fMRI Robotic Embodiment: A Pilot Study

Ori Cohen, Sébastien Druon, Sébastien Lengagne, Avi Mendelsohn,
Rafael Malach, Abderrahmane Kheddar and Doron Friedman

***Abstract*— We present a robotic embodiment experiment based on real-time functional Magnetic Resonance Imaging (rtfMRI). To our knowledge, this is the first time fMRI is used as an input device to identify a subject's intentions and convert them into actions performed by a humanoid robot. The process, based on motor imagery, has allowed subjects located in Israel to control a HOAP3 humanoid robot in France, experiencing the whole experiment through the eyes of the robot.**

I. Introduction

This work takes place in the context of the VERE project[1], a European research project that aims at dissolving the boundary between the human body and surrogate representations in immersive virtual reality and physical reality. By dissolving the boundary we mean that the subject is expected to have the illusion that his surrogate representation is his own body, and behave and think accordingly. This may help disabled humans to control an external device just by thinking, without any bodily movement being involved. As illustrated in Figure 1, our aim was to provide a subject located in Israel with the most intuitive thought-based control of a robotic avatar in France. To reach this goal, we have decided to focus on motor control, using an rtfMRI to detect the users movement intentions and translate them into actions performed by a HOAP3 humanoid robot. To our best knowledge, this pilot study is the first experiment of this type.

II. Previous works

Telerobotics is the technology that allows a human operator to steer robots at a distance. Telerobotic control strategies have evolved from the classical master-slave control to advanced supervisory control. Shared autonomy and the sophistication of the robotic control allows a telerobot to be steered by classical and modern input devices (such as keyboard, mouse, eye tracker, voice recognition systems, etc.), or through virtual reality functional intermediary [1]. But many recent works also

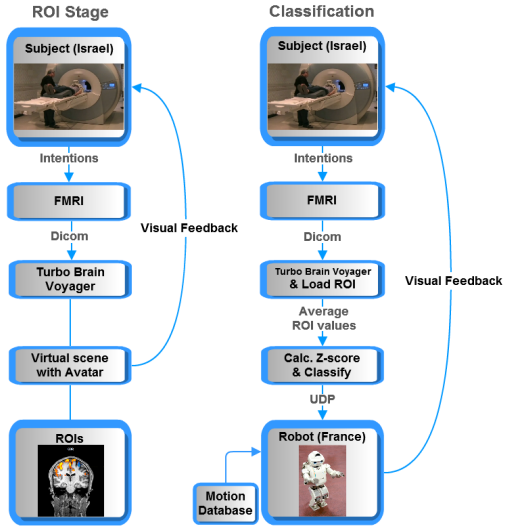

Fig. 1. Robotic Embodiment: General principle of data processing and experiment related tasks.

emphasize the possibility of using brain-computer interfaces (BCIs). BCIs allow a human to control a computer or a mechatronic device just by thinking, without any body movement being involved. While contemporary BCI systems are far from the interfaces imagined by Hollywood in movies such as Avatar[2] or Surrogates[3], there has been some progress made and a surge of interest in the last few years [2].

Most BCI systems intended for humans rely on the measurement of electroencephalogram (EEG) recorded from the scalp. BCI-controlled robots have been demonstrated using mainly three major EEG-based BCI paradigms: the steady state visually-evoked potential (SSVEP), the P300 wave, and motor imagery.

In SSVEP, a flickering visual stimulus is displayed to the subject. When the retina is excited by a signal ranging from 3.5Hz to 75Hz, the brain generates electrical activity at the same frequency of the visual stimulus, which can be detected in the EEG signal. SSVEPs are highly interesting for robot control due to their simplicity, their superior signal-to-noise ratio, and their high decision rate. Previous works include its use in the control of both mobile robots, e.g. [3][4],

S. Druon and A. Kheddar are with the CNRS-UM2 LIRMM UMR 5506, 161 rue Ada 34095 Montpellier Cedex 5, France, A. Kheddar is also with the CNRS-AIST Joint Robotics Laboratory, UMI3218, 305-8568 Tsukuba, Japan.

O. Cohen and D. Friedman are with the Interdisciplinary Center Herzliya (IDC H.), P.O. Box 167 Herzliya, 46150, Israel, O. Cohen is also with the Bar-Ilan University Ramat-Gan, 52900 Israel.

S. Lengagne is with the Karlsruhe Institute of Technology, Institute for Anthropomatics, Humanoids and Intelligence Systems Lab, Adenauerring 2, 76131 Karlsruhe, Germany.

A. Mendelsohn and R. Malach are with Weizmann Institute of Science, Department of Neurobiology, Rehovot 76100, Israel

[1] http://www.vereproject.eu

[2] http://www.imdb.com/title/tt0499549/

[3] http://www.imdb.com/title/tt0986263/

manipulators, e.g. [5] and recently, humanoids [6].

The P300 wave is an event-related potential (ERP) that appears 300ms after an infrequent task-related event. This ERP is now commonly used in BCI systems due to its reliability: the waveform is easily and consistently detectable, with little variation in measurement techniques. Even though the bitrate (i.e. the amount of commands which are sent to the external object in a second) is typically lower than SSVEP, it is still a reliable BCI system, included in many robotic control systems, see examples in [7] [8] [9] or [10].

Because they rely on visual or auditory stimuli, both SSVEP and P300 can be compared to a classical eye-tracking system in terms of input interface: they provide the same set of functionalities to the user and suffer from the same limitations. Mostly, the mapping between the user intentions and the functionality is arbitrary, in contrast to what we would expect from a "thought-based interaction paradigm.

Motor imagery has also been used for EEG-based BCI. Motor imagery is a mental process by which an individual rehearses or simulates a given action. As explained in [11], imagination of movement evokes brain networks that are similar to the networks evoked by real execution of the corresponding physical movement. A series of studies were carried out with motor-imagery based navigation of highly-immersive virtual reality [12][13][14] including experiments with a tetraplegic patient [15]. Motor-imagery requires more training and the bitrate is lower than P300 and SSVEP, but it is arguably based on a more intuitive mapping between the mental patterns and the resulting action taken by the system.

Since EEG is recorded from the scalp it suffers from high levels of noise and low spatial resolution as compared with other methods for recording brain activity. fMRI also has several drawbacks: it is expensive, less accessible, and has a low temporal resolution and a built-in delay because it is based on metabolic changes rather than on direct recording of electrical activity in the brain. However, due to its superior spatial resolution covering the whole brain simultaneously it holds much promise for completely new types of control paradigms.

We could find only one research project in the state-of-the-art presenting fMRI as an input device for a robotic hand control, and this was not applied to the embodiment problem. This work was performed by Honda Research Institute and Advanced Telecommunications Research (ATR) and concerned the control of a robotic hand with three predefined gestures; this non-published result was reported by [16].

Other Studies [17][18][19] used machine learning methods to decode brain patterns. In traditional fMRI experiments, we collect data from a subject's brain and have abundant time after the experiment is over to analyze the brain data. Additionally, algorithms do not need to be optimized for speed and for memory usage. Conversely, when dealing with real-time analysis and prediction, we need to use fast algorithms that can manipulate large datasets in fractions of a second. In our case the time between TRs, which is the time between our inputs, is 2s. In the current real-time experiments only three average raw values are calculated for the prediction, so one of the advantages of the simplicity of our method is its computational efficiency.

III. fMRI-BASED BCI

Our system is able to automatically identify a subjects intention based on motor imagery in real-time, classify brain activation patterns, and convert them into robotic actions performed by a small-size humanoid robot. The aim is to allow intuitive BCI control based on brain activity.

We present a set of preliminary successful BCI sessions preformed in an fMRI scanner. In this experiment subjects teleoperated a robot located in France from an fMRI scanner based in Israel. Each session lasted approximately 10–13 minutes.

A. 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) 2300ms, TE 2.98ms, $1mm^3$ voxels). For blood-oxygenation-level-dependent (BOLD) scanning, T2*-weighted images using echo planar imaging sequence (EPI) were acquired using the following parameters: TR 2000ms, TE 30ms, Flip angle 80, 35 oblique slices without gap, 20 towards coronal plane from Anterior Commissure-posterior Commissure (ACPC), $3 \times 3 \times 4$ mm voxel size, covering the whole cerebrum.

The data coming from the fMRI scanner is saved as Dicom files[4], and processed by Turbo BrainVoyager software (TBV) [20], which is a real-time processing, analysis, and visualization application that accepts input from an fMRI scanner. After processing the data, TBV saves the average raw data values for each region of interest (ROI) selected by the operator at each measured time point.

The fMRI scanner is located in Rehovot, Israel, and the robot in Béziers, France. The flow of commands was sent to the robot through a User Datagram Protocol (UDP) connection and the video flow was received through another network flow. The round trip time from transmission to

[4] http://medical.nema.org/

reception of data (ping) between Israel and France was between 100 to 150 milliseconds.

B. The ROI-based paradigm

Figure 2 depicts an image from TBVs view screen. 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 vs. right hand contrast as well as a legs vs. baseline contrast, using a general linear model (GLM) analysis.

A single experiment run is divided into three parts. In the first part, the data-gathering phase, the subject is given pseudo random motor-imagery instructions. The entire session is recorded for the purpose of finding ROIs. An ROI is an area in which the brain was more active in one experimental condition compared to the other condition. The operator places the ROIs inside the most saturated regions in yellow and blue where the event-related average signal for the current ROI is significantly higher than the other two ROIs, as seen in Figure 2. Figure 3 depicts the time-course of the contrast.

In the second part we instruct the subject to rest for one minute; this serves as a baseline resting period in which we collect the mean values for each ROI, by calculating the mean and standard deviation for each ROI for the entire baseline period.

In the third and last part, the task stage, we instruct the subject to imagine moving his limbs and collect the average values from each ROI every two seconds. A prediction is made using the Z-score formula: and is calculated for each measured value by using the mean and standard deviation from the baseline period:

$$z = \frac{x - \mu}{\sigma}. \tag{1}$$

where:

- x is the current measured average raw value in an ROI;
- μ is the mean raw value of the ROI in the rest period;
- σ is the standard deviation value of the ROI in the rest period.

The prediction chosen is the class corresponding to the ROI with the maximal Z-score value. The system then transmits the prediction to the HOAP3 robot located in France. Each ROI is mapped to a different action performed by the subject, which in turn activates a pre-computed robotic motion. Turning left, right or walking forward corresponds to left-hand, right-hand or legs imagery of movement.

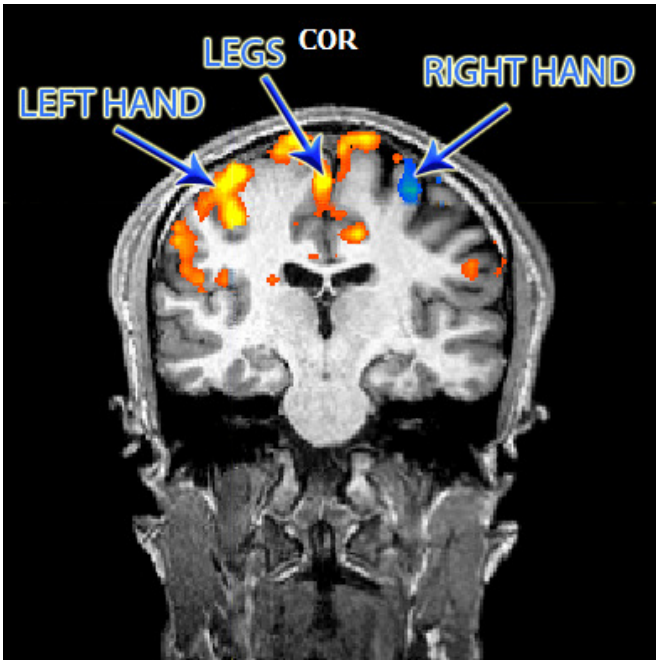

Fig. 2. An example of right-hand vs. left-hand contrast and legs vs. baseline contrast, taken from one subject over the first stage of the experiment. Since the contrast is very strong, even a simple classification based on the Z-score formula, as described here, works on a single trial basis.

C. Experimental setup

This study is a follow-up to a previous exploratory study, where a subject was able to control an avatar in a virtual environment, rather than a physical robot, using the same ROI paradigm [21]. In the study reported here we also used the avatar as a feedback in the first experimental stage, which was intended for defining three non-intersecting ROIs per subject. The subjects saw a virtual environment with an avatar standing in the bottom center of the space, and were instructed to imagine themselves as the avatar. The avatar would turn 90 degrees toward either the left or the right, or would walk 2 seconds when facing forward. In BCI we would like to achieve the most intuitive mapping between thought patterns and the resulting interaction [12]. Imagining hands for motion direction and feet for forward motion is not identical to the way you control your body when walking, but it clearly intuitive and not arbitrary.

In the next step the subjects viewed a live video feed through a camera located at the eyes of the robot. The robot was located inside a 9.6 meters by 5.3 meters room in France. They saw a technician who instructed them to move left, right or forward using his hand gestures. The objective given to the subjects in this experiment was to walk around two obstacles in an "eight"-shaped course, in approximate length of 1.5 meters. Each subject underwent between three to seven test sessions each lasting 13 minutes, in which the BOLD signal from the entire brain was measured every two seconds. At the end of this period we calculated two values: the mean signal and the standard deviation for the entire

rest period.

The third step included several task sessions, each session lasting 12 minutes (360 triggers). The system sent a nominal value to the robot every 2 seconds (corresponding to the ROI with the maximal Z-Score). The left or right commands initiated a thirty degrees turning sequence, and the forward command initiated a two-step forward walking sequence, both lasting between 8 and 14 seconds. The robot executed commands only after completing the previous command, i.e. many of the commands were ignored by the robot. In practice, while each command was based on a two second time window, the subjects focused on the same command (left, right, or forward) for the time it took the robot to perform the action.

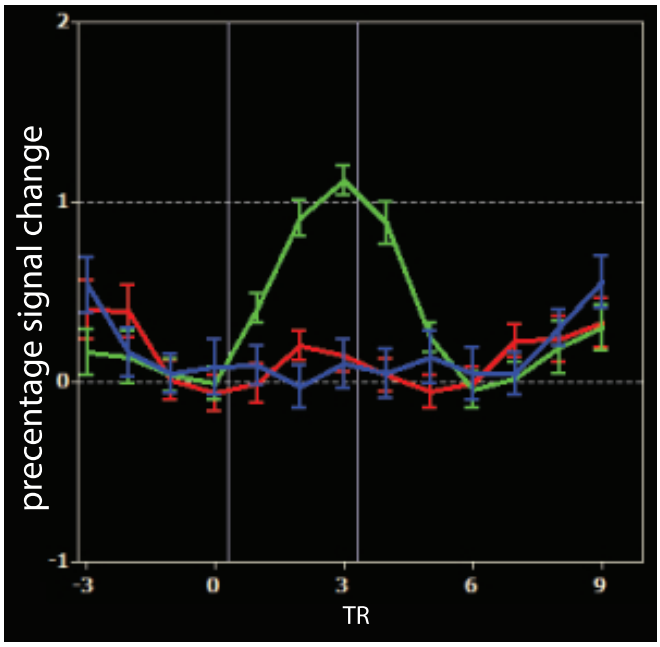

Fig. 3. An example of event-related averaging plot for left-hand ROI, taken from one subject. This plot depicts the average hemodynamic response evoked by the stimulus for an ROI over the first phase of the experiment. X and Y axes represent the TR position corresponding to the beginning of the event and the percentage signal change, respectively.

D. Limitations

One limitation concerns the ROI paradigm; using this method it is only possible to identify coarse ROIs, but it is impossible to identify more specific multi-voxel brain patterns that may lead to identifying more complex intentions; we are integrating such methods in our system for future studies.

We have tested the system with both motor imagery and motor movement; in the latter case subjects were instructed to move their fingers and toes in order to move, rather than the corresponding imagery tasks. In the case of motor movement the event-related brain signal is very strong; our experience indicates that any subject can perform the task without any difficulties and with literally no training. However, using motor imagery, the subjects need to be trained for several two hour sessions, because it is harder for them to find the right imagery strategy to activate the motor regions. Once they found their optimal strategy, we continue to use it in future sessions. For example, they imagine tapping with either right or left hand to turn, and imagine moving their feet back and forth to walk forward.

IV. Generation of the motion database for the robot

The use of a motion database was already described for a follower task with the HOAP-3 Robot [22]. In this paper we use the method presented in [23] to generate motions performing a sequence of contact stances and ensure the balance and the physical limits of the robot. Contrary to human-sized robots, HOAP-3 is a very stable robot so we are free to execute those motions with a local joint control loop, without using a balance stabilizer.

A. Motion optimization

As presented in [24], the goal is to generate motions that minimize a cost function and ensure a set of constraints relative to the balance and the limits of the robot and to the properties of the desired motion (e.g. position of the foot). In order to deal with the constraints classical optimization techniques revert to time discretization, even if they may produce unsafe motions where there are some constraint violations between the instant of the time-grid [22]. In order to avoid any constraint violation, this method considers a time-interval discretization that decomposes the motion into several intervals. We use a polynomial approximation over each time interval of any state variables of the robot, in order to easily take into account continuous inequality and equality constraints [23].

B. Motion properties

We create a database of motions in order for the robot to walk forward, turn to the left, or turn to the right. Each motion is decomposed into several contact phases, i.e., a lapse of time when no contacts are created or released. To ensure continuity, every motions starts and ends with the same posture. The turning motions are composed of five phases to perform a rotation of thirty degrees, whereas the walking motions are composed into nine phases and produce steps of five centimeters each.

During the optimization processes, we take into account the constraints regarding the balance, joint position, velocity and torque limits, and minimize the following function that produces a smooth and low-energy motion:

$$C(q) = a\int_0^T \sum_i \Gamma_i^2 dt + b\int_0^T \sum_i \ddot{q}_i{}^2 dt + cT \qquad (2)$$

with $a = 1e-2$, $b = 1e-5$ and $c = 4$, are the values we set heuristically to have human like walking motion.

V. Experimental Validation

Free choice scenarios allow for an experience of performing a task in a relatively natural and continuous mode, as opposed to trigger-based BCI experiences. A limitation of free-choice scenarios is that it is impossible to accurately measure success rates. In our experiment the subject was always successful in performing the task in the allocated time. Most trials were constructed so that time allowed for errors; in the scenario the subject was able to surround both obstacles by following visual instructions made by the technician. We cannot quantify success rates but it is clear that the probability of successfully completing the task with chance-level control is infinitesimally small.

We have validated this control system through a set of experiments where the subjects could steer the robot in three directions (Right, Left, Forward). The study included four right-handed participants, one male who performed the experiment several times using either exclusively motor imagery or motor movements, and three females who performed the motor movements.

The participants were asked to perform one of the three following missions:

- **Free navigation:** the user is allowed to visit the room freely
- **Seek and find:** an operator shows an object to the robot, then hides it. The subject has to navigate the room to locate it.
- **Follower:** an operator indicates by gestures to the robot the path that should be followed. This path was shaped like an eight around two obstacles in order to use all of the three basic movements.

The subjects gradually learned to control the robot during the session. As explained earlier (Section III.D) it is difficult to accurately assess BCI accuracy in free choice tasks. We have estimated the accuracy of the same method with an avatar-based experiment to be 100% for a two-class task (right hand vs. left hand) and 93% for the three-class task [21].

VI. Conclusions and future works

Our preliminary results indicate that subjects can learn to control a robot using either motor imagery or movement, classified by our system, in better-than-chance levels with very little training. Our aim is at allowing subjects to perform diverse tasks in the virtual or real environment, using a natural mapping of mental patterns to functionality. In the course of these studies, we intend to extend our method to use machine-learning techniques, and to explore how the sensation of agency and embodiment develop in the context of such BCI experiences.

VII. Acknowledgments

This research is supported by the European Union FP7 Integrated Project VERE (No 657295).

We would like to thank the subjects for helping us. We would also like to thank Dan Drai, and the Weizmann's institute technicians, Edna Furman-Haran and Fanny Attar for helping in this experiment.

References

[1] A. Kheddar, "Teleoperation based on the hidden robot concept," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 31, no. 1, pp. 1–13, 2001.

[2] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: Current state and future prospects," *IEEE reviews in biomedical engineering*, vol. 2, no. 1, pp. 187–199, 2009.

[3] R. Prueckl and C. Guger, "A brain-computer interface based on steady state visual evoked potentials for controlling a robot," in *Bio-Inspired Systems: Computational and Ambient Intelligence*, ser. Lecture Notes in Computer Science, J. Cabestany, F. Sandoval, A. Prieto, and J. Corchado, Eds. Springer Berlin / Heidelberg, 2009, vol. 5517, pp. 690–697.

[4] R. Ortner, C. Guger, R. Prueckl, E. Graenbacher, and G. Edlinger, "Ssvep based brain-computer interface for robot control," in *Computers Helping People with Special Needs*, ser. Lecture Notes in Computer Science, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds. Springer Berlin / Heidelberg, 2010, vol. 6180, pp. 85–90.

[5] H. Shen, L. Zhao, Y. Bian, and L. Xiao, "Research on ssvep-based controlling system of multi-dof manipulator," in *Advances in Neural Networks - ISNN 2009*, ser. Lecture Notes in Computer Science, W. Yu, H. He, and N. Zhang, Eds. Springer Berlin / Heidelberg, 2009, vol. 5553, pp. 171–177.

[6] P. Gergnodet, S. Druon, A. Kheddar, C. Hintermuller, C. Guger, and M. Slater, "Using brain-computer interface to steer a humanoid robot," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Karon Beach, Thailand, 7-11 December 2011, pp. 192–197.

[7] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. N. Rao, "Control of a humanoid robot by a noninvasive brain computer interface in humans," *Journal of Neural Engineering*, vol. 5, no. 2, p. 214, 2008.

[8] B. Rebsamen, E. Burdet, C. Guan, H. Zhang, C. L. Teo, Q. Zeng, M. Ang, and C. Laugier, "A Brain-Controlled Wheelchair Based on P300 and Path Guidance," in *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on*. IEEE, Feb. 2006, pp. 1101–1106.

[9] I. n. Iturrate, J. M. Antelis, A. Kübler, and J. Minguez, "A noninvasive brain-actuated wheelchair based on ap300 neurophysiological protocol and automated navigation," *Trans. Rob.*, vol. 25, pp. 614–627, June 2009.

[10] A. Lenhardt and H. Ritter, "An augmented-reality based brain-computer interface for robot control," in *Neural Information Processing. Models and Applications*, ser. Lecture Notes in Computer Science, K. Wong, B. Mendis, and A. Bouzerdoum, Eds. Springer Berlin / Heidelberg, 2010, vol. 6444, pp. 58–65.

[11] C. Neuper and G. Pfurtscheller, "Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates," *International Journal of Psychophysiology*, vol. 43, no. 1, pp. 41–58, 2001, thalamo-Cortical Relationships.

[12] D. Friedman, R. Leeb, G. Pfurtscheller, and M. Slater, "Human-computer interface issues in controlling virtual reality with brain-computer interface," *Human-Computer Interaction (HCI) Journal*, vol. 25, pp. 67–93, 2010.

[13] G. Pfurtscheller, R. Leeb, C. Keinrath, D. Friedman, C. Neuper, C. Guger, and M. Slater, "Walking from thought," *Brain Research*, vol. 1071, pp. 145–52, 2006.

[14] R. Leeb, C. Keinrath, D. Friedman, C. Guger, R. Scherer, C. Neuper, M. Garau, A. Antley, A. Steed, M. Slater, and G. Pfurtscheller, "Walking by thinking: The brainwaves are crucial, not the muscles!" *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 5, pp. 500–514, 2006.

[15] R. Leeb, D. Friedman, G. Muller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) bci control of a wheelchair in virtual environments: A case study with a tetraplegic," *Computational Intelligence and Neuroscience: special issue - Brain-Computer Interfaces: Towards Practical Implementations and Potential Applications*, 2007.

[16] T. Honryak, "Thinking of child's play," *Scientific American*, vol. 295, no. 3, pp. 30–30, 2006.

[17] J. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nature Reviews Neuroscience*, vol. 7, pp. 523–534, July 2006.

[18] Y. Kamitani1 and F. Tong, "Decoding the visual and subjective contents of the human brain," *Nature Neuroscience*, vol. 8, pp. 679–685, 2005.

[19] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fmri data," *Trends in Cognitive Science*, vol. 10, no. 9, pp. 424–430, 2006.

[20] "Turbo brainvoyager - netherlands," http://www.brainvoyager.com.

[21] O. Cohen, A. Mendelsohn, D. Drai, M. Koppel, and R. Malach, "Controlling an avatar by thought using real-time fmri," in *Israel Society For Neurosciences 2012*.

[22] S. Lengagne, N. Ramdani, and P. Fraisse, "Planning and fast replanning safe motions for humanoid robots," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1095 –1106, dec. 2011.

[23] S. Lengagne, A. Kheddar, S. Druon, and E. Yoshida, "Emulating human leg impairments and disabilities in walking with humanoid robots," in *IEEE International Conference on RObotics and BIOmimetics (IEEE-ROBIO)*, 2011.

[24] S. Lengagne, A. Kheddar, and E. Yoshida, "Considering floatting contact and un-modeled effects for multi-contact motion generation," in *Workshop on Humanoid Service Robot Navigation in Crowded and Dynamic Environments at the upcoming IEEE Humanoids Conference*, 2011.