

Real-Time Recognition of Feedback Error-Related Potentials during a Time-Estimation Task

Eduardo Lopez-Larraz, Iñaki Iturrate, Luis Montesano, Javier Minguez

Abstract—Feedback error-related potentials are a promising brain process in the field of rehabilitation since they are related to human learning. Due to the fact that many therapeutic strategies rely on the presentation of feedback stimuli, potentials generated by these stimuli could be used to ameliorate the patient's progress. In this paper we propose a method that can identify, in real-time, feedback evoked potentials in a time-estimation task. We have tested our system with five participants in two different days with a separation of three weeks between them, achieving a mean single-trial detection performance of 71.62% for real-time recognition, and 78.08% in offline classification. Additionally, an analysis of the stability of the signal between the two days is performed, suggesting that the feedback responses are stable enough to be used without the need of training again the user.

I. INTRODUCTION

Feedback is usually an event perceived by a person or an animal as a return of an executed task, given as a result of a conduct that was or not appropriate. Human learning mainly depends on the ability to distinguish between positive and negative feedbacks [1]. Furthermore, it is known that some skills do not develop properly if feedback inputs are absent. In the last few years, therapists have used positive/negative feedbacks to improve their practice and the motivation of patients whose advance is slow [2]. Recently, in the field of Brain-Computer Interfaces (BCI), there has been an increasing interest in their online detection. This is because they carry information to measure indirect parameters of the human learning process, that could be used to maximize the performance of the therapeutic strategy [3].

Event-related brain potentials (ERPs) are evoked responses to an internal or external event, in contrast with spontaneous brain activity [3]. Several types of ERPs have been described in the literature and one broad category are the error-related potentials (ErrPs) [4]. Some of these potentials have been characterized when the human realizes that an error has been committed by himself in a choice reaction task [5], by observation of another user committing an error [6], or by observation of a computer [4] or a simulated robot [7] in interaction or operation tasks. Several of these potentials have also been successfully recognized online in the context of BCI [4], [8].

E. Lopez-Larraz, I. Iturrate, L. Montesano and J. Minguez are with the Instituto de Investigación en Ingeniería de Aragón (I3A) and Dpto. de Informática e Ingeniería de Sistemas (DIIS), Universidad de Zaragoza, Spain. Email: {edulop, iturrate, montesano, jminguez}@unizar.es. This work has been partially supported by projects HYPER-CSD2009- 00067, DPI2009-14732-C02-01 funded by the Spanish Government and the Portuguese FCT project PTDC/EEA-ACR/70174/2006.

Another type of error-related potentials is produced when a subject is informed that he has committed an error (feedback ErrPs). Studies have shown that the typology of positive and negative feedbacks is different. Namely, an error-related negativity (ERN) occurs with higher amplitude (in absolute value) in the second case [1], [9]. As mentioned before, the online decoding of this feedback has a great value in therapies that involve learning new or lost skills.

This paper describes the design of a BCI for the online detection of these potentials, using the paradigm proposed in [9]. This paradigm is a time-estimation task, where the user has to estimate an interval of 1 second, receiving a positive or a negative response depending on his accuracy. The next section details the protocol and the analysis made to the EEG data, section III presents the results obtained, and section IV, draws the conclusions and future work.

II. METHODS

A. Instrumentation

The instrumentation used to record the EEG signals was a commercial gTec system (EEG cap, 32 gold EEG electrodes, 2 clip electrodes for the ears and a gUSBamp amplifier) connected to a computer via USB. The electrodes were placed following previous ERP studies [7] at FP1, FP2, F7, F8, F3, F4, T7, T8, C3, C4, P7, P8, P3, P4, O1, O2, AF3, AF4, FC5, FC6, FC1, FC2, CP5, CP6, CP1, CP2, Fz, FCz, Cz, CPz, Pz and Oz (according to international 10/10 system). The ground electrode was placed on FPz and we used a linked-ears montage for reference. The EEG was digitized at a sampling frequency of 256Hz, power-line notch-filtered to remove the 50Hz line interference, and bandpass-filtered between 0.5 and 10Hz. A Common Average Reference (CAR) filter was applied to remove any background activity detected on the signal. The signal recording, the processing, the visual application and the synchronization between the feedback stimuli and the EEG were developed under the BCI2000 platform [10].

B. Experimentation paradigm and protocol

The experimental protocol followed in this work was proposed by Miltner [9]. The subject had to estimate a given amount of time and received positive/negative feedback depending on his accuracy. The setting of the experiment was a person comfortably sat, observing a computer screen while the EEG was recorded (Figure 1). Each trial started with a visual cue to indicate that the subject had to press a button a given time later (1 second) and then, depending on



Fig. 1. Participant with the EEG system. The feedback shown on the computer screen produces a feedback ERP.

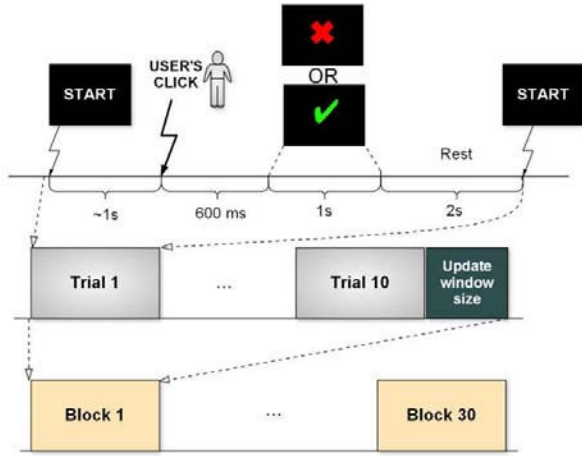


Fig. 2. Diagram of the experimentation protocol.

the proximity to this time, a positive/negative feedback was given 0.6 seconds later.

Five male, right-handed, 24-aged persons participated in the experiments. The participants were duly informed about the protocol. For each participant, the experiment was carried out in two sessions of 30 blocks, where each block had 10 trials (Figure 2). The time between the sessions was three weeks. To balance the number of signals corresponding to positive and negative responses, a time window was computed dynamically every 10 trials taking into account all the previous results (the window was increased as the subject's time-estimation performance improved and decreased as the performance deteriorated). With this strategy, about 150 positive and 150 negative feedback potentials were obtained for each participant and session.

The first experimentation session was done to obtain enough data to characterize the potentials and to have plenty of training examples so as to be able to recognize, in the second session, the signals online. So, the pathway of the first experiment done by each participant was exactly as in Fig. 2. In contrast, in the second, after each trial, a classifier previously trained determined, in real-time and only taking into account the EEG, if the feedback that the user had seen was positive or negative.

C. Analysis of the EEG data

The EEG was filtered with a threshold to eliminate the eye-blink artifacts. Then, the grand averages were computed for both conditions and for all participants. Figure 3(a) depicts the average potentials in the FCz electrode. Notice that the potentials for both conditions are different, showing different brain responses. This figure also shows the difference potential between both conditions, presenting three main peaks: first, a negative component at about 320ms, a prominent positive component around 430ms and another negativity at 530ms. This result is very similar to the one obtained in the original protocol [9].

Additionally, the sLORETA source location technique [11] was used to analyze the intracranial activity from the EEG recorded. Studies have shown that the error processing in the brain activates an area called Anterior Cingulate Cortex (ACC), corresponding to Brodmann areas 24 and 32 [1], [12], [6]. This localization was done on the second and third peaks of the averaged difference (positive minus negative feedback). The active Brodmann areas for the second peak were 24 and 32 (Figure 3(b)), and for the third peak were 6 and 24. Notice that the main focus of activity is close to the ACC, which suggests that exists an error detection and processing. These results agree with [1] and [12] in their studies about feedback ERPs.

III. RESULTS

The main objective of this work is to perform real-time online classification of the feedback ERPs described above as soon as they are recorded. The first step consists in carefully selecting the channels and the time window that contain all the information required for the classification task. To that effect, we downsampled the raw signals to 64Hz and carried out a r^2 analysis, broadly used in neurophysiology [3], [10], in order to determine the areas with the most statistical difference between the two classes. Figure 4 displays the r^2 coefficient for every channel during one second after the feedback presentation, averaged for the five participants. High statistical significance appears in the fronto-central channels, as expected from the analysis of the previous section. Based on Fig. 4, the channels chosen for classification were FC1, FC2, CP1, CP2, Fz, FCz, Cz and CPz, whereas the time interval was [200 – 600] milliseconds after the feedback. The resulting feature vector is the concatenation of the samples of all the selected channels within the time window, making a total of 208 features.

The second step is to feed the feature vectors obtained from the EEG measurements to train a classifier. Although the time window and channels selected were common for all the subjects, we trained a specific classifier for each participant with his own feature vectors. We have employed Support Vector Machines (SVM), since they have already been used to recognize Event-Related Potentials [13]. A key component of this classifier is the kernel function. In this paper, we report results for the ν -SVM [14] with a radial basis function, since it empirically provided the

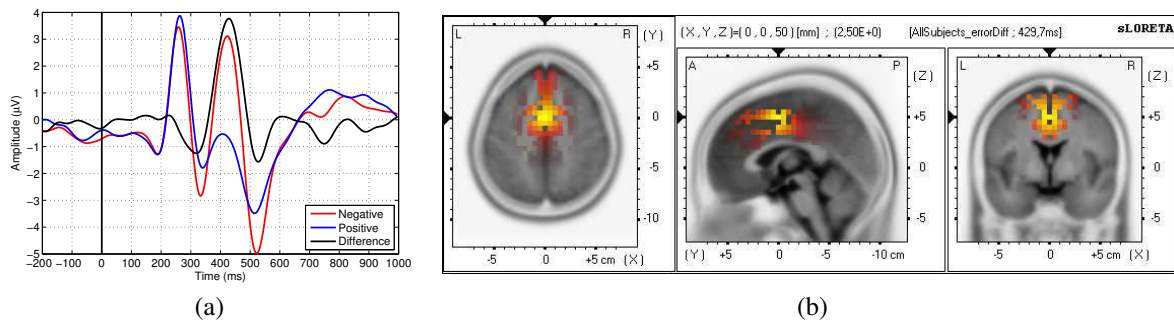


Fig. 3. (a) Grand Average signals for the five participants corresponding to positive, negative and difference between them (positive minus negative feedback) at FCz. Additionally, a baseline of 200ms before the feedback is displayed. (b) Source location at the positive peak (430ms). Yellow and red zones correspond to high activity.

best classification rates over other kernels such as linear, polynomial and sigmoid.

In the remainder of this section, we present classification results in two different settings. First, we analyze the performance of the classifier for single trial classification. Next, we analyze offline the performance of the classifier using cross-validation with all the EEG measurements.

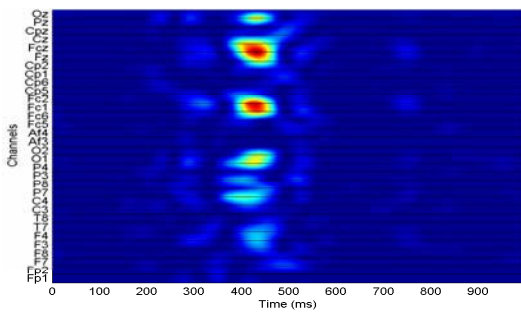


Fig. 4. r^2 of the positive versus negative feedback, averaged for all the participants. Yellow and red areas represent high statistical difference, whereas blue areas represent low or no statistical difference.

A. Real-time classification results

In a single trial classification framework, the objective is to classify online the EEG measurements using an already trained classifier. There are two important aspects that we have analyzed. First, the amount of training data required to achieve good classification rates. This is important to determine the duration of the training sessions, which can be expensive and tiring for EEG. Second, the performance of the classifier between different sessions. This helps to understand up to what extent it is necessary to re-train the classifiers for online recognition tasks. As described in Section II-B, for each participant we recorded two datasets (called Day 1 and Day 2, respectively) with a lapse of several days between them.

Figure 5 presents the results for single trial online recognition. Let us first discuss the left half of the plot corresponding to the classification rates obtained using only data from the first experiment session (Day 1). The figure shows the evolution of the classification rate, averaged over the five participants, as more data has been recorded and used for

training the SVM classifier. The x-axis indicates the percentage of data that was used for training and the percentage of data that was used for testing (e.g. 10-90% indicates that only the first 10% of the data was used to train the classifier, while the remainder 90% was used as test cases)¹. Results show that performance increases as more data is available from a 60% recognition rate to an 80% at the end of the session (the 90-10% case). Although the figure suggests that the accuracy of the classifier could still be improved with more data, in the next section we will show that it is close to its maximum performance.

Let us now analyze the classification of the second session (Day 2) using also the data from the first session. Results for this case are shown in the right part of Fig. 5. The accuracy rate at the black line (tagged as 100-100%) corresponds to the case where the classifier was trained with the whole dataset of Day 1 and the test is the whole dataset of Day 2. As mentioned in section II-B, this classification was done in real time while the participant was performing the task. It is interesting to note that there is a decrease on the classification rate with respect to the final score of Day 1, but it is better than the initial rates of Day 1. More precisely, the decrease in percentage was of 6.33% and 11.35% on negative and positive feedbacks respectively, suggesting that negative feedbacks are more stable across different days than positive feedbacks.

In order to recover from the degradation between Day 1 and Day 2, it is possible to train the classifier with data from both sessions. The right part of the figure shows the accuracies obtained for this case. The x-axis ranges from 110-90% to 190-10%. As before, the left number indicates the percentage of examples used for training (100% of Day 1 plus the corresponding percentage of Day 2) and the right number indicates the percentage of data of Day 2 used for testing. The results show an increase in accuracy, as the percentage of data from Day 2 included in the training set is increased. Furthermore, the accuracy remains stable after incorporating only the first 20% of Day 2 measurements

¹The SVM training time is under one millisecond for the entire dataset in a CoreDuo 2.4GHz. Classification requires just a few micro-seconds. These computational times make possible to even re-train the classifier for every single new measurement.

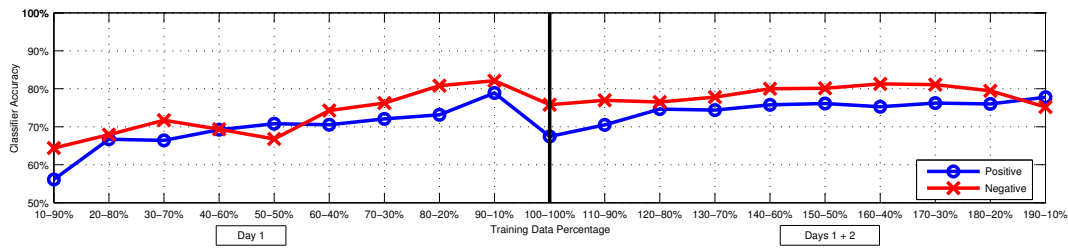


Fig. 5. Classification rates obtained sequentially using different percentages for train and test set from Day 1 (left) and blending data from both days using the whole dataset from Day 1 and a variable percentage of Day 2 (right). Values 10-90% to 90-10% represent data from Day 1, while values higher than 100% mean all the data from Day 1 and the remaining from Day 2.

(labeled as 120-80%). This fact points out that appending a very small number of new examples (60 trials recorded in about 5 minutes) is enough to obtain the maximum accuracy around 80%.

B. Offline classification

The previous analysis incorporated the data in a specific sequential way to analyze the evolution of the classifier analysis. This section presents results of an offline analysis of the data using cross-validation to characterize the classifier's behavior using all the available data. It provides information about the maximum accuracy that may be achieved by the classifier and allows us to compare performances among the different participants.

Table I shows for each session (Day 1 and Day 2) the classification rates obtained separately for each participant using 10-fold cross validation. The recognition percentage in all the cases is superior than a random classifier and, on average, gives a rate of 78.08%. The standard deviations were in all cases below 9%, implying a similar recognition rate for all the participants.

TABLE I
CLASSIFICATION RATES OBTAINED

	Day 1			Day 2		
	Positive	Negative	Average	Positive	Negative	Average
P.1	72.05	77.57	74.81	90.42	78.81	84.62
P.2	73.71	81.08	77.40	70.00	75.99	73.00
P.3	78.57	84.83	81.70	87.71	77.14	82.43
P.4	74.43	79.84	77.14	76.21	79.42	77.82
P.5	72.81	73.35	73.08	79.73	77.88	78.81
Average	74.31	79.33	76.82	80.81	77.85	79.33
std	2.54	4.26	3.25	8.35	1.36	4.48

IV. CONCLUSIONS AND FUTURE WORK

This work addresses the problem of single-trial real-time recognition of feedback potentials. A experiment of a one-second time estimation task was carried out with five participants. Results show that it is possible to achieve good recognition rates (around 80%) using a SVM classifier. Furthermore, we analyzed the requirements of the classifier in terms of the amount of training data, its performance among sessions and the possibility of fast re-training in order to achieve good performances using data from previous sessions.

As future work, we plan to investigate the case where feedback signals are not synchronized with the BCI (e.g. provided by a therapist). In this case it is necessary to detect these signals first and then classify them.

REFERENCES

- [1] S. Nieuwenhuis, C. B. Holroyd, N. Mol, and M. G. H. Coles, "Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance," *Neuroscience & Biobehavioral Reviews*, vol. 28, no. 4, pp. 441 – 448, 2004.
- [2] K. P. Tee, C. Guan, K. K. Ang, K. S. Phua, C. Wang, and H. Zhang, "Augmenting Cognitive Processes in Robot-Assisted Motor Rehabilitation," in *Proceedings of the 2nd Biennial IEEE/RAS-EMBS International*, 2008.
- [3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [4] P. Ferrez and J. d. R. Millan, "Error-Related EEG Potentials generated during simulated brain-computer interaction," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 923–929, 2008.
- [5] M. Falkenstein, J. Hoonmann, S. Christ, and J. Hohnsbein, "ERP components on reaction errors and their functional significance: a tutorial," *Biological Psychology*, vol. 51, no. 2-3, pp. 87 – 107, 2000.
- [6] H. van Schie, R. Mars, M. Coles, and H. Bekkering, "Modulation of activity in medial frontal and motor cortices during error observation," *Nature Neuroscience*, vol. 7, pp. 549–554, 2004.
- [7] I. Iturrate, L. Montesano, and J. Minguéz, "Robot Reinforcement Learning using EEG-based reward signals," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010.
- [8] G. Visconti, B. Dal Seno, M. Matteucci, and L. Mainardi, "Automatic recognition of error potentials in a P300-based brain-computer interface," in *Proceedings of the 4th International Brain-Computer Interface Workshop & Training Course*, 2008, pp. 238–243.
- [9] W. H. R. Miltner, C. H. Braun, and M. G. H. Coles, "Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task: Evidence for a Generic Neural System for Error Detection," *Journal of Cognitive Neuroscience*, vol. 9, no. 6, pp. 788–798, 1997.
- [10] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *Biomedical Engineering, IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, June 2004.
- [11] R. Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods and Findings in Experimental and Clinical Pharmacology*, pp. 5–12, 2002.
- [12] P. Ferrez and J. d. R. Millan, "Error-Related EEG Potentials in Brain-Computer Interfaces," *Ph.D. thesis, Ecole Polytechnique Federale de Lausanne*, 2007.
- [13] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain computer interfaces," *Journal of Neural Engineering*, vol. 4, June 2007.
- [14] B. Schlölkopf, A. Smola, R. Williamson, and P. Bartlett, "New support vector algorithms," *Neural Computation*, vol. 12, pp. 1207–1245, 2000.