IJRSET 2015 SPL Volume 2, Issue 11 Pages: 34-41
ISSN 2394-739X
International Journal for Research in Science Engineering and Technology

Noise Reduction Techniques and Algorithms For Speech Signal Processing

[1] M. A. Josephine Sathya, [2] Dr. S. P. Victor
[1] Guest Faculty & Ph.D Research Scholar, [2] Associate Professor & Head,
[1, 2] Department of Computer Science
[1] Mother Teresa Women's University, [2] St.Xavier's College (Autonomous),
[1] Kodaikanal,Tamil Nadu, India, [2] Palayamkottai, India.

Abstract:-

Acoustic problems in the environment have gained attention due to the tremendous growth of technology Exposure to high decibels of sound proves damaging to humans from both a physical and a psychological aspect. The problem of controlling the noise level in the environment has been the focus of a tremendous amount of research over the years. This paper describes a study of techniques for noise reduction which can be applied at the input to standard receivers trained on noise-free speech. In this review, we have classified the existing noise cancellation schemes and algorithms.

Keywords: - Noise reduction, Digital Signal processing, speech signal, Adaptive filters, Smoothing Algorithms.

1. INTRODUCTION

Noise can be defined as an unwanted signal that interferes with the communication or measurement of another signal. A noise itself is an information-bearing signal that conveys information regarding the sources of the noise and the environment in which it propagates. For example, the noise from a car engine conveys information regarding the state of the engine and how smoothly it is running, cosmic radiation provides information on formation and structure of the universe and background speech conversations in a crowded venue can constitute interference with the hearing of a desired conversation or speech.

The types and sources of noise and distortions are many and varied and include

(1) Electronic noise – such as the normal noise and shot noise

(2) Acoustic noise - emanating from moving , vibrating or colliding sources such as revolving machines, moving vehicles, keyboard clicks, wind and rain,

(3) Electromagnetic noise - that can interfere with the transmission and reception of voice, image and data over the radio-frequency spectrum,

(4) Electrostatic noise - generated by the presence of a voltage,

(5) communication channel distortion and fading and

(6) Quantization noise - lost data packets due to network congestion.

Signal distortion is the term often used to describe a systematic undesirable change in a signal and refers to changes in a signal due to the non-ideal characteristics of the communication channel, signal fading reverberations, echo, multipath reflections and missing samples. Noise and distortion are the main factors that limit the capacity of data transmission in telecommunication and the accuracy of results in signal measurement systems. Therefore the modeling and removal of the effects of noise and distortions have been at the core of the theory and practice of communications and signal processing. Noise reduction and distortion removal are important problems in applications such as cellular mobile communication, speech recognition, image processing, medical signal processing, radar, sonar, and in any application where the desired signals cannot be isolated from noise and distortion or observed in isolation[1].

2. INFLUENCE OF NOISE ON SPEECH SIGNAL APPLICATIONS

The performance of any speech signal processing system is degraded in the presence of noise (either additive or convolution). This is due to the acoustic mismatch between the speech features used to train and test this system and the ability of the acoustic models to describe the corrupted speech. When processing the speech signal, the quality of speech may be at risk from various sources of interference or distortions[2]. Typical sources of interference are:

- Background noise added to the speech signal: for example – environmental noise or engine noise when talking
on a mobile phone,
- Unintended echo occurring in closed spaces with bad acoustics,
- Acoustic or audio feedback: it occurs in two-way communication when the microphone in the telephone captures the actual speech of another person and the speech of the first person reproduced from loudspeakers, and sends them both back to the first person,
- Amplifier noise: an amplifier can produce additional thermal noise, which becomes noticeable during significant signal amplifications,
- Quantization noise created in the transformation of the analogue signal to digital: the interference occurs during sampling due to rounding up real values of the analogue signal,
- Loss of signal quality, caused by coding and speech compression. Due to numerous sources of interference influencing the speech signal, when designing the system for speech signal processing, it is necessary to apply the techniques of noise cancellation and speech quality improvement[2].

3. LINEAR FILTERING OF DIGITAL SIGNAL

Prior to processing, the analogue signal must be transformed into the digital form. The procedure of transforming the analogue speech signal into a digital one creates additional noise during sampling, called quantization noise. However, already at the sampling frequency of 8 kHz and 16-bit sample resolution, the intensity of quantization noise is neglectable in comparison to other noise sources (microphone amplifier noise, environmental noise). Once the analogue audio signal is transformed into a digital one, different techniques for noise cancellation and increasing speech signal quality are applied. The basic technique is linear filtering of the digital signal. Linear filtering encompasses signal processing in a time domain, reflected in a change of source signal spectrum content. The goal of filtering is to reduce unwanted noise components from the speech signal. Usually, linear digital filters consist of two types:

1. Finitive Impulse Response filters – FIR filters
2. Infinite Impulse Response filters – IIR filters.

In FIR filters, the output signal y[t] of a certain linear digital system is determined by convoluting input signal x[t] with impulse response h[t]:

$$Y[t] = x[t] * h[t] \qquad (1)$$

Where, t is the time domain value. Along with the time domain, digital filtering can also be conducted in the frequency domain. Digital filters in the frequency domain are divided into four main categories: low-pass, band-pass, band-stop and high-pass [3].

4. NOISE CANCELLATION IN FREQUENCY DOMAIN

The main procedure of filtering in the frequency domain i.e.spectral filtering consists of the input signal analysis, filtering and synthesis of the filtered signal. The input signal analysis consists of framing and unitary transform from a time domain to a transformdomain.

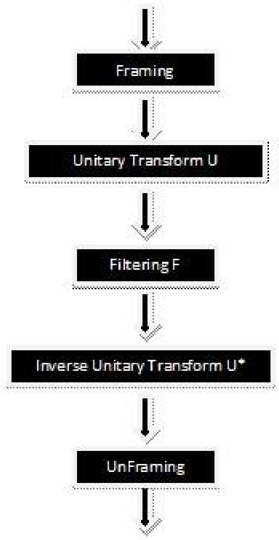

Figure 1: Spectral filtering procedure

The transform domain is most often the frequency domain. This is followed by filtering and return to the time domain, by the inverse unitary transform with unframing. Filtering primarily consists of the reduction of those frequencies whose power is below a certain threshold also called noise floor. The main goal of unitary transform is signal separation to a group of separate components, where it is easier to distinguish between the speech signal vector and the noise signal vector. Moreover, with the transform most of the speech signal energy is compressed into a relatively small number of coefficients, which facilitates processing. The most frequently used unitary transforms are the Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and the Karhunen-Loeve Transform (KLT) [4].

5. NOISE CANCELLATION USING ADAPTIVE FILTERING

Adaptive Noise Canceller (ANC) removes or suppresses noise from a signal using adaptive filters that automatically adjust their parameters .The ANC uses a reference input derived from single or multiple sensors located at points in the noise field where the signal is weak or undetectable. Adaptive filters then determine the input
signal and decrease the noise level in the system output. The parameters of the adaptive filter can be adjusted automatically and require almost neither prior signal information nor noise characteristics. However, the computational requirements of adaptive filters are very high due to long impulse responses, especially during implementation on digital signal processors.Convergence becomes very slow if the adaptive filter receives a signal with high spectral dynamic range such as in non-stationary environments and colored background noise. In the last few decades, numerous approaches have been proposed to overcome these issues. For example, the Wiener filter, Recursive-Least-Square (RLS) algorithm, and the Kalman filter were proposed to achieve the best performance of adaptive filters. Apart from these algorithms, the Least Mean Square (LMS) algorithm is most commonly used because of its robustness and simplicity. However, the LMS suffers from significant performance degradation with colored interference signals.[1]. Other algorithms, such as the Affine Projection algorithm (APA), became alternative approaches to track changes in background noise; but its computational complexity increases with the projection order, limiting its use in acoustical environments.

An adaptive filtering system derived from the LMS algorithm, called Adaptive Line

Enhancer (ALE), was proposed as a solution to the problems stated above. ALE is an adaptive self-tuning filter capable of,separating the periodic and stochastic components in a signal. The ALE detects extremely low-level sine waves in noise, and may be applied in speech with noisy environment. Furthermore, unlike ANCs, ALEs do not require direct access to the noise nor a way of isolating noise from the useful signal. In literature, several ALE methods have been proposed for acoustics applications. These methods mainly focus on improving the convergence rate of the adaptive algorithms using modified filter designs, realized as transversal Finite Impulse Response (FIR), recursive Infinite Impulse Response (IIR), lattice, and sub-band filters.

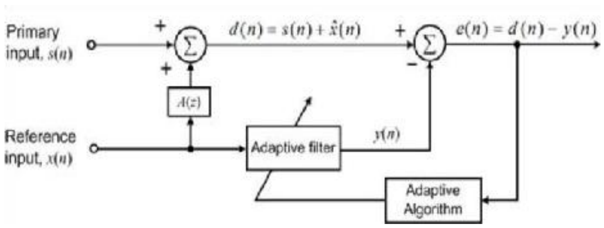

Figure 2: Block diagram of adaptive noise cancellation system

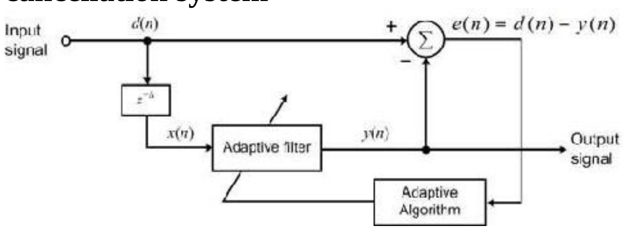

Figure 3: Block diagram of adaptive line enhancer

It is shown that for this application of adaptive noise cancellation, large filter lengths are required to account for a highly reverberant recording environment and that there is a direct relation between filter mis-adjustment and induced echo in the output speech. The second reference noise signal is adaptively filtered using the least mean squares, LMS, and the lattice gradient algorithms. These two approaches are compared in terms of degree of noise power reduction, algorithm convergence time, and degree of speech enhancement [5].

The effectiveness of noise suppression depends directly on the ability of the filter to estimate the transfer function relating the primary and reference noise channels. A study of the filter length required to achieve a desired noise reduction level in a hard-walled room is presented. Results demonstrating noise reduction in excess 10dB in an environment with 0dB signal noise ratio [6].

6. SMOOTHING ALGORITHMS

In many experiments in physical science, the true signal amplitudes (y-axis values) change rather smoothly as a function of the x-axis values, whereas many kinds of noise are seen as rapid, random changes in amplitude from point to point within the signal. In the latter situation it may be useful in some cases to attempt to reduce the noise by a process called smoothing. In smoothing, the data points of a signal are modified so that individual points that are higher than the immediately adjacent points (presumably because of noise) are reduced, and points that are lower than the adjacent points are increased. This naturally leads to a smoother signal. As long as the true underlying signal is actually smooth, then the true signal will not be much distorted by smoothing, but the noise will be reduced.

Most smoothing algorithms are based on the "shift and multiply" technique, in which a group of adjacent points in the original data are multiplied point-by-point by a set of numbers (coefficients) that defines the smooth shape, the products are added up to become one point of smoothed data, then the set of coefficients is shifted one point down the original data and the process is repeated. The simplest smoothing algorithm is the rectangular or unweighted sliding-average smooth; it simply replaces each point in the signal with the average of m adjacent points, where m is a positive integer called the smooth width. For example, for a 3-point smooth (m = 3):

$$S_j = \frac{Y_{j-1} + Y_j + Y_{j+1}}{3} \quad (2)$$

for j = 2 to n-1, where S_j the j^{th} point in the smoothed signal, Y_j the j^{th} point in the original signal, and n is the total number of points in the signal. Similar smooth operations can be constructed for any desired smooth width, m. Usually m is an odd number. If the noise in the data is "white noise" (that is, evenly distributed over all frequencies) and its standard deviation is s, then the standard deviation of the noise remaining in the signal after the first pass of an unweighted sliding-average smooth will be approximately s over the square root of m (s/sqrt(m)), where m is the smooth width.

The triangular smooth is like the rectangular smooth, above, except that it implements a weighted smoothing function. For a 5-point smooth (m = 5):

$$S_j = \frac{Y_{j-2} + 2Y_{j-1} + 3Y_j + 2Y_{j+1} + Y_{j+2}}{9} \quad (3)$$

for j = 3 to n-2, and similarly for other smooth widths .It is often useful to apply a smoothing operation more than once, that is, to smooth an already smoothed signal, in order to build longer and more complicated smooths. For example, the 5-point triangular smooth above is equivalent to two passes of a 3-point rectangular smooth. Three passes of a 3-point rectangular smooth result in a 7-point "pseudo-Gaussian" or haystack smooth, for which the coefficients are in the ratio 1 3 6 7 6 3 1. The general rule is that n passes of a w-width smooth results in a combined smooth width of n*w-n+1. For example, 3 passes of a 17-point smooth results in a 49-point smooth. These multipass smooths are more effective at reducing high-frequency noise in the signal than a rectangular smooth.In all these smooths, the width of the smooth m is chosen to be an odd integer, so that the smooth coefficients are symmetrically balanced around the central point, which is important because it preserves the x-axis position of peaks and other features in the signal. (This is especially critical for analytical and spectroscopic applications because the peak positions are often important measurement objectives).Note that we are assuming here that the x-axis intervals of the signal is uniform, that is, that the difference between the x-axis values of adjacent points is the same throughout the signal. This is also assumed in many of the other signal-processing techniques described in this essay, and it is a very common (but not necessary) characteristic of signals that are acquired by automated and computerized equipment.

Noise reduction Smoothing usually reduces the noise in a signal. If the noise is "white" (that is, evenly distributed over all frequencies) and its standard deviation is s, then the standard deviation of the noise remaining in the signal after one pass of a triangular smooth will be approximately s*0.8/sqrt(m), where m is the smooth width. Smoothing operations can be applied more than once: that is, a previously-smoothed signal can be smoothed again. In some cases this can be useful if there is a great deal of high-frequency noise in the signal. However, the noise reduction for white noise is less in each successive smooth. For example, three passes of a rectangular smooth reduces white noise by a factor of approximately s*0.7/sqrt (m), only a slight improvement over two passes. The frequency distribution of noise, designated by noise color, substantially effects the ability of smoothing to reduce noise. The Matlab/Octave function “NoiseColorTest.m” compares the effect of a 100-point boxcar (unweighted sliding average) smooth on the standard deviation of white, pink, and blue noise, all of which have an original unsmoothed standard deviation of 1.0. Because smoothing is a low-pass filter process, it effects low frequency (pink) noise less, and high-frequency (blue) noise more, than white noise.

Original unsmoothed noise	1
Smoothed white noise	0.1
Smoothed pink noise	0.55
Smoothed blue noise	0.01

End effects and the lost points problem. Note in the equations above that the 3-point rectangular smooth is defined only for j = 2 to n-1. There is not enough data in the signal to define a complete 3-point smooth for the first point in the signal (j = 1) or for the last point (j = n) , because there are no data points before the first point or after the last point. (Similarly, a 5-point smooth is defined only for j = 3 to n-2, and therefore a smooth cannot be calculated for the first two points or for the last two points). In general, for an m-width smooth, there will be (m-1)/2 points at the beginning of the signal and (m-1)/2 points at the end of the signal for which a complete m-width smooth cannot be calculated. What to do? There are two approaches. One is to accept the loss of points and trim off those points or replace them with zeros in the smooth signal. (That's the approach taken in most of the figures in this paper). The other approach is to use progressively smaller smooths at the ends of the signal, for example to use 2, 3, 5, 7... point smooths for signal points 1, 2, 3,and 4..., and for points n, n-1, n-2, n-3..., respectively. The later approach may be preferable if the edges of the signal contain critical information, but it increases execution time. The fast smooth function discussed below can utilize either of these two methods.

Examples of smoothing. A simple example of smoothing is shown in Figure 4. The left half of this signal is a noisy peak. The right half is the same peak after undergoing a triangular smoothing algorithm. The noise is greatly reduced while the peak itself is hardly changed. Smoothing increases the signal-to-noise ratio and allows the signal characteristics (peak position, height, width, area, etc.) to be measured more accurately by visual inspection.

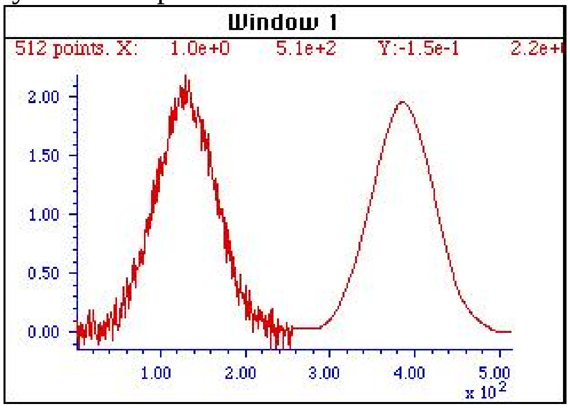

Figure 4. The left half of this signal is a noisy peak.

The right half is the same peak after undergoing a **smoothing** algorithm. The noise is greatly reduced while the peak itself is hardly changed, making it easier to measure the peak position, height, and width directly by graphical or visual estimation (but it does not improve measurements made by least-squares methods.

The larger the smooth width, the greater the noise reduction, but also the greater the possibility that the signal will be distorted by the smoothing operation. The optimum choice of smooth width depends upon the width and shape of the signal and the digitization interval. For peak-type signals, the critical factor is the smoothing ratio, the ratio between the smooth width m and the number of points in the half-width of the peak. In general, increasing the smoothing ratio improves the signal-to-noise ratio but causes a reduction in amplitude and in increase in the bandwidth of the peak.

CONCLUSION

The performance of any speech signal processing system is degraded in the presence of noise (either additive or convolution). This is due to the acoustic mismatch between the speech features used to train and test this system and the ability of the acoustic models to describe the corrupted speech. Various techniques for filtering the noise from a speech waveform has been studied. Most of these technique is based upon the concept tof adaptive filtering and takes advantage of the quasi-periodic nature of the speech waveform to supply a reference signal to the adaptive filter. Preliminary tests by authors indicate that the technique appears to improve the quality of noise speech and slightly reduce granular quantization noise. This technique also appears to improve the performance of the linear prediction analysis and synthesis of noisy speech. It is also found from studies that, for the lower order FIR adaptive filter, RLS algorithm produce highest SNR and it is superior to LMS in its performance. But LMS is converging faster that RLS for the Finite Impulse response (FIR) filter Taps. Optimum Mu (LMS) and Lambda (RLS) values have been obtained by fixing the FIR Tap weight. Acoustic noise cancellation ANC is best suited to remove ambient noise. The traditional wideband ANC algorithms work best in the lower frequency bands and their performance deteriorates rapidly as the bandwidth and the center frequency of the noise increases. Most noise sources tend to be broadband in nature and while a large portion of the energy is concentrated in the lower frequencies, they also tend to have significant high frequency components. Further, as the ANC system is combined with other communication and sound systems, it is necessary to have a frequency dependent noise cancellation system to avoid adversely affecting the desired signal.The major drawback of traditional single band ANC algorithms is that the performance deteriorates rapidly as the frequency of the noise increases. However, noise in real world conditions tends to be broadband with significant high frequency components. Adaptive filtering has been used for speech denoising in the time domain. During the last decade, wavelet transform has been developed for speech enhancement. Spectral analysis of non-stationary signals can be performed by employing techniques such as the Adaptive filters like LMS, NLMS, STFT and the Wavelet transform (WT), which use predefined basis functions. Empirical mode decomposition (EMD) performs very well in such environments. Also, Acoustic noise with energy greater or equal to the speech can be suppressed by adaptively filtering a separately recorded correlated version of the noise signal and subtracting it from the speech waveform. It is shown that for this application of adaptive noise cancellation, large filter lengths are required to account for a highly reverberant recording environment and that there is a direct relation between filter misadjustment and induced echo in the output speech. The second reference noise signal is adaptively filtered using the least mean squares, LMS, and the lattice gradient algorithms. These two approaches are compared in terms of degree of noise power reduction, algorithm convergence time, and degree of speech enhancement. Both methods were shown to reduce ambient noise power by at least 20 dB with minimal speech distortion and thus to be potentially powerful as noise suppression pre-processors for voice communication in severe noise environment.

REFERENCE

[1]. Saeed V. Vaseghi , “Advanced Digital Signal Processing and Noise Reduction”, Fourth Edition , 2008 ,John Wiley & Sons, Ltd. ISBN: 978-0-470-75406,Pp. 35 – 36.

[2]. Lakshmikanth .S, Natraj .K .R , Rekha .K .R, “Noise Cancellation in Speech Signal Processing-A Review”,2014, International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 1, January 2014, Pp.1 – 12.

[3]. M. Maletić, “Digital filtering audio signals”, ZEA – internal department material, FER, University of Zagreb, 2005.
[4]. J. Abraham, “Discrete Cosine Transforms”, course: Discrete Transforms and Their Applications, University Of Texas at Arlington,USA, 2009.Pp.5176 – 5185.
[5]. Boll, S., “Suppression of acoustic noise in speech using two microphone adaptive noise cancellation”, Acoustics, Speech and Signal Processing, IEEE Transactions on (Volume: 28, Issue: 6).
[6]. Pulsipher, D., “Reduction of non-stationary acoustic noise in speech using LMS adaptive noise cancelling”, Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '79, (Volume: 4).
[7] P. Kakumanu, A. Esposito, O. N. Garcia, R. Gutierrez-Osuna, "A comparison of acoustic coding models for speech-driven facial animation”, Elsevier, Speech Communication vol. 48, pp. 598–615, 2006.
[8] G. Zorić, I. S. Pandžić, “Real-time language independent lip synchronization method using a generic algorithm”, Signal Processing - Special section: Multimodal human-computer interfaces,Volume 86, Issue 12, pp. 3644-3656, 2006.
[9] Rabiner, L., Juang, B.H., “An introduction to Hidden Markov models”, IEEE, ASSP Magazine, and ISSN: 0740-7467 Vol. 3,Issue1, and Jan 1986.
[10]. 1. Lyzenga J, Festen J, Houtgast T: A speech enhancement scheme incorporating spectral expansion evaluated with simulated hearing loss of frequency selectivity. J Acoust Soc Am 2002; 112:1145–1157.