Enabling Bio-Feedback Using Real-Time fMRI

Henrik Ohlsson1, Joakim Rydell2,3, Anders Brun2,3, Jacob Roll1
Mats Andersson2,3, Anders Ynnerman3,4 and Hans Knutsson2,3

1 Division of Automatic Control, Department of Electrical Engineering, Linköping University, Sweden
2 Division of Medical Informatics, Department of Biomedical Engineering, Linköping University, Sweden
3 Center for Medical Image Science and Visualization, Linköping University, Sweden
4 Division of Visual Information Technology and Applications, Department of Science and Technology, Linköping University, Sweden

(ohlsson, roll)@isy.liu.se {joary, andbr, matsa, knutte}@imt.liu.se andyn@itn.liu.se

Abstract—Despite the enormous complexity of the human mind, fMRI techniques are able to partially observe the state of a brain in action. In this paper we describe an experimental setup for real-time fMRI in a bio-feedback loop. One of the main challenges in the project is to reach a detection speed, accuracy and spatial resolution necessary to attain sufficient bandwidth of communication to close the bio-feedback loop. To this end we have banked on our previous work on real-time filtering for fMRI and system identification, which has been tailored for use in the experiment setup.

In the experiments presented the system is trained to estimate where a person in the MRI scanner is looking from signals derived from the visual cortex only. We have been able to demonstrate that the user can induce an action and perform simple tasks with her mind sensed using real-time fMRI. The technique may have several clinical applications, for instance to allow paralyzed and "locked in" people to communicate with the outside world. In the meanwhile, the need for improved fMRI performance and brain state detection poses a challenge to the signal processing community. We also expect that the setup will serve as an invaluable tool for neuroscience research in general.

I. INTRODUCTION

Revealing the functionality of the human brain continues to be one of the grand scientific challenges. Although considerable effort has been made toward this end, many issues remain unresolved.

A new tool in this endeavor is functional Magnetic Resonance Imaging (fMRI). The aim in fMRI is to map cognitive, motor and sensor functions to specific areas in the brain [22]. The physical foundation for the method is the fact that oxygenated and deoxygenated blood have different magnetic properties. When a neuron in the brain is active it consumes oxygen, which is supplied by the blood. To compensate for the increased rate of oxygen consumption in an active brain area the blood flow is increased and the result is that the oxygenation level of the blood to this area is, in fact, increased. This increase, commonly known as the BOLD (Blood Oxygen Level Dependent) effect, can be measured in a magnetic resonance scanner. Thus, we can locate areas of brain activity indirectly by locating areas with elevated blood oxygen levels.

To map, for example, the sensory function area of a finger, one can stimulate the finger on a volunteer with a brush, while images of the brain are continuously acquired by the MR-scanner. During the stimulation of the finger there is an increase in image intensity (i.e. the active area becomes brighter) compared to a resting state. Thus, to detect activity we need to compare images where the finger is stimulated by the brush to images acquired in a resting state. The areas where the “activated” images are brighter than images acquired in the “rest” state indicate brain areas involved when the brush stimulates the finger.

In the project presented in this paper, we aim at using the estimates of brain activity for the purpose of bio-feedback, i.e. to use the information obtained in the fMRI scan to alter the stimuli generating the fMRI response and thus generating a feedback loop involving the brain. This requires that all parts of the loop, in particular the brain activity estimation, run in real-time. To capture real-time dynamics of the brain, we must acquire each image-slice rapidly. Unfortunately, this makes the images heavily contaminated with random noise. Hence, it is not enough to acquire just one image in activity and one in rest, as it is likely that we can not detect any significant change in intensity due to the high noise level. How the experiment and acquisition of the image volumes are performed is termed the paradigm and is, as a rule, a determining factor for success or failure.

Bio-feedback has since long been explored using electromyography (EMG), temperature and electroencephalography (EEG), see among others [12], [8], [10], [21], [2], [13], [15], [20], [19], [6], but is relatively new in the field of fMRI. Some of the most known examples are the one by DeCharms et al., who showed how patients suffering from chronic pain could learn how to control their pain by bio-feedback based on fMRI [4], and the one by Yoo et al., who made it possible to navigate throw a 2D maze through fMRI bio-feedback [24].

The long term vision behind the present project is to apply techniques used in system identification for the analysis and ‘control’ of brain activity. Potentially the ‘state of mind’ could be steered towards a goal state (activation pattern) by producing a sequence of stimuli that is dependent on the estimated activation pattern sequence. A dual view is that a person can be told to try to make the stimuli produced move towards a target stimulus by will. In the future it may in this way be possible to analyze certain brain functions in terms of brain state transition probability matrices.
However, being in the startup phase of the project, the goal of this first experiment has been to explore the response times that can be expected using fMRI for bio-feedback. We have chosen to work with measurements from the visual cortex, and based on those, track the sight of a person in the MRI scanner.

The paper is structured as follows: We start by formulating our problem in Section II and follow up by describing the experiments setup in Section III. The way we have chosen to solve the problem is presented in Section IV, followed by a description of obtained results in Section V. We finish with a discussion in Section VI.

Fig. 1. The MRI scanner used in the experiments.

II. PROBLEM DESCRIPTION

As an example of generating stimuli based on feedback from an fMRI signal, and thereby closing the loop, we here consider a visually-based experiment.

The stimuli are selected to consist of a flashing checkerboard, placed either on the left or the right of the screen. The aim of the experiment is to make the non-flashing part of the visual stimuli to follow the eye movements of the subject, i.e., to flash to the left if the subject is looking to the right, and to flash on the right side if the subject is looking to the left. Hence, the problem is to detect where the subject is looking at the moment, using the measured fMRI data. Once this is done, the stimulus is simply set to the opposite side.

To judge if the subject is looking to the left or to the right, we need to build a prediction model, with the measurements from the fMRI as the input, and the direction of the subject’s gaze as the output. This is a regression problem of high-dimensional nature. The input, i.e., the fMRI measurements, will typically be a signal of approximately 40000 elements or dimensions. Without any kind of regressor selection or regularization, we would therefore get a severe overfit to estimation data.

For the particular experiment setup described, we could use a two-class classifier to determine whether the subject is looking to the left or right. However, aiming at an extension where the stimulus can be moved more than to the left or the right side of the field of vision, regression was considered and not classification.

Previous attempts to handle fMRI data have used a range of various methods, from sliding-window general linear modeling (GLM) to support vector machines (SVM), see e.g. [17], [18], [3], [7], [5], [16]. A good overview is given in [1].

III. EXPERIMENT SETUP

As mentioned, the goal of this first real-time feedback experiment has been to create a simple eye-tracker, which will detect if the subject in the scanner is looking to the left or right and show a flashing checkerboard on the right or left 30% of the screen, respectively (see Figure 2, left figure).

The data was acquired using a 1.5 T Philips Achieva MR scanner, see Figure 1. The acquisition resolution was 80 by 80 pixels in each slice, and 7 slices were acquired. Field of view and slice thickness were chosen to obtain a voxel size of approximately 3 × 3 × 3 mm. The use of cubic voxels make three-dimensional signal processing (e.g. smoothing) viable. The acquired data cover the primary visual cortex, and a surface coil was used to provide an optimal signal-to-noise ratio within this region. To obtain high BOLD contrast, the echo time (TE) was set to 40 ms and the repetition time (TR) was set to 1000 ms. Hence we acquire one volume per second, which we consider to be sufficient to deliver close to real-time feedback to the subject.

The subject in the scanner was exposed to a visual stimulus through a pair of head mounted displays. The data processing was done in Matlab on a standard laptop.

Fig. 2. Visual stimuli used. Left figure: left 30% of the screen as a flashing checkerboard. Right figure: a centered vertical stripe, covering 100% vertically and 40% horizontally of the screen.

IV. TRAINING AND REAL-TIME fMRI

Before starting the real-time feedback phase, a training phase was performed to build a prediction model.

A. Training phase

During the training phase, two training data sets were gathered. First, the subject in the scanner was exposed to a flashing checkerboard, a centered vertical stripe covering 100% vertically and 40% horizontally of the screen. Figure 2 shows the visual stimulus used. Data was gathered for approximately 40 seconds.

The second training data set was gathered by instructing the subject in the scanner to look away from a periodically...
shifting flashing checkerboard (15 seconds flashing checkerboard on the left, 15 seconds flashing checkerboard on the right, see Figure 2). Data was gathered for approximately 90 seconds.

Using this last data set, 8 voxels were picked out correlating the best with the paradigm. The reason for not just using the two best correlating voxels was to be able to use the redundancy in data to reduce the impact of noise. The 8 voxels were picked out by first computing the correlation to a sine wave with a period of 30 seconds. This was done voxel-by-voxel. In order not to have to go through all possible phase shifts for the sine wave, to find the phase shift associated with the best correlation, canonical correlation analysis (CCA, [11]) was used. In this context, CCA has the property to automatically find the time delay in the sine wave giving the best correlation. Note that this usage of CCA would not be possible using a square wave. The voxel with the best correlation was chosen as the first of the 8 voxels. The three voxels with a phase within 90 degrees of the first one and with the highest correlations were also picked out. Finally, the 4 voxels correlating best with a sine wave at least 90 degrees out of phase compared to the best correlating voxel were chosen.

At this time, the voxel locations were verified to be within the visual cortex. This was done manually by inspection of a plot like the one shown in Figure 3. To further reduce noise and to gain some robustness against movements of the subject, the two training data sets were spatially smoothed. Note that this will turn the 8 chosen voxels into 8 neighborhoods, centered at the previously chosen voxels.

The 8 chosen neighborhood signals were then picked out from the two training data sets, detrended voxel-by-voxel, and merged together (90 seconds of data associated with the left-right stimuli followed by 40 seconds of data associated with the centered vertical flashing stripe). Finally, a linear predictor, using the 8 signals as regressors, was fit to a square wave, switching between −1 and +1 (in phase with first sine wave used above), and followed by zeros for the last 40 seconds. Hence, the predictor was expected to give −1 if the subject was looking to the left of the checkerboard, +1 if the subject is looking to the right, and zero otherwise.

The training phase is summarized in Algorithm 1.

**B. Real-time phase**

During the real-time data phase, the data was first spatially smoothed, just as the training data set. The signals from the 8 chosen neighborhoods were then detrended using a windowed least squares (WLS) approach, with a window size of 50 seconds. With \( \tilde{X}_i(t) \) being the data at time \( t \) from neighborhood \( i \), let

\[
\tilde{X}_i(t) = [\tilde{X}_i(t) \ \tilde{X}_i(t-1) \ \ldots \ \tilde{X}_i(t-50)].
\]

We can remove a linear trend in \( \tilde{X}_i(t) \) by subtracting the best fitted line

\[
\tilde{X}_i(t) = \tilde{X}_i(t) - \begin{bmatrix} \alpha_i & \beta_i \end{bmatrix} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ t & t-1 & \ldots & t-50 \end{bmatrix}.
\]

**Algorithm 1 Training phase**

Given data from a voxel \( i \) associated with stimulus on the left-right, \( X_{i}^{lr}(t), t = 1 \ldots 90 \), and from stimulus as a centered vertical stripe, \( X_{i}^{v}(t), t = 1 \ldots 40 \).

1) Use CCA to find how well \( X_{i}^{lr}(t) \) correlates to a sine wave with a period of 30 seconds.
2) Find the index for the voxel with the highest correlation.
3) Find the three voxels with the highest correlation but with a phase difference less then 90 degrees compared to the best correlating voxel.
4) Find the 4 voxels with the highest correlation having a phase difference of more than 90 degrees compared to the best correlated voxel.
5) Make sure that the chosen voxels are in the visual cortex.
6) For the chosen voxels, make a spatial smoothing using a Gaussian spatial filter to obtain \( \tilde{X}_{i}^{lr}(t) \) and \( \tilde{X}_i^{v}(t) \).
7) Detrend, voxel-by-voxel, the signals \( \tilde{X}_{i}^{lr}(t) \) and \( \tilde{X}_i^{v}(t) \) from the 8 chosen neighborhoods.
8) Concatenate the detrended \( \tilde{X}_{i}^{lr}(t) \) and \( \tilde{X}_i^{v}(t) \) to form \( \tilde{X}_i(t) \).
9) Find the \( \theta_i \) such that

\[
\sum_{t=1}^{130} [y(t) - \sum_{i=1}^{8} \theta_i X_i(t)]^2
\]

is minimized; \( y(t), t = 1 \ldots 90 \) being a \(-1/1\) square wave in phase with the best correlated voxel, and \( y(t) = 0, t = 91 \ldots 130 \).

where \( \alpha_i, \beta_i \) minimizes

\[
\left\| \tilde{X}_i(t) - \begin{bmatrix} \alpha_i & \beta_i \end{bmatrix} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ t & t-1 & \ldots & t-50 \end{bmatrix} \right\|^2.
\]

The first element in \( \tilde{X}_i(t) \) after the trend has been removed is used as input for the linear predictor. The resulting signal from this procedure will take values close to one when the subject is looking to the right and minus one when the subject is looking to the left. The flashing checkerboard was therefore moved to the left side when the predictor signal exceeded a certain threshold, and correspondingly for the right side.

For validation, the subject in the scanner was during the real-time phase instructed to keep its eyes on a moving point on the screen. In this way, we could keep track of where the subject was looking, which was used to validate the results.

The real-time phase is summarized in Algorithm 2.

**V. RESULTS**

Figure 3 shows the 8 voxels picked out in the training phase. The 4 voxels correlating best with the flashing checkerboard on the left are shown in the top row of Figure 3. The best correlation was computed for the voxel shown in the first column from the left, second best for the second column from the left and so on. A correlation of 0.6 was the highest correlation computed, and the signal from this voxel during the training phase is shown in the top figure of Figure 4.
Algorithm 2 Real-time phase

Given new data $X_i(t)$. Let $T$ be a threshold and assume that the $\theta_i$ and the 8 chosen neighborhoods are given from the training phase. Do the following:

1) For the chosen voxels, perform a spatial smoothing using a Gaussian spatial filter to obtain $\tilde{X}_i(t)$.
2) Detrend, voxel-by-voxel, the signals $\tilde{X}_i(t)$ from the 8 chosen neighborhoods to get $\tilde{\tilde{X}}_i(t)$.
3) Compute $\hat{y}(t) = \sum_{i=1}^{8} \theta_i \tilde{\tilde{X}}_i(t)$.
4) If $\hat{y}(t) < -T$: move the stimulus to the right side; if $\hat{y}(t) > T$: move the stimulus to the left side; and if $-T < \hat{y}(t) < T$: use the same stimulus as for $t-1$.

The bottom row of Figure 3 shows voxels with the highest correlation to stimuli on the right, arranged in the same way as the top row. As can be seen, the neighborhoods shown in the second row, columns 2–4, are not within the visual cortex. The signals from these neighborhoods were therefore not considered. The signal from the voxel correlating best (correlation 0.55) with stimuli on the right side is shown in the bottom figure of Figure 4.

The signal from the 5 remaining neighborhoods were weighted together to give an as good fit to the stimuli as possible (see Figure 5).

Figure 6 shows logged results from the real-time phase using the computed weighting and choice of neighborhood. The horizontal coordinate for the reference point where the subject in the scanner was aiming to look at is shown in the top subplot. The computed signal from the fMRI data is given in the middle subplot. The bottom subplot shows if the flashing checkerboard is to the left or the right ($-1$ if the checkerboard is to the left and +1 if it is to the right). It can be seen that, as the subject shifts focus from one side to the other, it takes between 2.5 and 7 seconds until the visual stimulus has changed.

VI. DISCUSSION

It should be emphasized that the purpose of this work has not been to introduce a method for an eye-gaze interface; the authors are well aware that there exist more simple, inexpensive and exact solutions for that specific purpose. The main contribution is instead the closing of the bio-feedback loop where the user experiences a real-time response from the state of his or her mind and is able to perform a simple task.

The choice of a visual stimulus is not of central importance for this work. A reason for choosing the specific experimental setup was that MR-compatible goggles provide a simple perception of a stimulus inside the MR-scanner, and the flashing checkerboard pattern enables a distinctive activation in the visual cortex due to both temporal variation and spatial high contrast edges.

The use of an MR-scanner as a Brain Computer Interface (BCI) in a real-time bio-feedback loop stresses the
Fig. 6. Logged results. Top figure: The reference signal showing where the subject should focus. A small value corresponds to the subject in the scanner looking to the left, while a high value corresponds to looking to the right. Middle figure: Computed signal from the fMRI measurements. Bottom figure: The location of the stimuli. Small value: flashing checkerboard on the left part of the screen; high value: checkerboard on the right part of the screen.

boundaries for image acquisition and signal processing to the absolute limit. In our current setup an average user experiences a response time of 5 seconds. However, we observed times down to 2.5 seconds. Similar results have recently been shown by LaConte et al [17]. Considering that the BOLD signal, in itself, has a response time of the same order, these response times can be seen as quite good results. However, it has been shown that it is possible to spot activity in the BOLD signal considerably earlier, see [14] and [23]. The question of whether these early signs of activity are large enough to be able to reliably detect activity is still open. MRI is continually improving with respect to acquisition time, SNR and resolution. A limiting factor for functional-MRI is the temporal dynamics of the BOLD response. For the visual cortex, stimuli like the flashing checkerboard pattern induce a BOLD response that is present for approximately 30 seconds [9]. During the first half, the BOLD signal increases in intensity apart from a very small initial dip. After that time, the blood oxygen control system of the brain compensates the blood oxygen distribution for this new state, and the BOLD response disappears.

An objective method to evaluate the performance of such a real-time fMRI system is to estimate the bandwidth in the bio-feedback loop. For the present setup the bandwidth is approximately 0.2 bits/s. A shorter acquisition time (currently about 1 s) will not by itself be a key factor to increase of the bandwidth above 1 bit/s limit, considering the temporal dynamics of the BOLD response. An improved SNR of the MRI would on the other hand provide the means to discern the BOLD response within the noise at a much earlier stage in the activation process, which has the potential to increase the bandwidth several orders of magnitude. This is a real future challenge both for the manufacturers of MRI equipment as well as for the signal processing community.

Although it is convenient to use visual stimuli inside the MR-scanner some issues must be considered. During the training phase, both unconscious and reflex-based eye movements degrade the training data. Using more advanced VR-goggles with an eye tracker device that fixate the stimuli at a local area in the visual cortex, independently of the eye motions of the user, would provide a significant improvement of the training data set. An additional problem using a gaze based BCI is that the user may unintentionally move the head a little synchronously to the movement of the gaze. These motion artifacts are the main reason why neighborhoods outside the visual cortex sometimes may provide high correlation to the paradigm. To detect and compensate for occasional head motions would improve the performance of the real-time phase. The head motion can be modeled as a rigid body motion and the new locations of the selected neighborhoods are straight forward to compute once the global head motion is estimated. To compensate for a user that continuously moves his or head is much more cumbersome due to the complex motion artifacts which are associated to MRI. Detection and compensation for small occasional head movements should be possible to perform within this setup.

A next step in our research is to extend the simple left/right response to a more complicated task involving a graded response. A possible task would be a virtual pole balancing problem. Such a graded response could be computed in different ways, but a straight-forward method is to apply a temporal integration on the present output signal.

A possible way to further increase the bandwidth in the bio-feedback loop would be to use parallel or sequential activation of different brain areas. Broca’s and Wernicke’s areas are e.g. activated in speech processing using language or signs. An activation in these areas could be deliberately induced by the person in the scanner by focusing the mind on a sentence, which can be done without any movement of the eyes. Activating several cortical areas at once will make the training phase more complex, and more advanced adaptive training methods will be required to fully explore these possibilities. To optimize the BCI bandwidth for a specific task, adaptation to each user’s own capabilities is necessary.

VII. ACKNOWLEDGMENTS

This work was supported by the Strategic Research Center MOVIII, funded by the Swedish Foundation for Strategic Research, SSF.
REFERENCES


