Acoustic Noise Classification and Characterization Using Statistical Properties

Prabhavathi C.N.1, Dr. K. M. Ravikumar2
1Research Scholar, Jain University, Bangalore, India
2Second Special Officer/P.G. Co-Ordinator, VTU Regional Office, Mysore, India

Abstract- Acoustical noises or background noises degrade the quality and intelligibility of the speech signal. The background noises are random in nature and it is important to classify such noises. Hence background noises are characterised based on statistical properties like mean, variance, co-variance etc. The autocorrelation function which is a time invariant function is used for classifying such noises. The SNR for various noises is estimated.

Keywords— Background noise, statistical properties, autocorrelation, SNR.

I. INTRODUCTION

The background noises like moving vehicle noise, babble noise, construction site noise, street noise, aircraft noise, fan noise etc [1] will degrade the quality and intelligibility of the speech signal during the communication from one end to other end. Hence the clean speech signal will be degraded and becomes a noisy speech signal. Due to these noisy signals the performance of the speech processing systems like speech recognition system, speaker recognition and identification system, speech coding system, hearing aid system [2] will degrade. Therefore it is of prime importance to eliminate or reduce the effect of noisy signals and improve the quality and intelligibility of noisy speech signal.

In order to reduce the effect of noise or to eliminate the noise in noisy speech signal it is required to characterise and classify various background noisy signals. Since the noisy signals are random in nature it is required to classify and characterise them using statistical parameters like autocorrelation, mean, variance, co-variance etc.

Representation of random signal can be done in i) Time domain ii) Frequency domain iii) time frequency domain. In time domain the random signal is analysed with various parameters like mean, standard deviation, variance, root mean square value. In frequency domain the power spectral density (psd) function is determined. In time frequency domain the spectrogram is analysed.

In this paper, noise classification is done using autocorrelation function. The noisy signal is said to be stationary if the autocorrelation function is time invariant, otherwise the noisy signal is said to be non stationary signal. Along with this the other statistical parameters are determined. The characterization of noise signal is done by determining the mean, variance and SNR.

The paper is organised as follows: Section II describes the related work, Section III describes the various statistical properties in time domain. Section IV describes the method used to determine the SNR. Section V describes the implementation steps. Section VI and VII describes results and conclusion respectively.

II. RELATED WORK


In this paper the basic parameters are considered for noise classification, categorization which are discussed in section III[8]. The purpose for considering these parameters is, since noise and noisy speech are random in nature. These parameters help to find the region where exactly the noise power or energy is concentrated in the noisy speech. The said parameters can be used under time domain and frequency domain.
Most of the parameter values estimated can be used to subtract / to cancel noise from the noisy speech which helps to enhance the degraded speech signal.

Also the effect of noise on various age group is considered for understanding purpose.

III. TIME DOMAIN APPROACH FOR NOISE CLASSIFICATION

The statistical characteristics of two random signals \( \{x_k(t)\} \) and \( \{y_k(t)\} \) are as follows [8]:

Mean Value (Expectation):

\[
\mu_x(t) = \frac{1}{N} \int_{-\infty}^{\infty} x_k(t) \, dt
\]

Variance:

\[
\sigma_x^2(t) = \mathbb{E}[x_k(t) - \mu_x(t)]^2 = \mathbb{E}[x_k^2(t)] - \mu_x^2(t)
\]

Co-Variance:

\[
e_{xy}(t) = \mathbb{E}[(x_k(t) - \mu_x(t))(y_k(t) - \mu_y(t))]
\]

IV. PROPOSED METHOD FOR NOISE CHARACTERIZATION

It is important to estimate the noise level in the noisy speech signal in terms of signal to noise ratio (SNR). This helps in the characterization of noise. The following equation is used to determine the SNR

\[
(SNR)_{SNR} = 10 \log \frac{\text{mean of signal power}}{\text{mean of noise power}}
\]

The SNR is also calculated by making use the variance as follows

\[
(SNR)_{SNR} = 10 \log \frac{\sigma^2_x - \sigma^2_n}{\sigma^2_n}
\]

Where \(\sigma^2_n\) is variance of noisy speech signal and \(\sigma^2_n\) is variance of noise signal [1].

V. IMPLEMENTATION STEPS

Noise database containing different background noises like white noise, traffic noise, canteen noise, mall noise and construction site noise are recorded using a microphone and an iphone for a duration of 60s.

Also a speech database is created by recording a sentence “A king ruled the state in the early days” which is taken from the TSP speech database [3]. The TSP speech database consists of 100’s of sentences.
The recording is done in a closed acoustic room for various age groups. The age groups are considered to be age less than 12 yrs, age between 12-19, 20-29, 30-40, and above 40. Each age group includes a male and a female. The SNR is calculated using “Eq. 14” and tabulated for various age groups across the various noises.

VI. RESULTS AND DISCUSSION

The plots of traffic noise, mall noise, fan noise and white noise is as shown in “Fig 1”.

From “Fig. 1” it is observed that white noise comes under stationary random signal as it is time invariant and the traffic noise, fan noise and mall noise also comes under stationary type of random signal. Whereas the gun shot noise, construction site noise, military noise can be put under the non stationary type of random signal. Since such signals has impulsive type of response Figure 2,3,4 and 5 shows the time plot for various age groups for the sentence ‘A king ruled the state in the early days’. The upper plot in each graph is of female and lower plot is of male.

Fig 1: Plot of various noises

Fig 2: Time domain plot for age <12yrs

Fig 3: Time domain plot for age between 15yrs and 21 yrs

Fig 4 : Time domain plot for age 40yrs
The tabulation of statistical parameters like mean, variance and standard deviation are tabulated as shown in “Table I” for various noise signals.

**Table I.**
Statistical Values Of Mean, Co-Variance And Standard Deviation

<table>
<thead>
<tr>
<th>Noise</th>
<th>Mean</th>
<th>Co-variance</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>-1.6801e-004</td>
<td>0.0012</td>
<td>0.0344</td>
</tr>
<tr>
<td>Traffic noise</td>
<td>-1.0997e-004</td>
<td>0.0021</td>
<td>0.0461</td>
</tr>
<tr>
<td>Mall noise</td>
<td>1.1993e-004</td>
<td>0.0070</td>
<td>0.0835</td>
</tr>
<tr>
<td>Fan noise</td>
<td>-2.6451e-005</td>
<td>2.3921e-006</td>
<td>0.0015</td>
</tr>
</tbody>
</table>

**Table II.**
Energy And Power Of Noises

<table>
<thead>
<tr>
<th></th>
<th>Traffic noise</th>
<th>Mall noise</th>
<th>Fan noise</th>
<th>White noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>0.8912</td>
<td>2.4956</td>
<td>4.5358e-004</td>
<td>0.4863</td>
</tr>
<tr>
<td>power</td>
<td>0.7942</td>
<td>6.2281</td>
<td>2.0573e-007</td>
<td>0.2364</td>
</tr>
</tbody>
</table>

The observation from the above tables is as follows:
The SNR’s are very low which implies the quality and intelligibility of the speech signal is totally lost. Hence an improvement of the SNR has to be done by suppressing the noise.

The improvement of the SNR can be done by subtracting the noise by its power / Energy from the degraded speech signal.
VII. SUMMARY AND CONCLUSIONS

For various noises the statistical properties were found out and tabulated. From the property of autocorrelation function, the noise signal was classified as stationary signal and non-stationary noise signal. SNR for various noises across various age groups were found out and tabulated. It has been concluded that since the snr values are very low, it is of primary importance to reduce the noise and to improve the performance of the system. Hence a robust algorithm has to be implemented to improve the snr for speaker independent and noise independent.

REFERENCES


[4] Ling MA, Been Milner and Dan smith 2006 Acoustic Environoment classification ACM transactions on speech and Language Processing


