Does Tactile Feedback Enhance Single-Trial Detection of Error-Related EEG Potentials?

Jacopo Tessadori, Lucia Schiatti, Giacinto Barresi and Leonardo S. Mattos
Biomedical Robotics Laboratory, Department of Advanced Robotics
Istituto Italiano di Tecnologia, Genova 16163, Italy
Email: {Jacopo.Tessadori, Lucia.Schiatti, Giacinto.Barresi, Leonardo.DeMattos}@iit.it

***Abstract*—Error-related electroencephalographic (EEG) potentials (ErrPs) have been explored to improve the reliability of modern Brain-Computer Interfaces (BCIs), thanks to the information they carry about user awareness of erroneous responses. ErrPs detection on a single-trial basis has been successfully demonstrated, and proved to effectively enhance human-computer interaction and BCI performance. Previous studies tested ErrPs elicited by providing either visual or tactile feedback, showing similar results for all feedback modalities. In the present work, we tested: 1) whether the addition of tactile feedback can improve the detection of ErrP, when used in combination and not alternatively to visual feedback; 2) whether a mismatch between the two different sensory channels can enhance ErrP detection. Results on a study carried out on 12 healthy subjects show that the addition of tactile stimuli significantly affects single-trial ErrP recognition (AUC increment of 4.3%) without significant difference in case of concordant or discordant visual and tactile stimuli.**

I. Introduction

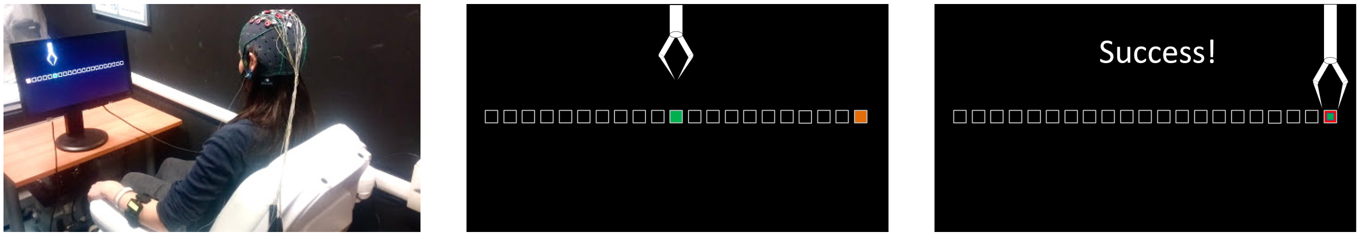
Brain-computer interfaces (BCIs) offer the possibility to improve the degree of autonomy of people suffering from severe motor disabilities, but without cognitive impairments, e.g. Amyotrophic Lateral Sclerosis (ALS) patients, by providing them with a non-muscular channel for communication with the outside world and control of external devices [1]. Usually these systems exploit non-invasive (mainly electroencephalographic - EEG) measures of brain signals, and extract significant features out of them, in order to infer the subject's intention. These features are then forwarded to a classifier, which translates them into a control signal for a communication interface or any other kind of assistive device [2]. Despite recent advances, BCIs still suffer from a lack of reliability, preventing them to be exploited as everyday life widespread products. Indeed, these systems, as well as other interaction technologies based on physiological signals, are prone to errors in recognition of subject's intent. One of the modern challenges of this research domain is consequently to increase the successful classification rate of subject's intention, without intensifying user mental effort or decreasing the communication rate (as would be the case for verification procedures where the output of classification is validated on two subsequent trials [3]).

A unique advantage of the use of brain data as input for interaction technologies is that they can provide at the same time both information from which mental control commands can be derived and information about cognitive states, that are passively elicited during human-machine interaction. Such information can be exploited to improve the interaction quality, as well as the rate of correct intention detection itself. To this aim, previous studies explored the use of one of these cognitive states, namely the awareness of error responses, as a way to improve the performance of BCIs [4].

The presence of error-related potentials in EEG recording right after people realize that they made an error was shown in different physiological studies in the last few decades [5]. Depending on the context, slightly different kinds of ErrPs have been observed: "response ErrP", arising after the subject's incorrect motor action (e.g. pressing the wrong command key) [6]; "feedback ErrP", in typical reinforcement learning tasks, following the presentation of a stimulus that indicates incorrect performance [7]. Beside ErrPs elicited in case of errors made by the subject himself, the presence of an "observation ErrP" was also proved when the subject is observing errors made by an operator during a choice reaction task [8]. The main components of response ErrP in the EEG are a negative potential at 80 ms followed by a larger positive peak between 200 and 500 ms after the incorrect response, while feedback and observation ErrPs are characterized by a negativity at 250 ms after presentation of incorrect performance feedback. All ErrPs have a fronto-central scalp distribution, and are probably generated in the anterior cingulate cortex (ACC), which regulates emotional responses [7].

Recent studies, exploring the feasibility of the inclusion of such ErrPs in BCI applications, demonstrated the existence of a so called "interaction ErrP", in which the error is not made or observed by the subject but by the interface, during the recognition of the subject's intention [9] [10]. This latter ErrP shows a similar shape to the one of response ErrP, with mainly a negative and a subsequent positive peak, whereas the timing is similar to the feedback and observation ErrPs, since the first negative peak occurs around 250 ms after presentation of error feedback.

The attempts of integration of ErrPs in the context of BCIs so far encompassed two main applications: correction of wrong action and the use of ErrP for error-driven learning. In the first case the interaction ErrP has been used e.g. as correcting signal in a motor imagery-based BCI to control one-dimensional step-wise movements of a cursor, by canceling the selected action if an error was detected [11]; or in a P300-based speller, to improve the spelling accuracy by means of online error correction [12]. The second approach has been applied to implement adaptive capabilities in BCIs, either



(a) (b) (c)

Fig. 1. Experimental protocol: (a) experimental setup; (b) scene at the beginning of a session; (c) scene at the end of a session.

updating the BCI classifier [13], or improving the behavior of a semi-autonomous system [14].

Some studies explored the patterns of interaction ErrP when using a different feedback modality, i.e. tactile instead of visual, to provide error signal [15]. In [16] ErrP classification results were evaluated through a comparison between visual and tactile feedback modalities, both for an active and a monitoring task. Erroneous and correct trials were classified above random level (ROC curves AUCs between 0.7 and 0.8) for all conditions, even if visual modality always resulted better than the tactile one.

In this study, we aimed at testing whether the use of a supplementary sensory feedback channel beside the visual one, namely tactile feedback, can improve the single-trial detection of the interaction ErrP for BCI applications. In this case, the tactile feedback is intended to increase the saliency of the visual one, and not to replace it, as investigated in the works described above, in which visual and tactile stimuli were never used together. To this purpose, we implemented a passive EEG acquisition protocol, where the subject was asked to monitor mono-dimensional cursor movements towards a target on a computer screen, with sporadic erroneous movements. The subject possibly received tactile stimulation on either forearm during cursor movements. Three conditions were tested: i) only visual stimuli; ii) tactile stimuli concordant with visual stimuli both in case of correct and erroneous behavior, and iii) tactile stimuli discordant with the visual stimuli during erroneous actions (i.e. tactile stimuli always in the direction of the correct action).

Our research questions are: 1) whether tactile feedback can improve the detection of interaction ErrP, thanks to a more engaging interaction in view of real BCI application for paralyzed patients; and 2) whether a mismatch between two different sensory channels can enhance ErrP detection. The latter is motivated in view of a possible future implementation of error-driven learning: e.g. a shared control scenario in which ErrPs are used to improve the behavior of a semi-autonomous system. In this case, visual feedback would be provided to the user by direct system observation (e.g. system movements direction), while tactile feedback could match some other input modality corresponding to user's high level intention, e.g. desired movement direction inferred by detection of motor imagery (exploiting the same BCI system used for ErrP), or gaze direction, with an eye-tracking system.

II. Methods

A. Experimental protocol

In order to test the influence of tactile feedback on interaction ErrP, we simulated errors made by a BCI in recognizing subject's intents in a one-dimensional control task. The implemented protocol was similar to the ones already presented in previous studies on ErrP inclusion in BCI applications [17]. Specifically, the acquisition protocol was developed to simulate a human-robot interaction task, where the user wishes to move the robot towards a target either on the right or left side of the working space (in this case the computer screen) by means of discrete steps. In our paradigm the subject did not actually control the interface; rather, the robot spontaneously moved towards the intended target with sporadic direction errors. We then tested the presence of ErrP generated by a movement in the wrong direction, simulating an error made by the interface in the recognition of subject's intent. The choice of not implementing a real motor intent detection in this experiment was motivated by the purpose of isolating the problem of tactile feedback effect on ErrP (i.e. our experimental question) from the more complex and general problem of implementing a complete BCI system.

The experimental setup is shown in Fig. 1a, and consisted of 16 active g.LADYbird (g.tec) gel electrodes located at Fz, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2 and CP4 according to the standard 10/20 international system, connected to a g.USBamp biosignal amplifier for EEG signals acquisition. Ground and reference were respectively placed on the forehead (AFz) and left ear lobe. Hardware filters were set to perform a bandpass between 0.1 and 30Hz. Signals were sampled at 256 Hz. Experiments were started only after impedance of all electrodes was stably under 5kΩ. Two Myo (Thalmic Labs) armbands positioned on the subject's forearms were used to provide tactile stimuli, consisting in 0.5 s of vibration on either armband, simultaneously to robot movements.

The graphical stimulation protocol was developed in Matlab, while data acquisition occurred through a Simulink model that handled the g.USBamp amplifier. The experiment scene is shown in Figure 1b and 1c, at the beginning and at the end of a session (robot reaching target), respectively. Each trial consisted of a cursor (green square) and robot arm movement between current position (among 21 possible positions) and

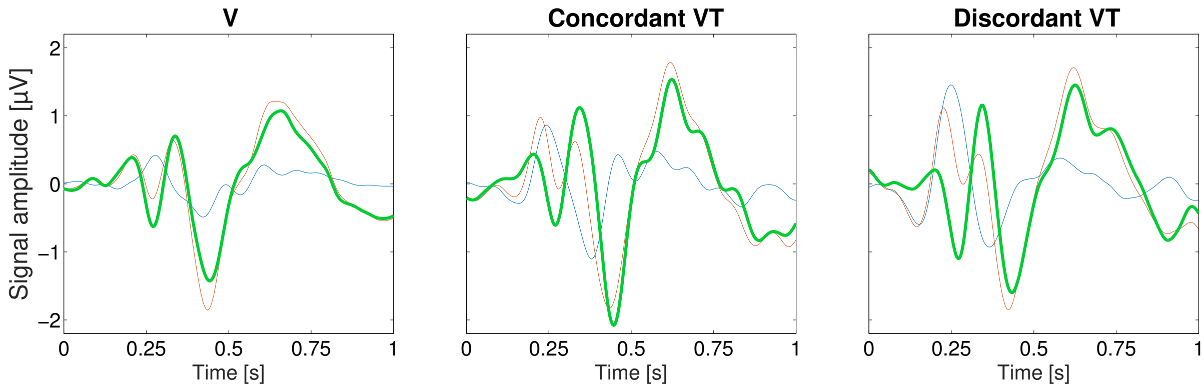


Fig. 2. Grand averages of filtered EEG data for one channel, FCz, in 1 s time window after stimulus onset. Columns refer to different stimuli paradigms (respectively, only visual V, concordant visuo-tactile VT, discordant visuo-tactile VT stimuli). Blue lines represent the average of all trials following a correct cursor movement, while red lines depict averages of erroneous trials. Thick green lines represent the difference of the two signals (erroneous minus correct).

the next one. Timing was set so that movement was followed by a pause of random duration between 1.5 and 2.5 s. Trials were repeated until the target (orange square) was reached. Then, the target was randomly positioned either at the far right or left side of the screen and cursor was placed at the middle position. At the end of each session a simple animation showed the robot extending and grasping the target. The task was performed in three conditions:

1) visual stimuli (V): the task was performed with no armband activation;

2) concordant visuo-tactile stimuli (concordant VT): the subject received a vibration from the armband placed at the side corresponding to cursor movement direction;

3) discordant visuo-tactile stimuli (discordant VT): the subject received a vibration from the armband placed at the side corresponding to target position, i.e. an erroneous movement caused cursor to move left (right), and right (left) armband to activate.

Twelve healthy subjects (31.9±4.2 y.o., 8 males and 4 females) participated in the study. Each subject tested a visual-only feedback condition and one of the two visuo-tactile conditions. Each condition lasted about half an hour, with pauses every 6 reached targets, and encompassed 750 trials with 20% of error probability (cursor movement in the opposite direction of target). Before the experiment started, all subjects agreed with experiment guidelines, and signed an informed consent document[1]. All anonymized experimental data are available online[2].

B. EEG signal processing

1) Error related potentials: EEG data were spatially filtered by 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 between 150 ms and 850 ms after each cursor movement were considered for subsequent analysis, after baseline correction (baseline estimated as the mean value in the 100 ms-long interval before stimulus). Since in this study we considered different stimuli conditions, possibly leading to ErrP variations, both time and frequency features have been computed. In fact it has been demonstrated that the addition of frequency features (theta band spectral power) increases the task-generalization capabilities of ErrP classification based on temporal features alone [18]. This is explained by the fact that ErrP mainly vary in latency but not in amplitude among different tasks. Time features were obtained by sub-sampling signal in time windows by a factor of 4. Frequency analysis was performed selecting non-overlapping 2 Hz-wide power spectrum bands between 2 and 10 Hz, on the same time window mentioned above, after filtering with a Blackman window. This procedure led to 45 time and 4 frequency features per channel.

2) Classification: Several different classifiers were tested (feature selection based on z-scores followed by linear discriminant analysis, learners trees, Gaussian mixture models) and the best single-trial classification results were achieved with a linear support vector machine (SVM). In order to compensate for the unbalanced number of samples in the correct and error classes, synthetic data of the minority class were generated with the SMOTE algorithm [19]. In line with the use of such an algorithm, the misclassification costs for the correct and error classes were respectively set to 0.4 and 0.6. The classifier was trained separately for each subject and type of stimulus. Classification performance was assessed using a 10-fold cross-validation procedure, and it is expressed by means of the correct identification ratios of correct and error classes, while the area under the ROC curves (AUC) was adopted as a single parameter expressing the overall classification accuracy. Early analysis proved that the random nature of both the SMOTE algorithm and partitioning of training and testing sets had a meaningful effect on the observed results. In order to reduce uncertainty, 100 hundred repetitions of the above procedure have been performed, and average classification results are shown and discussed in the following.

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

[2]http://teep-sla.eu/index.php/results/40-smc2017-dataset

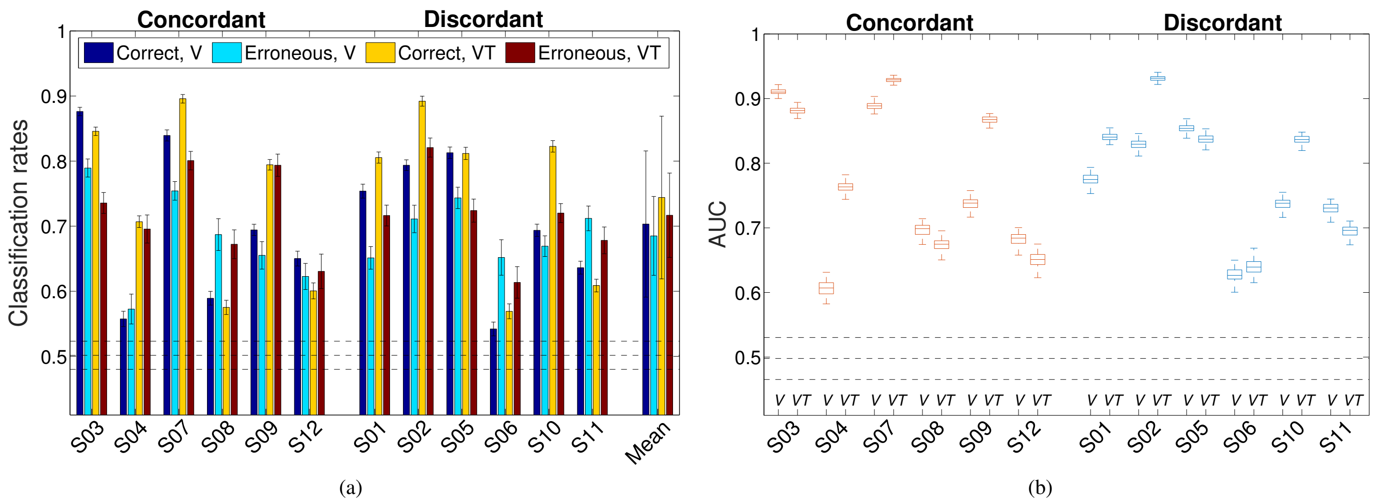


Fig. 3. Classification results: (a) Classification rates for correct and error classes, for all subjects during visual (V) and visuo-tactile (VT) stimuli sessions. Last group of bars represent mean classification rates across all subjects. Error bars represent standard deviations obtained from different repetitions of the classification algorithm (S01 to S12) or across subjects (mean group). (b) Box plots of observed AUC distributions for each subject, referring to visual only (V) and visuo-tactile stimuli (VT) conditions. Red (blue) box plots represent data obtained on subjects that underwent concordant (discordant) VT stimuli. Dashed lines in both panels represent average chance level classification rates or AUCs plus or minus one standard deviation. Chance level results were obtained by training the classifier on actual datasets with randomly shuffled labels.

TABLE I
CLASSIFICATION PERFORMANCE FOR CONCORDANT STIMULI

Subj.	V			VT		
	C[%]	E[%]	AUC[%]	C[%]	E[%]	AUC[%]
S03	88±1	79±1	91±1	85±1	74±2	88±1
S04	56±1	57±2	60±1	71±1	70±2	77±1
S07	84±1	75±1	88±1	90±1	80±1	93±1
S08	59±1	69±2	70±1	58±1	67±2	66±1
S09	69±1	66±2	75±1	79±1	79±2	87±1
S12	65±1	62±2	68±1	60±1	63±3	66±1
M±SD	**70±12**	**68±8**	**75±11**	**74±12**	**72±7**	**79±11**

TABLE II
CLASSIFICATION PERFORMANCE FOR DISCORDANT STIMULI

Subj.	V			VT		
	C[%]	E[%]	AUC[%]	C[%]	E[%]	AUC[%]
S01	75±1	65±2	76±1	81±1	72±2	85±1
S02	79±1	71±2	83±1	89±1	82±1	93±1
S05	81±1	74±2	86±1	81±1	72±2	84±1
S06	54±1	65±3	62±1	57±1	61±2	65±1
S10	69±1	67±2	73±1	82±1	72±1	84±1
S11	64±1	71±2	73±1	61±1	68±2	70±1
M±SD	**71±9**	**69±4**	**76±7**	**75±12**	**71±6**	**80±10**

III. Experimental Results and Discussion

A. Error related potentials

Figure 2 shows the grand averages of time features for one channel, FCz, under the three tested conditions (V, concordant VT and discordant VT) and for both classes (correct, C, and erroneous, E, cursor movement, respectively blue and red thin lines), as well as their difference (thick green line, erroneous minus correct). It is worth noting here that the left graph is obtained by averaging recorded data from 12 subjects, whereas the middle and right column include data from 6 subjects each. The difference between classes present a similar pattern in all considered cases, and coherent in the shape to what expected from the literature: a first negative peak centered around 260 after stimulus onset, and a positive one centered around 330 ms. A second deep negative peak at 440 ms after stimulus, followed by a wider positive one between 500 ms and 750 ms are also visible for all conditions, and match the shape of the interaction ErrP described in [10]. A slight difference is in that we observe a more evident last positive peak, as reported in [20], during observation of robot operation. Differently from what documented in [16], where the ErrP (E-C) signal presented a shift of the first positive peak from 300 ms in case of visual feedback, to 400 ms in case of tactile feedback, we did not observe differences in latencies of V or VT ErrPs. Nonetheless, both C and E trials signals show larger amplitude peaks for VT conditions, and the first small positive peak of E trials is delayed in presence of tactile stimulus.

B. Classification

For the considered tasks, namely error monitoring during either V or VT stimuli, we found that classification results computed on both time and frequency features sets did not change noticeably compared to those achieved using time features only. Therefore, only results obtained by analysis in the time domain are reported. Figure 3 reports bar plots of the classification rates (panel a) for all subjects, classes (C and E) and tested conditions (with and without tactile stimulation), as well as the corresponding AUCs distributions (panel b). Obtained results are, furthermore, summarized in Tables I and II. Specifically, table I lists the accuracy computed for each class (C and E), and the corresponding AUC value, for the

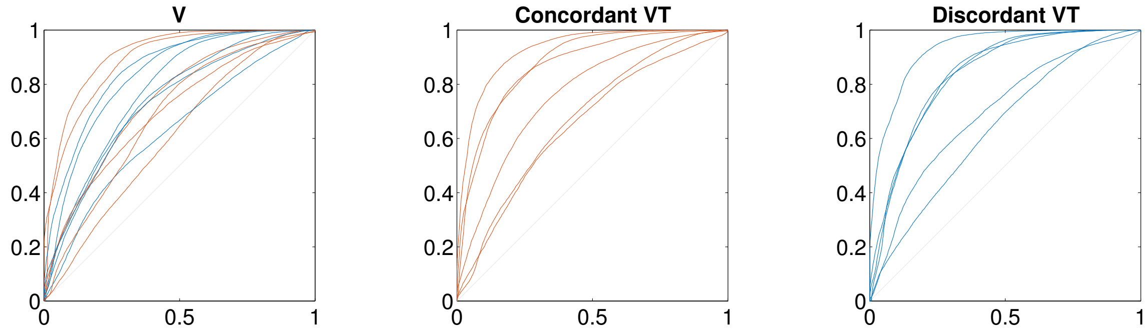


Fig. 4. Average ROC curves. Left graph reports curve obtained for trials during visual-only condition (V), middle and right graphs display ROC curves for data acquired during experiments with visuo-tactile stimuli (respectively, concordant VT and discordant VT). Red (blue) plots represent data obtained on subjects that underwent concordant (discordant) VT stimuli.

subjects that underwent concordant stimuli condition, during both V and VT experiments. Table II shows the same values obtained for discordant stimuli. Classification performance varies greatly among subjects, ranging from very accurate (90% and 82% respectively for C and E classes, in the best case) to barely above or chance level (54% and 57% in the worst case). The overall classification performance, expressed in terms of AUC, is on average between 75% and 76% for V conditions, and increases to 79% and 80% for concordant and discordant VT conditions respectively, with a standard deviation ranging from 7 to 11 percentage points.

To check the correlation between classification performances on the V and VT sessions of the same user, the R-squared of the linear model linking results (in terms of AUC) was computed. It resulted in a value of 0.55, suggesting that, while introduction of tactile stimulation can significantly change classification results, classification accuracy is subject- or montage-dependent, rather than a random fluctuation (i.e. subjects obtaining good results during V condition, tended to do the same in the VT condition).

The inter-subject variability of results is also evident from the ROC curves, shown in Figure 4, and corresponding AUC values reported in Figure 3b. Independently from the VT stimuli condition (concordant or discordant), the addition of tactile stimuli either leaves the performance mostly unchanged or it strongly improves classification performance. In particular, for the majority of subjects the classification performance increases several percentage points from V to VT condition, while for three subjects in the group of concordant stimuli and two subject in the group of discordant stimuli there is a slight decrease of classification accuracy.

The box plot depicted in Figure 5 illustrates the range of changes in classification performance (expressed in terms of AUC percentage points) following the addition of tactile stimulus to the visual one, and their statistical significance. The left box reports changes in AUC values for users subjected to the concordant tactile stimulation protocol, the middle one describes performance for discordant tactile stimulation, while the right column reports changes in AUC values for all VT conditions combined. Both tactile stimulation paradigms facilitate classification: discordant VT feedback causes an average increase of 4.6 points in AUC, while concordant stimulation changes mean AUC by $\sim$4%. One-sample t-tests confirmed results pointing towards a significant effect (mean greater than 0 at 5% significance level) for the first case (p=0.064) and with less evidence for the second (p=0.164). The difference in classification improvement, however, is not significant among the two tested conditions (p=0.688 from one-sided t-test on the hypothesis that AUC change from V to VT is greater for discordant than concordant stimuli condition). This can also be empirically observed in the EEG ErrP waveforms depicted in Figure 2: the amplitude of the peaks in the difference between correct and erroneous classes increases following the addition of tactile stimulus, but both curves for VT paradigms present very similar profiles. Considering all visuo-tactile sessions together, the mean AUC change compared to V condition results in a significant positive increase of 4.2 points (p=0.034).

Overall, these statistical results suggest that tactile stimulation can be helpful in increasing ErrPs classification accuracy. The described method of single-trial ErrP classification proved generally efficient, and the conducted experiments hint at the fact that providing mismatching sensory stimulation might be more effective than the simple involvement of an additional sensory pathway. However the number of subjects tested, given the high inter-subject variability observed, was too low to draw definitive conclusions. In fact, while top results (E and C classification rates$\sim$80) matched literature gold standards [4], recordings on several trials yielded classification results scarcely above chance level. A possible cause for such discrepancies is the low involvement level of the subject during the experiment. This will be the focus point of the follow-up of this work, in which similar experiments, i.e. monitoring of an autonomous agent's correct or erroneous behavior, will be performed using a video game-like or real scenario.

IV. Conclusion

Recent studies reported promising results on the feasibility of detecting single-trial erroneous responses during human-computer interaction, and the inclusion of ErrP detection was proven to lead to significant improvements in the performance

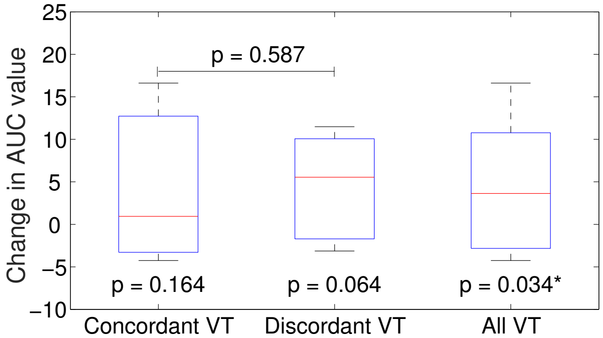


Fig. 5. Distributions of changes in classification performance (AUC) with the addition of tactile stimulus, for concordant VT (left box), discordant VT (middle box) and combined sets (right box). Red line indicates median, blue box extends to first and third quartiles, whiskers reach most extreme data outside of box. Reported p-values are resulting from one-sided t-tests (* indicating significance at 5% significance level).

of BCI systems. In particular, a new kind of error-related potential was found, called interaction ErrP, arising when the interface misinterprets the subject's intention. A still open issue is how to increase the recognition rate of single-trial erroneous and correct responses, and specifically which kind of feedback, whether visual, auditory, somatosensory, or a mix of them, can elicit the strongest interaction ErrP.

Our study aimed at investigating this problem, looking for the effect of a combined visual and tactile stimulation on single-trial detection of interaction ErrP. Furthermore, we explored the effect on ErrP of both concordant and discordant presentation of stimuli from the two sensory channels. The first hypothesis, i.e. whether adding tactile to visual stimuli enhances ErrP detection, was confirmed, as the addition of a tactile stimulus significantly increased ErrP classification performance. The second hypothesis, i.e. whether a mismatch between the two feedback channels can elicit a stronger ErrP, was only partly supported by our study: the discordant VT condition indeed produced improvements in classification performance over V condition, but not significantly differing from the ones caused by the concordant VT condition.

In conclusion, our study supports the idea that exploiting more sensory channels is a promising way to improve ErrP detection in real time BCI applications. This is in line with the research effort towards the inclusion of patients emotional and cognitive states into adaptive assistive interfaces, able to dynamically and automatically adjust their behavior to optimize performance and reliability, for out of the lab BCI technologies.

Acknowledgment

The research leading to these results has been partially funded by Fondazione Roma under the grant agreement supporting the TEEP-SLA project.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] J. d. R. Millán, "Brain-computer interfaces," The MIT Press, Tech. Rep., 2002.

[3] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "Eeg-based communication: improved accuracy by response verification," *IEEE transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 326–333, 1998.

[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] C. S. Carter, T. S. Braver, D. M. Barch, M. M. Botvinick, D. Noll, and J. D. Cohen, "Anterior cingulate cortex, error detection, and the online monitoring of performance," *Science*, vol. 280, no. 5364, pp. 747–749, 1998.

[6] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein, "Erp components on reaction errors and their functional significance: a tutorial," *Biological psychology*, vol. 51, no. 2, pp. 87–107, 2000.

[7] C. B. Holroyd and M. G. Coles, "The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity." *Psychological review*, vol. 109, no. 4, p. 679, 2002.

[8] H. T. van Schie, R. B. Mars, M. G. Coles, and H. Bekkering, "Modulation of activity in medial frontal and motor cortices during error observation," *Nature neuroscience*, vol. 7, no. 5, pp. 549–554, 2004.

[9] P. W. Ferrez and J. d. R. Millán, "You are wrong!—automatic detection of interaction errors from brain waves," in *Proceedings of the 19th international joint conference on Artificial intelligence*, no. EPFL-CONF-83269, 2005.

[10] ——, "Error-related eeg potentials generated during simulated brain–computer interaction," *IEEE transactions on biomedical engineering*, vol. 55, no. 3, pp. 923–929, 2008.

[11] ——, "Simultaneous real-time detection of motor imagery and error-related potentials for improved bci accuracy," in *Proceedings of the 4th international brain-computer interface workshop and training course*, no. CNBI-CONF-2008-004, 2008, pp. 197–202.

[12] P. Margaux, M. Emmanuel, D. Sébastien, B. Olivier, and M. Jérémie, "Objective and subjective evaluation of online error correction during p300-based spelling," *Advances in Human-Computer Interaction*, vol. 2012, p. 4, 2012.

[13] X. Artusi, I. K. Niazi, M.-F. Lucas, and D. Farina, "Performance of a simulated adaptive bci based on experimental classification of movement-related and error potentials," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 480–488, 2011.

[14] R. Chavarriaga and J. d. R. Millán, "Learning from eeg error-related potentials in noninvasive brain-computer interfaces," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 18, no. 4, pp. 381–388, 2010.

[15] M. Lehne, K. Ihme, A.-M. Brouwer, J. B. van Erp, and T. O. Zander, "Error-related eeg patterns during tactile human-machine interaction," in *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*. IEEE, 2009, pp. 1–9.

[16] R. Chavarriaga, X. Perrin, R. Siegwart, and J. d. R. Millán, "Anticipation-and error-related eeg signals during realistic human-machine interaction: A study on visual and tactile feedback," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. Ieee, 2012, pp. 6723–6726.

[17] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J. R. del Millan, "Latency correction of error potentials between different experiments reduces calibration time for single-trial classification," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 3288–3291.

[18] J. Omedes, I. Iturrate, L. Montesano, and J. Minguez, "Using frequency-domain features for the generalization of eeg error-related potentials among different tasks," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 5263–5266.

[19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[20] I. Iturrate, L. Montesano, and J. Minguez, "Single trial recognition of error-related potentials during observation of robot operation," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 4181–4184.