Flexible, Semi-Autonomous Grasping for Assistive Robotics

Jörn Vogel[1⋆], Katharina Hertkorn[1⋆], Rohit U. Menon[2], Máximo A. Roa[1]

***Abstract*— This paper proposes a scheme to provide flexible semi-autonomous grasping capabilities to an assistive robotic manipulator. The testbed consists of a five-finger robotic hand mounted on a robotic arm. During teleoperation the position of the hand is continuously controlled in the three translational degrees of freedom, and the user has no direct influence over the rotational behavior. The proposed semi-autonomy scheme assists the user for moving and orienting the hand towards the object, and automates the grasping process when it is triggered. The velocity commands issued by the user are enhanced using virtual fixtures, which are not preprogrammed to support one approach direction to the (known) object, but are adapted online according to the intended movement. The approach is validated with a psycho-physical user study where the participants grasp objects in a simulation environment using a SpaceMouse interface. This setting serves as a testbed for the target application in which disabled subjects will control the real robotic system with an interface based on bio-signals. The user study compares the semi-autonomous and the pure teleoperation modes in terms of objective and subjective measures, showing a performance increase (in terms of time to complete the task) and a workload decrease for the proposed semi-autonomous mode.**

I. INTRODUCTION

The use of assistive robotic manipulators for people with disabilities is becoming more common and affordable (e.g. the Jaco arm [1]). These manipulators are usually controlled through a human-machine interface (HMI), ranging from simple joysticks to sensors that detect remaining muscular activity [2], or invasive brain-computer interfaces (BCI) [3], [4]. While all these interfaces offer an intuitive scheme to control the robotic manipulator via velocity commands, they have drawbacks such as limited input capabilities or noisy control signals. Hence, the user needs to actively learn how to position the arm in order to grasp an object in a stable way, but it often happens that slightly inaccurate commands lead to grasp failures, converting grasping and manipulation of objects in a tedious task.

The use of autonomous components simplifies the task of grasping objects in this kind of teleoperated systems. Several solutions have been proposed to the problem of planning online the movements required to autonomously and successfully grasp an object [5]; in general, a set of grasps is computed for a given object, one grasp is selected according to some desired criterion, and a path planner generates the motion to execute the grasping task. Similar approaches have been implemented in the case of robots controlled via BCIs: the user selects a discrete grasping action, which is then executed in a completely autonomous way [6], [7], [8], [9]. However, this full autonomy reduces the flexibility of the system, as the target grasp is automatically selected. Moreover, it does not seem to be very satisfactory for people with upper limb paralysis using assistive technology [10]. Thus, it is desired to have the human involved in the control of the system and enhance the usability with shared control capabilities, which are called only when really needed. However, in most of these semi-autonomous approaches the user selects a grasp pose from a very restricted set of predefined grasps, and the motion of the robot is oriented towards that grasp pose [11]. Since the human operator is in the loop and has a task in mind, he knows best how the object needs to be grasped in order to successfully use it, but to use this information the user needs to have greater freedom for choosing potential grasps and directions to approach the object.

This paper proposes an assistance scheme based on shared-control, which allows the user to grasp objects in many different configurations, and is therefore flexible enough to closely follow the user's intention. To guarantee a large freedom for choosing target grasps, the system uses the "graspability map" [12], a dense grasp database that covers all possible hand poses that lead to force-closure precision grasps on the object. During runtime, the commanded velocities and the current position of the robot with respect to the object are used to search for a stable grasp pose in the database. To simplify the motion of the robot towards the object, the user-commanded velocity is gradually modified to converge to the selected grasp pose using virtual fixtures [13]. In the proposed system, the user only controls the translational motion of the robotic manipulator. The orientation of the end effector is autonomously controlled based on the grasp pose selected from the database. As the target grasp pose is continuously updated, the user can always influence the position of the hand relative to the object, i.e. he can move around the object to select a different configuration to grasp it, or can even withdraw the action.

The paper is organized as follows. After a brief overview of the integrated grasp database in Section II, the concept for realizing dynamic virtual fixtures is presented in Section III, along with the most relevant implementation details. In Section IV the assistive system is evaluated with a user study, and conclusions are presented in Section V.

[1]Jörn Vogel, Katharina Hertkorn and Máximo A. Roa are with the Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany. Email: firstname.lastname@dlr.de

[2]Rohit U. Menon is with the Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Bremen, Germany.

⋆These authors contributed equally to this work

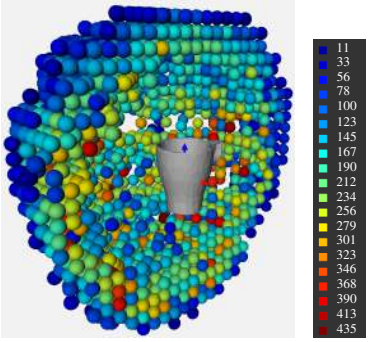


Fig. 1. Visualization of the Graspability Map Γ for a coffee mug. The color coding represents the size of the computed contact regions $\mathbf{s}_{\text{ricr}}$, which provides a measure of the robustness of the grasp (the larger the better).

II. GRASPABILITY MAP

In order to allow a flexible assistance that restricts the user and his intended motion as little as possible, we combine a dense grasp database with the concept of dynamic virtual fixtures. The integrated grasp database, or "graspability map" [12], is generated offline using a point cloud with associated normals that represents the object, the workspace Φ for the fingertips of the robotic hand, and a friction coefficient μ that describes the friction between the fingertips and the object.

The graspability map consists of a set $\Gamma = \{\mathbf{T}_0, \ldots, \mathbf{T}_m\}$ of homogeneous transformations $\mathbf{T}_i = [\mathbf{R}_i \ \mathbf{p}_i] \in \mathbb{R}^{3\times4}$, with $\mathbf{R}_i \in \mathbb{R}^{3\times3}$ a rotation matrix and $\mathbf{p}_i \in \mathbb{R}^3$ a translation vector. Each $\mathbf{T}_i$ describes a pose of the hand base frame, relative to the given object, where a force-closure (FC) precision grasp can be achieved. To compute the map we divide the space around the object into voxels with a user-defined resolution σ (for this work, $\sigma = 2\,\text{cm}$). The FC grasps are computed based on the concept of reachable independent contact regions [14], and provide a feasible contact area per finger (not only one contact point), which improves the robustness of the analytically computed grasps. Fig. 1 shows an example of the graspability map. For each pose $\mathbf{T}_i \in \Gamma$, the stored data are the grasp quality ε (largest minimum resisted wrench) [15], the size of the computed contact regions $\mathbf{s}_{\text{ricr}}$, and the finger configuration $\mathbf{q}_{\text{goal}}$ to perform the grasp. These data are later used to determine the target pose for the virtual fixtures, as described in the following section.

III. DYNAMIC VIRTUAL FIXTURES

To assist the operator in reaching a hand pose where an FC grasp can be executed, we use the method of dynamic virtual fixtures [13]. In this method, the target of the fixture is dynamically updated based on the robot position with respect to the object, and on a heuristic quality-criterion that rates the intuitiveness of the available grasps, for instance to prevent the hand from grasping an object in an unnatural upside-down pose. The developed method for dynamic virtual fixtures has two major components: determination of the target hand pose, and guidance of the end effector (depending on the user velocity) towards that pose.

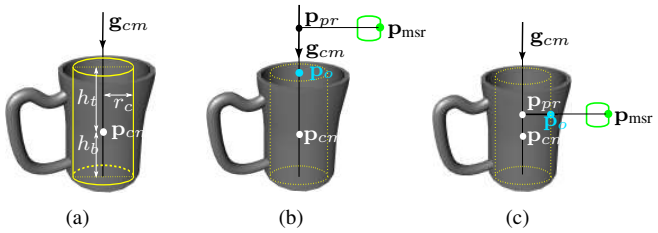


Fig. 2. Calculation of the point of interest $\mathbf{p}_o$ depending on the current position of the end effector $\mathbf{p}_{\text{msr}}$. a) Cylinder (depicted in yellow) inscribed in the object, with radius r_c and distances h_t and h_b from the center of mass $\mathbf{p}_{cm}$ to the upper and lower intersection point of the gravity direction $\mathbf{g}_{cm}$ with the cylinder. b) If the projection point $\mathbf{p}_{pr}$ is above the bounding box of the object, $\mathbf{p}_o$ is at the top of the object. c) If $\mathbf{p}_{pr}$ is within the object, $\mathbf{p}_o$ is at the side of the inscribed cylinder.

A. Determination of the target hand pose

The actual target pose $\mathbf{T}_{\text{targ},k}$ is obtained through an online rating of grasps at every time step k. The rating is defined to match the target pose with the user intent, so that the resulting movement of the robotic arm and the grasp configuration feel natural to the human operator. Thus, if the human approaches the object from above, it is very likely that it should be grasped from the top, and when the object is approached from the side, the grasp pose should also be on the object's side. Naturally, continuous variations occur in the grasping pose according to the approach direction. To achieve this, a point of interest $\mathbf{p}_o$ is defined on the object. It is determined from the projection of the current end effector position $\mathbf{p}_{\text{msr}}$ onto the gravity vector $\mathbf{g}_{cm}$ passing through the object's center of mass $\mathbf{p}_{cm}$, as illustrated in Fig. 2. This results in the projection point $\mathbf{p}_{pr} = \text{Pr}(\mathbf{p}_{\text{msr}}, \mathbf{g}_{cm})$. Depending on the location of $\mathbf{p}_{pr}$, the point of interest is

$$\mathbf{p}_o = \begin{cases} \mathbf{p}_{pr} + r_c & \text{if } \mathbf{p}_{pr} \text{ is within the object} \\ \mathbf{p}_{cm} - \frac{h_t}{\|\mathbf{g}_{cm}\|}\mathbf{g}_{cm} & \text{if } \mathbf{p}_{pr} \text{ is above the object} \\ \mathbf{p}_{cm} + \frac{h_b}{\|\mathbf{g}_{cm}\|}\mathbf{g}_{cm} & \text{if } \mathbf{p}_{pr} \text{ is below the object,} \end{cases} \tag{1}$$

where r_c is the radius, and h_t and h_b are the distances from the center of mass to the top and bottom of the cylinder inscribed in the object, respectively (Fig. 2(a)). This auxiliary cylinder is only needed to determine $\mathbf{p}_o$.

The point $\mathbf{p}_o$ helps to extract from the graspability map a set $\mathcal{T}_e \subset \Gamma$ of possible target grasp poses, valid in this time step: $\mathcal{T}_e = \{\mathbf{T}_{e,1}, \ldots, \mathbf{T}_{e,m}\}$. To achieve this, first the intersection point $\mathbf{p}_{\text{out}}$ between the commanded velocity direction $\mathbf{v}_{\text{user}}$ and the bounding box of the map BB_{map} is found (blue circle in Fig. 3). If the hand is already within the bounding box of the map, then $\mathbf{p}_{\text{out}}$ is just set equal to the current position of the end effector. Now we obtain $\mathcal{T}_e$ as the set of all hand poses that lead to a FC grasp along the line from $\mathbf{p}_{\text{out}}$ to the point of interest $\mathbf{p}_o$. To do this, the line $\mathbf{p}_o - \mathbf{p}_{\text{out}}$ is divided by the resolution σ of the map, and for each point on the line the nearest voxel of the graspability map is computed (red and green circles). In the case that one voxel does not contain valid FC grasps, the closest voxel with a valid FC grasp is picked up from the database (green circles with black lines in Fig. 3).

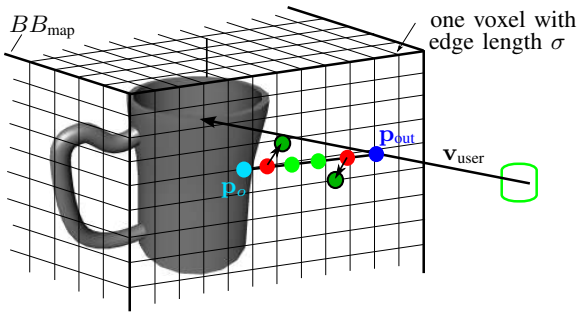


Fig. 3. Determination of valid hand poses in the set $\mathcal{T}_e$. The point $\mathbf{p}_{\text{out}}$ is obtained as the intersection of the user commanded velocity $\mathbf{v}_{\text{user}}$ with BB_{map}. The line from $\mathbf{p}_{\text{out}}$ to $\mathbf{p}_{\text{o}}$ crossses voxels with and without valid FC grasps (green and red circles, respectively).

In a second step, all poses $\mathbf{T}_{e,i} \in \mathcal{T}_e$ are rated to choose the most suitable grasp configuration at time k. The overall rating for each pose $\mathbf{T}_{e,i}$ considers both functional aspects and the naturalness of the grasp. The functionality of the grasp is considered in a function $F(\varepsilon, \mathbf{s}_{\text{ricr}})_i$, which makes a weighted sum of the grasp quality ε and the size of the reachable independent contact regions $\mathbf{s}_{\text{ricr}}$.

As the user does not have any additional visual information or input to choose a grasp, the rating also tries to favor the most "natural" hand poses. For instance, grasp poses where the hand is turned upside down may look awkward to the user (Fig. 4(c)). Furthermore, in our system, the user only has direct control over the translational movement of the robot. We therefore assume that minimizing the rotational movement needed to reach a valid grasp pose leads to a more transparent and natural (predictable) behavior of the robot. The naturalness N of a grasp is then defined here as a function that depends on three different variables: θ_{align}, θ_{tot}, and θ_{roll}.

The parameter θ_{align} evaluates the alignment of the hand with respect to the axis that defines the target orientation $\mathbf{a}$. This axis is defined depending on the location of the hand with respect to the object. To define it, we follow a procedure similar to the computation of the point of interest $\mathbf{p}_o$. The center $\mathbf{p}_{r,i}$ of the considered voxel is projected onto the direction of gravity $\mathbf{g}_{cm}$ to obtain the point $\mathbf{p}_{pr} = \text{Pr}(\mathbf{p}_{r,i}, \mathbf{g}_{cm})$. Thus, the axis $\mathbf{a}$ is given by

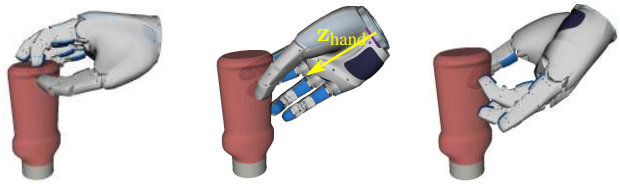
$$\mathbf{a} = \begin{cases} \mathbf{p}_{pr} - \mathbf{p}_{r,i} & \text{if } \mathbf{p}_{pr} \text{ is inside the object} \\ \mathbf{p}_{cm} - \mathbf{p}_{pr} & \text{if } \mathbf{p}_{pr} \text{ is outside the object.} \end{cases} \tag{2}$$

Then, θ_{align} is calculated as the deviation in orientation between the axis $\mathbf{a}$ of the current target pose and the axis $\mathbf{z}_{\text{hand}}$ of the hand, which points from the palm towards the fingertip (Fig. 4(b)), i.e. $\theta_{\text{align}} = \measuredangle(\mathbf{a}, \mathbf{z}_{\text{hand}}) \in [0, \pi]$.

The second parameter rates the total angular rotation θ_{tot} between the orientation $\mathbf{R}_{e,i}$ at the target grasp configuration and the orientation $\mathbf{R}_{\text{init}}$ at the initial pose of the hand. Both hand poses are compared using the angle-axis representation

$$\mathbf{R}_{12} = \mathbf{R}_{\text{init}}^{-1} \mathbf{R}_{e,i} \tag{3}$$

$$[\beta, \mathbf{b}] = \text{angle-axis}(\mathbf{R}_{12}), \tag{4}$$



(a) (b) (c)

Fig. 4. Three hand poses achievable from the same voxel. While a) and b) look like natural grasps, c) seems awkward. The overall rating for these cases is $\lambda_a = 1.20$, $\lambda_b = 1.74$, $\lambda_c = 0.70$ (the higher the better).

where $\mathbf{b}$ is the axis, consisting of the elements b_i, and β the angle spanned from the initial to the target orientation of the hand. The total angular rotation is then

$$\theta_{\text{tot}} = \sum_i (\beta b_i) \in [0, 2\pi]. \tag{5}$$

The third parameter is the required roll angle $\theta_{\text{roll}} \in [0, \pi]$ around the hand axis $\mathbf{z}_{\text{hand}}$, computed using the Euler angle representation. Though grasp configurations with a high θ_{roll} do not pose any problem in terms of grasp functionality, they might not seem natural, as illustrated in Fig. 4(c).

Finally, the overall rating λ_i for each relevant hand pose i is a combination of the functionality F and naturalness N:

$$\begin{aligned} \lambda_i &= F(\varepsilon, \mathbf{s}_{\text{ricr}})_i \cdot N(\theta_{\text{align}}, \theta_{\text{tot}}, \theta_{\text{roll}})_i \\ &= \exp\Big(-((1 - f_1(\varepsilon)) + (1 - f_2(\sum(\mathbf{s}_{\text{ricr}})))\Big) \cdot \\ &\quad \exp\Big(-(w_4|\theta_{\text{align}}| + w_5|\theta_{\text{tot}}| + w_6|\theta_{\text{roll}}|))\Big). \end{aligned} \tag{6}$$

The exponential function is chosen to ensure a low score for grasp configurations where at least one of the criteria has a low value. The functional parameters are normalized through the functions f_1 and f_2, such that $f_n(x) \in [0, 1]$, and factors w_j are used to guarantee equal weight for each criterion of naturalness. Finally, the pose with the highest λ_i is selected as the target pose $\mathbf{T}_{\text{targ},k}$ at the particular time step k.

B. Guidance of the end effector

To assist the user in grasping the object, our method continuously modifies the commanded translational velocity $\mathbf{v}_{\text{user}} \in \mathbb{R}^3$ at each time step k, to converge toward the desired target pose. The autonomous behavior of the system depends on the distance d_{msr} of the current position of the end effector with respect to the object: the closer the end effector is, the higher the role of the autonomy. For this purpose, we define four zones that surround the object, as shown in Fig. 5. In each zone, the influence of the autonomy is determined by a factor $\alpha \in [0, 1]$, explained in further detail below, and the resultant velocity $\mathbf{v}_{\text{res}}$ of the end effector is

$$\mathbf{v}_{\text{res}} = (1 - \alpha)\mathbf{v}_{\text{user}} + \alpha \mathbf{v}_{\text{auto}}, \tag{7}$$

where the autonomous component $\mathbf{v}_{\text{auto}}$ of the velocity is

$$\mathbf{v}_{\text{auto}} = \|\mathbf{v}_{\text{user}}\| \frac{\mathbf{p}_{\text{targ}} - \mathbf{p}_{\text{msr}}}{\|\mathbf{p}_{\text{targ}} - \mathbf{p}_{\text{msr}}\|} \tag{8}$$

and depends on the current position $\mathbf{p}_{\text{msr}}$ of the end effector and the position $\mathbf{p}_{\text{targ}}$ of the target hand pose.

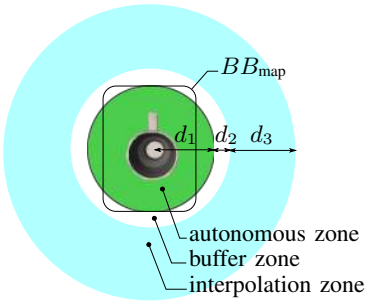


Fig. 5. Zones of Operation: green: autonomous zone $\alpha = 1$, white: buffer zone $\alpha = 1$, cyan: interpolation zone $\alpha \in [0, 1]$, and outer zone $\alpha = 0$.

In the outer zone ($d_{\text{msr}} \geq d_{\text{out}} = d_1 + d_2 + d_3$), where the hand is still far from the object, the user is in full control and there is no influence of the autonomy, therefore $\alpha = 0$. In the interpolation zone ($d_1 + d_2 \leq d_{\text{msr}} \leq d_{\text{out}}$) the orientation of the end effector is gradually changed to match the orientation of the target pose, as explained in the next subsection. Also, when the hand enters this area the velocity is gradually influenced by the autonomy, and α changes according to

$$\alpha = \min(e^{-(d_{\text{msr}} - d_1)/d_3}, 1). \tag{9}$$

When the hand enters the buffer zone ($d_1 \leq d_{\text{msr}} \leq d_1 + d_2$), its orientation is already the same as the orientation of the target pose. This buffer was added as a zone where the user can observe the hand orientation to decide if it matches his expectations; if not, the user can retract the hand from the object to try to find a better grasp configuration. Once the hand enters the autonomous zone ($d_{\text{msr}} \leq d_1$), $\alpha = 1$ and the velocity direction is autonomously set until the end effector reaches the target pose. This directional fixture towards the object is aborted only if the user wants to retract the hand. No further update of the target pose is required in this zone, and the selected hand orientation is held and can only be changed if the end effector point moves out of the autonomous zone. Finally, if the user triggers the grasp, the end effector moves autonomously to the selected grasp pose and executes the grasp.

C. Acquisition of the target orientation

The change in orientation of the end effector within the interpolation zone is carried out using spherical linear interpolation (SLERP) [16]. The rotation matrices for the initial and target poses are converted into quaternions (q_{init} and q_{targ}), and intermediate quaternions q_{inter} are given by

$$q_{\text{inter}} = \frac{\sin((1-t)\theta_{\text{tot}})}{\sin\theta_{\text{tot}}} q_{\text{init}} + \frac{\sin(t\theta_{\text{tot}})}{\sin\theta_{\text{tot}}} q_{\text{targ}}, \tag{10}$$

where t indicates the interpolation fraction, which depends on the distance of the hand with respect to the object:

$$t = \begin{cases} 0 & \text{if } d_{\text{msr}} > d_{\text{out}} \\ (d_{\text{msr}} - d_1 - d_2)/d_3 & \text{if } d_1 + d_2 < d_{\text{msr}} \leq d_{\text{out}} \\ 1 & \text{if } d_{\text{msr}} \leq d_1 + d_2 \end{cases} \tag{11}$$

Fig. 6 provides and example of such interpolation.

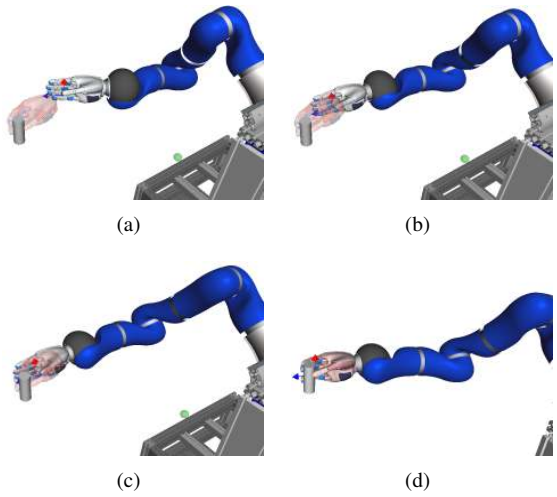


Fig. 6. Interpolation of the hand orientation, depending on the distance with respect to the object: a) $t = 0$, b) and c) $0 < t < 1$, d) $t = 1$.

To ensure that the change in orientation is achieved without unintended collisions, the interpolation needs to finish at a safe distance from the object. Thus, the lower limit of the interpolation zone, i.e. $d_1 + d_2$, should at least be equal to the length of the hand from its palm to the fingertips. This is guaranteed if $d_1 + d_2$ is equal to the radius of a sphere that circumscribes the bounding box of the map, BB_{map}. The distance d_1 is determined by the radius of a sphere inscribed in the bounding box of the graspability map, so that the autonomous zone encloses all the hand positions that allow an FC grasp on the object. The distance d_3, the only one that does not depend on the object itself, is a tradeoff between the desired speed of interpolation and the radius of influence of the autonomy on the movement of the hand.

D. Implementation Details

The implementation of the dynamic virtual fixtures is realized in two separated modules, one for the rating of grasp poses and another one for the velocity mapping. Thus, the velocity mapping (Sec. III-B) can be executed in the hard real-time control loop of the robot, in our case within $1\,\text{ms}$. The slower rating of poses (Sec. III-A) only needs to match the commanded user velocity, which in our system is sampled at a soft-realtime rate of $100\,\text{Hz}$. In order to provide updates to the target grasp pose without slowing down the user, we precompute offline the nearest FC pose for each voxel position that does not contain a valid FC grasp (red circles in Fig. 3). This is achieved by using a brute force search within the voxel space to look for the closest voxel (using the $\mathcal{L}_2$-norm) that contains a valid grasp configuration.

IV. EVALUATION

A virtual environment for the DLR assistive robotics setup [17] was used to evaluate the assistance for hand positioning with dynamic virtual fixtures. Usually this system is driven via a bio-signal based HMI that provides three Cartesian degrees of freedom and a discrete grasp trigger, as shown in Fig. 7 [2], [3]. The Cartesian degrees of freedom are mapped

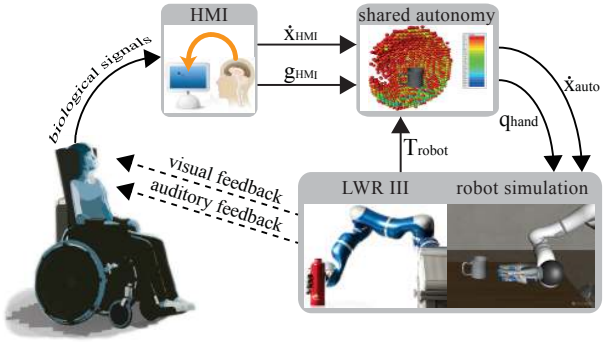


Fig. 7. Semi-autonomous grasping with the DLR Assistive Robotics Setup.

to the velocity of the end effector of the manipulator. The user shares the same workspace as the robotic manipulator, and therefore gets natural visual and auditory feedback.

In order to reduce the complexity of the system for this evaluation, we replaced the bio-signal based HMI with a SpaceMouse[1]. This allows a faster setup (as no training for the bio-signals is needed) and a reduction of side effects that might occur with the bio-signal based interface, such as differences in the decoding quality between subjects. However, since the SpaceMouse provides a cleaner signal than a bio-signal based HMI, a distortion function was applied to the SpaceMouse commands in one part of the experiment, to replicate conditions where the user intention is wrongly identified and the robot moves in a direction different from the commanded one. We evaluated the functionality of our approach based on three hypothesis:

- H1 Assistance increases the performance.
- H2 Assistance decreases the workload.
- H3 The effect of assistance is more evident for a distorted mapping.

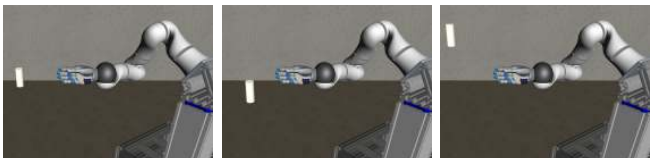
The experimental setup is described below, followed by the summary of results and a discussion on the findings.

A. Experimental Setup

Using a simulation instead of real-world experiments allows to specifically explore the effect of assistance, and minimizes the effect of possible errors coming from calibration and object detection in the real world. Thus, the chosen experimental setup leads to higher control over the experimental setup, and therefore to a more fair comparison of trials between the participants. The overall setup is shown in Fig. 8, where a participant sits in front of the simulation visualized on a 3D TV[2]. Participants wear active shutter glasses to ensure depth perception during the task. The visualization is provided via Instantplayer[3]. The virtual scene resembles the physical robotic setup, and consists of a DLR-HIT Hand II [18] mounted on a light-weight robot arm [19]. The scene is viewed on the screen from the perspective of a user sitting next to the robot base on the left (Fig. 8).

[1]SpaceMouse, from 3D Connexion, http://www.3dconnexion.de/
[2]Samsung UE46ES6300, 46-inch 3D LED TV
[3]Instantplayer, from Fraunhofer IGD, http://www.instantreality.org/

Fig. 8. Experimental setup to evaluate the assistance for hand positioning with virtual fixtures.



(a) (b) (c)

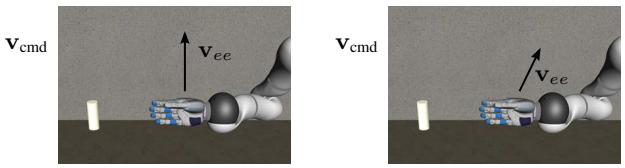
Fig. 9. Three different positions for the experiment: (a) “Easy”: the object lies in front of the hand; (b) “Medium”: object and hand lie on the same horizontal plane; (c) “Difficult”: object and hand are at different heights.

For the experiments, a cylindrical object must be grasped, and the distances that define the zones of operation are $d_1 = 0.1\,\mathrm{m}, d_2 = 0.03\,\mathrm{m}$, and $d_3 = 0.12\,\mathrm{m}$.

1) Sample: All participants were technical staff or students from the Institute of Robotics and Mechatronics at DLR. They were provided with a detailed description of the experiments, and gave written consent for their voluntary participation in this study. A total of 18 subjects (11 male, 7 female) with an average age of $M = 26.7$ y ($SD = 2.6$ y) participated, all had a background on engineering but none of them had previous experience with the setup. In fact, 12 of the participants do not work at all with CAD programs, and hence, had only limited or no experience using the SpaceMouse. The other 6 participants work on average 3.7 h per week with CAD programs.

Participants were asked to grasp the cylinder as fast as possible. The object is not placed on a table, but instead is free floating in space, to allow testing all possible approach directions to the object. The participants used the SpaceMouse to command the translational velocity of the hand, and with its buttons they could trigger the grasp. To introduce variations in the reaching motion towards the object, three different initial positions of the cylinder relative to the robotic hand were chosen, as depicted in Fig. 9. The difficulty of the positions increases from an “easy” position, where the object lies directly in front of the hand, over “medium” where the object is on the same horizontal plane as the hand, to “difficult”, where the object is at a different height with respect to the hand, and a motion in all three translational axis is required to grasp the object from the side. However, in all three conditions, the object was located at a distance of 0.3 m from the robots endefector.

The influence of mapping between the user command and the corresponding robot motion is also analyzed, as an HMI



(a) (b)

Fig. 10. Participants had to control the velocity of the end effector of the robot in two modes: (a) undistorted and (b) distorted.

based on bio-signals often results in unsteady data and hence, in a distorted mapping between the commanded velocity direction $\mathbf{v}_{\text{cmd}}$ and the resultant movement of the robot $\mathbf{v}_{ee}$. Therefore, all participants had to grasp the object in two modes, distorted and undistorted mapping:

$$\mathbf{v}_{ee} = \begin{cases} \mathbf{v}_{\text{cmd}} & \text{undistorted} \\ \begin{pmatrix} \gamma_{xx} & \gamma_{xy} & \gamma_{xz} \\ \gamma_{yx} & \gamma_{yy} & \gamma_{yz} \\ \gamma_{zx} & \gamma_{zy} & \gamma_{zz} \end{pmatrix} \cdot \mathbf{v}_{\text{cmd}} & \text{distorted.} \end{cases} \tag{12}$$

For the experiments, we choose $|\gamma_{i,j}| = 0.3$ for $i \neq j$ and $\gamma_{i,j} = 1$ for $i = j$. The signs of $\gamma_{i,j}$ are selected such that the distorted directions point away from the target object. The resulting distortion matrix is normalized to avoid effects on the magnitude of the velocity. Undistorted, the commanded velocity direction and the movement of the robot match, as illustrated for a velocity command in the z-axis in Fig. 10(a). In distorted mode, a command in one axis generates movements in all three translational axes (Fig. 10(b)).

Overall, three assistance conditions were tested: a no-assistance (manual) condition, and two conditions with assisted hand positioning. In the manual condition, the participant moves the end effector towards the object by commanding translational velocities, and the robot follows the commands according to the chosen mapping. When the end effector is within a distance of 0.05 m from the surface of the object, the object color changes to green, and the user may trigger the grasp to successfully finish the grasp trial. The other two conditions use assistance based on dynamic virtual fixtures. The two conditions differ only in the distance between hand and object that is allowed for triggering a grasp: in the condition "trigger-when-near", the end effector needs to enter the same area as in the "manual" condition, while in the condition "trigger-when-far", a distance of 0.13 m or smaller (i.e. entering the buffer zone) from the surface of the object is required. As soon as the hand is within the specified distance, the object color changes to green as in the manual condition, and when the grasp is activated, the motion to the target point is autonomously executed and finally the fingertips move to the planned contact points to form the grasp.

For each block of trials, for each assistance condition and each mode of mapping, subjects had to grasp the object in all three poses in increased order of difficulty, with two repetitions of the complete cycle. Altogether, a number of 6 blocks of grasp trials (3 assistance conditions x 2 modes of mapping) and 36 total grasp trials (6 blocks x 6 trials) had to be completed. The order of the assistance conditions was randomly permuted to control potential time effects like learning or fatigue. After completing 3 blocks of grasp trials with the undistorted mapping, the same order of assistance conditions was executed with the distorted mapping. Hence, all participants were already trained to use the system when they controlled it with the distorted mapping. The participant started each trial by pressing a button on the SpaceMouse, and then had a maximum of 20 s to reach the object and trigger the grasp. After each block, parts of the NASA-TLX questionnaire [20] regarding mental and physical demand of the task, as well as a specific questionnaire to evaluate naturalness of interaction, interactivity, and consistency of mapping [21] were administered.

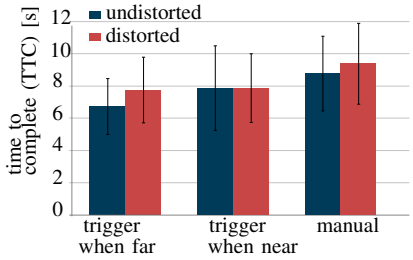


Fig. 11. Time to complete the task (TTC) for undistorted (blue) and distorted (red) control mode, in the three conditions: manual, "trigger-when-near" (autoclose), and "trigger-when-far" (autofar).

B. Results

Table I summarizes all the results. First, we present the findings using objective data, namely time to complete the task (TTC, time elapsed until the target voxel is reached), the overall translational movement of the end effector during the task (until the target voxel is reached), and the task success rate. Then, the subjective measures in terms of workload, sense of presence, consistency of interaction, and usefulness of the conditions for completing the task are reported.

1) Time to Complete (Fig. 11): A repeated measures analysis of variance (ANOVA) with *assistance* (manual vs. trigger-when-near vs. trigger-when-far) as repeated measures was performed on the TTC for each grasp, averaged over all object positions. A highly significant main effect occurred for both control modes (undistorted and distorted) ($Fs(2, 34) > 7.0$; $ps < 0.005$). Paired difference t-tests (two-sided) revealed a significantly shorter TTC for the distorted mode, when performing the task with assistance compared to no assistance ($M = 9.39$ s). The effect was higher for assistance far ($M = 7.75$ s, $t(17) > 2.88$, $p < 0.01$) than for assistance near ($M = 7.87$ s, $t(17) > 2.70$, $p < 0.015$). No effect between the assistance modes was found ($t(17) = 0.45$, n.s.- non significant). Interestingly, in the undistorted control mode only the TTC of the trigger-when-far condition ($M = 6.73$ s) is significantly shorter than the manual condition ($M = 8.77$ s, $t(17) = 3.79$, $p < 0.005$). Comparing trigger-when-near to the manual condition revealed no significant effects ($M = 7.87$ s, $t(17) = 1.84$, n.s.).

TABLE I

OVERVIEW OF EVALUATION RESULTS WITH MEANS (STANDARD DEVIATION)

	Undistorted			Distorted		
	Manual	Trigger-when-near	Trigger-when-far	Manual	Trigger-when-near	Trigger-when-far
Time to Complete, TTC [s]	8.8 (2.3)	7.9 (2.6)	6.7 (1.7)	9.4 (2.5)	7.9 (2.1)	7.7 (2.0)
Translational Movement [m]	0.8 (0.2)	0.7 (0.2)	0.6 (0.1)	0.9 (0.2)	0.7 (0.2)	0.6 (0.1)
Task Success Rate [%]	99.1 (0.2)	97.2 (0.4)	97.2 (0.4)	100 (0)	94.4 (0.6)	94.4 (0.6)
Mental Workload (very low: 0; very high: 20)	5.7 (3.8)	4.5 (3.7)	3.3 (2.9)	7.2 (4.1)	5.0 (3.5)	4.6 (3.8)
Easy to Grasp (not true: 1; true: 7)	4.3 (1.6)	5.8 (1.0)	6.2 (0.7)	4.0 (1.8)	5.2 (1.2)	5.8 (0.9)
Usefulness (very low: 1; very high: 7)	4.0 (1.5)	5.2 (1.2)	5.3 (1.5)	4.2 (1.5)	5.8 (0.8)	6.3 (0.7)

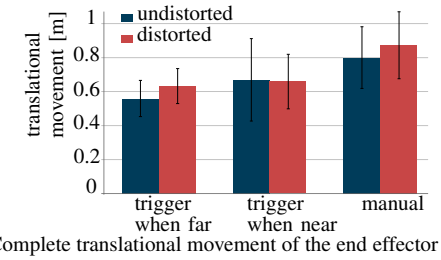


Fig. 12. Complete translational movement of the end effector for the undistorted (blue) and the distorted (red) control mode, in the three conditions: manual, "trigger-when-near" (autoclose), and "trigger-when-far" (autofar).

2) Translational Movement (Fig. 12): The analysis revealed a significant main effect of the assistance for both the undistorted ($F(2,34) = 9.96$, $p < 0.001$) and the distorted control mode ($F(2,34) = 23.45$, $p < 0.001$). In both control modes, the manual condition revealed significantly more translational movement ($M_{\text{undist}} = 0.80\,\text{m}$, $M_{\text{dist}} = 0.87\,\text{m}$) than the trigger-when-far condition ($M_{\text{undist}} = 0.56\,\text{m}$, $M_{\text{dist}} = 0.63\,\text{m}$, $ts(17) > 4.96$, $ps < 0.001$). In the distorted mode, also the trigger-when-near condition ($M = 0.66\,\text{m}$) reduced the translational movement significantly compared to the manual condition ($t(17) = 4.38$, $p < 0.001$). Between the assistance conditions, no significant effects could be found for both modes ($ts < 1.85$, n.s.).

3) Task Success Rate: Repeated measures ANOVA revealed no significant effects regarding the task success rate ($Fs < 3.3$, n.s). Sometimes participants could not finish the task within the 20 s, which happened in almost all conditions and modes. This occurred eight times in total over all 648 trials, and at most twice during the 12 grasps a participant had to execute in a specific mode. Only in distorted, manual control mode all grasp trials succeeded.

4) Workload: The mental and physical workload of the participants were separately assessed. Both measures show significant results with repeated measures ANOVA, although the effect is stronger for the distorted mode. In undistorted control mode, the effect for mental workload is $F(2,34) = 7.20$, $p < 0.005$. The mental demand in the trigger-when-far condition ($M = 3.28$) is significantly lower than in the manual condition ($M = 5.67$, $t(17) = 3.72$, $p < 0.005$) as well as in the trigger-when-near condition ($M = 4.50$, $t(17) = 2.93$, $p < 0.01$). No significant difference was found between the assistance conditions ($t(17) = 1.5$, n.s.).

The effects for the distorted control mode show the same pattern. Mental demand reduced significantly comparing the manual condition ($M = 7.22$) to the trigger-when-far ($M = 4.56$, $t(17) = 3.76$, $p < 0.005$) and the trigger-when-near condition ($M = 5.0$, $t(17) = 3.08$, $p < 0.01$). No significant difference between the assistance conditions could be found ($t(17) = 0.88$, n.s.).

Additionally, participants rated the workload in the different conditions with the following question "It was easy for me to grasp the object" (not true at all: 1; fully true: 7). In both control modes, highly significant effects are found ($F_{\text{undist}}(2,34) = 16.86$, $F_{\text{dist}}(2,34) = 14.64$, $ps < 0.001$). T-tests revealed that grasping with assistance is significantly easier compared to the manual condition (Undistorted: $M_{\text{manual}} = 4.33$, $M_{\text{t-near}} = 5.83$, $M_{\text{t-far}} = 6.17$, $ts(17) > 3.79$, $p < 0.005$; Distorted: $M_{\text{manual}} = 4.0$, $M_{\text{t-near}} = 5.17$, $M_{\text{t-far}} = 5.83$, $ts(17) > 3.60$, $ps < 0.005$).

5) Usefulness: Finally, participants assessed the usefulness with the question "Rate the usefulness of the conditions to grasp an object" (not at all: 1; very much: 7). For both control modes, highly significant effects were revealed ($F_{\text{undist}}(2,34) = 13.16$, $F_{\text{dist}}(2,34) = 40.27$, $ps < 0.001$). In the undistorted control mode, the manual condition ($M = 4.0$) was rated less useful than the assistance conditions ($M_{\text{t-near}} = 5.17$, $M_{\text{t-far}} = 5.33$, $ts(17) > 3.17$, $ps < 0.01$). The same effect was reported in the distorted mode ($ts(17) > 4.42$, $ps < 0.001$). Additionally here, the trigger-when-far condition was significantly more useful than the trigger-when-near condition ($t(17) = 4.12$, $p < 0.001$).

C. Discussion

In summary, all three hypotheses could be confirmed with the results of the user study. Performance significantly increased with assistance as shown with a decreased TTC and translational movement (H1 confirmed). The effect was significant for both assistance conditions in the undistorted mode, but only for the trigger-when-far condition in distorted mode. Thus, the higher level of assistance was especially necessary when control over the system was less intuitive (due to the distortion). No significant effect for the task success rate was found. In manual control, grasps mainly failed because participants were not familiar with the system. For the assistance conditions, failures occurred because the workspace of the arm was not taken into account when selecting the target grasp pose: the autonomy sometimes chose grasps poses for the hand that the arm could not

reach. The workload of participants decreased with the use of assistance, especially for the trigger-when-far assistance; thus, H2 is supported. Also, H3 can be confirmed: the positive effects of assistance are stronger for the distorted control mode, as detected especially with the TTC and translational movement. In the subjective measures, particularly with usefulness, participants also found assistance more important for the distorted mode. There, the assisted motion is significantly more consistent to their expectation, in contrast to the undistorted mode. In general, the study confirmed a higher workload of the user with a distorted mapping, which could be significantly reduced using assistance for hand positioning. These findings stress the importance of assistance in the real setup, as signals obtained from a bio-HMI are often distorted.

V. CONCLUSION

This paper introduced a new approach to provide semi-autonomous grasping capabilities for an assistive robotic arm. In the target application, the user would control the system via a bio-signal based HMI using continuous velocity commands. As the user should freely define the task, the main goal of the approach is to provide a flexible and transparent grasp support functionality. This is achieved by using a densely sampled grasp database, namely the graspability map. The target grasp is selected from this database according to functional aspects (grasp quality, size of the contact regions) and the perceived naturalness of the movement of the robotic hand (overall rotation angles). In order to move the robot to the selected grasp configuration, we provide assistance to the user by means of virtual fixtures. In our approach, the virtual fixtures are adaptively created by searching through the graspability map for the most suitable grasp according to the commanded velocity and approach direction given by the user. The geometric constraint for the virtual fixture is a simple point constraint, i.e., the target grasp. One of the advantages of the proposed approach is that this point is not static, but allows deviations from the desired path and even retraction from the object. This gives operational flexibility to the user, while at the same time provides him assistance for reaching the target grasp. In the presented user study, grasp failures occurred due to unreachable grasp target points; this problem will be tackled by including a reachability check in the pose selection algorithm, so that non reachable grasps are discarded at runtime. Ruling out grasp configurations that cannot be reached due to collisions with the environment can be achieved using the same technique. The next step in this line of research is using the proposed method on the real robotic system driven via bio-signal based HMIs. The attached video shows some examples in simulation environment, as well as a preliminary application of the approach to the real robot.

REFERENCES

[1] V. Maheu, P. Archambault, J. Frappier, and F. Routhier, "Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities," in *Proc. IEEE Int. Conf. on Rehabilitation Robotics (ICORR)*, 2011, pp. 1–5.

[2] J. Vogel, J. Bayer, and P. van der Smagt, "Continuous robot control using surface electromyography of atrophic muscles," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2013, pp. 845–850.

[3] L. Hochberg, D. Bacher, B. Jarosiewicz, N. Masse, J. Simeral, J. Vogel, S. Haddadin, J. Liu, S. Cash, P. van der Smagt, and J. Donoghue, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, pp. 372–377, 2012.

[4] J. Collinger, B. Wodlinger, J. Downey, W. Wang, E. Tyler-Kabara, D. Weber, A. McMorland, M. Velliste, M. Boninger, and A. Schwartz, "High-performance neuroprosthetic control by an individual with tetraplegia," *The Lancet*, vol. 381, no. 9866, pp. 557–564, 2013.

[5] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-driven grasp synthesis - a survey," *IEEE Trans. on Robotics*, vol. 30, no. 2, pp. 289–309, 2014.

[6] M. Palankar, K. 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 *Proc. IEEE Int. Conf. on Robotics and Biomimetics (ROBIO)*, 2008, pp. 348–353.

[7] E. You and K. Hauser, "Assisted teleoperation strategies for aggressively controlling a robot arm with 2D input," in *Proc. Robotics: Science and Systems (RSS)*, 2011.

[8] K. Hauser, "Recognition, prediction, and planning for assisted teleoperation of freeform tasks," *Autonomous Robots*, vol. 35, no. 4, pp. 241–254, 2013.

[9] D. McMullen, G. Hotson, K. Katyal, B. Wester, M. Fifer, T. McGee, A. Harris, M. Johannes, R. Vogelstein, A. Ravitz, W. Anderson, N. Thakor, and N. 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 Trans. on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 784–796, 2014.

[10] D.-J. 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," *IEEE Trans. on Systems, Man, and Cybernetics; Part A: Systems and Humans*, vol. 42, no. 1, pp. 2–14, 2012.

[11] K. Mulling, A. Venkatraman, J.-S. Valois, J. Downey, J. Weiss, S. Javdani, M. Hebert, A. Schwartz, J. Collinger, and A. Bagnell, "Autonomy infused teleoperation with application to BCI manipulation," in *Proc. Robotics: Science and Systems (RSS)*, 2015.

[12] M. Roa, K. Hertkorn, F. Zacharias, C. Borst, and G. Hirzinger, "Graspability map: A tool for evaluating grasp capabilities," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2011, pp. 1768–1774.

[13] S. Bowyer, B. Davies, and F. Rodriguez y Baena, "Active constraints/virtual fixtures: A survey," *IEEE Trans. on Robotics*, vol. 30, no. 1, pp. 138–157, 2014.

[14] M. Roa, K. Hertkorn, C. Borst, and G. Hirzinger, "Reachable independent contact regions for precision grasps," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2011, pp. 5337–5343.

[15] M. Roa and R. Suarez, "Grasp quality measures: review and performance," *Autonomous Robots*, vol. 38, no. 1, pp. 65–88, 2015.

[16] K. Shoemake, "Animating rotation with quaternion curves," *Proc. Conf. on Computer Graphics and Interactive Techniques*, vol. 19, no. 3, pp. 245–254, 1985.

[17] J. Vogel, S. Haddadin, B. Jarosiewicz, J. Simeral, D. Bacher, L. Hochberg, J. Donoghue, and P. van der Smagt, "An assistive decision-and-control architecture for force-sensitive hand–arm systems driven by human-machine interfaces," *Int. J. Robotics Research*, vol. 34, no. 6, pp. 763–780, 2015.

[18] H. Liu, K. Wu, P. Meusel, N. Seitz, G. Hirzinger, M. Jin, Y. Liu, S. Fan, T. Lan, and Z. Chen, "Multisensory five-finger dexterous hand: The DLR/HIT hand II," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2008, pp. 3692–3697.

[19] A. Albu-Schäffer, S. Haddadin, C. Ott, A. Stemmer, T. Wimböck, and G. Hirzinger, "The DLR lightweight robot - lightweight design and soft robotics control concepts for robots in human environments," *Industrial Robot Journal*, vol. 34, no. 5, pp. 376–385, 2007.

[20] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Human Mental Workload*, 1988.

[21] R. Scheuchenpflug, "Measuring presence in virtual environments," in *HCI International*, 2001, pp. 56–58.