A Multimodal Assistive-Robotic-Arm Control System to Increase Independence After Tetraplegia

Taylor C. Hansen, Troy N. Tully, V. John Mathews, *Life Fellow, IEEE*, and David J. Warren, *Life Senior Member, IEEE*

***Abstract*— Following tetraplegia, independence for completing essential daily tasks, such as opening doors and eating, significantly declines. Assistive robotic manipulators (ARMs) could restore independence, but typically input devices for these manipulators require functional use of the hands. We created and validated a hands-free multimodal input system for controlling an ARM in virtual reality using combinations of a gyroscope, eye-tracking, and heterologous surface electromyography (sEMG). These input modalities are mapped to ARM functions based on the user's preferences and to maximize the utility of their residual volitional capabilities following tetraplegia. The two participants in this study with tetraplegia preferred to use the control mapping with sEMG button functions and disliked winking commands. Non-disabled participants were more varied in their preferences and performance, further suggesting that customizability is an advantageous component of the control system. Replacing buttons from a traditional handheld controller with sEMG did not substantively reduce performance. The system provided adequate control to all participants to complete functional tasks in virtual reality such as opening door handles, turning stove dials, eating, and drinking, all of which enable independence and improved quality of life for these individuals.**

***Index Terms*— Assistive robotic technology, electromyography, spinal cord injury, usability study.**

I. Introduction

APPROXIMATELY 276,000 individuals in the United States live with spinal cord injury (SCI), making it the second leading cause of paralysis behind stroke [1]. Of these, nearly 60% result in tetraplegia, defined as impaired mobility of upper and lower limbs, pelvic organs, and trunk [1], [2]. Individuals with tetraplegia have indicated in surveys that restoration of hand and arm function would most improve their quality of life [3].

Various approaches have been taken to address this priority for this patient population. Typical techniques can include physical and occupational therapy, surgical procedures for the affected nerves, and cell therapy [4]. One investigational approach aims to provide control of the paretic limbs by decoding neuronal motor commands acquired proximal to the injury site. For example, intracortical microelectrodes have been used to record and interpret the neural activity associated with upper-limb movements [5], [6]. The decoding signals can be used in a brain-machine interface (BMI) to control an assistive device, such as a robotic arm [7], [8], or to bypass the lesion site and artificially produce biomimetic muscle contractions in the paretic limbs via functional electrical stimulation (FES) [9], [10]. Despite their potential, there are a number of concerns about intracortical BMIs, FES, and related methods, including invasiveness, cost, and questionable device longevity due to the foreign body response.

Another approach is remapping residual volitional functions to control assistive devices via non-invasive human-machine interfaces (HMIs). These approaches are by far more common and traditionally have included input devices such as sip-and-puff devices, chin joysticks, eye trackers, and voice control. Recently, non-invasive HMIs have gained much interest, with electroencephalography (EEG) perhaps receiving the most widespread attention [11], [12]. Other signal sources include surface electromyography (sEMG), inertial measurement units (IMUs), electrooculography, and tongue interfaces [13]. These non-invasive HMIs can be employed to control powered wheelchairs, keyboards, and computer cursors [14], [15].

More recently, assistive robotic manipulators (ARMs) such as JACO (Kinova, Quebec, Canada) and iARM (Exact Dynamics, Almere, Netherlands) have been developed to address impaired upper-limb mobility. The gripper-like end effector of an ARM can be translated in all three axes (x, y, z), rotated in all three axes (roll, pitch, yaw), and opened and closed. However, the typical ARM interface is a joystick or keyboard that precludes their use by individuals with tetraplegia.

Manuscript received 28 November 2023; revised 21 April 2024; accepted 23 May 2024. Date of publication 3 June 2024; date of current version 7 June 2024. This work was supported by NSF under Grant 1901236 and Grant 1901492. *(Corresponding author: Taylor C. Hansen.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of the University of Utah under Application No. 98851.

Taylor C. Hansen was with the Department of Biomedical Engineering, University of Utah, Salt Lake City, UT 84112 USA. He is now with Verily Life Sciences, South San Francisco, CA 94080 USA (e-mail: taylorhansen@google.com).

Troy N. Tully and David J. Warren are with the Department of Biomedical Engineering, University of Utah, Salt Lake City, UT 84112 USA (e-mail: troy.tully@utah.edu; david.warren@utah.edu).

V. John Mathews is with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97331 USA (e-mail: mathews@oregonstate.edu).

Digital Object Identifier 10.1109/TNSRE.2024.3408833

As such, some efforts have been undertaken to develop HMIs for ARM control by this patient population, but they lack widespread adoption due to issues with responsiveness, intuitiveness, and customizability [13]. Voice control has been noted as too exhausting for continuous movements, although it could be utilized for discrete, button-type functionality [16], [17]. EEG-based BMIs have had mixed results, and a common complaint is that they are too slow and unreliable [11], [16]. Wheelchair-mounted head arrays and chin joysticks have been used, but they too have become inadequate due to their limited degrees of freedom for control.

Another non-invasive control option is eye-trackers that estimate the user's 3D-gaze point to guide the ARM's end effector to the desired target [18], [19], [20]. These are ideal for ARM translation as eye-based control is inherently intuitive, but other ARM modes such as gripper orientation quickly become onerous. Other uses of eye-tracking involve button-type commands sent via gaze dwell times and winking [17], [21].

One input modality that has been underutilized in these efforts is sEMG from heterologous muscles. Many individuals with tetraplegia retain volitional muscle control of neck and shoulder muscles such as the trapezius, deltoid, platysma, sternocleidomastoid, and even biceps [22]. In the simplest implementation, the contraction of a given volitional muscle can be remapped to a degree of freedom of an ARM. An inherent advantage of sEMG is its proportionality, where the level of contraction can be mapped to a continuously variable degree of freedom, such as the openness of the ARM's gripper. The difficulty with sEMG is that of ensuring the independence of signal sources. However, several muscles, such as the trapezius and platysma, are independent enough for heterologous sEMG control [23].

One shortcoming of commonly used ARMs is that not all available functions of an ARM can be intuitively mapped to a single input source. Consequently, researchers have fused signal sources together into hybrid HMIs [24]. To combat the slow speeds of EEG-based BMIs, computer vision and eye-tracking have been explored as secondary command sources [12], [25]. Head-motion-based systems using IMUs have been paired with either sEMG or eye-trackers to increase the intuitiveness of the system and boost performance [26], [27], [28]. Augmenting residual sEMG signals with eye-tracking has previously shown promise in cursor control tasks [29]. More recently, a similar system was advantageous over sEMG alone in ARM-assisted reaching tasks [30]. It remains to be seen if more complex tasks are feasible with these disparate input modalities.

Here, we introduce a novel, modular ARM control system that uses a gyroscope, eye-tracker, and heterologous sEMG simultaneously for completing complex activities of daily living (ADLs). This multimodal system can be customized on-the-fly to accommodate a user's input-source preferences and capabilities. We implemented the system for two participants with tetraplegia and corroborated our findings with a cohort of non-disabled participants.

Fig. 1. Equipment setup for controlling the virtual reality ARM. (A) Front view showing the VR headset which was equipped with eye-tracking and a gyroscope. Pairs of sEMG electrodes were placed along the bellies of the left and right upper trapezius and platysma muscles to record muscle activity. (B) Side view showing the headphones used to give auditory cues during the experiments. (C) Back view showing the locations of the sEMG ground and reference electrodes.

II. Methods

A. Human Participants

Ten non-disabled participants (3 females and 7 males; 23.8 ± 5.2 years old) participated in the study. A 29-year-old female (P1) diagnosed with a complete C4 American Spinal Injury Association (ASIA) Grade A SCI 13 years prior due to trauma and a 21-year-old male with a complete C6 ASIA Grade A SCI two months prior due to trauma also participated. Neither P1 nor P2 could control a 3D joystick, which is commonly used to control a wheelchair-mounted ARM. (See Section V for ethics statement.)

B. Input Signal Acquisition

While seated, participants donned a Virtual Reality (VR) head-mounted display (HMD, HTC Vive Pro Eye, HTC Corporation, Xindian, Taiwan) that was equipped with a gyroscope and an eye-tracking system (Tobii AB, Danderyd, Sweden). After calibration, the headset reported head rotations in all three axes (roll, pitch, yaw) with sub-degree precision at a 90 Hz rate. The eye-tracking system estimated the gaze endpoint and provided the open or closed state of each eye at a 120 Hz rate. It was calibrated per the manufacturer's documentation, resulting in an accuracy of 0.5 to 1.1° [31].

Surface EMG was acquired from four muscles, left and right platysma at the base of the neck ("grimace" muscles) and left and right upper trapezius ("shoulder shrug" muscles). These muscles were unlikely to involuntarily contract during head movements. A pair of bipolar electrodes (Cardinal Health Inc., Dublin, OH, USA, part #H124SG) were placed on the skin overlying each muscle, parallel to the muscle action, and roughly centered over the belly of the muscle. A reference and a ground electrode were placed on either side of the spinous process of C2 (Fig. 1).

The sEMG signals were sampled at 1 kHz using a Summit Neural Interface Processor (Ripple Neuro Med, LLC, Salt Lake City, UT, USA) and filtered with a 2nd-order Butterworth lowpass filter (3 dB cutoff frequency = 375 Hz), a 6th-order Butterworth highpass filter (3 dB cutoff frequency = 15 Hz), and notch filters at 60, 120, and 180 Hz. All pairwise combinations of the 8 sEMG channels (2 per muscle) were used to create a total of 28 differential pairs. Combining these with the original 8 single-ended channels yielded a total of

36 data channels. Forward selection with a Gram-Schmidt orthogonalization step then algorithmically selected the 8 most functionally useful sEMG channels [32]. Mean absolute values (MAVs) for each selected channel were computed 30 times per second by rectifying the signals and averaging the result over 300-ms intervals, similar to our previous work [33].

A modified Kalman filter (mKF) [33] was trained to predict the normalized contraction levels (0 to 1 range) of the four muscles using MAVs. The sEMG data were recorded as participants followed commanded contraction trajectories. The commanded trajectory was provided by four bar charts, with the length of each bar indicating the desired level of contraction of the corresponding muscle. A level of 0 instructed the participant to leave the muscle at rest, and a level of 1 instructed the participant to maximally contract the muscle, which was gauged subjectively by the participant. Typically, the desired trajectory increased from rest to a maximum over 0.7 s, was held at maximum for 2 s, returned to rest over 0.7 s, and then stayed at rest for 1 s; this trajectory was repeated five times. The advantage of the mKF over a standard linear Kalman Filter is the addition of an ad-hoc activation threshold and gain as detailed in [33]. If control of the mKF was inadequate for a channel, the default gain (1.0) and threshold (0.2) were modified to provide reliable, control of each muscle as in [33]. Briefly, the default threshold was increased ad-hoc in 0.05 increments until baseline jitter was reduced to an imperceptible level when the muscles were at rest.

C. VR Environment

The VR environment was developed using Unity 2020.3.16f1 and the native Unity physics engine (Unity Technologies, San Francisco, CA, USA). The Tobii XR SDK (Tobii AB, Danderyd, Sweden) and Vive Eye and Facial Tracking SDK (HTC Vive, HTC Corporation, Xindian, Taiwan) for Unity enabled the eye-tracking capabilities of the HMD. We created three VR scenes, corresponding to three different experiments (Fig. 2). Each scene consisted of a room with a table placed in front of the participant, with the participant's arms at the height of the table; one of three tasks was arranged on the table: (1) a four-object task board; (2) a self-feeding task; or (3) a drinking task. The latter two tasks have been indicated as high-priority by the patient population [3], [34].

A virtual robotic hand (Auto Hand – VR Physics Interaction, version 2.1, 2021), controlled by HMIs (see Section II-D), was placed in each scene to simulate an ARM. ARM position was constrained to X: [−1.9, 1.9], Y: [1.8, 3.8], and Z: [2.5, 4.3] Unity units. Maximum velocities were 1 unit/s and 100 deg/s for translation and rotation, respectively, to mimic constraints of physical ARMs. Similar to most ARMs, at any given time the participants could control either the position or the orientation of the virtual hand. Mode switching was used to change between controlling position or orientation.

We also implemented four button-type functions for the VR hand: "confirm gaze selection," "grasp," "change mode," and "lock." The "confirm gaze selection" button was used to move the VR hand semi-autonomously to an object. Fixating upon an interactable VR object would be detected by the eye tracker, causing the object to be highlighted in red. If the "confirm gaze selection" button was pressed while the object was highlighted, the VR hand would move semi-autonomously towards the object. The hand would stop short of the object without making contact with it, thereby putting the participant in control of interacting with the object. The participant could change course during this automated trajectory by fixating on a new object and pressing the "confirm gaze selection" button or by directly controlling the position. To stop an automated movement, the participant could fixate away from all interactable objects and press the "confirm gaze selection" button.

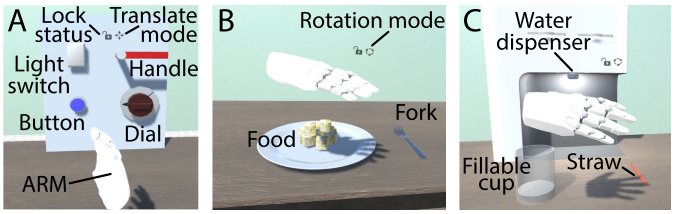

Fig. 2. VR environments were created to explore functional activities of daily living. (A) The TO-PET consisted of four tasks: a button, light switch, dial, and handle. When the participant fixates on an interactable object, it turns red, like the handle in this scene. (B) The self-feeding task required participants to pick up a gazeable fork, load the fork with food, and bring it to their mouth. (C) The drinking task involved picking up a gazeable straw, placing it into a cup, picking up the cup, filling the cup with water at the water dispenser, and bringing the cup to the mouth. ARM = assistive robotic manipulator.

The "grasp" button was used to both grasp and release objects. This button accepted proportional values between 0 (open) and 1 (closed). If the grasp button was pressed above a threshold (default 0.5) while the hand was in proximity to an object, the hand would automatically grasp the object, similar to real-life analogs in [7] and [35]. Once the object was grasped, the grasp level would be maintained, even if the "grasp" button was released. To release the object, the "grasp" button could again be activated above the threshold.

The "change mode" button enabled toggling the VR hand between translation and rotation modes. The "lock" button switched the VR hand between a locked and unlocked state. When locked, the hand was unable to move. The current hand mode and the lock status were displayed at the top of a heads-up display of the VR headset (Fig. 2).

D. Control Mapping

The four button-type functions of the VR hand could be mapped to some combination of sEMG, winking, and a hand-held game controller (HTC Vive, HTC Corporation, Xindian, Taiwan). Although the game controller excluded participants with tetraplegia, it established a performance baseline in the absence of an sEMG source. With sEMG commands, the output from the mKF for each muscle was mapped to one of the four buttons. Button-type functions are amenable to short contractions and thus avoid muscle fatigue. A given button was considered "pressed" if the mKF output exceeded 0.5 continuously for 100 – 750 ms. Since the muscles used could be activated independently, simultaneous sEMG button presses were possible. With wink commands, a button was

TABLE I
CONTROL MAPPINGS

Button/Movement	**Mapping Name** (Abbreviation) sEMG Mapping (SM)	sEMG + Winking Mapping (SWM)	Controller + Winking Mapping (CWM)*
Confirm gaze selection	L platysma (grimace)	L platysma (grimace)	Controller thumb button
Grasp (proportional)	R platysma (grimace)	R platysma (grimace)	Controller trigger
Lock	L trapezius (shoulder shrug)	L eye wink	L eye wink
Change mode†	R trapezius (shoulder shrug)	R eye wink	R eye wink
X Position (L/R)		Gyroscope yaw (rotate head L/R)	
Y Position (U/D)		Gyroscope pitch (tilt head F/B)	
Z Position (F/B)		Gyroscope roll (tilt head L/R)	
Roll		Gyroscope roll (tilt head L/R)	
Pitch		Gyroscope pitch (tilt head F/B)	
Yaw		Gyroscope yaw (rotate head L/R)	

L = left; R = right; U = up; D = down; F = forward; B = backward
*Not used by participants with tetraplegia
†Mode switching required to use gyroscope for either translation or rotation commands

"pressed" if the corresponding eye had been closed for at least 100 ms before opening again. Simultaneous commands were not permitted (*i.e.*, blinking was ignored). When needed, the thresholds for winking and sEMG button mappings were modified for each participant to provide comfortable button pressing.

The control mappings used for this work are summarized in Table I and consisted of sEMG Mapping (SM), sEMG + Winking Mapping (SWM), and Controller + Winking Mapping (CWM). Thus, in order to complete tasks with the VR hand, a combination of multiple signal modalities was required. The gyroscope was used to manually position or orient the virtual hand in space for all mappings. Movement commands were relative to the coordinate system of the participant. The speed of movements was set to be proportional to the gyroscope's angle of deflection, which was constrained between 15° and 30°. Because P1 had difficulty tilting her head, the lower threshold was reduced to 12° for her comfort.

E. Task-Oriented Performance Evaluation Tool

The first experiment used a subset of the tasks from the real-world Task-Oriented Performance Evaluation Tool (TO-PET) designed for use with ARMs [36]. The subset consisted of four of the original six tasks: a light switch, a button, a door handle, and a dial (Fig. 2A). To complete the light switch and button tasks, the participants moved the hand in close proximity to the object with a button press and then manually translated the hand forward to press the object. Hence only translations needed to be performed. For the door handle and dial tasks, the hand had to be properly oriented before grasping and twisting the object. This more complex maneuvers required switching between modes and grasping button presses. Prior to training, participants were shown a video demonstrating each of the TO-PET tasks.

The order of the control mappings (Table I) was shuffled for all participants. They were given ample time to practice hand control and develop a strategy for task completion for each mapping until they felt ready to proceed with the experiment. P2 disliked winking commands after his practice period, and requested to complete the TO-PET and subsequent experiments using only the SM.

Each participant performed a timed version of the TO-PET consisting of five attempts (trials) for each of the four tasks, for a total of 20 trials per mapping, with the order of the tasks shuffled for each repeat. All trials with a given mapping were attempted before moving to the subsequent mapping.

At the start of each trial, the VR hand was automatically reset to a default "home" position and orientation, the hand mode was set to "translate," and the participant was given a recorded auditory cue that indicated which task was to be performed. They had 45 seconds to complete the trial, with an audible warning 10 seconds before time expired. If they finished before time expired, an auditory cue was given to indicate success. Each trial, regardless of the outcome, was followed by a five-second enforced rest period.

The per-trial performance metrics for each of the four TO-PET tasks included attempt time (up to 45 s), success rate, number of "confirm gaze selection" button presses, number of "grasp" button presses, number of mode switches, number of hand locks, and participants' subjective ranking of the control mappings.

F. Self-Feeding Task

The second experiment was the self-feeding task (Fig. 2B). This task involved grasping a fork, loading the fork with a piece of food, and then bringing the food to their mouth. The task was deemed successful only if the fork was grasped and held at least one piece of food when it was brought into the proximity of the mouth. Successful completion of this task required, at a minimum, grasping one object and six mode changes between translation and rotation. For this task, the participants used their preferred control mapping from the TO-PET experiment. If the preferred mapping was CWM for non-disabled participants, the second-ranked mapping preference was used.

After viewing a video of the task and sufficient practice, the participants performed five timed trials. The hand state and movement mode were initialized as for the TO-PET task, and the fork and nine pieces of food were set to default locations on the table. The performance metrics of the task were the same as for the TO-PET tasks, except that the participants had 3 minutes to complete the task, and the number of fork

drops was measured. A fork drop occurred when a "grasp" command happened while the fork was already held (false-positive grasp command). If the fork was dropped and landed on the table, it was allowed to be recovered. If it fell off the table or all the food was pushed off the table, the trial was deemed unsuccessful.

G. Drinking Task

The third experiment consisted of a drinking task (Fig. 2C) using the same control mapping as the self-feeding experiment. Successful task completion required grasping a straw, inserting it into a cup, grasping the cup, filling it with water using a hands-free dispenser, and bringing the cup to their mouth. The task was marked successful only if the straw had been inserted into the cup, the cup was actively grasped, and the water in the cup was above a threshold level when it was brought into the proximity of the mouth. Successful completion required sequentially grasping two objects and switching modes a minimum of four times.

After viewing a video of the task and sufficient practice, participants attempted five timed trials. The hand state and movement mode were initialized as for the TO-PET task, and the cup and straw were set to default locations on the table. The performance metrics for this task were similar to the TO-PET tasks, except that the duration was up to 5 minutes and additional metrics involving the number of straw drops, the number of cup drops, and the number of cup spills were used. If the cup or straw was dropped and landed on the table, it was allowed to be recovered. If either fell off the table, the trial was deemed unsuccessful.

H. Statistical Analysis

Data from non-disabled participants were analyzed for statistical significance between control mappings using version 12.1 of the Statistics and Machine Learning Toolbox in MATLAB R2021a (The Mathworks, Inc., Natick, MA, USA). Given the small sample size of participants with tetraplegia, the data collected from these participants were not subjected to inferential statistical analysis but are instead overlaid on the figures presented in Section III.

Because all metrics showed evidence of non-normality (per Shapiro-Wilk test), all inferential statistics were calculated with non-parametric tests with a significance level of 0.05. For the TO-PET task, we used Friedman's test with five replicates per cell (MATLAB function `friedman`) for each performance metric. If a significant difference was found, post hoc multiple comparisons were performed with a Bonferroni correction to determine which control mappings were different from the others (MATLAB function `multcompare`). For the self-feeding and drinking tasks, we used a Wilcoxon rank sum test to detect significant differences between the two control mappings (SM vs. SWM) for each performance metric (MATLAB function `ranksum`). Results are reported as median (interquartile range) for all metrics and summarized visually with box plots (MATLAB function `boxplot`) with default settings. In all figures the data points in the box plots represent the median performance within each participant, the y-axis arrows indicate the direction of best performance, and statistical significance is * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

III. Results

A. Task-Oriented Performance Evaluation Tool

Participants attempted each of the four TO-PET tasks a total of five times with each of the three control mappings (sEMG Mapping (SM), sEMG + Winking Mapping (SWM), and Controller + Winking Mapping (CWM, only non-disabled participants); see Table I). We found that at least one control mapping with sEMG (SM or SWM) yielded on-par performance with the mapping using a handheld controller (CWM) (Fig. 3). For the majority of metrics, all three control mappings were statistically indistinguishable. The general exceptions were attempt time, grasp button presses, and SM against CWM for the handle success rate. This suggests that a control scheme based solely on residual volitional functions can perform as well as handheld controls.

1) Across All Control Mappings, Attempt Times Were Longer and More Variable for the Dial and Handle Than for the Button and Light Switch: (Fig. 3A). Of the four TO-PET tasks, only the two easiest tasks demonstrated a significant difference in attempt time between the CWM and SWM (button and light switch), but three of the tasks demonstrated a significant difference between CWM and SM (button, dial, and handle). There were no significant differences between SWM and SM. When using SM, median performance for the two participants with tetraplegia (P1 and P2) generally fell within the IQRs of the non-disabled participants. However, P1's attempt times with SWM were well above Q3 for all but the button task. The best attempt times for P1 and P2 were for the button using SM (6.9 s and 6.5 s, respectively). The worst times were the dial and handle with SWM for P1 (45 s each) and the handle with SM for P2 (33.6 s).

The success rates for the TO-PET tasks were similarly high across tasks and control mappings (Fig. 3B). For the more complex tasks, increasing levels of sEMG input tended to decrease the success rate, and this was significant for the handle where SM had less success than CWM ($p < 0.05$; 70% (60%) vs. 90% (20%), respectively). Both P1 and P2 had at least 80% success rates for all tasks when using SM, but P1's success rates were 20% and 0% with SWM for the dial and handle, respectively.

The number of grasp button presses increased for SM compared with CWM and was also higher for SWM during the dial task (Fig. 3C). On the two tasks that required grasping (dial and handle), all individual participants, except for P1 with SWM, had a median count of at least two presses per task. P1's count with SWM was lower because she struggled to complete each of these tasks and usually pressed the grab button at some point while trying to get the ARM into position.

The number of mode switches showed a tendency to increase with the increased use of sEMG (Fig. 3D) but the observed increase was not statistically significant among the control mappings. Notably, P1 and P2 used an ideal number of mode switches for the dial and handle tasks (1 and 2 presses, respectively) when using SM, but P1 had many more when

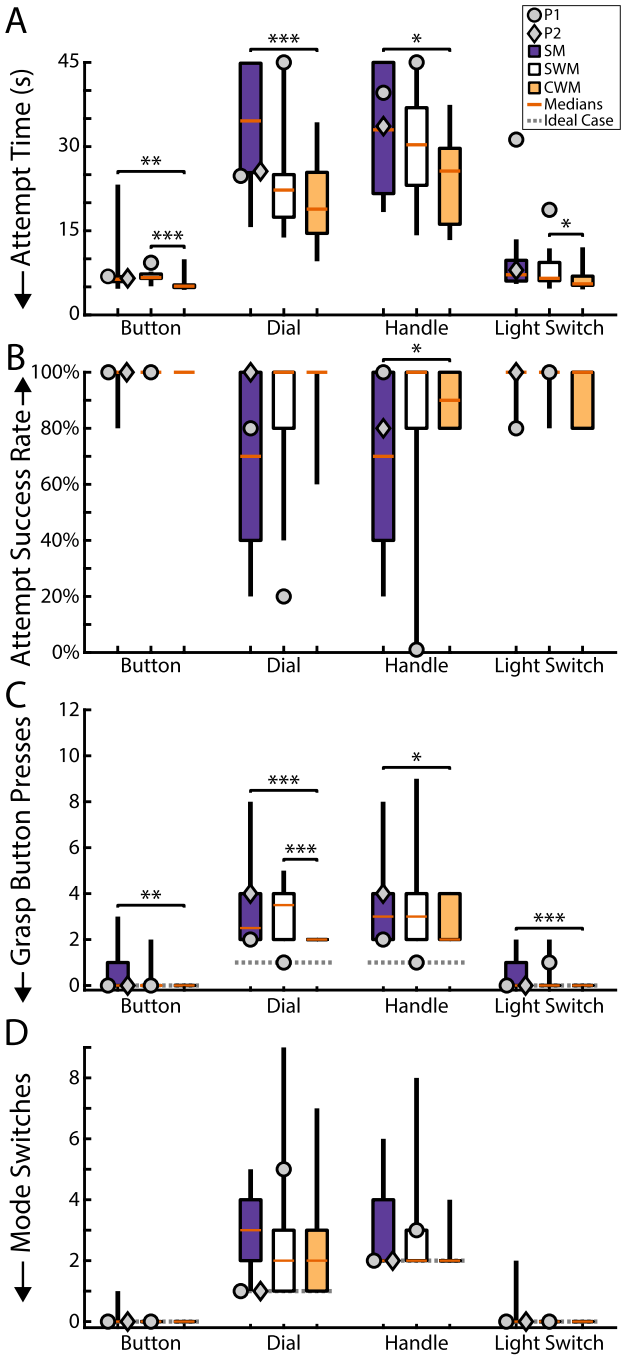

Fig. 3. Participants' outcomes with the task-oriented performance evaluation tool (TO-PET) varied with the control mapping and the task. (A) Attempt times differed between the control mappings for each task, with SM generally being longer than CWM. (B) Median success rates remained comparable across control mappings and tasks, except for the handle task, where SM was less successful than CWM. (C) For both the handle and the dial, all participants except P1 pressed the button more than necessary for an ideal trial (dashed gray line). (D) The number of mode switches did not differ significantly among control mappings. N = 10 for each box plot across all metrics. SM = sEMG; SWM = sEMG + Wink; CWM = Controller + Wink.

using SWM (5 and 3, respectively). We also recorded the number of times that participants pressed the "confirm gaze selection" button, but there were no differences noted among control mappings. This was also true for the number of times

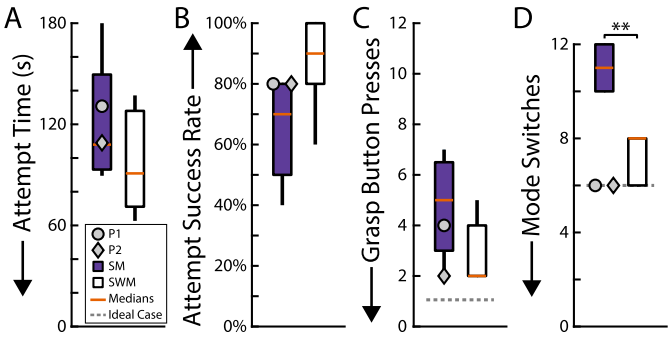

Fig. 4. Both control mappings enabled the successful completion of the multi-step self-feeding task. (A) Attempt times were comparable between SM and SWM. Most trials did not require the full time allotted. (B) SM and SWM did not have statistically different success rates, and P1 and P2 were both quite successful at completing the task. (C) All participants pressed the grasp button more than needed for an ideal case (dashed gray line), but the number of presses was similar between SM and SWM. (D) Participants using SWM used significantly fewer mode switches than those using SM, and many participants completed the task with a minimal number of mode switches (dashed gray line). N = 6 for each SWM boxplot and N = 4 for each SM boxplot across all metrics.

that participants locked or unlocked the hand with the lock button.

Following the TO-PET, non-disabled participants were asked to rank the three control mappings. Four participants ranked CWM as their favorite, three ranked SWM as their favorite, and three ranked SM as their favorite. Of the four that had CWM as their top choice, three chose SWM as their next choice, and one chose SM. Thus, for the self-feeding and drinking tasks, six participants used SWM, and four used SM. P1 and P2 were also asked to indicate their preferred control mapping after practice with the TO-PET. Both favored the maximal amount of sEMG-based control (SM) and used it to complete the self-feeding and drinking tasks, which was in line with four of the non-disabled participants.

B. Self-Feeding Task

All participants attempted to perform a multi-step self-feeding task using their preferred control mapping choice between SM and SWM. We found that both of these mappings were successful for completing this task. The only metric that significantly differed was the the number of mode switches (Fig. 4D). When considering the total attempt time for the task, the four non-disabled participants using SM took 108 s (56 s), and the six using SWM took 91 s (57 s). For comparison, P1 and P2's median attempt times were 131 s and 109 s, respectively (Fig. 4A). A real-world analog of this task performed with the native hand would take approximately 5 s.

Success rates between the two control mappings were not significantly different, although SM trended lower than SWM for non-disabled participants ($p = 0.15$; Fig. 4B). Success rates were 70% (30%) and 90% (20%) for SM and SWM, respectively. Of note, P1 and P2 both had success rates of 80% with the self-feeding task. All participants pressed the grasp button more than would be required in an ideal trial, regardless of control mapping, indicating a high number of false-positive "grasp" commands (Fig. 4C).

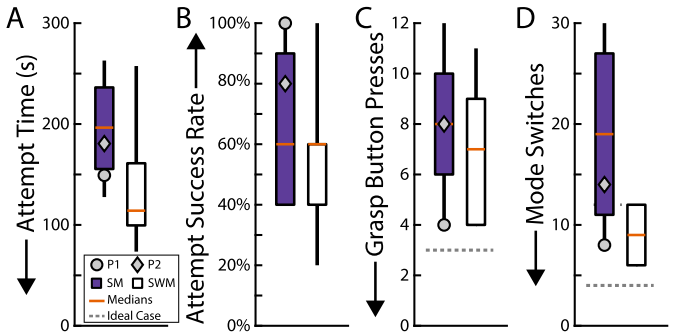

Fig. 5. Participants successfully carried out the drinking task with their preferred control mapping, but neither mapping was detectably superior. (A) Attempt times were not significantly different between the two control mappings, and the maximum times were below the allotted time of 300 s. (B) Success rates were comparable between control mappings, with P1 and P2 being top performers with SM. (C) The number of grasp button presses was nearly identical between control mappings, but both were more than an ideal case (dashed gray line). (D) The number of mode switches was more variable for SM than for SWM, but the medians were not significantly different. All participants switched modes more than required (dashed gray line). N = 6 for each SWM boxplot and N = 4 for each SM boxplot across all metrics.

The number of mode switches used by non-disabled participants was significantly higher when using SM compared with SWM ($p < 0.01$; Fig. 4D). Participants with SM switched 11 (2) times, and those using SWM switched 8 (2) times. P1 and P2 had notably fewer mode switches than their non-disabled counterparts, with a median count of six each, which matched the ideal case and was lower than the median for SWM.

Similar to TO-PET, the number of "confirm gaze selection" button presses was not significantly different between the two control mappings, although SM trended higher than SWM (5.5 (7.5) and 1 (1), respectively; $p = 0.06$). P1 and P2 had a median of 5 and 2 gaze selection button presses per trial, respectively. Across all four non-disabled participants using SM (N = 20 trials), the fork was dropped a total of 16 times or 0.8 times per trial. For the six non-disabled participant trials with SWM (N = 30 trials), the fork was dropped a total of 10 times, or 0.3 times per trial. P1 dropped the fork three times (0.6 times per trial), and P2 dropped the fork four times (0.8 times per trial). Finally, the number of times participants locked the hand with the lock button did not vary between control mappings.

C. Drinking Task

In the final experiment, participants completed a multi-step drinking task using the same control mapping as the self-feeding task. All participants were able to complete the task with their preferred control mapping, and we found no evidence for meaningful differences between the control mappings (Fig. 5). Attempt times were not significantly different for SM compared with SWM (114 s (62 s) and 196 s (81 s), respectively; $p = 0.48$; Fig. 5A). P1 and P2 completed the task with median times of 149 s and 181 s, respectively, using SM. All participants had median attempt times below the 300 s of allotted time. For comparison, a similar real-world task would take approximately 10 s to perform.

Median success rates between SM and SWM were statistically indistinguishable from one another for non-disabled participants (60% (20%) and 60% (50%), respectively; Fig. 5B). Notably, P1 had a 100% success rate with the drinking task, and P2 was successful in four of the five trials. His one missed trial was due to a cup drop right as he was bringing it to his mouth. As with the self-feeding task, all participants pressed the grasp button more than necessary, regardless of control mapping (Fig. 5C). However, the number of presses was not significantly different between SM and SWM (8 (4) and 7 (5), respectively; $p = 0.75$). P1 had a low median number of grasp button presses of four, which was one more than ideal. P2, on the other hand, had a median count of eight.

SM did not have a significantly different number of mode switches than SWM, but the IQR was larger and the number trended higher for SM (Fig. 5D; $p = 0.09$). Non-disabled participants using SM switched 19 (8) times, and those using SWM switched 9 (6) times. Both P1 and P2 switched modes fewer than the majority of their non-disabled counterparts, with a median count of 8 and 14, respectively. All participants switched modes more than needed for an ideal trial.

As with the first two experiments, the number of "confirm gaze selection" button presses did not significantly differ between control mappings (5.5 (8.5) for SM vs. 3 (5) for SWM; $p = 0.59$). The median number for P1 and P2 was 2 and 3, respectively, which was at or lower than the median for SWM). As in all the experiments, the number of hand locks was not different between SM and SWM.

The number of times the straw was dropped was recorded, and we did not observe statistical differences between the control mappings. In an ideal case, the straw would have been dropped only once when it was inserted into the cup. Additional drops indicate either missing the cup or dropping it unintentionally before arriving at the cup. For non-disabled participants using SM, for all trials (N = 20), the straw was dropped 35 times, or 1.8 times per trial. For the six participants using SWM, across all trials (N = 30), the straw was dropped a total of 47 times, or 1.6 times per trial. P1 dropped the straw a perfect 5 times, or once per trial. P2 dropped the straw once more than needed for a total of 6 times, or 1.2 times per trial.

The number of times the cup was dropped was recorded for each trial and did not differ significantly between control mappings. Ideally, the cup would have never been dropped, although a dropped cup resulted in a failed trial only if it fell off the table. With SM (N = 20 trials), the cup was dropped 13 times, or 0.7 times per trial. With SWM (N = 30 trials), the cup was dropped 45 times, or 1.5 times per trial. P1 dropped the cup three times (0.6 per trial), and P2 dropped it five times (once per trial).

Finally, we counted the number of times the water was spilled from the cup, requiring it to be refilled. This occurred eight times with SM (0.4 per trial) and three times with SWM (0.1 per trial). Both P1 and P2 spilled the water once during their five trials, or 0.2 spills per trial.

IV. Discussion

This study demonstrated that a gyroscope, eye-tracking, and heterologous muscle sEMG can be successfully combined and controlled simultaneously in a non-invasive HMI for ARM control during six ADL tasks. In particular, participants with

tetraplegia had $\geq$ 80% success rates with these increasingly complex tasks when they used their preferred control mapping. The evidence presented here suggests that the success of such a system depends on being customizable to the individual and motivates incorporating all three input signal modalities into the ARM. To our knowledge, this is the first time these three input modalities have been combined into a single HMI. We designed the system to be entirely modular such that custom control mappings can be assigned based on the user's abilities and preferences. Although some significant differences were found between control mappings for the TO-PET, these differences largely dissipated in the other tasks when participants were able to choose the control mapping of their preference and allotted task times were longer. This indicates that a system that can incorporate disparate control inputs based on residual volitional functions gives the advantage of increased customizability, thereby allowing each individual to competently wield a preferred control mapping for increasingly complex tasks, even after minimal training. We expect that with increased use, an individual's performance with the system would continue to improve.

Although the present system is a successful proof-of-concept, VR comes with inherent limitations, such as the obvious question of how well the presented results will translate to the real world [37]. With this in mind, we designed our system to utilize input sources using devices identical, or at least analogous, to those that are available in a real-world system. For example, the gyroscope could be readily replaced with a set of wireless IMUs while maintaining sufficiently accurate head orientation estimates. In particular, the real-world analog for the "confirm gaze selection" functionality is not yet commercially available, although research-grade systems for end-point detection via gaze are continually refined [18], [20], [28], [38]. A potential limitation of the present system is the reliance on single-use sEMG electrodes. A long-term solution must circumvent this constraint and simultaneously provide an accessible way for individuals to don the electrodes. One promising avenue, albeit in research stages, is fabric with embedded dry electrodes or conductive polymers that could be worn around the base of the neck, potentially in a form factor such as a scarf [39].

Other groups have explored the utility of head-based movements for control of an ARM in a research setting by those with tetraplegia [26], [27]. Using solely head-based movements requires excessive mode switching between the 2D-planes of control as well as an on-screen GUI to visualize the various ARM functions. Additionally, mode switching in such a system requires quick "head gestures" which in our experience may dislodge head-worn equipment, such as eye-trackers. Further, it may be impossible or painful to perform these head gestures for participants with SCI. The recent addition of eye-based movements to the head-movement-based system has allowed a participant with tetraplegia to perform a simple drinking task which involved using cursor control to grasp a cup and bringing it to their mouth [28]. Efforts are beginning to fuse eye-tracking and head-motion data, but implementation and evaluation in the target population is as yet unexplored by other groups [40]. Importantly, an sEMG-based trigger for certain ARM actions, such as button presses, has been recently suggested as a possible next step to improve a user's capabilities [28].

In this study, we focused on two pre-determined control mappings using a gyroscope, an eye-tracker, and heterologous sEMG that would be feasible for those with tetraplegia. For TO-PET tasks, we noted that participants typically used a similar strategy, regardless of which control mapping was utilized. They first fixated on the indicated object (e.g., dial) that was detected by the eye-tracker. They then pressed the "confirm gaze selection" button to semi-autonomously move the ARM into proximity with the object, and finally, they used the gyroscope to fine-tune the position and orientation of the ARM manipulator.

Participant P1 did quite poorly when controlling the ARM with the sEMG + Winking Mapping (SWM) because she struggled to reliably and intentionally switch modes using winking. Due to difficulty winking and having false eyelashes, her wink button presses resulted in either no mode switches or a burst of them. Her poor experience using winking for buttons and ability to perform all tasks using sEMG highlights the advantage of having alternative input modalities to choose from. Participant P2 also struggled with the winking commands using SWM and requested that he not do the timed version of TO-PET with wink-based commands. Similar to P1, he was able to perform all tasks using sEMG and his performance was usually comparable to or better than the non-disabled participants using this control mapping.

The phenomenon of inadvertent button presses was not so readily observed with P1 and P2, which suggests that they were more deliberate with their button presses. The explanation for this could be increased motivation to use the ARM system reliably, as has been suggested with other systems [41].

During practice trials for all tasks, most participants opted to set high thresholds for the hand lock button, making accidental locking extremely unlikely. During timed trials, no hand locking was observed, suggesting this functionality was not useful when the objective is to complete a task quickly. Nevertheless, we believe that such a locking function would have utility with a physical ARM to allow users to preclude inadvertent ARM movements when it is not needed as an assistive device.

Overall, more significant differences were observed during the TO-PET than during the self-feeding and drinking tasks, even for metrics that were similar among the experiments and for which significance was established with TO-PET (e.g., attempt time). This could be for three reasons. First, the complexity of the later experiments had a larger influence on performance than the control mapping. Second, the unbalanced nature of the later tasks led to reduced statistical power since there was not an equal number of participants who used each control mapping; more participants or enforced counterbalancing may have revealed additional performance differences between the control mappings instead of prioritizing insights from subjective participant preferences. Third, given the choice of which control mapping to use, participants were equally adept at these later tasks due to their own perceived ability

with the control mapping and increasing familiarity with the mapping. In other words, comparable performance between the two control mappings could simply be because the two control mappings are comparable when used by experienced and confident users.

For the TO-PET tasks, P1 and P2 frequently outperformed the majority of non-disabled participants across all metrics, with lower attempt times, higher success rates, and a lower number of inadvertent button presses. Between the two, P2's performance followed similar trends to those of the non-disabled participants, while P1's performance was unique, even for the sEMG Mapping (SM), which didn't use the winking commands that were especially troublesome for her. Another factor that explains this performance difference is her decreased range of roll neck motion. We did our best to accommodate this by reducing the angle of initiation for roll-type gyroscope movements from 15° to 12°. However, it is clear from the proportion of time spent in the "Bring to Mouth" steps of self-feeding and drinking that head tilting, required to move the object in the z-plane, was still quite difficult and slow for her. Despite this difference in ability, it did not appear to affect her motivation or her success rates.

P1's performance differences, as well as the underutilization of the lock button, indicate other button-type functions that should be explored. Foremost among them would be a "home" button to return the ARM back to a desired neutral position. The JACO ARM has such a button, but we did not explore its use in the present study. Additionally, it would be worth exploring whether a total of six buttons is feasible to control using the four muscles and individual eye winking. Furthermore, it is possible to encode additional functions as combinations of two distinct control inputs (e.g., shoulder shrug and right wink) or by "double-clicking" (e.g., two shoulder shrugs in rapid succession).

In piloting the present study, we required participants to perform extended muscle contractions for ARM control (e.g., translation commands). As expected, prolonged contractions were physically taxing, and we opted to instead map sEMG commands to control button-type functions. Subsequent participants did not raise concerns about muscle fatigue in this paradigm. We also piloted the use of heterologous sEMG from 32 single-ended electrodes which were attached to additional volitionally-controlled muscles in the neck. The control afforded by additional muscles could potentially be used for continuous control of the three axes for position or orientation, similar to the gyroscope used in the present study. We found that this extended control was challenging and frustrating, even for participants experienced with sEMG control, primarily due to high levels of crosstalk among anatomically overlapping muscles. In particular, we found that the use of the neck muscles to estimate 3D rotations of the head was unreliable, even with non-disabled subjects. Of course, another option to combat the crosstalk would be to implant electrodes into the muscle bellies [23], which is less invasive than cortical electrodes, but may not be desirable for some potential users [42].

Although this study was centered on our system's utility for ARM control, it could potentially be used to manipulate any number of assistive devices that might benefit from an adaptive and customizable input system.

V. Conclusion

We demonstrated functional control of a virtual assistive robotic arm (ARM) by two participants with tetraplegia via a novel combination of heterologous muscle signals, eye-tracking, and a head-mounted gyroscope. Similar performance was underscored by additional non-disabled participants. In our modular system, participants with tetraplegia strongly preferred the control mapping with the most sEMG control, and all participants were able to complete increasingly complex everyday tasks such as pressing buttons, opening doors, eating, and drinking. The high success rates of the participants with tetraplegia imply that they were particularly motivated by these more complex tasks. The intuition afforded to the user by this system to harness diverse residual volitional movements may drive its adoption within this patient population. Further work should be done to validate these promising results with a physical ARM using our multimodal control system.

Author Contributions

Taylor C. Hansen wrote the manuscript, designed the software system and experiments, recruited participants, and collected and analyzed the data. Troy N. Tully assisted in designing the software system and collecting the data. V John Mathews and David J. Warren oversaw the research and secured grant funding for this work. All authors contributed to the revision of the manuscript.

Ethics Review and Approval

All study procedures, including obtaining written informed consent, were conducted under a protocol approved by the Institutional Review Board of the University of Utah (Protocol No. 98851, expires 16 January 2024). Participants with tetraplegia were compensated for their time and travel.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgment

The authors would like to thank Tyler Davis and Ahmad Alsaleem for software contributions and discussions that led to the eventual design of the control system. They would also like to thank Jacob George for the use of his laboratory for these experiments.

References

[1] (2015). *Spinal Cord Injury (SCI) Facts and Figures at a Glance*. Accessed: Apr. 4, 2022. [Online]. Available: https://www.nscisc.uab.edu/Public/Facts

[2] R. J. Marino et al., "International standards for neurological classification of spinal cord injury," *J. Spinal Cord Med.*, vol. 26, no. sup1, pp. S50–S56, 2003, doi: 10.1080/10790268.2003.11754575.

[3] J. L. Collinger, M. L. Boninger, T. M. Bruns, K. Curley, W. Wang, and D. J. Weber, "Functional priorities, assistive technology, and brain–computer interfaces after spinal cord injury," *J. Rehabil. Res. Develop.*, vol. 50, no. 2, pp. 145–160, 2013.

[4] S. Thuret, L. D. F. Moon, and F. H. Gage, "Therapeutic interventions after spinal cord injury," *Nature Rev. Neurosci.*, vol. 7, no. 8, pp. 628–643, Aug. 2006.

[5] J. L. Collinger et al., "Neuroprosthetic technology for individuals with spinal cord injury," *J. Spinal Cord Med.*, vol. 36, no. 4, pp. 258–272, 2013.

[6] S. K. Wandelt et al., "Decoding grasp and speech signals from the cortical grasp circuit in a tetraplegic human," *Neuron*, vol. 110, no. 11, pp. 1777–1787, Jun. 2022.

[7] J. E. Downey et al., "Blending of brain-machine interface and vision-guided autonomous robotics improves neuroprosthetic arm performance during grasping," *J. Neuroeng. Rehabil.*, vol. 13, no. 1, pp. 1–12, Dec. 2016.

[8] L. R. Hochberg et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012.

[9] E. K. Chadwick et al., "Continuous neuronal ensemble control of simulated arm reaching by a human with tetraplegia," *J. Neural Eng.*, vol. 8, no. 3, Jun. 2011, Art. no. 034003.

[10] A. B. Ajiboye et al., "Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: A proof-of-concept demonstration," *Lancet*, vol. 389, no. 10081, pp. 1821–1830, 2017.

[11] G. Onose et al., "On the feasibility of using motor imagery EEG-based brain–computer interface in chronic tetraplegics for assistive robotic arm control: A clinical test and long-term post-trial follow-up," *Spinal Cord*, vol. 50, no. 8, pp. 599–608, 2012.

[12] H. Zeng et al., "Semi-autonomous robotic arm reaching with hybrid gaze–brain machine interface," *Frontiers Neurorobot.*, vol. 13, Jan. 2020, Art. no. 111.

[13] J. Lobo-Prat, P. N. Kooren, A. H. Stienen, J. L. Herder, B. F. Koopman, and P. H. Veltink, "Non-invasive control interfaces for intention detection in active movement-assistive devices," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, p. 168, 2014.

[14] A. Cruz, G. Pires, A. Lopes, C. Carona, and U. J. Nunes, "A self-paced BCI with a collaborative controller for highly reliable wheelchair driving: Experimental tests with physically disabled individuals," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 2, pp. 109–119, Apr. 2021.

[15] W.-K. Tam, T. Wu, Q. Zhao, E. Keefer, and Z. Yang, "Human motor decoding from neural signals: A review," *BMC Biomed. Eng.*, vol. 1, no. 1, pp. 22–43, Dec. 2019.

[16] A. Graser et al., "A supportive FRIEND at work: Robotic workplace assistance for the disabled," *IEEE Robot. Autom. Mag.*, vol. 20, no. 4, pp. 148–159, Dec. 2013.

[17] B. Noronha, S. Dziemian, G. A. Zito, C. Konnaris, and A. A. Faisal, "'Wink to grasp'—Comparing eye, voice & EMG gesture control of grasp with soft-robotic gloves," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, Jul. 2017, pp. 1043–1048.

[18] Y. L. Cio, M. Raison, C. L. Ménard, and S. Achiche, "Proof of concept of an assistive robotic arm control using artificial stereovision and eye-tracking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 12, pp. 2344–2352, Dec. 2019.

[19] A. Shafti and A. A. Faisal, "Non-invasive cognitive-level human interfacing for the robotic restoration of reaching & grasping," in *Proc. 10th Int. IEEE/EMBS Conf. Neural Eng. (NER)*. Washington, DC, USA: IEEE Computer Society, May 2021, pp. 872–875.

[20] M.-Y. Wang, A. A. Kogkas, A. Darzi, and G. P. Mylonas, "Free-view, 3D gaze-guided, assistive robotic system for activities of daily living," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 2355–2361.

[21] Y. S. Pai, T. Dingler, and K. Kunze, "Assessing hands-free interactions for VR using eye gaze and electromyography," *Virtual Reality*, vol. 23, no. 2, pp. 119–131, Jun. 2019.

[22] L. Fonseca et al., "Assisted grasping in individuals with tetraplegia: Improving control through residual muscle contraction and movement," *Sensors*, vol. 19, no. 20, p. 4532, Oct. 2019.

[23] M. R. Williams and R. F. Kirsch, "Evaluation of head orientation and neck muscle EMG signals as three-dimensional command sources," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 25, 2015.

[24] B. Z. Allison et al., "Toward smarter BCIs: Extending BCIs through hybridization and intelligent control," *J. Neural Eng.*, vol. 9, no. 1, Feb. 2012, Art. no. 013001.

[25] Y. J. Kim et al., "Vision-aided brain–machine interface training system for robotic arm control and clinical application on two patients with cervical spinal cord injury," *Biomed. Eng. OnLine*, vol. 18, no. 14, pp. 1–21, Dec. 2019.

[26] C. Lauretti et al., "Comparative performance analysis of M-IMU/EMG and voice user interfaces for assistive robots," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*. Washington, DC, USA: IEEE Computer Society, Jul. 2017, pp. 1001–1006.

[27] N. Rudigkeit and M. Gebhard, "AMiCUS—A head motion-based interface for control of an assistive robot," *Sensors*, vol. 19, no. 12, p. 2836, Jun. 2019.

[28] S. Stalljann, L. Wöhle, J. Schäfer, and M. Gebhard, "Performance analysis of a head and eye motion-based control interface for assistive robots," *Sensors*, vol. 20, no. 24, p. 7162, Dec. 2020.

[29] C. A. Chin, A. Barreto, J. G. Cremades, and M. Adjouadi, "Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities," *J. Rehabil. Res. Develop.*, vol. 45, no. 1, pp. 161–174, Dec. 2008.

[30] E. A. Corbett, N. A. Sachs, K. P. Körding, and E. J. Perreault, "Multimodal decoding and congruent sensory information enhance reaching performance in subjects with cervical spinal cord injury," *Frontiers Neurosci.*, vol. 8, May 2014, Art. no. 82873.

[31] *VIVE Pro Eye Specs & User Guide*. Accessed: Apr. 4, 2022. [Online]. Available: https://developer.vive.com/resources/hardware-guides/vive-pro-eye-specs-user-guide/

[32] J. Nieveen, M. Brinton, D. J. Warren, and V. J. Mathews, "A nonlinear latching filter to remove jitter from movement estimates for prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2849–2858, Dec. 2020.

[33] J. A. George, T. S. Davis, M. R. Brinton, and G. A. Clark, "Intuitive neuromyoelectric control of a dexterous bionic arm using a modified Kalman filter," *J. Neurosci. Methods*, vol. 330, Jan. 2020, Art. no. 108462.

[34] C.-S. Chung, H. Wang, and R. A. Cooper, "Functional assessment and performance evaluation for assistive robotic manipulators: Literature review," *J. Spinal Cord Med.*, vol. 36, no. 4, pp. 273–289, Jul. 2013.

[35] T. C. Hansen, M. A. Trout, J. L. Segil, D. J. Warren, and J. A. George, "A bionic hand for semi-autonomous fragile object manipulation via proximity and pressure sensors," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 6465–6469.

[36] C.-S. Chung, H. Wang, M. J. Hannan, D. Ding, A. R. Kelleher, and R. A. Cooper, "Task-oriented performance evaluation for assistive robotic manipulators: A pilot study," *Amer. J. Phys. Med. Rehabil.*, vol. 96, no. 6, pp. 395–407, 2017.

[37] J. Xie and X. Hu, "Virtual reality for evaluating prosthetic hand control strategies: A preliminary report," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Nov. 2021, pp. 6263–6266.

[38] D. Kim et al., "Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view," *Sci. Robot.*, vol. 4, no. 26, Jan. 2019, Art. no. eaav2949.

[39] C. M. Vidhya, Y. Maithani, and J. P. Singh, "Recent advances and challenges in textile electrodes for wearable biopotential signal monitoring: A comprehensive review," *Biosensors*, vol. 13, no. 7, p. 679, Jun. 2023.

[40] L. Wöhle and M. Gebhard, "Towards robust robot control in Cartesian space using an infrastructureless head-and eye-gaze interface," *Sensors*, vol. 21, no. 5, pp. 1–28, 2021.

[41] D. P. McMullen et al., "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. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 784–796, Jul. 2014.

[42] C. H. Blabe, V. Gilja, C. A. Chestek, K. V. Shenoy, K. D. Anderson, and J. M. Henderson, "Assessment of brain–machine interfaces from the perspective of people with paralysis," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 043002.