Silent Speech Recognition Interface for differently Abled

[1]Gopika Mohan, [1]Muhammad Ali N, [1]Nimmy Saji, [1]Sreeja J, [2]Rejeenth VR

[1]UG Scholar, Department of Computer Science and Engineering,

[2]Asst. Prof, Department of Computer Science and Engineering,

UKF College Of Engineering And Technology, Parippally, Kollam, Kerala, India

Abstract : Speech is the most preferred means of communication. Silent speech thus can be analyzed using lip-reading or using facial muscle movements. These signals are processed and matched with the already trained phoneme and then converted into speech. Thus, a silent speech interface has the potential to enable a differently-abled person to speak and interact with people to ease their lives.This project monitors the EMG Outputs so that a confidential and disturbance-less communication happens.Recognition of facial muscle movement and the conversion of these muscle movement into electrical signal relies on EMG sensors.

***IndexTerms* – Silent Speech Recognition, EMG, Faicial Muscle Movements.**

I. Introduction

Speech is the foremost convenient and natural style of human communication. Unfortunately, for everyone, normal language is not possible. As an example, persons that suffer traumatic injuries, laryngeal cancer, or neurodegenerative disorders may lose the ability to speak. The prevalence of this type of disability is vital, as evidenced by several studies. For instance, approximately 0.4% of the EU population have a speech impediment, while a survey conducted in 2011 concluded that 0.5% of persons in Europe presented 'difficulties' with communication. The American Speech-Language-Hearing Association (ASHA) reports that almost 40 million US citizens have communication disorders, costing the US approximately $154-186 billion annually. The World Health Organization (WHO), in its World Report on Disability derived from a survey conducted in 70 countries, concluded that 3.6% of the population had severe to extreme difficulty with participation within the community, a condition which has communication impairment as a specific case. Speech and language problems have a deep impact on day to day life for those who suffer,which make them a hurdle on a daily basis.Even most of the services are not trained to communicate with people with speech disabilities which feels so uncomfortable. As a result, people with speech impairments often develop feelings of personal isolation and social withdrawal, which can cause major affective disorder. Furthermore, a variety of those people also develop feelings of loss of identity after losing their voice. Communication impairment can even have important economic consequences if they cause occupational disability. Various methods can be used for the restoration of communication in the absence of clinical processes. One such is assistive technology. The U.S. National Institute on Deafness and Other Communication Disorders (NIDCD) defines this as any device that helps a private with deafness or a voice, speech, or language disorder to talk. For the precise case of communication disorders, devices used to supplement or replace speech are called augmentative and alternative communication (AAC) devices. These devices are so many and they will range from simple paper and pencil to visualization or text-to-speech (TTS) software. From an economic standpoint, the worldwide marketplace for AAC devices is predicted to grow at an annual rate of 8.0% during the subsequent five years, from $225.8 million in 2019 to $307.7 million in 2025. AAC users include individuals with a variety of conditions, whether congenital (e.g., encephalopathy, intellectual disability) or acquired (e.g., laryngectomy, neurodegenerative disease, or traumatic brain injury).By these days, a new AAC approach has emerged which is silent speech interfaces. SSIs are assistive devices to revive language by decoding speech from non-acoustic biosignals generated during speaking. There are many techniques to take the biosignals for the face, like surface electromyography (sEMG), which helps to capture the electrical activity of the facial muscles using a surface electrode. Since SSIs enable speech without relying on the acoustic signal, they supply a fundamentally new means of restoring communication capabilities to persons with speech impairments. But clinical uses, other applications of this include providing more privacy, enabling phone conversations to be held without being overheard by others, and enhancement of normal communication in noisy environments. These applications are only possible because these signals from the body are large to environmental noise and are independent of the acoustic signal.

II. Literature Survey

From a baby's birth, mother and the child will begin to exchange words that lay the basement for their relationship. As their children grow, they will learn and understands more about what others are talking and finds ways to communicate their needs and wants to others. Time before formal languages became popular, childrens communicate with their parents through preverbal methods. Classifying of speech into formal is very difficult and a herculean task. By classification, only 5% to 10% of the total population has a completely normal speaking condition and has a healthy voice; all of the others people suffer from some disorder. Speech disorders are problems that occur when producing the sounds via vocal chord or with the quality of human voice. and For those people to be apart of the society, in which others can communicate is literally a difficult task .Because they are not able to convey their message to other properly since, others don't know their sign languages and in most case people are running behind the time ,so no one will wait for them to see what they are trying to convey

III. Related Work

Data-Glove approach

The data-glove approach uses an assembled electronic glove which has been equipped with certain sensors that are utilized to differentiate the hand gesture.commonly used sign language systems uses the data glove method,as it is very efficient to maintain and use the substructure requires less force. The glove is equipped with ten flex sensors where each finger has to flex sensors. The flex sensors work as variable resistance sensors that alter resistance as given by flexing.These sensors send the data to the microcontroller from the bending point of every joint of fingers. At the combined joint of fingers and palm to fingertips it is covered over the outer layer of the data glove. To recognize the acceleration of hand movement at distinctive bearings a 3-axis accelerometer is used and this accelerometer is placed over the back portion of the glove.The data glove is exceptionally handling good in both fingerspelling and sign motion,which make the data glove more expensive.When the data gloves are less expensive then it is more weak to noise. When the quantity of the sensors are compromised then it will lead to the loss of data.Which will lead to loss of correctness in sign interpretation.

IV. Proposed System

Beforeyoub Here a system on the basis of facial muscle movements, called lip reading is proposed. By using Electromyography, activity of muscle tissues is recorded. The Electrodes are placed on the face and neck as shown. These electrodes monitor the tiny muscular movements that occur when humans speak. These signals are converted into electric pulses. This signal will be subjected to a classification problem to get an accurate letter prediction. Once a person speaks or prepares to speak, small contractions on speech articulatory muscles are generated. These contractions can be measured by electrodes between different areas on the skin surface.Before processing the collected data , it has to be preprocessed.Band-pass filters are used to filter any type of noises and then the effective segment information can be extracted. After preprocessing, the data is ready to get extracted and arranged as per clusters to get identified.[5] The proposed system is divided into two phases.First phase deals with creation of EMG recorder and the second phase deals with the EMG signal input and features.The first phase says that when humans speak, there will be movements on facial muscles and each letter will have a different type of facial muscle movement. These Muscle Movements have to be captured.In Order to capture the movements of the facial muscles, an EMG Sensor can be used .EMG Sensor records the Muscular activities and converts the muscular movements to Electrical signals.

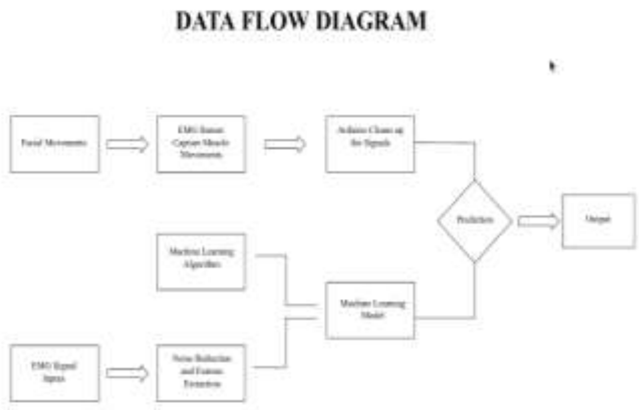

Fig. 1. Data Flow diagram

These Electrical Analog Signals are then converted to Digital Signals. In the second phase The Signal from the EMG Sensor is transmitted to the Ardiuno and then the Arduino will log the data into an SD Card. This data logged into the SD Card is used to train the Machine Learning Model.An arduino equipped EMG recorder incorporated with muscle signal are used to record the electrical signals from the throat and which is then saved to an sd within an sd card module,the saved dataset then undergoes ML training model.Six electrodes were placed to the throat. A sum of 1020 recordings were collected,40 for each of the 26 Alphabets. After the electrodes are placed, the letter is spoken through the device. A button is fixed to control the recording of the EMG signal.This process was repeated a total of 1020 times to obtain 1020 strings of EMG values.The data is taken and imported into the PC as a .CSV file.This CSV file is imported to excel. An SVM ML Model is then trained with the EMG dataset. After the completion of the model training,the ML model obtains an accuracy of 60%.This shows that the aid developed will be able to translate signals from the throat into letters with an accuracy of 60%.An interface build using stremlit was used to visualize the produce letter along the voice produced.

V. Conclusion

We would like to thank everyone who has contributed to the successful completion of this project. We would like to express our gratitude to our project guide Prof. Rejeenth V.R Assistant Professor in the Computer science and Engineering Department who gave valuable suggestions and guidance for our project. We express our deep-felt gratitude to beloved HOD Dr. Ramani, Head of the Department for providing necessary information regarding the project and also her support in completing it. We also thank our project coordinator, Prof. Jithin Jacob, assistant professor who gave expert supervision, encouragement, and constructive criticism amidst his busy schedule throughout the project. We are also grateful to all the authors of books and papers which have been referred to publishing this paper to express.

References

[1]National Institute on Deafness and Other Communication Disorders: Retrieved from https:// www.nidcd.nih.gov/health/statistics/quick-statistics-voice-speech-language (2010).

[2] Juang, B.H.: On the hidden Markov model and dynamic time warping for speech recognition— a unified view. AT&T Bell Laboratories Tech. J. 63(7), 1213–1243 (1984.)

[3] Wand, M., Schultz, T.: Session-Independent EMG-Based Speech Recognition (2011).