International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

AI SPEECH RECOGNITION SYSTEM IN FACEBOOK APPLICATION FOR VISUALLY IMPAIRED PERSONS

[1]**Dr.P.Kavitha** *M.E., Ph.D., Dept.of Computer Science Engineering, P.S.R.ENGINEERING COLLEGE, Tamilnadu, India.*

[2]**K.Amuthachenthiru** *M.E., Dept. of. Computer Science Engineering, P.S.R. Engineering College, Tamilnadu, India.*

---***---

Abstract **-** Internet is one of the basic luxury for daily living. Every person is using the facts and information on internet. On the other hand, blind people face difficulty in accessing the text resources. The development in computer based handy systems has opened up numerous oppurtunities for the visually disabled across. Audio response based virtual environment, the screen readers are helps blind people a lot to use internet applications. This project introduces the Voicemail system structural design that can be used by a blind person to access E-Mails easily. The involvement of research is helping blind individual to send and receive voice based mails messages in their inhabitant language with the help of a computer.

1. INTRODUCTION

Speech Recognition (SR) is the ability to translate a dictation or spoken word to text.

Speech Recognition known as "automatic speech recognition" (ASR),or speech to text(STT)

Speech recognition is the process of converting an acoustic signal, captured by a microphone or any peripherals , to a set of words .

To achieve speech understanding we can use linguistic processing

The recognized words can be an end in themselves, as for applications such as commands & control data entry and document preparation.

In the society every one either human or animals wish to interact with each other and tries to convey own message to others . The receiver for messages may get the exact and full idea of the senders, or may get the partial idea or sometimes can not understand anything out of it.

In some cases may happen when there is some lacking in communication (i.e when a child convey message, the mother can understand easily while others can not)

2.The process of conversion from speech to words is complex and varies slightly between systems.

It consists of three steps :

(1) Feature extraction – Pre-processing of the speech signal, extracting the important features into feature vectors.

(2) Phoneme recognition – bases on a statistically trained phoneme model (HMM) the most likely sequence of phoneme is calculated.

(3) Word recognition – Based on statistically trained language model similar to the phoneme model, the most likely sequence of word are calculated.

Speech dictation process :

After the preparation of master database of features of Gujarati Alphabets, the researcher has proposed the dictation model from where the actual human- machine interaction starts in the form of speech dictation. The researcher has divided the model into five different steps (Fig:9.4) i.e. (1) Input acquition

(2) Front end (3) Feature extractor (4) local match and (5) character printing.

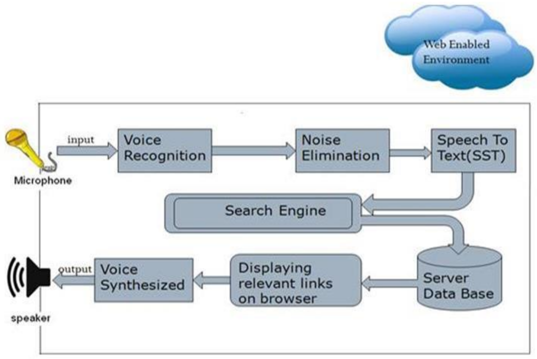

Fig 1. Text To Speech Diagram

The planned system is relies on a very fresh plan and obscurity just like the accessible mail systems. The foremost necessary facet that erstwhile unbroken in brain whereas developing the planned system's accessibility.

The present systems don't give this much convenience. So the systems present have a tendency to area unit developing is totally dissent from this system. In contrast to present system which emphasize more on user easiness of naive users, this system focus more on user easiness of all kind of folks including naive folks visually disabled people as well as uneducated people.

The entire structure is based on IVR- interactive voice response. When using this system the computer will prompt the client to perform precise operations to gain relevant services and if the client needs to way in the relevant services then they need to perform that particular operation.

One of the most important recompense of this system is that user will not need to use the keyboard. All operations will be based on voice proceedings.

3. Technique to extract speech features from speech signals

Speech recognition is highly affected by the type of speech i.e. isolated word Vs continuous. One of the hardest problems in speech recognition is determining when one word ends and the next one begin. In order to side step the problem, most systems force the user to issue single word-at-a-time commands. Typically, words must be separated by a gap of on the order of 300 milliseconds. Since this is unnatural, speech recognition systems that require multi-word commands may requires special training on the part of the users. One perspective on this is presented in Biermann et al. (1985).

Various approaches or recognition system are used to extract features of speech. The recognition system describe here, classifies the different features of Gujarati alphabet produce by a speaker. The system has a number of characteristic features. The researcher has performed explicit segmentation of speech signal into phonetic categories. Explicit segmentation allows using segment duration to discriminate letters, and to extract features from specific regions of the signal. Finally speech knowledge is used to design a set of features that work best for Gujarati letters.

Speech separation from noise, given a-priori information, can be viewed as a subspace estimation problem. Some conventional speech enhancement methods are spectral subtraction, Wiener filtering , blind signal separation and hidden Markov modeling.

4. Storage of speech files and its feature in traditional flat file format

The process of data storage in traditional flat file format consists two or more type of the files. Each prompted utterance is stored within a separate file in any valid audio file format. The stored speech file for each utterance is processed with the speech processing tools (i.e. software) and the corresponding features for each utterance is extracted and these processed outcome is stored in the other flat file, which is accompanying each utterance file. For the storage of the features we may use many different approaches as follow:

(1) One may take the separate file for each utterance i.e. for pitch of all the utterance one separate file, for the frequency of all the utterance one separate file and so on. In each feature file each row represents the different utterance. The affiliation of each row with the accompanying utterance must be previously determined. And for every feature file this affiliation remains same. Suppose there are 36 utterances and 10 features then there are 46 files.

(2) All the features (i.e. pitch, frequency and many more) for the one utterance are stored in the one file. For the second utterance again all the features (i.e. pitch, frequency and many more) are stored in the other file and so on. In this approach, every file is named in such way so that it accompanying the utterance. Suppose there are 36 utterances then there are 72 files (i.e. 36 – for utterance and 36 – for features of the accompanying utterance).

(3) All the features (i.e. pitch, frequency and many more) for the one utterance is stored in one line of the flat file separated by either comma or space, second line again stores the same features for the second utterance and so on. In the file format each column represents the same feature for all the utterance and each row represents different features for the one utterance. The affiliation of each feature with the column and affiliation of each utterance with each row must be previously determined. In this approach if there are 36 utterances then there are 37 files.

Fig 2 Main Form

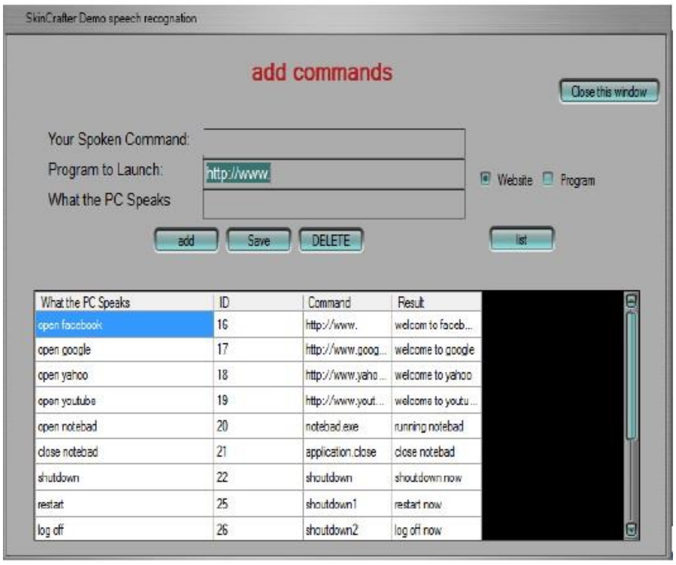

Fig 3 Add Commands

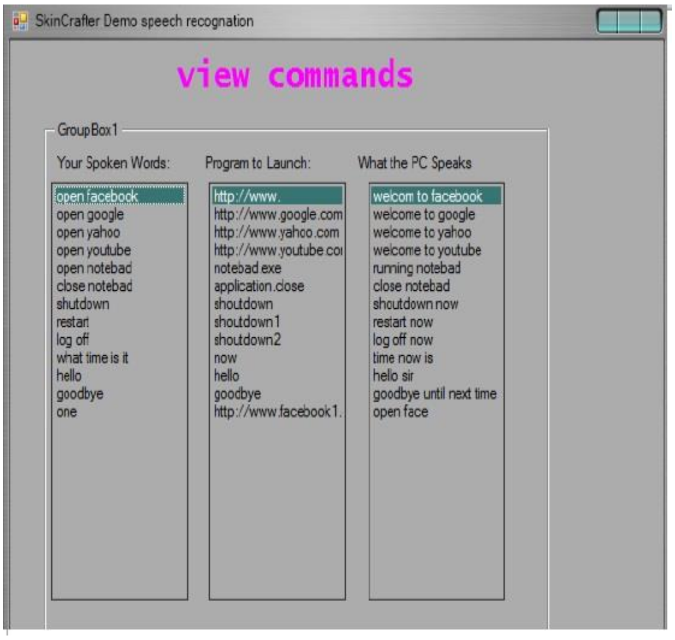

Fig 4 View commands

Fig 5 Grammar is loading

3. CONCLUSIONS

Speech signal is a highly redundant non-stationary signal. These attributes make this signal very challenging to characterise. It should be possible to recognize speech directly from the digitized waveform. However, because of the large variability of the speech signal, it is a good idea to perform some form of feature extraction that would reduce that variability. Applications that need voice processing (such as coding, synthesis, recognition) require specific representations of speech information. For instance, the main requirement for speech recognition is the extraction of voice features, which may distinguish different phonemes of a language. From a statistical point of view, this procedure is equivalent to finding a sufficient for statistic to estimate phonemes. Other information, not require for this aim, such as phonatory apparatus dimensions (that is speaker dependent), the speaker's moods, sex, age, dialect inflexions, and background noise etc., should be overlooked. To decrease vocal message ambiguity, speech is therefore filtered before it arrives at the automatic recognizer. Hence, the filtering procedure can be considered as the first stage of speech analysis. Filtering is performed on discrete time quantized speech signals. Hence, the first procedure consists of analog to digital signal conversion. Then, the extraction procedure of the significant features of speech signal is performed.

REFERENCES

[1] L. Mary Gladence 1, Shubham Melvin Felix2,Aatisha Cyril, "Text Reader forBlind: Text-To-Speech", International Journal of Pure and Applied Mathematics, Volume 117 No. 21 2017, 119-125,2017

[2]D. Freitas, "Speech Technologies for Blind and Low Vision Persons", Technology and Disability, 20(2):135-156,2008

[3] Mrs.Shilpa Reddy K, Mounika S.K,Pooja K, Sahana , "Text to Speech for the Visually Impaired", International Research Journal of Computer Science(IRJCS) ISSN: 2393-9842Issue 05, Volume 4,2017

[4] Samruddhi Deshpande ; Revati Shriram, "Real time text detection and recognition on hand held objects to assist blind people", 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT),2016

[5] Sanjana.B1, J.RejinaParvin,"Voice Assisted Text Reading System for Visually Impaired Persons Using TTS Method", IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 6, Issue 3, Ver. III (May. -Jun. 2016), PP 15-23, 2016