

Signal detection of salient visual features by the early visual pathway

Nicholas A. Lesica and Garrett B. Stanley
Division of Engineering & Applied Sciences, Harvard University

Abstract—The role of the lateral geniculate nucleus (LGN) of the thalamus in visual encoding remains an open question. Visual stimulation of the LGN produces two distinct types of responses: tonic and burst. It has been suggested that these two modes of response could serve to transmit high fidelity features of the visual world (tonic) or robustly detect salient features of the visual world (burst). Here, we use a combination of experimental and simulated LGN responses to natural scene movies to investigate the specificity of burst generation, and their role in signaling salient features of the visual scene. LGN responses to excitatory stimulus features were simulated with and without bursts and analyzed using signal detection theory. We find that bursts enhance detection of the onset of excitatory features as compared to purely tonic firing, and that algorithms inspired by this type of framework can enhance the transmission of visual information in situations with limited bandwidth.

I. INTRODUCTION

The lateral geniculate nucleus (LGN) of the thalamus is the gateway between the visual system’s peripheral sensors in the retina and its computational center in the cortex. During visual stimulation, neurons in the LGN exhibit two distinct types of responses known as tonic and burst (for review, see [12], [5]). The response type is determined by the state of low-threshold voltage-dependent calcium channels known as T channels. When the membrane is depolarized, the T channels are inactivated and the neuron fires tonic spikes. After a prolonged hyperpolarization, the T channels are de-inactivated. Subsequent depolarization opens the T channels causing a slow calcium spike which, in turn, causes a high frequency burst of action potentials.

Bursts were originally observed during sleep and were thought to represent a decoupling of the LGN from its retinal input [7], [15]. However, recent studies have demonstrated that burst events do indeed encode stimulus-related information, and that the percentage of bursts in the LGN response can be modulated by the statistical properties of the stimulus [10], [11], [2]. The relationship between the visual stimulus and the burst response is highly nonlinear, as demonstrated by observing the ratio of the first harmonic to the mean ($F1/F0$) and frequency-dependent transfer characteristics in LGN responses to drifting gratings during prolonged hyperpolarization [3], [8]. It has been suggested that bursts serve

to signal the appearance of a salient stimulus (detection), while tonic firing relays detailed features of the stimulus (transmission) [1], [3]. A study of LGN responses in awake behaving cats showed that the response to a drifting grating was characterized by an initial period of bursting at the onset of the stimulus, followed by tonic firing for the duration of the stimulus [4]. This implies that bursts may serve as a wake-up call, alerting the visual cortex to the presence of a stimulus in the receptive field (RF) and signaling the beginning of tonic relay [13]. Indeed, bursts were shown to be more effective than tonic spikes at eliciting responses in their cortical targets [16]. While the results of these studies suggest that LGN bursts may be an important part of the neural code during natural vision, this issue has not been thoroughly investigated.

We investigated the utility of the burst response mode in detecting the appearance of stimulus features typical of the natural environment. We characterized the stimulus features that elicited tonic and burst responses in a population of cat LGN X cells during natural stimulation. To explicitly characterize how the burst mechanism affects detection of these stimulus features, the LGN response to the presence and absence of these features must be compared with and without bursts. Performing such a comparison experimentally, by blocking T channels *in vivo*, is currently not possible, as blockers that abolish the activity of T channels without disrupting the other response properties of the cell have not yet been developed [9]. However, these experiments can be simulated using a biophysically inspired integrate-and-fire-or-burst (IFB) model, which accurately captures LGN encoding during both tonic and burst firing (see [6] for details of IF and IFB models).

Here, we simulated the LGN response to natural stimulus features with and without bursts using the IF and IFB models, and compared the results using signal detection theory. The results demonstrate how the characteristics of the burst response mode (high SNR, low-threshold, and ability to nonlinearly amplify weak stimuli) enhance the ability of an LGN neuron to detect specific stimulus features and suggest that bursts could be used as a reliable signal to direct the deployment of attentional resources to a behaviorally relevant area of the visual field. Finally, we develop a detect/transmit framework for the relay of visual information based on the tonic and burst properties of the LGN. The framework consists of a multi-sensor transmitter (LGN) and receiver (cortex) that are connected by a channel with limited total bandwidth (attention). Each sensor in the transmitter has two states: tonic and burst. In burst mode, a sensor transmits

This work was supported by National Geospatial-Intelligence Agency Grant HM1582-05-C-0009

N. A. Lesica is with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA, lesica@fas.harvard.edu

G. B. Stanley is with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA, gstanley@deas.harvard.edu

only one bit of information corresponding to the absence or presence of a salient stimulus. In tonic mode, a sensor attempts to faithfully relay the visual input with as many bits as are available. The mode of each sensor is determined by the salience of the recent visual input. To evaluate the detect/transmit framework, we compare video reconstructed from the outputs of detect/transmit sensors with that reconstructed from the outputs of transmit only sensors. The results demonstrate that the detect/transmit framework can significantly increase the fidelity of relay by dynamically allocating bandwidth to the most salient areas of the visual field.

II. RESULTS

A. Different features of natural stimuli trigger tonic and burst responses

To characterize the features of natural stimuli that triggered tonic and burst responses, we analyzed the single-unit responses of 58 LGN ON-center X cells to a series of natural scene movies. For each cell, we calculated the average stimulus preceding a burst event, known as the burst-triggered average (BTA), and the average stimulus preceding a tonic spike, known as the tonic spike-triggered average (TTA) [6]. We aligned the RF center of each cell and averaged the BTAs and TTAs across the population.

Figure 1A shows the spatial profile of the BTA in 64 ms increments relative to the onset of the burst event, as well as the temporal profile of the BTA averaged across all pixels in the RF center. The BTA gradually decreases well below the mean luminance from 300 to 100 ms before the burst and quickly rises back to the mean between 100 and 50 ms before the burst. As the BTA never rises above the mean value, it may be surprising that such a feature could evoke any response at all in an ON-center cell. However, the transformation from the stimulus feature described by the BTA to a burst response is apparent after considering retinal processing and the dynamics of burst generation. The transformation from the TTA, shown in figure 1B, to tonic spiking in an ON-center cell is more transparent, as the TTA quickly rises above the mean luminance from 100 to 50 ms before the response and gradually returns to the mean value. Taken together, these results suggest that bursts signal significantly different features of the visual scene and may be utilized for detecting gross changes in the luminance associated with the appearance or disappearance of objects in the visual field.

B. LGN bursts enhance detection of specific stimulus features

The LGN burst mechanism provides an amplified response to specific stimuli. To determine how the burst mechanism affects the ability of the LGN to signal the appearance of specific stimulus features, we can use the IF and IFB models [6] to simulate the LGN response with and without bursts, and compare the results using signal detection theory. In a natural setting, the LGN must signal the appearance of stimulus features that are embedded in noise, for example,

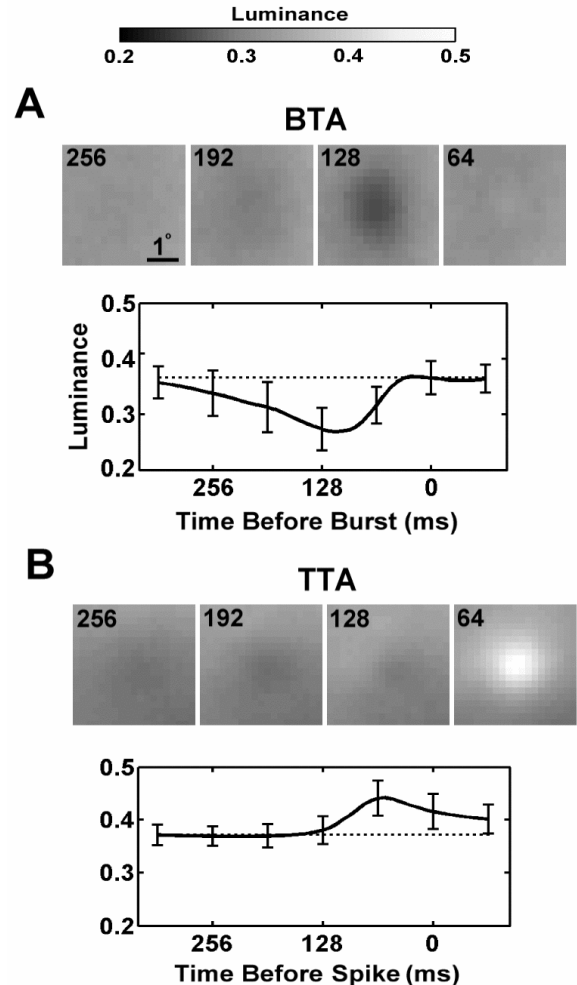


Fig. 1. The features of natural scene movies that trigger burst and tonic responses. **A.** The spatial profile of the burst-triggered average (BTA) in 64 ms increments relative to the onset of the burst event. The BTA shown is the average of the BTAs of 58 ON-center LGN X cells, after their RF centers were aligned. Also shown is the temporal profile of the BTA average across all pixels. The error bars represent one standard deviation. Similar results are shown for the tonic spike-triggered average (TTA) in **B.**

the activation of a vehicle's brake lights viewed through a rainy windshield. To simulate this situation, we presented various stimulus features in combination with additive noise. Each trial consisted of 256 ms of only the noise, followed by a 256 ms stimulus feature superimposed on the noise. The contrast of the feature was chosen randomly on each trial, and the contrast of the noise was varied relative to the contrast of the feature, so that the difference in the responses of the IF and IFB models as a function of the stimulus signal to noise ratio (SNR) of the stimulus could be examined.

To effectively signal the appearance of an excitatory stimulus feature, the response of the neuron in the presence of the feature must be significantly larger than the response in its absence. To compare the responses of the models (firing rate, denoted r) in the presence and absence of a stimulus feature, the response during all intervals which did not contain the excitatory rise of the feature (denoted S_0) and the 64 ms

interval containing the excitatory rise of the feature (denoted S_1) were separated for comparison. A typical trial for the appearance of a bright feature with $\text{SNR} = 2$ is shown in figure 2A, where the actual stimulus is shown in black, and the underlying mean intensity is shown in gray.

The intervals corresponding to the stimulus conditions S_1 (gray) and S_0 (black) are shown. For these simulations noise was also added to the model membrane potential to make the response stochastic. Figure 2B shows the membrane potentials of the IF and IFB models resulting from the example stimulus. The probability distributions of the response of the IF and IFB models under the two stimulus conditions for features embedded in noise with stimulus $\text{SNR} = 2$ are shown in Figure 2C. Distributions were estimated using 1000 trials. Because of the noise in the stimulus, there is no clear separation between the responses under the two stimulus conditions. However, the low threshold and nonlinear amplification properties of the burst mechanism can produce a larger response to the presence of the stimulus and than that of the IF model, helping to separate the response distributions of the IFB model under the S_1 and S_0 conditions. While not every presentation of an excitatory stimulus will trigger a burst in the IFB model (for example, excitatory noise can interrupt the prolonged hyperpolarization necessary to deactivate the burst mechanism) the response of the IFB model to the presence of the feature is, on average, larger than that of the IF model. The results shown here suggest that the nonlinear amplification associated with bursting activity may be a useful mechanism for the detection of salient features of the visual stimulus in the presence of noise.

C. A detect/transmit framework for the relay of visual information

Based on the tonic and burst properties of the LGN that facilitate the detection and transmission of visual input, we have developed a framework for the high fidelity relay of visual information over a channel with limited bandwidth. The framework consists of a multi-sensor transmitter with tonic and burst modes, and a receiver that decodes the transmitted signal and controls the mode of each sensor in the transmitter, designed to mimic cortical feedback control of LGN response mode. A schematic diagram of the framework is shown in figure 3.

The intensity of the visual stimulus (s) is specified by P pixels per frame. The transmitter contains P sensors, each of which corresponds directly to one pixel of the visual input. The transmitter sends the output of each sensor to the receiver once per frame via a noise-free, lossless channel. The bandwidth limit on the channel (for all sensors combined) is specified as β_{max} bits/sec, which, for a frame rate of F frames/sec, corresponds to $\beta_{max}/F = \beta_{frame}$ bits/frame. Each sensor in the transmitter can operate in either tonic or burst mode. In tonic mode, the sensor will attempt to transmit detailed features of the visual stimulus with as many bits as are available. In burst mode, the sensor will signal either the absence or presence of a salient stimulus with only one bit. Following the relay of each frame, the

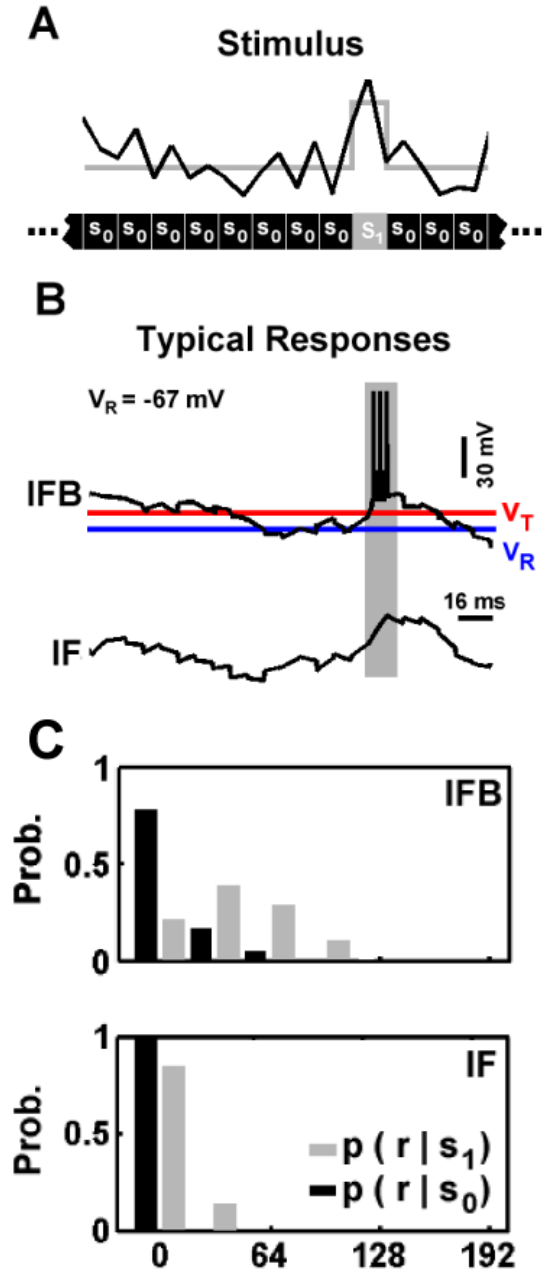


Fig. 2. The features of natural scene movies that trigger burst and tonic responses. **A.** Example of a stimulus embedded in noise. **B.** Voltage trace from integrate and fire (IF) and integrate and fire or burst (IFB) models in response to the stimulus in A. The rest voltage (V_R) and T-channel activation threshold (V_T) are shown for reference. **C.** Histograms of firing rate response from the two models, where the probability of response given presence (gray) and absence (black) of the stimulus feature are separated.

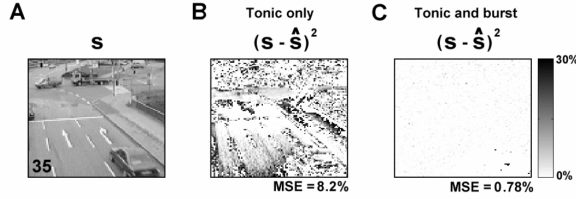
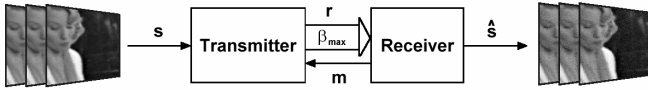


Fig. 3. An LGN inspired scheme for the relay of visual information. Below shows the actual frames from the traffic video and the error in the reconstructions. Each frame consisted of 100×100 8-bit grayscale pixels. **A.** Actual frames 35, 50, and 65. **B.** Squared error in the reconstructed frames (% variance of intensity of actual frame) from tonic only sensors. The MSE of each reconstructed frame is shown. **C.** Squared error in the reconstructed frames from tonic and burst sensors.

receiver determines the mode (m) of each sensor for the next frame based on the saliency of the recent visual input and sends the modes back to the transmitter (Note that the P bits/frame required to send the mode signal back to the transmitter is additional and is not included in constraint β_{max}).

We designed the detect/transmit framework to mimic the ability of the mammalian visual system to efficiently transmit visual information based on ‘bottom-up’ control of attention in response to changes in the external environment. However, ‘bottom-up’ control of attention is only one of many strategies that the visual system has developed to improve the transmission of visual information. Other strategies, such as spatial and temporal decorrelation, separate ON and OFF channels, and mechanisms for task dependent ‘top-down’ control of attention are not included in the model. Correspondingly, in evaluating the framework, we assumed that the goal of the transmitter is to send a representation of the visual stimulus with minimal mean-squared error (MSE). Thus, our model neglects any other features of the neural response that may be important, such as sparseness or redundancy [14].

Example frames of the actual video and the error in the reconstructions are shown in figure 3. Figure 3A shows actual frames 35, 50, and 65 of the video. Figure 3B shows the squared error in the reconstructed frames (as a percent of the variance of the intensity of the actual frame) from relay with tonic only sensors with bandwidth limited to 3 bits/frame per sensor ($\beta_{frame}/P = 3$). The MSE of each reconstructed frame is also shown. Figure 3C shows the squared error in the reconstructed frames from relay with tonic and burst sensors with $\beta_{frame}/P = 3$ and $\sigma_{tonic} = \sigma_{burst} = 2$. The

reconstructions from the signals of the detect and transmit sensors are superior to those from the transmit only sensors, as indicated by the decreased MSE.

III. CONCLUSIONS

In summary, we show here that 1) bursting activity in the visual thalamus is visually evoked, 2) that it signals salient features of the visual scene that are distinct from tonic spiking activity, 3) that bursting activity is associated with enhanced ability to detect signals in a noisy environment, and that 4) this type of transmit/detect framework can be implemented in a very simple algorithm designed to transmit information with high fidelity with limited bandwidth.

IV. ACKNOWLEDGMENTS

The authors would like to thank Yang Dan for the use of the cat data.

REFERENCES

- [1] F. Crick. Function of the thalamic reticular complex: The searchlight hypothesis. *Proc. Natl. Acad. Sci. USA*, 81:4586–4590, 1984.
- [2] K. S. Denning, R. C. Reid, and P. Reinagel. LGN burst state is controlled by visual signals. *Soc. Neurosci. Abstr.*, 699:12, 2003.
- [3] W. Guido, S. Lu, J. W. Vaughan, D. W. Godwin, and S. M. Sherman. Receiver operating characteristic (ROC) analysis of neurons in the cat’s lateral geniculate nucleus during tonic and burst firing mode. *Vis. Neurosci.*, 12:723–741, 1995.
- [4] W. Guido and T. Weyand. Burst responses in thalamic relay cells of the awake behaving cat. *J. Neurophysiol.*, 74:1782–1786, 1995.
- [5] R. Krahe and F. Gabbiani. Burst firing in sensory systems. *Nature Reviews: Neuroscience*, 5:13–23, 2004.
- [6] N. A. Lesica and G. B. Stanley. Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus. *J. Neurosci.*, 24:10731–10740, 2004.
- [7] M. S. Livingstone and D. H. Hubel. Effects of sleep and arousal on the processing of visual information in the cat. *Nature*, 291:554–561, 1981.
- [8] P. Mukherjee and E. Kaplan. Dynamics of neurons in the cat lateral geniculate nucleus: In vivo electrophysiology and computational modeling. *J. Neurophys.*, 74:1222–1243, 1995.
- [9] D. M. Porcello, S. D. Smith, and J. R. Huguenard. Actions of u-92032, a t-type ca^{2+} channel antagonist, support a functional linkage between i-t and slow intrathalamic rhythms. *J. Neurophysiol.*, 89:177–185, 2003.
- [10] P. Reinagel, D. Godwin, S. M. Sherman, and C. Koch. Encoding of visual information by LGN bursts. *J. Neurophysiol.*, 81:2558–2569, 1999.
- [11] C. Rivadulla, L. Martinez, K. L. Grieve, and J. Cudiero. Receptive field structure of burst and tonic firing in the feline lateral geniculate nucleus. *J. Physiol.*, 553:601–610, 2003.
- [12] S. M. Sherman. Tonic and burst firing: dual modes of thalamocortical relay. *TINS*, 24:122–126, 2001.
- [13] S. M. Sherman. A wake-up call from the thalamus. *Nature Neurosci.*, 4:344–346, 2001.
- [14] E. P. Simoncelli and B. A. Olshausen. Natural image statistics and neural representation. *Annu. Rev. Neurosci.*, 24:1193–1216, 2001.
- [15] M. Steriade, D. A. McCormick, and T. J. Sejnowski. Thalamocortical oscillations in the sleeping and aroused brain. *Science*, 262:679–685, 1993.
- [16] H. A. Swadlow and A. G. Gusev. The impact of ‘bursting’ thalamic impulses at a neocortical synapse. *Nature Neurosci.*, 4:402–408, 2001.