Research Paper

on

INTERACTIVE VOICE RESPONSE SYSTEM WITH SPEECHRECOGNITION

Submitted in partial fulfillment of the requirement for the award of the degree of B.tech CSE

B.Tech in Computer Science Engineering

GALGOTIAS UNIVERSITY
(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision of Name of Supervisor:

Dewan Imdadul Islam

Submitted By:

Name – Priya Tiwari	**Admission No. -** 22SCSE1010536
Name - Rishav Raj	**Admission No. -** 22SCSE1010527
Name – Moh. Yusuf Ali	**Admission No. -** 22SCSE1010545
Name – Vaibhav Singh	**Admission No. -** 22SCSE1010550

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING DEPARTMENTOF COMPUTER SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY, GREATER NOIDA

INDIA

IV semester, October, 2022-2023

Voice Recognition: A research Paper

1 INTRODUCTION

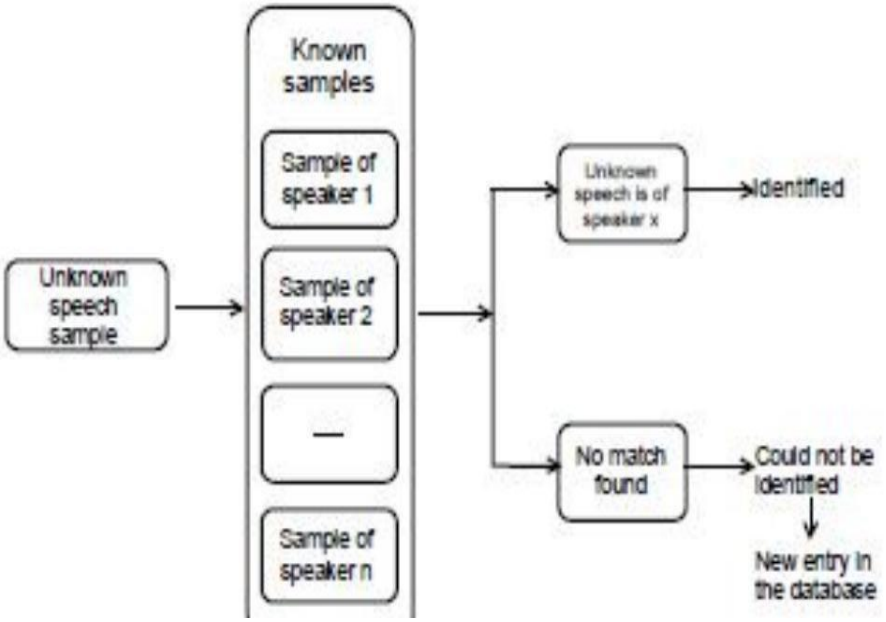
Voice is a basic, common and effective way for people to communicate with one another. Today speech technology is available in a limited but admirable range of functions. This technology allows machines to respond appropriately and reliably to human voices and provide useful and valuable services.

As computer communication is faster than using a keyboard, then people will prefer that program. This can be done by developing a voice recognition system: computer-to-speech that allows the computer Translating voicerequest and calling from text. Voice recognition system: process-to-text is a process of to convert an acoustic signal that is held using a microphone to a set of words. Fundamentally, Speech Recognition, also known as the Automatic Speech Recognition (ASR) process alters speech signal. The speech recognition platform aims to develop speech recognition strategies. The early computer programs were limited to scale and power. But a change in computer technology is already underway the default speech recognition field. Now a day it's easy to keep the big details of speech recognition due advances in computer technology. There are only a few languages that speech recognition systems have developed. Thus, many dimensions exist to construct the native language expressions. Automatic Expression recognition has limited human efforts in many fields, such as automated telephone processing on the telephone network, data entry, voice calling; is a question based on recycled travel and booking information, an understanding of natural languages translators etc. This paper highlights the basic building blocks of speech recognition systems, technology progress and problems in the automatic speech recognition system.[2] [3]

1.1 .About Voice Recognition Program:

The voice recognition system is broadly divided into two ways
A. To identify the Speaker
B. Speaker authentication
The process of identifying the voice of a speech given to a given group of speakers is called a speaker identification a speaker with its highest vibrations is the same as the voice that has been served a speaker with its unmatched voice features eligible for a new entry into the database. Known sets of the sounds programmed in two edges are called Open-set mode and Open-set mode. In open-set mode the speaker does not need to be a part of some known speakers. This is used in the event of a specific criminal activity where it originated Identifying multiple suspects in close-parameter mode, the speaker is part of the others well-known voices are already available in the database. This method is used for biometric proofing purposes security to identify an authorized person. Conversely, speaker authentication the process is the process of accepting or rejecting a speaker's claim. It is used to convince a person's request for authenticity. The Speaker's verification procedure is usually referred to as the open set mode it needs to check the authenticity of the voice from the set of speakers. The most important part of voice recognition software for the identity verification system of any speaker that emphasizes authorization of a particular service. One more the division of the voice recognition scheme is subject to Text and independent text recognition. This is the distinction is included in the text referred to by the speaker. If the text spoken by the speaker is the same the text stored during the training is called the Text Recognition Program. On the other hand, if any The random text presented by the speaker through voice identification is called the Voice

independent text program / program. So there are three ways to distinguish a reliable voice recognition system - The text depends and Text is independent, Open set and Close set and voice recognition and voice verification. The sound of talking says was captured using a microphone to turn it into an electrical signal. Purpose of sound card inside computer Convert an analog signal to a digital signal. The sound card has the ability to store and play this speech signal. [3]



The following are the building blocks for a common speech recognition system.
A. Signal preprocessing
B. Facility extraction
C. Language model
D. decoder
E. Speech recognition [4]

1.2 Types of speech

Speech recognition is divided by what kind of words they can understand. They categorized as:
1) Remote Name: Remote Identifier sometimes requires each spoken word to be silent (lack of sound)signal) in the bot
2) Connected Name: Same with a different name, but allows different expressions to work together "contains at least a break between them.
3) Continuous Speech: allows users to speak naturally and similarly the computer will decidecontent.
4) Sound Speech: It is a kind of speech that is loud and ineffective.
Speech recognition systems can be of different types depending on the types of words to be observed. Theseare the various types are distinguished as follows:

A. Isolated Words: Single word subjects often find that each expression is peaceful on both sides of the sample window. These programs usually have two listening / Voice listening regions, where the speaker has to wait at midnight speeches. The interval between statements is used to process speech signals.
B. Linked names: Linked words are the same as single words with a slight pause difference among them.
C. Continuous Words: Continuous speech recognition involves a natural way of speaking. It's difficult design that continued to speak because it looked for specific ways to determine the boundaries of speech.

D. Default Names: Default expression covers non-verbal, non-verbal and false statements that are difficult to read. the ASR program under this section deals with various aspects such as the words are grouped together like "mask" and "ahs".[1] [3]

2. AVAILABILITY OF KEEPING OF WORDS

This section provides a review of the development of speech recognition technologies. There is also a brief discussion of Various methods has been used to improve the recognition process. A survey of the last few years involved technology has reported significant improvements in the sector.

• The first attempts to develop automatic talk recognition was made in the 1950s when many researchers tried to explore the basic acoustic-phonetic concept.

• In 1952 Davis et.al (1952), set out to establish a motive for digital recognition. The proposed plan was in placewe use the concept of spectacle to show the digital vowel region.

• Olson and Belar (1956), attempted it find ten different characters for a single speaker composed of 10 monosyllabic words.

• In 1959, Fry and Denes tried to build phonetic identifier for vowels and analytical concrete. the screen and thesimilarity. In 1960, a hardware-based approach emerged, where several Japanese laboratories had entered this field.

• Suzuki and Nakata (1961), developed hardware for the recognition of vowels.

• Sakoe and Chiba in Japan propose the importance of robust synchronization systems. In 1970, the recognition of speech-based speech in individual words was a major focus of investigators. Line Predictive Coding (LPC) was optimized for low-level coding and it worked to use speech recognition systems for optimizing its visual parameters.

• Pruthi et al. (2000), developed a one-time word-dependent word recognition. Continuous HMM was used to see a different name for that Hindi language in 2006 was designed by Gupta. Al- Qatab et. al (2010), done is an Arabic speaking system that uses HTK that can recognize both isolated and continuous words.

• R. K. Aggarwal and M. Dave (2011), proposed the recognition of Hindi speech using Gaussian reflective mixtures great accuracy.

• Z. Yu et. al (2014), presented a teaching test for speech recognition based on the Hidden Markov Model (HMM) outlining the HMM speech recognition principle, the process of initiating speech recognition program. [3]

3. PROBLEMS IN TRANSFER RECOMMENDATIONS

The effectiveness of the speech recognition system is dependent on its ineffectiveness of the surrounding environment. There are many factors that affect the level of active recognition such as environmental factors, speaker / independent dependency, level of expression and channel variability. But with a speech recognition system it is necessary to match the changes that occur to fill the gaps.

4. APPLICATIONS

• Search for reports or documents on your computer.
• Create graphs or tables using data.
• State the information you want to include in the document.
• Print documents on request.
• Start video conferencing.
• Schedule meetings.
• Record minutes.
• Make travel arrangements

5. PERFORMANCE EVALUATION

The effectiveness of the speech recognition system is usually measured in terms of accuracy and speed. Program accuracy is evaluated according to Word error rate (WER). Therefore, performance is defined in WordRecognition Rate (WRR) is part of comparing Word error rate (WER). Name errors can be divided by the number of artifacts, replacements and deletions while seeing the speech. These two The performance aspect can be measured in terms of statistics,

Word Error Rate (WER) = I + S + D / N

Where 'I' is the input number, S is the input number, D is the subtraction number and N is the number

words with words. Word Recognition Rate (WRR) = 1-WER of speed, Actual time limit specified; is calculated by following the equation,

Real Time Factor (RTF) = T / D

When T is a time series and D is a duration.[4]

6. CONCLUSION

In this paper the basics are discussed and their recent progress is investigated. Different ways available by building a voice recognition system based on a modified feature translation process and speech recognition language system are compared in this paper. Voice recognition is a computer analysis of the human voice, mainly aimed at translating words as well phrases and regular identification of the speaker on the basis of the individual details included the waves of speech. This process makes it possible to use the voice of the presenter and is easy to verify personality. It provides access to control over various services such as google voice, e-commerce, window talk recognition, m-commerce, automation, home automation and safety management etc. This paper provides Review of various word recognition systems and speakers. Speech is a basic form of human interaction creatures, so it's a very user-friendly interface. Although the sector has gained a lot permission to modify apps and applications but there are a few parameters that affect accuracy as well the effectiveness of the speech recognition system. The most varied of speech involves the level of speech, of naturesituations, channel and context of speech. The robustness of the speech system is subject to certain / Speech signal features. To improve the power of the speech recognition system, it is necessary to design the speech they see in local languages. Multilingualism is a revolutionary new field in the field of speech recognition.

There are many development and research in the field of foreign languages but to improve its power and usefulness to indigenous peoples, it is important to use this technology in the indigenous languages. This paper offers various voice and speaker reviews recognition programs.

7. REFERENCES –

[1] VOICE RECOGNITION SYSTEM: SPEECH-TO-TEXT Prerana Das, Kakali Acharjee, Pranab Das and Vijay Prasad JAFS|ISSN 2395-5554 (Print)|ISSN 2395-5562 (Online)|Vol 1(2)|November 2015

[2] (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 3, 2009 Speech Recognition by Machine: A Review M.A.Anusuya and S.K.Katti

[3] Volume 5 Issue V, May 2017 IC Value: 45.98 ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET) ©IJRASET: All Rights are Reserved 262 Voice Recognition Technique: A Review Nisha

[4] Speech Recognition: A Complete Perspective Ashok Kumar, Vikas Mittal international

Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-6C, April 2019