PROJECT “AWAAZ”-

Developing an extremely low cost, non-movement restrictive plugin for hand movement to speech conversion

Debargha Ganguly.

Abstract— Project “AWAAZ” essentially uses a set of strategically (based on crypto graphical analysis of words) placed momentary tactile switches on a wearable system that sends characters or phrases to an Arduino board that processes the respective hand movement to form letters and words out of it. The Arduino board is programmed in such a way that it is capable of emulating an HID keyboard of the device that is attached. The letters are then sent to my app that converts the text to speech using standard android based text to speech conversion system that doesn’t require any internet connection through a third party word prediction software (reduces time by 32%). The system is designed to draw charge directly from the phone battery hence it doesn’t need any external power source, has a 100% accuracy rate and doesn’t have any limitations regarding connectivity. It also allows easy hand mobility with low finger stress, easily replaceable parts, and is developed inside a budget of 450 Rupees.

1 INTRODUCTION

1. INTRODUCTION-

1.4 percent of the population suffers from sorts of speech disorders and they experience a lower life-expectancy in part due to lack of expression. Improvement of lives of millions of people can be done by giving them a means of communicating.. Most literate people now have smartphones which are essentially processing powerhouses. A huge percentage of people from the middle and lower classes of developing countries are unable to buy standard devices for speech production because they retail at over thousands of dollars. Most low cost prototypes are created in such a way that they either have high independent costs (For just the system) or they are restrictive for motion and aren’t accurate enough for day to day usage. The lowest priced prototype right now costs INR 12,000+ and the lowest priced commercial alternative costs INR 30,000+.

2. METHODS AND IMPROVEMENTS

The feasibility of using gesture mapping was checked first. Sign languages use a lot of different parameters while communicating, which would create a lot of parameters and sensor data to be processed. The amount of data to be processed was becoming so much that low cost microprocessors wouldn’t be able to process that much seamlessly, and was significantly increasing the cost of the device. (30+ locations were to be mapped at once).

I wanted to use flex sensors initially, because they are easier to integrate into wearable electronics, however, since accurate character reproduction is required, plotting and tracking movements on the X-Y-Z axes require powerful processors that further increase the cost of the project. A system with flex sensors and powerful enough processors would cost over a 5000 Rupees, well above what the target demographic would have access to, or would be willing to spend. Flex sensors also create a lot of sensor noise, that would have reduced the accuracy rate of the device. Tactile momentary switches are much better because they are much cheaper, easily replaceable and extremely easy to work with producing predictable results. (Debounce time of 20ms)

While coding, the usage of the serial functions would seem like the easier option however it would result in the display of the results on a separate terminal window, which is highly inconvenient for any text to speech engine to read. The keyboard library functions were used instead because it allows HID device functionality, solving a lot of problems at once.

Arduinos can also be attached with certain shields that make them capable of text to speech capabilities, however it isn’t feasible because it requires preprogramed text strings to be sent, and it would be impractical too as the shield cost almost the double the price of most lower end smartphones.

The device is able to log 60- 120 keystrokes per minute. Although this is fast for a AAC, it still isn’t as fast as normal speech. If required, the user can use third party text prediction services like SwiftKey to engage auto correction, hence vastly increasing the speed, including being able to type words by just a single keystroke. The code for the Arduino is written in such a way that the software detects it’s a physical keyboard, and hence, is compatible with most auto correct engines. The system was also checked with a third party keyboard software that detected it as a hardware keyboard (SwiftKey) and it was found that the system used 67.82% of the total time without it, which is a substantial saving in terms of time if the person is using it for daily usage.

Debargha Ganguly – 801, B, Capri, Hiranandani Estate, Thane West, Mumbai, Maharashtra, India

A bluetooth or Wi-Fi based data transfer system wasn't used because it requires an additional power source whereas using an USB OTG cable network allows the charge to be drawn directly from the phone battery. This also removes problems like restrictions caused due to range of the connections.

3. DISCUSSION AND INNOVATION-

AAC devices have been produced before, however they are very expensive. Project AWAAZ was developed with a budget of 450INR which is about 1% of the cost of some of the devices that cost 1000s of dollars.

Inaccurate/ Inconsistent readings- Accuracy and consistency are extremely important for any machine that is used in everyday life. Gesture mapping systems don't provide predictable or repeatable results due to sensor noise. Attempts were made to reduce moving parts and sensors to the bare minimum to produce high levels of accuracy. In the testing phase, project awaaz was able to produce 100% consistent results over several cycles of testing.

Gesture analysis is processor time intensive because of the number of complex algorithms running on the processor, therefore results usually aren't instant. Due to rapid and exponential simplification of the code, the algorithm for the detection of letters in project awaaz is incredibly simple, allowing it to run on lower end microprocessors efficiently.

The system is wearable as a glove, and the phone can also be mounted on the hand, therefore, the device is non-restrictive for movements which is an essential criterion to be met for people who really want to use it in their day to day lives. Most systems require a start-up, boot up, or initialisation time This isn't needed by "AWAAZ" because it can just be plugged into a power source and used. Since the code doesn't need any operating system, it runs as soon as the device is connected to a power source. This also means that it doesn't need calibration every time it is used.

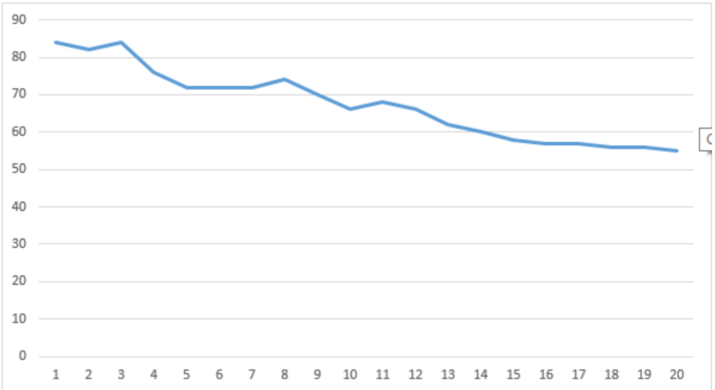
According to finger flexibility and strain tests conducted with the machine, the switches have been placed on the four fingers except the thumb, so that it can have the maximum reach and the minimum possible strain, giving the user the ability to use it throughout the day without causing inconvenience. The letters of the English language were also analysed cryptographically to figure out the letters that are used the most in the English alphabet, and these letters were made the most easily accessible for user friendliness. Apart from this, two phrases "Hello, how are you?" and "I need help, please help!" are also given their corresponding button, so that the user can ask others for assistance in the eventuality of an emergency. The 12 buttons can also be configured in different ways (22/30/36/40/42 unique combinations) as per the usage to add more phrases, or letters in case any other language support is needed.

4. RESULTS-

The device is able to successfully convert motor movements to speech, but however, that's just the technical aspect. Real life, day to day usage requires the brain to accept the machine as a body part that it can provide output through, and only then will it be able to speak fluently through it. This test is to examine whether cross modal linking, brain plasticity and neuron training can be used to change the ordinary pathway for the electrical signals through the neurons so that the brain adapts to using the device as ordinary speech. With conditional treatment, the brain is bound to adapt to using the gloves as an alternative, but it should be able to use it day to day and fast enough. So essentially the test is to determine whether the nervous system can accept the machine as a viable text to speech engine or not.

The following tests were carried out on the device

1.) Learning curve /Typing speed improvement with time/ Neuron training –
This test was conducted by using the following phrase "The quick brown fox jumps over the lazy dog" which is an English language pangram and measuring the time taken to type it out over time to test the adaptability of the nervous system by neuroplasticity to using the system as a viable replacement.



The figure shows the time taken to type the pangram using the system ,and its progression with time. (every 10 Minutes)

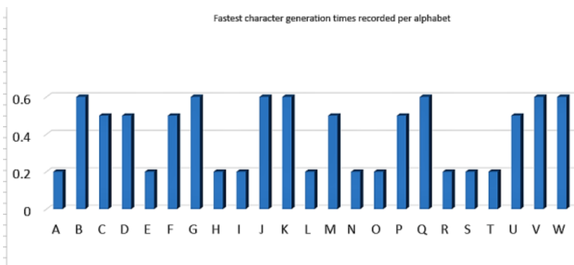
Data for the time required to type in the pangram through the AWAAZ suggesting neuron training helping the user get accustomed to using the system. The values for this will change subjectively from one person to the other according to age and other factors, however this suggests that the system can be used to type faster with time. The data was taken in intervals of 10 minutes. This is by no way the fastest the system can perform, with adequate training, the fastest time possible for this same task is around 30-35 seconds based on autocorrect performance.

2.) Comfort – Strain reduction –
According to angle strain and flexibility tests conducted with different orientation of the switches (all 5 fingers, base of the fingers only and using the tips of the fingers to press the respective buttons and 4 fingers and usage of the thumb to press the buttons) , the switches were mounted on the 4 fingers except the thumb because the thumb has the maximum flexibility and reach.

3.) Prediction and sentence completion- The plugin was used to type in 50 random words with and without the third party prediction engine. It was found that when it was used with the auto-correct engine, it was able to type the same words within 67.82% of the time required without the auto-correct and prediction enabled.
(This doesn't consider the time saving caused by changing the positions of the keys.)

4.) Replicability of the parts – The most expensive part of the system is the Arduino board that can be protected from external elemental damage so that it doesn't get damaged. Switches, wires, and gloves are easily replaceable because they can be found at any hardware store, and relatively cheap to replace when compared to other sensors like flex sensors or cameras.

5.) The fastest character generation times were recorded per alphabet and are displayed in seconds in the graph.



6.) Endurance of the system against external stress – The Arduino board was protected by wrapping it in layers of polythene, to protect it from accidental water damage. The entire system including the phone can be mounted on one hand too

7.) Battery life- The system was checked for battery life by connecting it to a 2200mAh battery (fully charged). Being continuously connected it drained the battery out completely in 20 hours when in stand-by mode. The battery consumption when the phone Is being used with the plugin depends on various other factors such as screen brightness, background processes and speaker volume. Therefore, the plugin draws 2200mAh charge in 20 hours.

5. *DISCUSSION AND CONCLUSION*-

The device was able to convert the hand movements to speech efficiently with a 100% accuracy rate therefore making it a viable alternative communication device for the speech disabled person. However the time required for the person to get used to the device is still quite high (currently more than a day) to use the system to it's full efficiency. The system is compatible to any kind of typing speed ,and the fastest possible typing speed can be changed by changing the delay at the end of the code. The future goals for the project is to shut the processor of the Arduino down to save power when not required, increase user friendliness, shorten the learning curve further so that the person can start using it more efficiently and increase the compatibility of the gloves to support all devices like Symbian phones (which are mostly found among lower income groups). It also needs to go through more rigorous tests for performance and durability before commercialisation. Anyways, this device shouldn't be compared with AAC devices like TALK because this is made to be extremely user friendly and simple for mass production and maximum reach impact. This module can provide it's usage only to any person who has muscle motor activity control (Including people with developmenta disabilities).

6. Algorithm-

Step 1.) Identify the key that has been pressed.
Step 2.) Identify if the control or alt key has been pressed, accordingly selecting the correct alphabet or letter to be sent
Step 3.) The data goes to the phone through the USB OTG Cable and the phone thinks that an external keyboard is sending that data
Step 4.) The data passes through the Swiftkey keyboard autocorrect engine for physical keyboards, if a space is pressed the closest prediction will be written onto the app that converts the text to speech.
Step 5.) A simple offline android text to speech engine converts the text to speech and then speaks it out through the phone speaker.

7. ACKNOWLEDGEMENTS

I would like to thank my research mentor – Dr Krishnendu Chatterjee for his constant help throughout this research.

8. REFERENCES

[Debargha Ganguly. (2016); DEVELOPING AN ECONOMIC SYSTEM THAT CAN GIVE A BLIND PERSON BASIC SPATIAL AWARENESS AND OBJECT IDENTIFICATION. Int. J. of Adv. Res. 4 (11). 2003-2008] (ISSN 2320-5407).