NON-INVASIVE HAND FREE CONTROL OF A ROBOTIC ARM

By

Derek Gin

A Thesis Submitted to the Faculty of the

DEPARTMENT OF BIOMEDICAL ENGINEERING

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In the Graduate College

THE UNIVERSITY OF ARIZONA

2022

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by: Derek Gin
titled: Non-Invasive Hand Free Control of a Robotic Arm

and recommend that it be accepted as fulfilling the thesis requirement for the Master's Degree.

Andrew Fuglevand — Date: Dec 13, 2022

Erika D Eggers
Erika Eggers — Date: Dec 14, 2022

Nima Toosizadeh
Nima Toosizadeh — Date: Dec 13, 2022

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

Andrew Fuglevand — Date: Dec 13, 2022
Thesis Committee Chair
Department of Physiology

Table of Contents

Contents

Abstract

Thousands of spinal injuries occur every year, resulting in total or partial quadriplegia. This debilitation robs people of their autonomy, in addition to the large financial cost of care and lost work. Some devices to return some autonomy exist, however, most of them require some movement or involve pre-programmed actions for a set environment. This experimental study seeks to design and compare novel methods of controlling a robotic arm pointer to a baseline, hand-controlled, mode. The modalities compared include a heads-position using a motion sensor to directly map the tip of the robotic arm to the position of the head, head velocity to create a vector to control the direction and speed based on the position of the head, and voice control which causes the tip to move in a singular direction or stop based on specific vocal input. Head position was found to be similar to hand control, performing significantly better than head velocity and voice control when observing movement time to target and throughput. Path length saw no significant differences between baseline and the three experimental modalities, and the NASA TLX showed a noteworthy dislike for head velocity mode. While this study lacks any form of gripping mechanism, it lays the groundwork for head position mode to be a novel method of control, for individuals with partially or totally limited body movement.

Introduction

Each year, approximately 17,000 new cases of spinal cord injuries (SCI) are identified (National, 2020). Forty-seven percent of these cases result in partial quadriplegia, and another 13 percent result in total quadriplegia (National, 2020). Of those injured, less than 1 percent makes a full recovery (National, 2020). SCIs are debilitating and result in a complete reliance on caretakers for basic activities of living, such as eating and bathing, leading to a sharp decrease in quality of life (QoL) and loss of autonomy. This leads to a massive financial burden to both the affected individual and society due to healthcare costs and loss of productivity. Fortunately, there have been promising assistive technologies to mitigate the effects. Previous surveys have indicated that people affected by quadriplegia desire devices that will assist with movement (Orejuela-Zapata et al. 2019). Examples include smart wheelchairs, computer use, and environmental manipulation with a device such as an assistive robotic manipulator (ARM).

Because control of ARMs requires the hand to manipulate a joystick, they are unavailable for use by many high-level tetraplegics. This obstacle could be partially overcome if there were alternate means by which a user could convey desired actions to a robotic device. One notable approach is to use brain computer interfaces (BCI), which uses an array of electrodes implanted in the brain to detect user intention and translate it into movement of an assistive device, such as an ARM (Collinger et al. 2013). Drawbacks of BCIs include the inherent risks that accompany the need for surgical implantation, and reduced efficacy of the implant over time.

Most high-level tetraplegics, however, retain the ability to voluntarily move their head and can speak. Therefore, the present study sought to evaluate head movements and voice to control an ARM in a reaching task. Two types of head control were evaluated: position and velocity. In position control, the position of the head was mapped onto the position of the robot arm - like a computer mouse dictating cursor position on a screen. In velocity control, the position the head controlled the velocity of the ARM - like a gas pedal in a car. As far as we are aware, no previous studies have evaluated these different control types in the context of assistive robotic arm control. Voice uses discrete rather than continuous control of ARM movement. For benchmark comparison, we also had subjects perform ARM reaching using their own hand. These methods were compared using objective measures for efficacy that can easily be replicated in other studies.

Methods

Six able-bodied individuals (3 female) participated in this study ranging in age from 25 to 65 years old. Each participant completed the same reaching task using four control methods: head position, head velocity, voice, and hand position. Ideally these were conducted on four

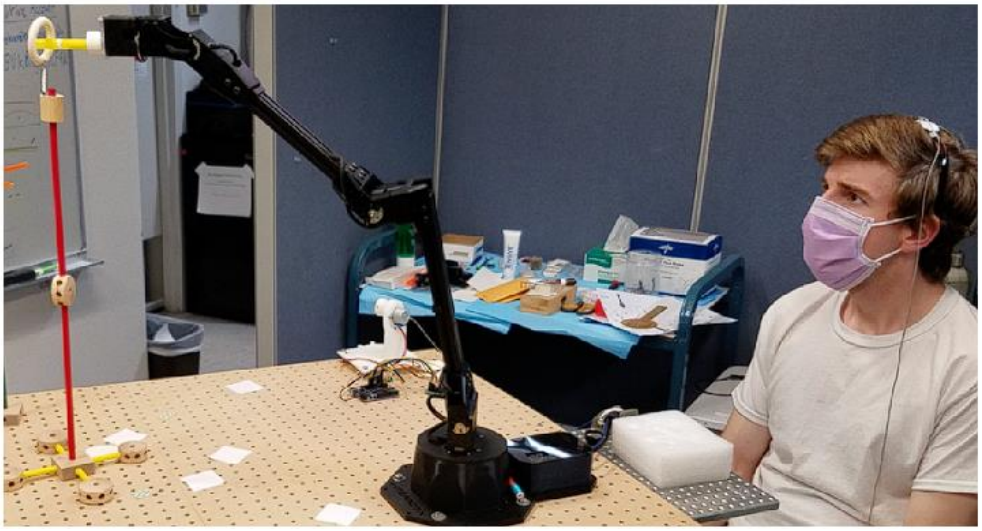

Figure 1: Example setup for head control, including target, arm, and headband.

separate days, one for each modality with the head position and velocity control on the same day. Due to time constraints, two participants completed the head controls and the voice control on the same day, totaling three sessions instead of four. Participants sat at a table behind and outside the reach of a Trossen ViperX 300 brand Robotic Arm (Figure 1). The robotic arm (75 cm maximum reach) was mounted at chest height and programmed to move with 4 degrees of freedom within a predetermined area. The gripper was removed and replaced with a pointer (~ 4 cm length). At the beginning of each session, sensors were applied to the subject and a brief calibration period was used to map actions of the head or voice onto movements of the robotic arm (see below). Subjects were given up to five minutes to familiarize themselves with each method prior to its testing.

For each control modality, subjects performed a set of 72 trials, involving 6 reaches to 12 different targets presented in an initially randomized order, which was then set to be identical across participants and sessions for consistency. Physical targets (2.5-cm diameter wooden rings) were mounted on supports of different heights. Target locations varied in height, medial-lateral, and forward-back positions spanning the reach space of the robot arm (Figure 2). On each trial,

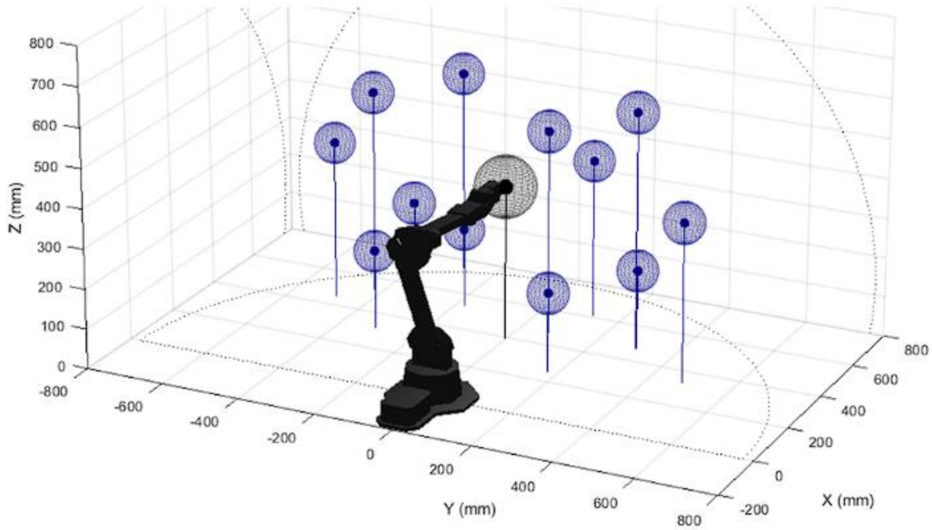

Figure 2: A graphical representation of the robotic arm (solid black object), the location of each target (blue spheres), and the home position (Black sphere)

the experimenter placed a target at one of the designated locations. For a successful reach, the tip of the pointer needed to be situated and held for 1 s within a virtual 10-cm diameter sphere centered in the middle of the ring, whereupon a tone was played indicating that the target was acquired. If the target was not acquired within one minute the trial was scored a failure. The robot then automatically moved to the next target location to assist the experimenter in placing the target ring at the correct location. The robot then returned to the start position. A tone was then played, and the subject commenced the next trial.

After completing a set (~ 30 - 45 minutes), participants completed the NASA Task Load Index (TLX) survey, which involves subjective numerical scoring of: mental demand, physical demand, temporal demand, performance, effort, and frustration (Hart et al. 1988). These scores are averaged with a lower score being considered preferable.

Hand Position Mode

For the hand position mode two Polhemus sensors, a small sensor detecting 6 movement degrees of freedom, were used to detect movement. One sensor was fastened to the back of the hand, while a second was attached to the shoulder. After the sensors were attached, the maximum reach of the robotic arm was calibrated with the extended arm, while the home location was the halfway point between the shoulder and the tip of the extended arm. The robotic arm would then mirror the movements of the hand. Movement of the hand is natural and intuitive, allowing this modality to be utilized as a benchmark comparison for the other modalities. A hand velocity mode was considered as well, but preliminary results showed that its performance was substantially inferior compared to the hand position mode. Therefore, it was not included in the study. Participants completed modalities in the order of hand, head, then voice. While the study originally intended to have the order of modalities randomized, time constraints and a delayed readiness of some control methods led to the consistent ordering of the methods. Half of the participants completed the head position mode before head velocity, while the other half completed head velocity prior to head position mode.

Head Control Mode

Both head control modes utilized the same Polhemus sensors as the hand control, where one was attached to the top of the head via a headband worn by the participant (see Figure 1), while the other was fixed to the neck near the C7 vertebrae. To calibrate the system for head control, the participant was asked to move their head around, including protracting and retracting, before coming to rest at a comfortable middle position where their head was neither protracted or retracted. In addition, the participant was shown a marker on the screen, which they used to ‘paint’ in a circle. The marker denoted a visual representation of the pitch (nodding the head up and down) and yaw (rotating the head side-to side) angles of the head.

Rotation angles of the head about the yaw and pitch axes were used to characterize the orientation of the robot arm (Figure 3). The extent of the reach of the arm was determined by the magnitude of protraction and retraction of the head. For the head position control, a custom

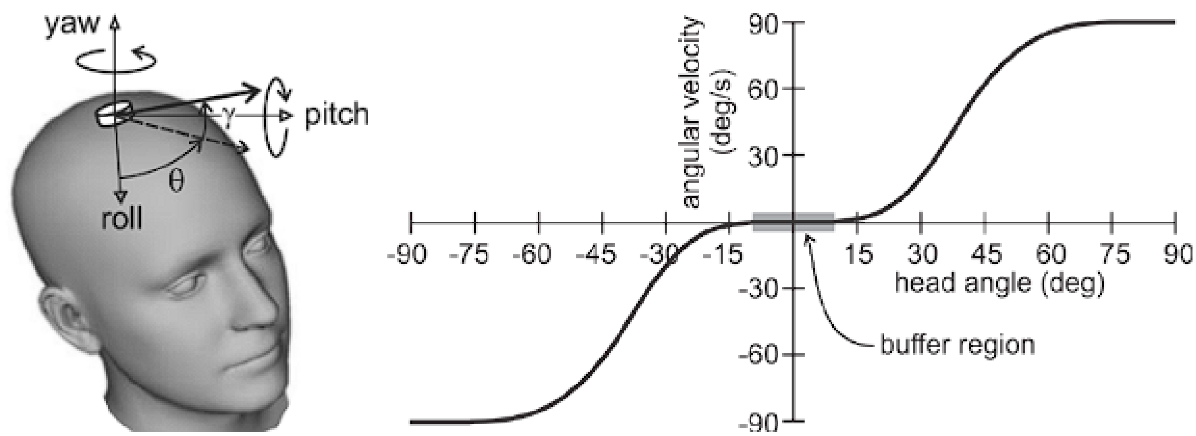

Figure 4: Left side (A) indicates the planes for yaw, pitch and roll relative to the human head. Right side (B) shows the velocity of the robotic arm relative to the change in angle of the controlling head.

Matlab script performed inverse kinematics to continuously convert the head angles and degree of pro-retraction into the associated joint robot arm angles needed to drive the pointer on the robotic to the designated 3D location. In essence, subjects were instructed to simply point their nose at the target and to push out or pull back their head along the pointing direction to control the extent of arm reach.

In velocity mode, the head acted like a “joystick,” where the rotation of the head and degree of pro-retraction affected the velocity of the robot pointer. Figure 3B shows the relationship between the change of angle and the velocity of the pointer. The neutral position of ± 10 degrees existed as a buffer to prevent unintended movement from slight shifts, and larger head movements resulted in faster movements up to a maximum velocity. This allowed for quicker movement over longer distances, and slower, more precise movements as the pointer approached and entered the target area.

Voice Control

Participants used single-word commands via microphone to control the robotic pointer. The voice signals were delivered and translated by the VoiceBot application, which locally utilizes Windows Voice Recognition. Once the participant spoke, it would be identified and delivered to a MatLab application to determine the next action of the robotic arm. Predefined words were “left,” “right,” “up,” “down,” “in,” “out,” and “stop.” Unlike the previous modalities, voice control could only move the robot arm in a single Cartesian direction at any given time. Movements would continue until the participant gave the “stop” command, or the arm extended to its maximum reach.

Data Analysis

The primary outcomes included movement time, path efficiency, and throughput. Movement time was the time elapsed from the start cue until the success or time-out cue. Path efficiency was calculated based on the robotic arm trajectory. The minimum possible distance was divided by the actual distance traveled and averaged for each modality. Throughput was calculated based on the Index of Difficulty (ID) relative to the movement time. ID was calculated as $\log_2((D/W) +1)$, where D is the straight-line distance between the two targets, and W is the width of the target. For this study, the straight-line distance was considered to be where the robotic arm left the home position area to the entry point of the target upon success as seen in the black line of Figure 4, and the red line is an example of the actual distance traveled. While determining D may vary, throughput allows for standardized comparison across other studies without having the exact same setup.

All test measures were taken, and a repeated measures one-way ANOVA test was performed for each outcome using each control modality as a factor to determine whether any of the modalities

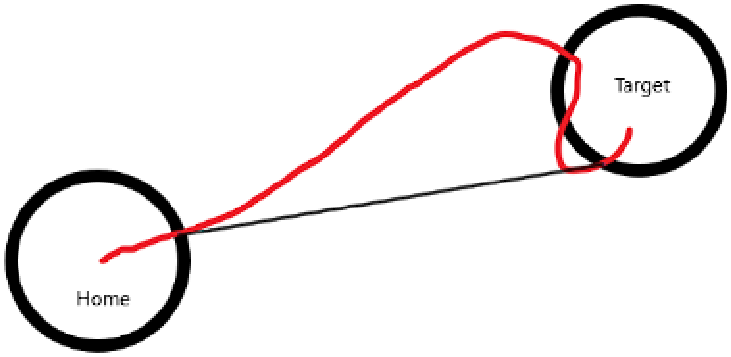

Figure 4: Example of how path efficiency is calculated. Red line indicates example movement from the participant. Black line indicates the shortest possible path from start to finish location. Distance of black line divided by the red line's exit from home to the final entrance into the target finds path efficiency.

performed significantly better than the others. The ANOVA test was performed on the NASA TLX results, as well.

Results

All six participants completed all four modalities. Examples of robotic arm control for each modality can be seen in Figure 5. In the left column (panels A, C, E, G) the magenta lines indicate the control inputs, while the black lines indicate the displacement of the robot endpoint from the starting position for each axis. All modalities, except for voice control, allow for multi-axis movements. On the right side of Figure 5, B, D, F, H are visualized 3D samples of the path taken by participants. This is most easily observed in figure 5E, voice control. The single command input is seen as a stepwise change in the magenta line, with the first command being 'right'. The solid black line depicts the movement of the arm, moving at a consistent speed in the specified direction until a new command is given. The new 'down' command ends the 'right' command returning it to baseline, while the black line then holds its position depicting the displacement.

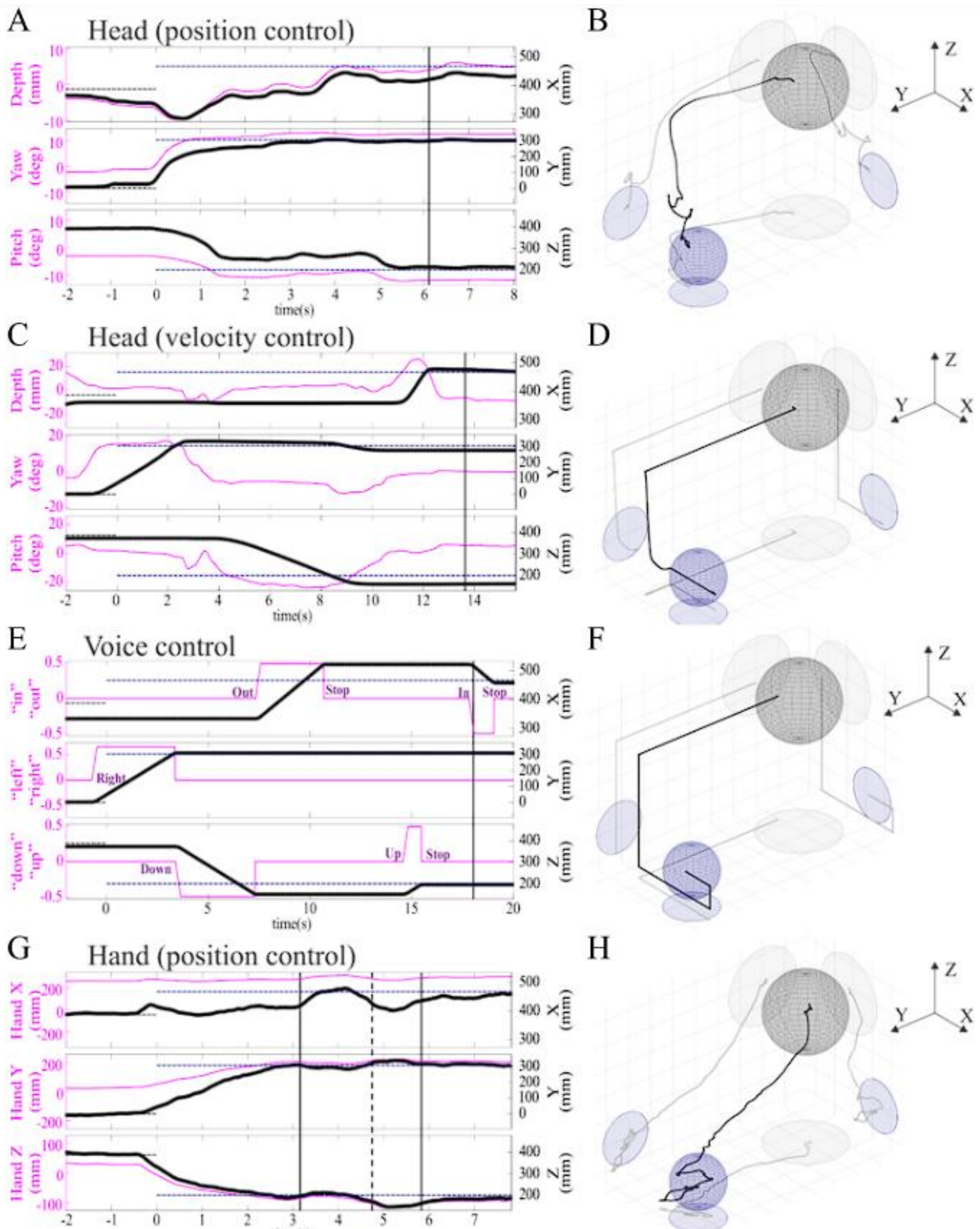

Figure 5: Left Column (ACEG) shows change in movement for a singular direction. Magenta line indicates the current input, depending on the type of control. The black line indicates the total displacement from the home/origin in that particular direction. Right column (BDFH) shows real 3D examples of participant movement paths from home to target for each control modality.

This continues until each dimension gets closer to the target displacement, represented by the dotted line. The black vertical line shows where the robotic arm end has entered the target area by correcting an overshoot with the command 'in.' The 'stop' command is then given, which halts all movement as the arm rests within the target area.

Movement time resulted in hand and head position modes having similar scores, and both significantly better compared to head velocity and voice control, p-value < 0.001. Movement time also benefited from practice, as can be seen between the first half and second half of the values. The practice effect was significant across all modalities, p-value < 0.001.

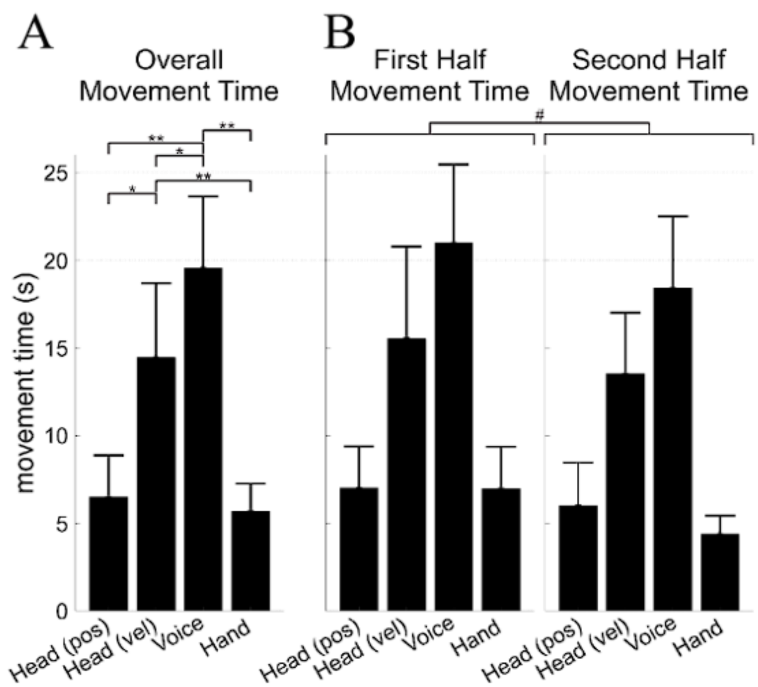

Figure 6: Average movement time across all subjects for each control mode. Section A shows the overall times. Section B shows the differences with the first and second halves of the targets are split, in case of practice effects. Error bars are ±1 SD, * p < .06, ** p < 0.001, # p < 0.001 comparing the two halves.

Path efficiency did not see any significant differences, with values of 0.55 ± 0.09 for hand, 0.46 ± 0.08 for head position, 0.51 ± 0.11 for head velocity, and 0.53 ± 0.05 for voice control with a p-value = 0.36. Throughput saw significant differences in hand control (0.33 ± 0.09 bits/s) and head position control (0.29 ± 0.07 bits/s) compared to head velocity control (0.15 ± 0.04 bits/s) and voice control (0.13 ± 0.02 bits/s) with a p-value < 0.001. NASA TLX saw a significantly higher value for head velocity at 12.76, p-value = 0.039, compared to the other three values 7.07, 6.43 and 10.36 for hand, head position, and voice control respectively.

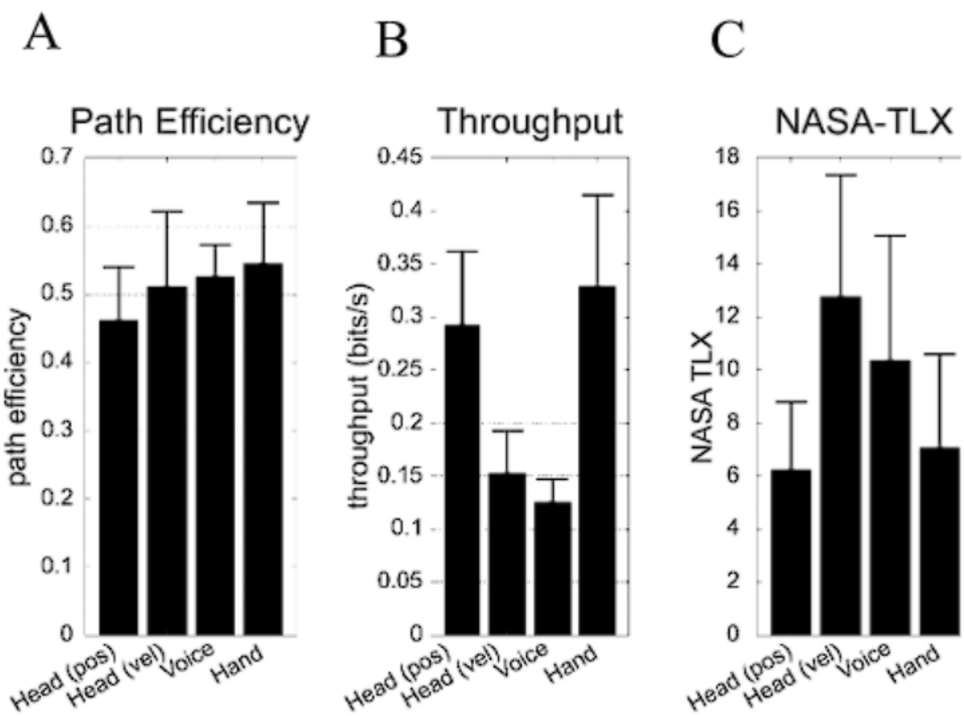

Figure 7: Compares scores for Path Efficiency (A), Throughput (B), and NASA-TLX scores (C). Error bars represent ± 1 Standard Deviation.

Discussion

Options exist for those with limited movement in their extremities for robotic arms to assist them in everyday activities. Unfortunately, no such options currently exist for those with complete paralysis. Here, we explored the possibility of using head movements and voice to control movements of a robotic arm that would enable those with high-level paralysis to benefit from this technology We found that head position mode was equally effective in controlling the robot arm as hand control by nearly every analysis. This method seemed intuitive in that robot arm reaching was produced in the same direction that the subject faced, resulting in it being the preferred method by participants. This can be seen per the NASA TLX score, where it was lower than even hand control. Importantly, head position control also allows for full movement, allowing its directional usage to remain flexible to the environment.

Head velocity and voice control methods were not as successful. Head velocity proved to be notably less intuitive. The buffer region allowed for a more stable method of holding still, which seemed useful once the arm neared and entered the target. Nevertheless, head velocity was found to be the least preferred per the NASA TLX. One contributor to this may be that during calibration it is imperative that the resting position is easily maintained over a long period of time. While most positions of the head do not cause discomfort in the short term, by the end of the study, some participants reported some degree of exhaustion and discomfort while in their resting position.

Voice control is widely utilized during everyday life for many people. While it may be useful under certain conditions, a few aspects prevented it from performing properly to control a robotic pointer. Expected issues that are common for voice-controlled devices were present, including the incorrect recognition of words and presence of background noise. While participants were guided to speak clearly and concisely, the recognition program appeared to perform inconsistently for some participants with accents or higher pitched voices. When using a voice-controlled phone search or turning a device on, repeating oneself is not typically an issue. During a more time sensitive activity, or an activity involving fine movements, having to repeat oneself can lead to overshooting of the robotic arm due to unresponsiveness. A second problem was the processing delay. From voice input to arm movement, or cessation of movement, was approximately between .8 and 1.2 seconds. This forced participants to preemptively give commands and predict where the arm would be once the command enacted the robotic arm’s response. Using a different, or creating a more specific and robust recognition program, could reduce these two issues. Lastly, the voice program utilized did not have a function to increase or decrease speed like the other modalities. In hindsight, it may be a useful tool to include voice controls such as “fast” and “slow.” This could allow the robotic arm to move quickly across larger areas, while adding precision once it gets closer to the target by making the delay less impactful. Speed up and slow down functions would be more flexible but may be difficult to control via voice if several speed up or slow down inputs need to be given. As it stands, voice control may be better for other processes, such as turning the robotic arm on and off, or perhaps for a gripper.

The primary difference between this study and previous studies is the increase in degrees of freedom. Wurth and Hargrove conducted a study that utilized throughput and path length as their outcomes, but they studied a myoelectric prosthesis with two degrees of freedom (Wurth et al. 2014). Lau and O’Leary conducted a comparison study between a mouth stick, sip puff device, and a tongue keypad to control a keyboard (Lau and O’Leary, 1993). This again involves a limited 2D interactive space, opposed to a fluid 3D space. Williams and Kirsch studied EMG of neck muscles vs head orientation for the control of a computer mouse, found that head orientation was significantly more effective than EMG (Williams et al. 2015). This study lends to the idea that head position control is an intuitive, non-invasive method of control. The study that is most similar to the present study was conducted by Mohammadi et al. Utilizing an intraorally mounted tongue control system, their system can interact with a 3D environment completing tasks such as pouring water (Mohammadi et al. 2021). While functional, the movement tends to be very slow. Similar to the present studies’ voice control, only one direction or process is feasible at a time. Very few studies interact with a 3D space, and those that do seek to simplify and thereby limit the functionality by only controlling a single direction at a time, or pre-programming movements and routes to a keyboard equivalent.

One limitation of this study lies within the calibration of the Polhemus system. Despite best efforts, the sensors are prone to shifts and displacement leading to an improper response from the robotic arm. This could be difficult to determine as well, as the displacement could result in a negligible difference, rendering the device completely unusable, or anywhere in between. This requires a decision from the person conducting the study on whether to continue or reset the calibration and current testing, which adds bias. Additionally, the voice control utilized was originally intended for general speech recognition. This resulted in commands being interpreted as words that are not commands (e.g., ‘down’ misinterpreted as ‘drown’.) It’s possible this issue could be resolved by utilizing more specific voice-recognition software. As mentioned within the methods section, the control modalities were not randomized. It is possible that the learning effects between entirely different methods (e.g. hand to voice) were minimal, but given the substantial learning effects that occurred within each control method it is likely there is some.

This study only lays the groundwork for which modality may be best suited for the positioning of a robotic arm in space. The addition of a gripping mechanism will be required for real world utility. Nonetheless, this study gives insight into which modalities may be best suited for basic movements. There is still further work to be done, but this advances our understanding of hands-free robotic arm control to allow people with high-level quadriplegia to experience greater autonomy in their everyday lives.

References

Binary_Fortress_Software. (n.d.). *VoiceBot*. Retrieved Dec. 7th, 2022, from https://www.voicebot.net/

Collinger, J. L., Boninger, M. L., Bruns, T. M., Curley, K., Wang, W., & Weber, D. J. (2013). Functional priorities, assistive technology, and brain-computer interfaces after spinal cord injury. *The Journal of Rehabilitation Research and Development*, *50*(2), 145. https://doi.org/10.1682/JRRD.2011.11.0213

Dziemian, S., Abbott, W. W., & Aldo Faisal, A. (2016). Gaze-based teleprosthetic enables intuitive continuous control of complex robot arm use: Writing & drawing. *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, *2016-July*, 1277–1282. https://doi.org/10.1109/BIOROB.2016.7523807

Folan, A., Barclay, L., Cooper, C., & Robinson, M. (2015). Exploring the experience of clients with tetraplegia utilizing assistive technology for computer access. *Disability and Rehabilitation: Assistive Technology*, *10*(1), 46–52. https://doi.org/10.3109/17483107.2013.836686

Gundogdu, K., Bayrakdar, S., & Yucedag, I. (2018). Developing and modeling of voice control system for prosthetic robot arm in medical systems. *Journal of King Saud University - Computer and Information Sciences*, *30*(2), 198–205. https://doi.org/10.1016/j.jksuci.2017.04.005

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Jackowski, A., Gebhard, M., & Gräser, A. (2016). A novel head gesture based interface for hands-free control of a robot. *2016 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2016 - Proceedings*. https://doi.org/10.1109/MeMeA.2016.7533744

Jain, S., Farshchiansadegh, A., Broad, A., Abdollahi, F., Mussa-Ivaldi, F., & Argall, B. (2015). Assistive robotic manipulation through shared autonomy and a Body-Machine Interface. *IEEE International Conference on Rehabilitation Robotics*, *2015-Septe*, 526–531. https://doi.org/10.1109/ICORR.2015.7281253

Johansson, K., Seiger, Å., Forsén, M., Holmgren Nilsson, J., Hartelius, L., & Schalling, E. (2018). Assessment of voice, speech and communication changes associated with cervical spinal cord injury. *International Journal of Language and Communication Disorders*, *53*(4), 761–775. https://doi.org/10.1111/1460-6984.12380

Kryger, M., Wester, B., Pohlmeyer, E. A., Rich, M., John, B., Beaty, J., McLoughlin, M., Boninger, M., & Tyler-Kabara, E. C. (2017). Flight simulation using a Brain-Computer

Interface: A pilot, pilot study. *Experimental Neurology*, *287*, 473–478. https://doi.org/10.1016/j.expneurol.2016.05.013

Kurtzke, J. F. (1977). Epidemiology of spinal cord injury. *Neurologia, Neurocirugia, Psiquiatria*, *18*(2-3 Suppl), 157–191.

Lau, C., & O'Leary, S. (1993). Comparison of Computer Interface Devices for Persons With Severe Physical Disabilities. *The American Journal of Occupational Therapy*, *47*(11), 1022–1030. https://doi.org/10.5014/ajot.47.11.1022

Lobo-Prat, J., Kooren, P. N., Stienen, A. H., Herder, J. L., Koopman, B. F. J. M., & Veltink, P. H. (2014). Non-invasive control interfaces for intention detection in active movement-assistive devices. *Journal of NeuroEngineering and Rehabilitation*, *11*(1). https://doi.org/10.1186/1743-0003-11-168

Mohammadi, M., Knoche, H., & Struijk, L. N. S. A. (2021). Continuous Tongue Robot Mapping for Paralyzed Individuals Improves the Functional Performance of Tongue-Based Robotic Assistance. *IEEE Transactions on Biomedical Engineering*, *68*(8), 2552–2562. https://doi.org/10.1109/TBME.2021.3055250

Moon, I., Kim, K., Ryu, J., & Mun, M. (2003). Face direction-based human-computer interface using image observation and EMG signal for the disabled. *Proceedings - IEEE International Conference on Robotics and Automation*, *1*, 1515–1520. https://doi.org/10.1109/robot.2003.1241806

National Spinal Cord Injury Statistical Center, Facts and Figures at a Glance. (2020). Birmingham, AL: University of Alabama at Birmingham

Noronha, B., Dziemian, S., Zito, G. A., Konnaris, C., & Faisal, A. A. (2017). "Wink to grasp" - Comparing eye, voice & EMG gesture control of grasp with soft-robotic gloves. *IEEE International Conference on Rehabilitation Robotics*, 1043–1048. https://doi.org/10.1109/ICORR.2017.8009387

Orejuela-Zapata, J. F., Rodríguez, S., & Ramírez, G. L. (2019). Self-Help Devices for Quadriplegic Population: A Systematic Literature Review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *27*(4), 692–701. https://doi.org/10.1109/TNSRE.2019.2901399

Rigby, P., Ryan, S. E., & Campbell, K. A. (2011). Electronic aids to daily living and quality of life for persons with tetraplegia. *Disability and Rehabilitation: Assistive Technology*, *6*(3), 260–267. https://doi.org/10.3109/17483107.2010.522678

Schäfer, J., & Gebhard, M. (2019). Feasibility analysis of sensor modalities to control a robot with eye and head movements for assistive tasks. *ACM International Conference Proceeding Series*, 482–488. https://doi.org/10.1145/3316782.3322774

Williams, M. R., & Kirsch, R. F. (2015). Evaluation of head orientation and neck muscle EMG signals as three-dimensional command sources. *Journal of NeuroEngineering and Rehabilitation*, *12*(1), 485–496. https://doi.org/10.1186/s12984-015-0016-6

Wurth, S. M., & Hargrove, L. J. (2014). A real-time comparison between direct control, sequential pattern recognition control and simultaneous pattern recognition control using a Fitts' law style assessment procedure. *Journal of NeuroEngineering and Rehabilitation*, *11*(1), 1–13. https://doi.org/10.1186/1743-0003-11-91