All Sciences Proceedings
http://as-proceeding.com/

1st International Conference on Pioneer and Innovative Studies

June 5-7, 2023 : Konya, Turkey

https://as-proceeding.com/index.php/icpis

Comparison of Energy Rate of a Level 3 Speech Signal using DWT with the Mother Wavelet Haar , meyer , coiflets , symlets , daubechies , biorthogonal

Abdelkader AMRAOUI[1], Slami SAADI[2]

[1]*Laboratory of Automation and Applied Industrial Diagnostics (LAADI), Faculty of Sciences and Technology*
[2]*Faculty of Exact Sciences & Computers, Ziane Achour University of Djelfa (UZAD), BP3117, Algeria*

kader2717yahoo.fr

Abstract – The objective of this work is to compare the energy rate of a level 3 speech signal using the discrete wavelet transform (DWT) of various mother wavelets such as Haar, Meyer, biorthogonal, Coiflets, Symlets, and Daubechies, to determine the range of values for both genders of the speaker. The speech signal is filtered by an 8th order Butterworth filter.

We perform the discrete wavelet transform (DWT) on the mother wavelet to obtain the energy values of each mother wavelet at the third level of the decomposition for male and female speakers. We then compare the energy values of speech signals between men and women speaking the same sentence, This comparison is conducted on several databases containing repeated sentences for the same individual or both genders, the results are found to be acceptable for further analysis.

Keywords – DWT; Energy; Wavelets

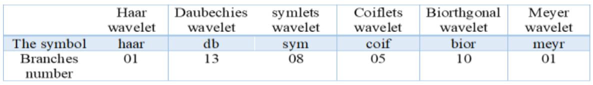
I. INTRODUCTION

Speech signal is a non-stationary, complex, and time-varying signal [01]. Furthermore, the highest frequency component in a speech signal is typically around 5 kHz. The differences in linguistic characteristics among individuals are characterized by variances in energy for both genders. Generally, speech signals have three main characteristics: pitch, fundamental frequency, and energy. The energy component, in particular, tends to be lower in women compared to men, indicating smaller energy levels in female speech signals.

II. MATERIALS AND METHOD

Experiments were conducted on 12 original signals extracted from a database prepared by us, in the form of audio sounds in WAV format. There were 9 signals from female speakers and 3 signals from male speakers, comprising different phrases and sometimes repeated for the same speaker or different speakers. We used a group of 6 mother wavelets and 38 branches, resulting in a total of 456 experiments to measure the percentage of retained energy between the original signal and the energy in the approximate coefficients of the third level of the discrete wavelet transform (DWT), which are divided according to the branches and mother wavelets as represented in Table (01).

Table (01): Wavelets used, and the number of branches for each one



	Haar wavelet	Daubechies wavelet	symlets wavelet	Coiflets wavelet	Biorthgonal wavelet	Meyer wavelet
The symbol	haar	db	sym	coif	bior	meyr
Branches number	01	13	08	05	10	01

It should be noted that in our experiments, the K4 Video Downloader program was used to download video clips of English phrases (audio clips) in MP4 format from the internet. These clips were then converted to (WAV) format using online converters such as (Cloudconvert en ligne) . The clips were further segmented into specified length phrases using the (Audacity 2.4.2) program, with a sampling frequency of fe =11025 Hz. These audio signals were analyzed and studied using the wavelet toolbox of (MATLAB a13) software on a computer with the following specifications:

PC Acer x67 , Processor : Intel® Core™ i3-2348M , CPU :2.30GHz, Version SMBIOS 2.7 , Operating system WIN 10.

III- SHANNON LAW

The operations performed on speech signals, such as distortion-free sampling and others, involve avoiding spectrum overlap and thus preventing information degradation. To achieve this, frequencies are required in accordance with Shannon's law. In order for the spectrum of the sampled signal not to overlap with the spectrum of the analog signal, it is necessary that (Fe - Fmax) be greater than Fmax, which leads to Shannon's theorem. Therefore, the sampling frequency should be at least twice the maximum frequency of the original signal under study. A low-pass filter is used with a cutoff frequency (Fc) of Fe/2 to ensure a bounded spectrum. The sampling of a signal is governed by Shannon's theorem, which can be expressed by the following relationship.

$$f_e \geq 2\, f_{max} \qquad (01)$$

f_e : The sampling frequency
f_{max} : The frequency of the original

IV- FILTERING

We use the widely used linear Butterworth low-pass filter with an order of n=8 and a slope of 48dB/decade, which is equivalent to 160dB/ décennie. This filter has a similar shape with a difference in the cutoff frequency band, set at f_c=600 hz , in order to reduce the power to 71% of the original signal (-3dB) equals -10 log_2$\boxed{f0}$10 . The amplitude of this filter is greater and more stable within the passband frequency. In the transition region, there is a moderate reduction. The selection of this filter is based on amplitude resolution. The order of the filter represents the number of columns in the filter's pass region [04], [05], [06].

V- THE DISCRETE WAVELET TRANSFORM (DWT)

The discrete wavelet transform (DWT), is a technique similar to the developed subband coding, which is named hierarchical coding [07], [08], [09], and known by the relation (02)[10] . The latter depends in its mathematical relation on convolution. [11] . DWT decomposes the original signal into detailed coefficients by passing the original signal x(n) through high-pass filters h(n) , and into approximate coefficients by passing the original signal x(n) through low pass filters. (DWT) use different frequencies, and different accuracies to analyze the original signal x(n) in different frequency bands, because the approximate coefficients have the highest amplitude, and low frequency components, and have the most of the energy, while the detailed coefficients have the lowest amplitude, and the highest frequency components, and the latter shares with the noise which is concentrated in the high frequency of the original signal [10].

$$X_{DWT}[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] * h[n-k] \qquad (02)$$

Figure (01) presents the wavelet analysis model:

Figure (01): Signal decomposition with DWT [12], [21]

DWT uses digital filtering techniques, where the signal to be analyzed is passed through filters with different frequencies, and different scales. The original signal is divided, into detail coefficients, and approximation coefficients, often denoted by ak and dk respectively, where d is a letter denoting the detail coefficients, and a denotes the approximate coefficients, while k denotes the analytical level, as shown in Fig (01), the two parameters are the result of the convolution of the original signal x(n) with the filter [13][05].

$$a1 = \sum_{k=-\infty}^{\infty} X(k)y[2n-k] \qquad (03)$$

$$d1 = \sum_{k=-\infty}^{\infty} X(k)z[2n-k] \qquad (04)$$

In waveform analysis, the signal is analyzed and synthesized through a step-by-step process that begins by dividing the signal x(n) until the desired level is reached. Then, the signal is reconstructed starting from the last level where we stopped, all the way back to the beginning of the first level. This process is performed in a reverse manner, hence vice versa, where the signal is synthesized by progressively combining the sub-signals obtained at each level until the original signal is reconstructed [05].

VI- VOCAL SIGNAL CHARACTERISTICS

1- FUNDAMENTAL FREQUENCY

The fundamental frequency, also known as the voice fundamental frequency, is the lowest frequency of a periodic sound that is perceived as the pitch of the voice. In the case of human voice, the fundamental frequency corresponds to the vocal cord vibration frequency during sound production. It is measured in Hertz (Hz) and is generally represented by the letter "f0." The fundamental frequency is an important parameter in the acoustic analysis of the voice and is used in various applications such as speech synthesis, speech recognition, and emotion analysis.

2 - AMPLITUDE

Amplitude is a measure of the energy of a wave, such as sound or light. For a sound wave, amplitude is related to the acoustic pressure of the wave, which corresponds to the variation in air pressure around its equilibrium state. Amplitude is usually measured in units such as pascal (Pa) for sound waves and volt (V) for electromagnetic waves. In the case of sound, amplitude can be used to determine the loudness or perceived volume level. The larger the amplitude, the louder the sound. Amplitude can also be used to determine the sound quality, especially for musical instruments, where amplitude can affect the richness and tonal color of the produced sound.

3 - ENERGY

Energy is represented by the intensity of sound, which is related to the air pressure upstream of the larynx. It is stronger for voiced sounds than for voiceless sounds. The energy of a signal is a measure of the amount of energy contained within the signal. It can be used to characterize a signal in the time domain. The energy of a signal can be used to determine whether the signal is of a transient nature (impulse signal) or a periodic nature (sinusoidal signal). A transient signal has finite energy, while a periodic signal has infinite energy. Calculating the total energy of a speech signal is not useful because we know that speech has a capacity and energy that vary over time. Therefore, what is important in the case of speech production is a tool that provides information about the energy evolution over time. Thus, a different method of speech processing is required. More specifically, the tools can always assume that the signal being processed is stationary [05], [14], [15], [16],[17]. The energy in a frame of duration N samples is given by:

$$E_X = \sum_{n=0}^{N-1} S^2 (n) \qquad (05)$$

X : level (DWT)

The energy of a signal is also useful for determining the average power of a signal. The average power of a signal is the amount of energy transmitted per unit of time. It is calculated by dividing the energy of the signal by the duration of the time period T. The average power of a signal is often used to characterize periodic signals such as audio signals and electrical signals.

VII - THE STUDY

To compare the energy present in the third level of the discrete wavelet transform (DWT) for all the mentioned mother wavelet branches for speech signals of both sexes, using the same spoken phrase "Do you speak English?", we followed the steps outlined in the flowchart (organigram) No. (01).

In the first step, the signal under study was prepared with a sampling frequency Fe = 11025 Hz, which is greater than twice the frequency of the speech signal, which is 5 kHz, as mentioned earlier. This frequency was chosen based on Shannon's law mentioned in equation (01) above.

In the second step, the signal was filtered using one of the Butterworth linear filters, which is the most commonly used filter with a similar shape but with a difference in the cutoff frequency range. This filter has the highest amplitude, more frequency stability, and a moderate reduction in the transition domain. The filter is selected based on the accuracy of the amplitude (35), and its filter order is n=8 with a cutoff frequency fc=600 Hz.

The third step involves applying the third level of the discrete wavelet transform (DWT) to the filtered signal using all the mother wavelets and their branches mentioned in Table (01). This is applied to different signals for speakers of both sexes. Finally, we extract the energy values present at this level for the signal under study, with the same phrase pronounced by two different sexes for comparison, as shown in the comparison table No. (02).

Flowchart (01): presents the study stages

VIII - PREVIOUS RESEARCH AND STUDIES:

Previous studies on signal processing using wavelet analysis have focused on comparing the most suitable wavelets. The comparison of the retained energy in the third level of DWT used in our research is similar to other criteria used in certain studies, such as the study on the best mother wavelet for speech signal processing [21], the study "A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing" [05], and the study on the best mother wavelet in speech signal processing, investigating the criterion of Percentage Root Mean Square Difference by Zhitao, Dong Youn, and William A (2000) [18]. These studies assess the differences and comparisons between the signals used.

Other relevant studies include the research by Tajane and Pitale (2014) [19] and Shweta and M.P (2013) [20], which compared different wavelets for analyzing ECG signals. The results of these studie, after comparing different mother wavelets, concluded that the best branches retaining energy are found in the third level of wavelet analysis.

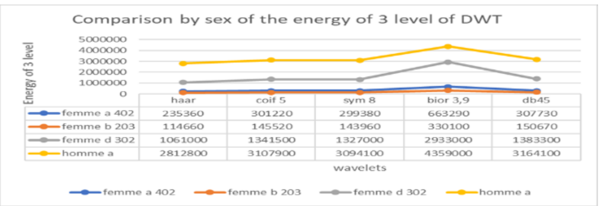
IX- RESULTS

Table (02): comparison of the energy values of a sentence (Do you speak English) of two sexes using The third level DWT

	Women a402	Women b203	Womend302	Man a102
haar	2.3536 e+05	1.1466 e+05	1.0610 e+06	2.8128 e+06
Coif 5	3.0122e+05	1.4552 e+05	1.3415 e+06	3.1079 e+06
Sym8	2.9938 e+05	1.4396 e+05	1.3270 e+06	3.0941 e+06
bior 3.9	6.6329 e+05	3.3010 e+05	2.9330 e+06	4.3590 e+06
Db45	3.0773 e+05	1.5067 e+05	1.3833 e+06	3.1641 e+06

Table (03):Comparing the highest energy across all branches of the wavelets using the third level of DWT for the same sentence (Do you speak English).

	Haar	meyer	daubechies	symlets	biorthogonal	coiflets
haar	haar					
Meyer		Myer				
db			db45			
sym				Sym8		
bior					bior3.9	
coif						Coif5



	haar	coif 5	sym 8	bior 3,9	db45
femme a 402	235360	301220	299380	663290	307730
femme b 203	114660	145520	143960	330100	150670
femme d 302	1061000	1341500	1327000	2933000	1383300
homme a	2812800	3107900	3094100	4359000	3164100

Figure (02): global comparison of a sentence (do you speak english) of sex 03 women (a ,b ,d) using DWT

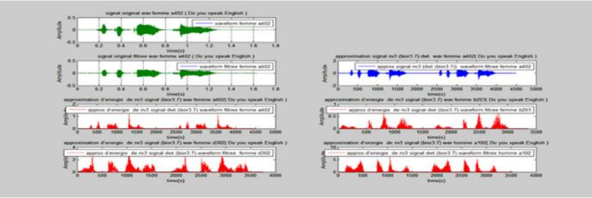


Figure (03): global comparison of a sentence (do you speak english) of sex 03 women (a ,b ,d) using DWT

X- CONCLUSION

Through the following experiments, we observe that the highest energy in the third level of discrete wavelet transform (DWT) varies from one mother wavelet to another and from one branch to another within the wavelets. We noticed that the highest energy in branch 8 for men in the DWT of the symlet mother wavelet is greater compared to all branches for both genders in the same sentence. On the other hand, the highest energy in the coiflets is found in branch 5 for men compared to other branches for both genders in the same sentence and the same third-level DWT. Furthermore, the energy in the biorthogonal mother wavelet is higher for men in branch 3.9 compared to the rest of the branches for both genders. For the highest energy in the daubechies mother wavelet, it is found in branch 45 for men compared to other branches for both genders in the same sentence and third level of (DWT). Based on the comparison table No. (02), we conclude that the energy for men is higher than that for women in the third level of discrete wavelet transform (DWT) for all branches and mother wavelets. This conclusion is drawn after comparing the same sentence for a group of women and men.

REFERENCES

[01] : AJGOU Riad . 11 /03 /2010. Techniques de détection de la période du pitch par lesméthodes temps fréquence et temps échelle.Université de Biskra

[02] :: https://www.elettroamici.org/fr/teorema-di-nyquist-shannon/.

[03] :https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_d%27%C3%A9chantillonnage

[04]: :https://fr.wikipedia.org/wiki/Fréquence_de_coupure

[05] : Abdelkader AMRAOUI1 , Slami SAADI2 . A Novel Approach for Speaker Gender Identification and Verification using DWT First Level Energy and Zero Crossing . Engineering, Technology & Applied Science Research Vol. 12, No. 6, 2022, 9570-9578 . p 9570 - 9578.

[06] : Par Art Pini , Avec la contribution de Rédacteurs nord-américains de Digi-Key , 24-03-2020 .Notions de base sur les filtres passe-bas antirepliement.

[07] : Deepika Sripathi - 2003 - Efficient Implementations of Discrete Wavelet Transforms UsingFpgas -- Florida State UniversityLibraries .

[08] : Polikar .R , 2006 -Second Edition– The Engineer'sUltimate Guide to WaveletAnalysis - the Wavelet Tutorial. . College of Engineering Web Servers: Rowan University.

[09] : Eva Hostalkova- ATHENS Nov 2009 - WaveletTransform-- Dept of Computing and Control Engineering- Institute of ChemicalTechnology, Prague .

[10] : Polikar .R , 2006 -Second Edition– The Engineer'sUltimate Guide to WaveletAnalysis - the Wavelet Tutorial. . College of Engineering Web Servers: Rowan University.

[11] : Karanpreet, K., & Ankush, K. (2014). Performance Analysis of Convolutional Interleaved DWT based OFDM system. International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 3(5), 6601-6604.

[12] : A. Sumithra M G , B. Thanushkodi K -- December,2009-- Performance Evaluation of DifferentThresholdingMethods in Time Adaptive WaveletBased Speech Enhancement , IACSIT International Journal of Engineering and Technology Vol.1,No.5,December,2009.

[13] : Karanpreet Kaur ^1, AnkushKansal ^2-- May 2014 -- Performance Analysis of ConvolutionalInterleaved DWT based OFDM system-- International Journal of Advanced Research in Computer and Communication Engineering-- Vol. 3, Issue 5, May 2014.

[14] : https://vlab.amrita.edu/?sub=3&brch=164&sim=857&cnt=1]

. [15] : Maripo Tsivery Tanjona .Soutenu le 23 Avril 2010 .Traitement du signal et la reconnaissance de la parole .

[16] : Ranganadh Narayanam* AssistantProfessor, Eece, Its, Ifhe – India . Narayanam * et al., 6(9): September, 2017 . voiced and univoiced separation in speech auditory brainstem responses of human subjects using zero crossing rate (zcr) and energy of the speech signal . Iternational journal of engineering sciences & research technology.

[17] : D.S.Shete , Prof. S.B. Patil ,Prof. S.B. Patil . Volume 4, Issue 1, Ver. I (Jan. 2014), PP 01-05 . IOSR Journal of VLSI and Signal Processing (IOSR-JVSP).Zerocrossing rate and Energy of the Speech Signal of Devanagari Script .

[18] : Zhitao, L., Dong Youn, K., & William A, P. (2000, July). Wavelet compression of ECG signals by the set partitioning in hierarchical trees (SPIHT) algorithm. IEEE Transactions on Biomedical Engineering, 47(7), 849 - 856.

[19] : Tajane, K. e., & Pitale, R. (2014, january). Review Paper :Comparative Analysis Of Mother Wavelet Functions With The ECG Signals. journal of Engineering Research and Applications, 4(1), 38-41.

[20] : Shweta, J., & M.P, P. (2013, october). Comparative Analysis of Various Wavelet Families used for R-Wave Detection of ECG Waveforms. International Journal of Innovative Research in Science,Engineering and Technology (IJIRSET), 2(10), 5177-5181.

[21] : Amraoui Abdelkader 1, Saadi Slami 2 . La Revue Numéro 6 . 2023 . Revue des sciences et technologies . Conseil Supérieur de la Langue Arabe . ISSN 2716-7674 .