Applied Soft Computing 18 (2014) 158–166

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

# Nonstationary power signal time series data classification using LVQ classifier

B. Biswal [a,*], M. Biswal [b], S. Hasan [c], P.K. Dash [c]

[a] GMR Institute of Technology, Rajam, Srikakulam (District), 532127 AP, India
[b] Silicon Institute of Technology, Bhubaneswar, India
[c] S'O'A University, Bhubaneswar, India

ARTICLE INFO

*Article history:*
Received 8 January 2010
Received in revised form 23 June 2012
Accepted 18 January 2014
Available online 28 January 2014

*Keywords:*
Non-stationary power signals
Power quality (PQ)
Wavelet packet decomposition (WPD)
Learning vector quantization

ABSTRACT

A new approach to time–frequency analysis and pattern recognition of non-stationary power signals is proposed in this paper. In this manuscript, visual localization, detection and classification of non-stationary power signals are achieved using wavelet packet decomposition and automatic pattern recognition is carried out through learning vector quantization neural network. The wavelet packet decomposition (WPD) of the non-stationary power signals is carried out to extract the coefficients at multiple level of decomposition. The relevant features for pattern classification are derived from the time-scale information obtained by WPD. The extracted features are used to classify different power quality disturbances by using learning vector quantization neural net. Various non-stationary power signal waveforms are considered to verify the applicability of the proposed technique.



## 1. Introduction

Non-stationary power signal disturbances and power quality (PQ) has been a cause for concern for both the utilities and users due to the use of many types of electronic equipments [1]. Harmonics, voltage swell, voltage sag, transients, and momentary interruptions can adversely affect these equipments. These disturbances cause several problems, such as overheating, failure of motors, disoperation of sensitive and protective equipment and inaccurate metering. Voltage transients can occur due to lightning, capacitor switching, motor starting, nearby circuit faults or accidents, and can also lead to power interruptions. Harmonic currents due to nonlinear loads throughout the network can also degrade the quality of services to sensitive high-tech customers. To distinguish the non-stationary disturbance signal patterns in the normal sinusoidal signal frequencies, advanced signal processing techniques along with pattern recognition approach play a vital role in classifying the patterns of the non-stationary power signal disturbances. Consequently solutions to electrical network disturbance problems involve continuous monitoring of electrical network voltage and current waveforms at certain customer sites and discretizing these signals into digital data to identify the different disturbance patterns using pattern recognition techniques. When a huge electrical network signal database is used for identifying the similar patterns and its localization in time, it constitutes an important area of non-stationary time series analysis.

To ensure a good power quality, a system must be able to monitor, locate, and classify disturbances by measurement approaches and instruments. These instruments must collect large amounts of measured data such as voltages, currents, and occurrence items. However, they do not automatically classify disturbances and they require offline analysis from the recorded data. Fast Fourier transformation (FFT) [2] has been applied for harmonic analysis in power system but short-time duration disturbances require short-time Fourier transformations (STFT) to aid the analysis. The choices for size of window affect both the frequency and time resolution when using STFT. On the other hand another transform like wavelet transforms (WT) [3–10] is a popular time–frequency analysis tool. It is very suitable for analyzing non-stationary signals and has been successfully applied to detection of PQ disturbances. But the WT is only suitable to identify the information in low-frequency band and unsuitable to identify the information in high-frequency band. So when the disturbances are present in high-frequency bands the WT fails to detect them. Hence wavelet packet decomposition (WPD) [11–14] is a finer tool for signal analysis and is a powerful analyzing tool which can avoid the demerits of WT. Because the WPT not only identify the disturbance at high frequency band, but also provide better visual detection and localization of the power signal disturbances than the wavelet transform.

* Corresponding author. Tel.: +91 9000405565.
*E-mail addresses:* birendra_biswal1@yahoo.co.in (B. Biswal), milan.biswal@gmail.com (M. Biswal), shasan_uce@rediffmail.com (S. Hasan), pkdash.india@gmail.com (P.K. Dash).

1568-4946/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2014.01.017

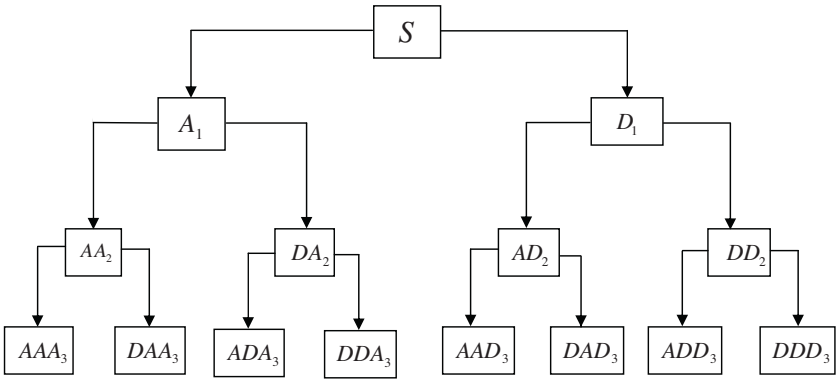



**Fig. 1.** Structure diagram of three-stage wavelet packet decomposition.

Wavelet packet atoms are waveforms indexed by three naturally interpreted parameters: position, scale (as in wavelet decomposition), and frequency. For a given orthogonal wavelet function, we generate a library of bases called wavelet packet bases. Each of these bases offers a particular way of coding signals, preserving global energy, and reconstructing exact features. We then select the most suitable decomposition of a given signal with respect to energy, entropy, standard deviation based criterion. But in the wavelet packet transform, each detail coefficient vector is also decomposed into two parts using the same approach as in approximation vector splitting. Wavelet packets offer a more complex and flexible analysis, because in wavelet packet analysis, the details as well as the approximations are split. The idea of this decomposition is to start from a scale-oriented decomposition, and then to analyze the obtained signals on frequency sub bands. It not only has the ability of local time–frequency analysis like wavelet transform but also can further decompose high-frequency bands and make the frequency resolution improved. Therefore, WPD has wider application value and foreground. Recently several methods of automatic recognition and classification based on artificial intelligence, such as artificial neural network (ANN), have been studied [15,16]. ANN is a popular method of recognition and classification. Because of the excellent time–frequency analysis of WPD and the excellent power signal classification ability of learning vector quantization algorithm (LVQ), we propose a novel method based on WPD and LVQ [17–19] for detection and classification of PQ disturbances. In this the extracted features from the time-series data is applied to the LVQ algorithm to group the data into clusters and thereby identifying the class of the data.

## 2. Wavelet packet signature

Wavelet transform is only suitable to identify the information in low-frequency band and unsuitable to identify the information in high-frequency band. The wavelet packet transform is suitable to finely identify the information in both high and low frequency bands, which is an ideal processing tool of non-stationary signal. In other words, it has higher frequency resolution and lower time resolution in low-frequency part while it has higher time resolution and lower frequency resolution in high-frequency part. In Fig. 1, $S$ represents the original signal. $A_1$ represents the low-frequency component of the original signal $S$. $D_1$ represents the high frequency component of original signal $S$. $AA_1$ represents the low-frequency component of $A_1$. $DA_2$ represents the high frequency component of $A_1$ and so on.

After three-stage wavelet packet decomposition the original signal can be expressed as

$$S = AAA_3 + DAA_3 + ADA_3 + DDA_3 + AAD_3 + DAD_3 + ADD_3 + DDD_3$$

$V_j \oplus W_j \oplus W_{j-1} \oplus W_{j-2} \ldots \oplus W_{j-m} = V_{j-m}$ expresses $A_J$, $W_J$ expresses $D_J$, $V_J$ is scale space, $W_J$ is wavelet spacei.e.

$$V_J \oplus W_j = V_{j-1} \quad j \in Z \tag{1}$$

Obviously,

$$V_j \oplus W_j \oplus W_{j-1} \oplus W_{j-2} \cdots \oplus W_{j-m} = V_{j-m} \tag{2}$$

for multi-resolution analysis, $L^2(R) = \bigoplus_{j \in Z} W_j$. It indicates the Hilbert space $L^2(R)$ is the sum of all orthogonal subspace $W_j (j \in Z)$ in accordance with the different scales of $j$ where $W_j$ is the closure space of wavelet function $\phi(t)$. Wavelet subspace $W_j$ can further subdivide according to binary in order to increase frequency resolution. A new space $U_j^n$ can unify scale space $V_j$ and wavelet space $W_j$

$$U_j^0 = V_j, \quad U_j^1 = W_j \quad j \in Z \tag{3}$$

The Hilbert space orthogonal decomposition $V_{J-1} = V_j \oplus W_j$ can be unified by the decomposition of $U_j^n$,

$$U_{j+1}^n = U_j^{2n} \oplus U_j^{2n+1} \quad j \in Z \tag{4}$$

$U_j^n$ is the closure space of function $u_n(t)$; $U_j^{2n}$ $u_n(t)$; $U_j^{2n}$ is the closure space of function $u_{2n}(t)$. $h(k)$ is the base function expansion coefficient of space $V_{-1}$; $g(k)$ is the orthogonal base function expansion coefficient of space $V_{-1}$. $u_n(t)$ meets the following equation:

$$u_{2n}(t) = \sqrt{2}\sum_{k \in Z} h(k) u_n(2t-k), \quad u_{2n+1}(t) = \sqrt{2}\sum_{k \in Z} g(k) u_n(2t-k) \tag{5}$$

In formula above, $g(k) = (-1)^k h(1-k)$.

When $n = 0$ the formula (5) are given as

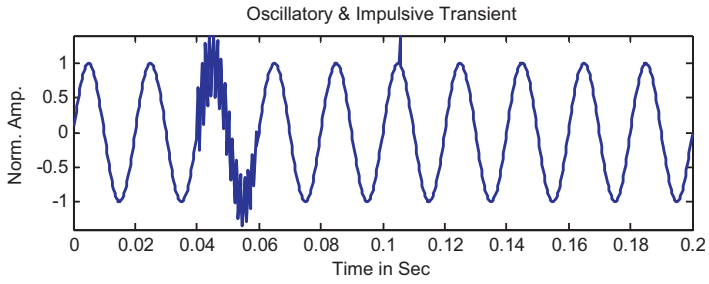
$$u_0(t) = \sqrt{2}\sum_{k \in Z} h(k) u_0(2t-k), \quad u_1(t) = \sqrt{2}\sum_{k \in Z} g(k) u_0(2t-k) \tag{6}$$

For Multi-resolution analysis, $\phi(t)$ and $\varphi(t)$ meet the following equation:

$$\phi(t) = \sqrt{2}\sum_{k \in Z} h(k) \phi(2t-k), \quad \varphi(t) = \sqrt{2}\sum_{k \in Z} g(k) \phi(2t-k) \tag{7}$$

In comparison, it is clear that $u_0(t)$ and $u_1(t)$ degradation scaling function $\phi(t)$ and wavelet function $\varphi(t)$. Eq. (7) is equivalent to Eq. (5). This equivalent extends to $n \in Z^+$. Eq. (4) is expressed as

$$U_{j+1}^n = U_j^{2n} \oplus U_j^{2n+1}, \quad j \in Z;\ n \in Z^+ \tag{8}$$




(a)

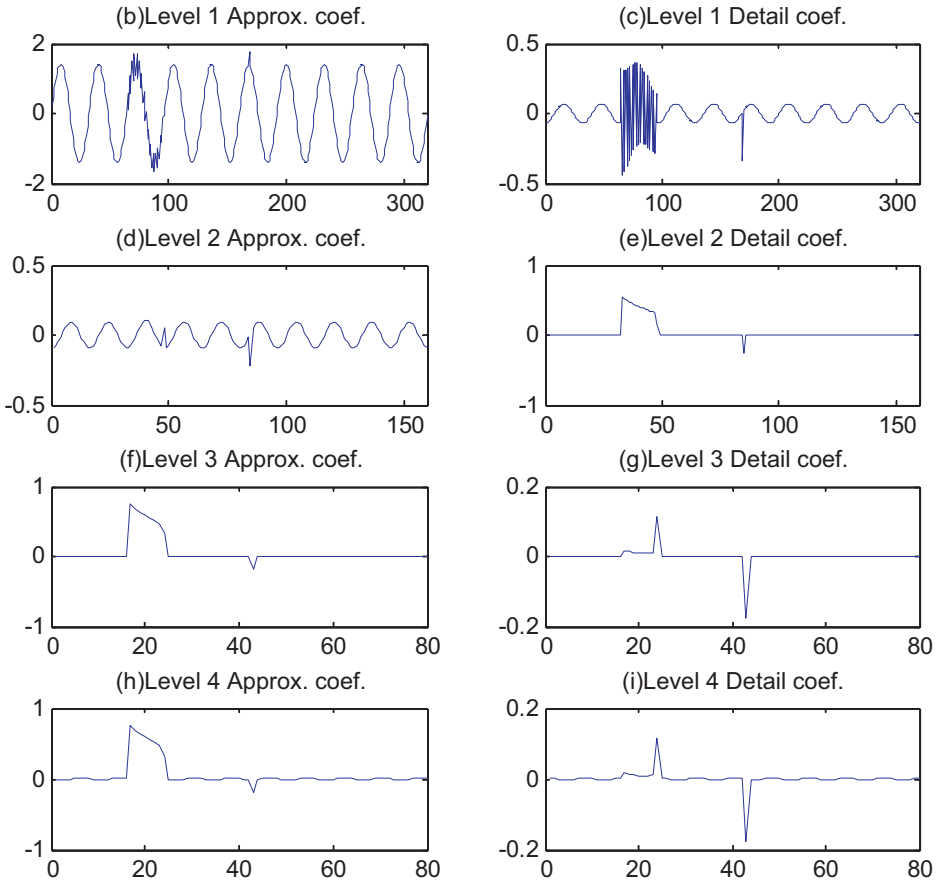



**Fig. 2.** (a) Oscillatory transient followed by impulsive transient; (b) level 1 approximate coefficients; (c) level 1 detail coefficients; (d) level 2 approximate coefficients; (e) level 2 detail coefficients; (f) level 3 approximate coefficients; (g) level 3 detail coefficients; (h) level 4 approximate coefficients; and (i) level 4 detail coefficients (using WPT).

The sequence $g_j^n(t) \in U_j^n$, $g_j^n(t)\{u_n(t)\}_{n \in Z^+}$, which is constructed by Eq. (8) is orthogonal wavelet packet based on the orthogonal basis function $u_0(t) = \phi(t)$.

Located $g_j^n(t) \in U_j^n$, $g_j^n(t)$ can be defined as

$$g_j^n(t) = \sum_l d_l^{j,n} u_n(2^j t - l) \tag{9}$$

Wavelet packet decomposition algorithm,

$$d_l^{j+1,n} = \sum_k [h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1}], \quad d_l^{j,2n+1} = \sum_k b_{k-2l} d_k^{j+1,n} \tag{10}$$

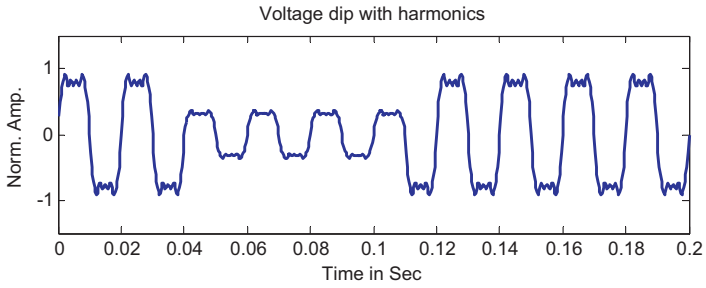
Wavelet packet reconstruction algorithm,

$$d_l^{j+1,n} = \sum_k [h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1}] \tag{11}$$

where space $U_j^n$ is one closure space of the function $u_n(t)$.

## 3. Time–frequency analysis using WPD

The proposed techniques are used to analyze the power signal disturbance in a realistic power network simulated by power system block set supported by MATLAB software. The sampling rate for the collection of power quality data is taken to be equal to 3.84 kHz. In our study we have discussed different types of non-stationary power signal problems. The coefficients after wavelet packet decomposition are shown in Figs. 2 and 3 for a signal with oscillatory & impulsive transient and a signal with voltage dip & harmonics. The Haar wavelet was used as the mother wavelet throughout the simulations in this paper. WPD provides excellent detection, visual localization and classification of the power signal disturbances.




(a)

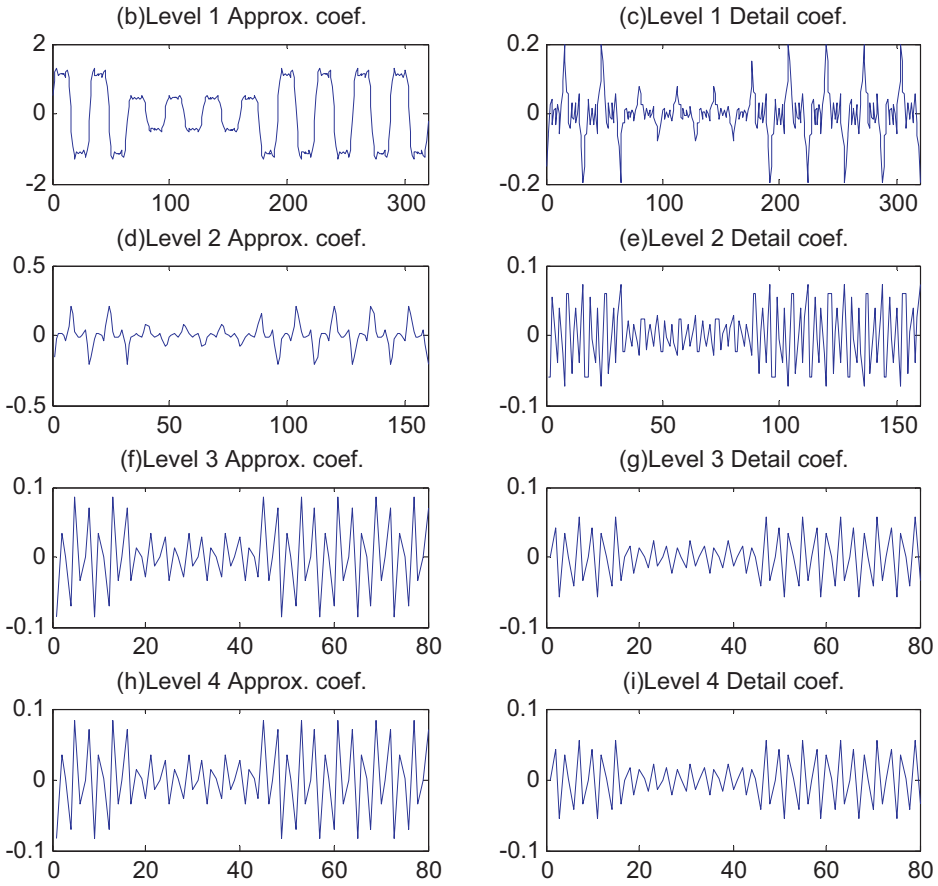



**Fig. 3.** (a) Voltage dip with harmonics; (b) level 1 approximate coefficients; (c) level 1 detail coefficients; (d) level 2 approximate coefficients; (e) level 2 detail coefficients; (f) level 3 approximate coefficients; (g) level 3 detail coefficients; (h) level 4 approximate coefficients; and (i) level 4 detail coefficients (using WPT).

## 4. Feature extraction

The power signal disturbances are non-stationary in nature and require a pattern recognition approach for determining the class to which each of these belongs. In our work we have considered seven classes (power signal disturbances) for classification using LVQ neural network mentioned below in Table 1.

**Table 1**
PQ disturbances with respective classes (g).

| Disturbance | Class |
|---|---|
| Oscillatory transient | $C_1$ |
| Impulsive transient (spike) | $C_2$ |
| Multiple notches | $C_3$ |
| Voltage rise (swell) | $C_4$ |
| Voltage dip (sag) | $C_5$ |
| Voltage rise with harmonics (swell harmonic) | $C_6$ |
| Voltage dip with harmonics (sag harmonic) | $C_7$ |
| Momentary interruption | $C_8$ |

In the wavelet packet domain the basis functions are obtained by translation and scale change. They remain well localized in both time (spatial) and frequency domains and thus represent scale and spatial information. Thus, a complete tree presents the distribution of a signal within a scale space continuum. The total number of coefficients in complete tree decomposition is exactly equal to the number of points (pixels) in an original signal. Our strategy was to first compute the energy associated within each wavelet packet. Since wavelet packets form orthogonal bases, their decompositions will preserve energy. It is easy to show that

$$\sum_k \left(S^p_{n,k}\right)^2 = \sum_l \left(S^{p-1}_{2n,l}\right)^2 + \sum_l \left(S^{p-1}_{2n+1,l}\right)^2 \quad (12)$$

**Table 2**
Extracted feature for signals belonging to classes $C_1$–$C_3$ shown in Fig. 4, for each level.

| | $C_1$ | $C_2$ | $C_3$ |
|---|---|---|---|
| Feature (level 1) | | | |
| $F_1$ | 0.1335 | 0.1172 | 0.2434 |
| $F_2$ | 0.6001 | 0.2822 | 0.1162 |
| $F_3$ | 0.1564 | 0.1907 | 0.2807 |
| Feature (level 2) | | | |
| $F_1$ | 0.1977 | 0.2131 | 0.3930 |
| $F_2$ | 0.4251 | 0.2681 | 0.1264 |
| $F_3$ | 0.1687 | 0.2322 | 0.2648 |
| Feature (level 3) | | | |
| $F_1$ | 0.3100 | 0.3231 | 0.4338 |
| $F_2$ | 0.6001 | 0.3781 | 0.2031 |
| $F_3$ | 0.1687 | 0.2919 | 0.2553 |
| Feature (level 4) | | | |
| $F_1$ | 0.6321 | 0.3295 | 0.5332 |
| $F_2$ | 0.8115 | 0.6785 | 0.3032 |
| $F_3$ | 0.5236 | 0.5484 | 0.3443 |

Therefore, if we define a energy measure as $E_n^p = \sum_k (S_{n,k}^p)^2$, then

$$E_n^p = E_{2n}^{p-1} + E_{2n+1}^{p-1} \tag{13}$$

Standard deviation of a signal is defined as the square root of the variance. The standard deviation is calculated as,

$$\sigma_j = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_j(n) - \mu_j)^2} \tag{14}$$

The entropy was calculated at each decomposed level of wavelet tree. The entropy, defined by,

$$H(x) = -\sum_k |x_k|^2 \log |x_k|^2 \tag{15}$$

Tables 2 and 3 summarize the value of the three features considered in this paper for the disturbances shown in Fig. 4. The features are energy, entropy and standard deviation denoted by $F_1$, $F_2$, $F_3$.

The normalized energy, entropy and standard deviation are calculated by dividing the energy, entropy and standard deviation by its maximum value respectively. The extracted features can be divided into two different types of class of disturbances. High frequency disturbances (Category-1) consist of oscillatory transient ($C_1$), spike ($C_2$) and multiple notches ($C_3$). The low frequency disturbances (Category-2) consist of voltage rise ($C_4$), voltage dip ($C_5$), voltage rise with harmonics ($C_6$), voltage dip with harmonics ($C_7$) and momentary interruption ($C_8$). The features have been extracted and the extracted features are normalized to classify the nature of power signal disturbances. It can be observed that the entropy value for low frequency disturbances like voltage sag, voltage swell, momentary interruption and the pure undistorted sinusoidal is minimum, harmonics including sag with harmonics and swell with harmonics have a comparatively higher value of entropy while disturbances like voltage flicker and flicker involving harmonics show maximum entropy value in the class of steady state PQ disturbances. Similarly in case of the short duration PQ disturbances, voltage notches and spikes have very low entropy values, while transients have relatively higher values. So this table helps in visual classification of PQ disturbances.

**Table 3**
Extracted feature for signals belonging to classes $C_4$–$C_8$ shown in Fig. 4, for each level.

| | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ |
|---|---|---|---|---|---|
| Feature (level 1) | | | | | |
| $F_1$ | 0.2845 | 0.1880 | 0.1564 | 0.0220 | 0.2651 |
| $F_2$ | 0.3337 | 0.5832 | 0.6001 | 0.5832 | 0.0591 |
| $F_3$ | 0.0135 | 0.2360 | 0.2166 | 0.0557 | 0.9304 |
| Feature (level 2) | | | | | |
| $F_1$ | 0.3086 | 0.2404 | 0.2238 | 0.2701 | 0 |
| $F_2$ | 0.5434 | 0.3832 | 0.6325 | 0.5827 | 0.0522 |
| $F_3$ | 0.0403 | 0.2540 | 0.2229 | 0.0472 | 0.9662 |
| Feature (level 3) | | | | | |
| $F_1$ | 0.5323 | 0.3234 | 0.3278 | 0.3627 | 0 |
| $F_2$ | 0.5211 | 0.4332 | 0.6927 | 0.5832 | 0.0927 |
| $F_3$ | 0.0332 | 0.2894 | 0.3221 | 0.0672 | 0.9682 |
| Feature (level 4) | | | | | |
| $F_1$ | 0.9218 | 0.6593 | 0.5279 | 0.4212 | 0.5641 |
| $F_2$ | 0.5623 | 0.5832 | 0.6529 | 0.6432 | 0.1030 |
| $F_3$ | 0.6312 | 0.3892 | 0.4232 | 0.0923 | 0.9782 |

**Table 4**
Cumulative features for signals belonging to classes $C_1$–$C_4$ shown in Fig. 4.

| Feature | $C_1$ | $C_2$ | $C_3$ | $C_4$ |
|---|---|---|---|---|
| $F_1$ | 1.2733 | 0.9829 | 1.6034 | 2.0472 |
| $F_2$ | 2.4368 | 1.6069 | 0.7489 | 1.9605 |
| $F_3$ | 1.0174 | 1.2632 | 1.1451 | 0.7182 |

The extracted features from wavelet packet domain have been added level wise for classification. The level wise addition was done because LVQ network is based on Kohonen's learning rule and according to Kohonen learning rule not only the winning neuron weights but also the neighborhood neurons are updated. This creates a problem because the normalized input features (data) are very closely spaced so there is a great chance of misclassification as the neighbor class neurons are updated which give rise to number of misclassification. To overcome this, the generated input feature vectors are added level wise for accurate power signal event classification. A long feature vectors were obtained and distinguished feature vectors (hundreds of each disturbance) have been given to the LVQ neural network for power signal disturbance classification. A sample of extracted features from each class is given below in Tables 4 and 5 respectively.

## 5. Learning vector quantization

LVQ procedures are easy to implement and intuitively clear. The classification of data is based on a comparison with a number of so-called prototype vectors. The similarity is frequently measured in terms of Euclidean distance in feature space. Prototypes are determined in a training phase from labeled examples and can be interpreted in a straightforward way as they directly represent typical data in the same space. Prototypes are updated according to their distance from a given example in a sequence of training data. Schemes in which only the winner, i.e. the currently closest prototype is updated have been termed Winner-Takes-All algorithms. The ultimate goal of the training process is, of course, to find a classifier which labels novel data correctly with high probability, after training. This so-called generalization ability will be in the focus of our analysis in the following. Clearly, a better theoretical understanding of the training algorithms should be helpful in improving their performance and in designing efficient schemes. In this work we employ a theoretical framework which makes possible a systematic investigation and comparison of LVQ training procedures.

**Table 5**
Cumulative features for signals belonging to classes $C_4$–$C_7$ shown in Fig. 4.

| Feature | $C_5$ | $C_6$ | $C_7$ | $C_8$ |
|---|---|---|---|---|
| $F_1$ | 1.4111 | 1.2359 | 1.076 | 0.8292 |
| $F_2$ | 1.9828 | 2.5782 | 2.3923 | 0.307 |
| $F_3$ | 1.1686 | 1.1848 | 0.2624 | 3.843 |

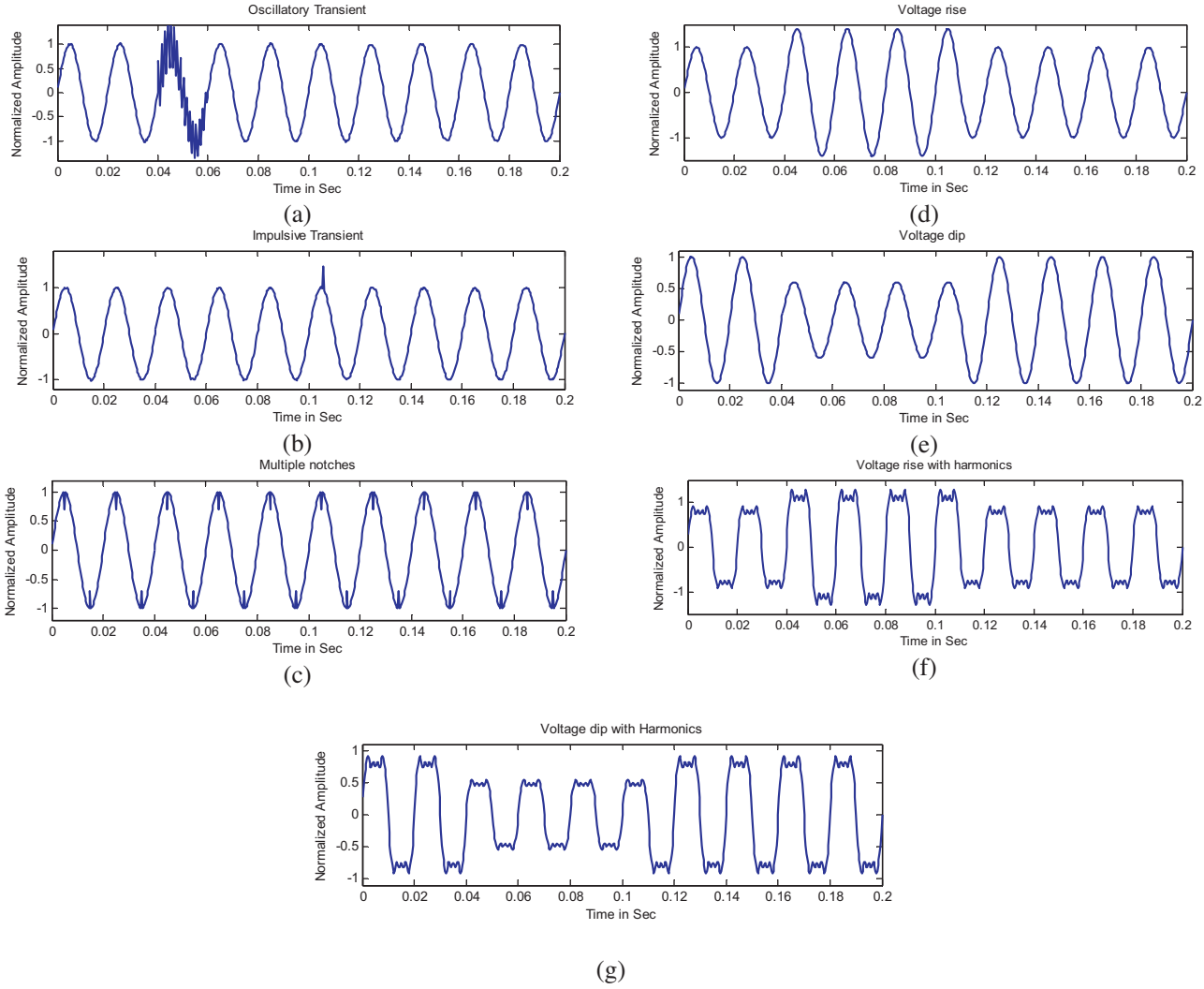



**Fig. 4.** Voltage waveform with (a) oscillatory transient; (b) impulsive transient; (c) multiple notches; (d) voltage rise; (e) voltage dip; (f) voltage rise with harmonics; and (g) voltage dip with harmonics.

We consider training from a sequence of training data which are generated from wavelet packet feature extraction procedure and these features are used as input to LVQ network.

### 5.1. Training LVQ network with learning rule

The learning vector quantization (LVQ) network is a continuous valued, multi-input and multi-output single-layer network, utilizing a combination of unsupervised and supervised training phase. The algorithm of a normal unsupervised training phase, similar to the Kohonen self organized feature map is summarized as follows:

(i) Initialization: at $t = 0$, choose random values for initial $W_k(0)$; $k = 1, 2, \ldots, m$ ($m$: the size of the network), with the magnitudes of the weights within a certain range in accordance with the inputs. The weight vector is given by

$$W_k(t) = [W_{k1}(t), W_{k2}(t), W_{k3}(t), \ldots, W_{kn}(t)]^T, \tag{16}$$

$n$ = dimension of the weight vector.

(ii) Sampling: a sample '$X$' is to be drawn from the date which represents in this case a vector.

(iii) Similarity matching: the best matching (winning) neuron for input $X(t) = [x_1(t), x_2(t), x_3(t), \ldots, x_n(t)]$ is obtained by using Euclidean distance as

$$Y(X) = arg_k \min\{||X - W_i||\} \tag{17}$$

The synaptic weight vectors of all neurons are updated using the update formula as

$$\begin{aligned} W_k(t+1) &= \{W_k(t) + n[x(t) - W_k(t)], \quad k \in C \\ &= \{W_k(t), \quad \text{otherwise} \end{aligned} \tag{18}$$

where $n$ = learning rate and $C$ denotes that winner.

Step (ii) is repeated until no significant change is observed in the weight values. The second training phase uses the LVQ technique to optimize the weights as

$$\begin{aligned} W_c(t+1) &= W_c(t) + \alpha(t)u_i(X)(X - W_c(t)) && \text{if } C_i = C_j \\ &= W_c(t) - \alpha(t)u_i(X)(X - W_c(t)) && \text{if } C_i \neq C_j \\ &= W_i(t) && \text{if } i \neq C \end{aligned} \tag{19}$$

In the above equations, $C$ is again used to denote the winning neuron, $C_i$ is the class assigned to the winner, and $C_j$ is the class the

**Table 6**
Variation in number of epochs.

| No. of classes | No. of competitive neuron | Total no of test input | No of epochs | Learning rate | No of misclas-sification | Efficiency in percentage |
|---|---|---|---|---|---|---|
| 2 | 4 | 15 | 50 | 0.0005 | 9 | 40.00 |
| 3 | 6 | 15 | 100 | 0.0005 | 9 | 40.00 |
| 4 | 8 | 25 | 150 | 0.0005 | 11 | 56.00 |
| 5 | 10 | 25 | 300 | 0.0005 | 10 | 60.00 |
| 6 | 12 | 30 | 350 | 0.0005 | 10 | 66.67 |
| 7 | 14 | 30 | 500 | 0.0005 | 11 | 63.33 |

input $X$ belongs to, $\alpha(t)$ is the learning rate. Further the following values are chosen for $\alpha(t)$ and $u_i(X)$

$$\alpha(t) = \alpha_0 \left(\frac{1-t}{T}\right) \tag{20}$$

$$u_i(t) = \frac{\exp(-||X - W_c||^2/2\sigma^2)}{\sum_{i=1}^{K_{\exp}}(-||X - W_c||^2/2\sigma^2)} \tag{21}$$

The value of $\alpha_0$ and $\sigma$ are chosen as $0 \le \alpha_0 \le 1$, $\sigma > 0$. The similarity measure is used after the final weights are obtained to obtain the distance of the feature vector $X$ to a set of weights as

$$d = (||X - W||) \tag{22}$$

## 6. Training results

The primary factors, which control the behavior of the LVQ network, are the number of hidden units, learning rate and training time. The number of hidden units had a greater effect. Keeping the learning rate fix to 0.0005, number of epochs to 100 and assuming that at least 2 competitive neurons are associated with 1 class, so there are '2*n*' number of competitive neurons if total '*n*' numbers of classes are present. Again if the number of competitive neuron is more then to train the network becomes very difficult. So the number of competitive neuron was taken to be at least 2 for each class. We first started with constant learning rate and different epoch number. The result for learning rate 0.0005, is given below in Table 6. Figs. 5–8 describes the mean square error (performance curve) for different epochs.

It is observed from Table 6 that there is no improvement in efficiencies and the number of misclassification increases due to fluctuation in the performance curve.

Hence the learning rate has a greater impact on the classification accuracy. At a particular learning rate network will show maximum efficiency and minimum misclassification. Keeping number of epochs fixed to 100, learning rate variation starts from 0.0005 to 0.05 for 2, 3, 4, 5, 6 and 7 classes (power signal disturbances) to be classified. If the learning rate is very less it will take more time to converge and the fluctuation in MSE will be more. At the same time if the learning rate is very high the network will be overloaded. There fore only at a particular range of learning rate the network will give maximum efficiency. Table 7 highlights the efficiency in percentage for different learning rate for different power signal classes (disturbances).

## 7. Testing results

Finally to classify all the seven classes (power signal events) with learning rate 0.04–0.055 the following observations are obtained, which is given below in Table 8. The means square error (performance curve) for six and seven number of classes (power signal

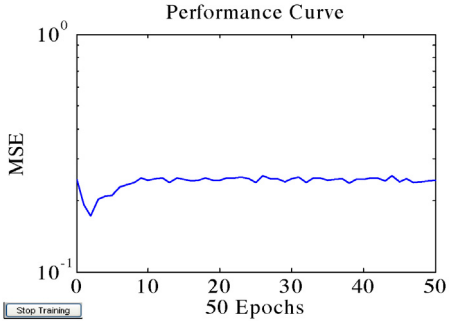



**Fig. 5.** MSE for 50 epochs.

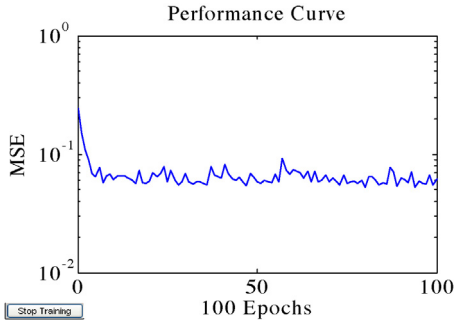



**Fig. 6.** MSE for 100 epochs.

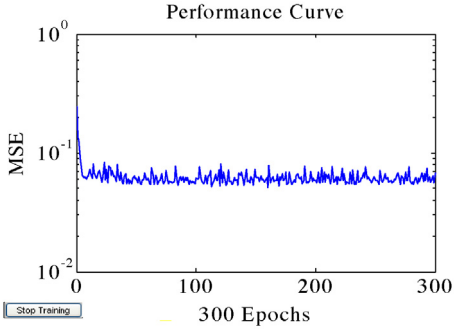



**Fig. 7.** MSE for 300 epochs.

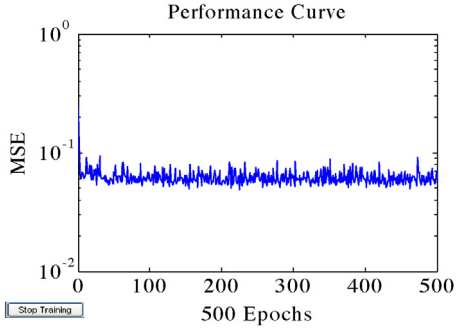



**Fig. 8.** MSE for 500 epochs.

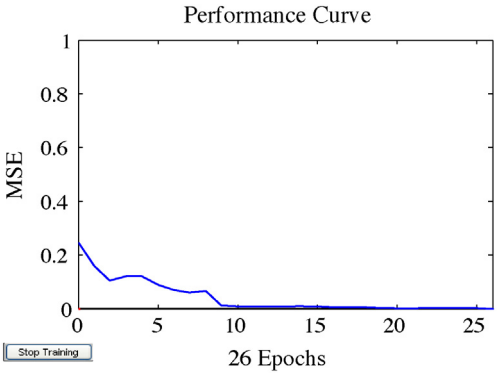



**Fig. 9.** MSE with learning rate 0.005.

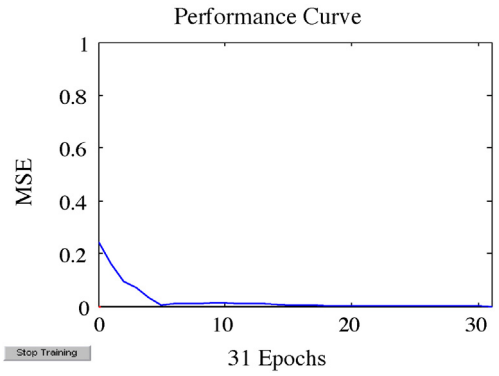



**Fig. 10.** MSE with learing rate 0.0047.

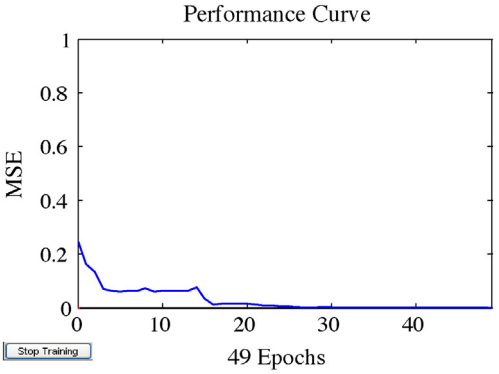



**Fig. 11.** MSE for learning rate 0.05.

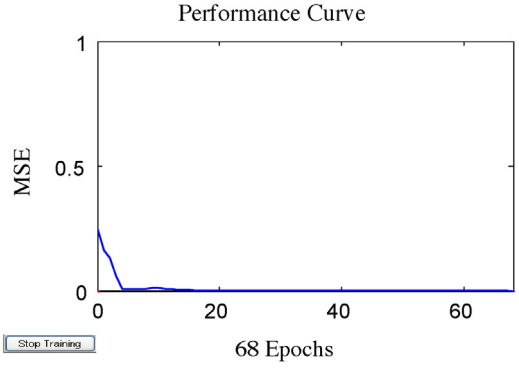



**Fig. 12.** MSE for learning rate 0.046.

**Table 7**
Testing results for slight variation in LR.

| No. of classes | Learning rate | Total no of epochs | No. of epochs after which MSE plot is stable |
|---|---|---|---|
| 6 | 0.005 | 100 | 18 |
| 6 | 0.0049 | 100 | 26 |
| 6 | 0.0047 | 100 | 31 |
| 7 | 0.05 | 100 | 41 |
| 7 | 0.048 | 100 | 49 |
| 7 | 0.047 | 100 | 51 |
| 7 | 0.046 | 100 | 68 |

**Table 8**
The comparison of classification accuracy with other classifiers.

| Proposed by | Overall accuracy % |
|---|---|
| Ghosh and Lubkeman (1995) [20] | 72 |
| Perunicic et al. (1998) [21] | 89 |
| Wijayakulasooriya et al. (2002) [22] | 98 |
| Kaewarsa et al. (2004) [23] | 90.1 |
| G.-S. Hu et al. (2008) [24] | 98.4 |
| S. Ekici (2009) [25] | 95.6 |
| Proposed | 99.14 |

disturbances) is shown below in Figs. 9 and 10 and in Figs. 11 and 12, respectively.

### 7.1. Testing results for classification

The LVQ neural net after training is tested for different power signal classification. In which the test vectors have been given as input to LVQ classifier for each of the seven classes. Some of the classification result with respect to different classes is presented below from Figs. 13–16. The position of the star indicate that the given test vector belongs to oscillatory transient, voltage spike, voltage notch and voltage swell respectively.

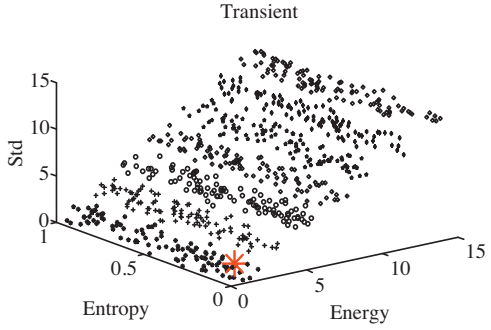



**Fig. 13.** Testing result for transient class.

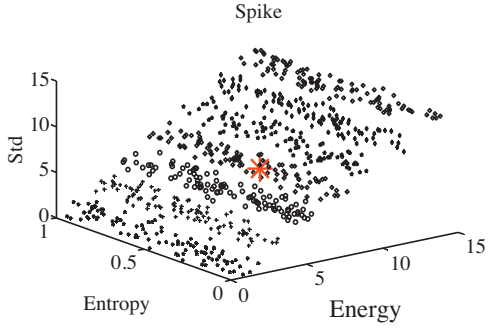



**Fig. 14.** Testing result for spike class.

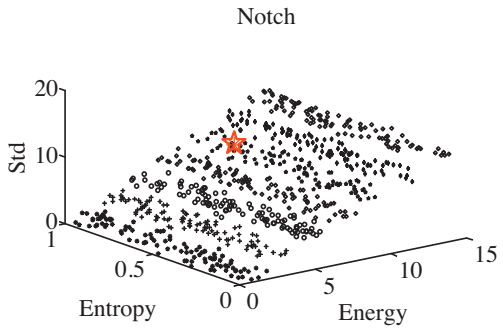



**Fig. 15.** Testing result for notch class.

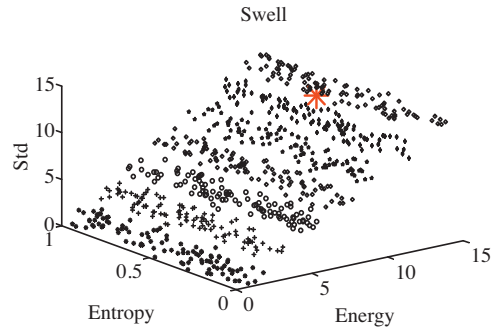



**Fig. 16.** Testing result for swell class.

## 8. Comparison

The proposed classifier was subjected to a test set consisting of 700 patterns with 100 patterns from each type of power quality disturbance.

Each pattern was first passed through wavelet packet decomposition and the features $F_1$, $F_2$ and $F_3$ were extracted. These features were acted as input to the trained LVQ neural network. The overall classification accuracy was obtained as 99.14%. The proposed classifier was compared with the other classifiers in literature in Table 8. As it can be observed, the proposed classifier has a higher accuracy than others.

## 9. Conclusion

In this work, the method to decompose and extraction of features based on wavelet packet transform is studied. During the decomposition process it is found that the detection and localization are more prominent in detail coefficients. Wavelet packet decomposition method is used to extract the features of various power quality disturbance signals and the extracted features are used by the LVQ classifier to classify the different types of disturbances. The concept of local energy of frequency bands based on wavelet packet decomposition and the methods to calculate them are presented, which is useful in non-stationary signal analysis. LVQ type learning models are preferred due to their simple learning rule, their intuitive formulation of a classifier by means of prototypical locations in the data space, and their efficient applicability to any given number of classes. Our rigorous studies along this line show remarkable differences in the generalization ability and convergence properties of such variations of LVQ-neural networks. Forthcoming investigations will concern, for instance, the extension to situations where the model complexity and the structure of the data do not match perfectly. This requires the treatment of scenarios with more than two prototypes and cluster centers. We also intend to study the influence of different variances within the classes and non-spherical clusters, aiming at a more realistic modeling. We are convinced that this line of work has lead to a deeper understanding of LVQ type training and has facilitated the development of efficient algorithms.

## References

[1] C. Loredana, F. Alessandro, S. Simona, A distributed system for electric power quality measurement, IEEE Trans. Instrum. Meas. 51 (4) (2002) 776–781.
[2] Z. Fusheng, G. Zhongxing, G. Yaozhong, FFT algorithm with high accuracy for harmonic analysis in power system, Proc. CSEE 19 (3) (1999) 63–66.
[3] P. Olivier, R. Pascal, M. Michel, Detection and measurement of power quality disturbances using wavelet transform, IEEE Trans. Power Deliv. 15 (3) (2000) 1039–1044.
[4] S. Santoso, E.J. Powers, W.M. Grady, A.C. Parsons, Power quality disturbance waveform recognition using wavelet-based neural classifier. I. Theoretical foundation, IEEE Trans. Power Deliv. 15 (2000) 222–228.
[5] D. Borras, et al., Wavelet and neural structure: a new tool for diagnostic of power system disturbances, IEEE Trans. Ind. Appl. 37 (1) (2001) 184–190.
[6] A.M. Gaouda, M.M.A. Salama, S.H. Kanoun, A.Y. Chikhani, Pattern recognition applications for power system disturbance classification, IEEE Trans. Power Deliv. 17 (3) (2002) 677–682.
[7] Z.L. Gaing, Wavelet-based neural network for power disturbance recognition and classification, IEEE Trans. Power Deliv. 19 (October (4)) (2004) 1560–1568.
[8] M. Wang, A.V. Mamishev, Classification of power quality events using optimal time–frequency representations. Part 1: Theory, IEEE Trans. Power Deliv. 19 (2004) 1488–1495.
[9] Y. Zhen, H. Cheng, Classification of power quality disturbances in noisy conditions, IEE Proc. Gen. Trans. Distrib. 153 (6) (2008) 728–734.
[10] C.-H. Lin, C.-H. Wang, Adaptive wavelet networks for power-quality detection and discrimination in a power system, IEEE Trans. Power Deliv. 21 (July (3)) (2006) 1106–1113.
[11] M. Fligge, S.K. Solanki, Noise reduction in astronomical spectra using wavelet packet, Astron. Astrophys. Suppl. Series 124 (1997) 579–587.
[12] C.V. Lambrecht, M. Karrakchou, Wavelet packets-based high resolution spectral estimation, Signal Process. 47 (1995) 135–144.
[13] J. Chung, E.J. Powers, J. Lamoree, S.C. Bhatt, Power disturbance classifier using a rule based and wavelet-packet based hidden Markov model, IEEE Trans. Power Deliv. 17 (2002) 233–241.
[14] A. Laine, J. Fan, Texture classification by wavelet packet signatures, IEEE Trans. Pattern Anal. Mach. Intell. 15 (November (11)) (1993) 1186–1191.
[15] S. Santoso, E.J. Powers, W.M. Grady, A.C. Parsons, Power quality disturbance waveform recognition using wavelet-based neural classifier. Part 1. Theoretical foundation, IEEE Trans. Power Deliv. 15 (January (1)) (2000) 1347–1348.
[16] S. Santoso, M.W. Grady, J.E. Powers, Power quality disturbance waveform recognition using wavelet-based neural classifier. Part 2: Application, IEEE Trans. Power Deliv. 15 (1) (2000) 229–235.
[17] N.B. Karayiannis, J.C. Bezdek, An integrated approach to fuzzy learning vector quantization and fuzzy c—Means clustering, IEEE Trans. Fuzzy Syst. 5 (November (4)) (1997) 622–628.
[18] N.B. Karayiannis, P.-I. Pai, A family of fuzzy algorithms for learning vector quantization, in: C.H. Dagli, B.R. Fernandez, J. Ghosh, R.T. Kumara (Eds.), Intelligent Engineering Systems through Artificial Neural Networks, vol. 4, ASME Press, New York, 1994, pp. 219–224.
[19] N.B. Karayiannis, J.C. Bezdek, N.R. Pal, R.J. Hathaway, P.-I. Pai, Repairs to GLVQ: a new family of competitive learning schemes, IEEE Trans. Neural Netw. 7 (May) (1996) 1062–1071.
[20] A.K. Ghosh, D.L. Lubkeman, The classification of power system disturbance waveforms using a neural network approach, IEEE Trans. Power Deliv. 10 (1) (1995) 109–115.
[21] B. Perunicic, M. Malini, Z. Wang, Y. Liu, Power quality disturbance detection and classification using wavelets and artificial neural networks, in: Proceedings of the International Conference on Harmonics and Quality of Power, Athens, Greece, 1998, pp. 77–82.
[22] J.V. Wijayakulasooriya, G.A. Putrus, P.D. Minns, Electric power quality disturbance classification using self-adapting artificial neural networks, IEE Proc. Gen. Trans. Distrib. 149 (1) (2002) 98–101.
[23] S. Kaewarsa, K. Attakitmongcol, Wavelet-based neural classification of power quality disturbances, in: Proceeding of International Symposium on Intelligent Signal Processing and Communication Systems, Seoul, Korea, 2004, pp. 299–304.
[24] G.-S. Hu, F.-F. Zhu, Z. Ren, Power quality disturbance identification using wavelet packet energy entropy and weighted support vector machines, Expert Syst. Appl. 35 (2008) 143–149.
[25] S. Ekici, Classification of power system disturbances using support vector machines, Expert Syst. Appl. 36 (2009) 9859–9868.