qrAs per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology
Volume 10, Issue 03, pages 143-151.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol10issue03.html
DOI: https://doi.org/10.46501/IJMTST1003022

Performance analysis of techniques in Automatic Speech Recognition System

Shaik Mastan Basha, Repalle Likhithasri, Shaik Nagoor Basha, Koduri Srinivasula Reddy, Bhashapogu Prem Kumar

Department of Electronics and Communication Engineering, Chalapathi Institute of Technology, Guntur, India.

To Cite this Article
Shaik Mastan Basha, Repalle Likhithasri, Shaik Nagoor Basha, Koduri Srinivasula Reddy, Bhashapogu Prem Kumar, Performance analysis of techniques in Automatic Speech Recognition System, International Journal for Modern Trends in Science and Technology, 2024, 10(03), pages. 143-151.https://doi.org/10.46501/IJMTST1003022

Article Info
Received: 30 January 2024; Accepted: 22 February 2024; Published: 01 March 2024.

ABSTRACT

A new improved method for speech/speaker recognition is presented in this paper using combination of discrete wavelet transform (DWT) and Relative Spectra Perceptual Linear Prediction (RASTA-PLP) for feature extraction. A graphical processing unit (GPU) is used for increasing the speed of computations. Wavelet transform is applied to the speech signal. Then RASTA-PLP coefficients of the wavelet channels are calculated. Two new sets of features are generated by concatenating the two extracted features and taking the average of the concatenated vector. The obtained feature vector is then fed to NN for classification where parallel computation for NN is introduced using GPU. The objective of the proposed method is to enhance the performance by introducing more features from the signal and applying parallel computations technique leading to an improvement in both the recognition rate and the computational speed whether for a clean or a noisy speech signal for the two proposed methods in comparison to using DWT and Mel-Frequency Cepstral Coefficients (MFCCs).

Keywords :Discrete Wavelet Transform (DWT), RASTA-PLP, MFCCS

1. INTRODUCTION

Speech has been one of the most potent tools at man's disposal since ancient times. Humans have constantly evolved and expressed themselves through speech. There is plethora of languages used and spoken by man throughout the entire world. [3] Computer speech recognition or Automatic Speech recognition system is a process in which the words spoken by humans or speech signals are translated into words. The words that are recognized by the machine can be the final output or these words can act as input to natural NLP i.e. natural language processing, it also uses algorithm, which is implemented as computer program. [15] The major goal of Automatic Speech recognition is to characterize, extract and recognize these words as of some language. Speech technology focuses on development and improvisation of techniques that enables the computers to identify speech as an input effectively. Speech recognition technology has enormously evolved over the past few decades, we can now find it in phones, like features such as Siri, plethora of phone applications that are automated used by offices and airlines, and even applications that can be easily used at homes for instance Dragon Dictate. The speech signals, which are spoken, are actually quasi-stationary signals; hence we need to extract the features from speech on basis of short-term amplitude spectrum. Feature extraction is the most integral phase of speech recognition. It is mainly used to reduce the noise and distortions in speech so that it can be effectively converted into text. [16].

2. LITERATURE REVIEW

Identification of fruits and vegetables are implemented in different areas. The most common areas are identification in the retail business and in areas where the purpose is to ease the harvest in the perspective of agriculture. Mostly, the identification is done manually by a cashier or via the self-service systems in a store. In this section, different methods of identifying fruits and vegetables will be presented.A company which has made great progress in its technical evolution when it comes to artificial intelligence, image recognition and automating physical work is Amazon. Amazon developed a product, called Amazon Go2, which enabled a shopping experience without cashiers or self-service checkouts. The company built the store where the customers check in with a smart phone using the application Amazon Pay3 . The store is set up with a large number of cameras and sensors. Thanks to computer vision and deep learning algorithms, Amazon managed to create a store where technology identifies the products the customers choose. No checkout is required, the chosen products are debited from Amazon Pay account that the costumer checked in with.StrongPoint is a company, with its headquarters in Norway, offering technical solutions to the retail business. Strong Point recently released an identification system called Digi5 . Digi consists of a user interface displayed on a touch screen, a scale, a camera and a label printer. The software is implemented with image recognition in the identification process and can be compared to the existing counterpart of this project. Digi is new to the market; hence it is not used in many stores. It is an economic issue of the retail business whether the business will change the existing systems or not.Related work including image recognition has been done in the purpose of controlling the vegetation and harvest of fruitage and other growths at fields of farmers. The technology has been used to automate the yield with the help of robotic harvesting However, the issue of creating a fast and reliable fruit detection system persists. This is due to large variation in the appearance of the fruits in field, including color, shape, size and texture properties.

3.PROPOSED METHODOLOGY

Automatic speech recognition system enables a machine to hear, comprehend and respond to the speech signal, which is provided by the speaker as input. The methodology of the speech recognition system can be understood by the following flowchart i.e. Figure 1. The input is given to the system; this input is in the form of speech signal that is then divided into smaller components known as frames, also known as the analysis part. Then the feature extraction is performed and useful features are extracted out of the voice signal. [19] After feature extraction the feature vectors are divided into two categories, training and testing. [27] In training speech modeling is performed and in testing pattern matching is done. During recognition phase the test speech data is matched with the training models and the result is given according to the best match. Thus, the speech recognition system can be classified mainly into four phases.Modeling method is the process of generating speaker models using feature-oriented feature vectors. The two major classifications of speaker modeling techniques are speaker identification and speaker recognition, which in turn is divided further into speaker independent and speaker dependent techniques. In the case of feature identification, speaker of the input signal is simply identified based on its specific features.

3.1 RASTA (Relative Spectral Algorithm)

RASTA or Relative Spectral Algorithm as it is known is a technique that is developed as the initial stage for voice recognition. This method works by applying a band-pass filter to the energy in each frequency sub-band in order to smooth over short-term noise variations and to remove any constant offset.In voice signals, stationary noises are often detected. Stationary noises are noises that are present for the full period of a certain signal and does not have diminishing feature. Their property does not change over time. The assumption that needs to be made is that the noise varies slowly with respect to speech. This makes the RASTA a perfect tool to be included in the initial stages of voice signal filtering to remove stationary noises. The stationary noises that are identified are noises in the frequency range of 1Hz - 100Hz

3.2 Formant Estimation

Formant is one of the major components of speech. The frequencies at which the resonant peaks occur are called the formant frequencies or simply formants . The formant of the signal can be obtained by analyzing the vocal tract frequency response

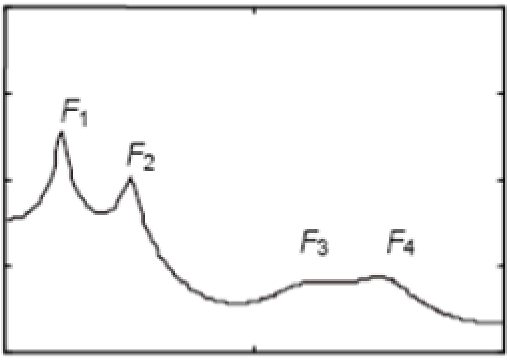

Figure 1: Vocal tract frequency response

Figure 2 shows the vocal tract frequency response. The x-axis represents the frequency scale and the yaxis represents the magnitude of the signal. As it can be seen, the formants of the signals are classified as F1, F2, F3 and F4. Typically, a voice signal will contain three to five formants. But in most voice signals, up to four formants can be detected. In order to obtain the formant of the voice signals, the LPC (Linear Predictive Coding) method is used. The LPC (Linear Predictive Coding) method is derived from the word linear prediction. Linear prediction as the term implies is a type of mathematical operation. This mathematical function which is used in discrete time signal estimates the future values based upon a linear function of previous samples.

$$\hat{x}(n) = -\sum_{i=1}^{p} a_i x(n-i)$$

where $\hat{x}(n)$ is the predicted or estimated value and $x(n-i)$ is the pervious value. By expanding this equation

$$\hat{x}(n) = -a(1)x(n-1) - a(2)x(n-2) - a(3)x(n-3) ,$$

The LPC will analyze the signal by estimating or predicting the formants. Then, the formants effects are removed from the speech signal. The intensity and frequency of the remaining buzz is estimated. So by removing the formants from the voices signal will enable us to eliminate the resonance effect. This process is called inverse filtering. The remaining signal after the formant has been removed is called the residue. In order to estimate the formants, coefficients of the LPC are needed. The coefficients are estimated by taking the mean square error between the predicted signal and the original signal. By minimizing the error, the coefficients are detected with a higher accuracy and the formants of the voice signal are obtained.

3.3 The Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) involves choosing scales and positions based on powers of two . so called dyadic scales and positions. The mother wavelet is rescaled or dilatedby powers of two and translated by integers. Specifically, a function f(t) ∈ L2(**R**) (defines space of square integrable functions) can be represented as

$$f(t) = \sum_{j=1}^{L} \sum_{k=-\infty}^{\infty} d(j,k)\psi(2^{-j}t-k) + \sum_{k=-\infty}^{\infty} a(L,k)\phi(2^{-L}t-k) \qquad (2.4)$$

The function ψ(t) is known as the mother wavelet, while φ(t) is known as the scalingFunction. The set of functions

$$\{\sqrt{2^{-L}}\phi(2^{-L}t-k), \sqrt{2^{-j}}\psi(2^{-j}t-k) \mid j \le L, j,k,L \in Z\},$$

Where Z is the set of integers is an orthonormal basis for L2(**R**).The numbers a(L, k) are known as the approximation coefficients at scale L,while d(j,k) are known as the detail coefficients at scale j.The approximation and detail coefficients can be expressed as:

$$a(L,k) = \frac{1}{\sqrt{2^L}} \int_{-\infty}^{\infty} f(t)\phi(2^{-L}t-k)dt \qquad (2.5)$$

$$d(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t-k)dt \qquad (2.6)$$

To provide some understanding of the above coefficients consider a projection fl(t) of the function f(t) that provides the best approximation (in the sense of minimum error energy) to f(t) at a scale l. This projection can be constructed from the coefficients a(L, k), using the equation

$$f_l(t) = \sum_{k=-\infty}^{\infty} a(l,k)\phi(2^{-l}t-k).$$

As the scale l decreases, the approximation becomes finer, converging to f(t) as $l \rightarrow 0$. The difference between the approximation at scale l + 1 and that at l, fl+1(t) - fl(t), is completely described by the coefficients d(j, k) using the equation

$$f_{l+1}(t) - f_l(t) = \sum_{k=-\infty}^{\infty} d(l,k)\psi(2^{-l}t-k).$$

Using these relations, given a(L, k) and {d(j, k) | j ≤ L}, it is clear that we can build the approximation at any scale. Hence, the wavelet transform breaks the signal up into a coarse approximation fL(t) (given a(L, k)) and a number of layers of detail {fj+1(t)-fj(t)| j < L} (given by {d(j, k) | j ≤ L}). As each layer of detail is added, the approximation at the next finer scale is achieved.

3.3.1 Vanishing Moments

The number of vanishing moments of a wavelet indicates the smoothness of the wavelet function as well as the flatness of the frequency response of the wavelet filters (filters used to compute the DWT).Typically a wavelet with p vanishing moments satisfies the following equation .

$$\int_{-\infty}^{\infty} t^m \psi(t)dt = 0 \qquad \text{for } m = 0, \ldots, p\text{-}1,$$

or equivalently,

$$\sum_k (-1)^k k^m c(k) = 0 \quad \text{for } m = 0, \ldots, p\text{-}1.$$

For the representation of smooth signals, a higher number of vanishing moments leads to a faster decay rate of wavelet coefficients. Thus, wavelets with a high number of vanishing moments lead to a more compact signal representation and are hence useful in coding applications.However, in general, the length of the filters increases with the number of vanishing moments and the complexity of computing the DWT coefficients increases with the size of the wavelet filters.

3.4 The Fast Wavelet Transform Algorithm

The Discrete Wavelet Transform (DWT) coefficients can be computed by using Mallat's Fast Wavelet Transform algorithm. This algorithm is sometimes referred to as the two-channel sub-band coder and involves filtering the input signal based on the wavelet function used.

3.4.1 Implementation Using Filters

To explain the implementation of the Fast Wavelet Transform algorithm consider the following equations:

$$\phi(t) = \sum_k c(k)\phi(2t-k) \tag{2.7}$$

$$\psi(t) = \sum_k (-1)^k c(1-k)\phi(2t-k) \tag{2.8}$$

$$\sum_k c_k c_{k-2m} = 2\delta_{0,m} \tag{2.9}$$

The first equation is known as the twin-scale relation (or the dilation equation) and defines the scaling function φ. The next equation expresses the wavelet ψ in terms of the scaling function φ. The third equation is the condition required for the wavelet to beOrthogonal to the scaling function and its translates. The coefficients c(k) or {c0, .., c2N-1} in the above equations represent the impulseresponse coefficients for a low pass filter of length 2N, with a sum of 1 and a norm of ½. The high pass filter is obtained from the low pass filter using the relationship

g ( )k c(k) k = −1 1− , where k varies over the range (1 . (2N . 1)) to 1.Equation 2.7 shows that the scaling function is essentially a low pass filter and is used to define the approximations.The wavelet function defined by equation 2.8 is a high pass filter and defines the details. Starting with a discrete input signal vector s, the first stage of the FWT algorithm decomposes the signal into two sets of coefficients. These are the approximation coefficients cA1 (low frequency information) and the detail coefficients cD1 (high frequency information), as shown in the figure below.

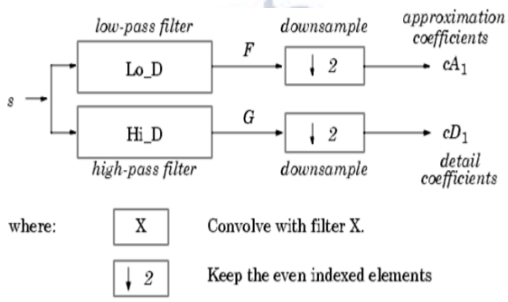

Figure 2: Filtering operation of the DWT

The coefficient vectors are obtained by convolving s with the low-pass filter Lo_D for Approximation and with the high-pass filter Hid for details. This filtering operation is Then followed by dyadic decimation or down sampling by a factor of 2. Mathematically the two-channel filtering of the discrete signal s is represented by the expressions:

$$cA_1 = \sum_k c_k s_{2i-k}\,, \qquad cD_1 = \sum_k g_k s_{2i-k} \qquad (2.10)$$

These equations implement a convolution plus down sampling by a factor 2 and give the forward fast wavelet transform.If the length of each filter is equal to 2N and the length of the original signal s is equal to n, then the corresponding lengths of the coefficients cA1 and cD1 are given by the formula:

$$\text{floor}\left(\frac{n-1}{2}\right) + N \qquad (2.11)$$

This shows that the total length of the wavelet coefficients is always slightly greater than the length of the original signal due to the filtering process used.

3.4.2 Multilevel Decomposition

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolutionComponents. This is called the wavelet decomposition tree.

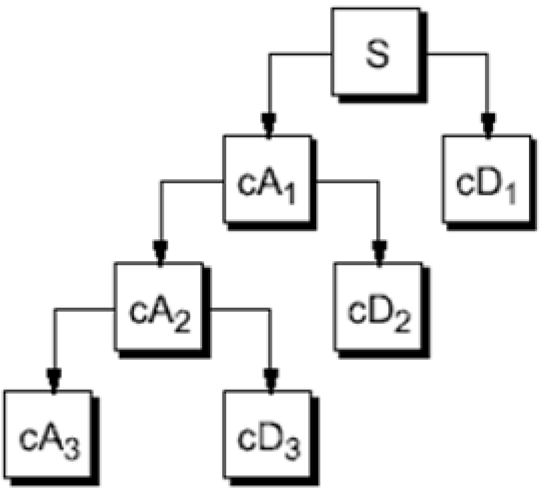

Figure 3: Decomposition of DWT co-efficient

The wavelet decomposition of the signal s analysed at level j has the following structure[cAj, cDj, ..., cD1]. Looking at a signals wavelet decomposition tree can reveal valuable information.The diagram below shows the wavelet decomposition to level 3 of a sample signal S

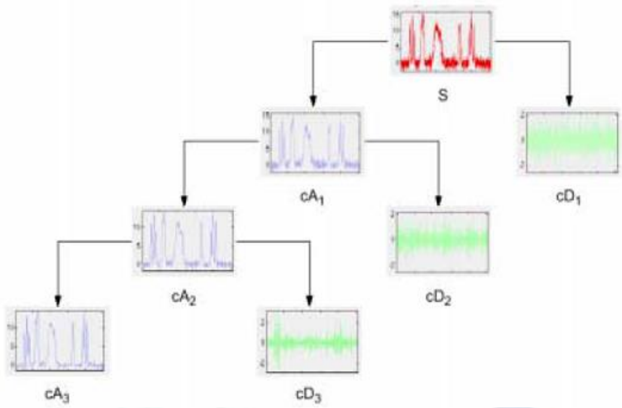

Figure 4: Decomposition of sample signal

Since the analysis process is iterative, in theory it can be continued indefinitely. In reality, the decomposition can only proceed until the vector consists of a single sample. Normally, however there is little or no advantage gained in decomposing a signal beyond a certain level. The selection of the optimal decomposition level in the hierarchy depends on the nature of the signal being analyses or some other suitable criterion, such as the low-pass filter cut-off.

3.5 Signal Reconstruction

The original signal can be reconstructed or synthesized using the inverse discrete wavelet transform (IDWT). The synthesis starts with the approximation and detail coefficients cAj and cDj, and then reconstructs cAj-1 by up sampling and filtering with the reconstruction filters.

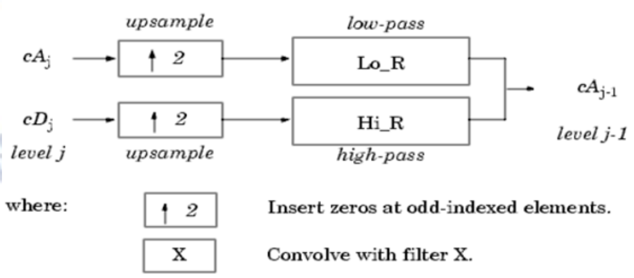

Figure 5: Wavelet reconstruction

The reconstruction filters are designed in such a way to cancel out the effects of aliasing introduced in the wavelet decomposition phase. The reconstruction filters

(Lo_Rand Hi_R) together with the low and high pass decomposition filters, forms a system known as quadrature mirror filters (QMF). For a multilevel analysis, the reconstruction process can itself be iterated producing successive approximations at finer resolutions and finally synthesizing the original signal.

3.6 Wavelet Energy

Whenever a signal is being decomposed using the wavelet decomposition method, there is a certain amount or percentage of energy being retained by both the approximation and the detail. This energy can be obtained from the wavelet bookkeeping vector and the wavelet decomposition vector. The energy calculated is a ratio as it compares the original signal and the decomposed signal.This is determined through numerous trial and errors. The coefficients that are extracted from the wavelet decomposition process is the second level coefficients as the level two coefficients contain most of the correlated data of the voice signal.The data at higher levels contains very little amount of data deeming it unusable for the recognition phase. Hence for initial system implementation, the level two coefficients are used.

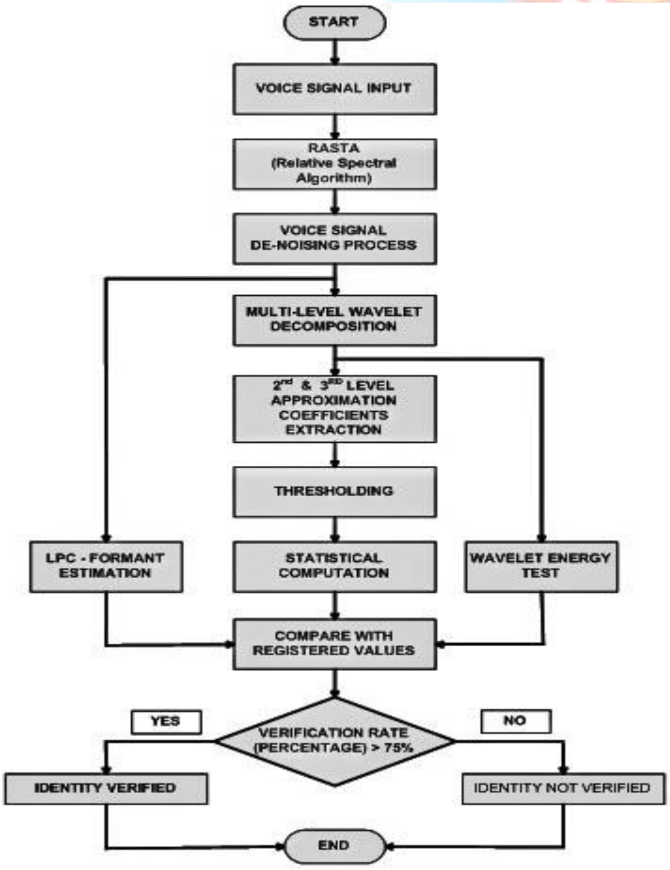

Figure 6:Complete System Flowchart

The coefficients are further threshold to remove the low correlation values, and using this coefficients statistical computation is carried out. The statistical computation of the coefficients is used in comparison of voice signal together with the formant estimation and the wavelet energy. All the extracted information acts like a 'fingerprint' for the voice signals. The percentage of verification is calculated by comparing the current values signal values against the registered voice signal values. The percentage of verification is given by: Verification % = (Test value / Registered value) x100. Between the tested and registered value, whichever value is higher is taken as the denominator and the lower value is taken as the numerator. Figure 3 shows the complete flowchart which includes all the important system components that are used in the voice verification program

4. RESULTS& DISCUSSION

GUI (Graphical User Interface)

Figure 7 shows the GUI (Graphical User Interface) implementation for the Voice Registration Section. This GUI enables the user to register an individual's voice signal using the pre-loaded voice signals that are saved in the program. The STEP 1 panel shows the pre-loaded voice signals contained in the program. The voice signals are obtained from a clean and noise-free environment. The user can select the available voice signal from the popup menu. The Show Voice Plot button enables the user to view the voice plot in the graph shown in the panel. The STEP 2 panel contains the function that enables the user to de-noise the signal. By pressing the De- Noise button, the user will be able to de-noise the signal and view the de-noised signal in the graph shown in the plot. The Extract Coefficients button enables the user to view the DWT (Discrete Wavelet Transform) Coefficients Detail and Approximation plot. The STEP 3 panel shows the extracted coefficients plot which is the 2nd level approximation & detail and the 3rd level approximation & detail. The STEP 4 panel shows the recognition methodology of the program. The Compute Values button performs the statistical computation on the 2nd level approximation and the 3rd level approximation and displays thevalues. At the same time the formant values and the wavelet energy of the voice signal is calculated and shown.

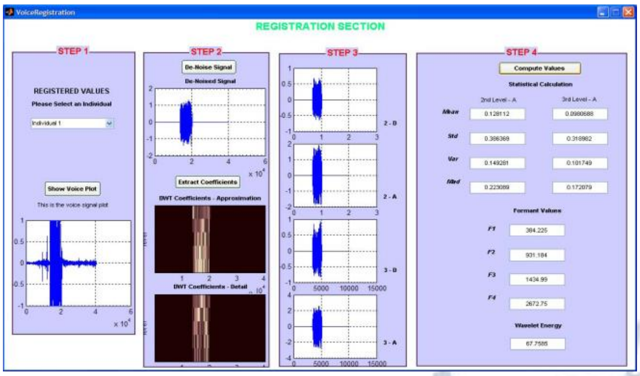

Figure 7: Voice Registration GUI

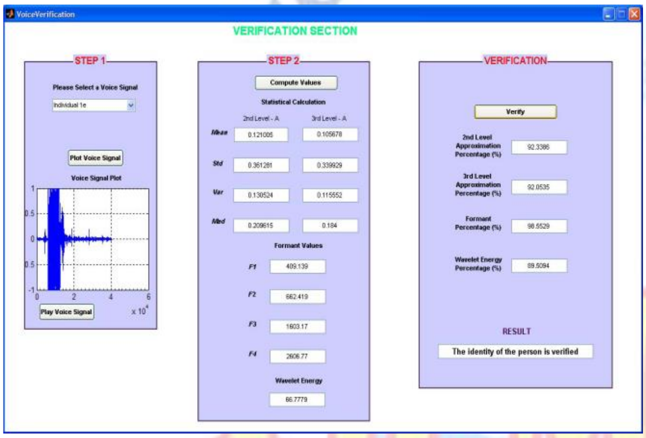

Figure 8: Voice Verification GUI

Figure 8 shows the GUI (Graphical User Interface) implementation for the Voice Verification Section. This GUI enables the user to verify an individual's voice signal using the pre-loaded voice signals that are saved in the program. The STEP 1 panel shows the pre-loaded voice signals contained in the program. These voice signals are recorded from different individuals saying their own name. The user can select the available voice signal from the pop-up menu. The Plot Voice Signal button enables the user to view the voice plot in the graph shown in the panel. The STEP 2 panel shows the recognition methodology of the program. The Compute Values button performs the statistical computation of the 2nd level approximation and the 3rd level approximation and displays these values. At the same time the formant values and the wavelet energy of the voice signal is calculated and shown.

Table 1: comparison test 1

Individual	1	2	3	4	5
1	Verified	Not Verified	Not Verified	Not Verified	Not Verified
2	Not Verified	Verified	Not Verified	Not Verified	Not Verified
3	Not Verified	Not Verified	Verified	Not Verified	Not Verified
4	Not Verified	Not Verified	Not Verified	Not Verified	Not Verified
5	Not Verified	Not Verified	Not Verified	Not Verified	Verified

Table 2: comparison test 2

Individual	1	2	3	4	5
1	Verified	Not Verified	Not Verified	Not Verified	Not Verified
2	Not Verified	Not Verified	Not Verified	Not Verified	Not Verified
3	Not Verified	Not Verified	Verified	Not Verified	Not Verified
4	Not Verified	Not Verified	Not Verified	Verified	Not Verified
5	Not Verified	Not Verified	Not Verified	Not Verified	Verified

Table 3: comparison test 3

Individual	1	2	3	4	5
1	Verified	Not Verified	Not Verified	Not Verified	Not Verified
2	Not Verified	Verified	Not Verified	Not Verified	Not Verified
3	Not Verified	Not Verified	Verified	Not Verified	Not Verified
4	Not Verified	Not Verified	Not Verified	Verified	Not Verified
5	Not Verified	Not Verified	Not Verified	Not Verified	Verified

Table 4: comparison test 4

Individual	1	2	3	4	5
1	Verified	Not Verified	Not Verified	Not Verified	Not Verified
2	Not Verified	Verified	Not Verified	Not Verified	Not Verified
3	Not Verified	Not Verified	Not Verified	Not Verified	Not Verified
4	Not Verified	Not Verified	Not Verified	Verified	Not Verified
5	Not Verified	Not Verified	Not Verified	Not Verified	Verified

From the Tables above of the verification result shows from the five random tests carried out, at any one given time, the program can successfully verify 4 out 5 persons accurately. The complete systems which constitutes all the system components for the recognition methodology is one of the main reasons for the high accuracy of the system. Currently, the percentage of verification is set at an average value of 78.75%. The verification rate can be further increased or decreased by adjusting the percentage of verification to a higher or lower value. By substituting a lower value, the system will be less secure while a higher value could jeopardize the accessibility rate of the system because of the certain level of tolerance is required for the voice signal as it tends to change with internal and external factors.The VERIFICATION panel shows the verification process of the system. The overall percentage values of the statistical computation, formant values and the wavelet energy are displayed

4.CONCLUSIONS

The Voice Recognition is carried out by estimating the formant and detecting the pitch of the voice signal by using LPC (Linear Predictive Coding). The voice recognition system that is developed is word dependant voice verification system used to verify the identity of an individual based on their own voice signal using the statistical computation, formant estimation and wavelet energy. By using the fifty preloaded voice signals from five individuals, the verification tests have been carried and an accuracy rate of approximately 80 % has been achieved.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Suitable Feature Extraction and Speech Recognition Technique for Isolated Tamil Spoken Words Vimala.C, Radha. DepartmentofComputerScience, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

[2] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi, "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", Journal of Computing, Vol. 2, ISSUE 3, pp. 138-143,MARCH 2010.

[3] Hajer Rahali, ZiedHajaiej, and Noureddine Ellouze, "Robust Features for Impulsive Noisy Speech Recognition Using Relative Spectral Analysis", International Journal of Computer, Information, Systems and Control Engineering Vol.8 No.9, pp. 1421-1426, 2014.

[4] Ingyin Khaing, "Myanmar Continuous Speech Recognition System Based on DTW and HMM", International Journal of Innovations in Engineering and Technology (IJIET), Vol. 2 Issue 1, pp.78-83, February 2013.

[5] Vibha Tiwari, "MFCC and Its Applications in Speaker Recognition", International Journal on Emerging Technologies, Vol.1 Issue 1, pp. 19- 22, 2010.

[6] Young-Giu Jung, Mun-Sung Han, and Sang Jo Lee, "Development of an Optimized Feature Extraction Algorithm for Throat Signal Analysis", Electonics and Telecommunication research Institute Journal, Volume 29, Number 3, pp.292-299, June 2007.

[7] Bhupinder singh, Rupinder Kaur, Nidhi Devgun, and Ramandeep Kaur, "The process of feature extraction in automatic speech recognition system for computer Machine interaction with humans: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, Issue 2, Feb 2012.

[8] Ram Singh, and Preeti Rao, "Spectral Subtraction Speech Enhancement with RASTA Filtering", Proceding of National Conference on Communications (NCC), Kanpur, India, 2007.

[9] Thiang, and SuryoWijoyo, "Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot", International Conference on Information and Electronics Engineering IPCSIT vol.6, IACSIT Press, pp.179-183, Singapore, 2011.

[10] Mathias De Wachter, Mike Matton, Kris Demuynck, Patrick Wambacq, Member, Ronald Cools, and Dirk Van Compernolle, "Template Based Continuous Speech Recognition", IEEE Transactions On Audio, Speech & Language Processing, pp. 898-901, 2007.

[11] Hajer Rahali, ZiedHajaiej, and Noureddine Ellouze, "Robust Features for Speech Recognition using Temporal Filtering Technique in the presence of Impulsive Noise", I.J. Image, Graphics and Signal Processing, Vol.11, pp.17-24, 2014.

[12] Shumaila Iqbal, Tahira Mahboob, and Malik Sikandar Hayat Khiyal, "Voice Recognition using HMM with MFCC for Secure ATM", International Journal of Computer Science Issues (IJCSI), Vol. 8, Issue 6, No 3, pp.297-303, November 2011.

[13] Anuradha S. Nigade, and J. S. Chitode, "Throat Microphone Signals for Isolated Word Recognition Using LPC", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, Issue 8, pp.906-913, August 2012.

[14] Eslam Mansour mohammed, Mohammed Sharaf Sayed, Abdallaa Mohammed Moselhy, and Abdelaziz Alsayed Abdelnaiem, "LPC and MFCC Performance Evaluation with Artificial Neural Network for Spoken Language Identification", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 6, No. 3,pp. 55-66, June, 2013.

[15] Bhupinder singh, Rupinder Kaur, Nidhi Devgun, and Ramandeep Kaur, "The process of feature extraction in automatic speech recognition system for computer Machine interaction with humans: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, Issue 2, Feb 2012.

[16] Kashyap Patel, and R. K. Prasad, "Speech Recognition and Verification Using MFCC & VQ", International Journal of Emerging Science and Engineering (IJESE), ISSN: 2319–6378, Volume-1, Issue-7, pp. 33-37, May 2013.

[17] Network Classifiers", International Journal of Modern Engineering Research (IJMER) Vol.2, Issue.3, pp-854-858 ISSN: 2249- 6645,pp. 854- 858, May-June 2012.

[18] E. Chandra, K. Manikandan, and M. Sivasankar, "A Proportional Study on Feature Extraction Method in Automatic Speech Recognition System", International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, Vol. 2, Issue 1, pp. 772-775, January 2014.

[19] Handy K. Elminir, Mohamed Abu El-Soud, and L. M. Abou ElMaged, "Evaluation of Different Feature Extraction Techniques for Continuous Speech Recognition", International Journal of Science and Technology, ISSN 2224-3577, Volume 2 No.10, October 2012.

[20] Thiang, and SuryoWijoyo, "Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot", International Conference on Information and Electronics Engineering IPCSIT vol.6, IACSIT Press, pp.179-183, Singapore, 2011.

[21] Garima Vyas, and Barkha Kumari, "Speaker Recognition System Based On MFCC and DCT", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Vol.2, Issue-5, pp. 145-148, June 2013.

[22] Shumaila Iqbal, Tahira Mahboob, and Malik Sikandar Hayat Khiyal, "Voice Recognition using HMM with MFCC for Secure ATM", International Journal of Computer Science Issues (IJCSI), Vol. 8, Issue 6, No 3, pp.297-303, November 2011.

[23] Santosh Gaikwad, Bharti Gawali, and Pravin Yannawar, "Performance Analysis of MFCC & DTW for Isolated Arabic Digit", International Journal of Advanced Research in Computer Science, Vol. 2 (1), pp. 513-518 Jan. –Feb, 2011.

[24] Serajul Haque, Roberto Togneri, and Anthony Zaknich, "ZeroCrossings with Adaptation for Automatic Speech

Recognition", Proceeding of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson.University of Auckland, New Zealand. December 6-8, pp. 199-204, 2006.

[25] J. Manikandan, and B. Venkataramani, "Design of a real time automatic speech recognition system using Modified One Against All SVM classifier" Microprocessors and Microsystems, ELSEVIER, Vol.35, pp. 568–578, 2011.

[26] Jun Wang, Dong Wang, Ziwei Zhu, Thomas Fang Zheng and Frank Soong, at Center for Speaker and Language Technologies (CSLT), on I-vectors, a Discriminative Scoring for Speaker Recognition Based, 2014.

[27] Yuxuan Wang, Kun Han, and DeLiang Wang, Fellow, IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, Based Speech Segregation, 2012.

[28] Ravikiran, D. N., & Dethe, C. G. (2018). Improvements in Routing Algorithms to Enhance Lifetime of Wireless Sensor Networks. International Journal of Computer Networks & Communications (IJCNC), 10(2), 23-32.

[29] Ravikiran, D. N., & Dethe, C. G. Fuzzy Rule Selection using LEACH Algorithm to Enhance Life Time in Wireless Sensor Networks. Advances in Wireless and Mobile Communications. ISSN, 0973-6972.

[30] Rajesh, G., Thommandru, R., & Subhani, S. M. DESIGN AND IMPLEMENTATION OF 16-BIT HIGH SPEED CARRY SELECT PARALLEL PREFIX ADDER.

[31] Polanki, K., Purimetla, N. R., Roja, D., Thommandru, R., & Javvadi, S. Predictions of Tesla Stock Price based on Machine Learning Model.

[32] Thommandru, R. A PROSPECTIVE FORECAST OF BRAIN STROKE USING MACHINE LEARNING TECHNIQUES.

[33] Rajesh, G., Raja, A., & Thommandru, R. OPTIMIZATION OF MINIATURIZED MICROSTRIP PATCH ANTENNAS WITH GA.

[34] Vellela, S. S., & Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.

[35] Vellela, S. S., & Balamanigandan, R. (2024). Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. Multimedia Tools and Applications, 83(3), 7919-7938.

[36] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., & Swapna, D. (2022, November). An Adaptive Load Balancing Technique for Multi SDN Controllers. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1403-1409). IEEE.

[37] Priya, S. S., Vellela, S. S., Reddy, V., Javvadi, S., Sk, K. B., & Roja, D. (2023, June). Design And Implementation of An Integrated IOT Blockchain Framework for Drone Communication. In 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-5). IEEE.

[38] Vellela, S. S., & Balamanigandan, R. An intelligent sleep-awake energy management system for wireless sensor network. Peer-to-Peer Netw. Appl.(2023).

[39] Addepalli, T., Babu, K. J., Beno, A., Potti, B. M. K., Sundari, D. T., & Devana, V. K. R. (2022). Characteristic mode analysis of two port semi-circular arc-shaped multiple-input-multiple-output antenna with high isolation for 5G sub-6 GHz and wireless local area network applications. International Journal of Communication Systems, 35(14), e5257.

[40] Srija, V., & Krishna, P. B. M. (2015). Implementation of agricultural automation system using web & gsm technologies. International Journal of Research in Engineering and Technology, 04 (09), 385-389.

[41] Potti, D. B., MV, D. S., & Kodati, D. S. P. (2015). Hybrid genetic optimization to mitigate starvation in wireless mesh networks. Hybrid Genetic Optimization to Mitigate Starvation in Wireless Mesh Networks, Indian Journal of Science and Technology, 8(23).

[42] Potti, B., Subramanyam, M. V., & Prasad, K. S. (2013). A packet priority approach to mitigate starvation in wireless mesh network with multimedia traffic. International Journal of Computer Applications, 62(14).

[43] Potti, B., Subramanyam, M. V., & Satya Prasad, K. (2016). Adopting Multi-radio Channel Approach in TCP Congestion Control Mechanisms to Mitigate Starvation in Wireless Mesh Networks. In Information Science and Applications (ICISA) 2016 (pp. 85-95). Springer Singapore.

[44] S Phani Praveen, Sai Srinivas Vellela, Dr. R. Balamanigandan, "SmartIris ML: Harnessing Machine Learning for Enhanced Multi-Biometric Authentication", Journal of Next Generation Technology (ISSN: 2583-021X), 4(1), pp.25-36 . Jan 2024.