Brain-Computer Interface in Multimedia Communication

Touradj Ebrahimi, Jean-Marc Vesin,
and Gary Garcia

CIRCUIT ©1997 JOHN FOXX IMAGES, HEAD: ©1995 PHOTODISC, INC.

Human-computer interface (HCI) has been a growing field of research and development in recent years [1]-[4]. Most of the effort has been dedicated to the design of user-friendly and ergonomic systems by means of innovative interfaces such as voice, vision, and other input/output devices in virtual reality [5]-[15]. Direct brain-computer interface (BCI) adds a new dimension to HCI [16]-[23]. Interesting research in this direction has already been initiated, motivated by the hope of creating new communication channels for persons with severe motor disabilities. In this article, we approach the problem of BCI from the viewpoint of interactions in a multimedia-rich environment for the general consumer market. However, this is by no means incompatible with applications for motor impaired subjects.

There is a general consensus that BCI represents a new frontier in science and technology. One of the challenging aspects is the need for multidisciplinary skills to achieve this goal. The growing field of BCI is in its infancy, and a significant amount of research is still needed to answer many questions and to resolve many complex problems.

This article raises various issues in the design of an efficient BCI system in multimedia applications. The main focus will be on one specific modality, namely electroencephalography (EEG)-based BCI. In do-

ing so, we provide an overview of the most recent progress achieved in this field, with an emphasis on signal processing aspects.

Human-Machine Interface in Multimedia Communication

Communication—A Definition

One of the main characteristics of humans, as opposed to animals, resides in their extended ability to communicate. The central role of communication in society, culture, and economy increases ceaselessly. Within this context, a first question is "What defines communication?" Several answers can be envisioned. Here, we use a simple one which refers to communication as a process to express and share experiences between humans. Such an experience can be either real or imaginary. Communication took place from the early days of cave men when they told stories around the fire and drew sketchy figures on the walls of their caves. Modern communication is not different, except in its sophistication and efficiency. An evolution in communication would consist of improving the quality of the expression and that of the shared experience without inherently modifying its underlying nature. Examples of such evolution are the moves from black-and-white television to color, to high definition, to stereo, to three-dimensional (3-D) holographic images. Following this logic, a revolution in communication would refer to a change at the fundamental level, either due to new modalities or to the addition of new dimensions. A trivial example is that of written book and press, which allowed for communication to take place in an asynchronous way. Anyone can read and enjoy the writings of his favorite author at any time. The same is true for photography, which can grasp a scene and immortalize a moment.

Based on the above, a second question is "What will be the next revolution in communication?" This difficult and often posed question is not easy to answer. As in the past the fate of communication will be shaped by many trends. Probably the two most significant will be those in media technologies on one hand and the interface between humans and machines on the other. Below we briefly describe each trend but will focus on the second trend.

New Media

An essential element in any communication is the content. Content can be seen as the recipient in which we express ourselves. In the history of mankind, we have been successful in finding multiple ways of expression. The obvious audio (music, song) and visual (painting, photography, film) expressions are by no means the only channels available to us. Mechanisms of expression stimulating all our senses, such as haptics (sculptures, clothing), olfaction (perfumes), and taste (cuisine) have always had a strong impact in culture and art. Electronic and digital communication has not been very successful

Examples of Revolutions and Evolutions in Communications	
Story telling and cave drawing	Books and written press
Photography	Telegraph
Telephone	Radio and music recording
Cinema	Television and video recording
Internet	Mobile communication

in taking full advantage of content beyond audiovisual modalities as vectors of expression due to limitations in technology and science for acquisition, representation, manipulation, and generation of such content.

In the meantime, efforts have been made to further the limits of expression as far as audiovisual content is concerned. Progress in virtual, augmented, and mixed reality is a good indication that science and technology can contribute in defining new paradigms of expression [64]-[66]. Advances in electronics, software, and communication technologies not only allow for the more efficient production, distribution, and consumption of traditional media but also will extend what is defined as media today. New media include modalities such as 3-D video, photo-realistic 3-D, and animated 3-D models and even scenes built from synthetic and natural objects coming from multiple sources. In the near future, digital media will eventually include totally new sensory information such as those mentioned above, aiming at stimulating all our senses and bringing us new possibilities of expression and experiences.

The Weak Link in Communication

Somewhere in the tumbling evolutions of communication, tools found their place in between communicating humans. Pen and paper are simple examples of such tools. These tools were further improved to become machines, of which computers are an important representative. Because of the existence of machines and computers in the chain of communication between humans, a new problem was identified: how to improve the communication between machines. This was a central focus of science and technology in the 20th century and continues to be so, leading to various disciplines and technologies which are still evolving. The last two decades have witnessed tremendous progress in communications. This progress has succeeded to make communication between machines the strong link in the information chain, somewhat undermining, if not ignoring, the importance of the links between humans and machines. In fact, despite the above-mentioned progress in connecting machines, we still use rather primitive interfaces to communicate with them. The several-fold increase in computer performance and communication channel bandwidth does not seem to

properly reflect in human-machine interfaces, and the good old keyboards, pointing devices, and displays still assure the majority of our interactions with computers. It seems as if Moore's Law has not applied with the same rigor to interfaces between humans and computers. Recently, other more natural communication modalities such as speech recognition and synthesis have paved the way out of this situation. Research and development in these areas indicate that there is hope that human-machine interaction can extend to other modalities such as interaction through smart cameras (vision), haptics (touch), olfaction (smell), and others. Most current efforts aim at adapting the interface to our natural senses or to replace them. Visual sensors mimic the human eye in machines to allow them to see or at least to interpret visual cues emanating from a user. Haptic devices allow machines to feel and measure pressure for machines to react in a more natural way to commands and actions. Likewise, they can produce haptic or even olfactory feedback to provide users with a richer experience, exciting senses beyond those of conventional audiovisual information. Is there a way to extend the man-machine interface beyond this? Are there ways to altogether bypass the natural interfaces of a human user such as his muscles and yet establish a meaningful communication between man and machine? Many in the research and development community assert that this is within possibilities of today and near future state of the art.

All proposed solutions seem invariantly to point to the source of our senses and emotions: the human brain.

Brain Activity—A New Communication Modality

Recent progress in technology allows us to probe and monitor physiological processes inside our body, for which no natural interfaces exist. In particular we can measure our blood pressure, heart rate variability, muscular activity, and brain electrical activity in efficient and noninvasive ways. It is natural to assume that such activity can be used as information in new communication channels. In this article, we focus on brain electrical activity and review methods and procedures aiming at detecting and interpreting such signals for the purpose of command and control in a multimedia environment. A variety of noninvasive methods are now available to monitor brain functions. These include EEG, magentoencephalography (MEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) [23]. PET, fMRI, and MEG are expensive and complex to operate and therefore not practical in most applications. At present, only EEG, which is easily recorded and processed with inexpensive equipment, appears to offer the practical possibility of a noninvasive communication channel. Furthermore, EEG signals are rather well studied, and there is evidence that subjects can control them to some extent in a voluntary manner.

An Overview of BCIs
Background

Although the EEG is an imperfect, distorted indicator of brain activity, it remains nonetheless its direct consequence. Also, it is based on a much simpler technology and is characterized by much smaller time constants when compared to other noninvasive approaches such as MEG, PET, and fMRI. When it became possible to process digitized EEG signals on a computer, the temptation was great to use EEG as a direct communication channel from the brain to the real world. Early research programs in that direction were sponsored by defense agencies. One often cites the project headed by Dr. J. Vidal, director of the Brain-Computer Interface Laboratory at UCLA, as the first successful endeavor at building a BCI [24]. Some years later, works such as [25] confirmed that it was indeed possible to determine reasonably well what mental task (out of a specified small set of tasks) was performed by a subject on the basis of his EEG. In this reference, mental task discrimination used a feature compounding asymmetry between right and left hemispheres in all frequency bands and on all electrodes.

The last ten years have witnessed an explosion in the area of BCI research. A cornerstone was the first workshop on BCI technology that took place in 1999, in Albany, New York, where 22 research groups presented their work [21]. Among them were the Graz group [18], the Neil Squire foundation [19], the Wadsworth Center [17], the Tuebingen group [22], and the Beckman Institute [26]. A formal definition of the term BCI has been set forward [23]: "A brain-computer interface is a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles." The framework is now clearly defined, and the critical mass is reached. An excellent review of all the efforts to make BCI technology a mature field of research can be found in [23].

The main motivation today is to develop replacement communication and control means for severely disabled people, especially those who have lost all voluntary muscle control (locked in). As expressed by [26], the potentially huge human impact of such developments does not result in high levels of funding or commercial interest. BCIs are classified as dependent or independent. Dependent BCIs rely upon voluntary ocular activity to generate a specific EEG pattern, i.e., a visual evoked potential (VEP) provoked by redirection of subject's gaze [27], [28] and as such do not fully satisfy the above definition. On the contrary, independent BCIs do not imply recourse to muscular intervention of any kind and constitute by far the principal topic of investigation.

General Structure of an Independent BCI

To create a communication channel from the brain to a computer, a BCI must first acquire signals generated by cortical activity. Most existing BCIs are based on an EEG.

Figure 1 shows an example of a typical EEG acquisition device in BCI applications. Some preprocessing is generally performed due to the high levels of noise and interference usually present. Then features related to specific EEG components must be extracted. Finally, some mapping from the feature vectors to device commands is exerted. In a vast majority of existing BCIs these commands take the form of a selection of a specific letter, in a letter set, or cursor displacement on a computer screen (Figure 2).

Clearly, to obtain an effective communication, a feature-to-command translation must be established. In almost all BCIs this is done through one or more training sessions composed of trials. During each trial the subject is asked to perform some mental task (typically from a set of two to five mental tasks), and the EEG features corresponding to this task are extracted. After enough trials, a classification algorithm can take care of the feature-to-command translation. In test sessions, the BCI system initiates a succession of trials. During each trial the subject performs the mental task corresponding to the commands he wants executed. A feedback is usually provided, i.e., the extracted feature vector is translated into a command that is executed (for instance a cursor moves to the left). This synchronous mode of operation may be unfit in control applications such as mechanical device control. Reference [29] presents an interesting approach for an asynchronous BCI, in which an additional EEG feature set is used by the subject to initiate a new trial. Note also what is called in [30] the man-machine learning dilemma. Feedback constitutes effectively an important aspect of human learning. And in a BCI, the subject uses feedback information to improve the production of an EEG activity such that fewer errors occur in feature-to-command translation. However, repeated failures during a session may induce subject's frustration and modification of his mental state. Additionally, the feedback itself, which is often visual, can cause a perturbation in EEG activity.

Signal Processing Aspects

Movement-induced perturbations in the EEG can be avoided by a careful monitoring of the experiments in terms of subject discomfort, so electrooculogram (EOG) and electromyographic (EMG) artifacts constitute the major source of trouble in EEG processing. In some works such as [31] and [32], the contaminated EEG trials are detected and discarded. In [33] a nonlinear filtering approach is proposed for EMG artifact cancellation. The EOG can be estimated by placing a pair of electrodes near one eye. This additional information is first used for direct subtraction from the EEG [34]. A more elaborate technique based on adaptive interference cancellation is proposed in [35]. It seems, however, that global techniques [36], that is, using simulta-

neously the information from all electrodes, represent the most promising approach. In [37] principal component analysis (PCA) is used to decompose a multielectrode EEG trial in linearly uncorrelated components, and then reconstruction is performed by omitting unwanted components such as EOG. A now familiar improvement consists of using independent component analysis (ICA) [38], [39] instead of PCA [40]-[42]. A quantitative evaluation [43] confirms the superiority of ICA with respect to other techniques. Another interest of ICA lies in the fact that it can also be used for event related potentials (ERPs) extraction [44].

Many existing BCIs use features related to EEG events known to be representative of specific brain activities. Those events may be characterized in the time domain, such as the slow cortical potentials (SCPs) or P300 evoked potentials. SCPs are slow voltage changes occurring on time scales from 0.5 to 10 seconds and associated with movement and cortical activation [45], [46]. SCPs have been shown to be controllable by a subject and have been used in the so-called thought translation device [22]. The evoked potential "P300," a positive peak, takes place about 300 ms after an infrequent or significant stimulus imbedded in a
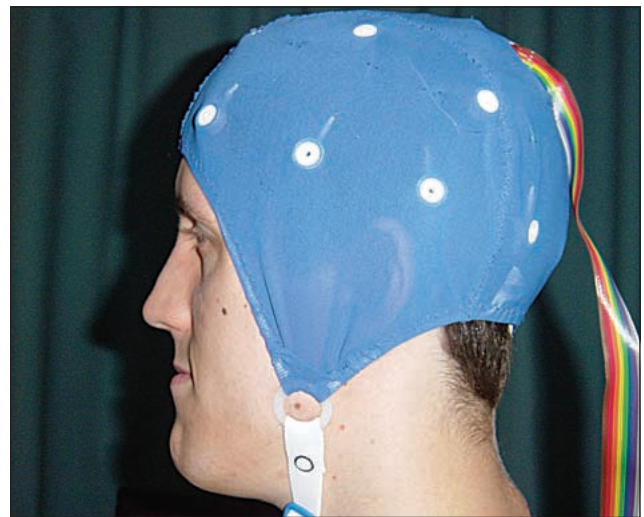
▲ 1. Example of an EEG electrode cap.
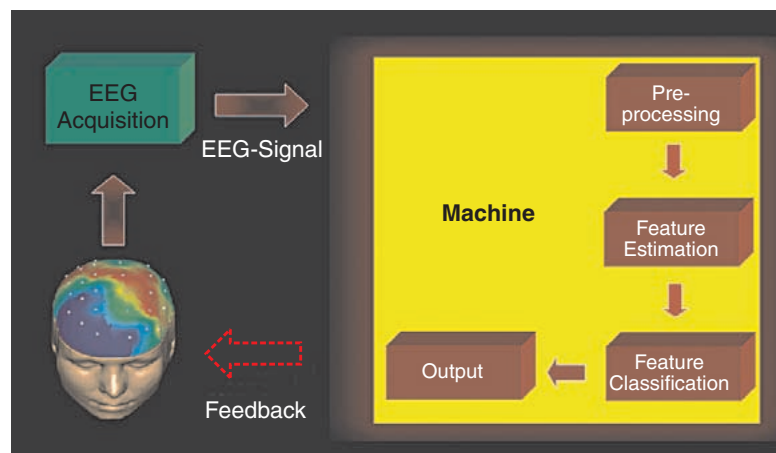
▲ 2. General block diagram of EEG-based BCI.

The Electroencephalogram

An electroencephalogram (EEG) is a recording of the very weak (in the order of 5-100 μV) electrical potentials generated by the brain on the scalp. EEG has been the subject of much fantasy in some popular books and movies. Its origin seemed to imply it was a direct expression of the cerebral processes, but unfortunately (by some aspects one could say fortunately) reality is quite different. The layers of cerebrospinal fluid, bone, and skin that separate the electrodes from the brain itself induce much attenuation and volume conduction. The former causes a poor signal to interference ratio and the latter an ample smoothing of the electrical activity of the cortex. Also, many physiological factors influence the EEG.

The pioneering work of Berger in the 1930s led him to discover the existence of structured patterns in the EEG, still under study today. Later, analysis of EEG activity, performed visually on paper records, was aimed at clinical purposes such as detection of epilepsy and other severe pathologies. The huge advances in acquisition devices and computers in the last 20 years have triggered a host of research on the EEG. If no mind-reading machine is likely to appear, some of the links between EEG activity and sensorimotor and mental processes are now at least partially understood.

The EEG is recorded as a potential difference between a signal electrode placed on the scalp and a reference electrode (generally one ear or both ears electrically connected). A conductive paste is used to decrease contact impedance and electrode migration. Due to signal levels, high-gain, high-quality amplifiers are placed between the electrodes and the acquisition devices. The International Federation in Electroencephalography and Clinical Neurophysiology defined a standard called the 10-20 electrode placement system (see Figure 5). The electrode names are standardized too and correspond to their location on the scalp. For instance, the two occipital electrodes are called O1 and O2. The average frequency content of the EEG imposes a sampling rate in the range 100-200 Hz, typically 128 Hz.

It is common practice to consider specific frequency bands thought to be associated with specific brain rhythms: the alpha (8-13 Hz), beta (13-30 Hz), delta (0.5-4 Hz), and theta (4-7 Hz) bands. Alpha waves, of moderate amplitude, are typical of relaxed wakefulness (idling). Lower-amplitude beta waves are more prominent during intense mental activity. Theta and delta waves are normal during drowsiness and early slow-wave sleep. Recently a distinction has been made between idling activity focused over visual cortex, still called alpha rhythm, and idling activity focused over motor cortex, now called mu rhythm. The latter seems to be associated to beta rhythm.

Another modality of EEG that is largely used in studies of human cognitive processing is the evoked potential (EP), or, to use a more neutral denomination, the event-related potential (ERP). EP is the voltage change in EEG activity recorded (typically during a one-second trial) after a stimulus has been presented to the subject. This stimulus is in most cases an auditory or a visual one. In the latter case one often uses the term visual evoked potential (VEP).

The main causes for EEG perturbation, particularly at frontal, temporal, and occipital electrodes, are the electrooculogram (EOG) and the electromyogram (EMG). These perturbations are most often called artifacts in the relevant literature. EOG is due to fluctuating electrical fields generated by eye movement, with the eyeball constituting an electrical dipole. EMG is due to the electrical activity of scalp muscles.

Further useful references on EEG and ERP are [61] and [68]-[70].

---

series of routine stimuli. As such, it can be used to detect the subject's choice by proposing in turn the possible options. This approach has been implemented in [26]. But much emphasis in the BCI literature has been put on frequency events, namely mu and beta rhythms (see "The Electroencephalogram"). Those rhythms are linked to motor output cortical activity but also to preparation of movement, which makes them suitable for BCI design. In several research works such as [18], [17], [31], [32], [47], and [48], the mental tasks consist of imagined movements, typically right- and left-hand imagined movement in a binary choice BCI. Spatial filtering such as surface Laplacian [31] or the method of common spatial patterns (CSPs) [32], [48] has also been employed to enhance those frequency features.

An interesting approach to frequency component selection, based on an extension of Kohonen's LVQ [49], is described in [50]. Finally, some BCIs use features not directly related to specific brain activities, such as those proposed in [51] that are based on autoregressive parameters.

The choice of the classification method for the feature-to-command translation, while important, is probably less crucial for a successful BCI operation. A decision tree is used in [31], a local neural network in [53], a Bayesian network in [54], and in [55] a time-dependent multiplayer perceptron (MLP) is proposed as an improvement upon conventional MLP [55]. What matters, however, is the adaptability of the classifier, which permits adjustment of the BCI to short- and

long-term variations in a subject's EEG activity. Typically adaptation should take place during test sessions.

Example of a Working BCI System
Architecture and Operational Modes

This section presents the overall architecture of a BCI system which was built to carry out research and development in BCIs for multimedia applications [71]. The system has been designed to be very modular and flexible, so as to exploit it in a large number of BCI applications. In a typical application, the resulting EEG pattern of a specific mental activity (MA) is first learned by the computer in an initial training session and automatically recognized in ulterior sessions.

The training process is mutual as the human subject and the computer learn how to produce and how to recognize a given EEG pattern corresponding to an MA.

Real-time interaction between the subject and the computer is therefore an essential part of the system. For reasons of efficiency, the BCI system has been designed around five operational modes (OMs) going from simple to more sophisticated.

Visualization OM

In this OM, the user can see a visual representation of her/his EEG activity in real time. Particular features associated with EEG signals, such as the power values in the typical frequency bands (delta, theta, alpha, and beta), inter-electrode coherences, and total power at a given electrode, are mapped to a 3-D virtual environment and regularly updated. The objectives of this OM are calibration and familiarization of the subject with the system.

Training Without Feedback OM

In this OM, audio or visual cues are presented to the subject for him to perform predefined mental activities. This OM allows the computer to learn those EEG patterns associated with particular mental activities. The learning process is carried out offline. Reference models are built for each predefined MA.

Training with Feedback OM

In this OM, the subject is asked to perform an MA and feedback is provided. The feedback is positive when the computer recognizes the MA and negative otherwise. This recognition is based on the reference models of the mental activities calculated during the training without feedback OM. This OM allows the subject to modulate his electrical brain activity to minimize the misclassification rate. As a matter of fact, several series of training
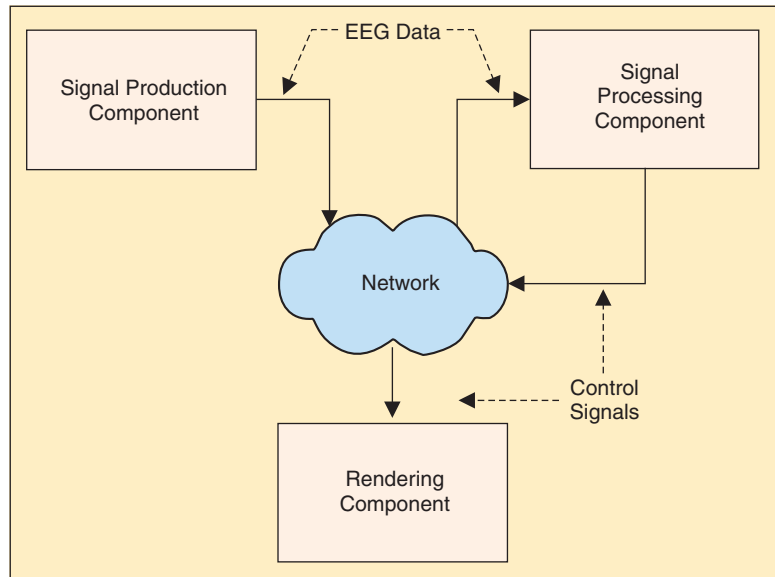
▲ 3. BCI system architecture.

with and without feedback are necessary to achieve good recognition rates.

Control OM

As the result of the previous experiments is a model of MA, the subject can start to control the system by producing the mental activities for which the system has been trained. Thus, visual or sound cues are no longer necessary.

Multiuser OM

This OM exists in a multiuser game where the goal is to gain control of an object by performing an MA. This kind of training paradigm was chosen because of its more stimulating effect when compared to a simple feedback.

The above experiments require a carefully designed BCI system in terms of flexibility, speed, and EEG real-time processing. A solution fulfilling these requirements is a distributed system in which each component provides specific services to the others in an efficient and transparent way. Figure 3 depicts the general block diagram of this BCI system.

Three component types can be distinguished:

▲ a signal production component whose responsibility is to digitize the acquired EEG signals and transmit them to the processing units

▲ a signal processing component that is in charge of the signal preprocessing, feature extraction, model building, and classification

▲ a rendering component that is used to display visual or auditory cues and for rendering possible feedback.

The communication rules between these components were designed according to the CORBA specifications and the algorithms were implemented in Java, C, and MATLAB.

Signal Preprocessing

As said before, EEG signals can be contaminated by ocular and muscular artifacts. Ocular artifacts are particularly undesirable because of their large amplitude. Therefore, we implemented an online elimination of those EEG segments that are contaminated by ocular artifacts.

The EEG signal is segmented into half-second segments and the variations in the signal power at electrodes Fp1 and Fp2 (Figure 4) are evaluated. Whenever the power values at both electrodes change abruptly the corresponding EEG segment is discarded (Figure 5).

Feature Extraction and Classification

The BCI system receives commands sent by the user in the form of EEG patterns produced during MAs. Typically the BCI system analyzes EEG segments (EEG trials) and extracts features from them that are compared with MA models computed during the training phase. The features extracted from EEG trials, the models computation, and the comparison methods are interdependent and constitute one of the most important design aspects in a BCI system.

The single EEG trial-classification problem can be stated as follows. Given a labeled set of EEG trials (train-

▲ 4. International 10-20 system placement for an EEG acquisition.

▲ 5. Example of preprocessing of EEG signals.

ing set) that were recorded during the performance of known MAs, we wish to characterize each MA by a model so that we can compare an unknown EEG trial to each MA model and assign it to its nearest MA model.

The features that are extracted from EEG trials determine the nature of the MA models and the comparison method. Among the possible choices, we focused on two approaches based on a multivariate autoregressive analysis of EEG trials and on a joint time-frequency-spatial analysis of the correlations between the univariate components of EEG trials.

Multivariate Autoregressive Analysis

The multivariate autoregressive (MVAR) representation of a discrete time N-multivariate signal $S(k)$ consists of a weighted linear combination of past observations plus a random, uncorrelated input

$$S(k) = \sum_{l=1}^{p} M(l)S(k-l) + e(k). \tag{1}$$

In (1), $S(k)$ and $e(k)$ are vectors of dimension N and the $M(l), l = 1, ..., p$ are matrices of dimension $N \times N$ whose coefficients are called MVAR coefficients. The MVAR coefficients are estimated with a least squares procedure.

The MVAR coefficients were chosen because they reflect the auto and cross-spectral relationships between the components of a multivariate signal $S(k)$.

In the case of EEG trials, the set of MVAR coefficients calculated for each trial constitutes its feature vector. We used a single layer neural network (NN) with two possible output values (0 or 1) for each MA pair. The output of the NN corresponding to an MA pair indicates to which MA the input EEG trial is closest. With this method, the classification of an unknown EEG trial S is performed in the following way. The MVAR coefficients of S are estimated and grouped into a feature vector which is fed into different NNs. A majority rule is then used on outputs to classify the EEG trial.

Joint Time-Frequency-Space Correlation (TFSC) Analysis

In this approach, the EEG trials are analyzed with respect to their correlative time-frequency representation (CTFR). The CTFRs provide a measure of the interaction strength between groups of neurons as a function of the time and frequency. Previous studies emphasized the importance of these parameters to classify brain activity [57]. The CFTR is commonly characterized by the ambiguity function [58]. In the following we limit ourselves to the continuous time and frequency representations for sake of clarity. A discrete domain formulation can be found in [72].
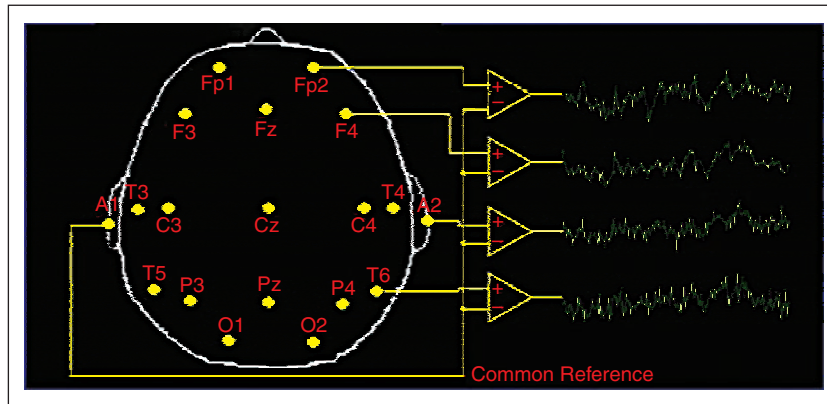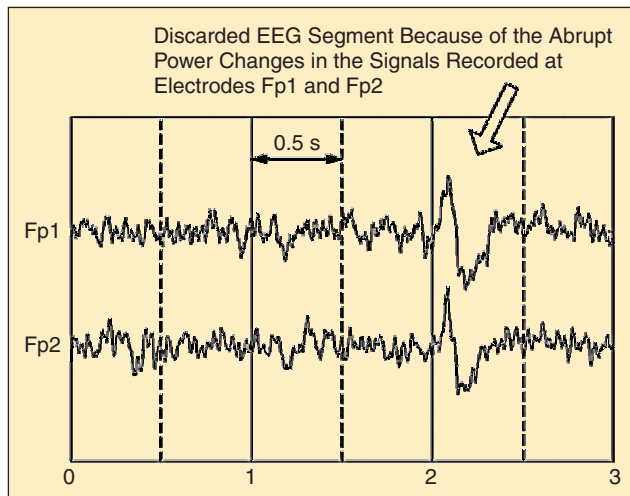
The ambiguity function of a multivariate signal $S(t) = [s_1(t) \ s_2(t) \ \ldots \ s_N(t)]^t$ is defined in (2). The univariate signals $s_\ell(t); \ell = 1, \ldots, N$ are the spatial components of $S(t)$

$$A_S(\theta, \tau) = \int S\left(t + \frac{\tau}{2}\right) S^H\left(t - \frac{\tau}{2}\right) {}^{j\theta t} dt \tag{2}$$

where ^{H} stands for the conjugate transpose and θ, τ are the frequency and time lags, respectively. Equation (2) can be written as

$$A_S(\theta, \tau) = [A_{mn}(\theta, \tau)] \quad 1 \leq m, \ n \leq N$$
$$A_{mn}(\theta, \tau) = \int s_m\left(t + \frac{\tau}{2}\right) s_n{}^*\left(t - \frac{\tau}{2}\right) {}^{j\theta t} dt. \tag{3}$$

The diagonal terms of $A_S(\theta, \tau)$ are called the auto ambiguity functions and the off-diagonal terms the cross-ambiguity functions.

In [59] Cohen defined the characteristic function in the univariate case as the product between the ambiguity function and a two-dimensional function called the kernel. In an analog way we can define the multivariate characteristic function as

$$M_S(\theta, \tau) = [\phi_{mn}(\theta, \tau) A_{mn}(\theta, \tau)] \quad 1 \leq m, \ n \leq N. \tag{4}$$

The two-dimensional functions $\phi_{mn}(\theta, \tau)$ are the kernels. Results in the univariate case [60] demonstrated that it is possible to design a kernel that is optimized for the classification in the $\theta - \tau$ plane. We generalized these results to the multivariate case by considering N^2 radially Gaussian kernels parameterized by their shape [60].

The characteristic function of a multivariate signal provides information on its joint time, frequency, and space correlations. The off-diagonal terms of the matrix $M_S(\theta, \tau)$ (4) represent the spatial correlations. If we transform the multivariate signal to obtain spatially decorrelated components, we then reduce the number of kernels to optimize from N^2 to N [36]. The decorrelating transformation is designed to obtain transformed components (directions) that maximally separate the classes [67].

When working with transformed components only N kernels have to be determined. They are separately optimized using an iterative procedure that updates their shape so as to enhance those regions in the $\theta - \tau$ plane where class separation is maximal [36]. The resulting characteristic functions are linearly separable.

Results and Discussions

Three types of mental activities, mental counting (MA1) and imagined left and right index movements (MA2 and MA3, respectively), were selected. Three healthy male subjects (S1, S2, and S3) participated in the experiments.

Five 20-minute sessions were carried out. In the first training session no feedback was provided. In the next sessions the model references were updated at the end of each session and used in the next session to provide feedback.

The protocol used in a session where feedback was provided is shown in Figure 6. During the first five minutes the visualization OM was applied to calibrate the system and familiarize the user with the system. The rest of the session was divided into three five-minute slices. These five-minute slices were divided into rest and recording periods. During a recording period a feedback was provided each half-second. The half-second segment immediately after the visual cue was discarded because of the influence of the visual evoked potentials.

EEG signals were recorded from electrodes Fp1, Fp2, C3, C4, P3, P4, O1, and O2 of the 10-20 international system [61]. Among these signals the last six were used for the
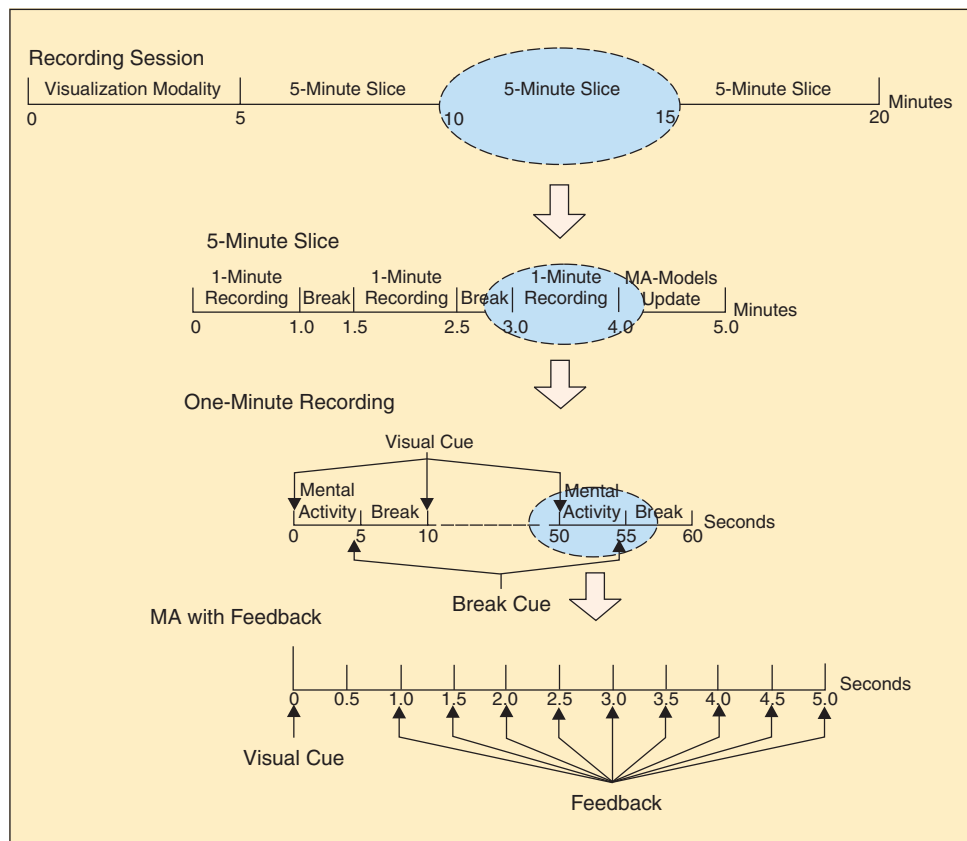
▲ 6. Protocol of a training session where feedback was provided.

classification, and Fp1, Fp2 for artifact detection and removal [36].

After the first training session where no feedback was provided, the weights of the NNs and the optimal value [62] for the model order p for the MVAR approach and the kernel parameters for the time-frequency-space correlation (TFSC) analysis were determined. From sessions two to five a continuous feedback was provided to the subject.

The results in terms of error rate are shown in Figures 7 and 8. As can be observed from these figures, the classification error rate decreased over the sessions for both approaches. This indicates that the feedback provided to the user and the update of the MA models at the end of each session positively influenced the overall performance of both user and BCI. The same trend can be observed in other studies where the feedback strategy was evaluated.

The second approach was better in terms of error rate. This last point indicates that the analysis of the coherent neuronal activity in time, frequency, and space appears to be more appropriate for the recognition of mental activities.
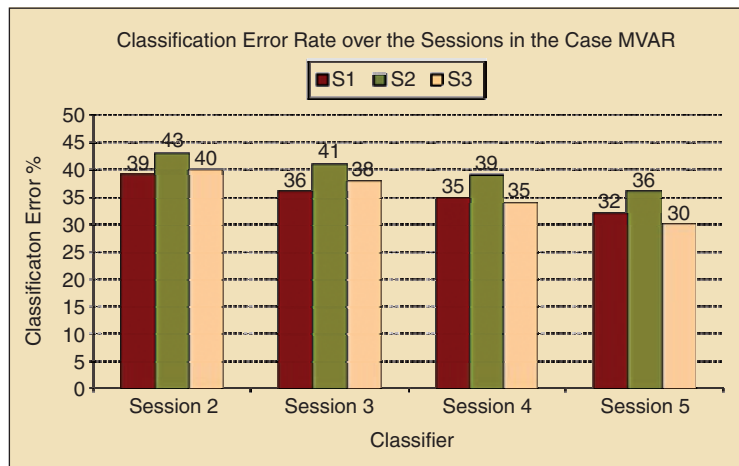
▲ 7. Evolution of the classification error rate over sessions two to five with MVAR features.
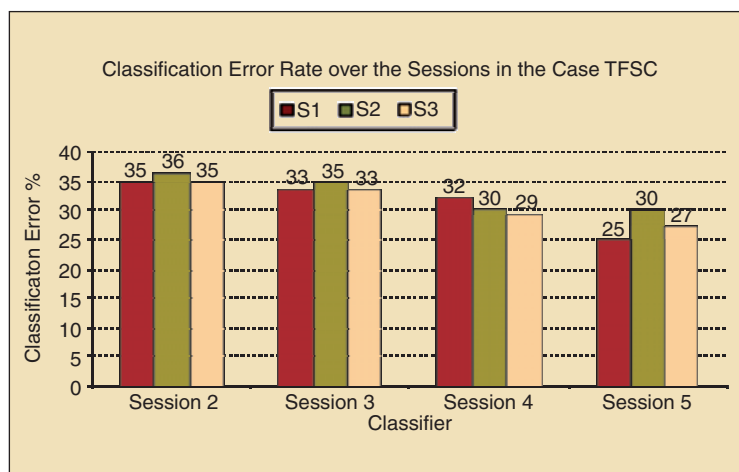
▲ 8. Evolution of the classification error rate over sessions two to six with TFSC features.

Perspectives and Challenges

BCI technology, by its intimate connection with the wonder of human thought processes, is a fascinating research field. The existing BCI systems have demonstrated that direct information transfer from the brain to a device is indeed possible, but many lines of investigation are wide open for multidisciplinary teams. One of the first problems to address is the limitation of the information transfer rate, which is at best currently 20 bits/min. It seems dubious that BCI protocols based on mental task classification can improve this figure by much. A simple reason for this is the practical impossibility for a subject to switch between mental tasks, even simple ones such as imagined movements, at time intervals smaller than two or three seconds. Finer spatial and temporal scales of analysis as well as determination of more stable and more reliable features could cope with more complex mental tasks. Other cortical activities such as emotions could be employed to initiate information transfer. In the framework of performance improvement, let us mention the very sensible advocacy in [23] for the design of reliable tools for performance assessment, in order to assure scientific credibility. Information transfer rate [23], [63] is a first step in this direction.

Another very challenging aspect in BCI technology is the interplay between the BCI system and the subject, in terms of adaptation and learning. Ideally, the BCI system should adapt to changes in subject EEG activity and the subject should also learn to better control his EEG activity, but these conjugate tasks should be as harmonious as possible. Finally let us note that the limitations of EEG-based BCI systems seem less important in a multimodal HCI. The EEG activity is, by its very nature, "orthogonal" to more classical modes. Inclusion of a BCI in a multimodal interface thus results in a net gain of information transfer capability.

Touradj Ebrahimi received M.Sc. and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology, Lausanne (EPFL), in 1989 and 1992, respectively. In 1993, he was a research engineer at the Corporate Research Laboratories of Sony Corporation, Tokyo. In 1994, he was as a researcher at AT&T Bell Laboratories. He is currently a titular professor at the Signal Processing Institute of the School of Engineering at EPFL, where he is involved in research and teaching for multimedia information processing and coding. In 2002, he founded Emitall, a research and development company in electronic media innovations. He was the recipient of the IEEE and Swiss national ASE award in 1989 and the winner of the first prize for the best paper appearing in *IEEE Trans-*

actions on Consumer Electronics in 2001. In 2001 and 2002, he received two ISO awards for contributions to MPEG-4 and JPEG 2000 standards. He is author or co-author of over 100 scientific publications and holds a dozen patents.

Jean-Marc Vesin graduated from the Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble (ENSIEG, Grenoble, France) in 1980. He received his M.Sc. from Laval University, Québec City, Canada, in 1984, where he spent four years on research projects. After two years in the industry, he joined the Signal Processing Institute (ITS) of the Swiss Federal Institute of Technology, Lausanne, Switzerland, where he obtained his Ph.D. in 1992. His primary interests are nonlinear signal modeling and analysis, adaptive systems, and evolutionary algorithms. From an application viewpoint, he works mainly in biomedical signal processing and in computer modelling of biological tissue electrical activity.

Gary Garcia received his M.Sc. degree in electrical engineering from the Swiss Federal Institute of Technology (EPFL) in 2001. Since April 2001 he has been a Ph.D. student at EPFL where he is involved in several projects in image, video and biomedical signal processing. Currently he is developing an adaptive direct brain-computer communication device with strong emphasis on the models and interpretations of synchronized neural activity.

References

[1] V.I. Pavlovic,. R. Sharma, and T. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 677-695, July 1997.

[2] Z. Duric, W.D. Gray, R. Heishman, L. Faylin, A. Rosenfeld, M.J. Schoelles, C. Schunn, and H. Wechsler, "Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction," *Proc. IEEE*, vol. 90, pp. 1272-1289, July 2002.

[3] W. Ying and T.S. Huang, "Nonstationary color tracking for vision-based human-computer interaction," *IEEE Trans. Neural Networks*, vol. 13, pp. 948-960, July 1997.

[4] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 1998.

[5] J. Grey, "Human-computer interaction in life drawing, A fine artist's perspective," in *Proc. IEEE 6th Int. Conf. Inf. Visualisation*, London, 2002, pp. 761-770.

[6] Y. Daabaj, "An evaluation of the usability of human-computer interaction methods in support of the development of interactive systems," in *Proc. IEEE Int. Conf. System Sciences*, Hawaii, 2001, pp. 1830-1839.

[7] L.S. Chen and T.S. Huang, "Emotion expressions in audiovisual human computer interaction," in *Proc. IEEE Int. Conf. Multimedia Expo.*, New York, 2000, vol. 1, pp. 423-426.

[8] D.W. Massaro, "Perceptual interfaces in human computer interaction," in *Proc. IEEE Int. Conf. Multimedia Expo.*, New York, 2000, vol. 1, pp. 563-566.

[9] H. Hongo, M. Ohya, M. Yasumoto, and K. Yamamoto, "Face and hand gesture recognition for human-computer interaction," in *Proc. IEEE 15th Int. Conf. Pattern Recognition*, 2000, vol. 2, pp. 921-924.

[10] J.A. Landay, "Informal user interface for natural human-computer interaction," *IEEE Intell. Syst.*, vol. 13, pp. 14-16, May-June 1998.

[11] J. Segen and S. Kumar, "Human-computer interaction using gesture recognition and 3D hand tracking," in *Proc. IEEE Int. Conf. Image Processing*, Chicago, IL, 1998, vol. 3, pp. 188-192.

[12] K.-H. Englmeier, C. Krapichler, M. Haubner, M. Seemann, and M. Reiser, "Virtual reality and multimedia human-computer interaction in medicine," in *Proc. IEEE 2nd Workshop Multimedia Signal Processing*, Redondo Beach, CA, 1998, pp. 193-202.

[13] R. Stiefelhagen and J. Yang, "Gaze tracking for multimodal human-computer interaction," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, 1997, vol. 4, pp. 2617-2620.

[14] C.M. Jones and S.S. Dlay, "Human-computer interaction and animation system for simple interfacing to virtual environments," in *Proc. IEEE Int. Conf. Compt. Cybernetics Sim.*, Cambridge, U.K., 1997, vol. 5, pp. 4242-4247.

[15] N. Kayian, "Exploratory study of implicit theories in human computer interaction," in *Proc. IEEE 6th Australian Conf. Computer-Human Interaction*, New Zealand, 1996, pp. 338-339.

[16] C. Guger, A. Schlögl, C. Neuper, D. Walterspacher, T. Strein, and G. Pfurtscheller, "Rapid prototyping of an EEG-based brain-computer interface (BCI)," *IEEE Trans. Rehab. Eng.*, vol. 9, pp. 49-58, Mar. 2001.

[17] J.R. Wolpaw, D.J. McFarland, and T.M. Vaughan, "Brain-computer interface research at the Wadsworth Center," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 222-226, June 2000.

[18] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in Graz brain-computer interface (BCI) research," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 216-219, June 2000.

[19] G.E. Birch and S.G. Mason, "Brain-computer interface research at the Neil Squire Foundation," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 193-195, June 2000.

[20] J.D. Bayliss and D.H. Ballard, "A virtual reality testbed for brain-computer interface research," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 188-190, June 2000.

[21] J.R. Wolpaw, N. Birbaumer, W.J. Heetderks, D.J. McFarland, P. Peckman, G. Schalk, E. Donchin, L.A. Quatrano, C.J. Robinson, and T.M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 164-173, June 2000.

[22] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (TTD) for completely paralyzed patients," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 190-193, June 2000.

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

[24] J.J. Vidal, "Real-time detection of brain events in EEG," *Proc. IEEE*, vol. 65, pp. 633-664, May 1977.

[25] Z.A. Keirn and J.I. Aunon, "A new mode of communication between man and his surroundings," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 1209-1214, Dec. 1990.

[26] E. Donchin, K.M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 174-179, June 2000.

[27] E.E. Sutter, "The brain response interface: Communication through visually-induced electrical brain responses," *J. Microcomput. Appl.*, vol. 15, pp. 31-45, 1992.

[28] M. Middendorf, G. McMillan, G. Calhoun, and K.S. Jones, "Brain-computer interfaces base on steady-state visual-evoked response," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 211-214, June 2000.

[29] S.G. Mason and G.E. Birch, "A brain-controlled switch for asynchronous control applications," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1297-1307, Oct. 2000.

[30] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proc. IEEE*, vol. 89, pp. 1123-1134, July 2001.

[31] J. del R. Millan, M. Franzé, J. Mouriño, F. Cincotti, and F. Babiloni, "Relevant EEG features for the classification of spontaneous motor-related tasks," *Biol. Cybern.*, vol. 86, pp. 89-95, 2002.

[32] H. Ramoser, J. Mueller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 441-446, Dec. 2000.

[33] L.P. Panych, J.A. Wada, and M.P. Beddoes, "Practical digital filters for reducing EMG artifacts in EEG seizure recordings," *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, pp. 268-272, 1989.

[34] P.M. Quilter, B.B. McGillivray, and D.G. Wadbrook, "The removal of eye movement artefact from the EEG signals using correlation techniques," *Proc. Inst. Elec. Eng. Conf. Publ.*, vol. 159, pp. 93-100, 1977.

[35] K.D. Rao and D.C. Reddy, "On-line method for enhancement of electro-encephalogram signals in presence of electro-oculogram artifacts using non-linear recursive least squares technique," *Med. Biol. Eng. Comput.*, vol. 33, pp. 488-491, 1995.

[36] G. Garcia and T. Ebrahimi, "Time-frequency-space kernel for single EEG-trial classification," presented at NORSIG Conf., Norway, 4-7 Oct. 2002.

[37] T.D. Lagerlund, F.W. Sharbrough, and N.E. Busacker, "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," *J. Clin. Neurophysiol.*, vol. 14, pp. 73-82, 1997.

[38] A. Hyvärinen and E Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, pp. 411-430, 2000.

[39] T.W. Lee, *Independent Component Analysis. Theory and Applications*. Boston, MA: Kluwer, 1998.

[40] T.-P. Jung, S. Makeig, C. Humphries, T.W. Lee, M.J. McKeown, V. Iragui, and T.J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163-178, 2000.

[41] R. Vigario, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 589-593, 2000.

[42] S. Vorobyov and A. Cichoki, "Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis," *Biol. Cybern.*, vol. 86, pp. 293-303, 2002.

[43] L. Vigon, M.R. Saatchi, J.E.W. Mayhew, and R. Fernandes, "Quantitative evaluation of techniques for ocular filtering of EEG waveforms," *Proc. Inst. Elec. Eng.-Sci. Meas. Technol.*, vol. 147, pp. 219-228, Sept. 2000.

[44] T.-P. Jung, S. Makeig, M. 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.

[45] N. Birbaumer, "Slow cortical potentials: Their origin, meaning, and clinical use," in *Brain and Behavior Past, Present, and Future,* G.J.M. van Boxtel and K.B.E. Böcker, Eds. Tilburg, The Netherlands: Tilburg Univ. Press, 1997.

[46] A. Kübler, B. Kotchoubey, J. Kaiser, J.R. Wolpaw, and N. Birbaumer, "Brain-computer communication: Unlock the lock-in," *Psych. Bull.*, vol. 127, pp. 358-375, 2001.

[47] J.R. Wolpaw and D.J. McFarland, "Multichannel EEG-based brain computer communication," *Electroencephalogr. Clin. Neurophysiol.*, vol. 90, pp. 444-449, 1994.

[48] C. Guger, H. Ramoser, and G. Pfurtscheller, "Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI)," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 447-456, Dec. 2000.

[49] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, pp. 1464-1480, 1990.

[50] M. Pregenzer and G. Pfurtscheller, "Frequency component selection for an EEG-based brain to computer interface," *IEEE Trans. Rehab. Eng.*, vol. 7, pp. 413-417, Dec. 1999.

[51] C.W. Anderson, E.A. Stolz, and S. Shamsunder, "Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 277-286, Mar. 1998.

[52] J. del R. Millan, J. Mouriño, M. Franzé, F. Cincotti, M. Varsta, J. Heikkonen, and F. Babiloni, "A local neural classifier for the recognition of EEG patterns associated to mental tasks," *IEEE Trans. Neural Networks*, vol. 13, pp. 678-686, 2002.

[53] W.D. Penny, S.J. Roberts, E.A. Curran, and M.J. Stokes, "EEG-based communication: A pattern recognition approach," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 214-215, June 2000.

[54] E. Haselsteiner and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 457-463, Dec. 2000.

[55] B.D. Ripley, *Pattern Recognition and Neural Networks.* Cambridge, U.K.: Cambridge Univ. Press, 1996.

[56] N. Saiwaki, N. Kimura, and S. Nishida, "An analysis of EEGS based on information flow with SD method," in *Proc. 1998 IEEE Int. Conf. Syst., Man, Cybernetics*, San Diego, CA, Oct. 1998, pp. 4115-4119.

[57] P.L. Nunez, R. Srinivasan, A.F. Westdorp, R.S. Wijesinghe, D.M. Tucker, R.B. Silberstein, and P.J. Cadusch, "EEG coherency I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, pp. 499-515, 1997.

[58] M. Coates, "Time-frequency modelling," Ph.D. dissertation, Dept. Engineering, Univ. of Cambridge, Cambridge, U.K., Aug. 1998.

[59] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[60] M. Davy, "Noyaux optimises pour la classification dans le plan temps-frequence—Proposition d'un algorithme constructif et d'une reference bayesienne basee sur les methodes MCMC—Application au diagnostic d'enceintes acoustiques," Ph.D. dissertation, Institut de Recherche en Communications et Cybernetique de Nantes, Université de Nantes, Nantes, France, Sept. 2000.

[61] H.H. Jasper, "The ten-twenty electrode system of the international federation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, pp. 371-375, 1958.

[62] A. Neumaier and T. Schneider, "Estimation of parameters and eigenmodes of multivariate autoregressive models," *ACM Trans. Math. Softw.*, vol. 27, no. 1, pp. 27-57, Mar 2001.

[63] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller, "Information transfer rate in a five-classes brain-computer interface," *IEEE Trans. Rehab. Eng.*, vol. 9, pp. 283-288, Sept. 2001.

[64] H. Tamura, H. Yamamoto, and A. Katayama, "Mixed reality: Future dreams seen at the border between real and virtual worlds," *IEEE Comput. Graph. Appl.*, vol. 21, pp. 64-70, Nov.-Dec 2001.

[65] E. Badique, "New imaging frontiers: 3D and mixed reality," in *Proc. IEEE 3D Data Proc. Vis. Trans. Conf.*, 2002, pp. 296-304.

[66] H. Yamamoto, "Case studies of producing mixed reality worlds," *IEEE Trans. Syst. Man Cybernet.*, vol. 6, pp. 42-47, 1999.

[67] C.W. Therrien, *Decision Estimation and Classification*. New York: Wiley, 1989.

[68] J.D. Bronzino, "Principles of electroencephalography," in *The Biomedical Engineering Handbook*, J.D. Bronzino, Ed. Boca Raton, FL: CRC, 1995.

[69] K.A. Kooi, R.P. Tucker, and R.E. Marshall, *Fundamentals of Electroencephalography,* 2nd ed. Hagerstown, MD: Harper & Row, 1978.

[70] M.D. Rugg and M.G.H. Coles, Eds., *Electrophysiology of Mind-Event-Related Brain Potentials and Cognition*. Oxford, U.K.:Oxford Univ. Press, 1995.

[71] G. Garcia, T. Ebrahimi, and J.-M. Vesin, "Classification of EEG signals in the ambiguity domain for brain-computer interface applications," in *IEEE Int. Conf. Digit. Sig. Proc.*, Santorini, Greece, vol. 1, 1-3 July 2002, pp. 301-305.

[72] L. Atlas, P. Droppo, and J. McLaughlin, "Optimizing time-frequency distributions for automatic classification," *Proc. SPIE,* vol. 3162, 1997.