International Journal of Innovative
Computing, Information and Control
Volume 8, Number 3(A), March 2012

ICIC International ©2012 ISSN 1349-4198
pp. 1523–1539

PSO BASED KERNEL PRINCIPAL COMPONENT ANALYSIS AND MULTI-CLASS SUPPORT VECTOR MACHINE FOR POWER QUALITY PROBLEM CLASSIFICATION

Jonglak Pahasa[1] and Issarachai Ngamroo[2]

[1]Department of Electrical Engineering
School of Engineering
University of Phayao
Muang, Phayao 56000, Thailand
j_pha@hotmail.com

[2]Department of Electrical Engineering
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
Chalongkrung Rd., Ladkrabang, Bangkok 10520, Thailand
ngamroo@gmail.com

Received November 2010; revised April 2011

Abstract. *Electric power quality (PQ) problems are very important aspects due to the increase in the number of loads which are sensitive to power disturbances. One of the important issues in the PQ problems is to detect and classify disturbance waveforms automatically in an efficient approach, because the possible solutions can be determined after the disturbance types are detected. This paper proposes a particle swarm optimization (PSO) based kernel principal component analysis (KPCA) and support vector machine (SVM) for PQ problem classification. Wavelet based multiresolution analysis (MRA) is utilized to extract features for various PQ disturbances. Dimension of these features are then reduced by KPCA so that the noise has less impact on the classification results. The multi-class SVM is used to classify the PQ problem using the dominant KPCA. The PSO is applied to optimize the KPCA and SVM parameters in order to improve the classification performance. The classification process implemented with various PQ events shows that the proposed technique provides more accuracy than the conventional technique under both noisy and noiseless environments.*

Keywords: Kernel principal component analysis, Support vector machine, Power quality classification, Multiresolution analysis, Particle swarm optimization

1. **Introduction.** Due to the increase in the number of loads which are sensitive to power disturbances, electric power quality (PQ) problems are very important aspects in recent times [1]. Poor electric PQ results from various power line disturbances such as voltage sag, swell, harmonics and outage. Detection and classification disturbance waveforms automatically in an efficient approach is one of the important issues in the PQ problems, because the possible solutions can be determined after the disturbance types are detected. However, in practice, the PQ events are often corrupted with noise due to monitoring devices. These make the detection and classification task more difficult.

Many research works have been carried out in the classification of PQ events using intelligent techniques such as neural networks [2], Fuzzy classifiers [3] and support vector machines (SVMs) [4]. Wavelet based multiresolution analysis (MRA) is utilized for the extraction of features from PQ events which are used as the input to these techniques. The MRA decomposes and represents the energy of the distorted signals at different frequency ranges. The PQ features are constructed from the decomposed signals. The MRA is able to extract the important information from the distorted signals. However, using all features extracted via MRA has some drawbacks, because all features may not correlate to the disturbance types. In addition, some correlate features may be sensitive to noisy condition. Besides, the feature selection is a well-research problem, which can improve the classification performance, speed up the training of the model and reduce the computation effort [5,6]. Thus, to improve the classification performance of PQ problem, the appropriate feature selection is highly expected.

Principal component analysis (PCA) is a powerful technique used for the selection of good features in the classification problem [7]. The PCA is able to handle high-dimensional, noisy and highly correlated data by projecting the data onto a lower-dimensional subspace that constrains most of the variance of the original data. Besides, the PCA is readily performed by solving an eigenvalue problem, or by using iterative algorithms which estimate the principal components. Application of PCA to PQ problems has been successfully proposed in [8]. However, PQ problem has a nonlinear input/output mapping function. Linear correlation among the variables is assumed in PCA, which degrades the performance of PCA in a nonlinear problem.

To solve this drawback, the kernel principal component analysis (KPCA) has been proposed in [9]. KPCA is the extended version of PCA for tackling the nonlinear problem. Two major advantage of KPCA are that the linearity assumption is relaxed and the noise has less impact on the results [10]. Basic idea of KPCA is to first map the input space into a feature space via nonlinear mapping and then compute the principal components (PCs) in that feature space.

In addition, in the recognition or classification stage, the dominant features are used as the input of the classification scheme. The SVM is a novel machine-learning tool and implemented successfully in the classification and regression with small sample cases. The SVM has been proved less vulnerable to over-fitting problem and higher generalization ability since SVM is designed to minimize structural risk [11-13]. In previous works, researchers have employed the SVM in many applications such as motion control of a two wheeled mobile robot [14], stator fault diagnosis for induction motors [15], estimation of automotive engine torque [16], the classification of natural spearmint essence [17], music annotation [18], automatic text summarization [19] and the detection of human CDNA expressions for ovarian carcinoma [20]. Besides, the SVM is successfully applied to the PQ problem such as in [4]. The combination of KPCA and SVM for enhanced statistical analysis of nonlinear processes is proposed in [21]. Therefore, in this work the SVM is adopted as a classification tool.

Besides, both KPCA and SVM have adjusting parameters which significantly affect the classification accuracy. Selecting appropriately these parameters can improve the classification performance. To achieve the suitable parameters, the particle swarm optimization (PSO) [22] is used. The PSO is a novel population based metaheuristic, which utilizes the swarm intelligence generated by the cooperation and competition between the particles in a swarm [23,24]. The PSO has been applied to the optimal tuning of the SVM parameters in [25] as well as the improvement of optimization synthesis and the speed of algorithm convergence.

In this paper, a new approach using KPCA and multi-class SVM (KPCA-SVM) for improving the classification of PQ disturbance signals is presented. The PQ features are extracted directly using the MRA and are applied to the KPCA-SVM for classification of seven types of PQ disturbances. The KPCA is utilized to reduce the feature dimension by projecting the MRA features into the KPCA spaces and then compute kernel principal components (KPCs). The multi-class SVM is employed to classify the PQ problem using dominant KPCs. The PSO is applied to optimize the KPCA and SVM parameters so as to improve the classification performance. As a result, the KPCA-SVM classifier provides better accuracy in the PQ problem classification than the conventional SVM.

The rest of this paper is organized as follows. First, the kernel principal component analysis is explained in Section 2. Next, SVM for classification is described in Section 3. The proposed PSO based KPCA-SVM for PQ problems is explained in Section 4. In Section 5, the experimental results are shown. Finally, the conclusion is provided.

2. **Kernel Principal Component Analysis.** Principal component analysis (PCA) achieves data compression and the extraction of relevant information by projecting the original data sets into the new orthogonal space [7]. Coordinate directions in this new space are obtained from covariance analysis of the original data and are known as principal components (PCs). Dimension reduction is accomplished by selecting the appropriate number of PCs and discarding the uninformative ones. Thus, the data can be represented in the reduced space without losing information.

An extension of PCA, especially developed to deal with nonlinear data distributions, is known as kernel principal component analysis (KPCA) [9]. Basic idea of KPCA is shown in Figure 1. KPCA transforms the nonlinear data into higher dimensional feature space, and then performs the conventional PCA in this space.

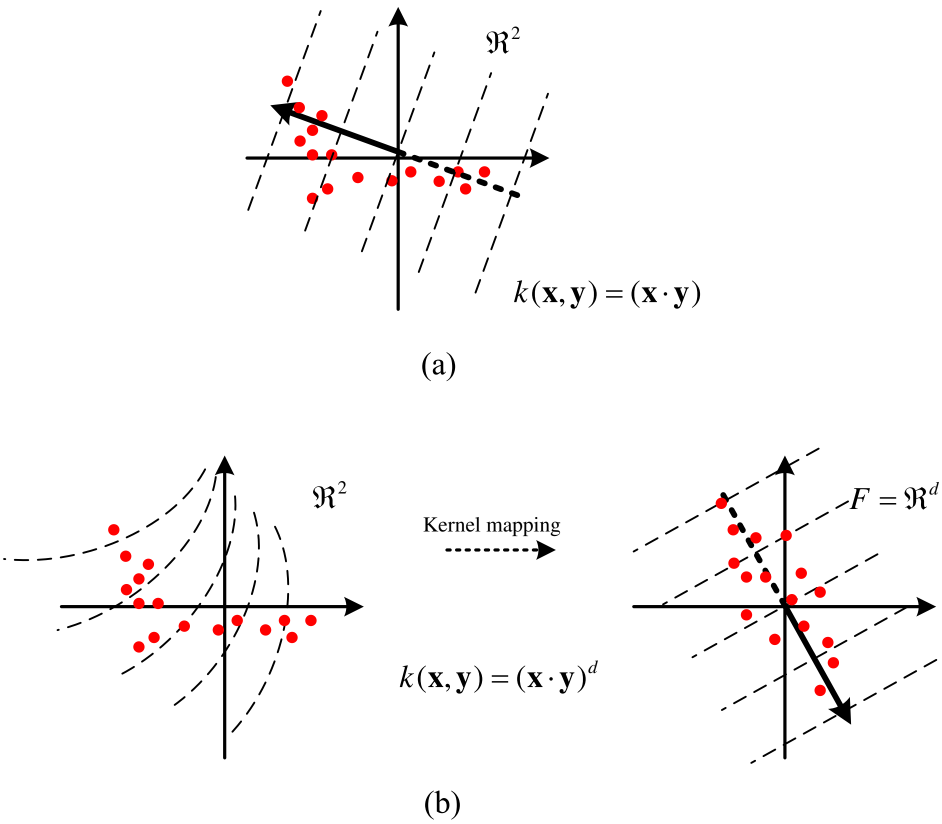

FIGURE 1. Basic idea of KPCA: (a) linear PCA and (b) kernel PCA

To derive KPCA, first the data $\mathbf{x}_k \in \Re^m$, $k = 1, \ldots, N$ are mapping into a feature space F where N is the number of samples. Then, the covariance matrix can be computed from

$$C^F = \frac{1}{N} \sum_{j=1}^{N} \Phi(\mathbf{x}_j)\Phi(\mathbf{x}_j)^T \tag{1}$$

where $\Phi(x_j)$ is the centered nonlinear mapping of the input variables. Then, the principal components are computed by solving the eigenvalue problem as

$$\lambda \mathbf{v} = C^F \mathbf{v} \tag{2}$$

where eigenvalues $\lambda \geq 0$ and $\mathbf{v} \in F \backslash \{0\}$. Here, $C^F \mathbf{v}$ can be represented as follows:

$$C^F \mathbf{v} = \frac{1}{N} \sum_{j=1}^{N} \big\langle \Phi(\mathbf{x}_j), \mathbf{v} \big\rangle \Phi(\mathbf{x}_j) \tag{3}$$

where $\langle \mathbf{x}, \mathbf{y} \rangle$ is the inner product between $\mathbf{x}$ and $\mathbf{y}$. This implies that all solutions $\mathbf{v}$ with $\lambda \neq 0$ must lie in the span of $\Phi(x_1), \ldots, \Phi(x_N)$. Hence, $\lambda \mathbf{v} = C^F \mathbf{v}$ is equivalent to

$$\lambda \langle \Phi(\mathbf{x}_k), \mathbf{v} \rangle = \langle \Phi(\mathbf{x}_k), C^F \mathbf{v} \rangle, \quad k = 1, \ldots, N \tag{4}$$

and there exists coefficients α_i $(k = 1, \ldots, N)$ such that

$$\mathbf{v} = \sum_{i=1}^{N} \alpha_i \Phi(\mathbf{x}_i) \tag{5}$$

Combining (4) and (5), obtains

$$\lambda \sum_{i=1}^{N} \alpha_i \big\langle \Phi(\mathbf{x}_k), \Phi(\mathbf{x}_i) \big\rangle = \frac{1}{N} \sum_{i=1}^{N} \alpha_i \Big\langle \Phi(\mathbf{x}_k), \sum_{j=1}^{N} \Phi(\mathbf{x}_j) \Big\rangle \big\langle \Phi(\mathbf{x}_j), \Phi(\mathbf{x}_i) \big\rangle \tag{6}$$

for all $k = 1, \ldots, N$.

Now, let us define an $N \times N$ matrix $\mathbf{K}$ by $[\mathbf{K}]_{ij} = K_{ij} = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$. As a result,

$$\lambda N \mathbf{K} \alpha = \mathbf{K}^2 \alpha \tag{7}$$

where $\alpha = [\alpha_1, \ldots, \alpha_N]^T$. To find solutions of (7), the eigenvalue problem (for nonzero eigenvalues) is solved by

$$N \lambda \alpha = \mathbf{K} \alpha. \tag{8}$$

Then, performing PCA in the feature space F is equal to resolving the eigen-problem of (8). This yields eigenvectors $\alpha_1, \alpha_2, \ldots, \alpha_N$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. Dimensionality can be reduced by retaining only the first p eigenvectors. Parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$ are normalized by requiring that the corresponding vectors in F be normalized, i.e.,

$$\big\langle \mathbf{v}_k, \mathbf{v}_k \big\rangle = 1, \quad \text{for all} \quad k = 1, \ldots, p \tag{9}$$

Using $\mathbf{v}_k = \sum_{i=1}^{N} \alpha_i^k \Phi(\mathbf{x}_i)$, (10) leads to

$$1 = \lambda_k \langle \alpha_k, \alpha_k \rangle \tag{10}$$

The principal components $\mathbf{t}$ of a test vector $\mathbf{x}$ are then extracted by projecting $\Phi(\mathbf{x})$ onto eigenvectors $\mathbf{v}_k$ in F, where $k = 1, \ldots, p$.

$$\mathbf{t}_k = \langle \mathbf{v}_k, \Phi(\mathbf{x}) \rangle = \sum_{i=1}^{N} \alpha_i^k \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}) \rangle \tag{11}$$

To solve the eigen-problem of (8) and to project from the input space to the KPCA space using (11), one can avoid the need for both performing the nonlinear mappings and computing both inner products in the feature space through the introduction of a *kernel function*, that is, $k(\mathbf{x}, \mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle$.

If one has nonlinear information of process, it could be used to select the kernel function among kernels in KPCA. Before applying KPCA, mean centering and variance scaling in high-dimensional space should be performed. Mean centering can be done by substituting the kernel matrix $\mathbf{K}$ with

$$\tilde{\mathbf{K}} = \mathbf{K} - 1_N\mathbf{K} - \mathbf{K}1_N + 1_N\mathbf{K}1_N \tag{12}$$

where $1_N = \dfrac{1}{N}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \in R^{N\times N}$.

Variance scaling can then be done by

$$\tilde{\mathbf{K}}_{scl} = \frac{\tilde{\mathbf{K}}}{trace(\tilde{\mathbf{K}})/(N-1)} \tag{13}$$

Examples of kernel functions are given in Table 1, where the parameters ρ and γ are determined by the user.

Table 1. Examples of kernel functions

Kernel type	Function
Linear	$k(x, y) = \langle x, y \rangle$
Polynomial	$k(x, y) = \langle x, y \rangle^{\rho}$
Radial basis function (RBF)	$k(x, y) = \exp\left(-\frac{\lVert x-y \rVert^2}{\gamma}\right)$

3. **Support Vector Machine.** Consider a training set of N points $\{(x_i, y_i)\}_{i=1}^{N}$ with data $x_i \in \Re^n$ and the corresponding class labels $y_i \in \{-1, +1\}$. The basic idea of SVM is to map the training data from the input space into a higher dimensional feature space via kernel function and then construct a separating hyperplane with maximum margin in the feature space [11]. The SVM problem can be formulated as a quadratic programming optimization problem that will find the weight parameter $\mathbf{w}$ and the bias parameter b. These two parameters will maximize the margin while ensuring that the training samples are well classified. Thus the SVM computes the optimal separating hyperplane by solving the following optimization problem:

$$\min \ ||\mathbf{w}||^2/2 + C\sum_{i=1}^{N} \xi_i \tag{14}$$

$$\text{Subject to } \ y_i(\mathbf{w}^T \cdot \varphi(x_i) + b) \le 1 - \xi_i$$
$$\xi_i \ge 0, \quad i = 1, \dots, N.$$

where $\xi_i \ge 0$, $i = 1, \dots, N$ are slack variables, parameter C is used to tune the trade-off between the amount of errors accepted, and $\varphi(\cdot)$ is transformation function. Using the Lagrange multipliers $\alpha_i \ge 0$, $i = 1, \dots, N$ to solve the quadratic programming problem, the dual problem can be derived as:

$$\max \sum_{i=1}^{N} \alpha_i - \left(\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \right) \Big/ 2 \tag{15}$$

$$\text{Subject to } \sum_{i=1}^{N} \alpha_i y_i = 0$$

$$0 \le \alpha_i \le C, \quad i = 1, \ldots, N.$$

For nonlinear classification problems, the data are first mapped into a higher dimensional feature space F by the transformation function $\phi : \mathbf{x} \longrightarrow \phi(\mathbf{x}) \in F \subset R^p$. Hence, the resulted decision function is given by

$$f(\mathbf{x}) = \text{sign} \left(\sum_{\alpha_i > 0} y_i \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b \right) \tag{16}$$

where $k(\mathbf{x}, \mathbf{x}_i)$ is a kernel mapping function between sampling $\mathbf{x}$ and support vector $\mathbf{x}_i$.

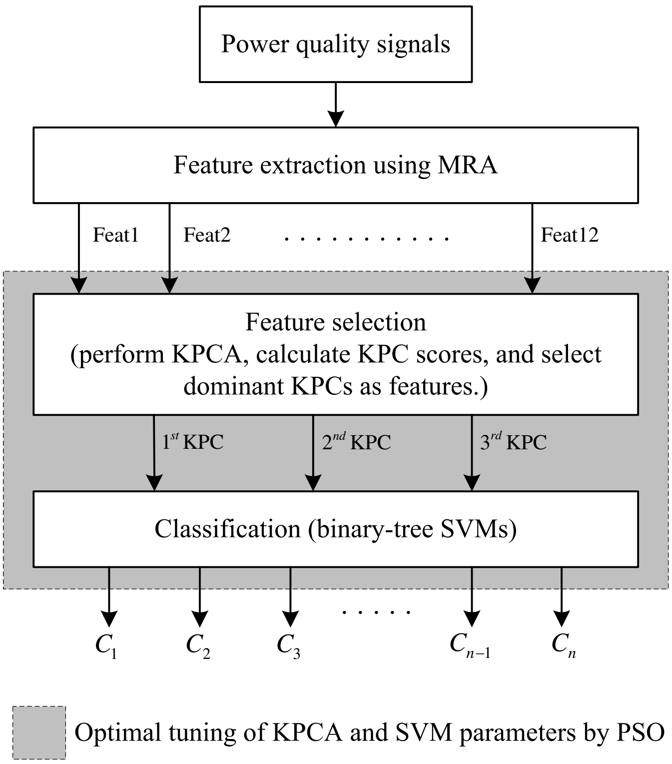

FIGURE 2. The PSO based KPCA-SVM for power quality problem classification

4. **PSO Based KPCA and Multi-class SVM for PQ Problems.** PQ problems can be described as any variations in the electrical power services resulting in mis-operation or failure of end-use equipment [1]. The proposed PSO based KPCA and multi-class SVM (called KPCA-SVM) for PQ problem classification is shown in Figure 2. In the model, it is assumed that the PQ signals are continuously recorded using PQ monitoring systems. The MRA is used for features extraction from the recorded PQ signals. The KPCA is utilized in the feature selection stage in order to reduce the feature dimensions. The

binary-tree SVM is employed to classify the PQ problem using dominant KPCs. Besides, the PSO is applied to optimize the KPCA-SVM parameters.

The unique features and the main advantages of the proposed PSO based KPCA-SVM for PQ classification problem are

1) The proposed KPCA-SVM is very accurate because the proper selection of features vectors by KPCA.

2) The proposed KPCA-SVM is robust because it can maintain the same accuracy even the different signal-to-noise ratios appear. This will be shown in the simulation result.

3) Without trial and error, all adjusting parameters, i.e., SVM and KPCA parameters are optimized by a PSO. The optimization objective is to maximize the accuracy of classification results for all training samples in the training of SVM using a 2-fold cross-validation technique.

4.1. **MRA based feature extraction.** In the feature extraction, the appropriate features are identified. This step is important for prediction algorithm as it affects the classification performance significantly. In this work, the standard deviation of MRA as described in [1] is used for feature extraction. The strategy of construction of MRA features is given as follows:

1) Use MRA to decompose the distorted signal into different resolution levels. The number of levels is selected to cover the highest frequency band of interest.

2) Find the standard deviation for each detail version at different resolution levels of the distorted signal. The standard deviation (Std.) for non-distributed frequency data, is defined as

$$\text{Std.} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})^2} \tag{17}$$

where $\mathbf{x}_i$ is non-distributed frequency data, $\bar{\mathbf{x}}$ is mean of data, and N is number of data.

3) Construct the standard deviation of MRA curve by plotting the standard deviation for each resolution level.

4.2. **KPCA based feature selection.** As shown in Figure 2, the KPCA is utilized in the feature selection stage to reduce the feature dimensions. The strategies of KPCA for feature selection are described as follows:

1) Compute the KPCA from the MRA features. Here, the MRA data is projected onto the KPCA space. The data in this new space is represented as a set of eigenvectors which are specified the data main directions.

2) Calculate the KPC scores. The KPCA technique allows the definition of significant values (eigenvalues) that weight the spread of the data sample through the main directions. These weight values are called the scores of KPC.

3) Select the dominant KPCs as the classification features. The obtained KPCs features are selected by considering the level of KPC scores. Besides, the number of KPC features used is selected by the classification accuracy obtained.

4.3. **Multi-class SVM for PQ classification.** As described, the SVMs are basically binary classifiers where the class labels can take only two values, i.e., +1 and −1. However, the classification of PQ problems often involves the simultaneous discrimination of numerous classes. In order to face this issue, a number of multi-class classification strategies can be adopted such as one-against-all, one-against-one and binary-tree classification. Among these approaches, the binary tree is promising choices. Because, this method uses the least number of classifier and repeated training sample as well as the improvement in speed of training and classification.

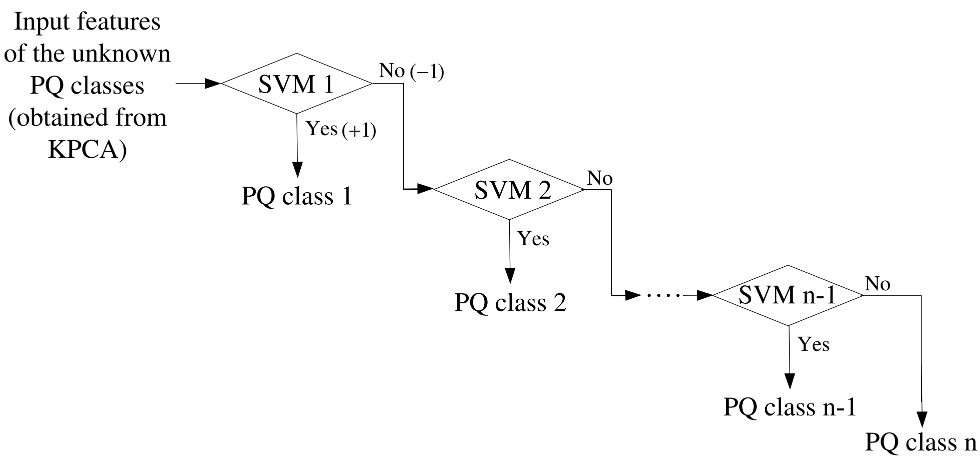

FIGURE 3. Power quality classification method using binary-tree SVMs

In this paper, the binary-tree classification is adopted as shown in Figure 3. The detail steps are as follows. First, select the PQ problem having highest frequency as the first class in n number of PQ problems while the $n-1$ number of PQ problems as the other class to build a binary classification SVM model. Do the same things until the last two classes are used to set up binary classification SVM model and all the history PQ problems are classified.

4.4. **PSO-based KPCA-SVM parameters.** PSO is the optimization techniques based on the observations of the social behavior of animals and the swarm theory [22]. The PSO has the advantage of being very simple in concept, easy to implement and computationally efficient algorithm. The PSO uses the concept of population and a measure of performance similar to the fitness value used with evolutionary algorithms. Population consists of potential solutions called particles. In each iteration, each particle remembers the best solution found by itself (personal best value: *pbest*), and by the whole swarm along the search trajectory (global best value: *gbest*).

In this work, the PSO is adopted to tune the KPCA and SVM parameters in order to improve the classification performance. Besides, during PSO optimization, the 2-fold cross-validation (CV) technique has been applied to evaluate the performance of the proposed method. It can detect and prevent over-fitting in a model. In this technique, all the training samples are divided randomly into two groups. The training and testing procedures are performed twice, i.e., the training in the first group and the testing in the second group, and vice versa. The accuracy of the model for optimization is computed using the classification results of all training samples. Therefore, an optimization problem based on the 2-fold cross validation is defined as:

$$\begin{aligned} &\text{Maximize } \frac{1}{T}\sum_{n=1}^{T-1} J_n \\ &\text{Subject to } \rho_{n,\min} \le \rho_n \le \rho_{n,\max}, \\ &\qquad C_{n,\min} \le C_n \le C_{n,\max}, \\ &\qquad \gamma_{n,\min} \le \gamma_n \le \gamma_{n,\max}, \quad n = 1, 2, \ldots, T-1 \end{aligned} \tag{18}$$

where J_n is the number of correct-classified in the PQ problem class n, $\rho_{n,\min}$ and $\rho_{n,\max}$ are the minimum and maximum of kernel parameter in KPCA of PQ problem class n, respectively, $C_{n,\min}$ and $C_{n,\max}$ are the minimum and maximum of SVM adjustable constants of PQ problem class n, respectively, $\gamma_{n,\min}$ and $\gamma_{n,\max}$ are the minimum and maximum of kernel parameter in SVM of PQ problem class n, respectively. In (18), it can be seen that sets of parameters ρ, C and γ are formulated as a particle in the PSO optimization. Based on the above mentioned illustrations, the proposed PSO-based KPCA-SVM parameters technique can be described as follows.

1) Specify the parameters of PSO such as population size of swarm, lower and upper bound values of problem space ($\rho_{n,\min}$, $\rho_{n,\max}$, $C_{n,\min}$, $C_{n,\max}$, $\gamma_{n,\min}$, $\gamma_{n,\max}$), minimum and maximum velocity of particles ($v_{\min}$, $v_{\max}$), minimum and maximum inertia weights ($w_{\min}$, $w_{\max}$), maximum iteration ($iter_{\max}$), and acceleration coefficients (c_1, c_2).

2) Randomly initialize 1^{st} particles.

3) Evaluate the fitness function as in (18) of each particle, find the best position found by particle i, call it as *pbest*, and find the best position found by swarm, call it as *gbest*.

4) Update the inertia weight (w) as follows:

$$w = \frac{w_{\max} - w_{\min}}{iter_{\max}} \times iter \tag{19}$$

where $w_{\max}$ and $w_{\min}$ are the initial and final weight, respectively, $iter_{\max}$ and $iter$ are the maximum and current iteration number, respectively.

5) Update the velocity (v) and position (u) of each particle,

$$v_i^{k+1} = (w \times v_i^k) + \left(c_1 \times rand_1 \times \left(pbest_i - u_i^k\right)\right) + \left(c_2 \times rand_2 \times \left(gbest_i - u_i^k\right)\right) \tag{20}$$

$$u_i^{k+1} = u_i^k + v_i^{k+1} \tag{21}$$

where v_i^k is velocity of agent i at iteration k, w is weighting function, c_1 and c_2 are the relative weights of the personal best position and global best position, respectively, *rand* is random number between 0 and 1, u_i^k is current position of agent i at iteration k, $pbest_i$ is *pbest* of particle i, $gbest_i$ is *gbest* of the group.

6) Increment the iteration for a step ($k = k+1$). If the current iteration is the maximum iteration $k = iter_{\max}$, stop. If not, go to Step 3.

5. **Experimental Results.** In this work, seven classes of PQ problems are investigated, i.e., pure sine (normal), harmonic distortions, voltage sag, voltage swell, outage, sag with harmonics and swell with harmonics, as described in [26]. The training and test data are generated from the parametric equations with different parameters as shown in Table 2. The simulated signals are sampled at 256 points/cycle and the normal frequency is 50 Hz.

However, in real electric power systems, the signals have usually noises. In order to test the sensitivity of the proposed method under different noise conditions, different levels of noises are added. In the research of PQ issues, the commonly noise considered is the additive white Gaussian noise (AWGN) [2]. Here, the different levels of noises with signal to noise ratio (SNR) values of 20, 30, 40 and 50 dB are added. The SNR is defined as

$$\text{SNR(dB)} = 10 \log \left(\frac{P_s}{P_n}\right) \tag{22}$$

where P_s and P_n are the power (variance) of the signal and noise, respectively.

Examples of the PQ signals integrated with noise values of 20 and 40 dB are shown in Figures 4-9. It can be seen that the noises are affected to the PQ signal significantly. These can degrade the PQ problem classification accuracy.

For each class of PQ disturbances (C_1 to C_7 in Table 2), 200 cases were generated for training and other 500 cases were generated for testing. Thus, the total signals used in this work are $(200 + 500) \times 7 = 4900$ cases. Note that, the data in both training and test set are added with noise. More details of data generated for training and testing are shown in Table 3.

TABLE 2. Parametric equation for simulation of disturbed signals

PQ signal	Class symbol	Equation and parameters variations
Pure sine	C_1	$f(t) = \sin(wt)$
Harmonics	C_2	$f(t) = (\alpha_1 \sin(wt) + \alpha_3 \sin(3wt) + \alpha_5 \sin(5wt) + \alpha_7 \sin(7wt))$ $0.05 \le \alpha_i \le 0.15,\ i = 1, 3, 5, 7,\ \sum \alpha_i^2 = 1$
Sag	C_3	$f(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) \sin(wt)$ $0.1 \le \alpha \le 0.9;\ T \le t_2 - t_1 \le 9T;\ t_1 \le t_2$
Swell	C_4	$f(t) = (1 + \alpha(u(t - t_1) - u(t - t_2))) \sin(wt)$ $0.1 \le \alpha \le 0.8;\ T \le t_2 - t_1 \le 9T;\ t_1 \le t_2$
Outage	C_5	$f(t) = (1 - \alpha(u(t - t_1) - u(t - t_2))) \sin(wt)$ $u(t) = 1, \text{ for } t \le 0, \text{ and } u(t) = 0, \text{ for } t < 0$ $0.9 \le \alpha \le 1.0;\ T \le t_2 - t_1 \le 9T;\ t_1 \le t_2$
Sag with harmonics	C_6	$f(t) = (1 - \alpha(u(t - t_1) - u(t - t_2)))(\alpha_1 \sin(wt) + \alpha_3 \sin(3wt) + \alpha_5 \sin(5wt))$ $0.1 \le \alpha \le 0.9;\ T \le t_2 - t_1 \le 9T;\ t_1 \le t_2$ $0.05 \le \alpha_i \le 0.15,\ i = 1, 3, 5,\ \sum \alpha_i^2 = 1$
Swell with harmonics	C_7	$f(t) = (1 + \alpha(u(t - t_1) - u(t - t_2)))(\alpha_1 \sin(wt) + \alpha_3 \sin(3wt) + \alpha_5 \sin(5wt))$ $0.1 \le \alpha \le 0.8;\ T \le t_2 - t_1 \le 9T;\ t_1 \le t_2$ $0.05 \le \alpha_i \le 0.15,\ i = 1, 3, 5,\ \sum \alpha_i^2 = 1$

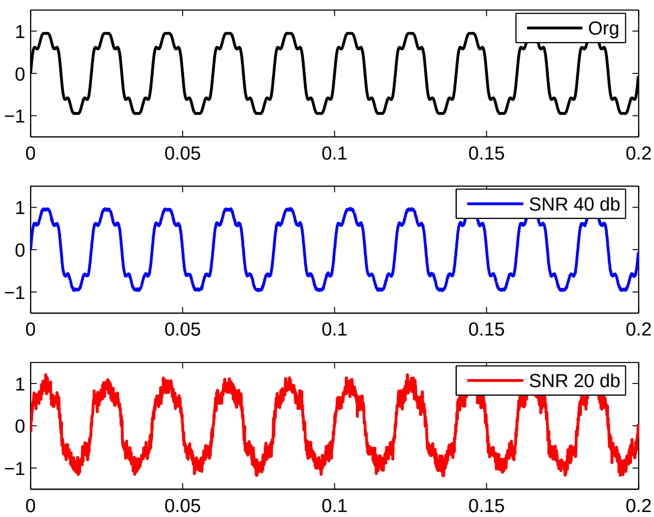

FIGURE 4. PQ signal integrated with noise values of 20 and 40 dB: Harmonics

TABLE 3. The number of data generated for training and testing (samples)

Data type	Training data with noise added					Testing data with noise added					Total
	0 dB	20 dB	30 dB	40 dB	50 dB	0 dB	20 dB	30 dB	40 dB	50 dB	
One class	40	40	40	40	40	100	100	100	100	100	700
All 7 classes	280	280	280	280	280	700	700	700	700	700	4900

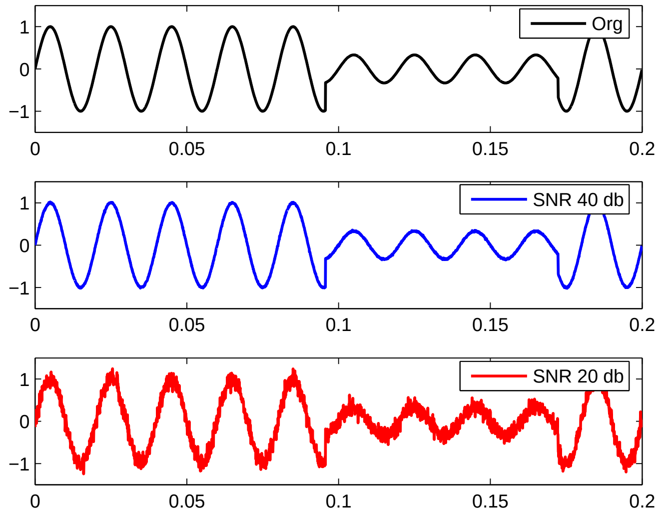

Figure 5. PQ signal integrated with noise values of 20 and 40 dB: Sag

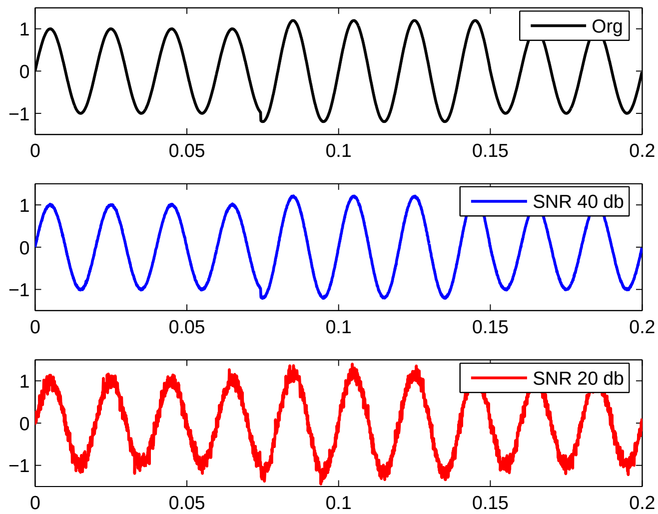

Figure 6. PQ signal integrated with noise values of 20 and 40 dB: Swell

In the MRA feature extraction, the Daubechies 4 wavelet with 12 levels of decomposition is utilized [1]. There are 12 features after the MRA procedure. These MRA features are used as the input of the KPCA-SVM.

For all KPCA-SVM experiments, the polynomial KPCA and radial basis function (RBF) type of the binary-tree SVM are adopted for classifying seven PQ disturbances. Because the SVM with RBF kernel function can provide promising results in the PQ classification problems [4]. The KPCA and SVM parameters are determined using PSO. The objective is to choose the model that produces maximum correct classification. Besides, the number of KPCs used in the classification model, are also investigated.

In the PSO-based KPCA-SVM parameters, the ranges of search parameters for PSO are set as follows: $\rho \in [0.01 \;\; 3.5]$, $C \in [1.0 \;\; 10^6]$ and $\gamma \in [1.0 \;\; 20]$, respectively. The

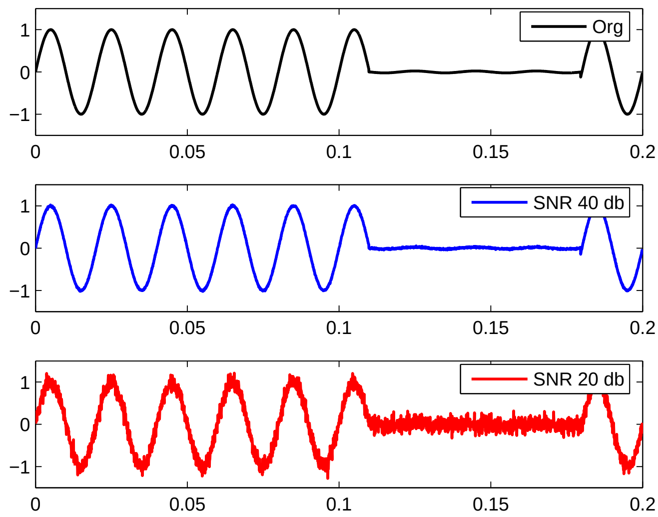

FIGURE 7. PQ signal integrated with noise values of 20 and 40 dB: Outage

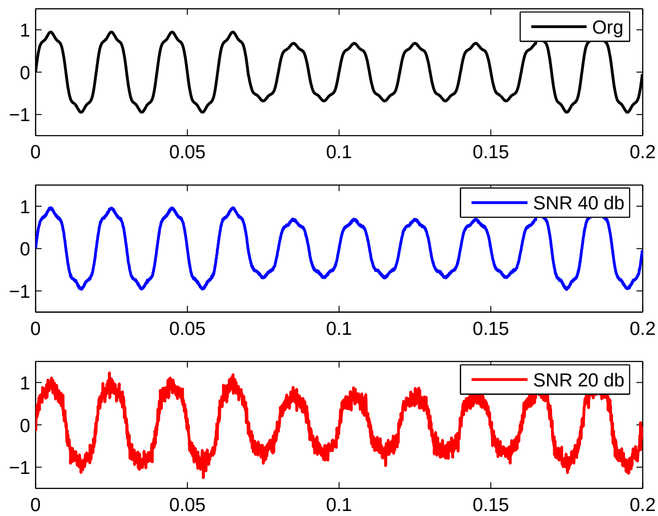

FIGURE 8. PQ signal integrated with noise values of 20 and 40 dB: Sag with harmonics

PSO parameters are set as follows: maximum particle velocity = 4, population size = 24, $w_{\max} = 0.9$, $w_{\min} = 0.4$, $c_1 = 2$, $c_2 = 2$ and $iter_{\max} = 600$.

Figure 10 shows percent explained of the KPCA eigenvalues. The score of the eigenvalue represent the information of the original data which is placed in the associated KPCA component. It can be seen that, the first component of KPCA has very high score (90.33%). This implies that the information of the original data is approximately placed in the first KPCs.

Percent explain of the eigenvalues in Figure 10 also indicates that the variables in PQ data are correlated (i.e., the largest four eigenvalues of the PQ data are significantly larger than those of the remaining eigenvalues).

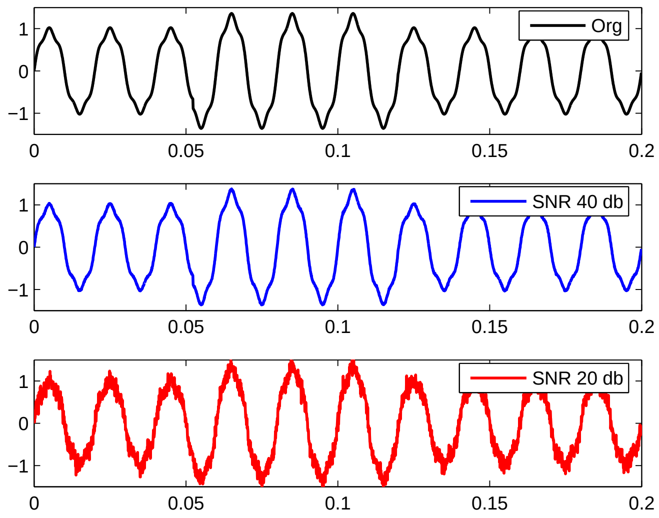

FIGURE 9. PQ signal integrated with noise values of 20 and 40 dB: Swell with harmonics

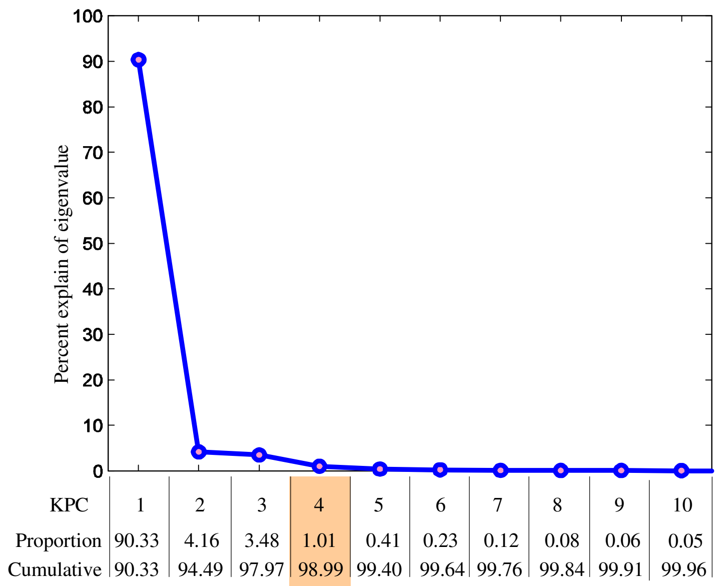

FIGURE 10. Percent explain of the eigenvalue of the KPCA components

TABLE 4. Classification accuracy using different number of KPCs retrained

Number of KPCs retrained	Classification accuracy
2	75.00
3	87.74
4	**96.09**
5	92.49
6	90.51
7	91.94
8	93.34

The classification results with difference number of KPCs retrained are shown in Table 4. It can be seen that using the first 4 KPCs produces the best accuracy which is consistent to the cumulative value in Figure 10. The first 4 KPCs not only cause the cumulative close to the remainder number of KPCs retained but also provide good classification performance. Besides, the convergence curve of the optimized KPCA-SVM method using the first 4 KPCs is shown in Figure 11.

Table 5 shows the KPCA-SVM parameters optimized by PSO. The 4 KPCs retained model is selected for further analysis based on the aforementioned tests.

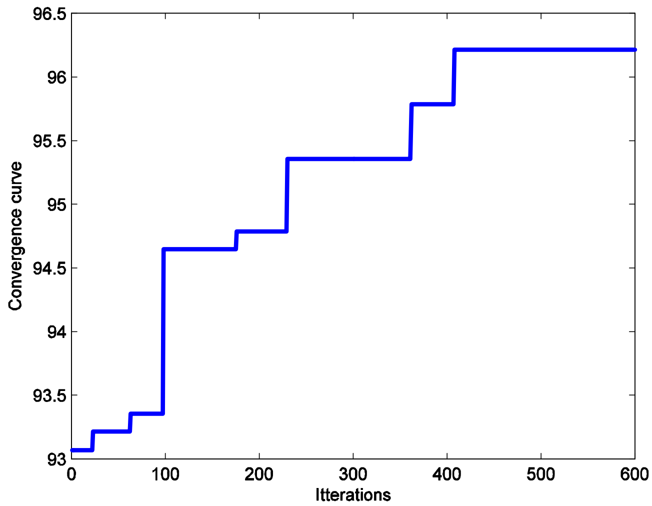

Figure 11. Convergence curve of PSO in KPCA-SVM using 4 KPCs retained

Table 5. KPCA-SVM parameters optimized by PSO in the 4 KPCs retained model

Model	KPCA parameter (ρ)	SVM parameters	
		C	γ
SVM_1	0.72	7949	5.61
SVM_2	0.17	5704	9.09
SVM_3	0.21	8867	7.04
SVM_4	0.50	9992	3.10
SVM_5	0.38	7102	2.60
SVM_6	0.14	9993	4.68

Table 6 shows the confusion matrix of the correct classification of the proposed KPCA-SVM without noise. Note that, the test signals are not included in the training stage. Table 7 shows the percentage of correct classification results of KPCA-SVM under different SNR values.

In order to evaluate the effectiveness of the proposed method, comparison studies with other different schemes, i.e., the conventional SVM (SVM) and the results of classification scheme in previous works, are performed. The strategy of constructing the SVM is the same as the KPCA-SVM. Thus, all 12 MRA features are used as the input of the SVM directly.

Table 6. Confusion matrix of the correct classification in case without noise based on KPCA-SVM

Actual class	Estimated class							Acc
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	
C_1	100	0	0	0	0	0	0	100
C_2	0	100	0	0	0	0	0	100
C_3	4	0	83	0	12	1	0	83
C_4	0	0	0	100	0	0	0	100
C_5	0	0	5	0	93	2	0	93
C_6	0	0	0	0	2	98	0	98
C_7	0	0	0	0	0	0	100	100
Average								**96.29**

Table 7. Percentage of correct classification under different SNR values based on KPCA-SVM

PQ problem	20 dB	30 dB	40 dB	50 dB	Without noise
C_1	100	100	100	100	100
C_2	98	100	100	100	100
C_3	82	85	83	83	83
C_4	100	100	100	100	100
C_5	91	89	93	93	93
C_6	99	98	98	98	98
C_7	99	100	100	100	100
Average	**95.57**	**96.00**	**96.29**	**96.29**	**96.29**
Average					**96.09**

Table 8 shows the comparison in terms of the percent classification accuracy between the results of this study (SVM and KPCA-SVM) and the classification scheme in [27,28] in case without noise. It can be seen that, the proposed KPCA-SVM method achieves high classification success over other methods.

Table 9 shows the comparison in terms of the percent classification accuracy between the results of this study and the classification scheme in [2]. It can be seen that, the proposed KPCA-SVM method not only withstands noise but also achieves high classification success.

6. **Conclusion.** An application of PSO based KPCA-SVM for PQ problem classification has been presented in this paper. The best characteristics of KPCA and SVMs theories are utilized for detecting as well as classifying nonlinear behavior of PQ events. The MRA is used to extract the features of PQ signals. The KPCA is applied to the selection of dominant MRA features which are used as the input of the binary-tree SVM. The PSO is adopted for optimizing KPCA and SVM parameters simultaneously. The proposed PSO based KPCA-SVM, not only reduces the features dimension but also robustly tolerates the noisy environments. Experimental results with seven types of PQ signals, i.e., pure-sine, harmonics, sag, swell, outage, sag with harmonics and swell with harmonics, demonstrate that the proposed method can improve the classification accuracy under both noiseless and noisy conditions.

Table 8. Performance comparison in terms of percentage of correct classification in case without noise

PQ problem	T. K. Abdel-Galil et al. [27]	H. He et al. [28]	SVM	KPCA-SVM
C_1	100	100	100	100
C_2	100	100	100	100
C_3	76.5	87	79	83
C_4	97	100	98	100
C_5	90	80.5	86	93
C_6	71.5	97	95	98
C_7	98	100	100	100
Average	**90.43**	**94.93**	**94.00**	**96.29**

Table 9. Performance comparison in terms of percentage of correct classification

Condition	M. Uyar et al. [2]	SVM	KPCA-SVM
Without noise	95.71	94.00	96.29
With noise: SNR 50 dB	95.14	94.00	96.29
With noise: SNR 40 dB	93.64	94.57	96.29
With noise: SNR 30 dB	91.85	94.00	96.00
With noise: SNR 20 dB	89.92	90.42	95.57
Average	**92.50**	**93.40**	**96.09**

Nevertheless, the weakness of the proposed technique is the strategy to extract the features from the PQ signals. The advent of PQ feature extraction methode can be adopted here such as energy difference of multiresolution, S-transform algorithm, hyperbolic S-transform and TT-transform. Integrating these transformation techniques into the feature extraction stage, is able to achieve more classification accuracy because all important information of PQ signals can be represented actually.

Acknowledgment. This work was supported by the King Mongkut's Institute of Technology Ladkrabang Research Fund.

REFERENCES

[1] A. M. Gaouda, M. R. Salama, M. R. Sultan and A. Y. Chikhani, Power quality detection and classification using wavelet-multiresolution signal decomposition, *IEEE Transactions on Power Delivery*, vol.14, no.4, pp.1469-1476, 2002.

[2] M. Uyar, S. Yildirim and M. T. Gencoglu, An effective wavelet-based feature extraction method for classification of power quality disturbance signals, *Electric Power Systems Research*, vol.78, pp.1747-1755, 2008.

[3] S. K. Meher and A. K. Pradhan, Fuzzy classifiers for power quality events analysis, *Electric Power Systems Research*, vol.80, pp.71-76, 2010.

[4] H. Eristi and Y. Demir, A new algorithm for automatic classification of power quality events based on wavelet transform and SVM, *Expert Systems with Applications*, vol.37, no.6, pp.4094-4102, 2010.

[5] G.-J. Mun, B.-N. Noh and Y.-M. Kim, Enhanced stochastic learning for feature selection in intrusion classification, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(A), pp.3625-3635, 2009.

[6] S. Hengpraprohm and P. Chongstitvatana, Feature selection by weighted-SNR for cancer microarray data classification, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(A), pp.4627-4635, 2009.

[7] C.-Y. Chang and H.-R. Hsu, Application of principal component analysis to a radial-basis function committee machine for face recognition, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(B), pp.4145-4154, 2009.

[8] Y. Y. Hong and B. Y. Chen, Locating switched capacitor using wavelet transform and hybrid principal component analysis network, *IEEE Transactions on Power Delivery*, vol.22, no.2, pp.1145-1152, 2007.

[9] B. Scholkopf, A. Smola and K. R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Compute*, vol.10, no.5, pp.1299-1319, 1998.

[10] A. M. Jade, B. Sridanth, V. K. Jayaraman, B. D. Kulkarni, J. P. Jog and L. Priya, Feature extraction and denoising using kernel PCA, *Chem. Eng. Science*, vol.58, pp.4441-4448, 2003.

[11] S. Gunn, Support vector machines for classification and regression, *ISIS Technical Report ISIS-1-98*, Image Speech & Intelligence Systems Research Group, University of Southampton, 1998.

[12] Z. Hua, X. Zhang and X. Xu, Asymmetric support vector machine for the classification problem with asymmetric cost of misclassification, *International Journal of Innovative Computing, Information and Control*, vol.6, no.12, pp.5597-5608, 2010.

[13] H. Tamura and K. Tanno, Midpoint validation method for support vector machines with margin adjustment technique, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(A), pp.4025-4032, 2009.

[14] L. Jiang, M. Deng and A. Inoue, Obstacle avoidance and motion control of a two wheeled mobile robot using SVR technique, *International Journal of Innovative Computing, Information and Control*, vol.5, no.2, pp.253-262, 2009.

[15] C. Manop, V. Kinnares and S. Chunwiphat, SVM-based stator fault diagnosis for induction motors using dq0 voltage components, *International Journal of Innovative Computing, Information and Control*, vol.6, no.10, pp.4539-4553, 2010.

[16] C. Ran, D. Sun and D. Qin, Estimation of automotive engine torque using support vector machines, *ICIC Express Letters*, vol.4, no.3(A), pp.757-762, 2010.

[17] X. Song, S. K. Halgamuge, D. Chen, S. Hu and B. Jiang, The optimized support vector machine with correlative features for classification of natural spearmint essence, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.1089-1099, 2010.

[18] G. Chen, T.-J Wang, L.-Y. Gong and P. Herrera, Multi-class support vector machine active learning for music annotation, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.921-930, 2010.

[19] N. Begum, M. A. Fattah and F. Ren, Automatic text summarization using support vector machine, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.1987-1996, 2009.

[20] C.-H. Lai, S.-S. Yu, J.-D. Chang and M.-H. Tsai, Statistical and SVM-based on cogene detection of human CDNA expressions for ovarian carcinoma, *International Journal of Innovative Computing, Information and Control*, vol.5, no.10(A), pp.3157-3177, 2009.

[21] Y. Zhang, Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM, *Chem. Eng. Science*, vol.64, no.5, pp.801-811, 2009.

[22] J. Kennedy and R. C. Eberhart, Particle swarm optimization, *Proc. of IEEE Int. Conf. on Neural Net*, pp.1942-1948, 1995.

[23] J.-F. Chang, A performance comparison between genetic algorithms and particle swarm optimization applied in constructing equity portfolios, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(B), pp.5069-5079, 2009.

[24] C.-C. Lai, C.-H. Wu and M.-C. Tsai, Feature selection using particle swarm optimization with application in spam filtering, *International Journal of Innovative Computing, Information and Control*, vol.5, no.2, pp.423-432, 2009.

[25] S.-W. Lin, K.-C. Ying, S.-C. Chen and Z.-J. Lee, Particle swarm optimization for parameter determination and feature selection of support vector machines, *Expert Systems with Applications*, vol.35, no.4, pp.1817-1824, 2008.

[26] IEEE recommended practice for monitoring electric power quality, *IEEE Std. 1159-1995*, 1995.

[27] T. K. Abdel-Galil, M. Kamel, A. M. Youssef, E. F. El-Saadany and M. M. A. Salama, Power quality disturbance classification using the inductive inference approach, *IEEE Transactions on Power Delivery*, vol.19, no.4, pp.1812-1818, 2004.

[28] H. He and J. A. Starzyk, A self-organizing learning array system for power quality classification based on wavelet transform, *IEEE Transactions on Power Delivery*, vol.21, no.1, pp.286-295, 2006.