Navigation Assistance for a BCI-controlled Humanoid Robot

Damien Petit, Pierre Gergondet, Andrea Cherubini, Maxime Meilland,
Andrew I. Comport, Abderrahmane Kheddar

▶ **To cite this version:**

HAL Id: lirmm-01222966
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01222966

Submitted on 9 Nov 2015

Navigation Assistance for a BCI-controlled Humanoid Robot

Damien Petit[1,2]   Pierre Gergondet[1,2]   Andrea Cherubini[2]   Abderrahmane Kheddar[1,2]   Maxime Meilland[3]
Andrew I. Comport[3]

Abstract— In this paper, we present a navigation scheme designed to control a humanoid robot via BCI in order to let it interact with the environment and with humans. The interface is based on the well-known steady-state visually evoked potentials (SSVEP) and the stimuli are integrated into the live feedback from the robot embedded camera displayed on a Head Mounted Display (HMD). One user has controlled the HRP-2 humanoid robot in an experiment designed to measure the performance of the new navigation scheme based on visual SLAM feedback. The new navigation scheme performance is tested in an experience where the user is asked to navigate to a certain location in order to perform a task. It results that without the navigation assistance it is much more difficult to navigate to the appropriate pose for performing the task. The detailed results of the experiments are reported in this paper, and we discuss the possible improvements of our novel scheme.

I. INTRODUCTION

Brain-Computer Interfaces (BCI) [1] allow one to bypass the usual physical pathways to control a wide range of systems [2], including humanoid robots [3]. Commands are extracted from the user brain signals interpretation and transmitted to the controlled systems. This new kind of interface allows disabled people to regain some mobility by means of a robotic surrogate. This is called *embodiment* and it is the research topic tackled within the VERE project[1], and the context of the work presented here.

In our recent work [4], we proposed an architecture that allows the user to achieve humanoid whole-body control using an electroencephalography (EEG) non-invasive brain-computer interface. In that work, we instructed the user to grasp an object within the ones in the field-of-view of the robot, then control the robot walk to reach a table where the user finally dropped the object on a selected spot. One challenge encountered in [4] was that steering the robot by BCI is not accurate enough to place the robot in a precise position in space. To overcome this problem, we choose to assist the user during the steering when he needs to place the robot in a specific position to realize a task. Moreover, we wanted closer interaction between robot and human, e.g., through touch or handing over objects.

To realize these goals, that pioneer work was refined over the past year to improve the navigation and interaction

[1]CNRS-AIST Joint Robotics Laboratory (JRL), UMI3218/CRT, Tsukuba, Japan
[2] CNRS-UM2 LIRMM, Interactive Digital Human group, UMR5506, Montpellier, France
[3] CNRS-I3S Laboratory, University of Nice Sophia Antipolis surname@i3s.unice.fr

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

with the environment and human. In this paper, we focus on the integration of a localisation component and on the implementation of a new navigation scheme for BCI control of a humanoid robot. These two components are validated in a demonstration that highlights the new capacities of the system.

The aim of this new demonstration shown in Figure 3 is to navigate the robot towards the user in order to physically interact with him/her. The user wears a Head Mounted Display (HMD) and a EEG cap. The HMD is used to display the live video feed from the robot head cameras in order to improve the user feeling of embodiement. In this kind of interaction it is very important to arrive accurately, without collision at the appropriate position for the robot to interact. It is in this context that the assistive navigation scheme has been developed.

The paper is organized as follows. In the first section, we recall the previous work done in [4], with its main components, and highlight the particular improvements that allow the new demonstration. The following section introduces the SLAM (Simultaneous Localisation and Mapping) system that has been integrated into the demonstration and the control laws that have been devised to use the output as a guideline for the robot walking trajectory. We then present the integration of the SLAM in the BCI user interface to allow the target-oriented navigation introduced in the previous section. Finally, we discuss the results of trial experiments performed with the system and discuss future improvements to the demonstration.

II. BRAIN COMPUTER INTERFACE SYSTEM

In this section, we present the different key elements of a BCI system. The formalism introduced by BCI architecture frameworks such as OpenViBE [5], BCI2000 [6] or TOBI hBCI [7] divides the BCI in three key elements as seen in Figure 1, i.e.: signal acquisition, feature extraction and user application.

A. Signal acquisition

Signal acquisition designates the technology that is used to acquire the signal from the user's brain activity. In this work, we chose to use the popular non-invasive electroencephalography to monitor the user brain activity. Despite poor spatial localization accuracy and low signal to noise ratio, the real-time capacity, cheap cost and non-invasive nature of EEG has made it the most popular technology in BCI.
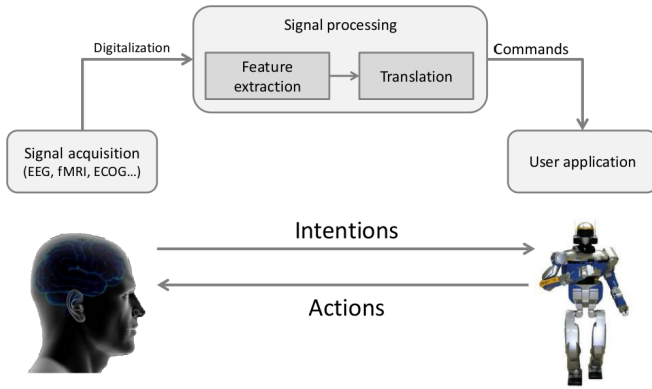
Fig. 1. General design of a BCI system

B. Feature extraction

Feature extraction describes the method used to extract user intentions from the signals acquired in the previous step. We chose to rely on the steady-state visually evoked potentials (SSVEPs). SSVEPs are elicited when the user observes a flickering stimulus. Indeed, we can retrieve the frequency of the stimulus flickering in the spectrum of the brain signal acquired above the visual cortex of the brain, as seen in Figure 2. The method we use to extract the SSVEP from the brain signals is the minimum energy classifier introduced in [8]. It was extended to provide a zero-class implementation, which detects that the user is not attending any stimulus. The classifier is able to operate at high level of accuracy – above 80% – after a short training of about 6 minutes – and a new command can be provided every 200ms [9]. This represents good performance for an SSVEP-based BCI system [10].
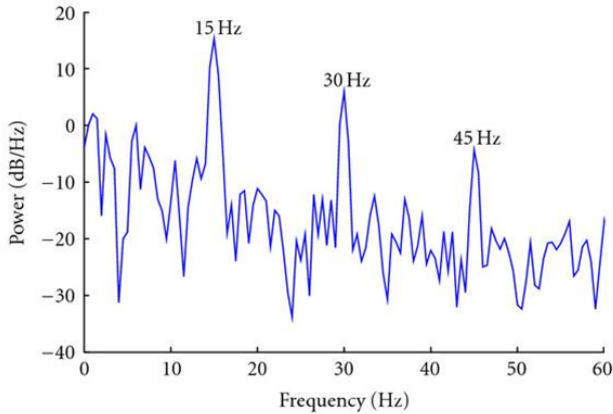

Fig. 2. Spectrum of brain signal when attending a 15Hz visual stimulus

In our previous work, we relied on 4 classes to perform the ensemble of the tasks proposed to the user. However, we trained the classifier to recognize 5 classes in this work to extend the capabilities of the interface. The frequencies of the stimuli were 6, 8, 9, 10 and 14 Hz.

C. User application

User application is the application that is driven by the BCI. Here, we use a humanoid robot, namely the HRP-2 [11] which has one RGB-D camera (Asus Xtion) mounted on its head. To allow task selection we use the aforementioned SSVEP paradigm that we adapted for humanoid control (as we will show in Section IV).

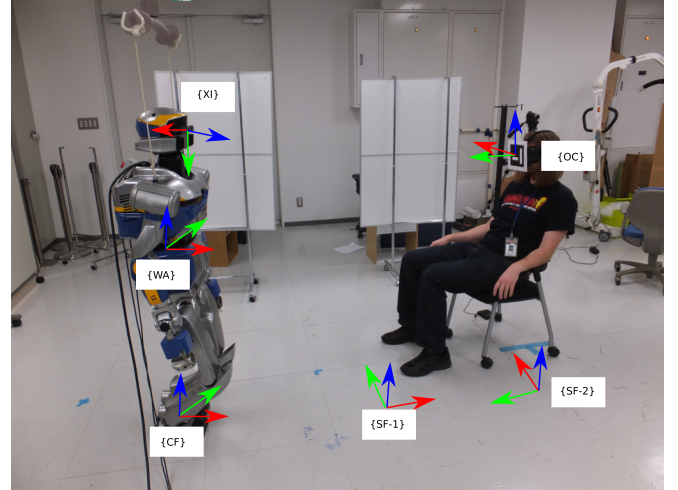## III. OUTLINE OF THE BCI-CONTROLLED SELF-INTERACTION FRAMEWORK

Fig. 3. Demonstration set up

In this section, we introduce the different phases of the demonstration. Further details on the framework are given in the next section.

The demonstration is driven by a finite state machine (FSM) in order to transit between the different stages. Because of the limitations of the BCI, part of the tasks is executed autonomously by the robot. For example, in the human-interaction phase, the only input from the user is the interaction selection; afterwards, the robot autonomously executes the movement. Therefore, the transitions in the FSM are triggered either by the user or by the robot as we can see in Figure 4.
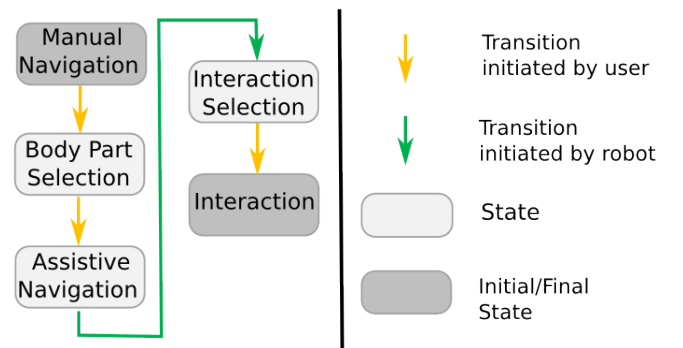

Fig. 4. Finite State Machine for self-interaction

The demonstration is divided in five phases that are reflected in the FSM and outlined below.

1) Manual navigation phase. In general, the robot may not have the user in its field of view. The aim of the first phase is to look for the user by steering the robot: through the feedback from the on-board camera to the HMD, the user can trigger the next phase as soon as he "sees himself through the robot eyes".

2) Body part selection phase. Once the user appears in the robot field of view, he/she interrupts the manual navigation phase to initiate the body part selection phase. The user body parts start to flicker, and the user selects his/her part of the body he wishes the robot to interact with. When one of the body parts is selected, the next phase is initiated.

3) Assistive navigation phase. The robot navigates autonomously in front of the selected user body part, in order to interact with it. The next phase is initiated when the robot reaches its destination.

4) Interaction selection phase. The user selects the task (i.e., type of interaction) to realize on the selected body part.

5) Interaction phase. The robot realizes the selected task.

IV. FRAMEWORK DETAILS

Many adaptations of the original scheme developed in our previous work [4] were required to improve the robot autonomy and enable interaction with a human. For a robot to safely interact with humans, proper BCI intention detection must be guaranteed. To this end, we have modified the task selection as explained in IV-B. To localize the user and its body parts in the scenario, augmented reality markers have been used and tracked using the software presented in [12]. Finally, to enable autonomous navigation to the destination pose, we rely on a SLAM algorithm and on simple velocity control scheme to drive the robot steps. All these changes are detailed in the rest of this section, after a brief part dedicated to the definition of the reference frames used in this work.

A. Reference Frame Definitions

The reference frames used in this work are shown in Figure 3. These are, the RGBD optical camera frame XI, the center of mass of the robot WA, the HMD frame OC, the predefined frame used to interact with the user SF-2, a waypoint to avoid collision with the user SF-1. SF-1 and SF-2 are part of a frame sequence that constitute the path of the robot. SF-i represents the current frame of that sequence, so SF-i is either SF-1 or SF-2. In this work, 3D points are represented in upper-case, using the homogeneous representation. Coordinate frames are specified in superscript, such as $^A\mathbf{P}$, and the homogeneous transformation matrix $^B\mathbf{T}_A$ transforms points from frame A to B. The transformation $^B\mathbf{T}_A$ is characterized by translation $^B\mathbf{P}_A = (^B X_A, {}^B Y_A, {}^B Z_A)$ and rotation $^B\mathbf{R}_A$.

B. Improved BCI task Selection for human-robot interaction

We rely on the SSVEP paradigm to perform humanoid control. However, this paradigm, like any other paradigm used in BCI, suffers from a major drawback from a control perspective, i.e., possible interpretation errors (false positives). Errors are prone to happen during a session: the user might get distracted or introduce noise in the EEG data by moving involuntarily and thus provoking a false positive. While this was not a major issue in our previous work, where the BCI was used for object-oriented tasks (e.g., grasping), it becomes crucial in this application, where the robot must interact safely and precisely with a human being. To reduce interpretation errors, we devised a solution that we call "enforced selection" and that is presented hereby.

The SSVEP classification algorithm delivers a new classification every 200ms. In the enforced selection paradigm, the user has to sustain his/her intention for a "given" time before it is validated by the system. This "given" time can be tuned to fit the user performance: for example, for an accustomed user only one second of sustained attention is enough, while a novice user may require three seconds.

This "enforced" selection is used when a misclassification of the user intention would trigger an action that cannot be recovered by the robot: for example, pouring water into a glass. However, it is not used when a wrong selection has little impact; for example, when steering a robot freely, giving a wrong speed input for 200 ms has little impact on the movements of the robot.

C. Manual navigation

In the first phase of the FSM, the user must manually control the robot until he is visible in the robot field of view. The interface shown to the user during this phase is fairly simple, and similar to the one introduced in [4]:

- three arrows are flickering in the robot field of view: one is used to move the robot in the forward direction while the other two are used to turn the robot left or right;
- a fourth flickering stimulus is used to stop the robot using the enforced SSVEP selection mechanism described earlier.

D. Body part selection phase

The manual navigation phase is stopped as soon as the user appears in the robot field of view. Since the user is equipped with a HMD, we have decided to customize it to ease the person detection by the robot. Hence, we have placed an Augmented Reality (AR)-code marker [12] on the HMD worn by the user, as shown in Fig. 3. As soon as the Aruco tracker detects and localizes the marker (i.e., as soon as the pose of frame OC in camera frame XI, $^{XI}\mathbf{T}_{OC}$, is known), the robot stops and the body part selection phase begins. To this end, the user body parts are projected in the image and start to flicker, to provide a new set of SSVEP stimuli on the user screen. For this, the poses of the body parts (for the moment, we use only the user arms) are assumed at a fixed, known pose, in the OC frame. In future work, we plan to

develop a human body tracking algorithm to generalize the approach to the different subjects biometrics.

The interface is initially similar to that in [4], with the addition of the navigation assistance. However, as the user is recognized within the field of view of the robot, the user arms will be displayed in the interface as seen in Figure 6.

E. Assistive navigation phase

As soon as the user has selected a body part, by triggering the corresponding SSVEP response, assistive navigation starts. To realize this, the pose of the robot with respect to the user must be updated in real-time and fed back to a controller for navigating to the desired body part. We hereby detail these two modules.

1) Localization: The real-time SLAM system [13] we adopted provides the 3D pose of the robot camera frame XI in the HMD frame OC, i.e., $^{XI}\mathbf{T}_{OC}$.

2) Assistive navigation: When navigation assistance is triggered, the control of the robot steering is centered around the target frame SF-i with i $\in$ (1..2) in the sequence associated to the selected arm: SF-1 then SF-2, hence only two arrows in the interface are used:

- The "up" arrow is not used to move the robot *forward* but to move the robot *towards* the current frame in the sequence;
- The "stop" button is used to switch back to the global frame.

The SF-1 frame is used as a predefined waypoint ensuring that the robot will not enter in collision with the user during the assistive navigation phase. A simple control scheme (S_1) is adopted to control the robot toward the current frame in the sequence. Once the final position is reached, the robot is realigned on SF-i before going further in the sequence.

$$(S_1) \begin{cases} \alpha_{(t)} = \arctan\left(^{WA}Y_{CF}, ^{WA}X_{CF}\right) \\ v_{x(t)} = V \\ v_{y(t)} = 0 \\ w_{z(t)} = K * \alpha_{(t)} \end{cases}$$

With V and K constants tuned for the robot and $v_{x(t)}$, $v_{y(t)}$, $w_{z(t)}$, the desired speed in WA sent to the humanoid robot pattern generator [14].
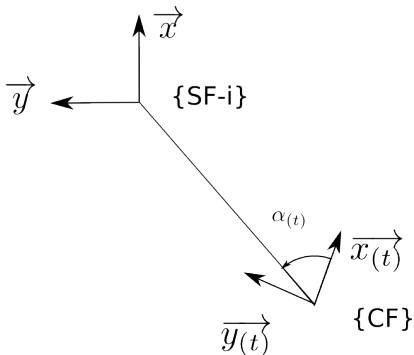
Fig. 5.   Navigation assistance

F. Interaction selection phase

When the robot reaches its destination, the user can trigger the touch task by focusing his intention on a SSVEP stimulus located over his arm. Since the localisation of the arm must be accurate, to realize the task another AR marker is placed on the user arms.
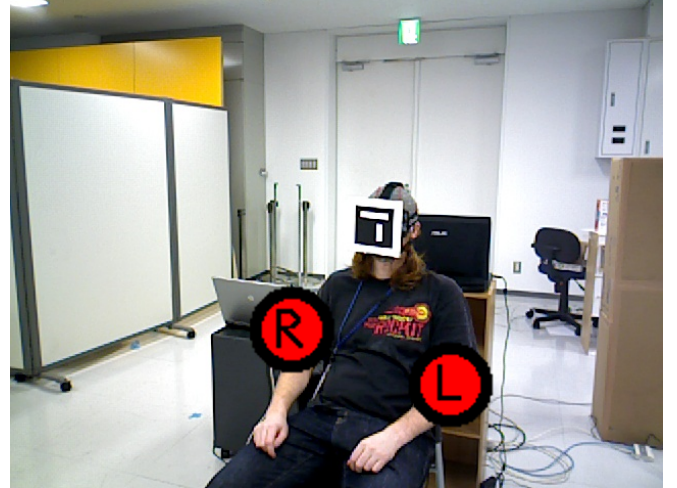
Fig. 6.   SSVEP arms in interface

G. Interaction phase

The interaction phase begin by adjusting the robot pose with small steps in order to perform the touch task. This pose adjustment is made with a different pattern generator than the one used previously to navigate the robot. This patter generator allows to control the poses of the robot feet. After the adjustment, the robot arm start his motion to touch the user arm.

V. EXPERIMENT

In this section we introduce the experiment we deviced to assess the new demonstration. Videos of the experiment are shown here:
https://www.dropbox.com/sh/
2zml5n89hlalbov/SkMmdL0osU

A. Subjects & Material

One healthy volunteer participated in the experiments, One males, age 24±1.4. The subject had no experience of BCI usage.

We use a g.USBamp (24 Bit biosignal amplification unit, g.tec Medical Engineering GmbH, Austria) to acquire the EEG data from the user's brain at a sampling frequency of 256Hz, bandpass filtered between 0.5 and 60Hz with a notch filter at 50Hz to get rid of the power line noise. We use 8 Ag/AgCl active electrodes. The electrodes are placed on the POz, PO3, PO4, PO7, PO8, O1, O2 and Oz positions of the international 10-20 system [15], Fpz is used as the ground electrode and the earlobe as a reference.

The frequencies of the stimuli are 6, 8, 9, 10 and 14Hz. Those were carefully selected to avoid common first or

second harmonics, while staying below 20Hz, to minimize the risk of eliciting an epileptic crisis in healthy subjects as advised in [16].

The experiments are conducted with a HRP-2 humanoid robot. During the experiment, the user is comfortably seated in an armchair, and wears an HMD (Oculus Rift) upon which the user interface is displayed at a refresh rate of 60Hz.

B. Experiment Protocol

The user is asked to steer the robot near one of his arms in order to touch it. This task is realized two times, one time using only the manual navigation and one time using the assistive navigation scheme.

VI. Results

Without assistive navigation the user steered the robot near his arm with an increasing difficulty when the robot get close to him. This can be explained by the distraction caused by the robot going near the user and the fear of not being able to stop the robot before collision. Both the distraction and fear causes the performance loss of the BCI system. Hence leads to a difficulty to steer the robot.

The final position of the robot was too close to the user. The marker placed on the user arm could not be detected. So the final pose of the robot did not permit to perform the interaction phase.

Using the assistive navigation the user first steered the robot to look for the marker placed on the HMD then stop the robot in order to initialize the localisation. This phase is unnecessary when using manual steering only. Once the localisation is initialized the user only needed to focus on one stimuli to go near his selected arm. The BCI performance were less affected by the fact that the robot went near him knowing that he was free from collision using the assitive navigation. The final pose of the robot in this case was accurate enough to allow the detection of the marker on the user arm allowing the interaction phase.

VII. Perspectives and Conclusion

The experience proved that due to the difficulty of an accurate manual navigation, the navigation assistance is necessary when we need to perform an interaction with the environment.

However, an issue that come with using an assistive navigation scheme is the decrease of the embodiement feeling caused by forcing the user to take the computed direction to go to the chosen interaction area. A solution to this issue could be to update the assistive control to offer the same kind of motion than the manual navigation but guiding the user and not constrain him. This could be done by dynamiclly adjust the gains on each proposed motion to guide the user to the computed direction and not forced him to this direction. This assistive control still needs to ensure a collision free navigation in order to not disturbed the performance of the BCI system as shown in the experiment. With this we could ensure to arrived at the interaction area safely without breaking the embodiement.

Other improvements could also be made to offers the possibility to detect potential interaction area without marker and stopping the robot walk. Using the map built in real time by the SLAM system we could detect objects during walking phase. An other way to detect such potential interaction area could be to detect the plane in the map or camera field of view. Indeed day to day interactable objects such as desk, table, fridge, drawer, wall, door, cubicle are all made of planes. Once one of those planes is detected we could use the same assitive navigation scheme to navigate toward the detected interaction area.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, pp. 767–791, 2002.

[2] J. del R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia, "Combining Brain-Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges," *Frontiers in Neuroprosthetics*, vol. 4, p. 161, 2010.

[3] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. N. Ra, "Control of a humanoid robot by a noninvasive brain-computer interface in human," *Journal of Neural Engineering*, vol. 5, pp. 214–220, 2008.

[4] P. Gergondet, A. Kheddar, C. Hintermüller, C. Guger, and M. Slater, "Multitask Humanoid Control with a Brain-Computer Interface: user experiment with HRP-2," *International Symposium on Experimental Robotics (ISER)*, 2012.

[5] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer, "OpenViBE: An Open-Source Software Platform to Design, Test and Use Brain-Computer Interfaces in Real and Virtual Environments," *Presence Teleoperators & Virtual Environments*, vol. 19, no. 1, pp. 35–53, 2010.

[6] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering*, vol. 51, p. 1034, 2004.

[7] G. R. Müller-Putz, C. Breitwieser, F. Cincotti, R. Leeb, M. Schreuder, F. Leotta, M. Tavella, L. Bianchi, A. Kreilinger, A. Ramsay, M. Rohm, M. Sagebaum, L. Tonin, C. Neuper, and J. del. R. Millán, "Tools for brain-computer interaction: a general concept for a hybrid BCI," *Frontiers in Neuroinformatics*, vol. 5, 2011.

[8] O. Friman, I. Volosyak, and A. Gräser, "Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 742–750, April 2007.

[9] R. Prueckl and C. Guger, "A brain-computer interface based on steady state visual evoked potentials for controlling a robot," *Lecture Notes in Computer Science*, vol. 5517, pp. 690–697, 2009.

[10] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichockia, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Progress in Neurobiology*, vol. 90, pp. 418–438, 2010.

[11] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi, "Humanoid robot HRP-2," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, no. April, pp. 1083–1090 Vol.2, IEEE, 2004.

[12] S. Garrido-Jurado, R. Muñoz Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, 2014.

[13] M. Meilland and A. I. Comport, "On unifying key-frame and voxel-based dense visual SLAM at large scales," in *International Conference on Intelligent Robots and Systems*, (Tokyo, Japan), IEEE/RSJ, 2013.

[14] A. Herdt, H. Diedam, and P. Wieber, "Online walking motion generation with automatic footstep placement," *Advanced . . .* , vol. 6, pp. 5–6, 2010.

[15] F. Sharbrough, G.-E. Chatrian, R. Lesser, H. Lüders, M. Nuwer, and T. Picton, "American Electroencephalographic Society guidelines for standard electrode position nomenclature," *Journal of Clinical Neurophysiology*, vol. 8, pp. 200–202, 1991.

[16] R. S. Fisher, G. Harding, G. Erba, G. L. Barkley, and A. Wilkins, "Photic- and pattern-induced seizures: a Review for the Epilepsy Foundation of America Working Group," *Epilepsia*, vol. 46, pp. 1426–1441, 2005.