Proceedings of the 5th International Conference on Advances in Power System Control, Operation and Management, APSCOM 2000, Hong Kong, October 2000.

Wavelet Feature Vectors for Neural Network Based Harmonics Load Recognition

W.L. Chan
Dept. of Electrical Engineering
Hong Kong Polytechnic University
Hong Kong
Email: eewlchan@polyu.edu.hk

A.T.P. So
Dept. of Building &Construction
City University of Hong Kong
Hong Kong
Email: bcalbstp@cityu.edu.hk

L.L. Lai
Dept. of Electrical, Electronic & Information Engineering,
City University, London
Email: l.l.lai@city.ac.uk

ABSTRACT

Power quality embraces problems caused by harmonics, over or under-voltages, or supply discontinuities. Harmonics are caused by all sorts of non-linear loads. In order to fully understand the problems, an effective means of identifying sources of power harmonics is important. In this paper, we make use of new developments in wavelets so that each type of current waveform polluted with power harmonics can well be represented by a normalised energy vector consisting of five elements. Furthermore, a mixture of harmonics load can also be represented by a corresponding vector. This paper describes the mathematics and algorithms for arriving at the vectors, forming a strong foundation for real-time harmonics signature recognition, in particular, useful to the re-structuring of the whole electric power industry. The system performs exceptionally well with the aid of an artificial neural network.

1. INTRODUCTION

Power quality has become an important concern both to utilities and their customers. End use equipment is more sensitive to disturbances that arise both on the supplying power system and within the customer facilities. Power quality embraces problems caused by harmonics, over or under-voltages, or supply discontinuities. Harmonics are caused by variable speed drives or computers, over-voltages by capacitor switching, under-voltages by nearby faults or motor starting. While the effects of harmonics on the power system can be of a steady state nature, the effects of capacitor switching or nearby faults are generally more transitory, i.e., of a short-term duration. Power quality concerns for power distribution systems generally fall into two basic categories. The first, which has received most of the attention, is the problem of harmonic distortion. These variations must be measured by sampling the voltage and/or current over time. The second area of concern is short-term and long-term disturbances such as voltage dips, impulses, sags, and surges. Disturbances are measured by triggering on an abnormality in the voltage or the current. Transient voltages may be detected when peak magnitude exceeds a specified threshold. RMS voltage variations may be detected when they exceed a specified level.

To provide a secure service and to improve the electric power quality, sources of disturbances must be known and controlled. This can be done by first detect, localise, and classify different disturbances and then quantify the amount of distortion [1]. Most monitoring systems worked on stationery waveforms of different loads using Fourier Analysis. However, real waveforms always consist of noise in terms of transient spikes, making the conventional Fourier Analysis approach not so accurate under certain circumstances. In this paper, discrete wavelet transforms are adopted to carry out the recognition of load harmonics patterns. If there is an instantaneous impulse disturbance happening at a certain time interval, it may contribute to the Fourier Transform but its location on the time axis cannot be identified. With wavelet transform, both time and frequency information can be obtained [2]. Current waveforms of typical loads on the power system are sampled and converted into a sequence of digital values. Discrete Wavelet Transform is then applied to these values. A Daubechies N=4 or D wavelet is selected as the mother wavelet for the DWT. The DWT is calculated for the signal and the signal of interest is then represented by a total of 5 coefficients (components of a normalized energy vector) for each type of load, the highest frequency being rejected. When a real current waveform recorded from the power system is available, the distinctive signature can be generated and recognize by a three-layer feed-forward neural network. In this way, we have been able to find out the different types of load which contribute electric power harmonics to the power system. The results can be used to assist different power generation companies and consumers to improve the power quality of the whole system.

2. THEORY

The Fourier Transform X(f) of a continuous-time signal x(t) is given by

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt$$

The continuous function X(f) is the frequency domain representation of x(t) obtained by summation of an infinite number of complex exponentials. To find X(f) on a digital computer with sampled and finite length signals, the Discrete Fourier Transform (DFT) is used. It is defined as

$$X[k] = \sum_{n=0}^{N-1} x[n]\, e^{-\frac{j2\pi n}{N}}$$

where x[n] is a sequence obtained by sampling the continuous signal x(t). The DFT produces a sequence of complex values X[k] whose magnitudes are those of discrete frequencies in x[n]. The representation of a signal by the DFT is reserved for periodic signals. The sampling rate should be at least twice as fast as the highest frequency in the signal.

If there is an impulse disturbance happens at a certain time interval of a signal, the entire frequency spectrum of the signal can be affected. However, the location of the disturbance on the time axis will be lost. A solution commonly used is to window the signal into sequence of snapshots, each sufficiently small that the waveform approximates a stationary waveform. The short-time Fourier transform uses a window to localise sharp transitions and minimise leakage. The STFT will work well provided the window is short enough compared with the fluctuation rate. The STFT is similar to the Fourier transform except that the input signal x(t) is multiplied by a window function w(t) whose position is translated in time by b:

$$STFT(f,b) = \int_{-\infty}^{\infty} x(t) w(t-b)\, e^{-j2\pi ft}\, dt$$

For digital implementation, the Windowed Discrete Fourier Transform (WDFT) is defined as

$$WDFT[k,m] = \sum_{n=0}^{N-1} x[n] w[n-m]\, e^{-j\frac{2\pi kn}{N}}$$

where w[n] are samples of the window function. For each window w[n-m], the WDFT produces a sequence of complex values whose magnitudes are those of the discrete frequencies of the input at time m. The WDFT of a signal can be represented in a two-dimensional grid where the divisions in the horizontal direction represent the time extent of each window w[n-m]; the divisions in the vertical direction represent the frequencies k. The DWFT contains frequency information as well as the location of time. Multiple resolution in time and frequency for signals containing a fundamental frequency superimposed with transients. Fine time resolution for short duration and high frequency signals, and fine frequency resolution for long duration and lower frequency signals are needed.

Wavelets are mathematical functions with advantages over Fourier methods for analysis of signals with transient features. Wavelet analysis is based on the decomposition of a signal according to scale rather than frequency, using basis functions with adaptable scaling properties. This method is generally referred to as multi-resolution analysis. A multi-resolution representation provides a hierarchical framework for interpreting signal structure and involves a coarse-to-fine transformation of the discrete-time data. Wavelet transform expands a signal not in terms of a trigonometric polynomial but by wavelets, generated using the translation (shift in time) and dilation (compression in time) of a fixed wavelet function. The wavelet function is localised in time and frequency yielding wavelet coefficients at different scales. The time-frequency localisation property means that more energetic wavelet coefficients are localised in the transform domain, thus providing a basis for data compression. This property is also useful for feature extraction.

The wavelet transform can be accomplished in two different ways depending on what information is required out of this transformation process. The first is a Continuous Wavelet Transform where one obtains a surface of wavelet coefficients, W(a,b), for different values of scaling factor a and translation b. The original signal x(t) has one independent variable t, but the wavelets have two independent variables a and b. The magnitude of the wavelet coefficients provides information on how close the scaled and translated wavelet is to the original signal. The Continuous Wavelet Transform of a continuous signal x(t) is defined as:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\overline{\psi}\left(\frac{t-b}{a}\right) dt$$

For computer implementation, the discrete wavelet transform is used. In the discrete wavelet transformation, the scale and translation variables are discretised. Any wavelet coefficient can be described by two integers, m and n. If a_0 and b_0 are the segmentation step sizes for the scale and translation respectively, the scale and translation in terms of these parameters will be $a = a_0^m$ and $b = nb_0a_0^m$. The discrete wavelet coefficients W(m,n) are given by:

$$W(m,n) = a_0^{-\frac{m}{2}} \int_{-\infty}^{\infty} x(t)\overline{\psi}(a_0^{-m}t - nb_0)\, dt$$

By careful selection of a_0 and b_0, a family of dilated mother wavelets which are orthonormal basis can be arrived. One of the simplest choice would be $a_0=2$ and $b_0=1$. With this choice of the scaling and translation constants we get dyadic-orthonormal wavelet transform. This choice of scaling and translation constants removes information redundancy of the Continuous Wavelet Transform.

An orthonormal compactly supported wavelet basis is formed by the dilation and translation of a single function $\psi(t)$, called the mother wavelet and is given by:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(\frac{t - 2^j k}{2^j})$$

where i,j are integers. The function $\psi(t)$ has M vanishing moments up to order M-1 and it satisfies the following two-scale difference equation:

$$\psi(t) = \sqrt{2} \sum_{k=0}^{L-1} g_k \phi(2t-k)$$

where $\phi(t)$ is called scaling function that forms a set of orthonormal bases as given below:

$$\phi_{j,k}(t) = 2^{\frac{j}{2}} \phi(\frac{t - 2^j k}{2^j})$$

The scaling function satisfies

$$\int_{-\infty}^{\infty} \phi(t) dt = 1$$

and the two-scale recursive difference equation

$$\phi(t) = \sqrt{2} \sum_{k=0}^{L-1} h_k \phi(2t-k)$$

Two coefficient sets $\{g_k\}$ and $\{h_k\}$ have the same length L for a certain basis, where L is related to the number of vanishing moments M in $\psi(t)$. L equals 2M in the Daubechies' wavelets [2]. In the wavelet representation , $\{h_k\}$ behaves as a low-pass filter and $\{g_k\}$ behaves as a high-pass filter to input signals. They are related by:

$$g_k = (-1)^k h_{L-k},\ k = 0,1,\ldots,L-1$$

and are called quadrature mirror filters.

For the Daubechies 4 (D4) wavelet, the dilation equation has the form:

$$\phi(t) = h_0 \phi(2t) + h_1 \phi(2t-1) + h_2 \phi(2t-2) + h_3 \phi(2t-3)$$

$$\text{where } h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}},\ h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}$$

$$h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}},\ h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

It is not possible to solve for $\phi(t)$ directly but can be solve iteratively with $\phi_0(t)=1$ and the above dilation equation. The D4 wavelet function $\psi(t)$ is given by:

$$\psi(t) = -h_3 \phi(2t) + h_2 \phi(2t-1) - h_1 \phi(2t-2) + h_0 \phi(2t-3)$$

With a signal x(t) applied by a D4 wavelet, the wavelet expansion of x(t) becomes:

$$x(t) \approx a_0 \phi(t) + a_1 \psi(t) + \ldots + a_{2^j+k} \psi(2^j t - k) + \ldots$$

$$\text{where } a_{2^j+k} = 2^j \int_0^l x(t)\psi(2^j t - k)dt,\ a_0 = \int_0^l x(t)\phi(t)dt$$

The coefficients a_j represent the amplitudes of the contributing wavelets in a similar manner as the Fourier coefficients. Mallat's pyramid algorithm [4] solves for the coefficients without finding the scaling and wavelet functions directly. For a signal of length N, the number of operations is approximately proportional to N for Mallat's algorithm and N $\log_2$N for the fast Fourier transform.

A wavelet becomes smoother as its number of coefficients increases. The width of each wavelet depends on the number of its coefficients. A wavelet with M coefficients is M-1 intervals wide. The D4 wavelet is more localised in time than D20 wavelet. For this reason, the D4 wavelet is a good choice for accurately detecting fast transient such as impulses. Actual implementation of the discrete wavelet transform, involves successive pairs of high-pass and low-pass filters at each scaling stage as shown in Figure 1. This can be thought of as successive approximations of the same function, each approximation providing the incremental information related to a particular scale, the first scale covering a broad frequency range at the high frequency end of the frequency spectrum however with progressively shorter bandwidths. Conversely, the first scale will have the highest time resolution, higher scales will cover increasingly longer time intervals.

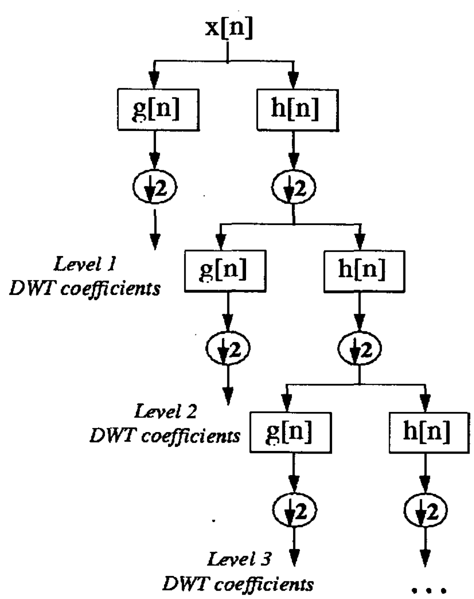

Figure 1 DWT decomposition

3. RESULTS

10 kHz sampling rate has consistently been used to take records of current waveforms of the following three types of common non-linear loads in buildings with a total period of 51.2 ms, i.e. 512 points per record.

1. ordinary 486 personal computer
2. fluorescent tube with electronic ballast
3. 200 W dimmer for incandescent light

Five level decomposition of wavelets has been applied for analysis. Although it is more desirable to have higher levels of wavelets, we have given up the idea due to computational efficiency and practical consideration that five levels have been good enough to represent the whole picture. Figure 2 shows the current waveform and wavelet decomposition of the personal computer. The first curve shows the current waveform; the second one shows a level-5 coarse approximation of the waveform; the third one shows the level-5 wavelet coefficient; the fourth, fifth, sixth and seventh curves show the level-4, level-3, level-2 and level-1 coefficients accordingly. Figure 3 shows the corresponding waveforms and coefficients of the electronic ballast. Figure 4 shows those belong to the 200 W dimmer.

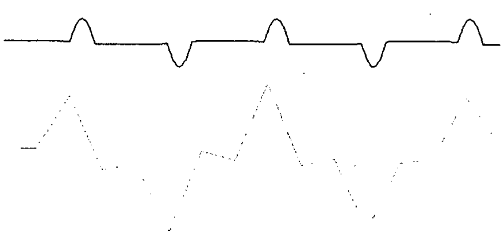

Figure 2 Wavelet Decomposition of PC Current Waveform

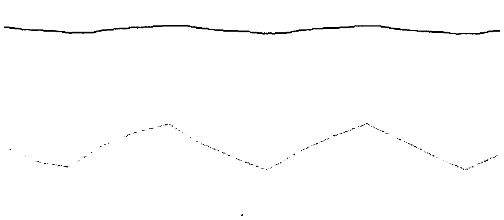

Figure 3 Wavelet Decomposition of Fluorescent Tube Current Waveform

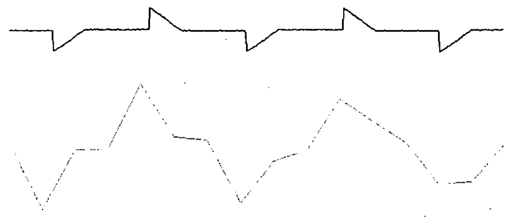

Figure 4 Wavelet Decomposition of 200 W Dimmer Current Waveform

Figure 5 shows the corresponding waveform and coefficients of a mixture of personal computer and fluorescent with electronic ballast. Figure 6 shows a mixture of electronic ballast and the 200 W dimmer.

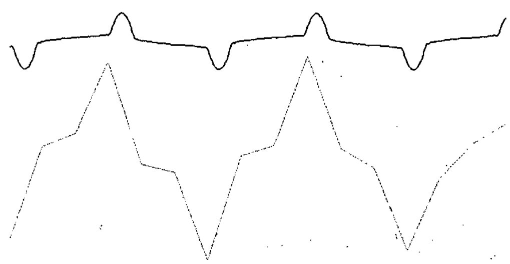

Figure 5 Wavelet Decomposition of Personal Computer and Fluorescent Tube

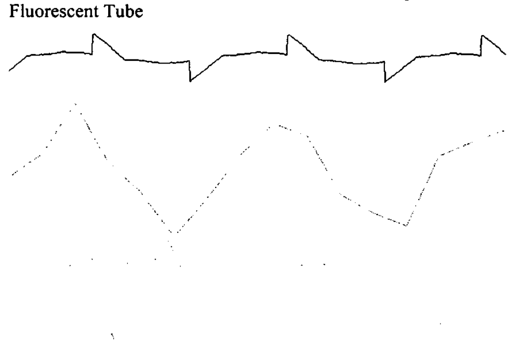

Figure 6 Wavelet Decomposition of 200 W Dimmer and Fluorescent Tube

Figure 7 shows a mixture of dimmer and personal computer. Figure 8 shows a mixture of all three types of non-linear loads. It can be seen that for each type of load or their combinations, the coefficient curves give distinctive signatures and we used these signatures to identify the types of load connected to the circuit. The method is straight forward while the computational loading is not excessive. The lowest curve within each figure gives the coefficient of highest frequency energy. It is used to capture transients of higher frequency, i.e. fast transients. Hence, normally, its magnitude is generally relatively small. Our experience has shown that it is often less than 1% of total energy. Therefore, we are able to ignore its existence. As it can give the finest details of the waveform, it can be used to theoretically discriminate the existence of different loads of the same type, e.g. two different brands of personal computers etc. As the level is increasing, the frequency content is reduced by half for each level and each curve can be used to illustrate different features of the same load, i.e. to characterise different types of loads.

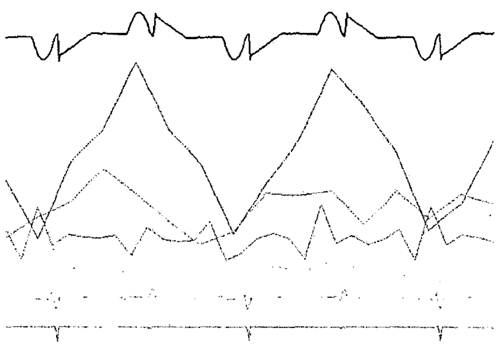

Figure 7 Wavelet Decomposition of 200 W Dimmer and Personal Computer

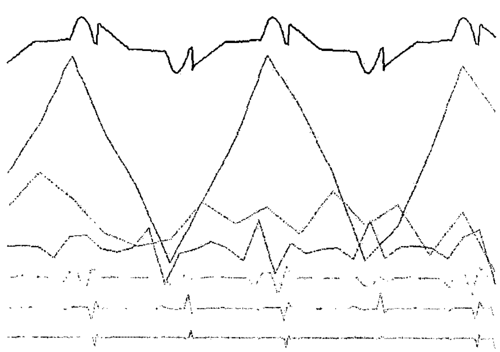

Figure 8 Wavelet Decomposition of all three Non-linear Loads

As mentioned above, the curve of the coefficient of each level represents a particular feature of the current waveform. The whole curve is summarised by one numerical value, i.e. the normalised energy. Five energy values altogether constitute an normalised energy vector with five elements of a particular load. Table 1 shows the summary of normalised energy vector of each load or its combination, i.e. "pc" for personal computer alone, "ec" for electronic ballast alone, "di" for dimmer alone, "ecdi" for electronic ballast plus dimmer, "pcec" for personal computer plus electronic ballast, "dipc" for dimmer plus personal computer and "pcecdi" for the combination of all three types of load. Hence, it can be visualised that each type of load is represented by a vector of five values. Such values can be fed into an artificial neural network for training purpose and later for fast signature recognition. As an illustration, an ANN of 5 input nodes, 2 hidden notes and 3 output nodes was used. After 15,000 iterations, the outputs of the ANN are shown in Table 2. That shows good performance of this method when aided by a neural network for fast signature recognition. For verification and testing purpose, waveforms of all non-linear loads plus a new fluorescent tube using conventional ballast have been recorded, as shown in Figure 5. Data of such new fluorescent tube has not been included in the original ANN training. It can be seen, from Table 2, that the ANN is still able to recognise the combinations.

4. CONCLUSIONS

A new method of identifying harmonics signature of different types of non-linear loads has been developed involving discrete wavelet transforms. Each type of load is well represented by five levels of wavelet decomposition. The result is the creation of a normalised energy vector for each type of load, each vector consisting of five elements. These five elements can characterise a particular type of load. Intelligent and automatic signature recognition can be achieved in future with the aid of artificial neural networks. Then, optimised harmonics compensation can be carried out even under a noisy environment due to the rejection of level-1 wavelet coefficient which represents the higher frequency component. It is hoped that this method can be further developed to identify and analyse transient disturbance within a power system when non-periodic waveforms are involved.

5. REFERENCES

[1] Chan W.L. and So A.T.P., "Power harmonics pattern recognition by solving fuzzy equations", Fuzzy Sets and Systems, Vol. 67, 1994, pp. 257-266.

[2] Strang G. and Nguyen T., Wavelets and Filter Banks, Wellesley Cambridge Press, Boston, 1996.

[3] Paubechies I., Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

[4] Mallat S., "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Trans. Pattern Anal. And Mach. Intell., Vol. 11, July, 1989, pp. 674-693.

Table 1 Wavelet Energy Vectors for different Load Combinations

Type of Load	Level 2	Level 3	Level 4	Level 5	Coarse Approximation	Total Energy
Pc	0.21%	1.79%	13.98%	17.30%	66.69%	251
Ec	0.08%	0.18%	0.60%	3.16%	95.89%	24
Di	2.25%	4.72%	9.76%	18.81%	63.20%	182
Ecdi	2.84%	3.04%	7.49%	23.31%	61.71%	242
Pcec	0.48%	1.10%	7.68%	14.26%	76.30%	427
Dipc	1.19%	4.03%	10.37%	10.73%	72.93%	411
Pcecdi	1.59%	2.73%	7.64%	11.35%	76.16%	592
Pcftecd	0.79%	1.58%	4.41%	3.18%	89.59%	982

Table 2 Output of Neural Network for Recognition

Type of Load	Coarse	Output 1	Output 2	Output 3
pc	66.69%	1.0000	0.0006	0.0006
Ec	95.89%	0.0006	1.0000	0.0008
Di	63.20%	0.0008	0.0009	1.0000
Ecdi	61.71%	0.0000	1.0000	0.9997
Pcec	76.30%	1.0000	1.0000	0.0000
Dipc	72.93%	1.0000	0.0000	1.0000
Pcecdi	76.16%	0.9991	0.9989	1.0000
Pcftecd	89.59%	0.9995	1.0000	0.9991