International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

A TECHNICAL REVIEW ON ESTIMATION OF NOISE PARAMETERS

Pankaj Kushwaha[1], Siddharth Ratna[2] ,Vikramjeet Singh[3]

[1, 2]Students, Electrical Engineering Department Greater Noida Institutes of Technology, Gr.Noida, (India)

[3] Assistant Professor, Electrical Engineering Department Greater Noida Institutes of Technology, Gr. Noida, (India)

ABSTRACT

In this paper, authors investigate estimation of noise parameters which can be used for background noise classification and mainly presents a criterion to group a large range of noise into a reduced set of classes of noise with similar speech characteristics along with a set of robust acoustic parameters. Background noise falls under two main categories, which are environmental noise and mechanical noise [1].Fortynoise samples were downloaded through internet from website www.partnersinrhyme.comwith the help of a microphone connected to personal computer & stored in memory as a noise database. User defined program was written in MATLAB for Mel Frequency Cepstral coefficient (MFCC) while built-in programs for Linear Predictive Coding (LPC), Real Cepstral Parameter (RCEP) and power spectrum have been explored in MATLAB to estimate speech parameters which may be utilized for speech analysis through any one of the soft computing techniquesviz. neural networks, fuzzy logic, genetic algorithms or a combination of these. Ten samples, each of four commonly encountered noises (Office, Fight scene, Crowd & City) i.e. 40 noises in total have been considered in our study for estimation of three coefficients viz. Mel Frequency Cepstral coefficient, Linear predictive coding and Real Cepstral Parameter. Our experimental results show that Mel Frequency Cepstral Frequencies are robust features in noise parameter estimation. 27 filter banks were used and filter bank output along with power spectrum was obtained in MATLAB. By trial & error method, it was found that the best result was obtained at maximum difference of 2.0416, 10.2613, 5.1617 & 7.4529 when average of highest & second highest MFCC coefficients was taken of Office, Fight scene, Crowd & City noises respectively.

Index Terms: Mel Frequency Cepstral Coefficient (MFCC), Linear PredictiveCoding (LPC), Real Cepstral Parameter (RCEP).

I. INTRODUCTION

Since over two decades, several techniques and algorithms have been proposed by many researchers to classify environmental noise using parameters such as line spectral frequency (LSF), log area ratio (LAR) coefficients, zero crossing rate (ZCR) and power spectral density (PSD). But none of the previously developed techniques have proven to be highly effective because of their own inherent limitations associated with each technique. Man-machine interaction has increased the demand for advanced speech processing algorithms capable of providing good performance levels [2].Recently, different research groups have carried out studies on new methods and algorithms for environmental noise classification; but in this paper, authors have tried to explore parameter estimation for speech analysis. MFCC estimation is an important parameter, put forward by Davis and Mermelstein, describes energy distribution of speech signal in a frequency field. A number of studies best support MFCCs and it produces good results in most of the situations. Integration of phase information with MFCC may be quite useful in speech synthesis [3]. Emotions may also be incorporated since in real world situation, human beings seldom communicate with neutral speech. In our daily life, we encounter different types and levels of environmental acoustical noises like traffic noise, car noise, office noise etc. In various speech-processing systems such as speech coding, speech recognition and

speaker verification, the unwanted noise signals are picked up along with the speech signals which often cause degradation in the performance of communication systems. By modifying the processing according to the type of background noise, the performance can be enhanced. This requires noise classification based on estimation and characterization of speech parameters. Environmental noise classifier can be used in various fields as, speech recognition and coding being the main ones. The acoustic features can be adapted to the type of environmental noise by choosing the most appropriate set to ensure reparability between phonetic classes. As low cost DSP's are increasingly becoming popular, therefore, the next generation of speech coders and intelligent volume controllers is likely to include classification module in order to improve robustness to environmental noise.

II. NOISE CLASSIFICATION METHODOLOGY

Two of the most promising techniques found for audio classification are Artificial Neural Networks (ANN) and Hidden Markov Models (HMM)[4]. The methodology that can be adopted for environmental noise classification through parameter estimation is based on exploring any one or a few of the environmental noise parameters viz .Linear Predictive Coding, Mel-cepstral based parameters, Real Cepstrum based parameters, line spectral frequencies coefficients, log area ratio coefficients, zero crossing rate and power spectral density. From these noise parameters, we have explored and analyzed two main parameters linear predictive coding, Mel frequency cepstral coefficients and one allied parameter i.e. real cepstrum parameter in this paper. Noise database created can be explored on basis of noise classes as follows:

i. Automobiles noise class (ANC): Cars, trucks, buses, trains, ambulance, police cars etc.
ii. Babble noise class (BNC) : Cafeteria, sports, stadium, office etc.
iii. Factory noise class (FNC): Tools such as drilling machines, power hammer etc.
iv. Street noise class (SNC):Shopping mall, market, fight, busy street, gas station etc.
v. Miscellaneous noise class (MNC): Aircraft noise, thunder storm etc.

Out of these noise classes, only three noise classes have been considered viz. office noise from babble noise class (BNC) and city, fight and crowd noise from street noise class (SNC).

III. ANALYSIS OF NOISE PARAMETERS

Speech parameters have been analyzed by acoustic-phonetic approach after spectral analysis. The first step in speech processing is feature measurement which provides an appropriate spectral representation of the characteristics of the time-varying speech signal by filter bank method implemented in MATLAB. Signal representation of downloaded office (briefcase) noise is as follows:

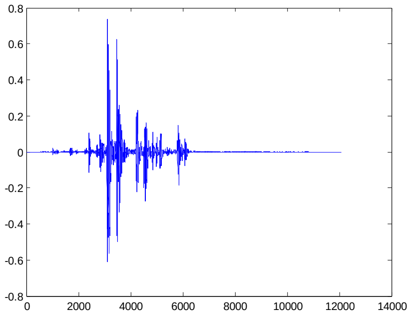


Figure 1(a): Office noise signal representation in MATLAB.

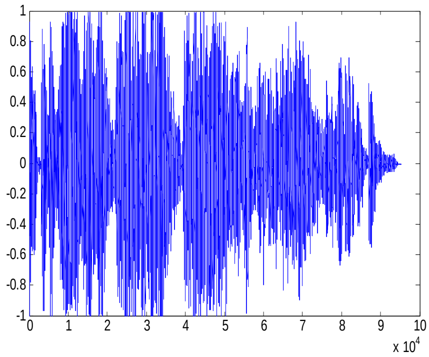


Figure 1(b): Fight noise signal representation in MATLAB

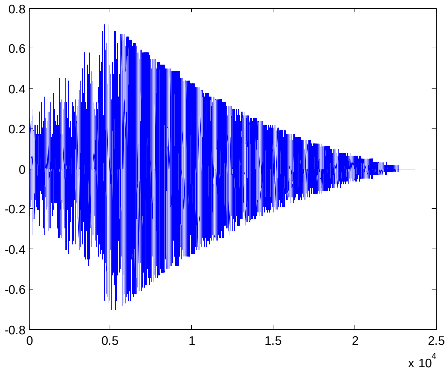


Figure 1(c): Crowd noise signal representation in MATLAB

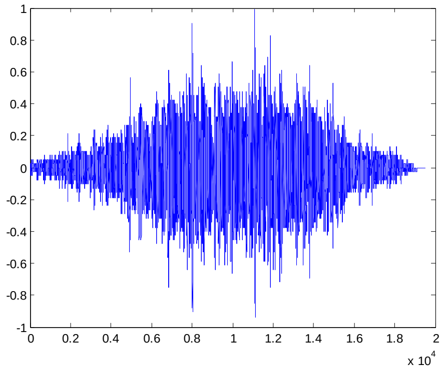


Figure 1(d): City noise signal representation in MATLAB

Similarly, signal representation of other noises has been recorded. The most common type of filter bank used for speech analysis is the uniform filter bank for which the center frequency, fi, of the i^{th} band pass filter is defined as

Fi = Fs i, 1 < i < Q, N

where Fs is the sampling rate of the speech signal, and N is the number of uniformly spaced filters required to span the frequency range of the speech. The actual number of filters used in the filter bank, Q, of our work satisfies the relation

Q < N / 2 < 54/2 < 27

with equality meaning that there is no frequency overlap between adjacent filter channels, and with inequality meaning that adjacent filter channels overlap. (If bi < Fs/ N, then certain portions of the speech spectrum would be missing from the analysis and the resulting speech spectrum would not be considered very meaningful).The digital speech signal, s(n), was passed through a bank of 27 band pass filters whose coverage spans the frequency range of interest in the signal (e.g., 100-3000 Hz for

telephone-quality signals, 100-8000 Hz for broadband signals) & output in MATLA Bis as follows-

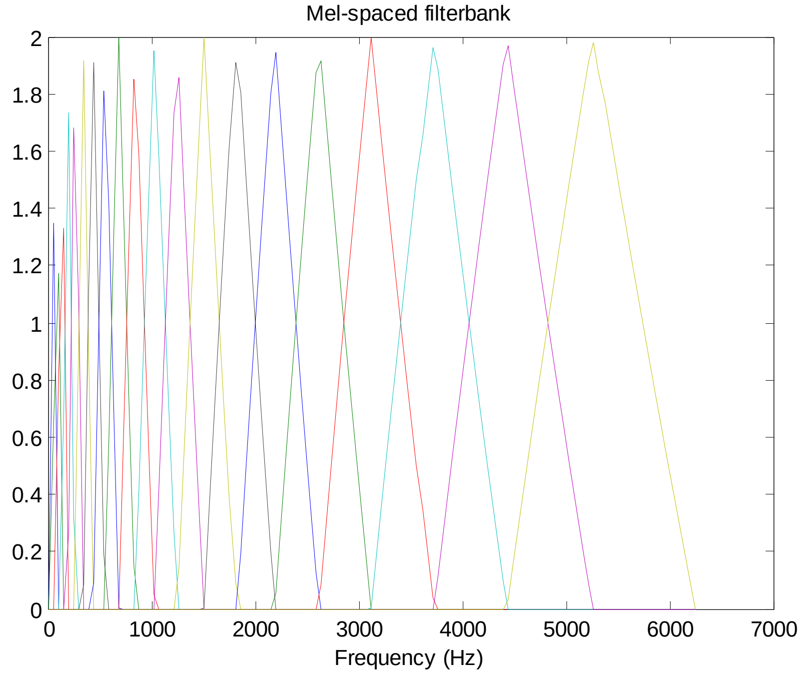
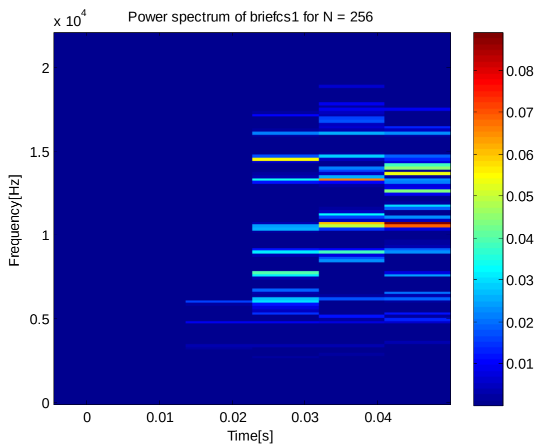


Figure 2: Filter-bank output of briefcase noise in MATLAB.

Similarly, filter bank outputs were obtained for other noises. Power spectrum output of all noises were obtained in MATLAB and that of briefcase noise obtained is as follows-



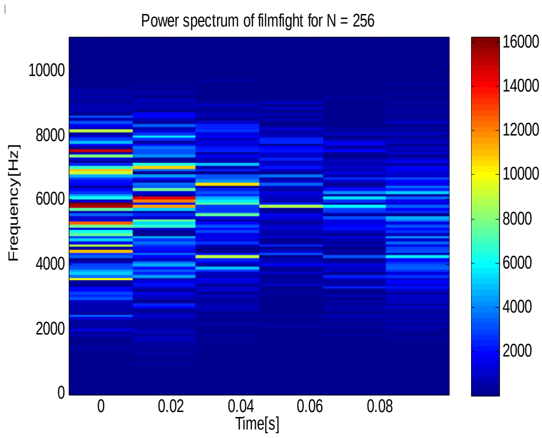


Figure 3(a): Power spectrum output of Office noise in MATLAB.

Figure 3(b): Power spectrum output of fight noise in MATLAB

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE ISSN 2321 - 2055

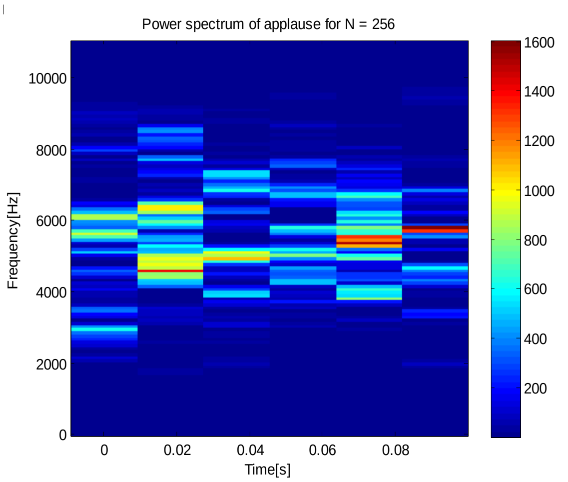


Figure 3(c): Power spectrum output of Crowd noise in MATLAB

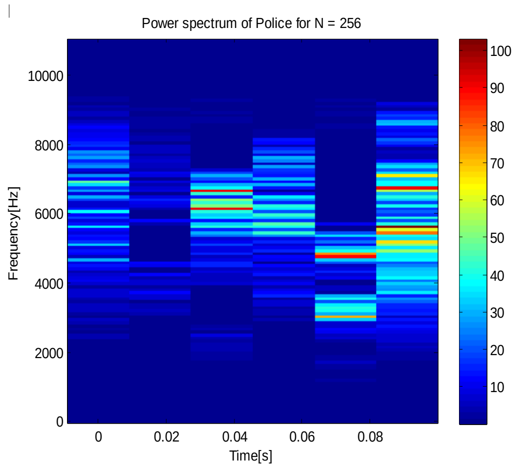


Figure 3(d): Power spectrum output of City noise in MATLAB

IV. SPECTRAL MODELS USED FOR ENVIRONMENTAL NOISE CLASSIFICATION

The effectiveness of models and its authenticity depends on to what degree the modelis adequate and how precisely corresponds to it. A neuro-fuzzy computing maybe used to provide system identification and interpretability of models. Following models have been for environmental noise classification.

V. LPC MODEL

The choice of signal features is usually based on previous knowledge of the nature of the signals to be analyzed . Speech synthesis based on LPC model in vocal tract of human throat may be assumed as follows in figure 4

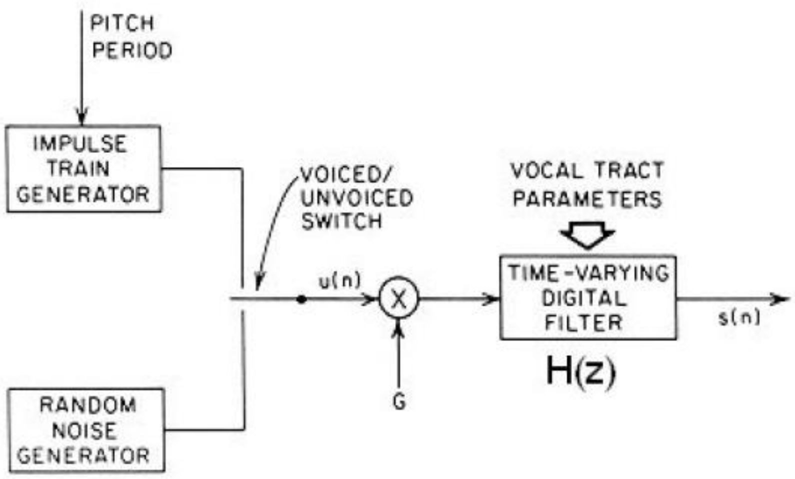


Figure 4: Speech synthesis based on LPC model in human throat.

A linear predictive (LP) analysis of a speech signal is based on the all-pole model [5]. 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 speech sample can be approximated as a linear combination of past speech samples. By minimizing the sum of the squared differences (over a finite interval) between the actual speech 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 speech (speech frames) and is constrained to be of the following form, where H (z) is a p^{th} order polynomial with z-transform and the coefficients a1,a2, …, ap are assumed to be constant over the speech analysis frame

H(z) = ! + a1z-1 + a2 z -2 + a3 z -3 + … + ap z-p

Here the order 'p' is called the LPC order. Thus the output of the LPC spectral analysis block is a vector of coefficients (LPC parameters) that specify(parametrically) the spectrum that best matches the signal spectrum over the period of time in which the frame of speech sample was accumulated.

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

S(n) ≈ a1s(n-1) + a2s(n-2) + aps(n-p),…… (1)

where the coefficients a1, a2,…, ap are assumed constant over the speech analysis frame. We convert eq. (1) to an equality by including an excitation, G u(n), giving:

Ps(n) = Σ ais(n-i) + G u(n), ……(2)

i=1

where u(n) is a normalized excitation and G is the gain of the excitation [6]. By expressing eq (2) in the z-domain we get the relation

S(z) = Σ ai z-i S(z) + G U(z),….. (3)

i=1, leading to the transfer function

H(z) = S(z) = 1 = 1……(4)

G U(z) p H(z)

Because speech signals vary with time, this process is done on short chunks of the speech signal, which are called frames. Usually 30 to 50 frames per second give intelligible speech 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.

VI. MFCC MODEL

MFCCs are widely used as acoustic features in speech analysis as they provide a compact representation of log-spectral envelope of speech signals [7]. Human perception of the frequency content of sounds, either for pure tones or for speech signals, does not follow a linear scale. This research has led to the idea of defining subjective pitch of pure tones. Thus for each tone with an actual frequency, f, measured in Hz, a subjective pitch is measured on a scale called the "mel" scale. As a reference point, the pitch of a 1 KHz tone, 40 dB above the perceptual hearing threshold, is defined as 1000 mels .Other subjective pitch values are obtained by adjusting the frequency of a tone such that it is half or twice the perceived pitch of a reference tone (with a known mel frequency).A filter bank, in which each filter has a triangular band pass frequency response, and the spacing as well as the bandwidth is determined by a constant mel frequency interval. (The spacing is approximately 150mels and the width of the triangle is 300 mels). Mel scale cepstral analysis use scepstral smoothing to smooth the modified power spectrum. This is done by direct transformation of the log power spectrum to the cepstral domain using an inverse Discrete Fourier Transform (DFT).The modified spectrum of S(w) thus consists of the output power of these filters when S(w) is the input. Denoting these power coefficients by Sk, k = 1, 2, . . . , K, wecan calculate what is called the mel-frequency cepstrum, Cn,k

$Cn = \Sigma$ (log Sk) $\cos[\, n\, (k - 1/2)\, \pi/K]$,

k=1

n = 1, 2. . . L,

where L is the desired length of the cepstrum. The first 12 coefficients (1st frame) can be discarded since they are the mean of the signal and hold little information. Hence13th coefficient (1st frame) is usually considered. Conventional MFCCs based on Short-Time Fourier Transform (STFT) are not necessarily sufficient for speech synthesis. The difference between the cepstrum and the Mel-frequency cepstrum is that in the Mel frequency cepstrum, the frequency bands are positioned logarithmically (on the Mel scale) which approximates the human auditory system's response more closely than the linearly-spaced frequency bands obtained directly from the FFT or DCT. DCT may be used as a hiding domain during steg analysis [8]. This can allow for better processing of data, for example, in audio compression. However, unlike the sonogram, MFCCs lack an outer ear model and, hence, cannot represent perceived loudness accurately. Thus, in the sound processing, the Mel-frequency cepstrum is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency.

VIII. STEPS IN MFCC EXTRACTION ARE AS FOLLOWS

The mel-frequency cepstrum has proven to be highly effective in recognizing structure and modeling the subjective pitch and frequency content of audio signals[[9].Much processing is needed between acquiring the voice signal and achieving speech recognition, which constitutes the speech recognition procedure. MFCCs are short-term spectral features.[10] They are calculated as follows:

Frame Blocking: In this step the continuous speech signal is blocked into frames of N samples, with adjacent frames being separated by M (M<N). Thus, audio noise signal s is blocked into frame of N samples shifting every M sample for each frame[11]. According to characteristics of the background noise it is possible to adapt eachor some blocks dynamically, so as to optimize their performance by selecting the best configuration for that type of noise. Human speech is a non-stationary signal, but when segmented into parts ranging from 10-40 msec, these divisions are quasi-stationary. For this reason the human speech input is to be divided into frames before feature extraction takes place. The selected properties for the speech signals are a sampling frequency of 16 kHz, 8-bit monophonic PCM format in WAV audio. The chosen frame size is of 256 samples, resulting in each frame containing 16 msec portions of the audio signal. It seems that a value of 256 for N is an acceptable compromise. Furthermore the number of frames is relatively small, which will reduce computing time.

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

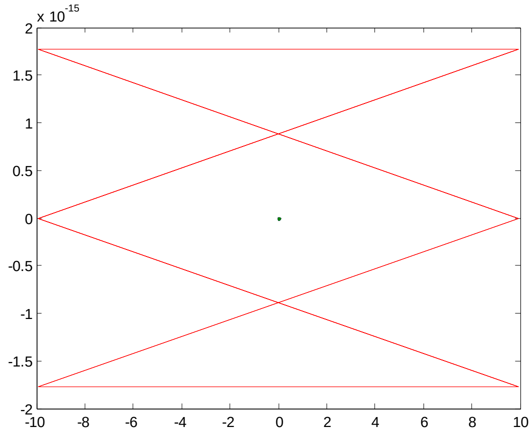


Figure 5: Frame of noise in MATLAB.

Windowing: The use of the window function reduces the frequency resolution by40%, 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[12]. The windowing serves a second purpose and that is the reduction of the spectral distortion that arises from the windowing itself.

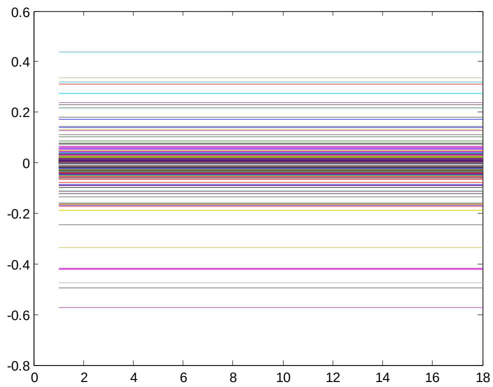


Figure 6(a): Office noisewindowed data after Hamming in MATLAB.

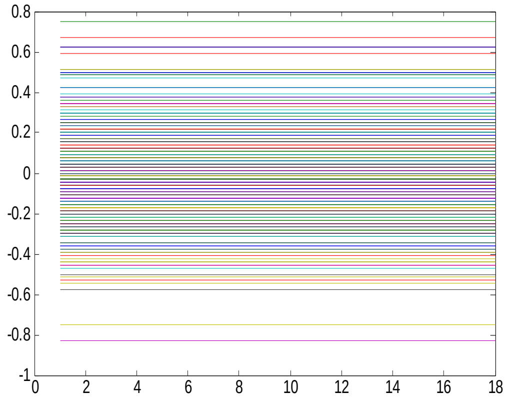


Figure 6(b): Fight noise windowed data after Hamming in MATLAB

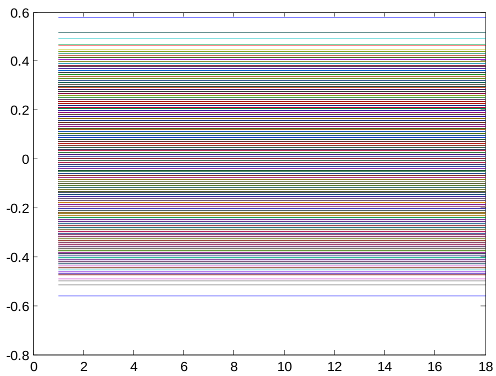


Figure 6(c): Crowd noise windowed data after Hamming in MATLAB.

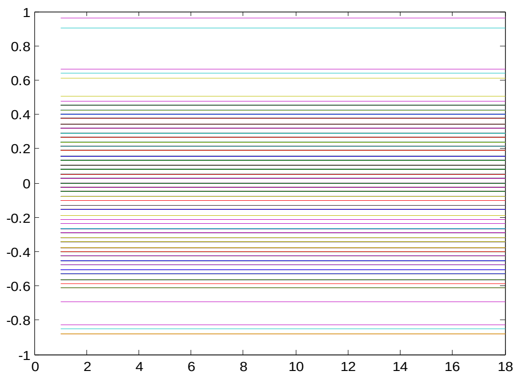


Figure 6(d): Crowd noise windowed data after Hamming in MATLAB.

Fast Fourier Transform: FFT is an effective tool for transforming signal into frequency domain. The frame size is not a fixed quantity and therefore can vary depending on the resulting time portion of the audio signal. The reason that the authors selected number of samples as 256 is that it is a power of 2, which enables the use of the Fast-Fourier Transform. The FFT is a powerful tool since it calculates the DFT of an input in a computationally efficient manner, saving processing power and reducing computation time. The operation results in the spectral coefficients of the windowed frames. It is desirable to derive an approximated version of Wang and Shamma's early auditory (EA) model in the frequency domain, where FFT algorithms are available[13].

Mel-scale Filter bank Frequency Transformation: Mel-cepstral coefficients are the features that will be extracted from speech during our work.[12,14] The key difference between MFCCs and cepstral coefficients lies in the processing involved when extracting each of these characteristics of a speech signal. The process of obtaining Mel-cepstral coefficients involves the use of a Mel-scale filter bank. Interpolation on log filter bank features can be more effective since the correlation is much stronger in the filter bank domain.[15] The spectral 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 perceives sound. The filter bank is composed of triangular filters[16] that are equally spaced on alogarithmic scale[17]. The Mel-scale warping is approximated and represent by the following Mel(f) = 2595 log10 (1 + f / 700), where f is frequency.

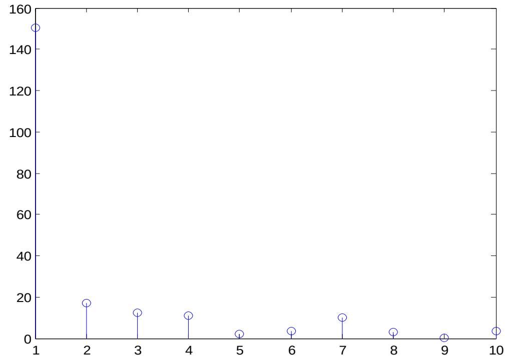


Figure 7(a): Mel-Spectral Coefficients of Office noise in MATLAB.

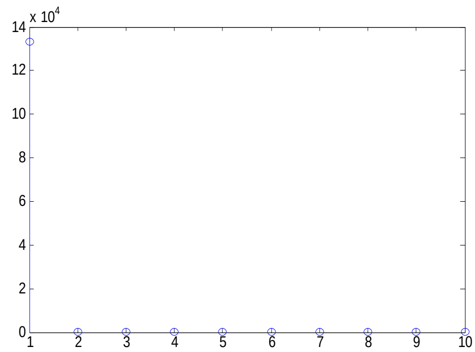


Figure 7(b): Mel-Spectral Coefficients of Fight noise in MATLAB.

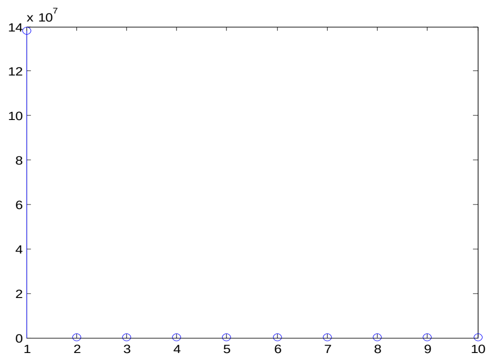


Figure 7(c): Mel-Spectral Coefficients of Crowd noisein MATLAB.

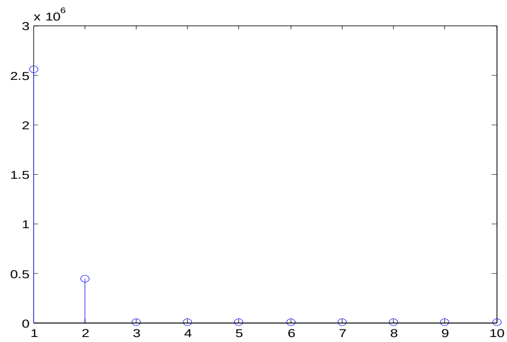
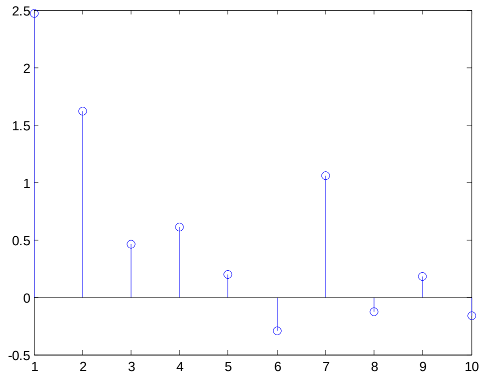


Figure 7(d): Mel-Spectral Coefficients of City noise in MATLAB

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

Discrete Cosine Transform: The Discrete Cosine Transform is applied to the log of the Mel-spectral coefficients to obtain the Mel-Frequency Cepstral Coefficients. Only the first 12 coefficients of each frame are kept, since most of the relevant information is kept amongst those at the beginning. The first 12 coefficients (1st frame) can be discarded since they are the mean of the signal and hold little information. Hence 13th coefficient (1st frame) is usually considered and the use of the DCT minimizes the distortion in the frequency domain.



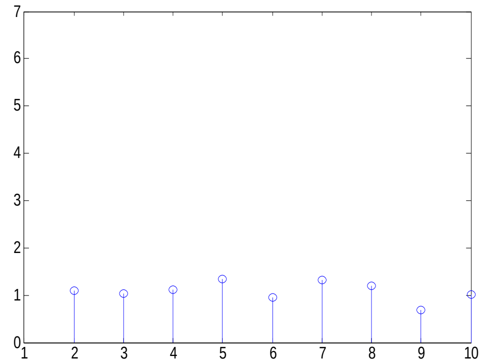
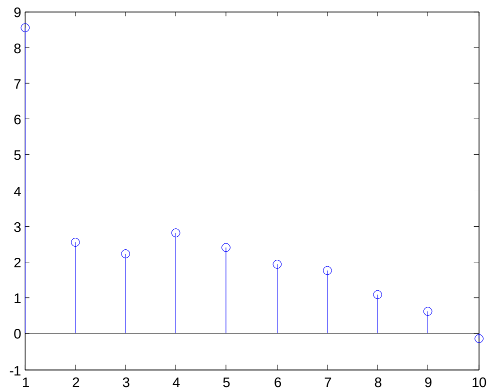


Figure 8(a):Office noise plot of DCT in MFCC **Figure 8(b):Fight noise plot of DCT in MFCC**



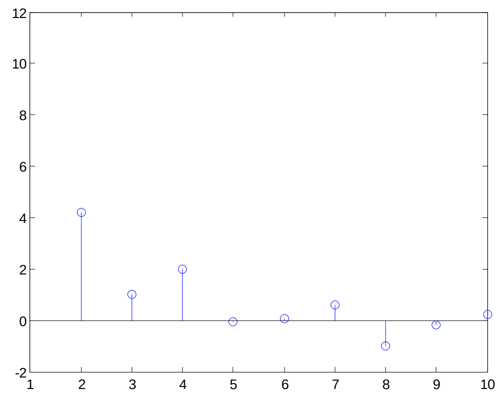


Figure 8(c):Crowd noise plot of DCT in MFCC **Figure 8(d):City noise plot of DCT in MFCC**

IX. RCEP MODEL

Real Cepstrum (RCEP) was first introduced by as part of homomorphic analysis of speech signals, in order to estimate the transfer function the vocal tract and the glottal pulse, under the assumption that pitch can be modeled as an impulse train[18]. From the theoretical point of view, the Cepstrum is defined as the inverse Fourier transform of the real logarithm of the magnitude of Fourier transform.[19] Therefore, by keeping only the first 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. This denotes the Fourier Transform of x and hence real cepstrum as a real-valued function can be used for the separation of two signals convolved with each other. Thus, RCEP is a cepstrum-based technique for determining a Harmonics-to-Noise Ratio (HNR) in Speech Signals and is a valid technique for determining the amount of

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

spectral noise, because it is almost linearly sensitive to both noise and jitter for a large part of the noise or jitter continuum. Thus real cepstrum block gives the real cepstrum output of the input frame and is also a popular way to define the prediction filter. Last, the line spectrum frequencies [20](a.k.a. line spectrum pairs) are also frequently used in speech coding. Line spectrum frequencies are another representation derived from linear predictive analysis which is very popular in speech coding.

IX. RESULTS OBTAINED IN MATLAB (UPTO TENTH ORDER FOR FIVE SAMPLES OF FOUR NOISES)

MFCC

Office briefcs1	Office cash reg	Office coin	Office elevator	Office faxmachine
0.6010	**0.9618**	**0.7044**	**1.6651**	**2.0590**
Fight filmfight	**Fight gunfight2**	**Fight gunfight3**	**Fight punch**	**Fight punch1**
2.3086	**1.5006**	**2.1490**	**0.8588**	**1.6766**
Crowd applause	**Crowd baskball**	**Crowd boooo**	**Crowd cheers**	**Crowd groan**
1.0022	**1.0480**	**1.6299**	**1.6689**	**1.6104**
City Police	**City Firecar**	**City Car rev.**	**City Carpenter**	**City Ambulance**
1.7570	**1.9231**		**0.5962**	**1.86971**

LPC

Office briefcs1	Office cash reg	Office coin	Office elevator	Office faxmachine
6.9394e-004	**0.0079**	**0.0050**	**0.0101**	**5.5141e-004**
Fight filmfight	**Fight gunfight2**	**Fight gunfight3**	**Fight punch**	**Fight punch1**
0.0174	**0.0215**	**0.0024**	**0.0300**	**0.0066**
Crowd applause	**Crowd baskball**	**Crowd boooo**	**Crowd cheers**	**Crowd groan**
0.0057	**0.0062**	**0.0065**	**0.0094**	**6.502[**
City Police	**City Firecar**	**City Car rev.**	**City Carpenter**	**City Ambulance**
0.0019	**0.0068**	**0.0137**	**8.0761e-004**	**0.0143**

RCEP

Office briefcs1	Office cash reg	Office coin	Office elevator	Office faxmachine
-3.1464	2.4564	0.4147	1.6377	1.4654
Fight filmfight	**Fight gunfight2**	**Fight gunfight3**	**Fight punch**	**Fight punch1**
0.4712	0.1696	0.4019	0.2923	0.2502
Crowd applause	**Crowd baskball**	**Crowd boooo**	**Crowd cheers**	**Crowd groan**
0.1399	0.3801	0.1215	0.2235	0.0365
City Police	**City Firecar**	**City Car rev.**	**City Carpenter**	**City Ambulance**
0.2090	2.8307	2.7555	0.1720	1.5375

Averages of Coefficients

MFCC COEFFICIENT	C1	C2	C3	C4	C5
Office noise (1-5)	7.7834	0.8202	0.3242	0.7064	1.0960
Fight noise (1-5)	10.5817	1.2234	1.2120	1.0187	0.7566
Crowd noise(1-5)	10.0298	0.4107	0.9704	0.2983	0.8025
City noise(1-5)	9.0532	2.2542	1.2018	1.1513	0.5801

LPC COEFFICIENT	C1	C2	C3	C4	C5
Office noise (1-5)	1.0000	-1.0936	0.8458	-0.3932	0.1799
Fight noise (1-5)	1.0000	-1.1821	0.8716	-0.4722	0.1652
Crowd noise(1-5)	1.0000	-1.7156	1.5275	-1.1130	0.9742
City noise(1-5)	1.0000	-1.0723	0.7006	-0.2870	0.1512

RCEP COEFFICIENT	C1	C2	C3	C4	C5
Office noise (1-5)	0.5656	0.1558	0.1018	0.0074	0.1065
Fight noise (1-5)	3.0748	-0.0763	-0.0417	-0.0287	0.0285
Crowd noise(1-5)	2.0536	-0.0048	-0.0731	-0.1067	0.0804
City noise(1-5)	1.5009	-0.1337	-0.1043	0.0100	-0.0686

X. TRIAL & ERROR

Sl No Trial method (for MFCC) Difference

Noise Samples	Office Noise	Fight Noise	Crowd Noise	City Noise
Average of first ten coefficients	0.6010	2.30864	1.0021	1.7570
Average of maximum & minimum coefficients	1.0843	9.2912	4.4319	4.8564
Average of second highest & third highest coefficients	1.3361	1.3838	0.6025	3.0846
Average of highest & second highest coefficients	2.0416	10.2613	5.1617	7.4529

XI. CONCLUSION

Our MATLAB results show that out of three noise parameters under consideration, Mel Frequency Cepstral Frequencies are robust features in noise parameter estimation and its characterization. By trial & error method, it was found that the best result of MFCC was obtained at maximum difference of 2.0416, 10.2613, 5.1617 & 7.4529 when average of highest & second highest MFCC coefficients was taken of Office, Fight scene, Crowd & City noises respectively since scaling becomes easier at maximum difference while undergoing defining membership in fuzzy logic operation for noise classification. Smart Volume Controllers may be designed after background noise classification based on these aspects.

REFERENCES

[1] Jidong Yang, Buket D. Barkana, ”The Acoustic Properties of Different Noise Sources”, Sixth International Conference on Information Technology: New Generations, Computer Society IEEE, 2009.

[2] Francesco Beritelli, Salvatore Casale& Salvatore Serrano, “Adaptive Robust speech processing Based on Acoustic Noise Estimation and Classification”, International Symposium on Signal Processing and Information Technology, IEEE, 2005.

[3] Longbiao Wang, Shinji Ohtsuka & Seiichi Nakagawa,”High Improvement of speaker Identification by Combining MFCC and Phase Information”, ICASSP, IEEE, 2009.

[4] C. Freeman, R.D. Dony& S.M. Areibi, “Audio Environment Classification for Hearing Aids Using Artificial Neural Networks with Windowed Input”, Proceedings of the Symposium on Computational Intelligence in Image and Signal processing (CIISP), IEEE, 2007.

[5] Ill-Conditioning and Bandwidth Expansion in “Linear Prediction of Speech”, Peter Kabal, Electrical & Computer Engineering, McGill University, Montreal, Quebec H3A 2A7.

[6] IEEE Transactions on,”Speech and Audio Processing”, vol.3. no.2, March 1995 Ravi P. Rramachandran, Mihailo S. Zilovic and Richard mammone

[7] IEEE Transactions on audio, “speech and language processing”, VOL. 20, NO. 7, September 2012TomiKinnunen, Member, IEEE, Rahim Saeidi, Member, IEEE, Filip Sedlák, Kong Aik Lee, Johan Sandberg, Maria Hansson-Sandsten, Member, IEEE, and Haizhou Li, Senior Member, IEEE.

[8] Liu Jing & Tang Guangming,”A New Method for Data Hiding Domain Classification”, International Forum on Information Technology andApplication, Computer Society, IEEE, 2009.

[9] “Automatic Music Summarization In Compressed Domain”.Xi Shao# *, Changsheng Xu#, Ye Wang* , Mohan S Kankanhalli*.

[10] “Mel Frequency Cepstral Coefficients for Music Modeling”, Beth Logan-Cambridge Research Laboratory, Compaq Computer Corporation, One Cambridge Center, Cambridge MA 02142. USA. Email-Beth.Logan@compaq.com.

[11] Cherdchai Eamdeelerd & Kraisin Songwatana, “Audio Noise Classification using Bark scale features and K-NN Technique”, IEEE, 2008.

[12] International Journal of Computer Applications (0975 – 8887)Volume 2 – No.5, June 2010 “Recognition of Human Speech using q-Bernstein Polynomials”.

[13] Wei Chu & Benoit Champagne, “A Noise-Robust FFT-Based Auditory Spectrum With Application in Audio Classification”, Transactions on Audio, Speech, and Language Processing, Vol. 16, No. 1, IEEE, January 2008.

[14] “Evaluation Of A Speech Recognition / Generation Method Based On Hmm And Straight” T. Irino*,+, Y. Minami*, T. Nakatani*, M. Tsuzaki**,+, and H. Tagawa*

[15] “Channel Error Protection Scheme For Distributed Speech Recognition”, Zheng-Hua Tan and Paul Dalsgaard Center for Person Kommunikation, Aalborg University, Denmarkzt, pd@cpk.auc.dk

[16] International Journal of Computer Applications (0975 – 8887) Volume 17– No.1, March 2011 "Vector Quantization Approach for Speaker Recognition using MFCC and Inverted MFCC".

[17] "Increased Mfcc Filter Badwidth For Noise-Robust Phoneme Recognition". Mark D. kowronski and John G. Harris.

[18] Gabriel Zigelboim and DrIlan D. Shallom, "A comparison Study of Cepstral Analysis with Applications to Speech Recognition", International Conference on Information technology,2006.

[19] IEEE Transactions on"Circuits And Systems—Ii": EXPRESS BRIEFS, VOL. 53, NO. 10, OCTOBER 2006 Soo-Chang Pei, Fellow, IEEE, and Huei-Shan Lin.

[20] Chen Zhe.,Wang Hongy Qiu Tianshuang. "Fractional Cepstrum and applying to Blind Chirp Signal Processing" Proceedings ICSP'O4.