We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Wavelet Transform for Signal Processing in Internet-of-Things (IoT)

Indrakshi Dey and Shama Siddiqui

Abstract

The primary contribution of this chapter is to provide an overview of different denoising methods used for signal processing in IoT networks from the perspectives of physical layer in the network. The chapter starts with the introduction to different kinds of noise that can be encountered in any kind of wireless communication networks, different kinds of wavelet transform and wavelet packet transform methods that can be used for denoising sensor signals in IoT networks and the different processing steps that are needed to be followed to accomplish wavelet packet transform for the sensor signals. Finally, a universal framework based on energy correlation analysis has been presented for denoising sensor signals in IoT networks, and such a framework can achieve considerable improvement in denoising performance reducing the effective noise correlation coefficient to 0.00001 or lower. Moreover, this method is found to be equally effective for Gaussian or impact noise or both.

Keywords: denoising, sensor signals, Internet-of-Things (IoT), wavelet transform, wavelet packet transform, energy correlation analysis

1. Introduction

Internet of Things (IoT) refers to a network of diverse range of smart devices used in the domains of healthcare, industry, vehicles, homes, agriculture, retail, poultry and farming, and many more. Typical equipment supporting the IoT functionality include lightning, thermostats, TVs, sensors, mobile phones, speakers, voice assistants, cameras, video cameras, etc. These devices are basically deployed to facilitate the processes of monitoring and automation by transmitting and receiving information via internet. Undoubtedly, IoT has emerged as a rapidly growing ecosystem that promises to deliver unmatched global coverage, quality-of-service (QoS), scalability, security and flexibility to handle different requirements for a comprehensive list of use-cases. This has resulted in increasing number of IoT devices (relays, sensors, transceiver, actuators etc.) being deployed in in all types of urban, suburban and rural environments to cater to the innovative and emerging applications.

Since more devices and appliances have been transforming into their smarter version, we now have the applications such as smart cars with features of smart dashboards, GPS, smart doors and auto-route designed to reduce the accidents. Such applications clearly require high number of connected devices; in fact, it has

been forecasted by International Energy Agency that the estimated number of connected devices which was 15 billion in 2018 shall reach 46 billion in 2030 [1]. In addition to the IoT devices, the evolution of IoT networking technologies has also been remarkable over the past decade, where more and more IoT devices have been shifting from using Long Term Evolution (LTE) to Narrowband-IoT (NB-IoT) which offers a cost-effective and energy efficient solution for continued operation of these systems. Naturally, the connected devices are expected to transmit large volumes of heterogeneous data at high data rates, and we will be required to deal with ever-increasing radio frequency noise.

The signals carrying IoT data are highly likely to face numerous obstacles and can be corrupted by significant amount of noise present in the environment. White Gaussian model has been commonly been used to quantify the noise faced by [1]. The types of noise which have been found to degrading the quality of IoT signals vary from the impact noise resulting from high frequency interference and instantaneous disturbance on the initialization of large equipment to changing connections around the participating IoT devices [2]. All these kinds of noise negatively influence the multi-device information fusion system [3]. Such noises should be filtered out and the transmitted signal should be reconstructed back to its actual form to ensure the accuracy and reliability of the transmitted information. Here, accuracy of IoT solutions is measured in terms of the number of packets reporting correct information, deviation between the reported and actual results and the delivery to correct destination timely. Similarly, the reliability of IoT is measured using information such as failure rate of the IoT devices, average time between two consecutive failures, average repair time and probability for needing to change a component within a certain time-frame.

Although this chapter mainly deals with algorithms for signal denoising, they can be also be applied for image denoising, as images can be represented as two-dimensional signals. Consequently, signal processing techniques applicable to signals can be modified for images.

2. Noise consideration

The process of removing the noise while retaining and not distorting the quality of the received signal or image is referred to as denoising. The traditional way of denoising is to use a low or band-pass filter with cut-off frequencies. However, the traditional filtering techniques are able to remove out-of-band noise. Therefore many denoising techniques are proposed to overcome this problem.

Denoising is also an indispensable link in speech signal processing owing to the varying origins and non-stationarity, and difficulty in modeling the noise affecting the signal. Assuming that the received signal is affected by white additive Gaussian noise (AWGN) which is also stationary in nature, the received signal $y(i)$ can be represented as,

$$y(i) = x(i) + \sigma\varepsilon(i), \quad i = 0, 1, \ldots, n-1 \tag{1}$$

where $x(i)$ is the noise-free transmitted signal, $\varepsilon(i)$ representing independent normally distributed random variable and σ representing the intensity of the noise affecting $y(i)$. Reconstruction of the original signal $x(i)$ from the instantaneous set of $y(i)$ values without actual assuming a specific model for $x(i)$ or $y(i)$ is the primary aim of the process of 'Denoising'. The most common approach is to recognize noise components as the high frequency components present in the corrupted received signal, apply Fourier transform and then filter out the high frequency components.

Therefore, the most traditional way of denoising signals is based on Fourier analysis and Fourier transform.

Another common denoising method is the modulus maxima method [4]. It is based on the concept that signal and noise exhibit different characteristics when projected to their maxima in space divided in multiple scales. Magnitude scales increasing with decreasing extreme value points are filtered out to remove noise and the extreme value points themselves are reconstructed back [5]. The modulus maxima method in addition to the noise effect is better than any other method when mixed with white noise and singular information is significant, but the computational complexity is quite high. However, Fourier transform based denoising is restricted due to its weakness in obtaining partial characteristic of the transmit signals and possible Gibbs phenomenon [6]. If the signal has the same frequency as the noise, filtering out those frequency components will cause noticeable loss of information of the desired signal when considering the frequency representation of the signal.

3. Wavelet transform

Wavelet Transform (WT) has emerged as a powerful tool for signal and image denoising and processing, that have been successfully used in many scientific fields such as signal processing, image compression, computer graphics and pattern recognition [7, 8]. On contrary to the traditional Fourier transform, WT is particularly suitable for application of non-stationary signals which may instantaneously vary in time. Primarily, the received signal is divided into different frequency components using wavelets. The basis function of WT is scaled based on frequency and a subset of small waves (known as mother wavelet) is used for implementing WT [9]. The mother wavelet is a time-varying window function used for decomposition of $x(i)$ into weighted sets of scaled versions of $y(i)$. Consequently, using wavelet transform in signal processing is the process of the partial transformation of the spatial domain and the frequency domain, in order to get useful information accurately from it though corrupted with noise.

Since different frequency levels are used for WT, it is quite convenient for analyzing the signal characteristics at different frequencies and detecting removing corrupting noise. Broadly, there are two types of WT, Continuous WT (CWT) and Discrete WT (DWT).

3.1 Continuous wavelet transform (CWT)

CWT measures the congruence between an analyzing function and actual signal by calculating the inner product and then integrating the product. The mother wavelet window function can be shifted and moved over the time-axis by changing scale and position parameters, thereby including different frequency components at the different locations. Mathematically CWT can be represented as,

$$\text{CWT}(a, b; x(i), \psi(i)) = \int_{-\infty}^{\infty} x(i) \frac{1}{a} \psi^* \left(\frac{i - b}{a} \right) di \qquad (2)$$

where $x(i)$ is the transmit signal, $\psi(i)$ is the analyzing function (wavelet), a is the scale parameter, b is a position in time and $*$ represents complex conjugate. Considering $\psi(i)$ as the band-pass impulse response, scaling the wavelet varies the bandwidth of the band-pass filter. CWT allows changing the support of the wavelet to get better resolution in frequency domain. CWT can be realized on computer and

the computation time can be significantly reduced if the redundant samples are removed after using the sampling theorem.

3.2 Discrete wavelet transform (DWT)

If suitable transformation is applied to a group of selected wavelet, a collection of orthogonal real-valued wavelets will be generated, a representation of the received signal referred to as wavelet expansion. In this case, the properties of the generated wavelets depend on the features of the mother wavelet. Since the newly generated wavelets are a group of orthogonal wavelets, they provide a time-frequency localization of the actual input signal, thereby concentrating the signal energy over a few frequency coefficients. Scaling and translation of the mother wavelet generated. If the scaling factor is a power of two, the wavelet transform technique is referred to as the dyadic-orthonormal wavelet transform [10]. If the chosen mother wavelet has orthonormal properties, there is no redundancy in the discrete wavelet transforms. In addition, this provides the multiresolution algorithm decomposing a signal into scales with different time and frequency resolution [9].

DWT is an implementation of WT using mutually orthogonal set of wavelets defined by carefully chosen scaling and translation parameters (*a* and *b*), such that the normalized area between the analyzing functions is unity, leading to a very simple and efficient iterative scheme for doing the transformation [11]. The translation equation can be expressed as,

$$\text{DWT}\big[n, a^j\big] = \sum_{m=0}^{N-1} x[m]\psi_j^*[m-n]; \quad \psi_j[n] = \frac{1}{\sqrt{a^j}}\psi\left(\frac{n}{a^j}\right) \tag{3}$$

where n is the time delay introduced, N is the signal length and ψ is the discrete mother wavelet windowing function. DWT operates on discrete wavelet sets thereby yielding signal compression and reducing the computational complexity considerably. Moreover, DWT provides better spatial and frequency localization, as compared to other multi-scale signal maxima representation, thereby eliminating redundancy. In DWT, signal is decomposed into 'approximation' and 'detail' coefficients at each level [12].

The process is repeated at multiple levels, a technique equivalent to consecutive iterations of low pass and high pass filtering. As a result, the low frequency and high frequency components of $x(t)$ yield the approximation and detailed coefficients respectively, which can be mathematically expressed as,

$$x(t) = \sum_{m=1}^{L}\left[\sum_{k=-\infty}^{\infty} D_m(k)\psi_{m,k}(t) + \sum_{k=-\infty}^{\infty} A_l(k)\phi_{l,k}(t)\right] \tag{4}$$

Where $D_m(k)$ is the detailed coefficient, $A_l(k)$ is the approximation coefficient, $\psi_{m,k}(t)$ is 2^m-scale discrete analyzing function, and $\phi_{l,k}(t)$ is the 2^l-scale scaling function. After scaling and wavelet filtering, we get [13].

$$h(n) = 2^{-1/2} <\phi(t), \phi(2t-n)>$$
$$g(n) = 2^{-1/2} <\psi(t), \phi(2t-n)> = (-1)^n h(1-n) \tag{5}$$

The approximation and the detailed coefficients are compared by applying FIR filter bank. The filter bank uses a low-pass filtering h for generating the approximation coefficients and high-pass filtering g for generating the detailed coefficients,
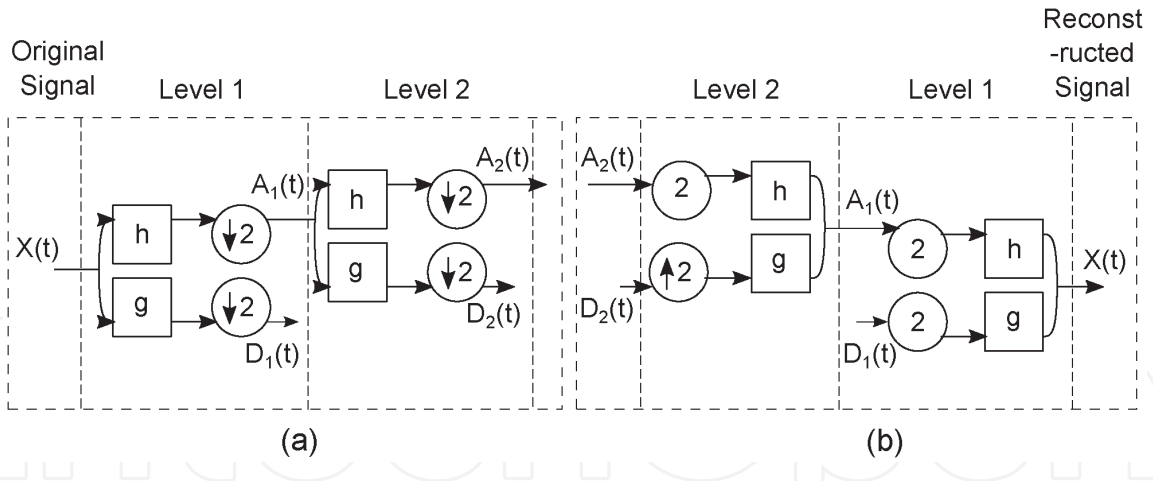
Figure 1.
The DWT decomposition and reconstruction steps of a 1D signal for level of 2; (a) decomposition, (b) reconstruction.
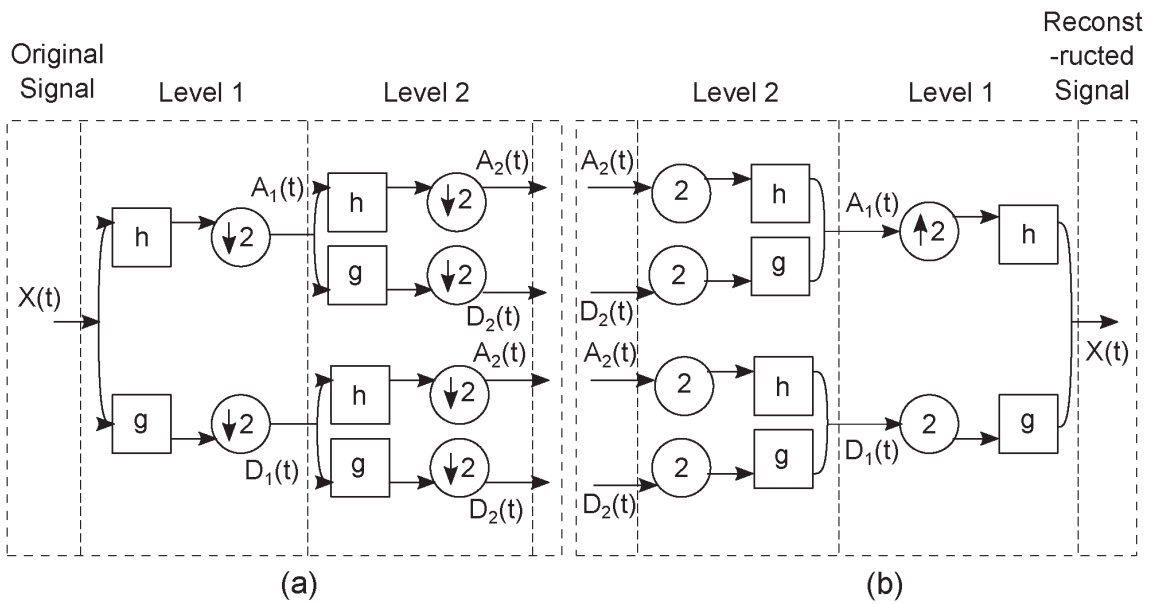

Figure 2.
The wavelet packet decomposition and reconstruction steps of a 1D signal for level of 2; (a) decomposition, (b) reconstruction.

followed by down-sampling by a factor of 2 at each scale level. The entire process is referred to as sub-band coding. The resultant tree structure is presented in **Figure 1**, where, $\downarrow 2$ and $\uparrow 2$ represents the processes of down-sampling and up-sampling respectively. The DWT decomposition process can be applied on both sub part of the signal, approximation coefficients and detail coefficients. This kind of decomposition is referred to as wavelet packet transform or wavelet packet tree decomposition. **Figure 2** represents the wavelet packet decomposition and reconstruction process.

3.3 Wavelet packet transform

Wavelet Packet Transform (WPT) is another powerful denoising tool. WPT is a generalized form of DWT, in which both smooth and details parts are subject to further transforms. A full transformed matrix contains $j(= \log_2 N)$ transform levels for searching for the best basis. The best basis can be chosen using different criteria. Shannon entropy is a very common one, which is defined as,

$$S = -\sum_{j} p_j \log\left(p_j\right) \qquad (6)$$

for which $p_j = |x_j|^2 / \|x\|^2$ and $p \log p = 0$ for $p = 0$. The optimized basis function will be a combination of both approximated and detailed coefficients and minimum entropy which can be obtained by comparing all the possible combinations of wavelet coefficients at different levels, minimizing $\sum \log |x_j|$, numbers larger than t and Stein's unbiased estimate of risk (SURE) [14].

Wavelet packet transform (WPT) has several advantages over WT (continuous and discrete) as it sets no requirements of mother wavelet windowing function [15], wavelet packet basis function [16], and selection of the number of decomposition levels [17] and threshold [18]. WPT is introduced in [19] for denoising and harmonic detection by computing the difference between the noise and the desired signal. The effectiveness is also experimentally verified in [20] and tested against dynamics of Electro-encephalogram (EEG) and Electro-cardiogram (ECG) measurements in [21]. Image denoising is implemented by using an adaptive anisotropic dual-tree complex WPT on a bivariate stochastic signal model in [21].

DWT has become a powerful tool for denoising experimental data over the past few years. Original data is decomposed into a series of wavelets at different scales and intensities. Using WT, where the signal is multiplied by a transformation matrix; the detailed and the smooth parts are separated and the process is repeated over $\log_2 N$ iterations. Depending on the length of the filtering steps, we can have different types of wavelets. If the number of steps vary from 4 to 20, the wavelets are referred to as Daublets. The Haar transform is a special case of Daublet 2. There can also be multiple filters, each with different filter lengths. If there are 5 filters, the wavelets are known as Coiflets, where each filter length is a multiple of 6. If there are 7 filters, the wavelets are known as Symmlets, where each filter length is a multiple of 2.

4. DWT for denoising data

The DWT denoising procedure consists of three steps. In the first step, if the length of the data stream is of length of the order of power of two, it is transformed to the wavelet domain. In the second step, coefficients with either zero magnitude or criterion-based minimized values are selected. In the third or final step, the minimized coefficients are reverted back to the original domain from the wavelet domain to extract the denoised data. DWT-based denoising techniques can be broadly classified into two categories - linear and non-linear. In linear DWT, signal and noise are assumed to be belonging to the smooth and the detailed part of the wavelet domain, where high frequency components are attenuated. While in non-linear DWT, the filter removes the coefficients selected in the second stage with amplitudes less than the threshold. In practicality, non-linear DWT is always preferred over linear DWT, as linear DWT introduces error due to the retention of noise components and loss of signal components owing to wavelet filtering.

Whether linear or non-linear DWT denoising technique is used, performance depends on the choice of the wavelet family and the length of the filter. The traditional way for making this choice is based on visual inspection of the data, for example, daublets are implemented when the data appears smooth in the wavelet domain, while Haar or other wavelets are used when the data appears bursty and discontinuous in the wavelet domain. In order to overcome the problems with DWT denoising, correlation denoising method was introduced in [11]. Correlation denoising method implements wavelet transformation and filtering in a way such

that the correlation between wavelet coefficients of the signal part and the noise part is different at each level. However, correlation denoising in its original form is computationally complex. In order to reduce computational complexity, wavelet threshold denoising method was proposed by [12]. The method is simple to calculate and the noise can be suppressed to a large extent. At the same time, singular information of the original signal can be preferred well, so it is a simple and effective method. A brief overview of what happens when DWT is applied for denoising is demonstrated in **Figure 3**.

The four major components of the DWT denoising technique are: wavelet-type selection, threshold selection, threshold function selection and threshold application to the wavelet coefficients.

1. **Wavelet Selection** - There is a wide variety of wavelets that can be used for denoising. Selecting the optimum one depends on the selection of the matching wavelet filter. Out of different wavelet transform based denoising methods, only minimum description length (MDL) method has the flexibility of choosing the filter type.

2. **Threshold Selection** - There are four basic types of threshold selection, minimax, Stein's unbiased estimate of risk (SURE), and minimum description length (MDL). The *Universal* threshold is computed using,

$$t = \sigma \times \sqrt{2 \times \ln(N)} \qquad (7)$$

for which N is the length of the signal data array, and σ is the standard deviation of noise. In practicality, in most cases, σ is unknown, but can be estimated using the first detailed part of the wavelet coefficient x_i through the expression,

$$\sigma_{\text{estimate}} \approx \frac{\text{median}(|x_i|)}{0.6745}. \qquad (8)$$

In the case of *Minimax* criterion using the estimates of the minimax risk bounds for the transformed wavelets, a table is generated for threshold values corresponding to each set of given data lengths. These threshold values are always smaller than the universal threshold. The noise level estimates are calculated using (8) and signal components are retained along with a few number of noise components.
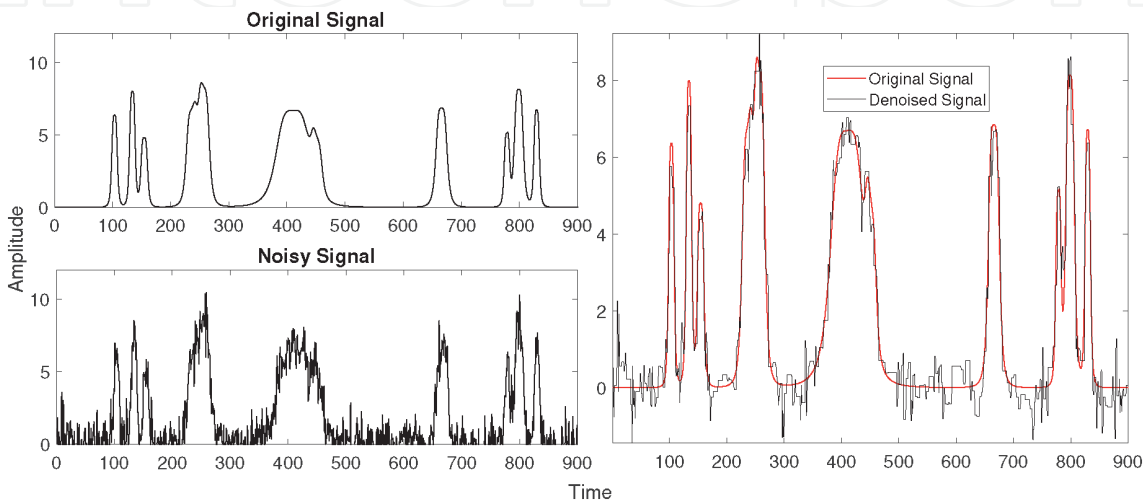

Figure 3.
Denoising with DWT.

Stein's unbiased estimate of risk (SURE) is used to obtain an unbiased estimate of the variance between the filtered and unfiltered data. SURE is defined as

$$\text{SURE}(t, x) = N - 2 \times M_{i:|x_i| \leq t} + \sum_{i=1}^{N} \left(|x_i|^t \right)^2 \qquad (9)$$

for which t, x_i, N and M refer to the candidate threshold, wavelet coefficient, data length and number of data points less than t. The value of t that minimizes the SURE value is selected as the threshold value while the final term of the SURE function represents the residual energy left after thresholding. The SURE threshold can be modified to yield global thresholds rather than local ones by combining SURE method with cycle-spinning technique; a method referred to as SPINSURE.

The *Minimum description length (MDL)* method for threshold computation can be expressed as,

$$\text{MDL}(k*, m*) = \min\left(\frac{3}{2} k \log(N) + \frac{N}{2} \log\left(\sum (x_m^2 - x_{mk}^2) \right) \right) \qquad (10)$$

for which k, m, x_m, and x_{mk} represent the number of largest coefficients retained after filtering, the filter type, the wavelet coefficients from m-type wavelet transform, and the k largest coefficients in amplitude respectively. Here k^* and m^* are the optimized values for the MDL criterion for threshold selection, where k^* is selected as the threshold for the corresponding wavelet coefficient. The $3/2k \log(N)$ term represents the penalty function with value proportional to the number of retained wavelet coefficients. The $\frac{N}{2}\left(\sum(x_m^2 - x_{mk}^2)\right)$ characterizes the error between the reconstructed and the original signal components.

3. **Selecting threshold function** - whether wavelet threshold denoising method is good or bad depends on two decisive factors; one is the threshold λ and the other important factor is the selection of the threshold function. The most basic threshold functions are the hard and soft threshold functions, comparative performance of which is presented in **Figure 4**.

The *Hard Threshold Function (HTF)* nullifies the decomposition coefficients to zero if they are less than the threshold and retains the coefficients if they are more than the threshold [22]. The HTF preserves the local properties of a signal with a few discontinuities introduced by the variations in the reconstructed signals. HTF can be expressed as,

$$\overline{\omega_{j,k}} = \begin{cases} \omega_{j,k}, & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \qquad (11)$$

The *Soft Threshold Function (STF)* [23] selects the threshold value such that all decomposition coefficients are nullified to zero. A major drawback with this technique is that a part of the high frequency components is lost owing to their location above threshold. STF can be mathematically expressed as,

$$\overline{\omega_{j,k}} = \begin{cases} \text{sgn}\left(\omega_{j,k}\right)\left(|\omega_{j,k}| - \lambda\right), & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \qquad (12)$$
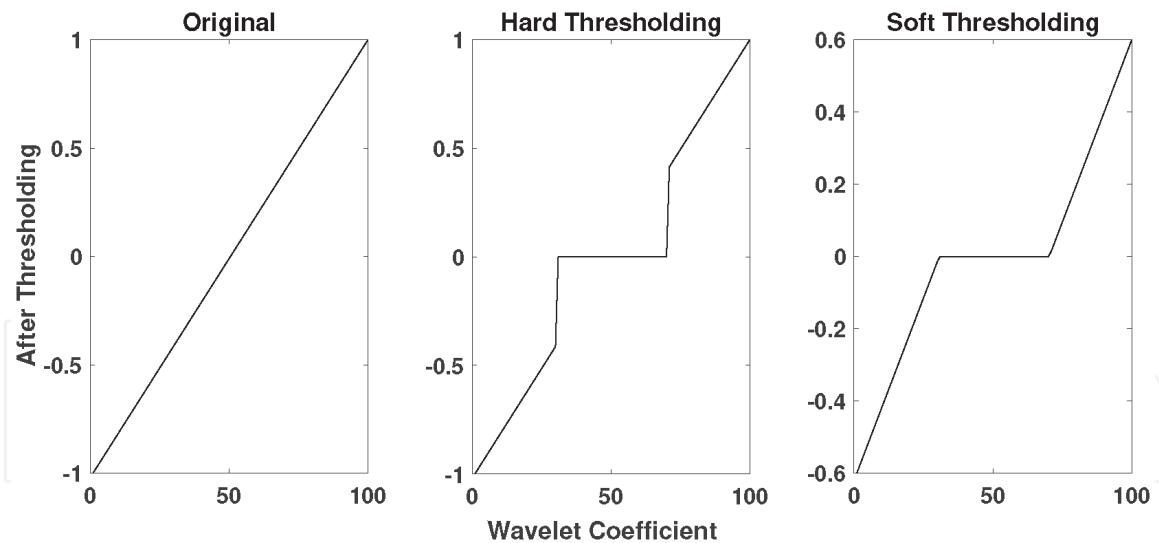
Figure 4.
Comparative hard and soft thresholding when implemented for DWT.

where $\overline{\omega_{j,k}}$, $\omega_{j,k}$, λ, and sgn() denotes the estimated wavelet coefficients, post-decomposition wavelet coefficients, threshold and symbolic piece-wise function respectively [24].

Garrote Threshold Function is proposed in [25] to improve the drawbacks of HTF and STF, whose denoising effect is better than the above two methods with respect to continuity of expressions,

$$\overline{\omega_{j,k}} = \begin{cases} \omega_{j,k} - \dfrac{\lambda^2}{\omega_{j,k}}, & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \qquad (13)$$

The continuity in the soft threshold function is much better, but it has a constant deviation. So, in order to overcome its shortcomings, the soft and hard threshold algorithms are compromised process by the literature; the ***semisoft threshold function*** [26].

$$\overline{\omega_{j,k}} = \begin{cases} \mathrm{sgn}\left(\omega_{j,k}\right)\left(|\omega_{j,k}| - T\lambda\right), & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \qquad (14)$$

It is worth-mentioning here, that the values of the threshold T is fixed with values between 0 and 1 in the case of HTF, STF, Garrote Threshold Function and semi-threshold function.

Another variation is the ***Improved Threshold Function*** which can be given by,

$$\overline{\omega_{j,k}} = \begin{cases} \mathrm{sgn}\left(\omega_{j,k}\right)\left(|\omega_{j,k}| - \dfrac{\lambda}{\exp^{-3}\left[\alpha(|\omega_{j,k}| - \lambda)/\lambda\right]}\right), & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \qquad (15)$$

The adjustment factor of the new function is different from the semisoft threshold function. It consists of a complex exponential function $\exp^{-3}\left[\alpha(|\omega_{j,k}| - \lambda)/\lambda\right]$ which has more adaptability; α is the normal number which can be adjusted freely and the values of α are different with the different signal. When $|\omega_{j,k}| = \lambda$, $\overline{\omega_{j,k}} \to \lambda$, $\overline{\omega_{j,k}} \to 0$. Therefore, continuously in place of

λ, the improved threshold function has the characteristics of soft threshold function; when $\overline{\omega_{j,k}} \to \infty$, $\overline{\omega_{j,k}} \to \omega_{j,k}$ improved threshold function based on $\overline{\omega_{j,k}} = \omega_{j,k}$ as the asymptotic line; it can be seen that, with the increase of $\omega_{j,k}$, $\omega_{j,k}$ will gradually be close to $\omega_{j,k}$; when $\omega_{j,k}$ becomes infinite, $\omega_{j,k} \approx \omega_{j,k}$. The choice of α is crucial for the success of the technique and the variation in α affects the denoising effect. When $\alpha = 0$, improved threshold function reduces to STF and when $\alpha = \infty$, improved threshold function reduces to HTF.

4. ***Thresholding or threshold application*** - thresholding is defined as the ways in which threshold is applied for modifying wavelet coefficients. DWT is a multi-level wavelet transform technique with different thresholds being applied at different level of coefficients

   Global Thresholding - This technique assumes the corrupting noise as Gaussian distributed with amplitude and frequency distributions same for all orthogonal bases for the entire data space. Global thresholding can be implemented using either hard, soft, Garrote or firm-threshold functions, expressed as,

   • Hard:

   $$x_i^* = \begin{pmatrix} 0, & \text{if } |x_i| \leq t \\ x_i, & \text{if } |x_i| > t \end{pmatrix} \tag{16}$$

   • Soft:

   $$x_i^* = \begin{cases} 0, & \text{if } |x_i| \leq t \\ \text{sign}(x_i)(|x_i| - t), & \text{if } |x_i| > t \end{cases} \tag{17}$$

   • Garrote:

   $$x_i^* = \begin{cases} 0, & \text{if } |x_i| \leq t \\ x_i - t^2/x_i, & \text{if } |x_i| > t \end{cases} \tag{18}$$

   • Firm:

   $$x_i^* = \begin{cases} 0, & \text{if } |x_i| \leq t_1 \\ \text{sign}(x_i)t_2(|x_i| - t_1)/(t_2 - t_1), & \text{if } t_1 < |x_i| \leq t_2 \\ x_i & \text{if } |x_i| > t_2 \end{cases} \tag{19}$$

   for which x_i and x_i^* represents the wavelet coefficients pre- and post-thresholding respectively. HTF partitions the wavelet coefficients into two parts by the selected threshold eliminating coefficients with low magnitude. STF reduces all coefficients by a factor equal to the threshold eliminating smaller coefficients. Similarly, Garrote thresholding reduces all large coefficients by a factor of a non-linear continuous function. Firm thresholding reduces only the middle coefficients while eliminating small and retaining large coefficients.

   Level-Dependent Thresholding - This technique uses different thresholds at each level of wavelet transformations. It uses a combination of SURE and global thresholding techniques to initiate a hybrid method. In this case, if the sample variance at each level is sparse, global thresholding is applied, while SURE thresholding is applied otherwise.

Data-Dependent Thresholding - A Data-dependent threshold (DDT) technique selects a threshold such that empirical wavelet coefficients are shrunk. The thresholding is achieved through statistical tests of hypotheses like linear regression. The level of this statistical test is adjusted to control the smoothness of the resulting estimator such that a good mean-squared error (MSE) performance is achieved for different data analysis settings with smoothness in estimator response. The main aim of this technique is to eliminate a group of wavelet coefficients that exhibit characteristics of pure noise.

Cycle-Spin Thresholding - It combines the process of subspace identification, projecting denoising and averaging of the projections. The subspace mentioned here refers to the region where most of the energy of the signal is concentrated and signal corrupted with noise is projected on to this subspace.

5. Signal denoising for IoT networks

The huge amount of sensor data generated in an IoT network are used to take decisions on a certain observation/ phenomenon based on real-time processing. The decision-making procedure often involves detecting the signal energy level transmitted from the sensors. If the received energy level is higher than a predefined threshold, the target is detected to be present phenomenon and vice-versa. However, the sensor data gets crippled with noise contributed by the wireless environment and the internal electronics of the sensors, on its way to the data center for processing. The WPT method will be the best option in this case for denoising the sensor data, where the original signal coefficients are preserved while removing the noise within the signal. The WPT method can decompose a signal in both scale and wavelet space thereby revealing more details about both the sensor signals and the crippling noise. If energy correlation analysis is used in conjunction with WPT, signal energy from the sensor data can be analyzed and noise can be eliminated by zooming into the signal characteristics at different time scales. Advantages of WPT over WT is evident in **Figure 5**. Hence, in this section, a universal framework is presented for denoising sensor signals in IoT networks. The framework is based on energy correlation analysis and combines the processes of WP decomposition, coefficient modification and WP reconstruction. The functional block diagram for this framework is presented in **Figure 6**.
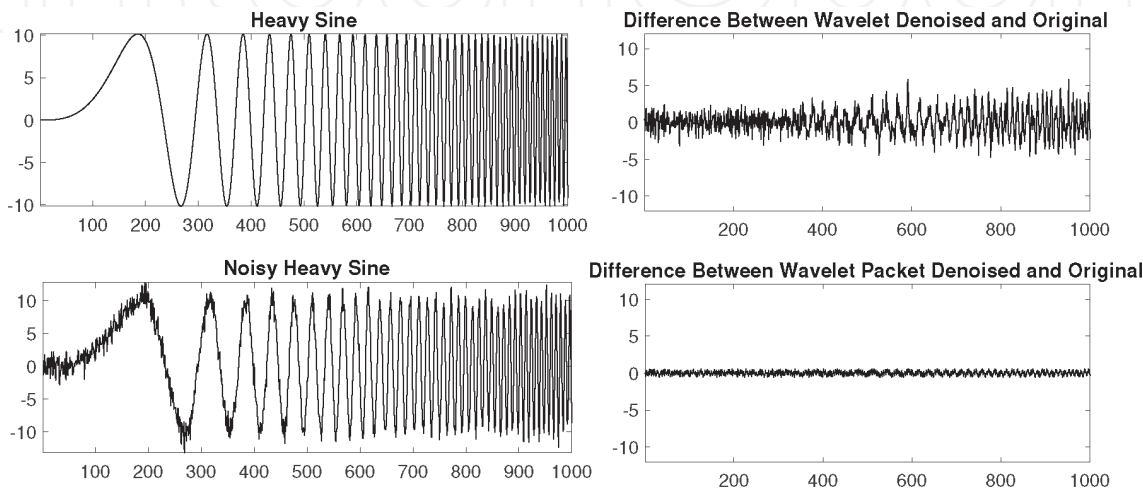

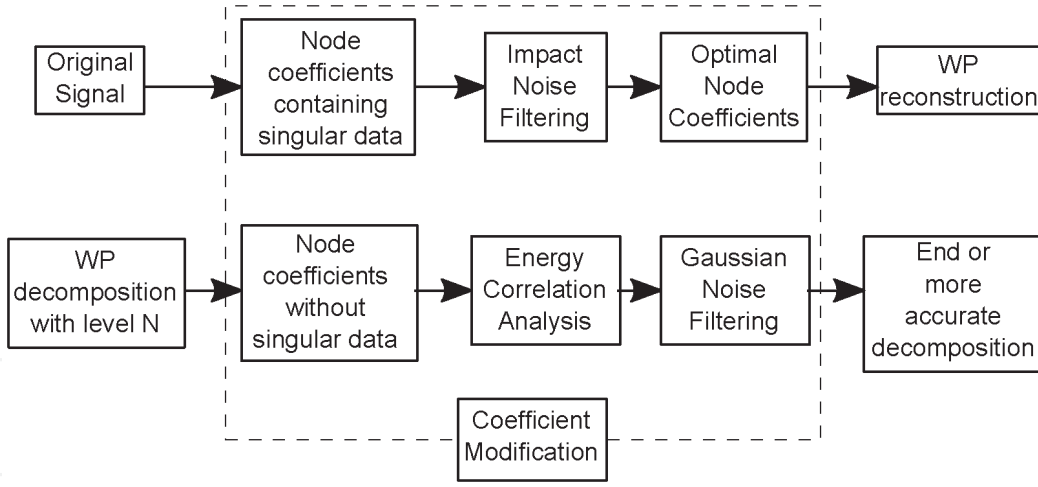**Figure 5.**
Comparative performance of WPT and WT.

Figure 6.
Architecture of the universal framework.

5.1 Wavelet packet transfer for IoT

In WPT for IoT networks, for a given for a given orthonormal scaling function $\phi(t)$ and wavelet function $\psi(t)$ the double scale Eq. [14] can be described as follows:

$$\phi(t) = \sqrt{2}\sum_{k}h_{0k}\phi(2t - k), \psi(t) = \sqrt{2}\sum_{k}h_{1k}\phi(2t - k) \qquad (20)$$

where h_{0k} and h_{1k} are a pair of conjugate orthogonal filter coefficients. WP functions for $n = 0, 1, \ldots$ can be defined as follows,

$$w_{2n}(t) = \sqrt{2}\sum_{k\in Z}h_{0k}w_n(2t - k), w_{2n+1}(t) = \sqrt{2}\sum_{k\in Z}h_{1k}w_n(2t - k) \qquad (21)$$

When $n = 0, w_0(t) = \phi(t), w_1(t) = \psi(t)$. $\{w_n(t)\}_{n\in Z}$ represents the wavelet packet assuming standard orthogonal wavelet basis can be constructed from the scaling function. Scaling and wavelet functions generated as a result of this process satisfy the properties of orthogonality over both scale and translation,

$$\langle w_n(t - k) \cdot w_n(t - l)\rangle = \delta_{kl} \quad k, l \in Z$$
$$\langle w_{2n}(t - k) \cdot w_{2n+1}(t - l)\rangle = 0 \quad n = 1, 2, \ldots \qquad (22)$$

In the process of WP decomposition, scale space $\{V_j\}_{j\in Z}$ composed of scaling functions and wavelet space $\{W_j\}_{j\in Z}$ composed of wavelet functions can be expressed in a unified way as follow:

$$U_j^0 = V_j, U_j^1 = W_j \quad j \in Z \qquad (23)$$

From $V_j = V_{j+1} \oplus W_{j+1}$, then

$$U_j^0 = U_{j+1}^0 \oplus U_{j+1}^1, U_j^n = U_{j+1}^{2n} \oplus U_{j+1}^{2n+1} \quad j \in Z, n \in Z^+ \qquad (24)$$

where, U_j^n denotes the closed subspace of square and integrable space $L^2(R)$ generated by the linear combination of wavelet packet w_n after translation and scaling operation.

During the procedure of multi-resolution analysis, objective function is decomposed into the subspace $\{V_j\}_{j \in Z}$, $\{W_j\}_{j \in Z}$ in $L^2(R)$ carried out further decomposition according to binary mode as follows:

$$W_j = U_j^1 = U_{j+1}^2 \oplus U_{j+1}^3, U_{j+1}^2 = U_{j+2}^4 \oplus U_{j+2}^5, U_{j+1}^3 = U_{j+2}^6 \oplus U_{j+2}^7$$

$$W_j = U_{j+2}^4 \oplus U_{j+2}^5 \oplus U_{j+2}^6 \oplus U_{j+2}^7 \tag{25}$$

Consequently,

$$W_j = U_{j+k}^{2^k+1} \oplus U_{j+k}^{2^k+2} \oplus \dots \oplus U_{j+k}^{2^{k+1}-1} \tag{26}$$

Finally, the wavelet packet coefficients can be computed [27] as follows:

$$d_k^{j+1,2n} = \sum_l h_{0(2l-k)} d_l^{j,n}, d_k^{j+1,2n+1} = \sum_l h_{1(2l-k)} d_l^{j,n} \tag{27}$$

where

$$d_k^{j+1,n} = \sum_k \left[h_{0(2l-k)} d_k^{j,2n} + h_{1(2l-k)} d_k^{j,2n+1} \right] \tag{28}$$

Following this technique of WPT, the efficiency of the denoising process improves quite a bit over the case where just WT is used for denoising the signals, as is evident in **Figure 5**.

5.2 Energy correlation analysis

Digital signal energy computation is achieved by extracting and squaring signal amplitude at different locations in the time domain and then adding them together [28]. The influence of relative large energy is eliminated using normalization technique [29]. This normalization can be avoided by selecting the sum of absolute values of amplitudes at each sampling points as approximations for evaluating energy; the mathematical formulation for which can be represented as:

$$e = \sum_{n=1}^{N} |f(n)|, \qquad n = 1, 2, \dots, N. \tag{29}$$

Any kind of non-deterministic relationship existing between two or more variables can be exploited and formalized using correlation analysis. Thus, different kinds of signals can be differentiated by exploring the internal relation with correlation analysis. x_i and y_i denote two random variables, respectively; the calculation formula of correlation coefficient can be given as follows:

$$r = S_{xy} / \sqrt{S_{xx} S_{xy}}, \qquad -1 \le r \le 1, \tag{30}$$

where $S_{xx} = \sum_{i=1}^{N} (x_i - \bar{x})^2$, $S_{yy} = \sum_{i=1}^{N} (y_i - \bar{y})^2$ and $S_{xy} = \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$.

The correlation coefficient r is referred to as "Pearson product-moment correlation coefficient," or Pearson's r and is used to estimate the relative relationship between variables using the following principles.

1. The closer the absolute value of Pearson's r to 1, more is the correlation and closer is the Pearson's r to 0, less is the correlation between the variables.

2. The polarity of the coefficient determines the direction of correlation, with plus-sign representing positive and minus-sign representing negative correlation.
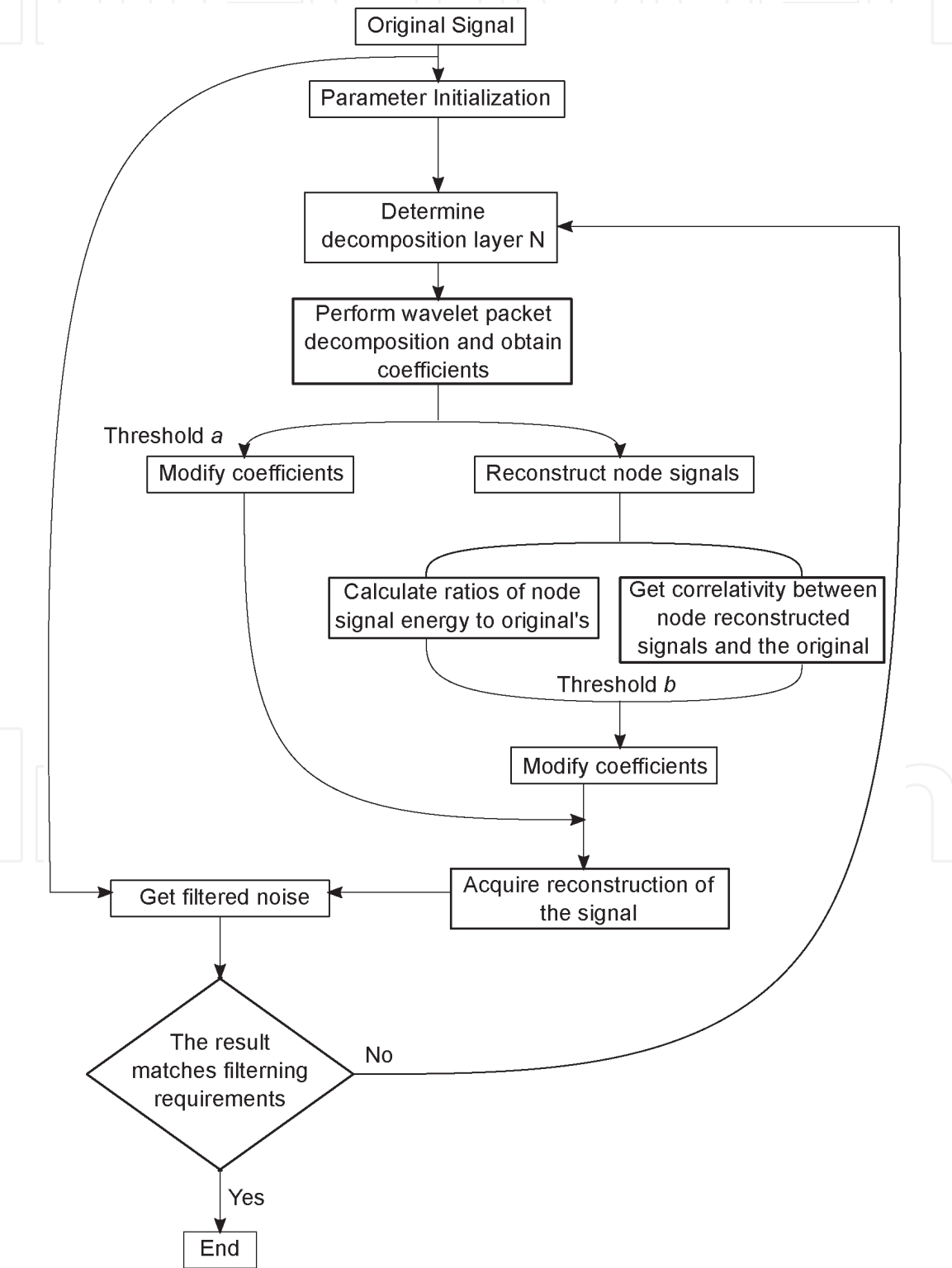

Figure 7.
Flowchart of wavelet packet coefficients based on energy-correlation analysis.

5.3 Processing method for WP coefficients based on energy-correlation analysis

An online filtering process capable of denoising both Gaussian and impact noise is presented below based on the energy correlation between signal components reconstructed from WP coefficients.

Step 1 - Obtain WP decomposition coefficients through the application of appropriate decomposition level and mother wavelet.

Step 2 - Compare WP coefficients in each subspace to eliminate singular data based on a pre-selected threshold through the application of multi-resolution analysis.

Step 3 - After reconstructing WP node signals from real coefficients, compute the ratios of the energy of the reconstructed signal components to the actual signal components to obtain the correlation between them. Subspace unsatisfied coefficients are processed through the use of a different threshold resulting in a series of new coefficients.

Step 4 - Using the new set of modified coefficients on each node, signal components are reconstructed and noise is eliminated. If the filtering requirements are not satisfied, repeat steps to step 4 after increasing the decomposition level. A flow-diagram for energy correlation analysis based WP coefficient processing is depicted in **Figure 7**.

6. Performance analysis of denoising techniques

The best way to denoise a signal is to assume that the noise signal is Gaussian distributed with values that are independent and identical real values. The performance of the denoising process can be evaluated by comparing the quality of the denoised signal with that of the original transmit signal. A variety of methods have been proposed over years to measure the performance of denoising; the most common of which are the metrics of SNR and the peak SNR (PSNR), generally accepted to measure the quality of signal and images respectively. For 1-D signal, measuring the performance of the denoising method by calculating the residual SNR is given by, $SNR = 10 \log_{10} \left(\sum_{n=0}^{N-1} x^2(n) / \sum_{n=0}^{N-1} (\overline{x}(n) - x^r(n))^2 \right)$ where $x(n)$ is the original signal, $x^r(n)$ is the denoised signal and $\overline{x}(n)$ refers to the mean value of $x(n)$.

In order to measure the quality of image, PSNR is generally used, which is given by $PSNR = 10 \log_{10} \left(L / \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} (\overline{x}(n,m) - x^r(n,m))^2 \right)$, where L, $x(n)$, $\overline{x}(n,m)$ and $x^r(n,m)$ refer to the quantized gray level of images, original image, mean value of $x(n)$ and the reconstructed image respectively. However, the choice of the noise power is absolutely crucial for visible performance difference. SNR is more important as compared to noise power when evaluating performance and with SNR above 3 dB, it is quite easy to isolate visible corruption.

7. Conclusions

Decomposition in time and frequency domain for Fourier Transform is replaced by decomposition in space domain for WT thereby removing any ambiguity related to time and frequency and offering high flexibility and quality to the overall denoising process. Different threshold estimation methods, wavelet types, threshold types and thresholding functions can be used for implementing WT depending on the application scenario, network architecture, the kind of signal transmitted

and the kind of noise commonly observed in the considered application scenario. However, comparing performances of different thresholding methods, wavelet types or threshold types when applied for the WT reveal that the number of decomposition levels are more crucial to the denoising performance than the types of wavelets or thresholds.

If the application scenario is considered to be an industrial IoT network, WPT method is preferred over simple WT for denoising sensor signals. This is because in WPT, signal is decomposed into an approximation and a detail component at each layer of each decomposition level, therefore resulting in 2^n number of components at n decomposition levels in contrast to just 2 components at each of the n decomposition levels of WT. Moreover, WT decomposes only the low frequency components in contrast to WPT which considers both low and high frequency components at each level. If WPT is combined with energy correlation analysis, effectiveness of the denoising process increases manifold owing to its immunity to diversity of signals in an IoT network. Integration of energy and correlation can be used to modify wavelet packet coefficients for eliminating Gaussian and impact noise efficiently.

A. Appendix A

....
Example MATLAB codeset for signal denoising
....
—————————————————-
Signal Generation
—————————————————-

```
N = 2048*2;
name = 'piece-regular';
f0 = load_signal(name, N);
f0 = rescale(f0,.05,.95);
sigma = 0.05;
f = f0 + randn(size(f0))*sigma;
figure(1)
subplot(2,1,1); plot(f0); axis([1 N 0 1]);
title('Clean signal');
subplot(2,1,2);
plot(f); axis([1 N 0 1]);
title('Noisy signal');
```
—————————————————-
Thresholding
—————————————————-

```
Theta0 = @(x,T)x.* (abs(x)¿T);
Theta1 = @(x,T)max(0, 1-T./max(abs(x),1e-9)).* x;
t = linspace(-3,3,1024)'; T = 1;
figure(2)
plot( t, [Theta0(t,T), Theta1(t,T)], 'LineWidth', 2 );
axis('equal'); axis('tight');
legend('Θ⁰', 'Θ¹');
```
—————————————————-
Wavelet-Threholding
—————————————————-

```
options.ti = 0; Jmin = 4;
```

```
W = @(f) perform_wavelet_transf(f,Jmin,+1,options);
Wi = @(fw)perform_wavelet_transf(fw,Jmin,-1,options);
x = W(f);
x1 = Theta0(x, 3*sigma);
figure(3)
subplot(2,1,1);
plot_wavelet(x,Jmin); axis([1N -11]);
title('W(f)');
subplot(2,1,2);
plot_wavelet(Theta0(W(f),T),Jmin); axis([1N -11]);
title('Θ⁰(W(f))');
f1 = Wi(x1);
figure(4)
subplot(2,1,1);
plot(f); axis([1 N 0 1]);
title('f');
subplot(2,1,2);
plot(f1); axis([1 N 0 1]);
title('f'₁);
x = W(f);
reinject = @(x1)assign(x1, 1:2^Jmin, x(1:2^Jmin));
Theta0W = @(f,T)Wi(Theta0(W(f),T));
Theta1W = @(f,T)Wi(reinject(Theta1(W(f),T)));
_____-

TI WT
_____-

options.ti = 1;
W = @(f) perform_wavelet_transf(f,Jmin,+1,options);
Wi = @(fw)perform_wavelet_transf(fw,Jmin,-1,options);
fw = W(f);
nJ = size(fw,3)-4;
figure(5)
subplot(5,1, 1);
plot(f0); axis('tight');
title('Signal');
i = 0;
for j=1:3
i = i+1;
subplot(5,1,i+1);
plot(fw(:,1,nJ-i+1)); axis('tight');
title(strcat(['Scale=' num2str(j)]));
end
subplot(5,1, 5);
plot(fw(:,1,1)); axis('tight');
title('Low scale');
```

Author details

Indrakshi Dey[1,2]* and Shama Siddiqui[3]

1 National University of Ireland, Maynooth, Ireland

2 Trinity College Dublin, University of Dublin, Ireland

3 DHA Suffa University, Karachi, Pakistan

*Address all correspondence to: deyi@tcd.ie; indrakshi.dey@mu.ie

IntechOpen

References

[1] A. A. Brincat, F. Pacifici and F. Mazzola. IoT as a Service for Smart Cities and Nations. IEEE Internet of Things Magazine, 2(1), pp. 28–31, 2019.

[2] Z.-C. Liu, X.-G. Chen, and Y.-F. Li. Detection and identification of abrupt changes for on-line sensor output signal. Transaction of Beijing Institute of Technology, 26(12), pp. 1104–1108, 2006.

[3] A.M. Rao and D. L. Jones. A denoising approach to multisensor signal estimation. IEEE Transactions on Signal Processing, 48(5), pp. 1225–1234, 2000.

[4] B. L. Jin, H. Li, N. J. Zhao et al. A new denoising algorithm for wavelet thresholding. Journal of Missile and Missile, 31 (1), pp. 167–169, 2011 (Chinese).

[5] D. L. Donoho. De-noising by soft-thresholding. IEEE Transactions on Information Theory, 41(3), pp. 613–627, 1995.

[6] M. Ding and H. Zhu. Two-Dimensional gibbs phenomenon for fractional fourier series and its resolution. Artificial Intelligence and Computational Intelligence, vol. 7530 of Lecture Notes in Computer Science, pp. 530–538, Springer, Berlin, Germany, 2012.

[7] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 36(5), pp. 961–1005, 1990.

[8] A. S. Lewis and G. Knowles. Image compression using the 2-D wavelet transform. IEEE Transactions on Image Processing, 1(2), pp. 244–250, 1992.

[9] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. IEEE Transactions on signal processing, 41(12), pp. 3397–3415, 1993.

[10] X. He and M. S. Scordilis. Psychoacoustic music analysis based on the discrete wavelet packet transform. Research Letters in Signal Processing 2008, pp. 1–5.

[11] D. L. Donoho and J. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81(3), pp. 425, 1994.

[12] D. L. Donoho. Denoising by soft-thresholding. IEEE Trans. Inform. Theory, 41(3), pp. 613–627, 1995.

[13] S. G. Mallat. A wavelet tour of signal processing. Academic Pr., 1999.

[14] J. Gubbi, A. Khandoker, and M. Palaniswami. Classification of sleep apnea types using wavelet packet analysis of short-term ECGsignals. Journal of Clinical Monitoring and Computing, 26(1), pp. 1–11, 2012.

[15] F. Adamo, G. Andria, F. Attivissimo, A. M. L. Lanzolla, and M. Spadavecchia. A comparative study on mother wavelet selection in ultrasound image denoising. Measurement, 46(8), pp. 2447–2456, 2013.

[16] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. IEEE Transactions on Information Theory, 38 (2), pp. 713–718, 1992.

[17] A. M. Hasan, K. Samsudm, A. R. Ramli, and R. S. Azmir. Wavelet-based pre-filtering for low cost inertial sensors. Journal of Applied Sciences, 10 (19), pp. 2217–2230, 2010.

[18] D. L. Donoho. De-noising by soft-thresholding. IEEE Transactions on Information Theory, 41(3), pp. 613–627, 1995.

[19] P. Mercorelli. Denoising and harmonic detection using non-orthogonal wavelet packets in industrial applications. Journal of Systems Science Complexity, 20(3), pp. 325–343, 2007.

[20] Y. Li, T. Zhang, L. Deng, B. Wang and M.Nakamura. Denoising and rhythms extraction of EEG under +Gz acceleration based on wavelet packet transform. Proceedings of the 7th ICME International Conference on Complex Medical Engineering (CME '13), pp. 642–647, Beijing, China, May 2013.

[21] J. Yang, W. Xu, Y.Wang, and Q. Dai. 2-D anisotropic dual-tree complex wavelet packets and its application to image denoising. Proceedings of the 15th IEEE International Conference on Image Processing (ICIP '08), pp. 2328–2331, October 2008.

[22] C. He, J. C. Xing, and Q. L. Yang. Optimal wavelet basis selection for wavelet denoising of structural vibration signal. Applied Mechanics and Materials, 578-579, pp. 1059–1063, 2014.

[23] J.-Y. Tang, W.-T. Chen, S.-Y. Chen, and W. Zhou. Wavelet-based vibration signal denoising with a new adaptive thresholding function. Journal of Vibration and Shock, 28(7), pp. 118–121, 2009 (Chinese).

[24] S. Badiezadegan and R. C. Rose. A wavelet-based thresholding approach to reconstructing unreliable spectrogram components. Speech Communication, 67, pp. 129–142, 2015.

[25] 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, 12(3), pp. 1257–1265, 2015.

[26] K. L. Yuan. Wavelet denoising based on threshold optimization method. Engineering Journal of Wuhan University, 48(1), pp. 74–80, 2015.

[27] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Transactions on Image Processing, 12 (11), pp. 1338–1351, 2003.

[28] K. Zhang, B.-J. Pang, and M. Lin. Wavelet packet analysis for acoustic emission signals caused by debris cloud impact. Journal of Vibration and Shock, 31(12), pp. 125–128, 2012.

[29] X.-H. Gu, G.-X. Zhang, D.-B. Hou, and Z.-K. Zhou. Detection of water pipe leak location using wavelet packet decomposition and power feature extraction. Journal of Sichuan University, 37(6), pp. 145–149, 2005.