Ultra-High Frequency Partial Discharge Signal Recognition in GIS based on Fisher Linear Discriminant Theory

Pei Xudong¹*, Pei Xubing², Xu Haiqi³

¹Gansu Electrical Apparatus Research Institute, Lanzhou, 730000, China;
²CGN New Energy Holdings Co., Ltd, Shenzhen, 518000, China;
³State Grid Urumqi Power Supply Company, Urumqi, 830011, China
*peixudong@tsccs.com
*Corresponding author’s e-mail: peixd_2020@163.com

Abstract. In this paper, dual tree complex wavelet transform (DT-CWT) is used to decompose UHF partial discharge (PD) signal in multi-scale. The optimal decomposition level of complex wavelet is solved. The wavelet energy features of real part and imaginary part of UHF-PD signal under the optimal decomposition scale are extracted. Fisher linear discriminant method is used to select the features of energy features and to identify the PD types. The results show that the optimized real part and imaginary part high frequency wavelet energy features can effectively identify four kinds of typical insulation defects, and the recognition rate can reach to 92.5%. Meanwhile, the optimal complex wavelet energy (OCWEF) feature has better sensitivity and recognition effect in PD type identification.

1. Introduction

The health state of GIS can be controlled by partial discharge (PD) detection of gas switchgear (GIS)[1-2]. The extracted PD signal features include statistical features, digital image features, waveform features, etc. [3], which can effectively identify the internal insulation defects of GIS. The feature extraction of PD waveform requires little computation, and no phase information is required to describe the features of PD signal directly, which has become a common feature parameter for PD identification.

However, wavelet transform has excellent time-frequency domain multi-scale decomposition ability of waveform, which has been widely applied in mechanical [6], biomedical [7], electrical [4-5] and other fields, and is the most widely used PD waveform analysis method at present.

At present, some scholars have extracted such statistical features as skewness, steepness and moment of the wavelet coefficient of PD signal for PD identification [10]. However, the proposed small baud feature has large redundancy and unclear physical meaning. The multi-level decomposition of PD signals was carried out, and the energy characteristics of the high and low frequency layers were extracted. However, it is simply believed that the higher the energy extreme value is, the more PD information it represents, and a lot of energy information is discarded without a detailed analysis of the ability to characterize the energy characteristics of the high and low frequency layers. It is proposed to extract the amplitude and frequency characteristics of PD signal by cross wavelet transform to identify transformer internal faults, but the problem of wavelet decomposition scale selection and redundant data processing is not solved [8]. The number of wavelet packet decomposition layers is optimized, but the...
extracted energy entropy of 32 sub-bands is used as the characteristic parameter of PD type recognition, and the data redundancy is large[9].

In this paper, the dual-tree complex wavelet transform (DT-CWT) is used to decompose ultra-high frequency (UHF) PD signal. A wavelet energy entropy is constructed, and fisher linear discrimination method is used to extract the energy spectrum characteristics of the optimal real and imaginary parts in the optimal decomposition scale, and the optimal complex wavelet energy features (OCWE) eigenvector is constructed. The experimental results show that four typical defects in GIS can be identified effectively by using the optimal wavelet energy characteristics of DT-CWT high frequency layer in real and imaginary parts.

2. Feature extraction and feature selection

2.1. Determination of the number of decomposition layers

When the signal is decomposed by dt-cwt, as the number of decomposition layers increases, the more detailed and accurate the time-frequency decomposition of the signal is, and the richer the detailed information will be. However, the longer the number of decomposition layers, the shorter the length of the signal will be. When the number of decomposition layers increases and the increment of information resolved to the newly generated layer is basically unchanged, it is no longer necessary to continue decomposition of the signal. Therefore, when the energy entropy of PD signal does not increase with the number of dt-cwt decomposition layers or the increase is not obvious, it means that the number of decomposition layers is optimal.

$$R_{error} = \frac{S_j - S_{j-1}}{S_{j-1}} \leq \varepsilon$$

where $S$ is the total energy entropy; $j$ is the number of decomposition layers; $R_{error}$ is the energy entropy error.

Generally, the smaller the value of $\varepsilon$ is, the more decomposition layers are, the more detailed information about the signal is resolved. However, the more decomposition layers are, the more data processing capacity is.

2.2. Feature extraction

Due to the different distribution of electric field under different defects in GIS, the time-frequency information of UHF PD signals excited by this is also different, and the discharge mechanism of the same type of defects has similar characteristics. Therefore, the characteristics of UHF PD signals at different decomposition scales can be used to identify the types of internal insulation defects in GIS. Assuming the optimal decomposition scale is $J_{best}$, then the energy matrix of UHF PD signal at the optimal scale can be expressed by

$$E = \begin{bmatrix}
E_{Re(1,1)} & E_{Re(2,1)} & \cdots & E_{Re(J_{best},1)} & E_{Re(J_{best},0)} \\
E_{Im(1,1)} & E_{Im(2,1)} & \cdots & E_{Im(J_{best},1)} & E_{Im(J_{best},0)} \\
E_{Z(1,1)} & E_{Z(2,1)} & \cdots & E_{Z(J_{best},1)} & E_{Z(J_{best},1)}
\end{bmatrix}$$

where $E_{Re(j,1)}$ represents the real part energy of the low frequency layer with complex wavelet coefficient, $E_{Im(j,1)}$ represents the imaginary part energy of the low frequency layer with complex wavelet coefficient, $E_{Z(j,1)}$ represents the total energy of the low frequency layer with complex wavelet coefficient.

2.3. Feature selection method based on Fisher's linear discrimination

When using PD signal features for pattern recognition, because different feature parameters have different ability to describe the features of different defects, and the same feature parameters may have different ability to describe different defects, the resolution ability of each feature to defects is different. If the whole feature is selected as the input of the recognizer, it will inevitably cause the dimension
disaster of the recognizer. Therefore, in order to improve the efficiency of defect diagnosis, PD characteristic parameters should be optimized, and the characteristic parameters with strong resolution and great contribution can be selected.

Fisher linear discriminant is a commonly used feature selection method. Its basic principle is to get the best discrimination vector by selecting Fisher linear discriminant function, and map the sample data to the best discrimination vector, so as to get the maximum and minimum class dispersion. Finally, the discriminant function values of different characteristic parameters (J) are calculated and sorted, as shown in formula (3) - (4).

\[
J(f_i) \geq J(f_j) \geq \cdots \geq J(f_n)
\]

\[
J_f(i, j) = \frac{|\mu_{i,j} - \mu_{j,i}|^2}{\sigma_{i,j}^2 + \sigma_{j,i}^2}
\]

where \( f_n \) is the feature subset and \( n \) is the dimension of the feature vector. In formula (4), \( i \) and \( j \) represent two different sample categories, \( \mu_{i,j} \) and \( \mu_{j,i} \) represent the mean value of feature subset \( f_l \), \( \sigma_{i,j}^2 \) and \( \sigma_{j,i}^2 \) represent the variance of feature subset \( f_l \).

There are many kinds of internal insulation defects in GIS. It is necessary to combine two types of defects and construct "fault pair" to optimize the characteristics by Fisher linear discriminant method. Assuming that there are \( k \) types of typical defects in GIS, the "fault pair" can be expressed as:

\[
\Theta(i, j) = \{(i, j) | i = 1, 2, \ldots, k-1; j = i+1, i+2, \ldots, k\}
\]

According to equation (3), the relation between the value of discriminant function (J) corresponding to \( n \) characteristic parameters is

\[
J_f(i, j) \geq J_f(i, j) \geq \cdots \geq J_f(i, j) \geq \cdots \geq J_f(i, j)
\]

Finally, \( d \) characteristic parameters with larger value of discriminant function are selected as the optimal feature subset of "fault pair" \( \Theta(I, J) \) can be represented as

\[
F_{ij} = \{f_l | l = 1, 2, \ldots, d\}
\]

Therefore, the final feature library \( F_{\text{final}} \) is

\[
F_{\text{final}} = \left\{ \bigcup_{i=1}^{l-1} \bigcup_{j=1}^{l} F_{i,j} \right\}
\]

3. Experimental data collection and PD defect identification

3.1. Test platform and data collection

Long term operation experience shows that there are four kinds of defects in GIS, i.e. metal protrusion defect (N (needle)), Metal contamination on insulator surface (M (metal)), air gap defect (G (GAP)) and free metal particle defect (P (particle)). Therefore, four kinds of defect models are designed, according to these four kinds of defect characteristics, as shown in Fig.1.
In order to capture the PD signal of defect model, the experimental circuit is set up, as shown in Fig. 2. A Tektronix 7100 oscilloscope with a maximum sampling rate of 20 GS/s and storage depth of 48 MB is utilized. Also a UHF sensor with 300–800 MHz bandwidth is configured.

In the experiment, four kinds of artificial defects were placed on 126kV GIS partial discharge experiment simulation platform and filled with 0.4mpa SF6 gas. The sampling rate was set at 5Gs/s, and the signal length was set at 50,000 sampling points. In the experiment, 50 groups of UHF PD signal effective discharge samples were collected under different discharge conditions (different voltage and defect model size), and 300 groups of discharge sample data were collected for each defect type.

3.2. Energy feature extraction

According to the curve of energy entropy of UHF PD signal with decomposition layers under various defects, it makes $\varepsilon = 0.03\%$. When the decomposition layers are equal to 8, $R_{error} \leq \varepsilon$. Therefore, this paper takes $J_{best}=8$.

Using DT-CWT to decompose the UHF PD signal in 8 layers, the 8-layer high frequency coefficient and 1-layer low frequency coefficient are obtained, and the energy matrix is E. As shown in Figure 3, the PD signals of high frequency layer (1-8 layers) and low frequency layer obtained from 8 layers of defect decomposition of G, M, N and P. As can be seen from Fig.3, the UHF PD signal of different defect types has different energy distribution at each decomposition scale after multi-level decomposition. Therefore, we can extract wavelet energy features of different scales to identify PD defect types. In addition, it is found that the energy value of four UHF PD signals in the low-frequency layer is far greater than that in the high-frequency layer. When the high-frequency energy and the low-frequency energy are taken as features at the same time, the low-frequency energy features will weaken the feature representation ability of the high-frequency layer, resulting in the feature representation ability of the high-frequency layer energy features can not be fully expressed, and the classifier can not fully embed the low-frequency and high-frequency energy features into the network training model, and more feature information of high frequency layer is lost.

![Fig.2 Schematic diagram of PD experiment](image)

(a) G-type defect
Fig. 3 PD waveform in each level after DT-CWT

Fig. 4 shows the comparison curve of real part energy feature and virtual part energy feature in 95% confidence interval. From Fig. 4, it can be seen that the energy distribution of PD signals of four typical defects on each layer is different after DT-CWT decomposition. The energy distribution of real part and imaginary part of M-type defects and P-type defects in the first three high-frequency layers shows a trend of attenuation, but the energy decay rate of P-type defects is low, and the energy of the upper layer is smaller than that of the upper layer, while the ability of M-type defects in the fifth layer is higher than that of the fourth and sixth layers. For class G defects, the energy distribution of real part and imaginary part fluctuates in different scales. Among them, the first, second, seventh and eighth layers of the virtual part have higher energy, while the first, second, sixth, seventh and eighth layers of the real part have higher energy, and the maximum energy of the virtual part appears in the seventh layer, while the real part has the same energy and maximum energy in the second and seventh layers. The energy of the real part and the virtual part of the third layer of N-type defects are the largest, the energy of other layers is relatively small, and the energy difference between other layers is relatively small.
The energy distribution characteristics of each scale represent the energy distribution characteristics of UHF PD signal in each frequency domain. For the frequency band with large energy distribution, the energy of the real part and the virtual part are also large in general. The difference between the real part and the imaginary part energy of the same defect layer is due to the fact that DT-CWT uses orthogonal wavelet basis to analyze PD signal. The difference between the real part and the imaginary part energy of the same defect layer is due to the fact that DT-CWT uses orthogonal wavelet basis to analyze PD signal, and the difference between real part and imaginary part energy is due to the different distribution of wavelet energy in two orthogonal directions.

### 3.3. Energy feature selection

Six pairs of "fault pairs" which are \( \Theta(G, M) \), \( \Theta(G, N) \), \( \Theta(G, P) \), \( \Theta(M, N) \), \( \Theta(M, P) \) and \( \Theta(N, P) \) respectively can be constructed by feature selection with Fisher linear discriminant method. Because the energy value of the low-frequency layer is far greater than that of the high-frequency layer, the energy characteristics of the low-frequency layer will reduce the feature representation ability of the high-frequency layer, lose the feature information of the high-frequency layer, and reduce the identification rate of the PD. Therefore, this paper uses the energy characteristics of the high-frequency layer of the PD signal as the analysis object, as shown in Tab.1.

**Tab.1 The energy feature space of UHF PD signals**

| Wavelet tree | Characteristic parameter |
|--------------|--------------------------|
| imaginary part | \( E_{\text{Im}(1, 1)} \sim E_{\text{Im}(8, 1)} \) |
| real part | \( E_{\text{Re}(1, 1)} \sim E_{\text{Re}(8, 1)} \) |

The Fisher linear discriminant method is used to optimize the high frequency energy features of DT-CWT, specifically selecting three feature quantities with larger discriminant function value (J) under each "fault pair", constructing the final feature set (Ffinal), and merging them. The feature selection results are shown in Tab.2.

**Tab.2 The energy features of the imaginary and real part after selection**

| Fault pairs | Energy characteristics of imaginary part | Energy characteristics of real part |
|-------------|-----------------------------------------|-----------------------------------|
| \( \Theta(G, M) \) | \( E_{\text{Im}(1, 1)} \) | \( E_{\text{Im}(2, 1)} \) | \( E_{\text{Im}(3, 1)} \) | \( E_{\text{Im}(5, 1)} \) | \( E_{\text{Im}(7, 1)} \) | \( E_{\text{Im}(8, 1)} \) | \( E_{\text{Re}(1, 1)} \) | \( E_{\text{Re}(2, 1)} \) | \( E_{\text{Re}(3, 1)} \) | \( E_{\text{Re}(5, 1)} \) | \( E_{\text{Re}(7, 1)} \) | \( E_{\text{Re}(8, 1)} \) |
| \( \Theta(G, N) \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) |
| \( \Theta(G, P) \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) |
| \( \Theta(M, N) \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) |
| \( \Theta(M, P) \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) |
| \( \Theta(N, P) \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) |

\( F_{\text{final}} \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \) | \( \checkmark \)
3.4. Identification of defect types

Support vector machine (SVM) is a machine learning method based on VC dimension theory and SRM principle, which is especially suitable for pattern recognition in high dimension space. In this paper, "one-to-one" SVM is selected to identify the UHF PD signals of four typical defects. Radial basis function (RBF) is chosen as the kernel function of SVM. The RBF function adopts the system default value. In this paper, 100 groups of data are randomly selected from four defect feature sets as training samples, and the rest 200 groups of data are selected as test samples. As a contrast, this paper simultaneously extracts (real wavelet transform, RWT) energy feature vector, high-frequency and low-frequency energy (H-LE) vector of DT-CWT real part and imaginary part for PD identification. The identification results are shown in Tab.3.

| Defect types | RWT          | OCWE         | HLE          | DT-CWT       |
|--------------|--------------|--------------|--------------|--------------|
| G            | 67.50%       | 96.00%       | 83.00%       |              |
| M            | 77.50%       | 92.50%       | 93.00%       |              |
| N            | 84.50%       | 96.50%       | 92.50%       |              |
| P            | 81.50%       | 98.00%       | 89.00%       |              |
| Average value| 77.75%       | 95.75%       | 89.375%      |              |

According to Tab.3, DT-CWT energy feature has a higher recognition rate than RWT energy, and the overall recognition rate reaches 95.75%. Compared with the PD identification using DT-CWT to extract complex wavelet energy features at all scales in the literature, the OCWE feature of high frequency layer proposed in this paper has a better recognition rate, especially for G-type defects, the recognition rate of this method reaches 96.00%, except for M-type defects, the recognition rate of other three types of defects is higher than the PD recognition rate using low frequency layer and high frequency layer features at the same time. In addition, for G defects, when the low-frequency energy feature is added, the recognition rate is significantly reduced to 83%; while for M defects, the recognition rate increases slightly. The reason for the high recognition rate of feature redundancy is that after multi-scale decomposition of UHF PD signal, the wavelet energy resolved to the low-frequency layer is far greater than the energy value of the high-frequency layer, so that the energy features of the high-frequency layer can not be fully expressed in SVM classification model. In addition, from the frequency domain analysis, the details of UHF PD signals generated by different defects are included in the energy distribution of signals in each frequency band. After entering the multi-level DT-CWT, the details of UHF PD signals are gradually resolved to each high-frequency layer, while the low-frequency layer contains the contour information of UHF PD signals.

Fig.5 shows the receiver operating characteristic curve (ROC) for PD type diagnosis with three different characteristics (RWT, OCWE and HLE). The area under the ROC curve (AUC) of each feature is 0.979, 0.931 and 0.867, respectively. It can be seen that the OCWE feature adopted in this paper has better sensitivity and can identify PD defect types more efficiently.
4. Conclusion

In this paper, the optimal scale decomposition of UHF PD is carried out, and the energy features of each scale are extracted, and Fisher linear discriminant method is used to optimize the energy features. Finally, SVM classifier is used to classify PD signals under four kinds of insulation defects. The results show that the energy feature optimized by Fisher linear discriminant method has a good effect on GIS internal insulation sink, and the overall recognition rate reaches 95.75%. Moreover, through ROC curve analysis, it is found that OCWE has better PD type discrimination sensitivity and better classification effect than RWT and HLE energy features. At the same time, there is no positive correlation between the classification ability of wavelet energy features and the size of energy features. The larger the energy features, the stronger the resolution ability is, and the better the effect of PD identification using high-frequency energy features is than that using low-frequency and high-frequency energy features at the same time.

References

[1] SCHAIK N V, CZASZEJKO T. Conditions of discharge-free operation of XLPE insulated power cable systems [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2008, 15(4): 1120-1130

[2] Hu Wentang, Gao Shengyou, YU Shaofeng, et al. Application of statistic parameters in recognition of partial discharge in transformers[J]. High Voltage Engineering, 2009, 35(2): 277-281.

[3] CHEN Pan, YAO Chenguo, LIAO Ruijin, et al. Application of signals separated band energy spectrum and mahalanobis clustering algorithm for switchgear partial discharge pattern recognition[J]. High Voltage Engineering, 2015, 41(10): 3332-3341.

[4] SHANG Haikun, YUAN Jinsha, WANG Yu, et al. Feature Extraction for Partial Discharge Based on Cross-Wavelet Transform and Correlation Coefficient Matrix[J]. TRANSACTIONS OF CHINA ELECTROTECHNICAL SOCIETY, 2014, 29(4): 274-281.

[5] TANG Ju, FAN Lei, ZHANG Xiaoxing, et al. Multi-scale feature parameters extraction of GIS partial discharge signal with harmonic wavelet packet transform[J]. TRANSACTIONS OF CHINA ELECTROTECHNICAL SOCIETY, 2015, 30(3): 250-257.

[6] SHI Mingjiang, LUO Renze, FU Yuanhua. Fault diagnosis of ratating machinery based on wavelet and energy feature extraction[J]. JOURNAL OF ELECTRIONIC MEASUREMENT AND INSTRUMENTATION, 2015, 29(8): 1114-1120.

[7] HU Wenlong, QIAO Xiaoyan. Ankle motion classification based on wavelet packet statistic features of sEMG signal[J]. JOURNAL OF TEST AND MENSUREMENT TECHNOLOGY, 2017, 31(2): 100-106.

[8] Umamaheswari R, Sarathi R. Feature Extraction of UHF PD Signals by Wavelet Packet Based MRSD Analysis For Defect Identification in Gas Insulated Systems[C]//India Conference (INDICON), 2012 Annual IEEE. IEEE, 2012: 1218-1222.

[9] Suryavanshi H, Velandy J, Surendran J. Wavelet Transform Coherence for Magnitude and Phase Spectrum Prediction From High Frequency Transient Signals: Partial Discharge in Transformers[C]. IEEE International Conference on Power Electronics, Drives and Energy Systems. India, 2016:1-6.

[10] Ming-Shou Su, Chung-Chu Chia, Chien-Yi Chen, et al. Classification of Partial Discharge Events in GILBS Using Probabilistic Neural Networks And the Fuzzy c-means Clustering Approach[J]. Electrical Power and Energy Systems, 2014, 61, 173-179.