TEST Engineering & Management

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

Speak Via Fingers Working Model Final Output

[1]K. Keran Grace, [2]Dass

[1,2]Department of Electronics and Communication, Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai-602105, Tamil Nadu, India
[1]kerangrace767@gmail.com, [2]dass@saveetha.com

Article Info
Volume 83
Page Number: 5303-5308
Publication Issue:
May - June 2020

Article History
Article Received: *19 November 2019*
Revised: *27 January 2020*
Accepted: *24February 2020*
Publication: *16 May 2020*

Abstract

The famous mode of communication for the mute is through sign languages. The main challenge faced by the mute is to convey what they think perfectly to normal people who lack knowledge in sign languages. To overcome this challenge, we have come up with a technological method where the alphanumeric code can be sensed from the inputs given through the fingers and processed to produce the desirable output. These codes are generally represented in numbers from 0 - 9 associated with the respective alphabets. The code can be generated for each alphabet and this can be processed and converted to voice. The output can be obtained in the speaker. This is more like a mute person being able to speak through their fingers. This method takes less time to produce the output. This can help the mentally challenged to express their views as well. This project would enhance the employment opportunities, communicating abilities for the people with multiple disabilities.

Keywords**:** *Text to speech, disability communicator, Translate, SMS, LCD*

1. Introduction

The statistics states that the ratio of people who are mute, blind and lack the capability to hear is 10:1million. That is approximately about 11% of the worldwide population. Communication is the most important aspect to convey our views or thoughts to another person which is very much essential. People with multiple disabilities have scarce languages or modes of communication, hence they find it difficult to communicate between themselves and with the others. Sign languages are mostly evolved from North American regions and are largely used by the mute and deaf. It is a mode generally expressed by gestures, actions and lip movement which is quite understandable for the opposite person. However, the types of sign languages that have been evolved is still tentative and it varies with region to region. The sign language is being widely taught in deaf and mute schools. And there has been no easier alternative for sign languages yet. There are basically 3 sign languages which are widely accepted. They are American Sign Language (ASL), Pidgin Signed English (PSE), Signing Exact English (SEE).

To overcome this methodology of sign languages which cannot be understood by the other humans, the solution is formulated. This proposal is a technological aid which helps the blind, deaf and mute to communicate in an easier way using the latest technology. Inputs are given by tapping fingers which are processed to produce the output. The format of keypad used here is 'Alphanumeric keypad'. This is somewhat like keypads that can be found in old 'type' mobile phones. Since the alphanumeric keys are quite familiar and can be easily taught to the user, it has been considered. The alphanumeric keypad has been used over years in many applications like laptops, phones, ATM pins etc. The alphanumeric key dictionary has been defined and it is interfaced with the GPIO pins of the Raspberry Pi ZERO. This process is based on giving input by two methods.

The First method is called multi tap keypad method. The alphanumeric keypad used here consists of 12 keys which are assigned for a function, say, the number 3 has three alphabets 'd', 'e' & 'f'. Depending upon the number of time's the button is pressed, the corresponding alphabet is obtained. To obtain 'e' the user must press 3 (twice).

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

The second method is called 'T9Predictive text' method. T9's objective is to make it easier for the user to give the textual input. It allows words to be entered by a single key press for each letter, as opposed to the multi-tap approach used in conventional mobile phone text entry, in which several letters are associated with each key, and selecting one letter often requires multiple key presses. But here, a combination of single taps of desired numbers (that contains desired alphabets) is enough to create the whole word.

There are cases where the user needs to communicate to a person who is not in his/her locality or is very far from the user. To deal with this, the system is also provided with an option of sending SMS to any desired phone. The user can follow the same procedure to give the textual input a she/she did before. The message can be sent to any phone number specifies by the user. This process does not require any extra hardware component. It, however, requires internet connection.

The RPi ZERO is a mini computer which has AVD processor. The input taken from user (by tapping the buttons) will be decoded from numbers to alphabets by programming code. These alphabets will be processed in the 'Text-to-Speech' application which is developed by Google, where the text will be synthesized to speech. The translator option is also provided where English can be translated to any regional language and be delivered as speech so that even a person who can't understand English can understand what the multi disabled person is trying to express.

If the user is in a noisy environment, the opposite person cannot earth e voice output. So, to overcome this problem, a small is display is attached to the device. Thereby, the opposite person can read the text output and understand what the user is trying to express.

2. Existing Systems

There are few concepts which are theoretically coincident to this proposal. Though the methodologies used are quite different. The relevant documents proposed by the other researchers are discussed below.

Sign language hand gestures, orientation and movements of the hands, arms or body, facial expressions and lip patterns are used for conveying messages. These are computed by image processing techniques and converted to readable formats [3]. The work focuses on finding a unique technique that aids the visually impaired by letting them hear what is represented as text and it is achieved by the technique that captures the image through a camera and converts the text available as voice signals. All these activities are coordinated with the use of Raspberry Pi. The visually impaired people are helped by the process in which the image to text and text to speech is given by the Tesseract OCR (online character recognition)[2]. Another paper presents a hand gesture recognition system or camera to translate the predefined hand gestures into Japanese. The hand gestures are defined based on simplified Japanese sign language. The KNN method is adopted to the pattern recognition according to the bimanual sensing data[4].

A sort of Chinese language Phonetic info yield framework code, phonetic code is utilized as information and yield signal. The client of the framework can express the underlying and find ofa syllable, tone, nourishing and other. Phonetic data by making an appendages development towards input gadget. The originator of the framework expects to let hard of hearing quiets input the Chinese phonetic code to phoneme-shaping gear to converse with others normally, character information can be done if the phonetic code contribution to world-processor[6].

In the existing systems, the main challenges are as follows:

- The problem is the placement of the whole device. The camera needs to be in front of the user at all time to read the input. It is not portable and the time of processing of image is very long producing output at a delayed time. There are chances that the camera captures any other object instead of the hand gestures resulting in wrong output. But our device doesn't require camera. It can be carried anywhere and easily portable which can produce output with maximum accuracy, provided that the user gives the input correctly.
- The already proposed theory uses the image processing technique, where the image can be converted to desired binary format which is processed accordingly. The image processing technique is generally time consuming and its accuracy is not up to the mark. The system must be trained before recognising various objects.
- The gesture sensing involves sensing of movements of hands used in sign languages to convert it system understandable format. The main demerit of gesture recognition is that the input sensing accuracy is very acute. Sign languages have similar gestures for different meanings, hence it can misinterpret the input. And each word must be trained with suitable gestures which is a tedious process.

The basic operation of the existing system is given below. (refer Fig 1)

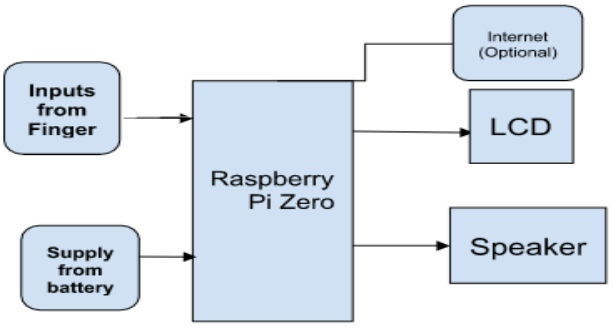

Figure 1: Basic operation of the existing system

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

3. Principle

Alphanumeric keypad is basically used in telephones and old mobile phones for sending text messages. To type a letter, the user must press the respective number buttons. The number of times the user should press the button, depends on the position of the alphabet in the number. This method of a text entry system is generally used for phones. The alphabet is printed under each key (beginning on "2") in a three-letter sequence as follows: ABC under 2 key, DEF under 3 key, etc. Exceptions are the "7" key, which adds a letter ("PQRS"), and the "9"key which includes "Z". Spaces are given by pressing"1"button.'#'isused for deleting an input. The system is used by repeatedly pressing the same key to cycle through the letters for that key. For example, pressing the "3" key twice would indicate the letter "E". After entering all the desired input, "0"is pressed for further processing of input, where the numbers are decoded into alphabets and later processed into speech. It is commonly used in conjunction with text messaging services.

The layout of the digit keys is different from that commonly appearing on calculators and numeric keypads.

Number	Letter
0	Enter
1	SPACE
2	ABC
3	DEF
4	GHI
5	JKL
6	MNO
7	PQRS
8	TUV
9	WXYZ
#	DELETE

Figure 2: The layout of digital keys

(I) T9 Predictive:

This method uses a special algorithm called 'Trie Algorithm'. It combines the groups of letters on each phone key with a fast-access dictionary of words. It looks up in the dictionary all words corresponding to the sequence of keypresses and orders them by frequency of use. As T9 "gains familiarity" with the words and phrases the user commonly uses, it speeds up the process by offering the most frequently used words. When certain numbers are entered, T9 looks up words in its fast access dictionary. When a numerical sequence could yield various words. For example, "4663" could also spell "hood,". Each time a key (1-9) is pressed (when in a text field), the algorithm returns a guess at what letters are most likely for the keys pressed to that point. Another example is to enter the word 'the', one would press 8 then 4 then3.

4. Components and Software

(I) Raspberry PiZERO:

The Raspberry Pi (RPi ZERO) is a scratch and dent section estimated, streak drive measured PC which can be effectively stopped to a presentation. We use a standard console. It is brief gadget that empowers individuals of any age to look into processing, and to figure out how to program in dialects like Scratch and Python. It can do all that we would expect a work area data preparing framework to perform, surfing from the net and playing superior quality recordings, to making databases, word-handling, and live games. RPi has the ability to intercommunicate with the outside world and has been utilized in an extensive cluster in computerized ventures, from music gear and sensors to climate stations and twittering perch rooms with infrared cameras.

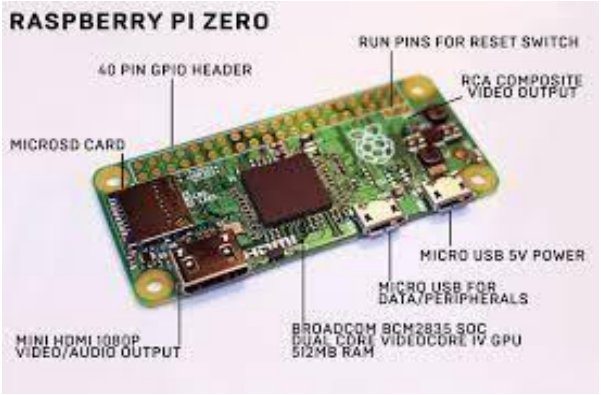

Figure 3: Raspberry pie zero

The Pi ZERO is a lot littler than the other RPi models. The RPi ZERO utilizes a Broadcom BCM2835 SoC with a 1 GHz single-center ARM1176JZF-Sprocessor, with 512KB shared L2 reserve. The RPi was intended for the Linux working framework, and numerous Linux disseminations currently have a rendition advanced for the RPi. This chip can perform numerous functionalities. For our situation we utilize RPi for the accompanying reasons, Accepting numeric contributions from the client by methods for GPIO sticks and putting away it in its memory. Decoding the arrangement of numbers put away to real letter sets by utilizing the proposed calculation. Storing the decoded English message to a book document. Performing content to discourse investigation on the content record. Storing the dissected discourse yield sound to a document.

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

Playing the yield record by methods for sound intensifier (Speakers). Sending the literary information to associated LCD. Sending SMS containing content to any ideal number utilizing web association

(II) Text-to-Speech:

Google TTS (Text-to-Speech) publicly released API empowers engineers to change over content to sound by applying amazing neural system models in a simple to utilize API. The API perceives more than 110 dialects and variations, to help your worldwide client base. The goal of this API is to give a simple method to utilize content to-discourse yield by Google on your RPi. The content backings perusing from standard info, plain content documents, and featured content. Google forces a 100-character limit on their discourse union assistance that makes it difficult to utilize their TTS framework for something besides short sentences.

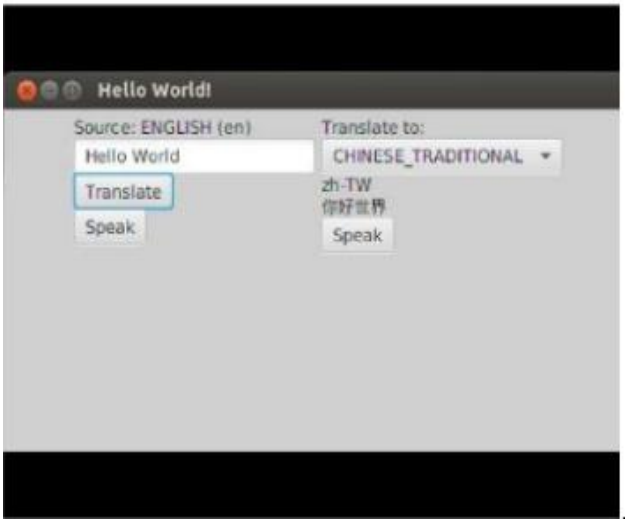

Figure 4: Translate working

Be that as it may, our API works around this impediment by separating the content contribution to suitable pieces. These pieces are set cleverly dependent on accentuation and linguistic structure of the content. Having prepared all pieces, it then connects the discourse sections into one sound record while shortening portions of quietness toward the beginning and end of each part. The one residual issue with this methodology is that the holding up time between client info and voice playback scales radically with the length of the content. In this way, rather than passing the content legitimately, we first separate the contribution to sections. The passages are then handled individually with each section being played back while the following one is blended. Any length of content can be parsed with sensible speed right now.

This API, as said previously, perceives more than 110 dialects. Thus, by conjuring legitimate directions, one can orchestrate content (From the accessible 110 dialects) into discourse with characteristic voice. The voice model can likewise be tweaked.

(III) PicoTTS:

The PicoTTS arrange uses disengaged Content to Discourse engine to scrutinize a book with normal sounding voices. This is a port of the separated substance to-talk engine from Google's Android working structure to the RPi. Pico TTS is expressly expected for embedded devices making it ideal for memory and getting ready constrained stages like the RPi. PicoTTS is SVOX's open source allocation of their disengaged substance to-talk engine. PicoTTS manages the yield when the RPi is detached Decipher:

Decipher Shell is an order line interpreter. Decipher shell API gives a basic automatic interface to making an interpretation of a discretionary string into any upheld language utilizing Neural Machine Translation. Interpret shell API is profoundly responsive, so RPi can incorporate with Translate shell API for quick, unique interpretation of source content from the source language to an objective language (e.g., English to Tamil). This API filters the decoded content (which might be in any language) from the RPi and makes an interpretation of it to the objective language. The handled yield is put away in another document.

(IV) SMS:

For the most part, SMS is sent utilizing GSM or by utilizing some other equipment segment. Yet, here the framework utilizes just an outsider API to send the ideal instant message to any ideal telephone number at zero expense. The gadget is customized to acknowledge both instant message and telephone number through the gloves catches. Be that as it may, this procedure requires web association.

5. Structure of the Model

A glove is worn by the client. The glove is fitted with many number of catches. The catches are set in every one of the three sections of fore finger. Thus, all different catches are set in all fingers. The catches are wired to the GPIO pins of RPi ZERO. RPiZERO is set some place close to the wrist. Each catch means specific number. Squeezing the thumb finger against wanted finger will enlist the relating contribution to the RPi ZERO for additional handling. This gadget does exclude an outer presentation to check if the info button is squeezed effectively. Along these lines, to conquer this issue a little vibrator, associated with RPiZERO, is set on our hand. The framework is customized so that, at whatever point the client contacts the info button, the vibrator vibrates for a milli-second asserting that the ideal information is enlisted to the RPi ZERO effectively. Right now, dazzle individual will realize that the catch is squeezed accurately. A smaller than usual speaker is set close by for playing the prepared yield. A little LCD is additionally set adjacent to the RPi ZERO to show the literary yield.

Both the previously mentioned procedures occur all the while.

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

The entire procedure doesn't utilize console, mouse or screen. The gadget works headless. Each phase of the interpreting and orchestrating process is summoned naturally and doesn't require human mediation. The framework is intended to begin the necessary projects at boot. When the necessary projects are booted at fire up, we hear an asserting message through speakers. This will be helpful for the client to affirm that he/she is prepared to give the contribution for handling

6. Working

As the RPi boots up, a program is begun which is planned to acknowledge numeric contributions from the client. The framework is furnished with a switch, where the client can choose any mode or strategy for giving info (Multi tap technique or T9 prescient content technique). The gadget is likewise encouraged with the choice to choose any language of voice and show yield. When the numeric catch is squeezed by the client, the framework is customized to such an extent that the squeezed number is perceived and kept in touch with a book record. "Spaces" between words are given reasonably by squeezing "1" button. Subsequently, as the client enters a mix of numbers (With reference of the word reference gave), the content record (with numbers) is presently fit to be decoded. In the event that the client nourishes the framework with undesired or wrong data sources, it very well may be erased by squeezing the 'Erase' Key. Utilizing the decoder program, the framework disentangles the alphanumeric numbers into significant English sentences. This yield is spared to another content record. Alternatively, as referenced prior, if the client needs the yield discourse and show to be in some other language, the decipher program is summoned. This will utilize the "Decipher shell" Programming interface. This Programming interface will take the decoded message put away in content record as the info. The content is then transferred to its servers and the made an interpretation of content is come back to RPi. It is put away to another content document.

The RPi tests whether the gadget is associated with web. Assuming Genuine, the "Google TTS" is activated. The (Content in Decoded content record) is sent to Google's servers to produce the discourse document which is then come back to your RPiand played utilizing player. This implies you will require a web association for it to work, yet the discourse quality is magnificent. In the event that there is no web association, "PicoTTS" is activated. Pico TTS has similar functionalities of Google TTS, yet it doesn't require web association. The nature of voice is somewhat diminished. The documents required for TTS preparing is locally accessible. In this way, the procedure is significantly quicker than GoogleTTS. The yield is played through speakers associated with the RPi. What's more, moreover, the relating English content yield is shown in LCD.

On the off chance that the client needs to send the decoded instant message to any telephone number, he/she needs to press the "SMS" button. When the "SMS" button is squeezed, the framework will be prepared to acknowledge the beneficiary telephone number. In the wake of entering the telephone number, utilizing comparative info technique as utilized previously, the 'Send' button is squeezed. The SMS with the instant message is sent to the ideal telephone number.

7. Flowchart

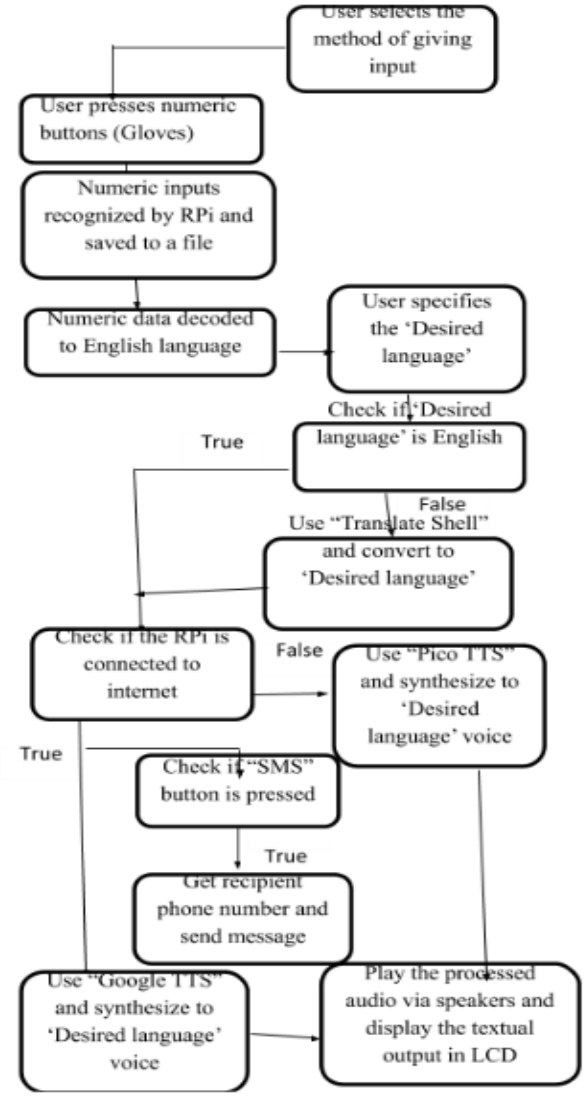

Figure 5: Flow chart

8. Conclusion

This project would enhance the employment opportunities, communicating abilities for the people with multiple disabilities. This method takes less time to produce the output. This can help the mentally challenged to express their views as well. The code will be generated for each alphabet and this can be processed and converted to voice.

May-June 2020
ISSN: 0193-4120 Page No. 5303-5308

The output can be obtained in the speaker. This is more like a mute person being able to speak through their fingers

References

[1] Kathawut Rojanasaroch, Teeravisit Laohapensaeng, School of Information Technology, "Communication aid device for illness deaf-mute"

[2] Shiyam Raghul M, Surendhar K , Suresh N, R. Hemalatha, Electronics and Communication Engineering, Sri Ramakrishna Engineering Colleg "Raspberry-Pi Based Assistive Device for Deaf, DumbPeople"

[3] Nikita P. Nagori, Vandana Malode , Jawaharlal Nehru Engineering College, Aurangabad "Communication Interface for Deaf-Mute People using Microsoft Kinect"

[4] Kengo Kuroki, Yiming Zhou, Graduate School of Computer Science and Engineering "A remote conversation support system for deaf-mute persons based on bimanual gestures recognition using finger-worn devices"

[5] Xu Wang, Lifang Xue, Dan Yang "Speech Plot Display for the Deaf-mute based on Combined Characters Encoding of Speech Signal"

[6] Jie Huang, Wengang Zhou, Houqiang Li, Weiping Li "Sign Language Recognition using 3D convolutional neural networks"

[7] Shujjat Khan, Donald Bailey, Gourab Sen Gupta Xu Wang, Lifang Xue, Dan Yang "Speech Plot Display for the Deaf-mute based on Combined Characters Encoding of Speech Signal"