8.07f /W(Jul162004) +model NEUCOM : 10600 Prod.Type:FTP pp.1–5(col.fig.:NIL) ED:SreejaGA PAGN:Sakthi SCAN:

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Neurocomputing ▮ (▮▮▮▮) ▮▮▮–▮▮▮
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Importance of electrophysiological signal features assessed by classification trees

Andreea Lazăr[a,d], Raul Mureşan[b,d], Ellen Städtler[c], Matthias Munk[c], Gordon Pipa[b,c,*]

[a] *Frankfurt International Graduate School for Science, Johann Wolfgang Goethe University, Max-von-Laue-Str. 1, 60438 Frankfurt am Main, Germany*
[b] *Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Max-von-Laue-Str. 1 60438 Frankfurt am Main Germany*
[c] *Max-Planck-Institute for Brain Research, Deutschordenstraße 46, D-60528 Frankfurt am Main, Germany*
[d] *Center for Cognitive and Neural Studies (Coneural), Str. Saturn 24, 400504 Cluj-Napoca, Romania*

Abstract

Sustained activity in prefrontal cortex is associated with the maintenance of information during short-term memory (STM). We have used impurity reduction criteria of classification trees to investigate how the behavioral performance of a monkey during STM is reflected in the information content of three features of recorded signals: rates of individual neurons, oscillations in the LFP, and oscillations in the spiking activity. The LFP power in all bands, but in the α and β bands in particular, is more informative than the firing rate of neurons and the spike power with respect to the monkey's performance.

Keywords: Classification trees; Impurity reduction; Short-term memory; Oscillations; Monkey prefrontal cortex

1. Introduction

Sustained activity in prefrontal cortex is associated with the maintenance of information during short-term memory (STM) [3–7]. Motivated by the controversial discussion of the rate coding hypothesis [10] and the assembly hypothesis [12] we investigated how the behavioral performance of a monkey performing an STM paradigm reflects in the information content of three features: rates of individual neurons, oscillations in the LFP, and oscillations in the spiking activity. For this purpose we used classification trees, a method that identifies structure in the feature space and ranks features according to their information content Fig. 1.

2. Behavioral task and electrophysiological data

We recorded multi-unit ('MUA', 32 kHz sampling rate) and field potential ('LFP', 1 kHz sampling rate) activities simultaneously from up to 16 fiber microelectrodes arranged in a 4×4 matrix with 500 µm spacing in the prefrontal cortex of two monkeys. Signals were filtered (.5−5 kHz (MUA) and 5−150 Hz (LFP) 3 dB/octave) and digitized, preprocessed by rejecting artifacts (movements, licking) and removing line noise at $50 \pm .5$ Hz. In total we analyzed four sessions with 1319 trials altogether (Sessions (1) 227; (2) 505; (3) 332; (4) 255).

The behavioral task of the monkeys was a visual short term memory task. The task consisted of a sample period (first 500 ms) during which a sample stimulus was presented, followed by 3 s of delay. After the delay a test stimulus that was either a matching or non-matching visual object to the sample was presented. The monkey's task was to discriminate between matching and non-matching stimuli and indicate its decision by a button press (match = left, non-match = right) on each trial. On average, the monkeys gave correct responses in 80% of the trials.

3. Method

We compared neuronal activity recorded during trials in which the monkey gave a correct response with activity

*Corresponding author. Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Max-von-Laue-Str. 1 60438 Frankfurt am Main Germany.
E-mail address: pipa@mpih-frankfurt.mpg.de (G. Pipa).

0925-2312/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.neucom.2006.10.136

Please cite this article as: A. Lazăr, et al., Importance of electrophysiological signal features assessed by classification trees, Neurocomputing (2006), doi:10.1016/j.neucom.2006.10.136

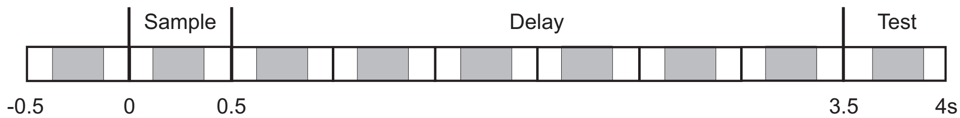

Fig. 1. Time course of the visual short term memory task: after a $-0.5-0$ s baseline, a sample stimulus is presented for 500 ms(starting at time 0), followed by a 3 s delay, and a second matching/nonmatching test stimulus. Time line is divided in 10 consecutive windows of 500 ms each, from which the central 300 ms are used (gray rectangles) to extract features of the signal: rate, spike power and LFP power.

recorded during trials with a wrong response. For each recording session we matched trials with correct and incorrect responses with respect to their number and temporal proximity. We divided the signal of each trial into 10 windows of 500 ms length from which we only considered the central 300 ms to make the windows mutually independent (200 ms sliding window). For each window, we derived the spike rate by dividing spike counts by the window size (300 ms), and computed the spike power and LFP power based on a multi-taper method for frequencies of interest between 5 and 100 Hz ([8], frequency steps 5 Hz, smoothing frequency of 10 Hz). Based on the spectral power estimated for each frequency of interest we extracted the average power in four frequency bands (1) 5–10 Hz; (2) 15–25 Hz; (3) 30–50 Hz; (4) 55–100 Hz). Thus, we derived a total of nine different features for each channel, trial and temporal window: rate, four bands of LFP power and four bands of spike power. To allow for compatibility between sessions we ensured the same number of extracted features per session by randomly selecting the smallest number of channels existing in all four experiments, which was seven, leading in total to 63 features per session.

To assess the discriminative performance concerning correct and incorrect responses per session of each feature, we employed classification trees based on the Gini Index and entropy estimation [1,9]. Classification trees have been widely recognized as an effective techniques for classification in data mining. They were designed to explore data in search of consistent patterns and relationships between variables. A tree is constructed by recursively partitioning a learning sample of data. Making use of the class information for this learning sample, the splits are selected in such a way, that for each step, the maximum separation between different classes is achieved. The ideal split would divide the data so that all items belonging to one class would be completely separated from the items belonging to other classes.

Different measurements have been proposed for evaluating splits [1,9,11,13], but they all have the same basic goal which is to favor homogeneity within each child node and heterogeneity between the child nodes. The goal of splitting is to produce child nodes with minimum impurity (heterogeneity within a node) so that the difference between the impurity of the parent node and those of the children (impurity reduction: Γ) is maximized

$$\Gamma = I_{\text{parent}} - \sum_{i=1}^{k} p(i) I(i), \tag{1}$$

where k is the number of child nodes ($k = 2$ for a binary tree) and $p(i)$ is the fraction of items belonging to each child node after the split. In our case, the impurity was assessed using the Gini index (I_{G}) and entropy (I_{E}) measurements

$$I_{\text{G}}(i) = 1 - \sum_{j=1}^{C} f(i,j)^2, \tag{2}$$

$$I_{\text{E}}(i) = -\sum_{j=1}^{C} f(i,j) \log_2 f(i,j), \tag{3}$$

where C is the number of classes and $f(i,j)$ is frequency of value j in node i.

One of the main advantages of using a decision tree technique for classification is that this method inherently estimates the suitability of features for the separation of items belonging to different classes. This property can be easily exploited when aiming for feature selection. Our goal was to determine a ranking of the features of our extracted signals in different periods of the STM task. Selection and ranking of features is emergent from the classification tree structure.

We investigated whether any of the described features was informative in distinguishing between the two classes: correct and incorrect responses. To this end we built a classification tree for each of the 10 analysis windows that were 300 ms long and covered the baseline, sample, delay and test period of the task. For each time window, one decision tree was used to classify all trials according to the two classes.

To build the classification tree, our approach was to use one set of observations (learning sample) and cross validate it with another completely independent set (testing sample). The prediction for the testing sample gives information on the generality of the tree. Having a similar performance in training and test implies the extraction of the relevant information describing the data and prevents overfitting. To find the ‘right size’ of the tree that maximizes generality we started with a maximally detailed tree and used pruning of leaves until classification performance during test was close to the one during training.

Please cite this article as: A. Lazăr, et al., Importance of electrophysiological signal features assessed by classification trees, Neurocomputing (2006), doi:10.1016/j.neucom.2006.10.136

4. Results

Since our results of the optimal tree size based on pruning and cross validation indicated that only small trees can be reliably built on the analyzed data set, we assessed the impurity reduction (Γ) based on the first level of the trees corresponding to the first split. We assessed Γ for each of 63 features of the spike and LFP signal in four individual sessions with each 200 to 500 trials and in nine epochs that covered the baseline, sample, delay and test periods. Results based on the Gini Index and entropy were similar.

In Fig. 2, results from one of the four sessions are presented. The session contains 505 trials (258 correct and 247 incorrect). For each of the nine temporal windows (time line on the vertical as indicated on the right side of the image), we computed the impurity reduction using rate (left column), LFP power (center), and spike power (right column). Since we considered seven randomly selected recording channels, we obtain seven Γ values for rate, 7×4 Γ values for LFP power and 7×4 Γ values for spike power. The four frequency bands for power considered are: (a) 5–10 Hz; (b) 15–25 Hz; (c) 30–50 Hz and (d) 55–100 Hz.

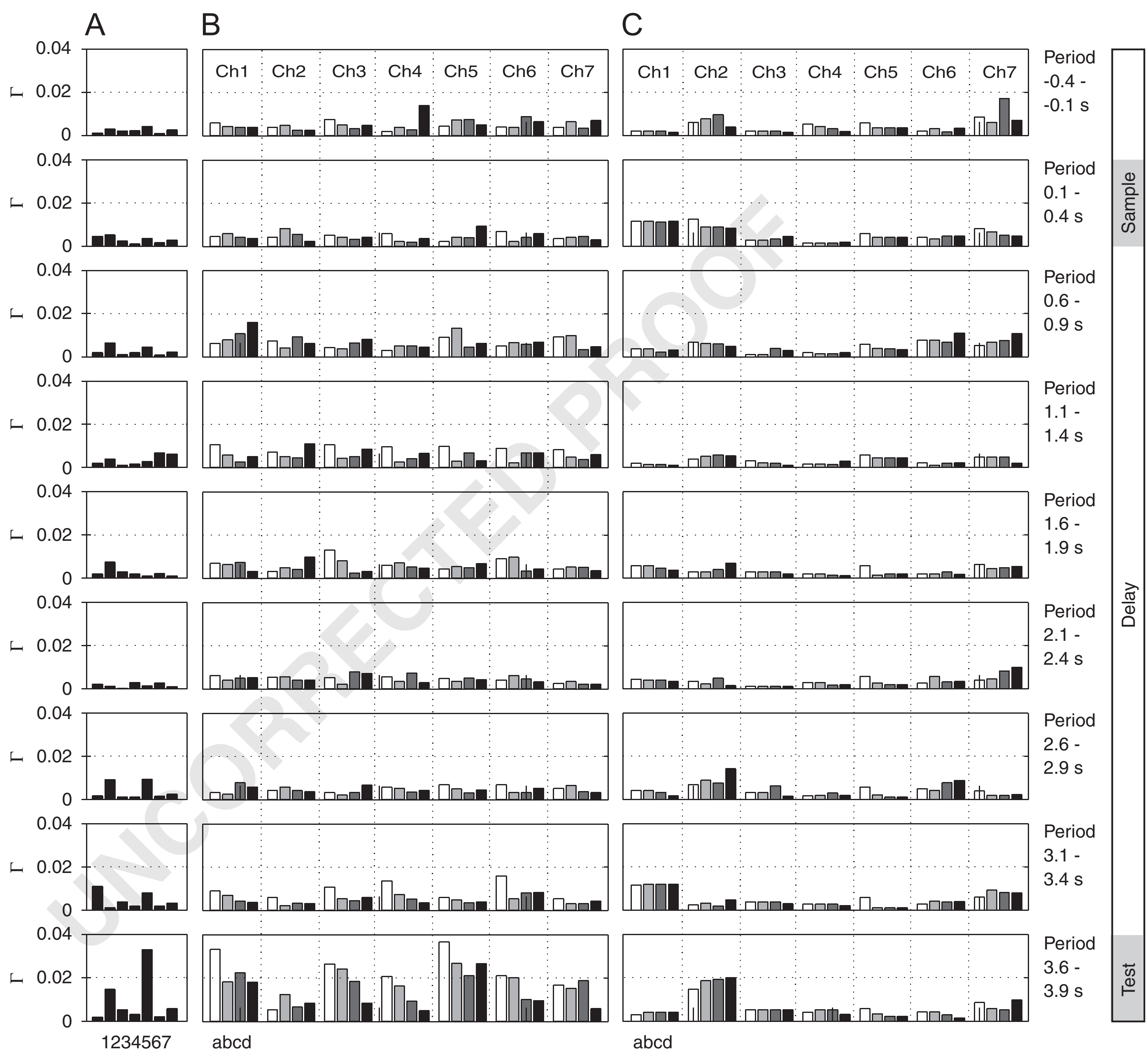

Fig. 2. Impurity reduction, Γ, for separation between correct/incorrect responses for nine periods in time (see Fig. 1) and for signal features of seven Channels (A) rate, (B) LFP power, and (C) spike power. In A individual channels are indexed from 1 to 7. In B and C, four different frequency bands are displayed (a) 5–10 Hz; (b) 15–25 Hz; (c) 30–50 Hz; (d) 55–100 Hz).

Please cite this article as: A. Lazăr, et al., Importance of electrophysiological signal features assessed by classification trees, Neurocomputing (2006), doi:10.1016/j.neucom.2006.10.136

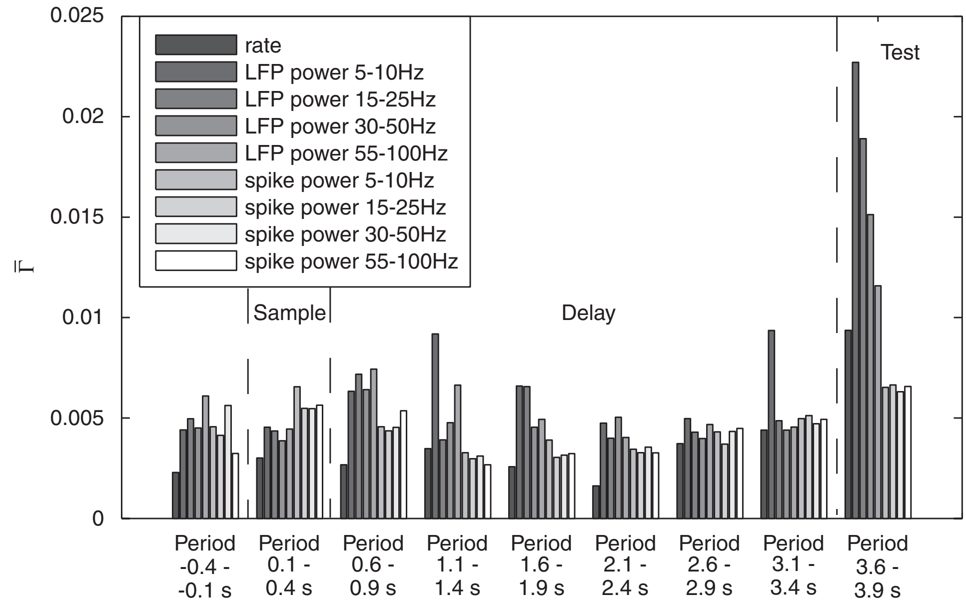

Fig. 3. Average impurity reduction, $\bar{\Gamma}$, (using Gini index) for separation between correct/incorrect responses assessed for nine features and nine periods in time (see Fig. 1)

Thus, we have in total 63 Γ values for each of the time windows.

For individual channels and features the impurity reduction ranges between 0 and 0.035 across baseline and all periods of the task. Impurity reduction (Γ) values increased during the test period for most LFP frequency bands and channels, but also for rate on channel 5, and for spike power on channel 2. Impurity reduction of different features analyzed for the same channels are more similar than Γ of the same feature from different channels. This implies that all considered features reflect on some level the same underlying mechanisms. Overall, low frequency oscillations of the LFP have the highest discriminative power between correct and incorrect trials.

To compare the discriminative power across different periods of the task irrespective of the channel's identity we assessed the average impurity reduction $\bar{\Gamma}$ across all channels (Fig. 3). In two of the four sessions we observed an increased average impurity reduction $\bar{\Gamma}$ during the test period for LFP power between 5 and 50 Hz. The effect is strongest for α and β bands and more moderate for the γ-band while the impurity reduction for the spike rate and spike power features are comparably low. Remarkably, the time course of $\bar{\Gamma}$ indicates that the impurity reduction during the test period is 4 to 5 times higher than during the rest of the task.

5. Conclusions

By using classification trees we intended to rank the importance of signal features for discriminative performance. We found that the reliable size of extracted decision trees was rather small (in the order of one to a few splits), and that the classification performance was low. The reason was, first, that the data did not form well separated clusters, second, that single features were not highly discriminative, and third, that differences between features were rather small. Taken together this might either mean that encoding and maintenance of information in the prefrontal cortex relies on complex signals that express multiple features which are only weakly modulated by behavior, or, that signals which are strongly correlated to behavior were not described by the signal we extracted and analyzed.

6. Uncited reference

[2].

Acknowledgments

This work was in part funded by the Hertie Foundation (AL, RM, GP) and the Volkswagen Foundation (MM, GP).

References

[1] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees, Rapport technique, Wadsworth International, Monterey, CA, 1984.
[2] G. Buzsaki, A. Draguhn, Neuronal oscillations in cortical networks, Science 304 (2004) 1926–1929.
[3] A. Compte, N. Brunel, P.S. Goldman-Rakic, X. J Wang, Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model, Cereb. Cortex 10 (2000) 910–923.
[4] J.M. Fuster, G. Alexander, Neuron activity related to short-term memory, Science 173 (1971) 652–654.
[5] J.M. Fuster, Memory in the Cerebral Cortex, MIT Press, Cambridge, MA, 1995.

Please cite this article as: A. Lazăr, et al., Importance of electrophysiological signal features assessed by classification trees, Neurocomputing (2006), doi:10.1016/j.neucom.2006.10.136

[6] P.S. Goldman-Rakic, Cellular basis of working memory, Neuron 14 (1995) 477–485.
[7] E.K. Miller, C.A. Erickson, R. Desimone, Neural mechanisms of visual working memory in prefrontal cortex of the macaque, J. Neurosci. 16 (1996) 5154–5167.
[8] P.P. Mitra, B. Pesaran, Analysis of dynamic brain imaging data, Biophys. J. 76 (1999).
[9] J.R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo, CA (1993).
[10] M.N. Shadlen, W.T. Newsome, Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey, J. Neurophysiol. 86 (2001).
[11] Y.S. Shih, Families of splitting criteria for classification trees, Statist. Comput. 9 (1999) 309–315.
[12] W. Singer, Neuronal synchrony: a versatile code for the definition of relations, Neuron 24 (1999).
[13] P.C. Taylor, B.W. Silverman, Block diagrams and splitting criteria for classification trees, Statist. and Comput. 3 (1993) 147–161.

Please cite this article as: A. Lazăr, et al., Importance of electrophysiological signal features assessed by classification trees, Neurocomputing (2006), doi:10.1016/j.neucom.2006.10.136