International Journal of Applied Mathematics Electronics and Computers

ISSN:2147-8228 http://dergipark.org.tr/ijamec **Original Research Paper**

Selection of Optimum Mother Wavelet Function for Turkish Phonemes

Özkan Arslan[1], Erkan Zeki Engin*[2]

Received : 22/04/2019 Accepted: 11/07/2019

Abstract: In this paper, we propose the selection of most suitable mother wavelet function for Turkish phonemes using discrete wavelet transform. The determination of most similar mother wavelet function to the signal has been a challenge in speech processing. The optimum mother wavelet function for Turkish phonemes have been determined by using quantitative measures which were energy and Shannon entropy, information theoretic measures which were joint entropy, conditional entropy, mutual information, and relative entropy from wavelet coefficients of the phonemes. In this study, 101 potential functions were investigated to determine the most appropriate mother wavelet. Experimental results showed that the most appropriate wavelet functions for /ç/ and /ş/ phonemes which are unvoiced fricatives have been found as Reverse Bi-orthogonal 3.9 (rbio3.9) and Reverse Bi-orthogonal 5.5 (rbio5.5), respectively. By considering all the results, it was seen that the Bi-orthogonal 3.1 (bior3.1) and Discrete Meyer (dmey) wavelet functions were the most suitable mother wavelets for all other phonemes.

Keywords: *Discrete wavelet transform, energy, entropy, information theoretic measure, mother wavelet selection, Turkish phonemes*

1. Introduction

One of the most important characteristics which distinguish human beings from all other living being is the ability of communicating using speech sounds. The evaluation process of human voice is quite complicated due to the diversity of physical properties and complex nature of the speech signals. Phonetics objectively examines the quality of sound and physical characteristics without taking into account the meaning attributed to the words and how the voice organs act during the formation of voices [1]. Also, phonetics is divided into three main branches as articulating, acoustic, and auditory [2-3]. Phonetics determines the sounds (phonemes) necessary for defining the audible expressions of all the world's languages in an accurate and consistent manner. The phonemes are the smallest unit of sound that are used to separate words semantically and focus on the natural characteristics of sound in the languages. In addition to all these, the phoneme is not an inter-lingual unit based on an understanding and each language has its own phonemes.

The analysis of the speech signals is performed by acoustic analysis on the basis of phonemes. The acoustic analysis is defined as the study of speech signals in electronic media. The analysis is based on objective parameters which are fundamental frequency, pitch perturbation (jitter) and amplitude perturbation (shimmer) and the analysis can be easily repeated [4]. In recent years, the acoustic analysis parameters are commonly used in the evaluation of medical and surgical outcomes [5], speaker and speech recognition [6], speech analysis and synthesis, linguistic and phonetic knowledge acquisition [7], diagnosis of sound diseases [8], and planning of treatment and monitoring of treatment processes [9]. Wavelet transform which is a multi-resolution analysis technique is one of the most widely used methods in the acoustic analysis [10]. However, the basis function that is most appropriate to the structure of the signal should be determined. The basic features of wavelet basis functions such as regularity, symmetry, orthogonality, compact support, and the vanishing moments have been largely used in determining the proper wavelet for specific applications in various fields [11-12]. To determine the most suitable basis wavelet function, regularity and symmetry properties were used for analysing Auditory-Evoked Potentials (AEP) and the optimum mother wavelet function was found as Bi-orthogonal 5.5 [13]. Bhatia et. al. [10] reported that the Morlet, Gaussian, Paul-4, and B-spline wavelets were determined to be the most appropriate mother wavelet functions for event detection and segmentation in electrocardiogram (ECG) signals after taking into account reality, complexity, and orthogonality properties.

In studies on selecting the appropriate wavelet family, the use of only the basic features of wavelets is not reliable and it is difficult to apply to all wavelet families. In addition to these basic features, the quantitative and information theoretic measures can be used in order to determine the most appropriate wavelet. The quantitative measure which is the sum of absolute values of wavelet coefficients was used to choose the most appropriate mother wavelet for analysis of EEG and EMG signals [14-15].

In [16], the most appropriate main wavelet selection was investigated for various biomedical signal classification problems. In this study, genetic algorithm was used to optimize the determination of the best main wavelet function. In the experimental evaluation, the proposed solutions for the best main wavelet selection were compared with the manual hit-and-trial methods. According to the results, sym8 and db8 for EEG signals, sym6 and sym9 for ECG signals and bior2.6 and db8 wavelets for EMG signals were determined as the most suitable main wavelets. The results show that automatic main wavelet selection algorithm solutions are consistent with manual selection of wavelet functions.

In another study [17], the most suitable main wavelet selection was

[1] *Dept. of Electronics and Communication Eng., Çorlu Engineering Faculty, Tekirdağ Namık Kemal University, Tekirdağ, Turkey*
ORCID ID : 0000-0003-1949-3688
[2] *Dept. of Electrical &Electronics Eng., Engineering Faculty, Ege University, İzmir, Turkey*
ORCID ID : 0000-0002-2249-3139
** Corresponding Author: Email: erkan.zeki.engin@ege.edu.tr*

performed and automatic feature extraction system was proposed for gear and bearing fault diagnosis using wavelet based signal processing. 324 candidate main wavelets were studied in the selection of the main wavelet function and it was found that the db44 wavelet function had the most similar form in both gear and bearing vibration signals. The results also show that although the db44 wavelet is the most similar main wavelet function for vibration signals, it is not a suitable function for all wavelet-based processes [17]. In a similar study, to select the appropriate main wavelet for the analysis of the acoustic emission (AE) signal, the AE signal was obtained from the faulty control valve that had three leakage stages; small, critical and damaged. The analysis using Wavelet Packet Transform (WPT) is expected to reveal AE properties in different leaks. The selection criteria for the main wavelet to be used in WPT are based on the Energy / Shannon entropy ratio. Four main wavelet families were used for selection: Daubechies, Symlet, Biorthogonal and Coiflet. For further analysis, the highest main wavelet with energy/Shannon entropy ratio is selected [18].

In a study in which the most suitable wavelet function selection was used to classify hyperspectral image classification, two criteria were applied in the selection of optimal base wavelet for three wavelet types such as Daubechies (db), Symlet (sym) and Coiflet (coif). The energy criterion includes entropy factor and energy/Shannon entropy ratio, while the matching criterion works according to the correlation coefficients. It is recommended that coif1, db3 and db7 main wavelets should be used in hyperspectral image classification for both energy and shape criteria [19].

In a study of voice processing, /a/ vowels which were recorded from healthy subjects and patients with Unilateral Vocal Fold Paralysis were analysed using discrete wavelet transform [20]. They used energy and Shannon entropy measures to determine the suitable base wavelet and the appropriate mother wavelet function was found as Daubechies 6. Agbinya [21] intended to find the optimum wavelet for voiced-unvoiced detection, pitch determination, and speech compression. The most appropriate wavelet was determined by using the energy calculated from approximation coefficients and Battle-Lemarie wavelet function was found as the optimum base wavelet for speech analysis and synthesis. In another study, they used mean square error (MSE) measure between the original and reconstructed signals to determine optimal wavelet for speech signals [22]. They concluded that Discrete Meyer was the most suitable base wavelet for speech signals. In determining the optimum wavelet function for bearing vibration signals detection, they recommended the use of energy to Shannon entropy ratio and information theoretic measures [23]. According to the results of this study, Bior5.5 and rBior 5.5 wavelet functions were found as the most suitable base wavelet at Discrete Wavelet Transform (DWT), respectively. Also, complex Morlet was determined as the best function for bearing vibration signals detection using continuous wavelet transforms.

In this study, the most suitable mother wavelet functions for Turkish phonemes were determined by using DWT. Daubechies, Coiflet, Discrete Meyer, Symlet, Bi-orthogonal, and Reverse Bi-orthogonal wavelet families were selected as candidate wavelet functions. In order to determine the most appropriate wavelet function, the quantitative measures which are energy and Shannon entropy, information theoretic measures which are joint entropy, conditional entropy, and mutual information, and the relative entropy were computed by using the detail coefficients of DWT. Finally, as a key contribution of our work, it is aimed to analyse the speech signals more effectively by determining the appropriate wavelet functions for Turkish phonemes.

The rest of this paper is organized as follows. In Section 2, we present the recording procedure, a brief introduction to the wavelet transform, wavelet selection measures and procedure of mother wavelet selection. In Section 3, we present the experimental results for Turkish phonemes. Finally, Section 4 contains some concluding remarks.

2. Materials and Methods

2.1. Experimental Data

2.1.1. Subjects

Our subjects were 15 male and 15 female native Turkish speakers. Their ages ranged from 20 to 30 with an average of 25.33±2.31 for male and 24.13±2.5 for female. No history of hearing or speech disorders was known for any of the speakers.

2.1.2. Speech Stimuli

The choice of monosyllabic words for an audio corpus should be made considering that the corpus will represent the general phonetic properties of the language. One of the most important problems encountered while creating audio corpus is that there are too many monosyllabic words in Turkish language. Therefore, monosyllabic words which are frequently used in Turkish language are chosen when the audio corpus is designed. The Turkish alphabet consists of 29 letters with 8 vowels and 21 consonants. The alphabet lacks the /q/, /w/ and /x/ of English whereas it includes letters with a diacritic, such as /ç/, /ş/, and /ğ/. Each subjects uttered the monosyllabic words as given in Table 1. These monosyllabic words are commonly used in Turkish language. The monosyllable words have three phonemes which are initial consonant, vowel, and final consonants. Also, vowel acts as the core of words.

2.1.3. Recording Procedure

The recordings were made in a sound treated booth using a Shure PG58 microphone and a computer with Creative Sound Blaster THX sound card. The distance between microphone and mouth was set to 20 cm. The syllables were recorded at 44100 Hz sampling rate and 24-bit resolution. The list of syllables was in front of the subjects during the recording and there was approximately five seconds between each word.

Table 1. List of Turkish monosyllabic audio corpus

Turkish words	In English	Turkish words	In English
bağ	Link	maç	match
can	Soul	not	note
çok	Very	pul	stamp
diş	Teeth	rol	role
fen	Science	sır	secret
gün	Day	şen	cheerful
hak	Right	tıp	medicine
jön	Artist	var	exist
kat	Floor	yol	way
leş	Carrion	zor	difficult

2.2. Wavelet Transform

In recent years, the wavelet transform has been a signal processing technique widely used in several fields such as data compression, fingerprint verification, speaker and face recognition, DNA and protein analysis, blood pressure, heart rate and ECG analysis, and speech de-noising [24-25].

In particular, the Wavelet Transform (WT) is widely used in the analysis of non-stationary signals. In contrast to the Short-Time Fourier Transform (STFT), Wavelet Transform (WT) uses short windows at high frequencies and long windows at low frequencies [26]. This property is obtained by looking at the similarities between scaling and shift of a base wavelet. While scaling offers ideas about local regularity, time refers to the moment of the formation of wavelet. Also, time is an important factor that determines sudden changes occurring in the wavelet analysis [23]. Wavelet transform is divided into two groups according to the method of operation as Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The computation of the wavelet coefficients for each scale is difficult and time consuming in CWT. Therefore, DWT has been used more frequently for non-stationary signal analysis.

Discrete Wavelet Transform can be computed by using filters developed by [26]. This method, also known as Mallat algorithm, is a fast filtering algorithm with the two channels of sub-band coders. According to the Mallat algorithm, DWT is based on the principle of filtering the signal from high-pass and low-pass as seen in Fig. 1. In each scale, detail (D) and approximation (A) coefficients length are halved with frequency range. DWT is a special case of wavelet transform and can be expressed as:

$$W(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \psi\left(2^{-j}n - k\right) \tag{1}$$

where ψ, j, and k represents mother wavelet, scaling and time-shifting, respectively.

According to Mallat algorithm, approximation ($a_{j,k}$) and detail ($d_{j,k}$) coefficients consist of the inner product of scaling function $\phi(t)$ and wavelet function $\psi(t)$ with the signal, respectively. The approximation and detail coefficients can be expressed as:

$$a_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle = \int x(t)\, 2^{j/2} \phi^*\left(2^j t - k\right) dt \tag{2}$$

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = \int x(t)\, 2^{j/2} \psi^*\left(2^j t - k\right) dt \tag{3}$$

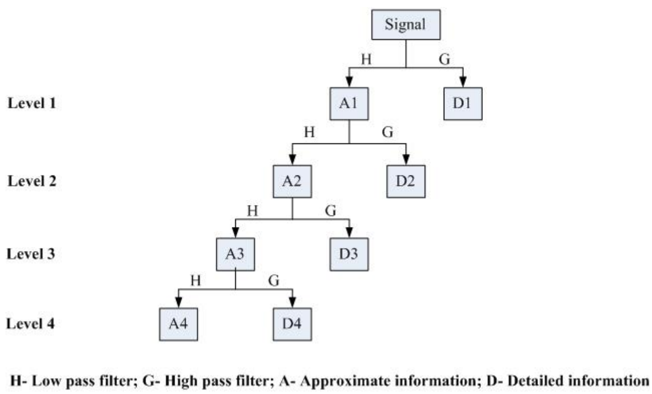

Fig. 1. Procedures of a four level signal decomposition using discrete wavelet transform

2.3. Wavelet Selection Measures

The optimum mother wavelet can be determined by using the properties of mother wavelet or similarity between the signal and mother wavelet. In wavelet analysis, the mother wavelets will give different results at analyzing the same signal. The mother wavelets are usually characterized by properties such as orthogonality, compact support, symmetry and vanishing moment. The properties of mother wavelet are considered in determining the optimum mother wavelets [27]. However, more than one mother wavelet with the same properties often exists. To overcome this problem, the similarity between the signal and mother wavelet are considered in determining the most suitable mother wavelet [28].

In this study, the quantitative measures which are energy and Shannon entropy, information theoretic measures which are joint entropy, conditional entropy, and mutual information, and the relative entropy were computed to select the most suitable mother wavelet.

2.3.1. Quantitative Measures

The quantitative measures such as energy and Shannon entropy can be used as well as qualitative measures in determining the optimum wavelet.

2.3.1.1. Energy

Energy is calculated by taking the sum of squares of wavelet coefficients in scales [29]. If the dominant frequency of the signal corresponds to a certain scale, the energy is high on that scale. The energy of a signal can be determined from the wavelet coefficients as expressed in (4).

$$E(s) = \sum_{i=1}^{N} |wt(s,i)|^2 \tag{4}$$

where N represents the number of the coefficients in scales.

2.3.1.2. Shannon Entropy

The energy distribution of the coefficients can be identified with entropy [23]. There are various way of measuring entropy such as Shannon, Minimum, and Collision. Shannon entropy is one of the most widely used types and is calculated as follows:

$$H = -\sum_{i=1}^{N} p_i \, log_2 p_i \tag{5}$$

where p denotes the probability distribution of the signal.

Shannon entropy is calculated according to two different methods of computing the probability distribution. In the first method, the entropy is calculated for each scale and the average entropy of scales is used to determine the optimum wavelet. In this method, the probability distribution is determined as follows [23]:

$$p_i(s) = \frac{|wt(s,i)|^2}{E(s)} \tag{6}$$

In the second method, the energy is calculated by (4) for each scale. Then, the probability of each scale is computed by dividing the energy of each scale to the total energy of scales [23]. The probability distribution can be defined as:

$$p(s) = \frac{E(s)}{\sum E(s)} \tag{7}$$

2.3.2. Information Theoretic Measures

Only wavelet coefficients are used in the computation of the energy and entropy measures. However, wavelet coefficients naturally depend on the analysed signal. Information theoretic measures that define the relationship between sequences are used in determining the optimum wavelet function. In information theoretic measures, the data sequence X expresses the voice signal. Similarly, the data sequence Y represents the wavelet coefficients in the related scales.

2.3.2.1. Joint Entropy

Joint entropy, $H(X,Y)$, describes the information between the signal and coefficients, and can be expressed as follows [30]:

$$H(X,Y) = \sum_{x\epsilon X}\sum_{y\epsilon Y} p(x,y)log_2 p(x,y) \tag{8}$$

where $p(x,y)$ represents the joint probability distribution of data sequences.

2.3.2.2. Conditional Entropy

The amount of information contained in the wavelet coefficients is measured by the conditional entropy $H(Y|X)$ when the probability distribution of the analysed signal is known [31]. The conditional entropy can be expressed as:

$$H(Y|X) = -\sum_{x\epsilon X} p(x)\sum_{y\epsilon Y} p(y|x)log_2 p(y|x) \tag{9}$$

In (9), $p(x)$ and $p(y|x)$ represent the probability distribution of the signal and the conditional probability distribution of the wavelet coefficients, respectively. The conditional probability can be defined as:

$$p(y|x) = \frac{p(x,y)}{p(x)} \tag{10}$$

As a result, conditional entropy is further defined as:

$$H(Y|X) = -\sum_{x\epsilon X}\sum_{y\epsilon Y} p(x,y)log_2 \frac{p(x,y)}{p(x)} \tag{11}$$

$$H(Y|X) = H(X,Y) - H(X) \tag{12}$$

2.3.2.3. Mutual Information

The mutual information, $I(X;Y)$, measures shared information between the analysed signal and wavelet coefficients [30]. The mutual information is defined as:

$$I(X;Y) = \sum_{x\epsilon X}\sum_{y\epsilon Y} p(x,y)log_2 \frac{p(x,y)}{p(x)p(y)} \tag{13}$$

$$I(X;Y) = -H(X,Y) + H(X) + H(Y) \tag{14}$$

2.3.2.4. Relative Entropy

The relative entropy defines the distance between probability distributions of voice signal and wavelet coefficients [31]. The relative entropy is expressed as:

$$D(X\|Y) = \sum_{x\epsilon X} p(x)log_2 \frac{p(x)}{p(y)} \tag{15}$$

The relative entropy is always non-negative and if the probabilities of the data sequences are equal, the relative entropy will be zero. The lowest relative entropy criteria should be used to determine the most appropriate mother wavelet [23].

2.3.3. Probability Estimation

Non-parametric density estimation is an important tool for statistical analysis of data. Non-parametric estimators can be used to evaluate data access to the multi-mode structure. The superiority of the non-parametric approach offers more flexibility compared to classical approaches for a given data set in modelling. Currently, the most popular non-parametric density estimation approach is the kernel density estimation (KDE) [32]. Kernel estimators are found with the kernel functions at each point of data and these estimators will smooth each data point from the local neighbourhood data points. The estimation is dependent on the shape and bandwidth of kernel function. The estimated density at any point is calculated as given in (16).

$$\tilde{f}(x) = \frac{1}{N}\sum_{i=1}^{N} K\left(\frac{x - x(i)}{h}\right) \tag{16}$$

where K and h denote kernel function and bandwidth, respectively. K is typically selected as a smoothed single-mode function. Although there are various kernel functions, Gaussian kernel functions are the most commonly used [33].
In this study, it is intended to select suitable mother wavelets for Turkish phonemes. For this purpose, the probability density functions of the wavelet coefficients and voice signals are calculated using KDE method. Also, two-dimensional KDE is calculated for the joint probability density function between the voice signals and the wavelet coefficients.

2.4. Procedure of Mother Wavelet Selection

The sampling frequency of the records used in this study is 44100 Hz and it has been reduced to 11025 Hz since the human ear responds to frequencies between 20 Hz and 20 kHz. The number of levels in the DWT was chosen as seven, taking into consideration the frequency range in Table 2. Then, the detail coefficients were selected to determine the optimum wavelet for phonemes.

Table 2. The frequency range according to the decomposition levels

Level (L)	Freq. Range (Hz)	Level (L)	Freq. Range (Hz)
1	2756-5512	5	172-343
2	1378-2755	6	86-171
3	689-1377	7	43-85
4	344-688		

The syllables were split into phonemes using voice activity detection (VAD) method. The voice activity was declared if the measured values higher or lower than the thresholds. The syllable signal was calculated by the short term energy and zero crossing rates for end point detection by VAD. If zero crossing rates was small and energy was high, we defined the signal as voiced, otherwise it was unvoiced. The phonemes were analysed by DWT for six different mother wavelet families. The phonemes were decomposed into seven levels by DWT and the quantitative measures, the information theoretic measures, and relative entropy were computed for six wavelet families. Then, the most suitable mother wavelet for each phoneme was determined according to the quantitative and information theoretic measures by taking the average of subject values. Daubechies, Symlets, Coiflets, Bi-orthogonal, Reverse Bi-orthogonal and Discrete Meyer wavelet families were examined in this paper. Each wavelet family could be expressed by its filter order. Bi-orthogonal wavelets can use filters with similar or dissimilar orders for decomposition (Nd) and reconstruction (Nr). The mother wavelet families used in this study are given in Table 3.

Table 3. Studied wavelet families

No.	Wavelet Family (short form)	Order
1-45	Daubechies (db)	db1-db45
46-60	Bi-orthogonal (bior)	bior1.1-bior6.8
61-75	Reverse Bi-orthogonal (rbio)	rbio1.1-rbio6.8
76-80	Coiflet (coif)	coif1-coif5
81	Discrete Meyer (dmey)	dmey
82-101	Symlet (sym)	sym2-sym20

Maximum energy, minimum Shannon entropy, minimum joint entropy, minimum conditional entropy, maximum mutual information, and minimum relative entropy criteria have been used in determining the best appropriate mother wavelet function belonging to six wavelet families for Turkish phonemes.

3. Experimental Results

In this paper, we have made an attempt to analyze various basis functions of discrete wavelet transform for Turkish phonemes. In the analysis, Shannon entropy was calculated according to two different methods of computing the probability. In the first method, the Shannon entropy values were found for each scale and then averaged for each subject. According to simulation results, it was observed that the average entropy values of subjects were close to each other in all phonemes. It has been difficult to decide the appropriate mother wavelet for the relevant phonemes from these entropy results. In the second method, the probability of each scale was computed by dividing the energy of each scale to the total energy of scales. The average entropy values of subjects for the phonemes yielded successful results in the determination of the appropriate wavelet function. Fig. 2 shows the average entropy values of subject which was computed by the second method for vowel /a/.

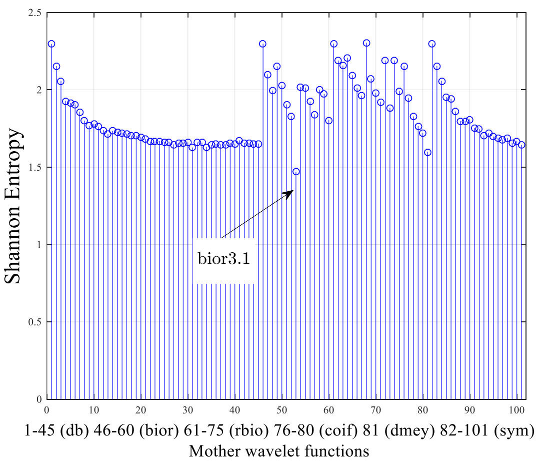

Fig. 2. The average Shannon entropy values of vowel /a/ by using the second method

According to the results obtained in our study, the joint entropy and conditional entropy values of the phonemes have not been considered because these entropy values belonging to six wavelet families were close to each other. Fig. 3 shows the energy values of vowel /e/ and consonant /t/. As can be seen from Fig. 3, it was observed that the most suitable mother wavelet for all phonemes of male and female subjects was found as bior3.1 wavelet function according to the energy measure criterion.

Fig. 4 shows the energy, Shannon entropy, relative entropy, and mutual information values of phoneme /a/. As can be seen from Fig. 4, it was inferred that the bior3.1 wavelet function was selected as the most appropriate wavelet function for vowel /a/ according to the high-energy, low-Shannon entropy, low relative entropy, and high mutual information criteria.

Similarly, Discrete Meyer (dmey) and Reverse Bi-orthogonal (rbio3.9) wavelet functions were selected as the optimum mother wavelet for phoneme /j/ and /ç/ as shown in Fig. 5 and Fig. 6, respectively. In Fig. 5, although the most appropriate mother wavelet was bior3.1 for phoneme /j/ according to the energy measure, the Discrete Meyer (dmey) mother wavelet was selected as the most appropriate wavelet function since the dmey wavelet for this phoneme was suitable according to other criteria. Similarly, as can be seen in Fig. 6, the rbio3.9 mother wavelet was selected as the most appropriate wavelet among all candidate wavelet families when all criteria were taken into account.

Finally, Table 4-8 present the most appropriate mother wavelets for all phonemes in Turkish language for male, female and the average of both gender when the quantitative, mutual information, and relative entropy criterion were considered.

Table 4. Optimal mother wavelet for Turkish vowel phonemes

Vowel Phonemes	Male	Female	Average
/a/	bior3.1	bior3.1	bior3.1
/e/	bior3.1	dmey	bior3.1
/ı/	bior3.1	dmey	dmey
/i/	dmey	dmey	dmey
/o/	bior3.1	bior3.1	bior3.1
/ö/	bior3.1	dmey	bior3.1
/u/	bior3.1	dmey	dmey
/ü/	dmey	dmey	dmey

Table 5. Optimal mother wavelet for Turkish nasal phonemes

Nasal Phonemes	Male	Female	Average
/m/	dmey	dmey	dmey
/n/	dmey	dmey	dmey

Table 6. Optimal mother wavelet for Turkish fricative phonemes

Fricative Phonemes	Male	Female	Average
/c/	dmey	dmey	dmey
/j/	dmey	dmey	dmey
/v/	bior3.1	bior3.1	bior3.1
/z/	bior3.1	dmey	dmey
/ç/	rbio3.9	rbio3.9	rbio3.9
/f/	bior3.1	bior3.1	bior3.1
/h/	bior3.1	bior3.1	bior3.1
/s/	bior3.1	bior3.1	bior3.1
/ş/	rbio5.5	rbio5.5	rbio5.5

Table 7. Optimal mother wavelet for Turkish plosive phonemes

Plosive Phonemes	Male	Female	Average
/b/	bior3.1	dmey	bior3.1
/d/	bior3.1	dmey	dmey
/g/	bior3.1	dmey	bior3.1
/p/	bior3.1	dmey	bior3.1
/t/	bior3.1	bior3.1	bior3.1
/k/	bior3.1	bior3.1	bior3.1

Table 8. Optimal mother wavelet for Turkish liquid phonemes

Liquid Phonemes	Male	Female	Average
/ğ/	bior3.1	dmey	bior3.1
/l/	dmey	dmey	dmey
/r/	bior3.1	dmey	bior3.1
/y/	dmey	dmey	dmey

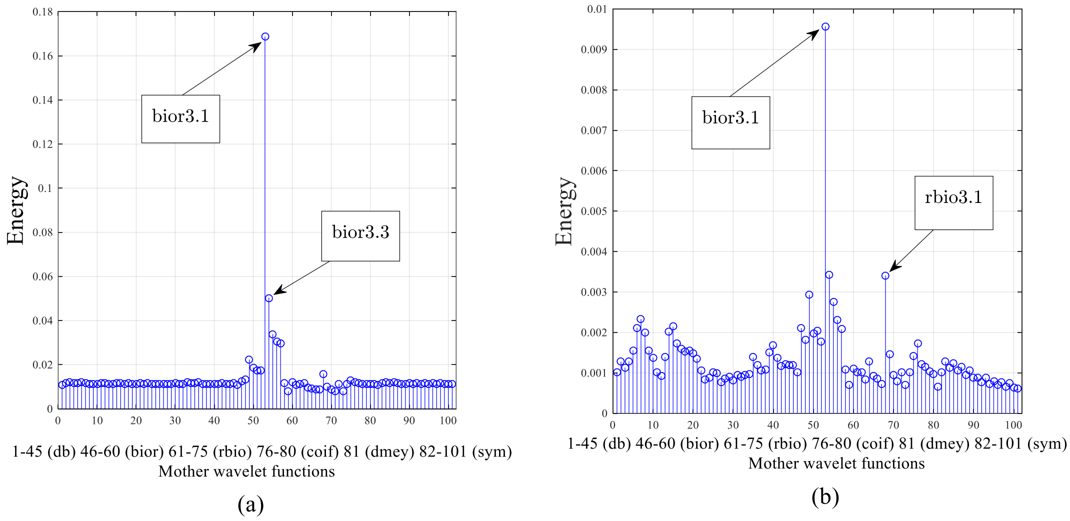

Fig. 3. The average energy values of: (a) vowel /e/, (b) consonant /t/

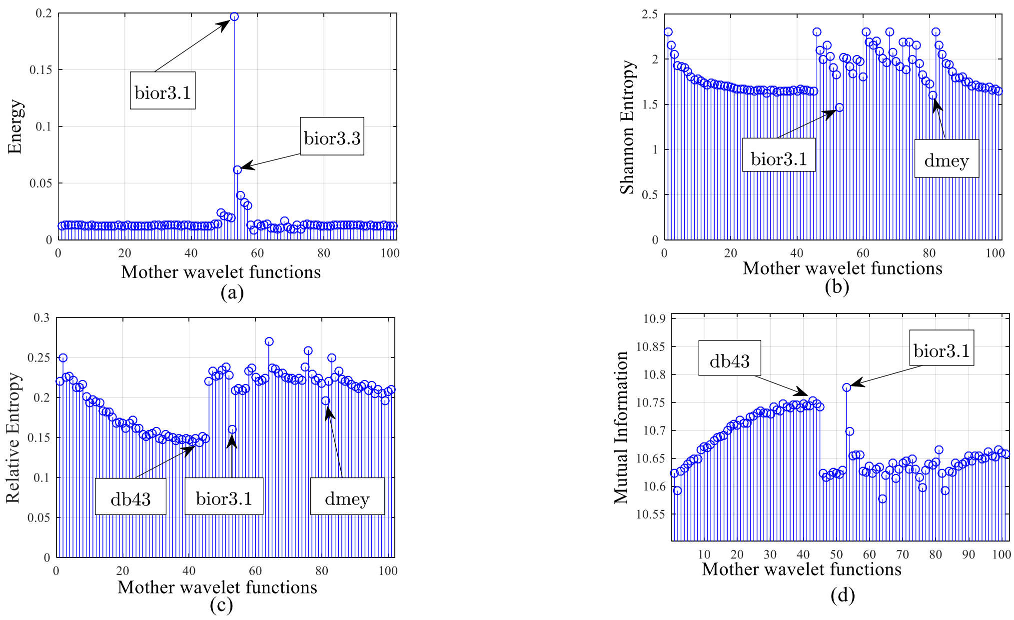

Fig. 4. The average values of vowel /a/: (a) Energy, (b) Shannon entropy, (c) Relative entropy, (d) Mutual information

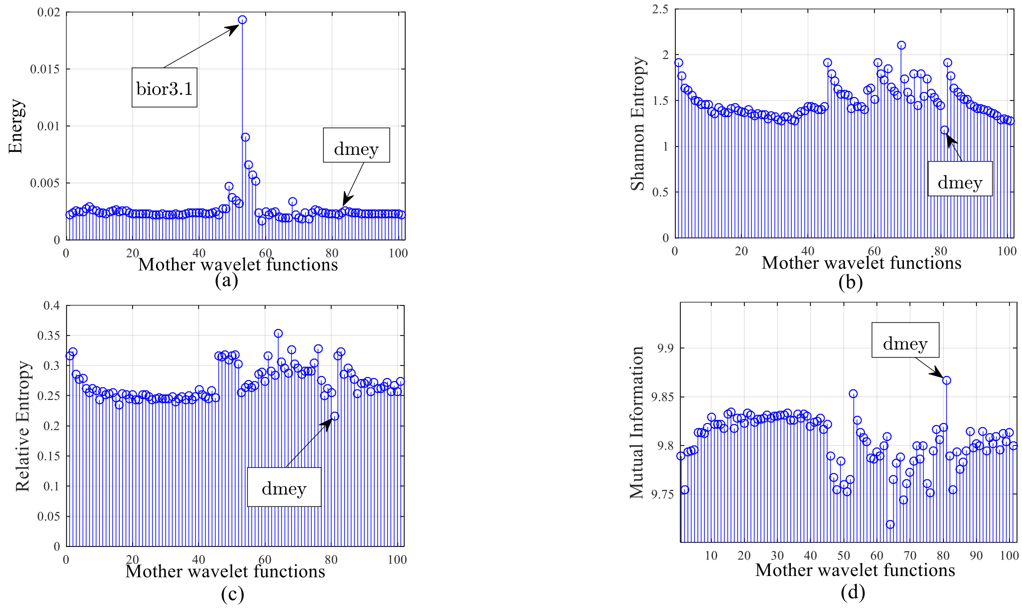

Fig. 5. The average values of consonant /j/: (a) Energy, (b) Shannon entropy, (c) Relative entropy, (d) Mutual information

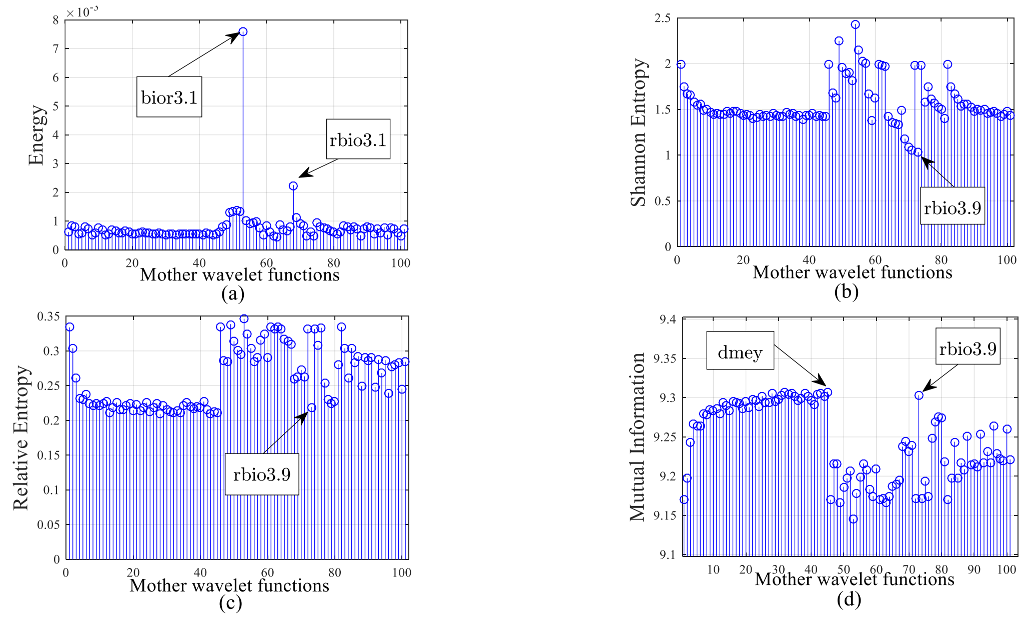

Fig. 6. The average values of consonant /ç/: (a) Energy, (b) Shannon entropy, (c) Relative entropy, (d) Mutual information

4. Conclusions

This study concentrated on the selection of the best mother wavelet function among 101 different wavelets for Turkish phonemes. The quantitative, information theoretic and relative entropy measure results were used to determine the best mother wavelet. Based on these selection criteria results, it was observed that the Bi-orthogonal (bior3.1) and Discrete Meyer (dmey) wavelet functions were the best candidate mother wavelets. The information theoretic measures which are joint and conditional entropy have not been significant difference among mother wavelets for Turkish phonemes. Therefore, these measures were not used to determine the appropriate wavelet. However, the mutual information and relative entropy have been observed that they could be used in determining the appropriate wavelet.

The determination of suitable base wavelet for phonemes was examined for three cases which take into account male, female, and their average. Based on the simulation results, it was observed that some phonemes have different optimum mother wavelet for male and female subjects. The most appropriate wavelet function has been determined as bior3.1 for most phonemes of male subjects while dmey has been determined for most phonemes of female subjects. In order to determine a common suitable wavelet for each phoneme objectively without gender consideration, the average of the criteria values for both male and female was taken and the dominant base wavelet was determined as the optimal wavelet function.

The experimental results showed that the optimal mother wavelet was not same for all Turkish phonemes. To overcome this problem, three suitable wavelets that yielded the best results for each measure of a phoneme were chosen as candidate mother wavelets. Final decision for the choice of mother wavelet was made among these sets of candidate wavelets which provide the majority of the high energy, low Shannon entropy, low relative entropy and high mutual information criteria. As can be seen in Table 6, rbio3.9 and rbio5.5 were determined as the most suitable base wavelets for /ç/ and /ş/ phonemes, respectively. These unvoiced fricative phonemes have different appropriate wavelets from the other phonemes and these differences may have occurred due to the articulation position of these phonemes.

So far, the most suitable wavelet function determination studies have been obtained depending on the application area and different conditions. The most appropriate wavelet function for vowel /a/ in English was determined as db6 in diseased voice analysis [20], while Battle-Lemarie and Discrete Meyer wavelet functions were determined as optimal wavelets in speech analysis and synthesis for English [21-22]. However, there is no study in the literature which selects the most appropriate wavelet function based on phoneme for all phonemes of Turkish language. Therefore this study will fill this gap.

Finally, in this study, it was aimed to select the optimum mother wavelets for all phonemes in Turkish. These optimum mother wavelets can be used for voice processing, speech and speaker recognition, speech enhancement and voice activity detection studies in Turkish language.

References

[1] K.N. Stevens, "Acoustic phonetics", vol. 30, MIT press, 2000.

[2] J. Laver, "Principles of phonetics", Cambridge University Press, 1994.

[3] J. Clark, C. Yallop and J. Fletcher, "An introduction to phonetics and phonology", Foreign Language Teaching and Research Press, 2007.

[4] S. Hadjitodorov and P. Mitev, "A computer system for acoustic analysis of pathological voices and laryngeal diseases screening", *Medical Engineering & Physics*, vol. 24, pp. 419–429, 2002.

[5] A. Acar, M. Cayonu, M. Ozman and A. Eryilmaz, "Changes in acoustic parameters of voice after endoscopic sinus surgery in patients with nasal polyposis", *Indian Journal of Otolaryngology and Head & Neck Surgery*, vol. 66, no. 4, pp. 381–385, 2014.

[6] Y. R. Oh, J. S. Yoon and H. K. Kim, "Acoustic model adaptation based on pronunciation variability analysis for non-native speech recognition", *Speech Communication*, vol. 49, no. 1, pp. 59–70, 2007.

[7] T. O. Lentz and R. W. Kager, "Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired", *Language and speech*, vol. 58, no. 3, pp. 387–413, 2015.

[8] J. J. G. Meil an, F. Martinez-Sanchez, J. Carro, D.E. Lopez, L. Millian Morell and J. M. Arana, "Speech in alzheimer's disease: Can temporal and acoustic parameters discriminate dementia?", *Dementia and Geriatric Cognitive Disorders*, vol. 37, no. 5-6, pp. 327–334, 2014.

[9] S.-M. Lee and J.Y. Choi, "Analysis of acoustic parameters for consonant voicing classification in clean and telephone speech", *The Journal of the Acoustical Society of America*, vol. 131, no. 3, pp. 197-202, 2012.

[10] P. Bhatia, J. Boudy and R. V. Andreao, "Wavelet transformation and pre selection of mother wavelets for ecg signal processing", in *Proc. of the 24th IASTED international conference on Biomedical Engineering*, 2006, pp. 390–395.

[11] Z. Chen, Y. Liu, Z. Liu and H. Tang, "The selections of wavelet function in singular signal detection", in *Proc. of the 2nd international conference on computer science and electronics engineering*, Atlantis Press, 2013.

[12] Z. Peng and G. Wang, "Study on optimal selection of wavelet vanishing moments for ECG denoising", *Scientific Reports*, vol. 7, no.1, pp.45-64, 2017.

[13] A. P. Bradley and W. Wilson, "On wavelet analysis of auditory evoked potentials", *Clinical neurophysiology*, vol. 115, no. 5, pp. 1114–1128, 2004.

[14] J. Rafiee, M. Rafiee, N. Prause and M. Schoen, "Wavelet basis functions in biomedical signal processing", *Expert Systems with Applications*, vol. 38, no. 5, pp. 6190–6201, 2011.

[15] C. Guo, J. Yu, L. Wu, Y. Liu, C. Jia, and Y. Xie, "Analysis and feature extraction of EEG signals induced by anesthesia monitoring based on wavelet transform", IEEE *Access*, vol. 7, pp. 41565-41575, 2019.

[16] G. Garg, "A signal invariant wavelet function selection algorithm", *Med Biol Eng Comput,* vol.54, pp.629–642, 2016.

[17] J. Rafiee, M.A. Rafiee, P.W. Tse, "Application of mother wavelet functions for automatic gear and bearing fault diagnosis", *Expert Systems with Applications*, vol.37, pp. 4568–4579, 2010.

[18] J.M. Juhani, and I. Rosdiazli, "Mother wavelet selection for control valve leakage detection using acoustic emission", IEEE *Conference on Systems, Process and Control (ICSPC)*, 2018, pp. 224-227.

[19] M. Abdolmaleki, M. Tabaei, N. Fathianpour, B.G.H. Gorte, "Selecting optimum base wavelet for extracting spectral alteration features associated with porphyry copper mineralization using hyperspectral images", *International Journal of Applied Earth Observation and Geoinformation*, vol. 58, pp.134–144, 2017.

[20] R. Behroozmand and F. Almasganj, "Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients speech signal with unilateral vocal fold paralysis", *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 474-485, 2007.

[21] J. I. Agbinya, "Discrete wavelet transform techniques in speech processing", in *Proc. IEEE TENCON*, 1996, pp. 514–519.

[22] Y. Long, L. Gang and G. Jun, "Selection of the best wavelet base for speech signal", in *Intelligent Multimedia, Video and Speech Processing*, 2004, pp. 218–221.

[23] R. X. Gao and R. Yan, "Wavelets: Theory and applications for manufacturing", Springer Science & Business Media, 2010.

[24] P. D. Swami, R. Sharma, A. Jain and D. K. Swami, "Speech enhancement by noise driven adaptation of perceptual scales and thresholds of continuous wavelet transform coefficients", *Speech Communication*, vol. 70, pp. 1–12, 2015.

[25] M. A. Oktar, M. Nibouche and Y. Baltaci, "Speech denoising using discrete wavelet packet decomposition technique", in *Signal Processing and Communication Application Conference (SIU)*, 2016, pp. 817–820.

[26] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[27] N. Ahuja, S. Lertrattanapanich and N. Bose, "Properties determining choice of mother wavelet", *IEE Proceedings Vision, Image and Signal Processing*, vol. 152, no. 5, pp. 659–664, 2005.

[28] W. K. Ngui, M. S. Leong, L. M. Hee and A. M. Abdelrahman, "Wavelet analysis: Mother wavelet selection methods applied mechanics and materials", Trans *Tech Publ*, pp. 953–958, 2013.

[29] H. Hong, T. Yonghong, Y. Wang, "Optimal base wavelet selection for ECG noise reduction using a comprehensive entropy criterion", *Entropy*, vol. 17, pp. 6093-6109, 2015.

[30] R.M. Gray, "Entropy and Information Theory", 2nd ed.; Springer: New York, NY, USA, 2011.

[31] T. M. Cover and J. A. Thomas, "Elements of information theory", John Wiley & Sons, 2012.

[32] Z. I. Botev, J. F. Grotowski, D. P. Kroese et al., "Kernel density estimation via diffusion", *The Annals of Statistics*, vol. 38, no. 5, pp. 2916-2957, 2010.

[33] J.N. Hwang, S.R. Lay and A. Lippman, “Nonparametric multivariate density estimation: a comparative study”, *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2795–2810, 1994.