iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62450

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Human Voice Recognition for Authentication

Shreya M Tantry[1], Sankalp[2], Prasha S[3], Nisha Dsouza[4], Mr. Ganesh Shetty[5]
ECE Department, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Udupi, India

Abstract: Human voice plays a very important role as a vital biometric parameter for authentication and Identification. Voice recognition is a biometric technology used to identify one particular person's voice. It provides Enhanced security, convenient authentication and considerable cost saving. It can be performed using the concept of processing using Technique. Voice recognition is a software program or hardware device that can decode the human voice. It can receive and interpret dictation, understand and perform spoken commands, and identify a person based on their unique voiceprint. Voice recognition has become more popular in recent years with the rise of Artificial Intelligence (AI) and intelligent assistants, such as: Amazon's Alexa, Apple's Siri, Google Assistant. Voice recognition is used to control a smart home ,Instruct a smart speaker, Command phones and tablets ,Voice recognition is also known as Automatic speech recognition (ASR), Computer speech recognition.

I. INTRODUCTION

Voice recognition is a software program or hardware device that can decode the human voice. It can receive and interpret dictation, understand, and identify a person based on their unique voiceprint. Voice recognition has become more popular in recent years with the rise of Artificial Intelligence (AI) and intelligent assistants, such as: Amazon's Alexa, Apple's Siri, Google Assistant. Voice recognition is used to control a smart home ,Instruct a smart speaker, Command phones and tablets , Voice recognition is also known as Automatic speech recognition (ASR), Computer speech recognition.

One of the key components in voice recognition systems is the use of signal processing techniques, such as the Fast Fourier Transform (FFT). The FFT is a mathematical algorithm that transforms a time-domain signal into its frequency-domain representation, enabling the analysis of the signal's frequency components.

Voice recognition using FFT is a powerful approach because it allows the system to focus on the frequency characteristics of the speech signal, which are essential for distinguishing between different phonemes and words. This method forms the foundation for many modern voice-controlled systems and virtual assistants, contributing to the seamless interaction between humans and machines through spoken language

II. LITERATURE REVIEW

1] In this paper, Improvement of speech recognition is done in noisy device itself. The performance of speech recognition deteriorates in noisy device because it depends on the level of signal to noise. Normally, noise reduction techniques are applied in post processing with recorded voice signal with noise. But it is effective only when the level of voice is higher than those of noise. The condition is not satisfied in heavy noisy devices such as robot vacuum cleaner and so on. Under this background, a method of adaptive noise cancellation is proposed for improving speech recognition in noisy device itself. As a case study, a robot vacuum cleaner is investigated with respect to what the main noise source is and how the noise propagates. An adaptive noise cancellation is applied with the placement of reference sensors and development of adaptive algorithm.

2] In this research, They have proposed to review several voice algorithms in terms of detection accuracy and processing overhead and to identify the optimal voice recognition algorithm that can give the best trade-offs between processing cost (speed, power) and accuracy. Also, to implement and verify the chosen voice recognition algorithm using MATLAB. Ten words were spoken in an isolated way by male and female speakers (four speakers) using MATLAB as a simulation environment, these word were used as a reference signal to trained the algorithm, for evaluating phase, all algorithms dictates to subject them to similar test criteria. From the simulation results, the Wiener Filter algorithm outperform the other four algorithms in terms of all measure of performance, and power requirement with the moderate complexity of the algorithm and its prospective implementation as a hardware.

3] This paper identifies a novel speech signal representation method is presented. The proposed method is based on the fractional Fourier transform (FrFT), which is a generalization of the classical Fourier transform (FT). Even though we use FrFT in feature extraction for speech recognition, it can very well be used in other areas such as enhancement, verification, and synthesis, where parametric representation of speech is needed. Experimental results conducted on the Aurora 2 database show significant improvements over MFCC at high Signal to noise ratio(SNR) conditions.

©IJRASET: All Rights are Reserved

4] This paper describes a phoneme segmentation algorithm that uses Fast Fourier Transform (FFT) spectrogram. The algorithm has been implemented and tested for utterances of continuous Arabic speech of 10 male speakers that contain almost 2346 phonemes in total. The recognition system determines the phoneme boundaries and identifies them as pauses, vowels and consonants. The system uses intensity and phoneme duration for separating pauses from consonants. Intensity in particular is used to detect two specific consonants (/r/, /h/) when they are not detected through the spectrographic information. Segmentation accuracy of 95.39% for the overall system has been achieved.

5] This paper presents different approximate designs for computing the FFT. The tradeoff between accuracy and performance is achieved by adjusting the word length in each computational stage. Two algorithms for word length modification under a specific error margin are proposed. The first algorithm targets an approximate FFT for an area-limited design compared to the conventional fixed design; the second algorithm targets performance so it achieves a higher operating frequency. Both of the proposed algorithms show that an efficient balance between hardware utilization and performance is possible at stage-level. The proposed approximate FFT designs are implemented on FPGA; experimental results show that hardware utilization using the first approximate algorithm are reduced by at least nearly 40%. The second algorithm increases performance of the designs by over 20%. Fine granularity design is also investigated, where the FPGA resources for a 256-point FFT computation can be further reduced by nearly 10% compared to a coarse design. Finally, the proposed approximate designs are applied to a feature extraction module in an isolated word recognition system.

6] Speech recognition is one of the current topics of discussion in the twenty-first century. The advent of technological gadgets in modern society has become rampant through vigorous efforts made by scientists in realizing their aim of developing an algorithm that will allow machines to interact with human beings. It is obvious that this situation previously thought of as fiction, has been achieved. Hence, there are several applications that enhance speech recognition. This paper aims to study some models of the speech recognition system, its classification of speech, its significance, and its application.

7] This paper presents different approximate designs for computing the FFT. The tradeoff between accuracy and performance is achieved by adjusting the word length in each computational stage. Two algorithms for word length modification under a specific error margin are proposed. The first algorithm targets an approximate FFT for an area-limited design compared to the conventional fixed design; the second algorithm targets performance so it achieves a higher operating frequency. Both of the proposed algorithms show that an efficient balance between hardware utilization and performance is possible at stage-level. The proposed approximate FFT designs are implemented on FPGA; experimental results show that hardware utilization using the first approximate algorithm are reduced by at least nearly 40%. The second algorithm increases performance of the designs by over 20%. Fine granularity design is also investigated, where the FPGA resources for a 256-point FFT computation can be further reduced by nearly 10% compared to a coarse design. Finally, the proposed approximate designs are applied to a feature extraction module in an isolated word recognition system.

8] Voice recognition consists of detecting a user' s identity based on characteristics of their voice. It is a widely applied form of biometric recognition in the world, particularly in fields where security has a high priority. The deep neural networks were used as feature extractor alongside classifiers, but they haven't been completely trained due to the success of deep learning. While such methods are extremely efficient, they still require manual attention. Especially in DNN, interactivity between people and machines is essential. This is where the art of voice recognition comes from. In addition to their application in speech recognition, deep neural networks have demonstrated their potential to be used for voice recognition as well. They provide an efficient implementation of complex nonlinear models for learning unique and invariant data structures. The main contribution of this work is to provide a brief overview of the field of deep neural networks and voice recognition, describing its system, underlying approaches, and challenges.

9] In this paper, they developed a robotic system which take voice commands from user in real time & can perform accordingly. Each word vibrates at different frequencies. So, they performed digital signal processing tool such as FFT (Fast Fourier Transform) on a input command to convert the signal from time domain into frequency domain. Then they extracted the features of input commands from its frequency domain to classify the input commands into a given set of commands by using PNN (probability neural Network) classifier. Once the input command is identified, they communicated with the robotic system serially by using GSM module which is controlled by using Arduino.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

III. PROPOSED METHODOLOGY

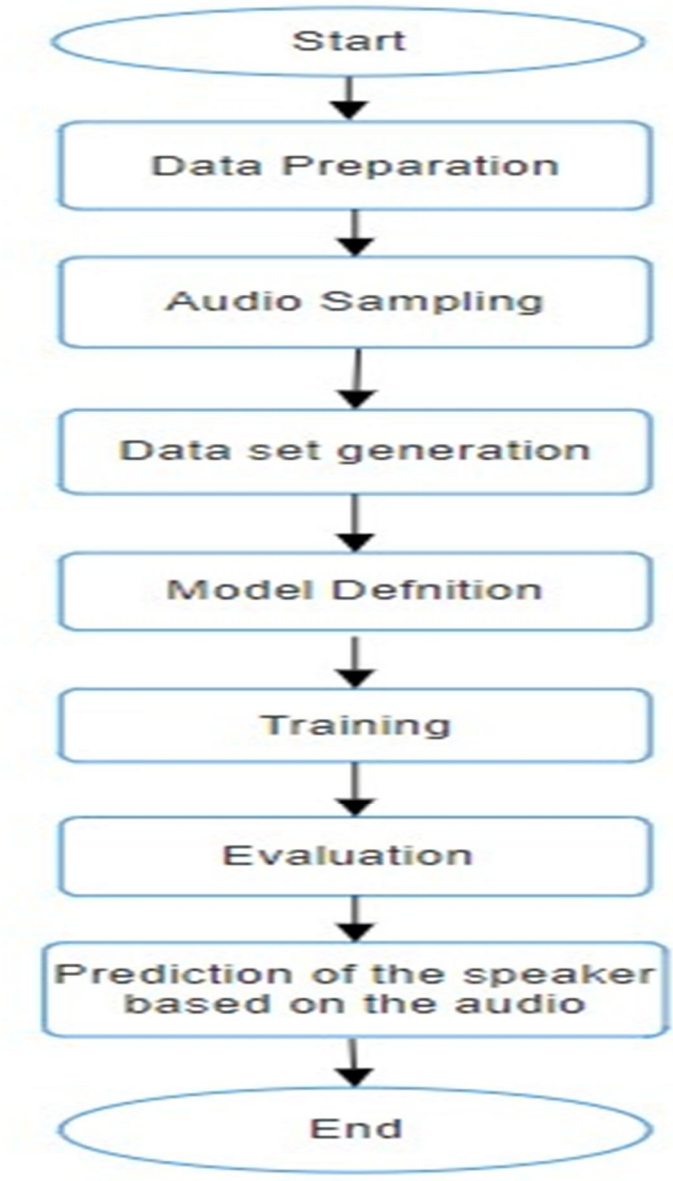


Fig 3.1: Flow diagram of the proposed system

IV. WORKING

1) The multiple samples of voices from different individuals are fetched, storing them in a dictionary where the keys are the different voice samples of the individual and the values are lists of audio recordings.

2) It then trains speaker models using the recorded voice samples. For each individual, it extracts MFCC features from the audio data, applies PCA to reduce the dimensionality of the feature space, and trains a KNN classifier to recognize the speaker.

3) After training, the program listens for speech input from the microphone. Once it receives audio input, it extracts MFCC features, applies PCA transformation, and classifies the speaker using the trained models.

4) Finally, if the speaker is recognized, it greets the identified person using text-to-speech synthesis. Otherwise, it informs the user that the speaker couldn't be recognized.

Overall, the program demonstrates a basic implementation of speaker recognition using MFCC features and machine learning classification techniques

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. FUTURE WORK

1) Explore advancements in signal processing techniques, including advanced methods beyond FFT, to further enhance the extraction of meaningful features from voice signals.
2) Improve the device's ability to adapt to various environmental conditions, such as noise, echoes, and background sounds, to maintain accurate recognition in different contexts.

VI. EXPECTED OUTCOME

1) It recognizes the human voice with the database given.
2) Displays the name of the identified person through a speaker.

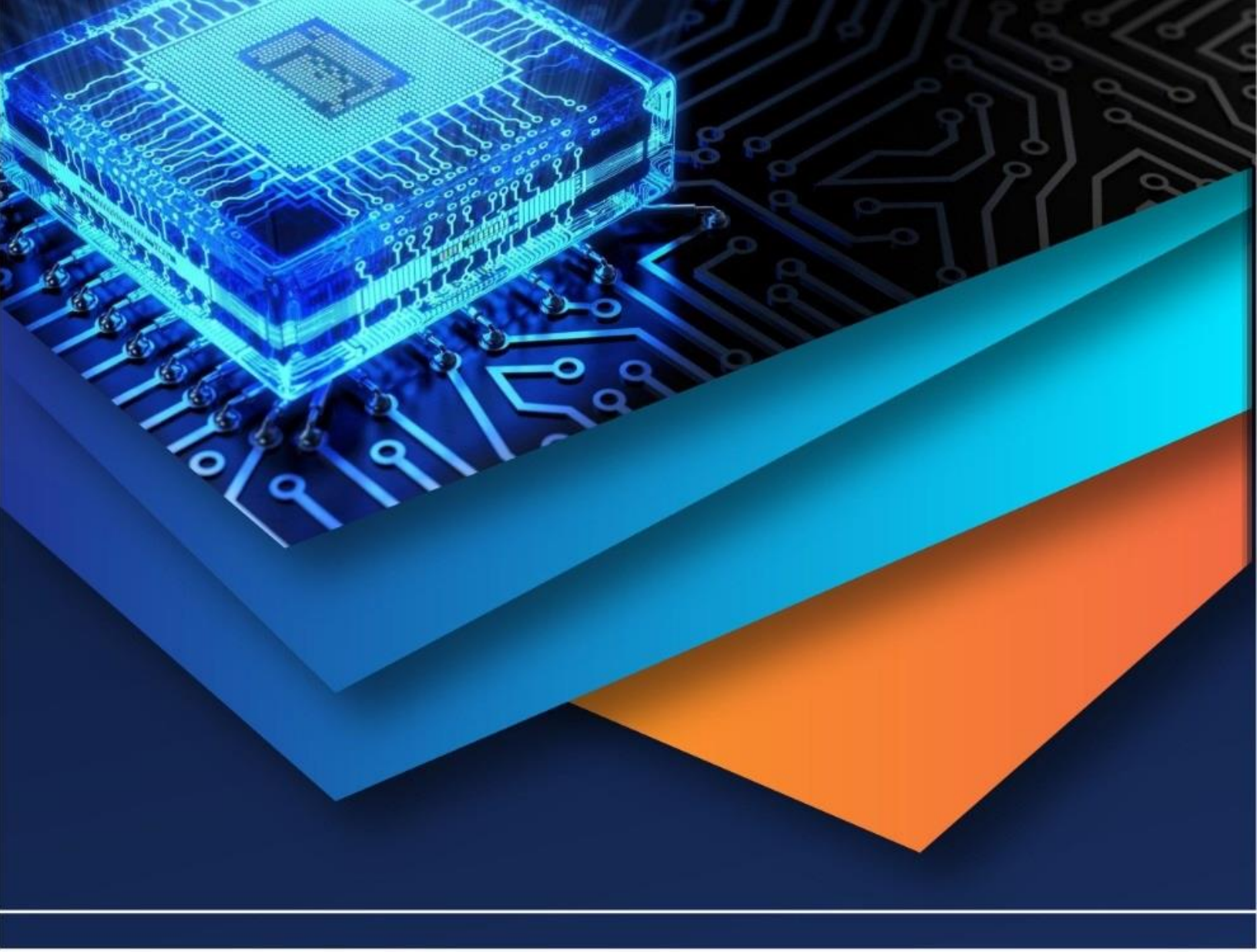
VII. CONCLUSION

Human voice recognition it offers significant benefits in security and user experience, challenges such as speaker variability and privacy concerns remain. The technology's future looks promising, but responsible development and ethical considerations are essential for its continued success.

REFERENCES

[1] Atheer Tahseen Hussein Department of Electrical, Electronic and Systems Engineering University Kebangsaan Malaysia Malaysia, 43600 Bangi, Selangor, Malaysia International journal of Engineering research and technology, August-2015
[2] Sairaj L Burewar Department of Electronics & Telecommunication Engineering Shri Guru Gobind Singhji Institute of Engineering & Technology Nanded – (431606), India sairajburewar@gmail.com 2018
[3] Adaptive noise cancellation for improving speech recognition 2019 Lee, Nokhaeng1 Samsung Electronics Co., Ltd. 129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Korea
[4] S. P. Gupta, S. V and S. G. Koolagudi, "Noise Cancellation by Fast Fourier Transform for Wav2Vec2.0 based Speech-to-Text System," 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp. 1-4, doi: 10.1109/I2CT57861.2023.10126221.
[5] E. H. Kim, K. H. Hyun, S. H. Kim and Y. K. Kwak, "Speech Emotion Recognition Using Eigen-FFT in Clean and Noisy Environments," RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju, Korea (South), 2007, pp. 689-694, doi: 10.1109/ROMAN.2007.4415174 2017, 147, 88–102
[6] D. Zhang et al., "Fast Fourier Transform (FFT) Using Flash Arrays for Noise Signal Processing," in IEEE Electron Device Letters, vol. 43, no. 8, pp. 1207-1210, Aug. 2022, doi: 10.1109/LED.2022.3183111.
[7] M. M. Rahman, A. K. M. F. Haque, M. Hasan, N. Sultana and M. Z. Islam, "Designing and development of voice to machine interfacing technique," 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Savar, Bangladesh, 2015, pp. 1-4, doi: 10.1109/ICEEICT.2015.7307456.
[8] M. M. Awais, Waqasahmad, S. Masud and S. Shamail, "Continuous Arabic Speech Segmentation using FFT Spectrogram," 2006 Innovations in Information Technology, Dubai, United Arab Emirates, 2006, pp. 1-6, doi: 10.1109/INNOVATIONS.2006.301939.
[9] W. Liu, Q. Liao, F. Qiao, W. Xia, C. Wang and F. Lombardi, "Approximate Designs for Fast Fourier Transform (FFT) With Application to Speech Recognition," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 12, pp. 4727-4739, Dec. 2019, doi: 10.1109/TCSI.2019.2933321

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)