Enhancing the Robot Avateering Metaphor Discreetly with an Assistive Agent and its Effect on Perception

Seng Lee Koh[1], Kevin Pfeil[2], and Joseph J. LaViola Jr.[3]

Abstract— We present a modeling approach to develop an agent that assists users discreetly in teleoperation when avateering a robot via an inexpensive motion sensor. Avateering is a powerful metaphor, and can be an effective teleoperating strategy. Avateering a Humanoid Robot (HR) with no wearable device encumberment, such as using the popular Kinect/NUI motion sensor, is also desirable and very promising. However, this control scheme makes it difficult for the slave robot to make contact and interact with objects with high accuracy due to factors such as viewpoint, individually-unique and unilateral human-side control imprecision, and lack of informative tactile feedback. Our research explores the addition of an assistive agent that arbitrates user input without disrupting the overall experience and expectation. Additionally, our agent assists with maintaining a higher level of accuracy for interaction tasks, in our case, a grasping and lifting scenario. Using the Webots robot simulator, we implemented 4 assistive agents to augment the user in avateering the Darwin-OP robot. The agent iterations are described, and results of a user study are presented. We discuss user perception towards the avateering metaphor when enhanced by the agent and also when unassisted, including perceived easiness of the task, responsiveness of the robot, and accuracy.

I. INTRODUCTION

Humanoid Robots possess enormous potential to perform as surrogates in telepresence or teleaction scenarios, especially in-lieu of real emergency personnel for disaster response under hazardous environments. Perusing the sixteen participants at DARPA Robotics Challenge (DRC) 2013, nearly every entrant (with the exception from Team CHIRON) involved humanoid-hybrids that include Team Schaft's winning HRP-2 based biped. Secondly, teleoperating humanoids with little or no device encumberment, such as with skeletal-joint tracking using Microsoft's Kinect or Asus's Xtion depth sensors, offer an interesting and cost-effective approach to allow for a more natural, intuitive, and non-obtrusive interface. With the close locomotion resemblance of HRs' shared with humans, teleoperating such robots with full-body human-motion and without encumberment is becoming an interesting research topic within the robotics community.

We observe from existing literature there are essentially three research directions that address the use of human-motion for humanoid teleoperation; The first covers primarily human-motion mapping and prototyping;

For instance, human-motion data represented as sets of discriminant features can be mapped to humanoid teleoperation commands [1][2][3][26][27]. In recent years, using the inexpensive Kinect motion sensor has been used to perform direct markerless human-imitation or manipulator control [4][7] without device encumberment. We can state these teleoperation techniques as puppeteering or avateering [5]. The second direction investigates design of usage metaphors from the Human-Robot Interaction (HRI) perspective [8][9], leveraging on latent human intuition to control the robot correctly. By inference, we can also metaphorize avateering; human-imitation is an effective teleoperational strategy as it provides intuitive control by leveraging on human-robot embodiment similarities. We can state that:

Avateering is a powerful metaphor that enables intuitive teleoperation of dexterous robots, especially if the robot, or manipulators themselves, are close models of the human locomotion system.

The third direction investigates the development of intelligent agents under the RAP system (Robots, Agents, and People) [12] to enhance the teleoperator's control of the robot. Especially when teleoperating with human-motion capture, the user interface associated with them are often inadequate, rendering simple manipulation tasks often tedious and sometimes impossible. Agents help users accomplish their tasks by predicting their intent, and using the predictions to augment their input into the robot [13]. In our own studies and to the best of knowledge, we observe it is non-trivial to design a teleoperation strategy that suitably applies the state-of-the-art from all directions. It turns out that often, the principles each purportedly contradict one another. For instance, avateering with motion capture on a HR can be intuitive, but the technique becomes cumbersome when attempting to accomplish simple manipulation tasks, such as grasping [11][17]. Agents can be applied to assist in the hardest tasks, such as manipulator position and force augmentation [10][13][27]. However, prediction and application of user intent to arbitrate user-input is more of an art, and sometimes more often than not, the agent has to take complete control of the robot for arbitrary time intervals. In these instances, agent assistance breaks the avateering metaphor, adversely affecting the controlling user's perception and expectation. It is hard to design an assistive agent that not only predicts what users may want, and how it may assist, but also arbitrate inconspicuously such that users may perceive they have full

[1]Department of Computer Science, University of Central Florida, Orlando FL, 32819 ksenglee@knights.ucf.edu
[2]Department of Computer Science, University of Central Florida, Orlando FL, 32819 kevin.pfeil@knights.ucf.edu
[3]Department of Computer Science, University of Central Florida, Orlando FL, 32819 jjl@eecs.ucf.edu

avateering control at all times using unencumbered motion capture.

We contribute an analysis of modeling such an agent that activates and assists discreetly at suitable timings to allow users to perceive avateering control, while teleoperating a HR to grasp, move, and release an object to a goal location. We constrain our discussion to humanoids with end-effectors that initiate non-prehensile grasping (no wrist), and state modeling axioms to build such an agent. For demonstration, we applied the axioms to our agent and present 4 techniques to assist a user avateering a DARWIN-OP robot to complete a simple trial. The trial is also set-up such that algorithms that aid force augmentation upon contact on the target object, such as force-feedback, will not be required.

II. RELATED WORK

Assistive teleoperation research has seen a great variety of methods proposed to augment their imprecise command inputs into robots. Dragan et. al. noted that research on this topic has ranged "from the robot having full control over all or some aspect of the motion, to taking control (or releasing it) at some trigger to never fully taking control" [13]. Usually, the assistance performed is premeditated, based on a pre-conceived notion about what user intent would be. For instance in [3], the robot takes full control of the manipulator to complete the grasp if it is close enough to the target. Dragan arbitrates user input controlling the manipulator's path by predicting the user's intended object to grasp.

Similarly, our work arbitrates user input to the motions of the robot arms; however, user intent for avateering is already implied by natural movements of user arms. We aim at predicting the user's target, but only execute arbitration before and upon grasping the object. This maintains the robot's accuracy, but also gives the user a sense of full control when manipulating the robot arms.

There are also many works that address force augmentation upon making contact with the target object, such as force-feedback, or modeling manipulator contact as an impedance-admittance to aid force correction upon the object being interacted with [29]. Kim et. al. [28] describes a shared compliant control used to tune the grasping force exerted by the manipulator so as to mitigate input error induced by communication time delays between local and remote sites. Weber [10] presents haptics assistance functions for a telepresence system that correct force and position of a screw-driver during a work-task controlled remotely by a teleoperator to assess their effect on stability and safety without negatively impacting realism. Our work here instead sets up the study such that adequate grasping force is applied on the object, but addresses the problem of detecting and executing user-intent of grasping, maintaining grasp on the object while in transit, and releasing.

The work in [18] details a method to enable autonomous robots to pickup specific objects. The algorithm involves a database of objects, which assists with determining the proper grasping technique. The discussed robot is designed to help in assistive scenarios, and is completely autonomous,

but its interaction is limited to objects it can recognize in its database. Other work by [19] describes methodology to enable autonomous robot arms to grasp novel objects, or objects that are not already known about. This helps by extending the usability of the grasping robot to more scenarios. The experiments indicated a high success rate for picking up certain objects of various shapes and sizes, but there were some that did see poor rates.

In both of the above papers, the robotic platforms were autonomous. However, assistive robots do not necessarily need to be autonomous. In [20], a household robot is described. It is primarily aimed at monitoring the elderly by allowing a remote user to drive around and inspect the person, including their biometrics. It is driven around using a waypoint system, and the authors have plans for incorporating an arm onto the platform. This type of robot would enable users to interact with the environment remotely while being in full control of the system. For instance, the remote user could bring medicine to the person, pickup objects from the floor, or find and fetch needed items.

Although humans may implicitly know how to interact with the environment, using robots as a surrogate agent through telepresence or remote control is still a difficult task, compared to manual interaction. [16] used adjustable autonomy to assist users in completing tasks involving robots. In their experiment, the user's skill level determined the robot's level of autonomy. Novice users received much help from the robot and were able to complete the task swiftly, but the sub-tasks they were less skilled in were handled by the robot. While this concept is useful to help reduce task completion time, it reduces the level of control that the user has over the robot.

Lewis et al used their robotic platforms for multiple scenarios, including [21], a robot team implementation, where two independent robots paired to lift objects. The robot arms featured a claw, which allows for a prehensile power grasp [22]. The grasping of objects was controlled by the user manually, and no assistive algorithm was used. When completing a task, one of the robots automatically drove near the other and mirrored the other robot's actions. For instance, when the operator raised one robot arm, the other also move upwards; when moving one robot forward, the other kept pace with it. This technique alleviated the user's cognitive load, and essentially turned two robots into one lifting entity.

[17] describes a technique to assist robot arms with a grasping task. In their scenario, a sensor tracks the user's arm motion. This input guides the robot arm towards an object. Since the user is in full control of the robot arm motion, there is no need for a vision system to assist in guiding the arm towards the goal. When the robot is near the object to pickup, the user makes a "clenching" gesture that invokes a computation of Inverse Kinematics for each digit on the manipulator wrist to grasp the object from its current position.

In the following section, we present modeling principles stated as axioms to design our agents, along with simplifying assumptions which enable their development tractable.
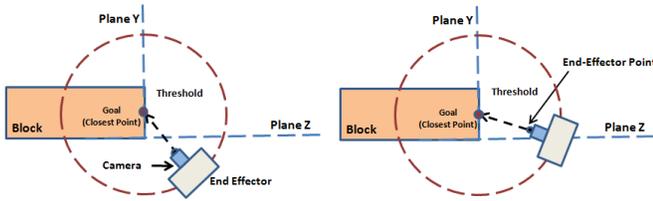
Fig. 1. The end-effector shown on the left acquires a potential goal point, but the agent is not activated as it does not meet the locality requirement, since it is located behind Plane Z. At the next tick, the user moves the end-effector that meets the locality requirement (in front of Planes Y & Z, and within the threshold), which then activates the agent to take control and move the end-effector to the last-acquired goal point.
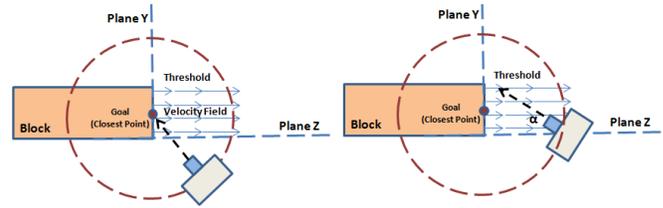

Fig. 2. Similar to Figure 1, end-effector shown on the left acquires a potential goal point, but the agent is not activated. At the next tick, the user moves the end-effector that meets the locality requirement, but the agent will only activate and take control if it approaches the object plane the point resides on at an angle smaller or equal to a set α, and its velocity exceeds a set threshold.

III. PROPOSED AGENT MODELING AXIOMS

We divide the agent's arbitration phases into two stages; i) Before, and ii) After making contact with the target object. We also state that the agent has perfect information about the operating space. For instance, the humanoid's end-effectors' would have eye-in-hand cameras for identifying and localizing possible target placements, and LiDAR for Point Clouds that extrapolate useful geometries from the environment, such as object planes and normals.

A. Before Contact

The main idea behind the agent letting the user maintain avateering control consists of determining its i) Timing of Activation, and ii) Type of Arbitration, based on predicting the users' expectation of the robot's manipulator trajectories around objects before and after making contact.

1) Agent Activation: A straight-forward heuristic: Target proximity, if localized appropriately, can be a simple and effective heuristic to determine where and when the agent should activate. Tracing a ray from the center of the eye-in-hand camera, the point it hits is usually the target (or the closest point). Figure 1 illustrates how we can take the ray-trace distance as an input to a Euclidean radial-distant activation threshold. The agent can activate whenever this threshold is breached. We state this heuristic with the following axiom:

Axiom (1) : The closest target point, based upon the pose of the manipulator (which directly affects the eye-in-hand camera pose), is usually, the intended target from the user.

Figure 1 also illustrates how the end-effector can be poorly located away from the region where the user would prefer the agent to be activated, which in this case is by the right side of the block. A better localization heuristic is to check further whether the end-effector is between planes Y and Z. Used together with a radial threshold, the agent can be activated where the end-effector is both properly localized and orientated towards a valid target point on the object.

2) Improving the base heuristic on activation timing: Axiom (1) addresses agent activation locality ('where') sufficiently, but not necessarily 'when'. Assume that upon activation, the agent takes full control and moves the end-effectors to their target placements, localization is not sufficient to hide the noticeable arbitration to a keen eye. It is further exacerbated when considering the operator's attentive field-of-view (FOV) would usually be upon the task-space regions where the agent activates. This observation postulates to two additional axioms as we model alternative agent arbitrative behavior:

Axiom (2) : The agent should blend its activation-arbitrative actions into the robot's current executing action or operating environment.

Axiom (3) : The agent should divert some or all activation-arbitrative actions away from the user's main focal point of attention (Sleight-Of-Hand Principle).

An application of Axiom (2) is to allow free avateering and delay activation till the agent detects an expected action triggered by the end-effector in order to make contact with the object; we can state this criteria as Just-In-Time (JIT) Activation. One such instance exists is where the user sends the end-effectors, at some reasonable velocity and angle to the normal of the object planes, towards their target points. The agent takes control only when the end-effector velocity with respect to the plane exceeds a stated threshold. Additionally, the corrective action it performs also blends and conceals within the momentum moving the end-effector towards the plane. We can model a threshold as a velocity vector field, growing at increasing length from the surface using its plane normal (Figure 2). The changing field implicitly models how a user would approach the object surface, given that the end-effector velocity usually does reflect a user's level of intent about approaching the object; For instance, a user who feels confident enough would send the end-effectors towards the object plane from far at greater speeds, and at a vector that is closely parallel to the plane normal (e.g smaller than α, Figure 2). At slow speeds, the agent can activate (if the dot-product threshold is met) since the velocity-field vector gets shorter with decreasing distance to the object contact

surface. Linear and rotational velocities of the end-effectors can be computed using the geometric Jacobian with their associated Denavit-Hartenburg (DH) coordinate frames, and the resulting dot product between the end-effector's linear velocity and velocity-field vector dependent upon its position in the activation space serves as input to decide activation. For instance using Figure 2, once the locality requirement is met by the end-effector, when the dot-product computation between the end-effector's linear velocity and velocity-field vector it intersects inside the field exceeds a set value, the agent will take full control at the next time-step.
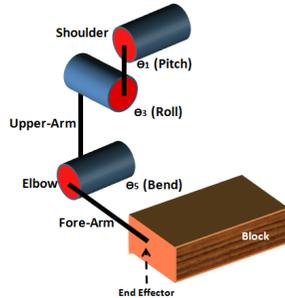

Fig. 3. A 3 Degree-Of-Freedom Arm Manipulator applying non-prehensile grasping at the side of the block

3) Alternative Arbitration Strategy After Activation: Axiom (2) models the agent to surreptitiously arbitrate user input by quickly taking and relinquishing full control when the end-effector meets the activation requirement based on locality and approach. Axiom (3) models an alternative input mediation policy that hypothesizes i) whether the assistive agent, upon activation, can allow the user a measure of full control (e.g. user can voluntarily move the end-effector towards or away from the target points) but maintains its arbitration role through partial control, ii) the option of deactivating the agent, and iii) whether the agent can accomplish both i) and ii) without the user noticing. One idea is to leverage on the redundancy of the manipulator, letting the user control the most significant degrees-of-freedom (DOF) that accomplishes the task from the user's point of view that dominates attention, and letting the agent control the remaining DOF to correct the trajectory error. Figure 3 illustrates such an example using a DARWIN-OP arm with 3DOF. Upon meeting the agent activation requirement, it is sufficient for the user to control only θ_3 (roll) in order to send the end-effector to make contact with the right side of the block, while the agent sets the required pitch (θ_1) and elbow bend (θ_5). Additionally, if the user rolls the manipulator to move the end-effector away from the activation threshold, the agent deactivates and the user regains full DOF control. The manipulator pose, end-effector location, and nature of the task (in this case, applying contact to the sides of the block) also helps to conceal the arbitrative action the agent performs, as the block constrains the available operating space the end-effector can move. Of course, the assumption

that the user is sufficiently attentive to the manipulator roll operation also helps to perpetuate the concealment.

B. After Contact

User intent is hard to predict after the end-effector makes contact. We state the following:

Axiom (4) : The agent should assist in maintaining a suitable contact force between the end-effectors and object as the user desires. Otherwise, the user regains full avateering control.

In many task scenarios, it is very useful for the agent to help the user maintain an arbitrary force contact on the object based on the task objectives. Force input augmentation, such as force-feedback, can be suitably applied to maintain a steady grasping force on the object if the user desires so. However, we observed that users have difficulty holding on to the block with avateering (see Section: Agent & Trial Description), via many occasions accidentally moving the end-effectors away from the grasp points as they attempt to move the plank away from the crate with the humanoid. The main difficulty lies in determining a user's intent to either maintain or release grasp. A feedback system presenting a representation of the force contact on the target object, either visually or by tactile, can be helpful for the user. However, the lack of physical constraints in between the user's hands afforded by the real object manipulated remotely by the humanoid prevents the user from adequately ascertaining and maintaining the grasp points (Figure 4). There are many heuristics which can be applied to evaluate how the agent can assist here upon contact, and in fact, the agent can opt to do nothing and de-activate with the user having full control. One strategy however, is to take a holistic approach, and consider the user (or even robot) state and predict the next action the user chooses to take. Gesture and pose recognition with a-posteriori prediction can be encoded into the agent, for instance using HMM [14] and POMDP models [6], and based upon results from the prediction the agent can choose whether to maintain the contact or deactivate.

IV. AGENT & TRIAL DESCRIPTION

We developed a set of 4 agents using the modeling principles outlined from Axiom (1) to Axiom (4). Each agent corresponds to a distinctive technique, which we state as i) No-Degrees-Of-Freedom (No-DOF) ii) No-Degrees-Of-Freedom Just-In-Time (No-DOF-JIT) iii) One-Degree-Of-Freedom (1-DOF) and iv) Two-Degrees-Of-Freedom (2-DOF). We used Webots EDU to devise a simple trial to test our main hypotheses that:

Hypothesis 1: User perception of Avateering can co-exist with the presence of an assistive agent in the background that arbitrates the user's input.

Hypothesis 2: User can be agnostic to the presence of an assistive agent when the corrective action it performs
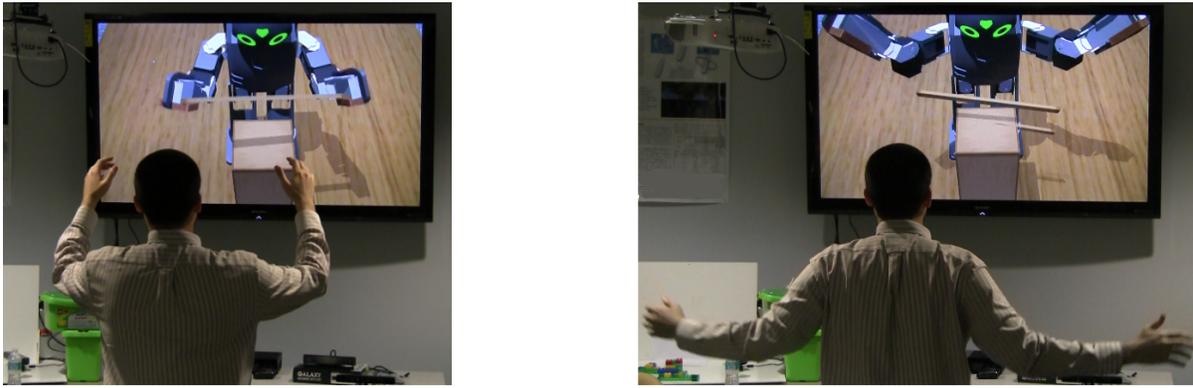
Fig. 4. Left Image: User attempts to grasp a block by the sides and lift it away from the crate. Right Image: Agent deactivates with a 'release' gesture; Darwin's arms recalibrate back slowly to user-space

blends with users' expectation of the end-effectors' behavior during avateering.

Each technique prescribes to a different heuristic in their approach on interacting with the object. However for the study, all techniques shall share the same strategy in deciding for the user whether to maintain the grasp for transport after contact. We devised a scenario for the trial where avateering only the robot's right-arm is required, while its left-arm mirrors the joint angles of the right-arm. We state however, that our findings can be generalized to any arbitrary manipulator which are fairly anthropomorphic, and applies non-prehensile grasping when interacting with the objects. The trial requires the user to avateer a DARWIN-OP humanoid robot (Darwin) to complete the tasks in the following order; (i) Grasp a plank by the sides off the crate, (ii) Lift the plank, (iii) Move the plank away from the crate (iv) Drop the plank off the side of the standing platform. The user can choose to turn and walk the humanoid either to its left or right for (iii). Even though only the user's right arm is required, the user has to avateer with both arms, as the encoded gesture recognition heuristics require movements from both arms to decide whether to deactivate the agent after contact (Figure 4). Walking and turning Darwin is also by gestures. End-effector platform attachments 5) and arm joint constraints ensure that force correction will not be required and bounded in order for the Darwin-OP to grasp by the sides, and only contact prehensile grasping force is sufficient to lift the plank off the crate during free avateering.

A. No-DOF Technique

No-DOF is designed with the 'naive' approach, applying the modeling principle of Axiom (1) alone. The agent takes full control of Darwin's arms after their end-effectors meet the locality requirement, performing the arbitrative action by moving the arms to the goal points, and relinquishes back control. As the name of the technique implies, the user has no input during this arbitration phase. In the robotics-literature, corrective action is applied to all entries of the vector [q] of the manipulator, where [q] represents the joint variables of Darwin's three degree-of-freedom right-arm. With exception to the 2-DOF technique, all techniques used the iterative Damped-Least-Squares (DLS)[24] algorithm to

compute the manipulators' corrective paths to their target points. The set of Forward-Kinematics (FK) equations tightly couple between the joint-angle variables in [q] ($\theta_1, \theta_3, \theta_5$), rendering their solutions highly non-linear. Without stating the associated DH parameters, the FK system of equations (position) for Darwin's right arm are:

$$\{X\} = \begin{pmatrix} L_1 c_1 - L_3 s_3 - L_H s_1 s_5 + c_1 c_3 (L_2 + L_H c_5) \\ L_3 c_1 + L_1 s_1 + L_H c_1 s_5 + c_3 s_1 (L_2 + L_H c_5) \\ L_2 s_3 + L_H c_5 s_3 \end{pmatrix} (1)$$

where $\{X\}$ represents the end-effector point in 3D space, L_1 & L_3 represent upper-limb link offsets introduced to mitigate singularities, L_2 & L_H represent upper & lower arm lengths respectively, and the following abbreviations were used: $c_i = cos(\theta_i)$, $s_i = sin(\theta_i)$, $i = 1, 3, 5$

Upon agent activation, projected operation and task-space of Darwin's end-effectors around the plank usually provide good seeding data, allowing the algorithm to converge to a valid Inverse Kinematics (IK) solution in a low-order of iterations. If no valid solution is obtained after a set amount of iterations, user avateering input is used instead. After contact is made, the agent maintains a locking grasp on the object, and deactivates (unlocks) upon recognition of a 'release' gesture made by the user (Figure 4). For the remaining techniques, the agent's arbitrative strategy after contact is made with the object remains constant.

B. No-DOF Just-In-Time (JIT) Technique

No-DOF JIT is similar to No-DOF, but includes the modeling principles of Axiom (2). Besides locality, the agent activates if the end-effectors exceed the velocity and angle-of-approach threshold with respect to the object plane where the goal point resides.

C. 1-DOF Technique

1-DOF applies the alternative arbitration principle supported in Axiom (3). The user retains control of Darwin's upper-arm roll (Figure 3) when the agent activates (Axiom (1)), and can choose to deactivate the agent before contact by rolling the end-effector away from the locality threshold. Additionally, computing IK solutions can be less expensive
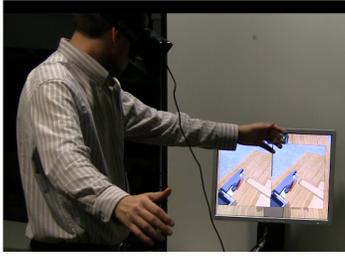
Fig. 5. User attempting the trial in a telepresence setting

(agent becomes more efficient during arbitration), as unique geometric solutions are possible if one of the elements in [q] is known.

D. 2-DOF Technique

2-DOF JIT is similar to 1-DOF, but the user retains control of the additional degree of elbow-bend (i.e. agent controls only the upper-arm pitch). Besides control, an additional difference involves the agent's arbitrative behavior. Once the 1-DOF agent acquires the goal point, it can choose either to set the elements in [q] it controls immediately, or to interpolate values in-between to set a path the manipulator takes to make contact with the object plane. Instead, 2-DOF sets the path frame-by-frame after the goal point is acquired, using an iterative approach similar to the Cyclic-Coordinate Descent (CCD) algorithm [25]. The object plane that the end-effector intends to make contact serves as a very useful constraint, assuring that user-set inputs of the upper-arm's roll and elbow-bend (θ_3,θ_5) will reach the goal point's z-coordinate. Based on Darwin's right-arm FK, the agent only needs to adjust the upper-arm pitch (θ_1) to move the end-effector closer to the goal-point in coordinates x-y with respect to the user's inputs for θ_3 and θ_5 in that current frame.

V. USER STUDY

The following sections describe the virtual environment setup, task objectives, devices and software, and participant demographics.

A. PARTICIPANTS

We recruited the help of 15 students from the University of Central Florida. Our pool consists of 12 males and 3 females, ages between 18 and 41.

B. DEVICES AND SOFTWARE

We ran our trial as a simulation using Webots EDU installed on Ubuntu 12.04 LTS. Webots provides a realistic model of the Darwin-OP since it is commonly used for Robocup [1]. However, we modified the end-effectors to allow for easier grasping of virtual objects. We used the Kinect motion sensor, with OpenNI 2.2[2] and NiTE 2.2.0.11[3] middleware to extract skeletal-joint data from the user for avateering.

[1]http://www.robocup2013.org/472-2/

[2]http://www.openni.org/

[3]http://www.openni.org/files/nite/

C. STUDY DESIGN

Darwin was placed on a platform along with a crate, upon which rested a plank (Figure 4). The participant was tasked with guiding Darwin's arms to grasp and lift the plank, navigate it to an edge of the platform, and release it. Each participant was asked to try five times to complete the trial, per technique, which excludes the control technique. The control is essentially free avateering, with no agent assisting the user to complete the tasks within the trial.

Since Darwin does not feature fingers, it cannot execute precision grasps [22]. We instead required the robot to use a non-prehensile power grasp by using both of the arms to press against the plank, securing it between the arms. We did not allow scooping of the plank from underneath.

Prior to beginning the tasks, we allowed the user a few moments to become familiar with the simulation. We asked the user to try commanding the arms as well as navigating the robot around an open area. Once the user felt comfortable with the commands, we proceeded with the actual trial. The order of techniques executed at each trial are randomized but ensured to be unique. If Darwin successfully picked up the plank, the user was required to turn and walk Darwin via body gestures to transport the plank to a side of the platform and release it. Once both conditions were met, or if the plank was knocked off the crate or dropped too early, the scenario was reset.

D. QUANTITATIVE METRICS

Since the difference between techniques were fairly subtle, we opted to not record the task completion time. Rather, we wanted to track the number of successes and failures for picking up the plank, measuring ''Grasp" and ''Release" rates for all users. Our hypothesis is that there will be superior techniques against the control ('Free' Avateering), if an agent did benefit the users.

E. QUALITATIVE METRICS

After each technique was used five times, we asked the participant to fill out a small survey consisting of seven questions about the experience. The posed questions are as follows:

- I quickly understood how to control the entire robot
- It was easy to pickup the plank from the crate
- The robot arms moved according to my expectation
- As a whole, performing the task was easy
- It was easy to make the robot walk
- The robot walked according to my expectations
- Overall, the robot moved according to my expectations

The questions were measured on a 7-point Likert scale. Although the participant was unaware of which technique they were using at any given time, our questionnaires were usable to measure perception of the technique. In the case of the techniques involving the agent, we attempted to determine if the user felt in-control of the robot during the trial.
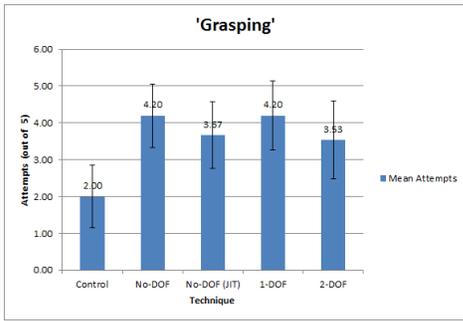
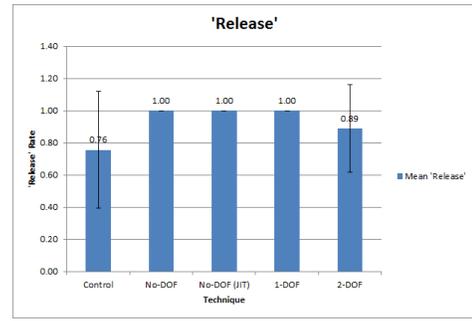Fig. 6.   Pickup ('Grasping') Success Rate.

Fig. 7.   Let-Go ('Release') Success Rate.

VI. RESULTS AND DISCUSSION

We targeted specific quantitative and qualitative metrics to analyze the difference between unassisted and assisted avateering. The following sections describe the results from the user study.

A. QUANTITATIVE ANALYSIS

Using a repeated measures ANOVA test, significant difference was found between the techniques ($F_{4,14} = 15.852; p < .0001$), with the sphericity assumption maintained. With the control technique removed, no significant difference was found across the techniques ($F_{3,14} = 2.040; p = 0.123$). Figure 6 illustrates the task completion rate for picking up the plank. It is very clear to see that the avateering metaphor benefited from the inclusion of an agent. We performed T-tests between all pairs of techniques for a total of 10 comparisons. Using Holm's Sequential Bonferroni's Adjustment[23] we controlled Type I errors. Significant difference was found between the control technique and the other four involving the agent:

- No-DOF: $t_{14} = -9.886; p < .005$
- No-DOF JIT: $t_{14} = -5.229; p < .0071$
- 1-DOF: $t_{14} = -7.432; p < .0056$
- 2-DOF: $t_{14} = -5.996; p < .00625$

As expected, the avateering metaphor greatly benefited from an assistive agent when grasping the plank. The four agent-assisted techniques allowed the participant to pickup the plank twice as often as the control base metaphor, proving that free avateering is not sufficiently efficient by itself.

In addition to helping users pickup the plank, two of the techniques assisted the user in maintaining a grasp on the plank until the end of the trial. Figure 7 illustrates the success rate for maintaining the grip and dropping the plank when needed. This success rate is dependent upon the prior task of successfully lifting the plank from the crate. The large standard deviation of this success rate suggests that the presence of expert users who accustomed quickly to avateering do not benefit greatly from the agent's assistance. This also suggests that the agent is useful for non-expert or inattentive users, or to alleviate the inadequacy of the motion sensor interface either by noise, or by inadvertent arm movements during the transport.

We also observed that users who are more confident with their movements scored better with the 'No-DOF JIT' & '2-

DOF' techniques, while users who are feeling more careful excelled with 'No-DOF' & '1-DOF'. This suggests that the threshold parameters across the techniques can be pre-calibrated according to a user's affect before undergoing the trials.

B. QUALITATIVE ANALYSIS

We used Friedman tests in order to determine statistical significance between the techniques. In the event of significance, we used Wilcoxon's signed rank tests between all techniques, for a total of 10 comparisons. We were unable to find statistical significance between the control technique and the others in regards to user perception. This aligns with our expectations; the users were relatively unable to determine that the robot arms were being guided by the agent.

However, the participants did respond via our questionnaires that it was easier to pickup the plank from the crate using 3 of the 4 techniques:

- No-DOF - Mean: 5.1; Median: 5
- No-DOF JIT - Mean: 4.8; Median: 5
- 1-DOF - Mean: 5.0; Median: 6
- 2-DOF - Mean: 3.9; Median: 3
- Control - Mean: 4.1; Median: 4

Second, the users responded that the robot arms moved more accordingly to their expectations, when using a technique involving the agent:

- No-DOF - Mean: 5.6; Median: 6
- No-DOF JIT - Mean: 5.0; Median: 5
- 1-DOF - Mean: 5.1; Median: 5
- 2-DOF - Mean: 4.1; Median: 4
- Control - Mean: 4.7; Median: 5

This positive outcome reveals that although the users did not have knowledge of the assistive agent arbitrating their arms, they still felt more in control than free avateering. Additionally, not only do the arms move according to their expectation, their perception is similar to the agent-assisted techniques.

VII. CONCLUSION & FUTURE WORK

Agent-assisted avateering is definitely preferred to non-assisted free avateering, however, it must not come at a loss of user expectation to the control metaphor. Our results indicate that though users were oblivious to the agent's correction across the assisted techniques, task completion was achieved at a significantly higher rate compared with unassisted. In this paper, we demonstrated and evaluated

an effective modeling paradigm to design four assistive agents that enabled the teleoperator to outperform the control technique, and also meeting their expectation and perception of avateering.

Our arbitration agents work effectively to ensure task accuracy with humanoids having non-prehensile grasping, and complements well with a simple heuristic that recognizes a 'release' gesture to deactivate the agent. In the future, we plan to extend this work to include interaction across the arm manipulator and wrist, and beyond non-prehensile grasps. One possible application includes medical robot manipulators with high redundancy, which are currently used to perform serious surgeries that require utmost precision. In order to execute a precision grasp that aligns with the user's expectations, the arm and fingers would need to translate to one of many possible positions. The agents currently use a simple gesture recognition heuristic to predict user intent for object 'release', and is triggered only when the robot end-effectors make contact. A more powerful approach would be to formalize the coordination between the gestural grammar with agent-assisted avateering in order to define a more complete avateering control model.

VIII. ACKNOWLEDGMENT

REFERENCES

[1] Pierce, R.M.; Kuchenbecker, K.J., "A data-driven method for determining natural human-robot motion mappings in teleoperation," Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on , vol., no., pp.169,176, 24-27 June 2012

[2] Koenemann, J.; Bennewitz, M.; "Whole-body imitation of human motions with a nao humanoid." In Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI '12). ACM, New York, NY, USA, 425-426.

[3] Kofman, J.; Xianghai Wu; Luu, T.J.; Verma, S., "Teleoperation of a robot manipulator using a vision-based human-robot interface," Industrial Electronics, IEEE Transactions on , vol.52, no.5, pp.1206,1219, Oct. 2005

[4] Nguyen, V.; Lee, J.; "Full-body imitation of human motions with kinect and heterogeneous kinematic structure of humanoid robot," System Integration (SII), 2012 IEEE/SICE International Symposium on, vol., no., pp.93,98, 16-18 Dec. 2012

[5] Hoffman, G.; Kubat, R.; Breazeal, C., "A hybrid control system for puppeteering a live robotic stage actor," Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on , vol., no., pp.354,359, 1-3 Aug. 2008

[6] Pineau, J; Gordon, G, "POMDP planning for robust robot control", The Twelveth International Symposium on Robotics Research, 2005

[7] Suay, H.B.; Chernova, S., "Humanoid robot control using depth camera," Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on , vol., no., pp.401,401, 8-11 March 2011

[8] Guo, C.; Sharlin E., "Exploring the use of tangible user interfaces for human-robot interaction: a comparative study." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08). ACM, New York, NY, USA, 121-130.

[9] Wai Shan Ng; Sharlin, E., "Collocated interaction with flying robots," RO-MAN, 2011 IEEE , vol., no., pp.143,149, July 31 2011-Aug. 3 2011

[10] Weber, C.; Nitsch, V.; Unterhinninghofen, U.; Farber, B.; Buss, M., "Position and force augmentation in a telepresence system and their effects on perceived realism," EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint , vol., no., pp.226,231, 18-20 March 2009

[11] Khokar, K.H.; Alqasemi, R.; Sarkar, S.; Dubey, R.V., "Human motion intention based scaled teleoperation for orientation assistance in pre-shaping for grasping," Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on , vol., no., pp.1,6, 24-26 June 2013

[12] Scerri, P.; Johnson, L.; Pynadath, D.; Rosenbloom, P.; Schurr, N.; Si, M.; Tambe, M., "Getting robots, agents, and people to cooperate: An initial report." In AAAI Spring Symposium on Human Interaction with Autonomous Systems in Complex Environments, 2003.

[13] Dragan A., Srinivasa S.; 2013. A policy-blending formalism for shared control. Int. J. Rob. Res. 32, 7 (June 2013), 790-805

[14] Hye Sun Park, Eun Yi Kim, Sang Su Jang, Se Hyun Park, Min Ho Park, and Hang Joon Kim. 2005. "HMM-Based gesture recognition for robot control," In Proceedings of the Second Iberian conference on Pattern Recognition and Image Analysis - Volume Part I (IbPRIA'05), Jorge S. Marques, Nicols Prez de la Blanca, and Pedro Pina (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 607-614.

[15] Crandall, J.W.; Goodrich, M.A., "Characterizing efficiency of human robot interaction: a case study of shared-control teleoperation," Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on , vol.2, no., pp.1290,1295 vol.2, 2002

[16] Lewis, B.; Tastan, B.; Sukthankar, G.; "An adjustable autonomy paradigm for adapting to expert-novice differences," Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on , vol., no., pp.1656,1662, 3-7 Nov. 2013.

[17] Nutan Chen, Keng Peng Tee, Chee-Meng Chew, and Rui Yan. 2012. "Intuitive interaction for robotic grasping," In Proceedings of the Workshop at SIGGRAPH Asia (WASA '12). ACM, New York, NY, USA, 105-111.

[18] Huebner, K.; Welke, K.; Przybylski, M.; Vahrenkamp, N.; Asfour, T.; Kragic, D.; Dillmann, R., "Grasping known objects with humanoid robots: A box-based approach," Advanced Robotics, 2009. ICAR 2009. International Conference on , vol., no., pp.1,6, 22-26 June 2009.

[19] Saxena, A., Driemeyer, J.; & Ng, A. Y. (2008)."Robotic grasping of novel objects using vision," The International Journal of Robotics Research, 27(2), 157-173.

[20] Hening, S., Cottrell, P., Teoderescu, M., Kurniawan, S., & Mantey, P. (2013, May). "Assistive living robot: a remotely controlled robot for older persons living alone," In Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments (p. 10). ACM.

[21] Lewis, B.; Sukthankar, G.; "Two hands are better than one: Assisting users with multi-robot manipulation tasks," Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on , vol., no., pp.2590,2595, 25-30 Sept. 2011

[22] Cutkosky M., Howe. R.; "Human grasp choice and robotic grasp analysis. In Dextrous robot hands," S. T. Venkataraman and T. Iberall (Eds.). Springer-Verlag New York, Inc., New York, NY, USA 5-31 1990.

[23] Holm, S. "A simple sequentially rejective multiple test procedure." Scandinavian journal of statistics (1979): 65-70.

[24] Buss, S.; Kim, J., "Selectively Damped Least Squares for Inverse Kinematics," J. Graphics Tools 10(3):37-49 (2005)

[25] Kenwright, B., "Inverse Kinematics Cyclic Coordinate Descent (CCD)," Journal of Graphics Tools, 16:4, 177-217 (2012)

[26] Rennuit, A.; Micaelli, A.; Merlhiot, X.; Andriot, C.; Guillaume, F.; Chevassus, N.; Chablat, D.; Chedmail, P., "Passive control architecture for virtual humans," Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on , vol., no., pp.1432,1437, 2-6 Aug. 2005

[27] Mingxing Liu; Micaelli, A.; Evrard, P.; Escande, A.; Andriot, C., "Interactive dynamics and balance of a virtual character during manipulation tasks," Robotics and Automation (ICRA), 2011 IEEE International Conference on , vol., no., pp.1676,1682, 9-13 May 2011

[28] Kim, W.S.; Hannaford, B.; Fejczy, A.K., "Force-reflection and shared compliant control in operating telemanipulators with time delay," Robotics and Automation, IEEE Transactions on , vol.8, no.2, pp.176,185, Apr 1992

[29] "Experimental Study on Remote Manipulation Using Virtual Reality," PRESENCE, vol.2, no.2, pp.112-124, 1993.