

An Efficient Feature Extraction Method Based on Entropy for Power Quality Disturbance

¹P. Kailasapathi and ²D. Sivakumar

¹Department of Electrical Engineering, ²Department of Instrumentation Engineering,
Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India

Abstract: Feature extraction of a disturbed power signal provides information that helps to detect the responsible fault for power quality disturbance. A precise and faster feature extraction tool helps power engineers to monitor and maintain power disturbances more efficiently. Firstly, the decomposition coefficients are obtained by applying 10 level wavelet multi resolution analysis to the signals (normal, sag, swell, outage, harmonic and sag with harmonic and swell with harmonic) generated by using the parametric equations. Secondly, a combined feature vector is obtained from standard deviation of these features after distinctive features for each signal are extracted by applying the energy, the Shannon entropy and the log-energy entropy methods to decomposition coefficients. Finally, the entropy methods detect the different types of power quality disturbance.

Key words: Wavelet transform, harmonic, sag, multilevel, Shannon entropy and log-energy entropy

INTRODUCTION

The power quality problems have become more pronounced in recent days as power electronic devices and microprocessor based systems widely used in modern industrial control system. The proliferation of highly sensitive computerized equipments place more stringent demands on the quality of electric power supplied to the customer.

The quality of power supplies has major concern for electric utilities and end users due to the competitive electricity markets under deregulation. To study, the power quality problems, it is required to monitor the voltage and current waveforms and identify the different disturbances from these measurements. After detecting the disturbance, it is required to identify the source of disturbance. Power quality problems involve voltage sag, voltage swells, notch, spike, switching transient, impulses, flicker, harmonics, etc.

Voltage sag, swell is not a complete interruption of power. It is a partial change in waveform structure. When the normal voltage signal increases by 10-90%, it is known as voltage swell. Most voltage sags do not go below 50% of the nominal voltage and they normally last from 3-10 cycles or 50-170 msec. Voltage sag is a problem that occurs due to a fault, switching of heavy load or starting of large motors.

There may be sudden changes in voltage signal due to internal switching or sudden switch off inductive loads or capacitive banks or similar phenomena which persists

for a short duration known as transient. Chirp occurs when the signal varies with rapid change in frequency for a short duration.

WAVELET TRANSFORM

A wavelet is a wave-like oscillation with amplitude that starts out at zero, increases and then decreases back to zero. It can typically be visualized as a “brief oscillation” like one might see recorded by a seismograph or heart monitor. Wavelets can be combined, using a “shift, multiply and sum” technique called convolution with portions of an unknown signal to extract information from the unknown signal (Gaouda *et al.*, 2002).

A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as “daughter wavelets”) of a finite-length or fast-decaying oscillating waveform (known as the “mother wavelet”). Wavelet transforms have advantages over traditional fourier transforms for representing functions that have discontinuities and sharp peaks and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals. Wavelet transform is a transform which is capable of providing the time and frequency information simultaneously, hence giving a time-frequency representation of the signal. Wavelet transforms are classified into Discrete Wavelet Transforms (DWTs) and Continuous Wavelet Transforms (CWTs). Both DWT and CWT are continuous-time (analog) transforms. They can be used to represent continuous-time (analog) signals.

CWTs operate over every possible scale and translation whereas DWTs use a specific subset of scale and translation values or representation grid (Santoso *et al.*, 1996).

Continuous wavelet transform: Where $x(t)$ is the signal to be analyzed, $\psi(t)$ is the mother wavelet or the basis function. All the wavelet functions used in the transformation are derived from the mother wavelet through translation (shifting) and scaling (dilation or compression):

$$X_{wT}(\tau, s) = \frac{1}{\sqrt{|S|}} \int x(t) \cdot \psi \left(\frac{t - \tau}{s} \right) dt$$

The mother wavelet used to generate all the basic functions is designed based on some desired characteristics associated with that function. The translation parameter τ relates to the location of the wavelet function as it is shifted through the signal. Thus, it corresponds to the time information in the wavelet transform. The scale parameter s is defined as $|1/\text{frequency}|$ and corresponds to frequency information. Scaling either dilates (expands) or compresses a signal. Large scales (low frequencies) dilate the signal and provide detailed information hidden in the signal while small scales (high frequencies) compress the signal and provide global information about the signal. Notice that the wavelet transform merely performs the convolution operation of the signal and the basis function. The above analysis becomes very useful as in most practical applications, high frequencies (low scales) do not last for a long duration but instead, appear as short bursts while low frequencies (high scales) usually last for entire duration of the signal.

Discrete wavelet transform: The Discrete Wavelet Transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations obeying some defined rules. In other words, this transform decomposes the signal into mutually orthogonal set of wavelets which is the main difference from the Continuous Wavelet Transform (CWT) or its implementation for the discrete time series sometimes called Discrete-Time Continuous Wavelet Transform (DT-CWT) (Hamid and Kawasaki, 2002).

The wavelet can be constructed from a scaling function which describes its scaling properties (Sushama *et al.*, 2007). The restriction that the scaling functions must be orthogonal to its discrete translations implies some mathematical conditions on them which are mentioned everywhere, e.g., the dilation equation.

Figure 1 and 2 show the input signal and the corresponding Wavelet Transformed (WT) signal (Burrus *et al.*, 1998).

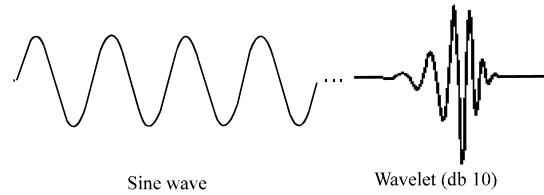


Fig. 1: Input signal

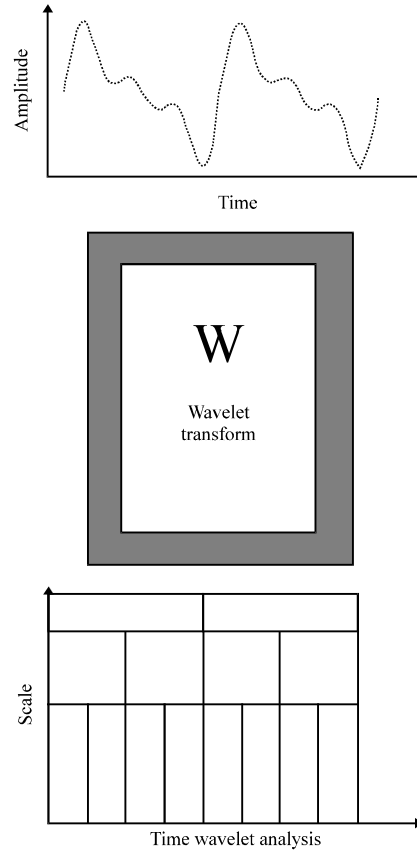


Fig. 2: WT signal

The frequency axis in plot is labeled as scale. It should be noted that the scale is inverse of frequency. That is, high scales correspond to low frequencies and low scales correspond to high frequencies. Consequently, the little peak in the plot corresponds to the high frequency components in the signal and the large peak corresponds to low frequency components (which appear before the high frequency components in time) in the signal (Fig. 3).

The parameter scale in the wavelet analysis is similar to the scale used in maps. As in the case of maps, high scales correspond to a non-detailed global view (of the signal) and low scales correspond to a detailed view. Similarly, in terms of frequency, low frequencies (high

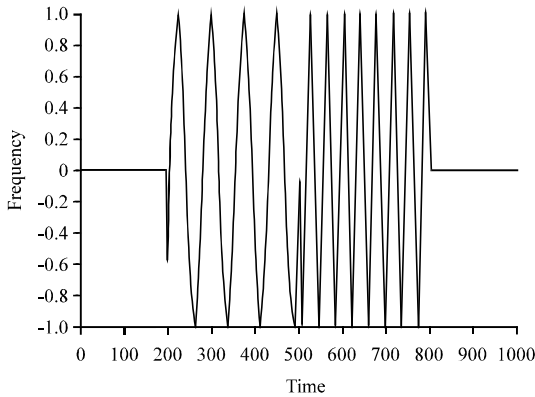


Fig. 3: Sinusoidal signal having two different frequency components at two different times

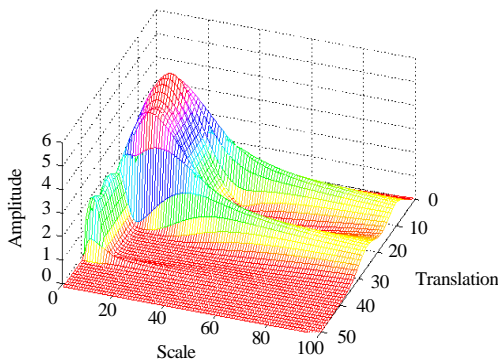


Fig. 4: Continuous wavelet transform of the sinusoidal signal

scales) correspond to a global information of a signal (that usually spans the entire signal) whereas high frequencies (low scales) correspond to a detailed information of a hidden pattern in the signal (that usually lasts a relatively short time) (Fig. 4).

It might be puzzled from the frequency resolution shown in the plot, since it shows good frequency resolution at high frequencies. It should be noted however that it is the good scale resolution that looks good at high frequencies (low scales) and good scale resolution means poor frequency resolution and vice versa.

THE PROPOSED ENTROPY-BASED FEATURES ANALYSIS APPROACH

Entropy-Based Features Method: Entropy as a thermodynamic state variable was introduced into physics by German physicist Rudolf Clausius in second half of 18th century. It was originally defined as:

$$dS = \frac{\delta Q}{T} \tag{1}$$

Where:

- dS = An elementary change of entropy
- Q = A reversibly received elementary heat
- T = An absolute temperature

Of course such a definition has no sense for signal processing. However, it started a diffusion of entropy as a term into the other areas. The entropy as a measure of system disorganization appeared for the first time in connection with the first postulate of thermodynamics: “Any macroscopic system which is in time t0 in given time-invariant outer conditions will reach after a relaxation time the so-called thermodynamic equilibrium. It is a state in which no macroscopic processes proceed and the state variables of the system gains constant time-invariant values”. The entropy of a system is maximal when the system has reached the thermodynamic equilibrium.

The above depicted key idea promoted entropy to a generic measure of system disorganization.

From the above analysis researchers can see, the characteristics of the original waveform can be reflected in different scales after the MRA decomposition. Based on this observation, researchers can construct the feature vector to detect different kinds of PQ disturbances. This idea is shown in Fig. 5. The sampled waveform was decomposed into different resolution levels (i) according to MRA. Then, the entropy of the detail information at each decomposition level (i) is calculated according to the following equation:

$$H(x) = - \sum_{i=1}^N p(x_i) \log_{10} p(x_i) \tag{2}$$

In Eq. 2 where $x = \{x_1, x_2, \dots, x_N\}$ is a set of wavelet (detail) coefficients in wavelet decomposition from level 1 to level l. N is the total number of the coefficients at each decomposition level and E(i) is the entropy of the detail at decomposition level (i). In order to identify different kinds of PQ disturbances, the Entropy Difference (ED) at each decomposition level is calculated which is the difference of the entropy E(i) with the corresponding entropy of the reference (normal) waveform at this level Eref(i):

$$ED(i) = E(i) - E(ref)(i) \tag{3}$$

By observing this ED(i) feature vector at different resolution levels and following the criterion proposed later in this study, one can effectively detect, localize and classify different kinds of PQ disturbances. This method

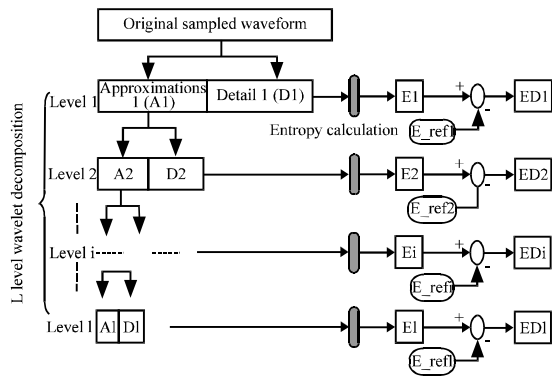


Fig. 5: Entropy-based features system architecture

is named as Entropy Difference of MRA Method. The major advantages of this method include two aspects. The first one is that by using this method, one can significantly reduce the dimensionality of the analyzed data. As researchers can see from Fig. 5, for a 1 levels multi-resolution decomposition, only an 1-dimensional feature vector need to be observed. This is a significant reduction compared to the original sampled waveform.

The second advantage is that this method keeps all the necessary characteristics of the original waveform for analysis. Different PQ characteristics are represented by the energy difference at different resolution scale which provides an effective way for different types of PQ detection.

Figure 6 shows a normal pure sine wave (60 Hz) and its four types of typical PQ disturbances: low frequency distortion, high frequency distortion, voltage sag and voltage swell. Sampling frequency used is 5 kHz. These PQ disturbance models are based on the IEEE Standard 1159-1995 (IEEE Recommended Practice for Monitoring Electric Power Quality) (Gaouda *et al.*, 1999) which are widely adopted in the academic and industry community (Hamid and Kawasaki, 2002; He and Starzyk, 2006). Specifically, for the short duration variations, the typical duration for voltage sag is from 0.5-30 cycles with the voltage magnitude between 0.1 and 0.9 p.u. while the typical duration for voltage swell is from 0.5-30 cycles with the voltage magnitude between 1.1 and 1.8 p.u. For the frequency distortions, the typical spectral content for low frequency distortion is <5 kHz with the voltage magnitude of 0-4 p.u. while for the high frequency distortions, the typical spectral content is between 0.5 and 5 MHz with the voltage magnitude of 0-4 p.u. A detailed discussion of these typical PQ characteristics can be found by Gaouda *et al.* (1999) and detailed mathematical models can also be found by He and Starzyk (2006). In the current study, researchers use the Daubechies4 (Db4) wavelets and 4 levels decomposition for analysis.

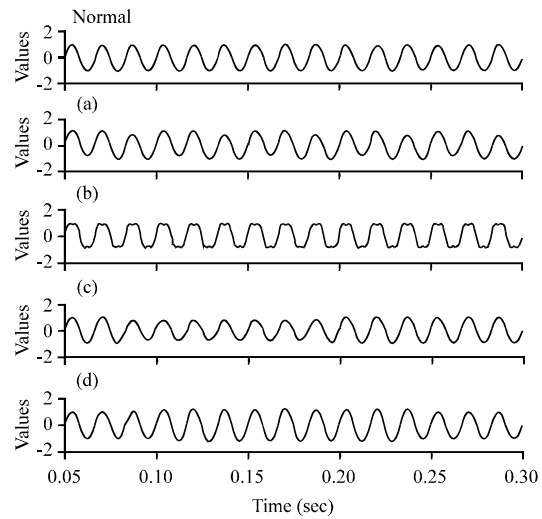


Fig. 6: Original normal waveform and its distortion signal; a) low frequency distortion; b) high frequency distortion; c) voltage sag from 0.08-0.196 sec and d) voltage swell. From 0.104-0.24 sec

Since, it is well known that wavelet transform can localize the time information for PQ disturbances, researchers will focus on the characteristics analysis and performance evaluation of the proposed method. Interested audiences can refer to study (IEEE Std. 1159, 1995; O'Shea, 2002) for the detection of the beginning and ending time of the power quality disturbance. Figure 7 give the entropy-based features analysis result to the signals in Fig. 6a-d where the horizontal axis represents the decomposition level (scale) and the vertical axis is the entropy difference as defined in Eq. 3. Based on the analysis result in Fig. 7, the following criteria are proposed for detecting and classifying different kinds of PQ disturbances.

Conjecture 1:

- If the peak-value of the ED is located at scale 4 (curve c and d), it is an amplitude distortion which means either swell or sag disturbance. Otherwise, it is a frequency distortion
- If the triangle (peak-value) is concave downward (negative ED), the distortion is a swell. If the triangle is concave upward (positive ED), the distortion is sag
- If the peak-value is at scale smaller than 6 (curve b), it is a high-frequency disturbance
- If the peak-value is at scale higher than 6 (curve a), it is a low-frequency disturbance

One thing should be noted here is that the reference scale 6 is related to the sample frequency f_s and normal frequency of the power signal (50 or 60 Hz).

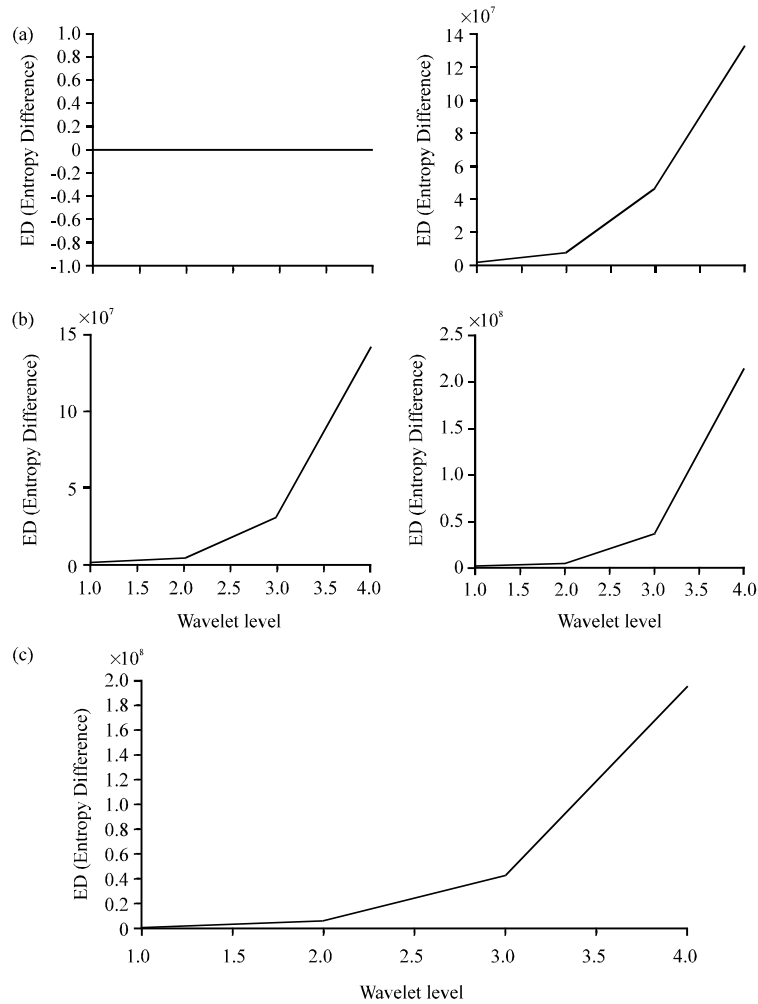


Fig. 7: Entropy Based Feature Method with different wavelet family; a) Entropy Based Feature Method entropy based feature method sine wave -db4; b) Entropy Based Feature Method sym8 wave and coif5 and c) Entropy based feature method-bior-3.9

Determination of the MDL: Using the methodology presented so far, researchers can detect, localize and classify different kinds of PQ disturbances based on the proposed method. However, how many levels of decomposition are enough for the proposed method to be effective? Obviously, more levels of decomposition will increase the computational cost. In this study, researchers aim to find the Minimum Decomposition Level (MDL) for the proposed method and modify the above evaluation criterions to be universal.

In MRA, since both the high pass filter and the low pass filter are half band, the decomposition in frequency domain for a signal sampled with the sample frequency f_s can be demonstrated in Fig. 8 assuming the total decomposition levels for entropy-based features method is 1.

The 6.1 sag signal is applied on wavelet transform al is first level approximation coefficient d1, d2, d3 and d4 are details coefficient (Fig. 9).

Table 1 shows the frequency range at each decomposition level based on the frequency decomposition shown in Fig. 8. In the proposed Entropy-Based Features Method, only the detail information at each decomposition level is needed (Zhang *et al.*, 2003). Assuming that the normal reference frequency of the power signal is Freq. (50 or 60 Hz) and researchers want to locate the energy of this reference signal at level N. According to Table 1, researchers have:

$$\frac{f_s}{2^{N+1}} \leq f_{ref} \leq \frac{f_s}{2^N} \quad (4)$$

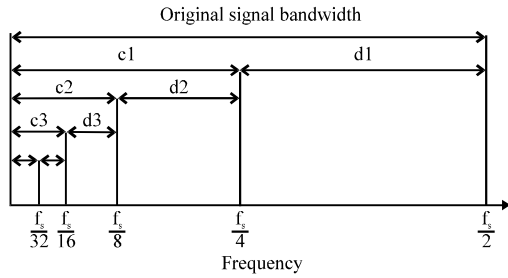


Fig. 8: The wavelet decomposition in the frequency domain

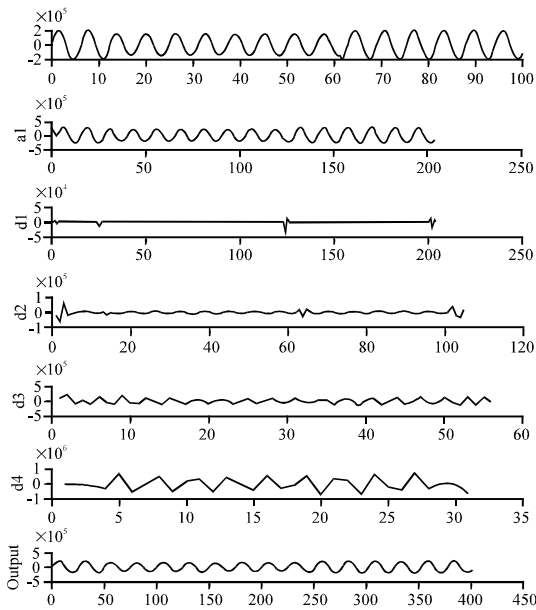


Fig. 9: 6.1 sag signal is applied on wavelet transform a1 is first level approximation coefficient d1, d2, d3 and d4

Table 1: Frequency range of the MRA decomposition

Decomposition level (l)	Frequency range	
	Approximation information (A)	Detail information (D)
1	$0-f_s/2^2$	$f_s/2^2-f_s/2^1$
2	$0-f_s/2^3$	$f_s/2^3-f_s/2^2$
...
n	$0-f_s/2^{n+1}$	$f_s/2^{n+1}-f_s/2^n$

From Eq. 4, researchers get:

$$\log_2 \frac{f_s}{f_{ref}} - 1 \leq N \leq \log_2 \left(\frac{f_s}{f_{ref}} \right) \quad (5)$$

Since, researchers need to locate the energy of the reference frequency at the center of the final result (level 6 as shown in Fig. 7), the MDL for the proposed methods, denoted as N_{min} is found by:

$$N_{min} = 2 \times N \quad (6)$$

For instance, in Fig. 6, sampling frequency $f_s = 5000$ Hz, $f_{ref} = 60$ Hz, according to Eq. 5, researchers get: $5.3808 < N < 6.3808$. Researchers should choose $N = 6$. According to Eq. 6, researchers get $N_{min} = 2 \times N = 12$. This is the reason why researchers choose 12 levels of decomposition in Fig. 7 and why the evaluation reference scale is 6 in the criteria proposed in study.

For any actual sampled signal, Eq. 5 and 6 give the most efficient way of finding the decomposition levels for the Entropy-Based Features Method. For example, if researchers sample the same waveform in Fig. 6 by sampling frequency $f_s = 2500$ Hz, researchers find only 10 levels of decomposition is necessary according to Eq. 5 and 6. Figure 10 gives the analysis result of 10 levels decomposition for the same signal as shown in Fig. 4. Comparing the analysis results in Fig. 11 and 5, researchers can see that 4 levels decomposition is enough for classification varies type of PQ disturbances in this situation ($f_s = 2500$ Hz). In this way, researchers saved about 17% of the computational time compared to the 12 levels of decomposition. Now, researchers can modify the evaluation criterions in study to a universal form:

Conjecture 2:

- If the peak-value of the ED is located at scale of $N_{min} = 2$, the disturbance is either sag or swell (negative ED means swell and positive ED means sag)
- If the peak-value is at a scale smaller than $N_{min} = 2$, the disturbance is high frequency distortion; otherwise, it is a low frequency distortion

The regions of this classification are shown in Fig. 11.

Choice of a suitable wavelet: It is well known that the choice of the appropriate wavelet is very important for all the wavelet based PQ analyses (Santoso *et al.*, 1996; Burrus *et al.*, 1998). In this study, researchers investigate the influence of different kinds of wavelets to the proposed Entropy-Based Features Method.

Four commonly used wavelets, namely, Daubechies wavelets, Symlets and Coif lets wavelet are taken into account. Notice that the Haar wavelets are Daubechies wavelets with $N = 1$. Figure 10a-d is the entropy-based features analysis result to the previous distorted signal in Fig. 11 ($f_s = 5000$ Hz, MDL = 12). Table 2 shows their corresponding features and from the Fig. 10a-d, researchers can see that although different wavelets have some influence of the final analysis result all these wavelets keep good resolution to classify various type of

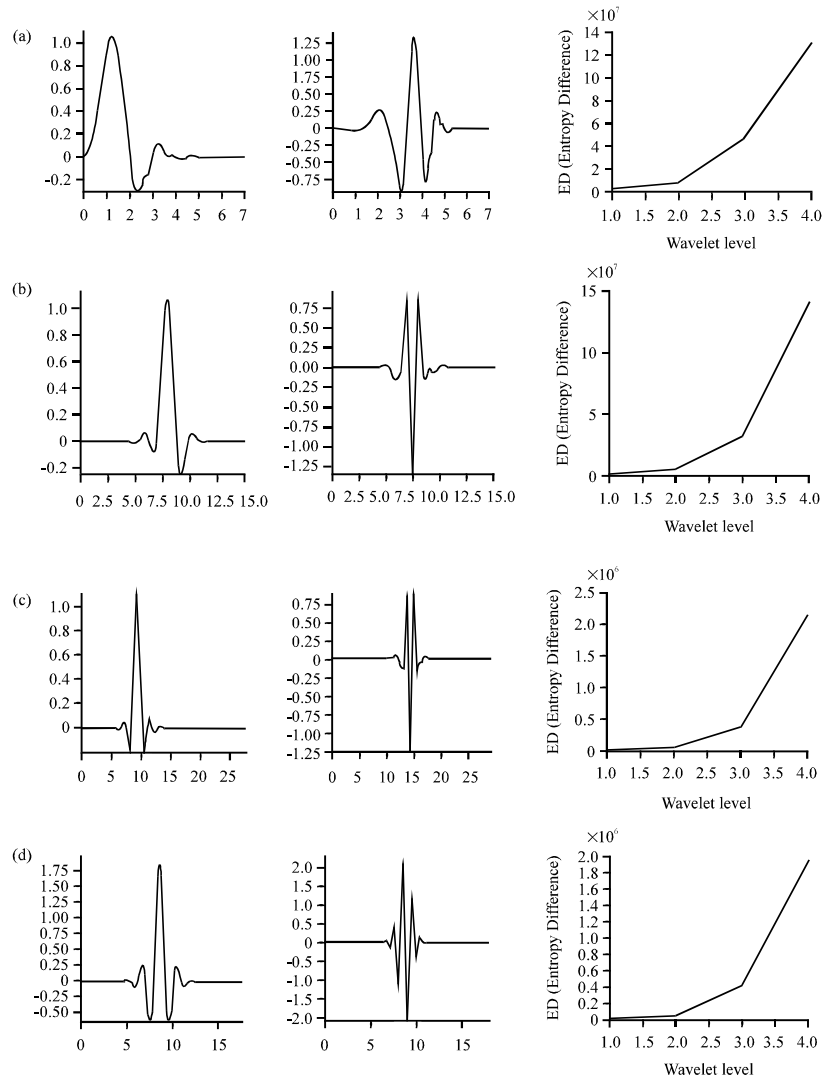


Fig. 10: Entropy-based features analysis; a) db4 wavelet scaling function and wavelet function-entropy-based features analysis result by db4; b) Sym8 wavelet scaling function and wavelet function-entropy-based features analysis result by sym8; c) Coif5 Wavelet scaling function and wavelet function-entropy-based features analysis result by coif5 and d) Bior3.9 wavelet scaling function and wavelet function-entropy-based features analysis result by bior3.9

PQ disturbances. By these means, researchers can claim that the Entropy-Based Features Method has good robustness characteristics. However, it is very important to evaluate corresponding computational cost of different Wavelets. This will provide useful information for actual application, considering the tradeoff between performance and time cost.

Assume that the input discrete signal $x(t)$ is represented by a vector of length $N = 2K$. The DWT using a wavelet with M filter coefficients, researchers need to compute:

$$y_A[n] = \sum_{k=-\infty}^{\infty} x(k)c[2n-k] = \sum_{k=0}^{M-1} c[k]X[2n-k] \quad (7)$$

$$y_D[n] = \sum_{k=-\infty}^{\infty} x(k)h[2n-k] = \sum_{k=0}^{M-1} h[k]X[2n-k] \quad (8)$$

Where:

y_A = The low pass component (approximation)

y_D = The high pass component (details)

$c[k]$ = Low pass (approximation) filter

$d[k]$ = High pass (detail) filter coefficients as defined in the dilation equations:

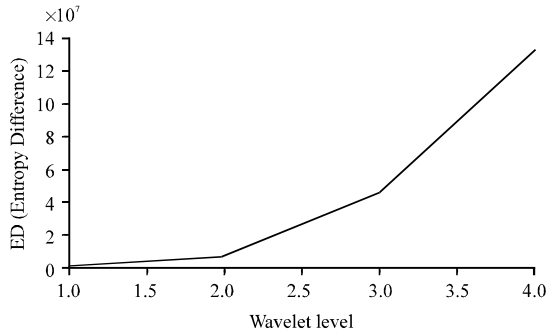


Fig. 11: Analysis result in band limited noise environment (SNR = 20 dB)

Table 2: Wavelet characteristics

Wavelet name	Orthogonal	Compact support	Support width	Filters length	Symmetry
Haar	Yes	Yes	1	2	Yes
Daubechies	Yes	Yes	2N-1	2N	Far from
Coif lets	Yes	Yes	6N-1	6N	Near from
Symlets	Yes	Yes	2N-1	2N	Near from

$$\phi(t) = \sum_{k=0}^{M-1} c(k)\phi[t - k] \text{ and } \psi = \sum_{k=0}^{M-1} h[k]\phi[t - k] \quad (9)$$

where, $h[k] = (-1)^k c[M-k]$. M is the filter length (total number of filter coefficients). Notice that when the DWT applied to discrete signal (a vector), the computation is simply the convolution of two vectors, the signal and the filter coefficients. After a type of wavelets is chosen, the length of wavelet filter will be kept the same at all levels. For this reason, the computation complexity is mainly dependent on the length of the wavelet filters. Since, the mother wavelet produces all wavelet functions (via the dilation equations) used in the transformation through translation and scaling, it determines the characteristics (such as smoothness, symmetry) of the:

$$\Delta t\% = \frac{t_{\text{wavelets}} - t_{\text{Haar}}}{t_{\text{Haar}}} \quad (10)$$

Where:

t_{wavelets} = The Entropy-Based Features Method computational time by the specific choosing wavelets

t_{Haar} = The reference computation time of the Entropy-Based Features Method by Haar wavelet

$\Delta t\%$ = How much more computational time should the specific choice of wavelet needed for Entropy-Based Features Method compared to that of Haar wavelet

Based on the above definition, although, Haar wavelet has the smallest computational cost, further

investigation shows that Haar resulting wavelet transform. Therefore, the details of a particular application should be taken into account and the appropriate mother wavelet should be chosen in order to use wavelet transform effectively.

For example, the Coif lets have important near symmetry property which is highly desired in image processing, since they correspond with nearly linear phased filters. Certainly, if the computational complexity is the only factor under consideration, the filter length should be chosen as short as possible. Researchers see that, the time complexity is proportional to the number of filter coefficients, thus wavelets with larger number of coefficient take longer time to compute. To gain empirical Understanding of the computational complexity of different wavelet families, researchers conducted the experiments in Matlab 7.5 by Intel Pentium 4, 1.8 GHz processor using a stopwatch timer.

The computational time for Haar wavelet is chosen as the base for evaluation and researchers define the following variable $\Delta t\%$ for reference wavelet has difficulty to localize the time information for the beginning and ending time of the disturbance. Therefore, in the situation that the classification of different kinds of PQ disturbances is the only concerned issue, researchers can choose Haar wavelet for entropy-based features method to avoid unnecessary computational cost. Otherwise, based on the data in Table 3, researchers recommend using wavelets with shorter filter length such as Db2, Sym2 or Coif1 wavelets for practical application since they have relatively smaller computational cost as well as more reliable analysis performance. Of course, this kind of selection should also be based on their noise tolerance performance in actual applications which will be discussed.

NOISE TOLERANCE ANALYSIS FOR ENTROPY-BASED FEATURES METHOD

Since, noise is omnipresent in electrical power distribution networks, researchers analyze whether the proposed entropy-based features method is still effective in a noisy environment. Two types of noise, namely Gaussian white noise and band limited spectrum noise are considered.

Gaussian white noise: Gaussian white noise was considered in papers (O’Shea, 2002; Sushama *et al.*, 2007; Yang and Liao, 2001) for power quality disturbance analysis. In this study, researchers suppose that the noise riding on the sampled signal for entropy-based features analysis is white Gaussian distribution. Here, researchers

focus on the detection and classification of different kinds of PQ disturbances under different Signal to Noise Ratio (SNR). A detailed discussion about the localization of the beginning and ending time of the disturbances in noise environment can be referred in study (O’Shea, 2002) in which an adaptive threshold of wavelet analysis is proposed to eliminate the noise influence.

Figure 12 shows the entropy-based features analysis result for the signal in Fig. 4 in the noisy environment with SNR = 20 dB. Comparing Fig. 12 with Fig. 5 researchers can see, the Entropy-Based Features Method has good anti-noise performance and allow us to correctly classify different kinds of PQ disturbances In the noisy environment, to test the entropy-based features method performance in different noise environments, researchers use Monte-Carlo Method to get the average correct classification rate when SNR varied from 20-50 dB. The value of SNR is defined as follows:

$$SNR = 10 \log \left(\frac{P_s}{P_n} \right) \text{dB} \quad (11)$$

Where:

P_s = The power (variance) of the signal

P_n = The noise

For each PQ disturbance, 100 cases with different parameters were simulated for each choice of wavelets. The average correct classification rate according to the evaluation criteria proposed in study is calculated for different wavelets under different SNR.

Band limited spectrum noise: In real electrical distribution networks, noise caused by power electronic devices, control circuits, loads with solid-state rectifiers and switching power supplies are not Gaussian white noise. It has been shown that the power quality noise is defined as electrical signals with broadband spectral content <200 kHz superimposed upon the signal (Dugan *et al.*, 2002). In this research, researchers consider a band limited noise spectrum close to the fundamental frequency (60 Hz). Figure 11 shows the entropy-based features analysis (Db4 wavelet, SNR = 20 dB) Result for the distorted signal combined with a band limited noise. As researchers can see here, the entropy-based features method still shows good performance in this kind of noise environment.

To test the performance of different wavelets in different SNR. Within the band limited noise, Fig. 11 shows the Monte-Carlo Method for the average correct classification probability for different SNR.

Based on the analysis of the experimental results in Fig. 11 and 12, researchers conclude that the

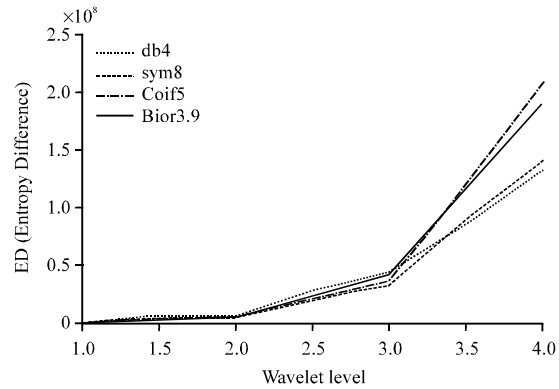


Fig. 12: Entropy-based features all wavelet method

Entropy-Based Features Method is not noise sensitive and performed well in the noisy environments. Although, different types of noise and SNR have some influence on its performance, Entropy-Based Features Method always can achieve high correct classification probability as presented in Fig. 11 and 12.

CONCLUSION

This study considers an efficient feature extraction approach to classifying the PQ disturbances. In this approach, three feature vectors for each disturbance signal are firstly obtained by using different feature extraction methods and are examined for evaluating of classification performance. Then, a combined feature vector is obtained from standard deviation of features belonging to these methods. The experimental results show that the proposed combined feature vector has effectively a classification capability the PQ disturbances. Moreover, the proposed method can reduce the quantity of extracted features of disturbance signal without losing its property. Thus, the classifier system based on proposed Feature Extraction Method needs less memory space and less computing time at both the training and testing processes. Entropy Feature Method can more efficiently separate the critical frequency components of a power quality signal and appears to perform better as a feature extraction tool than other signal analysis tools.

REFERENCES

- Burrus, C.S., R.A. Gopinath and H. Guo, 1998. Introduction to Wavelets and Wavelet Transforms: A Primer. Prentice Hall, USA., ISBN-13: 9780134896007, Pages: 268.

- Dugan, R.C., M.F. McGranaghan, S. Santoso and H.W. Beaty, 2002. *Electrical Power Systems Quality*. McGraw-Hill, New York, pp: 28-29.
- Gaouda, A.M., M.A. Salama, M.R. Sultan and A.Y. Chikhani, 1999. Power quality detection and classification using wavelet-multi resolution signal decomposition. *IEEE Trans. Power Delivery*, 14: 1469-1476.
- Gaouda, A.M., S.H. Kanoun, M.M.A. Salama and A.Y. Chikhani, 2002. Wavelet-based signal processing for disturbance classification and measurement. *IEE Proc. Gener. Transm. Distrib.*, 149: 310-318.
- Hamid, E. Y. and Z.I. Kawasaki, 2002. Wavelet-based data compression of power system disturbances using the minimum description length criterion. *IEEE Trans. Power Delivery*, 17: 460-466.
- He, H. and J.A. Starzyk, 2006. A self-organizing learning array system for power quality classification based on wavelet transform. *IEEE Trans. Power Delivery*, 21: 286-295.
- IEEE Std. 1159, 1995. IEEE recommended practice for monitoring electric power quality. The Institute of Electrical and Electronics Engineers Inc., USA. <http://www.techstreet.com/products/26423>.
- O'Shea, P., 2002. A high-resolution spectral analysis algorithm for power-system disturbance monitoring. *IEEE Trans. Power Syst.*, 17: 676-680.
- Santoso, S., E.J. Powers, W.M. Grady and P. Hofmann, 1996. Power quality assessment via wavelet transform analysis. *IEEE Trans. Power Delivery*, 11: 924-930.
- Sushama, M., G.T.R. Das and A.J. Laxmi, 2007. Mitigation of voltage sags with dynamic voltage restorer using wavelets. *Proceedings of the International Conference on Recent Advancements and Application of Computer in Electrical Engineering*, March 22-24, 2007, Rajasthan, India.
- Yang, H.T. and C.C. Liao, 2001. A de-noising scheme for enhancing wavelet-based power quality monitoring system. *IEEE Trans. Power Delivery*, 16: 353-360.
- Zhang, H., P. Liu and O.P. Malik, 2003. Detection and classification of power quality disturbances in noisy conditions. *IEE Proc. Gener. Transm. Distrib.*, 150: 567-572.