ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

Intelligent Volume Controller In Presence Of Background Noise

Ms. Aarti Neema[1], Mr. Umang Rastogi[2]

Ms. Aarti Neema Mtech Student EC Department SIMS, Indore

Mr. Umang Rastogi Assistant Pro. CS Department AIMT, Greater Noida

Abstract-**This paper proposes the use of intelligent volume controller for 2G, 2.5G, and 3G mobile phones based on fuzzy logic & its soft computing technique for improving voice quality in the presence of background noise. The IVC makes use of the noise level and class information generated by a system for fuzzy pattern classification of background noise.. At the highest level, FNC system contains two main modules: feature extraction and feature matching. For this feature extraction work, one hundred noise data samples were collected for each of the four types of audio data (internet) noises i.e. car, train, office and market noise from website www.partnersinrhyme.com on internet as well as real data (original) noises from the environment, with the help of a microphone connected to a personal computer and thus, a total of eight hundred noise samples were stored in the computer memory for noise analysis and simulation results in Matlab were recorded after implementation and compiled in this paper. Fuzzy feature matching involves the actual procedure to identify the unknown noise by comparing extracted features from noise input with the ones from a set of known noises in the database through fuzzy rules. These stored noise data samples were processed by extracting their features in the form of coefficients representing the noise parameter estimates through three models viz, Mel Frequency Cepstral Coefficient (MFCC) model, Linear Predictive Coding (LPC) model and Reflection Coefficient Efficiency Parameters (RCEP) model out of approximately twenty seven models. Two models i.e. LPC and RCEP were based on built-in and the third model MFCC on a user-defined program. Based on the coefficient outputs obtained through above three models; three background noise classifiers were designed using predictive modeling after training in neural network for all the eight hundred noise data samples and their performance was compared on the basis of their classification accuracies. Here, FLCs/ NNCs are applied to design an intelligent volume controller (IVC) for mobile phone. The fuzzy logic/ neural network based controller do not require any identification of mobile phone dynamics to control the volume and also assures the disturbance rejection with high robustness. Finally, intelligent volume controller (IVC) was implemented through Matlab program using active noise control & results were compared for internet noise data set and original noise data sets for above three models of noise parameter estimates independently. The output of three independent intelligent volume controllers was obtained based on the output of three independent noise classifiers.**

During the entire research, we have achieved classification accuracy up to 95% for noise classifiers using RCEP, LPC and MFCC based feature sets, respectively. Also, improvement in intelligent volume controller was achieved in terms of noise attenuation level of up to 0.05db, 0.012db and 0.017db. These results were further verified for 800 noise data samples of audio data (internet) noises as well as real data (original) noises from the environment for classifiers & controllers. It was found that the results of classifiers & controllers varied by at most 1% for the two different category of noises since noise parameter estimates varied by at most 1% only when internet noise samples were compared to those of original noise samples.

Keywords- **Neural Network Controller (NNC), Fuzzy logic controller (FLC), Intelligent volume controller (IVC), Mel frequency cepstral coefficients (MFCC), Linear predictive coefficients (LPC), Real cepstral parameter coefficients (RCEP).**

I. INTRODUCTION

Whenever we are having **a** conversation on a mobile phone, if the background noise level is high, we usually ask the speaker on the other end to speak up, or we may increase the volume. Also, during high background noise levels, users tend to bring their mobiles very close to their ears. Quality of service *(00s)* is improved by providing an intelligent volume controller that intelligently changes the volume level based on the background noise levels and classes. Background noise levels can he high while in busses, trains, planes, markets, sporting venues, and other public places. Several methods have been investigated for background noise classification. Background noise classification information is very useful and can he used to dynamically adapt acoustic volume levels to suit **a** particular type of noise, to improve the understandability of sounds in different noise environments, volume levels can he

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

adjusted automatically. Car, bus, and train noises fall into the low-frequency noise category. The spectrum of some classes of noise remain constant in time (stationary noise), whereas others vary suddenly (non stationary).

The fuzzy volume controller for mobile phones adjusts the volume according to the background noise level and noise class by a fuzzy system. The inputs to the fuzzy volume controller are the noise level derived from the voice activity detector **(VAD)** present within the speech codec, the current volume level, and the noise class derived from a system for fuzzy, pattern classification of background noise. By intelligently adjusting the volume level, the QoS is improved for both stationary and non stationary background noise in mobile environments. We refer to **a** mobile phone which uses an IVC in this article as a mobile phone (SCP):

Hearing loss in individuals can be gradual, and quality of hearing can vary from person to person. Hearing loss can also result in difficulty for individuals in understanding speech in the presence of background noise. Hence, the IVC needs to be personalized based on the requirements of an individual's hearing requirements.

II. INTELLIGENT VOLUME CONTROLLER

The volume control knob is ubiquitous and is as old as electronic systems. It is the most used control in any electronic system be it a television, CD player or a mobile phone. The manual volume control has several shortcomings. For a sophisticated mobile phone, the user needs to be trained to use volume control to get the best of the sound quality. The volume control needs to take into account the changes in the background noise levels; the user's hearing requirements, and several other factors. ***A volume controller which can intelligently adjust the volume levels is referred to as Intelligent Volume Controller (IVC).***

The IVCs have to be designed based on very complex models of hearing mechanism. Also the measurement of inputs required for IVC such as background noise level, loudness level etc are based on subjective measures. In this context of uncertainty and imprecision, the soft computing techniques become useful for IVC implementation. Based on the human like information processing model, the intelligent computing techniques lead to robust, High Machine IQ (HMIQ) and computationally inexpensive IVCs Several soft computing techniques have been explored for IVC implementation.

Intelligent volume control is for overcoming most of the problems faced by the user while using manual volume control. The IVC takes into account the background noise level noise classes the speaker profile the hearing profile of the user and many more factors while adjusting the volume levels. The IVC also will have learning and generalization capability from the usage patterns. Intelligent volume control has several advantages such as:

a) The user need not have to adjust the volume levels repeatedly.
b) The volume control is transparent and thus the mobile phone with IVC delivers an improved speech quality and service.
c) The IVC takes into account the hearing requirements of the user and hence is able to perform better than a normal mobile phone.
d) The user's preferences for noise class and noise levels may be encoded into the IVC.

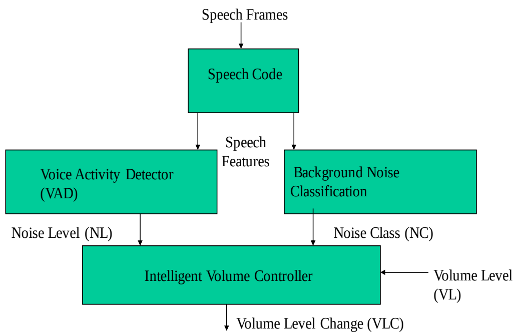
The intelligent volume controller design may have several benefits over the conventional controllers in cellular phones. There may be improved quality of speech for stationary and non-stationary noise in mobile environments, as the controller uses information on background noise level and classes to adjust the volume level. Noise classes like car noise fall into low-frequency noise. They do not affect the intelligibility of speech so much as factory noise. Therefore, the use of background noise class information can provide effective volume adjustment.

The study in this paper includes a comparison between the performance of noise classifiers based on classification accuracy & performance of intelligent volume controllers based on noise attenuation level to ascertain their applications for noise data classification. In certain situations and for some datasets, the established techniques do not yield very good classification accuracy. Depending upon the characteristics of the data; the quest for a suitable model is a pertinent challenge. This research work focuses to explore models for best noise parameter estimate to design an efficient noise classifier for intelligent volume controller through active noise control technique as a variety of real world problems.

Rate of volume control change is also an important parameter; too fast a change can interfere with intelligibility itself. Hence, a comfortable volume change control rate is to be adapted based on mobile phone user surveys. The block diagram of the IVC is as shown Fig

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

BLOCK DIAGRAM OF INTELLIGENT VOLUME CONTROL IN PRESENCE OF BACKGROUND NOISE



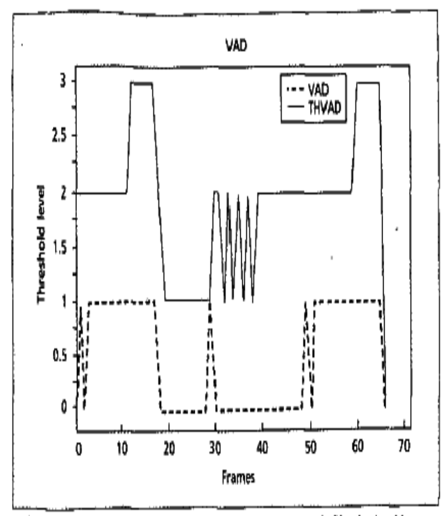
III. SOFT COMPUTING FOR VOLUME CONTROL

Soft computing techniques may be used to approach problems with human information processing model. The soft computing technique comprises of fuzzy logic, neural network, evolutionary-computing, machine learning and probabilistic-chaotic computing. The intelligent computing control algorithms which are well suited for control are:

- **a) Fuzzy Control**: This controller uses Fuzzy logic concepts. The Fuzzy rules are defined by an expert and can be easily interpreted. The Fuzzy rule base can be personalized by using the audiogram as an input. The fuzzy logic controller requirements and hence well suited for the real time embedded applications.
- **b) Neural Control**: Neural networks are mathematical models of highly interconnected neurons. These networks can be trained to recognize a pattern in the data. A neural network can be used for IVC. The neural network is trained by set of sample data and works as a black box controller.
- **c) Neuro-Fuzzy Control**: This controller uses hybrid of neural networks and fuzzy logic. The neural network is able to learn from training data while the fuzzy logic captures the expert knowledge to provide a high performance controller.
- **d) Genetic-Fuzzy Control:** This is a hybrid controller which use genetic and fuzzy logic algorithm. The genetic algorithm is used for the i/o scaling, membership functions and rule base of the fuzzy controller.

IV. MEASURING BACKGROUND NOISE LEVELS

Several techniques have been proposed for measuring the noise levels. The technique proposed for IVC is to use the VAD embedded in second-generation ***(2G), 2.5G,*** and ***3G*** mobile phones. An adaptive noise suppressor filter is used to filter the input signal frame. The coefficient of this filter is computed during noise-only periods determined by special measures taken to identify noise-only frames. These include signal stationary and periodicity measures. Several algorithms have been proposed to improve the performance of Global System for Mobile Communications (GSM) VAD for stationary and non stationary noises. The threshold value computed by the VAD can be use as crisp noise level input for the fuzzy controller. A fuzzy set for noise level (NL) is {Very Low, Medium Low, Low, Zero, High, Medium High, very High). Figure **1** shows the variation of threshold values for a test speech file for different speech frames (the x-axis indicates the number of speech frames).



Voice activity detection (also known as speech activity detection or, more simply, speech detection) is an algorithm used in speech processing wherein the presence or absence of human speech is detected in regions of audio. The main uses of VAD are in speech coding and speech recognition. Voice activity detection (VAD) may not only indicate the presence or absence of speech, but also other aspects of the speech, for example whether the speech is voiced, unvoiced or sustained. Voice activity detection is usually language independent.

While talking to someone, there will be silent periods when we are not talking. A VAD feature in VOIP can disable the silence packets and use the silent period to transmit some traffic other than voice.

The process of separating conversational speech from silence, music, noise or other non-speech signals is called voice activity detection (VAD). The primary

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

function of a voice activity detector is to provide an indication of the presence of speech in order to facilitate speech processing as well as possibly providing delimiters for the beginning and end of a speech segment. It was first investigated for use on time-assignment speech interpolation (TASI) systems. VAD is an important enabling technology for a variety of speech-based applications. For these purposes there have been proposed various VAD algorithms that trade off delay, sensitivity, accuracy and computational cost.

V. NOISECLASSIFICATION METHODOLY

The methodology adopted is based on a fuzzy approach in which the matching phase is performed using a set of fuzzy rules. The rule base may be tuned using evolutionary techniques. During the period of conversation, the surrounding noise level may vary, so the SVT adjusts the volume. Hence, for a user of a mobile phone, QoS is improved as volume control is transparent. The fuzzy system was automatically extracted from examples. After every 10-ms of waveform, the feature extraction module computes the set of parameters adopted. Out of 27 noise parameter available for exploring environmental noise classification, ***we have explored and analyzed two main parameters LPC, MFCC and one allied parameter RCEP only in this paper***.

VI. NOISE DATABASE SOURCE

Under mentioned noises have been collected from following sources for the formation of noise database-

A) Signals (Internet)

(a) s1 (Human voice) 5 samples, (b) s2 (Car noise) 5 samples, (c) s3 (Office noise) 5 samples

(d) s4 (Market noise) 5 samples, (e) s5 (Train noise) 5 samples

B) Signals (Original)

(a) s1(Human voice) 5 samples, (b) s2(Car noise) 5 samples, (c) s3 (Office noise) 5 samples

(d) s4 (Market noise) 5 samples, (e) s5 (Train noise) 5 samples

In this work, we used a database containing a large number of different types of environmental noise which can be explored on basis of environmental ***background noise classes*** as follows:

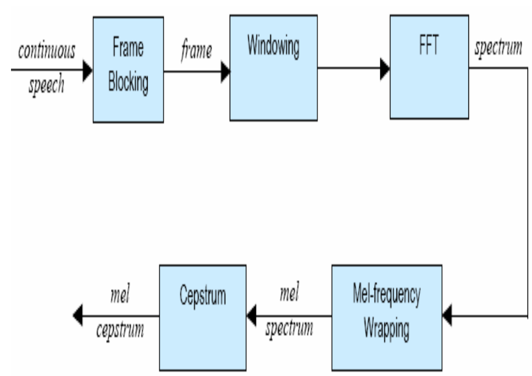
a) Automobiles noise class (ANC): Cars, trucks, buses, trains, ambulance, police cars etc
b) Babble noise class (BNC): Cafeteria, sports, stadium, office etc
c) Factory noise class (FNC): Tools such as drilling machines, power hammer etc.
d) Street noise class (SNC):
e) Shopping mall, market, busy street, bus station, gas station etc.
f) Miscellaneous noise class (MNC): Aircraft noise, thunder storm etc

From these background noise classes, four noise classes have been considered in this paper: 1) Automobiles (car & train); 2) babble (office); 3) street (market); 4) biotic (human).

VII. FEATURE EXTRACTION

a) MFCC MODEL

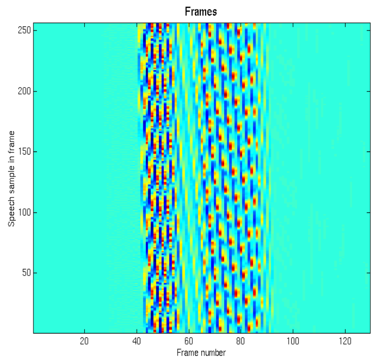
The extraction of the best parametric representation of acoustic signals is an important task to produce a better recognition performance. The efficiency of this phase is important for the next phase since it affects its behavior. MFCC is based on human hearing perceptions which cannot perceive frequencies over 1Khz. In other words, in MFCC is based on known variation of the human ear's critical bandwidth with frequency. MFCC has two types of filter which are spaced linearly at low frequency below 500 Hz and logarithmic spacing above 500Hz. A subjective pitch is present on Mel Frequency Scale to capture important characteristic of phonetic in speech. The overall process of the MFCC is shown in Figure



A) FRAME BLOCKING

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

Framing is the first applied to the speech signal of the speaker. The signal is partitioned or blocked into N segments (frames).



B) WINDOWING

The next step in the processing is to window each individual frame so as to minimize the signal discontinuities at the beginning and end of each frame. Hamming window is used as window shape by considering the next block in feature extraction processing chain and integrates all the closest frequency lines. The Hamming window equation is given as: If the window is defined as W (n), 0 ≤ n ≤ N-1 where N = number of samples in each frame, then the result of windowing is the signal

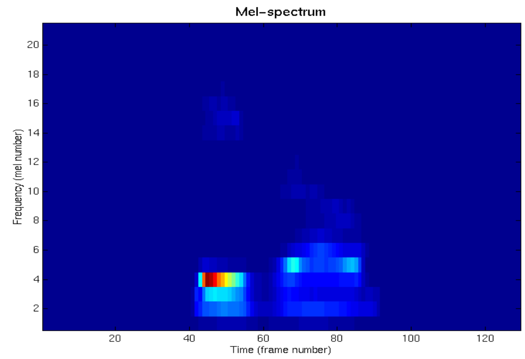
$$y_l(n) = x_l(n)w(n) \quad 0 \le n \le N\text{-}1$$

$y_l(n)$ = Output signal
$x_l(n)$ = input signal
w (n) = Hamming window,

Typically the *Hamming* window is used, and then the result of windowing signal is shown below:

$$W(n) = 0.54\text{-}0.46\cos\left(\frac{2\pi n}{N-1}\right)$$
$$0 \le n \le N\text{-}1$$

Use of the window function reduces the frequency resolution by 40%, so the frames must overlap to permit tracing and continuity of the signal. The motive for utilizing the windowing function is to smooth the edges of each frame to reduce discontinuities or abrupt changes at the endpoints. The windowing serves a second purpose and that is the reduction of the spectral distortion that arises from the windowing itself.



C) FAST FOURIER TRANSFORM

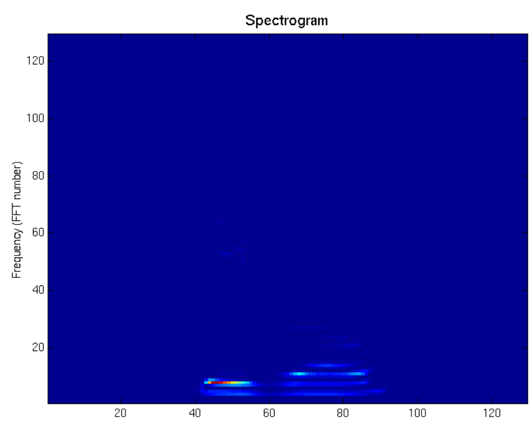
Next step is the Fast Fourier Transform which converts each frame of N samples in time domain to frequency domain. To convert each frame of N samples from time domain into frequency domain The Fourier Transform is to convert the convolution of the glottal pulse U[n] and the vocal tract impulse response H[n] in the time domain. This statement supports the equation below:

$$Y(w) = FFT[h(t)*X(t)] = H(w)*W(n)$$

If X (w), H (w) and Y (w) are the Fourier Transform of X (t), H (t) and Y (t) respectively.

FFT is used for doing conversion from the spatial domain to the frequency domain. Each frame having N samples is converted into frequency domain. Fourier transformation is a fast algorithm to apply Discrete Fourier Transform (DFT), on the given set of Nm samples shown below:

$$D_k = \sum_{n=0}^{N-1} D_n e^{\frac{-j2\pi kn}{N}}$$

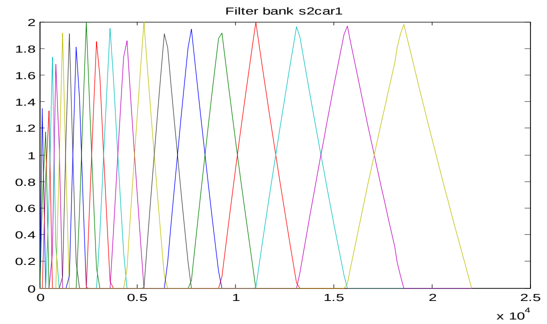


D) MEL-FREQUENCY WRAPPING

The spectrum obtained from the above step is Mel Frequency Wrapped; the major work done in this process is to convert the frequency spectrum to Mel spectrum. The process of obtaining Mel-cepstral coefficients involves the use of a Mel scale filter bank. The spectral

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

coefficients of each frame are then converted to Mel scale after applying a filter bank. The Mel-scale is a logarithmic scale resembling the way that the human ear perceivessound. The filter bank is composed of triangular filters that are equally spaced on a logarithmic



MEL-SPACED FILTER BANK (K =20) PLOT FOR ORIGINAL CAR NOISE IN MATLAB

scale. After that the following equation is used to compute the Mel for given frequency f in HZ:

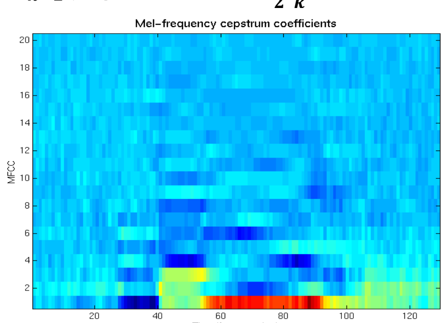
$$m_f = 2595 \log_5 \left[\frac{f}{700}+1\right]$$

Thus, with the help of Filter bank with proper spacing done by Mel scaling it becomes easy to get the estimation about the energies at each spot and once this energies are estimated then the log of these energies also known as Mel spectrum can be used for calculating first 13 coefficients using DCT. Hence, first 13 coefficients are calculated using DCT and higher are discarded.

E) CEPSTRUM

In this final step, we convert the log Mel spectrum back to time. The result is called the Mel frequency Cepstrum coefficients (MFCC). We can calculate what is called the mel-frequency Cepstrum, Cn,

$$C_n = \sum_{k-1}^{k} (\log D_k) \cos\left[m\left(k-\frac{1}{2}\right)\frac{\pi}{k}\right]$$



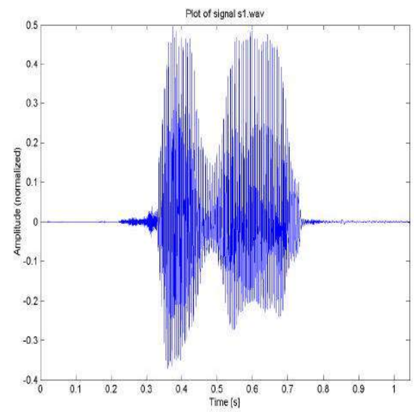
Where m = 0, 1… k- 1

Where Cn represents the MFCC and m is the number of the coefficients here m=13 so, total number of coefficients extracted from each frame is 13.

In the last step, the log of this spectrum is processed in order to get the cepstral coefficients by doing DCT (discrete Cosine Transform). For implementation, in this paper only first 13 coefficients are taken.

F) PROCESS VISUALIZATION

In this section, we will visualize the results obtained from some of the main parts of the feature extraction process, having recently viewed the details of the procedure. Following figure will serve as the audio signal intended for the analysis in this case.



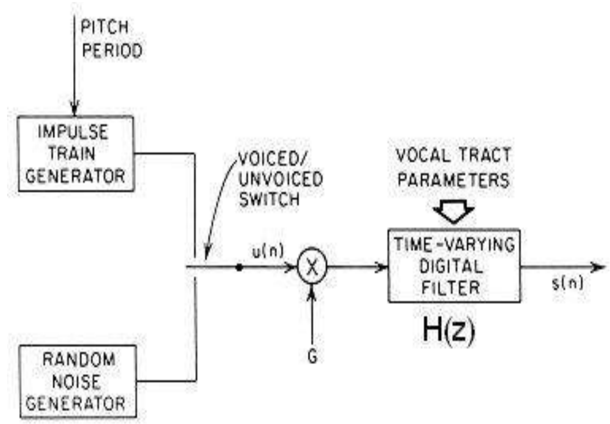
DIGITAL AUDIO SIGNAL

This audio sample represents the audio signal recorded. It will serve as the input of the feature extraction in order to visualize the results of the processing involved. The next step is to frame the audio sample into portions of a predetermined size. This is done to process the frames taking advantage of the quasi-stationary property of the frames, given a properly selected frame size.

b) LPC Model

The choice of signal features is usually based on previous knowledge of the nature of the signals to be analyzed. Noise synthesis based on LPC model is comparable to vocal tract of human throat.

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014



LPC MODEL IN HUMAN THROAT

The object of linear prediction is to form a model of a Linear Time Invariant (LTI) digital system through observation of input and output sequences. The basic idea behind linear prediction is that a noise sample can be approximated as a linear combination of past noise samples. By minimizing the sum of the squared differences (over a finite interval) between the actual noise samples and the linearly predicted ones, a unique set of predictor coefficients can be determined.

If u (n) is a normalized excitation source and being scaled by 'G', the gain of the excitation source, then LPC model is the most common form of spectral analysis models on blocks of noise (noise frames) and is constrained to be of the following form, where H (z) is a pth order polynomial with z-transform and the coefficients a_1, a_2, …, a_p are assumed to be constant over the noise analysis frame

$$H(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \ldots + a_p z^{-p}$$

Here the order 'p' is called the LPC order.

If 'N' is the number of samples per frame and 'M ' is the distance between the beginnings of two frame, then for a given noise sample at time 'n'; S(n), can be approximated as a linear combination of the past 'p' noise samples, such that

$$s(n) = a_1 s(n-1) + a_2 s(n-2) + \ldots + a_p s(n) \quad (1)$$

Where the coefficients a_1, a_2… a_p are assumed constant over the noise analysis frame. We convert eq. (1) to an equality by including an excitation, G u (n), giving

$$s(n) = \sum a_i s(n-i) + Gu(n), \quad 1 \le i \le p \quad (2)$$

Where u (n) is a normalized excitation and G is the gain of the excitation. By expressing eq (2) in the z-domain, we get the relation as follows in (3)

$$S(z) = \sum a_i z^{-i} S(z) + G\,U(z), \; 1 \le i \le p \qquad (3)$$

Leading to the transfer function as given in (4)

$$H(Z) = \frac{S(Z)}{G\,U(Z)} = \frac{1}{P} = \frac{1}{H(Z)}$$

Because noise signals vary with time, this process is done on short chunks of the noise signal, which are called frames. Usually 30 to 50 frames per second give intelligible noise with good compression. When applying LPC to audio at high sampling rates, it is important to carry out some kind of auditory frequency warping, such as according to mel or Bank frequency scales.

c) RCEP MODEL

As per theoretical point of view, the Cepstreal logarithm of the magnitude of few cepstral coefficients and setting the remaining coefficients to zero, it is possible to smooth the harmonic structure of the spectrum. Cepstral coefficients are therefore very convenient coefficients to represent the speech spectral envelope. Hence, the following function calculates the real Cepstrum of the signal x.

$$y = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log \left|X(e^{j\omega t}|\right) e^{-j\omega t}\, d\omega$$

This denotes the Fourier Transform used for the separation of two signals convolved with each other based technique for determining a Harmonics valid technique for determining the sensitive to both noise and jitter for a large part of the noise or jitter. Thus real Cepstrum block gives the real Cepstrum way to define the prediction filter. Last, the line spectrum frequencies (a.k.a. line spectrum pairs) are also frequently used in speech representation derived from linear predictive analysis. Cepstrum is defined as the inverse Fourier transform of the Fourier transform. Therefore, by keeping only the first of x and hence real Cepstrum as a real-valued function can be. Thus, RCEP is a C Harmonics-to-Noise Ratio (HNR) in Speech Signals and amount of spectral noise, because it is almost linearly continuum output of the input frame and is coding. Line spectrum frequencies are another analysis which is very popular in speech coding.

VIII. RESULT-ANALYSIS

RESULTS OBTAINED IN MATLAB FOR FIVE SAMPLES OF FOUR INTERNET NOISES

A) MFCC

SAMPLES	CAR	OFFICE	MARKET	TRAIN
S1	0.7606	1.1829	0.8646	0.0271
S2	0.8497	0.2051	0.4135	0.5599
S3	0.1915	0.6141	0.8271	0.1922
S4	0.5952	0.7134	0.0903	0.9966
S5	0.6787	0.2297	0.1616	0.9129

B) LPC

SAMPLES	CAR	OFFICE	MARKET	TRAIN
S1	0.2164	0.5474	0.1504	0.6579
S2	0.1270	0.5195	0.1527	0.6629
S3	0.2298	0.2179	0.1558	0.7030
S4	0.0988	0.1775	0.1181	0.6006
S5	0.1835	0.2018	0.1645	0.6627

C) RCEP

SAMPLES	CAR	OFFICE	MARKET	TRAIN
S1	0.0011	0.0007	0.0006	0.0012
S2	0.0009	0.0010	0.0003	0.0005
S3	0.0003	0.0009	0.0001	0.0013
S4	0.0004	0.0001	0.0007	0.0008
S5	0.0000	0.0000	0.0002	0.0017

D) AVERAGE OF COEFFICIENT

MFCC COEEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MAR KET (S1-S5)	TRA IN (S1-S5)
C1	16.613	0.978	2.040	1.572
C2	0.978	1.074	1.302	1.154
C3	2.040	1.787	1.748	1.687
C4	1.572	1.397	1.377	1.437
C5	2.101	1.331	1.718	1.594

MFCC COEEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MARKE T (S1-S5)	TRA IN (S1-S5)
C1	16.723	19.669	19.297	18.984
C2	0.981	1.151	1.412	1.244
C3	2.170	1.827	1.751	1.696
C4	1.586	1.399	1.146	1.517
C5	2.241	1.411	1.729	1.599

LPC COEEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MAR KET (S1-S5)	TRA IN (S1-S5)
C1	1.000	1.000	1.000	1.000
C2	0.497	0.590	0.298	1.432
C3	0.546	0.430	0.553	0.005
C4	0.281	0.172	0.619	0.769
C5	0.543	0.099	0.209	0.731

RECP COEEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MAR KET (S1-S5)	TRA IN (S1-S5)
C1	7.997	8.522	8.808	8.433
C2	0.004	0.000	0.799	0.001
C3	0.004	0.005	0.000	0.003
C4	0.004	0.002	0.002	0.003
C5	0.002	0.006	0.004	0.001

RESULTS OBTAINED IN MATLAB FOR FIVE SAMPLES OF FOUR ORIGINAL NOISES

A) MFCC

SAMPLES	CAR	OFFICE	MARKET	TRAIN
S1	0.7503	1.1929	0.8749	0.0219
S2	0.8399	0.2091	0.4179	0.5629
S3	0.1924	0.6541	0.8418	0.1969
S4	0.5953	0.7234	0.0916	0.9993
S5	0.6887	0.2397	0.1713	0.9279

B) LPC

SAMPLES	CAR	OFFICE	MARKE T	TRAIN
S1	0.2267	0.5584	0.1518	0.6749
S2	0.1287	0.5185	0.1687	0.6636
S3	0.2399	0.2279	0.1564	0.7180
S4	0.0998	0.1785	0.1219	0.6017
S5	0.1848	0.2188	0.1658	0.6377

C) RECP

SAMPLES	CAR	OFFICE	MARKE T	TRAIN
S1	0.0181	0.0015	0.0146	0.0029
S2	0.0012	0.0130	0.0017	0.0175
S3	0.0133	0.0017	0.0191	0.0025
S4	0.0012	0.0191	0.0014	0.0168
S5	0.0120	0.0063	0.0115	0.0029

D) AVERAGE OF COEEFICIENT

LPC COEEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MARK ET (S1-S5)	TRA IN (S1-S5)
C1	1.000	1.000	1.000	1.000
C2	0.499	0.596	0.388	1.572
C3	0.636	0.520	0.561	0.016

ijournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 4, April 2014

C4	0.294	0.183	0.729	0.839
C5	0.613	0.119	0.218	0.881

RECP CO EEFI CIENT	CAR (S1-S5)	OFF ICE (S1-S5)	MARKET (S1-S5)	TRAIN (S1-S5)
C1	7.998	8.632	8.814	8.513
C2	0.114	0.005	0.829	0.018
C3	0.0013	0.0185	0.0017	0.143
C4	0.104	0.016	0.182	0.016
C5	0.0013	0.146	0.011	0.121

IX. CONCLUSION

From simulation results in Matlab, it was found that volume controller performed better for MFCC model as compared to other two models (LPC & RCEP) of original noise database as compared to internet noise database. This was due to the fact that classification accuracy (based on classification confusion matrix) was the highest for MFCC model parameter estimate viz 95% as compared to other two models (LPC & RCEP) viz. 94.05% as found earlier. It has been observed that the classification accuracy is 1% higher for Original noise dataset w.r.t. Internet noise dataset in classifier as well as Intelligent Volume Controller.[3]

Finally, intelligent volume controller was implemented using active noise control & volume controller results were compared for internet noise data set and original noise data sets for above three models of noise parameter estimates in dependently. We have achieved classification accuracies of upto 93.01%, 94.05% and 95.00% using RCEP, LPC and MFCC based feature sets, respectively. Also, improvement in intelligent volume controller was achieved in terms of noise attenuation level of upto 0.05db, 0.012db and 0.017 db using RCEP, LPC and MFCC based feature sets, respectively. Thus, the highest classification accuracy upto 95.00% in noise classifier and maximum improvement in terms of noise attenuation level upto 0.017 db in intelligent volume controller was attained using MFCC based feature set through active noise control.

The real challenge for designing Intelligent Volume Controller (IVC) is generally based on background noise classification accuracy which is hard to achieve since it deals with the variability of the environments and channels from where the noise was obtained However, under mismatched conditions and noisy environments, often expected in real-world conditions, the performance of classification systems degrades significantly, far away from the satisfactory level and hence effects the performance of further stages. Therefore, robustness becomes a crucial research issue in noise classification field to improve the performance of an Intelligent Volume Controller.

REFERENCES

1) Schafer, R. and Rabiner, L.Digital Representation of Speech Signals.. Proceedings of the IEEE 63 (1975): 662-677.
2) Gray, R.M.Vector Quantization.. IEEE ASSP Magazine 1 (1984): 4-29.
3) Schafer, R. and Rabiner, L Systems for Automatic Formant Analysis of Voiced Speech..Journal of the Acoustical Society of America 47 (1970): 634-648.
4) Tokhura, Y. A weighted cepstral distance measure for speech recognition..IEEE Transactions on acoustics, speech and signal processing 35 (1987): 1414-1422.
5) Fujimura, O.Analysis of nasal consonants.. Journal of the Acoustical Society of America 34 (1962): 1865- 1875.
6) Hughes, G. and Halle, M. .Acoustic Properties of Stop Consonants.. Journal of the Acoustical Society of America 30 (1957): 07-116.
7) Atal, B.S. Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. Journal of the Acoustical Society of America 55 (1974): 1304-1312.
8) Furui, Sadaoki. Digital Speech Processing, Synthesis, and Recognition. New York: Marcel Dekker, 2001