Journal of Neuroscience Methods 221 (2014) 32–40

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Multivariate synchronization index for frequency recognition of SSVEP-based brain–computer interface

Yangsong Zhang, Peng Xu, Kaiwen Cheng, Dezhong Yao*

Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, 610054, China

HIGHLIGHTS

- A novel frequency recognition method (MSI) was proposed.
- MSI showed better performance than traditional method when using short length data.
- MSI showed better performance when using small number of channels.
- MSI was successfully used in online BCI experiment.

ARTICLE INFO

Article history:
Received 20 March 2013
Received in revised form 23 July 2013
Accepted 28 July 2013

Keywords:
Steady-state visual evoked potential (SSVEP)
Brain–computer interface (BCI)
Multivariate synchronization index (MSI)

ABSTRACT

Multichannel frequency recognition methods are prevalent in SSVEP-BCI systems. These methods increase the convenience of the BCI system for users and require no calibration data. A novel multivariate synchronization index (MSI) for frequency recognition was proposed in this paper. This measure characterized the synchronization between multichannel EEGs and the reference signals, the latter of which were defined according to the stimulus frequency. For the simulation and real data, the proposed method showed better performance than the widely used canonical correlation analysis (CCA) and minimum energy combination (MEC), especially for short data length and a small number of channels. The MSI was also implemented successfully in an online SSVEP-based BCI system, thus further confirming its feasibility for application systems. Because fast and accurate recognition is crucial for practical systems, we recommend MSI as a potential method for frequency recognition in future SSVEP-BCI.

1. Introduction

Interest in brain–computer interface (BCI) research has increased because of its application value in neural engineering and neuroscience (Jensen et al., 2011; Nicolas-Alonso and Gomez-Gil, 2012). BCIs can provide online communication between a human or animal brain and external devices without depending on the normal output pathways of peripheral nerves and muscles. BCI applications include communication devices for disabled people, neuroprostheses and games (Allison et al., 2007; Nicolas-Alonso and Gomez-Gil, 2012). Because of its high temporal resolution, relatively low cost and high portability, scalp electroencephalography (EEG) is the most widely used modality in BCI systems (Nicolas-Alonso and Gomez-Gil, 2012). Many EEG signals could serve as control signals in BCI systems; examples of these signals include slow cortical potentials, sensorimotor rhythms, P300 evoked potentials, and steady-state visual evoked potentials (SSVEPs) (Nicolas-Alonso and Gomez-Gil, 2012; Wolpaw et al., 2002). Of these, SSVEPs have received widespread attention for their high signal-to-noise ratio (SNR). SSVEPs are actually near-sinusoidal waveforms; the SSVEP frequency is the same as the fundamental frequency and harmonics of the driving stimulus, and SSVEPs generally occur in the occipital and parietal lobes (Regan, 1989). Due to their high information transfer rate and minimal training, SSVEP-based BCIs have become a crucial branch of BCIs (Vialatte et al., 2010). In an SSVEP-BCI system, the targets are encoded by a single frequency or various combinations of frequencies (Cheng et al., 2002; Shyu et al., 2010; Srihari Mukesh et al., 2006; Zhang et al., 2012). Different commands can be transmitted by shifting the subject's attention to the coded targets.

However, users have shown large inter-variation in the SSVEP amplitude and distribution (Srinivasan et al., 2006, 2007). Parameter optimization (e.g., channel selection and appropriate data length) is required for each user for better BCI performance (Wang et al., 2006; Wu and Yao, 2008). These calibrations are always necessary, especially when some traditional signal analysis methods are used, such as power spectral density analysis (PSDA) (Wang

* Corresponding author at: #4, Section 2, North JianShe Road, ChengDu, SiChuan 610054, China. Tel.: +86 028 83201018; fax: +86 028 83208238.
E-mail address: dyao@uestc.edu.cn (D. Yao).

0165-0270/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2013.07.018

et al., 2006) and stability coefficients (SC) (Wu and Yao, 2008). These optimization requirements limit the practical application of SSVEP-based BCIs (Bin et al., 2009). Because any BCI performance strongly depends on quickly and correctly decoding the user' intentions, it is critical to employ a reliable signal processing method to detect and extract SSVEPs without a troublesome optimization step. The methods that require no or few optimizations and provide high recognition performance may make BCI applications more convenient (Bin et al., 2009; Friman et al., 2007).

In the past few years, a few multichannel detection methods for SSVEP-BCIs have been proposed (Friman et al., 2007; Lin et al., 2006). These methods that extract more information from multichannel EEGs can overcome the above mentioned disadvantages and achieve higher detection accuracy. Two multichannel detection methods for frequency recognition are minimum energy combination (MEC) and canonical correlation analysis (CCA). MEC is a method that finds spatial filters to combine a multi-electrode EEG into virtual channel signals in which the nuisance signals and noise are reduced as much as possible (Friman et al., 2007), and MEC has been used in several SSVEP-based BCI systems (Cecotti, 2010; Volosyak, 2011; Volosyak et al., 2011). The CCA-based method uses the multivariable CCA statistical method to calculate the correlation coefficients between the multiple electrodes of an EEG and the reference signals (Lin et al., 2006). This method shows better detection accuracy than that of PSDA; thus, it has been used in some SSVEP-based BCI systems (Bin et al., 2009; Wang et al., 2011). In addition, CCA shows better performance than that of MEC (Nan et al., 2011).

In this work, we proposed another novel multichannel frequency recognition method known as the multivariate synchronization index (MSI). The S-estimator was adopted to estimate the synchronization index, and the new method (MSI) was compared with CCA and MEC using simulated and experimental data.

2. Materials and methods

This study was approved by the Institution Research Ethics Board at University of Electronic Science and Technology of China. All participants were asked to read and sign an informed consent form before participating in the study. After the experiment, all the participants received monetary compensation for their time and effort.

2.1. MSI-based frequency recognition

In this work, we aimed to estimate the synchronization between the actual mixed signals and the reference signals as a potential index for recognizing the stimulus frequency. Briefly, we proposed the use of the S-estimator as the index. In fact, the S-estimator was previously designed to measure the amount of synchronization over a single or two regions of the cortex; the S-estimator is based on the entropy of the normalized eigenvalues of the correlation matrix of multivariate signals (Carmeli, 2006; Carmeli et al., 2005; Joudaki et al., 2012). In this method, the measured synchronization is inversely proportional to the embedding dimension. The dimensionality of the data is maximal for totally uncorrelated time series but minimal for perfectly synchronized data (Carmeli et al., 2005; Walker et al., 2010). The data dimensionality is a good measure of synchronization because the variance of an eigenspectrum correlates with the dimensionality. In fact, the more disperse the original eigenspectrum, the higher the entropy of the normalized eigenvalues. In contrast, the more concentrated the original eigenspectrum, the lower the entropy of the normalized eigenvalues (Carmeli et al., 2005). In this study, we assumed that the reference signal derived from the stimulus would be synchronized to the mixed EEG recordings; thus, the S-estimator can provide a synchronization index for frequency detection.

The MSI must also create a reference signal from the stimulus frequencies used in an SSVEP-based BCI system, similarly to CCA and MEC.

Let us denote the EEG signals by a matrix X of size $N \times M$ and the reference signals by a matrix Y of size $2N_h \times M$. Here, N is the number of channels, M is the number of samples, and N_h is the number of harmonics for the sine and cosine components. Without loss of generality, X and Y are normalized to have a zero mean and unitary variance. Then, a correlation matrix is calculated as

$$c = \begin{bmatrix} C11 & C12 \\ C21 & C22 \end{bmatrix} \tag{1}$$

where

$$C11 = \frac{1}{M} XX^T \tag{2}$$

$$C22 = \frac{1}{M} YY^T \tag{3}$$

$$C12 = C21 = \frac{1}{M} XY^T \tag{4}$$

The matrix C includes both the autocorrelation and cross-correlation of X and Y, and the autocorrelation will influence the synchronization measure (Joudaki et al., 2012). To reduce these influences, the following linear transformation is adopted:

$$U = \begin{bmatrix} C11^{-(1/2)} & 0 \\ 0 & C22^{-(1/2)} \end{bmatrix} \tag{5}$$

Then, the transformed correlation matrix is

$$R = UCU^T = \begin{bmatrix} I_{N\times N} & C11^{-(1/2)}C12C22^{-(1/2)} \\ C22^{-(1/2)}C21C11^{-(1/2)} & I_{2N_h\times 2N_h} \end{bmatrix} \tag{6}$$

where $I_{N\times N}$ is the identity matrix of dimension N, and $I_{2N_h\times 2N_h}$ is that of dimension $2N_h$. The autocorrelation is canceled out in formula (6) after the linear transformation of (5) (Joudaki et al., 2012).

Let $\lambda_1, \lambda_2, \ldots, \lambda_P$ be the eigenvalues of matrix R. Then, the normalized eigenvalues are calculated as follows:

$$\lambda'_i = \frac{\lambda_i}{\sum_{i=1}^{P} \lambda_i} = \frac{\lambda_i}{tr(R)} \tag{7}$$

where $P = N + 2N_h$. Then, the synchronization index between two sets of signals can be calculated as follows:

$$S = 1 + \frac{\sum_{i=1}^{P} \lambda'_i \log(\lambda'_i)}{\log(P)} \tag{8}$$

If the two sets of signals are uncorrelated, $C12 = C21 = 0$, and R will be diagonal. Then, $\lambda'_i = 1/P$, and consequently, S should be zero. Alternatively, if the two sets of signals are perfectly correlated, R will have ones on the main diagonal and zeros elsewhere. Only one normalized eigenvalue is one, and the other eigenvalues are zeros. Therefore, S should be one. For other situations, the S value should range from zero to one.

To implement the MSI for frequency detection, we must also create reference signals, which are also calculated for CCA and MEC. If K stimulus frequencies $(f_1, f_2, \ldots, f_K)$ are used in the SSVEP-based

BCI, then the reference signals for each stimulus frequency are computed as follows:

$$R_{fi} = \begin{bmatrix} \sin(2\pi \cdot f_i \cdot t) \\ \cos(2\pi \cdot f_i \cdot t) \\ \vdots \\ \sin(2\pi \cdot N_h \cdot f_i \cdot t) \\ \cos(2\pi \cdot N_h \cdot f_i \cdot t) \end{bmatrix}, \quad t = \frac{1}{F_s}, \frac{2}{F_s}, \ldots, \frac{M}{F_s} \tag{9}$$

where F_s is the sampling rate.

Next, we can calculate the synchronization index between the signals from the multiple EEG electrodes and each reference signal Y ($Y \in \{R_{f1}, R_{f2}, \ldots, R_{fK}\}$) and then obtain K indices $S_1, S_2, \ldots, S_K$. The stimulus frequency f coding the target at which the user gazing is the frequency that satisfies the following:

$$T = \max_i S_i, \quad i = 1, 2, \ldots, K \tag{10}$$

2.2. CCA-based frequency recognition

Canonical correlation analysis (CCA) is a multivariable statistical method for seeking linear combinations that maximize the correlation between two sets of data. CCA extends ordinary correlation to two sets of variables (Friman et al., 2001; Lin et al., 2006). When using CCA for frequency recognition, we also require the reference signals described in formula (9).

With CCA, we can find the weight vectors W_x and W_y to obtain the maximum canonical correlation between $x = X^T W_x$ and $y = Y^T W_y$ by solving the following optimization problem:

$$\max_{W_x, W_y} \rho(x, y) = \frac{W_x^T X Y^T W_y}{\sqrt{W_x^T \cdot XX^T \cdot W_x \cdot W_y^T \cdot YY^T \cdot W_y}} \tag{11}$$

For each reference signal Y ($Y \in \{R_{f1}, R_{f2}, \ldots, R_{fK}\}$), we can obtain a maximum canonical correlation and then use these coefficients to recognize the target of the user's gaze in the SSVEP-based BCI system as follows:

$$T = \max_i \rho_i, i = 1, 2, \ldots, K \tag{12}$$

where ρ_i are the CCA coefficients obtained with the K reference signals.

Although we can also obtain the smaller correlation coefficients, only the maximum correlation coefficient has been used in the classification in most studies (Bin et al., 2009; Lin et al., 2006; Wang et al., 2011). CCA is discussed in more detail by Lin et al. (2006) and Friman et al. (2001).

2.3. MEC-based frequency recognition

Minimum energy combination (MEC) is a method of finding combinations of electrode signals that remove strong noise and nuisance signals for EEG data (Friman et al., 2007). A variant version of MEC that has been adopted in the online system is presented in this section (Cecotti, 2010; Volosyak, 2011). For SSVEP data stimulated by the frequency f, we can model the response by adding the noise as follows:

$$X^T = Y^T A + E \tag{13}$$

where X is the EEG signal, and Y is the reference signal, as given above. A is the amplitude matrix of size $2N_h \times N$ for all electrode signals, and E is the noise matrix of size $M \times N$.

The signals from multiple electrodes should be combined to extract the discriminative features. This combination can be achieved by linear transformation of X. Suppose that W is a matrix of size $N \times N_s$ that contains the combination weights, where Ns is determined by formula (19). Then, we obtain

$$S = X^T W \tag{14}$$

Next, the core goal is finding the optimal W such that the nuisance signals and noise can be reduced as much as possible. Several methods to determine W are proposed, but MEC is the best and is widely used (Cecotti, 2010; Friman et al., 2007; Volosyak, 2011). MEC is based on PCA. First, an orthogonal projection is used to remove any potential SSVEP components from X:

$$\widetilde{X} = X^T - Y^T (YY^T)^{-1} Y X^T \tag{15}$$

Then, W is found by optimizing the following formula:

$$\min \|\widetilde{X}^T W\|^2 = \min W^T \widetilde{X} \widetilde{X}^T W \tag{16}$$

The optimization procedure can be achieved by applying PCA to $\widetilde{X}$.W is chosen based on the eigenvalues ($\lambda_1 \le \lambda_2 \le \lambda_3 \ldots$) and the corresponding eigenvectors ($v_1, v_2, v_3, \ldots$) as follows:

$$W = \left[\frac{v_1}{\sqrt{\lambda_1}}, \frac{v_2}{\sqrt{\lambda_2}}, \ldots, \frac{v_{Ns}}{\sqrt{\lambda_{Ns}}} \right] \tag{17}$$

N_s is determined by finding the smallest N_s satisfying the following condition:

$$\frac{\sum_{i=1}^{N_s} \lambda_i}{\sum_{j=1}^{N} \lambda_j} > 0.1 \tag{18}$$

After obtaining W, the SSVEP power is calculated as follows:

$$\widetilde{P} = \frac{1}{N_s N_h} \sum_{l=1}^{N_s} \sum_{k=1}^{N_h} \|Y_k S_l\|^2 \tag{19}$$

Y_k is the kth harmonic component of Y, and S_l is the lth combination signal. For all the reference signals of formula (9), a power estimation for each stimulus frequency ($f_1 f_2, \ldots f_K$) can be calculated. Then, these power estimations are normalized as follows:

$$P_i = \frac{\widetilde{P_i}}{\sum_{j=1}^{K} \widetilde{P_i}}, \quad i = 1, 2, \ldots, K \tag{20}$$

The target at which the user gazes in the SSVEP-based BCI system can be determined as follows:

$$T = \max_i p_i, \quad i = 1, 2, \ldots, K \tag{21}$$

For details about MEC, please refer to the references mentioned above.

2.4. Simulation

To test the performance of the MSI, we computed the recognition accuracies of the three methods for various SNRs. We chose three groups of frequencies for the simulation. The first group included 27 Hz, 29 Hz, 31 Hz, 33 Hz, 35 Hz, 37 Hz, 39 Hz, 41 Hz and 43 Hz, the frequencies adopted by Lin et al. (2006). The second group included 8 Hz, 9 Hz, 10 Hz, 11 Hz, 12 Hz, 13 Hz, 14 Hz, and 15 Hz, the frequencies chosen by Nan et al. (2011). The third group included 6.7 Hz, 7.5 Hz, 8.6 Hz, 10 Hz, 12 Hz, and 15 Hz, the frequencies used by Bin et al. (2009). For each frequency, we generated 4 sinusoidal signals to simulate 4 channels of SSVEPs. The signals were created using the following formula:

$$a \sin(2\pi \cdot f \cdot t + \phi_1 + \sin(2\pi \cdot 2 \cdot f \cdot t + \phi_{2)} \tag{22}$$

The sampling rate was 250 Hz, and the signals lasted for 10 s. Then, Gaussian white noise was added to the sinusoidal signals to simulate the noise-contaminated signals. To observe the influence of the SNR on the accuracy of these methods, SNRs ranging from −7 db to −20 db were chosen. The general definition of SNR is as follows:

$$\mathrm{SNR} = 10\log\frac{P_{\text{signal}}}{P_{\text{noise}}} = 10\log\frac{(A/\sqrt{2})}{\sigma^2} \tag{23}$$

where P_{signal} and P_{noise} are the power of the signal and the power of the noise, respectively. A is the amplitude of the sinusoidal signals, and σ^2 is the variance of the noise (Lin et al., 2006; Nan et al., 2011).

Finally, we performed frequency recognition using the 3 methods for each group of frequencies. The data length for recognition was 1 s. Therefore, each frequency was presented ten times during each 10 s long data, resulting in 90, 80, and 60 total recognition operations for the three frequency groups. For each frequency group, the evaluation index of the recognition performance was the accuracy, which was defined as the ratio of the number of correct recognition operations to the number of total recognition operations. This procedure was repeated 50 times for each SNR level, and then the average accuracy and standard deviations were calculated for each SNR level using each of the 3 methods.

2.5. *Offline experiment*

Four frequencies, 7.5 Hz, 8.6 Hz, 10 Hz, and 12 Hz, were chosen for the offline experiment. The flickering stimulus was controlled using a computer through a control program written in C++ Builder and based on the Windows DirectX API. A laptop with a 13″ screen and a 60 Hz refresh rate was used to present the stimuli. The subjects were instructed to gaze binocularly at each frequency stimulus for 30 s, followed by a rest period of approximately 1–2 min. Eleven paid healthy right-handed subjects (two female and nine male, age ranging from 21 to 25 years) participated in this study. All subjects had normal or corrected-to-normal vision. These subjects had no risk of epileptic seizure. Six were naive to the SSVEP-based BCI equipment and paradigm.

The experiment was performed in a normal room without electromagnetic shielding. The subjects were seated in a comfortable armchair, 60 cm away from the center of the laptop monitor. Eight Ag/AgCl recording electrodes with gel were used for the EEG recordings with a Symtop Amplifier (NIL System, Chengdu, China). Fcz was adopted as the reference, and Afz was adopted as the ground. The location of these electrodes is shown in Fig. 1.

Data were recorded at a 1000 Hz sampling rate with a band-pass filter from 0.5 to 30 Hz and a 50 Hz notch filter for the line frequency interference (50 Hz in China). Impedances were kept below 5 kΩ. The EEG signals were read, stored and further processed using the self-developed software programmed in C++ Builder. All the raw EEG signals were recorded for further offline analyses.

During this offline experiment, we evaluated the performances of the three methods for various conditions, including different numbers of channels and different time window lengths. Three groups with 4, 6 and 8 channels were selected, and the placements of these channels were similar to those in previous studies (Friman et al., 2007; Nan et al., 2011; Zhang et al., 2012). For each channel group, we repeated the frequency recognition procedure using five time window lengths, i.e., 0.5 s, 0.8 s, 1 s, 1.5 s and 2 s. For each time window length, we extracted non-overlapping segments from the 30 s data of each frequency and pooled all the segments for the four frequencies together. Afterward, we used the 3 methods to conduct the frequency recognition. The accuracy, which was the ratio of the number of segments correctly classified to the number of total segments, was used to evaluate the performances of the three methods.

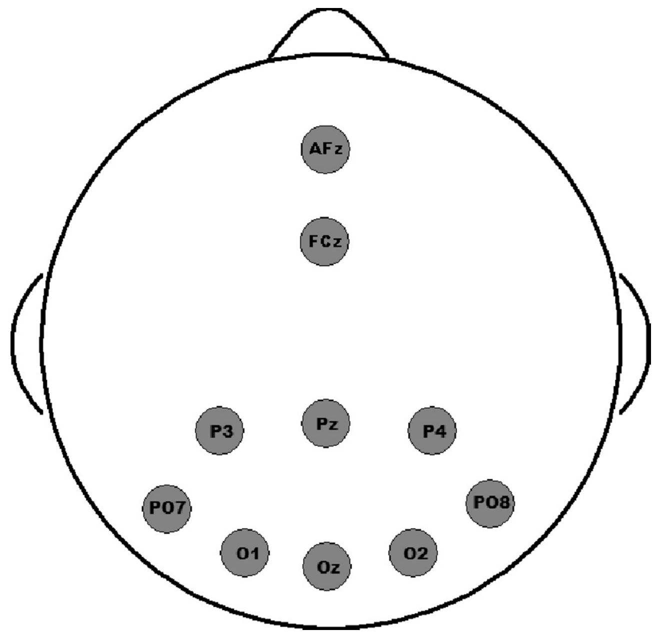

Fig. 1. Electrode placement on the scalp during our experiment.

2.6. *Online SSVEP-BCI experiment*

In our online Chengdu BCI system, an LCD monitor (Acer 17″) with a 60 Hz refresh rate was used to display the stimulus targets and the application interface. Four targets were coded at 12.8 Hz and 8 Hz using the multiple frequencies sequential coding (MFSC) protocol (Zhang et al., 2012). For this protocol, each target was coded by one permutation of the two frequencies. In this study, the four permutation sequences that coded the four targets were 8–8 Hz, 8–12.8 Hz, 12.8–8 Hz, and 12.8–12.8 Hz, respectively. For implementing MFSC, a stimulus coding cycle was introduced, which was defined as a period in which the permutation sequences code the targets. Each cycle was further divided into two equal time epochs for two frequencies in one permutation to present sequentially and code a target. Between each cycle, there was an interval for the users to shift their gaze. During this interval, no flickering stimuli were presented. In the practical application, the frequency recognition procedure was completed during each blank interval to provide a result identifying the target at which the user gazed during last cycle. In this study, the duration of each epoch in one cycle was 1.5 s, and the interval between cycles was 1 s.

The application controlled a robot using our SSVEP-based BCI. The four targets coded the left hand, right hand, left foot and right foot of the robot. When a target was recognized, the corresponding limb of the robot showed movement. A picture of the application interface is shown in Fig. 2.

The settings for acquiring the EEG were the same as those for the offline experiment. In the online system, the MSI was implemented as the frequency recognition method, and the entire BCI system was developed by our lab using C++ Builder. After several minutes of training, the subjects had to run two sessions of online tests. In each session, the subjects randomly controlled the robot' limbs twenty times. The times for selecting the four limbs were balanced during this process.

Ten healthy right-handed subjects (one female and nine male, age ranging from 21 to 25 years) were recruited to participate in this study. All subjects had normal or corrected-to-normal vision. These subjects had no risk of epileptic seizure. Five were naive to the SSVEP-BCI paradigm.

Fig. 2. A picture of the application interface.

The accuracy, which was the correct recognition commands out of 40 desired commands, was used to evaluate the BCI performance. In addition, we also calculated the information transfer rate (ITR), which has been widely used to evaluate the performance of BCI systems (Wolpaw et al., 2002).

Suppose the following: an SSVEP-BCI system has N possible targets; each target has the same probability of being desired by the users; the probability P that the desired target will actually be selected is always the same; and each of the other (i.e., undesired) targets has the same probability of being selected (i.e., $(1-P)/(N-1)$). Then, the bit rate (in bits min^{-1}) can be computed as follows:

$$\frac{\text{Bits}}{\text{Target}} = \log_2 N + P\log_2 P + (1-P) \times \log_2 \left[\frac{1-P}{N-1}\right] \quad (24)$$

$$\text{BitRate} = \frac{\text{Bits}}{\text{Target}} \times N_T \quad (25)$$

where N_T is the number of correct commands in 1 min.

3. Results

3.1. Simulation

For all three methods, two main parameters must be set, i.e., the number of harmonics (N_h) and the number of channels (N). For the number of harmonics (N_h), the studies using CCA and MEC have demonstrated that an N_h of 2 consistently generates satisfying results (Bin et al., 2009; Volosyak, 2011). The number of channels varies from 4 to 9 in SSVEP-BCI systems (Bin et al., 2009; Cecotti, 2010; Friman et al., 2007; Lin et al., 2006; Nan et al., 2011; Volosyak, 2011; Volosyak et al., 2011; Wang et al., 2011; Zhang et al., 2012). In our simulation, the number of harmonics was set to 2, and the number of channels was set to 4.

We adopted three frequency groups, as mentioned above, to test the new method. One was the high frequency group, another one was the low frequency group, and the last one included harmonic frequencies. The average recognition accuracies of the three methods at different SNRs for the three conditions are shown in Fig. 3(a)–(c).

For the three conditions, the MSI always showed higher accuracy and robustness to decreased SNRs. This finding indicated that the MSI might have broad applicability to different types of frequencies. When the SNR was larger than −12 db, the MSI significantly differed from CCA and MEC for the high and low frequencies conditions. In addition, the accuracies of all methods were less than 100% when the frequency group including harmonic frequencies, but MSI still showed the best performance and significantly differed from the other methods for −7 db.

Furthermore, we also studied the effect of the number of channels on the three methods. We chose the high frequencies to implement this simulation and set the SNR to −15 db. Fig. 4 illustrates that more channels improve the accuracies of the three methods and that the MSI also shows higher accuracy and smaller standard deviation. There are no differences between the three methods when only one signal channel is used, as shown in Fig. 3.

3.2. Offline EEG test

The simulation data may not reproduce the complexity of the EEG signals. We used the offline experimental data from our lab to further evaluate the MSI performance. Eight channels were used. These electrodes were symmetrically distributed on the two hemispheres and covered the parietal-occipital region, which was widely used in other SSVEP-BCI systems. To test the MSI with these real SSVEP data, we selected three groups of channels based on the settings in other studies (Friman et al., 2007; Nan et al., 2011; Zhang et al., 2012). The first group of 4 channels consisted of P3, P4, O1 and

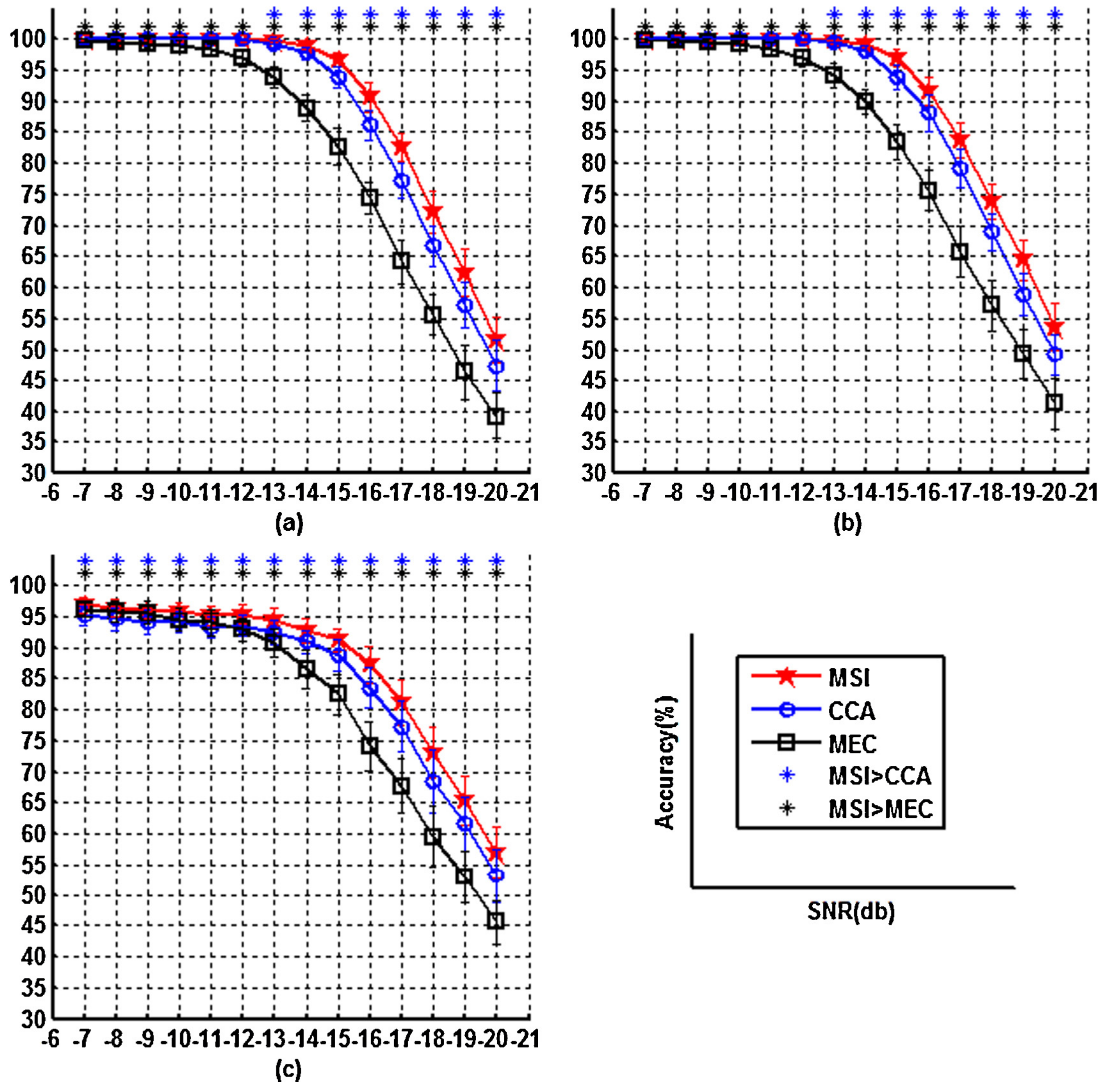

Fig. 3. Simulation recognition accuracies and standard deviation of the three methods at different SNR for three conditions. (a) High frequency group, (b) low frequency group and (c) frequency group that containing harmonics. The asterisk denotes a significant difference between two methods (paired *t*-test, $p < 0.05$). The error bars represent standard deviations. A data length of 1 s was used for recognition.

O2 (Zhang et al., 2012). The second group of 6 channels consisted of P3, P4, O1, O2, Pz and Oz (Friman et al., 2007; Nan et al., 2011). The third group of 8 channels consisted of P3, P4, O1, O2, Pz, Oz, PO7 and PO8. The offline data were resampled to 250 Hz for the simulation test. 4 data lengths, i.e., 0.5 s, 0.8 s, 1 s, 1.5 s and 2 s, were used during this test.

Fig. 5 presents the results of three methods for different combinations of the number of channels and the time window length. Similarly to the simulation, the MSI always performed better than CCA and MEC. The MSI was particularly better than the other methods when a short data length was used. Considering the tests with simulated and real data, we believe that the MSI might be a promising new candidate for the frequency recognition method in SSVEP-BCI systems.

3.3. Online SSVEP-BCI test

From results shown in Fig. 5, the average accuracies were above 85% when the time window length was 1.5 s and all eight electrodes were used. This result was used as a guideline when we ran the online system to set the following parameters: the duration for each frequency in one cycle (1.5 s) for coding the targets with MFSC and the interval between cycles that was provided to the subject to shift his gaze (1 s). The results from ten subjects are shown in Table 1. The average accuracy of all subjects was 86%, and the ITR was 19.83 bit min^{-1}. For most subjects, the MSI could decode their intentions. Some subjects performed worse than others. This difference across subjects may be due to different attention strategies, fatigue, or other factors. In addition, some subjects, i.e., S2, S4, S5,

Table 1
Performance of each subject in the online test.

Subject	Accuracy(%)	ITR(bit min^{-1})
S1	100	30.00
S2	75	11.89
S3	100	30.00
S4	60	5.93
S5	95	24.52
S6	95	24.52
S7	75	11.89
S8	90	20.59
S9	80	14.42
S10	90	20.59
Average(std)	86(13.08)	19.43(8.16)

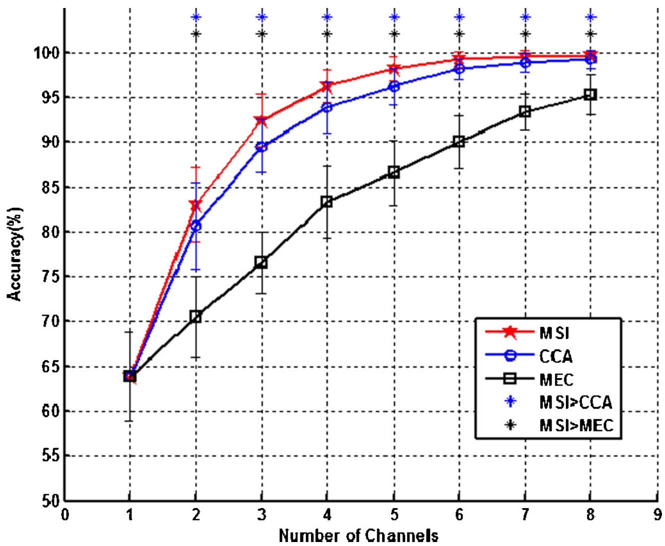

Fig. 4. Simulation recognition accuracies and standard deviation of the three methods with different numbers of channels for the high frequency conditions. The error bars represent the standard deviations. The asterisk denotes a significant difference between the two methods. (paired *t*-test, $p < 0.05$). The error bars represent standard deviations. The data length is 1 s.

S7, and S9, were naive to BCI applications; thus, the performance between these subjects varied greatly.

4. Discussion

A good detection method is a core component in a high performance BCI system. Multichannel detection methods benefit from an optimized combination of multiple signals and have greater robustness against noise, thus improving the results. Furthermore, these methods require minimal or no parameter optimization, making the applications convenient.

In this paper, a novel multiple channel frequency recognition method (MSI) is proposed. For the first time, to the best of our knowledge, the S-estimator algorithm is adopted to estimate the synchronization between EEGs and reference signals in the MSI for frequency recognition in SSVEP-BCI. The S-estimator is a nonlinear measure of synchronization based on dynamical systems theory. This algorithm is robust and highly sensitive even with a reasonably small amount of data for assessing the synchronization, i.e., few channels and short measurements (Carmeli, 2006; Jalili et al., 2007). There characteristics may guarantee better performance with a small number of electrodes and short data lengths, as shown in Figs. 4 and 5. The small number of electrodes will make the user more comfortable. In addition, accurately detecting the intention of the user with short data lengths is crucial for developing a high-performance system for SSVEP-based BCI applications (Wu et al., 2011; Wu and Yao, 2008). Furthermore, short data acquisition can prevent fatigue to some extent because of shorter gazing time. For CCA and MEC, the potential disadvantage is the linear combination used in both methods. The linear combinations in CCA and MEC may miss useful data descriptors (Zhang and Deng, 2012). The advantage of the MSI may be that it avoids the linear combinations. In the future, we will elaborate on the theoretical analysis of those methods. In addition, other synchronization estimation methods must be tested to search for the most efficient MSI implementation.

MSI and CCA can be considered tools for measuring the synchronization or correlation between two sets of vectors. When the two sets of vectors are independent, the indices of MSI and all the canonical correlations of CCA are zeros, and CCA and MSI present the same results. When the two sets of vectors are not independent, CCA and MSI may provide different measures to evaluate the relationships. Another potential measure for capturing the relationships between two signal sets is the log-likelihood ratio (LLR), which can be used to test whether two sets of vectors are independent (Anderson, 2003). The LLR may be another important tool for frequency recognition in SSVEP-based BCIs. Developing a novel frequency recognition method based on LLR would be a valuable topic for future study.

MSI, CCA and MEC all create reference data from the stimulus frequencies used by the system. The number of harmonics (N_h) is a key parameter when creating the reference data. However, the number of harmonics does not significantly influence the performance of the reference matrices, and using two harmonic components in the reference data is sufficient for desirable outcomes in some actual systems (Bin et al., 2009; Volosyak, 2011; Volosyak et al., 2011; Zhang et al., 2012). With this consideration, we adopted two harmonics in our current study. A major advantage of the multiple channels methods is the minimal work required for parameter optimization and electrode selection.

One question we considered was whether MSI could be suitable for different frequency combination groups when designing the system, e.g., high frequencies, low frequencies and frequencies that included harmonics. To test that idea, we chose three types of frequency groups in the simulation. These simulations were run with a short data length (1 s) and a small number of channels (4 channels). The results showed that the MSI performed better than the other methods, even when the harmonic frequencies were used. Increasing the number of channels and the data length can improve the results, as shown in Figs. 4 and 5. The real EEG signals were more complex than the simulation data and may have been contaminated by noise. Thus, more channels may be needed to achieve a better result. In Fig. 5, when the number of channels increased from 4 to 8, the average accuracies also increased for all methods. The MSI and the other two methods did not differ significantly when more channels were used. However, MSI differs significantly from CCA when using eight channels, but not six channels, for the 1.5 s condition. Using signals from a broader area may increase the risk of significant noise, which negatively affects frequency detection. For very short (0.5 s or 0.8 s) data lengths, the SNR of the data may be quite low. Thus, there may be no significant differences between MSI and CCA. However, these phenomena may always indicate that MSI is more robust to noise. This property may stem from the robustness of the S-estimator to dynamical noise (Jalili et al., 2007). Increasing the data length can increase the SNR, which will improve the recognition accuracy, as shown in Fig. 5(a)–(c).

The ITR in Table 1 may be smaller than that of some systems (Cecotti, 2010; Volosyak, 2011). The ITR depends on the number of targets in the BCI system and the user's performance (Wolpaw et al., 2002). In our online test, we used only four targets, fewer than those used in other systems. Accuracy may also reduce the ITR. The frequencies used to encode the targets may influence the recognition performance to some degree because different subjects may have preferred frequencies (Srinivasan et al., 2006, 2007; Zhang et al., 2012). In the current work, we did not optimize the frequency for each subject and only used 12.8 Hz and 8 Hz for all subjects. As shown in Table 1, the accuracy of S2 was lower than that of the other subjects; this result might be because the selected frequencies cannot evoke the strongest SSVEPs for this subject. The poor performance resulted in a very low ITR, but this phenomenon has also been observed in BCI research (Cheng et al., 2002). In spite of this phenomenon, our online test demonstrated the feasibility of MSI. In the future, we will apply this new method to designing practical systems, such as a speller system with the MFSC protocol. It is likely that we can obtain an ideal ITR system using this frequency recognition method.

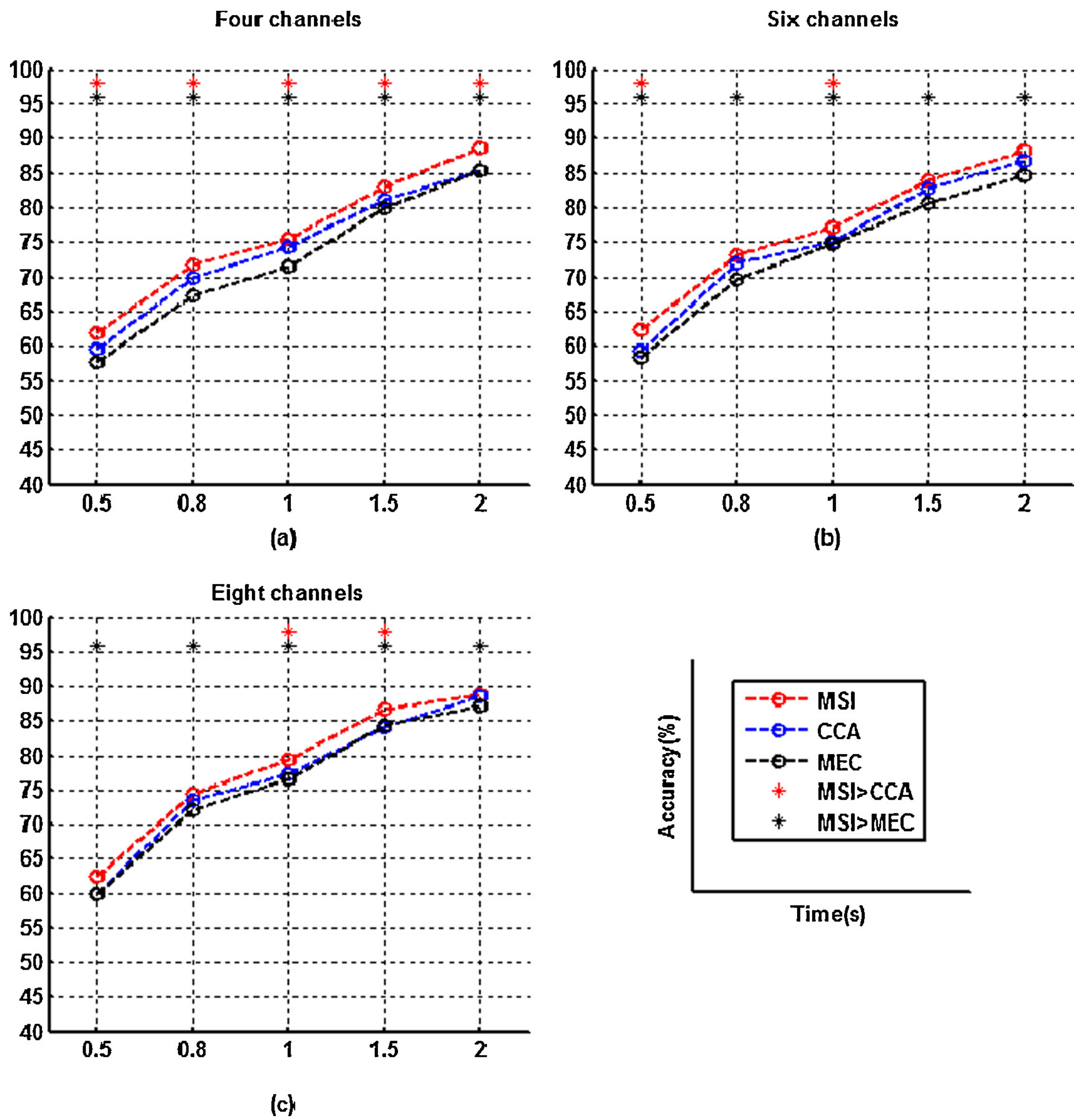

Fig. 5. Average accuracies of the three methods for different time window lengths and different numbers of channels. (a) 4 channels (P3, P4, O1 and O2), (b) 6 channels (P3, P4, O1, O2, Pz and Oz), (c) 8 channels (P3, P4, O1, O2, Pz, Oz, PO7 and PO8). The asterisk denotes the significant difference between two methods (paired *t*-test, $p < 0.05$).

5. Conclusion

We proposed a new frequency recognition method based on a multivariate synchronization index (MSI) and verified its efficiency with both simulation data and offline real EEG data. The results indicated that the MSI performed better than the widely used CCA and MEC. Using these results, we developed a simple online SSVEP-BCI system to further test the performance of the MSI. The MSI not only has the advantages of no calibration data, no channel selection and no parameter optimization but also demonstrated higher detection accuracy and robustness, which could potentially enhance the ITR of BCI systems. The MSI may be a promising alternative for frequency recognition in the future.

Acknowledgments

This work was supported by the 973 project (2011CB707803), the 863 project (2012AA011601), and the NSFC (Nos. 31070881 and 91232725) and partially supported by the 111 project and PCSIRT.

We thank Chi Man Wong (University of Macau, Macau SAR, China) and Cristian Carmeli (Neuroimaging Research Lab, Lausanne University Hospital) for helpful discussions. We also thank the reviewers for their insightful comments.

References

Allison BZ, Wolpaw EW, Wolpaw JR. Brain–computer interface systems: progress and prospects. Expert Rev Med Devices 2007;4:463–74.

Anderson TW. An introduction to multivariate statistical analysis. 3rd ed. Hoboken, NJ: John Wiley & Sons; 2003.

Bin G, Gao X, Yan Z, Hong B, Gao S. An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. J Neural Eng 2009;6:046002.

Carmeli C. Assessing cooperative behavior in dynamical networks with applications to brain data. Lausanne: EPF Lausanne; 2006.

Carmeli C, Knyazeva MG, Innocenti GM, De Feo O. Assessment of EEG synchronization based on state-space analysis. Neuroimage 2005;25:339–54.

Cecotti H. A self-paced and calibration-less SSVEP-based brain–computer interface speller. IEEE Trans Neural Syst Rehabil Eng 2010;18:127–33.

Cheng M, Gao X, Gao S, Xu D. Design and implementation of a brain–computer interface with high transfer rates. IEEE Trans Biomed Eng 2002;49:1181–6.

Friman O, Cedefamn J, Lundberg P, Borga M, Knutsson H. Detection of neural activity in functional MRI using canonical correlation analysis. Magn Reson Med 2001;45:323–30.

Friman O, Volosyak I, Graser A. Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces. IEEE Trans Biomed Eng 2007;54:742–50.

Jalili M, Lavoie S, Deppen P, Meuli R, Do KQ, Cuenod M, et al. Dysconnection topography in schizophrenia revealed with state-space analysis of EEG. PLoS ONE 2007;2:e1059.
Jensen O, Bahramisharif A, Oostenveld R, Klanke S, Hadjipapas A, Okazaki YO, et al. Using brain–computer interfaces and brain-state dependent stimulation as tools in cognitive neuroscience. Front Psychol 2011;2:100.
Joudaki A, Salehi N, Jalili M, Knyazeva MG. EEG-based functional brain networks: does the network size matter? PLoS ONE 2012;7:e35673.
Lin Z, Zhang C, Wu W, Gao X. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Trans Biomed Eng 2006;54:1172–6.
Nan W, Wong CM, Wang B, Wan F, Mak PU, Mak PI, et al. A comparison of minimum energy combination and canonical correlation analysis for SSVEP detection. In: The 5th international IEEE EMBS neural engineering conference (NER2011); 2011. p. 469–72.
Nicolas-Alonso LF, Gomez-Gil J. Brain computer interfaces, a review. Sensors (Basel) 2012;12:1211–79.
Regan D. Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. Elsevier: New York; 1989.
Shyu KK, Lee PL, Liu YJ, Sie JJ. Dual-frequency steady-state visual evoked potential for brain computer interface. Neuroscience Letters 2010;483:28–31.
Srihari Mukesh TM, Jaganathan V, Reddy MR. A novel multiple frequency stimulation method for steady state VEP based brain computer interfaces. Physiological Measurement 2006;27:61–71.
Srinivasan R, Bibi FA, Nunez PL. Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency. Brain Topogr 2006;18:167–87.
Srinivasan R, Fornari E, Knyazeva MG, Meuli R. Maeder P fMRI responses in medial frontal cortex that depend on the temporal frequency of visual input. Exp Brain Res 2007;180:677–91.
Vialatte FB, Maurice M, Dauwels J, Cichocki A. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. Progr Neurobiol 2010;90:418–38.
Volosyak I. SSVEP-based Bremen–BCI interface – boosting information transfer rates. J Neural Eng 2011;8:036020.
Volosyak I, Moor A, Gräser A. A dictionary-driven SSVEP speller with a modified graphical user interface. Adv Comput Intell 2011;6691:353–61.
Walker DM, Carmeli C, Pérez-Barbería F, Small M, Pérez-Fernández E. Inferring networks from multivariate symbolic time series to unravel behavioural interactions among animals. Anim Behav 2010;79:351–9.
Wang Y, Wang R, Gao X, Hong B, Gao S. A practical VEP-based brain–computer interface. IEEE Trans Neural Syst Rehabil Eng 2006;14:234–9.
Wang YT, Wang Y, Jung TP. A cell-phone-based brain–computer interface for communication in daily life. J Neural Eng 2011;8:025018.
Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clin Neurophysiol 2002;113:767–91.
Wu CH, Chang HC, Lee PL, Li KS, Sie JJ, Sun CW, et al. Frequency recognition in an SSVEP-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing. J Neurosci Methods 2011;196:170–81.
Wu Z, Yao D. Frequency detection with stability coefficient for steady-state visual evoked potential (SSVEP)-based BCIs. J Neural Eng 2008;5:36–43.
Zhang Y, Xu P, Liu T, Hu J, Zhang R, Yao D. Multiple frequencies sequential coding for SSVEP-based brain–computer interface. PLoS ONE 2012;7:e29519.
Zhang ZM, Deng ZD. A Kernel canonical correlation analysis based idle-state detection method for SSVEP-based brain–computer interfaces. Adv Mater Res 2012;341:634–40.