An Analysis on Wavelet Applications as Speech Data Mining Tools

Senthil Devi K A[#1], Dr. Srinivasan B[*2]

[#1]Assistant Professor, Gobi Arts & Science College, Tamil Nadu, India.
[#2]Associate Professor, Gobi Arts & Science College, Tamil Nadu, India.

Abstract *Recently there has been significant development in the use of wavelet methods in various data mining and signal processing applications. Fourier Transform methods are not well suited for detection and classification of speech signals which possess non-stationary characters. It has been shown that wavelets can approximate time varying non-stationary signals in a better way than the Fourier transform representing the signal on both time and frequency domains. Furthermore, wavelet decomposition allows analyzing a signal at different resolution levels. This paper presents general overview of wavelets and their applications in speech as data mining tools. It first presents a data mining framework in which the overall process is divided into smaller components. It discusses the impact of wavelets on speech and data mining research.*

Keywords- *Wavelet theory, wavelet transformation, wavelet packet decomposition, multiwavelet, multiwavelet packet, data mining and speech processing.*

I. Introduction

The wavelet transform is a synthesis of ideas that emerged over many years from different fields. Generally speaking, the wavelet transform is a tool that partitions data, functions, or operators into different frequency components and then studies each component with a resolution matched to its scale [1]. Nowadays many fast and efficient software packages are developed to perform wavelet transforms. Wavelets have quickly gained popularity among scientists and engineers, both in theoretical research and in applications due to its easy accessibility.

Wavelet theory could naturally play an important role in data mining because wavelets could provide data presentations that enable efficient and accurate mining process and they can also could be incorporated at the kernel for many algorithms [1]. In this paper, we present a general overview of wavelet methods in data mining and speech processing with relevant mathematical foundations and of research in wavelets applications.

II. A FRAMEWORK FOR SPEECH DATA MINING PROCESS

Speech data mining is the process of searching and analysing the contents of huge speech data for identifying patterns and associations, retrieving keywords and useful information. Data mining is an iterative process which consists of the following 4 major components: data management, data preprocessing, core mining process and post-processing [1]. In data management, the mechanism and structures for accessing and storing data are specified. The subsequent data preprocessing step is an important step, which ensures the data quality and improves the efficiency and ease of the mining process. Real-world data tend to be incomplete, noisy, inconsistent high dimensional and multi-sensory etc. and hence are not directly suitable for mining. Data preprocessing includes data cleaning to remove noise and outliers, data integration to integrate data from multiple information sources, data reduction to reduce the dimensionality and complexity of the data, and data transformation to convert the data into suitable forms for mining. Core mining refers to the essential process where various algorithms are applied to perform the data mining tasks. The discovered knowledge is refined and evaluated in post-processing stage.

Speech has become one of the most important human communication media. Due to advancements in recent technology, very large quantities of speech can be collected and stored digitally. This large volume of speech cannot be efficiently reviewed by human beings. New technologies are required to provide intelligent access to obtain potential knowledge from this large speech collection. Speech data mining is defined as the nontrivial extraction of hidden and useful information from the masses of speech data.

Wavelet theory has become one of the most important and powerful tool of signal representation [1]. Wavelet transforms are suitable tools for analysing non-stationary signals. Since speech consists of both high and low frequency components, short and long duration sounds, the wavelet transform is well suited to this type of analysis. In many speech related applications, wavelets are used to perform preprocessing (e.g. noise filtering), dimensionality reduction and data transformation.

III. WAVELET THEORY ANALYSIS

Wavelet theory provides a unified framework for a number of techniques that have been developed in various signal processing applications. For example, Multi resolution signal processing, subband coding, speech and image compression. Wavelet Theory can be used to improve Speech processing performance through two approaches. In the first approach, it can be used as back-end to remove noises. In the second approach, wavelet-based features can be added to other successful features to improve speech processing [3]. In this section, bases of wavelet theory are discussed.

A. What is Wavelet?

A wave is an oscillating periodic function of time or space. In contrast, wavelets are localized waves. They have their energy concentrated in time or space and are suited to analysis of transient signals. While Fourier transform uses waves to analyse signals, the wavelet transform uses wavelets of finite energy. The plots wave and wavelet are shown in the Fig. 1. The wavelet theory deals with the properties of wavelets.

Fig.1. Plots of Wave and Wavelet

B. Fourier Transform Vs. Wavelet Transform

Fourier transform (FT) is used to represent a signal as the sum of a series of sines and cosines. While the FT is useful for analysing the spectral content of a stationary signal and it transforms difficult operations into very simple ones in the Fourier dual domain [4]. But it cannot be used for the analysis of non-stationary signals or for real time applications. The main disadvantage of a Fourier Transform however is that it has only frequency resolution and no time resolution. This means that although FT can be able to determine all the frequencies present in a signal, it does not know when they are present. To overcome this problem in the past decades several solutions have been developed which are more or less able to represent a signal in the time and frequency domain at the same time. Wavelet theory extends the ideas of the traditional Fourier theory. The wavelet transform or wavelet analysis is the most recent solution to overcome the shortcomings of the Fourier transform. Fig. 2a shows a sinusoidal signal. Its corresponding Fourier Transform and Wavelet Transform are shown in Fig. 2b.

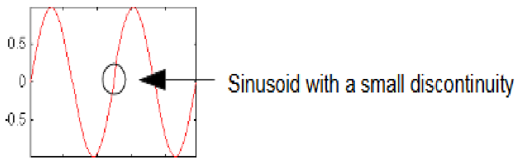

Figure 2a. A signal

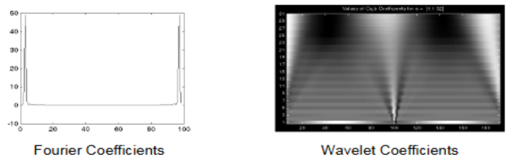

Fig. 2b. Plots of Fourier Transform and Wavelet Transform of the above signal.

C. Wavelet Transform

The wavelet transform is a type of signal transform that is commonly used in speech compression. The wavelet transform can be viewed as transforming the signal from the time domain to the wavelet domain. This new domain contains more complicated basis functions called wavelets, mother wavelets or analysing wavelets [2]. A wavelet prototype function at a scale s and a spatial displacement u is defined as:

$$\Psi_{u,s}(t) = 1/\sqrt{s}\,\Psi\left((t-u)/s\right) \quad (u \in \mathbb{R},\ s \in \mathbb{R}^{*}_{+})$$

This localization feature, along with wavelets localization of frequency, makes many functions and operators using wavelets sparse when transformed into the wavelet domain. This sparseness, in turn results in a number of useful applications such as data compression and detecting features in signals.

D. Continuous Wavelet Transform

The Continuous Wavelet Transform (CWT) is used to decompose a signal into wavelets, small oscillations that are highly localized in time. Whereas the Fourier transform decomposes a signal into infinite length sines and cosines, effectively losing all time-localization information, the CWT's basis functions are scaled and shifted versions of the time localized mother wavelet. The CWT is used to construct a time-frequency representation of a signal that offers very good time and frequency localization [6].

E. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently [7]. For many signals, the low-frequency part contains the most important part. It gives an identity to a signal. Consider the human voice. If we remove the high-frequency components, the voice sounds different, but we can still tell what's being said. In wavelet analysis, we often speak of approximations and details. The approximations are the high scale, low frequency components of the signal. The details are the low-scale, high frequency components [8]. The DWT is defined by the following equation:

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

where ψ (t) is a time function with finite energy and fast decay called the mother wavelet. The DWT analysis can be performed using a fast, pyramidal algorithm related to multirate filterbanks. In the pyramidal algorithm the signal is analysed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step. This is achieved by successive high pass and low pass filtering of the time domain signal and is defined by the following equations:

$$y_{\mathrm{low}}[n] = \sum_{k=-\infty}^{\infty} x[k]\, g[2n-k]$$

$$y_{\mathrm{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]\, h[2n-k]$$

F. Multi-Resolution Analysis (MRA) using Filter Banks

Filters are one of the most widely used signal processing functions. Wavelets can be realized by iteration of filters with rescaling. the filtering operation is used to determine the resolution of the signal, which is a measure of the amount of detail information in the signal and the scale is determined by upsampling and downsampling operations[8]. The DWT is computed by successive lowpass and highpass filtering of the discrete time-domain signal as shown in Fig. 3 and Fig. 4.

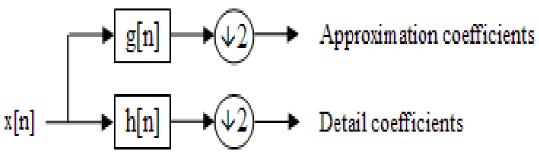

Fig. 3. DWT high pass and lowpass filters

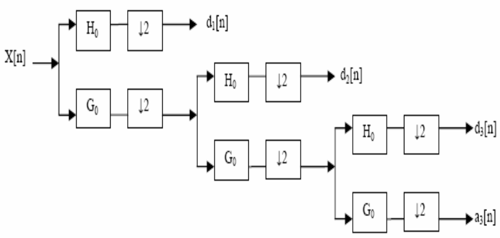

Fig. 4. 3 level Wavelet Decomposition

G. Wavelet Packet Decomposition

A more comprehensive form of the standard wavelet transform is the wavelet packet, which decomposes both the high and low frequency bands at each level [8]. A pair of low- and high pass filters is used to recognize two sequences capturing dissimilar frequency sub-band features of the original signal. These sequences are then decimated. This process can be repeated to partition the frequency spectrum into smaller frequency bands for obtaining different features while detecting the temporal information. Wavelet packet atoms are waveforms indexed by three naturally interpreted parameters, position, scale, frequency. WPT features have better presentation than the DWT. The wavelet packet decomposition is shown in the Fig. 5.

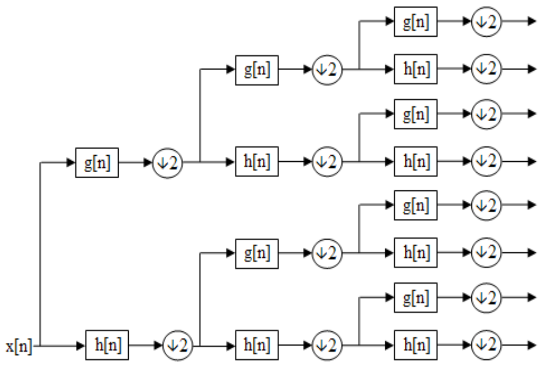

Fig. 5. 3 level Wavelet Packet Decomposition Tree

H. Multiwavelet Transform

A newer alternative to the wavelet transform to represent a signal is the multiwavelet transform. Multiwavelets are very similar to wavelets but have some important differences. In particular, whereas wavelets have an associated scaling function and wavelet function, multiwavelets have two or more scaling and wavelet functions [8]. Instead of one scaling function and one wavelet, multiple scaling functions and wavelets are used. This leads to a more degree of freedom in constructing wavelets. A single level of standard wavelet decomposition splits the input signal into lowpass and highpass coefficients through filtering and downsampling. The fundamentals properties of signal such as compact support, orthogonally, symmetry, vanishing moments and short support can be gathered simultaneously in multiwavelets.

I. Multiwavelet Packet

Just as with scalar wavelets, the multiwavelet filter bank procedure involves iterating the filtering operations on the low pass channel of the filter bank. And just as with scalar wavelets, iterating on the high pass channel as well can produce new basis function. Multiwavelet packet combines the wavelet packet decomposition with multiwavelet filter [9]. The construction of multiwavelet packet is similar to that wavelet packet.

IV. WAVELET FAMILIES

There are several types of wavelet families whose qualities vary according to several criteria such as: the support of the mother wavelets, the symmetry, the number of vanishing moments and the regularity. These criteria are associated with two properties: the existence of a scaling function and the orthogonality or the biorthogonality of the resulting analysis. Some of the wavelets are discussed below.

A. Haar wavelet

The Haar wavelet is a sequence of rescaled, square-shaped functions which together form a wavelet family or basis. The wavelet is the simplest type of wavelet . In discrete form, Haar wavelets are related to a mathematical operation called the Haar transform. The Haar transform decomposes a discrete signal into two sub signals of half its length. A HT decomposes each signal into two components, one is called average (approximation) and the other is known as difference (detail) [10], [11]. The technical disadvantage of the Haar wavelet is that it is not continuous. This property can, however, be useful for the analysis of signals with sudden transitions, such as monitoring of tool failures in machines. A plot of Haar wavelet is shown in Fig. 6.

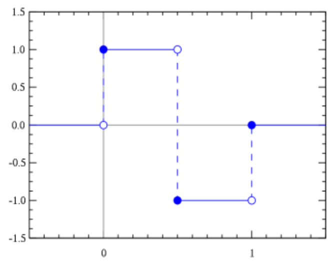

Fig.6. A plot of Haar wavelet

B. Daubechies Wavelets

The Daubechies wavelet is one of the popular wavelets and has been used for speech recognition [3]. Daubechies wavelet family is the most widely used orthogonal wavelet family and is named in the honour of its inventor, the Belgian physicist and mathematician Ingrid Daubechies. They represent a collection of orthogonal mother wavelets with compact support, characterized by a maximal number of vanishing moments for some given length of the support. Corresponding to each mother wavelets from this class, there is a scaling function (also called father wavelet) which generates an orthogonal MRA. They provide us with a set of powerful tools for performing basic speech processing tasks. These tasks include compression and noise removal for audio signals and speech recognition.

C. Symlets

Symlets (symN, where N is the order), also known as Daubechies least asymmetric mother wavelets. They are compact supported, orthogonal, continuous, but only nearly symmetric mother wavelets. The purpose was to create wavelets with the same size and same number of vanishing moments as Daubechies, but with near linear phase filters. Symlets have the highest number of vanishing moments for a given support width. Symlets have N/2 vanishing moments, support length N - 1 and filter length N.

D. Coiflets

Coiflets (coifN, where N is the order) are discrete wavelets designed by Ingrid Daubechies and named in the honor of Ronald Coifman who was another researcher in the field of wavelets theory. Ronald Coifman suggested the construction of a orthonormal wavelets family with the same number of vanishing moments as the scaling functions they came from. Coiflets are compactly supported wavelets and were designed to be more symmetrical than Daubechies mother wavelets to have a support of size N -1 and filter length N. The number next to the wavelet's name represents the number of vanishing moments, related to the number of wavelet coefficients.

E. Biorthogonal

Biorthogonal family of wavelets exhibits the property of linear phase, which is needed for signal and image reconstruction [2]. By using two wavelets, one for decomposition and the other for reconstruction instead of the same single one, interesting properties are derived.

F. Morlet

The Morlet wavelet (or Gabor wavelet) is a wavelet composed of a complex exponential multiplied by a Gaussian window. This wavelet is closely related to human perception, both hearing and vision.

V. WAVELET APPLICATIONS IN SPEECH MINING

Real world data sets are usually not directly suitable for performing speech data mining algorithms. They contain noise, missing values and may be inconsistent. In addition, real world speech corpus tend to be too large and high-dimensional. Wavelets provide a way to estimate the underlying function from the speech. By retaining selective wavelet coefficients, wavelet transform could then be applied for various speech processing such as noise filtering and dimensionality reduction. Moreover, since wavelet coefficients are generally decorrelated. We could transform the original data into wavelet domain and then carry out speech data mining tasks.

A. Denoising

Removing noise from data can be considered as a process of identifying outliers from available noisy data [12], [13]. Wavelet techniques provide an effective way to denoise and have been successfully applied in various areas especially in speech processing.

B. Dimensionality Reduction

The goal of dimension reduction is to express the original speech content using some smaller set of speech content with or without a loss of information [1]. Dimensionality reduction can be thought as an extension of the data transformation. While data transformation just transforms original data into wavelet domain without discarding any coefficients, dimensionality reduction only keeps a collection of selective wavelet coefficients.

C. Clustering

The multi-resolution property of wavelet transforms inspires the researchers to consider algorithms that could identify clusters at different scales in speech content. From a signal processing perspective, Applying wavelet transform on a speech decomposes it into different frequency sub-bands [1]. Hence identifying clusters is converted to finding connected components in the transformed feature space. Moreover, application of wavelet transformation provides multi resolution data representation and hence finding the connected components could be carried out at different resolution levels.

D. Speech classification

Classification problems aim to identify the characteristics that indicate the group to which each instance belongs. The classification methods can be applied on the wavelet domain of the original speech or selective dimensions of the wavelet domain. The wavelet based speech classification methods are mostly used in Speaker Identification. This classification method achieves a significant speedup over traditional speaker classification methods [3].

E. Regression

Regression uses existing values to forecast what other values will be and it is one of the fundamental tasks of data mining. The basic approach of using wavelets for nonparametric regression is to consider the unknown function expanded as a generalized wavelet series and then to estimate the wavelet coefficients from the data [1]. Hence the original nonparametric problem is thus transformed to a parametric One.

F. Similarity Search

The problem of similarity search in data mining is to find similar patterns in the data set based on some similarity measures. This task is most commonly used for time series, speech, image and text data sets. The similarity search conducted on the wavelet transformed domain could be more efficient. Wavelets also have widespread applications in content-based similarity search in image or audio databases. Keyword spotting is one of the content-based similarity search in speech mining which is able to localize the multiple occurrences for a word in continuous speech signal of any length [14].

G. Speech segmentation

In many speech processing applications, the speech signals are segmented using constant time segmentation. Constant segmentation needs to use windows to decrease the boundary distortions [15], [16]. A more natural approach is to segment the speech signals on the basis of time-frequency analysis. Boundaries are assigned in places where some energy of a frequency band rapidly changes. Wavelet transformation of speech is an efficient way to perform time-frequency based segmentation. In wavelet based segmentation, the energy in different frequency subbands gives an excellent opportunity to distinguish the beginning and the end of phonemes.

VI. CONCLUSIONS

Wavelet techniques have a lot of advantages and there already exists numerous successful applications in data mining and signal processing. The object of this paper is to increase familiarity with basic wavelet applications in speech as data mining tools. General overview of wavelets and their applications used in speech and data mining are discussed. The impact of wavelets on speech data mining research is clearly discussed. Hence that it is concluded that wavelet techniques have a lot of advantages and there already exists numerous successful applications in speech data mining. It goes without saying that wavelet approaches will be of growing importance in speech data mining.

REFERENCES

1. Tao Li, Sheng Ma, Mitsunori Ogihara, Wavelet methods in data mining, Chapter 27, Data Mining and Knowledge Discovery Handbook , Springer, 2005, pp 603-626.
2. Olivier Rioul and Martin Vetterly, *Wavelet and Signal Processing*, IEEE Signal processing magazine, October 1991.
3. M. H. Farouk, *Applications of Wavelets in Speech Processing*, Springer Briefs in Electrical and Computer Technology, 2013.
4. Cristina Laura Stolojescu, *A Wavelets Based Approach for Time Series Mining*, Ph.D thesis, Politehnica University of Timisoara, 2011.

5. Michel Misiti, Yves Misiti, Georges Oppenheim, Jean-Michel Poggi, *Wavelet Toolbox Computing Visualization Programming User's Guide*, version 1, 1996.
6. Salhi L, Talbi M, and Cherif A, *Voice Disorders Identification using Hybrid Approach: Wavelet Analysis and Multilayer Neural Networks*, World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:2, No:9, 2008
7. George Tzanetakis, Georg Essl, Perry Cook, *Audio Analysis using the Discrete Wavelet Transform Organized sound*, Vol. 4(3), 2000.
8. Vimal Krishnan V.R, Babu Anto P, *Features of Wavelet Packet Decomposition and Discrete Wavelet Transform for Malayalam Speech Recognition*, International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009.
9. Jian Ping Li, Stephane Jaffard, C Y Suen, John Daugman, Victor Wickerhauser, Bruno Torresani, John Yen, Ning Zhong, Sankar K Pal, *Wavelet Analysis amd Active Media Technology*, Proceedings of the International Computer Congress 2004, Logistical Engineering University, P R China, 28 – 30 May 2004.
10. S.S. Tamboli1, Dr. V. R. Udupi., *Image Compression using Haar Wavelet Transform*, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 8, 2013.
11. Kannan K., Arumuga Permal, Arumuga Perumals, Arulmozhih , *Area level fusion of Multi-focused Images using Multi-Stationary Wavelet Packet Transform,* International Journal of Computer Applications (0975 – 8887) Volume 2 – No.1, May 2010.
12. Stephane Mallat, *A Wavelet Tour of Signal Processing a sparse way*, book published by Academic Press, October, 2008.
13. Farooq O, Datla, *Wavelet – based denoising for robust feature extraction of speech recognition*, Electron, lett. 39(1), 163-165, 2003.
14. Senthil devi K.A., Dr.Srinivasan B, *A novel Keyword Spotting Algorithm in speech mining using wavelet*", International Journal of Current Research Vol. 8, Issue, 08, pp.36943-36946, August, 2016.
15. B.T. Tan, M. Fu, A. Spray, P. Dermody, *The use of wavelet transform for phoneme recognition*, Proceedings of the 4th International Conference of Spoken Language Processing Philadelphia, Vol. 4, USA, October 1996, pp.2431-2434
16. Bartosz Siolko, Suresh Manandhar, Richard C.Wilson and Mariusz Ziolko, *Wavelet Method of Speech Segmentation*, IEEE Xplore, Proceeding in Signal Processing Conference, 2006.