A Comparative Psychophysical and EEG Study of Different Feedback Modalities for HRI

Xavier Perrin
ASL Lab, ETHZ
Zürich, Switzerland
xavier.perrin@mavt.ethz.ch

Ricardo Chavarriaga
IDIAP Research Institute
Martigny, Switzerland
rchava@idiap.ch

Céline Ray
ASL Lab, ETHZ
Zürich, Switzerland
celine.ray@mavt.ethz.ch

Roland Siegwart
ASL Lab, ETHZ
Zürich, Switzerland
rsiegwart@ethz.ch

José del R. Millán
IDIAP Research Institute / EPFL
Martigny, Switzerland
jose.millan@idiap.ch

ABSTRACT

This paper presents a comparison between six different ways to convey navigational information provided by a robot to a human. Visual, auditory, and tactile feedback modalities were selected and designed to suggest a direction of travel to a human user, who can then decide if he agrees or not with the robot's proposition. This work builds upon a previous research on a novel semi-autonomous navigation system in which the human supervises an autonomous system, providing corrective monitoring signals whenever necessary.

We recorded both qualitative (user impressions based on selected criteria and ranking of their feelings) and quantitative (response time and accuracy) information regarding different types of feedback. In addition, a preliminary analysis of the influence of the different types of feedback on brain activity is also shown. The result of this study may provide guidelines for the design of such a human-robot interaction system, depending on both the task and the human user.

Categories and Subject Descriptors

H.1.2 [**Models and principles**]: user/machine systems; H.5.2 [**Information interfaces and presentation**]: user interfaces; I.2.9 [**Artificial intelligence**]: Robotics

General Terms

Human Factors

Keywords

Auditory feedback, visual feedback, vibro-tactile feedback, robot navigation, brain-computer interface, multimodal interaction

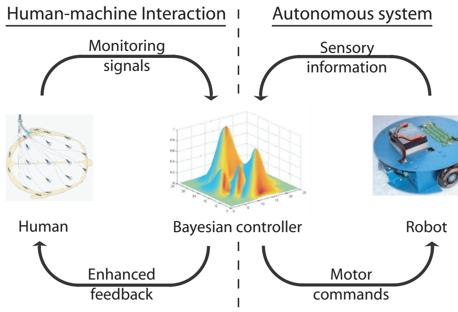


Figure 1: Scheme of the proposed semi-autonomous navigation concept.

1. INTRODUCTION

Robots interacting with humans should be able to receive commands as well to communicate relevant information such as perceptual input or internal states back to the human. This is even more relevant for robots helping humans, as in service robotics. We have proposed a new concept for semi-autonomous navigation for disabled people where the user relies mainly on the machine and provides only corrective monitoring signals when needed [17]. In this approach, shown in Figure 1, the robot is endowed with autonomous capabilities (depicted on the right part) and can interact with the human in order to reach the user's desired goal. For instance in navigational tasks, at each relevant place in the environment (e.g., crossroads), the robot chooses a direction of travel according to local environmental information and to the previously learned human-machine interaction. This choice is then submitted to the human user for approval before its execution. In general, the human monitors the activity of the robot and provides a corrective signal whenever the robot proposition differs from the user's desired action.

In this approach designed for daily use, it is then crucial to reliably communicate the controller's decisions to the human user. This paper seeks to compare different ways to provide such communication when completing a navigational task. We will explore the three major feedback modalities, namely visual, auditory, and tactile feedback. Three different kinds of visual feedbacks (i.e., icons, spatially located squares, and text), two of auditory (i.e., spatially located sounds and voice) and one tactile (i.e., spatially located actuators) are tested. Both quantitative (accuracy and rapidity) and qualitative (user feelings about the feedback types) information was used to assess their suitability.

In the long term, the described approach for sustained human-robot interaction might also be implemented using Brain-Computer Interfaces (BCI), where brain signals are translated onto robot commands. To this end, we perform a preliminary analysis of electroencephalographic (EEG) signals generated while performing the described task, with a particular focus on signals generated when human subjects supervise the robot's decisions.

In the next section, a brief review of the state of the art of different feedback principles and input systems is proposed. Then, in section 3, we will describe our experimental setup and the protocol used for collecting information. Section 4 will present the results, which are discussed in section 5. Finally, we will conclude with a summary and an outlook on future work.

2. STATE OF THE ART

The proposed semi-autonomous navigation system implies a closed human-machine communication loop. On one side, the robot's decisions are sent to the user (i.e., machine-to-human communication), and on the other side, the human sends monitoring signals to the controller (i.e., human-to-machine communication). This section contains a brief review of related work with a focus on their application in brain computer interfaces.

2.1 Input systems

Common input systems for human-machine interaction range from keyboards, joysticks, and touch screens to devices more adapted to disabled persons, like voice commands, eye-tracking, or sip and puff systems [20, 24].

In recent years, a novel technology has been studied, namely brain-computer interfaces (BCI). Non-invasive, electroencephalography (EEG) based BCIs rely on the decoding of brain activity in order to manipulate robotic devices, virtual keyboards, or more general computer applications [15, 22]. It allows the use of a new communication channel, the brain, instead of limb or eye movements or voice commands.

The work done by Ferrez and Millán [7, 8] about error potentials is a recent addition to the available decoded brain-commands for human robot interaction. This potential indicates the human awareness of an erroneous response made by the system when classifying the user intention. In the experiments we are presenting, we will study the influence of the different feedback modalities on the recognition of this error potential.

The choice of an input system depends on the human user and also on the targeted application. Information from the input system can be either discrete, for buttons or voice command, or continuous, for joysticks or eye-tracking systems. BCI systems can be designed to provide both types of inputs. As opposed to continuous input systems, discrete input systems typically encode fewer control commands, thus having a lower information bit rate, therefore they are used to convey high-level information, as in our proposed human-robot interaction scheme. The input processing has to be designed accordingly, as well as the information provided to the user. Moreover, we have shown that a semi-autonomous approach where the user emits corrective actions yields higher information transfer rates than explicit navigational discrete commands [17].

2.2 Feedback systems

In Human-Computer Interactions, Brewster have proposed the transcription of visual information to auditory or tactile representation. His so-called *earcons* [4] or *tactons* [3] try to imitate the use of visual icons to represent files, folder, menus, or actions. Specific tacton (respectively earcon) patterns are created by modifying the frequency (pitch), amplitude (intensity), duration, rhythm, or body location (timbre) of the stimuli. Vibro-tactile stimuli have also been studied for providing spatial information to the human about directions to explore in a building-clearing task [13].

EEG studies have largely used stimulus presentation in order to provide feedback of the subject's performance of a task or to provide a cue to react to. Visual feedback is widely used, as it is considered a natural communication channel. Auditory feedback is a good alternative or a complement to the visual one. Vibro-tactile (haptic) feedback is nowadays getting more and more attention due to the novelty and to the potential applications it has. Feedback has also an influence on brain activity and has to be carefully designed for optimal usage [14].

When providing visual feedback about EEG signal classification, the performance of a user can be displayed either with bars, lines, moving cursor, or icons [7, 11, 23]. Hinterberger et al. experimented the usage of audio melodies for indicating the output of the EEG classification [10, 11] or for representing the brain activity itself with the sonification of the EEG in real time [9]. Vibro-tactile systems have recently been studied in comparison with visual feedback [12]. They show no significant difference for the realisation of the task, but do allow the visual channel to be freed up or to complete the information flow.

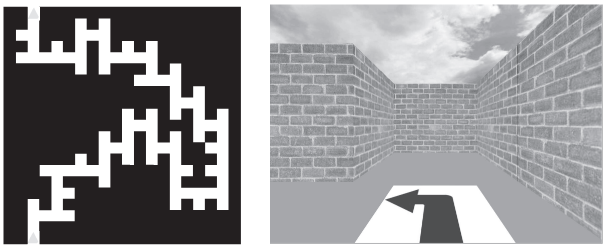
Some brain signals, the so-called evoked potentials, appear as responses to external stimuli. For instance, the P300 signal is a positive EEG deflection 300 ms after stimulus onset, elicited when a significant, rare stimulus appears in a sequence of frequent other stimuli. There are numerous applications on the use of P300 using both visual [5, 18] and auditory [19] stimuli.

In recent years, there has been an increased interest in studying brain activity during real-world experiences. In particular, virtual reality techniques have been applied to both functional magnetic resonance imaging (fMRI) [21] and EEG studies [2]. This work follows the same approach and aims to exploit the high temporal resolution of EEG, as opposed to fMRI higher spatial resolution, in realistic situations.

3. METHODS

Twenty-two subjects (6 women) aged 24-52 (mean 30.18, std. dev. 5.87) participated in the experiment. All participants gave informed written consent before the experiment. In addition, EEG activity was recorded for four of these subjects while they performed the task.

During the experiment, subjects are asked to monitor the decisions of a simulated robot navigating in a virtual maze (Figure 2). The correct trajectory is indicated by arrows drawn on the floor of the corridors. At each junction, the controller decision is presented to the user using one of six



(a) (b)

Figure 2: (a) Example of maze with T crossings only. (b) Subject's view of the maze. The visual cue on the floor shows the correct direction.

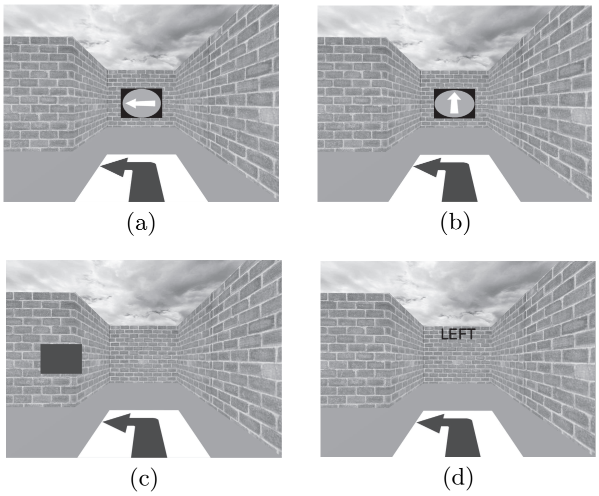


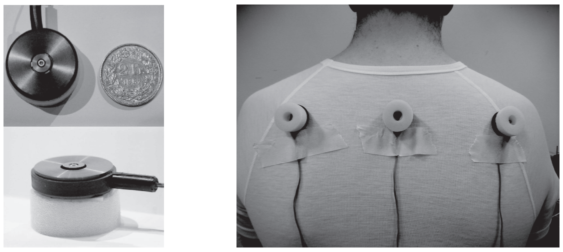
Figure 3: Examples of visual feedback. Pictograms (V1); (a) correct and (b) erroneous feedback. (c) Squares (V2) and (d) text feedbacks (V3).

different feedback modalities. The subject is asked to press a key whenever an erroneous decision is presented. Each modality was tested separately and user responses and reaction times were recorded for each condition. Moreover, verbal reports were acquired before, during, and after the experiment to assess the user's preference among the different modalities.

3.1 Types of Feedback

Six different types of feedback were tested to convey the robot's navigation decisions, i.e., *turn left*, *turn right*, and *forward*. At each point, one of three decisions is presented to the subject. The tested feedback conditions were:

- *Visual pictograms (V1).* An icon containing an arrow (pointing *left*, *right*, or *up*), similar to a traffic sign, is shown in the center of the screen (c.f. Figure 3a,b).
- *Visual squares (V2).* Colored squares are shown at the *left*, *right*, or *center* of the screen (c.f. Figure 3c).
- *Visual text (V3).* The words *left*, *right*, and *forward* are shown in the center of the screen (c.f. Figure 3d).



(a) (b)

Figure 4: Vibro-tactile actuator (a) front and side view and (b) placed on the body.

- *Auditory tones (A4).* Sound tones, spatially localized to the *left*, *right*, or *center* of the user, were played back through stereo headphones. The same tone was used for the three conditions so as the information was solely provided by the spatial localization.
- *Auditory words (A5).* A pre-recorded voice informed the user about the controller's decision pronouncing the words *left*, *right*, and *forward*.
- *Tactile (T6).* Vibro-tactile actuators[1] were located in the upper back of the user (c.f. Figure 4). The electromagnetic devices provide a short vibration to the subject, i.e., an oscillation at 200 Hertz during 50 milliseconds.

The motivation for the selection of these feedback types was to cover the major human sensory channels (sight, hearing, touch) and to make the association stimulus-command evident. This avoids the need for the subjects to learn the association and reduces the risk of misinterpretation.

3.2 Experimental Protocol

Throughout the experiment, subjects were asked to grade the different types of feedback with respect to the following adjectives: *(i)* Natural, *(ii)* Understandable, *(iii)* Not Disturbing, and *(iv)* Pleasant. The whole experiment consists of three phases and lasts around 90 minutes [2].

3.2.1 Preliminary Measures

This phase is intended to establish a priori feedback preferences of the subjects before actually experiencing the maze navigation. After hearing the description of the experiment, subjects were asked which type of feedback ranks first and last for the above mentioned criteria. Moreover, the response time needed for pressing a key as soon as a stimulus was presented on a black screen was measured.

A control situation was also included where the subject actively drives the robot through the maze by pressing one of the arrow keys according to the visual cue on the floor. This helps the user to get used to the 3D maze environment and the cues to follow. This manual driving is similar to the semi-autonomous navigation strategies encountered in robotics: at each relevant place, i.e., crossings, the robot

[1] From Engineering Acoustics, INC., Fl, USA [6]

[2] Experimental sessions involving EEG recordings lasted about 60 minutes more to account for electrode placement, subject preparation, and additional trials per condition.

waits for a direct order from the user [1, 24]. This test serves as a basis of comparison for the reaction times and for the percentage of correct orders.

3.2.2 Semi-Autonomous Maze Navigation

In this phase the subjects have to monitor the robot's decisions while navigating inside the virtual maze with a first-person view. In order to focus our attention on the human-robot interaction, the robot is following predefined paths. This prevents a long learning period of the rather complex environment by the subjects. This allows furthermore to correctly label the sample, to run more experiments than with a real robot, and even to involve novice users without previous experience with robots (which otherwise might appear to drive chaotically and thus distract their attention). Three different mazes were designed, all of them consisting of 36 binary intersections. The robot controller was set up so as to take erroneous decisions in 40% of the trials. The speed of the robot was set so as to spend three seconds between two successive intersections.

When the robot arrives at an intersection, it proposes an action using one of the feedback types described in Section 3.1, waiting then for one second for the user's response. The subject is expected to press a key whenever the controller makes an error. The system also informs the user when he/she does not respond adequately (i.e., either by pressing the key when there is no error -false positive- or not pressing when an error occurs -true negative). If the subject does not press a key within one second after an error, the trial is also counted as a failure. In all cases, the robot will follow the correct trajectory towards the exit of the maze.

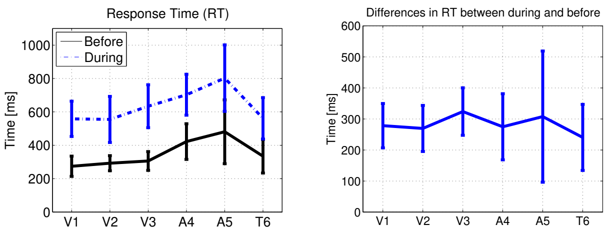
Subjects whose EEG brain activity was recorded did the three mazes for each condition (resulting in about 65 correct and 43 erroneous feedbacks) while the rest of the subjects choose two out of the three mazes (43 correct and 29 erroneous feedbacks). The order of mazes and feedback types changed randomly across different subjects. After the subject completed the task for each condition, the operator asked him/her to grade it, again according to the four criteria already mentioned (i.e., Natural, Understandable, Not Disturbing, and Pleasant). The scale ranged from one for a perfect fit to the criteria (e.g., very pleasant), to five for the opposite (e.g., very unpleasant), three being neutral.

3.2.3 Final Debriefing

Once the task has been completed for all the six different types of feedback a post-experimental debriefing takes place. The subject is asked to rank all the conditions in order to obtain a posteriori preferences.

3.3 EEG recordings

We recorded EEG activity during the experiment for four of the subjects (mean age 26.5; SD 1.0) using a Biosemi acquisition systems (www.biosemi.nl). Data was acquired with a sampling rate of 512 Hz using 64 electrodes according to the standard 10/20 international system. Signals were then re-referenced to the common average reference (CAR) and a 1-10Hz band-pass filter was applied. External triggers were sent to the acquisition board of the EEG system by the experimental software to timestamp relevant events (i.e., experiment start/stop, feedback delivery, and user's response)



(a) (b)

Figure 5: (a) Response time in milliseconds for the preliminary tests and during the experiments. (b) Difference between the response times calculated for each subject. See section 3.1 for the meaning of the feedback labels.

EEG data was segmented into epochs corresponding to each maze intersection and baseline activity computed in the 100 milliseconds *prior* to the feedback onset. Epoch activity (in the time domain) was then used to classify correct and erroneous trials. We use a linear logistic classifier trained using an iterative recursive least square algorithm [16]. This classifier is trained using EEG samples from a particular time window (where the phenomena are expected to occur) where each sample is considered independent.

Separate classifiers were designed per subject and type of feedback. Classification performance was assessed using 10-fold cross validation. Moreover, in order to emulate realistic operating conditions of a BCI device, no artifact rejection was applied and all the trials were included in the analysis.

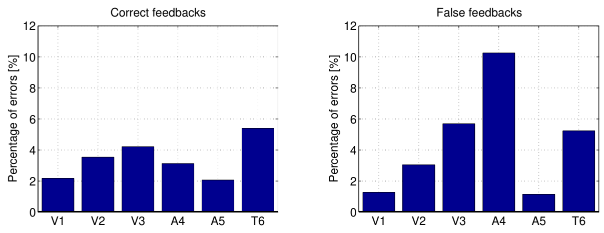
4. RESULTS

This section presents the different performance measures results. Reaction time, user performance, user feedback evaluations, and EEG signals are analysed in order to assess the suitability of each feedback type.

4.1 Reaction time

Figure 5a shows the reaction time during both the preliminary measures and during maze navigation for all types of feedback. For the preliminary measures, the visual feedback types have the shortest response time and also the most uniform (small standard deviation), followed by vibrotactile feedback (T6). No significant difference was found between them (ANOVA test, $\alpha < 0.05$, followed by a multiple-comparison Bonferroni correction). The two auditory feedback types have the longest response time, the voice cues being the slowest and having the biggest standard deviation. This last result can be explained by the fact that subjects reacted either as soon as they heard something or at the end of the word.

When we look at the reaction times during the experiments, we can see that V1, V2 and T6 elicit the quickest responses when performing the monitoring task (no statistical difference). A possible explanation is that the position of the squares on the screen (V2) reflects immediately the proposed direction. As the subject already knows where to go, the decision to accept or reject the proposition is immediate. When displaying pictograms (V1), the subject has to do one more step, i.e., a pattern matching process between the requested direction and the proposed one. When providing



(a) (b)

Figure 6: Percentage of erroneous responses made by the subjects for (a) correct and (b) false robot propositions. See section 3.1 for the meaning of the feedback labels.

the information by mean of vibrations (T6), the processing is similar to that in V2 but slightly slower. Having to decode text (V3), auditory cues (A4) or words (A5) in order to compare with the desired direction takes significantly more time. As before, some subjects also waited until the end of the word before acting and some not, thus explaining the large standard deviation. The two auditory feedback types are statistically the slowest.

Figure 5b represents the difference between the mean response time before and during the experiments, calculated for every subject. It highlights the larger variability in the time required to process the voice feedback (A5) and the change in the feedback's processing. But there is no statistical difference among the different conditions.

4.2 Understandability - User performance

Figure 6 gives us insights into the understandability of each feedback type by displaying the percentage of erroneous responses made by the subjects for each feedback type, given a correct or erroneous proposition from the robot. The poorest performance corresponds to 3D sound (A4). Users reported difficulty in distinguishing the center tone from the side ones, thus explaining the percentage of false answers. A different tone per cue would have helped.

The performance with the voice feedback (A5) is the best, closely followed by the pictogram cues (V1). Although the auditory proposition takes longer to process, the user are more self-confident about the answer. The traffic signs have two characteristics which make them easy to process adequately: they are well-known by the subjects (from their car-driving experiences) and they are similar to the cues contained in the maze. The result of the text feedback (V3), with the second worst performance, differs greatly from the voice feedback. A possible explanation is the fact that the text was not distinguishable enough from the background.

It can be argued that subjects would make more mistakes at the beginning of a new maze because they would have to remember the task to be solved or to adapt themselves to the new type of feedback. Comparison of the error percentages in the full maze (as shown in figure 6), or after removing the first five, ten, and fifteen crossings showed no significant difference (data not shown). This may suggest that no specific training was required for the different feedback types.

The performance of the subjects when providing a navigational command at each crossing reached 99.6%. It can be deduced that the cues placed in the maze and the first-

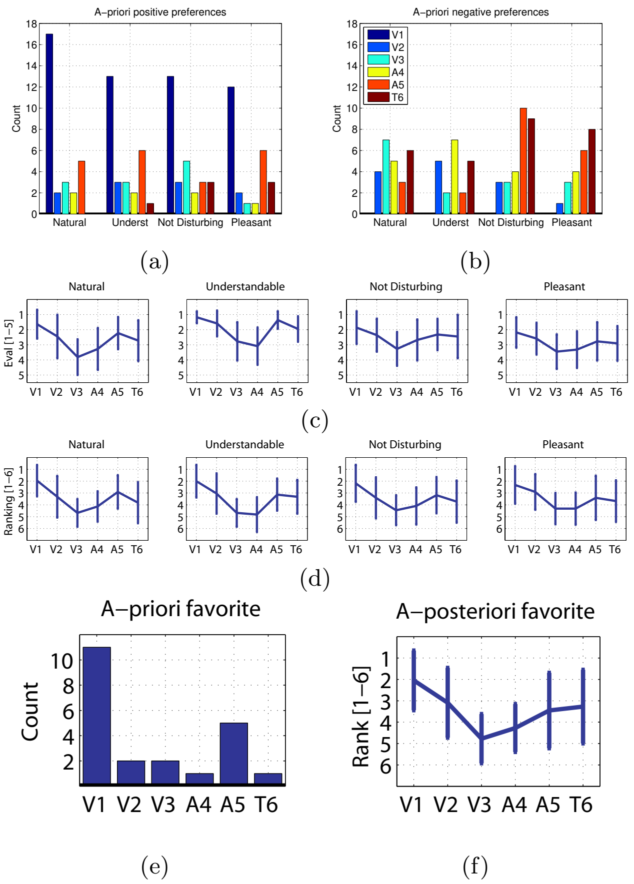


Figure 7: Qualitative assessment of the different feedback types: a-priori evaluation of (a) positive or (b) negative preferences, (c) evaluation during the tests and (d) a-posteriori rank. (e) A-priori preferred feedbacks and (f) a-posteriori preferences. A ranking or an evaluation of 1 corresponds to the best grade. See section 3.1 for the meaning of the feedback labels.

person view were well assimilated by the subjects and do not perturb the navigational decision process.

4.3 Verbal reports

When asked for their a-priori best and worst feedback types according to our criteria, the participants in the tests widely agreed on the fact that the visual feedback using signs (V1) should be the best: it is natural, understandable, not very disturbing, and pleasant (see figure 7a,e). Although the voice (A5) is ranked as second favorite, second most natural, and second most understandable, people would find it disturbing and unpleasant for daily use (figure 7b). Tactile feedback (T6) is little known and sometimes dreaded. Its scores were the most negative. Users didn't show any particular trend with respect to other feedback types.

If we compare the feedback types as evaluated during the tests (figure 7c), one can see that V1 is still strongly preferred. Voice (A5), square (V2), and tactile feedback (T6) are statistically the next-best rated feedback types. The lowest scores, according to our four criteria, were given to text (V3) and sound feedback (A4). This is mostly due to the

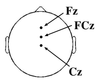
Subject	Time Window	Electrodes	Elect. location
1	[150ms, 450ms]	Fz, FCz, Cz	
2	[200ms, 450ms]	Fz, FCz	
3	[200ms, 450ms]	Fz, FCz	
4	[200ms, 450ms]	Fz, FCz, Cz	

Table 1: Selected electrodes and time windows used for EEG classification. Rightmost column shows the location of the electrodes used for classification on the subject's scalp.

fact that the people had to concentrate in order to properly read and understand what was written on the screen and that it was difficult to discriminate the different auditory stimuli solely by spatial location. Some subjects reported their wish to have a different sound for each stimulus.

The a-posteriori ranking of the different feedback types (Figure 7d) reflects their evaluation during the tests but strengthens their differences. From the statistics and figure 7f, showing the overall a-posteriori ranking of the six feedback types, we can order the feedbacks by preference: pictograms (V1), followed by squares (V2) and voice (A5), followed by tactile (T6), followed at the end by text (V3) and sounds (A4).

4.4 EEG recordings

We want to assess whether event related potentials elicited by the different types of feedback can be used to classify erroneous and correct trials. Based on previous studies [7], we focus the analysis on electrodes located in fronto-central areas (i.e. electrodes Fz, FCz and Cz). Moreover, to discard the use of motor activity in the classification, the signal of the first 150ms after the feedback onset was not taken into account. Independent classifiers were designed for each individual. Table 1 show the selected electrodes and time windows chosen for each subject. Except for subject one, classification was based on the time signal from 200ms to 450 ms after the feedback onset.

Classification performance for each condition and subject is shown in Figure 8. Each point in the figure, represents the mean classification performance (10-fold cross validation) in the ROC (*Receiver Operating Characteristic*) space. In this space, the x-axis corresponds to the false positive rate, FPR (i.e., misclassified error trials), while the y-axis corresponds to the true positive rate, TPR (i.e., properly classified correct trials). The performance of a perfect classifier corresponds to the point (0,1), i.e. upper left corner in the plot, while random classification lies along the diagonal line.

It can be observed that, with the exception of subject 4, classification above random levels is obtained for several types of feedback; in general, a higher classification rate was obtained for correct rather than for error trials. The best classification rates were obtained for the tactile feedback (T6) in subject 2 ($TPR = 0.76$; $FPR = 0.30$); and visual squares (V2) in subject 3 ($TPR = 0.71$; $FPR = 0.40$).

In contrast, text feedback (V3) yields near-random performance in three of the subjects. This type of feedback was lowest ranked and most unnatural during the verbal reports, several subjects pointing out that this stimulus was not salient enough and the interpretation of the text message required extra cognitive processes.

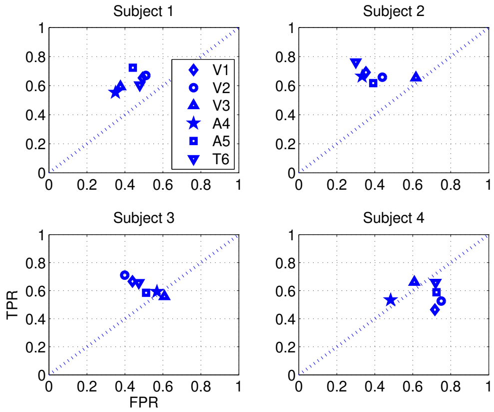


Figure 8: EEG classification performance for all subjects and types of feedback. Mean classification performance (10-fold cross validation) is shown in the ROC space. TPR = True positive rate, FPR = False positive rate. See section 3.1 for the meaning of the feedback labels.

These preliminary results suggest that, in certain conditions, it is possible to recognize EEG activity related to the recognition of errors. It should be noticed that a simple classifier was used in this study, and more powerful techniques might provide better classification results. In particular, we plan to apply Gaussian classifiers which have been previously used to recognize error potentials in BCI applications [7].

5. DISCUSSION

Using a virtual reality environment, we have presented a qualitative and quantitative evaluation of different types of feedback used to communicate a robot's navigational decisions to a human user. The human subject acts as a *critic* of the robot, sending corrective signals whenever the robots makes a wrong decision. Moreover, preliminary analysis of EEG signals elicited in this task is also provided as a first step towards future implementation using Brain-Computer Interfaces.

In general, the visual pictograms (V1) is the most liked feedback and the one providing the quickest and best answers. It is not a surprising result given the structured world we live in. "A picture is worth a thousand words": carefully selected pictograms transmit immediately the desired information. This is especially true in our test environment, were the cues pointing to the exit are the same as the provided feedback. This facilitates decision making, but not detract from it in a normal environment where subjects know where they want to go. Furthermore, it is easy to design new pictograms for other navigational commands, like entering a door, docking at a desk, or making a U-turn for example. Pictograms are therefore the most convenient feedback types from a user and communication point of view.

Spatially placed squares (V2) and text (V3) elicit quick responses as well but are less accurate. They are also less appreciated by the users, the text having the worst a-posteriori rank. An extension of the available feedbacks to new commands is easy for the text, but the design of this feedback type should be improved in order to have more distinguishable stimuli and reduce the amount of erroneous responses. Concerning the squares, new commands would imply specific new locations or specific successions of squares, thus needing some learning for the human subject. Pictograms and text feedback could fit on a relative small display, whereas localized squares (or any type of icons) may require a bigger display. Alternatively, squares could be replaced by lights when transferred onto a real mobile platform, thus simplifying the interface from a technical point of view and clearing the field of view. However, for a larger set of commands, squares might not be an appropriate type of feedback.

Voice feedback (A5) did not leave anybody indifferent. Its appreciation and the time used to process the auditory information differed greatly from subject to subject. Interestingly, it had the best percentage of correct responses, arguably in association with its largest reaction time. As noted by some subjects, this feedback doesn't have a pop-up effect that might trigger false responses. Thus it could be a precise but slow feedback system, easily extendible to further navigational commands. Subjects had also different strategies in processing the stimuli. Some reacted at the beginning of the words as a result of a short learning phase. The present study aimed also at understanding the processing of spatially places tones (A4) and the results are worse than expected, mainly due to the fact that the different stimuli were not distinguishable enough. Subjects took more time before reacting in order to be sure of their answer, which was often even then wrong. Therefore they mostly agreed on the second-worst rank of this feedback type. A different tone per stimulus would have helped a lot but would have required a learning phase, which we wanted to avoid. This learning phase may also be required if the number of commands increases, demanding a higher number of tones, as is the case for the *earcons* [4].

Similarly ranked than voice feedback (A5), vibro-tactile feedback (T6) performed worst in terms of the amount of erroneous responses. Some people reported difficulties in feeling the actuator placed in the center of the back and suggested placing it on the chest. However, reaction time for vibro-tactile feedback was the second best after visual stimuli. Thus, if we remove the visual feedbacks from the available ones in order not to clutter the visual channel, to leave intact the visual field of view, or to ease the technical apparatus, tactile and voice feedback seem to be a sensible choice. The user would also have more freedom to interact with people or to enjoy the surroundings if he/she does not have to concentrate on a screen. On the other hand, the extension of navigational commands would lead to more complex haptic stimuli, like the *tactons* [3], requiring a learning process. It would thus give an advantage to voice feedback despite its slower reaction time.

The comparison between a-priori feelings and a-posteriori ranking reveals a real change in opinion about vibro-tactile (T6) and text feedback (V3). Subjects were not familiar with vibro-tactile feedback, thus producing a poor a-priori evaluation. After the experiments, they discovered that it was more suitable for such a human-robot interaction than thought. Although not enough female subjects were included in the study to assess gender specific preferences, we found that men tend to positively evaluate the tactile feedback more than women. However, a further study would be required to confirm this finding. As mentioned before, text feedback was not well appreciated because of the poor saliency of the stimuli, thus resulting in the change in opinion. In a similar way, there was a large variability in the a-priori evaluation of voice feedback (A5), but it decreased after the experiment. In the end, A5 was selected as one of the best feedback types other than the visual ones.

It should be noted that the task to be solved by the subject was not very natural to them at the beginning. People reported the desire to provide an input for agreeing instead of disagreeing with the robot's proposition. One could have thus expected worst performances at the beginning of a new modality or of a new maze, which didn't appear significantly in the results.

Additionally, we present preliminary results on the classification of EEG signals elicited by the different types of feedback. Consistent with the subject's qualitative evaluations, visual text feedback (V3) has the poorest classification performance. The fact that the stimuli were not salient enough and the required decoding process may cause event-related potentials to not be well synchronized across trials, makes its recognition more difficult.

Although the classification performance obtained is not very high (especially for erroneous trials), these results, using a simple classifier with no artifact rejection, constitute a promising basis to further explore the use of BCI systems in this type of human-robot interaction. We plan to extend the present study by including a larger number of subjects and comparing different classification algorithms.

Finally, a fine adaptation of every feedback system presented in this paper to each user is required, as there is a large inter-subject perceptual variability. Adaptation to possible sensory impairments as well as to the personal feelings of the human is also required. For example, a female voice could be preferred by a male user, a male one by a female user, or a different tone intensity in each ear. The more adapted the human-robot interaction is, the better the results.

In summary, the present paper provides a detailed comparison between the user's perception of different feedback modalities for human-robot interaction, confirming some de facto hypotheses but also providing new information about less common types of feedback, like vibro-tactile actuators. Some guidelines for the design of feedback systems or for increasing the number of the available commands are also brought to light. The next step of the study will be to assess the learning curves for each modality as another aspect of a feedback's adequacy for human-robot interaction. Having more subjects will contribute to the refinement on the psychophysical aspects of the findings, but especially for the recognition of the error potential in brain activity. As previously mentioned, more women participating in the experiments could bring better insights into the inter-subject or even the inter-gender variability of the perception of the different stimuli. We will also perform experiments in an office environment where subjects will have to visit different places, either along self-generated or fixed paths.

The present study used explicit cues to signal the correct path to follow in order to have a well-controlled experimental setup, i.e. the correct labeling of erroneous robot decisions. In future work, we will reproduce this experiment using the user's own representation of the correct path (e.g. the subject learns the correct trajectories prior to the experiments), which is closer to realistic situations.

6. ACKNOWLEDGMENTS

Virtual mazes were created using the Virtual Environments Library (veLib) developed at the Max Planck Institute for Biological Cybernetics (http://velib.kyb.mpg.de). Localized audio stimuli was generated using openAL 3D audio API (http://www.openal.org). This work has been supported by the Swiss National Science Foundation NCCR 'IM2' and by the EC-contract number BACS FP6-IST-027140. This paper only reflects the authors' view and funding agencies are not liable for any use that may be made of the information contained herein.

7. REFERENCES

[1] A. Argyros, P. Georgiadis, P. Trahanias, and D. Tsakiris. Semi-autonomous Navigation of a Robotic Wheelchair. *Journal of Intelligent and Robotic Systems*, 34:315–329, 2002.

[2] J. D. Bayliss and D. H. Ballard A virtual reality testbed for brain-computer interface research. *IEEE Trans Rehab Eng*, 8(2):188–190, 2000.

[3] S. Brewster and L. Brown. Tactons: Structured Tactile Messages for Non-Visual Information Display. In *Proceedings of AUIC 2004*, pages 15–23, Dunedin, New Zealand, 2004.

[4] S. Brewster, P. Wright, and A. Edwards. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of InterCHI'93*, pages 222–227, Amsterdam, 1993.

[5] E. Donchin, K. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Trans Rehab Eng*, 8(2):174–179, 2000.

[6] EAI. Engineering Acoustics, INC. *http://www.eaiinfo.com*, 2005.

[7] P. Ferrez and J. d. R. Millán. You Are Wrong!—Automatic Detection of Interaction Errors from Brain Waves. In *Proc of the 19th Int Joint Conf on Artificial Intelligence*, Edinburgh, UK, 2005.

[8] P. Ferrez and J. d. R. Millán. Error-Related EEG Potentials Generated During Simulated Brain-Computer Interaction. *IEEE Trans Biomed Eng*, To appear.

[9] T. Hinterberger and G. Baier. Parametric Orchestral Sonification of EEG in Real Time. *IEEE MultiMedia*, 12(2):70–79, 2005.

[10] T. Hinterberger, G. Baier, J. Mellinger, and N. Birbaumer. Auditory feedback of human EEG for direct brain-computer communication. In *Proc of the 10th Meeting of the Int Conf on Auditory Display (ICAD 04)*, 2004.

[11] T. Hinterberger, N. Neumann, M. Pham, A. Kübler, A. Grether, N. Hofmayer, B. Wilhelm, H. Flor, and N. Birbaumer. A multimodal brain-based feedback and communication system. *Experimental Brain Research*, 154(4):521–526, 2004.

[12] L. Kauhanen, T. Palomäki, P. Jylänki, F. Aloise, M. Nuttin, and J. d. R. Millán. Haptic Feedback Compared With Visual Feedback For BCI. In *Proc of the 3rd Int Brain-Computer Interface Workshop & Training Course*, Graz, 2006. Verlag der TU Graz.

[13] R. Lindeman, J. L. Sibert, E. Mendez-Mendez, S. Patil, and D. Phifer. Effectiveness of Directional Vibrotactile Cuing on a Building-Clearing Task. In *Proceedings of the ACM CHI 2005*, Portland, 2005.

[14] D. McFarland, L. McCane, and J. Wolpaw. EEG-based communication and control: short-term role of feedback. *IEEE Trans Rehab Eng*, 6(1):7–11, 1998.

[15] J. d. R. Millán, F. Renkens, J. Mourino, and W. Gerstner. Non-invasive brain-actuated control of a mobile robot by human EEG. *IEEE Trans Biomed Eng*, 51(6):1026–1033, 2004.

[16] L. C. Parra, C. D. Spence, A. D. Gerson and P. Sajda. Recipes for the linear analysis of EEG. *Neuroimage*, 28(2):326–341, 2005.

[17] X. Perrin, R. Chavarriaga, R. Siegwart, and J. d.. Millán. Bayesian Controller for a Novel Semi-Autonomous Navigation Concept. In *3rd European Conference on Mobile Robots (ECMR 2007)*, Freiburg, Germany, 2007.

[18] B. Rebsamen, E. Burdet, C. Guan, H. Zhang, C. L. Teo, Q. Zeng, M. Ang, and C. Laugier. A Brain-Controlled Wheelchair Based on P300 and Path Guidance. In *IEEE Int Conf on Biomedical Robotics and Biomechatronics (BioRob'06)*, pages 1101–1106, 2006.

[19] E. W. Sellers and E. Donchin. A P300-based brain-computer interface: Initial tests by ALS patients. *Clinical Neurophysiology*, 117(3):538–548, 2006.

[20] R. Simpson and S. Levine. Adaptative Shared Control of a Smart Wheelchair Operated by Voice Control. In *Proc of the IEEE/RSJ Int Conf on Intelligent Robots and Systems (IROS '97)*, volume 2, pages 622–626, 1997.

[21] H. J. Spiers and E. A. Maguire. Decoding human brain activity during real-world experiences. *Trends in Cognitive Sciences*, 11(8):356–365, 2007.

[22] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-Computer Interfaces for Communication and Control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

[23] J. Wolpaw, H. Ramoser, D. McFarland, and G. Pfurtscheller. EEG-based communication: improved accuracy by response verification. *IEEE Trans. on Neural Systems and Rehabilitation*, 6(3):326–333, 1998.

[24] H. A. Yanco. Wheelesley, A Robotic Wheelchair System: Indoor Navigation and User Interface. In V. O. Mittal, H. A. Yanco, J. M. Aronis, and R. C. Simpson, editors, *Assistive Technology and Artificial Intelligence*, Lecture Notes in Artificial Intelligence, pages 256–268. Springer-Verlag, 1998.