International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 13 (2019) pp. 3081-3096

A Study on Wavelet Threshold Functions in Denosing

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

Every scientist or engineer dealing with the real world data knows very well that, in general, signals do not exist without noise - Gaussian Noise, Impulsive Noise(salt-and- pepper noise or spike noise), Speckle Noise. So to obtain the correct information, the signals have to be denoised for which there are abundance of methods available in the literature. Among these methods, the *wavelet threshold denoising methods* are very effective and have been very successfully applied in many areas of science and technology. The *wavelet threshold denoising* methods involve three steps - a linear forward wavelet transformation, a nonlinear thresholding step and a linear inverse wavelet transform. The second step uses a function, called as threshold function. Donoho and Johnstone proposed two types of threshold functions - hard and the soft thresholding functions. However, these thresholding techniques suffer from certain drawbacks and hence a number of wavelet threshold functions have been introduced by the researchers since then. In this paper, we shall present a detailed survey on the various wavelet threshhold functions available in the literature.

Keywords: Denoising, wavelet transform, inverse wavelet transform, MRA, wavelet threshold function, threshold value.

1. INTRODUCTION

Every scientist or engineer dealing with the real world data knows very well that, in general, signals do not exist without noise - Gaussian Noise, Impulsive Noise(salt-and- pepper noise or spike noise), Speckle Noise. So to obtain the correct information, the signals have to be denoised. For some decades now, developing efficient denoising methods has been a very challenging area of research at the point of intersection of the two main areas of Mathematics – Functional Analysis and Statistics. Among many denoising methods available in the literature – such as the Gaussian smoothing model **[1]**, the anisotropic filtering model **[2]**, the Rudin-Osher-Fatemi total variation model **[3]**, the Yaroslavsky neighbourhood filters**[4, 5]**, the SUSAN filter **[6]**, the Wiener local empirical filter as implemented by Yaroslavsky **[7]**, DUDE – the discrete universal denoiser **[8]** and the UINTA – Unsupervised Information Theoretic Adaptive Filtering **[9]**, the non local means (NL-means) algorithm [**10],** Wavelet threshold denoising method method **[11],** the translation invariant wavelet thresholding method**[12]**, methods combining wavelets and the concept of total variation **[13, 14, 15],** Meyer method **[16]**, methods based on bandlets **[17]** and curvlets **[18]** etc – the wavelet(a particular part of Functional Analysis) based denoising methods stand distinguished because of some very useful properties that wavelets own, and among the wavelet based methods, *wavelet threshold denoising method* stands out distinctly because of its simplicity and good effect.

The wavelet threshold denoising method is actually based on the Mallat's theory, according to which we can reconstruct a signal completely from its low-frequency approximation part and high-frequency detail part **[19, 20]**. In fact, suppose an original discrete $s(n)$ is given by $s(n)=\sum_{j,k\in\mathbb{Z}} c_{j,k}\varphi_{j,k}(n)+\sum_{i=1}^{j}\sum_{k\in\mathbb{Z}} d_{i,k}\psi_{i,k}(n)$, where $c_{j,k}=\left\langle s(n),\varphi_{j,k}(n)\right\rangle$ denotes the approximate coefficients, j denotes the decomposition level, φ denotes the scaling function, $d_{i,k}=\left\langle s(n),\psi_{i,k}(n)\right\rangle$ denotes the detail coefficient, and $\psi_{i,k}$ stands for the wavelet basis function. Here, low frequency information, of the original signal $s(n)$, is contained in the approximate coefficients $c_{j,k}$ and the high frequency information is contained in the detail coefficients $d_{i,k}$. The *Wavelet threshold denoising method* is mainly based on the assumption that the low-frequency part of the signal $\sum_{j,k\in\mathbb{Z}} c_{j,k}\varphi_{j,k}(n)$ represents the major profile of the original signal, whereas high-frequency part $\sum_{i=1}^{j}\sum_{k\in\mathbb{Z}} d_{i,k}\psi_{i,k}(n)$ represents its details. Further, it is assumed that details of each level contain the noise information.

In the process of the execution of the wavelet based thresholding denoising method, the detail coefficients $d_{i,k}$ at each level are modified and reconstructed with the approximate coefficients $c_{j,k}$ of the last level.

There are two factors that are important in the wavelet decomposition and reconstruction process – one is the selection of a suitable wavelet basis $\{\psi_{i,k}\}_{i,k\in\mathbb{Z}}$ and the other is the choice of decomposition level j. Also, there are two important things involved in the modification of wavelet coefficients – one is the selection of a suitable number, known as the threshold value, in reference to which the wavelet coefficients are been modified and the other is the selection of a function, known as threshold function, which describes the process of modification of the wavelet coefficients.

As for as the selection of the wavelet is concerned, one has to keep in mind some properties of the wavelets as outlined in section three below – such as length of the support, symmetry, vanishing moment regularity etc **[21]**. A method of determination of maximum decomposition level has been suggested in **[22],** in which the authors determined it on the basis of the minimum frequency of the desired signal.

The main aim of this paper is to survey various threshold functions involved in the wavelet based denosing methods available in the literature.

In section 2, we present a brief historical development from Fourier Analysis to Wavelet Analysis where as in section 3, we present various definitions and results connected with Fourier and wavelet Analysis. In section 3, we present the *wavelet threshold denoising method*. In section 4, we present the concept of *wavelet threshold denoising method*. Then in section 5, we present a detailed survey of the threshold functions available in the literature highlighting the methods of determining the threshold values and evaluation criterions (that is, means to measure performance of the method) used by the authors. In section 6 and 7, we present a brief explanation of the methods of determining threshold values and evaluation criterions available in the literature.

2. JOURNEY FROM FOURIER TO WAVELET ANALYSIS: HISTORICAL FRAMEWORK

Discovery of the Fourier's idea: It is a known fact that process of understanding becomes easier if complicated structures are synthesized by using simpler ones**[23]**. For example, the process of studying complex numbers becomes simpler if we firstly understand real numbers and then study complex numbers. Similarly, the process of understanding matter becomes simpler if we start understanding quarks, then atoms, then molecules and then finally the matter. Jean Baptiste Joseph Fourier did the same thing in the early part of the nineteenth century, when he was studying the most burning problem of that time: "*How heat diffuses in a continuous medium*".

During his attempt of putting forward a solution, an idea , which later on proved to be one of the most important ideas in the history of science and technology, of synthesizing a function with the help of simpler functions –Sines and Cosines –came in his mind. As a result, on December 21,1807, Fourier submitted a manuscript to "institute de France" in which he claimed that "*every periodic function can be expressed as a weighted sum of sines and cosines*".

Rejection of the Fourier's Idea: His(Fourier's) manuscript went to a committee consisting of Laplace, Langrange, Lacroix and Monge (At that time Poisson was only the committee's clerk). As the Fourier's work was based on intuitions and no rigorous approach was followed, in particular with respect to the convergence of the series, the committee rejected Fourier's paper for publication(actually, Langrange opposed most strongly among the members of the committee). Thus, finally in 1808 Poisson (clerk of the committee at that time) put forward the committee's report which was: "*Rejected on the grounds that it contained nothing new or interesting*".

Resubmission of the Fourier's idea: This rejection did not make big set back to the Fourier's mission, probably he knew the tradition of this world: "Whenever some body comes with a new idea ,people ,especially prominent ones ,oppose it".

One more rejection of the Fourier's idea: He once again came, with few modifications, particularly in rigouring his idea, in 1811, essentially to the same committee ,as `a candidate for the "Grand pring de mathematiques" for 1812. This time again the committee did not allow his paper for publication in "Memories de e'Academie des sciences", although they awarded him the prize.

Publishing of the Fourier's idea: In 1824, things changed and Fourier became the secretary of the Academie and then he published his work , without practically changing his 1811 version.

Problem of convergence: Although Fourier's idea later revolutionized the science and technology, but the problem of convergence, highlighted by Lagrange, really serious. To this end, Poisson, Cauchy, Drichlet and many other mathematicians remain engaged. In 1826 Cauchy published a proof of convergence of Fourier series, but there was a flaw in his proof which, three years later, a 23-years old boy Drichlet, highlighted. He (Drichlet) himself gave some sufficient conditions for the convergence of Fourier series (Drichlet point wise convergence theorem). Later in 1881, Jorden improved these sufficient conditions. In the last century (i.e, 20^{th} century) there had been many new sufficient conditions for the convergence of Fourier series, put forward by many mathematicians. For more details, see **[23, 24, 25]**.

Emergence of new ideas because of Fourier's idea*:* Because of the introduction of the Fourier series many new concepts either popped up or were made more rigorous. For example, Riemann discovered the idea of integral (now known as Riemann integral) only after getting motivated by Fourier series. Latter Lebesgue, also inspired by some problems concerning Fourier series, generalized the concept of Riemann integral to that what is now known as Lebesgue integral. For more details see **[23, 24, 25]**.

Extension of the Fourier's idea: After Fourier coined his idea of Fourier series, some related ideas such as Fourier transform, discrete Fourier transform, Fast Fourier transform also came up during the course of time. Among these it is been said that "Fast Fourier Transform" is one of the best algorithm of all times in science and technology. It was proposed by Cooley and Tukey in 1965 **[26]** and due to this algorithm "Fourier Transform" became the "king of all transforms".

Applications: 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.

Inadequacy of Fourier Analysis*:* If a function is approximated or represented by a sum (weighted) of some simpler functions, which in turn are obtained by a single simple function, then such approximations or representations of functions offer many advantages such as: • making simpler the analysis of complicated functions; • compressing significantly the representations of the original function if, for some reasons, only a few simpler functions renders good approximation of the original function. Fourier in his representation of functions, used "sinusoids of different frequencies" as simpler function. As said above, although his representation of functions have been in use, since the time of Fourier, in various fields of science and technology. "Fourier representation of functions", with the help of sinusoids , has two major drawbacks such as: (i) Sinusoids (Sine and Cosine functions) does not have compact support in "time domain"(although they have perfectly compact support in frequency domain). Putting in other words, this means that in time domain they stretch out to infinity. This makes their "*non applicability to non-stationary signals*."

(ii) Fourier representation provides spectral content without the time localization . This means that: "*non - stationary signals whose spectral content varies with time can't be analysed with the help of Fourier Analysis*".

Windowed Fourier Transforms: Once the above weakness in Fourier representation for non - stationary signals was realized by the Applied mathematicians , they started to modify it. In this context the first modification came by Gabor in 1946, who, was studying the representation of a communication signal , in a time - frequency plane , by using oscillatory basis functions . He modified the concept of Fourier transform to " Short Time Fourier Transform" (STFT) by using a Window - Gaussian function. The trick behind the STFT is: "*to represent the signal by using a time - localized window and then doing the analysis for each segment*".

This modification provides a true time - frequency representation of the signal as in this case we compute Fourier transform for every window (i.e, time localized) segment of the signal. One year later in 1947, Jean Ville proposed, a similar transform, Winger - Ville transform for the representation of energy of a signal in the time frequency plane . In fact during the period 1940 to 1970 , many similar transforms were proposed . To cite a few: • Cohen Distribution; • Wigner- Ville- Rihaczek Distribution. All these Windowed Fourier transforms differ only as far as the choice of the window function is concerned.

Drawbacks in Windowed Fourier transforms: Although introduction of Windowed Fourier transforms served some purpose, but when we have to analyse a signal which has • *high frequency components with short time spans* or • *low frequency components with long time spans*, then we need a narrow window to handle first case and a wide window to analyse the second case , and since a Windowed Fourier transform uses only a single window to analyse the entire signal, Windowed Fourier transform is inadequate to handle such signals. Thus, "*using of a signal window function for the entire signal* " is a major drawback in Windowed Fourier transform.

The discovery of J.Morlet: In 1970's a geophysical engineer at the French oil company Elf Acquitaine, J. Morlet, was analysing a signal which had: • *high frequency components with short time spans* and • *low frequency components with long time spans*. As said above, Windowed Fourier transforms can't handle this situation, so Morlet tried to discover some thing new to handle this situation and he came up with a brilliant idea of using "different window functions for analysis different frequency bands in a signal". Also, the different window functions, that he used to analyse a signal of above type, are derived from a single function - Gaussian function - by dilation and compression. Due to the " small and oscillatory" nature of these window functions he named them as: " wavelets of constant shape."

Joining Hands with Grossman: Following the tradition of Mankind: "offering opposition to every new idea ", Morlet, just like Fourier, faced much criticism from his contemporaries, in particular because of lack of mathematical rigour. Morlet, in his search to find a mathematical rigorous foundation to his ideas, met a theoretical Physicists of quantum mechanics, Grossman, and discussed his ideas with him. Grossman, in Morlet's work, find some thing similar to the coherent states formalism , a technique he had been using in Quantum mechanics - and so he shown a lot of interest in it. After some time, Grossman succeeded in formalising Morlet's ideas and also devised an exact inversion formula for "Morlet's integral transform " and did a lot of applications together with Morlet.

Meyer's discovery: In the spring of 1985, Y . Meyer, while waiting on a line to photocopy some papers, heard about Grossman and Morlet's work . After going into their work, Meyer recognised that Morlet and Grossman's analysis and inversion formula is , actually , a rediscovery , of a formula in Harmonic Analysis, introduced by A. Calderan in 1960's. Y. Meyer not only recognised the similarity of Morlet and Grossman's work with A .Calderan but also found that "there is a great deal of redundancy in wavelets". Inspired by this, Meyer started working for developing wavelets with better localization properties and he succeeded in producing an "orthogonal wavelet basis" with nice time and frequency localization. Surprisingly, again the Meyer's orthogonal wavelet basis turned out to be a rediscovery of J.O. Stromberg - a Harmonic analyst - who discovered the same basis five years earlier to Meyer.

Original discoverer: It is rather more surprising that the art of constructing orthogonal basis did not start from Stromberg or Meyer but it dates back to 1909, when Alfred Haar, a German Mathematician, constructed an orthonormal wavelet basis- although at that time the name wavelet was not in use. It was also discovered, later, that Haar's work of constructing orthonormal basis were expanded by Paul Levy in 1930 , when he was engaged in studying " Brownian motion " and also independently by Paley and Littlewood .

Wavelet Frames: Meanwhile Daubechies, a student of Grossman introduced the concept of "wavelet Frames " for the purpose of discretizing the time and scale parameters of the wavelet transform . This new concept offered more liberty in the matter of " choice of basis", although, at the cost of some redundancy.

Multiresolution Analysis: In 1986, the concept of Multiresolution Analysis for discrete wavelet transform was introduced by Mallat and Meyer . This development , later in 1988, became Ph.D. thesis of Mallat. The idea of Mallat was: "Decomposition of a discrete signal into its dyadic frequency bands by a series of high pass and low pass filters to compute its DWT from the approximation of these various scales ."

It is worth noting that the same ideas were familiar to electrical engineers under the name of "quadrature Mirror Filters " (QMF) and sub band filtering, for about 20 years earlier to Mallat.

Final word: In 1988 with the development of Daubechies wavelet orthonormal basis of compactly supported wavelets, the foundations of " Modern wavelet theory " were laid. In last 30 years or so, 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 **[23, 24, 25, 26, 27]** and the references therein.

Applications of Wavelets: We now list some areas of applications of wavelet: • Data compression • Denoising • Source and Channel Coding • Biomedical Engineering • Non destructive Evaluation • Study of Distant Universe • Wavelet Networks • Zero Crossing Representation • Fractals • Turbulence Analysis • Financial Analysis • Medicine • Seismology • Computer graphics • Digital communication • Pattern recognition • Approximation theory • Sampling theory • Statistics • Numerical analysis • Operator theory • Computer vision • Differential equations • Natural scenes • Mammalian visual systems.

New Arrivals: As it is well known fact "nothing is perfect in this world", wavelets too face some problems while dealing with objects in more than one dimension. So in last 15 years or so, some new concepts emerged. To name a few: • Multi directional wavelets • Complex wavelets • Curvelets • Shearlets • Composite wavelets • Bandelets • Grouplets.

3. JOURNEY FROM FOURIER TO WAVELET ANALYSIS: TECHNICAL FRAMEWORK

$L^p[-\pi,\pi]$ Spaces, $1 \le p < \infty$: $L^p[-\pi,\pi]$ consists of all 2π- periodic complex valued measureable functions f on $[-\pi,\pi]$ satisfying $\left[\int_{-\pi}^{\pi}|f(x)|^p\,dx\right]^{1/p} < \infty$.

Fourier coefficients: For $f \in L^1[-\pi,\pi]$, we call $\hat{f}_c(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)cos(kx)\,dx$

and $\hat{f}_s(k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)sin(kx)\,dx$ as real k^{th} Fourier coefficients of f, $k \in \mathbb{Z}^+$.

Also, we call $\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)exp(-ikx)\,dx$, $k \in \mathbb{Z}$ as complex k^{th} Fourier coefficients of f.

Fourier Series: For $f \in L^1[-\pi,\pi]$, the trigonometric series $\frac{1}{2}\hat{f}_c(0) + \sum_{k=1}^{\infty}[\hat{f}_c(k)\,cos(kx) + \hat{f}_s(k)sin(kx)]$ is known as real as the real Fourier series of f and the series $\sum_{k\in\mathbb{Z}}\hat{f}(k)e^{-ikx}$ is known as the complex Fourier series of f. Fourier series offers an effective tool to analyse the "frequency" properties of a periodic function. If one thinks a periodic function 'f' as a record of piece of a music, then Fourier series of f represents the tones in the music and shows the strength of each tone - the larger the value of $|\hat{f}(k)|$, the stronger the tone is. Similar to the concept of Fourier series, there are concepts of Fourier Sine series and Fourier Cosine series. For details See **[23]**

Finite Fourier transform: The map that maps the function $f \in L^1[-\pi,\pi]$, to the sequence$\{\hat{f}(k)\}_{k\in\mathbb{Z}}$ is known as the Finite Fourier transform of f. Some times $\hat{f}(k)$ is also known as "the value of Finite Fourier transform of f at k". For basic properties such as linearity theorem, convolution theorem, Parseval's formula etc of finite Fourier transform one can see **[23].**

Convergence of Fourier series: There are several results such as •Dini's Test • Dirichlet's Point wise Convergence Theorem • Jorden theorem etc, which provide us certain sufficient conditions for the convergence(point wise or uniformly) of Fourier series of a function f to f. The details of these and many more results, providing the sufficient conditions for the convergence(point wise or uniformly) of Fourier series of a function f to f, can be seen in **[23, 28]**. It is important to mention here some of the concepts and results which pave the way to these convergence results of Fourier series. These are: • Dirichlet kernel • Continuous Fourier kernel • Riemann- Lebesgue Property for Dirichlet kernel • Discrete Fourier kernel • Riemann's Localization Principle. Moreover, by using the principle of Uniform boundedness, it can be shown that "If $f \in L^1[-\pi,\pi]$ is continuous, even then Fourier series of f may not converge, even point wise , at many points of $[-\pi,\pi]$.

Relation between Trigonometric series and Fourier series: In general every trigonometric series need not be a Fourier series of a function. For example, the trigonometric series $\sum_{n\ge2} a_n\,sin(nt) + \sum_{n=1}^{\infty} b_n\,cos(nt)$, with $a_0 = 0$; $a_1 = 0$; $a_n = \log(n)$; $b_n = 0$, is not a Fourier series of any $f \in L^1[-\pi,\pi]$. However, we have have the result **[23]**: If the trigonometric series $\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\,sin(nt) + \sum_{n=1}^{\infty} b_n\,cos(nt)$ converges uniformly on $[-\pi,\pi]$, then the limit is a 2π −Periodic , continuous function whose Fourier coefficients are a_n , b_n .

$L^p(\mathbb{R})$ Space, $1 \le p < \infty$: $L^p(\mathbb{R})$ consists of all complex valued measureable functions f on $\mathbb{R}$ satisfying: $\left[\int_{-\infty}^{\infty}|f(x)|^p\,dx\right]^{1/p} < \infty$.

Fourier transform of a function $f \in L^1(\mathbb{R})$: For $f \in \mathbb{L}^1$, we call the function $\hat{f}$ defined as $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt$, $\omega \in \mathbb{R}$, as the "Fourier transform of f" and the map which maps f to $\hat{f}$ as the "Fourier transform of f". Fourier transform of a function $f \in L^1(\mathbb{R})$, actually, denotes the continuous version of Fourier series. For basic concepts and properties of Fourier transform such as • Scaling • Time Shift → Frequency Modulation • Exponential Modulation→ Frequency shift • Change of roof • Convolution theorem etc one can see **[23].**

Inverse Fourier transform: If $L^1(\mathbb{R})$, then the integral $\frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{i\omega t}\,d\omega$ is known as the inverse Fourier transform of f and is denoted by $\check{f}$. For more details, see **[23].** Similar to the concept of Fourier transform of $f \in L^1(\mathbb{R})$, we have the concept of Sine and Cosine Fourier transforms of $f \in L^1(\mathbb{R})$. For Details , See **[23]**

Fourier transform of a function $f \in L^2(\mathbb{R})$: If $f \in L^2(\mathbb{R})$, then Fourier transform $\hat{f}$ of f is defined as $\hat{f}(\omega) = \mathbb{L}^2 - \lim_{N\to\infty}\int_{-N}^{N} f(x)exp(-ix\omega)\,dx$, where an expression like $f(t) = \mathbb{L}^2 - \lim_{n\to\infty} f_n(t)$ means $\lim_{n\to\infty}\|f_n - f\|_2 = 0$. For various properties connected to Fourier transform of $f \in L^2(\mathbb{R})$ such as • Reciprocity theorem •Inversion Formula etc, one can see **[23].**

The space Z_N: We define $\mathbb{Z}_N$ as the space $\{0,1,2,\dots\dots\dots.N-1\}$ equipped with the $\sigma-$Algebra as the set of all of its subsets and the " measure" as counting measure , so that every function $f: \mathbb{Z}_N \to \mathbb{C}$ is measureable. We also define $L^1(\mathbb{Z}_N) =$ The set of all functions from $\mathbb{Z}_N$ to $\mathbb{C}$ **.**

Discrete Fourier transform: Let $f \in \mathbb{L}^1(\mathbb{Z}_N)$**.** Then "discrete Fourier transform of f" is denoted by Df and is a function from $\mathbb{Z}_N$ to $\mathbb{C}$ defined as: $(Df)(n) = \sum_{k=0}^{N-1} f(k)e^{-2\pi ikn/N}$.

Discrete Fourier transform operator: The map $F: L^1(\mathbb{Z}_N) \to L^1(\mathbb{Z}_N)$ defined as $F(f) = Df$ is known as " Discrete Fourier transform operator ".

Matrix associated with discrete Fourier transform of $L^1(\mathbb{Z}_N)$: We can construct a matrix M_N of order N by N such that $Df = M_N f$, where $f = [f(0), f(1), \dots..f(N-1)]^T$, $Df = [Df(0), Df(1), \dots.Df(N-1)]^T$. For getting introduced with some properties of , one can See [6]. After going through these properties one may observe that they are similar to those of Fourier transform in case of $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ functions.

Windowed Function: Let $g \in L^2(\mathbb{R})$. Then we say, g is a Window function if $tg \in L^2(\mathbb{R})$.

Windowed Fourier transform: Let $g \in L^2(\mathbb{R})$ be a fixed window function and b is a fixed real number. Let $f \in L^2(\mathbb{R})$. Then "Windowed Fourier transform of f" (induced by g) is denoted by $T_{g,b}f$ and is a function on $\mathbb{R}$ defined as $(T_{g,b}f)(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\overline{g(t-b)}\,dt$. If $g(t) = 1, \forall\, t \in \mathbb{R}$, then " Windowed Fourier transform " of f reduces to " Fourier transform of f " .

Gabor transform: If in a " Windowed Fourier transform " the Window ' g' is taken as Gaussian function, $g_a(t) = \frac{1}{2\sqrt{\pi a}}e^{-t^2/4a}$, then the "Windowed Fourier transform of f" is known as "Gabor transform of f".

Short - Time Fourier transform of f: If in case of a "Windowed Fourier transform", the window g is such that $\hat{g}$ is also a window, then the "windowed Fourier transform of f " is usually known as "Short - Time Fourier transform of f " .

Fast Fourier Transform: In many application in image Processing, optical geology signal processing etc, a large number of bits are to be processed in a very short span of time. For example a colour TV picture requires around 8 million of bits to be processed in only a second. So if we want to process, by using "Discrete Fourier transform", 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 **[26].**

Multi Resolution Analysis(MRA): A sequence of closed subspaces $(V_j)_{j\in z}$ in $L^2(\mathbb{R})$ is known as MRA if it satisfies the following properties: (i)(Increasing)$(0) \subset \dots \subset V_{-2} \subset V_{-1} \subset V_{-0} \subset V_1 \subset V_2 \subset \dots \subset L^2(\mathbb{R})$

(ii)(Separation) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ (iii) (Density) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$ (iv) (Scaling) $f(t) \in V_j$ if and only if $f(2t) \in V_{j-1}$.(iv)(Orthonormal basis)There exists $\varphi \in L^2(\mathbb{R})$, called scaling function, such that $\{\varphi(t-n): n \in \mathbb{Z}\}$ is an orthonormal basis for V_0.

Scaling function(father wavelet): The function φ appearing in above definition is known as "scaling function of MRA" or the "father wavelet.

An Example of MRA: For each $j \in \mathbb{Z}$, if we define,

$$V_j = \{f: f \in \mathbb{L}^2(\mathbb{R}) \text{ and } f \text{ is constant on } [2^{-j}n,\ 2^{-j}(n+1)] \text{ for all } n \in \mathbb{Z}\},$$

then $\{V_j\}_{j\in\mathbb{Z}}$ is an MRA with scaling function as $\varphi(x) = \begin{cases}1 & if\ x \in [0,1) \\ 0 & otherwise\end{cases}$.

The conditions in the definition of MRA are not independent – in fact the separation condition is implied by increasing ,scaling and orthonormality conditions**[28]**.

Decomposition of $L^2(\mathbb{R})$ [23]: If $\{V_j\}_{j\in\mathbb{Z}}$ is an MRA with scaling function φ, it can be shown that $\mathbb{L}^2(\mathbb{R}) = \bigoplus_{n\in\mathbb{Z}} W_n$, where $W_n = V_n^{\perp}$ in V_{n+1}. This means that every MRA produces an orthogonal direct sum decomposition of the Hilbert space $L^2(\mathbb{R})$.

Mother wavelet: Let $\{V_j\}_{j\in\mathbb{Z}}$be an MRA and $\psi \in V_1$. Then we say: ψ is a mother wavelet if $\{\psi(t-n): n\in\mathbb{Z}\}$ is an orthonormal basis of $W_O = V_0^{\perp}(\text{in}V_1)$. In **[28],** it has been shown that "if ψ is a mother wavelet, then $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$is an "orthonormal basis" for $L^2(\mathbb{R})$, where $\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k)$.

Orthonormal wavelets: If for any ψ in $L^2(\mathbb{R})$, $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2(\mathbb{R})$, then ψ is usually known as "orthonormal wavelet"**[23].** Thus**,** if ψ is a mother wavelet, then ψ is also an orthonormal wavelet**.** Orthonormal wavelets have been characterized in **[28].**

Role of MRA: (i) MRA is used to construct orthonormal wavelets.

(ii) MRA can also be used to approximate $L^2(\mathbb{R})$ 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**[29].**

(iii) With the aid of MRA, a "Fast wavelet transform algorithm" called as "pyramid algorithm" was developed by Mallat **[29]** to determine discrete wavelet coefficients and hence discrete wavelet transform of a function**.**

Filter (Low pass Filter): Let $\{V_j\}_{j\in\mathbb{Z}}, \varphi)$ be an MRA and $g \in V_1$, then the function m_g defined as $m_g(w) = \sum_{k\in\mathbb{Z}} \frac{b_k}{\sqrt{2}} e^{-ik\omega}$, where b_k satisfies $g(t) = \sum_{k\in\mathbb{Z}} \sqrt{2}\, b_k \varphi(2t - k)$ with $\sum_{k\in\mathbb{Z}} |b_k|^2 < \infty$, is known as "filter associated with g". Further, the filter m_φ, associated with the scaling function φ, is known as "Low-pass filter".

Filter equality: If $\{V_j\}_{j\in\mathbb{Z}}, \varphi)$ is an MRA then for any $g \in V_1$, it can be shown that $\hat{g}(2w) = m_g(w)\hat{\varphi}(w)\ a.e.$ **[23].** This equality is known as Filter equality.

Mother wavelet theorem: Let $\{V_j\}_{j\in\mathbb{Z}}, \varphi)$ be MRA, and consider $\psi(t) = \sum_{k\in Z}(\bar{c}_{1-k} - 1)^k(\sqrt{2}\phi(2t-k) \in V_0, \{c_k\} \in l^2(\mathbb{Z})$, such that $\hat{\psi}(w) = e^{-i\left(\frac{\omega}{2}+\pi\right)} m_\varphi\left(\frac{\omega}{2}+\pi\right)\hat{\varphi}\left(\frac{\omega}{2}\right)$, then ψ is a mother wavelet.

Continuous Wavelet Transform: Let $h \in \mathbb{L}^2(\mathbb{R})$ be fixed .Then the "continuous wavelet transform of any $f \in L^2(\mathbb{R})$", induced by h is denoted by $\mathcal{W}_h f$ and is a function from $\mathbb{R}^* \times \mathbb{R}$ to $\mathbb{C}$ defined as:

$$(\mathcal{W}_h f)(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t).\overline{h\left(\frac{t-b}{a}\right)}\, dt$$

where $\mathbb{R}^* = \mathbb{R} - \{0\}$. Also, $(\mathcal{W}_h f)(a,b)$ are known as "continuous wavelet coefficients of f ".

Admissibility condition: As with " Fourier transform", we have an Inversion formula and a Parsevel's formula, for wavelet transform as well. But for this, h should satisfy a condition known as admissibility condition: $C_h = \int_{-\infty}^{+\infty} \frac{|\hat{h}(\omega)|^2}{|\omega|} d\omega < \infty$. Another consequence of this admissibility condition is: If h , $\hat{h}$ are windows, then $\int_{-\infty}^{+\infty} h(t)dt = 0$.

Discrete wavelet transform: If we take $a = a_0^j\,; b = kb_0a_0^j$, where $a_0 > 0\,;\, b_0 \in \mathbb{R}\,; j, k \in \mathbb{Z}$, then from the definition of continuous wavelet transform, we can write: $(\mathcal{W}_h f)(a_0^j, kb_0a_0^j) = \frac{1}{\sqrt{a_0{}^j}} \int_{-\infty}^{+\infty} f(t)h(a_0^{-j}t - kb_0)dt$

Usually , the left hand side of the above equation is denoted by $(\mathcal{W}_h f)(j,k)$. Thus , now we can think $\mathcal{W}_h f$ as a function on $\mathbb{Z}\times\mathbb{Z}$ defined as : $(\mathcal{W}_h f)(j,k) = a_0^{-j/2} \int_{-\infty}^{+\infty} f(t)\ \overline{h(a_0^{-j}t - kb_0)}dt$. In this case, $\mathcal{W}_h f$ is known as " Discrete wavelet transform of f" induced by $'h'$. The numbers $(\mathcal{W}_h f)(j,k)$ 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 **[30]**. 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.

Now, let us address following two questions: • *Can we recover f, if we know only " Discrete wavelet coefficients of f"* ? • *If f gets recovered by knowing only " Discrete wavelet coefficients of f", then "is it numerically stable"*.

To answer these questions, we need to study the concept of frames in a Hilbert space.

Frame, Frame Operator, Frame analysis operator and Frame synthesis operator: A sequence $\{x_n\}$ is a Hilbert space H is said to be a frame if there exists two positive constants $A, B > 0$ such that for all $x \in H$, $A\|x\|^2 \leq |\langle x, x_n\rangle|^2 \leq B\|x\|^2$. Here, A, B are known as frame bounds. If $A = B$, then the frame $\{x_n\}$ is known as exact frame. If in a frame $\{x_n\}$ removal of just one term leaves it no longer a frame, then we say the frame is exact. Every orthonormal basis of H (Hilbert space)is a frame. A frame need not be an orthonormal basis of H (Hilbert space). If $\{x_n\}$ is a frame in a Hilbert space , then the operator $S: H \to H$ defined $S(x) = \sum_{n=1}^{\infty}\langle x, x_n\rangle\, x_n$ is known as " frame operator". It is bounded, positive and satisfies $AI \leq S \leq BI$. If S is the frame operator , then S^{-1} exists; is positive on H;satisfies $B^{-1}I \leq S^{-1} \leq A^{-1}I$ and $\{S^{-1}(x_n)\}$ is a frame (known as dual frame) with bounds $\frac{1}{B}$ and $\frac{1}{A}$. If $\{x_n\}$ is a frame in a Hilbert space , then for any $x \in H$, $x = \sum_{n=1}^{\infty}\langle x, S^{-1}(x_n)\rangle x_n = \sum_{n=1}^{\infty}\langle x, x_n\rangle S^{-1}(x_n)$. The operator T defined by $T: \mathcal{H} \to \ell^2(\mathbb{Z}), x \mapsto \{\langle x, x_i\rangle\}_{i\in\mathbb{Z}}$ is called the analysis operator. The adjoint T^* of the analysis operator is known as synthesis operator and is a function from $\ell^2(\mathbb{Z})$ to H defined as $T^*(\{c_{i\in\mathbb{Z}}\}) = \sum_{i\in\mathbb{Z}} c_i\, x_i$. If $\{h_{j,k}\}_{j,k\in\mathbb{Z}}$ forms a frame in $L^2(\mathbb{R})$ for some $h \in \mathbb{L}^2(\mathbb{R})$, then the map $T: L^2(\mathbb{R}) \to l^2(\mathbb{Z})$ defined as $:T(f) = \{\langle f, h_{j,k}\rangle\}_{j,k\in\mathbb{Z}}$ is one - to -one , linear with both T and T^{-1} as continuous.

Recovery of f: Since in the above paragraph, T is injective, f is completely determined by its " discrete wavelet coefficients" . So the answer to the first question is Yes .

Numerical stability: Also, since in the above paragraph, T^{-1} is continuous, the determination of f by its" discrete wavelet coefficients" can be done in a "numerically stable" way. So the answer to the above raised second question is also Yes.

Fast wavelet transform: Similar to " Fast Fourier Transform ", there is " Fast wavelet Transform". Let us comment on it briefly . If $\left(\{V_j\}_{j\in\mathbb{Z}}, \varphi\right)$ is an MRA with ψ as Mother wavelet, then we know that $\{\psi_{j,k}\}$ is an orthonormal basis of $L^2(\mathbb{R})$ so that for any $f \in L^2(\mathbb{R})$, we have $f = \sum_{j,k\in\mathbb{Z}}\langle f, \psi_{j,k}\rangle\psi_{j,k}$. Thus we can know f, by knowing $\langle f, \psi_{j,k}\rangle$. But usually it takes a lot of time and effort to calculate these coefficients $\langle f, \psi_{j,k}\rangle$. Mallat in 1989 discovered a fast way to compute these coefficients $\langle f, \psi_{j,k}\rangle$. The method that he proposed is known as " Fast wavelet transform" algorithm. For a detailed account of this concept, see **[28]**.

Properties of wavelets that make them extremely useful in real life[29]: Wavelets find their applications in many areas of science and technology. The main properties of wavelets that make them extremely useful are: **(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**; (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.

4. THE WAVELET THRESHOLD DENOISING METHOD: The story of wavelet based denoising methods began in early nineties when Mallat **[31]** introduced a denosing algorithm for removing white noise from signals and images by using the evolution of the wavelet transform maxima across scales. But the real surge in the research work of denoising methods, by using wavelets, began after the introduction of the *wavelet threshold denoising method* due to Donoho and Johnstone in 1994**[32],** who used a very simple approach in contrast to the Mallat's method of using wavelet maxima across the different scales**[31]**.

The concept of wavelet threshold denoising method [33]: *Wavelet threshold denoising method* tries to eliminate the noise present in a signal, but at the same time preserving the signal characteristics, regardless of its frequency content. The method involves three steps: **Step 01:** A linear forward Wavelet transform. **Step 02:** A Nonlinear thresholding step. **Step 03:** A linear inverse Wavelet transform.

More explicitly, we can put the *Wavelet threshold denoising method* as: If $Z(t)$ represents the observed signal, $U(t)$ represents the uncorrupted signal, N(t) represents the additive noise, so that we have $Z(t) = U(t) + N(t)$, $W(.)$ represents the forward wavelet transform, $W^{-1}(.)$ represents the inverse wavelet transform and $D(.\,,\lambda)$ denotes the threshold function with threshold λ, then our aim of removing the noise from $Z(t)$ to recover the signal $\widehat{U}(t)$ as an approximation of the original signal $U(t)$, can be achieved by going through the following three steps: **Step 01:** $Y = W(Z)$. **Step 02:** $X = D(Y, \lambda)$. **Step 03:** $\widehat{U} = W^{-1}(X)$.

In step one, what actually we do is: we obtain wavelet decomposition of the signal as said in the first section of this paper. Then in step 02, we modify the wavelet coefficients (detail coefficients, see section one of this paper). To modify the coefficients, firstly we select a number λ, known as threshold value, by using some standard techniques (some are given in section 5 and some are explained in section 6 of this paper) and then we select the modification method(depending on the threshold value λ), known as the threshoding function $D(.,\lambda)$. Finally, in step 03, we reconstruct the *denoised signal* by using inverse wavelet transform.
Whenever**,** we obtain the denoised signal by using the above method, we also measure the performance of the method by using some standard methods (some are explained in the section 6 of this paper).

Basic assumptions while applying thresholding [33]: (i) In the wavelet domain, noise remains dominated in the small wavelet coefficients while the coefficients having big absolute values carry less noise and more signal information.

(ii) In the wavelet domain, a sparse signal is obtained because of the decorrelating property of wavelet transform.

(iii) Noise is spread out equally along all wavelet coefficients.

(iv) In the observed data, the noise level is not too high so that the wavelet coefficients that carry signal information can be separated from the noisy ones.

5 THRESHOLD FUNCTIONS

In this section we shall present a survey on the various threshold functions available in the literature since the proposal of two classical threshold functions – *Hard threshold function* and *soft threshold function*. We shall also present the algorithm followed by the authors for the determination of the *threshold values* and the evaluation criterions followed by the authors for checking the strength of the method.

● As said earlier, in 1994, Dhono and Johnstone actually proposed the above wavelet thresholding denoising method. They proposed two threshold functions - *hard threshold function* and *soft threshold function* - for this purpose**[34]**. They defined the so called hard threshold function as: $D(x,\lambda) = x \; if \; |x| > \lambda$ and $D(x,\lambda) = 0 \; if \; |x| \le \lambda$, with x, λ representing wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denoting the wavelet coefficient obtained after processing the wavelet coefficient x by the hard thresholding. They defined the soft threshold function as: $D(x,\lambda) = sgn(x)(|x| - \lambda) \; if \; |x| > \lambda$ and $D(x,\lambda) = 0 \; if \; |x| \le \lambda$, with, as earlier, x, λ representing wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denoting the wavelet coefficient obtained after processing the wavelet coefficient x by the soft thresholding. Also, $sgn(x)$ denotes the sign of x. Note that Hard thresholding is actually a "keep or kill" procedure. It kills the wavelet coefficients, with absolute values less than or equal to the threshold value λ and retains the other coefficients. On the other hand, the Soft thresholding actually throws away the wavelet coefficients x, having absolute values less than or equal to the threshold value λ and shrinks the nonzero coefficients towards zero by an amount of λ(positive). Further, note that unlike the hard thresholding function, the soft thresholding function is continuous, although has a discontinuous derivative. If one looks at the above two thresholding functions, hard and soft, it looks that hard thresholding is a natural choice over the soft thresholding, but it is not the case as the continuity of the soft thresholding function makes it more valuable. Actually, Soft thresholding makes algorithms theoretically more tractable **[35, 36].** Moreover, there are some algorithms such as the GCV procedure **[37, 36],** with which hard thresholding does not even work, where as Soft thresholding can be used with all the algorithms because of its achieving nearly minimax rate over a huge number of Besov spaces**[35]**. Moreover, sometimes, it may happen that some pure noise coefficients may pass the hard threshold and appear as annoying 'blips' in the output, where as Soft thesholding shrinks these false structures **[36]**. It has been observed that the hard thresholding may be unstable and sensitive even to small changes in the signal, where as the soft thresholding can create unnecessary bias when the true coefficients are large **[38]**. Due to the discontinuity of the Hard thresholding function, the wavelet coefficients obtained, after processing by hard thresholding function, have a bad continuity, which causes the oscillation and very poor smoothness of the reconstructed signal. On the other hand, there are deviations between the estimated wavelet coefficients processed by soft thresholding function and the actual wavelet coefficients, which brings out the distortion of there constructed signal **[39, 19]**. To overcome above highlighted deficiencies in hard and soft thresholding functions, a lot of new thresholding functions have been introduced by the researchers. Below, we will list many of them.

● As said above, when the wavelet coefficients get reduced by the threshold value, then, in soft threshholding, many times the reconstructed image gets motivated by some deviations **[36]**. To get rid off these drawbacks, in **1998,** Gao **[40]** used the following Breiman's non-negative Garrote function as thresholding function. $D(x,\lambda) = 0 \; if \; |x| \le \lambda$ and $(x,\lambda) = x - \lambda^2 x^{-1} \; if \; |x| > \lambda$, where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by the proposed thresholding. This function is continuous and shrinks or kills the coefficients and offers samples of advantages. In the paper, the author showed that estimates yielded by this thresholding function shares the same asymptotic convergence rate as the hard and the soft shrinkage estimates do. Moreover, by using simulations, the author has shown that thresholding function offers more advantages over both hard shrinkage and soft shrinkage. The author studied minimax thresholds and the threshold selection procedure based on Stein's Unbiased Risk Estimate (SURE) for this thresholding function. The author also proposed a threshold selection procedure by combining Coifman and Donoho's cycle-spinning and SURE and called the procedure SPINSURE. With the help of examples, the author has shown that SPINSURE is more stable than SURE. The parameter used by the authors, for evaluating the success of the method, is Signal Noise Ratio(SNR).

● In the process of getting rid of shortcomings of hard and soft thresholding functions, in **1998**, **Zhang and Desai[41],** inspired by differentiable Sigmoid function proposed following threshold function that has a continuous derivative.

$$D(x,\lambda) = \begin{cases} x + \lambda - \lambda(2k+1)^{-1} : x \in (-\infty, \lambda) \\ (2k+1)^{-1} \lambda^{-2k} x^{2k+1} : |x| \le \lambda \\ x - \lambda + \lambda(2k+1)^{-1} : x \in (\lambda, +\infty) \end{cases},$$

where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, k is a positive integer. Moreover, since $D(\omega_{j,k},\lambda) \to 0$ as $|\omega_{j,k}| \to \lambda$ and $D(\lambda,\lambda) = 0$, it follows that the thresholding function is continuous at λ, thus having the same properties which the soft thresholding function has. Also, as $k \to \infty$, then the threshold function tends to the soft threshold function introduced above, so that we can say that this function is asymptotically closer to the soft threshold function. The authors have shown that the proposed threshold functions made possible the construction the adaptive algorithm. For the selection of the threshold value, the authors used Stein's unbiased risk estimate (SURE). The authors have used MSE method for evaluating the performance of the method. With the help of many numerical examples, the authors justified that the proposed denoising method is wonderfully effective in adaptively finding the optimal solution in a MSE sense.

● In **2001**, **Zhang [42]** proposed a type of Thresholding Neural Network (TNN) for adaptive noise reduction. The author proposed following thresholding functions to serve as the activation function of the TNN. $D(x,\lambda)=x+0.5\sqrt{(x-\lambda)^2+m}-\sqrt{(x+\lambda)^2+m}$, and $D(x,t)=\left(\left(1+e^{\frac{-x+\lambda}{\mu}}\right)^{-1}-\left(1+e^{\frac{-x-\lambda}{\mu}}\right)^{-1}+1\right)x$, where x, λ represent wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, m and μ are the user – defined fixed parameter. Observe that, in both the cases, $D(x,\lambda)$ has all higher order drrivitaives for positive real values of m and μ. The first function he called as *a new type of soft thresholding function*, because if we set $m=0$ then the function becomes identical to the standard soft thresholding function. Also, the author called the second function as *a new type of hard thresholding function* because if we allow $\mu\to\infty$, then the function becomes identical to the standard hard thresholding function. Since to have better numerical properties in learning algorithms of a neural network, the network activation function (thresholding function) should have higher order derivatives, so the proposed thresholding functions prove to be better than the usual hard and soft thresholding functions − which do not have higher order derivatives. Moreover, on using the gradient of the above proposed thresholding functions, w. r. t. λ, in the learning algorithm, better adjustability is achieved since the functions have nonzero derivatives for all λ. This is one of the main reasons that above proposed thresholding functions offer better numerical properties in comparison to standard thresholding functions (hard and soft). As can be seen from the above proposed thresholding functions, large values of m and μ, offer more adjustability to the thresholding functions, because they possess higher order derivatives for $|x|<\lambda$. However, too large values of m and μ, decrease the thresholding ability because when m and μ approch infinity, then the thresholding functions become linear. Thus a smaller values of m and μ are suggested so that all good properties of standard thresholding techniques are been preserved. The author discussed in detail the optimal solution of the TNN in the sense of MSE and established that number of optimal solutions are not more than one in case of above proposed soft threhholding TNN. The author also analyzed optimal performances of both the above proposed thresholding TNNs and compared the performance with the linear noise reduction method. The author presented Gradient-based adaptive learning algorithms - supervised and unsupervised batch learning as well as supervised and unsupervised stochastic learning - for obtaining the optimal solution. Using some numerical results, the author showed that the TNN is highly efficient in calculating the optimal solutions of thresholding methods in an MSE sense and more often shows better performance than other denoising methods.

● In the process of getting rid of shortcomings of hard and soft thresholding functions, in **2004**, Fang and Huang [**43**] proposed following threshold function for dnoising purposes. $D(x,\lambda)=\begin{cases} x\left(|x|^{\alpha}-\lambda^{\alpha}\right)|x|^{-\alpha} : |x|\geq\lambda \\ 0 : otherwise\end{cases}$, where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also α is a parameters which when approaches to one, the proposed function approaches to soft thresholding function and when approaches to infinity, the proposed function approaches to hard thresholding function. The authors with the help of experimental results and power spectral analysis, have shown that this threshold function produces better results in comparison to hard and soft thresholding functions.

● In **2007**, Sahraeian et al. **[44]** proposed a type of Thresholding Neural Network (TNN) for image noise reduction. The author proposed a threshold function to serve as the activation function of the TNN, which reads as: $D(x,\lambda)=\begin{cases} a\left(e^{b|x|}-1\right)\operatorname{sgn}(x) : |x|\leq\lambda \\ \left(|x|+ce^{-b|x|}\right)\operatorname{sgn}(x) : x\notin[-\infty,+\infty]\end{cases}$, where x, λ represent wavelet coefficient of the noisy signal and the threshold value, respectively, and $\hat{}D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Here, a,b,c are parameters with b determining the degree of threshold effect and a,c are such that the threshold function and its derivative are continuous at the threshold value λ. Also, as said by the authors the above proposed function keeps the advantageous properties of standard hard thresholding function such as preservation of the detail characteristic of image edges. The authors also used a cycle-spinning-based technique for the reduction of image artifacts. The experimental results performed by the authors show that the proposed by the authors offers better results than other denoising techniques, in terms of PSNR and visual quality.

● While studying Digital subtraction angiography (DSA), in **2008**, Zhang et. al. [**45**] proposed following threshold function for dnoising purposes . $D(x,\lambda)=\begin{cases} \operatorname{sgn}(x)\sqrt[u]{\left(|x|^{u}-\alpha|\lambda|^{u}\right)} : |x|\geq\lambda \\ 0 : otherwise\end{cases}$, where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, α and u are parameters which are be freely adjustable, whose appropriately chosen values can provide an effective denoising. The exponent u appearing in the expression of the above threshold function has been used for the purpose of increasing the difference between the signal and the noise. It can be observed that the above function becomes the hard threshold function if we take $\alpha = 0$. It has been observes that α is carefully selected between 0 and 1, then the use of this function produces optimal results**[46]**.

• In the process of getting rid of shortcomings of hard and soft thresholding functions, in **2009**, Huaigang et al. [**47**] proposed following threshold function for dnoising purposes.

$$D(x,\lambda)=\begin{cases} x+\lambda(k-1)-0.5k\lambda^m x^{1-m} : x\in(\lambda,+\infty) \\ 0.5k|x|^{\left[m+k^{-1}(2-k)\right]}\lambda^{-\left[m+k^{-1}(2-k)\right]}\operatorname{sgn}(x) : |x|\le\lambda \\ x-\lambda(k-1)-0.5k\left(-\lambda^m\right)x^{1-m} : x\in(-\infty,-\lambda) \end{cases},$$

where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, m, k are turning parameters. On turning parameter k, the above proposed function can be made in between the hard- and the soft-thresholding functions and on tuning the other parameter m, the near-optimum thresholding function is adjusted to the optimum one by applying small changes. As claimed by the authors, the optimization of the parameter k works similar to a global search and optimization of the parameter m works like a local search in finding the best thresholding function. Note that the hard and soft threshold functions replace the coefficients with mod below the threshold value to zero, while the proposed function replaces them by a polynomial. Moreover, the coefficients that were below the threshold value and close to it were attenuated to a value less than the far coefficients . With the help of many numerical examples, the authors justified that with a change in parameters, the proposed denoising method produces better results than the traditional hard and soft thresholding methods. The authors have used RMSE and SNR for evaluating the performance of the method.

• **In 2009, Nasri and Nezmabadi – Pour[48],** proposed following thresholding function for removing noise in images.

$$D\left(t_{j,k},\lambda_j\right)=\begin{cases} t_{j,k}-0.5\lambda_j^{\,2}\left(t_{j,k}\right)^{-1} : \left|t_{j,k}\right|\ge\lambda_j \\ 0.5t_{j,k}^3\lambda_j^{\,-2} : t_{j,k}\in\left(-\lambda_j,+\lambda_j\right) \end{cases},$$ where

$t_{j,k}$, λ_j represent wavelet coefficient of the noisy signal and the threshold value, respectively, and $D\left(t_{j,k},\lambda_j\right)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient $t_{j,k}$ by thresholding. In order to improve further the efficiency of denoising method, the authors used the above proposed thresholding function in a new subband-adaptive thresholding neural network. The authors proposed many new adaptive learning methods and used them to suppress two very important type of noises - Gaussian and speckle – in case of natural images, ultrasound images and SAR images. The simulation results, conducted by the authors, show that the above proposed thresholding function offers better results in comparison to conventional methods when used with the proposed adaptive learning types.

• In **2010**, Cai-lian et al. [**49**] proposed following threshold function for dnoising purposes

$$D(x,\lambda)=x\left[1-e^{-\left(x\lambda^{-1}\right)^2}\right]\left[1+e^{-\left(x\lambda^{-1}\right)^2}\right]^{-1},$$ where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Observe that this function is continuous and thus will share sproperties of soft threshold functions also. By using some numerical examples, the authors have shown that the usage of the above threshold function produces better results than some of the other classical threshold functions. The authors used PSNR and MSE for the evaluation of the method.

• In the process of getting rid of shortcomings of hard and soft thresholding functions, in **2010**, Lin and Cai [**50**] proposed following threshold function for dnoising purposes.

$$D(x,\lambda)=\begin{cases} \sqrt{1-\beta^2}\,sig(n)|x-\lambda|+\beta x : |x|\ge\lambda \\ 0 : otherwise \end{cases},$$ where x, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also $\beta=\exp\left[-\left(|x|-\lambda\right)^2 k^{-1}\right]$ and k is a positive real number. According to the authors this threshold function is better than the traditional soft- and hard-threshold functions – it overcomes the problem of an invariable dispersion, between the modified wavelet coefficients and the decomposed wavelet coefficients, which arises in case of using the soft threshold function and is also more elastic than the hard soft- and hard-threshold functions. Moreover, the simulation conducted by the authors tells us that the removal of noise by using this new threshold function suppresses the Pseudo-Gibbs phenomena which arise near the singularities of the signal. Moreover, the authors have shown that the use of this new function produces better SNR than the use of the traditional hard- and soft-threshold functions.

• In **2011**, **Norouzzadeh and Rashidi [51]** proposed a threshold function which being continuous and differentiable is suitable for gradient decent learning methods such as TNN. This function is used by the TNN and threshold values for wavelet sub-bands are estimated according to least mean square (LMS) algorithm. The experimental results show improvement in noise reduction from images based on visual assessments and PSNR comparing with well-known thresholding functions. The authors used the proposed function with the TNN and threshold values for wavelet sub – bands as per the LMS algorithm. The experiment conducted by the authors show that this proposed function offers better results in comparison to the classical threshold functions.

● In **2015**, Chen et. al. **[52]** proposed following, (so called Semisoft threshold function) threshold function for dnoising purposes: $D(x,\lambda) = sgn(x)(|x| - T\lambda)\ if\ |x| \geq \lambda$ and $D(x,\lambda) = 0\ if\ |x| < \lambda$, where**,** T lies in the interval [0, 1]. Observe that when T = 1, the function reduces to soft threshold function and when T = 0, the function reduces to hard threshold function.

● In case of image processing, although the standard hard and soft thresholding functions are the oldest one and has been used widely, but it has been seen, in case of application of soft thresholding, that the edges of the image get blurred because of the shrinking of some of the wavelet coefficients, where as in case of application of hard threshoding, the image shows visual distortion because of the discontinuity of the hard thresholdig function at the threshold. So in **2015**, **Wang et. al. [53]**, proposed following thresholding function.

$$D(\omega_{ij},\lambda)=\begin{cases}\mathrm{sgn}(\omega_{ij})\left[|\omega_{ij}|-\sin\left(2^{-1}\pi|\omega_{ij}^{-1}\lambda|^{n}\right)\lambda\right]:\omega_{ij}\notin(-\lambda,+\lambda)\\ 0:\omega_{ij}\in[-\lambda,+\lambda]\end{cases},$$

where ω_{ij},λ represent wavelet coefficient of the noisy image and the threshold value, respectively, and $D(\omega_{ij},\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient ω_{ij} by the thresholding function. Also, $\mathrm{sgn}(\omega_{ij})$ denotes the sign of ω_{ij} and n is the adjustment factor. It can be seen that $D(\omega_{ij},\lambda)$ approaches to zero if $|\omega_{ij}|$ approachs to the threshold value λ, proving that $D(\omega_{ij},\lambda)$ is continuous at λ and hence Psudo Gibbs phenomenon gets avoided resulting in to the avoidance of image distortion. Also, it can be seen that $D(\omega_{ij},\lambda)$ approaches to ω_{ij} if $|\omega_{ij}|$ approaches to infinity, and hence the problem of fixed deviation (in case of soft thresholding) gets solved resulting resulting in to the avoidance of blurring of image edges. Also, the value of the parameter n, can be adjusted depending on the aprticulat application and choice of the threshold value. The authors have used the threshold value given by the expression $\lambda_j = \delta_j\sqrt{2lb(m)}\,(j+2)^{-1}$, where λ_j represents the threshold of the level j, m represents the total number of the wavelet coefficients and j represents the level of the decomposition. Also, δ_j represents the $(0.6745)^{-1}median|\omega_{ij}|$ with ω_{ij} as the high frequency coefficients of the horizontal, vertical, and diagonal directions of the layer . The parameters used by the authors, for evaluating the success of the method, is Peak Signal Noise Ratio(PSNR).

● **In 2015, He et al[54],** proposed following thresholding function for removing noise:

$$D(x,\lambda)=\begin{cases}x-\lambda+\lambda\tanh\left(\alpha\lambda^{-1}(x-\lambda)\right):x\in(\lambda,+\infty)\\ 0:|x|\leq\lambda\\ x+\lambda+\lambda\tanh\left(\alpha\lambda^{-1}(x+\lambda)\right):x\in(-\infty,-\lambda)\end{cases}$$

, where x, λ represent wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(x,\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, α is a positive real number, called as shape parameter. To determine the optimal value of the shape parameter α, the authors used a strategy based on artificial fish swarm algorithm. Since $D(x,\lambda)\to 0$ as $|x|\to\lambda$ and $D(\lambda,\lambda)=0$, it follows that the above proposed thresholding function is continuous at λ, thus having the all the properties which the soft thresholding function has. Also, it is clear that the above thresholding function is monotonic w. r. t. x, ensuring that the above thresholding function has variation trend same to the hard and soft thresholding functions. Also, due to existence of higher order derivatives, w. r. .t. x, of the above proposed thresholding function in the intervals $(-\infty, -\lambda]$ and $[-\lambda, \infty)$, the denoised signal will have enough smoothness. Observe that if $\alpha \to 0$, the above proposed function approaches to the soft thresholding function and when $\alpha \to \infty$, then the above proposed function approaches to the hard thresholding function. The authors have obtained the threshold value λ by using the universal threshold formula $\lambda=\sigma\sqrt{2\log(N)}$, where σ^2 stands for the average noise variance, and N represents the signal length. The parameter used by the authors, for evaluating the success of the method, is Mean square Error(MSE). The simulation conducted by the authors, showed that the denoising done by the above thresholding function achieves better results in comparison to Hard and soft thresholding functions.

● In the year **2016**, the authors of **[55]** have proposed following thresholding function for denoising.

$$D(\omega_{j,k},\lambda)=\begin{cases}\mathrm{sgn}(\omega_{j,k})\left\lfloor\omega_{j,k}-\exp^{-3}\left\lfloor\alpha\left(|\omega_{j,k}|-\lambda\right)\lambda^{-1}\right\rfloor\lambda\right\rfloor:|\omega_{j,k}|\geq\lambda\\ 0:\omega_{j,k}\in(-\lambda,+\lambda)\end{cases}$$

, where $\omega_{j,k}$, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(\omega_{j,k},\lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient $\omega_{j,k}$ by thresholding. Observe that the adjustment factor $\exp^{-3}\left\lfloor\alpha\left(|\omega_{j,k}|-\lambda\right)\lambda^{-1}\right\rfloor$ of the above thresholding function is quite different from the semi soft threshold function and has more adaptability. Also, α is known as the normal number and can be adjusted freely and has different values for the different signal. Moreover, since $D(\omega_{j,k},\lambda)\to 0$ as $|\omega_{j,k}|\to\lambda$ and $D(\lambda,\lambda)=0$, it follows that the thresholding function is continuous at λ, thus having the same properties which the soft thresholding function has. Also, as $|\omega_{j,k}|\to\infty$, then $D(\omega_{j,k},\lambda)\to\omega_{j,k}$. This not only reduces the discreteness of the hard threshold function, but can also avoid the constant deviation problem in soft thresholding function at the same time. The variable α appearing in the thresholding function is very crucial and a variation in its value can directly effect the noise. Observe that if $\alpha = 0$, the the above function becomes the soft thresholding function and when $\alpha = \infty$, then the above function becomes hard thresholding function. The authors have obtained the threshold value by using the formula $\lambda = \sigma_n \sqrt{2\log(N)}\left(\log_2(j+1)\right)^{-1}$, where σ_n denotes the noise standard deviation, N stands for the length of the signal and j denotes the decomposition level. This threshold determination formula is actually an improvement over the commonly used formula $\lambda = \sigma_n \sqrt{2\log(N)}$ (which uses the same threshold value for each decomposition level j and hence enhancement effects are not obtained as one desires for) and uses different threshold values for different decomposition levels j. It can be seen that as the decomposition level j increases, value of λ gradually reduces, which is in agreement with the propagation characteristics of the noise at different scales, thus ensuring that the denoising is effective. The parameters used by the authors, for evaluating the success of the method, are Signal Noise Ratio(SNR) and Root mean squared Error(RMSE). The simulation conducted by the authors, showed that the denoising done by the above thresholding function achieves better results in comparison to Hard, soft and Garrote thresholding methods.

● In the year **2017**, the authors of **[19]** have proposed following thresholding function for denoising of "Vehicle Platform Vibration Signals(VPVS)" taking in to consideration the characteristics of the VPVS and the activation function of Threshold Neural Network (TNN).

$$D\left(t_{j,k},\lambda_j\right) = \begin{cases} t_{j,k} - 0.5\lambda_j - e^{t_{j,k}\lambda_j^{-1}-1} 0.2\lambda_j : t_{j,k} \geq \lambda_j \\ 0.3\lambda_j^{-3}\left|t_{j,k}\right|^4 \operatorname{sgn}\left(t_{j,k}\right) : t_{j,k} \in (-\lambda_j, +\lambda_j) \\ t_{j,k} + 0.5\lambda_j + e^{t_{j,k}\lambda_j^{-1}+1} 0.2\lambda_j : t_{j,k} \leq -\lambda_j \end{cases}, \text{ where}$$

$t_{j,k}$, λ_j represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D\left(t_{j,k},\lambda_j\right)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient $t_{j,k}$ by thresholding. In fact, while proposing the above threshold function, they considered that VPVS has a main vibration frequency range; it is mixed in two main kinds of noises – trend and random noise, the trend is been distributed in the low frequency whereas the random noise is been distributed in the high frequency and the activation function of TNN is continuous and derivable. Note that the wavelet coefficients having absolute values less than the threshold value are modified by a power function, instead of setting them to zero as in the case of hard or soft thresholding functions, and hence prevents the useful information loss. Moreover, Furthermore, unlike the hard thresholding function, the above mentioned thresholding function and its derivative are both continuous and resulting in to the prevention of oscillation and bad smoothness of the reconstructed signal. The authors have obtained the threshold value through Unsupervised Learning of TNN. The parameters used by the authors, for evaluating the success of the method, are Signal Noise Ratio(SNR) and Root mean squared Error(RMSE). The simulation conducted by the authors, showed that the denoising done by the above thresholding function achieves better results in comparison to five other denoising methods(Hard and soft thresholding methods, the denoising methods proposed by the authors of **[42], [48]** and**[53].**

● **In 2017, Dedha and Melkemi [36],** proposed following thresholding function.

$$D(x,\lambda) = \begin{cases} x - \lambda + \dfrac{2\lambda}{\sqrt{\pi}} \displaystyle\int_0^{\alpha\left(\frac{x-\lambda}{\lambda}\right)} e^{-t^2} dt : x \in (\lambda, \infty) \\ 0 : |x| \leq \lambda \\ x + \lambda - \dfrac{2\lambda}{\sqrt{\pi}} \displaystyle\int_{\alpha\left(\frac{x+\lambda}{\lambda}\right)}^{0} e^{-t^2} dt : x \in (-\infty, -\lambda) \end{cases}, \text{ where } x \ , \lambda$$

represent wavelet coefficient of the noisy signal and the threshold value, respectively, and $D\left(x,\lambda\right)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by thresholding. Also, α is known as the shape parameter and can be adjusted freely. Moreover, since $D\left(x,\lambda\right) \to 0$ as $|x| \to \lambda$ and $D\left(\lambda,\lambda\right) = 0$, it follows that the thresholding function is continuous at λ, thus having the all the properties which the soft thresholding function has. Also, if $|x| \to \infty$, then $D\left(x,\lambda\right) \to x$, the above proposed function has the same asymptotic convergence rate as the Hard and Soft thresholding function. Observe that if $\alpha \to 0$, the above proposed function approaches to the soft thresholding function and when $\alpha \to \infty$, then the above proposed function approaches to the hard thresholding function. The above proposed thresholding function is continuous on $(-\infty, \infty)$, monotonic on $(-\infty, \infty)$, has higher order derivatives in $(-\infty, -\lambda]$ and $[-\lambda, \infty)$, thus having much better properties than the classical hard and soft thresholding functions. The authors have obtained the threshold value by using the formula $\lambda = \sigma\sqrt{2\log(N \times M)}$, where σ^2 stands for the noise variance, and $N \times M$ represents the size of the image. The parameters used by the authors, for evaluating the success of the method, is Peak Signal to Noise Ratio(PSNR). The simulation conducted by the authors, showed that the denoising done by the above thresholding function achieves better results in comparison to Hard, soft and Garrote thresholding methods.

● In **2017**, Golilarz et al. **[56]** proposed following threshold function to denoise images (two dimensional signals) This function is

$$D\left(x,\lambda\right) = \begin{cases} \operatorname{sgn}\left(x\right)\left[|x| - \cos\left(5^{-1}\pi\left|x^{-1}\lambda\right|\right)\lambda\right] : x \notin \left(-\lambda, +\lambda\right), \\ 0.19x^3\lambda^{-2} : x \in \left[-\lambda, +\lambda\right] \end{cases}$$

where x, λ represent wavelet coefficient of the noisy image and the threshold value, respectively, and $D(x, \lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient x by the thresholding function. The authors combined this threshold function with the Un - decimated Wavelet Transform to denoise images. The authors claimed that the above proposed function is data dependent as well in addition to its smoothness and non – linearity properties. The parameters used by the authors, for evaluating the success of the method, is Peak Signal Noise Ratio(PSNR) and claimed that with the above proposed methodology, a higher value of PSNR is obtained in comparison to the other classical methods – hard and soft thresholding.

● In the year **2018,** to Estimate Dynamic Motion Parameters, **Wang et al**, **[57]** have proposed following thresholding function for denoising

$$D(\omega_{j,k}, \lambda) = \begin{cases} \operatorname{sgn}(\omega_{j,k})\left[|\omega_{j,k}| - e^{-\beta|\omega_{j,k}|}e^{\beta\lambda}\lambda\right] : |\omega_{j,k}| \geq \lambda \\ 0 : \omega_{j,k} \in (-\lambda, +\lambda) \end{cases},$$

where $\omega_{j,k}$, λ represent a wavelet coefficient of the noisy signal and the threshold value, respectively, and $D(\omega_{j,k}, \lambda)$ denotes the wavelet coefficient obtained after processing the wavelet coefficient $\omega_{j,k}$ by thresholding. Observe that the adjustment factor $e^{-\beta|\omega_{j,k}|}e^{\beta\lambda}$ of the above thresholding function is quite different from the semi soft threshold function and has more adaptability. Also, β is a non – negative real number, which when taken close to zero, the above proposed thresholding function approximates the soft thresholding function and when it is taken sufficiently high, the thresholding function approximates the hard thresholding function. The authors have obtained the threshold value by using the formula $\lambda = \sigma \frac{\sqrt{2\log(N)}}{\log_2(j+1)}$, where σ denotes the noise standard deviation, N stands for the length of the signal and j denotes the decomposition level. This threshold determination formula is actually an improvement over the commonly used formula $\lambda = \sigma\sqrt{2\log(N)}$ (which uses the same threshold value for each decomposition level j and hence enhancement effects are not obtained as one desires for) and uses different threshold values for different decomposition levels j. It can be seen that as the decomposition level j increases, value of λ gradually reduces, which is in agreement with the propagation characteristics of the noise at different scales, thus ensuring that the denoising is effective. The parameters used by the authors, for evaluating the success of the method, are Signal Noise Ratio(SNR) and Root mean squared Error(RMSE). The simulation conducted by the authors, showed that the denoising done by the above thresholding function achieves better results in comparison to Hard, soft and semi – soft thresholding. The authors have also combined the above proposed thresholding function with an inter- scale correlation method to further improve the accuracy of judgment of the wavelet coefficients near the threshold value. They have shown that the simulation conducted in this manner improves the results further.

For more threshold functions available in the literature, we refer the references given in the list of the publications cited in this paper.

6. DETERMINATION OF THE THRESHOLD VALUE:

In the processs of denoising by wavelet thresholding method, determination of a good threshold value is very important – a small threshold value may produce a result which is a very good approximation of the original signal, but may contain still a large amount of noise, where as a large threshold value, on the other hand, produces a sufficiently good smooth signal, but may destroy details, which in image(a two dimensional signal) processing may result in to blur and artifacts. So determination of the threshold value should be done very carefully. There are many methods of determining threshold value available in the literature, but the most popular and classical algorithms are minimax, universal and rigorous sure threshold estimation techniques **[34, 35, 58, 59].** We will explain here each one of these very briefly.

Minimax Threshold value: The minimax threshold value λ_M is defined as: $\lambda_M = \sigma\lambda_n^*$, , where λ_n^*, is given by $\lambda_n^* = inf_\lambda sup_d\{R_\lambda(d)[n^{-1} + R_{Oracle}(d)]^{-1}\}$ with $R_\lambda(d) = E(\delta_\lambda(d) - d)^2$ and $R_{Oracle}(d)$ named as oracle used to account for the risk associated to the modification of the value of a given wavelet coefficient. The authors considered two oracles – one is the Diagonal Liner Projection (DLP) and the other is the Diagonal Linear Shrinker (DLS) . The ideal risks for these oracles are defined as: $R_{oracle}^{DLP}(d) = min\{d^2, 1\}$ and $R_{oracle}^{DLS}(d) = d^2(d^2+1)$. Also, σ is given by $\sigma = (0.6745)^{-1} median\ (d_{L-1}, k)$ with L denoting the number of dcompositiion level. This median selection is made on the detail coefficient of the signal under consideration.

Universal threshold Value: The universal threshold or global threshold λ_{UNIV} value is given by the formula $\lambda_{UNIV} = \sqrt{2\log(N)}\,\sigma$, where N represents the signal length and σ^2 represents the noise variance given by the formula $\sigma = (0.6745)^{-1} median\ (d_{L-1}, k)$. An obvious advantage of this formula for the determination of threshold value is in software implementation because of its easy to remember and coding. Further, this formula for the determination of threshold value guarantees us that any sample in the wavelet transform for which the signal under consideration is zero is reconstructed as zero. Also, since in the asymptotic sense, Universal threshold is the optimal threshold, so it may produce better reconstructed signals when used in combination with the soft thresholding function **[ ]** if the number of samples are quite large . Moreover, it has been observed that, in general, this formula of determination of threshold value is not good to determine a threshold, but it helps to have a starting value when nothing is known about the signal condition.

Rigorous Sure (rigresure) threshold Value: Another threshold value determination method, known as Rigorous

Sure (rigresure) threshold, is in use and is given by the formula $\lambda_s = argmin_{0<\lambda<\lambda_U} Sure\ (\lambda, S(a,b)\sigma^{-1})$, where we define $Sure\ (\lambda, X) = n - 2.\Theta\ \{i: |X_i \leq \lambda|\} + [min(|X_i|, \lambda)]^2$ with the operator Θ returning the cardinality of the set $\{i: |X_i \leq \lambda|\}$. This threshold formula gives us a method in which we use a threshold λ at each resolution level of the wavelet coefficient. This threshold value uses Stein's Unbiased Risk Estimate (SURE) criterion to obtain unbiased estimate**[D5]**.Other threshold determination methods that are in use are *Universal threshold level dependent, Universal modified threshold level dependent, Exponential threshold, Exponential threshold level dependent, The modified unified threshold* **[60].**

7. EVALUATION PARAMETERS

To check the performance of a denoising algorithm, there are many methods available in the literature. But the most common in use are "Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR). The former we use in case of one dimensional signals and the the later one in case of images (two dimensional signals). We define SNR as: SNR = $10\log_{10}\left[\left(\sum_{n=0}^{N-1} x^2(n)\right)\left(\sum_{n=0}^{N-1}\left(\bar{x}(n) - x^r(n)\right)^2\right)^{-1}\right]$, where $x(n)$ represents the original signal, and $x^r(n)$ denotes the denoised signal and $\bar{x}(n)$ represents the Expectation of $x(n)$. Also, we define PSNR as: SNR = $10\log_{10}\left[L\left(\sum_{n=0}^{N-1}\sum_{m=0}^{m-1}\left(\bar{x}(n,m) - x^r(n,m)\right)^2\right)^{-1}\right]$, where $x(n,m)$ represents the original image, and $x^r(n,m)$ denotes the denoised image, $\bar{x}(n,m)$ represents the Expectation of $x(n,m)$ and L represents the quantized gray level of the image **[59].** Other methods that are being used to measure the performance of a wavelet threshold denoising method are via the calculation of Mean Square Error(MSE), defined as

MSE = $N^{-1}\sum_{N=0}^{N-1}\left(s(n) - \tilde{s}(n)\right)^2$; Root Mean Square Error(RMSE), defined as RMSE = $\sqrt{(2N)^{-1}\sum_{N=0}^{N-1}\left(s(n) - \tilde{s}(n)\right)^2}$ and Percentage Root Mean Square Difference given by $\sqrt{\left[\sum_{N=0}^{N-1}\left(s(n) - \tilde{s}(n)\right)^2\right]\left[\sum_{N=0}^{N-1}\left(s(n)\right)^2\right]^{-1}} \times 100\%$ **[61].**

REFERENCES

[1] **M. Lindenbaum, M. Fischer, and A.M. Bruckstein**, On gabor contribution to image enhancement, Pattern Recognition, 27 **(1994)**, pp. 1–8.

[2] **P. Perona and J. Malik**, Scale space and edge detection using anisotropic diffusion, IEEE Trans. Patt. Anal. Mach. Intell., 12 **(1990)**, pp. 629–639.

[3] **L. Rudin, S. Osher, and E. Fatemi**, Nonlinear total variation based noise removal algorithms, Physica D, 60 **(1992)**, pp. 259–268.

[4] **L. P. Yaroslavsky**, Digital Picture Processing - An Introduction, Springer Verlag, **1985**.

[5] **L. Yaroslavsky and M. Eden**, Fundamentals of Digital Optics, Birkhauser, **1996**.

[6] **S. M. Smith and J.M. Brady**, Susan - a new approach to low level image processing, International Journal of Computer Vision, 23 (**1997)**, pp. 45–78.

[7] **L. P. Yaroslavsky**, Digital Picture Processing - An Introduction, Springer Verlag, **1985**.

[8] **E. Ordentlich, G. Seroussi, M. Weinberger S. Verd´u, and T. Weissman**, A discrete universal denoiser and its application to binary images, in Proc. IEEE ICIP, vol. 1, **2003**, pp. 117–120.

[9] **S.A. Awate and R. T. Whitaker**, Image denoising with unsupervised, information-theoretic, adaptive filtering, Technical Report UUCS-04-013, UUCS, **2004**.

[10] **A. Buades, B. Coll and J. M. Morel**, A review of image denoising algorithms, with a new one. SIAM Journal on Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, **2005**, 4 (2), pp.490-530.

[11] **D. Donoho and I. Johnstone**, Ideal spatial adaptation via wavelet shrinkage, Biometrika, 81 **(1994)**, pp. 425–455.

[12] **R.R. Coifman and D. Donoho**, Wavelets and Statistics, Springer Verlag, **1995**, Translation invariant de-noising, pp. 125–150.

[13] **T.F. Chan and H.M. Zhou**, Total variation improved wavelet thresholding in image compression, in Proceedings of ICIP, **2000**, vol. 2, **2000**, pp. 391–394.

[14] **S. Durand and J. Froment**, Reconstruction of wavelet coefficients using total variation minimization, SIAM Journal on Scientific computing, 24 **(2003)**, pp. 1754–1767.

[15] **S. Lintner and F. Malgouyres**, Solving a variational image restoration model which involves l∞ constraints, Preprint 2003-18, University Paris 13, **2003**.

[16] **Y. Meyer**, Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, vol. 22, AMS University Lecture Series, **2002**.

[17] **E. Le Pennec and S. Mallat**, Geometrical image compression with bandelets, in Proceedings of the SPIE 2003, vol. 5150, 2003, pp. 1273–1286.

[18] **J. Starck, E. Cands, and D. Donoho**, The curvelet transform for image denoising, IEEE Transactions on image processing, 11 **(2000)**, pp. 670–684.

[19] **M. Li, Z. Wang, J. Luo, Y. Liu, and S. Cai**, Wavelet Denoising of Vehicle Platform Vibration Signal Based on Threshold Neural Network, Hindawi Shock and Vibration Volume **2017**, Article ID 7962828, 12 pages.

[20] **T. W. Yang, P. F. Chen, J. D. Han, and W. L. Xu,** "Wavelet transform analysis of optic fiber force sensing signals at puncture needle tip, "Optics and Precision Engineering, vol.23, no.8, pp. 2149 – 2157, **2015.**

[21] **R. Yan, R. X. Gao, and X. Chen**, "Wavelets for fault diagnosis of rotary machines : a review with applications, "Signal Processing, vol.96, pp.1 – 15, **2014**.

[22] **Y.-P. Zang, D.-J. Zhang, and W.-Z. Wang, "**Per-level threshold de-noising method using wavelet and its application in engine vibration analysis," Journal of Vibration and Shock, vol. 28, no. 8,pp.57–60,**2009.**

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

[24] **Z. Abbas and P. Raina (2017),** A Study on Applications of Wavelets to Data Mining, International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 13, Number 12 (2018) pp. 10886-10896.

[25] **Z. Abbas and P. Raina (2016),** Shannon Wavelet Analysis with Applications: A Survey, IOSR Journal of Mathematics, Volume 12, Issue 5 Ver. VIII, e: 2278 – 5728, P: 2318 – 765X, 08 – 20.

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

[27] **Loknath and Debnath, (2002),** Wavelet Transforms and their Applications , Birkhauser

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

[29] **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.

[30] **L. Daubechies,** 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.

[31] **S. Malllat and W. L. Hwang,** Singularity detection and processing with wavelets, IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 38, NO. 2, **1992,** PP 617 – 643.

[32] **D. L. Donoho and J. M. Johnstone,** "Ideal spatial adaptation by wavelet shrinkage, "Biometrika, Vol.81, no.3, pp.425–455, **1994.**

[33] **R. Rangarajan, R. Venkataramanan and S. Shah(2002),** Image Denoising Using Wavelets — Wavelets & Time Frequency.

[34] **D. L. Donoho and J. M. Johnstone,** "Ideal spatial adaptation by wavelet shrinkage, "Biometrika, Vol.81, no.3, pp.425–455, **1994.**

[35] **D. L. Donoho,** DeNoising by Soft Thresholding, IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 41, NO. 3, MAY **1995,** PP 613 – 627.

[36] **B. Dehda and K. Melkemi,** Image Denoising using new Wavelet Thresholding Function, Journal of Applied Mathematics and Computational Mechanics **2017**, 16(2), 55-65**.**

[37] **M. Jansen**, Noise Reduction by Wavelet Thresholding, volume 161. Springer Verlag, United States of America, 1 edition, **2001**.

[38] **N. Mustafa, J. Ping, S. A. Khan, M. Giess,** Medical Image De-Noising Schemes using Wavelet Transform with Fixed form Thresholding**,** International Journal of Advanced Computer Science and Applications, Vol. 6, No. 10, **2015.**

[39] **J. Li, C. Cheng, T. Jiang, and S. Grzybowski, "**Wavelet de- noising of partial discharge signals based on genetic adaptive threshold estimation," IEEE Transactions on Dielectrics and ElectricalInsulation,vol.19,no.2,pp.543–549,**2012.**

[40] **H.Y. Gao,** Wavelet shrinkage denoising using the non-negative garrote, J. Comput. Graph. Statist. **1998**, 7, 469-488.

[41] **X.-P. Zhang and M. Desai, "**Adaptive denoising based on SURE risk," IEEE Signal Processing Lett., vol. 10, no. 5, pp. 265–267, Oct. **1998.**

[42] **X.-P. Zhang, "**Thresholding neural network for adaptive noise reduction," IEEE Transactions on Neural Networks, vol. 12, no. 3,pp.567–584, **2001.**

[43] **H. T. Fang and D. S. Huang,** "Wavelet de-noising by means of trimmed thresholding," in Proceedings of the 5th World Congress on Intelligent Control and Automation (WCICA '04), pp. 1621–1624, Hangzhou, China, June **2004**.

[44] **S. M. E. Sahraeian, F. Marvasti, N. Sadati**, Wavelet Image Denoising Based on Improved Thresholding Neural Network and Cycle Spinning. In: Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing, (**2007**).

[45] **G. Zhang, J. Wu, and Z. Cui**, "Application of wavelet thresholding de-noising in DSA," in Proceedings of the International Symposium on

Information Science and Engineering (ISISE '08), pp. 130–134, IEEE, December **2008**.

[46] **G. S. Vijay, Kumar H. S., Srinivasa Pai P., Sriram N. S. and Raj B. K. N. Rao ,** Evaluation of Effectiveness of Wavelet Based Denoising Schemes Using ANN and SVM for Bearing Condition Classification, Hindawi Publishing Corporation Computational Intelligence and Neuroscience Volume **2012**, Article ID 582453, 12 pages doi:10.1155/2012/582453.

[47] **Z. Huaigang, W. Zhibin, and Z. Ying**, "Analysis of signal denoising method based on an improved wavelet thresholding," in Proceedings of the 9th International Conference on Electronic Measurement and Instruments (ICEMI '09), pp. 1987–1990, IEEE, August **2009.**

[48] **M. Nasri and H. Nezamabadi-pour, "**Image denoising in the wavelet domain using a new adaptive thresholding function," Neurocomputing,vol.72,no.4–6,pp.1012–1025,**2009.**

[49] **L. Cai-lian, S. Ji-xiang, and K. Yao-hong,** "Adaptive image denoising by a new thresholding function," in Proceedings of the 6th International Conference on Wireless Communications, Networking and Mobile Computing, pp. 1–5, IEEE, September **2010.**

[50] **Y. Lin and J. Cai(2010)**, "A new threshold function for signal denoising based on wavelet transform," in Proceedings of the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA '**10**), pp. 200–203, IEEE.

[51] **Y**. **Norouzzadeh and M. Rashidi,** Image denoising in wavelet domain using a new threshold function, in Proceedings of International Conference on Information Science and Technology, **2011**, 721 – 724.

[52] **X. Chen, S. Li, and W. Wang**, "New de-noising method for speech signal based on wavelet entropy and adaptive threshold," Journal of Information and Computational Science, vol. 12, no. 3, pp. 1257–1265, **2015.**

53] **X. Wang, X. Ou, B. W. Chen and M. Kim,** Image Denoising Based on Improved Wavelet Threshold Function for Wireless Camera Networks and Transmissions, Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume **2015**, Article ID 670216, 8 pages.

[54] **C. He, J. Xing, J. Li, Q. Yang and R. Wang,** A New Wavelet Thresholding Function Based on Hyperbolic Tangent Function, Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2015, Article ID 528656, 10 pages.

[55] **L. Jing-yi, L. Hong, Y. Dong and Z. Y. Sheng**, A New Wavelet Threshold Function and Denoising Application, Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2016, Article ID 3195492, 8 pages.

[56] **N. Amiri Golilarz, N. Robert, J. Addeh, A. Salehpour**, Translation Invariant Wavelet Based Noise Reduction using a New Smooth Nonlinear Improved Thresholding Function, Computational Research Progress in Applied Science and Engineering 3 (**2017**) 104–108.

[57] **Y. WANG, B. ZHANG1, F. DING and H. REN,** Estimating Dynamic Motion Parameters With an Improved Wavelet Thresholding and Inter-Scale Correlation, IEEE Access, Digital Object Identifier 10.1109/ACCESS.2018.2846290.

[58] **D. L. Donoho and I. M. Johnstone (1998)**. "Minimax estimation via wavelet shrinkage." Annals of statistics: 879-921.

[59] **Burhan Ergen,** Signal and Image Denoising Using Wavelet Transform, Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology, ISBN 978-953-51-0494-0 Hard cover, 634 pages Publisher In Tech, Published online 04, April, 2012, Published in print edition April, 2012.

[60] **O. E. Charri, R. Latif, K. Elmansouri, A. Abenaou and W. Jenkal,** ECG signal performance de‑noising assessment based on threshold tuning of dual‑tree wavelet transform, Bio Medical Engineering On – Line, El B'charri et al. BioMed Eng OnLine (2017) 16:26 DOI 10.1186/s12938‑017‑0315‑1.

[61] **G. G. Tsaneva, K. Tcheshmedjiev,** Denoising of Electrocardiogram Data with Methods of Wavelet Transform, International Conference on Computer Systems and Technologies - CompSysTech'13.