*Journal of Computing Research and Innovation (JCRINN) Vol 2, No 2 (2017)*
*https://crinn.conferencehunter.com*

# Article 9
# Noise Reduction in Audio

Zulfikri Paidi, Nurhasyimah Idrus
Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA Perlis Branch, Malaysia

***Abstract***
*Noise reduction in audio is a process of minimizing the presence of noise in the audio signal. Noise reduction approach helps user understand how to remove noise from the signal. Audio is very susceptible with noise and it presence will give impact on two major ways; first it will degrade the quality of information being transfer and second limit the capacity of data being transmit. There are lot of noise reduction techniques been proposed as to remove noise from audio signal such as wiener filter and average filter. In this project, wavelet theory implementation is proposed for reducing noise from audio signal. The main objective is to identify and remove the presence of noise in the signal. Furthermore, measurements of noise signal pattern are recorded before and after the noise reduction process. To be specific, this project will focus on speech type of audio. Speech acquisition of the participant were recorded in two types of environments; controlled and uncontrolled environments. Types of speech recorded in each environment are normal, whispering and angry speech. Speech signal patterns are analysed using cross-correlation between three wavelets families consists of Coiflet 5, Daubechies 10 and Daubechies 9 in order to determine best fit the original signal.*

***Keywords*** *-Audio, Noise, AWGN, Wavelets, Cross-correlation*

## Introduction
Noise is referred to signal, which is unable to convey any useful information arising due to some sort of disturbance and cause the interference in the communication or the measurement of signal (Villanueva-Luna, Adrian, Jaramillo-Nuñez, Alberto, Sanchez-Lucero, Daniel, Ortiz-Lima, Carlos, Aguilar-Soto, Gabriel, Flores-Gil, Aaron, & May-Alarcon, 2011). In order to acquire the required information without the presence of noise especially in audio files such as speech recording files, the noise reduction must be used. There are various kinds of tools and software available which can be used to improve the signal contain of noise in the audio files such as Wiener filter, Average filter, and subspace transformation (Benesty, Jingdong, Chen, & Huang, 2008). In this project, MATLAB is used to develop a system that can remove the noise since the result of the technique recovers signal with higher correlation.

In ideal situation, audio recordings are free from the distorted or noisy speech signal. Besides, the quality of information signal is not corrupted and the capacity of information transmission is high in audio recordings. However, like any signal, audio recordings are very susceptible with noise. Noise is determined as a corruption in the communication that degrades the quality of information signal and limits the capacity of data transmission (Goel, Roopali, Jain, & Ritesh. (2013). Without the reduction of noise in the signal, it will limit the ability of user to identify the information transfer in audio files contain of noise (JaiShankar, & Duraiswamy, 2012).

The importance of this project is to implement the noise reduction technique in audio files contain of speech. In addition, noise reduction can solve major problems such as unclear audio files and in many applications such as cellular mobile communication, speech recognition, image processing, medical signal processing, radar, sonar, and any other applications where the desired signals cannot be isolated from noise.

**Noise Reduction in Audio Theory**

There are large amount of research and engineering concern in noise reduction in order to obtain a clean speech without interference of noise (Villanueva-Luna, et. al., 2011); (Aggarwal, Rajeev, Karan, Jai, Kumar, Vijay, Rathore, Sanjay, Tiwari, Mukesh, Khare, & Anubhuti. 2011); (Benesty, et al., 2008). As well as this put an interest to those researchers, this project comprehends the theory and implementation of the noise reduction technique. First and foremost, this project describes the audio theory continues with the noise theory and wavelets theory.

A. *Audio Theory*

In the age of digital information, audio data has been becoming an essential part in various modern computer applications. Audio classification has become a focus in the research of audio processing and pattern recognition (Dhanalakshmi, Palanivel, Ramalingam, & Vennila, 2009). Audio is a representation of sound in which sound is referred to vibrations that travel through the air or another medium whether solid, liquid, or gaseous and can be heard by human ears (McLoughlin, 2009). All vibrations of sound are known as sound waves. The frequency of the waves is dependent on the frequency of the vibrating source. The high frequency of the vibrating source impresses the sound wave to have a high frequency (Vijaykumar, Vanathi, & Kanagasapabathy, 2007). In addition, frequency is defined as rate or the number of times per second, which a sound wave cycles from positive to negative and to positive again (Villanueva-Luna, et. al. (2011). Frequency is measured in cycles per second or Hertz (Hz).

Meanwhile the audio signal is the form of sound that represents longitudinal variations of pressure in a medium to be synthesized directly or originate at a transducer such as a microphone to convert the audio signal into the electrical signal (Villanueva-Luna, et. al. (2011). Transducer functions as a device used to convert the signal from one form of energy into another form of energy (Seiça, 2012). The main advantage of converting the signal into electrical signal is the signal can then be processed or also known as signal processing.

Signal processing occurs as audio signal can be represented in digital or analogue signal (Mihov, Doychev, & Ivanov, 2009). Digital signal is converted into encoded digital data stream by using a digital-analogue converter (DAC) through an amplifier and a speaker. Digital audio in general is categorized as speech, music and noise which in this project analyzing on speech signal since speech is important in delivering information in communication (Vijaykumar, et al., 2007). Alternately, analogue signal applied analogue-digital converter (ADC) to convert into encoded digital data stream.

Speech is the most basic means to human communication (Gold, Morgan, & Ellis, 2011). Throughout the years, speech has been well applied into many areas of technology include database access service, speaker verification, identification of vocal-related illnesses and telephone banking (Muangjaroen, & Yingthawornsuk, 2012). In natural environment, speech signal generally immersed in noise and it is essential for speech processing system to implement noise reduction techniques to extract the desired speech signal (Russo, Stella, & Rozic, 2012). There are some characteristics of speech signals such as auditory quality, characterization of different sounds and problems that might occur. For further information, vowel "a" and "n" have a relatively regular shape, while "t" pronounced as a single letter and especially the sibilant "s" which consists of fast oscillating parts that look very similar to noise. Therefore, it is clear that noise reduction will be very sensitive for speech parts (Wieland, Urban, & Funken, 2009).

### B. *Noise Theory*

Generally, noise often present in transmission of signals, which can cause major problems in communication (Goel, et al., 2013). Noise is defined as an unwanted signal that interfere the communication or measurement of another signal (Villanueva-Luna, et. al. (2011). Noises present in communication channels are disturbing the original signals from path without any noise (JaiShankar, & Duraiswamy, 2012). Besides, noise disturbs in the communication which degrades the quality of information signal (Goel, et al., 2013). Over the past few decades there is an increase in the level of environmental noise which is due to growth of technology whereby the noise coming from various factors such as vehicles, noisy engines, and heavy machines (Chavan, Chavan, Manjusha, & Gaikwad, 2010). Noise is further classified into several categories depending on its frequency spectrum or time characteristics (Di Bert, Luca, Caldera, Peter, Schwingshackl, David, Tonello, & Andrea 2011). Several categories are further shown in Table 1.

Table 1: Categories of Noise

| Types of Noise | Characteristics |
|---|---|
| White noise | A pure random noise that has an autocorrelation function and consists of frequencies in a flat power spectrum |
| Band-limited noise | Noise that has same characteristics as the white noise with a limited bandwidth that covers the limited spectrum of signal and contains a sin-shaped autocorrelation |
| Narrowband-noise | A type of noise that contain a narrow bandwidth such as 50/60 Hz from the electrical supply |
| Colored-noise | A non-white noise or wide-band noise that has a non-flat shape of spectrum. Examples are brown noise and pink noise |
| Impulsive noise | Noise that contain a short-duration pulse of random amplitude, time of occurrence and duration |
| Transient noise pulses | A noise that consists of long duration noise pulses such as clicks and burst noise |

Although there are various categories of noise, this project is only investigating on removal of white noise in speech signal due to white noise is a common source of noise which consists in all frequencies (O'Haver, Prof. Tom. 2013). White noise is defined as the random signal contains equal power of spectrum within a fixed bandwidth at any centre frequency which refers to a statistical model for signals and signal sources rather than any specific signal (Aggarwal, et al., 2011), (Vijaykumar, et al. 2007). White noise is a type of noise which presents in which all frequency components, ranging from zero frequency (DC) to infinite frequencies (Anwar, Umer Hassan, & Sabieh, 2008).

Additive White Gaussian Noise Model (AWGN) is a channel model in which the only impairment to communication, an addition of independent White Gaussian Noise samples to the input signals

(MathWorks, 2014a). Besides, an AWGN channel adds White Gaussian Noise to the signal that passes through this AWGN model.

### C. *Wavelets Theory*

Since there is a problem in the presence of noise in the speech signal, wavelets can help to minimise the problem (Aggarwal, et al., 2011). Over a decade, various de-noising techniques have been proposed for the removal of noises from audio signal such as Wiener filter, Average filter, VisuShrink and BayesShrink(Aggarwal, et al., 2011), (Benesty, et al., 2008). However, recently, wavelet methods are mostly used in removal of noise in the signal processing due the high effectiveness in the reduction of noise in the signal (Aggarwal, et al., 2011). Wavelets are referred to nonlinear functions and do not remove noise by low-pass filtering like many traditional methods where this low-pass filtering can blur the sharp features in a signal and makes it difficult to separate the noise from the signal (Aggarwal, et al., 2011). Furthermore, wavelets are able to significantly reduce noise from speech without degrading the quality of speech signal or introducing any audible artefacts (Nongpiur, 2008).

Wavelets consists of a huge number of wavelets families which have several different wavelets having a high number of vanishing moments and capable of representing complex polynomials (Ratnakar, Sunil. Singh, & Nitisha 2009). Wavelets with more vanishing moments should be selected as it provides better reconstruction quality and introduce less distortion into processed speech and concentrate more signal energy in few coefficients.

There are three wavelets families; Daubechies 10, Daubechies 9 and Coiflet 5 that have high vanishing moments. Theoretically, vanishing moment is referred to scaling function which is able to represent complex signal accurately (Villanueva-Luna, et. al., 2011). Daubechies wavelets are widely used in solving a broad range of problem such as signal discontinuities and fractal problems (Ratnakar, Sunil. Singh, & Nitisha 2009). Daubechies wavelets are family of orthogonal wavelets of discrete wavelet transforms which is characterized by a maximal number of vanishing moments for some given support.

Coiflet 5 is discrete wavelets function to have scaling functions with vanishing moments. This Coiflet 5 wavelet is generally used in many applications using Caldeon-Zygmund Operators. These wavelets families are implemented in measurement of the project by using cross-correlation in order to determine best fit between the original signal and the ones that have been processed. Cross-correlation is referred to measure of similarity of two waveforms as a function of a time-lag applied to one of them. This concept of cross-correlation is a need and widely used for signal processing applications in digital communications (Anwar, et al. 2008).

## Methodology

The initial phase for this project is information gathering on noise reduction in audio whereby the objective, method, project requirement, and the significance of the project were determined. Speech was prepared and recorded. A participant was given the same speech and recorded speech was collected as samples of speech data. Speech acquisition of the participant were recorded in two types of environments; controlled and uncontrolled environments. Types of speech recorded in each environment are normal, whispering and angry speech. The overall samples used in this project are 48 speech data. Graphical User Interface Development Environment (GUIDE) in MATLAB was used for developing noise reduction system in speech development.

The speech signals can be processed, transmitted or stored once in digital domain. Since the Audio Device Block enables experimentation and processing of digital signals, From Audio Device in MATLAB Simulink was implemented to read audio data from an audio device in real time. From Audio Device Block was selected with To Multimedia File in MATLAB Simulink for the purpose to save the audio acquired by a given time. The From Audio Device Block buffers the data from audio device using the process illustrated in Figure 1.

The Figure 1 shows the From Audio Device GUI, where this project selected a 5 seconds queue period. The audio device begins writing the input data to a buffer at the start of the simulation. This input data has data type specified by the Device data type parameter. When the buffer is full, the From Audio Device block writes the contents of the buffer to the queue. The size of the queue specified by using the Queen duration (seconds) parameter. The From Audio Device Block pulls data from the top of the queue to fill the Simulink frame as the audio device appends audio data to the bottom of the queue. This data has the data type specified by the Output data type parameter. This file is used in this project to experience de-noise method using wavelets.

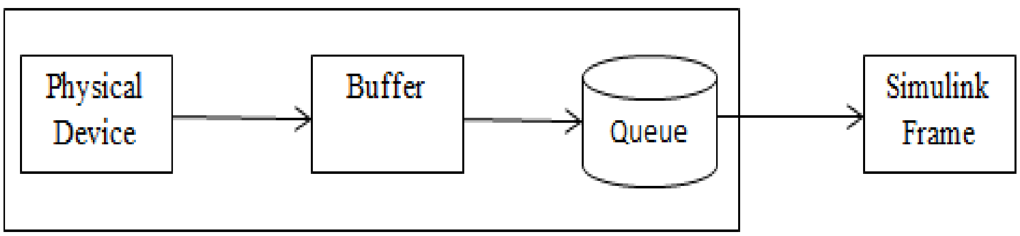


Figure 1: The From Audio Device Block

In addition, the Multimedia File Block write video frames, audio frames or both to a multimedia (.avi, .wav, .wma, .mp4, .ogg, .flac, or .wmv) file. Cross-platform supported file formats for audio files are WAV, FLAC, OGG, and MPEG4 (only for Window 7 and MAC OS X)In this project, MATLAB programming used the *wavread* function to load a WAVE file specified by the string filename which returning the sampled data in y. If the filename does not include an extension, *wavread* appends a .wav extension.

In this project, MATLAB programming used the *wavread* function to load a WAVE file specified by the string filename which returning the sampled data in y. If the filename does not include an extension, *wavread* appends a .wav extension. Furthermore, Waveform Audio File Format (WAVE) or commonly known as WAV due to its filename extension is a Microsoft and IBM audio file format standard for storing an audio bit stream on personal computer (PC). This audio file format is selected for this project since it is the main format used in Windows systems for raw and typically uncompressed audio and consists of simple structure, which is easy to exchange sound files among different programs. Some radio broadcasters especially which adopted the tapeless used uncompressed WAV files for transmitter simulation and receiver testing.

**Experimentation and Results**

Experiments were constructed using different types of speech signal; normal, whispering and angry signal in each control and uncontrolled environments. Three wavelet families of Coiflet 5, Daubechies 10 and Daubechies 9 had been analysed through the experiments to de-noise the presence of noise in three speech signals. From each experiment, the main objective is to first identify noise from the speech recorded and then remove the presence of noise in speech using wavelets techniques. In addition, White Gaussian Noise was added at level AWGN 0, AWGN 5, AWGN 10 and AWGN 15 into the speech signals before the noise reduction take places. Patterns of speech signal were collected and recorded inside spreadsheet.

Analysis of results obtained from the experiments is further shown in Figure 2. This figure showed that when the value of AWGN is increased until it reached the maximum AWGN 15, the pattern of signal become constant. The signals de-noised by Coiflet 5, Daubechies 10 and Daubechies 9 showed the constant result when reached AWGN 15.

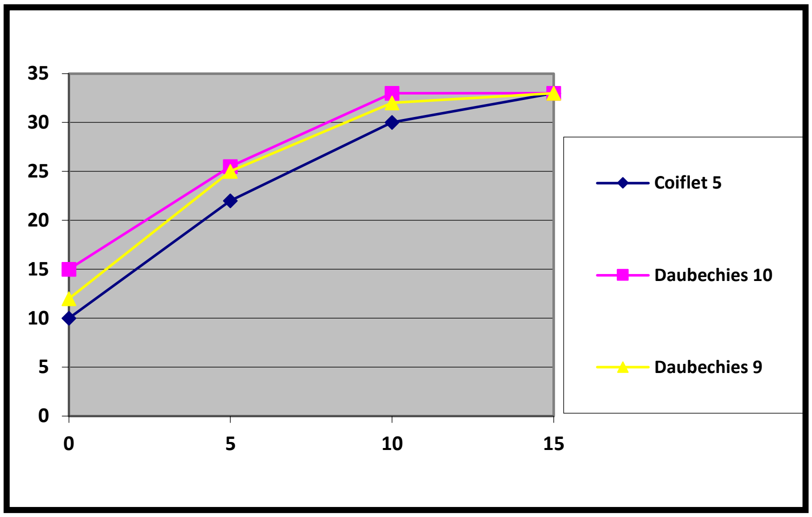

Figure 2: Signal and AWGN

Another finding acquired from the experiments is different expression of speech voices from recording which consists of normal, whispering and angry speech did not affect the AWGN. When the AWGN were reaching the maximum AWGN 15, the pattern of signal remain constant whether the experiments are using normal, whispering or angry speech signal.

**Conclusions**

This project had achieved its two objectives; to identify and remove the presence of noise in normal, whispering and angry signal in controlled and uncontrolled environments. Then, second objective is to measure the pattern of the normal, whispering, and angry signal before and after the noise reduction using cross-correlation technique. By comparing the experiment result between normal, whispering and angry signal in controlled and uncontrolled environments, it can be further concluded that Daubechies 10 and Daubechies 9 shows better result when de-noise the signal. In addition, the higher the amount of White Gaussian Noise (AWGN) is added to the signal, the better the signal can be de-noised. Based on the observation and analysis from the experiments conducted, the uncontrolled environment shows better result in de-noising process compared to in controlled environment.

## References

Aggarwal, Rajeev, Karan Singh, Jai, Kumar Gupta, Vijay, Rathore, Sanjay, Tiwari, Mukesh, & Khare, Anubhuti. (2011). Noise reduction of speech signal using wavelet transform with modified universal threshold. *International Journal of Computer Applications, 20*(5), 14-19.

Anwar, Umer Hassan and Sabieh. (2008). Statistical properties of White Noise (Electronics & Signal Processing).

Benesty, J., Jingdong, Chen, & Huang, Y. A. (2008). On the Importance of the Pearson Correlation Coefficient in Noise Reduction. *Audio, Speech, and Language Processing, IEEE Transactions on, 16*(4), 757-765. doi: 10.1109/TASL.2008.919072

Chavan, Mahesh S, Chavan, Mrs Manjusha N, & Gaikwad, MS. (2010). Studies on implementation of wavelet for denoising speech signal. *International Journal of Computer Applications, 3*(2).

Dhanalakshmi, P, Palanivel, S, & Ramalingam, Vennila. (2009). Classification of audio signals using SVM and RBFNN. *Expert Systems with Applications, 36*(3), 6069-6075

Di Bert, Luca, Caldera, Peter, Schwingshackl, David, & Tonello, Andrea M. (2011). *On noise modeling for power line communications.* Paper presented at the Power Line Communications and Its Applications (ISPLC), 2011 IEEE International Symposium

Goel, Roopali, & Jain, Ritesh. (2013). Speech Signal Noise Reduction by Wavelets. *International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN-2278-3075, 2*.

Gold, Ben, Morgan, Nelson, & Ellis, Dan. (2011). *Speech and audio signal processing: processing and perception of speech and music*: Wiley. com.

JaiShankar, B, & Duraiswamy, K. (2012). AUDIO DENOISING USING WAVELET TRANSFORM.

MathWorks. (2014a). AWGN Channel. from http://www.mathworks.com/help/comm/ug/awgn-channel.html

McLoughlin, Ian. (2009). *Applied speech and audio processing*: Cambridge University Press.

Mihov, Slavy, Doychev, Doycho, & Ivanov, Ratcho. (2009). *Practical Investigation of Specific Types of Noise Signals for the Purpose of Suppression in Hearing-Aid Devices.* Paper presented at the XLIV International Conference ICEST-2009, B.

Muangjaroen, Supavit, & Yingthawornsuk, Thaweesak. (2012). *A Study of Noise Reduction in Speech Signal using FIR Filtering.* Paper presented at the International Conference on Advances in Electrical and Electronics Engineering.

Nongpiur, R. C. (2008, March 31 2008-April 4 2008). *Impulse noise removal in speech using wavelets.* Paper presented at the Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference

O'Haver, Prof. Tom. (2013). Signals and noise. From http://terpconnect.umd.edu/~toh/spectrum/SignalsAndNoise.html

Ratnakar Madan, Prof. Sunil Kr. Singh, and Nitisha Jain2 (2009). Signal Filtering Using Discrete Wavelet Transform. *2*.

Russo, Mladen, Stella, Maja, & Rozic, Nikola. (2012). Noise reduction in speech signals using a cochlear model. *Advances in Smart Systems Research, 2*(1), 7.

Seiça, Álvaro. (2012). The Transducer Function: An Introduction to a Theoretical Typology in Electronic Literature and Digital Art. *Journal of Science and Technology of the Arts, 4*(1), 71-79.

Soto, J Gabriel, Flores-Gil, Aaron, & May-Alarcon, Manuel. (2011). De-Noising Audio Signals Using MATLAB Wavelets Toolbox.

Vijaykumar, VR, Vanathi, PT, & Kanagasapabathy, P. (2007). Modified adaptive filtering algorithm for noise cancellation in speech signals. *Electronics and Electrical Engineering.–Kaunas: Technologija*(2), 74.

Villanueva-Luna, Adrian E, Jaramillo-Nuñez, Alberto, Sanchez-Lucero, Daniel, Ortiz-Lima, Carlos M, Aguilar- Wieland, Bernhard, Urban, Karsten, & Funken, Stefan. (2009). *Speech signal noise reduction with wavelets.* Diplomarbeit an der Universität Ulm.