Combining Eye Gaze Input with a Brain-Computer Interface for Touchless Human-Computer Interaction

Thorsten O. Zander[2], Matti Gaertner[2], Christian Kothe[2] and Roman Vilimek[1]

[1] Siemens AG, Corporate Technology, User Interface Design
Otto-Hahn-Ring 6, 81730 Munich, Germany

[2] Technische Universitaet Berlin, Chair of Human-Machine Systems, Team PhyPA
Franklinstr. 28-29, 10587 Berlin, Germany
thorsten.zander@mms.tu-berlin.de

Running Title: Combining Eye Gaze Input with a Brain-Computer Interface

Abstract. A Brain-Computer Interface (BCI) provides a new communication channel for severely disabled people that have completely lost the control over all muscular activity (locked-in syndrome). Furthermore, applications for people that are not in a complete locked-in state have been proposed (e.g. controlling a wheelchair). It is questionable whether a BCI is the best choice for these patients because using the remaining muscular activity might be more efficient. For example, gaze-based interfaces can be utilized as a communication channel for people that are still able to control their eye movements. Anyhow, these interfaces suffer from the lack of a natural degree of freedom for the selection command, e.g. a mouse click. Currently, the workaround for this problem in most eye-controlled applications is based on so-called dwell times. This interaction technique can easily lead to errors if the user does not pay close attention to where they are looking. In the work presented here, we developed a multimodal interface using eye movements to determine the object of interest and a Brain-Computer Interface (BCI) to simulate the activation command. Experimental results show, that the resulting hybrid BCI is a robust and intuitive device for touchless interaction in HCI. Although it is somewhat slower in comparison to standard dwell time eye gaze interfaces, it reliably leads to less errors. Furthermore, the BCI activation command allows to deal with different stimulus complexities[1].

Keywords: Brain-Computer Interaction, BCI, multimodal, eye tracking, eye controlled applications, HCI

1 Motivation

A classical electroencephalogram (EEG) based Brain-Computer Interface (BCI) is as system which "gives their users communication and control channels that do not depend on the brain's normal output channels of peripheral nerves and muscles" (Wolpaw et al. 2002). Its main purpose is to provide assistance in communication to severely paralized patients –

[1] *Parts of this study have been presented at the HCII 2009 conference in San Diego and have already been published in the proceedings of that conference.*

especially those who suffer from amyotrophic lateral sclerosis (ALS). Applications for BCI based systems also have been proposed for patients who are not in a completely locked-in state - e.g. controlling wheelchairs (Galán et al. 2008). It has to be questioned whether a BCI should be the method of choice, for patients that still have got some control over parts of their periphal nervous system. For these patients other physiological parameters might provide a more efficient information channel for interaction. Especially, when being able to control eye movements, the usage of an eyetracker seems to be a promising option because the achievable bitrate in communication is higher, as the information is not restricted to binary commands. Furthermore, the preparation time is definitely lower and eye gaze based systems are in the same price range as EEG systems. Nevertheless, some crucial problems remain to be solved in eye gaze based interaction in order to make it an efficient communication channel. One major problem that is tackled in the work presented here is the definition of a proper selection command.

Among others, Bolt introduced the use of eye movements as input modality for search-and-select tasks already in 1982 (Bolt, 1982). His idea of "eyes as output" was intended to facilitate HCI. Since then, numerous studies showed that the users' gaze can be used to efficiently solve search tasks (Engell-Nielsen, Glenstrup, Hansen, 2003; Nilsson, Gustafsson, Carleberg, 2007; Jacob, 1993; Murata, 2006; Hutchinson, 1993). However, whereas moving the mouse cursor with eye movements is quite intuitive, it is more difficult to find a proper mechanism for performing the click operation. Most solutions today are based on so-called dwell times, i.e. the user has to fixate an item for a pre-defined period of time in order to activate it. This technique has to face the inherent problem of finding the optimal dwell time. If it is too short, click events will be carried out unintentionally and thus lead to errors. If the dwell time is too long, fewer errors will be made but more experienced users will get annoyed and demotivated. Especially, in scenarios where the complexity of the provided stimuli is varying over time, there is no possibility of defining an optimal dwell time.

Adding a BCI to eye gaze based interaction could solve the described problem by providing an additional binary communication channel. On the one hand a BCI command is under voluntary control and on the other it is independent from stimulus complexity. Even though the dipolar properties of human eyes have a strong impact on the EEG signal (Lins et al. 1993), modern machine learning based algorithms are able to extract features independent from the users' eye movements. By combining BCI with eye gaze control we implemented a hybrid BCI (Pfurtscheller et al. 2006) which allows for an intuitive and efficient device for touchless interaction in HCI. Such a hybrid BCI can be a useful tool for both medical application and healthy users as well.

2.1 Gaze controlled User Interfaces

Eye gaze interaction can be a convenient and -- with certain restrictions -- a natural addition to the interaction with technical systems. The eye gaze of humans is basically an indicator for a person's attention over time (Kahneman, 1973). For human-computer interaction this means that the mouse cursor and visual focus usually correspond to each other which implies an intuitive substitution of the conventional mouse control by eye movements.

However, this rule does not always apply. The design of gaze-based systems has to consider unintentional fixations and sporadic dwellings on objects that typically occur during visual search or when people are engaged in demanding mental activity (cf. Yarbus, 1967). This fact is known as the "Midas Touch" problem: Although it may be helpful to simply look at an object and have the corresponding actions occur without further activity, it soon becomes annoying as it gets almost impossible to let the gaze rest anywhere without issuing a command (Jacob, Legett, Myers & Pausch, 1993). The problem directly points to the challenge of defining the mouse click operation in gaze-controlled environments.

In the past research dwell-time based solutions proved to be the best technique that can establish an even faster interaction process than using a mouse (Sibert & Jacob, 2000; Jacob, 1991). However, choosing a dwell time duration is always a trade-off between speed and accuracy. Furthermore, a well defined feedback informing the user about the current state of the activation progress is crucial, but can be difficult to design (Beinhauer, Vilimek & Richter, 2005). Even with adaptive algorithms, like e.g. shortening the dwell time period with growing user experience (Salvucci & Anderson, 2000), one major problem remains: The system can not know whether the user fixates a command button to trigger an action or for a different reason. For example the user could also have difficulties to read the description on the button, or reflect about the corresponding systems action, or he / she tries to understand the meaning of a complex icon. It is simply not possible to find a perfect relation between gaze duration and user intention. Additionally, for the user the activation of an item is inherently a different action than searching for it. From the perspective of distributed cognition (Hollan, Hutchins & Kirsh, 2000), it is reasonable to define somehow orthogonal input methods for these tasks. As dwelling is an inherent part of searching, this might lead to a high workload induced by dwell time based systems. Thus, it would be beneficial to replace this implicit way of issuing a command with a concurrent and directly controllable user action.

2.2 Brain-Computer Interfaces

A Brain-Computer Interface provides new information channels for human-machine interaction which are based on changes in the user's cognition. Though there are several methods for recording brain activity, the EEG is most promising for this approach as it provides a high temporal resolution and is comparably easy to apply. Methods from statistical machine learning are used to extract features of the EEG which allow for deducing context relevant information. But the capability of that information flow is usually restricted to binary commands and a low bit rate in communication.

At present, most BCI research focuses on solutions for the medical care sector where significant contributions were made in assisting people with massively restricted motor abilities (Birbaumer, Ghanayim, Hinterberger, Iversen, Kotchoubey & Kübler et al., 1999; Leeb, Friedman, Mueller-Putz, Scherer, Slater & Pfurtscheller, 2007; Birbaumer, 2006). Recently, some approaches evolved for augmenting this research, like the definition of BCIs in a broader context. From that perspective we propose a new, more open definition of BCI-based systems:

> *A BCI is a system to provide computer applications with access to real-time information about cognitive state, on the basis of measured brain activity.*

From this we derived a threefold categorization of types of BCIs (Zander & Jatzev, 2009):

> **Active BCI** *An active BCI is a BCI which derives its outputs from brain activity which is directly consciously controlled by the user, independently from external events, for controlling an application.*
>
> **Reactive BCI** *A reactive BCI is a BCI which derives its outputs from brain activity arising in reaction to external stimulation, which is indirectly modulated by the user for controlling an application.*
>
> **Passive BCI** *A passive BCI is a BCI which derives its outputs from arbitrary brain activity without the purpose of voluntary control, for enriching a human-computer interaction with implicit information.*

Another approach to use BCIs in different applications is the definition of hybrid BCIs for combining BCIs with other input modalities (Allison, Wolpaw & Wolpaw, 2007).

In this study we combine an active BCI approach with input deduced from an eye tracking device resulting in a new hybrid BCI.

2.3 Combining Eye Gaze Input and BCI

In the research reported here we evaluated the combination of a selection command through a BCI and spatial navigation through eye gaze input, which defines a hybrid BCI. A two dimensional cursor control is realized by tracking the user's eye gaze and a dedicated control thought detectable by a BCI is mapped on the activation of objects.

The eye movements themselves, however, are quite a challenge for an EEG based BCI. The eye, when moving, is a powerful dipole that disturbs the detection of the much lower potentials of the brainwaves. The recognition algorithm has to be capable to deal with that noise.

In this investigation, we focused on learning more about the potential of the here presented hybrid BCI as a multimodal BCI/Eye Gaze Interface. In typical Human-Computer Interaction the information presented by technical systems varies in complexity over time. One example is the complexity of the information encoded in icons displayed on a computer screen. Some are simple in structure and easy to read while others are ambiguous or complex. This is also influenced by subjective factors depending on experience of users or physiological factors, like quality of eye-sight. Hence, in these scenarios the choice of a proper dwell time gets even harder and might be even impossible. From this perspective, the system presented here would clearly outperform a dwell time based system, if the following requirements on the performance can be fulfilled.

First, anybody should be able to use the system after a short training session. Therefore, contrary to most experiments on BCIs, our participants had no working experience with a BCI before. Second, the replacement of dwell times by BCI need to prove to be a promising solution to the Midas Touch problem by not yielding higher error rates in the selection tasks. Task completion times of the BCI selection method should also be lower or at least comparable to dwell times. Finally, using the new interface must at least be as convenient as a solely gaze-based interface. Thus, the workload associated with the presented hybrid BCI should not be higher and its usage should be preferred in comparison to conventional eye tracking interfaces.

3 Experimental Evaluation

This study evaluates a search-and-select task in a 2D environment.We intend to mimic an everyday HCI. Therfore we use a typical personal computer setup and display an abstraction of item selection in HCI. As the quality of the gaze tracking is not a factor of investigation we decided to minimize its influence, by minimizing the chances of erroneous input.

This experiment compares a BCI based activation of targets in an eye controlled selection task against two conventional dwell time solutions with different activation latencies. The study aims to determine the degree to which a BCI can match or even outperform dwell time activation in respect to the factors effectiveness, efficiency and demands regarding cognitive resources in "clicking" the target stimulus.

Task difficulty in the selection task was varied by showing either simple visual stimuli with only a few random characters or by presenting more complex visual stimuli featuring a higher number of characters. Two different dwell times, a short one and a long one, were chosen for a better representation of the range of typical interaction situations with gaze-controlled applications. Assuming that generating the activation thought and processing feature extraction and classification still require a substantial minimal duration, it does not seem very likely that subjects will be able to complete tasks with the BCI faster. The question of interest here is whether they are significantly slower with a BCI than with dwell times. Additionaly we investigated the impact of stimulus complexity on the different activation types. Here, we expect a clear advantage of the BCI based solution.

The activation thought via BCI is a conscious, explicit command - in contrast to the implicit commands of dwell time solutions. Thus, the error rate in the BCI condition should be substantially lower, especially for difficult selection tasks. Also, this fact should lead to a lower workload and higher acceptance of the user, as the interaction should be more intuitive.

3.1 Methods

3.1.1 Participants

Ten participants (five female, five male) took part in the present study. They were monetarily compensated for their participation. Their age ranged from 19 to 36 years. Before engaging in the experiment subjects were screened for shortness of sleep, tiredness, and alcohol or drug consumption. All participants reported normal or corrected-to-normal vision.

3.1.2 Tasks

The participants had to perform a search-and-select task. They were presented with stimuli consisting of four characters in the "easy" condition and seven characters in the "difficult" condition. The reference stimulus was displayed in the center of the screen. Around this item twelve stimuli were shown in a circular arrangement, eleven distractors and one target stimulus, which was identical to the reference stimulus. The radial arrangement of search stimuli ensured a constant spatial distance to the reference stimulus. All presented strings consisted of consonants only to avoid similarities to known words. The distractors shared a constant amount of characters with the target. Examples of the search screens are shown in figure 1.

Subjects had to select the target stimulus by either fixating it for the given dwell time or by thinking the activation thought. It was not possible to use standard suggestions for dwell time durations from literature (e.g. Jacob, 1993), because the difficulty levels of the search task are not directly comparable to search tasks on a graphical user interface (GUI) in terms of absolute time needed for identification. The stimuli rather were chosen to be easily kept in working memory in the "easy" condition and to almost exceed its storing abilities in the "difficult" condition. To make sure that the dwell times match stimulus complexity, different versions were tested in pre-experiments. The selection criterion was that the short version is still well controllable and that the long activation latency is not perceived as slowing down the user. The short dwell time was 1.000 milliseconds, the long dwell time 1.300 milliseconds. To keep control over the duration of the experiment, a trial terminated automatically after 15 seconds if the user was not able to select an item during this time range. These trials were excluded from further analysis

Fig. 1. Examples for easy (left) and difficult (right) search tasks.

3.1.3 Apparatus

Brain activity was recorded by a 32 channel EEG system (Brain Products BrainAmp DC, actiCap). Electrodes were positioned according to the 10-20 system with 128 possible positions covering all relevant areas of the scalp. Signal processing was focussed on the sensomotric areas around C3 and C4. Grounding was established with electrode Fz. Eye movements were tracked with an infrared camera equipped remote eye tracker (SensoMotoric Instruments, iView X RED). Lighting conditions were held constant during the experiment.

3.1.4 The PhyPA BCI

An active BCI session as it is conducted by Team PhyPA, is usually divided into four stages (Zander, Kothe, Welke & Roetting, 2008). In the first stage, which is called *user training*, the user gets familiar with the system and the experimental task. In the second stage, labelled EEG data is recorded for calibrating a classifier on the task-relevant EEG features. This stage is called *machine training*. Next follows the *convergence phase*, where the user tries to control a simple feedback. Depending on the performance of and feedback from the user, parameters of the classifier can be readjusted. This stage should lead to a convergence of two learning systems, the user and the technical pattern recognition system. The last stage is the *application phase*, in which the user carries out the experimental task by using the previously calibrated BCI system.

3.1.5 BCI classes for binary control

In the BCI session of this experiment the task for the user was to select an object on the screen by an imagined hand movement. Users had to imagine wringing out a towel with both hands by turning them into opposite directions. We chose this type of imagined hand movements because it represents a dedicated activity, which seemed to be reasonable from the perspective of embodied cognition (Wilson, 2002). During the user training, users were assisted in how to imagine this movement. Many users found the imagination task easier after they were given the hint to try to feel sensations on their hands instead of visually imagining the movement.

When the users had learned to perform the imagination task, the machine training started. In order to train a robust classifier, the experimental design of the machine training has to be chosen carefully. Most importantly, the training phase has to be as similar as possible to the experimental condition. Nevertheless, there is an inherent difference between training phase and application phase, because the first contains no feedback from the system. Also the data from the training phase should not contain too much task-relevant noise, so-called artifacts. In the case of this experiment these constraints have been fulfilled by adopting the search-and-select task of the experimental condition for the training paradigm. In the training phase of this experiment, subjects were viewing the same circle of strings as in the application phase. In contrast to the application phase a box containing the word ‘search’ was jumping randomly from string to string while covering the letters behind. Subjects were instructed to follow the box with their eyes until the word ‘select’ appeared. When this occurred, they were instructed to perform the imagined hand movement for three seconds. Two classes were defined for classification. One class was chosen as the period of time when subjects were following the ‘search-box’ with their eyes. The other class was defined as the period of time when subjects were instructed to perform the imagined hand movement. The training session lasted approximately 15 minutes and 40 trials of training data for each condition were recorded.

When the training session was finished, a classifier was trained on selected features of the previously collected data. These features have been extracted from the spectral domain, as it is a well known phenomenon that the imagination of hand movements causes reduced amplitude of the sensori motor rhythm (SMR), 7 – 13 Hz, over sensorimotor brain regions (Pfurtscheller, Neuper, Flotzinger & Pregenzer, 1997).

For feature extraction we used the SpecCSP algorithm (Tomioka, Dornhege, Nolte, Blankertz, Aihara & Mueller, 2006), which automatically generates weights for each electrode and frequency by maximizing the variance between classes and by finding optimal linear combinations of electrodes and frequencies. Using this method, the dimensionality of the data is reduced drastically while still containing the relevant information by generating meta-channels through a linear combination of the weighted channels and frequencies. This method helps to avoid overfitting resulting from a bad trial-to-dimensionality ratio. The mean electrode and frequency weighting over all subjects is shown in figure 2.

The selected features are based on two different types of signals. The first class of features is based on the event related desynchronization over the motor cortex. The second class is defined as noise in the search condition, which should not be dependent on any specific process. The latter preliminary is hard to control, as there might be factors in the search condition inducing specific class relevant features which are detectable by SpecCSP. This has to be controlled by validating the features selected by the classifier, which will be done in section 3.2.

Fig. 2: mean electrode and frequency weighting over all subjects calculated by the SpecCSP algorithm. The units of all weights shown are arbitrary.

In the second step, a linear classifier (Linear Discriminant Analysis) was generated from the extracted features. The classification accuracy was validated via a 3x10 cross-validation. When the error rate was below a threshold of 40 %, the classifier was saved for later usage in the experiment.

3.1.6 Design and Procedure

After making sure that all EEG electrodes were in place and working, additional electromyogram (EMG) electrodes were attached to the participants' arms to monitor for muscular activity which might correlate to activation command. Before and during the technical preparations subjects received a general overview on the procedure of the experiment and their tasks. A complete and summarized presentation of the test setting was given afterwards. Then the user training and machine training stages followed (cf. 3.1.5). To finalize this preparation phase, subjects practiced using the BCI command and the experimentor readjusted the bias of the classifiers hyperplane in the convergence phase. When the training was successful, a short calibration of the eye tracker followed and the experiment started. Two different levels of search difficulty (easy, difficult) and three levels of activation technique (dwell time short, dwell time long, BCI) were varied in a 2×3 factorial design, with repeated measurements on both factors. Participants went through the levels of the factor activation technique in separate blocks. The order of these blocks was counterbalanced across subjects. Subjects completed 30 trials per condition. The experiment itself took about one hour, the whole test procedure about 2.5 hours. Effectiveness was measured in terms of errors in task completion. Efficiency was defined as the time needed to complete a search task. Mental workload was assessed with the unweighted version of the

NASA Task Load Index – the so-called Raw Task Load Index, RTLX (Byers, Bittner & Hill, 1989).

The NASA TLX was completed by the participants after each condition. At the end of the experiment the participants had the opportunity to discuss their experiences with the experimenter and were asked to rate the activation techniques according to their preferences.

3.2 Results and Discussion

The crossvalidated accuracy of the classification during the training phase lies at 89.0 % averaged over all subjects. The standard deviation is 10.3 %. One subject could be identified as an outlier as it stated clear doubts on the functionality of BCI based systems. Hence, its motivation for this experiment is questionable. Consistently, the crossvalidation showed an accuracy of 60.9 %. The averaged crossvalidation accuracy without the outlier lies at 92.1 % (std 3.1 %). The recorded EMG shows no significant correlation to the classification relavant classes.

Time needed for task completion and accuracy (data on errors) were averaged across all subjects for each selection method and level of search difficulty. Trials with errors were not included in the analysis of response time. First, an analysis of variance was conducted on the results. The alpha level for significance was chosen to be 5 %. In a second step, the data of the easy and difficult condition were pooled for each selection method. This allows taking a closer look in pairwise comparisons between BCI vs. long dwell time and BCI vs. short dwell time. To avoid the problem of multiple comparisons, the alpha level was Bonferroni corrected for these tests.

Fig. 3. Percentage of correct selections: Brain-Computer Interface (bci), long dwell times (dwl) and short dwell times (dws).

The accuracy data are summarized in Figure 3. The "easy" condition yielded 88.0% correct selections when using the BCI. In 93.8% of all tasks correct answers were produced in the "dwell time long" (dwl) condition, in the "dwell time short" (dws) condition 83.8%. Fewer correct selections were made in the "difficult" condition. Remarkably, the BCI leads to the best results with 78.7% correct selections, although the difference to the long dwell time, 75.6%, is only marginal. The short dwell time condition, however, leads to a strong negative effect on performance as the percentage of correct answers dropped to 51.1%. This change in the result pattern in the difficult condition is reflected in a significant search condition activation technique interaction ($F(2,18) = 13.30$, $p < .001$). An analysis of the main effects confirms general differences between the activation techniques ($F(2,18) = 12.47$, $p < .001$) and that the difficult search condition leads to more errors ($F(1,9) = 38.37$, $p < .001$).

The pooled BCI accuracy average is 83.3% correct selections, the corresponding values for dwell time long and dwell time short are 84.7% and 67.4%. Pairwise t-tests reveal that the better performance of the BCI compared to "dwell time short" is significant ($t(9) = 3.66$, $p = .005$). The small differences between BCI and "dwell time long" is not reliable ($t(9) = 0.33$, $p = .75$). As expected, the BCI allows users to activate (click) GUI items more precisely than a dwell time solution with short latencies. Long dwell times are suited for precise object activation but do not prove to be substantially better than BCI based selection.

Task completion was fastest in both search conditions with short dwell times (easy: 3.98 s; difficult: 5.38 s). Next was dwell time long (4.79 s; 7.37 s), leaving BCI the slowest method of activation (5.90 s; 8.84 s). This general difference between the input methods is statistically confirmed ($F(2,18) = 56.25$, $p < .001$). The results are depicted in Figure 4.

Fig. 4. Task completion times: Brain-Computer Interface (BCI), long dwell times (dwl) and short dwell times (dws).

Looking at these results the data also shows that the difficult search task leads to longer search times, which is only of minor interest ($F(1,9) = 102.38$, $p < .001$). The significant interaction of the factors ‘search condition’ and ‘activation technique’ reflect larger differences in the ‘difficult’ compared to the ‘easy’ condition ($F(2,18) = 7.46$, $p < .01$). The pairwise comparisons support the view that BCI selection was slowest (bci: 7.37 s; dwell time long: 6.08 s; dwell time short: 4.68 s). These differences are significant (bci - dwl: $t(9) = 4.31$, $p = .002$; bci - dws: $t(9) = 13.57$, $p < .001$).

The overall TLX results show no differences in workload between the activation techniques. Nevertheless, in one subscale of the NASA RTLX – the question that was concerned about the amount of frustration of the users – the BCI method was rated significantly lower ($p < .05$) than both dwell times. This finding is consistent with the preferences ratings at the end of the experiment. Here, nine out of ten participants preferred using the combined BCI/Eye Gaze Interface over the standard gaze-based interface. In the concluding questenaire many subjects stated, that they followed an avoidance strategy in the dwell time approach. They moved their eyes shortly on an item and then quickly into a ‘safe area’ at the border of the screen to avoid any miss selection. Then they recalled the previously seen image from their memory and decided, wheter this item is a target or not.

As stated in 3.1.5 the paradigm of the machine training stage has to be quite similar to the task in the application stage, but may be contaminated by artifacts. Hence, it is necessary to evaluate on which type of features the classification is based on. In this case, the question has to be answered, whether the BCI classification was based on the intended features of the EEG signal or whether it was influenced by artifacts. As parts of the features have been chosen automatically by the SpecCSP algorithm, it would be possible, that these do not reflect the intended aspects of the EEG signal, namely imagined hand movements and absence of imagined hand movements. The definition of the machine training stage also induces another class relevant factor which is recorded by the EEG, namely not moving the eyes and moving the eyes. If features induced by eyemovements are more relevant for classification than those resulting from motor imagery, one class will be defined by the existence of eye movements while the other is defined by the absence of eye movements. Hence, the resulting selection method would behave similar to the dwell time based approach. This has to be investigated. There is strong evidence that supports the assumption that classification fundamentally was based on the imagination task. On the one hand the SpecCSP frequency- and electrode weightings (s. Fig. 2) show that for the selection-class electrodes over the motor cortex have the highest weights. Also the frequency range with the highest weights lies in the alpha band, the frequency band of the sensorimotor rhythm. A classification on features reflecting the absence of eye movements would induce higher weights on frontal electrodes and on a broader frequency band. For the search-class frontal electrodes are not weighted high either. These facts support the assumption that classification was not based on eye movements. From a practical point of view, classification on the absence of eye movements would face the same problems as the dwell time approach. An item is activated if there are no eye movements. Hence, we asked the subjects to fixate a string of letters without imagining the activation command during the convergence phase (stage 3). In nine of ten subjects the strings were not selected until subjects were asked to perform the imagination task.

4 Conclusions and Outlook

Our results show that it is possible to perform more accurate selections with the BCI activation command than with the solution that is based on short dwell times. Also the strong user preference for using a BCI instead of dwell times for the activation of selected objects and the lower rating in the frustration scale of the BCI method are quite remarkable results. However, using BCI is still somewhat slower. Nonetheless, although statistically significant, the magnitude of the difference between BCI and the dwell time solutions is notably small. An explanation for that might lie in the change from an implicit to an explicit activation command. When using a BCI, subjects could take the time they needed for solving the task and did not have to choose an avoidance strategy as in the dwell time approach. The above mentioned findings of user prefence and low frustration ratings also point to the assumption that users where taking the time they needed for task completion during the BCI selection without being in a rush because of a restricted dwell time. Taken together, these findings show that a BCI has successfully proven to be a real competitor for dwell time activations already in the current state of technological development. The concluding statement is that the presented system is a good match to existing dwell time based solutions regarding the factors investigated in the experimental setup. With the benefit of voluntary controlling the activation command, it is especially applicable in environments with fluctuating stimulus complexities. Fixed dwell times are not applicable in such environments because the time that is needed for object selection varies from trial to trial. This is underlined by the results from the factor combination of short dwell times and difficult strings.

This fact, as well as the other results of the presented study, clearly indicate that the presented approach for activation is a forthcoming technology for multimodal interfaces that may help to optimize touchless interaction for healthy users as well as in medical applications.

Acknowledgements: We gratefully thank Sandra Troesterer, Eva Wiese, Johann Habakuk Israel and Matthias Roetting for their beneficial comments and support.

References

Allison, B. Z., Wolpaw, E.W., Wolpaw J.R., (2007). Braincomputer interface systems: progress and prospects. *Expert Review of Medical Devices*, 4(4):463-474.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. (2001) Recent Advances in Augmented Reality *Computer Graphics*, 21(6), 34-47.

Beinhauer, W., Vilimek, R., Richter, A. (2005). Eye-Controlled Applications - Opportunities and Limits. *Proceedings of Human-Computer Interaction International 2005*. Mahwah, NJ: Lawrence Erlbaum Associates.

Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kbler, A., Perelmouter, J., Taub, E., Flor, H. (1999). A Spelling Device for the Paralyzed. *Nature*, 398, 297-298.

Birbaumer, N. (2006). Breaking the Silence: Brain-Computer Interfaces (BCI) for Communication and Motor Control. *Psychophysiology*, 43, 517-532.

Bolt, R.A. (1982). Eyes at the Interface. *Proceedings of the 1982 Conference on Human Factors in Computing Systems* (pp. 360-362). New York: ACM Press.

Bowman, D. A. & Hodges, L. F. (1997) An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: *Tagungsband Symposium on Interactive 3D graphics.* Providence, Rhode Island, USA, ACM Press, 35-38.

Bowman, D.A. (1999). *Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application*. Ph.D. thesis, Virginia Polytechnic & State University.

Bowman, D. A., Kruijff, E., LaViola Jr., J. J. & Poupyrev, I. (2005) *3D User Interfaces: Theory and Practice.* Reading, Massachusetts: Addison-Wesley.

Byers, J.C., Bittner, A.C., Hill, S.G. (1989). Traditional and Raw Task Load Index (TLX). Correlations: Are Paired Comparisons Necessary? In Mital, A. (Ed.). *Advances in Industrial Ergonomics and Safety* (pp. 481-485). London, UK: Taylor & Francis.

Drascic, D. & Milgram, P. (1996) Perceptual issues in augmented reality. *SPIE Stereoscopic Displays and Applications VII, Virtual Reality Systems III* 2653, 123–134.

Engell-Nielsen, T., Glenstrup, A.J., Hansen, J.P. (2003). Eye Gaze Interaction: A New Media - Not Just a Fast Mouse. In Itoh, K., Komatsubara, A., Kuwano, S. (Eds.), *Handbook of Human Factors / Ergonomics* (pp. 445-455). Tokyo, Japan: Asakura Publishing.

Galán, F., Nuttin, M., Lew, E., Ferrez, P. W., Vanacker, G., Philips, J., Millán, J. R. (2008).A brain-actuated wheelchair: asynchronous and non-invasive Brain-computer interfaces for continuous control of robots. Clinical Neurophysiology, 119(9), 2159-2169.

Hinckley, K., Pausch, R., Gobble, J. & Kassell, N. (1994). Passive real-world interface props for neurosurgical visualization. In: *Tagungsband CHI'94.* ACM Press, 452–458.

Hollan, J., Hutchins, E., Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interact*ion, 7(2):174-196.

Hutchinson, T.F. (1993). Eye Gaze Computer Interfaces: Computers That Sense Eye Positions on the Display. *Computer*, 26, 65-67.

Insko, B. E. (2001) *Passive Haptics Significantly Enhances Virtual Environments.*Chapel Hill: University of North Carolina.

Iwata, H. (2003) Haptic Interfaces. In: Jacko, J. A. & Sears, A. (Hrsg.) *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. 206-219.

Jacob, R.J.K. (1991). The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look at Is What You Get. *ACM Transactions on Information Systems*, 9, 152-169.

Jacob, R.J.K. (1993). What You Look at Is What You Get. *IEEE Computer*, 26, 65-66.

Jacob, R.J.K., Legett, J.J., Myers, B.A., Pausch, R. (1993). Interaction Styles and Input/Output Devices. *Behaviour & Information Technology*, 12, 69-79.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Krause, F.-L., Israel, J. H., Neumann, J. & Feldmann-Wüstefeld, T. (2007) Usability of Hybrid, Physical and Virtual Objects for Basic Manipulation Tasks in Virtual Environments. In: *Tagungsband IEEE Symposium on 3D User Interfaces (3DUI).* Charlotte, NC, USA, ACM Press, 87-94.

Leeb, R., Friedman, Mueller-Putz, G.R., Scherer, R., Slater, M., Pfurtscheller, G. (2007). Self-Paced (Asynchronous) BCI Control of a Wheelchair in Virtual Environments: A Case Study with a Tetraplegic. *Computational Intelligence and Neuroscience*, 7, 1-8.

Lins, O. G., Picton, T. W., Berg, P., Scherg, M. (1993). Ocular artifacts in EEG and event-related potentials. Brain Topography, 6(1), 51-63.

Massie, T. H. (1996) *Initial haptic explorations with the PHANToM: Virtual touch through point interaction.* Master Thesis, Bosoton: Massachusetts Institute of Technology.

Murata, A. (2006). Eye Gaze Input Versus Mouse: Cursor Control as a Function of Age. *International Journal of Human-Computer Interaction*, 21, 1-14.

Nilsson, S., Gustafsson, T., Carleberg, P. (2007). Hands Free Interaction with Virtual Information in a Real Environment (pp. 53-57). *Proceedings of COGAIN 2007*, Leicester, UK.

Pfurtscheller, G., Neuper, C., Flotzinger, D., & Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, 103(6), 642-651.

Pfurtscheller, G., Scherer, R., Müller-Putz, G. (2006). Heart Rate-Controlled EEG-Based BCI: The Graz Hybrid BCI. In Proceedings of the 3rd International Brain-Computer Workshop and Training Course 2006. Graz University of Technology Publishing House, Graz, Austria.

Pierce, J., Stearns, B. & Pausch, R. (1999) Voodoo dolls: seamless interaction at multiple scales in virtual environments. In: *Tagungsband Symposium on interactive 3D graphics.* Atlanta, GA, USA, ACM Press, 141 - 145.

Poupyrev, I., Billinghurst, M., Weghorst, S., Ichikawa, T. (1996). Go-Go interaction technique: Non-linear mapping for direct manipulation in VR (pp. 79-80). *Proceedings of UIST*. New York: ACM Press.

Salvucci, D.D., Anderson, J.R. (2000). Intelligent Gaze-Added Interfaces. *CHI '00. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 273-280). New York: ACM Press.

Srinivasan, M. A., Basdogan, C. (1997). Haptics in virtual environments: Taxonomy research status, and challenges. *Computers & Graphics*, 21(4), 393-404.

Sibert, L.E., Jacob, R.J.K. (2000). Evaluation of Eye Gaze Interaction. *CHI '00. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281-288). New York: ACM Press.

Tomioka, R., Dornhege, G., Nolte, G., Blankertz, B., Aihara K., Mueller, K.R. (2006). Spectrally weighted common spatial pattern algorithm for single trial EEG classification.Technical Report, 21. Dept. of Mathematical Engineering. Tokio, Japan: University of Tokyo.

Ware, C., Rose, J. (1999). Rotating Virtual Objects with Real Handles. *ACM Trans. CHI,* Vol. 6, No. 2, 162-180.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4):625-636.

Yarbus, A.L. (1967). Eye Movements During Perception of Complex Objects. In Riggs, L.A. (Ed.). *Eye Movements and Vision* (pp. 171—196). New York: Plenum Press.

Zander, T.O., Kothe, C., Welke, S., Roetting, M. (2008). Enhancing Human-Machine Systems with Secondary Input from Passive Brain-Computer Interfaces. *Proceedings of the 4th International BCI Workshop & Training Course*. Graz, Austria: Graz University of Technology Publishing House.

Zander, T.O., Jatzev, S. (2009). Detecting affective covert user states with passive Brain-Computer Interfaces. In *Proceedings of the ACII 2009*. Los Alamitos, CA: IEEE Computer Society Press.

Figures

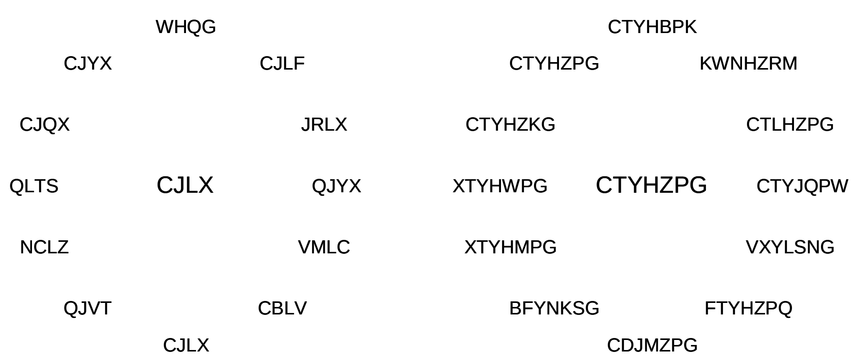

Fig 1: Examples for easy (left) and difficult (right) search tasks.

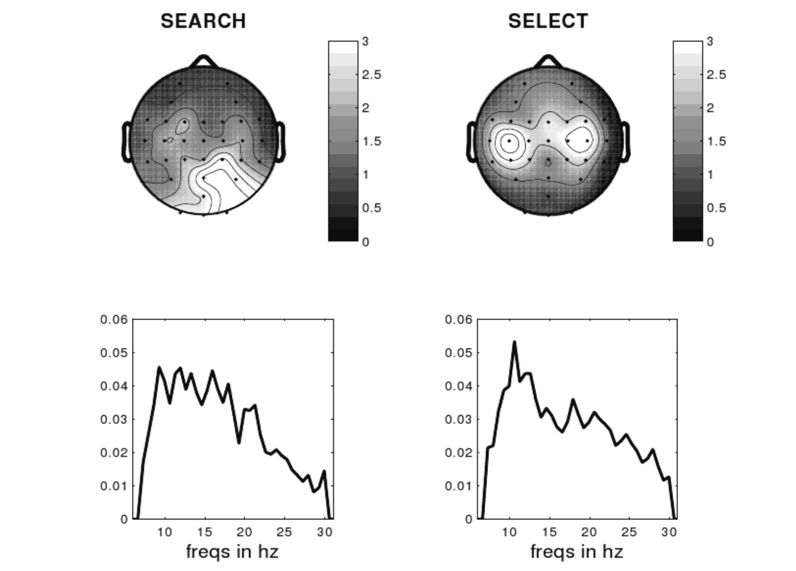

Fig 2: mean electrode and frequency weighting over all subjects calculated by the SpecCSP algorithm. The units of all weights shown are arbitrary.

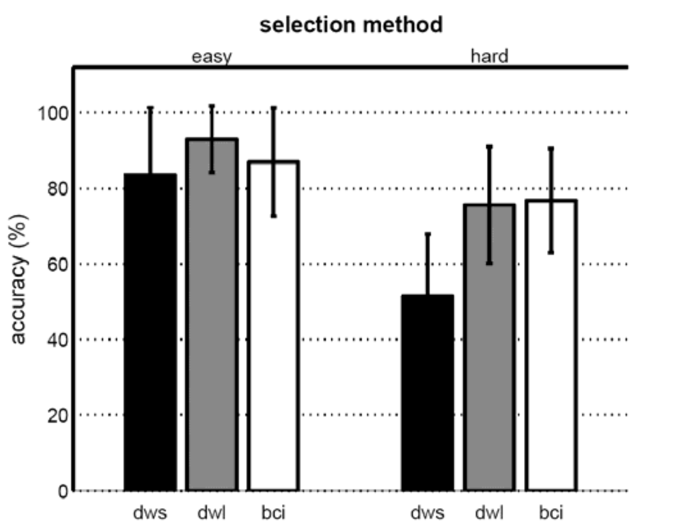

Fig 3: Percentage of correct selections: Brain-Computer Interface (bci), long dwell times (dwl) and short dwell times (dws).

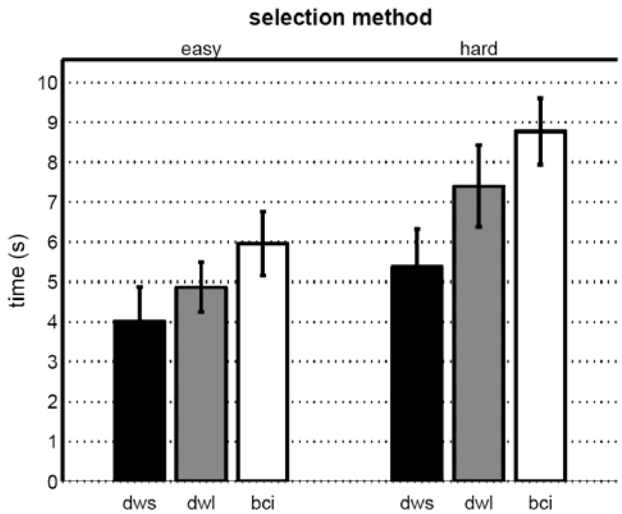

Fig 4: Task completion times: Brain-Computer Interface (bci), long dwell times (dwl) and short dwell times (dws).