The InfoPhase Method or How to Read Neurons with Neurons

Raul C. Mureşan[1,2], Wolf Singer[2,3], and Danko Nikolić[2,3]

[1] Center for Cognitive and Neural Studies (Coneural), Str. Cireşilor nr. 29,
400487 Cluj-Napoca, Romania
muresan@coneural.org,
WWW home page: http://www.coneural.org/muresanlab
[2] Max Planck Institute for Brain Research, Deutschordenstrasse 46,
60528 Frankfurt am Main, Germany
[3] Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1,
60438 Frankfurt am Main, Germany

Abstract. We present a novel method that can be used to characterize the dynamics of a source neuronal population. A set of readout, regular spiking neurons, is connected to the population in such a way as to facilitate coding of information about the source in the relative firing phase of the readouts. We show that such a strategy is useful in revealing temporally structured processes in the firing of source neurons, which have been recorded from cat visual cortex. We also suggest extensions of the method to allow for the direct identification of temporal firing patterns in the source population.

Key words: Spiking neurons, temporal coding, readouts, visual cortex.

1 Introduction

Information coding in the brain remains, to this date, the subject of lively debate. It has been proposed that in addition to the traditional firing-rate coding [1], the temporal structure of spike-trains also carries a significant amount of information [2]. A special case of temporal code is represented by neuronal synchronization, which has been recently shown to correlate with the perception of brightness [3], and also with conscious perception [4]. Beyond synchronization, temporal codes might assume more generalized forms, whereby spikes are not perfectly aligned, but arranged into temporal patterns. Polychronization, which relies on synaptic delays, is such an example [5]. A different putative mechanism organizing spikes into temporal patterns is represented by timed inhibition, precisely controlled by fast oscillatory rhythms in the gamma band [6]. In this latter case, temporal phase patterns may even be expressed within a single gamma cycle, with more excited neurons firing earlier relative to the onset of the cycle than less excited ones. Furthermore, we now know that there is a tight interplay between the expression of neuronal oscillations and temporal coding in the cortex [7], with rhythms expressed in different frequency bands [8]. Even further,

temporal codes could be defined beyond single-spike timing. For example, bursts or very fast rate fluctuations could be timed relative to each other, as dictated either by internal network constraints, or by the temporal dynamics of the stimulus. It is very likely that multiple timescales play important roles [9] and the relative timing of processes evolving on these different timescales, is of crucial importance to the brain. Identifying these processes and their specific combinations, which carry information about stimuli, remains however a tremendous challenge. Various methods have been put forward, so far, for the detection and proper interpretation of synchronous spike patterns [10, 11] or temporal structure of membrane potential fluctuations [12, 13]. However, the difficulty remains to identify, in a flexible way, more general activity patterns (beyond synchrony) across a large population of neurons. There are two categories of problems: First, one cannot always know the timescale on which the relevant processes should be searched since these are likely to coexist on multiple timescales. Second, it is unclear how much history dependence one should take into account when detecting candidate correlation patterns. Nontrivial combinations of single spikes, bursts or even fast rate fluctuations, could be co-occurring in a stimulus specific manner and precisely timed relative to each other. The number of possible combinations far exceeds the analysis capabilities of most modern techniques. Thus, one needs to narrow down the search to a specific temporal window and a specific process/timescale. We wanted to break this limitation by considering that in the brain, neuronal information is read out by other neurons which integrate complex input patterns. Following a similar previous approach [14], we considered a readout set of simulated regular spiking (RS) neurons (observers) that preferentially sample subpopulations of a larger neuronal pool (the source neurons), the latter being recorded from the brain. We attempt to recover the information coded in the source neurons' dynamics by observing the relative firing phase of the readouts. At least to a first degree of approximation, interesting events can be assumed to span the timescale defined by the time constant of the neuronal membrane and the dynamics of synaptic currents. Except for the implicit time constants of readouts and their corresponding synapses, we did not restrict the information extraction to a fixed timescale, allowing for a broad range of temporal processes to be observed in a biologically relevant fashion. Moreover, the activation of readouts could reveal nontrivial combinations in the activity of source subpopulations, even when having source processes delayed with respect to each other on a broad temporal interval.

2   Methods

Since the method we describe here relies on information encoded in the relative firing phases of artificial readout neurons, we called the method "InfoPhase".

2.1   Experimental procedures, stimulation and recording

The experiment was performed in the visual cortex of a lightly anesthetized, paralyzed cat. Anesthesia was induced with ketamine (Ketanest, Parke-Davis,

10 mg kg^{-1}, intramuscular) and xylazine (Rompun, Bayer, 2 mg kg^{-1}, intramuscular) and maintained with a mixture of 70% N_2O and 30% O_2 supplemented with halothane (0.5%-1.0%). The animal was paralyzed with pancuronium bromide (Pancuronium, Organon, 0.15 mg kg^{-1} h^{-1}). Visual stimuli were presented binocularly on a 21 inch computer screen (HITACHI CM813ET) with 100 Hz refresh rate. Data were recorded from area 17 by inserting multiple silicon-based multi-electrode probes (16 channels per electrode) supplied by the University of Michigan (Michigan probes). Signals were amplified 10,000× and filtered between 500 Hz and 3.5 kHz for extracting multi-unit activity. Offline spike-sorting techniques were next applied and 61 single units were extracted. The investigated neuronal activity was acquired in response to sinusoidal gratings moving in 12 directions in steps of 30°. For each condition, 20 trials have been recorded, yielding a total of 240 trials. Each trial had a length of 4800 ms. After a spontaneous activity lasting for 1000 ms, the stimulus was presented for 3500 ms, and subsequently, after the stimulus was removed, 300 ms additional OFF response activity was recorded.

2.2   The readout model

For the artificial readouts we used the two-dimensional phenomenological model of Izhikevich [15]. Each neuron is described by a set of differential equations:

$$dv/dt = 0.04v^2 + 5v + 140 - u + I \tag{1}$$

$$du/dt = a \cdot (bv - u) \tag{2}$$

where v - membrane potential, u - recovery variable, I - total post-synaptic current ($I = \sum psc_i$), a,b - parameters. When the membrane potential reaches a value larger than 30 mV, a spike is recorded and the membrane potential is reset to its resting value (-65 mV), while the recovery variable is increased by a given amount (8 for the RS neuron).

We used the parameter settings for the RS neurons, as they best represent the dynamics of pyramidal neurons in the cortex [15]. We next connected 10 RS readouts to the recorded population of cortical neurons (61 source neurons). This was the most important step, as we wanted to obtain the following: reduce the dimensionality of the data by projecting the activity of the source neurons onto the readouts, preserve as much temporal information as possible, and finally, reduce the influence of the firing rate. To this end, we created balanced synapses for each readout: Every readout was connected with all source neurons via excitatory synapses that produced exponentially decaying post-synaptic currents [16]:

$$psc_i = A \cdot W_i \cdot g_i \cdot (E_{syn} - v_{post}) \tag{3}$$

$$dg_i/dt = -g_i/\tau_{syn} \tag{4}$$

where, psc_i - post-synaptic current contributed by synapse i, A - an amplitude parameter determining the maximal amplitude of psc (global constant that allows scaling all synapses), W_i - synaptic strength, g_i - instantaneous synaptic

conductance, E_{syn} - reversal potential of the synapse (taken 0 mV for excitatory synapses), v_{post} - membrane potential of the post-synaptic neuron, τ_{syn} - time constant for the decay of the synaptic conductance (here 30 ms).

The synapses of neurons are balanced in the sense that the total synaptic gain (sum of all input weights) is similar for all readouts. Each connection is instantiated such that it represents a relatively small random fluctuation ($<$ 28%) on top of a constant baseline:

$$W_i = 0.9 + rand_{[-0.25...+0.25]} \tag{5}$$

where $rand_{[]}$ represents the uniform random function in a given interval.

The instantiation scheme for the synapses is crucial. Each readout neuron is receiving roughly the same amount of excitation but samples slightly differently the source population of cortical neurons. By choosing a small enough value for A (here 0.05) the readouts will engage into quasi-periodic firing, with slight differences in their phases (Fig. 1). Because all synapses have very similar strengths, except for the small random fluctuation ($< 28\%$), when one source cortical neuron increases its firing rate, all readout neurons receive simultaneously increased excitation, and hence tend to shift their phase together. The exact amount of phase shift, for each readout, will depend on the particular sampling given by the synaptic distribution, and the activity of the underlying source subpopulations. Each readout can be considered as an independent observer that integrates, over time, the activity of preferred source subpopulations.
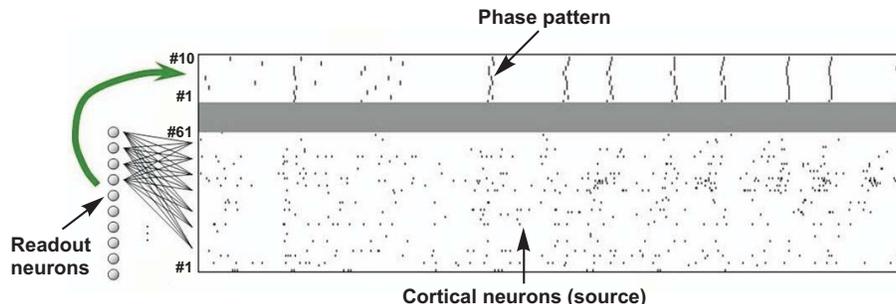
Fig. 1. Schematic representation of the InfoPhase readout method. Readout neurons are connected via conductance-based synapses to recorded cortical neurons, and produce quasi-periodic phase patterns.

We can now go one step further and extract the phase relationships among readout neurons (Fig. 2). These relative phases represent differences in the activation of various source subpopulations (randomly selected by the particular synaptic instantiation) that were integrated over time by the readouts and translated into small phase differences. Thus, we are observing non-trivial combinations of spatio-temporal patterns evolving in the source population (cortical neurons) and reflected in the phases of the readouts.

2.3 The classifier

To assess the amount of information that can be extracted by the readout population from the cortical source, we next constructed a simple classifier, in three steps: extraction of phase vectors, clustering, and training. First, the phase patterns of the readouts are isolated using a sliding window for detection (Fig. 2A) and then converted into phase vectors by computing the relative phases with respect to neuron 1 (Fig. 2B). We obtained a sequence of phase vectors that represents the mapping of the cortical firing onto the activity of the readouts (Fig. 2C). Each experimental trial can then be represented as a sequence of phase vectors.
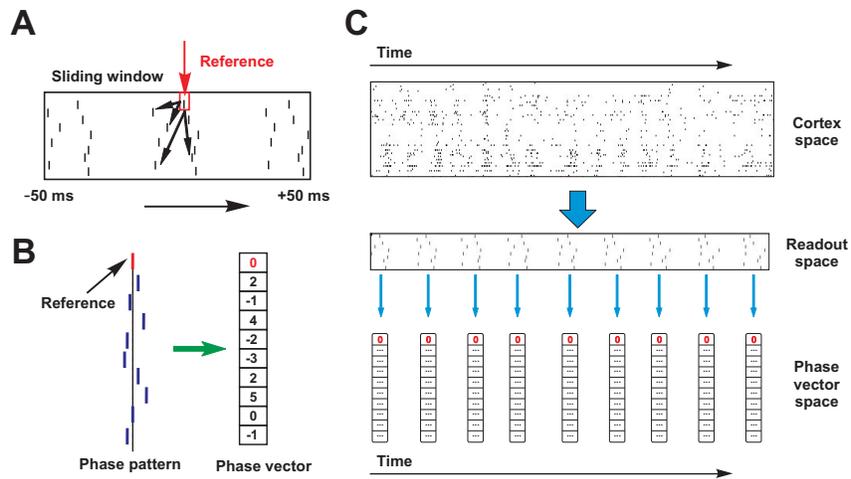
Fig. 2. Building of phase vectors. A. A sliding window is used to detect and isolate the firing phase patterns of the readouts. B. The isolated phase pattern is converted into a phase vector by computing the relative phases with respect to a reference neuron. C. The information from the cortical population is reflected in the readout phase patterns which are then converted into a sequence of phase vectors. Schematic illustration only.

Many phase vectors that represent a trial are similar to each other such that the vector space appears clustered. There are a few dense clouds and many scattered points. To have a good representation of this space we first applied a K-Means clustering with K being determined empirically (for this dataset the optimal $K=40$). Clustering, ensures that dense clouds are represented only by one phase vector (cluster center), and hence, increases robustness against noise. We should mention, however, that the method functions reasonably well without this step. For simplicity, we chose a classifier that memorizes, for each stimulus class, a set of representative, specific, phase vectors, called model vectors. The model vectors are computed during the training phase. We used a split-half

procedure by randomly selecting, for each condition, half the trials (10 trials) for training and half for testing. From each training trial, the phase vectors are considered one by one. For a given vector, the closest phase vector (in Euclidian distance) is searched in the other training trials (of all conditions). If the closest vector belongs to the same stimulus condition, then the vector is marked as being stimulus specific and added to the list of model vectors for the corresponding stimulation condition. After training, each stimulation condition (class) will be represented by a variable number of model vectors, assigned by the previously described procedure. During testing, for each test trial, we first compute the predicted stimulus class by applying a scoring procedure. The score is computed as follows: For each phase vector from the trial to be classified, we search for the closest model vector among all models. For the matching condition, a value is added to the corresponding global scoring:

$$S_k = S_k + 1/NrModels(k) \tag{6}$$

where, S_k - global score corresponding to the matching condition k, $NrModels(k)$ - the number of model vectors describing stimulus class k.
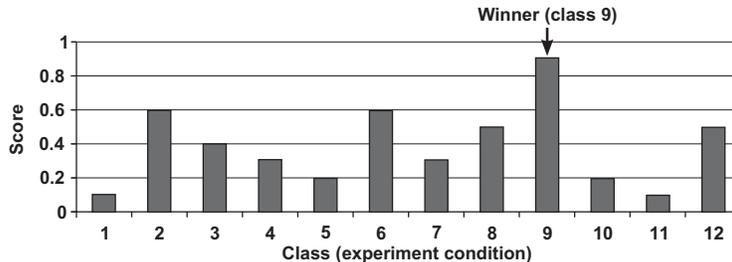
Fig. 3. A typical result of the scoring procedure for a trial. The trial is assigned the predicted class (stimulus) with the highest score.

The scoring protocol normalizes against the number of model vectors of each class such that classes having different number of models are treated equally. All phase vectors of the trial to be classified are scored in a similar manner. Finally, the predicted class of the trial is computed as the class with the maximum score (Fig. 3). After classifying all the test trials, we computed the classification performance as being the percent correctly classified trials from the testing set. Trials for which there were two or more classes sharing the same maximum score were considered as unclassifiable and were not included in the computation of the classification performance. This case appeared however very rarely.

3   Results

We applied the InfoPhase method to a dataset recorded from cat area 17, stimulated with drifting sinusoidal gratings moving in 12 different directions ($30°$

steps, see Methods). We classified the test trials, computed the classification performance, and tried to identify the time scale of the information that has been extracted from the cortical population's activity.

3.1 Classification performance

After training on 120 trials and classifying the other 120 trials we obtained a classification performance of 57%.
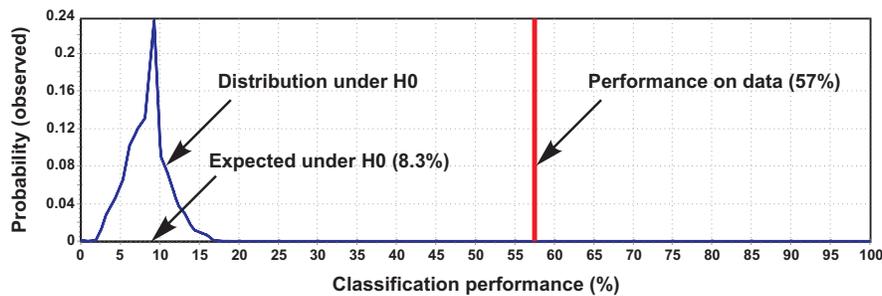
Fig. 4. Classification performance on the original data (57%) and distribution of classification performances when the labels (classes) are randomly permuted (H0).

We next wanted to determine whether this performance was obtained by chance or it reflected legitimate information about the stimuli that was extracted by the method from the cortical spikes. We hypothesized that the observed performance was obtained by chance (H0) and computed the distribution of the performance when the trial labels (classes) were randomly permuted. We used 5000 permutations to estimate the distribution. The expected performance under H0 was (as can also be computed theoretically) 8.3%. More importantly, the performances obtained by chance never exceeded 18% (Fig. 4). Thus, the difference between chance performance and the observed classification performance of 57% on the original, unpermuted data, was highly significant ($p \ll 0.001$).

3.2 Temporal dynamics of cortical neurons

Sinusoidal gratings, which were used in the experimental protocol, are known to produce robust rate responses [17]. By integrating 3.5 seconds of activity across the 61 source neurons and computing the mean firing rate, one can classify the test trials 98.3% correct. The mean firing rate code is completely insensitive to the temporal structure of the spike trains. At the other extreme, there might be fast codes, where information is coded in up to tens of milliseconds of activity. We estimated the sensitivity of the readouts to the temporal structure of the source spike trains by progressively altering the original cortical spike times, in
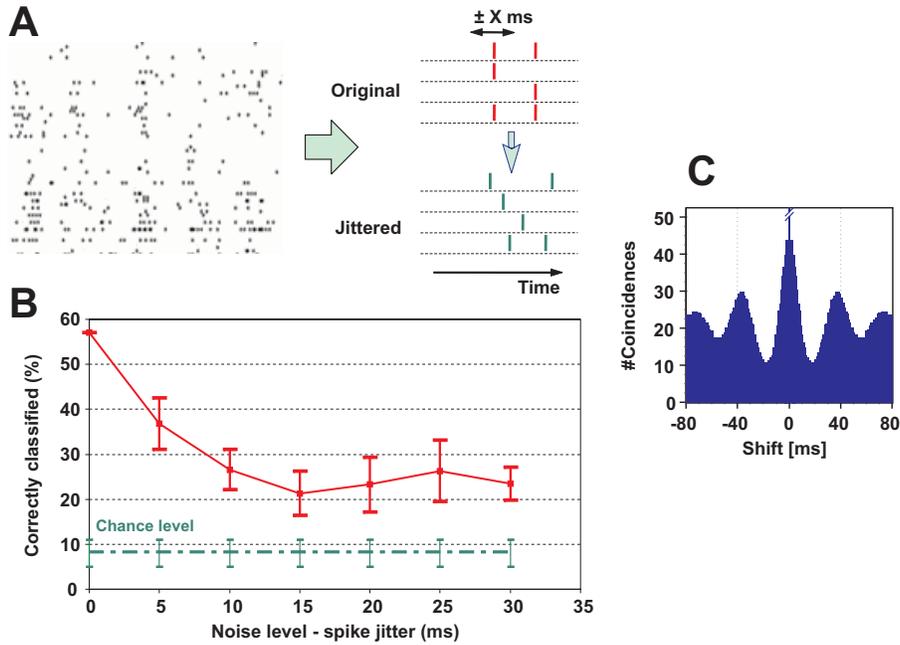
Fig. 5. Determining the temporal scale of the information extracted by the readout population. A. The temporal structure of the source spike trains can be progressively altered by jittering individual spike times. B. The classification performance drops steeply with the amount of jitter, indicating sensitivity of the method to the temporal structure of the source spike trains (error bars are SD). C. Autocorrelation histogram on the spike train of a source unit.

a way that preserves mean firing rates but destroys the fast temporal structure of the data.

We applied progressive jittering of spike times in the source dataset (Fig. 5A) by independently shifting each spike with a value drawn from a Gaussian distribution with 0 mean and standard deviations between 5 and 30 ms, changed in steps of 5 ms. The classification performance dropped rapidly with the amount of jitter applied to the source spike trains (Fig. 5B), hinting towards the possibility that the method extracts information that is encoded on fast temporal scales. Also note that the classification performance remained significantly better than chance, regardless of the amount of jitter. This suggests that, despite using phase readout patterns to classify, the readouts can also take advantage of the firing rate codes in the source dataset and thus the influence of the rate code cannot be completely eliminated. In addition, autocorrelation analysis of the source spike trains (Fig. 5C) revealed oscillatory activity with an oscillation period around 25-27 ms (beta-high band [8]) suggesting that the oscillation cycles might play an important role in enabling the coding of fine-grained temporal information in the source neurons [6].

4 Discussion

As we have seen, artificial spiking neurons can be successfully used to read out information from cortical spike trains. The method we presented here relies on regular spiking neurons that represent biased observers, sampling subpopulations of the cortical neuron set. The bias is minimal such that differences in the activation of the respective subpopulations are only enough to bring readouts out of phase but not enough to change their mean firing rate. The information can be thus extracted without imposing a rigid, fixed time scale for the observers. Moreover, the specificity of phase differences among the readouts seems to correlate with the fine temporal structure of the cortical spike trains. However, there are several aspects that need to be addressed in future studies.

First, the synapses of the readouts are biased in a random way. Thus, each readout represents an independent observer of the source population, transforming the input in a non-linear fashion, with a fixed amount of memory that is given by the membrane and synaptic time constants. However, the collective memory (phase relations between readouts) might go well beyond the individual time constants, as the advancement/delay of readouts relative to each other could last for much longer durations than the individual time constants. A future improvement of the method should consider some synapse training procedure allowing the readouts to optimally encode information in their relative firing phases.

Second, we need to address more in depth the exact signature of the information that is extracted by the readouts from the cortical spike trains. The high sensitivity of the classification performance with respect to the temporal structure of the cortical source suggests that the extracted information is coded on very fast time scales. Nonetheless, there is always the possibility that jittering affects only the capability of the method to correctly extract the information, even if the latter is coded on slow time scales and is thus insensitive to jittering. However, this possibility is remote since we always retrain the classifier after each jittering of the inputs. Retraining insures that the best possible model vectors are built, given the structure of the source spike trains. To find the exact source of the information and the relevant time scale, we need to go back to the original spike trains, guided by the occurrences of the phase patterns that were specific to each class (model vectors). A possibility would be to compute "phase pattern-triggered average" of the source spike trains. In this case, one computes an average of the source spike trains, around the time stamps where a specific pattern occurs, thus identifying the combination of source spikes that produced the pattern. If, indeed, a specific constellation of input spikes produced a specific readout pattern, then the next question is how fast can the readouts extract information? Or how much time is needed to integrate the relevant input and reach a decision? Even more so, what is the role of oscillatory activity in structuring these input patterns?

Finally, we conclude that using artificial spiking neurons to read out information from cortical neurons can be extremely fruitful and opens the path for a new generation of studies that might help us reveal the nature of the neuronal code.

Acknowledgments

References

1. Adrian, E.D.: The basis of sensation. W.W. Norton, New York (1928)
2. Singer, W.: Neuronal Synchrony: A Versatile Code for the Definition of Relations? Neuron **24** (1999) 49–65
3. Biederlack, J., Castelo-Branco, M., Neuenschwander, S., Wheeler, D.W., Singer, W., Nikolić, D.: Brightness Induction: Rate Enhancement and Neuronal Synchronization as Complementary Codes. Neuron **52** (2006) 1073–1083
4. Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., Rodriguez, E.: Synchronization of Neural Activity across Cortical Areas Correlates with Conscious Perception. Journal of Neuroscience **27** (2007) 2858–2865
5. Izhikevich, E.M.: Polychronization: Computation With Spikes. Neural Computation **18** (2006) 245–282
6. Fries, P., Nikolić, D., Singer, W.: The gamma cycle. Trends in Neuroscience **30** (2007) 309–316
7. Buzsáki, G.: Rhythms of the brain. Oxford University Press, Oxford (2006)
8. Mureşan, R.C., Jurjuţ, O.F., Moca, V.V., Singer, W., Nikolić, D.: The Oscillation Score: An Efficient Method for Estimating Oscillation Strength in Neuronal Activity. Journal of Neurophysiology **99** (2008) 1333–1353
9. Abbott, L.F., Regehr, W.G.: Synaptic Computation. Nature **431** (2004) 796–803
10. Grün, S., Diesmann, M., Aertsen, A. Unitary events in multiple single-neuron spiking activity: I. Detection and significance. Neural Computation **14** (2002) 43–80
11. Pipa, G., Wheeler, D.W., Singer, W., Nikolić, D.: NeuroXidence: A Non-parametric Test on Excess or Deficiency of Joint-Spike Events. Journal of Computational Neuroscience DOI 10.1007/s10827-007-0065-3 (2008)
12. Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., Yuste, R.: Synfire Chains and Cortical Songs: Temporal Modules of Cortical Activity. Science **304** (2004) 559–564
13. Mokeichev, A., Okun, M., Barak, O., Katz, Y., Ben-Shahar, O., Lampl, I.: Stochastic emergence of repeating cortical motifs in spontaneous membrane potential fluctuations in vivo. Neuron **53** (2007) 413–425
14. Mureşan, R.C., Pipa, G., Florian, R.V., Wheeler, D.W.: Coherence, Memory and Conditioning. A Modern Viewpoint. NIP-LR **7(2)** (2005) 19–28
15. Izhikevich, EM.: Simple Model of Spiking Neurons. IEEE Transactions on Neural Networks **14** (2003) 1569–1572
16. Mureşan, R.C., Ignat, I. The "Neocortex" Neural Simulator. A Modern Design. In: Proceedings of the International Conference on Intelligent Engineering Systems, Cluj-Napoca (2004)
17. Hubel, D., Wiesel, T.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology London **160** (1962) 106–154