An improved GUI based Voice Identification with Applied RASTA-PLP

K.Vidyasagar[1], T.Chaitanya[2] , K.Swathi[3] and K.Sudharshanarao[4]
Sai Spurthi Institute of Technology, B-Gangaram,Satuupalli,India
Email: K.vidyasagar818@gmail.com
Email: tchaitanya31@gmail.com, swathi7201@gmail.com,siddhu.486@gmail.com

Abstract — **Speech signal processing and voice identification is having its own significance in critical investigation process. Voice signal are non stationary. Wavelet transforms used for processing non stationary voice signals. Word dependent voice recognition system is proposed with the combination of Relative Spectral Algorithm (RASTA) for signal filtering and linear predictive coding (LPC) for detecting the voice signals. Wavelet coefficients are derived for statistical analysis. The proposed methodology comprises Voice registration process and Voice verification process. This process is simulated using MATLAB. The mean, standard deviation, variance, median statistical parameters are considered to verify the test voice signal with the registered voice signal.**

Index Terms — **RASTA , LPC, Wavelets, mean, variance, standard deviation, median**

I. Introduction

Speech recognition and processing is having significance in critical investigation process. Word based identification is one phenomena to verify the identity of the speaker. The speech signal consists of resonant frequencies of vocal cards labeled as formants. Stationary and Non Stationary noise signals are also having wide scope of influence on the original signal. Multiple Wavelet decomposition is one mechanism Applied on the speech signal. Successive approximation based iterative wavelet decomposition is applied to decompose into low frequency components. Inverse discrete wavelet transform is used to reconstruct the signal and to remove the aliasing effect due to signal decomposition. The relative spectral algorithm (RASTA) is applied to remove the non stationary noise signal using mean subtraction.

II. Literature Survey

This paper focused on feature extraction of the speech signal using linear predictive codes (LPC), perceptual linear prediction (PLP) , Mel frequency cepstral coeffients (MFCC) .PLP and

MFCC is used to extract the features of the voice signal. LPC is used for future features prediction. The Author considered four features to the input layer of the feed forward back propagation neural network [1].

This paper proposed MFCC mechanism, multiple features are clustered using K-means Algorithm and created a VQ code book. Summary report of different mechanisms for speech features discussed [2] .

This paper studied the feature vector sets using MFCC and RASTA –PLP algorithm. Support vector machine and navie bayes mechanisms tested on the voice signals. This paper concludes the GMM super vector level fusion produces reliable results when augmenting with RASTA-PLP and MFCC features [3].

This paper investigates the significance of BPN. The features are extracted using RASTA- LPC and continuous wavelet transform. Principal component analysis (PCA) is applied to obtain conservative signal information [4].

This paper describes the classification and recognition of speeches .MFCC features are extracted for classification. Gaussian mixture model (GMM) is also used in their work. In order to obtain the maximum likelihood maximization algorithm is used [5].

RASTA-PLP algorithm is used to eliminate the noise and LPC is used to detect the resonance of the signal [6].

This paper proposed a novel method of statistical feature extraction (SFX) mechanism. This paper focused on the influence of data mining algorithms on multiple data sets[7].

This paper described a mechanism to recognize the speech in real time. The features are extracted using MFCC. Half raised sine function is applied on the resultant signal. This paper also compares the improved DWT with the conventional DWT. The entire system is implemented on Field programmable gate array [8].

III. Methodology

A voice signal is fed as an input signal to the proposed system for registration process. The output signals are plotted after de noising the test and registered voice signal. The wavelet energy is estimated along with, mean, variance, standard deviation and median values of the inputted signal is observed. These observed values are registered in data base for comparison. The results are plotted for analysis using graphic user interface (GUI). The original inputted signal is plotted in order to compare with the test signal. The mean, standard deviation, variance and median statistical values for both the signals will be computed in the second and third successive levels. In the verification process the wavelet energy level percentage, formant percentage, second and third level percentages are computed to determine the verified status of the voice signal.

Wavelet decomposition is carried until the vector contains a single value. While decomposing the wavelets by considering the approximation and detail, specific amount of energy is retained in the process of decomposition. This energy is the ratio of original signal and decomposed signal labeled as wavelet energy.

Level 2 and Level 3 coefficients were extracted using the wavelet decomposition process. In the recognition phase Level 3 coefficients contain minimal correlation information. Most of the correlated information contains in the Level 2 coefficients. The low correlation information is further threshold in order to estimate the mean, standard deviation, variance and median parameters, formant estimation and wavelet energy. Test signal values compared with the registered signal values for verification process. The verification is the ratio of test value and the registered value observed in percentage.

Relative spectral algorithm (RASTA) is proposed to remove the non stationary noise in the speech signal varies with time. Since the stationary noise does not change for a period of time, detection and removal of stationary noise is not a typical one to deliberate. Hence RASTA is applied to filter the non stationary noise signal.

a) Algorithm

Step 1: Read the Test input voice signal
Step 2: Apply "Rasta" algorithm
Step 3: De noise the voice signal for smoothing
If
The non stationary noise is removed
Then Go to Step4
Else
Go to step 2
Step 4: Select the wavelet bases and

Determine the decomposition level then
Apply discrete wavelet transform (DWT)
Step 5: Determine the threshold value and threshold function
Step 6: De noise the signal
Step 7: Apply inverse DWT
Step 8: Reconstruct the Signal
Step 9: Estimate the formants using LPC Algorithm
Step10: Decompose the wavelets until a single value
Step 11: Extract the Level 2 and Level 3 coefficients
Step 12: Estimate the wavelet energy
Step 13: Apply 'Threshold' for smoothing the signal
Step 14: Estimate the statistical values.
Step15: Formants and wavelet energy values compare with register values
Step16: If the verification rate is above 75 percentages
Then
Identity verified
Else
Identity not verified.
Step17. Stop

IV. RESULTS AND DISCUSSION

The verification process is preceded with different voice signals. Using the MATLAB. The test signals are compared with the registered voice signals. The formants and mean standard deviation variance and median parameters are estimated for comparison. Figure1 represents the voice signal registration process. The signal de noised and approximated signals are displayed. The signal Approximation when compared with the original signal, the entire noise is suppressed and more supportive to estimate the statistical values. The RASTA algorithm in association with linear predictive coding is more prudent mechanism for Speech process recognition applications. Figure 2 and Figure 3 represents the Voice verification stage.
While analyzing the vocal tract from Figure 2 and Figure 3, resonant peaks are observed labeled as formant frequencies (F). Format 1 (F1) , format 2 (F2), format 3 (F3), format 4 (F4) are considered for the voice signals. Linear predictive coding (LPC) derived from the word linear prediction is adopted to detect the formants. The predicted formants using the LPC will be removed by applying the Inverse filtering mechanism. Predicted signal and test signal mean square error is estimated to remove the formants. Hence the resonant effect is curtailed from the test signal. The residue signal (filtered) is analyzed. For the test signal 1(Figure 2), F1 and F2 is zero and for F3 and F4 the signal levels observed as 2310.47 peak values. For test signal 2 (figure 3), F1 is observed as 0 and for F2 and F3 the peak signal levels are 508.8. For F4 is 4000.
Table 1, Table 2, table 3, table 4 represents the verification process carried on different sets of the voice signal.

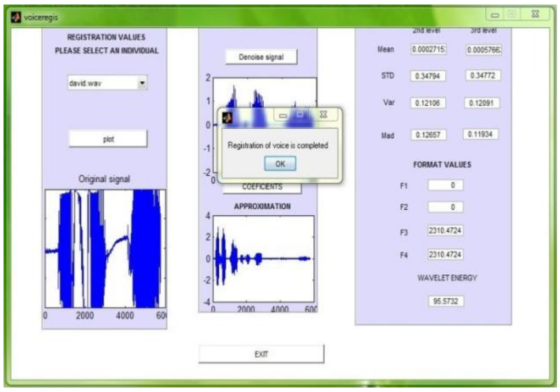

Figure 1. Voice Registration values

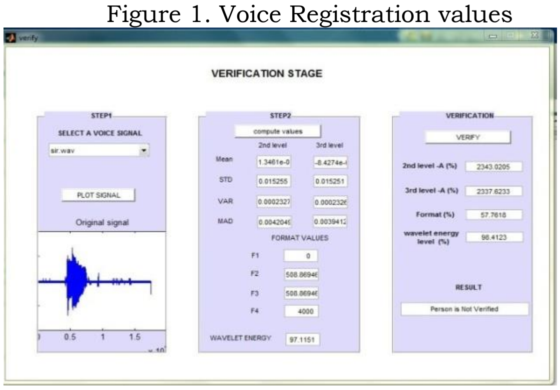

Figure 2 . Voice verification not matched

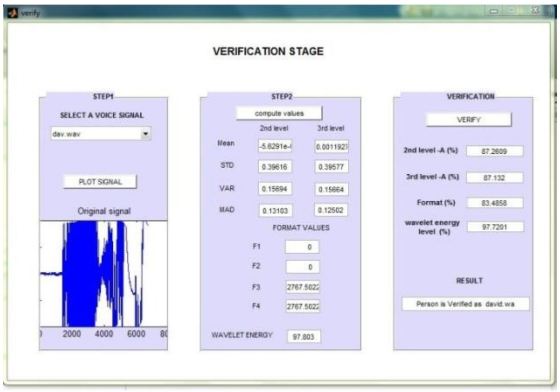

Figure 3. Voice verification matched

Table1. Test results set 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	Verified	Not Verified
5	Not Verified	Not Verified	Not Verified	Not Verified	Verified

Table2. Test results set 2

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

Table3. Test results set 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

Table4. Test results set 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	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

Table5. Test results set 5

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

V. CONCLUSION

An improved GUI based Voice Identification with Applied RASTA-PLP algorithm is tested on five sets of samples. Perhaps Non stationary Noise signals are curtailed using the RASTA algorithm, certain stationary noise signals may influence on statistical results analysis. Applying the Discrete Wavelet Transform removes the remaining noise signal. Applying LPC removed the formants present in the Continuous speech (voice) Signal.

In future the improved Algorithm is augmented with Back Propagation neural network (BPN) in order to enhance the accuracy of the verification process. The resulted signal of the proposed algorithm output is applied as test signal to the BPN to iterate the process until the resultant signal closes to the original signal. With the proposed method '87'percent of the signal closes to the original signal is deliberated as verified. This percent may be enhanced to considerable value using BPN.

VI. REFERENCES

[1] Namrata Dave, "Feature Extraction Methods LPC, PLP and MFCC in Speech Recognition" , International journal for advance research in engineering and technology, Vol. 1(6): page 1-5, July 2013.

[2] Amit kumar singh,rohit singh, ashutosh dwivedi, "evolvement and recent research in parametric representations of speech features for automatic speaker recognition", International Journal of Electrical, Electronics and Data Communication,Vol.2 (1):Pages 11-15, Jan,2014.

[3] Imen Trabelsi and Dorra Ben Ayed, " A Multi Level Data Fusion Approach for Speaker Identification on Telephone Speech" , International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 6(2): pages 33-42, April 2013.

[4] Mohamed Hassine, Lotfi Boussaid, Hassani Massouad, "Hybrid Techniques for Arabic Letter Recognition", International Journal of Intelligent Information Systems, Vol.4 (1): pages 27-34, February 2015.

[5] S G Bagul & Prof R.K.Shastri, Text Independent Speaker Recognition System using GMM, International Journal of Scientific and Research Publications, Volume 2, Issue 10, October 2012

[6] Rajparthiban Kumar, et.al, "Development of a Novel Voice Verification System using Wavelets", Proceedings of the International Conference on Computer and Communication Engineering, 2008 May 13-15, 2008

[7] Simon Fong, Kun Lan, and Raymond Wong, "Classifying Human Voices by Using Hybrid SFX Time-Series Preprocessing and Ensemble Feature Selection", BioMed Research International, Vol.2013 (2013): pages 1-27, 2013.

[1] xiaohui hu, et.al, isolated word speech recognition system based on FPGA", journal of computers, vol. 8(no. 12) , December 2013.

ACKNOWLEDGEMENT

The authors like to express sincere thanks to the management and principal of Sai Spurthi Institute of Technology for providing the fund and necessary infrastructure.

AUTHOR PROFILE

Prof. K. Vidyasagar: Received B. Tech degree in Instrument Technology from Andhra University College of Engineering Visakhapatnam, M.E from P. S. G. Tech Coimbator. He is now a research scholar under the guidance of Dr. A. Bhujangarao, Andhra University. His current research interests include image processing in biomedical instrumentation and related embedded systems.

K.Swathi: Received B.Tech degree from Sai Spurthi Institute of technology, Sattupalli, M.Tech From Mother Theresa institute of Science and technology. Her current research interests are communication systems and related applications.

T. Chaitanya : Received B. Tech degree in Electronics and Communication Engineering from Bapatla Engineering College Bapatla, M. Tech Systems and Signal processing from Adams Engineering College, Paloncha. Her current research interests including VLSI Signal Processing and related embedded systems.

K.Sudharsanarao: Received B.Tech degree from Sai Spurthi Institute of technology, Sattupalli, M.Tech From Mother Theresa institute of Science and technology. His current research interests are communication systems and related applications.