Research Article

Power Transformer Partial Discharge Fault Diagnosis Based on Multidimensional Feature Region

Rong Jia, Yongtao Xie, Hua Wu, Jian Dang, and Kaisong Dong

1Xi'an University of Technology, Xi'an, Shaanxi 710048, China
2State Grid Gansu Province Electric Power Research Institute, Gansu 730000, China

Correspondence should be addressed to Yongtao Xie; 18792511340@163.com

Received 16 March 2016; Revised 7 June 2016; Accepted 3 July 2016

Academic Editor: Wenyu Zhao

Copyright © 2016 Rong Jia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Effectively extracting power transformer partial discharge (PD) signals feature is of great significance for monitoring power transformer insulation condition. However, there has been lack of practical and effective extraction methods. For this reason, this paper suggests a novel method for the PD signal feature extraction based on multidimensional feature region. Firstly, in order to better describe differences in each frequency band of fault signals, empirical mode decomposition (EMD) and Hilbert-Huang transform (HHT) band-pass filter wave for raw signal is carried out. And the component of raw signals on each frequency band can be obtained. Secondly, the sample entropy value and the energy value of each frequency band component are calculated. Using the difference of each frequency band energy and complexity, signals feature region is established by the multidimensional energy parameters and the multidimensional sample entropy parameters to describe PD signals multidimensional feature information. Finally, partial discharge faults are classified by sphere-structured support vector machines algorithm. The result indicates that this method is able to identify and classify different partial discharge faults.

1. Introduction

Power transformer is one of the most critical and expensive electrical pieces of equipment in power system, whose safety and reliability are closely related to the operation condition of the whole system. Furthermore, in the process of actual operation, power transformer is unavoidable to subject to such outside factor influence as electricity, machinery and heat, and so forth further causing its winding insulation deterioration to produce partial discharge (PD) phenomena, threatening the safety of operation of the whole system. Therefore, it is essential to monitor the insulation condition and provide a proper maintenance action for in-service power transformers [1–3]. A CIGRE international survey indicates that most of power transformers life is 30 years or so [4]. However, the majority of currently in-service power transformers were installed prior to 1980 and as a result the bulk of the population is approaching or has already exceeded its design life [5, 6], which leads to a significant risk for power system stakeholders. Consequently, monitoring and diagnostic techniques are of important significance for improving operational reliability of the in-service power transformers.

At present, various monitoring and diagnostic techniques have been adopted for power transformers such as visual inspection, infrared scanning, tanδ measurement, PD measurement, and oil meteorologic chromatogram analysis [7, 8]. Among them, PD measurement is considered as an effective diagnostic tool to assess and monitor insulation condition for in-service power transformers. PD is a type of slight failure that does not usually lead to serious insulation damage before electrical tree occurs. Therefore, it is very critical for stakeholders to determine the insulation quality if PD activity can be detected in its early stage.

Accordingly, this paper proposes a new method for power transformer partial discharge fault diagnosis. And this paper content is arranged as such. Section 2 summarizes related background information about how PD fault is diagnosed and states current challenges of PD signals feature extraction.
2. Overview of PD Fault Diagnosis

Transformer PD signal is of strong nonlinearity and time-variation. And in the in situ detection, it is frequently subject to the overlap of many interference signals [9, 10]. Statistics [11] indicate that major interference signals are discrete spectral interferences, stochastic shaped interferences, and periodic pulse shaped interferences, which leads great difficulties to signals feature extraction and fault diagnosis. Therefore, how to accurately extract signals feature is the key to PD faults identification.

Up to now, feature extraction methods mainly concentrate on the analysis of statistic graphic spectrum and wave shapes [12, 13]. Among them, methods based on statistic graphic spectrum analysis require the high sampling rate and a large number of samples, so it is unfavorable for online detection of insulation, whereas methods based on wave shape analysis require the low quantity of data samples. So it used to extract the time-variant signal feature, whose main methods are time-frequency analysis, Hilbert-Huang transform (HHT), wavelet theory, fractal theory, Chaos theory, and so forth [14–16]. Frequency and time features are important character for signals. Those can be obtained by using fast Fourier transform algorithm (FFT) or short time Fourier transform algorithm (STFT). However, both FFT and STFT are not suitable for time-variation signals. Because HHT has better self-adaptability, it is suitable for analyzing the partial dynamic feature of nonsmooth signals such as literatures [17] and [18]. But HHT requires a priori knowledge about the sensitive frequency bands of fault signals. Moreover, there are the end effects and the mode mixing in HHT. It is pointed out in literatures [19–21] that the frequency spectrum envelop, time-frequency energy entropy information, and so forth are able to depict the features of nonstable time-varying signals but, in practical use, they are easy to be subjected to noise interference so that the effects are unrealistic.

Unlike classical time-frequency analysis methods using a series of sinusoidal functions to describe a signal, wavelet transform decomposes a signal into wavelet coefficients of various time scales. And it is considered as one of the most powerful techniques for faults signal denoising and extracting transient feature [22, 23]. For example, the feature extraction of PD signals is achieved based on the cross-wavelet transform and relevant coefficient matrix in literature [24]. Unfortunately, how to determine an optimal mother wavelet is a big challenge for faults signal feature extraction. In literature [25–27], the fractal theory is successfully used to extract the feature of GIS PD signals. However, how to select fractal element shapes is unclear, and fractal theory is still immature.

Because different insulation defect faults have different partial discharge principle and there are strong randomness and dispersion in its phase distribution, feature frequency, and pulse magnitude, those classical feature extraction methods are not well suitable for online PD faults diagnosis. Moreover, the same insulation defect fault has some similarities on its frequency spectrum envelop and the frequency band energy distribution, whose features have an obvious probability distribution in certain frequency band. So this frequency band can describe different faults signal feature.

The magnitude of energy and the complexity of frequency components are two different parameters, which can represent one frequency band from different angles. And literature [28] points out that sample entropy value magnitude can reflect the complexity of the system. Therefore, this paper uses sample entropy and energy to establish one feature plane to describe the essential feature information of the fault. So this paper proposes a new method for power transformer PD signal feature extraction and fault diagnosis based on multidimensional feature region. Firstly, in order to obtain components of raw signal on different frequency bands, EMD and HHT band-pass filter wave for PD signal is carried out. Secondly, the sample entropy value and the energy value of each frequency band component are calculated. And then the feature region is established by using the sample entropy value and the energy value of each frequency band component to describe multidimensional features of PD signal. Finally, partial discharge faults are classified by sphere-structured support vector machines algorithm.

3. Frequency Band Component Extraction

In order to better describe differences in each frequency band of fault signals, the frequency band component of raw signal should be extracted. Now, EMD and HHT band-pass filter are used to extract frequency band component.

3.1. EMD. EMD is served as a kind of self-adaptive decomposition algorithm without obtaining a priori knowledge of raw signals in advance. And it avoids the optimum base function selection problem of wavelet decomposition [29]; the decomposition process of any signals $s(t)$ is shown in Figure 1.

Analog signals given in literature [30] are shown in Figure 2. Taking these analog signals as the example, the analysis is carried out. In the literature, single exponent attenuation signal is shown as in formula (1), and single exponent vibration signal is shown as in formula (2):

$$s_1(t) = Ae^{-t/\tau}, \quad (1)$$

$$s_2(t) = Ae^{-t/\tau} \sin(2\pi f_s t). \quad (2)$$

It can be known from Figure 3 that after two types of signals are decomposed by EMD, 7-order intrinsic mode function (IMF) components (imf1, imf2, ..., imf7) and 1 residual component (res.) are obtained. Then Hilbert time-frequency distribution of the 2nd-order IMF component is calculated. The result is shown in Figure 4.
It can be known from Figure 4 that the 2nd-order IMF frequency band of single exponent attenuation PD signals is 3.8–4.6 MHz, while the 2nd-order IMF frequency band of single exponent vibration PD signals is 2.2–4.2 MHz. Accordingly, the frequency band range of the same order IMF decomposed by different signals is different.

3.2. Extract Different Frequency Band Components of Raw Signal. The order of IMF obtained from PD signals through EMD decomposition is closely related to partial discharge type. And even if they are at the same order of IMF, their frequency band still has the obvious differences.

Aiming at this phenomenon, it is here that HHT band passing filter is adopted to solve the problem. The concrete procedures are as follows:

(1) Sampling frequency of signals $s(t)$ is $F_s$. The number of sampling points is $n$. After EMD decomposition, the $m$ pieces of IMF components are obtained. The magnitude of each component Hilbert time-frequency spectrum is $m \times n$, which should be classified as $l$ frequency bands and the set of instant amplitude values on the $k$ frequency band can be marked as $H(k)$ ($k = 1, 2, \ldots, l$).

(2) Zero set all the instant amplitude values on the Hilbert time-frequency spectrum except for frequency band $k$ and mark as $H^*(k)$.

(3) Zero set the IMF component point values which is corresponding to the instant amplitude value being set as zero in $H^*(k)$.

(4) Each IMF component treated using the filter can be reconstructed to obtain signals $s(t)$ component on the frequency band.

(5) As to the rest of frequency bands, the above procedures are repeated. That is, signals $s(t)$ component on the different frequency bands should be obtained.

Signals $s(t)$ are treated through HHT band passing filter in order to obtain $l$ components. And matrix $A$ is constructed.
using $l$ components. In other words, matrix $A$ can cover the
information of signal $s(t)$ under each frequency band:

$$
A = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{l1} & a_{l2} & \cdots & a_{ln}
\end{bmatrix},
$$

(3)

where $a_{ij}$ represents raw signals $s(t)$ data on each frequency band.

4. PD Signals Feature Region

4.1. Multidimensional Energy Parameters. Different insulation
defect faults have different partial discharge principle. And their
discharge pulse has obvious difference on the wave shape and the frequency band energy distribution, while the same insulation defect fault has strong similarity on the wave shape and the frequency distribution [31]. For this reason, this paper suggests that multidimensional energy parameters are used to describe PD signals feature quantity. The calculation process of multidimensional energy parameters is given as follows:

(1) Each row datum of time-frequency matrix $A$ “$a_{i1}, a_{i2}, \ldots, a_{in}$” stands for an independent frequency band, being marked as “$X_k, k = 1, 2, \ldots, l$.”

(2) Energy value $E(k)$ of each frequency band is calculated. And then the normalization of $E(k)$ is carried out:

$$
E(k) = \sum_{i=1}^{n} X_k(i)^2, \quad E^*(k) = \frac{E(k)}{\sum_{k=1}^{l} E(k)}.
$$

(4)

The multidimensional energy parameters of PD signals obtained from the above steps are marked as $E = [E_1^*, E_2^*, \ldots, E_l^*]$.

4.2. Multidimensional Sample Entropy Parameters. Sample entropy algorithm is a kind of theory originating from nonlinear dynamics [28], whose value magnitude can reflect the complexity of the system. It is likely to provide a kind of fresh solution for the nonstable signals analysis.

Sample entropy of each row datum in time-frequency matrix $A$ is calculated, whose steps are as follows:

(1) It is necessary to determine the dimensional number $m$ and the threshold value $r$. How to choose these parameters is discussed in [32]. In this paper, $m = 2$ and $r = 0.2$. Row data in the time-frequency matrix $A$ is marked as “$x_1, x_2, \ldots, x_n$.” And then “$x_1, x_2, \ldots, x_n$” should be converted into one group $m$ dimensional vectors; see the following equation:

$$
X(i) = [x(i), x(i+1), \ldots, x(i+m-1)],
$$

(5)

where $i = 1, 2, \ldots, n - m + 1$.

(2) The maximum difference value between the corresponding elements from $X(i)$ and $X(j)$ is defined as the distance between $X(i)$ and $X(j)$; namely,

$$
d(i,j) = \max_k |x(i+k) - x(j+k)| \quad (k = 0, 1, \ldots, m - 1).
$$

(6)

(3) According to the given threshold $r$, count the number of $d(i,j)$ which makes $d(i,j) < r$ the ratio of this number and the total vector $n - m + 1$ is defined as $B^m(r)$:

$$
B^m(r) = \frac{N}{n-m+1},
$$

(7)

where $1 \leq j \leq n - m$ and $j \neq i$.

(4) Take the average of $B^m(r)$ for every $i$, which is denoted by $B^m(r)$:

$$
B^m(r) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} B^m_i(r).
$$

(8)

(5) Change the embedding dimension from $m$ to $m+1$, repeating Step (1) to Step (4), and then $B^{m+1}(r)$ is obtained.

(6) The sample entropy of this row data in the time-frequency matrix $A$ is obtained:

$$
\text{SampEn}(m, r, n) = -\ln \frac{B^{m+1}(r)}{B^m(r)}.
$$

(9)

(7) Repeating Step (1) to Step (6) for the rest of other row data in the time-frequency matrix $A$, the multidimensional sample entropy parameters of PD signals can be obtained. Then, it is marked as follows:

$$
P = [p_1, p_2, \ldots, p_l],
$$

(10)
Empirical mode decomposition, $\psi/2\pi = 1$

Figure 3: The PD analog signals of EMD decomposition results. Note: in the figure, the scale of the vertical axis (imf1, imf2, ..., imf7) is the same as the raw signal; the horizontal axis does not have unit, the numerical meaning only on behalf of the sample points.

where $p_i (i = 1, 2, ..., l)$ represents the sample entropy of the row $i$.

4.3. PD Signal Feature Region. A feature plane is established to describe the PD signals feature from different dimensions. This feature plane consists of the multidimensional energy parameters and the multidimensional sample entropy parameters.

The energy parameter $E(k)$ and the sample entropy parameter $p(k)$ of the same frequency band would be defined as one feature point $s_k = (p(k), E(k))$. And then all the feature points of PD signals are marked as one set $S$; namely,

$$S = \{(P(1), E(1)), (P(2), E(2)), \ldots, (P(l), E(l))\}. \quad (11)$$

Finding a minimum circle $(O, R)$ on the feature plane includes all elements of the feature point set $S$. And then
5. Fault Diagnosis Based on Sphere-Structured Support Vector Machine

Support vector machine algorithm is a new machine learning method developed on the statistical study theory. It avoids the network structure selection, overlearning and underlearning, and other problems in artificial neural network algorithm. However, standard SVM is a binary classifier so that it cannot effectively solve multiclass classification [33]. Therefore, to overcome the shortages of standard SVM classifiers, many researchers tried to modify and improve SVM, such as “one-against-one” [34], “one-against-all” [35], and “decision tree” [36] which is suitable for multiclass classification. But the essences of these methods need to solve a large number of quadratic programming problems [37]. In order to reduce computational complexity, this paper suggests using sphere-structured SVM to deal with the multiclass classification.

One multiclassification problem is expressed as follows: $k$ several $n$ dimensional space element set $A^m, m = 1, \ldots, k$. Each set $A^m$ stands for the one element, including $l$ several sample feature points $x_i, i = 1, \ldots, l$. As far as possible, to find a spherical surface includes all the elements of the set $A^m$. This circle $(O, R)$ is defined as the feature region of PD signals on the feature field. It describes the PD signals feature information on the different frequency bands. Its calculation process is shown in Figure 5. And the feature region of fault signal is shown in Figure 6.

![Diagram](image-url)
In order to avoid some rough points impact on the algorithm, it is just here that a slack variable $\xi^m$ is introduced; see the following equation:

$$\|x_i - a^m\|^2 \leq R^m + \xi^m, \quad \xi^m \geq 0, \quad i = 1, \ldots, l,$$

(12)

where $a^m$ represents the center of sphere and $R^m$ represents the radius of sphere.

The objective function of the above-mentioned problems should be defined as

$$\min F (R^m, a^m, \xi^m) = R^m + C^m \sum_{i} \xi^m,$$

(13)

where $C^m$ represents the penalty coefficient.

Each classification can be described as similar to the quadratic programming problem. Solving this quadratic programming problem can obtain one sphere. And this sphere represents this class. Points on the spherical surface play a key role in spherical determination, called the support vector, as shown in Figure 7.

One set $X = \{x_1, x_2, \ldots, x_l\}$ represents one testing sample. $x_i$, $i = 1, 2, \ldots, l$, represents one feature point of the sample. To judge which class sample $X$ belongs to, the process is as follows.

Firstly, it is necessary to calculate $\sum_{i=1}^{l} \|x_i - a^m\|^2$, which describes the quadratic sum of distance between $x_i$ in the sample $X$ with each spherical center. And then, to compare the magnitude of $\sum_{i=1}^{l} \|x_i - a^m\|^2$ with $l \cdot R^m$, let $d$ represent the number of $\sum_{i=1}^{l} \|x_i - a^m\|^2 \leq l \cdot R^m$.

1. For $d = 0$, this means that most of feature points of sample $X$ are not located in any sphere. So it is necessary to find out one sphere which is the nearest to the sample $X$. If $\sum_{i=1}^{l} \|x_i - a^m\|^2 - l \cdot R^m = \min \sum_{i=1}^{l} \|x_i - a^m\|^2 - l \cdot R^m$, the sample $X$ belongs to class $p$.

2. For $d = 1$, the sample $X$ belongs to the class that this sphere represents.

3. For $d > 1$, this means the sample $X$ is located in multiple spheres intersection area. Then, taking Figure 8 as the example, consider the following:

1. All spheres including sample $X$ are marked as one set $I$.
2. $\forall i, j \in I$, calculate $\vec{a}^i D$ and $\vec{a}^i C$, respectively, where $\vec{a}^i D$ represents the projection of $\vec{a} x$ in $\vec{a}^i a^j$ and $\vec{a}^i C$ represents the projection of $\vec{a} A$ in $\vec{a}^i a^j$:

$$a^i D = \frac{\vec{a} x \cdot \vec{a} a^j}{\|\vec{a}^i a^j\|},$$

$$a^i C = \frac{\vec{a} A \cdot \vec{a} a^j}{\|\vec{a}^i a^j\|}.$$

(14)

3. Compare the magnitude of $a^i D$ with $a^i C$, for $a^i D < a^i C$; the sample $X$ belongs to class $i$, or else the sample $X$ belongs to class $j$.

6. Experimental Analysis

Based on power transformer structure and the different discharge forms of the different insulation defect, transformer partial discharge forms can be divided into the following three types: insulation internal defect (e.g., there are bubbles in insulating oil), surface discharge (e.g., the insulator surface flashover phenomenon), and electrode tip discharge (e.g., the winding tip discharge). The in situ data of 330 kV transformer stations in Gansu province, China, are taken as a real example for carrying out the analysis. Figure 9 shows the site condition of the 330 kV transformer stations. Figure 10 is a set of sample signals of different faults.

1. Extract UHF PD Signal Feature Region. EMD and HHT band-pass filters are carried out for different faults UHF PD signals. And then the PD signals components on different
frequency bands should be restructured. Figure 11 shows the PD signals after filtering and reconstruction.

Since the strongest noise in transformers station mainly concentrates into 10 kHz, the transformers iron-core magnetic noise mainly concentrates into the range of 10–70 kHz [38], so the paper chooses the subband over 1 MHz to be served as the feature subband. The signals frequency band energy and sample entropy parameters on each subband are calculated. And then the feature filed is established to extract the feature region PD signals. Figure 12 shows the feature region within 90% confidence interval of different UHF PD signals. Among them, 1# represents the feature region of the surface discharge defect, 2# represents the feature region of the electrode tip discharge defect, and 3# represents the feature region of the insulation internal defect.

It can be seen from Figure 12 that there exist obvious differences in the feature region of the different defects. The differences mainly focus on the different frequency bands energy value and sample entropy value. The sample entropy value of 2# and 3# is greater than 1#. It indicates that the discharge mechanism of 2# and 3# is more complicated and more random than 1#. The signals energy value of 1# and 2# is greater than 3#. It indicates that UHF PD signals amplitude of 1# and 2# is bigger than 3#. And their discharge process can release more energy causing more serious damage to the power transformer insulation.

(2) Partial Discharge Fault Diagnosis. The sphere-structured support vector machine algorithm is used to identify different partial discharge faults caused by different defects. Accordingly, 20 groups faults sample data are randomly selected to
carry out the training. The rest of the faults sample data are mixed randomly. Through many time testing experiments, the final statistic results are shown in Table 1.

It can be seen from Table 1 that the feature region of PD signals is able to describe the feature information of different partial discharge faults. And it is of better fault identification resolution. The average identification rate of different faults is more than 95%.

Then, two popular approaches of faults classification [39], PCA-SVM [40] and wavelet-SVM [41], are used to make a contrast analysis.

As one can notice from Table 2, the fault classification accuracy of PCA-SVM is lower. The fault classification accuracy of wavelet-SVM and the approach proposed in this paper is almost the same. So PCA is unsuitable for feature extraction of time-variation signals. The wavelet approach can better analyze time-variation signals, but how to determine an optimal mother wavelet is a big challenge for faults signal feature extraction. On the whole, the method proposed in this paper has more advantages in feature extraction of time-variation signals and faults classification.

7. Conclusion

Power transformer PD signals contain a large amount of insulation state information of power transformer. Effectively extracting signals feature information is of great significance for power transformer insulation online monitoring. In order to effectively extract PD signals feature, this paper suggests a new method. Firstly, using UHF PD signals, different frequency band components construct the time-frequency matrix. And then signals feature region is established by multidimensional energy parameters and multidimensional sample entropy parameters. It is confirmed in Section 6 that this approach can describe PD signals feature information. Finally, spherical structure support vector machine algorithm is used to identify different partial discharge faults.

Competing Interests

The authors declare there are no competing interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51279161). Furthermore, the authors are grateful for the staff of China Electric Power Research Institute and State Grid Gansu Province Electric Power Research Institute.

References


