Grasping With Your Face

Jon Weisz, Ben Shababo, Lixing Dong and Peter K. Allen

Abstract BCI (Brain Computer Interface) technology shows great promise in the field of assistive robotics. In particular, severely impaired individuals lacking the use of their hands and arms would benefit greatly from a robotic grasping system that can be controlled by a simple and intuitive BCI. In this paper we describe an end-to-end robotic grasping system that is controlled by only four classified facial EMG signals resulting in robust and stable grasps. A front end vision system is used to identify and register objects to be grasped against a database of models. Once the model is aligned, it can be used in a real-time grasp planning simulator that is controlled through a non-invasive and inexpensive BCI interface in both discrete and continuous modes. The user can control the approach direction through the BCI interface, and can also assist the planner in choosing the best grasp. Once the grasp is planned, a robotic hand/arm system can execute the grasp. We show results in using this system to pick up a variety of objects in real-time, from a number of different approach directions, using facial BCI signals exclusively. We believe this system is a working prototype for a fully automated assistive grasping system.

1 Motivation

With recent advances in robotics and computer vision, it is possible to imagine a robotic system to assist people with severely limiting disabilities in activities of daily living (ADL), improving their quality of life. ADL frequently require the user to grasp an object stably in a context aware way. Complex hands and manipulators increase the flexibility and grasping capabilities of a robotic assistant, but at the cost of requiring more complex control of many DOFs. Our goal is to create a robust system that can control a dexterous grasping system in real-time using only a small number of signals from an inexpensive and non-invasive BCI device with minimal training.

Department of Computer Science, Columbia University

Online, interactive control of robotic arms and hands for grasping in natural environments is a difficult problem. Typically, most systems use simple, parallel jaw grippers which simplifies the grasping process. Using more complex and higher DOF robotic hands increases the versatility of the system but at the cost of higher complexity for control. This generally requires more input from the user. More input, however, requires more channels, more processing, higher latencies, and generally higher cost. Our solution to this trade-off is to use a higher DOF hand with an accompanying high level interface that offloads the complexity of input requirements for the user to a simple and intelligent user interface. To avoid the issue of increasing the complexity of the input to the hand but maintain the flexibility of the more complex hand, the user needs a high level and intuitive interface.

2 Prior Work

A significant proportion of BCI-robotics research has focused on manipulating robotic arms and hand and different strategies for Simplementing solutions via electrophysiological signals have been investigated. Vogel et al. [26] showed that using the BrainGate cortically implanted electrode a subject was able to exercise Cartesian velocity control over an end effector and control opening and closing of the hand. However. this approach requires an invasive device capable of recording a large number of high quality signals.

It has also been shown that non-invasive devices can exercise effective control over robotic arms. For example, distal limb surface EMG signals have been used to control robotic arms in several applications [1, 4]. However, thus far accurate real time control has only been demonstrated for simple trajectory tracking tasks while using a large number of signals.

In order to reduce the number and quality of human signals needed, some intermediate level abstractions have been used. In [23], Shenoy et al. demonstrated basic cartesian control of a robot arm and gripper from 9 forearm electrodes to perform basic pick and place tasks. Another approach to abstraction is the use of forearm EMG signals to quickly switch between preset discrete states of various robotic hands [29, 28, 10, 6, 13]. In many situations in which assistive BCI-robotics would be applicable, limb EMG signals may not always be available or convenient. Therefore, various authors have proposed control schemes using face and head EMG signals to control robotic arms and grippers [20, 8, 18].

EEG has also been developed for BCIs to control robotic arms and hands in simple tasks. In [24, 9] BCI signals are used to control functional electrical stimulation to close and open a subject's wrist. In [17], surface electrode signals related to eye gaze direction are used to control 2D arm position and EEG signals are used to detect eye blinks to control gripper closing. In [11] hand opening/closing and elbow flexion/extension are controlled by EEG signals.

The majority of this previous work concentrates on trajectory control. However, it has been shown that users find BCI control easier using even higher level, goal oriented paradigms [19]. We have begun to see work that attempts to exploit higher level abstractions to allow users to perform more complex tasks with robotic arms. In [2], EEG signals were used to select targets for pick and place operations for a small humanoid robot. Waytowich et al. [27] used EEG signals to control pick and place operations of a 4-DOF Stäubli robot. Bryan et al. [12] presented preliminary work extending this approach to a grasping pipeline on the PR2 robot. In that work, a 3D perception pipeline is used to find and identify target objects for grasping and EEG signals are used to choose between them. In [15], grasping is decomposed to a 4 stage pipeline where EEG signals are used to control transitions between stages. And in [22], the authors demonstrate an interface to navigate in two dimensions and select goals in a complex virtual environment and propose a hierarchical control scheme for learning high level tasks dynamically.

While previous work has shown that complex interactions can be mediated by BCI signals, *thus far a fully developed, end-to-end, real-time, BCI-based grasping system for complex hands has not been demonstrated*. A full grasping pipeline, which our system addresses, must integrate target selection and localization, multi-DOF hand configuration planning, and approach trajectory planning with user intent decoded from a noisy low dimensional BCI signal.

3 Technical Approach

Figure 1 is an overview of our system. A 3D range camera is used to image an object to be grasped. This range image is used to both identify and align the object from a database of existing models [7]. Once the model is chosen and aligned, we show that four simple signals are sufficient to supervise an online grasp planning system resulting in robust grasps using real sensor data for object localization. By carefully reducing the configuration space of our planning and control architecture, our system enables subjects to select grasps appropriate to the desired task. The user interacts in real-time with the robot through a kinematic simulator which allows the user to visualize and supervise all of the elements necessary to plan a task specific grasp. The goal of this approach is to make simple, inexpensive BCI devices powerful enough to allow users to grasp many objects that are important for everyday living in a context aware way.

The grasping pipeline is divided up into four stages: object identification and alignment, grasp planning, grasp review, and grasp execution, which are described below. This pipeline is controlled using only four facial gestures. The use of these gestures in each stage of the pipeline is explained in Table 1. In general, gesture 1 serves as a ”click” and moves the user through the pipeline. The exception to this is at points of decision for the user. In these cases, gesture 2 serves as the ”YES” option and gesture 1 becomes ”NO” and returns the user to an earlier point in the pipeline. Because false positive readings of gesture 1 and 2 have strong consequences, we found that both are best associated with a concise and strong gesture such as closing one eye or clenching the jaw. Gestures 3 and 4 control the approach direction of the hand relative to the object during the grasp planning stage of the pipeline. These gestures can be maintained to generate *continuous* motion of the hand over two degrees of freedom and therefore are best associated with gestures that can be contracted for several seconds without too much twitching or fatigue.

Gesture	Run Planner	Review Grasps	Execution
1	start/stop planner	cycle through grasps	restart
2	n/a	select grasp	confirm grasp
3	rotate around x-axis	n/a	n/a
4	rotate around z-axis	n/a	n/a

Table 1 A description of the user interface as the user progresses through phases of the pipeline.

Object Identification and Localization: Our grasp planner requires a complete description of the geometry and location of the target object. The raw point cloud data is gathered by a Microsoft Kinect. In this work, we assume the target is a member of a known set of objects. We use the method described in [16] to identify and localize the target object in the scene. Briefly, this method generates features from pairs of oriented points on the surface of the object. Prospective models are processed offline and put in to a hash table. Features are sampled from the sensor data and tested for collision in the hash table. If a sufficient number of collisions occurs with points on the same model, a varient of RANSAC is used to test the hypothesis that a set of points in the sensor data corresponds to a particular model at a particular location. Fig. 2 shows a correctly chosen model aligned with the range scan. This method is robust and fast enough to demonstrate the efficacy of our BCI-grasping pipeline.

Grasp Planning Phase: In this work we use the Eigengrasp Grasp Planner developed by Ciocarlie and Allen, the details of which can be found in [5]. In this system, grasps are planned through stochastic optimization using simulated annealing. Recent advances in neuroscience research have shown that control of the human hand during grasping is dominated by movement in a configuration space of highly reduced dimensionality[25, 21]. The Eigengrasp Grasp Planner uses these insights to reduce the the dimensionality of the hand's joint postures to a subspace with a lower dimensionality. This allows the planner to function in real time with a human user in the loop by guiding the planner by partial demonstration. The planner allows the operator to specify how much the stochastic optimization process can vary each parameter that describes the demonstration pose against the gradient of the grasp energy function in searching for a valid, complete grasp. The planner optimizes a grasp energy function as the sum of two parts, a measure of how nearly a hand conforms to the object, and a continuous approximation of the canonical Ferrari-Canny grasp quality measure. Hand configurations with a high enough quality function are marked as potentially good grasps. These configurations will put the hand in a position which is near the object but where the finger configuration is restricted to the postural subspace described above. The quality of these configurations as starting grasp poses is then reassessed after a kinematic simulation of grasping. The hand

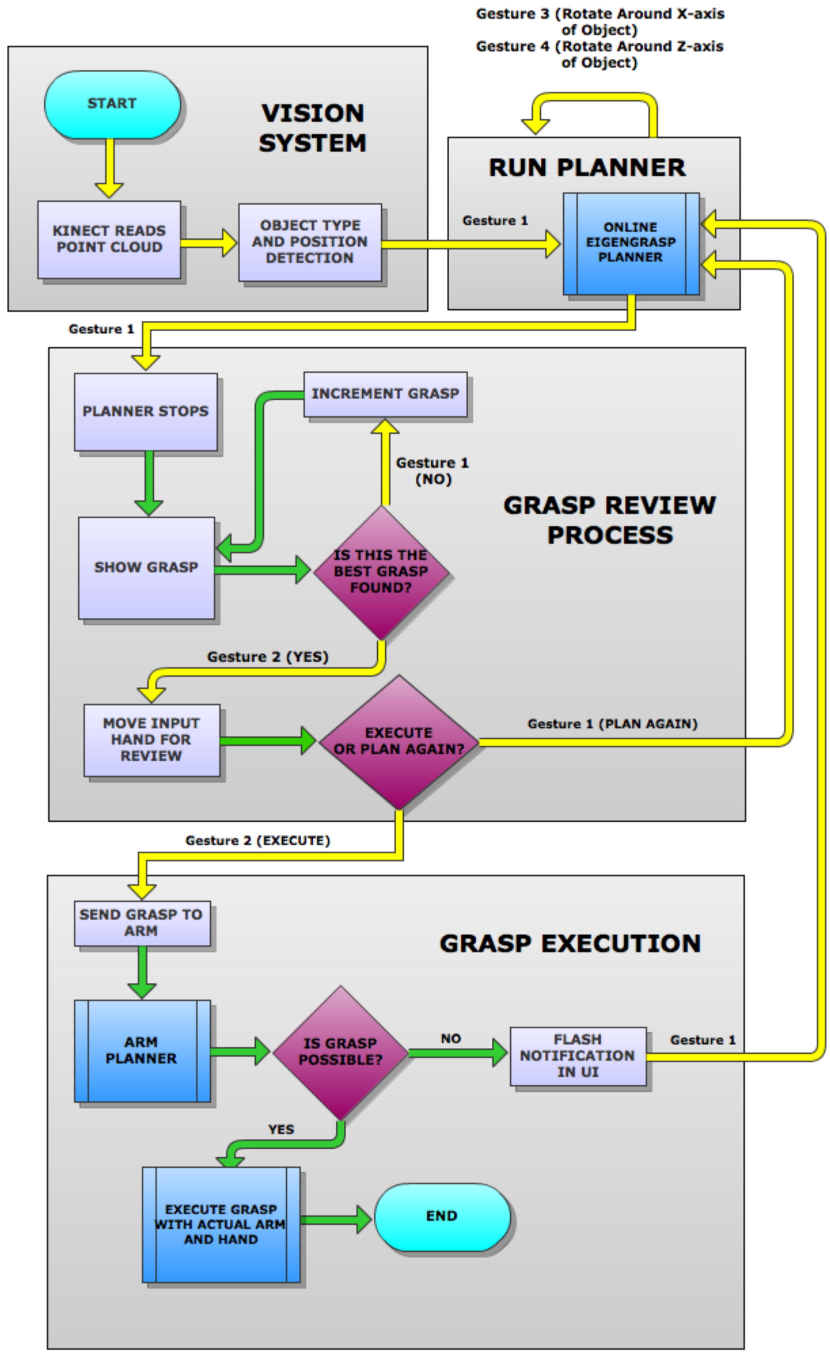

Fig. 1 An overview of the grasping pipeline. In the first phase, the system uses the kinect to identify and localize the object in the scene. In the second phase, the user guides the planner to an appropriate approach direction. In the third phase, the user stops the planner and reviews the available grasps. In the fourth phase, the user sends the grasp to the robot for execution.

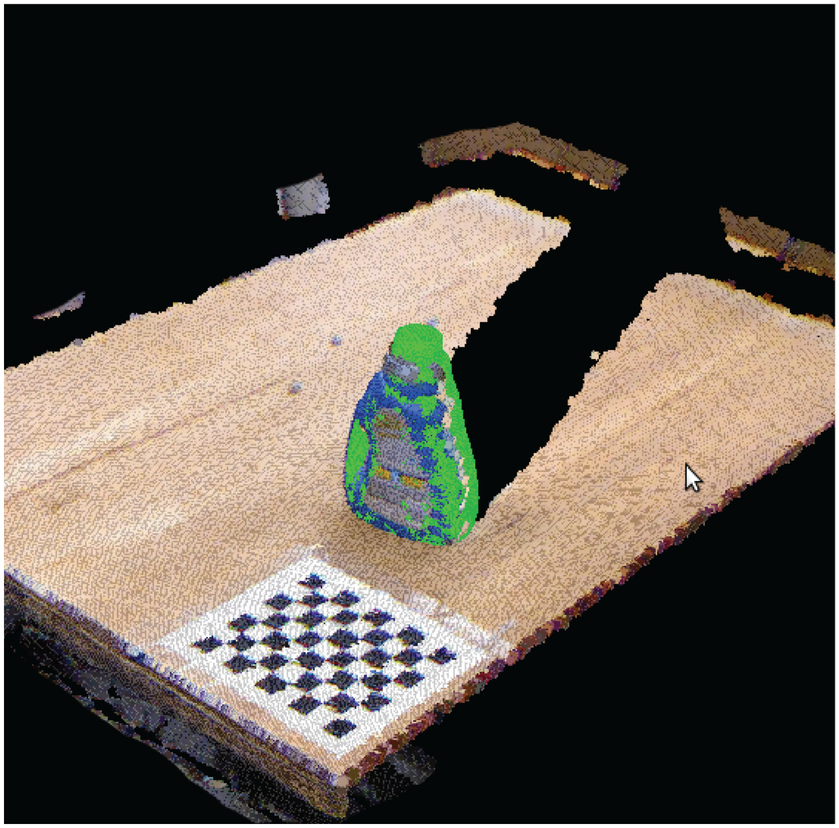

Fig. 2 The point cloud with RGB texture from the vision system. The blue bottle is the object point cloud, while the green represents the detected model overlain on the scene.

approaches the object and closes all of its joints, leaving the postural subspace as necessary to conform to the object.

The user is presented with a world in our grasping simulator, *GraspIt!* [14], that contains the hand, the object detected by the vision system, and the surface on which the object sits. Elements of the user interface for the planning phase are shown in Fig. 3. When planning begins, two clones of the hand are placed into the world, and the transparency of each of the three hands identifies it to the user. The most transparent hand represents the planning process and shows samples of the grasp space region the planner is exploring. This online feedback helps the user choose an approach direction that best communicates their intent to the planner.

The input hand has intermediate transparency, and the user can control the rotation of the input hand around the x and z axis of the object, respectively. This effectively moves the hand around the surface of a sphere while maintaining that the approach direction of the hand faces the center of the object. The third hand during the planning phase is the solution hand which is fully opaque. As the planner

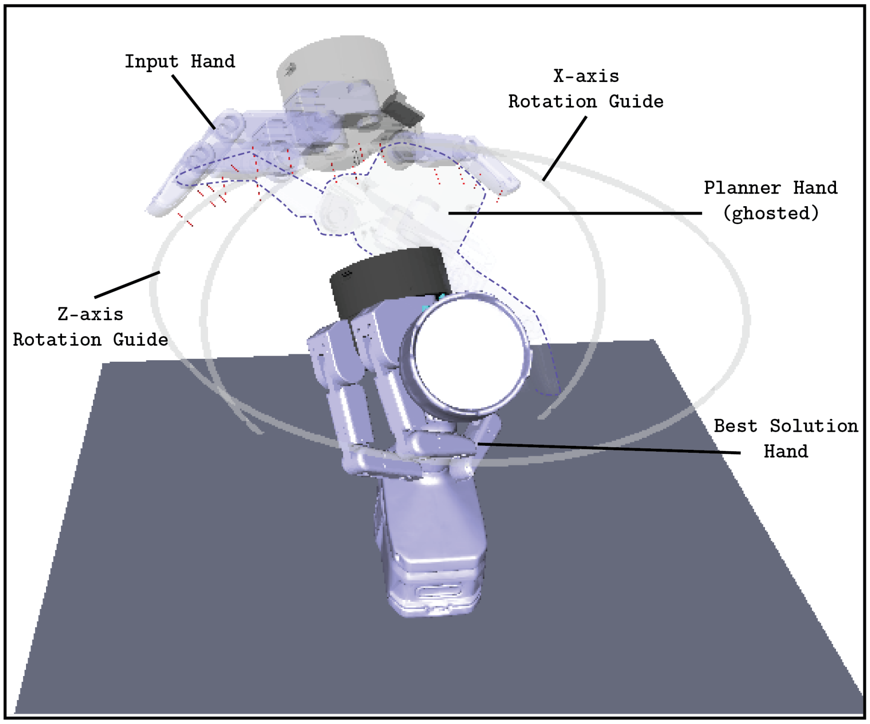

Fig. 3 An illustration of the grasp planning user interface in GraspIt!. Here we demonstrate the three hands of the system. The 'Planner Hand,' which is most transparent hand, demonstrates the current state of the planner. The 'Input Hand,' which is of intermediate transparency, is the hand through which the user directs the planning system. Here you can see the rotational guides which allow the user to visualize their available control directions. The 'Solution Hand', which is fully opaque, demonstrates the best grasp currently available. This is the grasp which is closest to the approach direction that the Input Hand is demonstrating and which also has the minimal grasp energy.

runs, it stores the ten best grasps in a list. The solution hand demonstrates the current best grasp. The list of best grasps is sorted such that preference is given to solutions that reflect the user's desired approach. When the user is satisfied, they can stop the planner and progress into the review process using gesture 1.

Review Phase: Once the planner is stopped the user has an opportunity to review the list of best grasps and choose one for execution. As in the previous stage, the solution hand is used to display the grasps to the user. At any point, the user can select a grasp which removes the solution hand from the world and closes the input hand into the chosen grasp. Now the user can evaluate the grasp more closely and examine the quality metrics for the grasp. If the user is satisfied, they can confirm selection of the grasp and send it to the next stage of the pipeline. If the user does not want to execute any of the found grasps, they can select the grasp that is closest to what they have in mind and restart the planning process. Importantly, the next

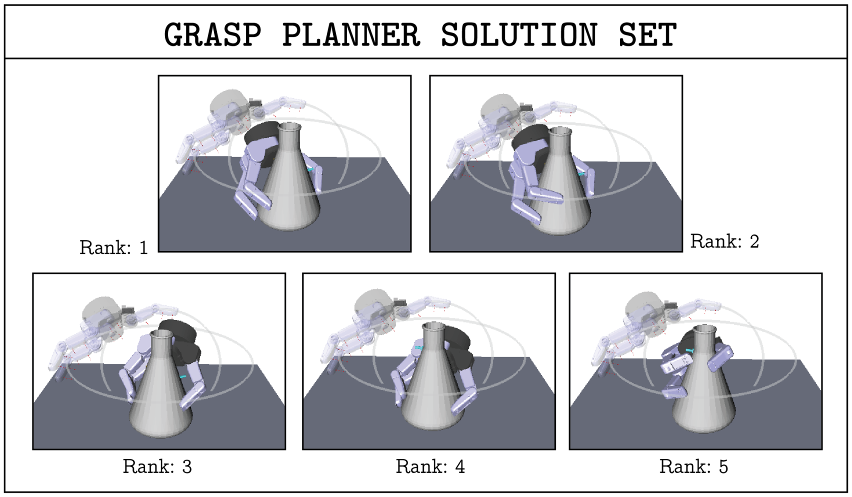

Fig. 4 An illustration of the review phase. Using one of the facial gestures, the user can cycle through the available grasps and visualize them in the simulator. The grasps are sorted by their closeness to the last demonstrated approach vector. If none of the grasps are suitable, the user can go back to the planning phase with a second facial gesture. Otherwise, the user selects a grasp to be executed using a third gesture.

iteration of planning will not only take the input hand's position as a constraint but also the eigengrasp values of this selected grasp.

Execution Phase Once the robotic control software receives the selected grasp which is planned relative to the object's coordinate system, the arm planner decides if that grasp is achievable given the kinematic constraints of the arm and the vision system's estimate of the object's position in the world. If the planner cannot find a solution to execute the grasp, it will send a message back to GraspIt! which then notifies the user. If this occurs, the user can restart the planning phase to find a new grasp. If the planner can find a solution, then the grasp is executed with the actual arm and hand.

4 Experimental Results

4.1 Brain-Computer Interface

Hardware: The BCI hardware used for this experiment was the Emotiv EPOC which is is a low-cost 16 electrode headset. Of the 16 electrodes, 14 are positioned at the International 10-20 locations corresponding to AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2. The remaining 2 channels are references and are positioned

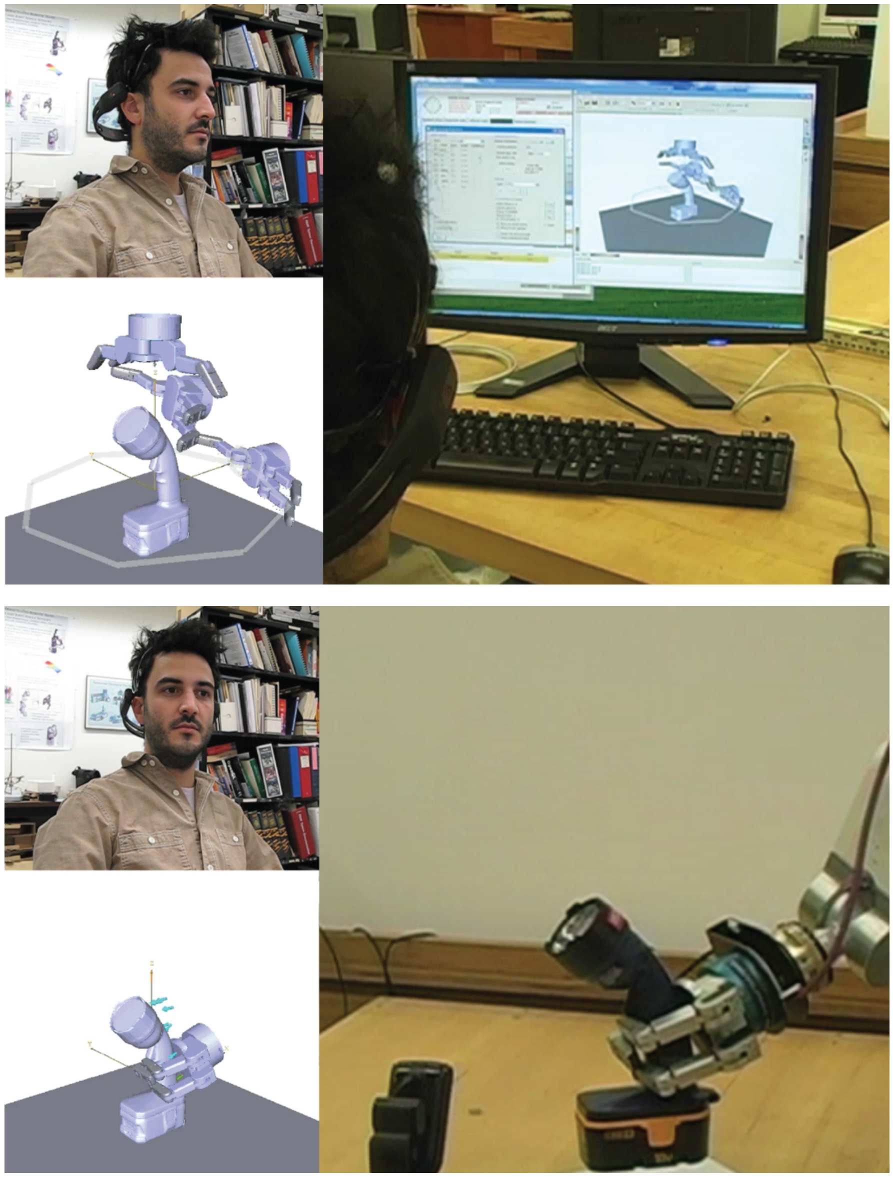

Fig. 5 Two sets of video stills from the actual execution of the system. **Top:** Clockwise from the top-left are an image of the user wearing the Emotiv headset and operating the system, the computer monitor with user interface, and a screen capture of the simulation. **Bottom:** User watching the robotic arm implementing the selecting grasp as shown in the simulator window on the lower left.

at P3 and P4. Each electrode has a sampling rate of roughly 128 hz and a resolution of 16 bits (14 bits effective). The headset communicates with the computer via a bluetooth connection.

Software: We used the Cognitiv Suite and Expressiv Suite software that come bundled with the Emotiv Epoc to process the electrophysiological signals. The Cogntiv Suite is designed to be trained on different cognitive states, usually associated with mental imagery or imagined movement. However, because the EMG signals from facial gestures are much stronger and more easily produced in a consistent manner, we chose to train the Cogntiv Suite on these types of signals instead of cognitive states. First, the system is trained on a neutral signal to define a baseline signal, then a set of facial gesture states are learned. In the examples presented here, gesture 1 is mapped to clenching the jaw muscles and is detected through the Expressiv Suite. The remaining gestures are detected through the Cognitiv Suite. Gesture 2 is mapped to closing the right eye, gesture 3 is mapped eyebrow movement, and gesture 4 is mapped to tensing the muscles around the ear. This combination of classifiers and gestures was found by trial and error to be reliable for the particular user discussed in this paper, but the system as a whole is in no way dependent on which particular gestures are used.

4.2 Robot Setup

Hardware: Our grasping platform is comprised of a 280 model BarrettHand mounted on a Stäubli TX60L 6-DOF robotic arm. A Microsoft Kinect sensor is used to generate point clouds of the object.

Software: The Barrett hand is commanded in position control mode using the OpenWAM driver. Planning for the motion of the arm is done in OpenRave using a bidirectional random tree planner by Berenson et al. [3], and small linear motions near the object are planned using the built-in inverse kinematics planner on the Stäubli TX60L arm.

4.3 Training

One subject completed 10 successful training periods on the neutral signal and each of the three facial gestures that would be processed by the Cogntiv Suite. Each training period lasts 8 seconds, and after each period the subject is asked whether or not they were able to maintain the appropriate gesture for the entire time. Only sessions in which the subject answered yes are considered successful and included in the training data. During training, the subject is given visual feedback via the motion of a cube and a power meter both of which represent the strength of current classifica-

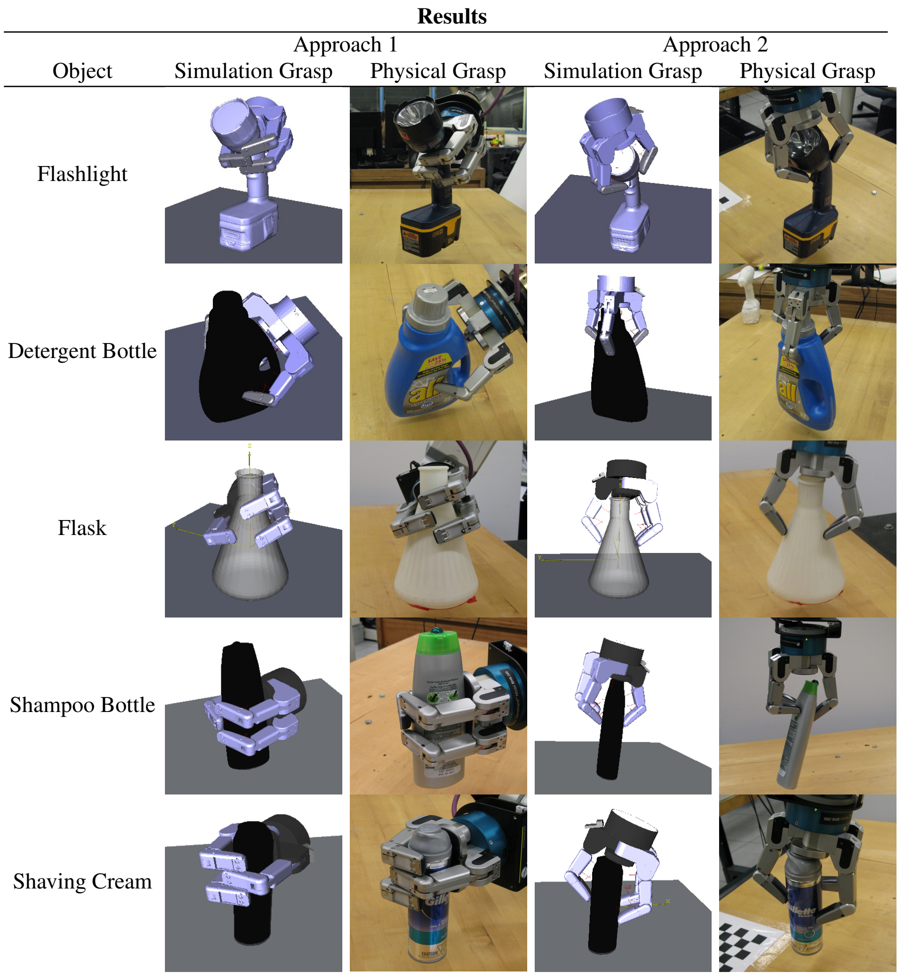

Fig. 6 The results of ten different attempts to grasp five objects using two different user-selected approach directions per object.. In each case, the user was able to generate an appropriate grasp using the simulator which was then transfered to the robot.

tion. The Expressiv Suite does not need any training, however, the sensitivity of the classifier for jaw-clenching was adjusted for best performance for the subject.

4.4 Grasping Experiments

The subject used the system to grasp and pick up five common objects: a flashlight, a flask, a bottle of laundry detergent, a bottle of shampoo, and a canister of shaving cream. For each setup, the Stäubli arm starts such that it is completely vertical with the BarrettHand at the top. A simulation world consisting of the hand, the object, and a surface is presented to the subject who can move the input hand and start the planner at any time. For each object, the subject executed grasps from two different approach directions.

In Fig. 6, we show the results of ten different attempts to grasp five objects using two different approach directions per object. In each case, we demonstrate the grasp planned in simulation and the final grasp achieved by the physical robot. We found that using this grasp planning environment, we were easily and reliably able to grasp objects. In each case, the physical grasp was successful on the first attempt. Planning the grasp takes on the order of 10-30 seconds. The review process takes an additional 10-15 seconds. We refer the reader to the video at `http://robotics.cs.columbia.edu/jweisz/bciGraspingISER2012` for an example of the entire procedure.

4.5 Discussion

In this work we have described an end-to-end system for grasping objects using a low-cost, non-invasive BCI device. Although further refinements and user studies still need to be done, we have shown three important ideas in this paper. First, we have shown that a user interface using only two user controlled dimensions is expressive enough to demonstrate the user's intent by controlling the approach direction of the hand and that the Eigengrasp planner is able to produce reliable and appropriate grasps using this information alone, vastly simplifying the grasp planning process. Second, we have shown that a low cost, non-invasive, noisy, low-bandwidth BCI interface is sufficient to guide the grasp planner. By reducing the complexity of the grasp planning process, we can accomodate such a BCI device. Third, a central tenet of this work is that by keeping a human-in the-loop, we are able to support context-aware grasping, spanning different grasps for different tasks. Although the grasp quality measure used by the planner is sufficient to produce a reasonable list of plausible grasps, it is not sufficient to chose the most reliable and appropriate grasp among them, which can lead to failures. With a human adding their own intuition to select a grasp from the candidate list, we have never seen a failure in grasping the object.

We believe our user interface works well, but it is yet to be tested with multiple subjects, which is the subject of ongoing research in our lab. As part of this work, we are also analyzing how much training of the BCI control is needed to create a competent user. Making the system easy to learn and use is an important goal of this work. Finally, we are also experimenting with using EEG signals from the Emotiv EPOC instead of using facial gestures which generate EMG signals.

References

1. Artemiadis, P.K., Kyriakopoulos, K.J.: A switching regime model for the EMG-based control of a robot arm. IEEE transactions on systems, man, and cybernetics **41**(1), 53–63 (2011)
2. Bell, C.J., Shenoy, P., Chalodhorn, R., Rao, R.P.N.: Control of a humanoid robot by a noninvasive brain-computer interface in humans. Journal of neural engineering **5**(2), 214–20 (2008)
3. Berenson, D., Srinivasa, S.S., Kuffner, J.: Task Space Regions: A framework for pose-constrained manipulation planning. The International Journal of Robotics Research (2011)
4. Castellini, C., van der Smagt, P.: Surface EMG in advanced hand prosthetics. Biological cybernetics **100**(1), 35–47 (2009)
5. Ciocarlie, M.T., Allen, P.K.: Hand posture subspaces for dexterous robotic grasping. The International Journal of Robotics Research **28**(7), 851–867 (2009)
6. Cipriani, C., Zaccone, F., Micera, S., Carrozza, M.: On the Shared Control of an EMG-Controlled Prosthetic Hand: Analysis of User Prosthesis Interaction. IEEE Transactions on Robotics **24**(1), 170–184 (2008)
7. Goldfeder, C., Ciocarlie, M., Peretzman, J., Dang, H., Allen, P.K.: Data-driven grasping with partial sensor data. In: IROS'09: Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems, pp. 1278–1283. IEEE Press, Piscataway, NJ, USA (2009)
8. Gomez-Gil, J., San-Jose-Gonzalez, I., Nicolas-Alonso, L.F., Alonso-Garcia, S.: Steering a Tractor by Means of an EMG-Based Human-Machine Interface. Sensors **11**(7), 7110–7126 (2011)
9. Hazrati, M.K., Erfanian, A.: An on-line BCI for control of hand grasp sequence and holding using adaptive probabilistic neural network. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. **2008**, 1009–12 (2008)
10. Ho, N.S.K., Tong, K.Y., Hu, X.L., Fung, K.L., Wei, X.J., Rong, W., Susanto, E.A.: An EMG-driven exoskeleton hand robotic training device on chronic stroke subjects: Task training system for stroke rehabilitation. In: 2011 IEEE International Conference on Rehabilitation Robotics, pp. 1–5. IEEE (2011)
11. Horki, P., Solis-Escalante, T., Neuper, C., Müller-Putz, G.: Combined motor imagery and SSVEP based BCI control of a 2 DoF artificial upper limb. Medical & biological engineering & computing **49**(5), 567–77 (2011)
12. M. Bryan, V. Thomas, G. Nicoll, L. Chang, J.S., Rao, R.: What You Think is What You Get: Brain-Controlled Interfacing for the PR2. Tech. rep., Iros 2011: The PR2 Workshop, San Francisco (2011)
13. Matrone, G., Cipriani, C., Carrozza, M.C., Magenes, G.: Two-channel real-time EMG control of a dexterous hand prosthesis. In: 2011 5th International IEEE/EMBS Conference on Neural Engineering, pp. 554–557 (2011)
14. Miller, A.T., Allen, P.K.: Graspit!: A versatile simulator for robotic grasping. IEEE Robotics and Automation Magazine **11**, 110–122 (2004)
15. Müller-Putz, G.R., Scherer, R., Pfurtscheller, G., Rupp, R.: EEG-based neuroprosthesis control: a step towards clinical practice. Neuroscience letters **382**(1-2), 169–74 (2005)
16. Papazov, C., Burschka, D.: An efficient ransac for 3d object recognition in noisy and occluded scenes. In: Computer Vision ACCV 2010, vol. 6492, pp. 135–148 (2011)
17. Postelnicu, C.c., Talaba, D., Toma, M.i.: Controlling a Robotic Arm by Brainwaves and Eye. Ifip International Federation For Information Processing pp. 157–164 (2011)
18. Ranky, G.N., Adamovich, S.: Analysis of a commercial EEG device for the control of a robot arm. In: Proc. IEEE Northeast Bioengineering Conference (NEBEC), pp. 1–2. New York, NY (2010)
19. Royer, A.S., Rose, M.L., He, B.: Goal selection versus process control while learning to use a brain-computer interface. Journal of neural engineering **8**(3), 036,012 (2011)
20. Sagawa, K., Kimura, O.: Control of robot manipulator using EMG generated from face. In: ICMIT 2005: Control Systems and Robotics, vol. 6042, pp. 604,233–604,233–6 (2005)
21. Santello, M., Flanders, M., Soechting, J.F.: Patterns of hand motion during grasping and the influence of sensory guidance. The Journal of Neuroscience **22**(4), 1426–1435 (2002)

22. Scherer, R., Friedrich, E.C.V., Allison, B., Pröll, M., Chung, M., Cheung, W., Rao, R.P.N., Neuper, C., Pr, M.: Non-invasive brain-computer interfaces: enhanced gaming and robotic control. In: Advances in Computational Intelligence, vol. 6691/2011, pp. 362–369 (2011)
23. Shenoy, P., Miller, K.J., Crawford, B., Rao, R.N.: Online electromyographic control of a robotic prosthesis. IEEE transactions on bio-medical engineering **55**(3), 1128–35 (2008)
24. Tavella, M., Leeb, R., Rupp, R., Millan, J.D.R.: Towards natural non-invasive hand neuroprostheses for daily living. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. **2010**, 126–9 (2010)
25. Tsoli, A., Jenkins, O.C.: 2d subspaces for user-driven robot grasping. In: RSS Workshop on Robot Manipulation: Sensing and Adapting to the Real World. Atlanta, GA (2007)
26. Vogel, J., Haddadin, S., Simeral, J.D., Stavisky, S.D., Bacher, D., Hochberg, L.R., Donoghue, J.P., Van Der Smagt, P.: Continuous Control of the DLR Light-weight Robot III by a human with tetraplegia using the BrainGate2 Neural Interface System. In: International Symposium on Experimental Robotics (2010)
27. Waytowich, N., Henderson, A., Krusienski, D., Cox, D.: Robot application of a brain computer interface to staubli TX40 robots - early stages. World Automation Congress (WAC), 2010 pp. 1–6
28. Woczowski, A., Kurzyski, M.: Human-machine interface in bioprosthesis control using EMG signal classification. Expert Systems **27**(1), 53–70 (2010)
29. Yang, D., Zhao, J., Gu, Y., Jiang, L., Liu, H.: EMG pattern recognition and grasping force estimation: Improvement to the myocontrol of multi-DOF prosthetic hands. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 516–521. IEEE (2009)