# Grasping with your brain: a brain-computer interface for fast grasp selection

Robert Ying and Jonathan Weisz and Peter K. Allen

***Abstract*— We present a shared control online grasp planning framework that encorporates an advanced EEG-based interface. Unlike commonly used paradigms, the EEG interface we incorporate allows online generation of a flexible number of options. This online planning framework allows the user to direct the planner towards grasps that reflect their intent for using the grasped object. This paradigm is integrated with a full grasping paradigm, and has demonstrated robust control in our investigations thus far.**

## I. INTRODUCTION

In previous work [8], [1], [2], we have presented a series of refinements of a user interface for enabled a shared-control online grasp planner to function using a variety of input devices and paradigms. In this work, we further extend that UI to an EEG paradigm that allows an, potentially large number of slightly ambiguous options to be presented to the user. The user acts as a filter for the planner - directing the planner to a desired approach direction and filtering proposed candidates until a reasonable one is found. We have found this paradigm to be successful in somewhat cluttered, natural scenes using a wide variety of interfaces.

## II. METHODS

### A. One-of-many selection

The key development in this system is a reliable event-related potential (ERP) detector which can be used to provide a one-of-many selection between various options, combined with a closed-loop online grasp planner that takes advantage of visual ambiguity to quickly converge to correct, reachable grasps.

To be precise, we have created a process by which a service (ROS, GraspIt!, or otherwise) can query a user to select a specific option out of a set of between one and sixteen available options, or no option at all. The user only needs to pay attention to a stream of images while wearing an electroencephalographic (EEG) headset.

*1) EEG input:* Our current implementation uses a B-Alert X10 EEG system from Advanced Brain Monitoring (Carlsbad, CA), which provides 9 electrodes positioned according to the 10-20 system and a pair of reference channels. The EEG data is acquired at 256 Hz, with 60 Hz notch filter and 0.5 Hz high-pass filters applied before any additional processing. The EEG interest metric is based on that described in [3], [4], [5], with some additional normalization and post-processing.

*2) Choosing options:* When the EEG system is sent a list of image options from the grasping pipeline, it first presents them to the user in a grid for between three and five seconds. During this time, the user is asked to determine which of the available options (if any) he or she would like to select. The user is instructed to mentally count the number of times that particular option appears.

The system then presents the options in a Rapid Serial Visual Presentation (RSVP) paradigm [6]. This is used instead of the more common Farwell-Donchin P300 speller paradigm [7] because it allows the system to take advantage of visual ambiguity to converge on final grasp selections more quickly, while still allowing for an approximation of the P300 speller if so requested by the grasp planner [1].

To generate the RSVP sequence, the system randomly selects each option to appear between three and seven times. The sequence is then randomly ordered, with the constraint that the same option does not appear in two consecutive image presentations. This method has, in experimental data, been sufficient to trigger the "oddball" response that is necessary for the P300 signal. If there are less than five options, the system will automatically fill in distractor image options to make this constraint more feasible.

The images are each presented at 4 Hz, and preliminary EEG scores $e_i$ are assigned. We then aggregate each of the $n$, images by their option, and determine whether or not the user has made a selection.

To test if the user has consciously selected any of the images, we sort the images by their EEG scores, and then split it into a group of size $x$ and $n-x$. We vary $x$ so as to maximize the change in the average measured EEG score:

$$x^* = \arg\max_{x\in[1,n]} \left( \frac{1}{n}\sum_{i=1}^{n} e_i - \frac{1}{n-x}\sum_{i=x+1}^{n} e_i \right) \qquad (1)$$

If $x^* > \max(0.1n, 5)$, we determine that the user had not made a choice. In practice, this is a highly reliable means of checking whether the user was paying attention and attempting to make a selection.

If $x^* \leq \max(0.1n, 5)$, we use each of the images which are sorted in the top $x^*$ positions as a vote; the option with the most votes is selected. If there are options with an equal number of votes, the tie is broken by the option with a better average EEG score.

*This work was not supported by any organization

[1]Robert Ying, Jonathan Weisz, and Peter K. Allen are with the Computer Science Department, Columbia University, New York, New York. robert.ying@columbia.edu, {jweisz, allen}@cs.columbia.edu

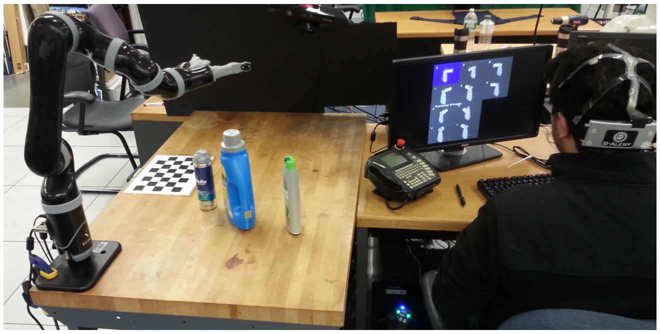

Fig. 1. The subject guiding the system through the grasp selection stage. In this example, none of the grasps found in the database for the object were reachable. The user will need to select a good starting point, and allow the automated planner to generate new candidates.

### B. Grasping Paradigm

We extend the grasping paradigm established in [8] to take advantage of the one-of-many RSVP-based selection system described above.

There are four states that the user progresses through when attempting to formulate a grasp:

*1) Object selection state:* The object recognition pipeline is initialized upon entering the object selection state, using a Microsoft Kinect to provide the RGBD data for the recognition [9]. Each of the objects in the scene is then highlighted, with the camera held to a fixed point and perspective, such that all of the objects in the scene are visible. Between the various images only the highlighted object changes; the subject need only focus on the object that he or she would like to manipulate for the system to be effective, similar to the P300 speller paradigm. This allows the user to select the correct object with high confidence.

*2) Grasp selection state:* Once the object is selected, the system moves into the grasp selection state. A series of grasps are selected from a precomputed grasp database, beginning a closed loop online grasp planning process. In this stage each of the grasps is visually distinct, and supplies the planner with an approach angle and rough grasp location to start with. Here, we take advantage of visual ambiguity – as the user needs to remember which option looks most like the grasp that he or she wanted to select, visually similar grasps are functionally equivalent for the purposes of seeding our online grasp planner. An example of the subject guiding the system through this state can be seen in **??**

*3) Grasp refinement state:* In the grasp refinement state, the grasp planner searches for reachable grasps which are similar to those that the user selects. Simultaneously, the user is presented with the currently available grasp options.

As the grasps from the database are often not reachable in the initial case, the grasping pipelines waits until there are reachable grasps that are sent to the user. The user is asked to select the closest grasp to the one that is desired in each trial, which the planner then uses as a basis to search from. This allows the user to effectively walk the grasping point and direction to a desired location, even if that approach vector is not in the database.

The current grasp is highlighted in each RSVP sequence; if the user selects that option, we exit the grasp refinement state.

*4) Confirmation state:* In the confirmation state, the user is given a chance to confirm that the selected grasp is the one that he or she would like to execute. Once selected, the system sends the desired parameters to MoveIt! for execution.

At any point, the user is also presented with the option of going backwards in the pipeline, so as to select a different option. In experimental results, this is not often necessary.

### C. Experiment

The user is asked to grasp a container shaped object from among several presented. The objects are placed so that they are a somewhat natural disarray, but do not fully occlude other objects from the perspective of the camera. Our user was able to proceed through the grasping pipeline 5 times without incident, and each attempt ended in a successful grasp of the object. These experiments were conducted under Columbia IRB protocol IRB-AAAJ6951.


## ACKNOWLEDGMENT

References are important to the reader; therefore, each citation must be complete and correct. If at all possible, references should be commonly available publications.


## REFERENCES


[1] J. Weisz, B. Shababo, L. Dong, and P. K. Allen, "Grasping with your face," *Springer Tracts in Advanced Robotics*, pp. 435–448, 2013. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-00065-7_30

[2] J. Weisz, A. G. Barszap, S. S. Joshi, and P. K. Allen, "Single muscle site semg interface for assistive grasping," *IROS*, September 2014.

[3] E. A. Pohlmeyer, J. Wang, D. C. Jangraw, B. Lou, S.-F. Chang, and P. Sajda, "Closing the loop in cortically-coupled computer vision: a braincomputer interface for searching image databases," *J. Neural Eng.*, vol. 8, no. 3, p. 036025, May 2011. [Online]. Available: http://dx.doi.org/10.1088/1741-2560/8/3/036025

[4] P. Sajda, E. Pohlmeyer, J. Wang, L. Parra, C. Christoforou, J. Dmochowski, B. Hanna, C. Bahlmann, M. Singh, and S.-F. Chang, "In a blink of an eye and a switch of a transistor: Cortically coupled computer vision," *Proceedings of the IEEE*, vol. 98, no. 3, p. 462478, Mar 2010. [Online]. Available: http://dx.doi.org/10.1109/JPROC.2009.2038406

[5] A. Gerson, L. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, p. 174179, Jun 2006. [Online]. Available: http://dx.doi.org/10.1109/TNSRE.2006.875550

[6] M. C. Potter and E. I. Levy, "Recognition memory for a rapid sequence of pictures." *Journal of Experimental Psychology*, vol. 81, no. 1, p. 1015, 1969. [Online]. Available: http://dx.doi.org/10.1037/h0027470

[7] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, p. 510523, Dec 1988. [Online]. Available: http://dx.doi.org/10.1016/0013-4694(88)90149-6

[8] J. Weisz, C. Elvezio, and P. K. Allen, "A user interface for assistive grasping," *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov 2013. [Online]. Available: http://dx.doi.org/10.1109/IROS.2013.6696813

[9] C. Papazov and D. Burschka, "An efficient ransac for 3d object recognition in noisy and occluded scenes," in *Computer Vision–ACCV 2010*. Springer, 2011, pp. 135–148.