Towards error-free interaction

Tsvetomira Tsoneva[1] Jordi Bieger[1,2] Gary Garcia-Molina[1]

***Abstract*— Human-machine interaction (HMI) relies on pattern recognition algorithms that are not perfect. To improve the performance and usability of these systems we can utilize the neural mechanisms in the human brain dealing with error awareness. This study aims at designing a practical error detection algorithm using electroencephalogram signals that can be integrated in an HMI system. Thus, real-time operation, customization, and operation convenience are important. We address these requirements in an experimental framework simulating machine errors. Our results confirm the presence of brain potentials related to processing of machine errors. These are used to implement an error detection algorithm emphasizing the differences in error processing on a per subject basis. The proposed algorithm uses the individual best bipolar combination of electrode sites and requires short calibration. The single-trial error detection performance on six subjects, characterized by the area under the ROC curve ranges from 0.75 to 0.98.**

I. INTRODUCTION

Advanced human-machine interaction (HMI) relies on pattern recognition algorithms, which are not error free. Machine errors reduce the overall performance of the system and can be particularly annoying for the user. The human brain has developed complex neural and cognitive mechanisms that deals with error awareness. We can utilize these mechanisms to improve the performance and the usability of HMI systems like brain-computer interfaces (BCIs). Different neurophysiological studies have shown the presence of error-related responses in human electroencephalogram (EEG), called error-related potentials (ErrP). ErrPs are associated with the anterior cingulate cortex (ACC) [1], which is also responsible for regulating emotional responses.

Different studies have shown the presence of ErrP emerging shortly after an error made by the subject [2]. Usually they use a choice reaction task, which requires quick response to a stimulus. This type of ErrP is also known as error-related negativity (ERN) or response ErrP. Other studies have reported a different type of ErrP during a reinforcement learning task [3]. This type of ErrP appear after feedback indicating an erroneous response from the subject; hence it is known as feedback ErrP. Whereas these neural correlates of error awareness are manifested after errors committed by the subjects themselves, ErrPs are also present after an observation of an error, for example committed by the interface the subject is interacting with. These are known as interaction ErrP. ErrPs of this kind have been observed during simulated or actual brain-computer interaction [4][5].

1. Philips Research, HTC 34, 5656 AE Eindhoven, The Netherlands
2. Radboud University, 6525 HR Nijmegen, The Netherlands
Please direct all correspondence to: tsvetomira.tsoneva@philips.com

Our objective is to explore the brain mechanisms dealing with awareness of erroneous responses and to design a practical solution that can be integrated in any HMI system. That is why a number of requirements should be satisfied. Real-time operation is necessary prerequisite if we want to integrate such a solution in already real-time HMIs. Thus, we aim for computationally efficient signal processing and robust detection of erroneous responses. Individual specificities should be considered as different users might have different physiological responses. The solution must adapt to the user, ideally after a short calibration procedure. Convenience is also essential for a system working in a real-life environment, thus, we want to use only few measurement sites.

In this paper we report our approach for machine error detection following the above mentioned requirements. The paper is organized as follows. We first introduce our experimental setup in Section II. Then we report the results we have obtained in Section III. We address certain points of discussion in Section IV and conclude the paper in Section V.

II. EXPERIMENTAL SETUP

A. Participants

Six volunteers (3 females and 3 males) aged between 23 and 29, took part in the experiment. All participants were healthy, right-handed and had normal or corrected-to-normal vision. They signed an informed consent form before the start of the experiment.

B. Task

To minimize the influence of external factors and to isolate the response to errors made by the interface, we designed a relatively trivial experimental paradigm in the form of a game. The game was very simple, so that it was very unlikely that the subjects would make an error. The goal of the game was to move a square (the stimulus) horizontally from one side of the screen to the other by a single key press. The total length of the path was 14 squares, the last one being the target. The subject was given up to 7 moves to successfully complete the path. The stimulus could be moved with a step of one or three squares. The subject had to develop an efficient strategy in order to reach the target within the given number of moves.

Figure 1 illustrates one trial in the experimental protocol. At first the stimulus (a gray square) appears on screen. After 1700 ms a numerical and a visual indication of the suggested step size and the expected next position are presented. The step size can be either one or three (selected at random), therefore the expected next position is either one or three squares further from the current position of the stimulus. If

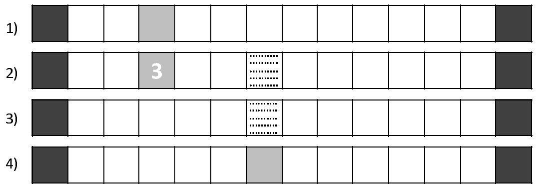

Fig. 1. A correct trial in the experiment protocol: 1) The stimulus (gray square) appears on screen for 1700 ms. 2) A numerical and a visual (square with dots) indication of the expected next position are presented for up to 2000 ms. 3) The expected position of the stimulus (the square with dots) stays on screen for 1000 ms after a key is pressed. 4) The stimulus has moved to the new position.

the subject likes the suggested step s/he is expected to press a key. If a key is not pressed in the next 2000 ms a new suggestion is given. After the key press, the visual indication of the expected position stays on screen for another 1000 ms, after which the stimulus moves to the new position. The new position of the stimulus may be the expected one (correct trial) or not (error trial). In an error trial the stimulus moves either one or two squares back from the current position. In order to finish the game, the last move should end exactly on the target. In case of selecting a step bigger than the remaining squares to the target, the stimulus jumps back to the beginning of the path. The game continues until the target is reached. The game is won if the subject reaches the target within the given number of moves, and is lost otherwise.

C. Experimental procedure

The subjects were told that they would play about 30 games and they were instructed to try to win as many as possible. All subjects played the game for about 30 minutes, completing between 23 and 31 games (mean: 28 games). There were two modes of the game: *correct mode* and *error mode*. In the correct mode the system did not commit errors. In the error mode, error moves appeared with a probability of 25%. The session started with two games in correct mode. Then a correct mode game was played every 6 games to reduce the effect of habituation. The rest of the games were played in error mode. The mean number of trials per game was 7, resulting in an average of 198 trials per subject. The mean number of error trials per subject was 40, and the mean number of correct trials was 158.

D. Signal acquisition

Continuous EEG from 32 scalp electrodes, digitized at 2048 Hz, was acquired using a BioSemi ActiveTwo system [6]. The electrodes were positioned according to the international 10/20 standard and were uniformly distributed over the scalp (see the axes of Fig. 3 for the electrode positions). The signals were subsampled to 256 Hz and then bandpass filtered in the 0.5-25 Hz band using a Butterworth filter. In order to minimize the effect of any background neuronal activity in the area of interest and to emphasize the differences in neural responses to error and correct trials we performed an exhaustive search for the best bipolar combination (BBC) of electrodes on a per subject basis. This procedure is explained in detail in Section III-C.

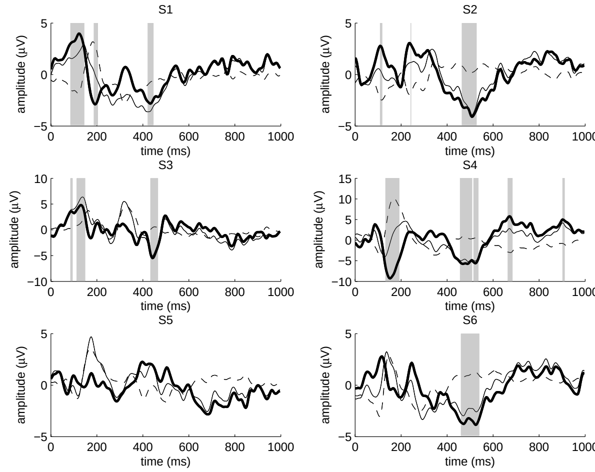

Fig. 2. Difference (bold line) between potentials for error (solid line) and correct (dashed line) trials for all subjects (bipolar combination: Cz-Pz). Gray areas show the statistical significance ($p < 0.0002$).

III. Experimental results

A. Error-related potentials

We first checked whether there were any differences in the brain responses to error and correct trials. Fig. 2 shows the brain potentials for error and correct trials as well as the difference (error-minus-correct) for the bipolar combination Cz-Pz. ErrPs are associated with the ACC located in the fronto-central sites of the brain and the choice of Cz is common in the literature (see Section I). For five of our subjects (S1, S2, S3, S4 and S6) a first positive peak is observable around 150 ms after the stimulus movement, followed by a negative peak around 200 ms and a positive peak around 300 ms. Finally, a broader negative peak occurs around 450 ms. This ErrP shape is similar to previously reported results [5]. Subject S5 did not exhibit this response. In order to check the significance of the difference between the responses to error and correct trials we performed a paired t-test with Bonferroni correction. The areas where the difference is significant are shaded in gray in Fig. 2. As it can be seen, the resulting error potentials differ between subjects, some of them exhibiting higher amplitudes than others. Furthermore, the statistical significance of the difference between error and correct trials is highly variable, as it could be expected from the low signal-to-noise ratio of EEG signals. If we want to use ErrP to improve HMI, we need to be careful as inappropriate adaptation of the interface could further frustrate the user. Emphasizing the individual differences in error processing could lead to a more robust solution.

To investigate the effect of frequency of error occurrence on error processing we invited two of our participants for a second experiment. We used the same experimental protocol but this time the probability of errors was 50%. The ErrPs from the second experiment were very similar (almost identical) to the ones from the first experiment, suggesting that frequency of errors does not affect the observed neural responses.

B. Error detection algorithm

Templates of the brain responses to erroneous and correct outcomes (*error template* and *correct template* respectively) for each subject can be obtained from a training set. In this experiment the training set was composed of a random selection of the recorded potentials. In practice the training sets results from a calibration procedure. The templates are one-second long averages, time-locked to the stimulus movement. The templates are univariate signals, which result from a bipolar EEG combination. To decide if a given trial is erroneous, a score is calculated as follows.

$$s = \sum_i (1 - p_i)(|x(i) - \tau_e(i)| - |x(i) - \tau_c(i)|), \qquad (1)$$

where x is the brain response to the current trial, τ_e and τ_c are the error and correct templates respectively, i is the sample index, and p_i is the significance level (determined by a paired t-test) of the difference between the templates at sample index i.

If the score is higher than a previously selected threshold (e.g. zero), then the trial is considered erroneous. The selection of the threshold depends on the application. Thus, a balance between the true positive and false positive error detection rates needs to be sought. A practical manner to assess the threshold selection impact on the system consists in drawing the ROC curve (see Fig. 5), which represents the true positive versus the false positive error detection rate for different threshold choices. The area under the ROC curve (AUC) provides an indication of the error detection performance.

C. Best bipolar combination

Given a set of M recorded EEG sites, $M(M-1)/2$ possible bipolar combinations exist. An exhaustive search for the BBC can be done by computing the average AUC (over 50 random selections of train/test sets) for all possible combinations. The result of this procedure is visualized in the colormap of Fig. 3, where the gray level of a cell is proportional to the AUC associated with the corresponding bipolar combination. For convenience of visualization, the AUC values of the diagonal elements has been artificially set to 0.5 (random level). By definition the colormap is symmetric with respect to the diagonal. Certain regions in the map, especially the ones representing fronto-parietal combinations, exhibit larger AUCs. Yet, these regions are not common to all subjects.

To better visualize the individual differences and to facilitate physiological interpretation, we represent in Fig. 4 the bipolar combinations corresponding to the largest twenty AUCs. The electrodes of each combination are connected through a line whose thickness is proportional to its AUC. For convenience of visualization the AUCs are quantized into three levels. For all subjects but S3, we observe strong connections between right and central sites. The electrode site F8 is of particular relevance for subjects S1, S4, and S6.

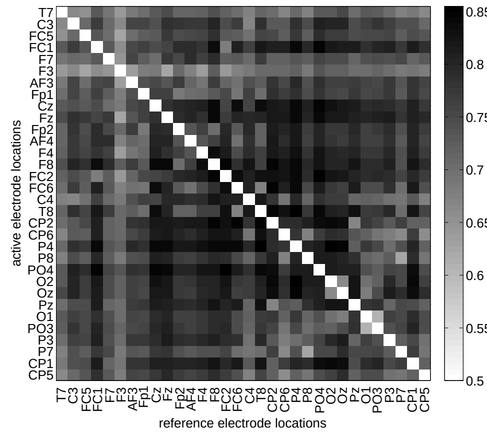

Fig. 3. Average AUC across all subjects for all possible bipolar combinations. The gray level of a cell is proportional to the AUC of the corresponding bipolar combination, darker cells indicating better performance.

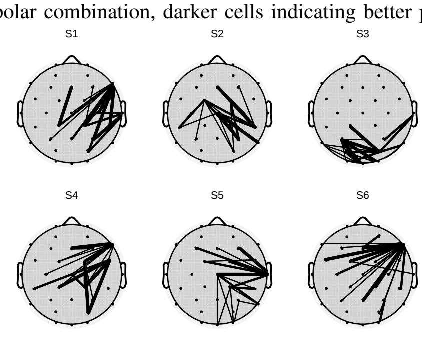

Fig. 4. Bipolar combinations for the largest twenty AUCs for each subject. The thickness of a line is proportional to its AUC value (quantized in three levels).

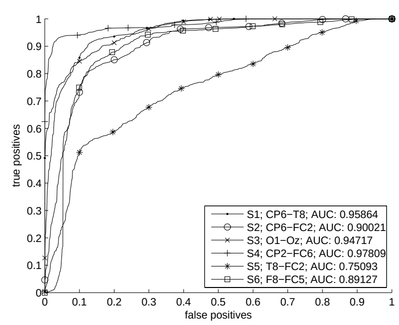

Fig. 5. ROC curves for the BBC of each subject.

The ROCs associated with the BBC for each subject are depicted in Fig. 5. All subjects but S5 have average AUCs above 0.89. The particularly low performance of subject S5 is expected given the lack of significant differences between responses to error and correct trials presented in Fig. 2.

D. Training set size

The number of elements in the training set needs to be sufficient to ensure a desired level of detection performance. It is generally true that a larger training set will result in better performance, yet we would like to limit the duration of the calibration procedure. To determine the influence of the training set size on performance, the following analysis was conducted. For each subject, the set of responses to error trials was randomly divided into a training and a testing set of equal sizes, along with an equal number of responses to correct trials. The detection algorithm was then run for the

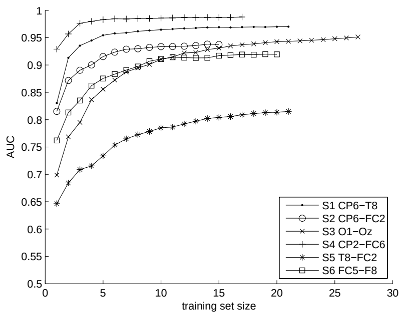

Fig. 6. Performance versus used number of training samples for the BBC of each subject.

subject's BBC with the first N elements (both erroneous and correct) of the training set, for N ranging from 1 to the size of the testing set for the particular subject. To estimate the average level of performance, this process was repeated 100 times with different random choices for the elements in the testing and training sets.

Fig. 6 shows how increasing the size of the training set improves the detection performance. The figure indicates that increasing the size of the training set beyond a certain size (about 10 responses for error and correct trials) leads to marginal gain in performance.

IV. Discussion

The suggested approach is suitable for real-time operation. This was tested in a demonstrator that used the paradigm explained in Section II-B. The signals were recorded between electrodes Fz and Cz (the BBC for the tested subject), analyzed real-time and whenever an error was detected the latest move was undone.

It is important to mention the time invariability of the reported potentials. The ErrPs of the second experiment examining the influence of frequency of error occurrence were very similar to the potentials from the original experiment, although the actual recordings were taken a few weeks apart. This was further confirmed by the fact that the real-time demonstrator functions with more than three months old data.

One could argue that the suggested experimental framework does not fully represent the spontaneity of machine errors in real-life. However, we did not find an effect of habituation when comparing the ErrP from the beginning of the experiment with the ones from the end of the experiment. Furthermore, considering the triviality of the task, the motivation of the participants could be to complete the whole session as fast as possible. So that, even if they do not care about winning or loosing after a certain point, every move backwards brings them a step further away from their goal, hence increasing their frustration. A few participants reported after the end of experiment that they found the task particularly annoying.

The effect of frequency of error occurrence was already addressed in the second experiment explained in Section III-A. Although we do not exclude the possibility that the element of surprise attendant to error trials might play a role in the error processing, the observed potentials do not seem to be due to the infrequency of the stimulus.

Hypothetically, the horizontal movement of the eyes could cause the differences between responses to error and correct trials. In that case one would expect the BBC to be defined by inter-hemispheric fronto-polar sites [7]. Yet, our results show involvement of right fronto-parietal sites, indicating that the reported differences are not caused by horizontal eye movement.

V. Conclusions and future work

In this paper we presented an approach for automatic detection of the neural correlates of error awareness in the human brain with the goal of improving the performance and usability of HMI systems. We have set a number of requirements for a practical solution that can be easily integrated in an existing HMI system. These requirements were real-time operation, accounting for individual specificities, and convenience of operation. Six subjects participated in an experiment, in which machine errors were simulated. Our results confirmed the presence of EEG potentials related to processing of machine error. We have implemented an error detection algorithm that achieves high error detection performance given by AUCs ranging from 0.75 to 0.98. The proposed solution sought the best individual bipolar combination of electrodes that emphasizes the differences in error processing. The impact of the training set size on the detection performance was investigated and it appears that only a few examples of brain responses to error and correct trials are sufficient for high performance. The feasibility of the proposed solution was tested using a real-time demonstrator.

More research is needed to evaluate the time invariability of the best bipolar combination for a given subject. In view of standard positioning of the measurement sites it is necessary to test for existence of a bipolar combination that guarantees reasonably high error detection rates for all users. To achieve this a larger group of subjects has to be involved.

VI. Acknowledgments

We thank Bas Zoetekouw for his help reviewing this paper and our participants for their kind collaboration.

References

[1] 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, pp. 747–749, May 1998.

[2] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, "Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 6, pp. 447 – 455, 1991.

[3] C. B. Holroyd and M. G. Coles, "The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity," *Psychol Rev*, vol. 109, pp. 679–709, Oct 2002.

[4] G. Schalk, J. R. Wolpaw, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: presence of an error potential," *Clin Neurophysiol*, vol. 111, pp. 2138–2144, Dec 2000.

[5] P. W. Ferrez and J. del R Millan, "Error-related EEG potentials generated during simulated brain-computer interaction," *IEEE Trans Biomed Eng*, vol. 55, pp. 923–929, Mar 2008.

[6] BioSemi. BioSemi, Amsterdam, The Netherlands. [Online]. Available: http://www.biosemi.com/

[7] I. Erkens, "Artifact detection and correction in Neurofeedback and BCI applications," Philips Research, Tech. note TN-PR 2008/00409, 2008.