1 Title page

Name of the project: ArmBand

inverse reinforcement learning to control a robotic arm using a Brain-Computer Interface

Principal investigator:

Dr. Laurent Bougrain

Abstract

The goal of this project is to use inverse reinforcement learning to better control a JACO robotic arm developed by Kinova in a Brain-Computer Interface (BCI). Asynchronous BCI such as motor imagery based-BCI allows the subject to give orders at any time to freely control a device. But using this paradigm, even after a long training, the accuracy of the classifier used to recognize the order is not 100%. While a lot of studies try to improve the accuracy using a preprocessing stage that improves the feature extraction, we propose to work on a post-processing solution. The classifier used to recognize the mental commands will provide as outputs a value for each command such as the posterior probability. But the executed action will not only depend on this information. A decision process will also take into account the position of the robotic arm and previous trajectories. More precisely, the decision process will be obtained applying an inverse reinforcement learning on a subset of trajectories specified by an expert.

2 Project's objectives

Brain-Computer interfaces (BCI) [1] interpret brain activity to produce commands on a computer or other devices like a robotic arm (see figure 1). A BCI therefore allows its user, and especially a person with high mobility impairment, to interact with its environment only using its brain activity.

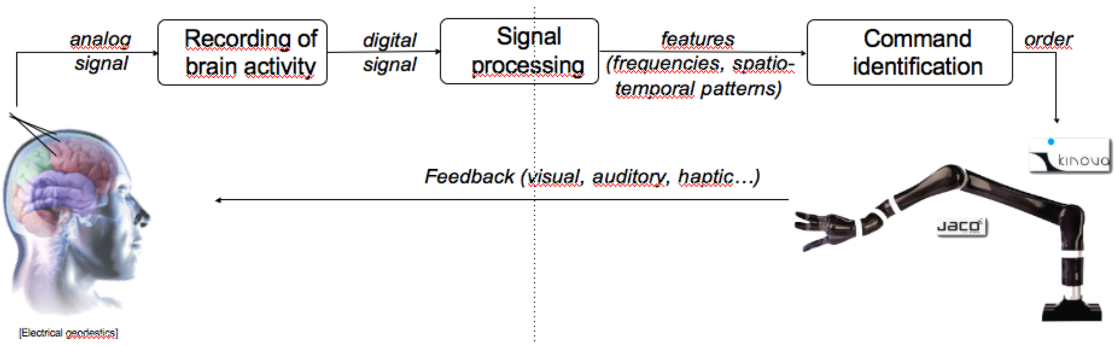

Figure 1: The Brain-Computer Interface loop : from electroencephagraphic signals acquisition, feature extraction and classification to feedback. Our project will add a decision process based on an inverse reinforcement learning in the command identification module.

Overcome the variability of the mental command

A major difficulty to properly interpret the mental command lies in the fact that brain activity is very variable even if a particular task is reproduced identically. Beyond the noise acquired by the recording system, background brain activity, concentration, fatigue or medication of the subject are the source of this variability. This variability makes it difficult for the classifier to recognize the different mental commands. Specific preprocessings such as common spatial pattern filter [2] are useful to help distinguish the mental command. However, this effort is not always sufficient. It therefore becomes necessary to explore new solutions to address this variability.

Interest in reinforcement learning

Thus, it is now necessary to make decision systems able to deal with this variability. This is why some projects introduce a reinforcement learning in their BCI system such modifying the classifier [3]. We propose to use reinforcement learning in a broader context.

Proposed study

We plan to show in this project that a reinforcement learning improves the control of a robotic arm. More precisely, the decision process will take into account a subset of trajectories specified by an expert and the position of the robotic arm in addition to the usual outputs of the mental commands classifier.

3 Background information

Electroencephalography

Electroencephalography (EEG) is a brain imagery technique based on a simple principle. Tiny electrical signals are emitted by groups of neurons that fire almost simultaneously. These discharges create potential differences which can be measured by electrodes distributed over the scalp. EEG has several advantages compared to other brain imagery techniques. It is a non invasive technique. The system is small, light and cheap. It is portable and can indeed be embedded in a wheelchair. The installation needs few minutes. The instantaneous readings can observe brain function in real time. However, it has some defects such as a bad spatial resolution (centimeter), a noisy signal and needs an everyday calibration.

Motor imagery based-BCI

EEG expresses various brain activities through physiological phenomena associated with them: a variation in amplitude for the evoked potentials or a synchronization/desynchronization for a motor imagery. More precisely, in the latter case, a motor execution (and planning) changes the amplitude spectrum, especially in the μ (8-12 Hz) and β (14-25 Hz) bands, in the motor cortex area devoted to the body part moved (figure 2):

- μ waves get out-of-sync 2.5s before movement performed to resynchronize in 2.5 seconds.
- β waves are out-of-sync 1s before the movement and synchronize strongly 1s after movement.

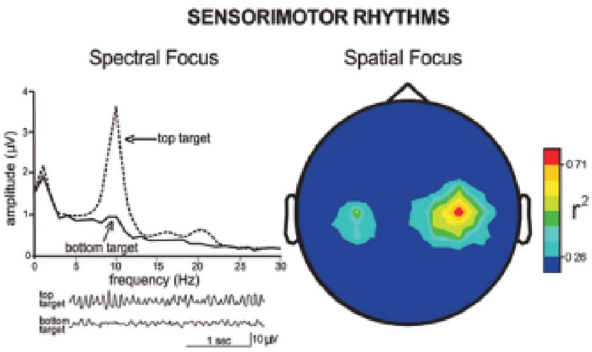

Figure 2: Energy modulations observed in the motor cortex (left) of subjects after training during a motor imagery. The location of these changes correspond to the motor area of the left hand (right) [4].

These frequencies waves also vary during a movement imagination. By detecting these variations in the motor areas of the subject including a patient suffering of a motor impairment, a BCI can detect the imagination a movement. The paradigm developed by Pfurtscheller et al. [5] is the basis for our BCI. An already implemented protocol using OpenViBE[1] will allows the participant to record signals corresponding to the imagination of several motor actions (left hand, right hand, feet…) (see figure 3). A classifier will be trained to discriminate the various motor imageries. A posterior probability will be associated to each class. Thus even if the accuracy is not 100%, these informations will guide the decision system to propose the best next action taking also into account the history.

Inverse reinforcement learning

Inverse reinforcement learning (IRL) is a particular case of learning from demonstration. The aim of an IRL-agent is to learn to perform a specific task by observing an expert demonstrating the task. In the IRL framework, the expert is assumed to act so as to be rewarded when the task is correctly completed [6]. The environment in which the expert evolves is modeled as a Markov Decision Process (MDP), that is a set of state in which the agent can perform actions. After each action the state changes according to a Markovian transition probability [7].

From the agent perspective, the reward is unknown and has to be inferred from the demonstration of the expert. Demonstrations are seen as trajectories in the state-action space of the MDP. After having discovered the reward, the agent can then act so as to maximize the same reward as the expert. The main problem of IRL is that the problem is ill-posed and that an infinite number of reward functions can be the cause of the behavior

[1] http://openvibe.inria.fr

of the expert, including the uniform zero function (any behavior is optimal according to a null reward) [8]. Many researches have been conducted during the last decade to solve this problem by adding constraints [9] or modeling the problem as a classification problem [10] and many others.

In this project, IRL will be used to discover the reward that makes a movement natural according to the human. This reward will be used to correct BCI errors in a way that the generated motions will look like the human user would have generated it. By observing a person moving the arm using a joystick, the reward function will be inferred and used as a mean to correct the control of the arm.

4 Detailed technical description

(a) Technical description

WP 1 : Motor imagery paradigm and its use for 2D/3D control.

The first task will be to review the paradigm proposed by Graz on motor imagery [11] i.e. the features extracted from the EEG signals and the protocol parameters (stages, duration. . .). Particular importance will be attached to know in the literature the largest number of motor imageries used until now in relation to the accuracy obtained by different classifiers [12, 13].

The participants will be able to experiments the Motor Imagery protocol using an existing OpenViBE scenario (see figure 3). This step will allow us to know the performance of a basic motor imagery based BCI using different movement imaginations (left hand, right hand, feet. . .) and various classifiers (linear discriminant analysis, support vector machine. . .).

WP 2: Which GUI to control a robotic arm?

Based on preliminary studies [14, 15] and taking into account the number of movements selected for controlling the robotic arm, an interface will be done to provide a first brain-machine interface (see figure 4). This solution will provide a baseline performance when no reinforcement learning is used. More precisely, the following steps have to be done:

1. Record the EEG responses to the Graz protocol
2. Compute the Common Spatial Pattern and the classifier from the recorded EEG signal.
3. Use this classifier to control on line the robotic arm

WP 3: Add the inverse reinforcement learning

A general implementation of the inverse reinforcement learning is available in C++. The different state variables have to be defined to apply the algorithm on our specific problem. The state variables will contain the joint robotic arm and the outputs of the classifier. More precisely, the following steps have to be done:

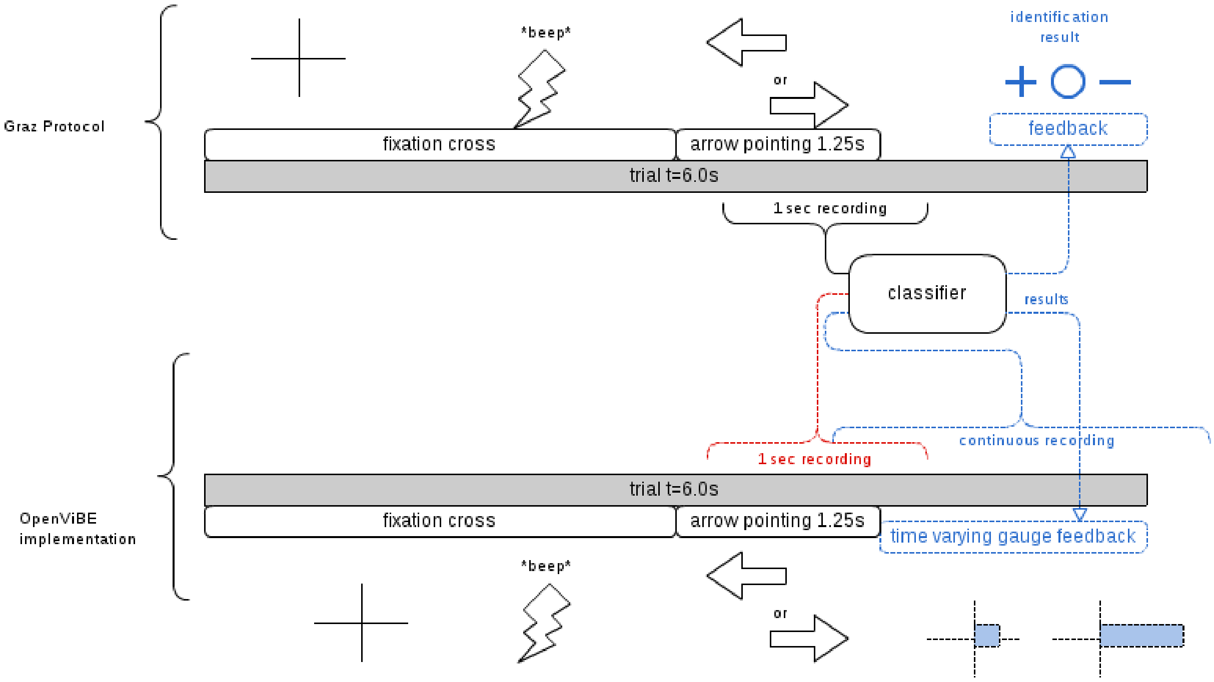

Figure 3: Adapted Graz BCI protocol used within OpenViBE.

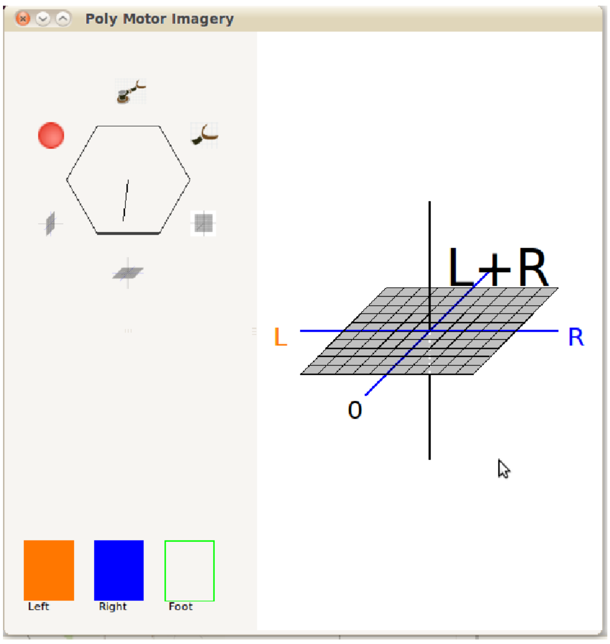

Figure 4: A possible interface to control the robotic arm [15].

1. Record the trajectories of the robotic arm and the values of the joints robotic arm (while pressing a button using a joystick from different starting positions for example)
2. For each trajectory, build a sequence of BCI commands (right, left, forward. . . .)
3. Replay the recorded trajectories and record the brain activities associated to the BCI commands
4. Compute from the recorded brain activity the outputs of the classifier
5. Train the inverse reinforcement learning algorithm using the joints robotic arm, the outputs of the classifier and the correct action. The Brain-Machine interface is completed now.
6. Control on-line the robotic arm using the Brain-Machine Interface and compare the performances (accuracy, time. . .) with a classic BCI system (see WP 2).

(b) Resources needed

The project will provide a Jaco robotic arm by Kinova, a Micromed EEG amplifier (SD LTM 32) and computers with the OpenViBE software and several scenarios for controlling the robotic arm. In addition, a C++ library including several inverse reinforcement learning algorithms will be available.

Jaco robotic arm JACO is a robotized arm developed by Kinova. It is a 6-axis robotic manipulator with a three fingered hand (http://www.youtube.com/watch?v=e65PLikG1lk). This upper arm has 6 degrees of freedoom for the arm and 3 fingers that can be opened and closed altogether or just using the index finger and the thumb linked. Originally it is possible to use a 3 axis joystick to control the arm. The arm has sensors (force, position and accelerometer) and can reach an object up to 90 cm at a maximum speed of 30 cm/sec. This arm can be added to a wheelchair and is especially suitable for a person with a disability of the upper limbs.

OpenVibe OpenViBE is a C++ open-source software devoted to the design, test and use of Brain-Computer Interfaces [16]. The OpenViBE platform consists of a set of software modules that can be integrated easily and efficiently to design BCI applications. Key features of the platform are its modularity, its high-performance, its portability, its multiple-users facilities and its connection with high-end/Virtual Reality displays. The "designer" of the platform enables to build complete scenarios based on existing software modules using a dedicated graphical language and a simple Graphical User Interface (GUI). This software is available on the INRIA Forge under the terms of the LGPL-V2 license (http://openvibe.inria.fr).

reinforcement learning program Several IRL algorithms have already been coded by Supélec IMS team members (especially Edouard Klein, member of the project). The code library, written in C++, will be made available to the personal affected to the project. It contains algorithms as [9] or [17]. More details can be provided upon request.

Figure 5: Jaco robotic arm by Kinova.

5 Work plan and implementation schedule

	week 1					week 2					week 3					week 4				
	2	3	4	5	6	9	10	11	12	13	16	17	18	19	20	23	24	25	26	27
WP1: Motor imagery paradigm																				
Task 1.1 Review of the Graz BCI																				
Task 1.2 Experiment the Graz BCI on OpenViBE																				
Task 1.3 Ajust the OpenViBE scenario (nb of motor imageries)																				
WP2: Which GUI to control a robotic arm?																				
Task 2.1: Study various GUI used for Graz BCI																				
Task 2.2: Propose and build with OpenViBE a GUI for the robotic arm																				
Task 2.3: Record the EEG responses to the Graz BCI using the GUI																				
Task 2.4: Build classifier from the recorded EEG signal.																				
Task 2.5: Use this classifier to control on line the robotic arm																				
WP3 : Add the inverse reinforcement learning																				
Task 3.1 Review the inverse reinforcement learning																				
Task 3.2 Record the trajectories of the robotic arm																				
Task 3.3 Build sequences of BCI commands																				
Task 3.4 Record the brain activities associated to the BCI commands																				
Task 3.5 Compute the outputs of the BCI classifier																				
Task 3.6 Train the inverse reinforcement learning algorithm																				
Task 3.7 Control on-line the robotic arm using the BCI																				
Task 3.8 Compare the performances (accuracy, time. . .)																				
WP4 : Valorisation																				
Task 4.1 Write a report/article																				
Task 4.2 Make a video																				
Task 4.3 Build a webpage																				

6 Benefits of the research

The main benefits will be to show how an inverse reinforcement learning can help to share the control of a device with the user.

The results of the project will be presented in a specific webpage on the site of the workshop. This webpage will contain an overview of the method and a video of the control of the arm by a BCI using an inverse reinforcement learning.

The full method will be presented in an article and will be submitted to a review such as Frontiers in Neuroprosthetics, IEEE Transactions on Biomedical Engineering or Journal of Neural Engineering.

7 Profile of the team:

(a) Leader

Laurent Bougrain is an associate professor at the university of Lorraine (France). He is a member of the Inria team CORTEX dedicated to computational neuroscience at LORIA[2]/INRIA Nancy grand Est. He has been working for more than a decade on time series analysis with a focus on experimental data obtained during neuroscientific experiments. In recent years, he has dedicated his research to Brain-Computer Interfaces (BCI). He is working on template-based classifier for single trial detection using multichannel denoising techniques. He is the winner of the international BCI competition IV of the challenge about predicting the finger flexion from ECoG in 2008 [18]. He is currently working on a project on reinforcement learning to control a robotic arm and a wheelchair from EEG. He also collaborates to the worldwide BCI-software OpenVibe (http://openvibe.inria.fr).
E-mail: bougrain@loria.fr
Web: `http://www.loria.fr/~bougrain`,

(b) Staff proposed by the leader

Other researchers needed (describing the required expertise for each, max 1 page)

Edouard klein: Ph.D Student in inverse reinforcement learning.
Edouard Klein was born in 1987. He received the Electrical Engineering degree of PHELMA (Grenoble, France) in 2010. Since that, he has worked on reinforcement learning and learning from demonstration in the IMS research group of Supélec. He is also a PhD student co-supervised by Yann Guermeur (ABC team, CNRS), Matthieu Geist (IMS research group, Supélec) and Olivier Pietquin (IMS research Group, Supélec and UMI 2958 (GeorgiaTech - CNRS)). The topic of is PhD is automatic feature selection in inverse reinforcement learning.

Participants should have ideally expertise in the field of brain-computer interfaces or at least one of the following domain: machine learning, Human-Computer Interface, electroencephalography and C++.

References

[1] J. R. Wolpaw, *et al.*, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002. [Online].

[2] Lorraine Research Laboratory in Computer Science and its Applications

Available: http://www.sciencedirect.com/science/article/B6VNP-45HFKTC-2/2/070472147433d00168e8d54909b982d2

[2] F. Lotte and C. Guan, "Spatially regularized common spatial patterns for EEG classification," in *International Conference on Pattern Recognition (ICPR)*, 2010.

[3] J. Fruitet, *et al.*, "Automatic motor task selection via a bandit algorithm for a brain-controlled button," INRIA, Research Report 7721, 2011. [Online]. Available: http://hal.inria.fr/inria-00624686/fr/

[4] J. N. Mak and J. R. Wolpaw, "Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects." *IEEE Rev Biomed Eng*, vol. 2, pp. 187–199, 2009. [Online]. Available: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=20442804

[5] G. Pfurtscheller, *et al.*, "Eeg-based discrimination between imagination of right and left hand movement," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 6, pp. 642 – 651, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0013469497000801

[6] S. Russell, "Learning agents for uncertain environments (extended abstract)," in *Proc. COLT*, 1998, p. 103.

[7] R. Sutton and A. Barto, *Reinforcement learning.* MIT Press, 1998.

[8] A. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proc. ICML*, 2000, pp. 663–670.

[9] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. ICML*, 2004.

[10] N. Ratliff, J. Bagnell, and M. Zinkevich, "Maximum margin planning," in *Proc. ICML*, 2006.

[11] C. Neuper, *et al.*, "Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain-computer interface." *Clinical Neurophysiology*, vol. 120, no. 2, pp. 239–247, 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19121977

[12] J. Philips, *et al.*, "Adaptive shared control of a brain-actuated simulated wheelchair," in *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, june 2007, pp. 408 –414.

[13] A. S. Royer, *et al.*, "EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies." *IEEE Trans Neural Syst Rehabil Eng*, vol. 18, no. 6, pp. 581–9, 2010.

[14] B. Blankertz, *et al.*, "The berlin brain-computer interface presents the novel mental typewriter hex-o-spell," in *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course*, 2006, pp. 108–109.

[15] L. Botrel, "Interface cerveau-ordinateur basée sur l'imagination motrice pour le contrôle 3d d'un bras robotique," Master de Sciences Cognitives et Applications, Université Nancy 2, 2011.

[16] Y. Renard, *et al.*, "Openvibe: An open-source software platform to design, test and use brain-computer interfaces in real and virtual environments," *Presence : teleoperators and virtual environments*, vol. 19, no. 1, 2010.

[17] E. Klein, M. Geist, and O. Pietquin, "Batch, Off-policy and Model-Free Apprenticeship Learning," in *Proc. EWRL*, 2011.

[18] N. Liang and L. Bougrain, "Decoding finger flexion from band-specific ecog signals in human," *Frontiers in Neuroprosthetics*, (submitted).

8 Additional information (mandatory)