

Improved Dynamic Speaker Recognition System using NLMS Adaptive Filter

P. Hema Kumar
P.G Scholar
Dept of ECE
SIET, Narasapur.

V. Srinivas
Dept of ECE
SIET, Narasapur,
India.

T. Madhu
Professor in ECE
SIET, Narasapur,
India.

ABSTRACT

In a controlled environment, we can implement a speaker recognition system using MFCC and Vector Quantization. So, the main objective of this paper is to develop a speaker recognition system using MFCC and Vector Quantization(VQ) in a noisy environment, when the input speech utterance is given through a microphone. Normalised Least Mean Square Adaptive(NLMS) Filter is used to improve the performance of the system in noisy environment. So the NLMS Adaptive filter is used to reduce the background noise from input speech signal and then the filtered signal is given to the Feature Extraction phase. For implementation simplicity, it is developed as Text- Dependent Speaker Recognition System with 10 speakers, each speaker locally recorded database is used for training. The performance of the proposed system tested in noisy environment with and without using the NLMS adaptive filter and improved recognition evaluated using Equal Error Rate (EER).

Keywords

Least Mean Square, NLMS Adaptive Filter, Vector Quantization, Equal Error Rate (EER).

1. INTRODUCTION

Speaker Authentication is one of the biometric applications that takes unknown's voice for verification purpose. The main objective of the Speaker Recognition System is to identify the unknown from the features of his/her voice which in turn makes an excellent tool to operate the devices for many security applications.

The main problem in speaker recognition system is identifying the unknown speaker in noisy environment. Generally we recorded the speech signal in different environmental conditions, so the undesired features of a speaker results and this leads to degrade in the performance of the system drastically. That is, recording the unknown speaker voice through microphone dynamically and comparing the features with features of reference model in trained database leads to incorrect results especially in noisy environments.

So, the main objective of this paper is developing a system using MFCC for feature extraction and Vector Quantization for Feature Matching in a noisy area such that the noise from the input speech signal is cancelled by using Adaptive filtering techniques. The noise cancellation from input speech utterance can be done by using Recursive Least Square (RLS), Normalised Least Mean Square (NLMS) and Least Mean Square (LMS) Adaptive Filters. So, the noise gets reduced and filtered signal is given for feature extraction. So, ultimately the performance of the system gets improved when we are using the NLMS Adaptive filter compared to that of the system without using filter in a disturbed environment. The complete processing is done through MATLAB.

2. REVIEW OF LITERATURE

Dr. H B Kekre, Dr. V A Bharadi, A R Sawant, OnkarKadam, PushkarLanke, RohitLodhiya proposed a Speaker recognition System based on MFCC and Vector Quantization. Calculated Euclidean distances between two voice samples for testing this method. The testing code is written in MATLAB. They proposed use of a MFCC feature extraction with KMCG for feature vector generation & matching in speaker recognition system. They concluded that current system has great potential to be used for spectral feature extraction.

Speaker Recognition using MFCC and Vector Quantisation by GeetaNijhawan, Dr. M.K Soni given that background noise influences the overall efficiency of speaker recognition system and is still considered as one of the most challenging issue in Speaker Recognition System (SRS). Feature Extraction of the input speech can be done using MFCC and then vector quantization of the extracted MFCC features is done using VQLBG algorithm and concluded that Voice Activity Detector (VAD) can be used which dominantly improves the performance of the system in noisy conditions.

Prof.Ch.Srinivasa Kumar and Dr. P. Mallikarjuna Rao proposed "Automatic Speaker Recognition System using MFCC and VQ. The system was tested many times with various databases and found to be very reliable.

The analysis of the speech signal in different environments and studied the accuracy and performance of the speaker recognition system was presented by Satyanand Singh , Dr.E.G Rajan. They analysed about the mismatch that affecting of speaking style, utterance quality and also considering surrounding noise.

The performance of the LMS Adaptive filter is used for noise cancellation in voice signal was illustrated by Prof.Akhilesh Singh Thakur and Dr.Vibha Tiwari. With the variation in the Step Size, noise can be reduced but steady state error increases, that algorithm was developed or improved by changing the values of parameters separately or combinely.

Jyoti Dhiman, Shadab Ahmad, Kuldeep Gulia proposed a paper which illustrates the implementation aspects of LMS and NLMS adaptive filters and the performance of these filters with respect to computational complexity and Signal to Noise Ratio.

"Design of an Adaptive Filtering Algorithm for Noise Cancellation" was a paper presented by Roshny Jose George . In this paper, they focussed on implementation of Least Mean Square(LMS) Adaptive filter and measured the improvement

by calculating Signal to Noise Ratio of noisy signal and filtered signal. It clearly said that human speech signals, acquired through microphone, were given as input signal and the output signals from the filter was noise free. Also the results showed that the algorithm gives the better result for different speech signals.

3. PROPOSED METHOD

MFCC provides the efficient features in the clean environment, here we are not modifying the architecture of MFCC. Instead we are processing the input speech signal before going to Feature Extraction (MFCC) Phase. The pre-processing proposed technique that we are using here is LMS Adaptive Filter. LMS Adaptive Filter is used to reduce or cancel the noise from noisy signal. So, the input speech signal is directly applying to a Adaptive filter and then the cleaned signal or filtered signal is given to the further stage called Feature Extraction Phase. So that MFCC gives improved or exact features of the unknown person even in the presence of surrounding noise.

So that the features of the unknown speaker that is acquired through microphone for testing is compared with the features of Reference model (Trained Database) using Feature matching algorithm. The codebook of each speaker can also be derived. Each time the Euclidian Distance is computed for the two speakers, such that the speaker corresponding to the VQ Codebook with smallest total distortion (min. VQ distance) is identified as the speaker of input speech signal.

The unique representation of each speaker in an efficient manner, this process is called Vector Quantization. It produces small set of feature vectors from a large set of vectors such that to minimize the average distance to every other point. So, for determining the similarity of vectors, the overall computation can be reduced and often reduced the table lookup of similarities between code book vectors.

The proposed method is given in below figure.

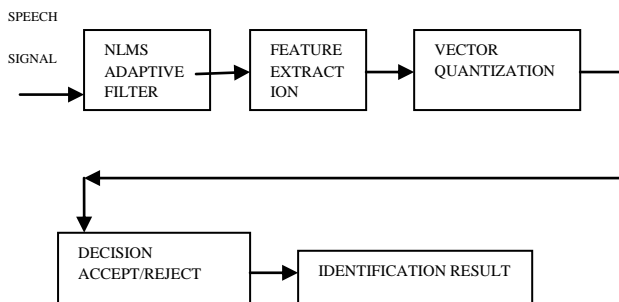


Fig 1: Block Diagram of Proposed method

3.1 Least Mean Square (LMS) Adaptive Filter

The most important method of computing a signal a signal corrupted with some additive noise is to allow the signal to pass through a filter which suppresses noise and does not change the signal. Specifically, Adaptive filters are capable of adjusting the parameters automatically and its design does not require the knowledge of previous characteristics.

Least Mean Square (LMS) filter is a kind of adaptive filter used to make a desired filter by finding filter coefficients that produces the least mean squares of the error signal. Adaptive filter does not have any constant filter coefficients and automatically the parameters are adjustable. So, such self adjustment in their parameters minimizes the

error signal. Also the main challenge is operating under varying conditions, this can be fulfilled by NLMS Adaptive filter. Ofcourse, Adaptive filter can be used in many signal processing areas such as echo cancellation, signal prediction etc., but here we used Adaptive filter in this paper for Echo Cancellation.

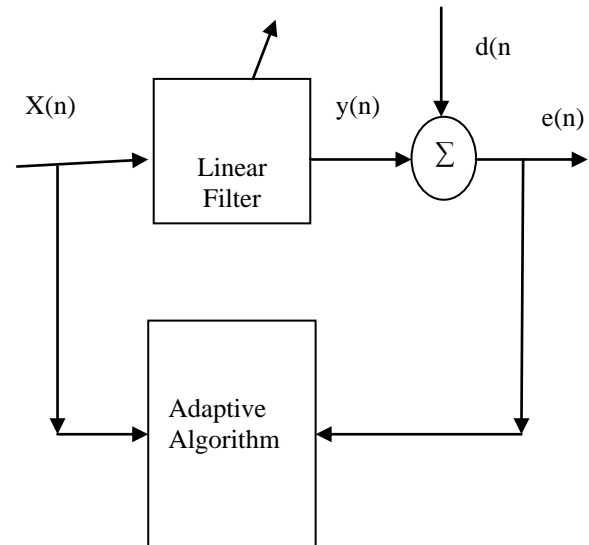


Fig 2. Typical Adaptive Filter

There are number of parameters in Adaptive filters that plays an important role in reducing or eliminating the noise but noise cancellation can be obtained by variation in step size and thus adaptive filter deals with noise cancellation for signal corrupted with white Gaussian noise. For the small amount of changes in weight, it changes about the optimal weights and thus if variance changes large extent due to weights change the convergence in mean would be misleading. So, the proper selection of step size is very much needed in order to avoid the problem.

The main drawback of LMS algorithm is it is very sensitive to scaling of its input. The normalised Least mean square adaptive filter is another class of LMS algorithm that clearly overcome the drawback of LMS algorithm by normalising the power of the input. Also the range of step size and mean square deviation depends on the power of input signal, this makes clearly the NLMS adaptive filter can be effectively used than that of LMS Adaptive filter. The convergence speed of NLMS adaptive filter depends upon the statistics of input signal.

4. EXPERIMENTS

Experiments were conducted for speaker recognition using a single word database. 20 Speakers were selected to be used as speaker model.

The experiment of the speaker verification was conducted in noisy environment. Then the speech signal was passes through the Feature extraction phase, to get the Mel Frequency Cepstral(MFCC) Coefficients and code book for VQ models were generated during the training phase. During the evaluation phase, the Euclidean distance of each speaker for corresponding VQ codebook will be calculated.

Experiments in noisy environments with using LMS adaptive filter were also conducted and following the same procedure as above in noisy environment without using LMS adaptive filter.

The filtered signals and the speech signal without filtered also observed as shown below.

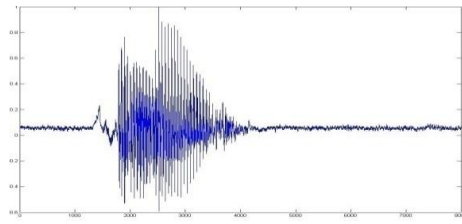


Fig 3: Speech signal without filter

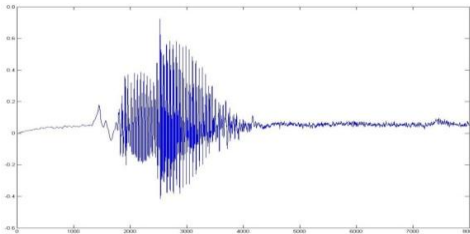


Fig 4: Speech signal with filter

Also the experiments in noisy environments with using NLMS adaptive filter also conducted and also observed the various performance metrics that shows the improvement in present system when compared to that previous experiments.

5. RESULTS

In this Speaker Recognition system, we used FAR, FRR and EER for the performance evaluation. The speaker that is going to be identified is matched against every known template, yielding a Euclidean distance describing the similarity between the speaker and the template. Also EER is a threshold independent performance measure such that lower the EER is, better is the System's performance.

The Equal Error Rate (EER) of 53% was achieved in noisy environment without using any filter before the feature extraction phase. The overview for 3 repetitions of the same utterance by a speaker as shown in Table 1.

Table 1. Output for Noisy Environment

Speaker	Euclidian Distance	Result
1 st time	1.09	Speaker does not found
2 nd time	1.11	
3 rd time	1.21	

The False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (ERR) of 33% was achieved in noisy environment using LMS Adaptive Filter. By using this filter, we have the improvement in Euclidean distance and it is better than that of without using any filter. This LMS adaptive filter improves the performance of the system as shown in Table 2 and figure 6 shows the plot between FAR and FRR.

Table 2. Output for Noisy Environment using LMS Adaptive Filter

Speaker	Euclidian Distance	FRR	FAR	EER
1 st time	0.95	31.81	41.23	42.26
2 nd time	0.922	12.72	38.14	38
3 rd time	0.94	20.58	39.17	33

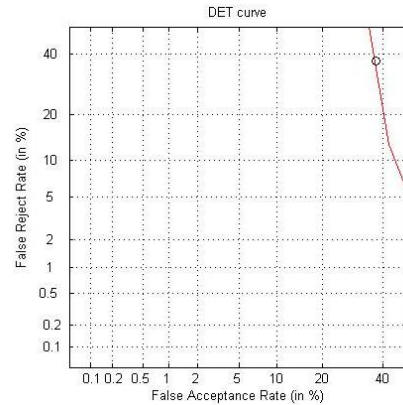


Fig 5: Plot for DET Curve

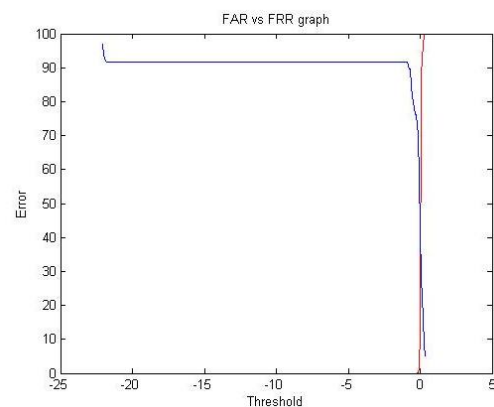


Fig 6: Plot for FAR and FRR

The Equal Error Rate (EER) of 27% was achieved in noisy environment using NLMS Adaptive filter. This significantly improves the performance of the system more than that of using LMS Adaptive Filter. The improvement in Euclidean distance is clearly observed in the below table when compared with the method of using LMS Filter and also without using any filter. The overview for 3 repetitions of the same utterance by a speaker as shown in Table 3 and figure 8 shows the plot between FAR and FRR.

Table 3. Output for Noisy Environment using NLMS Adaptive Filter

Speaker	Euclidian Distance	FRR	FAR	EER
1 st time	0.81	13.26	31	27.58
2 nd time	0.84	28.34	36	29.28
3 rd time	0.88	44.12	41	30.45

If the distributions are overlap, the FAR and FRR intersect at a certain point. The value of FAR and FRR at this point is called the Equal Error Rate (ERR) and this is clearly shown in below figure 8. The lower the EER is, the better is the System's performance and this is shown in above table also.

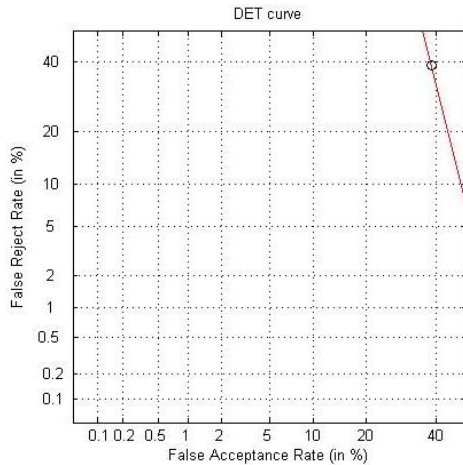


Fig 7. Plot for DET Curve

If the distributions are overlap, the FAR and FRR intersect at a certain point. The value of FAR and FRR at this point is called the Equal Error Rate (ERR) and this is clearly shown in below figure 8.

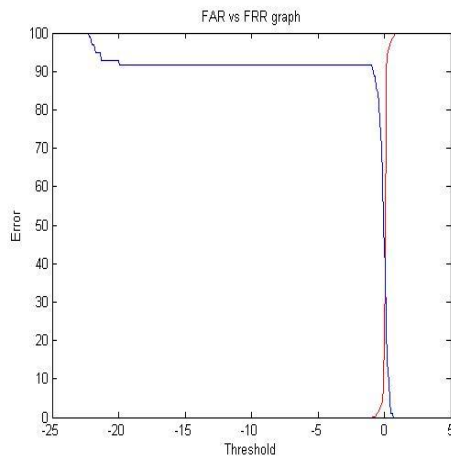


Fig 8. Plot for FAR and FRR

6. CONCLUSION

This paper clearly gives the idea how NLMS adaptive filter improves the performance of the System in noisy environments. The Equal Error Rate (EER) of 23% was achieved using this method. However, the accuracy decreases as number of feature vectors increases. Furtherly, the work can be extended for increasing the performance of the speaker recognition system using another class of adaptive filter called Recursive Least square(RLS) filter.

7. REFERENCES

- [1] Salina Abdul Samad, Aini Hussain, Khairul Anuar Ishak "Improved Hybrid Speaker Verification in Noisy Environments Using Least Mean -Square Adaptive Filters".
- [2] Roshny Jose George "Design Of An Adaptive Filtering Algorithm For Noise Cancellation" International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 04 | July-2015.
- [3] John Creighton and R.Doraiswami "Real Time Implementation Of an Adaptive Filter For Speech Enhancement" 2004 IEEE.
- [4] Geeta Nijhawan, Dr. M.K Soni "Speaker Recognition Using MFCC and Vector Quantisation" Int. J. on Recent Trends in Engineering and Technology, Vol. 11, No. 1, July 2014.
- [5] Signal Enhancement Using LMS and Normalized LMS - MATLAB & Simulink - MathWorks India.
- [6] Michael Lutter "Mel Frequency Cepstral Coefficients (feature extraction/ Mfcc)" The Speech Recognition Wiki25 November 2014.
- [7] J.P.Campbell, "Speaker Recognition: A Tutorial", Proc. Of the IEEE, Vol 85, No. 9, September 1997, pp. 1437-1462.
- [8] Vibha Tiwari "MFCC and its applications in speaker recognition" International Journal on Emerging Technologies 1(1): 19-22(2010) ISSN : 0975-8364.
- [9] Prof. Ch.Srinivasa Kumar, Dr. P. Mallikarjuna Rao "Design Of An Automatic Speaker Recognition System Using MFCC, Vector Quantization And LBG Algorithm" International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 8 August 2011, ISSN : 0975-3397.
- [10] Prof. Vaishali M. Karne, Prof. Akhilesh Singh Thakur , Dr. Vibha Tiwari "Least Mean Square (LMS) Adaptive Filter For Noise Cancellation" International Journal of Application or Innovation in Engineering & Management (IJAIEM) , ISSN 2319 – 4847.
- [11] Dr. H B Kekrel, Dr. V A Bharadi, A R Sawant "Speaker Recognition using Vector Quantization by MFCC and KMCG Clustering Algorithm" 2012 International Conference on Communication, Information & Computing Technology (ICCICT), Oct. 19-20.
- [12] Sheng Zhang, student Member, IEEE, Jiashu Zhang, and Hongyu Han "Robust Variable Step-Size Decorrelation Normalized Least-Mean Square Algorithm and its Application to Acoustic Echo Cancellation" IEEE/ACM Transactions on Audio, Speech, and Language Processing.
- [13] A. Srinivasan "Speaker Identification and Verification using Vector Quantization and Mel Frequency Cepstral Coefficients" Research Journal of Applied Sciences, Engineering and Technology 4(1): 33-40, 2012 ,ISSN: 2040-7467.
- [14] Jayant M. Naik "Speaker Verification: A Tutorial" January 1990 - IEEE Communications Magazine.
- [15] Dr. Sadaoki Furui "Speaker recognition" Sada oki Furui (2008), Scholarpedia ,3 (4):3 7 1 5
- [16] "Performance Analysis and Enhancements of Adaptive Algorithms and Their Applications." A thesis submitted to the Nanyang Technological University by Shengkui zhao.
- [17] "A Novel Windowing Technique for Efficient Computation of MFCC for Speaker Recognition" Md Sahidullah, Student Member, IEEE, Goutam Saha, Member, IEEE.
- [18] Paresh M. Chauhan , Nikita P. Desai "Mel Frequency Cepstral Coefficients (MFCC) Based Speaker Identification in Noisy Environment Using Wiener Filter"

- [19] Jorge MARTINEZ*, Hector PEREZ, Enrique ESCAMILLA , Masahisa Mabo SUZUKI “ Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques” -1-61284-1325-5/12/\$26.00 ©2012 IEEE.
- [20] Jyh-Min CHENG and Hsiao-Chuan WANG “A Method Of Estimating The Equal Error Rate For Automatic Speaker Verification” 0-7803-8678-7/04/\$20.00 82004 IEEE.
- [21] JYOTI DHIMAN, SHADAB AHMAD, KULDEEP GULIA “ Comparison between Adaptive filter Algorithms (LMS, NLMS and RLS)” International Journal of Science, Engineering and Technology Research (IJSETR) Volume 2, Issue 5, May 2013, ISSN: 2278 – 7798.
- [22] “Technical Document About FAR, FRR and EER Version 1.0” © 2004 by SYRIS Technology Corp.