

Controlling an avatar by thought using real-time fMRI

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 J. Neural Eng. 11 035006

(<http://iopscience.iop.org/1741-2552/11/3/035006>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 132.68.4.120

This content was downloaded on 26/05/2014 at 07:40

Please note that [terms and conditions apply](#).

Controlling an avatar by thought using real-time fMRI

Ori Cohen^{1,2}, Moshe Koppel², Rafael Malach³ and Doron Friedman¹

¹ The Interdisciplinary Center Herzliya, 46150, Israel

² Bar-Ilan University, Ramat-Gan 52900, Israel

³ Weizmann Institute of Science, Department of Neurobiology, Rehovot 76100, Israel

E-mail: orioric@gmail.com, moishk@gmail.com, rafi.malach@gmail.com and doronf@idc.ac.il

Received 5 February 2014, revised 25 March 2014

Accepted for publication 26 March 2014

Published 19 May 2014

Abstract

Objective. We have developed a brain–computer interface (BCI) system based on real-time functional magnetic resonance imaging (fMRI) with virtual reality feedback. The advantage of fMRI is the relatively high spatial resolution and the coverage of the whole brain; thus we expect that it may be used to explore novel BCI strategies, based on new types of mental activities. However, fMRI suffers from a low temporal resolution and an inherent delay, since it is based on a hemodynamic response rather than electrical signals. Thus, our objective in this paper was to explore whether subjects could perform a BCI task in a virtual environment using our system, and how their performance was affected by the delay. **Approach.** The subjects controlled an avatar by left-hand, right-hand and leg motion or imagery. The BCI classification is based on locating the regions of interest (ROIs) related with each of the motor classes, and selecting the ROI with maximum average values online. The subjects performed a cue-based task and a free-choice task, and the analysis includes evaluation of the performance as well as subjective reports. **Main results.** Six subjects performed the task with high accuracy when allowed to move their fingers and toes, and three subjects achieved high accuracy using imagery alone. In the cue-based task the accuracy was highest 8–12 s after the trigger, whereas in the free-choice task the subjects performed best when the feedback was provided 6 s after the trigger. **Significance.** We show that subjects are able to perform a navigation task in a virtual environment using an fMRI-based BCI, despite the hemodynamic delay. The same approach can be extended to other mental tasks and other brain areas.

Keywords: fMRI, BCI, avatar

(Some figures may appear in colour only in the online journal)

1. Introduction

Brain–computer interfaces (BCIs) allow people to interact with external devices using their thought alone. Most BCI research is aimed at helping patients with severe nervous-system damage including spinal cord injuries and stroke, and the goal is to provide such patients with some levels of communication, control of external devices, and mobility.

BCI can be invasive, using electrocorticography (e.g. [1]) and intracortical neural implants (e.g. [2–4]). Non-invasive BCIs are often based on electroencephalography (EEG). Non-invasive BCI-controlled navigation have been demonstrated using mainly three major EEG-based BCI paradigms: the steady state visually-evoked potential (SSVEP), the P300

wave, and motor imagery (MI). SSVEP-based BCIs were used to control virtual environments [5, 6], mobile robots [7, 8], manipulators [9], a wheelchair [10], and recently, humanoid robots [11, 12]. P300 is typically used for spelling application, but was also used in several virtual environment studies [13, 14].

Because they rely on visual evoked responses, both SSVEP and P300 can be compared to eye-tracking systems: they provide similar functionality and suffer from similar limitations. MI has also been used for EEG-based BCI; imagination of movement evokes brain networks that are similar to the networks evoked by real execution of the corresponding physical movement [15]. A series of studies were carried out with MI based navigation of highly-immersive

virtual reality [16–19] including experiments with a tetraplegic patient [20].

fMRI-based BCI is promising for several reasons. Although it is expensive, fMRI is a risk free procedure, unlike invasive methods. Due to the superior spatial resolution fMRI may be used to classify a much wider set of mental patterns than EEG, and thus fMRI-based BCI may allow exploring new BCI paradigms. While the temporal resolution of fMRI is significantly lower than EEG, in practice the difference might not be so dramatic. EEG has a high temporal resolution, but due to the low spatial resolution the classification requires a time window, which in practice results in a delay of several seconds. For example, SSVEP-based BCIs, which have the highest bit rate among EEG-based methods, provide the best results with a 3.4 s delay [21]. If successful, attempts can be made to localize underlying brain patterns and transfer the paradigms back to other, more accessible signals, such as EEG, near-infrared spectroscopy (NIRS) [22, 23], and hybrid EEG-NIRS-based BCIs [24]. fMRI-based BCI can also be used for training patients in BCI (e.g. before surgery), for rehabilitation sessions, or for next generation neurofeedback—in all these cases very specific brain areas may be targeted. In addition, smaller, less expensive and portable fMRI devices may become available [25].

Real-time fMRI has been suggested for various applications [26]. Typically, real-time fMRI is used as a form of neurofeedback, i.e., the raw signal values from a specific region of interest (ROI) in the brain are visualized on the screen, either as a bar or as a time-course plot. The subject uses a mental strategy to increase or decrease the activity in the target brain region. Berman *et al* [27] found that subjects quickly gained the ability to self-modulate their primary motor cortex by using finger tapping, in contrast to using tapping imagery. Similar strategies involved using continuous feedback [28–30] or intermittent feedback [31]. Such neurofeedback sessions are different from BCI in several ways. First, the goal is different: in neurofeedback the goal is to train the subjects to modulate their brain activity, whereas in BCI the goal is to allow subjects to control an external device by thought. In BCI the subject has to toggle among two or more different mental patterns, in a relatively fast pace, in order to control the environment. In neurofeedback most of the effort is done by the subject, and the system is used only for visualizing the brain signals, whereas BCI systems include algorithms for processing the brain signals and mapping them into specific actions taken by the external device, in real time. In the last few years, there have been a few attempts at fMRI-based BCI systems, beyond neurofeedback. Subjects were able to move a computer cursor [32, 33], balance an inverted pendulum [34], control a robotic arm [35–37], and a wheeled robot [38]. Our main contributions in this paper are the comparison of cue-based and free-choice tasks, the systematic analysis of several time delays, and the analysis of subjective responses to controlling an avatar by thought.

Using the same approach and system as described here, we have performed a pilot study whereby subjects were able to control a humanoid robot [39]. In this study we go beyond the robotic pilot study, which was intended to verify technical

feasibility. We provide a subject with the sense of being embodied in a 3D avatar while performing a task in a virtual environment. Since one of the main hurdles in fMRI is the temporal delay, we introduce the concept of time to feedback (TTF) and study it systematically. Since the BOLD signal lags after the mental activity by a few seconds, we assume that a classification result obtained at time t relates to the subject's intentions at time $t - d$, where d is the magnitude of the delay. Thus, even though we can provide feedback to the subject following every scan (in our case every 2 s), we also explored slower control schemes, such that each action is based on durations of 2, 4, 6, 8 or 10 s; we refer to these durations as TTFs.

BCI systems rely not only on the engineered components, but also on the mental effort of the participants. There has been surprisingly little research on the subjective and psychological aspects of using BCI as an interaction device. Ohara *et al* [40] observed gamers' body language when playing a MindFlex⁴ game. They point out that people want to compensate for the lack of bodily manifestations to intentionality, and thus performed a qualitative analysis of the body gestures during interaction. Friedman *et al* [41] describe the subjective experience of controlling a highly-immersive virtual reality using an EEG-based BCI. In follow up work they also describe controlling an avatar by thought, with some insights into human-computer interface issues arising in such experiences [16].

2. The system

Imaging was performed on a 3T Trio Magnetom Siemens scanner, and all images were acquired using a 12 channel head matrix coil. Three-dimensional T1-weighted anatomical scans were acquired with high resolution 1 mm slice thickness, 3D MP-RAGE sequence, repetition time (TR) 2300 ms, TE 2.98 ms, 1 mm³ voxels). For blood-oxygenation-level-dependent (BOLD) scanning, T2*-weighted images using echo planar imaging sequence (EPI) were acquired using the following parameters: TR 2000 ms, TE 30 ms, flip angle 80, 35 oblique slices without gap, 20 toward coronal plane from anterior commissure–posterior commissure, 3 × 3 × 4 mm voxel size, covering the whole cerebrum. Visual feedback is provided by a mirror, placed 11 cm from the eyes of the subject and 97.5 cm from a screen, which results in a total distance of 108.5 cm from the screen to the eyes of the subject.

There is a tradeoff between the scanning rate and the number of slices scanned, and we have opted for a scan time (TR) of 2000 ms. A lower TR of 1000 ms is possible if we scan a smaller portion of the brain (i.e., less slices), but this would prevent us from performing a full brain analysis in the future.

Our system is based on Turbo Brain Voyager (TBV, Brain Innovation, Netherlands)⁵. Dicom files⁶ from the scanner are processed by TBV, which computes the average raw data values for each ROI selected by the operator online. Using

⁴ <http://mindflexgames.com/>

⁵ www.brainvoyager.com/

⁶ <http://medical.nema.org/>

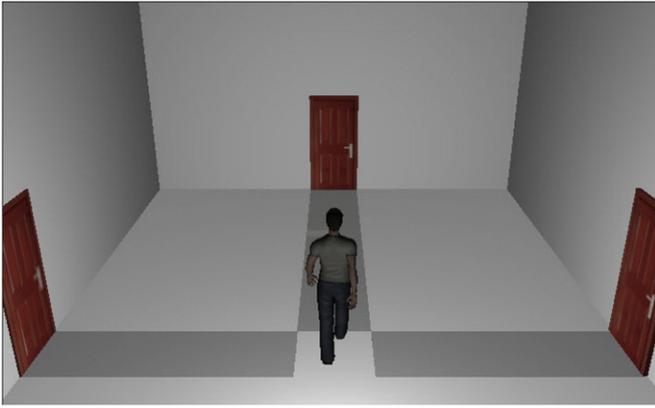


Figure 1. The subject sees an avatar in the center of a three-door room and needs to rotate or move the avatar toward the corresponding door, using either motor imagery or finger movement.

TBV all subject scans are auto aligned by a real-time on-line algorithm that uses a statistical atlas to automatically position the scanned slices [42] and by a real-time three-dimensional motion correction algorithm, the prospective acquisition correction algorithm, which adjusts slice position and orientation in order to reduce motion artifacts [43]. This pre-processing is applied at the initial scan, between every two scans for subsequent subject movement, and when subjects return for additional scans, on different days.

Our system includes a complete tool for running a wide range of real-time fMRI studies with different experimental protocols, different analysis methods, and different virtual environments, using the Unity game engine⁷ (Unity Technologies, California). The values that take part in the computation can be inspected online, during the experiment, and are logged offline for further analysis. The system allows easily configuring classification and interaction parameters during an experiment, playing back experimental sessions, and interfacing with external devices. In this paper we describe the first analysis method that we have used, which was deliberately kept simple, based on comparing activation levels of ROIs.

3. The ROI-based paradigm

An experiment is divided into three parts. The first part is intended for localization of brain areas. The subject sees an avatar standing in the center of a three-door room, as seen in figure 1. The subject is given pseudo-random instructions ('left', 'right', and 'forward'). Six seconds after each action the subject is instructed to rest and during that time the avatar executes a pre-determined command that corresponds with the instruction. The rest duration is 8–10 s; we avoided a fixed rest period in order to avoid habituation and anticipation of the next cue. The 'right' and 'left' commands cause the avatar to turn toward the right and left doors correspondingly and the 'forward' command causes the avatar to move toward the door on the front. The entire session is recorded for the purpose of finding ROIs: selecting a group of voxels that are

⁷ <http://unity3d.com/>

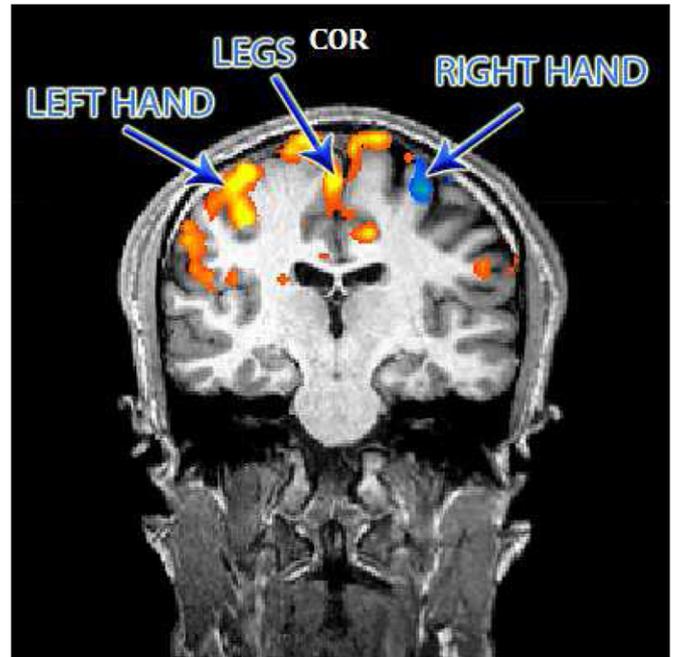


Figure 2. An example taken from one subject over the first stage of the experiment, intended for localization of ROIs. The three regions (from left to right) represent the three areas correspondingly: left-hand, legs, and right-hand, in the primary motor cortex, and are delineated by a left versus right-hand contrast as well as a legs versus baseline contrast, using a GLM analysis.

more active in one experimental condition compared to the other conditions, as detected by a general linear model (GLM) analysis. The experimenter manually marks the ROIs inside the most saturated regions in yellow and blue for the three classes, as seen in figure 2. Figure 3 depicts the event-related average time-course of the contrast. We assume that there are inter-subject differences in the size and position of specific ROIs. For every subject we calibrate the ROIs only once at the beginning of the experiment, and continue to use them in every subsequent run, over multiple days.

In the second part of the experiment we instruct the subject to rest for 1 min, and the system collects the mean and standard deviation (STD) for each of the three ROIs for the entire baseline period.

In the third and last part, the task stage, the system collects the average values from each ROI every 2 sec. A classification is made every TR t using the Z-score formula:

$$z_t = \frac{x_t - \mu}{\sigma}. \quad (1)$$

where:

- x_t is the average raw value in an ROI at TR t ;
- μ is the mean raw value of the ROI in the baseline period; and
- σ is the STD value of the ROI in the baseline period.

The selected class is the one corresponding to the ROI with the maximal Z-score value, and the system transmits the classification result to the Unity engine for execution. Each ROI is mapped to a different action performed by the subject: turning left, right, or walking forward corresponds to left-hand, right-hand, or leg imagery, respectively. Each action executed by the avatar is a fixed step that takes under 2 s to complete.

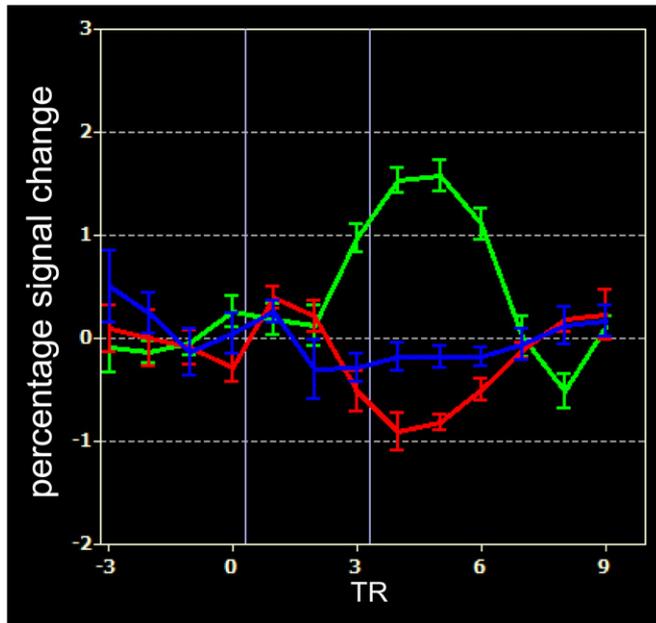


Figure 3. An example of an event-related averaging plot for left-hand ROI, from a complete run, taken from one subject. The green curve represents the hemodynamic response evoked by the left-hand cue at the left-hand ROI and the red and blue curve represent the right-hand and legs respectively. The cue is provided in TR 0, and peak occurs at TR 5.

4. Cue-based experiment

4.1. Method

We first performed a cue-based experiment: the subject is given auditory cues that instruct him what to ‘think’, in pseudo-random order. This has the advantage that we can compute the accuracy of the BCI, but it does not provide the subject with a sensation of controlling an interface.

We tested two conditions: MI and motor movement (MM). In the latter case we allow subjects to move their toes and fingers. When using fMRI it is easy to verify that all information comes from the brain, and in our studies we make sure that the classification is based only on motor areas, and is not based on auditory processing (e.g., responding to the auditory cues) or visual areas (e.g., looking at the direction that you expect the avatar to move to). Eventually, BCI is intended for paralyzed patients. When our system would be used by paralyzed patients we expect things to be different in several respects: on the one hand, they would not need to actively suppress using their muscles and body; on the other hand, depending on the nature of their illness or injury the brain activity may be significantly different than that of healthy subjects. In this experiment we compare both cases of control. We allowed two groups to repeat the task until reaching reasonable performance, and thus there were differences in the number of runs performed by the subjects.

Written informed consent was obtained from all volunteers. The study was approved by the Ethics Committee of the Weizmann Institute of Science, which complies with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Six subjects participated in the experiment.

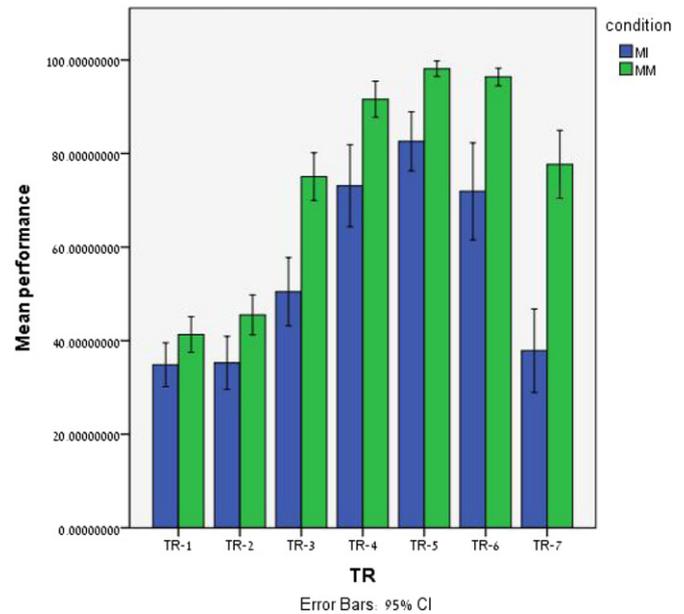


Figure 4. A comparison between the average classification accuracy of groups A (MM) and B (MI). Each TR is 2 s, and the TTF is 10 s.

Group A consisted of five subjects who performed the experiment by using MM (three females, mean age 27). Group B consisted of three subjects: two who also belonged to group A and an additional subject (overall two females, mean age 26). All subjects were healthy and right handed. We have verified that subjects were not moving their body during the MI experiments by visual inspection (the subject is visible to the fMRI operators through a transparent window and is covered by two video cameras). In addition, subject S1 performed one session with simultaneous electromyogram (EMG) recording. Data (bandpass 1–5000 Hz; sampling rate 10 000 Hz) obtained from the hand and leg (extensor digitorum communis and tibialis anterior) was collected in one experimental session to validate that no muscle activity is involved in MI. A comb band stop filter was used with a fixed value of 16 Hz to remove repetitive noise that came from scanning 32 slices every 2000 ms.

The structure of this step is identical to the structure of the localization stage, with the only difference being that the feedback is based on classification result rather than on the trigger. In each run of the cue-based experiment the subjects received 30 triggers, ten from each class, in a pseudo-random order. The trigger was given by a voice command. The subject was expected to keep the mental imagery in his mind (or perform a physical action in the MM condition) for three TRs (6 s), until he heard a ‘rest’ cue. This was followed by either four or five TRs of rest in between triggers; the rest between triggers was pseudo random, in order to avoid expectation and habituation. The classification and the feedback took place 10 s after voice trigger.

4.2. Results

The chance level in this 3-class task is 33.3%. Figure 4 describes the average classification results in the first seven

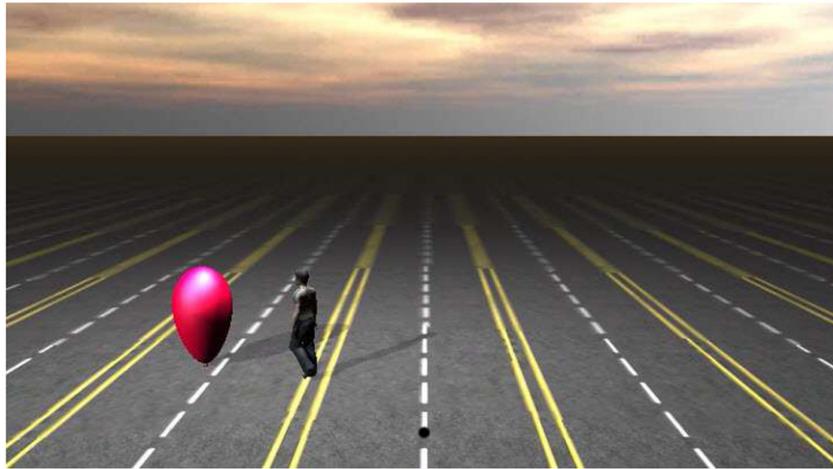


Figure 5. The subject sees an avatar standing in the center of the space (on top of the black circle) and needs to freely navigate the avatar toward the red balloon.

TRs following a trigger. Classification accuracy coincides with the hemodynamic response: in TR 1 and TR 2 the accuracy is around chance level, then it gradually increases with the best accuracy in TR 4 to TR 6 (8–12 s after the trigger), and gradually drops to chance level. The classification accuracy in the MM condition was significantly higher ($p < 0.00001$) than in the MI condition. A GLM analysis using Bonferroni corrections indicates that the performance after the first two TRs was significantly lower than in all the other TRs ($p < 0.00001$), and the performance in TRs 3 and 7 is significantly higher than TRs 1 and 2 ($p < 0.00001$) and significantly lower than the performance in TRs 4, 5, and 6 ($p < 0.00001$).

5. Free-choice experiment

5.1. Method

In the free-choice task the subject sees an avatar standing in the center of a space (figure 5). The subject is instructed to control the avatar using MM or MI, and the subject has to reach the balloon by using a minimal number of steps (figure 6). At the beginning of each trial a red balloon appears in front of the subject, in a different location. The subject then hears a voice command that instructs him to start the trial. During the trial the subject does not know when the trial ends, is not limited to a fixed path, and can choose his actions freely. Each run is composed of six trials, and each trial lasted 1 min and 48 s, with 12 s rest between trials. Each subject performed several such 12 min runs per session.

Performance is affected by the nature of the task. There would be a great difference between navigation tasks that require frequent control changes (such as ‘left’, ‘forward’, ‘right’) and tasks that require long sequences of the same command. In principle, only tasks of the first kind would pose a substantial test of the subject’s ability to control the interface and change commands quickly. Our tasks were relatively simple, as in principle they did not require many changes; the subject was expected to rotate to the correct angle and then continuously move forward. In practice, however, most subjects did not have a 100% control of the interface.

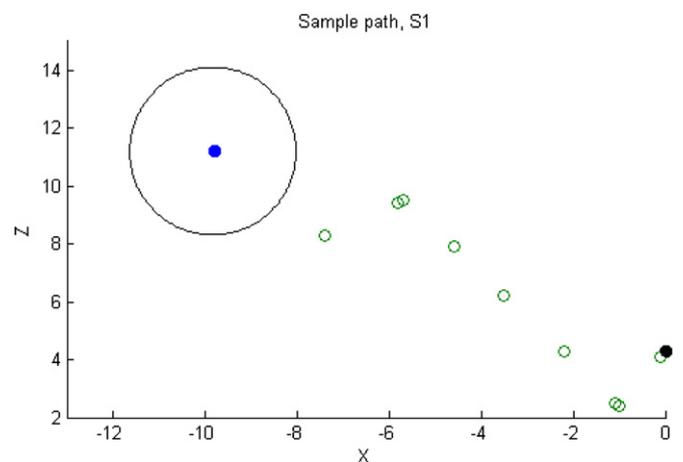


Figure 6. A birds-eye view of a near optimal trial performed by subject S1. The black spot indicates the starting point. The subject needs to guide the avatar toward the balloon, which is indicated as a blue spot. When the avatar reaches the surrounding perimeter around the balloon, indicated as a circle, we consider the trial to be successful.

Thus, in practice, even in this simple navigation task, subjects often had to switch commands in order to correct for their previous errors. For example, unpracticed subjects would often overshoot the rotation due to the delay, and would then have to keep rotating in both directions till they faced the balloon directly.

Each subject started performing the task with a TTF of 8 s (four TRs). When the subject was able to reach the balloon in at least half of the trials he continued to perform the same task with a smaller TTF, including three, two, and one TRs (6, 4, and 2 s, correspondingly). In all cases the feedback was based on the classification result of the last TR within the TTF duration. Only one command was sent to the avatar at each TTF interval, and the avatar’s walking distance and rotation had fixed values. The time for each trial was fixed so the number of possible steps varies as we varied the TTF. In larger TTFs we expected the subjects to have a more accurate control, but they were given less opportunities for errors.

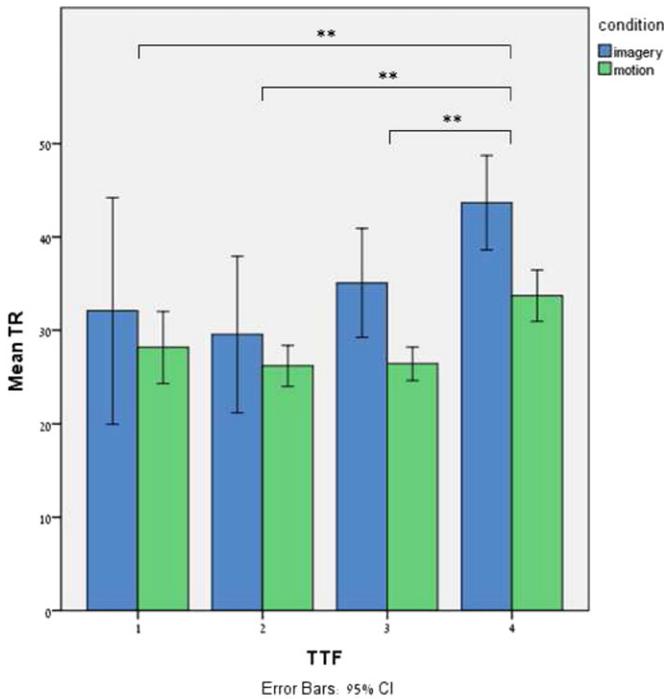


Figure 7. Average time to reach the balloon in both imagery (MI) and motion (MM), in the four different TTFs: 1–4 TRs, or 2–8 s (** indicates $p < 0.01$).

Seven subjects participated in the experiment. Group A, using MM, consisted of the initial five subjects of group A with one additional subject. Group B, using MI, was the same as in the MI cue-based experiment. All subjects were healthy and right-handed. As in the cue-based experiments we have verified that subjects were not moving their body during the experiments by visual inspection and EMG, and that the selected voxels in the ROIs are not taken from auditory or visual areas.

5.2. Results

In free-choice tasks it is impossible to know for certain what the subject was trying to achieve. Therefore, it is not possible to provide a direct comparison of free-choice versus cue-based performance, and in the free-choice task we provide two estimates of the performance. The shortest path to the balloon was either 5 or 6 steps. Figure 7 summarizes the time it took the subjects to reach the balloons. If a subject failed to reach the balloon in the allocated time then we assign the maximum time (108 s) for that trial. A two-way factorial ANOVA with factors condition (MM versus MI) and TTF, and Bonferroni correction for multiple comparisons, indicates that MI required significantly more time to reach the balloons than MM ($p < 0.00001$). The time required to reach the balloons when the TTFs were 2, 4 and 6 was significantly shorter than with TTF 8 ($p = 0.001$). There were no significant differences between TTFs 2, 4, and 6 and no significant interaction effect was found between condition and TTF.

The second measurement, performance, was calculated as follows:

$$\gamma = \frac{\alpha}{\beta} \quad (2)$$

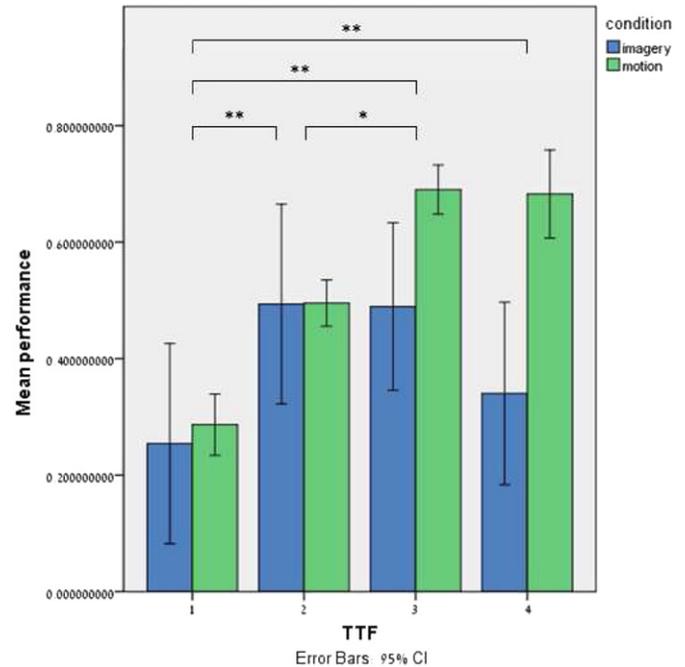


Figure 8. Average performance in both conditions: imagery (MI) and motion (MM), in the four different TTFs: 1–4 TRs, or 2–8 s (* indicates $p < 0.05$ and ** indicates $p < 0.01$).

where:

- α is the optimal (hypothetic) number of steps required to reach the balloon in the specific trial.
- β is the actual number of steps taken by the subject in the specific trial.
- γ is the inverse of the overhead. If a subject was not able to reach the balloon we assign a value of 0.

Figure 8 summarizes the subjects' performance. A two-way factorial ANOVA with factors condition (MM versus MI) and TTF, and Bonferroni correction for multiple comparisons, indicates that the performance in the MM condition was significantly better than the MI condition ($p < 0.00001$). The performance with a TTF of 6 s was the best; this was significantly better than TTFs 2 ($p < 0.00001$) and 4 ($p = 0.03$); the difference between TTFs 6 and 8 was not significant ($p = 0.061$). Rather, an interaction between condition and TTF was found significant for TTF 8 ($p < 0.00001$); i.e., the difference between MM and MI was especially large in TTF 8.

5.3. Subjective reports

The fMRI environment is far from natural and subjects need to get used to it: they are expected not to move their head during scans and stay still for a long duration of 1 or 2 h. They can only communicate with the experimenter during breaks between sessions, and the scanner is very noisy during the scan. When they get used to the scanner, the subjects report finding our tasks entertaining. Although we have not used stereo (depth) display, the subjects report being highly immersed in the virtual scene, since this is the only thing they can see, and it fills in a large part of their field of view. Some of the subjects were able to reach the target quickly, and then

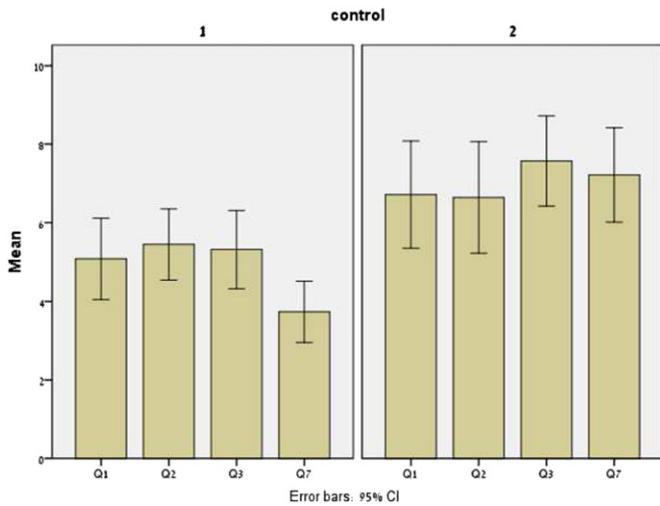


Figure 9. Reported degree of embodiment, in the free-choice experiment, using (1) motor imagery and (2) motor movement, for questions 1, 2, 3 and 7.

continued to explore the environment; e.g., they have tried to rotate around the balloon or even take their avatar outside the frame.

The subjects were asked to fill questionnaires and were interviewed multiple times, at the end of most of the free-choice sessions. The questionnaires included 14 questions with 10-step Likert scale. The analysis is based on 52 questionnaire instances.

The questionnaires included questions about the degree of feeling embodied in the avatar (four questions), the degree of the sense of control (four questions), and additional questions that we ignore here since they did not reveal significant effects. A multivariate analysis of variance (MANOVA) indicated that all factors had significant effects: subject ($p = 0.00005$), control type (motion versus imagery) ($p = 0.01$) and TTF

($p = 0.05$). There were no significant interaction effects. Further ANOVA tests with Bonferroni corrections indicated that the reporting of the sense of embodiment with the avatar was higher when controlling the avatar using imagery than when the avatar was controlled by motion (figure 9). This was significant in two out of the four questions ($p = 0.02$ and $p = 0.0001$) and nearly significant in the other two questions ($p = 0.063$ and $p = 0.076$). ANOVA tests also revealed that subjects reported a significantly lower sense of control in the shortest TTF (2s) as compared to the other TTF conditions (figure 10). This was significant in three questions ($p < 0.01$) and nearly significant in the fourth question ($p = 0.062$).

6. Discussion

An important issue with fMRI-based BCI is the delay between initiating a mental pattern ('thought') and the amount of time to detect it in the BOLD signal. The cue-based task results indicate that we can expect the highest accuracy 8–12 s following a cue. However, in the free-choice task the best performance was achieved 6 s following the cue. Our explanation for this difference is that subjects can adapt to the delay, probably by planning ahead their activity. A previous study have found out that subjects perform better neurofeedback with intermittent feedback than with continuous feedback [31]. Our study suggests that subjects can perform well with continuous feedback, given appropriate context and a long enough TTF.

Performing the task using imagery is more difficult than performing it when being allowed to move the fingers and toes. Our performance measurement is more sensitive than measuring the time, and it reveals that the difference between motion and imagery is most evident when the TTF is 8 s; i.e., it seems that subjects cannot 'keep the imagery' for more than a few seconds.

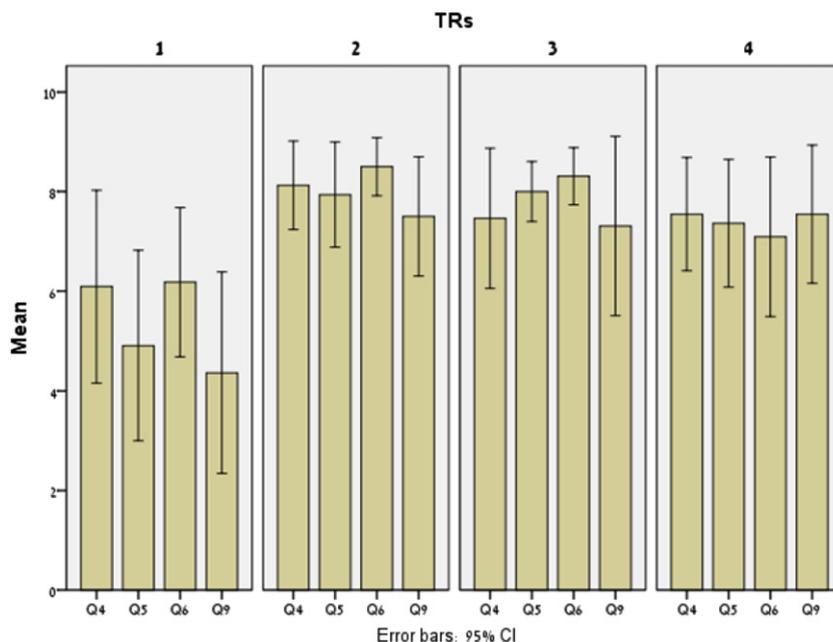


Figure 10. Reported degree of control with the different TTFs of 1–4 TRs, in the free-choice experiment, for questions 4, 5, 6 and 9.

When comparing the objective performance results with the subjective reports an interesting incongruence relates to a TTF of 4 s. The degree of control that the subjects reported is significantly higher than that reported for a TTF of 2 s, and seems to be as high as for TTFs of 6 and 8 s. However, the performance in TTF 4 was significantly lower than the performance in TTFs 6 and 8.

Subjects using imagery felt a higher sense of being embodied in the avatar. This could be a result of the increased effort and mental focus. Regardless of the explanation, this raises hope for rehabilitation applications of BCI, which lack proprioceptive feedback.

7. Conclusion and future work

The results of this study indicate that subjects can learn to control an avatar using motor imagery or movement, classified by our system online from fMRI data. We note that high performance is achieved with very little training, whereas motor-imagery based BCI using EEG typically requires extensive training. fMRI-based BCI may provide an essential tool in the future, when preparing for invasive BCI surgery, or for rehabilitation. fMRI can also aid in recognizing new mental patterns for BCI and developing new BCI paradigms, and the patterns can then be searched for in other types of signals, such as EEG, fNIRS, and invasive methods.

The ROI-based method we have presented here is simple and computationally efficient; we plan to extend it using machine learning techniques in order to identify more specific multi-voxel brain patterns that may lead to identifying more complex intentions. We also hope to explore global brain activity rather than only the activity in specific ROIs, taking advantage of the fact that the whole brain is scanned. Finally, in the course of these studies we also intend to further explore how the sensation of agency and embodiment develop in the context of such BCI experiences.

Acknowledgments

This research is supported by the European Union FP7 Integrated Project VERE (No 657295), www.vereproject.eu. We would like to thank the subjects for their participation, and the Weizmann Institute fMRI scanner staff Edna Furman-Haran, Nachum Stern and Fanny Attar for their help in this experiment.

References

- [1] Leuthardt E C, Schalk G, Wolpaw J R, Ojemann J G and Moran D W 2004 A brain–computer interface using electrocorticographic signals in humans *IEEE Trans. Systems Man Cybern. Part A* **1** 63
- [2] Donoghue J P, Nurmikko A, Black M and Hochberg L R 2007 Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia *J. Physiol.* **579** 603–11
- [3] Kim S, Simeral J D, Hochberg L R, Donoghue J P, Friehs G M and Black M J 2011 A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair *IEEE Trans. Neural Syst. Rehabil. Eng.* **19** 193–203
- [4] Hochberg L R et al 2012 Reach and grasp by people with tetraplegia using a neurally controlled robotic arm *Nature* **485** 372–5
- [5] Faller J, Müller-Putz G, Schmalstieg D and Pfurtscheller G 2010 An application framework for controlling an avatar in a desktop-based virtual environment via a software ssvep brain–computer interface *Presence: Teleoperators Virtual Environ.* **19** 25–34
- [6] Touyama H and Hirose M 2007 Steady-state VEPs in CAVE for walking around the virtual world *Universal Access in Human-Computer Interaction. Ambient Interaction (Lecture Notes in Computer Science vol 4555)* pp 715–7
- [7] Prueckl R and Guger C 2009 A brain–computer interface based on steady state visual evoked potentials for controlling a robot *Bio-Inspired Systems: Computational and Ambient Intelligence (Lecture Notes in Computer Science vol 5517)* ed J Cabestany, F Sandoval, A Prieto and J Corchado (Berlin: Springer) pp 690–7
- [8] Ortner R, Guger C, Prueckl R, Graenbacher E and Edlinger G 2010 SSVEP based brain–computer interface for robot control *Computers Helping People with Special Needs (Lecture Notes in Computer Science vol 6180)* ed K Miesenberger, J Klaus, W Zagler and A Karshmer (Berlin: Springer) pp 85–90
- [9] Shen H, Zhao L, Bian Y and Xiao L 2009 Research on SSVEP-based controlling system of multi-DoF manipulator *Advances in Neural Networks—ISNN 2009 (Lecture Notes in Computer Science vol 5553)* ed Y Wen, H Haibo and N Zhang (Berlin: Springer) pp 171–7
- [10] Long J, Li Y, Wang H, Yu T, Pan J and Li F 2012 A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair *IEEE Trans. Neural Syst. Rehabil. Eng.* **20** 720–9
- [11] Gergondet P, Druon S, Kheddar A, Hintermuller C, Guger C and Slater M 2011 Using brain–computer interface to steer a humanoid robot *ROBIO '11: IEEE Int. Conf. on Robotics and Biomimetics* pp 192–7
- [12] Gergnodet P, Kheddar A, Hintermuller C, Guger C and Slater M 2013 Multitask humanoid control with a brain–computer interface: user experiment with HRP-2 *ISER: 13th Int. Symp. on Experimental Robotics* (Berlin: Springer) pp 215–29
- [13] Bayliss J D 2003 Use of the evoked potential P3 component for control in a virtual apartment *IEEE Trans. Rehabil. Eng.* **11** 113–6
- [14] Piccione F, Priftis K, Silvoni S, Piron L, Merico A, Tonin P, Vidale D and Furlan R 2007 Integration of a p300 brain computer interface into virtual environment *Virtual Rehabilitation, 2007* (Piscataway, NJ: IEEE) p 88
- [15] Neuper C and Pfurtscheller G 2001 Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates *Int. J. Psychophysiol.* **43** 41–58
- [16] Friedman D, Leeb R, Pfurtscheller G and Slater M 2010 Human-computer interface issues in controlling virtual reality with brain–computer interface *Hum. Comput. Interact.* **25** 67–93
- [17] Pfurtscheller G, Leeb R, Keinrath C, Friedman D, Neuper C, Guger C and Slater M 2006 Walking from thought *Brain Res.* **1071** 145–52
- [18] Leeb R et al 2006 Walking by thinking: The brainwaves are crucial, not the muscles! *Presence: Teleoperators Virtual Environments* **15** 500–14
- [19] Royer A S, Doud A J, Rose M L and He B 2010 EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** 581–9
- [20] Leeb R, Friedman D, Muller-Putz G, Scherer R, Slater M and Pfurtscheller G 2007 Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: A case

- study with a tetraplegic *Comput. Intell. Neurosci.* **2007** 79642
- [21] Bakardjian H, Tanaka T and Cichocki A 2010 Optimization of SSVEP brain responses with application to eight-command brain-computer interface *Neurosci. Letters.* **469** 34–8
- [22] Coyle S M, Ward T E and Markham C M 2007 Brain-computer interface using a simplified functional near-infrared spectroscopy system *J. Neural Eng.* **4** 219
- [23] Sitaram R, Caria A and Birbaumer N 2009 Hemodynamic brain-computer interfaces for communication and rehabilitation *Neural Netw.* **22** 1320–8
- [24] Pfurtscheller G, Allison B Z, Brunner C, Bauernfeind G, Solis-Escalante T, Scherer R, Zander T O, Mueller-Putz G, Neuper C and Birbaumer N 2010 The hybrid BCI *Front. Neurosci.* **4** 3
- [25] Wienbruch C, Candia V, Svensson J, Kleiser R and Kollias S S 2006 A portable and low-cost fMRI compatible pneumatic system for the investigation of the somatosensory system in clinical and research environments *Neurosci. Lett.* **398** 183–8
- [26] deCharms R C 2008 Applications of real-time fMRI *Nature Rev. Neurosci.* **9** 720–9
- [27] Berman B D, Horovitz S G, Venkataraman G and Hallett M 2012 Self-modulation of primary motor cortex activity with motor and motor imagery tasks using real-time fMRI-based neurofeedback *Neuroimage* **59** 917–25
- [28] deCharms R C, Christoff K, Glover G H, Pauly J M, Whiteld S and Gabrieli J D 2004 Learned regulation of spatially localized brain activation using real-time fmri *Neuroimage* **21** 436–43
- [29] Rota G, Sitaram R, Veit R, Erb M, Weiskopf N, Dogil G and Birbaumer N 2009 Self regulation of regional cortical activity using real-time fmri: the right inferior frontal gyrus and linguistic processing *Hum. Brain Mapp.* **30** 1605–14
- [30] Caria A, Sitaram R and Birbaumer N 2012 Real-time fmri: a tool for local brain regulation *Neuroscientist* **18** 487–501
- [31] Johnson K A, Hartwell K, LeMatty T, Borckardt J, Morgan P S, Govindarajan K, Brady K and George M S 2012 Intermittent real-time fmri feedback is superior to continuous presentation for a motor imagery task: a pilot study *J. Neuroimaging* **22** 58–66
- [32] Yoo S S, Fairney T, Chen N K, Choo S E, Panych L P, Park H W, Lee S Y and Jolesz F A 2004 Brain-computer interface using fMRI: Spatial navigation by thoughts *Neuroreport* **15** 1591
- [33] LaConte S M, Peltier S J and Hu X P 2007 Real-time fMRI using brain-state classification *Human Brain Mapp.* **28** 1033–1044
- [34] Eklund A, Ohlsson H, Andersson M, Rydell J, Ynnerman A and Knutsson H 2009 Using real-time fMRI to control a dynamical system by brain activity classification *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009* (Berlin: Springer) pp 1000–8
- [35] Hornyak T 2006 Thinking of child's play *Sci. Am.* **295** 30
- [36] Lee J H, Ryu J, Jolesz F A, Cho Z H and Yoo S S 2009 Brain-machine interface via real-time fMRI: preliminary study on thought-controlled robotic arm *Neurosci. Lett.* **450** 1–6
- [37] Minati L, Nigri A, Rosazza C and Bruzzone M G 2012 Thoughts turned into high-level commands: Proof-of-concept study of a vision-guided robot arm driven by functional MRI (fMRI) signals *Med. Eng. Phys.* **34** 650–8
- [38] Andersson J P 2012 Visuospatial attention based brain-computer interfacing using real-time fMRI *PhD Thesis* Utrecht University, The Netherlands
- [39] Cohen O, Druon S, Lengagne S, Mendelsohn A, Malach R, Kheddar A and Friedman D 2012 fMRI based robotic embodiment: a pilot study *IEEE/RAS-EMBS Int. Conf. on Biomedical Robotics and Biomechatronics (BioRob)* (Roma, Italy, 24–28 June 2012) pp 314–9
- [40] O'Hara K, Sellen A and Harper R 2011 Embodiment in brain-computer interaction *ACM: Proc. SIGCHI Conf. on Human Factors in Computing Systems* pp 353–62
- [41] Friedman D, Leeb R, Guger C, Steed A, Pfurtscheller G and Slater M 2007 Navigating virtual reality by thought: What is it like? *Presence: Teleoperators Virtual Environments* **16** 100–10
- [42] van der Kouwe A, Gicquel S, Chen G, Schmitt F, Harder M, Salat D, Sorensen A G, Fischl B and Dale A 2003 On-line automatic slice positioning and between-scan correction for brain mr protocols *Proc. 11th Annu. Meeting of ISMRM*, vol 11 (Toronto, Canada) p 797
- [43] Mueller E, Thesen S, Heid O and Schad L R 2000 Prospective acquisition correction for head motion with image-based tracking for real-time fMRI *Magn. Reson. Med.* **44** 457–65