The Effect of Vibrotactile Feedback on ErrP-based Adaptive Classification of Motor Imagery

L. Schiatti[1], G. Barresi[2], J. Tessadori[2], L. C. King[2] and L.S. Mattos[2]

***Abstract*— This work presents an implementation of Error-related Potential (ErrP) detection to produce progressive adaptation of a motor imagery task classifier. The main contribution is in the evaluation of the effect of vibrotactile feedback on both ErrP and motor imagery detection. Results confirm the potential of self-adaptive techniques to improve motor imagery classification, and support the design of vibratory and in general tactile feedback into Brain-Computer Interfaces to improve both static and adaptive performance.**

I. INTRODUCTION

The development of Brain-Computer Interfaces (BCIs) offered the possibility to provide a non-muscular channel for communication and control to people with severe motor disabilities [1]. The need of a portable and acceptable technology for the patients boosted the interest towards non-invasive BCIs, mainly relying on electroencephalographic (EEG) signals. In this context, different adaptive signal processing approaches were developed, aiming at improving the system's performance and robustness, thus allowing BCIs effective integration into current Assistive Technologies [2].

One of the limitations of EEG-based BCIs focusing on voluntary control, e.g. motor imagery (MI), is related to the non-stationarity of data, which leads to a decrease in performance from the training to the testing phase. In order to overcome this problem, many approaches have been proposed based on adaptive classification: assuming that the labels of incoming trials are known, it was proven that proper updates in classifier parameters can improve the performance of the static classifier [3]. In a real BCI application, though, user intention, corresponding to the class label, is usually unknown. One interesting approach to cope with this lack of knowledge is to exploit the same neural channel both for extraction of an active control signal (the output of the BCI classifier) and for retrieving information on user's awareness of a misinterpreted intention. The latter can be achieved by detecting a passive EEG feature, the so called Error-related Potential (ErrP), which is evoked in user's fronto-central activity, and has been proven to be detectable on a single-trial basis with a sufficiently high (approximately 80%) accuracy [4]. Since ErrP is an evoked potential, contrary to sensorimotor rhythms exploited to detect motor imagery, features allowing its detection are quite stationary over time. This characteristic makes it suitable to be exploited in a reinforcement learning framework, to improve the performance of an adaptive classifier used for the detection of the active EEG pattern, i.e. the one underlying motor imagery. This concept was implemented by [5], on magnetoencephalographic (MEG) data recorded during a two-class covert attention paradigm, with potential applications in BCI tasks, like mental typewriting [6].

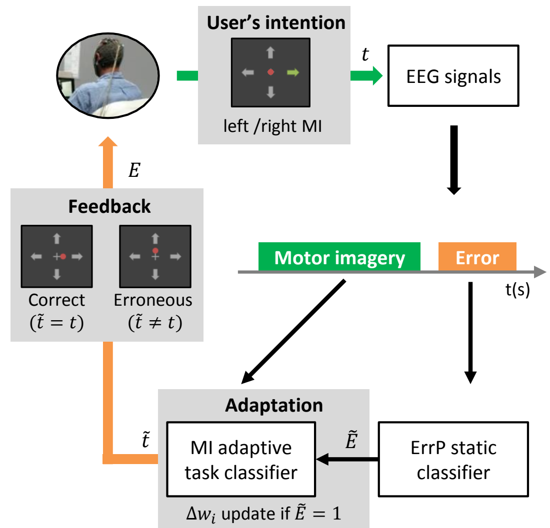

Fig. 1: Scheme of the MI adaptive classifier's functioning. EEG signals are used both to extract an active control signal from a motor imagery task (green phase), and a passive information (Error) on user's evaluation of the MI classifier's output, after feedback is provided (orange phase). The output of a static ErrP classifier is then used to update the MI adaptive classifier's parameters.

In the present work, the method proposed by [5] was applied to EEG data recorded during a four-classes motor imagery task, as shown in Fig. 1. During the experiment, a classification feedback was simulated by means of a virtual cursor movement shown after each motor imagery phase, in order to collect a realistic dataset for both motor imagery and ErrP detection. Only data related to left and right hand motor imagery were considered in this preliminary study, in a binary task classifier implementation. In the described setting, ErrP detection greatly depends on how feedback is designed. Previous studies supported the evidence that using a tactile feedback channel to close the control loop between user and the assistive interface/device, can greatly improve MI-based BCIs performance [7]–[9]. Following this approach, this study presents an evaluation of the effect of tactile feedback on both static and adaptive motor imagery classification and on ErrP detection. Given the high subject-

[1]Unit for Visually Impaired People, Istituto Italiano di Tecnologia, Genoa, Italy (email: lucia.schiatti@iit.it).

[2]Biomedical Robotics Lab, Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genoa, Italy.

specificity of features related to motor imagery, a within-group experimental design was chosen to evaluate the effect of a different feedback. Each subject repeated the experiment in two different conditions, (i) with visual and (ii) with visuo-tactile stimulation during motor imagery (user's intention) and feedback phases (see Fig. 1).

II. METHODS AND MATERIALS

A. Adaptive Classification

The adaptive classifier implemented in this study was introduced by [5], and proved to be effective in improving the binary classification of left and right hand motor imagery on MEG data. The working principle of this adaptive algorithm is based on the updating of the MI classifier's parameters when an error is detected after the MI task classification.

Specifically, labeling as $t \in \{0,1\}$ the true target class, i.e. left or right, corresponding to subject's intention, the output of the MI task classifier can be denoted as $\widetilde{t}$. From the computational point of view, a logistic regression model is used to compute $\widetilde{t}$, in terms of the probability:

$$p(t=1|x,w) = \sigma(x,w) = \frac{1}{1+e^{-\sum_{i=0}^{n} w_i x_i}} \tag{1}$$

where $x = (x_1, ..., x_n)$ is a vector of feature values extracted from the user's EEG activity, which is relevant to discriminating between the two considered classes. The vector $w \in \mathbb{R}^{n+1}$ is the vector of weights, with $x_0 = 1$ accounting for the bias term. The output of the MI task classifier is defined by the function: $\widetilde{t} = \chi\left((t=1|x,w) > \frac{1}{2}\right)$, where χ returns 1 if the argument is true, and 0 otherwise.

The learning rule for the classifier parameters w consists of an update of w in the direction of the gradient of the prediction error, quantified by the log-likelihood function:

$$\varDelta w_i = \frac{\partial G(x,w,t)}{\partial w_i} = \eta(t - \sigma(x,w))x_i \tag{2}$$

where η is the learning rate. In a real system, the subject's real intention t is unknown. Updates $\varDelta w_i$ occur only when the static ErrP classifier detects an error ($\widetilde{E} = 1$), in which case it can be assumed that the observed output $\widetilde{t}$ is incorrect, and $t = 1 - \widetilde{t}$. Therefore, the learning rule for the adaptive MI classifier can be written as:

$$\varDelta w_i = \eta\widetilde{E}(1 - \widetilde{t} - \sigma(x,w))x_i \tag{3}$$

Ideally, if the static ErrP classifier had a perfect behavior (i.e. the output $\widetilde{E}$ is always corresponding to the perceived error E), the learning would happen only when true errors, i.e. true positives (TP), are detected at the output of MI task classification. In practice, the performance of the adapting rule is affected by false positives (FP), i.e. correct trials wrongly detected as erroneous, which cause the MI classifier to learn from incorrectly labeled data, and false negatives (FN), i.e. erroneous trials wrongly classified as correct, preventing the algorithm from performing a correct update.

B. Experimental protocol

The experimental protocol was designed in order to simulate errors made by a BCI in recognizing subject's intents during a bi-dimensional control task. Six healthy subjects (27.7±4.6 y.o., all males) participated in the study, after agreeing with the experiment's guidelines, and signing an informed consent document[1]. Subjects did not actually control the interface, rather they performed a four-classes motor imagery task (both hands, right hand, feet, left hand), with the goal of moving a virtual cursor in one of the four directions (up, right, down, left). Each trial encompassed a fixation cross appearing in a black screen (1.5 s), followed by a green arrow pointing in one of four directions, indicating to the subject which motor imagery task to perform (see Fig. 1). A red dot superposed to the fixation cross was also shown, simulating a virtual cursor. The MI task lasted for 4 s, and it was followed by a feedback simulating the result of MI classification, presented for 2 s after a break of 1 s (black screen). The feedback consisted in the movement of the red cursor in either the direction pointed by the arrow (correct trial) or one of the other three directions (erroneous trial), with 40% of error rate. The experiment was repeated by each subject in two conditions: either receiving only visual feedback (V) as described above, or a visuo-tactile feedback (VT). The latter consisted of a vibrotactile stimulation provided by means of bands placed on the wrists an ankles, for the entire duration of the motor imagery task (4 s), and for 1 s starting from the feedback presentation, according to the cursor movement. Vibrations on both wrists, right wrist, both ankles and left wrist were provided for up, right, down, and left directions respectively. Each experiment encompassed 8 sessions (half in V and half in VT condition). Only trials related to left hand and right hand MI were considered in the present study, leading to around 45 trials per session for each subject.

C. Experimental setup

The experimental setup consisted of 20 active gel electrodes (g.LADYbird from g.tec) located at F1, Fz, F2, FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2 and CP4 according to the standard 10/20 international system. Ground and reference were respectively placed on the forehead (AFz) and left ear lobe. Electrode locations were chosen in order to cover the motor cortex area where changes related to movement imagination are known to happen, and in the frontal area for error potential detection. Experiments were started only after impedance of all electrodes was stably under 5 k$\varOmega$. EEG signals were acquired using a g.USBamp biosignal amplifier at a sampling frequency of 512 Hz. The graphical protocol was developed in Matlab, while data acquisition occurred through a Simulink model that handled the g.USBamp amplifier. Four custom-made silicone rubber (ACC Silicone M230) cuffs

[1]IIT ADVR TEEP01 protocol, approved by the Ethical Committee of Liguria on June 14th, 2016.

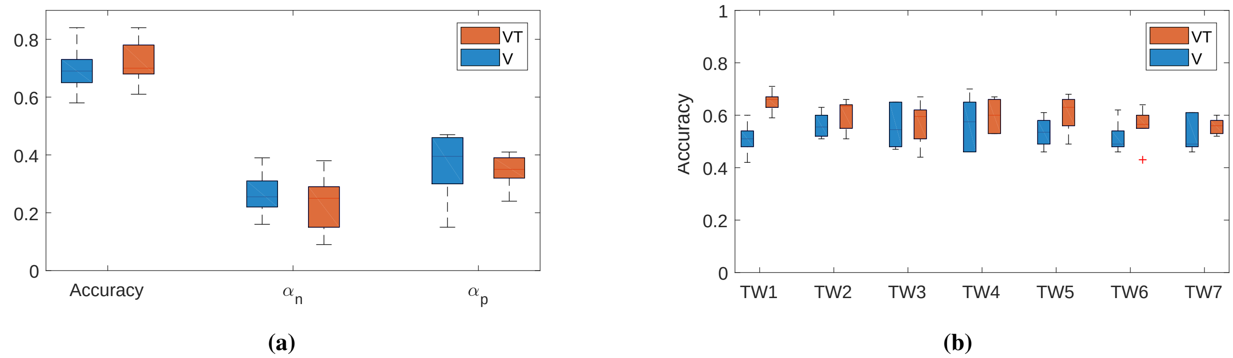

Fig. 2: Static classification results (averages among subjects) comparing V and VT conditions. On the left (a) ErrP classification accuracy, α_n and α_p; on the right (b) MI classification accuracy in different time windows (TW). On each box, the central mark indicates the distribution median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The outliers are plotted individually ('+' symbol).

with embedded vibration motors (Precision Microdrives 304-116) were cast in order to provide tactile stimuli. Control occurred through serial communication, implemented over a direct USB connection[2].

D. ErrP classification

For ErrP classification, EEG data was spatially filtered by means of common-average re-referencing (CAR), then band-pass filtered between 1 and 10 Hz with a 4th order Butterworth filter, since EEG error correlates are known to be slow potentials. Time windows of length 1 s were extracted from recorded data, starting from feedback presentation. Both time and frequency features were extracted and used for ErrP classification. Time features were obtained by sub-sampling signals to 64 Hz. Frequency analysis was performed selecting non-overlapping 0.5 Hz-wide power spectrum bands between 1 and 10 Hz, after applying the FFT. Features were ranked by means of z-score, and only features with score above the threshold of 0.25 were selected. A Support Vector Machine (SVM) with second order polynomial kernel was used for classification, using a leave-one-out cross validation procedure. These classification results, in terms of false negative and false positive rate (α_n and α_p) were then used for the adaptive MI classification, to simulate the output of the static ErrP classifier and the functioning of the updating algorithm in realistic conditions for each subject.

E. MI task classification

For the MI task classification, data was filtered between 8 and 30 Hz with a zero-phase FIR filter of order 20, and spatially filtered by means of small Laplacian [10]. Seven partially overlapping (overlap 0.5 s) time windows (TW) of 1 s length were extracted starting from each trial in the interval 1.5-5.5 s, corresponding to the MI task. Log band-power features were then computed in each TW, considering the three 4 Hz frequency bands: 8-12 Hz, 12-16 Hz, 16-20 Hz. For static MI classification, the logistic regression model presented in Eq. 1 was used, exploiting a 10-fold cross-validation scheme. For adaptive MI classification, data from the first session was used as training set, while data from the remaining three sessions was used as testing set (to simulate a realistic scenario). The training set was exploited to select the 10 most significant features based on the z-score, and to compute the initial weights values for the logistic regression model. The value of static MI classification accuracy was used as ground-truth to evaluate the adaptive classifier performance. This latter was implemented using the same logistic regression model, starting from the weights values of the static MI classifier tuned on the training set, and applying the learning rule in Eq. 3 to update classifier's weights after each new testing example.

III. RESULTS AND DISCUSSION

A. ErrP classification

In Fig. 2a results related to static ErrP classification accuracy are reported, for the V and VT conditions. A mean accuracy of 0.7 is obtained in both cases, regardless of the presence of tactile feedback. This result differs from what was obtained in previous works [11], in which an improvement in ErrP detection was observed when tactile feedback was used. Such an observation could be motivated by different factors. The first one is the fact that in the present experiment, differently from the cited one, the task is bi-dimensional, so three different actions can be considered as an error after each trial. Another reason for the ineffectiveness of tactile feedback on ErrP classification could be the high error ratio (40%), that in general could lead to worse detection accuracy. The α_n and α_p are slightly higher than in [5] (see Table I), probably due to the lower ErrP classification performance. Furthermore, the α_p is greater than α_n for almost all subjects. This is an undesirable condition, since FP cause a wrong update of the classifier's parameters.

B. MI task static classification

Fig. 2b shows results related to static MI classification, in terms of accuracy for each subject and time window, for V and VT feedback conditions. When only visual feedback is provided, the mean accuracy among all subjects and TW is generally close to the chance level, with top results (0.65 and 0.70) between 1 and 2.5 s after the stimulus onset.

[2]Open hardware haptic-driver shield and open source firmware available at: `https://github.com/mrkaroshi/haptic_shield`.

TABLE I: Change in classification accuracy ($\Delta A\%$) from static to adaptive classifier, in V and VT conditions, in case of ideal 100% ErrP accuracy (ΔA_i), and for real values of FN rate (α_n) and FP rate (α_p), ΔA_r.

	V				VT			
Subj.	α_n	α_p	ΔA_i	ΔA_r	α_n	α_p	ΔA_i	ΔA_r
S01	0.26	0.47	14%	6%	0.29	0.36	8%	5%
S02	0.39	0.46	4%	3%	0.27	0.39	13%	7%
S03	0.16	0.15	9%	6%	0.09	0.24	9%	10%
S04	0.31	0.37	9%	-1%	0.15	0.32	9%	4%
S05	0.22	0.42	14%	7%	0.38	0.41	6%	8%
S06	0.25	0.30	13%	4%	0.23	0.34	9%	0%

Adding a tactile stimulation during motor imagery increases the classification accuracy of 2% to 28% on average.

To statistically evaluate the effect of feedback, a 2x7 within-subjects design was adopted. The first within factor is constituted by the 2 levels of absence/presence of feedback (V/VT), while the second one represents the 7 levels of repeated measures. Being the ANOVA assumptions checked, no significant effect of the 7 repeated measures was observed - $F(6, 30)$=1.468 with p=0.223 - and a significant effect of the feedback was found - $F(1, 5)$=12.16 with p=0.0175 - as expected according to the research hypothesis. The significant improvement observed in MI classification in presence of tactile feedback leads to another consideration about the absence of the same result on ErrP classification. Indeed, the use of an identical kind of feedback (only differing in its duration) could not be appropriate in order to simultaneously enhance the detection of an active and a passive feature. A different feedback design could be necessary to draw more solid conclusions to this regard.

C. MI task adaptive classification

To quantify the performance improvements achievable in a real setting, the adaptive classification accuracy was computed considering FP and FN ratios previously assessed for each subject. The adaptive classification was computed in the TW with the best adaptive behavior (maximum improvement in classification accuracy, compared to the static classification), and suitable learning rate η (values chosen between 0.01 and 0.45) for each subject and for both conditions V and VT. Table I reports changes in classification performance (ΔA_r), and the comparison with values achievable in an ideal setting. i.e. 100% of ErrP detection accuracy (ΔA_i). The adaptive algorithm allows for a MI classification performance improvement ranging from few percentage points to 14% in the ideal case. These values decrease when taking into account the FP and FN rate. While no significant difference is observed among results in the ideal conditions for V and VT conditions (ΔA=10% on average), the decrease of performance occurring in the realistic case is generally lower for VT condition. These results may suggest that the addition of tactile feedback, beside the beneficial effect on MI detection, could also increase the benefits of adaptive techniques. Further analysis on a larger sample of subjects will be necessary to assess whether the addiction of tactile feedback could also stabilize relevant features across sessions.

IV. CONCLUSIONS

This work evaluated the impact of designing tactile vibratory feedback into a BCI architecture for adaptive motor imagery classification based on ErrP detection. Static classification results confirmed and extended previous literature findings on benefits coming from the integration of haptic biofeedback in MI-based rehabilitation and neuroprosthetic control. Further investigation will be needed to evaluate if tactile feedback could help in generating MI features that are more stationary over time, thus allowing a better performance of adaptive algorithms for automatic MI classifier retuning. To improve the adaptive classification results, more complex classifiers, such as recurrent neural networks, will be exploited in the next step of this study to implement the explicit update rule. Differently from previous work, the addition of vibratory feedback did not result in an improved ErrP detection. To clarify this point, future work will be devoted to test alternative feedback designs, differentiating the tactile stimulation provided during the active control (MI) and feedback (ErrP) phases.

ACKNOWLEDGEMENT

This research has been partially funded by Fondazione Roma under the TEEP-SLA project's grant agreement.

REFERENCES

[1] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nature Reviews Neurology*, vol. 12, no. 9, p. 513, 2016.

[2] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, *et al.*, "Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges," *Frontiers in neuroscience*, vol. 4, 2010.

[3] P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, and K.-R. Müller, "Towards adaptive classification for bci," *Journal of neural engineering*, vol. 3, no. 1, p. R13, 2006.

[4] R. Chavarriaga, A. Sobolewski, and J. del R Millán, "Errare machinale est: The use of error-related potentials in brain-machine interfaces." *Frontiers in Neuroscience*, vol. 8, 2014.

[5] A. Llera, M. A. van Gerven, V. Gómez, O. Jensen, and H. J. Kappen, "On the use of interaction error potentials for adaptive brain computer interfaces," *Neural Networks*, vol. 24, no. 10, pp. 1120–1127, 2011.

[6] N. M. Schmidt, B. Blankertz, and M. S. Treder, "Online detection of error-related potentials boosts the performance of mental typewriters," *BMC neuroscience*, vol. 13, no. 1, p. 19, 2012.

[7] A. Chatterjee, V. Aggarwal, A. Ramos, S. Acharya, and N. V. Thakor, "A brain-computer interface with vibrotactile biofeedback for haptic information," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, no. 1, p. 40, 2007.

[8] M. Gomez-Rodriguez, J. Peters, J. Hill, B. Schölkopf, A. Gharabaghi, and M. Grosse-Wentrup, "Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery," *Journal of neural engineering*, vol. 8, no. 3, p. 036005, 2011.

[9] C. Jeunet, C. Vi, D. Spelmezan, B. NKaoua, F. Lotte, and S. Subramanian, "Continuous tactile feedback for motor-imagery based brain-computer interaction in a multitasking context," in *Human-Computer Interaction*. Springer, 2015, pp. 488–505.

[10] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for eeg-based communication," *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[11] J. Tessadori, L. Schiatti, G. Barresi, and L. S. Mattos, "Does tactile feedback enhance single-trial detection of error-related eeg potentials?" in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2017, pp. 1417–1422.