DOI:10.3217/978-3-85125-378-8-13

Including temporal dynamics in single trial Motor Imagery classification using Empirical Mode Decomposition

Simon R.H. Davies[1] and Christopher J. James[2]

[1] Institute of Digital Healthcare, University of Warwick, Coventry, Warwickshire, U.K.
davies_s@wmg.warwick.ac.uk
[2] School of Engineering, University of Warwick, Coventry, Warwickshire, U.K.
c.james@warwick.ac.uk

Abstract

This paper describes the inclusion of temporal signal information in the process of Empirical Mode Decomposition (EMD) applied to the electroencephalogram (EEG). Key information in the classification of motor imagery EEG is very frequency centric but also contains strong temporal cues due to the nature of the experimental paradigm. Taking the temporal dynamics of the EEG into account, multiple snapshots of the signal in time are input into a multi-variate EMD process. Features were created from the processed signal using Common Spatial Patterns and these were input to a Support Vector Machine for classification. The results showed that the added temporal dynamics gave no major improvement to the sensitivity or specificity compared to regular EMD.

1 Introduction

Motor Imagery (MI) works by using the changes in brain activity over the sensorimotor cortex during imagined localised limb movements [5]. These changes include the suppression of the μ rhythm on the contralateral hemisphere. The μ rhythm is limited to a frequency band of 8-13 Hz, predominantly around 10 Hz. There is also weaker resonance activity in the beta bands around 20 Hz. During the commencement of MI there will be a large decrease in potential followed immediately by a large increase and then a return to pre-MI levels in the contralateral hemisphere as the limb being imagined moved. The lateral hemisphere will see a similar effect but to a much lesser degree, so called Event Related Desynchronisation/Event Related Synchronisation (ERD/ERS).

Empirical Mode Decomposition (EMD) is a process that can decompose non-stationary and non-linear signals into a group of frequency harmonics called Intrinsic Mode Functions (IMFs) and residual signal noise [2]. This makes it readily applicable to electroencephalogram (EEG) recordings. In this paper we work with the specific Brain-Computer Interface (BCI) [5] paradigm of MI. EMD has already been shown to be a useful signal processing technique for MI [1]. Here we introduce a new version of EMD that uses Taken's theorem [**?**] to incorporate temporal data into the EMD sifting process. This new algorithm will be tested in classifying MI trials, compared to standard EMD.

Any continuous single-spatial channel signal can be converted into a multi-temporal channel signal using Taken's theorem as described in reference [3]

$$x(t) = (x(t-\tau), x(t-2\tau), ..., x(t-(m-1)\tau) \in R,$$

where τ is the lag and m is the embedding dimension. A similar method was used to apply Independent Component Analysis (ICA) to a single EEG channel [3]. The embedding dimension

 DOI:10.3217/978-3-85125-378-8-13

needs to be greater than $2D + 1$, where D is the number of signal sources. From [3] we set the dimension size of $m = 30$ and $\tau = 1$.

2 Methodology

2.1 Dataset

The EEG dataset contains 90 MI trials from 11 different subjects. Each trial is 8.2 seconds long, consisting of 4.1 seconds neutral activity, a stimulus cue that is randomly either left or right, and 4.1 seconds of sustained MI. The recordings were made with a 64-channel EEG running BCI2000 software with a sampling frequency of 160 Hz [6]. The electrodes selected for analysis were FC3, FC4, C5, C3, Cz, C4, C6, CP3, CP4, T7 and T8 according to the 10-10 system and were determined to cover the motor cortex. This is the same electrode selection as previous EMD studies [4].

2.2 Feature Extraction

EMD works by applying an iterative sifting process to the data, where the mean envelope of each channel is subtracted repeatedly until the data has been decomposed into a group of oscillations of varying frequencies that are symmetrical with respect to the x axis. However, Park et al [4] came up with an enhanced version that uses spatial information to help form the envelopes, causing each channel's IMFs to be equal in number to the rest and occupy the same frequency bands, making them consistent across all channels and far easier to analyse, this is called Multi-variate EMD (MEMD) and was shown to perform well compared to existing standard methods such as wavelet transforms and the application of a Butterworth filter.

MEMD works by calculating the mean envelope of n-directional signals, instead of just one signal in isolation, by projecting the signal along different directions in n-variate spaces and then averaging the projections to get the local mean. Low discrepancy Hammersley sequences are used to obtain quasi-uniform points on high dimensional spheres to form the projection vectors. In this case instead of applying MEMD to EEG data of multiple channels, we use dynamical embedding to create a multi-temporal signal and apply MEMD to that. The aim of this is to use temporal data to form more accurate and consistent IMFs from a signal channel using the temporal data, this is called Temporal MEMD (T-MEMD). Channels FC3, FC4, C5, C3, Cz, C4, C6, CP3, CP4, T7 and T8 according to the 10-10 system were selected for analysis.

The resulting IMFs for each trial, for either EMD or T-MEMD, are un-embedded and then selected or discarded based on two different methods, one knowledge-based and one performance-based. The knowledge-based method attempts to select relevant IMFs based on previous knowledge of their features. A Fast-Fourier Transform (FFT) is calculated for the duration of the signal that was recorded after the cue as it contains the MI. IMFs with at least 5% of their total power between 8-13 Hz (the μ rhythm) are considered to have possible MI relevant data and are kept. The performance-based method uses a brute force method to identify the combination of IMFs that gives the best result per subject in terms of the number of trials correctly classified. As with this dataset the number of IMFs per channel never exceeds 13 and the IMFs containing useful information never exceeds 5 in number, it means all possible combinations of IMFs (maximum number of combinations: 2379) can be tested in a practical amount of time.

The selected IMFs are summed to form the processed signal and features are created for each trial using Common Spatial Patterns (CSP). CSPs calculate a set of spatial filters that

 DOI:10.3217/978-3-85125-378-8-13

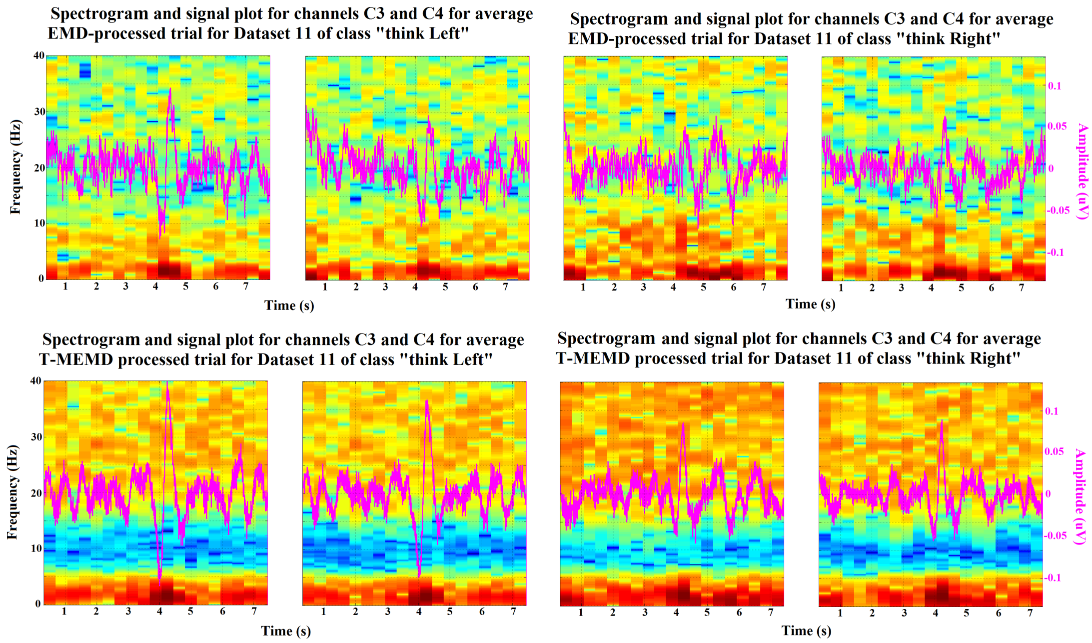

Figure 1: This figure shows the spectrograms for the average processed signal with the waveform itself superimposed in purple. The plots are for electrodes C3 and C4, for the classes "think Left" and "think Right" respectively for the EEG data of User 11. The first row is regular EMD and the bottom row is T-MEMD with the stimulus occurring exactly half way across the x-axis. In this case, T-MEMD has created clearer ERD/ERS peaks.

Method	Dataset 11		Avg., all channels			Avg., C3			Avg., C4		
	Sens.	Spec.	Sens.	Spec.	S.D.	Sens.	Spec.	S.D.	Sens.	Spec.	S.D. (%)
EMD knowledge-based	62.2	53.3	50.3	49.2	10.6	46.5	38.7	24.9	52.1	28.9	30.6
T-MEMD knowledge-based	35.6	37.8	53.2	49.0	11.6	37.5	44.2	28.2	45.6	30.1	29.3
EMD brute force	73.3	75.6	68.9	68.9	5.3	69.0	53.2	16.2	68.4	50.6	15.2
T-MEMD brute force	62.2	71.1	70.7	69.6	6.8	65.0	46.5	20.3	66.8	51.9	14.4

Table 1: The sensitivity, specificity and standard deviation of each EMD and T-MEMD method for a single dataset applied to all channels, the average sensitivity, specificity and standard deviation of each EMD and T-MEMD method for all datasets applied to all channels and applied to a single channel, C3 and C4 respectively.

maximise the variances of one class and minimises them in the other [7]. The features are then input to a Support Vector Machine (SVM). As there were only 90 trials per user it was decided to use the Leave One Out method (LOO). None of the CSP features or brute force classifier made use of the test trials.

3 Results & Analysis

EMD and T-MEMD both achieved very similar performance with the difference between the two being well within their respective standard deviations in every case. The brute force method

 DOI:10.3217/978-3-85125-378-8-13

returned significantly higher performance than the knowledge based method for both decomposition methods. Whilst the IMFs chosen with the brute force method were not consistent between users or decomposition methods, they did focus on the frequency bands known to contain MI information that were identified in the introduction (Figure 1). A single channel analysis of the high-performing brute force method was carried out to further see if embedding temporal data added anything to the EMD process. As CSPs need multiple channels to function the variance of each processed trial was used as the input of the SVM.

As Table 1 shows, there is negligible difference in performance indicating that the added temporal dynamics do not contain any new information. In part, this may be due to the fact that the underlying processes for MI affect all channels, and whilst ERD/ERS and the μ rhythm are expected to be stronger on one side of the brain versus the other, the changes still occur simultaneously for both hemispheres. This implies that the information content for MI is inherent in the lateralisation of the changes - i.e. spatially. In [3] adding temporal dynamics to ICA had a greater impact than with EMD because ICA contains zero temporal information, whilst EMD still uses the data laid out in chronological order. EMD can also only discard background noise if it is unstructured due to its criteria for identifying IMFs.

4 Conclusion

Ultimately the added temporal dynamics did not significantly improve the classification performance. However it might still have some effect on performance if applied to an EEG signal with significant temporally independent features, which is not the case in MI. Another way to enhance the method could be to incorporate both spatial and temporal information by decomposing all channels in their multi-temporal form simultaneously using MEMD.

References

[1] S.R.H. Davies and C.J. James. Novel use of empirical mode decomposition in single-trial classification of motor imagery for use in brain-computer interfaces. In *Engineering in Medicine and Biology Soc. (EMBC), 2013 35th Annual Intl. Conf. of the IEEE*, pages 5610–5613, July 2013.

[2] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.

[3] C. James and D. Lowe. Using dynamical embedding to isolate seizure components in the ictal EEG. *Science, Measurement and Technology, IEE Proceedings -*, 147(6):315–320, Nov 2000.

[4] Cheolsoo Park, D. Looney, N. ur Rehman, A. Ahrabian, and D.P. Mandic. Classification of motor imagery bci using multivariate empirical mode decomposition. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(1):10–22, Jan 2013.

[5] G. Pfurtscheller, C. Brunner, A. Schlgl, and F.H. Lopes da Silva. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, 31(1):153 – 159, 2006.

[6] Gerwin Schalk, Dennis J. Mcfarl, Thilo Hinterberger, Niels Birbaumer, and Jonathan R. Wolpaw. Bci2000: A general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51, 2004.

[7] Zhisong Wang, Alexander Maier, Nikos K. Logothetis, and Hualou Liang. Single-trial classification of bistable perception by integrating empirical mode decomposition, clustering, and support vector machine. *EURASIP J. Adv. Sig. Proc.*, 2008.