Proceedings of the 4th International
IEEE EMBS Conference on Neural Engineering
Antalya, Turkey, April 29 - May 2, 2009

Thalamic Bursts Mediate Pattern Recognition

Magnus Jändel
Swedish Defence Research Agency
Stockholm, Sweden
magnus.jaendel@foi.se

***Abstract*—A new functional model for burst firing in the dorsal thalamus is proposed where thalamocortical pattern recognition systems, based on kernel machine principles, are connected by burst signaling. The systems include input trapping in the dorsal thalamus, cortical learning state memory and processing in the thalamic reticular nucleus. Misclassified events are captured as training examples in the waking state and the pattern recognition systems are trained by extensive thalamic bursting in deep sleep.**

Keywords-thalamus; thalamic reticular nucleus; neocortex; pattern recognition; kernel machine; support vector

I. WHAT IS THE FUNCTION OF THALAMIC BURSTS?

The thalamus routes sensory data to the neocortex and relays also data between cortical domains. Relayed signals are filtered and shaped in the thalamus (see [1] for an in-depth review of the thalamus). Thalamic relay neurons have two firing modes. Tonic firing has an effectively linear relation between inputs and outputs. The burst mode is an all or nothing response where subthreshold inputs have no effect but suprathreshold inputs triggers a long burst of dense spikes. Both modes convey information efficiently but the coding is very different with less noise in the burst mode but high fidelity in the tonic mode [2]. Thalamic burst firing is common during slow-wave sleep but appears also in the waking state. Bursts deactivate spontaneously after ~100 ms and are followed a quiescent period of about 100 ms [3] [4]. Bursting relay neurons resemble a digital register that traps the input and holds a snap-shot for the duration of the burst.

The precise function of the burst relay mode is not known. It has been suggested that bursts control attention [5] and that bursts are wakeup signals for alerting cortical centers to qualitative new sensory input [6]. The tonic mode allows detailed analysis of detected new features once attention has been summoned by burst firing. Bursting in slow-wave sleep is explained [1] as a signal of the idle state of the thalamus.

In-depth understanding of brain systems requires a close interplay between theoretical and experimental work where mathematical modeling inspires comprehensive explorations. In that spirit, we launch a new hypothesis about the role of thalamic bursts – they drive kernel machine pattern recognition in the thalamocortical system.

II. THE PATTERN RECOGNITION HYPOTESIS

We hypothesize that one of the functions of the thalamocortical system is to perform pattern recognition on the many different signals that are conveyed by the system. The neocortex includes memory for the learning state, the dorsal thalamus provides a register for inputs and the thalamic reticular nucleus (TRN) is the processing engine of the pattern recognition machine. In the following "thalamus" shall mean dorsal thalamus and "cortex" shall mean neocortex. The operation of the system is governed by higher-order brain systems (HOBS) including elements of the cortex and the limbic system.

There are three modes of the pattern recognition system - all requiring that the relevant parts of the thalamus shift to burst mode. Qualitatively new sensory data trigger the classification mode. Sensory inputs and learning state data from the cortex are processed in the TRN and the resulting classification is forwarded to HOBS. Misclassifications cause the anomalous input patterns to be impressed in the cortical memory. The system is tuned and optimized in deep sleep. Remembered burst patterns are repeatedly copied from the cortical memory to the thalamic register thus simulating significant experiences. Feedback from the TRN modifies the strength of the memories thereby optimizing the system.

III. KERNEL MACHINE MODEL

To provide a mathematical framework for the outlined model we conjecture that the thalamocortical system implements support vector machines (SVM) [7]. SVM are efficient pattern recognition algorithms that work by implicitly projecting inputs to a large-dimensional feature space where linear classifiers are applied. The solution is a hyperplane in feature space that separates training example classes with a maximal margin. This approach strikes a fine balance between learning the training examples accurately and good generalization when faced with new examples. For simplicity we shall only consider binary classifications. Multi-value classifications can readily be produced by a bank of binary classifiers.

Zero-bias *v*-SVM is a special case of a support vector machine that is uniquely apt for biological implementation [8]. This section provides a bare-bone definition. Consider at set of *m* training examples $(\mathbf{x}_i, y_i)$ where $\mathbf{x}_i$ is an input vector with binary or real-valued components and $y_i \in \{1, -1\}$ is the correct binary classification of the example. The support vector machine classifies a test input vector $\mathbf{x}$ as positive if and only if $f(\mathbf{x}) \geq 0$ where

This work was supported by the Swedish Foundation for Strategic Research.

$$f(\mathbf{x}) = \sum_{i=0}^{m} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) . \quad (1)$$

Zero-bias means that there is no constant factor in (1) as for most support vector machines. The classification function f(**x**) depends of the training examples, the weights α_i and the non-linear symmetric kernel function K. The weights defines the solution to the optimization problem where the dual objective function,

$$W(\boldsymbol{\alpha}) = -\frac{1}{2} \sum_{i,j=1}^{m} y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) , \quad (2)$$

is maximized subject to,

$$0 \le \alpha_i \le \frac{1}{m} , \quad (3)$$

and

$$\sum_{i=1}^{m} \alpha_i = \nu . \quad (4)$$

The parameter $0<\nu<1$ controls the trade-off between accuracy and generalization. This model is a specialization to zero-bias of the ν-SVM described by [9] where the constraint (4) is used as suggested by [10].

There are no local optima so the solution to equations (2)-(4) is found by gradient ascent in the hyperplane defined by (4). Gradient ascent can be implemented so that each weight α_i incrementally is updated (subject to (3)) according to,

$$\Delta\alpha_i \sim \frac{1}{m} \sum_{s=1}^{m} C_s - C_i , \quad (5)$$

where $\Delta\alpha_i$ is the increment of α_i and C_i is given by,

$$C_i = y_i \sum_{j=1}^{m} y_j \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) . \quad (6)$$

Conceptually C_i is understood as the classification margin of the i:th example.

The solution partitions the training examples in two groups: support vectors with weights $\alpha_i > 0$ and trivial examples with $\alpha_i = 0$. Only support vectors contribute to classifications. Support vectors are found close to the boundary between positive and negative domains. They are extreme cases that are hard to classify correctly. Trivial examples are mainstream members of their class. Memory-saving algorithms where trivial examples are discarded from the training set have been shown to be efficient [11].

IV. THALAMOCORTICAL SUPPORT VECTOR MACHINES

The thalamocortical system includes, according to our hypothesis, many different pattern recognition systems that may apply to e.g. visual, somatosensory or corticocortical inputs. This section describes how one of these pattern recognizers is realized as a neural zero-bias ν-SVM. Note that anatomical facts described in the following are accurate but computational functions are speculations.

The functional parts of one of our pattern recognizers are,

- A relay matrix module in the thalamus (RM) for holding the stimuli to be evaluated.

- A processing module (CL) in the TRN.

- A memory for support vectors in the cortex (OM).

The thalamocortical system is mapped so that an organized field of inputs is related to a well-defined set of relay cells in the thalamus, a matching topological map in the thalamic reticular nucleus (TRN) and corresponding maps in the cortex. The RM, CL and OM are hence mapped to each other by driving neural projections (see Fig. 1). HOBS manages the pattern recognizer, activates the three processes of the system and incites the RM to operate in burst mode.

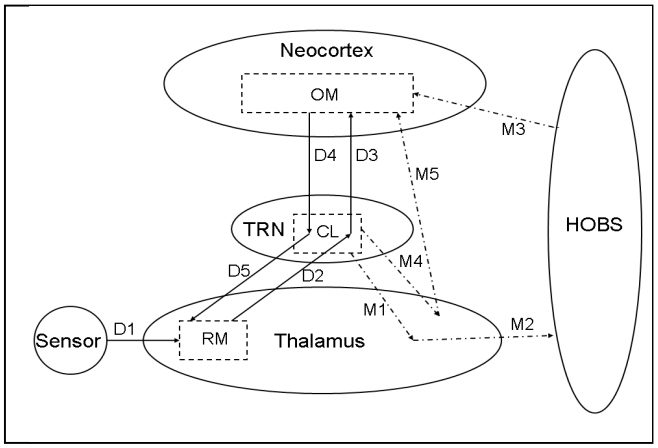

Figure 1. Outline of one of many thalamocortical pattern recognition systems. Solid ovals stand for known brain components where higher-order brain systems (HOBS) is a place holder for management functions in the cortex and in the limbic system. The pattern recognition system consists of hypothetical modules RM, CL and OM that are shown as dashed rectangles. RM is a part of the thalamic relay matrix, CL is the core pattern recognition logic in the TRN and OM is a cortical associative memory. Solid lines (D1, D2, D3, D4, D5) are broad driving projections carrying immediate or recalled sensory data. Dot-dashed lines (M1, M2, M3, M4, M5) are other narrow connections. Afferents (D1) from sensors (or the cortex) drive the relay matrix of the thalmus. Driving pathways (D2) from mapped thalamic relay cells innervate the corresponding map in the thalamic reticular nucleus (TRN) and project (D3) to the OM. Corticothalamic projections innervate the TRN (D4) and carries "dream stimuli" from the OM to the thalamic relay matrix (D5). Classifications are computed in the TRN and are forwarded (M1) to the thalamus for relay (M2) to HOBS. HOBS initiates learning of initially misclassified examples (M3). Learning feedback for adjusting memory weights relays (M4) via the thalamus (M5) to the OM. The connection pattern is anatomically feasible but the detailed function is speculative.

A. The classification process

We shall first consider the classification process in a trained system that receives sensory data (D1). See Fig. 1 for references to neural connections. HOBS activates the pattern recognizer by triggering the burst mode of the RM. Each relay cell in the RM is either quiescent or fires a long spike train. The resulting pattern is the input vector **x** of the SVM.

The OM is an unstable associative memory for support vectors. It oscillates rapidly between support vector states. The average duration of each state $\mathbf{x}_i$ is proportional to the SVM weight α_i. The OM oscillation frequency is much faster than the rate of burst episodes in the thalamus. See [12] and [13] for simulations showing the feasibility of such oscillating memories and [8] for an in-depth analysis of the OM.

Projections (D4) from the cortical OM and relays of stimuli (D2) from the thalamic RM adjoin in the CL of the TRN. The kernel function $K(\mathbf{x}_i,\mathbf{x})$ is calculated in the dendritic trees of reticular cells where afferent relay projections are combined with efferents from the cortex. The classification function (1) is approximated by temporal summation in cells of the CL where the duration of the cortical signal provides for the weight α_i and the valence y_i of the support vector is reflected as an inhibitory or excitatory contribution to the temporal sum. The CL performs hence a temporal stochastic integration. The refractory period of ~100 ms following a burst inhibits any interference between classification episodes.

Since TRN cells project no axons to the cortex but connect abundantly to the thalamus, we assume that the output classifications are relayed (M1, M2) to HOBS via diffusively projecting cells in the thalamus.

B. Surprise Learning

New training examples are engraved in the OM when classifications fail and the animal is surprised. This process is managed by HOBS (M3). The extended duration of thalamic bursts means that the stimuli $\mathbf{x}$ still is projected (D3) to the cortex when the failed classification is detected so that the surprising event can be imprinted in the OM by an appropriate spurt of neuromodulators. The emotional valence of the surprise determines the labeling y of the example. The weight of the new memory is not yet adjusted to reflect its relative importance. All the weights must typically be tuned. The individual that just has suffered a surprise may e.g. overreact to similar stimuli because the pattern recognizer is unbalanced.

C. Importance Learning

A good night's sleep puts new experiences in context. Deep slow-wave sleep is characterized by synchronized bursting of the thalamic relay cells in a TRN-induced ~10 Hz rhythm [14] [15]. Thalamic relay cells are not sensitive to sensory inputs during rhythmic bursting [16]. Contreras et al. suggest that significant experiences are imprinted during slow-wave cycles [17]. We expand on this hypothesis by proposing that the cortex learns support vector weights in deep sleep. While classification episodes are comparatively rare in the waking state and in REM sleep, the deeply sleeping brain tunes the pattern recognition machines by simulating such events. The detailed mechanism may work as follows.

The OM oscillates incessantly between support vector states. The duration of each such transient memory pattern is proportional to the weight α_i. The probability of presenting a support vector $\mathbf{x}_i$ at any given time is hence proportional to α_i.

Projections (D4, D5) from the cortex to the relay matrix of the thalamus are here conjectured to serve as relays for OM memories. The RM accepts "dream stimuli" from the OM since the normal sensory input that dominates in the waking state is quiescent in deep sleep. Dream input is coded as bursts since the thalamus is enticed to rhythmic bursting in deep sleep. The rhythm synchronizes the relay cells of the RM.

A dream stimulus is a copy of an old percept $(\mathbf{x}_i, y_i)$ that has been preserved as a memory in the OM. The relay matrix module will, after each burst episode, lock on the support vector that happens to be projected at the onset of the new burst cycle. The OM oscillates much faster than the TRN ~10 Hz cycle. The RM presents therefore support vectors with a distribution such that the probability of each support vector is proportional to the support vector weight.

The CL in the TRN generates the SVM kernel function while calculating the classification function. This process continues in deep sleep using dream stimuli as inputs. We assume that a weighted version $y_j K(\mathbf{x}_i,\mathbf{x}_j)$ of the kernel function is forwarded to the OM via the thalamus (M4, M5). Note that $(\mathbf{x}_j, y_j)$ is the example that is trapped in the thalamic relay matrix and $(\mathbf{x}_i, y_i)$ is the current state of the OM. The OM receives hence a feedback signal (M5) that, weighted with the label y_i of the present OM state and the probability distribution of $\mathbf{x}_j$, on average is proportional to the quantity C_i of (6).

The sleeping OM is in a slightly plastic phase where support vector memories on average are boosted or diluted according to the rules:

1) Depress the current memory pattern in proportion to C_i.

2) Potentiate all memory patterns in proportion to $\frac{1}{m}C_i$.

When applied in sufficiently small steps in a long series of randomly selected dream stimuli, these learning rules correspond to a gradient ascent solution of the zero-bias ν-SVM problem according to (5). The details of this argument are spelled out in [8]. This hypothetical but biologically feasible learning process finds a good approximation of the optimal support vector weights. The weight for a trivial pattern decays to zero and such patterns are therefore eventually erased from memory.

V. Discussion and Conclusions

The explicit model for thalamocortical kernel machines suggested here is anatomically feasible but the assumed mapping of model to brain is not the only possible choice. For brevity we present here the configuration that appears to be best supported by the admittedly scant evidence. As an example of an alternative mapping we can mention the possibility that corticothalamic driving pathways that have been interpreted as a higher-order relay between cortical centers in fact are the D4 and D5 connections in Figure 1. This option would be more credible if it turns out that all thalamic nuclei mix first and higher order relays as surmised in [1]. The temporary summation in the computation of the classification function can also be located to the thalamus rather than to the TRN. The association with thalamic bursts appears, however,

to be essential. The key advantages of the burst mode as compared to the tonic mode are that i) bursts trap inputs thus allowing an extended evaluation period during which the OM can oscillate many times, ii) the burst mode minimizes interference with normal perception in the tonic mode and iii) the burst mode signals to all sub-systems that the pattern recognition mode is on. A tentative evolutionary path for the suggested apparatus and a neural representation of the parameter ν in (4) are found in [8]. Further evidence of relevance for suggested model is discussed in the following.

Thalamic bursts in the waking state are triggered by novel sensory inputs as suggested by the higher frequency of burst in the lateral geniculate nucleus (LGN) of cats viewing natural scenes as compared to cats viewing featureless artificial environments [18]. Bursts are in particular associated with sudden movement of the viewed objects.

Functionally related thalamocortical and corticothalamic connections are coupled in the TRN by a widespread network of axons and dendrodendritic synapses that would be capable of correlating information over an extended local map [1] [19]. This is precisely what is needed for computing the kernel function of our model.

A thalamic relay cell can innervate several cortical areas and would therefore be able to supply burst signals to the OM as well as tonic signals to other brain centers [1].

Corticothalamic synapses are more peripheral than afferent sensory synapses in the dendritic arbors of thalamic relay cells [1]. Real sensory inputs would therefore override dream inputs in the waking state. Dream stimuli are expressed only when sensory inputs are silent in deep sleep as required by our model. The long time constant of the metabotropic receptors that connect corticothalamic projections to thalamic relay cells makes sense if these are used for setting up dream stimuli in the relay matrix.

Corticothalamic connections to first-order nuclei are usually thought to be modulators and not drivers. The evidence for the modulatory nature of these connections is based on experiments where the originating cortical area is disabled and yet no loss of receptive field is recorded [1]. This is, however, also in line with the present hypothesis where sensory drivers dominate in the waking state and cortical drivers only take over in deep sleep.

There is an order of magnitude more axons from the cortex to thalamus than vice versa [20]. This can be explained if many support vector machines operate on the same input channels. The same visual map could e.g. be scanned for familiar faces, dangerous flying objects or edible berries depending on the present focus of attention. The high corticothalamic to thalamocortical axon ratio may reflect the wide range of possible contexts for pattern recognition.

The following predictions, suggested by the present hypothesis, are open for experimental validation or falsification.

1) "Modulatory connections" from the cortex to the thalamus are in fact driving and capable of inducing meaningful burst patterns in the thalamic relay matrix during deep sleep or if the burst mode is induced while afferent drivers are disabled.

2) The cortex holds many associative memory modules that incessantly cycles trough the impressed patterns with an oscillation frequency of ~100 Hz. Corticothalamic projections from these nodes innervate the TRN.

3) Some of the signals from the TRN to the thalamus are forwarded to the cortex by diffusively projecting relay cells.

4) Damage to the TRN without impact to thalamocortical relay can yet destroy low-level pattern recognition skills.

REFERENCES

[1] S. M. Sherman and R. W. Guillery, Exploring the thalamus and its role in cortical function, 2nd ed., Cambridge, MA: MIT Press, 2006.

[2] P. Reinagel, D.W. Godwin, S.M. Sherman and C. Koch, "Encoding of visual information by LGN bursts," J. Neurophys., vol. 81, 2558-2569, 1999.

[3] H. Jahnsen and R. Llinas, "Electrophysiological properties of guinea-pig thalamic neurons: an in vitro study," J. Physiol., vol. 349, pp. 205-226, 1984.

[4] S.M. Sherman, "Tonic and burst firing: dual modes of thalamocortical relay," Trends in Neurosci., vol. 24, pp. 122-126, 2001.

[5] F. Crick, "Function of the thalamic reticular complex: the searchlight Hypothesis," Proc. Nat. Acad. Sci. USA, vol. 81, pp. 4586–4590, 1984.

[6] S.M. Sherman, "A wake-up call from the thalamus," Nat. Neurosci., vol. 4, pp. 344 –346, 2001.

[7] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based methods, Cambridge: Cambridge University Press, 2000.

[8] Magnus Jändel, "Neural support vector machine", unpublished, 2008.

[9] B. Schölkopf, A.J. Smola, R.C. Williamson and P.L. Bartlett, "New support vector algorithms," Neural Computation, vol 12, pp. 1207-1245, 2000.

[10] C-C. Chang and C-J. Lin, "Training ν -Support vector classifiers: theory and algorithms," Neural Computation, vol 13, pp. 2119-2147, 2001.

[11] B. Schölkopf and A.J. Smola, Learning with kernels, Cambridge, MA: MIT Press, 2002.

[12] L. Pantic´, J.J. Torres, H.J. Kappen and S. Gielen, "Associative Memory with Dynamic Synapses," Neural Computation, vol. 14, pp. 2903-2923, 2002.

[13] D. Horn and M. Usher, "Neural Networks with Dynamical Thresholds," Physical Review A, vol. 40(2), pp. 1036-1044, 1989.

[14] M. Steriade and R. W. McCarley, "Brainstem control of wakefulness and sleep," New York: Plenum Press, 1990.

[15] M. Steriade and D. Contreras, "Relations between cortical and thalamic cellular events during transition from sleep patterns to paroxysmal activity," J. Neurosci., vol. 15, pp. 623-642, 1995.

[16] D.M. McCormick and T. Bal, "Sleep and arousal: Thalamocortical mechanisms," Ann. Rev. Neurosci., vol. 20, pp. 185-215, 1997.

[17] D. Contreras, A. Destexhe, T.J. Sejnowski and M. Steriade, "Spatiotemporal patterns of spindle oscillations in cortex and thalamus," J. Neurosci., vol. 17(3) , pp. 1179-1196, February 1997.

[18] N.A. Lesica and G.B. Stanley, "Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus," J. Neurosci., vol. 24, 10731-10740, November 2004.

[19] E.G. Jones, "Thalamic circuitry and thalamocortical synchrony," Phil. Trans. of the Royal Society of London, Series B, Biological Sciences, vol. 357, pp.1659-1673, 2002.

[20] G.M. Shepherd, ed., The synaptic organization of the brain, 4:th ed., New York: Oxford University Press, 1998.