Power quality disturbances classification based on S-transform and probabilistic neural network

Nantian Huang\textsuperscript{a,b,*}, Dianguo Xu\textsuperscript{a}, Xiaosheng Liu\textsuperscript{a}, Lin Lin\textsuperscript{b}

\textsuperscript{a} Department of Electrical Engineering, Harbin Institute of Technology, Harbin, China
\textsuperscript{b} College of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin, China

Abstract

Classifying power quality (PQ) disturbances is one of the most important issues for power quality control. A novel high-performance classification system based on the S-transform and a probabilistic neural network (PNN) is proposed. The original power quality signals are analysed by the S-transform and processed into a complex matrix named the S-matrix. Eighteen types of time–frequency features are extracted from the S-matrix. Then, after comparing the classification abilities of different feature combinations, a selected subset with 2 features is used as the input vector of the PNN. Finally, the PNN is trained and tested with simulated samples. By reducing the number of features in the PNN's input vector, the new classification system reduces the time required for learning and the computational costs associated with the features and the PNN's memory space. The simulation results show that 8 types of PQ disturbance signals with 2 types of complex disturbances were classified precisely and that the new PNN-based approach more accurately classified PQ disturbances compared to back propagation neural network (BPNN) and radial basis function neural network (RBFNN) approaches.

1. Introduction

With the increasing use of modern equipment and automation in electrical power systems, power quality (PQ) disturbances and their automatic analysis are becoming challenging issues for power engineers [1]. PQ disturbances are identified during monitoring by the exceedance of a defined threshold and are characterised by a set of appropriate parameters, including rms or peak voltage magnitude and duration, for each disturbance event [2]. Typical types of disturbances include voltage sag, swell, interruption, flicker, harmonic, occasional transient and some types of complex disturbances. PQ disturbances interrupt sensitive manufacturing devices and produce serious consequences. To improve PQ, the disturbances should first be classified, but identifying and controlling PQ disturbances with short durations and multiple types are complicated tasks.

Classifying power quality disturbance events is the foundation of PQ disturbance analysis and control. In recent decades, researchers have emphasised the study of classifying PQ disturbances. In addition to classifying accuracy requirements, the process of identifying disturbances must be efficient to satisfy the real-time demands of power system analysis. Further, the effects of noise signals must be removed to satisfy the real environment requirements. The traditional identification process includes 2 steps: feature extraction and pattern recognition.

The features extracted from disturbance signals that distinguish original signal properties and noise immunity serve as the basis for classification systems in practical applications. Time–frequency analysis methods can be used to extract detailed features from both the time and the frequency domains. Time–frequency transform methods commonly used to extract disturbance features include the short time Fourier transform (STFT), discrete wavelet transform (DWT), wavelet packet transform (WPT) and S-transform (ST) methods. The ST [7] is an extension of the wavelet transform and is based on a moving and scalable localising Gaussian window. The ST has anti-noise abilities and excellent time–frequency characteristics. Compared to the STFT, DWT and WPT methods, the ST is superior and can be used to classify PQ signals, especially in high-noise environments.

Neural networks (NNs) are strongly fault tolerant and robust. Processing disturbance signals with neural networks is similar to the process of human thinking and is satisfied for complex disturbance...
The classification systems based on the ST and different neural network structures, such as the radial basis function neural network (RBFNN) and the back propagation neural network (BPNN) [8–10], have been used to identify disturbances. However, problems remain in the existing NN-based systems and affect classification accuracy and efficiency. Typical problems include the setting of initial weights, the number of hidden layer nodes and the efficiency of the learning process. A probabilistic neural network (PNN) is a supervised neural network that has been widely used for pattern recognition [11–14]. PNNs do not require initial weights, a learning process or the modification of weights between the inference vector and the target vector. Based on these characteristics, PNNs have better classification efficiencies than other types of neural networks and can be used to analyse PQ in real time. However, the space complexity of PNNs is greater than that of other neural networks. Mishra et al. [14] used 4 features to classify PQ disturbance signals and compared the classification results of a PNN to those of an FFML network and a learning vector quantisation (LVQ) network. Different neural networks were trained with 25 samples and tested with 100 samples, and the results indicate that the PNN more accurately and efficiently classified the signals compared to the other neural networks. PNNs require only 1 training epoch to construct the neural network. The CPU training and testing times are reduced from 75 s and 0.06 s to 0.9 s and 0.002 s, respectively. Huang et al. [16] distinguished 6 types of disturbances using 3 PQ signal features. However, complex disturbances were not recognised [16]. These authors did not consider the classification ability of other feature combinations [14,16].

To summarise the available literature, a classifier composed of the ST and a PNN can accurately recognise PQ disturbances and is immune to the influence of noise. One disadvantage of PNN-based classifiers is their spatial complexity. To reduce the space requirements of PNN-based classifiers and to improve classification accuracy, the feature combination that has the lowest number of features and acceptable classification ability should be selected from the original feature set extracted from disturbance signals.

In this research, a statistical feature selection method was used to identify the feature combination with the best classification accuracy and the least dimension. After comparing the classification abilities of different feature combinations, a new approach based on the ST and a PNN and requiring only 2 features to construct the PNN’s input vector was developed. In the method, 8 types of disturbance signals, including voltage sag (C1), voltage interruption (C2), voltage swell (C3), voltage flicker (C4), transient (C5), harmonic (C6), harmonic with sag (C7) and harmonic with swell (C8), are first established by a mathematical model. Simulated signals are used for training and verifying the classifier. The appropriate features are then extracted from the ST result. After the classification abilities of feature combinations composed of 18 different features extracted from the ST have been compared, the combination with the best classification ability and composed of 2 features is used as the input vector for the PNN. Finally, the trained classifier based on the PNN is used to recognise PQ disturbances. PNNs have the advantage of fast learning and require only a single-pass network training stage and no iterations to adjust weights. Furthermore, the feature selection step reduces the computational time and spatial costs of the PNN. Simulation results show that 8 types of disturbances were identified accurately by the proposed approach and that the PNN-based classifier more accurately classified disturbances compared to BPNN and RBFNN approaches.

1.1. The S-transform

The S-transform (ST) is an extension of the continuous wavelet transform (CWT). Its frequency-dependent resolution has a direct relationship with the Fourier spectrum. The advantages of the ST arise because the modulating sinusoids are fixed with respect to the time axis, whereas the localising scalable Gaussian window dilates and translates [7].

The ST of a time series $h(t)$ is defined as

$$S(t,f) = \int_{-\infty}^{\infty} h(t)g(t-t,f)e^{-j2\pi ft}dt$$

(1)

where $f$ is the frequency, and $t$ and $t$ denote time.

The Gaussian modulation function $g(t,f)$ is defined as

$$g(t,f) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2}{2\sigma^2}}$$

and $\sigma = \frac{1}{f}$. The final expression is defined as

$$S(t,f) = \int_{-\infty}^{\infty} h(t)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2}{2\sigma^2}}\right)e^{-j2\pi ft}dt$$

(3)

The CWT $W(t,d)$ of a $h(t)$ function is defined as

$$W(t,d) = \int_{-\infty}^{\infty} h(t)\psi(t-d)dt$$

(4)

The S-transform is obtained by multiplying the CWT with a phase factor:

$$S(t,f) = e^{j2\pi ft}W(t,d)$$

(5)

The final form of the continuous S-transform is obtained as

$$S(t,f) = \int_{-\infty}^{\infty} h(t)f\left(\frac{1}{\sigma(f)}\right)e^{-\frac{(t-\tau)^2}{2\sigma^2}}e^{-j2\pi ft}dt$$

(6)

The width of the Gaussian window is $\sigma(f) = T = \frac{1}{f^2}$. The inverse ST is

$$h(t) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} S(t,f)df\right]e^{j2\pi ft}dt$$

(7)

The discrete Fourier transform of the time series $h(t)$ is obtained as

$$H[n] = \frac{1}{N} \sum_{k=0}^{N-1} h(k) e^{-j2\pi nk}$$

(8)

where $n=0,1,\ldots,N-1(N \geq 1)$.

The ST of a discrete time series $h(t)$ is derived by letting $\tau \rightarrow jT$ and $f \rightarrow n/N$:

$$S[jT,n] = \sum_{m=0}^{n-1} \frac{1}{N} H[m+n][G(m,n)e^{j2\pi jm/N}]$$

(9)

where $G(m,n) = e^{-2\pi^2 m^2 n^2/NT}$, $j,m,n = 0,1,\ldots,N-1$.

Based on the DFT, the discrete inverse of the ST is

$$h[k] = \frac{1}{N} \sum_{n=0}^{N-1} \left[\sum_{j=0}^{N-1} S[jT,n]e^{j2\pi nk}\right]$$

(10)

where $j$ and $n=0,1,\ldots,N-1$.

The ST result is a complex matrix named the S-matrix. Each row of the matrix corresponds to a particular frequency, and each column corresponds to a specific sampling point. The elements of the S-matrix are complex amplitude values. The ST-amplitude (STA) matrix is obtained as $STA(jT,f) = |S[jT, n/N]|$.

The ST has better noise immunity than other time–frequency analysis methods, such as the wavelet transform [14], and it precisely extracts PQ signal features without wave filtering. Furthermore, both time–amplitude and frequency–amplitude features can be extracted for disturbance diagnosis. Therefore, the ST can classify PQ disturbances in practical environments.
2. Probabilistic neural network (PNN)

A probabilistic neural network (PNN), proposed by Specht [11], is an outgrowth of the Bayesian classifier with a Parzen window. A typical PNN consists of an input layer, a pattern layer (hidden layer) and a competitive output layer. Unlike other neural networks, PNNs do not require a learning process or initial weights. In PNNs, there is no relationship between the learning process and the recalling process [11–14].

In signal-classification applications, the training examples are classified according to the values of the probabilistic density function (PDF), which is the basic principle of PNNs. A simple PDF is defined as

\[
f_s(x) = \frac{1}{N} \sum_{j=1}^{N_j} \exp\left(-\frac{|x - x_{ij}|^2}{2\sigma^2}\right)
\]

(11)

Modifying and applying (11) to the output vector \( H \) of the hidden layer in the PNN yields

\[
H_h = \exp\left(-\frac{\sum_{k} (x_k - W_{hk}^{by})^2}{2\sigma^2}\right)
\]

(12)

\[
\text{net}_j = \frac{1}{N_j} \sum_{h} W_{hyj} H_h \quad \text{and} \quad N_i = \sum_{h} W_{hyj}
\]

(13)

\[
\text{net}_j = \max(\text{net}_j), \quad \text{then} \quad y_j = 1, \quad \text{else} \quad y_j = 0
\]

(14)

here, \( i \) is the neuron number of the input layers; \( h \) is the neuron number of the hidden layers; \( j \) is the neuron number of the output layers; \( k \) is the number of training examples; \( N \) is the number of classifications (clusters); \( \sigma \) is the smoothing parameter (standard deviation); \( x \) is the input vector; \( W_{hyj} \) is the connection weight between input layer \( x \) and hidden layer \( H \); \( W_{hyj}^{by} \) is the connection weight between hidden layer \( H \) and output layer \( Y \); and \( |x - x_{ij}| \) is the Euclidean distance between vectors \( x \) and \( x_{ij} \), where \( |x - x_{ij}| = \sum_k^s (x_k - x_{ij})^2 \).

Each pattern unit contributes a signal equal to the probability to its associated category unit. The test point is generated by a Gaussian centred on the associated training point. The sum of these local estimates, computed at the corresponding category unit, gives the discriminant as \( \text{net}_j = \max(\text{net}_j) \), which is the Parzen window estimation of the underlying distribution. The \( \max(\text{net}_j) \) operation gives the desired category for the test point.

The architecture of the PNN is described in Fig. 1.

![Fig. 1. Architecture of PNN.](image)

2. PQ disturbance signal analysis and classification

PQ disturbance signals are complicated, and obtaining them is difficult. Researchers normally use software to simulate disturbance signals and to produce test samples [2–10]. In this research, 8 types of PQ disturbance were simulated in Matlab 7.0. The simulated samples were analysed by the ST.

2.1. Simulation of PQ disturbance signals

Eight types of PQ disturbances were simulated in Matlab 7.0. The equations of each disturbance are shown in Table 1. The equations are based on those given in Refs. [8,15]. The standard amplitude \( (A) \) is 1 pu, the fundamental frequency \( (f) \) is 50 Hz and \( u(t) \) is a step function.

2.2. PQ disturbance signal analysis based on S-transform

The ST performs a multi-resolution analysis on a time varying signal as its window width varies inversely with frequency. The transform has a high time resolution at high frequencies and a high frequency resolution at low frequencies.

In this research, the sampling rate was 3.2 kHz, equal to 64 points per circle. In the disturbance analysis process, the harmonic was less than 13 times (650 Hz), and the frequency range of the transient was less than 1300 Hz [15]. According to the

<table>
<thead>
<tr>
<th>Disturbance class</th>
<th>Equation</th>
<th>Equations’ parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal signal</td>
<td>( v(t) = A \cos(\omega t) )</td>
<td>( A = 1 \text{pu}; \quad f = 50 \text{Hz}; \quad \omega = 2\pi f; \quad u(t) = \begin{cases} 1 &amp; t \geq 0 \ 0 &amp; t &lt; 0 \end{cases} )</td>
</tr>
<tr>
<td>Voltage sag</td>
<td>( v(t) = A\left(1 - k(u(t_2) - u(t_1))\right) \cos(\omega t) )</td>
<td>( 0.1 &lt; k &lt; 0.9; \quad T &lt; t_2 - t_1 &lt; 9T )</td>
</tr>
<tr>
<td>Voltage interruption</td>
<td>( v(t) = A\left(1 - k(u(t_2) - u(t_1))\right) \cos(\omega t) )</td>
<td>( 0.9 &lt; k &lt; 1; \quad T \geq t_2 - t_1 &lt; 9T )</td>
</tr>
<tr>
<td>Voltage swell</td>
<td>( v(t) = A\left(1 + k(u(t_2) - u(t_1))\right) \cos(\omega t) )</td>
<td>( 0.1 &lt; k &lt; 0.9; \quad T \leq t_2 - t_1 &lt; 9T )</td>
</tr>
<tr>
<td>Flicker</td>
<td>( v(t) = A\left(1 + 2 \cos(2\pi t + \phi)\right) \cos(\omega t) )</td>
<td>( 0.1 &lt; k &lt; 0.2; \quad 5 \leq \beta \leq 20 )</td>
</tr>
<tr>
<td>Transient</td>
<td>( v(t) = A\left(1 + k(u(t_2) - u(t_1))\right) \cos(\omega t) + \left(k_0 \cos(\omega t_0) + k_1 \cos(10\omega t)\right) )</td>
<td>( k = 0.7; \quad \omega = 0.0015; \quad \omega_0 = 2\pi f_0; \quad 900 \text{Hz} \leq f_0 &lt; 1300 \text{Hz} )</td>
</tr>
<tr>
<td>Harmonic structure</td>
<td>( v(t) = A\left(1 - k(u(t_2) - u(t_1))\right) \cos(\omega t) + h_1 \cos(3\omega t) + h_2 \cos(5\omega t) )</td>
<td>( 0 &lt; h_1 &lt; 0.25; \quad i = 3.5, 7 ); ( \omega = 2\pi f_0; \quad 900 \text{Hz} \leq f_0 &lt; 1300 \text{Hz} )</td>
</tr>
<tr>
<td>Harmonic with sag</td>
<td>( v(t) = A\left(1 - k(u(t_2) - u(t_1))\right) \cos(\omega t) + h_1 \cos(3\omega t) + h_2 \cos(5\omega t) )</td>
<td>( 0.1 &lt; k &lt; 0.9; \quad 0.05 \leq h_1 \leq 0.15; \quad i = 3.5, 7 )</td>
</tr>
<tr>
<td>Harmonic with swell</td>
<td>( v(t) = A\left(1 + k(u(t_2) - u(t_1))\right) \cos(\omega t) + h_1 \cos(3\omega t) + h_2 \cos(5\omega t) )</td>
<td>( 0.1 &lt; k &lt; 0.9; \quad 0.05 \leq h_1 \leq 0.15; \quad i = 3.5, 7 )</td>
</tr>
</tbody>
</table>
Shannon theorem, the sampling rate of 3200 Hz was satisfied for PQ disturbance analysis.

Fig. 2 shows the simulated normal voltage signal with its time–frequency, time–amplitude, frequency–amplitude and frequency-standard deviation plots generated from the S-matrix. Figs. 3–10 present the original disturbance signals and their analysis results based on the ST. From these figures, important magnitude and frequency information can be extracted.

The figures are composed of 5 parts. The first part of Fig. 2 shows a normal voltage signal without any disturbance. The first parts of Figs. 3–10 show the plots of the disturbance signals. The second parts of Figs. 2–10, called time–frequency contours (TF-contours), present frequency values versus time values for the S-matrix. The third parts of Figs. 2–10, called time–maximum amplitude plots (TmA-plots), present maximum amplitudes versus time values. The values in these plots are obtained by searching the columns of the STA at every frequency. In addition, the TmA-plot displays the STA at the fundamental frequency. For a normal voltage signal, as shown in the third part of Fig. 2, the amplitude has a constant value. The fourth parts of Figs. 2–10, called frequency–maximum amplitude plots (FmA-plots), present maximum amplitudes versus normalised frequency values, and the values in these plots are obtained by searching the rows of the STA at every frequency. The plot shows the disturbance’s frequency components and their maximum amplitude. If there is one peak on the FmA-plot, there is a fundamental frequency component in the signal. The fifth parts of Figs. 2–10, called frequency-standard deviation plots (Fstd-plots), show standard deviations versus normalised frequency values, and the values in these plots are obtained by searching the rows of the STA at every frequency.

![Fig. 2. Normal voltage signals and its S-transform analysis.](image)

![Fig. 3. Voltage swell and its S-transform analysis.](image)
Fig. 4. Voltage sag and its S-transform analysis.

Fig. 5. Voltage interruption and its S-transform analysis.

Fig. 6. Voltage flicker and its S-transform analysis.
Comparing Figs. 2 and 3–10 reveals the difference between normal signals and disturbance signals. The ST result of the normal signal with no disturbance, described in the second part of Fig. 2, is mainly in the fundamental frequency area (50 Hz). The TmA-plot and the Fstd-plot do not have peaks in the time and frequency domains. The peak of the FmA-plot appears at the fundamental frequency, and there is no other peak in the plot.

Figs. 3–5 show the features of voltage swell, voltage sag and voltage interruption. The energy is mainly focused in the fundamental frequency area. The TmA-plot shows the time at which the disturbances occurred; the FmA-plot and the Fstd-plot show the frequency at which the disturbances occurred; and the peaks of the 2 plots appear at 50 Hz, denoting the disturbance frequency.

Fig. 6 shows the details of voltage flicker and its ST result. The amplitude of the fundamental frequency changed periodically. The FmA-plot shows that the maximum amplitude of flicker occurred at the fundamental frequency and that the maximum amplitude was larger than the normal signal. The Fstd-plot shows 2 peaks near the fundamental frequency.

Figs. 7–9 show the harmonic, harmonic with swell and harmonic with sag features. Unlike the aforementioned disturbances, the FmA-plot of harmonic has several peaks corresponding to different harmonics. The frequency area of the peaks is between 50 Hz and 650 Hz. The pure harmonic signals have no curve mutation in the TmA-plot, but the complex disturbances have mutation in the TmA-plot, denoting sag or swell. The Fstd-plot has no peak at the fundamental frequency, and the peaks do not denote the harmonic frequency. However, the complex disturbances have peaks at the fundamental frequency, and the peaks of the Fstd-plot denote the harmonic frequencies.

According to Fig. 10, the transient distortion frequency was much higher than that of the other disturbances.
2.3. Feature extraction of power quality disturbances

After summarising the analysis results for different types of disturbance signals, the features were extracted not only from the whole time or frequency area but also from different frequency areas according to the characteristics of the ST results for each type of PQ disturbance signal. The features extracted from different frequency areas describe the characteristics of the original disturbance signals and reduce the computing complexity of the features.

The characteristics of the different disturbances are revealed below.

1. The amplitudes of the voltage sag, voltage swell and voltage interruption change significantly, and the frequency area of amplitude distortion is near the fundamental frequency.
2. The amplitude of flicker changes periodically, and the distortion frequency is near the fundamental frequency.
3. The amplitude of harmonic’s fundamental frequency is stable. The distortion is mainly focused on the high frequency area from 150 Hz to 550 Hz.
4. Two types of complex disturbances have the same harmonic characteristics, but their amplitudes change when voltage sag or swell occurs.
5. Transient distortion occurs in the highest frequency area above 900 Hz.

Therefore, the disturbance feature extraction should obey the following rules.

1. The features should be extracted from time, frequency and amplitude.
2. Features should be extracted from high frequency areas and low frequency areas separately.

In this research, 18 types of features were extracted from the S-matrix and the original signals. The features are described as follows.

Features 1–5 were extracted from the TmA-plot.

Feature 1 (F1): the maximum amplitude of the TmA-plot \( A_{\text{max}} \).
Feature 2 (F2): the minimum amplitude of the TmA-plot \( A_{\text{min}} \).
Feature 3 (F3): the mean of the TmA-plot (Mean).
Feature 4 (F4): the standard deviation of the TmA-plot (STD).
Feature 5 (F5): the amplitude factor \( A_f \), defined as \( A_f = A_{\text{max}} + A_{\text{min}} - 1/2 \) in the range \( 0 < A_f < 1 \).

Features 6–13 were extracted from the FmA-plot.

Feature 6 (F6): the standard deviation of the FmA-plot in the high frequency area above 100 Hz.
Feature 7 (F7): the maximum amplitude of the FmA-plot in the high frequency area \( A_{HF\text{Max}} \).
Feature 8 (F8): the minimum amplitude of the FmA-plot in the high frequency area \( A_{HF\text{Min}} \).
Feature 9 (F9): \( A_{HF\text{Max}} - A_{HF\text{Min}} \).
Feature 10 (F10): the skewness of the high frequency area.
Feature 11 (F11): the kurtosis of the high frequency area.
Feature 12 (F12): the standard deviation of the FmA-plot.
Feature 13 (F13): the mean of the FmA-plot.

Features 14–18 were extracted from the Fstd-plot.

Feature 14 (F14): the mean of the Fstd-plot.
Feature 15 (F15): the standard deviation of the Fstd-plot.
Feature 16 (F16): the standard deviation in the low frequency area, below 100 Hz, of the Fstd-plot.
Feature 17 (F17): the standard deviation in the high frequency area, above 100 Hz, of the Fstd-plot.
Feature 18 (F18): the total harmonic distortion (THD).

The amplitude of a sample point is defined as \( x_i \), where \( 1 \leq i \leq N \), and \( N \) is the number of sample points. The mentioned feature functions are defined as follows.

**Mean:** \( \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \).

**STD:** \( \sigma_{\text{STD}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2} \).

**Skewness:** \( \sigma_{\text{Skewness}} = \frac{1}{(N-1)\sigma_{\text{STD}}} \sum_{i=1}^{N} (x_i - \bar{x})^3 \).

**Kurtosis:** \( \sigma_{\text{Kurtosis}} = \frac{1}{(N-1)\sigma_{\text{STD}}} \sum_{i=1}^{N} (x_i - \bar{x})^4 \).

---

**Fig. 9.** Harmonic with swell and its S-transform analysis.
2.4. Feature selection of power quality disturbances

The 18 features extracted from the ST results describe the characteristics of the original disturbance signals. The type of the disturbance signal can be classified by combinations of these features. However, the computational time and spatial complexity increase with the number of features used. Thus, the dimension of

<table>
<thead>
<tr>
<th>Feature combination</th>
<th>Recognised disturbance type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1</td>
</tr>
<tr>
<td>F1F7</td>
<td>✓</td>
</tr>
<tr>
<td>F3F6</td>
<td>✓</td>
</tr>
<tr>
<td>F3F7</td>
<td>✓</td>
</tr>
<tr>
<td>F3F9</td>
<td>✓</td>
</tr>
<tr>
<td>F3F17</td>
<td>✓</td>
</tr>
<tr>
<td>F5F6</td>
<td>✓</td>
</tr>
<tr>
<td>F5F7</td>
<td>✓</td>
</tr>
<tr>
<td>F5F9</td>
<td>✓</td>
</tr>
<tr>
<td>F5F17</td>
<td>✓</td>
</tr>
<tr>
<td>F5F18</td>
<td>✓</td>
</tr>
</tbody>
</table>

Fig. 10. Voltage transient and its S-transform analysis.

Fig. 11. The scatter plot of F1–F7 combination.

Fig. 12. The scatter plot of F3–F6 combination.

Fig. 13. The scatter plot of F3–F7 combination.
the feature combination, which composes the input vector of the classifier, should be reduced by feature selection.

Feature selection is one of the most active topics in pattern recognition research. Although many effective methods have been designed for automatic feature selection [17,18], the statistical feature selection method is considered to be one of the most efficient and reliable methods.

Based on the previous analysis, the disturbance signals cannot be separated by a single feature [8]. Therefore, we compared the classification abilities of all combinations of 2 features. If a feature combination could not distinguish disturbance signals, we added new features to the feature pair with the best classification ability to obtain new feature combinations.

Fig. 14. The scatter plot of F3–F9 combination.

Fig. 15. The scatter plot of F3–F17 combination.

Fig. 16. The scatter plot of F5–F6 combination.

Fig. 17. The scatter plot of F5–F7 combination.

Fig. 18. The scatter plot of F5–F9 combination.

Fig. 19. The scatter plot of F5–F17 combination.
2-dimensional scatter plots of each feature combination were used to identify the optimum combinations. The scatter plots provided “distinctive regions” for the accurate separation and identification of each disturbance. If a sample set of one type of PQ disturbance signal had no cross with the sample sets of other types of disturbance signals in the 2-dimensional scatter plot of a specific feature combination, this type of disturbance could be classified with the current feature combination. Because 18 features were extracted from the original signals, there were 153 different combinations of 2 features.

To compare the classification abilities of the feature combinations, we simulated 100 samples of each type of disturbance with white noise randomly and used 2-dimensional scatter plots in our analysis. The SNR of the white noise was 50 dB. To identify the pair of features with the best distinguishing ability, the values of the sample’s disturbance parameters, such as distortion duration and amplitude variation, were chosen to be as comprehensive as possible. The combinations that identified more than 4 types of disturbances are shown in Table 2. The disturbance classes, which can be recognised, are labelled by a “/” in the table. Blank cells indicate overlapping disturbances. The 2-dimensional scatter plots of the feature pairs in Table 2 are shown in Figs. 11–20.

Comparing the scatter plots of different feature combinations, the combination of F5 and F6 had the best classification ability. Figs. 11–20 show that these combinations had better ability to separate and classify the disturbance samples.

As shown in Fig. 16, F5 and F6 had a unique ability to separate the disturbance regions. The only 2 disturbances with overlapping samples were voltage sag and interruption. This result was due to the similarity between the definitions of these two disturbances. Compared to the feature pairs that distinguished 6 types of disturbances, including F3–F9, F3–F17, F5–F9 and F5–F18, the samples distinguished by F5 and F6 were more convergent. With sag and swell as one class, disturbances can be precisely recognised by F5 and F6. With the aim to reduce the space complexity of PNN, F5 and F6 are chosen to compose the input vector of PNN. Voltage sag and swell are separated by other features after the classification process of PNN.

2.5. The design of the classifier

In Figs. 16 and 6 types of disturbances are distinguished perfectly by F5 and F6. The 2 types of disturbances with cross-samples are voltage sag and voltage interruption. Recalling the definitions of the two types of disturbances, the only difference is the degree of amplitude change following a distortion. Thus, feature 19 was extracted from the original disturbance samples to describe the degree of amplitude change. The feature is defined as follows.

**Feature 19 (F19):** the minimum of the maximum amplitudes in all cycles of the disturbance sample.

If 0.1 pu < F19 < 0.9 pu, the disturbance type is voltage sag; if 0 < F19 ≤ 0.1 pu, the type is voltage interruption.

The classifier includes 4 steps.

1. The original signals are analysed by the ST, and the features are extracted from the S-matrix.
2. The features are selected. F5 and F6 are used to compose the input vector of the PNN.
3. The disturbance samples are distinguished into 7 types; voltage sag (C1) and interruption (C2) are grouped into one class.
4. The voltage sag and interruption (C1 and C2) are separated by F19. All 8 types of disturbance are recognised.

The structure of the proposed classifier is described in Fig. 21. In the proposed classifier, the neuron number in the PNN input layer is 2; the neuron number in the hidden layer is determined...
by the number of input samples; and the neuron number in the output layer is equal to the number of disturbance types. Because voltage sag and interruption are recognised as one class, the neuron number in the output layer is 7 in this classifier. After sag and interruption are separated from the other types of disturbances, they are distinguished by feature 19 (F19).

3. Simulation and analysis

To improve the classification ability and to meet accuracy requirements following feature selection, the new classifier was constructed and tested for simulated disturbance signals. The classification results of the new PNN-based approach were also compared with those of classifiers based on the back propagation network (BPNN) and the radial basis function neural network (RBFNN).

These 3 classifiers had the same input vector, which was composed of feature 5 (F5) and feature 6 (F6). Thus, the neuron number on the input layer of each type of neural network was 2. The neuron number of the output layer was equal to the class number of the disturbance signals. Because voltage sag and voltage swell were classified as one class, the number of neurons in the output layer was 7 for all 3 classifiers. In the neural network training and testing stages, voltage sag and voltage interruption were treated as the same class. They were separated by feature 19 (F19) in a final step.

The architecture and parameters of the BPNN are described in Table 3. Detailed descriptions of BPNNs can be found in [8,19].

In the RBFNN method, the individual centres of the Gaussian radial-basis functions and their common width are indentified by the k-means clustering algorithm. The least-mean-square (LMS) algorithm is used to estimate the weights of the RBFNN’s output layer. More details about the learning and classification processes can be found in [19–21].

The PNN has 2 neurons on the input layer and 7 neurons on the output layer. The neuron number on the hidden layer is equal to the number of training samples. The spread of radial basis functions is defined by the “trial and error” method. The process of estimating the spread parameter is shown in Table 4. The training and testing data had no noise components.

Each neural network was trained with 25, 50, 75 and 100 input samples of each type without noise, and 100 samples of each class without noise were considered for testing. The testing results are shown in Table 5. The classification accuracy increased as the number of training samples increased, and the PNN was the most accurate of the 3 neural networks. The results also indicate that neural networks are more accurate with more training samples. The PNN met the classification accuracy requirements. In the additional experiments, the training sample number was 100 of each class.

The previous experiments indicate the superiority of the PNN for classifying pure signals. To test the noise immunity of the new approach, 100 samples of each type of disturbance with white noise were simulated in Matlab 7.0. The SNR of the white noise was 50 dB. These samples were used to train the BPNN, the RBFNN and the PNN. After the training process, 100 samples with white noise SNR values of 30 dB, 40 dB and 50 dB were simulated to test the classification ability of each classifier. The classification accuracy of each neural network for the various SNR values is shown in Table 6.

According to Table 6, the classification ability of the new approach was high, even with high noise levels. For varying SNR values, the results of the PNN were better than those of the BPNN and the RBFNN. The new approach is efficient and effective.

4. Conclusion

In this article, a new approach for classifying PQ disturbances was proposed. In this method, the original signals are analysed by the ST, and 18 features are extracted to identify different types of disturbances. After comparing the classification abilities of feature combinations on a scatter plot, F5 and F6 were selected to compose the input vector of the PNN classifier. The PNN classifies the original signals into 7 types, and sag and interruption are recognised as a class. F19 is then used to distinguish sag and interruption. This feature selection not only reduces the computational time and spatial costs of the PNN but also improves the classification efficiency of the proposed classifier.

As a result of the advantages of the probabilistic neural network, the new classifier is efficient and accurate. Results of a comparison test show that the PNN-based classifier was more
accurate than BPNN and RBFNN approaches and that it could be used in high-noise applications.

Acknowledgements

This work was supported by the National Nature Science Foundation of China (nos.50877017 and 60972065) and the Science and Technology Foundation of Department of Education of Jilin Province, China (2011, no. 262).

References


Nantian Huang received the B.S. degree from Harbin Institute of Technology, Harbin, China, in 2003, the M.S. degree from Dalian University of Technology, Dalian, China, in 2006. Now he is currently pursuing the Ph.D. degree in the Department of Electrical Engineering, Harbin Institute of Technology, Harbin, China. He is also a lecturer of the School of Information and Control Engineering, Harbin Institute of Chemical Technology. His main interests and research fields are power quality and intelligent signal processing.

Dianguo Xu received the B.S. degree in control engineering from Harbin Shipbuilding Engineering Institute, Harbin, China, in 1981, and the M.S. and Ph.D. degrees in electrical engineering from Harbin Institute of Technology, Harbin, in 1984 and 1990, respectively. Since 1994, he has been a Professor with the Department of Electrical Engineering, Harbin Institute of Technology. His current research interests include robotics, lighting electronics, power quality mitigation, consumer electronics, power electronics, and motor drives. Dr. Xu is a member of the China Electrotechnical Society and China Power Supply Society.

Xiaosheng Liu received the B.S. degree (in 1988), M.S. degree (in 1993) and Ph.D. degree (in 1999) from Harbin Institute of Technology (HIT), Harbin, China. Now he is a professor in the Department of Electrical Engineering, Harbin Institute of Technology. His main interests and research fields include power line communication and its routing methods, communication networks and control technology, and information and communication of Smart grids.

Lin Lin received the B.S. degree from Lanzhou Jiaotong University, Lanzhou, China, in 2003, the M.S. degree from Dalian University of Technology, Dalian, China, in 2008. Now she is a lecturer of the School of Information and Control Engineering, Jilin Institute of Chemical Technology. Her main interests and research fields are pattern recognition and signal processing.