Non-invasive brain-computer interfaces: Enhanced gaming and robotic control

Reinhold Scherer[12], Elisabeth C.V. Friedrich[3], Christa Neuper[13], Brendan Allison[1], Markus Pröll[1], Mike Chung[2], Willy Cheung[2], and Rajesh P.N. Rao[2]

[1] Institute for Knowledge Discovery, Graz University of Technology, Krenngasse 37, 8010 Graz, Austria
[2] Computer Science & Engineering, University of Washington, Box 352350, Seattle, WA 98195-2350, USA
[3] Department of Psychology,University of Graz, Universitätsplatz 2/III 8010 Graz, Austria
reinhold.scherer@tugraz.at

Abstract. The performance of non-invasive electroencephalogram-based (EEG) brain-computer interfacing (BCI) has improved significantly in recent years. However, remaining challenges include the non-stationarity and the low signal-to-noise ratio (SNR) of the EEG, which limit the bandwidth and hence the available applications. In this paper, we review ongoing research in our labs and introduce novel concepts and applications. First, we report on the long-term stability and robustness of detection of oscillatory components modulated by distinct mental tasks. Second, we present an enhancement of the 3-class self-paced Graz-BCI that allows interacting with the massive multiplayer online role playing game World of Warcraft. Third, we propose a scalable, adaptive learning framework, which allows users to teach the BCI new skills on-the-fly. Using this hierarchical BCI, we successfully train and control a humanoid robot in a virtual home environment.

1 Introduction

Brain-computer interface (BC) technology allows direct interaction with the environment by recording and translating the users brain activity in real-time. Electroencephalographic (EEG) signals, i.e., bioelectrical potentials recorded from the scalp, are the most common non-invasive source of brain signals in BCIs [1]. Major problems confronting BCI developers are the non-stationarity and inherent variability of EEG signals and the low signal-to-noise-ratio (SNR). These characteristics, among others, hamper reliable detection and translation of on-going EEG patterns into messages and hence limit the bandwidth. This is aggravated by the fact that the brain itself is a highly adaptive system. No a priori fixed mappings between the brain and the application exist. User feedback training and machine learning are required to optimize the interplay.

The two major types of EEG features used in BCIs are evoked potentials (EP) and oscillatory components (also known as event-related (de)synchronization, ERD/S). EPs are stereotypical brain responses to external perceptual events that are stable over time and require little adaptation on part of the user. EP-based BCIs do not require long training and achieve high information transfer rates (ITR, usually <30 bit/min), but rely on external stimuli to elicit EPs. Changes in oscillatory patterns can be induced internally, i.e., users intentionally perform distinct mental tasks (MT) to send specific messages to the BCI. ERD/S based systems require longer training and achieve lower ITRs, however, users can initiate communication whenever required.

In this paper, we review research currently being performed in our laboratories that aims to shorten ERD/S BCI training, reduce errors, and increase effective bandwidth through intelligent processing. By identifying the best control signals for each subject, and incorporating context awareness into more intelligent software, we can extend the control possible with BCIs. We identified the mental tasks that yield the best performance within subjects, and extended the self-paced 3-class Graz-BCI to a new BCI that can control the massive multi-player online role-playing game (MMORPG) World of Warcraft (Blizzard Entertainment, Inc.) We are also working in a new adaptive hierarchical architecture that allows user to teach the BCI new skills.

2 Evaluation of mental tasks for robust control

Kinesthetic motor imagery is the classical mental task (MT) used for controlling ERD/S BCIs. We started the systematic investigation of MTs of different modalities with the aim to identify MTs that are statistically stable over long periods of time (days) and thus may increase performance and reduce training time [2]. In order to identify such MTs, we recorded thirty channel EEG data from 9 able-bodied female subjects during 7-s cue-guided imagery trials on four different days. Electrodes were evenly distributed over the head. Thirty trials per class were recorded on each day in randomized order. MTs included:

1. Mental rotation (ROT): Visualization of 3-dimensional L-shaped figure rotating in the 3-d space.
2. Word association (WORD): Generation of as many words possible beginning with the presented letter (e.g. B = bank, bold, buy, etc).
3. Auditory imagery (AUD): Imagination of listening to a familiar tune (melody) without articulating the words.
4. Mental subtraction (SUB): Calculation of successive elementary subtractions from the presented problem (e.g. 105-6 = 99, 99-6 = 93, etc).
5. Spatial navigation (NAV): Imagination of navigating through a familiar house (flat) thereby focusing on orientation and furnishing.
6. Imagery of faces (FACE): Imagination of the face of the best female friend.
7. Motor imagery (MI) of the right hand: Imagination of repetitively squeezing a hand-sized ball with the own right hand.

The EEG was band pass filtered 0.5-100 Hz, amplified and digitized at 256 Hz. Recorded EEG signals were visually scored and trials with muscle or eye

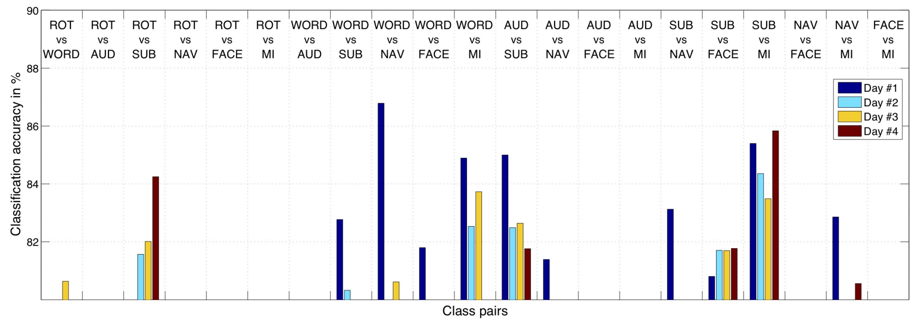

Fig. 1. Mean classification performance (10-times 10-fold cross-validation), averaged over nine subjects, based on classifying pairs of mental tasks on four different days.

movement activity within the imagery period were excluded from further analysis. Two distinct classes are required to transfer bits of information and consequently as first step we examined binary classification performances. The classical method of common spatial patterns (CSP) was used to design class specific spatial filters in the 8-30 Hz frequency band, and Fishers linear discriminant analysis (LDA) to classify the log-transformed normalized variance from 4 projections [3]. Each day was analyzed independently to rank the discriminability of the imagery pairs and to evaluate the variability between days. To get an overview of timing and dynamics of the induced EEG patterns, trials were subdivided into thirteen 1-s data segments with 0.5 s overlap. For each time segment and imagery pair, CSPs and LDA were computed and evaluated using a 10-times 10-fold cross-validation procedure.

Figure 1 summarizes initial results, including mean cross-validated accuracies, averaged over all participants, for each class pair and day. The mean accuracy for each day was computed by averaging subject-specific peak accuracies within the 7-s imagery period. Class pairs SUB vs. MI, AUD vs. SUB and SUB vs. FACE achieved binary classification accuracies >80% on every single day. The computed accuracies are comparable to accuracies computed between different motor imagery tasks reported in the literature. However, we found that performance over different days was less consistent when using distinct MI tasks. This is particularly evident during early training. Our off-line results suggest that the identified MTs induce EEG patterns that are statistically distinct and stable over time, and thus appropriate for BCIs [2]. We are now conducting feedback experiments to confirm these findings and evaluating multi-class discrimination performances.

3 Toward BCI-based gaming: World of Warcraft

Our comparison of different MTs for BCIs extends earlier work that also explored the best tasks for online BCI control (e.g. [4]). In [4], we also introduced a self-paced 3-class ERD/S BCI that allowed BCI user operating Virtual Google Earth

Fig. 2. BCI-based interaction with WoW. (a) Picture of the BCI player in front of the computer screen. (b) Game avatar (level 10 Tauren hunter) (c) The avatar in the starting area Mulgore. At the bottom of the screen, the in-game add-on shows the BCI-operated navigation arrows, and below that, the timeout bar for selecting critical events. In the screenshot, the avatar is walking forward. (d) The avatar is close to a non-playable character that offers him a quest. The timeout bar starts growing from left to right. If the user wants to accept the quest, he has to wait a predefined time. Otherwise, he can move away and cancel the action.

(Google, Inc., Mountain View, CA) [5]. We adapted this 3-class approach into a new system that lets the BCI user play the massive multiplayer online role playing game World of Warcraft (WoW, Blizzard Entertainment, Inc.).

WoW is a very popular video game that offers individual players or parties of players a virtual universe to explore and interact with. The main goal of the game is to team up with other players to fight monsters and complete various quests. Computer keyboard and mouse are used to control the avatar. In order to achieve hands-free control we developed a network controlled application, which simulates specific time based mouse and keyboard inputs. For visual feedback, the WoW native application-programming interface LUA was used to extend the standard game interface with an add-on that allows displaying BCI feedback. When the BCI detected left hand, right hand, or foot motor imagery (MI), the user received real-time feedback in form of arrows pointing to the left, right and forwards, respectively, which mapped to the navigation commands "rotate left", "rotate right" and "move forward" (Fig.2.a-c). The length of the arrow (4 increments) corresponds to the quality of detection, i.e., the BCI has to detect

MI for a configurable minimum time before the avatar starts moving. The avatar continues executing the selected action as long as the user performs MI. Users can start, stop and switch between MI tasks as required (self-paced mode).

In WoW, many actions involve specific points of interest (POIs). For example, the player must typically go to a certain POI to accept a quest from a non-player characters (NPC), engage specific enemies, gather collectible resources, or loot objects. With three directional control, a fourth, binary degree of freedom can be approximated by using a timeout, which we use for confirming critical actions, like attacking an opponent within range or accepting new quests. Whenever the player's avatar comes within the range of a POI, noncritical actions (e.g. getting a gatherable resource) are executed immediately. Critical actions can be accepted by staying within the range of the POI without moving for a predefined time period. During this period, the timeout bar located below the navigation arrows starts extending. Once fully extended, the action is performed (Fig.2.d).

This control strategy lets players explore and interact with the virtual game environment in many ways. Users can play hundreds of quests, engage a wide range of enemies in numerous environments, win loot from defeated enemies and awards from NPCs, earn a reputation in the game world, and improve their character in many ways. All of these actions could be implemented socially. Indeed, most players in WoW do not play alone, but conduct quests and other actions in cooperation with a party consisting of other human players. Hence, this new WoW BCI system had a much more complex and engaging task structure than our prior work, as well as more intelligent BCI processing software that incorporated context to enable more complicated, high-level control with a limited bandwidth.

4 Toward hierarchical adaptive BCIs for robotic control

There is also growing interest in intelligent and context-aware software for robotic devices, which could assist disabled individuals with various tasks in daily life. Many disabled people cannot use high bandwidth communication systems, and seemingly simple tasks like getting a glass of water may take far too long if the user must control low-level details of this task. Hence, intelligent robotic devices often allow users to accomplish high-level goals with a single command. BCIs have recently begun to adopt goal-directed protocols. In [6], we used a P300-based BCI to send high-level commands to a semi-autonomous humanoid robot that physical interacted with real objects in an augmented reality environment. To autonomously perform these tasks, the robot requires prior knowledge of the environment and substantial artificial intelligence. For example, to pour a glass of water from a bottle in a refrigerator, robots must navigate in the given environment, know the locations of cups and the bottle, manipulate objects, and heed safety protocols.

We are now developing a new generation of scalable, user-adaptive BCIs that combine the advantages of process-directed and goal-directed control [7, 8]. Users could utilize imagery to teach the BCI new commands, which are then made available for selection using evoked potentials (e.g., the P300). This leads to a hierarchical hybrid BCI wherein lower- level actions are first learned and later semi-autonomously executed using a higher-level command, thereby improving accuracy and freeing the user from tedious ongoing process-oriented control. Continuing the above example, the user first navigates the robot to the kitchen, and hence teaches the robot the task "Go to kitchen" from a specific starting point. To make this flexible and universally applicable, the robot has to generate an internal representation of the environment that allows finding the path to the kitchen from any location in the users environment. In a first experiment we used radial basis function models to learn navigation policies. Usually, however, not enough training data is available and so learned models may not reliably enough to predict the path. To overcome this problem, we introduced the use of uncertainty for guiding a BCIs behavior. We use Gaussian processes (GPs) for learning high-level commands and exploit the fact that they provide a measure of uncertainty in their output [8]. When the uncertainty in a given region of the task space is too high (e.g., due to lack of training in that area), the BCI switches to user control for further guidance rather than continuing to execute the unreliable and potentially dangerous high-level command. Such uncertainty-guided decision-making is critical for real-world BCI applications, such as BCI-control of a robotic wheelchair or helper robot, where user safety and the safety of those around the robot are of paramount importance.

The current prototype of the hierarchical BCI for robotic control is composed of the three components: A steady-state visual EP-based (SSVEP) BCI; a hierarchical menu and learning system; and the humanoid robot, i.e., a simulation of the robot that mimics the physics of the real world. The three components interact closely: The hierarchical adaptive menu system displays available commands as flashing stimuli for the user to choose using the SSVEP BCI. The user makes the desired selection by focusing on the desired command in the menu. The BCI detects the option the user is focusing on and sends its classification output to the hierarchical menu system, which in turn sends a command to the robot and switches to the next appropriate menu. The robot executes the command it receives, which can be either a lower-level command such as turn left/right or a higher-level learned command. Finally, the user closes the control loop by observing the simulated robots action and making the next desired selection based on the updated menu.

Up to three flickering stimuli (12 Hz, 15 Hz and 20 Hz) were presented on a TFT computer screen with a refresh rate of 60 Hz. The view of the robot in its environment was shown in a larger immersive setting above. One channel EEG was recorded (60 Hz notch; sampling rate 256 Hz) bipolarly from electrodes placed at Cz and Oz. SSVEPs were calculated by applying the Fast Fourier Transform to 1-s segments of EEG every 0.5 s. The frequency with the highest average power among the three target frequencies within the past 4-s was classified as the users choice. The structure of the hierarchical menu system is depicted in Fig.3.a. Users could choose to train either new or existing skills by navigating the robot through the virtual environment. During navigation, the robot moved

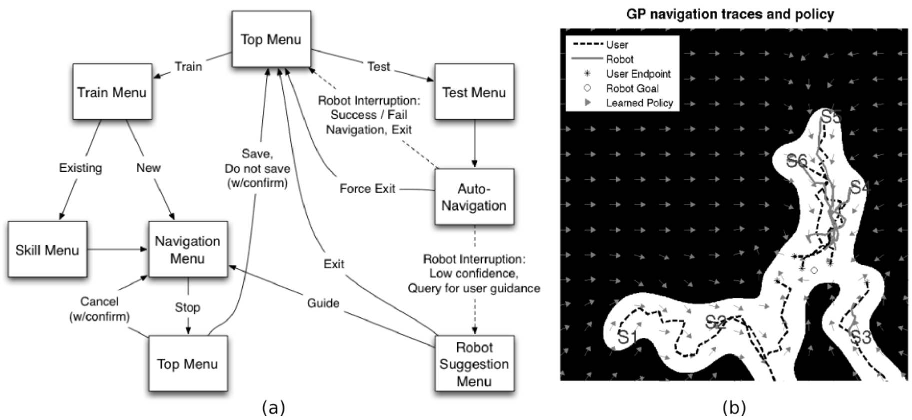

Fig. 3. Hierarchical learning framework. (a) Hierarchical menu structure (b) Navigation traces for GP model learning. Locations S2, S4 and S6 represent starting points of the robot during test mode (high-level command). S1, S3 and S5 represent starting points in training mode (low-level control). Dark areas mark regions of high uncertainty.

forward automatically. Available navigation commands included left, right and stop. After the user demonstrated and saved samples of the task, the BCI learns the task via GPs. Learned skills can be called and tested (Test Menu). The learned model guides the humanoid robot through the environment. Whenever the robots enters regions of uncertainty (high variance in the GP model), the BCI interrupts the operation and asks the user for assistance, i.e., to either guide the robot and collect more data, or to exit and return to the top-level menu. During execution of a high-level command, the robot queries the BCI for navigation direction based on the current position (GPS). If the GP model is used, one obtains both a predicted mean value as well as the variance of the prediction. This variance can be related to the confidence of the BCI in the learned model: high variance implies low confidence in the predicted navigational command and vice versa. For the current implementation, we used a simple threshold to decide when the robot should ask the user for guidance based on this confidence metric.

First results from four subjects using the hierarchical BCI indicate that the proposed system reduces the number of required selections and decreases the navigational time compared to low-level control [8]. Fig.3.b shows navigation traces of GP model learning and guidance.

5 Conclusion and Future direction

Non-invasive BCIs have evolved in recent years, and can now interact with a variety of applications and artificial devices. However, there is still room for improvement. We need better models of brain function and methods that predict brain activity, more practical recording equipment, and novel concepts on how to optimize brain-computer adaptation and extend progress to new applications and goals.

In this paper, we reviewed research in our labs that enhances the usefulness of BCIs and paves the way for new applications. We demonstrated that a proper combination of control strategies (i.e. mental tasks) increases classification results. We introduce a novel interface for the 3-class self-paced Graz-BCI that allows playing the massive multiplayer online role-playing game World of Warcraft. Players can not only navigate through the virtual world, but also interact with the virtual environment to talk to characters in the game world, acquire and complete game quests, and improve their character. We propose a new adaptive hierarchical architecture that allows a user to teach the BCI new skills on-the-fly. These learned skills are later invoked directly as high-level commands. We plan to continue this research direction, and explore new tasks and task combinations (such as in hybrid BCIs) and alternate applications and devices.

Acknowledgements

This work was supported in part by the NeuroCenter Styria, the ICT Collaborative Project BrainAble (247447), the National Science Foundation (0622252 & 0930908), the Packard Foundation, and the Office of Naval Research. We are grateful to Shane Dabiri and Tom Thompson from Blizzard Entertainment for supportive advice with WoW.

References

1. Mason, S.G., Bashashati, A., Fatourechi, M., Navarro, K.F, Birch, G.E.: A comprehensive survey of brain interface technology designs. Ann Biomed Eng **35(2)** (2007) 137–169
2. Friedrich E.V.C, Scherer, R., Neuper, C.: The effect of distinct mental strategies on classification performance of EEG-based brain-computer interfaces. Biol. Physiol. (to appear)
3. Ramoser, H., Müller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Biomed. Eng. **8(4)** (2000) 441–446
4. Scherer, R., Lee, F., Schlögl, A., Leeb, R., Bischof, H., Pfurtscheller, G.:Towards self-paced Brain-Computer Communication: Navigation through virtual worlds. IEEE Trans. Biomed. Eng, **55(2)** (2008) 675–682
5. Scherer, R., Schlögl, A., Lee, F., Bischof, H., Janša, J., Pfurtscheller, G.: The Self-Paced Graz Brain-Computer Interface: Methods and Applications. Computational Intelligence and Neuroscience. Article ID 79826 (2007) doi:10.1155/2007/79826
6. Bell, C.J., Shenoy, P., Chalodhorn, R., Rao, R.P.N.: Control of a humanoid robot by a noninvasive brain–computer interface in humans. J. Neural Eng. **5** (2008) 214–220
7. Chung, M., Scherer, R., Chaldohorn, C., Rao, R.P.N.: Towards Hierarchical BCIs: Combining Motor Imagery and Evoked Potentials for Robotic Control. 4th. Int. BCI Meeting 2010, Pacific Grove, CA, USA, May 31 - June 4, 2010
8. Chung, M., Cheung, W., Scherer, R., Rao, R.P.N.: A Hierarchical Architecture for Adaptive Brain-Computer Interfacing Proc. IJCAI 2011 (submitted)