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 using advanced EEG-based interface. This online planning framework allows the user to direct the planner towards grasps that reflect their intent for using the grasped object by successively selecting grasps that approach the desired approach direction of the hand. The planner divides the grasping task into phases, and generates images that reflect the choices that the planner can make at each phase. The EEG interface is used to recognize the user's preference among a set of options presented by the planner. The EEG signal classifier is fast and simple to train, requiring a half hour training period before the test session. The system as a whole requires almost no learning on the part of the subject.**

I. INTRODUCTION

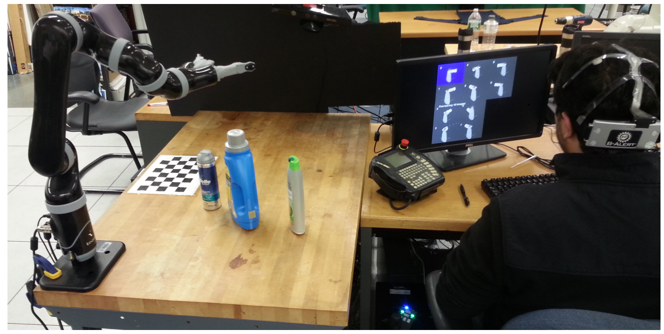
In previous work [1], [2], we have presented a series of refinements of a user interface for enabled a shared-control online grasp planner in moderately cluttered scenes. We have presented modifications of this system that integrate with a variety of facial EMG based devices. In this work, we further extend that UI to an EEG paradigm. This new device interface allows a 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 number of interfaces.

II. METHODS

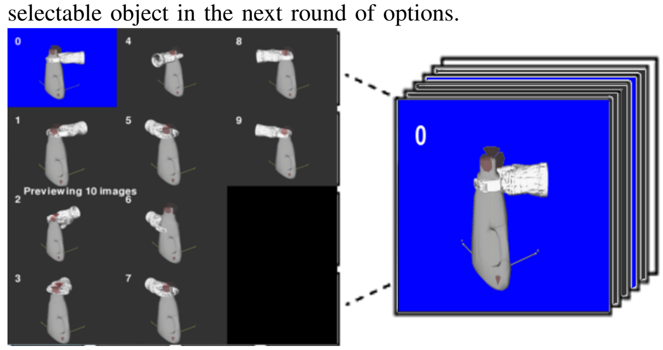
A. *One-of-many selection*

The EEG interface presented in this paper is based on an 'interest' detector which can be used to provide a one-of-many selection between various options. This 'interest' signal paradigm is based on the work in [**?**]. The options are presented as a stream of images, and the subject is primed to look for particular images that suit some criterion. This paradigm is known as *Rapid Serial Visual Presentation* (RSVP). Spikes in EEG activity which correlate with 'interest' are connected with the image that was presented at the time the EEG activity was evoked, which is then used to derive the user's desired input.

In order to grasp an object using this system, the grasping process is split into a multi-step pipeline where each step generates a visual representation of the options the user can take. Some options which cannot be visually represented, such as redoing a previous state, are presented as white text on a black background.



(a) The subject guiding the system through the grasp refinement stage. In this example, the blue grasp highlighted in the upper left is the highest scoring grasp, and is reachable. The user can select this option to stop the refinement stage and accept this grasp, or they can select a different grasp to bias the planner towards similar solutions. If this grasp is reachable, it will be sent as the highlighted, selectable object in the next round of options.



(b) The grasp planning system compiles a set of images representing potential actions, for example a set of grasps as seen in this image. The image options are tiled together to form the summary pane seen on the left, which lets the user pick out the one that reflects their desire. The images are then shuffled, with repetitions, into a stream that is serially presented to the user.

Previous work with this paradigm has asked the subject to look for objects of a particular category. In our system, the images represent a actions that are suggested by the grasp planner, which the subject may not have previous experience with. In this case, the subject must be given time to analyze the options and primed to find the features which make their desired option visually distinct from similar options. In Fig. 1b, we illustrate the *summary pane* containing a grid of all of the options, which are then shuffled and presented to the user. In Fig. 1a, you can see the subject reviewing the options in a *summary pane* before the serial presentation of them begins.

One major advantage of this paradigm is that it generalizes

*This work is support by NSF Grant IIS-1208153.

[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

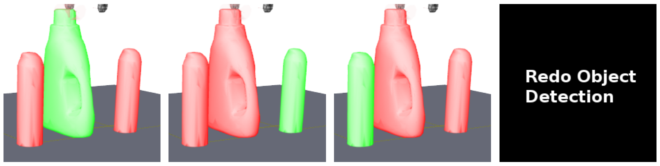


Fig. 2: The views presented in the object selection phase. The green object is the one which will be grasped if the image is selected. The text option will rerun the object detection.

a single interaction across all phases of the grasp planning pipeline. The system only has to be trained to recognize the 'interest' signal for each subject. Afterwards, the subject's interaction with each phase is the same, and the system does not require training for each phase.

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.

2) Choosing options: To generate the RSVP sequence, the system randomly selects each option to appear between three and seven times. The sequence is then randomly ordered so that the same option does not appear in two consecutive presentations. If there are less than five options, the system will automatically fill in distractor image options to make this constraint more feasible. These images th the word 'distractor'.

The images are each presented at 4 Hz, and EEG scores e_i are assigned. We aggregate each of the n images by their option, and determine whether the user has made a selection.

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

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

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.

B. Grasping Pipeline

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

1) Object selection state: In this stage, an object recognition system is used to retrieve models from a database that fit the scene. An image represented selection of each object is generated as shown in Fig. 2, with the target object highlighted green. Between the various images only the highlighted object changes. An additional state is presented that allows the user to run the recognition system again. The user simply thinks about the object they wish to select as the images are show.

2) Grasp selection state: Once the object is selected, the system moves into the grasp selection state. A set of preplanned grasps for the object are retrieved from a 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 to start with. Here, we take advantage of visual ambiguity – as the user as visually similar grasps are functionally equivalent for the purposes of seeding our online grasp planner. In Fig. 1a, a subject guiding the system through this state can be seen

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. Every 15 seconds, the available grasps are presented to the user. The user's selected grasp is then used to set the approach direction and hand state used to seed the planner. The reachable grasp with the heighest quality score is also sent as an option highlighted in blue. If no reachable option has been found yet, no highlighted grasp is option is presented. If this option is selected, the displayed grasp is *selected*, and the refinement state ends. The user 'walk' the grasping point and direction to a desired approach direction, even if it is not initially represented in the options.

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 grasp is executed by the robot.

At any point, the user is also presented with the option of going backwards in the pipeline to select a different option.

C. Experiment

The user i was to grasp a container shaped object from a table top with several objects. The objects are placed so that they are close together, 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

The authors would like to gratefully acknowledge the help of Dr. Paul Sajda and Dr. David Jangraw in this work.

REFERENCES

[1] J. Weisz, B. Shababo, L. Dong, and P. K. Allen, "Grasping with your face," *Springer Tracts in Advanced Robotics*, 2013.

[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," *Journal of Neural Engineering*, vol. 8, no. 3, p. 036025, 2011. [Online]. Available: http://stacks.iop.org/1741-2552/8/i=3/a=036025