LDA Optimized Multi-scale Texture Features Based Diagnosis Method of Defects inside Insulated Tubular Busbars

S Q Li\textsuperscript{1a}, L B Zhou\textsuperscript{1}, J H Liu\textsuperscript{1} and Y W Zhou\textsuperscript{2b}
\textsuperscript{1} Shanghai Maintenance Company, State Grid Corporation, Shanghai, 200240, China
\textsuperscript{2} Shanghai Judian Electric Technology Company Limited, Shanghai 200240, China
\textsuperscript{a}lisunqi@126.com; \textsuperscript{b}ywzhou01@163.com

Abstract. In this paper, a method for diagnosing defects inside insulated tubular busbars based on LDA optimized multi-scale texture features is proposed to help to guarantee stable operation of the tubular busbars and the whole power grid. Firstly, multi-scale PRPD spectrum space was built with the UHF discharge signals of different defects by image pyramid theory. Then first-order, second-order and higher-order texture statistics were extracted from each image in the multi-scale PRPD spectrum space to form multi-scale texture features and LDA algorithm was used to optimize the features. The method was used to make texture features contain more information about partial discharge and help to improve the accuracy of diagnosis. Experiments were conducted on a 40.5kV insulated tubular busbar and CART classification trees were established as a classifier to identify the types of defects. The results of experiments show that this method can identify the defects of insulated tubular busbars accurately.

1. Introduction
In the research of partial discharge features extraction and insulated defects identification of the electrical equipment, the selection of features and the method to extract features are important factors affecting the effect of defects diagnosis. At present, the most common features used for defect diagnosis include fractal features, statistical features, wavelet features, Weibull parameters, moment features and texture features of images \cite{1}. Among them, the texture features of images have the advantages strong resistance to external environmental interference, good discrimination, and high sensitivity so that texture features are widely used in defect diagnosis of power equipment \cite{2}. However, the diagnosis only based on single-dimensional texture features does not have enough accuracy which indicates that the message contained in single-dimensional texture features is insufficient to reflect the insulation state of power equipment. Therefore, a method for diagnosing defects based on LDA optimized multi-scale texture features is proposed in this paper to make texture features contain more information about partial discharge and help to improve the accuracy of diagnosis. Since the insulated tubular busbar is an important connecting conductor of the main transformer and the user side, its operation state directly affects the stability of the system and the reliability of the power supply \cite{3-4}. Using the method proposed in this paper can monitor the insulation state of the insulated tubular busbars in real time. When the defects are identified in the tubular busbar, it is convenient to arrange maintenance in the first time to guarantee the reliability and stability of the power supply.
2. Multi-scale texture features extraction

2.1. Gaussian pyramid

Image pyramid is a multi-scale representation technology of images [5]. The Gaussian pyramid is the most widely used in image processing, signal processing and other fields among all the image pyramids [6]. The process of Gaussian pyramid construction is as follows:

Assume the original image $G_0$ as the bottom of Gaussian pyramid. The $k$th layer of Gaussian pyramid can be calculated as (1). The $k-1$th layer makes convolution with a lowpass window $w(m, n)$ and the outcome of convolution needs downsampling.

$$G_k = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m, n) \ G_{k-1}(2l + m, 2j + n)$$

In which, $N$ is the number of the top of the Gaussian pyramid; $C_l$ is the column number of the $l$th layer and $R_l$ is the row number of the $l$th layer; $w(m, n)$ is a $5 \times 5$ size window function, shown as (2).

$$w(m, n) = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

After the steps above, Gaussian pyramid of the original image can be obtained. The number of pyramid layers selected in this patent is 4, that is, one PRPD spectrum can finally obtain four images including the original image. Considering the complexity of calculation, 4-layer-Gaussian pyramid is appropriate for feature extraction.

2.2. Multi-scale texture features

Texture features are widely used in the fields of scene target recognition, remote sensing image analysis, industrial detection, medical image analysis, text segmentation and other fields [6]. According to the order of texture features extracted from the target image, they can be divided into three categories: first-order, second-order and higher-order texture features.

2.2.1. First-order texture statistics. First-order texture statistics refer to the features extracted from feature histograms of the image. The grayscale histogram and the LBP feature histogram are selected and normalized to extract the first-order texture features. The extraction formulas for the normalized feature histogram are as follows:

$$d_1 = \sum_{i=0}^{255} iH(i)$$

$$d_2 = \sum_{i=0}^{255} (i - d_4)^2H(i)$$

$$d_3 = \frac{1}{d_2^2} \sum_{i=0}^{255} (i - d_4)^3H(i)$$

$$d_4 = \frac{1}{d_2^4} \sum_{i=0}^{255} (i - d_4)^4H(i) - 3$$
\[ d_5 = \sum_{i=0}^{255} H(i)^2 \]  
(7)

\[ d_6 = -\sum_{i=0}^{255} H(i) \log_2 H(i) \]  
(8)

where \( H(i) \) is the normalized feature histogram; \( d_1, d_2, d_3, d_4, d_5, d_6 \) are the mean, variance, skewness, kurtosis, energy, and entropy of the histogram respectively.

2.2.2. Second-order texture statistics. The second-order texture statistics are features extracted from image matrices such as co-occurrence matrix, gradient matrix, etc. The second-order texture features used in this paper are features extracted from Tamura texture, gray-scale co-occurrence matrix (GLCM) and Laws texture. For Tamura texture, features of roughness, contrast, and direction are extracted; for GLCM features, the extracted features are angular second-order moments (energy), contrast, correlation, and entropy [7,8]. According to several researches [9,10], the extraction steps of Laws can be concluded as follows:

1) Make mutual convolutions of 4 basic vectors representing 4 directions to form 16 texture templates [9]. The basic vectors are shown as (9). \( L_5, E_5, S_5 \) and \( R_5 \) represents level, edge, point and ripple features respectively. The obtained texture templates are listed in Table 1 [10].

\[
\begin{align*}
L_5 &= [1 \ 4 \ 6 \ 4 \ 1], \quad E_5 = [-1 \ -2 \ 0 \ 2 \ 1] \\
S_5 &= [-1 \ 0 \ 2 \ 0 \ 1], \quad R_5 = [1 \ -4 \ 6 \ -4 \ 1]
\end{align*}
\]  
(9)

| Table 1. Texture templates |
|---------------------------|
| \( L_5 \) | \( L_5 E_5 \) | \( L_5 S_5 \) | \( L_5 R_5 \) |
|\( E_5 L_5 \) | \( E_5 E_5 \) | \( E_5 S_5 \) | \( E_5 R_5 \) |
|\( S_5 L_5 \) | \( S_5 E_5 \) | \( S_5 S_5 \) | \( S_5 R_5 \) |
|\( R_5 L_5 \) | \( R_5 E_5 \) | \( R_5 S_5 \) | \( R_5 R_5 \) |

2) The target image \( G \) is convolved with 16 texture templates respectively to obtain 16 "micro windows" so that 16 texture images \( G_m \) (m represents the texture template used) are formed. Then make these texture images \( G_m \) pass through a "macro window" (size 15 × 15) to calculate the texture energy measurement (TEM) corresponding to each image. TEM is a matrix of image pixel dimensions and each element \( a(i, j) \) in the matrix can be calculated by (10). 16 TEMs were obtained for each of the texture images, of which 9 were TEMs with rotation invariance, denoted as \( R_{\text{tem}} \), and listed in Table 2.

\[
a(i, j) = \sum_{u=-7}^{7} \sum_{v=-7}^{7} |G_m(i + u, j + v)|
\]  
(10)

| Table 2. List of TEMs with rotation invariance |
|-----------------------------------------------|
| \( R_{\text{tem}}L_5E_5 \) | \( R_{\text{tem}}L_5S_5 \) | \( R_{\text{tem}}L_5R_5 \) |
| \( R_{\text{tem}}S_5R_5 \) | \( R_{\text{tem}}E_5S_5 \) | \( R_{\text{tem}}E_5R_5 \) |
| \( R_{\text{tem}}R_5R_5 \) | \( R_{\text{tem}}S_5S_5 \) | \( R_{\text{tem}}E_5S_5 \) |

3) Calculate the energy \( E_n \) and variance \( \text{Var} \) of \( R_{\text{tem}} \) by (11) and (12).
\[ E_n = \sum_i \sum_j a(i, j)^2 \]  
\[ \text{Var} = \frac{\sum_i \sum_j (a(i, j) - \sqrt{E})^2}{T} \]  

In which, \( i, j \) are the pixel coordinates; \( a(i, j) \) is the element in the \( R_{\text{tem}} \) matrix; \( E \) is the mean value of the pixels in the \( R_{\text{tem}} \), \( E = \frac{\sum_i \sum_j a(i, j)}{T} \); \( T \) is the total number of pixels.

2.2.3. Higher-order texture statistics. The higher-order statistics of images refer to higher-order moments, higher-order cumulants and their corresponding higher-order moment spectra and higher-order cumulant spectra, which are larger than the second-order statistics. Among these higher-order statistics, the k-order moments and k-order cumulants are selected to extract, denoted as \( p_k(\tau_1, \tau_2, ..., \tau_{k-1}) \) and \( c_k(\tau_1, \tau_2, ..., \tau_{k-1}) \), in which \( \tau_1, \tau_2, ..., \tau_{k-1} \) are delay sequences.

Considering the actual computational complexity, the third- and fourth-order statistics are extracted. Equations (13) and (14) are specific expressions for third-order moment and the third-order cumulant respectively. Equations (15) and (16) are specific expressions for fourth-order moments and fourth-order cumulants.

\[ p_3(\tau_1, \tau_2) = E[x(k)x(k + \tau_1)x(k + \tau_2)] \]  
\[ c_3(\tau_1, \tau_2) = E[x(k) - p_1][x(k + \tau_1) - p_1][x(k + \tau_2) - p_1] \]  
\[ p_4(\tau_1, \tau_2, \tau_3) = E[x(k)x(k + \tau_1)x(k + \tau_2)x(k + \tau_3)] \]  
\[ c_4(\tau_1, \tau_2, \tau_3) = p_4(\tau_1, \tau_2, \tau_3) - [p_2(\tau_1)p_2(\tau_3 - \tau_2) - p_2(\tau_2)p_2(\tau_3 - \tau_1) - p_2(\tau_3)p_2(\tau_2 - \tau_1)] \]  
\[ -p_1[3p_3(\tau_2 - \tau_1, \tau_3 - \tau_1) + 3p_3(\tau_2, \tau_3) + p_3(\tau_1, \tau_3) + 3p_3(\tau_1, \tau_2)] + 2p_1^2 \]  
\[ \cdot [p_2(\tau_1) + p_2(\tau_2) + p_2(\tau_3) + p_2(\tau_3 - \tau_1) + p_2(\tau_3 - \tau_2) + p_2(\tau_2 - \tau_1) - 6p_1^4] \]  

In which, \( x(k) \) is a one-dimensional sequence converted from a two-dimensional image; \( p_1 \) and \( p_2 \) are the first and second moments of the image respectively; \( E(\cdot) \) represents the mean value of the pixels in the matrix.

3. Linear Discriminant Analysis

While Multi-scale texture features can improve the description refinement of images, the large number of features will have a negative impact on the efficiency of subsequent pattern recognition operations. Linear Discriminant Analysis (LDA) algorithm can be used to reduce the dimensionality of multi-scale texture features in a supervised manner, helping classifiers identify the types effectively [11].

Assume the number of kinds of defects is \( c \) and the vector of texture features extracted is \( X_k = [x_1, x_2, x_3, ..., x_n] \) (\( k \) represents the sample number). The original dimension of texture features is \( n \). Divide whole samples into two kinds, training set and test set. Assume the number of samples in training set is \( M \) and the number of samples of defect \( i \) in training set is \( m_i \). Calculate the mean value of the samples of defect \( i \) [12]:

\[ u_i = \frac{1}{m_i} \sum_{X_k \in \text{class } i} X_k \]  

(17)

And the mean value of whole samples in training set is

\[ U = \frac{1}{c} \sum_{i=1}^{c} \sum_{X_k \in \text{class } i} X_k \]  

(18)

Calculate the divergence matrix between various types of defect samples \( S_b \) and the divergence matrix of each type of defect samples \( S_w \).
\[ S_b = \sum_{i=1}^{c} m_i (u_i - U) (u_i - U)^T \]

\[ S_w = \sum_{i=1}^{c} \sum_{X_k \in \text{class } i} (u_i - X_k) (u_i - X_k)^T \]

Set the projected line \( y = w^T x \) with the projecting direction \( w \) so original feature vectors can be projected to new vectors with less dimensions through \( w \). In order to make samples separate well, the distance of the projection points between classes of the samples must be maximized and the distance between the projection points of the same class of samples must be minimized. Therefore, Fisher criterion function is introduced:

\[ J(w) = \frac{w^T S_b w}{w^T S_w w} \]  

(21)

In which, \( J(w) \) is the objective function. In order to make the intra-class distance smaller and inter-class distance larger, \( J(w) \) should take the maximum which also makes \( S_b w = \lambda S_w w \). \( w \) is the eigenvector matrix to be solved. Assume the target number of dimensions is \( d \). Calculate all the eigenvalues of \( S_w^{-1} S_b \) and list them size down. The top \( d \) eigenvalues are selected and the corresponding eigenvectors are obtained to form the projecting matrix \( w \).

The original features of all the samples (including the test samples) can be transformed into the filtered feature vectors \( P_k \) through \( w \) expressed as

\[ P_k = w^T X_k \]  

(22)

4. Experiment

In this paper, UHF sensors were used to detect PD signals of insulated tubular busbars with three typical kinds of insulation defects: metal tip, internal defect and floating electrode. The experiment was conducted on a 40.5kV insulated tubular busbar. A total of 150 sets of data were collected, 100 for each defect. PRPD spectrum was made according to the collected UHF signals, shown as Figure 1. 20 sets of PRPD spectrum of each kind of defect were randomly selected as training set. The others were regarded as the test set.

![PRPD spectrum](image)

(a) Metal tip model (b) Internal defect model (c) Floating electrode model

Figure 1. PRPD of insulated tubular busbars with different typical defect

Then multi-scale texture features of each sample were extracted according to the method described in the paper. The total number of features is 164 which can be filtered to the most favourable features for the identification of insulated tubular busbars with LDA optimized features. The CART classification trees were established by the train set to identify the filtered multi-scale texture features and output the diagnosis results of the test set.

In order to verify the superiority of multi-scale texture features, first-order, second-order, and high-order texture features are separately applied to classify the defects with the CART classification trees. The accuracy of different texture features is listed in Table 3.
Table 3. The accuracy of different texture features

| Defect type          | Recognition accuracy/% |
|----------------------|------------------------|
|                      | First-order texture features | Second-order texture features | High-order texture features | Multi-scale texture features |
| Metal tip            | 78                      | 80                         | 82                         | 88                         |
| Internal defect      | 74                      | 74                         | 78                         | 86                         |
| Floating electrode   | 74                      | 78                         | 84                         | 92                         |

Comparing the recognition results in Table 3, it is obvious that the multi-scale texture features have higher recognition accuracy rate than using first-order, second-order, or high-order texture features alone, which indicates that multi-scale texture features are more suitable for defect diagnosis of the insulated tubular busbars. The results of experiments show that the extraction method of PRPD features proposed in this paper can help increase the accuracy of the defect diagnosis of insulated tubular busbars, which can improve the stable operation of the power system.

5. Conclusion
In order to improve the accuracy of defect diagnosis of insulated tubular busbars, a feature extraction method is proposed in this paper: Firstly, use UHF sensors to collect discharge signals of different defects to construct a PRPD spectrum and build multi-scale PRPD spectrum space based on the image pyramid theory. Then extract first-order, second-order and higher-order texture statistics from every image in the multi-scale PRPD spectrum space to form multi-scale texture features. After that, use LDA optimization method to refine the features more useful for defect classification. Finally, establish CART classification tree as a classifier to identify the types of defects of insulated tubular busbars. The results of experiments show that this method can accurately identify the defect of insulated tubular busbars.

References
[1] Tang J 2013 Research on Signal Noise Reduction and Characteristic Analysis Method in PD Detection Xidian University.
[2] Liu Li, Zhao Lingjun, Guo Chengyu 2018 Wang Liang, Tang Jun. Texture Classification: Status=of-the-art Methods and Prospects Acta Automatica Sinica 44 584-607.
[3] Ruan L, Li C, Yang F, Li P, Zhu S, Gao N, Jin H 2018 Development Status and Research Trend of Insulated Tubular Bus High Voltage Apparatus 54 43-53.
[4] Liu F, Xue Z, Deng Y, Ma Q 2017 Operating Characteristics and State Evaluation Methods for Insulated Tubular Bus-bar High Voltage Technology 43 4088-4095.
[5] Dai X, Li H, Yang H, Zhang J 2014 A Fast Image Enhancement Method by Virtual Image Pyramid Sequence Fusion Chinese Journal of Computers 37 602-610.
[6] Wang B, He Z 2015 Wu Xing, Jia Yuanyuan. New Diamond Search Algorithm Based on Gaussian Pyramid. Computer Engineering and Applications 51 174-178.
[7] Liu L, Kuang G 2009 Overview of Image Textural Feature Extraction Methods Journal of Image and Graphics 14 622-635.
[8] Gao C, Hui X 2010 GLCM-Based Texture Feature Extraction Application of Computer System 19 195-198.
[9] Qu Z, Zhou H 2015 Crack extraction algorithm combining texture feature and saliency Computer Engineering and Design 36 3056-3059+3081.
[10] Ge L, Zhu Q, Fu S 2009 Application of Laws’ Masks to Stereo Matching Acta Optica Sinica 29 2506-2510.
[11] Xu J, Wu Z, Xu Y, Zeng J 2019 Face Recognition Based on PCA, LDA and SVM Algorithms Computer Engineering and Applications 55 34-37.
[12] Zhang H, Niu Y 2018 WLAN Indoor Positioning Algorithm Based on Linear Discriminant Analysis and Gradient Boosting Decision Tree Journal of Instrumentation 39 136-143.