A MULTIPURPOSE ROBOTIC GLOVE DESIGNED TO TEACH SIGN LANGUAGE THROUGH GUIDED MANUAL MOTIONS

SohamJoshi
Raaghav Malik

Columbus Academy

Joshis22@columbusacademy.org; Malikr22@columbusacademy.org

Abstract: Projections show the numbers of deaf and deafblind individuals are quickly increasing, as both the number of deaf and blind individuals are expected to double within the next 50 years. Additionally, there is limited support for deafblind individuals, as they have little sensory access to the world. Most existing solutions are inadequate because they are either not portable, readily accessible or are very expensive. The goal of our project is to create a low cost, accessible solution for those with these disabilities: a portable device that utilizes the sense of touch to teach sign language, instead of conventional sign language learning methods that rely on sight. By doing this, a more immersive and personalized experience is created. The implementation of this solution is two-pronged: first, the physical portion of the solution is equipped with servo motors that control the pulling and release of a cord threaded through rings on the fingers that moves the user's hands into various sign language positions, which is closely modeled after human hand anatomy. Secondly, feasibility is determined for an AI algorithm to take in data from an external camera to efficiently add new signs to the glove, and also to track the user's sign language patterns and rate the user's accuracy with various signs. Through these mechanisms the user is engaged in learning new signs and expanding their sign language vocabulary, all without their sense of sight. We were successful in creating a working prototype analyzed through hand-tracking mechanisms, and achieving a 91-92% accuracy for the American Sign Language alphabet.

1. INTRODUCTION

1.1. Deafblindness

Deafblindness is a general term for anyone who suffers from some degree of hearing and visual loss. A person who suffers from deafblindness may be born blind and lose their hearing as they get older, or be born deaf and lose their vision as they get older. This combination of symptoms is disastrous, as it can vastly limit the abilities of people with this disability (Jaiswal et al., 2018).

In the United States, it is estimated that 70,000-100,000 people are deafblind (NARUC), and over a million exclusively blind and over a million exclusively deaf. Both of these numbers are expected to double in the next 40 years. Globally, approximately 97.6 million people experience some form of visual or hearing impairment, and this too is getting much worse. By 2050, estimates show that one in every ten people will have disabling hearing loss, while blindness rates are expected to triple (Bourne et al., 2017; Goman et al., 2017; World Health Organization, 2019). These statistics highlight the widespread nature of this problem. Even more astonishing facts lie underneath this data. The World Health Organization has estimated that deafness costs the world more than $21 billion annually in assistance and lost productivity (World Health Organization, 2019). In this way, helping individuals with deafness helps the economy of society as a whole. In addition, approximately 100,000 U.S. school children suffer from a serious visual problem, even though vision is one of the most important tools needed to learn (United States Census Bureau, 2007). As of 2007, Over 90% of the world's visually impaired live in developing countries, where they lack the necessary resources to function effectively (World Health Organization, 2007). Therefore, there is an increasing need for a cost-effective solution to this issue. In addition, global awareness of this condition is almost non-existent, especially within developing countries, where this condition is especially dangerous. Primary ear and hearing care has not yet been implemented in developing countries, and the support for secondary and tertiary care is minimal (WFDB, 2018). Also, there are not enough ear health and vision specialists in the world, so there is a need either to train more or find an alternative solution that does not depend on these specialists (Moss and Blaha). Furthermore, the effects of this condition are catastrophic. Not only does deafblindness limit speech and learning at an early age, it leads to slow progress in school and problems in securing a job. It leads to separation from the rest of society and extreme differences in the social and economic aspects (Miles, 2008). Additionally, research has shown that a multisensory, phonics-based approach at an early age is best for learning new concepts (van Staden, 2013). Hence, to create the best learning experience for the deafblind user, there is a need to utilize the various remaining senses, including touch, and whatever limited form of vision or hearing that remains.

Today, current communication technologies make maximum use of the remaining senses, especially the sense of touch (Mason, 2014). This allows the individual to communicate effectively with those around them by both receiving and giving information. As there are many different forms of deafblindness, the possible tools that may help someone who struggles from deafblindness vary greatly, with differing levels of technicality. Some common tools are described below. Magnification Devices: Depending on the level of hearing or vision loss, certain magnification devices, such as hearing aids or cochlear implants, can be used to recover some hearing. However, these depend on the type of hearing loss the user has and are not available in many parts of the world. Additionally, the cost of such technologies far exceed the amount of financial support given to deafblind individuals, and as a result, many cannot afford them (WFDB, 2018).

2. DESIGN OVERVIEW

The design of the device is analogous to that of the human hand, with a flexible frame to provide the underlying structure, similar to human bones, while keeping it flexible enough to comfortably fit around hands of various sizes. The bending and extension of the fingers are facilitated by the pulling and pushing of cords, representing the flexor tendons in the human hand, which also cause the bending and extension of each digit. This cord retracts and withdraws as a result of the rotation of the servo motors located in the upper forearm of the device, representing the flexor muscles in the human hand, which are attached to the flexor tendons in a similar way. The dorsal and palmar interossei muscles of the human hand, which control the abduction and adduction of the fingers, respectively, are represented by micro servos providing a 30 degree rotation to the metacarpophalangeal joints, which connects the digit to the main surface of the hand. One dissimilarity between the device and the human hand is the location of the flexor pollicis brevis and flexor digiti minimi brevis muscles, which are responsible for the primary bending of the thumb to the palm of the hand, and the primary bending of the little finger to the palm of the hand. Although these muscles are located in the thenar and hypothenar eminences of the human hand, respectively, which are the bulges of muscle at the base of the thumb and little finger, the servo causing the corresponding movements on the device are located in the upper forearm, next to the other servos representing the flexor muscles for the other fingers. This difference in design, however, leads to little difference in the overall movements of the device. Two other muscles in the thenar eminence, the abductor pollicis brevis and adductor pollicis, control the abduction of the thumb away from and towards the midline of the hand, respectively. To simulate these muscle movements, the device contains a larger servo located at the back of the hand, which is connected to a wheel that causes the corresponding rotation. The degree of rotation for this servo is 90 degrees, which is greater than that for the interossei muscles of the other fingers. To simulate the opposition of the thumb, the same cord for flexion will be used, in conjunction with abduction, allowing the thumb to rotate in a circle.

The proposed operation of the device is to start off by going through the alphabet as they are simpler, one-handed signs and easier to learn for the student. They will also provide a base for the later more complex signs, as the glove will spell out each new word before teaching the new gesture so the user can understand the glove exclusively through touch. The glove will progressively go through more complex signs, so by the end users can frame together sentences and effectively communicate with others. In addition to this teaching mode, there is also an assessment mode where users can practice the gestures they have learned which will be tracked either through a camera on a smartphone or laptop, while providing real-time feedback for improvement. The user will then have options to practice these signs, all controlled through the pro-tactile user-interface, using an external Braille keyboard or through Braille buttons on the glove.

3. APPROACH

3.1. Plan and Prototyping Stages

To reduce costs, the original hypothesis was that if we only use a single servo motor for each finger, instead of a servo motor for every joint in the finger, we can create a device designed to be much more affordable with no drawback in functionality, as many signs in American Sign Language can be formed by solely utilizing the flexor tendons and flexor muscles. To test this hypothesis, we proceeded to create a feasibility prototype which only allows for the flexion of each digit as a simpler model of the human hand. The structure of this prototype was created using 3D-printed parts created from polylactic acid (PLA) filament. This prototype is modeled after the inMoov project with modifications for ASL. To create the finger flexion and extension movements, two cords are threaded through each finger: one on the top, and one on the bottom. The flexion cord, located on the front of the finger, makes the finger flex when pulled, while the extension cord, located on the back of each finger, makes the finger extend when pulled. These cords are attached to a wheel driven by a servo motor that pulls the flexion cord while releasing the extension cord when turned in one direction, while releasing the flexion cord and pulling the extension cord when turned in the other direction. A simple circuit was used to train various signs on the Stage 2 Prototype. Using a potentiometer dial and a pushbutton, we are able to toggle between the various fingers and change the flexion of each one, while a second push button allows us to save the sign to the memory of the microcontroller. However, due to the rigid structure of this prototype, many flaws were noticed, and the glove was not able to perform many signs, including basic alphabet letters, so we proceeded to develop another prototype: a device that fully encompasses a user's hands with flexion in all fingers and abduction, adduction, and opposition in the thumb. In addition, the Stage 3 prototype is made out of a flexible rubber that fits precisely around each of the user's fingers, allowing for a more comfortable fit. It primarily consists of 3D printed rings with loops where a cord is threaded. This cord is attached to servo motors located on the forearm of the glove, which pulls and retracts, making the finger flex and extend, similar to the design of the feasibility prototype.

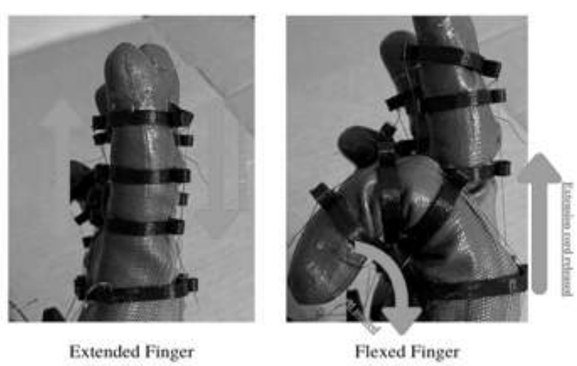

Figure 1. Stage 3 Prototype finger flexion and extension

One major difference is that the glove is wearable, unlike the Stage 2 Prototype, and rings are used to thread the flexion and extension cords instead of directly through the finger. Additionally, a larger servo fits on the back of the glove, creating the thumb abduction and adduction movements. To make the thumb both flex and abduct, a novel mechanism was used in which a flexible thumb guide was fabricated that moves the thumb away from and towards the rest of the fingers, while being flexible enough to allow the thumb to flex and extend like the other digits. By abducting the thumb away from the fingers while flexing it, we can also create the opposition motion of the thumb in this way.

Figure 2. Stage 3 Prototype thumb abduction, adduction, and flexion

4. HAND TRACKING FUNCTIONALITIES

Hand keypoint extraction was used for multiple different aspects of this project and implemented through the Google Mediapipe framework. This framework allows us to create and modify a machine learning pipeline that locates 21 specific keypoints on our hands as various gestures are performed in front of the camera by overlaying a wire model of the hand on top of the input image.

This algorithm was then modified to output the relative coordinates of the various keypoints, normalized in a 1x1 square centered at the palm, with (1, 1) as the bottom left coordinate and (0, 0) as the top right coordinate. Additionally, the z-coordinate was estimated based on lighting and hand anatomy through Mediapipe.

Figure 3. Normalization of coordinates for keypoint estimation and coordinate output

Once the locations of the keypoints were found, more information can be extrapolated. Firstly, the keypoint coordinates were used to calculate the various flexion and abduction angles between the fingers. To do this, the angles between two consecutive keypoints were found and summed for the whole finger to get the flexion angle for that finger. For example, to find the angle of rotation of the PIP joint on the index finger, the vector pointing from the PIP joint to the DIP joint and the vector pointing from the PIP joint to the MCP was found. Using the dot product formula where $a$and $b$are the two vectors this angle can be calculated. Repeating this process for every joint in the finger, the total flexion angle of the finger can be calculated. A similar process is repeated to find the abduction angle by finding the angle between the vector pointing to the MCP of the thumb from the base of the hand and the vector pointing to the MCP of the index finger from the base of the hand.

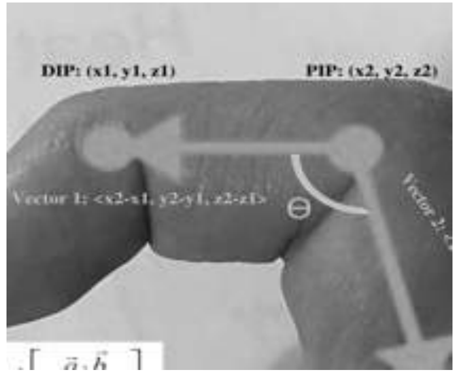

Figure 4. Finger flexion angle calculation

After the various flexion and abduction angles are found, they can be used to teach the glove new signs. For example, we can form a sign with our hands in front of the camera so that keypoints and angles are tracked, and saved to the memory of the Arduino microcontroller. The device can rotate its flexion and abduction servos to the corresponding angles to form the sign. This feature allows us to add new signs to the device with relative ease, instead of hard-coding each of the servos, and can easily be expanded as we continue to add more servos to the device for future prototypes. Keypoint extraction and analysis also helps with testing of the prototype, as it allows us to determine how inconsistent the human hand forming various signs is to the device forming various signs. We propose a third use of the hand tracking feature which we seek to include in future iterations: personalized feedback to the user. In the assessment mode of the device, the device can prompt the user (through the Braille display, the LCD display, and/or the speaker) to form a specific sign. Similar to how the device was tested, the error between when the user is forming a sign and when an expert is forming a sign, found in an online ASL dictionary, is determined through finding the distance between the respective keypoints, as well as the percent error between the coordinates. The average of this distance is then computed. The signs that have a high average distance and high percent error are the signs that the user needs the most assistance on, so the algorithm ranks the signs based on this characteristic. This allows the user to practice the signs that he/she struggles with the most. Our Stage 3 prototype cost us around $60 to develop. Including further manufacturing and development costs, a proposed selling price for the minimally viable product is $120. This is vastly cheaper than current BTSs, the cheapest of which sells for $1,000 today, validating one of our objectives. This objective was largely accomplished through minimizing the amount of servo motors on the device by using a single servo motor per finger, closely related to how the human hand only has one flexor muscle per finger, instead of a motor per joint in the hand.

5. RESULTS AND DISCUSSION

5.1. Stage 3 Prototype Testing

To test the Stage 3 Prototype, the hand tracking algorithm was utilized to extract various keypoints and angles on the glove and compare them to the theoretical sign keypoints and angles. Multiple tests were performed to compare the set of keypoints formed by the glove and the set of keypoints formed by a human: a percent error analysis and a distance error analysis. Firstly, the percent error was calculated between each coordinate formed by the glove and each coordinate formed by the human hand, and these were averaged between the X, Y, and Z coordinates to get the total error for a specific keypoint. Lastly, the percent error for each keypoint was averaged to get the average percent error for a single gesture, and subtracted from 100% to get the average accuracy. This process was repeated for every gesture in the ASL alphabet, excluding the letter “R,” which is explained below.

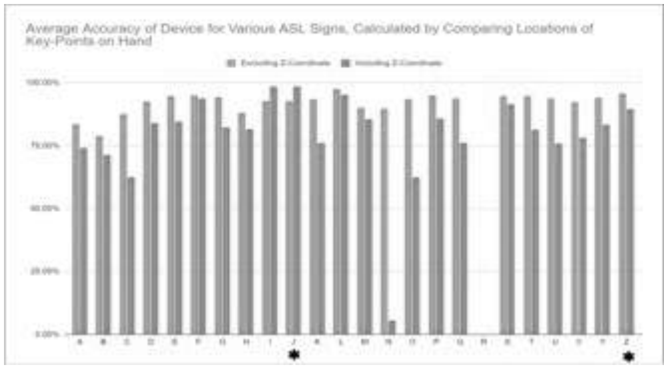

Figure 5. Percent error analysis for various ASL alphabets

This chart shows that when including the Z-coordinates, the accuracy of the gestures are significantly lower for most signs. The Z-coordinates were estimated based on lighting conditions, and was largely inconsistent, as the keypoint tracking does not take in an input of a depth sensor. This allows us to significantly reduce the costs for the user, as personalized feedback can be given without the reliance on the depth sensor, but also limits the extent of testing, as the Z-coordinate gives very little information. Still, only using the X-coordinates and the Y-coordinates, we can still evaluate the accuracy of the device, as it is difficult to change the Z-coordinate of a sign without also changing the X-coordinate and the Y-coordinate. Using the percent analysis approach, there is a greater than 75% accuracy for all signs, and the least accurate sign is the letter “B” with a 78.93% accuracy. The average accuracy for all letters (excluding the inconsistent Z-coordinate) is 91.86%. The letter “R” could not be shown in the results because the hand tracking algorithm does not recognize the hand when the index finger and the middle finger are crossed, which is used in this sign. Therefore, no result could be computed and this had to be omitted from the analysis. Additionally, the letters “J” and “Z” (marked with an asterisk) both involve motion, as the sign for “J” involves bringing the pinky down in the shape of a “J” and the sign for “Z” involves tracing a “Z” with the index finger in the air. To make the testing procedure the same for each sign, the motion aspect of these signs was omitted, and instead a static image of the gesture was used instead. Additionally, the distance between keypoints on the glove and keypoints on the human hand was also found, so that each keypoint on the glove was compared to the same keypoint on the human hand. For example, the coordinates of the tip of the index finger on the glove was found, and the coordinates of the tip of the index finger on a human hand performing the same sign was found, and the distance between these coordinates was calculated using the Pythagorean Theorem to determine the error for this coordinate. This process was repeated for each of the 21 keypoints on the hand and device.

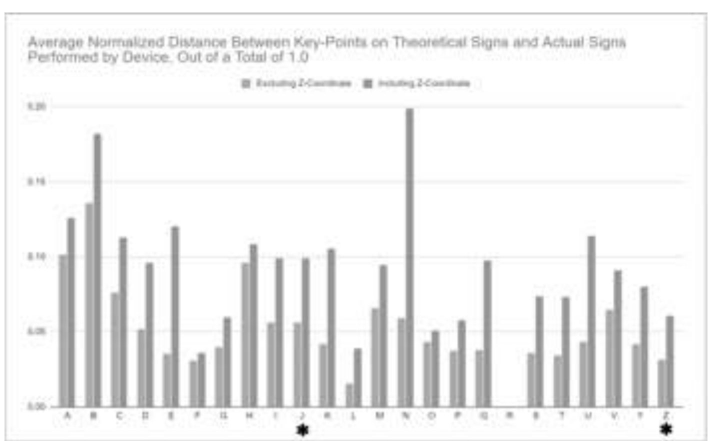

Figure 6. Average error in distance for all keypoints on various ASL alphabet signs

Again, the error with the Z-coordinate is much higher than when the Z-coordinate is not included, due to inconsistencies in the approximation of this coordinate. In future testing procedures, the depth sensor will need to be utilized in addition to current methods. Additionally, the sign for “R” could not be analyzed again and the signs for “J” and “Z” were limited to a static image. By dividing each of the distance errors by 1.0, which is the maximum distance error possible for each keypoint, a percentage is calculated that also shows the accuracy of this keypoint. For this test, the average accuracy was computed to be 94.64% for the ASL alphabet signs excluding the Z-coordinate, and 90.53% including the Z-coordinate. Because the hand tracking algorithm could not identify the glove due to differences in skin tone, a human hand imitating the glove as closely as possible was used, through anonymous images. This may lead to some human error, but the hand was not the experimenter’s own hand, to reduce this human error. For future studies, the training set of the hand tracking algorithm will be modified to include the device. Current devices designed to translate the ASL alphabet to speech do so at around a 90% accuracy, which our device surpasses (Medhi and Khan, 2012; Waldron and Kim, 1995). However, the objective for our device is not to translate ASL alphabet to speech but to teach sign language instead, which has not currently been implemented. We are expanding on this current research on analyzing sign language in a more quantitative way for use in robotics and machine learning.

6. CONCLUSION

Overall, we created a device that would be able to teach people sign language at a significantly lower cost, at a high level of accuracy. We were able to create a glove-like unit that replicated the movement of the human hand and managed to complete all ASL alphabet signs, except for the letter “R”. Through our testing we confirmed that our solution performs at a 92% accuracy. We believe that our product can help those who are deafblind easily communicate with the people around them. We have also started preliminary research on the Stage 4 Prototype, which will include additional movements for wrist flexion and extension and radial and ulnar deviations, and elbow flexion and extension, allowing us to have a greater accuracy for the full range of sign language gestures. Possible movements for the final design will include: flexion (bending) of each finger, including the thumb, abduction and adduction of each finger up to 30 degrees, abduction and adduction of the thumb up to 90 degrees, opposition of the thumb and the little finger, flexion and extension of the wrist, radial and ulnar deviation of the wrist, flexion and extension of the bicep, pronation and supination of the forearm, the vertical and horizontal flexion and extension of the shoulder, and internal and external rotation of the shoulder.

For more complex motions, such as shoulder rotation and flexion, vibration motors located on various parts of the device will be used that guide the user into the right position, without forcing their arm to move.

7. REFERENCES

[1] Ahmed, M. A., Zaidan, B. B., Zaidan, A. A., Salih, M. M., & Lakulu, M. (2018). A Review on Systems-Based Sensory Gloves for Sign Language Recognition State of the Art between 2007 and 2017. Sensors (Basel, Switzerland), 18(7), 2208. doi:10.3390/s18072208

[2] Assistive Technology. (2018, May 18). Retrieved June 22, 2021, from World Health Organization website:

[3] https://www.who.int/news-room/fact-sheets/detail/assistive-technology.

[4] At risk of exclusion from CRPD and SDGs implementation: Inequality and persons with deafblindness. (2018, September). Retrieved January 1, 2020, from World Federation of the Deafblind website: https://www.safo.no/wp-content/uploads/2018/10/2018-Global-rapport-om-d%C3%B8vblindhet.pdf

[5] Bourne, R. R. A., Flaxman, S. R., Braithwaite, T., Cicinelli, M. V., Das, A., & Jonas, J. B. (2017). Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health, 5(9). https://doi.org/10.1016/S2214-109X(17)30293-0

[6] Deafness and hearing loss. (2019, March 20). Retrieved January 1, 2020, from World Health Organization website: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[7] Gilbert, C., et al. (Eds.). (2007). Vision 2020: Global initiative for the elimination of avoidable blindness. Retrieved from World Health Organization website: https://www.who.int/blindness/Vision2020_report.pdf.

[8] Goman AM, Reed NS, Lin FR. Addressing Estimated Hearing Loss in Adults in 2060. JAMA Otolaryngol Head Neck Surg. 2017;143(7):733–734. https://doi.org/10.1001/jamaoto.2016.4642

[9] Jaiswal A, Aldersey H, Wittich W, Mirza M, Finlayson M (2018) Participation experiences of people with deafblindness or dual sensory loss: A scoping review of global deafblind literature. PLoS ONE 13(9): e0203772. https://doi.org/10.1371/ journal.pone.0203772

[10] Mason, A. (2014, October). Deaf-blind communication technology. Retrieved January 1, 2020, from National Federation of the Blind website: https://www.nfb.org/sites/www.nfb.org/files/images/nfb/publications/bm/bm14/bm1409/bm140906.htm

[11] Moss, K., & Blaha, R. (n.d.). The unique educational and services needs of children with deaf-blindness. SEE/HEAR. Retrieved from http://www.tsbvi.edu/seehear/archive/unique.html

[12] Taylor, C. L., & Schwartz, R. J. (1955). The anatomy and mechanics of the human hand. Artificial Limbs, 2(2), 22-35.

[13] United States Census Bureau. (2007, January 12). State and country quick facts (USA). Retrieved February 7, 2007 at http://quickfacts.census.gov/qfd/states/00000.html.

[14] van Staden, A. (2013). An evaluation of an intervention using sign language and multi-sensory coding to support word learning and reading comprehension of deaf signing children. Child Language Teaching and Therapy, 29(3), 305–318. https://doi.org/10.1177/0265659013479961

[15] Usher Syndrome. (2017, March 16). Retrieved January 1, 2020, from National Institute on Deafness and Other Communication Disorders website: https://www.nidcd.nih.gov/health/usher-syndrome