An Efficient Voice Based Mail for Differently-Abled Persons

Dr. V Srinadh[a], M. Shyam Sai Satish[b]

[a]Computer Science & Engineering, Assistant Professor,
GMR Institute of Technology,Rajam-532127, Jawaharlal Nehru Technological University,
Kakinada, Andhra Pradesh, INDIA, srinadh.v@gmrit.edu.in

[b] Computer Science & Engineering, Student,
GMR Institute of Technology,Rajam-532127, Jawaharlal Nehru Technological University,
Kakinada, Andhra Pradesh, INDIA, shyamsaisatish@gmail.com

Abstract: Pattern recognition is an overall innovation for discovery and arrangement of patterns in information. Pattern recognition is a human capacity. Patterns are characterized by individuals. Recognition is subsequently identified with finding ideas and the specialty of naming. Thus, Pattern recognition is an exceptionally fundamental space of science. Pattern recognition is field of examination, which is in presence from quite a while. One of the utilizations of Pattern recognition is voice recognition. Voice recognition is a significant space of improvement and is executed in advanced partners. Voice recognition has gotten a gigantic change the innovation. Voice recognition assisted a typical individual with communicating openly with a framework or a gadget. It doesn't need a specific ability to be drilled by the client. In this paper we use voice to produce mail for dazzle individuals. This will assist with blinding individuals to ease from the hardships looked during the way toward mailing.

KEYWORDS: Pattern, Recognition, Artificial intelligence, Voice recognition, Digital assistant.

1. INTRODUCTION

Pattern is everything around in this advanced world. A Pattern can either be seen genuinely or it very well may be noticed numerically by applying calculations. Pattern recognition is the way toward perceiving designs by utilizing AI calculation. Pattern recognition can be characterized as the arrangement of information dependent on information previously acquired or on measurable data separated from designs and their portrayal. In a common Pattern recognition application, the crude information is prepared and changed over into a structure that is amiable for a machine to utilize. Pattern recognition includes grouping and bunch of patterns.

Visually impaired and debilitated individuals are distinctively abled that they can have contrasts in connecting with normal things in day by day life. To facilitate their hardships many apparatuses have been created. Innovation has made an unrest in opportunities for debilitated students. Assistive innovation empowers understudies with inabilities to make up for specific hindrances. This particular innovation advances autonomy and diminishes the requirement for other help. Assistive advancements become more normal spot, individuals without handicaps are profiting with them. For instance, individuals for whom English is a subsequent language are exploiting screen perusers. More seasoned people are utilizing screen enlargers and magnifiers. One of these hardships are

getting to mail and playing out the mail exercises. Our device chiefly focuses on aiding these individuals utilizing straightforward voice orders.

2. LITERATURE REVIEW

Daze individual is somebody who has outwardly debilitation. Stick is an instrument help used to help their every day action. A traditional stick actually have restricted capacity in recognition impediment and opening so not all action are helped utilizing stick. One of troublesome experience is to decide the heading of qibla at whatever point they will supplicate. This examination foster capacity of stick as an apparatus that can decide the bearing of qibla by voice order intelligently [2].In this paper voice acknowledgment is utilized as an application in everyday life utilization of visually impaired individuals. It is utilized in strolling stick of visually impaired individuals. It can decide and direct headings utilizing a HMC5883L compass and voice colleague remembered for the stick. The correspondence is done in a wireless way. The clamor in the climate impacts the productivity of the framework radically.

In this day and age break of safety is a significant concern, voice based confirmation is one of the major biometric acknowledgment methods. It is partitioned into two phases, highlight extraction and element coordinating. The exhibition of the voice put together verification framework depends with respect to its exactness. In this paper voice acknowledgment is utilized as an instrument in the everyday life. The principle cycle of voice acknowledgment utilizes two techniques include extraction and element planning. Utilizing these two cycle voice is perceived and a few times changed over to message. Distinctive component extraction methods are looked at in this paper. Since voice acknowledgment includes recurrence as a property, the creator cited about voice quality [4].

The Voice is a sign of boundless data. Advanced handling of discourse signal is vital for high velocity and exact programmed voice acknowledgment innovation. These days it is being utilized for medical care, communication military and individuals with handicaps in this manner the advanced sign cycles, for example, Feature Extraction and Feature Matching are the most recent issues for investigation of voice signal. To separate important data from the discourse signal, settle on choices on the interaction, and acquire results, the information should be controlled and examined. Essential strategy utilized for removing the provisions of the voice signal is to discover the Mel recurrence cepstral coefficients [6].In this paper voice acknowledgment methods are been examined exhaustively. Voice acknowledgment methods include highlight determination and component extraction. It likewise portrays how we can extricate information from voice utilizing different procedures cost productively.

Solid discourse acknowledgment is a difficult issue, requiring a mix of numerous unpredictable strategies; anyway current techniques have had the option to accomplish a noteworthy level of exactness. Then again, today, the vast majority of the organizations or establishments are leading their assessments online to be a piece of this best truly developing world. In this framework client can give any accessible assessment at any open community according to his/her decision and authority additionally can consolidate labor and cycle defer overhead [7]. This paper depends on execution of voice acknowledgment in biometric apparatuses. The creator tells about the execution of voice innovation in online assessment framework for genuinely handicapped individuals.

Voice is the fundamental, normal and productive type of specialized strategy for individuals to cooperate with one another. Today discourse advancements are usually accessible for a restricted however fascinating scope of errand. These advancements empower machines to react effectively and dependably to human voices and offer helpful and important types of assistance. As speaking with PC is quicker utilizing voice as opposed to utilizing console, so individuals will incline toward such framework. Correspondence among the individual is overwhelmed by communicated in language, subsequently it is normal for individuals to expect voice interfaces with PC [8]. This paper bargains about discourse to message transformation method engaged with voice acknowledgment. The creator bargains about various kinds of discourse acknowledgment expressions like secluded word, associated word, consistent discourse and unconstrained discourse. The creator characterizes the models as speaker subordinate models and speaker autonomous models3.

3. METHODOLOGY

3.1 Speech Recognition:

Speech recognition is the capacity of a machine or program to recognize words and expressions in communicated in language and convert them to a machine-intelligible arrangement. Simple discourse acknowledgment programming has a restricted jargon of words and expressions, and it might possibly recognize these in case they are spoken plainly. More refined programming can acknowledge normal discourse.

Speech recognition works utilizing calculations through acoustic and language displaying. Acoustic displaying addresses the connection between etymological units of discourse and sound signs; language demonstrating matches sounds with word successions to help recognize words that sound similar hidden Markov models are utilized also to perceive worldly examples in discourse to further develop precision inside the framework. The most continuous utilizations of discourse acknowledgment inside the undertaking incorporate call steering, discourse to-message preparing, voice dialing and voice search.

Before any machine can decipher discourse, an amplifier should interpret the vibrations of an individual's voice into a wavelike electrical sign. This sign thusly is changed over by the framework's equipment for example, a PC's sound card into an advanced sign. It is the advanced sign that a discourse acknowledgment program investigates to perceive separate phonemes, the essential structure squares of discourse. The phonemes are then recombined into words. Be that as it may, many words sound the same, and, to choose the suitable word, the program should depend on the specific situation. Many projects build up setting through trigram examination, a technique dependent on an information base of regular three-word groups in which probabilities are allocated that any two words will be trailed by a given third word. For instance, if a speaker says "who am," the following word will be perceived as the pronoun "I" as opposed to the comparable sounding yet more outlandish "eye." Nevertheless, human intercession is in some cases expected to address mistakes.

Speech emotion recognition essentially comprises of three cycles, i.e., highlight extraction, include choice, and feeling characterization. at first list of capabilities including speaker-autonomous components and speaker subordinate elements is acquired get-togethers exaction. Besides, the

extricated highlights are thought about by connection examination and ideal element subset is acquired get-togethers choice. At last, discourse feeling is characterized into a few classes utilizing the chose features. In the element extraction, the blending highlight set including speaker-autonomous elements (i.e., enthusiastic elements, for example, normal change pace of major recurrence that could take out the effect of individual contrasts) and speaker subordinate components (i.e., passionate elements, for example, beat qualities that are handily impacted by the speaker's close to home attributes) is acquired from preprocessed discourse signal examples.

3.2 Voice Based Mail:

A voice based mail is an tool which can reduce the complexity of usage of mail by disabled people. We can use different modules to input speech commands and output the voice data. The user doesn't need to have any knowledge on the usage of the product. He just needs to follow the instructions carefully. He needs to have an idea about the language and should be able to understand the language. The user needs to give inputs using voice. So he needs to see that there should be no noise around him.

3.3 Speech to Text:

Speech to text transformation is the way toward changing over expressed words into composed writings. This cycle is additionally regularly called discourse acknowledgment. Albeit these terms are practically equivalent, Speech acknowledgment is now and again used to portray the more extensive interaction of separating importance from discourse, for example discourse understanding. The term voice acknowledgment ought to be kept away from as it is normal related to the way toward recognizing an individual from their voice, for example speaker acknowledgment.

All discourse to-message frameworks depend on somewhere around two models: an acoustic model and a language model. Furthermore enormous jargon frameworks utilize an elocution model. Comprehend that there is nothing of the sort as an all inclusive discourse recognizer. To get the best record quality, these models can be specific for a given language, tongue, application space, sort of discourse, and correspondence channel.

Like some other example acknowledgment innovation, discourse acknowledgment can't be without mistake. The discourse record exactness is exceptionally reliant upon the speaker, the style of discourse and the ecological conditions. Discourse acknowledgment is a harder interaction than people's opinion, in any event, for an individual. People are accustomed to getting discourse, not to interpreting it, and just discourse that is very much figured can be translated without vagueness.

According to the client's perspective, a discourse to-message framework can be sorted situated in its utilization: order and control, exchange framework, text correspondence, sound archive record, and so on Each utilization has explicit prerequisites as far as dormancy, memory requirements, jargon size, and versatile provisions.

As indicated by fig 3.1 Speech-to-message programming is a kind of programming that successfully takes sound substance and translates it into composed words in a word processor or other showcase objective. This kind of discourse acknowledgment programming is incredibly important to any individual who needs to produce a great deal of composed substance without a ton of manual

composing. It is likewise helpful for individuals with inabilities that make it hard for them to utilize a console. Discourse to-message programming may likewise be known as voice acknowledgment programming.

Despite the fact that discourse to-message programming is generally sold as an independent application, it has likewise been incorporated into fresher working frameworks for certain gadgets. Most discourse to-message programming programs pointed toward helping with record center around perceiving a wide scope of jargon from a solitary client or a restricted arrangement of clients, instead of perceiving a more modest scope of jargon from a bigger client base.

As far as specialized capacity, numerous discourse to-message programming programs separate verbally expressed word sound into short "examples" and partner those examples with basic phonemes or units of articulation. Then, at that point, complex calculations sort the outcomes to attempt to anticipate the word or expression that was said. Discourse to-message programming has improved a considerable amount in exactness and advanced in everyday usefulness to assume a bigger part in present day correspondences over computerized stages.

Fig 3.1 Speech to text

3.5 Text to Speech:

Text-to-speech is a type of assistive technology that reads digital text aloud. It's sometimes called "read aloud" technology.With a click of a button or the touch of a finger, TTS can take words on a computer or other digital device and convert them into audio. TTS is very helpful for people who struggle with reading. But it can also help people with writing and editing, and even focusing. TTS works with nearly every personal digital device, including computers, smart phones and tablets. All kinds of text files can be read aloud, including Word and Pages documents. Even online web pages can be read aloud.

The voice in TTS is computer-generated, and reading speed can usually be sped up or slowed down. Voice quality varies, but some voices sound human. There are even computer-generated voices that sound like children speaking. Many TTS tools highlight words as they are read aloud. This allows kids to see text and hear it at the same time. Some TTS tools also have a technology called optical character recognition (OCR). OCR allows TTS tools to read text aloud from images. For example, your child could take a photo of a street sign and have the words on the sign turned into audio. Text to speech systems were first developed to aid the visually impaired by offering a computer-generated spoken voice that would "read" text to the user. TTS should not be confused with voice response systems. Voice response systems synthesize speech by concatenating sentences from a database of prerecorded words and are used for different purposes than TTS systems, which form sentences and/or phrases based on a language's graphemes and phonemes.

Voice response systems are limited to synthesizing sentences that contain only words that have been predetermined by the system. TTS systems, in contrast, are theoretically capable of "reading" any string of text characters to form original sentences.

Types of text to speech tools :

1. Built in text to speech.

2. Web based tools.

3. Text to speech apps.

4. Chrome tools.

5. Text to speech software programs.

4. EXPERIMENTAL RESULTS

We use anaconda prompt to open our server and start the program using the below commands.

Fig 4.1 commands to run UI(Server)

4.1 Sending A Mail :

Fig 4.2 Sending a Mail

First we execute the complete program with the help of command (shift + enter). After entering the program we have to navigate the mail using the voice.

The program first gives a list of details (or) options like sending a mail, counting total number of mails, reading number of unread mails.

After inputting the data like send a mail, the program asks about the mail id of the recipient. The user should convey the data clearly along with the correctness of the data. The program then asks the user about the body of the mail.

As we input the body of the mail, we login with the users id and password. The session starts by using the different protocols that are already embedded in the program.

All the operations done by the program are safe and encrypted so that there is no illegal access of the data.

4.2 Reading Total Number Of Mails :

```
------------------------------------------------------------
        Project: Voice based Email for visually impaired
------------------------------------------------------------
You are logged In from : shyam
a. Compose a mail.
b. Check total no of mails.
c. Check unread mails.
Your choice:
ok done!!
You said : b
From : iB Hubs <info@ibhubs.co>

Subject : iB Hubs Startup School =?utf-8?Q?=E2=80=9820?= is accepting
 applications | Apply Now

From : shyamsai satish <shyamsaisatish@gmail.com>

Subject : Hi

From : Google <no-reply@accounts.google.com>

Subject : Security alert

From : ProtonVPN <contact@protonvpn.com>

Subject : All ProtonVPN apps are now open source

From : Google <no-reply@accounts.google.com>

Subject : Security alert for your linked Google Account
```

Fig 4.3 Reading Total Mails

```
 shyam, your new Xiaomi Redmi 5A doesn't have the latest Google apps

From : Google <no-reply@accounts.google.com>

Subject : Security alert

From : Google <no-reply@accounts.google.com>

Subject : Security alert

From : shyamsai satish <shyamsaisatish@gmail.com>

Subject : termpaper all

From : shyamsai satish <shyamsaisatish@gmail.com>

Subject : Fwd: termpaper all

56
```
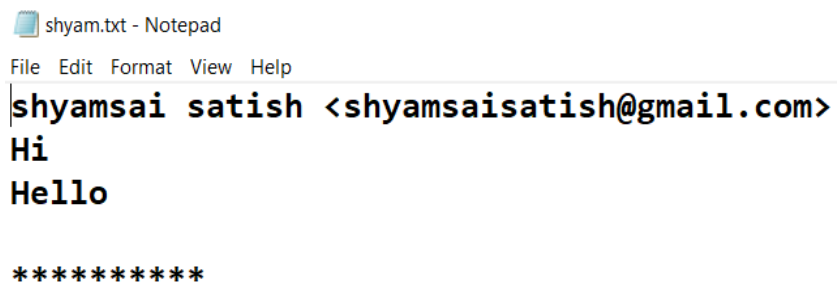
Fig 4.4 Reading total count of mails

This is similar to the previous module but it retrieves the data from the mail without any voice command and reads the mail according the order.

4.3 Check The Number Of Unread Mail:


```
-------------------------------------------------------------
        Project: Voice based Email for visually impaired
-------------------------------------------------------------
You are logged In from : shyam
a. Compose a mail.
b. Check total no of mails.
c. Check unread mails.
Your choice:
ok done!!
You said : c
0
```

Fig 4.5 Check Total Number of Unread Mails


```
shyam.txt - Notepad
File  Edit  Format  View  Help
shyamsai satish <shyamsaisatish@gmail.com>
Hi
Hello

**********
```

Fig 4.6 Saving Unread Mails In a File

This module reads the total number of unread mails in the drive and saves (or) creates a .txt file in the device. After completion of the process, the program automatically marks the mail as read.

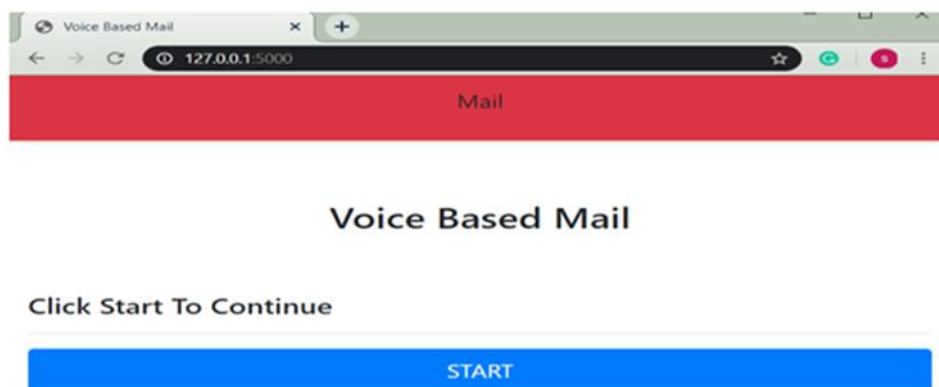The final UI(user interface) which can be used by the users freely is as follows:

Fig 4.7 Start UI

5. CONCLUSION AND FUTURE WORK

Since the visually impaired and handicapped individuals think that its hard to utilize mail service easily, this apparatus can help in facilitating the troubles looked by them. It doesn't need any specialized information for the client to utilize this apparatus. This can help in making a lot of utilization of the components accessible via the mail. The client needn't bother with any preparation to be taken to utilize the device. He/she can without much of a stretch discover the activities done by the apparatus. The client can comprehend the guidelines given by the instrument thinking about that he comprehends the language great. This email framework can be utilized by any client of all ages bunch easily of access. It has feature of speech to text as well as text to speech with speech reader which makes designed system to be handled by visually impaired person as well as blind and disabled persons. In future we can deploy this application on an server and can be used at any place.

6. REFERENCES

1. Zinchenko, K., Wu, C. Y., & Song, K. T. (2016). A study on speech recognition control for a surgical robot. IEEE Transactions on Industrial Informatics, 13(2), 607-615.
2. Asrin, A., Hapsari, G. I., &Mutiara, G. A. (2018, May). Development of Qibla Direction Cane for Blind Using Interactive Voice Command.In 2018 6th International Conference on Information and Communication Technology (ICoICT) (pp. 216-221).IEEE.
3. Gumelar, A. B., Kurniawan, A., Sooai, A. G., Purnomo, M. H., Yuniarno, E. M., Sugiarto, I., ...&Fahrudin, T. M. (2019, August). Human Voice Emotion Identification Using Prosodic and Spectral Feature Extraction Based on Deep Neural Networks. In 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH) (pp. 1-8). IEEE.
4. Hundal, J. K., &Hamde, S. T. (2017, September). Some feature extraction techniques for voice based authentication system. In 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI) (pp. 419-421). IEEE.
5. Xu, X., Li, Y., Xu, X., Wen, Z., Che, H., Liu, S., & Tao, J. (2014, September). Survey on discriminative feature selection for speech emotion recognition.In The 9th International Symposium on Chinese Spoken Language Processing (pp. 345-349).IEEE.
6. Bala, A., Kumar, A., & Birla, N. (2010). Voice command recognition system based on MFCC and DTW. International Journal of Engineering Science and Technology, 2(12), 7335-7342.
7. Rudrapal, D., Das, S., Debbarma, S., Kar, N., Debbarma, N., &Rudrapal, D. (2012). Voice recognition and authentication as a proficient biometric tool and its application in online exam for PH people. *International Journal of Computer Applications*, *39*(12), 6-12.
8. Prasad, V. (2015). Voice recognition system: speech-to-text. *Journal of Applied and Fundamental Sciences*, *1*(2), 191.
9. Dubagunta, S. P., Kabil, S. H., & Doss, M. M. (2019, May). Improving Children Speech Recognition through Feature Learning from Raw Speech Signal. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5736-5740). IEEE.
10. Jain, A. K., Ross, A., &Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, *14*(1), 4-20.

Dr. V Srinadhreceived the Ph.D degree in Computer Science from the University of GITAM University, Visakhapatnam in March, 2019. He is currently an Assistant professor of Computer Science and Engineering at GMRIT, Rajam, Andhra Pradesh. His areas of interests: Network Security, Cyber Security, Group communication and Machine Learning.

Dr. V Srinadh, M. Shyam Sai Satish

Mr. M. Shyam Sai Satish doing his B.Tech in Computer Science & Engineering branch inGMR Institute of Technology,Rajam-532127, affiliated to Jawaharlal Nehru Technological University, Kakinada, Andhra Pradesh, INDIA. His areas of interests: Network Security, Group communication Data Mining and Machine Learning.