Recognition of Temporal Event Sequences by a Network of Cortical Neurons

Maciej T. Lazarewicz, Sandhitsu Das, Leif H. Finkel
Department of Bioengineering
University of Pennsylvania, Philadelphia, PA 19104

Abstract

Recognition of ordered sequences of temporal events is central to many perceptual recognition tasks, from speech detection to analysis of biological motion. We describe a simple cortical network capable of recognizing event sequences through a process of encoding followed by detection. The network is composed of regular spiking and fast spiking neurons, with minimal connectivity. Ordered sequences of inputs occurring over tens-to-hundreds of milliseconds, are time-compressed by the network into tightly clustered spike outputs occurring over a few milliseconds. We investigate the ability of the network to accurately encode the input pattern, in the presence or absence of noise. We show that information about relative input timings are preserved in the output interspike intervals.

Keywords: Visual cortex, Biological motion detection, Spike-based computation

Introduction

Many instances of perceptual recognition involve the detection of an ordered sequence of events occurring on a time scale of tens-to-hundreds of milliseconds. Individual phonemes in spoken speech, or vocalizations in bird song, must occur in the correct order, and within certain time constraints to produce a recognizable word (or song fragment). We are particularly interested in the visual events that characterize biological motions, such as walking. Each walking cycle (~1 Hz) consists of an ordered sequence of events—starts and stops of motions of various limbs in various directions [5]. Part of what defines "walking" and distinguishes it from similar motions (running, skipping, etc.) is this spatiotemporal pattern of events.

The time scale over which these perceptual events occur poses a challenge for any cortical recognition mechanism. Information must be accumulated over tens or hundreds of milliseconds—enough time to observe several "events" (under some experimental conditions, psychophysical experiments show recognition of point-light walkers in 100-200 ms [6]). But cell time constants are on the order of 10 ms. Any mechanism must also be, to some degree, time-scale invariant [3] (to allow recognition of faster or slower walking), and stride-length invariant.

We describe simulations of a small network of biophysically-realistic cortical neurons that is capable of detecting the occurrence of the correct order of events.

Model Description

The network is composed of four regular spiking (RS) cortical neurons and four fast spiking (FS) interneurons. The cell models were taken from Destexhe, as posted on the SenseLab database [7]. Both cells are single compartment models expressing Na_t and K_{DR} channels, and additionally in the RS cell, K_M. AMPA and GABA synapses were taken from [1]. All simulations were carried out in Neuron [2].

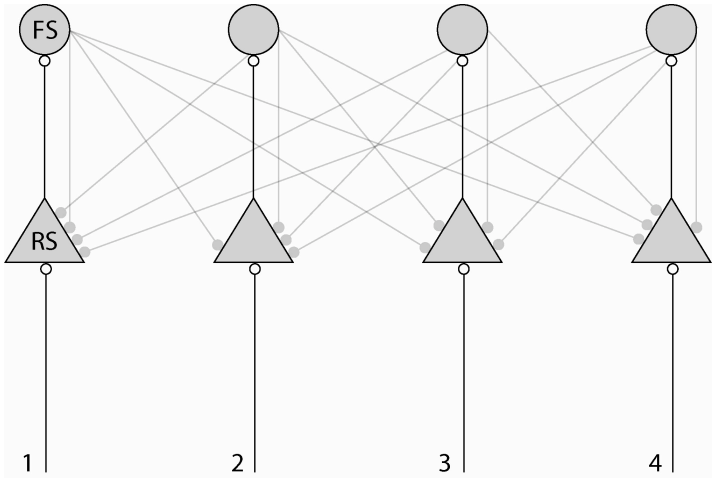

Figure 1. Each RS cell excites its associated FS cell, which in turn inhibits all of the RS cells. Inputs are numbered 1-4 for reference.

As shown in Figure 1, each RS cell excites its associated FS cell, which in turn inhibits all of the RS cells. There are no connections between RS cells or between FS cells. In other simulations, not described here, we investigated the role of these additional intracortical connection classes.

Each RS cell receives an input, delivered as an injection current. Each input represents the output of lower cortical centers corresponding to a particular "event", and is simulated by a linearly increasing current ramp. In all simulations described here, the ramp increased from 0-10 nA over 2 seconds. Although this is depolarizing current, it could reflect the net difference between total excitatory and inhibitory inputs to the RS cell (see discussion for possible mechanisms). Each RS cell receives a separate input injection, which begins at a certain time and continues for 2 seconds. We investigated the response of the network to input streams in which the cells were activated in various orders, in which the time delay between onset times was varied, and to which noise was added.

Results

Simulations reveal that the network transforms these ordered, continuous input streams into clusters of tightly ordered spikes. Figure 2A shows the output of the 4 RS cells in response to the input sequence 1-2-3-4. In response, each RS cell fires periodically at roughly 40 Hz (this rate depends on current injection amplitude), however, the spikes of all 4 RS cells are tightly clustered. Inspection of the spike order shows that the order of RS cell firing exactly corresponds to the order of the inputs (1-2-3-4). Figure 2B shows that the network reproduces the input sequence in spike order for all permutations of input onsets.

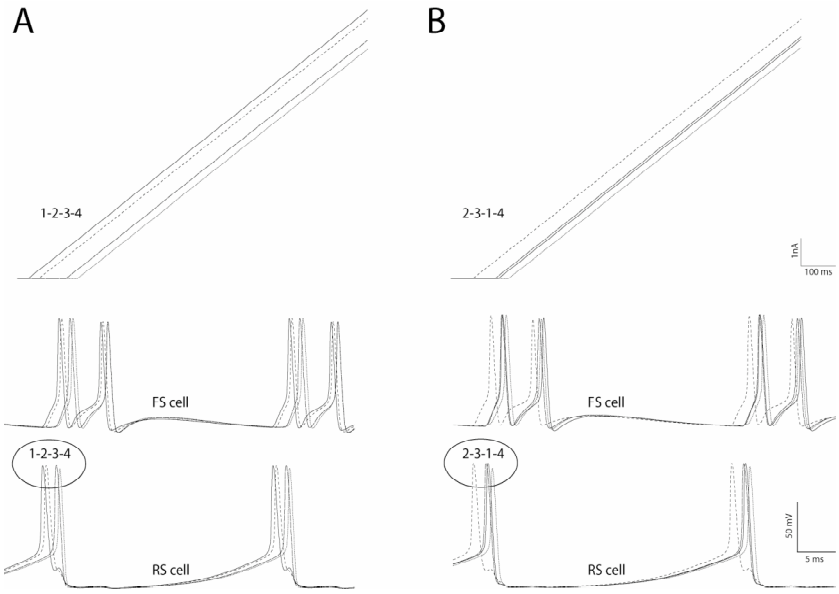

Figure 2. The output of the 4 RS cells in response to the input sequence 1-2-3-4 (A) and 2-3-1-4 (B)

Closer inspection of the interspike intervals between RS cell spikes shows a rough, quantitative preservation of the relative onset times of the inputs. Figure 3 shows results of 50 simulations in which the onset times of each input were varied over a range of 80 ms. Input 1 commenced at a randomly chosen time between 10-90 ms (after start of the simulation), input 2 commenced between 110-190 ms, input 3 between 210-290, and input 4 between 310-390 ms. Thus, the order of the inputs was constant (1-2-3-4) but the onset times varied significantly. Figure 3 shows histograms that compare the relative amplitudes of the inputs to the relative interspike intervals. The abscissa plots the ratio of the differences between inputs (A, 2^{nd} and 1^{st} to 3^{rd} and 2^{nd}; B, 3^{rd} and 2^{nd} to 4^{th} and 3^{rd}) divided by the ratio of the resulting interspike intervals. Because the inputs are linear and starts from zero, proportions between inputs are equivalent to proportions between the onset times. A value of 100 would indicate perfect preservation of the time of event information. Values for these simulations ranged between 90-120%, indicating a substantial amount of quantitative information is preserved.

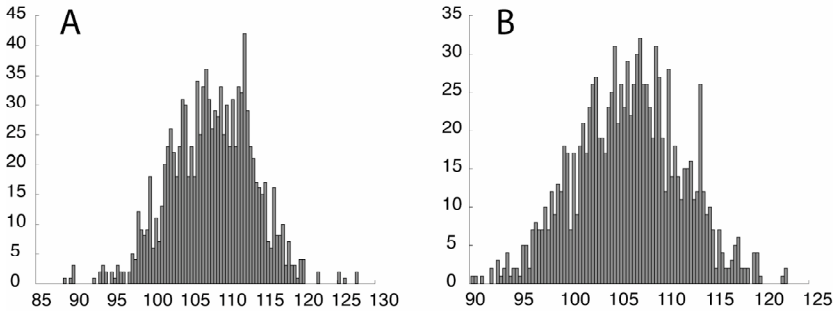

Figure 3, Histograms comparing the ratio of the differences between input amplitudes (A, 2^{nd} and 1^{st} to 3^{rd} and 2^{nd}; B, 3^{rd} and 2^{nd} to 4^{th} and 3^{rd}) to the ratio of the resulting interspike intervals. 100% would indicate perfect quantitative match. Simulation run without noise.

One would expect these results to be affected by the presence of noise. To test this assumption, every 3 ms we added an independent Gaussian random noise component

to each current injection (mean = 0 nA; σ=0.001 nA^2). This means that the ratio of inputs coming to RS cells will be distorted every 3 ms compared to the original values. Actual values for this distortion are given in figures 4C and 4D. The majority of input ratios are distorted between 40-160%. However, distortion of the output spikes is significantly less, in the range 80-140% (figure 4A and 4B). Which means that the circuit is able to provide reliable information about structure of the inputs even in the presence of noise. This possibly results from effective averaging of the white noise by the cell time constant.

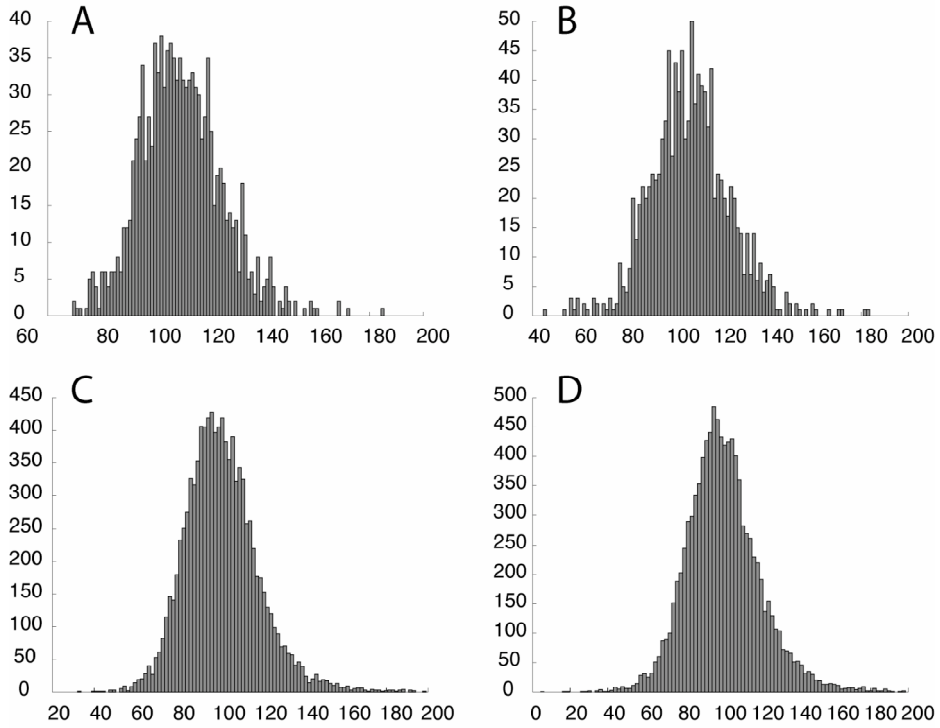

Figure 4. Simulations run with added noise. A, B - histograms of the input/output ratios, as in Figure 3. C, D - Measuring the distortion in inputs due to added noised. Histograms comparing the ratio of inputs without noise to the ratio of inputs with noise

Since the 4 RS cells receive significantly different levels of current injection, one might expect them to fire at different rates. The reason they fire synchronously (but with a critical few millisecond ISI) at the same rate is due to inhibition from the FS interneurons. As shown in figure 5, elimination of inhibition (as might occur with biccuculline), releases each RS cell to fire at a significantly different rate, in a non-correlated manner. The network contains no FS-FS connections or RS-RS connections to generate synchrony, and inspection of figure 2 reveals that the FS cells do not begin to fire until at least the 3rd RS cell has fired, thus they cannot be regulating the interspike intervals. Therefore, it appears, that the RS cells are firing as their current ramps bring cell voltages to threshold, that threshold is reached roughly 2 ms later in each subsequent RS cell (1-2-3-4), and that firing of the FS cells (each in turn) is responsible for the roughly 20 ms delay between spike clusters. Note that without this temporal gap, it would be difficult to distinguish spikes in one cluster from those in the next. In summary, the order and timing of the RS cell spikes directly follows from the relative levels of current injection.

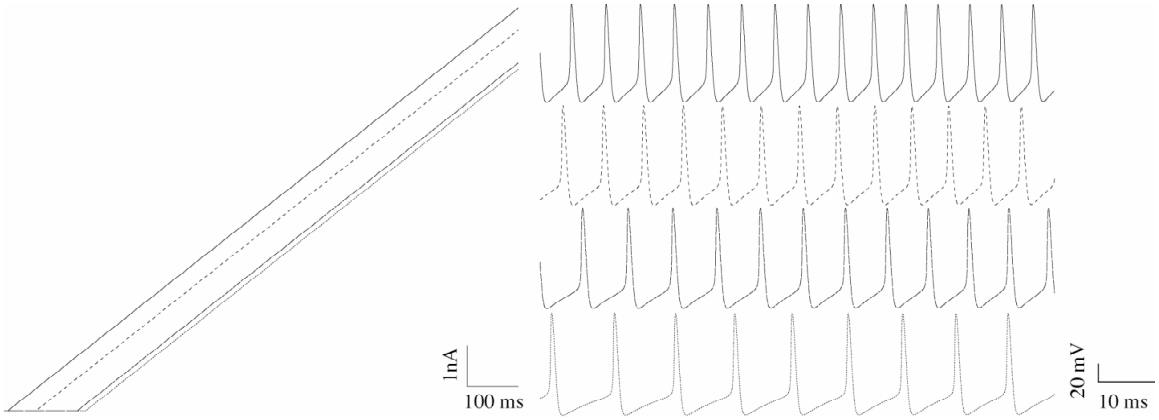

Figure 5. Elimination of inhibition releases each RS cell to fire at a significantly different rate, in a non-correlated manner

Discussion

We have shown that a profoundly simple network architecture (2 cell types, one class of cell-cell connection) can translate an ordered sequence of inputs into an ordered sequence of cell firings, preserving (to some degree) the quantitative ratio of input timings. The advantage of such an analog-to-digital type conversion is the time compression. Events which transpired over tens or hundreds of milliseconds are now represented in a set of inter-spike intervals occurring over a few milliseconds. This renders the problem of recognizing the sequence much more tractable for a cortical neuron (or network).

If this network generates an *encoding* of the event sequence, recognition still requires *detection* of the correct encoded sequence. We would propose that the detection process could involve coupling a specific set of transmission delays to a coincidence detection mechanism. For example, to detect the spike sequence 1-2-3-4, each spike would be delayed by 2 ms less than the previous spike, and then a single cell would be capable of detecting the temporal coincidence of the incoming spikes. One mechanism for generating such time delays could be axonal delays, however other mechanisms, including passive or active dendritic delays are conceivable. If the inputs to the coincidence detector simultaneously activate feedforward inhibition of the detector, then any set of (non-coincident) inputs that doesn't quickly activate the detector will evoke inhibition that quells the process. Thus, only the proper event sequence will be detected.

The model proposed here shares some features developed by previous investigators. Hopfield and Brody [3] identified the problem of recognizing event sequences (in speech patterns, etc.) and their model similarly initiates a decay (or accumulation) process with each event. Our mechanism differs in that it compresses the event timings into narrowly clustered spike times. Lisman and Idiart [4], in their CA3 model of short-term memory, developed the idea of time compressing spike representations. However, their mechanism requires an underlying voltage oscillation, and focuses more on stable repetition of the spike pattern rather than the input encoding process.

The major assumption of this model is existence of a steadily increasing current injection, initiated by each event. Such a process might reflect a mechanism at either the cellular or network level. For example, a controlled increase in the fraction of cells

activated in a local population, or a net shift in the excitation versus inhibition due to differential facilitation and/or depression of synaptic inputs. At the cellular level, possible mechanisms might include voltage-dependent (e.g., NMDA) or Ca^{++}-dependent synaptic inputs, or current due to persistent Na^{+} channels or metabotropic receptors.

An advantage of the proposed model is that by separating the recognition process into encoding and detection stages, a single circuit can be used to encode all input patterns, and separate cells/circuits are only required for the detection phase. Given the brief (~2 ms) interspike intervals, which as axonal conduction delays would correspond to a distance of ~200-600 μm of transmission along an unmyelinated cortical fiber) both processes of encoding and recognition could easily occur within the spatial domain of a cortical hypercolumn.

Acknowledgements
Supported by ONR MURI grant N00014-01-1062

References

[1] A. Destexhe, Z.F. Mainen and T.J. Sejnowski, An efficient method for computing synaptic conductances based on a kinetic model of receptor binding, 6 (1994) Neural Computation 10-14.
[2] M.L. Hines and N.T. Carnevale, The NEURON simulation environment, Neural Computation 9(6) (1997) 1197-1209.
[3] J.J. Hopfield and C.D. Brody, What is a moment ? Transient synchrony as a collective mechanism for spatiotemporal integration, PNAS 98(3) (2001)
1282-1287.
[4] J.E. Lisman and M.A. Idiart, Storage of 7 +/- 2 short-term memories in oscillatory subcycles, Science 267 (1995) 1512-1515.
[5] M. T. Lazarewicz, S.R. Das and L.H. Finkel, Event signatures in biological motion detection, Proc. Soc. Neurosci. New Orleans 2003.
[6] M. Pavlova, W. Lutzenberger, A. Sokolov and N. Birbaumer, Dissociable Cortical Processing of Recognizable and Non-recognizable Biological Movement: Analysing Gamma MEG Activity. Cereb Cortex. 14(2) (2004) 81-188.
[7] http://senselab.med.yale.edu/senselab/ModelDB/ShowModel.asp?model=3817