Using Empirical Mode Decomposition with Spatio-Temporal Dynamics to Classify Single-Trial Motor Imagery in BCI

Simon R. H. Davies, *Student Member, IEEE*, and Christopher J. James, *Senior Member, IEEE*

Abstract— **This paper introduces a new signal processing method called Spatio-Temporal Multivariate Empirical Mode Decomposition (ST-MEMD). It is a new variation of Empirical Mode Decomposition (EMD) that takes spatial and temporal information into account simultaneously rather than processing each signal source in isolation. The original and new methods were tested on single-trial electroencephalogram data with a two-class problem, in this case data using the Motor Imagery paradigm in brain-computer interfacing. However, whilst ST-MEMD retained the increase in sensitivity and specificity from adding spatial data, the new temporal data made no meaningful difference in terms of performance.**

I. INTRODUCTION

Empirical Mode Decomposition (EMD) applies an iterative sifting process to a signal in order to decompose it into a group of Intrinsic Mode Functions (IMFs) and residual noise (Figure 1) [1]. It can be applied to any non-linear and non-stationary signal allowing it to be used in many different fields, from analyzing mortgage rate data [2] to characterizing non-linear water waves [3]. It functions by subtracting the mean envelope of the signal repeatedly until it produces a signal with a mean envelope that is approximately parallel with the *x*-axis, i.e. the signal's oscillations are symmetrical. This is stored as an IMF and the envelope subtraction continues until no peaks or troughs are left, leaving the remainder of the signal to be classified as residual noise.

However it can only decompose one signal source at a time. Park et al [4] developed a new variation of EMD that can decompose multiple sources of a signal simultaneously called Multi-variate Empirical Mode Decomposition (MEMD), which produced significantly better performance. However, whilst this new method makes use of spatial data to decompose the signal, it does not include wider temporal dynamics. The novel method detailed in this paper uses both spatial and temporal data to inform the decomposition process.

To compare the methods, both were applied to a pre-recorded electroencephalogram (EEG) dataset of 990 trials divided approximately into two equal classes. EEG uses electrodes placed on the scalp to measure changes in voltage caused by brain activity [5]. EEG can be used to construct a Brain-Computer Interface (BCI). A BCI is a device that uses the brain-activity of a person as an input to select desired outputs on a computer [6]. In practice it is mostly used by patients with severe paralysis who cannot make use of keypads, joysticks or eye-gaze technology. The concept used in this dataset was a BCI paradigm called motor imagery. This is suitable as EMD has already been shown to be useful in processing motor imagery for classification [7]. As brain activity as measured at the scalp is of low amplitude and mixed in with other mental 'noise', concepts like motor imagery must be used to help decipher the user's intent. Motor imagery as a control signal exploits the left-right divide of the brain and limb control. When a person imagines moving a limb on the left side of their body there is heavily localized activity on the right side of their motor cortex and vice-versa [8]. This activity also occurs within specific frequency bands. A rhythmic signal in the 8-13 Hz band called the μ rhythm will be suppressed on the contralateral side of the motor cortex. There may also result in resonance activity in the 20 Hz band. At the onset of motor imagery there will be a large desynchronization and resynchronization of brain activity that takes approximately one second.

It is hoped that the novel method's added temporal information will improve performance as it is being applied to an EEG motor imagery dataset. Whilst spatial information is of key importance in EEG signals and motor imagery activity due to the strong spatial localization in the motor cortex, temporal information is also important as rhythmic activity is a dominant presence in EEG signals during motor imagery, making it important that the temporal dynamics of the signal are also captured.

II. METHODOLOGY

A. The Dataset

The EEG recordings used were a pre-recorded dataset of 90 motor imagery trials from 11 different subjects [9]. The recordings were made with a 64-channel EEG with a sampling frequency of 160 Hz. The experiment covered several different movements but this analysis used only the trials that involved imagining opening and closing the left or right fist. Each experimental run consisted of 4.1 seconds neutral activity, followed by a random visual cue on the left or right side of a computer screen. The subject imagines opening and closing the corresponding fist until the cue disappears after another 4.1 seconds. Subjects performed this action 90 times in total with an approximately equal divide between left and right imagery. Subjects eight to eighteen were chosen due to having the same trial sample length of

Research supported by the Warwick Manufacturing Group.

S. R. H. Davies is with the Institute of Digital Healthcare, International Digital Laboratory, University of Warwick, CV4 7AL, UK (Tel: +44 (0) 2476 151341; e-mail: davies_s@wmg.warwick.ac.uk).

C. J. James is with Warwick Engineering in Biomedicine, School of Engineering, University of Warwick, CV4 7AL, UK (Tel: +44 (0) 2476 151341; e-mail: c.james@warwick.ac.uk).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

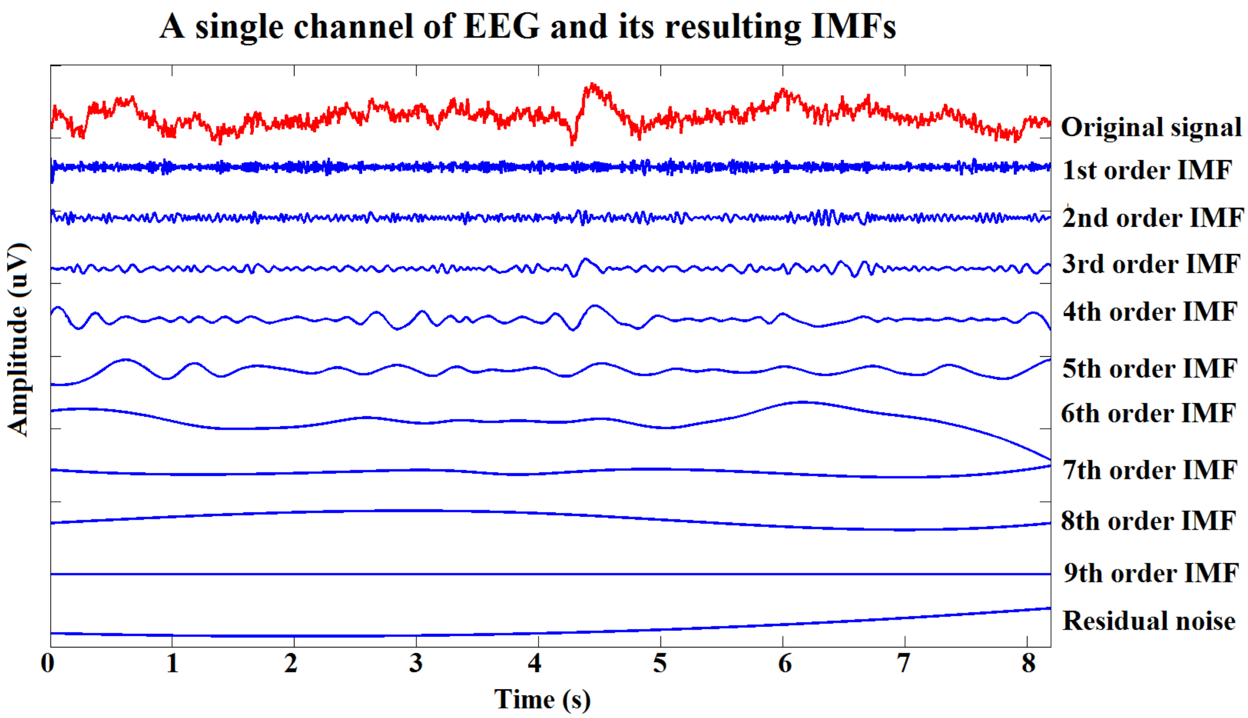

Figure 1. The average of all Dataset 11's "think Left" trial data from channel C3 with EMD applied to it. The IMFs are sorted by frequency in a descending order until only the residual noise is left. Should all the IMFs and noise be summed together they will reform the original signal. This process is useful for extracting features from a signal that are composed of a narrow range of frequencies, such as motor imagery's rhythmic 8-13Hz μ rhythm.

1312 (8.2 s) with the visual cue occurring on the 657th sample. This gives a total of 990 trials.

B. Feature Extraction

EMD works by applying an iterative sifting process using the following steps to the signal, $\mathbf{x}(t)$:

1. Identify the maxima and minima of the signal.
2. Interpolate between the maxima and minima to create upper and lower envelopes.
3. Calculate the mean between the two envelopes, $\mathbf{m}(t)$.
4. Subtract the mean from the signal to get an IMF candidate, $\mathbf{x}_{n+1}(t) = \mathbf{x}_n(t) - \mathbf{m}(t)$.
5. Check if $\mathbf{x}_{n+1}(t)$ is an IMF by calculating if it is symmetrical with respect to zero.
6. If $\mathbf{x}_{n+1}(t)$ is an IMF then store the IMF and return to step 1.) with the signal $\mathbf{x}(t) = \mathbf{x}_n(t) - \mathbf{x}_{n+1}(t)$, else discard and return to step 1.) with the signal $\mathbf{x}(t) = \mathbf{x}_n(t) - \mathbf{x}_{n+1}(t)$.
7. When there are less than two extrema left in the signal the remaining data is classified as the residual.

The new MEMD method applies the following steps to a multi-channel signal:

1. Choose a suitable point set for sampling on an $(n - 1)$ sphere.
2. Calculate a projection, denoted by $\{p^{\theta k}(t)\}_{t=1}^{T}$, of the input signal $\{\mathbf{v}(t)\}_{t=1}^{T}$ along the direction vector $\mathbf{x}^{\theta k}$, for all k (the whole set of direction vectors), giving $\{p^{\theta k}(t)\}_{k=1}^{K}$ as the set of projections.
3. Find the time instants $t_j^{\theta k}$ corresponding to the maxima of the set of projected signals $\{p^{\theta k}(t)\}_{k=1}^{K}$.
4. Interpolate $[t_j^{\theta k}, \mathbf{v}(t_j^{\theta k})]$ to obtain multivariate envelope curves $\{e^{\theta k}(t)\}_{k=1}^{K}$.
5. For a set of K direction vectors, the mean $\mathbf{m}(t)$ of the envelope curves is calculated as $\mathbf{m}(t) = 1/K\Sigma_{k=1}^{K} e^{\theta k}(t)$.
6. Extract the "detail" $\mathbf{c}_i(t)$ using $\mathbf{c}_i(t) = \mathbf{v}(t) - \mathbf{m}(t)$ (i is an order of IMF). If the "detail" $\mathbf{c}_i(t)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $\mathbf{v}(t) - \mathbf{c}_i(t)$, otherwise apply it to $\mathbf{c}_i(t)$.

This results in the channels being decomposed simultaneously, with each channel's IMFs occupying the same frequency band as the other channels'. This makes it much easier to analyse and compare. In this case we are applying MEMD to a dynamically embedded multi-channel signal. Taken's theorem is applied to each signal channel and this converts a single channel of data into multiple snapshots of the signal in time using a series of delay vectors [10],

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

where τ is the lag and m is the embedding dimension. MEMD is applied to the snapshots of every signal all at once, meaning it will use temporal as well as spatial information to decompose the channels. A similar method was used with Independent Component Analysis (ICA) to create an ICA method that could be applied to a single channel [11]. In that work it was found that the embedding dimension needs to be greater than $2D + 1$, where D is the number of signal sources. We know that muscle activity, left hemisphere motor imagery, right hemisphere motor imagery and a variety of background noise will be picked up by the electrodes. As the number of sources is unlikely to exceed single figures a dimension size of $m = 30$ was selected. For ease of operation $\tau = 1$ was used.

The resulting IMFs are un-embedded and two different methods are applied to identify IMFs containing possible motor imagery information – a knowledge-based method and a brute force method. The knowledge-based method tries to identify relevant IMFs in each channel using pre-existing knowledge about the μ rhythm. If 5% of the IMF's total power is in the 8-13 Hz frequency band it is considered as possible motor imagery related activity. The selected IMFs are then summed to form the processed signal. The brute force method works by trying every possible combination of IMF until the highest performance is found. As the maximum number of IMFs produced per channel in this case is 13, and the number of IMFs containing relevant information is usually less than 5, it is possible to process every possible combination of IMF (2379 combinations in this case) and record which resulting classifier gave the best performance.

Common Spatial Patterns (CSPs) are used to extract features from the processed signals of either method. CSPs calculate a set of spatial filters that maximize the variances of one class and minimize them in the other [12]. Finally the extracted features for each class are input into a Support Vector Machine (SVM), using a linear kernel as it is a two-class problem. As each user's dataset had only 90 trials it was decided to use the Leave One Out method (LOO) [13]. This is when $(n - 1)$ trials are used as training data and the n^{th} trial is used to test the resulting classifier. The result is recorded and the test trial is swapped with a training trial until results for all individual trials have been recorded.

III. RESULTS & ANALYSIS

The results in Table 1 for an individual dataset show that both EMD methods with knowledge-based selection result in moderate performance, with MEMD having slightly higher sensitivity and ST-MEMD having slightly lower specificity. This pattern is also present in a reduced fashion in the average of all eleven datasets, but the difference between the two methods is well within their respective standard deviations. The brute force selection method consistently

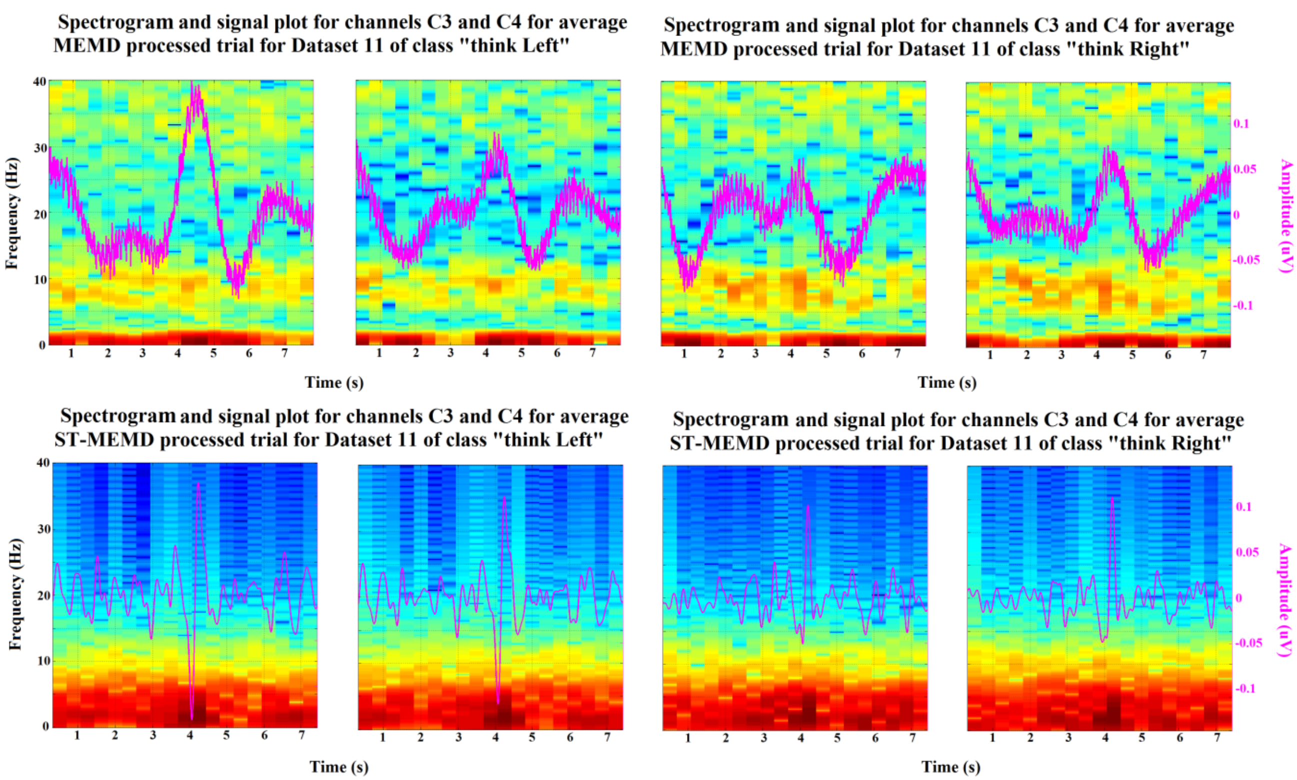

Figure 2. This figure shows the spectrograms for the average processed signal obtained with the brute force method with the waveform itself superimposed in purple. The plots are for electrodes C3 and C4, for the classes "think Left" and "think Right" respectively for the EEG data of User 11. The first row is the original MEMD method and the bottom row is the enhanced ST-MEMD method with the direction stimulus occurring exactly half way across the x-axis. In this case, ST-MEMD has created much clearer ERD/ERS peaks with all the power concentrated in the lower frequencies.

TABLE 1. SENSITIVITY, SPECIFICITY AND STANDARD DEVIATION OF MEMD AND ST-MEMD METHODS WITH EITHER KNOWLEDGE BASED OR BRUTE FORCE SELECTION METHOD APPLIED

Method	Dataset 11		Average of all datasets		
	Sens. (%)	Spec. (%)	Sens. (%)	Spec. (%)	S.D (%)
MEMD knowledge based	40.0	53.3	48.9	52.4	12.6
ST-MEMD knowledge based	60.0	48.9	53.5	50.9	15.1
MEMD brute force	62.2	77.8	76.0	74.7	6.5
ST-MEMD brute force	71.1	64.4	75.7	71.5	8.4

produces higher sensitivity and specificity. In the individual dataset MEMD again has higher sensitivity but lower specificity than ST-MEMD. In the average of all datasets MEMD with brute force is superior in both sensitivity and specificity but again the difference between the two EMD methods is within both their standard deviations.

Overall ST-MEMD achieved similar performance to MEMD with no noticeable improvements in sensitivity or specificity. The brute force method consistently outperformed the knowledge-based method. The IMFs selected by the brute force method were not consistent between datasets but unsurprisingly they did focus on the frequency bands known to contain motor imagery information. In Figure 2 it is possible to clearly identify the ERD/ERS event that occurs at the stimulus time in the ST-MEMD plots.

The difference between the two EMD methods was well within their respective standard deviations, meaning that ST-MEMD still retained the extra spatial information that was added. This suggests that the temporal data added nothing of value. This may be because the signal already contained some temporal data in that it was laid out in a chronological fashion, making the information obtained through the embedding process redundant. Comparatively, the single-channel ICA method mentioned earlier saw an increase in performance when applied to a dynamically embedded signal because the ICA algorithm did not use any temporal information in the extraction process at all.

The negligible performance difference could also be because the underlying motor imagery process affects all channels simultaneously and with the same signal morphology, with only a change in magnitude between channels. This reinforces the idea that the key content of motor imagery is spatial and housed in the lateralized changes in power.

IV. CONCLUSION

In this case there was no change in performance, with no increase in sensitivity or specificity, but no statistically relevant decrease either. This infers that ST-MEMD's IMFS contained no new information but did not lose any either. This is probably due to the temporal data already having been incorporated into both methods due to the envelopes and the resulting IMFs having the same temporal dynamics embedded in the signal.

EMD has many applications outside of EEG and any signals with strong temporal features may yet benefit from ST-MEMD. Alternatively, other multi-channel signal processing methods aside from MEMD and ICA may benefit from dynamical embedding to create new spatio-temporal processing methods. One beneficial outcome of this study is that it has further reinforced the idea that critical motor imagery information has a lot of spatial content, and is not overtly temporal.

REFERENCES

[1] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Royal Society London A*, vol. 454, pp. 903-995, November 1998.

[2] N.E. Huang, M.L. Wu, W. Qu, S.R. Long, and S.S.P. Shen, "Applications of Hilbert–Huang transform to non-stationary financial time series analysis," *Appl. Stochastic Models Bus. Ind.*, vol. 19, pp 245-268, October 2003.

[3] T. Schlurmann, "Spectral Analysis of Nonlinear Water Waves Based on the Hilbert-Huang Transformation," *J. Offshore Mech. Arct. Eng.*, vol. 124, pp 22-27, August 2001.

[4] C. Park, D. Looney, N. Rehman, A. Ahrabian, and D.P. Mandic, "Classification of Motor Imagery BCI using Multivariate Empirical Mode Decomposition," *IEEE Transactions On Neural Systems And Rehabilitation Engineering,* vol. 21, pp. 10-22, January 2013.

[5] S.J.M. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy," *J. Neurol. Neurosurg. Psychiatry*, vol. 76, pp 2-7, June 2005.

[6] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, pp. 767-791, March 2002.

[7] 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*", 35th Annual International Conference of the IEEE EMBC*, pp. 5610-5613, July 2013.

[8] G. Pfurtscheller, C. Brunner, A. Schlogl, and F.H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, pp. 153-159, May 2006.

[9] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, J.R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering,* vol. 51, pp. 1034-1043, June 2004.

[10] F. Takens, "Detecting strange attractors in turbulence," in *Lecture Notes in Mathematics*, D.A. Rand and L.-S. Young, Ed. Berlin: Springer, pp. 366-381, 1981.

[11] C. James and D. Lowe, "Using dynamical embedding to isolate seizure components in the ictal EEG". *Science, Measurement and Technology, IEE Proceedings*, vol. 147, pp. 315-320, November 2000.

[12] Z.J. Coles, M.S. Lazar, and S.Z. Zhou, "Spatial patterns underlying population differences in the background EEG", *Brain Topography,* vol. 2, pp. 275-284, June 1990.

[13] A. Elisseeff, and M. Pontil, "Leave-one-out error and stability of learning algorithms with applications," in *Learning Theory and Practice*, J. Suykens, G. Horvath, S. Basu, C. Micchelli, and J. Vandewalle, Ed. Amsterdam: IOS Press, 2002.