



UNIVERSITY OF LEEDS

This is a repository copy of *A new multivariate empirical mode decomposition method for improving the performance of SSVEP-based brain–computer interface*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/123046/>

Version: Published Version

Article:

Chen, Y-F, Atal, K, Xie, S-Q et al. (1 more author) (2017) A new multivariate empirical mode decomposition method for improving the performance of SSVEP-based brain–computer interface. *Journal of Neural Engineering*, 14 (4). 046028. ISSN 1741-2560

<https://doi.org/10.1088/1741-2552/aa6a23>

(c) 2017, IOP Publishing Ltd. This is an author-created, un-copyedited version of an article published in the *Journal of Neural Engineering*. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at: <https://doi.org/10.1088/1741-2552/aa6a23>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

1 Article

2 A new multivariate empirical mode decomposition 3 method for improving the performance of 4 SSVEP-based brain computer interface

5 Yi-Feng Chen ^{1,2}, Kiran Atal ², Sheng-Quan Xie ² and Quan Liu ^{1,3,*}

6 ¹ School of Information Engineering, Wuhan University of Technology, Wuhan, Hubei 430070 China;
7 cheniyifengwhut@126.com

8 ² Mechanical Engineering, University of Auckland, Auckland, New Zealand; kata530@aucklanduni.ac.nz;
9 s.xie@auckland.ac.nz

10 ³ Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education,
11 Wuhan University of Technology, Wuhan, Hubei 430070 China; quanliu@whut.edu.cn

12 * Correspondence: quanliu@whut.edu.cn

13 Academic Editor:

15 **Abstract:** *Objective.* Accurate and efficient detection of steady-state visual evoked potentials
16 (SSVEP) in electroencephalogram (EEG) is essential for the related brain-computer interface (BCI)
17 applications. *Approach.* Although the canonical correlation analysis (CCA) has been applied
18 extensively and successfully to SSVEP recognition, the spontaneous EEG activities and artifacts
19 that often occur during data recording can deteriorate the recognition performance. Therefore, it is
20 meaningful to extract a few frequency sub-bands of interest to avoid or reduce the influence of
21 unrelated brain activity and artifacts. This paper presents an improved method to detect the
22 frequency component associated with SSVEP using multivariate empirical mode decomposition
23 (MEMD) and CCA (MEMD-CCA). EEG signals from 9 healthy volunteers were recorded to
24 evaluate the performance of the proposed method for SSVEP recognition. *Main results.* We
25 compared our method with CCA and temporally local multivariate synchronization index (TMSI).
26 The results suggest that the MEMD-CCA achieved significantly higher accuracy in contrast to
27 standard CCA and TMSI. It gave the improvements of 1.34%, 3.11%, 3.33%, 10.45%, 15.78%,
28 18.45%, 15.00% and 14.22% on average over CCA at time windows from 0.5 s to 5 s and 0.55%,
29 1.56%, 7.78%, 14.67%, 13.67%, 7.33% and 7.78% over TMSI from 0.75 s to 5 s. The method
30 outperformed the filter-based decomposition (FB), empirical mode decomposition (EMD) and
31 wavelet decomposition (WT) based CCA for SSVEP recognition. *Significance.* The results
32 demonstrate the ability of our proposed MEMD-CCA to improve the performance of SSVEP-based
33 BCI.

34 **Keywords:** Brain computer interface; canonical correlation analysis; electroencephalogram;
35 multivariate empirical mode decomposition; steady-state visual evoked potentials.
36

37 1. Introduction

38 Brain computer interface (BCI) provides an advanced communication method to people
39 suffering from severe motor disabilities with external environment via measures of brain activity [1].
40 In the past few years, a range of BCI systems have been investigated including P300 [2, 3], steady
41 state visual evoked potential (SSVEP) [4, 5], motor imagery [6] and hybrid BCI system [7, 8]. In
42 particular, SSVEP-based BCIs have attracted widespread interest because of its high information
43 transfer rate (ITR), high signal-to-noise ratio (SNR), and minimal training [9, 10]. SSVEP is defined
44 as periodic evoked potentials elicited by rapidly oscillating visual stimulus, typically exhibiting the
45 same frequency as the target as well as its harmonics. SSVEP-based BCI system mainly contains
46 several visual stimuli flickering at distinct frequencies. When users focus attention on one of the

1 frequency-coded stimuli, there will be a corresponding component increase in SSVEP signals at the
2 same frequency. The target thus can be determined by analyzing the frequency information of
3 SSVEP signals.

4 In current SSVEP-based BCI systems, the most commonly used technique to obtain the SSVEP
5 is electroencephalogram (EEG) because of its very high temporal resolution and convenience [11].
6 The EEG noninvasively records electrical activity of the brain by placing the electrodes along the
7 scalp. Therefore, the SSVEP recognition highly depends on algorithms for features extraction from
8 recorded EEG signals. Various methods have been proposed to extract the SSVEP information from
9 EEG in the real BCI application [9, 12, 13]. The power spectrum density analysis (PSDA) was widely
10 used for frequency detection corresponding to SSVEP in BCI systems [14]. The spectrum of EEG can
11 be estimated using fast Fourier transform (FFT) method with low computational cost, where the
12 peak of the power spectrum density is considered as the SSVEP component induced by the visual
13 target. However, this method is sensitive to artifacts in EEG and needs long window size for
14 spectrum estimation with high frequency resolution [15]. Canonical correlation analysis (CCA) has
15 been validated as a more efficient technique than PSDA for frequency recognition [16]. CCA as a
16 multivariable statistical algorithm can estimate the degree of correlation within two multivariate
17 signals. The largest coefficient between recorded EEG and pre-constructed reference signals is used
18 to identify the target. The reference signals comprise a series of sine-cosine waves at the
19 fundamental and higher harmonic frequencies of stimulus. Multivariate synchronization index
20 (MSI) [17] detects the SSVEP by measuring the synchronization between recorded EEG and
21 pre-constructed reference signals with the same frequency of visual stimuli. The temporally local
22 MSI (TMSI) [4] improves the original MSI using temporally local information of signals. And
23 likelihood ratio test (LRT) [18] is another tool to measuring the correlation between two set of
24 multivariate signals and has been used in frequency detection of SSVEP-based BCIs. They all could
25 achieve better performance than CCA.

26 The CCA-based method provides an improvement in SNR of SSVEP signals, thus achieves
27 significantly higher recognition accuracy, and has been widely used in BCI system because of its
28 efficiency and robustness [19, 20]. However, reference signals of sinusoidal waves lack sufficient
29 information from real SSVEP. Therefore, several extensions of CCA algorithms have been
30 introduced to optimize reference signals with calibration in order to enhance the accuracy. Bin et al.
31 [21] developed a BCI system based on code modulation VEPs, in which VEP templates obtained by
32 averaging the EEG over multiple stimulus cycles were used as the references. A multiway CCA
33 (MCCA) method considered optimized reference signals from space and trial data modes instead of
34 the original sinusoidal reference signals [22]. A L1-regularized multiway canonical correlation
35 analysis (L1-MCCA) was further proposed for reference signal optimization [23]. In addition, a
36 multiset canonical correlation analysis (MsetCCA) [24] and a common feature analysis (CFA) [25]
37 achieved reference signal optimization by extracting the common features completely based on
38 training data. All these extended CCA approaches focus on optimizing reference signals and have
39 been successfully implemented to improve the identification accuracy compared with the standard
40 CCA in SSVEP-based BCI. Nevertheless, the performance depends not only upon reference signals
41 but also upon EEG data. The spontaneous EEG unrelated to stimulus could deteriorate the detected
42 accuracy. Besides, artifacts in EEG may inevitably result in misleading detection. Therefore, there is
43 a need to further improve the accuracy by optimizing the EEG signals.

44 In this study, we assume that the original EEG signals contain several non-overlapping
45 frequency sub-bands, and the SSVEP related brain activities are located within some sub-bands
46 rather than whole frequency band. The extracted signal occupying such sub-bands may bear a
47 higher SNR, thus performs more efficiently in using CCA for SSVEP recognition. The traditional
48 band-pass filters based on Fast Fourier Transform (FFT) divides a signal into several separate
49 frequency domains with predefined bandwidth. The wavelet transform (WT) successively
50 decomposes the signal into different frequency components. However, the EEG signals generated
51 by human brain, a complex nonlinear system, are essentially nonlinear and non-stationary. The
52 linear methods like FFT and WT are based on predefined basis functions and usually require

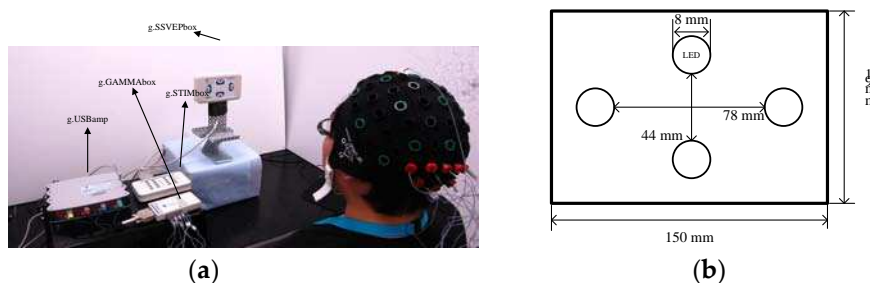
1 assumption of stationarity, thus they are inappropriate to deal with the highly complex EEG
 2 showing nonlinear dynamics and chaos, although they may work well for short time intervals.
 3 Therefore, new nonlinear methods are needed to analyze signals exhibiting nonlinear and
 4 non-stationary characteristics. Empirical mode decomposition (EMD) [26] is one of such techniques.

5 EMD is a completely data-driven technique based on local properties of the signal. It
 6 adaptively decomposes the signal into a set of components called intrinsic mode functions (IMFs)
 7 through a sifting process without prior assumptions on the data. The IMFs indicate oscillation
 8 modes in the signal and are complete and almost orthogonal, thus they can serve as the basis
 9 functions derived from the data itself. This decomposition has been widely used to analyze the
 10 nonlinear and non-stationary processes, such as bio-signals, climate and earthquake [27-29], etc.
 11 Corresponding analyses show that EMD acts essentially as dyadic filter bank within the
 12 decomposed IMFs. Choice of IMFs in which SSVEP is more prominent according to time–frequency
 13 analysis of EEG via the Hilbert transform [26] provides a potential optimization for features
 14 detection from frequency sub-bands. However, in EMD technique, there are two obstacles: mode
 15 mixing in single-channel decomposition and mode misalignment in multiple-channel
 16 decomposition. The multivariate empirical mode decomposition (MEMD) has been proposed to
 17 better align the corresponding IMFs of multichannel signals [30]. In addition, by adding extra white
 18 noise channels, the mode mixing problem is reduced and the IMFs are forced to follow a dyadic
 19 filter bank structure. It is the so-called noise-assisted MEMD (N-A MEMD) [31]. Taking into account
 20 both properties of narrowband IMFs and multichannel nature, we propose a MEMD-CCA
 21 approach to improve the accuracy of SSVEP recognition.

22 2. Materials and Methods

23 2.1. EEG collection

24 Nine healthy subjects (6 males and 3 females) with normal or corrected-to-normal vision from
 25 the University of Auckland participated in the study. Ethical approval was obtained from the
 26 University of Auckland Human Participants Ethics Committee (UAHPEC). Written informed
 27 consents were also received from all subjects. All experiments were managed by g.BCIsys (g.tec
 28 medical engineering GmbH, Austria), which is g.tec's BCI research environment as shown in Figure
 29 1. The bio-signal acquisition system g.USBamp was used to record the EEG data sampled at 256 Hz.
 30 The data were further passed through a band-pass filter with a bandwidth of 0.5-60 Hz, and a notch
 31 filter at 50 Hz. Eight EEG channels including “PO7”, “PO3”, “POz”, “PO4”, “PO8”, “O1”, “Oz” and
 32 “O2” according to the extended international 10–20 system were selected for analysis. The right
 33 earlobe electrode served as reference channel, and “FPz” was defined as ground. Visual stimuli
 34 were generated using four white colored LEDs with 8 mm diameter on the SSVEP box (i.e.,
 35 g.SSVEPbox) controlled by the Paradigm Control and g.STIMbox blocks at the flashing frequencies:
 36 10, 11, 12 and 13 Hz, respectively.



37
 38
 39 **Figure 1.** Experimental setup. (a) Overview of participant’s location at the table and g.BCIsys. The
 40 distance between the subjects and the target is about 70 cm. The g.USBamp is an amplifier and
 41 acquisition system. The g.STIMbox is used to generate and record trigger signals. The g.SSVEPbox is
 42 a stimulation device for SSVEP. And g.GAMMAbox is used for power supply and driver/interface
 43 box for 16 active electrodes. (b) The layout of SSVEP stimulation box (i.e., g.SSVEPbox) with four
 44 bright white LEDs (type WU-2-104WD, diameter 8 mm each with luminous intensity of 1500 mcd).

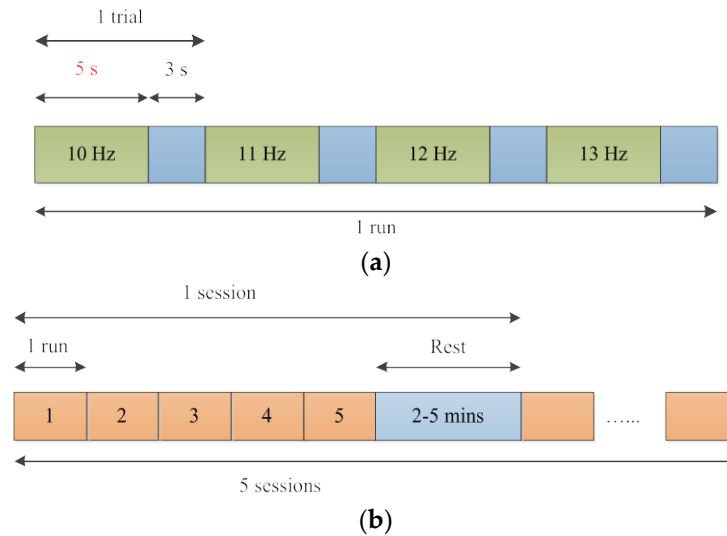


Figure 2. Illustration of the whole experiment. (a) The experiment process of 1 run. Each run contains 4 trials corresponding to 4 different frequencies, and in each trial, the subjects are asked to gaze at one of the targets for 5 s, rest for 3 s before shifting to the next target. (b) The experiment process of 1 session and the whole experiment. Each session includes 5 runs. Between sessions, the subjects can relax for 2-5 mins to avoid the visual fatigue. For each subject, the whole experiment consists of 5 sessions.

The whole experiment is presented in Figure 2. All participants were seated in a comfortable chair facing the SSVEP visual stimulation box with a distance of 70 cm. Each subject completed 5 sessions, each containing 5 runs with a rest between 2-5 mins as shown in Figure 2(b). After 5 runs, the subject has a 2-5 mins rest. Each run includes 4 trials associated with four stimulus frequencies, and the current target is marked by a small green LED. Subjects stare at the current stimulus for 5 s from the beginning of corresponding cue during each trial. They then have a 3 s rest before moving attention to the next target. The synchronized trigger signals indicating the subjects concentrate on a target LED or have a rest were also recorded on an event channel. A total of 100 trials (4 trials \times 5 runs \times 5 sessions) data were recorded from each subject.

2.2. Subbands decomposition

2.2.1. Fourier decomposition

The Fourier transform decomposes a time series into sinusoidal functions with different frequencies. A series of linear band-pass filters is able to divide the input signal into a number of sub-bands. The eighth-order type I Chebyshev filters [32] are designed in this study to implement the band-pass filters. The filtering process is applied to each channel of EEG separately. Then the components in the predefined pass-bands are extracted from original EEG signals. Since this type of filter as one of the classical infinite impulse response (IIR) filters has nonlinear phase distortion, so the function “filtfilt” in Matlab (v7.13, MathWorks Inc., Natick, MA, USA) is used to compensate for the distortion. This allows for design and implementation of zero-phase band-pass filters.

2.2.2. Wavelet decomposition

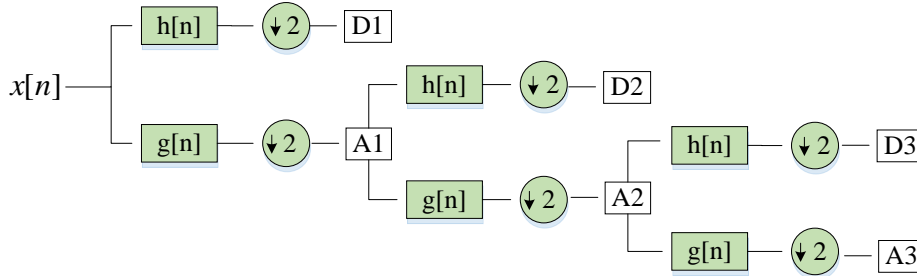
The Fourier transform decomposes a signal into sine and cosine waves, while wavelet transform (WT) decomposes the signal into coefficients with respect to the wavelet functions [33]. The wavelet coefficients are calculated to represent the correlation between the wavelets and signal. The WT can provide both time and frequency information simultaneously through translations and dilations of the mother wavelet and is defined as:

$$F(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \Psi_{(a,b)}^*(t) dt \quad (1)$$

1 with

$$2 \quad \Psi_{(a,b)}(t) = \Psi\left(\frac{t-b}{a}\right) \quad (2)$$

3 where $f(t)$ is the original signal, $*$ represents the complex conjugate function, $\Psi(x)$ is the mother
4 wavelet function. Wavelet basic functions $\Psi_{(a,b)}(t)$ are translation and dilation versions of the
5 mother wavelet at a and b .



6

7

Figure 3. Illustration of a 3 level decomposition of WT.

8 In practical applications, the discrete wavelet transform (DWT) using discrete scales and
9 translations is commonly employed to separate a given signal into orthogonal set of wavelets and
10 results in finite coefficient sets. Figure 3 presents typically dyadic wavelet decomposition. A series
11 of half band filters are used to compute the DWT of a given signal $x[n]$, $h[n]$ are the half band
12 high-pass filters and $g[n]$ are the half band low-pass filters, respectively. Half frequencies of the
13 input signal are removed after filtering, so half number of points is redundant according to
14 Nyquist's rule, and then the outputs are down-sampled by 2. The high-pass and low-pass filters
15 results in detail coefficients (i.e., D) and approximation coefficients (i.e., A) at each level. Each of the
16 coefficient sets represents activities of origin signal in certain sub-bands.

17 2.2.3. Empirical mode decomposition

18 Empirical mode decomposition (EMD) is a data-driven and adaptive algorithm for analyzing
19 nonlinear and non-stationary data according to local characteristic time scale [26]. This method
20 decomposes any complicated time series into a set of IMFs. Given any time series $x(t)$, the EMD
21 algorithm are outlined below:

22 (1) Identify all local maxima and minima of $x(t)$.

23 (2) Interpolate all maxima with cubic splines to obtain an upper envelope $e_{\max}(t)$.

24 (3) Interpolate all minima with cubic splines to obtain a lower envelope $e_{\min}(t)$.

25 (4) Compute local mean by averaging envelopes

$$26 \quad m_1(t) = (e_{\max}(t) + e_{\min}(t))/2 \quad (3)$$

27 (5) Subtract the local mean from the original data to extract the first component

$$28 \quad h_1(t) = x(t) - m_1(t) \quad (4)$$

29 (6) Check the properties of $h_1(t)$. If $h_1(t)$ fulfills the requirements of IMF, $h_1(t)$ is an IMF, it is
30 designated as

$$31 \quad c_1(t) = h_1(t) \quad (5)$$

32 If $h_1(t)$ does not satisfy an IMF, then $h_1(t)$ is treated as the original data, that is

$$33 \quad x(t) = h_1(t) \quad (6)$$

34 And repeat above procedure until $h_1(t)$ becomes an IMF.

35 (7) Separate the IMF $c_1(t)$ from the data, the residual is designated as

$$36 \quad r_1(t) = x(t) - c_1(t) \quad (7)$$

37 (8) Treat the residual $r_1(t)$ as the new data, that is

$$38 \quad x(t) = r_1(t) \quad (8)$$

(9) Repeat the procedure from (1) to (8) on all the subsequent residuals, until the residual satisfies the stopping criterion, the result is

$$r_2(t) = r_1(t) - c_2(t) = \cdots = r_n(t) - r_{n-1}(t) - c_n(t) \quad (9)$$

Finally, the signal $x(t)$ can be decomposed into n IMF and a residual expressed as follows:

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (10)$$

Generally, an effective IMF need to satisfy two conditions including: 1) the numbers of extrema and zero-crossings are the same or differ at most by 1, and 2) the average of upper envelope and lower envelope is zero at any point [26]. The first condition guarantees the traditional narrow band requirement, and the second condition innovatively defines a local requirement by local maxima and minima instead of the classical global one. With these two conditions, a meaningfully instantaneous frequency can be calculated to obtain localized time-frequency spectrogram through the Hilbert transform

$$H(c_i(t)) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{c_i(t')}{t-t'} dt' \quad (11)$$

from each IMF, where $c_i(t)$ is the i IMF, P is Cauchy principal value. Then the analytic signal $Z_i(t)$ is defined as

$$Z_i(t) = c_i(t) + jH(c_i(t)) = a_i(t)e^{j\theta_i(t)} \quad (12)$$

the instantaneous frequency is given by

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \quad (13)$$

2.2.4. Multivariate empirical mode decomposition

Multivariate empirical mode decomposition is a recently proposed method to extend the standard EMD for analyzing multivariate signal [30]. EMD separately decomposes each channel of a multivariate signal, but this decomposition may result in mode nonalignment which means that the total number of IMFs and scale properties in the same-index IMFs derived from different channels can be different. MEMD simultaneously decomposes multichannel data ensuring better alignment of corresponding IMFs from different channels, which will benefit narrowband SSVEP detection with broadband spontaneous EEG. For multivariate data, the local mean by averaging upper and lower envelopes cannot be defined directly due to no clear definition for local maxima and minima of complex signals. To solve this problem, projections of the input signal along different directions in multidimensional space are calculated. Then extrema of those projections are identified and interpolated to obtain the multivariate envelopes. Finally, local mean can be computed by averaging multiple envelopes. Given a n -channel multivariate signal $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots, x_n(t)\}$, MEMD algorithm is summarized as follows:

(1) Generate a suitable set of direction vectors \mathbf{s}^k using low discrepancy Hammersley sequences [30] on an $(n-1)$ sphere, that is

$$\mathbf{s}^k = \{s_1^k, s_2^k, \dots, s_n^k\} \quad k = 1, 2, \dots, K \quad (14)$$

where θ_k is the direction angle of corresponding direction vector, defined by

$$\theta^k = \{\theta_1^k, \theta_2^k, \dots, \theta_{n-1}^k\} \quad (15)$$

(2) Calculate projection of the input signal along each direction vector, then we obtain a set of projections for all k , denoted by $p^k(t)$, $k = 1, 2, \dots, K$.

(3) Identify all maxima and their corresponding time instants $t_i^{\theta_k}$ of the projection $p^k(t)$ for all k .

(4) Interpolate $[t_i^{\theta_k}, \mathbf{x}(t_i^{\theta_k})]$ for all k to generate multivariate envelope $\mathbf{e}^k(t)$, $k = 1, 2, \dots, K$.

(5) Calculate the mean of envelopes by

$$\mathbf{m}(t) = \frac{1}{K} \sum_{k=1}^K \mathbf{e}^k(t) \quad (16)$$

(6) Subtract $\mathbf{m}(t)$ from original signal to obtain the detail component $\mathbf{d}(t)$, given by

$$\mathbf{d}(t) = \mathbf{x}(t) - \mathbf{m}(t) \quad (17)$$

(7) Check the properties of $\mathbf{d}(t)$ to sift appropriate IMFs. If $\mathbf{d}(t)$ fulfills the requirements for multivariate IMF, repeat the above steps to $\mathbf{x}(t) - \mathbf{d}(t)$, otherwise repeat them to $\mathbf{d}(t)$ until all projected signals satisfy the stop criteria. The requirements for multivariate IMF are similar to definition in EMD, while the constraint for numbers of extrema and zero-crossings is not required.

To further circumvent the mode mixing problem in EMD. Extra channels containing multivariate independent white noise are introduced into original signal. Because of the broadband property of white noise and dyadic filter bank behavior of MEMD on white noise, IMFs corresponding to the original signal also exhibits a quasi-dyadic structure enforced by the extra noisy channels, in turn reducing the mode mixing problem and the spectral overlap between IMFs [31]. Therefore, MEMD provides a potentially useful technique for feature extraction and fusion of narrowband but non-stationary activities from multichannel signals, for instance SSVEP detection in multichannel EEG.

In fact, it should be named noise-assisted MEMD (N-A MEMD) method instead of the original MEMD by adding uncorrelated Gaussian white noise channel to the input multivariate signal in the process of decomposition. Different with original MEMD, in N-A MEMD, m -channel independent white noise is firstly created and added into the original n -channel inputs, obtaining a $(m + n)$ -channel combined signal. And then MEMD is used to decompose such a signal to get multivariate IMFs. Finally, the IMFs corresponding to the channels of original signal are extracted. However, it should be mentioned that the so-called noise-assisted methods also use the MEMD to decompose such a combined signal including original inputs and extra noise channels. So in this study, we use the name of MEMD instead of N-A MEMD in spite of adding extra noise in all decompositions.

2.3 Canonical correlation analysis

2.3.1. Standard CCA

Canonical correlation analysis is a tool for measuring the linear relationship between two sets of variables. Given two multivariate variables \mathbf{X} and \mathbf{Y} , their respective linear combinations designed as x and y can be given by

$$x = \mathbf{w}_x^T \mathbf{X} \quad (18)$$

and

$$y = \mathbf{w}_y^T \mathbf{Y} \quad (19)$$

where \mathbf{w}_x and \mathbf{w}_y are two corresponding weight vectors. The CCA seeks \mathbf{w}_x and \mathbf{w}_y to maximize the correlation between x and y . This indicates that the parameter of correlation to be maximized is

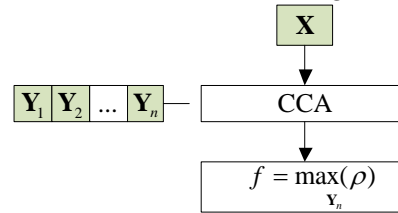
$$\rho(x, y) = \frac{E[xy^T]}{\sqrt{E[xx^T]}E[yy^T]} = \frac{E[\mathbf{w}_x^T \mathbf{X} \mathbf{Y}^T \mathbf{w}_y]}{\sqrt{E[\mathbf{w}_x^T \mathbf{X} \mathbf{X}^T \mathbf{w}_y]}E[\mathbf{w}_y^T \mathbf{Y} \mathbf{Y}^T \mathbf{w}_y]} \quad (20)$$

In SSVEP recognition using CCA [16], \mathbf{X} corresponds to the multichannel EEG and \mathbf{Y} is the predefined reference signal corresponding to the n th stimulus frequency. Assume the data length is L and the sampling rate is f_s , the reference signal can be defined as

$$\mathbf{Y} = \mathbf{Y}_n = \begin{bmatrix} \sin(2\pi f_n t) \\ \cos(2\pi f_n t) \\ \vdots \\ \sin(2\pi N f_n t) \\ \cos(2\pi N f_n t) \end{bmatrix}, \quad t \in \left[\frac{1}{f_s}, \frac{2}{f_s}, \dots, \frac{L}{f_s} \right] \quad (21)$$

where f_n is the fundamental frequency of a certain target, N is the number of harmonics. CCA computes the correlation between \mathbf{X} and \mathbf{Y} for all stimulus frequencies respectively, and then the

1 SSVEP frequency is recognized as that of reference with maximal correlation value. Figure 4
 2 presents the flowchart of the standard CCA for SSVEPs recognition.

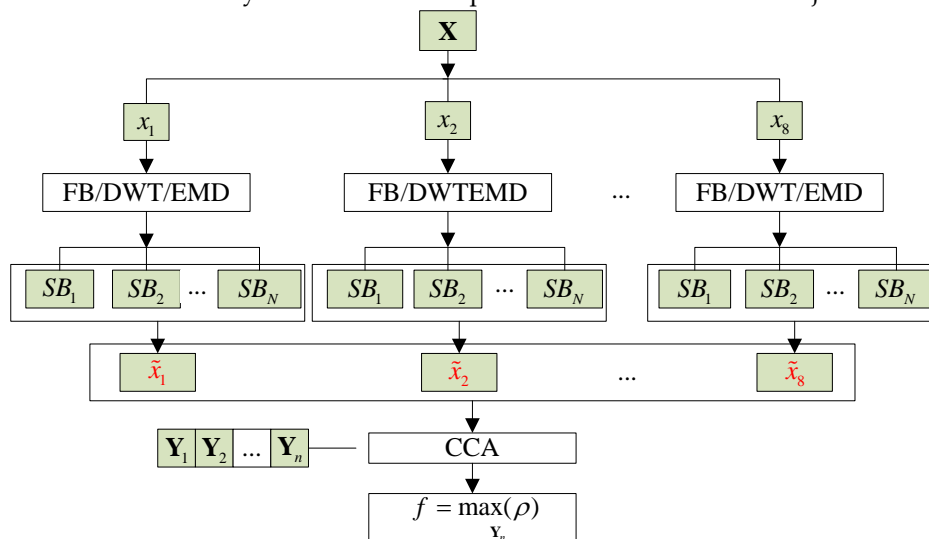


3

4 **Figure 4.** Flowchart of the standard CCA method for frequency detection of SSVEPs. \mathbf{X} is the
 5 recorded EEG data and \mathbf{Y}_n is the n th reference signal. The SSVEP frequency is recognized as the
 6 frequency of \mathbf{Y}_n that maximizes correlation.

7 2.3.2. Sub-band CCA

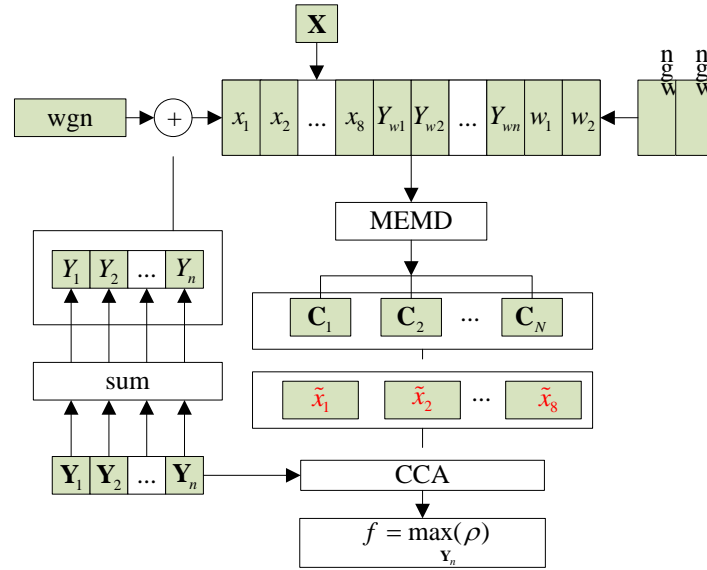
8 The standard CCA detects the frequencies of SSVEPs in the whole frequency band of EEG. In
 9 this case, the spontaneous EEG and artifacts reduce the SNR and recognition accuracy of SSVEPs.
 10 Therefore, a sub-band CCA provides a solution to further improve the performance in
 11 SSVEPs-based BCI. The specific benefits of certain sub-bands containing the discriminative
 12 information of SSVEPs are investigated. This study proposes a MEMD-CCA approach to improve
 13 SSVEP recognition accuracy in sub-bands instead of the whole band of EEG. In addition, we
 14 compare MEMD with traditional filter-based decomposition (FB), discrete Wavelet decomposition
 15 (WT) and EMD in sub-bands extraction and frequency detection in CCA. Figure 5 shows the
 16 flowchart of the FB-CCA, WT-CCA and EMD-CCA methods for frequency detection of SSVEPs.
 17 These methods all firstly decompose each channel of the multivariate EEG into several sub-bands,
 18 then the components corresponding to the SSVEPs are reconstructed for further frequency
 19 detection in CCA. Figure 6 gives the flowchart of the MEMD-CCA for SSVEPs recognition. Unlike
 20 FB-CCA, WT-CCA and EMD-CCA methods, MEMD-CCA simultaneously separates the
 21 multivariate signal as a whole into multivariate components. Moreover, MEMD considers the
 22 consistency of decomposition across channels by adding two extra channels containing white noise
 23 to align the separated components. In addition, by introducing another four reference channels,
 24 each of which compromises the sum of reference signal corresponding to target and white noise,
 25 MEMD ensures the consistency of extracted EEG patterns across trails and subjects.



26

27 **Figure 5.** Flowchart of the FB-CCA, WT-CCA and EMD-CCA methods for frequency detection of
 28 SSVEPs. Assume eight channel EEG signals are recorded and denoted as $\mathbf{X} = \{x_1, x_2, \dots, x_8\}$. Each of
 29 them is first decomposed into N sub-bands by FB, DWT and EMD methods, designed as
 30 SB_1, SB_2, \dots, SB_N . Then the sub-bands related to SSVEP are reconstructed for further frequency

1 detection via CCA. The reconstructed signal is defined as $\tilde{\mathbf{X}} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_8\}$. \mathbf{Y}_n is the predefined
 2 reference signal corresponding to the n th stimulus frequency defined in Eq(21).



3
 4 **Figure 6.** Flowchart of the MEMD-CCA method for frequency detection of SSVEPs.
 5 $\mathbf{X} = \{x_1, x_2, \dots, x_8\}$ is the recorded EEG, \mathbf{Y}_n is the n th reference signal and “wgn” is the white noise.
 6 A composite signal is first constructed by adding extra channels contain two channel white noise
 7 and n channel reference signals. Then MEMD is employed to decompose such combined signal. The
 8 IMFs corresponding to the original EEG are reserved to extract sub-bands related to SSVEP. Note
 9 that \mathbf{Y}_n may be a multivariate signal containing several harmonics, it is converted into a
 10 one-dimensional data Y_n by summing up all harmonic components. Then each of them is
 11 corrupted by added white noise to generate extra reference channels.

12 3. Results

13 3.1 Performance evaluation

14 The classification accuracy is calculated to evaluate the performance of our proposed
 15 MEMD-based method for SSVEP recognition. A comparison of our improved method with
 16 standard CCA and TMSI is performed. In addition, in order to illustrate the advantages of MEMD,
 17 we compare it with traditional filter-based decomposition, WT and EMD methods described in
 18 previous sections for SSVEP detection. The validation is estimated using 100 trials data from each
 19 subject. The paired sample t -test in IBM SPSS Statistics (Version 23.0, Armonk, NY: IBM Corp.) is
 20 used to indicate their statistically significant difference of averaged accuracy across all subjects.
 21 Furthermore, Bonferroni correction is employed for multiple comparisons. All tests are two tailed,
 22 and the statistical significance level is $p < 0.05$.

23 3.2. Time-frequency analysis of SSVEP

24 Owing to the dyadic structure of WT, EMD and MEMD on a decomposed signal, the original
 25 EEG is separated into seven components to match the frequency properties of SSVEP. Both WT and
 26 EMD act as a dyadic filter bank on one-channel signal while MEMD acts on each channel of
 27 multivariate signal as a dyadic filter bank. A collection of dyadic band-pass filters are also
 28 implemented with type I Chebyshev. Table 1 shows the corresponding frequency ranges of
 29 separated components from EEG signal in terms of rhythmic activity. The collected EEG signal is
 30 sampled at 256 Hz, so the highest frequency included in the signal is 128 Hz according to Nyquist
 31 sampling theorem. In this case, the dyadic tiling in frequency domain approximately reflects the
 32 typical bandwidths in EEG known as delta-band (0–4 Hz), theta-band (4–8 Hz), alpha-band (8–16

1 Hz), beta-band (16–32 Hz), gamma-band (32–64 Hz), and the EMG band (64-128 Hz). The exact
 2 boundaries of these rhythms are not consistent across studies, while in this study the frequency
 3 ranges in decomposition is conducted according to Gajic, et al. [34] and Nguyen-Ky, et al [35]. The
 4 components B1-B7 are decomposed by a set of non-overlap band-pass filters using type I
 5 Chebyshev, D1-D7 are decomposed by DWT method, while C1-C7 are decomposed by EMD or
 6 MEMD method.

7 **Table 1.** Frequencies of EEG corresponding to decomposition levels.

Sub-band	Clinical band	Frequency range(Hz)	Filter-based decomposition	Wavelet decomposition	EMD/MEMD
1	EMG	64-128	B1	D1	C1
2	Gamma	32-64	B2	D2	C2
3	Beta	16-32	B3	D3	C3
4	Alpha	8-16	B4	D4	C4
5	Theta	4-8	B5	D5	C5
6	Delta	2-4	B6	D6	C6
7	Delta	0-2	B7	D7	C7

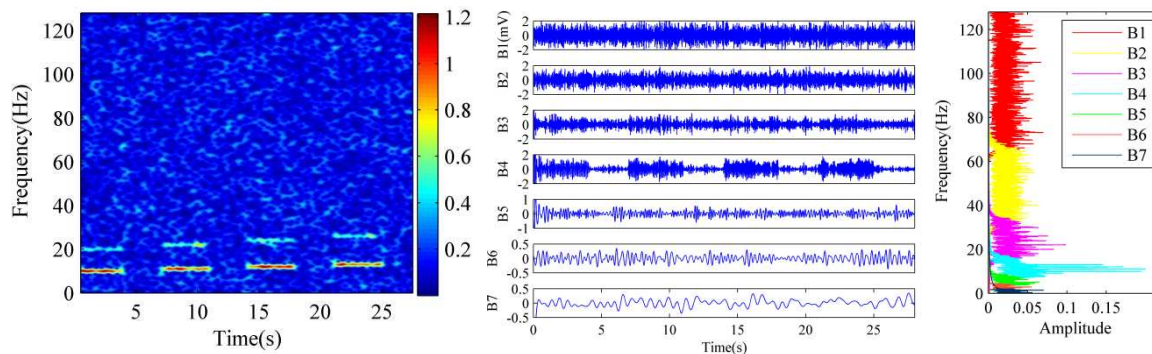
8

9 We illustrate the operation of filter-based decomposition (FB), discrete Wavelet decomposition
 10 (WT), EMD and MEMD in simulated data decomposition, the simulation signal is given by

$$11 \quad s(t) = s_1(t) + s_2(t) + s_3(t) + s_4(t) + wgn(t) \quad (22)$$

$$12 \quad \text{where} \quad \begin{cases} s_1(t) = \sin(2\pi * f_1 * t) + 0.5 * \sin(2\pi * 2f_1 * t) & 0 \leq t \leq 4 \\ s_2(t) = \sin(2\pi * f_2 * (t-7)) + 0.5 * \sin(2\pi * 2f_2 * (t-7)) & 7 \leq t \leq 11 \\ s_3(t) = \sin(2\pi * f_3 * (t-14)) + 0.5 * \sin(2\pi * 2f_3 * (t-14)) & 14 \leq t \leq 18 \\ s_4(t) = \sin(2\pi * f_4 * (t-21)) + 0.5 * \sin(2\pi * 2f_4 * (t-21)) & 21 \leq t \leq 25 \end{cases}, \quad f_1 = 10Hz, \quad f_2 = 11Hz, \quad f_3 = 12Hz \quad \text{and} \quad f_4 = 13Hz$$

13 $f_3 = 12Hz$ and $f_4 = 13Hz$ are four stimulus frequency of the target. The SSVEPs are simulated with
 14 four sinusoidal signals and its first harmonic at different time delay (i.e., 0 s, 7 s, 14 s and 21 s). And
 15 $wgn(t)$ is the white noise to simulate spontaneous EEG activities. For MEMD, the one channel
 16 signal is extended to eight channels and each channel is generated using Eq (22). Then two extra
 17 independent white noise channels are added into the eight-channel simulation signal. MEMD
 18 decomposes such a combined signal. The IMFs associated with the first channel are reserved for
 19 further analysis. A comprehensive comparative study of FB, WT, EMD and MEMD to produce
 20 spectrogram estimates is shown in Figure 7. The first column presents the time-frequency
 21 spectrograms of the simulation signal using short-time Fourier transform (STFT), continuous
 22 wavelet transform (CWT), the Hilbert–Huang spectrogram (HHS) based on EMD and MEMD. STFT
 23 is calculated using the Hamming window with 1s data length and 128 data points overlap. And the
 24 complex Morlet wavelet is applied to compute CWT. The wavelet has a default effective support of
 25 $[-8, 8]$ used in Matlab (v7.13, MathWorks Inc., Natick, MA, USA). It is constructed with a bandwidth
 26 parameter of 3 and a wavelet center frequency of 3. The length of the wavelet is taken to be 256. The
 27 time-frequency spectrograms of the simulation signal obtained using all four methods present
 28 obvious stimulus frequencies for the duration of the SSVEP task, while HHS based on EMD and
 29 MEMD is more localized. However, EMD exhibits more mode mixing.

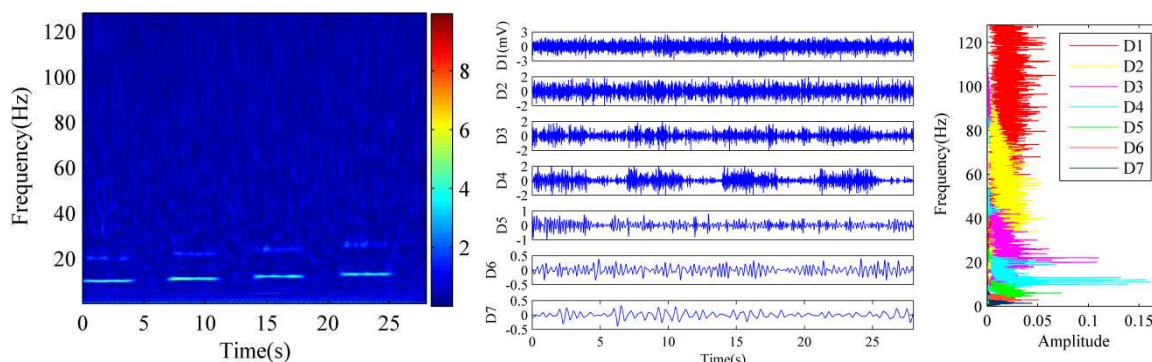


1
2

(a) STFT spectrogram

(b) Filter-based decomposition

(c) Power spectra

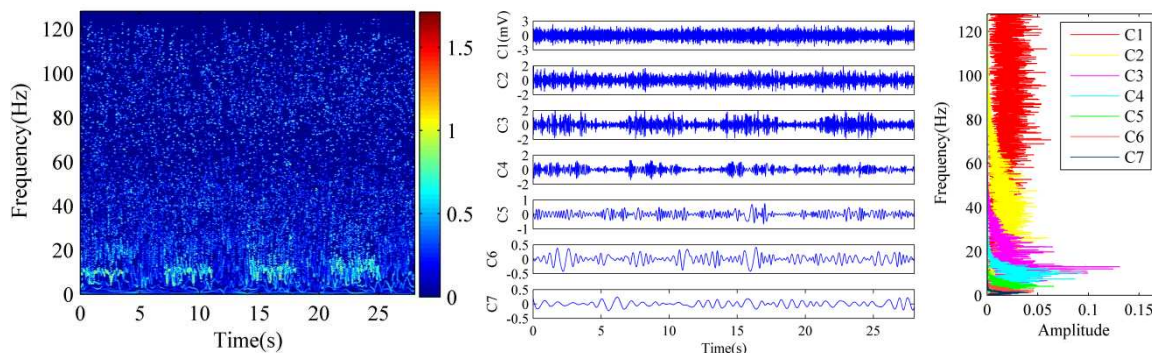


3
4

(d) Wavelet (Morlet) spectrogram

(e) Wavelet (sym4) decomposition

(f) Power spectra

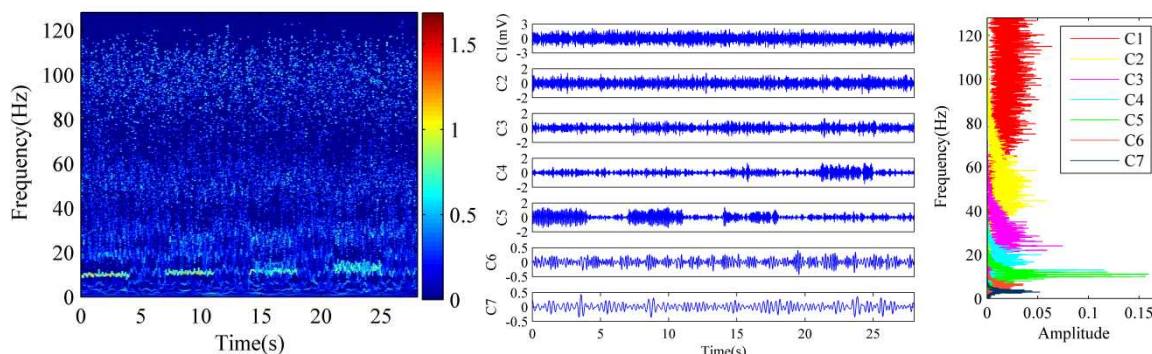


5
6

(g) EMD spectrogram by Hilbert transform

(h) EMD decomposition

(i) Power spectra



7
8

(j) MEMD spectrogram by Hilbert transform

(k) MEMD decomposition

(l) Power spectra

9
10

Figure 7. Time-frequency analysis using the Fourier transform method (a, b, c), wavelet (d, e, f), EMD (g, h, i) and MEMD (j, k, l) methods.

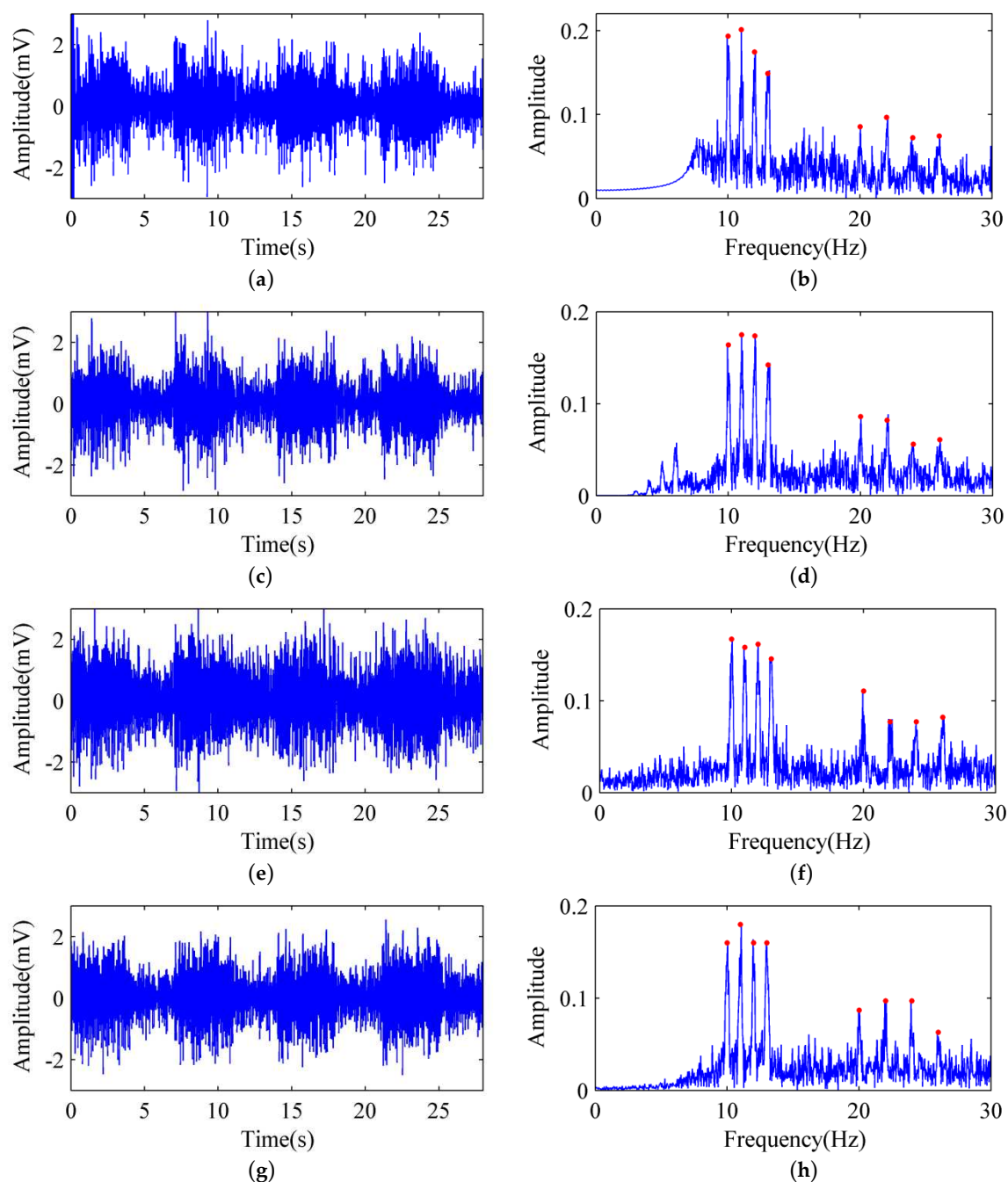
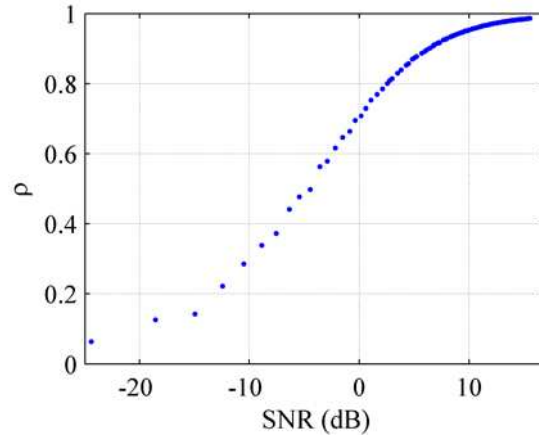


Figure 8. Reconstruction using the components decomposed by FB, WT, EMD and MEMD. (a), (c), (e) and (g) are the reconstructed signals and (b), (d), (f) and (h) are their corresponding spectrum. The red points indicate the frequency of the stimulus and their first harmonic.

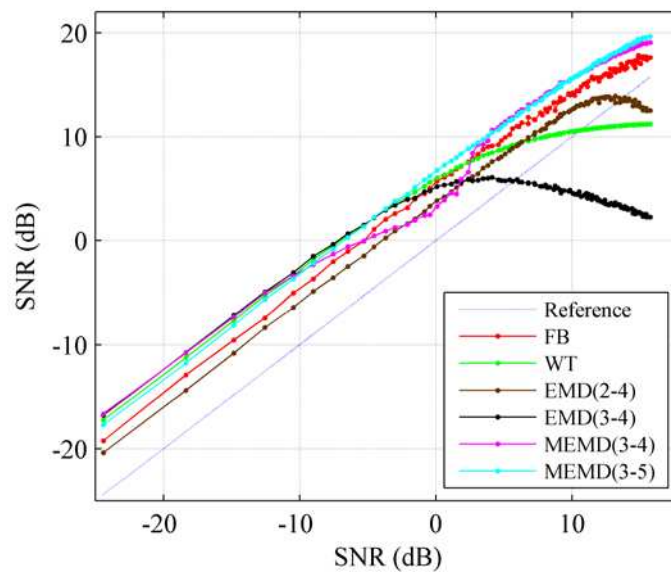
Subfigures in the middle column in Figure 7 show the components decomposed by FB, WT, EMD and MEMD method, while subfigures in the right column in Figure 7 present the spectra of the corresponding components. The WT decomposition is calculated using the sym4 mother wavelet. The time-frequency components of B4 decomposed by FB and D4 decomposed by WT method locate in the alpha band component and contain the basic frequency of SSVEP. And B3 and D3 within beta band contain the first harmonic of stimulus. The frequency ranges of the decomposition by FB and WT are consistent with classical sub-bands of EEG in Table 1. For EMD, the SSVEP signal is mainly located in the components of C3 and C4, while it is within C3, C4 and C5 decomposed by MEMD, and C4 mainly contains the 13 Hz component. Note that MEMD reduces the spectral overlap between two consecutive components and thus provides a higher frequency resolution compared with other methods. Furthermore, Figure 8 presents the extracted

1 sub-band components and their spectrum decomposed by FB, WT, EMD and MEMD. The
 2 components related to SSVEP are estimated from the simulation data at B3 and B4 derived from FB,
 3 D3 and D4 derived from WT, C3 and C4 using EMD, and C3, C4 and C5 using MEMD, respectively.
 4 All methods can reserve SSVEP rhythms. In particular, the MEMD retains less spontaneous EEG
 5 components, thus SSVEP is more prominent than those observed in other methods.



6
7
8
9
10
11

Figure 9. The influence of noise on the correlation of CCA between the simulated target signal (i.e., SSVEP) and its related contaminated signal via simulation. Larger correlation (i.e., ρ) indicates higher possibility in target recognition, so increasing the SNR may improve recognition accuracy. The SSVEP is simulated using sinusoidal waveforms and then the contaminated SSVEPs are generated at different SNR by adjusting the value of i defined in Eq (23).



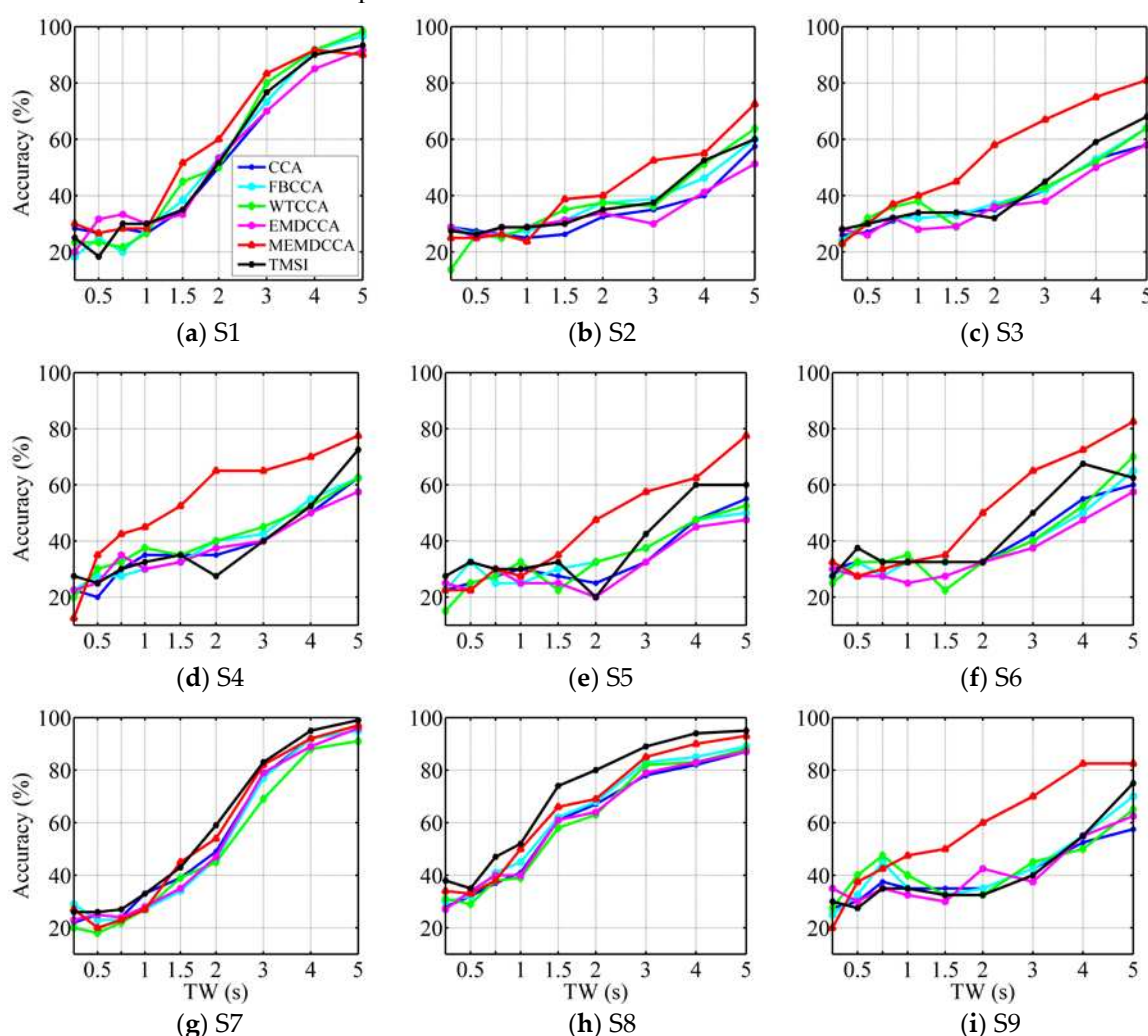
12
13
14
15
16
17

Figure 10. Performance comparison of the anti-noise capability of FB, WT, EMD and MEMD through simulation at various SNR. The FB method uses B3 and B4 to reconstruct the SSVEP, and WT method use D3 and D4. Note that EMD(2-4) indicates that the SSVEP is reconstructed using components C2, C3 and C4 decomposed by EMD, and MEMD(3-4) means that the SSVEP is reconstructed using components C3 and C4 decomposed by MEMD, and so on.

18 In order to illustrate the performance of the CCA approach for recognition of target signal at
 19 various signal-to-noise ratio (SNR), we generate the stimulated SSVEP at different levels with
 20 respect to the white noise in Eq (22), defined as

$$21 \quad s(t) = i * (s_1(t) + s_2(t) + s_3(t) + s_4(t)) + wg(t) \quad (23)$$

1 where $i=[0.1,0.2,\dots,10]$ defines the level of simulated SSVEP. Figure 9 shows the correlation of
 2 CCA for recognition of target signal at various SNR. When SNR gets lower, the correlation falls
 3 rapidly, whereas when SNR increases over approximately 10dB, the correlation increases slowly.
 4 The simulation study demonstrates that frequency feature of SSVEP is more prominent by
 5 increasing the SNR. Then the anti-noise capability of the four approaches is compared through
 6 simulation as shown in Figure 10. The reconstructed SSVEP using components C3 and C4 derived
 7 from EMD and MEMD perform better in lower SNR (< -10 dB) compared with FB and WT, while
 8 MEMD also perform better in high SNR (> 3 dB). The highest SNR is obtained with the MEMD.
 9 However, the anti-noise capability of summing C3 and C4 decomposed by MEMD decreases
 10 between approximate -10 dB and 3 dB. In this case, the SSVEP is mainly located in the components
 11 of C3, C4 and C5 as shown in subfigures in the bottom in Figure 7. It should be considered that
 12 SSVEP responses are extremely weaker than background EEG activity in real application and have
 13 lower SNR. Therefore, C3 and C4 corresponding to beta and alpha band in EEG will be used to
 14 reconstruct the SSVEP for frequencies detection.



21 **Figure 11.** SSVEP recognition accuracies with respect to time window length (i.e., 0.25 s, 0.5 s, 0.75 s,
 22 1 s, 1.5 s, 2 s, 3 s, 4 s and 5 s) for each of the nine subjects S1-S9 derived by the CCA, FB-CCA,
 23 WT-CCA, EMD-CCA, MEMD-CCA and TMSI, respectively. The components corresponding to the
 24 beta and alpha sub-bands in EEG are used for frequency detection. FB-CCA uses the sub-bands of
 25 B3 and B4, WT-CCA uses D3 and D4, EMD-CCA and MEMD-CCA use C3 and C4.

26 3.3. SSVEP recognition

27 The frequencies of SSVEP (i.e., 10, 11, 12 and 13 Hz) and its first harmonic (i.e., 20, 22, 24 and 26
 28 Hz) in this experiment are mainly within the frequency range of alpha-band (8–16 Hz) and

1 beta-band (16–32 Hz). In Table 1, it can be found that the third and fourth components decomposed
 2 by FB, WT, EMD and MEMD correspond to alpha and beta rhythms, respectively. And Figure 10
 3 indicates that the reconstructed signal using C3 and C4 have higher SNR than that using C3, C4,
 4 and C5 in the case of lower SNR. Therefore, C3 and C4 in EEG are empirically chosen to reconstruct
 5 the SSVEP of nine subjects for frequencies detection. According to the analysis above, the
 6 SSVEP-related sub-bands are estimated from the components of B3 and B4 derived from FB, D3 and
 7 D4 divided by WT, C3 and C4 using EMD/MEMD in real application, respectively. Figure 11 shows
 8 the SSVEP recognition accuracies for each subject at time windows of 0.25 s, 0.5 s, 0.75 s, 1 s, 1.5 s, 2 s,
 9 3 s, 4 s and 5 s. The fundamental frequency and first harmonic are used in the reference signals. The
 10 recognition accuracies increase with respect to time window length. Generally, the methods based
 11 on MEMD outperformed standard CCA when time window is more than 1 s. It also outperformed
 12 TMSI except S7 and S8. The FB-CCA and WT-CCA achieve higher accuracy than standard CCA in
 13 some cases, such as S1 and S2 in Figure 11, while EMD-CCA provides no improvement in
 14 recognition accuracy on most occasions. This indicates that the alpha and beta band related to SSVEP
 15 in EEG have been better decomposed by MEMD, and the extracted components including C3 and
 16 C4 are more consistent across all subjects. The EMD extracts and aligns the IMFs poorly in
 17 comparison with MEMD for multichannel EEG decomposition. In particular, an adaptively dyadic
 18 filter bank structure of MEMD on each channel of EEG signals promotes the consistency across the
 19 subjects. It is noted that the accuracy is absolutely poor at small time windows (< 1 s) for all
 20 methods.

21 **Table 2.** Comparison of the recognition accuracy across nine subjects between our proposed
 22 methods (i.e., FB-CCA, WT-CCA, EMD-CCA and MEMD-CCA) and existing methods (i.e., standard
 23 CCA, TMSI), respectively. The time window is from 0.25 s to 5 s. The results are presented as mean \pm
 24 standard deviation. The paired t -test is also used to determine whether the mean of recognition
 25 accuracy over all nine subjects are significant difference between MEMD-CCA and standard CCA,
 26 TMSI. The statistically significant difference ($p < 0.05$) is marked by asterisks (*) and daggers (†),
 27 respectively.

TW (s)	Accuracy (%)					
	CCA	FB-CCA	WT-CCA	EMD-CCA	MEMD-CCA	TMSI
0.25	26.33 \pm 2.96	25.44 \pm 3.91	22.11 \pm 5.58	26.67 \pm 4.50	25.33 \pm 6.61	28.78 \pm 3.73
0.5	27.44 \pm 3.97	29.33 \pm 3.91	28.44 \pm 6.39	27.56 \pm 3.71	28.78 \pm 5.83	28.78 \pm 6.02
0.75	30.00 \pm 4.72	29.78 \pm 8.27	31.67 \pm 8.47	31.78 \pm 4.68	33.11 \pm 7.36	32.56 \pm 5.90
1	32.56 \pm 4.75	31.44 \pm 5.98	34.00 \pm 5.22	29.78 \pm 4.58	35.89 \pm 9.99	34.33 \pm 6.93
1.5	36.22 \pm 10.08	36.33 \pm 9.87	35.56 \pm 11.01	33.89 \pm 10.60	46.67 \pm 9.94*†	38.89 \pm 13.63
2	40.22 \pm 12.78	42.56 \pm 11.54	41.22 \pm 10.05	40.89 \pm 12.75	56.00 \pm 8.99*†	41.33 \pm 18.77
3	51.33 \pm 18.47	53.11 \pm 18.66	53.11 \pm 18.50	49.44 \pm 20.31	69.78 \pm 11.34*†	56.11 \pm 20.68
4	62.00 \pm 18.92	64.00 \pm 19.58	63.33 \pm 18.45	60.67 \pm 19.19	77.00 \pm 13.25*†	69.67 \pm 18.12
5	69.78 \pm 16.98	72.56 \pm 16.83	72.89 \pm 15.45	67.89 \pm 18.49	84.00 \pm 7.83*†	76.22 \pm 15.55

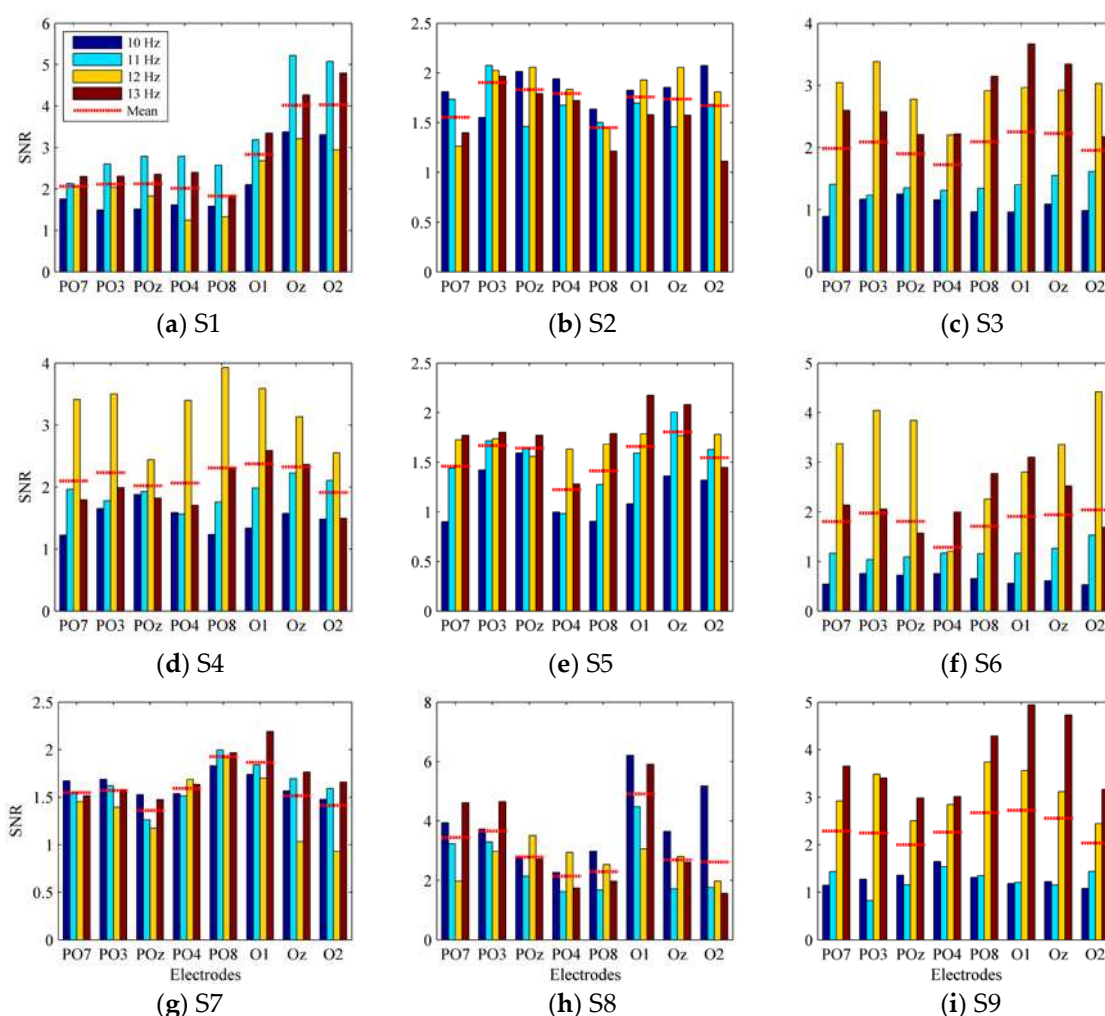
28 * Accuracy calculated from improved methods is significantly higher compared with that from standard CCA.

29 † Accuracy calculated from improved methods is significantly higher compared with that from TMSI.

30 The averaged accuracies of nine subjects are shown in Table 2. The FB-CCA and WT-CCA
 31 achieve higher averaged accuracies than standard CCA and EMD-CCA at time windows from 2 s to
 32 5 s. TMSI achieves higher averaged accuracies than standard CCA with all data length. Especially,
 33 MEMD-CCA yields the highest accuracy on average and gives the improvements of 1.34%, 3.11%,
 34 3.33%, 10.45%, 15.78%, 18.45%, 15.00% and 14.22% over standard CCA at time windows from 0.5 s to
 35 5 s. It also gives the improvements of 0.55%, 1.56%, 7.78%, 14.67%, 13.67%, 7.33% and 7.78% over
 36 TMSI at time windows from 0.75 s to 5 s. Furthermore, the paired-sample t -test is used to analyze the
 37 statistical difference between the standard CCA, TMSI, and each of the FB-CCA, WT-CCA,

1 EMD-CCA and MEMD-CCA over all subjects. The results indicate that MEMD-CCA achieves
 2 significantly higher accuracies than standard CCA and TMSI over 1.5-5 s data length, while the
 3 FB-CCA, WT-CCA and EMD-CCA method achieve no significantly higher accuracy than that of the
 4 standard CCA and TMSI methods.

5 Although the SSVEP responses with high amplitude can be record by electrodes located at the
 6 occipital areas, the best locations are subject specificity. Signal-to-noise ratio (SNR) can be used as
 7 the criterion to determine optimal electrode positions for each subject. It is estimated by a ratio of
 8 power at a given stimulation frequency with respect to the mean power of the adjacent points
 9 according to [36]. This estimator approximately represents the SNR of SSVEP. Figure 12 illustrates
 10 the SNR with different electrodes across the nine subjects. For example, the electrodes placed at the
 11 bottom region (i.e., "O1", "Oz" and "O2") have the best SNR for S1. While the middle region (i.e.,
 12 "PO3", "POz", "PO4", "O1", "Oz" and "O2") has relatively higher SNR for S2.



13
14
15
16
17
18
19 **Figure 12.** Signal-to-noise-ratio (SNR) of SSVEP at different stimulation frequencies with respect to
 20 EEG electrodes for each of the nine subjects S1-S9, respectively. The bar is averaged SNR over
 21 experiment realizations (5 runs \times 5 sessions for each frequency). The mean value across four
 22 frequencies for a given electrode is also indicated with red dashed line.

23 Furthermore, the computational time of selected six methods is compared as shown in Table 3.
 24 The time is calculated in Matlab (v7.13, MathWorks Inc., Natick, MA, USA) on a personal computer
 25 with the configuration of AMD A8-7410 APU @ 2.20 GHz, 8 GB RAM, 64 bit Win 10. The EMD-CCA
 26 and MEMD-CCA take larger time than CCA, FB-CCA, WT-CCA and TMSI. Especially,
 27 MEMD-CCA requires the longest computational time. In fact, although MEMD-based methods
 28 outperform the traditional algorithms in processing the nonlinear and non-stationary signals, they
 29 always require large computational resources.

Table 3. The computational time for CCA, FB-CCA, WT-CCA, EMD-CCA, MEMD-CCA and TMSI, respectively. Five session data from S1 are used. For each time window, there are total 100 computations. The results are presented as mean \pm standard deviation.

TW (s)	Computational time (s)					
	CCA	FB-CCA	WT-CCA	EMD-CCA	MEMD-CCA	TMSI
0.25	0.0043 \pm 0.0061	0.0144 \pm 0.0152	0.1571 \pm 0.0760	1.06 \pm 0.09	29.46 \pm 18.29	0.0159 \pm 0.0116
0.5	0.0036 \pm 0.0008	0.0126 \pm 0.0025	0.1448 \pm 0.0065	1.37 \pm 0.09	32.72 \pm 15.83	0.0175 \pm 0.0048
0.75	0.0037 \pm 0.0010	0.0120 \pm 0.0012	0.1482 \pm 0.0120	1.47 \pm 0.06	37.95 \pm 17.15	0.0191 \pm 0.0044
1	0.0037 \pm 0.0008	0.0120 \pm 0.0016	0.1470 \pm 0.0054	1.74 \pm 0.14	40.32 \pm 9.28	0.0238 \pm 0.0053
1.5	0.0040 \pm 0.0007	0.0126 \pm 0.0015	0.1491 \pm 0.0075	1.86 \pm 0.11	49.16 \pm 11.71	0.0496 \pm 0.0110
2	0.0047 \pm 0.0012	0.0129 \pm 0.0013	0.1509 \pm 0.0106	2.12 \pm 0.04	57.04 \pm 17.60	0.0915 \pm 0.0382
3	0.0056 \pm 0.0013	0.0144 \pm 0.0018	0.1516 \pm 0.0073	2.35 \pm 0.12	71.42 \pm 14.58	0.1547 \pm 0.0449
4	0.0062 \pm 0.0019	0.0160 \pm 0.0023	0.1479 \pm 0.0053	2.69 \pm 0.08	87.22 \pm 13.98	0.2666 \pm 0.0651
5	0.0063 \pm 0.0022	0.0163 \pm 0.0021	0.1531 \pm 0.0099	2.83 \pm 0.09	103.19 \pm 26.70	0.3865 \pm 0.0597

4. Discussion

Analysis of the SSVEP demonstrates that MEMD-CCA is able to achieve the highest accuracy compared with standard CCA, FB-CCA, WT-CCA, EMD-CCA and TMSI. While the FB-CCA, WT-CCA and EMD-CCA methods give a similar performance as the standard CCA method. The significant improvements for recognizing SSVEP suggest that the beta (i.e., 16-32 Hz) and alpha (i.e., 8-16 Hz) bands in EEG corresponding to C3 and C4 decomposed by MEMD show much better discriminatory properties in comparison with the whole frequency band. The empirical sub-bands extraction by MEMD for recognizing SSVEP approximates the ground truth. The SSVEP performs more prominent in the extracted sub-bands. The fundamental frequency components (i.e., 10, 11, 12, and 13 Hz) are mainly located in the alpha band, and their first harmonics (i.e., 20, 22, 24, and 26 Hz) are in the beta band. MEMD is more appropriate to analyze nonlinear and non-stationary signal like EEG compared with linear method like FFT and WT, therefore, it performs better than FB and WT in adaptive separation of alpha and beta bands. Although EMD can decompose each channel of EEG into several data-driven components, it misaligns the same-index IMFs across multiple channels, and thus it fails to achieve the effective fusion of multivariate signal. Unlike the standard one channel EMD algorithm, MEMD simultaneously separates the multivariate signal as a whole into multivariate IMFs, effectively aligns the frequency range of the IMFs with same index, and therefore solves the mode misalignment in multivariate IMFs. MEMD guarantees that the same-index IMFs contain the same information across the data channels and is suitable for the fusion applications to multiple components of a multivariate signal.

In this study, the components C3 and C4 are selected empirically, the high recognition accuracy suggests that the MEMD estimates for SSVEP is proper. However, the important thing is that there is no clue to choose the summation of only C3 and C4 to obtain the SSVEP-related sub-band. There is no any standard criterion for the selection. MEMD may produce different numbers of IMFs when it is applying on the same signal in different times because of mode mixing. Hence the choice is not always perfect to achieve the sub-band extraction goal. However, by adding two extra channels containing white noise and another four reference channels, each of which compromises the sum of reference signal corresponding to target and white noise, MEMD behaves as an adaptively dyadic filter bank structure on each channel of EEG signals and ensures the consistency of extracted EEG patterns across trails and subjects. The reference channels enable the enhanced decomposition accuracy by alignment with the stimulus frequency. Moreover, MEMD is able to reduce the mode mixing problem due to the broad band property of white noise, consequently decrease frequency overlapping of two consecutive IMFs and improve the frequency resolution.

1 Note that the designed stimulus frequencies of the target are located in the alpha band in this
 2 study, so the dyadic filter bank structure ensures that their first harmonics are located in the beta
 3 band. In practice, the stimulus frequencies may be in other bands, thus the IMFs can be differently
 4 selected towards the highest detection accuracy. Moreover, Bin et al. [37] has indicated that
 5 reference signals containing different numbers of harmonics give no significant difference for
 6 SSVEP recognition accuracy by CCA. Therefore, we only select the first harmonics in the reference
 7 signals for SSVEP recognition. In this study, only four targets are designed using LEDs to illustrate
 8 the performance of our proposed method. Just like CCA, the improved MEMD-CCA is also suitable
 9 for large number of targets with different frequencies. It will be better that the fundamental
 10 frequencies of stimulus are within a single sub-band as shown in Table 1. However, when using the
 11 traditional frame-based method to generate visual stimulation, the number is generally restricted
 12 by the LCD screen refresh rate. Therefore, over recent years, several methods have been proposed
 13 to encode more targets with limited stimulus frequencies [5, 19]. In [5], a novel stimulation method
 14 is proposed to overcoming the limitation of refresh rate of a monitor. The authors designed 9
 15 targets based on intermodulation frequencies on a LCD screen. In [19], any stimulation frequency
 16 up to half of the refresh rate can be implemented on a conventional LCD screen. The authors
 17 designed a SSVEP-based BCI platform using 12 targets including digits (i.e., 0-9), backspace, and
 18 enter keys in LCD screen. Their corresponding flashing frequencies are in the range of 9-11.75 Hz
 19 and the interval between two consecutive targets is 0.25 Hz. The frequency range (i.e., 9-11.75 Hz) is
 20 in the alpha-band (i.e., 8-16 Hz), and the first harmonic is in beta-band (i.e., 16-32 Hz). The design is
 21 the same as that in our study. So MEMD-CCA can be used to recognize the frequencies related to 12
 22 targets. However, it also should be noted that with the increase of the targets, MEMD-CCA will
 23 need more reference channels and thus consume more time.

24 EMD-based methods containing ensemble EMD (EEMD) and multivariate EMD (MEMD) are
 25 quite time consuming. Especially, MEMD takes even larger time due to introduction of white noise
 26 and reference channels including information of target frequency. It can be seen in Table 3,
 27 MEMD-CCA takes around 103 s to detect the SSVEP on an eight channel EEG with 1280 data points
 28 (time window 5 s). The enormous amount of computational time limits the applications in real-time.
 29 Nevertheless, advances in hardware solutions and parallel computing are expected to enable the
 30 online and real-time operation of MEMD-based algorithms.

31 For SSVEP detection or recognition in BCI applications, many researchers have confirmed that
 32 a sophisticated calibration with appropriate analysis method could significantly improve the
 33 accuracy [25]. The extended CCA methods containing MCCA [22], L1-MCCA [23], MsetCCA [24]
 34 and CFA [25] optimize the reference signal from the training data. The trained reference signals
 35 include more discriminative information in comparison with sine and cosine functions used in
 36 standard CCA. Therefore, these extend methods effectively improve the accuracy of SSVEP
 37 detection. However, the optimized reference signals are derived from the whole band of training
 38 data. In further study, the MEMD methods can be employed to further optimize the reference
 39 signals from extracted sub-bands of training data. The optimization process of reference signals is
 40 based on extracted sub-band corresponding to SSVEP from the training data and thus could yield
 41 further improvements.

42 Furthermore, other new techniques also have been proved to outperform the CCA method in
 43 SSVEP recognition, such as multivariate synchronization index (MSI) [17], temporally local MSI
 44 (TMSI) [4] and likelihood ratio test (LRT) [18]. However, the essence of these methods is the same as
 45 CCA, which is to investigate the relationships between two sets of variables. In SSVEP recognition,
 46 they all rely on the original EEG and pre-constructed sine-cosine waves instead of optimized
 47 signals. The MEMD-CCA estimates the SSVEP in an optimized narrowband of interest rather than
 48 original broadband activities. The results indicate that the MEMD-CCA also achieves higher
 49 accuracy on average than TMSI. Hence, the idea in our study is absolutely suitable for this class of
 50 methods for further improvement of SSVEP recognition.

51 In SSVEP-based BCI systems, the stimulation frequency is a crucial factor for achieving high
 52 performance. Generally, SSVEP respond to flickering stimuli at a wide range of frequency from 1 to

1 100 Hz [38]. The frequency can be divided into low-frequency range (up to 12 Hz),
2 medium-frequency range (12-30 Hz) and high-frequency range (above 30 Hz) [39, 40]. SSVEP
3 magnitudes strongly depend on the stimulation frequency. The low-frequency stimuli evoke higher
4 response magnitude and thus the SSVEPs have good SNR, while these responses significantly
5 decrease in the high-frequency range. Consequently, the low and medium frequency stimuli are
6 primarily used to elicit SSVEPs in majority of BCI systems [10, 13, 38, 40]. However, the stimulation
7 in these frequency ranges may cause visual fatigue or even seizure in subjects. High-frequency
8 visual stimuli can eliminate the above limitations for a more comfortable BCI system. Hence, many
9 studies have focused on the high-frequency SSVEPs, but limited success has been achieved because
10 of the rapid decrease of SNR with the increase of frequency [40]. The poor SNR and the
11 non-stationarity of EEG demand an efficient algorithm for recognition of high-frequency SSVEPs. In
12 this study, stimulation with frequencies of below 30 Hz is used to design the BCI system.
13 Theoretically MEMD-CCA can improve the SNR of SSVEPs regardless of the frequency ranges and
14 thus achieves higher accuracy. Therefore, this method may provide a potential solution to
15 implement a high-frequency SSVEP-based BCI system. For example, the stimulation frequencies are
16 in gamma-band (32–64 Hz), the components C1 and C2 can be chose to detect the SSVEPs. However,
17 due to dyadic filter bank structure of MEMD, C1 and C2 have wider band than the high level
18 components, hence there will retain relatively more spontaneous EEG activities. A higher frequency
19 resolution by increasing sample rate can eliminate this problem. Furthermore, because the strength
20 of spontaneous EEG activities also reduces at high frequency, the SNR is still appreciable. We will
21 further consider the frequency range and compare our results with high-frequency system using
22 MEMD-CCA.

23 The selection of a neutral reference in EEG recording is a critical issue. Currently, earlobe
24 reference, neck reference and mastoid reference etc have been used for EEG recording [41, 42]. The
25 g.USBamp system in this study for EEG measuring typically uses the right earlobe electrode as the
26 reference. However, because these reference locations are neither at neutral points nor zero of
27 potential, they may cause an undesired temporal bias in recordings and even results in errors in
28 EEG analysis. Many researchers have pay attention to this topic and proposed several solutions to
29 this problem. The average reference (AR) [43] using average potential across all electrodes provided
30 an available option. Yao [44, 45] introduced an approximate standardization of the reference in EEG
31 measuring called reference electrode standardization technique (REST). A scalp point or the
32 average reference is transformed to a reference point at infinity in REST. Both REST and AV have
33 been proved to outperform all other known references. However, it also should be noted that none
34 of these two methods are completely free of limitations [46]. The electrode density, electrode
35 coverage and head model etc can cause extra errors in EEG. The diversity of optimal electrode
36 positions across subject also should be considered in implementation of practical BCIs.

37 5. Conclusions

38 In this study, a novel MEMD method for improving the performance of SSVEP-based BCI is
39 introduced. MEMD provided a robust way to analyze nonlinear and non-stationary signals and
40 extract narrowband signal of interest from multichannel broadband signals. In a SSVEP-based BCI,
41 the MEMD algorithm enhanced multi-component extraction of SSVEP responses buried in
42 broadband background activities. In comparison with traditional linear method like FFT and WT,
43 MEMD performed more suitable to analyze nonlinear and non-stationary EEG and estimated more
44 localized in time-frequency analysis. Furthermore, unlike the standard single channel EMD method,
45 MEMD can more correctly decompose the common oscillatory modes across multichannel EEG.
46 The performance of the proposed approach for SSVEP detection in sub-bands exhibited significant
47 improvements over CCA and TMSI, while FB-CCA, WT-CCA and EMD-CCA provided no
48 significant difference with standard CCA. The further optimization of reference signals in a
49 particular sub-band of training data rather than the whole frequency band may give an
50 improvement, this will be investigated in the future work.

1 **Acknowledgments:** The authors would like to acknowledge the funding support from the Lottery Health
 2 Research Council of New Zealand and the Faculty of Engineering Research and Development Grant from the
 3 University of Auckland. Also the work is supported by National Natural Science Foundation of China (Grant
 4 Number: 51475342 and 51675389).

5 **Conflicts of Interest:** The authors declare no conflict of interest.

7 References

- 8 [1] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain-computer interfaces for
 9 communication and rehabilitation," *Nat Rev Neurol*, vol. 12, pp. 513-525, 2016.
- 10 [2] S. Halder, E. M. Hammer, S. C. Kleih, M. Bogdan, W. Rosenstiel, N. Birbaumer, *et al.*, "Prediction of
 11 Auditory and Visual P300 Brain-Computer Interface Aptitude," *PLOS ONE*, vol. 8, p. e53513, 2013.
- 12 [3] B. O. Mainsah, L. M. Collins, and C. S. Throckmorton, "Using the detectability index to predict P300
 13 speller performance," *Journal of Neural Engineering*, vol. 13, p. 066007, 2016.
- 14 [4] Y. Zhang, D. Guo, P. Xu, Y. Zhang, and D. Yao, "Robust frequency recognition for SSVEP-based BCI
 15 with temporally local multivariate synchronization index," *Cognitive neurodynamics*, vol. 10, pp.
 16 505-511, 2016.
- 17 [5] C. Xiaogang, W. Yijun, Z. Shangen, G. Shangkai, H. Yong, and G. Xiaorong, "A novel stimulation
 18 method for multi-class SSVEP-BCI using intermodulation frequencies," *Journal of Neural Engineering*,
 19 vol. 14, p. 026013, 2017.
- 20 [6] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, and B. He, "Noninvasive Electroencephalogram Based
 21 Control of a Robotic Arm for Reach and Grasp Tasks," *Scientific Reports*, vol. 6, p. 38565, 2016.
- 22 [7] M. Teng, L. Hui, D. Lili, Y. Hao, L. Xulin, L. Peiyang, *et al.*, "The hybrid BCI system for movement
 23 control by combining motor imagery and moving onset visual evoked potential," *Journal of Neural
 24 Engineering*, vol. 14, p. 026015, 2017.
- 25 [8] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A speedy hybrid BCI spelling approach
 26 combining P300 and SSVEP," *IEEE Transactions on Biomedical Engineering*, vol. 61, pp. 473-483, 2014.
- 27 [9] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," *IEEE
 28 Transactions on Biomedical Engineering*, vol. 61, pp. 1436-1447, 2014.
- 29 [10] A.-A. Mohammad and P. Angelika, "Advancing the detection of steady-state visual evoked potentials
 30 in brain-computer interfaces," *Journal of Neural Engineering*, vol. 13, p. 036005, 2016.
- 31 [11] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, "BCI demographics: how
 32 many (and what kinds of) people can use an SSVEP BCI?," *IEEE transactions on neural systems and
 33 rehabilitation engineering*, vol. 18, pp. 107-116, 2010.
- 34 [12] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: focus
 35 on essential paradigms and future perspectives," *Progress in neurobiology*, vol. 90, pp. 418-438, 2010.
- 36 [13] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A Comparison Study of Canonical Correlation
 37 Analysis Based Methods for Detecting Steady-State Visual Evoked Potentials," *PloS one*, vol. 10, p.
 38 e0140703, 2015.
- 39 [14] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi, S. W. Lee, and C.-H. Im, "Development of an SSVEP-based
 40 BCI spelling system adopting a QWERTY-style LED keyboard," *Journal of neuroscience methods*, vol. 208,
 41 pp. 59-65, 2012.
- 42 [15] Z. Wu and D. Yao, "Frequency detection with stability coefficient for steady-state visual evoked
 43 potential (SSVEP)-based BCIs," *Journal of Neural Engineering*, vol. 5, pp. 36-43, 2007.

- 1 [16] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis
2 for SSVEP-based BCIs," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 2610-2614, 2006.
- 3 [17] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition
4 of SSVEP-based brain-computer interface," *Journal of neuroscience methods*, vol. 221, pp. 32-40, 2014.
- 5 [18] Y. Zhang, L. Dong, R. Zhang, D. Yao, Y. Zhang, and P. Xu, "An Efficient Frequency Recognition
6 Method Based on Likelihood Ratio Test for SSVEP-Based BCI," *Computational and Mathematical Methods
7 in Medicine*, vol. 2014, p. 7, 2014.
- 8 [19] Y.-T. Wang, Y. Wang, and T.-P. Jung, "A cell-phone-based brain-computer interface for
9 communication in daily life," *Journal of neural engineering*, vol. 8, p. 025018, 2011.
- 10 [20] L. Ke, C. Andrea, W. Yijun, C. Xiaogang, G. Shangkai, and G. Xiaorong, "An online hybrid BCI system
11 based on SSVEP and EMG," *Journal of Neural Engineering*, vol. 13, p. 026020, 2016.
- 12 [21] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed BCI based on code modulation
13 VEP," *Journal of neural engineering*, vol. 8, p. 025015, 2011.
- 14 [22] Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. Jin, X. Wang, *et al.*, "Multiway canonical correlation analysis
15 for frequency components recognition in SSVEP-based BCIs," in *International Conference on Neural
16 Information Processing*, 2011, pp. 287-295.
- 17 [23] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical
18 correlation analysis for SSVEP-based BCI," *IEEE Transactions on Neural Systems and Rehabilitation
19 Engineering*, vol. 21, pp. 887-896, 2013.
- 20 [24] Y. U. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI
21 using multiset canonical correlation analysis," *International journal of neural systems*, vol. 24, p. 1450013,
22 2014.
- 23 [25] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "SSVEP recognition using common feature
24 analysis in brain-computer interface," *Journal of neuroscience methods*, vol. 244, pp. 8-15, 2015.
- 25 [26] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, *et al.*, "The empirical mode
26 decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,"
27 *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454,
28 pp. 903-995, 1998.
- 29 [27] V. Bajaj and R. B. Pachori, "Classification of seizure and nonseizure EEG signals using empirical mode
30 decomposition," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, pp. 1135-1142, 2012.
- 31 [28] J. Lee, D. D. McManus, S. Merchant, and K. H. Chon, "Automatic motion and noise artifact detection in
32 holter ECG data using empirical mode decomposition and statistical approaches," *IEEE Transactions on
33 Biomedical Engineering*, vol. 59, pp. 1499-1506, 2012.
- 34 [29] T. Lee and T. B. M. J. Ouarda, "Prediction of climate nonstationary oscillation processes with empirical
35 mode decomposition," *Journal of Geophysical Research: Atmospheres*, vol. 116, pp. n/a-n/a, 2011.
- 36 [30] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal
37 Society A: Mathematical, Physical and Engineering Science*, vol. 466, pp. 1291-1302, 2010.
- 38 [31] N. Ur Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode
39 decomposition," *IEEE Transactions on Signal Processing*, vol. 59, pp. 2421-2426, 2011.
- 40 [32] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for
41 implementing a high-speed SSVEP-based brain-computer interface," *Journal of neural engineering*, vol.
42 12, p. 046008, 2015.

- 1 [33] A. N. Akansu and R. A. Haddad, *Multiresolution signal decomposition: transforms, subbands, and wavelets*:
2 Academic Press, 2001.
- 3 [34] D. Gajic, Z. Djurovic, J. Gligorijevic, S. Di Gennaro, and I. Savic-Gajic, "Detection of epileptiform
4 activity in EEG signals based on time-frequency and non-linear analysis," *Frontiers in computational*
5 *neuroscience*, vol. 9, p. 38, 2015.
- 6 [35] T. Nguyen-Ky, P. Wen, Y. Li, and R. Gray, "Measuring and reflecting depth of anesthesia using
7 wavelet and power spectral density," *IEEE Transactions on information technology in Biomedicine*, vol. 15,
8 pp. 630-639, 2011.
- 9 [36] R. Kuś, A. Duszyk, P. Milanowski, M. Łabęcki, M. Bierzyńska, Z. Radzikowska, *et al.*, "On the
10 Quantification of SSVEP Frequency Responses in Human EEG in Realistic BCI Conditions," *PLOS ONE*,
11 vol. 8, p. e77536, 2013.
- 12 [37] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer
13 interface using a canonical correlation analysis method," *Journal of neural engineering*, vol. 6, p. 046002,
14 2009.
- 15 [38] K. No-Sang, M. Klaus-Robert, and L. Seong-Whan, "A lower limb exoskeleton control system based on
16 steady state visual evoked potentials," *Journal of Neural Engineering*, vol. 12, p. 056009, 2015.
- 17 [39] P. F. Diez, V. A. Mut, E. M. Avila Perona, and E. Laciari Leber, "Asynchronous BCI control using
18 high-frequency SSVEP," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, p. 39, 2011.
- 19 [40] I. Volosyak, D. Valbuena, T. Luth, T. Malechka, and A. Graser, "BCI Demographics II: How Many (and
20 What Kinds of) People Can Use a High-Frequency SSVEP BCI?," *IEEE Transactions on Neural Systems*
21 *and Rehabilitation Engineering*, vol. 19, pp. 232-239, 2011.
- 22 [41] F. Chella, V. Pizzella, F. Zappasodi, and L. Marzetti, "Impact of the reference choice on scalp EEG
23 connectivity estimation," *Journal of neural engineering*, vol. 13, p. 036016, 2016.
- 24 [42] Y. Qin, P. Xu, and D. Yao, "A comparative study of different references for EEG default mode network:
25 the use of the infinity reference," *Clinical neurophysiology*, vol. 121, pp. 1981-1991, 2010.
- 26 [43] M. Junghöfer, T. Elbert, D. M. Tucker, and C. Braun, "The polar average reference effect: a bias in
27 estimating the head surface integral in EEG recording," *Clinical Neurophysiology*, vol. 110, pp. 1149-1155,
28 1999.
- 29 [44] D. Yao, "A method to standardize a reference of scalp EEG recordings to a point at infinity,"
30 *Physiological measurement*, vol. 22, p. 693, 2001.
- 31 [45] D. Yao, L. Wang, R. Oostenveld, K. D. Nielsen, L. Arendt-Nielsen, and A. C. Chen, "A comparative
32 study of different references for EEG spectral mapping: the issue of the neutral reference and the use of
33 the infinity reference," *Physiological measurement*, vol. 26, p. 173, 2005.
- 34 [46] P. L. Nunez, "REST: A Good Idea but Not the Gold Standard," *Clinical neurophysiology : official journal of*
35 *the International Federation of Clinical Neurophysiology*, vol. 121, pp. 2177-2180, 06/02 2010.
- 36