Autonomy Infused Teleoperation
with Application to BCI Manipulation

Katharina Muelling\textsuperscript{1}, Arun Venkatraman\textsuperscript{1}, Jean-Sebastien Valois\textsuperscript{1}, John Downey\textsuperscript{4,6}, Jeffrey Weiss\textsuperscript{3}, Shervin Javdani\textsuperscript{1}, Martial Hebert\textsuperscript{1}, Andrew B. Schwartz\textsuperscript{1,5}, Jennifer L. Collinger\textsuperscript{2,3,4}, J. Andrew Bagnell\textsuperscript{1}

\textsuperscript{1}Carnegie Mellon University, Robotics Institute, Pittsburgh, PA, USA
\textsuperscript{2}VA Pittsburgh Healthcare System, Pittsburgh, PA, USA
\textsuperscript{3}University of Pittsburgh, Department of Physical Medicine and Rehabilitation, Pittsburgh, PA, USA
\textsuperscript{4}University of Pittsburgh, Department of Bioengineering, Pittsburgh, PA, USA
\textsuperscript{5}University of Pittsburgh, Department of Neurobiology and Bioengineering
\textsuperscript{6}Center for the Neural Basis of Cognition, Pittsburgh PA

Abstract—Robot teleoperation systems introduce a unique set of challenges including latency, intermittency, and asymmetry in control inputs. User control with Brain-Computer Interfaces (BCIs) exacerbates these problems through especially noisy and even erratic low-dimensional motion commands due to the difficulty in decoding neural activity. We introduce a general framework to address these challenges through a combination of Machine Vision, User Intent Inference, and Human-Robot Autonomy Control Arbitration. Adjustable levels of assistance allow the system to balance the operator’s capabilities and feelings of comfort and control while compensating for a task’s difficulty. We present experimental results demonstrating significant performance improvement using the shared-control assistance framework on adapted rehabilitation benchmarks with two subjects implanted with intracortical brain-computer interfaces controlling a high degree-of-freedom robotic manipulator as a prosthetic. Our results further indicate shared assistance mitigates perceived user difficulty and even enables successful performance on previously infeasible tasks. We showcase the extensibility of our architecture to quality-of-life tasks such as opening a door with a BCI, pouring liquids from a container with a dual-joystick game controller, and manipulation in dense clutter with a 6-DoF motion controller.

I. INTRODUCTION

Robust robotic teleoperation systems must be capable of mitigating latency \cite{4,22}, intermittency \cite{36}, difficulty in performance of high-precision tasks \cite{33,27}, and reconciling the discrepancy between input mechanisms of the user and the robot (e.g. joysticks versus motor torques) \cite{33}. These challenges are especially present when using Brain-Computer Interfaces (BCIs) as input device to teleoperate a robotic manipulator. The advent of BCIs holds promising opportunities for empowering those physically impaired, restoring their mobility and abilities in the control of wheelchairs \cite{42} and prosthetic limbs \cite{42,14,8}. However, the difficulty in mapping recorded neural activity to teleoperation motion commands compounds the difficulties present in conventional teleoperation, hindering its applicability in contexts requiring high precision and dexterity. Reduced integrity of neural signals caused by the degradation of invasive BCIs over time \cite{39} results in increased erraticity and noise in the interpreted motion commands. This effect also reduces the dimensionality of the decodable control input from the user. Additionally, the lack of haptic feedback can result in dangerous interactions between the robotic prosthetic and the environment.

The combination of autonomous robot technologies and direct user control in a shared-control teleoperation frameworks can help to overcome these limitations and decrease the demands on the user \cite{9}. Autonomy infused user control leverages the strengths of human-in-the-loop supervision for higher-level task planning and exploits the ability of robotic systems to reliably solve high-precision tasks. However, the blending (arbitration) between autonomy and direct user control is crucial as it balances the comfort and perceived control of the operator. Although a high amount of autonomy can outperform direct operator control \cite{24}, user studies indicate that the amount of autonomy should depend on both the user’s ability to deal with the difficulty of the task and the robot’s confidence of the inferred user goal \cite{18,11,47}.

We propose a framework that addresses the challenges of robot teleoperation with a focus on BCI manipulation tasks. Our shared-control teleoperation architecture is inspired by efforts...
in both autonomous robotics and human-robotic interaction, combining computer vision, user intent inference \[13, 11, 1, 49\], and human-robot arbitration \[21, 3, 45, 5\]. The value of computer vision in BCI controlled manipulation was demonstrated with a simple sphere-and-cylinder detection systems in \[29, 15\]. We extend upon these findings by utilizing a more advanced perception pipeline (Section II-D) in conjunction with a model library concept containing 3D object models with corresponding pre-labeled grasp set metadata. In contrast with the fixed distance based threshold in \[15\], we introduce capture envelopes, similar to gravity fields \[40\], for smooth and continuous user grasp inference that can vary based on the object and grasp from the model library (Section II-A). Utilizing prior work in value function based user intent inference (Section II-B), we circumvent the requirement of explicit user goal selection by eye-tracking or similar interfaces as used in \[29\] to determine the user’s desired goal. Finally, following the suggestions of user studies \[47, 18, 11\], we allow for adjustable assistance levels through the human-robot arbitration (Section II-C) – blending the needs of the task with the operator’s capabilities while balancing their comfort, and sense of control. Experimental results showcase significant performance improvements of using our shared-control framework to assist in control of a 7 DoF robotic manipulator in conjunction with an intracordical BCI on adapted rehabilitation benchmarks (Section III). To highlight the generality and extensibility of our architecture, we show applications to quality-of-life tasks such as opening a door and pouring liquids as well as the ability to use conventional teleoperation interfaces (Section IV). A complementary overview of related work in shared teleoperation and BCI technologies (Section V) is given prior to the conclusion (Section VI).

II. AUTONOMOUS ROBOT MANIPULATION ASSISTANCE

The combination of computer vision, user intent inference, arbitration between autonomy and human control, and a compliant controller define the major components of our autonomy infused shared-control teleoperation system illustrated in Fig. 2. Below, we detail and describe the involved modules.

A. Capture Envelopes: Grasp Inference

For an assistive telemanipulation system to be intuitive and transparent, the prediction of user intentions is crucial. Even when there is only one object in the environment, it is not clear how the human wants to grasp this object, as one object can be grasped from multiple positions and poses. To estimate the user’s intent, we assume that the user is an intent driven agent following a policy to minimize a cost function \(c_g\) while approaching a goal \(g\) (i.e., grasp pose). User actions, such as the commanded end-effector velocities, lead to lower cost for some grasp poses while simultaneously increasing the cost for others. The process of intent estimation then becomes the measure of the user’s progress on each possible grasp. In the following, we will consider the case of inferring the grasp pose with which the user intends to grasp a single object. We then extend this approach to handle multiple objects by utilizing the principle of maximum entropy in Section II-B.

To infer the intended grasp pose of a single object, we assume that the grasp pose is defined by the type of object and the direction from which the operator approaches this object. The object specific grasp poses and their approach vector are stored in the model library. Similar to the idea of gravity fields \[40\] where the end-effector is pulled towards a specific position, we define capture envelopes to account for the fact that a grasp pose cannot be approached from all directions without undesirable collisions with the target object.

Formally, let \(\mathcal{G}\) be the set of grasps associated with an object in the environment, we want to select the grasp pose \(g \in \mathcal{G}\)
that the user most likely aims for. Let each grasp \( g \) contain a desired end-effector pose \( G = (R_G, x_G) \in SE(3) \) as well as an approach direction \( d_g \in \mathbb{R}^3 \). The cost for a grasp \( g \in \mathcal{G} \) is then computed by:

\[
c_g(G, E) = \frac{k_t c_{\text{tran}}(x_G, x_E)}{k_r c_{\text{rot}}(R_G, R_E)} + c_{\text{kin}}(G),
\]

where \( E = (R_E, x_E) \in SE(3) \) is the current end-effector pose and \( k_t, k_r \) are gain factors. Intuitively, this cost relationship prefers grasps that are close by to the end-effector location while favoring orientations similar to that of the current rotation of the end-effector. Specifically, \( c_{\text{tran}} \) computes the translational distance between the current end-effector position \( x_E \in \mathbb{R}^3 \) and the grasp position \( x_G \in \mathbb{R}^3 \). The function \( c_{\text{rot}} \) is computed from the quaternion dot product and decreases with increasing difference between the orientations \( R_G \) and \( R_E \). Finally, we append a cost \( c_{\text{kin}} \) that implements the kinematic feasibility constraint for the grasp pose, returning zero if the pose is feasible for the manipulator and infinity otherwise.

To compute the translation cost \( c_{\text{tran}} \), we define a capture envelope (Fig. 3), represented as a truncated cone with an origin close to the target and aligned with the approach vector \( d_g \):

\[
c_{\text{tran}}(x_G, x_E) = \begin{cases} t & \text{if } x_E \text{ is in the cone} \\ \infty & \text{else,} \end{cases}
\]

where \( t \) captures the progress of the end-effector along the grasp’s defined approach vector \( d_g \). Specifically, it is the projection of the vector between \( x_G \) and the position \( x_E \) projected onto the approach vector:

\[
t = \frac{(x_G - x_E)^T(x_L - x_G)}{\|x_L - x_E\|^2},
\]

where the launch position \( x_L \) is set to a predefined distance along the approach direction from the goal position \( x_G \). The capture envelope, as defined, provides a major advantage over methods that would only consider the radial distance to the target because the hand is never guided backward into the object when it approaches the object from above or behind.

Finally, all grasps \( g \in \mathcal{G} \) are ranked based on their cost \( c_g \), filtered based on a threshold cost. The best ranked grasp is used as the automated desired pose \( A \).

B. User Intent Inference via Maximum Entropy

Although the development of capture envelopes allows for successful assistive grasping with a single object, many real-world scenarios contain multiple objects. Thus, it becomes necessary to reason about the probability of each being the intended target. A simple approach that selects the best capture envelope across objects as described above, is limited by the fact that it only takes into account the current end-effector pose, rather than the entire trajectory. In this section, we describe a method based on the principle of maximum entropy to overcome this limitation. The principle of maximum entropy intuitively allows us to reason about the probability distribution over goals while making minimum commitment beyond the information observed so far. After the most likely object is identified, we leverage the capture envelope ranking as described above to achieve successful grasping behaviors.

We assume the user is a rational agent running an optimal controller to minimize some cost-to-go (value) function towards their intended goal. By reasoning about the observed trajectory, generated as a result of user commands, we should be able to compute the likelihood of each possible goal (object) \( g_o \) given the value function \( c_{g_o} \) optimized by the intent-driven user (agent). In practice, it is difficult to find the true value function of the human operator and we instead use an approximate surrogate. Formally, let \( \xi_{X \rightarrow Y} \) denote a trajectory starting at pose \( X \) and ending at \( Y \). Using the principle of maximum entropy \([50, 51]\), we compute the probability of a trajectory for a specific goal (object) \( g_o \) as \( p(\xi\mid g_o) \propto \exp(-c_{g_o}(\xi)) \) trajectory decreases exponentially with cost. Following \([11]\), we use a first order approximation to address the difficulty in in computing the normalizing factor, the partition function, for the above conditional distribution. This gives us the probability over trajectories:

\[
p(\xi_{S \rightarrow E} \mid G) = \frac{\exp(-c_{g_o}(\xi_{S \rightarrow E}^* - c_{g_o}(\xi_{E \rightarrow G}^*))}{\exp(-c_{g_o}(\xi_{S \rightarrow G}^*))},
\]

where \( \xi_{X \rightarrow Y}^* \) is the optimal (minimum-cost) trajectory, \( S \) is the starting pose of the end-effector, \( E \) is the current end-effector pose, and \( G \) is the pose of the goal (object). Finally, Bayes’ rule gives the desired probability per goal:

\[
p(G \mid g_o) \propto p(\xi_{S \rightarrow E} \mid G)p(G)
\]

with a prior over goals \( p(G) \). Enumeration over discrete, finite goals normalizes the distribution. In the experiments discussed in Section \([13]\), we showcase the potential of this approach for multiple-object manipulation under BCI control. For these trials, \( c_{g_o} \) is an Euclidean distance between \( S \) and \( E \) and initializes with a uniform prior over the objects; however, this cost function may alternatively be learned from demonstrations \([50]\). At every intent-inference loop iteration, the optimal goal pose:

\[
G^* = \arg \max_{G \in \mathcal{G}} P(G \mid \xi_{S \rightarrow E})
\]

is selected. The final pose \( A \) is computed using the capture envelopes on the objects identified by \( G^* \) as discussed above.
C. Human-Robot Arbitration

Given an automated desired pose $A$ and the user commanded velocities $v_u$ (see Section II-A), the system needs to blend the two commands to generate new robot joint velocities. First, a desired pose $U \in SE(3)$ is generated from the user commanded velocities $v_u$ and the current end-effector pose $E$. Second, an arbitration scheme is needed. We decided to follow the suggestions of previous user studies [18, 11, 47] and keep the human operator in control to the largest extent possible and smoothly increase assistance in certain scenarios based on the confidence of the estimated intent. The arbitration between user commands and autonomous robot control is realized by a linear blending function [45, 5, 11]:

$$D = (1 - \alpha)A + \alpha U$$

where $\alpha$ is an arbitration factor that defines the amount of control given to the user. $\alpha = 1$ means the user has full control and $\alpha = 0$ means that robot assistance takes over completely. To compute $\alpha$ we use a sigmoid function that allows a smooth, continuous blending between the user and robot command:

$$\alpha = \frac{1}{1 + e^{-\sigma(I - t) + \sigma}}$$

where $I$ defines the confidence of the intent and $a$ and $o$ are parameters that ensure that $\alpha$ is in the range $[\alpha_{\text{min}}, 1]$. The value of $\alpha_{\text{min}}$ defines the minimal control contribution of the user and is adapted to the needs and ability of an individual user. Finally, to ensure the user’s ability to regain control in scenarios where there is a large discrepancy between the system’s assistive policy and the user’s command, the value $\alpha$ is increased above a safety threshold, allowing the user to break away.

In the experiments discussed in Section III and IV, the confidence of the intent $I$ is computed based on the progress of the end-effector towards a goal: $I = 1 - t$.

D. Computer Vision for Environment Perception

To achieve context-sensitive assistive teleoperation, the system requires knowledge about specific objects and obstacles present in the scene. Our perception module tackles both object recognition and localization in tabletop scenes using depth-image template matching, similar to the approach presented in [6] and the silhouette matching of [41]. Depth image template matching leverages efforts in 2D image template matching from the computer vision literature [25] while utilizing the scale disambiguation offered by range data.

We first filter out the planar supporting-surface (table) points with RANSAC. To prevent false detections of the robot arm, we first correct for joint errors then remove corresponding pixels in the depth image [20, 12]. The resulting depth image is segmented to find the potential locations of objects. Subsequently, the location of each object compared to pre-generated depth-image templates for each object type (e.g., block, ball, canteen, etc.) in various poses, with assignment to the best scoring template. The templates are created using 3D models from the model library, acquired using a 3D laser scanner. Due to sensor noise and matching error, iterative closest point (ICP) matching is used to finalize the object’s pose. Using this method, our system can recognize and localize multiple, physically separated objects of the same or different types.

E. Safe and Compliant Control

Given a desired motion in task space, the servo controller generates the required joint velocities by minimizing a cost function balancing the joint motion and error in the translation and rotation of the end-effector [30, 31, 37]. Due to redundancy in many robot manipulators, we can further add a secondary objective in the nullspace of the manipulator Jacobian [38]. In our servo controller, we aim to prevent collisions of the arm with the environment by using the nullspace objective to bias the arm towards preferred configurations away from the table. We also add a quadratic hinge penalty on the distance of the the elbow to the table applied in the nullspace of the joints above the elbow.

The lack of haptic feedback for external contact and incorrect or erratic movement commands from the operator can result in dangerous interactions with the world. We rely on software-based compliant control from force and torque sensors built into the robot to help prevent damage. The compliance provided by the joint velocity control software is realized by a constant monitoring of the operational space forces and torques from a sensor at the wrist of the robot arm. We apply intended force corrections to the joints using a procedure similar to [17, 44]. Since our joint-velocity PID controller uses the integral gain term, our end-effector compliance is implemented by computing a joint configuration offset to remove unwanted forces and torques applied at the end-effector $F_e$. The offset $q_e$ is given by $q_e = K J^T F_e$, where $K$ is a tuned gain term related to the proportional gain term in the velocity controller, and $J^T$ is the transpose of the Jacobian around the current joint configuration. Since the wrist-based force-torque sensor cannot detect collisions of the arm above the wrist, the described joint-offset controller becomes ineffective. To compensate, we introduce a stall detection mechanism within the control loop based on the build-up of the integral gain term. When a threshold is crossed, the controller forces for relevant joints are ramped down, releasing pressure at the contact points. The various thresholds, gains, and reference force-torque vectors are task dependent and stored in the model library. For example, activating a door lever handle requires a specific relaxation of forces along certain axis compared to free-motion, grasping, or pouring from a glass.

III. EXPERIMENTS

We evaluated the autonomy infused shared-control teleoperation framework on case-specific adaptations of two common rehabilitation benchmarks: the Action Research Arm Test [48, 26] and Box and Blocks [28, 46]. In addition to these two single object manipulation tasks, we tested a multi-object setup where the subject had to grasp one of two objects on the table corresponding to the one indicated by the experimenter to test the maximum entropy based user intent inference (Section II-B). During all experiments, two subjects controlled 4 DoF (3 DoF for end-effector translation velocity in task space and...
another DoF for a grasp velocity to open and close the hand) via BCI. A video showing the performance of the system can be found in the supplementary material of this paper.

A. Experimental Setup

The robotics platform used as a prosthetic consisted of a RGB-D camera mounted on a two-stage, four-axis neck, a seven DoF Barrett WAM arm equipped with a three-axis force-torque sensor at the wrist, and a four DoF three-fingered BarrettHand with pressure sensors in the palm and each finger. In all tests, a table was located directly in front of the robot. The subject was positioned next to the table and had a direct view of the task space (see Fig. 1).

We tested our setup with two tetraplegic subjects who had been using an intracortical microelectrode (Blackrock Microsystems, Salt Lake City UT) BCI for 2.5 years (Subject 1) or 2 months (Subject 2). While Subject 1 was able to practice the tasks prior to and during the development of the system, Subject 2 did not perform any of the tasks prior to the experimental sessions. This study was conducted under an Investigational Device Exemption (IDE) granted by the US Food and Drug Administration and was approved by the Institutional Review Boards at the University of Pittsburgh, the Space and Naval Warfare Systems Center Pacific, and Carnegie Mellon University. This trial is registered on clinicaltrials.gov (http://clinicaltrials.gov/ct2/show/NCT01364480). Both participants provided informed consent prior to participation.

To extract the user commanded velocities, intracortical recordings were made from two microelectrode arrays in Subject 1 and from four arrays in Subject 2. At the beginning of each session, thresholds were set on each channel at −4.5 times the root-mean-square (RMS) voltage for Subject 1, and -5.25 times RMS for voltage for Subject 2. The number of threshold crossings on each channel was recorded every 30 ms to generate a control signal that refreshed at 33 Hz for Subject 1. Threshold crossings were binned every 20 ms for Subject 2, resulting in an update rate of 50 Hz. The threshold crossings were filtered using a 450 ms window for Subject 1 and a 440 ms window for Subject 2. This filtered signal was used to decode the intended endpoint and grasp velocities using an indirect optimal linear estimator (OLE) decoder. It was trained prior to testing using the two step calibration method described in [8, 46].

The experiments were performed in two modes: Autonomy Infused Teleoperation (AIT) and Direct Control (DC). In the DC mode, decoded taskspace velocities were directly used for the servo and hand motion with almost no assistance. However, to ensure the safety of the robot, we applied workspace limits, elbow avoidance, and compliance control (Section II-E). In the AIT mode, we applied autonomous manipulation assistance utilizing computer vision, intent inference, and human-robot arbitration (Section II) and the safety measures used in DC.

Additionally, a hand control scheme that assisted the user with opening and closing motions during certain manipulation tasks was implemented. In particular, for the AIT experiments the following assistance was given: (i) when entering the capture envelope, the system ensures that the hand/finger positions are set as specified by the object grasp library and (ii) to avoid early grasp initiation, the system suppresses the user’s hand close signal in the capture envelope if the hand is not sufficiently close to the preset grasp position. Furthermore, the autonomy assist applied a squeeze while holding an object until it received a user commanded a release signal exceeding a pre-defined threshold.

B. Action Research Arm Test (ARAT)

To test the abilities of the subjects to control the robot arm via a BCI with Autonomy Infused Teleoperation, we used a subset of the Action Research Arm Test (ARAT) [48, 26]. ARAT was developed as a standardized test to assess upper limb impairments following a stroke. In the adapted experiments, the subjects had to grasp four different sized blocks (2.5 cm, 5 cm, 7.5 cm, 10 cm) and a ball. Starting from a neutral position, the subject was required to reach for the object, grasp it, and transport it from the left side of the workspace to a raised surface on the right side of the workspace (see Fig. 4). The task had to be completed within 2 minutes. If an object was dropped within the reachable workspace, the subject was allowed to re-grasp it. Otherwise, the trial was treated as timed out. An experimenter corrected any movement of the release platform by the robot arm.

The subjects were asked to perform the task three times in a row for each object in each control mode, AIT and DC. The subjects were not made aware of the operating mode of the robot. They were notified that each of the three attempts would be counted in their final score. After completing a group of three trials for a mode, the subject was asked to rate the difficulty of the task on a scale from 1 to 10: 10 if the task was extremely difficult to perform and 1 if it was extremely easy. The times and videos of each trial were recorded.
Fig. 5: Comparison of end-effector positions for the ARAT task for two sample trials. Observe the simpler trajectories with computer guided assistance. The Direct Control trials showcase the difficulty of standard BCI manipulation: noisy control, unintended early releases, and unstable erratic movement while grasping. As shown in Table I, this results in overall shorter times for successful task completion with AIT. Note that we show a 2D projection of the 6D end-effector motion onto the table surface.

| Subject | Object | Success Rate (n=3) | Completion Time (s) | Time to Grasp (s) | Number of Drops | Difficulty Rating (1-10) |
|---------|--------|-------------------|--------------------|------------------|----------------|------------------------|
|         |        | AIT | DC | AIT | DC | AIT | DC | AIT | DC | AIT | DC | AIT | DC | AIT | DC |
| Subject 1 | Block 10cm | 66% | 0% | 31 | - | 6.33 | 19 | 1.0 | 3.6 | 4 | 8 |
|          | Block 7.5cm | 100% | 0% | 9.2 | - | 3 | 28 | 2.0 | 1.0 | 1 | 9 |
|          | Block 5cm | 100% | 0% | 7.9 | - | 4.5 | 19.3 | 0.0 | 3.0 | 1 | 9 |
|          | Block 2.5cm | 100% | 0% | 42 | - | 4.6 | 5.3 | 0.0 | 3.5 | 3 | 9 |
|          | Ball | 66% | 0% | 42.3 | - | 21 | 10.5 | 1.0 | 1.0 | 3 | 9 |
| Subject 2 | Block 10cm | 100% | 0% | 30 | - | 3.5 | 5 | - | - | 4 | 6 |
|          | Block 7.5cm | 50% | 0% | 46.43 | - | 2.5 | - | - | - | 4.5 | /5 |
|          | Block 5cm | 66% | 0% | 26.6 | - | 15 | - | 0.3 | - | 2 | 9 |
|          | Block 2.5cm | 100% | 0% | 40.63 | - | 25.67 | - | - | - | 1 | 9 |

TABLE I: ARAT benchmark comparison with BCI implanted subjects with Autonomy Infused Teleoperation (AIT) and with Direct Control (DC). The data is averaged from three trials in a single session with an exception for Block 7.5 for Subject 2, averaged over two sessions with three trials each. No completion time could be reported for the DC experiments as there were zero successful trials. Also note the reduction in the time to grasp and the number of drops.

Results: The results of the test are summarized in Table I. A visual comparison of the trajectory of the robot’s end-effector in both modes for two objects is shown in Fig. 5. In the following section, we will discuss each subject individually.

The trial completion time of Subject 1 with the AIT mode varied between 7.1 s and 72.1 s. In the AIT mode, two out of 15 trials were failures. In both cases, the subject grasped the object successfully, but released it too early causing the object to go out of reach. With DC, the subject was unable to perform the complete task in any of the 15 trials. Here, the subject pushed the object away before grasping, could not stabilize the position long enough to complete the grasp, and after a successful grasp was unable to transport the object to the release platform without dropping it (see Fig. 5a and Fig. 5b.

Subject 1 grasped the object successfully in every trial with AIT. The average time in this mode of the first successful grasp was less than 6.4 s except for the ball. Here, the subject moved the object in a hard-to-reach area while approaching it from the side, requiring another grasp attempt in that area. The average time for the first successful grasp with DC varied between 5.3 and 28 s with an overall average of 16 s (13 of 15 trials grasped successfully). These results indicate the ability of AIT to reduce the time until first object possession. AIT was also able to reduce the number of drops after grasping for Subject 1. After each drop the subject was able to re-grasp the object quickly if the object was in reach for the manipulator. Subject 1 reported an average perceived difficulty of 2.4 when using the Autonomy Infused Teleoperation in comparison to the 8.8 average difficulty under Direct Control.

Subject 2 was not able to grasp any objects in the 15 trials with DC. Often, the subject hovered close to the object without being able to stabilize or was unable to approach the object. With AIT, Subject 2 was able to grasp the object in 13 of 15 trials. The subject was able complete 61.5% of the successful grasps trials. The task completion time varied between 11.5 s and 74 s. AIT reduced Subject 2’s average perceived difficulty from 7.8 with DC to 3.2.

C. Object Transfer Task: Box and Blocks

We performed an object transfer task based on the Blocks and Blocks experiment [23, 46]. This test required the subject to transport a 7.5 cm block from one side of the workspace to the other as many times as possible in two minutes. The subject was required to lift the block to at least a specific height given by an object placed at the edge of the table within the subject’s view. After each successful transfer, the object was reset by the experimenter until the 2 minute time limit was reached. Upon completion of one trial, the subject was asked to rate the difficulty on a scale from 1 to 10. The task was performed in each of the two modes AIT and DC without the subject being made aware of the mode of the trial.

Subject 1 performed the experiment over two different sessions and no trials were run with Subject 2. During the first session, the subject performed the task three times in each mode. During the second session, the subject performed the task two times in each condition. The experiments during the second session were additionally modified in three ways. 1) A low-pass filter was implemented on the grasp signal for the AIT mode. As a result, the robot only released the object for a continuous strong release signal. 2) The subject was asked...
to start the movement towards the object after the hand of the experimenter left the object in order to allow the computer vision system to detect the object. 3) The object was put at a random position on the left side of the workspace when placing it back.

**Results:** Table ?? summarizes the results of both sessions. In Session 1, the results were inconclusive. Here, the subject moved the arm to the position where the block was positioned after each successful transfer before the object was placed there. As a result, the arm was blocking the view of the object, or moving it constantly before the computer vision module could locate the object. In the second session, the subject was instructed to wait before moving the arm. Here, the AIT increased the successful transfers and decreased the difficulty rating compared to DC.

**D. Multi-Object Grasping with Intent Inference**

To evaluate whether the architecture can successfully infer intent in a multi-object setting, two 7.5 cm blocks were placed in various configurations 10 to 31 cm apart from each other. The subject was instructed to grasp one of the two objects chosen by the experimenter (yellow or wooden block). The experiment was performed only with Subject 2. Starting from a neutral position, the subject had to reach for the object, grasp it, and lift it from the table. The experiment was repeated 36 times solely with AIT as single object grasp results had been unsuccessful with DC. The subject was able to grasp 32 of 36 experimental indicated blocks successfully. Two failures were due to the object being moved out of reach, one failure was due to the subject being unable to lift the arm after grasping and one failure was due to a wrong object selection. The grasping time varied between 4.5 s and 110 s – averaging 17.61 s over the trials.

**IV. Extensions**

We show the extensibility of our architecture to other everyday living tasks such as pouring, door opening and manipulating unknown objects and the generality to multiple teleoperation control interfaces. While door opening was performed with Subject 1, the other two tasks were performed by an experienced operator using a game controller that could be operated with either the dual-joysticks or with its 6-DoF motion tracking capabilities.

**A. Door Opening**

To explore using the BCI together with assistive teleoperation, we investigated the task of door opening with Subject 1.

---

**TABLE II: BCI object transfer (Box and Blocks) benchmark results**

|   | Transfers | Difficulty rating (1-10) |
|---|-----------|--------------------------|
|   | AIT       | Direct Control           |
| Session 1 | 7         | 6                        |
|         | 2         | 6                        |
|         | 6         | 6                        |
| Session 2 | 4         | 1                        |
|         | 2         | 2                        |
|         | 7         | 0                        |

Average 5.2 2.6 3 6.6

---

Fig. 6: Extending the model library allows the user to engage in object-specific affordances such as pouring when grasping a soda can (Fig. 6a) and opening and closing a door when grasping the handle (Fig. 6b).

**B. Pouring**

The lack of robust rotation control with low-dimensional input complicates common daily living tasks, such as pouring from a glass. We augment our framework through the model library to assist a user with the required rotational motion and stability when near “fillable” containers. Specifically, we extended our approach to allow pouring from a soda can, glass or canteen into a bowl (Fig. 6a). We extend the approach of grasp poses in the motor library, by defining pouring poses that become active when approaching such a container with a “pourable” object. The end-effector is guided to the pour pose using capture envelopes (Section II-A). At the pouring pose, the user translational commands are mapped to corresponding rotational velocities for pouring. When the user breaks-away, the system ensures that the “pourable” object is re-oriented upright before allowing for translations away. An experienced operator controlled the arm using a dual-joystick game controller. The task involved grasping the “pourable” object, moving it to a bowl, and returning the object to the table. In all 10 trials, the contents were successfully poured into a bowl. In one trial, the object was not properly returned to the table after pouring.

**C. Beyond the Model Library: Novel Objects**

The computer vision component Section II-D and grasp inference methodology Section II-A relied upon prior knowledge of object models and preselected, predefined grasp poses. Though this approach enables greater reliability or specifics in regards to object grasping (e.g. grasp a mug by the handle), the introduction of novel objects exposes a potential flaw in singly relying upon access to a fixed library. The architecture
Vogel et al. address the lack of neural haptic feedback through certain limitations preventing seamless teleoperation via BCIs. These limitations have been evident in traditional teleoperation applications for upper limb prosthetics [14, 8]. However, BCIs, such as intracortical BCIs, provide superior bandwidth and have been used for continuous high degree of freedom control of upper limb prosthetics [35]. Embracing a variety of input methods such as intracortical arrays, binary switches [21], or a linear blend between the two, the operator can arbitrate between the user command and autonomous robot control. The integration of robot sensor data for compliant control enables safe interaction with the environment [43]. We achieve a similar effect through joint-stall detection and Jacobian transpose compliant control from a wrist mounted force-torque sensor. For BCI manipulation with non-human primates, [19] uses simple optical sensors to provide obstacle avoidance and grasping assistance. We extended this idea using a more complex computer vision system and model library augmented grasp inference and showed results on human subjects.

VI. Conclusion

Our shared-control assistive teleoperation framework provides an intuitive and responsive system, overcoming erratic, noisy, and low-dimensional user inputs by i) finding environmental context through computer vision ii) inferring the user’s intent from their motion commands, and by iii) dynamically arbitrating between the user command and autonomous robot control. In experiments with two subjects implanted with intracortical BCIs, we demonstrated the capability of autonomy infused user control for achieving a high dexterity level in manipulation tasks. This allowed the subjects to complete tasks that were unsuccessful through direct control.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding under the Defense Advanced Research Projects Agency’s (Arlington, VA, USA) Autonomous Robotic Manipulation Software Track (ARM-S) program and the Revolutionizing Prosthetics program (contract number N66001-10-C-4056). The material presented in this paper is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1252522. This study was completed under an investigational device exemption granted by the US Food and Drug Administration. We thank the study participants for their dedication and insightful discussions with the study team. The views expressed herein are those of the authors and do not represent the official policy or position of the Department of Veterans Affairs, Department of Defense, National Science Foundation, or the US Government. We thank Sidd Srinivasa for helpful conversations and Pedro Mediano for his work on the “Dragonfly” software bridge that enabled this effort.

Fig. 7: Using a model-free object perception and grasp point detection algorithms allows us to augment the shared-control framework to handle situations with novel objects (7a). Supervoxel segmentation (7b) and forward simulation of the robot hand (7c), are used to create grasp fixtures (7d). Clutter clearing arises when trying to find hidden objects lying underneath (7e). Due to limited time with the BCI, we were only able to show this extension as a proof-of-concept using a motion tracking game controller.
REFERENCES

[1] D. Aarno and D. Kragic. Motion intention recognition in robot assisted applications. Robotics and Autonomous Systems, 56:692–705, 2008.

[2] D. Aarno, S. Ekvall, and D. Kragic. Adaptive virtual fixtures for machine-assisted teleoperation tasks. In IEEE international conference on robotics and automation, 2005.

[3] P. Aigner and B. McCarragher. Human integration into robot control utilising potential fields. In IEEE International Conference on Robotics and Automation Albuquerque, pages 291–296, 1997.

[4] R. Ambrose, H. Aldridge, R.S. Askew, R. Burridge, W. Blaethmann, M. Diftler, C. Lovchik, D. Magruder, and F. Rehmark. Robonaut: NASA’s space humanoid. IEEE Intelligent Systems, 15(4):57–63, 2000.

[5] S. Anderson, S. Peters, and K. Iagnemma. Semi-autonomous stability control and hazard avoidance for manned and unmanned ground vehicles. In 27th army science conference, 2010.

[6] J.A. Bagnell, F. Cavalcanti, L. Cui, T. Galluzzo, M. Hebert, M. Kazemi, M. Klingensmith, J. Libby, T.Y. Liu, N. Pollard, M. Pivtoraiko, J.-S. Valois, and R. Zhu. An integrated system for autonomous robotics manipulation. In IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 2955–2962, October 2012.

[7] A. Boularias, J. Andrew Bagnell, and A. Stentz. Efficient optimization for autonomous robotic manipulation of natural objects. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, pages 2520–2526, 2014.

[8] J.L. Collinger, B. Wodlinger, J.E. Downey, W. Wang, E.C. Tyler-Kabara, D.J. Weber, A.J. McMorland, M. Veliste, M.L. Boninger, and A.B. Schwartz. High-performance neuroprosthetic control by an individual with tetraplegia. The Lancet, 381:557–564, 2013.

[9] Jacob W Crandall and Michael A Goodrich. Characterizing efficiency of human robot interaction: A case study of shared-control teleoperation. In Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on, volume 2, pages 1290–1295. IEEE, 2002.

[10] M. Desai and H. Yanco. Blending human and robot inputs for sliding scale autonomy. In International workshop on robot and human interactive communication, pages 537–542, 2005.

[11] A. Dragan and S. Srinivasa. A policy-blending formalism for shared control. The International Journal of Robotics Research, 32:790–805, 2013.

[12] D. Grest, J. Woetzel, and R. Koch. Nonlinear body pose estimation from depth images. In Pattern Recognition, pages 285–292. Springer, 2005.

[13] K. Hauser. Recognition, prediction, and planning for assisted teleoperation of freeform tasks. Autonomous Robots, 35:241 – 254, 2013.

[14] L.R. Hochberg, D. Bacher, B. Jarosiewicz, N.Y. Masse, J.D. Simeral, J. Vogel, S. Haddadin, J. Liu, S.S. Cash, P. van der Smagt, and J.P. Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature, 485(7398):372–375, 2012.

[15] K.D. Katyal, M.S. Johannes, S. Kellis, T. Aflalo, C. Klaes, T.G. McGee, M.P. Para, Y. Shi, B. Lee, K. Pejsa, C. Liu, B.A. Wester, F. Tenore, J.D. Beaty, A.D. Ravitz, R.A. Andersen, and M.P. McLoughlin. A collaborative bci approach to autonomous control of a prosthetic limb system. In Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on, pages 1479–1482. IEEE, 2014.

[16] D. Katz, A. Venkatraman, M. Kazemi, J.A. Bagnell, and A. Stentz. Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. In Robotics: Science and Systems Conference, June 2013.

[17] O. Khatib and J. Burdick. Motion and force control of robot manipulators. In Robotics and Automation. Proceedings. 1986 IEEE International Conference on, volume 3, pages 1381–1386. IEEE, 1986.

[18] D. Kim, R. Hazlett-Knudsen, H. Culver-Godfrey, G. Rucks, T. Cunningham, D. Portee, J. Bricout, Z. Wang, and A. Behal. How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 42(1):2–14, 2012.

[19] H.K. Kim, J. Biggs, D.W. Schloerb, J.M. Carmena, M.A. Lebedev, M. Nicolelis, and M.A. Srinivasan. Continuous shared control for stabilizing reaching and grasping with brain-machine interfaces. Biomedical Engineering, IEEE Transactions on, 53(6):1164–1173, 2006.

[20] M. Klingensmith, T. Galluzzo, C. Dellin, M. Kazemi, J.A. Bagnell, and N. Pollard. Closed-loop servoing using real-time markerless arm tracking. In International Conference on Robotics And Automation (Humanoids Workshop), May 2013.

[21] J. Kofman, X. Wu, and T. Liu. Teleoperation of a robot manipulator using a vision-based human-robot interface. IEEE Transactions on Industrial Electronics, 5:1206–1219, 2005.

[22] D. Kortenkamp, D. Keim-Schreckenghost, and R.P. Bonasso. “adjustable control autonomy for manned space flight”. IEEE aerospace conference proceedings”, pages 629–640, 2000.

[23] Q. V Le, D. Kamm, A. F Kara, and A. Ng. Learning to grasp objects with multiple contact points. In IEEE International Conference on Robotics and Automation (ICRA), pages 5062–5069. IEEE, 2010.

[24] A.E. Leeper, K. Hsiao, M. Ciocarlie, L. Takayama, Z. Wang, and A. Behal. How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 42(1):2–14, 2012.

[25] R.C. Lyle. A performance test for assessment of upper
limb function in physical rehabilitation treatment and research. *International Journal of Rehabilitation Research*, 4:483–492, 1981.

[27] P. Marayong, M. Li, A. Okamura, and G. Hager. Spatial motion constraints: Theory and demonstrations for robot guidance using virtual fixtures. In *2003 IEEE International Conference of Robotics and Automation*, pages 1954–1959, 2003.

[28] V. Mathiowetz, G. Volland, N. Kashman, and K. Weber. Adult norms for the box and block test of manual dexterity. *The American Journal of Occupational Therapy*, 39:386–391, 1985.

[29] D.P. McMullen, G. Hotson, K.D. Katyal, B.A. Wester, M.S. Fifer, T.G. McGee, A. Harris, M.S. Johannes, R.J. Vogelstein, A.D. Ravitz, W.S. Anderson, N.V. Thakor, and N.E. Crone. Demonstration of a semi-autonomous hybrid brain-machine interface using human intracranial eeg, eye tracking, and computer vision to control a robotic upper limb prosthetic. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):784–796, July 2014.

[30] R.M. Murray, Z. Li, and S.S. Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 1994.

[31] J. Nakanishi, R. Cory, M. Mistry, J. Peters, and S. Schaal. Operational space control: A theoretical and empirical comparison. *The International Journal of Robotics Research*, 27(6):737–757, 2008.

[32] M. Palankar, K.J. De Laurentis, R. Alqasemi, E. Veras, R. Dubey, Y. Arbel, and E. Donchin. Control of a 9-dof wheelchair-mounted robotic arm system using a p300 brain computer interface: Initial experiments. In *Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on*, pages 348–353. IEEE, 2009.

[33] S. Park, R.D. Howe, and D.F. Torchiana. Virtual fixtures for robotic cardiac surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 1419–1420. Springer, 2001.

[34] A. Saxena, J. Driemeyer, and A. Ng. Robotic grasping of novel objects using vision. *The International Journal of Robotics Research*, 27(2):157–173, 2008.

[35] A.B. Schwartz, D.J. Weber, and Moran D.W. Brain-controlled interfaces: movement restoration with neural prosthetics. *Neuron*, 52(1):205–220, 2006.

[36] Thomas B Sheridan. *Telerobotics, automation, and human supervisory control*. MIT press, 1992.

[37] B. Siciliano and O. Khatib. *Springer handbook of robotics*. Springer, 200.

[38] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo. *Robotics: Modelling, Planning and Control*. Springer Publishing Company, Incorporated, 1st edition, 2008. ISBN 1846286417, 9781846286414.

[39] J.D. Simeral, S.P. Kim, M.J. Black, J.P. Donoghue, and L.R. Hochberg. Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array. *Journal of neural engineering*, 8(2):025027, 2011.

[40] I.E. Sutherland. Sketchpad: A man-machine graphical communication system.

[41] A. Toshev, A. Makadia, and K. Daniilidis. Shape-based object recognition in videos using 3d synthetic object models. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 288–295. IEEE, 2009.

[42] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A.B. Schwartz. Cortical control of a prosthetic arm for self-feeding. *Nature*, 453(7198):1098–1101, 2008.

[43] J. Vogel, S. Haddadin, J.D. Simeral, S.D. Stuvisky, D. Bacher, L.R. Hochberg, J.P. Donoghue, and P. van der Smagt. Continuous control of the dlr light-weight robot ii by a human with tetraplegia using the braingate2 neural interface system. In *Experimental Robotics*, pages 125–136. Springer, 2014.

[44] R. Volpe and P. Khosla. A theoretical and experimental investigation of explicit force control strategies for manipulators. *Automatic Control, IEEE Transactions on*, 38(11):1634–1650, 1993.

[45] C. Weber, V. Nitsch, U. Unterhinninghofen, B. Faerber, and M. Buss. Position and force augmentation in a telepresence system and their effects on perceived realism. In *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 226–231, 2009.

[46] B. Wodlinger, J.E. Downey, E.C. Tyler-Kabara, A.B. Schwartz, ML Boninger, and J.L. Collinger. 10 dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations. *Journal of Neural Engineering*, 2015.

[47] E. You and K. Hauser. Assisted teleoperation strategies for aggressively controlling a robot arm with 2d input. In *Robotics: Science and Systems*, 2011.

[48] N. Yozbatiran, L. Der-Yeghiaian, and S.C. Cramer. A standardized approach to performing the action research arm test. *Neurorehabil Neural Repair*, 22:78–90, 2008.

[49] W. Yu, R. Alqasemi, R. Dubey, and N. Pernalete. Telemanipulation assistance based on motion intention recognition. In *IEEE International Conference on Robotics and Automation*, pages 1121 – 1126, 2005.

[50] B.D. Ziebart, A. Maas, J.A. Bagnell, and A. Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, July 2008.

[51] B.D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peeters, J.A. Bagnell, M. Hebert, A. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *IEEE/RSJ IROS, 2009.*