. This material is based upon work supported by the National

Institutes of Health under Grant No. R01-MH60364 and by the National Science Foundation under Grant No. PHY99-07949.

References

[1] C. C. Bell, Mormyromast electroreceptor organs and their afferent fibers in mormyrid fish. II. Intra-axonal recordings show initial stages of central processing, J. Neurophysiol. 63 (1990) 303–318.

[2] C. C. Bell, K. Grant, J. Serrier, Sensory processing and corollary discharge effects in the mormyromast regions of the mormyrid electrosensory lobe: I. Field potentials and cellular activity in associated structures, J. Neurophysiol. 68 (1992) 843–858.

[3] C. C. Bell, A. Caputi, K. Grant, J. Serrier, Storage of a sensory pattern by anti-Hebbian synaptic plasticity in an electric fish, Proc. Natl. Acad. Sci. USA 90 (1993) 4650–4654.

[4] C. C. Bell, D. Bodznick, J. Montgomery, J. Bastian, The generation and subtraction of sensory expectations within cerebellum-like structures, Brain. Beh. Evol. 50 (1997) 17–31, (suppl.1).

[5] C. C. Bell, V. Han, Y. Sugawara, K. Grant, Synaptic plasticity in a cerebellum-like structure depends on temporal order, Nature 387 (1997) 278–281.

[6] L. F. Abbott, S. B. Nelson, Synaptic plasticity: taming the beast, Nature Neurosci. (suppl.) 3 (2000) 1178–1183.

[7] K. Grant, Y. Sugawara, L. Gomes, V. Han, C. C. Bell, The Mormyrid electrosensory lobe *in vitro*: Physiology and pharmacology of cells and circuits, J. Neurosci. 18 (1998) 6009–6025.

[8] P. D. Roberts, C. C. Bell, Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation, J. Compu. Neurosci. 9 (2000) 67–83.

[9] P. D. Roberts, Dynamics of temporal learning rules, Phys. Rev. E 62 (2000) 4077–4082.

[10] A. Williams, P. D. Roberts, T. K. Leen, Stability of negative-image equilibria in spike-timing-dependent plasticity, Phys Rev E Stat Nonlin Soft Matter Phys 68 (2-1) (2003) 021923.

[11] P. D. Roberts, Recurrent biological neural networks: The weak and noisy limit, (in press) *Phys. Rev. E* (arXiv.org:cond-mat/0305515) (2004).