SINGULAR SPECTRUM ANALYSIS AS A PREPROCESSING FILTERING STEP FOR FNIRS BRAIN COMPUTER INTERFACES.

Loukianos Spyrou * †, *Yvonne Blokland* * †, *Jason Farquhar**, *Jorgen Bruhn* * †

* Radboud University Nijmegen
Donders Institute for Brain, Cognition and Behaviour
6500 GL Nijmegen, The Netherlands

† Radboud University Medical Centre
Department of Anaesthesiology
6500 HB Nijmegen, The Netherlands

ABSTRACT

Near Infrared Spectroscopy is a method that measures the brain's haemodynamic response. It is of interest in brain-computer interfaces where haemodynamic patterns in motor tasks are exploited to detect movement. However, the NIRS signal is usually corrupted with background biological processes, some of which are periodic or quasi-periodic in nature. Singular spectrum analysis (SSA) is a time-series decomposition method which separates a signal into a trend, oscillatory components and noise with minimal prior assumptions about their nature. Due to the frequency spectrum overlap of the movement response and of background processes such as Mayer waves, spectral filters are usually suboptimal. In this study, we perform SSA both in an online and a block fashion resulting in the removal of periodic components and in increased classification performance. Our study indicates that SSA is a practical method that can replace spectral filtering and is evaluated on healthy participants and patients with tetraplegia.

Index Terms— Singular spectrum analysis, NIRS, BCI

1. INTRODUCTION

Haemodynamic changes in the brain have been widely used in functional magnetic resonance imaging (fMRI) and more recently in functional near-infrared spectroscopy (fNIRS) techniques. They capture changes in blood oxygenation level and can be used to estimate differences between various mental tasks. The potential for their utilisation on fNIRS brain-computer interfaces has been studied more recently in movement-based paradigms [1, 2]. fNIRS technology works by placing an near-infrared transmitter on the scalp and measuring the amount of light that is received by a receiver placed on a different site. The amount of received light depends on its absorption by haemoglobin molecules, which in turn provides an estimate of the activation in various areas of the brain.

The most common application of movement based brain-computer interfaces (BCI) is for individuals who have lost the motor capabilities necessary to interact with the external world, by partly or completely omitting the regular motor output pathways. Instead, mental tasks of movement attempt or imagery are performed by the user and then decoded by a machine learning algorithm, such that the mental task can be translated into actual meaningful output like control of devices. For fNIRS, an increase in haemoglobin oxygenation (HbO) concentration coupled with a decrease in haemoglobin de-oxygenation (HbR) is observed during movement [2]. These patterns are exploited in order to facilitate the detection of movement. As is common for most BCIs, an algorithm is trained on subject specific examples of movement and subsequently used on new examples to predict whether the user is moving. Pattern classification techniques are used to train classifiers that can learn the prediction function that maps the observed data to the user intended task. Due to the nature of the NIRS signal, it contains non-task related biological artifacts that corrupt the acquired signal and make the classification task difficult. Undesired activity patterns such as slow drifts, respiratory, blood pressure and heart rate waves, movement artifacts and general instrumentation noise have been identified and methods have been devised to overcome them.

Blood pressure waves (Mayer waves $[0.05\ 0.2]$Hz), respiratory artifacts (around 0.4Hz) and heart rate (around 1Hz) are traditionally dealt with by spectral filtering the appropriate band of interest. Such methods suffer when a biological activity changes frequency throughout an experiment. In [3] a data driven method that removes periodic components from a fNIRS ensemble has been described, although it employs many channels to facilitate the process. Similarly, in [4] a method is described that employs the use of a reference channel which is placed on the scalp in such a way that doesn't pickup up brain activity but only the superficial non-brain artifact sources. Then, this activity is subtracted from the data channels resulting in reduced artifact noise. In [5], adaptive filtering methods have been developed which adapt the subtraction process in a temporal way based on the observed data, however, their efficacy is dependent on the level of the interfering artifacts. Slow drifts are usually dealt with by high pass filtering, detrending or more recently by wavelet minimum description length methods [6]. Moreover, similar

to fMRI research, the general linear model (GLM) has been used to model the NIRS responses and remove noise [7].

In this paper, we suggest a method based on Singular Spectrum Analysis (SSA) that is used for the smoothing of fNIRS haemodynamic responses. SSA has found many applications such as weather forecasting, economic time series prediction and medical data analysis [8]. It has been used in [9] for the removal of cardiac artifacts from electromyography signals in an adaptive way and it has found application in fMRI [10]. Our study focuses on the removal of periodic components and is proposed as a replacement of spectral filtering. It has the advantage that it is non-parametric, data driven and overcomes the problem of artifacts whose frequency band changes during the course of an experiment. The algorithm was evaluated on a dataset containing healthy participants and patients with tetraplegia.

2. SINGULAR SPECTRUM ANALYSIS

SSA is a non-parametric spectral estimation technique that decomposes a time-series $\mathbf{x} = [x_1 \ldots x_N]$ into various components such as trend, oscillatory waves and noise [8]. It consists of four steps:

A. Embedding The time series is converted to its trajectory matrix

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \ldots & x_K \\ x_2 & x_3 & \ldots & x_{K+1} \\ \ldots & \ldots & \ldots & \ldots \\ x_L & x_{L+1} & \ldots & x_N \end{bmatrix} \quad (1)$$

which is a Hankel matrix with equal antidiagonal elements.

B. Singular value Decomposition Perform the singular value decomposition of $\mathbf{X}$ and obtain $\mathbf{U}, \mathbf{\Sigma}$ and $\mathbf{V}$ as $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. The matrices $\mathbf{U}$ and $\mathbf{V}$ contain the left and right singular vectors and the matrix $\mathbf{\Sigma}$ is a diagonal matrix which contains the singular values in order of decreasing magnitude. Then, the matrix $\mathbf{X}$ is partitioned in the following way:

$$\mathbf{X} = \sum_i \mathbf{E}_i \quad (2)$$

where $\mathbf{E}_i = \sigma_i \mathbf{U}_i \mathbf{V}_i^T$ is an eigentriple and the index i corresponds to the i_{th} singular value or vector. Each matrix contains one component corresponding to a trend, oscillatory wave or noise.

C. Eigentriple grouping According to some criterion select a subset I containing M of the L eigentriples and form the resultant matrix:

$$\mathbf{X}_I = \sum_{i \in M} \mathbf{E}_i \quad (3)$$

D. Reconstruction This step brings the time series into its original form containing only the selected subset of eigentriples. The matrix $\mathbf{X}_I$ is diagonally averaged, or Hankelized, to obtain a time series of the same dimensions as the original series.

The most important parameters for SSA algorithms are the choice of the window length L and the the size of the time series. The largest periodicity T that can be separated depends on L by realising that the window length should be larger than the periodicity $L \geq T$. It has been shown [8] that values close to $L = N/2$ provide a low reconstruction error but for a limited type of processes. More recently [11], it has been shown that the rule of $L = log(N)^c$, where $c \in (1.5, 3)$ is more general and suitable for real-world signals. Also, for time series where the slowest periodicity T_s is much less than N, the requirements on L are more relaxed [8].

3. MATERIALS AND METHODS

3.1. Dataset

For these analyses we used two-channel fNIRS data collected from 8 healthy participants and 7 patients with tetraplegia performing a fingertapping task [2]. The NIRS optodes (interoptode distance of 35mm) were placed above the motor cortex near the C3 and C4 locations of the 10-20 EEG system. For each subject twelve 15-second trials (Fs=250Hz) were recorded per movement condition: actual movement, imagined movement and rest. As the patients were unable to perform actual movements they executed attempted movements instead.

3.2. SSA single-channel preprocessing

The optical signals from the fNIRS acquisition device were converted to haemoglobin changes using the modified Beer Lambert law [12]. This converts the optical density changes for each of the two channels to oxygenated (HbO) and deoxygenated (HbR) concentration changes. The resulting dataset contains 4 time series (2 channels x 2 haemoglobin types) and was further downsampled to 10Hz. The SSA algorithm was applied to each time series separately. The differential pathlength factor (DPF) was selected individually for each subject according to their age. To evaluate the smoothing properties of SSA we compared the following preprocessing procedures as an input to the classification algorithm:

- (M1) baseline case, no preprocessing
- (M2,M3) lowpass filters: a) 0.16Hz and b) 0.08Hz
- (M4) bandstop filter: [0.08 0.12]Hz
- (M2 & M4) together
- (M5) SSA in sequence batches
- (M6) SSA trial-based

The baseline method gives us a comparison reference to base the efficacy of the various methods. Methods M2,M3 are commonly found in the literature and M4 attempts to correct for Mayer Waves. The filters were applied to the total time series which was around 30 minutes long. M5 is performed by dividing the whole time series in six blocks ($N \approx 2500$) belonging to each of the six sequences (section 3.1) and performing SSA on each batch separately. This emulates an offline evaluation of the method. The value of N is also large enough to be a sizeable multiple of the expected periodicities. The expected approximate periodicities for our sampling frequency of 10Hz are $L_h = 10$ for heart rate, $L_r = 20$ for respiration and $L_m = 100$ for Mayer-waves. M6 is performed on a trial-by-trial basis. Each trial containing 18 seconds of data (see section 3.3) is smoothed using SSA using approximately the preceding three minutes ($N = 2000$). For both M5 and M6 the leading eigentriple (the one with the highest singular value) was selected to evaluate the method. This corresponds to extracting a general trend of the lowest resolution [8]. Using that eigentriple for the series reconstruction, the window length was selected by performing a grid search on a large number of window lengths ($10 \leq L \leq 300$). The window L with the highest classification performance on the training set was selected (section 3.3).

3.3. Classification

Trials were split into 3-second segments and the concentration changes were averaged within each segment. This increased the number of training examples to 60 per movement condition. A classifier using both the HbO and HbR average concentration changes had four features per segment (2 averages x 2 channels). The classifier was evaluated on the time period of 3 to 18 seconds, since this is when the response is expected. Performance of an L2-regularized linear logistic regression classifier was computed for three binary problems to distinguish each individual movement condition from the rest condition: 1) executed movement versus rest (controls only), 2) attempted movement versus rest (patients only) and 3) imagined movement versus rest (both groups). Classification performance was evaluated with a chronological (block-wise) 12-fold cross-validation where a block corresponded to one 15 second trial, i.e. five consecutive 3-second segments. For each fold 2 blocks (one per condition) were removed from the training set to make the test examples. For more details on the classification approach refer to [2]. The statistical significance between the preprocessing methods was evaluated with McNemar's test [13].

4. RESULTS

Selecting the leading eigentriple for the grouping and reconstruction phase (see section 2.C) was successful in producing a smooth series for all subjects and movement type. Including any eigentriples other than the leading one produced periodic components in the reconstructed series (results omitted due to space limitations). A grid search was performed for the window length for the following values: $10 \leq L \leq 300$. Highest classification performance on the test set was obtained for $L \approx 100$. In Figure 1 the resulting training and test (cross-validation) set classifier performance is shown over the whole dataset. The performance peaks around $L = 100$ close to the periodicity of Mayer waves. This also corresponds to $c \approx 2.27$ regarding the rule of $L = log(N)^c$ for $c \in (1.5, 3)$ ($N = 2000$). In Table 1, for $L = 100$, the resulting classification test set performance can be compared to spectral frequency methods as described in section 3.2. SSA performed significantly better than M2&M4 ($z - score = 3.6177$, $p < 0.01$) and than the baseline method ($z - score = 6.01$, $p \ll 0.01$) according to two-tailed McNemar's test over the whole dataset. In Figure 2 group averages are shown for actual movement (healthy group) and attempted movement (patient group). Group averages are computed by normalising each subject's average to have unit norm. Shaded regions display the variance between subjects as expressed for one standard deviation of the group average. In Figure 3 we show an example reconstructed trend for one of the subjects. For visualisation purposes, the grouping of eigentriples to identify Mayer waves and heart rate was performed by grouping the leading eigentriples (except the first) in terms of the peak at their frequency spectra. As is observed in Figure 3 (d), complete separation was not established between Mayer waves and heart rate.

Classification rate	overall	HAM	TAM	HIM	TIM
M1-Baseline	70	76	75	65	63
M2-LPF 0.16Hz	71	77	75	64	66
M3-LPF 0.08Hz	71	78	76	64	66
M4-BSF	72	78	77	67	65
M2&M4	73	80	78	67	67
M5-SSA batch	75	82	77	71	68
M6-SSA online	75	82	78	71	67

Table 1. Classification performance results for various preprocessing steps. Results are shown for healthy subjects actual and imagined movement (HAM,HIM) and for patients with tetraplegia attempted and imagined movement (TAM,TIM). For SSA $L = 100$ was used.

5. DISCUSSION

This study counts as a critical initial step towards the incorporation of SSA in NIRS-BCI research. It is shown that SSA can provide enhanced performance over spectral filtering methods for removing the undesired periodic components prevalent in NIRS signals. With no prior assumptions, the parameter of interest (window length) is estimated as to maximise class separability which is described by the increased classification per-
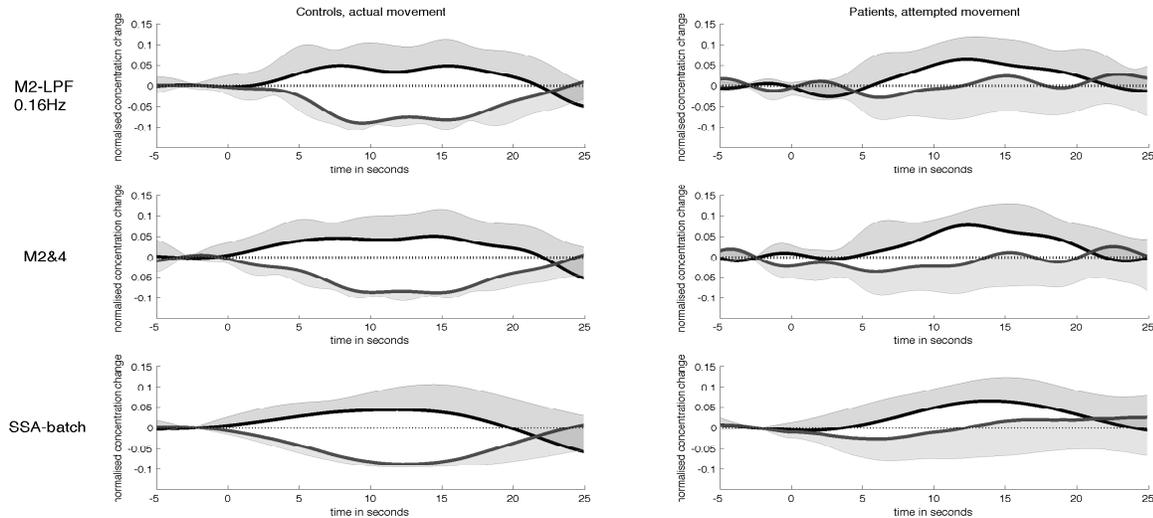
Fig. 2. Normalised group averages for HbO (black) and HbR (gray) for 3 different preprocessing methods. Actual movement for healthy participants is shown on the left and attempted movement for patients with tetraplegia on the right. The shaded area corresponds to one between-subject standard deviation from the mean. Movement is from 0 to 15 seconds. The SSA method shows less variance and less periodic fluctuations.
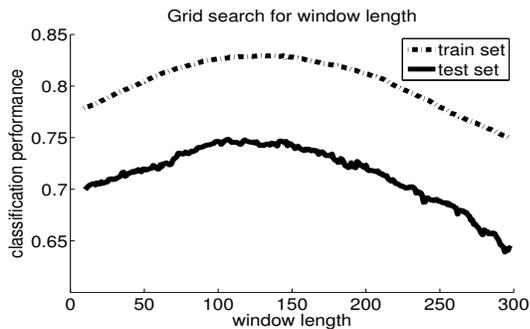

Fig. 1. Training (dashed) and test (solid) classification performance for various window lengths.

formance. The estimated window length coincides with the periodicity of Mayer waves suggesting that this artifact was successfully rejected since there is a spectral overlap between the task related response and Mayer waves. Selecting only the first eigentriple provided a general trend component for all subjects and conditions without additional grouping algorithms. However, complete separation between the different periodic components was not achieved. Sequential SSA can be employed for this purpose where different window lengths are used to extract specific components separately [14] . The reconstructed time series show less between-subject variance and fewer oscillations than with conventional spectral filtering even with frequency bands where the artifacts are expected. It has to be noted that due to the possible variability of the frequency of e.g. Mayer waves and heart rate, during the course of an experiment, a data-driven method such as SSA is pre-

ferred to a fixed spectral filter that takes into account the general population's expected frequency range. Also, by studying in detail the reconstruction properties in a subject specific way (e.g. adapting the window length and the selection of eigentriples) higher performance may be achieved. The algorithm was applied both in an offline and an online fashion. The equivalence of the results is indicative that the method can be used in online BCI experiments enhancing its utility.

The multichannel extension of SSA (MSSA) where multiple time series are grouped together and their global trend is separated is suggested as a next step. These multiple time series can be the data from multiple NIRS channels. In that light, the algorithm should be evaluated and compared within the context of other multichannel methods.

REFERENCES

[1] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, and B. Blankertz, "Enhanced performance by a hybrid NIRS-EEG brain computer interface.," *NeuroImage*, vol. 59, no. 1, pp. 519–29, Jan. 2012.

[2] Y. Blokland, L. Spyrou, D. Thijssen, T. Eijsvogels, W. Colier, M. Floor-Westerdijk, R. Vlek, J. Bruhn, and J. Farquhar, "Combined EEG-fNIRS Decoding of Motor Attempt and Imagery for Brain Switch Control: An Offline Study in Patients With Tetraplegia," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 1–8, 2014.
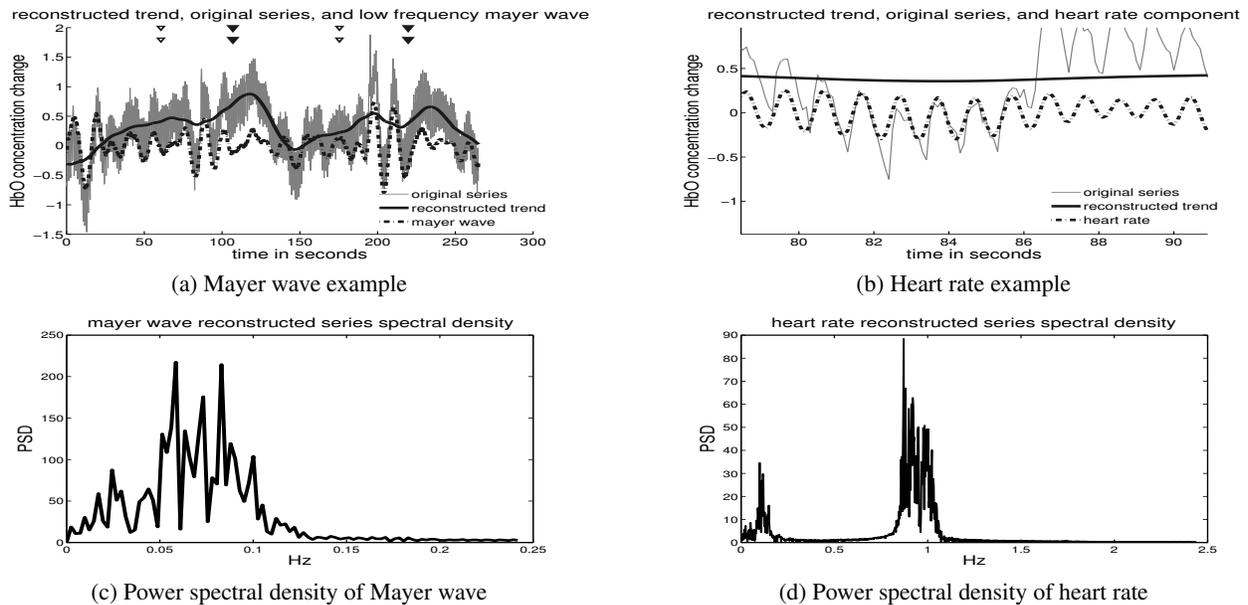
[3] T. Fekete, D. Rubin, J. M. Carlson, and L. R. Mujica-

(a) Mayer wave example

(b) Heart rate example

(c) Power spectral density of Mayer wave

(d) Power spectral density of heart rate

Fig. 3. Example of extracting the trend for one subject and one sequence. In a) the extracted trend is shown for the whole sequence together with the original time series and a group of eigentriples corresponding to the frequency range of mayer waves. Small triangles indicate the start of imagined movement while large triangles indicate actual movement. b) Zooms in to show a different group of eigentriples corresponding to heart rate. c) and d) show the power spectral densities of the the reconstructed mayer waves and heart respectively. Note that the eigentriple grouping for Mayer waves and heart rate was performed for illustration purposes only and has no effect on the classification performance.

Parodi, "The NIRS Analysis Package: noise reduction and statistical inference.," *PloS one*, vol. 6, no. 9, pp. e24322, Jan. 2011.

[4] R. B. Saager and A. J. Berger, "Direct characterization and removal of interfering absorption trends in two-layer turbid media.," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 22, no. 9, pp. 1874–82, Sept. 2005.

[5] Q. Zhang, G. E. Strangman, and G. Ganis, "Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: how well and when does it work?," *NeuroImage*, vol. 45, no. 3, pp. 788–94, Apr. 2009.

[6] K. E. Jang, S. Tak, J. Jung, J. Jang, Y. Jeong, and J.g C. Ye, "Wavelet minimum description length detrending for near-infrared spectroscopy.," *Journal of biomedical optics*, vol. 14, no. 3, pp. 034004, 2009.

[7] J. C. Ye, S. Tak, K. E. Jang, J. Jung, and J. Jang, "NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy.," *NeuroImage*, vol. 44, no. 2, pp. 428–47, Jan. 2009.

[8] Golyandina N. and Zhigljavsky A., *Singular Spectrum Analysis for Time Series*, SpringerBriefs in Statistics, 2013.

[9] S. Sanei, Tracey K. M. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis.," *IEEE transactions on bio-medical engineering*, vol. 59, no. 2, pp. 428–34, Feb. 2012.

[10] P. Paolo, M. Danilo, G. Claudio, H. Giacomo, G. Angelo, and L. Alberto, "Singular Spectrum Analysis and adaptive filtering enhance the functional connectivity analysis of resting state fMRI data," *International Journal of Neural Systems*, vol. 24, no. 3, May 2014.

[11] R. K. Atikur and D. S. M., Poskitt, "A Note on Window Length Selection in Singular Spectrum Analysis," *Australian & New Zealand Journal of Statistics*, vol. 55, no. 2, pp. 87–108, June 2013.

[12] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function.," *Trends in neurosciences*, vol. 20, no. 10, pp. 435–42, Oct. 1997.

[13] S. L. Salzberg, "On Comparing Classifiers : Pitfalls to Avoid and a Recommended Approach," *Data Mining and Knowledge Discovery*, vol. 327, pp. 317–327, 1997.

[14] Golyandina N., Nekrutkin V., and Zhigljavsky A., *Analysis of Time Series Structure: SSA and related techniques*, Chapman and Hall/CRC, 2001.