Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Speech Recognition Using Unbiased And Normalized Adaptation Noise Reduction (Unanr) Algorithm Using Adaptive Filtering And Naive Bayes Classifier

Kadam Sarika Shamrao[1] , Dr. A. Muthukumaravel[2]

[1]Research Scholar, Department of Computer Science, Under the faculty of Science and Humanities, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

[2]Dean, Faculty of Arts and Science, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

Abstract :- In the human communicating system, speech signals are significantly vital. In speech processing functionalities like Automated Speech Recognition (ASR) system, removing background noise is critical since it reduces the system's performance. The adaptive filtering strategy is often employed as a filtering technique which will help in noise reduction and also for speech enhancement to tackle this difficulty. The Least Mean Square (LMS) approach is commonly used in adaptive filters to remove noise, although in practice, this approach has several drawbacks in voice enhancement. In this study, a robust adaptive filtering method based on the Unbiased and Normalized Adaptive Noise Reduction (UNANR) algorithm is used to de-noise voice signals. The Naive Bayes classifier is used for classification in order to improve the recognition of speech. According to the results, the UNANR approach outperformed the LMS algorithm in accuracy after classification using the Naive Bayes classifier. The UNANR accuracy was around 89 % which is higher than the baseline LMS algorithm which produces accuracy of 86%.

Keywords- Automated Speech Recognition, adaptive filters, LMS algorithm, UNANR algorithm, Naive Bayes, Accuracy

I INTRODUCTION

Speech is a reliable and effective method of communication between humans. Automatic Speech Recognition, Cellular applications, teleconference platforms, hearing equipments, voice coders and forensics are all examples of speech processing applications. Despite their convenience, these apps' accuracy is hindered in real-time by a variety of sources of noise, including speakers, ambient noise, and communication channel distortion. The catching of good quality signals is a critical challenge in all of these cases. It is necessary to strengthen the signal in terms of noise reduction to increase the quality of a noise-corrupted speech waveform. The significant area of speech enhancement is the improvisation of speech that has been deteriorated by noise, in which noise and distortion are suppressed through a filtering technique. Filtering is a signal processing procedure that aims to modify the information found in a signal. Adaptive and non-adaptive filtering techniques are the two types of filtering techniques most commonly used.

Adaptive filtering has been one among the most prevalent and effective technologies for speech enhancement in recent times. The average weights of this filtering technique are usually a self adjustable as per an optimum algorithm or a predetermined parameter. Since the adaptive filters' settings are constantly changing in effort to match a performance criterion, they are time-varying. This filter manages signals using a recursive algorithm. The training and testing is compared which will lead to a error signal as an outcome, which is then utilized to adjust certain earlier assumed filter settings. This process of adjustment proceeds until the steady-state condition is reached under the influence of the incoming signal.

In this work, we describe an Automatic Speech Recognition (ASR) system which uses a unique Unbiased and Normalized Adaptive Noise Reduction (UNANR) system that uses a gradient steepest descent method to reduce noise in speech signals. The Naive Bayes classification method is adopted for the speech recognition purpose. The UNANR model-based adaptive noise-reduction system outperforms the least-mean-square (LMS) filter in removing noise from speech recordings. After the classification process, to demonstrate UNANR's capability in the ASR system, we compared its performance to LMS-based realizations in terms of SNR, PSNR, Mean Absolute Error, Mean Square Error, Correlation and Accuracy.

II LITERATURE REVIEW

To increase the accuracy of ASR schemes, Tomoko Kawase et al., (2017) devised and validated a strategy for automatically switching noise - reducing factors. The noise-power vector, which is derived from high noise speech phrases are intended to categorize the voice statements and allocate the appropriate metric set to get the best identification accuracy, will be used to quantify noise properties. Studies with speech phrases tainted by multiple noise sources indicated that the proposed strategy can lower ASR's word error rate [1].

The improved LMS-ANR procedure is presented by Kalamani et al., (2018) in this study effort for upgrading the noisy Tamil voice signal in diverse noise polluted situations. This algorithm automatically adjusts its coefficients in response to fluctuations in incoming signal. The investigations are being carried out for numerous noises with different input SNR readings from the sentences collection of the database. This is conducted to ensure that the existing and designed filter perform as expected. The results show that the LMS-ANR algorithm performs well than the previous LMS noise filter in varying noisy situations, as evaluated by metrics such as improvement of PSNR, SNR, and MOS value [2].

For filtering background noise from voice signals, many filters have been developed and implemented. Minajul et al. [2018] surveyed and discussed in this paper many approaches for eliminating or lowering noise, including various adaptive filters. Initially, both additive white Gaussian noise (AWGN) and real noise were subjected to adaptive filtering algorithms. For both kinds of noise, the optimal filter, i.e., Kalman filter, is applied following the adaptive filters. According to the research, the Kalman filter outperforms other adaptive filters such as the LMS and NLMS for AWGN noise, although this is not the case for practical noise [3].

Acoustic noises are generally prevalent with a speech signal. As a result, researchers have devised numerous strategies for removing sounds, which are referred to as filtering. Filtering strategies differ depending on the type of application. Hence these techniques are divided into two major categories: linear filtering and non-linear filtering. Several filtering approaches have been presented by Minajuland Kaustubh (2019) in this study focused on the two forms of noise, linear and non-linear. If the influencing noise is linear, linear filtering techniques such as LMS and NLMS are essential, whereas non-linear noise necessitates non-linear filtering approaches such as enhanced variants of Kalman filters, fuzzy oriented adaptive filter, ANFIS and Neural Networks. This study includes multiple voice feature extraction approaches in addition to filtering strategies that can be utilized for voice filtering and many other speech-based applications [4].

In this research Radek Martinek et al., explored a new way of analyzing speech signals that will be used in voice command control of technological functions. A programmer for operating the KNX technique was built and employed, along with a Microsoft recognizer to identify the voice commands. A sound card and a LabView SW programme, as well as a database were employed, to manage the operations in the SH spaces. The LMS along with ICA are implemented. Control signals for five forms of interference were examined during the experiment. The hybrid technique had a greater performance level than the LMS algorithm, according to the findings [5].

The incorporation of machine learning to a speech recognition system is the main emphasis of this paper discussed by Sunanda Mendiratta et al., [2019]. Multiple machine learning techniques, such as ELM, SVM, ANN, and kNN classifier, are used to improve the core three stages of ASR. The simulation results indicate that the ELM classifier produces the best accuracy. The optimization of pre-processing and feature extraction techniques is also given a priority here. The research demonstrates that typical classifier results can be enhanced beyond by combining them with distinctive optimizing techniques for a better outcome [8].

III PROPOSED METHODOLOGY

Speech recognition's ideal intention is to listen to spoken words together while recognizing certain predefined language. This paper proposes a new voice recognition system that focuses on noise reduction utilizing an adaptive filtering strategy for pre-processing. The UNANR system was used in pre-processing to filter out the noise from the voice signal. The noise level of a given audio signal is measured and fragmented. The essential audio signal features, such as pitch and word length measure, are then extracted for optimal classification. Finally, using a Naive Bayes classifier, specified elements from extracted parameters are used to train the classifier. To demonstrate the efficacy of the proposed UNANR approach, it is compared with the common filtering technique, LMS Algorithm in the speech process.

Adaptive Filtering:

The adaptive filter is designed to reduce noise by estimating the gradient vector from the provided noisy data. In speech enhancement applications, the primary input signal p(n) is the combination of the speech signal s(n) and additive noise signal a(n) as expressed below

$$p(n) = s(n) + a(n) - - - (1)$$

When the reference input r(n) is associated with a(n) in certain manner, the error signal obtained from the adaptive filter is expressed as ‘e’ along with the filter output u(n) as $e = \{s(n) + a(n)\} - u(n)$, then

$$e^2 = \{s(n) + a(n)\}^2 - 2u(n)\{s(n) + a(n)\} + u(n)^2$$

$$= \{a(n) - u(n)\}^2 + s(n)^2 + 2s(n)a(n) - 2u(n)s(n) - - - - - (2)$$

The mean-squared error (MSE) is calculated because the signal and noise are uncorrelated.

$E[e^2] = E[\{a(n) - u(n)\}^2] + E[s(n)^2] - - - (3)$

Least-Mean-Square (LMS) Algorithm:

The filtering and adaptive practices make up the LMS algorithm. In filtering, the evaluation of the outcome of a transversal filter generated by a collection of tap input and obtaining an estimation error by determining this output to a specified response are the major processes. Then the tap weight of the filter is regularly modified in line with the estimated errors in an adaptive process. As a result, the interaction of these two processes creates a feedback loop over the LMS algorithm.

Tap-weight adaptation of the LMS algorithm is given by

$w(n+1) = w(n) + \mu\, p(n)e(n) - - - (4)$

Where w(n) is the tap weighted vector and μ represents the step size

Unbiased and Normalized Adaptive Noise Reduction (UNANR):

The UNANR model's adaptation approach is intended to change the co-efficient that is convolved through the reference input so as to measure the amount of noise in a given speech signal. The UNANR model is a linear finite impulse response (FIR) filter, as shown in Fig. 1. The filter's response for every time instant n can be written as,

$$u(n) = \sum_{m=1}^{M} w_m(n) r(n-m+1) - - - - (5)$$

Where r(n - m + 1) signifies the reference input noise at the prevailing (m = 1) and subsequent (m – 1) input samples, and $w_m(n)$ denotes the UNANR coefficients with order M

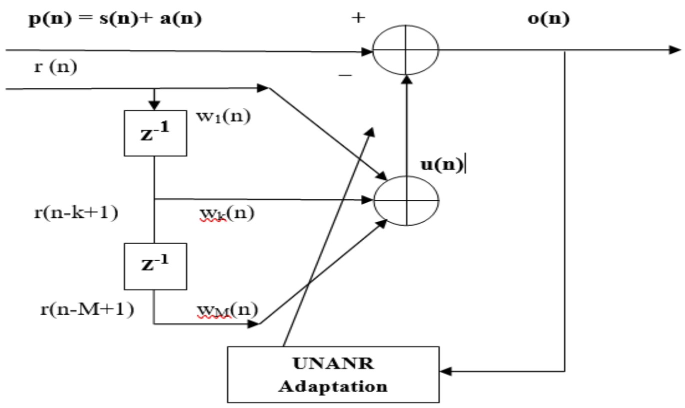

Fig: 1 Structure of UNANR

The UNANR model's adaptation phase is intended to change the coefficients that are convolved with the reference input in order to calculate noise included in the voice signal delivered.

$\hat{s}(n) = o(n) = p(n) - u(n) - - - (6)$

Squaring equation (6) on both sides and substituting equation (1) in that gives,

$\hat{s}^2(n) = p^2(n) + u^2(n) - 2p(n)u(n) - - - (7)$

$\quad = [s(n) + a(n)]^2 + u^2(n) - 2[s(n) + a(n)]u(n)$

$$= s^2(n) + 2s(n)a(n) + a^2(n) + u^2(n) - 2[s(n) + a(n)]u(n) - -(8)$$

Another purpose of the UNANR coefficient adaption strategy is to limit the error loss.

$$\varepsilon(n) = \hat{s}^2(n) - s^2(n)$$
$$= a^2(n) + 2s(n)a(n) + u^2(n) - 2[s(n) + a(n)]u(n) - - - (9)$$

Naive Bayes classifier:

Speech recognition is a multiclass classifying issue; hence Naive Bayes classifiers are chosen since they can support several classes. The supervised classification technique known as the Nave Bayes classifier is centered on the Bayesian theory, which is a basic and robust probability classification procedure. It primarily evaluates the probability values for every word in the text/sentence and outputs the term with the highest likelihood. It makes a prediction based on an event of previous occurrence. To determine the attributes for classification, it just requires a modest quantity of training data. Class conditional independence presumes that feature elements from one class are independent from other data points.

IV RESULTS AND DISCUSSIONS

Now let us see the results obtained by our proposed work. Fig 2 represents the overall segmented signals that is been given as input, Fig 3 represents single original signal, Fig 4 represents noise generation, Fig 5 represents noise removed speech, Fig 6 represents enhanced output speech signal, Fig 7 represents parameter analysis such as MAE, MSE, SNR, PSNR and Cross Correlation and at last Fig 8 represents accuracy analysis of UNANR and LMS algorithm.

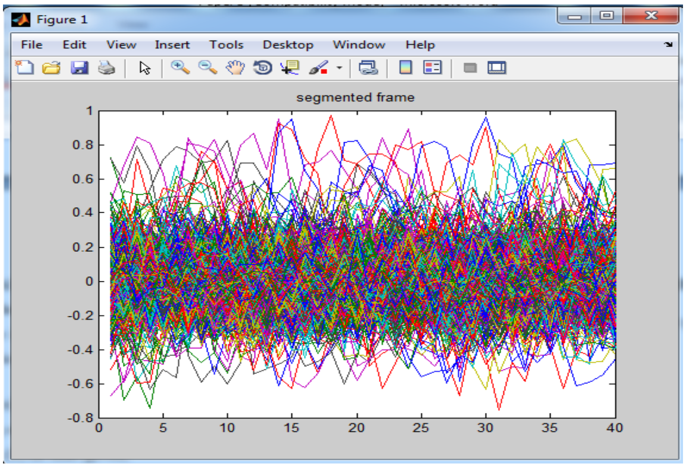

Fig 2: Overall Segmented Signals

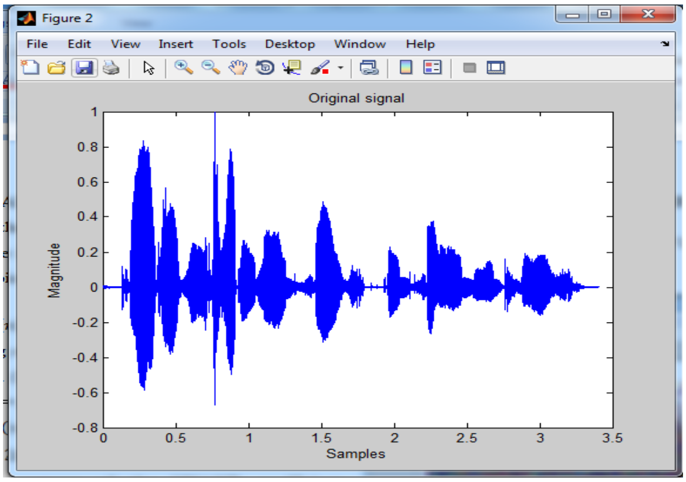

Fig 3: Original Input Speech Signal

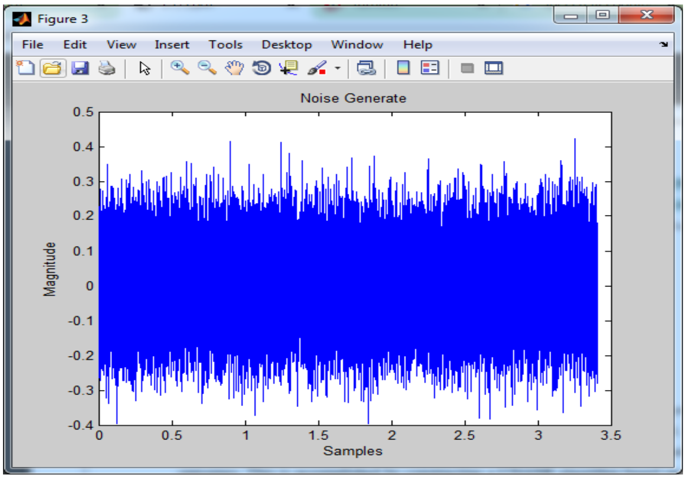

Fig 4: Noise Speech Signal

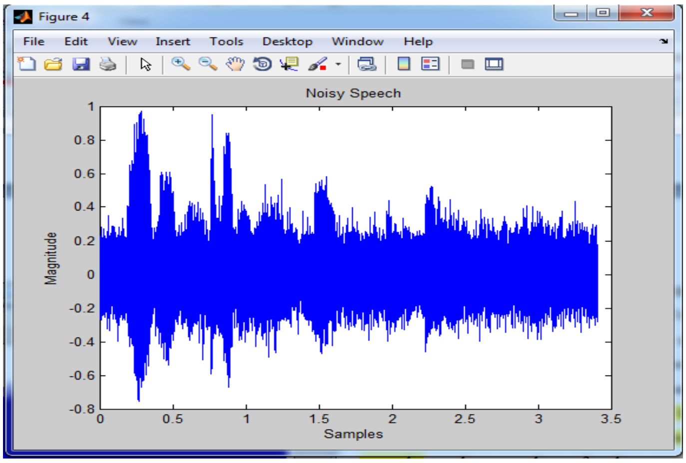

Fig 5: Noise Reduced Speech Signal

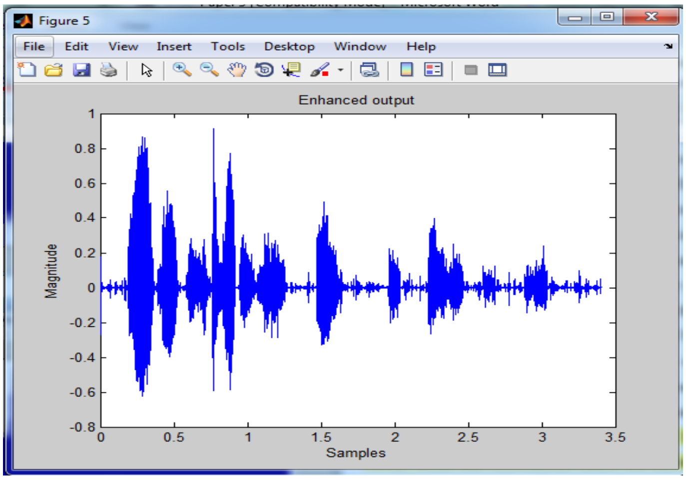

Fig 6: Enhanced Speech Signal

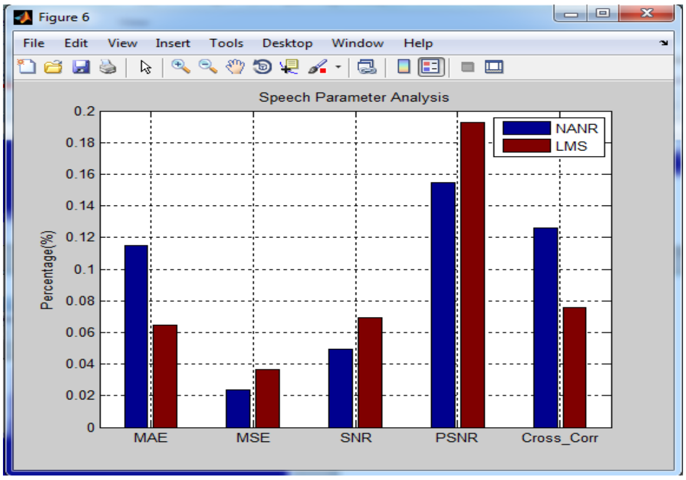

Fig 7: Speech Parameter Analysis

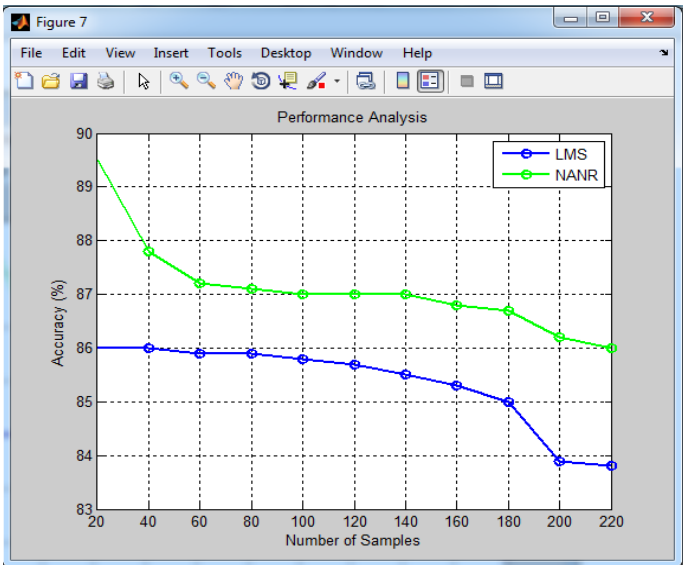

Fig 8: Accuracy Analysis

The following Table 1 represents Accuracy Comparison of LMS and UNANR algorithms.

Table 1 Accuracy Comparison Table

No of Samples	LMS Accuracy %	UNANR Accuracy %
20	86	89.4
40	86	87.8
60	85.9	87.2
80	85.9	87.1
100	85.7	87
120	85.5	87
140	85.5	87
160	85.3	86.9
180	85	86.8
200	84	86.1
220	83.9	86

V CONCLUSION

Speech recognition software automates human-computer interaction, making systems highly user-friendly. Background noise and echoes can induce speech to be distorted. As a result, experts have sought to improve the voice recognition technology in order to achieve better outcomes. This is accomplished by constructing a UNANR algorithm based on an adaptive filtering method in this case. Due to the fact that voice recognition is a pattern recognition issue, Naive Bayes classifiers are deployed for classification. The primary goal of a VRS is to achieve the highest level of recognition. In the final section of this study, the experimental findings achieved using the suggested algorithm is compared to those obtained using the well-known Least Mean Square (LMS) approach. The accuracy was found to be around 89%, and a number of other metrics were calculated to demonstrate the efficiency of the suggested system. Despite the fact that both methods provided good results, the proposed framework using the UNANR algorithm and the Naive Bayes classifier was able to recognize the speech signal effectively.

REFERENCES

[1] Kawase, Tomoko & Niwa, Kenta & Hioka, Yusuke & Kobayashi, Kazunori. (2017), “Automatic Parameter Switching of Noise Reduction for Speech Recognition”, Journal of Signal Processing. 21. 63-71. 10.2299/jsp.21.63.

[2] Kalamani M, Krishnamoorthi M, Valarmathi R S, "Modified Least Mean Square Adaptive Noise Reduction algorithm for Tamil Speech Signal under Noisy Environments" (2018), International Journal of Pure and Applied Mathematics. Volume 119 No. 16, 4461-4466, ISSN: 1314-3395.

[3] Haque M., Bhattacharyya K. (2018), "Speech Background Noise Removal Using Different Linear Filtering Techniques", Advanced Computational and Communication Paradigms. Lecture Notes in Electrical Engineering, vol 475. Springer, Singapore. https://doi.org/10.1007/978-981-10-8240-5_33.

[4] Haque, Minajul & Bhattacharyya, Kaustubh. (2019). "A study on different linear and non-linear filtering techniques of speech and speech recognition", ADBU Journal of Engineering Technology (AJET). 8. 1-13.

[5] Martinek, Radek, Vaňuš, Jan, Nedoma, Jan, Fridrich, Michael, Frnda, Jaroslav, Kawala-Sterniuk & Aleksandra. (2020), "Voice Communication in Noisy Environments in a Smart House Using Hybrid LMS+ICA Algorithm" Sensors. 20. 6022. 10.3390/s20216022.

[6] Mohammad, Zia Ur Rahman & Mohedden, Sk& Rao, Dr & Reddy, Y. & G.V.S, Karthik. (2011), "Filtering Non-Stationary Noise in Speech Signals using Computationally Efficient Unbiased and Normalized Algorithm" International Journal on Computer Science and Engineering.

[7] Lalit P. Patil, Prof. Amiteshwar Bhalavi, Prof. Rupesh Dubey and Mukesh Patidar, "Efficient algorithm for Speech Enhancement using Adaptive filter", International Journal of Electrical, Electronics and Computer Engineering 3(1): 98-103(2014),ISSN No.: 2277-2626.

[8] Mendiratta, Sunanda, Neelam Turk and D. Bansal. "A Robust Isolated Automatic Speech Recognition System using Machine Learning Techniques."(2019).International Journal of Innovative Technology and Exploring Engineering, ISSN: 2278-3075, Volume-8 Issue-10.