Wavelet based Human Voice Identification System

Maryam Mohammed Mubarak al Balushi[1], Vidhya Lavanya R[2], Sreedevi Koottala[3], Ajay Vikram Singh[4]

[1,2,3]Electronics and Communication, Department, Middle East College, Oman, Muscat
[4]Computer Science and Information Technology, Amity University, India, Noida
[1]malbalushi@mec.edu.om, [2]vidhya@mec.edu.om, [3]sreedevi@mec.edu.om, [4] ajayavs.iitr@gmail.com

Abstract: This paper investigates the use of wavelet transform in order to remove noise from the signals. One of the ongoing research in multimedia applications is speech signal processing. Wavelet denoising technique is attempted to reduce and remove noise from the audio signal. It is therefore required to transform audio signal to wavelet domain by using discrete wavelet transform followed by denoising algorithm. Both soft and hard thresholding of denoising technique is used to compare the performance of human noise identification. Denoising the signal is performed in the transformation domain and improvement in the denoising will be achieved in various families of wavelet transform. There are several types of wavelets such as Haar, Symlets, BiorSplines, Discrete Meyer (Dmey) and Reverse Biothogonal. In this paper, identification of human noise is verified with Dmey and Fejer- Korovkin wavelets. Comparative analysis is done to calculate SNR after the spectral subtraction. The quality of the audio signal is determined by Mean Square Error (MSE) and Signal to Noise Ratio (SNR) of the denoised signal. The simulation results are performed in Matlab program. From the result analysis, it is shown that the Fejer- Korovkin wavelet filter has a better performance as compared to other type of wavelet transform and can be able to identify and differentiate human voices with others.

Keywords: Wavelet filters, Denoising, Thresholding techniques, Signal to Noise Ratio (SNR),

I. INTRODUCTION

One of the major research area of Speaker recognition is the as identification of human voice by extracting its features. It's also, called voice recognition which is able to interpret and receive dictation through program or machine, where even carry out and understand commands of the speech. Today's technology is growing fast as converting the phrases or voice of human into electrical signal, while they can be transformed these electrical signal into coding patterns. Therefore, the major focus in this work will be to retrieve human voice from the natural noisy surroundings. Speech recognition is currently used in many real-time applications, such as cellular telephones, computers, and security system. However, these systems are far from perfect in correctly classifying human speech into words. In order to identify the human voice correctly, this paper proposes a new algorithm. Noises considered in this work are of two kinds; one, noise added to the human voice due to the surroundings and second, all audio signals other than human voice. The aim of the research work is to remove noise from an original signal. Noises are of many kinds such as pink, white even other kind of noise which found in speech or audio signal. There are many noises surrounding us. So, to get unique voice identification, noise must be removed. Once it is properly analysed, the noises can be removed and the voice can be identified. Suitable algorithm are used to remove the noises from the signals. Removal of noise and identification of the human voice is a challenging work in the speech signal processing field. Biometric is one of the methods to differentiate a human voice with a robotic voice and the comparative results can be stored in the data base. One of the methods for processing the audio signals is referred as wavelet transform, which provides performed outstanding results of audio signal. The noise can be added to the human signals with various kinds of devices. For example in vehicles, noisy engines, pumps etc where the noise probability can be increased and in turn reduces audio quality. This method of de-noising signals can be used in many applications such as in smart phones to control the device, in medical that can be help doctors to hear the heartbeat and distinguish the sounds of animals.

II. EXISTING WORK

There are several types of wavelet algorithms that were developed for denoising human voice signal. In spite of the fact that the literature covers an extensive diversity of such theories, this review will mainly focus on five major themes which appear frequently in particular the literature reviewed. The most pleasant approach to the issue of selection of the best wavelet based for speech signal has been proposed by [1] Long Y, Gang L and Jun Guo, presented the paper of simple algorithm to select suitable wavelet based on speech and it also explains the important role to solve problems of processing for speech signal. The work concludes Dmey wavelet as the best type for speech signal and however, the researchers have proved that the Haar wavelet for decomposition of speech signal is not suitable. In the recent paper [2] it was proposed that denoise of speech signal using suitable wavelet transform is important. Threshold values are maintained to denoise the speech signal and to recover the original one. In addition, a comparison is made between soft and hard types of thresholding, using matrix of Mean Opinion Score and Signal to Noise ratio to get average result for wavelet denoising and to retrieve unwanted and wanted signals. The researcher at the end have arrived that by producing suitable shrinkage wavelet denoising the human voice can be identified correctly. Recently, several authors at [3] have proposed methods of wavelet transform by using soft thresholding as better way for de-noising signal rather than hard thresholding. While, the authors stated that using frequency domain rather than time domain is better to differentiate high frequency noise and low frequency signal. In addition, they use multi-resolution analysis as practical application by using wavelet filter bank. Then, they discuss about modification of universal threshold and types of thresholding. It was observed in this paper and in related research that implementation is done by using algorithm of babble noise removal in matlab program which compares the original signal with denoising signal by assuming that the signal referring to coefficient amplitude of Discrete Wavelet Transform. On the other hand, in [4] it was proposed to use two types of wavelets to analyze speech signal such as Daubechies Wavelet Transforms (DB) and Symlets wavelets where, Daubechies is more pleasing of sonically in compression with other. They used both the thresholding types for comparison. The implementation of algorithm is done in three steps. First step is decomposition, then given detail for thresholding coefficients and reconstruction of signal. The authors concluded the paper by achieving the denoising of speech signal by using wavelets. Also, the Daubechies is requiring less decomposition levels for hard thresholding as compared to others wavelet. Also, it is mentioned that the best function is hard thresholding technique. As reported by [5, 6, 7, 8, 9] performance of denoising by using wavelet transform technique is best suitable for voice identification with the removal of different types of noise such as electromagnetic noise, electronic noise, quantization noise and acoustic noise which can be found in audio signal. On the other hand, the authors have used mathematical equations to determine cross correlation, Peak Signal to Noise Ratio (PSNR), and Mean Square Error. The researchers tested several families of wavelet and concluded that the best one to perform the speech signal is Daubechies10 which provided high signal to noise ratio.

III. PROPOSED WORK

Figure 1 shows the block diagram of proposed work which has undergone different phases. The block shown below consists of inputting the data, filtering, wavelet transform, denoising and identification of human voice.

Input data: Voices such as human voice, animal, robot or nature noise are given as input to analyse.

Wavelet transforms: Different set of the filter which translate and dilation of generating the wavelet is representing the wavelet theory. There are several types of wavelet called wavelet family; therefore it requires one filter of them to be designing.

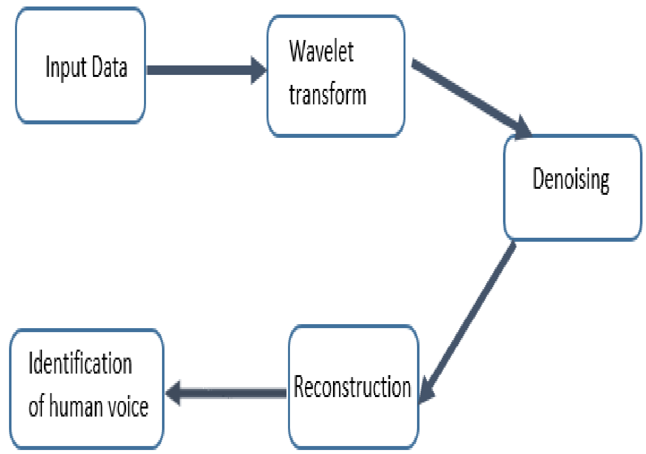

Fig. 1. Block Diagram

The wavelet transform (WT) is considered as a powerful tool due to its multi resolution possibilities of signal processing. In non-stationary signals application, the wavelet transform is very suitable with transitory phenomena, which frequency response differs in time comparing with Fourier transform. The wavelet coefficients referring to a measure of similarity in the frequency content of a chosen wavelet function and signal, where they can compute these coefficients as a convolution of the signal and the scaled wavelet function, on the other hand they can be interpreted like a dilated band pass filter due to its band pass like spectrum. So, in this part will describe the method which uses to analysis the data of input source. Then, it will convert the data of input source to wavelet while, apply this method is wavelet transform. Therefore will be use this technique in easy way to analysis voice sequences.

In analysis a discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location in time.

Denoising: This block analyses the data that comes from the transform. On the other hand, it represents the main section as it suppresses either the low frequency in the voice as detecting or enhancing in the voice whereas calculating the signal to noise ratio and the peak of a signal to noise ratio or high frequency as smoothing the voice. In addition, it could be filtered the voice either in the in the frequency domain. After applying the data in the system will require the filter to get pure signal data without any impurities.

Figure 2 explains the flowchart of the proposed work done for the identification of voice signal.

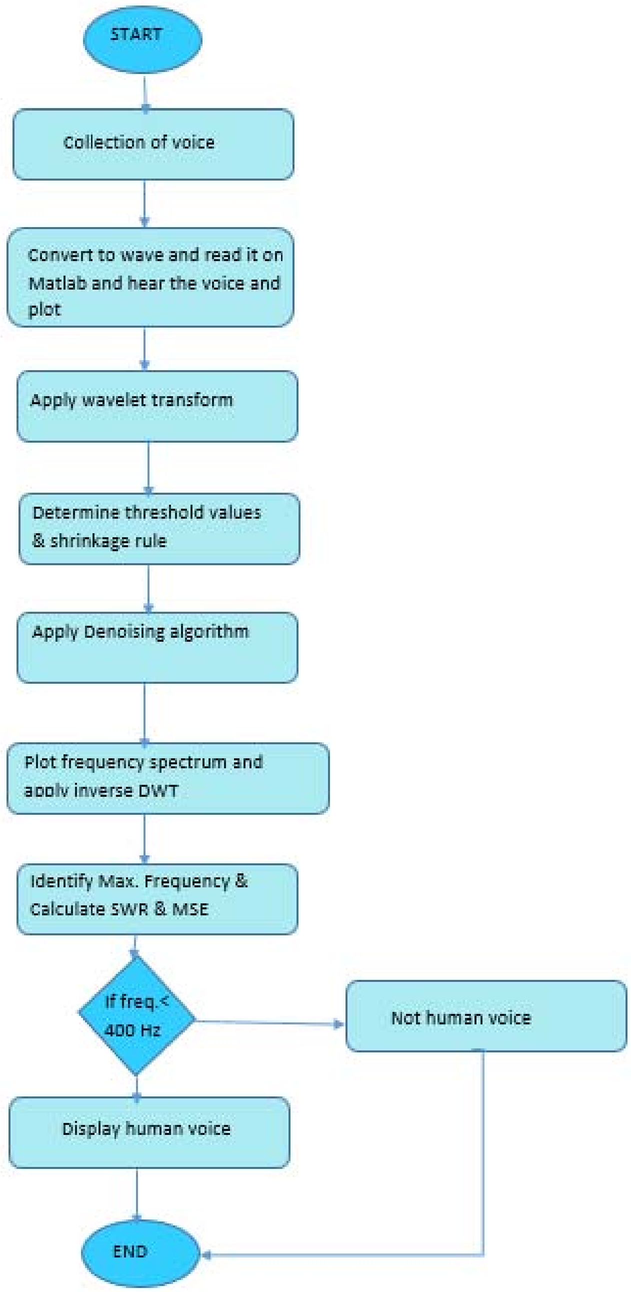

Fig. 2. Flow chart

IV. SIMULATION RESULTS

This section discusses in detail about the testing of the system step by step. In order to achieve this Mat lab program is used to apply the system and achieve the goals in sequence. The testing is done step by step to obtain the SNR values. Figure 3 represents the original signal, noisy signal added to the original and denoised signals. Figure 4 represents signal in left channel and reconstruction of signal with level 3 decomposition. Figure 5 represents the details of all the three levels. Figure 6 represents frequency spectrum of the signals.

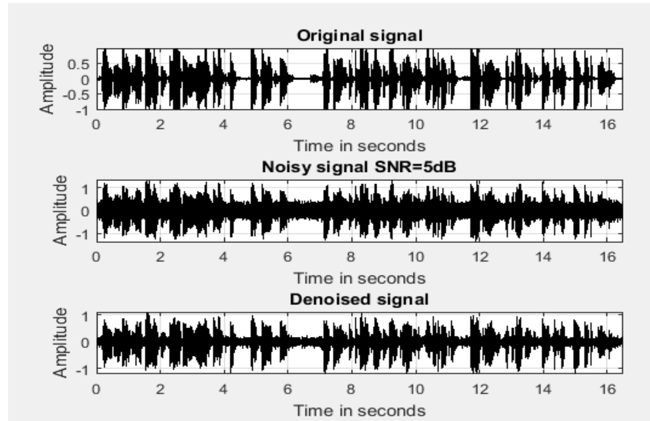

Fig. 3. Original signal, Noisy Signal (man1_nb) and Denoised signal

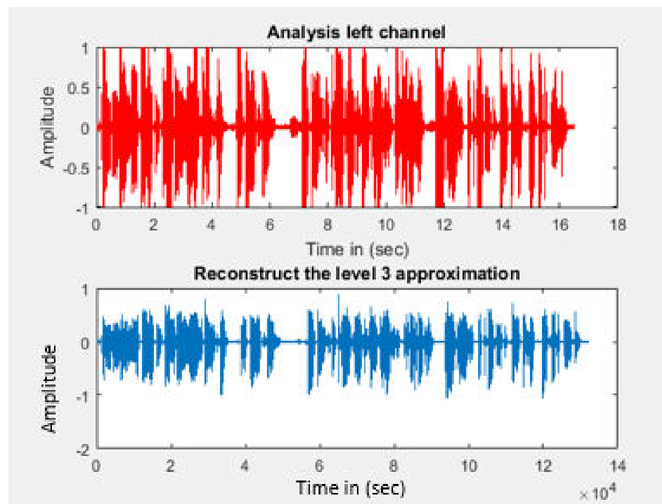

Fig. 4. Left channel signal with reconstruction level 3

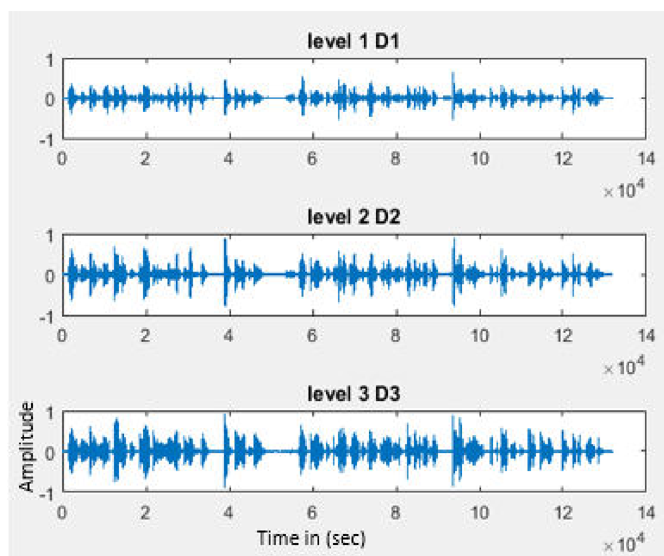

Fig. 5. Details for three levels

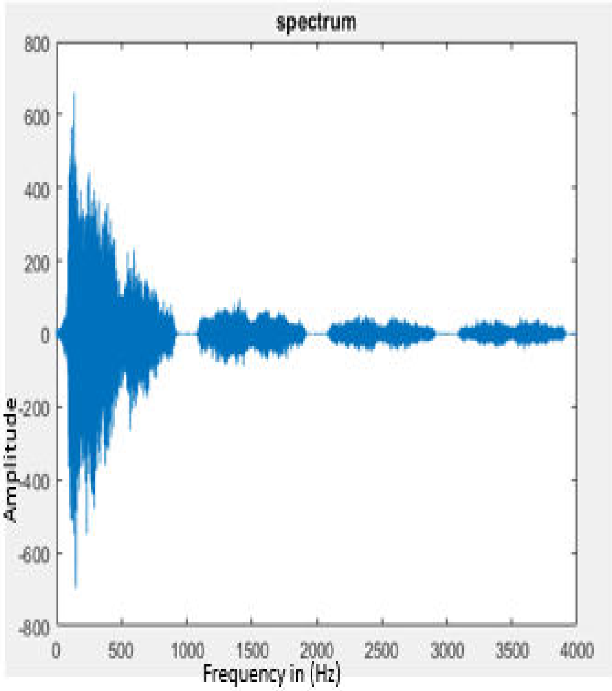

Fig. 6. Frequency spectrum of the signal

TABLE 1: SNR values of various voices using Demy wavelet with Heursure and Hard thresholding techniques.

Voice name	SNR	Error	Identification
man1_nb.wav	5.19	0.3755	Human voice
eerie-pitchy-voices.wav	5.9	0.8516	Not Human Voice
Duck-quacking animals038.wav	-0.9	0.6698	Not Human Voice
excuseme.wav	6.35	0.1132	Not Human voice
eng_m5.wav	3.35	1.3506	Human voice

Test with Dmey wavelet: The audio signals are collected from the online databases. Few of them are tested for Dmey wavelet and Fever Korovkin wavelet. Table 1 shows the SNR values of different voices. The shrinkage rules with the hard thresholding technique used in the proposed work is Heursure. The SNR values shows the ratio of the signal with the noise and the audio is identified as human or not human. The signal is identified correctly by using Dmey wavelet for all possible cases except for few. One such is listed in the table, signal name excuseme.wav. The denoising algorithm report it has high SNR but the identification went wrong.

TABLE 2: SNR values of various voices using Fejer- Korovkin wavelet with Heursure and Hard thresholding techniques.

Voice name	SNR	Error	Identification
man1_nb.wav	5.21	0.37	Human voice
eerie-pitchy-voices.wav	5.83	0.8501	Not Human Voice
Duck-quacking animals038.wav	1.00	0.6620	Not Human Voice
excuseme.wav	6.38	0.1111	Human voice
eng_m5.wav	3.39	1.3467	Human voice

Test with Fever-Korovkin wavelet: Table 2 represent the SNR values of the signals for Fever-Korovkin wavelet. Same database audio signals are used for the test. The SNR values are high when compared to table 1 and all the voices are identified correctly. For comparison purpose, Heursure shrinkage rule and hard threshold is used for both the wavelets.

V. CONCLUSION

Due to tremendous growth in digital technology an intense use of digital communications, it has become mandatory to make the signals free from noise. In speech signal processing, identification of human voice from the noise is a challenging work. This proposed work explains the algorithm for the retrieve of human voice with high SNR and correct identification. The results tabulated shows that the SNR is improved by a value of 5% using Fejer- Korovkin wavelet when compared with the Demy wavelet.

REFERENCES

[1] Y. Long, L. Gang and G. Jun, "Selection of the best wavelet base for speech signal, " in *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004.*, Hong Kong, China, China, 2004.

[2] B. J. Saikia and U. Baruah, "A GA-NN based wavelet method of speech signal denoising, " in *2015 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2015.

[3] Y. Tian and L. Wang, " An improved denoising method for speech signal, " in *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, Dengleng, China, 2011.

[4] S. Som, S. Sinha, R. Kataria (2016) "Study On SQL Injection Attacks: Mode, Detection And Prevention", International Journal of Engineering Applied Sciences and Technology, Indexed in Google Scholar, ICI etc., Impact Factor: 1.494, Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 23-29, June - July 2016.

[5] Q.-Y. Cai, G.-S. Wang and Y.-X. Yu, " Research of still image compression based on Daubechies 9/7 Wavelet transform, " in

2010 2nd International Conference on Future Computer and Communication, Wuha, China, 2010.

[6] T. Yadav and R. Mehra, "Denoising and SNR improvement of ECG signals using wavelet based techniques, " in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, Dehradun, India, 2016.

[7] Y. Cengiz and Y. D. U. Arıöz, "An Application for speech denoising using Discrete wavelet transform, " in *2016 20th National Biomedical Engineering Meeting (BIYOMUT)*, Izmir, Turkey, 2016.

[8] M. G. Sumithra and K. Thanuskodi, "Wavelet based speech signal de-noising using hybrid thresholding, " in *2009 International Conference on Control, Automation,Communication and Energy Conservation*, Perundurai, Tamilnadu, India, 2009.

[9] Y. Qian, "A Method of Improved Wavelet Threshold for Signal De-Noising, " in *2010 International Conference on Multimedia Technology*, Ningbo, China, 2010.

[10] K. Özkan, E. Seke and Ş. Işık, " A new approach for speech denoising, " in *Signal Processing and Communication Application Conference (SIU), 2016 24th*, Zonguldak, Turkey, 2016.

[11] R. Choudhary, S. Som *(2016), "Encryption Technique Using Dynamic Table and Block Division Process on Binary Field" International Conference on "Reliability, Infocom Technologies and Optimizations (Trends and Future Directions) ICRITO 2016, 7-9 September 2016, IEEE Conference, indexed with* ***SCOPUS****, Amity University Uttar Pradesh, India, p.p. 353-358.*

[12] Shobha Tyagi, Subhranil Som, Qamar Parvez Rana *(2017) "Trust based Dynamic Multicast Group Routing ensuring Reliability for Ubiquitous Environment in MANETs", International Journal of Ambient Computing and Intelligence (IJACI), Volume 8, Issue 1, Scopus Indexed, ISSN: 1941-6237, DOI: 10.4018/IJACI, Pages 70 – 97, January – March 2017. (http://www.igi-global.com/journals/abstract-announcement/158348)*

[13] Seema Nath, Subhranil Som (2017), *"Security and Privacy Challenges: Internet of Things", Indian Journal of Science and Technology, Scopus Indexed,* ***included in 'Web of Science'*** *and included in the list of journal recommended by UGC, Vol 10(3), DOI: 10.17485/ijst/2017/v10i3/110642, ISSN (Print) : 0974-6846 ISSN (Online) : 0974-5645, January 2017.*