ORIGINAL RESEARCH PAPER

Method of GIL partial discharge localization based on natural neighbour interpolation and ECOC-MLP-SVM using optical simulation technology

Yiming Zang¹ | Yong Qian¹ | Hui Wang¹ | Antian Xu² | Gehao Sheng¹ | Xiuchen Jiang¹

¹Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
²Department of Light Sources and Illuminating Engineering, Fudan University, Shanghai, China

Correspondence
Yong Qian, Department of Electrical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road Minhang, Shanghai, 200240, China.
Email: qian_yong@sjtu.edu.cn

Associate Editor: Chaoyang Li

Funding information
Key Technology for Environmentally Friendly Gas-In, Grant/Award Number: 2017YFB0902500

Abstract
Partial discharge (PD) is one of the main causes rendering the deterioration of the insulation state in gas-insulated transmission line (GIL). Accurate and timely localization of the PD source is essential to ensure the safe and stable operation of the GIL. At present, optical PD detection technology shows advantages in terms of high sensitivity and strong anti-interference performance. Therefore, an optical PD localization method based on optical simulation fingerprint database is proposed. The introduction of simulation conquers the difficulty of obtaining a PD fingerprint database in field experiments. This method constructs a simulation fingerprint database by performing PD simulation in a GIL simulation model of the same size as the actual GIL. The natural neighbour interpolation algorithm is applied to expand the simulation fingerprint database to cover all locations in the GIL. The two-level localization method proposed is to match the PD fingerprint to be tested with the expanded fingerprint database, which can reduce the amount of calculation while maintaining the localization accuracy. The experimental results show that the average localization error of this method is only 9.7 mm, and the localization time is reduced by about 11 times compared with the normal one-level localization method.

1 | INTRODUCTION

To meet the large-scale power transmission need and reduce transmission loss, gas-insulated transmission line (GIL) has been widely utilized worldwide since the 1970s [1–3]. During the operation of the GIL, the defects and deterioration of the insulation pose a significant threat to its safety [4], while partial discharge (PD) serves as a harbinger and manifestation of insulation degradation [5, 6]. Therefore, the localization of PD in the GIL functions as an important measure to improve the efficiency of maintenance and ensure a stable operation.

Till date, the PD localization for the GIL has mainly been based on the detection of ultra–high frequency (UHF) signals and ultrasonic signals [7–9]. By contrast, the optical-only PD localization method performs with high sensitivity, no electromagnetic interference and no mechanical vibration interference [10, 11], which has promising research prospects and application value.

Nevertheless, little research on optical-only PD localization for the GIL and even other gas-insulated equipment could be found, and available ones mainly focus on two aspects of research. One is using an optical array for PD localization; yet, the optical array can only detect PD in a rather small area and does not work when confronted with the GIL with a long-distance structure [12]. The other is applying fingerprint database and machine learning into the optical PD localization. The Biswendu Chatterjee team conducted a series of optical PD experiments at different locations in the actual experiment device to assemble data. They put the optical PD fingerprint database containing location information into machine learning algorithms for training, which is used for the PD localization by fingerprints matching [13–15]. However, this method is limited due to several factors such as the huge size of the actual equipment and the complexity of the field situation. In this way, difficulty in obtaining training data for machine learning through a large number of field PD experiments could be seen. Moreover, this method can only identify the PD source generated at the specific location where the PD experiment was performed, which results in substantial impacts on localization accuracy.
Therefore, we propose a novel optical PD localization method based on natural neighbour interpolation (NNI) and the Error Correcting Output Codes-Multilayer Perceptron-Support Vector Machine model (ECOC-MLP-SVM) applying optical simulation technology, designed to improve the detection range, accuracy and efficiency of the GIL PD localization. This method builds a GIL simulation model in Tracepro software [16] that is exactly the same as the actual GIL size and sensor layout, which can adjust according to the actual size. In the simulation model, the optical PD simulation is performed to obtain optical PD fingerprints at different locations in the GIL, in order to get over the difficulty of obtaining PD data in field experiments. To a certain extent, the higher the density of the fingerprint the database shows, the higher the localization accuracy is. Due to the large size of the GIL, it is impractical to simulate the PD signals at all locations. Therefore, a fingerprint expansion method to expand the limited optical PD fingerprints to PD fingerprints at all locations in the GIL through the NNI algorithm is proposed, which will boost the database density and make up for the defect that only a few specific PD locations can be identified. Meanwhile, in order to eliminate the impact of the actual PD signal intensity changes, the principal component analysis (PCA) algorithm is used to extract the feature of the difference between each optical signal irradiance as the PD fingerprint feature.

In this process, however, the large size of the GIL and the expanded fingerprint database aiming to improve localization accuracy will drastically increase the number of learning samples, resulting in a significant drop in fingerprint matching efficiency. Therefore, we propose a two-level PD localization method to reduce the amount of calculation in the location process, solving the problem of the dimensionality curse caused by too many fingerprints. In the first level (pre-localization), the improved ECOC-MLP algorithm is used to match on a fingerprint database with a large range but low density to obtain the approximate area where the PD source is located. Since the ECOC-MLP algorithm can simplify the multi-classification problem into multiple two-classification problems, the dimensionality curse problem can be settled by using the two-classification method [17]. Furthermore, the low-density fingerprint database used in the first level reduces the amount of calculation rather than directly using a normal density fingerprint database for localization. In the second level (final localization), a small-range fingerprint database that contains a limited number of fingerprints is established in the approximate small area confirmed in the first level. Finally, the SVM algorithm is used to match the detected PD fingerprint in the second-level fingerprint database to obtain the specific location of the PD source.

Optical PD simulation is applied instead of conducting an experiment to obtain PD fingerprint data. Then, the NNI algorithm is used to perform fingerprint expansion on the limited simulated fingerprint data, thereby establishing an optical PD fingerprint database covering all locations in the GIL. Finally, through the two-level matching model based on the ECOC-MLP-SVM, the PD source is accurately and efficiently located. The feasibility of this method is verified by the laboratory GIL experimental model.

2 | OPTICAL PD SIMULATION MODEL OF THE GIL

2.1 | Simulation settings of optical PD signals

In the optical PD simulation, the typical needle-plate defect is used to simulate the generation of PD, and other types of PD defect are also applicable. In the simulation, a spherical point light source is set as the PD source, located on the head of the needle defect. It is assumed that the optical PD signals emitted by the PD source are perpendicular to the PD source surface and are evenly distributed [18]. The optical radiation flux of the PD source is 100 watts. The total number of optical rays is 250,000. The insulating medium in the GIL simulation model is set to SF₆, whose optical refractive index is 1.000783. The absorption spectrum of SF₆ is mainly concentrated in the mid-infrared band, while the wavelength of the PD light radiation in SF₆ is mainly concentrated at approximately 500 nm. Thus, the absorption of light by SF₆ can be negligible, and we set the PD light source to green light with a wavelength of 546.1 nm [19].

In order to effectively simulate the received intensity of the optical sensor, the concept of light irradiance $E_o$ is proposed:

$$E_o = \frac{dI_o}{dS}$$  \hspace{1cm} (1)

where $I_o$ is the light radiation flux received by the optical sensor and $S$ is the receiving area of the sensor.

On the condition that the PD localization method is based on the distribution of optical signals between different sensors rather than the direct use of actual intensity, the relative irradiance $E_o$ can fully represent the characteristics of optical PD signals between each sensor.

2.2 | Simulation settings of GIL model material

The material parameters of the GIL simulation model exert a pivotal influence on the propagation and scattering of PD optical signals. The diffuse reflection model of the simulation GIL surface material is a bidirectional reflection distribution function (BRDF) [20]. In Figure 1, this model determines the reflection and scattering law of the PD light signal on the inner surface of the GIL model.

where $\alpha_r$, $\theta_r$ and $\phi_r$ represent the solid angle, elevation angle and azimuth angle of the reflected light, respectively. $\alpha_i$, $\theta_i$ and $\phi_i$ represent the solid angle, elevation angle and azimuth angle of the incident light, respectively.

The definition of the BRDF model is

$$f(\theta_i, \phi_i, \theta_r, \phi_r) = \frac{dL(\theta_r, \phi_r)}{dE(\theta_i, \phi_i)}$$  \hspace{1cm} (2)
where \( dE \) is the incident light irradiance in a direction per unit area and \( dL \) is the reflecting light irradiance in a direction per unit area.

The material of this simulation model is set to aluminum consistent with the experimental GIL tank material. Considering aluminum is a non-transmissive metal material, only the absorption, specular reflection and diffuse reflection of the tank need to be taken into account during the propagation of PD light in this simulation model. The relationship of these three parameters satisfies the following relationship:

\[
\alpha + R + D = 1
\]

(3)

where \( \alpha \) represents the absorption coefficient, \( R \) represents the specular reflection coefficient and \( D \) represents the diffuse reflectance coefficient \[18]. The GIL material used is polished and oxidized medium-smooth aluminium. Among them, the corresponding parameters of the polished and oxidized medium-smooth aluminium used in the GIL model are \( \alpha = 0.3 \), \( R = 0.2 \) and \( D = 0.5 \) \[21]\.

### 2.3 Simulation settings of GIL structure

Owing to the fact that long-distance GIL equipment currently in operation cannot be used for PD testing and lacks reserved optical sensor installation locations and allowing for the main structural characteristics of GIL, a test GIL tank was built in the laboratory and a GIL simulation model of the same size was built in the Tracepro software. The simulation model with the same size as the experimental GIL tank is shown in Figure 2.

The specific dimensions of the experimental GIL are as follows: the wall thickness is 10 mm, the inner radius is 90 mm, the internal height of the GIL tank is 310 mm and the inner conductor diameter of the axis is 25 mm. The axis conductor post is connected to a needle-plate defect model that can change the angle and radial length. The distance among the needle plates is always 6 mm, the angle of the tip cross section is 30°, the length of the needle tip is 25 mm and the radius of the grounding disk below is 10 mm. The occurrence of PD at different locations in the GIL is simulated by changing the location of the needle-plate defect.

For the collection of optical PD signals, according to the distribution of optical sensors on the experimental GIL simulation tank, nine optical signal detection points are also set on the inner wall of the simulated GIL model. Nine detection points are divided into 3 columns at 120° around the tank. The distance between the centre of the 3 detection points on each column from the top surface of the tank is 70, 160 and 250 mm, respectively. The detection point has a radius of 10 mm and is a fully transmissive volume model, which would not generate any catadioptric reflection or absorption effect on the light emitted by the PD source.

According to the aforementioned simulation settings, a PD source is randomly selected in the GIL simulation tank to perform an optical PD simulation experiment. The ray-tracing diagram and the light irradiance distribution through three cross sections are shown in Figure 3.

### 2.4 Establishment of optical PD fingerprint database

On the basis of the aforementioned optical PD simulation, the optical PD fingerprint database is set up, which figures out the problem of obtaining the fingerprint database on site and improves the resolution of the fingerprint database. The construction process of the fingerprint database presents as follows.

An optical PD simulation experiment is performed at \( N \) locations in the simulation tank. Let \( L_1 \) be the location of PD sources \( (j = 1, 2, \ldots, N) \). In every simulation experiment, there are \( M \) detection probes for optical signal acquisition. Let \( S_i \) be the detection probes \( (i = 1, 2, \ldots, M) \). When the simulated PD source is located at \( L_j \), the PD light irradiance detected by the detection probes \( S_i \) is expressed as \( \phi_{i,j} \).
In order to highlight the strength difference and distribution law of the PD signals among the sensors, and to avoid the adverse effects caused by the fluctuation in the intensity of the optical signal between different PD sources, we use the detection signal after normalization and PCA feature extraction as PD fingerprint.

First, the value of each detection probe is subtracted from each other.

\[
D = \frac{M!}{(M-2)! \times 2!} \tag{4}
\]

\[
\delta_{h,N} = \varphi_{a,N} - \varphi_{b,N} \quad \text{s.t.} \quad \begin{cases} h = 1, 2, \ldots, D \\ a < b; \\ a, b \in Z; \\ a, b \in [1, M]. \end{cases} \tag{5}
\]

where \(D\) stands for the fingerprint dimension after subtraction, \(\varphi_{a,N}\) and \(\varphi_{b,N}\) represent the light irradiance of the PD at \(N\) location detected by probes and \(\delta\) is the difference of the light irradiance between each two detection probes.

Second, all the values \(\delta\) of the same PD source are normalized to \([-1,1]\). The rule of normalization is

\[
\delta_{h,j} = \frac{2 \times (\delta_{h,j} - \min(\delta_{1,j}, \delta_{2,j}, \ldots, \delta_{D,j}))}{(\max(\delta_{1,j}, \delta_{2,j}, \ldots, \delta_{D,j}) - \min(\delta_{1,j}, \delta_{2,j}, \ldots, \delta_{D,j}))} - 1 \tag{6}
\]

where \(\delta_{h,j}\) is the normalized light irradiance value.

Third, the first \(P\) principal components are extracted from \(N\) vector \([\delta_{1,P}, \delta_{2,P}, \ldots, \delta_{D,P}]\) by the PCA algorithm, obtaining \(N\) final PD fingerprint \(\Psi_{j} = [\varphi_{1,j}, \varphi_{2,j}, \ldots, \varphi_{P,j}]^\top\).

Finally, all PD fingerprints \(\Psi_{j}\) are combined into an optical PD fingerprint database \(\Psi\):

\[
\Psi = \begin{bmatrix}
\varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,N} \\
\varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\varphi_{P,1} & \varphi_{P,2} & \cdots & \varphi_{P,N}
\end{bmatrix} \tag{7}
\]

where \(P\) is the fingerprint dimension \((P < D)\), \(N\) is the number of the simulated PD sources, the column vector \(\Psi_{j} = [\varphi_{1,j}, \varphi_{2,j}, \ldots, \varphi_{P,j}]^\top\) in the matrix represents the optical PD fingerprint of the PD source \(L_j\) and \(\varphi\) is the transformed fingerprint feature after PCA.

The simulated fingerprint database can be built in line with different structures and sizes of GIL equipment, replacing large-scale offline repetitive experiments and improving localization flexibility.

In the fingerprint database, each PD fingerprint represents the feature information related to location of one PD source. We match the PD fingerprint to be located with the fingerprints in the fingerprint database through a matching algorithm proposed, so as to obtain a final fingerprint, the most similar to the fingerprint to be located. The location corresponding to the final fingerprint is recognized as the localization result.
3 | PD FINGERPRINT EXPANSION: NNI

In the optical PD simulation, although we select as many PD source locations as possible in the GIL simulation tank for simulation, in light of the simulation time and efficiency, it is impossible to traverse all locations in the GIL. Therefore, a PD fingerprint expansion method is proposed to solve the problem of limited location of simulated PD sources. In this manner, the scale of fingerprint samples in the fingerprint database is enlarged, improving the range and accuracy of localization.

3.1 | Natural neighbour coordinates

The natural neighbour is based on the Voronoi tessellation (VT) and the Delaunay triangulation (DT). Although the problem of interpolation in three-dimensional (3D) coordinates is solved here, NNI has the same principle in two-dimensional (2D) and 3D, except that the 2D is extended to 3D Voronoi polyhedra and Delaunay tetrahedra [22]. Therefore, a more intuitive 2D theory is used here to explain.

In the NNI, the VT and DT of the scattered point set correspond to each other. The VT partitions the plane (or space) into a disjoint set of polygons (or polytopes) called tiles. Each tile \( T_i \) encloses one given point \( x_i \) in the scattered point set. The \( T_i \) represents the area (or volume) that is closer to the point \( x_i \) than to any other points of the set:

\[
T_i = \{ x \in \mathbb{R}^2 \mid d(x, x_i) \leq d(x, x_j) \forall j = 1, \ldots, n \} \quad (8)
\]

where \( d(x, x_i) \) represents the Euclidean distance between the points \( x \) and \( x_i \). If \( x_i \) and \( x_j \) have a common boundary or point of contact, the two points are called natural neighbour.

Before NNI, the construction of VT and DT of the given scattered point set is a necessary preprocessing step.

3.2 | Natural neighbour interpolation

Based on the VT and DT constructed, the concept of natural neighbour can be adapted to a newly inserted point \( x \) that is not in the present point set. The new point \( x \) creates a new Voronoi cell by stealing area (or volume) from the existing and surrounding natural neighbours, as shown in Figure 4 [23].

In Figure 4, bold lines represent the tile after the insertion of point \( x \). The dotted lines indicate the old edges before the insertion of point \( x \). The DT associated with the VT is shown by dashed lines, where the boundaries of the tile are the vertical bisector of the triangle boundaries. Therefore, the new tile of point defines as

\[
T(x) = \{ z \in \mathbb{R}^2 \mid d(z, x_i) \leq d(z, x_j) \forall j = 1, \ldots, n \} \quad (9)
\]

**FIGURE 4** Voronoi diagram and Delaunay triangulation (DT) with insertion point \( x \)

The intersections of new tile and old tiles are

\[
T_i(x) = T(x) \cap T_i \quad (10)
\]

Thus, representing the area of a tile \( T \) by \( S(T) \), the NNI at the point \( x \) defines as

\[
f(x) = \sum b_i(x) z_i \quad (11)
\]

where

\[
b_i(x) = \frac{S[T_i(x)]}{S(T)} \left( 0 \leq b_i(x) \leq 1, \sum_i b_i(x) = 1 \right) \quad (12)
\]

Through (6), it can show that the NNI is a weighted average of the natural neighbours of the interpolated point \( x \). However, (6) is only \( C^0 \) continuous. We must use a \( C^1 \)-continuous interpolant at each point in the set. Therefore, we denote the gradient \( \nabla z(x) \) as the function \( z(x) \) at the point \( x_i \) and replace the \( z_i \) in (6) by a first-degree polynomial \( g_i(x) \) shown as

\[
g_i(x) = z_i + \nabla z(x_i)^T (x - x_i) \quad (13)
\]

According to (8), we combine the values \( z_i \) and their corresponding gradients while using natural neighbour weights. Thus, the NNI value defines as

\[
f(x) = \sum_i w_i(x) g_i(x) = \sum_i \frac{b_i(x) d(x, x_i)^{-1}}{\sum_i b_i(x) d(x, x_i)^{-1}} g_i(x) \quad (14)
\]
where
\[ w_i(x) = \frac{\sum_i b_i(x)d(x,x_i)^{-1}}{\sum_i b_i(x)d(x,x_i)^{-1}} \] (15)

NNI is a method based on area-based weights, which is superior to distance-based weights, since NNI would compensate for the data density change, not just be sensitive to distance.

Through NNI, the fingerprint database \( \Psi \) is expanded to the fingerprint database \( \Psi_{NNI} \), which covers the PD fingerprints of all locations in the GIL tank. Finally, in response to the localization requirements of different accuracy, we use different sampling rates to sample fingerprints in \( \Psi_{NNI} \), so as to build fingerprint databases with different degrees of density.

4 | TWO-LEVEL PD LOCALIZATION METHOD: ECOC-MLP-SVM

Although the accuracy and range of localization are improved by the NNI fingerprint expansion, it also increases the burden of calculation in the matching process. Moreover, in view of the large size of the actual GIL device, the number of fingerprints will increase simultaneously while constructing the simulated fingerprint database. Therefore, a two-level fingerprint matching method using ECOC-MLP-SVM is proposed, which can reduce the amount of calculation and avoid the dimensionality curse caused by a large number of PD fingerprints.

4.1 | First-level PD localization: ECOC-MLP

The first level of localization is for the entire area of the GIL tank, but the fingerprint database is constructed at a lower sampling rate from \( \Psi_{NNI} \), that is, the low-density fingerprint database. This method can ease the burden of fingerprint calculation in the matching process and realize the pre-localization of the PD localization.

Although it is devised for the low-density fingerprint database, due to the large size of the GIL and the accuracy of pre-localization, a certain number of fingerprints still remain in the first-level fingerprint database. Therefore, we use the ECOC-MLP algorithm that can simplify the multi-classification problem into multiple binary classification problems to avoid the dimensionality curse in pre-localization.

The ECOC-MLP algorithm includes two main stages: training stage and test stage.

For the training stage, we define a \( C \times b \) code matrix, where \( C \) represents the number of classes. Each class label is encoded by the rows of the coding matrix. Using training samples to train each base classifier of the coding matrix, a classifier with \( b \) output nodes can be obtained.

For the test stage, an incoming test sample \( \mathbf{x} \) is applied to the trained classifier, creating an output vector:
\[ y = [y_1, y_2, \ldots, y_b]^T \] (16)

where \( y_j \) is the output of \( j \)th (\( j = 1, 2, \ldots, b \)) node.

For each class, the ECOC calculates the distance between the output vector and each label of class:
\[ L_i = \sum_{j=1}^b |Z_{ij} - Z_j| \] (17)

where \( Z_{ij} \) represents the value at \( i \)th row and \( j \)th column in code matrix.

Therefore, the decoding rule for test sample \( \mathbf{x} \) is
\[ i = \text{ArgMin}(L_i) \] (18)

Using the aforementioned decoding rules, the MLP learning network is used to transform the determination of each codeword in the coding matrix into the process of weight learning in the MLP network, which is more adapted to the ECOC method [24].

During the construction of the MLP network, a weight \( \omega \) is introduced to reflect the influence of error in the output vector produced by each incoming sample. By the adjustment of the weight \( \omega \), when the error on target codeword is high, produced error in each output node has more influence on the training cost function. This allows the weight parameters of the network to be effectively updated, reducing the error on total codeword by trained network. Therefore, we combine the weight with the MLP training form of minimizing the squared error cost function. The modified cost function \( F \) is
\[ F(\omega) = \frac{1}{N} \sum_{i=1}^N \omega_i(y(\omega, u_i) - d_i)^2 \] (19)

where \( \omega_i \) represents the weight of error produced by \( i \)th sample, \( u_i \) is the input vector, \( d_i \) is the desired network output and \( y(\omega, u_i) \) is the actual network output of the \( i \)th training vector. The summation of total errors produced by the \( i \)th sample in target codeword is as follows:
\[ \omega_i = \sum_{j=1}^b (y_{ij}(\omega, u_i) - d_j)^2 \] (20)

where \( j \) is the number of output nodes, \( b \) is the length of codeword.

Therefore, the first-level fingerprint database is matched by ECOC-MLP to get the first-level location of the PD source, which is prepared for the second-level localization.

4.2 | Second-level PD localization: SVM

The second level of localization is for a small range but a normal density PD fingerprint database. Although the density
of the fingerprint database is higher than that in the first level, the fingerprint database covers a smaller range. Therefore, the number of fingerprints will be far less than in the first level. Centring on the cross section where the first-level localization result locates, extending a certain distance (depends on specific size) to both sides of the GIL is selected as the range of fingerprint matching. In this limited range, $\Psi_{\text{NNI}}$ is sampled at a normal density to obtain the second-level fingerprint database. The selection of database range in two-level localization is shown in Figure 5.

SVM minimizes confidence and empirical risk, so that it can achieve a good classification effect in the case of small samples, while it will have the curse of dimensionality when the number of samples is large. Thus, for the second-level fingerprint database, we use the SVM algorithm to match the actually detected PD fingerprints with the second-level fingerprint database, obtaining the fingerprint most similar to the detected one. The matching function of the SVM algorithm is

$$f(x) = \text{sgn} \left( \sum_{i=1}^{n} \lambda_i y_i K(x_i, x) + b \right) (i = 1, 2, \ldots, n)$$  \hspace{1cm} (21)

where $n$ is the number of training samples, $\lambda$ is Lagrange multiplier, $y_i$ is the label of $x_i$, $b$ is threshold and $K(x_i, x)$ is the inner product kernel function. Here, the Radial Basis Function is used as the kernel function:

$$K(x_i, x_j) = \exp \left( -\frac{||x_i - x_j||^2}{2\sigma^2} \right)$$ \hspace{1cm} (22)

where $\sigma$ represents kernel function parameters. The parameters of SVM are determined through optimization.

Through determining the spatial position corresponding to the PD fingerprint identified by SVM, the final location of the PD source can be obtained in the second-level PD localization.

5 | EXPERIMENTAL VERIFICATION OF PD LOCALIZATION

For the experimental verification, this PD localization method based on the optical simulation mainly includes two parts: experimental detection stage and optical fingerprint localization stage. The overall process of the PD localization is shown in Figure 6.

5.1 | Experimental detection stage

In order to verify the PD localization proposed, we build a platform to perform actual PD experiments at different locations in the experimental GIL simulation tank, as shown in Figures 7 and 8. In Figure 8, the nine sensors are marked with serial numbers.

5.2 | Optical fingerprint localization stage

This stage is mainly divided into two parts: the establishment of the simulated fingerprint database and the matching of PD fingerprints for localization.

For the establishment of the fingerprint database, we select 1620 locations that are evenly distributed throughout the GIL simulation model for PD optical simulation, collecting the light irradiance value of nine detection probes for each simulated PD source. Given that the number of simulated PD sources is limited, to improve the localization accuracy, we use the NNI algorithm to expand the fingerprint database. Under the circumstances that the three rows of sensors are evenly...
distributed on the outer wall of the tank, we take the NNI expanded fingerprint map of one row of sensors as an instance. The schematic diagram is shown in Figures 9–11. For the NNI fingerprint map of each sensor, it represents the intensity of the light signal received by this sensor when a PD occurs at any position in the GIL simulation model, expressed in terms of relative light irradiance. As a result, the NNI fingerprint maps of the nine sensors are combined to form an expanded PD fingerprint database $\Psi_{NNI}$ for two-level fingerprint matching.

For the two-level PD localization method, we use two different sampling rates to sample $\Psi_{NNI}$. First, a low sampling rate is applied to uniformly sample $\Psi_{NNI}$, acquiring the first-level fingerprint database $\Psi_{first-level}$, which contains 1600 PD fingerprints (spatial resolution $= 6086.84 \text{ mm}^3$/per sampling point). The ECOC-MLP algorithm is used to match the detected PD fingerprints $\Psi_{detect}^j$ ($j = 1, 2, \ldots, 18$) with the fingerprints in $\Psi_{first-level}$, which can obtain 18 PD locations of pre-localization. The coding matrix of ECOC is one-versus-all. The number of hidden layer nodes of the MLP network is set to 50. The learning rate of the network is set to 0.005.
FIGURE 9 NNI fingerprint map of the sensor no. 4

FIGURE 10 NNI fingerprint map of the sensor no. 5

FIGURE 11 NNI fingerprint map of the sensor no. 6
According to the tank size used here, the cross section of each pre-located PD location is extended by 10 mm on both sides to form a cylindrical area with a height of 20 mm, which is the range of the fingerprint database for second-level localization. In these 18 small cylindrical areas, the fingerprint database \( W_{\text{fingerprint}} \) is sampled at a normal sampling rate, respectively, obtaining 18 second-level fingerprint databases that contain 722 PD fingerprints (spatial resolution = 870.25 mm\(^3\) per sampling point) in each of them. Among them, in the aforementioned experiment process, the spatial resolution of fingerprint databases is determined on the basis of comprehensive consideration of localization accuracy and localization efficiency. Finally, the SVM algorithm is used to match the 18 detected fingerprints \( \Psi_{\text{detect}}^{j} (j = 1, 2, \ldots, 18) \) with 18 corresponding two-level fingerprint databases to gain the final location of these PD sources.

### 5.3 Experimental results and analysis

The localization errors of the 18 detected PD sources are shown in Figure 12.

From Figure 12 it can be seen that the average localization error of the PD localization method proposed is 9.7 mm. Among them, 55.56% of the PD sources have a localization error of less than 10 mm, and only 11.11% of the PD sources face a localization error greater than 15 mm. In addition, the localization theory followed here is based on the characteristics of each PD fingerprint to identify, which is basically the difference in the distribution of the signal received by each optical sensor. Therefore, in the case of the same spatial resolution of the fingerprint database, the sparser the layout of the sensor (to ensure the detection effect within a certain range), the greater the difference between the PD fingerprints in the fingerprint database, which is more conducive to the identification of the fingerprint recognition algorithm.

Moreover, in the case of the same localization accuracy, we compare the calculation amount of the method (two-level method) proposed and the method (one-level method) that uses the PD fingerprint database with a normal sampling rate directly, as shown in Table 1. The result is run on the same computer.

It can be seen from Table 1 that the two-level localization method can reduce the amount of calculation. At the same time, the method proposed can greatly shorten the localization time.

Therefore, according to the results of the experimental verification, the PD localization method proposed has high localization accuracy and short localization time, which can adapt to GIL equipment of different sizes.

### 6 Conclusion

A GIL PD localization method based on an optical simulation fingerprint database is proposed. First, the PD fingerprint database constructed by optical simulation is expanded through the NNI algorithm. Second, according to the concept of the two-level localization method, the ECOC-MLP-SVM algorithm is used to match and locate the PD fingerprints. Finally, the feasibility of this PD localization method is verified through factual experiments. The following conclusions are drawn:

1. Optical simulation is introduced into the PD localization of GIL. The construction of PD fingerprint database through optical simulation solves the problem of constructing a fingerprint database in a field experiment, which can help obtain PD fingerprint information more conveniently. Meanwhile, the simulation model can be adjusted according to different GIL sizes, expanding the scope of application of the localization method.

2. The fingerprint expansion method based on NNI overcomes the difficulties of limited PD simulation times and broadens the scale of the fingerprint database. The expanded PD fingerprint database can cover all locations of the GIL tank and improve the localization accuracy.

3. The two-level localization method based on ECOC-MLP and SVM solves the problem of a sharp increase in the amount of calculation caused by too many fingerprint samples in the fingerprint database, which shortens the localization time while ensuring localization accuracy. Through the experimental verification, the average localization accuracy of this method can reach 9.7 mm. The
localization time under the same conditions is about 11 times shorter than that of normal one-level localization method.

ACKNOWLEDGEMENTS
This work was supported in part by the National Key R&D Program of China (2017YFB0902500), the National Natural Science Foundation of China under Grant No.62075045 and the State Grid Science & Technology Project (Key Technology for Environmentally Friendly Gas-Insulated Transmission Lines).

ORCID
Yiming Zang https://orcid.org/0000-0002-1370-8820

REFERENCES
1. Tu, Y., et al.: ±100-kV HVDC SF 6/N 2 Gas-Insulated Transmission Line. IEEE Trans. Power Deliv. 35(2), 735–744 (2019)
2. Poehler, S., Rudenko, P.: Directly buried gas-insulated transmission lines (GIL). In: PES T&D 2012, pp. 1–5. IEEE (2012)
3. Magier, T., Tenzer, M., Koch, H.: Direct current gas-insulated transmission lines. IEEE Trans. Power Deliv. 33(1), 440–446 (2017)
4. Wang, J., et al.: Motion analysis of spherical metal particle in AC gas-insulated lines: random effects and resistance of the SF 6/N 2 mixture. IEEE Trans. Dielectr. Electr. Insulation. 23(5), 2617–2625 (2016)
5. Srivastava, K., Morcos, M.: A review of some critical aspects of insulation design of GIS/GIL systems. In: 2001 IEEE/PES Transmission and Distribution Conference and Exposition. Developing New Perspectives (Cat. No. 01CH37294), vol. 2, pp. 787–792. IEEE (2001)
6. Li, C., et al.: Field-dependent interface charging phenomenon based on dominant charge behaviors. Appl. Phys. Lett. 114(20), 202904 (2019)
7. Zheng, Q., et al.: Intelligent learning approach for UHF partial discharge localization in air-insulated substations. High Volt. 5, 583–590 (2020)
8. Hui, Z., Po, H., Ke, T.: Long distance GIL PD Fault localization method based on amplitude difference and time difference calculation of UHF coupling signal. Automation, 39 (2015)
9. Wei, X., et al.: Application research of the partial discharge automatic detection device and diagnostic method based on the ultrasonic in long distance GIL. J. Phys. Conf. Series, 1213(5), 052088 (2019)
10. Ran, Z., et al.: Study on relationship between optical signals and charge quantity of partial discharge under four typical insulation defects. In: 2013 annual report conference on electrical insulation and dielectric phenomena, pp. 1209–1212. IEEE (2013)
11. Li, J., et al.: A novel GIS partial discharge detection sensor with integrated optical and UHF methods. IEEE Trans. Power Delivery. 33(4), 2047–2049 (2016)
12. Guzman, H., Kranz, H.-G.: Analysis of internal partial discharges in solid dielectrics using an electro-oscillation measurement system. In: Proceedings of 1995 IEEE 5th international conference on conduction and breakdown in solid dielectrics, pp. 230–243. IEEE (1995)
13. Biswas, S., et al.: A methodology for identification and localization of partial discharge sources using optical sensors. IEEE Trans. Dielectr. Electr. Insulation. 19(1), 18–28 (2012)
14. Baug, A., et al.: Identification of single and multiple partial discharge sources by optical method using mathematical morphology aided sparse representation classifier. IEEE Trans. Dielectr. Electr. Insulation. 24(6), 3703–3712 (2017)
15. Bag, S., et al.: S-transform aided random forest based PD location detection employing signature of optical sensor. IEEE Trans. Power Deliv. 34(4), 1261–1268 (2018)
16. Lin, N., et al.: GaN-based LEDs with air voids prepared by one-step MOCVD growth. J. Lightwave Technol. 29(18), 2831–2835 (2011)
17. Guruswami, V., Sahai, A.: Multiclass learning, boosting, and error-correcting codes. In: Proceedings of the twelfth annual conference on computational learning theory, pp. 145–155 (1999)
18. Xu, Y., et al.: Simulation analysis on the propagation of the optical partial discharge signal in I-shaped and L-shaped GILs. IEEE Trans. Dielectr. Electr. Insulation. vol. 25(4), 1421–1428 (2018)
19. Ren, M., et al.: Optical partial discharge diagnostic in SF6 gas insulated system via multi-spectral detection. ISA Trans. 75, 247–257 (2018)
20. Qi, C., et al.: Stray light suppression in BRDF measurement infrared optical system. Chinese Opt. Lett. vol. 1(7), 398–400 (2003)
21. Chen, X., et al.: Influence of different structure and specification parameters on the propagation characteristics of optical signals generated by GIL partial discharge. Energies. 15(12), 3241 (2020)
22. Liang, L., Hale, D.: A stable and fast implementation of natural neighbor interpolation. CWP, vol. 4(6), 657. Ctr. for Wave Phenomena, Colorado School of Mines, Golden. (2010)
23. Amidror, I.: Scattered data interpolation methods for electronic imaging systems: a survey. J. Electron. Imaging. 11, 157–176 (2002)
24. Hatami, N., Ebrahimpour, R., Ghaderi, R.: Ecoc-based training of neural networks for face recognition. In: 2008 IEEE Conference on Cybernetics and Intelligent Systems, pp. 450–454. IEEE (2008)
25. Cheung, J.Y., et al.: Low optical power reference detector implemented in the validation of two independent techniques for calibrating photon-counting detectors. Opt. Express. 19(21), 20347–20363 (2011)

How to cite this article: Zang Y, Qian Y, Wang H, Xu A, Sheng G, Jiang X. Method of GIL partial discharge localization based on natural neighbour interpolation and ECOC-MLP-SVM using optical simulation technology. High Voltage. 2021;6:514–524. https://doi.org/10.1049/hve2.12071