USABILITY OF A GESTURE CONTROL DEVICE FOR ROBOTIC PLATFORM OPERATION

Geoffrey White
Defence Science and Technology Group
Land Personnel Protection MSTC
Melbourne, Australia

Susan G. Hill*
US Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, Maryland, US

***Abstract*—This paper describes a study which examined the usability of a commercial-off-the-shelf (COTS) gesture recognition device, to control a small, Parrot® robotic platform. Results indicate that participants were able to easily learn and use the gesture device to control the robot for a simple movement task, although some usability issues were identified. The limitations discovered included the need for a neutral position between gestures, the lack of variable controls, the desire for more precision of control and the impact of fatigue on the participants' performance. These initial findings will inform future concept development and experimental research for human interaction with automated or autonomous technology.**

Keywords—robot; gesture control; usability; human-robot interaction

I. INTRODUCTION

Gesture is an important part of human-human communication and is an active area for research in human-robot communication. Gesture can be thought of as a movement by some part of the human body that, by itself, has meaning. Mitra & Acharya [1] define gestures as "…expressive, meaningful body motions involving physical movements of the fingers, hands, arms, head, face, or body with the intent of: 1) conveying meaningful information or 2) interacting with the environment." For example, military hand signals [2] are gestures that have specific meanings and are intended to convey information from one soldier to other soldiers.

The technology used to recognise gesture research for human-robot communication can be categorized into two main categories: instrumented recognition systems and camera-based recognition systems.[3][4][5] For this study, we used an instrumented gesture control device to capture the electromyographic (EMG) signals produced by muscle contraction in the arm to sense hand and finger movement. Forearm movement was interpreted using an inertial measurement unit (IMU).

The commercial product used for this experiment was the Myo™ armband by ThalmicLabs™ [6]. This device uses eight EMG sensors worn below the elbow on the forearm in a bracelet configuration. See Figure 1. It also uses a nine-axis inertial measurement unit (IMU), with a three-axis gyroscope, three-axis accelerometer, and three-axis magnetometer, as a means to sense arm movement, orientation and acceleration [7]. There is also vibro-tactile feedback available from the Myo armband itself.

Several research studies have used the Myo armband as a means for control for a variety of applications, including the control of a robot or drone [8][9][10][11][12][13].

Fig. 1. The Myo™ armband from Thalmic Labs [6] (image provided by Thalmic Labs™)

For example, [10] used the gesture device combined with speech to communicate with mixed ground and air robotic platforms in a search and rescue (SAR) scenario. The Myo™ armband was chosen to use in this current usability experiment as a low-cost, COTS technology that could be used to explore concepts of gestural control of a small robot.

II. METHOD

This usability study was conducted to explore the concept of using a gesture device to control a single robot and to identify areas where usability could be improved. The international standard, ISO 9241-11, provides guidance on usability and defines it as: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [15]. Three different display views were used by each participant, which allowed feedback on the impact of the display type on the usability of the gesture device. This in turn allowed examination for the ability to perform the task against the different views. The subjective assessments of the participants were of most interest, therefore performance was not quantitatively measured.

A. Participants

Fourteen volunteers participated in this usability study. All participants were from the Defence Science and Technology Group. All used their dominant arms to wear the gesture device and make gestures; only one used the left arm. Participants also were asked about their experiences with using various kinds of controls for automated systems.

*This work was conducted while the second author was an exchange defence scientist at DST Group as part of the U.S. Army Engineer and Scientist Exchange Program (ESEP).

B. Procedure

Upon completing the appropriate Human Ethics consent form, participants were given a brief introduction to the Myo armband. This included explaining how the system worked with their physiology, the technical functions of the system and how the Myo would be used for this study. The Myo was then fitted to the participants for calibration.

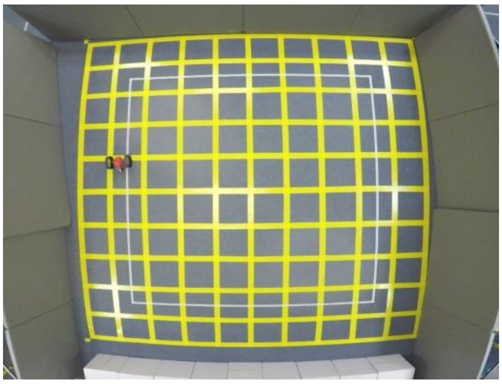

Fig. 2. The Overhead (OH) view showed the view of the environment as seen from an overhead camera. This provided a view that, for example, might be provided by an unmanned aerial vehicle (UAV).

The Myo has two different modes of calibration. The first mode is a general one which is agnostic to the human wearer. The second uses individual custom calibrations from each participant. Thalmic suggests a custom profile can improve reliability in cases where the Myo device will not be moved for the duration of the activity. In this study, a custom profile was developed for each participant. The custom profile links directly to each participant's unique EMG signals for each of the gestures used during the study. Verifying a working customised profile before running the activity, as suggested by the manufacturer, maximized the likelihood that the system would work as intended.

Upon completion of the calibration process, the user verified the operation of the device by making gestures and viewing immediate feedback on-screen as to what the Myo interpreted the human gesture to be. Directly following the profile creation, the participant tried out all of the five gestures to see if they were recognized correctly. In the few cases where gestures were not reliably recognized, a recalibration was conducted until recognition performance was satisfactory.

For practice, participants were given approximately ten minutes using a direct view (DV) of the robot to become accustomed to the operation of the system. After they were comfortable with using the system, the participants moved to a separate room, remote and with no direct visibility of the robot. Video camera equipment recorded three views simultaneously, including the Robot First Person (FP) view, the Overhead view (OH) and a view of the participants arm wearing the Myo. See Figures 2 and 3 for examples of the OH and FP, respectively. Only one view, either FP or OH was displayed to the participant at a time, using a 59 cm computer monitor. The video frames between cameras were synchronised by using a camera flash.

As mentioned, during the assessment a video camera was positioned to record a side-view, of the armband in use, referred to as the "Participant View". The video field of view was positioned so that only the torso and the gesturing arm of the person were visible; the face was not recorded. Participants were asked to control the Parrot Drone to follow the white taped square which was visible on the test environment floor using one of the display views (FP or OH). They performed this task for ten minutes.

After the ten minutes, participants were asked a series of questions to assess the usability of the system with the view just used. Following this, the participants performed the same task using the alternative view and were asked the same series of usability questions.

Fig. 3. The Robot First Person (FP) view was the view from an on-board camera showing the environment as seen by the Parrot drone.

The two displays were counterbalanced across participants – half used the OH display first; the other half used the FP display first. After both display view trials were complete, the participant was asked about the overall experiment experience. They were also asked to complete the System Usability Survey, adapted for the gesture device from the SUS by [14]. The entire experiment participation lasted about 75-90 minutes.

III. RESULTS AND DISCUSSION

In this study we were interested in exploring the concept of using gestures with this type of device. The key questions of interest were: 'what is easy?', 'what is difficult?', and 'do gestures support successful performance?' By identifying things that were usable and work well, and other things that were not usable and do not work well for human use, we were able to identify areas that are ripe for future research.

The results reported here are organized into several areas: participant experience; usability of gesture device; impact of display type; opinions on the device calibration, training and practice; and other relevant comments. The complete final quantitative results are available in the technical report on this work (White and Hill, in press.)[16]

A. Participant Experience

In response to questions about prior experience with remote driving and joystick type controls, a range of experience was noted, particularly for gaming applications. Observations did not lead to any conclusions on the impact of gaming and various controller experience on the usability assessments provided by the participants.

B. General Usability

Participants were very forthcoming in their opinions of using the gesture device to control the robot. Most found it easy to use for the limited set of gestures and relatively simple task they were asked to perform, but not all were equally adept at using the device. Many said it was not precise enough, did not allow for small refinements, and they did not see it as sufficiently reliable enough for use outside of a laboratory environment; in other words, “not ready for prime time.”

At the same time, many thought using gesture devices for controlling a robot had some promise. This kind of controller could be quiet and stealthy and have a relatively small size and weight.

There were several usability issues that stood out in being identified as problematic for almost all the participants:

- The requirement to “rest” or neutral position between gestures (as currently required in the implementation of the Myo armband.).
- The inability to make multiple gestures at one time, or at least sequentially, without the need for a “rest” in between gesture (e.g. inability to move forward and turn simultaneously).
- The desire for variable control, for example, acceleration or rate of turning.
- The desire to make “higher level” commands using gestures, rather than just simple, discrete movement commands (e.g. go to the corner of the road).
- Lack of precision of the movements controlled by the gestures.
- Potential impact of fatigue.

C. Physical Considerations

Most thought the gesture control armband was comfortable and the weight of the band was not a problem. However, several individuals thought there should be larger sizes available as they noted discomfort by the end of the experimental session. Another identified concern was of the robustness of the device when exposed to heat and sweat or if used in outdoor and more extreme environments of Army operations. In addition, activities such as bumping into obstacles might cause the armband to slip or dislodge and potentially cease to function. Such operational considerations would need to be addressed if the device was to be used outside of a laboratory environment. Arm fatigue was reported as a potentially significant issue.

D. Command Gesture Mappings

While most found the five gestures themselves easy to learn and use, there were some instances of confusion and lapses in correctly using the intended gesture. Some felt that the mapping of gestures to commands could be improved. For example, the fist was used as the gesture to move forward. While most found this easy to learn and use, some thought a fist perhaps would be better used for a stop gesture. The gesture to stop the robot was to rest the muscles (i.e., place the hand in a neutral position); several people expressed the desire for a positive, active motion for stopping rather than a lack of action. While the reasoning for this (having to do with the reading of the EMG signals) makes sense, it was not an intuitive way to implement a stop. In addition, several of the participants had difficulties in using the “rest” as a means to stop.

As mentioned in B, some participants also wanted the ability to combine gestures, for example combine the forward and turn right (or left) gestures so that they did not have to stop between gestures. Also several desired the ability for the gestures to be able to control speed and acceleration.

All of these comments suggest the importance of carefully identifying appropriate gestures that will satisfy a number of potential constraints, such as being reliably readable for the software/hardware device, be operationally meaningful, and be intuitive for users to learn and use.

E. System Feedback

Participants identified the importance of real-time feedback from the system. Many found the written status of what the robot was interpreting as the current command extremely useful. However, this status was visible only on the FP view. Several comments during the course of the experiment also identified the difficulty of knowing if the robot was interpreting the gesture command correctly, or whether the individual was not performing the gesture correctly for what was intended. We saw instances of both. However providing information to the participant about the state of the robot allowed for appropriate adjustment as they could then know if it was their actions or the robot interpretation which was problematic. When providing an opinion on the two different displays used (FP and OH), a number of participants noted the availability of written feedback of the interpreted command as one of the things they liked about the FP display.

The Myo armband includes a feature where the armband vibrates when the “double-tap” gesture that locks or unlocks the device (like a toggle switch) is performed. Therefore, the user should feel the vibration as a cue whenever the device toggles on/off. Several users mentioned that they felt the vibration of the Myo armband, while others apparently never noticed the tactile cue. They used this tactile feedback as a signal that their intended “double-tap” was read correctly or, occasionally, that some movement was (incorrectly) interpreted as an on/off gesture. If they felt the vibration cue, and knew what it meant, they could easily recover from the misinterpretation of the signal. However, if the participants did not feel or interpret the tactile feedback, then they might continue to make gestures and become frustrated with the lack of response, not realizing that the gesture device was locked/off. This is an example of the importance of feedback; the tactile feedback seemed to work well for several users.

F. Impact of Display View Types

The participants found advantages and disadvantages in the use of the two different displays during task trials and were about equally split in which display they preferred. For example, the OH provided a view that made it easier to align with the path, but made it difficult to identify which way the robot was facing. The FP display provided text feedback on which command was being executed which participants found very helpful, however, it was difficult to carry out path alignment. So there was no clearly better display type for this task. Comments suggested that the needs of the task will have a major impact on the display view preferred.

The DV was only used during the training and practice session. Some participants preferred the DV; they liked seeing, in person, the robot move in response to their gesture commands. The DV also provides auditory feedback during robot movement, which the displays in the remote location did not. The DV would be the view that robot users in close proximity to their robot would have. However, realistically, if the robot is remote from the user, then any view would be viewed through some computer-based display.

IV. Conclusion

This experiment has identified important aspects of usability for gesture control of a small robotic platform using an instrumented EMG device. Both positive aspects and limitations of the gesture device and display views were identified. Future work in this space will be targeted specifically at reducing the amount of low level operation of the platform using gestures in favour of a higher level autonomy. For example, we will develop automatic alignment or obstacle avoidance such that fine grained gestures are not required for successful control of the platform.

Acknowledgements

The authors sincerely appreciate the time our volunteers took in finding the time in their busy schedule to come and support our work.

References

[1] Mitra, S. and Acharya, T. (2007) Gesture recognition: a survey. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and reviews*, Volume **37** No. 3 311-324

[2] U.S. Department of the Army (1987) *Visual Signals Field Manual* (FM 21-60)

[3] Hill, S.G. (2017) *Human-Robot Bi-Directional Communications Literature Overview.* , DST-Group-TR-3416, Fisherman's Bend, VIC, AU, Defense Science and Technology Group..

[4] Elliott, L., Hill, S. and Barnes, M. (2016) *Gesture-Based Controls for Robots: Overview and Implications for Use by Soldiers.* ARL-TR-7715, Aberdeen Proving Ground, MD, U.S. Army Research Laboratory.

[5] LaViola Jr., J.J. (2013) 3D gestural interaction: the state of the field. *International Scholarly Research Notices Artificial Intelligence* Vol. **2013** Article ID 514641, 2013

[6] Thalmic Labs. 2018. ThalmicLabs. [ONLINE] Available at: https://www.thalmic.com. [Accessed 10 April 2018].

[7] Mulling, T. and Sathiyanarayanan, M. (2015, July) Characteristics of hand gesture navigation: a case study using a wearable device (MYO). In *Proceedings of the 2015 British HCI Conference* 283-284, ACM

[8] Barber, D. J., Howard, T. M., and Walter, M. R. (2016, May) A multimodal interface for real-time soldier-robot teaming. In *SPIE Defense+ Security, Unmanned Systems Technology XVIII* 98370M-98370M, International Society for Optics and Photonics

[9] Cacace, J., Caccavale, R., Finzi, A., and Lippiello, V. (2016, October) Attentional multimodal interface for multidrone search in the Alps. In *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, 001178-001183, IEEE

[10] Cacace, J., Finzi, A., and Lippiello, V. (2016) Multimodal interaction with multiple co-located drones in search and rescue missions. *arXiv preprint arXiv:1605.07316*.

[11] Sathiyanarayanan, M. and Mulling, T. (2015) Map navigation using hand gesture recognition: a case study using MYO connector on Apple maps. *Procedia Computer Science* **58** 50-57

[12] Sathiyanarayanan, M., Azharuddin, S., Kumar, S. and Khan, G. (2014) Gesture controlled robot for military purpose. *International Journal for Technological Research in Engineering 1*(11) 1300-1303

[13] Sathiyanarayanan, M., Mulling, T. and Nazir, B. (2015) Controlling a robot using a wearable device. *International Journal of Engineering Development and Research* **3**(3)

[14] Brooks, J. (1986) SUS: A quick and dirty usability scale. [ONLINE] Available at https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

[15] ISO 9241-11 (1998) Ergonomic requirements for office work with visual display terminals (VDTs) Part 11: Guidance on Usability. ISO.

[16] White, G. and Hill, SG. (in press). Usability of a Human Gesture Device for Robot Control. DST-Group-XX-XXXX.