Educational Administration: Theory and Practice
2024, 30(3), 2395 - 2399
ISSN: 2148-2403
https://kuey.net/

Research Article

Voice Recognition Techniques: A Review Paper

Aditya Kumar Jha[1*], Niraj Singhal[2], Arpit Chhabra[3]

[1*,2,3]Department of Computer Science and Engineering, Sir Chhotu Ram Institute of Engg & Technology, C.C.S. University, Meerut, U.P., India

***Corresponding author-**Aditya Kumar Jha
Email:-informatics.aditya@gmail.com

Citation:, Aditya Kumar Jha, et al ,(2024),.Voice Recognition Techniques: A Review Paper, Educational Administration: Theory and Practice, 30(3), 2395 - 2399,
Doi:. 10.53555/kuey.v30i3.5944

ARTICLEINFO

ABSTRACT

A certain voice can be recognised using voice recognition technology. The cornerstone for speaker identification is voice signals. Voice targeting is applicable to a wide range of applications, including voice mail, database access, phone banking, and phone purchasing. The ability to enter one's voice for verification is among the most potent uses of voice recognition for security. The fundamental means of interpersonal communication is speech. The technique of translating speech sounds into appropriate text is known as speech recognition. Over the past few years, speech recognition technology has advanced significantly. Nonetheless, there are other significant study obstacles, such as variations in speaker and language, ambient sound, word size, etc. This study aims to give a comprehensive overview of speech acceptance by summarising the several approaches utilised in the standard speech system and describing the numerous processes involved.

***Keywords*:** Speech recognition, modeling, speech processing, training and assessment.

1. INTRODUCTION

Voice is a fundamental, widely used, and efficient form of interpersonal communication. These days, voice technology can be used for a small but impressive number of purposes. Thanks to technology, machines can now react to human voices correctly and consistently and offer beneficial services. Since computer communication is faster than typing on a keyboard, people would favor such programs. This may be done by developing a voice recognition system for computers, which allows the computer to interpret voice instructions and make phone calls from text. Software for voice recognition: The "process-to-text" method converts an audio signal that was recorded using a microphone into a series of words. Voice signals are fundamentally altered by voice recognition, often known as the Automatic Speech Recognition (ASR) process.
Developing voice recognition strategies is the goal of the speech recognition platform. Scale and power constraints restricted the early computer programmes. However, the subject of default voice recognition is already seeing a shift in computer technology. Thanks to advancements in computer technology, voice recognition significant details are now easily retained. The languages in which speech recognition systems have evolved are quite limited. As a result, the construction of native language expressions involves numerous dimensions. Automatic expression recognition has decreased the need for human interaction in many domains, such as voice calling, data entry, and automated telephone processing on the telephone network. Natural language translators, queries based on repurposed travel and reservation data, etc. are a few examples of these fields. This essay focuses on the fundamental components of voice recognition systems, technological advancements, and issues with automatic speech recognition [2] [3].

1.1 AboutVoiceRecognitionProgram:

There are two main divisions inside the speech recognition system.

A. ToidentifytheSpeaker

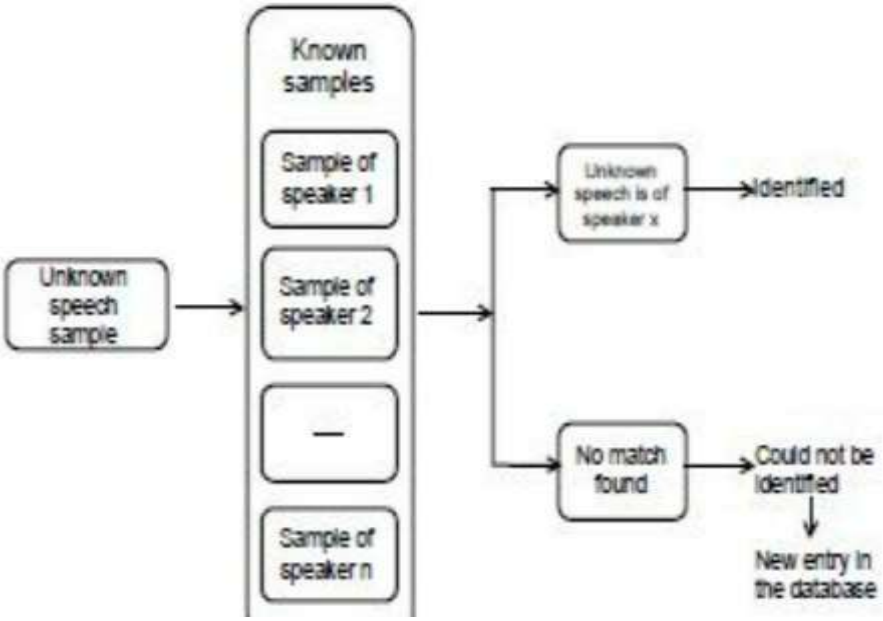
B. Speakerauthentication

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. Open-set mode and Open-set mode are known sets of sounds encoded in two edges. The speaker does not have to be a member of a group of well-

Copyright © 2024 by Author/s and Licensed by Kuey. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

known speakers to operate in open-set mode. This is applied in cases of particular criminal action from which it originates. Identifying several suspects in close-range mode, the speaker's voice is among the other recognizable voices previously included in the database. This technique is used to securely identify an authorized individual via biometric proofing.

On the other hand, the process of approving or disapproving a speaker's assertion is known as speaker authentication. It is employed to support someone's appeal for veracity. The process of verifying the authenticity of a voice from a set of speakers is commonly known as the "open set mode" in speaker verification.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. A microphone was used to record speech and convert it into an electrical signal. The purpose of the computer's sound card Transform a signal from analog to digital. This voice signal can be stored and played by the sound card. [3]



Thefollowingarethebuildingblocksforacommonspeechrecognitionsystem.

A. Signalpreprocessing
B. Facilityextraction
C. Languagemodel
D. decoder
E. Speechrecognition[4]

1.2 Typesof speech

Speechrecognitionisdividedbywhatkindofwordstheycanunderstand.Theycategorizedas:

1) RemoteName: RemoteIdentifiersometimesrequireseachspokenwordtobesilent(lackofsound) signal) in the bot

2) ConnectedName: Samewithadifferentname,butallowsdifferentexpressionstoworktogether" contains at least a break between them.

3) ContinuousSpeech:allowsuserstospeaknaturallyandsimilarlythecomputerwilldecide content.

4) SoundSpeech: Itisakindofspeechthatisloudand ineffective.Speechrecognition systems can be ofdifferenttypesdepending on thetypesofwordsto beobserved.These are the various types are distinguished as follows:

A. IsolatedWords: 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. Linkednames: Linkedwordsarethesameassinglewordswithaslightpausedifferenceamongthem.

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. Literature Review

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 todevelop automatic talk recognition was made in the 1950s whenmany researchers tried to explore the basic acoustic-phonetic concept.
- Davis et al. (1952) set out to create a motivation for digital recognition.The suggested strategy was implemented, and the digital vowel region is displayed using the spectacle concept.
- Olson and Belar (1956), attempted it find ten different characters for a single speaker composed of 10 monosyllabic words.
- Fry and Denestried (1959) attempted to construct a phonetic identifier for vowels and analytical concrete. the screen and the similarity. A hardware-based method surfaced in 1960 after numerous Japanese laboratories got involved in this area.
- SuzukiandNakata(1961),developedhardwarefortherecognitionofvowels.
- Sakoe and Chiba inJapanpropose the importanceof robust synchronizationsystems. 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- Qatabet. 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.
- In 2014, Z. Yu and colleagues introduced a voice recognition teaching exam that utilized the Hidden Markov Model (HMM). The test outlined the HMM speech recognition theory and the steps involved in starting a speech recognition program. [3]

3. RESEARCH GAP

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

- Searchforreportsordocumentsonyourcomputer.
- Creategraphsortablesusingdata.
- Statetheinformationyouwanttoincludeinthe document.
- Printdocumentsonrequest.
- Startvideoconferencing.
- Schedulemeetings.
- Recordminutes.
- Maketravelarrangements

5. METHODOLOGY

The effectiveness of the speech recognition system is usually measured in terms of accuracy and speed. Program accuracyis evaluated accordingtoWorderror rate (WER).Therefore, performanceis defined inWord Recognition 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 aspectcan be measured in terms of statistics,

WordErrorRate(WER)=I+S+D/N

Where'I' is the input number Sistheinputnumber,DisthesubtractionnumberandNisthe number words with words. Word Recognition Rate (WRR) = 1-WER of speed, Actual time limit specified; is calculatedby following the equation,

RealTimeFactor(RTF)=T/D

WhenT isatimeseries andDis a duration.[4]

6. CONCLUSION

This study examines their recent development and discusses the fundamentals. This research compares several approaches to developing a voice recognition system using a modified feature translation procedure and speech recognition language system.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. A review of many word recognition speakers and systems is given in this study. Speech is an extremely user-friendly interface since it is a fundamental method of human contact.

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 ofthespeech recognitionsystem.Themostvariedofspeechinvolvesthelevelofspeech,of nature situations, 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.

REFERENCES–

[1] VOICE RECOGNITION SYSTEM: SPEECH-TO-TEXT Prerana Das, KakaliAcharjee, 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.

[5] Wood, B., 2018. *Whitepaper: Voice & Visual Search – How to reach your customersintheevolvingsearchlandscape*. [ebook] Hallam, pp.6-16.Availableat:

[6] <https://www.hallaminternet.com/whitepaper-voice-visual-search-how-to-reach-your- customers-in-the-evolving-search-marketing-landscape/>[Accessed4June2021].

[7] Wood, B., 2020. *UX and SEO: Experience as the Fourth Pillar of SEO | Hallam*.[online] Hallam. Available at: <https://www.hallaminternet.com/ux-and-seo-experience-as-the-fourth-pillar-of-seo/>[Accessed7July2021].

[8] Dharamveer, Samsher, D. B. Singh, A. K. Singh, N. Kumar, Solar Distiller Unit Loaded withNanofluid- A Short Review. Lecture Notes in Mechanical Engineering, Springer, Singapore, (2019) 241-247, https://doi.org/10.1007/978-981-13-6577-5_24.

[9] S Kumar and D. Singh, Energy and exergy analysis of active solar stills using compound parabolic concentrator, International Research Journal of Engineering and Technology (IRJET), 6 (2019) 12.

[10] R. Shanker, D. Singh, D. B. Singh "Performance analysis of C.I. engine using biodiesel fuel by modifying injection timing and injection pressure" International Research Journal of Engineering and Technology (IRJET) 6 (2019) 12.

[11] A. K. Anup and D. Singh, FEA analysis of refrigerator compartment for optimizing thermal efficiency, International Journal of Mechanical and Production Engineering Research and Development, 10 (2020) 3, 3951-3972.

[12] S Kumar and D. Singh, Optimizing thermal behavior of compact heat exchanger, International Journal of Mechanical and Production Engineering Research and Development, 10 (2020) 3, 8113-8130.

[13] Dharamveer and Samsher, Comparative analyses energy matrices and enviro-economics for active and passive solar still, materialstoday: proceedings, https://doi.org/10.1016/j.matpr.2020.10.001 (2020).

[14] Dharamveer, Samsher, Anil Kumar, Analytical study of N^{th} identical photovoltaic thermal (PVT) compound parabolic concentrator (CPC) active double slope solar distiller with helical coiled heat exchanger using C_uO Nanoparticles, Desalination and water treatment, 233 (2021) 30-51, https://doi.org/10.5004/dwt.2021.27526

[15] Dharamveer,Samsher, Anil Kumar, Performance analysis of N-identical PVT-CPC collectors an active single slope solar distiller with a helically coiled heat exchanger using CuO nanoparticles, Water supply, October 2021, SCI-E Index, IWA Publication. I.F 1.275,https://doi.org/10.2166/ws.2021.348
[16] M. Kumar and D. Singh, Comparative analysis of single phase microchannel for heat flow Experimental and using CFD, International Journal of Research in Engineering and Science (IJRES), 10 (2022) 03, 44-58.
[17] Subrit and D. Singh, performance and thermal analysis of coal and waste cotton oil liquid obtained by pyrolysis fuel in diesel engine, International Journal of Research in Engineering and Science (IJRES), 10 (2022) 04, 23-31.