International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 12 (2018) pp. 10886-10896

A Study on Applications of Wavelets to Data Mining

Zaheer Abbas[1] and Princess Raina[2]

[1] *Department of Mathematical Sciences, Baba Ghulam Shah Badshah University Rajouri. Jand K, India.*

[2] *Department of Mathematical Sciences, Baba Ghulam Shah Badshah University Rajouri. Jand K, India.*

Abstract

Since the day wavelet came into the picture, the field of wavelets have been growing on both the theoretical and applications fronts. Among many areas where wavelets have been used to study real life problems, one is "Data Mining". Wavelets have been used, for many years now, to study various aspects of data mining. In this paper, after introducing fundamental ideas connected with wavelets and data mining, we shall present a survey of applications of wavelets to various aspects data mining.

Keywords: Wavelet transform, MRA, Data mining, Data management, Denoising, data transformation, dimensionality reduction, clustering, classification, similarly search , approximate query processing, visualization, neural network, principle/independent component analysis;

WAVELET ANALYSIS

If we want to study matter, the process becomes simpler if we start understanding quarks, then atoms, then molecules and finally matter. Similarly properties of integers can be more understood if we start working with primes and then proceed to analyse integers. Similarly one can understand the organisms if one firstly understands cells. As a general rule: the process of understanding becomes easier if complicated structures are synthesized by using simpler ones". In the first quarter of the 19^{th} century Jean Baptiste Joseph Fourier, while studying the problem of heat diffusion in a continuous medium, did the same thing by introducing the concept of Fourier series. In fact, he claimed that "every periodic function can be expressed as a weighted sum of sines and cosines. Initially, his idea could not appeal to his contemporary mathematicians (Laplace, Langrange, Lacroix and Monge etc.) as it was lacking rigor, in particular with respect to the convergence of the series. But later on due to the efforts of many other mathematicians (Poisson, Cauchy, Drichlet, Jorden etc.), Fourier's idea was made mathematically strong at all fronts and in the process many new concepts(Riemann integral, Lebesgue integral etc) either popped up or were made more rigorous(see [12]). Meanwhile, many related ideas such as Fourier transform, discrete Fourier transform, Fast Fourier transform also came up during the course of time and among these it is been said that "Fast Fourier Transform" is one of the best algorithm of all times that science and technology has ever witnessed. It was proposed by Cooley and Tukey in 1965 **[**29**]** and due to this algorithm "Fourier Transform" is known as "king of all transforms". Fourier analysis finds its application in many areas of science and technology such as, Electrical engineering, Crystallography, Telephone , X −ray machines, Harmonic signals, Quantum mechanics, Wave motion, Turbulence, Analysis of stationary signals, Real time signal processing. Although, Fourier analysis finds tremendous applications in science and technology but are inadequate to handle non-stationary signals mainly because of the two reasons: (i) Sinusoids does not have compact support and (ii) Fourier representation provides spectral content without the time localization. This lead to the engagement of the mathematicians for the modifications in Fourier analysis to make it able to handle non-stationary signals. As a result many windowed Fourier transforms such as "Short time Fourier transform (STFT)", Wigner-Ville transform, Cohen distribution, Wigner-Ville –Rihaczck distribution were introduced but none could handle the problem of non stationary signals which have either low frequency components with long time spans or high frequency components with short time spans efficiently, because they all use a single window to analyse the entire signal. Finally, the problem of handling non-stationary signals got solved when in 1970 Morlet came up with the idea of using " different window functions for analysis different frequency bands in a signal" by dilations and translations of a single Gaussian function. Due to the " small and oscillatory" nature of these window functions he named them as: " wavelets of constant shape". Thus wavelets came in the picture and then many mathematicians - Grossman, Meyer, Morlet, Daubechies, mallat etc- contributed for a rigorous setting of wavelets.

In the process, many new concepts such as "wavelet Frames" "Multiresolution Analysis(MRA)" were introduced. In 1988, with the development of Daubechies wavelet, , orthonormal basis of compactly supported wavelets, the foundations of " Modern wavelet theory " was laid. In last 30 years many new wavelet families such as: Daubechies wavlet family, Coiflet wavelet family, Block spline semi - orthogonal wavelet family, Battele - Lemarie's wavelet family, Biorthogonal wavelets of Cohen family, Shannon wavelet family, Meyer's wavelet family, and MRA algorithms have been introduced. For more details on historical development of wavelet analysis, see **[**34**]** and the references therein. Let us now introduce some mathematical frame work of wavelet analysis, which will help the reader to take more advantage from this paper.

Fourier series and Finite Fourier transform: For any 2π-periodic, measureable function f on $[-\pi ,\pi]$ and satisfying $\int_{-\pi}^{\pi}|f(x)|\,dx < \infty$, we call the trigonometric series $\frac{1}{2}\hat{f}_c(0) +$

$\sum_{k=1}^{\infty}[\hat{f}_c(k)\, cos(kx) + \hat{f}_s(k) sin(kx)]$ as the real Fourier series of f. Here, $\hat{f}_c(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) cos(kx)\, dx$ and $\hat{f}_s(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\ \ sin(kx)\, dx$ are known as real k^{th} Fourier coefficients of f. Also, we call the series $\sum_{k\in\mathbb{Z}} \hat{f}(k) e^{-ikx}$ as the (complex) Fourier series of f and here $\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) exp(-ikx)\, dx,$ is called as k^{th} Fourier coefficients of f. Finally, the map that maps relates the function f to the sequence $\{\hat{f}(k)\}_{k\in\mathbb{Z}}$ is known as the "finite Fourier transform". For basic properties (linearity theorem, convolution theorem, Parseval's formula etc) of finite Fourier transform and that of Fourier series (such as convergence), see [12].

Fourier transform: The space of all complex valued measureable functions on $\mathbb{R}$(the space of all real numbers) with $\int_{-\infty}^{\infty}|f(x)|\, dx < \infty$ is denoted by $\mathbb{L}^1(\mathbb{R})$ and the space of all complex valued measureable functions on $\mathbb{R}$(the space of all real numbers) with $\int_{-\infty}^{\infty}|f(x)|^2\, dx < \infty$ is denoted by $\mathbb{L}^2(\mathbb{R})$. Let f be a function in $\mathbb{L}^1(\mathbb{R})$. Then the function $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t}\, dt$, $\omega \in \mathbb{R}$, is known as the "Fourier transform of "and the map which maps f to $\hat{f}$ is known as Fourier transform on $\mathbb{L}^1(\mathbb{R})$(Note that co – domain of Fourier transform is not $\mathbb{L}^1(\mathbb{R})$). Also, if f is a function in $\mathbb{L}^2(\mathbb{R})$, then the function $\hat{f}(\omega) = \lim_{N\to\infty}\int_{-N}^{N} f(x) exp(-ix\omega)\, dx, \omega \in \mathbb{R}$ is known "Fourier transform of f" and the map which maps f to $\hat{f}$ is known as Fourier transform on $\mathbb{L}^2(\mathbb{R})$. For basic properties (linearity theorem, convolution theorem, Parseval's formula etc) of Fourier transform, one can see [12].

Discrete Fourier transform: Let f be a complex valued function on $\mathbb{Z}_N = \{0,1,2, \dots\dots\dots . N-1\}$. Then "discrete Fourier transform" of f is denoted by Df and is a function from $\mathbb{Z}_N$ to $\mathbb{C}$ (the set of all complex numbers) defined as $(Df)(n) = \sum_{k=0}^{N-1} f(k) e^{-2\pi ikn/N}$. Also, the map which maps f to Df is known as " Discrete Fourier transform operator ". For basic properties discrete Fourier transform one can see [12].

Fast Fourier transform: In order to find the discrete Fourier transform of a complex valued function on $\mathbb{Z}_N$, we require N^2 multiplication and $N(N-1)$ additions which is really a huge task if N is too large . "Fast Fourier transform "algorithm is the method which reduces the number of these operations (Multiplication and Addition) considerably. In particular, when $N = 2^k$, it reduces the number of multiplication operations from N^2 to something proportional to $N\log_2(N)$". This algorithm traces back its roots in the work of Gauss who according to Heideman et al, expressed this algorithm in a clumsy notation and it got published only Gauss's death . But the real break through came by the seminal article of Cooley and Tukey [29].

Mother Wavelet and Continuous Wavelet Transform: If $\psi \in \mathbb{L}^2(\mathbb{R})$, then the "continuous wavelet transform of any $f \in L^2(\mathbb{R})$ induced by ψ" is denoted by $\mathcal{W}_\psi f$ and is a function from $\mathbb{R}^* \times \mathbb{R}$ to $\mathbb{C}$ defined as: $(\mathcal{W}_\psi f)(a,b) = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} f(t)\ \overline{\psi\left(\frac{t-b}{a}\right)}\, dt = \int_{-\infty}^{+\infty} f(t)\overline{\psi_{a,b}(t)}\, dt = \langle f, \psi_{a,b}\rangle$, where $\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right)$ and $\mathbb{R}^* = \mathbb{R} - \{0\}$. We call ψ as "mother wavelet" and $(\mathcal{W}_\psi f)(a,b)$ as "continuous wavelet coefficients of f" [12]. Also, a and b are known as dilation and translation parameters.

Admissibility condition: In practice, the mother wavelet ψ is also assumed to satisfy the following admissibility condition:$C_\psi = \int_{-\infty}^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}\, d\omega < \infty$. Among many consequences of this admissibility condition on the mother wavelet the most important one's are existence of reconstruction formula; Parsevel's theorem and the condition $\int_{-\infty}^{+\infty}\psi(t)dt = 0$ [33]. For basic properties (linearity theorem, convolution theorem, Parseval's formula etc) of wavelet transform, one can see [12].

Discrete Wavelet Transform: If we take $a = a_0^j\, ; b = kb_0 a_0^j$, where $a_0 > 0\, ;\, b_0 \in \mathbb{R}\, ; j, k \in \mathbb{Z}$, then usually $(\mathcal{W}_\psi f)(a,b)$ is denoted by $(\mathcal{W}_\psi f)(j,k)$ and we can think $\mathcal{W}_\psi f$ as a function on $\mathbb{Z}\times\mathbb{Z}$ defined as $(\mathcal{W}_\psi f)(j,k) = a_0^{-j/2}\int_{-\infty}^{+\infty} f(t)\ \overline{\psi(a_0^{-j}t - kb_0)}dt$. In this case, $\mathcal{W}_h f$ is known as " Discrete wavelet transform induced by the mother wavelet ψ" and $(\mathcal{W}_h f)(j,k)$ as the "discrete wavelet coefficients of f". Also, in this case $\psi_{a,b}$ is denoted by $\psi_{j,k}$ so that $\psi_{j,k}(t) = a_0^{-j/2}\psi(a_0^{-j}t - kb_0)$ and $(\mathcal{W}_\psi f)(j,k) = \langle f, \psi_{j,k}\rangle$. The numbers $(\mathcal{W}_h f)(m,n)$ are known as " Discrete wavelet coefficients of f". In practice, we usually take $a_0 = \frac{1}{2}$ and $b_0 = 1$. Generally speaking, wavelet transform is a device that breaks the data (or functions) in to many frequency components and then studies each component with a resolution in accordance to its scale [33]. Thus wavelet transform offers both informative and economical representations of data of interest. Now a days there are many software packages like MATLAB that contain fast and efficient algorithms to find wavelet transforms.

Numerical stability and reconstruction formula: It is a well known fact that we can recover f, in a numerically stable manner if we know "Discrete wavelet coefficients of f" only provided that $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is either a frame or is an orthonormal basis of $\mathbb{L}^2(\mathbb{R})$ [33].

Orthonormal Wavelets: If for any ψ in $\mathbb{L}^2(\mathbb{R})$(not necessarily satisfying the admissibility condition), $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R})$, then ψ is usually known as "orthonormal wavelet"[51]. Orthonormal wavelets have been charecterized in [51]. Among the various ways of constructing othonormal wavelets one is by using the concept of Multiresolution Analysis(MRA) [51].

Multiresolution Analysis(MRA), father wavelet and construction of orthonormal wavelets: A sequence of closed subspaces $(V_j)_{j\in z}$ in $\mathbb{L}^2(\mathbb{R})$ is known as MRA if it satisfies the properties: (i) (Increasing)(0) $\subset \dots \subset V_{-2} \subset V_{-1} \subset V_{-0} \subset V_1 \subset V_2 \subset \dots \subset \mathbb{L}^2(\mathbb{R})$ (ii)(Separation) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;

(iii) (Density) $\overline{\cup_{j\in\mathbb{Z}} V_j} = \mathbb{L}^2(\mathbb{R})$; (iv) (Scaling) $f(t) \in V_j$ if and only if $f(2t) \in V_{j-1}$; (iv) (Orthonormal basis)There exists $\varphi \in \mathbb{L}^2(\mathbb{R})$such that $\{\varphi(t-n): n \in \mathbb{Z}\}$ is an orthonormal basis for V_0. The function φ, appearing above, is known as "scaling function of MRA" or the "father wavelet". There is a result, known as **"**Mother Wavelet theorem" [12] gives method of constructing orthonormal wavelets from MRA: If $\{V_j\}_{j\in\mathbb{Z}}, \varphi)$ is a Multiresolution analysis and $\psi \in \mathbb{L}^2(\mathbb{R})$ satisfies $\hat{\psi}(w) = e^{-i(w/2+\pi)}\overline{m_\varphi(2^{-1}w+\pi)}\hat{\varphi}(2^{-1}w)\ a.e,$ then $\psi(t) = \sum_{k\in Z} \bar{c}_{1-k}(-1)^k\sqrt{2}\ \ \varphi(2t-k) \in V_0$ $(\{c_k\} \in l^2(\mathbb{Z})$ is an orthonormal wavelet. Here, $m_\varphi(\omega) = \sum_{k\in\mathbb{Z}} 2^{-1/2} b_k\, e^{-ik\omega}$ and is known as "Low pass filter.

Further role of MRA: In addition to the role played in the construction of "orthonormal wavelets", MRA can also be used to approximate $\mathbb{L}^2(\mathbb{R})$ by its subspaces so that one can select a suitable subspace for a particular application task to ensure a balance between the efficiency and the accuracy**.** To say, bigger subspaces offer better accuracy but lesser efficiency and small subspaces although utilise lesser computing resources but may not provide good accuracy**[**63**]**.

Moreover, MRA also offer decomposition of $\mathbb{L}^2(\mathbb{R})$ in to various subspaces **[6],** $\mathbb{L}^2(\mathbb{R}) = \underset{n\,\in\,\mathbb{Z}}{\oplus} W_n$, where W_n is equal to the complement of V_n in V_{n+1}. Also, with the aid of MRA, a "Fast wavelet transform algorithm" called as "pyramid algorithm" was developed by Mallat to determine discrete wavelet coefficients and hence discrete wavelet transform of a function[63]**.** Some important examples of wavelets are: Haar wavelet, Shannon wavelet, Lemarie-Meyer-Wavelet, Franklin Wavelet, Spline Wavelet, Daubechies wavelet, Morlet wavelet, Mexican Hat wavelet, biorthogonal wavelet etc.

***Properties of wavelets that make them more applicable*:** Wavelets find their applications in many areas of science and technology, among which data mining is one. The main properties of wavelets that make them extremely useful for the "data mining" are [63]:

(i) *Computation complexity:* DFT requires $O(N^2)$ multiplications while multiplications required by FFT are $O(N\,log(N))$, but Fast wavelet transform, based on Pyramid algorithm of Mallat, requires only $O(N)$ multiplications [33]**;**

(ii) *Compact support:* Any complex valued function f on $\mathbb{R}$ is said to have compact support if the set $\{x: f(x) \neq 0\}$ lies in some closed and bounded interval. Many wavelets have this property. This property of wavelets ensures that while analysing a region by a wavelet, the data outside of this region does not get effected;

(iii) *Vanishing moments:* A bounded real function f is said to m vanishing moments if $\int x^k f(x) = 0$, for $k = 1, 2, 3, \ldots, m$. Wavelets have this property and this property of wavelets makes wavelets extremely useful for denoising and dimensionality reduction;(iv) *Decorrelated coefficients:* Wavelets reduce temporal correlations which results in to wavelet coefficients with much smaller correlation than the correlation of the corresponding temporal process. (v) *Parseval's theorem:* This ensures the preservation of distances between the objects when wavelet transform is applied on a signal, etc.

Data mining

The volume of the data that we have been storing in computer files and databases is increasing at a phenomenal rate and at the same time the users of these data sets want more adequate and comprehensive information from them. For example , a salesman is not only satisfied by knowing simple listing of his sale contacts, but he also wants to know about his past sales as well as his future sale predictions. This increased demands of adequate and useful information are not met by simple structured language(for example SQL) queries. Here enters the data mining to desire hidden and useful information from data sets. The process of extracting implicit, previously unknown and useful information from a large data is known as data mining (popularly known as knowledge discovery in databases [63]. It has been one of the fastest growing area in recent times. Data mining is a multidisciplinary subject borrowing ideas from various research areas of " artificial intelligence " such as machine learning knowledge acquisition expert systems, pattern recognition and of mathematics such as statistics, information theory uncertain inference. Knowledge discovery process can be viewed as "iterative sequence" of following four components[63]: Data manag – ement (Storage, Easy accessing etc); Data Pre-processing (Denoising, data transformation, dimensionality reduction; Data mining tasks and algorithms(clustering, classification, regression, distributed data mining, similarly search, approximate query processing, visualization, neural network, principle/independent component analysis); Post processing(Evaluation, interpretation documentation etc). During the last two decades or so wavelet methods have been very successfully used to study above mentioned components of knowledge discovery in data bases. In the next section, we shall survey the applications of wavelets to data mining.

APPLICATIONS OF WAVELETS TO DATA MINING

Data management:

As said earlier, the volume of the data that we have been storing is increasing at a phenomenal rate and so it becomes extremely important to manage the data in such a way that it can be accessed quickly and efficiently. In fact, for a given time series data, it is a frequent query to locate the falling and rising trends and sudden changes at different levels of abstractions[63]. Shahabi et al. ([85], [86]) introduced a wavelet based approach to handle such queries. They used a tree structure, TSA (Trend and Surprise Abstraction)-tree, to handle such queries. In fact, they suggested to decompose a given time series data, in to two parts – one is high frequency part (surprise) and other is low frequency part (trend), based on the discrete wavelet transform. Then the trend part is further to be decomposed in to two parts – low frequency part(trend) and high frequency part (surprise) and the process is kept on till we reach an appropriate level of decomposition.

Then since, in order to keep all the information intact, we need to store all the wavelet coefficients involved at each level of decomposition, so the amount of space required for storing the original data, is same as storing the data by this method. So they further, proposed two techniques of dropping some wavelet coefficients, so that lesser space is required to store the data (by their method) without loosing too much information.

Since now a days it is a tradition to work with the digital images, so managing the digital images data efficiently is the need of the hour. In this direction, Venkatesan et al. [96] proposed a wavelet based image indexing technique by introducing a hash function. The authors suggested, firstly, to decompose the given image by using wavelet methods and then tile each subband by rectangles in a random way. Then calculate mean or variance (as the case may be) of each rectangle and then quantize with the help of randomized rounding (i.e., probabilistic quantization) method. The quantized statistics are then input to the decoding stage of a suitably chosen error-correcting code to generate the final hash value. This step helps ensure that hash value remains the same if the decoder input remains within a suitably defined decoding region.

Santini and Gupta [83] also proposed a wavelet based model for the management of image data bases, by defining wavelet transform as a data type. They also introduced an algebra for the manuplation of wavelet data type. Subramanya and youseef [91] presented a wavelet based scheme to index audio data. They claimed that performance of their method is better than some of the other methods (not based on wavelets). Santini and Gupta [82], presented a data model which can be used to for the representation, manipulation and query of wavelet features in multimedia databases. They said that the representing features as complex data type and the possibility for the manipulation of their structure is extremely important for the design of powerful and efficient multimedia databases which are capable of transcending the limitations of the query-by-example model.
Using wavelet-based time-space partitioning (WTSP) tree, Chaoli wang and Han-Wei shen[99], presented a new scheme for the management and rendering large scale (time-varying) data. The authors used the hierarchical TSP tree data structure for capturing both spatial and temporal locality and coherence of the underlying time-varying data and exploited the wavelet transform to convert the data into a multiresolution spatio-temporal representation. A two stage wavelet transform process (3D+1D), in combination with the construction of TSP tree, has been applied to the data in the spatial and temporal domains, respectively, during preprocessing. To render, the wavelet-compressed data stream has been decompressed on-the-fly and rendered using 3D hardware texture mapping. Because of WTSP tree, random access of data, at arbitrary spatial and temporal resolutions, at run time has been possible. The authors claim that the user is provided with flexible error control of image quality and rendering speed trade-off.

Denoising

As recorded by Han and Kamber **[**48], noise is a random error or variance of a measured variable. There are many types of noises such as Gaussian Noise, Impulsive Noise(salt-and-pepper noise or spike noise), Speckle Noise. In a data, noise can be because of (i) measurement errors in the process of collecting data; (ii) errors occurring while entering the collected data in the computers; (iii) inefficient technology; (iv) natural causes (v) etc. To mine a given data efficiently, noise has to be removed. Removing of noise from a given data can be thought as a process of locating outliers or obtaining optimal estimates of the missing data from given noisy data[63]**.** To do this, there are many techniques (which do not use wavelets) available in the literature, but since all most all of these methods are not particularly manufactured to handle the noise problem, so they are not able to produce good results. But denoising of a given noisy data by wavelet methods is very effective and have been very successfully applied in many areas of science and technology.

As recorded in [63], the main idea behind any wavelet based denoising method is: transformation of the given noisy data into a wavelet basis, then considering the bigger coefficients as the useful information and the smaller coefficients as the noise and then finally removing the smaller coefficients (noise) by a suitable rule.

The most popular method, that uses wavelets for removing noise is the one proposed by Dhono and Johnstone [39] and is known as **"**Wave shrink method". If $x = (x_1, x_2, \dots x_n)$ is original signal (without noise) and $y = (y_1, y_2, \dots y_n)$ is the signal obtained by adding noise to x, then the wave shrink method consists of three steps: (i) Transformation of y to the wavelet domain (ii) Shrinking of empirical wavelet coefficients to zero and (iii) Transformation of the shrunk coefficients back to the data domain. Usually there are three shrinkage functions (although there are some more, for example see [104, 35]) which are more in use : (i) Hard Shrinkage function – $\delta_\lambda^H(x) = 0 \ if \ |x| \le \lambda$ and $\delta_\lambda^H(x) = x \ if \ |x| > \lambda$; (ii) Soft Shrinkage function – $\delta_\lambda^S(x) = 0 \ if \ |x| \le \lambda$, $\delta_\lambda^S(x) = x - \lambda \ if x > \lambda$ and $\delta_\lambda^S(x) = \lambda - x \ if x < -\lambda$ and (iii) nonnegative Garrote shrinkage function – $\delta_\lambda^G(x) = 0 \ if \ |x| \le \lambda$ and $\delta_\lambda^G(x) = x - \lambda^2 x^{-1} \ if \ |x| > \lambda$. Here, λ is called as threshold and various methods of fiunding it can be seen in [44, 104, 13] and the references there in.

Li et al. [61] applied wavelets to denoise the biological data and showed that if the localities of the data- the attributes - are strong enough, wavelet denoising is able to improve the performance. Prochazka and Knkal [80] applied discrete wavelet transform for signal preprocessing and signal segment feature extraction. He et al. [50] studied the reduction of noise produced by complex uncertainty, by using wavelet based multiresolution technique. The authors applied the developed approach to a river water quality simulation system for establishing its practicability in both data cleaning and parameter estimation. Anutam and Rajni [10] analysed the performance of various wavelet thresholding methods at different levels of decomposition for denoising of images. Ruikar and Doye [81] also proposed a method of denoising of images using wavelet methods, by introducing a new threshhold function. To improve the denoising effect of the existing methods, He et al.[49], in 2015, proposed a new wavelet based method for the determination of threshold, by exploiting inter - scale correlation. In fact, firstly the authors proposed a new correlation index, which is based on the propagation characteristics of the wavelet coefficients, then by using this new correlation index, a method to determine the threshold is established. In 2015, Madane [66] done a very brief survey on various denoising methods for ECG signals based on wavelets and also studied comparison among various available methods. Bahendwar and Sinha [13] used wavelet discrete transform to propose an algorithm for denoising of medical images. Dansena and Dewangan [32] offered a very nice comparison of various wavelet based denoising methods for medical images, by using thresh holding and optimization techniques. Hamsalakshmi and Kalaivani [47] did an analysis on denoising of one dimensional voice signal by using wavelets methods. Guzel et al [45] applied wavelets to denoise remote sensing very low frequency(VLF) signals obtained form the Holographic array for ionospheric research system and the Elazig VLF receiver syetem. AlMahamdy and Riley [7] studied denoising of ECG signals by using discrete wavelet transform methods and also compared it with some other methods such as adaptive filters method and Savitzky – Golay filtering method. Gao [43], proposed a denoising technique by combining Wave Shrink denoising technique [39] with Breiman's non-negative garrotte [18]. The authors have shown that the asymptotic convergence rate for non-negative garrote shrinkage is same as that for the hard and the soft shrinkage estimates. The authors have further used simulations to show that garrote shrinkage is more suitable than both hard shrinkage and soft shrinkage. The authors have also proposed a threshold selection procedure obtained by combining Coifman and Donoho's cycle-spinning and SURE(knowna as SPINSURE) and have shown through examples that it is better than SURE. Agarwal et al [5] studied denoising (speckle noise) of medical imaging resonance of brain by using wavelets. Li et al [64] proposed a method for denoising the Vehicle Platform Vibration Signal (VPVS) by combining threshold neural network (TNN) with the wavelet coefficients thresholding. In the proposed method the authors constructed a thresholding function and then selected its optimized threshold by unsupervised learning of TNN. Then the original Vehicle Platform Vibration Signal, polluted with trend and noise, is taken as VPVS model and then, on the basis of based on power spectral and energy distribution, a denoising flow has been proposed. The authors also compared the method, through simulation, with the previous methods and shown that their method is superior to other methods. Karim et al [56], applied Discrete Wavelet Transform for denoising the temperature data with the help of symlet 16 along with 32 corresponding filters (low and high-pass). Biswas and Om [16], proposed a new thresholding function to denoise images having Gaussiian white noise.

Data Transformation:

In data mining, there are many data transformation techniques such as Wavelet techniques, Data refinement transformation, Identity transform, Genetic Algorithm and Wrappers, Program synthesis and Feature Selection technique. Among these techniques, the techniques based on wavelets are more effective for many reasons such as: (i) in wavelet domain, performing operations on a given data enables the performance of the operations progressively in a coarse-to-fine fashion; (ii) in wavelet domain, we can perform operations at different resolutions; (iii) in wavelet domain, we can handle features at various scales; (iv) in wavelet domain, operations can be localised in both spatial and frequency domains; (v) in wavelet domain, temporal correlation gets reduced. In fact, many models which do not perform well in the original domain, work quite nicely in the wavelet domain. For example, Principal Component Analysis (PCA) is a very effective and famous algorithm for face recognition, but it has many limitations such as poor discriminatory power and huge computational load, but when studied in wavelet domain by Feng et al. [42], its limitations gone. In their work, the authors applied wavelet transform for decomposing an image into various frequency subbands, and then they used a mid-range frequency subband for PCA representation. Buccigrossi and Simoncelli (see reference 29 in [63]) used wavelet domain for developing a probability model for natural images. Homby et al. [53] used wavelet domain to analyse potential field data.

Dimension Reductionality (or feature selection):

Dimensionality reduction means to represent the original data set by using a smaller set of data without loosing useful information. Since wavelet transformation puts data as sum of some prototype functions, wavelets are very useful in doing this, by deleting certain coefficients. A little bit rigorously, the dimensionality reduction means to project tuples, say p-dimensional, on to a k – dimensional tuples $(k < p, k\ fixed)$ in such a way that the distances remain preserved. There are two ways of executing wavelet based dimensionality reduction - (i) Just keep the largest k coefficients and put rest of the coefficients to zero; (ii) Just keep the first k coefficients and put rest of the coefficients to zero. The process of holding first k coefficients means that we are assuming that all wavelet coefficients in the first k coarsest levels are only significant and the rest of the coefficients, that is, coefficients at a higher resolution levels, are negligible. Such a strong assumption demands the selection of k very carefully and absence of local singularities in the function under consideration [1]. In **[63]** et al has been recorded that if we consider an orthonormal basis in L^2 , then on holding the largest k wavelet coefficients, we get optimal k-term Haar approximation to the original signal. For more detail, see [63]. Chan and Fu [23] used the method of holding first k coefficients(Haar wavelet transform) of the original time series.

Agarwal et al [3], proposed a wavelet based method for dimension reduction of available hyperspectral images and has shown its superiority over the traditional Principal Component Analysis (PCA) technique. Tripathy et al [92] studied dimensionality reduction of Data Warehouse Using Wavelet Transformation. Hamdi et al [46], studied the classification of mammographic images in three steps. In step one, extraction of features which characterize the tissue areas was done; in step second dimension reduction was done by using wavelet methods (among other methods) and finally in step third classification was done. Usman and Mustapha [94] applied Discrete Wavelet Transform for the dimensionality reduction of satellite imagery. Padole [75] studied Dimensionality Reduction of DNA Sequences through Wavelet Transforms. Wright et al [102] used (and also compared) discrete cosine transform discrete wavelet transform for dimensionality reduction of observed rainfall time series for the 438 catchments in the Model Parameter Estimation Experiment (MOPEX) data set.

Clustring:

The situation of clustering of data is encountered in many fields of science and technology. Roughly speaking, we can put the problem of clustering as: Given a space ***P*** of data points, find a partition of ***P*** such that the data points lying in a given class are similar to each other. This problem of clustering of data has been a subject of discussion in machine learning, databases and statistics with different aims[63]. The MRA property of wavelet transforms makes wavelets extremely useful to study clustering. Sheikholeslani, et al. [87] used a multi-resolution clustering approach, for very large spatial data bases, by using wavelets. They proposed a algorithm which outperforms many other clustering methods like BIRCH [105] and CLARANS [74] . Zhang et al. [106], used wavelet transform and proposed a clustering method. Antoniadis et al [9] proposed two methods for identifying clusters time-dependent functional data(high dimensional). Misiti et al [68] proposed a wavelet method for clustering signals by considering a suitably chosen set of wavelet coefficients after going through wavelet based denoising and dimensionality reduction. Barraganet et al [15] used wavelet methods for clustering of a multivariate time series. Pierpaolo and Maharaj [77] also used wavelet methods for clustering of a multivariate time series. Cattani, and Ciancio [22] also studied clustering in time series by using wavelets.

Classification:

In classification problems, we try to locate the group to which a given instance belongs. Wavelets are very useful to study classification problems because of their various properties, in particular multiresolution property. This can be done in two ways[63] – (i) we can apply classification methods on the wavelet domain of the original data or on selective dimensions of the wavelet domain and secondly (ii) we can incorporate the multi-resolution property of wavelets into classification methods. In [19, 20, 21], Castelli et al. proposed classification methods which are based on wavelets for large digital images and it is found that this method is better than traditional Pixel-by-Pixel method [31]. Blume and Ballard[17] gave a method for the classification of image Pixels using learning vector quantization and Haar wavelet transform. Scheanders et al. [84] described a texture analysis of the images by using wavelets. For the classification of texture samples having small dimensions, Mojsilovic et al. [70] also proposed a classification method based on wavelets. In [25, 60] more applications of wavelets to texure classification can be seen. Tzanetakis et al [93] applied wavelets to extract feature set for representing music rhythm and surface information to build. Marcin and Sorin [67] proposed a setup for texture - based probabalistic classification and localization of three dimensional objects in two dimensional digital images. Villez etal [97] described two wavelet based methods for data mining time series of urban water networks. Sheerwood and Derakhshani [88] applied wavelets to classify non - stationary ECG signal for Brain - computer interface applications. Panda et al [76] combined vector support machine with wavelet transform to study classification of ECG signals. Ariwazhagan, and Ganesan [11] studied texture classification using wavelets. In fact, the authors studied the problem by using (i) wavelet statistical features, (ii) wavelet co-occurrence features and (iii) a combination of wavelet statistical features and co – occurrence features of one level wavelet transformed images with different feature databases and found that the third method gives good results.

Regression:

Forecasting of new values on the basis of existing values is known as regression. It is one of the most important data mining task and the denoising task is actually a subtask of this task. There are many referaences included in [63] in which wavelet methods have been used to discuss linear and non - linear regression and various components connected to regression. Amrani et al [8] used wavelets to study regression for Lossless Coding of Remote-Sensing Data. Agarwal [4] used wavelet methods for studying regression in connection with the detection of traffic incidents.

Distributed Data Mining:

Distributed data mining means to mine the data which is distributed, geographically , across multiple sites. In this connection, we need efficient algorithms which reduce the communication overhead, central storage requirements, and computation times[63]. Wavelets because of their orthogonal property play an extremely useful role in distributed data mining as the orthogonality property guarantees exact and independent analysis(local) which can further be used as a building block for a global model. Moreover, compact support property of wavelets plays an important role in designing parallel algorithms as this property makes it sure that processing of a region has no effect on the data out of this region.

Kangupta et al. [55, 52] proposed the idea of performing distributed data mining by using collective data mining(CDM) in wavelet domain . The major steps of their approach are (i) Selection of a suitable orthonormal representation for the type of the data model to be constructed, (ii) generation of othonormal basis coefficients(approximately) at each local site, (iii) selection of an approximate sample of the datasets from each site and putting of them at one site and then generation of the basis coefficients (approx) corresponding to non - linear cross terms, (iv)Combination of all the local models and then transformation of the model into the user described canonical representation and outputting the model. Hershberger et al. [52] used the method of wavelet based CDM to linear discriminant analysis and multivariate regression.

Similarity Search:

In data mining, the similarity search problem is: For a given pattern in a data, find similar patterns in the data set. There are three ways in which wavelets can be use in this data mining task: (i) By transforming the data in to wavelet domain, then doing dimensionality reduction and then studying similarity search; (ii) extracting compact feature vectors and defining new similarity measures ; (iii) supporting similarity search at different scales. Many authors [23, 54, 79, 103] over the period of time have used wavelet methods for similarity search in time series. An overview of similarity search by wavelet method can be found in [28, 71, 72] . Chan and Fu [23] used wavelets to describe time series matching strategy. In fact they combined Haar wavelet transform with the R - tree for similarity search. Wu et al. [103] compared DFT and DWT in time series matching very comprehensively. Struzik and Siebes [89, 90] also proposed similarity measures by using Haar wavelet transform. Natsev et al. [73] proposed wavelet - based retrieval of user specified scenes algorithms for similarity search in data bases. In fact many references can be seen in [63] for similarly search/ inducing using wavelets. Liabotis et al[65] proposed an algorithm for similarity retrieval in case of time series. The authors used wavelets to for dimensionality reduction, then by using X-Trees, the transformed series are indexed. Chan, et al [24] used Haar Wavelets for similarity search of Time-Series.

Approximate Query Processing:

There are a many occasions, when exact answers are not easy to get quickly and we become interested in a fast approximate answers. The process of getting fast approximate answers is known as Approximate Query Processing. Wavelet play a crucial role in this process as a data reduction mechanism to get wavelet synopses, on the basis of which approximate queries can be handled. Many references for approximate query processing can be seen in detail in [63].

Visualization:

Visualization is that task of data mining which helps us in better Understanding of the data. But for big datasets, it becomes very difficult to get do even easiest visualization task. Here wavelet methods, because of multiscale property of wavelet transform, helps viewing of the most important aspects of the data first. Valdivia et al [95], used wavelets to study the time varying data on the nodes of a graph. Miller et al [69] proposed a very interesting approach, based on wavelets, to explore unstructured text . In fact, they used wavelet transform to a custom digital signal which was constructed from words with in a document . The multi resolution energy that came up as a result of applying wavelets, is used to the study the features of narrative flow in the frequency domain. More references on visualization by using wavelets can be seen in [63].

Nueral Network:

A number of authors [27, 26, 36, 59, 30, 57, 58, 101] have combined multi scale wavelet decomposition with neural networks to study data. All these approaches do one of the following: (i) using wavelets as the neuron's activation functions (ii) using wavelets in pre processing phase by extracting feature from time scale data [26, 36, 101]. Bakshi et al. [14] discussed the advantage of using wavelet neural network over other artificial neural learning techniques. Li and Chon [65] discussed the use of wavelet transform and self - organizing map to study air pollution data. Alexendridis and Zapranis [6] studied wavelet nueral networks in detail for applications. Wen et al [100] studied SAR image segmentation by using wavelet nueral network.

Principal Component Analysis:

In [63] a lot of references have been recorded for showing use of wavelets in principle/independent component analysis.

REFERENCES

[1] Abramovich, F., Bailey, T. and Sapatinas, T. (2000) Wavelet analysis and its statistical applications. *JRSSD*, (48):1–30.

[2] Aggarwal, C. C. (ed.) (2007), Data Streams: Models and Algorithms, Springer.

[3] Agarwal, A., Ghazawi, T. E., Moigne, J.L. and Joiner, J. An Application of Wavelet Based Dimension Reduction to AIRS Data.

[4] Agarwal, S. Kachroo, P. and Regentov, E.(2016), A hybrid model using logistic regression and wavelet transformation to detect traffic incidents, IATSS Research 40, 56–63.

[5] Agarwal, S., Singh, O. P. and Nagarie, D. (2017), Analysis and comparison of wavelet transforms for denoising MRI image, Biomedical and Pharmacology Journal, 10(2): 831 – 836.

[6] Alexendridis, A. K. and Zapranis,A. D.(2013), Wavelet neural networks: A practical guide, Neural Networks Volume 42, June 2013, Pages 1-27.

[7] AlMahamdy, M. And Riley, H. B. (2014), Performance study of different denoising methods fro ECG signals, Procedia computer science, 37, 325 – 332.

[8] Amrani, N., Sagristà, J. S. Laparra, V., Marcellin, M. W and Malo, J(2016), Regression Wavelet Analysis for Lossless Coding of Remote-Sensing Data, IEEE Transactions on Geoscience and Remote

Sensing, Volume: 54, Issue: 9.

[9] Antoniadis, A., Brossat, X., Cugliari, J. and Poggi, J. M.,(2013) ,Clustering functional data using wavelets, International Journal of Wavelets, Multiresolution and Information Processing, Volume 11, Issue 01, January 2013.

[10] Anutam and Rajni (2014), Performance analysis of image denoising with wavelet thresholding methods for different levels of decomposition, The International Journal of Multimedia & Its Applications (IJMA) Vol.6, No.3.

[11] Ariwazhagan, S. and Ganesan, L. (2003), Texture classification using wavelet transform, Pattern Recognition Letters, Volume 24, Issues 9–10, Pages 1513-1521.

[12] Bachmana, N.L. and Beckenstein , E., (2000), Fourier and wavelet Analysis, springer - verlay New york , Inc .

[13] Bahendwar, Y. S. and Sinha, G. R.(2015) Efficient algorithm for denoising of medical images using discrete wavelet transforms, Mathematical Methods and Systems in Science and Engineering, ISBN: 978-1-61804-281-1, page 158-161.

[14] Bakshi ,B. R., Koulouris, A. and Stephanopoulos, G.(1994); Learning at multiple resolutions: Wavelets as basis functions inartificial neural networks and inductive decision trees. Wavelet Applications in Chemical Engineering. Kluwer Inc., Boston.

[15] Barragan, J. F., Fontes, C. H. and Embiruçu, M. (2016), A wavelet-based clustering of multivariate time series using a Multiscale SPCA approach, Computers & Industrial Engineering, Volume 95, Pages 144-155.

[16] Biswas, M. and Om, H. (2012), A new soft – Thresholding Image Denoising Method, 2nd International conference on Communication, computing and security, Procedia Technology 6(2012), 10 – 15.

[17] Blume ,M. and Ballard ,D.(1997); Image annotation based on learning vector quantization and localized Haar wavelet transform features, In Proc. SPIE 3077, 181–190.

[18] Breiman, L. (1995), Better subset regression using the non – negative Garrote, Technometrics, 37(4): 373 – 384.

[19] Castelli ,V. and Kontoyiannis, I.(1996); Wavelet-based classification :Theoretical analysis. Technical Report RC-20475, IBM Watson Research Center.

[20] Castelli ,V. and Kontoyiannis, I.(1997); An efficient recursive partitioning algorithm for classification, using wavelets. Technical Report RC-21039, IBM Watson Research Center.

[21] Castelli ,V., Li ,C., Turek ,J. and Kontoyiannis, I. (1996); Progressive classification in the compressed domain for large eos satellite databases, April.

[22] Cattani, C. and Ciancio(2005), A. Wavelet clustering in time series analysis, Balkan Journal of Geometry and Its Applications, Vol.10, No.2, 2005, pp. 33-44.

[23] Chan, K. P, and Fu, A. W.-C. (1999), Efficient time series matching by wavelets. In ICDE, pages 126–133.

[24] Chan, F. K. P and Fu, A. W. C(2003), Haar Wavelets for Efficient Similarity Search of Time-Series: With and Without Time Warping, IEEE Transactions on Knowledge and Data Engineering, 15(3), 686-705, 2003.

[25] Chang ,T. and Kuo, C.(1993);Texture analysis and classification with tree-structured wavelet transform. IEEE Trans. on Image Processing, 2(4):429–441.

[26] Chen ,B., Wang ,X.Z., Yang ,S., and Mc Greavy ,C.(1999); Application of wavelets and neural networks to diagnostic systemdevelopment,2,an integrated framework and its application. Computers and Chemical Engineering, (23):945–954.

[27] Chen ,B., Wang ,X.Z., Yang ,S., and Mc Greavy, C.(1999); Application of wavelets and neural networks to diagnostic systemdevelopment,1,feature extraction. Computers and Chemical Engineering, (23):899–906.

[28] Chiann ,C. and Morettin , P. A.(1999); A wavelet analysis for time series. Journal of Nonparametric Statistics, 10(1):1–46.

[29] Cooley , J., Tukey , J . (1965) An algorithm for the machine calculation of complex Fourier series , Math . Com . 19 , 297-301.

[30] Cristea ,P., Tuduce ,R. and Cristea ,A.(2000); Time series prediction with wavelet neural networks. In Proc. of the 5th Seminar on Neural Network Applications in Electrical Engineering, pages 5–10.

[31] Cromp ,R. F. and Campbell, W. J.(1993); Data mining of multidimensional remotely sensed images. In Proc. 2nd International Conference of Information and Knowledge Management ,,Arlington, VA, Nov.

[32] Dansena, P. and Dewangan, O. (2015), Comparision and analysis Of Wavelet based Image denoising techniques, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 4, Issue 2.

[33] Daubechies, I.,(2003), Ten Lectures on Wavelets, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 1992. Mechanics Research, vol. 30, no. 5, pp. 463–472.

[34] Debnath, L., (2002), Wavelet Transforms and their Applications , Birkhauser .

[35] Dehda, B. and Melkemi, K. (2017), Image Denoising using new wavelet thresholding functions, Journal of Applied Mathematics and computational aMechanics, 16(12), 55 – 65.

[36] Deschenes, C. J. and Noonan ,J. P.(1995); A fuzzy kohonen network for the classification of transients using the wavelet transform for feature extraction. Information Sciences, (87):247–266.

[37] Donoho ,D. L. and Johnstone, I. M.(1994); Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81(3):425–455.

[38] Donoho ,D. L. and Johnstone, I. M.(1995); Adapting to unknown smoothness via wavelet shrinkage. Journal of the American Statistical Association, 90(432):1200–1224.

[39] Donoho, D. L. and Johnstone, I. M.(1998); Minimax estimation via wavelet shrinkage. Annals of Statistics, 26(3):879–921.

[40] Donoho, D. L., Johnstone, I. M., Kerkyacharian ,G., and Picard, D.(1995); Wavelet shrinkage: Asymptopia? J. R. Statist. Soc. B.,57(2):301–337.

[41] Feng, L. and Lin, L. (2017), Comparative analysis of image denoising methods based on wavelet transform and threshold functions, new wavelet threshold function and denoising applications, IJE Transactions B; Applications Vol. 30, No. 2, 199 – 206.

[42] Feng ,G. C., Yuen, P. C., and Dai ,D. Q.(2000); Human face recognition using PCA on wavelet sub band. SPIE Journal of Electronic Imaging, 9(2).

[43] Gao, H. Y. (1997), Wavelet shrinkage denoising using the non – negative Garrote, MathSoft, Inc.

[44] Gao ,H.Y(1997). Threshold selection in Wave Shrink, theory for matlab wavelet toolbox on denoising.

[45] Guzel, E. , Canyilmaz, M and Turk, M. (2011), Application of wavelet based denoising techniques to remote sensing very low frequency signals, Radio Science, Vol. 46, RS 2013.

[46] Hamdi, N., Auhmani, K. and M. M. Hassani(2014), A comparative study of dimension reduction methods combined with wavelet transform applied to the classification of mammographic images, International Journal of Computer Science & Information Technology (IJCSIT), Vol 6, No 6, December 2014 DOI:10.5121/ijcsit.2014.6611 139.

[47] Hamsalakshmi, C. and Kalaivani, R. (2016), Analysis of wavelets based compression in 1D signals, 2D and 3d images, International Journal of Engineering research and Science, Volume 2, Issue 3, ISSN 2395 – 6992.

[48] Han, J. And Kamber, M.(2000), Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers.

[49] He, C., Xing, J., Li, J. , Yang, Q. and Wang, R.(2015) A New Wavelet threshold Determination Method Considering Interscale Correlation in Signal Denoising, Mathematical Problems in Engineering , Article ID 280251, 9 pages.

[50] He, L., Huang, G., Zeng, G. and Lu, H. (2008) "Wavelet-based multiresolution analysis for data cleaning and its application to water quality management systems" Expert Systems with Applications, 35 (3), 1301-1310.

[51] Hernandez, E. And Weiss, G.(1996), A first course on wavelets, CRC Press

[52] Hershberger ,D. E. and Kargupta ,H.(2001);Distributed multivariate regression using wavelet-based collective data mining. Journal of Parallel and Distributed Computing, 61(3):372– 400.

[53] Hornby, P., Boschetti, F. and Horowitz, F.(1997); Analysis of potential field data in the wavelet domain. In Proceedings of the 59th EAGE Conference, May 1997.

[54] Huhtala ,Y., Karkkainen ,J., and Toivonen, H.(1999); Mining for similarities in aligned time series using wavelets. In Data Mining and Knowledge Discovery: Theory, Tools, and Technology .SPIE Proc., 1999.

[55] Kargupta ,H. and Park ,B.(1998); The collective data mining: A technology for ubiquitous data analysis from distributed heterogeneous sites. Submitted to IEEE Computer Special Issue on Data Mining.

[56] Karim, S. A. A., Ismail, M. T., Hasan, M. K. and Sulaiman, J. (2013), Denoising the Temperature Data Using wavelet Transform, Applied Mathematical Sciences, Vol. 7, 2013, no. 117, 5821 – 5830.

[57] Kobayashi ,K. and Torioka ,T.(1994); A wavelet neural network for function approximation and network optimization. In Proceedings of ANNIE'94, AMSE Press., pages 505–510.

[58] Kobayashi ,K. and Torioka ,T.(1997); Designing wavelet networks using genetic algorithms. In Proceedings of EUFIT'97, volume1.

[59] Kostka, P., Tkacz ,E., Nawrat, Z., and Malota ,Z.(2000); An application of wavelet neural networks (wnn) for heart valve prostheses characteristic. In Proc. of the 22nd Annual International Conference of IEEE, Engineering in Medicine and Biology Society, volume 4, pages 2463–2465.

[60] Laine, A. and Fan, J.(1992); texture classification by wavelet packet signatures. Technical report, University of Florida.

[61] Li ,Q., Li, T. and Zhu, S.(2002); Improving medical/biological data classification performance by wavelet pre-processing. In ICDM 2002,Volume 4, Issue 2 - page 65.

[62] Li ,S.T. and Chou ,S.W.(2000); Multi-resolution spatio - temporal data mining for the study of air pollutant regionalization. In Proceedings of the 33rd Hawaii International Conference on System Sciences.

[63] Li,T., Li, Q., Zhu, S. and Ogihara. M(2002), A survey on Wavelet applications in Data mining, Vol. 4, Issue 2; pages 49-68.

[64] Li, M., Wang, Z., Luo, J., Liu, Y. and Cai, S. (2017), Wavelet denoising of Vechile plateform vibration signal based on threshold Neural network, Hindawi.

[65] Liabotis, I., Theodoulidis, B. and Saraee, M.(2006), Improving Similarity Search in Time Series Using Wavelets, International Journal of Data Warehousing and Mining 2(2):55-81.

[66] Madane M. R.(2015), ECG Denoising With Wavelet Transform Methods : A Survey, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Vol. 4 Issue 02,

[67] Marcin, G.and Sorin, S.(2010); Local Wavelet Features for Statistical Object Classification and Localisation. Multimedia, IEEE, 17, 1, 118−125 .

[68] Misiti, M., Misiti, Y. Oppenheim, J. and Poggi, J. M.,(2013), Clustering Signals Using Wavelets, Computational and Ambient Intelligence. IWANN 2007. Lecture Notes in Computer Science, vol 4507. Springer, Berlin, Heidelberg

[69] Miller, N. E., Wong, P. C., Brewster, M., and Foote, H.(1998);TOPIC ISLANDS - A wavelet-based text visualization system. InD. Ebert, H. Hagen, and H. Rushmeier, editors, IEEE Visualization'98, pages 189–196.

[70] Mojsilovic, A . and Popovic, M. v.(1997); Wavelet image extension for analysis and classification of infarcted myocardial tissue. IEEE Transactions on Biomedical Engineering, 44(9),September.

[71] Morettin, P. A.(1996); From fourier to wavelet analysis of time series .In A. Prat, editor, Proceedings in Computational Statistics, pages 111–122.

[72] Morettin, P.(1996);Wavelets in statistics. (3):211–272, 1997.

[73] Natsev, A., Rastogi, R., and Shim ,K.(1999); Walrus: a similarity retrieval algorithm for image databases. In Proceedings of ACM SIGMOD International Conference on Management of Data, pages 395–406. ACM Press.

[74] Ng, R. T. and Han ,J.(1994);Efficient and effective clustering methods for spatial data mining. 20th International Conference on Very Large Data Bases, September 12–15, Santiago, Chile proceedings, pages 144–155.

[75] Padole, M.C., Dimensionality Reduction of DNA Sequences through Wavelet Transforms, Recent Advances in Computer Science, ISBN: 978-960-474-354-4.

[76] Panda, R., Khobragade, P. S., Jambhule, P. D., Jengthe, S. N., Pal, P.R. and Gandhi, T. K.(2010), Classification of EEG signal using wavelet transform and support vector machine for epileptic seizure diction, International Conference on Systems in Medicine and Biology, DOI: 10.1109/ICSMB.2010.5735413,IEEE.

[77] Pierpaolo, D. and Maharaj, E. A. (2012), wavelet-based clustering of multivariate time series, Fuzzy sets and systems, Volume 193, Pages 33-61.

[78] Pong Chan, K. and Fu, A. W.C.(1999); Efficient time series matching by wavelets. In ICDE, pages 126–133.

[79] Popivanov, I. and Miller, R. J..(2002);Similarity search over time series data using wavelets. In ICDE.

[80] Prochazka, A.and Kukal ,J.(2008); Wavelet Transform Use for Feature Extraction and EEG Signal Segments Classification. Communications, Control and Signal Processing,719–722 .

[81] Ruikar, S.D. and Doye, D. D.(2011) Wavelet Based Image Denoising Technique, International Journal of Advanced Computer Science and Applications(IJACSA), Vol. 2, No.3,

[82] Santini ,S. and Gupta ,A.(2001); A data model for querying wavelet features in image databases. In Multimedia Information Systems, pages 21–30,.

[83] Santini, S. and Gupta, A.(2001); Wavelet data model for image databases. In IEEE Intl. Conf. on Multimedia and Expo, Tokyo, Japan, August.

[84] Scheunders,P., Livens,S., de Wouwer, G. V., Vautrot,P., and Dyck, D. V.(1998);Wavelet-based texture analysis. International Journal on Computer Science and Information Management, ijcsim, 1(2):22–34.

[85] Shahabi, C., Chung ,S., Safar ,M., and Hajj, G.(2001); 2d TSA tree: A wavelet-based approach to improve the efficiency of multi-level spatial data mining. In Statistical and Scientific Database Management, pages 59–68,.

[86] Shahabi, C., Tian, X. and Zhao, W.(2000); TSA-tree: A wavelet based approach to improve the efficiency of multi-level surprise and trend queries on time-series data. In Statistical and Scientific Database Management, pages 55–68.

[87] Sheikholeslami ,G., Chatterjee ,S., and Zhang, A.(1998); Wave Cluster: A multi-resolution clustering approach for very large spatial databases. In Proc. 24th Int. Conf. Very Large Data Bases, VLDB, pages 428–439.

[88] Sherwood, J.and Derakhshani, R.(2009);On Classifiability of Wavelet Features for EEG-Based Brain-computer Interfaces. IJCNN, 2895−2902.

[89] Struzik ,Z. R. and Siebes ,A.(1999); Measuring time series' similarity through large singular features revealed with wavelet transformation. In DEXA Workshop 1999, 162–166.

[90] Struzik, Z. R. and Siebes, A.(1999); The haar wavelet transform in the time series similarity paradigm. In Proceedings ofPKDD'99, pages 12–22.

[91] Subramanya ,S. R. and Youssef, A.(1998); Wavelet-based indexing of audio data in audio/multimedia databases. In IWMMDBMS, pages 46–53.

[92] Tripathy, A. , Das, K. and Swarnakar, T. (2011), Dimensionality Reduction of Data Warehouse Using Wavelet Transformation: An Enhanced Approach for Business Process, Computer Networks and Information Technologies, CNC 2011, Communications in Computer and Information Science, vol 142. Springer, Berlin, Heidelberg.

[93] Tzanetakis ,G., Essl, G., and Cook ,P.(2001); Automatic music algenre classification of audio signals. pages 205–210, Bloomington, IN, USA, October.

[94] Usman, B. and Mustapha, A. (2015), Satellite imagery dimensionality reduction using discrete wavelet transform for the extraction of cadastral features, International Research Journal of Computer Science (IRJCS) ISSN: 2393-9842 Issue 12, Volume 2 (December 2015).

[95] Valdivia, P., Dias, F. , Petronetto, Silva, C. T and Nonato, L. G.(2014), Wavelet-based Visualization of Time-Varying Data on Graphs, IEEE.

[96] Venkatesan, R., Koon, S., Jakubowski, M. and Moulin, P.(2000); Robust image hashing. In Current proceedings.

[97] Villez, K., Pelletier, G.and Rose, C.(2007); Comparison of Two Wavelet-Based Tools for Data Mining of Urban Water Networks Time Series, Water Science & Technology, 56, 6, 57−64 .

[98] Walden, A.G.,wavelet analysis discrete time series, online available.

[99] Wang, C. and Shen, H. W., A Framework for Rendering Large Time-Varying Data Using Wavelet-Based Time-Space Partitioning (WTSP) Tree.

[100] Wen, X. B., Zhang.H. and Wang, F. U. (2009), A Wavelet Neural Network for SAR Image Segmentation, Sensors 2009, 9, 7509-751.

[101] Woodford, B. J. and Kasabov, N. K.(2001); A wavelet-based neural network classifier for temporal data. In Presented at the5th Australasia-Japan Joint Workshop, University of Otago, Dunedin, New Zealand, November.

[102] Wright, A., Walker, J. P., Robertson, D. E. and Pauwels, V. R. N(2017), A comparison of the discrete cosine and wavelet transforms for hydrologic model input data reduction, Hydrol. Earth Syst. Sci., 21, 3827-3838.

[103] Wu, Y.L., Agrawal, D., and Abbadi, A. E.(2000); A comparison of DFT and DWT based similarity search in time-series databases. In CIKM, pages 488–495.

[104] Yi, L. J., Hong, L., Dong, Y. and Sheng, Z. Y. (2016), A new wavelet threshold function and denoising applications, Hindawi publishing corporation, Mathematical Problems in Engineering, Volume 2016, Article ID 3195492.

[105] Zhang ,T., Ramakrishnan, R. and Livny ,M.(1996), BIRCH: an efficient data clustering method for very large databases. In Proceedings of ACM SIGMOD, pages 103–114 .

[106] Zhang, H., Bao ,T., Zhang,Y.and Lin ,M.(2006);Unsupervised Feature Extraction for Time Series Clustering Using Orthogonal Wavelet Transform. Informatica, 30, 305−319.