Investigation of Effective Automatic Recognition Systems of Power-Quality Events

Ameen M. Gargoom, *Student Member, IEEE*, Nesimi Ertugrul, *Member, IEEE*, and Wen. L. Soong

Abstract—There is a need to analyze power-quality (PQ) signals and to extract their distinctive features to take preventative actions in power systems. This paper offers an effective solution to automatically classify PQ signals using Hilbert and Clarke Transforms as new feature extraction techniques. Both techniques accommodate Nearest Neighbor Technique for automatic recognition of PQ events. The Hilbert transform is introduced as single-phase monitoring technique, while with the Clarke Transformation all the three-phases can be monitored simultaneously. The performance of each technique is compared with the most recent techniques (S-Transform and Wavelet Transform) using an extensive number of simulated PQ events that are divided into nine classes. In addition, the paper investigates the optimum selection of number of neighbors to minimize the classification errors in Nearest Neighbor Technique.

Index Terms—Automatic recognition, Clarke transformation, Hilbert transform, power quality (PQ), S-transform, wavelet transform.

I. INTRODUCTION

THE utilizations of the nonlinear loads such as power electronic devices have become a common practice in power systems. Although these devices are manufactured individually according to the associated standards, the wide utilization of such loads can alter the quality of the power supply, which can appear in various forms of events such as voltage sags, swells, and harmonics. If such disturbances are not mitigated, they can lead to failures or malfunctions of various sensitive loads in power systems and may be costly. Therefore, there is a growing need to develop power-quality (PQ) monitoring techniques that can classify the potential sources of disturbances.

Before the implementation of a mitigation technique to reduce or eliminate the disturbances, it is required to monitor two primary indicators: voltage and current signals at the points of interest in the power system. In a monitoring system, such indicators that contain the unique features of power systems are acquired for distinguishing among disturbances. However, for an effective classification of the disturbances, there is a need to avoid the storage of a large number of data that may present in the case of multiple faults and disturbances. Moreover, if a large number of data is present, it is desirable to have a monitoring system that can automatically classify the disturbances. Therefore, there is a need for further improvement of the PQ moni-
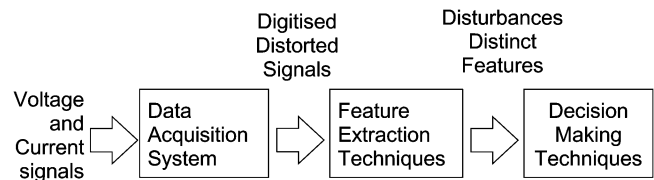
Fig. 1. Principle block diagram of an automatic recognition system.

toring systems, which can offer an effective classification technique and can be efficient in terms of computational resources.

Fig. 1 illustrates the principal diagram of an automatic monitoring process. As shown in the figure, the distorted signals are captured using voltage and/or current signals. This is based on various triggering strategies used by the data acquisition devices. Then, the features that distinguish the captured signals are extracted using a suitable signal-processing technique in the feature extraction block. Finally, the disturbances are separated automatically by applying the decision-making techniques on the relevant features.

This paper begins with a detailed literature review of monitoring and classification of PQ signals in Section II. Both S-Transform and Wavelet Transform are described in Section III together with two novel PQ monitoring approaches introduced in the paper. In Section IV, k-nearest neighbor technique (kNN) for automatic classification is explained. The classification results and the optimum number of neighbors of the kNN technique using selected feature extraction techniques are given in Section V. The paper concludes with the summary of the accuracy of the proposed classifiers.

II. LITERATURE REVIEW ON MONITORING AND CLASSIFICATION OF PQ SIGNALS

In order to extract distinctive features of power disturbances for automatic classification, a number of studies on signal processing techniques were reported in the literature.

Among these, the Fourier transform (FT) technique is commonly used in practice to provide harmonic information about the signals monitored. However, as reported in [1]–[3], FT alone is not sufficient for the feature extraction due to the transient natures of most of the PQ signals where the time information is required.

As an improvement to FT technique, the short time Fourier transform (STFT) technique was implemented in [4], where the voltage disturbances were analyzed in the time-frequency domain. In the paper, the signal studied was split into a set of output signals, which were obtained using bandpass filters centered at the harmonic frequencies. The paper provided a comparison of STFT and Wavelet multiresolution analysis. It was shown that,

once the center of the window function is selected, the center frequencies of the bandpass filters associated with the discrete STFT can be chosen freely. However, it was found out that the proposed technique requires a significant amount of computational resources. This is due to the fact that the proposed STFT needs the implementation of FT technique $(N-L)$ times, where N is the length of the signal and L is the window function length.

The wavelet transform (WT), was shown to be suitable for the PQ analysis, specifically for nonstationary signals. Due to the varying window function of WT at different frequency levels, it has been explored extensively in various studies as an alternative to STFT in [5]–[10], including continuous wavelet transform (CWT) and discrete wavelet transform (DWT) [also known as multiresolution analysis (MRA)]. Most of the previous works on CWT were based on visual detection of PQ events in the time-scale plane [5], [6], which has the capability of handling the noisy signals. However, the main disadvantage of CWT is its redundancy of using large number of scales.

In order to avoid the redundancy in the CWT, the MRA was adapted by many researchers. In [7], wavelet MRA has been applied on three types of PQ events (transformer energizing, converter operation, and capacitor energizing). The features of these events were extracted using the statistics of the squared wavelet coefficient at four decomposition scales. The reason of squaring wavelet coefficient was to reduce the effect of the noises in the signals.

In [8] and [9], the MRA and the Parseval's Theorem have been implemented to extract the disturbances features of the PQ signals in terms of their energies at different resolution levels. In both studies, it was shown that energy values of the decomposed signals at different levels are sensitive to the type of the disturbances. However, these studies did not investigate the phase shift effect of the signal, which has a significant effect on the energy values of the decomposed signals.

It can be concluded here that, WT can be considered as a powerful technique. However, it has limitations in detecting and extracting some PQ signals that have gradient changes such as sags or swells that start at the beginning of the half periods of the measured signal.

S-Transform (ST) was also introduced recently in [3] as a new PQ signal processing technique. In this study, the features of seven simulated signals were extracted by calculating the minima and maxima of the ST absolute matrix. It was shown that the use of such features with Neural Network classifier led to a very high degree of classification (99.28%) for signals contaminated with 40 dB noise (almost pure signals). However, the accuracy is dropped to 75% when the contaminated noise was 20 dB. Similar to the STFT, ST also requires significant amount of computational resources. This is due to the fact that the ST matrix is calculated by performing the inverse Fourier transforms for a number of iterations depending on the frequency resolution.

For automatic recognition of PQ data, the artificial neural network (ANN) technique was utilized in [9] and [10]. The advantage of this technique is its capability to handle easily the noisy data that is present in real-time measurements. However, the main drawback of ANN technique is the need of a large num-

bers of training cycles and the requirement of retraining the entire ANN for every new PQ event, as demonstrated in [11] and [12].

Fuzzy Logic (FL) technique was reported in [3] and [13] for automatic recognition of the PQ events. Although FL technique does not require training process for automatic recognition, where it uses simple "IF-THEN" relation, the technique has some limitation on some PQ events. This is specifically valid for transients and flickers that cannot be described easily by the explicit knowledge of such events.

In [1], k-Nearest Neighbor (kNN) pattern recognition technique was applied to classify the disturbances online. In this study, the kNN classifier demonstrated 88.42% of accuracy for a set of simulated data with 3.5% of noise level, without examining the number of neighbors.

As mentioned previously, the techniques studied so far are based on single-phase measurements. If a three-phase system monitoring is required, it is necessary to monitor each phase separately and then apply a suitable technique to each phase, which results in more computational recourses and poor accuracy. Therefore, this paper considers a new PQ monitoring technique that is introduced for monitoring the three-phase systems simultaneously. In addition, for single-phase systems, Hilbert transform (HT) technique is also examined as an alternative technique, and compared to the previous single-phase based techniques.

Although the kNN technique requires a large capacity of memory to store the training data, due to the recent developments in semiconductor memory technologies, it is a good candidate in automatic recognition of PQ events. Therefore, this paper also considers the kNN technique. The paper utilizes different signal processing techniques at different numbers of nearest neighbors to investigate the effect of number of neighbors on the classification accuracy.

III. FEATURE EXTRACTION TECHNIQUES

To be able to distinguish and classify the different types of disturbances in power systems, it is necessary to perform further processing on the original signals. However, due to the significant number of events and disturbances that may be present in modern AC power systems, it is important to study the unique features of each event and determine the effectiveness of a potential technique(s) in extracting its unique features.

Feature extraction can be defined as a unique process that transforms the raw signals from its original form to a new form so that suitable information can be extracted. The feature extraction step is crucial in an automatic classification system. This is because a classifier can operate reliably only if the features of each event are selected properly. A brief background of the feature extraction techniques that are used in this study will be given.

A. Wavelet Multiresolution Analysis (MRA)

The wavelet MRA is based on decomposition of the target signal, $v[k]$ into different signals at various levels of resolution. This can be achieved by convolving the signal with a low pass filter (also known as the scaling function, ϕ), and a high pass filter (known as the wavelet function, ψ), to obtain a set of

wavelet approximation coefficients, A, and detailed coefficients D. Hence, the target signal, $v[k]$, can be represented in terms of both the scaling and the wavelet functions as [20]

$$v[k] = \sum_n A_0(n)\phi(k-n)$$
$$+ \sum_n \sum_m D_m(n) - 2^{m/2} \cdot \psi(2^m k - n) \quad (1)$$

where A_0 is the approximation coefficient at decomposition level 0, and D_m is the detailed coefficient at level m. The letters n and m are integers and represent the number of samples and the level numbers, respectively.

For automatic classification, it is necessary to define a distinct feature of the signal. The distinguishing feature vector from wavelet MRA is defined in [8] as

$$X[k] = [\,\|D_0\|_2^2 \quad \|D_1\|_2^2 \quad \cdots \quad \|D_{m-1}\|_2^2\,] \quad (2)$$

where $\|D_{m-1}\|_2^2$ is the energy of the wavelet detailed coefficient at the levels $m-1$.

In this paper, the vector X in (2) is represented statistically by determining its standard deviation and its mean value, which will be referred to as the features $F1$ and $F2$, respectively

$$F1 = \frac{1}{M} \sum_{k=0}^{M} X[k] \quad (3)$$

$$F2 = \sqrt{\frac{1}{M} \sum_k \left(X[k] - \frac{1}{M} \sum_{k=0}^{M} X[k] \right)^2} \quad (4)$$

where M is the total number of decomposed levels.

B. S-Transform (ST)

The S-transform is defined by convolving the target waveform, $v[k]$, with a specific window function. In this technique, the window function is chosen to be a function of both time and frequency. This is an advantage over the Wavelet technique since it has a window function (known as Gaussian window) that is a function of the frequency of the signal.

The convolution is performed in the frequency domain by multiplying the Fourier transforms of the analyzed signal and the window function. Then, the inverse Fourier transform is performed to obtain the analyzed signal in the ST domain. For example, for a discrete signal with a Gaussian window, the ST of the signal $v[k]$ is calculated as [3]

$$V_{ST}\left[k, \frac{n}{N}\right] = \sum_{m=0}^{N-1} V\left[\frac{m+n}{N}\right]$$
$$\cdot \exp\left(-\frac{2\pi^2}{n^2}\right) \cdot \exp\left(j\frac{2\pi n}{N}k\right) \quad (5)$$

where k, m, and $n = 0, 1, \ldots, N-1$, $j = \sqrt{-1}$, and $V[(m+n)/N]$ is the Fourier transform of the analyzed signal $v[k]$, and N is the length of the Fourier transform (same as the signal).

As can be seen in (5), the output of the ST is a complex matrix. Its rows represent the frequency content of the signal, and its columns represent the corresponding time.

The distinct feature vector of the signal can be defined using the amplitude versus time vector, $S[k]$, which is calculated from the ST matrix as [3]

$$S[k] = \max\left|V_{ST}\left[k, \frac{n}{N}\right]\right|. \quad (6)$$

The distinguishing features that used with a FL classifier were the minima and maxima of the magnitude-time spectrum obtained from the ST matrix. In this study, the extracted features are the mean value and the standard deviation of the magnitude-time spectrum obtained from the ST matrix. These features are labeled as $F3$ and $F4$ respectively and calculated as

$$F3 = \frac{1}{N} \sum_k S[k] \quad (7)$$

$$F4 = \sqrt{\frac{1}{N} \sum_k \left(S[k] - \frac{1}{N} \sum_k S[k] \right)^2}. \quad (8)$$

C. Clarke Transformation (CT)

The Clarke Transformation is commonly utilized in real-time motor control applications. This is due to the fact that, in a three-phase system, the phase quantities are not independent variables. Therefore, it is possible to transform a three-phase system to an equivalent two-phase representation.

This paper applies this concept to offer an alternative method which is able to process all three phase signals of the power system simultaneously. Then, by using the CT, the three-phase system can be transformed into the two-phase system as follows [14]:

$$\begin{bmatrix} v_\alpha[k] \\ v_\beta[k] \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & \cos(\gamma) & \cos(2\gamma) \\ 0 & \sin(\gamma) & \sin(2\gamma) \end{bmatrix} \begin{bmatrix} v_a[k] \\ v_b[k] \\ v_c[k] \end{bmatrix} \quad (9)$$

where $v_a[k]$, $v_b[k]$, and $v_c[k]$ are the three-phase voltages (or currents) at instant k, $\gamma = 2\pi/3$, and $v_\alpha[k]$ and $v_\beta[k]$ are known as the two-phase Clarke's component vectors.

Similar to the previous discussions, a unique feature can be obtained using a space vector, S_V that is defined as

$$S_V = [\nu_\alpha[k] + j\nu_\beta[k]] \quad (10)$$

where $j = \sqrt{-1}$.

The magnitude of S_V is given by

$$|S_V[k]| = \left[\sqrt{(v_\alpha[k])^2 + (v_\beta[k])^2}\right]. \quad (11)$$

The primary advantages of this approach are its capability to analyze all three phases of the power system simultaneously and its simplicity. Therefore, this technique can offer faster processing time than the previous techniques while it can analyze all three phases of the system simultaneously.

In the proposed CT technique, this paper describes two features $F5$ and $F6$, which are utilized for the feature extraction.

The features $F5$ and $F6$ are the mean value and standard deviation of the magnitude of the space vector, respectively

$$F5 = \frac{1}{N} \sum_{k=1}^{N} |S_V[k]| \qquad (12)$$

$$F6 = \sqrt{\frac{1}{N-1} \sum_{k=1}^{N} \left(S_V[k] - \frac{1}{N} \sum_{k=1}^{N} |S_V[k]| \right)^2} \qquad (13)$$

where N represents the length of the three-phase signals.

D. Discrete Hilbert Transform (DHT)

The Discrete Hilbert Transform is a mathematical process used to generate a complex signals from real signals. It is obtained by convolving the real signal $x(t)$ with the function $(1/\pi t)$, as given in [15]

$$x_H(t) = x(t) * \frac{1}{\pi \cdot t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\lambda)}{t - \lambda} d\lambda. \qquad (14)$$

Since the output of the DHT is, in fact, a 90° phase—shifted version of the original signal $x(t)$, a complex signal (also known as analytical signal) that is associated with the original signal can be constructed as

$$x_c[k] = x[k] + j x_H[t]. \qquad (15)$$

The envelope of the original signal is then can be defined as

$$|x_A[k]| = \sqrt{x^2[k] + x_H^2[k]}. \qquad (16)$$

The features that were examined from the DHT are the mean value $(F7)$ and the standard deviation $(F8)$ of the vector $x_A[k]$

$$F7 = \frac{1}{M} \sum_{m} x_A[k] \qquad (17)$$

$$F8 = \sqrt{\frac{1}{M} \sum_{m} \left(x_A[k] - \frac{1}{M} \sum_{m} x_A[k] \right)^2} \qquad (18)$$

where M is the length of the target signal.

Table I summarizes the features defined for each transformation as described above.

IV. Nearest Neighbors Technique for Data Clustering

The main advantage of the nearest neighbor (NN) technique is that it is nonparametric technique which is simple, yet, effective in many cases [16], where no prior statistical knowledge about the data is required.

The classification decision in the NN technique is made based on a set of training features that represent all of the expected classes of the PQ events. In this technique, any unknown pattern, x, is compared with all the cases of the training features of data. This is done by calculating the distances between the unknown

TABLE I
EXTRACTED FEATURES USING DIFFERENT SIGNAL PROCESSING TECHNIQUES

Feature	Descriptions
F1	Mean value of norms of the WT detailed coefficient at different levels
F2	Standard deviation of norms of the WT detailed coefficient at different levels
F3	Mean value of the magnitude-time spectrum obtained from ST matrix
F4	Standard deviation of the magnitude-time spectrum obtained from ST matrix
F5	Mean value of the space vector of CT
F6	Standard deviation of the space vector of CT
F7	Mean value of the signal envelope using DHT
F8	Standard deviation of signal envelope using DHT

pattern and all other features, and then is assigned to a class with a minimum distance [17]

$$x \in W_i, \quad \mathbf{iff} \, d(t_j, x) = \min\left(d(t_j, x)\right) \qquad (19)$$

where W_i is the pattern class, the operator **iff** is a mathematical abbreviation for "if and only if," t_j is a training data sample, and d is the distance between the unknown pattern x, and t_j. The distance d is calculated as

$$d(t_j, x) = \sqrt{\sum_{j}^{N} (t_j - x)^2} \qquad (20)$$

where N is the total number of samples.

The k-Nearest Neighbor (kNN) technique is similar to NN technique described above, except that the algorithm finds the first k minimum distances d, and the decision is made based on the class which has the majority minimum distances. However, to apply the kNN, an appropriate number of neighbors should be chosen since the accuracy of the classification depends on this number.

To demonstrate the performance of the classifier, so called "confusion matrix" can be constructed using a set of test data such as

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

where the diagonal entry a_{ii} of the confusion matrix denotes the number of elements from the test data whose true class is "i" and n is the total number of classes [18].

V. Results

The accuracy of the selected features has been investigated using an extensive number of simulated PQ signals by applying the kNN technique. However, to ensure the highest accuracy, the kNN classifier has been applied by changing the number of neighbors from 1 to 50. Therefore, the results given below
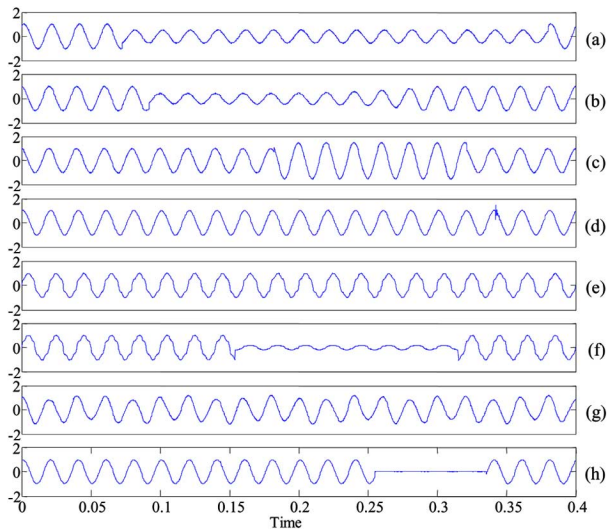
Fig. 2. Sample voltage waveforms (per-unit) that illustrate the time-domain characteristics of (a) sag, (b) motor sag, (c) swell, (d) switching transient, (e) harmonics, (f) sag with harmonics, (g) flicker, and (h) interruption.

Fig. 3. Simulated random changes for the sag parameters, which consist of 500 sag events.

demonstrate the optimal numbers of neighbors for a given features to minimize the classification error.

As will be described below, a total of 4500 events were considered in this study, which were obtained from nine different types of events, each containing 500 different cases.

A. Simulation of PQ Events

The PQ events studied in this paper have been classified into the following nine classes: sags (dips), motor sag, swells, switching transients, harmonic distortion, sags with harmonics, flicker and interruption, and pure signal. A set of sample voltage waveforms given in Fig. 2 demonstrates the characteristics of the first eight PQ events as a function of time.

Although the figures given above illustrate a typical profile of each event, in the practical power systems, the characteristics of each event may vary significantly. Therefore, to be able to study the effects of such variations, 500 different cases were generated using the simulated event data. This is performed for each class of events by randomly changing various parameters. The parameters used to vary the classes of events are the depth, the angle, the starting time and the duration of the events, which are defined below.

The depth of the event is defined as the change in the amplitude of a signal. The angle represents the phase shift at which the signal is captured. The starting time is the time at which the event starts. The duration is the time period of the event.

In this study the above described parameters were varied according to the IEEE recommended practice in [19]. For example, for the sag events, the four parameters were varied randomly as illustrated in Fig. 3.

As can be seen in the figure, the depth of the sag is varied from 10% to 90% of the magnitude of the pure sine waveform. The angle of the signal is varied from 0% to 100% of the entire period (which is a realistic assumption since the captured waveforms in a practical monitoring system could have a phase shift that
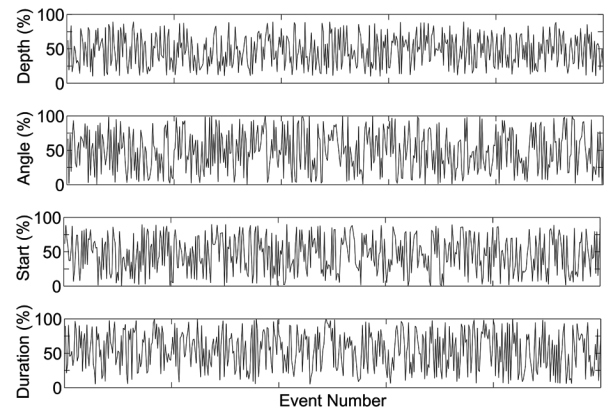
may vary from 0 to 2π). In addition, the starting time of the sag is varied from 0% to 80% of the total length of the signal. Moreover, the duration of the sag is varied from 5% to 100% of the total length of the waveform. In this study, each event accommodates 20 cycles of the captured signal. Therefore, a duration of 5% corresponds to one cycle of the waveform.

For the motor sag events, the recovery pattern of the sag was simulated using an exponential function of $A \exp(k^4)$ where A is a constant representing the depth of the motor sag, and k is an index that varies from 0 up to the length of the sag. The sag duration is assumed to be equal or greater than 10 cycles (which is a realistic value for industrial motor drives), and the depth of the sag was varied from 90% to 70%.

For the interruption and the swell events, the four parameters (depth, angle, starting time and duration) were also varied as described above. However, the magnitude of the pure sine waveform was increased between 10% and 90%.

For the harmonics events, 2nd, 3rd, 5th, 7th, 9th, and 11th harmonics are used to randomly contaminate the ideal waveforms. During the generation of such events, the total harmonic distortion (THD) of the waveform was kept greater than 5%, as suggested in [19]. If the THD of a waveform was less than 5%, it was regarded as a pure sine waveform.

In the case of the flicker events, the amplitude of the simulated signals was changed periodically to introduce the effect of a flicker. To achieve this, the magnitude of the target waveform was varied as a function of another sine wave. This results an oscillation in the amplitude of the target waveform, which varied randomly from 50% to 70% of the fundamental frequency.

The parameters which were changed randomly in the case of transient events are the oscillation frequency of the transients (varied from 10 times to 15 times of the fundamental frequency) and the amplitude of the overshoot (varied up to 150% of the amplitude of the pure signal).

B. Minimization of Classification Errors

In this paper, the kNN wavelet transform classifier utilized a total of 4500 events for the training purposes. In addition, the same number of events was generated randomly for testing the classifier. The performance of the classifier using the selected
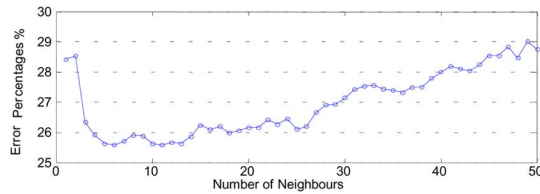
Fig. 4. Average classification errors of kNN-WT (%) using the features $F1$ and $F2$.
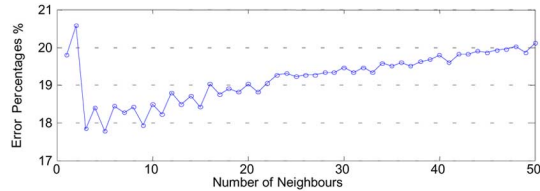

Fig. 5. Average classification errors of kNN-ST (%) using the features $F3$ and $F4$.
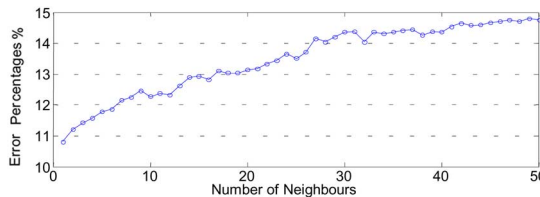

Fig. 6. Average classification errors of kNN-DHT (%) using the features $F7$ and $F8$.
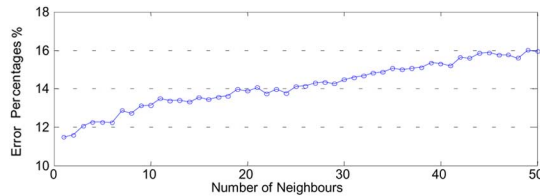

Fig. 7. Average classification errors of kNN-CT (%) using the features $F5$ and $F6$.

TABLE II
CONFUSION MATRIX OF kNN—WT CLASSIFIER WITH 5 NEIGHBORS

	Pure	Sag	Motor Sag	Swell	Harm-onics	Flicker	Trans-ient	Sag + Harm	Interr-uption
Pure	**478**	0	0	0	1	0	21	0	0
Sag	1	**190**	107	0	0	0	1	131	70
Motor Sag	0	82	**390**	0	0	0	0	27	1
Swell	0	0	0	**472**	0	28	0	0	0
Harm-onics	12	0	0	0	**459**	0	29	0	0
Flicker	0	0	0	4	0	**496**	0	0	0
Trans-ient	34	0	0	0	66	0	**400**	0	0
Sag + Harm	0	119	89	0	0	0	4	**200**	88
Interr-uption	0	88	40	0	0	0	0	111	**261**

TABLE III
CONFUSION MATRIX OF kNN—WT CLASSIFIER WITH 5 NEIGHBORS

	Pure	Sag	Motor Sag	Swell	Harm-onics	Flicker	Trans-ient	Sag + Harm	Interr-uption
Pure	**486**	2	0	12	0	0	0	0	0
Sag	16	**235**	3	159	0	1	18	11	57
Motor Sag	0	0	**456**	0	0	44	0	0	0
Swell	34	139	10	**308**	0	5	0	0	4
Harm-onics	0	0	0	0	**476**	0	0	24	0
Flicker	0	0	41	0	0	**459**	0	0	0
Trans-ient	0	11	0	0	0	0	**484**	0	5
Sag + Harm	0	8	0	0	80	0	0	**371**	41
Interr-uption	0	32	0	7	1	0	26	9	**425**

combinations of the features, as summarized in Table I, were given in Figs. 4–7.

Fig. 4 shows the average of classification errors for all the events using the kNN technique, where the features $F1$ and $F2$ were included and plotted from 1 to 50 neighbors. The results indicate that a minimum error of 25.58% is obtained when the classification decision was made based on six neighbors. However, it is known that an even number of neighbors may not indicate a confident classification. Therefore, the best number of neighbors in this case should be considered as five, which also indicates an error around the similar value 25.64%.

The confusion matrix at the best number of neighbors (five neighbors) obtained above is shown in Table II. In the table, the column headings represent the true classes of the PQ signals, and the raw headings represent the classification results of the classes. For example, the number 478 in the first column and raw indicates that among the 500 cases of pure sine wave signals, the kNN classifier can classify 478 cases correctly.

Although the classification error may be considered high in Fig. 4, as indicated above, this is an average error including

all the events, which could be reduced significantly, if similar events, such as sag, motor sag, and sag with harmonics, are excluded in the calculations of the classifier average error.

Fig. 5 is given to illustrate the classification error of the kNN classifier when the features of $F3$ and $F4$ are used from ST matrix, which is also plotted for different number of neighbors. In this case, a minimum average classification error of 17.78% occurred when the number of neighbors was five.

Similarly, the corresponding confusion matrix results were given in Table III for the classifier with five neighbors. Although the average error is still considered high, it can be noted that the kNN-ST classifier has a better classification results than the previous kNN-WT classifier. For example, in this classifier, the classification of the sags due to faults, the motor sags and the sags with harmonics have been improved significantly, while the classification error of the swell event has increased.

As stated previously, this paper offers two new techniques to improve the classification of the events combining the kNN technique with DHT and CT, which demonstrated further improvements on the average classification error for the events considered in the study.

TABLE IV
CONFUSION MATRIX OF kNN—DHT CLASSIFIER FOR $K = 1$

	Pure	Sag	Motor Sag	Swell	Harm-onics	Flicker	Trans-ient	Sag + Harm	Interr-uption
Pure	**500**	0	0	0	0	0	0	0	0
Sag	0	**264**	15	0	0	0	0	214	7
Motor Sag	0	6	**494**	0	0	0	0	0	0
Swell	0	0	0	**500**	0	0	0	0	0
Harm-onics	0	0	0	0	**494**	0	6	0	0
Flicker	0	0	0	0	0	**500**	0	0	0
Trans-ient	0	0	0	0	6	0	**494**	0	0
Sag + Harm	0	191	9	0	0	0	0	**274**	26
Interr-uption	0	0	0	0	0	0	0	6	**494**

TABLE V
CONFUSION MATRIX OF kNN—CT CLASSIFIER FOR $k = 1$

	Pure	Sag	Motor Sag	Swell	Harm-onics	Flicker	Trans-ient	Sag + Harm	Interr-uption
Pure	**500**	0	0	0	0	0	0	0	0
Sag	0	**250**	11	0	0	0	0	239	0
Motor Sag	0	1	**499**	0	0	0	0	0	0
Swell	0	0	0	**485**	15	0	0	0	0
Harm-onics	0	0	0	2	**498**	0	0	0	0
Flicker	0	0	0	0	0	**500**	0	0	0
Trans-ient	0	0	0	0	0	0	**500**	0	0
Sag + Harm	0	231	12	0	0	0	0	**252**	5
Interr-uption	0	0	0	0	0	0	0	1	**499**

The average classification error of the kNN-DHT at different number of neighbors is given in Fig. 6. As can be seen in the figure, the minimum average classification error is 10.8% while the decision is based on one neighbor only.

The results of the confusion matrix for 1 neighbor are shown in Table IV. It can be concluded in this table that most of the classification errors in the kNN-DHT are due to the sag only events and the sag with harmonics, while the errors in classifying the motor sag events were slightly increased. Moreover, the error in the swell events was reduced considerably compared to the kNN-ST classifier.

The average classification error of kNN-CT (with features $F5$ and $F6$) versus number of neighbor is shown in Fig. 7. The minimum average error in this classifier is 11.5% (for one neighbor). Similarly, the confusion matrix at the best selection of the number of neighbors (1 neighbor) is given in Table V.

Although the average error in this classifier is slightly higher than the previous kNN-DHT classifier (primarily due to misclassifying the sags and the sag with harmonics events), a better classification for the motor sags events were achieved.

TABLE VI
kNN PERFORMANCE WITH DIFFERENT FEATURE EXTRACTION TECHNIQUES

Class	Classification accuracy in %			
	kNN-WT	kNN-ST	kNN-HT	kNN-CT
Pure	95.6	97.2	**100**	**100**
Sag	38	47	**52.8**	50
Motor Sag	78	91.2	98.8	**99.8**
Swell	94.4	61.6	**100**	97
Harmonics	91.8	95.2	98.8	**99.6**
Flicker	99.2	91.8	**100**	**100**
Transient	80	96.8	98.8	**100**
Sag with Harmonics	40	**74.2**	54.8	50.4
Interruption	52.2	85	98.8	**99.8**
Number of neighbors	5	5	1	1

VI. CONCLUSION

The paper presented a detailed literature review of monitoring and classification of PQ signals. A brief description of the S-Transform and Wavelet Transform were described together with two novel PQ monitoring approaches. It was concluded that the kNN based techniques can offer an effective automatic classification of PQ events. However, it is required to identify the best number of neighbors in the classification of the events.

In the paper, two new feature extraction techniques were examined. The kNN-DHT method was developed for the single phase analysis of the power systems, and the kNN-CT was offered for three phase systems. The performance of these two techniques were also compared with the Wavelet and ST techniques, using k-Nearest Neighbor pattern recognition technique.

As demonstrated, to ensure a higher accuracy in each classifier, the number of neighbors was varied. This provided a pattern where a clear conclusion could be drawn based on increasing or decreasing profile of the average error. The results were given between 1 to 50 neighbors. It was observed that the above 50 neighbors, the profile of the average error followed the same pattern.

In the paper, a total of 4500 events were considered, which were obtained from nine different types of events, each containing 500 different cases. In the study, the motor sags was introduced as an event that had not been studied previously.

It was demonstrated that the proposed two new techniques (kNN-DHT and kNN-CT) were found to be very effective classifiers compared with the kNN S-Transform and kNN Wavelet transform, except in the case of the sag with harmonic events, where the kNN S Transform offers better solution.

The classification accuracy of each individual class of event using the techniques studied are summarized in Table VI. For example, 100% of accuracy in the table indicates that the confidence about the classification of an event using a given technique is 100%. The table also includes the best number neighbor for each classifier, which is always less than 5. The numbers in bold, in the table, indicate the highest accuracy that was achieved for a given event and a classifier. As it was expected, the classification accuracy of sag only events and

sag with harmonics were low, which due to the similarity of these two events. However, a significant improvement have been achieved with the two proposed (kNN-HT and kNN-CT) in distinguishing the sag only events as compared to the other classifiers. As it can be seen, while for the sag with harmonics events the kNN-ST has performed better than others, its performance with swell events was the lowest. Therefore, as an overall performance, the two new proposed classifiers, for single phase and three phase systems, have provided improvements in classifying the PQ events considered, except the sag with harmonic events.

REFERENCES

[1] A. M. Gaouda, S. H. Kanoun, and M. M. A. Salama, "On-line disturbance classification using nearest neighbor rule," *Int. J. Elect. Power Syst. Res.*, vol. 57, no. 1, pp. 1–8, 2001.
[2] R. A. Flores, "State of the art in the classification of power quality events, an overview," in *Proc. 10th Int. Conf. Harmonics Quality of Power*, 2002, vol. 1, pp. 17–20.
[3] M. V. Chilukuri and P. K. Dash, "Multiresolution S-transform-based fuzzy recognition system for power quality events," *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 323–330, Jan. 2004.
[4] Y. H. Gu and M. H. J. Bollen, "Time-frequency and time-scale domain analysis of voltage disturbances," *IEEE Trans. Power Del.*, vol. 15, no. 4, pp. 1279–1284, Oct. 2000.
[5] F. Jurado, N. Acero, and B. Ogayar, "Application of signal processing tools for power quality analysis," in *Proc. Canadian Conf. Electrical and Computer Engineering*, May 2002, vol. 1, pp. 82–87.
[6] O. Poisson, P. Rioual, and M. Meunier, "Detection and measurement of power quality disturbances using wavelet transform," *IEEE Trans. Power Del.*, vol. 15, no. 3, pp. 1039–1044, Jul. 2000.
[7] S. Santoso, W. M. Grady, E. J. Powers, J. Lamoree, and S. C. Bhatt, "Characterization of distribution power quality events with Fourier and wavelet transforms," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 247–254, Jan. 2000.
[8] A. M. Gaouda, S. H. Kanoun, M. M. A. Salama, and A. Y. Chikhani, "Pattern recognition applications for power system disturbance classification," *IEEE Trans. Power Del.*, vol. 17, no. 3, pp. 677–683, Jul. 2002.
[9] Z.-L. Gaing, "Wavelet-based neural network for power disturbance recognition and classification," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1560–1568, Oct. 2004.
[10] S. Santoso, E. J. Powers, W. M. Grady, and A. C. Parsons, "Power quality disturbance waveform recognition using wavelet-based neural classifier- part I: Theoretical foundation," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 222–228, Jan. 2000.
[11] J. Chung, E. J. Powers, W. M. Grady, and S. C. Bhatt, "Power disturbance classifier using a rule-based method and wavelet packet-based hidden Markov model," *IEEE Trans. Power Del.*, vol. 17, no. 1, pp. 233–241, Jan. 2002.
[12] A. M. Youssef, T. K. Abdel-Galil, E. F. El-Saadany, and M. M. A. Salama, "Disturbance classification utilizing dynamic time warping classifier," *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 272–278, Jan. 2004.
[13] W. R. A. Ibrahim and M. M. Morcos, "An adaptive fuzzy technique for learning power-quality signature waveforms," *IEEE Power Eng. Rev.*, vol. 21, no. 1, pp. 56–58, Jan. 2001.
[14] E. Clarke, *Circuit Analysis of AC Power Systems*. New York: Wiley, 1958, vol. 1.
[15] A. D. Poularikas, *The Transforms and Applications Handbook*. Boca Raton, FL: CRC, 2000, pp. 4–10.
[16] H. C. Entorf, F. J. Garijo, and J. C. Riquelme, *Advances in Artificial Intelligence*. Berlin, Germany: Springer-Verlag, 2002, p. 23.
[17] E. Micheli-Tzanakou and E. M. Tzanakou, *Supervised and Unsupervised Pattern Recognition*. Boca Raton, FL: CRC, 2000, pp. 18–18.
[18] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. New York: Wiley, 2004, pp. 10–10.
[19] *IEEE Recommended Practice for Monitoring Electric Power Quality*, IEEE/Std. 1159-1995, 1995.
[20] A. M. Gaouda, M. M. A. Salama, M. R. Sultan, and A. Y. Chikhani, "Application of multiresolution signal decomposition for monitoring short-duration variations in distribution systems," *IEEE Trans. Power Del.*, vol. 15, no. 2, pp. 478–485, Apr. 2000.

Ameen Gargoom (S'05) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the University of Garyounis, Benghazi, Libya, in 1994 and 2001, respectively, and is currently pursuing the Ph.D. degree at the University of Adelaide, Adelaide, Australia.

From 1995 to 2001, he was a Consultant Engineer with AL-Emara Co., Benghazi. He was then an Associate Lecturer with the University of Garyounis. His main research interests include power-quality monitoring and the application of signal processing technique to power systems.

Nesimi Ertugrul (M'95) received the B.Sc. and M.Sc. degrees in electrical and in electronic and communication engineering from the Istanbul Technical University, Istanbul, Turkey, in 1985 and 1989, respectively, and the Ph.D. degree from the University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, U.K., in 1993.

He has been with University of Adelaide, Adelaide, Australia, since 1994, where he is a Senior Lecturer. His primary research topics include sensorless operation of switched motors, fault-tolerant motor drives, condition monitoring, and electric vehicles. He is the author of a book *LabVIEW for Electric Circuits, Machines, Drives and Laboratories* (Prentice-Hall, 2002). He serves on the editorial advisory board for the *International Journal of Engineering Education*.

Wen L. Soong was born in Kuala Lumpur, Malaysia. He received the B.Eng. degree from the University of Adelaide, Adelaide, Australia, in 1989, and the Ph.D. degree from the University of Glasgow, Glasgow, U.K., in 1993.

He was an Electrical Engineer for four years in the Power Controls Program General Electric Corporate Research and Development, Schenectady, NY, before taking a teaching position with the Electrical and Electronic Engineering Department, University of Adelaide, in 1998. His research interests include permanent magnet and reluctance machines, renewable energy generation, magnetic levitation, and condition monitoring and diagnostics.