An efficient Video-Synopsis technique for optical recordings with application to the analysis of rat barrel-cortex responses

V. Tsitlakidis [1], N.A. Laskaris [1], G.C. Koudounis[2], E.K. Kosmidis[3]

[1] Laboratory of Artificial Intelligence & Information Analysis, Department of Informatics, AUTh, 54 124, Greece,
[2] Cardiology dept., General Hospital of Kalamata, Greece
[3]Laboratory of Physiology, School of Medicine, AUTh, 54 124, Greece

Abstract— **Optical imaging techniques are nowadays highly popular in neuroscience research, due to their high spatial and temporal resolution. Optical recordings data are coming in the form of sequential snapshots (i.e. videos) reflecting changes in neural activity. By adopting carefully designed stimulation paradigms, these signals constitute an invaluable source of information regarding the emerging spatiotemporal dynamics of brain's response. However, the volume of collected data can obscure the understanding of underlying mechanisms. In particular, the comparison between different recording conditions is a challenging task that is usually solved empirically.**

We introduce an algorithmic technique that identifies spatial domains of coherent evoked activity, produces a meaningful summary of the involved videos and facilitates the comparison of response dynamics. A self-organizing network (SON) lies in the core of the methodology and is responsible for the segmentation of imaged areas into disjoint, functional homogeneous regions. The obtained segments are then ordered according to the strength of response and the ones below an adaptively defined threshold are suppressed. In this way, regions of interest (ROIs) are defined automatically for each response individually and can subsequently be compared across-responses revealing the spatial aspects of neural code that usually give rise to functional maps.

Our technique is demonstrated using averaged responses from rat S1 somatosensory cortex. For the first time, some evidence is provided that the deflection direction of a single whisker might be reflected in the location of activation's first entry.

Keywords— **Optical recordings, Neural Gas, visual summaries.**

I. Introduction

Optical imaging is a relatively new recording technique that uses microscopes and cameras with high spatial and temporal resolution. This explains why only a few methodological papers have appeared so far, concerned with the signal understanding task in the collected data. The analysis usually starts, in an exploratory mode, with the displayed data observed by an experienced user and proceeds with the manual definition of ROIs and the extraction of associated time-courses of activity. For a more thorough treatment, Wavelet-analysis [1] and multivariate decompositions (PCA, spatial-ICA) [2] have been attempted on the raw data. More recently, a manifold learning approach has been introduced [3] with the advantage to produce an advanced parsimonious visualization of the data.

It is the scope of this work to introduce a novel methodology that fully automates the detection of ROIs, and hence can be repeatedly applied as a means of studying the putative temporal-dependent modulations of a complex, spatially-encoding scheme with which the brain differentiates the external world stimuli. Moreover, treating data from a control condition as surrogates for the response-related video-segmentation, we can define finely-tuned thresholds for the precise detection of significantly activated brain regions.

While the original motive stemmed from the recent advances in handling video databases [4], the realized algorithmic steps were borrowed from previous work on mining information from multisite encephalographic recordings [5]. In a nutshell, we first exploit the original Neural Gas algorithm so as to identify spatial domains of coherently-evoked neural activity, then derive a temporal course for each group of pixels and associate a response strength with it, based on a conventional SNR estimator. Using the obtained SNR measurements, the segments are ordered and color-coded. The procedure is repeated for different response latency-ranges and, by keeping a uniform scale for the color-code, the evolution of response dynamics is mapped in the most intelligible way, as a series of activation topographies. The distribution of SNR measurements from the spontaneous activity data can be exploited in a thresholding scheme that will reveal the segments of significant stimulus-evoked activations.

To introduce and demonstrate our methodology, we utilize data from a study targeting rat somatosensation. The scope of that study and some relevant background elements are provided in Section 2. Section 3 is devoted to the presentation of the video synopsis technique, while Section 4

P.D. Bamidis and N. Pallikarakis (Eds.): MEDICON 2010, IFMBE Proceedings 29, pp. 335–338, 2010.
www.springerlink.com

includes a brief report on some new experimental findings resulted from its application.

II. Experimental data

Rodent's whiskers are highly sensitive tactile detectors, similar to primate fingertips that are actively moved through space to extract information about the environment. A somatosensory stimulus evokes a topographical response in cortex, and a physical map of the whisker pad can be found in stained sections of rodent cortex [6]. It is proven that single whisker deflections evoke a specific response that rapidly spreads across the barrels [7] and beyond [8], especially when the animal is anesthetized [9] [10] [11] [12]. The responding neural activity gives the impressions that a particular cortical area responds exclusively to a specific stimulus [13]. However the way in which all this neural activity encodes the stimulus information, is still unknown.

In the particular experimental study performed on anesthetized rats, optical recording data were used to identify the characteristics of stimulus responses that code the direction of the stimulus. A Voltage Sensitive Dye (VSD) was first applied on animal cortex that transformed the intracellular voltages differences into optical signal that was then recorded with special CCD cameras.

The experiments were carried out at School of Medicine of Yale University (D.J. Davis, R. Sachdev and V.A. Pieribone, in preparation) and focused on the neural activation in layer 2/3 of cortex following single whisker movements (stimulation sweeps) [14] [15] or no deflection (background activity sweeps). The background activity sweeps were used to subtract the cardiac and dye bleaching artifacts from stimulation sweeps [16] and also to produce sweeps of spontaneous cortical activity (control sweeps).

Stimuli were whisker deflections on two different directions (caudal and rostral) and three different amplitudes. Here, we use only averaged videos corresponding to the maximum amplitude in both directions, and averaged videos corresponding to spontaneous activity (artifact-corrected data). Each video was one second in duration and had been sampled at a rate of 0.5 kHz. This means 500 frames for each video, with a frame of [80 x 80] pixels. Stimuli onset time (when applied) was at 150 ms (75th frame). The preprocessing of all the averaged videos included: 1) a simple algebraic transformation that associates each pixel (n,m) with a signal expressing the relative increase in fluorescence DF/F due to stimulus onset [3], 2) temporal band-pass filtering and spatial low-pass filtering. 3) definition of the useful part of Field Of View (FOV).

III. The method

A. Feature extraction

We first define the feature vectors that will be used in the subsequent clustering step. A conjunction of temporal and spatial domain features will be adopted, so that clustering will naturally result into connected segments consisting of pixels which reflect similar activation-timecources. Regarding the temporal domain, signal values at consecutive latencies of interest (LOIs) will form the first part of the feature vector that is associated with each pixel (the simplest way to select LOIs is based on the time interval around the peak(s) in the time-dependent curve of integrated activity from a response video (see, fig.1a). The second part of the feature vector is formed by the corresponding pair (x,y) of pixel coordinates in the FOV. Both parts of the feature vector are concatenated after proper scaling (to counterbalance for the range differences) and weighting (based on a factor $0\le\beta\le1$) so as to control the relative importance of the two domains in the overall representation.

With superscripts to denote the two different domains, the data matrix (containing all feature vectors from 'useful' pixels) takes the form:

$$\mathbf{X}^{\mathbf{data}}_{[N\times\ w+2]} = \beta\,.\,\frac{1}{\mathrm{w}}.\frac{\mathbf{X}^{A}_{[N\times w]}}{r^{\mathrm{A}}} \cup (1-\beta).\frac{1}{2}.\frac{\mathbf{X}^{B}_{[N\times 2]}}{r^{\mathrm{A}}} = \left[\ \mathrm{X}_1^A\ \mathrm{X}_1^B\ \mid \cdots \mid \mathrm{X}_i^A\ \mathrm{X}_i^B \mid \cdots \mid \mathrm{X}_N^A\ \mathrm{X}_N^B\ \right] \qquad (1)$$

Where N is the number of pixels (of actual FOV), w is the number of selected latencies and r denotes the range of values in either domain.

B. Spatial Segmentation via Neural-Gas based clustering of pixels

After the feature extraction step, the N formed patterns conveying the spatiotemporal response dynamics are fed to a clustering routine which take over their partition into homogenous groups representing well localized activities of similar temporal morphology. Neural-Gas network is employed to accomplish this task, due to its efficiency [17]. The algorithm is executed using the combined (temporal +spatial) representation, however the results are transferred to both domains and visualized separately.

Strictly speaking, the "Neural-Gas" algorithm is applied to the data matrix $X^{data}=[X_1 \mid X_2 \mid \ldots X_N]$. This algorithm is an artificial neural network model, which converges efficiently to a small, user-defined number $k<N$ of codebook vectors, using a stochastic gradient descent procedure with a ''soft-max'' adaptation rule that minimizes the average distortion error. The network ability to respect the intrinsic

data dimensionality is known [8] and this makes it the best candidate for our summarization purposes. The computed code vectors $O_j \in R^{w+2}$, j=1,2,…k are used in a simple encoding scheme: the nearest code vector are assigned to each X_i in X^{data}. This procedure divides the response manifold $V \subset R^{w+2}$ into k Voronoi-regions

$$\mathbf{V}_j = \left\{ X \in \mathbf{V} : \left\| X - O_j \right\| \le \left\| X - O_i \right\| , \forall i , \; i = 1,2,...k \right\}$$

From a more practical point of view, the bulk of information contained in the data matrix is represented, in a parsimonious way, by a (Nxk) partition matrix U, with elements u_{ij} such that

$$u_{ij} = \begin{cases} 1 & if \;\; X_i \in \mathbf{V_j} \\ 0 & if \;\; X_i \notin \mathbf{V_j} \end{cases} , \quad \sum_j^k \sum_{i=1}^N u_{ij} = \sum_j^k N_j = N \qquad (2)$$

Next, the computed k-partition is applied, individually, to both parts (temporal & spatial) of the patterns. Hence temporal signals and pixel-coordinates are grouped accordingly. By within-group averaging of the former, k response profiles are computed (segment-based averages). Each one serves as indicative response profile for the spatial domain contained in formed segment (i.e. the corresponding group of pixels).

C. Tracking response dynamics via piecewise implementation.

To study the evolution of response dynamics and in particular the spatial aspects of the dynamical changes, we perform repeatedly the above described segmentation procedure, using different (optionally overlapping) time-segments in the pixel-activity representation (eq.1). However, all the intermediate SNR-values obtained in each segmentation step are treated together during the ordering scheme. In this way, a common color map (Fig.2b) is obtained that can facilitate the comparative visualization of response topographies. Figure 2c, serves as a demonstration of this procedure. It includes the successive segmentations (using the overlapping time windows indicated in Fig,2a) of the 3 videos corresponding to the two somatosensory responses and the spontaneous activity.

D. Data-sieving.

To simplify the whole picture, we exploit the outcomes from the application of piecewise segmentation to spontaneous activity data and form an empirical distribution for the SNR-values. As can be seen in Fig.2d, there are a few segments that present high SNR-values even without stimulation. This experimental fact be explained via the tendency of the clustering algorithm (not in particular of Neural-Gas) to identify similar activations in combination with the tendency of SNR-estimator to recognize waveform-similarities as high signal content. Based on this empirical distribution, an SNR-value can be defined as threshold associated with a user-defined significance level. All the segments with response strength (i.e. SNR-value) below the thresholds can be considered as part of spatial domains non responsive to a particular stimulus. By the same token, the derived threshold can be used in the detection of stimulus first entry, by locating the first segment (during the post-stimulus latency range) that exceeds it.

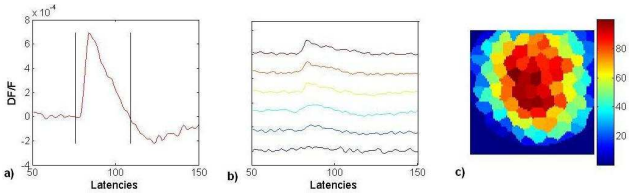

Figure 1: The main steps of the segmentation methodology. a) Feature extraction selecting LOIs based on the time interval around the peak. b) Temporal visualization of ordered segments. c) Spatial domain visualization of ordered segments.

.

IV. RESULTS

Using the described framework for video-synopsis, we contrasted the data corresponding to (averaged) responses from caudal and rostral deflection of the shame whisker and also compared them against spontaneous activity. Rostral deflection resulted in larger amplitude responses than caudal (Fig 2a). Using k=100, β=0.5 and the latency ranges shown in Fig.2a, we show that whisker stimulation induces a stimulus-specific pattern of dynamical changes (see Fig.2c), with the strength and the velocity of spatial spread being the most differentiating characteristics. To provide a fine description of the spatial aspects of stimulus encoding in S1 barrel cortex, we locate and compare the earliest entries of response (based on the above Data-sieving procedure). Figure 2e shows the response topographies after thresholding, and Figure 2f includes the detected top on SNR-rank segments from both deflection-directions. Figure 2f suggests that rostral and caudal deflections of the same whisker are also encoded topographically on the S1 having distinct areas of maximum activation.

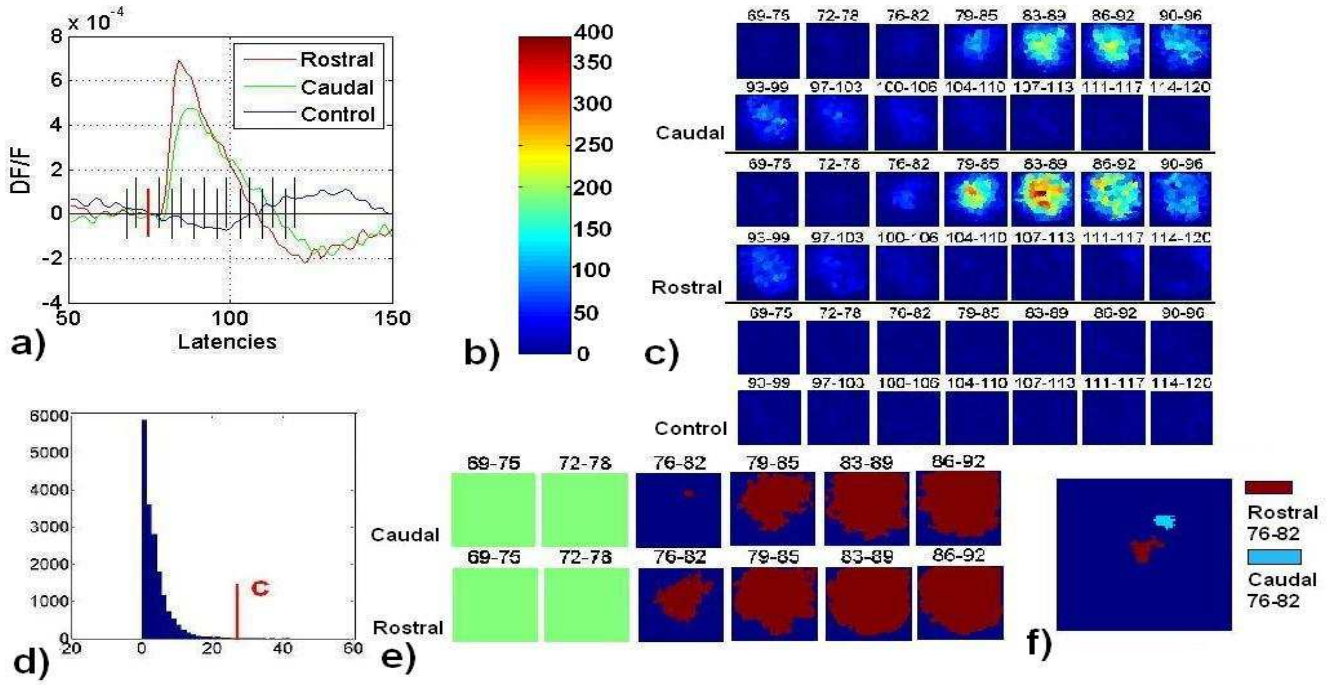

Figure 2: a) Overlapping time-windows used in the piecewise implementation of the main segmentation step. b) The ordered SNR-rank color code. c) Piecewise application of the segmentation step. d) Empirical Distribution of the SNR-values from the spontaneous activity data and the selected threshold c (corresponding to P-value 0.001). e) Response topographies after Data-Sieving. f) First entry response-locus (top SNR-rank segments) for both deflection directions

Conclusions

We introduce a novel method that summarizes optical recording data by identifying spatial domains of coherent evoked activations and present them in an orderly fashion. The obtained visualizations support comparisons across different recording conditions and facilitate the understanding of neural response dynamics at a single glance. Without limiting the applicability of the method, rat somatosensory responses were utilized for demonstration purposes. The preliminary results reveal some new aspects of somatosensory encoding on the cortex. We provide evidence that a spatiotemporal code exists at the cortical level for the directionality of whisker deflection.

References

1. Bathellier et al., (2007), NeuroImage, vol 34(3), pp.1020-1035, 2007
2. Reidl et al., (2007), Neurolmage, vol. 34, no.1, pp. 94-108, 2007.
3. Nikolaos A. Laskaris, Efstratios K. Kosmidis et al, (2008) IEEE Engineering In Medicine And Biology Magazine, March/April 2008.
4. Pritch et al., (2008), IEEE Trans. PAMI, Vol 30(11), pp. 1971-1984, 2008.
5. N. Laskaris, S. Fotopoulos, and A. Ioannides, (2004) IEEE Signal Processing Mag., vol. 21, no. 3, pp. 66–77, 2004.
6. Woolsey TA, Van der Loos H., (1970) The, Brain Res, 17:205-242, 1970.
7. Ferezou I, Bolea S et al, (2006) , Neuron 50:617-629, 2006.
8. Ferezou I, Haiss F, et al, (2007), Neuron 56:907-923, 2007.
9. Kleinfeld D, Delaney KR, (1996), J Comp Neurol 375:89-108, 1996.
10. Senseman DM, Robbins KA, (2002), Journal of Neurophysiology 87:1499-1514, (2002).
11. Petersen CC, Grinvald A, Sakmann B, (2003) J Neurosci 23:1298-1309, 2003.
12. Civillico EF, Contreras D, (2006) J Neurophysiol 96:336-351, 2006.
13. Welker C., (1971), Brain Res 26:259-275, 1971.
14. Simons DJ, (1978), J Neurophysiol 41:798-820, 1978.
15. Sachdev RN, Ebner FF, Wilson CJ, (2004) J Neurophysiol 92:3511-3521, 2004.
16. Orbach HS, Cohen LB, Grinvald A, (1985) J Neurosci 5:1886-1895, 1985.
17. Laskaris N, Fotopoulos et al ,(1997) Electroencephalogr Clin Neurophysiol 1997;104:151±156.

Author: Tsitlakidis Vasilios
Institute: Department of Informatics, AUTh
Street: Aristotle University Campus, 54124
City: Thessaloniki
Country: Greece
Email: vtsitlak@gmail.com