33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Accuracy of a BCI based on movement-related and error potentials

Xavier ARTUSI*, Imran Khan Niazi**, Marie-Françoise LUCAS*, Dario FARINA†

Abstract—**New paradigms for brain computer interfacing (BCI), such as based on imagination of task characteristics, require long training periods, have limited accuracy, and lack adaptation to the changes in the users' conditions. Error potentials generated in response to an error made by the translation algorithm can be used to improve the performance of a BCI, as a feedback extracted from the user and fed into the BCI system. The present study addresses the inclusion of error potentials in a BCI system based on the decoding of movement-related cortical potentials (MRCPs). We theoretically quantify the improvement in accuracy of a BCI system when using error potentials for correcting the output decision, in the general case of multiclass classification. The derived theoretical expressions can be used during the design phase of any BCI system. They were applied to experimentally estimated accuracies in decoding MRCPs and error potentials. The average misclassification rate (*n* = 6 subjects) of MRCPs associated to the imagination of elbow flexions at two speeds was 26%, with a bit transfer rate of 0.17. The inclusion of error potentials, experimentally recorded and classified with misclassification rate of 20%, led to a theoretical error rate of 14% with a bit transfer rate of 0.30.**

I. INTRODUCTION

A Brain-Computer Interface (BCI) is a system that allows communication between the brain and external devices, without the use of nerves or muscles. The inputs of non-invasive BCI systems are usually EEG signals. The output is a decision of action among a set of possible ones (a command to an external system). The core of a BCI is thus a classification algorithm, and a training session is used to build the decision rules that allow the decoding of the user's intention. Although the research efforts in this field have increased substantially in the last decade, the applications are still limited. New paradigms for brain computer interfacing (BCI), such as based on imagination of task characteristics, require long training periods, have limited accuracy, and lack adaptation to the changes in the users' conditions. Limited accuracy and robustness are specific problems in applications where the tasks to be classified are similar as, for example, for decoding actions corresponding to the same movement, which is imagined at different target torques and/or different rates of torque developments [1][2].

Several authors have shown that error potentials generated in response to an error made by the translation algorithm can be used to improve the performance of a BCI, as a feedback extracted from the user and fed into the BCI system [3][4][5]. The present study addresses the inclusion of error potentials in a BCI system based on the decoding of movement-related cortical potentials (MRCPs). We theoretically quantify the improvement in accuracy of a BCI system when using error potentials for correcting the output decision. This is done in the general case of multiclass BCI. The derived theoretical expressions can be used during the design phase of any BCI system. The improvement in the BCI transfer rate is quantified by the probability of error of the global system based on the error rates of the classification system and of the error potential detection process. The error rates of the classification system are obtained from the classification of experimental movement-related cortical potentials (MRCPs) corresponding to the imagination of different modalities of elbow flexion imagination. The accuracy in detection of error potentials in single trials is obtained from experimental EEG signals recorded in the same session as the MRCPs after the display of a pseudo feedback. The probability of error of the global system is estimated theoretically from the two sources of error determined experimentally and for a given decision strategy.

The paper is organized as follows. We will first present the proposed methods for classification of movement modalities and detection of error potentials (Section II-A). In Section II-B we provide expressions of the theoretical performance of the corrected system. Finally, in Section III we show the results obtained from experimental signals.

II. METHODS

A. Task classification and error potential recognition

The corrected system is composed of two blocks processing EEG signals. The first block decodes user intentions from signals noted EEG1. The second block performs an evaluation of this decision from signals (noted EEG2) supposed to contain an error potential if the displayed decision is wrong.

1) Task classification algorithm: The aim of the task classification algorithm is the decoding of the user intentions from the EEG1 signals. In our experimental protocol, it corresponds to the discrimination of kinetic parameters from single-trial MRCPs. For feature extraction, we use the marginals of the discrete wavelet transform (DWT) [1][2], that reflect the average signal intensity over dyadic frequency subbands. The dyadic decomposition is well suited to describe and discriminate signals whose discriminative information is mainly at low frequencies since the frequency resolution is higher for low frequencies than for high frequencies. The

* Institut de Recherche en Communication et Cybernétique de Nantes (IRCCyN) - Centrale Nantes, 1 rue de la Noë, 44321 Nantes, France

** Center for Sensory-Motor Interaction, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark (e-mail: df@hst.aau.dk)

† Department of Neurorehabilitation Engineering, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, Göttingen, Germany (dario.farina@bccn.uni-goettingen.de)

choice of marginals is motivated by the fact that the analyzed signals are generally not perfectly time-aligned in realistic (asynchronous) conditions and, as a consequence, the features used for classification should only be composed of frequency descriptors. In order to reduce the dimension of the representation space and according to the low frequency content of MRCPs, signals are described by the DWT marginals in the range $0H_z - 2H_z$ (4 descriptors). According to our previous work [1], the classification was performed from EEG recorded at C_z position, using $2s$ of signal starting just after the cue.

The classification was performed using a linear SVM [6] that is considered as a robust classifier even in rather challenging conditions. If the number of classes n is greater than 2, we use a one-versus-rest (OVR) procedure with n two-class linear SVMs, each of them separating one class against the rest of the population. This SVM used for the task classification is noted SVM_1.

2) Detection of error potentials: This section concerns the analysis of the user reaction to the displayed decision of the BCI (output from SVM_1), in order to estimate if it is correct according to the detection of an error potential. This is performed by analyzing the EEG2 signal recorded after a decision on the task classification is done. When the BCI system provides a wrong decision, an evoked potential appears between about 200 ms and 700 ms after the feedback is displayed [5]. Therefore, we consider the signal in a window between 150 ms and 650 ms after the display of the response to decide if an error potential is present. Two approaches are here proposed and compared: classification and detection, both in a supervised context. For each approach we use the channels FC_z and C_z, the error potential being characterized by a fronto-central distribution along the midline [5]. For each subject, we select the best combination of channels (FC_z, C_z, or $FC_z + C_z$) by optimizing the error rate on a learning set.

We denote x, a multichannel signal, *wrong* the class of signals recorded after a wrong decision (supposed to contain an error potential), *correct* the class of signals recorded after a correct decision, $\mathcal{X}_w$ the learning set of signals of the class *wrong*, and $\mathcal{X}_c$ the learning set of signals of the class *correct*.

a) Classification approach of error potentials: We use a linear SVM classifier (noted SVM_2). The signal is filtered by a low-pass filter at $0 - 10H_z$ and down-sampled from 1024 Hz to 64 Hz. The descriptors are the signal samples (32 descriptors per channel) [7]. When more than one channel is used, descriptors of each channel are concatenated.

b) Detection approach of error potentials: With this approach, when $FC_z + C_z$ are used, we convert multichannel signals into monochannel signals by averaging the two channels. Alternative approaches of multi-channel detection were tested (results not shown) but did not provide better results, therefore simple averaging of the channels was used.

For a given signal x, the detection index $\mathcal{I}(x)$ is defined from the crosscorrelation $\phi_x(\tau)$ between the signal and the averaged potential $\overline{x_w}$ of $\mathcal{X}_w$: $\mathcal{I}(x) = \max\{\phi_x(\tau), \tau \in [-50ms, 50ms]\}$. A strong correlation will translate the presence of an error potential (class *wrong*). This index is compared to two thresholds t_h, t_l:

- If $\mathcal{I}(x) \geq t_h$ the decision of the BCI is estimated as *wrong*.
- If $\mathcal{I}(x) \leq t_l$ the decision of the BCI is estimated as *correct*.

Between thresholds, the detector does not provide a response. The thresholds are fixed in order to limit under a given value the false alarm rate, as estimated on the learning set.

B. Theoretical improvement of the BCI system

To calculate the probability of error of the global system, it is necessary 1) to compute the *a posteriori* probability of a class knowing the responses of the task classifier (online SVM_1) and of the detector of error, and 2) to define a decision strategy from these responses. This is treated in the general multiclass case.

1) Notations: Figure 1 shows the notations used in this section.

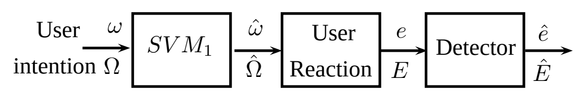

Figure 1: Formalization of the problem.

The problem is formalized using four random variables:

- Ω and $\hat{\Omega}$ with values in $\{\omega_1, \ldots, \omega_n\}$ representing respectively the movement intention and the intention decoded by the incremental SVM_1 (with n the number of classes), $\omega, \hat{\omega}$ being the corresponding realizations;
- E and $\hat{E}$ with values in $\{correct, wrong\}$ representing the accuracy of the SVM_1 decision and it's estimation by the detector, $e, \hat{e}$ being the corresponding realizations; $e = correct$ if $\hat{\omega} = \omega$ else $e = wrong$.

2) Posterior probability of a class: We denote:

- $P_\Omega(\omega) = \text{Prob}(\Omega = \omega)$: *a priori* probability of the class ω,
- $P_{\hat{\Omega}|\Omega}(\hat{\omega}, \omega) = \text{Prob}(\hat{\Omega} = \hat{\omega}|\Omega = \omega) = P_E(e)$: conditional probability of the SVM_1,
- $P_{\hat{E}|\hat{\Omega}\Omega}(\hat{e}, \hat{\omega}, \omega) = P_{\hat{E}|E}(\hat{e}, e) = \text{Prob}(\hat{E} = \hat{e}|E = e)$: conditional probability of the detector.
- $P_{\Omega|\hat{\Omega}\hat{E}}(\omega, \hat{\omega}, \hat{e}) = \text{Prob}(\Omega = \omega|\hat{\Omega} = \hat{\omega}, \hat{E} = \hat{e})$: *a posteriori* probability of the class ω.

The probability of the class ω knowing the response of the classifier ($\hat{\omega}$) and of the detector ($\hat{e}$) is given by the Bayes theorem [8]:

$$P_{\Omega|\hat{\Omega}\hat{E}}(\omega, \hat{\omega}, \hat{e}) = \frac{P_{\hat{E}|\hat{\Omega}\Omega}(\hat{e}, \hat{\omega}, \omega).P_{\hat{\Omega}|\Omega}(\hat{\omega}, \omega).P_\Omega(\omega)}{P_{\hat{\Omega}\hat{E}}(\hat{\omega}, \hat{e})} \quad (1)$$

with

$$P_{\hat{\Omega}\hat{E}}(\hat{\omega}, \hat{e}) = \sum_{i=1}^{n} P_{\hat{\Omega}|\Omega}(\hat{\omega}, \omega_i).P_{\hat{E}|\hat{\Omega}\Omega}(\hat{e}, \hat{\omega}, \omega_i).P_\Omega(\omega_i)$$

3) Strategy of decision and BCI error rate: The detector provides information on the accuracy of the response of the SVM_1, but not the class. If this response is detected as *wrong* in the general multiclass case, it is not possible to deduce the true class even with a perfect detector. So a natural strategy to integrate the detector in the full BCI system is as follows:

- if the detector provides an answer: decision on the class only if the SVM_1 response is estimated as *correct*. When the response is estimated as *wrong*, the subject repeats the imagination task.
- else, selection of the class according to the decision of the task classifier SVM_1.

According to this strategy, the probability of error of decision, when the detector provides an answer, corresponds to the probability that, when the detector indicates $correct$, the class $\hat{\omega}$ given by the SVM_1 is different from the true intention ω of the user. We denote it P_{Er_1}:

$$\begin{aligned} P_{Er_1} &= \sum_{\substack{i,j \\ i \neq j}} P_{\Omega\hat{\Omega}|\hat{E}}(\omega_i, \hat{\omega_j}, correct) \\ &= \sum_{\substack{i,j \\ i \neq j}} P_{\Omega|\hat{\Omega}\hat{E}}(\omega_i, \hat{\omega_j}, correct) P_{\hat{\Omega}|\hat{E}}(\hat{\omega_j}, correct) \quad (2) \\ &= \sum_{\substack{i,j \\ i \neq j}} \frac{P_{\hat{E}|\hat{\Omega}\Omega}(correct, \hat{\omega_j}, \omega_i).P_{\hat{\Omega}|\Omega}(\hat{\omega_j}, \omega_i).P_{\Omega}(\omega_i)}{P_{\hat{E}}(correct)} \end{aligned}$$

with $P_{\hat{E}}(correct) = \sum_e P_{\hat{E}|E}(correct, e).P_E(e)$ and $P_{\hat{E}|E}(\hat{e} \neq e)$ the probability of error of the detector.

By integrating P_{resp} the probability of response of the detector, the global error rate becomes:

$$P_{Er} = \frac{(1 - P_{resp}).P_E(wrong) + P_{resp}.P_{\hat{E}}(correct).P_{Er_1}}{(1 - P_{resp}) + P_{resp}.P_{\hat{E}}(correct)} \quad (3)$$

with $P_E(wrong)$ the probability of error of the SVM_1. In the case of the classification approach of detection, $P_{resp} = 1$ and $P_{Er} = P_{Er_1}$.

Given the previous decision strategy, the performance of the corrected system is also characterized by the probability that the subject has to imagine again the task; it corresponds to the probability that the detector estimates the SVM_1 decision as *wrong* when it gives an answer:

$$P_{repeat} = P_{resp}.P_{\hat{E}}(wrong) \quad (4)$$

And the probability for the user to repeat m times the same task is $(P_{repeat})^m$. The bit transfer rate [9], taking into account the above probabilities, is calculated for a n classes problem as:

$$\begin{aligned} BpT = [\log_2(n) &+ (1 - P_{Er}) \log_2(1 - P_{Er}) \\ &+ P_{Error} \log_2(\frac{P_{Er}}{n-1})](1 - P_{repeat}) \quad (5) \end{aligned}$$

These theoretical formulations will be used in Section III with realistic values for the sources of error as obtained by experimental recordings to provide numerical indications on the advantage in terms of bit transfer rate of including an error potential detector into the system.

C. Experimental protocol

In order to use realistic values for the sources of error, we analyze the performance of the task classification algorithm and error potential recognition on experimental data. Although the theoretical approach described above is general and can be applied to a multiclass problem, the experimental analysis is performed in the simplified case of a biclass BCI on six healthy subjects. Each subject was asked to perform two motor imagination tasks, slow and fast right arm flexion (with the same frequency), according to a visual information (on a computer screen) on when to imagine the movement. Each trial consisted of eight time periods, namely focus, preparation, imagination, hold, rest, error potential focus, and result, and rest period. After performance of each task the subjects were presented with a pseudo feedback (*slow/fast*) that randomly corresponded to a correct or wrong classification (75% correct and 25% wrong). A pseudo feedback allows to isolate the problem of error potential detection from other aspects (such as the variability in classification results among subjects). Despite the fact that the feedback was random, the subjects were informed that the feedback obtained was the result of the online processing of their mental states by a BCI system. In total, 120 trials were collected from each subject in two sets of 60 trials. Trials containing eye movement artefacts or EMG level above four times the standard deviation of noise were rejected.

The evaluation of the performance of the task classification and of the error detector was made by the leave-one-out procedure with the approaches described in Section II-B.

III. RESULTS

A. Classification of movements

The results of the task classification are presented in Table I. The error rate, similar for the two classes (*slow/fast*), was 26% on average over all subjects with a low standard deviation (4%). The average classification error obtained from these data will be used as conditional probabilities of the SVM_1 $P_{\hat{\Omega}|\Omega}(\hat{\omega} \neq \omega)$.

Table I: Results of the task classification.

Subjects	1	2	3	4	5	6	Average
Error rate	28%	24%	32%	30%	22%	22%	26%

B. Detection of error potentials

The threshold t_h (resp. t_l) for the detection approach was fixed to have less than 20% of $\mathcal{X}_c$ (resp. $\mathcal{X}_w$) trials detected as class *wrong* (resp. *correct*) on the learning set. Results are presented in Tables II and III in terms of rate of response of the detector (that will be used as an estimation of P_{resp} when the detection approach is used), and percentage of well classified as *wrong* (resp. *correct*) (that will be used as estimation of the conditional probabilities of the detector $P_{\hat{E}|E}(\hat{e} = e)$ in the next section).

For the two approaches the error rates averaged over the two classes were not very different (20% for the classification approach and 26% for the detection approach). However, the classification approach allows better estimation of the class *correct* whereas the presence or absence of an error potential are equally estimated by the detection method. In Section III-C we will discuss which approach provides the best performance at the output of the full adaptive system.

Table II: Results of the detection approach of error potentials. Rate of response and percentage of well classified (used as estimations of P_{resp} and $P_{\hat{E}|E}$).

Subjects	1	2	3	4	5	6	Av.
$\%responses$	63.3	50.8	85.0	81.7	61.7	56.7	66.5
class $correct$	74	63	78	77	78	70	73
class $wrong$	74	76	80	80	72	67	75

Table III: Results of the classification approach of error potentials. Percentage of well classified (used as estimation of $P_{\hat{E}|E}$). In this case $P_{resp} = 100\%$.

Subjects	1	2	3	4	5	6	Av.
class $correct$	92	91	93	87	83	80	88
class $wrong$	73	43	60	60	37	63	56

Table IV: Global error rate of the corrected system (to be compared to the error rate of the SVM_1 alone, Table I) and probability of repetition. Initial error rate is on average 0.26 and final error rate (P_{Er}) is 0.14.

Sub.	Approach	P_{Er}	P_{repeat}	Intial BpT	Final BpT
1	Detection	19.8%	25.0%	0.14	0.21
	Classification	10.0%	26.2%		0.39
2	Detection	19.6%	16.9%	0.20	0.24
	Classification	16.5%	17.2%		0.29
3	Detection	15.6%	34.4%	0.10	0.25
	Classification	16.8%	23.4%		0.26
4	Detection	15.5%	32.8%	0.12	0.25
	Classification	16.5%	27.1%		0.26
5	Detection	15.4%	20.3%	0.24	0.30
	Classification	17.6%	21.4%		0.26
6	Detection	17.4%	21.6%	0.24	0.26
	Classification	11.5%	29.5%		0.34
Av.	Detection	17%	25%	0.17	0.25
	Classification	14%	20%		0.30

C. BCI error rate improvement

Using the previous results, we compute the probability of error of the corrected system, the probability of repetition and the bit transfer rate. These are summarized in Table IV.

On average over all subjects, using the classification approach of detection of errors, the global error rate is 14% (17% for the detection approach) and the repetition rate is 20% (25% for the detection approach). These results must be compared to the 26% of error rate in the case of SVM_1 alone. The initial average bit transfer rate is 0.17 and is improved by 76% with the classification approach. These results demonstrate that even with a rather low detection rate of error potentials, their inclusion in the BCI system substantially improves the performance. They also indicate that the classification approach for error potentials is preferable over the detection approach, due to a better estimation of the class *correct*.

IV. Discussion and conclusion

In this study, we have quantified the improvement of a BCI system performance by including the detection of error potentials. According to a final decision strategy (repetition of the imaginary movement if the response of the detector is *wrong*), the theoretical performance of the global system was calculated in terms of probability of error, repetition rate and bit transfer rate, from experimental performance of the SVM_1 and of the error detector. The theoretical derivation can be used during the phase of designing of a BCI, without testing each time online the global system on experimental data, but testing only (online or offline) the part corresponding to the classification of a particular set of imaginary tasks.

The task classification algorithm is based on our previous work [1][2] and the classification results are in agreement with these previous studies. For detecting error potentials, two approaches were compared on experimental data. The results (Table IV) showed that the classification approach provides lower probability of error and probability of repetition. With this approach, on average 88% of *correct* (resp. 56% of *wrong*) responses of the SVM_1 are classified as *correct* (resp. *wrong*)). On average, with a repetition rate of 20%, the error rate is 14%, to be compared to 26% in the case of the SVM_1 alone. The more relevant performance index that takes into account the error rate and the repetition rate is the bit transfer rate, which was improved by 76% with the inclusion of error potentials.

These results show that including a detection of errors, although far from ideal, increases significantly the accuracy of the BCI system. The repetition of the imaginary movement on average once every 5 commands seems to be reasonable for clinical applications. Note that if the detection process was perfect, the probability of repetition would be equal to the probability of error of the SVM_1 (here 26%). Finally, the study also shows for the first time an experimental paradigm of error potentials extracted after the classification of MRCPs associated to the same imaginary task performed at two speeds. Despite this classification problem is more challenging that the classification of different tasks, the performance achieved with the inclusion of error potentials are promising for the use of this paradigm in online and clinical settings.

References

[1] D. Farina, O. Nascimento, M. Lucas, and C. Doncarli, "Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters," *Journal of neuroscience methods*, vol. 162, no. 1-2, pp. 357–363, 2007.

[2] D. Vautrin, X. Artusi, M. Lucas, and D. Farina, "A Novel Criterion of Wavelet Packet Best Basis Selection for Signal Classification With Application to Brain-Computer Interfaces." *IEEE transactions on biomedical engineering*, 2009.

[3] R. Chavarriaga and J. Millán, "Learning From EEG Error-Related Potentials in Noninvasive Brain-Computer Interfaces," *IEEE Transactions on Rehabilitation Engineering*, vol. 18, no. 4, pp. 381–388, 2010.

[4] P. Ferrez and J. Millán, "Simultaneous real-time detection of motor imaginary and error related potentials for improved BCI accuracy," pp. 197–202, Graz 2008.

[5] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online Detection of P300 and Error Potentials in a BCI Speller," *Computational Intelligence and Neuroscience*, 2010.

[6] N. Cristianini and J. Shawe-Taylor, *An introduction to support Vector Machines: and other kernel-based learning methods*. Cambridge Univ Pr, 2000.

[7] P. Ferrez and J. Millán, "You Are Wrong Automatic Detection of Interaction Errors from Brain Waves," *International Joint Conference On Artificial Intelligence*, vol. 19, p. 1413, 2005.

[8] R. Duda, P. Hart, and D. Stork, *Pattern classification*. Wiley, 2001.

[9] B. Blankertz, C. Schäfer, G. Dornhege, and G. Curio, "Single trial detection of EEG error potentials: A tool for increasing BCI transmission rates," *Artificial Neural Networks-ICANN 2002*, pp. 138–138, 2002.