ON THE NEED FOR ON-LINE LEARNING IN BRAIN-COMPUTER INTERFACES

José del R. Millán[a]

IDIAP-RR 03-30

JULY 2003

SUBMITTED FOR PUBLICATION

[a] IDIAP – BP 592, CH-1920 Martigny, Switzerland

IDIAP Research Report 03-30

ON THE NEED FOR ON-LINE LEARNING IN BRAIN-COMPUTER INTERFACES

José del R. Millán

JULY 2003

SUBMITTED FOR PUBLICATION

Abstract. In this paper we motivate the need for on-line learning in BCI and illustrate its benefits with the simplest method, namely fixed learning rates. However, the use of this method is supported by the risk of hampering the user to acquire suitable control of the BCI if the embedded classifier changes too rapidly. We report the results with 3 beginner subjects in a series of consecutive recordings, where the classifiers are iteratively trained with the data of a given session and tested on the next session. At the end of these sessions 2 of the subjects reach a suitable performance that is close to allow them to start operating a brain-actuated device.

1 Introduction

Over the last years we have developed a portable brain-computer interface (BCI) that analyzes on-line electroencephalogram (EEG) signals—i.e., the brain electrical activity recorded from electrodes placed onto the scalp—in order to determine the subject's mental state that is then mapped into actions such as selecting a letter from a virtual keyboard or driving a mobile robot [1]-[3]. Ours is one of the few BCIs that have been demonstrated in practice enabling subjects to control different devices such as a computer cursor, a virtual keyboard and even robots [4]-[8]. For reviews, see [9]-[10].

Despite these initial successful demonstrations of BCI, there are a number of issues that need to be solved before this technology can move outside the lab, especially to fulfill the hopes of disabled people [2]. One of these issues is how to scale up the number of recognizable mental states. A second one is to deal with the high variability of brain signals. In this paper we explore this second issue. When a subject is performing the mental tasks required for the operation of a BCI, his/her brain is also engaged in other ongoing activities. The measured brain signals are then the result of the combination of these two activities and successive recordings give rise to highly variable responses. Some well-known factors that condition this ongoing state are attention, fatigue, or motivation. Other factors that arise in natural conditions outside shielded rooms, like in our case, are ambient noise and light, among others. Still other factors, which occur especially at the beginning of the subject's training, are related to changes of the strategies the subjects carry out aiming at facilitating recognition of the mental tasks. These latter factors are more prominent when subjects received on-line feedback, as they verbally report, even though they are instructed not to change strategies along a day. In summary, brain activity changes naturally over time. In particular, this is the case from one session (with which data the classifier is trained) to the next (where the classifier is applied). Thus, there is a need for on-line adaptation of the classifier while the subject operates the BCI. Equally important, rapidly reaching a sufficiently good and stable performance is critical for a BCI, otherwise users can become frustrated and stop utilizing the interface.

To illustrate this need, Figure 1 shows brain power maps in the band 8-12 Hz for the three mental tasks of one of our experimental subjects for two consecutive recording sessions (top panels and bottom panels, respectively). These maps have been generated from the average of all the EEG samples of a given mental task. The three mental tasks are imagination of left hand movements (left panels), imagination of right hand movements (central panels), and generation of words beginning with the same random letter (right panels), respectively. Figure 1 illustrates how, for each mental task, the brain maps—and hence the EEG samples—differ in quite a few respects from one session to the next. This is particularly the case for the first two mental tasks, imagination of left and right hand movements, where even the power map of the "right" task in the second session is closer to the power map of the "left" task in the first session than to that of the same mental task. Altogether, this spontaneous variability of brain signals between sessions hinder correct on-line recognition with any classifier trained with the data of previous sessions.

Although different alternative on-line learning methods are available from the neural network, statistical, and computational learning disciplines (e.g., [11]), an important requirement for BCI (as for any other real-time learning task such as robotics [12]-[13]) is the use of computationally inexpensive training methods. This is a necessary constraint if we wish the system to adapt on the fly and respond quickly, on the order of a second, to make brain interaction of practical use. In this paper we use a simple statistical classifier to illustrate the benefits of on-line learning with three beginner subjects. After a few sessions, two of these subjects reached a performance that, according to our previous experience, is close to allow them to start operating a brain-actuated device [1]-[3], and this despite the subjects have been trained without feedback (see Section 3). To the best of our knowledge, the issue of on-line learning in BCI has not been explored previously.

In the next sections we briefly review those BCI systems that have been actually used for communication and control, highlighting the features of our own approach. Then, we describe the

experimental protocol and the statistical classifier, report the experimental results, and discuss some open issues for on-line learning in BCI.
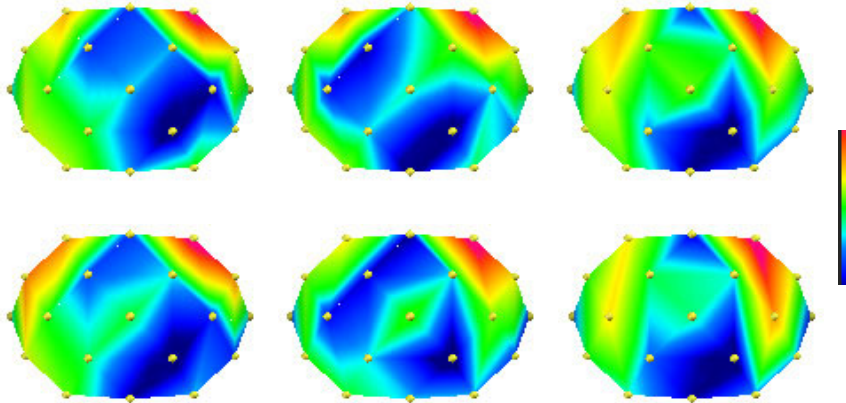

Figure 1. Power maps in the band 8-12 Hz for the three mental tasks of subject "B" for two consecutive recording. Filled circles indicate positions of the electrodes (frontal on top). See text for details.

2 BCI for communication and control

A BCI may monitor a variety of brainwave phenomena. Some groups exploit evoked potentials—the automatic responses of the brain to external stimuli. Evoked potentials are, in principle, easy to pick up but constrain the subject to synchronize themselves to the external machinery. A more natural and suitable alternative for controlling devices is to analyze components associated with spontaneous mental activity. One of such a kind of spontaneous signals is the direct activity of neurons in the motor cortex measured with implanted electrodes [4]-[5]. Such a direct measurement of brain activity may, in principle, enable fast recognition of mental states and even achieve complex interactions. For humans, however, non-invasive methods based on EEG signals are preferable, but they suffer from a reduced spatial resolution and increased noise due to measurements on the scalp. Thus, some researchers measure slow cortical potentials from EEG signals [6] and others look at local variations of EEG rhythms, mainly related to the imagination of body movements [7]-[8]. In our case, we use a combination of motor-related and cognitive mental tasks (such as arithmetic operations or language) as a number of neurocognitive studies have found that different mental tasks activate local cortical areas at different extents.

Most EEG-based BCIs rely on synchronous protocols where the subject must follow a fixed repetitive scheme to switch from one mental task to the next [6]-[8]. In these synchronous BCI systems, the EEG phenomena to be recognized are time-locked to a cue, and a trial typically lasts from 4 to 10 seconds. In contrast, our BCI is based on an asynchronous protocol in which the subject makes voluntary, self-paced decisions on when to stop performing a mental task and when to start the next one. This makes the system very flexible and natural to operate, and yields rapid response times (e.g., 0.5 s in our case).

Typically, EEG-based BCIs make binary decisions as they seek to recognize 2 different mental states. Our system tries to recognize 3 different mental tasks using a simple statistical classifier that may also respond "unknown" for uncertain samples. The incorporation of rejection criteria—a probability threshold in our case, see Section 4—to avoid making risky decisions is an important concern in BCI. From a practical point of view, a low classification error is a critical performance criterion for a BCI; otherwise users can become frustrated and stop utilizing the interface. On the other hand, the BCI should still respond rapidly, on the order of a second, to achieve a sufficient bit rate for

interaction. This is possible with an asynchronous BCI, as we illustrate below.

Based on these principles, we have developed two brain-actuated applications providing alternative channels for communication and control, respectively. Thus, the brain interface can be used to select letters from a virtual keyboard on a computer screen and to write a message [2]. Initially, the whole keyboard is divided in three parts, each associated to one of the mental tasks. Then, as the classifier recognizes what the subject is concentrated on, the keyboard is successively split in smaller parts until a letter is selected. This letter goes to the message and the whole process starts over again. The actual selection of a part of the keyboard incorporates some additional reliability measures (in addition to the statistical rejection criteria) to reduce the likelihood of wrong selections and correct them in case they occur. For our trained subjects, it takes 22.0 s on average to select the desired letter [2]. In preliminary work where those reliability measures were relaxed, one subject could write letters at an average speed of 7.0 s [1]. Other brain-actuated keyboards [6], [8] allow subjects to write a letter every 2 minutes and 1 minute, respectively.

Until now brain-actuated control has mainly relied on implanted electrodes [4]-[5], since EEG-based systems have been considered too slow for controlling rapid and complex sequences of movements. Indeed, so far control tasks based on human EEG have been limited to simple exercises such as moving a computer cursor or opening a hand orthosis [7]-[8]. However, recently we have shown that an EEG-based BCI can also make possible the continuous control of a mobile robot (emulating a motorized wheelchair) generating non-trivial trajectories among different rooms in a house-like environment [3]. The key idea here is that the user's mental states are associated with high-level commands (e.g., "turn right at the next occasion") and that the robot executes these commands autonomously using the readings of its on-board sensors. Another critical feature is that a subject can issue high-level commands at any moment. This is possible because the operation of the BCI is asynchronous and, unlike synchronous approaches, does not require waiting for external cues. Using this system, two subjects have succeeded in mentally driving the robot along non-trivial trajectories visiting 3 or 4 rooms in the desired order. Furthermore, experimental results show that mental control of the robot is only marginally worse than manual control for the same trajectories.

3 Data recording

In the experiments reported below, EEG potentials were recorded with a portable commercial system using a cap with 32 integrated scalp electrodes located at standard positions of the International 10-20 system. The sampling rate was 512 Hz. The raw EEG potentials were first spatially filtered by means of a surface Laplacian (SL) computed globally by means of a spherical spline of order 2 [14]-[15]. This spatial filtering yields new potentials that represent better the cortical activity due only to local sources below the electrodes. The superiority of this kind of SL-transformed potentials over raw potentials for the operation of a BCI has been demonstrated in different studies (e.g., [16]). Then, every 62.5 ms—i.e., 16 times per second—the power spectral density in the band 8-30 Hz was estimated over the last second of data with a frequency resolution of 2 Hz for the 8 centro-parietal channels C3, Cz, C4, CP1, CP2, P3, Pz, and P4. As a result, an EEG sample is a 96-dimensional vector (8 channels times 12 frequency components). No artifact rejection or correction was employed, and so all the samples were given to the classifier that tries to recognize the correct mental task of each of them.

Three volunteer healthy subjects "A", "B", and "C", without any previous experience with BCI or mental training, participated in three consecutive days of recordings where they carried out three mental tasks. These tasks are imagination of repetitive self-paced left and right hand movements, and generation of words beginning with the same random letter. Each day, subjects performed a number of consecutive recording sessions of 4 minutes, separated by breaks of 5-10 min. During each training

session subjects switched randomly every 20 s between the three mental tasks[1]. For the purpose of this study, and other parallel ones, subjects did not receive any feedback. Also, in this off-line study we only analyze the performance of the subjects during four consecutive sessions of the third day.

4 Statistical classifier

The different mental tasks are recognized by a Gaussian classifier trained to classify EEG samples as class #1, #2, #3 or "unknown". In this statistical classifier, every unit represents a prototype of one of the mental tasks (or classes) to be recognized. We use several prototypes per mental task in order to estimate the posterior class probability distribution for an EEG sample. We assume that the class-conditional probability density function of class C_k is a superposition of N_k Gaussians (or prototypes) and that classes have equal prior probabilities. In our case, all the classes have the same number of prototypes, namely 4. In addition, we assume that all four prototypes have an equal weight of 1/4. The number of prototypes has been chosen by cross-validation over all the subjects.

Usually, each prototype of a given class C_k has its own covariance matrix Σ_k^i. In our case, in order to reduce the number of parameters we restrict our model to a diagonal covariance matrix Σ_k that is common to all the prototypes of the class C_k.

To initialize the center of the prototypes, μ_k^n, of the class C_k we run a clustering algorithm—typically, self-organizing maps [17]. We then initialize the diagonal covariance matrix by setting

$$\left(\Sigma_k\right)_{mm} = \frac{1}{S_k} \sum_{n=1}^{S_k} (x^n - \mu_k^n)_m^2$$

where S_k denotes the number of training samples belonging to the class C_k and μ_k^n is the nearest prototype of this class to the sample x^n. The index m denotes the element of a vector, and mm the diagonal element of a matrix.

During learning we improve these initial estimations of μ_k^n and Σ_k^i by stochastic gradient descent so as to minimize the mean square error between the output of the classifier, the posterior class probability distribution, and the target vector in the form *1-of-c*. After updating μ_k^n and Σ_k^i for a given training sample, the covariance matrices of all the prototypes of the same class are averaged to obtain the common class covariance matrix Σ_k. This simple operation leads to better performance than if separate covariance matrices are kept for each individual prototype. It is also worth noting that given the relatively low number of parameters of the covariance matrices to be estimated, as compared to when we use the full matrices, usual regularization techniques (e.g., [18]) do not improve performance.

The brain-computer interface responds every 0.5 s. Firstly, during each frame of 62.5 ms it computes the class-conditioned probability for each class. Secondly, it averages the class-conditioned probabilities over 8 consecutive samples. Thirdly, it estimates the posterior probability based on the average class-conditioned probability of each class using Bayes' formula. Finally, the response is the class with the highest posterior probability provided that it is greater than a given probability threshold[2]; otherwise the response is classified as "unknown" so as to avoid making risky decisions for uncertain samples. This rejection criterion keeps the number of errors (false positives) low.

In on-line learning the learning rates, usually one for each free parameter of the model, are updated after the presentation of each training sample to better fit the local error surface, what accelerates convergence [11]. However, in a BCI context, the risk exists that since the subjects are trained in a

---

[1] While operating a brain-actuated application, the user does essentially the same as during the recording sessions. The only difference is that in the former case he/she switches to the next mental task as soon as the desired action has been performed.

[2] The choice of the probability threshold was based on a ROC study over all the subjects using the first recording session.

changing environment, a classifier that learns too fast may confuse them. As a consequence, the BCI acquires wrong classification rules, what may lead to a collapse if the user tries to follow those rules. For this reason, in the present study we have simply used fixed learning rates and still observe significant improvements in performance.

5 Experimental results

The classifiers have been trained iteratively with the EEG samples of each individual session. That is, the classifier was first trained off-line with the data of the first session (using 4-fold cross-validation) and tested on the second session. Then, the resulting classifier was trained off-line again with the data of the second session and tested on the third session, and so on. In addition, once the first classifier is available, this classifier and the successive ones have been also trained on-line during the first minute of each session—i.e., the samples are processed sequentially and only once—and tested on the remaining 3 minutes of that session.

Table 1 shows the performance of the three subjects over the consecutive recording sessions 2, 3, and 4. For each session, the first column gives the error rate and the second column indicates the rate of "unknown" responses. The probability threshold is 0.75 in all the cases. For each subject the first row gives the performance of the off-line classifier and the second raw that of the on-line classifier. To make a fair comparison, the off-line classifiers have been tested on the same data than the on-line ones; i.e., on the last 3 minutes of each session. For the off-line classifiers, the learning rates of the centers and diagonal covariance matrices were chosen from small sets of fixed values by 4-fold cross-validation. The usual values of the learning rates for the centers and covariances were 1e-8 in combination with 1e-10 or 1e-6 with 1e-8, respectively. For the on-line learning, the values of the learning rates were 100 times larger than for off-line learning.

Table 1: Performance of three experimental subjects over three consecutive recording sessions without feedback. For each session, performance is measured in terms of error (left column) and unknown response rates (right column). For each subject, the first row gives the performance of the classifier trained off-line and the second the performance of the classifier trained on-line on the first minute of the corresponding session.

Subjects	Sessions					
	2		3		4	
"A"	19.1%	28.1%	14.8%	23.8%	11.0%	25.3%
	16.3%	27.1%	14.6%	21.8%	9.4%	24.6%
"B"	25.9%	37.5%	20.5%	37.3%	16.6%	34.1%
	19.4%	32.9%	14.2%	32.0%	11.7%	34.4%
"C"	14.5%	56.8%	20.9%	58.6%	18.3%	55.5%
	14.0%	52.6%	17.4%	51.9%	17.4%	51.4%
	error	*reject*	*error*	*reject*	*error*	*reject*

These results indicate that on-line learning improves systematically the performance of the subjects. In particular, for subject "B", on-line learning reduces the error and unknown rates by 28% and 9%, respectively. For the other subjects, on-line learning yields a more modest improvement. However, taking into account the performances of the different training iterations, a Z-test across subjects on 2 independent proportions shows significant differences between the off-line and on-line learning. On the other hand, the classifier for subject "C" is still far from acceptable performance.

According to our previous experience, this is probably due to the lack of feedback that helps the user develop correct mental strategies. As for subjects "A" and "B", they reached a performance that should be close to allow them start operating a brain-actuated device. A clear indication for this is that they improve their performance from a recording session to the next.

To illustrate the challenge of training the classifier embedded in the BCI in this iterative way, which fits also better user's training in an online setting, we have trained a classifier for subject "B" with the conventional method based on a k-fold cross validation using all available data. In this case, we have merge the same four recording sessions used before and applied a 5-fold cross validation with the same probability threshold of 0.75. The average generalization error and unknown rates are 6.5% and 29.4%, far below the average rates of the three sessions in Table 1—21.0% and 36.3%, respectively. Furthermore, if we perform a ROC analysis to select the optimal probability threshold then errors and unknown responses fall to 2.1% and 36.3%, which mismatches completely the on-line performance.

5.1 Is on-line learning that better?

Given that the classifiers are trained with a limited amount of data, one may wonder if the better performance of on-line learning is due to either the quality of the new EEG samples (there actually exists a drift in the distribution) or just to the larger amount of samples (it uses one additional minute of date, although each sample is processed only once). A related question is whether or not the on-line algorithm is making an optimal (or, at least, good) use of the new data.

Table 2: Performance of the same off-line classifier on two different sets of EEG samples. For each subject, the first row gives the performance of the classifier trained on all the data of a given recording session X_i, and the second the performance of the classifier trained on with the first three minutes of session X_i plus the first minute of the following session X_j. For each session, performance is measured in terms of error (left column) and unknown response rates (right column).

	Subjects	Sessions					
		2		3		4	
off-line, X_i	"A"	19.1%	28.1%	14.8%	23.8%	11.0%	25.3%
off-line, X_i-Δ_i+Δ_j		12.0%	25.5%	11.3%	24.2%	07.6%	22.5%
off-line, X_i	"B"	25.9%	37.5%	20.5%	37.3%	16.6%	34.1%
off-line, X_i-Δ_i+Δ_j		20.3%	36.6%	17.0%	33.6%	12.6%	31.9%
off-line, X_i	"C"	14.5%	56.8%	20.9%	58.6%	18.3%	55.5%
off-line, X_i-Δ_i+Δ_j		09.3%	56.1%	13.0%	63.7%	15.6%	56.0%
		error	*reject*	*error*	*reject*	*error*	*reject*

To answer these questions, we have designed two experiments where, for each subject and recording session, the classifiers are trained on the same amount of EEG samples. In the first experiment we compare the performance of the above off-line classifier trained on all the data of a given recording session X_i (4 minutes in total), with that of the same classifier but trained on the first three minutes of session X_i plus the first minute of the following session X_j, what amounts to 4 minutes of data too. In both cases the classifiers have been tested on the last 3 minutes of the following recording session X_j. Table 2 shows the performance of these two classifiers. For each subject the first row gives the performance of the normal off-line classifier, X_i, and the second raw that of the second off-line classifier, X_i-Δ_i+Δ_j. The probability threshold and the learning rates are the same as before for

both classifiers. A Z-test across subjects shows that the off-line classifier trained on the second set of EEG samples is significantly better than the first, what yields a clear evidence of the drift in the distribution and so the potential advantage of on-line learning.

The second experiment seeks to compare the performance of the above on-line classifier with that of an off-line classifier trained on the same amount of data as the on-line one; i.e., all the EEG samples of session X_i plus the first minute of the following session X_j. Table 3 gives the results of this experiment, where a Z-test across subjects does *not* shows significant differences, except perhaps for subject "A", despite the fact that the on-line learning algorithm processes the samples of the session X_j sequentially and only once.

Table 3: Performance of the on-line and off-line classifiers on the same sets of EEG samples, namely all the data of session X_i plus the first minute of the following session X_j. For each subject, the first row gives the performance of the off-line classifier and the second that of the on-line classifier, which processes the samples of the following session only once—hence the symbol $\oplus$. For each session, performance is measured in terms of error (left column) and unknown response rates (right column).

	Subjects	Sessions					
		2		3		4	
off-line, $X_i+\Delta_j$	"A"	12.6%	31.4%	11.3%	24.2%	08.2%	20.1%
on-line, $X_i \oplus \Delta_j$		16.3%	27.1%	14.6%	21.8%	09.4%	24.6%
off-line, $X_i+\Delta_j$	"B"	20.3%	37.5%	13.0%	40.0%	10.4%	38.3%
on-line, $X_i \oplus \Delta_j$		19.4%	32.9%	14.2%	32.0%	11.7%	34.4%
off-line, $X_i+\Delta_j$	"C"	08.7%	61.4%	06.6%	60.0%	15.3%	59.3%
on-line, $X_i \oplus \Delta_j$		14.0%	52.6%	17.4%	51.9%	17.4%	51.4%
		error	*reject*	*error*	*reject*	*error*	*reject*

6 Conclusions

In this paper we have motivated the need for on-line learning in BCI and have illustrated its benefits with the simplest method, namely fixed learning rates. In particular, we have shown that if the classifier obtained from off-line training on a given recording session is then tuned on-line with a small amount of EEG samples of the next session, then we get a significant improvement in performance. Also, this performance is statistically similar to that obtained if we would train the classifier off-line with the same amount of data. The use of this basic on-line method is supported by the risk of hampering the user to acquire suitable control of the BCI if the embedded classifier changes too rapidly. Obviously, a topic of current research is the analysis of more sophisticated on-line methods to tune the learning rates dynamically that fit the BCI framework, especially when the users receive feedback to help them improve performance.

In this study we have assumed that the correct mental task for every EEG sample is immediately known and so we can perform standard supervised learning. This is the case while the subjects are being trained until they reach a stable suitable control of the BCI. But once the user starts operating a brain-actuated device, the subject's intention is not known instant by instant while still there is a need for on-line learning. In this case, we could resort to TD and reinforcement learning [19], especially if the subject is controlling robotic devices, a task in which these learning techniques have been demonstrated to be particularly effective [13].

Acknowledgments

References

[1] Millán, J. del R. & Mouriño, J. (2003) Asynchronous BCI and local neural classifiers: An overview of the Adaptive Brain Interface project. *IEEE Trans. on Neural Systems and Rehabilitation Engineering,* to appear.

[2] Millán, J. del R. (2003) Adaptive brain interfaces. *Communications of the ACM* **46**:74–80.

[3] Millán, J. del R., Renkens, F., Mouriño, J. & Gerstner, W. (2003) Non-invasive brain-actuated control of a mobile robot. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence,* to appear.

[4] Serruya, M.D., Hatsopoulos, N.G., Paninski, L., Fellows, M.R. & Donoghue, J. (2002) Instant neural control of a movement signal. *Nature* **416**:141–142.

[5] Taylor, D.M., Helms Tillery, S.I. & Schwartz, A.B. (2002) Direct cortical control of 3D neuroprosthetic devices. *Science* **296**:1829–1832.

[6] Birbaumer, N. et al. (1999) A spelling device for the paralysed. *Nature* **398**:297–298.

[7] Wolpaw, J.R. & McFarland, D.J. (1994) Multichannel EEG-based brain-computer communication. *Electroencephalography and Clinical Neurophysiology* **90**:444–449.

[8] Pfurtscheller, G. & Neuper, C. (2001) Motor imagery and direct brain-computer communication. *Proceedings of the IEEE* **89**:1123–1134.

[9] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. & Vaughan, T.M. (2002) Brain-computer interfaces for communication and control. *Clinical Neurophysiology* **113**:767–791.

[10] Millán, J. del R. (2002) Brain-computer interfaces. In M. A. Arbib (ed.), *Handbook of Brain Theory and Neural Networks,* 2nd ed, pp. 178–181. Cambridge, MA: MIT Press.

[11] Saad, D. (ed.) (1998) *On-line Learning in Neural Networks.* Cambridge, UK: Cambridge University Press.

[12] Schaal, S, Atkeson, C.G. & Vijayakumar, S. (2002) Scalable techniques from nonparametric statistics for real time robot learning. *Applied Intelligence* **17**:49–60.

[13] Millán, J. del R. (1996) Rapid, safe, and incremental learning of navigation strategies. *IEEE Trans. on Systems, Man, and Cybernetics—Part B* **26**:408–420.

[14] Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. (1989) Spherical spline for potential and current density mapping. *Electroencephalography and Clinical Neurophysiology* **72**:184–187.

[15] Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. (1990) Corrigendum EEG 02274. *Electroencephalography and Clinical Neurophysiology* **76**:565.

[16] Babiloni, F., Cincotti, F., Lazzarini, L., Millán, J. del R., Mouriño, J., Varsta, M., Heikkonen, J., Bianchi, L. & Marciani, M.G. (2000) Linear classification of low-resolution EEG patterns produced by imagined hand movements. *IEEE Trans. on Rehabilitation Engineering* **8**:186–188.

[17] Kohonen, T. (1997) *Self-Organizing Maps,* 2nd ed. Berlin: Springer-Verlag.

[18] Hastie, T., Tibshirani, R. & Friedman, J. (2001) *The Elements of Statistical Learning.* New York: Springer.

[19] Sutton, R.S. & Barto, A.G. (1998) *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.