A hybrid EEG-based BCI for robot grasp controlling*

Wenchang Zhang, Fuchun Sun, Senior member, IEEE, Chunfang Liu, Weihua Su, Chuanqi Tan, Shaobo Liu

Abstract—Brain–Computer Interfaces (BCI) can help disable people to improve human – environment interaction and rehabilitation. Grasping objects with EEG-based BCI has become a popular and hard research in recent years due to the high degree of freedom robot and complex grasp planning. Unlike commonly used paradigms, we propose a pipeline of hybrid EEG-based BCI for robot grasping by shared control to solve the key problems including target object selection, robot intelligent planning and shared control by both user intention and robot. Six experimental users could successfully use the system to grasp a number of objects in a test scene. The results of grasping experiment demonstrate that our method achieve an effective performance.

I. INTRODUCTION

Brain–computer interface (BCI), which can directly communicate brain activities and computer control signals, has been more and more applied to help people who has restricted mobility such as stroke, extremity disability and so on, for some activities of daily living and rehabilitation. With the development of robotic arm and hand, these persons can interact with their environment including manipulating cup to drink water, opening door and any other actions by equipped BCI systems. Furthermore, rehabilitation training by motor imagery (MI) based BCI and exoskeleton robot brings more benefit for stroke patients’ recovery of sensory-motor functions [1][2]. However, robot grasp controlling by BCI is still a hard problem because of high degree of freedom robot arm and hand.

In the EEG-based BCI field, most of related studies were focused on 1-D cursor control, which was generally implemented through detecting and classifying the changes of mu (8–12 Hz) or beta (13–28 Hz) rhythm during different motor imagery tasks, e.g., imagination of left- and right-hand movement [3]–[7]. In the past few years, other forms of 2-D BCI were also reported that they could be adopted a discrete control paradigm by using the steady-state visual evoked potential (SSVEP) [8]–[10]. Faller et al. [11] proposed a configurable application framework seamlessly integrate SSVEP stimuli within a desktop-based virtual environment. Furthermore, the concept of a hybrid BCI combining P300 and SSVEP was first proposed in [12]. This BCI detected at least two brain patterns in a simultaneous [13]–[18] or sequential manner [19]–[21]. Hybrid BCIs have become a hot research topic in recent years.

Allison et al. [22]–[24] demonstrated that the BCI could improve accuracy by combining multiple brain signals such as motor imagery and SSVEP, especially for users with poor performance.

In the past two years, EEG-based BCI for three-dimensional grasp controlling was paid more attention. Y. Li et al. [25] proposed a real-time BCI system that based on combining two brain signals including Mu/Beta rhythm during motor imagery and P300 potential to carry on an online experiment involving six subjects performing 3-D control tasks. Jianjun Meng et al. [26] used non-invasive BCI to control a robotic arm with high accuracy for performing tasks requiring multiple degrees of freedom by decomposing three-dimensional space into two sequential low dimensional virtual cursor controls. Instead of controlling robot hand movement step by step directly through BCI, Robert Ying et. al. [27] presented an online grasp planning framework by using an advanced EEG-based interface paradigm called Rapid Serial Visual Presentation (RSVP). This system was trained to recognize the “interest” signal for each object and calculates the position coordinate by robot vision. Although these researches achieved some outcome in some aspects such as object selection and robot movement controlling, there is still a great potential for the performance of robot grasp controlling by BCI.

In this work, we aim to illustrate the advantages of hybrid BCI and shared control for the task of grasping object by robot hand and BCI. Thus, we present a hybrid EEG-based BCI system for robot grasp controlling, which has many advantages. Firstly, we recognize objects by computer vision from Kinect, and select target object by SSVEP. It is very simple and performs high accuracy. Secondly, our shared-control paradigm makes grasp task of robot hand more safe and robust. The interference between hand and objects can be avoided. Finally, users can utilize this system with minimal training and achieve grasp task with high success rate.

The rest of the paper is organized as follows. The pipeline of EEG-based BCI for grasping is illuminated in Section II, which includes object selection by SSVEP, intelligent grasp planning, and shared control paradigm. Section III displays experiments configuration and provides the results of SSVEP recognition rate, MI recognition rate and grasp performance. Finally, we give a conclusion in Section IV.

II. PIPELINE OF EEG-BASED BCI FOR GRASPING

There are three progressing states when attempting a grasp by EEG-based BCI: object selection, grasp planning, and shared control. The pathway is illustrated in Figure 1. The first step is object recognition and selection. The objects in scene can be recognized by computer vision from Kinect including color,
shape and deep information. We select the target object that we want to grasp by SSVEP paradigm from EEG-based BCI. Then, the following step is shared control for robot hand grasping, which is combined by grasp intelligent planning and BCI control. The robot hand can run intelligent planning control based on point cloud from Kinect by optimization algorithms and some constraints. If that intelligent planning is not satisfied by the user, he generates intent control. The EEG command is immediately detected and BCI control step is triggered. Then, the user can control the direction of hand movement by MI and achieve grasp execution by electrooculography (EOG). After EEG option selection and shared control, the robot hand can grasp the target object. If some step fails, the program jumps to the beginning.

Figure 1. Pipeline of the EEG-based BCI for Grasping

A. Object selection by EEG-based BCI

In this stage, Kinect fixed just above the object is utilized for recognition by retrieving models from a database that fit the scene. After image representing selection of each object is generated by Kinect vision recognition, the user is instructed to just look for the object that they want to grasp by EEG. Then the target object highlights green in potential selection. Concretely it can be selected based on SSVEP. The subject focuses on the virtual LED image above each object image that oscillated at given frequency such as 5Hz, 7Hz, 11Hz, 13Hz. (Figure 2)

Figure 2. Object selection by SSVEP. There are four objects and four different frequencies flashing LED above them. When we focus on the target object 2, the algorithm detects 7Hz frequency and object 2 highlights green.

The SSVEP frequency recognition is based on canonical correlation analysis (CCA) that can capture the interrelationship between predictor and response variables. Considering these two variables $X, Y$ and their linear combinations $x = X^T w_x$ and $y = Y^T w_y$, CCA finds the weight vectors, $w_x$ and $w_y$, which maximize the correlation between $x$ and $y$.

$$
\rho(x, y) = \max_{w_x, w_y} \frac{E[w_x^T XX^T w_x]}{\sqrt{E[w_x^T XX^T w_x]E[w_y^T YY^T w_y]}}
$$

where $\rho(x, y)$ is the maximum canonical correlation.

We calculate CCA coefficients from multiple channels of EEG signals sampling as a set of all stimulus frequencies, and associate harmonics (second and third harmonics) as another set of variables. Therefore, the reference signals $Y_n$ can be represented as

$$
Y_n = \begin{bmatrix}
\sin(2\pi f_n t) \\
\cos(2\pi f_n t) \\
\sin(4\pi f_n t) \\
\cos(4\pi f_n t) \\
\sin(6\pi f_n t) \\
\cos(6\pi f_n t)
\end{bmatrix},
$$

where $f_n$ denotes the stimulation frequency and $f_s$ is the sampling rate. The frequency with the largest CCA coefficient is the stimulus frequency of the recorded SSVEP.

$$
\hat{f} = \arg\max_{f_n} \rho_{mn}, \quad m = 1,2,3.
$$

where $\hat{f}$ is the SSVEP frequency estimation.

B. Grasp planning

A stable and successful grasp of robot is determined by many properties, such as form-closure, force-closure, sufficient contact point and distance between the center of the object and the grasping point. However, these properties are difficult to calculate exactly. Therefore, optimization algorithms are usually applied to solve the grasp planning problem.

Kinect can recognize and generate point cloud data for target object and fix its position. For three-finger robot hand configuration, it is better that more enclosed points should be inside the volume enclosed by the fingers. In order to grasp object robustly, we capture a lot of features and constraints including the number of points located inside the volume enclosed within the finger-tips, the number of points contained in spheres of different sizes located at the hand’s center and the palm of the hand (Figure 3). These constraints are dependent on the task, attributes, experience, and robot hand. Task and attributes constraints enforce the search for a suitable grasp that matches constraints for the task. Experience and robot hand constraints seeks for a suitable grasp that matches the local shapes grasped before. For example, objects often share local shape, we can grasp similar shapes on other objects. Once the robot hand finds a good candidate grasp position and satisfy those constraints maximization, it begins to execute the grasp.

Figure 3. Grasp planning by point cloud simulation. In order to grasp a bottle, robot hand fingers should enclose more points to generate form-closure and force-closure for a stable grasp.
C. Shared control

The system is integrated by the intelligent robot hand and the hybrid EEG-based BCI. We apply shared control strategy [28] to control robot hand for grasping object. In order to grasp stably and successfully, the robot hand stops automatically once it is detected to fail connection with the BCI. Among the shared control, BCI system play the most major role. Therefore, when the subject perform EEG commands, the intelligent control system will decrease its role. If the shared control system has not detected BCI commands, the robot hand achieves assistance task consistently and take over control for grasp planning. Figure 4 depicts an architecture of the shared control for grasping based on BCI.

\[
\text{min} \quad \|D\|_2^* + \alpha \|E\|_1 \quad \text{s.t.} \quad O = D + E \quad (11)
\]

- We determine the orientation of the robot movement by calculating the maximum probability from P(C).

a. Control of horizontal and vertical movement based on motor imagery

In order to control robot hand movement by BCI, we applied event related desynchronisation (ERD) detection algorithm based on low-rank linear dynamical systems (LR-LDS) for motor imagery pattern recognition [29][30]. This algorithm can classify multi-pattern recognition without complex signal pre-processing and post-processing such as band-pass filtering, artifacts removing and classifier choosing.

Firstly, we apply linear dynamical systems (LDSs) to extract the spatio-spectral feature of EEG signals. Let \( \{Y(t)\}_{t=1}^n \), \( Y(t) \in R^m \) be a sequence of \( t \) EEG signal sample at each instant of time \( t \). We have \( x(t) = \sum_{i=1}^L A_i x(t-i) + B v(t) \) with \( A_i \in R^{m \times m}, B \in R^{m \times n_v} \), \( Y(t) \) can be represented approximately by function of dimensional hidden state \( x(t) \), \( y(t) = \varphi(x(t)) + \omega(t) \), we redefine the hidden state of \( x(t) \) to be \( [x(t)^T \ x(t-1)^T \ldots \ x(t-k)^T]^T \) and consider a linear dynamic system as an auto-regressive moving average process without firm input distribution.

\[
\begin{align*}
(x(t + 1) & = Ax(t) + B v(t) \quad v(t) \sim N(0, Q) \\
(y(t) & = C x(t) + \omega(t) + \bar{y} \quad \omega(t) \sim N(0, R), \quad x(0) = x_0)
\end{align*}
\]

(7)

where \( A \in R^{m \times m} \) is the transition matrix that describes the dynamics property, \( C \in R^{m \times n_v} \) is the measurement matrix that describes the spatial appearance, \( \bar{y} \in R^m \) is the mean of \( y(t) \), \( v(t) \) and \( \omega(t) \) are noise components.

So let \( Y = U \Sigma V^T \), we get the parameter estimate of \( \hat{C}, \hat{X} \)

\[
\begin{cases}
\hat{C} = U \\
\hat{X} = \Sigma V^T
\end{cases}
\]

(8)

where \( \hat{X} = [X(1), X(2), \ldots, X(\tau)] \). \( \hat{A} \) can be determined by Frobenius:

\[
\hat{A}(\tau) = \arg \min_A \|X_2(\tau) - AX_1(\tau-1)\|_F
\]

(9)

where \( X_2(\tau) = [X(2), X(3), \ldots, X(\tau)] \). So the solution is in closed-form using the state estimated:

\[
\hat{A}(\tau) = [X(:,2)X(:,3) \ldots X(:,\tau)] \ast [X(:,1)X(:,2) \ldots X(:,\tau-1)]^\dagger
\]

(10)

where \( \dagger \) denotes matrix pseudoinverse.

We can obtain the result \([A,C]\), a couple of spatio-temporal feature matrix.

Then, we approximate the extended observability by taking the L-order observability matrix, i.e. \( O(n,L) = [C^T, (CA)^T, \ldots, (CA^{L-1})^T]^T \). In this way, an LDS model can be alternately identified as an n-dimensional subspace of \( R^{lm} \).

Based on low rank and sparse matrix decomposition, observability matrix \( O \) can be decomposed into \( D+E \) as following formulation:

\[
\min_{D,E} \|D\|_1 + \alpha \|E\|_1 \quad \text{s.t.} \quad O = D + E
\]

(11)
where D is a low-rank matrix and E is the associated sparse error. The Inexact ALM Method can be also used to solve this optimization problem. The output D represents low-rank descriptor for LDSs and can be employed for classification of EEG trails.

Finally, Mahalanobis distance can describe the distance between two feature spaces of EEG trails of LR-LDS. We can classify EEG signals by comparing Mahalanobis Distance between training data and testing data. Nearest two samples mean that they may be same class. Finally, the MI-EEG pattern can be recognized by the classification.

In this work, the user can image four different motors including right hand movement, left hand movement, tongue movement and feet movement to indicate four corresponding directions right, left, up and down (Fig 5). In order to increase the classification accuracy, six subjects are ranged for training. Each of them is trained for four sessions and all the training sessions for a subject are arranged in several consecutive weeks. In addition, we use deep information from Kinect to calculate the depth of the robot hand movement for grasping.

**b. Control of grasp execution and stop based on EOG**

We apply EOG to give command of grasp execution and stop. EOG is a technique for measuring the potential of eye by the electrodes placed around the eye. When the eye-gaze changes or eye blinks, the voltages of eye potential cause variations. For eye movement events, such as blink and saccade, they can be recognized by pattern recognition algorithms. This paradigm assumes single strong eye blinking events for robot hand closure and double eye blinking events for open. The lowest and highest absolute values of eye potential are rejected from recorded parameters for each event separately. A strong blink has an amplitude higher than any other eye activity. For double blink events are identified two sets of values, one set for each blink. A double blink event is chosen because this event is executed only when the user has the intention to execute it.

**II. EXPERIMENT**

**A. EEG-based BCI Performance**

We utilized Neruoscan EEG system with 32-channel electric potentials from the scalp of the user wearing an EEG cap and 4-channel electrooculography representing eye movements.

SSVEP detection for object selection was performed every 200 ms. We chose EEG signals from eight channels ("P7," "P3," "Pz," “P4,” “P8,” “O1,” “Oz,” and “O2”) and filtered signals within the range of 3–20 Hz by Butterworth filter. Then we extracted 1.6 s period (400 data points) for CCA. Six subjects performed the object selection experiment by SSVEP and the accuracies are shown in Table I.

**TABLE I. RESULTS IN SSVEP DETECTION**

<table>
<thead>
<tr>
<th>Subjects</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trails</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>90</td>
<td>87.5</td>
<td>92.5</td>
<td>90</td>
<td>92.5</td>
<td>90</td>
</tr>
</tbody>
</table>

MI-based EEG signals were used to control the four directions (right, left, up, down) of robot hand movement. We applied low-rank linear dynamical systems algorithm to recognize the four MI pattern (right hand movement, left hand movement, tongue movement, feet movement). There are two steps: training paradigm and test paradigm. The aim of training paradigm was to learn the different features of subjects’ EEG signals. We labeled the classifications of different trails from subjects and built a training dataset. Figure 6 depicts the training paradigm.

**TABLE II. RESULTS IN THE MI TEST (ACCURACY %)**

<table>
<thead>
<tr>
<th>Subjects</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trails of each MI</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Right accuracy</td>
<td>70</td>
<td>65</td>
<td>75</td>
<td>70</td>
<td>75</td>
<td>65</td>
</tr>
<tr>
<td>Left accuracy</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>75</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Up accuracy</td>
<td>50</td>
<td>60</td>
<td>70</td>
<td>60</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>Down accuracy</td>
<td>60</td>
<td>70</td>
<td>65</td>
<td>60</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>Average accuracy</td>
<td>61.25</td>
<td>65</td>
<td>68.75</td>
<td>66.25</td>
<td>68.75</td>
<td>68.75</td>
</tr>
</tbody>
</table>

**B. Control robot hand to grasp object**

The experiment platform was based on Barrett 3-finger hand and SCHUNK 7-joint arm (Figure 7). Our goal was to evaluate the contribution of shared control for grasping object.
In the applied shared control paradigm, the robot performed grasp planning intelligently by computer vision and experience at low-level. The user asynchronously sent high-level commands by BCI. The experiment was run under three conditions: BCI control with shared control (BWS), BCI control without shared control (BOS), and robot intelligent control without shared control (ROS). We chose four objects (a bottle of water, a cup, a beer can, a box of milk) and placed them into 5 arbitrarily scenes within the field of view of the Kinect camera. The whole test time for grasp task was 240s. We considered the task was failed if it was timeout. Table III presented the results of grasp success rate.

We found that grasp task was easy to fail when the arm and hand interfered to other object in the case of ROS. On the other hand, it cost too much time by BOS and the user was very tired in this whole task. By the compared experiment, grasp task by BWS performed the best because its grasp success rate was the highest and the average times was short.

### Table III. Result of Grasp Success Rate

<table>
<thead>
<tr>
<th>Subjects</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenes for condition</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Failure number</td>
<td>BWS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>BOS</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>ROS</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Success rate (%)</td>
<td>BWS</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>BOS</td>
<td>80</td>
<td>100</td>
<td>80</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>ROS</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>Average Times (s) of success</td>
<td>BWS</td>
<td>128</td>
<td>123</td>
<td>132</td>
<td>126</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>BOS</td>
<td>195</td>
<td>188</td>
<td>203</td>
<td>210</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>ROS</td>
<td>107</td>
<td>102</td>
<td>123</td>
<td>116</td>
<td>136</td>
</tr>
</tbody>
</table>

### III. Conclusion

In this paper, we propose a pipeline of a hybrid EEG-based BCI for grasping object by shared control. The results of experiment demonstrate that the hybrid EEG-based BCI with shared control can achieve selection and grasping task efficiently. Grasping object problems with multi-degree of freedom and complex grasp planning are solved effectively.

During the experiment, we found that SSVEP performed higher accuracy than MI and the 4-class MI pattern recognition had a low success rate. We can increase the number of training subjects to boost accuracy in our low-rank linear dynamical systems algorithm.

In the current implementation, the subjects should make decisions at fixed pipeline. We attempt to integrate the EEG data in a more real-time strategy in the future work to make the user in the control loop online.

### Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant Nos. 91420302 and 91520201) and Innovation Cultivating Fund Project17-163-12-ZT-001-019-01.

### References


3282


[26] Yanjun Meng, Shuying Zhang, Angeliki Bekyo, Jaron Olsoe, Bryan Baxter, & Bin He1, Noninvasive Electroencephalogram Based Control of a Robotic Arm for Reach and Grasp Tasks, Scientific Reports 6, Article number: 38565 (2016), http://dx.doi.org/10.1038/srep38565


