Implementation of Power Disturbance Classifier Using Wavelet-Based Neural Networks

ZWE-LEE GAING

Abstract- In this paper, a wavelet-based neural network classifier for recognizing power quality disturbances is implemented and tested under various transient events. The discrete wavelet transform (DWT) technique is integrated with the probabilistic neural network (PNN) model to construct the classifier. First, the multi-resolution analysis (MRA) technique of DWT and the Parseval’s theorem are employed to extract the energy distribution features of the distorted signal at different resolution levels. Second, the PNN classifies these extracted features to identify the disturbance type according to the transient duration and the energy features. Since the proposed methodology can reduce a great quantity of the features of distorted signal without losing its original property, less memory space and computing time are required. Various transient events are tested, the results show that the classifier can detect and classify different power disturbance types efficiently.

Keywords- power quality, wavelet transform, Parseval’s theorem, probabilistic neural network

I. INTRODUCTION

Poor power quality (PQ) may cause many problems for the affected loads, such as malfunctions, instabilities, short life-time and so on. Poor quality of electric power is normally caused by power line disturbances such as impulses, notches, glitches, momentary interruptions, wavefaults, voltage swell/sag, and harmonic distortion, resulting in misoperation or failure of end-use equipment. In order to improve power quality, the sources and causes of such disturbances must be known before appropriate mitigating actions can be taken. A feasible approach to achieving this goal is to incorporate detection capabilities into monitoring equipment so that events of interest will be recognized, captured, and classified automatically. Hence, good performance monitoring equipment must have functions which involve the detection, localization, and classification of transient events. In particular, when the disturbance type has been classified accurately, the power quality engineers can define the major effects of the disturbance at the load and analyze the source of the disturbances, so that an appropriate solution can be formulated [1-2].

Due to the wavelet analysis block can transform the distorted signal into different time-frequency scales, so it is a well-suited tool for analyzing high-frequency transients in the presence of low-frequency components such as non-stationary and non-periodic wide-band signals. Currently, the wavelet transform (WT) technology has often been employed to capture the time of transient occurrence and extract frequency features of power disturbance [3-15].

Mo et al. [11] demonstrated how to extract the features from the wavelet transform coefficients at different scales as inputs to neural networks for classifying the non-stationary signal type. Santoso et al. [3][18][20] proposed to extract the squared wavelet transform coefficients (SWTC) at each scale as the inputs to the neural networks for classifying the disturbance type. Perunicic et al. [16] employed the DWT coefficients as inputs to a single-layer self-organizing map neural network to train and classify the transient disturbance type. Elmittawy et al. [17] used the preprocessed DWT coefficients as inputs to a refined neuro-fuzzy network to train and classify the power system disturbance type. Angrisani et al. [12] proposed to employ the continuous wavelet transform (CWT) to estimate the disturbance time duration and the DWT to estimate the disturbance amplitude. The two features thus obtained are then used to classify the transient disturbance type. Santoso et al. [13-14] presented a wavelet-based neural classifier integrating the DWT, learning vector quantization (LVQ) neural network, and decision-making scheme to become an actual power disturbance classifier. The classifier employed the DWT coefficients as inputs to multiple LVQ neural networks to train and perform waveform recognition, and use the decision-making schemes to classify the transient disturbance type. Chung et al. [19] presented a novel classifier using a rule-based method and a wavelet packet-based hidden Markov model (HMM). The rule-based method is employed to classify the time-characterized-feature disturbances, while the wavelet packet-based HMM is utilized to categorize the frequency-characterized-feature power disturbances.

As seen in the above studies, the DWT technology has often been employed to capture the time of transient occurrence and extract frequency features of power disturbance. Integrating the DWT technology with the artificial intelligence method or expert system to become a practical power disturbance classifier for recognizing accurately the disturbance has attracted much research interest. However, two practical problems must be overcome in the above methods. One is that adopting directly the DWT coefficients as inputs to the neural networks requires a large memory space and much learning time, and the other is that the decomposition level with the number of extraction features must be reduced to enhance computing efficiency and accuracy of recognizing the disturbance type.

This paper presents a novel classifier consisting of two methods. First, the wavelet MRA technique and the Parseval’s theorem are employed to extract the energy distribution features of the distorted signal at different resolution levels. Second, the PNN is employed to classify disturbance types according to the transient duration and the detailed energy distribution. By using the Parseval’s theorem, the number of features of distorted signal can be reduced...
without losing its property. Because the efficiency of learning and recall of the PNN is fast, it is suitable for real-time operation for fault diagnosis and signal classification problems.

Sixty distorted signals that were generated by Power System Blockset in Matlab, such as momentary interruption, capacitor switching, voltage sag/swell, and harmonic distortion, to test the proposed classifier. The accuracy rate and computation time of the proposed method are measured. The experimental results showed that the proposed method could analyze the signals efficiently, thus enhancing the performance of power transient recorder.

II. WAVELET TRANSFORM

The wavelet transform (WT) uses the wavelet function \( \phi \) and scaling function \( \phi \) to perform simultaneously the multi-resolution analysis (MRA) decomposition and reconstruction of the measured signal.

### 2.1 Multi-resolution Analysis and Decomposition

The first main characteristic in WT is the MRA technique that can decompose the original signal into several other signals with different levels (scales) of resolution. From these decomposed signals, the original time domain signal can be recovered without losing any information.

The recursive mathematical representation of the MRA is as follows:

\[
V_j = W_{j+1} \oplus V_{j+1} = W_{j+1} \oplus W_{j+2} \oplus \cdots \oplus W_{j+n} \oplus V_n \quad (1)
\]

where

- \( V_{j+1} \) is the approximated version of the given signal at scale \( j+1 \),
- \( W_{j+1} \) is the detailed version that displays all transient phenomena of the given signal at scale \( j+1 \),
- \( \oplus \) denotes a summation of two decomposed signals,
- \( n \) is the decomposition level.

### 2.2 Mathematical Model of DWT

Before the WT is performed, the wavelet function \( \phi(t) \) and scaling function \( \phi(t) \) must be defined. The wavelet function serving as a high-pass filter can generate the detailed version of the distorted signal, while the scaling function can generate the approximated version of the distorted signal. In general, the discrete \( \phi(t) \) and \( \phi(t) \) can be defined as follows:

\[
\phi_{j,n}[t] = 2^{j/2} \sum_n c_{j,n} \phi[2^j t - n],
\]

\[
\phi_{j,n}[t] = 2^{j/2} \sum_n d_{j,n} \phi[2^j t - n].
\]

Simultaneously, the two functions must be orthonormal and satisfy the properties as follows:

\[
\langle \phi, \phi \rangle = \frac{1}{J^2},
\]

\[
\langle \phi, \phi \rangle = \frac{1}{J^2},
\]

\[
\langle \phi, \phi \rangle = 0.
\]

where

- \( c_j \) is the scaling coefficient at scale \( j \),
- \( d_j \) is the wavelet coefficient at scale \( j \).

Assuming the original signal \( x_j[t] \) at scale \( j \) is sampled at constant time interval, thus \( x_j[t] = (x_{0}, x_{1}, \ldots, x_{N-1}) \), the sampling number is \( N = 2^j \). \( j \) is an integer number. For \( x_j[t] \), its DWT mathematical recursive equation (as \( V_j = V_{j+1} \oplus W_{j+1} \)) is presented as follows:

\[
DWT(x_j[t]) = \sum_k x_j[k] \phi_{j,k}[t],
\]

\[
= 2^{(j+1)/2} \left( \sum_n u_{j+1,n} \phi[2^{j+1} t - n] \right),
\]

\[
\begin{array}{c}
0 \leq n \leq N/2^j - 1.
\end{array}
\]

where

\[
\begin{align*}
\phi_{j+1,n} &= \sum_k c_{j,k} V_{j,k+2n}, \quad 0 \leq k \leq N/2^j - 1, \\
\phi_{j+1,n} &= \sum_k d_{j,k} V_{j,k+2n}, \quad 0 \leq k \leq N/2^j - 1, \\
d_k &= (-1)^k c_{2p-1-k}, \quad p = N/2^j, \\
\end{align*}
\]

Figure 1 illustrates the three decomposed/reconstructed levels of the DWT algorithm. At each decomposition level, the length of the decomposed signals (e.g. \( u_{j+1} \) and \( w_{j+1} \)) is half of that of the signals \( (x_0) \) in the previous stage.

According to the orthonormal wavelet functions and Equation (5), the signal \( x_j[t] \) can be reconstructed from both \( u_{j+1} \) and \( w_{j+1} \) coefficients using the inverse discrete wavelet transform (IDWT) (as \( V_{j+1} \oplus W_{j+1} = V_j \)).

![Figure 1 Three decomposed/reconstructed levels of DWT](image)

### III. PARSEVAL’S THEOREM IN DWT APPLICATION

In Parseval’s theorem, assuming a discrete signal \( x/n \) is the current that flows through the 1 \( \Omega \) resistance, then the consumptive energy of the resistance is equal to the square sum of the spectrum coefficients of the Fourier transform in the frequency domain.

\[
\frac{1}{N} \sum_{n=-N}^{N} |x[n]|^2 = \sum_{k=-N/2}^{N/2} |a_k|^2
\]

where \( N \) is the sampling period, and \( a_k \) is the spectrum coefficients of the Fourier transform.

To apply this theorem to the DWT, we use Equations (5) and (9) to obtain Equation (10) that is the Parseval’s theorem in the DWT application.

\[
\frac{1}{N} \sum_j |x[j]|^2 = \frac{1}{N_j} \sum_k |u_{j,k}|^2 + \sum_{j=1}^{J} \left( \frac{1}{N_j} \sum_k |w_{j,k}|^2 \right)
\]
Hence, through the DWT decomposition, the energy of the distorted signal is shown by Equation (10). The first term on the right of Equation (10) denotes the average power of the approximated version of the decomposed signal, while the second term denotes that of the detailed version of the decomposed signal. The second term giving that the energy distribution features of the detailed version of distorted signal will be employed to extract the features of power disturbance.

IV. FEATURE EXTRACTION AND CLASSIFICATION

4.1 Duration of Electromagnetic Transients

In general, when a transient disturbance occurs, the power signal will generate a discontinuous state at the start and end points of the disturbance duration. Employing the DWT technique to analyze the distorted signal, through 1-level decomposition of the MRA, will cause the wavelet coefficients \( w_j \) at the start and end points of the disturbance to generate severe variation. Therefore, we can easily obtain the start time \( t_s \) and end time \( t_e \) of the disturbance duration from the variations in absolute wavelet coefficients \( w_j \), and calculate the disturbance duration \( t_d \).

\[
t_d = |t_e - t_s|
\]  

To filter the noise and correct \( t_d \), the \( w_j \) coefficients need to be modified by subtracting the standard deviation of the absolute wavelet coefficients from the absolute wavelet coefficients.

4.2 Detailed Energy Distribution

As seen in Equation (10), the energy of the distorted signal can be partitioned at different resolution levels in different ways depending on the power quality problem. Therefore, we will examine the coefficient \( w_j \) of the detailed version at each resolution level to extract the features of the distorted signal for classifying different power quality problems. The process can be represented mathematically by Equation (12).

\[
P_j = \frac{1}{N_j} \sum_{k} |w_{jk}|^2 = \frac{\|w_j\|^2}{N_j}
\]  

where \( \|w_j\| \) is the norm of the expansion coefficient \( w_j \).

Four special properties in Equation (12) need further explanation:

- The Daubanchie ‘db4’ wavelet function was adopted to perform the DWT, thus resulting in the energy distributions of the decomposition levels 6, 7, and 8 are larger. However, using different wavelet function will generate different result.
- The energy distribution unaffected by the time of disturbance occurrence.
- The same outline of energy distribution despite variations in the vibration amplitude of the same disturbance type.
- The low-level energy distribution will show obvious variations while the distorted signal with the high-frequency elements. Contrary, the high-level energy distribution will show obvious variations while the distorted signal with the low-frequency elements.

To display clearly the characteristics of above properties, we make the Equation (12) to normalize by Equation (13).

\[
P^D_j = (P_j)^{1/2}
\]

4.3 Probabilistic Neural Network (PNN) [21]

The PNN model is one of the supervised learning networks, and has many features distinct from those of other networks in learning processes. They are as follows:

- it is implemented based on the probabilistic model, such as Bayesian classifiers,
- A PNN is guaranteed to converge to a Bayesian classifier providing it is given enough training data.
- requires no learning processes,
- no need to set the initial weights of network,
- no relationship between learning processes and recalling processes,
- do not use the differences between the inference vector and the target vector to modify the weights of network.

The learning speed of the PNN model is very fast, making it suitable for fault diagnosis and signal classification problems in real-time. Figure 2 shows the architecture of a PNN model that is composed of the radial basis layer and the competitive layer [21].

In the signal classification application, the training examples are classified according to the their distribution values of probabilistic density function (PDF), which is the basic principle of the PNN. A simple PDF is defined as follows:

\[
f_k(X) = \frac{1}{N_k} \sum_{i=1}^{N_k} \exp(-\frac{\|X - X_{ik}\|^2}{2\sigma^2}).
\]

Modifying and employing Equation (14) to the output vector \( H \) of the hidden layer in the PNN is as below:

\[
H_h = \exp(-\frac{\|X - W_{ih}\|^2}{2\sigma^2}),
\]

The algorithm of the inference output vector \( Y \) in the PNN is as follows:

\[
net_j = \frac{1}{N_j} \sum_k W_{kj} \cdot H_k \quad \text{and} \quad N_j = \sum_k W_{kj},
\]

if \( net_j = \max \{ net_k \} \) then \( Y_j = 1 \), else \( Y_j = 0 \).
is the number of input layers,
h is the number of hidden layers,
j is the number of output layers,
k is the number of training examples,
$N_k$ is the number of classifications (clusters)
$\sigma$ is the smoothing parameter (standard deviation), $0.1 < \sigma < 1$. In general, $\sigma$ is set to 0.5.

$X$ is the input vector,
$$\|X - X_{kj}\|$$ is the Euclidean distance between the vectors $X$ and $X_{kj}$, i.e.,
$$\|X - X_{kj}\| = \sum_i (X_i - X_{kj})^2,$$
$W_{ih}$ is the connection weight of between the input layer $X$ and the hidden layer $H$,
$W_{hy}$ is the connection weight of between the hidden layer $H$ and the output layer $Y$.

The learning and recalling processes of the PNN for classification problems can refer to [21].

$\text{Voltage Wave Shape Sampling Data}$

$\begin{array}{c}
\text{Discrete Wavelet Transform (DWT)} \\
\text{Multi-resolution Analysis Deposition}
\end{array}$

$W_1, W_2, \cdots, W_{13}$

$\begin{array}{c}
\text{Disturbance Duration} \\
\text{obtained by Equation (11)}
\end{array}$

$\begin{array}{c}
\text{Energy Feature Extraction by Parseval’s} \\
\text{Theorem Using Equations (12) and (13)}
\end{array}$

$I_I$

$P_{11}^D, P_{21}^D, \cdots, P_{13}^D$

$\begin{array}{c}
\text{Transient Signal Classification using Probabilistic Neural} \\
\text{Network}
\end{array}$

$\begin{array}{c}
\text{Harmonic} \\
\text{Capacitor} \\
\text{Switching} \\
\text{Voltage Swell} \\
\text{Voltage Sag} \\
\text{Interuption}
\end{array}$

Figure 3  Diagram of the proposed classifier

### 4.4 Classification of Transient Signals Using PNN Model

Though the PNN has the some disadvantages, such as large memory requirement and the recalling time is proportional to the quantity of training samples, however, we can overcome these drawbacks that employ the Parseval’s theorem reduced the training input.

In this study, we will perform a 13-level decomposition of each discrete distorted signal to obtain the detailed version coefficients $w_i \sim w_{13}$. Using Equation (11), we can obtain the disturbance duration $t_i$ by the squared wavelet coefficients of 1-level decomposition. Simultaneously, with Equations (12) and (13), we can obtain each detailed energy distribution $(P_1^D \sim P_{13}^D)$. These features would be applied to the PNN for classifying the distorted signals. The calculation procedures of the proposed classifier are shown in Figure 3.

### V. APPLICATIONS AND RESULTS

#### 5.1 Validation and Feature Extraction

To verify the feasibility of the proposed method, we used the Power System Blockset Toolbox in Matlab to generate one pure sine wave signal (frequency = 60Hz, amplitude = 1 p.u.) and five transient distorted signals. These distorted signals included momentary interruption, capacitor switching, voltage sag/swell, and harmonic distortion. The sampling rate of the digital recorder was 256 points/per-cycle. The Daubechies 'db4' wavelet was adopted to perform the DWT.

The PNN model was provided by the Neural Network Toolbox in Matlab. The proposed method was written in Matlab language and executed on a Pentium III 550 personal computer with 256MB RAM.

Figures 4 to 9 shows the detailed version of a 3-level decomposition $(w_1 \sim w_3)$ and the detailed energy distribution $(P_1^D \sim P_{13}^D)$ of each given distorted signal. In the harmonic distortion case, the disturbance duration will be set to be very short because the voltage wave shape is disturbed periodically. The simulation results are summarized in Table 1.

From Figures 4 to 9 and Table 1, we can categorize three properties of energy distribution of the given distorted signals. These properties become the foundations for classifying the disturbance type.

- When a sag or swell or interrupt occurs, $P_6^D$, $P_{11}^D$, and $P_{13}^D$ will show great variations.
- When the voltage suffers a transient disturbance of the high-frequency elements such as capacitor switching and harmonic distortion, $P_3^D$, $P_8^D$, and $P_{12}^D$ will show obvious variations.
- When the voltage suffers a transient disturbance of the low-frequency elements such as flicker, $P_5^D$, $P_{10}^D$, and $P_{11}^D$ will show obvious variations.

Figure 10 shows orderly the energy distributions of six signals on same 3-D coordinate axis (using piecewise style). Thus, we can clearer observe the differences of energy distribution between different signals.

#### 5.2 Training Examples and Signal Classification

Table 1 also shows the format of training examples of the PNN model. Each training data involves thirteen energy values $(P_1^D \sim P_{13}^D)$, one disturbance duration time $(t_i)$, and one expected classification type.

Because the time of occurrence, duration, and amplitude magnitude of electromagnetic disturbance in a power system are random, we created randomly 10 voltage waves of each given distorted signal that had different time of occurrence, duration, and amplitude magnitude from the distorted signal property as the training examples. There were 60 distorted signals generated in this paper (involving the pure sine wave). Using Equations (11), (12) and (13), we can obtain the features of the 60 signals. They are rearranged as the format shown in Table 1 to be the training examples for training the PNN model. The numbers of input and output layers in the PNN model were set to be 14 and 6, respectively.
First, we employed 30 (a half of all training examples) and 60 training examples to train the PNN model, respectively. Then, we also created randomly 20 distorted voltage waves to test the proposed approach. The experimental results are shown in Table 2. As can be seen, the training examples were 30, the classified accuracy rate of the distorted signals of the proposed approach was 80% (four distorted signal failures). When the training examples were 60, the classified accuracy rate was 90% (Two distorted signal failures). The results show that the more training examples, the better accuracy rate.

Because the PNN model requires short learning time, the proposed approach is suitable for employing the real-time processing in a modern digital recorder with requirements. The learning time and recalling time of the PNN model are also shown in Table 2.

1. **Pure Sine Wave**

2. **Momentary Interruption ($t_i=0.133$sec.)**

3. **Voltage Sag ($t_i=0.15$sec.)**

4. **Voltage Swell ($t_i=0.15$sec.)**

5. **Capacitor Switching ($t_i=0.0457$sec.)**
Table 1 Energy features and disturbance durations of distorted signals.

<table>
<thead>
<tr>
<th>Signal type</th>
<th>$P_D^1$</th>
<th>$P_D^2$</th>
<th>$P_D^3$</th>
<th>$P_D^4$</th>
<th>$P_D^5$</th>
<th>$P_D^6$</th>
<th>$P_D^7$</th>
<th>$P_D^8$</th>
<th>$P_D^9$</th>
<th>$P_D^{10}$</th>
<th>$P_D^{11}$</th>
<th>$P_D^{12}$</th>
<th>$P_D^{13}$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sine Wave</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0008</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0012</td>
<td>0.0013</td>
<td>0.01</td>
</tr>
<tr>
<td>Intermittent</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0008</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0012</td>
<td>0.0013</td>
<td>0.1</td>
</tr>
<tr>
<td>Saw</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0008</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0012</td>
<td>0.0013</td>
<td>0.1</td>
</tr>
<tr>
<td>Sinc Switching</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0008</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0012</td>
<td>0.0013</td>
<td>1</td>
</tr>
<tr>
<td>Harmonic</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0008</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0012</td>
<td>0.0013</td>
<td>1</td>
</tr>
</tbody>
</table>

6. Harmonic Distortion ($t_r = 0.0001$ sec.)

VI. DISCUSSION AND CONCLUSION

This paper proposed a wavelet-based neural network classifier for power quality disturbance recognition. The proposed method can reduce the quantity of extracted features of distorted signal without losing its property, thus requiring less memory space and computing time for proper classification of disturbance types. The experimental results showed that the proposed classifier have the ability of recognizing and classifying different power disturbance types efficiently, and it has the potential to enhance the performance of power transient recorder with real-time processing capability. Because the distorted signals in this study were generated by simulation software, employing real distorted signals measured by the digital recorder to improve the proposed classifier is one of our future works.

VII. ACKNOWLEDGEMENT

The author gratefully acknowledges the financial support from the National Science Council of Taiwan and the technical support from Taiwan Power Company.

Table 2 Test results of the proposed method

<table>
<thead>
<tr>
<th>Number of training examples</th>
<th>30</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of testing examples</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Learning time (sec.)</td>
<td>0.018</td>
<td>0.035</td>
</tr>
<tr>
<td>Recall time (sec.)</td>
<td>0.2</td>
<td>0.39</td>
</tr>
<tr>
<td>Accuracy rate</td>
<td>80%</td>
<td>90%</td>
</tr>
</tbody>
</table>

REFERENCES


10. David C. Robertson, Octavio I. Camps, Jeffrey S. Mayer, and William


Zwe-Lee Gaing received his M.S. and Ph.D. degree from National Sun Yat-Sen University in Taiwan in 1996. He works as an associate professor of the Electrical Engineering of Kao-Yuan Institute of Technology, Kaohsiung, Taiwan. His research interests are in the filed of artificial intelligence with application to power system operation and control.