1 Direct brain-computer communication through EEG signals

GARY N. GARCIA MOLINA, TOURADJ EBRAHIMI,
ULRICH HOFFMANN and JEAN-MARC VESIN
Swiss Federal Institute of Technology
EPFL

1.1 INTRODUCTION

Automatic systems capable of understanding different facets of human communication will be at the heart of human-computer interfaces (HCI) in the near future.

An HCI which is built on the guiding principle: "think and make it happen without any physical effort" is called a brain-computer interface (BCI). Indeed, the "think" part of this principle involves the human brain, "make it happen" implies that an executor is needed (here: a computer) and "without any physical effort" means that a direct interface between the brain and the computer is required.

To make the computer understand what the brain intends to communicate necessitates monitoring the brain activity. Among the possible brain monitoring methods, the scalp recorded electroencephalogram (EEG) constitutes an adequate alternative because of its good time resolution, relative simplicity and noninvasiveness when compared to other methods such as: functional magnetic resonance imaging, positron emission tomography, magnetoencephalography and electrocorticogram. Furthermore, there is clear evidence that observable changes in EEG result from performing given mental activities [1]. In the following, an EEG based BCI will be simply called a BCI.
Current BCIs use the following EEG signals:

Event Related Potentials (ERPs)

ERPs are transient signals which are characterized by a voltage deviation in the EEG and are caused by external stimuli or cognitive processes triggered by external events.

When the user pays attention to a particular stimulus, presented by the BCI an ERP that is time locked with that stimulus appears in her EEG. The changes induced by the ERP in the EEG can be detected by the BCI. Therefore, by focusing her attention to the adequate stimuli, the user can command the BCI.
The advantage of an ERP based BCI resides in the fact that little training is necessary for a new user to gain control of the BCI. Nonetheless, the communication is slow since the user must wait for the relevant stimulus presentation [2] [3].

Steady State Visual Evoked Responses (SSVERs)

SSVERs are elicited by a visual stimulus that is modulated at a fixed frequency. SSVERs are characterized by an increase in the EEG activity at the stimulus frequency. Through feedback, users learn to voluntarily control their SSVER amplitude, whose variations can be detected by the BCI [4].

Slow Cortical Potential Shifts (SCPSs)

SCPSs are shifts of cortical voltage, lasting from a few hundred milliseconds up to several seconds. Users can learn to produce slow cortical amplitude shifts in an electrically positive or negative direction for binary control. This skill can be acquired if the users are provided with a feedback on the course of their SCPS production and if they are positively reinforced for correct responses [5].

Oscillatory sensorimotor activity

The 8-12 Hz and 18-26 Hz activities recorded over the motor cortex exhibit noticeable changes during movement, preparation for movement and imagined movement [6]. Indeed, such activities decrease in the hemisphere that is opposite (contralateral) to the movement and increase in the other hemisphere (ipsilateral). The frequency ranges and the magnitude of the changes are user dependent; if trained, a BCI can detect these changes and react according to a previously established protocol [7] [8].

Spontaneous EEG signals

These signals are recorded during the performance of mental activities other than imagined motor tasks and are not elicited by external stimuli (e.g. mental counting, mental rotation of an object, etc.). The BCI can function with spontaneous signals if the patterns characterizing the corresponding mental activities are learned by the BCI in a training phase [9].

Current BCIs are mainly designed for people with severe motor disabilities in order to provide them with new ways of communication which do not depend on the brain's normal output channels of peripheral nerves and muscles [10] [11]. Furthermore, BCIs can be used as a complement to other HCI devices to enrich the interaction between humans and computers [12] [13].

The achievement of a successful BCI system depends on design factors as well as on the user who should acquire a new skill consisting in controlling her EEG in order to interact with the system. Therefore, human factors such as fatigue, stress or boredom are of great influence.

As mentioned in [11], effective BCIs adapt to each user on three levels.

- First, when a new user accesses the BCI, it adapts to that user's signal characteristics.
- Second, periodic adjustments are necessary to reduce the impact of EEG variations that results from different user's mind states.
- Third, the adaptive capacity of the brain is engaged in the sense that, through feedback the brain activity will modify so as to produce those EEG patterns that best control the BCI.

In this chapter we present a complete BCI system that uses spontaneous and oscillatory EEG signals and implements the three levels of adaptation mentioned above.

1.2 GENERAL CONCEPTS

A BCI can be defined as a communication system in which the messages or commands that the user sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles [11]. The user communicates with the BCI by performing mental activities (MAs) that the BCI is able to recognize because of previous training. An unknown MA triggers a neutral answer from the BCI, the nature of this answer depends on the context in which the BCI is used. By convention we denote by MA_0 an unknown MA.

Our BCI consist of three units, namely EEG acquisition, processing and output (Fig. 1.1). While the user is performing the MAs that control the BCI, her EEG signals are acquired, digitized and sent to a computer where the signals are processed and translated into commands which generate ***actions*** in the output unit. These actions, that can be noticed by the user constitute a feedback that the user exploits so as to modulate her mental activity to make the BCI accomplish her intents.

The correspondence between the patterns present in the EEG, resulting from a mental activity, and the actions generated by the output unit is determined using machine learning techniques. Therefore a training phase, in which the user is asked to perform the MAs that will be used to command the BCI, is carried out.

The training phase is composed of several sessions and results in models associated with each MA, that serve as reference for their respective MA. The models are built through several training sessions in order to take into account different external (environmental) and internal (user related) conditions which can induce variations in the EEG.

The MA models resulting from the training phase are validated in a training with feedback (TRWF) phase in which the user is asked to perform the MAs and

a feedback, telling her if the MA she just performed was successfully recognized (positive feedback) or not (negative feedback) is provided. This implies simultaneous learning of both, the user (who adapts her EEG depending on the feedback) and the BCI (who updates its MA models).

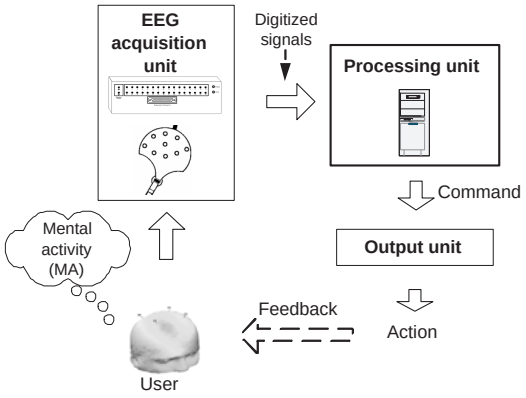

Fig. 1.1 BCI architecture

When the TRWF phase is completed the user can start controlling the BCI by performing the MAs for which the BCI was trained (application phase).

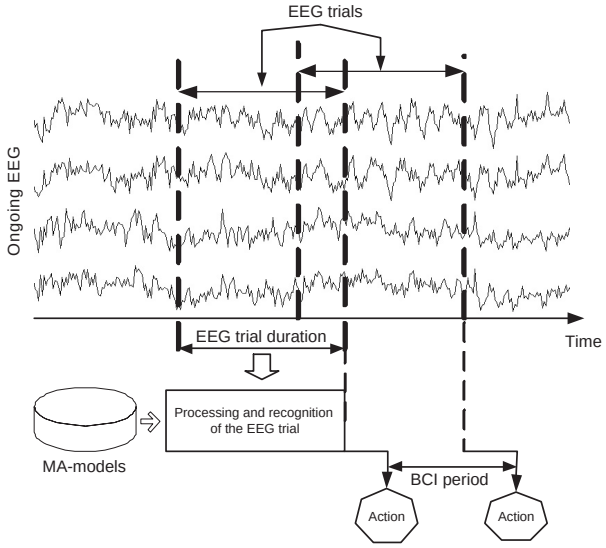

Fig. 1.2 Application phase: Time scheduling

The time schedule of the application phase is depicted in Fig. 1.2. The amount of EEG that is analyzed by the BCI to generate an action, the ***EEG trial duration*** and the interval between two consecutive actions, the ***BCI period*** depend on the type of MAs and on the user. These parameters are determined during the training phase, in which some typical values for these parameters are tested and those values that provide the best results in terms of recognition error are finally selected.

The above mentioned process, namely training, training with feedback and application is carried out several times throughout the use of the BCI in order to satisfy the three levels of adaptation mentioned in Section 1.1.

The extraction of the patterns that characterize a given MA and the model construction are presented in Sections 1.4 and 1.5 respectively.

1.3 EEG ACQUISITION

EEG signals are measured at the scalp by affixing an array of electrodes (Fig. 1.3) positioned according to the 10-20 international system [14] and with reference to digitally linked ears (DLE). DLE referenced voltages are obtained by using the average of voltages at both ear lobes as reference. The ear lobes are selected because they constitute an almost quiet reference. As a matter of fact, they present small influences due to temporal activity [15].

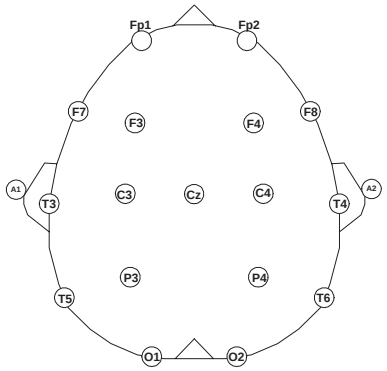

Fig. 1.3 International 10-20 system of electrodes-placement.

The electrode labels correspond to their position with respect to the brain zones, i.e. Fronto-polar(Fp), Frontal (F), Central (C), Temporal (T), Parietal (P) and Occipital (O). Odd indexes are located in the left hemisphere and even ones in the right hemisphere .

If V_e is the voltage at electrode e, V_{A1} and V_{A2} the voltages at left and right ear lobes respectively and if we arbitrarily choose V_{A1} as the physical reference, then the DLE referenced voltage of electrode e is

$$V_e^{DLE} = (V_e - V_{A_1}) - \frac{1}{2}(V_{A_2} - V_{A_1}) = V_e - \frac{1}{2}(V_{A_1} + V_{A_2}) \tag{1.1}$$

1.4 SIGNAL PROCESSING

The purpose of signal processing is twofold - firstly, maximization of the signal-to-noise ratio (preprocessing) and secondly, effective characterization of the MAs (feature extraction).

1.4.1 Preprocessing

Noise sources that affect EEG can be nonneural (50 Hz power-line, eye blinks and muscular activity) or neural (EEG features other than those used for BCI control) [10].

In this research, we centered our analysis on nonneural noise. A 50 Hz notch filter [16] was used to suppress power line interference during the measurements.

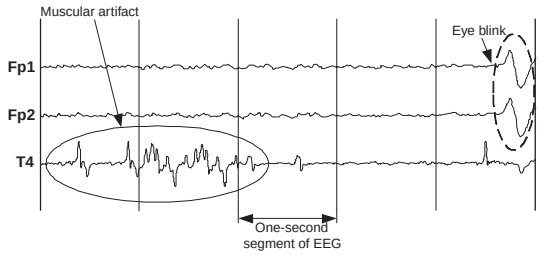

Fig. 1.4 Muscular and eye blink artifacts in EEG.

Any EEG trial containing an eye blink or muscular artifact is considered as unknown to the BCI and consequently elicits the neutral action (Section 1.2).

Eye blinks are characterized by an abrupt increase of amplitude which is more noticeable in electrodes Fp1 and Fp2 (Fig. 1.4). Current methods used to detect such artifacts are basically offline as they were designed for clinical research [17]. We designed an online detection procedure based on the frequency content of the signals recorded at Fp1 and Fp2.

During a calibration procedure the user is asked to blink her eyes a couple of times, the resulting EEG is then segmented into one-second long blocks whose frequency contents, for electrodes Fp1 and Fp2 are calculated from 2 to 40 Hz in two-Hz wide bands. The 38 power values corresponding to a one-second segment are then grouped into a vector (Fp-frequency vector). The complete set of vectors obtained at the end of the calibration procedure, is presented to a k-means algorithm [18] to compute

two centers[1], associated with segments containing ocular artifacts and clean ones. By computing the distance between these centers and the Fp-frequency vector of a segment, one can decide if such a segment contains an eye-blink artifact or not. The accuracy of the centers is then tested in a second calibration procedure where visual feedback is provided whenever the BCI detects an eye blink.

Muscular artifacts are characterized by high amplitude high frequency bursts in EEG (Fig. 1.4) which are more noticeable at electrodes F7, F8, T3, C3, C4 and T4. The centers characterizing the EEG segments containing muscular artifacts with respect to clean ones are determined in the same way as for ocular artifacts. As for ocular artifacts, the accuracy of the centers is tested in the second calibration procedure where a visual feedback (different from that of eye blinks) is provided when the BCI detects a muscular artifact.

The calibration procedure which results in references for the detection of eye blinks and muscular artifacts is repeated before each recording session to take into account the particular environmental conditions. The total duration of the calibration is usually less than ten minutes.

1.4.2 Feature extraction

An EEG trial X belongs to[2] $\mathbb{R}^{N_e \mathrm{x} T_{trial}}$ where N_e is the number of electrodes and T_{trial} is the number of samples in the EEG trial. The rows of X: $X(1), ..., X(N_e)$ contain the signals recorded at different electrodes.

The goal of feature extraction is to map the EEG trial space, $\mathbb{R}^{N_e \mathrm{x} T_{trial}}$ into a feature space, $\mathbb{R}^{D_f}$ (D_f is the dimensionality of the feature space) which is suitable for the discrimination of EEG trials resulting from the performance of different MAs.

The mapping function is usually determined using the physiological knowledge of the MAs that are used to control the BCI. Common mappings result from time-space [19] [20], frequency (parametric [21] [22] and nonparametric [6]) and time-frequency analysis of EEG trials [23].

In this study, the feature vectors are obtained using a set of space-frequency filters which are determined for each MA. Below we present the method, based on the approach presented in [20] used to determine those filters for a particular MA (targeted MA).

Consider a set of EEG trials[3], Ξ which is labeled (i.e. the MA associated with each element in Ξ is known), the sub-indices i and j are used for EEG trials belonging (targets) and not belonging (nontargets) to the targeted MA respectively. Then,

$$\Xi = \{ X_i | \, 1 \leqslant i \leqslant N_I \} \cup \{ X_j | \, 1 \leqslant j \leqslant N_J \}$$

[1]One of the two centers determined by the k-means algorithm is generally associated with clean segments whereas the other one with perturbed segments. Indeed, the frequency content of eye blinks and muscular artifacts is different enough from that of clean EEG

[2]The space of real matrices of dimension N_e x T_{trial}

[3]The temporal mean is removed from each EEG trial

where N_I and N_J are respectively, the number of targets and nontargets.

The mean autocorrelation matrices of targets and nontargets are [4]

$$R_T = \frac{1}{N_I} \sum_{i=1}^{N_I} X_i X_i^t = \left\langle X_i X_i^t \right\rangle \tag{1.2}$$

$$R_{NT} = \frac{1}{N_J} \sum_{j=1}^{N_J} X_j X_j^t = \left\langle X_j X_j^t \right\rangle \tag{1.3}$$

The sum of the autocorrelation matrices, being positive semi-definite can be diagonalized and its eigenvalues are positive, i.e. the diagonal elements of D_R in Eq. 1.4 are positive.

$$R_T + R_{NT} = V D_R V^t \tag{1.4}$$

The following matrices are defined[5]: $S_T = D_R^{-\frac{1}{2}} V^t R_T V D_R^{-\frac{1}{2}}$ and $S_{NT} = D_R^{-\frac{1}{2}} V^t R_{NT} V D_R^{-\frac{1}{2}}$. One can easily check that $S_T + S_{NT} = I$, where I is the identity matrix.

The matrices S_T and S_{NT} can be diagonalized, have the same eigenvectors and their eigenvalues are in $]0;1[$. Then, $S_T = U D_S U^t$ and $S_{NT} = U(I - D_S)U^t$. From the last equalities it can be inferred that

$$D_S = U^t D_R^{-\frac{1}{2}} V^t R_T V D_R^{-\frac{1}{2}} U = P \left\langle X_i X_i^t \right\rangle P^t = \left\langle P X_i (P X_i)^t \right\rangle = \left\langle Z_i Z_i^t \right\rangle \tag{1.5}$$

$$I - D_S = U^t D_R^{-\frac{1}{2}} V^t R_{NT} V D_R^{-\frac{1}{2}} U = P \left\langle X_j X_j^t \right\rangle P^t = \left\langle P X_j (P X_j)^t \right\rangle = \left\langle Z_j Z_j^t \right\rangle \tag{1.6}$$

where $P = U^t D_R^{-\frac{1}{2}} V^t$, $Z_i = P X_i$ and $Z_j = P X_j$. Z_i and Z_j can be seen as transformed EEG trials which result from the projection, by P of X_i and X_j.

The elements of the n^{th} row of P are the coefficients of the n^{th} projection. The n^{th} row of Z_i is $Z_i(n) = \sum_{k=1}^{N_e} P_{n,k} X_i(k)$.

If the n^{th} eigenvalue (λ_n) of S_T is large, the eigenvalue (1-λ_n) of S_{NT}, associated with the same eigenvector is small. The latter amounts to say that, if $\langle Z_i(n) Z_i^t(n) \rangle$ is large then $\left\langle Z_j(n) Z_j^t(n) \right\rangle$ is small and vice versa. It follows that the mean energy of the projections whose eigenvalues are the largest or the smallest, constitute suitable features to classify an EEG trial as belonging to the targeted MA or not.

Only some rows of P (relevant components) need to be selected to build a feature vector. As a matter of fact the n^{th} row is selected if the corresponding eigenvalue λ_n satisfies:

$$\min(\lambda_n, 1 - \lambda_n) \leq \rho \tag{1.7}$$

[4]The upper index ^{t} stands for the transpose operator

[5]The notation $D^{-\frac{1}{2}}$ means that the square root of the inverse of each element in the diagonal of D is taken

where ρ is a parameter whose typical values are in $[0.1; 0.2]$.

As it can be easily noticed, the projection matrix P was built using the information in the time domain only. In order to take into account the information encoded in EEG frequencies, projection matrices are built for different frequency bands. The frequency bands can be the classical four bands used in medical research, namely δ (1-4 Hz), θ (4-7 Hz), α (8-13 Hz) and β (15-20 Hz). Since the choice of frequencies bands depends on the MA and on the user we did not use the typical bands, instead we considered 4-Hz bands from 1 to 41 Hz, i.e. 10 frequency bands.

In order to prevent any phase distortion of the signals when filtered into the above mentioned bands, we used linear-phase finite impulse response filters [24].

The feature vector x associated with an EEG trial X is computed in the following way: X is filtered into the 10 frequency bands, then each of the filtered trials is transformed applying the respective projection matrix. Finally, the mean energies corresponding to each projection compose the feature vector. The dimension of the feature vector D_f is equal to $\sum_{b=1}^{10} \eta_b$ where η_b is the number of relevant components associated with the b^{th} frequency band.

Since the determination of the feature space is optimized for a given MA and user, this procedure implements the first level of adaptation (see Section 1.1).

1.5 LEARNING MENTAL ACTIVITIES

We consider a BCI that is commanded by N_{MA} MAs, each of them associated with an action. The set of actions that the BCI is able to perform is noted as $\{A_0, A_1, ..., A_{N_{MA}}\}$ where A_0 is the neutral action that corresponds to an unknown MA (Section 1.2).

BCI functioning is modeled as depicted in Fig. 1.5. Each BCI period the MA generator (which is composed of the user's brain, the acquisition device and the feature extractor) produces a vector[6] $x(m)$ drawn from a probability density function (PDF) $p(x(m))$ which depends on on the action at time $(m-1)$, $A(m-1)$ and on extra-system factors (Ext) that can be external (noise and environmental conditions) and internal (user's state of mind).

The vector $x(m)$ is then processed by the recognition algorithms whose result is the vector $L(m)$ containing N_{MA} scores that measure the likelihood that $x(m)$ corresponds to each MA.

The taken action can be noticed by the MA generator who adapts its behavior consequently (feedback). Additionally, the current action $A(m)$ depends on $L(m)$ and the last τ actions $A(m-1), ..., A(m-\tau)$, this dependency is presented in Section 1.5.2.

In order to produce the desired actions, the MA generator should adapt its PDF through reinforcement learning. While external factors can be controlled, internal ones are harder to control even by the user herself. It becomes then necessary to

[6] the feature vector whose computation is explained in Section 1.4.2

introduce a mechanism that allows the recognition algorithms to adapt to the user's current PDF, this mechanism is presented in Section 1.5.1.

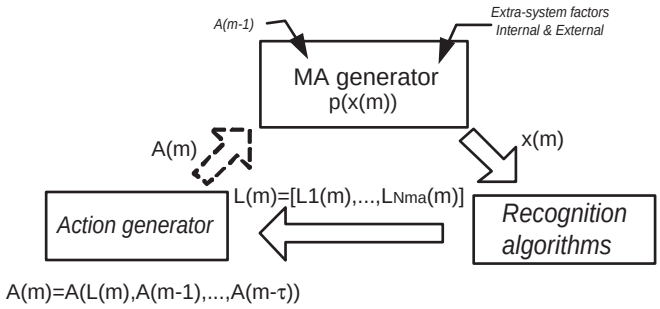

Fig. 1.5 BCI model.

1.5.1 Dynamic characterization of the PDF associated with an MA

The parameters of recognition algorithms need to be flexible in order to dynamically characterize the PDF associated with each MA under different extra-system conditions. The characterization process is carried out in several training sessions and it is continuously updated under the assumption that the extra-system conditions remain constant during each session.

The first characterization of MA_q, that is denoted by $\Phi_{MA_q}(0)$ is obtained after a training without feedback session (open loop) in which the user is asked, by a visual cue, to perform the MA_q while the produced feature vectors are stored to compute $\Phi_{MA_q}(0)$ (Fig. 1.6).

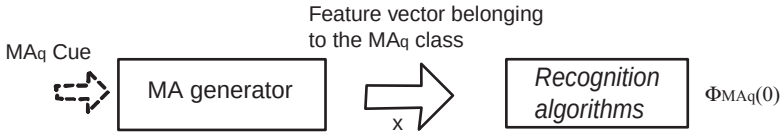

Fig. 1.6 Training without feedback (open loop).

Subsequent characterizations, $\Phi_{MA_q}(k)$ for $k > 0$, are obtained after training with feedback sessions (closed loop), in which the user is asked to perform MA_q and the system uses the last characterization to tell the user if the current MA was successfully identified (Fig. 1.7).

An effective way to integrate the current characterization with the precedent one consists in considering Φ_{MA_q} as a set composed of those vectors that define the boundaries of the PDF_q. Such a set is built using the approach presented in [25] for novelty detection. In the following we briefly present this method for the first

training session (without feedback) and we generalize it for the subsequent training-with-feedback sessions.

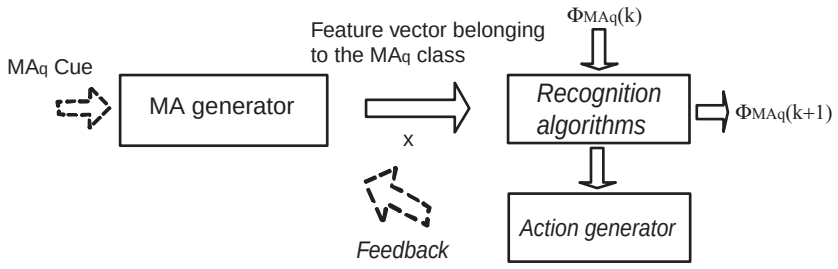

Fig. 1.7 Training with feedback (closed loop).

A transformation F_H that maps the feature vector space, $\mathbb{R}^{D_f}$ into a high (possibly infinite) dimensional space H, is determined through a Kernel function that is the internal product in H and can be expressed in terms of dot products in $\mathbb{R}^{D_f}$. The advantage of this mapping resides in the fact that simple data representations in H can model complex distributions in $\mathbb{R}^{D_f}$. The choice of the Kernel function depends on the application, here we used a Gaussian Kernel as suggested in [25], which makes H of infinite dimension [26].
If x and y belong to $\mathbb{R}^{D_f}$, the Gaussian Kernel is defined as follows:

$$K(x,y) = \exp\left(-\frac{x\cdot x - 2x\cdot y + y\cdot y}{\sigma^2}\right) \tag{1.8}$$

where "$\cdot$" is the dot product of x and y, and σ is the width parameter that is determined by cross-validation.

Let Υ_q be the set containing the feature vectors belonging to class MA_q, $\bar{\Upsilon}_q$ the complement of Υ_q and $F_H(\Upsilon_q)$, $F_H(\bar{\Upsilon}_q)$ the images of Υ_q and $\bar{\Upsilon}_q$ under F_H.

We assume that elements in $F_H(\Upsilon_q)$ are tightly bounded by a sphere of radius R and centered in $\Omega(\in H)$ while the vectors in $F_H(\bar{\Upsilon}_q)$ are outside that sphere. Therefore, the value of R and the center Ω are found by solving the optimization problem consisting in minimizing R.

In a real application, some of the feature vectors can be in the wrong set (training errors) because of perturbations or user's lack of concentration during the measurements. In order to take the training errors into account and to avoid overfitting it is necessary to introduce slack variables. Then, the optimization problem becomes:

$$\text{minimize}\left(R^2 + C_1\sum_i \xi_i^{(x)} + C_2\sum_j \xi_i^{(y)}\right) \tag{1.9}$$

under the constraints
$\|u_i - \Omega\|^2 \leqslant R^2 + \xi_i^{(x)}\ ;\ u_i = F_H(x_i)\ ;\ x_i \in \Upsilon_q$

$\|v_j - \Omega\|^2 > R^2 + \xi_j^{(y)}$; $v_j = F_H(y_j)$; $y_j \in \bar{\Upsilon}_q$
$\xi_i^{(x)} \geqslant 0$; $\xi_j^{(y)} \geqslant 0$
where C_1 and C_2 are penalization constants whose values are determined as explained later in the text. The index i is used for elements in Υ_q and $F_H(\Upsilon_q)$, and the index j for elements in $\bar{\Upsilon}_q$ and $F_H(\bar{\Upsilon}_q)$

Problems of this kind are handled by introducing Lagrange multipliers $\alpha_i^{(x)} \geqslant 0$, $\alpha_j^{(y)} \geqslant 0$, $\gamma_i^{(x)} \geqslant 0$, $\gamma_j^{(y)} \geqslant 0$ and a Lagrangian.

$$\Lambda_P = R^2 + C_1 \sum_i \xi_i^{(x)} + C_2 \sum_j \xi_j^{(y)} - \sum_i \gamma_i^{(x)} \xi_i^{(x)} - \sum_j \gamma_j^{(y)} \xi_j^{(y)} \tag{1.10}$$
$$- \sum_i \alpha_i^{(x)} \left(R^2 + \xi_i^{(x)} - \|u_i - \Omega\|^2 \right) - \sum_j \alpha_j^{(y)} \left(\|v_j - \Omega\|^2 - R^2 + \xi_j^{(y)} \right)$$

The Lagrangian Λ_P has to be minimized with respect to the primal variables: R, $\xi_i^{(x)}$, $\xi_j^{(y)}$, Ω and maximized with respect to the dual variables: $\alpha_i^{(x)}$, $\alpha_j^{(y)}$, $\gamma_i^{(x)}$, $\gamma_j^{(y)}$ (in other words, a saddle point has to be found).

Taking derivatives of Λ_P with respect to the primal variables and setting them to zero leads to the following results:

$$\sum_i \alpha_i^{(x)} - \sum_j \alpha_j^{(y)} = 1 \tag{1.11}$$

$$\Omega = \sum_i \alpha_i^{(x)} u_i - \sum_j \alpha_j^{(y)} v_j \tag{1.12}$$

$$C_1 - \gamma_i^{(x)} - \alpha_i^{(x)} = 0 \Rightarrow 0 \leqslant \alpha_i^{(x)} \leqslant C_1 \tag{1.13}$$

$$C_2 - \gamma_j^{(y)} - \alpha_j^{(y)} = 0 \Rightarrow 0 \leqslant \alpha_j^{(y)} \leqslant C_2 \tag{1.14}$$

when Equations 1.11, 1.12, 1.13 and 1.14 are substituted in 1.10 one obtains the dual problem

$$\Lambda_D = 1 - \sum_{i1,i2} \alpha_{i1}^{(x)} \alpha_{i2}^{(x)} K\left(x_{i1}, x_{i2}\right) \tag{1.15}$$
$$+2 \sum_{i,j} \alpha_i^{(x)} \alpha_j^{(y)} K\left(x_i, y_j\right) - \sum_{j1,j2} \alpha_{j1}^{(y)} \alpha_{j2}^{(y)} K\left(y_{j1}, y_{j2}\right) = 1 - \Omega \cdot \Omega$$

where
$K\left(x_{i1}, x_{i2}\right) = u_{i1} \cdot u_{i2}$, $u_{i1} = F_H\left(x_{i1}\right)$ and $u_{i2} = F_H\left(x_{i2}\right)$
$K\left(x_i, y_j\right) = u_i \cdot v_j$, $u_i = F_H\left(x_i\right)$ and $v_j = F_H\left(y_j\right)$
$K\left(y_{j1}, y_{j2}\right) = v_{j1} \cdot v_{j2}$, $v_{j1} = F_H\left(y_{j1}\right)$ and $v_{j2} = F_H\left(y_{j2}\right)$

The dual Lagrangian, Λ_D should be maximized with respect to the $\alpha_i^{(x)}$'s and the $\alpha_j^{(y)}$'s. The last amounts to say that $\Omega \cdot \Omega$ should be minimized (Eq. 1.15); therefore, the sphere in the space H is close to the origin.

Furthermore, the Karush-Kuhn-Tucker (KKT) conditions [25] imply that at the optimum Eqs. 1.16 and 1.17 hold.

$$\alpha_i^{(x)} \left(R^2 + \xi_i^{(x)} - \|u_i - \Omega\|^2 \right) = 0 \tag{1.16}$$

$$\alpha_j^{(y)} \left(\|v_j - \Omega\|^2 - R^2 + \xi_j^{(y)} \right) = 0 \tag{1.17}$$

The value of $\alpha_i^{(x)}$ determines the position of u_i with respect to the sphere.

$$\text{if } \begin{cases} \alpha_i^{(x)} = 0, & u_i \text{ is inside the sphere} \\ 0 < \alpha_i^{(x)} < C_1, & u_i \text{ is in the boundary of the sphere} \\ \alpha_i^{(x)} = C_1, & u_i \text{ is outside the sphere (false negative)} \end{cases}$$

Reciprocal results are found for the $\alpha_j^{(y)}$ and their corresponding v_j. Because of the KKT conditions the solution of the problem in Eq. 1.15 is sparse, i.e. the majority of α's are equal to zero at the optimum value of Λ_D. The characterization Φ_{MA_q} is composed of those vectors[7] (in the feature space $\mathbb{R}^{D_f}$) whose $alpha$'s are different from 0.

The radius of the sphere R is determined by computing the distance between the center Ω (Eq. 1.12) and any of the u_{i*}'s such that its corresponding α_{i*} is in $]0; C_1[$, i.e. u_{i*} is at the boundary of the sphere (Eq. 1.18).

$$\begin{aligned} R^2 &= \|u_{i*} - \Omega\|^2 \\ &= 2 \left(1 - \sum_i \alpha_i^{(x)} u_i \cdot u_{i*} + \sum_j \alpha_j^{(y)} v_j \cdot u_{i*} \right) - \Lambda_D \\ &= 2 \left(1 - \sum_{i | x_i \in \Phi_{MA_q}} \alpha_i^{(x)} K(x_i, x_{i*}) + \sum_{j | y_j \in \Phi_{MA_q}} \alpha_j^{(y)} K(y_j, x_{i*}) \right) - \Lambda_D \end{aligned} \tag{1.18}$$

From the above results and Eq. 1.11 it can be stated that

$$C_1 \leqslant \frac{1}{\text{number of } u\text{'s outside the sphere (false negatives)}} \tag{1.19}$$

$$C_2 \leqslant \frac{1}{\text{number of } v\text{'s inside the sphere (false positives)}} \tag{1.20}$$

In the first session, C_2 was determined using the equality in 1.20, under the assumption that the false positive rate was equal to the rate of nondetected artifacts, and C_1 was determined using 1.19 under the assumption that the false negatives rate was in the order of 1%.

[7]The vectors whose α is different from zero are usually called support vectors

In subsequent sessions, since reference models exist to provide feedback, the values of C_1 and C_2 were set depending on the number of false negative and false positive feedbacks respectively, according to Eqs. 1.21 and 1.22.

$$C_1 = \frac{1}{(1-\eta)N_f^-} \tag{1.21}$$

$$C_2 = \frac{1}{(1-\eta)N_f^+} \tag{1.22}$$

where N_f^- and N_f^+ are the numbers of false negative and false positive feedbacks respectively, and η is a learning factor that characterizes the amount of novelty in a given training with feedback session. Typical values for η are in $[0; 0.3[$.

The characterization at the end of the k^{th} session, $\Phi_{MA_q}(k)$ is composed of the vectors that are in the boundary of the sphere. In the $(k+1)^{th}$ session, the new characterization, $\Phi_{MA_q}(k+1)$ is obtained adding the characterization set of the k^{th} session to the new training vectors and solving the optimization problem (Eqs. 1.9 to 1.17).

It can be shown that if the boundary of the PDF_q changes, the characterization adapts to that change. In this way we can achieve the second level of adaptation (see Section 1.1).

As feedback is provided during the training sessions, (except the first one) the brain can adapt itself in order to produce the right MAs. In this way, the third level of adaptation is achieved (see Section 1.1).

The model of MA_q at the end of the k^{th} session is composed of $\Phi_{MA_q}(k)$, the values of the α's associated with the elements in $\Phi_{MA_q}(k)$ and the radius of the sphere corresponding to the MA_q at the end of the k^{th} session: $R_q(k)$.

For an unknown vector, $x_?$ its score, $\varsigma_q(x_?)$ with respect to the MA_q characterization is computed as the ratio between the distance of $F_H(x_?)$ to the center of the sphere associated with MA_q, and the radius R_q.

$$\begin{aligned}\varsigma_q(x_?) &= \frac{\|F_H(x_?) - \Omega\|^2}{R_q^2} = \frac{2\left(1 - F_H(x_?)\cdot\Omega\right)}{R_q^2} \\ &= \frac{2\left(1 - \sum\limits_{i|x_i\in\Phi_{MA_q}} \alpha_i^{(x)} K(x_i, x_?) + \sum\limits_{j|y_j\in\Phi_{MA_q}} \alpha_j^{(y)} K(y_j, x_?)\right)}{R_q^2}\end{aligned} \tag{1.23}$$

The smaller $\varsigma_q(x_?)$ the larger the likelihood that $x_?$ was produced during the performance of $(MA)_q$.

If the score is larger than 1, $F_H(x_?)$ is outside the sphere, thus the likelihood of $x_?$ being produced during the performance of MA_q is zero.

When a feature vector $x(m)$ is presented to the recognition algorithms (see Fig. 1.5) the result is a vector:

$L(m) = \left[\ L_1(m) \quad ... \quad L_{N_{MA}}\ (m)\right]$

where $L_q = \max(-\log(\varsigma_q(x(m))), 0)$ indicates the likelihood that $x(m)$ was produced during the performance of MA_q.

1.5.2 Action generation

The action generator is a state machine whose $(N_{MA} + 1)$ states correspond to the N_{MA} actions and the neutral one. In the *ideal scheme* the action at time m, depends only on $L(m)$ according to:

$$\text{Ideal scheme criteria} \begin{cases} \text{if } L_q(m) = 0, \forall q \text{ then } A(m) = A_0 \\ \text{otherwise } A(m) = A_{\arg\max_q(L_q(m))} \end{cases}$$

Since the recognition of MAs is not perfect, the action generator is not allowed to abruptly change its state. Instead, a dynamic transition mechanism is introduced that depends on the recognition error. As the user performance increases,i.e. the recognition error rate decreases, the mechanism converges to the ideal scheme.

Consider that at time m the BCI is executing the action A_{q1} and, according to the ideal scheme criteria $A(m+1) = A_{q2}$. The transition towards A_{q2} is allowed if, under the ideal scheme criteria, $A(m+1) = A(m+2) = ... = A(m+\tau_{q2q1}) = A_{q2}$ otherwise the BCI keeps executing A_{q1}.

The value of τ is determined such that

$$\tau_{q2q1} \geqslant \frac{\log \epsilon}{\log p_{q2q1}} \tag{1.24}$$

where p_{q2q1} is the probability that a vector x, generated during the performance of MA_{q1} was recognized as belonging to the MA_{q2} class and ϵ is a constant that characterizes the degree of confidence on the transition from A_{q1} to A_{q2}, e.g. if 99% confidence is required then, the value of ϵ is set to 0.01.

One can easily verify that $\tau_{q2q1} \xrightarrow[p_{q2q1} \to 0]{} 0$ (Eq. 1.24), i.e. the smaller the probability of confusion between MA_{q2} and MA_{q1} the shorter the transition duration.

Therefore, as the user performance improves, the transition scheme evolves towards the ideal scheme.

1.6 EXPERIMENTAL METHODS AND PROTOCOL

Two male and healthy users (S1, S2) 25 and 27 years old participated in twelve sessions of about 30 minutes. The activities in these sessions were organized as follows: training without feedback in the first three sessions, free control of an object on a computer screen in the sixth and the last two sessions, and training with feedback in the remaining ones. The users were comfortably sitting in an armchair and placed in front of a computer screen. The experimentation room was quiet and slightly illuminated.

The signals from electrodes Fp1, Fp2, F7, F3, F4, F8, T3, C3, C4, T4, A2, T5, P3, P4, T6, O1 and O2 were recorded with reference to A1, and the DLE referenced voltages were then calculated (see Section 1.3).

Each session started with the calibration procedure in which the parameters for the rejection of artifacts were set (see Section 1.4.1).

The duration of EEG trials and of the BCI period were set at the end of the first session as explained in Section 1.2.

The users were asked to perform three MAs to control the BCI, namely imagined movement of left and right index finger (MA_1 and MA_2 respectively), and successive mental subtractions by 3 (MA_3). Visual cues were used to indicate which MA to perform, for MA_1 and MA_2 a pointing object was displayed in the screen and for MA_3 a random 3-digit number was displayed to indicate the starting value for the successive subtractions.

1.6.1 Protocol of a training without feedback session

The first ten minutes were spent in the calibration procedure and the remaining time was divided into two five minutes measurements separated by about two minutes break.

Each five-minute measurement time (active thinking period) was organized as follows (Fig. 1.8): a visual cue is displayed indicating the MA that should be executed during the next five seconds. At the end of this interval a break cue appears, telling the user that she should perform MA_0, i.e. any mental activity but the three ones that are trained. The duration of the break was randomly chosen between 4,5 or 6 seconds. This process is repeated 30 times (10 times, in a random order, per MA).

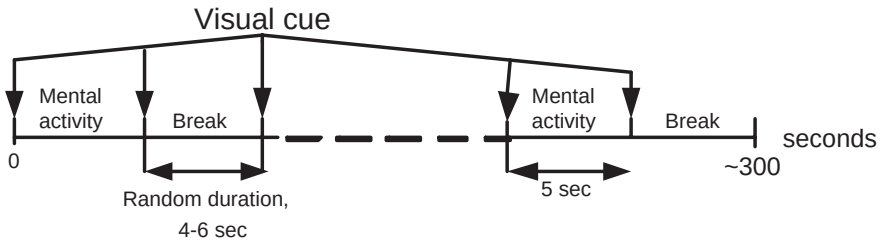

Fig. 1.8 Active thinking period corresponding to a training without feedback session

1.6.2 Protocol of a training with feedback session

The general time organization was identical to that of the training without feedback sessions.

The difference between the two kinds of training sessions resides in the fact that during the periods of active thinking a feedback is provided to the user. When the

feedback is positive, the action to which the current MA is associated is executed by the BCI and the neutral action is executed if the feedback is negative.

The interval between two successive feedbacks is equal to the BCI period which was determined after the first training (without feedback) session.

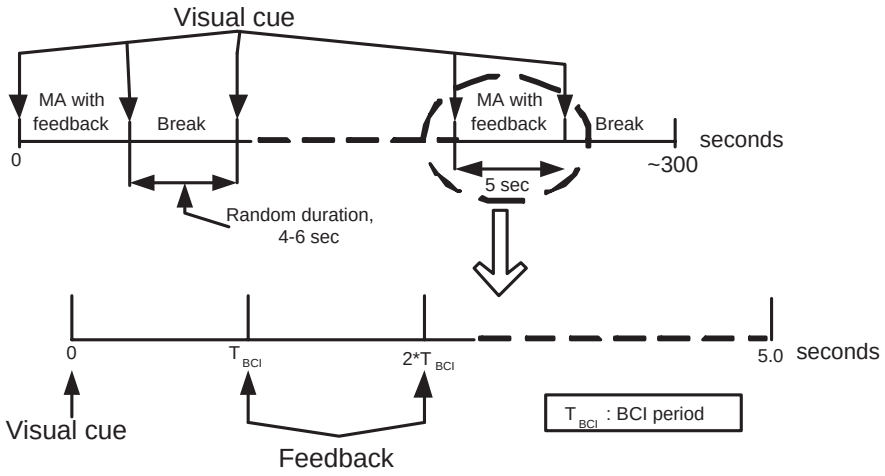

Fig. 1.9 Active thinking period corresponding to a training with feedback session

1.6.3 Protocol of a free control session

Two objects are placed on the computer screen, one of them is the target and the other the mobile object that is controlled by the user thoughts in order to reach the target object. The distance between the two objects is the same for each free control session, the relative position of the objects is nevertheless changed to avoid habituation. The "game" was repeated three times per session.

1.7 RESULTS AND DISCUSSION

Preprocessing The results of the preprocessing procedure (Section 1.4.1) are reported in Fig. 1.10 in terms of the rate of clean EEG segments that were detected as artifacts (false artifacts) and the rate of segments containing artifacts accepted as clean (false clean segments).

For our purposes it is more important to have a low *false clean rate* in order to accurately describe each MA. As it can be seen in Fig. 1.10 this rate is below 3% for both users.

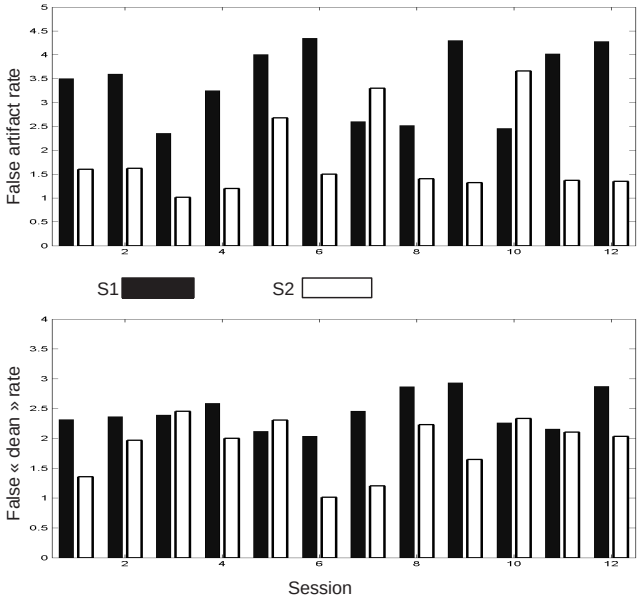

Fig. 1.10 Top: False artifact detection rate (%). Bottom: Rate of artifacts detected as clean signals (%).

Training without feedback - Sessions: one to three For the training sessions without feedback, different durations of *EEG trial duration/BCI period* were considered [8]. These values were evaluated according to the recognition error rate which was determined using the ideal scheme criteria (Section 1.5.2) and taking about 60% of the EEG trials for training the MA models[9] and the remaining trials to determine the recognition error. In order to have more reliable results, the recognition error was averaged over twenty different choices of the training and testing set.

The results, reported in Table 1.1 show that, the longer the duration of the EEG trial the better is the recognition accuracy. For both users, the EEG trial duration and the BCI period were set to 2 seconds. Even if the values 3/2 and 3/3 provided slightly better results we preferred to avoid the overlap (in the 3/2 case) and an excessive delay in the feedback (in the 3/3 case).

User	$0.5/0.5$	$1.0/0.5$	$2.0/1.0$	$2.0/2.0$	$3.0/2.0$	$3.0/3.0$
S1	40.54	35.33	31.25	30.95	30.45	29.22
S2	42.57	36.33	33.63	31.82	30.97	31.57

Table 1.1 Global recognition error rate (%), in the first three sessions for different durations (in seconds) of EEG trial/BCI period

As explained in the feature extraction part (Section 1.4.2) the number of relevant spatial filters for each frequency band should be determined according to Eq. 1.7, in this study we set $\rho = 0.15$.
The obtained results are reported in Table 1.2. For MA_1 and MA_2 the frequency bands from 9 to 25 Hz have the largest numbers of relevant spatial filters while for MA_3 the band that has the largest number of relevant spatial filters is user dependent (13-17 Hz for S1 and 5-9 Hz for S2). Furthermore, the dimension of the feature space depends on the MA and on the user.

In Fig. 1.11 we present the main spatial filter associated with the most important frequency band (the one that has the largest number of relevant filters) for each MA and user.

Training with feedback - Sessions: four to five and seven to ten In Tables 1.3 and 1.4 we reported the evolution, over the sessions of the transition parameters τ for user S1 and S2 respectively. τ_{q2q1} corresponds to the transition from A_{q1} to A_{q2}. The value of ϵ in Eq. 1.24 was set to 0.01.

For both subjects the mean transition duration is two BCI periods, namely 4 seconds. One can see that the transition duration follows smoothly the variations of the error rates (see Fig. 1.12 and 1.13) . Indeed, τ was modified if the user performance was significantly changed. This behavior is due to the log function in

[8] When the EEG trial duration is longer than the BCI period an overlap between the trials exist.
[9] Every model was built using the same number of trials

Frequency band	S1 MA1	S1 MA2	S1 MA3	S2 MA1	S2 MA2	S2 MA3
1-5 Hz	0	2	2	2	1	3
5-9 Hz	2	1	3	1	3	5
9-13 Hz	5	2	3	2	2	2
13-17 Hz	4	6	5	4	3	1
17-21 Hz	6	3	4	3	4	3
21-25 Hz	5	2	1	0	1	2
25-29 Hz	1	3	2	1	5	1
29-33 Hz	2	2	4	2	2	2
33-37 Hz	1	1	3	2	3	3
37-41 Hz	0	0	2	1	0	2
Feature space dimension	26	22	29	18	24	24

Table 1.2 Number of relevant spatial filters for each frequency band, MA and user ($\rho = 0.15$ see Section 1.4.2)

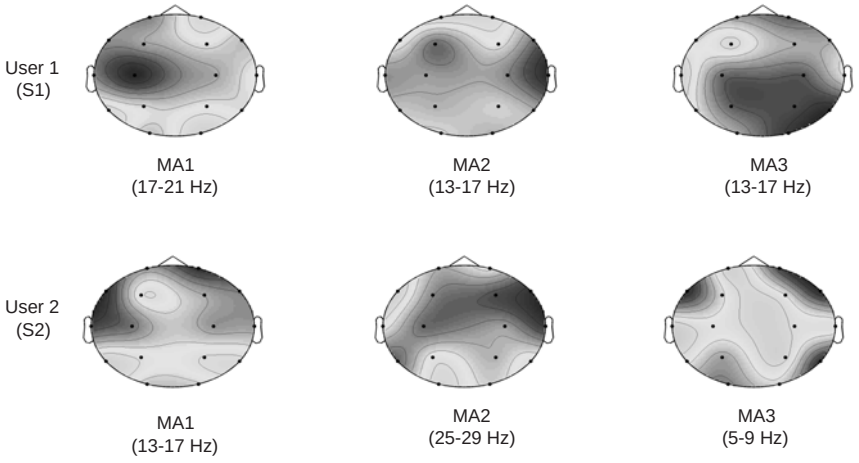

Fig. 1.11 Representation of the space filters associated with the most important frequency bands for each MA. Top: User1, Bottom: User 2

the definition of τ (Eq. 1.24). Thus, if the user improved her performance, the BCI rewarded her by shortening the duration of the transition.

Figures 1.12 and 1.13 depict the evolution of the users performance over the sessions, where feedback was provided. In the horizontal axis we reported the true positives rate for each MA, i.e. the percentage of trials that were correctly identified as belonging to their respective MA class and in the vertical axis the false positives rate, i.e. the percentage of trials that were wrongly identified as belonging to each MA class. The numbers above the markers indicate the number of the session.
As it can be observed, the trends indicate that the true positive rate was globally improved for both users. Indeed, the true positive rates are larger in the last session (were feedback was provided) for each MA. Nonetheless, a more erratic evolution appears for the false positive rate, for S1 by instance the false positive rates for MA_3 and MA_1 are worst in the last session and for S2, MA_1 presents the same particularity.

The way in which the false positive rate is handled depends on the application, in our case it was more important to have a high true positive rate so as to avoid user frustration that can appear if the false negative rate is high. For other applications, such as driving a wheelchair it could be more important to have a small false positive rate.

τ	4^{th} session	5^{th} session	7^{th} session	8^{th} session	9^{th} session	10^{th} session
τ_{01}	4	3	3	3	2	3
τ_{02}	3	3	2	2	3	2
τ_{03}	2	3	3	2	3	2
τ_{10}	3	3	3	2	3	3
τ_{12}	3	3	3	2	2	3
τ_{13}	2	2	2	2	2	3
τ_{20}	2	2	2	2	2	2
τ_{21}	3	2	2	2	2	2
τ_{23}	2	2	2	2	2	2
τ_{30}	2	2	2	2	2	2
τ_{31}	0	2	2	2	2	2
τ_{32}	2	3	2	2	2	2

Table 1.3 Evolution of the transition parameter τ for the training without feedback sessions for user S1

Free control - Sessions: six, eleven and twelve In Tables 1.5 and 1.6 we report the time, in seconds that each user took to reach the target object. The shortest possible time to reach the target was 80 seconds, i.e. 40 EEG trials correctly recognized. If the object is moved outside the limits of the screen the target was considered as missed.

τ	4^{th} session	5^{th} session	7^{th} session	8^{th} session	9^{th} session	10^{th} session
τ_{01}	3	4	3	2	2	3
τ_{02}	2	0	2	2	2	0
τ_{03}	2	3	3	0	2	2
τ_{10}	3	3	3	3	3	3
τ_{12}	2	3	2	2	2	2
τ_{13}	2	2	2	2	2	2
τ_{20}	2	2	2	3	2	2
τ_{21}	3	2	3	3	2	2
τ_{23}	3	2	2	3	2	2
τ_{30}	3	2	2	2	2	2
τ_{31}	3	2	2	3	2	2
τ_{32}	3	2	2	2	3	2

Table 1.4 Evolution of the transition parameter τ for the training without feedback sessions for user S2

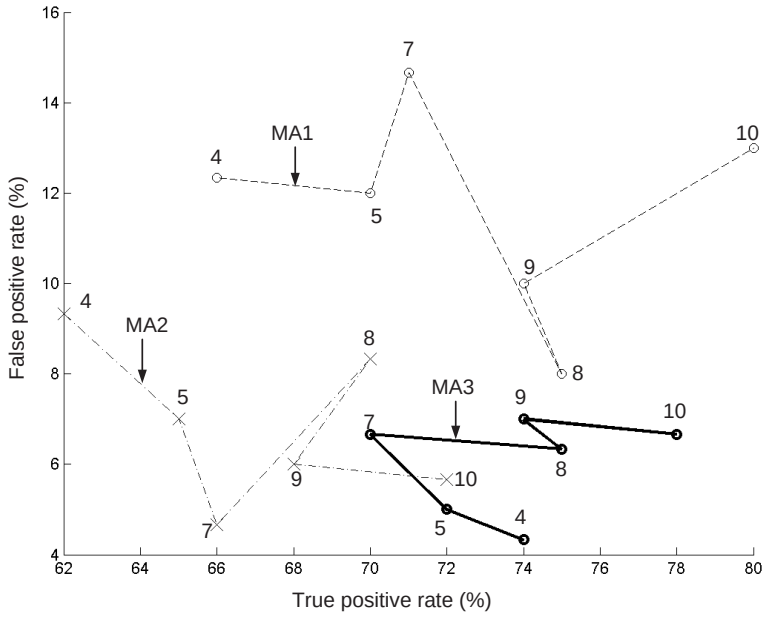

Fig. 1.12 Evolution, for user S1 of the true positive and false positive rate for each MA. The numbers above the markers indicate the session number

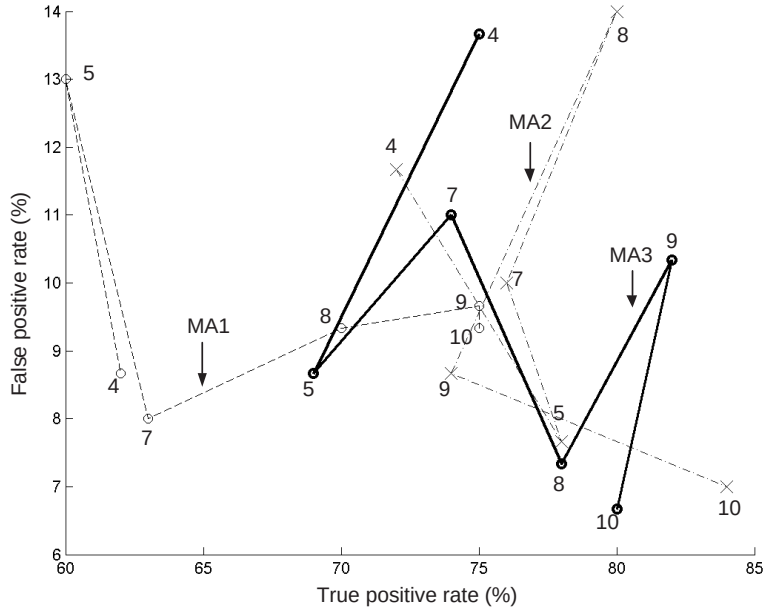

Fig. 1.13 Evolution, for user S2 of the true positive and false positive rate for each MA. The numbers above the markers indicate the session number

From the reported results it appears that, when the users reached the target, the time required is much longer than the optimal time. The best times are more than six and seven times the optimal one for S1 and S2 respectively. Nonetheless, the trends indicate that the times became shorter as the users gained more experience with the BCI. One can imagine that through training, the time can be further improved.

Game	6^{th} session	11^{th} session	12^{th} session
1	Missed	533	486
2	582	564	492
3	Missed	627	516

Table 1.5 Time (in seconds) to reach the target in the free control sessions for user S1. If the object went outside the limits of the screen the target was considered as missed.

Game	6^{th} session	11^{th} session	12^{th} session
1	Missed	656	Missed
2	Missed	Missed	626
3	Missed	634	582

Table 1.6 Time (in seconds) to reach the target in the free control sessions for user S2. If the object went outside the limits of the screen the target was considered as missed.

1.8 CONCLUSION

In this chapter we presented the three adaptation levels that are required from a BCI system and a model that can implement those levels. The first adaptation level is achieved by computing a set of space frequency filters whose parameters are determined for each user and each MA. The resulting feature vectors are then characterized by a set whose elements are those vectors that are at the boundary of the probability density functions associated with each MA. This characterization is periodically updated using an efficient algorithm which easily integrates the "knowledge" gained in a training session in the next session (second adaptation level). The obtained results show that the user performance tends to improve through the training sessions because of the feedback. This implements the third level of adaptation as it engages the adaptive capabilities of the brain.

References

1. A. Gevins. The future of electroencephalography in assessing neurocognitive functioning. *Electroencephalography and Clinical Neurophysiology*, 1(106):165–172, 1998.

2. E. Donchin, K.M. Spencer, and R.S. Wijesinghe. The mental prosthesis: Assessing the speed of a p300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8:174–179, 2000.

3. J. Bayliss. *A Flexible Brain-Computer Interface*. PhD thesis, Department of Computer Science, University of Rochester, 2001.

4. M.S. Middendorf, G.R. McMillan, G.L. Calhoun, and K.S. Jones. Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8:211–214, 2000.

5. N. Birbaumer, A. Kübler, 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 Transactions on Neural Systems and Rehabilitation Engineering*, 8:190–193, 2000.

6. G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89:1123–1134, 2001.

7. 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 Transactions on Neural Systems and Rehabilitation Engineering*, 8:216–219, 2000.

8. J.R. Wolpaw, D.J. McFarland, and T.M. Vaughan. Brain-computer interface research at the wadsworth center. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 8:222–226, 2000.

9. J.d.R. Millán. A local neural classifier for the recognition of eeg patterns associated to mental tasks. *IEEE Transactions on Neural Networks*, 13:678–686, 2002.

10. 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 Transactions on Neural Systems and Rehabilitation Engineering*, 8:164–173, 2000.

11. J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, pages 767–791, 2002.

12. G.L. Calhoun, G.R. McMillan, D.F. Ingle, and M.S. Middendorf. Eeg-based control: Neurologic mechanisms of steady-state self-regulation. Technical Report AL/CF-TR-1997-0047, Wright-Patterson Air Force Base, 1997.

13. T. Ebrahimi, J.-M. Vesin, and G.N. Garcia. Brain-computer interface in multimedia communication. *IEEE Signal Processing Magazine*, pages 14–24, 2003.

14. H.H. Jasper. The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 1(10):371–375, 1958.

15. J.R. Evans and A. Abarbanel. *Introduction to Quantitative EEG and Neurofeedback*. Academic Press, 1999.

16. R.W. Hamming. *Digital Filters*. Prentice-Hall, 1983.

17. L. Vigon, R. Saatchi, J.E.W. Mayhew, and R. Fernandes. A quantitative evaluation of techniques for ocular artefact filtering of eeg waveforms. *IEEE Proc. Science Measurement and Technology*, 147:219–228, 2000.

18. K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 1990.

19. A.B. Barreto, L.M. Vicente, and I.K. Persad. Spatio-temporal eeg patterns associated with voluntary motion preparation. *IEEE Proc. Int. Conf. Eng. Medecine and Biology Society*, 2:857–858, 1995.

20. J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filters for single-trial eeg classification in a movement task. *Electroenceph. Clin. Neurophysiol.*, 110:787–798, 1999.

21. C.W. Anderson, E.A. Stolz, and S. Shamsunder. Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. *IEEE Transactions on Biomedical Engineering*, 45:277–286, 1998.

22. G. Pfurtscheller, C. Neuper, A. Schloegl, and K. Lugger. Separability of eeg signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 6:316–325, 1998.

23. G.N. Garcia, T. Ebrahimi, and J.-M. Vesin. Joint time-frequency-space classification of eeg in a brain-computer interface application. *EURASIP Journal on Applied Signal Processing*, 1(7):713–729, 2003.

24. L.F. Lind. Linear phase filter design in the time domain. *IEE Colloquium on Digital and Analogue Filters and Filtering Systems*, pages 1–4, 1990.

25. D.M.J. Tax. *One-class classification*. PhD thesis, Technische Universiteit Delft, 2001.

26. B. Schölkopf and A. Smola. *Learning with Kernels*. MIT Press, 2002.