A Novel Neurotechnology and Computational Intelligence Method Applied to EEG-based Brain-Computer Interfaces

Chin-Teng Lin[1,2,*], Yu-Ting Liu[1,2], Shang-Lin Wu[2], Zehong Cao[1,2], Yu-Kai Wang[2], Chih-Sheng Huang[2], Jung-Tai King[2], Shi-An Chen[2], Shao-Wei Lu[2], Chun-Hsiang Chuang[1,2,*]

***Abstract*—Significant advances in neuroscience, sensor technologies, and efficient signal processing algorithms have greatly facilitated the transition from laboratory-oriented neuroscience research to practical applications. Brain-computer interfaces (BCIs) represent major strides in translating brain signals into actionable decisions and primarily consist of hardware and software that guide the communications between users and systems. This article presents several current neurotechnologies and computational intelligence methods applied to EEG-based BCIs. In the hardware aspect, novel portable EEG devices featuring dry electrodes are introduced as substitutes for traditional BCIs with wet electrodes and its bulky size. With these advantages, these novel EEG devices can acquire real-time EEG signals for operational workplaces without requiring conductive gel/paste or scalp preparations. As for the software aspect, blind source separation, artificial neural networks, effective connectivity measurements and information fusion techniques are introduced to address the technical issues of artifact removal, rapid event-related potential detection, complex brain network description, and decision fusion, respectively. For instance, information fusion technique has been utilized to attack the individual differences problem of motor imagery applications in the real-world environment. With continuous improvements in the development of a convenient approach to record brain signals and extract knowledge regarding intentions, BCI techniques are envisioned to lead to a wide range of real-life applications in the near future.**

This work was supported in part by the UST-UCSD International Center of Excellence in Advanced Bio-engineering sponsored by the Taiwan Ministry of Science and Technology under Grant Number partially by MOST 104-2627-E-009-001 and MOST 105-2221-E-009-191, and partially by the Aiming for the Top University Plan of National Chiao Tung University and the Ministry of Education, Taiwan under Contract 105W963. Research was also sponsored in part by the Army Research Laboratory and performed under Cooperative Agreement Number W911NF-10-2-0022. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the US Government. The US Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein.

Chin-Teng Lin, Yu-Ting Liu, Zehong Cao, Chun-Hsiang Chuang are with Faculty of Engineering and Information Technology, University of Technology Sydney, Australia and Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan (e-mail: ctlin@mail.nctu.edu.tw; tingting76319@gmail.com; zhcaonctu@gmail.com; chchuang@ieee.org).

Shang-Lin Wu, Yu-Kai Wang, Chih-Sheng Huang, Jung-Tai King, Shi-An Chen, Shao-Wei Lu are with the Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan (e-mail: slwu19870511@gmail.com; yukaiwang@cs.nctu.edu.tw; chih.sheng.huang821@gmail.com; jtchin2@gmail.com; sachen@nctu.edu.tw; swlucifer@gmail.com).

***Index Terms*—Brain-computer interfaces, Biosensors, Computational intelligence, Biomedical signal processing**

I. Introduction

Recently, brain-computer interfaces (BCIs) [1], [2] have been shown to be the most promising conduits for individuals with disabilities or reduced mobility to communicate with external environments or trigger surrounding devices. BCIs have also been shown to be successful in a wide range of applications, such as personal authentication or identification [3], [4], assessment of emotional disorders [5], games [6], and accident prevention [7-10]. However, several technical issues in signal acquisition, signal preprocessing, feature extraction, and signal translation must be addressed to facilitate the transition of laboratory-oriented neuroscience research to practical BCI devices (see Fig. 1).

Monitoring the neurophysiological activities involved with motion in a naturalistic environment using vibration-sensitivity equipment, such as functional magnetic resonance imaging [11] or positron emission tomography [12], represents a significant measurement challenge. The electroencephalogram (EEG) is currently the preferred device for non-invasively imaging humans' brains in BCIs as they performing tasks that involve natural movements in a real-world environment [13]. However, in conventional EEG devices, placing the electrodes on the scalp with a conductive gel or paste is one of the common ways to measure the brain's electrical activity. Nevertheless, such electrodes, termed "wet electrodes", require a time-consuming preparation process; therefore, BCI systems are difficult to be applied outside of laboratory-scale experiments.

Regarding data quality, the measured brain signals are easily contaminated with artifacts originating from non-cerebral origins. The amplitude of these artifacts commonly generated by ocular and muscle activities can be quite large and may thus mask the cortical signals of interest, bias the analysis and interpretation, and affect the performance of the BCI [14], [15]. Several blind source separation techniques have been proposed for signal preprocessing to remove such artifacts. Blind source separation is called "blind" because the axes of projection and

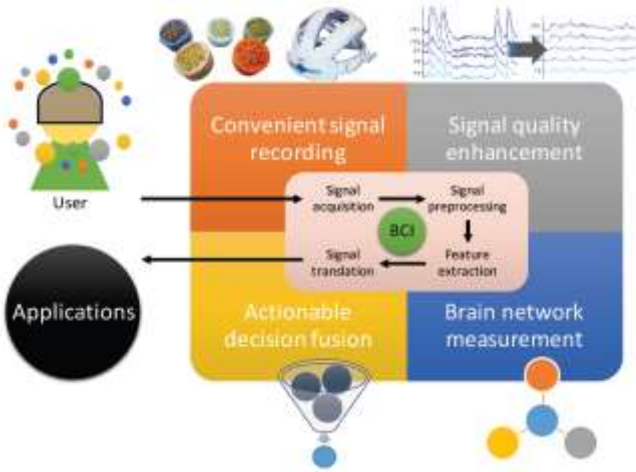

Fig. 1. Current neurotechnology and computational intelligence methods applied to enhance BCI performance.

therefore the sources are determined through the application of an internal measure, which are without using any prior knowledge of the data structure. For example, independent component analysis (ICA) [16-18], which produces the maximally temporally independent signals available in the EEG recording, is a powerful tool for suppressing artifacts. However, the iterative process of measuring the independence within multichannel recordings is computationally intractable. In addition, manually excluding the independent components related to artifacts is still a time-consuming and offline process.

The use of event-related potentials (ERPs) in BCIs is an effective method of basic communication [19]. The classification accuracy of the ERP-based BCIs is reliant on the number of trials used to analyze the data. The influence of infrequent noises can be eliminated from the recorded EEG data by averaging the ERPs of a large number of repeated trials. The signal-to-noise ratio (SNR) can be increased by averaging a progressively larger number of trials. However, the large number of trials to reduce the computational speed of the BCIs. An efficient algorithm to speed up the convergence estimation of ERPs is highly desirable for BCI applications.

For a BCI system, it is essential to efficiently extract informative features from multi-channel EEG signals. Many useful approaches for analyzing the rhythmic pattern of EEG signals and extracting quantitative EEG features, such as the amplitude values of EEG signals, band powers, power spectral density values, and auto regressive parameters, have been introduced to design BCIs. In the neuroscience field, there has been increasing interest in studies mapping the human brain connectivity in recent years. The importance of this research topic was emphasized in the Human Connectome Project [20], which is devoted to investigating the knowledge of the human brain network. The effective connectivity [21] is one of the most commonly used measurements to identify the causal and directional relationship between different brain regions. It is reasonable to assume that the coupling between spatially separated brain areas can provide complementary information for BCIs.

A considerable amount of multi-modality information is often simultaneously employed and recorded in a so-called hybrid BCI system [22]. These distinct information sources provide various estimations of decision and action from multi-aspect data, which may help improve the system performance. Pelletier et al. [23] illustrate that there is beneficial using multimodal information according to the limitations of one modality, which can often be offset by the strengths of another. In addition, optical signals (i.e., NIRS) is one of the optimal solution for a multimodal approach since its signals do not interfere with electric or magnetic fields [24]. There are several types of EEG signals used to design and operate BCIs, such as P300 event-related potentials (ERPs), and steady-state visually evoked potentials (SSVEPs). Among them, P300 and SSVEP signals have become extremely popular due to the high information transfer rate (ITR) they produce and their minimal user training requirement [25, 26]. Each collected signal possesses its own properties and potential uncertainties to describe the underlying cognitive states. A comprehensive analysis of multiple sources is needed to reduce individual uncertainty and improve the system performance reliability. Therefore, developing an effective approach to integrate multi-modal information is an important and urgent issue.

In this article, current neurotechnology and computational intelligence methods are introduced as possible solutions to address the aforementioned technical issues. Section II introduces a series of EEG capture devices developed to measure EEG signals with channels ranging from 4 to 64 channels. In contrast to conventional wet electrodes, dry electrodes exhibit the electronic characteristics of electrically conductive materials. They obtain high quality signals without skin abrasion or preparation. In these novel EEG devices, dry electrodes act as substitutes for traditional wet electrodes; these dry electrodes can acquire real-time EEG signals for operational workplaces without requiring conductive gel/paste or scalp preparation in BCI applications. An online artifact removal technique based on canonical correlation analysis (CCA) [15] as a blind source separation used to remove artifacts is presented in Section III.A. The feasibility of rapid P3 detection using a radial basis function network (RBFN) [27] under a small number of EEG trials is demonstrated in Section III.B. In Section III.C, one of the commonly used measurements of brain effective connectivity, Granger causality analysis, is introduced to extract detailed changes in the brain network during sustained-attention driving. In Section IV, Dempster-Shafer (D-S) theory [28], [29] is used to aggregate pieces of evidence from multiple information sources and exploit redundancy and complementariness between sources in global information. When applied to integrate different physiological signals, this fusion technique can improve the quality of final decisions and facilitate the optimal estimation of objects.

II. EEG-BASED NEUROIMAGING TECHNOLOGY FOR BCIS

Conventional wet electrodes are commonly used to measure EEG signals. These electrodes provide excellent EEG signals with the proper skin preparation and conductive gel application; however, the skin must be prepared prior to

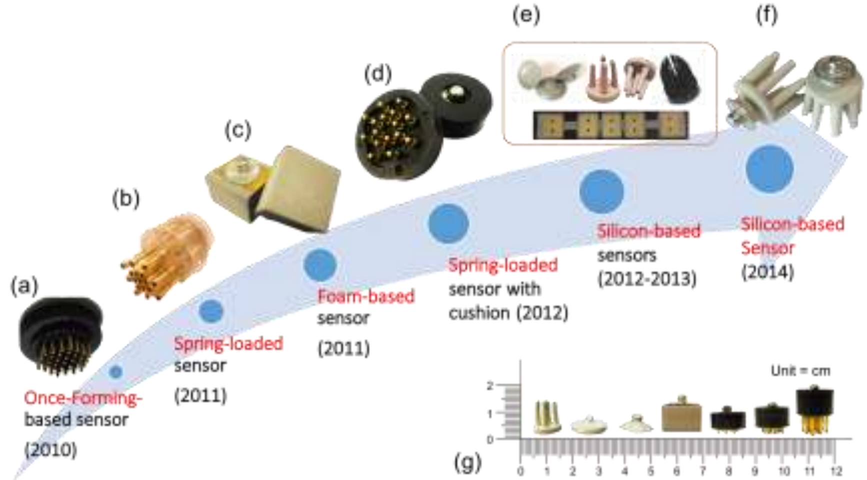

Fig. 2. Novel dry-contact sensors for measuring scalp EEG signals. (a)-(f) Different types and (g) sizes of novel dry sensors have been developed for various purposes in the past years.

applying the wet electrodes, which is typically problematic for users. To overcome these drawbacks, we have developed several types of novel dry-contact EEG sensors that can efficiently reduce the preparation time without conductive gel. Fig. 2 lists several types of dry sensors, including spring-loaded [30], foam-based [31], and silicon-based sensors [32]. The dry foam electrode is fabricated by an electrically conductive polymer foam covered with conductive fabric and can be used to measure bio-potentials without skin preparation or conduction gel. Moreover, the foam substrate of the dry electrode enables a high geometric conformity between the electrode and irregular scalp surface to maintain a low skin-electrode impedance, even under motion. The spring-loaded sensor was proposed for potential operations in the presence of hair and without any skin preparation or conductive gel usage. Each probe was designed to include a probe head, plunger, spring, and barrel. The 17 probes were inserted into a flexible substrate using a one-time forming process in an established injection molding procedure. With 17 spring contact probes, the flexible substrate allows for a high degree of geometrical conformity between the sensor and irregular scalp surface to maintain low skin-sensor interface impedance. Additionally, the flexible substrate also initiates a sensor buffer effect, thereby eliminating pain when force is applied. Most importantly, the data quality obtained with these dry electrodes [31] is comparable to that obtained with wet-electrode systems while avoiding the need for skin abrasion, preparation or gels.

In conventional EEG devices, the measured brain activity is transmitted through a cable connected between the EEG cap and computer, which limits the application and usability of BCIs in real life. To overcome this connection limitation, the developed EEG hats (Fig. 3) include a wireless transmission module and a chargeable battery, which allow recordings to be made without being tethered to a computer; thus, subjects are able to move freely around the room/office. For instance, we can use these wireless and wearable EEG devices to conduct complex experiments, such as drowsy driving [33-40], distracted driving [41], [42], motion sickness [43], [44], or navigation [45] in a motion simulator (Fig. 4(a-b)) or real-world driving environment (Fig. 4(c)). This advantage of the convenient EEG acquisition offers the opportunity to improve our understanding of complex coordinated and multi-joint naturalistic behaviors in operating environments.

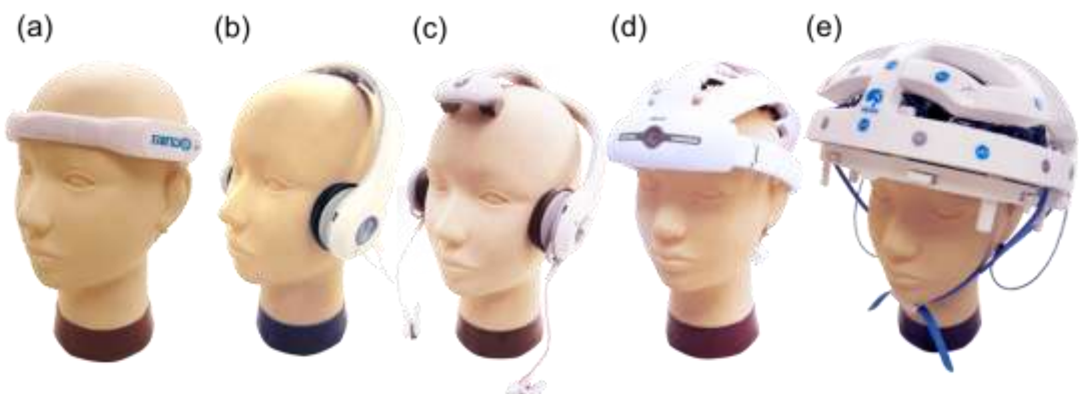

Fig. 3. Different EEG hats carrying different numbers of sensors. (a) Four-channel system that uses silicon-based dry electrodes to measure brain activity on the forehead area. (b) Earphone-like EEG system that uses spring-loaded dry electrodes to measure brain activity on the hairy area. (c) The X-shaped EEG system uses a combination of spring-loaded sensor and foam-based sensor novels to monitor the brain activity at Fp1, Fp2, Pz, and Oz. (d)-(e) High-density EEG systems with 32 and 64 dry electrodes, which are placed based on the international 10-20 system of electrode placement.

Fig. 4. Wearable and wireless EEG devices for convenient EEG recording in operating environments. (a-b) Simulated driving environment. (c) Real-world driving environment.

III. Computational Intelligence for EEG Signal Processing

A. Removal of Artifacts to Enhance Signal Quality

Figure 5(a) presents a flowchart of the artifact removal process used to enhance the EEG signal quality. One can apply a band-pass filter (1-50 Hz) to eliminate high-frequency noise and the DC drift. An artifact-free EEG can be reconstructed after removing the artifacts. Several blind source separation (BSS) techniques [46] have been proposed for artifact removal. One of the BSS techniques is canonical correlation analysis (CCA) [15], which separates the m-channel EEG signals $\mathbf{X}(t) = [\boldsymbol{x}_1(t), \boldsymbol{x}_2(t), \cdots, \boldsymbol{x}_m(t)]^T$ into maximally autocorrelated and mutually uncorrelated sources $\mathbf{S}(t) = [\boldsymbol{s}_1(t), \boldsymbol{s}_2(t), \cdots, \boldsymbol{s}_m(t)]^T$, assuming that the EEG signals are a linear combination of the sources. The linear combination of sources can be represented by the mixing system $\mathbf{X}(t) = \mathbf{A} \cdot \mathbf{S}(t)$, where $\mathbf{A} \in \mathbb{R}^{m\times m}$ is the unknown mixing matrix. The unknown source signals $\mathbf{S}(t)$ can be derived by introducing the de-mixing matrix $\mathbf{W} \in \mathbb{R}^{m\times m}$ such that $\mathbf{W} \cdot \mathbf{X}(t) = \hat{\mathbf{S}}(t)$, where $\hat{\mathbf{S}}(t) \approx \mathbf{S}(t)$. Ideally, $\mathbf{W}$ is the inverse of the unknown mixing matrix $\mathbf{A}$.

The goal of BSS-CCA is to find the matrices $\mathbf{w_x} = \left[w_{x_1} w_{x_2} \cdots w_{x_m}\right]$ and $\mathbf{w_y} = \left[w_{y_1} w_{y_2} \cdots w_{y_m}\right]$ that maximize the correlation ρ between two canonical variates $\mathbf{U}(t) = \mathbf{w}_\mathbf{x}^T\mathbf{X}(t)$ and $\mathbf{V}(t) = \mathbf{w}_\mathbf{y}^T\mathbf{Y}(t)$ as follows:

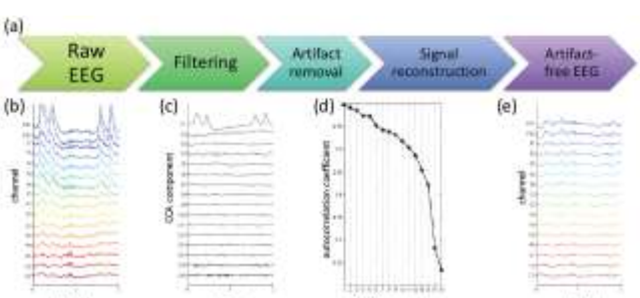

Fig. 5. Artifact-free EEG reconstruction. (a) Flowchart of artifact removal. (b) Raw EEG signals contaminated with eye blinks and muscle noises. (c) CCA components ordered in terms of (d) autocorrelation coefficients. (e) Artifact-free EEG signals.

$$\max_{\mathbf{w_x},\mathbf{w_y}} \rho(\mathbf{U},\mathbf{V}) = \frac{\mathbf{w}_\mathbf{x}^T\mathbf{C_{xy}}\mathbf{w_y}}{\sqrt{(\mathbf{w}_\mathbf{x}^T\mathbf{C_{xx}}\mathbf{w_x})(\mathbf{w}_\mathbf{y}^T\mathbf{C_{yy}}\mathbf{w_y})}}, \tag{1}$$

where $\mathbf{Y}(t) = \mathbf{X}(t-1)$ is the instantly delayed signals of the observed EEG signals, $\mathbf{C_{xx}}$ and $\mathbf{C_{yy}}$ are auto-covariance matrices, and $\mathbf{C_{xy}}$ is the cross-covariance matrix. After calculating the partial derivative with respect to $\mathbf{w_x}$ and $\mathbf{w_y}$, the optimal problem of Eq. (1) is equivalent to the following eigenvalue problem:

$$\begin{cases} \mathbf{C_{xx}^{-1}C_{xy}C_{xy}^{-1}C_{yx}w_x} = \rho^2\mathbf{w_x} \\ \mathbf{C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy}w_y} = \rho^2\mathbf{w_y} \end{cases}, \tag{2}$$

where $\mathbf{w_x}$ and $\mathbf{w_y}$ are eigenvectors and the canonical autocorrelation coefficient ρ^2 is the eigenvalue.

Figure 5(b) shows a 2 s EEG recording contaminated with eye blinks and muscle noises. Figure 5(c) displays the extracted time series of CCA components, which are ordered in terms of autocorrelation coefficients from high to low in Fig. 5(d). Compared to the brain activity components, the components with lower autocorrelation coefficients, i.e., the 15th and 16th CCA components, correspond to muscle artifacts because the broad frequency spectrum of the muscle noise in EEG recordings resemble temporally white noises. By contrast, the 1st CCA component with a relatively higher autocorrelation coefficient corresponds to eye artifacts because eye movements and eye blinks typically produce low-frequency, high-amplitude signals that are highly auto-correlated with time. An artifact-free EEG $\mathbf{X}'(t)$ is reconstructed by removing these artifact components by setting $[\boldsymbol{s}_1(t), \boldsymbol{s}_{15}(t), \boldsymbol{s}_{16}(t)] = 0$ and operating $\mathbf{X}'(t) = \mathbf{A} \cdot \mathbf{S}(t)$, as shown in Fig. 5(e).

B. Radial Basis Function Network for Tracking Evoked Potentials

As shown in Fig. 6(a), many ERP-based BCI applications are designed based on a rapid serial visual presentation paradigm (RSVP), such as image search [47] and auto typing [48]. However, due to the nonstationarity of brain activity, an accurate estimation of the ERP requires the system to average over a large number of trials. In [27], a nonlinear adaptive algorithm referred to as a data-reusing RBFN (DR-RBFN) was proposed to not only estimate the latency and amplitude of brain dynamics but also increase the convergence rate considerably.

Given K previous epochs $\{d(k) \in \mathbb{R}^M | k = 1,2,\cdots,K\}$, the output of the RBFN, denoted as $y(k)$, can be calculated as follows:

$$\begin{gathered} \mathrm{y}(k) = \sum_{j=1}^{N} w_j(k)h_j(x) \\ h_j(x) = \exp\left(-\frac{\left\|x-c_j\right\|^2}{\sigma_j}\right), \end{gathered} \tag{3}$$

where N is the number of hidden units, w is the weight between the hidden layer and output layer, and $\sigma_j = \beta(M-1/N-1)$. In this study, β is set to 0.8. In the kernel function h, the kernel center c can be calculated as follows:

$$c_j = (j-1)\frac{M-1}{N-1} + 1. \tag{4}$$

$w(k)$ can be updated by the least mean squares algorithm as follows:

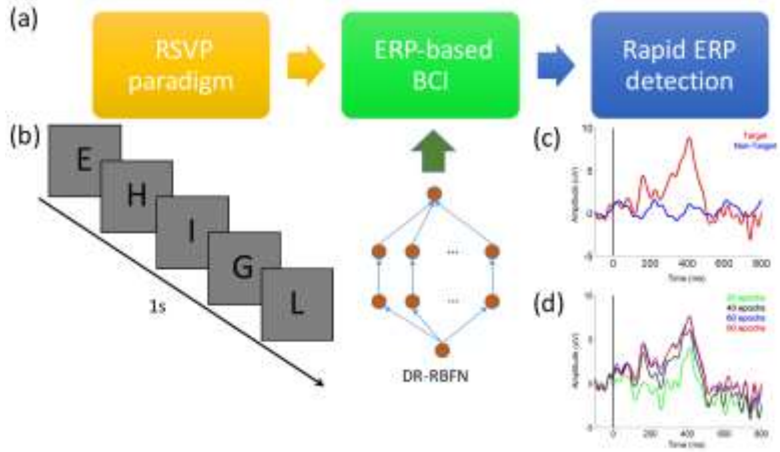

Fig. 6. Rapid ERP detection. (a) RBFN applied to ERP-based BCIs. (b) RSVP paradigm. Uppercase letters were randomly presented to subjects, who were instructed to respond to the target, i.e., the letter "G". Comparison of the ERPs estimated by the EA and DR-RBFN algorithms. (c) ERP estimated by applying EA to 80 epochs. The red and blue traces represent the target- and non-target-evoked ERPs, respectively. (d) Target-evoked ERP estimated by applying the DR-RBFN to 20 (green), 40 (black), 60 (blue), and 80 (red) epochs.

$$w_j(k+1) = w_j(k) + \mu H^T (HH^T + \varepsilon I)^{-1} \cdot e(k)$$

where $e(k) = d(k) - y(k)$ is the error signal, μ is the learning rate, ϵ is a positive constant, and

$$H = \begin{bmatrix} h_1(x_1) & \cdots & h_N(x_1) \\ \vdots & \ddots & \vdots \\ h_1(x_M) & \cdots & h_N(x_M) \end{bmatrix}. \quad (5)$$

Six subjects with normal or correct-to-normal vision participated in the RSVP experiment. As shown in Fig. 6(b), uppercase letters were used as visual stimuli. The letter "G" was predefined as the target, and the other letters were non-targets. All stimuli were presented at a frequency of 5 Hz (200 ms per letter). The interval between two sequential targets is 20-25 non-target stimuli such that the occurrence rate of target is approximately 5%. Subjects were instructed to use their right index finger to press the response button while they detected the target on the screen. In each session, the experiment would end when the subjects detected 80 targets.

The EEG data were recorded using a dry EEG device at a sampling rate of 250 Hz. A low-pass filter with a cut-off frequency of 30 Hz and high-pass filter with a cut-off frequency of 0.5 Hz were applied to remove the line noise and DC drift, respectively. Each EEG epoch of 900 ms began 100 ms before and ended 800 ms after the stimulus onsets were selected from the continuous EEG recordings. Baseline wander was removed by subtracting the mean of the data before stimulus onset. Then, the target-evoked P3 wave estimated by the DR-RBFN [27] was compared with that estimated by ensemble averaging (EA), a conventional approach to assess the ERP. In the DR-RBFN, the number of reused data was set to 3 and the number of hidden nodes was set to 50. The learning rate and positive constant were set to 0.1 and 0.001, respectively. We utilized grid search for the parameter learning in this study. The data-reusing least-mean-square (DR-LMS) algorithm is exploited for real-time implementation of the DR-RBFN. The DR-LMS algorithm reuses data pairs from previous iterations to generate the new gradient estimates that are in turn used to update the adaptive weight vector. This algorithm operates in real-time and has a fast convergence rate and can, thus, track signal variations across trials. The determined parameters in the training stage were further exploited to optimize the proposed DR-DBFN, and enhance the system performance.

The red trace shown in Fig. 6(c) is the average ERP of 80 target epochs in the Pz site estimated by EA. Compared with the average ERP of non-target epochs (blue trace), the P3 elicited by the targets can be easily detected at approximately 400 m. The green, black, blue, and red traces shown in Fig. 6(d) represent the average ERPs of 20, 40, 60, and 80 target epochs, respectively, in the Pz site estimated by the DR-RBFN. Increasing the number of epochs used leads to a higher SNR and more stable ERP. The average ERP of 40 target epochs approximated the results obtained by EA, indicating that the DR-RBFN led to a considerably higher convergence rate. This property of the DR-RBFN is advantageous for real-time BCI applications, as it helps reduce the number of trials required for an accurate estimation and to precisely track potentials.

C. *Causality Analysis for Assessing Brain Connectivity*

The feature extraction is crucial to the system performance of the BCI. Rather than extract the brain activity from a single brain region, the brain network that characterizes some coordinated activity within a network of functionally distinct regions can provide a more detailed description of complex behaviors. Graph theory [49], the dynamic causal model [50], and Granger causality (GC) [51] are the most widely used measures for studying effective connectivity. Take GC for example. GC refers to the fact that signal X_1 can lead to another signal X_2 if the information in the past of X_1 helps predict the future of X_2. We can represent the multivariate process at time t as a stationary autoregressive process of order p. Consider, for example, two signals ($n = 2$).

$$\begin{aligned} X_1(t) &= \sum_{i=1}^{p} A_{11}(i) X_1(t-i) + \sum_{i=1}^{p} A_{12}(i) X_2(t-i) + \xi_1(t) \\ X_2(t) &= \sum_{i=1}^{p} A_{21}(i) X_1(t-i) + \sum_{i=1}^{p} A_{22}(i) X_2(t-i) + \xi_2(t) \end{aligned}, \quad (6)$$

where $t \in \{p+1, p+2, \cdots, T\}$ is the current time point and T is the length of the signal. The model order p is typically obtained by minimizing information criteria, such as the Akaike Information Criterion or Bayesian Information Criterion, to accurately model the data. The parameters A and ξ

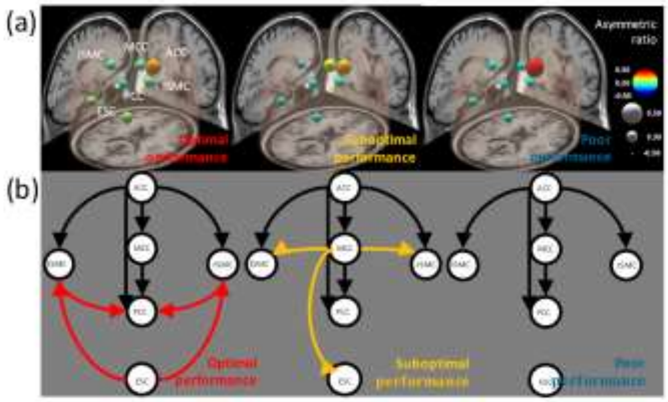

Fig. 7. Effective connectivity between independent EEG processes measured by GC at different levels of behavioral performance. (a) Asymmetric ratios of causal flow (i.e., the difference between the outflow and inflow) (b) Causal connectivity (i.e., the causal interaction between brain regions). (ACC: anterior cingulate cortex, MCC: midcingulate cortex, lSMC: left sensorimotor cortex, rSMC: right sensorimotor cortex, PCC: posterior cingulate cortex, ESC: extrastriate cortex)

model the coefficient matrix and prediction error, respectively, which can be estimated using an ordinary least squares approach [52].

To determine the causal effect of X_2 on X_1, the prediction error is re-estimated using a submodel that excludes signal X_2. Then, the error $\overline{\xi}_1(t)$ is estimated and compared with $\xi_1(t)$ obtained in the full model.

$$X_1(t) = \sum_{i=1}^{p} \overline{A}_{11}(i) X_1(t-i) + \overline{\xi}_1(t). \quad (7)$$

The strength of the effect of X_2 on X_1 (i.e., causal magnitude) was determined as $GC(2 \rightarrow 1) = \ln \overline{\Sigma_{11}} / \Sigma_{11}$, where Σ_{11} and $\overline{\Sigma_{11}}$ are the variances of ξ_1 and $\overline{\xi}_1$, respectively.

Here, we applied GC to independent EEG processes that were collected from a simulated driving experiment in which participants performed a sustained-attention driving task [53]. The asymmetric ratio of the causal flow (Fig. 7(a)) and the significant connectivity of the brain network (Fig. 7(b)) varied with changes in behavioral performance, which were measured by the reaction time in response to unexpected events. During

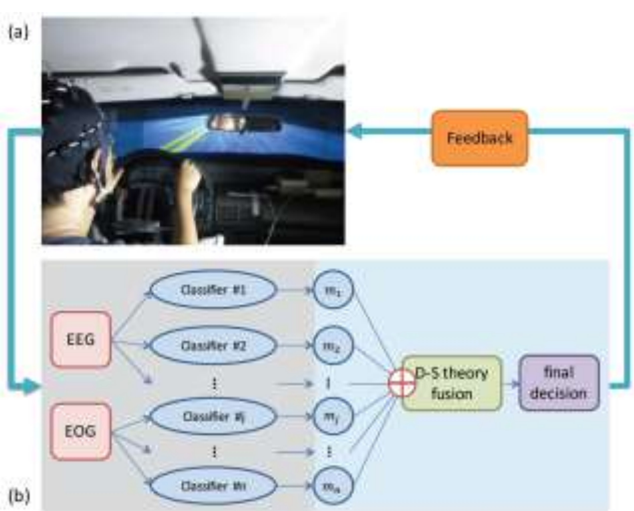

Fig. 8. Multimodal fusion of EEG and EOG signals for a hybrid BCI using D-S theory.

TABLE I
CLASSIFICATION ACCURACIES OBTAINED USING INDIVIDUAL CLASSIFIERS AND ENHANCED BY D-S FUSION.

Modality	**EEG**		**EOG**		**Fusion**
Classifier	NBC	SVM	NBC	SVM	D-S theory
Mean±Std (%)	55.7±5.1	70.9±3.8	59.6±4.8	60.3±2.1	75.1±3.6

the transition from optimal to poor task performance, participants suffered from declining vigilance and fatigue and struggled to avoid behavioral lapses. Under such circumstances, more efforts were needed by subjects to keep themselves engaged in the task, as evidenced by the new connectivity from the MCC to the L/R SMC (Fig. 7B). Fig. 7C shows that PCC- and ESC-related links vanished, which might be related to the fading of consciousness [54]. These reductions of cortico-cortical connectivity produce a cortical gate that disconnects the brain from the external environment and blocks sensory inputs [55]. These results provide new neural markers of behavioral lapses to help neuroengineers design a driving assistance system.

D. Decision Fusion Technique for Multimodal Information Translation

The fusion technique plays an important role in hybrid BCIs that combine two or more sub-BCI systems with different input signal sources. One distinguished approach in the fusion research community is D-S theory [28], [29], a promising approach used to make a final decision from multi-aspect information. D-S theory is a generalized variant of Bayesian probability theory that introduces the notion of assigning beliefs and plausibilities to possible hypotheses of each decision-maker along with the required combination rule to fuse multi-modality information. D-S theory allows each source to incorporate information in different levels of detail. This property is advantageous for assigning a possibility mass to sets or intervals; hence, the fusion system can efficiently consider both stochastic (or objective) uncertainty and epistemic (or subjective) uncertainty.

Consider, for example, two basic probability assignments, m_1 and m_2. An optimal decision, m, can be made by integrating various information from m_1 and m_2 via Dempster's rule [28], [29] as follows:

$$m_1 \oplus m_2(A) = \frac{1}{1-\kappa} \sum_{B \cap C = A} m_1(B) m_2(C), \quad (8)$$

where all $\{A, B, C\} \subseteq 2^{\Theta}, A \neq \emptyset$ and $m_1 \oplus m_2(\emptyset) = 0$. The conflict coefficient $\kappa = \sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)$ measures the degree of conflict between m_1 and m_2. A larger value of κ indicates greater conflict between two sources.

The efficacy of D-S theory in multi-aspect data fusion is demonstrated in a typical BCI application, namely, detecting whether the cognitive state of participants is alert or not during a realistic sustained-attention driving task [41]. Distracted driving experiment consists of an unexpected deviation (swerving) of the car and the presentation of mathematical equations. A flowchart of the proposed system is shown in Fig. 8. The simultaneously recorded EEG and electrooculography (EOG) signals were used to build an ensemble of support vector machines (SVMs) and Naïve Bayes classifiers (NBCs).

Table I provides the classification results of different comparative models reported by five-fold cross-validation. The average accuracies of the NBC and SVM using EEG signals alone are 55.7±5.1% and 70.9.3±3.8%, respectively. The average accuracies of the NBC and SVM using EOG signals alone are 59.6±4.8% and 60.3±2.1%, respectively. When using D-S theory to fuse the outcomes derived from distinct classifiers, the classification accuracy can reach an average value of 75.1±3.6%. These results suggest that multi-modality information with D-S theory fusion can effectively enhance the performance of BCIs.

IV. Conclusions

This article presents the latest BCI-related researches done in our group. Our previous work applied computational intelligence technology in BCIs (i.e., drowsy and distracted driving applications [10]) to inspire detailed investigations of practical issues in real-life applications. Novel EEG devices featuring dry electrodes facilitate and speed up electrode positioning before recording and allow subjects to move freely in operational environments. We also demonstrate the feasibility of applying CCA, RBFNs, effective connectivity measurements and D-S theory to help BCIs extract informative knowledge from brain signals. Two recent trends in research in the computational and artificial intelligence community, Big Data and Deep Learning, are expected to impact the direction and development of BCIs. Those ongoing studies will enable the next generation of BCIs.

References

[1] S. Machado, F. Araujo, F. Paes, B. Velasques, M. Cunha, H. Budde, *et al.*, "EEG-based brain-computer interfaces: an overview of basic concepts and clinical applications in neurorehabilitation," *Rev. Neurosci.*, vol. 21, pp. 451-68, Dec. 2010.

[2] P. Brunner, L. Bianchi, C. Guger, F. Cincotti, and G. Schalk, "Current trends in hardware and software for brain-computer interfaces (BCIs)," *J. Neural Eng.*, vol. 8, pp. 1-7, Apr. 2011.

[3] S. Marcel and J. D. R. Millan, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, pp. 743-752, Apr. 2007.

[4] R. Palaniappan and D. P. Mandic, "Biometrics from brain electrical activity: a machine learning approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, pp. 738-742, Apr. 2007.

[5] R. Sitaram, S. Lee, S. Ruiz, M. Rana, R. Veit, and N. Birbaumer, "Real-time support vector classification and feedback of multiple emotional brain states," *NeuroImage*, vol. 56, pp. 753-765, May 2011.

[6] L.-D. Liao, C.-Y. Chen, I.-J. Wang, S.-F. Chen, S.-Y. Li, B.-W. Chen, *et al.*, "Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors," *J. Neuroeng. Rehabil.*, vol. 9, pp. 1-11, Jan. 2012.

[7] C. Papadelis, Z. Chen, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, E. Bekiaris, *et al.*, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents," *Clin. Neurophysiol.*, vol. 118, pp. 1906-1922, Sep. 2007.

[8] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE Trans. Circuits Syst. Reg. Papers*, vol. 59, pp. 2044-2055, Sep. 2012.

[9] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 121-131, Sep. 2011.

[10] C. T. Lin, L. W. Ko, and T. K. Shen, "Computational intelligent brain computer interaction and its applications on driving cognition," *IEEE Comput. Intell. M.*, vol. 4, pp. 32-46, Nov. 2009.

[11] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner, "Movement-related effects in fMRI time-series," *Magn. Reson. Med.*, vol. 35, pp. 346-355, Mar. 1996.

[12] S. A. Nehmeh and Y. E. Erdi, "Respiratory motion in positron emission tomography/computed tomography: A review," *Semin. Nuclear Med.*, vol. 38, pp. 167-76, May 2008.

[13] L.-D. Liao, C.-T. Lin, K. McDowell, A. E. Wickenden, K. Gramann, T.-P. Jung, *et al.*, "Biosensor technologies for augmented brain-computer interfaces in the next decades," *Proc. IEEE*, vol. 100, pp. 1553-1566, Mar. 2012.

[14] J. N. Knight, "Signal fraction analysis and artifact removal in EEG," Colorado State University, 2003.

[15] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 2583-2587, Nov. 2006.

[16] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, pp. 1129-59, Nov. 1995.

[17] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proc. Natl. Acad. Sci.*, vol. 94, pp. 10979–10984, Sep. 1997.

[18] T.-P. Jung, S. Makeig, M. J. McKeown, A. J. Bell, T. W. Lee, and T. J. Sejnowski, "Imaging brain dynamics using independent component analysis," *Proc. IEEE*, vol. 89, pp. 1107-1122, July 2001.

[19] E. W. Sellers, D. J. Krusienski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance," *Biol. Psychol.*, vol. 73, pp. 242-52, Oct. 2006.

[20] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, *et al.*, "The WU-Minn human connectome project: An overview," *NeuroImage*, vol. 80, pp. 62-79, Oct. 2013.

[21] K. J. Friston, "Functional and effective connectivity in neuroimaging: A synthesis," *Hum. Brain Mapp.*, vol. 2, pp. 56-78, 1994.

[22] S. Amiri, R. Fazel-Rezai, and V. Asadpour, "A review of hybrid brain-computer interface systems," *Adv. Hum.-Comput. Interact.*, vol. 2013, p. 8, Jan. 2013.

[23] I. Pelletier, H. C. Sauerwein, F. Lepore, D. Saint-Amour, and M. Lassonde, "Non-invasive alternatives to the Wada test in the presurgical evaluation of language and memory functions in epilepsy patients," *Epileptic Disorders*, vol. 9, pp. 111-126, Jun. 2007.

[24] M. C. Shirley, E. W. Tomás, and M. M. Charles, "Brain–computer interface using a simplified functional near-infrared spectroscopy system," *J Neural Eng*, vol. 4, pp. 219-226, Sep. 2007.

[25] E. Yin, T. Zeyl, R. Saab, D. Hu, Z. Zhou, and T. Chau, "An Auditory-Tactile Visual Saccade-Independent P300 Brain–Computer Interface," *Int J Neural Syst*, vol. 26, pp. 1650001-1650016, Feb. 2016.

[26] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A Hybrid Brain-Computer Interface Based on the Fusion of P300 and SSVEP Scores," *IEEE Trans. Neural Syst. Rehabil. Eng*, vol. 23, pp. 693-701, Jul. 2015.

[27] W. Qiu, C. Chang, W. Liu, P. W. Poon, Y. Hu, F. K. Lam, *et al.*, "Real-time data-reusing adaptive learning of a radial basis function network for tracking evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 226-37, Feb. 2006.

[28] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.

[29] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Stat.*, vol. 38, pp. 325-339, Apr. 1967.

[30] L.-D. Liao, I.-J. Wang, S.-F. Chen, J.-Y. Chang, and C.-T. Lin, "Design, fabrication and experimental validation of a novel dry-contact sensor for measuring electroencephalography signals without skin preparation," *Sensors*, vol. 11, pp. 5819-5834, May 2011.

[31] C.-T. Lin, L.-D. Liao, Y.-H. Liu, I. J. Wang, B.-S. Lin, and J.-Y. Chang, "Novel dry polymer foam electrodes for long-term EEG measurement," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 1200-1207, May 2011.

[32] Y.-H. Yu, S.-W. Lu, L.-D. Liao, and C.-T. Lin, "Design, fabrication, and experimental validation of novel flexible silicon-based dry sensors for electroencephalography signal measurements," *IEEE J. Transl. Eng. Health Med.*, vol. 2, pp. 1-7, Nov. 2014.

[33] C.-T. Lin, L.-W. Ko, I.-F. Chung, T.-Y. Huang, Y.-C. Chen, T.-P. Jung, *et al.*, "Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks," *IEEE Trans. Circuits Syst. Reg. Papers,* vol. 53, pp. 2469-2476, Nov. 2006.
[34] C.-T. Lin, I.-F. Chung, L.-W. Ko, Y.-C. Chen, S.-F. Liang, and J.-R. Duann, "EEG-based assessment of driver cognitive responses in a dynamic virtual-reality driving environment," *IEEE Trans. Biomed. Eng.,* vol. 54, pp. 1349-1352, July 2007.
[35] C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, and T.-P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Trans. Circuits Syst. Reg. Papers,* vol. 52, pp. 2726-2738, Dec. 2005.
[36] C.-T. Lin, C.-H. Chuang, Y.-K. Wang, S.-F. Tsai, T.-C. Chiu, and L.-W. Ko, "Neurocognitive characteristics of the driver: A review on drowsiness, distraction, navigation, and motion sickness," *J. Neurosci. Neuroeng.,* vol. 1, pp. 61-81, June 2012.
[37] C.-T. Lin, K.-L. Lin, L.-W. Ko, S.-F. Liang, B.-C. Kuo, and I.-F. Chung, "Nonparametric single-trial EEG feature extraction and classification of driver's cognitive responses," *EURASIP J. Adv. Signal Process.,* vol. 2008, pp. 1-11, Mar 2008.
[38] C.-T. Lin, K.-C. Huang, C.-F. Chao, J.-A. Chen, T.-W. Chiu, L.-W. Ko, *et al.*, "Tonic and phasic EEG and behavioral changes induced by arousing feedback," *NeuroImage,* vol. 52, pp. 633-642, Aug. 2010.
[39] S. W. Chuang, L. W. Ko, Y. P. Lin, R. S. Huang, T. P. Jung, and C. T. Lin, "Co-modulatory spectral changes in independent brain processes are correlated with task performance," *NeuroImage,* vol. 62, pp. 1469-77, Sep. 2012.
[40] C. T. Lin, C. H. Chuang, C. S. Huang, S. F. Tsai, S. W. Lu, Y. H. Chen, *et al.*, "Wireless and wearable EEG system for evaluating driver vigilance," *IEEE Trans. Biomed. Circuits Syst.,* vol. 8, pp. 165-76, Ap. 2014.
[41] Y.-K. Wang, S.-A. Chen, and C.-T. Lin, "An EEG-based brain–computer interface for dual task driving detection," *Neurocomputing,* vol. 129, pp. 85-93, Apr. 2014.
[42] C. T. Lin, S. A. Chen, T. T. Chiu, H. Z. Lin, and L. W. Ko, "Spatial and temporal EEG dynamics of dual-task driving performance," *J. Neuroeng. Rehabil.,* vol. 8, pp. 1-13, Feb. 2011.
[43] Y.-C. Chen, J.-R. Duann, S.-W. Chuang, C.-L. Lin, L.-W. Ko, T.-P. Jung, *et al.*, "Spatial and temporal EEG dynamics of motion sickness," *NeuroImage,* vol. 49, pp. 2862-2870, Feb. 2010.
[44] C. T. Lin, S. F. Tsai, and L. W. Ko, "EEG-based learning system for online motion sickness level estimation in a dynamic vehicle environment," *IEEE Trans. Neural. Netw. Learn. Syst.,* vol. 24, pp. 1689-700, Oct. 2013.
[45] T.-C. Chiu, K. Gramann, L.-W. Ko, J.-R. Duann, T.-P. Jung, and C.-T. Lin, "Alpha modulation in parietal and retrosplenial cortex correlates with navigation performance," *Psychophysiology,* vol. 49, pp. 43-55, Jan. 2012.
[46] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology,* vol. 37, pp. 163-78, Mar. 2000.
[47] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," *IEEE Trans. Neural Syst. Rehabil. Eng.,* vol. 16, pp. 432-41, Oct 2008.
[48] U. Orhan, K. E. Hild, 2nd, D. Erdogmus, B. Roark, B. Oken, and M. Fried-Oken, "RSVP keyboard: An EEG based typing interface," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.,* Kyoto, Mar. 2012, pp. 645-648.
[49] L. Waldorp, I. Christoffels, and V. van de Ven, "Effective connectivity of fMRI data using ancestral graph theory: Dealing with missing regions," *NeuroImage,* vol. 54, pp. 2695-705, Feb. 2011.
[50] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage,* vol. 19, pp. 1273-302, Aug. 2003.
[51] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica,* vol. 37, pp. 424-438, Aug. 1969.
[52] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Berlin: Springer, 2005.
[53] C. H. Chuang, L. W. Ko, T. P. Jung, and C. T. Lin, "Kinesthesia in a sustained-attention driving task," *NeuroImage,* vol. 91, pp. 187-202, May 2014.
[54] M. Massimini, F. Ferrarelli, R. Huber, S. K. Esser, H. Singh, and G. Tononi, "Breakdown of Cortical Effective Connectivity During Sleep," *Science,* vol. 309, pp. 2228-2232, Sep. 2005.
[55] S. K. Esser, S. Hill, and G. Tononi, "Breakdown of Effective Connectivity During Slow Wave Sleep: Investigating the Mechanism Underlying a Cortical Gate Using Large-Scale Modeling," *J. NeuroPhysiol.,* vol. 102, pp. 2096-2111, Oct. 2009.

Dr. Chin-Teng Lin received the B.S. degree from National Chiao-Tung University (NCTU), Taiwan in 1986, and the Master and Ph.D. degree in electrical engineering from Purdue University, USA in 1989 and 1992, respectively. He is currently the Chair Professor of Faculty of Engineering and Information Technology, University of Technology Sydney, Chair Professor of Electrical and Computer Engineering, NCTU, International Faculty of University of California at San-Diego (UCSD), and Honorary Professorship of University of Nottingham. Dr. Lin was elevated to be an IEEE Fellow for his contributions to biologically inspired information systems in 2005, and was elevated International Fuzzy Systems Association (IFSA) Fellow in 2012. He is elected as the Editor-in-chief of IEEE Transactions on Fuzzy Systems since 2011. He also served on the Board of Governors at IEEE Circuits and Systems (CAS) Society in 2005-2008, IEEE Systems, Man, Cybernetics (SMC) Society in 2003-2005, IEEE Computational Intelligence Society (CIS) in 2008-2010, and Chair of IEEE Taipei Section in 2009-2010. Dr. Lin is the Distinguished Lecturer of IEEE CAS Society from 2003 to 2005, and CIS Society from 2015-2017. He served as the Deputy Editor-in-Chief of IEEE Transactions on Circuits and Systems-II in 2006-2008. Dr. Lin was the Program Chair of IEEE International Conference on Systems, Man, and Cybernetics in 2005 and General Chair of 2011 IEEE International Conference on Fuzzy Systems. Dr. Lin is the coauthor of Neural Fuzzy Systems (Prentice-Hall), and the author of Neural Fuzzy Control Systems with Structure and Parameter Learning (World Scientific). He has published more than 200 journal papers (Total Citation: 20,155, H-index: 53, i10-index: 373) in the areas of neural networks, fuzzy systems, multimedia hardware/software, and cognitive neuro-engineering, including approximately 101 IEEE journal papers.

Yu-Ting Liu received the B.S. degree in Department of Electrical and Control Engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 2010. He is currently pursuing the Ph.D. degree with the Institute of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan.

His current research interests include fuzzy logic theory, machine learning, computer vision, EEG analysis and application.

Shang-Lin Wu was born in Kaohsiung, Taiwan, in 1987. He received the B.S. and M.S. degree in Department of Computer Science and Information Engineering from National Chiayi University, Chiayi, Taiwan, in 2009 and 2011, respectively. He is currently working toward the Ph.D. degree in Institute of Electrical Control Engineering at National Chiao Tung University (NCTU), Hsinchu, Taiwan.

His research interests are in the areas of fuzzy theory, transfer learning, biomedical signal processing, biofeedback control, and embedded system.

Zehong Cao received the B.E. degree in Electronic and Information Engineering from the Northeastern University (NEU), China, in 2012, and the Master degree in Electronic Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong, 2013. He is currently pursuing the dual Ph.D. degree from the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS), Australia, and Institute of Electrical Control Engineering, National Chiao Tung University (NCTU), Taiwan. His research interests are in the area of biomedical signal processing, brain-computer interface, and machine learning.

Yu-Kai Wang received the B.S. degree in mathematics education from National Taichung University of Education, Taichung, Taiwan, in 2006, and the M.S. degree in biomedical engineering from National Chiao Tung University, Hsinchu Taiwan, in 2009. I received Ph.D. degree in Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan, in 2015. Currently, I am a postdoctoral research. In the past years, I am working on dynamic cognitive science and neuro-engineering. To achieve better performance of EEG-based application, the current interests of my research is to integrate the computational intelligence (CI) with neuroscience knowledge for addressing cognitive functions, discovering the relationships between brain dynamics and cognitive variables, and diagnosing neurological diseases.

Chih-Sheng Huang received the B.S. degree from Fu Jen Catholic University, Taipei, Taiwan, in 2007; the M.S. degree from the National Taichung University, Taichung, Taiwan, in 2010; and the Ph.D. degree from the Institute of Electrical Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2015.

Currently, he is Principal Engineer of Research & Development division at Singular Wings Medical, Hsinchu, Taiwan. His current research interests include statistical learning, pattern recognition, machine learning, bio-signal processing, brain-computer interface.

Chun-Hsiang Chuang received the B.S. degree in mathematics education from Taipei Municipal Teachers College in 2004, the M.S. degree in educational measurement and statistics from National Taichung University in 2009, and the Ph.D. degree in Electrical Engineering from National Chiao Tung University (NCTU) in 2014. He was a visiting scholar at Swartz Center for Computational Neuroscience, University of California in 2012-2013. Currently, he is a postdoctoral researcher of Brain Research Center at NCTU, Taiwan.

His research interests include machine learning, computational neuroscience, biomedical signal processing, and brain-computer interface.