A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2007 J. Neural Eng. 4 R32

(http://iopscience.iop.org/1741-2552/4/2/R03)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 130.113.111.210
The article was downloaded on 21/03/2012 at 14:58

Please note that terms and conditions apply.
TOPICAL REVIEW

A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals

Ali Bashashati\(^1\), Mehrdad Fatourechi\(^1\), Rabab K Ward\(^{1,2}\) and Gary E Birch\(^{1,2,3}\)

\(^1\) Department of Electrical and Computer Engineering, The University of British Columbia, 2356 Main Mall, Vancouver, V6T 1Z4, Canada
\(^2\) Institute for Computing, Information and Cognitive Systems, 289-2366 Main Mall, Vancouver, BC V6T 1Z4, Canada
\(^3\) Neil Squire Society, 220-2250 Boundary Rd, Burnaby, BC, V5M 4L9, Canada

E-mail: alibs@ece.ubc.ca

Received 11 October 2006
Accepted for publication 16 February 2007
Published 27 March 2007
Online at stacks.iop.org/JNE/4/R32

Abstract

Brain–computer interfaces (BCIs) aim at providing a non-muscular channel for sending commands to the external world using the electroencephalographic activity or other electrophysiological measures of the brain function. An essential factor in the successful operation of BCI systems is the methods used to process the brain signals. In the BCI literature, however, there is no comprehensive review of the signal processing techniques used. This work presents the first such comprehensive survey of all BCI designs using electrical signal recordings published prior to January 2006. Detailed results from this survey are presented and discussed. The following key research questions are addressed: (1) what are the key signal processing components of a BCI, (2) what signal processing algorithms have been used in BCIs and (3) which signal processing techniques have received more attention?

This article has associated online supplementary data files

1. Introduction

The ultimate purpose of a direct brain–computer interface (BCI) is to allow an individual with severe motor disabilities to have effective control over devices such as computers, speech synthesizers, assistive appliances and neural prostheses. Such an interface would increase an individual’s independence, leading to an improved quality of life and reduced social costs.

A BCI system detects the presence of specific patterns in a person’s ongoing brain activity that relates to the person’s intention to initiate control. The BCI system translates these patterns into meaningful control commands. To detect these patterns, various signal processing algorithms are employed.
processing methods employed in different BCI systems, and consequently to identify the methods that have not yet been explored, (b) to form a historical reference for new researchers in this field and (c) to introduce a possible taxonomy of signal processing methods in brain–computer interfaces.

The organization of the paper is as follows. In section 2, the general structure of a BCI system and the current neuromechanisms4 in BCI systems are presented. Section 3 details the procedure we followed to conduct this study. Results, discussion and conclusion are in sections 4–6, respectively.

2. General structure of a BCI system

Figure 1 shows the functional model of a BCI system (Mason and Birch 2003). The figure depicts a generic BCI system in which a person controls a device in an operating environment (e.g., a powered wheelchair in a house) through a series of functional components. In this context, the user’s brain activity is used to generate the control signals that operate the BCI system. The user monitors the state of the device to determine the result of his/her control efforts. In some systems, the user may also be presented with a control display, which displays the control signals generated by the BCI system from his/her brain activity.

The electrodes placed on the head of the user record the brain signal from the scalp, or the surface of the brain, or from the neural activity within the brain, and convert this brain activity to electrical signals. The ‘artifact processor’ block shown in figure 1 removes the artifacts from the electrical signal after it has been amplified. Note that many transducer designs do not include artifact processing. The ‘feature generator’ block transforms the resultant signals into feature values that correspond to the underlying neurological mechanism employed by the user for control. For example, if the user is to control the power of his/her mu (8–12 Hz) and beta (13–30 Hz) rhythms, the feature generator would continually generate features relating to the power-spectral estimates of the user’s mu and beta rhythms. The feature generator generally can be a concatenation of three components—the ‘signal enhancement’, the ‘feature extraction’ and the ‘feature selection/dimensionality reduction’ components, as shown in figure 1.

In some BCI designs, pre-processing is performed on the brain signal prior to the extraction of features so as to increase the signal-to-noise ratio of the signal. In this paper, we use the term ‘signal enhancement’ to refer to the pre-processing stage. A feature selection/dimensionality reduction component is sometimes added to the BCI system after the feature extraction stage. The aim of this component is to reduce the number of features and/or channels used so that very high dimensional and noisy data are excluded. Ideally, the features that are meaningful or useful in the classification stage are identified and chosen, while others (including outliers and artifacts) are omitted.

The ‘feature translator’ translates the features into logical (device-independent) control signals, such as a two-state discrete output. The translation algorithm uses linear classification methods (e.g., classical statistical analyses) or nonlinear ones (e.g., neural networks). According to the definition in Mason and Birch (2003), the resultant logical output states are independent of any semantic knowledge about the device or how it is controlled. As shown in figure 1, a feature translator may consist of two components: ‘feature classification’ and ‘post-processing’. The main aim of the feature classification component is to classify the features into logical control signals. Post-processing methods such as a moving average block may be used after feature classification to reduce the number of error activations of the system. The components between the user and control interface can be treated as a single component, a BCI transducer, which functions in a manner similar to physical transducers like a dial or switch. The role of the BCI transducer is to translate the user’s brain activity into logical (or device-independent) control signals.

The control interface translates the logical control signals (from the feature translator) into semantic control signals that are appropriate for the particular type of device used.

---

4 According to the Merriam-Webster Medical Dictionary, a bodily regulatory mechanism based in the structure and functioning of the nervous system is called a neuromechanism.
Finally, the device controller translates the semantic control signals into physical control signals that are used by the device. The device controller also controls the overall behavior of the device. For more detail, refer to Mason and Birch (2003).

Table 1 provides a simplified description of the BCI transducer components.

### 2.1. Electrophysiological sources of control in current BCIs

In BCI systems, electrophysiological sources refer to the neurological mechanisms or processes employed by a BCI user to generate control signals. Current BCIs fall into seven main categories, based on the *neuromechanisms and recording technology* they use. In Wolpaw et al (2002) BCI systems are categorized as five major groups. These categories are sensorimotor activity, P300, VEP, SCP and activity of neural cell (ANC). In this paper, two other categories were added: ‘response to mental tasks’ and ‘multiple neuromechanisms’. BCI systems that use non-movement mental tasks to control a BCI (e.g. Anderson et al (1995b) and Millan et al (1998)) assume that different mental tasks (e.g. solving a multiplication problem, imagining a 3D object, or mental counting) lead to distinct, task-specific EEG patterns and aim to detect the patterns associated with these mental tasks from the EEG. BCI systems based on multiple neuromechanisms (e.g. Gysels et al (2005)) use a combination of two or more of the above-mentioned neuromechanisms in a single design of a BCI system.

Table 2 shows these categories with a short description of each. Note that although the designs that use direct cortical recordings are included as a separate group, direct cortical recording is a recording technology and not a neuromechanism. As shown in table 2, BCI designs that use sensorimotor activity as the neural source of control can be further divided into three sub-categories: those based on changes in brain rhythms (e.g. the mu and beta rhythms), those based on movement-related potentials (MRPs) and those based on other sensorimotor activity.

### 3. Methods

The BCI designs selected for this review include every journal and conference paper that met the following criteria:

(1) One or more of the keywords BCI, BMI, DBI appeared in its title, abstract or keyword list.

(2) The work described one or more BCI designs (the minimum design content that met the criteria was a *BCI transducer* as described in section 2). There were a few papers that only reported pre-processing techniques specifically designed for brain–computer interfaces that use neural cortical recordings. These papers were reported in the pre-processing techniques. Papers that presented tutorials, descriptions of electrode technology, neuroanatomy, and neurophysiology discussions that might serve as the basis for a BCI were not included.

(3) Only papers published in English and in refereed international journals and conference proceedings were included.

(4) Designs that use functional magnetic resonance imaging (fMRI) (Weiskopf et al 2003, 2004, Yoo et al 2004), magneto-encephalography (MEG) signals (Georgopoulos et al 2005), near-infra-red spectrum (NIRS), auditory evoked potentials (Hill et al 2004, Su Ryu et al 1999) and somatosensory evoked potentials (Yan et al 2004) were not included in this paper.

(5) Papers were published prior to January 2006.

Although no paper meeting the five criteria explained above was omitted from the analysis, some papers may have been missed unintentionally. The current work should thus be regarded as an initial step to build a public database that can be updated and evolved with time.

In tables 3–8, we categorized the papers according to the signal processing methods used. For each of the design blocks of a BCI system shown in figure 1 (signal enhancement, feature extraction, feature selection/dimensionality reduction, feature classification, and post-processing), we created a table that reports the signal processing techniques used in that block. Since most of the BCI designs that use neural cortical recordings do not contain a feature extraction component, we generated a separate table (table 7) to report the translation schemes used in these designs.

Feature extraction methods used in BCI systems are closely related to the specific neuromechanism(s) used by a BCI. For example, the feature extraction algorithms employed in VEP-based BCIs are used to detect the visual evoked potentials in the ongoing EEG. In BCI systems that operate on slow cortical potentials (SCP), the extracted features are mostly used for the purpose of identifying this specific...
phenomenon in the brain signal. Thus in table 5, we categorize the feature extraction algorithms based on the seven neurological sources described in table 2. For example, the methods used in VEP-based BCI s are assembled under a different category from those used in SCP-based BCIs. A more detailed version of table 5 can be found in appendix B at stacks.iop.org/JNE/4/R32.

The different feature classification algorithms used in BCI systems are shown in table 6. As feature classification algorithms are also closely related to the type of the features that they classify, the feature classification algorithms are also categorized based on the feature extraction methods. This can be found in the supplementary data at stacks.iop.org/JNE/4/R32, which also contains a detailed version of table 7, where the classification algorithms for BCIs that use cortical neural recordings are shown.

Categorizing the feature classification methods based on the feature extraction methods used does not necessarily limit the use of a specific feature classification to a specific feature extraction method. The same applies to the categorization of the feature extraction methods based on neuromechanisms used in BCI systems. The aim here is to provide as specific

<table>
<thead>
<tr>
<th>Neuromechanism</th>
<th>Short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensorimotor activity</td>
<td>Changes in brain rhythms (mu, Beta, and gamma)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Mu rhythms in the range of 8–12 Hz and beta rhythms in the range of 13–30 Hz both originate in the sensorimotor cortex and are displayed when a person is not engaged in processing sensorimotor inputs or in producing motor outputs (Jasper and Penfield 1949). They are mostly prominent in frontal and parietal locations (Kozelka and Pedley 1990, Kubler et al 2001a, Niedermeyer and Lopes da Silva 1998). A voluntary movement results in a circumscribed desynchronization in the mu and lower beta bands (Pfurtscheller and Aranibar 1977). This desynchronization is called event-related desynchronization (ERD) and begins in the contralateral Rolandic region about 2 s prior to the onset of a movement and becomes bilaterally symmetrical immediately before execution of movement (Pfurtscheller and Lopes da Silva 1999). After a voluntary movement, the power in the brain rhythms increases. This phenomenon, called event-related synchronization (ERS), is dominant over the contralateral sensorimotor area and reaches a maximum around 600 ms after movement offset (Pfurtscheller and Lopes da Silva 1999). Gamma rhythm is a high-frequency rhythm in the EEG. Upon the occurrence of a movement, the amplitude of gamma rhythm increases. Gamma rhythms are usually more prominent in the primary sensory area.</td>
</tr>
<tr>
<td>Movement-related potentials (MRPs)</td>
<td>MRP s are low-frequency potentials that start about 1–1.5 s before a movement. They have bilateral distribution and present maximum amplitude at the vertex. Close to the movement, they become contralaterally preponderant (Babiloni et al 2004, Deecke and Kornhuber 1976, Hallett 1994).</td>
</tr>
<tr>
<td>Other sensorimotor activities</td>
<td>The sensorimotor activities that do not belong to any of the preceding categories are categorized as other sensorimotor activities. These activities are usually not restricted to a particular frequency band or scalp location and usually cover different frequency ranges. An example would be features extracted from an EEG signal filtered to frequencies below 30 Hz. Such a range covers different event-related potentials (ERPs) but no specific neuromechanism is used.</td>
</tr>
<tr>
<td>Slow cortical potentials (SCPs)</td>
<td>SCPs are slow, non-movement potential changes generated by the subject. They reflect changes in cortical polarization of the EEG lasting from 300 ms up to several seconds. Functionally, a SCP reflects a threshold regularization mechanism for local excitory mobilization (Neumann et al 2003, Wolpaw et al 2002).</td>
</tr>
<tr>
<td>P300</td>
<td>Infrequent or particularly significant auditory, visual, or somatosensory stimuli, when interspersed with frequent or routine stimuli, typically evoke in the EEG over the parietal cortex a positive peak at about 300 ms after the stimulus is received. This peak is called P300 (Allison and Pineda 2003, Kubler et al 2001a).</td>
</tr>
<tr>
<td>Visual evoked potentials (VEPs)</td>
<td>VEP s are small changes in the ongoing brain signal. They are generated in response to a visual stimulus such as flashing lights and their properties depend on the type of the visual stimulus (Kubler et al 2001a). These potentials are more prominent in the occipital area.</td>
</tr>
<tr>
<td>Response to mental tasks</td>
<td>BCI systems based on non-movement mental tasks assume that different mental tasks (e.g., solving a multiplication problem, imagining a 3D object, and mental counting) lead to distinct, task-specific distributions of EEG frequency patterns over the scalp (Kubler et al 2001a).</td>
</tr>
<tr>
<td>Activity of neural cells (ANC)</td>
<td>It has been shown that the firing rates of neurons in the motor cortex are increased when movements are executed in the preferred direction of neurons. Once the movements are away from the preferred direction of neurons, the firing rate is decreased (Donoghue 2002, Olson et al 2005).</td>
</tr>
<tr>
<td>Multiple neuromechanisms (MNs)</td>
<td>BCI systems based on multiple neuromechanisms use a combination of two or more of the above-mentioned neuromechanisms.</td>
</tr>
</tbody>
</table>

<sup>a</sup> In Ramachandran and Histein (1998), references regarding the similarity between attempted movements and real movements in the ERD of mu patterns are provided. An attempted movement occurs when a subject attempts to move some part of his/her body, but because of either a disability or the experiment control, the actual movement does not happen. Similarly, Gervins et al (1989) have shown that imaginary (attempted) movements generate movement-related potentials (MRPs) similar to those generated by actual movements. Thus, neuromechanisms corresponding to attempted movements are grouped in the same category of real movement.
information as possible about signal processing in current BCI designs and the researchers can combine any feature extraction method and/or feature classification method from different categories if necessary.

Each table includes major classes corresponding to each design block. These classes were initially determined by our team and then refined after an initial pass through the selected papers. In some cases, each major class was further divided into more specific categories. The full classification template with all the major classes and sub-classes of each design component is listed in the left column of tables 5–7, and in the left two major columns of tables 5–7. Note that in this paper, major classes are written in **bold** type and sub-classes are represented in **bold-italic** type. For example, in table 5 or appendix B (in supplementary data), VEP-based BCI designs that use some type of power-spectral parameters of the EEG are categorized under the VEP-spectral parameters class, while a BCI design that is based on the movement-related potentials (MRP) and that uses the same method is categorized under the sensorimotor activity-spectral parameters and sensorimotor activity-MRP-spectral parameters classes in table 5 and appendix B, respectively. As an example from table 6, BCI designs that use linear discriminant analysis (LDA) classifiers are categorized under LDA. Supplementary data at stacks.iop.org/JNE/4/R32, which has a more detailed version of table 6, categorizes BCI designs that use PSD features and LDA classifier under PSD-LDA class.

The category for each BCI design was determined by selecting the closest sub-class in the classification template. For the papers that reported multiple designs multiple classifications were recorded. The designs were categorized based only on what was reported in each paper. No personal knowledge of an authors’ related work was used in the classification.

In some cases, it was difficult to differentiate between the signal enhancement, feature selection and feature extraction design components of a brain–computer interface. Based on the definitions in table 1, the methods that satisfied the following four criteria were considered to be signal enhancement methods:

1. The method was implemented to improve the signal-to-noise ratio of the brain signal.
2. The output of the block had the same nature as the input brain signal (i.e. the output stayed in the temporal domain).
3. The algorithm was directly performed on the brain signal and not on the features extracted from the brain signal.
4. The method did not handle artifacts.

The common spatial patterns (CSP) method is an example of a method that satisfies the above-mentioned criteria and was thus categorized as a signal enhancement method. The principle component analysis (PCA) method is another example that sometimes satisfies the above four criteria and was categorized as a signal-enhancement method. In the cases where PCA is applied after feature extraction to reduce the dimensionality of the extracted features, it is categorized as feature selection/dimensionality reduction method. Only designs that incorporated signal enhancement algorithms other than the general band-pass filtering of the EEG, the power-line-effect rejection and the traditional normalization of the signal were reported in the signal enhancement section of this paper.

4. Results

The detailed classification results of the survey are summarized in tables 3–8 (refer to supplementary data at stacks.iop.org/JNE/4/R32 for more detailed versions of tables 5–7). As mentioned in section 3, these six tables address the signal enhancement, feature selection/dimensionality reduction, feature extraction, feature classification and post-processing methods used. The references listed for each sub-class category represent all the papers that reported on designs related to that sub-class. As such, one can find all the designs that have specific attributes of interest. For example, if one is interested in all BCI technology designs that have used parametric modeling (and specifically extracted AR parameters of the signal) to detect the sensorimotor activity, then all the references to the relevant papers can be found in table 5 under sensorimotor activity—parametric modeling (AR, AAR and ARX parameters). Alternatively, if one is looking for designs that do not have a feature extraction block but directly apply the support vector machine (SVM) classification method on the brain signal, then these papers can be easily located in the none-SVM class in supplementary data at stacks.iop.org/JNE/4/R32. Similarly, the designs that use the SVM classification method, regardless of the feature extraction technique used, are categorized under the SVM class in table 6.

To enhance the clarity of tables 3–8, the following notations are used:

(a) BCI designs based on multiple neuromechanisms (as defined in table 2) are presented in separate categories such as **MN: sensorimotor activity + response to mental tasks**, which show that the design is based on the sensorimotor activity and the response to mental tasks neuromechanisms.

(b) Two or more methods that are consecutively used in a design block are separated by ‘+’. As an example, **CSP—log transformation** denotes a design that first applies common spatial patterns (CSP) on the signal and then applies a logarithmic function on the resulting time-series.

(c) Two methods that are applied simultaneously in a design component are separated by ‘×’. For example, **AR parameters × PSD parameters** corresponds to a design that uses both autoregressive (AR) and power-spectral-density (PSD) features in the feature extraction block.

(d) **LRP + {CSP — log-transformation}** denotes designs that use the two kinds of feature extraction methods separated by ‘×’. The first method is based on the extraction of lateralized readiness potentials (LRP), and the second feature extraction method is based on consecutively applying CSP followed by a logarithm function on the signals that are grouped in ‘{’.

(e) To facilitate readability, we have provided an index of terms in appendix A.
Table 3. Pre-processing (signal enhancement) methods in BCI designs.

<table>
<thead>
<tr>
<th>Pre-processing method</th>
<th>Reference ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined CSP and PCA</td>
<td>(Xu et al. 2004b)</td>
</tr>
<tr>
<td>Singular value decomposition (SVD)</td>
<td>(Trejo et al. 2003)</td>
</tr>
<tr>
<td>Common spatio-spatial patterns (CSSP)</td>
<td>(Lemm et al. 2005)</td>
</tr>
<tr>
<td>Local averaging technique (LAT)</td>
<td>(Peters et al. 2001)</td>
</tr>
<tr>
<td>Robust Kalman filtering</td>
<td>(Bayliss and Ballard 1999, 2000a)</td>
</tr>
<tr>
<td>Wiener filtering</td>
<td>(Vidal 1977)</td>
</tr>
<tr>
<td>Sparse component analysis</td>
<td>(Li et al. 2004a)</td>
</tr>
<tr>
<td>Maximum noise fraction (MNF)</td>
<td>(Peterson et al. 2005)</td>
</tr>
<tr>
<td>Spike detection methods</td>
<td>(Oheid and Wolf 2004)</td>
</tr>
<tr>
<td>Neuron ranking methods</td>
<td>(Sanchez et al. 2004)</td>
</tr>
</tbody>
</table>

Table 4. Feature selection/dimensionality reduction methods in BCI designs.

<table>
<thead>
<tr>
<th>Feature selection/dimensionality reduction method</th>
<th>Reference ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequential forward feature selection (SFFS)</td>
<td>(Fabiani et al. 2004, Keirn and Aunon 1990)</td>
</tr>
<tr>
<td>Grid search method</td>
<td>(Glassman 2005)</td>
</tr>
<tr>
<td>Relief method (Kira and Rendell 1992)</td>
<td>(Millan et al. 2002a)</td>
</tr>
<tr>
<td>Support vector machine (SVM)-based recursive feature elimination</td>
<td>(Gysels et al. 2005)</td>
</tr>
<tr>
<td>Stepwise discriminant procedure</td>
<td>(Vidal 1977)</td>
</tr>
<tr>
<td>Linear discriminant analysis (LDA)</td>
<td>(Graimann et al. 2003b)</td>
</tr>
<tr>
<td>Fisher discriminant analysis (dimensionality reduction)</td>
<td>(Wang et al. 2004d)</td>
</tr>
<tr>
<td>Zero-norm optimization (l0-opt)</td>
<td>(Lal et al. 2004)</td>
</tr>
<tr>
<td>Orthogonal least square (OLS1) based on radial basis function (RBF)</td>
<td>(Xu et al. 2004b)</td>
</tr>
</tbody>
</table>

5. Discussion

Several points raised in the previous section deserve further comment. Figure 2 summarizes the information in tables 3, 4 and 8, which respectively address signal enhancement, feature selection/dimensionality reduction and post-processing algorithms in BCI designs. Specifically, this figure shows the number of BCI designs that use specific signal enhancement, feature selection/dimensionality reduction and post-processing techniques.
Table 5. Feature extraction methods in BCI designs. Refer to appendix B in supplementary data for a more detailed version of this table.

<table>
<thead>
<tr>
<th>Neuro-mechanism</th>
<th>Feature extraction method</th>
<th>Reference ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Signal envelope — cross-correlation</td>
<td>(Wang et al 2004a, 2004b)</td>
</tr>
<tr>
<td></td>
<td>Mixed filter</td>
<td>(Hinterberger et al 2003)</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>(Hinterberger et al 2003, Schroder et al 2003)</td>
</tr>
<tr>
<td>P300</td>
<td>Cross-correlation</td>
<td>(Bayliss and Ballard 1999, 2000a, 2000b, Farwell and Donchin 1988)</td>
</tr>
<tr>
<td></td>
<td>Stepwise discriminant analysis</td>
<td>(Donchin et al 2000, Farwell and Donchin 1988)</td>
</tr>
<tr>
<td></td>
<td>Matched filtering</td>
<td>(Serby et al 2005)</td>
</tr>
<tr>
<td></td>
<td>PPM</td>
<td>(Jansen et al 2004)</td>
</tr>
<tr>
<td></td>
<td>Area calculation</td>
<td>(Farwell and Donchin 1988)</td>
</tr>
<tr>
<td></td>
<td>Area and peak picking</td>
<td>(Kaper and Ritter 2004b, Xu et al 2004a)</td>
</tr>
<tr>
<td></td>
<td>Not mentioned (calculated P300 but details not mentioned)</td>
<td>(Bayliss 2003, Polikoff et al 1995)</td>
</tr>
<tr>
<td></td>
<td>Asymmetry ratio of different band powers</td>
<td>(Su et al 1999)</td>
</tr>
</tbody>
</table>
Table 5. (Continued.)

<table>
<thead>
<tr>
<th>Neuro-mechanism</th>
<th>Feature extraction method</th>
<th>Reference ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response to mental tasks</strong></td>
<td>Cross-correlation</td>
<td>(Sutter 1992)</td>
</tr>
<tr>
<td>Response to mental tasks</td>
<td>Amplitude between N2 and P2 peaks</td>
<td>(Lee et al 2005)</td>
</tr>
<tr>
<td>Response to mental tasks</td>
<td>None</td>
<td>(Guan et al 2005, Vidal 1997)</td>
</tr>
<tr>
<td><strong>ANC</strong></td>
<td>None</td>
<td>(Anderson et al 1995a, Panuccio et al 2002)</td>
</tr>
<tr>
<td>ANC</td>
<td>Cross-covariance — PCA</td>
<td>(Isaacs et al 2000)</td>
</tr>
<tr>
<td>ANC</td>
<td>LBG vector quantization (VQ)</td>
<td>(Darmanjian et al 2003)</td>
</tr>
<tr>
<td>MN: SCP + other brain rhythms</td>
<td>SCP calculation + power spectral parameters</td>
<td>(Gysels and Celka 2004, Gysels et al 2005)</td>
</tr>
</tbody>
</table>

*Designs that differentiate between relaxed state and movement tasks are considered in ‘sensorimotor activity + response to mental tasks’ category.

In the remainder of this section we highlight the top three or four methods that have been used in the signal processing blocks of BCI systems (as introduced in section 2).

Of the 96 BCI designs that employ signal enhancement techniques before extracting the features from the signal, 32% use surface Laplacian (SL), 22% use either principal component analysis (PCA) or independent component analysis (ICA), 14% use common spatial patterns (CSP) and 11% use common average referencing (CAR) techniques. Thirty-eight of the reported BCI designs employ feature selection/dimensionality reduction algorithms; 26% of these 38 designs use genetic algorithms (GA), 24% use distinctive sensitive learning vector quantization (DSLVQ), and 13% use PCA.

Of the 30 BCI designs that use post-processing algorithms to reduce the amount of error in the output of the BCI system, 57% use averaging techniques and consider rejecting activations that have low certainty, 27% consider using the debounce block (or refractory period) to deactivate the output for a short period of time when a false activation is detected, and 16% use event-related negativity (ERN) signals to detect error activations.

Figure 3 summarizes the results presented in table 5, and shows the number of BCI designs that are based on...
### Table 6. Feature classification methods in BCI designs that use EEG and ECoG recording technology.

<table>
<thead>
<tr>
<th>Feature classification method</th>
<th>Reference ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Committee of MLP NN</strong></td>
<td>(Millan et al. 2000b, Varsta et al. 2000)</td>
</tr>
<tr>
<td><strong>FIR-MLP NN</strong></td>
<td>(Haselsteiner and Pfurtscheller 2000)</td>
</tr>
<tr>
<td><strong>Committee of Platt's RAN algorithm (Platt 1991)</strong></td>
<td>(Millan et al. 1998)</td>
</tr>
<tr>
<td><strong>Committee of NNs trained with Adaboost</strong></td>
<td>(Boostani and Moradi 2004)</td>
</tr>
<tr>
<td><strong>Committee of single perceptrons with no hidden layers</strong></td>
<td>(Peters et al. 2001)</td>
</tr>
<tr>
<td><strong>TRNN</strong></td>
<td>(Barreto et al. 1996a, 1996b)</td>
</tr>
<tr>
<td><strong>kMeans – LVQ</strong></td>
<td>(Borisoff et al. 2004)</td>
</tr>
<tr>
<td><strong>Growing hierarchical SOM</strong></td>
<td>(Liu et al. 2005)</td>
</tr>
<tr>
<td><strong>Custom designed local NN</strong></td>
<td>(Cincotti et al. 2003b)</td>
</tr>
<tr>
<td><strong>Fuzzy ARTMAP</strong></td>
<td>(Palaniappan et al. 2002)</td>
</tr>
<tr>
<td><strong>Single layer NN</strong></td>
<td>(Garcia et al. 2002)</td>
</tr>
<tr>
<td><strong>RBF-NN</strong></td>
<td>(Hung et al. 2005)</td>
</tr>
<tr>
<td><strong>Static neural classifier (Adaline)</strong></td>
<td>(Barreto et al. 1996a, 1996b)</td>
</tr>
<tr>
<td><strong>Gamma NN</strong></td>
<td>(Barreto et al. 1996a, 1996b)</td>
</tr>
<tr>
<td><strong>Sparse FLD</strong></td>
<td>(Blankertz et al. 2002a)</td>
</tr>
<tr>
<td><strong>Nonlinear discriminant function</strong></td>
<td>(Fabiani et al. 2004)</td>
</tr>
<tr>
<td><strong>Bayes quadratic classifier</strong></td>
<td>(Keim and Aunon 1990)</td>
</tr>
<tr>
<td><strong>Linear Bayesian decision rule</strong></td>
<td>(Vidal 1977)</td>
</tr>
<tr>
<td><strong>Linear classifier based on time-warping</strong></td>
<td>(Mason and Birch 2000)</td>
</tr>
<tr>
<td><strong>Logistic regression</strong></td>
<td>(Parra et al. 2002, 2003a)</td>
</tr>
<tr>
<td><strong>Linear classifier (no details)</strong></td>
<td>(Ramoser et al. 2000)</td>
</tr>
<tr>
<td><strong>Single layer Perceptron model (a linear classifier)</strong></td>
<td>(Li et al. 2004b, Wang et al. 2004d)</td>
</tr>
<tr>
<td><strong>Two-dimensional linear classifier trained by a non-enumerative search procedure</strong></td>
<td>(Cheng et al. 2004)</td>
</tr>
<tr>
<td><strong>ZDA</strong></td>
<td>(Hinterberger et al. 2003)</td>
</tr>
<tr>
<td><strong>LDS</strong></td>
<td>(Lee and Choi 2002)</td>
</tr>
<tr>
<td><strong>SOM-based SSP</strong></td>
<td>(Millan et al. 2000b, 2002b)</td>
</tr>
</tbody>
</table>
sensorimotor activity, SCP, VEP, P300, activity of neural cells, ‘response to mental tasks’ and multiple neuromechanisms and use different feature extraction techniques.

Based on the results of figure 3, 41% of the BCIs that are based on the sensorimotor activity use power-spectral-density features, 16% rely on parametric modeling of the data, 13% use time–frequency representation (TFR) methods and 6% do not employ any feature extraction methods. 74% of the SCP-based BCI designs calculate SCP signals using low-pass filtering methods, and 64% of the VEP-based BCIs use power-spectral features at specific frequencies. 26% of the BCIs based on P300 calculate the peaks of the signal in a specific time window to detect the P300 component of the EEG; 22% use TFR-based methods, 22% use no feature extraction method, and 15% use cross-correlation with a specific template. 41% of the BCI designs that use mental tasks to control a BCI use power-spectral features and 37% use parametric modeling of the input signal. As most of the BCI designs that are based on neural cortical recordings mainly try to model the direct relationship between the neural cortical recordings and movements, they do not use a feature-extraction algorithm. 45% of the BCI designs that are based on multiple neuromechanisms rely on power-spectral features, 17% use parametric modeling, and 17% use time–frequency representation (TFR) methods.

Summarizing tables 6 and 7, the number of BCI designs that use different feature classification algorithms are shown in figure 4. About 75% of the BCI designs use classification schemes that are not based on neural networks (NN). These are composed of those methods that

<table>
<thead>
<tr>
<th><strong>Feature classification method</strong></th>
<th><strong>Reference ID</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>HMM-based techniques</td>
<td>Rezek et al 2003</td>
</tr>
<tr>
<td>CHMM</td>
<td>(Blankertz et al 2002a, Ivanova et al 1995)</td>
</tr>
<tr>
<td>AR HMM</td>
<td>(Panuccio et al 2002)</td>
</tr>
<tr>
<td>HMM + SVM</td>
<td>(Cincotti et al 2003b, Allinson and Pineda 2003b)</td>
</tr>
<tr>
<td>HMM</td>
<td>(Lee and Choi 2003a, Panuccio et al 2003b)</td>
</tr>
<tr>
<td>SVM</td>
<td>(Blankertz et al 2002a, Ivanova et al 1995)</td>
</tr>
<tr>
<td>Threshold detector</td>
<td>(Blankertz et al 2002a, Panuccio et al 2003b)</td>
</tr>
<tr>
<td>Linear combination — threshold detector</td>
<td>(Townsend et al 2003b)</td>
</tr>
<tr>
<td>Continuous feedback + threshold detector</td>
<td>(Kaiser et al 2001a)</td>
</tr>
<tr>
<td>Linear combination — continuous feedback</td>
<td>(Fabiani et al 2003, Hinterberger et al 2003)</td>
</tr>
<tr>
<td>Continuous feedback using MD</td>
<td>(Schloegl et al 2003)</td>
</tr>
<tr>
<td>Continuous audio feedback</td>
<td>(Kaiser et al 2001, 2003)</td>
</tr>
<tr>
<td>Static classifier that is inferred with sequential variational inference</td>
<td>(Curran et al 2004, Sykacek et al 2004)</td>
</tr>
<tr>
<td>Random forest algorithm</td>
<td>(Neuper et al 1999)</td>
</tr>
</tbody>
</table>

# Table 6 (Continued.)

* Regularization may be applied before LDA classification scheme.
use threshold detectors as the feature classifier or as part of the feature classification scheme (27%), linear discriminant (either LDA or FLD) classifiers (26%), those that show continuous feedback of the extracted features (16%), and those that use support vector machines (SVM) (11%). 27% of the neural-network-based classifiers are based on the multi-layer perceptrons (MLP) neural network and 39% are based on learning-vector-quantization (LVQ) classification scheme. 

During our analysis of the literature, a number of salient points about the signal processing methods emerged. We think that some of these points are worth sharing with the BCI research community. In the following sections we summarize some of them. Note that these observations are based on comments made by the researchers in their published papers and are also based on the trends in the literature; they do not cover all the methods reported in the literature.
5.1. Signal enhancement

Signal enhancement (pre-processing) algorithms have been used for brain–computer interfaces that are based on the EEG and the activity of the neural cells (ANC), but no signal enhancement algorithms have been applied on electrocorticogram (ECoG)-based brain–computer interfaces. Given the huge difference in the characteristics of EEG and ANC, the signal enhancement algorithms used in EEG-based and ANC-based BCIs have very little overlap; only PCA has been used in both groups. While ANC-based BCIs mostly aim at spike detection, ranking, and sorting neuron activities, EEG-based ones mostly transform or select EEG channels that yield better performance. Overall, the use of a pre-processing stage before feature extraction (if applied) has been proven to be useful. The choice of a suitable pre-processing technique however is dependent on several factors such as the recording technology, number of electrodes, and neuromechanism of the BCI. Next, we discuss some of the techniques used in signal enhancement in EEG-based BCI systems. Specifically, a discussion of spatial filtering including referencing methods and common spatial patterns (CSP) is presented. These methods are among the most used techniques that have become increasingly popular in BCI studies.

5.1.1. Referencing methods

Referencing methods are considered as spatial filters. The proper selection of a spatial filter for any BCI is determined by the location and extent of the control signal (e.g. the mu rhythm) and of the various sources of EEG or non-EEG noise. The latter two are not completely defined and presumably vary considerably between different studies and both across and within individuals (McFarland et al. 1997).

Figure 2. Signal enhancement, feature selection/dimensionality reduction and post-processing methods in BCI designs.

For BCIs that use the mu and beta rhythms, the common average referencing (CAR) and Laplacian methods are superior to the ear reference method. This may be because these methods use high-pass spatial filters and enhance the focal activity from the local sources (e.g. the mu and the beta rhythms) and reduce the widely distributed activity, including that resulting from distant sources (e.g. EMG, eye movements and blinks, visual alpha rhythm). Comparing the two variations of the Laplacian filtering methods (the large Laplacian and the small Laplacian), it is shown that the large Laplacian method is superior to the small Laplacian method in BCI systems that use the mu rhythm (McFarland et al. 1997).

Although an accurate Laplacian estimate from raw potentials requires many electrodes, one study showed that the recognition rates were increased by using a small number of electrodes only (Cincotti et al. 2003a). In this case, the linear combination of channels implementing the Laplacian estimation was likely to have caused a favorable transformation of the signals to recognize different patterns in the ongoing EEG.

5.1.2. Common spatial patterns (CSP). CSP is a signal enhancement method that detects patterns in the EEG by incorporating the spatial information of the EEG signal. Some of its features and limitations include the following.

An advantage of the CSP method is that it does not require the a priori selection of subject-specific frequency bands. Knowledge of these bands, however, is necessary for the band-power and frequency-estimation methods (Guger et al. 2000b). One disadvantage of the CSP method is that it requires the use of many electrodes. However, the inconvenience of applying more electrodes can be rationalized by improved performance (Guger et al. 2000b, Pfurtscheller et al. 2000). The major problem in the application of CSP is its sensitivity to artifacts in the EEG. Since the covariance matrices are used as the basis for calculating the spatial filters, and are estimated with a comparatively small number of examples, a single trial contaminated with artifacts can unfortunately cause extreme changes to the filters (Guger et al. 2000b, Ramoser et al. 2000). Since the CSP method detects spatial patterns in the EEG, any change in the electrode positions may render the improvements in the classification accuracy gained by this method useless. Therefore, this method requires almost identical electrode positions for all trials and sessions which may be difficult to accomplish (Ramoser et al. 2000).

5.2. Feature extraction

In this section we discuss some of the feature extraction techniques that have received more attention in BCI systems. Specifically, time and/or frequency representation methods, parametric modeling and specific techniques of modeling neural cortical recordings are discussed.
5.2.1. **Time and/or frequency methods.** A signal, as a function of time, may be considered as a representation with perfect temporal resolution. The magnitude of the Fourier transform (FT) of the signal may be considered as a representation with perfect spectral resolution but with no temporal information. Frequency-based features have been widely used in signal processing because of their ease of application, computational speed and direct interpretation of the results. Specifically, about one-third of BCI designs have used power-spectral features. Due to the non-stationary nature of the EEG signals, these features do not provide any time domain information. Thus, mixed time–frequency representations (TFRs) that map a one-dimensional signal into a two-dimensional function of time and frequency are used to analyze the time-varying spectral content of the signals. It has been shown that TFR methods may yield performance improvements comparing to the traditional FT-based methods (e.g. Qin et al. (2005) and Bostanov (2004)). Most of the designs that employ TFR methods use wavelet-based feature extraction algorithms. The choice of the particular wavelet used is a crucial factor in gaining useful information from wavelet analysis. Prior knowledge of the physiological activity in the brain can be useful in determining the appropriate wavelet function.

Correlative TFR (CTFR) is another time–frequency representation method that, besides the spectral information, provides information about the time–frequency interactions between the components of the input signal. Thus, with the CTFR the EEG data samples are not independently analyzed (as in the Fourier transform case) but their relationship is also taken into account. One drawback of the CTFR resides in its relative high sensitivity to noise. Consequently, the most important values of the CTFR in terms of classification must be selected (Garcia et al. 2003a, 2003b).

5.2.2. **Parametric modeling.** Parametric approaches assume the time series under analysis to be the output of a given linear mathematical model. They require an *a priori* choice of the structure and order of the signal generation mechanism model (Weitkunat 1991). The optimum model order is best estimated not only by maximizing the fitness but also by limiting the model’s complexity. For noisy signals, if the model’s order is too high, spurious peaks in the spectra will result. On the other hand, if the order is too low, smooth spectra are obtained (Kelly et al. 2002a, Polak and Kostov 1998, Weitkunat 1991).

For short EEG segments, parametric modeling results in better frequency resolution and a good spectral estimate. Note that parametric modeling may yield poor estimates if the length of the EEG segments processed is too short (Birch 1988). For such modeling, there is no need for *a priori* information about potential frequency bands, and there is no need to window the data in order to decrease the spectral leakage. Also, the
frequency resolution does not depend on the number of data points (Guger et al. 2003a, Polak and Kostov 1998, Weitkunat 1991). Estimating these parameters, however, is very sensitive to artifacts (Birch 1988, Guger et al. 2003a).

Special attention should be paid to the choice of the sampling rate in parametric modeling (Weitkunat 1991), since severely oversampled signals tend to show only very small amplitude differences between successive samples. Hence, low-order models produce small prediction errors, giving the false illusion that an adequate model has been obtained. The sampling rates dictated by the Nyquist criterion are recommended.

5.2.3. Modeling the neural firing rates. Extraction algorithms for motor control operate on spike trains, recorded from a population of cortical units, mostly with the purpose of predicting arm trajectories. Several extraction methods such as linear filtering methods and neural networks have been used to determine arm movement trajectories from neural firing rates. We summarize below a few important issues in modeling the neural firing rates. A more detailed critical discussion of extraction algorithms for cortical control of arm prosthetics can be found in Schwartz et al. (2001).

One limitation of linear filter methods is that they rely on an a priori model of movement-related neuronal responses. Artificial neural network (ANN) solutions can optimize each cell’s contribution to the population prediction (Schwartz et al. 2001). Several of the algorithms used are based on the position of the moving limb. In the primary motor cortex at least, this parameter is more poorly represented than the velocity during movement. With most algorithms, the different sources of variability need to be specified explicitly because some sort of optimal function is being modeled to the cell response (Schwartz et al. 2001).

The performance of the modeling techniques is constrained by their training sets and may be limited, both in terms of extrapolation beyond and interpolation within the training set when new data are applied. The success of the linear filters is due to the underlying linearity of the relationship between firing rate and movement direction. These filters are limited by the conditions used to fit their coefficients and may suffer from the same training constraints as ANNs (Schwartz et al. 2001).
5.3. Feature selection/dimensionality reduction

Feature selection algorithms are used in BCI designs to find the most informative features for classification. This is especially useful for BCI designs with high dimensional input data, as it reduces the dimension of the feature space. Since feature selection block reduces the complexity of the classification problem, higher classification accuracies might be achieved. The experiments carried out in Flotzinger et al (1994) and Pregenzer and Pfurtscheller (1999) show that when feature selection is used, the classification accuracy is better than when all the features are used.

Principal component analysis (PCA) and genetic algorithms (GA) are among the mostly used feature selection and/or dimensionality reduction methods in BCIs. PCA has also been widely used in pre-processing stage of BCI designs. PCA is a linear transformation that can be used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the 'most important' aspects of the data. PCA has the distinction of being the optimal linear transformation for keeping the subspace that has largest variance. PCA only finds linear subspaces, works best if the individual components have Gaussian distributions, and is not optimized for class separability. One other possible application area of PCA is in classification stage, in which, PCA is applied for weighting input features. While a standard neural network, such as the multi-layer perceptrons (MLP), can do the necessary classification itself, in some cases doing a PCA in parallel and weighting input features can give better results as it simplifies the training of the rest of the system.

Unlike PCA, GAs are heuristic search techniques in the problem space. GAs typically maintain a constant-sized population of individuals which represent samples of the space to be searched. Each individual is evaluated on the basis of its overall fitness with respect to the given application domain. New individuals (samples of the search space) are produced by selecting high performing individuals to produce ‘offspring’ which retain many of the features of their ‘parents’. This eventually leads to a population that has improved fitness with respect to the given goal. Genetic algorithms have demonstrated substantial improvement over a variety of random and local search methods (De Jong 1975). This is accomplished by their ability to exploit accumulating information about an initially unknown search space in order to bias subsequent search into promising subspaces. Since GAs are basically a domain-independent search technique, they are ideal for applications where domain knowledge and theory is difficult or impossible to provide (De Jong 1975). An important step in developing a GA-based search is defining a suitable fitness function. An ideal fitness function correlates closely with the algorithm’s goal, and yet may be computed quickly. Speed of execution is very important, as a typical genetic algorithm must be iterated many, many times in order to produce a usable result for a non-trivial problem. Definition of the fitness function is not straightforward in many cases and often is performed iteratively if the fittest solutions produced by a GA are not what is desired.

5.4. Feature classification

Linear classifiers are generally more robust than nonlinear ones. This is because linear classifiers have fewer free parameters to tune, and are thus less prone to over-fitting (Muller et al 2003a). In the presence of strong noise and outliers, even linear systems can fail. One way of overcoming this problem is to use regularization. Regularization helps limit (a) the influence of outliers and strong noise, (b) the complexity of the classifier and (c) the raggedness of the decision surface (Muller et al 2003a).

It is always desirable to avoid reliance on nonlinear classification methods, if possible, because these methods often involve a number of parameters whose values must be chosen appropriately. However, when there are large amounts of data and limited knowledge of the data, nonlinear methods are better suited in finding the potentially more complex structure in the data. In particular, when the source of the data to be classified is not well understood, using methods that are good at finding nonlinear transformation of the data is suggested. In these cases, kernel-based and neural-networks-based methods can be used to determine the transformations. Kernel-based classifiers are classification methods that apply a linear classification in some appropriate (kernel) feature space. Thus, all the beneficial properties of linear classification are maintained, but at the same time, the overall classification is nonlinear. Examples of such kernel-based classification methods are support vector machines (SVMs) and kernel Fisher discriminant (KPD) (Muller et al 2003a). For a more detailed critical discussion regarding linear and nonlinear classifiers in brain–computer interfaces, refer to Muller et al (2003a).

Some BCI designs have used classification algorithms such as FIR-MLP and TBNN that utilize temporal information of the input data (Haselstein and Pfurtscheller 2000, Ivanova et al 1995). The motivation for using such classifiers is that the patterns to be recognized are not static data but time series. Thus, the temporal information of the input data can be used to improve the classification results (Haselstein and Pfurtscheller 2000). Utilizing the temporal information of features is not necessarily performed directly in the classification stage, and can be done with a static classifier like MLP and a mapping of the temporal input data to static data. However, using classifiers such as FIR-MLP and TBNN that directly utilize temporal information may yield better performances as they are much better suited for exploiting temporal information contained in the time series to be classified. Regardless of the method that is used for exploiting temporal information, these approaches are preferred over static classification as they may increase the performance of BCI systems.

Using a group (committee) of classifiers rather than using a single classifier might also yield to better performances of
BCI systems. Only a few BCI designs have employed such an approach in classifying features and achieved performance improvements (Millan et al 2000b, 2002b, Peters et al 2001, Varsta et al 2000). The classification accuracy of the committee depends on how much unique information each committee member contributes to classification. A committee of classifiers usually yields better classification accuracy than any individual classifier could provide, and can be used to combine information from several channels, i.e., from different spatial regions (Peters et al 2001).

As the number of epochs available for evaluating a BCI system is small, using a technique that reduces the bias of the estimated performance on a specific dataset is highly recommended. This is especially important when different architectures of a certain design are being compared. K-fold cross-validation and statistical significance tests are especially useful for these cases (e.g. refer to Anderson et al (1998), Kelly et al (2002b), Lalor et al (2005), Obermaier et al (2001d) and Peterson et al (2005)). K-fold cross-validation can be used simply to estimate the generalization error of a given model, or it can be used for model selection by choosing one of several models that has the smallest estimated generalization error but it is not suitable for online evaluations. A value of 5 to 10 for K is recommended for estimating the generalization error. For an insightful discussion of the limitations of cross-validatory choice among several evaluation methods, see Stone (1977).

5.5. Post-processing

Post-processing techniques can be utilized in most of the BCI designs to decrease the error rates. Some post-processing techniques can be designed specifically for a target application. For example, when a BCI system is used to activate a spelling device, some letters can be omitted without losing information. The system can also take into consideration the conditional probabilities of letters provided by one or two preceding letters and make corresponding suggestions to the patient (Kubler et al 1999). Such techniques may also be feasible for other applications and consequently increase the performance of the BCI systems.

There is a possibility that just after the end of a trial, some features of the brain signal reveal whether or not the trial was successful (that is, whether the outcome was or was not what the subject desired). These features are referred to as error potentials and can be used to detect errors in a BCI system and void the outcome. This error detection approach was encouraged by evidence that errors in conventional motor performances have detectable effects on the EEG recorded just after the error occurs (Falkenstein et al 1995, Falkenstein et al 2001, Gehring et al 1995). Whatever the nature of the error potential, the central decision for a BCI is how useful the error potential can be in detecting errors in single trials, and thereby improving accuracy. While its signal-to-noise ratio (SNR) is low, the error potential can improve the performance of a BCI system. In the meantime, better methods for recognizing and measuring the error potential could substantially improve its SNR, and thereby increase its impact on accuracy of a BCI system. Such error potentials have been used in a few BCI systems to increase the performance (Bayliss et al 2004, Blankertz et al 2002b, 2003, Parra et al 2003b, Schalk et al 2000).

Another useful technique in decreasing false activations of BCI systems is to consider a measure of confidence in classification. In such a case, the output of the system can only be activated when the probability of the output being in an active state is greater than a given probability threshold or some criterion. Otherwise, the response of the BCI is considered ‘unknown’ and rejected to avoid making risky decisions. This is a useful way of reducing false decisions of the system (e.g., Cincotti et al (2003b), Millan et al (1998) and Penny et al (2000)) and might be used in any BCI design.

Considering mechanisms like debouncing the output of BCI designs also can reduce the number of false activations (Bashashati et al 2005, Borisoff et al 2004, Fatourechi et al 2004, 2005, Muller-Putz et al 2005b, Obeid and Wolf 2004, Pfurtscheller et al 2005, Townsend et al 2004). These methods are specifically useful for so-called asynchronous (self-paced) BCIs. Since false positives could happen in periods longer than just a few samples, using the debouncing technique in a manner similar to the debouncing of physical switches is expected to improve false activation rates (but with a cost in decreased re-activation time). The debounce component continuously monitors the output of the classifier. After an activation is detected (e.g. a change in logical state from ‘0’ to ‘1’ in a binary classifier), the output is activated for one time sample, then the output is forced to an inactive state for Td − 1 time samples, where Td is the debounce time period in samples. In some studies this time period is referred to as refractory period. As the debounce period is increased, the false activation rate is decreased for a given true positive rate. However, with increasing this time period, the re-activation time of the BCI system is impacted. The trade-off is clear and one needs to consider this for a given application.

6. Conclusions

We have completed the first comprehensive survey of signal processing methods used in BCI studies and published prior to January 2006. The results of this survey form a valuable and historical cross-reference for methods used in the following signal processing components of a BCI design: (1) pre-processing (signal enhancement), (2) feature selection/dimensionality reduction, (3) feature extraction, (4) feature classification and (5) post-processing methods. This survey shows which signal processing techniques have received more attention and which have not. This information is also valuable for newcomers to the field, as they can now find out which signal processing methods have been used for a certain type of a BCI system.

Many signal processing methods have been proposed and implemented in various brain–computer interfaces and comparison of these methods for different BCI applications
would be a useful task. However, at this point we cannot perform this task given the diversity of brain–computer interface systems from different aspects such as target application, neuromechanism used, amount of data tested, number of subjects and the amount of training they have received, recording systems, and experimental paradigms. Also acknowledged in McFarland et al (2006), we think that comparison of methods would be possible in well-designed systematic studies (Jackson et al 2006) and on established datasets such as the BCI Competition datasets (Blankertz et al 2004, 2006). We believe that, for a fair comparison of methods, more data would be needed as comparing methods with the data of one or two subjects does not necessarily guarantee the same findings on a larger subject pool. Jackson et al (2006) have provided a step toward this goal by proposing some ways to establish a systematic study both in design and reporting the results and we think that this task would only be possible with the collective help of all the researchers in this field.

We hope that this study will spawn further discussion of signal processing schemes for BCI designs. The proposed taxonomy and classes defined in tables 3–8 represent a proposed set of subcategories, not a final one, and we encourage others to revise or expand upon this initial set. Our direction in the future is to establish an online public-accessible database where research groups will be able to submit their signal processing designs as well as propose revisions/expansions of the proposed definitions and categories presented in this paper.

Appendix A. Index of terms

<table>
<thead>
<tr>
<th>Index term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAR</td>
<td>Adaptive auto-regressive</td>
</tr>
<tr>
<td>AEP</td>
<td>Auditory evoked potential</td>
</tr>
<tr>
<td>AGR</td>
<td>Adaptive Gaussian representation</td>
</tr>
<tr>
<td>ALN</td>
<td>Adaptive logic network</td>
</tr>
<tr>
<td>ANC</td>
<td>Activity of neural cells</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial neural networks</td>
</tr>
<tr>
<td>AR</td>
<td>Auto-regressive</td>
</tr>
<tr>
<td>ARTMAP</td>
<td>Adaptive resonance theory MAP</td>
</tr>
<tr>
<td>ARX</td>
<td>Auto-regressive with exogenous input</td>
</tr>
<tr>
<td>BPF</td>
<td>Band-pass filter</td>
</tr>
<tr>
<td>C4.5</td>
<td>–</td>
</tr>
<tr>
<td>CAR</td>
<td>Common average referencing</td>
</tr>
<tr>
<td>CBR</td>
<td>Changes in brain rhythms</td>
</tr>
<tr>
<td>CCTM</td>
<td>Cross-correlation-based template matching</td>
</tr>
<tr>
<td>CER</td>
<td>Coarse-grained entropy rate</td>
</tr>
<tr>
<td>CHMM</td>
<td>Coupled hidden Markov model</td>
</tr>
<tr>
<td>CN2</td>
<td>–</td>
</tr>
<tr>
<td>CSP</td>
<td>Common spatial patterns</td>
</tr>
<tr>
<td>CSSD</td>
<td>Common spatial subspace decomposition</td>
</tr>
<tr>
<td>CSP</td>
<td>Common spatio-spectral patterns</td>
</tr>
<tr>
<td>CTFR</td>
<td>Correlative time–frequency representation</td>
</tr>
<tr>
<td>CTF SR</td>
<td>Correlative time–frequency–space representation</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier transform</td>
</tr>
<tr>
<td>DSLVQ</td>
<td>Distinctive sensitive learning vector quantization</td>
</tr>
<tr>
<td>ERD</td>
<td>Event-related desynchronization</td>
</tr>
<tr>
<td>ERN</td>
<td>Event-related negativity</td>
</tr>
</tbody>
</table>

Index term Description
ERS | Event-related synchronization |
FLD | Fisher’s linear discriminant |
FFT | Fast Fourier transform |
Freq-Norm | Frequency normalization |
GA | Genetic algorithm |
GAM | Generalized additive models |
GLA | Generalized linear models |
GPER | Gaussian process entropy rates |
HMM | Hidden Markov model |
ICA | Independent component analysis |
IFFT | Inverse fast Fourier transform |
k-NN | k-nearest neighbor |
LDA | Linear discriminant analysis |
LDS | Linear dynamical system |
LGM | Linear Gaussian models implemented by Kalman filter |
LMS | Least mean square |
LPC | Linear predictive coding |
LPF | Low-pass filter |
LRP | Lateralized readiness potential |
LVQ | Learning vector quantization |
MD | Mahalanobis distance |
MLP | Multi-layer perceptron neural networks |
MN | Multiple neuromechanisms |
MNF | Maximum noise fraction |
MRA | Movement-related activity |
NID3 | Non-negative matrix factorization |
NN | Neural networks |
OLS1 | Orthogonal least square |
OPM | Outlier processing method |
PCA | Principal component analysis (a.k.a. Karhounen–Loeve transform) |
PLV | Phase locking values |
PPM | Piecewise Prony method |
PSD | Power-spectral density |
RBF | Radial basis function |
RFE | Recursive feature/channel elimination |
RNN | Recurrent neural network |
SA–UK | Successive averaging and/or considering choice of unknown |
SCP | Slow cortical potentials |
SE | Spectral-entropy |
SFFS | Sequential forward feature selection |
SL | Surface Laplacian |
SOFFN | Self-organizing feature neural network |
SOM | Self-organizing map |
SSEP | Somatosensory evoked potential |
SSP | Signal space projection |
SSVEP | Steady state visual evoked potential |
STD | Standard deviation |
SVD | Singular value decomposition |
SVM | Support vector machine |
SVR | Support vector machine regression |
SWDA | Stepwise discriminant analysis |
TBNN | Tree-based neural network |
TFR | Time–frequency representation |
VEFD | Variable epoch frequency decomposition |
VEP | Visual evoked potential |
WE | Wavelet entropy |
WK | Wiener–Khinchine |
ZDA | Z-scale-based discriminant analysis |
References


Allison B Z and Pineda J A 2005 Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system Int. J. Psychophysiol. 59 127–40


Babiloni F, Bianchi L, Cattini A, Salinari S, Marciani M G and Cincotti F 2001a Mahalanobis distance-based classifiers are able to recognize EEG patterns by using few EEG electrodes Proc. 23rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Istanbul) pp 651–4


Bayliss J D and Ballard D H 1999 Single trial P300 recognition in a virtual environment Proc. Int. ICSC Symp. on Soft Computing in Biomedicine (Genova, Italy)


Bostanov V 2004 BCI competition 2003–data sets ib and IIIb: feature extraction from event-related brain potentials with the


Burke D, Kelly S, de Chazal P and Reilly R 2002 A simultaneous filtering and feature extraction strategy for direct brain interfacing Proc. 24th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Houston, TX) pp 279–80


Cheng M, Gao S 1999 An EEG-based cursor control system Proc. 21st Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society & Annual Fall meeting of the Biomedical Engineering Society (Atlanta, GA) p 669


Cheng M, Xu D, Gao X and Gao S 2001 Brain–computer interface with high transfer rates Proc. 8th Int. Conf. on Neural Information Processing (Shanghai, China)


De Jong K A 1975 An analysis of the behaviour of a class of genetic adaptive systems PhD Thesis University of Michigan, Ann Arbor, MI (Diss. Abstr. Int. 36(10), 5140B, University Microfilms No. 76–9381.44)


Donoghue JP 2002 Connecting cortex to machines: recent advances in brain interfaces Nat. Neurosci. 5 1085–8


IEEE Int. Conf. on Acoustics, Speech and Signal Processing (Philadelphia, PA) pp 345–8

Flotzinger D, Pregenzer M and Pfurtscheller G 1994 Feature selection with distinction sensitive learning vector quantisation and genetic algorithms Proc. IEEE Int. Conf. on Neural Networks (Orlando, FL) pp 3448–51

Fukada S, Tatsumi D, Tsujimoto H and Inokuchi S 1998 Studies of input speed of word inputting system using event-related potential Proc. 20th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Hong Kong) pp 1458–60


Garcia G, Ebrahimi T and Vesin J M 2002 Classification of EEG signals in the ambiguity domain for brain–computer interface applications Proc. IEEE 14th Int. Conf. on Digital Signal Processing (Santorini, Greece) pp 301–305


Garcia G, Ebrahimi T and Vesin J M 2003b Support vector EEG classification in the四级ier and time-frequency correlation domains Proc. 1st IEEE-EMBS Conf. on Neural Engineering (Capri Island, Italy) pp 591–4


Guan J, Chen Y, Lin J, Yun Y and Huang M 2005 N2 components as features for brain–computer interface Proc. 1st Int. Conf. on Neural Interface and Control (Wuhan, China) pp 45–9


Guger C, Edlinger G and Pfurtscheller G 2003b How many people are able to operate an EEG-based brain–computer interface (BCI)? Presented at the 2nd Int. Meeting on Brain–Computer Interfaces for Communication and Control (Albany, NY)

Guger C, Muller G, Neuper C, Krausz G, Niedermayer I and Pfurtscheller G 2000a Brain–computer communication system: the EEG-based control of a hand orthosis in a quadriplegic patient Presented at the Int. Conf. on Computers Helping People with Special Needs (Karlruhe, Germany)


Hinterberger T and Baier G 2005 Parametric orchestral sonification of EEG in real time IEEE Multimedia 12 70–9


A multimodal brain-based feedback and communication system

Exp. Brain Res. 154 521–6


Huan N J and Palaniappan R 2005 Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals Proc. 27th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Shanghai, China) pp 633–6

Huggins J E et al. 2003 Electroecographic as the basis for a direct brain interface: Opportunities for improved detection accuracy Proc. 1st IEEE-EMBS Conf. on Neural Engineering (Capri Island, Italy) pp 387–90


Ivanova I, Pfurtscheller G and Andrew C 1995 AI-based classification of single-trial EEG data Proc. 17th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Montreal, Canada) pp 703–4


Kelly S, Burke D, Chazal P d and Reilly R 2002b Parametric models and spectral analysis for classification in brain–computer interfaces Proc. 14th IEEE Int. Conf. on Digital Signal Processing (Santorini, Greece) pp 307–10


Kira K and Rendell L A 1992 The feature selection problem: traditional methods and a new algorithm Proc. of the 10th National Conf. on Artificial Intelligence (San Jose, CA) pp 129–34


Kostov A and Polak M 1997 Prospects of computer access using voluntary modulated EEG signal Proc. ECPD Symposium on Brain & Consciousness (Belgrade, Yugoslavia) pp 233–6


Lee H and Choi S 2002 PCA-based linear dynamical systems for multichannel EEG classification Proc. 9th Int. Conf. on Neural Information Processing (Singapore) pp 745–9


Millan J R 2004 On the need for on-line learning in brain–computer interfaces Proc. Annual Int. Joint Conf. on Neural Networks (Budapest, Hungary)
Millan J, Franze M, Mourino J, Cincotti F and Babiloni F 2002a Relevant EEG features for the classification of spontaneous motor-related tasks Biol. Cybern. 86 89–95
Mourino J, Chiappi S, Jane R and Millan J R 2002 Evolution of the mental states operating a brain–computer interface Proc. of the Int. Federation for Medical and Biological Engineering (Vienna, Austria) pp 600–1
Obermaier B, Guger C, Neuper C and Pfurtscheller G 2001a Hidden markov models for online classification of single trial EEG data Pattern Recognit. Lett. 22 1299–309
Pfurtscheller G and Neuper C 2001 Motor imagery and direct


Peters B O, Pfurtscheller G and Flyvbjerg H 2001 Automatic


Platt J 1991 A resource-allocating network for function interpolation Neurocomputing 2 247–72


Pregenzer M and Pfurtscheller G 1995 Distinction sensitive learning vector quantization (DSLQ)—application as a classifier based feature selection method for a brain–computer interface Proc. 4th Int. Conf. on Artificial Neural Networks (Cambridge, UK) pp 433–6


Qin L, Deng J, Ding L and He B 2004a Motor imagery classification by means of source analysis methods Proc. 26th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (San Francisco, CA) pp 4356–8

Qin L, Ding L and He B 2004b Motor imagery classification by means of source analysis for brain–computer interface applications J. Neural Eng. 1 335–41

Qin L and He B 2005 A wavelet-based time-frequency analysis approach for classification of motor imagery for brain–computer interface applications J. Neural Eng. 2 65–72


Ramachandran V S and Histine W 1998 The perception of phantom limbs Brain 121 1603–30


Reger B D, Fleming K M, Sanguineti V, Alford S and Musa-Itvaldi F A 2000a Connecting brains to robots: the development of a hybrid system for the study of learning in neural tissues Proc. 7th Int. Conf. on Artificial Life (Portland, OR) pp 263–72


Rezek I, Roberts S and Sykacek P 2003 Ensemble coupled hidden markov models for joint characterization of dynamic signals Proc. 9th Int. Workshop on Artificial Intelligence and Statistics (Key West, FL) ed C M Bishop and B J Frey
evoked potential Proc. 1st Int. Conf. on Neural Interface and Control (Wuhan, China) pp 37–9


Wu R C, Liang S F, Lin C T and Hsu C F 2004 Applications of event-related-potential-based brain–computer interface to intelligent transportation systems Proc. IEEE Int. Conf. on Networking, Sensing and Control (Taipei, Taiwan) pp 813–8


Xu W, Guan C, Siong C E, Ranganatha S, Thulasidas M and Wu J 2004b High accuracy classification of EEG signal Proc. 17th Int. Conf. on Pattern Recognition (Cambridge, UK) pp 391–4

Yan W, Sutherland M T, Sanfretello L L and Tang A C 2004 Single-trial classification of ERPs using second-order blind identification (SOBI) Proc. IEEE Int. Conf. on Machine Learning and Cybernetics (Shanghai, China) pp 4246–51


Yu Z, Mason S G and Birch G E 2002 Enhancing the performance of the LF-ASD brain–computer interface Proc. 24th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Houston, TX) pp 2443–4