Visual and Auditory Brain-Computer Interfaces

Shangkai Gao*, Fellow, IEEE, Yijun Wang, Member, IEEE, Xiaorong Gao, Member, IEEE, Bo Hong, Member, IEEE

Abstract—Over the past several decades, electroencephalogram (EEG)-based brain-computer interfaces (BCIs) have attracted attention from researchers in the field of neuroscience, neural engineering, and clinical rehabilitation. While the performance of BCI systems has improved, they do not yet support widespread usage. Recently, visual and auditory BCI systems have become popular because of their high communication speeds, little user training, and low user variation. However, building robust and practical BCI systems from physiological and technical knowledge of neural modulation of visual and auditory brain responses remains a challenging problem. In this paper, we review the current state and future challenges of visual and auditory BCI systems. First, we describe a new taxonomy based on the multiple access methods used in telecommunication systems. Then, we discuss the challenges of translating current technology into real-life practices and outline potential avenues to address them. Specifically, this review aims to provide useful guidelines for exploring new paradigms and methodologies to improve the current visual and auditory BCI technology.

Index Terms—brain-computer interface (BCI), visual BCI, auditory BCI, multiple access
I. INTRODUCTION

Brain–computer interfaces (BCIs) establish a direct communication channel between a brain and a computer or external device. The primary aim of BCI research is to create a non-muscular communication channel so that people with severe motor disabilities can use it for communication and control. BCI has rapidly developed into a highly recognized field of biomedical engineering in the past few decades. Among different brain imaging techniques that have been applied to BCIs, electroencephalography (EEG) is the most commonly used method and the only type we studied in this paper.

While performance of EEG-based BCI systems is slowly increasing, the current levels of BCI performance do not yet support widespread usage. Accordingly, visual and auditory BCI systems (hereinafter referred to as v-BCI and a-BCI respectively) that exhibit high communication speed and classification accuracy have become more popular in recent BCI research. Specifically, the v-BCI and a-BCI systems covered in this review only include BCI-based on brain responses to exogenous visual or auditory stimuli (e.g., steady-state visual evoked potentials (SSVEP) and auditory steady-state responses (ASSR)) and endogenous potentials linked to the reaction to the visual or auditory stimuli (e.g., visual and auditory P300 event-related potentials (ERPs)). Exogenous and endogenous brain responses generally represent sensory-specific stimulus processing and non-sensory-specific mental processing respectively. Note that, this review does not include BCIs based on other EEG signals (e.g., BCIs based on motor imagery [1], and BCIs based on slow cortical potentials (SCP) [2]), although most of them use cues and feedbacks in visual or auditory modalities.

A. Historical review

The term “brain–computer interface” first appeared in 1970s. Vidal used the term to express the concept of putting observable electrical brain signals to work as carriers of information in man–computer communication or for the purpose of controlling external apparatus [3, 4]. In the following decades, several pioneers developed many of the classical v-BCI paradigms. In 1988, Farwell and Donchin reported a BCI paradigm based on P300 evoked potentials [5]. This 6×6 matrix visual speller demonstrated the promising prospect of real BCI applications. In the early 1990s, a number of new BCI paradigms were proposed. An efficient visual-evoked potential (VEP)-based BCI system was presented by Sutter in 1992 [6]. The 8×8 speller determined the user’s intents by recognizing the direction of eye gaze using VEP recorded from the visual cortex. This study reported the first clinical application of v-BCI which obtained a communication speed above 10 words per minute in an amyotrophic lateral sclerosis (ALS) patient. In addition to these well-known studies, there were others in this period that received less attention by the BCI community. For example, Principe’s group had proposed a novel system based on the cognitive response to congruent or incongruent words in a sentence in 1990 [7]. The studies in this period laid important groundwork for the field.

In the first decade of the new century, v-BCI and a-BCI research achieved rapid development. The numbers of research groups and scientific publications increased tremendously. Advanced signal processing and machine learning techniques have been applied to system implementation [8, 9]. Many new BCI paradigms, such as steady-state visual evoked potential (SSVEP) based BCI [10, 11], motion onset VEP based BCI [12], as well as auditory BCIs [13, 14], emerged and gradually matured. Meanwhile, the early developed paradigms (e.g. BCIs based on P300 and VEPs) were significantly improved and resulted in initial clinical trials. These systems were proved applicable to patients with ALS, stroke, and spinal cord injury [14, 15, 16, 17].

During the past several years, the clinical application of v-BCIs and a-BCIs has received increased attention [18, 19]. Seller et al. [18] tested the P300-based BCI with an ALS patient during long-term independent home use. Recently, the a-BCI system has been further tested in patients with disorders of consciousness [19]. With the rapid development of the v-BCI and a-BCI technology, researchers in the broader scientific and medical communities have become involved. The potential applications go far beyond the initial clinical settings. Recently, new subtypes of BCI (e.g. hybrid BCI [20], passive BCI [21], emotional BCI [22], and collaborative BCI [23]) have appeared in various publications. The v-BCI and a-BCI technology contributes a lot to these new BCI paradigms. There is no general consensus about whether the new types conform to the original BCI definition. However, the relaxed restrictions of the BCI definition has broadened its applications and will hopefully lead to further advances in the coming decades.
B. About this review

The v-BCI and a-BCI systems have become more popular in recent BCI research. However, building robust and practical systems from physiological knowledge of the modulation of neural responses to visual and auditory stimulus still poses a great challenge to researchers. In the BCI literature, the published review papers tend to focus on specific engineering aspects such as signal processing [24, 25], classification [26, 27, 28], general system design [29, 30, 31, 32, 33], or applications [34, 35]. A methodological review of v-BCI and a-BCI systems that describes their current stage as well as future challenges is missing. This topic is of significant importance for the following reasons:

1. The v-BCI and a-BCI can be categorized into gaze dependent and gaze independent systems. The gaze dependent v-BCI systems take advantage of high signal-to-noise ratio (SNR) in EEG recordings and high information transfer rates (ITR). The gaze independent v-BCI and a-BCI systems can provide relatively high system performance for locked-in patients who cannot use gaze dependent BCIs. Since high BCI performance relies on reliable, repeatable, and distinguishable brain signals, the v-BCI and a-BCI systems can provide robust system performance. Other advantages might include fewer electrodes, less user training, and lower user variation [36, 37]. All of these reasons make v-BCI and a-BCI systems a good candidate for real-life applications.

2. The current v-BCI and a-BCI systems lack a unified system framework, in part due to the fact that they have been studied separately since their conception. A summary of the general paradigms and methodologies developed in the v-BCI and a-BCI systems will facilitate future development and improvement. Their common properties such as the multiple target coding methods and the challenges in signal processing and classification have never been systematically reviewed or summarized.

The present review will focus on the current state and future challenges of the v-BCI and a-BCI systems. To put all varieties of v-BCI and a-BCI systems in a unified framework, we borrow the concept of signal modulation and multiple access (MA) methods [38] from the telecommunication systems. The remaining parts of this review are organized as follows. Section II introduces the general methods for brain signal modulation and the commonly used brain signals in current v-BCI and a-BCI systems. Section III describes the information stream followed by a taxonomy summary of v-BCI and a-BCI paradigms under a unified framework based on the multiple access methods. Section IV explores the challenges and strategies to cope with them. Finally, a brief summary is given in Section V.

II. BRAIN SIGNALS IN V-BCI AND A-BCI

A. Brain signal modulation

Brain signals could be modulated by exogenous stimuli or endogenous mental activities. As shown in Fig. 1, the exogenous stimuli in v-BCI and a-BCI systems are visual and auditory stimuli, while endogenous components could be induced by users’ covert attention or mental tasks. These brain responses can happen at sensation, perception, or cognition levels. Sensation is the processing of senses by the sensory system to external stimulus signals. Evoked potentials (EP) produced by visual and auditory stimuli reflect typical sensation processes. Perception deals with the organization, identification, and interpretation of sensory information. A sensory perception at conscious level allows the individuals to sense the environment around them. The process of cognition contains attention, learning, reasoning, decision making and so on. In a BCI system, the brain responses of the above three stages can be modulated by voluntary attention of the subject, thus the conveyed information can be encoded. The features in the modulated brain signals can be extracted in time, frequency, or space domains. The combination of features from different domains can substantially improve classification accuracy and thereby enhance the BCI performance.
the original messages

signals must be modulated

information

A.

attention has

signals, endogenous attention has been widely used in v

to words and other meaningful stimuli, typically

which have a parietal maximum,

evaluation, selective attention, and conscious discrimination

deviant stimulus

fronto

v

2)

which can be recorded from the scalp with maximum amplitude at the vertex

been

pseudo

VEP

Brain signals modulated by exogenous stimulu

which are brain’s response to visual stimuli, which can be recorded with maximum amplitude over the occipital region on the scalp [40]. The subtypes of VEPs in current v-BCIs include: (1) transient VEP (TVEP) under low-rate stimulus condition (<2Hz); (2) steady-state VEP (SSVEP) under high-rate stimulus condition (>6Hz); (3) motion VEP, which reflects visual motion processing; and (4) code-modulated VEP, which can be elicited by a pseudo-random stimulus sequence [6]. Auditory steady state response (ASSR) is an exogenous brain signal that has been used in current a-BCIs. ASSR is an auditory evoked potential (AEP) in response to rapid auditory stimuli, which can be recorded from the scalp with maximum amplitude at the vertex [41].

2) Endogenously modulated brain signals

- response to oddball stimulus (auditory Mismatch Negativity(MMN) [42], N200 and P300[43])
- response to mental tasks (Late positive components (LPC)) [44]
- response inhibition (No-Go N2) [45]
- semantic processing (N400) [46]
- attention modulated brain signals (SSVEP, ASSR) [47, 48]

Endogenous ERP signals play important roles in v-BCIs and a-BCIs. Major ERP components used in current v-BCIs and a-BCIs include MMN, N200, P300, LPC, No-Go N2, and N400. The auditory MMN, which is a fronto-central negative potential originating from the auditory cortex, peaks at 150-250 ms from the onset of the deviant stimulus [42]. N200 and P300 ERP components, maximal over the central and parietal areas, reflect stimulus evaluation, selective attention, and conscious discrimination in oddball tasks [43]. Late positive components (LPC), which have a parietal maximum, reflect cognitive response selection process in mental tasks [44]. No-Go N2, which is mainly distributed over the frontal-central area, reflects inhibitory response control [45]. N400 is a brain response to words and other meaningful stimuli, typically showing a centro-parietal scalp distribution [46]. In addition to ERP signals, endogenous attention has been widely used in v-BCIs and a-BCIs. Selective attention such as spatial attention has been found to significantly modulate the amplitude of SSVEP and ASSR [47, 48].

III. MULTIPLE TARGET CODING IN V-BCI AND A-BCI

A. Information stream in v-BCI and a-BCI

The technologies in the telecommunication system can inspire new train of thoughts in BCI designs. Essentially, information stream in a BCI is quite similar to a telecommunication system. To express different intents, brain signals must be modulated in a certain way so that the intent embedded EEG signals can be then demodulated into the original messages. Meanwhile, to avoid the mutual interference, the modulated brain signals for different intents
should be orthogonal or near orthogonal to each other. For this purpose, the modulated brain signals could be arranged by time/frequency/code/space divisions. This strategy is similar to the multiple access (MA) technology that allows multiple users to simultaneously share the bandwidth with least performance degradation in telecommunication systems [38]. There are four basic multiple access schemes: Time division multiple access (TDMA), Frequency division multiple access (FDMA), Code division multiple access (CDMA), and Space division multiple access (SDMA). In TDMA, the users are allotted different time slots during which they have the entire channel bandwidth. In FDMA, the entire bandwidth is divided into a number of partial frequency bands and distributed among users. In CDMA, the users are assigned separate codes to modulate their signals, which differentiate them from each other. SDMA divides the geographical space into smaller spaces and discriminates users based on their spatial locations. Details of these methods can be found in [38]. In fact, we can find analogies to all these methods in v-BCI and a-BCI systems [49]. The basic principles of the multiple target access methods used in BCIs are described in Table I. In most ERP-based BCIs, multiple targets appear at different time slots following the principle of TDMA. The SSVEP-based BCI is a typical FDMA system in which each target occupies its own frequency band without overlap. The BCI based on pseudorandom code modulated VEP works in a similar way to the CDMA method. The SDMA method has been applied to the designs of v-BCI, in which the EEG signals are modulated by different target locations in the visual field. In addition, the hybrid multiple access (HMA) method has recently been employed in the v-BCI studies to improve system performance [50, 51, 52, 53, 54].

B. Taxonomy of v-BCI and a-BCI

Current BCI systems could be classified by operation manner (such as dependent/independent BCIs, and synchronous/asynchronous BCIs) or the brain signals employed in the system (e.g., SSVEP and P300) [55]. Here, to highlight the nature of BCI as a novel communication system, we propose a new taxonomy to sort the existing v-BCI and a-BCI according to the multiple target access methods (see Table 1). In our previous study, the VEP-based BCI systems were categorized using this classification method [49]. This study further extends the taxonomy to classify all v-BCI and a-BCI systems in a comprehensive and systematic way. Similar to the classification of the telecommunication systems, the v-BCI and a-BCI systems can be sorted into the following five groups: (1) TDMA, (2) FDMA, (3) CDMA, (4) SDMA, and (5) HMA. In this way, v-BCI and a-BCI systems can be examined under a unified framework based on multiple target access methods. There are three primary advantages to this categorization. First, it simplifies the understanding of the design and implementation of v-BCI and a-BCI systems, making it easier for BCI researchers to incorporate existing technologies from traditional communication into these systems. For example, system design optimization and system performance evaluation methods in the telecommunication systems are easily transferable. Second, it facilitates the comparison between v-BCI and a-BCI as well as between systems using different EEG signals. For example, the auditory and visual P300 BCIs are grouped together into the TDMA category. In this way, methods and techniques applied separately in v-BCI and a-BCI systems can be better understood and then integrated to improve system performance. Third, it can help to transfer the existing methodologies and techniques in communication systems to improve system performance of the current v-BCI and a-BCI systems. For example, the signal modulation and demodulation methods in telecommunications can be adopted to develop new BCI paradigms with more robust system performance [56].

Table 1 lists stimulus, brain response, and representative publications of the v-BCI and a-BCI systems according to the proposed taxonomy. The following findings illustrate the characteristics of the multiple target access methods in the current v-BCI and a-BCI technology. First, it is clearly shown that TDMA and FDMA are the two most popular methods in system design. Specifically, TDMA has been widely used in the P300-based BCI systems, while FDMA has been applied to the SSVEP- and ASSR- based BCI systems. Second, CDMA has rarely been used. However, the highest ITR in current BCI systems was obtained by the code-modulated VEP-based BCI system [57]. In general, CDMA requires rapid brain responses to the stimulus sequence. In practice, brain signal analysis and stimulus design pose large challenges for any implementation of CDMA in a BCI system. Third, the HMA method was recently introduced into v-BCI studies. The representative studies show its potential for improving BCI performance. In addition to the common characteristics in the v-BCI and a-BCI systems, Table 1 also indicates the different properties of v-BCI and a-BCI systems. First, the number of studies on v-BCI is much larger than a-BCI. Also, v-BCI and a-BCI systems specialize in different aspects. The gaze dependent v-BCI systems can reach very high ITR specifically due to the advantage of high SNR of VEP signals. However, for the independent BCI systems
where the user cannot use muscle control such as eye gazing to operate the system, the a-BCI and v-BCI systems show comparable performance. Second, the multiple target access methods for v-BCI are more diverse than a-BCI. Currently, the CDMA, SDMA, and HMA methods are missing in the a-BCI systems. It is obvious that v-BCI has been more thoroughly investigated since it has been studied for much longer time than the a-BCI systems. It is generally more difficult to implement CDMA and SDMA in a-BCI systems. Compared with VEP signals, sequential coding of ERP signals in a-BCIs using a CDMA paradigm will be much slower due to a larger inter-stimulus interval (ISI). In a-BCIs that use stimuli from multiple locations, spatial modulations of EEG signals are generally weak and difficult to detect. Although some a-BCIs use spatial location to enhance the ERP signals [58, 59], this review attributes these systems to the TDMA category since no space-specific information was used for target identification. This finding might also suggest that there is significant room for improvement in a-BCI systems. For example, the HMA method that combines TDMA and FDMA might be useful to improve overall system performance.
<table>
<thead>
<tr>
<th>Frame structure</th>
<th>v-BCI</th>
<th>a-BCI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time division</strong> (TDMA)</td>
<td>Flicker</td>
<td>N1, P1, N2, P2 [3, 60]</td>
</tr>
<tr>
<td></td>
<td>Moving line</td>
<td>N2, P2 [12, 61, 62, 63]</td>
</tr>
<tr>
<td></td>
<td>Visual oddball</td>
<td>P300 [5, 16, 18, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76]</td>
</tr>
<tr>
<td></td>
<td>Meaningful words</td>
<td>N400 [7]</td>
</tr>
<tr>
<td></td>
<td>Go/noGo</td>
<td>N2 [77]</td>
</tr>
<tr>
<td><strong>Frequency division</strong> (FDMA)</td>
<td>Flicker / checkerboard SSVEP</td>
<td>[10, 11, 15, 89, 90, 91, 92, 93, 94, 95]</td>
</tr>
<tr>
<td></td>
<td>Flicker (phase) SSVEP</td>
<td>[96, 97, 98]</td>
</tr>
<tr>
<td></td>
<td>Multi-frequency stimulus SSVEP</td>
<td>[99, 100, 101, 102]</td>
</tr>
<tr>
<td></td>
<td>Flickering dots, grating, flicker SSVEP</td>
<td>[103, 104, 105]</td>
</tr>
<tr>
<td><strong>Code division</strong> (CDMA)</td>
<td>Pseudo-random code (m sequence) eVEP</td>
<td>[6, 57]</td>
</tr>
<tr>
<td></td>
<td>Code words SSVEP</td>
<td>[56]</td>
</tr>
<tr>
<td><strong>Space division</strong> (SDMA)</td>
<td>Checkerboard VEP</td>
<td>[4]</td>
</tr>
<tr>
<td></td>
<td>Flicker SSVEP</td>
<td>[52]</td>
</tr>
<tr>
<td><strong>Hybrid multiple access</strong> (HMA)</td>
<td>Time + frequency SSVEP, SSVEP + P300</td>
<td>[50, 51]</td>
</tr>
<tr>
<td></td>
<td>Space + frequency SSVEP</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td>Hybrid signals</td>
<td>N170+ P300 [53]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P300+ SSVEP [54]</td>
</tr>
</tbody>
</table>
IV. CHALLENGES

A. Addressing BCI-related electrophysiological issues

This review considered BCI systems as communication channels and categorized them according to the multiple target access method used in system design. This new taxonomy facilitates the comparison of v-BCI and a-BCI systems and also simplifies the understanding of the design and implementation of a BCI system from an engineering perspective. However, as the most complex biological system, the human brain is much more complicated than a telecommunication system. Its intrinsic properties such as non-linearity and non-stationarity pose big challenges when trying to implement a robust BCI system. These electrophysiological issues seriously affect BCI performance. To build a practical system, these issues must be taken into account when designing and implementing a system.

1) Non-linearity in EEG

The brain is a nonlinear system in which population dynamics of neural ensembles can be observed \[108\]. Its activities such as EEG signals can be better characterized by nonlinear dynamic methods than linear methods \[109, 110\]. The non-linearity in EEG signals has to be treated carefully in the v-BCI and a-BCI systems. In general, it could affect BCI performance in two opposite ways. First, it could lead to additional information for improving classification. For example, non-linearity exists in signal generation of SSVEPs in human visual cortex. The nonlinear resonance phenomena of SSVEPs can be distinctly characterized by brain responses at frequencies identical, harmonic, and subharmonic to the stimulus frequency \[111\]. Therefore, the harmonic components can provide useful information additional to the fundamental frequency component for detecting the stimulus frequency. The efficiency of combining multiple harmonic components has been well demonstrated in the SSVEP-based BCIs \[11, 37, 95\]. Second, some nonlinear properties in the brain could degrade the task-related EEG signals and thereby deteriorate the BCI performance. This situation commonly exists in ERP signals, which are highly sensitive to neurophysiological parameters \[112\]. For example, there is a limit to how fast the brain can process incoming stimuli. When two stimuli have to be processed within a short interval, the response to the second stimulus is significantly slowed (Psychological Refractory Period). Similarly, when two visual targets are presented in rapid succession, people will often fail to detect the second target (Attentional Blink). These effects show nonlinear modulation of the P3 component \[113, 114\]. To build a robust BCI system, avoiding these events when they occur must be included in system design.

2) Non-stationarity in EEG

The non-stationarity of brain activity in association with diverse mental and behavioral states occurs continuously over time \[115\]. It can be caused by the brain’s internal factors such as variabilities in neurophysiological states and psychological parameters, as well as external factors such as changes of electrode contact and electrode position, movement artifacts, and environmental noises. Similar to other BCIs, a major challenge in the v-BCI and a-BCI systems is the inter-session non-stationarity in the EEG data that often leads to deteriorated BCI performance. Specifically, inter-session non-stationarity in EEG classification can be attributed to the differences between a training session and an online session, and the changes across multiple online sessions \[116\]. To address this problem, adaptive classification methods that can automatically update the classifiers during online BCI operations have been developed \[117, 118\]. In addition, zero-training methods have attracted increasing attention in recent BCI studies \[76, 119, 120\]. The zero-training methods aim to solve the non-stationarity problem in feature extraction and classification by integrating information across multiple sessions or subjects. Another challenge within a smaller time scale is the inter-trial non-stationarity in EEG signals. The trial-to-trial variability can lead to variation of SNR in single-trial EEG signals \[121\]. Therefore, optimizing parameters for single-trial EEG signals plays an important role in EEG classification. For example, a typical problem in the ERP-based BCI is the number of trial repetitions required for target identification, which is crucial for reducing target detection time. In addition, the non-stationary problem may be alleviated by using advanced data analysis methods. For example, Stationary Subspace Analysis (SSA), which can decompose multivariate time series into stationary and non-stationary components, has been found applicable to BCI data \[122\].

B. Improving information transfer rate (ITR)

One of the major challenges in advancing v-BCI and a-BCI technology is the performance bottleneck, which is mostly attributed to the poor signal-to-noise ratio of EEG signals \[55\]. A variety of metrics have been proposed to
evaluate the performance of BCI systems [123]. In current systems, information transfer rate (ITR) is the most widely used metric. The ITR (in bits/minute) defined by Wolpaw et al. [55] is calculated as follows:

$$ITR = \left\{ \log_2 M + P \log_2 P + (1 - P) \log_2 (1 - P) \right\} \frac{60}{T}$$

where $M$ is the number of choices, $P$ is the accuracy of target detection, and $T$ (in seconds/selection) is the average time for a selection. More details of ITR estimation in online BCIs can be found in [124]. According to (1), the methods to improve ITR can be considered regarding to $M$, $P$, and $T$ separately. Although these three parameters always interact with each other in real systems, to facilitate summarizing the existing studies, this study considers the three factors separately.

1) Improving target detection accuracy

In general, the target detection accuracy can be improved in two different ways: (1) improving the SNR of task-related EEG signals, and (2) maximizing the separability of multiple classes. To achieve these goals, efforts have been made to increase the amplitude and dimension of features in the task-related EEG signals. Besides, advanced data analysis techniques such as signal processing and machine learning approaches have also been widely employed in current BCI systems [24, 26, 27, 125, 126].

a) Signal-to-noise ratio (SNR)

Improving the SNR of EEG signals is done by increasing the signal level and/or decreasing the noise level. First, SNR can be improved through applying advance signal processing methods. Trial averaging, commonly used to improve the SNR in ERP analysis, has been widely used in current v-BCI and a-BCI systems [55]. Recently, trial averaging across subjects has been applied in a collaborative BCI to improve the performance of an individual BCI [23, 77]. Spatial filtering can be used to project multi-channel EEG data into a low-dimensional spatial subspace to eliminate task-irrelevant components and improve the SNR of task-related EEG signals. For example, the Canonical Correlation Analysis (CCA) approach significantly improved the frequency detection of SSVEP [89, 97, 127]. CCA maximizes the correlation between the SSVEP signals and predefined reference signals [127]. Another widely used spatial filtering method is Independent Component Analysis (ICA) [64, 128]. ICA enhances the SNR of EEG signals by separating task-related EEG components from the task-irrelevant EEG components and the artifactual components [129, 130].

SNR can also be improved by eliciting enhanced task-related EEG signals. The amplitude of ERP signals always correlates to the user’s cognitive states associated with attention and emotion. Therefore, cognitive tasks can be employed in the stimulus design to generate more robust ERP signals. This concept has been proved highly efficient in recent studies. For example, compared with a cue without spatial properties, the combination of both pitch properties and spatial location of the stimulus in a discriminating cue significantly improved the system performance of a multi-class a-BCI [58]. In another a-BCI using active mental response, the subject’s voluntary recognition of the property of the target digits (e.g., left vs. right side; male vs. female voice) enhanced the ability to discriminate between brain responses (N2 and LPC components) to target and non-target digits [39, 78]. In a motion VEP-based BCI, the involvement of counting in target identification showed significantly improved amplitude of motion VEP signals compared to gazing [12]. In a recent study, Lakey et al. [72] manipulated attention in a P300-based BCI using a short mindfulness meditation induction (MMI) and found MMI subjects produced significantly larger P300 amplitudes than control subjects. Belitski et al. [87] employed simultaneous visual and auditory stimuli to enhance the P300 and thereby improved the performance of the classic visual P300 speller. In an affective SSVEP-based BCI, visual stimuli using emotional human faces significantly enhanced the amplitude of the SSVEP signals compared with the checkerboard stimuli [131].

b) Separability of multiple classes

Target detection accuracy depends on the separability of multiple classes. Machine learning techniques have been widely used to improve target detection accuracy in BCI systems [132]. The techniques used in current BCIs include diverse methods for feature selection, feature combination, and classification [26, 132]. In system design, separability of multiple classes can be improved by increasing the dimension of informative features in the task-related EEG signals. For example, in the SSVEP-based BCI using frequency coding, multiple-frequency coding has been used to build complex flickering stimuli to improve the separability of multiple classes [101, 102]. Recently, the method of hybrid EEG signals has been proposed, combining multiple EEG signals that carry independent information. For example, Yin et al. [51] integrated random flashes and flickers to simultaneously evoke the P300 and SSVEP signals. In another study, stimuli of facial images were used in the oddball paradigm to elicit
the face-sensitive ERP component N170, which was combined with P300 to enhance target detection [53]. In addition, another efficient way is to use complex coding methods such as the CDMA technology in system design. The BCI based on code-modulated VEP used orthogonal m-sequences to elicit VEP signals that could be easily discriminated by cross-correlation analysis [6, 57].

2) Increasing number of classes

The number of classes plays an important role in a BCI system. In general, the BCIs with high ITR have a large number of classes [5, 61, 133]. Compared to other BCIs, v-BCI and a-BCI systems are more capable of providing a large amount of classes to implement complex applications. The P300 BCI and the SSVEP BCI are two systems that can realize a large amount of classes [36, 37]. Obviously, multiple access methods facilitate the implementation of a large number of classes in these two types of BCIs. The P300 BCI systems typically use the TDMA method to code target stimuli. A row/column flashing approach has been commonly used to implement a stimulus matrix such as the well-known P300 speller using a 6×6 matrix [5]. Recently, other stimulus presentation methods have been developed to improve the row/column approach. For example, the method of flashing quasi-random groups of characters realized a 7×12 matrix speller [68]. Several stimulus coding methods including FDMA and CDMA have been adopted in the VEP-based BCI systems. For example, the BCI system using code-modulated VEP can realize paradigms with 32 or 64 classes [6, 57]. Currently, frequency coding is the most widely used method in the SSVEP BCI. Multiple-frequency coding methods have also been used to increase the number of classes [99, 100]. In addition, mixed coding approaches such as frequency-phase mixed coding [96] and frequency-space mixed coding [52] have been developed in recent studies.

3) Reducing target detection time

Generally, target detection time can be reduced by considering the following aspects. First, single-trial classification can be much more efficient than trial averaging. Currently, machine learning based single-trial analysis is widely used [126]. Second, adaptive methods can reduce target detection time. For example, the method of adaptive number of trial repetitions, which is called ‘dynamic stopping’ in [134], can significantly improve performance of the ERP-based BCIs [66, 73]. The ‘smart stopping’ method, in which the time to stop trial repetitions was determined based on the real-time monitoring of the SNR of the ERP signals, functioned in a similar way [62]. In the SSVEP-based BCIs, the data length of frequency detection was adjusted automatically to meet the target detection criterion in each selection [11, 15]. Third, optimized stimulus presentation can reduce target detection time. This method has been well studied in the P300 BCIs. One straightforward way is to reduce the duration of the ISI between two flashes in stimulus presentation [65]. In practice, a tradeoff between ISI and accuracy needs be made towards higher system performance. Another way is to optimize the stimulus coding method. For example, the traditional row/column coding method can be improved by coding quasi-random groups of characters [68]. In this way, the number of flashes per trial for identifying a target character can be significantly reduced.

C. Implementing real-life applications

Currently, moving a BCI system from the laboratory into real-life applications poses severe challenges for the BCI community [35, 135, 136]. Usability and user experience will play a key role in widening the application of the BCI technology. The following issues need to be addressed in a practical BCI system: (1) ease of use, (2) low-cost hardware and software, and (3) robust system performance [137]. Compared to other BCIs that don’t require external stimulation, the v-BCI and a-BCI systems pose more challenges in terms of system design and implementation. Tackling these two topics for practical systems, this review focuses on three major challenges: (1) the development of a mobile BCI platform, (2) methods to reduce fatigue, and (3) the design of asynchronous system control. Then it summarizes potential applications of the v-BCI and a-BCI systems.

1) Mobile system design

A mobile BCI platform technology can enable and facilitate numerous BCI applications in real-world environments. The implementation of a mobile BCI system should consider the following three major challenges. First, a mobile BCI requires mobile hardware solutions for EEG device, data processing platform, and stimulation device. Using bulky, expensive, wired hardware components will not only cause discomfort and inconvenience, but will also affect the ability of users to perform routine tasks. Recently, researchers have made rapid progress in the mobile EEG technology featuring miniature wireless EEG amplifier and dry electrode [138, 139, 140]. As a result, mobile BCIs have emerged rapidly [141]. For example, a cell-phone based mobile BCI system was demonstrated
with a phone-dialing program using SSVEPs [137]. A mobile P300-based a-BCI was demonstrated while subjects walked outdoors [86]. Second, the number of electrodes needs to be reduced to facilitate system use and reduce system cost. Different electrode selection methods have been proposed in BCI studies [37, 71, 142]. For example, in the SSVEP BCI, the selection of a bipolar electrode efficiently extracted SSVEPs with a high SNR [15]. Third, the system needs to be capable of solving the problem of artifacts in EEG signals since movement artifacts and ambient noises are much more severe in real-world environments. The emerging mobile brain imaging (MOBI) technology [143] could help solve this problem.

2) Fatigue reduction

Mental fatigue refers to a temporary inability to maintain optimal cognitive performance resulting from prolonged periods of cognitive activity. Mental fatigue can cause discomfort and decreased attention, and thereby degrades the amplitude of EEG signals [144]. Since visual or auditory stimulations are required in v-BCI and a-BCI systems, mental fatigue should be reduced as much as possible so that the system will remain practical for daily use. In general, this can be done by optimizing the physical properties of the stimulus. Currently, visual fatigue is one of the biggest disadvantages of v-BCI systems, significantly hindering their use in real-life applications. To solve this problem, researchers have made great efforts in optimizing the physical properties of the visual stimulus to reduce the discomfort. For example, different types of stimulus patterns such as high-frequency stimulus [93], high duty-cycle stimulus [94], and image-based stimulus [131] have been proposed for reducing visual fatigue while maintaining robust performance in the SSVEP-based BCIs. In another study, Hong et al. [61] investigated fatigue effect in two v-BCI systems using N200 (i.e., motion-onset VEP) and visual P300 respectively. The N200 was found insensitive to fatigue caused by trial repetitions, whereas the visual P300 showed significant amplitude decrease associated with visual fatigue. Recently, stimulus optimization has also been employed in the a-BCI systems. In one instance, because selective listening to natural stimuli is much more comfortable than artificial auditory stimuli such as pure tones, natural syllables were used to build an a-BCI [83].

3) Asynchronous system design

Most current v-BCIs and a-BCIs use synchronous control protocols where the period of control is initiated by the system. However, asynchronous control protocols, in which the user makes self-paced decisions on when to start or stop using the system [145], are more flexible and natural. An important issue in asynchronous control is detecting idle states. Several methods have been developed to solve this problem. First, detecting an idle state can be improved by adding additional EEG features into stimulus design. For example, in an SSVEP BCI, Cheng et al. [11] designed an ON-OFF button for activating/deactivating the visual stimuli so that the system could switch between the idle and control states. Similarly, a brain switch based on ERD/ERS or brain hemodynamic response was designed to turn on/off an SSVEP BCI within a hybrid BCI system [20]. In an SSVEP-based brain switch, the discrimination of idle and control states was improved by adding additional stimuli with different frequencies to areas around the target stimulus [146]. In an N200 BCI, the spatial information of the speller matrix was integrated to provide a more precise description of the motion VEP response patterns, which then could be used to detect the non-control state effectively [63]. Second, idle state detection could also be improved by developing effective computational approaches for distinguishing between EEG signals in idle and control states. For example, in a P300 BCI, Zhang et al. [74] proposed a computational approach to model target P300, non-target P300, and non-control EEG signals and then derived a recursive algorithm to detect control states based on likelihood.

4) Clinical applications

Due to the advantages such as high ITR and little user training, the v-BCI and a-BCI systems have been applied to many clinical applications to help patients with motor disabilities to communicate with their environments [147]. Most v-BCI systems depend on the muscle control of eye to gaze at the target during system use. For patients who are able to move their eyes (e.g., patients with spinal cord injury), these gaze-dependent systems provide an alternative solution to conventional assistive devices such as eye-tracking systems. Although current gaze-dependent BCIs show lower communication speeds than the eye-tracking systems, they have some distinct properties that make them attractive to users. For example, the SSVEP-based BCI, which is capable to have a large number of classes, can be totally calibration free [37]. As the performance of v-BCIs continues to improve, the gaze-dependent BCIs could provide high communication speed comparable to the eye-tracking technologies in the near future. For totally locked-in patients, only the independent BCI systems can satisfy their needs. The typical independent v-BCI and a-BCI systems include v-BCI systems using selective visual attention and a-BCI systems using selective listening.
Currently, gaze-independent v-BCIs and a-BCIs provide comparable BCI performance in terms of number of targets and ITR [33]. Due to a limited number of classes, most gaze-independent BCIs used a two-level selection procedure in complex applications such as spelling. This procedure introduces an additional workload and therefore limits the communication speed.

Although the v-BCIs and a-BCIs have mainly been developed towards clinical applications, very few studies have been carried out in patients [6, 14, 15, 16, 17, 18, 19, 71]. Currently, there are several reasons that limit the applicability of the v-BCIs and a-BCIs in clinical applications. First, conventional assistive technologies such as eye-tracking systems can provide more efficient control than gaze-dependent BCIs. Second, gaze-independent BCIs based on SCP and motor imagery provide alternative BCI solutions to locked-in patients. Third, totally locked-in patients typically have difficulties in learning how to use the BCI system [147]. Joint efforts between researchers and clinicians are required to promote the development of v-BCIs and a-BCIs more applicable for clinical uses.

5) Other applications

The v-BCI and a-BCI systems also have potential in many non-clinical applications [148]. Recently, the concept of using BCI to improve human performance has been demonstrated by several studies. For example, the P300 BCI using a rapid serial visual presentation (RSVP) paradigm was used to improve human performance in target detection [70]. Other non-clinical applications include mental state monitoring [136] and video gaming [149]. By solving the challenges discussed above, the v-BCI and a-BCI technology could benefit a much larger population whether they are patients with disabilities or not.

V. SUMMARY

In essence, a BCI is a system that aims to read the activity of the human brain, commonly thought to be the most complex biological system in the world. Although knowledge of the human brain has gradually increased, we still know very little about how it works. The lack of knowledge of the underlying neural mechanisms continues to be a challenge when building and studying BCI technology. By conducting an in-depth analysis of brain signal modulation, multiple access methods, and the practical challenges of v-BCI and a-BCI systems, this review aims to provide useful guidelines for exploring the new paradigms and methodologies that are being used to improve current technology.

The original purpose of the BCI technology was to provide a tool to help the patients with motor disabilities to communicate with their environments [55]. From a technical point of view, a real-time platform for brain-computer interaction is a more general definition of BCI. Under this definition, BCI includes all technologies that use online brain signal analysis to influence human interactions with computers, their environments, and even other humans [30]. Compared with the term ‘interface’, ‘interaction’ puts more emphasis on the process of mutual action and influence between the brain and the computer. This interactive platform can read user’s intentions and emotions in real time and thereby improve the traditional human-computer interaction (HCI) technology. By monitoring the user’s cognitive state in real time, it is possible to make prompt and effective interventions to prevent declines in cognitive and behavioral performance. In addition, by actively exercising the brain, a BCI intervention can also facilitate clinical practice in rehabilitation. Furthermore, BCI can even be applied to improving behavioral performance for healthy people. Indeed, emerging applications such as passive BCI [21], emotional BCI [22], and BCI-based neuro-rehabilitation [150] have shown great potential in the last few years.

Finally, it has to be noted that there is still a long way to go before the BCI technology can be effective, reliable, and affordable enough to benefit a large population in daily life. Future scientific and technical breakthroughs, which require collaborative efforts among multidisciplinary teams of experts in neuroscience, engineering, and clinical rehabilitation, will be the key to achieving the goal.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Alexander Maye and Prof. Fusheng Yang for their critical comments on the manuscript.


S. G. Mason and G. E. Birch, “A brain
A. S. Boksem, T. F. Meijman, and M. M. Lorist, “Effects of mental fatigue on attention: An ERP study,”
C. Grozea, C. D. Voinescu, and S. Fazli, “Bristle
Z. Lin, C. Zhang, W. Wu, and X. Gao, “Frequency recognition based on canonical correlation analysis for SSVEP
Neuroimage
P. Yuan, X. Gao, B. Allison, Y. Wang, G. Bin, and S. Gao, “A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces,”
Z. Lin, C. Zhang, W. Wu, and X. Gao, “Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs,”
H. Bakardjian, T. Tanaka, and A. Cichocki, “Emotional faces boost up steady-state visual responses for brain-computer interface,”
G. Dornhege, B. Blankertz, G. Curio, and K. -R. Müller, “Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms,”
Y. Wang, X. Gao, B. Hong, and S. Gao, “Practical designs of brain–computer interfaces based on the modulation of EEG rhythms,”
“The Berlin Brain-Computer Interface: Non-Medical Uses of BCI Technology,”
C. Grozea, C. D. Voinescu, and S. Fazli, “Bristle-sensors—low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications,”
Int. J. Psychophysiol., vol. 73, no. 2, pp. 95–100, 2009.
A. S. Boksem, T. F. Meijman, and M. M. Lorist, “Effects of mental fatigue on attention: An ERP study,”
G. G. Mason and G. E. Birch, “A brain-controlled switch for asynchronous control applications,”
H. Pan, Y. Q. Li, R. Zhang, Z. H. Gu, and F. Li, “Discrimination Between Control and Idle States in Asynchronous SSVEP-Based Brain

Shangkai Gao (SM’94, F’07) graduated from the Department of Electrical Engineering of Tsinghua University, Beijing, China, in 1970, and received the M.E. degree of biomedical engineering in 1982 in same department of Tsinghua University. She is now a professor of the department of Biomedical Engineering in Tsinghua University. Her research interests include neural engineering and medical imaging, especially the study of brain-computer interface. She is also a fellow of American Institute for Medical and Biological Engineering (AIMBE). She is now the Editorial Board Member of IEEE Transactions on Biomedical Engineering, Journal of Neural Engineering and Physiological Measurement, as well as the senior editor of IEEE Transactions on Neural System and Rehabilitation Engineering.

Yijun Wang (M’11) received his B.E. and Ph.D. degrees in biomedical engineering from Tsinghua University, Beijing, China, in 2001 and 2007, respectively. He is currently an assistant project scientist at the Swartz Center for Computational Neuroscience, University of California San Diego. His research interests include brain-computer interface, biomedical signal processing, and machine learning.

Xiaorong Gao (M’04) received the B.S. degree in biomedical engineering from Zhejiang University in 1986, the M.S. degree in biomedical engineering from Peking Union Medical College in 1989, and the Ph.D. degree in biomedical engineering from Tsinghua University in 1992. He is currently a professor of the Department of Biomedical Engineering, Tsinghua University. His current research interests are biomedical signal processing and medical instrumentation, especially the study of brain-computer interface.

Bo Hong (M’04) received his B.S. and Ph.D. degree of biomedical engineering from Tsinghua University, in 1996 and 2001, respectively. From 2004 to 2005, he was a visiting faculty in the department of biomedical engineering and the center for neural engineering at Johns Hopkins University, USA. Since 2005, he has been associate professor with Department of Biomedical Engineering, School of Medicine, Tsinghua University. His research interests are brain computer interface and neural information decoding.