International Journal of Bioelectromagnetism *www.ijbem.org*
Vol. 14, No. 4, pp. 234 - 239, 2012

Denoising Near Infrared Spectroscopy Signals with Principal Component Analysis Improves the Detection of Metabolic Foci

A M Bianchi[a], E Molteni[a], D Contini[b], S Franceschetti[c], A Torricelli[b], E Visani[c], and S Cerutti[a]

[a] *Bioengineering Department, Politecnico di Milano, Italy*
[b] *Physics Department, Politecnico di Milano, Italy*
[c] *IRCCS Neurologico Besta, Milano, Italy*

Correspondence: S Cerutti, Bioengineering Department, Via Golgi 39, Politecnico di Milano, I-20133 Milano, Italy. E-mail: sergio.cerutti@polimi.it, phone +39 02 2399 3339.

Abstract. In this work Principal Component Analysis (PCA) was applied, in order to denoise Near-InfraRed Spectroscopy data recorded from 10 subjects performing visual tasks at different flickering frequencies. Due to the poor signal quality, indeed, the sole application of General Linear Model had not been able to extract any metabolic focus of activation, at the group level. Nevertheless, after pre-processing with PCA, the same dataset could provide stable, bilateral occipital pattern of activation at group level, by means of the same General Linear Model used before. Overall, the employment of Principal Component Analysis during the preprocessing stage provided improved detection of the metabolic foci.

Keywords: Principal Component Analysis, Near Infrared Spectroscopy, Visual task, General Linear Model, Time-resolved Spectroscopy

1. Introduction

Near-infrared spectroscopy (NIRS) technique allows the measurement of hemodynamic and metabolic neuronal responses to brain activation with inexpensive and portable instrumentation. These capabilities are making NIRS, in its present technological state, an important tool in the neurosciences.

Theoretically, NIRS signal consists in a low-variance, slow-dynamic, stable and free-from-artifact track: it is poorly sensible to movement artifacts, it is not invalidated by radiofrequencies nor by the magnetic fields, and it only suffers from environmental red light interferences. Despite the good premises, NIRS tracks often show superimposed optical noise, and massive processing is needed for interpretation. Among the wide choice of methods for signal denoising, Zhang et al. [2004] used a principal component analysis (PCA) to determine the principal spatial components of the spatial-temporal covariance of baseline optical data, and then they used it to filter systemic signal variation from optical data of brain activation. On the other hand, Virtanen et al. [2009] compared PCA with Independent Component Analysis for the extraction of artifactual contributions to the NIRS signal.

Last, a very recent study by Ebihara et al. [2012] employed PCA for the creation of NIRS topograms, to be compared with maps obtained by means of single-photon-emission computed tomography (SPECT).

In the present study, the denoising of NIRS signals by using PCA allowed for the recovery of a NIRS database, which would have been wasted otherwise. Indeed, after PCA application, General Linear Model analysis was performed on data: it could provide group results which had not been obtained with the employment of GLM only.

2. Methods

2.1. Visual task and data recording

The subjects were seated on a comfortable chair with their upper limbs laid on the armrest and with an elbow angle of 90 degrees. A 17” touch-screen monitor was placed in front of them. The test consisted in 3 conditions: (1) checkerboard reversal, at 4 Hz frequency, (2) checkerboard reversal, at 8 Hz frequency and (3) checkerboard reversal, at 12 Hz frequency. The first condition started after 80 s initial baseline. Then, one stimulation period lasting 20s followed, in which subjects were asked to look at the computer screen, presenting a checkerboard with pattern reversal at 4 Hz. Then, 20s rest was designed. This combination of 20s stimulation and 20s rest periods was repeated 10 times. Afterwards, 10s final rest ended the test. This structure was maintained for the checkerboard reversal at 8 Hz and 12 Hz.

The time resolved fNIRS prototype used for this study was provided by the Dipartimento di Fisica (Politecnico di Milano, Italy); time resolved NIRS sampling frequency was 1 Hz. The detailed hardware description is reported in Contini et al. [2006]. Each subject wore a fiber-holder over the head. 16-channels NIRS recordings were obtained.

2.2. Pre-processing with Principal Component Analysis

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is *less than* or *equal to* the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (i.e., uncorrelated with) the preceding components [Abdi & Williams, 2010].

In this work, a synthetic channel, describing the on-of time course of the visual task was added to the NIRS raw signal track of each subject. Then, principal components were extracted by means of a MATLAB routine. Afterwards, components were visually inspected for contributions removal, and then the signal was reconstructed back (fig.1). After signal reconstruction, the new denoised version of NIRS tracks were processed by means of GLM (fig.2).

2.3. NIRS data processing with Generalized Linear Models

Generalized linear model (GLM) approach was applied for data analysis [Friston et al., 1995]. Data processing was performed by means of NIRS-SPM v.3.1 software [Ye et al., 2009]. An hemodynamic response function (HRF) low-pass filter was chosen. The wavelet-MDL detrending algorithm (4 coefficients) was applied, aiming at avoiding the removal of task-related oscillations. No correction for serial correlations was performed. A design matrix (DM) was used, containing three regressors modeling: (1) the rest periods interleaving the stimulation blocks, and (2) the stimulation blocks of the visual test. A contrast array was designed to investigate the relationship between the visual activation and rest. The interpolated t-statistic maps were obtained for each subject first (1^{st} level analysis), and then for the patients’ group (2^{nd} level analysis).

GLM provides intrinsic statistical significance of the maps shown, and it allows the mapping of t-statistics with reference to a threshold, which we set at p-value = 0.05 for the 1^{st} level analysis, and at p-value = 0.001 for the 2^{nd} level analysis.

GLM analysis was done twice: (1) on the raw NIRS tracks making up the data set and (2) on the same NIRS data set, after denoising of the raw tracks by Principal Component Analysis employment.

2.4. Subjects

Ten healthy volunteers (3 males and 7 females), with a mean age of 27.15 years (SD 2.01 years, range 23-30 years), took part in the present study. All volunteers were native Italian speakers and were not paid for their participation. They all had normal vision (or corrected to normal vision) and had no history of psychiatric disorders. The participants were screened thoroughly for neurological symptoms: they all did not show neuro-psychological illness; cognitive level, attentive capability and memory skills were in normal state. Written informed consent was obtained from all volunteers after the examination and test procedure had been explained. The study was approved by the “Istituto Besta” Ethics Committee and it was conducted according to the latest version of the Helsinki Declaration.

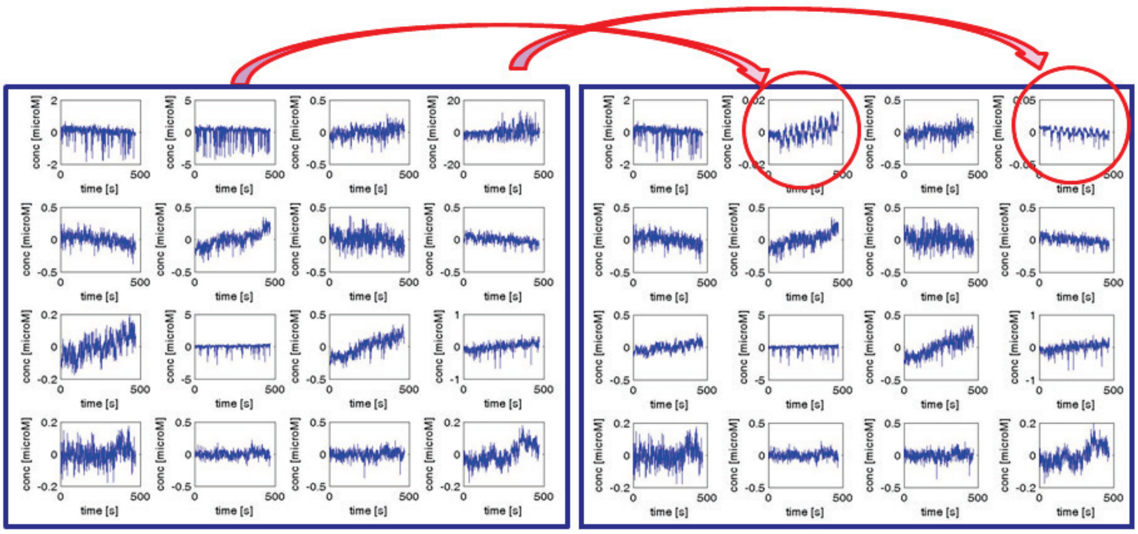

Figure 1. Left panel shows the normalized raw NIRS data, recorded over the occipital cortex of subject n°10. Channels 1 to 16 are depicted from top to bottom, from left to right. The reported signal amplitudes are reduced with respect to the original measurements, since the signal underwent normalization, in order to enter PCA processing. Right panel depicts the same recording, after PCA processing. The reader can notice that channels 2 and 4 have run into substantial modifications. After the employment PCA, channels 2 and 4 show reduced amplitude, but the residual activity clearly unravels an alternating trend, synchronous to the task (protocol design). As a result, remarkable denoising effect is obtained for such channels, resulting in enhancement of the following processing steps, such as General Linear Model employment (see figure 2).

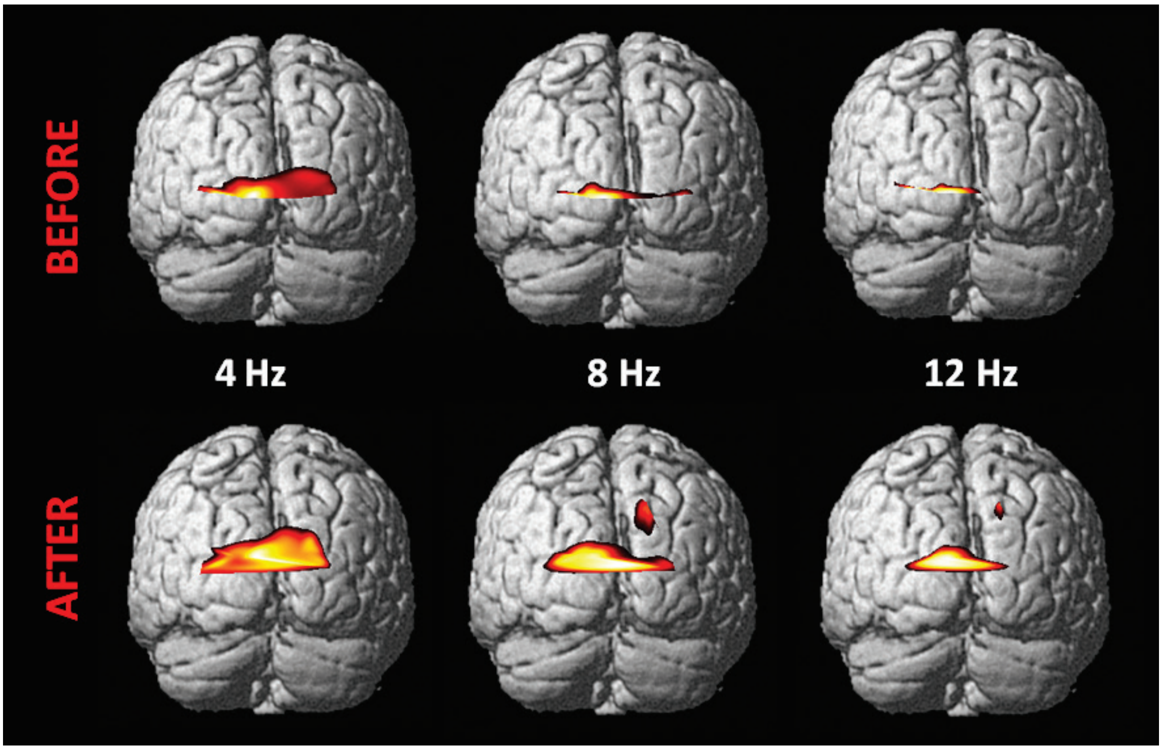

Figure 2. HbR foci over the occipital cortex of subject n°10. Top row: HbR foci have been obtained by applying GLM on the raw NIRS data. Bottom row: HbR foci have been obtained by applying GLM on the NIRS data, after preprocessing with Principal Component Analysis. P-value is 0.05 for GLM statistics. Left column depicts results for 4 Hz stimulation, middle column shows activations for 8 Hz stimulation and right columns reports activation with 12 Hz flickering.

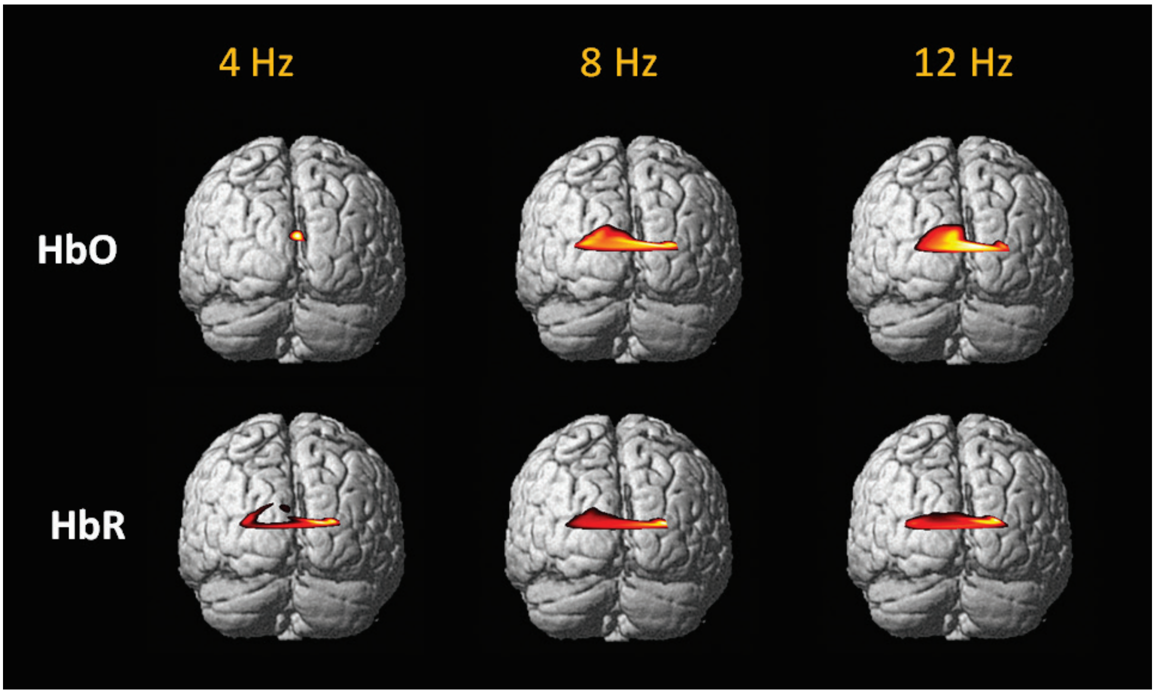

Figure 3. Group analysis (at second level), obtained applying General Linear Model after Principal Components denoising to NIRS data. P-value is 0.001 for GLM statistics. Top row shows HbO maps; bottom row shows HbR maps. Left column depicts results for 4 Hz stimulation, middle column shows activations for 8 Hz stimulation and right columns reports activation with 12 Hz flickering.

Apart from HbO focus in 4 Hz condition, showing improbable inter-hemispheric location and small size, a bilateral region of activation, steady in shape, is recovered for the other conditions and for HbR.

3. Results

The employment of GLM on raw NIRS data did not allow the extraction of significant voxels of activation at the group level (second level analysis): nor setting the significance threshold at p-val=0.001, nor at p-val=0.05. Indeed, the noise superimposed to the majority of NIRS tracks prevented from the extraction of a common pattern of activation.

First level analysis, conducted on the single subjects of the study, revealed variable and often slender patterns of activation. A metabolic pattern with correct occipital location, bilateral extension and stable shape across conditions could be extracted for one subject only.

For this reason, Principal Component Analysis was applied to raw NIRS data, in order to denoise tracks destined to subsequent GLM processing (fig. 1).

After pre-processing with PCA, raw data were newly studied with the same General Linear Model used before. At the single subjects level (first level), in all cases but one a stable pattern could be obtained for both HbO and HbR in the correct (bilateral occipital) location. An example of result improvement is shown in figure 2 for HbR. Figure 3 depicts the group analysis, obtained from the 10 subjects. Group analysis shows consistency of the activated region of interest (ROI) across conditions for HbR, and for 8Hz and 12Hz conditions for HbO. On the other hand, HbO activation in the 4Hz condition seems to be mislocated. Overall, in this second case a group analysis, otherwise unavailable, could be obtained.

4. Discussion

Principal Component Analysis (PCA) has been extensively used in medicine [Joliffe & Morgan, 1992], and has found fertile ground in the electroencephalography and evoked-potentials processing [Chapman & McCrary, 1995; Pourtois et al., 2008]. Despite the large number of applications, PCA is a relatively new method in the optical spectroscopy field, which historically has favored Independent Component Analysis (ICA), leaving PCA aside. Though, after the pioneering application of Zhang et al. [2004], Virtanen et al. [2009], recently compared the performance of both PCA and ICA in disentangling the contributions embedded in NIRS signal, also proving that both methods were able to reduce undesired contributions to the signals, and finally stating that PCA typically performs equal to or better than ICA. Differences between PCA and ICA could be attributed primarily to different criteria for identifying the specific contributions. PCA, moreover, appeared to be especially suited for use in NIRS applications, where the cerebral activation is diffuse.

Last, an additional, very recent application of PCA should also be mentioned: indeed, PCA has been successfully applied for comparing NIRS data with SPECT results in cerebral ischemia [Ebihara et al., 2012]. In this case, PCA allowed the decomposition of NIRS data in focused spatial contributions, and the weight maps, also called topograms, showed high rates of agreement with SPECT data acquired from the same ischemic patients.

To our knowledge, though, the combined use of PCA and GLM on NIRS signals has never been attempted before. The present work attests the advantage deriving from the combined use of these two processing techniques, that in the study described above allowed for the retrieval of a database, which would have been wasted otherwise. Indeed, the use of GLM only had provided inconsistent activations, which completely faded away; on the other hand, preprocessing with PCA allowed some noise reduction, and the recovery of a common functional activation, at a group level.

5. Conclusions

The employment of Principal Comonent Analysis for artifact removal and for denoising of cerebral Near Infrared Spectroscopy signal is proving to be advantageous. PCA denoising seems to be even more helpful in the case of combined use with other statistical methods, such as – in the case shown – General Linear Model. PCA thus candidates as a method for cleaning NIRS data affected by poor signal-to-noise ratio.

Acknowledgements

The research leading to these results has received funding from the European Community’s Seventh Framework Programme under the nEUROPt Project (grant agreement FP7-HEALTH-2007-201076).

References

Abdi H, Williams LJ. "Principal component analysis.". Wiley Interdisciplinary Reviews: *Computational Statistics*, 2: 433-459, 2010.

Chapman RM, McCrary JW. EP component identification and measurement by principal components analysis. *Brain Cognition*. Apr;27(3):288-310, 1995.

Contini D, Torricelli A, Pifferi A, Spinelli L, Paglia F, Cubeddu R. Multi-channel time-resolved system for functional near infrared spectroscopy. *Optics Express*. 14, 5418-5432, 2006.

Ebihara A, Tanaka Y, Konno T, Kawasaki S, Fujiwara M, Watanabe E. Evaluation of cerebral ischemia using near-infrared spectroscopy with oxygen inhalation. *Journal of Biomedical Optics*. Sep 1;17(9):96002, 2012.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* 2(4), 189-210, 1995.

Joliffe IT, Morgan BJ. Principal component analysis and exploratory factor analysis. *Statistical Methods for Medicine Research*. 1(1):69-95, 1992.

Pourtois G, Delplanque S, Michel C, Vuilleumier P. Beyond conventional event-related brain potential (ERP): exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topography*. Jun;20(4):265-77, 2008.

Virtanen J, Noponen T, Meriläinen P. Comparison of principal and independent component analysis in removing extracerebral interference from near-infrared spectroscopy signals. *Journal of Biomedical Optics*. Sep-Oct;14(5):054032, 2009.

Ye JC, Tak SH, Jang KE, Jung JW, Jang JD. NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. N*euroImage* 44, 428–447, 2009.

Zhang Y, Brooks, DH, Franceschini MA, Boas DA. Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *Journal of Biomedical Optics*. Jan-Feb;10(1):11014, 2005.